

Are American rivers Tokunaga self-similar? New results on fluvial network topology and its climatic dependence

S. Zanardo,¹ I. Zaliapin,² and E. Foufoula-Georgiou^{1,3}

Received 21 February 2012; revised 20 December 2012; accepted 21 December 2012; published 1 March 2013.

[1] The topology of river networks has been a subject of intense research in hydrogeomorphology, with special attention to self-similar (SS) structures that allow one to develop concise representations and scaling frameworks for hydrological fluxes. Tokunaga self-similar (TSS) networks represent a particularly popular two-parameter class of self-similar models, commonly accepted in hydrology but rarely tested rigorously. In this paper we (a) present a statistical framework for testing the TSS assumption and estimating the Tokunaga parameters; (b) present an improved method for estimating the Horton ratios using the Tokunaga parameters; (c) evaluate the proposed testing and estimation frameworks using synthetic TSS networks with a broad range of parameters; (d) perform self-similar analysis of 408 river networks of maximum order $\Omega \geq 6$ from 50 catchments across the US; and (e) use the Tokunaga parameters as discriminatory metrics to explore climate effects on network topology. We find that the TSS assumption cannot be rejected in the majority of the examined river networks. The theoretical expression for the Horton ratios based on the estimated Tokunaga parameters in the TSS networks provides a significantly better approximation to the true ratios than the conventional linear regression approach. A correlation analysis shows that the Tokunaga parameter c , which determines the degree of side-branching, exhibits significant dependence on the hydroclimatic variables of the basin: storm frequency, storm duration, and mean annual rainfall, offering the possibility of relating climate to landscape dissection. While other possible physical controls have been neglected in this study, this result is intriguing and warrants further analysis.

Citation: Zanardo, S., I. Zaliapin, and E. Foufoula-Georgiou (2013), Are American rivers Tokunaga self-similar? New results on fluvial network topology and its climatic dependence, *J. Geophys. Res. Earth Surf.*, 118, 166–183, doi:10.1029/2012JF002392.

1. Introduction

[2] The branching structure of river networks has been actively studied since the 1960s, to address a broad range of hydrological, geomorphological, and environmental problems. Specifically, the attention of the hydrologic community has been traditionally drawn by the connections between the river network topology and the hydrologic response. An extensive literature therefore exists in this area, starting with the work of *Kirkby* [1976], and *Rodriguez-Iturbe and Valdes* [1979] and followed by numerous other studies reviewed in *Rodriguez-Iturbe and Rinaldo* [1997]. The river network structures have also been shown to constitute a dominant control for other observed natural processes, such as

stream biodiversity, riparian vegetation functioning, bed load sediment size distribution, and food web structures [e.g., *Muneepeerakul et al.*, 2008; *Power and Dietrich*, 2002; *Kiffney et al.*, 2006; *Sklar et al.*, 2006; *Stewart-Koster et al.*, 2007].

[3] The ability to characterize a river network via its topological properties is a powerful tool to approach the above problems. It allows one to quantitatively describe the connections between the network properties and various observed processes that operate on the network, as well as to perform comprehensive numerical simulations aimed at hypothesis testing and improving the understanding of the network dynamics. For example, the observed scaling relationships between hydrologic (e.g., annual peak flow) and geomorphic (e.g., drainage area, width function peak) variables have been frequently studied through ensemble simulations of synthetic river networks [e.g., *Menadbe et al.*, 2001; *Veitzer and Gupta*, 2001].

[4] A commonly accredited property of river networks, based on empirical observations, is the so-called Tokunaga self-similarity (TSS) [*Tokunaga*, 1966, 1978], which constitutes a standard assumption in river network modeling. The Tokunaga self-similar model has two assumptions: (1) the mean number T_{ij} of branches of order i that merge with a randomly selected branch of order j does not depend on the branch orders, only on the difference $j-i$; and therefore

All Supporting Information may be found in the online version of this article

¹National Center for Earth-Surface Dynamics, University of Minnesota, Minneapolis, Minnesota, USA.

²Department of Mathematics, University of Nevada, Reno, Nevada, USA.

³Department of Civil Engineering, University of Minnesota, Minneapolis, Minnesota, USA.

Corresponding author: S. Zanardo, National Center for Earth-Surface Dynamics (NCED), University of Minnesota, Minneapolis (MN), USA. (szanardo@umn.edu)

$T_{i(i+k)} = T_k$ for any $i, k \geq 1$; and (2) the numbers T_k obey the exponential relationship $T_k = ac^{k-1}$ for two positive topological constants a and c . Since Peckham [1995a, 1995b], the Tokunaga model has been gaining increasing popularity, not limited to the hydrologic literature, and constitutes now a "benchmark criterion" for current network modeling approaches [e.g., Tarboton, 1996; Newman et al., 1997; Dodds and Rothman, 1999; Cui et al., 1999; McConnell and Gupta, 2008].

[5] The Tokunaga self-similarity assumption of river networks is generally accepted in the literature. However, to the best of our knowledge, a data-based support for TSS has been provided only in a few studies [i.e., Tokunaga, 1966; Peckham, 1995a, 1995b; Peckham and Gupta, 1999; Tarboton, 1996] and for a very limited number of basins; most of the subsequent works on the Tokunaga model refer to these studies. Moreover, when the TSS property was reported, data limitations often precluded a rigorous analysis and hypothesis testing. For example, Peckham [1995a] noted that the values of T_k from two real networks "appear to fluctuate about fairly stable values". While this observation suggested that the considered networks are self-similar, no formal test was applied to investigate whether or not the reported fluctuations were statistically significant.

[6] Studies that statistically confirm (or reject) the TSS property over a range of different climatic and topographic regions are still lacking. An important related question is whether the Tokunaga parameters a and c show a characteristic value or a range of values. For example, Cui et al. [1999] argued that these parameters can be interpreted as representing the effects of regional controls. Moreover, as noticed by McConnell and Gupta [2008], it is reasonable to expect some physical restrictions on the values of a and c , based on the typically observed Horton ratios (i.e., common descriptors of the topological structure of river networks based on a hierarchical ordering of their tributaries) [Horton, 1945]; however such restrictions are yet to be explored.

[7] Motivated by this growing interest in the Tokunaga model and by the lack of a rigorous, data-based testing procedure for the Tokunaga self-similarity, the current study has the following main goals: (1) To identify a set of formal statistical methods, that allows to analyze the topology of the river networks and to estimate the accuracy of these methods; (2) to evaluate, on the basis of an extensive dataset, whether the Tokunaga self-similarity assumption for river networks holds across different climatic and topographic regions; (3) to evaluate the range of the Tokunaga parameters (a, c) in the analyzed Tokunaga self-similar river networks; (4) to investigate whether the distribution of side-branches is geometric as suggested by previous theoretical [Burd et al., 2000] and empirical [Troutman, 2005; Mantilla et al., 2010] studies; and (5) to explore whether the Tokunaga parameters can serve as discriminatory metrics to understand the controls on landscape dissection and river network topology. It should be clear that the focus of this work is solely on the topology of river networks, as we do not consider any geometrical characteristics such as channel lengths or drainage areas. As extensively discussed in section 2, when a river network is found to be TSS, its topology can be completely characterized, in an average sense, by the values of the two Tokunaga parameters, which also define analytically the topological Horton ratios. Moreover, if the distribution of side-branches is geometric

(or any other one-parameter distribution), the branching structure of the TSS networks admits also a rigorous probabilistic characterization, based solely on the Tokunaga parameters. Therefore, not only will these results improve our understanding of the topological structure of river networks, but also they will give more confidence (on the basis of the extensive dataset analyzed) in the parameterization of current river network models. As illustrated in section 3, the topological properties are evaluated using well known statistical methods, whose performance is studied using numerical simulations in section 4 to address goal 1. In particular, in section 4 we apply the proposed tests to a set of synthetic TSS trees for which we know a priori the true TSS parameters. We then assess the ability of the proposed tests to accurately estimate these properties, as well as define the confidence level associated with these tests. In section 5 the proposed statistical methodologies are applied to 408 real river networks extracted from 50 catchments across the continental United States (goals 2, 3, and 4).

[8] Recently, Mantilla et al. [2010] have analyzed 30 river basins to test a particular form of statistical self-similarity used in the Random Self-similar Network (RSN) model of Veitzer and Gupta [2000], as well as to evaluate the range of variability of the RSN parameters. The RSN and Tokunaga models use different mechanisms of generating a self-similar topology: the Tokunaga model uses an appropriate random sampling of side branches [e.g., Cui et al., 1999], while the RSN model uses an iterative replacement of randomly sampled network generators [Veitzer and Gupta, 2000]. Our results hence complement those of Mantilla et al. [2010] in building a solid empirical basis for the theoretical results on river network topology.

[9] The rather wide range of the TSS parameter c found from a large number of catchments across the US prompts the question as to what physical parameters might affect the topological structure of a river network and whether the Tokunaga parameters can serve as metrics to explore this question (goal 5). Although a complete answer to this question would require extensive study of climatic, geologic, ecologic, and soil properties of the catchments, we present in this study (section 6) a preliminary analysis of the dependence of the parameter c on a range of hydroclimatic variables and report significant dependence. We interpret this result as encouraging, prompting further study on the connection between landscape forming processes and fluvial network topology. Discussion of our results and suggestions for future research are given in section 7.

2. Review of Horton and Tokunaga Self-Similarity

2.1. Horton-Strahler Orders and Tokunaga Indices

[10] A stream network is represented here by a planar rooted tree T , as illustrated in Figure 1. It is composed of sources (channel heads), vertices (stream junctions), edges (stream links) and a root (basin outlet).

[11] The Horton-Strahler (HS) ordering of river tributaries was initially outlined by Horton [1945] and later refined by Strahler [1957]. The HS order is defined to be the same for a stream junction and the immediate (unique) downstream link. The HS ordering of the streams in a tree is performed in a hierarchical fashion, from the sources to the outlet, as follows [Horton, 1945; Strahler, 1957; Newman et al., 1997; Burd et al., 2000]: (1) each source has order r

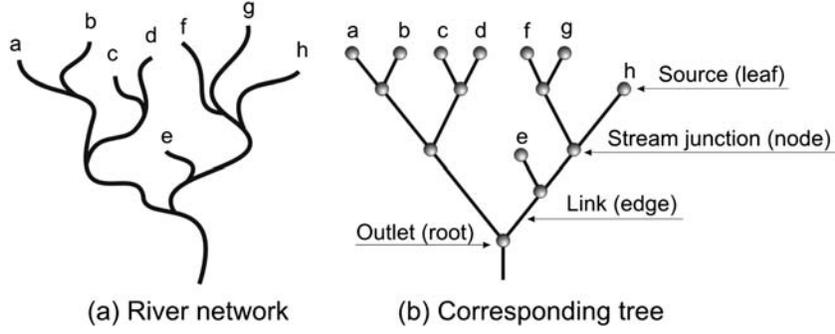


Figure 1. A river basin (panel a) and its representation by a planar rooted tree (panel b).

(source) = 1; (ii) when two streams, c_1, c_2 , of the same order r meet, they form a stream p of order $r(p) = r + 1$; (iii) when two streams of different orders meet, they form a stream p of the highest order of the two. Figure 2a illustrates this definition, which can be formally written as

$$r(p) = \begin{cases} r(c_1) + 1 & \text{if } r(c_1) = r(c_2), \\ \max\{r(c_1), r(c_2)\} & \text{if } r(c_1) \neq r(c_2). \end{cases} \quad (1)$$

is called the *initial junction*, the junction farthest from the root is called the *terminal junction*. The order $\Omega(T)$ of a finite network T is the order of its root, or, equivalently, the maximal order of its links (or junctions). The *magnitude* m_i of a branch i is the number of sources upstream to its initial junction. In what follows, N_r will denote the total number of branches of order r , M_r the average magnitude of branches of order r and C_r the average number of stream links in a branch of order r in a finite tree T .

[13] The Tokunaga indexing [Tokunaga, 1966, 1978; Peckham, 1995a, 1995b; Newman et al., 1997] extends the Horton-Strahler ordering by cataloging *side-branching*, that is the merging of streams of different orders; it is illustrated

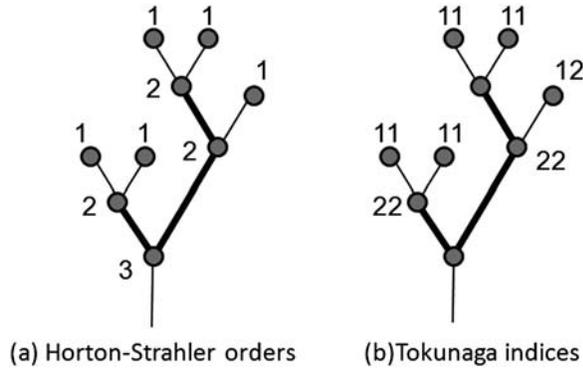


Figure 2. Example of (a) Horton-Strahler ordering, and of (b) Tokunaga indexing. Two order-2 branches are depicted by heavy lines in both panels. The Horton-Strahler orders refer, interchangeably, to the stream joints or to the immediate upstream links. The Tokunaga indices refer to the entire branches, and not to individual joints or links.

in Figure 2b. Let $\tau_{ij}^{(l)}$, $1 \leq l \leq N_j$, $1 \leq i < j \leq \Omega$ denote the number of branches of order i that join the nonterminal junctions of the l -th branch of order j . We define $N_{ij} = \sum_l \tau_{ij}^{(l)}$, $j > i$ as the total number of branches of order i that join a branch of order j in a tree T . In a finite tree of order $\Omega \geq j$, the Tokunaga index T_{ij} is defined as the average number of branches of order $i < j$ per branch of order j :

$$T_{ij} = \frac{N_{ij}}{N_j}. \quad (2)$$

2.2. Horton Self-Similarity

[14] The topological *Horton laws*, widely observed in hydrological and biological networks [Turcotte et al., 1998; Horton, 1945; Veitzer and Gupta, 2000; Dodds and Rothman, 2000], state, in their ultimate form, the equality of the ratios of various branch statistics for two consecutive orders. For instance, the commonly studied Horton laws include

$$\frac{N_r}{N_{r+1}} = R_B, \quad \frac{M_{r+1}}{M_r} = R_M, \quad \frac{C_{r+1}}{C_r} = R_C, \quad (3)$$

for an appropriate range of orders r and some positive constants R_B, R_M , and R_C . Recall that N_r and M_r are, respectively, the total number and average magnitude of branches of order r , and C_r is the average number of links within a branch of order r in a finite tree of order Ω . The Horton laws imply, in particular,

$$N_r \sim M_r^{-\alpha} \quad \text{as } r \rightarrow \infty, \quad \alpha = \frac{\log R_B}{\log R_M}, \quad (4)$$

which explains why the Horton laws can also be called *Horton self-similarity*. McConnell and Gupta [2008] emphasized the approximate, asymptotic nature of the empirical statements (3) and the necessity to consider an appropriate limit of the ratios of the branch statistics. Often, the *Horton laws* are stated as the convergence of the ratios of the branch statistics as the tree order increases [Peckham, 1995a; McConnell and Gupta, 2008]:

$$\frac{N_r}{N_{r+1}} \rightarrow R_B > 0, \quad \text{as } \Omega - r \rightarrow \infty, \quad (5)$$

$$\frac{M_{r+1}}{M_r} \rightarrow R_M > 0, \quad \text{as } r, \Omega \rightarrow \infty, \quad (6)$$

$$\frac{C_{r+1}}{C_r} \rightarrow R_C > 0, \quad \text{as } r, \Omega \rightarrow \infty. \quad (7)$$

[15] Here the limit constants R_B, R_M, R_C are called *Horton ratios*. Notice that the convergence in (5) is seen for the small-order branches, while the convergence in (6) and (7) for large Ω and large-order branches.

[16] In a probabilistic set up, one considers a space of finite binary trees with an appropriate probability measure [e.g., *Burd et al.*, 2000; *Veitzer and Gupta*, 2000; *Zaliapin and Kovchegov*, 2011]. Then the statistics $N_r, M_r, C_r, \tau_{ij}^{(l)}, N_{ij}$, and T_{ij} become random variables and the convergence in (5)–(7) should be understood in a suitable probabilistic sense.

2.3. Tokunaga Self-Similarity

[17] In a deterministic setting, we call a tree T of order Ω a *self-similar tree* if its side-branch structure is (1) the same for all branches of a given order:

$$\tau_{ij}^{(l)} = T_{ij}, \quad 1 \leq l \leq N_j, \quad 1 \leq i < j \leq \Omega, \quad (8)$$

and (2) invariant with respect to the branch order:

$$T_{i(i+k)} := T_k \quad \text{for } 2 \leq i + k \leq \Omega. \quad (9)$$

A Tokunaga self-similar (TSS) tree obeys an additional constraint first considered by *Tokunaga* [1978]:

$$T_{k+1}/T_k = c \quad \Leftrightarrow \quad T_k = ac^{k-1} \quad a, c > 0, \quad 1 \leq k \leq \Omega - 1. \quad (10)$$

[18] Here, the pair (a, c) is called *Tokunaga parameters*.

[19] In a random setting, we say that a random tree T of order Ω is *self-similar* if

$$E[\tau_{i(i+k)}^{(j)}] := T_k \quad \text{for } 1 \leq j \leq N_{i+k}, 2 \leq i + k \leq \Omega. \quad (11)$$

[20] A self-similar random tree is called *Tokunaga self-similar* if, furthermore, the condition (10) holds.

[21] In a deterministic tree that satisfies both the Horton and Tokunaga self-similarity laws, one has [*Tokunaga*, 1978; *Peckham*, 1995a]:

$$R_B = \frac{2 + c + a + \sqrt{(2 + c + a)^2 - 8c}}{2}, \quad (12)$$

$$R_C = c. \quad (13)$$

[22] *Peckham* [1995a] has noticed that in a deterministic Tokunaga tree of order Ω one has $N_r = M_{\Omega-r+1}$, which implies that the Horton law (6) for magnitudes M_r follows from the Horton law (5) for the counts N_r , and vice versa, with $R_M = R_B$. He also conjectured that $R_C < R_B$. *McConnell and Gupta* [2008] have demonstrated that both the asymptotic Horton laws (5), (6) with $R_B = R_M$ follow from the Tokunaga self-similarity (10).

[23] Given the random nature of natural processes, in the analysis of real river networks it is appropriate to use the probabilistic definitions of the Horton and Tokunaga self-similarity. This means that a tree T that describes a given stream network is considered as a *random tree* from

a space \mathcal{T} of finite binary trees with some probability measure \mathbf{P} . The next section describes the statistical inference about random trees.

3. Statistical Inference About Random Trees

[24] This section introduces statistical tests for self-similarity, Tokunaga self-similarity, and the distribution of the side branch counts $\tau_{ij}^{(l)}$ in a finite random tree T as well as the estimation procedures for the Tokunaga parameters and Horton ratios.

3.1. Testing the Self-Similarity Hypothesis

[25] The definition (11) of self-similarity in a random tree deals with the side-branch counts $\tau_{ij}^{(l)}$, which may have a joint distribution with distinct marginals and complicated correlation structure. Accordingly, a direct statistical test of (11) in a single tree will hardly have a reasonable power. To address this issue in a practical fashion, we introduce the following two assumptions:

- (A1) the side-branch counts $\tau_{ij}^{(l)}$, for different values of the triplet (i, j, l) , are independent random variables;
- the side-branch counts $\tau_{ij}^{(l)}$ for a fixed pair (i, j) and $1 \leq l \leq N_j$ can be considered a sample (i.e., a set of independent identically distributed random variables) from a random variable that we denote by τ_{ij} .

[26] Under the assumptions (A1)-(A2) the tree self-similarity is equivalent to the null hypothesis

$$H_0 : E[\tau_{ij}] =: T_{j-i}, \quad 1 \leq i < j \leq \Omega, \quad (14)$$

which can be tested using the analysis of variance (ANOVA) framework. Specifically, the above composite null hypothesis is composed of $\Omega - 2$ individual hypotheses, each of which corresponds to a fixed difference $1 \leq j - i = k \leq \Omega - 2$ and can be verified by an ANOVA test with the P -value P_k :

$$H_{0,1} : E[\tau_{12}] = E[\tau_{23}] = \dots = E[\tau_{(\Omega-2)(\Omega-1)}] = E[\tau_{(\Omega-1)\Omega}],$$

$$H_{0,2} : E[\tau_{13}] = E[\tau_{24}] = \dots = E[\tau_{(\Omega-2)\Omega}],$$

...

$$H_{0,(\Omega-2)} : E[\tau_{1(\Omega-1)}] = E[\tau_{2\Omega}]. \quad (15)$$

[27] The case $j - i = \Omega - 1$ includes a single random variable $\tau_{1\Omega}$ and hence is not tested.

[28] The overall decision in the above multi-comparison test is based on the analysis of the multiple P -values, P_k , $k = 1, \dots, \Omega - 2$. Specifically, we fix the confidence level $0 < \beta_s < 1$ and reject the null hypothesis if at least one of the P -values is below β_s , that is if at least one of our individual hypotheses $H_{0,k}$ is rejected at level β_s . In short, the null hypothesis is rejected if $\min_k P_k < \beta_s$. This procedure results in rejecting proportion α_s of true null hypotheses, where $\alpha_s \geq \beta_s$ with equality attained only for $\Omega = 3$ when we work with a single null hypothesis $H_{0,1}$. In case of a composite null, $\Omega > 3$, there exist various approaches to find an appropriate *corrected level* β_s for a given level α_s . The simplest is the *Bonferroni correction*,

which suggests $\beta_s = \alpha_s / (\Omega - 2)$; see *Westfall et al.* [2011] and references therein for an overview on multiple comparison tests. In this study we will use simulations to find β_s for a given α_s . The qualitative conclusions of this study remain valid when using Bonferroni or other conventional corrections. For a tree that passes the above test, we will say that the hypothesis of self-similarity cannot be rejected at a confidence level α_s .

[29] A couple of comments are in order about the properties of the proposed test and the validity of its main assumptions. First, we notice that the random variables $\tau_{ij}^{(l)}$ may have different variance for different l . This effect is called *heteroskedasticity*, and it is known to affect the ANOVA test. However, it is also known that ANOVA is robust with respect to a mild heteroskedasticity [*Turner and Thayer*, 2001]. The results of *Peckham and Gupta* [1999], furthermore, support the assumption that the variance of $\tau_{ij}^{(l)}$ is not varying much for a given $(j-i)$.

[30] The proposed test relies on the validity of the assumptions (A1)-(A2), which is hardly possible to test in real basins. These assumptions, however, seem to be not completely unrealistic in the light of the existing empirical evidence [*Peckham*, 1995a; *Peckham and Gupta*, 1999; *Mantilla et al.*, 2010]. Moreover, in the above hypothesis testing the primary concern is about the independence of the samples within the same hypothesis $H_{0,k}$. We notice that those samples are always collected over *distinct*, physically separated stream branches, which provides an intuitive justification for our assumptions.

[31] Finally, our test is applied to discrete random variables, with approximately geometric distribution [see section 3.4 and *Peckham and Gupta*, 1999], while the classical ANOVA test is designed for continuous Normal random variables. Our experiments with synthetic SST trees with geometric number of side branches in section 4 demonstrate that the values of P_k are (1) independent and identically distributed, so that each individual null hypothesis has the same chance to be rejected, and (2) the nominal (β_s) and actual ($\hat{\beta}_s$) levels in each individual test are very close for $\beta_s > 0.05$, while for smaller values, $\beta_s < 0.05$, the actual levels are higher, $\hat{\beta}_s > \beta_s$ (not shown). In other words, the numerical experiments produce a larger number of very small P -values than expected in a Normally distributed population. This effect is due to the discrete nature of the distribution of side-branch numbers; it motivates us to rely on numerically estimated significance levels β_s in this study.

3.2. Testing the Tokunaga Self-Similarity and Estimating the Tokunaga Parameters

[32] If a tree T passes the self-similarity test described above, the next question is whether the tree is also Tokunaga self-similar, according to the condition (10). We notice that even if the null hypothesis (14) is not rejected, one does not know the values T_k of the Tokunaga coefficients. Hence, testing the Tokunaga constraint (10) in a self-similar tree is a complementary statistical problem. The first step to its solution is the pooled estimation of the Tokunaga coefficients, which will be done by the maximal likelihood method:

$$\hat{T}_k = \frac{1}{n_k} \sum_{i=1}^{\Omega-k} N_{i(i+k)} = \frac{1}{n_k} \sum_{i=1}^{\Omega-k} \sum_{l=1}^{N_{i+k}} \tau_{i(i+k)}^{(l)}, \quad 1 \leq k \leq \Omega - 1, \quad (16)$$

where $n_k = \sum_{i=k+1}^{\Omega} N_i$ is the number of branches of order equal or larger than $k+1$. The testing of the Tokunaga constraint (10) is done here by analysis of the goodness of fit of an additive linear model

$$\log_{10} \hat{T}_k = (k-1) \log_{10} \hat{c} + \log_{10} \hat{a} + e_k, \quad k = 1, \dots, \Omega - 1$$

for a suitably chosen parameter estimation (\hat{a}, \hat{c}) . The goodness of fit is measured by the *coefficient of determination*

$$R^2 = 1 - \frac{\sum_k e_k^2}{\sum_k (\hat{T}_k - \bar{T})^2}, \quad (17)$$

where

$$\bar{T} = 1/(\Omega - 1) \sum_{k=1}^{\Omega-1} \hat{T}_k.$$

[33] The coefficient R^2 can take values in the interval $[0, 1]$ with $R^2 = 1$ indicating a perfect linear fit.

[34] A tree T will be called Tokunaga self-similar if the respective coefficient of determination R^2 is above a pre-defined value R_0^2 , which is determined on the basis of extensive numerical simulations of synthetic TSS trees, as discussed in section 4.

[35] It seems natural to estimate the parameters (a, c) in the above model by the *least squares method* (LSM), that is minimizing the sum of squared residuals $\sum_k e_k^2$ in the log-log form of equation (10). However, as demonstrated below, this results in dependent estimators (\hat{a}, \hat{c}) , which is inconsistent with the Tokunaga self-similarity set-up. In fact, the regular LSM gives the same weight to all the indices \hat{T}_k , while the number of branches used to calculate the indices according to (16) exponentially decreases with k .

[36] This motivates us to use a *weighted least squares method* (WLSM), according to which the estimated \hat{a} and \hat{c} minimize the quantity $\sum_k w_k e_k^2$ where w_k is a weight determined by the number z_k of branches used to estimate \hat{T}_k :

$$z_k = \sum_{i=1}^{\Omega-k} \sum_{l=1}^{N_{i+k}} 1. \quad (18)$$

[37] The application of this method to TSS trees (see section 4) suggests that $w_k = \sqrt{z_k}$ represents a good choice for the weights. It can be shown that these weights produce the same result as the least square linear regression that would use z_k values \hat{T}_k at point k . The respective WLSM significantly reduces the dependence of the pair (\hat{a}, \hat{c}) and leads to a smaller variance of the estimated parameters than the one produced with unweighted LSM.

3.3. Estimating the Horton Ratios

[38] The Horton ratios are commonly estimated [*Peckham*, 1995a; *Peckham and Gupta*, 1999; *Furey and Troutman*,

2008], from the best linear interpolation to the logarithms of the branching statistics. This approach is based on an alternative form of the Horton laws, which we give here only for N_r :

$$N_r \sim N_0 R_B^{\Omega-r}, \quad \Omega - r \rightarrow \infty,$$

where $x_r \sim y_r$ stands for $\lim_{r \rightarrow \infty} x_r/y_r = 1$. For instance, the estimator of R_B is constructed as $\hat{R}_B = 10^b$, where b is the slope of the best least squares fit to the linear model

$$\log_{10} N_r = A - br + e_r.$$

[39] Here we also test an alternative approach that directly uses the ratios of the branch statistics. We call this approach *sample average ratio method* as we consider the following estimators of the Horton ratios in a finite tree T of order Ω :

$$\begin{aligned} \hat{R}_B &= \frac{1}{\Omega-1} \sum_{r=1}^{\Omega-1} \frac{N_r}{N_{r+1}}, & \hat{R}_M &= \frac{1}{\Omega-1} \sum_{r=1}^{\Omega-1} \frac{M_{r+1}}{M_r}, \\ \hat{R}_C &= \frac{1}{\Omega-1} \sum_{r=1}^{\Omega-1} \frac{C_{r+1}}{C_r}. \end{aligned} \quad (19)$$

[40] Since the convergence happens for high orders (equations (6), (7)), we expect these methods to be biased as they use all the orders. Therefore, these estimated ratios will be also compared with those theoretically predicted in a TSS tree by the Tokunaga parameters according to the equations (12), (13).

3.4. Testing the Distribution of Side-Branch Counts

[41] Beyond testing the river network topology, our analysis is further extended to testing the distribution of the number of side-branches. In particular, we test whether the distribution of the side-branch counts $\tau_{i(i+k)}^{(l)}$, for each fixed order difference $1 \leq k \leq \Omega - 1$, is geometric. The geometric hypothesis is motivated by the works of *Burd et al.* [2000], who proved that for the critical binary Galton-Watson tree, also known as Shreve's random topology model, the distribution of side-branch count is geometric; *Mantilla et al.* [2010], who have shown qualitatively that the number of internal and external nodes of the branches in real stream networks follows a geometric distribution; and *Peckham and Gupta* [1999], who showed empirical distributions for the side-branch counts in real rivers that are reminiscent of the geometric distribution.

[42] The goodness of fit of the geometric distribution is sometimes evaluated using the χ^2 test, which is known to have low power, especially when sample sizes are small [*Bracquemond et al.*, 2011]. In this study, the hypothesis of geometrically distributed side-branch counts is tested on the basis of the *generalized Smirnov transformation* (GST) [*Nikulin*, 1992], which allows the transformation of a geometric sample into an exponential sample. This method was applied by *Bracquemond et al.* [2011] to the goodness of fit for the geometric distribution and can be summarized as follows. The samples to be tested are first transformed using the GST, then the well-known Kolmogorov-Smirnov test is applied to test the exponential null hypothesis in the transformed samples. The Kolmogorov-Smirnov test is run for each of the order differences, $1 < k < \Omega - 1$, thus

producing $(\Omega - 1)$ individual P -values. The null hypothesis is rejected when the smallest P -value is smaller than the corrected level β_g for a suitable significance level α_g . The corrected level β_g is determined by numerical experiments in SST trees with geometric number of side branches. The qualitative conclusions of this study remain the same if one will use the Bonferroni or any other conventional level correction.

4. Monte Carlo Evaluation of the Statistical Inference Methods

[43] This section applies the inference methods introduced in section 3 to synthetic Tokunaga trees with geometric distribution of the side-branch counts $\tau_{ij}^{(l)}$. Our goal is to evaluate the statistical tests illustrated in section 3. To this end, in what follows we will apply the statistical tests to the synthetic TSS trees and evaluate the ability of the methods to capture the analyzed properties (i.e., self-similarity, Tokunaga self-similarity and geometric distribution of side-branches). In this way, for each test we can also estimate the actual significance level to be used in the analysis of real river networks as well as determine the accuracy in the estimation of the Tokunaga parameters and the Horton ratios.

[44] A synthetic TSS tree of order Ω with the Tokunaga parameters (a, c) is constructed using the following recursive procedure. A tree of order $\Omega = 1$ consists of a single root vertex and a root edge. To construct a tree of order $\Omega > 1$, start with a *perfect binary tree* of order Ω ; this tree has $2^{\Omega-1}$ leaves of the same depth $\Omega - 1$. Assign Horton-Strahler orders to the vertices of this tree; in this simple case the order r is related to the vertex depth d via $r = \Omega - d + 1$. Each branch in this tree consists of a single edge. To each branch of order $2 \leq \Omega' \leq \Omega$ attach a random number of side-branches of smaller orders $\Omega'' = \Omega' - k$, $k = 1, \dots, \Omega' - 1$. The random number of side-branches of order $\Omega' - k$ is drawn from the geometric distribution with mean T_k , which is given by (10). The order of attachment of side-branches to a given branch is random: all permutations of the side branches are equally likely. Each side branch of order $\Omega'' < \Omega$ is constructed using the same procedure. The ability to construct a tree of order $\Omega = 1$ ensures that this recursive procedure is well defined.

[45] For this analysis we considered 70 pairs of the Tokunaga parameters (a, c) that uniformly cover the range $0.9 \leq a \leq 1.3$, $2 \leq c \leq 3.4$. This range is chosen to surround the values of the Tokunaga parameters found for the real river networks: $(a, c) \approx (1.0, 2.5)$ [see *Peckham*, 1995a; *Burd et al.*, 2000; *Peckham and Gupta*, 1999; and section 5 below]. Contour plots shown in Figure 4 show how 70 pairs of Tokunaga parameters are indeed enough to assess the accuracy of the analyzed methods. For each pair of the Tokunaga parameters we analyzed 1000 independent Tokunaga trees, as for a higher number of trees the results did not change significantly. It was observed that results slightly vary with the order of the synthetic networks and in particular the efficacy of the statistical tests increases with the order, or, in other words, the larger the order the stricter the test. This implies that for a conservative assessment of the test efficiency we would need to use synthetic trees with large order; however, given the small number of order 8 and

9 river networks in our data set (see section 5) we used synthetic trees of order $\Omega=7$, which was the order of the majority of the catchments considered. During the self-similarity analysis, in simulations as well as in observations, we do not consider branch statistics related to the highest tree order: N_Ω , $N_{k,\Omega}$, and $T_{k,\Omega}$. Hence, instead of working with a tree of order $\Omega=7$ we work with a forest of trees of order $\Omega=6$. This truncation is motivated by the fact that the branch of the largest order might behave statistically different from the theoretical predictions because of the finite-size effects.

4.1. Self-Similarity and Geometric Distribution

[46] The main goal here is to compare the nominal and actual significance levels in testing the self-similarity and geometric null hypothesis via the ANOVA and the Kolmogorov-Smirnov approaches, respectively. Specifically, we used the simulations to compare the nominal significance level α and the actual proportion of experiments when the true null hypothesis is rejected. For each Tokunaga pair (a,c) we found the levels $\beta_s(a,c)$ and $\beta_g(a,c)$ that correspond to rejecting the true null (multi-comparison) hypothesis in 5% of the examined trees (50 out of 1000). Recall that the null is rejected when $P_0 = \min_k P_k < \beta$. Our simulations resulted in a set of β_s values (one for each pair of Tokunaga parameters) with the average of $\bar{\beta}_s = 0.0052$ and sample standard deviation of $\sigma_{\beta_s} = 0.002$; and a set of β_g values with the average of $\bar{\beta}_g = 0.004$ and sample standard deviation of $\sigma_{\beta_g} = 0.002$. These levels are used in the data analysis of section 5 for the tests on network self-similarity and geometric distribution of side-branches, respectively.

4.2. Tokunaga Self-Similarity and Parameters

[47] We start by comparing the properties of the estimates of the Tokunaga parameters using the least squares method and weighted least squares method introduced in section 3.2. Figure 3 shows the parameters (a,c) estimated by the LSM (Figure 3a) and by the WLSM (Figure 3b). It is clear that the LSM leads to a significant correlation

between the parameters, which is largely reduced with the WLSM. Moreover, the plots show that WLSM clearly outperforms the LSM in terms of accuracy. Further analyses (not shown) suggested that the square root of the set sizes considered to estimate the Tokunaga indices T_k (equation (18)) is a suitable choice for the weights. In this study we will therefore use the weighted least squares estimation procedure.

[48] The mean errors $\langle a - \hat{a} \rangle$ and $\langle c - \hat{c} \rangle$, where $\langle \cdot \rangle$ indicates the sample average, were statistically indistinguishable from zero in all experiments. In particular, for all the pairs of Tokunaga parameters the mean errors of \hat{a} ranged from -3×10^{-4} to -9×10^{-4} , whereas the mean errors of \hat{c} ranged from 1.6×10^{-2} to 2.2×10^{-2} . The reported deviations from zero are orders of magnitude smaller than the respective standard deviations, shown in Figure 4. This suggests that the proposed estimators are unbiased; or at least asymptotically unbiased if one takes into account the fact that all errors for each estimator have the same sign. Figure 4a shows the sample standard deviation $s_{\hat{a}}$ of \hat{a} , which is always relatively small, regardless of the Tokunaga parameters (a,c) . We also notice that $s_{\hat{a}}$ seems to be independent of a and it decreases as c increases. The sample standard deviation $s_{\hat{c}}$ of \hat{c} (panel b) is generally larger, ranging between 3% and 8% of the true value of c . The standard deviation $s_{\hat{c}}$ seems to increase as a function of $c-a$.

[49] Figure 4c shows the threshold value of the coefficient of determination, R_0^2 , that corresponds to the Monte Carlo confidence level of 5%. This means that the rejection criterion $R^2 < R_0^2$ resulted in rejecting the true null hypothesis about the Tokunaga self-similarity in 5% of the examined trees (50 out of 1000). The threshold value varies between 0.6 and 0.9; it tends to increase as $a+c$ increases. This behavior of R_0^2 reflects the fact that smaller values of a and c correspond to smaller sample sizes for $\tau_{i(i+k)}^{(l)}$ and hence to a larger scatter of the estimated coefficients \hat{T}_k . For the rejection criterion applied to real river networks in section 5 we will refer to $R_0^2 = 0.8$ as this was the average value calculated in these numerical simulations.

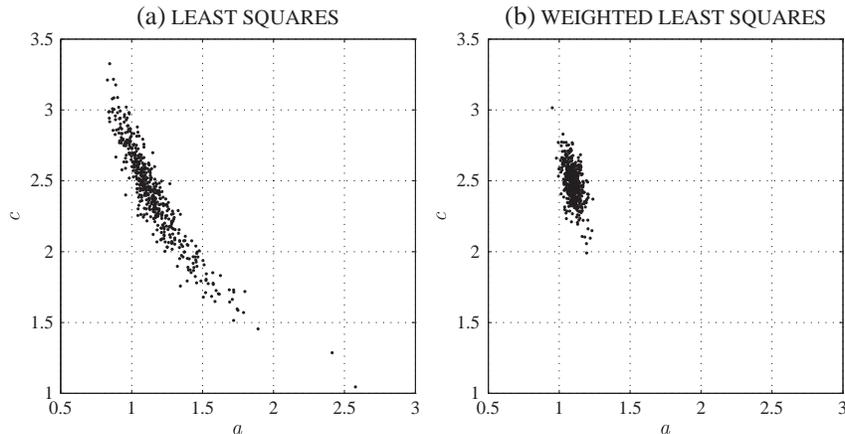


Figure 3. Estimation of the Tokunaga parameters (a, c) via the least squares method (panel a) and the proposed weighted least squares method (panel b). The weighted least squares method significantly reduces the dependence between a and c , as well as the variance in the estimations. True values of $a = 1.1$ and $c = 2.5$ were used for this simulation.

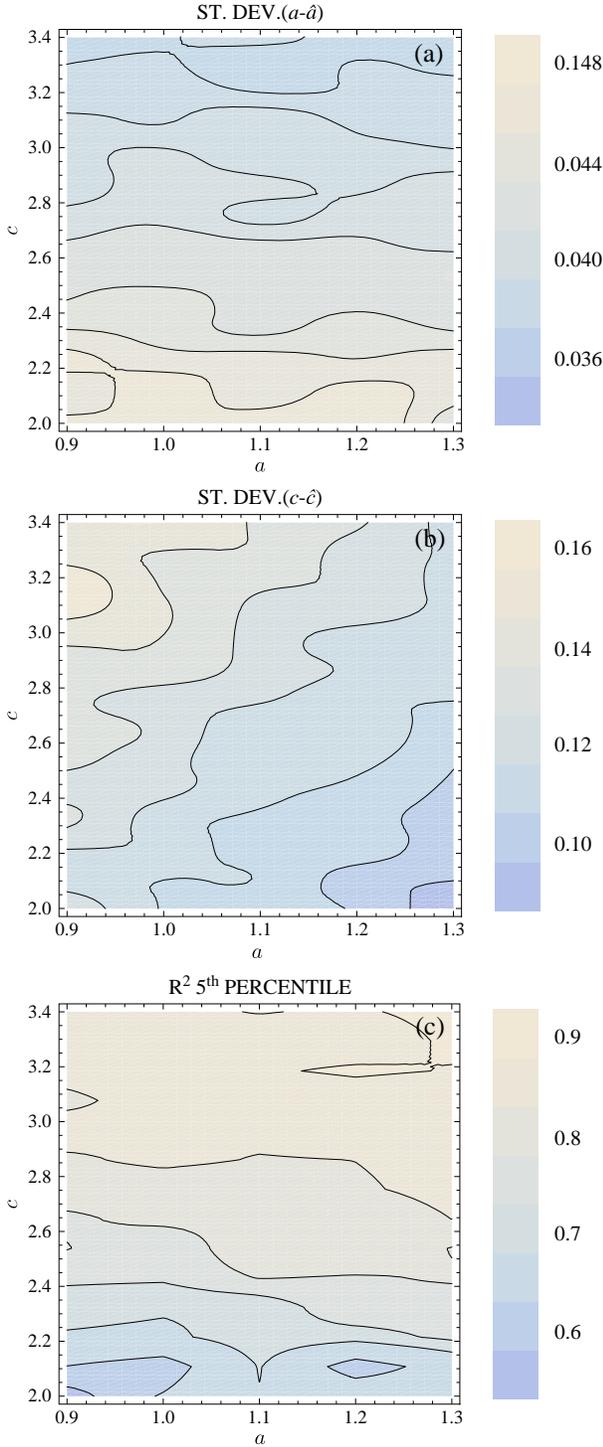


Figure 4. Standard deviations of the residuals ($a-\hat{a}$) and ($c-\hat{c}$), as well as the 5th percentile of the R^2 (equation (17)) estimated from simulations of synthetic TSS trees of order $\Omega = 7$ for a range of true parameters a and c .

4.3. Horton Ratios

[50] Figure 5 compares the distributions of the estimated Horton ratios in the simulated networks with the respective theoretical values. We used here the three estimation methods described in section 3.3: regression estimation, sample

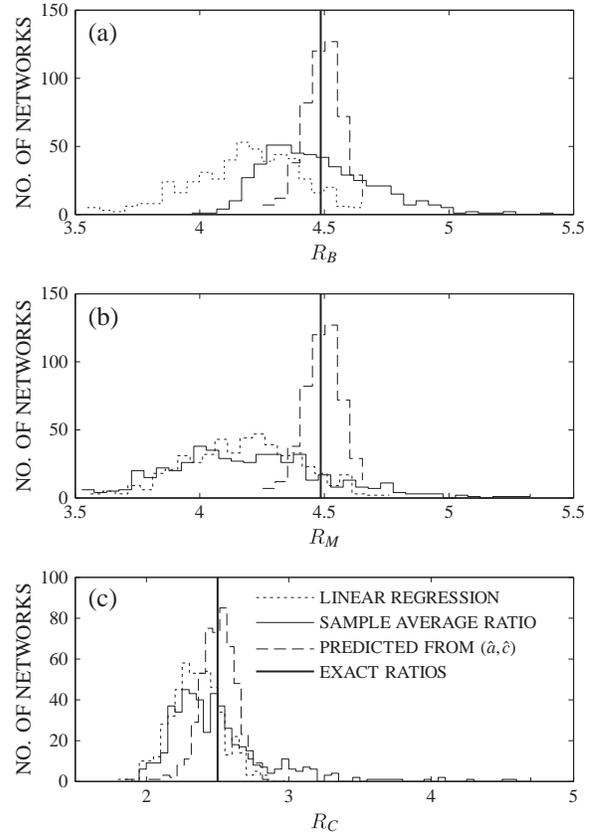


Figure 5. Distribution of the estimated Horton ratios for the simulated Tokunaga networks. The estimation was done with the regression method (dotted lines), the sample average ratio method (solid lines), and by applying the estimated Tokunaga parameters to the theoretical expressions of the Horton ratios (equations (12) and (13)) (dashed lines). The thick solid lines represent the exact Horton ratios obtained using the true values of a and c . See Section 3.3 for details on the methods.

average ratio estimation, and prediction from the estimated Tokunaga parameters. For this simulation we used 1000 Tokunaga networks with a single Tokunaga pair $a=1.1$ and $c=2.5$.

[51] The regression and average ratio methods seem to be equally effective and they both underestimate the theoretical Horton ratios. At the same time, the Horton ratios predicted from the estimated parameters \hat{a} and \hat{c} are significantly more accurate. This suggests that if a river network passes the TSS test, more accurate estimates of the Horton ratios of this network can be obtained by equations (12) and (13) than the conventional log-log regression.

[52] Figure 6 shows the estimated ratios N_{r+1}/N_r , M_{r+1}/M_r , and C_{r+1}/C_r as functions of the order r , for the 1000 synthetic Tokunaga networks. The figure illustrates that the uncertainty in the Horton ratios varies with the order r and clearly shows the bias in the estimation of R_M and R_C for low orders. As pointed out before, this bias is expected, since the convergence to these Horton ratios happens for high orders, see (6), (7). However, as noticed also by Peckham [1995a], the actual rate of convergence in the Horton laws is fast, and the limiting value is reached as fast as for $r \geq 3$. The deviations of the

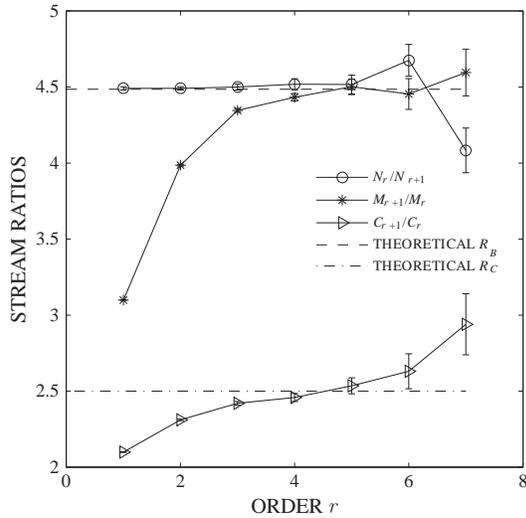


Figure 6. Estimated ratios N_r/N_{r+1} , M_{r+1}/M_r , and C_{r+1}/C_r as functions of the order r , for 1000 simulated Tokunaga networks with $a = 1.1$ and $c = 2.5$.

estimated R_C from the theoretical values for large orders is explained by the fact that the number of high-order branches is too small to produce sufficiently large samples of branch statistics. This effect further contributes to the uncertainty in the estimation of the Horton ratios.

[53] Results produced by these simulations will guide the calculation of the Horton ratios for real river networks (section 5). In particular, we will compare the Horton ratios obtained on the basis of the Tokunaga parameters with those obtained with the regression method as the sample average estimated method is shown to produce similar results.

5. Analysis of Real River Networks

5.1. Data

[54] We have analyzed 50 watersheds whose locations sample different climatic and geographic regions of the continental United States, as illustrated in Figure 7. The drainage areas of the examined basins range from 290 km² to 7200 km², and the Horton-Strahler orders of their stream networks range from 6 to 9. All the catchments belong to the MOPEX database; more information can be found at <ftp://hydrology.nws.noaa.gov/pub/gcip/mopex/>. For each watershed, we extracted the stream network from the 30 m Digital Elevation Models (DEMs) available at <http://seamless.usgs.gov/>.

[55] *Gesch* [2007] estimated a mean relative accuracy of 1.64 meters, based on 13,000 high precision survey points in the National Elevation Dataset. However, the relative accuracy for the DEMs we used was not reported except for a few cases. Studies have shown that, as the resolution of the DEM increases, so does the uncertainty of the elevation heights [e.g., *Thompson et al.*, 2001; *Erskine et al.*, 2007]. Therefore, while a higher resolution may increase the ability to capture small-scale terrain features, the uncertainty in their estimation may increase due to DEM’s errors. Besides, high resolution DEMs (e.g., from LiDAR) are not available yet for the large number of watersheds used in this study, rendering such a detailed analysis infeasible. In the following analysis it is shown that the extraction of topological characteristics of river networks does not require the identification of small-scale landscape features, such as channel heads, and that the estimated Tokunaga parameters are pretty robust to the DEM’s resolution.

[56] The river network extraction was obtained by applying the well-known criterion of the minimum contributing area threshold [e.g., *Tarboton et al.*, 1991], which determines whether a pixel corresponds to hillslope or stream depending

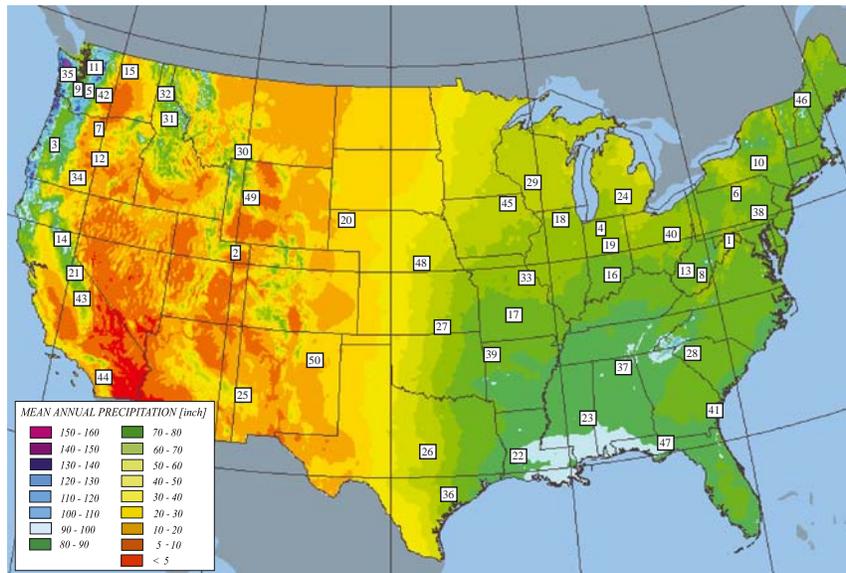


Figure 7. Location of the examined catchments. The spatial variability of the underlying climatic regimes is schematically represented through the mean annual precipitation. The map was obtained from The National Atlas of the United States of America, <http://nationalatlas.gov/>. Climatic and geographic characteristics for each catchment are reported in the supplementary material.

on whether its contributing area exceeds a fixed threshold. This method is generally accepted for determining channel heads from 30 m DEMs [e.g., *O’Callaghan and Mark, 1984; Band, 1986; Mark, 1988; Tarboton et al., 1991; Gardiner et al., 1991*]. However, especially when higher resolution DEMs are available, other methods have also been used to account for other geomorphological characteristics such as terrain slope and curvature [e.g., *Peucker and Douglas, 1975; Montgomery and Dietrich, 1992; Montgomery and Foufoula-Georgiou, 1993; Orlandini et al., 2003; Passalacqua et al., 2010*]. The main challenge with these methods is to predict the exact location of channel heads on the basis of certain thresholds, that vary both in space and time and whose estimation has been proven daunting. To test the effect of different thresholds on our study we repeated the analysis for threshold areas ranging from 0.05 to 0.5 km². Interestingly, we found that our results are only slightly affected by the area threshold, therefore we will show here only the results relative to a threshold of 0.1 km², while results relative to other thresholds can be found in the auxiliary material. The consistency of our results across a range of thresholds may be attributable to the fact that our primary focus is on the river network topology; therefore we do not necessarily need the exact location of channel heads, but, for our purposes, it suffices to capture channelized valleys regardless of where exactly the channels begin. Moreover, while it is reasonable to expect that random gain (or loss) of extracted low-order channels due to the variation of a threshold might affect the identification of the analyzed network properties, it should be noticed that the gain (or loss) due to the variation of the area threshold is not completely random, given the documented scaling of the drainage area distribution in river basins and its relationship with numerous network characteristics, such as Horton-Strahler order and magnitude [e.g., *Rodriguez-Iturbe and Rinaldo, 1997*]. Indeed, the properties we are analyzing (i.e., Tokunaga self-similarity, Tokunaga parameters, distribution of side-branches) are average scaling properties and our analysis shows that their estimation is robust to the scale (threshold area) used to define the first-order basins.

5.2. Results

5.2.1. Self-Similarity

[57] The self-similarity test was applied to each river network extracted from all the basins or subbasins of order $\Omega \geq 6$. Overall we analyzed 408 basins, of which 305 have order $\Omega=6$, 78 have order $\Omega=7$, 22 have order $\Omega=8$, and 3 have order $\Omega=9$. Figure 8a shows the distribution of the minimal P -values, P_0 , obtained in the ANOVA test of the self-similarity hypothesis.

[58] Accordingly, approximately 96% of the examined stream basins can be considered self-similar. The detailed test results are reported in Table 1; the values are obtained using the simulated corrected significance levels. It is interesting to note that as we increase the area threshold the fraction of networks that can be considered self-similar slightly increases (see also the auxiliary material).

5.2.2. Tokunaga Self-Similarity

[59] The networks that passed the ANOVA self-similarity test were further tested for the Tokunaga self-similarity. Figure 8b shows the distribution of the values of the coefficient of determination, R^2 . The rejection threshold was set to $R_0^2 = 0.8$ as this was the average value calculated from the

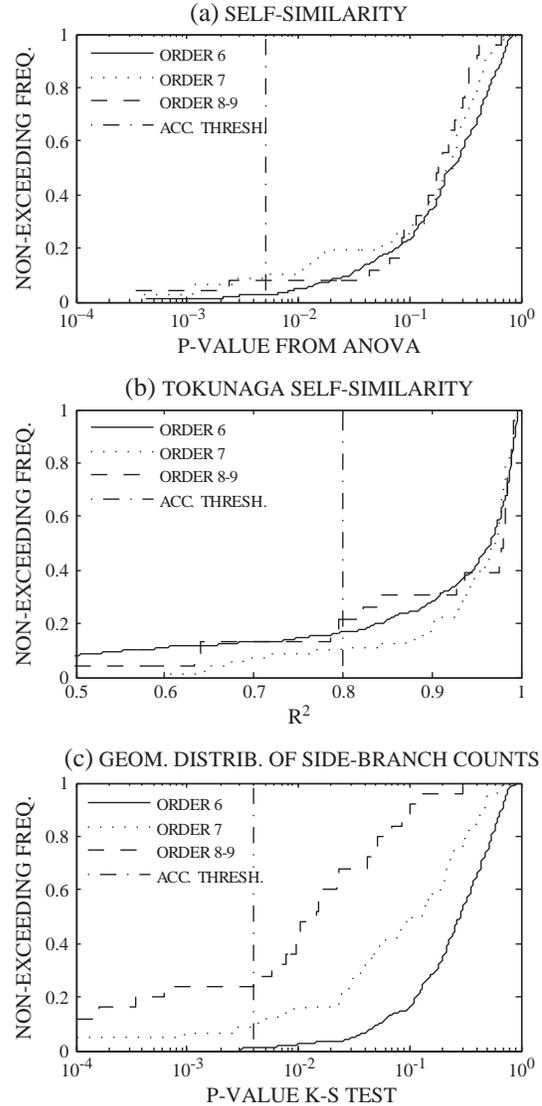


Figure 8. Empirical cumulative distribution function of (a) the P -values obtained in the ANOVA test of self-similarity, (b) the coefficients of determination, R^2 , obtained in the test for Tokunaga self-similarity, and (c) the P -values obtained with the Kolomogorv-Smirnov test of geometric side-branch counts. Results are reported separately for different network orders. The acceptance thresholds (ACC. THRESH. in the legend), shown in vertical dash-dotted lines, were obtained from the numerical simulations of TSS synthetic networks (with $\Omega = 7$) and were chosen as the 5th percentile of the statistics (i.e., ANOVA p-values, R^2 , and K-S p-value) calculated from the simulated networks.

Table 1. Number and Percentage of Self-Similar (SS) and Tokunaga Self-Similar (TSS) Networks, According to the Tests Described in Section 3

	# total	# SS	# TSS	# SS	# TSS
Order 6	305	298	247	98	83
Order 7	78	71	63	91	89
Order 8	22	20	16	90	80
Order 9	3	3	2	100	67
All orders	408	392	328	96	84

numerical simulations (see section 4 and Figure 4c). The detailed results of the test are summarized in Table 1. Analogously to the results on the self-similarity, the condition of TSS was accepted for a large percentage of the self-similar networks (approximately 84 %).

5.2.3. Tokunaga Parameters

[60] The Tokunaga parameters \hat{a} and \hat{c} , estimated using the weighted least squares method for the networks that have passed the Tokunaga self-similarity test, are shown in Figure 9; the respective mean and the standard deviation are reported in Table 2. The variability of the estimators is relatively low for the parameter a , which shows a standard deviation significantly lower than that of c . Moreover, as the basin order increases, the variability in the estimation of both a and c decreases. Values of a and c do not vary significantly as the area threshold changes (see auxiliary material) and, on average, we consistently observed a value of a around 1.1 and a value of c ranging between 2.6 and 2.7, although the spread around the mean value of c is larger for smaller-order basins. These are consistent with the few values reported in the literature; for example, *Tarboton* [1996] found ($a=0.98$; $c=2.89$) for the Buck Creek (NC) and *Peckham* [1995a] found ($a=1.2$; $c=2.4$) and ($a=1.2$; $c=2.7$) for the Kentucky River (KY) and the Powder River (WY), respectively.

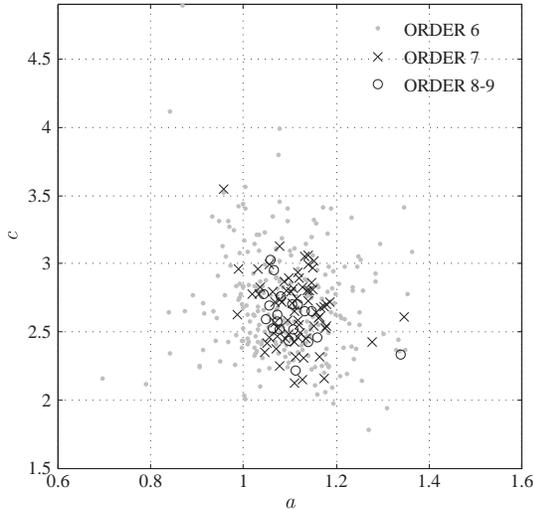


Figure 9. Parameter space of the estimated Tokunaga parameters \hat{a} and \hat{c} for the 50 catchments considered. Estimates are stratified by network order and the results indicate that as the order increases the parameter space reduces to a narrower range.

Table 2. Mean and Standard Deviation (in Parentheses) of the Tokunaga Parameters \hat{a} and \hat{c}

	\hat{a}	\hat{c}
Order 6	1.1(0.1)	2.7(0.40)
Order 7	1.1(0.06)	2.7(0.26)
Order 8	1.1(0.07)	2.6(0.20)
Order 9	1.1(0.01)	2.6(0.08)

5.2.4. Distribution of Side-Branch Counts

[61] Figure 8c shows the distribution of the P -values obtained with the Kolmogorov-Smirnov test for the geometric distribution of the side-branch counts $\tau_{ij}^{(l)}$. It is interesting to observe that the fraction of networks with geometrically distributed side-branches clearly decreases as the order of the network increases. This result may be due to spatial restrictions that come into play in large river basins, where physical constraints may be encountered, thus producing larger heterogeneity in sub-basin network topology. Moreover, the elongation of the basin which, according to the well known Hack’s law [Hack, 1957], occurs as the drainage area increases, might also limit the free development of network side-branching.

[62] It should be pointed out that, in the majority of the networks analyzed, the lowest P -value corresponds to the count n_1 and in many cases this was the only P -value lower than the selected threshold, while all the other side-branch counts of the network successfully passed the Kolmogorov-Smirnov test. Therefore, if the count n_1 was to be neglected, significantly more networks would pass the test. Indeed, as we increase the area threshold in the network extraction, which generally implies the removal of the smaller-order links, we observe that for a larger number of networks the hypothesis of geometrically distributed side-branches is satisfied (see the auxiliary material). Table 3 summarizes the results of the Kolmogorov-Smirnov test. This result is important as it implies that, for the majority of the river networks, the only difference in the distribution of side-branches is attributable to the Tokunaga parameters. Moreover, this observation justifies the use of a geometric distribution of side-branches for the simulation of synthetic river networks.

5.2.5. Horton Ratios

[63] For the networks that passed the Tokunaga self-similarity test, we estimated the Horton ratios R_B , R_M and R_C using the regression approach (the sample average estimated method was shown to produce similar results to the regression approach, thus it will not be used here, see section 3.3) and compared these estimates with the ratios computed from the estimated Tokunaga parameters using (12) and (13). The distributions of the ratio estimates with the two methods are shown in Figure 10.

[64] As expected from the numerical simulations, the regression-based estimates tend to be lower than those computed from the parameters \hat{a} and \hat{c} . The latter, according to the simulations, are closer to the true values of the ratios, that is the values that are computed from equations (12) and (13) using the assigned values of a and c .

[65] Based on this analysis, the most reliable ranges for the stream ratios (i.e., the ones produced by the Tokunaga parameters) are $3.7 \leq R_B = R_M \leq 5.7$ and $1.8 \leq R_C \leq 4.2$. It is

Table 3. Number and Percentage of Networks With Geometrically Distributed Side-Branches (NGDS)

	# total	# NGDS	% NGDS
Order 6	305	301	99
Order 7	78	71	91
Order 8	22	18	82
Order 9	3	0	0
All orders	408	390	96

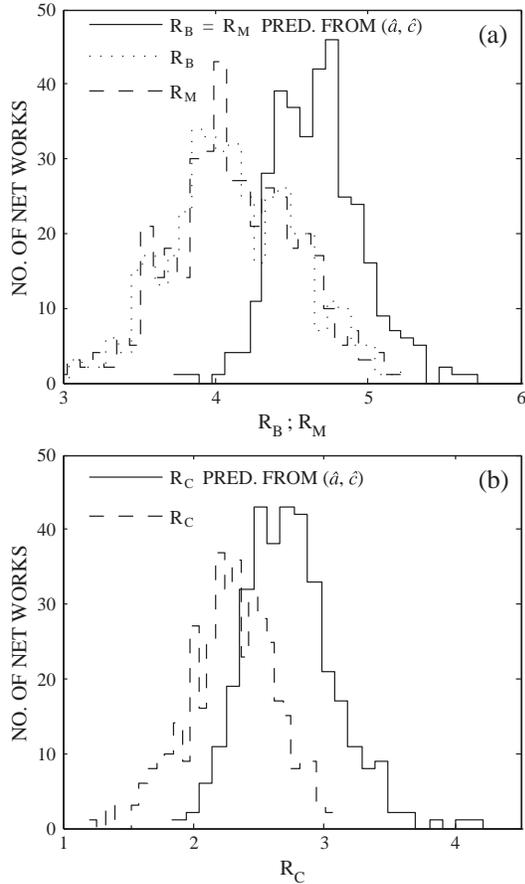


Figure 10. Frequency histograms of the Horton ratios estimated by the least squares method (dashed and dotted lines) and predicted from the estimated Tokunaga parameters (solid lines) for the 408 networks considered. Panel (a) reports $R_B = R_M$, panel (b) reports R_C .

interesting to compare our ranges to those estimated by *Mantilla et al.* [2010] using the expressions obtained by *Veitzer and Gupta* [2000] and *Troutman* [2005] in their random self-similar model for river stream networks. The ranges reported by *Mantilla et al.* [2010] (i.e., $4.3 \leq R_B \leq 4.8$, $2.3 \leq R_C \leq 2.7$), which are based on 30 networks of orders 7 and 8 across the US, are comparable with ours. The larger variability of our estimators is possibly due to the fact that we considered a larger number of networks and also many networks of order $\Omega=6$, which have been shown to bear a larger topological variability. If we remove the 6-order networks from our analysis the ranges of the stream ratios become $4.2 \leq R_B = R_M \leq 5.1$ and $2.1 \leq R_C \leq 3.5$. The mean of $R_B = R_M$ is 4.64 with a standard deviation of 0.19 and the mean of R_C is 2.67 with a standard deviation of 0.27.

6. Climatic Dependence of Tokunaga Parameters

[66] A question of significant interest is what physical properties of a basin determine the river network structure. While some studies have addressed this question focusing on geometric characteristics of river networks such as the drainage density [e.g., *Leeder*, 1993; *Tucker and Bras*, 1998], no studies so far have reported relations between physical properties of a basin and river network topology.

Indeed, different basin properties may affect how river networks are topologically structured such as topography, geology, soil composition, vegetation, and climate. A full exploration of these controls requires an analysis that is beyond the goals of this paper. However, given the availability of extensive climatic data for the MOPEX catchments, in this section we explore whether a correlation exists between climate and the Tokunaga parameters.

[67] A major effort of the MOPEX project has been to assemble high quality historical hydrometeorological datasets for a wide range of river basins [*Schaake et al.*, 2006]. The selected basins had to fulfill rigorous requirements and a critical aspect was to have research quality estimates for mean areal precipitation; for example, a minimum number of rain gauges, as a function of the basin area, was defined according to *Schaake et al.* [2000]. The trustworthiness of this data-set is further strengthened by its use in many recent studies [e.g., *Zanardo et al.*, 2012 and references therein].

[68] The hydro-climatic properties we focus on in this analysis are the average basin wetness at the mean annual scale and the climatic storminess at the storm-event scale. The following indicators were used as surrogates for these hydro-climatic characteristics: (1) the mean annual rainfall volume, P ; (2) the mean storm frequency, λ (i.e., the inverse of the average number of days between the beginning of consecutive storm events); and (3) the mean storm duration, Δ . These quantities were computed for each basin using 40-50 years of daily rainfall series available in the MOPEX data-set for each of the catchments considered.

[69] Interestingly, it was found that the parameter a did not exhibit significant correlation with any of these climatic indicators. This is not surprising given that a (i.e., the mean number of k -order branches that drain into $(k+1)$ -order branches) expresses only the 1-order difference between merging channels (i.e., the most basic and robust property of a directional hierarchical tree), and that it varies within a pretty narrow range even for the large range of networks analyzed (see Figure 9 and Table 2).

[70] On the contrary, the parameter c expresses the disproportionality between the number of low- and high-order channels, and it thus represents both the small- and large-scale features of the network. It is important to note that a large value of c indicates a large number of low-order channels with respect to the number of high-order channels and therefore it characterizes a more “feathered” network. Given the importance of the lower-order channels in the value of c and given that their configuration involves smaller time scales and is more dependent on external forcing compared to that of higher order channels, it is reasonable to expect a dependence between c and the hydro-climatology of the basin, which is what our analysis indicates. Specifically, Figure 11 shows the relationship between each of the examined hydro-climatic indicators and the Tokunaga parameter c estimated for all the analyzed basins. The results were stratified to depict (a) only networks where the TSS hypothesis was accepted with an R^2 larger than 0.98, (b) only networks where the TSS was accepted with $R^2 > 0.9$, and (c) all the networks that exhibited Tokunaga self-similarity in our analysis (based on the threshold of 0.8 identified in the simulations). This was done to differentiate between three levels of confidence with which the TSS hypothesis was accepted. It should also be noted that for this

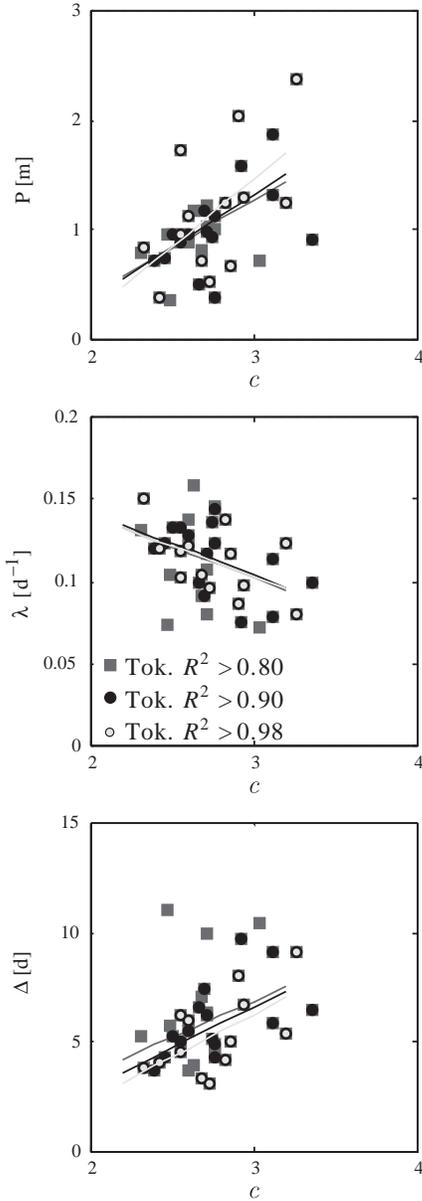


Figure 11. Correlations between the Tokunaga parameter c and the three climatic variables considered: (a) mean annual rainfall, P , (b) mean storm frequency, λ , (c) mean storm duration, Δ . The solid lines represent the linear interpolation of each set of points.

analysis we only considered 6-order networks. Given that there is only one rainfall data-set for each basin, each point in the plots of Figure 11 represents the average c among the values computed for all the 6-order sub-basins in each basin. This choice is motivated by the fact that larger networks are more likely to encounter physical constraints but also to cover different geological or tectonic regions, which might somehow cloud the climatic control that is explored in this analysis.

[71] Figure 11 suggests that c might have a nonnegligible correlation with the three hydro-climatic variables P , λ and Δ . Table 4 reports the values of the Pearson’s linear correlation coefficient for each of the cases considered as well as the associated P -values, which express the probability of

Table 4. Pearson’s Linear Correlation Coefficients, r , and the Associated P -values (in Parentheses) for the Tokunaga Parameter c and the Three Climatic Variables Considered: Mean Annual Rainfall Volume, P ; Mean Rainfall Frequency, λ ; and Mean Storm Duration, Δ ^a

	P	λ	Δ
$R^2 > 0.80$	0.51(0.001)	-0.41(0.012)	0.40(0.013)
$R^2 > 0.90$	0.53(0.004)	-0.51(0.005)	0.57(0.002)
$R^2 > 0.98$	0.58(0.037)	-0.51(0.072)	0.59(0.03)

^aThe correlation coefficients are computed after progressively refining the dataset for an increasing value of the coefficient of determination, R^2 , calculated to test the TSS hypothesis (section 3.2).

obtaining the observed nonzero correlation coefficient in the case when the correlation is actually zero. The coefficients were found to be significant at 5% level for all the cases considered except for the dependence on λ in the networks with $R^2 > 0.98$. Interestingly, as the confidence in the TSS of the network increases (i.e., as we increase the lower bound of R^2), the correlation increases.

[72] While the correlation values are not extremely high, further statistical tests show that they are in fact considerably robust. In particular, the significance of these correlations was confirmed using the following three approaches: (1) bootstrap, (2) permutation, and (3) randomization analysis. Details of this analysis are given in Appendix 1 and attest the statistical significance as well as the robustness of the reported correlations.

[73] It should be noticed that, in general, climate can have a long term-effect on topographic characteristics, such as slope and elevation, which, in turn, might affect the river network topology. Therefore, the observed correlation between climate and network topology might in fact include, or act as a surrogate, of the possible dependence of network topology on topography. To isolate the direct effect of topography on topology, we explored the dependence of the parameter c on six catchment-averaged topographical characteristics. The six characteristics used are: mean, standard deviation and maximum value of elevation and slope. Results are reported in the auxiliary material. Interestingly, while these topographic indicators are correlated with the climatic variables, their correlations with the parameter c are smaller (and less significant) than those of the climatic variables. This further suggests that climate may leave its distinct signature not only on the average features of a landscape but also on the finer dissection of the landscape as expressed by the topological structure of river networks.

[74] It is important to point out that the highest values of the correlation coefficients (ranging from 0.44 to 0.8) are obtained using the area threshold of 0.09 km² for the network extraction. As the area threshold increases, these values tend to decrease and eventually, for the area threshold larger than 0.3 km², no apparent correlation is observed between c and the climatic variables considered. This is not surprising as the effect of climate on topology is more likely associated with the organization of channels at the smaller (say 1 to 3 order watersheds) rather than larger spatial scales; a large threshold misses these small scales and makes it hard to discern possible climatic controls.

[75] The presented results are in agreement with one’s intuitive expectations. Indeed, the results show that in the

catchments where rainfall is distributed in a few strong events we observe higher values of c compared to the catchments with less rainfall that is more uniformly distributed in time. Recall now that as c increases, the number of low-order branches that merge to the high-order branches also increases; this creates an enhanced “feathering” of the network. We find it reasonable that a more regular climatic forcing generates a more regular landscape while scattered strong events may cause a disproportional hierarchy of the network branching.

[76] To better illustrate the concept of feathering described by the parameter c we show in Figure 12 the comparison between a 6-order network with $c=1.8$, located within the Kickapoo river basin (catchment number 29 in Figure 7, WI), and a 6-order network with $c=4.9$, located within the Methow river basin (catchment number 15 in Figure 7, WA). The climatic variables are $\Delta=3.8$ d, $\lambda=0.15$ d⁻¹, and $P=830$ mm for the former, and $\Delta=6.5$ d, $\lambda=0.1$ d⁻¹, and $P=916$ mm for the latter. We chose this pair of networks because their difference is readily visible, however, this difference is probably attributable not only to climate. Indeed the two landscapes have also different physical characteristics: the former, which is underlain by sedimentary rock, is relatively flatter and with lower elevation than the latter, which is underlain by plutonic rock and has steeper slopes. We can observe two general features common to many of the networks considered: (a) the merging angle between two branches is usually wider for larger c 's and (b) when the values of c are higher, there is usually more space between branches of the same order. Both these observations suggest that when c is large, the branches of the same order that run relatively parallel and close to each other are rare. Therefore, according to the observed relationship between c and the climatic variables it is reasonable to make the hypothesis that

strong storm events, in the long term, might cause (or be one of the causes for) channels that run by each other to collapse, so that only one survives and the room left by the other one is filled by smaller order channels, which in turn increases the value of c . A comprehensive understanding of these effects as well as proving (or disproving) this hypothesis might require analysis of landscape evolution under controlled laboratory experiments [e.g., *Bonnet and Crave, 2003*] or numerical experiments via landscape evolution models, which is beyond the scope of this study.

7. Discussion and Conclusions

[77] The study was motivated by the growing interest in the Tokunaga model for river network topology. This model, while commonly accepted, usually lacks a proper justification based on formal tests and extensive data analysis. We propose here (i) rigorous tests for self-similarity and Tokunaga self-similarity of a tree, and (ii) accurate estimation procedures for the Horton ratios and the Tokunaga parameters (a, c). The proposed tests are based on the classical concepts of the analysis of variance (ANOVA) [*Turner and Thayer, 2001*] and the least squares goodness-of-fit. To the best of our knowledge this is the first time a formal set of techniques is suggested (and validated in synthetic tree simulations) for analysis of essential topological properties of trees including the Horton laws, topological self-similarity, and the Tokunaga self-similarity. Our results are thus not limited to hydrogeomorphic applications, but they are readily applicable to a wide range of phenomena that obey the Tokunaga self-similarity. Such phenomena include, but are not limited to, vein structure of botanical leaves [*Newman et al., 1997; Turcotte et al., 1998*], diffusion limited aggregation [*Ossadnik, 1992; Masek and Turcotte, 1993*], percolation

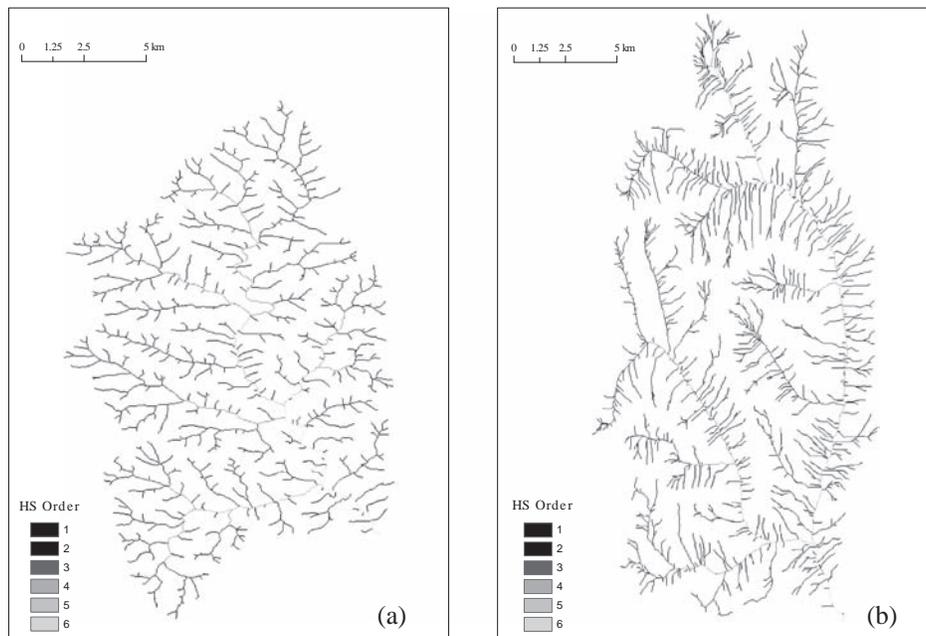


Figure 12. Example of two 6-order networks with drastically different values of the Tokunaga parameter c . Panel (a) shows a network with $c = 1.8$; panel (b) shows a network with $c = 4.9$. The two basins are located within the Kickapoo river basin (catchment number 29 in Figure 7, WI) and the Methow river basin (catchment number 15 in Figure 7, WA), respectively.

[Turcotte *et al.*, 1999; Yakovlev *et al.*, 2005; Zaliapin *et al.*, 2006a, 2006b], branching processes [Burd *et al.*, 2000], as well as nearest-neighbor clustering in Euclidean spaces and tree representation of time series [Zaliapin and Kovchegov, 2011, and references therein].

[78] In this work, we used the proposed statistical methodologies to evaluate the assumptions of self-similarity (SS) and Tokunaga self-similarity (TSS) and estimate several topological properties for 50 catchments across the continental United States. At each catchment, we analyzed all basins and sub-basins of orders equal or higher than $\Omega = 6$. Overall, we analyzed 408 individual networks. An important result of this study is that both the SS and TSS hypotheses cannot be rejected for the majority of the examined networks and the levels of acceptability of the SS and TSS hypothesis are independent on the network order. On the other hand, the analysis of the side-branch distribution shows that the percentage of networks with geometrically distributed side-branches clearly decreases as the order increases. A similar observation was reported by Jarvis and Sham [1981], who consistently detected a downstream change in network structure in all the networks they analyzed. As they argued, this kind of pattern reflects the spatial requirements of different tributary sizes: in large networks, the competition between tributaries of different sizes becomes an important control parameter in the development of a network. Moreover, the basin elongation prescribed by the Hack’s law, might itself constitute a significant limitation to the free development of side-branches as the area of the basin increases.

[79] The estimated Tokunaga parameters (\hat{a} , \hat{c}) demonstrate a relatively low variability, which increases as the basin order decreases. The variability of \hat{a} is significantly smaller than that of \hat{c} . Recall that a is the mean number of side-branches of order i that merge with a branch of order $(i + 1)$; its relative constancy across a large number of networks suggests that a represents a “global” property of river networks, as opposed to the parameter c , which seems to be more related to local characteristics. In particular we showed how the value of c defines the “feathering” property of river networks, which is well illustrated in Figure 12. While this property is not related to the network geometrical features, such as the network shape, the parameter c can be used to classify river networks on the basis of their topology.

[80] Our results may be useful for improving current numerical network modeling. For instance, the Random Self-similar Network (RSN) model introduced by Veitzer and Gupta [2000] is probably the most versatile among the existing models. Based on recursive local replacement of the network generators, it reproduces a broad range of network properties, including, under particular conditions, the Tokunaga self-similarity. However, in the Tokunaga networks simulated by the RSN model, the Tokunaga parameters are constrained by the relationship $c - a = 1$ [Veitzer and Gupta, 2000]. It is therefore important to observe that, according to our analysis, (i) the difference $c - a$ does not seem to be constant, and (ii) the average difference between the Tokunaga parameters is approximately 1.55.

[81] We have found that for a relatively high percentage of the networks analyzed, the hypothesis of geometric distribution of side-branch counts cannot be rejected. This finding appears particularly important from the perspective of stochastic network modeling. Cui *et al.* [1999] proposed a

stochastic generalization of the deterministic Tokunaga model by assuming a random number of side-branches extracted from a negative binomial distribution (NBD), which is a generalization of the geometric distribution. While the mean of the NBD is completely defined by the Tokunaga parameters, the distribution also depends on a third parameter which, according to Cui *et al.* [1999], represents the spatial variability in the network topology. The fact that a geometric distribution (completely specified by a single parameter) of side-branches is not inconsistent with the majority of the examined networks is indeed promising as, under the geometric assumption, the modeling framework is completely characterized by the two Tokunaga parameters.

[82] We tested the relationship between climate and the topological properties of the river networks and found statistically significant correlations between certain hydro-climatic variables (namely, mean annual rainfall, mean storm frequency, and mean storm duration) and the Tokunaga parameter c . Correlation values are not very high ($r \approx 0.4 - 0.6$) but various numerical tests establish their high statistical significance. Moreover, the interpretations of the observed correlations are in agreement with one’s physical intuition. It is worth noticing that this is a preliminary result that shows the possibility of the Tokunaga parameters to serve as metrics for establishing connections between processes forming the landscape and the topology of the developed fluvial networks. This certainly calls for further analysis, as for example in this study we did not consider other controls such as tectonics, geology or soil composition, which possibly have a competing effect to climate forcing on the resulting river network topology.

[83] The reported correlations between the topological structure of river networks and climate might have a direct application in problems where climatic data are not available whereas landscape data are. An outstanding example is the so-called asymmetric seasonality between Titan’s hemispheres hypothesized by Aharonson *et al.* [2009]. In particular the authors suggest that a difference in the satellite’s climate (wetter in the Northern hemisphere and drier in the Southern hemisphere) might be the cause of the observed different geomorphology of the two hemispheres. Presence of drainage networks has been repeatedly observed in images collected by the Huygens probe [Tomasko *et al.*, 2005] as well as by the Cassini spacecraft [Elachi *et al.*, 2005]. Given our results, a topological analysis of these drainage networks as suggested by the present study might provide further insight onto Aharonson *et al.*’s [2009] hypothesis.

Appendix A: Tests on the Significance of the Correlations Between c and the Climatic Variables

[84] To further confirm the significance of the correlations between the TSS parameter c and the climatic variables considered, we performed simulations via three approaches: (1) bootstrap, (2) permutation, and (3) randomization analysis. Specifically, let c_i and V_i denote, respectively, the estimated Tokunaga parameter \hat{c} and a particular climatic variable V_i for the i -th basin, and $r_V = \text{corr}(c_i, V_i)$ be the respective estimated correlation, for $i = 1, \dots, N$, N being the number of basins considered.

[85] Bootstrap analysis estimates the uncertainty of the observed correlations [Efron and Tibshirani, 1994]. A set of

$N_{\text{boot}} = 10,000$ random paired samples of size N is constructed by random drawing *with replacement* from the observed pairs (c_i, V_i) . The random samples hence deviate from the original one in that they might have some repeating pairs (c_i, V_i) . The correlation $r_j^{\text{boot}}, j = 1, \dots, N_{\text{boot}}$ is computed for each simulated bootstrap sample. The sample variance of the bootstrap sample, $\text{Var}(r_i^{\text{boot}})$, is a good approximation for the variance of the correlation r_i . Figure 13a shows the histogram of the bootstrap correlations for the mean storm frequency. The 95% confidence interval of the bootstrap correlations is well separated from 0. This suggests that the observed correlation $r = 0.4$ is significant.

[86] Permutation analysis evaluates how likely it is to obtain the observed correlation by chance. A set of $N_{\text{perm}} = 10,000$ random paired samples of size N is constructed by randomly

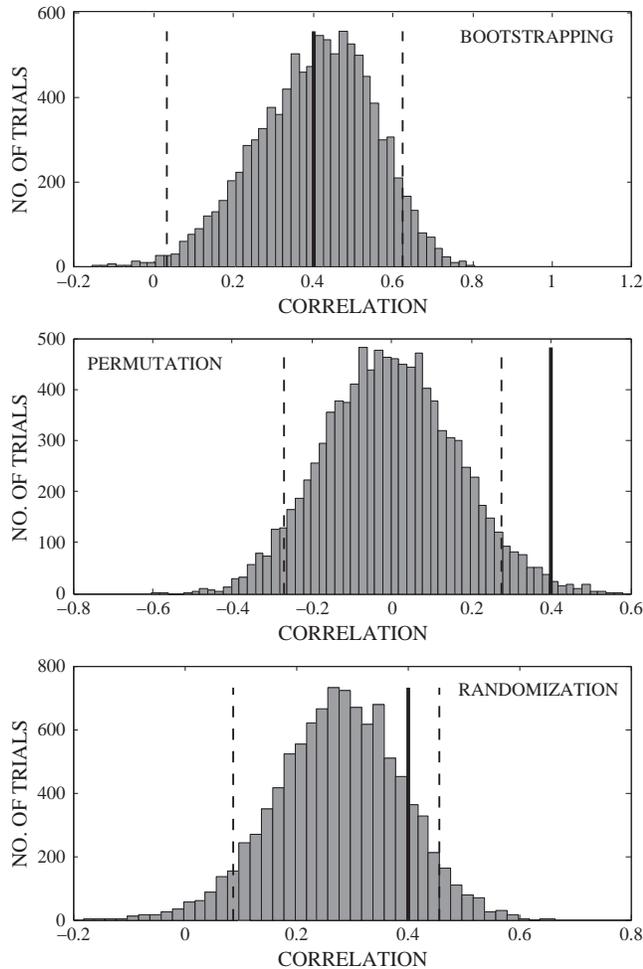


Figure 13. Tests on the robustness of the linear correlation between the Tokunaga parameter c and the mean rainfall frequency, λ , for all the networks that resulted Tokunaga self-similar from our analysis. The tests used are bootstrapping (top panel), permutation (middle panel) and randomization of c (bottom panel). Dashed lines represent the 95% confidence interval; the solid lines represent the correlation value obtained from data. These methods were used also to test the robustness of the linear correlation between c and mean annual rainfall and between c and the mean storm duration, analogous results were obtained.

matching the values of c_i and $V_i, i = 1, \dots, N$. The correlation $r_j^{\text{perm}}, j = 1, \dots, N_{\text{boot}}$, is computed for each simulated permutation sample. Since this method destroys possible relationships between the paired values (c_i, V_i) , the average permutation correlation should be zero. Figure 13b shows the histogram of the permutation correlations for the mean storm frequency. The probability of obtaining the permutation correlation r^{perm} greater than the observed correlation $r = 0.4$ is 0.08; the 95% confidence interval for the permutation correlation is well separated from the observed value. This suggests that the observed correlation $r = 0.4$ is significant.

[87] The uncertainties (variances) of the bootstrap and permutation correlation help one to evaluate how likely it is to obtain the observed correlation r_i by chance in a situation when c_i and V_i are in fact independent. Both analyses suggest that such a possibility is negligible.

[88] Randomization analysis evaluates the effect of the uncertainty in the estimation of the Tokunaga parameter c on the correlation r_i . For each basin we generated $N_{\text{rand}} = 10,000$ values for c from a Normal distribution with the mean equal to the observed value c_i and standard deviation estimated from the numerical simulations in section 2, and evaluated the correlation of such values with the climatic variable. Figure 13c shows the histogram of the randomized correlations for the mean storm frequency. The 95% confidence interval for the randomized correlation is well separated from 0; this interval, moreover, does cover the observed correlation value $r = 0.4$.

[89] The results for other examined climatic variables are very similar to those shown in Figure 13 for the mean storm frequency, therefore they are not shown.

[90] The results of Figure 13, which suggest that the observed correlations are statistically significant and can be explained neither by sample fluctuations (bootstrap and permutation analyzes) nor by the uncertainties in the estimation of c_i (randomized analysis). We hence assume that the correlation are due to the actual dependence between the Tokunaga parameter and the climatic characteristics. While the observed correlations are not extremely high (especially when considering all the TSS networks), they are quite robust thus providing a solid justification for a climatic effect on river network topology.

[91] **Acknowledgments.** Bill Dietrich and Dave Milledge are gratefully acknowledged for numerous discussions and insightful feedbacks. We would like to thank the participants of the 2012 Oregon State University Workshop on Mathematical Problems in the Environmental Sciences for their comments on the earlier version of this paper. This work has been supported by NSF's CMG collaborative research project "Envirodynamics on River Networks", by grants EAR-0934628 (to E.F.G.), and EAR-0934871 (to I.Z.), as well as by the National Center for Earth surface Dynamics (NCED), a Science and Technology Center funded by NSF under agreement EAR-0120914. E.F.G also acknowledges the support via the Ling Professorship at the University of Minnesota.

References

- Aharonson, O., A. Hayes, J. I. Lunine, R. D. Lorenz, and C. Elachi (2009), An asymmetric distribution of lakes on Titan as a possible consequence of orbital forcing, *Nat. Geosci.*, 2, 851–854.
- Band, L. E. (1986), Topographic partitioning of watersheds with digital elevation models, *Water Resour. Res.*, 22, 15–24.
- Bonnet, S., and A. Crave (2003), Landscape response to climate change: Insights from experimental modelling and implications for tectonic versus climatic uplift of topography, *Geology*, 31(2), 123–126.

- Bracquemond, C., E. Cretois, and O. Gaudoin (2011), A comparative study of goodness-of-fit tests for the geometric distribution and application to discrete time reliability. <http://www-ljk.imag.fr/SMS/ftp/BraCreGau02.pdf>. [Online; accessed 11-May-2011].
- Burd, G. A., E. C. Waymire, and R. D. Winn (2000), A self-similar invariance of critical binary Galton-Watson trees, *Bernoulli*, 6, 1–21.
- Cui, G., B. Williams, and G. Kuczera (1999), A stochastic Tokunaga model for stream networks, *Water Resour. Res.*, 35(10), 3139–3147.
- Dodds, P. S., and D. H. Rothman (2000), Scaling, universality, and geomorphology, *Ann. Rev. Earth Planet. Sci.*, 28, 571–610, doi:10.1146/annurev.earth.28.1.571.
- Efron, B., and R. J. Tibshirani (1994), *An Introduction to the Bootstrap*, 456, CRC, Boca Raton, Fla.
- Elachi, C., et al. (2005), Cassini Radar views the surface of Titan, *Science*, 308, 970–974.
- Erskine, R. H., T. R. Green, J. A. Ramirez, and L. H. McDonald (2007), Digital elevation accuracy and grid cell size: effects on estimated terrain attribute, *Soil Sci. Soc. Am. J.*, 71, 1371–1380.
- Furey, P. R., and B. M. Troutman (2008), A consistent framework for Horton regression statistics that leads to a modified Hack's law, *Geomorphol.*, 102(3–4), 603–614, doi:10.1016/j.geomorph.2008.06.002.
- Gardiner, T. W., K. C. Sasowsky, and R. L. Day (1991), Automated extraction of geomorphic properties from digital elevation data, *Z. Geomorphol. Suppl.*, 80, 57–68.
- Gesch, D. B. (2007), The National Elevation Dataset, in *Digital Elevation Model Technologies and Applications—The DEM Users Manual*, 2nd ed., edited by D. Maune, chap. 499–118, Am. Soc. for Photogram. and Remote Sensing, Bethesda, Md.
- Hack, J. T. (1957), Studies of longitudinal stream profiles in Virginia and Maryland, *U.S. Geol. Surv. Prof. Pap.*, 294-B, 1–97.
- Horton, R. E. (1945), Erosional development of streams and their drainage basins: Hydrophysical approach to quantitative morphology, *Geol. Soc. Am. Bull.*, 56, 275–370.
- Kiffney, P. M., C. M. Greene, J. E. Hall, and J. R. Davies (2006), Tributary streams create spatial discontinuities in habitat, biological productivity, and diversity in mainstream rivers, *Can. J. Fish. Aquat. Sci.*, 63, 2518–2530, doi:10.1139/F06-138.
- Kirkby, M. J. (1976), Tests of the random network model, and its application to basin hydrology, *Earth Surf. Processes*, 1, 197–212.
- Leeder, M. R. (1993), Tectonic controls upon drainage basin development, river channel migration and alluvial architecture: Implications for hydrocarbon reservoir development and characterization, in *Characterisation of Fluvial and Aeolian Hydrocarbon Reservoirs*, edited by C. North and J. Prossor, Geol. Soc. London Spec. Publ., 73, 9–24.
- Mantilla, R., B. M. Troutman, and V. K. Gupta (2010), Testing statistical self-similarity in the topology of river networks, *J. Geophys. Res.*, 115, F03038, doi:10.1029/2009JF001609.
- Mark, D. M. (1988), Network models in geomorphology, in *Modeling Geomorphological Systems*, edited by M. G. Anderson, 73–97, John Wiley, New York.
- Masek, J. G., and D. L. Turcotte (1993), A diffusion limited aggregation model for the evolution of drainage networks, *Earth Planet. Sci. Lett.*, 119, 379.
- McConnell, M., and V. Gupta (2008), A proof of the Horton law of stream numbers for the Tokunaga model of river networks, *Fractals*, 16, 227–233.
- Menadbe, M., S. E. Veitzer, V. K. Gupta, and M. Sivapalan (2001), Tests of peak flow scaling in simulated self-similar river networks, *Adv. Water Resour.*, 24(9), 1037–1049.
- Montgomery, D. R., and W. E. Dietrich (1992), Channel initiation and the problem of landscape scale, *Science*, 255, 826–830.
- Montgomery, D., and E. Foufoula-Georgiou (1993), Channel network source representation for digital elevation models, *Water Resour. Res.*, 29(12), 3925–3934.
- Muneepeeraikul, R., E. Bertuzzo, H. J. Lynch, W. F. Fagan, A. Rinaldo, and I. Rodriguez-Iturbe (2008), Neutral metacommunity models predict fish diversity patterns in Mississippi-Missouri basin, *Nature*, 453, 220–222, doi:10.1038/nature06813.
- Newman, W. I., D. L. Turcotte, and A. M. Gabrielov (1997), Fractal trees with side branching, *Fractals*, 5, 603–614.
- Nikulin, M. S. (1992), Gihman statistic and goodness-of-fit tests for grouped data, *C. R. Math. Rep. Acad. Sci. Can.*, 14(4), 151–156.
- O'Callaghan, J., and D. Mark (1984), The extraction of drainage networks from digital elevation data, *Comput. Vision Graphics Image Process.*, 28, 328–344.
- Orlandini, S., G. Moretti, M. Franchini, B. Aldighieri, and B. Testa (2003), Path-based methods for the determination of nondispersive drainage directions in grid-based digital elevation models, *Water Resour. Res.*, 39(6), 1144, doi:10.1029/2002WR001639.
- Ossadnik, P. (1992) Branch order and ramification analysis of large diffusion limited aggregation clusters, *Phys. Rev. A*, 45, 1058.
- Passalacqua, P., T. Do Trung, E. Foufoula-Georgiou, G. Sapiro, and W. E. Dietrich (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, S. (1995a), New results for self-similar trees with applications to river networks, *Water Resour. Res.*, 31(4), 1023–1029.
- Peckham, S. (1995b), Self-similarity in the three-dimensional geometry and dynamics of large river basins, PhD thesis, Univ. of Colo., Boulder, Colorado.
- Peckham, S., and V. K. Gupta (1999), A reformulation of Horton's laws for large river networks in terms of statistical self-similarity, *Water Resour. Res.*, 35(9), 2763–2777.
- Peucker, T. K., and D. H. Douglas (1975), Detection of surface-specific points by local parallel processing of discrete terrain elevation data, *Comput. Graphics Image Process.*, 4, 375–387.
- Power, M. E., and W. E. Dietrich (2002), Food webs in river networks, *Ecol. Res.*, 17, 451–471.
- Rodriguez-Iturbe, I., and A. Rinaldo (1997), *Fractal River Basins: Chance and Self-Organization*, Cambridge Univ. Press, New York.
- Rodriguez-Iturbe, I., and J. B. Valdes (1979), The geomorphologic structure of hydrologic response, *Water Resour. Res.*, 15, 1409–1420, doi:10.1029/WR015i006p01409.
- Schaake, J. C., Q. Duan, M. Smith, and V. Koren (2000), Criteria to select basins for hydrologic model development and testing, Preprints, 15th Conf. On Hydrology, 10–14 January 2000, Paper P1.8, Am. Meteorol. Soc., Long Beach, Calif.
- Schaake, J., S. Cong, and Q. Duan (2006), The US MOPEX data set, IAHS-AISH Publ., 307, 9–28.
- Sklar, L. S., W. E. Dietrich, E. Foufoula-Georgiou, B. Lashermes, and D. Bellugi (2006), Do gravel bed river size distributions record channel network structure?, *Water Resour. Res.*, 42, W06D18, doi:10.1029/2006WR005035.
- Stewart-Koster, B., M. J. Kennard, B. D. Harch, F. Sheldon, A. H. Arthington, and B. J. Pusey (2007), Partitioning the variation in stream fish assemblages within a spatio-temporal hierarchy, *Mar. Freshwater Res.*, 58, 675–686.
- Strahler, A. N. (1957), Quantitative analysis of watershed geomorphology, *Trans. AGU*, 38, 913–920.
- Tarboton, D. (1996), Fractal river networks, Horton's laws and Tokunaga cyclicity, *J. Hydrol.*, 187, 105–117.
- Tarboton, D. G., R. L. Bras, and I. Rodriguez-Iturbe (1991), On the extraction of channel networks from digital elevation data, *Hydrol. Processes*, 5, 81–100.
- Thompson, J. A., J. C. Bell, and C. A. Butler (2001), Digital elevation model resolution: Effects on terrain attribute calculation and quantitative soil-landscape modeling, *Geoderma*, 100, 67–89, doi:10.1016/S0016-7061(00)00081-1.
- Tokunaga, E. (1966), The composition of drainage network in Toyohira river basin and valuation of Horton's first law (in Japanese with English summary), *Geophys. Bull. Hokkaido Univ.*, 15, 1–19.
- Tokunaga, E. (1978), Consideration on the composition of drainage networks and their evolution (1978), Geograph. Rep. 13, 1–27 Tokyo Metro. Univ., Japan.
- Tomasko, M. G., et al. (2005), Rain, winds and haze during the Huygens probe's descent to Titan's surface, *Nature*, 438, 765–778.
- Troutman, B. M. (2005), Scaling of flow distance in random self-similar channel networks, *Fractals*, 13(4), 265–282.
- Tucker, G. E., and R. L. Bras (1998), Hillslope processes, drainage density, and landscape morphology, *Water Resour. Res.*, 34(10), 2751–2764.
- Turcotte, D. L., J. D. Pelletier, and W. I. Newman (1998), Networks with side branching in biology, *J. Theor. Biol.*, 193, 577–592.
- Turcotte, D. L., B. D. Malamud, G. Morein, and W. I. Newman (1999), An inverse cascade model for self-organized critical behavior, *Physica A*, 268, 629–643.
- Turner, J. R., and J. F. Thayer (2001), *Introduction to Analysis of Variance: Design, Analysis and Interpretation*, 192, Sage Publ., Thousand Oaks, Calif.
- Veitzer, S., and V. K. Gupta (2000), Random self-similar river networks and derivations of generalized Horton laws in terms of statistical simple scaling, *Water Resour. Res.*, 36, 1033–1048.
- Veitzer, S., and V. K. Gupta (2001), Statistical self-similarity of width function maxima with implications to floods, *Adv. Water Resour.*, 24(9–10), 955–965, doi:10.1016/S0309-1708(01)00030-6.
- Westfall, P. H., R. D. Tobias, and R. D. Wolfinger (2011), *Multiple Comparisons and Multiple Tests Using the SAS System*, 2nd ed., SAS Publ., Cary, N. C.
- Yakovlev, G., W. I. Newman, D. L. Turcotte, and A. Gabrielov (2005), An inverse cascade model for self-organized complexity and natural hazards, *Geophys. J. Int.*, 163, 433–442.

- Zaliapin, I., and Y. Kovchegov (2011), Tokunaga and Horton self-similarity for level set trees of Markov chains, *Chaos Solitons Fractals*, *45*, 358–372, doi:10.1016/j.chaos.2011.11.006.
- Zaliapin, I., H. Wong, and A. Gabrielov (2006a), Inverse cascade in percolation model: Hierarchical description of time-dependent scaling, *Phys. Rev. E*, *71*, 066118.
- Zaliapin, I., H. Wong, and A. Gabrielov (2006b), Hierarchical aggregation in percolation model, *Tectonophysics*, *413*, 93–107.
- Zanardo, S., C. Harman, P. A. Troch, S. C. Rao, and M. Sivapalan (2012), Intra-annual rainfall variability controls on inter-annual variability of catchment water balance: A stochastic analysis, *Water Resour. Res.*, *48*(6), doi:10.1029/2010WR009869.