Mathematical characterization of fractal river networks

Author: Akira Terui[1,*]

Email: hanabi0111@gmail.com

[1]Department of Biology, University of North Carolina at Greensboro, Greensboro, NC 27410, USA

Maintext

Rivers form complex branching networks, and the ecological implication of river network complexity has gained a great interest over the past few decades[1–3]. To this end, there have been concerted efforts to construct virtual river networks, aiming to provide theoretical insights into how river network structure controls riverine ecological dynamics. Carraro and Altermatt[4] made a great contribution to this research area by comparing the scaling properties of virtual river networks produced by three different simulation methods - balanced binary trees (BBTs)[5], random branching networks (RBNs)[6,7], and optimal channel networks (OCNs)[8]. The first two methods have two parameters that control river network size (order or the number of nodes) and complexity (branching probability). In the meantime, OCNs are spanning trees that suffice a local minimum of a function describing total energy expenditure across the network. Carraro and Altermatt[4] performed extensive analysis and concluded that: (1) OCNs best represent the scaling properties of river networks, such as Horton's law and the exceedance probability of catchment area; (2) branching ratio is a scale-dependent quantity as the value changes across spatial resolutions at which river networks are extracted (expressed as the threshold catchment area A_T that initiates channels or pixel size l); (3) OCNs most accurately predicted the metapopulation stability and capacity. I agree that OCNs are capable of reproducing the scaling properties of real river networks. However, the supporting ground for the rest of the conclusions is, unfortunately, seriously flawed or rather inconclusive due to the improper mathematical definition or use of scale invariance, dimensions, and units.

First, the term "scale invariance" was falsely used in their article. To explain this, let y^* and y be the perimeter length of a perfect circle and the length of a coastline, respectively, measured as multiples of a ruler with a unit length x (the "observation scale") – a classical comparison between "scale-dependent" and "scale-invariant" objects used in Mandelbrot[9]. The approximated length measured by the ruler is $q(x)x$, where $q(x)$ is the number of dividing steps with a given ruler length x. In the case of perimeter length y^* (scale-dependent), one will obtain the exact value as the ruler length x approaches zero because a sufficiently short ruler can characterize the smooth curve of a perfect circle (the existence of characteristic scale). Clearly, $\lim_{x \to 0} q(x)x = 2\pi r$ (r is the radius) for the perimeter of a perfect circle. In contrast, the length of a complex coastline y (scale-invariant) reaches infinity as the scale x approaches zero because a shorter ruler can capture similar, but finer-scale complexity that is ignored by a longer ruler (the lack of characteristic scale; see Figure 1 in Mandelbrot[9] for the graphical comparison between the two types of objects). This property, i.e., the part is a reduced-scale image of the whole, is referred to as "statistical self-similarity" that underpins the mathematical characteristics of scale-invariance in the coastline length y. Such a scale-invariant structure can be eptiomized by the power-law function[10,11]:

$$y = f(x) = cx^z$$

where c is the scaling constant and z is the scaling exponent. The scale invariance can be easily proved by multiplying the scaling factor λ:

$$y' = f(\lambda x) = c\lambda^z x^z = \lambda^z f(x) = \lambda^z y$$

Thus, the observed object $y = f(x)$ is said to be scale-invariant because the multiplicative extension/shrink of observation scale x results in the same shape of the original object y but with a different scale[12].

Carraro and Altermatt[4] provided some evidence that branching ratio p_r follows a power law along the axes of observation scale A_T and pixel size l (i.e., length on a side) using OCNs (Equation 1 in the original article):

$$\mathrm{E}(p_r) \approx 1.531 A_T^{-0.523} A^{-0.032}$$

where A is the entire watershed area expressed in the number of pixels; thus, $A = A'l^{-2}$ (A' is the watershed area measured in the unit of l^2, such as km^2). Also, the authors empirically showed "visual" examples that the relationship between p_r and A_T follows a power-law using the data from 50 real river networks (Figure 3a in the original article). Provided that this relationship statistically holds true (see the next section for this issue), branching ratio has a property of scale-invariance. Despite this, the authors claimed that "*Here we show that an alleged property of such random networks (branching probability) is a scale-dependent quantity that does not reflect any recognized metric of rivers' fractal character...*" (Abstract) This interpretation is the opposite as it has been defined in the literature of scale-invariance[10–13], including the author's previous publication[14]. Therefore, the sentences that build upon this interpretation are inappropriate and misleading (**Table 1**). Importantly, many of these sentences are the concluding sentence of a paragraph. As such, the progression of supporting sentences in these paragraphs is no longer logical, unfortunately.

Second, dimensions and units are improperly treated in their analysis. A dimension is the power of an axis along which a physical quantity is measured, and a unit is a way to assign a number to a particular dimension to make it relative. For example, length is a dimension and meter is a unit of length. Throughout the article, the authors used the number of pixels to measure the river length N, the total catchment area A, and the threshold catchment area A_T. There is no issue to use the number of pixels as a unit. However, a critical problem in their article is that they obscured the dimensions of pixels. For example, they made very unclear that the river length N has a dimension of pixel "length", while the total catchment area A and the threshold catchment area A_T have a dimension of pixel "area". In particular, the authors incorrectly defined p_r as a dimensionless quantity (Methods) despite its unit is [pixel length^{-1}] with a dimension of length.

Their analysis is acceptable for OCNs in which pixel size is constant. However, the improper use of units and dimensions causes serious problems in the analysis of "real" river networks; the authors used different pixel sizes (length on a side) among watersheds (103 m to 1268 m; Supplementary Table 2 in the original article), meaning that the same number of pixels translates into very different lengths and areas. For example, in Figure 3 in the original article, the the observation scale A_T ranges $20 - 500$ pixels. This pixel range translates into $0.2 - 5.3$ km^2 for the Toss watershed (with smallest pixel size) and $32.2 - 803.9$ km^2 for the Stikine watershed (with largest pixel size). Further, the branching ratio p_r is also affected by this variation in pixel size as its unit is [pixel length^{-1}]. When the pixel length was converted to a unit of km^{-1}, the unit ranges from 0.1 km^{-1} (Toss) to 1.3 km^{-1} (Stikine). Hence, the authors compared values that are uncomparable.

To explore the consequences of this improper use of pixels, I re-analyzed their data with MERIT Hydro that has a constant pixel size of 3 arc-second (~90 m at the equator) across the globe[15]. I extracted river networks in R[16] using the R package `whitebox` with ten values of A_T [km^2] ($A_T = 1, ..., 1000$ with an equal interval at a $\log_{10}$ scale). Following the author's definition of branching ratio, I calculated branching ratio p_r as $\frac{N_L}{L}$ (N_L the number of links in a network [dimensionless]; L the total river length [km]) for each river network extracted at a given observation scale A_T. In the original article, the authors did not perform any statistical analysis to prove the power-law despite the fact that the estimates of p_r are subject to statistical uncertainty, just like any other geospatial layers. Instead, they relied on a subjective argument by "picking" a specific value of p_r estimated in a given watershed at a given observation scale A_T (Table 1), claiming the rank of p_r varies across scales. To avoid such a subjective argument, and to statistically substantiate the power-law of p_r along the axis of A_T, I fitted the following log-linear models (i represents an individual data point of p_r estimated in a given watershed at a given scale A_T).

$$\log_{10} p_{r,i} = c_{w(i)} + z \log_{10} A_{T,i} + \varepsilon_i \qquad \text{(M0)}$$
$$\log_{10} p_{r,i} = c_{w(i)} + z_{w(i)} \log_{10} A_{T,i} + \varepsilon_i \quad \text{(M1)}$$

The first model (M0) assumes the "universal" scaling with the constant scaling exponent z across watersheds; i.e., the branching ratio at all the 50 watersheds follows the same power law with the watershed-specific scaling constant $c_{w(i)}$ ($w(i)$ is watershed w for a data point i). The second model (M1) assumes the "local" scaling with the watershed-specific scaling exponent $z_{w(i)}$; i.e., the branching ratio at all the 50 watersheds follows the same power law. I compared the evidence ratio of the two models using the approximated Bayes Factor (BF)[17], which defined as $\text{BF} = \exp\left(\frac{\text{BIC(M1)} - \text{BIC(M0)}}{2}\right)$. In this definition, a value of $\text{BF} > 1$ gives the support for the universal model M0 over M1; for example, if $\text{BF} = 2$, the model M0 is twice as likely as the model M1.

a

recreated Figure 3a using km^2 as a unit of the threshold catchment area[18].

A dimension is a discrete value (e.g., one or two) when we measure a perfect geometry: the length of a straight line has a dimension of one, and the area of a squire has a dimension of two. In contrast, fractal objects exhibit non-discrete values of dimensions[9,10]; for example, in the case of Koch curve, it has a dimension $D \approx 1.26$[10].

This correction clarified that the authors compared branching probability extracted at observation scales that differ two orders in km^2 (Figure 1). As a result, the observation scale A_T rarely changed the rank of branching probability (compare p_r with the same A_T in km^2), as opposed to the author's statement "*by extracting different river networks at various scales (i.e., various A_T values) and assessing the rivers' rank in terms of p_r, one observes that rivers that look more "branching" (i.e., have higher p_r) than others for a given A_T value can become less "branching" for a different A_T value (Fig. 3).*" Therefore, the reported results in the original article are merely an artifact of inconsistent units of A_T across watersheds. Importantly, this serious issue is pertinent to their metapopulation simulations, as they used pixel size l as a unit of river distance. In their real river examples, the distance of 10 pixels varies from 1030 m to 12680 m among watersheds. The results of their metapopulation simulations rest heavily on the measured distance between nodes because they used exponential curves (i.e., sensitive to units) to approximate distance decay in population synchrony and dispersal. Therefore, it is very likely that substantial errors are contained in Figure 6 and Supplementary Figures 4 – 9 of the original article that underpin their main conclusion that OCNs most accurately predicted the metapopulation stability and capacity. Although the authors explained that the number of pixels among real watersheds were aligned to make them comparable to virtual river networks (Methods), they instead could have tweaked the virtual river networks to avoid this fundamental problem.

Carraro and Altermatt[4] offered an important perspective to the use of the different classes of virtual river networks. Therefore, a full re-interpretation/re-analysis is warranted to provide valid insights into how we assess river network structure.

Data availability

There is no new data associated with this manuscript. Codes are available at https://github.com/aterui/public-proj_carraro-cee

Reference

1. Grant, E. H. C., Lowe, W. H. & Fagan, W. F. Living in the branches: Population dynamics and ecological processes in dendritic networks. *Ecology Letters* **10**, 165–175 (2007).

2. Carrara, F., Altermatt, F., Rodriguez-Iturbe, I. & Rinaldo, A. Dendritic connectivity controls biodiversity patterns in experimental metacommunities. *Proceedings of the National Academy of Sciences* **109**, 5761–5766 (2012).

3. Mari, L., Casagrandi, R., Bertuzzo, E., Rinaldo, A. & Gatto, M. Metapopulation persistence and species spread in river networks. *Ecology Letters* **17**, 426–34 (2014).

4. Carraro, L. & Altermatt, F. Optimal Channel Networks accurately model ecologically-relevant geomorphological features of branching river networks. *Communications Earth & Environment* **3**, 1–10 (2022).

5. Yeakel, J. D., Moore, J. W., Guimarães, P. R. & de Aguiar, M. A. M. Synchronisation and stability in river metapopulation networks. *Ecology Letters* **17**, 273–283 (2014).

6. Terui, A. *et al.* Metapopulation stability in branching river networks. *Proceedings of the National Academy of Sciences* **115**, E5963–E5969 (2018).

7. Terui, A., Kim, S., Dolph, C. L., Kadoya, T. & Miyazaki, Y. Emergent dual scaling of riverine biodiversity. *Proceedings of the National Academy of Sciences* **118**, e2105574118 (2021).

8. Carraro, L. *et al.* Generation and application of river network analogues for use in ecology and evolution. *Ecology and Evolution* **10**, 7537–7550 (2020).

9. Mandelbrot, B. How long is the coast of Britain? Statistical self-similarity and fractional dimension. *Science* **156**, 636–638 (1967).

10. Rodríguez-Iturbe, I. & Rinaldo, A. *Fractal River Basins: Chance and Self-Organization.* (Cambridge University Press, 2001).

11. Brown, J. H. *et al.* The fractal nature of nature: Power laws, ecological complexity and biodiversity. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* **357**, 619–626 (2002).

12. Proekt, A., Banavar, J. R., Maritan, A. & Pfaff, D. W. Scale invariance in the dynamics of spontaneous behavior. *Proceedings of the National Academy of Sciences* **109**, 10564–10569 (2012).

13. Rinaldo, A., Rigon, R., Banavar, J. R., Maritan, A. & Rodriguez-Iturbe, I. Evolution and selection of river networks: Statics, dynamics, and complexity. *Proceedings of the National Academy of Sciences* **111**, 2417–2424 (2014).

14. Giometto, A., Altermatt, F., Carrara, F., Maritan, A. & Rinaldo, A. Scaling body size fluctuations. *Proceedings of the National Academy of Sciences* **110**, 4646–4650 (2013).

15. Yamazaki, D. *et al.* MERIT hydro: A high-resolution global hydrography map based on latest topography dataset. *Water Resources Research* **55**, 5053–5073 (2019).

16. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. (2021).

17. Wagenmakers, E.-J. A practical solution to the pervasive problems of p values. *Psychonomic Bulletin & Review* **14**, 779–804 (2007).

18. Carraro, L. CompareRiverNetworks. https://zenodo.org/record/6472920. (2022).