Trees, networks, and hydrology

Andrea Rinaldo,[1] Jayanth R. Banavar,[2] and Amos Maritan[3,4]

Received 14 March 2005; revised 8 July 2005; accepted 20 October 2005; published 11 April 2006.

[1] This paper reviews theoretical and observational material on form and function of natural networks appeared in somewhat disparate contexts from physics to biology, whose study is related to hydrologic research. Moving from the exact result that drainage network configurations minimizing total energy dissipation are stationary solutions of the general equation describing landscape evolution, we discuss the properties and the dynamic origin of the scale-invariant structure of river patterns and its relation to optimal selection. We argue that at least in the fluvial landscape, nature works through imperfect searches for dynamically accessible optimal configurations and that purely random or deterministic constructs are clearly unsuitable to properly describe natural network forms. We also show that optimal networks are spanning loopless configurations only under precise physical requirements that arise under the constraints imposed by continuity. In the case of rivers, every spanning tree proves a local minimum of total energy dissipation. This is stated in a theorem form applicable to generic networks, suggesting that other branching structures occurring in nature (e.g., scale-free and looping) may possibly arise through optimality to different selective pressures. We thus conclude that one recurrent self-organized mechanism for the dynamic origin of fractal forms is the robust strive for imperfect optimality that we see embedded in many natural patterns, chief and foremost hydrologic ones.

Citation: Rinaldo, A., J. R. Banavar, and A. Maritan (2006), Trees, networks, and hydrology, *Water Resour. Res.*, *42*, W06D07, doi:10.1029/2005WR004108.

1. Introduction

[2] We address the problem of the dynamic origin of network structures that we observe in nature, whose recurrent patterns of organization are the subject of much recent research [e.g., *Watts and Strogatz*, 1998; *Watts*, 1999; *Barabási and Albert*, 1999; *Albert and Barabási*, 2002; *Barabási et al.*, 2000, 2002; *Dorogovtsev and Mendes*, 2002; *Newman*, 2003; *Dodds et al.*, 2003; *Song et al.*, 2005]. It seems appropriate, in this context, to review results relevant to hydrology. We note, in fact, that the line of reasoning proposing the key role of the selection of locally optimal (and thus dynamically accessible) structures originated from the study of rivers [*Rodriguez-Iturbe et al.*, 1992b, 1999c; *Rinaldo et al.*, 1992; *Rodriguez-Iturbe and Rinaldo*, 1997], and that several related developments have appeared in physical and biological contexts, thus possibly perceived as unrelated to hydrology. An annotated review, providing a coherent framework for seemingly disparate results and addressing open questions also for hydrologic research, is thus possibly useful.

[3] Our basic claim, supported by a number of facts rooted in theory and observation that we briefly recall in this paper, is that the hydrologic case for the selection of tree-like networks (as opposed to looping) is rather compelling [*Rodriguez-Iturbe and Rinaldo*, 1997]. Moreover, the beauty, diversity and deep symmetries of fractal river networks are fascinating, and well-understood, signatures of how nature works [*Mandelbrot*, 1977, 1983; *Bak*, 1996], and thus the search for analog mechanisms at work in other physical and biological contexts seems valuable. Natural and artificial network patterns, however, show a great variety of forms and functions and many do not show the tree characters (i.e., a unique path from any site to the outlet) that rivers exhibit. One thus wonders what is the basic dynamic reason for radically different forms and functions.

[4] Figures 1, 2, 3, and 4 illustrate a sample of the above variety. A reference framework for different types of hydrologic networks is meant to show that departures from trees and tree-like networks spanning a given area (that is, L^D sites with $D = 2$) commonly arise. A choice of real and abstract structures relevant to hydrology is proposed. A real fluvial network (Figure 1a) of the Dry Tug Fork river (CA) is suitably extracted from a digital terrain map [*Rodriguez-Iturbe and Rinaldo*, 1997]. Notice its clear tree-like structure, usual in the runoff production zone of the river basin. Its morphological features (like aggregation and elongation) are typical of fluvial patterns and recurrent modules appear regardless of the scale of total contributing area, such that the parts and the whole are quite similar notwithstanding local signatures of geologic controls, here marked by a fault line clearly visible across the landscape.

[1] Dipartimento di Ingegneria Idraulica, Marittima, Ambientale e Geotecnica and International Centre for Hydrology "Dino Tonini," Universitá di Padova, Padua, Italy.
[2] Department of Physics, Pennsylvania State University, University Park, Pennsylvania, USA.
[3] Dipartimento di Fisica "Galileo Galilei," Università di Padova, Padua, Italy.
[4] Also at Istituto Nazionale per la Fisica della Materia, Genoa, Italy.

0043-1397/06/2005WR004108

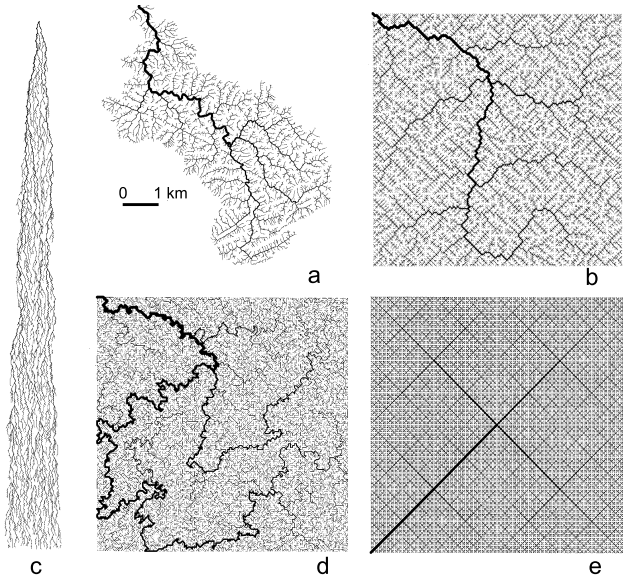

Figure 1. Samples of trees where a unique pattern links any inner site to the outlet of the network: (a) A real river network, the Dry Tug Fork (California), suitably extracted from digital terrain maps [*Rodriguez-Iturbe and Rinaldo*, 1997]; (b) a single-outlet optimal channel network (OCN) selected starting from an arbitrary initial condition by an algorithm accepting random changes only of lowering the total energy dissipation of the system as a whole, thus incapable of reaching the ground state and settling in a local minimum dynamically accessible [*Rodriguez-Iturbe et al.*, 1992b]; (c) Scheidegger's construct [*Scheidegger*, 1967]; (d) a "hot" OCN where any arbitrary change randomly assigned to an evolving network is accepted provided it maintains a tree-like form [*Rodriguez-Iturbe and Rinaldo*, 1997]; (e) Peano's construct [*Peano*, 1890; *Mandelbrot*, 1983; *Marani et al.*, 1991].

Scheidegger's [1967, 1970] directed network (Figure 1c) is constructed by a stochastic rule; with even probability, a walker chooses between right or left forward sites only. The model was devised with reference to drainage patterns of an intramontane trench and maps exactly into a model of random aggregation with injection or voter models [e.g., *Takayasu et al.*, 1988, 1991a, 1991b] and also describes the time activity of a self-organized critical [*Bak et al.*, 1987] (Abelian) avalanche [*Dhar*, 1999]. *Peano's* [1890] network (Figure 1e), is a deterministic fractal [*Mandelbrot*, 1983], whose main topological and scaling features, some involving exact multifractals [*Marani et al.*, 1991], have been solved analytically [*Marani et al.*, 1991; *Colaiori et al.*, 1997]. The basic prefractal is a cross seeded in a corner of the square domain that covers the cross and its ensuing iterations. All subsequent subdivisions cut in half each branch to reproduce the prefractal on four, equal subbasins. Here the process is shown at the 11th stage of iteration. Optimal channel networks (OCNs)(Figure 1b) are described in section 3. They hold fractal characteristics that are obtained through a specific selection process from which one obtains a rich structure of scaling optimal forms that are known (as discussed in section 3) to closely conform to the scaling of real networks even in the case of unrealistic geometric boundaries [*Rodriguez-Iturbe et al.*, 1992b, 1992c]. To design a very inefficient tree, we have constructed a nondirected structure constrained to be tree-like (Figure 1d) by using a Metropolis algorithm (see section 3).

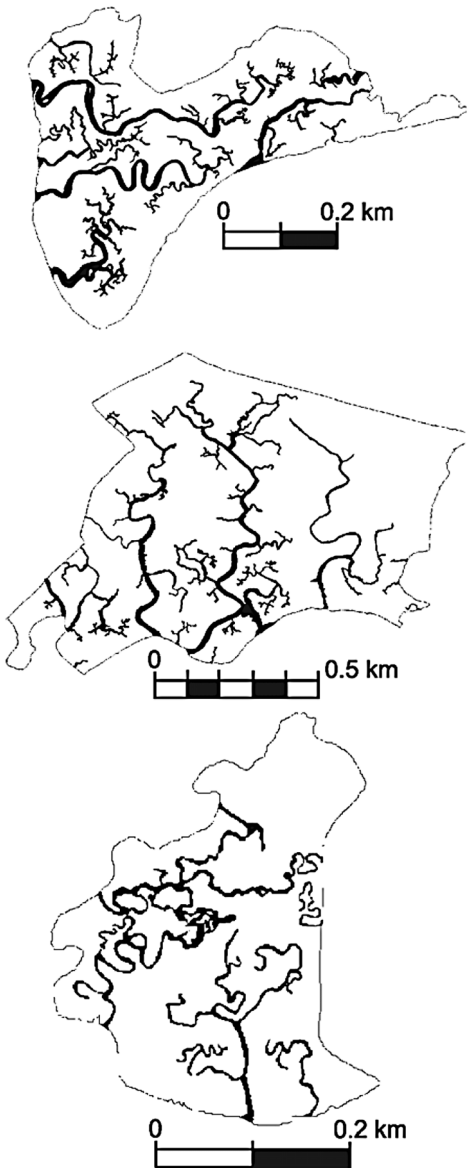

Figure 2. Drainage patterns that form in the tidal landscape are much diverse according to local conditions and develop over a different range of scales. In addition, they may or may not develop loops. Here we show three tidal networks developed within saltmarshes of the lagoon of Venice suitably extracted from remote sensing [*Feola et al.*, 2005]. The tidal networks are observed a few km away in space and roughly within the same microtidal range, have similar scales, and have very different aggregation. Sinuosities of the tidal meanders vary greatly from site to site, possibly reflecting the age of the saltmarsh [*Marani et al.*, 2002], and the drainage density describing the average distance one has to walk before encountering a creek within a saltmarsh is widely different from site to site [*Marani et al.*, 2003]. The presence of loops much depends on local conditions, and is not necessarily affected by flood or ebb dominance.

Figure 3. A small-scale drainage pattern developed within a macrotidal environment, here the Eden estuary in Scotland. Channeled pathways are extracted by suitable digital image processing techniques (courtesy of E. Belluco). Here we see an organized sequence of regular ditches reminiscent of the organization of trenches draining into a complex, looping network structure. Even at eyesight, one catches the lack of a fundamental similarity of the parts and the whole that characterizes river networks.

This basically corresponds to an algorithm that accepts any change attributed sequentially at random sites of an evolving spanning network; an existing link is disconnected at a random site and rewired randomly to another nearest neighbor provided the change maintains a tree-like structure [*Rodriguez-Iturbe and Rinaldo*, 1997]. Hot OCNs (Figure 1d),so-called because they correspond to high temperatures of a Metropolis scheme where every spanning tree seeded in the outlet is equally likely, are thus abstract forms meant to reproduce a rather undirected tree.

[5] Loops are also observed in the fluvial landscape. Powerful or weak tidal forcings (Figures 2 and 3) introduce preferential scales related to crossovers of processes operating with comparable ranges into an otherwise aggregated pattern. However (Figure 4), deltaic networks arising from the interplay of distributive drainage patterns and of marine and littoral processes can substantially alter the similarity of the parts and the whole. One notes, in fact, that the structure of deltaic networks shows major loops, differently from rivers in runoff-producing areas. Note also that differences in the sinuosity of the branches prevent any detailed statistical similarity of the parts and the whole across scales [*Marani et al.*, 2003; *Feola et al.*, 2005].

[6] Legitimate questions naturally arise by looking at the structures in Figures 1–4: why should loopless trees develop? Are the observed landforms random structures? Are different network configurations equally probable? If not, to what selective pressure do they respond? Are there universal features shown by fluvial landforms, and what is their proper characterization? The fluvial network context alone thus proposes the need for some settlement with respect to the general dynamic origin of network scale invariance, possibly toward an understanding of a general framework for the processes of network growth and selection.

[7] This paper is organized as follows. An introductory section (section 2) recalls the basic theoretical background for landscape evolution models from which fluvial networks are extracted, jointly with a brief review of the foremost observational results that act as benchmarks for comparative studies of natural and generated networks. Section 3 discusses how optimal channel networks are seen as the byproduct of stationary solutions of the general landscape evolution equations. In particular, we shall discuss important differences in the statistical structures of global and local (i.e., dynamically accessible) minima of the exact functional derived from the landscape evolution equation at stationarity. Section 4 addresses the origin of loops in a connected structure and shows that every tree is a minimum of total energy dissipation where the local physics commands a concave mathematical form of the functional to be minimized. Section 5 explored the impact of the hydrologic results on a class of networks that are supposedly relevant to biological studies. Section 6 then explores the impact of hydrologic findings related to the shape of the functional to be minimized in the general field of network selection, with interesting inferences on the origin of general network

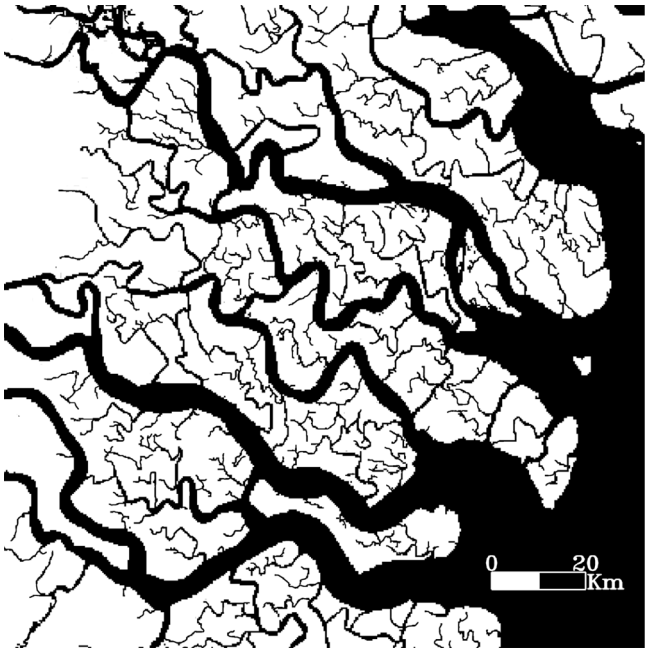

Figure 4. A large-scale, spaceborn image of the Brahmaputra-Ganges deltaic network. The original NASA image is taken from http://www.visibleearth.nasa.gov/, and the channelized pattern is extracted via suitable image processing techniques that recognize the spectral signatures of water (courtesy of E. Belluco). The complex interplay of the distributional characters typical of deltaic patterns and of the drainage patterns affected by strong tidal forcings (emphasized by the pronounced gradients of channel widths) produces loops appearing on all scales.

shapes. A set of conclusions closes then the paper with a look at future work in this area.

2. Theoretical and Observational Background

[8] Branching river networks are striking examples of natural fractal patterns which self-organize, despite great diversities in forcing geologic, lithologic, vegetational, climatic and hydrologic factors, into forms showing deep similarities of the parts and the whole across up to six orders of magnitude, and recurrent patterns everywhere [*Rodriguez-Iturbe and Rinaldo*, 1997]. Form and function coevolve. Interestingly, the drainage network in a river basin shows tree-like structures that provide efficient means of transportation for runoff and sediment and show clear evidence of fractal behavior. Numerous efforts to model the production zone of a river (where the system is open, i.e., water is more or less uniformly injected in space and later collected through the structure of the network implying that landscape-forming processes are well defined) have focused on reproducing the statistical characteristics of the drainage network. Much attention has also been paid to the temporal behavior and to the evolution of the topography of the basins, the so-called landscape evolution problem [see, e.g., *Willgoose et al.*, 1991; *Dietrich et al.*, 1992; *Howard*, 1994; *Banavar et al.*, 1997] and references therein).

[9] Our observational capabilities are also noteworthy. Accurate data describing the fluvial landscape across scales (covering up to 5 orders of magnitude) are extracted from digital terrain maps remotely collected and objectively manipulated. Raw data consist of discretized elevation fields $\{z_i\}$ on a lattice. The drainage network is determined assigning to each site i a drainage direction through steepest descent at i, i.e., along $\vec{\nabla} z_i$. Multiple flow directions in topographically convex sites, and their derived hydrologic quantities, are also easily tackled (D. G. Tarboton, TARDEM programs for the analysis of DTMs, available at http://www.engineering.usu.edu/dtarb). Many geomorphological features are then derived and analyzed. To each pixel i (the unit area on the lattice) one can associate a variable that gives the number of pixels draining through i i.e., following the flow directions. This quantity represents the total drainage area (total contributing, or accumulated, area) a_i at the point i, expressed, e.g., in pixel units, via $a_i = \sum_j w_{j,i} a_j + 1$ where $w_{j,i}$ is the element of an adjacency matrix, i.e., $w_{j,i} = 1$ if $j \rightarrow i$ and 0 otherwise and 1 represents the unit area of the "pixel" unit that discretizes the surface. In the case of uniform rainfall injection, a_i provides a measure of the flow at point i. In fact, the steady state flow J_i, can be defined as the sum of the injections over all the points upstream of site i, site i included, and thus obeying to the equation $J_i = \sum_j w_{j,i} J_j + r_i$, where r_i is the distributed injection. In the case of constant injection ($r_i \equiv$ constant = 1), the steady flow at i is proportional to the area a_i drained in that point allowing to use these two quantities interchangeably. Assuming $J_i \sim a_i$ is an accepted hydrologic assumption for landscape-forming discharges [*Leopold et al.*, 1964].

[10] Drainage directions determine uniquely network lengths. Downstream lengths L_i (i.e., from a site i to the outlet following the largest topographic gradient, i.e., steepest descent) can be computed easily to derive their distributions which clearly show the characters of finite-size scaling [*Maritan et al.*, 1996a]. The upstream length l_i at point i is defined as the distance, measured along the stream, from the farthest source draining into i. Overall, channelized patterns are now reliably extracted from $\{z_i\}$ fields through the exceedence of geomorphological thresholds, and have thus much improved our ability to describe objectively natural forms over several orders of magnitude [*Montgomery and Dietrich*, 1988, 1992]. For a review, see *Rodriguez-Iturbe and Rinaldo* [1997].

[11] Large-scale observations have allowed thorough comparisons across scales defining fractal river basins [*Mandelbrot*, 1977, 1983]. One outstanding example of fractal relation is Hack's law [*Hack*, 1957; *Mandelbrot*, 1983; *Rigon et al.*, 1996] relating the upstream length l_i at a given position i to the total cumulative area a_i at that position, seen quite early as a signature of fractal geometry. Contributing area a_i at any point is related to the gradient of the height (the topographic slope) of the landscape at that point: $|\vec{\nabla} z_i| \propto a_i^{\gamma-1}$ with a numerical value of γ around 0.5 [e.g., *Tarboton et al.*, 1989; *Montgomery and Dietrich*, 1992]. This slope-discharge relation proves a powerful synthesis of the local physics. The distributions of cumulative areas a_i and upstream lengths l_i are characterized by power law distributions (with the expected finite size corrections) with exponents in the narrow (and related [*Rinaldo et al.*, 1999a]) ranges 1.40–1.46 [*Rodriguez-Iturbe et al.*, 1992a] and 1.67–1.85 [*Maritan et al.*, 1996a], respectively. It is particularly revealing, in this context, that the finite-size scaling ansatz provides a most stringent observational proof of self-similarity and a strong form of Hack's law [*Rigon et al.*, 1996]. Scaling in the river basin has been documented in many other geomorphological indicators (including exact limit scalings [*Banavar et al.*, 1999]), making the case for the fractal geometry of nature particularly compelling [*Rodriguez-Iturbe and Rinaldo*, 1997]. Further proofs have been found by the striking invariance of probability distributions of length and area under coarse graining of the elevation field. The case of rivers is thus a solid starting point for other queries about the possible consilience of natural mechanisms.

[12] A major challenge lies in the explanation of the dynamic origin of fractal forms [e.g., *Bak et al.*, 1987]. Considerable efforts have been devoted to define static or dynamic models able to reproduce the statistical characteristics of fluvial patterns, and general concepts like self-organized criticality have been explored in this context [*Rinaldo et al.*, 1993]. It should be observed that real drainage basins are not static but usually evolve on extremely long timescales. Nevertheless, statistical properties seem to be preserved during most of the evolutionary process of a basin; most features characterizing the river basin morphology are irrespective of age. Some geomorphological signatures like valley densities (the relative extent of unchanneled concave areas), however, reflect climate changes without appreciable changes in the basic scaling features of aggregated area and length [*Rinaldo et al.*, 1995].

[13] It is worthwhile reviewing the theoretical background of landscape evolution models, because river networks may be defined by nodes on a regular lattice representing the elevation field, and links determined by steepest descent on the topography whose evolution deter-

mines the structure. Our aim is to find the simplest model that simulates the dynamical evolution of morphologically realistic landscapes and that preserves certain features during evolution. The equation that has been proposed to describe the evolution of the landscape, among a realm of somewhat less refined but essentially equivalent forms [e.g., *Howard*, 1994], is the following:

$$\dot{z}(t, \underline{x}) = -\alpha J(t, \underline{x}) \, |\vec{\nabla} z(t, \underline{x})|^2 + D\nabla^2 z(t, \underline{x}) + U, \qquad (1)$$

where z again denotes the elevation at the point $\underline{x} = (x, y)$ of the substrate plane and J is the modulus of the water flux (the landscape-forming discharge at a site) at that point and at time t. The first term is, of course, an erosional term proportional to the flux, the second is a diffusive term, known [e.g., *Dietrich et al.*, 1992] to portray hillslope evolution; the third term is a constant term modeling uplift. The existence of an uplift originating in tectonic forces is a well known fact in geomorphology [e.g., *Scheidegger*, 1991]: a landscape represents the instantaneous equilibrium of two concurrently active processes, uplift (endogenic) and degradation (exogenic). A stationary state results from the exact balance of these two agents. A simple argument leading to equation (1) follows from the general form $\dot{z} = F(\vec{\nabla} z, \nabla^2 z, \ldots, J)$, where an explicit dependence on z is excluded because it would break translational invariance. Note also that landscape-forming fluxes $\vec{J}$ (a two dimensional vector) become scalars i.e., $J = |\vec{J}|$ because $\vec{J}$ is assumed parallel to $\vec{\nabla} z$ [*Banavar et al.*, 1997, 2001].

[14] The diffusive term acts on the surface even at points with zero contributing areas unlike the first term which vanishes when the flux becomes zero. In the absence of the diffusive term the presence of maxima on the surface will cause the formation of singularities during the evolution, because e.g., points at the top of a hill will never be eroded by the first term (both J and $\vec{\nabla} z$ vanish). The presence of at least an infinitesimal diffusive term is then essential in eliminating these singularities. In the discretized version of the model each site (pixel) collects at least an unit area and thus no singularities due to a vanishing contributing area appear even in the absence of the diffusive term. Moreover, the discretization implicitly introduces a diffusive effect because it smoothes z on distances of the order of the lattice length and also prevents the occurrence of singularities due to a vanishing $\vec{\nabla} z$ when $D = 0$. Versions of equation (1) in which the activation of erosive actions depends on a value of $\nabla z\sqrt{J}$ exceeding a given threshold have also been studied [*Rinaldo et al.*, 1993; *Rigon et al.*, 1994]. Finally, a common version of the basic equation is the following:

$$\frac{\partial z(t, \underline{x})}{\partial t} = -\alpha J(t, \underline{x}) \, |\vec{\nabla} z(t, \underline{x})|^2 + U + \eta(t, \underline{x}). \qquad (2)$$

where U is a specifically defined as the geologic forcing, i.e., tectonic uplift, and $\eta(t, \underline{x})$ is a noise term with zero mean portraying the interplay of endogenic and exogenic factors [e.g., *Scheidegger*, 1991; *Dietrich et al.*, 1992; *Howard*, 1994; *Rodriguez-Iturbe and Rinaldo*, 1997]. Equation (2) has been solved numerically to study the coupled network growth and hillslope evolution [e.g., *Willgoose et al.*, 1991] (see also *Rodriguez-Iturbe and Rinaldo* [1997, chap. 1, and references therein] for a review of the subject).

[15] Suffice here to focus on the simplified version of equation (1) in the discretized lattice form obtained by putting $D = 0$ and removing noise terms. In fact, due to the coarse grained scale of the elevation field, the effect of the diffusive term would be negligible because it is not relevant to the large-scale behavior. This is consistent with established geomorphological tenets: fluvial processes (whose transport rates are defined by $\alpha J(t, \underline{x}) \, |\vec{\nabla} z(t, \underline{x})|^2$) are responsible for the imprinting of the network through randomly arriving (and practically instantaneous) landscape-forming events, whereas hillslope processes, whose net transport is defined by the divergence of a diffusive flux yielding the term $D\nabla^2 z(t, \underline{x})$ in equation (1), act on different timescales by smoothing out landscapes without the capability of altering their basic structure [*Rinaldo et al.*, 1995]. Thus we will consider as basic equation shaping the surface whose gradients produce collection networks, the following amended version on a lattice:

$$\dot{z}(t, \underline{x}) = -\alpha J(t, \underline{x}) \, | \, \vec{\nabla} z(t, \underline{x}) \, |^2 + U. \qquad (3)$$

which contains all ingredients needed for our current purposes. Moreover, when not explicitly stated otherwise, the flux ($J(t, \underline{x}) \sim J(\underline{x})$) will be taken to be proportional to the drained area as postulated above, and the terminologies "flux" and "drained area" will be used interchangeably. As noted above, this would in principle restrict the range of our conclusions because it corresponds to the assumption of an uniform rainfall acting on the surface during landscape-forming events. It has been noted, however, that relaxing this assumption, on assuming, say, heterogeneous patterns of rainfall in space and time, does not alter the generated structures unless for particular cases where the macroscale of the forcing heterogeneity is much smaller than the basin scale [*Rodriguez-Iturbe and Rinaldo*, 1997]. Note also that the case of heterogeneous rainfall corresponds to an uneven injection term in the definition of total cumulative fluxes J (and hence areas), an interesting case of multiplicative noise in equation (1) and its analogs. Thus we can safely conclude that in cases of practical interests our assumptions stand and the ensuing results worth generalizing.

[16] It should be noted that often geomorphologists [e.g., *Dietrich et al.*, 1992; *Howard*, 1994] employ a different version of equation (3). In fact, by noting that general empirical evidence suggests slope-area relationships of the type $|\nabla \bar{z}| \propto a^{-m/n}$, with $m/n \sim 0.5$, in many fluvial regimes, it was empirically concluded that the proper landscape evolution equation should be in the form $\dot{z} = -\alpha a^m |\nabla z|^n$, possibly complemented by a diffusive term portraying the inferences of hillslope transport. We see no contradiction with our main tenet, which assumes, in the small gradient approximation, that when fluxes substitute total contributing areas, the proper exponents should be $m = 1$, $n = 2$. More generally, however, landscape-forming discharges yield $J \propto a^{m'}$ with $m' < 1.0$, a well known empirical fact in hydrology often associated with the return period of bankfull discharges (1–2 years) [*Leopold et al.*, 1964]. Less frequent discharges have usually a similar scaling behavior yet with a lower exponent m'. It is also interesting to note that empirical slope-area relationships significantly different

from the 1/2 slope (at least in gently sloping landscapes) would indicate, within the validity of our scheme, a geomorphologic signature of climate which has been partly exploited [*Rinaldo et al.*, 1995] and perhaps worth future research.

[17] In spite of its simplicity, the model in equation (3) shows many interesting features. Note that the stationary solutions of equation (3) (i.e., $\dot{z} = 0$) are such that

$$| \vec{\nabla} z | \propto J^{-1/2}. \tag{4}$$

which implies $|\vec{\nabla} z| \propto a^{-1/2}$ if $J \propto a$. This is, indeed, the previously mentioned slope-discharge relation which is a well-known empirical fact. Notice also that a significant field of geomorphological research, both empirical and theoretical, deals with slope-area relations, their implications on hillslope/channel transition processes, seen especially in the light of deviations from (4) [e.g., *Tarboton et al.*, 1989; *Dietrich et al.*, 1992].

[18] Numerical simulations of equation (3) have shown that landscape evolution is characterized by two distinct timescales (as was noted by *Rinaldo et al.* [1993, 1995]). The soil elevations are lowered in a nonuniform way by erosion, causing variations in the drainage directions during the evolution. In a lattice model, at any given time, one may represent the drainage directions at all sites by means of a two dimensional map. After a first characteristic time, some sort of ''freezing'' time, the spanning graph determining the drainage directions in the basin does no longer change. Erosion keeps acting on the landscape and changes the soil height, but preserves the drainage structure. The second characteristic time, which is much longer, is the relaxation time at which the profile reaches its stable shape. Because many of the measured quantities, such as the distributions of drained areas and mainstream lengths, depend only on the two dimensional map, the existence of a freezing time much smaller than the relaxation time may provide an explanation for the fact that several statistical properties are found to be almost the same for many rivers, irrespective of their age.

3. Optimal Channel Networks

[19] In this section we review the basis for the claim that tree-like fluvial structures are a natural by-product of some optimization of form and function peculiar to the physics of rivers.

[20] To exploit some properties of stationary states of equation (1), we start from the review of the static model of river networks known as the optimal channel network (OCN) [*Rodriguez-Iturbe et al.*, 1992b, 1992c] and its implications [*Rinaldo et al.*, 1992, 1996, 1999a; *Rigon et al.*, 1996]. The OCN model was originally based on the ansatz that configurations occurring in nature are those that minimize a functional describing the dissipated energy and on the derivation of an explicit form for such a functional. A major step was the later proof [*Banavar et al.*, 1997, 2001] that optimal networks are exactly related to the stationary solutions of the basic landscape evolution equation, i.e., (3). In particular, any configuration that minimizes total energy dissipation, within the framework of general dynamical rules, corresponds, through the slope discharge relation, to an elevation field that is a stationary solution of the basic landscape evolution equation. Thus spanning, loopless (see also section 4) network configurations characterized by minimum energy dissipation are obtained by selecting the configuration, say s, that minimizes

$$E(s) = \sum_i a_i^{\gamma} \tag{5}$$

where i spans the lattice and a_i and γ are as defined above. Given that $a_i = \sum_j w_{j,i}\, a_j + 1$ where $w_{j,i}$ is the element of the adjacency matrix spanning the connectivity of every node j to i, the configuration s determines uniquely, on a spanning tree, the values of a_i. It is crucial, as we shall see later, that one has $\gamma < 1$ directly from the physics of the problem.

[21] The global minimum (i.e., the ground state) of the functional in equation (5)) is exactly characterized by known mean field exponents [*Maritan et al.*, 1996b], and one might expect to approach this mean field behavior on trying to reach stable local minima on annealing of the system. This is in fact the case. The proof of the above is not trivial: any stationary solution of equation (3) must locally satisfy the relationship $|\vec{\nabla} z| \propto J^{-\frac{1}{2}}$ between flux and gradient at any point. In a discrete version, the landscape is described by a field of elevations $\{z_i\}$ and gradients of Δz_i obtained by the biggest drop in elevation about site i. We thus maintain that the drainage basin can be reconstructed using the rule of steepest descent; that is, the flux in a point has the direction of the maximum gradient of the elevation field (the direction toward the lowest among all its nearest neighbors). Moreover, the channelized part of the landscape is necessarily (but not sufficiently) identified by concave areas where the above assumption holds strictly. One can thus uniquely associate any landscape with an oriented spanning graph on the lattice, i.e., an oriented loopless graph passing through each point. Identifying the flux in a point with the total area drained in that point, one can reconstruct the field of fluxes $\{J_i\}$ corresponding to a given oriented spanning graph. From the fluxes, a new field of elevation can be defined using equation (4).

[22] To elucidate further the physical nature of the functional in equation (5), we note that to each landscape $\{z_i\}$ defined on a lattice we associate a dissipation energy via

$$E = \sum_i k_i J_i \Delta z_i \tag{6}$$

where i spans all sites, Δz_i is the height drop along the drainage direction, J_i is the flow through the site i, and k_i is a quantity related to the soil properties measuring the heterogeneity of erosion-controlling features such as vegetation cover, exposed lithology, degree of saturation etc. For relatively homogeneous basins, one can assume $k_i = 1$ without losing sight of the main problem, i.e., the self-organized formation of aggregated patterns [*Rinaldo et al.*, 1992]. Moreover, observational evidence shows that the flow velocity tends to be constant throughout the network and thus the energy dissipated to maintain the water flow is proportional to the potential energy associated with the landscape [*Rinaldo et al.*, 1993]. The power in a link is $J_i \Delta z_i$ and (6) represents the power expenditure in the system as a whole. Equation (5) is recovered through $\Delta z(i) \sim J_i^{\gamma - 1}$, with $\gamma \simeq 0.5$, a corollary of equation (4) and

its parent equation (3), which, as seen above, are established empirical facts. Note that $\gamma = 0.5$ derives from the equation in the low-gradient approximation (where in steady state the leading term is $J(\underline{x})|\vec{\nabla} z(\underline{x})|^2 \sim$ constant) under the further assumption of exchanging J_i with a_i. Thus documented alterations in the slope-area relation (4) toward values of $\gamma \neq 0.5$ ought to be related, at least in gently sloping topographies, to uneven distributions of rainfall in landscape-forming events. This is precisely why we must limit our attention to patterns embedded in runoff-producing areas within the above framework to confine the problem within tractable limits.

[23] Equation (6), for a configuration s leading to aggregation yielding fluxes $J_i \propto a_i$, can thus be rewritten as

$$E(s) = \sum_i J_i(s)^\gamma \sim \sum_i a_i^\gamma, \tag{7}$$

where $J_i(s)$ is the set of fluxes defined by the connectivity of a spanning, loopless configuration s, and $\gamma \simeq 0.5$. Note that we wish to emphasize the dependence on the configuration s, an oriented spanning graph associated with the landscape topography z through its gradients ∇z. An interesting question is how networks resulting from the erosional dynamics are related to the optimal networks arising from the minimization of the dissipated energy. Specifically, we require that any landscape reconstructed from an optimal configuration using the slope-discharge relation is a stationary solution of the evolution equation. Superficially, this may seem to be a trivial fact because the relation between gradients and flows is verified by construction, but one should notice that the slope discharge relation alone does not implies stationarity, because the flow may not be (and in general is not) in the direction of the steepest descent in the reconstructed landscape. Thus optimal channel networks consist of the configurations s which are local minima of (7) in the sense specified below: two configurations s and s' are close if one can move from one to the other just by changing the direction of a single link (i.e., the set of links $s \cup s'$ represent a graph with a single loop). A configuration s is said a local minimum of the functional (7) if each of the close configurations s' corresponds to greater energy expended. Note that not all changes are allowed in the sense that the new graph again needs to be loopless. Thus a local minimum is a stable configuration under a single link flip dynamics, i.e., a dynamics in which only one link can be flipped at a given time, and is flipped only when the move does not creates loops and decreases the functional (7). Any elevation field thus obtained by enforcing the slope-area relation to a configuration minimizing at least locally the functional (7) is a stationary solution of equation (3), i.e., the landscape reconstructed from an optimal drainage network with the slope-discharge rule is consistent with the fact that the flow must follow steepest descent.

[24] The proof is derived considering a configuration realizing a local minimum of the dissipated energy, and a site i. The link issuing from i will join one of the nearest neighbors of i, say k. Let j be one of the remaining nearest neighbors such that changing the link from $i \rightarrow k$ to $i \rightarrow j$ one still gets an allowed configuration. Paths emerging from k and j will intersect downstream in a given point w, or will never intersect until they reach their outlets. Let S_{k_j} denote the set of all points in the path from k to w in the first case and from k to its outlet in the second. Likewise, one may define S_{j_k}. Changing the link from $i \rightarrow k$ to $i \rightarrow j$ will cause only the areas of sites belonging to the sets S_{k_j} and S_{j_k} to change. In particular all areas in the set S_{j_k} will be increased of an amount equal to the area a_i contributing to the flow through i, and all areas in the set S_{kj} will be decreased by that amount. Thus such a change will cause a change $(\Delta E)_{k \rightarrow j}$ in the dissipated energy equal to

$$(\Delta E)_{k \rightarrow j} = \sum_{x \in S_{j_k}} \left[(a_x + a_i)^\gamma - a_x^\gamma\right] + \sum_{x \in S_{k_j}} \left[(a_x - a_i)^\gamma - a_x^\gamma\right]. \tag{8}$$

where $\gamma = 1/2$ and a_x are the flows before the flip. The condition for a configuration to be a local minimum of E translates into the set of conditions

$$(\Delta E)_{k \rightarrow j} > 0 \tag{9}$$

for each i and j such that j is a nearest neighbor of i and gives rise to a loopless configuration. Conditions (9) imply that the elevation field determined by the local minimum configuration using the slope-area relation represents a stationary solution of equation (3). In order to be a stationary solution, the elevation field determined by a graph using slope-area relations must be such that the drainage directions derived with the steepest descent rule yield again the graph from which the elevation field originated. This would imply that if $i \rightarrow k$ is the drainage direction in the point i, the biggest drop in elevation from i to its nearest neighbors is in the direction of k. This condition reads

$$z(j) > z(k) \tag{10}$$

for any j that is a nearest neighbor of i and different from k. The proof that equation (9) implies (10) is elsewhere [*Banavar et al.*, 2001], also in the general case $0 < \gamma < 1$. The converse is not true, however, i.e., a stationary solution of equation (3) is not necessarily a local minimum of the dissipated energy under the single link flip dynamics.

[25] The OCN model has been thoroughly analyzed [*Maritan et al.*, 1996b; *Banavar et al.*, 2001]. In particular, the scaling behavior of the global minimum has been worked out analytically and it has been found to yield mean field exponents. Interestingly, local minima also exhibit critical behavior but are characterized by different nontrivial scaling exponents of key probability distributions describing, e.g., drained area, channeled length, elongation [*Takayasu et al.*, 1988, 1991a, 1991b; *Maritan et al.*, 1996b; *Rodriguez-Iturbe and Rinaldo*, 1997; *Rinaldo et al.*, 1999b].

[26] Figures 5a–5c show examples of local and global minima of OCNs (here chosen in a multiple-outlet configuration). Figure 2a shows the result obtained by "Eden" growth generated by a self-avoiding random walk, which is known to lead to suboptimal structures [*Rodriguez-Iturbe and Rinaldo*, 1997]. It is interesting to use Eden structures as benchmarks because their chance-dominated selection principle (no necessity is implied by the random walk

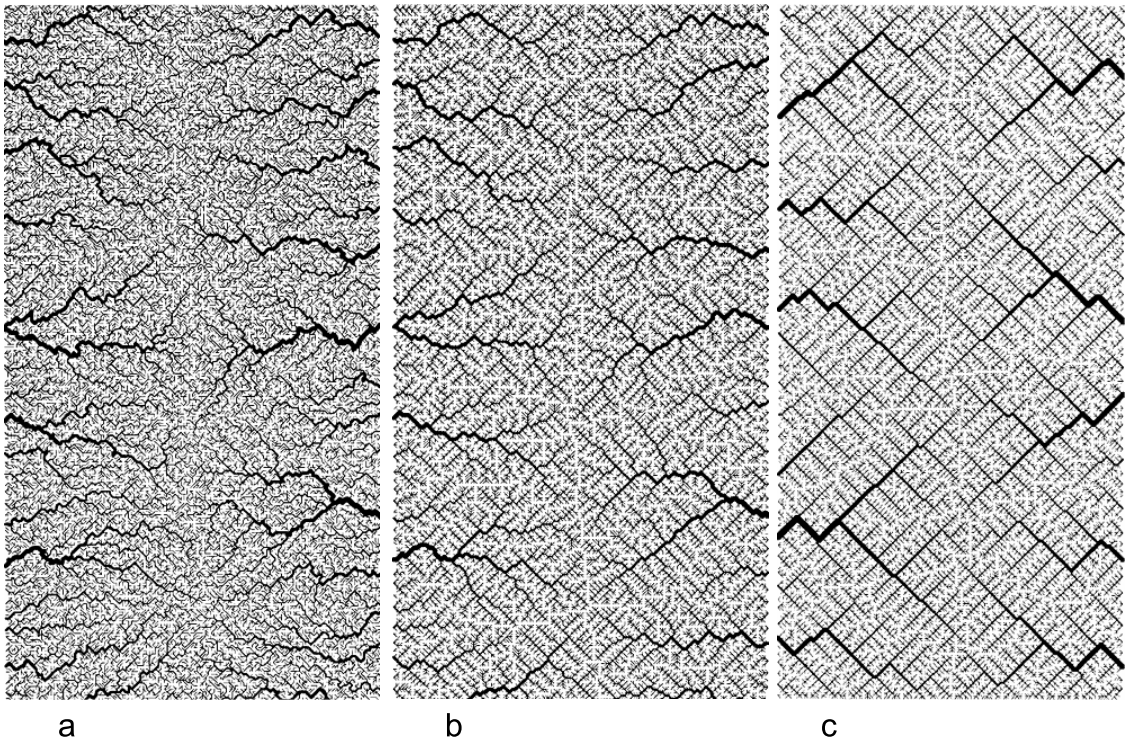

Figure 5. Multiple-outlet networks obtained in the same rectangular domain by (a) Eden growth patterns of self-avoiding random walks filling the domain, (b) an imperfect (i.e., $\mathcal{T} = 0$) optimal channel network (OCN) leading to a local minimum of total energy dissipation, and (c) ground state OCNs obtained through simulated annealing using a very slow schedule of decreasing temperatures $\mathcal{T}$. In Figure 5b note that OCNs bear long-lived signatures of the initial condition owing to the myopic search procedure but actually reproduce perfectly the aggregation and elongation structure seen in real river landscapes. The reaching of the ground state is confirmed by the matching of the exact mean field exponents with those calculated for Figure 5c [after *Rigon et al.*, 1998].

dynamics, and tree-like structures are selected because of the self-avoiding nature imposed on the process) because such structures were initially thought of as capturing the essentials of natural selection [*Leopold et al.*, 1964]. That turned out to be an artifact of nondistinctive tests of the network structure [*Rinaldo et al.*, 1999a], nicely termed the ''statistical inevitability'' of Horton's laws [*Kirchner*, 1993]. Indeed, if topological measures alone (e.g., Horton numbers, Tokunaga matrices) are used to sort out the fine properties of networks, one can be hugely misled into finding spurious similarities with natural forms, as one would sometimes safely conclude even at eyesight: Compare, for example, Figure 5a with Figure 5b. Topological features like, for example, those based on Strahler's ordering like Horton ratios or Tokunaga's matrices are indistinguishable in the two cases shown [*Rinaldo et al.*, 1999a]. Yet these are very different networks, as (linked) scaling exponents of areas, lengths and elongation clearly reveal. If eyesight and common sense would not suffice, exact proofs are available, like in the case of Peano's basin (Figure 1e) where topological measures match perfectly those of real basins and of OCNs, but fail to satisfy the strict requirements of aggregation and elongation. More subtle but equally clear is the failure of random walk type models [*Leopold et al.*, 1964] or topologically random networks [*Shreve*, 1966, 1967, 1969, 1974; *Mesa and Gupta*, 1987] to comply with exhaustive comparisons [*Rinaldo et al.*, 1999a]. Notice that the latter models were extremely influential in suggesting that chance alone was behind the recurrence of natural patterns, because of the equal likelihood of any network configurations implied by the topologically random model. Instead their purported similarity with natural patterns is now seen as an artifact of lenient comparative tools, and the statistical properties and ''laws'' derived in that context are almost inevitable for spanning trees.

[27] Necessity is instead at work in the selection of natural networks. Figure 5b [after *Rigon et al.*, 1996] shows a local minimum of E in equation (5), whose fine features match perfectly those found in nature [*Rinaldo et al.*, 1999a]. These results (Figure 5b) are obtained moving from an initial configuration s. A site is then chosen at random, and the configuration is perturbed by disconnecting a link, which is reoriented to produce a new configuration s'. If the new configuration lowers total energy dissipation i.e., $E(s') \leq E(s)$, the change is accepted and the procedure is restarted. Figure 5c is obtained through the same procedure used to obtain Figure 5b, where an annealing procedure has been implemented, i.e., unfavorable changes may also be accepted with probability $\propto \exp(-(E(s) - E(s'))/\mathcal{T})$ where $\mathcal{T}$ assumes the role of temperature in a gas or a spin glass. It is rather instructive to compare Figure 5b with Figure 5c, where a ground state is reached by very careful annealing using a schedule of slowly decreasing temperatures. This

state is characterized by mean field scaling exponents (here matched perfectly), and overall all too regular and straight to reproduce, even at eyesight, the irregular and yet repetitive vagaries of nature.

[28] Several random constructs have been thoroughly analyzed [*Rinaldo et al.*, 1999a], in a few cases through exact results [e.g., *Colaiori et al.*, 1997], comparing them with optimal ones obtained through minimization of total energy dissipation. Random network forms range from self-avoiding random walks like Eden growth patterns, topologically random or Leopold-Langbein constructions, to the so-called Scheidegger network which is a directed random aggregation pattern with injection. Deterministic fractals like Peano's networks have been also exactly analyzed (for a perspective of dynamic bridges of determinism and stochasticity in the geophysical context, see, e.g., *Sivakumar* [2004]).

[29] Thus many misleading similarities are inferred from the matching of topological measures like Horton's ratios. These turn out to be too lenient measures, as they occur almost inevitably for spanning loopless networks and thereby do not distinguish the structure of the aggregations patterns. On this basis alone it was shown that topological similarities are to be interpreted as necessary, rather than sufficient, conditions for comparison of network structures. A distinctive comparison of network structures stems from the matching of several scaling exponents which characterize the finite-size scaling forms of the distributions of length, aggregated area and elongation. At times, perfect matching of aggregation (underlined by Horton's ratios of bifurcation, length and area indistinguishable from those observed in natural structures) proves inconsistent with the structure of channeled lengths, like in the case of self-avoiding random walks.

[30] Moreover, suboptimal networks, that is, those derived by imperfect search of the type perceived as dynamically feasible, match all the features of the networks observed in the fluvial landscape, and thus pass the most thorough screening differently from all chance-dominated constructs. We thus conclude that the hydrologic context strongly suggest a case for optimal selection of network structures in nature.

4. Minimum Energy and Loopless Structures

[31] We shall now define precisely the selective advantages of trees in the fluvial physics [*Banavar et al.*, 2000]. Consider a square lattice. Fix an orientation for all lattice bonds. On each bond b a flux J_b is defined (notice that in the general context of networks it is illegal to identify the flux from node, say, i with J_i because it is generally not unique, differently from the case of trees). We shall assume that $J_b > 0$ if it is flowing along an assigned orientation. Uniform (unit) injection is equivalent to the set of constraints $(\partial J)_x = 1$ where ∂ is a discrete version of the divergence, and is a measure of the net outflow from a site:

$$(\partial J)_x = \sum_{b \in x} J_b \theta(b,x) \; = \; 1 \tag{11}$$

where the unit value is the model injection, constant for every node in the simplest case; b spans all bonds (links) concurring on node x, and $\theta(b, x) = 1(-1)$ if b is oriented outward (inward) node x. Any local minimum of the function

$$E \; = \; \sum_b \mid J_b \mid^{\gamma} \tag{12}$$

when $0 < \gamma < 1$, corresponds to $J_b \neq 0$ only on the bonds of a spanning tree. The main point [*Banavar et al.*, 2000] is in the proof that the networks that correspond to local minima of the dissipated energy are loopless and tree-like. The tree must be spanning due to the constraints (11): one cannot have $J_b = 0$ for all b's connected to a site so that there must be at least one outlet from each site x. Some site (or sites) must also be declared to be the global outlet. We shall show that loopless structures emerge as optimal solutions of equation (12) with the constraint (11), which is precisely the case for river networks.

[32] Figure 6 illustrates an extremely simple example with just four sites: Figure 6a shows the setup for the elementary four-bond network. The dot is the outlet. Here the current a is taken as the parameter regulating the entire distribution of fluxes owing to continuity. Figure 6b illustrates the only loopless configurations of the system generated by integer values of a. Figure 6c shows the plot of the function E versus a from the following equation (13) with $\gamma = 0.5$:

$$E = \mid a \mid^{\gamma} + \mid a + 1 \mid^{\gamma} + \mid 1 - a \mid^{\gamma} + \mid 2 - a \mid^{\gamma} . \tag{13}$$

which is derived from equation (12) after implementation of (11). In particular, Figure 6b shows the plot $E(a)$ where one notices that there are local minima in correspondence with one of the four currents being zero ($a = 2, 1, 0, -1$), corresponding to the four trees shown in Figure 3c. The explanation is simple. Suppose that $a \sim 0$ (the other cases are equivalent). All the terms in (13) but $|a|^{\gamma}$ can be expanded in Taylor series around $a = 0$. Thus, locally, one has

$$E = 2 + 2^{\gamma} + \mid a \mid^{\gamma} + \mathcal{O}(a) \tag{14}$$

which has a cusp-like behavior because $0 < \gamma < 1$. Notice that $\partial E/\partial a|_{a=0^{\pm}} = \pm\infty$ and thus one cannot find the minima simply by imposing the condition $\partial E/\partial a = 0$. If $a \neq 0, \pm 1, 2$, $\partial^2 E/\partial a^2 < 0$ and there are no other minima of E (only maxima) [*Banavar et al.*, 2000]. The proof for the general case is elsewhere [*Banavar et al.*, 2001].

[33] Figure 6c shows the function E versus a plotted for various values of γ (specifically, for $\gamma = 0.25, 0.5, 1$ and 2). Note that for $\gamma = 1$ all directed (with the currents going in the positive directions) configurations, loopless or not, have the same energy. The case $\gamma = 2$ corresponds to the resistor network case for which there is just one minimum at $a = 1/2$ [*Doyle and Snell*, 1989]. Note that since there is one unknown current for each bond and one continuity equation for each site the number of independent variables is given by the number of bonds minus the number of sites (excluding the outlet), which for the simple topologies considered is equal to the number of elementary loops (this is a particular case of the Euler theorem).

[34] Since we have seen that local minima occur in singular configurations where some currents are zero, we

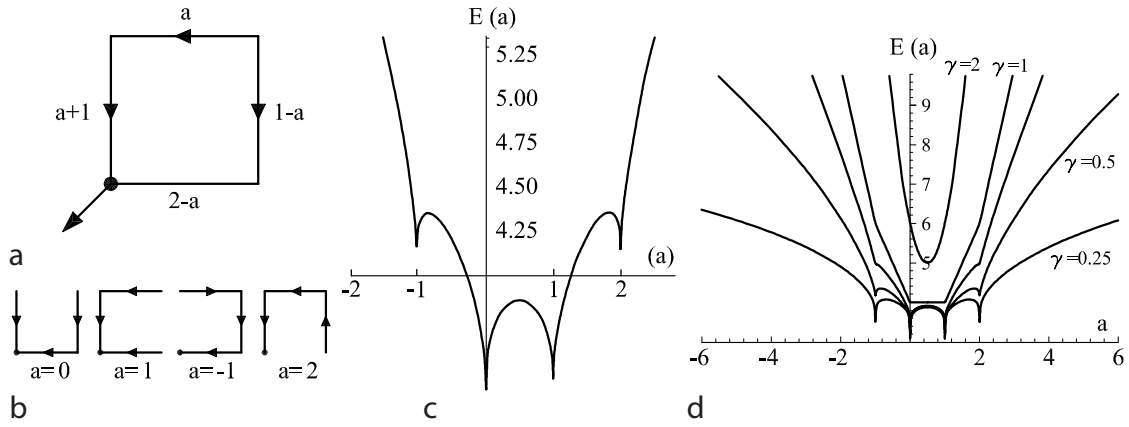

Figure 6. Four-bond lattice. (a) Four-node arrangement, with indications on the currents that respect continuity (note that a unit flux is injected at each node); (b) the only four possible trees, corresponding to cases $a = 0, -1, 1, 2$. (c) Energy function $E(a)$ versus a in equation (13). (d) Energy functions $E(a)$ versus a for the cases $\gamma = 0.5, 0.75, 1, 2$ [after *Banavar et al.*, 2002].

cannot introduce the standard technique of Lagrange multipliers to find the minima of E with the constraint (11). In order to be able to do that E must be regularized as $E = \sum_b (J_b^2 + \varepsilon^2)^{\gamma/2}$ in the limit $\varepsilon \to 0$. The general proof is beyond the scope of this paper and given elsewhere [*Banavar et al.*, 2000] for an arbitrary graph, where the number l of independent loops is given by the number of bonds minus the number of sites plus the number of connected components. Note that for the particular case where the graph must be a spanning structure the number of connected components is unit (for example, in the case of an $n \times m$ rectangular lattice one has $l = nm - n - m + 1$).

[35] Obviously, for the set of dynamical rules postulated above, the energy landscape is riddled with a large number of local minima characterized by a range of similar values of E. In single realizations, boundary and initial conditions affect the feasible (i.e., dynamically accessible) optimal state to different degrees depending on their constraining power. This fact matches the observation [*Maritan et al.*, 1996a] that scaling exponents are coherently linked in a range of values, narrow enough but significantly different from the ground state [*Maritan et al.*, 1996b]. The truly important implications are twofold: on one side, in fact, all local optima are trees; on the other, imperfect optimal search procedures are capable of obtaining suboptimal networks which nevertheless prove statistically indistinguishable from the forms observed in nature and quite different from the absolute minima [*Rodriguez-Iturbe and Rinaldo*, 1997]. Indeed, we believe that the worse energetic performance and yet the better representation of the patterns of nature are thought of as mimicking the myopic tinkering of evolutionary processes; indeed, you may not have a blueprint, but you jump on a selective advantage once you feel it however myopic the "watchmaker" may be (as opposed to the blind watchmaker, whose aimless tinkering inspired a beautiful metaphor [*Dawkins*, 1988]).

5. On a Few (Biological) Implications

[36] Here we show that some of the properties of spanning networks that we have studied may, indeed, have biological implications. In particular, we address typical allometric scaling relationships, that is, power laws (with noninteger scaling exponent) relating measures of metabolic rates of living organisms with their body mass. This is done by analyzing the general features of spanning networks serving an assigned volume [*West et al.*, 1997, 1999a, 1999b; *Banavar et al.*, 1999, 2002, 2003; *Brown et al.*, 2000; *Maritan et al.*, 2002].

[37] We consider a single network source that services L^D sinks uniformly distributed in a D-dimensional space of volume V and mass M. Each site is connected to one or more of its neighbors resulting in a transportation network that spans the system. Such a network may be a well-connected one with loops (Figure 7a) or merely a spanning tree (Figure 7b), extreme examples of which are star-like (or explosion) or spiral structures (Figure 8) [*Banavar et al.*, 1999]. We begin with the situation in which each sink, x, is supplied by the source at a steady rate, r_x, no less than a positive value $r_{\min}$ and no larger than a value $r_{\max}$ ($r_{\max} \geq r_{\min}$) both of which do not depend on L. A simple special case would correspond to a uniform constant rate for all sinks. Such a system could represent a biological organism which needs a steady supply of nutrients to all its parts [*McMahon and Bonner*, 1983; *Peters*, 1983; *Calder*, 1984; *Damuth*, 1998]; the sinks represent, e.g., the cells served by capillaries. The metabolic rate of such a biological organism is given by $B = \sum_{i \in V} r_i$ and simply scales as the volume V or mass M, i.e., $\propto L^D$. Note that another example is the inverse problem of the drainage network river where the sinks represent source areas, r_i is the net injection rate and the network provides routes for transport. Any transportation network must provide a route from the source to all the L^D sinks and consists of interconnected links in each of which, in steady state, the flow rate does not change with time. Each link starts or terminates at the source or a sink. Let, as usual, the scalar quantity $|J_b|$ represent the magnitude of the flow on the bth link. The source has an outward flow, whose rate exactly equals the sum of all the flow rates into the sinks. At a junction of the links, a conservation law (exactly of the type employed in equation (11)) for the net flow holds; the inflow must exactly balance the outflow plus

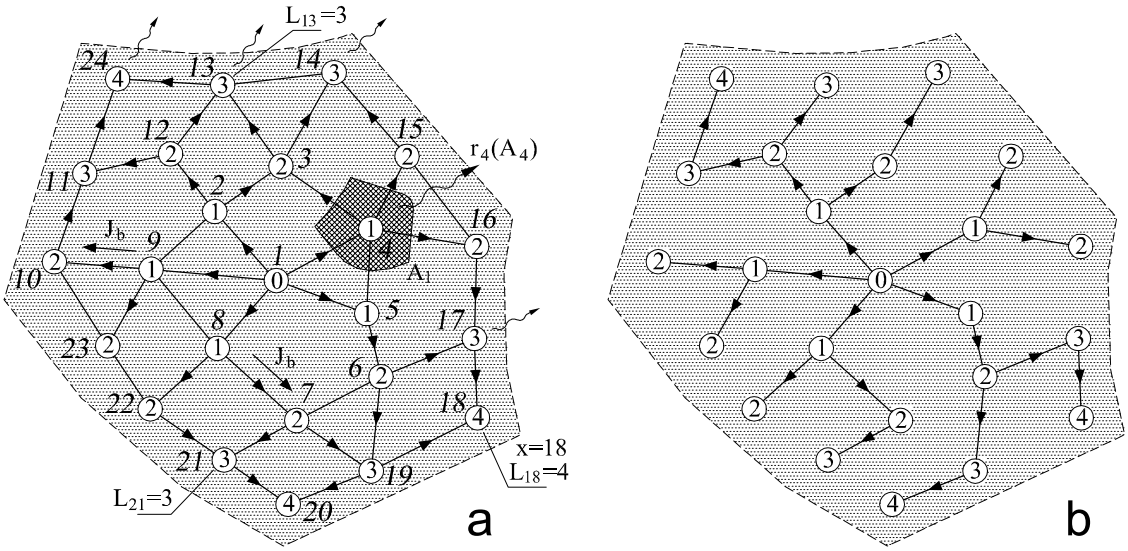

Figure 7. Network types for the study of biological analogies: (a) well connected with loops and (b) a subset of Figure 7a which is tree-like and spanning all $L^D = L^2$ sites. The cross-hatched area about node 4 indicates some domain A_4 serviced by node 4 (an area in $D = 2$ or a volume for $D = 3$) which has a biological meaning described elsewhere. In hydrologic settings it is simply the elementary pixel area that covers the topographic surface and, in general, works as a supply or as a drainage domain. The term $r_4 = r_4(A_4)$ represents the injection/delivery proper of node 4. The number circled at every node represents the topological distance (in link units) to the seed (the sink if the network collects; the source if the network distributes). L_{18} refers to the topological distance of node 18 from the source/sink. Here J_b indicates the flux along the link [after *Banavar et al.*, 1999].

the amount supplied to the sink. This conservation law does not uniquely determine the flow on each link for an arbitrary network, as seen above. The degrees of freedom in the choice of the flow pattern are controlled by the number of independent loops, equal to (the number of links) − (the number of sites) +1. The total quantity of nutrients in the network at any instant of time, C, is simply given by (i.e., is proportional to) $\sum_b |J_b|$.

[38] We begin our discussion by considering a maximal network (Figure 7a) in which every site (sources plus sinks) is connected to each of its neighbors. The distance, L_x, between the source (denoted by O) and any given sink x is defined to be the minimum number of sinks encountered among all the routes along the maximal network from O to x (the standard Euclidean distance from O to x, say $|x|$, scales as L_x). As a consequence, two neighboring sites x and y satisfy the constraint that $|L_x - L_y| = 0$ or 1. We now define a convention for the orientation of each of the links. The link between sites x and y is defined to have an orientation that is directed away from the source. For cases in which $L_x = L_y$, the orientation is chosen to be one of the two possibilities. We define a flow J_b as having a magnitude equal to $|J_b|$ and a sign that is positive if the flow direction coincides with the link orientation and negative otherwise. We note that

$$\sum_b J_b(L_y - L_x) = \sum_x L_x r_x \tag{15}$$

where b is a link oriented from x to y, the summation on the left-hand side is over all links of the maximal network, r_x is the flow into x minus the flow out of x (and is a positive quantity) and the sum on the right-hand side is over all sinks. The left hand side is also equal to

$$\sum_b J_b(L_y - L_x) = \sum_b^{directed} J_b, \tag{16}$$

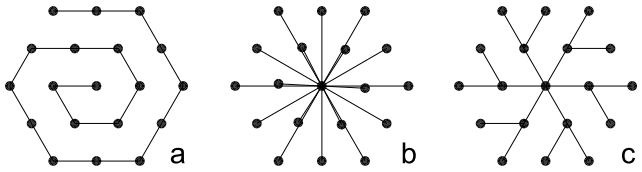

Figure 8. Spanning tree types, connecting N nodes on a triangular (unit) lattice. (a) Spiral pattern. Its total length L_T is defined as the total (along-link) "road" length (computed keeping as unit the distance between nodes regardless of orientation), which turns out to be quite small whereas the average length $\overline{L}$ from any node to the seed is very large ($\overline{L} = (1 + 2 + \cdots + N)/(N - 1)$); it greatly penalizes the individual while considerably helping the whole if one thinks, say, of traveltimes to a target in the sink when particles are seeded in any node at $t = 0$ and are moved along the network with constant velocity. In this case the average length is the mean traveltime;.(b) Star-like (i.e., explosion) pattern producing a very large L_T because of the large extent of unshared road and a very small average distance to the sink, $\overline{L}$. Explosion patterns by design favor the individual benefit against the collective one as mean traveltime to the target is at a minimum. (c) Typical aggregation tree, spanning and loopless, short as a whole (small L_T) as well as direct for the individual (small $\overline{L}$) [after *Stevens*, 1974].

where the restricted sum on the right hand side is only over links for which L_x is not equal to L_y. These links comprise what we define to be a maximal directed network. All viable directed networks are made up of links which are a subset of the links in the maximal directed network and yet span the entire system (or provide routes to each of the sinks from the source). The following inequalities follow [*Banavar et al.*, 1999]:

$$C = \sum_b |J_b| \geq \sum_b^{directed} |J_b| \geq |\sum_b^{directed} J_b| = \sum_x L_x r_x, \qquad (17)$$

where in the last step, we use (15) and (16) and the fact that the RHS of equation (15) is positive. The last term in (17) is a quantity that is independent of the choice of the flow configurations. Notice that in equation (17) the equality holds for the case $J_b \geq 0$ on the links belonging to the maximal directed network, and $J_b = 0$ otherwise. Furthermore,

$$\sum_x L_x r_x \geq r_{\min} \sum_x L_x = r_{\min} L^D \langle L_x \rangle, \qquad (18)$$

where $\langle L_x \rangle = \sum_x L_x / \sum_x 1$. Because the maximal network is space-filling and $\langle L_x \rangle \propto \langle |x| \rangle$ scales as L, the key result is obtained.

[39] Among spanning trees, B scales at most as $C^{D/(D+1)}$ and at least as $C^{1/2}$. In fact, the upper limit, $C^{D/(D+1)}$ is realized for directed spanning trees as shown above. In order to obtain the lower bound, let C_N be the amount of nutrients for an arbitrary spanning tree serving N sinks. Eliminating a sink at the maximum distance from the source would change each J_b by, at most, an amount $r_{\max}$. Thus

$$C_N \leq C_{N-1} + (N-1) r_{\max} \leq \cdots \leq \frac{N(N-1)}{2} r_{\max}. \qquad (19)$$

Because $N \propto L^D \sim B$, B scales at least as $C^{1/2}$. This case is described by a spiral pattern (Figure 8b).

[40] For an efficient network (see, e.g., Figure 7b), for which C is as small as possible for a given service volume V, both B and V scale as $C^{D/(D+1)}$. Thus, for $D = 3$ one obtains "quarter power" scaling via

$$B \propto M^{\frac{D}{D+1}} \qquad (20)$$

This efficient network has all links directed away from the source (or toward a collection point or the outlet for the river basin). Indeed, M scales as L^{D+1} for all directed networks (the flow in all but the links of the directed maximal network are necessarily zero in these cases) independent of whether they have loops or a tree-like structure, as long as J_b is nonnegative on each link. Such solutions do, indeed, exist and include all directed trees. These solutions belong to the class of the most efficient networks in that they lead to the smallest value of $C \sim M$. The application of this result to the problem of allometric scaling in living organisms, which span masses that range over 21 orders of magnitude [*McMahon and Bonner*, 1983; *Peters*, 1983; *Calder*, 1984], is straightforward [*Banavar et al.*, 1999].

[41] Interestingly, an application of the limit scaling properties as a test of optimality has been been proposed for food webs [*Garlaschelli et al.*, 2003] whose network patterns have interesting properties [e.g., *Montoya and Solé*, 2002].

[42] In spite of an impressive array of scales and the accompanying diverse requirements in the resources needed for sustaining the organism, a robust feature is that a variety of biological quantities (generically denoted in the equation below as B, be it the heartbeat frequency, the lifespan or any measure of metabolic rates) that are related to blood circulation, scale algebraically with the mass of the organism, M, as $B \propto M^{1/\alpha}$ where α is a scaling exponent and the constant of proportionality depends on the given organism [*McMahon and Bonner*, 1983; *Peters*, 1983; *Calder*, 1984; *Schmidt-Nielsen*, 1984; *Damuth*, 1998; *Brown*, 1995; *West et al.*, 1997, 1999a, 1999b]. The exponent α is usually and consistently found to be obtained from the fraction 1/4, i.e., quarter powers.

[43] In the above analysis, the mass of an organism, M, ought to scale at least as the flow volume, C, so that in the simplest and most efficient scenario, $B \sim M^{3/4}$, which is the central result of allometric scaling. The analysis shows, however, that the basic result does not require any assumptions regarding the hierarchical nature of the network nor demands a tree-like structure (like in Figure 7b) even though the presence of a tree would greatly shorten the total length of the network, thereby increasing its viability and efficiency. Note that all directed networks, of which fractal trees are a subset, obey the general rule termed quarter power law for bodies in space ($D = 3$) spanned by a distribution network. This generalizes and reinforces the result of *West et al.* [1997] who find that the allometric property responds to particular optimal selections. Thus on assuming that network flows yield a proxy of mass and of total metabolic rates [*West et al.*, 1997, 1999a, 1999b; *Enquist et al.*, 1998], we predict that quarter power allometric scaling in living organisms does not need specific assumptions on the fractal-like nature of the network because the key size-form relations are applicable to most transportation networks; they only need to be directed.

[44] We shall later investigate the degree of inefficiency in the network structure that quarter power laws can tolerate. It is worthwhile to mention here, however, ongoing discussions about the so-called big picture [*West et al.*, 1999a] on the origins and the departures from allometric scaling in living organisms. The above exact results do not pretend to explain the complexity of living organisms. They simply suggest that complex hierarchical models of tree-like structures are not needed to produce allometric scaling once one assumes that flow volumes supplying metabolites to a given number of "sites" provide a proxy for mass and metabolic rates. Thus hierarchical constructs at best explain the 3/4 law for very particular cases. Incidentally, the commonly observed fluctuations in the range 2/3–3/4 in allometric scaling exponents for different macroecological data sets (that prompted doubts on the very existence of universal allometric scaling [see, e.g., *Dodds et al.*, 2001]) have been suggested to stem from the relative balance (or lack of it thereof) of supply and demand of metabolites delivered as [*Banavar et al.*, 2002]

$$B \propto \left(\frac{\text{supply}}{\text{demand}} M \right)^{\frac{D}{D+1}} \qquad (21)$$

which obviously decays into equation (20) if supply and demand are balanced. Directed transportation of metabolites, however, is seen as the key and common reason for the occurrence of quarter power scaling.

6. Hydrologic Allometry

[45] We now turn to a test of the theoretical tools within the hydrologic context of river networks. How strong should the departure from directedness be to be able to alter the central allometric tendency of the network? In rivers, the role of the metabolic rate is taken by total contributing area, a_x, at any site x within the basin. As seen in the introduction, at the xth location, area is defined by the recursion relation $a_x \sim \sum_{j \in nn(x)} a_j + 1$, where $nn(x)$ are the nearest neighbors of x that drain into x through appropriate steepest descent drainage directions (note that this is an alternative way of defining total contributing area, avoiding the introduction of adjacency matrices). Note also that the added unit is the area of the elementary pixel, i.e., in the simplest picture area simply surrogates the total number of nodes connected to x. Thus area a_x at any site x (or the total number of connected sites) plays the role of the basin metabolic rate, B, whereas the analog of mass, M, is defined by the quantity

$$M \propto \sum_{y(x)} a_y \tag{22}$$

where $y(x)$ indexes the collection of all sites y connected to x [*Maritan et al.*, 2002]. Note that $\sum_{y(x)} a_y$ does not add to a_x, nor does it include it. Allometric plots relate M and B via a power law, $M \propto B^{\gamma}$, and their hydrologic counterpart yields

$$\sum_{y(x)} a_y \propto a_x^{\gamma} \tag{23}$$

We recall that for directed networks in two dimensions ($D = 2$), the key prediction [*Banavar et al.*, 1999] is that the lowest attainable value is $\gamma = (D + 1)/D = 3/2$. Our result technically predicts that, for efficient drainage basins, a log-log plot of $M \propto \sum_{y(x)} a_y$ versus $B \propto a_x$ ought to have a slope of about 3/2 because $D = 2$.

[46] Note that the fractal nature of river networks would stem from the fact that embedded within any basin are other subbasins with similar features reflected in linked scaling exponents [*Rinaldo et al.*, 1999a]. The mass M in equation (22) associated with any site x relates to its area a_x via [*Maritan et al.*, 2002]

$$\sum_{y(x)} a_y \propto a_x \langle L_x \rangle \tag{24}$$

where $\langle L_x \rangle$ is the mean distance of the sites within the subbasin to their outlet x measured along the network; thus it is only from this point that we restrict our attention to the case of spanning trees. In this case, if y is upstream of x there is only one path joining y to x along the network. Such being the case, we argue that

$$\langle L_x \rangle \propto a_x^h \tag{25}$$

where h is the so-called Hack's exponent relating the upstream length to the total contributing area. Hack's law, whose validity and meaning have been much debated in the scientific literature [e.g., *Mandelbrot*, 1983], is commonly defined by relating mainstream length, say L, to drainage area A at the closure rather than everywhere within the basin. The validity of equation (25) also within nested subbasins has been suggested to constitute a strong version of Hack's law and a proof of the embedded self-similarity of the network structure resulting in a fractal structure of river basins [*Rigon et al.*, 1996]. Thus the mainstream, sometimes rather arbitrarily defined but most commonly taken as the longest flow path length (and thereby a single flow path), is proportional to the mean length upstream of x, i.e., $L \propto \langle L_x \rangle$. From equations (24) and (25) one obtains [*Maritan et al.*, 2002]

$$\gamma = 1 + h \tag{26}$$

which exceeds the limit scaling $\gamma = 3/2$ whenever $h > 1/2$. Notice that Hack's exponent h would be equal to 1/2 only if geometric similarity is to be preserved as a basin increases in area while preserving its shape, and this corresponds to the limit scaling for most efficient networks. Typical observational values range about $h \sim 0.57$ [*Hack*, 1957] and this was seen as a clear indication of the fractal nature of rivers [*Mandelbrot*, 1983]. Note that OCNs have, indeed, $h = 0.57 \pm 0.02$ and thus feasible optimality has been suggested to imply Hack's law [*Rigon et al.*, 1998].

[47] Figure 9 shows typical allometric plots for four river basins of various sizes, geology, vegetational state and digital terrain map properties. From top to bottom: Guyandotte (WV), area 2088 km^2, mainstream length $L = 145.1$ km, $\gamma = 1.56 \pm 0.02$; Tirso (Italy), 2090 km^2, $L = 103.0$ km, $\gamma = 1.53 \pm 0.02$; Johns Creek (KY), 484 km^2, $L = 68.8$ km, $\gamma = 1.59 \pm 0.02$; Moshannon Creek (PA), 393 km^2, $L = 49.7$ km, $\gamma = 1.52 \pm 0.01$ [after *Maritan et al.*, 2002]. The observed values of γ range from 1.50 to 1.59, and the scatter of the individual curves (which we term intranetwork scaling to suggest that the noise within same "species" is studied), relative to the nested subbasins of the same basin, is remarkably small, and yet the differences on the scaling exponents are noticeable. From an extended survey of field data, it clearly emerges that individual networks conform to scaling laws that can significantly differ from the lower bound $\gamma = 3/2$ [*Maritan et al.*, 2002].

[48] To investigate the extent of the deviations of γ from its lower limit we have tested a broad class of statistical and deterministic network models, some amenable to exact solution. These include, for comparison, a real network, the Dry Tug Fork river (CA) (Figure 1a) whose allometric exponent is $\gamma = 1.57$ (hence $h = 0.57$), matching standard observations [*Hack*, 1957]. We have studied stochastic constructs such as the Scheidegger network (Figure 1c), whose scaling exponents are known exactly [*Takayasu et al.*, 1988; *Huber*, 1991] to be $\gamma = 5/3$ whereas the computed value is $\gamma = 1.67 \pm 0.01$. *Peano*'s [1890] network (Figure 1e) is a deterministic fractal whose main topological and scaling features have been solved analytically [*Marani et al.*, 1991; *Colaiori et al.*, 1997]. The exact value is $\gamma = 3/2$, nicely reproduced by computations at iterative stages of construction larger than 10. Optimal channel networks (Figure 1b;

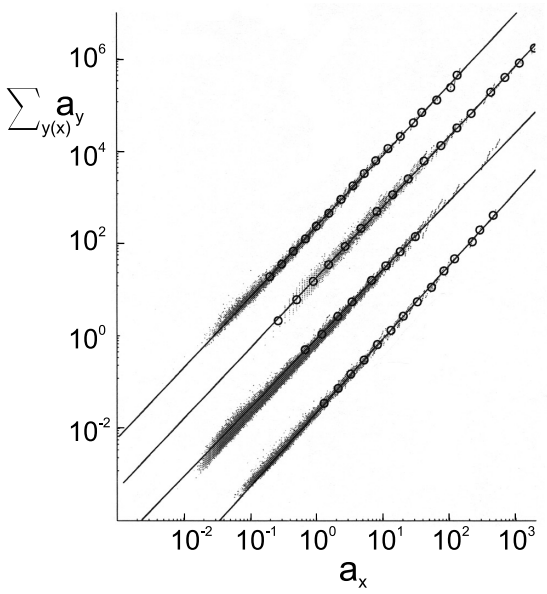

Figure 9. Network allometry: Log-log plots of $\sum_{y(x)} a_y$ versus a_x for four different basins. These are analog to allometric plots relating M and B via a power law, $M \propto B^{\gamma}$. Here $\gamma = 1 + h$ where h is Hack's exponent [after *Maritan et al.*, 2002]. Plots have been arbitrarily shifted vertically to distinguish the different basins.

see also Figure 5b), described in section 3, have fractal characteristics that are obtained through a specific network selection process (see, e.g., equation (5). Indeed, one obtains a rich structure of scaling optimal forms that are known (section 3) to closely conform to the scaling of real networks, even in the case of unrealistic geometric boundaries (here $\gamma = 1.57 \pm 0.01$ [*Rodriguez-Iturbe et al.*, 1992b; *Maritan et al.*, 1996a, 1996b; *Rodriguez-Iturbe and Rinaldo*, 1997]). Hot OCNs (Figure 1e) are rather undirected trees. They yield $\gamma = 1.68 \pm 0.1$ [*Rodriguez-Iturbe and Rinaldo*, 1997]. Other examples have been studied and are dealt with elsewhere [*Maritan et al.*, 2002].

[49] In all cases, excellent agreement is found between the directly determined value of the allometric scaling exponent and its relationship to Hack's exponent. Moreover, in all cases $\gamma \geq 3/2$ as predicted [*Banavar et al.*, 1999]. We thus confidently conclude that the allometric property is, indeed, distinctive of the network structure, differently from topological measures (note, for instance, that Peano's Horton ratios and Tokunaga matrices are computed exactly and are indistinguishable from those of OCNs or real rivers) whose matching ought to be interpreted as a necessary, rather than sufficient, condition when comparing network structures.

[50] An analog of interspecies allometric scaling corresponds, in this context, to an ensemble average of data from different populations of networks. In river basins, we have seen that γ varies in the range 1.50–1.60 with relatively small but detectable scatter in the individual curves (i.e., for a single fluvial network). The ensemble average built by mixing different subbasins nested in the same basin with other basins (specifically, those shown in Figures 1a–1e) and their subbasins, is shown in Figure 10. Though the scatter is higher, mimicking that of most macroecological data sets, the mean value of γ is statistically indistinguishable from 3/2. Thus quarter power scalings (of which the scaling exponent 3/2 is the analog in two dimensions) simply emerge as the central tendency. This surprising result [*Maritan et al.*, 2002] matches an important, probably overlooked result. In fact, it has been shown that the equivalent of an ensemble average of Hack's exponents (deduced from the mainstream length-total area plots of a very large number of different basins and their nested subbasins regardless of whether channeled or hillslope lengths were accounted for, and covering over 11 orders of magnitude) yields values indistinguishably close to $h = 1/2$ [*Montgomery and Dietrich*, 1992], a fact that puzzled investigators for some time [*Rodriguez-Iturbe and Rinaldo*, 1997]. This effectively corresponds to an ensemble average of different network shapes, yielding an "interspecies" allometric scaling exponent $\gamma = 1 + h \sim 3/2$. Notice that infinite topologically random networks also have asymptotically $h = 1/2$ and hence $\gamma = 3/2$ [*Mesa and Gupta*, 1987]. Effects of ensemble averaging of networks both in the bulk or at the boundaries of multiple-outlet optimal networks where competition for drainage occurs because of the constraint of the fixed total area being drained have also been studied [*Maritan et al.*, 2002] and the above results are unambiguously confirmed.

[51] Thus individual network forms exhibit allometric exponents that are sensitive probes of the network structure, which are directly related to the underlying fractal structure of the network (i.e., intraspecies scaling). Ensemble averages, the analog of interspecies scalings, smooth out details, enhance the scatter and lead to an γ exponent that approaches the limiting value obtained for directed networks. Our results demonstrate the robustness of the central tendency of allometric scaling in network structures. However, the sensitivity in probing the geometrical variability of network shapes is much refined when studying homogeneous geometries reflected in consistent deviations of the allometric scaling exponent from the limit values $\gamma = 3/2$ ($D = 2$) for planar networks or $\gamma = 4/3$ ($D = 3$) in plants and living organisms. We thus suggest that unavoidable fluctuations in the geometrical arrangements of the parts and the whole of a living network do not alter the basic tendency provided by biological needs. Thus the ubiquity

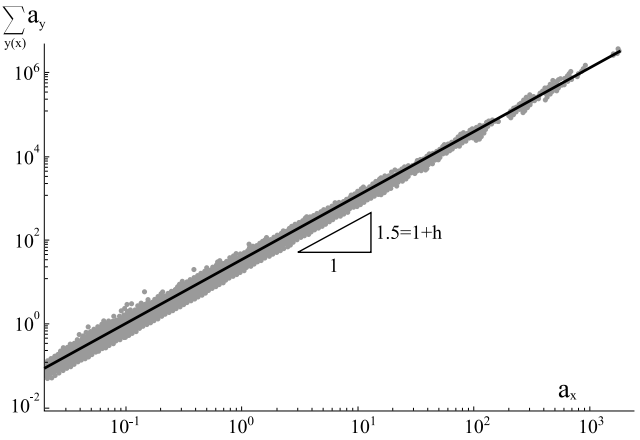

Figure 10. Ensemble average of all networks in Figures 1a–1e. Note the limited scatter around the slope $\gamma = 1 + h = 3/2$, which shows that the ensemble average of h, is indeed, $h = 0.5$ [after *Maritan et al.*, 2002].

of the so-called quarter power law may be a consequence of the robustness of network properties with respect to geometrical variations in systems where supply rates are independent of body mass. Thus the purported recurrence of the so-called 3/4 law may consist of chance, owing to the robustness to noisy geometry and topology, and necessity dictated by supply-demand balance [*Banavar et al.*, 2002].

7. Network Structures From Other Selection Principles

[52] The rules investigated in the previous sections suggest that the convexity of the function defining the selective advantage of different hydrologic network structures matters. In the case of fluvial basins, the basic concavity of the energy E is provided directly by the physics of the landscape evolution problem. We wonder whether similar principles may apply to the selection of different network structures by natural processes of different nature. Indeed, many complex systems, from the Internet to metabolic or ecological ones, can be mathematically described by networks of interacting elements, and loops are ubiquitous [e.g., *Bollobas*, 2001; *Barabási and Albert*, 1999; *Albert and Barabási*, 2002; *Barabási et al.*, 2000, 2002; *Montoya and Solé*, 2002; *Dorogovtsev and Mendes*, 2000, 2002; *Newman*, 2003; *Song et al.*, 2005]. We thus wonder whether meaningful topologies of complex networks may emerge as the result of selection principles.

[53] Here we suggest, following *Colizza et al.* [2004], that a class of optimal models evolved by local rules and chosen according to global properties of the aggregate yields unexpected behavior in the transition from different types of optimal topologies. Random or scale-free arrangements and small-world phenomena [*Watts and Strogatz*, 1998; *Watts*, 1999] are then seen as particular cases emerging from selective pressures toward connectivity and/or directedness.

[54] To select arbitrary structures, the evolving network maintains a fixed number of nodes n, as well as the number of links l within a context where the degree of outgoing links from every node is arbitrary. The "energy" function E of our selective algorithm introduces a new definition of distance on a graph that accounts not only for the length of the shortest path between two nodes, but also for the degrees k_p of the nodes p encountered along such path, i.e., the number of nodes to which p is connected [*Colizza et al.*, 2004]:

$$E = \sum_{i<j} d_{ij}, \tag{27}$$

where i and j are nodes of the system, and

$$d_{ij} = \min_{P} \sum_{p \in P: i \to j} k_p^{\alpha}. \tag{28}$$

P is a path from site i to site j of the system, and p is any node belonging to such path; the distance d_{ij} is the minimum of the sum of connectivities k_p^{α}, evaluated along the path P from i to j, over all the possible paths connecting i to j. Note that in the particular case of loopless tree-like structures, such path is unique and $d_{ij} = \sum_{p \in P: i \to j} k_p^{\alpha}$. The above, new definition of weighted graph distance is meant to reproduce the conflict between two opposite and competitive trends: on the one extreme ($\alpha \to 0$), the highest connectivity among the elements of the system minimizes distances regardless of "traffic" to simply reduce the distance between vertices; on the other, the necessity to avoid (or, on the contrary to favor, e.g., for $\alpha < 0$) problems arising from highly connected nodes (some sort of bottlenecks) along the path from i to j. The weighted distance grants the functional, depending on α the concave or convex character that proved so important for the tree/network character of hydrologic networks (section 4).

[55] A continuous transition is warranted by the parameter α whose value controls the convexity of the functional. As above, chance is assumed to act through local, random changes of connectivity and necessity acts through a selection based on the global E value. All relevant features are computed via the proper adjacency and the related Laplacian matrices [*Newman*, 2003], including spectral properties of their eigenvalues, degree distributions, clustering coefficients and so on. Suffice here to mention that scale-free networks are defined by the proportion $P(k)$ of nodes having k links decaying like $P(k) \sim k^{-\gamma} F(k/K)$ where K is a suitable cutoff related to the largest hub [*Barabási and Albert*, 1999], whereas small-world constructs exhibit very small path lengths coupled with large clustering [*Watts and Strogatz*, 1998]. Many recent contributions have addressed the related features of observed network structures both natural and artificial, like, e.g., the Internet [*Barabási et al.*, 2000; *Yook et al.*, 2001], cellular and metabolic networks [e.g., *Barabási*, 2002], food webs [*Montoya and Solé*, 2002; *Garlaschelli et al.*, 2003] and the Web of scientific collaborations [*Barabási et al.*, 2002; *Newman*, 2003] among others.

[56] Notice that scale-free networks from optimal design have already been studied [*Valverde et al.*, 2002; *Venkatasubramanian et al.*, 2004] regardless of whether optimal design would pertain to stationary states of some general dynamics. Loops complicate the general picture, but the suggestion that an optimal balance of directedness (i.e., network diameter) and connectivity may, indeed, yield to optimal scale-free constructions is rather revealing [*Valverde et al.*, 2002]. Consider, for example, a model of transportation network. One would like to achieve the highest possible connectivity between the nodes, connecting each one to each other, in order to reduce the distance between them and, consequently, the costs of transportation. At the same time, it is wasteful to wire everything with everything else, because of the emergent problems of traffic passing through the highly connected nodes (or hubs), with the result of slowing down the transportation and so reducing its efficiency. Indeed, there could be cases in which it is more convenient to cover longer paths along the graph, instead of the shortest one (the graph distance, i.e., the minimum number of links necessary to move from one site to another), in order to avoid hubs along the path. The energy function E in equation (27) takes into account all these aspects, considering distance based on the sum of the nodes' degrees to the power α encountered along a path connecting two sites. The usual graph distance is recovered in the case $\alpha = 0$.

[57] The optimization method used in the numerical simulations is a regular Metropolis scheme at zero temperature of the type employed to generate the OCNs in Figure 4.

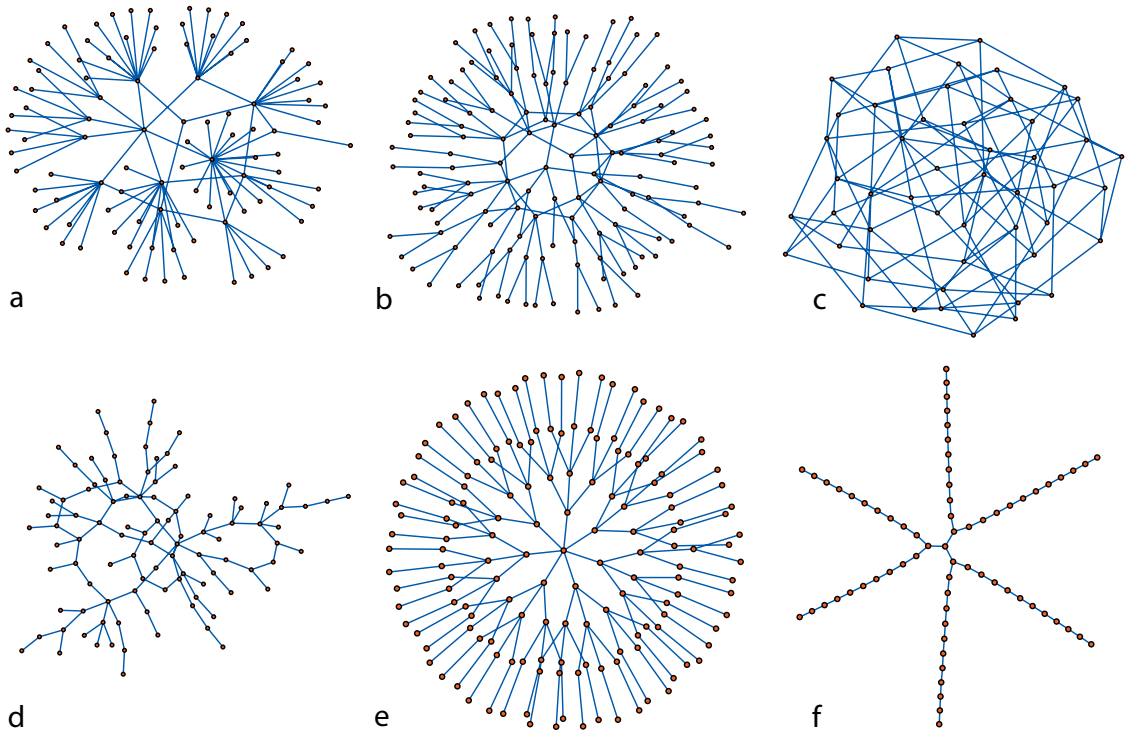

Figure 11. Sample of optimal networks: (a) $\alpha = 0.4$, $l/n = 1.05$, $n = 100$; (b) $\alpha = 0.7$, $l/n = 1.05$, $n = 140$; (c) $\alpha = 0.5$, $l/n = 2.0$, $n = 50$; (d) $\alpha = 4.0$, $n = 68$, $l = n - 1$; (e) $\alpha = 0.9$, $n = 156$, $l = n - 1$; (f) $\alpha = 2.0$, $l/n = 1.05$, $n = 100$ [after *Colizza et al.*, 2004].

It implies the following steps: (1) generation of a random initial configuration (say s) with fixed n and l, (2) random rewiring ($s \rightarrow s'$), (3) connectedness control, and (4) energetic control. For step 2, specifically, a link connecting the sites i and j is randomly chosen and substituted with a link from i to a site k, not already connected to i, extracted with uniform probability among the sites of the system, yielding a new configuration s'. In this way, the number of links l, as well as the size of the system, remains constant during the minimization. At step 3, if the graph is not connected after rewiring, step 2 is repeated. During step 4, the new value of $E(s')$ is calculated. The new configuration is accepted only if it is energetically favorable, i.e., only if $E(s') < E(s)$; otherwise the change is rejected, and we go back to step 2. Notice that the zero temperature setting (no unfavorable changes may be accepted) insures feasible optimality of the emerging network structure. The minimization algorithm stops after F consecutive failed changes on the network; here $F = n(n - 1)$, so that, on average, every pair of vertices is allowed to change its state (connected or not) twice. Every simulation is repeated 200 times, starting with different random initial configurations and varying the size n of the system and, for each size, different values of the ratio l/n.

[58] Let now consider networks with loops, so that $l > n - 1$. The optimal topologies display distinct structures as the network parameter l/n varies. Two different regimes are observed. The first occurs for values of $r = l/n$ close to 1: the system displays a critical behavior and the type of network obtained is a truncated scale-free network with a sharp cutoff at some characteristic scale dependent on the size n of the system. The second regime occurs at larger values of l/n, and the optimal degree distribution tends to peak around the average value of k, $\langle k \rangle$. Samples of network topologies obtained through such an optimization are shown in Figure 11 for different values of α and l/n [*Colizza et al.*, 2004].

[59] Thus, depending on l/n, one may select scale-free networks that display the presence of some highly connected nodes together with many peripheral and relatively unconnected sites (Figures 11a and 11c), or a network in which almost every node has the same degree $k = \langle k \rangle$ (Figures 11d and 11f). The nature of this transition can be investigated by looking at the clustering coefficient $C = \langle C_i \rangle$, defined as the average over the whole network of the clustering coefficient C_i of every node, $C_i = l_i/(k_i(k_i - 1)/2)$ where l_i is the number of existing links between the neighbors of node i, and $k_i(k_i - 1)/2$ is the total number of pairs between the neighbors. For $\alpha > 1$, the system undergoes a clear phase transition as the value of the ratio l/n increases from a regime characterized by zero clustering, to one in which clustering exist (Figure 12). This feature parallels exact results found by Erdos and Renyi for random networks [*Bollobas*, 2001].

[60] Although the ratio l/n can assume very small values, the network does not display the critical behavior observed for $\alpha < 1$, where topologies are characterized by the presence of trafficked and interconnected hubs because of the large value of α. In this regime, the competition between the minimization of the graph distance regardless of the increase in node connectivity, on the one hand, and the minimization of node connectivity on the other, is dominated by the latter. The system therefore tries to minimize the degree of each node and the result is a peaked distribution around the mean value $\langle k \rangle$, with nontrivial topology characterized by zero clustering. The network, in fact, self-

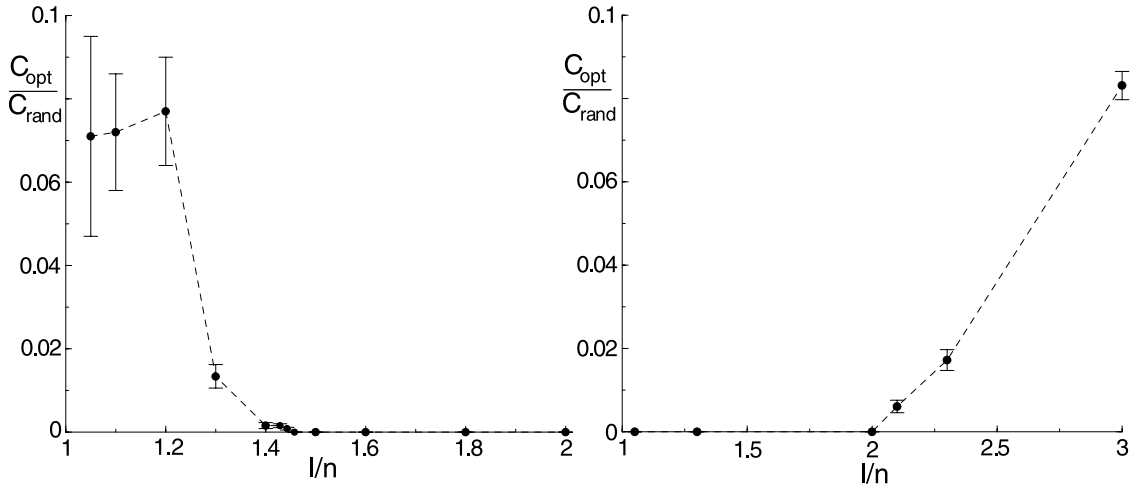

Figure 12. Phase transition in the clustering coefficient of the optimal configuration C_{opt}, normalized to the random one C_{rand}. Results are for network size $n = 70$ and values of (left) $\alpha = 0.4$ and (right) $\alpha = 2.0$ [after *Colizza et al.*, 2004].

organizes in such a way as to avoid clustering, since this would result in an increase of E. When the ratio l/n reaches a critical value (dependent on α), the system has no more advantage in creating loops of length greater than 3, and is forced to select triangles resulting in a clustering coefficient small but different from zero. Overall, optimal networks characterized by a clustering coefficient different from zero are selected as opposed to ones with no clustering at all.

[61] Note that in all regimes obtained, the optimal topology observed differs from a random network [e.g., *Bollobas*, 2001]. In fact, the characteristic path length L of optimal networks probes thoroughly structural properties. L is defined as the average, over all the possible pairs in the system, of the graph distance between pairs of nodes. In the whole interval of α investigated by *Colizza et al.* [2004], the characteristic path length of the optimal configuration, say L_{opt}, is comparable to, or smaller than, the random analog L_{rand}. Comparing the results obtained for the clustering coefficient C and for the characteristic path length L, it was also suggested that there are intervals of α and l/n in which L_{opt} is very small with respect to L_{rand} while $C_{opt} \gg C_{rand}$, so that the network displays the so-called small-world effect [*Watts and Strogatz*, 1998].

[62] Note, finally, that the idea behind some kind of selection of natural network forms is reinforced by the recent finding that they often consist of self-repeating patterns on all length scales [*Song et al.*, 2005]. This result is achieved by the application of a renormalization procedure that coarse grains the network. Interestingly, models of preferential attachment [*Barabási and Albert*, 1999; *Yook et al.*, 2001] do not reproduce the invariance properties observed in nature [*Song et al.*, 2005], whereas the optimized structures described in this section do [*Colizza*, 2004] very much like in the river case [*Rodriguez-Iturbe and Rinaldo*, 1997, section 5.10]. For a comment relevant to the possible role of selection, see also *Strogatz* [2005].

8. Conclusions

[63] The models analyzed in this paper, in spite of their simplicity, seem to capture several features of networks in nature. Though by no means exhaustive, our results show consistently that selective criteria (either derived from the physics of the phenomena or purely speculative) blend chance and necessity as dynamic origins of recurrent network patterns. The main conclusions are as follows.

[64] 1. Networks resulting from erosional dynamics in fluvial landscapes are exactly related to system configurations arising from the minimization of total energy dissipation. A remarkable feature of optimal channel networks is that they originate from an elevation field such that a local slope-area relation holds at any point, and the set of drainage directions corresponds to a given loopless network. Each loopless spanning tree, in fact, proves a local minimum of the total energy dissipation;

[65] 2. The role of selective pressures as a possible cause of emergence of the features observed in complex networks has been reviewed both in the cases of loopless and looping networks. In the latter case we show that several distinct networks structures may emerge, among which scale-free arrangements appear. Degree distributions, clustering coefficients and average path lengths of the selected aggregates show collectively the emergence of nontrivial phase transitions with increasing links-to-nodes ratios. Different features like scale-free or small-world networks are obtained for particular cases.

[66] Our main conclusion is thus that the emergence of the structural properties observed in natural network patterns may not be necessarily due to embedded rules for growth, but may rather reflect the interplay of dynamic mechanisms with an evolutionary selective process. This has implications even for hydrologic research, because many landforms originated by the collection or the distribution of hydrologic runoff (from riverine to tidal or deltaic patterns) might, indeed, be classified according to the compliance to the above mechanisms.

[67] **Acknowledgments.** This work was supported by INFM, NASA, NATO, NSF, RIMOF (Fondazione Cassa di Risparmio di Verona, Vicenza, Belluno e Ancona), COFIN 40% (Idrodinamica e Morfodinamica a Marea), and, in particular, the EU Projects (1) TIDE EU RTD Project (EVK3-CT-2001-00064) and (2) AQUATERRA (GOCE - CT 505428). The authors wish to thank Enrica Belluco and Marco Marani for their help.

References

Albert, R., and A. L. Barabàsi (2002), Statistical mechanics of complex networks, *Rev. Mod. Phys.*, *74*, 47–97.

Bak, P. (1996), *How Nature Works: The Science of Self-Organized Criticality*, Springer, New York.

Bak, P., C. Tang, and K. Wiesenfeld (1987), Self-organized criticality, *Phys. Rev. Lett.*, *59*, 381–384.

Banavar, J. R., F. Colaiori, A. Flammini, A. Giacometti, A. Maritan, and A. Rinaldo (1997), Sculpting of a fractal river basin, *Phys Rev. Lett.*, *78*, 4522–4526.

Banavar, J. R., A. Maritan, and A. Rinaldo (1999), Size and form in efficient transportation networks, *Nature*, *399*, 130–134.

Banavar, J. R., A. Maritan, and A. Rinaldo (2000), Topology of the fittest network, *Phys. Rev. Lett.*, *84*, 4745–4748.

Banavar, J. R., F. Colaiori, A. Flammini, A. Maritan, and A. Rinaldo (2001), Scaling, optimality and landscape evolution, *J. Stat. Phys.*, *104*, 1–33.

Banavar, J. R., J. Damuth, A. Maritan, and A. Rinaldo (2002), Supply-demand balance and metabolic scaling, *Proc. Natl. Acad. Sci. U. S. A.*, *99*(16), 10,506–10,509.

Banavar, J. R., J. Damuth, A. Maritan, and A. Rinaldo (2003), Allometric cascades, *Nature*, *421*, 713–714.

Barabási, A. L. (2002), *Linked*, Perseus Books, New York.

Barabási, A. L., and R. Albert (1999), Emergence of scaling in random networks, *Science*, *286*, 509–512.

Barabási, A. L., R. Albert, and H. Jeong (2000), Scale-free characteristics of random networks: The topology of the World-Wide Web, *Physica A*, *281*, 69–77.

Barabási, A. L., H. Jeong, Z. Neda, E. Ravasz, A. Schubert, and T. Vicsek (2002), Evolution of the social network of scientific collaborations, *Physica A*, *311*, 590–614.

Bollobas, B. (2001), *Random Graphs*, Cambridge Univ. Press, New York.

Brown, J. H. (1995), *Macroecology*, Univ. of Chicago Press, Chicago, Ill.

Brown, J. H., G. B. West, and B. J. Enquist (2000), Scaling in biology: Patterns and processes, causes and consequences, in *Scaling in Biology*, edited by J. H. Brown and G. B. West, pp. 1–24, Oxford Univ. Press, New York.

Calder, W. A. (1984), *Size, Function and Life History*, Harvard Univ. Press, Cambridge, Mass.

Colaiori, F., A. Flammini, A. Maritan, and J. R. Banavar (1997), Analytical solutions for optimal channel networks, *Phys. Rev. E*, *55*, 1298–1302.

Colizza, V. (2004), Statistical mechanics approach to complex networks: From abstract to biological networks, Ph.D. dissertation, Int. Sch. for Adv. Stud., Trieste, Italy.

Colizza, V., J. R. Banavar, A. Maritan, and A. Rinaldo (2004), Network structures from selection principles, *Phys. Rev. Lett.*, *92*, 198701, doi:10.1103/PhysRevLett.92.198701.

Damuth, J. H. (1998), Common rules for animals and plants, *Nature*, *395*, 115.

Dawkins, R. (1988), *The Blind Watchmaker*, Oxford Univ. Press, New York.

Dhar, D. (1999), The Abelian sandpile and related models, *Physica A*, *263*, 4–28.

Dietrich, W. E., C. J. Wilson, D. R. Montgomery, and J. McKean (1992), Erosion thresholds and land surface morphology, *J. Geol.*, *3*, 161–173.

Dodds, P. S., D. H. Rothman, and J. S. Weitz (2001), Re-examination of the "3/4-law" of metabolism, *J. Theor. Biol.*, *28*, 571–583.

Dodds, P. S., R. Muhamad, and D. J. Watts (2003), An experimental study of search in global social networks, *Science*, *301*, 827–830.

Dorogovtsev, S. N., and J. F. F. Mendes (2000), Scaling behavior of developing and decaying networks, *Europhys. Lett.*, *52*, 33–39.

Dorogovtsev, S. N., and J. F. F. Mendes (2002), Evolution of networks, *Adv. Phys.*, *51*, 1079–1187.

Doyle, P. G., and J. L. Snell (1989), *Random Walk and Electric Networks*, Am. Math. Soc., New York.

Enquist, B. J., J. H. Brown, and G. B. West (1998), Allometric scaling of plant energetics and population density, *Nature*, *395*, 163–167.

Feola, A., A. D'Alpaos, E. Belluco, S. Lanzoni, M. Marani, and A. Rinaldo (2005), A geomorphological study of lagoonal landforms, *Water Resour. Res.*, *41*, W06019, doi:10.1029/2004WR003811.

Garlaschelli, D., G. Caldarelli, and L. Pietronero (2003), Universal scaling relations in food webs, *Nature*, *423*, 165–168.

Hack, J. T. (1957), Studies of longitudinal profiles in Virginia and Maryland, *U.S. Geol. Surv. Prof. Pap.*, *294-B*, 52 pp.

Howard, A. D. (1994), A detachment-limited model of drainage basin evolution, *Water Resour. Res.*, *30*, 2261–2285.

Huber, A. (1991), Scheidegger's rivers, Takayasu's aggregates and continued fractions, *Physica A*, *170*, 463–469.

Kirchner, J. (1993), Statistical inevitability of Horton's laws and the apparent randomness of stream channel networks, *Geology*, *21*, 591–594.

Leopold, L. B., M. G. Wolman, and J. P. Miller (1964), *Fluvial Processes in Geomorphology*, W. H. Freeman, New York.

Mandelbrot, B. B. (1977), *Fractals: Form, Chance and Dimension*, W. H. Freeman, New York.

Mandelbrot, B. B. (1983), *The Fractal Geometry of Nature*, W. H. Freeman, New York.

Marani, A., R. Rigon, and A. Rinaldo (1991), A note on fractal channel networks, *Water Resour. Res.*, *27*, 3041–3049.

Marani, M., A. D'Alpaos, E. Belluco, S. Lanzoni, and A. Rinaldo (2002), Tidal meanders, *Water Resour. Res.*, *38*(11), 1225, doi:10.1029/2001WR000404.

Marani, M., S. Lanzoni, E. Belluco, A. D'Alpaos, A. Defina, and A. Rinaldo (2003), On the drainage density of tidal networks, *Water Resour. Res.*, *39*(2), 1040, doi:10.1029/2001WR001051.

Maritan, A., A. Rinaldo, R. Rigon, I. Rodriguez-Iturbe, and A. Giacometti (1996a), Scaling in river networks, *Phys. Rev. E*, *53*, 1501–1513.

Maritan, A., F. Colaiori, A. Flammini, M. Cieplak, and J. R. Banavar (1996b), Disorder, river patterns and universality, *Science*, *272*, 984–988.

Maritan, A., R. Rigon, J. R. Banavar, and A. Rinaldo (2002), Network allometry, *Geophys. Res. Lett.*, *29*(11), 1508, doi:10.1029/2001GL014533.

McMahon, T. A., and J. T. Bonner (1983), *On Size and Life*, Sci. Am. Libr., New York.

Mesa, O. J., and V. K. Gupta (1987), On the main channel length-area relationships for channel networks, *Water Resour. Res.*, *23*, 2119–2122.

Montgomery, D. R., and W. E. Dietrich (1988), Where do channels begin?, *Nature*, *336*, 232–236.

Montgomery, D. R., and W. E. Dietrich (1992), Channel initiation and the problem of landscape scale, *Science*, *255*, 826–830.

Montoya, J. M., and R. V. Solé (2002), Small world patterns in food webs, *J. Theor. Biol.*, *214*, 405–412.

Newman, M. E. J. (2003), The structure and function of complex networks, *SIAM Rev.*, *45*, 167–256.

Peano, G. (1890), Sur une courbe qui remplit toute une air plane, *Math. Ann.*, *36*, 157–160.

Peters, R. H. (1983), *The Ecological Implications of Body Size*, Cambridge Univ. Press, New York.

Rigon, R., A. Rinaldo, and I. Rodriguez-Iturbe (1994), On landscape self-organization, *J. Geophys. Res.*, *99*(B6), 11,971–11,993.

Rigon, R., I. Rodriguez-Iturbe, A. Maritan, A. Giacometti, D. G. Tarboton, and A. Rinaldo (1996), On Hack's law, *Water Resour. Res.*, *32*, 3367–3374.

Rigon, R., I. Rodriguez-Iturbe, and A. Rinaldo (1998), Feasible optimality implies Hack's law, *Water Resour. Res.*, *34*, 3181–3188.

Rinaldo, A., I. Rodriguez-Iturbe, R. Rigon, R. L. Bras, E. Ijjasz-Vasquez, and A. Marani (1992), Minimum energy and fractal structures of drainage networks, *Water Resour. Res.*, *28*, 2183–2191.

Rinaldo, A., I. Rodriguez-Iturbe, R. Rigon, E. Ijjasz-Vazquez, and R. L. Bras (1993), Self-organized fractal river networks, *Phys. Rev. Lett.*, *70*, 1222–1226.

Rinaldo, A., W. E. Dietrich, G. K. Vogel, R. Rigon, and I. Rodriguez-Iturbe (1995), Geomorphological signatures of varying climate, *Nature*, *374*, 632–636.

Rinaldo, A., A. Maritan, F. Colaiori, A. Flammini, R. Rigon, I. Rodriguez-Iturbe, and J. R. Banavar (1996), Thermodynamics of fractal networks, *Phys. Rev. Lett.*, *76*, 3364–3368.

Rinaldo, A., R. Rigon, and I. Rodriguez-Iturbe (1999a), Channel networks, *Annu. Rev. Earth Planet. Sci.*, *26*, 289–306.

Rinaldo, A., W. E. Dietrich, S. Fagherazzi, S. Lanzoni, and M. Marani (1999b), Tidal Networks: 2. Watershed delineation and comparative network morphology, *Water Resour. Res.*, *35*, 3905–3917.

Rodriguez-Iturbe, I., and A. Rinaldo (1997), *Fractal River Basins: Chance and Self-Organization*, Cambridge Univ. Press, New York.

Rodriguez-Iturbe, I., E. Ijjasz-Vasquez, R. L. Bras, and D. G. Tarboton (1992a), Power-law distributions of mass and energy in river basins, *Water Resour. Res.*, *28*, 988–993.

Rodriguez-Iturbe, I., A. Rinaldo, R. Rigon, R. L. Bras, and E. Ijjasz-Vasquez (1992b), Energy dissipation, runoff production and the three dimensional structure of channel networks, *Water Resour. Res.*, *28*, 1095–1103.

Rodriguez-Iturbe, I., A. Rinaldo, R. Rigon, R. L. Bras, and E. Ijjasz-Vasquez (1992c), Fractal structures as least energy patterns: The case of river networks, *Geophys. Res. Lett.*, *19*, 889–893.

Scheidegger, A. E. (1967), A stochastic model for drainage patterns into an intramontane trench, *Bull. Assoc. Sci. Hydrol.*, *12*, 15–20.
Scheidegger, A. E. (1970), Stochastic models in hydrology, *Water Resour. Res.*, *6*, 750–760.
Scheidegger, A. E. (1991), *Theoretical Geomorphology*, 3rd ed., Springer, New York.
Schmidt-Nielsen, K. (1984), *Scaling: Why is Animal Size so Important?*, Cambridge Univ. Press, New York.
Shreve, R. L. (1966), Statistical law of stream numbers, *J. Geol.*, *74*, 17–37.
Shreve, R. L. (1967), Infinite topologically random channel networks, *J. Geol.*, *77*, 397–414.
Shreve, R. L. (1969), Stream lengths and basin areas in topologically random channel networks, *J. Geol.*, *77*, 397–414.
Shreve, R. L. (1974), Variation of mainstream length with basin area in river networks, *Water Resour. Res.*, *10*, 1167–1177.
Sivakumar, B. (2004), Chaos theory in geophysics: Past, present and future, *Chaos Solitons Fractals*, *19*(2), 441–462.
Song, C., S. Havlin, and H. A. Makse (2005), Self-similarity of complex networks, *Nature*, *433*, 392–395.
Stevens, P. S. (1974), *Patterns in Nature*, Little, Brown, New York.
Strogatz, S. H. (2005), Romanesque networks, *Nature*, *433*, 365–366.
Takayasu, H., I. Nishikawa, and H. Tasaki (1988), Power-law mass distribution of aggregation systems with injection, *Phys. Rev. A*, *37*, 3110–3117.
Takayasu, H., M. Takayasu, A. Provata, and G. Huber (1991a), Steady-state distribution of generalized aggregation systems with injection, *J. Stat. Phys.*, *65*, 725–739.
Takayasu, H., M. Takayasu, A. Provata, and G. Huber (1991b), Statistical models of river networks, *J. Stat. Phys.*, *65*, 725–745.
Tarboton, D. G., R. L. Bras, and I. Rodriguez-Iturbe (1989), Scaling and elevation in river networks, *Water. Resour. Res*, *25*, 2037–2051.
Valverde, S., R. Ferrer Cancho, and R. V. Solè (2002), Scale-free networks from optimal design, *Europhys. Lett.*, *60*, 512–517.
Venkatasubramanian, V., S. Katare, P. R. Patkar, and F. P. Mu (2004), Spontaneous emergence of complex optimal networks through evolutionary adaptation, *Comput. Chem. Eng.*, *28*(9), 1789–1798.
Watts, D. J. (1999), *Small Worlds: The Dynamics of Networks Between Order and Randomness*, Princeton Univ. Press, Princeton, N. J.
Watts, D. J., and S. H. Strogatz (1998), Collective dynamics of "small world" networks, *Nature*, *393*, 440–442.
West, G. B., J. H. Brown, and B. J. Enquist (1997), A general model for the origin of allometric scaling laws in biology, *Science*, *276*, 122–126.
West, G. B., J. H. Brown, and B. J. Enquist (1999a), The fourth dimension of life: Fractal geometry and allometric scaling of organisms, *Science*, *284*, 1677–1679.
West, G. B., J. H. Brown, and B. J. Enquist (1999b), A general model for the structure and allometry of plant vascular systems, *Nature*, *400*, 664–667.
Willgoose, G. R., R. L. Bras, and I. Rodriguez-Iturbe (1991), A coupled channel network growth and hillslope evolution model: 1. Theory, *Water Resour. Res.*, *27*, 1671–1684.
Yook, S. H., H. Jeong, A. L. Barabási, and Y. Tu (2001), Weighted evolving networks, *Phys. Rev. Lett.*, *86*, 5835–5838.

J. Banavar, Department of Physics, Pennsylvania State University, University Park, PA 16802, USA. (jayanth@phys.psu.edu)

A. Maritan, Dipartimento di Fisica "Galileo Galilei," Universitá di Padova, via Marzolo 9, I-35131 Padova, Italy. (amos.maritan@padova.infm.it)

A. Rinaldo, Dipartimento di Ingegneria Idraulica, Marittima, Ambientale e Geotecnica, Universitá di Padova, via Loredan 20, I-35131 Padova, Italy. (rinaldo@idra.unipd.it)