

University of Montana

ScholarWorks at University of Montana

Graduate Student Theses, Dissertations, &
Professional Papers

Graduate School

2006

Allometric scaling in river basin models

Ted F. Cosart

The University of Montana

Follow this and additional works at: <https://scholarworks.umt.edu/etd>

Let us know how access to this document benefits you.

Recommended Citation

Cosart, Ted F., "Allometric scaling in river basin models" (2006). *Graduate Student Theses, Dissertations, & Professional Papers*. 8001.

<https://scholarworks.umt.edu/etd/8001>

This Thesis is brought to you for free and open access by the Graduate School at ScholarWorks at University of Montana. It has been accepted for inclusion in Graduate Student Theses, Dissertations, & Professional Papers by an authorized administrator of ScholarWorks at University of Montana. For more information, please contact scholarworks@mso.umt.edu.



**Maureen and Mike
MANSFIELD LIBRARY**


The University of
Montana

Permission is granted by the author to reproduce this material in its entirety, provided that this material is used for scholarly purposes and is properly cited in published works and reports.

****Please check "Yes" or "No" and provide signature****

Yes, I grant permission _____

No, I do not grant permission _____

Author's Signature: 

Date: 5/30/2006

Any copying for commercial purposes or financial gain may be undertaken only with the author's explicit consent.

ALLOMETRIC SCALING IN RIVER BASIN MODELS

by

Ted Cosart

B.A., University of Montana, 1983

presented in partial fulfillment of the requirements

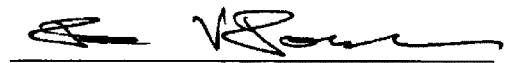
for the degree of

Master of Science

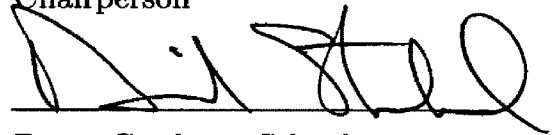
The University of Montana

May, 2006

Approved by:



Chairperson



Dean, Graduate School

5/26/06

Date

UMI Number: EP38802

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



UMI EP38802

Published by ProQuest LLC (2013). Copyright in the Dissertation held by the Author.

Microform Edition © ProQuest LLC.

All rights reserved. This work is protected against unauthorized copying under Title 17, United States Code



ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 - 1346

Allometric Scaling in River Basin Models

Chairperson: Dr. Jesse Johnson 

River basins scale allometrically, flow rate to mass of material transported. They exhibit a scaling index near a known theoretical limit for efficient transportation networks. Models called optimal channel networks (OCN's) also approximate the scaling index using minimum energy dissipation. OCN's based on an initial loopless network and optimization with simulated annealing show this allometric scaling with an index very close to that predicted for efficient transportation networks. Models based on the same initial conditions, and optimized to reach the scaling index itself, rather than minimum total energy dissipation, show a much higher total energy dissipation than do the OCN models, an unnatural direction of flow away from outlets in a multi-outlet configuration, and unnaturally circuitous flow directions in the single-outlet model. The results suggest that minimization of material in an allometric relationship is an insufficient condition for the creation of efficient transportation network structures found in nature.

ACKNOWLEDGMENTS

This thesis owes its motivating idea to Dr. Jesse Johnson, who saw interesting questions lurking unanswered in the literature. Gratefully I acknowledge his generosity with ideas and time, as I implemented the tools we used to explore some of those ideas. Any progress made in addressing them is due in large part to his insight and good judgment.

I thank Dr. Alden Wright and Dr. Leonid Kalachev for their participation on my thesis committee, as well as for their corrections of and insights into the work.

For encouragement and advise in persuing a Master's degree, I thank Professor Darrell Hagen of Carroll College, friend and mentor.

I thank my brothers and J K Weiss for sharing their enthusiasm and skills with computers and programming, providing me a gateway into Computer Science.

I also seize this opportunity to pay respect to what was bequeathed to my brothers and me by our parents: respect for learning and intellectual persuit, whatever the field.

TABLE OF CONTENTS

ACKNOWLEDGMENTS	iii
CHAPTER 1 INTRODUCTION	1
1.1 The Forces Behind Networks	1
1.2 Allometric Scaling Limits	2
1.3 Self Similarity in River Basins	4
1.4 Allometry in River Networks	5
1.5 Models of River Basin Geometry	8
1.6 Goals of This Study	9
CHAPTER 2 METHODS	10
2.1 The River Basin Models	10
2.1.1 Space-filling Grid	10
2.1.2 OCN Models	11
2.1.3 Scaling-optimized Models	12
2.2 Optimization with Simulated Annealing	12
2.2.1 The Principle Behind Simulated Annealing	12
2.2.2 Simulated Annealing Algorithm	14
2.2.3 Annealing parameters	16
2.3 Measurements and Error	18

2.3.1	Total Energy Dissipation	18
2.3.2	The Scaling Index	19
CHAPTER 3	RESULTS	22
3.1	Network Structures	22
3.2	Area to Mass Scaling	23
3.3	Cumulative Distributions	37
3.3.1	Total Contributing Area	37
3.3.2	Total Contributing Area of Drainage Basins	40
3.3.3	Energy Dissipation per Unit Channel Link Length	42
3.4	Total Energy Dissipation	44
3.5	Mean Distance of Site to Outlet	47
3.6	Summary of Power Laws	49
CHAPTER 4	DISCUSSION	50
4.1	Assumptions About the Models	50
4.2	The SON Model and the Theoretical Scaling Limit	51
4.3	Future Directions with New Models	54
APPENDIX A	Table of Most-Used Variables	61
APPENDIX B	Random Number Generation	62
APPENDIX C	A Note on Performance with Different C Compilers	64
BIBLIOGRAPHY	66

LIST OF TABLES

Table 3.1	Hack's Exponent from α versus Mean Distance	49
Table 3.2	Summary of the Scaling Indices of Power Laws	49
Table A.1	Most-used Variables	61

LIST OF FIGURES

Figure 1.1	Example of steepest-descent flow on a 3×3 hexagonal grid.	7
Figure 2.1	Triangular lattice of 8×8 sites.	11
Figure 2.2	Bin sizes for binned plots of $\log_{10}A_X$ vs. $\log_{10}M_X$	20
Figure 3.1	Eden Growth Model, the Random Initial Configuration . . .	24
Figure 3.2	Detail, Lower Left Corner, Full Resolution, Eden Model . . .	25
Figure 3.3	Multi-outlet OCN	26
Figure 3.4	Detail, Lower Left Corner, Full Resolution, Multi-outlet OCN	27
Figure 3.5	Single-outlet OCN	28
Figure 3.6	Detail, Lower Left Corner, Full Resolution, Single-outlet OCN	29
Figure 3.7	Multi-outlet SON	30
Figure 3.8	Detail, Lower Left Corner, Full Resolution, Multi-outlet SON	31
Figure 3.9	Single-outlet SON	32
Figure 3.10	Detail, Lower Left Corner, Full Resolution, Single-outlet SON	33
Figure 3.11	Networks Sampled During OCN Optimization	34
Figure 3.12	Networks Sampled During SON Optimization	35
Figure 3.13	Networks Sampled During SON Optimization, Continued . .	36
Figure 3.14	Network Mass versus Total Contributing Area.	38
Figure 3.15	Network Mass versus Total Contributing Area, Unbinned Data.	39

Figure 3.16	Cumulative Contributing Area	41
Figure 3.17	Cumulative Contributing Area of Drainage Basins	43
Figure 3.18	Cumulative Energy Dissipation per Unit Channel Link Length	45
Figure 3.19	Total Energy Dissipation over Annealing Stages	46
Figure 3.20	Mean Distance to Outlet	48
Figure 4.1	Scaling Index α for Large Basin Areas, SON	52
Figure 4.2	Sub-basins Extracted by Size in Single-outlet Models	58
Figure 4.3	New Model Structure	59
Figure 4.4	Preliminary Results for Proposed New Model, Scaling Measurements. The dashed lines in the area and dissipation plots show the slope that gives the scaling index in real river basins.	60

CHAPTER 1 INTRODUCTION

1.1 The Forces Behind Networks

Recent literature across the sciences investigates the driving forces behind the formation of networks. Conceptualized as the set of links between nodes, the networks studied are either those for which the cost of a link is mostly independent of physical length or those for which that cost mostly depends on the links physical length. Links between sites on the world wide web typify the first kind, while river basins, animal circulatory systems, and city bus routes are examples of second. How the latter networks assume the forms we see qualitatively and assess geometrically is a problem whose solution offers considerable benefit. For example, understanding river networks aids management of scarce water resources. Ability to assess the anticipated form of a blood vessel network enhances medical diagnosis.

This work explores the formation of river basin networks in light of 2 well known observations: they exhibit minimal energy dissipation on all scales, and a minimal amount of transported material on all scales. Minimal energy dissipation networks have been modeled and the results reported extensively (Rodriguez-Iturbe, *et. al.* [6], Maritan, *et. al.* [4], Meakin, *et. al.* [7]). The second observation is newer, and has not been used as the constructive force in modeling river basins. This work reproduces known models of minimum energy dissipation, and makes new models

based on minimization of transported material. Several measurements are made on the models to explore their differences and similarities.

1.2 Allometric Scaling Limits

Central to this work is the finding in Banavar, *et. al.* [2] that some natural transportation networks show an allometric scaling relationship, transportation-rate to amount of material transported, close to a theoretical limit of minimum transported material. Banavar, *et. al.* [2] classify such networks as most efficient, since the material transported is as small as possible for the area the network is serving. They establish a theorem that gives the limit of such scaling. With C as the amount of transported material¹, L as the linear size of the area served by the network they give the following theorem,

*For any spanning network in D dimensions,
 C scales at least L^{D+1} and at most as L^{2D} .
for large L .*

The theorem leads to a minimum and maximum scaling relationship between material transported and the linear size of the area served:

$$C \propto \begin{cases} L^{2D} & \text{Maximum Material} \\ L^{D+1} & \text{Minimum Material} \end{cases} \quad (1.1)$$

¹ C is first defined in Banavar, *et. al.* [2] as the total blood volume in an organism's metabolic system, and then redefined, generalized in the context of transportation networks to be the "quantity of nutrients in the network at any instant of time." The meaning of "nutrients" in this general context is unclear.

Given these limits and the relationship between rate of transport B and linear size,

$$B \propto L^D, \quad (1.2)$$

the authors show that the scaling relationship between rate-of-transport and network mass is constrained. In the least-efficient case,

$$B \propto L^D \quad (1.3)$$

$$C \propto L^{2D} \quad (1.4)$$

$$B \propto C^{\frac{1}{2}} \quad (1.5)$$

The least-efficient network is exemplified by a “space-filling spiral ” in which flow accumulates along a single path that covers the entire area served.

At the other extreme are networks like river basins. In such networks the flow along each link is “directed ”, that is, the material is moving along shortest paths away from a single source (in circulatory systems) or toward a single sink (in a river basin)². Here, a minimal amount of material is used for the given service area, leading to the classification in Banavar, *et. al.* [2] as “most efficient”. Such networks in D dimensions, with transport rate B and amount of material transported C satisfy,

$$B \propto L^D \quad (1.6)$$

$$C \propto L^{D+1} \quad (1.7)$$

$$B \propto C^{\frac{D}{D+1}} \quad (1.8)$$

²See Banavar, *et. al.* [2], supplementary material, for a more rigorous definition of a “directed” spanning tree, and a proof of their theorem.

1.3 Self Similarity in River Basins

River basins networks share with other fractal objects self-similar properties evinced in scaling relationships given by power laws of the form

$$g(x) = ax^\alpha \tag{1.9}$$

where a and α are constants, x represents a measured property such as area and $g(x)$ a function of the property, such as the cumulative distribution of sub-basin areas. The relationship in Equation 1.9 holds when x is multiplied by a constant (Rodriguez-Iturbe, *et. al.* [6]). This means that any portion of the basin shows the same relationship between $g(x)$ and x (Rodriguez-Iturbe, *et. al.* [6]). The fixed proportionality across all scales earns such congruencies the terms “scale free” and “scale invariant”.

River basins are evaluated for self-similarities through analysis of flow networks constructed from a digital elevation map (DEM). The pixel gives an area of uniform elevation (30-200 square meters are common resolutions). Flow networks are constructed on the grid, commonly through steepest descent directions from each pixel to one of eight surrounding pixels (2 horizontal, 2 vertical, 4 diagonal neighbors). The flow accumulation along the elevation gradients can then be calculated, and related to area and other properties³. From these measurements the scale invariant properties, upholding Equation 1.9 are revealed.

Self-similar properties in river basins often complementary cumulative probability

³For a discussion of the geomorphologic concepts and the difficulties related to these map extractions, see Rodriguez-Iturbe, *et. al.* [6], Section 1.2.9.

distributions expressed in the form

$$P(X > x) \propto x^\alpha \tag{1.10}$$

$$\tag{1.11}$$

where X is a random variable representing an instance from the range of values of x , so that $P(X > x)$ gives the frequency with which an instance of X exceeds x . Other examples of self-similarity are found in size relationships and the quantities (like flow-rate) that can be related to size. The theorem of allometric scaling described in Section 1.2 gives such a relationship.

1.4 Allometry in River Networks

In confirming the predictions of the theorem, Banavar, *et. al.* [2] measure DEM-derived river basin networks and report the expected relationship. In most-efficient transportation networks, where B is the flow rate into a given site, M is the amount of material transported by the network, here called network mass ⁴, and D is the network dimension,

$$M \propto B^{\frac{D+1}{D}} \tag{1.12}$$

$$D = 2 \tag{1.13}$$

$$M \propto B^{\frac{3}{2}} \tag{1.14}$$

⁴Dreyer, [3], in deriving separately from Banavar, *et. al.* [2] the same allometric scaling relationship between rate and amount of material transported, computes the quantity giving the amount of material as exactly the mass of the transported material. Thus, following Dreyer, [3], M will be called network mass.

with the relationship holding across all scales, as indicated by allometry.

In the DEM measurements flow rate B is replaced by a proxy, *total contributing area* (Banavar, *et. al.* [2])⁵. As described in Banavar, *et. al.* [2], the total contributing area A_X , at a given site X can be computed as the sum of the areas of its contributing nearest neighbors, nn , plus the single unit of area representing the DEM pixel⁶ Adopting the terms in Banavar, *et. al.* [2] flow, as area, is expressed recursively:

$$A_X = \sum_{Z \in nn(X)} A_Z + 1 \quad (1.15)$$

Again following Banavar *et. al.* [2] The network mass, M , is expressed here as the summed total contributing areas of the collection γ of sites that drain in to any site X . The area (or, equivalently, flow) at site X itself is excluded from the sum, as M denotes the material of the network that empties into X .

$$M_X = \sum_{Z \in \gamma} A_Z \quad (1.16)$$

Figure 1.1 illustrates the calculations of total contributing area and network mass. With proxies for flow and network mass, total contributing area and sum of total contributing area, then, the allometric scaling relation given by Equation 1.12 can be expressed:

$$M_X \sim A_X^\alpha, \alpha = \frac{3}{2} \quad (1.17)$$

⁵The validity of total contributing area as a proxy for flow is demonstrated in Rodriguez-Iturbe, *et. al.* [6], Section 4.6.

⁶The single unit of area is analogous to a single unit of precipitation injected at site, leading to sites with increasing accumulations of flow equal to their increasing total contributing areas along flow paths from sources to downstream sites.

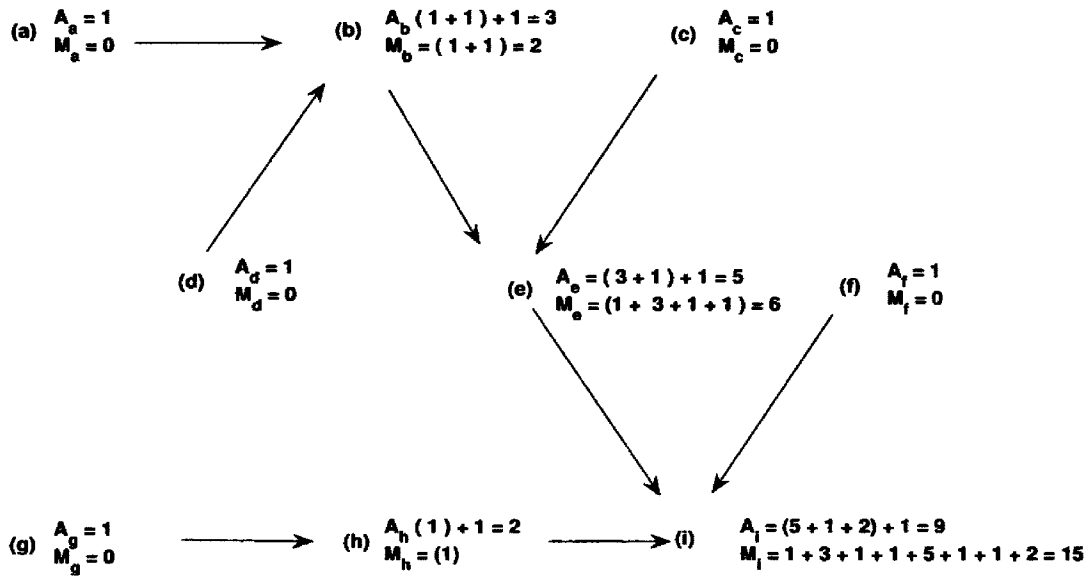


Figure 1.1 Example of steepest-descent flow on a 3 × 3 hexagonal grid.

Each letter names a site on this single-outlet network. Each site is centered on a pixel, an area of uniform elevation. The arrows show the direction of steepest-descent flow. To the right of each letter is the sum that gives that pixel's total contributing area (A_x) and total material transported by the sub-network that empties in to that pixel (mass, M_x). Note that sites a, c, d, f, g represent ridges, points of highest elevation relative to neighboring pixels. As such they receive no upstream flow, which gives the networks they empty zero mass and a single unit of contributing area, that of the pixel itself.

1.5 Models of River Basin Geometry

River basins have been modeled using principles of minimum energy dissipation. These models are called optimal channel networks (OCN's). In Rodriguez-Iturbe, *et. al.* [6], it is shown that over any link in a river channel the optimal energy expenditure P is related to mean annual flow Q by,

$$P = kQ^{0.5}L, \quad (1.18)$$

with k a constant and L the length of the link. Given this link between flow and optimal energy dissipation, Rodriguez-Iturbe, *et. al.* [6] further analyze the relationship between discharge and structure of basins. Both empirical evidence and their analysis lead to 3 postulates, given a set of known discharge values throughout a basin:

1. Minimum energy expenditure in any link of the network.
2. Equal energy expenditure per unit area of channel anywhere in the network.
3. Minimum energy expenditure in the network as a whole.

Meakin, *et. al.* [7] present a simulation of multiple river basins distributed over an area with a square boundary, created according to these 3 principles of minimum energy dissipation. As indicated by principle 1, Meakin, *et. al.* [7] describe energy dissipation at any link i in the network by $P_i = kQ_i^{0.5}L_i$, with Q_i as the mean annual flow in the link and L as the link length. Energy dissipation for a basin (the area drained by a collection of sites flowing into a single outlet) is given by the sum:

$$P_j = \sum_i kQ_i^{0.5}L_i \quad (1.19)$$

An area of basins optimized to minimum energy dissipation should show a minimum value E for the sum:

$$E = \sum_j P_j \quad (1.20)$$

Single outlet basins, reproduced by several investigators (Rodriguez-Iturbe, *et. al.* [6], Banavar, *et. al.* [2], Maritan, *et. al.* [4]) and the multi-outlet model described in Meakin, *et. al.* [7], show power-law scaling properties similar to those in real river basins.

1.6 Goals of This Study

While river network models based on minimum energy dissipation have been measured for the mass-to-area scaling index given by Equation 1.17 (see Maritan, *et. al.* [4]), the implications of the difference between the ideal index and the measured value is explored by creating a network model whose formation is based on achieving the allometric scaling index $\alpha = 1.50$. This model is compared with minimum energy dissipation models, to show the consequence of mass-minimization as the driving force behind network formation.

Also, since the results of the work on the multi-outlet OCN model in Meakin, *et. al.* [7] were published before Banavar, *et. al.* [2] published their theorem about allometric scaling, this study also tests the multi-outlet model for the expected scaling relationship of flow-rate to material-transported.

CHAPTER 2 METHODS

2.1 The River Basin Models

2.1.1 Space-filling Grid

Most of the models of river basin geometry used in this thesis are based on that described in Meakin, *et. al.* [7], using a triangular lattice of network sites. In graph terms the network model created on the lattice is a tree or forest, depending on whether a single or multiple outlets are used. A river basin is represented by a spanning tree whose edges are links representing flow from one network site (node) to one of its neighbors, the flow terminating in an outlet, or a sink. A given site's neighborhood is the hexagonal constellation of sites nearest it. Border sites, corner sites, and some sites adjacent to border sites do not have a full complement of neighbors.

Each site receives one unit of flow as "precipitation", representing a mean measure of rainfall into the system. Flow accumulates along the edges from sources (sites that receive no flow from a neighbor) to the outlet. The flow at a site is calculated as the sum of the flow values of contributing neighbors (if any) plus the one unit of "precipitation", following the recursive equation for total contributing area (Equation 1.15). Figure 2.1, shows the arrangements of sites and borders on the triangular lattice. The single-outlet model differs only in that all flow from all sites is emptied into one outlet, usually a border site. Note that the border sites along the vertical edges of the grid, as shown in Figure 2.1, number less than the horizontal, as they

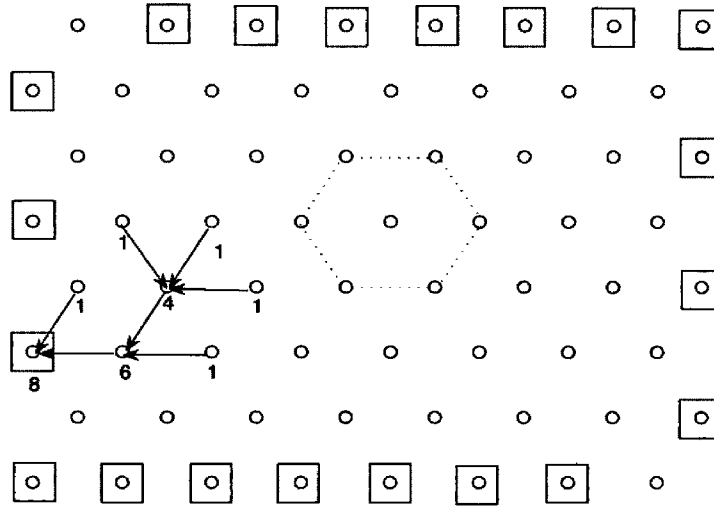


Figure 2.1 Triangular lattice of 8×8 sites.

One hexagonal neighborhood is shown with dotted lines. Arrows indicate flow from sources (arbitrarily chosen) into an outlet, the path labeled with each site's flow value, or, equivalently, total contributing area. All border sites are in boxes. The distance between columns is defined as 1 unit. The distance between rows is reduced to $\frac{\sqrt{3}}{2}$ units, so that the length of a link from a site to any of its neighbors is 1 unit.

alternate with the offset rows. This scheme may be different from the border-outlet scheme given in Meakin, *et. al.* [7], which does not specify whether all the sites along the vertical edges of the grid are border sites.

2.1.2 OCN Models

Minimum energy dissipation models, also known as optimal channel networks (OCN's), with one and with multiple outlets, were created by optimizing the total energy dissipation of the system using simulated annealing, detailed in Section 2.2. While the border outlets may differ in number from the models described in

Meakin, *et. al.* [7] (the authors do not specify which sites are boundary sites), the multiple-outlet model used here attempts otherwise to implement that described in Meakin, *et. al.* [7]. The optimized value in either a multi- or single-outlet OCN is the sum of the energy dissipation of all sites i on the grid, the quantity E in Equation 1.20, with $j = 1$ in the case of the single-outlet model.

2.1.3 Scaling-optimized Models

Scaling-optimized networks (SON's) differ from the OCN models in the optimization function and the parameters used in simulated annealing. The target function is that which gives the scaling index α in the scaling relation (see Equation 1.17). The index is computed as the slope of the least-squares-fit line of the double logarithmic plot of Equation 1.17 at each site. It is compared with the target slope, the theoretical limit of $\frac{3}{2}$. The optimality of a network configuration is given by the proximity of its scaling index to the ideal slope.

2.2 Optimization with Simulated Annealing

2.2.1 The Principle Behind Simulated Annealing

Press, *et. al.* [5] notes that simulated annealing is a good method for optimizing large discrete configuration spaces with many poor local extrema that can hide good global ones, and with factorially large numbers of configurations. It has (as have other methods) effectively solved the traveling salesman problem, and is used to optimize complex circuit designs (Press, *et. al.* [5]). The algorithm uses a process analogous to thermodynamic processes, seen in “the way liquids freeze and crystallize, or metals cool and anneal” Press, *et. al.* [5]. It is an optimizing process in that, in slowly cooling systems, “nature is able to find [a] minimum energy state” Press, *et. al.* [5].

A system probabilistically changes energy states, such that, even at a relatively low temperature, there is still a (small) probability that a component may be in a state of high energy. The Boltzmann distribution describes the probability distribution as a relationship between temperature T and energy E , a constant k :

$$D(E) = e^{\frac{-E}{kT}} \quad (2.1)$$

When used as a computational algorithm to search for global optima in a solution space, the general application of simulated annealing includes:

1. Establishment of a “cooling” schedule for the reduction of T in Equation 2.1. The schedule tells how slowly T , an analog to temperature in the thermodynamic system, will be lowered, that is, how slowly to lower the probability that the current solution will be replaced by a counter-optimal solution.
2. Between temperature reductions, a series of random solutions are generated. To each new solution the Boltzmann probability distribution is applied, such that solutions more optimal than the current one are always accepted (the new solution becomes the current one), and solutions less optimal are accepted when the Boltzmann equation yields a higher probability than a random probability. Otherwise the current solution does not change.

Describing their solution to the traveling salesman problem, Press, *et. al.* [5] note that experimentation is needed in finding the correct schedule. Such trial-and-error adjustments in the schedule improved the non-OCN models described here. Unlike the OCN models (with parameters supplied by Meakin, *et. al.* [7]) their optimization procedure lacked a proven set of parameters.

2.2.2 Simulated Annealing Algorithm

Meakin, *et. al.* [7] gives the following annealing algorithm, used as described here for creating optimal channel networks (OCN), and modified as noted in creating slope-optimized networks (SON):

1. Make an Eden growth model¹ of the network, using boundary sites as collection sites, until the whole grid is filled.
2. Calculate the total contributing area for each site: the sum of the contributing area of neighbors plus 1 (see Equation 1.15). Given the area calculation, ascertain other quantities according to the desired optimization:
 - (a) For OCN's, calculate the total energy dissipation, E (see Equation 1.20), with $k = L = 1$. As seen in Figure 2.1, the length of each link is 1, and k , as a constant of proportionality, can be set to 1 without loss of accuracy in relative total energy dissipation, the minimized quantity.
 - (b) For SON's, calculate the scaling index α (see Equation 1.17) with least squares regression on the set of all lattice points with mass greater than zero.
3. Randomly select one site in the grid and change its flow direction.
4. If the change creates a loop in the network, repeat (3).

¹The Eden model is a loopless tree or forest produced by a random walk, producing one link per step. The randomness is dampened by the constraint that a given new link must connect a non-linked site with a linked site, so that, at a given step only neighbors to currently linked sites are eligible to become part of a network. Thus, for example, the initial link in the Eden model must represent flow from a site neighboring an outlet into the outlet, the second link must either connect to the just-linked non-outlet site, or a border site. This constraint establishes the local connectivity and loopless (tree) topology of river basin networks.

5. Given the valid perterbation (change of flow in one site), recalculate the flow for the new configuration, and then the appropriate value:
 - (a) For OCN's, total energy dissipation, E'
 - (b) For SON's, scaling index α'
6. If the new configuration is better, adopt it as the current:
 - (a) for OCN's, if $E' < E$ adopt the new configuration.
 - (b) For SON's, in most cases, the squared difference between α' and α was the minimized quantity, so that, with α_{target} as the sought scaling index, if $(\alpha' - \alpha_{target})^2 < (\alpha - \alpha_{target})^2$, new configuration was adopted. ²
7. Otherwise compute the Boltzmann probability p for the current temperature T .
 - (a) For OCN's, as given by Meakin, *et. al.* [7], the analog to the Boltzmann probability is,

$$p = e^{-(E'-E)/T} \quad (2.2)$$
 - (b) For SON's as an aid to tuning the annealing schedule (discussed below in 2.2.3) the constant k in the original Boltzmann (see Equation 2.1) was included, and given a value less than unity to attenuate T , with 0.6 the best found, so that the probability calculation becomes:

$$p = e^{-(\alpha'-\alpha)^2/0.6T}. \quad (2.3)$$

²In the early attempts at tuning the SON optimization, a simpler comparison was used, such that, similar to the OCN comparison, a smaller or larger α' was a sufficient test for a better configuration. This was possible because the progress of the scaling index approached the target strictly from above or below, deepening on the relationship being expressed in the optimization. SON's optimized to reach an index if $3/2$, to satisfy $A_X \sim M_X^{3/2}$ never achieved a value under $3/2$, while the scaling index in those optimized to the inverted index to satisfy the relationship, $M_X \sim A_X^{2/3}$ never exceeded $2/3$. The latter were earlier models, and the former were preferred as the results were closer to the ideal given the same annealing parameters.

Given p , sample a random number r from a uniform distribution between 0 and 1, and compare r and p . If $r < p$ accept the new configuration. Otherwise, reject it.

8. Repeat (3) - (7) many times, each iteration comprising a step at a given temperature.
9. After a given number of steps (discussed below in 2.2.3), and while the temperature is above some set minimum, lower the temperature, and repeat (3) - (8).

2.2.3 Annealing parameters

The following parameters were set using Meakin, *et. al.* [7] as the guide:

1. **The number of steps.** This gives the number of new configurations generated between temperature reductions. The authors used $256^2 = 65,536$ steps for their grid of size $l = 256$. In this study, this proportion was adopted for all sizes, so that all schedules set l^2 steps between temperature reductions for a network of size $l \times l$.
2. **Initial temperature, $T(0)$:** Initial temperature was determined according to,

$$T(0) = \lambda P(0), \tag{2.4}$$

where $P(0)$ is the total energy dissipation in the original configuration, and the constant λ is adjusted according to the model.

- (a) OCN models: Meakin, *et. al.* [7] chose $T(0) = 12,800$. This is about $\frac{1}{20}$ of the total energy dissipation of an initial eden model of the multi-outlet

model used by the Meakin, *et. al.* [7], of size $l = 256$ (the average for 22 examples of multi-outlet eden models of size $l = 256$ was about 12,752). While different values for λ were tried³, this standard seemed to produce the best results in OCN models, and the $T(0)$ was set with Equation 2.4 and $\lambda = 0.05$, for all of the OCN networks produced.

- (b) SON models: Trial and error, after initially using the same initial temperature as that for the OCN's, showed that a much larger larger proportion of the initial energy dissipation, with the other parameters set as noted, gave a better $T(0)$, the best results achieved with $\lambda = 0.7$.

3. Temperature adjustment: In Meakin, *et. al.* [7] the temperature for the n^{th} stage, set at the end of stage number $(n - 1)$ is given by

$$T(n) = \beta^n T(0), \tag{2.5}$$

with β set according the type of network:

- (a) OCN models: Following Meakin, *et. al.* [7], a reduction schedule with $\beta = 0.982$ was adopted for all.
- (b) SON models: Experimentation showed that results improved inversely with the steepness of temperature change. Long, gradual series of temperature reductions, all other parameters being equal, gave the best results. In the most effective set of parameters, the value was $\beta = 0.996$.

³Trials on OCNs using a constant $T(0) = 12800$ for grid sizes smaller than $l = 256$ showed that T should be adjusted to reflect the initial energy dissipation value, since relatively large T yields small values of the exponent in the Boltzmann equation and so gives relatively large Boltzmann probability values, so that the "temperature" stays "hotter" longer.

4. **The minimum temperature** is the value for T which signals the end of the annealing process. It was set according to network type:

- (a) OCN models were optimized using the minimum temperature after Meakin, *et. al.* [7], of 0.0004
- (b) SON models, as noted above, benefited from longer annealing scheduling, and trial and error produced a most effective set of parameters with the minimum temperature set at 4×10^{-11} .

2.3 Measurements and Error

2.3.1 Total Energy Dissipation

As seen in the description of the simulated annealing algorithm, before beginning the optimization process, the total energy dissipation of an OCN model is computed. This initial energy calculation is a summing of energy dissipation of all the sites in the model, producing the quantity given in Equation 1.20, with $k = L = 1$ (see 2.2.2, item 2). Thereafter the total is adjusted according to each new configuration during the annealing process. Each change is a perturbation of a single site's direction of flow, so that only the flow values of sites "downstream" in the old and new directions need be diminished and augmented accordingly. The total energy dissipation, then, is adjusted by subtracting from the current total the individual energy dissipation values in the affected sites, and adding to the total their new individual energy dissipation values. Given roundoff error and other problems associated with repeated computations using 64-bit IEEE floating point operations, a check for accuracy in the final result was made using 20 runs of multiple-outlet OCN models. The final total energy dissipation at program's end resulting from all the adjustments made over the annealing process

(here called the program total), was comparing with a new calculated total over all the flow values in the sites in the final model configuration. The 2 totals differed in magnitude by no more than 0 ± 10^{-4} for all 20 models, whose total dissipation values (as measured by the program total) averaged 1.9571×10^5 , with a standard deviation of about 307.

2.3.2 The Scaling Index

As noted in the description of the models in Section 2.1.3 the allometric scaling index, α in Equation 1.17, was computed as the slope of the line of least squares fit to the double logarithmic plot of the total contributing area (see Equation 1.15) versus the total network mass (see Equation 1.16).

In computing α during simulated annealing on the SON models, The least-squares line was fit to the all of the data points, $\log_{10}A_X$ and $\log_{10}M_X$ (A_X is the total contributing area at each site, and M_X is the network mass for the sub-basin that empties into the site, given in Equation 1.16).

Binning Procedure In plotting the index of the final models it was noted that in OCN models the raw data yielded slopes that deviated significantly from the examples of real river basins and OCN models in the literature (Banavar, *et. al.* [2], Maritan, *et. al.* [4]). It was also noted that the plots in the literature used a binning method, such that the plotted data points are, as explained in Banavar, *et. al.* [2], “obtained by binning total areas, and computing the ensemble average of the sum of the inner areas for each sub-basin within the binned interval.” Maritan, *et. al.* [4] also used this method to plot the data obtained for real river basins. The method used in measuring the index in the models is not specified. In this study the binning method just described was used to obtain the data points in the optimized

models. More precisely, $\log_{10}A_X$ was binned into 4 bins per decade. Per-bin averages of $\log_{10}A_X$ and $\log_{10}M_X$ were used to obtain a least-squares fit line, its slope giving the index α . Besides the best fit line, $\log_{10}M_X$ is plotted against the midpoint of each bin (see Figure 3.14. Bins with less than 10 data points were not used in the calculation.

The number of sites in a size range (as measured by total contributing area) declines as the site-size increases, as expected, since sites of relatively large flow accumulation must number fewer than the sites whose flow is accumulated therein. Figure 2.2 shows the bin sizes for each bin for both OCN networks and SON networks. I

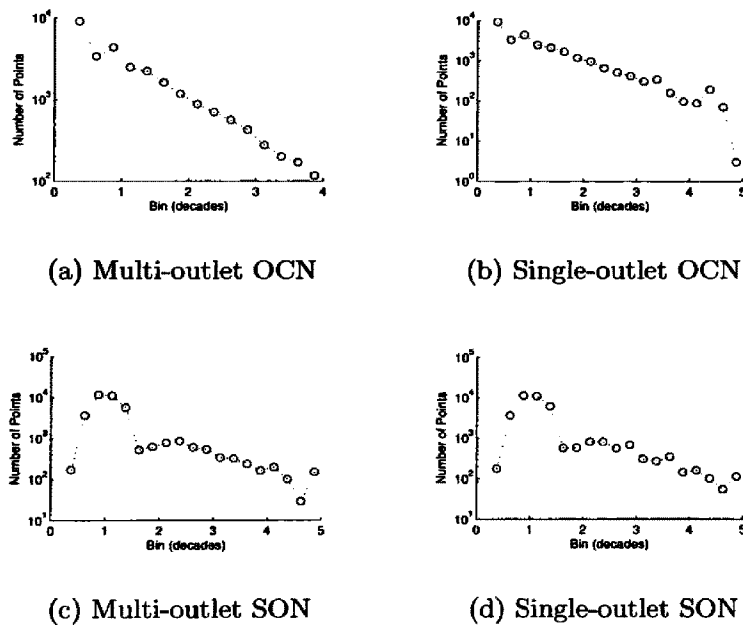


Figure 2.2 Bin sizes for binned plots of $\log_{10}A_X$ vs. $\log_{10}M_X$

The plots show the number of sites vs. the bins of $\log_{10}A_X$. Note that in all examples the first bin, containing flow values below 2, less than $\log_{10}2 \approx 0.3$, is not plotted because it is empty, as sites with area $A = 1$ receive only their single unit of injected flow, and are non-massive networks, that is, $M = 0$, (see Equation 1.16). The least massive sub-basins have $A = 2$.

2.3.2.1 The Least Squares Method

The least squares calculations used to determine the the scaling index α was computed using the procedure described by Gould and Tobochnik [1]. The error noted on the plots shown throughout the discussion of the models is the most probable error in the slope. The slopes generated by this method showed, in about 10 tests, to match that generated by the software package Matlab, in its linear plot fitting utility.

CHAPTER 3 RESULTS

3.1 Network Structures

Figures 3.1 - 3.9 show the structures of the initial Eden model and the resulting optimized networks. The channels are the collection of links along flow paths. Link width is proportional to the log of the area of the basin whose outflow the link represents. In all of the depictions of the full model output, sites with areas below a threshold $A_x \leq 6$ are not plotted. For depicting the whole network, trial and error showed that excluding this range of small sites best reduces the clutter of smallest channels, producing white space borders that highlight the various subbasins on a reasonably fine scale. Each picture of the full model is followed by a detail from the lower left corner of the whole, with full resolution.

OCN structures show the branching patterns similar to the OCN models in the literature (Meakin, *et. al.* [7], Rodriguez-Iturbe, *et. al.* [6], Maritan, *et. al.* [4]). The branching pattern of the SON models in Figures 3.7 and 3.9, for which no examples were found in the literature, is typical for all the SON models created for this study. In the multi-outlet model almost all of the border sites become detached from other sites during optimization, and a single large site (large total contributing area) receives most of the flow, resulting in a structure similar to the single-outlet SON model (The plots of cumulative distributions presented below show the sparse distribution of large-area drainage basins in the multi-outlet SON model).

The SON models are also striking for their unnatural flow direction, that, unlike the OCN models (Figures 3.3 and 3.5) they are characterized by paths of flow that converge into a main channel in the middle of the basin area, so that even sources adjacent to the outlet flow away from the outlet, toward the middle of the drainage area. These circuitous paths are seen in all the examples of both single- and multi-outlet SON models. The large channels turn into themselves, in some cases with diagonal turns that suggest the hexagonal arrangements in the underlying triangular lattice. While such a solution is not explored here, the idea that the SON model's main channels may be dividing and spiraling into ever-shrinking miniatures of a single kind of polygon is reminiscent of the form of an exact fractal structure like the Peano basin (discussed in Rodriguez-Iturbe, *et. al.* [6], Section 2.4).

The emergence of the final network forms can be seen in Figures 3.11, 3.12, and 3.13. They reveal that for both OCN's and SON's, the initial Eden configuration is first transformed by simulated annealing to a highly dissipative system, distinguishable as having, qualitatively, an irregularly sinuous structure. With the steep decline in energy dissipation the main channels to be featured in the final form appear amid smaller channels yet retaining the sinuousness of the "hot," chaotic-looking earlier networks. In the SON case the sinuosity of smaller channels is retained to the final form, while the OCN models resolve, even to the smallest basins into straight channels.

3.2 Area to Mass Scaling

Figure 3.14 shows the indices α for the mass-to-area scaling index (see Equation 1.17), as measured using bins. Figure 3.15 shows the fit of the data without bins. In the binned plots the single outlet OCN matches the index given in Maritan, *et. al.* [4]. No measure for α is given for the multi-outlet OCN model in any of the literature

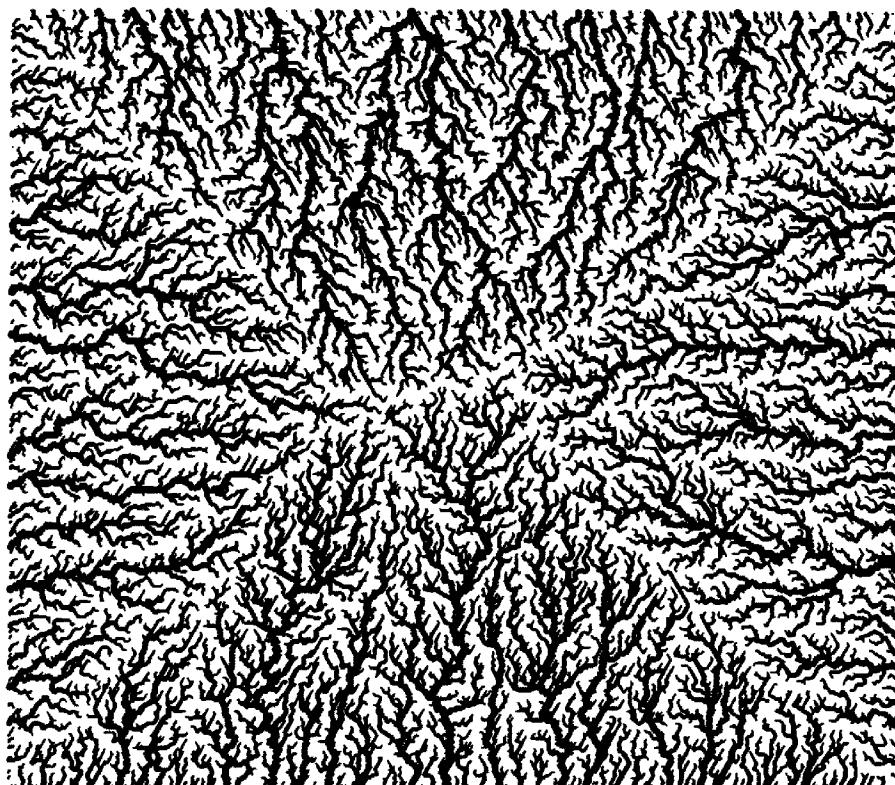


Figure 3.1 Eden Growth Model, the Random Initial Configuration

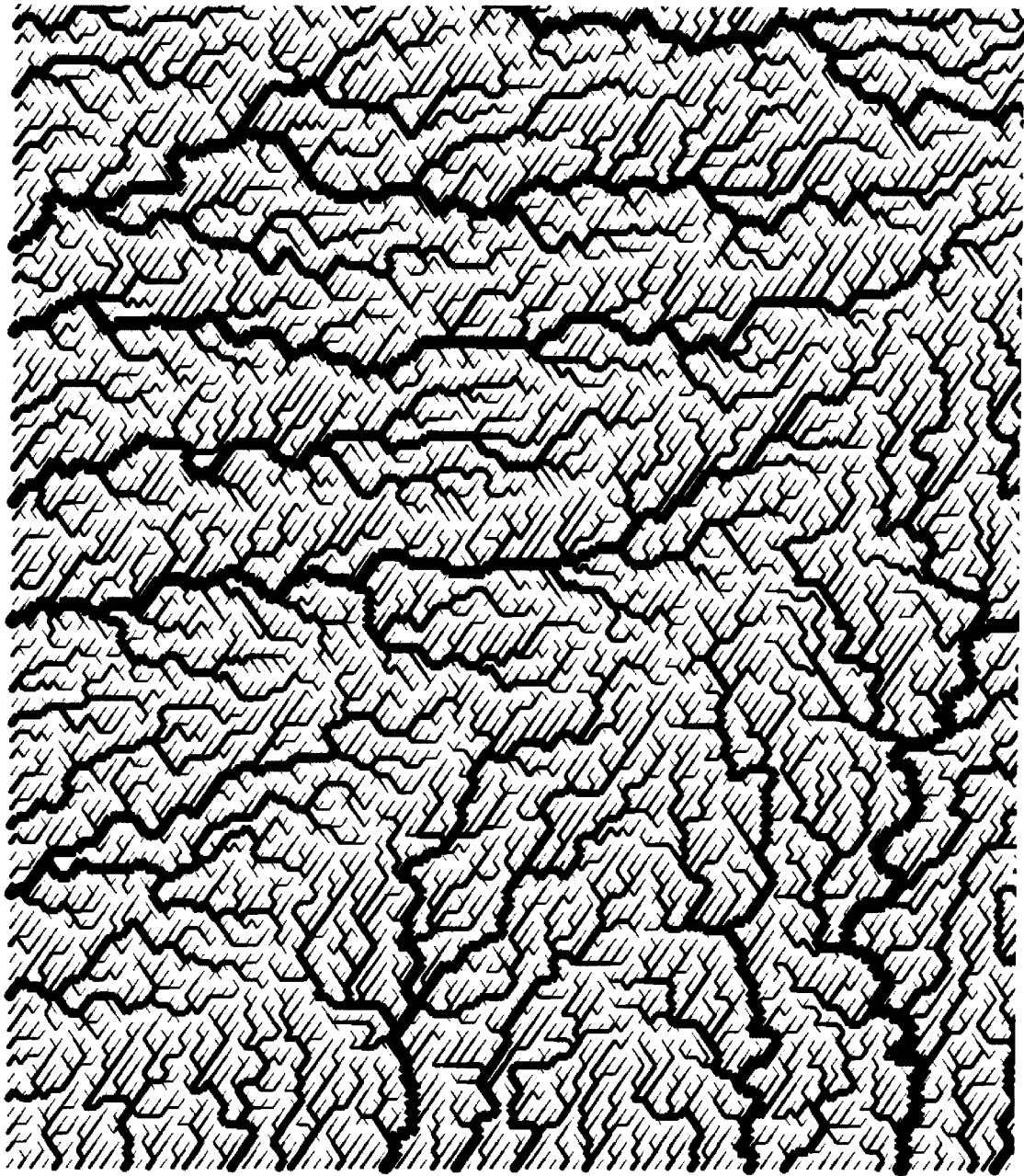


Figure 3.2 Detail, Lower Left Corner, Full Resolution, Eden Model

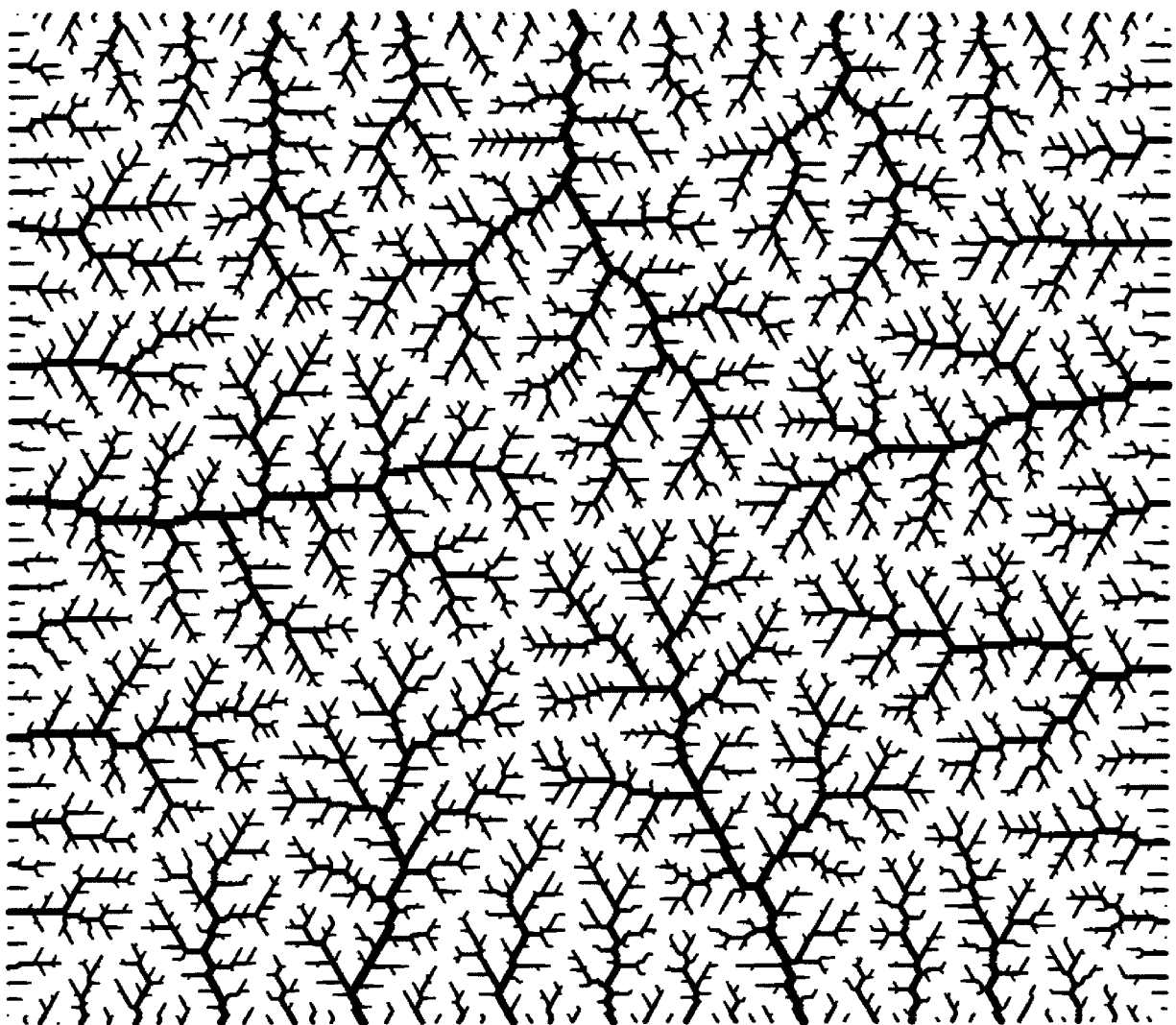


Figure 3.3 Multi-outlet OCN

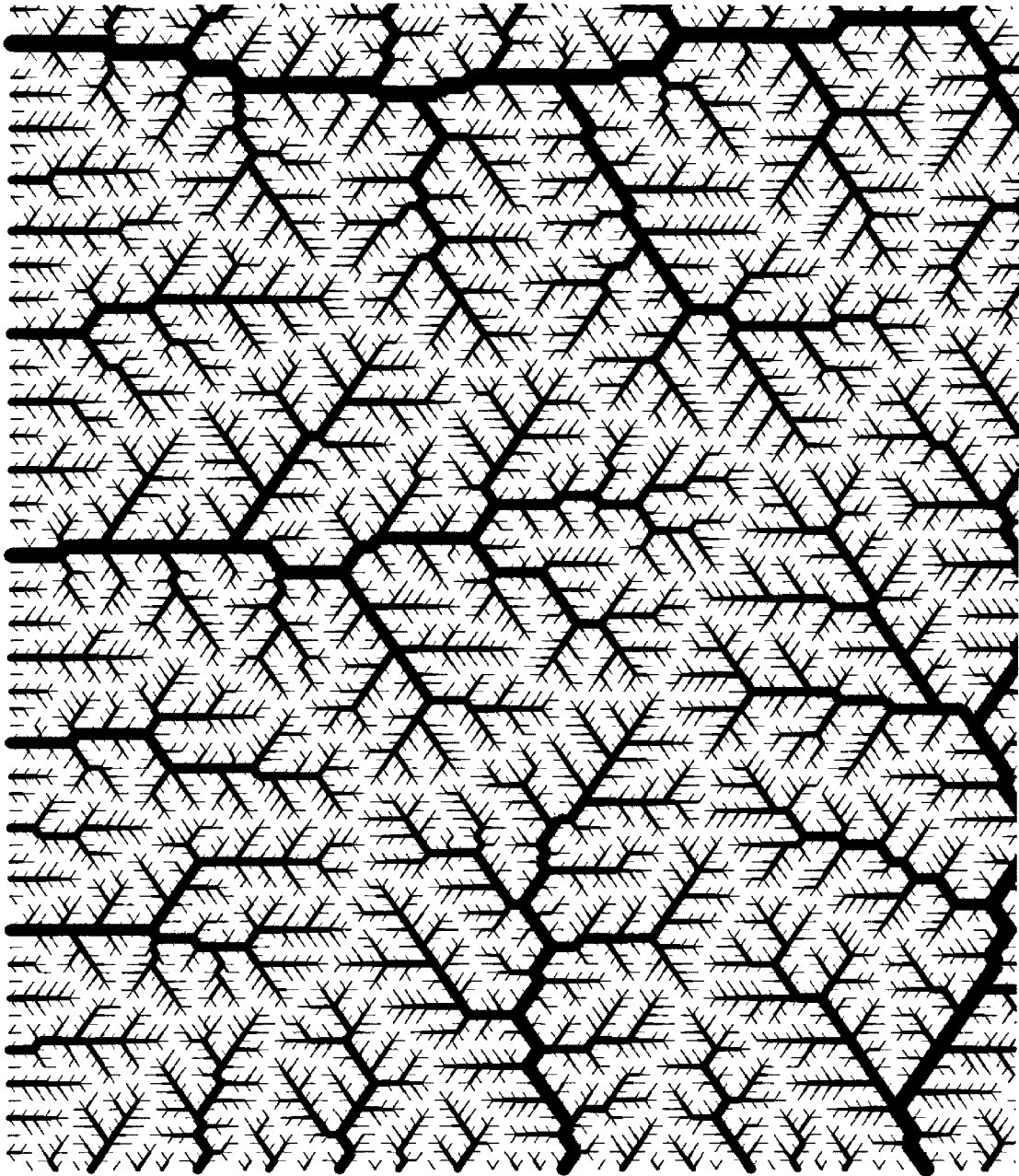


Figure 3.4 Detail, Lower Left Corner, Full Resolution, Multi-outlet OCN

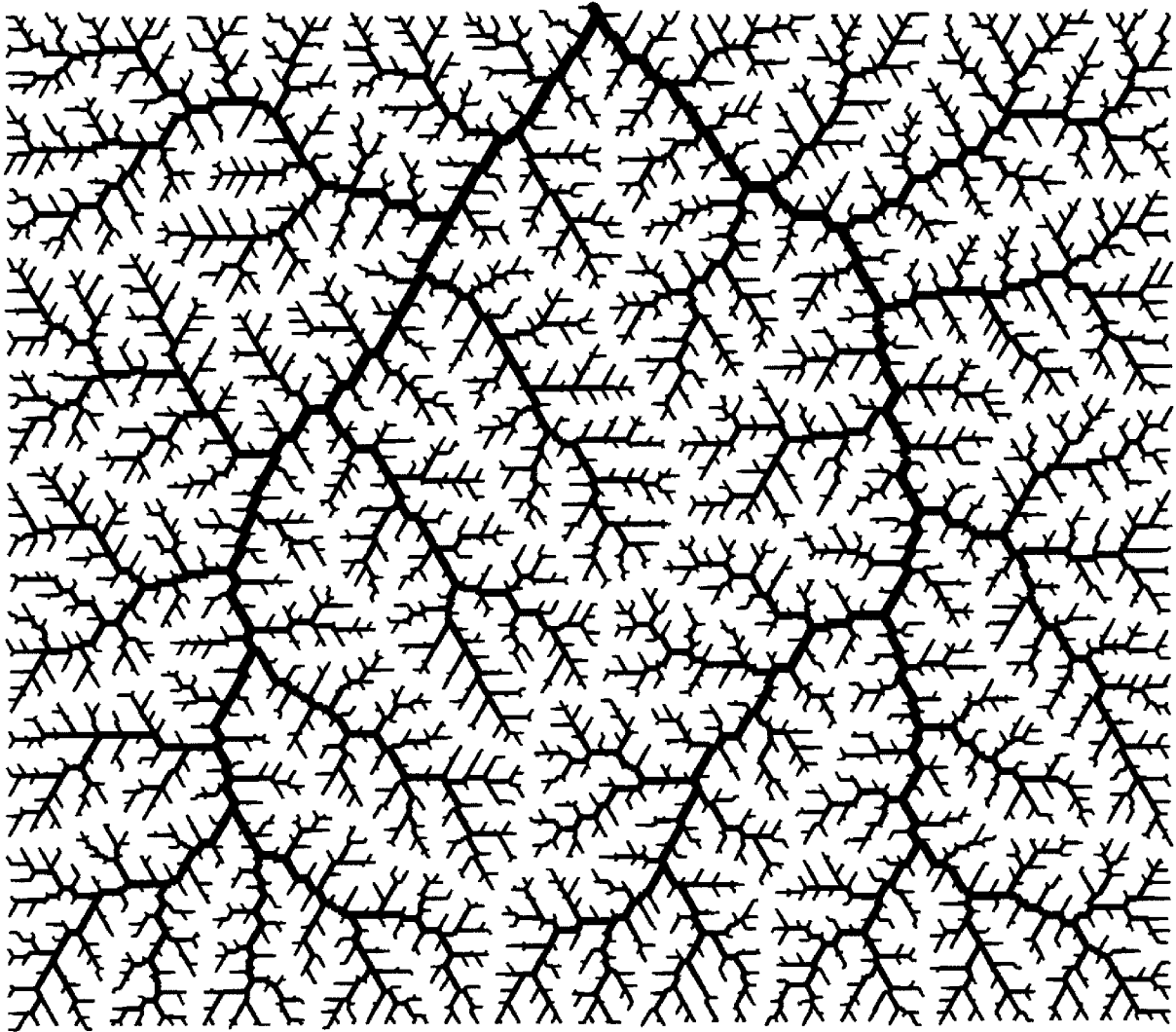


Figure 3.5 Single-outlet OCN

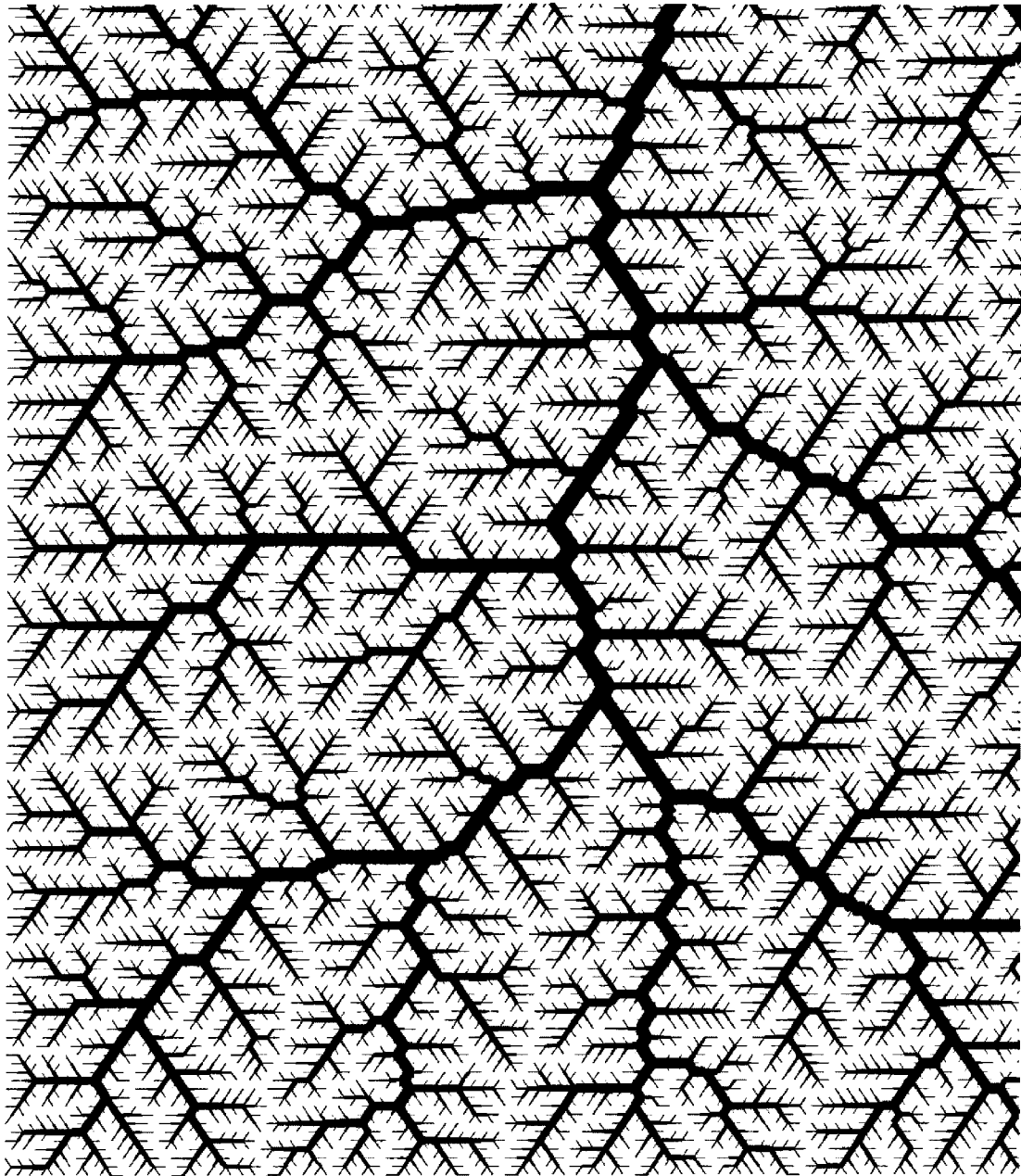


Figure 3.6 Detail, Lower Left Corner, Full Resolution, Single-outlet OCN

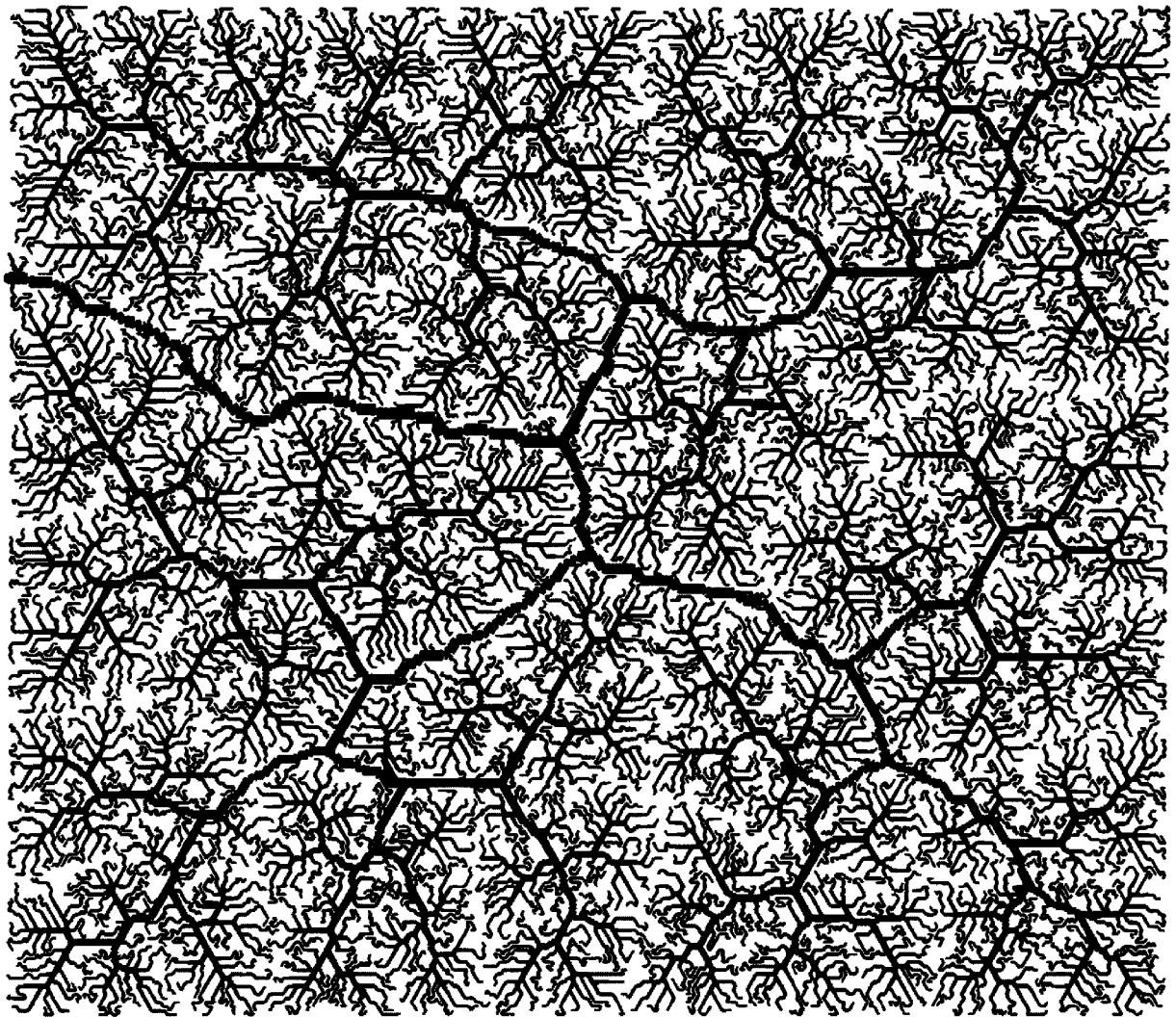


Figure 3.7 Multi-outlet SON

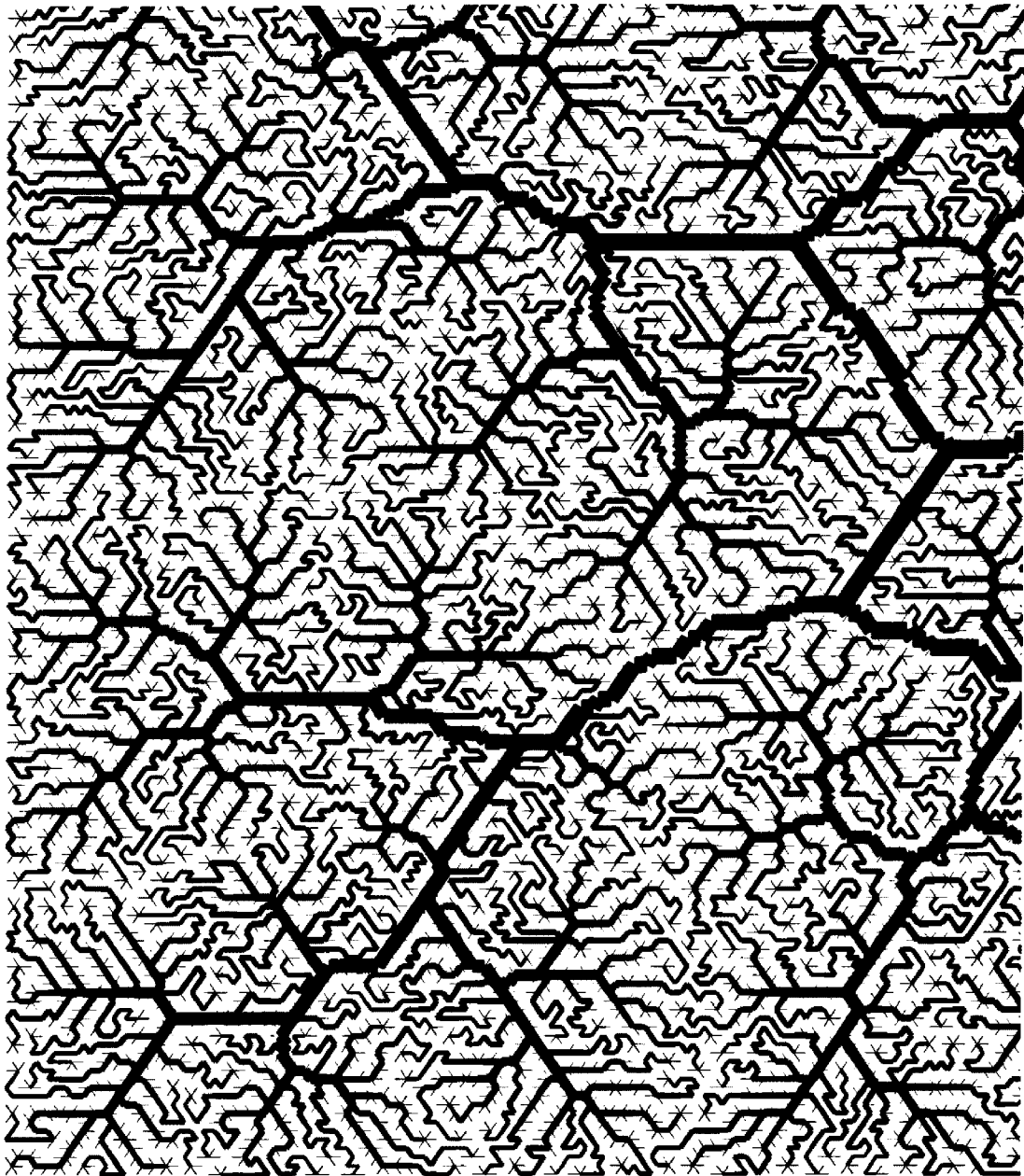


Figure 3.8 Detail, Lower Left Corner, Full Resolution, Multi-outlet SON

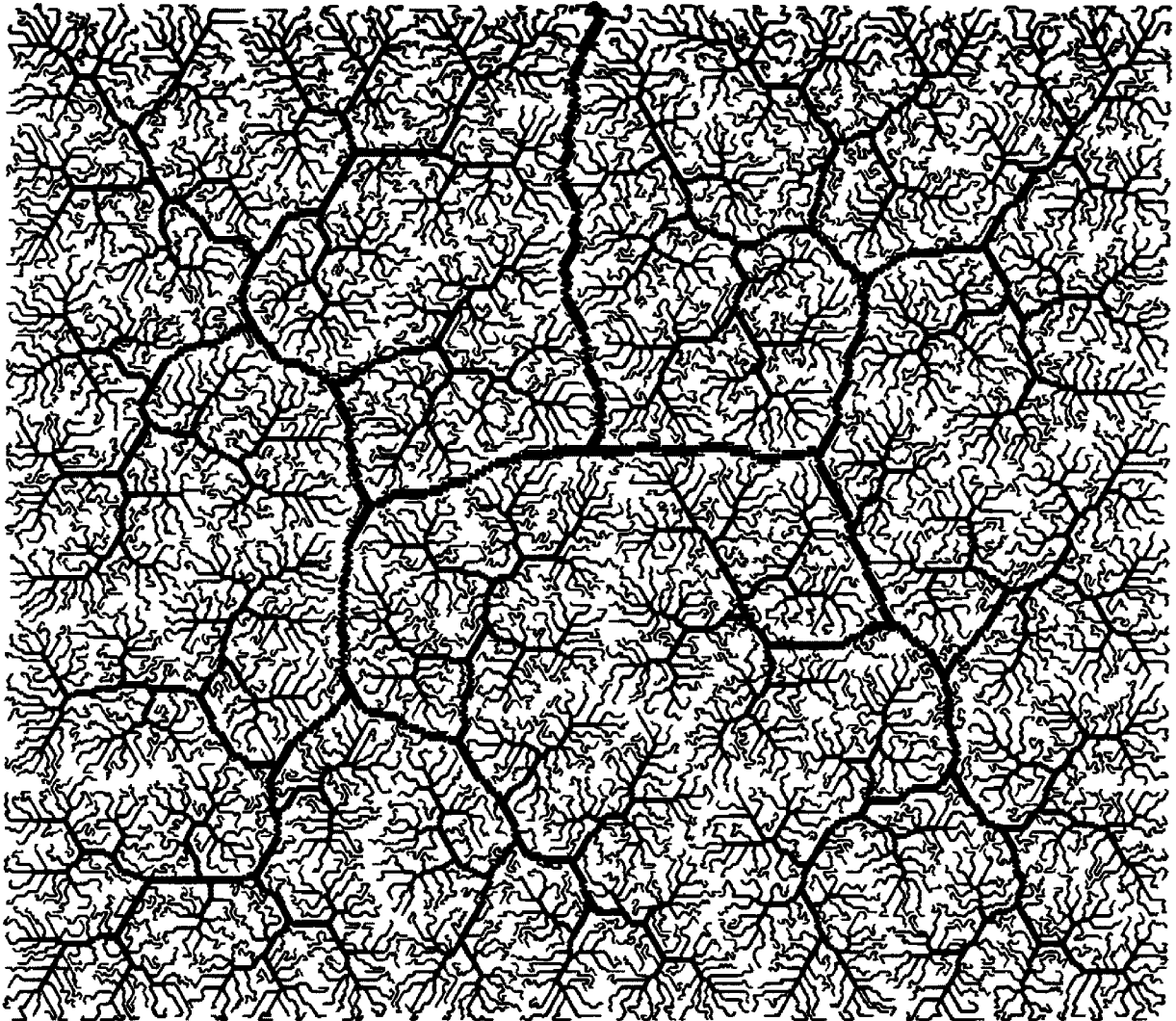


Figure 3.9 Single-outlet SON

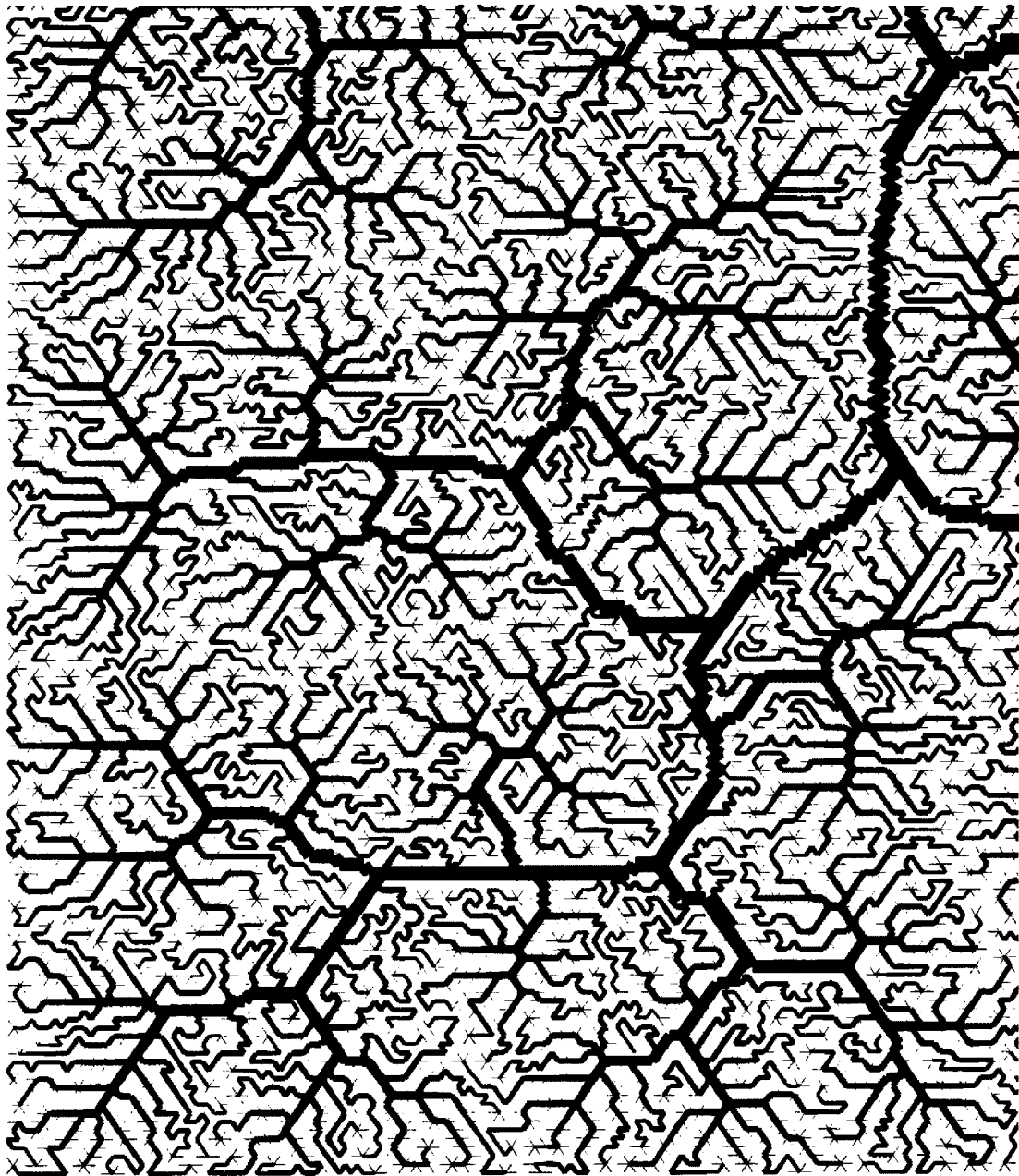


Figure 3.10 Detail, Lower Left Corner, Full Resolution, Single-outlet SON

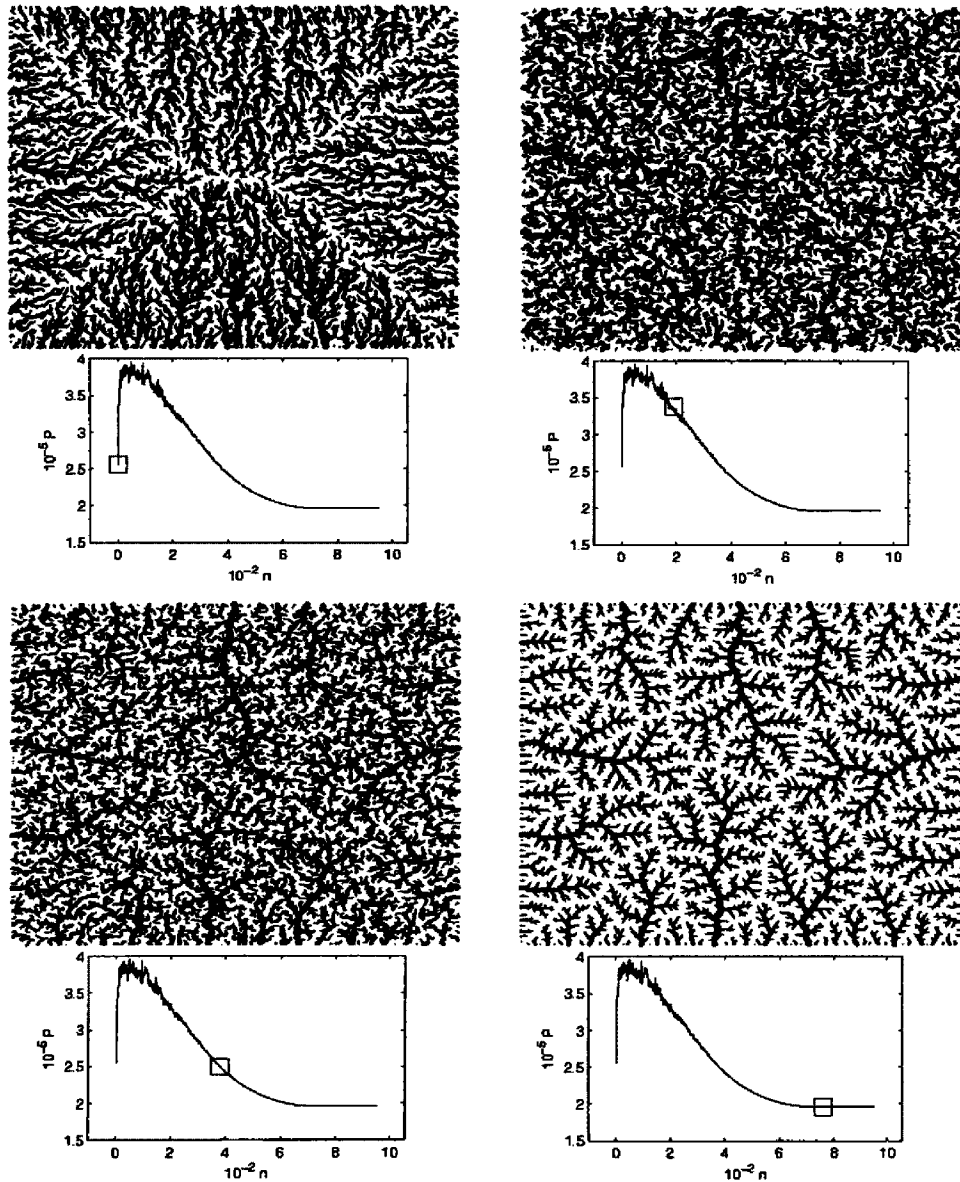


Figure 3.11 Networks Sampled During OCN Optimization

A multi-outlet area of basins, sampled during different stages of annealing for an OCN. Below each network is the plot of total energy dissipation versus annealing stages. The square in each plot indicates the annealing stage at which the corresponding network was the current configuration.

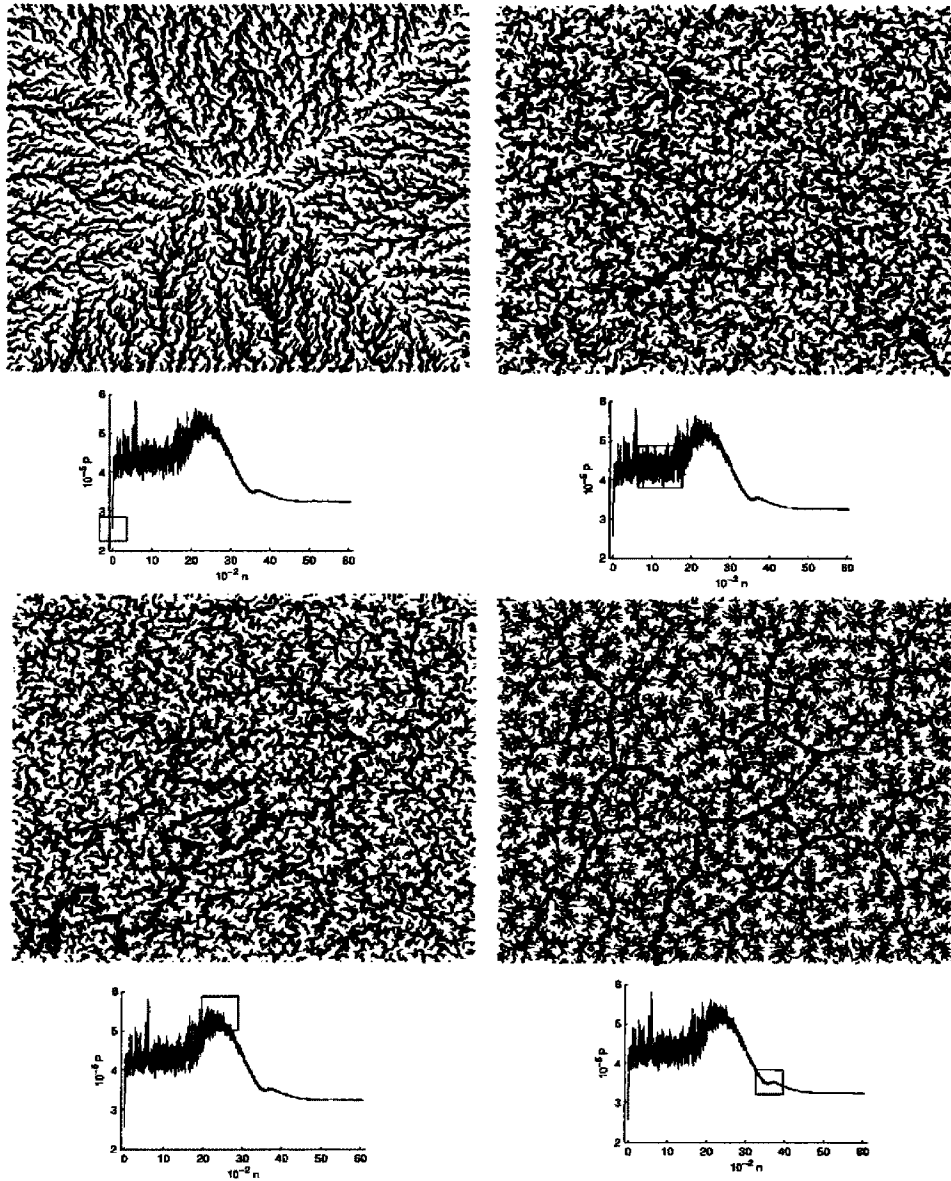


Figure 3.12 Networks Sampled During SON Optimization

A multi-outlet area of basins, sampled during different stages of annealing for an SON. Below each network is the plot of total energy dissipation versus annealing stages. The square in each plot indicates the annealing stage at which the corresponding network was the current configuration. The final structure is shown in Figure 3.13.

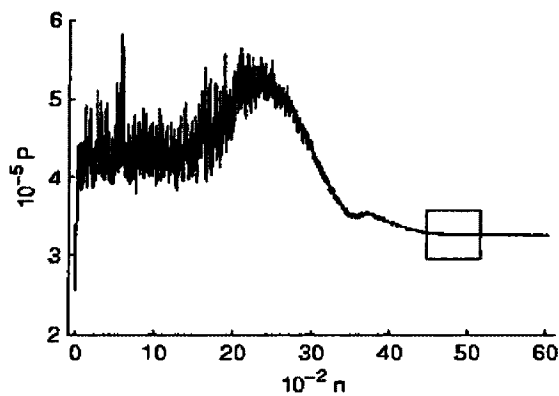
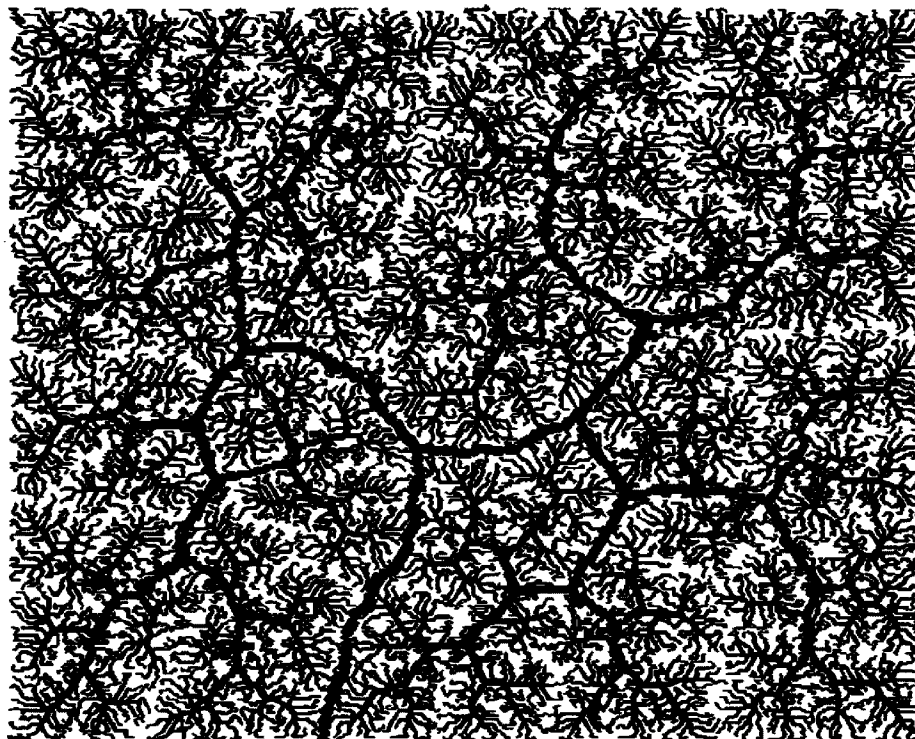


Figure 3.13 Networks Sampled During SON Optimization, Continued

The final network sampled during different stages of annealing for a multi-outlet SON model. (See the Series in 3.12).

reviewed for this study, so that its higher index cannot be compared to any other examples. The SON models, having been optimized to match the limiting value of $\alpha = 1.5$, closer to that value.

Unbinned, in the OCN case the scaling index is considerably higher than in the binned plots, indicating that the averaging process is attenuating some scatter within some ranges of basin-area size. In the case of the SON models, although the binned data shows a scaling index further from the target value of $\alpha = 1.5$, the paucity of data points mutes the sharply different scaling trend at the smaller scales. As the unbinned SON plots show, sub-basins with areas less than about $10^{1.5}$ are trending to a steeper slope, indicating a relatively high mass-to-area scale. Sinuosity in flow paths increases mass relative to flow¹ and so the tendril-like smaller channels (see Figures 3.7 and 3.9) scale heavier on the mass side of the scaling relation.

3.3 Cumulative Distributions

3.3.1 Total Contributing Area

The power law for the cumulative distribution of total contributing area², $P[A \geq a]$, that a randomly selected site has a total contributing area greater than size a , has been found to scale with a , in real river basins (Rodriguez-Iturbe, *et. al.* [6]) as,

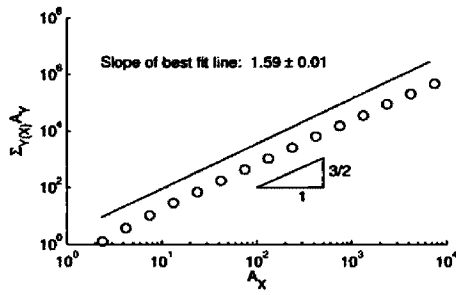
$$P[A \geq a] \propto a^{-\beta}, \beta = 0.43 \pm 0.02 \quad (3.1)$$

For OCN networks optimized by simulated annealing, Rodriguez-Iturbe, *et. al.* [6] found the index³ β to be nearer 0.50. The latter value is shown analytically in

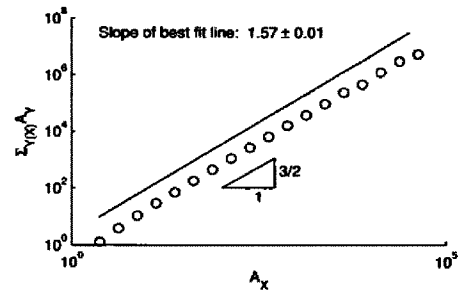
¹As discussed in Banavar, *et. al.* [2], the least efficient transportation network is the “space-filling spiral,” which is maximally indirect and is the most massive network.

²See Equation 1.15 for the recursive definition of total contributing area.

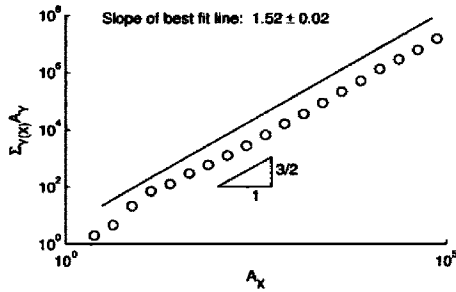
³Rodriguez-Iturbe, *et. al.* [6], Section 4.33. The index was measured by extracting the largest



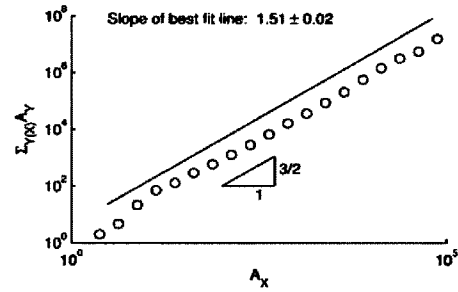
(a) Multi-outlet OCN



(b) Single-outlet OCN



(c) Multi-outlet SON



(d) Single-outlet SON

Figure 3.14 Network Mass versus Total Contributing Area.

The plot is shown for data binned as described in Section 2.3.2

Rodriguez-Iturbe, *et. al.* [6] Section 4.23, to be the “exact” solution for a networks with globally minimal total energy dissipation. Simulated annealing, then, insofar as it attains the global minimum total energy dissipation, should yield networks with $\beta \approx 0.5$.

A more natural $\beta \approx 0.43$ was achieved by Rodriguez-Iturbe, *et. al.* [6] with a hill climbing algorithm, a version of the simulated annealing in which the temperature is kept at zero, so that the only accepted changes in network configuration are those that

network in a multi-outlet basin with periodic boundary conditions, on the theory that, having developed within such an environment, the networks development would be free of the influence of boundary constraints.

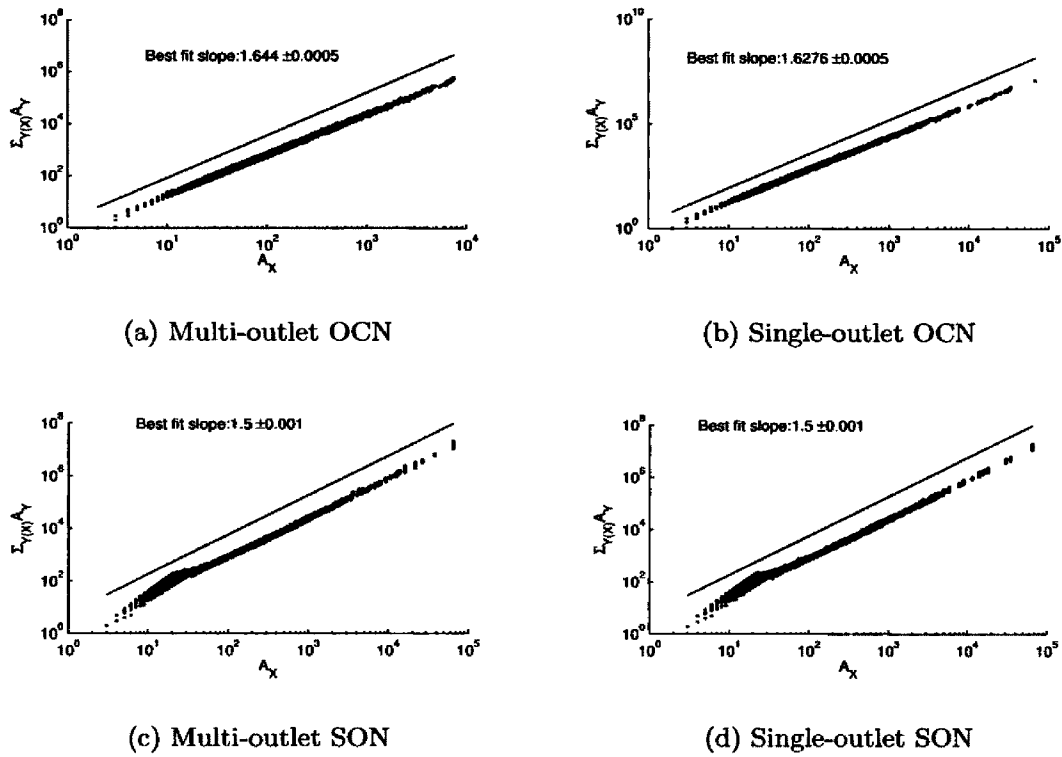


Figure 3.15 Network Mass versus Total Contributing Area, Unbinned Data.

lower the total energy dissipation. This procedure produces more dissipative networks (the network configuration converges to local optima, globally sub-optimal), but with several power law relationships, including this one, closer to natural river basins than are those optimized to a better minimum with a temperature reduction schedule. Figure 3.16 shows the cumulative area distributions for OCN and SON created with the annealing schedule as outlined in Section 2.2. The OCN examples as expected exceed the natural $\beta \approx 0.43$. The best fit portion of the data shows values for β larger than 0.5, unlike the OCN models in Rodriguez-Iturbe, *et. al.* [6] (Chapter 4), which do not exceed 0.50. The exceedence here may be caused by finite size effects in the range of data chosen for the best-fit line. The larger areas steepen the slope as linearity of the plot decays with increasing $\log_{10}a$, seen most dramatically in the

plot of the multi-outlet OCN. The effect becomes especially obvious in the tails of the plots, as the sites with large contributing areas are few in number and the cumulative distribution frequency falls dramatically for each successively larger value of a .

While showing similar slopes to the OCN models in their linear midsections, plots of the SON models show a non-linear head and tail, such that there is about $2\frac{1}{2}$ orders of magnitude of linearity, as opposed to 3 to 4 orders in the OCN examples. The SON examples also show a sporadic distribution of the largest sites, reflecting the consolidation of flow from most sites into few main channels, rather than in the more balanced distribution in the OCN structure.

3.3.2 Total Contributing Area of Drainage Basins

Meakin, *et. al.* [7] measure the cumulative distribution of the number of drainage basins with a given total contributing area. A drainage basin is a network that drains into one outlet in a multi-outlet model. As measured here for both OCN's and SON's in Figure 3.17, the scaling relationship is,

$$N(A > a) \propto a^{-\beta}, \quad (3.2)$$

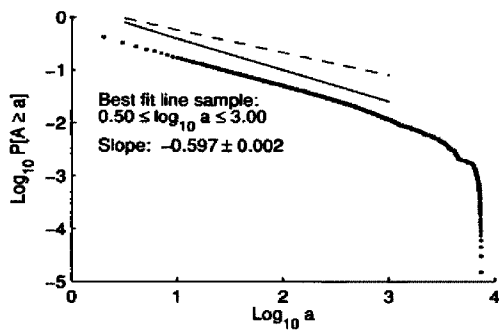
with $N(A > a)$ as the number of basins with area greater than a .

Meakin, *et. al.* [7] found a slope of $\beta \approx 0.51$, compared to the $\beta \approx 0.55$ shown here, and notes that this indicates a power law distribution of the number of basins of size A as,

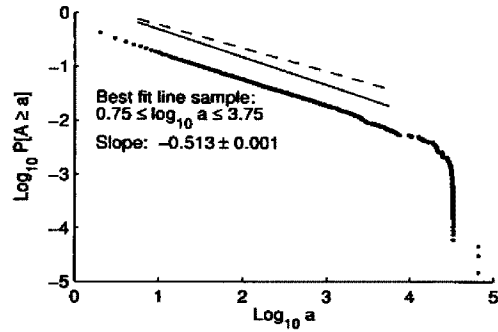
$$N(A) \sim A^{-\tau}, \quad (3.3)$$

with $\tau \approx 1.51$ for results in Meakin, *et. al.* [7], and in this study $\tau \approx 1.55$, this discrepancy possibly due to different border schemes (discussed in 1.5).

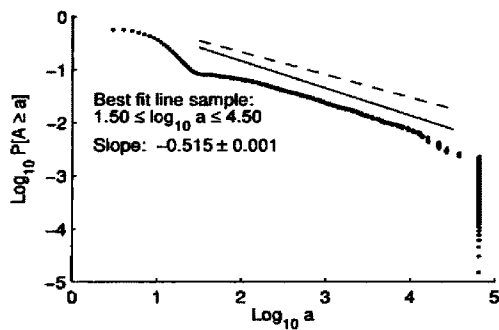
Although Meakin, *et. al.* [7] reports no field data measuring a collection of



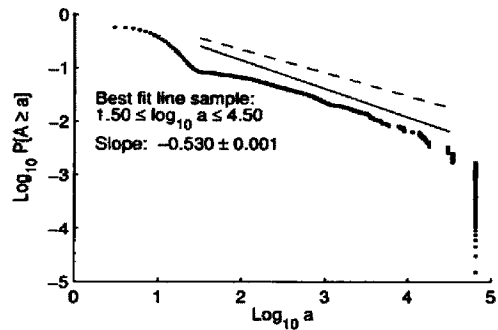
(a) Multi-outlet OCN



(b) Single-outlet OCN



(c) Multi-outlet SON



(d) Single-outlet SON

Figure 3.16 Cumulative Contributing Area

Log, log plot of the cumulative probability for contributing area versus contributing area size. The solid line has the slope of the least-squares best fit of $P[A \geq a]$ to a range of $\log_{10} a$ as indicated. The range was chosen by eye to find the most linear-looking, contiguous 3 orders of magnitude. The dashed line has slope -0.43 , the index found in real river basins (± 0.02).

drainage-basins over a contiguous area, the slope $\beta \approx 0.51$ is close to the 0.50 noted above in 3.3.1, the cumulative distribution of total contributing area found in OCN models optimized with simulated annealing to obtain a global minimum (compared to the 0.43 in natural basins, which are mimiced by OCN models optimized with hill climbing). It is interesting to note that, for the OCN models, the cumulative distribution of drainage areas scales similarly to the cumulative distribution of areas for a single network, while, in a single-outlet model, the finite grid-size effects on the sub-basins of size similar to drainage basins, do not scale in line (see Figure 3.16).

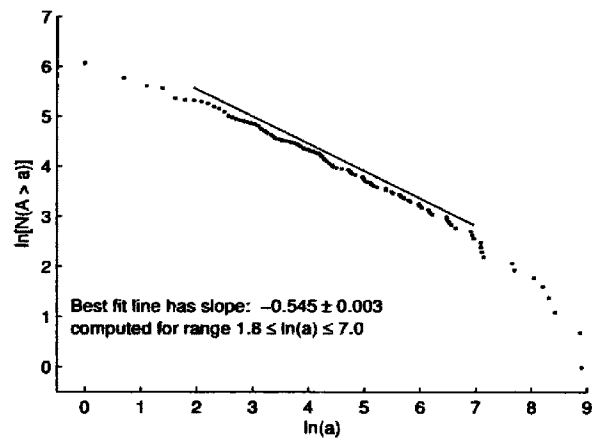
In stark contrast to the OCN models, the SON model shows that the scaling-index optimization has reduced border outlets to only 4 sizes, other than than the main outlet that receives the bulk of the flow. The SON imperative of minimum network mass on all scales would seem to undermine the OCN imperative of minimum energy dissipation in the multi-outlet context: "optimal drainage structure of an area tends to have a few large river networks together with a distribution of smaller networks covering the whole area" (Meakin, *et. al.* [7]).

3.3.3 Energy Dissipation per Unit Channel Link Length

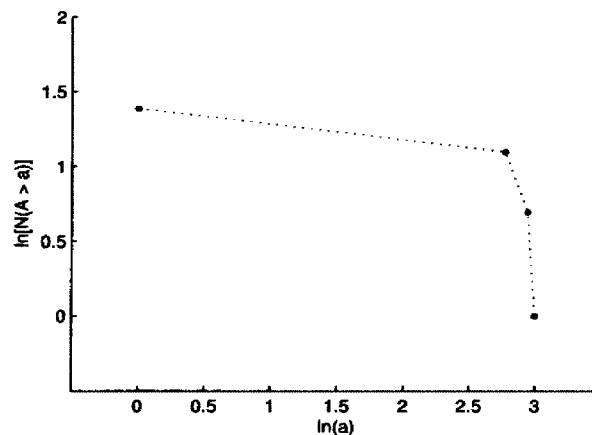
Meakin, *et. al.* [7] note that "uniform energy input to a dissipative system often results in a power law distribution of energy dissipation," and that in river basins the scaling index for the cumulative distribution of energy dissipation per unit channel length, p , is given by

$$\Pi(p > p^*) \sim p^{*-v} \quad (3.4)$$

where, in natural basins v is between 0.90 and 0.93. p in this case can be represented by the energy dissipation along a given link i (see equation 1.18), that is, the unit link of outflow from site i (recall that all sites in the model are connected by links of



(a) Multi-outlet OCN



(b) Multi-outlet SON

Figure 3.17 Cumulative Contributing Area of Drainage Basins

Log, log plot of the cumulative distribution of the size of drainage basins (sub-basins that empty into a sink on the border). The line has the slope of the least-squares best fit of $P[A \geq a]$ to a range of $\ln a$ values as indicated. The range was chosen by eye to fit the most linear-looking section. Given its paucity of site sizes and non-linear arrangement, the SON model's data was not fit with a line.

length 1). Unlike all other sites, basin outlet sites, since they have no flow, do not represent a unit channel link.

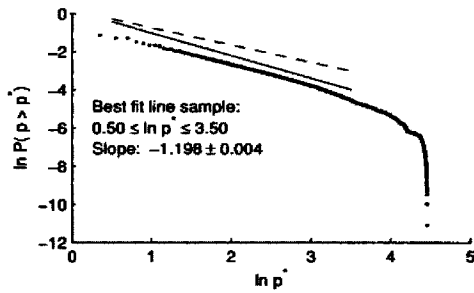
For this measure, in this study, all of the models exceed the expected values, with multi-outlet OCN showing $v \approx 1.2$, much larger than the expected and larger than its counterpart in Meakin, *et. al.* [7]. The reason for this large discrepancy is unknown, but may reflect a binning procedure used in Meakin, *et. al.* [7], though it is not described.

3.4 Total Energy Dissipation

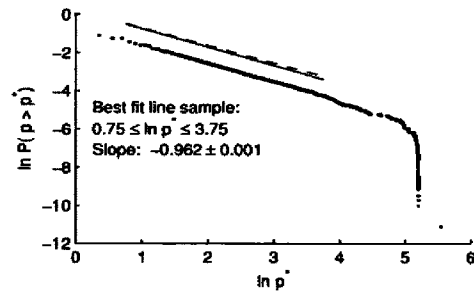
Figure 3.19 shows the total energy dissipation for each annealing stage. The multi-outlet OCN example replicates the overall character and final value for total energy dissipation of the plot shown in Meakin, *et. al.* [7], whose model is the basis for this example. One striking difference between this plot and that of Meakin, *et. al.* [7] is seen in the initial, large jump in energy dissipation from the value of the original configuration in this example (common to all the examples created for this study), while the plot in Meakin, *et. al.* [7] shows no such immediate rise. In Meakin, *et. al.* [7] the initial value is shown to be around 3.8 units, as opposed to the initial value here of about 2.5.

The final total is higher in the single-outlet OCN example, while its reduction relative to the initial configuration, about 7%, is close to the reduction in the multi-outlet example, about 8%. This suggests that for a given area, multiple basins can achieve more optimal energetic configurations than possible with a single outlet configuration.

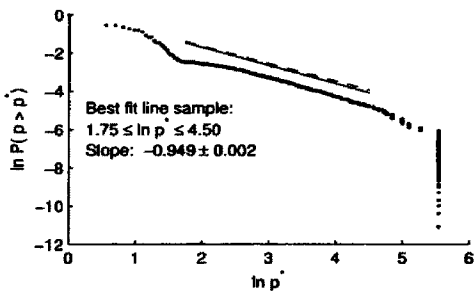
The SON models show a higher total energy dissipation than that in the OCN networks. Optimization on the scaling index plainly does not result in minimal (optimal) total energy dissipation. That both the multi-outlet and the single-outlet models



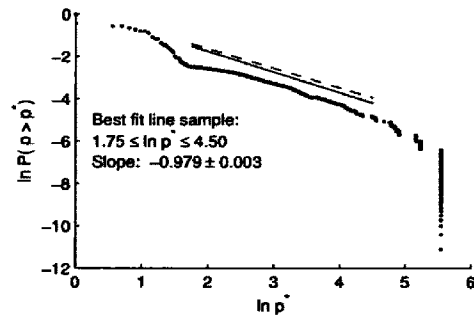
(a) Multi-outlet OCN



(b) Single-outlet OCN



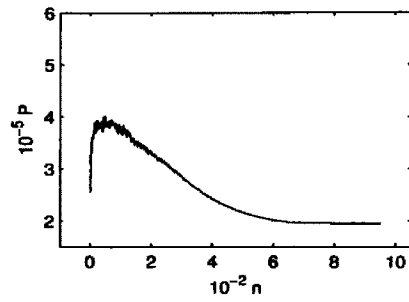
(c) Multi-outlet SON



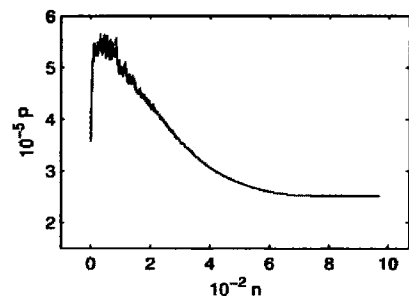
(d) Single-outlet SON

Figure 3.18 Cumulative Energy Dissipation per Unit Channel Link Length

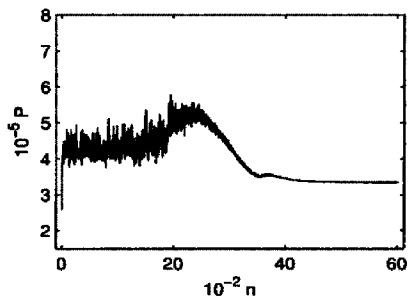
Log, log plot of the cumulative probability of energy dissipation per unit channel link length. The solid line has the slope of the best-fit line to a data-range of $\ln(a)$ as indicated. The range was chosen by eye to find the most linear-looking portion of the plot. The dashed line has slope -0.92 , the approximate index found in real river basins (Meakin, *et. al.* [7]).



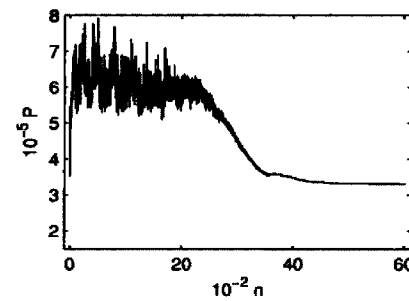
(a) Multi-outlet OCN



(b) Single-outlet OCN



(c) Multi-outlet SON



(d) Single-outlet SON

Figure 3.19 Total Energy Dissipation over Annealing Stages

have the same final total (within 1% of each other) reflects the transformation in all examples of the SON optimization, into networks with the same essential structure (see Figures 3.7 and 3.9). It also suggests that the optimized SON network will yield the same total dissipation from many different initial configurations.

3.5 Mean Distance of Site to Outlet

Figure 3.20 shows the mean distance of a site to the outlet of a sub-basin versus the size of the outlet. As argued in Maritan, *et. al.* [4], this relationship should be

$$\langle L \rangle \propto A_X^h \quad (3.5)$$

where $\langle L_X \rangle$ is the mean distance from site to the outlet of a sub-basin A_X is the total contributing area of the outlet, and h is Hack's exponent, that relates the length of the longest stream L_X in a sub-basin to the total contributing area.

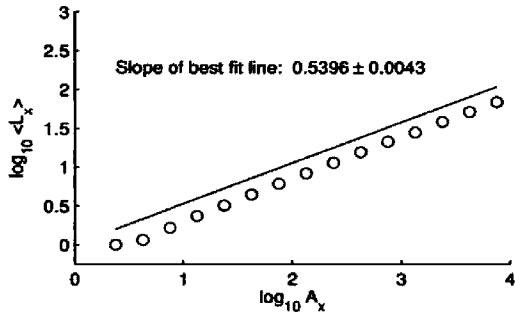
Maritan, *et. al.* [4] assume that $L_X \propto \langle L_x \rangle$, which leads⁴ to a relationship between the mass-to-area scaling index α and h , such that

$$\alpha = 1 + h. \quad (3.6)$$

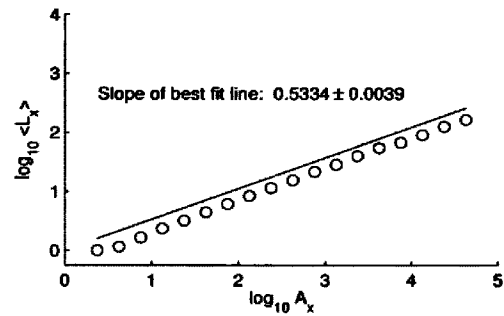
Whereas α and $1 + h$ match within 0.002 in single-outlet OCN models in Maritan, *et. al.* [4], models in this study show a larger discrepancy between mass-to-area scaling index α and $1 + h$, with h derived from Equation 3.5. In Maritan, *et. al.* [4] h is computed directly using the length L , the distance from the outlet to its most distant contributing source rather than the mean length, $\langle L \rangle$.

Table 3.5 shows a comparison of h as derived from the scaling index α and from the relationship of Equation 3.5.

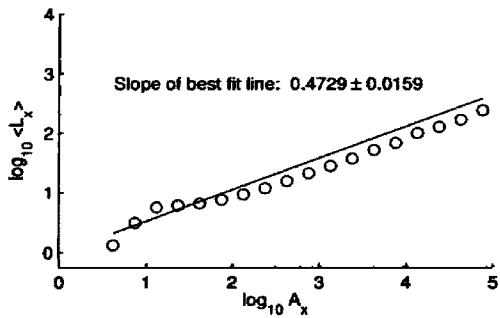
⁴The following proportionality, shown by Maritan, *et. al.* [4], relates Hack's exponent to the mass-to-area scaling index α through the mean distance value. Given a sub-basin with outlet at site X , with a constant injection into the system (for example constant mean annual precipitation at all sites), then network mass M_x , total contributing area A_X , and mean-distance-to-outlet $\langle L_X \rangle$ are related by $M_X \propto A_X \langle L_X \rangle$.



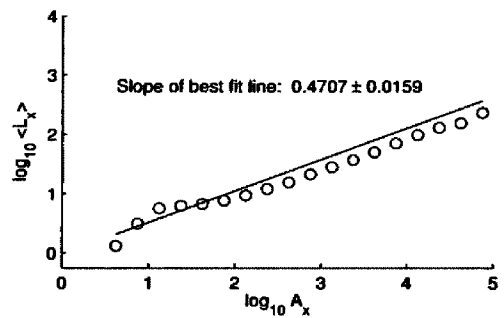
(a) Multi-outlet OCN



(b) Single-outlet OCN



(c) Multi-outlet SON



(d) Single-outlet SON

Figure 3.20 Mean Distance to Outlet

Log, log plot of the mean distance of sites in a sub-basin to the outlet versus the outlet size (total contributing area). The data has been binned as described for the mass-to-area scaling index (see Figure 3.14).

Model	$h = \alpha - 1$	Error	$h : \langle L \rangle \propto A^h$	Error
Multi-outlet OCN	0.59	0.01	0.54	0.00
Single-outlet OCN	0.57	0.01	0.54	0.00
Multi-outlet SON	0.52	0.02	0.47	0.02
Single-outlet SON	0.51	0.02	0.47	0.02

Table 3.1 Hack's Exponent from α versus Mean Distance

3.6 Summary of Power Laws

The table 3.6 summarizes the power-law scaling results discussed above, showing the indices for the models created here, as well as those for models published elsewhere, and the indices for real river basins. Table entries without error estimates reflect the lack of an error estimate in the literature providing the index.

	Mass-to-area Scaling	Cumulative Probability Distribution, Basin Area	Cumulative Probability Distribution, Energy Dissipation
Network	$M \propto A^\alpha$	$P[A \geq a]$	$\Pi[P \geq p^*]$
Multi-outlet OCN	1.59 ± 0.01	0.597 ± 0.002	1.198 ± 0.004
Single-outlet OCN	1.57 ± 0.01	0.513 ± 0.001	0.962 ± 0.001
Multi-outlet SON	1.52 ± 0.02	0.515 ± 0.001	0.949 ± 0.002
Single-outlet SON ¹	1.51 ± 0.02	0.530 ± 0.001	0.979 ± 0.003
Published, Multi-outlet OCN	NA	0.50	0.92
Published, Single-outlet OCN	1.57 ± 0.02	0.50	NA
Real River Basins	1.50 to 1.59 ± 0.02	0.43 ± 0.02	0.92

Table 3.2 Summary of the Scaling Indices of Power Laws

CHAPTER 4 DISCUSSION

4.1 Assumptions About the Models

If the results in Chapter 3 are to be seen as relevant to real river networks, these assumptions about the models are necessary:

1. Energy dissipation is correctly modeled by the Equation 1.18.
2. The scaling of network mass to flow rate, and hence the proximity to the most efficient class of networks as described in Banavar, *et. al.* [2], is correctly measured by the slope of the least-squares fit line to the log, log plot of the quantities in Equation 1.17, in its non-binned form. This is the measure used to calculate the value of the current configuration being optimized by simulated annealing to create a scaling-optimized network (SON).
3. For the respective target functions (total energy dissipation, scaling index), simulated annealing is finding a local optimum in the solution space that is close to the global optimum. In the case of multi-outlet OCN's the final total dissipation values match that in the literature (Meakin, *et. al.* [7]), while all other's have no confirmation from other investigations.
4. The Eden model provides a sufficiently random configuration that annealing can properly search the configuration space.

5. The random number generators used in the models produce statistically random series throughout a given simulated annealing session (see the Appendix for a discussion of random number generators).
6. The calculations of total contributing area and network mass are valid measurements of their counter parts in real river basins.
7. The hexagonal grid provides enough freedom of direction to emulate continuous space, for the purposes of accurately measuring fluvial geometric properties.
8. The size of the grid is large enough that its limits do not invalidate the data it provides, so that the single flow direction behaves like multi-directional flow.
9. The constant, uniform injection of material (the model of precipitation, as noted in 2.1.1), and constant velocity throughout the network correctly model real river basins (see Rodriguez-Iturbe, *et. al.* [6], Section 4.3).

4.2 The SON Model and the Theoretical Scaling Limit

This study's SON networks are optimized to scale as closely as possible to $\alpha = 1.50$ (Equation 1.17), the theoretical limit in efficient transportation networks. They differ both qualitatively and in scaling properties from the OCN models. Qualitatively, as noted in Section 3.1, they do not show the directed flow expected in the class of most efficient networks (c.f. Banavar, *et. al.* [2]). Compared to the OCN models, in the double log plots of cumulative distributions (Figures 3.16 and 3.16) the SON model's apparent self-similarity, as indicated by the most linear portion of the plotted data, is restricted to a comparatively narrow range of scales. The double \log_{10} plot of cumulative distribution of total contributing area, for example (Figure 3.16), shows this narrow range to be about $2\frac{1}{2}$ orders of magnitude versus 3 or more for the OCN

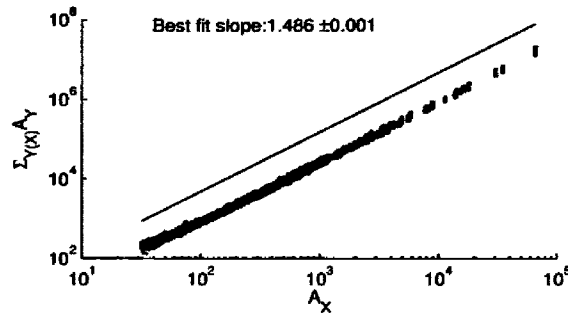


Figure 4.1 Scaling Index α for Large Basin Areas, SON

The log, log plot of network mass versus total contributing area in a single-outlet SON model. The data plotted is for sub-basins with area larger than $10^{1.5}$.

examples. This narrowed range is also seen in the log, log plots giving α (Figures 3.14 and 3.15), which show the same non-linear features over the same range of small-area basins. Even though α is closer to the ideal limit of 1.50 in the SON models, there is a distinctly steeper linear trend indicating a larger value for α on the small scale. Thus the remaining data, in compensation, has an index smaller than the ideal, as seen in Figure 4.1, with $\alpha \approx 0.48$.

Figure 4.2 isolates the structure of the small sub-basins, corresponding to the non-linear-trending small-scale portion of power-law plots, and the midsize sub-basins corresponding to the linearly trending portion of the power law plots of the single-outlet SON models. For comparison the sub-basins of corresponding size are shown for the single-outlet OCN. The largest sub-basins, corresponding to the tails of the power-law plots, clearly the effect of finite grid size in both OCN and SON networks, are not shown.

The small sub-basins in the OCN model are relatively straight channels. The small sub-basins in the SON model have the sinuous character of a whole network in a “hot” phase (characterized by high energy dissipation) of the simulated annealing process

(Figures 3.11 and 3.12). Besides high energy dissipation, sinuosity also indicates mass-heavy scaling ($\alpha \gg 1.50$), as indicated by the theoretical upper limit of $\alpha = 2.00$ seen in a single spiral flow path filling an entire area, and accounts for the steep linear trend of the small area data in the SON plot for α . During the optimization, grid size, or the method by which α (the slope of the least-squares fit line) is calculated, may be forcing the 2-phase scaling and structure. Whatever the cause, the consistent structure and measurements for all the SON models created for the study indicates that this particular optimization scheme might stubbornly disfavor self-similar scaling across all basin-sizes, while the OCN models show no such division in the plots of α . There may be a temperature reduction schedule, or change in the way the optimization procedure measures α , that would allow a “straightening” of the small-area basins, and a smoother allometric plot. Depending on the final overall structure, such a correction could also significantly lower the total energy dissipation, bringing the SON model closer to the minimal energy dissipation condition common to the OCN’s and real river basins.

For the (larger) portion of the SON plots whose trend attenuates α , the differences in network structure of the 2 models are less pronounced (Figure 4.2), though the SON model lacks the regular, elongated main channels fed by short sub-channels, seen in the OCN.

Overall, the SON model scales closer to the ideal $\alpha = 1.50$ than does the OCN model, while the index in real river basins shows more variation between individual networks. In a summary of scaling in 15 river networks, Maritan, *et. al.* [4] show α ranging from 1.50 to 1.59, with little “noise” in the plots of the individual basins. Plotting α using ensemble averaging of many river basins gives an α “statistically indistinguishable” from 1.50. Maritan, *et. al.* [4] conclude that there is a “robust central tendency” of allometric scaling in network structures toward the limit, while

individual variations, given that their scaling is not “noisy”, are “sensitive probes of the network structure”. The meaning of such variability may be informed by the SON model. It achieves a scaling index α close to 1.50, but shows a high (sub-optimal) energy dissipation, and an unnatural network structure. Though the discontinuity in the scaling trend of its power law measurements suggests the small-area basins could be made more energetically efficient in the SON model, the extent to which allometric scaling is achieved without minimal total energy dissipation suggests that the imperative of minimized network mass in allometric scaling is not a sufficient cause of natural transportation network structure.

While there is no certain formulation providing for the driving forces behind natural river basin structure, the OCN’s scaling similarities, paired with principles of least energy expenditure provides, as noted in Rodriguez-Iturbe, *et. al.* [6], a “comprehensive theoretical framework supporting the likelihood” of river basins as natural OCN’s. Though imperfectly realized by the SON model in this work, a thorough exploration of non-energetic optimization according to the theorem that establishes allometric scaling as an observed condition of river basin networks, could, depending on the relative energy dissipation of the resulting network, suggest a tie between the SON and OCN imperatives, with implications for the formation of other transportation networks, like animal circulatory systems, for which the minimization of transported material is *prima facie* advantageous.

4.3 Future Directions with New Models

When this thesis was presented orally, members of the thesis committee suggested that the discontinuity in the scaling trend of the SON model might be remedied by a target function in the simulated annealing process that minimized the sum over each

site X of the squared difference,

$$\Delta M = \sum_X \left(M_X - kA_X^{\frac{2}{3}} \right)^2, \quad (4.1)$$

where M_X is the mass (amount of transported material) associated with site X , and A_X the total contributing area of site X . k is the constant of proportionality that defines the mass-to-area power law (see Equation 1.17) as the equivalence:

$$M_X = kA_X^{\frac{3}{2}} \quad (4.2)$$

With k known the difference in Equation 4.1 is the difference between each site's mass and the most-efficient mass for the area given by that site in the most-efficient mass-to-area scaling relationship. If the sum of the square of the difference over all sites could be sufficiently minimized with simulated annealing the resulting network would closely approximate the allometric mass-to-area scaling, the same goal sought for the SON. It was hoped that this new way of calculating the relative optimality would eliminate the discontinuity discussed above in Section 4.2.

Because the value of k in Equation 4.2 is not known for the grid, it was suggested that linear regression could be used to obtain k for a given configuration. During simulated annealing, when the network is perturbed and a new configuration is to be tested for its optimality, an approximation of k could be obtained from finding the y-intercept b of the least squares linear fit of $\log_{10} M_X$, $\log_{10} A_X$, for all sites X . The surrogate k can then be computed as 10^b . During the early stages this value would reflect a log-space linear fit to data far from the desired values. However, the exact exponent, $\frac{3}{2}$, will have the greater influence on values of ΔM , so that reductions in ΔM might produce values for k closer and closer to the actual value of k for exact

allometric scaling on the grid.

The hope was that this target function, by avoiding the least-squares calculations, might avoid the discontinuity in the mass-to-area scaling between the large-area and small-area basins (discussed above in Section 4.2).

Time did not allow for the program changes needed to efficiently optimize a 256×256 grid. A smaller version on a 64×64 grid was created to see if the allometric scaling was smoother, and to have a look at the topology of the resulting network. Figure 4.3 shows the structure of a single-outlet model optimized to Equation 4.1. Figure 4.4 shows various metrics, including area-to-mass scaling.

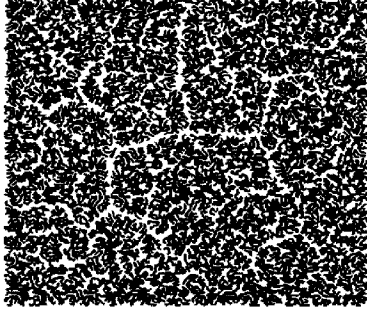
Immediately striking is the more direct flow paths than those exhibited in the SON structures (see Figures 3.7 and 3.9). Its mass-to-area scaling is farther from the theoretical most-efficient scaling index, $\alpha = 1.50$, than is that seen in the SON, in both raw and binned measurements. Compared to the OCN's, however, the raw data plot is closer to the target and for the binned plots about the same (see Figures 3.15 and 3.14).

The model also displays a discontinuity in the area-to-mass scaling similar to that in the SON model. Further, a linear data trend is missing in the log, log plots of the cumulative probability distributions for area and energy dissipation, indicating that the model shows less self-similarity than do the OCN and SON models (see Figures 3.16 and 3.18). This may be caused by a combination of the smaller grid size and the short time spent tuning the simulated annealing algorithm, resulting in a network reflecting a local optimum with a significantly higher ΔM (see Equation 4.1) than the global minimum.

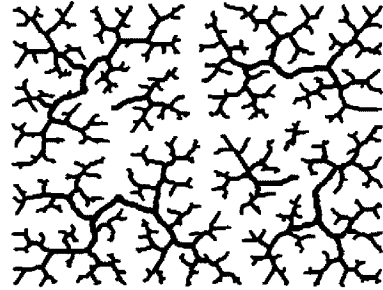
The suggestive finding is the combination of the directed, efficient-looking structure (also reflected in the mass-to-area scaling index matching that in the OCN) with higher-than-optimal energy dissipation (about 15,000 in this model, and 9,000 for

an OCN of the same size). This suggests that a efficient structure (none of the indirection of flow seen in the SON), with a condition of efficient mass-to-area scaling and inefficient energy expenditure. Arguing against drawing any conclusions are the discontinuity in the mass-to-area scaling and the lack of self-similarity in the other scaling measurements. These speak to the need for further refinement in the model.

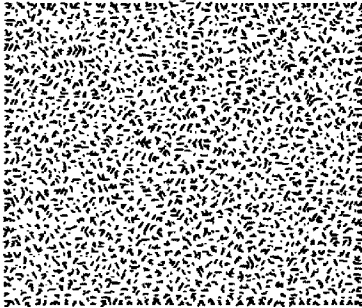
Suggestions for an improved model based on Equation 4.1 include the addition of binning during the annealing process. At each step the site to perturb is chosen randomly from within a bin, the bin itself chosen according to a random number weighted by its value for $\left(M_X - kA_X^{\frac{3}{2}}\right)^2$, where X is a bin, and M_X , A_X the average mass and area, respectively, of the sites within the bin. Such a scheme might allow a more uniform adjustment across all sites during simulated annealing, eliminating the discontinuity in the final model that inhibits smooth allometric mass-to-area scaling across all sub-basins.



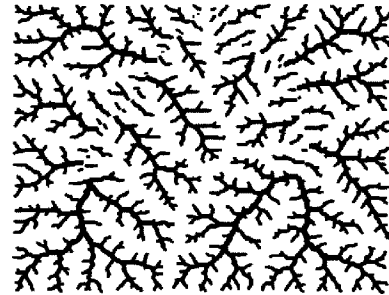
(a) Small Sub-basins, SON



(b) Midsize Sub-basins, SON



(c) Small Sub-basins, OCN



(d) Midsize Sub-basins OCN

Figure 4.2 Sub-basins Extracted by Size in Single-outlet Models

In these sub-basins of single-outlet models, small sub-basins X have total contributing areas $A_X < 10^{1.5}$. Midsize sub-basins X shown here have areas in the range $10^{1.5} \leq A_X \leq 10^4$. The midsize range corresponds to the most linear looking portions of the log, log plots of the cumulative distribution of total contributing area in SON models. The small size range corresponds to the SON's sinuous data trend over small scales in the log, log plots.

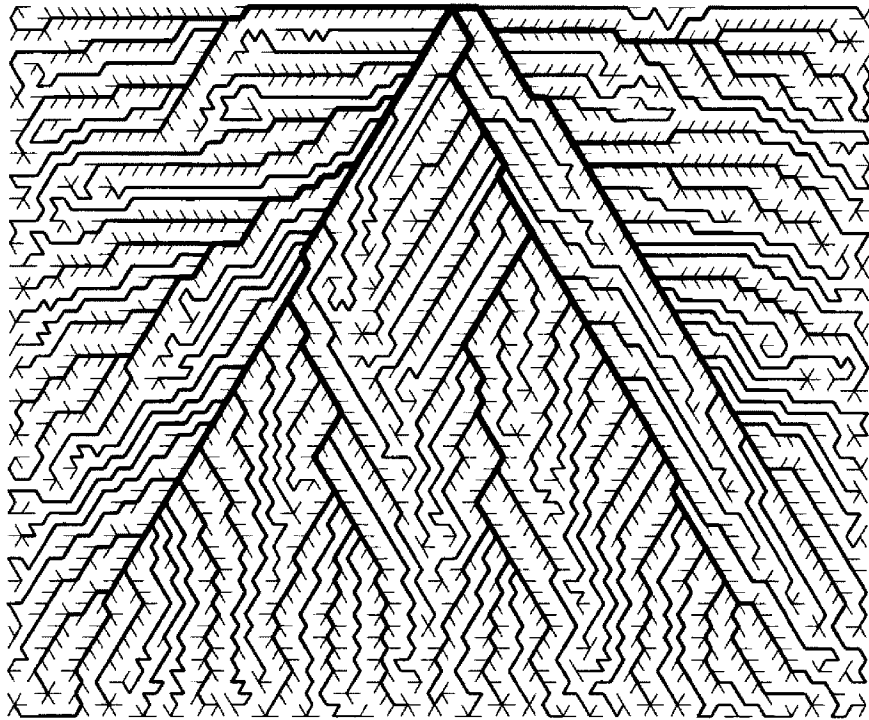
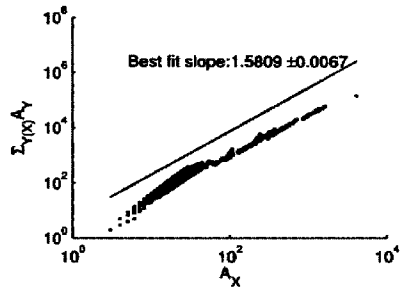
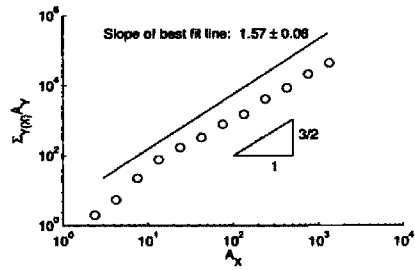


Figure 4.3 New Model Structure

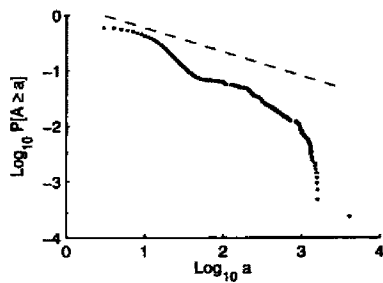
Structure of the Model Optimized to the Squared Difference (Equation 4.1)



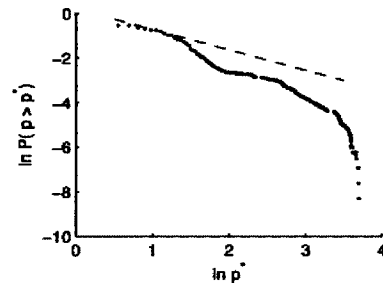
(a) Mass-to-area Scaling, Raw Data



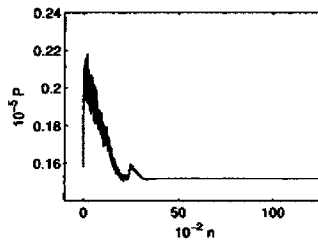
(b) Mass-to-area Scaling, Binned



(c) Cum. Probability Distribution, Area



(d) Cum. Probability Distribution, Energy Dissipation



(e) Energy Dissipation over Annealing Stages

Figure 4.4 Preliminary Results for Proposed New Model, Scaling Measurements. The dashed lines in the cumulative area and dissipation plots show the slopes that gives the scaling indicies for each distribution in real river basins.

APPENDIX A Table of Most-Used Variables

Variable	Definition
<i>L</i>	Linear size of a basin area
<i>D</i>	The number of dimensions of the network
<i>B</i>	Rate of transport, or, proportionately, the number of transport sites.
<i>A</i>	Total contributing area.
<i>M</i>	Network mass, or total material transported.
α	Scaling index in the allometric relationship, $M \propto A^\alpha$, with $A \sim B$

Table A.1 Most-used Variables

APPENDIX B Random Number Generation

The simulated annealing algorithm used to generate these models relied on a pseudo random number generator to provide uniform random deviates for the stepwise perturbation of the network configuration (see Section 2.2.2, item 3), as well as for evaluating the Boltzmann probability (Section 2.2.2, item 7). The number generator used here is that given by Press, *et. al.* [5], Section 7., called *ran2*. Initially the *ran1* was used, about which the authors advise “we do not know of any statistical test that *ran1* fails to pass, except when the number of calls starts to become on the order of the period m , say $> 10^8$ ”. It was discovered that the OCN models approach the period of *ran1*, while the SON models surpass the period. Tests were performed comparing results from the 2 generators.

5 program runs showed that the procedure that produces the OCN models makes, on average, 2.1696^8 calls¹ per run, on the upper end of the period, but still within the stated range for statistically effective random number generation. The SON models, however, often surpass the period of the generator: the best SON results² make calls numbering about 1.5×10^9 . New models, created with the *ran2* generator, with its period on the order of 10^{18} were created to compare with the former results, for both OCN’s and SON’s. Comparisons of 7 OCN models generated using *ran2* showed an average total energy dissipation of 195,620, and a standard deviation of 231.86,

¹The standard deviation of this 5-run sample is about 1.6×10^4 .

²Models with a scaling index, in raw data terms, of 1.50 and an error less than 10^{-2}

compared to an average of 195,718 for 25 runs with *ran1*, and a standard deviation of 307. 3 SON models (2 multi-outlet, 1 single-outlet) were run with *ran2* and compared with 3 models of the same type run with *ran1*. All 6 models had a scaling index of $1.500 \pm .001$. The *ran2* models had an average total energy dissipation of 332,400, with a standard deviation of 1433. The *ran1* models had an average total energy dissipation of 337,700, with a standard deviation of $7.0123e+03$. Qualitatively all the networks looked similar by type.

One possible effect of the misuse of *ran1* can be seen in the total energy dissipation of the SON models, which were about 1.5 % higher in the (small) sample tested, and had a much higher variability between models, when *ran1* was used. While the small samples do not allow a definitive connection with the random number generator, the variability in energy dissipation between the *ran1* and *ran2* SON models, versus the uniform slope index reached in all the models testifies to the variety of dissipative states that provide the “correct” scaling relationship, a finding that should be seen in the light of the model’s constraints: finite size, the square grid boundary, the triangular lattice, and the initial eden growth model from which the result is derived.

All models pictured and measured in this work, when not otherwise stated, were created with the *ran2* generator.

APPENDIX C A Note on Performance with Different C Compilers

The simulated annealing program was written in the C programming language. One version of the program makes many calls to the *pow* function, while another uses the *sqrt* instead, both functions declared in the standard C library's *math.h* file. The *gcc* version 3.4.5 proved to have a slow implementation of the *pow* functions, compared to the Intel *icc* compiler version 9.0.

A program run that took 1700 seconds with a *gcc* compiled version took about 87 seconds on the Intel compiler. The running-time profiler *gprof* showed that the bottleneck was the loop shown below. The function that contains the loop takes up over 89% of total running time, as sampled by the profiler (the test ran for 1 annealing stage of 2,562 steps). These are the offending lines, as shown by the profiler's breakdown:

```
for( i = 0; i < i_n_squared; i ++ )
{
    d_returned_diss = d_returned_diss
        + pow( (double) pi_flow[ i ], DISS_EXP);
} /*end "for each site"*/
```

The *pow* function seems to contribute significantly to this difference. A simple program that computes the *pow* function 600×10^6 times, in six loops of 100×10^6

iterations, took less than a millisecond using *icc*, but 111.930 seconds with *gcc*. When the *pow* calls were replaced by *sqrt*, *gcc* took 9.717 seconds, and *icc* less than a millisecond. When the body of the loop contained no line of code at all the *gcc* compiled program ran in about half a second (as expected, the *icc* program in less than a millisecond), suggesting that the compilation of loop-intensive code is done better by *icc*. All the executables were compiled with the optimization given by the *-O2* flag.

BIBLIOGRAPHY

- [1] *An Introduction to Computer Simulation Methods*. Addison-Wesley Publishing Company, 1996.
- [2] J.R. Banavar, A. Maritan, and A. Rinaldo. Size and form in efficient transportation networks. *Nature (London)*, 398:130–132, 1999.
- [3] Olaf Dreyer. Allometric scaling and central source systems. *Physical Review Letters*, 87(3), 2001.
- [4] A. Maritan, R. Rigon, J. R. Banabar, and A. Rinaldo. Network allometry. *Geophysical Research Letters*, 29(11), 2002.
- [5] William H. Press, Saul A. Teukolsky, William T. Vetterling, and Brian P. Flannery. *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge University Press, 1992.
- [6] I. Rodriguez-Iturbe and A. Rinaldo. *Fractal River Basins: Chance and Self-Organization*. Cambridge University Press, 1997.
- [7] T. Sun, P. Meakin, and T. Jossang. Minimum energy dissipation model for river basin geometry. *Phys. Rev.*, 49(6):4865, 1994.