A PHYSICALLY BASED CHANNEL NETWORK AND CATCHMENT EVOLUTION MODEL

by

GARRY RAYMOND WILLGOOSE

Bachelor of Engineering. (Hons 1), Civil Engineering
University of Newcastle, Australia.
(1981)

Bachelor of Science, Mathematics
University of Newcastle, Australia
(1981)

Master of Science, Water Resources
Massachusetts Institute of Technology
(1987)

Submitted to the Department of Civil Engineering
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy
in Civil Engineering

at the

Massachusetts Institute of Technology
April, 1989

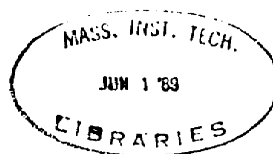© Massachusetts Institute of Technology.

Signature of Author

_____
Department of Civil Engineering
April 10, 1989

Certified by_____

_____
Rafael L. Bras
Thesis Supervisor

Accepted by_____   _____
Ole S. Madsen
Chairman, Departmental Committee on Graduate Students

Vol. 1

A PHYSICALLY BASED CHANNEL NETWORK AND CATCHMENT EVOLUTION MODEL

by

Garry Raymond Willgoose

Submitted to Department of Civil Engineering
on April 10, 1989 in partial fulfillment of the
requirements for the degree of Doctor of Philosophy in
Civil Engineering

ABSTRACT

A catchment evolution and channel network growth model is presented. Elevations within the catchment are simulated by a sediment transport continuity equation applied over geologic time. Sediment transport may by modelled by both fluvial (e.g. Einstein–Brown) and mass movement (e.g. creep and landsliding) mechanisms. An explicit differentiation between the channel and the hillslope is made with different transport processes in each regime. The growth of the channel network is governed by a physically based threshold, which is nonlinearly dependent on discharge and slope and thus governed by hillslope form. Hillslopes and the growing channel network interact through the different sediment transport processes and the preferred drainage to the channels to produce the long term form of the catchment.

General requirements for network formation in physically based models are examined by use of a previously reported leaf vein growth model. Elements of chaos were discovered that result in apparently random networks being generated. It was argued that this is also true for the catchment and connections with topologically random networks were provided.

Synthetic catchments were simulated using a numerical implementation of the model and statistics for the catchments are analyzed. Drainage density and elevation characteristics are correlated with nondimensional numbers arising from a nondimensionalization of the governing equations. These nondimensional numbers parameterize rates of tectonic uplift, sediment transport, both in the channel and the hillslope, channel growth, and resistance to channelization. Runoff rate, erodability and flood frequencies arise explicitly in these numbers. A fundamental measure of catchment dissection based on one of the nondimensional numbers is proposed. It follows that drainage density and hillslope length are dependent, in a well defined way, on runoff rate, slopes and catchment erodability.

Simulation results are compared with reported field data and small scale experimental catchment evolution studies and found to be consistent. A linear log–log relationship between channel slope and area, observed in the field, is also observed in the simulation data at dynamic equilibrium. An explanation based on model physics is proposed, a central feature being the balance between tectonic uplift and fluvial erosion at dynamic equilibrium. This explanation also accounts for observed deviations from the linear log–log relationship where slopes are reduced for small areas; these small areas are dominated by diffusive transport processes in the hillslope. The channelization threshold based on discharge and slope is compared with recently reported data of hillslope slopes and contributing areas at

channel heads; the threshold is consistent with the field data.

Observed differences between hypsometric curves, previously attributed to catchment age, are found to result from differences in the tectonic uplift regime. A scheme for landscape classification, based on the nondimensional numbers, is proposed which is more consistent with the governing physical processes than previous work. A one—dimensional advection—diffusion reformulation of the sediment transport equation is proposed that predicts rates of hillslope retreat and hillslope degradation, and provides a link to observed hillslope transport mechanisms.

Thesis Supervisor: Dr. Rafael L. Bras

Title: Professor of Civil Engineering

TABLE OF CONTENTS

4

LIST OF TABLES

LIST OF FIGURES

13

ACKNOWLEDGMENTS

"When we try and pick out anything by itself, we find it hitched to everything else in the Universe"

John Muir

I can't hope to thank all the people that have helped in the course of this research, there have been too many, I can only hope that people omitted here understand.

My thesis advisor, Prof. Rafael L. Bras, is thanked for his guidance and for providing unswerving support for the topic and approach of this research, even in the face of apparently insurmountable skepticism. Prof. Ignacio Rodriguez–Iturbe is thanked for his inexhaustible enthusiasm and exceptional generosity intellectually and for originally suggesting that I look at the work of Meinhardt. Without both of them, and the intellectual environment they provided at Parsons Lab, this work would never have come to fruition. The other members of my doctoral committee. Profs. Peter S. Eagleson and Michael A. Celia have also, through their thoughtful comments, criticisms, and encouragement, contributed substantially to the form of this work.

The Ippen Fellowship Fund provided opportunities to travel to AGU meetings; these meeting have been a highlight of my stay at MIT. In particular, it is there that I met Prof. Bill Dietrich, Berkeley, who, though initially doubtful about the approach of my work, has been consistently encouraging and has pointed out geomorphology literature of which I was unaware.

At the Lab, fellow students, past and present, partook in many interesting discussions (or arguments depending on viewpoint) without which life at the Lab would have been much less enjoyable. Thanks to the Americans who didn't take offense at my political views. In my research group, David Tarboton and Laurens van der Tak, among others, are thanked for their helpful comments and criticisms. Don Polmann and I shared many treasonable thoughts about computers. Pat Dixon, administrative assistant, was always there to talk to, and help if necessary, as were Elaine Healy and Carole Solomon (thanks for the loan of the short–wave). Elaine deserves special thanks for typing the drafts of this thesis, while still maintaining a sense of humor (and reminding me I should keep mine).

At home, the residents of Conner 2 were always there for cards, Star Trek or commiseration about the latest Celtics defeat. Their appetite for chocolate chip cookies has to be seen to be believed! It was fun being Graduate Resident and I'll miss you all. Veronica, my wife, is thanked for taking time out of her career to accompany me to the United States.

CHAPTER 1

INTRODUCTION

1.1 The Scope of this Work

For many years hydrologists and geomorphologists have been fascinated by questions about catchment form. Why is it that channels form branched networks reminiscent of blood veins, leaf veins, plant roots, and branches? Why is it that the hillslopes adopt the form they do, draining directly to the channels with characteristic patterns of runoff distribution? How do those properties interact with the hydrologic response of the catchment? It is to these types of questions that this work is addressed.

That networks occur in many different physical settings suggests that the processes that create them should bear qualitative similarities that go beyond the details of the specific physics. One of the first tasks of this work is to understand the qualitative processes that govern the development of networks from initially uniform conditions.

From an understanding of the qualitative processes the goal is then to develop a quantitative understanding of how channel networks and hillslopes evolve with time. Through a quantitative understanding of the evolution process general statements may be made about the catchment form and hydrologic response at any point in time. Catchment form is then seen in the context of the complete history of erosional development of the catchment leading up to the present time.

The quantitative understanding of catchment evolution processes and their effect on catchment form are the main thrusts of this work. Using a model of erosion processes that has been theoretically and experimentally verified at small scales, together with a physically based conceptualization of the channel growth process, a large scale model of catchment evolution involving channel network growth and

elevation evolution has been developed. Insights into the importance of catchment scale interactions to overall catchment form are developed through the incorporation of large scale interactions between the extending channel network and the subsequently evolving hillslopes . It will be shown that neither the properties of the channel network nor the properties of the hillslopes can be viewed in isolation but must be viewed as components of a complicated large scale nonlinear system; the drainage basin.

Using a nondimensionalization of the physical equations a number of nondimensional numbers that govern catchment form are proposed. These numbers are dependent upon such physical inputs such as climate and geological conditions. The basic tenet of the work is that it is necessary to understand the physics of the catchment processes to be able to fully understand the catchment form. The intention is to " ... identify linked process equations and so define geomorphological systems in such a way that an analytical, predictive approach can be used ... " (p 48, Huggett, 1988). It is not claimed, nor was it intended, that the proposed model accounts for all the processes occurring in the catchment. Rather a general model framework is presented which is both physically realistic and provides a vehicle by which the important interactions within the catchment can be examined.

There exists very little data at the basin scale with which a large scale model of the kind proposed here can be completely verified. There exists channel network data with which to verify, by themselves, the generated channel networks. There also exists, independently of these data, hillslope data with which to verify the generated hillslopes. Such comparisons of the individual components are successfully made. Some of the limited data that geomorphologists have collected at the basin scale is reinterpreted and alternative explanations for observed trends are offered which are consistent with the presented model.

In summary, the guiding philosophy of this work is to develop a physically

based model at the basin scale that incorporates the important interactions, with time, between the two domains of the catchment; the channel network and the contributing hillslopes.

1.2 Report Outline

This report can be broadly divided into three main sections, each section consisting of several chapters.

The first section is a literature review and analysis of previous work regarding network generation models. Chapter 2 presents a review of channel network and hillslope statistics and models. Chapter 3 briefly reviews physically based network models from outside of the hydrology and geomorphology literature. One particular nonhydrologic model, a model of leaf veins in leaves, is analyzed in detail in Chapter 4 so as to understand the processes that govern network form and growth.

The second section of this report consists of the development of the physically based channel network model. Chapter 5 presents the governing equations and provides a justification of the physics adopted. A nondimensionalization of these equations is presented in Chapter 6 and some of the geomorphological implications are explored. Appendix A details the numerical solution technique for the governing equations. Appendix B presents a reformulation of the sediment transport equations used in the model and compares this new formulation with other work. Appendix C derives the equations describing the sediment transport and channel initiation mechanisms in greater detail than does Chapter 5.

The third section of this report is devoted to the analysis of the channel networks and catchments simulated by this model. Chapter 7 presents an analysis of simulation results and correlates these results with the physics adopted in the model. Chapter 8 compares three sets of field data with results obtained from the simulations. Appendix D gives the complete set of simulations performed and their corresponding

parameter set.

Chapter 9 summarizes this work and suggests further avenues for research.

CHAPTER 2

EMPIRICALLY AND PHYSICALLY BASED CATCHMENT MODELS

2.1 Introduction

This chapter will summarize research related to the description of channel networks and hillslopes. A prodigious amount of work has been done in this area. Only that work that has relevance to the interpretation and analysis of the catchment simulations, that form the bulk of the results of this work, will be introduced here. A more comprehensive treatment of channel network properties can be found in Abrahams (1984).

Historically researchers have examined the channel networks within catchments, and the hillslopes that contribute to these channels, separately. Work has been reported on the form of the channel networks within the catchment independently of the hillslopes. In a complementary fashion work has been reported in the form of the hillslopes independently of the form of the channel network. This distinction is artificial but has been forced by the complexity of the interaction between the regimes.

The treatment in this chapter will reflect this disparate research heritage. The first half of the chapter details work on the nature and form of channel networks. The second half of the chapter details work on the nature and form of the hillslope. A short section at the end discusses some of the problems of parameterizing the interaction between the two regimes. There is also a natural distinction between that work that simply describes the catchment form at a given time and that which attempts to describe how the catchment evolves to a particular form at a given time. Hence each section on channel networks and hillslopes has been divided between that work that describes the catchment characteristics independently of time, and that work that explicitly addresses catchment evolution. The description of the time dependent physical characteristics and the processes that control them is very important. These

processes are at the heart of the physically based catchment model described in Chapter 5.

2.2 Models of Channel Characteristics

2.2.1 Descriptive, Time Independent Channel Models

2.2.1.1 Ordering Models

The most outstanding characteristic of the drainage channels in a catchment is that they form a tree—like network. Proceeding upstream from the basin outlet, each channel bifurcates into two smaller streams, each of these streams breaks into two smaller streams, etc., until the upper reaches of the network are reached and the channels terminate. An alternative, equally valid, view is that as you proceed downstream, channels merge, two at a time, until the catchment outlet is reached. In topology nomenclature the channel network forms a β (i.e., binary) tree; each bifurcation produces two branches.

The ordering models can be classified into two kinds: those that apply their hierarchy from downstream in the upstream direction, and those that apply it from the upstream source areas in a downstream direction. The former have generally been unsuccessful and will be briefly discussed first.

Gravelius in 1914 proposed a system (Figure 2.1) where the channel at the basin outlet is assigned an order of 1. The classification proceeds upstream. At the first bifurcation upstream the minor stream is assigned an order of 2 and main stream upstream of the bifurcation continues with order 1. In any general stream of order (i) at the next upstream bifurcation the minor stream is assigned an order of (i+1) and the major stream an order of (i). The selection of which stream is major and which is minor is subjective. The main problem with Gravelius classification scheme is that source streams will have different orders, depending upon their topological position in the network. This is contrary to experimental evidence which points towards source
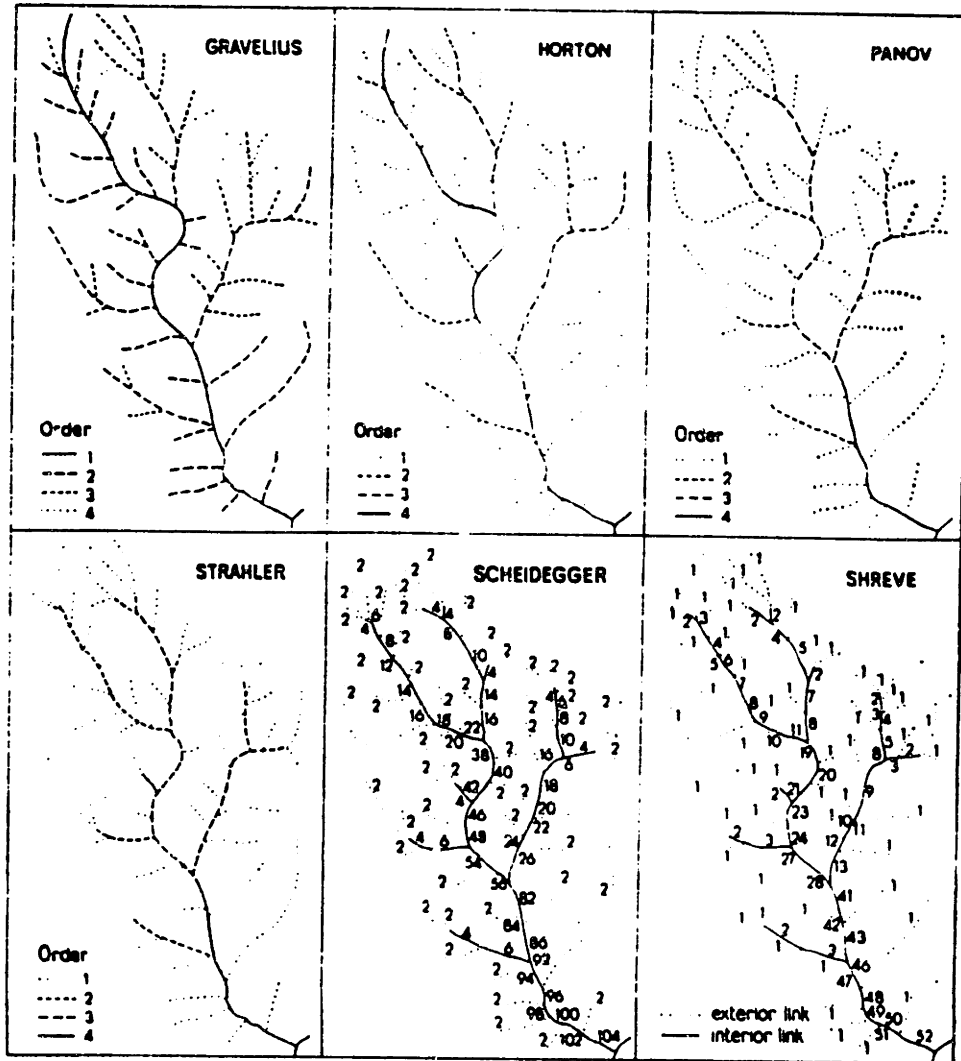
Figure 2.1: Stream classification systems
(from Zavoianu, 1985)

streams having similar characteristics. In addition, if two nested catchments are analyzed, then the same link in the field will have a different order depending upon which of the nested catchments is being analyzed. Thus conditions downstream evidently have an effect on the ordering system in the upstream reaches. This ordering scheme is thus unsatisfactory.

Horton (1945) took a fundamentally different approach to stream ordering. He ordered his streams beginning at the sources proceeding downstream. This idea underpins all contemporary network ordering schemes. As with Gravelius, Horton makes a distinction between main streams and tributary streams. Proceeding upstream on a stream of order (i), the mainstream upstream of the bifurcation is assigned the order (i) and the tributary stream order (i−1). The order of a tributary stream that is also a source (i.e., it terminates rather than bifurcates), is assigned an order of 1. The Horton scheme has the same failings as the the Gravelius scheme. Firstly, the distinction between main and tributary streams is subjective. Secondly, for two nested catchments, it is possible for a given stream in the field to be assigned a different order, depending upon the network topology downstream of that stream.

Strahler (1964) proposed an ordering scheme that removed many of the inconsistencies inherent in the Horton scheme. Zavoianu (1985) claims a Soviet researcher, B. Panos, proposed essentially the same scheme in 1948. In the scheme all source links are assigned an order of 1. When two links of the same order (i) merge, the downstream link is assigned an order one higher (i+1). When two links of unequal order merge, the order assigned to the link downstream is that of the higher order upstream link. Essentially the lower order link's topological contribution to the higher order link is ignored. However, the contribution is reflected in the physical properties of the streams that are correlated with order.

The most important channel network properties described by Strahler statistics are listed in Table 2.1. Typical ranges observed for these statistics are also listed.

TABLE 2.1

Observed Strahler Statistics

Statistic	Definition[*]	Observed Range
Bifurcation Ratio	$R_b = \dfrac{N_i}{N_{i+1}}$	$4.79^{(1)}, 4.87^{(3)}$ $5.74 \pm 1.2^{(1)}$ $4.92 \pm 0.8^{(1)}$ $3 - 5^{(2)}$ $2 - 6^{(4)}$
Stream Length Ratio	$R_\ell = \dfrac{L_{i+1}}{L_i}$	$2.1 - 2.9^{(2)}$ $2.37^{(1)}$
Area Ratio	$R_A = \dfrac{A_{i+1}}{A_i}$	$4.3 - 6.9^{(1)}$
Stream Slope Ratio	$R_S = \dfrac{S_i}{S_{i+1}}$	$1.86^{(1)}$ $1.44 - 3.46^{(1)}$

[*] i = order

N_i = total number of streams order i;

L_i = average length of all order i streams;

S_i = average slope of all order i streams;

A_i = mean area at downstream end of all order i streams

[1]Zavoianu (1985); [2]Strahler (1964); [3]Schumm (1956)

[4]Shreve (1966)

Many researchers have claimed that these ratios are independent of order i for any given catchment. Others have claimed that they should be invariant over all catchments and all orders. Shreve (1966), using topologically distinct random networks (see Section 2.2.1.2), showed that the value of the bifurcation ratio is constrained by topological arrangement of a branched network. Small sample sizes and systematic effects have made it impossible to statistically validate these claims.

Shreve (1966) and Smart (1967) noted a slight decrease in the bifurcation ratio with order. Tokunaga (1978) introduced a new stream number law, based on scale invariance of network topology, that predicts such a trend in bifurcation ratio. This new scheme has conceptual advantages when examining channel network growth, a central purpose of this work.

Tokunaga adopts the Strahler ordering scheme but abandons the stream number law based on bifurcation ratio. Instead he describes the ratios of numbers of lower order streams (i.e., (i–1), (i–2), ...) flowing laterally into the higher order stream of order (i). These ratios are assumed constant irrespective of order. These ratios are

$$\epsilon_1 = {}_i\epsilon_{i-1} \tag{2.1}$$
$$\epsilon_2 = {}_i\epsilon_{i-2}, \quad \text{etc, for all orders (i).}$$

where   i = order

    ${}_i\epsilon_j$ = average number of streams of order j flowing laterally into a single stream of order i.

The ratios ϵ_1, ϵ_2, ... are considered to be independent of order. Tokunaga defines a second statistics K, that relates these ratios so that

$$K = \frac{\epsilon_2}{\epsilon_1} = \frac{\epsilon_3}{\epsilon_2} = \dots = \frac{\epsilon_{j+1}}{\epsilon_j} \tag{2.2}$$

The statistic K is also considered to be independent of order. Tokunaga (1978) shows that the assumption of K and ϵ_1 independent of order is inconsistent with the bifurcation ratio being independent of order. As will be shown below the effect on bifurcation ratio is, for practical purposes, insignificant, but the conceptual value of the Tokunaga stream number scaling will be seen to be important later.

An outline of the nonstationary bifurcation ratio result follows. For a basin of order λ_t the average number of streams of order λ is given by (Tokunaga, 1978)

$$N(\lambda,\lambda_t) = \frac{Q^{\lambda_t-\lambda-1} - P^{\lambda_t-\lambda-1}}{Q - P} \ Q \ (2 + \epsilon_1 - P) + P^{\lambda_t-\lambda-1} \ (2 + \epsilon_1)$$

$$(2.3)$$

where

$$P = \frac{2 + \epsilon_1 + K - \sqrt{(2 + \epsilon_1 + K)^2 - 8K}}{2}$$

$$Q = \frac{2 + \epsilon_1 + K + \sqrt{(2 + \epsilon_1 + K)^2 - 8K}}{2}$$

Thus the bifurcation ratio for streams of order $(\lambda-1)$ flowing into stream of order λ in a basin of order λ_t is

$$R_b(\lambda_t-\lambda) = \frac{N((\lambda-1), \lambda_t)}{N(\lambda,\lambda_t)} \qquad\qquad (2.4)$$

25

Substituting Equation (2.3) into Equation (2.4) yields the following expression for the bifurcation ratio

$$R_b(\lambda_t - \lambda) = \left[\frac{Q}{P}\right]\left[1 + \frac{A\left[\frac{P}{Q} - 1\right]}{\left[\frac{Q}{P}\right]^{\lambda_t - \lambda - 1} + A}\right] \qquad (2.5)$$

where

$$A = \frac{\left[\frac{Q}{P} - 1\right]\left[\frac{P}{Q}\right]}{\left[1 - \frac{P}{2 + \epsilon_1}\right]} - 1$$

Tokunaga shows that the only case where the bifurcation ratio is constant is when A = 0. This is possible only in the case of K = 0 and $\epsilon_1 \neq 0$; a degenerate case that Tokunaga calls, for obscure reasons, "structurally Hortonian networks". Additionally, it may be noted that Q > P so that the bifurcation ratio increases with increasing $(\lambda_t - \lambda)$. The asymptotic result for bifurcation ratio as $(\lambda_t - \lambda)$ tends to infinity is

$$\lim_{(\lambda_t - \lambda) \to \infty} R_b(\lambda_t - \lambda) = \frac{Q}{P}$$

$$= 2K_\epsilon^2 - 1 + 2K_\epsilon\sqrt{K_\epsilon^2 - 1} \qquad (2.6)$$

where

$$K_\epsilon = \frac{2 + \epsilon_1 + K}{\sqrt{8K}}$$

A plot of this asymptotic bifurcation ratio versus values of K_ϵ is given in Figure 2.2. Tokunaga shows that the parameter values $K = 2$, $\epsilon_1 = 1$ correspond to the most probable networks for Shreve's topologically distinct random networks (see section 2.2.1.2). The asymptotic bifurcation ratio in this case is $R_b(\infty) = 4$. This point is plotted on Figure 2.2. In addition, in Figure 2.3 the variation of bifurcation ratio with $(\lambda_t - 1)$ is plotted (with parameters $K = 2$, $\epsilon_1 = 1$). It is apparent that the variation of R_b with order is small, particularly in comparison with possible variations in K and ϵ_1. Trends in R_b with order are likely to be difficult to distinguish from normal random variations.

As previously noted, the main advantage of the Tokunaga scaling hypothesis is that it characterizes networks with the order independent parameters K and ϵ_1, and that it solves a fundamental scaling problem with the bifurcation ratio which becomes apparent when describing growing channel networks. A more detailed consideration of the advantages of the K, ϵ_1 statistics over the R_b statistic will be provided in Section 2.2.2, where the temporal aspects of channel network growth will be considered.

Finally, Tarboton, et al. (1988), have, using fractal arguments, postulated a relationship between the Strahler bifurcation and length ratios for space filling networks. They assert that

$$R_b = R_\ell^2 \qquad (2.7)$$

This result is only true asymptotically as the order of the catchment approaches infinity. Thus this result is not inconsistent with the Tokunaga scaling hypothesis.

2.2.1.2 Topologically Distinct Random Networks

Complementary to the classification of networks by order is the classification of
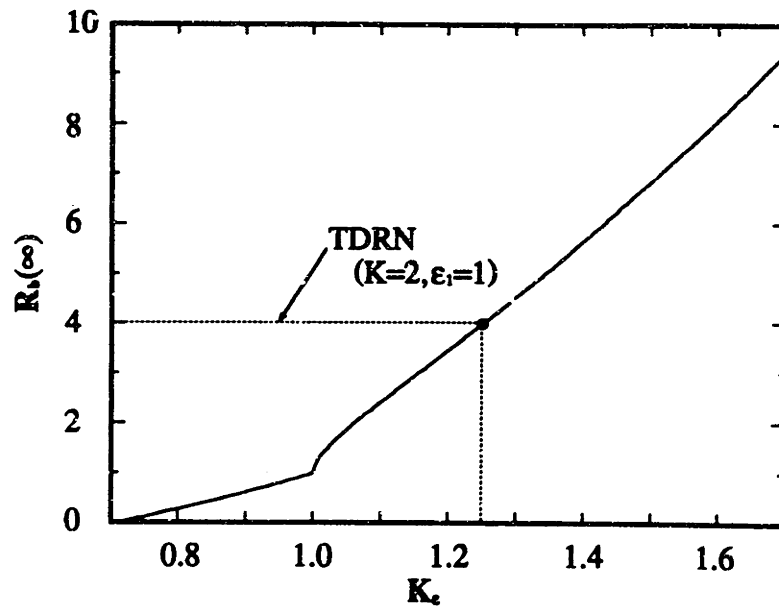
Figure 2.2: Variation of the bifurcation ratio
for an infinite catchment using
Tokunaga's number law
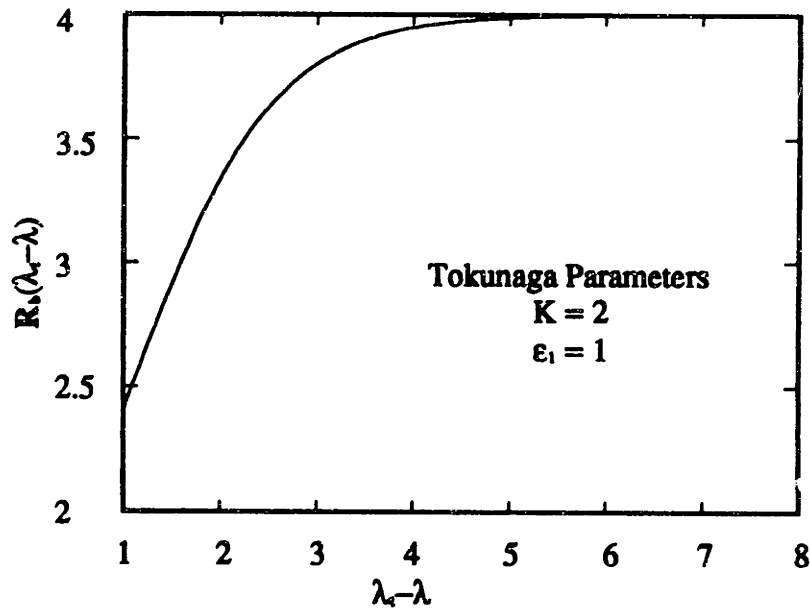
Figure 2.3: Variation of the bifurcation ratio
with catchment order using the
Tokunaga number law

networks by magnitude. Magnitude is typically discussed in the context of topologically distinct random networks (TDRN), first discussed by Shreve (1966, 1967, 1969) and further developed by other researchers (e.g., Mock, 1971). The concept of magnitude will be discussed first.

The idea of link magnitude is similar to that of Strahler order (Figure 2.1). A link is defined as that length of stream between two bifurcations. Each source link is assigned a magnitude of 1. Each downstream link is assigned a magnitude equal to the sum of the magnitudes of its upstream tributary links. A link's magnitude can then be interpreted as the number of source links contributing to that link. The total number of links, n, in a network is related to the magnitude of the highest magnitude link, m, by $n = 2m-1$.

Shreve (1967) distinguished between source links and all other links. Source links are called "exterior links", all other links "interior links". Shreve (1969), using blue lines from 1:24000 topographic maps, observed a significant difference in the mean length of interior and exterior links. Shreve (1974) presents data for Eastern Kentucky that shows mean interior and exterior link lengths to be approximately equal. Using digital elevation data Tarboton (unpublished data) analyzed a number of catchments in the United States and found that mean link lengths for interior and exterior links were not significantly different.

Mock (1971) further differentiated link types. Two types of exterior links were defined:

1. Source link: A link draining into a link downstream of magnitude 2.

2. Tributary source link: A link draining into a link downstream of magnitude greater than 2. In this case the other stream merging at the outlet is of higher magnitude and the other stream may be considered the main stream.

Four types of interior link were defined.

3. <u>Bifurcating link</u>: A link with equal magnitude links upstream and merging with a lesser magnitude link downstream.

4. <u>Tributary bifurcating link</u>: A link with equal magnitude links upstream and merging with a higher magnitude link downstream.

5. <u>Tributary link</u>: A link with unequal magnitude links upstream and merging with a higher magnitude link downstream.

6. <u>Cis–trans link</u>: A link with unequal magnitude links upstream and merging with a lower magnitude link downstream.

The tributary notation refers to the case where the link being classified is the lesser (on the basis of magnitude) of the two streams at its downstream confluence. Abrahams (1984) makes a further distinction between cis and trans links. A <u>cis</u> link is where the upstream lower magnitude link and the lower magnitude link downstream both enter on the same side of the river. A <u>trans</u> link is where they enter from opposite sides. Abrahams asserts that because of the space filling characteristics of networks, where tributaries on one side of a stream are roughly uniformly spaced, that the length distribution of cis and trans links should be different.

The notion of a topologically distinct random network (TDRN) was first introduced by Shreve (1966). Two networks of equal magnitude are topologically distinct if the arrangement of links in the networks is topologically different. They are the same if the topological arrangement of the links is the same. Shreve (1967) presents a strict definition of this concept. He symbolically represents a network by a string of E's and I's. Two networks of equal magnitude are topologically distinct if their EI string is different. The EI string is constructed as follows.

1. Start at the outlet of the basin. Travel upstream on the network at each bifurcation always taking the leftmost branch that has not already been traversed.

2. Score an I the first time a particular interior link is traversed.

30

3.   Score an E the first time a particular exterior link is traversed.
Two networks of equal magnitude are topologically distinct if their EI strings are
different.

To obtain mean properties of the networks from the TDRNs (e.g. mean width
function, mean length of Strahler streams) it is necessary to postulate a probability
distribution of occurrence for each individual TDRN. Shreve (1966) proposed that
each TDRN of a given magnitude should have an equal likelihood of occurrence. Mock
(1971) grouped together TDRNs exhibiting lateral symmetry into what he referred to
as ambilateral groups. He calculated mean properties assuming that each ambilateral
group was equally likely. These mean properties included the mean number of links of
a particular type (e.g. Source, Tributary Source) conditioned on various topological
statistics of the catchment (e.g. magnitude). Two TDRNs are members of the same
ambilateral group if their only topological difference is that a subnetwork is switched
left to right (i.e. the subnetworks look the same if one subnetwork is viewed in a
mirror).

Neither Shreve nor Mock provided any evidence that justified their postulate for
the probability distribution for the occurrence of TDRNs. Proving their hypotheses has
been difficult because of the large data requirements. Indeed it may be that the
probability of occurrence of TDRNs is conditioned on the prevailing catchment
geometry or systematic geologic effects

Shreve (1966) noted that the bifurcation ratio for individual networks varied
between 2 and 5. His random networks showed a slight upward curvature in the mode
bifurcation ratio with order, which is consistent with the Tokunaga (1978) scaling
hypothesis.

It seems reasonable to assume that if link and Strahler stream distributions are
different for Shreve's and Mock's equal likelihood postulates, then derived properties

31

such as travel time distributions and average width function will be dependent upon the postulate used. Whether the effect is important hydrologically or is verifiable in the field is an open question.

2.2.1.3 Applications of Ordering and TDRN Models

The Strahler/Horton/Tokunaga ordering and the TDRN/magnitude topological classification models for channel networks do not explicitly address temporal components of catchment form. Despite this these models have proved to be useful descriptors of catchment form at any given point in geologic time. In addition, a number of researchers have used statistics derived from these classifications schemes to develop models for the instantaneous unit hydrograph (IUH).

For many years it has been recognized that catchment geomorphology statistics can be used as predictors of catchment flood properties. The literature is replete with examples of simple applications of this kind (Chow, 1964). For instance, independent variables such as mean channel slope and mean channel length being commonly used to describe the "time of concentration" of catchments.

The first major step beyond this type of statistical study was the development of the geomorphologic instantaneous unit hydrograph (GIUH) by Rodriguez–Iturbe and Valdes (1979). In this work the IUH was interpreted as being the probability distribution that any raindrop that fell in the catchment would reach the outlet of the catchment at a given time. Roughly speaking the runoff process was conceptualized as a series of linear reservoirs, each reservoir corresponding to the area of the catchment that drained into channels of the various Strahler orders. Discharge out of the catchment resulted from a drop having a random chance of proceeding from a low order reservoir to a higher order reservoir (Figure 2.4).
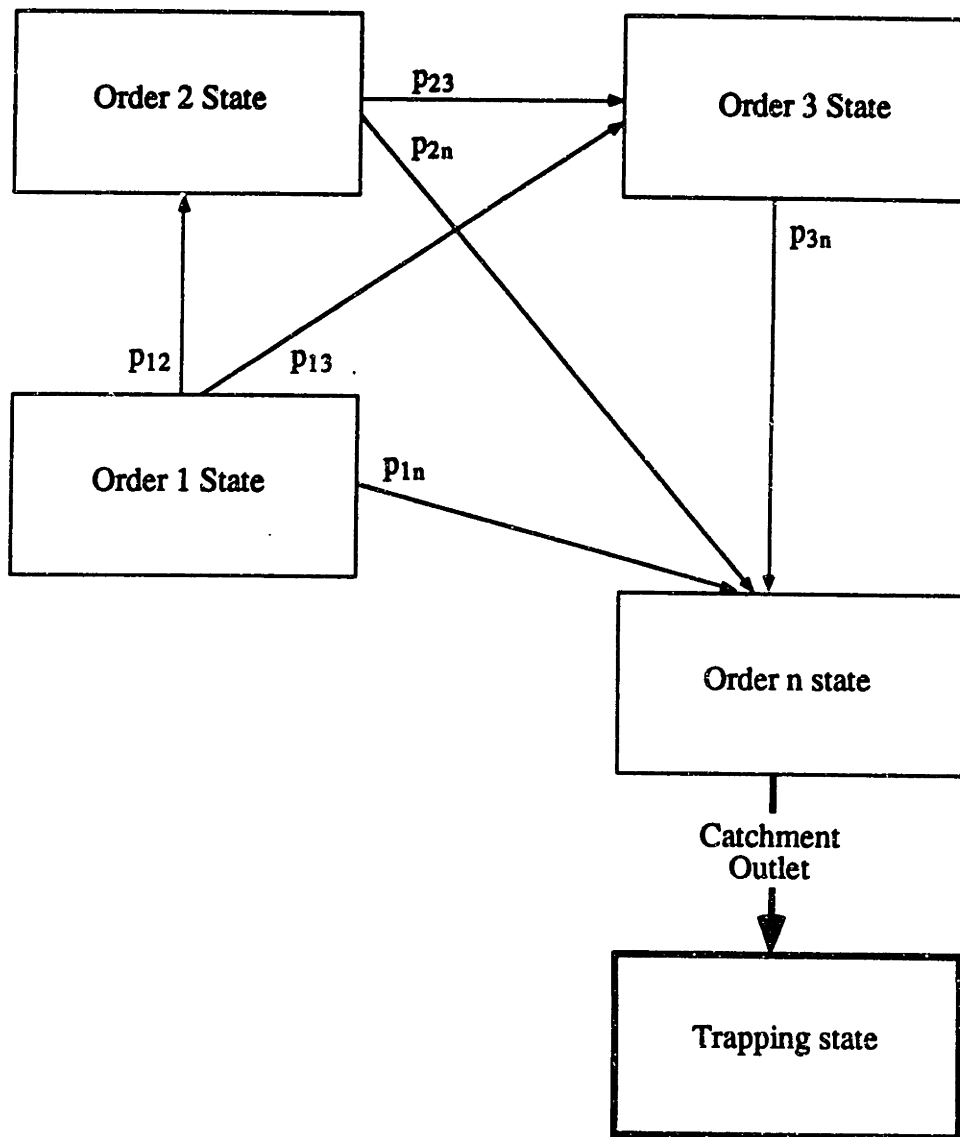
Figure 2.4: Schematic of the probabilistic Markov chain
conceptualization of the GIUH

Rodriguez–Iturbe and Valdes derived the following relationships for the peak discharge and the time to peak of the IUH

$$Q_p = \frac{1.31}{L_\Omega} R_\ell^{0.43} V \tag{2.8}$$

$$T_p = 0.44 \frac{L_\Omega}{V} R_b^{0.55} R_A^{-0.55} R_\ell^{-0.38} \tag{2.9}$$

where

R_b, R_ℓ, R_A = Strahler bifurcation, length and area ratios respectively.

L_Ω = length scale of the catchment, length of the order Ω stream, where the order of the basin is Ω.

V = velocity of flood wave peak, assumed constant over both the hydrograph and the catchment.

The important contribution of this work was the conceptual link between the statistics parameterizing the catchment geomorphology and the IUH. Later work by Gupta and Waymire (1983) and van der Tak (1988) has corrected and clarified a number of minor details in the methodology. A major assumption is the constant velocity assumption. There exists evidence to both support and refute this assumption; i.e. that flow velocities are constant within the catchment (e.g. Pilgrim, 1977) or that the wave speed is variable over the period of a single hydrograph and varies substantially with discharge (e.g. Price, 1975; Wong and Laurenson, 1983). There have been a number of attempts to remove this dependency on constant velocity (Gupta, et al., 1980; Rodriguez–Iturbe, et al., 1982).

The other important application of geomorphology is in the determination of the so-called "width function" (e.g. Surkan, 1969; Troutman and Karlinger, 1984). The roots of idea are based in the TDRN concept developed by Shreve. The width

function at some distance x is defined as the number of links whose outlets are (x–1) links distant from the outlet. Thus the width function represents the topological width of the catchment with distance. This can easily be related to the physical width of the basin. If the mean link length is independent of order or magnitude and if the mean area draining to any link is also independent of both order and magnitude, then the width function is directly proportional to the mean area draining at a given distance from the catchment outlet, the distance–area diagram. Thus the width function is equivalent to the distance–area diagram commonly used in synthetic unit hydrograph studies. If the assumption is made that the flood wave velocity is constant in both space and time, then the time–area diagram and IUH follow directly.

Mesa (1986) extended the width function by modeling the channel network as a spatially homogeneous birth/death Markov process with increasing topological distance from the catchment outlet (each link from the outlet is considered to be a generation in the birth death process). He then developed expressions for the width function conditioned on magnitude. This is equivalent to conditioning the generated network on the expected area of the catchment since (Shreve, 1969)

$$
\begin{aligned}
\text{E [Area]} \quad &= \text{E [mean link area]} \cdot \text{number of links} \\
&= \text{E [mean link area]} \cdot (2 \text{ x magnitude} - 1)
\end{aligned}
$$

Mesa compares the expected width function, conditioned on magnitude and maximum length, independently, and finds that they both give satisfactory approximations (± 50%) to the actual width functions at 6 field sites.

As an aside it may be noted that Mesa (1986) modelled the topologically random networks as birth/death Markov processes with each generation being interpreted as increasing topological distance from the catchment outlet. At any distance there is a probability of bifurcation or death (i.e., to channel bifurcation or

source area). This process, which models TDRNs conditioned on magnitude, implies some distribution of likelihood for the individual TDRNs of any given magnitude. The expected or mean width function results from an averaging over the population distribution of TDRNs generated. This can be compared with Shreve (1966) who considered each TDRN of set magnitude to be equally likely, or Mock (1971) who considered each ambilateral class to be equally likely. Gupta and Waymire (1983) showed that the birth/death modeling of Mesa (1986) is equivalent to assuming the TDRNs to be equally likely, as assumed by Shreve (1966). Thus the mean statistics of Shreve's work and Mesa's work are comparable. The same cannot be said of Mock's statistics.

In conclusion, the application of geomorphological principles to catchment response is, as yet, in a rudimentary stage. Several ideas from the literature have been briefly discussed. The techniques described above use as inputs the measured characteristics of the channel networks and the surrounding hillslopes. That the landscape results from erosion during flood events is an accepted concept. It thus seems reasonable that the catchment, sediment transport and flood hydrograph form should be fundamentally linked. An approach unifying the geomorphology and the response could potentially generalize the geomorphological instantaneous unit hydrograph and width function. Such a unification may also answer outstanding questions about the flood wave velocity distribution in space and perhaps lead to techniques relating remotely sensed elevation data to catchment response.

2.2.2 Models for Channel Network Evolution

2.2.2.1 Introduction

The channel network classification models described in Section 2.2.1 have one property in common; they all describe the channel network form independent of temporal aspects. They ignore how the network developed and look at the network at

a given time. Any attempt at geographic regionalization across a broad range of catchments implicitly assumes one of two things

1. That the characteristics described by the ordering/TDRN models are independent of time.

2. That the characteristics described by the ordering/TDRN models are dependent on time, but that all catchments are at dynamic equilibrium.

Neither of these points have been well addressed in the literature. The independence of time of ordering/TDRN statistics have never been satisfactorily examined. Many people have claimed the landscape to be in dynamic equilibrium, but as will be discussed in Chapter 6, there is considerable argument over this point. Nobody has looked at the time scales of landscape adjustment and compared them to time scales of variation of tectonic uplift or climate to see if the landscape has enough time to adjust to changing geologic and climatic conditions. The reason for this paucity of study is obvious: except at very small scales, the time scales of change are too long to make a study of timescales a realistic proposition. In addition, the statistical scatter from basin to basin makes it difficult to assert with any certainty that descriptive statistics are even regionalizable.

So far all descriptive models of channel network growth have been based on some conceptualization of reality. A major limitation is that they only model planar form of the catchment, without regard to the elevation characteristics of the network. Energy considerations are ignored. For instance, the outflow of sediment from a catchment causes a net reduction in the elevation potential energy of the catchment, yet no attempt is made to model the processes dissipating this energy. Consequently, the mechanism for growth of the network is at best conceptual. At worst there is no physical justification of the growth mechanism.

The major models that simulate network growth will be described in the following sections. All of these models consider network growth in the plane, without

regard to elevation. One model for channel network evolution will not be considered here. This model is the channel elevation, heat–engine, analogy of Leopold and Langbein (1962). This model is fundamentally different from the ones discussed in this section. It does not model network growth. Rather the elevation characteristics of a fully grown network are analyzed, on the basis of a thermodynamic analogy. It displays substantial similarities with the hillslope evolution models to be examined in Section 2.3.3 and will be discussed with these models.

2.2.2.2 Headward Growth Model

Conceptually headward growth is both the easiest to understand and the easiest to ascribe a physical interpretation. It was introduced by Howard (1971). Recently van der Tak (1988) reexamined the characteristics of the Howard model.

The principle of the headward growth model is simple. The network is simulated on a rectilinear grid of points. The channel network is a connected set of these points. This connection process of these points will now be described. A node (or nodes) on the boundary is selected as the outlet of the catchment. This point is used as the beginning of the network simulation process. The simulation proceeds in generations or "time steps". At each generation the existing network of points is examined and active growth sites, where growth may potentially occur in the coming generation, are identified. The criteria for deciding which sites are active are

1.   The active point is on a channel

2.   There is at least one empty node adjacent (i.e., not occupied by a channel) to the active site, so that growth is physically possible.

For any given simulation a growth probability, p_g, for each active site is assigned. By random selection this probability is used to determine which active sites will actually grow by one grid length in the next generation. At any growth site the actual growth direction is randomly determined. Various ad hoc rules need to be

defined to ensure that, for instance, two growing sites do not grow into the same grid point, and form a closed loop in the process.

Growth of the network stops when either of two conditions are satisfied

1.   all nodes are occupied by channels, the ultimate drainage density is achieved,

2.   a predetermined drainage density has been achieved (i.e., a given number of nodes are occupied by channels).

The model is called headward growth since the heads of the channels grow into the "hillslopes" from the growing sites.

There are a number of implementation details of the model that have important consequences for the form of the simulated networks. These details will also be significant when this technique is compared with other network simulation techniques described in later chapters. All of these details involve aspects of how the network grows. Howard (1971) and van der Tak (1988) looked solely at the form of the networks at ultimate drainage density. The final network generated follows directly from the growth processes. Once a stream segment is generated by the simulation, it is fixed in space forever. Both Howard and van der Tak note significant differences in the qualitative form of generated networks, depending on the probability parameters used in the simulation. The discussion that follows will concentrate on more subtle aspects of the growth process that are implicit in the model structure. These issues will become important in later chapters.

The first important point is the question of the locality of the growth process. Active growth nodes are determined solely on the basis of whether adjacent nodes are already occupied by channels or not. If adjacent nodes are not occupied, then the direction of growth at that node, if it is chosen to grow in the next generation, is totally random. The selection of growth direction pays no heed to the existing pattern of the network. It is only required that other channels not occupy nodes adjacent to

the active site. For instance, two channels may grow towards each other, totally "unaware" of each other's existence, until they are adjacent. At this time they cannot continue growth in the same direction and they must turn away from each other in some fashion (Figure 2.5). In diffusion limited aggregation research (see Chapter 3) this property is sometimes called locality since growth conditions are determined solely on the conditions of the network in the adjacent nodes, rather than some more global characteristic such as the overall network pattern. This locality property has its most apparent effect on channels near the domain boundaries, with the channels strongly reflecting the boundary geometry.

The second important point is the spatial distribution of the points selected for growth in the next generation. In Howard's model all active growth sites are considered equally likely to grow in the next generation. Thus an active site at the base of the network, and an active node at the outer extremities of the network are equally likely to grow. This effect is important when the pattern of network growth with time is considered. The question of the distribution of growth sites in space is addressed in detail in Chapter 3 where the concept of growth site screening is introduced. Screening effectively reduces the probability of growth for interior nodes. This is a consequence of the network around that point suppressing growth. The result of the lack of screening in the Howard model is that at intermediate times the drainage density at the root of the network will be higher than at the extremities. Ultimately the drainage density will be constant throughout the catchment — that is a requirement of the model — but at intermediate times the drainage density will be higher at the network root. This situation can be contrasted with that in the physically based network models discussed in later chapters where screening is observed.

The third important point is about the rate of growth of the network. Neither Howard nor van der Tak were concerned with how fast networks grow. They were

generation (i)

Growth equally likely in
all directions since the channel
head of A is not adjacent
to channel B

generation (i+1)

Growth in channel A
restricted from horizontal
growth because the channel
head of A is adjacent to
channel B

● Channel nodes
in this generation

○ Channel head
where growth
occurs

Existing channels
in this generation

generation (i+2)

Growth in channel A
biased away from channel B
(i.e branch head repelled)
whenever the channel head
of A is adjacent to channel B

⇨ Equally likely growth
directions in the
next generation

generation (i+3)

Growth directions of
channel A are no longer
constrained because
the channel head of A is not
adjacent to channel B

Figure 2.5: Localised repulsion of channels:
Howard's headward growth model

41

solely concerned with the ultimate form of the network. Interest in the use of headward growth models in more general channel/hillslope models is developing (Shaw and Mooers, 1988). These authors modelled channel and hillslope processes as spatially diffusive processes with greater diffusivity in the channel. The relative rates of channel growth to channel and hillslope elevation response are important in determining the transient form of catchment elevations.

The mean rate of growth of a single active site in the Howard model is given by

$$\frac{d\ell}{dt} = p_g$$

where ℓ = length of an individual channel, units of the grid interval

   t = time, units of generations

Some first order estimates of rates of <u>network</u> growth for the Howard model will now be determined. As a simple initial case consider the case where all nodes on the channel network are active sites. That is, if the network has a total length of ℓ_n grid units, there are ℓ_n active sites. This is a reasonable approximation to the early stages of network growth when drainage density is low. In this case the rate of growth of the network is

$$\frac{d\ell_n}{dt} = \ell_n \frac{d\ell}{dt}$$

where ℓ_n = total length of all streams in the network, units are the grid interval.

An <u>upper</u> bound on the rate of growth of the network at early times and low drainage densities is then

$$\ell_n = e^{p_g t} \tag{2.10}$$

Some estimate can be obtained of the effect of inactive nodes by examining the

situation near ultimate drainage density or later times. If all remaining unfilled nodes are isolated (i.e., not clustered together) then the potential sites to be filled can be given by

$$\frac{d\ell_n}{dt} = (\ell_u - \ell_n)\frac{d\ell}{dt}$$

where $\quad \ell_u \quad$ = length of stream at the ultimate drainage density

$$= (n \times m) \text{ for a } n \times m \text{ rectangular grid}$$

Near the ultimate drainage density an upper bound on the rate of growth of the network is given by

$$\ell_n = \ell_u - \alpha\, e^{-p_g t} \tag{2.11}$$

Development of an expression for the rate of growth of the drainage network at intermediate times and drainage densities is complicated by the inability to obtain explicit expressions for the number of active sites at any stage in network development. An upper bound on this rate at early times is given by Equation (2.10). An upper bound or this rate at later times is given by Equation (2.11). A postulated variation of network length with time is given by Figure 2.6. At early times the growth curve is concave upwards, reflecting the form of Equation (2.10). At later times the growth curve is concave downward reflecting the form of Equation (2.11). At some intermediate time there is an inflection point in the growth curve. Whether the curve exhibits a period of constant growth rate will be addressed later in work, in the context of the physically based channel growth model.

Note that the ultimate drainage density (and thus the mean hillslope length) of

43
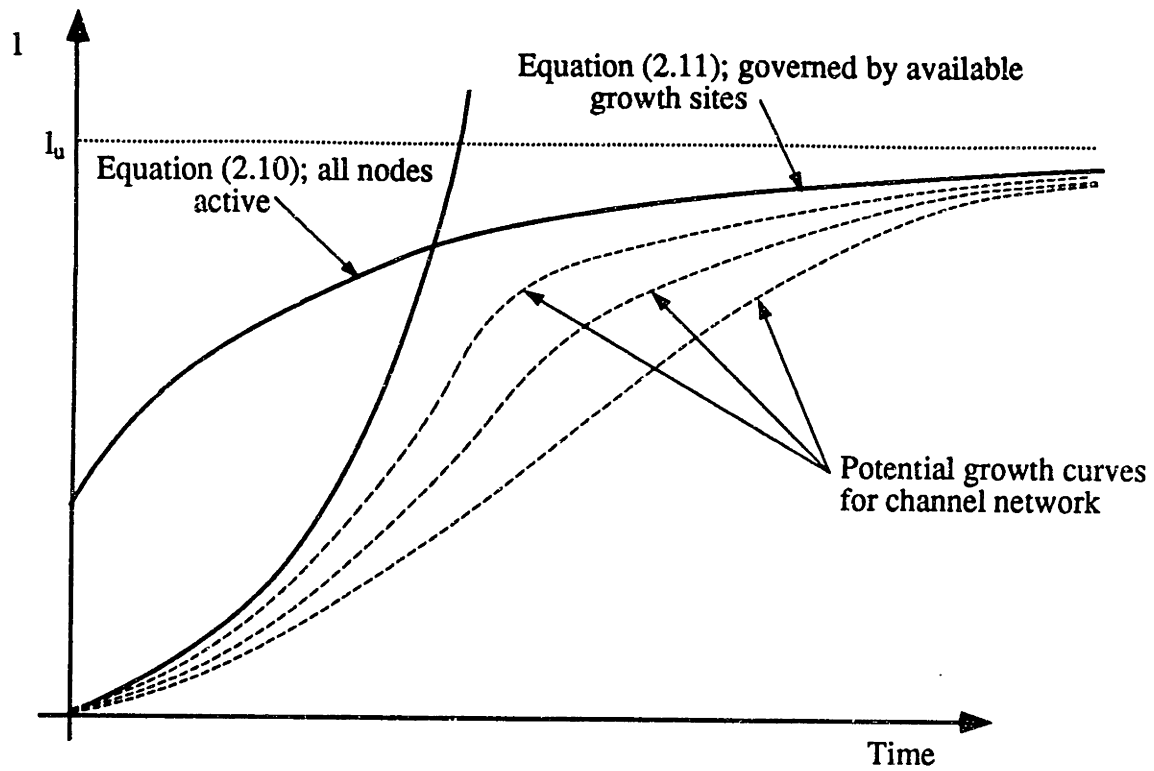
Figure 2.6: Network growth rates: Howard's
headward growth model

44

the simulated network is fixed by the grid size and spacing. The irregularity of the channel is also fixed by the grid. An alternative way to simulate networks of variable drainage is to apply an area threshold on the ultimate network, ala Tarboton, et al. (1988). Nodes with drainage areas less than the area threshold are defined to be hillslopes. A growth process could be simulated by gradually reducing the area threshold with time. This is a form of allometric growth model, to be discussed in Section 2.2.2.4.

2.2.2.3 Random Walk Models

A simulation model that is opposite in philosophy to the headward growth model is the random walk model. First introduced by Leopold and Langbein (1962) it amounts to modeling the channel network in a downward direction from the runoff source areas to the catchment outlet.

In this model source areas are considered to be distributed randomly in space, with a Poisson distribution. The model tracks the random flow downhill from the source until it either hits the boundary of the domain, in which case the boundary point is considered to be a catchment outlet, or it hits another stream, in which case it is considered to be a tributary of the stream it hits.

The single conceptual advantage of this model over the headward growth model is that it includes the idea of runoff source areas, even if somewhat loosely. A key issue is the assumption that source areas are uniformly distributed in space. The network is assured to be space filling, this follows from the Poisson distribution of source areas in space.

As with the headward growth model, this model is a local model for channel growth. Every channel grows independently of any other channel, unless they are adjacent, in which case the two channels join. As with the headward growth model, this means that channel growth occurs independently of any global drainage pattern,

imposed by the pre—existing channel network.

A difference of the random walk model from the headward growth model is the role screening plays in the form of the simulated network. Attributing a growth process with time to the random walk is clearly unreasonable. This makes it more difficult to conceptualize how the preexisting network may influence the "future" network form. It might be reasonable to condition the simulated network upon some required magnitude (it may then be interesting to look at the network form with variable magnitude). As far as screening is concerned the random walk is fundamentally different from the headward growth model. In this respect the random walk model exhibits similarities to the DLA models (see Chapter 3) because both grow by a random walk like process.

The fractal dimension of the channels in the simulated networks will be that of a random walk; that is, a fractal dimension of 1.5. The available experimental evidence for river channels indicates that this value is too high (Tarboton, et al., 1988). Thus the channels simulated by the random walk model are too irregular.

2.2.2.4 Allometric or Topological Growth

The idea of allometric growth of networks is deeply rooted in the scaling ideas of Horton (1945) and Strahler (1964). The allometric growth process may be loosely demonstrated by taking a given channel network and averaging (blurring) the network at a large horizontal scale and then gradually reducing the scale of averaging. Initially a very coarse network will be observed with more and more detail becoming visible with time as the blurring is reduced. Allometric growth can be modelled by allowing a topological representation of the network, either Strahler order or topologically distinct random networks, to grow a single link at a time. This is related to Horton and Strahler's work because they envisaged a scale independence of their topological statistics. A third order basin can be interpreted as an early time version of a fourth

46

order basin, with all the first order links removed. Woldenburg (1966) was the first to formalize this notion of topological growth of networks.

It is important to recognize that the allometric growth process is fundamentally different to either the headward growth or the random walk model. The latter models simulate the physical dimensions of the network. The allometric growth models are topological growth models. Strahler first order streams spawn lower order streams; these lower order stream even lower order streams, etc. Physical properties are not considered except that a priori distribution on link or Strahler stream characteristics may be applied. Conceptually the growth models are more closely related to the infinite topologically distinct random networks of Shreve (1967).

Very little work has been done with allometric growth simulation models. The exception is a landscape simulation study of Kelley, et al. (1988). The main reason for discussing the allometric growth is to note the conceptual advantages of the Tokunaga (1978) stream number law over that of Horton or Strahler.

Tokunaga (1978) points out that the Horton/Strahler bifurcation ratio of any particular order of stream is dependent upon the order of the catchment. This characteristic was first observed by Smart (1967), though Shreve (1966) noted an order dependence in a slightly different context. The Tokunaga stream numbers hypothesis, parameterized by constants K and ϵ_1, has been described earlier. These constants are independent of catchment order and predict a bifurcation ratio dependent on catchment order. The Tokunaga model gives the number of (i–1), (i–2),...,(i) streams that will be tributaries to an order (i) stream. All the stream orders may be incremented by 1, and new first order streams added and the Tokunaga scheme will still be applicable to the new allometrically grown network. The same is not true for the Horton/Strahler bifurcation ratio. Thus the Tokunaga stream numbering scheme exhibits scale independence in the growth processes (at least in the mean stream numbers), whereas the bifurcation ratio does not. It must be noted, however, that

47

both the bifurcation ratio and Tokunaga's number rule are relationships for the mean number of streams; they do not parameterize the variability about the mean. It will be shown in Chapter 7 that neither number law is particularly good for normal sized networks because of the large variability around the mean. In the mean it would appear that the Tokunaga hypothesis is more appealing, simply because of this scale independence, but for practical purposes neither measure is particularly satisfactory because of this variability.

2.3 Models of Hillslope Characteristics

2.3.1 Introduction

The hillslope is defined as the intervening area between the channel network. This, rather vague, definition of what constitutes hillslope is taken as fact in much work on hillslope geomorphology. Traditionally geomorphology has been artificially divided into two areas: channel network geomorphology and hillslope geomorphology. The interaction between the two regimes has, until recently, been largely ignored so that accurate, and general, definitions of where channels end and hillslope begins, have not been necessary.

For the purposes of this section, a channel will be defined as a well contained flow of limited extent perpendicular to the flow direction and with a width of flow comparable to the depth of flow. Hillslope will constitute everything else. Hillslope processes may or may not be dominated by either Horton (i.e., overland flow) or Dunne (i.e., subsurface saturation) runoff production mechanisms. Figure 2.7 illustrates what this channel definition may look like in nature. This is a rather more pragmatic, less accurate, definition of hillslope and channel than some consider acceptable (e.g. Kirkby,1988). Nevertheless this definition will suffice for the discussion of hillslope processes in this section. It will also allow the discussion of what constitutes hillslope and channel in a separate, self—contained section (Section 2.4)
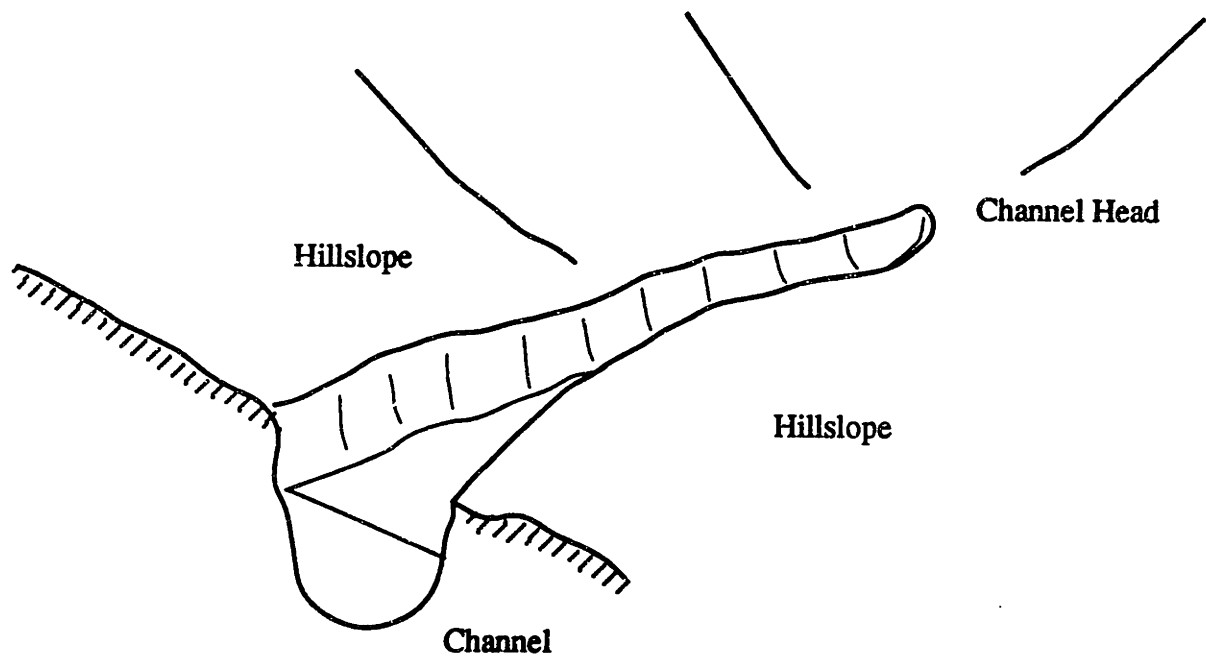
Figure 2.7: Schematic of channel and hillslopes

devoted to hillslope channel interactions.

Analogously to the treatment of channel networks in the previous section, this section will be divided into two parts. The first part will concentrate on those descriptive measures of the hillslope that attempt to characterize hillslope form at any give time, i.e., no explicit consideration of temporal effects. The second part of this section will be devoted to models of hillslope evolution. In contrast to the channel network case, most hillslope research has been devoted to understanding temporal aspects of hillslope form.

2.3.2 Descriptive, Time Independent Hillslope Models

Rather surprisingly, considering the amount of research that has been devoted to hillslope hydrology and hillslope geomorphology, there is a dearth of statistics for summarizing hillslope form. This may reflect the infant state of knowledge about hillslope runoff processes. It may also reflect the relative ease with which channel properties may be identified, even from a low resolution topographic map. Hillslope form is a distributed property, channel are easily identified lines. Hillslope runoff is also much more complicated than channel flow. Most of the statistics to be described below provide no direct insight into hillslope runoff processes. The work of Beven (1979) and O'Loughlin (1981) on subsurface saturation is an obvious exception to this gross generalization.

Mandelbrot (1983) triggered an interest in fractal characterization of landscape and hillslopes with his beautiful three—dimensional pictures of mountain landscapes, simulated using fractal techniques. Numerous other authors have duplicated this work. These techniques will not be discussed here, nor will the techniques for measuring fractal dimensions. Detailed discussion of these results will be delayed until Section 2.3.4 because the interpretation of these measurements will require some consideration of hillslope evolution processes like creep, erosion, and rockfall, which are discussed in

50

Section 2.3.3.

Horton (1945) introduced the most fundamental statistic for hillslope characterization, the drainage density. The drainage density is defined as

$$D_d = \frac{\ell}{A}$$

where D_d = drainage density

ℓ = total length of channels in the catchment

A = total area of the catchment

and it is related to the mean hillslope length by the relationship

$$\ell_h = \frac{1}{2D_d} \tag{2.12}$$

where ℓ_h = mean hillslope length.

Considerable variation of drainage density among catchments has been observed. Horton interpreted the mean hillslope length as that length of overland flow that was just sufficient to initiate erosion. Once erosion began channels were formed. This surface runoff dominated transport process for channel formation was the predominate interpretation of drainage density and hillslope length until Dunne (1969) showed that groundwater runoff processes (e.g., subsurface saturation) can be dominant in many catchments. It will be shown in Section 5.5 that the subsurface saturation concept is consistent with the notion of a fixed support area, or source area, for channel formation (Figure 2.8). That is, the area contributing to the channel head is fixed. Effectively this means that the drainage density may be considered to be a function of the channel support area and the planar geometry of the groundwater flow
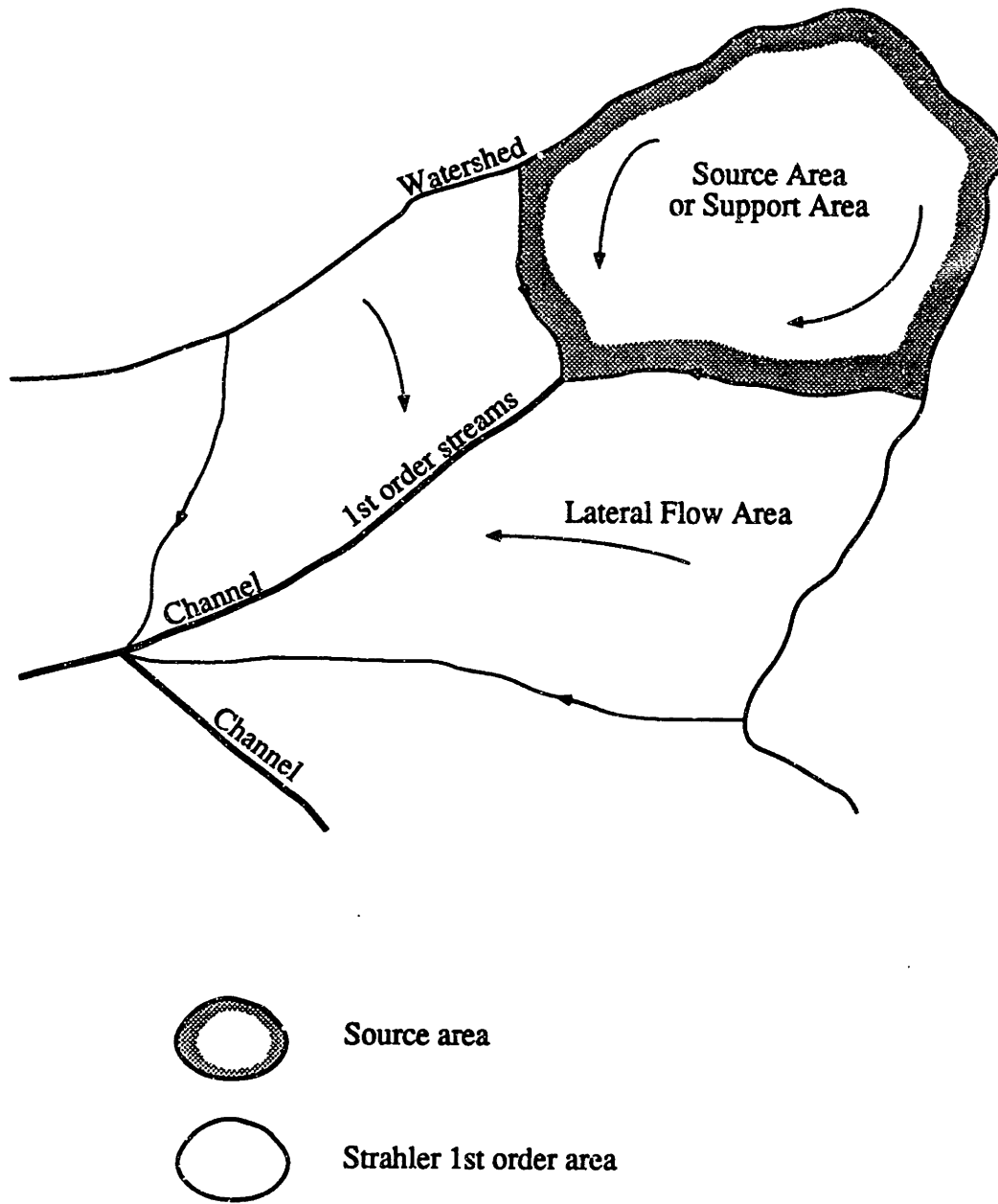
Figure 2.8: Schematic of hillslope regions in
a Strahler first order area

on the hillslopes around the channel head.

Although the runoff processes proposed by Horton and Dunne are different the effect on the interpretation of the drainage density is the same; it is the inverse of the mean hillslope length. The only important difference between the Horton and Dunne interpretation is the physical process that governs hillslope runoff.

The research relating drainage density to catchment conditions is unfortunately unable to differentiate between these two opposing runoff processes. Gregory and Walling (1968) and Gregory (1976) summarized the important physical inputs to drainage density. Increased drainage density has been positively correlated with

1.   increased extreme runoffs

2.   increased mean runoffs

3.   reduced soil permeability

4.   reduced vegetative cover

5.   increased relief

6.   increased sediment yield

For instance, many authors have noted a relationship between drainage density and the mean annual peak discharge of the form (Gregory, 1976)

$$Q_p = \alpha \, D_d^{\beta}$$

where the power on the drainage density, β, is typically of the order of 2. Rodriguez–Iturbe and Escobar (1982) examined this relationship between drainage density and extreme runoffs on the basis of energy conservation principles. They concluded that "drainage density is both the cause and effect of energy expenditure of the effective rainfall" (p. 137). Thus the drainage density and extreme runoffs should be viewed as being interactive variables.

The other main statistic that has been used to characterize hillslope form is the

hypsometric curve. This curve is a function relating the amount of area in a catchment above a given elevation, in a non–dimensional form. Schumm (1956) examined the hypsometric curve over a range of catchments in a landfill in New Jersey (Figure 2.9). He attributed differences in the hypsometric curve for different catchments to differences in ages of the catchments. It must be cautioned, however, that hypsometric curves for a single catchment through time have never been plotted, so that such assertions about age differences are somewhat speculative at this stage.

Strahler (1952, 1964) present a number of statistics to summarize the form of the hypsometric curve.

1.   Relative area lying beneath the curve.

2.   Slope of the hypsometric curve at the inflection point.

3.   Degree of sinuosity.

He notes that, though did not explain why, many hypsometric curves can be fitted by the curve

$$y = \left[\frac{d - x}{x} - \frac{b}{d - b} \right]^{z}$$

where y = relative height = $\frac{h}{H - H_0}$

x = relative area = $\frac{a}{A}$

b, d, z = fitting coefficients

a = area with elevation greater than elevation h

A = area of the catchment

H, H_0 = maximum and minimum elevations within the

catchment

It is generally believed that the hypsometric curve tends, with age, to a characteristic form called the monadnock stage. In this stage the catchment consists of

(a) Elevations normalized against
the current catchment relief
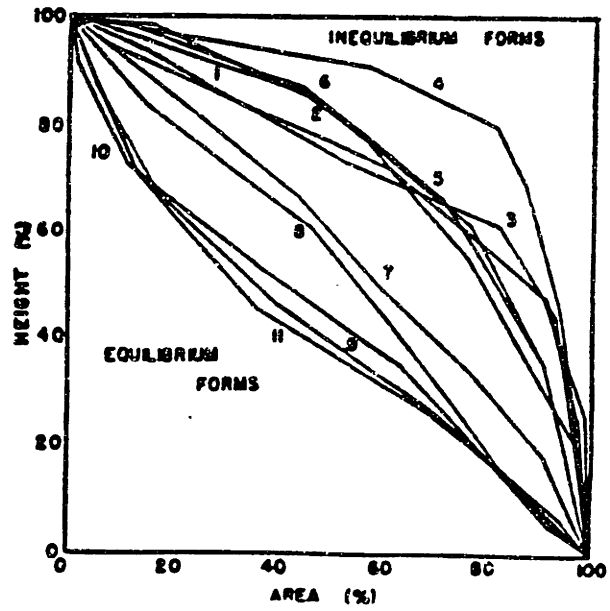
(b) Elevations normalized against
the original catchment relief
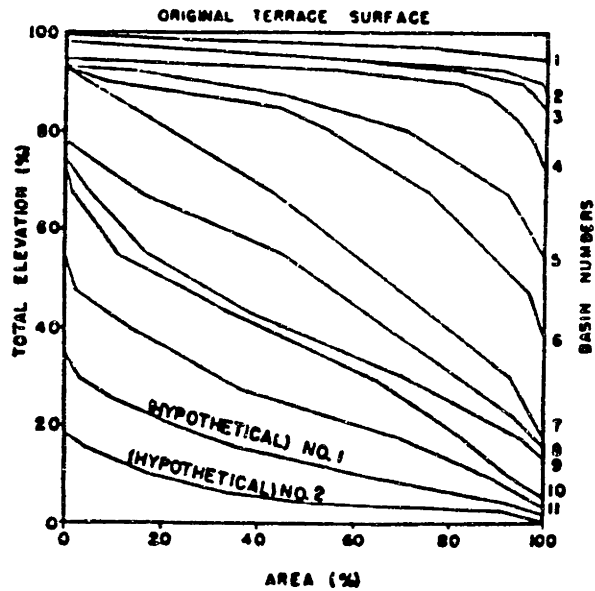
Figure 2.9: Hypsometric curves, Perth Amboy, New Jersey.
(from Schumm, 1956)

55

a number of isolated bodies of resistant rock (called monadnocks) surrounded by a "generally subdued surface" (p. 4–69, Strahler, 1964).

Strahler (1964) defines a ruggedness number for the hillslopes as

$$R_h = \frac{H\ D_d}{S_g} \qquad (2.13)$$

where
R_h = hillslope ruggedness number

H = mean hillslope drop from the divide to the stream

D_d = drainage density

S_g = mean slope of the hillslope.

For a uniform plane the hillslope ruggedness number is 0.5. Strahler (1958) found that the average value for several catchments ranged from 0.3 to 1.0. No explanation was offered as to how this variation reflected differences in the hillslope form. An explanation for the variation will be offered in Section 7.5.

Strahler (1964) related hillslope and channel slopes. This relationship was of the form

$$S_h = \alpha\ S_c^{\beta} \qquad (2.14)$$

where
S_h, S_c = hillslope and channel slopes, respectively

α, β = fitting coefficients, 4 and 0.8, respectively.

Finally Beven and Kirkby (1979) and O'Loughlin (1981) present models for predicting regions of subsurface saturation. This distribution of saturated areas is parameterized as that area where

$$\frac{q}{KS} > \text{threshold constant} \qquad\qquad (2.15)$$

where   q = discharge = RA

         R = rainfall depth

         A = upstream contributing area

         K = subsurface conductivity

         S = hydraulic slope of groundwater table

O 'Loughlin (1981) defines the threshold constant in this equation as

$$\text{Threshold constant} = h_0 + x_s(M - S) \qquad\qquad (2.16)$$

where h_0 = depth of water table

         x_s = width of the saturated region around the channel

         M = surface slope of the hillslope

This equation defines the point at which the convergence of flow lines will result in the groundwater table reaching the surface. Where M = S this definition is essentially that of Beven and Kirkby (1979). A number of authors have noted the patterns of subsurface saturation excess predicted by these models are qualitatively consistent with those field workers. In addition, recent work in forested catchments in South East Australia, indicates a positive correlation between the wetness coefficient, defined as

$$WC = \frac{q}{KS}$$

and surface soil moisture content (O 'Loughlin, personal communication). Such a result is consistent with Beven (1983), where he parameterizes equation (2.15) in terms of a moisture deficit. It would seem that measures of the spatial distribution of these

57

wetted areas would be a useful measure of the form of the hillslope. Such work has not been done.

This concludes the treatment of statistics used to describe hillslope form. The following section will concentrate on models of hillslope evolution and the processes that are important controlling agents in the evolution of the characteristics of hillslopes:

2.3.3 Models for Hillslope Evolution

This section describes some models that have been proposed to simulate the mass movement process on hillslopes. These models are of fluvial overland flow erosion, creep, rainsplash and rockfall. One of the models discussed, that of Leopold and Langbein (1962), was originally presented in the context of elevation evolution in channels, rather than elevation evolution in hillslopes. It is discussed in this section because conceptually it displays greater similarities to the hillslope models below than it does to the channel network models discussed earlier. The Leopold and Langbein heat equation model discussed below, unlike the channel evolution models, does not model the growth of channel networks, but rather the elevation changes once they have grown. In a similar way the hillslope models are solely concerned with elevation changes with time.

There is general agreement in the literature that the elevation in the channel networks and hillslopes may be modelled as an "open dissipative system" (Leopold and Langbein, 1962; Scheidegger, 1970; Thornes, 1983; Huggett, 1988). The general form of the governing equation for elevation changes in an open dissipative system is

$$\frac{\partial z(\underline{x})}{\partial t} = \text{sources} - \text{sinks} + \text{spatial coupling} \qquad (2.17)$$

where   z = elevation

$\underline{x}$ = horizontal dimension, $\underline{x} \in \Omega$ where Ω is the catchment

t = time

This equation expresses the continuity of sediment, and thus elevation, within the catchment. Sediment mass may be considered as elevation per unit area so that a continuity equation for sediment is equivalent to a state equation for elevation as in equation (2.17). The equation is called a dissipative system because of the sink term. There is an outflow of sediment; elevation is being dissipated away. For a catchment this dissipation term will be the sediment transport process. The sediment transport process will also be the coupling term since changes in spatial patterns of elevations will cause changes in drainage pattern and sediment transport and thus further changes in the spatial pattern of elevation.

Thus the sources, sinks and spatial couplings are the physical processes that will sculpt the landscape. This section is devoted to the discussion of these processes. There is some disagreement about which processes dominate at the catchment scale, and thus need to be modelled. Given the right conditions, though, all the processes to be discussed have the potential to dominate at some scale. The difficult, and largely unanswered, question is to determine what those "right" conditions are. A nondimensional analysis in Chapter 6 will begin to address this question.

The source term in Equation (2.17) is important since without a source term the only steady state solution for the elevation will be zero everywhere, i.e., a flat plain. In the catchment setting this source term will be tectonic uplift. It is difficult to conceive of any other physical process that would result in a net increase in elevation throughout the whole catchment from watershed to outlet. Because of the processes of aggradation and degradation erosion processes can produce localized elevation increases but the net elevation over the whole catchment must be decreased.

The spatial coupling and sink terms are potentially the most important terms of the equation. It is through these terms that the spatial pattern of hillslopes and channel network come into being. The remainder of this section will be devoted to discussing the physical processes that are operative at the catchment scale which can be thought as being spatial coupling and sink terms.

Initially, fluvial sediment transport will be discussed. They typically operate as either sheet or rill erosion. Other, secondary, processes that act at the hillslope scale will then be discussed. These processes include creep, rainsplash, and rockfall. These processes are typically modelled by a Fickian diffusion term, and it will be shown that they act in a fundamentally different fashion to fluvial sediment transport.

This section will close with a consideration of some of the applications of these models. The development of characteristic hillslope cross–sections will be discussed in this context. A number of simulation models will be described and the implication of their results to interpretations of hillslope form discussed. Some important failings of these models will be identified.

Kirkby (1971), in a seminal work on hillslope form, proposed a general framework within which one–dimensional transport by Hortonian overland flow on the hillslopes could be viewed. He used a continuity equation for sediment transport of similar form to Equation (2.17). For low slopes he proposed a general formulation of the sediment transport–spatial coupling term of

$$\frac{\partial z}{\partial t} = \frac{\partial}{\partial x}(Q_s)$$

(2.18)

where Q_s is the transport law for the hillslope.

The sediment transport law was then related to the appropriate governing process dominant in the hillslope. Kirkby suggested that a transport law of the following form was appropriate.

$$Q_s \ \alpha \ Q^m \ S^n \qquad\qquad\qquad (2.19)$$

where 　　　Q 　= discharge

　　　　　　S 　= slope

　　　　　　m, n = coefficients dependent on the governing process

　　　　　　　　= 0 and 1, respectively, for soil creep

　　　　　　　　= 0 and 1–2, respectively, for rainsplash

　　　　　　　　= 1.3–1.7 and 1.3–2, respectively, for soil wash

　　　　　　　　= 2–3 and 3, respectively, for fluvial transport in channels.

Though very comprehensive, Kirkby's model was by no means complete. Better understanding of some of the physical processes has since been obtained; for example, rock fall and rainsplash. In addition, Kirkby only considered one–dimensional flow. More sophisticated treatments, all based on Hortonian flow runoff mechanisms, will now be considered.

The case of overland flow sediment transport will be dealt with first. To model the slope development of a two–dimensional unchannelized basin, which can be considered a hillslope, Luke (1974) proposed the continuity equation.

$$\frac{\partial z}{\partial t} = \frac{\partial}{\partial x}\left[q_s \frac{\frac{\partial z}{\partial x}}{S}\right] + \frac{\partial}{\partial y}\left[q_s \frac{\frac{\partial z}{\partial y}}{S}\right]$$

where 　　　S 　= maximum downhill slope

$$= \left[\left[\frac{\partial z}{\partial x}\right]^2 + \left[\frac{\partial z}{\partial y}\right]^2\right]^{1/2}$$

　　　　　　q_s 　= sediment flux/unit width.

This equation is simply a restatement of Equation (2.18) in two–dimensional

form. Using a transport equation of the generic form

$$q_s = S\ f(Q)$$

where $f(\cdot)$ = monotonically increasing function of Q.

Luke derived characteristic forms for the hillslopes based on some simple boundary conditions. Slope profiles were consistent with those proposed by Kirkby (1971).

Smith and Bretherton (1972) looked at two hillslope characteristics.

1.   The tendency of a one–dimensional hillslope to converge to a characteristic form.

2.   The tendency of a one–dimensional hillslope to channelize.

This latter work was carried out by deriving the flow and sediment transport continuity equations for a smooth, one–dimensional hillslope, and applying a two–dimensional elevation perturbation. If the perturbation grew, then this indicated a tendency for the hillslope to rill or channelize. The governing equations used were the continuity equations for Hortonian flow and sediment transport

$$\frac{\partial Q}{\partial x} = \alpha \tag{2.20}$$

$$\frac{\partial Q_s}{\partial x} = \frac{\partial z}{\partial t}$$

where α = runoff rate

Q_s = sediment transport equation = $\beta\ Q^n\ S^m$.

After applying the two–dimensional perturbation to these governing equations, a perturbation equation for elevation was obtained. This equation governed the growth of the perturbations with time and was

$$\frac{\partial}{\partial x}\left[\frac{1}{B_1}\frac{\partial z'}{\partial t}\right] = \frac{\partial}{\partial x}\left[\frac{A_1}{B_1}\frac{\partial^2 z'}{\partial y^2}\right] + \frac{\partial}{\partial x}\left[\frac{1}{B_1}\frac{\partial}{\partial x}\left[\frac{\partial Q_s}{\partial S}\frac{\partial z'}{\partial x}\right]\right] + \frac{1}{S}\frac{\partial^2 z'}{\partial y^2}$$

where

z' = elevation perturbation

       $= (z - z^*)$

z^* = the nominal elevation about which the governing equations were linearized

x, y = streamwise and cross–stream directions, respectively

A_1, B_1 = constants dependent on the mean slope S, and the mean sediment transport, Q_s.

They examined the stability of this perturbation and found that it was always unstable; that is, any initial perturbation in elevation always grew. They also found that the narrower the perturbation was in the cross–stream direction, the faster it grew. An important conclusion of this work was that whenever fluvial sediment transport is the dominant transport process then uniform overland flow is unstable; rilling will always occur. This conclusion is of some concern since it is contrary to observation where uniform overland flow and sheet erosion have been observed with no tendency toward rilling. Dunne and Aubrey (1986) demonstrated that in hillslope regions where rainsplash, a stabilizing process, dominates fluvial transport the tendency to rill is suppressed, but downslope as fluvial transport begins to become important, rilling begins to occur.

There are a number of secondary processes that occur on the hillslopes that may need to be considered, depending on the situation. Many of these processes have been modelled by a Fickian diffusion term of the form

$$\frac{\partial z}{\partial t} = D_z \frac{\partial^2 z}{\partial x^2} = D_z \frac{\partial S}{\partial x} \qquad (2.21)$$

where $D_z^.$ = diffusivity

S   = slope

As previously noted Kirkby (1971) modelled creep and rainsplash with such a term. Dunne (1980) supported the idea of modeling rainsplash by Equation (2.21) but Dunne (1988) noted that the proportion of the hillslope over which rainsplash is important, at the field scale, is quite small. Coventry, et al. (1988) suggested that rainsplash may be an important process for detachment of particles which are then transported by traditional fluvial mechanisms. This process they called rainflow. However, the rates of rainflow transport reported were 1–10 mm/1000 years, very small by hillslope erosion standards, where sheet wash erosion may be as much as 1--2 mm/year, particularly in steep tropical environments. Culling (1963) first proposed that soil creep could be modelled by a diffusive term as in Equation (2.21).

Leopold and Langbein (1962) pursued the idea of an analogy between the heat equation and elevation. They based their argument on the second law of thermodynamics; entropy in a closed system should always increase. On the basis of their ideas, Scheidegger (1970) proposed that elevation could be modelled by

$$\frac{\partial z}{\partial t} = D \frac{\partial^2 z}{\partial x^2} \qquad (2.22)$$

While this equation is the same as Equation (2.21), the underlying principle is different. In this latter case, the researchers appealed to the principle that energy dissipation within the catchment should be uniformly distributed throughout the catchment.

The unifying principle of all the processes described above is that diffusivity of elevation is spatially constant. Appendix B shows that it is possible to reformulate the fluvial sediment transport formula of Equation (2.19) in the form of an

advection–diffusion equation. In one dimension, that is

$$\frac{\partial z}{\partial t} = -v\,\frac{\partial z}{\partial x} + D\,\frac{\partial^2 z}{\partial x^2}$$

where v = advection velocity of an elevation perturbation

        D = diffusivity of an elevation perturbation.

In this case of fluvial sediment transport, however, neither v or D are constant and vary in a predictable fashion in space, dependent on both slope S and discharge Q.

Finally Kirkby (1971) and Andrews and Bucknam (1987) propose a model for debris flows that for high slopes is in the form of nonlinear diffusion and for low slopes becomes linear. For high slopes the equation is

$$\frac{\partial z}{\partial t} = K_0\,\frac{\partial}{\partial x}\left[\frac{S}{1 - (\frac{S}{\theta})_r}\right] \tag{2.23}$$

where K_0 = diffusivity

        θ_r = angle of repose of the debris material.

The expression simplifies to the following equation for low slopes $(\frac{S}{\theta}_r < 0.1)$

$$\frac{\partial z}{\partial t} = K_0\,\frac{\partial^2 z}{\partial x^2} \tag{2.24}$$

For a typical angle of repose of 30°, the latter equation is applicable to slopes up to about 7°. Fitting fault scarps with these equations and paleodating them, Andrews and Buckman derived a value for K_0 of 0.61 m^2/1000 year.

Geomorphologists have been interested in determining characteristic hillslope forms. This is the nondimensional shape that a hillslope will achieve with time, given

uniform inputs. Some of the studies that have examined the issue of characteristic profiles are Scheidegger (1970), Smith and Bretherton (1972), Kirkby (1971), Luke (1974), Ahnert (1976), and Andrews and Bucknam (1987). The major differences between the studies are the form of transport law used and whether the hillslope was considered one dimensional or two dimensional.

Scheidegger (1970) obtained an analytic expression for hillslope evolution using Equation (2.22) as the governing equation. The solution for elevation is

$$z = \alpha_1 \, e^{\alpha_2 t} \, e^{\alpha_3 x}$$

In nondimensional form this can be written as

$$z' = e^{\alpha_3' x'} \tag{2.25}$$

where
$$z' = \frac{z}{\alpha_1 \, e^{\alpha_2' t'}} = \text{nondimensional elevation}$$

$$t' = \frac{t}{T} = \text{nondimensional time}$$

$$x' = \frac{x}{L_x} = \text{nondimensional horizontal distance}$$

T = time scale

L_x = horizontal length scale

$\alpha_2' = \alpha_2 \, T$

$\alpha_3' = \alpha_3 \, L_x$

The importance of Equation (2.25) is that it shows that the hillslope has a characteristic longitudinal profile which simply declines uniformly in space with time. In some sense we might consider the nondimensionalized form of the elevation to be in

dynamic equilibrium and to have some characteristic form. The important topic of dynamic equilibrium will not be pursued further here but will be dealt with in detail in Chapter 6.

Both Kirkby (1971) and Ahnert (1976) present qualitative indications of characteristic hillslope profiles depending upon the form of the governing transport process. Kirkby classifies his profiles on the basis of the transport law in Equation (2.19). Ahnert (1976) presents his on the basis of the different types of physical processes (e.g. creep, landslide, erosion). Both sets of results are consistent. This can be seen in the one–dimensional hillslope illustrated in Figure (2.10). If tectonic uplift is in equilibrium with the erosion, then, using Kirkby's formulation of transport from Equation (2.19), the rate of change of elevation becomes

$$\frac{\partial z}{\partial t} = 0 = T - \beta\, Q^m\, S^n \tag{2.26}$$

where
$$Q = R\, x$$
$$S = \frac{\partial z}{\partial x}$$
$$T = \text{rate of tectonic uplift}$$

Solving this differential equation yields the following solution for elevation along the hillslope

$$z = \left[Z_0 - \left[\frac{T}{\beta R^m}\right]^{1/n} \frac{n}{n-m+1} x^{\left(\frac{n-m+1}{n}\right)} \right] \tag{2.27}$$

where Z_0 = the elevation of the watershed

Thus if the transport law is such that $m > 1$, then the final profile is concave up as shown in Figure (2.10b). If $m < 1$, then the final profile is concave down. Equation (2.27) is consistent with the results presented by both Kirkby and Ahnert.

67
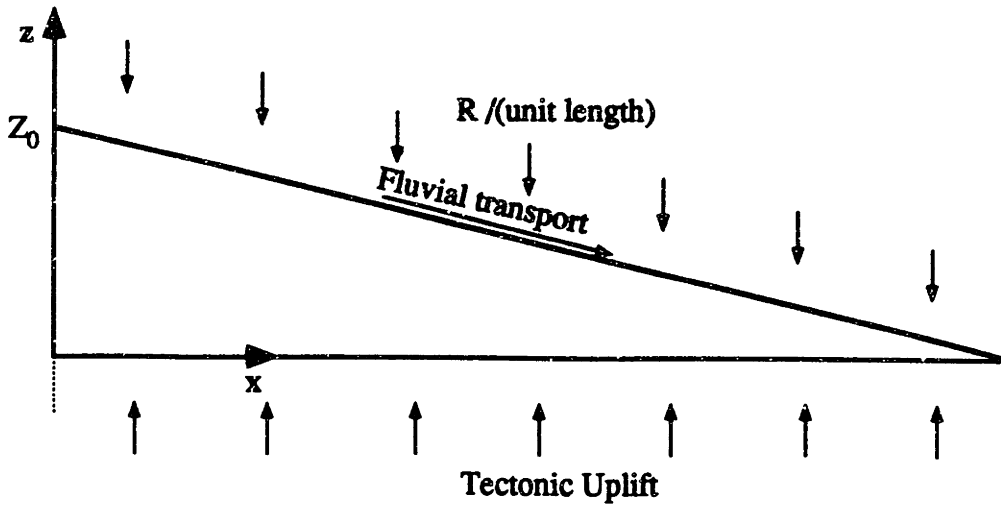
(a) Initial (t=0) Profile
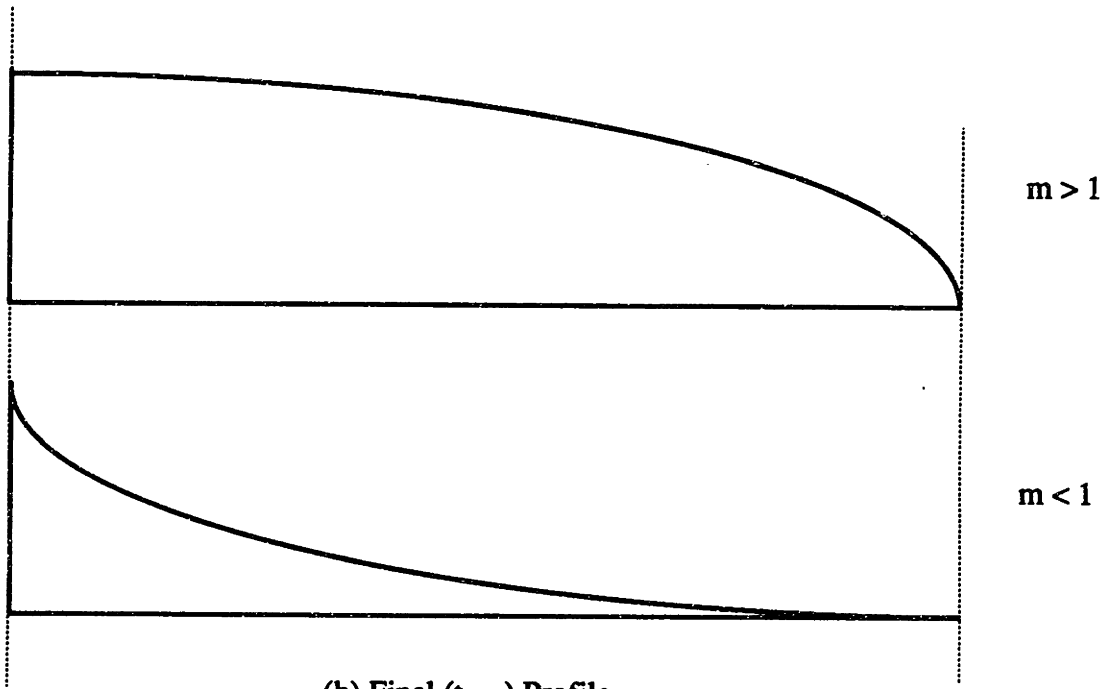
(b) Final (t=∞) Profile

Figure 2.10: Classification of characteristic hillslope profiles

Geomorphologists have also argued about the relative contributions of hillslope retreat and hillslope degradation to the hillslope form and how it develops. There has been no quantitative analysis of this. The question is addressed in Appendix B. Suffice to say that rates of hillslope retreat and hillslope degradation may be obtained directly from the governing equations on the hillslope.

The aforementioned results are only true for one—dimensional hillslopes or for very restrictive conditions in two dimensions. To generalize these results computer simulations are necessary.

The two main approaches to catchment simulation are that of Ahnert (1976) and that of Cordova, et al. (1982), and Roth, et al. (1989). The former author developed a comprehensive hillslope simulation model that accounts for a large number of transport processes including tectonism, weathering, rainsplash, fluvial erosion, plastic and viscous flows, and debris slides, the general form of which are in line with the parameterizations given above. He used this model in two dimensional modeling of terrain to demonstrate the effects of geological inhomogeneity. As previously noted he had also examined characteristic profiles for hillslopes.

The studies of Cordova, et al. and Roth, et al. modelled solely fluvial transport. They modelled flow, via Mannings equation, and elevation evolution, via continuity of sediment transport. The transport law used was the Einstein—Brown equation (see Section 5.5 for a discussion of fluvial transport laws). The stated intention of this simulation model was to produce channel networks. The model, however, does not explicitly model the development of channel—like features. For instance, it does not model the concentration of flow along preferred directions characteristic of channel networks. The "channel network" is determined a posteriori from an analysis of elevations and flow directions. That channel networks can be inferred is a direct consequence of the interpretation of drainage directions; each node has many nodes flowing into it, and only one flowing out of it. Because the model considers the

surface, and all derived properties, to be smooth, the model could be better described as a hillslope model. At the scale of the hillslope, a channel looks more like a line or point, a dirac function. To model this requires a much greater concentration of flow than the Cordova model can provide. That a Dirac function represents a channel's properties better than the smooth representation of preferred drainage directions and flow used by Cordova is a central tenet of the model developed in Chapter 5. It will be the preferential erosion in the channel, represented as line, that causes the convergence of flow in the hillslopes around the channel network.

This concludes the consideration of time dependent processes and their effect on the form of hillslopes. It has been shown that a wide diversity of governing processes may be parameterized in a simple power law dependent solely on discharge and local slope. The promise of computer simulation has been highlighted, particularly in the case of two–dimensional flow problems. The importance of the channel–hillslope interactions on hillslope form was also pointed out.

2.3.4 Fractal Characterization of Landscape and Hillslopes

As previously noted, interest in fractal characterization of the landscape has increased in recent years. This interest has largely been a result of the stunning landscape pictures of Mandelbrot (1983) and others. The discussion in the previous section of the governing processes on the hillslope provides the groundwork for an examination of the research about fractals in the landscape. The mechanics of estimation of the fractal dimension of data will not be treated here. A treatment of this may be found in Tarboton, et al. (1988) or Mandelbrot (1983).

Ahnert (1984), Culling (1986), and Culling and Datko (1987) have studied the fractal characterization of landscape. Ahnert found that the cumulative elevation was fractal with a dimension between 0.5 and 1.0. This corresponds approximately to a fractal dimension for elevations along a transect of between 1.5 and 2.0 and for

70

elevation over an area of between 2.5 and 3.0. Culling (1986) and Culling and Datko (1987) found the fractal dimensions of transects of landscape to be between 1.2 and 1.5. Chase (1988) demonstrated, with a landscape evolution model, that depending on the dominant erosional regime the landscape areal fractal dimension can vary between 2.2 and 2.9.

Culling and Datko noted a change in the fractal dimension for horizontal scales less than 400m. For length scales less than 400m, the estimated fractal dimension was about 0.1 higher than the fractal dimension for length scales larger than 400m. Thus at low length scales, the landscape was qualitatively rougher than at longer length scales.

Davis, et al. (1988) also noted a break in the fractal dimension at a horizontal length scale around 500m. In their case they found the fractal dimension to be lower at short length scales. This result contradicts the findings of Culling and Datko. However, Davis, et al. obtained their elevation data from contour maps. A possible explanation is that this horizontal length scale corresponds to that length below which significant smoothing of point elevations occurs due to the linear interpolation between contours.

Culling (1986) attempts to explain the fractal nature of the landscape in terms of Gaussian random fields. These landscapes were called "diffusion degradation regimes" and are modelled by a diffusion equation for elevation; that is

$$\frac{\partial z}{\partial t} = D_z \frac{\partial^2 z}{\partial x^2}$$

The previous section noted that this equation adequately models a number of landscape forming processes (e.g., creep, rockfall, rainsplash).

A number of assumptions in Culling's analysis invalidate his conclusion for

scales less than hillslope scale. He assumes mean slopes are zero; a satisfactory assumption at longer length scales than the hillslope, but not satisfactory at the hillslope scale. He assumes slope increments (i.e., the slope over the ruler length used for determination of the fractal dimension; see Tarboton, et al., 1988) are independent. This is clearly incorrect at the hillslope scale since sediment transport and flow continuity require that the slope increments be correlated, otherwise continuity cannot be maintained.

Culling appears to recognize this limitation on his results to scales greater than the hillslope by noting that "A transect taken across a landscape can be divided into a series of interfluves." (page 236)

In conclusion, it appears that natural landscape seems to have a fractal dimension of between 2.0 and 2.5, for scales greater than the hillslope. The question of whether a fractal dimension exists, either the same or different, at horizontal scales less than the hillslope scale appears to be an open question.

2.4 Coupled Hillslope and Channel Evolution

The models presented above for channel network evolution and hillslope evolution have an important feature in common. They ignore the coupling between the hillslope and the channel. An important motivation of this work is to develop a realistic model for the evolution of both the channel network and the hillslope; that is, recognizing the unity of the catchment. In this context a number of important questions arise that in previous work could be simply ignored or glossed over.

The first question is that the interest is in a model consisting of two states (i.e. hillslope and channel), what constitutes a channel? This may seem a somewhat vacuous question and indeed in many cases the distinction between channel and hillslope is obvious. For instance, a river bed (i.e., channel) is a drainage path where depths of flow are large, velocities high, and the flow is typically bounded by steep

banks. However, near the source the distinction may be less obvious. The well defined channel may become a series of poorly defined depressions and springs (e.g., McHugh and Prestergaard, 1988). Alternatively, it could be a well defined channel head in a fixed location (e.g., Coelho–Netto, et al., 1987). This distinction between channel and hillslopes <u>at the channel head</u> is crucial since it is the position of the channel heads that determines the drainage density of the catchment.

These considerations lead to the more difficult question of what are the physical processes necessary to form a channel, and what governs the partition between the hillslope and the channel. Dunne (1989) suggests a subsurface saturation criteria dependent on the groundwater flow head gradient. Montgomery and Dietrich (1988) have noted a relationship between source area and the local hillslope slope at the channel heads in a number of catchments in California.

The question of what is "the" channel network is further complicated by strong transient effects at short timescales. The portion of the channel network flowing varies with a number of climatic variables and varies from storm to storm and within storms (Gardiner and Gregory, 1981). For a long–term geomorphology model, it is necessary to average out these short–term effects. It is necessary to consider the average wetted length of the channel network. If the geomorphology results largely from the sediment transport in flood events, then the effective network is the average wetted length of channel network during these geomorphologically effective events; the wetted length between flood events is largely irrelevant since dry periods do not provide significant erosive potential.

The average wetted length of channel can be determined from consideration of the different types of hillslope processes. For overland flow, Horton postulated that channels form when the tractive shear stress exceeded a threshold. The more intense the rainfall, the shorter will be the hillslope length. The average hillslope length will then be averaged over the distribution of geomorphologically effective events. For

73

subsurface saturation runoff a subsurface saturation criteria, such as developed by Beven or O'Loughlin, may be used. As in the case of overland flow, the region of saturation (i.e., the channel) will increase with increased rainfall. Again averaging is necessary to determine the mean channel network. In addition the effectiveness of flood events needs to be considered. Since sediment transport is nonlinear with runoff, large runoff events are disproportionately more important than small runoff events in determining the landscape form. These major runoff events should be weighted more than the minor runoff events. Thus there are two ways of defining the average channel network which potentially yield different results.

1. Where runoff is the important issue, the network during runoff events is averaged, weighting by the amount of runoff.

2. Where geomorphology is the important issue, the network during geomorphologically effective runoff events is averaged, weighting by the effectiveness of each event in sculpting the landscape form.

This idea of weighting on the basis of geomorphologic effectiveness dates back to Wolman and Miller (1960). The application of this idea to determination of the average channel network is obvious conceptually, but nontrivial practically.

A field orientated approach is to develop a measure of the mean hillslope length and thus the drainage density. An example would be the fractal measures of the previous section. The network that is found will be the geomorphologically effective network rather than the effective network for runoff. Unfortunately this technique does not answer the question of what are the physical runoff processes that create the channel network, so is of limited usefulness in designing a simulation model.

Another important question is how to distinguish between rills and channels, if there is a distinction, and what physical process defines the distinction if there is.. Horton (1945) in his explanation of channel formation from overland flow did not discuss rilling. Smith and Bretherton (1972) note that for overland flow when

sediment transport occurs, then channel/rill formation must occur. Neither view is realistic; rills do occur in some cases, and hillslope sheet erosion has been observed without rill formation. It was previously noted that rainsplash can suppress the formation of rills by diffusive mechanisms. Channel formation occurs when fluvial transport dominates the diffusive rainsplash effects. Another possible explanation is that channel formation is suppressed by groundcover vegetation (e.g., grasses). Only when the shear stresses are high enough to disrupt this mat and penetrate to the friable underlying soil will channel formation begin. Either way channel formation is controlled by more than simple shear stress or flow velocity thresholds. The controlling process for channel head development will thus result from the interaction of a number of poorly understood processes.

Another important issue for a coupled hillslope—channel model is determining how fast a channel grows when the prevailing channel network is out of equilibrium with the landscape and climate. This disequilibrium may occur when the network is growing initially, under changes in runoff conditions (e.g., urbanization, climatic changes), or changes in geologic conditions (e.g., after an uplift event). The literature is notably silent on this issue. Existing discussions of channel growth are purely qualitative with little connection to physical principles (e.g., Schumm, 1956; Morisawa, 1964).

In conclusion, some of the important issues encountered in the coupling of the hillslope and channel have been discussed. The most significant problem is how to distinguish between channels and hillslopes. Temporal averaging is necessary because of the variation of wetted channel length over time. The concept of the geomorphologically effective channel network has been introduced. This geomorphologically effective network may be different from the effective channel network active in runoff processes. Questions about channel growth rates were raised. Many of the ideas presented here will form the basis of the coupled channel

network–hillslope model developed in Chapter 5 and later chapters. Some of these same concepts will also arise in discussion of the nonhydrologic network models discussed in Chapters 3 and 4.

CHAPTER 3

REVIEW OF NON–HYDROLOGIC NETWORK MODELS

3.1 Introduction

Branched networks are ubiquitous in nature. They occur in metal solidification, polymer growth, two–phase flow in groundwater, leaf veins, capillaries in lungs and, of course, in the branches and roots of trees, from where they obtained their name.

That networks are so common, over such a broad range of physical phenomena, suggests that there must also be a qualitative similarity in the mechanisms that control the growth of these networks. Certainly the qualitative form of the networks generated in each case is similar. For instance, there is a partitioning of the domain into two regimes; one regime representing the network or aggregate, the other regime representing a substrate from where the network developed. All of these phenomena form networks that, outwardly at least, appear random in nature. Many of the networks provide a transport capacity for the substrate from the tips of the network to the root. It therefore seems logical to suggest that the governing processes must be qualitatively similar.

In this chapter a number of physically based models that produce networks will be briefly discussed. It will be shown that the governing physics for all these models presented have qualitative, even quantitative similarities. This chapter cannot go into details of these models, instead their essence will be presented.

The search for qualitative characteristics of the network formation process that are essential to network generation will begin in this chapter, and be completed in the next chapter. In the hydrologic setting the use of topological and qualitative characteristics of networks has been extremely useful, as noted in Chapter 2. Similar studies of the dominant qualitative characteristics of general network growth will similarly be useful. It is believed that the essential topological requirements for

network growth discovered here are preserved in the physically based model of river network growth that will be presented in Chapter 5.

3.2 Diffusion Limited Aggregation (DLA) Models

The concept of diffusion limited aggregation (DLA) originated with Witten and Sanders (1981). It attempts to model the development of branched clusters of material. A related technique is cluster–cluster aggregation. Both DLA and cluster–cluster aggregation form fractal clusters. However, only DLA forms clusters that appear network like; the discussion that follows is thus restricted to DLA.

DLA works in the following fashion. A grid is defined (typically triangular, square, or hexagonal) and atoms are randomly placed at the nodes of the grid, uniformly in space. This is the initial condition. At each time step, each atom is given an independent (both in space and time) random perturbation to an adjacent node. If the adjacent node is empty, that particle moves to the adjacent node, if not then collision rules are invoked to decide the subsequent motion of both of the colliding particles. With correct choice of collision rules global momentum is conserved. This description is one of a standard cellular automata (Rothman and Gunstenson, 1988). In DLA, one particular atom is chosen to be a seed, and its position, in time, fixed. If another atom hits the seed, or hits an atom connected to that seed, it sticks and becomes part of the stationary aggregate. Thus, with time, a network–like aggregate will grow (Figure 3.1).

This process is diffusive and is a discrete simulation solution to the two–dimensional isotropic diffusion equation (Witten and Sanders, 1981)

$$\frac{\partial c}{\partial t} + D\left[\frac{\partial^2 c}{\partial x^2} + \frac{\partial^2 c}{\partial y^2}\right] = 0 \qquad (3.1)$$
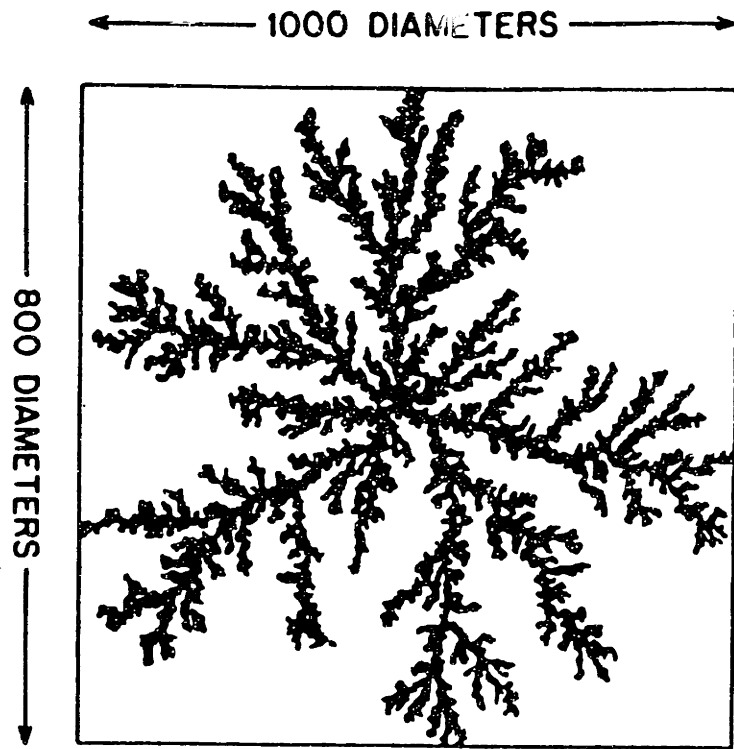
Figure 3.1: Example of network generated using DLA.
(from Meakin, 1986)

where   D   = diffusivity, related to the rate of movement of particles

             c   = average concentration of particles/unit area.

with the initial conditions

$$c(0) = c_0 = \frac{N}{XY}$$

where   N   = initial number of atoms

          X, Y   = dimensions of the grid

A first type boundary condition, fixed concentration, at the edges may be applied by embedding the solution domain within an even larger domain. In the intervening region the number of atoms is kept constant, simulating an infinite store at the fixed concentration.

The network grows out into the substrate region. The network grows by atoms sticking to the network, referred to as the aggregate. It is important to note that atoms, once stuck to the network, never return to the substrate. Because of this, the network can be considered to be a constant concentration boundary condition to the substrate region (i.e. $c=0$ along the network); the concentration is zero at the interface with the network (Ball, 1986). Within the network the concentration is actually $c = 1$ (1 atom/site) but as far as the substrate is concerned is only sees that the transport of particles from the network back into the substrate is zero, which can only be modelled by a $c = 0$ boundary condition. This inconsistency between the actual concentration in the network ($c = 1$) and the concentration seen by the substrate ($c = 0$) is a conceptual problem with the way that boundary conditions are applied in the model rather then in the transport processes within the substrate region. So the governing equations are effectively

$$\frac{\partial c}{\partial t} = D \left[\frac{\partial^2 c}{\partial x^2} + \frac{\partial^2 c}{\partial y^2} \right] \qquad : \text{in the substrate}$$

$$c = 0 \qquad\qquad\qquad : \text{in the network}$$

As the network grows, then this boundary condition at the network on the substrate region also grows.

Since most transport processes are driven by concentration gradients the zero concentration in the network means that DLA is unable to model transport processes within the network. This is a major failing of the DLA model if an analogy is sought between the substrate concentration and the catchment elevation, and between the aggregate and the channel network. DLA would model the channels as flat. Energy is expended when water flows downhill. The DLA models, as described, cannot model the balance of potential energy expenditure for flow and the sediment transport, partitioned between the hillslope and the channel. Because DLA models the "channels" as flat then all the energy expenditure occurs in the hillslopes.

An important characteristic of network growth with DLA that has recently been identified (Stanley, 1986) is the property of network growth site screening. Loosely speaking, screening is the capability of the existing network to bias the growth in the various parts of the network. In DLA, the existence of the network surrounding a potential growth site suppresses growth there. Fastest growth occurs in the extremities of the network. In addition, Stanley proposed that the fractal characteristics of any segment will be modified by the existence of the network around it. The net result of this is that the fractal dimensions of the final network will vary depending upon where and under what screening conditions that network branch grew; thus the time history of network is crucial to describing the final form of the network. A mathematical description of how screening governs the growth rate follows.

Network growth in DLA occurs by the transport of particles across the interface

between the substrate and the aggregate. The more particles that are transported across the interface, the faster the network grows. DLA models this transport in a discrete fashion with the aggregate/network growing by 1 particle when 1 particle crosses the interface from the substrate to the aggregate and sticks to the aggregate. In a continuum form, this transport across the interface can be modelled as

$$R = T = D \frac{\partial c}{\partial n}$$

where   T   = transport/unit width

n   = direction perpendicular to the aggregate interface with the substrate

R   = rate of growth/unit width of the aggregate perpendicular to the interface.

This equation gives the rate of growth of a unit width of the interface. The growth rate for the total aggregate or network follows directly and is

$$R_{aggregate} = \oint_{\underline{A}} D \frac{\partial c}{\partial n} ds$$

where   s   = direction parallel to interface

$\underline{A}$ = the interface of the aggregate

Thus the relative rate of growth of some unit length of the aggregate interface compared to the rate of growth of the total aggregate is

$$R' = \frac{R}{R_{aggregate}} = \frac{\frac{\partial c}{\partial n}}{\oint_{\underline{A}} \frac{\partial c}{\partial n} ds} \tag{3.2}$$

This expression is equivalent to the probability that, in the discrete DLA simulation, the interface will grow by one unit or one atom in a unit width in a unit period of time. Since $\frac{\partial c}{\partial n}$ varies throughout the domain, the rate of growth of network at various points in the network will also vary accordingly.

As previously mentioned in DLA, the network effectively acts as a zero–concentration boundary condition. In addition, because the network grows outward from the initial seed or root of the network, concentrations of substrate are lowest near the root of the network. This is because the zero concentration boundary conditions have greater time to diffuse outward from the network near the root of the network than they have had near the extremities. The net effect of this is that $\frac{\partial c}{\partial n}$ is highest at the extremities of the network so that the extremities of the network will grow faster than the interior of the network. This is the essence of screening; the existence of the network around growth sites near the root of the network reduces the grow rate there. Alternatively if growth is viewed in the DLA sense as occurring an atom at a time, the probability of growth is highest near the extremities of the network.

If we return to the potential analogy between DLA and the catchment where concentration is considered analogous to elevation, then $\frac{\partial c}{\partial n}$ is the slope in the hillslope perpendicular to the channel. Channel growth occurs proportionally to the hillslope slope. The concentration within the network is analogous to the elevation of the channel (considered fixed by DLA). The transport within the substrate region is of sediment, which is analogous to reduction of elevation. The most important thing to note, however, is the existence of the screening effect on network growth. Since the screening effect on channel network growth has never been experimentally measured, it can only be assessed indirectly. For instance, the observation that drainage density is constant spatially (Abrahams, 1984) suggests the existence of some screening effects.

This might be contrasted for instance with the case of the Howard (1971) network headward growth model described in Section 2.2.2.2. Superficially the Howard headward growth model appears similar to DLA. Growth occurs in generations with each generation representing the growth of a single link connecting two node points; this can be considered analogous to DLA's growth one atom at a time. However, Howard's model displays deep dissimilarities in the mechanisms that govern the network growth rate. Howard assigns each legal growth site an equal probability of growth; nodes at the root of the network have an equal probability of growth as nodes at the extremities of the network; there is no sort of screening. As noted in Section 2.2.2.2, this means that at intermediate times the network drainage density is high around the root of the network and low at the extremities. Screening effects will be returned to in the following chapters and will assume some significance in the interpretation of experimental observations by Montgomery and Dietrich (1988) in Chapter 8.

Consider a variant of the classical DLA problem, where concentration is not zero _within_ the network, but varies from zero at the root of the network to some concentration c at the network extremities (assume c is less than the mean substrate concentration). In this case if transport within the network is driven by a concentration gradient, transport can be modelled. Let us now compare $\frac{\partial c}{\partial n}$ in the substrate with that of the classical DLA solution. In the new situation the differences between the concentration in the network and concentration in the substrate at the extremities of the network are reduced. At the root since the network concentrations are essentially unchanged the concentration gradient at the network, $\frac{\partial c}{\partial n}$, is unchanged. Compared with the classical DLA solution $\frac{\partial c}{\partial n}$ at the network extremities is reduced relative to the gradient at the network root. Thus in the new case the growth rate at the extremities of the network is reduced, while that at the network root will be unchanged. In a relative sense (i.e. Equation 3.2) network growth occurs faster at the

network root than at the extremities — the screening effect of the network is reduced.

The importance of the screening property cannot be overstressed. It is through screening that the spatial distribution of network growth is controlled. The functional form of the screening will govern how fast the extremities of the network grow compared with other parts of the network. These differences in the relative growth rate of different parts of the network govern the distribution of the drainage density of the network. And if Stanley (1986) is correct it will also govern the fractal characteristics of the channel network. In addition, it should be noted that screening results from the interaction of the network and the substrate. The substrate gradients give the screening effect, but the substrate gradients result from the interaction of the concentration boundary condition at the network and the substrate concentrations.

Thus there are a number of problems with using the standard DLA model to simulate network growth in a catchment. These problems are:

1. The fixed concentration boundary condition at the aggregate means that transport within the network cannot be modelled. Thus energy expenditure cannot occur in the channels; a DLA model of catchment erosion would force all erosional energy expenditure to occur on the hillslope. The distribution of growth sites in the network is also biased towards growth at network extremities.

2. DLA, as currently implemented by a cellular automata, is restricted to linear diffusion. Diffusivities may vary in space and time (Toffoli and Margolus, 1987, for instance, model refraction by using a spatially variable diffusivity) but the diffusivity cannot depend upon the solution. Thus nonlinear diffusion, with diffusivity dependent upon concentration, cannot be modelled. Appendix B shows that sediment transport is modelled by a nonlinear diffusive process.

Neither of these problems is essential to the DLA conceptualization, but both

arise out of the current computational techniques for cellular automata.

Despite these difficulties DLA has been seen as a useful model for a number of physical process that create networks including diffusion limited deposition of gaseous metals on surfaces, solidification of metals (Ball, 1986), growth of polymers from the basic monomer (Daoud, 1986), floculation and gelation (Kolb, et al., 1986) and dielectric breakdown (e.g. lightning). The solidification processes, for instance, results from modeling the heat equation (of the form of Equation 3.1), where the diffusivity is a function of the conductivity and specific heat of the material.

3.3 Viscous Fingering in Porous Media

This section will briefly discuss aspects of viscous fingering (Figure 3.2). It is commmonly observed that network–like fingering occurs where

1.   a less viscous fluid is injected into a more viscous fluid

2.   a fluid is injected or infiltrates into a unsaturated media.

Similarities exist in the governing equations for fluid flow in a porous media in the above cases and the DLA solution technique described in the previous section. Darcy's law for flow in a porous media is (in tensor notation).

$$q_j = K_{ij} \frac{\partial h}{\partial x_i} \qquad (3.3)$$

where q_j = specific discharge in direction j

$\frac{\partial h}{\partial x_i}$ = hydraulic gradient in direction i

K_{ij} = conductivity tensor

so that it follows that

$$S_s \frac{\partial h}{\partial t} = \frac{\partial}{\partial x_i} \left[K_{ij} \frac{\partial h}{\partial x_j} \right] \qquad (3.4)$$
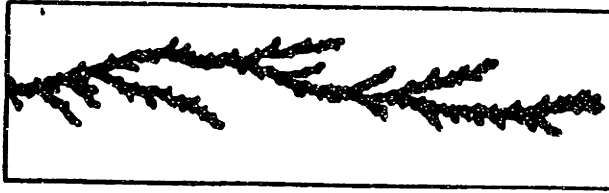
Typical viscous fingers created by water advancing into a linear Hele Shaw cell filled with a polymer solution (scleroglucan).
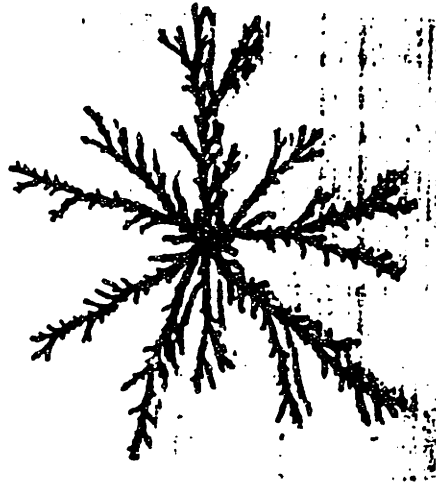
Typical viscous fingers created by water advancing into a radial Hele Shaw cell filled with a polymer solution (scleroglucan).

Figure 3.2: Examples of viscous fingering:
a less viscous fluid is injected into a
more viscous background fluid
(from Nittmann, et al, 1986)

where S_s = storativity of the aquifer

For a constant conductivity this gives

$$\frac{\partial h}{\partial t} = \frac{K_{ij}}{S_s} \frac{\partial^2 h}{\partial x_i \partial x_j} \tag{3.5}$$

The similarity of Equations (3.4) and (3.5) with the governing equation for DLA, Equation (3.1), should be noted where the head, h, above is analogous with the concentration in DLA.

If a very low viscosity miscible fluid is injected into a high viscosity background fluid then it has been shown that fractal networks are generated (Nittmann, et al., 1985). The governing equations for the two phases are

$$\frac{\partial h}{\partial t} = D_{ij} \frac{\partial^2 h}{\partial x_i \partial x_j} \qquad \text{: injected fluid}$$

$$\frac{\partial h}{\partial t} = D'_{ij} \frac{\partial^2 h}{\partial x_i \partial x_j} \qquad \text{: background fluid}$$

where $\quad \| D_{ij} \| << \| D'_{ij} \|$

If the fluid flow resistance is low then D_{ij} is high. For instance, in a Hele–Shaw flow apparatus, as experimentally studied by Nittmann, et al. (1985, 1986) D_{ij} is inversely proportional to viscosity. In the groundwater case D'_{ij} it is just the conductivity K_{ij} again inversely proportional to viscosity. In the limit as the viscosity ratio tends to infinity, and for perfectly miscible fluids, this is analogous to the DLA model (Ball, 1986). In this case the pressure drop in the injected fluid (here the a   te or

network) from the injection point to the interface with the background fluid (the substrate) is negligible in comparison with the pressure drop in the background fluid because the viscosity of the injected fluid is much smaller. For DLA to be an analogy perfect miscibility of the fluids is required so that there are no capillary pressures at the interface to remove small scale irregularities at the interface.

In the terminology of the DLA discussion of the last section, the head variation along the network or injected fluid is virtually zero so that the network may be considered to be a constant head, but moving boundary condition, to the head in the background fluid or substrate. Thus, the validity of the DLA approximation is a function of the viscosity ratios, and is measured by how much the boundary head varies within the network compared to the background fluid; as the viscosity ratio drops the head variation will be increased and the DLA approximation becomes less valid. As noted in the previous section the distribution of growth sites (i.e. viscous fingers) will also depend on the amount of head variation within the aggregate due to the screening effects on head.

Invasion of an unsaturated porous media by a wetting fluid will result in qualitatively similar, if not quantitatively identical, patterns. In the saturated region behind the wetting front the conductivity K_{ij} is much higher ($\|K_{ij}\|_{invading} \gg \|K_{ij}\|_{background}$), than in front of the wetting front. The analogy with the low viscosity invading fluid case is direct.

In conclusion, the important points to note from this section are:

1. The similarity of the porous media governing equations and those for the DLA problem. Both involve gradient driven transport within the background substrate (concentration in DLA, head in groundwater).

2. When the ratio of the conductivity in the porous media for the invading fluid to the background fluid is infinite groundwater flow is modelled by DLA. These large conductivity differences can be caused by variable

viscosity or saturation.

3. Where the differences between the invading and background fluids are not so strong head variation within the invading fluid may be of the same order as in the background fluid. Because head variations within the network are not negligible screening effects will be different to those exhibited by DLA. Network characteristics should be different from those exhibited by DLA.

3.4 Mitchison Leaf Vein Model

Mitchison (1980) proposed a model for chemical transport in leaf veins that exhibits qualitative similarities to Equation (3.3) for transport by viscous fingering. The flux of auxin (a growth chemical) is modelled by

$$\varphi = D \frac{\Delta c}{\Delta x} \tag{3.6}$$

where
φ = auxin flux

D = diffusivity

Δc = change in auxin concentration between two adjacent cells

Δx = length of the cell

For vein formation to occur Mitchison notes that $\frac{|\varphi|}{D}$ must decrease as $|\varphi|$ increases and suggests that suitable formulations for this dependency are of the form

$$D \sim \varphi^n \quad ; n = 2 \tag{3.7}$$

A vein is "identified" as being a pathway in which the flux of auxin is significantly greater than in the rest of the leaf.

The governing equation for the leaf is, in differential form

$$\frac{\partial}{\partial x} \left[D(\varphi) \frac{\partial c}{\partial x} \right] = 0 \qquad\qquad (3.8)$$

This is the steady state analog of Equation (3.4). Combining Equations (3.6) and (3.7), the diffusivity in Equation (3.8) can be expressed as a function of concentration

$$D(\varphi) \sim D\left(\left(\frac{\partial c}{\partial x}\right)^{-1} \right)$$

which clearly shows that Equation (3.8) is nonlinear diffusion in the concentration. Since $D(\varphi)$ is variable in space, both within the leaf vein and within the background substrate, then the equations are qualitatively but not quantitatively analogous to DLA and two phase porous media flow.

The vein formation process proceeds as follows. As $\frac{\partial c}{\partial x}$ decreases, the diffusivity increases which further reduces the gradient. In Equation (3 6) the transport in increased by these interactions. Thus preferred transport paths are created in which $\frac{\partial c}{\partial x}$ is low and D is high so that chemical transport is concentrated along these paths. It is important to note that there exists a gradient of concentration along these paths. Thus if the transport paths, here considered to be the network, are considered to be fixed concentration boundary conditions with time on the remainder of the domain, the applied concentration will be variable in space, lowest on the root, highest at the branch tips. Again we note the difference with the DLA boundary condition at the network of constant concentration. The higher the diffusivity in the preferred transport path, or the greater the nonlinearity of Equation (3.7), the more valid the DLA approximation will be because the concentration gradient will be lower for equivalent fluxes. Mitchison showed that networks of preferred drainage paths were created, but did not analyze them. The equations the investigator proposed show

strong similarities to those of both DLA and viscous fingering problem.

3.5 Conclusions

The models presented in this chapter exhibit a number of qualitative and quantitative similarities. Most importantly they all generate networks in space where there are two phases, one phase being the network itself and the other phase being the surrounding regions of substrate material. This chapter has attempted to quantify the similarities on the basis of the governing equations of the physical processes involved. It has been shown that these models, though from different fields of research, are variants of the same problem.

Important points in the preceding discussions are:

1. All the models simulate some distributed property (e.g., concentration, pressure) using either a linear or nonlinear diffusion process.

2. All the models are autocatalytic in the transport of this distributed property. Autocatalysis is a form of positive feedback where a small change in the property results in, through interactions in the nonlinear system, an amplification of that change. In DLA, though the governing equations in the substrate and the network are linear, nonlinearity arose from the differentiation in transport processes between the network and the background substrate. This property exhibits itself as preferred transport pathways for the modelled properties.

3. In all models the generated network may be considered to be nearly a constant concentration boundary condition on the substrate part of the domain. In the case of groundwater flow and leaf vein growth, the head and concentrations (respectively) in the network result from the interaction of the transport processes in the network and the substrate. In DLA the concentration is zero in the network.

4. All models exhibit some form of growth site screening governed by the interaction between the network and the surrounding substrate. The detailed physics of the screening process is different for each model, but the concept that certain parts of the network are more likely to grow or that parts of the network grow faster than others, and that these rates are determined by substrate characteristics, is common to all models. The distribution of drainage density during active network growth is strongly influenced by the screening properties of the model.

With the exception of DLA, there is no explicit differentiation made between the areas of preferred transport (the network) and the remainder of the region; in other methods the network must be interpreted a posteriori in a somewhat arbitrary fashion (e.g., transport rate above a threshold). It is the intention of this work to construct an analogy between the preferred transport paths and the channel network and an analogy between the remainder of the domain and the hillslopes. Channels, after all, are just preferred transport paths for water and sediment. A means to make this differentiation explicit, rather than implicit, is desired. The next chapter will examine another non-hydrologic network model. This model displays deep similarities to the models in this chapter, yet makes the differentiation between network and substrate, and channel and hillslope, explicitly. In addition, its simple and explicit physics makes it easy to examine some fundamental theoretical characteristics of network generation in physical models.

CHAPTER 4

THE MEINHARDT LEAF VEIN MODEL

4.1 Introduction

Building on the ideas of the previous chapter we present a model first introduced by Meinhardt (1976), and further explored by Meinhardt (1982), that simulates the growth of leaf veins. These vein cells demonstrate network–like patterns within the simulated leaf.

The system of differential equation that Meinhardt (1976) presented were chosen for detailed study for a number of reasons:

1. The model makes an explicit differentiation between two states in the system; leaf vein cells and normal leaf cells. The leaf vein cells form a network pattern in space. A qualitative analog can be seen between this and the drainage basin where the leaf veins might be considered channels, and leaf cells the surrounding hillslope.

2. The leaf vein networks grow headward in time in a similar fashion to channel networks.

3. The growth process is governed by physical interactions, following directly from the system of differential equations. The simplicity of the governing equations provides the opportunity for deeper understanding of the network growth processes; even if only qualitatively.

This chapter will first present, and discuss, the equations of Meinhardt (1976). The complicated nonlinear interactions between the various components of the equation will be discussed, always searching to classify, simplify and generalize the network producing behavior.

It will be demonstrated that there are three qualitative characteristics of the equations that create the differentiated pattern of leaf vein networks. Using sensitivity

studies it will be argued that these three characteristics are not specific to Meinhardt's equations, but are, in fact, powerful prerequisites of any network growth model based on physical mechanisms.

The chapter will conclude with a discussion of the random properties displayed by networks. Using concepts from chaos theory it will be shown how the apparently random properties follow from the modelled physics. Some tentative hydrologic analogies will be made with the three necessary conditions for network formation referred to above. Consequently it will be suggested that the chaotic behavior observed in the Meinhardt system of equations is a property common to many systems that exhibit networking and that this may be an explanation for why channel networks appear random.

4.2 The Governing Equations

Meinhardt (1976, 1982) presented a system of partial differential equations that simulate the growth of leaf veins. A slightly more general form of the equations will be discussed to show the generality of the qualitative growth mechanisms of the system. These equations, hereafter called the Meinhardt equations, are:

$$\frac{\partial a}{\partial t} = \frac{c \, a^{m_1} z}{h} - \mu a + D_a \frac{\partial^2 a}{\partial x_i^2} + \rho_0 Y \tag{4.1a}$$

$$\frac{\partial h}{\partial t} = c_3 \left[c_2 \, ca^{m_1} z - \nu h + D_h \frac{\partial^2 h}{\partial x_i^2} + \rho_1 Y \right] \tag{4.1b}$$

$$\frac{\partial z}{\partial t} = c_0 - \gamma \, c_0 z - \xi \, Y z^{m_2} + D_z \frac{\partial^2 z}{\partial x_i^2} \tag{4.1c}$$

95

$$\frac{\partial Y}{\partial t} = c_1 a - 0.1Y + \frac{Y^2}{1 + 9Y^2} \tag{4.1d}$$

Meinhardt solved these equations on a rectangular grid with sides L_{x_1} nodes by L_{x_2} nodes. The boundary conditions applied to the equations are

$$\frac{\partial a}{\partial x_1} = \frac{\partial h}{\partial x_1} = \frac{\partial z}{\partial x_1} = \frac{\partial Y}{\partial x_1} = 0 \quad \text{at } x_2 = 0,\ x_2 = L_{x_2}$$

$$\frac{\partial a}{\partial x_2} = \frac{\partial h}{\partial x_2} = \frac{\partial z}{\partial x_2} = \frac{\partial Y}{\partial x_2} = 0 \quad \text{at } x_1 = 0,\ x_1 = L_{x_1}.$$

These equations are identical to those of Meinhardt (1976) if

$$m_1 = 2; \quad c_2 = 1; \quad c_3 = 1; \quad \gamma = 1; \quad m_2 = 1$$

Of the four states modelled by these equations, three of them (a, h, z) model chemical transport processes in the leaf. The fourth state, Y, models the spatial pattern of leaf vein cells. From our perspective the latter state, differentiation, Y, is the most interesting. Differentiation will be conceptualized as the channelization in the catchment. In Meinhardt's work the four states represent:

1.   a: activator, high values of activator tend to start or trigger the differentiation process.

2.   h: inhibitor, high values of inhibitor tend to stop production of activator. Inhibitor is created by higher activator levels.

3.   z: substrate, the substrate is consumed by the differentiation process.

4.   Y: differentiation, a 0–1 variable that signifies the spatial pattern of

veins within the leaf.

Of the four states, differentiation is the most important. The other three states are useful simply because their interactions produce spatial patterns of differentiation that appear similar to channel networks. The exact physics of activator, inhibitor, and substrate will later be shown to be computationally and conceptually convenient but practically unimportant.

Differentiation, Y, simulates the spatial pattern of leaf veins. If $Y = 0$ at a point, then that point is a normal leaf cell. If $Y = 1$ at that point, then that point is a leaf vein cell. The Equations (4.1) ensure that the points where $Y \approx 1$ form a network—like pattern in space (Figure 4.1).

The network simulation process proceeds as follows: Initially activator, inhibitor and differentiation are small (approximately 0.0) everywhere and · the substrate is $z = 1$ everywhere. This represents an undifferentiated leaf, with a chemical substrate ready to be consumed. In addition, Meinhardt (1976) assumes that coefficient c is a random field so as to produce random networks. It will be shown later that this random field is unnecessary. At the start of the simulation a single point is chosen as the seed or starting point of network. Differentiation is set to $Y = 1$ at this point. The nonlinear dynamics of Equation (4.1) is then sufficient to produce a network of differentiated points, $Y = 1$, that grow headward with time eventually filling the domain (Figure 4.1).

The process by which a single point in space differentiates in time will now be explained. Initially the point will have a value for differentiation of $Y = 0$, i.e., undifferentiated. Finally, if the point lies on the network, it will have a value of differentiation of $Y = 1$, i.e., differentiated. At the time the point differentiates there will be a short period of transition where the value of differentiation is between 0 and 1. Once a point differentiates, it remains differentiated for all time; differentiation is a one—way process that cannot be reversed. The process of differentiation is triggered by
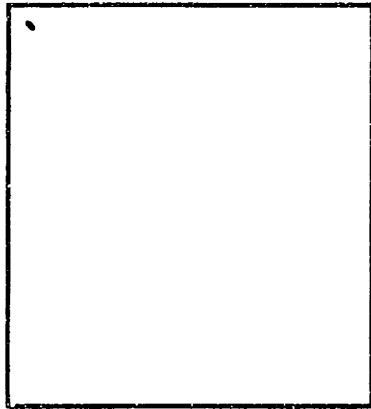
t = 0
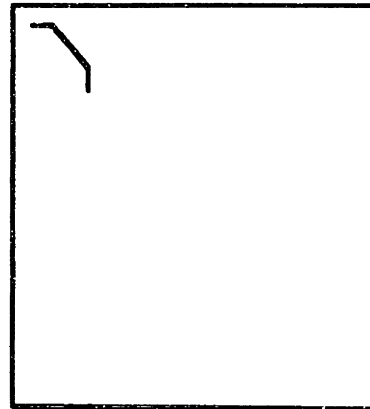
t = 1000
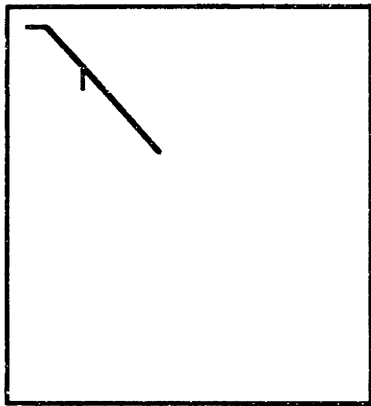
t = 2000
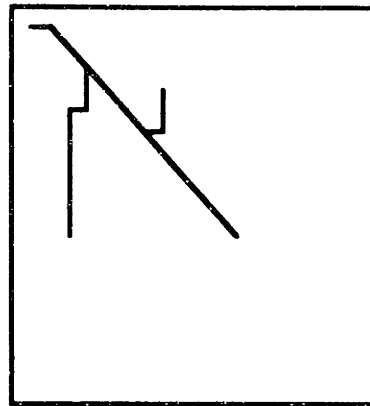
t = 3000
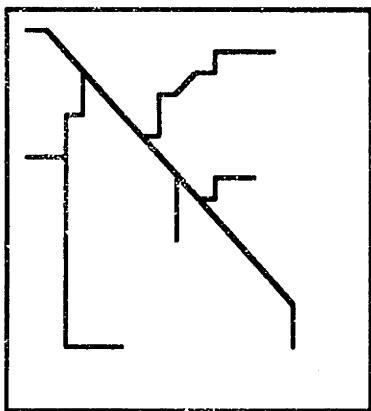
t = 4000
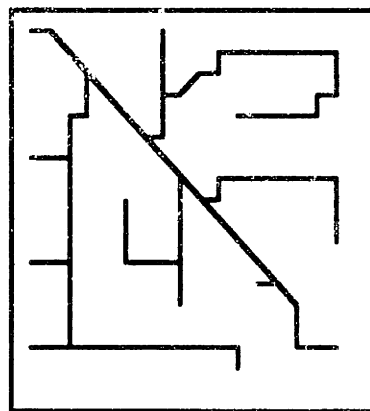
t = 9000

Figure 4.1: Sample Meinhardt network showing headward growth with time

activator exceeding a threshold. Inhibitor and substrate values are irrelevant here because they do not appear <u>directly</u> in the differentiation equation (Equation 4.1d). They are only important insofar as they produce high levels of activator. This differentiation process for a single point with time is shown in Figure 4.2.

How the differentiation process is controlled by the differentiation equation is demonstrated in Figure 4.3. Figure 4.3 shows the differentiation equation for three values of the activator term. The value of zero activator (Curve C) represents the situation either at very early time (when $Y = 0$) or at very late time (when $Y = 1$). There are three critical points (i.e. $\frac{dY}{dt} = 0$), two stable ($Y = 0$, $Y = 1$), and one unstable point ($Y = 0.1$). Points at $Y = 0$ or $Y = 1$ are stable against fluctuations in Y, or small fluctuations in activator. All values of Y are attracted to, and will eventually converge on, $Y = 0$, or $Y = 1$.

Curve A in Figure 4.3 shows the differentiation equation for a value of activator above the activator threshold. For the case where the activator term is greater than 0.0025 (Curve B) there is only one critical point, $Y = 1$. Values of Y are attracted to the value $Y = 1$. Thus an undifferentiated point ($Y = 0$) will tend to become differentiated (Y is attracted to 1). Once Y is greater than 0.1, the activator value may decline to zero, but the value of Y is now in the basin of attraction of $Y = 1$, and so that point will tend inevitably to differentiate.

Other than the activator–differentiation connection there are other well defined and important connections between Equations (4.1a), (4.1c) and (4.1d). These connections are shown in Figure 4.4. Some connections are not shown because they are of secondary importance. The three important connections to note are that the activator is an input to differentiation, differentiation is an input to substrate, and substrate is an input to the activator and inhibitor equations. Thus there are three conceptual components to Equation (4.1).

1.   A component that triggers differentiation (i.e., activator)
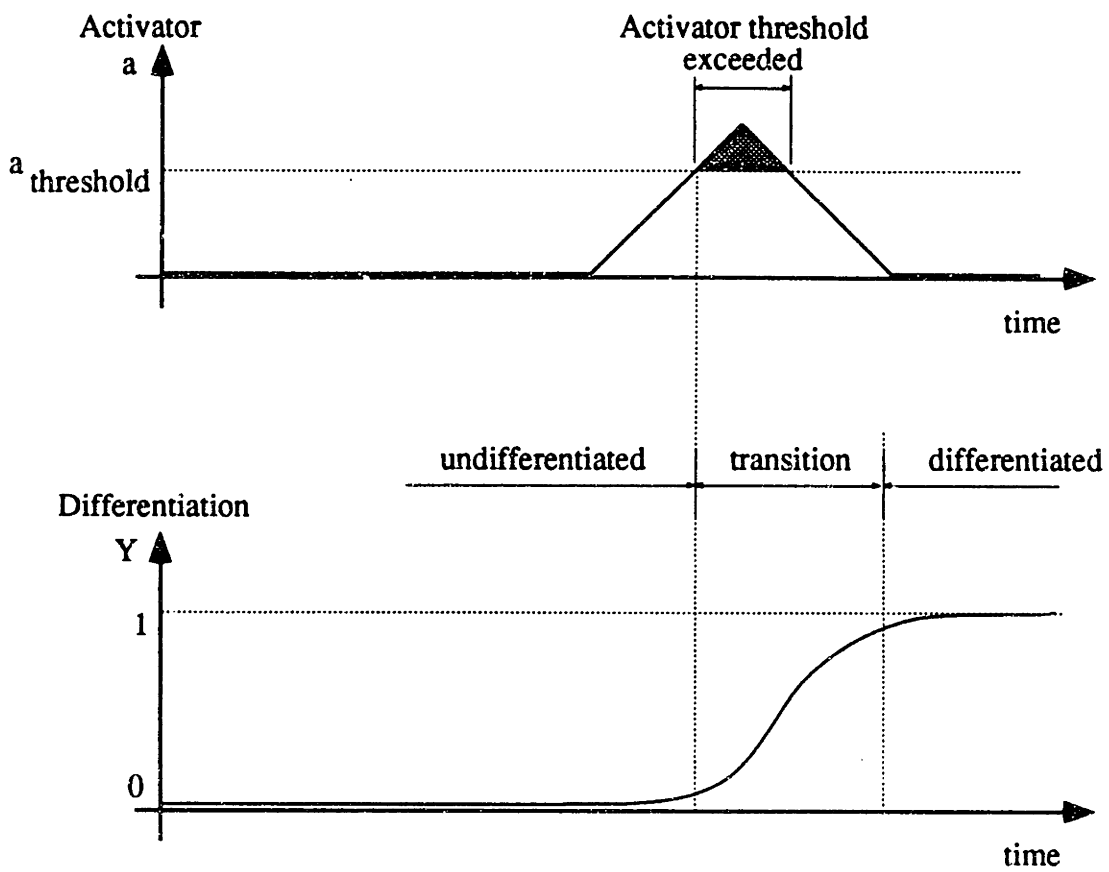
Figure 4.2: Schematic of Activator and Differentiation evolution with time
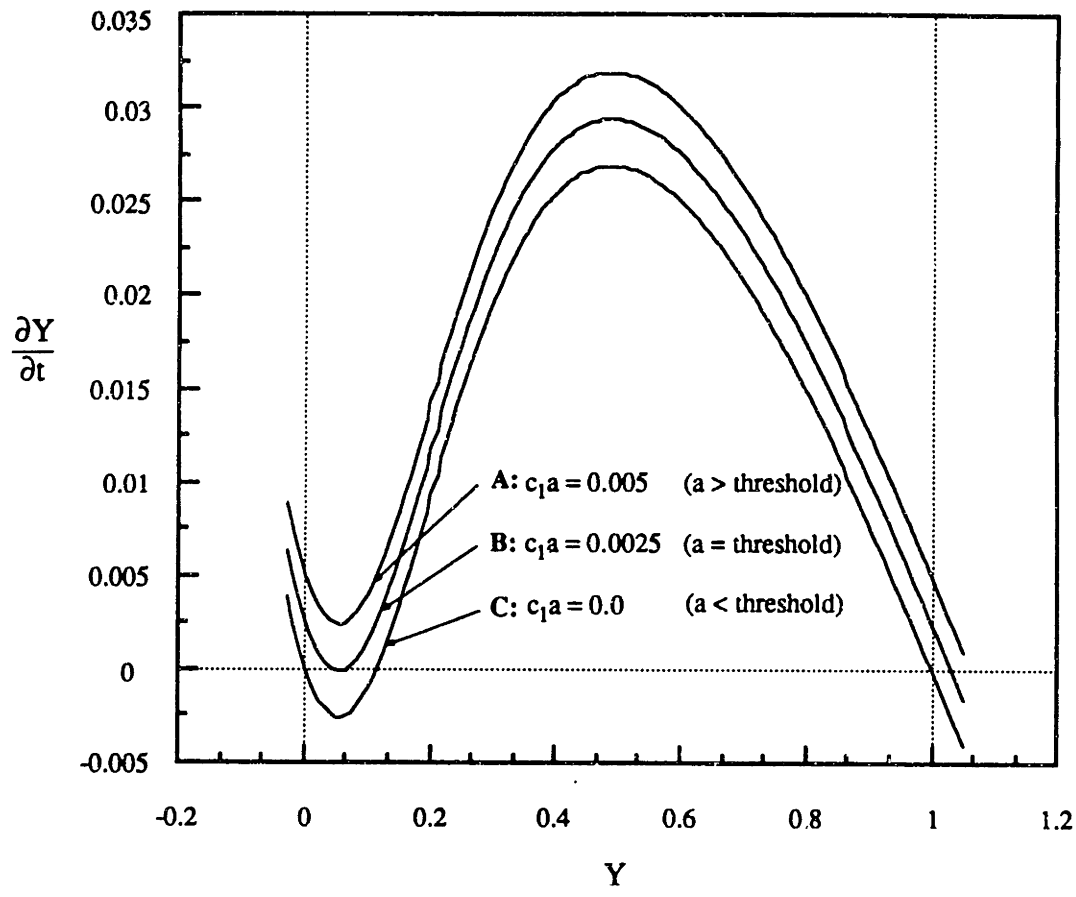
Figure 4.3: The Activator threshold and the Differentiation
equation representation

$$\frac{\partial a}{\partial t} = \boxed{\frac{c\, a^{m_1} z}{h}} - \mu\, a + D_a \frac{\partial^2 a}{\partial x^2} + \rho_0\, Y$$

$$\frac{\partial h}{\partial t} = c_3 \left(\boxed{c_2\, c\, a^{m_1} z} - \nu\, h + D_h \frac{\partial^2 h}{\partial x^2} + \rho_1\, Y \right)$$

①

$$\frac{\partial Y}{\partial t} = \boxed{c_1\, a} - 0.1\, Y + \frac{Y^2}{1 + 9\, Y^2}$$

②

$$\frac{\partial z}{\partial t} = c_0 - \gamma\, c_0\, z - \boxed{\xi\, Y z} + D_z \frac{\partial^2 z}{\partial x^2}$$

③
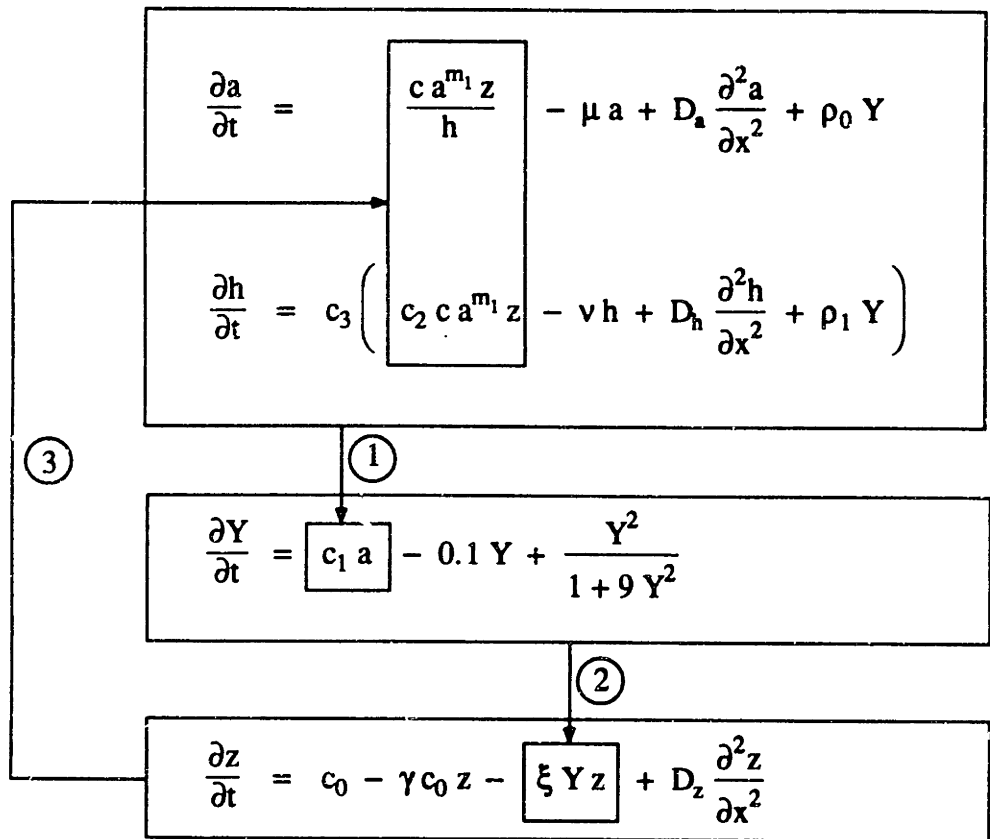
①   Activator ⟶ Differentiation

②   Differentiation ⟶ Substrate

③   Substrate ⟶ Activator

Figure 4.4: Conceptual connections between the states
in the Meinhardt equations

2. A component that is consumed by differentiation and which reinforces the component that triggers differentiation (i.e., substrate).

3. A component that describes the spatial pattern of differentiation (i.e., differentiation).

These conceptual components will be important in later chapters where qualitative analogies are made between the biological model and the physically based channel and hillslope model described in Chapter 5. In this latter model an analogy will be made so that the three components of the developed models are velocity or shear stress, elevation, and channelization, respectively.

A number of authors have analyzed the mathematics of the Meinhardt equations. The techniques used have concentrated on stability analyses at a point and how patterns of activator and inhibitor develop in one dimension (e.g. Haken and Olbrich, 1978; Briere, 1983; Segel, 1984). Bifurcations in the stable states with time have been examined by Granero, et al. (1977) and Nicolis and Prigogine (1977). While many of the concepts and techniques presented in these works are important, the analyses, by necessity, were restricted to a single point in space or properties along a line. The two dimensional, time varying effects that are central to network growth were not studied because of their mathematical complexity.

Why the Meinhardt equations generate networks is the subject of the following section. Some qualitative characteristics of the channel network growth process will be identified and some general characteristics of a physically based network growth model will be proposed.

4.3 Network Growth in the Meinhardt Equations

This section will examine the mechanisms whereby Equations (4.1) generate networks in space. Some general, qualitative, rules will be proposed for space filling network growth on the basis of the governing physical processes.

It is important, first of all, to understand how a single point differentiates, before it will be possible to explain network growth spatially. As previously noted, an individual node differentiates because activator exceeds on activation threshold. Thus an important question to be asked is under what conditions can the activator become large? Consider the activator equation (4.1a)

$$\frac{\partial a}{\partial t} = \frac{ca^{m_1}z}{h} - \mu a + D_a \frac{\partial^2 a}{\partial x^2} + \rho_0 Y$$

It can be shown by scaling analysis that, for the parameters of Meinhardt (1976), the last two terms on the right−hand side are of secondary importance, compared with the first two terms. The activator equation can be approximated by.

$$\frac{\partial a}{\partial t} = \frac{ca^{m_1}z}{h} - \mu a$$

It follows that activator can only grow ($\frac{\partial a}{\partial t} > 0$), if

$$\frac{ca^{m_1-1}z}{\mu h} > 1 \tag{4.2}$$

This threshold is a reliable indicator of growth potential for activator. Figure 4.5 is a schematic cross−section along a line of differentiated nodes (i.e. along a branch). It illustrates how the states in Equation (4.2) vary along and in front of the branch. It is seen that that the criteria of Equation (4.2) is highest just in front of the branch head. This effect is seen in plan also, where a region of high activator potential is seen in front of the branch head.

The process of branching occurs by a similar process. The networks gain their
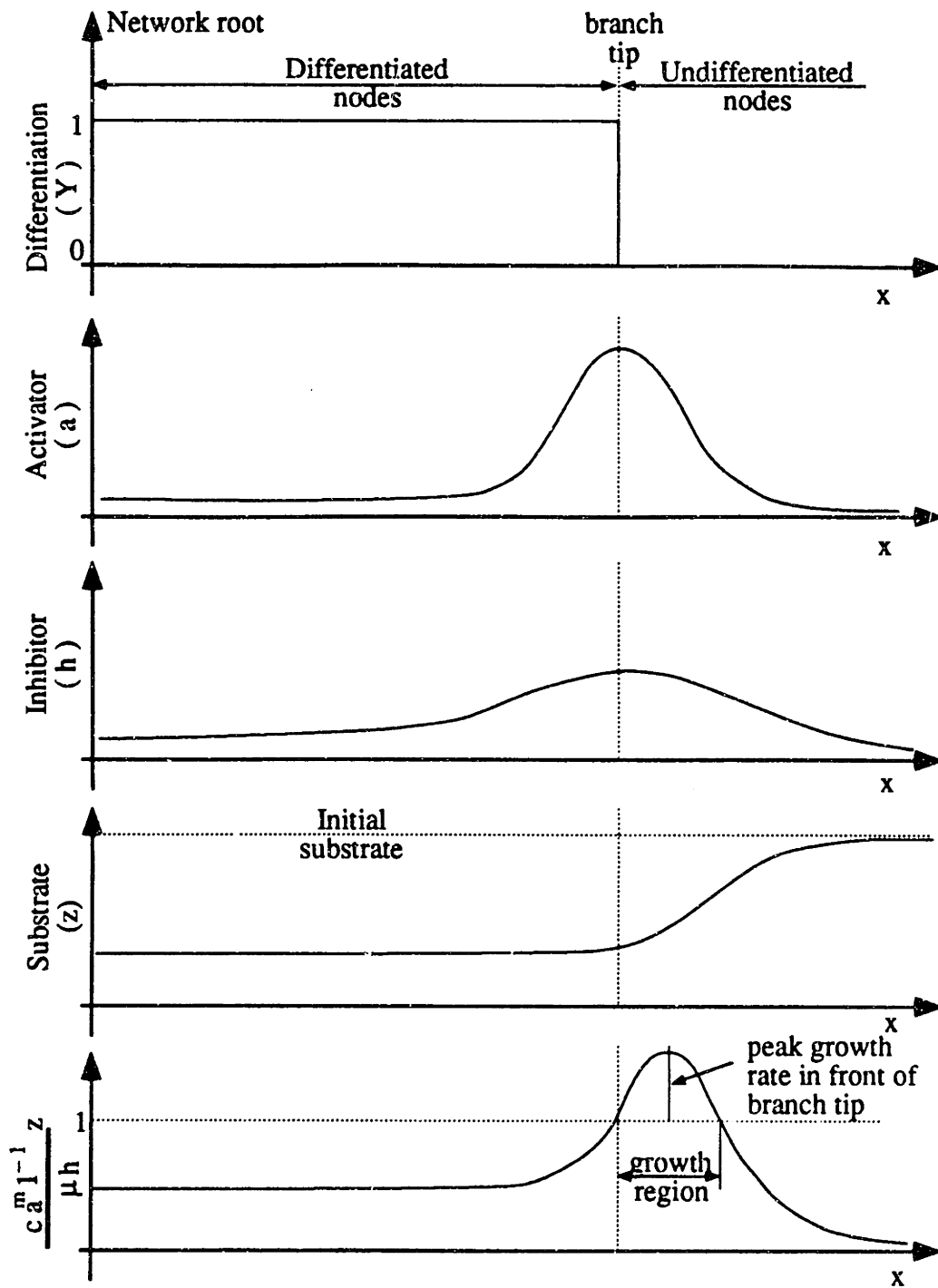
Figure 4.5: Distribution of state's values along a branch in a network: Growth region corresponds to shaded regions in Figures 4.6, 4.7 and 4.8

tree structure from branching or budding from the main stream; branch tip or branch head splitting into two parts is never observed. The mechanism by which networking occurs is thus lateral branching. Buds off the branch occur far behind the growing tip and result from asymmetries of the activator, inhibitor and substrate. As in branch head growth, the region of activator growth is given by Equation (4.2). The distribution in space of the states determines the distribution of the branching points. Because of the screening near the network root, network growth sites are less likely near the network root. Near the network root substrate levels are low and inhibitor levels are relatively high. Thus there is general inhibition of growth or lateral branching behind the growing branch heads. This is just the screening effect discussed in Section 3.2; the physics that causes the screening is somewhat different here but the qualitative effect on the spatial distribution of network growth is similar.

There are three prerequisites on the qualitative behavior of the activator physics. These prerequisites ensure that network–like growth of differentiation occurs.

The first prerequisite is that the region of high activator must be localized around the growing branch tip. The subtle asymmetries in the spatial distribution of activator, inhibitor and substrate around the growing tip result in high values of activator at small distances in front of the growing tip. This effect is illustrated in Figure 4.6.

It is important to note that the actual position and shape of the "activated" region depends on the physics of the processes acting in the undifferentiated portion of the leaf. The driving force for these processes is the line of differentiated nodes. Thus there is a complicated interaction between the existing pattern of differentiation, the processes in the leaf proper and thus the shape of the activated region, and the future pattern of differentiated nodes.

The second prerequisite is that the region of high activator must move with the growing branch tip. This movement of the activated region as the branch grows in
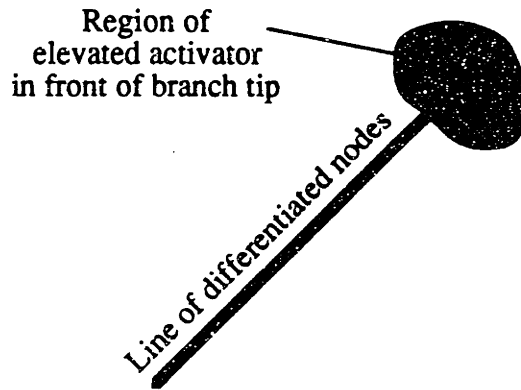
Figure 4.6: Schematic of the mechanism for branch tip growth: Activator is increased in front of the growing branch tip by substrate gradients

BEFORE TIP GROWTH   AFTER TIP GROWTH
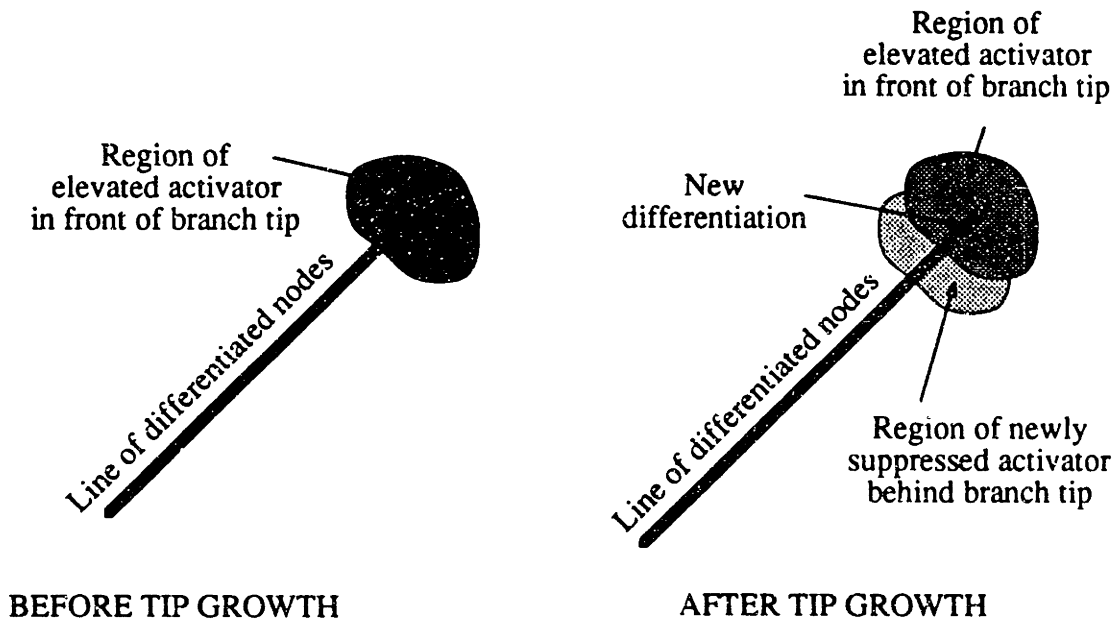
Figure 4.7: Schematic of the mechanism for maintaining linear growth of branchs: Activator is locally reduced behind the growing tip by substrate depletion to ensure growth occurs in only the forward direction

time results from the interaction between differentiated and undifferentiated regions. The process of channelization consumes substrate and this lowers the activator behind the tip so that the region of high activator moves with and in front of the tip. This process is demonstrated in Figure 4.7. If the activator were not locally suppressed behind the growing tip then blobs of differentiation would form.

The third prerequisite of the activator distribution is that regions of high activator must "repel" each other so that the branch tips grow away from each other. This produces space filling characteristics, similar to those observed in channel networks (e.g., Abrahams, 1984). The repulsion of growing tips and repulsion from boundaries is shown in Figure 4.8. This effect arises from complex interactions between the spatial distribution of activator and inhibitors driven by the diffusion processes in the undifferentiated portions of the leaf.

Gierer (1981) proposed that the system of equation with the correct characteristics to produce spatial patterns, of the form in Figure 4.1, is more general than that proposed by Meinhardt (1976). His conclusions were reached for an activator–inhibitor system rather than the activator–inhibitor–substrate system of Meinhardt (1976). Since activator is the only direct input into the differentiation equation the differences between the system of equations Gierer studied and the Meinhardt equations are relatively minor.

Gierer (1981), used general equations of the form

$$\frac{\partial a}{\partial t} = f(a,h) + D_a(a)$$

$$\tag{4.3}$$

$$\frac{\partial h}{\partial t} = g(a,h) + D_h(h)$$

where $D(\cdot)$ is a general diffusion–like coupling in space. He proposed six conditions required for pattern formation.

Future Growth

Activator suppressed between tips

Activator elevated in front of and outside branch tips

Line of differentiated nodes

Line of differentiated nodes

Future Growth
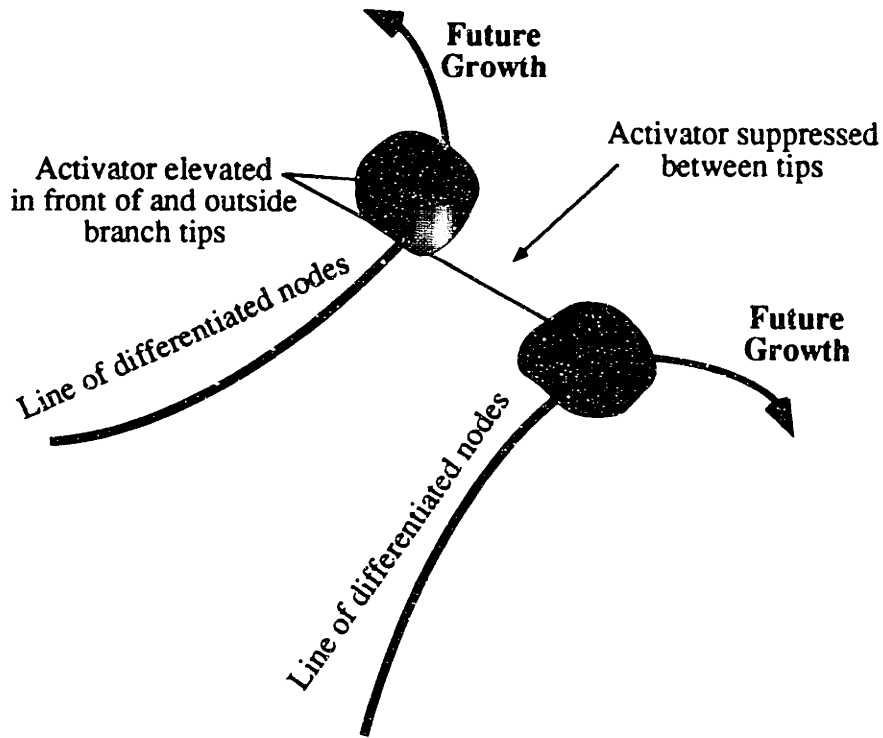
(a) Two tips near each other

Activator elevated in front of tip and opposite side to boundary

Future growth

Activator suppressed on boundary side

Real branch

Line of differentiated nodes

Impermeable boundary

Image branch
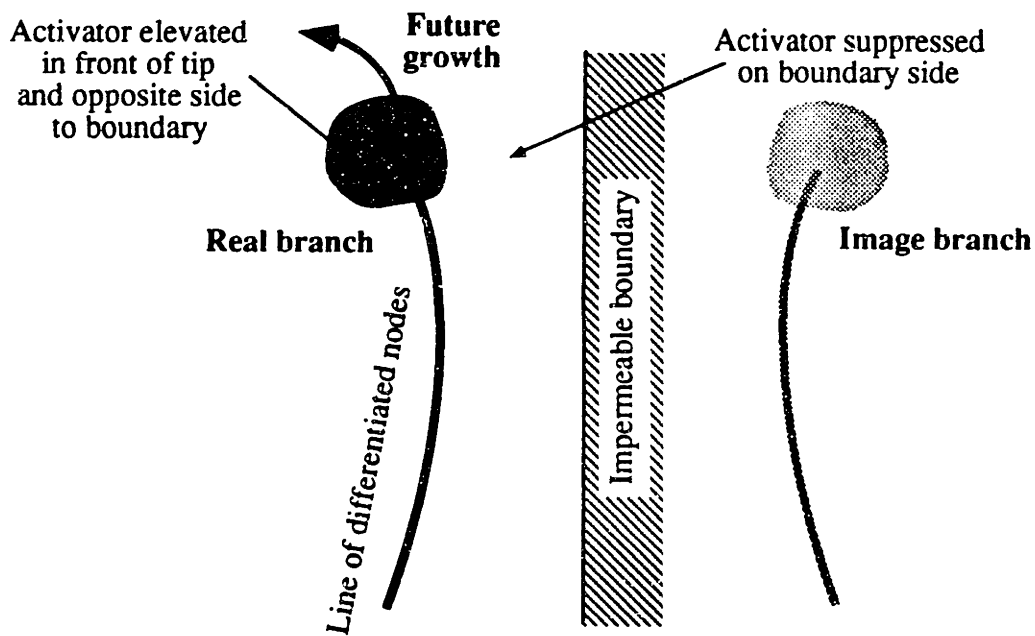
(b) Tip near boundary

Figure 4.8: Schematic of the mechanism to ensure space filling by the networks
(a) Mutual repulsion of growing tips
(b) Repulsion from the boundary

These conditions are:

1. the existence of an autocatalytic term (i.e., $\frac{\partial a}{\partial t} \approx ca^m$, m > 1, for some value of states).

2. the existence of an inhibitory (i.e., h) component.

3. the inhibitory effect must be strong enough to ensure that stable, bounded solutions exist for activator.

4. the length scale of activation must be less than the length scale of the total field size (otherwise boundary conditions stabilize the activator).

5. the inhibitory time scale must be less than the activating time scale (so that activation does not run away from inhibition).

6. the length scale of inhibition must be greater than the length scale of activation (so that activation is confined to areas near those already activated).

These conditions can be linked to the three prerequisites on the activator distribution that have been noted above (Figures 4.6, 4.7, and 4.8). To ensure a localized region of activator around the branch head requires an autocatalytic activator (point 1) with an inhibitor away from the branch head to suppress long range activation (points 2, 3, 4, and 6). To ensure that the region of activation moves with the branch head requires all of the above effects, plus it also requires that the inhibitor acts faster than the time scale of branch head advance. Since branch head movement cannot proceed faster than the activator growth of the branch head, then the inhibitor must act at timescales less than the activator timescale (point 5). The repulsion of branch heads results from the longer range inhibition and shorter range activation (point 6).

Gierer's conditions suggest that the form of the equations that can generate networks may be more general than those presented by Meinhardt (1976).

To test whether Gierer's conditions on activator and inhibitor in Equation (4.3)

are sufficient to generate networks, a sensitivity study of the parameters used by Meinhardt (1976) in Equation (4.2) was performed. The parameters were systematically varied to determine their effect on the networks generated. Two related subjects were of particular interest:

1. Testing Gierer's six hypotheses regarding the activator and inhibitor system in the context of the more general equations of Equation (4.1).

2. Determining parameter windows outside of which networks could not be generated.

The parameters to be discussed below are all parameters in Equation (4.1). To test the requirement for autocatalytic behavior (Gierer's point 1), m_1 was varied over a broad range. Networks could be generated for values of $m_1 > 1$, provided compensatory changes were made in parameter c so that activator magnitudes were maintained at around the same level. As m_1 approached 1, the rate at which the network propagated declined reflecting the lower autocatalytic strength.

To test Gierer's points 2 and 3, the parameter c_2 was varied. This change modified the relative balance between activator and inhibitor generation. If c_2 was reduced by more than 25%, so reducing inhibition, runaway activation occurred and the whole region differentiated into a blob—like structure. If c_2 was increased by more than about 25%, network growth ceased. Thus it appears that the balance that exists between activator and inhibitor magnitudes is very important.

To test Gierer's point 6, regarding the relative length scales of activator and inhibitor, the diffusivities of activator and inhibitor were varied. This variation was equivalent to a change in the length scales of diffusion. Meinhardt's ratio of diffusivities was $\frac{D_h}{D_a} = 10$. Reduction of D_h was possible until D_h was of comparable magnitude to D_z (Meinhardt's $\frac{D_h}{D_z} = 3$) at which value a network was difficult to generate. Changes in D_h were negatively correlated to changes in drainage density,

thus being in line with the changes in diffusion length scales. Meinhardt (1976) chose the values of D_a and μ with an eye to maintaining connected networks. The length scale of activator diffusion ($\equiv \frac{D_a}{\mu}$) was about 0.5 of a grid spacing. If either, D_a was increased, or μ was decreased, thus increasing the activator length scale, "networks" could still be generated but disconnected regions were common. That is, adjacent nodes were not necessarily differentiated. This was a numerical, rather than a conceptual, problem.

To test Geirer's point 5, regarding the inhibitor timescale being less than the activator timescale, the parameter c_3 was reduced. A reduction in c_3 results in an increased reaction time for inhibitor. In a similar fashion to reductions in c_2, reductions in c_3 resulted in runaway activation. An interesting interaction with the diffusive length scales occurred if c_3 was increased since $c_3 D_h$ is the inhibitor diffusivity. As c_3 was increased drainage density was reduced. With pronounced increases in c_3 (> 50%) network generation was suppressed; this resulted from the inhibitor being increased compared to the activator.

Other sensitivities examined were for parameters c_0, γ, ϵ, m_2, D_z, and c_1. The substrate equation (4.1c) was relatively insensitive to changes in any of the parameters; some processes were not even required. An exception was c_0, the source term for substrate, which had to be non–zero for network formation. There had to be a source of substrate to counteract the consumption of substrate by other terms in the substrate equation.

In the differentiation equation, c_1 controlled the drainage density of the network. If c_1 was small, so that the activator threshold was high, then for a given distribution of activator a node was less likely to differentiate, and vice versa. Thus changes in c_1 and drainage density were positively correlated.

In general, it seems that Gierer (1981) was correct. Provided that his generic

rules are satisfied, then networks may be formed by a wide range of physical processes. We did find, however, that the values of the coefficients on these physical processes needed to be "optimized" to generate the best looking networks; the form of the processes, however, was not critical.

In conclusion, it has been shown that it is possible to generate space filling networks from a very generic set of equations featuring diffusive coupling of processes in space and autocatalysis in time. Three generic characteristics that must be shown by the spatial pattern of activation were also proposed. It is believed, and supported by the sensitivity studies, that these characteristics are generic requirements of all network generating processes.

4.4 Input Randomness, Chaos and Random Networks

4.4.1 Introduction

All the network patterns presented here and in Meinhardt (1976, 1982) have a random appearance. No two branches are alike and there are no obvious symmetries. Nevertheless the qualitative form of the networks appears to be similar despite this randomness. It is the purpose of this section to explore the reasons for this randomness, and to demonstrate that the randomness in the networks arises because of a phenomenon called transient chaos, a form of chaos that has only recently been recognized in the mathematical literature (Moon, 1987).

The most obvious random input to Equation (4.1) is that Meinhardt (1976, 1982) applied a small random field to parameter c. These random fluctuations in c were less than 1% of the mean value. No other random inputs were applied. However, it should be noted that numerical errors add a small random input. It will be shown that both these sources can be important in creating the observed randomness of the networks. The important question that will be answered below is how these small random fluctuations grow to dominate the overall form of the generated networks.

The relevance of randomness in networks can be motivated by the simple example in Figure 4.9. Here six different networks are presented. All six networks were generated using identical parameters but varying the random inputs and the way that the numerical calculations were performed. Figure 4.9a is the baseline case. It uses the random field input on parameter c (Equation 4.1a), and the equations are solved in single precision on a microVAX (32 bit, 7 significant digits). The network is clearly random with no apparent lateral symmetries resulting from the rectangular domain. Figure 4.9b is the same as Figure 4.9a except that parameter c is now deterministic, a constant across the domain. There is some lateral symmetry in the parts of the network developed at early time (the diagonal branch splitting the domain in two) but it is lost in that part of the network that develops at later times. Figure 4.9c is the same as Figure 4.9b except that the internal calculation of the 5 point centered finite difference for the activator diffusion term is slightly modified; the order of addition of the nodes is anti–clockwise instead of clockwise. The two finite difference approximations were:

$$\left[\frac{\partial^2 z}{\partial x_1^2} + \frac{\partial^2 z}{\partial x_2^2}\right]\bigg|_{(i,j)} = \frac{z_{i+1,j} + z_{i,j-1} + z_{i-1,j} + z_{i,j+1} - 4z_{i,j}}{(\Delta x)^2}$$

$$\left[\frac{\partial^2 z}{\partial x_1^2} + \frac{\partial^2 z}{\partial x_2^2}\right]\bigg|_{(i,j)} = \frac{z_{i+i,j} + z_{i,j+1} + z_{i-1,j} + z_{i,j-1} - 4z_{i,j}}{(\Delta x)^2}$$

where (i,j) = node coordinates for evaluation of the second derivative

                     approximation.

It is apparent that this extremely small change in numerics has a significant effect on the network form at later times. At early times, when numerical errors have not had a chance to propagate, the networks are identical. Numerical errors thus seem
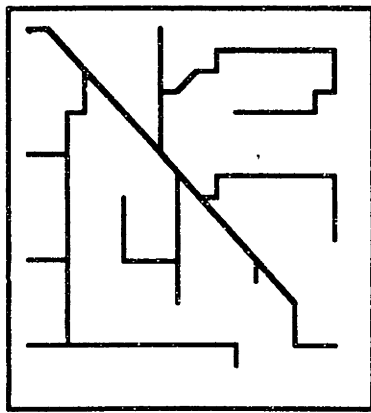
(a) Baseline
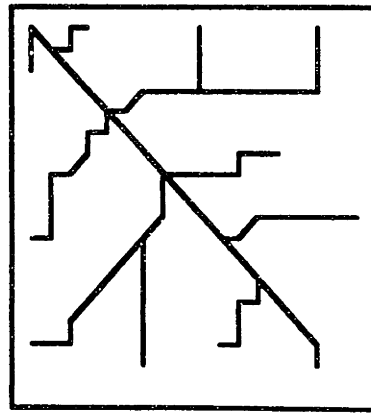
(b) Single precision,
no randomness
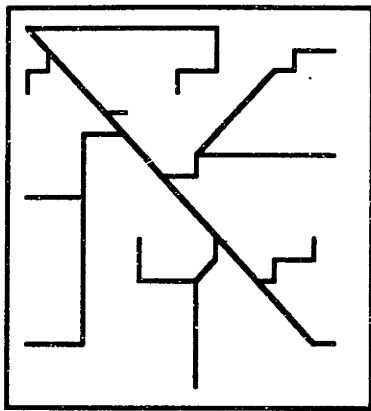
(c) New diffusion,
no randomness
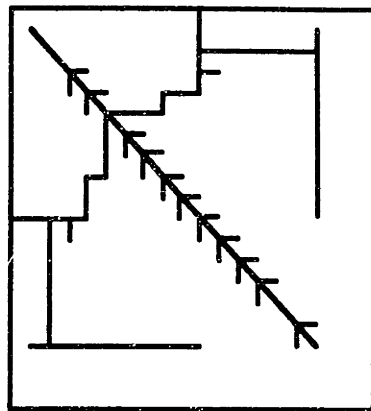
(d) Double precision,
no randomness
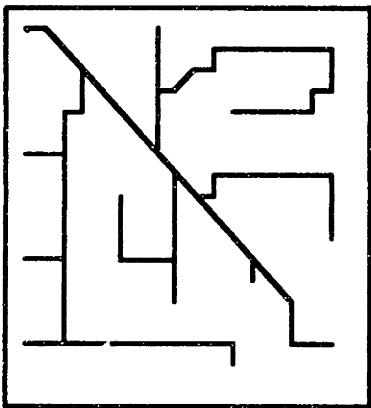
(e) New diffusion term,
baseline randomness
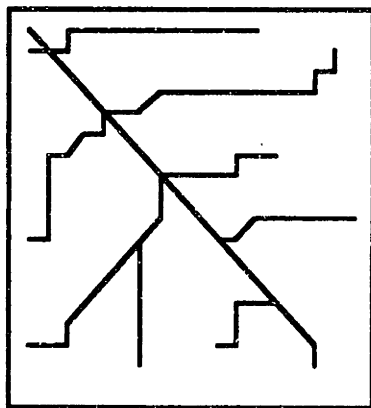
(f) Randomness in the
initial conditions on
activator

Figure 4.9: Illustrating the effect of randomness
on networks from the Meinhardt equation :
Input randomness and numerical noise

115

to be responsible for part of the randomness of the networks. Figure 4.9d confirms this impression by solving the same problem as Figure 4.9b in double precision (64 bits, 15 significant digits). The network exhibits lateral symmetry and close examination indicates no randomness in any of the states of the equation (i.e., activator, inhibitor, substrate or differentiation).

The random effects due to the numerics are smaller than the input randomness on parameter c. Figure 4.9e shows a network with the diffusion term calculated as for Figure 4.9b but with the input randomness of Figure 4.9a. That Figure 4.9a and 4.9e are indistinguishable indicates that the numerical randomness is small in comparison with the applied randomness in this simulation. In a large enough simulation, where the network is actively growing over many timesteps, it is possible that the cumulative effects of numerical errors may become important.

Finally, although the results are not presented here, not all random effects result in random networks. Minor perturbations were applied to the initial substrate values. There was no perceptible differences with the network generated without the substrate perturbation. On the other hand minor random perturbations on the initial activator levels resulted in random networks comparable in randomness with the networks resulting from the random parameter c (see Figure 4.9f).

There are four important features of this example that will be noted

1. Randomness in parameter c has a very important effect on the randomness of the simulated network.

2. Not all input randomness is created equal. Some input randomness has no apparent effect on the networks generated.

3. Numerical error due to minor round–off errors in the space and time discretization can propagate nonlinearly in time to dominate the network form at later times.

4. · Despite this, parameter randomness is <u>not</u> essential for generation of

116

networks from Meinhardt's equations. Parameter randomness modifies the form of the final network. In addition, once the parameter randomness is determined (i.e., a random field realization of the parameter is chosen), in the absence of other perturbations, the resulting network is fixed.

The reasons for this behavior will now be explored, but since the concepts that are to be used are new, then some explanation of the methodology is in order.

4.4.2 Basics of Nonlinear Systems and Chaos

Consider, initially, a first order, single state linear differential equation

$$\frac{ds}{dt} = bs \qquad\qquad s = s_0 \text{ at } t = 0 \tag{4.4}$$

The exact solution of this equation is

$$s = s_0 \, e^{bt} \tag{4.5}$$

If the system starts at a slightly perturbed initial condition $s = s_0 + \Delta s_0$ at time zero, then the solution is

$$s = s_0 \, e^{bt} + \Delta s_0 \, e^{bt} \tag{4.6}$$

Whether the initial perturbation Δs_0 grows with time or not depends on the sign of b. If b is negative, then after a long enough time the exact and the perturbed answers will be indistinguishable. If b is positive, however, the two answers will diverge for all time. In the nomenclature of the chaos literature the solution of Equation (4.4), for b positive, is sensitively dependent on the initial conditions.

Unfortunately the linear system of Equation (4.4) is too simple to demonstrate everything needed. Although, for b positive, the initial random fluctuation in Equation (4.6) grows with time, it never dominates the exact solution, Equation (4.5), because the exact solution grows at the same rate. Consideration of a nonlinear equation similar to Equation (4.4) will demonstrate a case where both the exact solution and perturbation grow, but at different rates. Consider the differential equation

$$\frac{ds}{dt} = bs(1-s) \qquad\qquad s = s_0 \text{ at } t = 0$$

Linearizing this equation about $s = s^*$, the exact solution at time t^*, using a Taylor Series expansion up to the first order terms,

$$\frac{d(s^* + \Delta s)}{dt} = bs^*(1-s^*) + (b - 2bs^*)\Delta s \qquad\qquad (4.7)$$

The equation for the growth of the perturbation Δs, applied at time t^*, is then

$$\frac{d\Delta s}{dt} = (b - 2bs^*)\Delta s \qquad\qquad (4.8)$$

This equation is linear in Δs so that, as in Equation (4.4), the perturbation Δs grows if $(b - 2bs^*) > 0$. The perturbation will grow faster than the exact solution if

$$\frac{(1 - 2s^*)}{s^*(1 - s^*)} > 1$$

Under these conditions it is possible for the perturbation to grow to dominate the exact solution. For a system of nonlinear differential equations, the situation is

slightly more difficult. Consider the general n dimensional system with states $s_1, ..., s_n$ of the form

$$\frac{d\underline{s}}{dt} = \underline{f}(\underline{s}) \qquad (4.9)$$

where $\qquad \underline{s}(0) \quad = \quad \underline{s}_0$

$$\underline{s} \quad = \quad \begin{bmatrix} s_1 \\ \vdots \\ s_n \end{bmatrix} = \text{states of the equation}$$

To obtain the differential equation governing the perturbation growth, the system is linearized in a similar fashion as was Equation (4.7), now using an n–dimensional Taylor Series so that

$$\frac{d(\underline{s}^* + \Delta \underline{s})}{dt} = \underline{f}(\underline{s}^*) + A(\underline{s}^*)\Delta \underline{s} \qquad (4.10)$$

where $\qquad A(\underline{s}^*) \quad = \quad \begin{bmatrix} \dfrac{\partial f_1}{\partial s_1} & \dfrac{\partial f_1}{\partial s_2} & \cdots & \dfrac{\partial f_1}{\partial s_n} \\ & \vdots & & \\ \dfrac{\partial f_n}{\partial s_1} & & \cdots & \dfrac{\partial f_n}{\partial s_n} \end{bmatrix} \Bigg|_{\underline{s} = \underline{s}^*}$

$$\Delta \underline{s} \quad = \quad \begin{bmatrix} \Delta s_1 \\ \vdots \\ \Delta s_n \end{bmatrix} = \text{vector of perturbations on the states } \underline{s}.$$

Subtracting Equation (4.9) evaluated at $\underline{s} = \underline{s}^*$ from Equation (4.10) gives the differential equation governing the growth of the perturbations

$$\frac{d(\Delta \underline{s})}{dt} = A(\underline{s}^*)\Delta \underline{s} \qquad (4.11)$$

The solution of this equation, for small Δs, is a simple result from linear algebra (e.g. Hirsch and Smale, 1974)

$$\Delta \underline{s} = e^{A(\underline{s}^*)t} \Delta \underline{s}_0 \qquad (4.12)$$

The way to characterize whether the individual perturbations Δs_i grow or diminish with time is more complicated than in the one—dimensional case of Equation (4.4). The answer involves determining the eigenvalues and eigenvectors of $A(\underline{s}^*)$. A strict, though difficult, derivation of this answer may be found in Hirsch and Smale (1974). The derivation that follows is somewhat less strict and more intuitive.

The matrix exponential can be expanded as a series

$$\Delta \underline{s} = \left[I + A(s^*)t + ...\right] \Delta \underline{s}_0 \qquad (4.13)$$

or rearranging the truncated form of this equation

$$(\Delta \underline{s} - \Delta \underline{s}_0) = A(s^*) \Delta \underline{s}_0 t \qquad (4.14)$$

The perturbations $\Delta s_1, ..., \Delta s_n$ will grow with time when the corresponding term on the right—hand side of Equation (4.14) is positive. This is characterized by the eigenvalues $\lambda_1, ..., \lambda_n$ and the corresponding eigenvectors $\underline{e}_1, ..., \underline{e}_n$ of $A(s^*)$.

It is a basic result of linear algebra that the matrix $A(s^*)$ can be expressed in terms of its eigenvalues and eigenvectors, (neglecting the s^* dependence for clarity)

$$A = T^{-1} D T$$

where   T   = (n x n) matrix, columns are the eigenvectors

D   = (n x n) diagonal matrix of the eigenvalues

$$= \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \cdot & \\ & & & \cdot \\ 0 & & & \cdot \lambda_n \end{bmatrix}$$

Substituting this result into Equation (4.14) yields

$$(\Delta \underline{s} - \Delta \underline{s}_0) = T^{-1} D T \Delta \underline{s}_0 t$$

Multiplying through by T, on the left—hand side, and transforming the equation so that the basis of the vectorspace is the eigenbasis, yields the equation

$$(\Delta \underline{s}' - \Delta \underline{s}'_0) = D \Delta \underline{s}'_0 t \tag{4.15}$$

where

$$\Delta \underline{s}' \qquad\qquad = T \Delta \underline{s} \tag{4.16}$$

                = the original perturbations $\Delta \underline{s}$, transformed into the eigenspace

$$\Delta \underline{s}'_0 \qquad\qquad = T \Delta \underline{s}_0$$

Noting the definition of the diagonal matrix D, it is now clear that if an individual eigenvalue $\lambda_i > 0$, then the corresponding perturbation $|\Delta s'_i - \Delta s'_{0i}|$ will be positive so that the perturbation $\Delta s'_i$ grows with time. This is clearer if the steps

that took us from Equation (4.12) to (4.13) are reversed and applied to Equation (4.15) so that

$$
\begin{aligned}
\Delta s' \quad &= e^{Dt} \, \Delta \underline{s}'_0 \\[2mm]
&=
\begin{bmatrix}
e^{\lambda_1 t} & & & 0 \\
& e^{\lambda_2 t} & & \\
& & \ddots & \\
0 & & & e^{\lambda_n t}
\end{bmatrix}
\begin{bmatrix}
\Delta s'_{01} \\
\Delta s'_{02} \\
\vdots \\
\Delta s'_{0n}
\end{bmatrix}
\end{aligned}
\qquad (4.17)
$$

Taking an individual perturbation $\Delta s'_i$ then

$$
\Delta s'_i = e^{\lambda_i t} \, \Delta s'_{0i}
$$

Thus if $\lambda_i > 0$, the corresponding perturbation $\Delta s'_i$ will grow with time. The chaos literature (e.g., Moon, 1987) defines the Lypanov exponent, loosely speaking, the measure of how fast small perturbations grow or decline, as

$$
\sigma_i = e^{\lambda_i}
\qquad (4.18)
$$

where σ_i = ith Lypanov exponent.

If the ith Lypanov exponent is greater than 1, then the ith perturbation will grow with time; if it is less than 1, the corresponding perturbation will decay with time.

It is important to note that Equations (4.17) and (4.18) only give information about whether perturbations in the eigenspace grow or decay, not whether and how much the original perturbations Δs_1, ..., Δs_n grow or decay. How the original untransformed perturbations, $\Delta \underline{s}$, grow and decay depends on the linear combination of

the eigenvectors and eigenvalues implied by Equation (4.16).

In addition, for a non—linear system, the linearization that gives the perturbation Equation (4.11) changes with time, dependent on the value of $\underline{s}^*$ with time. Thus, just as the linearization changes with time, so do the eigenvalues and eigenvectors. This is why the definition of Lypanov coefficient above is only loosely correct; the chaos literature uses a time averaged value removing the dependency of the linearization on time (Holden, 1987).

The relevance of the preceding discussion to the Meinhardt equations will now be presented. The perturbations that are important are the initial random field input into the coefficients in the activator equation and the random numerical noise that inevitably arises in the numerical solution of the equations. How they propagate into the states of Equation (4.1) will be shown.

4.4.3 Analysis of the Chaotic Growth of the Meinhardt Equation

This section will analyze the Meinhardt equations to determine regions of chaotic growth, where small perturbations can grow and eventually dominate the solution of the equations. This will demonstrate how random errors, either input randomness or numerical errors, may modify the form of the network generated so that it appears random.

Using the techniques described in the previous section, the Meinhardt equations will be linearized around a nominal trajectory. This linearization gives the differential equation of the perturbations from which the eigenvalues, eigenvectors, and Lypanov exponents will be obtained. The distribution of positive eigenvalues (Lypanov coefficients greater than 1) will be examined. These regions of positive eigenvalues will be asserted to be regions of transient chaos, where small perturbations may grow and dominate the solution of the equations.

The Meinhardt equation (Equation 4.1) may be formulated as in the form of the

general nonlinear system of Equation (4.9) as

$$\frac{\partial \underline{s}}{\partial t} = \underline{f}(\underline{s})$$

(4.19)

where

$$\underline{s} = \begin{bmatrix} a \\ h \\ z \\ y \end{bmatrix} \qquad \underline{s}(0) = \begin{bmatrix} a_0 \\ h_0 \\ z_0 \\ y_0 \end{bmatrix}$$

$$\underline{f}(\underline{s}) = \begin{bmatrix} f_a(\underline{s}) \\ f_h(\underline{s}) \\ f_z(\underline{s}) \\ f_y(\underline{s}) \end{bmatrix} = \begin{bmatrix} \dfrac{ca^{m_1}z}{h} - \mu a + D_a \dfrac{\partial^2 a}{\partial x_i^2} + \rho_0 Y \\[2ex] c_3 (c_2 ca^{m_1}z - \nu h + D_h \dfrac{\partial^2 h}{\partial x_i^2} + \rho_1 Y) \\[2ex] c_0 - \gamma c_0 z - \epsilon Y z^2 + D_z \dfrac{\partial^2 z}{\partial x_i^2} \\[2ex] c_1 a - 0.1Y + \dfrac{Y^2}{1 + 9Y^2} \end{bmatrix}$$

It follows that the equations for the perturbations $\Delta\underline{s}$ can be obtained from Equation (4.19) by use of the linearizations in Equations (4.10) and (4.11) so that

$$\frac{\partial(\Delta\underline{s})}{\partial t} = A(\underline{s}^*)\Delta\underline{s}$$

(4.20)

where

$\underline{s}^*$ = nominal states at $t = t^*$, the point around which the linearization is performed.

$\Delta\underline{s}$ = perturbations on the states $\underline{s}$.

$A(\underline{s}^*)$ = (nxm) state matrix for the perturbations

$$= \begin{bmatrix} \dfrac{\partial f_a}{\partial a} & \dfrac{\partial f_a}{\partial h} & \dfrac{\partial f_a}{\partial z} & \dfrac{\partial f_a}{\partial Y} \\[2mm] \dfrac{\partial f_h}{\partial a} & \dfrac{\partial f_h}{\partial h} & \dfrac{\partial f_h}{\partial z} & \dfrac{\partial f_h}{\partial Y} \\[2mm] \dfrac{\partial f_z}{\partial a} & \dfrac{\partial f_z}{\partial h} & \dfrac{\partial f_z}{\partial z} & \dfrac{\partial f_z}{\partial Y} \\[2mm] \dfrac{\partial f_Y}{\partial a} & \dfrac{\partial f_Y}{\partial h} & \dfrac{\partial f_Y}{\partial z} & \dfrac{\partial f_Y}{\partial Y} \end{bmatrix} \Bigg|_{\underline{s} = \underline{s}^*}$$

The individual derivatives in the matrix $A(\underline{s}^*)$ are

$$\frac{\partial f_a}{\partial a}\bigg|_{\underline{s}=\underline{s}^*} = \frac{cm_1(a^*)^{m_1-1} z^*}{h^*} - \mu + D_a \qquad (4.21a)$$

$$\frac{\partial f_a}{\partial h}\bigg|_{\underline{s}=\underline{s}^*} = -\frac{c(a^*)^{m_1-1} z^*}{(h^*)^2} \qquad (4.21b)$$

$$\frac{\partial f_a}{\partial z}\bigg|_{\underline{s}=\underline{s}^*} = \frac{c(a^*)^{m_1}}{h^*} \qquad (4.21c)$$

$$\frac{\partial f_a}{\partial Y}\bigg|_{\underline{s}=\underline{s}^*} = \rho_0 \qquad (4.21d)$$

$$\frac{\partial f_h}{\partial a}\bigg|_{\underline{s}=\underline{s}^*} = c_3 c_2 c\, m_1 (a^*)^{m_1-1} z^* \qquad (4.21e)$$

$$\frac{\partial f_h}{\partial h}\bigg|_{\underline{s}=\underline{s}^*} = -\nu + D_h \qquad (4.21f)$$

$$\frac{\partial f_h}{\partial z}\bigg|_{\underline{s}=\underline{s}^*} = c_3 c_2 c(a^*)^{m_1} \qquad (4.21g)$$

$$\left.\frac{\partial f_h}{\partial Y}\right|_{\underline{s}=\underline{s}^*} = \rho_1 \tag{4.21h}$$

$$\left.\frac{\partial f_z}{\partial a}\right|_{\underline{s}=\underline{s}^*} = \left.\frac{\partial f_z}{\partial h}\right|_{\underline{s}=\underline{s}^*} = 0 \tag{4.21i}$$

$$\left.\frac{\partial f_z}{\partial z}\right|_{\underline{s}=\underline{s}^*} = -\gamma c_0 - \epsilon m_2 Y^* (z^*)^{m_2-1} + D_z \tag{4.21j}$$

$$\left.\frac{\partial f_z}{\partial Y}\right|_{\underline{s}=\underline{s}^*} = -\epsilon (z^*)^{m_2} \tag{4.21k}$$

$$\left.\frac{\partial f_Y}{\partial a}\right|_{\underline{s}=\underline{s}^*} = c_1 \tag{4.21l}$$

$$\left.\frac{\partial f_Y}{\partial h}\right|_{\underline{s}=\underline{s}^*} = \left.\frac{\partial f_Y}{\partial z}\right|_{\underline{s}=\underline{s}^*} = 0 \tag{4.21m}$$

$$\left.\frac{\partial f_Y}{\partial Y}\right|_{s=s^*} = -0.1 + \frac{2Y^*}{(1 + 9Y^{*2})^2} \tag{4.21n}$$

The expressions for derivatives in Equation (4.21) are applicable at a point in space and to determine them all that are needed is the values of the nominal states at that point; the derivatives are independent of the values of the states at surrounding points. Thus Equation (4.20) is not a partial differential equation. This is an important point since the original problem (Equation 4.19) was a partial differential equation and to solve it discretization in space is required. The original problem is coupled in space. However, due to the linearization the perturbations are uncoupled in space; the solution for perturbations can be determined independently of the states at

nearby grid points. Thus the perturbation equation can be solved at each grid point independently of the other grid points. Thus for a 20 x 20 grid four hundred 4 x 4 problems must be solved to determine the eigenvalues and eigenvectors. If the perturbations were coupled in space a single, 1600 x 1600 problem must be solved. Hence determining the eigenvalues and eigenvectors of $A(s^*)$ for the former problem is reasonable with existing computer equipment while the latter is not.

The reason for this decoupling in space of the perturbations is the linear spatial coupling in the original state equations (Equation 4.19). The diffusion term $D_a \frac{\partial^2 a}{\partial x^2}$, $D_h \frac{\partial^2 h}{\partial x^2}$, $D_z \frac{\partial^2 z}{\partial x^2}$ are linear. If these terms were not linear, then the perturbation would not be decoupled. It would then be necessary to linearize the diffusion terms to decouple them. Thus the following eigenvalue and eigenvector analysis is exact and involves no approximation.

To demonstrate the stability properties of the Meinhardt equations, the states from a partially grown network were used. These states were used as the nominal states for the linearization. The parameters used to generate the sample data set were those of Meinhardt (1982).

Contours of the four states activator, inhibitor, substrate and differentiation are provided in Figure 4.10. Figure 4.10a shows contours of activator. The four, perhaps five, peaks correspond to points near the network where active growth of activator is occurring around the branch heads. The pattern of differentiation and branch heads is shown in Figure 4.10d. The eigenvalues of $A(s^*)$ from the perturbation equations are contoured in Figure 4.11. As noted in the previous section, these eigenvalues correspond to the eigenvectors of $A(s)$, not the original states. Not only are the eigenvectors linear combinations of the states, but they are linear combinations that vary from point to point in the grid, since the eigenvalues and eigenvectors vary from point to point depending upon the linearization at that point.
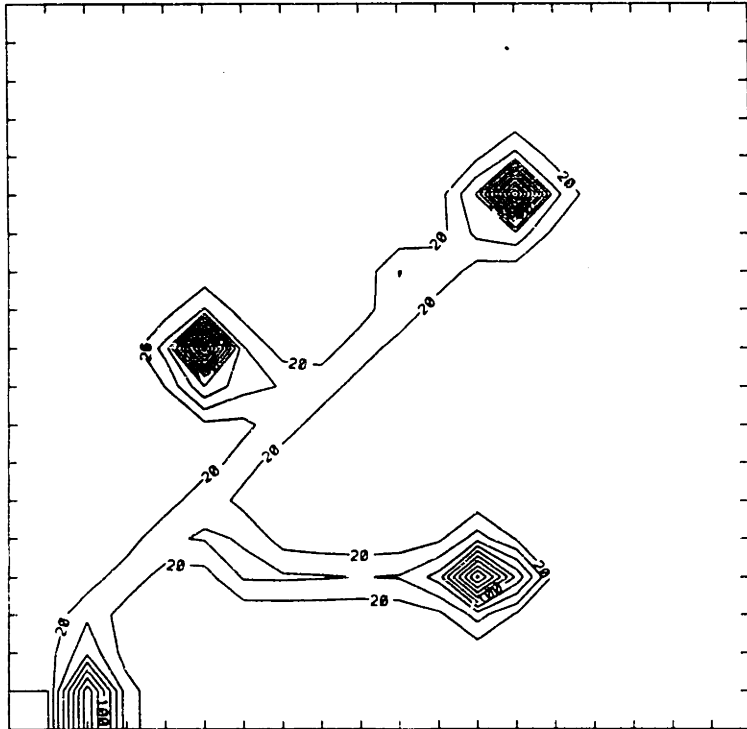
(a) Activator
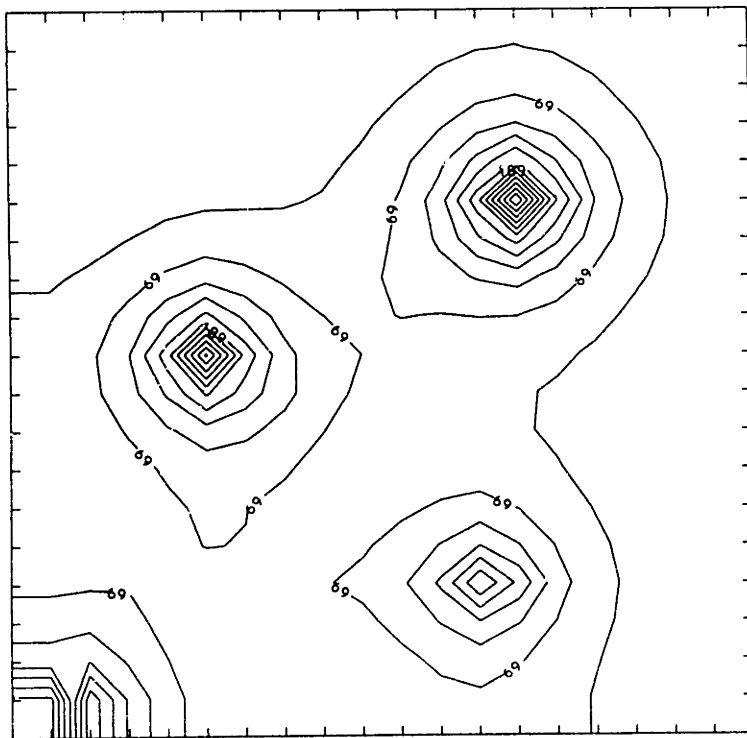
(b) Inhibitor

Figure 4.10: Contours of the states of the Meinhardt equations:
time = 3000
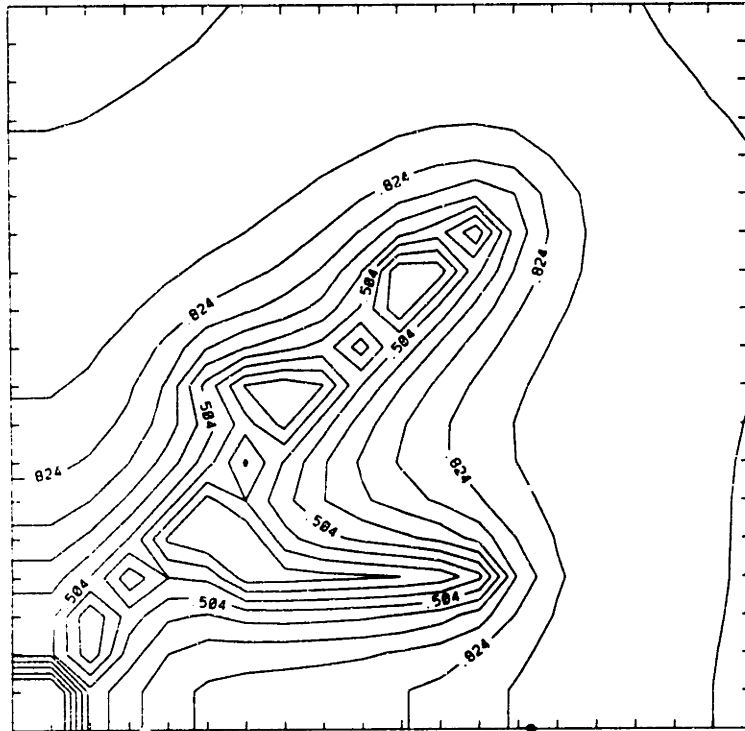
(c) Substrate
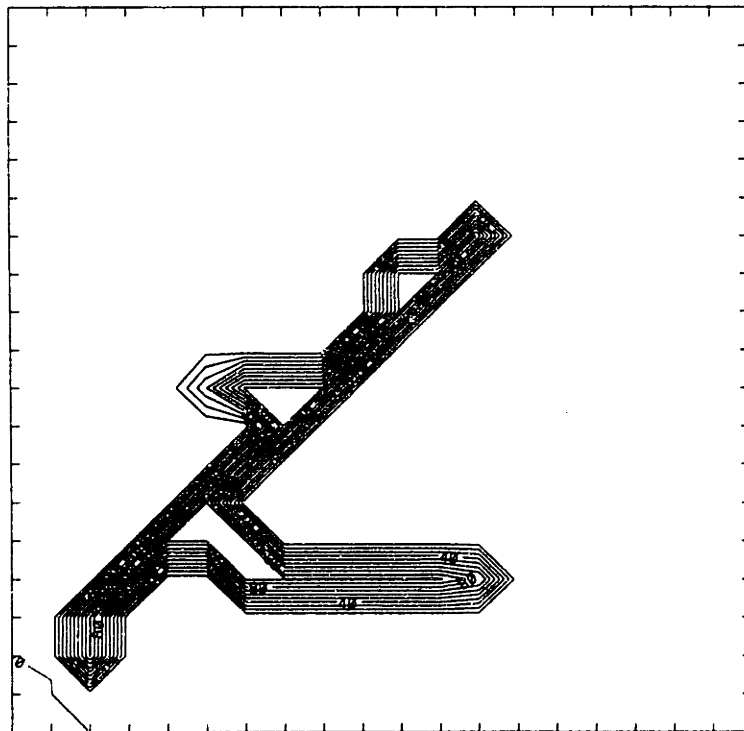
(d) Differentiation

Figure 4.10 (ctd): Contours of the states of the Meinhardt equations:
time = 3000
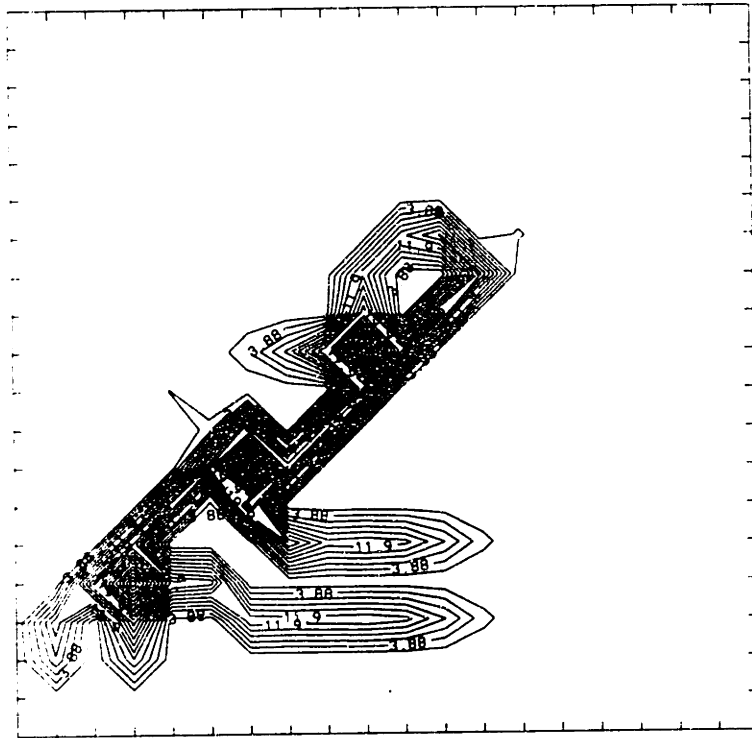
(a) Eigenvalue 1
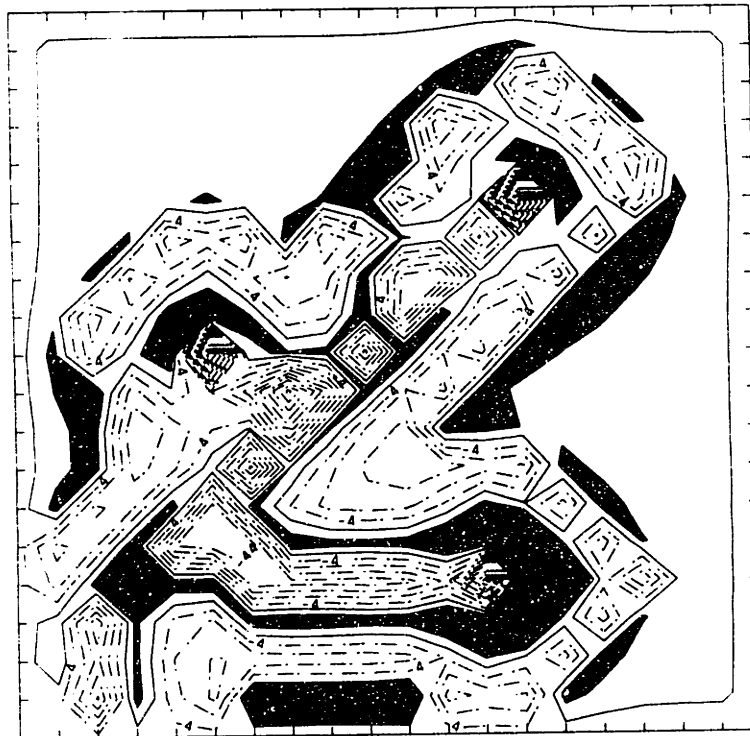
(b) Eigenvalue 2

■ **Positive eigenvalues
(unstable)**

Figure 4.11: Contours of the eigenvalues of the linearized
Meinhardt equations: time = 3000
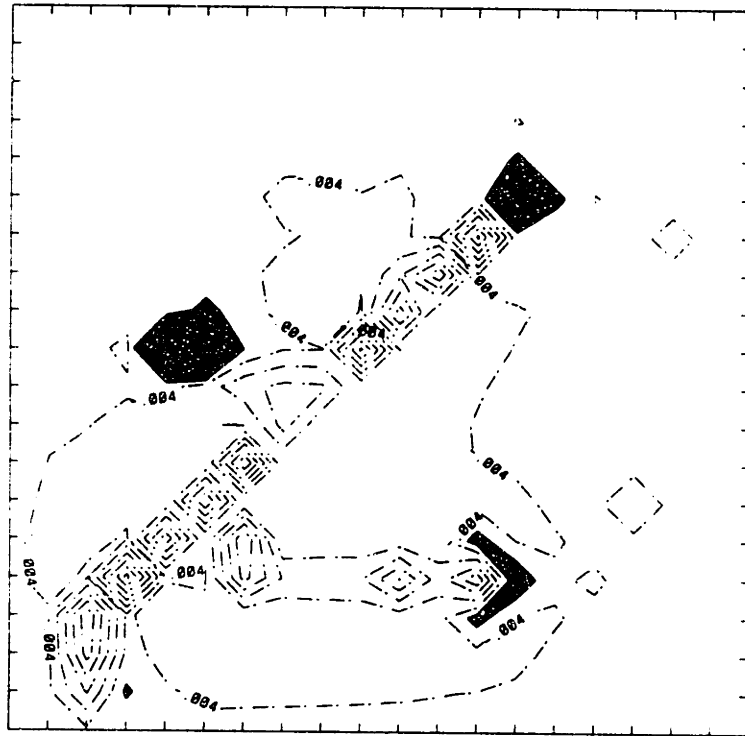
(c) Eigenvalue 3   ■ **Positive eigenvalues**
(unstable)
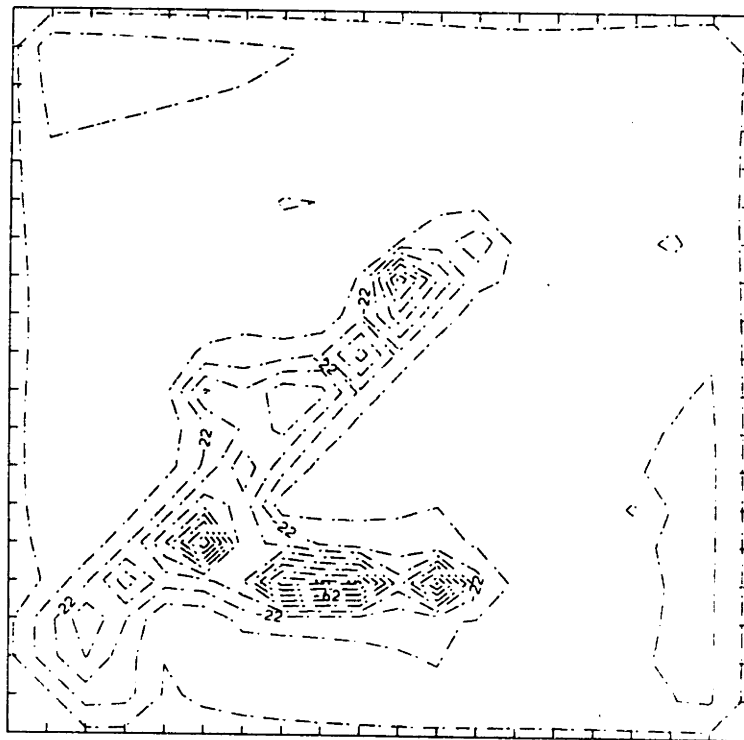
(d) Eigenvalue 4

Figure 4.11 (ctd): Contours of the eigenvalues of the linearized
Meinhardt equations: time = 3000

131

While each eigenvalue corresponds to an eigenvector that is a linear combination of the four states, and this linear combination changes in space, so that a positive value for a particular eigenvalue cannot be categorically attributed to any particular state, it is believed that Figure 4.11a largely corresponds to activator, Figure 4.11b to inhibitor, Figure 4.11c to substrate, and Figure 4.11d to differentiation. Negative eigenvalues correspond to stable Lypanov coefficients (i.e., σ_i < 1) while positive eigenvalues correspond to unstable Lypanov coefficients. The most obvious result is the existence of large regions of stability in the undifferentiated localities and the existence of instability for the points where the network is currently growing for activator and differentiation. The eigenvalue roughly corresponding to substrate is positive everywhere and that for inhibitor negative everywhere. Thus at the tips random effects are very important, which is in accord with observations and explains the sensitivity of the networking process to applied randomness.

We are now in a position to discuss propagation of randomness in the Meinhardt equations. Figure 4.11a indicates that any random fluctuations in activator near the branch head will propagate unstably. Outside this region around the branch head activator fluctuations are suppressed.

Far from differentiated points because the equations are stable the initial perturbations almost disappear, but not entirely. When the lines of differentiation pass through these areas, the residual effects (though extremely small) are propagated unstably in time and are sufficient to result in different patterns being formed for different realizations of the random fields (Figure 4.9e).

Randomness also has other important effects. Since activation, and thus differentiation, are dependent on activator exceeding the activator threshold of Equation (4.2), then addition of random perturbations results in better activation. Where activation would not have occurred, it has, and where it has, it occurs faster when random perturbations are applied. Networks with random components

differentiate and proceed to equilibrium faster than they would otherwise. Also random effects mean that nodes exceeding the activator threshold are less determined by boundary conditions and more by the spatially random perturbations on the autocatalysis parameter, c, which propagate as chaotic fluctuations in activator.

The chaotic effects described in this section are believed to be the cause of the random behavior of the networks exemplified in the introduction to this section. The random effects on parameter c obviously feed directly into activator and thus into the differentiation function. The numerical effects on the other hand are always feeding into the activator equation, but at early times, their cumulative effect is small. Thus the network symmetry at early times. At later times there is sufficient time for chaotic effects to act on the accumulation of errors to send the solution a trajectory entirely different to that of the exact solution (i.e., double precision solution). The effect of the different diffusion formulations only serves to reinforce the idea that accumulation of random effects can dominate network form.

A very important, but unanswered, question is whether the chaotic systems mechanisms that generate the randomness in the networks are specific to the Meinhardt equations or whether they are more generally applicable. Of course, there is no way to be certain of this without testing more general forms of equations but some general and powerful comments can be made.

The differentiation equation (Equation 4.1d) is driven by activator exceeding an activator threshold. Three generic properties that the activator must possess, for networks to form, were proposed in Section 4.3. It has also been noted that the only direct input into the differentiation equation is activator. It is apparent from Equation (4.21a) that for $m_1 > 1$, the autocatalytic behavior of activator will result in Lypanov exponents for activator perturbations being greater than 1. When activator is growing, perturbations in activator are also growing. Thus a result of the autocatalytic behavior is the existence of a sensitive dependence on initial conditions, a central

feature of chaotic systems. If the activator and inhibitor equations are replaced with some other form of activator function, provided that this new activator is also autocatalytic, it is suggested that this new activator function will also show sensitive dependence to initial conditions. If a small change in activator propagates through the system so as to produce a self–reinforcing change in activator, then that activator will be sensitive to random effects.

It is believed that the physical system that will be proposed in Chapter 5 displays this behavior, and that this autocatalytic behavior is an explanation for the random form of channel networks in the field.

4.5 <u>Conclusions</u>

This chapter examined some of the important characteristics of a set of equations developed by Meinhardt (1976, 1982) that have been shown to exhibit networking behavior. The governing physics of these equations is a chemical transport model and the equations were first presented in the context of leaf vein growth within a leaf.

The most important characteristic of these equations is that there is an explicit consideration of a differentiation process in space; the differentiation between normal leaf cells and leaf vein cells. A qualitative analogy between leaf veins and channel networks appears to exist. The differentiation process is shown to be governed by some very complex interactions of chemical transport properties in the leaf cells surrounding the leaf veins. This chapter identified how the underlying physics governed the form of the networks generated. This work suggests that interactions between the channel and the hillslope are central to the growth and form of the channel networks

Using sensitivity studies, it was shown that the physics of the Meinhardt equations were not unique in generating networks of differentiated nodes. All that is

required is that the physics comply with some very general rules. On the basis of the sensitivity studies, three general rules were proposed on the characteristics of the growth mechanisms of a line of differentiated nodes. They were:

1.   A potential growth region must surround and precede the end of a growing branch. This allows the branch to grow.

2.   The potential growth region must move with the branch tip and the growth potential must be actively suppressed immediately behind the growing tip. This forces the branch to grow as a line rather than as a blob.

3.   The potential growth regions of two growing branches must repel each other, and the domain boundaries must repel the potential growth region. This assures that the resultant network is space filling.

Finally, the observed randomness of the generated networks was explored in terms of transient chaos. It was shown that while input randomness results in random networks, the input randomness is not essential for the network growth process. Networks can grow in the absence of random input. The means by which random input comes to dominate the network form was explained in terms of transient chaos near the growing branch tips. Maps of Lypanov coefficients demonstrated that chaotic behavior occurs in regions of active growth and showed how random effects may grow and dominate the form of the generated networks. It was argued, on the basis of the sensitivity studies, that this chaotic behavior may not arise from the specific form of the Meinhardt equation but may, in fact, be a general characteristic of all differential systems that exhibit autocatalysis in the state that governs network growth.

CHAPTER 5
A PHYSICALLY BASED CHANNEL NETWORK
AND CATCHMENT EVOLUTION MODEL

5.1 Introduction

In this chapter a model is presented that realistically simulates the growth and evolution of channel networks and their contributing hillslopes. The realism results from the incorporation of continuity for flow and sediment transport. The constitutive relation relating flow and sediment transport incorporates generally accepted physics. The model explicitly describes the differentiation between channel and hillslopes, and their responses to runoff and sediment transport. This differentiation is the most important physical characteristic of catchments. The interaction of the hillslope and channel regimes over long time scales is also taken into account. The dynamics of channel and network growth are modelled using physically based mechanisms for channel initiation and growth. In addition, other important effects such as tectonic uplift, rockslide, and soil creep are described with their own physics, independently of the sediment transport mechanism.

It is believed that the model presented in this chapter is a good representation of the minimal physics necessary for a branched network of channels to be formed. This will be argued on the basis of the qualitative understanding developed in Chapter 4 regarding the important processes involved in creating networks in the Meinhardt equations. Where necessary the intuition thus gained will be used in formulating the processes that are modelled in this and later chapters.

The philosophical justifications for selectively incorporating processes, rather than incorporating all possible processes, are threefold:

1. The smaller the number of processes involved in the evolution of the networks and catchments, the easier it is to understand the complex

nonlinear interactions that occur.

2.   Perhaps paradoxically, the smaller the number of processes involved, the more general are the results. For each independent process that is modelled, a new nondimensional parameter is created. To compare two catchments using similitude, all nondimensional parameters must be equated. If only dominant processes are modelled, then the number of nondimensional parameters to be matched is minimal, and greatest flexibility is obtained in scaling two catchments for similarity. These similitude conditions are dealt with in detail in Chapter 6.

3.   The smaller the number of processes, the less the computational problems in solving the resulting equations. As will be seen later, the computational problems of solving the equations presented here are quite severe.

The major disadvantage of choosing a minimal physical system is that it is probable that different processes may dominate at different scales. It is believed that all of the important processes at the catchment scale are included, so that this scale problem is not an issue. In particular, a number of processes that are only important at the hillslope scale, or in steep rocky catchments with thin soils, have been ignored. The most important limiting assumption at the hillslope scale is that sediment transport will be considered to be transport limited rather than source limited. That is, the limiting condition on how much sediment is transported is the transport capability, rather than the existence or lack of transportable material.

5.2 The Governing Equations of the Physical Model

This section presents the governing equations that will be used in this work. The equations are used to simulate the growth and evolution of the channel networks and the contributing hillslopes. Justification of the form of the crucial terms in these

equations will be provided in following sections. The formulation of the governing equations implicitly assumes that the equations will be solved on a grid of equally spaced points in the plane, in an analogous fashion to the finite difference solution technique for partial differential equations.

Two variables are solved for in the plane, elevation and an indicator function that identifies where channels exist in space. On the basis of the direction of steepest slope, a flow or drainage direction is assigned to each node. These drainage directions are then used to determine area contributing to (i.e., flowing into) each node. From these areas, and the steepest slopes at the nodes, continuity equations for flow and sediment transport are written. These areas and steepest slopes are also used to obtain overland flow velocities, or the activator, which are then used in the channelization function, as indicators of potential for channelization. Details of the numerical implementation of the solution technique for these equations can be found in Appendix A.

The governing differential equations are:

$$\frac{\partial z_j}{\partial t} = c_{0j} + \frac{1}{\rho_s(1-n)\ L_g^2} \sum_i [(Q_i)^{m_1}(S_i)^{n_1}\ I_{ij}\ f(Y_i)] + D_z\ L_g\ \frac{\partial^2 z_j}{\partial x_i^2} \tag{5.1a}$$

$$\frac{\partial Y_j}{\partial t} = d_t\ [0.0025\ c_1 a_j + (-0.1Y_j + \frac{Y_j^2}{1 + 9Y_j^2})] \tag{5.1b}$$

and the constitutive equations are:

$$a_j = \beta_5 Q_j^{m_5} S_j^{n_5} \tag{5.2a}$$

$$Q_{sj} = f(Y_j) Q_j^{m_1} S_j^{n_1} \tag{5.2b}$$

$$Q_j = \beta_3 A_j^{m_3} \tag{5.2c}$$

$$f(Y_i) = \begin{cases} \beta_1 O_t & Y_i \approx 0 \\ \beta_1 & Y_i \approx 1 \end{cases} \tag{5.2d}$$

where the equations are solved on the domain Ω. A rectangular domain is used in much of the following work so that

$$\Omega = [0, L_1] \times [0, L_2] \tag{5.2e}$$

The variables in Equation (5.1) are

z_j = elevation, at node j

Y_j = indicator variable for channelization, at node j

$\approx \begin{cases} 0 & \text{hillslope node} \\ 1 & \text{channel node} \end{cases}$

a_j = activator of channelization, at node j

Q_{sj} = sediment transport, at node j (units of mass/time)

Q_j = discharge, at node j

t = time

x_i = horizontal distance, in direction i

L_g = grid spacing (equal in both planar directions)

c_{0j} = tectonic input

ρ_s = density of eroded material

n = porosity of material before erosion and after deposition

S_j = slope in steepest downhill direction, at node j

D_z = diffusivity for diffusive transport processes

I_{ij} = indicator function for whether node i drains into or out of node j

$$= \begin{cases} -1 & i=j, \text{ drainage of node } j \\ 1 & \text{node i drains into node } j \\ 0 & \text{node i does not drain into or out of node } j \end{cases}$$

β_1 = rate constant for sediment transport in channels

O_t = ratio of rate constant for sediment transport in the hillslope region to sediment transport in the channels

m_1, n_1 = powers of Q and S in the sediment transport equation

d_t = rate constant for channel growth

$1/c_1$ = threshold coefficient on the activator of channelization

β_5 = multiplicative constant on activator

m_5, n_5 = powers on Q and S, respectively, in the activator equation

β_3, m_3 = multiplicative constant and power, respectively, for relating the characteristic discharge to the characteristic area.

5.3 Explanation of the Governing Equations

The governing equations (5.1) are nonlinear differential equations with two states distributed in space. These two states are elevation and the indicator function for channelization. The most important qualitative characteristic of a catchment, the branched network of channels that form the backbone of the drainage system of a basin, is thus explicitly modelled. There are five important properties distributed in space that are derived directly from these two states. The first three are the steepest downhill slope, contributing area and discharge. The other two important properties are derived from the slopes and discharges and are the distribution of activator and sediment transport in space. The activator and sediment transport distributions feed back into the two state equations for elevation and channelization as inputs. Thus there is a nonlinear interaction between the elevation and channelization, and the

140

activator and channelization distributions in space. This interaction is the central feature of the model.

The differential equation for elevation (5.1a) is a continuity equation in space for sediment transport. Three physically based transport processes are modelled. The most important process is the continuity term for fluvial sediment transport, which is the second term in Equation (5.1a). This term is dependent on discharge and the magnitude of the slope in the steepest downhill direction. Additionally, the magnitude of the fluvial transport is dependent upon whether that point in space is channelized or not. Typically, sediment transport on the hillslopes is much less than that in the channel. This effect is parameterized by the $f(Y)$ term of Equation (5.2d), where O_t is much less than 1. The physics underlying the formulation of the fluvial transport term of Equation (5.2c) is explained more fully in Section 5.4. The form of the summation in Equation 5.2 follows from summing all the sediment inflows and outflows (expressed as mass/(unit volume)), converting them to an effective volume (using the density and porosity) and applying the sediment transport imbalance over the area associated with the node, L_g^2. That is

$$\frac{\partial z}{\partial t} L_g^2 = \frac{\text{sediment inflow} - \text{sediment outflow}}{\rho_s \, (1-n)}$$

The second important term in the elevation equation is the tectonic input; the first term in Equation (5.1a). The form of this term may be quite general with variability allowed both in space and time. For instance, an uplift event, such as may result from an earthquake (see, for example, Morisawa, 1964) can be described by

$$c_0(\underline{x}, t) = \overline{c}_0(\underline{x}) \, \delta(t - t_0)$$

where

$\overline{c}_0$ = the uplift resulting from the tectonic event

t_0 = time at which the event occurred

$\delta(t)$ = dirac delta function

Uplift that occurs continuously with time, but that may be variable in space, such as may result from continuous bulging of the continental crest (see, for example, Havlena and Gross, 1988) can be described by

$$c_0(\underline{x}, t) = \overline{c}_0(\underline{x})$$

where $\overline{c}_0$ = the uplift rate resulting from bulging, variable in space, constant in time.

This tectonic term is important since it is the only mechanism to oppose the continual downwasting of fluvial sediment transport. As will be described in detail in Chapter 6, the so–called "dynamic equilibrium" of landscapes cannot occur without the opposing processes of tectonic uplift and fluvial erosion.

The third term of the elevation evolution equation is the diffusive transport term. As noted in Section 2.3, a number of hillslope transport processes can be modelled by use of a spatially constant diffusion term. The processes that are lumped together in this diffusion term include hillslope soil creep, rainsplash, and rockslide. The literature indicates that the first two of these processes are only important in the regions of the hillslopes near the watershed where discharge, and thus fluvial sediment transport, is small. Rockslide, the third of the diffusive processes, is much more uncertain. Little work has been done on determining the importance of this process at long time scales. Since the adapted model was not primarily formulated for modeling rocky, source limited transport environments, the diffusive process will be largely

ignored in this work, even though the conceptual framework of Equation (5.1a) allows it to be modelled. In Chapters 7 and 8, where the results of computer simulations will be discussed, studies of the sensitivity to this term will be performed.

The channelization equation (5.1b) is the important equation with respect to the development of channels and the extension of the networks. As noted in Chapter 4, the effect of the form of Equation (5.1b) is to make two values of Y stable attractors, 0 and 1. Initially the catchment starts with $Y = 0$ everywhere, a situation corresponding to no channels, only hillslope. When the value of the activator a exceeds a critical threshold called the activation threshold, the value of $Y = 0$ becomes unstable and Y goes into a transition state where it is increasing to $Y = 1$, i.e., that spot in space is in transition from hillslope to channel. When Y reaches a value of 1, Y remains at 1 forever, since the value of $Y = 1$ is stable irrespective of the value of activator. That is, once a channel is formed, a channel cannot be unformed, irrespective of what happens in the catchment after that. The effect of the activator function is solely to trigger the beginning of the channelization process once a given threshold is exceeded. The rate at which a point is channelized once the activator threshold is exceeded is determined by the parameter d_t; a large value of d_t results in the point differentiating quickly. The form of the channelization process in Equation (5.1b) is inspired by the form of the differentiation function in the Meinhardt equations.

The details of how this channelization process leads to network extension and pattern development is a central justification for the work described in Chapter 4. While the quantitative definition of activator in the Meinhardt equations and that of Equation (5.2a) is quite different, the qualitative behavior is very similar. This is particularly true of the spatial distribution of activator around the growing channel heads, and its influence upon the growth of the spatial pattern of channelization or differentiation. In Section 4.3 three conditions were asserted to be necessary conditions

on the qualitative, or topological, spatial distribution of activator if networks were to be formed. These conditions will be restated in the context of Equations (5.1) and (5.2).

The first condition is that regions of high activator must be formed around the growing channel tip. As is shown in Figure 5.1, these regions of high activator result from the localized high slopes and the convergence of the drainage flow patterns around the channel head. Both of these effects result from the larger erosion along the channels, compared to the hillslopes. This preferential erosion in the channels results from the higher sediment transport rates in the channels since $O_t \ll 1$. Simply put, channels erode faster than hillslopes and this behavior is necessary for a channel to grow. If channels do not erode faster than hillslopes, then there is no preferential drainage to incipient channels, and no encouragement through autocatalysis for channelization to occur.

The second condition for network formation was that the region of activation at the channel head must move with the channel head. As shown in Figure 5.2, this results from the capturing of the flow directions around the channel head. In this way discharges behind the channel head are diminished, even though slopes are of comparable magnitudes. Thus activator is lower behind the channel head than at the channel head.

The third condition for network formation was that growing channel heads should repel each other and that growing channel heads should be repelled by a boundary. This requirement ensures that the resulting network is space filling, a commonly assumed, if not observed, feature of channel networks (Abrahams, 1984). The repulsion of growing channel heads, as shown in Figure 5.3, results from an interaction between the drainage patterns and erosion. Provided that in the sediment transport Equation (5.2b) $m_1 > 1$, then the rate of elevation change is positively correlated with discharge. Thus everything else being equal, the region between the
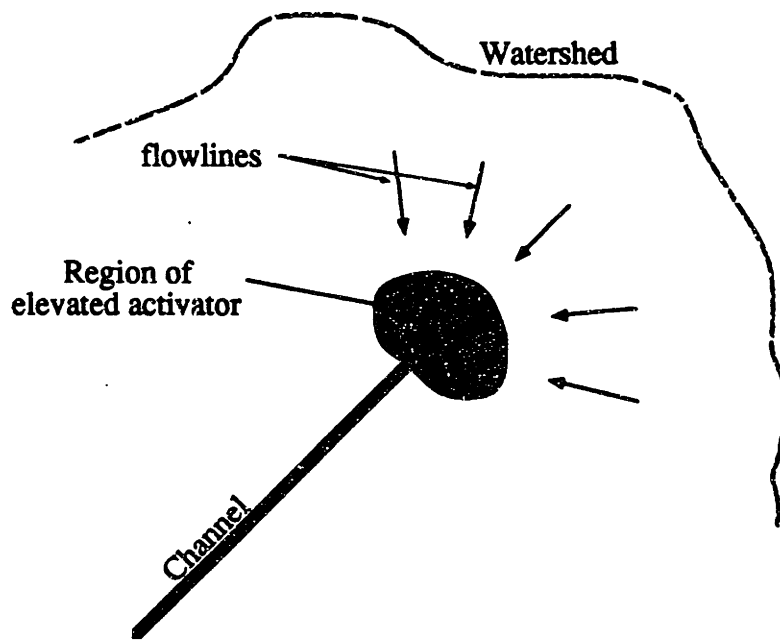
Figure 5.1: High Activator distribution around
the channel head
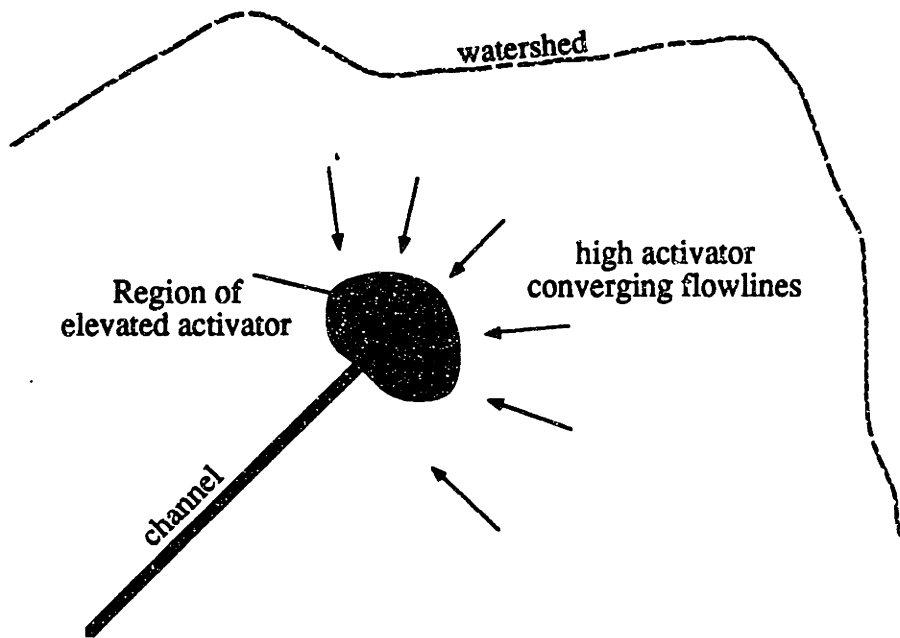
**BEFORE BRANCH HEAD
GROWTH**
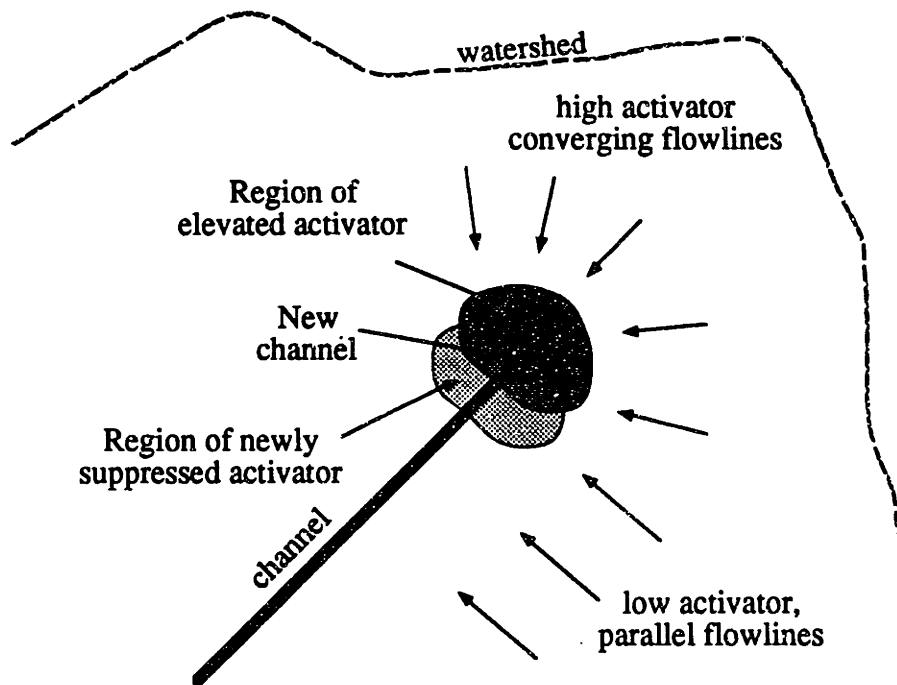
**AFTER BRANCH HEAD
GROWTH**

Figure 5.2: Mechanism for linear growth of channels:
Activator movement with the channel head
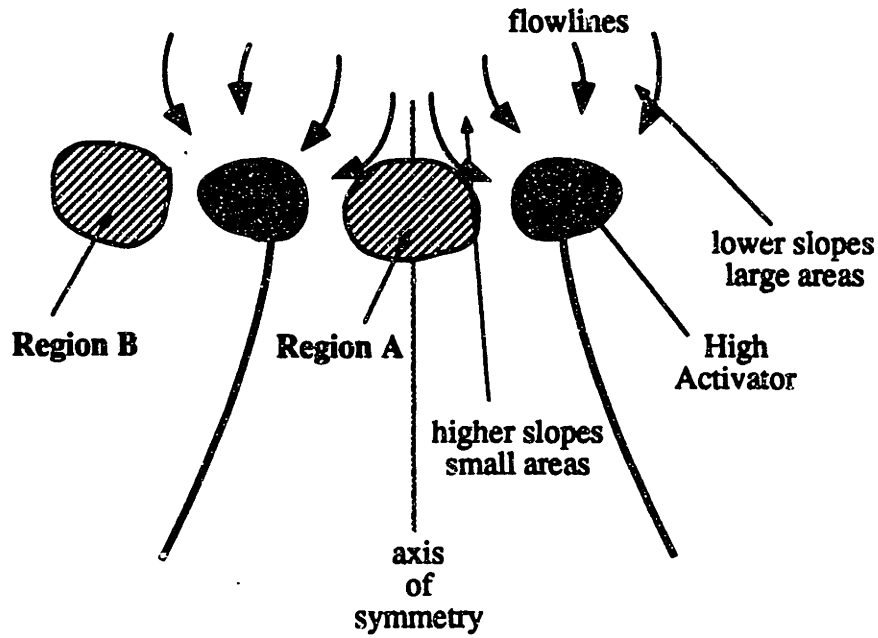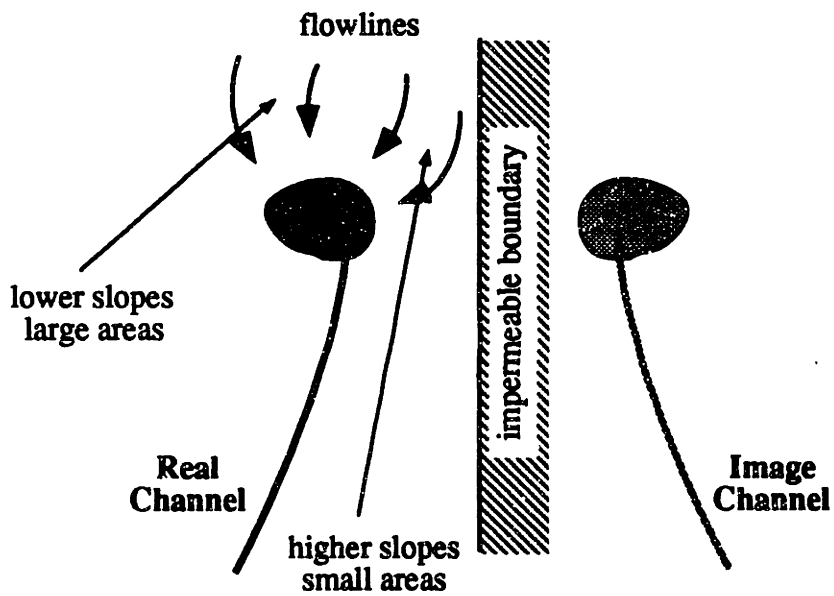
Figure 5.3: Mechanism for space filling by channels
(a) Mutual repulsion by channel heads
(b) Repulsion from boundaries

growing tips (region A) which has lower discharges, will have a slower rate of change of elevation than the region outside the growing tips (region B). Since discharges follow the steepest downhill direction, that means that discharges will be highest on the outside of the growing tip (region B). On the other hand, slopes will be highest inside the growing tips, in region A. It appears that provided $m_1 \approx n_1$ that the higher discharge in region B dominates the higher slope in region A. Thus the highest activator is typically in region B so that growing tips repel each other.

The argument for repulsion from the boundaries proceeds in a similar fashion (Figure 5.3). The boundary is modelled as a zero slope condition perpendicular to the boundary, i.e.,

$$\frac{\partial z}{\partial x_1} = 0 \qquad\qquad x_2 = 0, L_2$$

$$\tag{5.3}$$

$$\frac{\partial z}{\partial x_2} = 0 \qquad\qquad x_1 = 0, L_1$$

for a rectangular solution domain, $\Omega = [0,L_1] \times [0, L_2]$.

Postulate an image channel an equal distance from the boundary, a mirror image of the actual channel. The boundary is a similar flow divide to that which develops naturally in Figure 5.3. The argument for mutual repulsion between the real channel and the image channel proceeds exactly as that above for the repulsion of two real channels. The boundary conditions of Equation (5.3) mean that the boundary of the domain is simply an artificial watershed, identical in all but one respect to that developed naturally in the interior of the domain. The one respect where the watersheds differ is that the domain boundary, which is a straight line, artificially constrains the form of the channel alongside it and reduces its irregularity. This is a common problem of channel network generation schemes that apply fixed boundary

conditions (e.g. van der Tak, 1988).

There is another boundary condition which is important. This is the elevation condition imposed at the catchment outlet. All elevations in this work are defined relative to the fixed elevation imposed at the outlet of the catchment (i.e. the elevation of the notch). Thus the tectonic uplift rate, c_0, is defined as the uplift relative to the elevation of the outlet. As an example consider a small catchment with an outlet on a the flood plain of a very large river. The outlet elevation may be dominated by elevation changes in the floodplain in the large river; i.e. from the point of view of the small catchment the elevation condition at the outlet is externally imposed and variable in time. In this case c_0 for the small catchment is the tectonic uplift relative the flood plain of the large river (i.e. the catchment outlet elevation) not relative to sea level.

The preferential erosion that occurs in the channel and the autocatalytic effect this has on the activator is central to the branch head growth. It also has an interesting consequence for hillslope erosion which is by definition unchannelized. For the hillslope to be unchannelized, it must be that there is no autocatalytic tendency to channelize. Something must be happening on the hillslope to stabilize the tendency to channelization. This may result from there being no tendency for channels to erode faster than the hillslope. If channels do not erode faster, then there is no positive reinforcement for channelization. This may result from increased inputs to sediment transport from rainsplash (a diffusive process in Equation 5.1) of the same order of the fluvial erosion. Thus when fluvial erosion on hillslopes is comparable to the stabilizing rainsplash effects, the self—reinforcing channel erosion is dominated by the channel destroying diffusion. This mechanism is consistent with the first condition for network formation, and the observed physics (Dunne and Aubrey, 1986). It would also explain the disturbing conclusion of Smith and Bretherton (1972) who found, in a model without diffusion, that channelization must occur in all regions that are undergoing

active erosion.

It was noted in Chapter 4 that the idea of Y=0 representing hillslope and Y=1 representing channels is only approximate. In particular, at a growing channel head there is a period of time when the hillslope is in transition from hillslope to channel; i.e. Y is between 0 and 1. Points intermediate between hillslope and channel have sediment transport properties that are intermediate between that for hillslope and that for channel. The adopted transition is of the form

$$
f(Y) = \begin{cases} \beta_1 \, O_t & \alpha Y \leq 0.1 \\ \beta_1 \, [\; O_t + (1-O_t)(Y-0.1) \frac{Y}{0.9}] & \alpha Y \leq 1 \\ \beta_1 & \alpha Y \geq 1 \end{cases}
$$

Where α = a model parameter greater than 1

The model is insensitive to changes in α. A detailed discussion of numerical issues related to the solution of the governing equations is provided in Appendix A.

5.4 Physical Justification of the Sediment Transport Equation

The generic sediment transport formula used in this work is

$$
Q_s = \beta_1 \, Q^{m_1} S^{n_1} \tag{5.4}
$$

where

Q_s = sediment transport, in mass/time

Q = discharge

S = channel or hillslope slope

Though a sediment transport equation of this form has been used by geomorphologists in previous work (Smith and Bretherton, 1972), this formulation of the sediment transport equation is unconventional compared with the form used by specialists in fluvial sediment transport (e.g., Vanoni, 1975). The unconventional formulation was adopted because it has a number of significant computational advantages over the traditional formulation (see Appendix A for details). This section aims to show how the new formulation may be obtained from the Einstein–Brown equation, a commonly accepted fluvial sediment transport formula. It will be shown that a minimal number of simplifying assumptions are required.

In addition, it will be shown how the Einstein–Brown equation, an instantaneous sediment transport relation, can be converted into a mean temporal sediment transport relation for large timescales. It will be shown that the simple form of Equation (5.4) can be maintained under this temporal averaging, and that only the coefficient β_1 is modified with coefficients m_1, n_1 unchanged. Thus the temporal averaging results in the rate of sediment transport being changed; its form dependent on m_1, n_1 is unchanged.

This work adopted the Einstein–Brown sediment transport equation. This equation captures the important dependencies of sediment transport; that is, its dependence on depth, velocity, sediment size, and channel geometry.

The Einstein–Brown equation is expressed in terms of a non–dimensional sediment transport ϕ, and a nondimensional shear stress, $\frac{1}{\psi}$. Vanoni (1975) gives the governing equation as

$$\phi = 40 \left(\tfrac{1}{\psi}\right)^3 \tag{5.5a}$$

where

$$\phi = \frac{q_s}{\gamma_s \; F_1 \; \sqrt{g(s - 1) \; d_s^3}} \qquad (5.5b)$$

$$\frac{1}{\varphi} = \frac{\tau_0}{\gamma(s - 1) \; d_s} \qquad (5.5c)$$

$$F_1 = \sqrt{\frac{2}{3} + \frac{36\nu^2}{gd_s^3(s - 1)}} - \sqrt{\frac{36\nu^2}{gd_s^3(s - 1)}} \qquad (5.5d)$$

and the notation used is:

q_s = sediment discharge, mass/time/(unit width)

s = specific gravity of sediment

γ = ρg = specific weight of water

d_s = a representative diameter for the sediment particle. Normally d_{50}, the 50 percentile diameter, is used.

g = acceleration due to gravity

τ_0 = γRS = bottom shear stress

R = hydraulic radius

S = bed slope

ν = kinematic viscosity of water.

If the sediment is considered homogeneous throughout the catchment this equation may be simplified to yield

$$q_s = F_2 \; (RS)^3 \qquad (5.6a)$$

where

$$F_2 = 40\, \gamma_s\, F_1 \sqrt{g(s-1)\, d_s^3} \left[\frac{1}{(s-1)\, d_s}\right]^3 \tag{5.6b}$$

$$= \text{constant}$$

Experimental evidence suggest that the exponent in Equation (5.6a) is variable so that a more general formulation is used below where

$$q_s = F_2(RS)^p \tag{5.7}$$

Equation (5.7) is not in the form of Equation (5.4) so that this equation must be reformulated so that it is dependent on discharge Q rather than hydraulic radius. This is achieved by use of Mannings equation for discharge

$$Q = \frac{R^{5/3} S^{1/2} P}{n} \tag{5.8}$$

where   P   = wetted perimeter of flow

           n   = Mannings roughness coefficient

Eliminating the hydraulic radius R from Equations (5.7) and (5.8) yields a sediment transport equation that is true for all cross—sections

$$Q_s = F'(P)\, Q^{\frac{3p}{5}}\, S^{\frac{7p}{10}} \tag{5.9}$$

where $F'(P) = F_2 \left(\frac{n}{P}\right)^{\frac{3p}{5}}$

Note that the multiplicative constant F' is dependent upon the wetted perimeter. The form of this dependence depends on the channel geometry of the flow.

A number of different channel geometries have been examined including

1.   a wide channel with uniform depth across the cross–section

2.   overland flow/unit width

3.   a triangular channel with sideslopes a_1

4.   a general channel cross–section of the form $y = a_1 |x|^{b_1}$, where a_1 and b_1 are variable.

The detailed derivations for each of these cases are presented in Appendix C. The simplest case, a wide channel, will be used to illustrate the techniques involved. In this case the governing sediment equation for the wide channel is

$$Q_s = F' \left(Q^{\frac{3p}{5}} S^{\frac{7p}{10}} \right) \tag{5.10}$$

where $\qquad F' = F_2 \left(\frac{n}{w}\right)^{\frac{3p}{5}}$

$\qquad\qquad w$ = width of the channel.

The multiplicative constant F' is dependent, in a well defined way, on flow geometry and sediment characteristics. F' is constant because the wetted perimeter is independent of flow depth. For the specific example of the Einstein–Brown equation ($p = 3$), Equation (5.10) simplifies to

$$Q_s = F' Q^{1.8} S^{2.1} \tag{5.11}$$

where $F' = [F_2 \left(\frac{n}{w}\right)^{1.8}]$

Note that for the wide channel Equations (5.10) and (5.11) are exact and require no approximation. This is also the case for overland flow/unit width and for the triangular channel. However, some very small approximations are used to

reformulate the sediment transport for the general cross–section into the form of Equation (5.4).

For comparison purposes, the exact sediment transport formulae for the wide channel and a triangular channel (with side slope a_1) are tabulated in Table 5.1 together with those for a generalized cross–section. The channel cross–section has very little effect on the functional dependence of sediment transport to channel slope. The cross–section dependence is most apparent in the functional form of the discharge, with a lesser dependence on the multiplicative constant. Table 5.1 indicates the range of values that m_1 and n_1 may take in Equation (5.4). That is

$$m_1 \quad \epsilon \quad [0.375p, 0.6p]$$

$$n_1 \quad \epsilon \quad [0.7p, 0.813p]$$

$$\frac{m_1}{n_1} \quad \epsilon \quad [0.413, 0.857]$$

Moore and Burch (1987) used unit stream power theory and the experimental data of Mosley (Schumm, et. al., 1987) to derive a sediment transport equation for rills and hillslopes. Their equation was

$$Q_s \quad \begin{aligned} &\sim Q^{1.6} S^{1.8} \\ &\sim (Q^{0.53} S^{0.6})^p \end{aligned}$$

where $p = 3$.

This equation corresponds well with a case intermediate between the wide channel and the variable geometry channel (Table 5.1).

The derivations summarized above apply only to the instantaneous sediment discharge this being the typical use of the Einstein–Brown equations. It remains to be

TABLE 5.1

Sediment Transport Equations

Channel Geometry Parameters a_1	b_1	Sediment Transport Formulas
1	2	$\left[F_2 \dfrac{n^{0.421p}}{2.24^p} \right] Q^{0.421p} S^{0.789p}$
1	3	$\left[F_2 \dfrac{n^{0.467p}}{2.29^p} \right] Q^{0.467p} S^{0.767p}$
1	4	$\left[F_2 \dfrac{n^{0.495p}}{2.36^p} \right] Q^{0.495p} S^{0.753p}$
3	2	$\left[F_2 \dfrac{n^{0.444p}}{2.46^p} \right] Q^{0.444p} S^{0.778p}$
3	3	$\left[F_2 \dfrac{n^{0.487p}}{2.67^p} \right] Q^{0.487p} S^{0.757p}$
3	4	$\left[F_2 \dfrac{n^{0.512p}}{2.91^p} \right] Q^{0.512p} S^{0.744p}$
10	2	$\left[F_2 \dfrac{n^{0.455p}}{2.66^p} \right] Q^{0.455p} S^{0.773p}$
10	3	$\left[F_2 \dfrac{n^{0.496p}}{3.21^p} \right] Q^{0.496p} S^{0.752p}$
10	4	$\left[F_2 \dfrac{n^{0.518p}}{3.66^p} \right] Q^{0.518p} S^{0.741p}$

Triangular Channel (a_1=sideslope)	$\left[F_2 \dfrac{n^{0.375p}}{2.38} \dfrac{a_1}{1+a_1^2} \right] Q^{0.375p} S^{0.813p}$
Wide Channel	$\left[F_2 \left(\dfrac{n}{w}\right)^{0.6p} \right] Q^{0.6p} S^{0.7p}$
Henderson (1966)	$\beta Q^2 S^2$
Moore and Burch (1987)	$\beta Q^{1.6} S^{1.8}$

demonstrated that Equation (5.4) is also a satisfactory representation of the mean temporal sediment discharge over long time scales.

The time scales of interest in this work are typically of the order of thousands of years. By temporal averaging over the distribution of flood hydrographs, where Equation (5.4) describes the instantaneous sediment discharge, a modified version of Equation (5.4) can be obtained for the mean temporal sediment discharge. In the process a new value of the multiplicative constant β is obtained that is dependent on the moments of the distribution of flood events and the discharge, Q, used is the mean annual discharge from a flood frequency analysis.

Hereafter, to avoid confusion of notation we will use the parameters of Equation (5.4), i.e., β_1, m_1 and n_1 to represent those parameters applicable to the mean temporal sediment discharge equation and the equation

$$Q_s = \beta_2 \, Q^{m_2} S^{n_2} \tag{5.12}$$

with its parameters β_2, m_2 and n_2 to represent those parameters applicable to the instantaneous sediment discharge equation. What follows is a brief summary of the averaging over time of the instantaneous sediment transport to obtain the mean temporal sediment transport equation. The complete details are provided in Appendix C.

Consider a single flood hydrograph described by the discharge with time Q(t), with a characteristic duration T_p, and a characteristic discharge Q_p (Figure C.3), so that

$$t' = \frac{t}{T_p}$$
$$Q'(t) = \frac{Q(t)}{Q_p}$$

The sediment load of a single hydrograph can then be expressed as

$$Q_s = [\, \beta_2 \, T_p \int_{-\infty}^{\infty} (Q'(t'))^{m_2} \, dt' \,] \, Q_p^{m_2} \, S^{n_2}$$

By considering the peak discharge, Q_p, and the time of duration of the flood, T_p, as random variables then the sediment transport per hydrograph, Q_s, is a random variable. If, in addition, the rate at which the hydrographs occur with time is also a random variable, independent and Poisson distributed in time, then the average sediment transport rate is given by

$$\overline{Q}_s = [\beta_2 \, \overline{T}_p \, \lambda \int_{-\infty}^{\infty} (Q'(t'))^{m_2} \, dt'] \cdot$$

$$\left[1 + m_2(m_2 - 1) \frac{\sigma_{Q_p}^2}{\overline{Q}_p^2} + m_2 \frac{\sigma_{Q_p T_p}^2}{\overline{Q}_p \, \overline{T}_p} \right] \overline{Q}_p^{m_2} \, S^{n_2} \qquad (5.13)$$

where $\overline{T}_p$ =   mean length of hydrographs

λ =   rate at which hydrographs arrive with time

$\overline{Q}_p$ =   mean peak discharge over all the hydrographs that carry significant sediment load. Derived from flood frequency analysis

$\sigma_{Q_2}^2$ =   the variance of the peak discharge, derived from flood frequency analysis

$$\sigma^2_{Q_p T_p} \quad = \quad \text{covariance between the peak discharge and the length of}$$

the flood hydrograph.

It should be noted that this equation is in the form of the generic sediment transport equation where the correspondence between coefficients in the instantaneous equation (5.12) and the time averaged equation (5.4) is as follows.

$$m_1 = m_2$$

$$n_1 = n_2$$

$$\beta_1 = \beta_2 \left[\overline{T}_p \lambda \int_{-\infty}^{\infty} (Q'(t'))^{m_2} dt' \right] \cdot$$

$$\left[1 + m_2(m_2-1) \frac{\sigma^2_{Q_p}}{\overline{Q}_p^2} + m_2 \frac{\sigma^2_{Q_p T_p}}{\overline{Q}_p \overline{T}_p} \right]$$

A more complete expression for β_1 that allows for skewed distributions of peak discharges (e.g. the log–Pearson Type III distribution) is given by

$$\beta_1 = \beta_2 \left[\overline{T}_p \lambda \int_{-\infty}^{\infty} (Q'(t'))^{m_2} dt' \right] \cdot$$

$$\left[1 + m_2(m_2-1) \frac{\sigma^2_{Q_p}}{\overline{Q}_p^2} + m_2(m_2-1)(m_2-2) \frac{\gamma_{Q_p} (\sigma^2_{Q_p})^{3/2}}{\overline{Q}_p^3} + \right.$$

$$\left. m_2 \frac{\sigma_{Q_p T_p}}{\overline{Q}_p \overline{T}_p} \right] \qquad\qquad (5.14)$$

where

$$\gamma_{Q_p} = \text{skewness coefficient of } Q_p$$

$$= \frac{\dfrac{1}{N} \displaystyle\sum_{i=1}^{N} (Q'_p)_i^3}{(\sigma_{Q_p}^2)^{3/2}}$$

N = total number of flood events

Note that the incorporation of skewness only modifies the value β_1, and not the functional form (i.e. m_1 and n_1) of the sediment transport equation.

The major difference between the instantaneous and time averaged equations is the definition of the discharge. The dependence of the mean temporal sediment transport on the mean peak discharge is important. The discharge, $\overline{Q}_p$, in Equation (5.13) is the average of all the peak discharges of flood events that carry significant sediment loads. Thus $\overline{Q}_p$ is the mean peak discharge from a flood frequency analysis based on exceedance series, with the lower cutoff on discharge being that below which sediment load is insignificant.

Finally the differentiation between the rates of sediment transport in the channels and on the hillslopes, factor O_t, needs to be justified. From Mannings equation for discharge (Equation 5.8) if the discharge is fixed then an increase in flow depth is balanced by a reduction in the wetted perimeter (all other factors being equal). This decrease in wetted perimeter causes an increase in sediment discharge (Equation 5.9). Thus as the flow concentrates and becomes deeper the sediment transport increases. It is this effect that is conceptually modeled by the coefficient O_t. Equation (5.9) indicates that the two important physical determinants of O_t are flow cross–section geometry and sediment properties (parameterized by F_2). The latter may in turn be affected by the flow velocities. Clearly determining O_t apriori is a difficult problem.

Some experimental evidence exists to support the use of $O_t < 1$. Priest, et al. (1975) measured sediment transport rates on the hillslope immediately upstream of a gully head and then again in the gully immediately downstream of the head. These data suggest a value of O_t about 0.3 for an actively cultivated catchment. They found that hillslope conservation measures (e.g. vegetation growth) did not significantly reduce the gully sediment transport. If, in this case, hillslope sediment transport rates are reduced then this would suggest that O_t in natural catchments should be less than 0.3.

The interpretation of discharge, Q, varies from the channel to the hillslope. For channel it is just the discharge. For the hillslope it is the discharge/(unit width) times the effective width of a hillslope node, L_g. The numerical code implicitly assumes $L_g=1$.

In conclusion, this section has shown three things:

1. A commonly accepted instantaneous total load sediment transport equation, Einstein–Brown, may be reformulated into the functional form used in this work and thus this functional form may be related to experimentally measurable quantities.

2. The instantaneous sediment transport equation, in the functional form adopted in this work, can be time averaged to give a mean temporal sediment transport equation, again in the functional form used in this work. The discharge in the new mean temporal equation is the mean peak discharge. The transport coefficient β is dependent upon the distribution of flood peaks, parameterized by mean annual peak discharge.

3. There is evidence to support the use of different sediment transport rate. coefficients in the channel and the hillslope parameterized in the model by coefficient O_t.

5.5 Physical Justification of the Activator Mechanism

This section will provide a linkage between the types of processes that are observed to trigger or activate channelization in the field, and Equation (5.2a) which is used to represent the activator mechanism in this work. It should be stated at the outset that the state of the art understanding of these processes is, at best, primitive. It may be stated that a qualitative understanding of the processes at work is developing at the current time. Quantitative understanding of the range of processes is less advanced, and is either the subject of current research, such as groundwater induced stream growth (Dunne, 1988), or has yet to be fully addressed, such as is the case for overland flow induced channel growth.

The generic equation used to represent the channelization activator is

$$ a = \beta_5 \, Q^{m_5} S^{n_5} \tag{5.15} $$

As noted in Section 5.3, the purpose of the activator in the channelization equation (5.1b) is to trigger the one way process modeling the transition from hillslope to channel. This process is triggered at the time when the activator exceeds the channelization threshold at that point. Once this threshold is exceeded, the channelization process proceeds at a rate governed by the timescale for channelization, d_t. The importance of the activator is in the triggering of the channelization, rather than in governing the rate of channelization. Once channelization is triggered, growth occurs independently of the activator level.

The question that this section addresses is: What physical processes can Equation (5.15) be claimed to simulate given the use activator and its threshold are put to in the model of Equation (5.1)? Some examples of physical processes follow.

A number of different physical processes will be examined that could trigger channel head advance. One of the most common criteria for the design of erosion works is overland flow velocity. Many engineering handbooks give tabulations of allowable velocities for various forms of ground cover or erosion protection (e.g., Henderson, 1966). We will show that overland flow velocity can be expressed in the generic form of activator in Equation (5.15).

If the wide channel assumption is made so that hydraulic radius, R, is equal to flow depth, y, and the wetted perimeter, P, is independent of discharge, then Mannings equation can be written as

$$v = \frac{R^{2/3} \, S^{1/2}}{n} \qquad (5.16)$$

and the discharge for a wide channel of width w can be written as

$$Q = \frac{R^{5/3} \, S^{1/2} \, w}{n} \qquad (5.17)$$

Combining these equations yields

$$v = \beta_5 \, Q^{0.4} \, S^{0.3} \qquad (5.18)$$

where $\beta_5 = \dfrac{1}{w^{2/5} n^{3/5}}$

A similar expression for velocity in a triangular channel may also be derived where the exponents are different.

Another important concept in modern sediment transport theory is the concept of a threshold bed shear stress, below which no sediment transport takes place. This

concept was first clearly expressed by Shields (1936) (Vanoni, 1975) in his now famous Shields diagram which relates bottom shear stress with sediment transport. Thresholds on shear stress are included in transport formulae due to Duboys (in 1897), Shields (in 1936) and Laursen (in 1958), among others.

For a wide channel, or overland flow region, the bottom shear stress τ_* is given by

$$\tau_* = \gamma \, R \, S \qquad\qquad (5.19)$$

In a similar fashion to the derivation for overland flow velocity we obtain an activator expression of

$$\tau_* = \beta_5 \, Q^{0.6} \, S^{0.7} \qquad\qquad (5.20)$$

where $\beta_5 = \gamma \left(\frac{n}{w}\right)^{3/5}$

A similar expression for shear stress in a triangular channel may also be derived.

Dunne (1969) proposed a conceptualization of a groundwater process where groundwater streamtubes converged onto a seepage face at a channel head causing channel erosion. This conceptualization of gully advance is supported by other field work (e.g. Priest, et al., 1975). Dunne (1989) suggested a threshold on hydraulic gradient above which erosion at the seepage force will occur by piping.

$$\left[\frac{dH}{dx}\right]_{\text{threshold}} = (s{-}1)(1{-}n) \qquad\qquad (5.21)$$

where

$\dfrac{dH}{dx}$ = groundwater hydraulic gradient at the seepage face

s = specific gravity of the sediment material

n = porosity

Using Darcy's law for groundwater flow at the seepage face yields an activator formulation of

$$\frac{dH}{dx} = \left[\frac{1}{Khw}\right] Q \tag{5.22}$$

where

K = hydraulic conductivity

h = height of the seepage face

w = width of the seepage face

In summary three different and physically based mechanisms for controlling channel growth have been examined. All three of these mechanisms can be formulated in the form of the generic activator equation of Equation (5.15). In each case the nonlinear dependence of the activator mechanism on discharge and slope followed directly from the physics of the activation mechanism. In all cases the nonlinear dependence on discharge and slope was different.

The differentiating feature between the different activation mechanisms is the ratio (m_5/n_5). In Appendix C techniques developed in Chapter 6 are used to show that activator functions with the same value of (m_5/n_5) are equivalent and differ only in a transformation on the activator threshold above which channelization occurs. Table 5.2 tabulates the activator mechanisms that have been examined, their governing equations, and their (m_5/n_5) values.

An important observation must be made about the definition of discharge. A single definition is used in both the activator and sediment transport equations. In

TABLE 5.2

Activator Equations $(a = \beta_5 \, Q^{m_5} \, S^{n_5})$

Mechanism	Governing Equation [*]	m_5/n_5
1. Overland flow velocity (wide channel)	$v = \left[\dfrac{1}{w \, n^{3/5}} \right] Q^{0.4} \, S^{0.3}$	1.33
2. Overland flow velocity (/unit width)	$v = \left[\dfrac{1}{n^{3/5}} \right] q^{0.4} \, S^{0.3}$	1.33
3. Overland flow[+] velocity (triangular channel)	$v = \left[\dfrac{a_1}{4(1+a_1^2)n^3} \right]^{0.25} Q^{0.25} \, S^{0.3}$	0.67
4. Overland flow shear stress (wide channel)	$\tau = \left[\gamma(\tfrac{n}{w})^{3/5} \right] Q^{0.6} \, S^{0.7}$	0.86
5. Overland flow shear stress (/unit width)	$\tau = \left[\gamma n^{3/5} \right] q^{0.6} \, S^{0.7}$	0.86
6. Overland flow shear[+] stress (triangular channel)	$\tau = \left[\dfrac{n a_1 \gamma^{8/3}}{4(1+a_1^2)^{1/2}} \right]^{3/8} Q^{0.375} \, S^{0.813}$	0.46
7. Groundwater stream sapping	$\dfrac{dH}{dx} = \left[\dfrac{1}{Khw} \right] Q$	∞
8. Montgomery and Dietrich (1988)	$a'_{threshold} = Q \, S^{2.5 m_3}$	$\dfrac{0.40}{m_3}$

[*] For notation, see the text.

[+] a_1 = sideslope of the triangular channel

Section 5.4 it was noted that for sediment transport this discharge may be interpreted in two different ways, depending on the interpretation to be placed on the elevation evolution equation (5.1a). They are:

1. If Equation (5.1a) is to be interpreted as the instantaneous change of elevation with time, then Q is the instantaneous discharge.

2. If Equation (5.1a) is to be interpreted as the mean change of elevation with time, elevation perturbations being averaged out, then Q is the mean of all peak discharges of all significant flood events, $\overline{Q}_p$.

Activator must be interpreted in an analogous fashion. In the former case, activator is considered to be that activator occurring instantaneously in time. In the latter case, activator is considered to be the mean effective activator occurring at a point over many storm events. The mean effective activator must be parameterizable on the mean peak discharge $\overline{Q}_p$. The question of what constitutes the mean temporal activator of channelization, over many flood events, is very poorly understood, and has not been satisfactorily addressed in the literature. Calver (1978) provided a preliminary understanding on the basis of a conceptualization of the instantaneous channel growth mechanism. The lack of theoretical basis for the conceptualization, and lack of any experimental evidence, make this work somewhat speculative.

There is some experimental evidence to support a channelization activation mechanism of the functional form proposed here. Montgomery and Dietrich (1988) present data relating source area versus local slope on the surrounding hillslope for a large number of channel heads in California. This data is fitted very well by an equation of the form

$$A\,S^{2.5} = C \qquad\qquad (5.23)$$

where C   = constant

The ratio of m_5/n_5 for this data $0.4/m_3$. Recall that m_3 is the power on area in the relationship between discharge and area (Equation 5.2c). If $m_3 = 1$ then this value is substantially lower than any of the unrilled surfaces, but close to values for a triangular channel. A detailed discussion of the data of Montgomery and Dietrich (1988) is presented in Section 8.4.

If the proposed activator and autocatalytic channelization model is correct all hillslopes will have a value for activator which is less than the activator threshold. Digital terrain data for a catchment was analyzed to determine if the channel network could be identified from the areas and slopes of the terrain data. The techniques used for determining areas and slopes are those described in Tarboton, et al. (1988) and are a variant of the techniques described in Appendix A for determination of area and slope in the simulation model. The catchment analyzed was W15, a $23km^2$ catchment in Walnut Gulch, Arizona, used by Tarboton, et al. Figure 5.4a shows that region of the catchment where

$$A\ S^{0.75} \qquad \geq 3.85 \qquad\qquad\qquad (5.24)$$

which corresponds to the region that exceeds the overland flow activator threshold when discharge is proportional to area (Equation 5.18). Figure 5.4b shows that region of the same catchment where

$$A\ S^{1.17} \qquad \geq 2.79 \qquad\qquad\qquad (5.25)$$

which corresponds to the region exceeding the overland shear stress activator threshold when discharge is considered proportional to area (Equation 5.20)

A rough network can be discerned in a background of disconnected pixels in
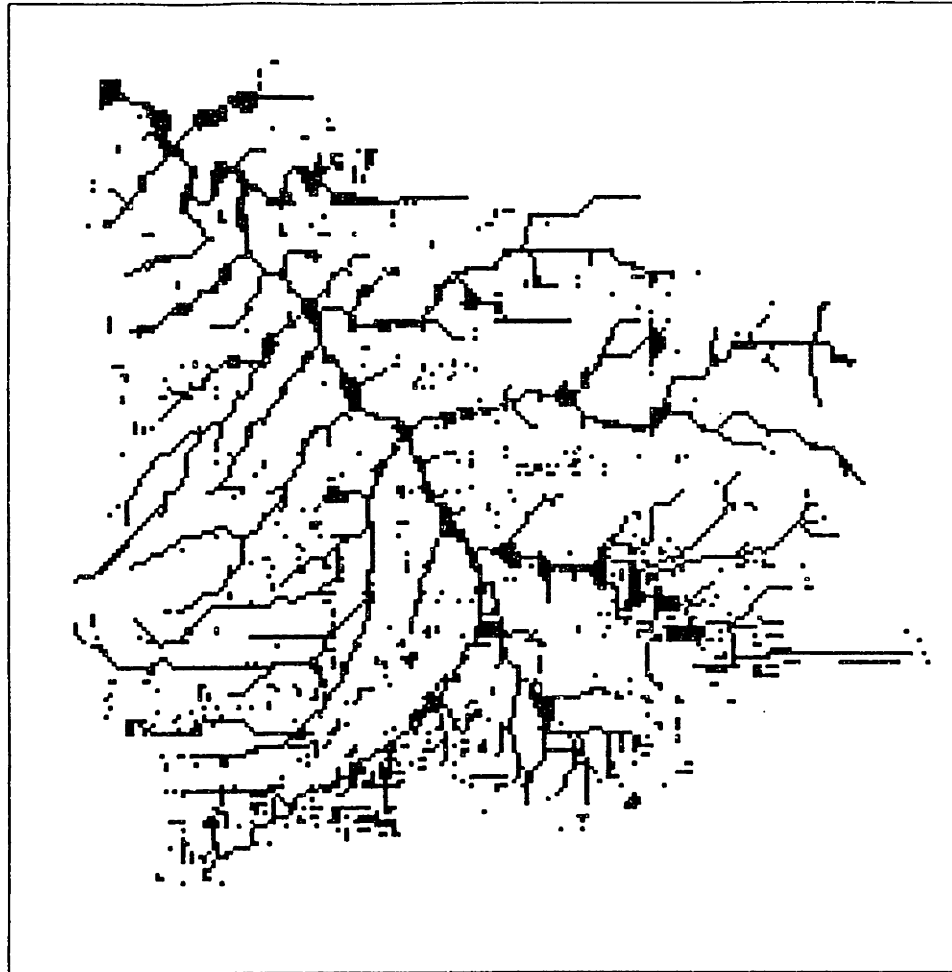
(a) Regions where $A S^{0.75} \geq 3.85$
(activator $\equiv$ average velocity)

Figure 5.4: Activator distribution for digital terrain data of catchment W15
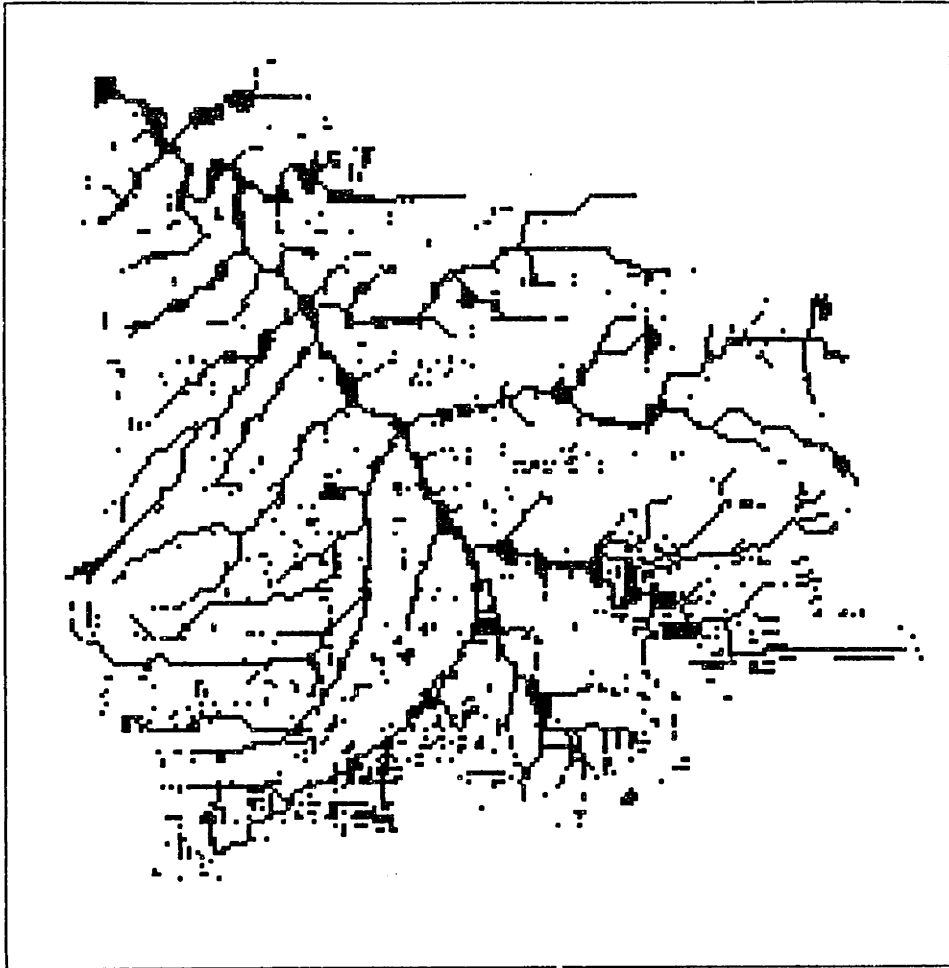
(b) Regions where $A S^{1.17} \geq 2.79$

(activator $\equiv$ bottom shear stress)

Figure 5.4 (ctd): Activator distribution for digital terrain data of catchment W15

these figures. The disconnected pixels arise from noise in the elevations in the USGS data elevation error (nominally ± 8m on a 30m square grid). This noise creates randomly distributed pixels with anomalously high slopes so the the activator for a pixel may fall above the threshold. This observation is true even if the thresholds in Equations (5.24) and (5.25) are varied so the random background noise is not a characteristic of the thresholds used to plot the figures. For instance, lowering the threshold does make the network more detailed but it also increases the number of background pixels above the threshold.

Tarboton, et al. (1988) used area above a threshold (called the support area) to identify where his channels started. A channel began when

$$A \geq \text{support area}$$

Note that for uniform runoff over the catchment (so that $Q = R\,A$) this definition is equivalent with the Dunne activator mechanism in Equation (5.22).

5.6 Sample Results

This section presents some sample results of the application of the computer model documented in Appendix A, based on the theory described in this chapter. This section is not intended to be a comprehensive consideration of all aspects of the model and the simulations; that will be provided in Chapters 7 and 8. Rather, using a single simulation through time, typical characteristics of the generated catchments will be noted.

The simulation run discussed in this section is CR2–3, and the parameters used for this, and all other, runs are listed in Appendix D. The results of the simulation presented here are typical of results obtained in this work. Figures 5.5 through 5.10 show the spatial distribution of various properties for selected times.

Figure 5.5 shows the simulated channel network. It demonstrates the headward growth of the channels from the initial seed on the bottom left–hand corner of the grid. The directions of overland flow are also shown and they demonstrate the convergence of flow directions on the hillslopes around the channel heads illustrated in Figure 5.1. The branching pattern of the network resulting from lateral branching is qualitatively similar to branching in the stream sapping hypothesis of Dunne (1969) and the pattern of future branching is mirrored by the pattern of hillslope flow directions (Figure C.4).

As the network grows, it erodes valleys along the channels. This results from the preferential erosion in the channels compared with the hillslopes. Figure 5.6, contours of elevation, clearly shows this characteristic. It is the valley that results in the convergence of hillslope flow directions noted above. These valleys result in the preferred hillslope flow directions being towards the channel network. An alternative view of this valley formation with time is given by Figure 5.7, which is an isometric view of elevations within the catchments.

The network growth process is dominated by the spatial distribution of the activator on the hillslopes. Figure 5.8 shows this spatial distribution of hillslope activator. The spatial distribution complies with the three conditions on activator distribution that have been proposed (see Figures 5.1 to 5.3). Most importantly the activator is concentrated around the channel heads and moves with the channel heads. It is apparent in Figure 5.8c that the highest peaks of the activator are at the growing channel heads, and that other peaks within the interior parts of the network are considerably lower. Verifying that the activator region repel each other, and that boundaries repel them is more difficult. Figure 5.8a shows some repulsion from the boundaries, the highest values of activator are away from the boundary so that future channel growth is into the unchanneled catchment interior. At later times the channels and high activator regions are relatively uniformly spaced which is consistent with the repulsion hypothesis.
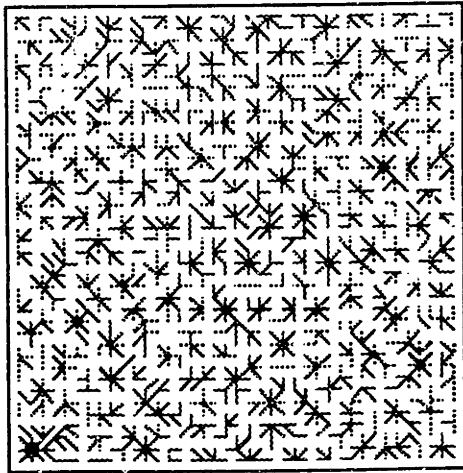
(a) $t' = 100$
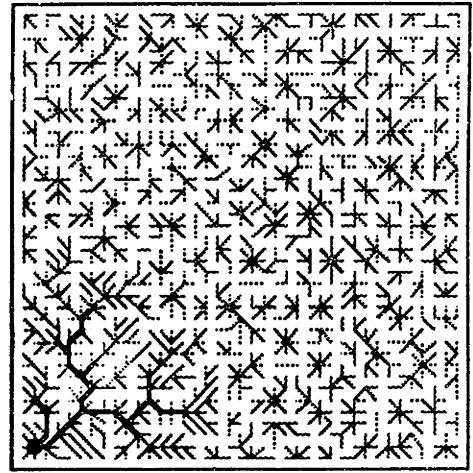
(b) $t' = 2000$
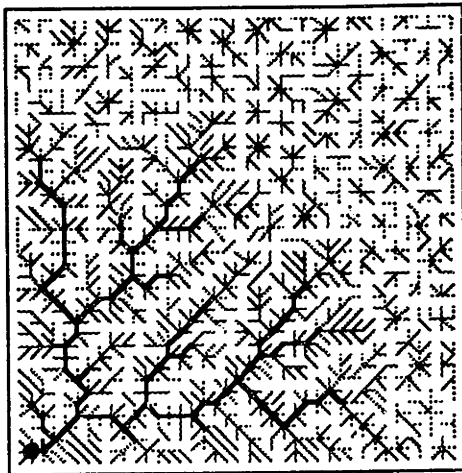
(c) $t' = 6000$
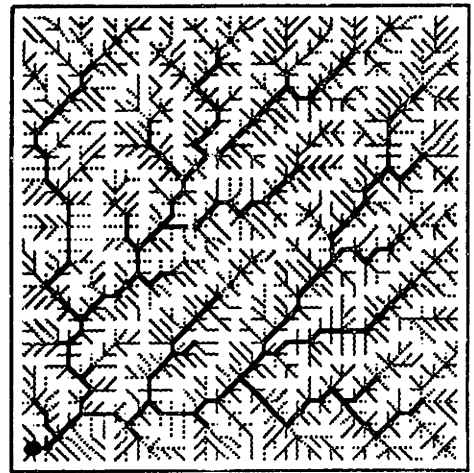
(d) $t' = 13000$

↗ channel   ↗ hillslope   ● Catchment outlet

Figure 5.5: Channel network and hillslope flow directions with time: CR2-3 simulation

173

(a) $t' = 100$

(b) $t' = 2000$

(c) $t' = 6000$

(d) $t' = 13000$

Max.　　　　　Min.
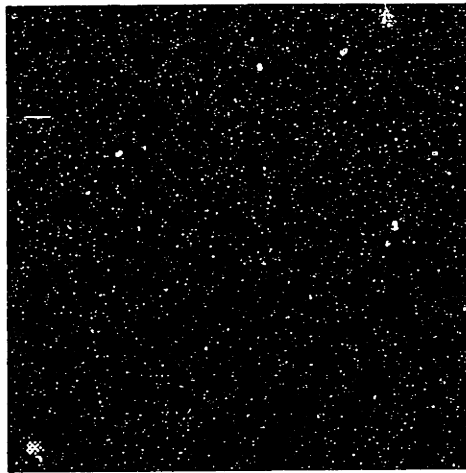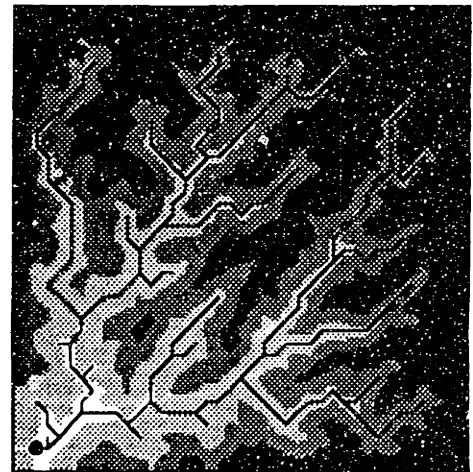
● Catchment outlet

Figure 5.6: Elevations with time: CR2-3 simulation
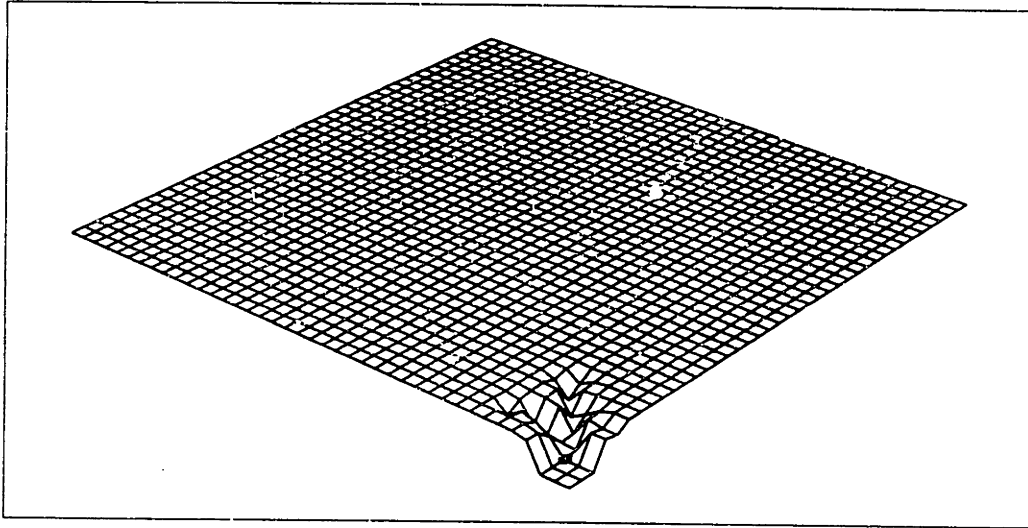
(a) $t' = 500$
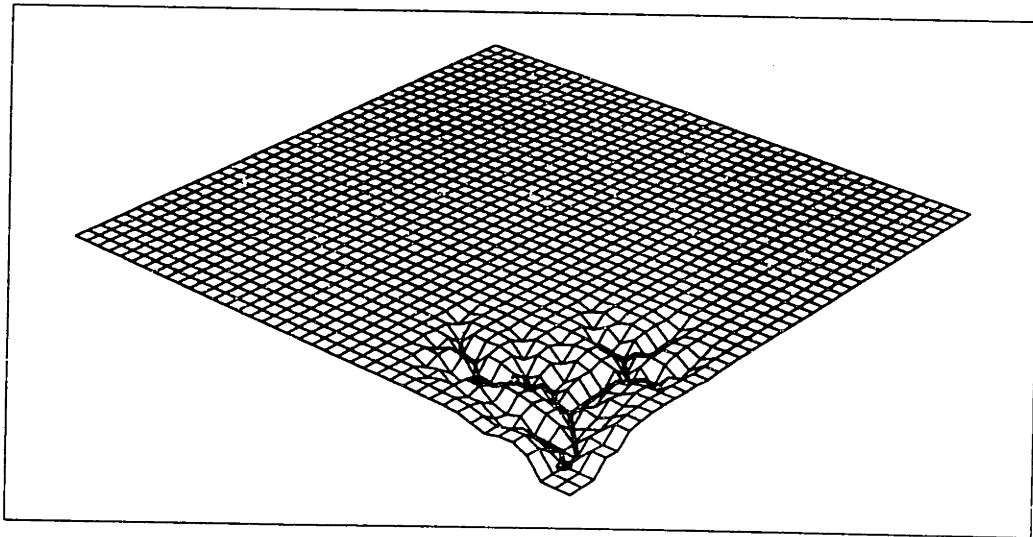
(b) $t' = 2000$

Figure 5.7: Isometric view of catchment with time:
Simulation CR2-3
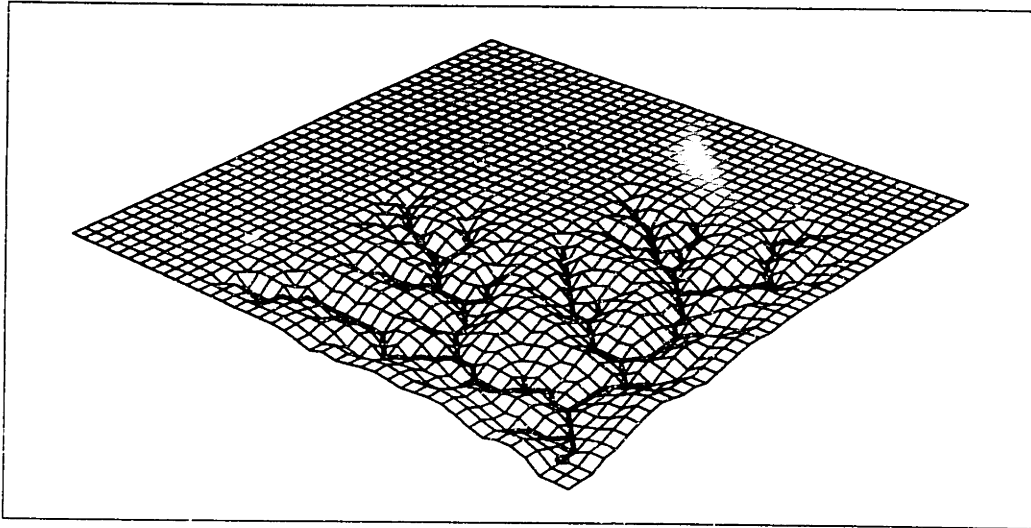
(c) $t' = 6000$
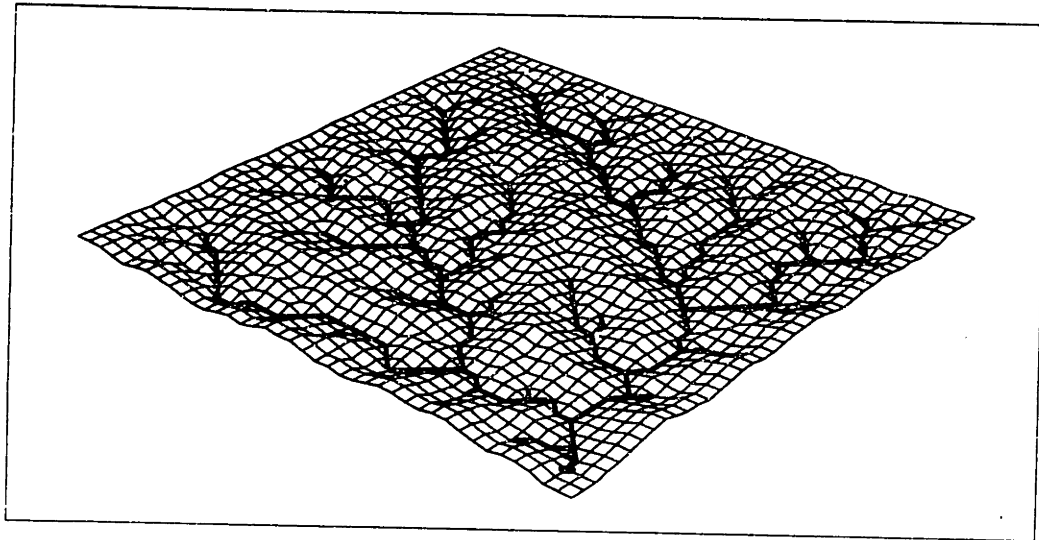
(d) $t' = 13000$

Figure 5.7 (ctd): Isometric view of catchment with time:
Simulation CR2-3

(a) $t' = 100$

(b) $t' = 2000$

(c) $t' = 6000$

(d) $t' = 13000$

Max.   Min.
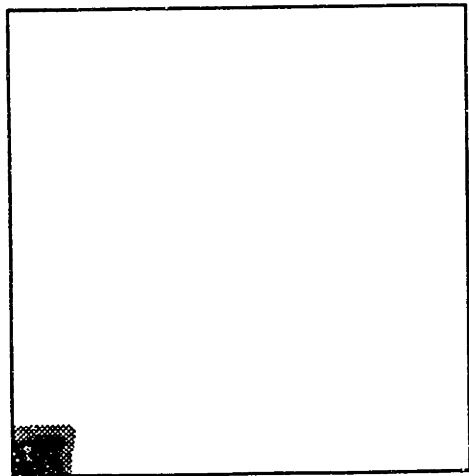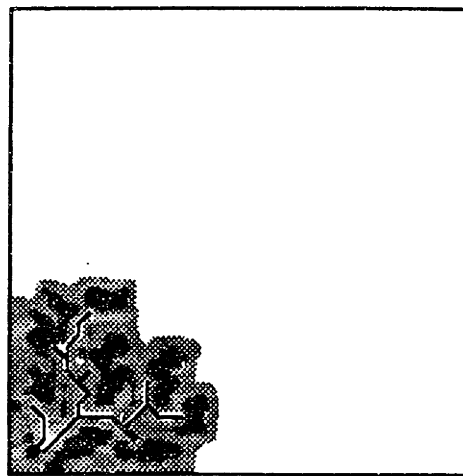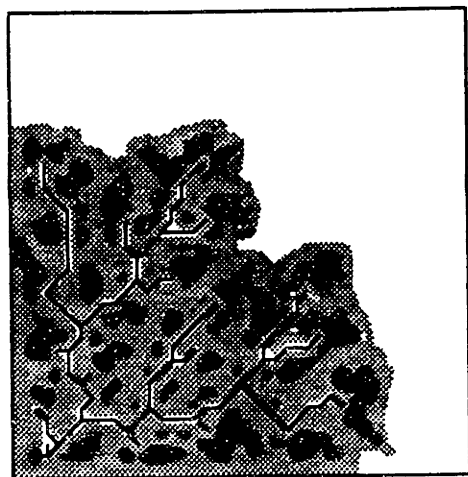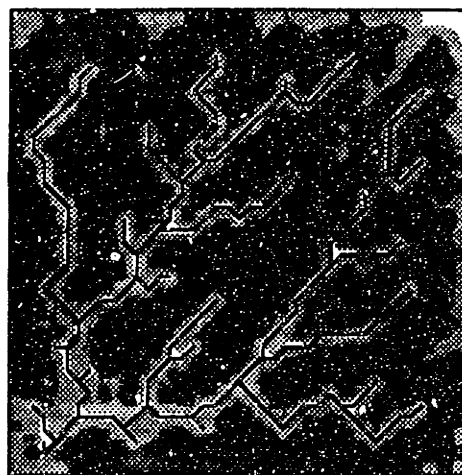
● Catchment outlet

Figure 5.8: Activator with time: CR2-3 simulation

Figure 5.9 is a plot of contributing area to the network. Not surprisingly the high areas are concentrated along the channel networks. This follows directly from the valley development and consequent preferential drainage to the channel network. The blotchy appearance of Figure 5.9a results from the initial minor random perturbations on elevation which create large number of interior draining regions on the initially flat surface.

Contours of hillslope slope are provided in Figure 5.10. The most interesting characteristic of this plot is that the steepest slopes do not occur around the growing head. The steepest slopes are on the laterally draining valley sides; the slopes draining down the valley to the channel heads are quite low by comparison. Figure 5.10d shows that as the catchment evolves with time the highest slopes are in the upstream reaches of the catchment, with lower slopes downstream. These lower slopes result from the hillslope erosion that has taken place in the older, root sections of the catchment of the bottom left—hand corner (see Figure 5.6).

Activator can be high even if the slopes at the channel head are low. In Figure 5.5 the hillslopes contributing to the channel head are long compared to the hillslopes draining laterally to the channel. Thus in the activator formulation of Equation (5.2a) the area contribution overwhelms the slope contribution. This indicates that network screening, which is primarily a result of reduction of contributing areas near the root of the network, is an important process in governing the distribution of growth sites.

Figure 5.11 gives the hypsometric curves for the catchment. The shape and trends with time of this curve are consistent with the interpretation of field data proposed by Schumm (1956). Hypsometric curves are discussed in more detail in Chapter 7.

Figure 5.12 shows the elevations of all the streams, normalized against both distance and elevation, for a variety of times both before and after the network has stopped growing. The curvature of the profile is reasonable and consistent with

(a) $t' = 100$

(b) $t' = 2000$

(c) $t' = 6000$

(d) $t' = 13000$

Max.   Min.
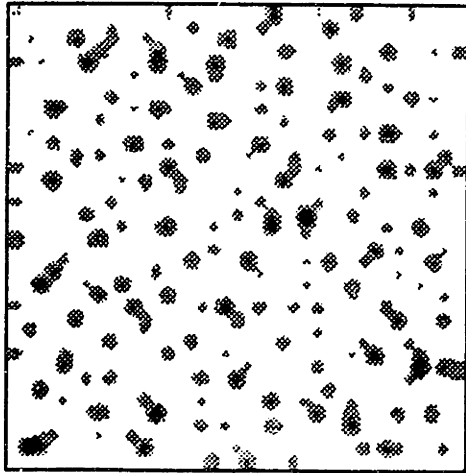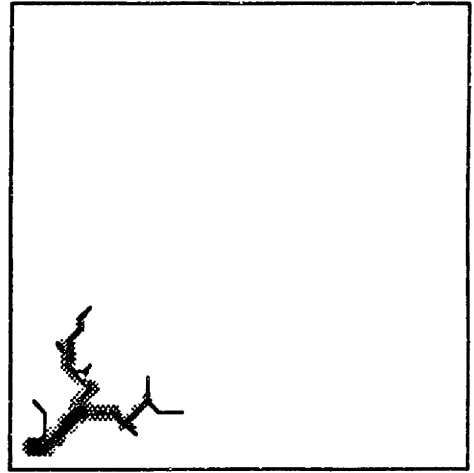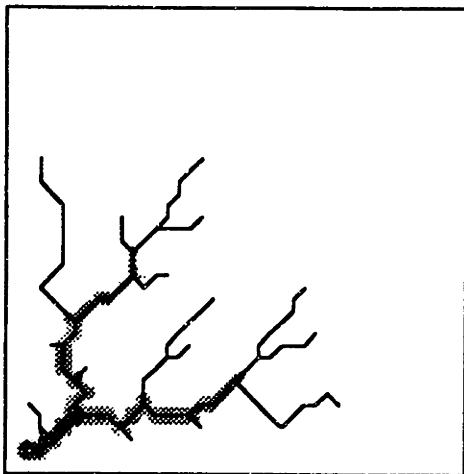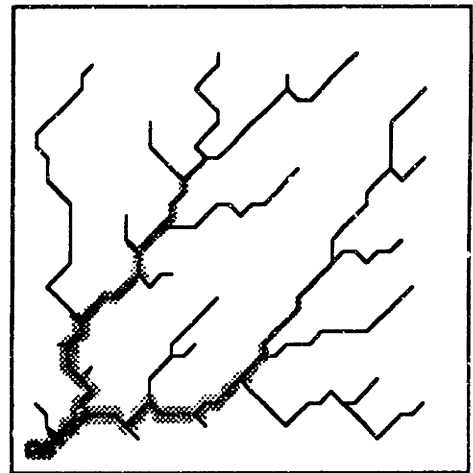
● Catchment outlet

Figure 5.9: Contributing area with time: CR2-3 simulation

(a) $t' = 100$

(b) $t' = 2000$

(c) $t' = 6000$

(d) $t' = 13000$

Max.   Min.
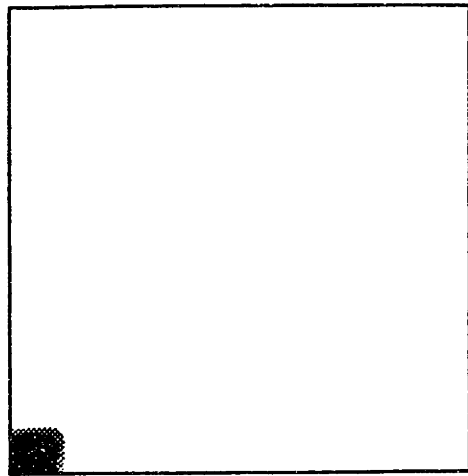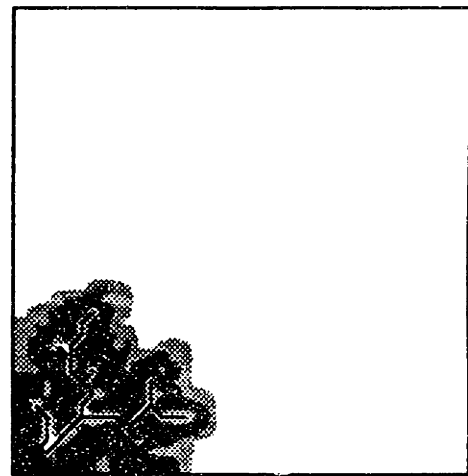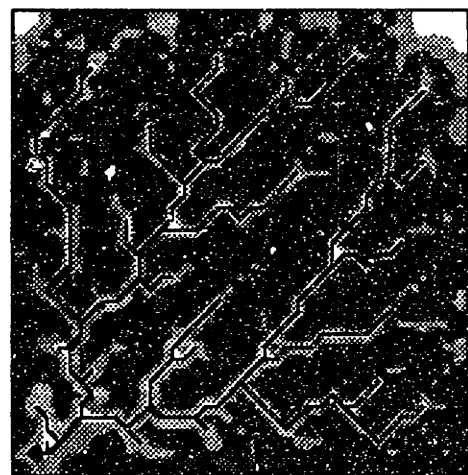
● Catchment outlet

Figure 5.10 Hillslope slopes with time: CR2-3 simulation
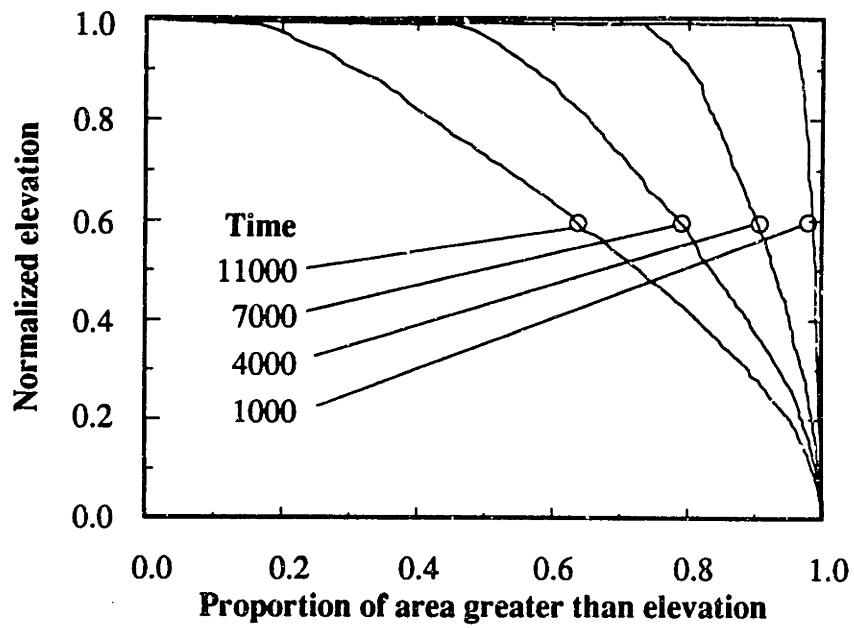
Figure 5.11: Hypsometric curve with time: CR2-3
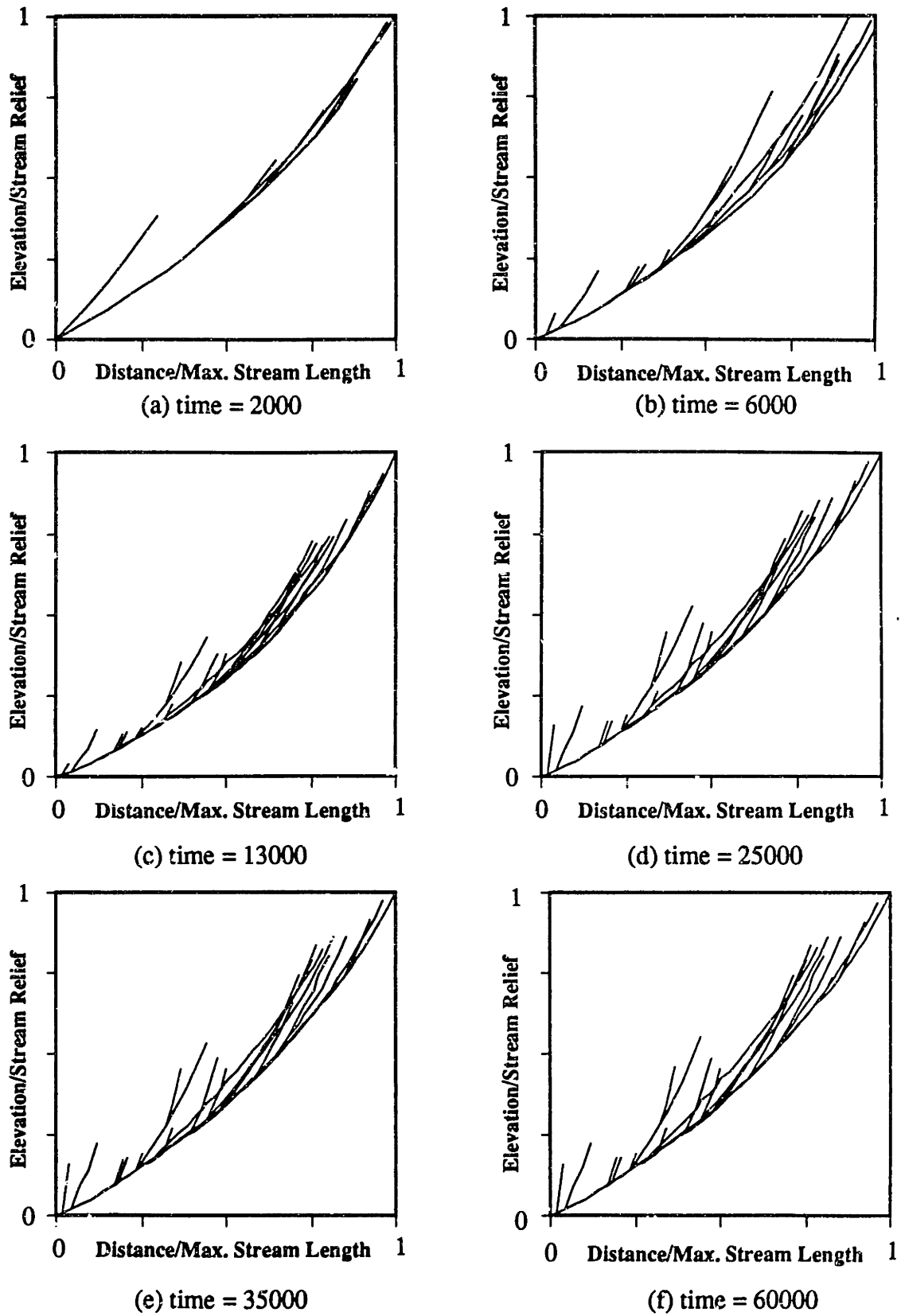
Figure 5.12: Longitudinal profiles for channels, channel elevations normalized by stream relief, horizontal distance from catchment outlet normalized by the length of the longest stream: simulations CR2-3 and CR8-1

(time = 60000 ≡ dynamic equilibrium)

observed data.

Table 5.3 list some sample statistics for the catchment for the instant in time at which the network stops growing. These statistics will be discussed in greater detail in Chapter 7. Suffice to say at this stage that the statistics are realistic for natural catchments.

In conclusion, a sample result has been shown and the characteristics of a typical simulation have been illustrated. The networks that are generated are qualitatively realistic both in planar and elevation profile properties. The drainage directions on the hillslopes are shown to be consistent with hypotheses of previous researchers and consistent with the stream sapping, headward growths mechanism.

TABLE 5.3

Sample Statistics: Simulation CR2–3

Statistic		Statistic	
R_b	5.20	R_ℓ	2.85
R_s	1.73	R_A	6.61
$R_b(1-2)$	5.50	$R_b(2-3)$	4.00
$R_\ell(1-2)$	3.29	$R_\ell(2-3)$	1.12
$R_s(1-2)$	1.70	$R_s(2-3)$	1.88
$R_A(1-2)$	7.15	$R_A(2-3)$	4.46
K	1.78	ϵ_1	2.25
D_d'	6.80	magnitude	22

Mean catchment relief	9.90
Mean hillslope relief	4.18
Mean stream relief	6.41
Mean hillslope slope	1.92

(1–2)≡statistics for Strahler order 1 to order 2

CHAPTER 6

NONDIMENSIONALIZATION AND SCALING

OF THE GOVERNING EQUATIONS

6.1 Introduction

The scaling and similitude properties of the system of governing equations presented in Chapter 5 will be considered here. Similitude is an important concept that allows the quantification of the similarities of, and differences between, two catchments. Through the use of a small number of nondimensional numbers the governing physics of a catchment may be summarized. These nondimensional numbers lead to similarity conditions that allow the quantitative comparison of data obtained at the field scale and at the controlled experimental scale.

This chapter is primarily concerned with derivation of the nondimensional numbers applicable to the governing equations, and with explanation of some of the more important consequences of this nondimensionalization to catchment geomorphology concepts such as "dynamic equilibrium." This chapter is not concerned with an exploration of the sensitivity of catchment form to these nondimensional numbers; the effect of changes in the nondimensional numbers on catchment form is the subject of Chapters 7 and 8.

The discussion below is divided into four main sections. The first deals solely with the mechanics of the nondimensionalization under both transient and steady state conditions. A linkage to the concept of "dynamic equilibrium" is developed. The following section applies this nondimensionalization to the problem of comparing two catchments, i.e., scaling one catchment so that its governing equations are in some sense the same as that of another catchment. The third section addresses the question of what are suitable scales to use in nondimensionalization; measures of catchment form previously used by geomorphologists are examined for their suitability as

catchment scales. The central problem of determining appropriate scales from field data is addressed and a number of open questions posed.

The final section deals with an application of the nondimensional equations to the problem of landscape classification. Landscape classification attempts to define the differences between, and similarities of, the landscapes of different regions. This final section will, by the use of similitude, challenge current ideas in this field. A more general, and what is believed to be more correct, framework for classifying landscapes will be presented.

6.2 The Mechanics of the Nondimensionalization

The equations (5.1a–e) define the governing physics of the model discussed below. They do not, however, provide an explicit understanding of the physical behavior of the results under scale transformation. For instance, how do the results vary if the catchment is twice as large, the sediment transport rates are halved, etc. A nondimensional analysis of these equations will explicitly consider, and predict, the effect of these scale transformations.

Examination of these equations, and consideration of their simulated solutions, indicate that there are five fundamental scales; all others can be derived from these scales. These nondimensional parameters and the scales that they are based upon are

$$z' \quad = \frac{z}{L_z} \qquad : \text{nondimensional elevation} \qquad (6.1)$$

$$x' \quad = \frac{x}{L_x} \qquad : \text{nondimensional horizontal distance}$$

$$(6.2)$$

$$R' \quad = \frac{T_R R}{L_R} \qquad : \text{nondimensional runoff rate} \qquad (6.3)$$

$$t' \quad = \frac{t}{T} \qquad : \text{nondimensional time} \qquad (6.4)$$

where L_z = vertical length scale (e.g. m)

          L_x = horizontal length scale (e.g. m)

186

L_R = runoff length scale (e.g., mm depths)

T_R = runoff rate time scale (e.g., hours)

T = catchment evolution timescale (e.g. years).

At this stage physical interpretations of these scales will not be formulated. These interpretations, based on measurable properties, will be made in Section 6.4.

Note that the vertical and horizontal length scales are distinct. Differences in the scaling of these lengths will prove important later in this chapter. For instance, the vertical length scale, L_z, may be the hillslope drop from the watershed to the channel while the horizontal length scale, L_x, may be the mean hillslope length. It will be shown later that these two scales are independent. It should also be pointed out that there are other nondimensional parameters that are important and that capture the form of tectonic effects, sediment transport, etc. One such nondimensional parameter, (m_5/n_5), the parameter governing the form of the activator equation is described in Section C.2.3. Other important nondimensional parameters will be described as the need arises.

The form of the nondimensional runoff rate in Equation (6.3) arises from the dependence of discharge on runoff rate and drainage area.

$$Q = RA \hspace{4cm} (6.5)$$

The form of the nondimensionalization in Equation (6.3) assumes that the runoff rate is independent of area. For field scale problems this is a simplification of reality. For experimental studies, as described by Schumm, et al. (1987), this assumption is in line with the experimental inputs, spatially and temporally uniform rainfall. In the field, the situation is more complicated. Section 5.4 addressed the interpretation of discharge in the temporally averaged sediment transport equation. It was asserted that in this case the discharge to be used is the mean peak discharge. Practitioners (Leopold, et al., 1964) have in the past found this discharge to be related to area by

$$Q_p = \beta_3 A^{m_3} \qquad\qquad m_3 \, \epsilon \, [0.5, 0.8]$$

or

$$Q_p = RA$$

where

$$R = \frac{\beta_3}{A^{(1-m_3)}}$$

The areal dependence follows from the areal dependency of rainfall, and the effect of travel times within the catchment upon the coincidence of peak discharges from tributaries as the catchment becomes bigger. The importance of this latter timing effect is dependent upon the areai dependence of the rainfall, the timing effect being greatest when the areal dependence of the rainfall is least (i.e., rainfall is uniformly spread over the catchment). It should then be clear that determination of the relative contribution of the vertical and horizontal scales to discharge is difficult, and lies at the heart of current work on runoff processes (Pilgrim, 1983; Milly and Eagleson, 1987,1988). Rather than make assumptions in this regard we have chosen to look at the simpler situation where runoff is independent of area. It is believed that the qualitative conclusions reached in later sections will not be substantially affected by such a simplification.

Using the basic nondimensional parameters of Equations (6.1) to (6.4) and the runoff relationship of Equation (6.5), the nondimensionalization of some important derived properties is

$$A' = \frac{1}{L_x^2} A \qquad\qquad : \text{nondimensional area} \qquad\qquad (6.6)$$

$$Q' = \frac{T_R}{L_x^2 \, L_R} Q \qquad\qquad : \text{nondimensional discharge} \qquad\qquad (6.7)$$

$$S' = \frac{L_x}{L_z} S \qquad \qquad : \text{ nondimensional slope} \qquad (6.8)$$

$$a' = \frac{T_R^{m_5}}{L_z^{n_5} L_R^{m_5} \bar{\beta}_5 L_x^{2m_5 - n_5}} a \qquad \qquad : \text{ nondimensional activator} \qquad (6.9)$$

$$Q_S' = \frac{T_R^{m_1}}{L_z^{n_1} L_R^{m_1} \bar{\beta}_1 L_x^{2m_1 - n_1}} Q_S \qquad \qquad : \text{ nondimensional sediment discharge} \qquad (6.10)$$

where

$$\beta_1'(\underline{x}) = \frac{\beta_1(\underline{x})}{\bar{\beta}_1} \qquad \qquad : \text{ nondimensional transport coefficient}$$

$$\beta_5'(\underline{x}) = \frac{\beta_5(\underline{x})}{\bar{\beta}_5} \qquad \qquad : \text{ nondimensional activator coefficient}$$

From this we define a nondimensional channelization indicator function

$$f'(Y) = \frac{f(Y)}{\bar{\beta}_1}$$

The nondimensional coefficients β_1' and β_5' are introduced for two reasons. First, β may vary in space within a catchment. Some results could be as much related to spatial patterns of β as to the actual magnitude of β. The spatial pattern is

captured by β' while the magnitude is captured by the scale $\bar{\beta}$. Second the scales $\bar{\beta}_1$ and $\bar{\beta}_5$ can not be related to the scales of Equation (6.1) to (6.4) because the inputs to the β's are sediment diameters and mannings n, which are independent of the catchment scales used in Equations (6.1) and (6.4). Sections 5.4 and 5.5 indicate, however, that β_1 and β_5 are not necessarily independent.

For a similar reason it is convenient to introduce a nondimensional tectonic uplift of the form

$$c_0'(n) = \frac{c_0(x)}{\bar{c}_0} \qquad \text{:nondimensional tectonic uplift coefficient}$$

$$(6.11)$$

Substituting the nondimensional relationships of Equations (6.1) through (6.11) into the governing equations of Equation (5.1) and collecting terms yields the nondimensionalized governing equations

$$\frac{\partial z_j'}{\partial t'} = TT \cdot c_{0j}' + TS \cdot \sum_i [(Q_i')^{m_1} (S_{ij}')^{n_i} I_{ij} f'(Y_i)] + TD \cdot \frac{\partial^2 z'}{\partial x_i'^2}$$

$$(6.12a)$$

$$\frac{\partial Y_i}{\partial t'} = TC \cdot \left[0.0025 \, TA \cdot a_j' + [-0.1 Y_j + \frac{Y_j^2}{1 + 9 Y_j^2}] \right]$$

$$(6.12b)$$

where

TT, TS, TD, TC, TA = nondimensional numbers defined below.

In Equation (6.12) we define a number of nondimensional numbers that govern the relative magnitudes of the physical processes in the model. They are:

Transient Tectonic Uplift Number:

$$TT = \frac{\text{Rate of tectonic uplift}}{\text{Total rate of elevation change}}$$

$$= \frac{T \bar{c}_0}{L_z} \tag{6.13a}$$

Transient Sediment Transport Number:

$$TS = \frac{\text{Rate of elevation change due to sediment transport}}{\text{Total rate of elevation change}}$$

$$= \frac{T \, L_x^{2m_1-n_1} \, L_z^{n_1-1} \, L_R^{m_1} \, \bar{\beta}_1}{L_g^2 \, T_R^{m_1} \, \rho_s (1-n)} \tag{6.13b}$$

Transient Diffusion Number:

$$TD = \frac{\text{Rate of elevation change due to diffusive transport}}{\text{Total rate of elevation change}}$$

$$= \frac{L_g T D_z}{L_x^2} \tag{6.13c}$$

Transient Channelization Rate Number:

$$TC = \frac{\text{Timescale of elevation change}}{\text{Timescale of channel growth}}$$

$$= (T d_t) \tag{6.13d}$$

Transient Activator Number:

$$TA = \frac{\text{Magnitude of channel initiation process}}{\text{Threshold on channel initiation}}$$

$$= \frac{L_z^{n_5} \, L_x^{2m_5-n_5} \, L_R^{m_5} \, c_1 \, \bar{\beta}_5}{T_R^{m_5}} \tag{6.13e}$$

191

The Equations (6.12a) and (6.12b) define the conditions of _transient_ similarity. Transient conditions are when systematic changes in either elevations or the channel network are observed, so that the right hand side of Equation (6.12a) and (6.12b) can be considered non—zero at the time scales of interest. Short term fluctuations in elevations and the channel network are ignored for the purposes of defining transient conditions since we consider the system of Equations (5.1) to be mean temporal equations.

The nondimensional solution domain is given by Ω'. It is possible to define the rectangular solution domain that is used in terms of L_x so that Equation (5.2e) becomes

$$\Omega' = [0, \frac{L_1}{L_x}] \times [0, \frac{L_2}{L_x}]$$

However, it is more convenient to introduce another length scale, the catchment length scale L_d, representing the length scale of the domain

$$\Omega' = [0, \frac{L_1}{L_d}] \times [0, \frac{L_2}{L_d}] \tag{6.14}$$

where $L_d = \alpha_1 L_x$.

The nondimensional number α_1 is used to represent the catchment horizontal scale, relative to the fundamental horizontal scale, L_x (i.e the hillslope scale). The nondimensional number α_1 will called the _Catchment Scale Number_ since it defines the size of the catchment relative to the hillslope scale. The reasons for this formulation will be addressed in Section 6.4.

For two catchments to be in transient similitude, the nondimensional

parameters defined for each catchment must be equal. Formally this can be expressed as (subscripts (1) and (2) refer to catchment (1) and (2), respectively)

1.	$TT_{(1)} = TT_{(2)}$	: Tectonic uplift similarity	
2.	$TS_{(1)} = TS_{(2)}$	: Sediment transport similarity	
3.	$TD_{(1)} = TD_{(2)}$	: Diffusion similarity	(6.15)
4.	$TC_{(1)} = TC_{(2)}$	: Channelization rate similarity	
5.	$TA_{(1)} = TA_{(2)}$	: Activator similarity	
6.	$\alpha_{1(1)} = \alpha_{1(2)}$	: Catchment scale similarity	

Furthermore, there are conditions on the nondimensional catchment properties so that

7.	$Q'(\underline{x}')_{(1)} = Q'(\underline{x}')_{(2)}$	: Discharge distribution	
8.	$f'(Y(\underline{x}'))_{(1)} = f'(Y(\underline{x}')_{(2)}$	: Sediment transport distribution	
			(6.16)
9.	$z'(\underline{x}')_{(1)} = z'(\underline{x}')_{(2)}$	: Elevation distribution	
10.	$c_0'(\underline{x}')_{(1)} = c_0'(\underline{x}')_{(2)}$	: Tectonics distribution	

The conditions of Equation (6.16) amount to forcing the two catchments to be exactly the same, after scaling. This shall be termed <u>deterministic similarity</u>. It may be that, after scaling, statistics that capture the average behavior over the catchment of properties in Equations (6.16) are the same. This type of similarity will be called <u>statistical similarity</u>. This distinction is made because it has been noted by previous researchers (Schumm, et. al., 1987), and it will be noted here also, that even though there exists <u>physical similarity</u> (Equation 6.15), the sensitive dependence of networks on minor fluctuations during their growth, means that under virtually no conditions

will Equation (6.16) be satisfied, except in some average, statistical sense. Two networks on physically identical catchments will never be the same because of the effect of minor input fluctuations. This issue was discussed in Section 4.4, for the Meinhardt equations, and it will be returned to in Chapter 7 when simulation results are interpreted.

An important concept in geomorphology is that of dynamic equilibrium. While a catchment may exhibit short term, episodic, fluctuations in form, catchments in dynamic equilibrium do not exhibit any systematic changes with time. The definitions of both "dynamic equilibrium" and "systematic change" in the literature are vague. There are three possible definitions of dynamic equilibrium depending upon viewpoint.

1.   Other than short term episodic fluctuations, the channel network and the surrounding hillslopes exhibit no change in time. Elevations are constant, and the channel network is fixed in space. Hereafter this case will be referred to as <u>deterministic dynamic equilibrium</u>.

2.   Other than short term episodic fluctuations, statistics describing the channel network and the surrounding hillslopes, exhibit no change in time (i.e., statistical stationarity). For instance, Strahler slope ratios may fluctuate over the short term, but their temporal mean value should be constant with time. Hereafter this case will be referred to as <u>statistical dynamic equilibrium</u>.

3.   Other than short term episodic fluctuations, the nondimensional form of the channel network and the surrounding hillslopes exhibits no change with time. For instance, slope magnitudes may change systematically with time, but the nondimensional slopes (i.e., $S'(\underline{x})$) do not change with time; in the case of slopes the vertical length scale will change. Hereafter this case will be referred to as <u>nondimensional deterministic dynamic equilibrium</u>. We might talk of stationarity of the statistics of the

nondimensionalized catchment. This case will be referred to as nondimensional statistical dynamic equilibrium.

These definitions of dynamic equilibrium are all fundamentally different. The strictest definition is the first, that of deterministic dynamic equilibrium. Here the form of the catchment is totally fixed and no systematic variations are allowed in time in any part of the catchment. The definition of statistical dynamic equilibrium is more flexible, in that while it restricts the variation of statistics, it allows variations within the catchment, provided that the mean properties do not change. Field studies that assert, for instance, that the Strahler slopes do not vary with time are asserting statistical dynamic equilibrium, since the Strahler slope is a statistic describing the average slope of a number of channels with the same order.

The third definition, of nondimensional dynamic equilibrium, is the most interesting since it has been used implicitly, but has never been explicitly described. In the same sense that the similitude conditions were asserted in Equation (6.15), then having nondimensionalized slopes, $S'(\underline{x})$, that are the same should constitute some form of equilibrium. The slope magnitudes or scales are dependent on time (via the vertical scale L_z which is now variable with time), but the distribution of slope within the catchment is independent of time (i.e. nondimensional slopes do not change with time). Typically the concept of nondimensional dynamic equilibrium arises most frequently in dealing with characteristic profiles of hillslopes or channels (e.g., Smith and Bretherton, 1972; Kirkby, 1971; Ahnert, 1976). Three examples of nondimensional dynamic equilibrium are:

1. Strahler slope ratios independent of time.

2. Hypsometic integrals converging to the Monadanock stage/profile (Strahler, 1964).

3. The nondimensional hillslope profile described by Equation (2.25)

The preceding discussion was aimed at eliminating confusion over what is meant by dynamic equilibrium. By and large, the derivation that follows assumes statistical dynamic equilibrium. We concern ourselves primarily with ensuring that physical similitude conditions based on nondimensional numbers are obeyed (Equation 6.15), and not with similitude conditions for the nondimensional catchment properties (Equation 6.16). We will assert, however, that the time derivatives of elevation and channel growth are zero; a somewhat stricter requirement than statistical dynamic equilibrium. The nondimensionalized governing equations at dynamic equilibrium are then

$$\text{DET} \cdot c'_{0j} + \sum_i [(Q'_i)^{m_1} (S'_j)^{n_1} I_{ij} \, f'(Y_i)] + \text{DED} \cdot \frac{\partial^2 z'_j}{\partial x'^2_i} = 0$$

(6.17a)

$$0.0025 \, \text{DEA} \cdot a'_j + [-0.1 Y_j + \frac{Y^2_j}{1 + 9Y^2_j}] = 0$$

(6.17b)

In an analogous fashion to the transient case nondimensional numbers arise that govern the physics of the model at dynamic equilibrium. They are

Dynamic Equilibrium Tectonic Uplift/Transport Number:

$$\text{DET} = \frac{TT}{TS}$$

$$= \frac{\text{Rate of tectonic uplift}}{\text{Rate of elevation change due to sediment transport}}$$

$$= \frac{\overline{c}_0 \, T_R^{m_1} \, \rho_s (1-n) \, L^2_g}{\overline{\beta}_1 \, L_x^{2m_1 - n_1} \, L_z^{n_1} \, L_R^{m_1}}$$

(6.18a)

Dynamic Equilibrium Diffusion/Transport Number:

$$DED = \frac{TD}{TS}$$

$$= \frac{\text{Rate of elevation change due to diffusive transport}}{\text{Rate of elevation change due to sediment transport}}$$

$$= \frac{D_z \, T_R^{m_1} \, \rho_s \, (1-n) \, L_g^3}{\bar{\beta}_1 \, L_x^{2m_1-n_1+2} \, L_z^{n_1-1} \, L_R^{m_1}} \qquad (6.18b)$$

Dynamic Equilibrium Activator Number:

$$DEA = TA = \frac{L_z^{n_5} \, L_x^{2m_5-n_5} \, L_R^{m_5} \, c_1 \, \bar{\beta}_5}{T_R^{m_5}} \qquad (6.18c)$$

The <u>Catchment Scale Number</u>, α_1, is as defined in the transient case. Note that the dynamic equilibrium nondimensional numbers are expressible in terms of the transient nondimensional numbers defined in Equation (6.12) and (6.13). The tectonic uplift/transport number is simply the relative rates of tectonic uplift to sediment transport. The diffusion transport number is simply the relative rates of diffusion transport to fluvial sediment transport. The activator number for dynamic equilibrium is identical to that for the transient case.

In a similar fashion to the transient case, it is possible to define the conditions for similitude to apply between two catchments ((1) and (2)) at dynamic equilibrium.

1. $DET_{(1)} = DET_{(2)}$: Tectonics/Transport similarity

2. $DED_{(1)} = DED_{(2)}$: Diffusion/Transport similarity

$$(6.19)$$

3. $DEA_{(1)} = DEA_{(2)}$: Activator similarity

In addition, for deterministic dynamic equilibrium to apply it is required that

197

4. $Q'(\underline{x}')_{(1)} = Q'(\underline{x}')_{(2)}$

5. $f(Y(\underline{x}'))_{(1)} = f(Y(\underline{x}'))_{(2)}$ (6.20)

6. $z'(\underline{x}')_{(1)} = z'(\underline{x}')_{(2)}$

For statistical dynamic equilibrium only the statistics of properties in Equation (6.20a) as well as Equation (6.19) need be the same for the two catchments.

A final note will be made in conclusion of this section about the nondimensionalization of the governing equation. The nondimensionalizations of Equations (6.12) and (6.19) are not unique. For instance, it may be more convenient to replace the nondimensional elevation z' with a nondimensional slope S' (defined as z'/x'), since slopes are a commonly used means of parameterizing catchment relief properties. The definition of nondimensional slope in Equation (6.8) can be used to replace L_z in the nondimensional governing equations. Define the slope scale and nondimensional slope as

$$S' = \frac{L_x}{L_z} S = \frac{S}{L_S}$$

where L_S = slope scale

The nondimensionalized transient governing equations are identical to those presented before (Equation 6.12) but the definitions of the transient nondimensional numbers are modified to

$$TT = \frac{T \, \overline{c}_0}{L_S \, L_x} \qquad\qquad (6.21a)$$

$$TS = \frac{T\, L_x^{2m_1-1}\, L_S^{n_1-1}\, L_R^{m_1}\, \bar{\beta}_1}{T_R^{m_1}\, \rho_s(1-n)\, L_g^2}$$

(6.21b)

$$TD = \frac{L_g\, TD_z}{L_x^2}$$

(6.21c)

$$TC = Td_t$$

(6.21d)

$$TA = \frac{L_x^{2m_5}\, L_S^{n_5}\, L_R^{m_5}\, c_1\, \bar{\beta}_5}{T_R^{m_5}}$$

(6.21e)

The nondimensionalized dynamic equilibrium equations are also identical to those previously presented (Equation 6.17) with modified definitions for the nondimensional numbers so that

$$DET = \frac{\bar{c}_0\, T_R^{m_1}\, \rho_s(1-n)\, L_g^3}{\bar{\beta}_1\, L_x^{2m_1}\, L_S^{n_1}\, L_R^{m_1}}$$

(6.22a)

$$DED = \frac{D_z\, T_R^{m_1}\, \rho_s(1-n)\, L_g^2}{\bar{\beta}_1\, L_S^{n_1-1}\, L_x^{2m_1+1}\, L_R^{m_1}}$$

(6.22b)

$$DEA = \frac{L_x^{2m_5}\, L_S^{n_5}\, L_R^{m_5}\, c_1\, \bar{\beta}_5}{T_R^{m_5}}$$

(6.22c)

6.3 Similarity Between Catchments Using Nondimensionalization

The previous section described the nondimensionalization of the governing equations. Some similarity conditions were briefly outlined. This section expands on those similarity conditions, and demonstrates their use by deriving the conditions under which two catchments are similar. This will illustrate how field data from one catchment may be compared with field data from another catchment, and how small scale experimental data and numerical simulations may be scaled up to field scale.

Deterministic Dynamic Equilibrium

Consider initially the case of two catchments at deterministic dynamic equilibrium. The similarity conditions for this case are summarized by Equations (6.19) and (6.20). Similarity for two cases will be considered.

Case 1: Dynamic Equilibrium Similarity Ignoring Diffusion

In the case where diffusive effects in the catchment are negligible and may be safely neglected, it is only necessary to consider Tectonic/Transport similarity and Activator similarity, parameterized by the nondimensional numbers **DET** and **DEA** respectively. If the scales of the two catchments being compared are differentiated by the subscripts (1) and (2) respectively, then the similitude conditions are

$$DET_{(1)} = DET_{(2)}$$
$$DEA_{(1)} = DEA_{(2)}$$

or substituting the definitions of the nondimensional numbers from Equation (6.18)

$$\frac{\overline{c}_{0(1)} \; T_{R(1)}^{m_1} \; \rho_{(1)}^{(1-n_{(1)})} \; L_g^2}{\overline{\beta}_{1(1)} \; L_{x(1)}^{2m_1-n_1} \; L_{z(1)}^{n_1} \; L_{R(1)}^{m_1}} = \frac{\overline{c}_{0(2)} \; T_{R(2)}^{m_1} \; \rho_{(2)}^{(1-n_{(2)})} \; L_g^2}{\overline{\beta}_{1(2)} \; L_{x(2)}^{2m_1-n_1} \; L_{z(2)}^{n_1} \; L_{R(2)}^{m_1}}$$

<div align="right">(6.23a)</div>

$$\frac{\overline{\beta}_{5(1)} L_{x(1)}^{2m_5-n_5} \; L_{z(1)}^{n_5} \; L_{R(1)}^{m_5} c_{1(1)}}{T_{R(1)}^{m_5}} = \frac{\overline{\beta}_{5(2)} L_{x(2)}^{2m_5-n_5} \; L_{z(2)}^{n_5} L_{R(2)}^{m_5} c_{1(2)}}{T_{R(2)}^{m_5}}$$

<div align="right">(6.23b)</div>

These equations can be rearranged into a form that is reminiscent of the original
nondimensional number so that the similitude conditions are

$$\frac{\overline{c}_{0(1,2)} \; T_{R(1,2)}^{m_1} \; \rho_{s(1,2)}^{n'_{(1,2)}} \; L_{g(1,2)}^2}{\overline{\beta}_{1(1,2)} \; L_{x(1,2)}^{2m_1-n_1} \; L_{z(1,2)}^{n_1} \; L_{R(1,2)}^{m_1}} = 1$$

<div align="right">(6.24a)</div>

$$\frac{\overline{\beta}_{5(1,2)} \; L_{x(1,2)}^{2m_5-n_5} \; L_{z(1,2)}^{n_5} \; L_{R(1,2)}^{m_5} \; c_{1(1,2)}}{T_{R(1,2)}^{m_5}} = 1$$

<div align="right">(6.24b)</div>

where the scaling coefficients are defined as

$$\overline{c}_{0(1,2)} = \frac{\overline{c}_{0(1)}}{\overline{c}_0(2)}$$

$$T_{R(1,2)} = \frac{T_{R(1)}}{T_{R(2)}}$$

$$\overline{\beta}_{1(1,2)} = \frac{\overline{\beta}_{1(1)}}{\overline{\beta}_{1(2)}}$$

$$L_{x(1,2)} = \frac{L_x(1)}{L_{x(2)}}$$

$$L_{z(1,2)} = \frac{L_{z(1)}}{L_{z(2)}}$$

$$\rho_{s(1,2)} = \frac{\rho_{s(1)}}{\rho_{s(2)}}$$

$$n'_{(1,2)} = \frac{1 - n_{(1)}}{1 - n_{(2)}}$$

$$c_{1(1,2)} = \frac{c_{1(1)}}{c_{1(2)}}$$

$$\overline{\beta}_{5(1,2)} = \frac{\overline{\beta}_{5(1)}}{\overline{\beta}_{5(2)}}$$

$$L_{g(1,2)} = \frac{L_{g(1)}}{L_{g(2)}}$$

The important difference between Equations (6.18) and (6.24) can be demonstrated taking one parameter as an example. The parameter $\overline{c}_{0(1)}$ is a characteristic scale of catchment (1) with the units of $\overline{c}_0$. On the other hand $\overline{c}_{0(1,2)}$ is a scaling coefficient relating the characteristic scale $\overline{c}_{0(1)}$ of catchment (1) with the characteristic scale of $\overline{c}_{0(2)}$ of catchment (2), and $\overline{c}_{0(1,2)}$ is dimensionless.

Case 2: Full Dynamic Equilibrium Similarity

In the case where diffusive effects are not negligible, so that tectonics, sediment transport and diffusion are of a comparable magnitude, or where individual components may be dominant in different parts of the catchment, so that none may be ignored, then the similitude conditions for catchments (1) and (2) are

$$DET_{(1)} = DET_{(2)}$$
$$DEA_{(1)} = DEA_{(2)}$$

$$DED_{(1)} = DED_{(2)}$$

Substituting for the definitions of the nondimensional variables from Equation (6.18) and (6.24) yields

$$\frac{\bar{c}_{0(1,2)}\, T_{R(1,2)}^{m_1}\, \rho_{s(1,2)}\, n'(1,2)\, L_{g(1,2)}^2}{\bar{\beta}_{1(1,2)}\, L_{x(1,2)}^{2m_1-n_1}\, L_{z(1,2)}^{n_1}\, L_{R(1,2)}^{m_1}} = 1 \qquad (6.25a)$$

$$\frac{\bar{\beta}_{2(1,2)}\, L_{x(1,2)}^{2m_5-n_5}\, L_{z(1,2)}^{n_5}\, L_{R(1,2)}^{m_5}\, c_{1(1,2)}}{T_{R(1,2)}^{m_5}} = 1 \qquad (6.25b)$$

$$\frac{D_{z(1,2)}\, T_{R(1,2)}^{m_1}\, \rho_{s(1,2)}\, n'(1,2)\, L_{g(1,2)}^3}{\bar{\beta}_{1(1,2)}\, L_{x(1,2)}^{2m_1-n_1+2}\, L_{z(1,2)}^{n_1-1}\, L_{R(1,2)}^{m_1}} = 1 \qquad (6.25c)$$

where $D_{z(1,2)} = \dfrac{D_{z(1)}}{D_{z(2)}}$ is the diffusion scaling coefficient analogously to those coefficients defined in Equation (6.24).

The similitude conditions are the same as for case 1, transport and activator similitude, except that an additional condition, Equation (6.25c), is required for diffusion similitude.

Transient Conditions

The conditions for similitude for two catchments under transient conditions will be treated in a similar fashion to the treatment of conditions at dynamic equilibrium. The main complication of the transient case is that the total number of similitude conditions is increased from 3 to 5. The transient similitude conditions were

summarized in Equation (6.15) and (6.16). As was done for dynamic equilibrium, two cases will be considered, that where diffusion is not an important process and that where it is.

Case 3: Transient Similarity Ignoring Diffusion

In the case of where diffusive effects in the catchment are negligible, it is only necessary to consider similarity conditions for tectonics, transport, channel growth and activator, parameterized by the nondimensional numbers TT, TS, TC, and TA, respectively. Using the same notation as in Equations (6.24) and (6.25), the similarity conditions for two catchments (1) and (2) are

$$TT_{(1)} = TT_{(2)}$$
$$TS_{(1)} = TS_{(2)}$$
$$TC_{(1)} = TC_{(2)}$$
$$TA_{(1)} = TA_{(2)}$$

and substituting for the definitions of the nondimensional numbers from Equations (6.12) and (6.13) yields

$$\frac{T_{(1,2)} \bar{c}_{0(1,2)}}{L_{z(1,2)}} = 1 \tag{6.26a}$$

$$\frac{T_{(1,2)} L_{x(1,2)}^{2m_1-n_1} L_{z(1,2)}^{n_1-1} L_{R(1,2)}^{m_1} \bar{\beta}_{1(1,2)}}{T_{R(1,2)}^{m_1} \rho_{s(1,2)} n'_{(1,2)} L_{g(1,2)}^2} = 1 \tag{6.26b}$$

$$T_{(1,2)} d_{t(1,2)} = 1 \tag{6.26c}$$

$$\frac{L_{x(1,2)}^{2m_5-n_5}\, L_{z(1,2)}^{n_5}\, L_{R(1,2)}^{m_5}\, c_{1(1,2)}\, \bar{\beta}_{5(1,2)}}{T_{R(1,2)}^{m_5}} = 1 \qquad (6.26d)$$

where

$$T_{(1,2)} = \frac{T_{(1)}}{T_{(2)}}$$

$$d_{t(1,2)} = \frac{d_{t(1)}}{d_{t(2)}}$$

The equilibrium similitude conditions in Equation (6.24) are a subset of the transient similitude conditions above. The rearranged equations are

$$\frac{\bar{c}_{0(1,2)}\, T_{R(1,2)}^{m_1}\, \rho_{s(1,2)}\, n'_{(1,2)}\, L_{g(1,2)}^{2}}{\bar{\beta}_{1(1,2)} L_{x(1,2)}^{2m_1-n_1}\, L_{z(1,2)}^{n_1} L_{R(1,2)}^{m_1}} = 1$$

$$(6.27a)$$

$$\frac{\bar{\beta}_{5(1,2)} L_{x(1,2)}^{2m_5-n_5}\, L_{z(1,2)}^{n_5} L_{R(1,2)}^{m_5} c_{1(1,2)}}{T_{R(1,2)}^{m_5}} = 1 \qquad (6.27b)$$

$$\frac{T_{(1,2)}\, \bar{c}_{0(1,2)}}{L_{z(1,2)}} = 1 \qquad (6.27c)$$

$$T_{(1,2)}\, d_{t(1,2)} = 1 \qquad (6.27d)$$

The first two equations, Equations (6.27a) and (6.27b), are the similarity conditions for dynamic equilibrium. Equation (6.27c) is a similarity condition on the

timescale reflecting the rate of tectonic uplift, c_0, divided by the rate of elevation change within the catchment, L_z/T. Equation (6.27d) is a similarity condition on the timescale of the channel growth process, d_t, compared to the timescale of elevation change. Thus the transient similarity conditions are just those for dynamic equilibrium with two additional constraints on timescales within the model.

Case 4: Full Transient Similarity

In the case where diffusive effects are not negligible, so that tectonics, sediment transport and diffusion are of comparable magnitude, or where individual components may dominate in different parts of the catchment, then the similitude condition for two catchments (1) and (2) are

$$TT_{(1)} = TT_{(2)}$$
$$TS_{(1)} = TS_{(2)}$$
$$TC_{(1)} = TC_{(2)}$$
$$TA_{(1)} = TA_{(2)}$$
$$TD_{(1)} = TD_{(2)}$$

Substituting for the definitions of the nondimensional numbers from Equations (6.12) and (6.13) yields

$$\frac{\overline{c}_{0(1,2)} \; T_{R(1,2)}^{m_1} \; \rho_{s(1,2)} \; n'_{(1,2)} \; L_{g(1,2)}^2}{\overline{\beta}_{1(1,2)} L_{x(1,2)}^{2m_1-n_1} L_{z(1,2)}^{n_1} L_{R(1,2)}^{m_1}} = 1$$

(6.28a)

$$\frac{\overline{\beta}_{5(1,2)} \; L_{x(1,2)}^{2m_5-n_5} \; L_{z(1,2)}^{n_5} L_{R(1,2)}^{m_5} c_{1(1,2)}}{T_{R(1,2)}^{m_5}} = 1 \tag{6.28b}$$

$$\frac{D_{z(1,2)} \; T_{R(1,2)}^{m_1} \; \rho_{s(1,2)} \; n'_{(1,2)} \; L_{g(1,2)}^3}{\overline{\beta}_{1(1,2)} \; L_{x(1,2)}^{2m_2-n_1+2} \; L_{z(1,2)}^{n_1-1} \; L_{R(1,2)}^{m_1}} = 1 \tag{6.28c}$$

$$\frac{T_{(1,2)} \; \overline{c}_{0(1,2)}}{L_{z(1,2)}} = 1 \tag{6.28d}$$

$$T_{(1,2)} \; d_{t(1,2)} = 1 \tag{6.28e}$$

As was the case for Equations (6.27), the equations above have been rearranged to highlight the similarity between Cases 2 and 4, the dynamic equilibrium and transient cases. Again the difference between the transient case and the dynamic equilibrium case is the addition of two similitude conditions on the timescale of catchment evolution process, T, and the rate of channel growth, d_t.

The previous discussion in this section highlights two important points. These points have important consequences in analyzing field and numerical data.

The first point is that as the number of processes to be modelled increases, so does the number of similitude conditions. With more similitude constraints, there is less freedom in the choice of the values of $L_{z(1,2)}$, $L_{x(1,2)}$, $L_{R(1,2)}$, etc., that may be chosen to scale two catchments so that they look like each other. That is, as the number of independent similitude conditions increases, the likelihood of two catchments being similar (i.e., all nondimensional numbers equal) decreases. Computer generated and experimentally obtained results have less generality since the

207

degrees of freedom in the scaling are reduced. But, in reality, this lack of generality may be illusory. It may be that some processes have no significant effect on certain aspects of catchment form. Requiring similarity of these insignificant processes is an unnecessary constraint, and only serves to reduce the generality of the data under consideration. This is the rationale for developing the "minimal" model that fits the observed data, as noted in the introduction to this chapter. The less processes that are found to be important in generating the catchment, and controlling its form, the greater the generality of the results. This is a strong, pragmatic incentive to reduce the number of processes modelled to a minimum.

The second important point to note is that at different scales, different processes may be dominant. For instance, Schumm, et al. (1987) noted that at the small scale of the Parker (1977) experiments, the overland transport processes are dominated by rainsplash. At the typical field scale, however, hillslope processes are dominated by fluvial transport and rainsplash effects are insignificant (Dunne, 1988). Thus measurements made at one scale may be dominated by a process unimportant at another scale. In the parameter estimation literature this problem is referred to as the parameter observability problem; the data does not contain sufficient information about the processes of interest. The relative magnitude of rainsplash and sediment transport will be examined by use of the scales assigned in the nondimensional analysis. From Equation (6.13a) the relative magnitude of the diffusion and fluvial sediment transport may be expressed in terms of scales as

$$\left[\frac{\text{diffusion}}{\text{transport}}\right] \sim \frac{TD}{TS} = DED = \left(\frac{L_g TD_z}{L_x^2}\right)\left(\frac{T_R \, \rho_s (1-n) \, L_g^2}{T \, L_x^{2m_1-n_1} \, L_z^{n_1-1} \, L_R^{m_1} \, \overline{\beta}_1}\right)$$

$$= \left(\frac{T_R \, \rho_s (1-n) \, L_g^3}{L_x^{2m_1-n_1+2} \, L_z^{n_1-1} \, L_R^{m_1} \, \overline{\beta}_1}\right)\left(\frac{D_z}{\overline{\beta}_1}\right) \qquad (6.29)$$

The relative magnitudes of diffusion and fluvial transport in two different catchments (1) and (2) may be compared by using the scaling techniques just described, so that

$$\frac{\left[\dfrac{\text{diffusion}}{\text{transport}}\right]_{(1)}}{\left[\dfrac{\text{diffusion}}{\text{transport}}\right]_{(2)}} = \left(\frac{T_{R(1,2)} \, \rho_{s(1,2)} \, n'(1,2) \, L_{g(1,2)}^3}{L_{x(1,2)}^{2m_1-n_1+2} \, L_{z(1,2)}^{n_1-1} \, L_{R(1,2)}^{m_1}}\right)\left(\frac{D_z}{\overline{\beta}_1}\right)$$

$$(6.30)$$

We have already noted, in Chapter 5, that typical values of m_1 and n_1 are 1.5 and 2, respectively, so that the relative magnitudes of diffusion and transport can be approximated by the expression

$$\frac{\left[\dfrac{\text{diffusion}}{\text{transport}}\right]_{(1)}}{\left[\dfrac{\text{diffusion}}{\text{transport}}\right]_{(2)}} = \left(\frac{T_{R(1,2)} \, \rho_{s(1,2)} \, n'(1,2) \, L_{g(1,2)}^3}{L_{x(1,2)}^3 \, L_{z(1,2)} \, L_{R(1,2)}^{1.5}}\right)\left(\frac{D_{z(1,2)}}{\overline{\beta}_{1(1,2)}}\right)$$

$$(6.31)$$

From this expression it can be seen that diffusion will be much less important than fluvial transport in catchment (1) compared to catchment (2) if either $L_{x(1,2)}$, $L_{z(1,2)}$, or $L_{R(1,2)}$ are large. That is, if catchment (1) is larger, steeper or wetter than

catchment (2), then diffusion process will be less influential. The relative sensitivities to these three effects is shown by the power of the appropriate scaling coefficient. In the case of the Parker (1977) results the most important effect is the relative magnitude of the horizontal scales of catchments. This is a very important effect when interpreting small scale experimental data. It is believed that these qualitative differences in dominant runoff processes in the Parker experiments are due to the differences in the applicable horizontal length scales; in this case, the length of the hillslope. This will be demonstrated with some hypothetical, but realistic, calculations. Figure 6.1 shows the experimental catchment of Parker (1977) with the scales of some typical experimentally generated networks (from Schumm, et al. 1987). This figure also shows a hypothetical field scale catchment with some typical scales that might be observed in this case. Substitution of these results into Equation (6.31) yields

$$\frac{\left[\dfrac{diffusion}{transport}\right]_{(1)}}{\left[\dfrac{diffusion}{transport}\right]_{(2)}} \sim \frac{1}{10^{11}}$$

where   (1) $\equiv$ hypothetical field catchment

             (2) $\equiv$ Parker (1977) experimental catchment

This result indicates that diffusive effects are 10^{11} times less important in the hypothetical field catchment than they were in the Parker (1977) experiments; a significant difference by any standard.

6.4 Determination of the Scales Used in Nondimensionalization

The previous sections of this chapter have outlined the mechanics and use of the nondimensionalization of the governing equations used in this work. It has also
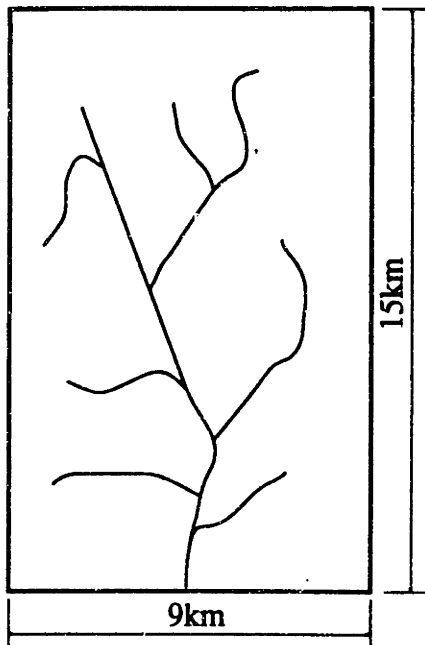
(a) Field scale catchment

Runoff Rate : $L_R \approx 10\text{mm/hr}$

Catchment Slope : S ≈ 0.0015

Horizontal Length Scale : $L_x \approx 15000$ m
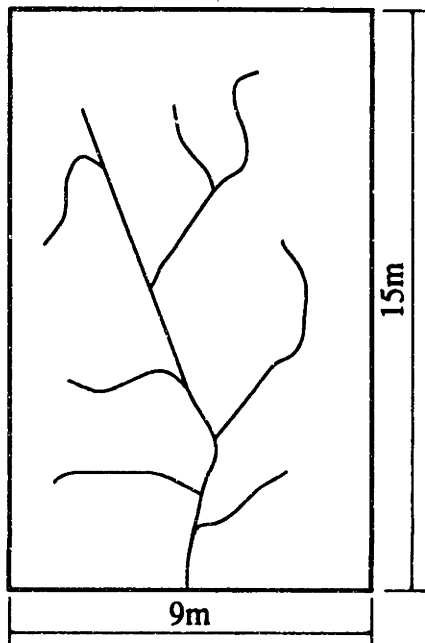
(b) Parker and Mosley model catchments

Runoff Rate : $L_R \approx 50\text{mm/hr}$

Catchment Slope : S ≈ 0.1

Horizontal Length Scale : $L_x \approx 15$ m

Figure 6.1: Comparison of catchment scales for field and plot scale catchments

demonstrated some basic properties of the scaling behavior of these equations for different physical situations. In all of these discussions, however, a detailed explanation of the physical interpretation of the scales used in the nondimensionalization has been postponed until later.

This section addresses the issue of the selection of appropriate scales in the nondimensionalization. The discussion that follows concentrates on relating the scales introduced in the previous sections to the descriptive statistics that are commonly used in the geomorphology literature, e.g. drainage density, relative relief. As a practical matter the connection between the nondimensionalized equations and the catchments at the field scale will be formalized. The horizontal length scales will be discussed first, followed by the vertical length scales; these two length scales being the most commonly discussed in the literature. The runoff length scale will be briefly discussed. This section will close with a discussion of timescales, an area that has received scant attention in the literature.

An important question that will arise in this section is whether there exists a single horizontal length scale and a single vertical length scale. Several candidates for each of these scales will be identified. That there should exist a single fundamental length scale from which all other are derived is very important. This is because the nondimensionalization of Section 6.2 assumes a single horizontal length scale, L_x, and a single vertical length scale, L_z. If there were more than one fundamental length scale, not only would the nondimensionalization have to be changed to reflect this fact but also the relative contribution of each of the scales in each of the physical terms would need to be determined. This latter issue would be a major stumbling block. Just as the nondimensional numbers of equation (6.15) and (6.19), and thus the similitude conditions, depend upon the physics of equations (5.1) so the similitude conditions would depend upon the relative contribution of the scales. Thus proving the existence of a single horizontal and a single vertical length scale is a crucial component to

justifying the nondimensionalization presented in this chapter.

It is commonly perceived that the mean hillslope length and hillslope drop are invariant within the catchment. This is not true. In fact, the activator number defined using the hillslope length scales, TA_h, is the invariant measure of the landscape. Because of this the mean hillslope length is inversely related to the runoff rate and the hillslope slope, so that the hillslope length may vary even within a geologically homogeneous catchment. This question is addressed in greater detail in Section 7.3.1.

6.4.1 Horizontal Length Scales

The most fundamental statistic describing the planar form of networks is drainage density. Drainage density is defined as

$$D_d = \frac{\ell}{A}$$ (6.32)

where ℓ = total length of channels in the region Ω

  A = area of region Ω

Drainage density is a dimensional property with dimension $\frac{1}{L}$. High drainage density means that the drainage network is dense in the region, while low drainage density means that the network is sparse. Horton (1945) suggested that drainage density is a fundamental scale of the catchment since half its inverse is the mean hillslope length. From this argument it follows that the hillslope length scales of this catchment (denoted by (1) and (2)) are related by

$$\frac{D_{d(1)}}{D_{d(2)}} = \frac{L_{x(2)}}{L_{x(1)}} = \frac{1}{L_{x(1,2)}}$$ (6.33)

where $\qquad L_{x(1,2)} \qquad$ = horizontal length scale factor (see Equation (6.24b)

$\qquad\qquad L_{x(1)}, L_{x(2)} \qquad$ = horizontal length scales of the two catchments (1)

$\qquad\qquad$ and (2) respectively.

In this case L_x is considered to be the mean hillslope length (see Figure 6.1).

Another characteristic horizontal scale is the size of the catchment. This could be the maximum length of the catchment, or the maximum stream length, or square root of the catchment area to name a few (Figure 6.1). These are all valid horizontal length scales, but they are fundamentally different to the hillslope length scale. They are not fundamental to the landscape form because they measure the scale, or size, of the study catchment. The size of the catchment can vary, but with the drainage density constant, all that would change would be the Strahler order, magnitude or some other measures of the scale of the channel network. The length scale L_d was introduced in Equation (6.13) to represent the length scale of size of the domain. The hillslope length scale, L_x, and the catchment length scale, L_d, were related by the equation

$$L_d = \alpha_1 L_x \qquad\qquad (6.34)$$

where α_1 can be considered to be analogous to the order or magnitude of the drainage network and was previously given the name Catchment Scale Number. Consider how this α_1 is related to the length scale L_x as defined by drainage density. Equation (6.34) may be expressed as

$$L_d = \frac{\alpha_1}{2D_d} \qquad\qquad (6.35)$$

where $\qquad L_x \qquad = \frac{1}{2D_d} =$ mean hillslope length

It is possible to relate α_1 and the magnitude, m, of the network if it is noted that the total stream length of a network of magnitude m is

$$\ell = (2m - 1)\,\ell_\ell \tag{6.36}$$

where ℓ_ℓ = mean link length, assumed constant and independent of magnitude

         m   = magnitude of the network

The total area of a catchment draining to a network of magnitude m can be expressed as

$$A = (2m - 1)\,A_\ell$$

where A_ℓ = mean area draining to a link, assumed constant and independent of magnitude

so that

$$D_d = \frac{\ell_\ell}{A_\ell}$$

Using these expressions for network area and length and the definition of drainage density yields

$$\alpha_1 = 2L_d\,\frac{\ell_\ell}{A_\ell} \tag{6.37}$$

and for the case of a square catchment of side length L_d (i.e $A_\ell = L_d^2 / (2m-1)$) this yields

$$\alpha_1 = 2(2m-1) \left(\frac{\ell_\ell}{L_d}\right) \qquad (6.38)$$

and if the mean link length ℓ_ℓ and the mean hillslope length L_x are related (i.e. $\ell_\ell = \eta L_x$) then from equation (6.34)

$$\alpha_1 = \sqrt{2(2m-1)\eta}$$

Thus α_1 is just a nondimensional number measuring the scale (e.g. magnitude) of the channel network so that the length scale of the catchment, L_d, and the hillslope length scale, L_x, can be related by the channel network from Equation (6.34) as

$$L_d = \sqrt{2(2m-1)\eta} \; L_x$$

In a similar fashion an expression can be obtained relating the scale factor α_1 and Strahler order. The total stream length of a complete network of Strahler order n is given by

$$\ell = \frac{\ell_1}{(1 - R_\ell R_b)} (1 - (R_\ell R_b)^n)$$

where ℓ_1 = mean length of lst order Strahler stream

       R_b, R_ℓ = Strahler bifurcation and length ratios respectively.

The total area of the complete nth order Strahler basin is

$$A = \frac{A_1}{R_A - 1} (R_A^{n-1} - 1)$$

where A_1 = the mean contributing area of 1st order Strahler stream

R_A = Strahler area ratio

Using these expressions and the definition of drainage density yields the relationship between the scale factor α_1 and the order n of

$$\alpha_1 = 2L_d \left(\frac{\ell_1}{A_1}\right) \left(\frac{R_A - 1}{R_\ell R_b - 1}\right)\left(\frac{(R_\ell R_b)^{n-1} - 1}{R_A^{n-1} - 1}\right) \qquad (6.39)$$

and if $R_\ell R_b > 1$ and $R_A > 1$ which is normally true, then the limit for an infinite order catchment is

$$\lim_{n \to \infty} [\log(\alpha_1)] = \log[2\, L_d \left(\frac{\ell_1}{A_1}\right) \left(\frac{R_A - 1}{R_\ell R_b - 1}\right) \left(\frac{R_A}{R_\ell R_b}\right)] + n \log \left(\frac{R_\ell R_b}{R_A}\right) \qquad (6.40)$$

Note that for α_1 to increase with order, as expected, then $R_\ell R_b \geq R_A$. For a square catchment with length of sides L_d, Equation (6.40) simplifies to

$$\log \alpha_1 = \log [2\, \frac{\ell_1}{A_1^{1/2}} \frac{(R_A - 1)^{1/2}}{(R_\ell R_b - 1)} \frac{R_A^{1/2}}{R_\ell R_b}] + n \log \left(\frac{R_\ell R_b}{R_A^{1/2}}\right) \qquad (6.41)$$

It should be emphasized that Equations (6.39) through (6.41) are only true for complete Strahler networks. A complete network of order n is one that occurs immediately upstream of the confluence of two n order streams or occurs when an nth order stream flows into a higher order stream.

It has been shown how the two horizontal scales of a catchment, one the mean length of hillslopes, the other the size of the catchment can be related through the catchment scale number, α_1. This scale number α_1 has been shown to be directly related to both the magnitude of the network and the Strahler order of the network, two common measures for the topological size of the catchment.

6.4.2 Vertical Length Scales

Another fundamental feature of catchments and their channel networks is their elevation properties. Many statistics parameterize the elevation characterized in terms of slopes, which involve both the horizontal and vertical length scales, rather than elevation alone. Simpler statistics using elevation alone will be dealt with first. But first some preliminaries.

For horizontal length scales it has been noted that there are, in fact, two important length scales. These length scales were the fundamental horizontal length scale, the mean hillslope length, and the catchment basin length scale parameterizing the size of the study catchment. A similar situation exists for the vertical length scales. The first vertical length scale, believed to be the fundamental vertical length scale of the landscape is the vertical length scale of the hillslopes. This may be parameterized by, for instance, mean hillslope drop. The second vertical length scale is that of the study catchment, or channel network. This second length is a function of the size of the catchment. Presumably these two length scales are related, the catchment vertical scale being a function of order or magnitude and the hillslope vertical scale. Very little work has been done in this area, and the issue of relating the two vertical scales will be returned to later, after discussion of the various measures of vertical scale used by geomorphologists.

In the work that follows the catchment vertical length scale (e.g. catchment

relief) will be represented by L_e and the fundamental vertical length scale (e.g. mean hillslope drop) by L_z.

One vertical length scale involving elevation alone is the maximum basin relief H. The maximum basin relief is defined as the difference between the elevation at the mouth of the catchment and the highest point on the catchment perimeter. This is a catchment vertical scale. The scaling relationship, for two catchments (1) and (2), using the catchment vertical length scale factor $L_{e(1,2)}$ is

$$H_{(2)} = \frac{1}{L_{e(1,2)}} H_{(1)}$$

Another simple statistic involving elevation alone and which has attracted recent attention is the Strahler stream drop (Yang, 1971; Tarboton, et al., 1988). The former author found that the mean vertical drop of a Strahler stream of any given order was constant, and independent of order. The latter authors, using channel networks derived from digital terrain data, found that there exists a threshold on drainage density above which this stream drop constancy rule was no longer satisfied. They interpreted this threshold drainage density as a fundamental horizontal length scale. In the context of this work a horizontal length scale factor based on this scale could be used in Equations (6.33) as previously described. The Strahler stream drop can be used as the catchment vertical length scale and a catchment vertical length scale factor $L_{e(1,2)}$ derived so that

$$\Delta z_{(2)} = \frac{\Delta z_{(1)}}{L_{e(1,2)}}$$

where $\quad \Delta z_{(1)}, \Delta z_{(2)} \quad$ = Strahler stream drop for two catchments (1) and

(2) respectively.

$L_{e(1,2)}$ = catchment vertical length scale factor.

This length scale measures the vertical scale of the channel network. Potentially the drop could be related to the hillslope vertical scale, the fundamental vertical scale, through the catchment magnitude. Some relationship between the channel and the hillslope vertical length scales would clearly be needed to achieve this. A relationship based on analyses of catchment simulations will be proposed in Chapter 7

There are a number of other statistics that are used to characterize the vertical scale, that implicitly involve a slope in some generic way. The incorporation of slopes means that these statistics characterize the ratio of the vertical and horizontal scales (L_z/L_x) rather than the vertical scale alone. If the appropriate horizontal scale can be determined by techniques previously described, then these statistics can be used to determine the vertical scale. It is important, however, that the horizontal length scale used to determine the horizontal length scale factor is consistent with that used in the definition of the "slope". It is here that the distinction between catchment scales and the hillslope scales must be clearly made. In addition, it will be shown in Chapter 7 that there are a number of measures of hillslope horizontal scale, all of which give different values.

Schumm (1956) defined the relief ratio (in percent) as

$$R_h = \frac{H}{L} \tag{6.42}$$

where $\quad$ H $\quad$ = maximum basin relief

L $\quad$ = maximum length of the basin parallel to the principal drainage direction

R_h $\quad$ = Schumm (1956) relief ratio

If the maximum length of the basin, L, and the catchment scale, L_d, are linearly

related then this catchment relief ratio can be scaled by the relationship

$$R_{h(2)} = \frac{L_{e(1,2)}}{L_{d(1,2)}} R_{h(1)}$$

(6.43)

Melton (1957) provided an alternate definition of the relief ratio as

$$R_{hp} = 100 \frac{H}{P}$$

(6.44)

where   P   = basin perimeter

        R_{hp} = Melton (1957) relief ratio

Dimensionally this equation appears identical to Equation (6.36), so that the scaling of the relief ratio is the same as Equation (6.43). It should be noted, however, that the basin perimeter, P, has been noted to have a fractal characteristic (Strahler, 1964) in a similar fashion to that observed for the stream channels (Hack, 1957) so that if

$$P \sim X^d; \quad d > 1$$

(6.45)

where   X   = ruler length as defined by Mandelbrot (1983)

        d   = fractal dimension

then the scaling relationship for Melton's relief ratio would be

$$R_{hp(2)} = \frac{L_{e(1,2)}}{L_{d(1,2)}^d} R_{hp(1)}$$

(6.46)

Strahler (1964) introduces a "nondimensional" ruggedness number to describe

how rugged the combination of channel networks and hillslopes are. It is defined as

$$R_g = H_g D_d$$
(6.47)

where $\quad R_g \quad$ = Strahler (1964) ruggedness number

$\quad\quad\quad\quad D_d \quad$ = drainage density

$\quad\quad\quad\quad H_g \quad$ = basin relief

Dimensionally the ruggedness number, R_g, is a slope and the catchment and hillslope slope scales are confused (drainage density is dependent on the hillslope horizontal scale, L_x, while the basin relief is dependent on the catchment vertical scale, L_e), so that the scaling relation is

$$R_{g(2)} = \frac{L_{e(1,2)}}{L_{x(1,2)}} R_{g(1)}$$
(6.48)

Clearly from the previous discussion R_g is dependent on the magnitude of the basin, through the relationship between L_e and L_z. As the magnitude of the network, and consequently L_e, increases, R_g increases.

Strahler (1964) defines another (this time truly nondimensional) ruggedness number, which is dependent on the hillslope scales. In this definition ruggedness number is defined by Equation (6.47) is divided by a hillslope slope. The basin relief used is the difference in elevation between the river and the watershed so that the ruggedness number characterizes the hillslopes rather than the channel network. The equation is

$$R_g = \frac{H_g D_d}{S_h}$$
(6.49)

where S_h = some representative hillslope slope (e.g maximum hillslope slope, mean slope).

This ruggedness number is truly nondimensional and magnitude independent because all of the inputs are hillslope scales. Since this ruggedness number is truly nondimensional, it does not scale and is constant across all basins with similar hillslopes. It is thus not useful for determining vertical scaling relationships though it may be useful for classifying landscapes (see Section 6.5).

It is clear that there are variety of ways of parameterizing the vertical relief of a catchment. It is equally clear that there has, in the past, been no clear distinction between the catchment vertical scale and the fundamental vertical scale. There is no clear consensus on this distinction. This discussion has blurred the distinction of the fundamental horizontal and vertical scales for the purposes of simplicity rather than from any experimental evidence. It will now be shown that there does appear to be a connection between the catchment vertical length scale and the hillslope vertical length scale.

In the previous section a derivation relating the catchment and hillslope horizontal scales and the Strahler order was presented. In a similar fashion a speculative relation between vertical length scales will now be derived. Strahler (1964) noted a relationship between hillslope slope and channel slope of the form

$$\theta_g = 4\ \theta_c^{0.8}$$

where θ_g = slope of hillslope region

θ_c = slope of the channel draining the same hillslope region

If this relationship is applied to the lst order Strahler area, then

$$\frac{L_z}{L_x} = 4 \, S_1^{0.8} \tag{6.50}$$

where $\quad$ S_1 $\quad$ = mean channel slope of Strahler 1st order streams

$\qquad\qquad$ L_z $\quad$ = hillslope vertical length scale

$\qquad\qquad$ L_x $\quad$ = hillslope horizontal length scale.

If it is now assumed that the maximum drop from channel source to catchment outlet results from the stream that flows from order 1 to order 2 to ... to order n, in an nth order catchment, then the value of catchment relief can be related to the channel. The catchment relief R is given as

$$R = \sum_{i=1}^{n-1} \left(\frac{R_\ell}{R_s}\right)^{i-1} S_1 \, \ell_1 + L_z = L_e \tag{6.51}$$

where $\qquad$ S_1, ℓ_1 $\qquad$ = average slope and length, respectively, of the 1st order

$\qquad\qquad\qquad\qquad$ Strahler stream

$\qquad\qquad$ R_s, R_ℓ $\qquad$ = Strahler ratios for slope and length respectively.

Evaluating the summation in this equation and substituting for the 1st order Strahler slope of Equation (6.50) yields a relationship between the catchment vertical length scale and the hillslope vertical length scale.

$$L_e \quad = L_z \left[L_z^{0.25} \left| 1 - \frac{R_\ell}{R_s} \right|^{n-1} \left(\frac{\ell_1}{5.62 \, L_x^{1.25} \left(1 - \frac{R_\ell}{R_s}\right)} \right) + 1 \right]$$

$$\quad = \alpha_2 \, L_z \tag{6.52}$$

where

$$\alpha_2 = [L_z^{0.25} \, |1 - \frac{R_\ell}{R_s}|^{n-1} \, (\frac{\ell_1}{5.62 \, L_x^{1.25} \, (1 - \frac{R_\ell}{R_s})}) + 1]$$

If $R_\ell < R_s$ then α_2 increases with catchment order. In the case where $R_\ell = R_s$, so that the mean Strahler stream drops are independent of order, the case described by Tarboton, et al. (1988b), the relationship between the catchment and hillslope vertical length scales is different and given by

$$L_e = L_z \, (\frac{n \, L_z^{0.25}}{5.62 \, L_x^{0.25}} + 1) = \alpha_2 L_z \qquad (6.53)$$

where
$$\alpha_2 = (\frac{n \, L_z^{0.25}}{5.62 \, L_x^{0.25}} + 1).$$

The relationships of Equation (6.52) and (6.53) relate the catchment and hillslope length scales on the basis of the catchment Strahler order. An equivalent relationship based on catchment magnitude has not been previously presented, however, a proposed relationship between catchment slopes, areas and magnitudes will be discussed in Section 8.3

It is important to note that Equations (6.52) and (6.53) are speculative. There does not exist any experimental work in the literature to substantiate the exact form of these relationships. Nevertheless these equations and Equations (6.35b) and (6.38) which relate the catchment and hillslope horizontal length scales via Strahler order and magnitude are important conceptually. These relationships indicate that the statistics of drainage density, catchment size, catchment relief, hillslope relief and Strahler order or magnitude are all related. We have already shown the relationship

$$L_d = \alpha_1 \text{ (order/magnitude)} \cdot L_x \qquad (6.54a)$$

$$L_e = \alpha_2 \text{ (order/magnitude, } L_x, L_z) \cdot L_z \qquad (6.54b)$$

Consider a case where the drainage density and catchment size are given. The relationship of Equation (6.54a) yields the order and the magnitude of the channel network. The relationship of Equation (6.54b) indicates that only one of L_z or L_e need to be measured, the other follows from the order and magnitude of the channel network. Thus of the four degrees of freedom in the horizontal and vertical length scales, Equation (6.54) implies that there are only 3 independent degrees of freedom. This issue will be returned to in Chapter 7 where further relationships will be developed on the basis of the computer simulations.

The discussion thus far has restricted itself to determining the characteristic horizontal and vertical length scales and using them as input for nondimensionalization of the governing equations. It is these length scales that have been most thoroughly quantified in the literature. Other important scales in Equations (6.1) through (6.5) are runoff rate, differentiation/channel growth timescale, and elevation change timescale. These will be dealt with briefly below.

6.4.3 Runoff Length Scales.

The runoff rate is, superficially, an easy property for which to determine the scaling coefficient L_R. The runoff rates for the two catchments are determined and Equation (6.3) is used to determine L_R. This assumes that the dominant discharge is proportional to area as in Equation (6.5). Realistic catchments that have these characteristics would be those formed in the rainfall–erosion facility (REF) at Colorado State University (Parker, 1977; Schumm, et al., 1987).

Natural catchments are unlikely to obey such a simple relationship between discharge and area. If the dominant discharge is considered to be the bankful

discharge, then a more reasonable relationship would be (McDermott and Pilgrim, 1983)

$$Q = \beta_4 A^{m_4} \qquad m_4 \, \epsilon \, [0.5, 0.8] \qquad\qquad (6.55)$$

where β_4 = frequency factor for bankful discharge

The factor β_4 can be considered to be equivalent to the runoff rate but the dimensions are different due to the areal dependence of rainfall. As previously discussed this areal dependence is poorly understood so this case will not be further discussed.

This concludes the discussion of transport processes that are important for <u>both</u> the transient and the dynamic equilibrium cases.

6.4.4 <u>Timescales of Elevation Change.</u>

The timescales T in Equation (6.4), and d_t in Equation (6.12b) are relevant only for the transient case of Equation (6.12). Except for Brunsden (1980) and Schumm, et al. (1987), there has been an almost complete absence in the literature of explicit consideration of transient effects in network growth and catchment evolution. Dynamic equilibrium is typically implicitly assumed. As a result there are no commonly accepted statistics for parameterizing the two transient effects; elevation change and network/channel growth. The former of the two transient effects is easiest to parameterize from the governing equations and will be dealt with first.

To parameterize the elevation change transients, it is convenient to reformulate Equation (6.12a) as

$$\frac{\partial z'}{\partial t'} = \frac{1}{L_z}(T\, c_0) + \frac{1}{L_x^2}(T\, D_z)\frac{\partial^2 z'}{\partial x_i'^2}$$

$$+ \left(\frac{L_x^{2m_1-n_1}\, L_z^{n_1-1}\, L_R^{m_1}}{T_R^{m_1}\, \rho_s\,(1-n)\, L_g^2}\right)(T\,\beta_1)\sum_i \left[(Q_i')^{m_1}(S_i')^{n_1}\, I_{ij}\, f'(Y_i)\right]$$

<div align="right">(6.56)</div>

or

$$\frac{\partial z'}{\partial t'} = \left(\frac{1}{L_z}\right) c_0' + \left(\frac{L_x^{2m_1-n_1}\, L_z^{n_1-1}\, L_R^{m_1}}{T_R^{m_1}\, \rho_s\,(1-n)\, L_g^2}\right)\beta_1' \sum_i \left[(Q_i')^{m_1}(S_i')^{n_1}\, I_{ij}\, f'(Y_i)\right]$$

$$+ \left(\frac{1}{L_x^2}\right) D_z'\, \frac{\partial^2 z'}{\partial x_i'^2}$$

<div align="right">(6.57)</div>

where $\quad c_0' = T\, c_0$

$\qquad\qquad \beta_1' = T\,\beta_1$

$\qquad\qquad D_z' = T\, D_z$

It is important not confuse c_0', β_1', and D_z' with the nondimensional parameters; for these three parameters the prime simply signifies a transformation. It follows from Equation (6.57) that, for example, a doubling of the timescale in any given catchment results in a requisite doubling of all the transport coefficients, if similitude is to be maintained. Alternatively, two catchments with transport coefficients $c_{0(1)}$, $\beta_{(1)}$, $D_{z(1)}$, $c_{0(2)}$, $\beta_{(2)}$, $D_{z(2)}$, related by a constant of proportionality $T_{(1,2)}$ (see Equation 6.25 for notation) are similar and only differ in the timescale $T_{(1)}$ and $T_{(2)}$ at which the processes are measured. This follows from Equation (6.57). For example, consider · a catchment A with transport coefficients double that of some other catchment B, but otherwise identical. Catchments A and B are similar except that time appears to

<div align="center">228</div>

proceed twice as fast in catchment A than in catchment B. In terms of nondimensional age, t', catchment A has a nondimensional age twice that of catchment B.

6.4.5 Timescales of Network Growth

The other major timescale of importance is the timescale of network growth, $1/d_t$. Parameterizing the timescale of network growth from the governing equations is not possible. It should be noted that $1/d_t$, the inverse of the rate constant d_t in the governing Equations (6.12), is the timescale of <u>channel</u> growth rather than the timescale of <u>network</u> growth. This distinction is important since network growth consists of two interactive processes.

1.   The rate at which the high activator region advances in front of the growing channel head.

2.   The rate at which the channel grows at any particular point in space once the activator threshold is exceeded.

The second of these two processes is parameterized by the channel growth timescale $1/d_t$.

The first of these processes is primarily dependent on the rate of overland erosion. The channel growth rate plays a minor role. With very low overland erosion rates, the high slope (and thus high activator) region cannot proceed very far in advance of the growing channel head so that the channel extends in space slowly. With very high overland erosion rates, the overland area is eroded so quickly that slopes are never high enough to exceed the activation threshold even though the slope advance is very fast. Again the channel extends in space very slowly. At some, as yet undefined, intermediate overland erosion rate, the channel extension rate and network growth rate are maximum. It is apparent that parameterization of network growth

rates will need to be performed on simulated networks rather than from the governing equations and the network growth rate thus obtained will be empirical in nature (see Chapter 7).

One possible way to parameterize network growth is by the rate of change of drainage density, so that

$$\left(\frac{dD_d'}{dt_N'}\right)(2) = L_{x(1,2)} \; T_{N(1,2)} \left(\frac{dD_d}{dt}\right)(1) \tag{6.58}$$

where the network growth timescale, T_N, is defined in an analogous fashion to the elevation change and channel growth timescales so that

$$t_N' = \frac{t}{T_N} \qquad : \text{Nondimensional network growth timescale}$$

$$T_{N(1,2)} = \frac{T_{N(1)}}{T_{N(2)}} \tag{6.59}$$

Another definition of network growth is simply in terms of the rate of increase in the length of the network so that

$$\frac{dL_N'}{dt_n'} = \frac{T_N}{L_x} \frac{dL_N}{dt} \tag{6.60}$$

This relationship assumes that channel length is not fractal. As noted previously a definition based on channel length suffers from the difficulty of determining whether the channel length is fractal or not, and if so determining the dimensionality.

6.5 Similitude and Geomorphological Landscape Classification

Finally, another application of the nondimensionalization of this chapter is to the problem of the classification of landscapes. That is, which landscapes are similar, and which are fundamentally different. This is a subject that has held the interest of geographers and geomorphologists since the turn of the century (Davis, 1924; Scheidegger, 1987). It has often been the case that these researchers have provided conflicting interpretation of what seem to be similar observed data. A possibility is that these data, and the interpretations based on them, are not necessarily conflicting but result from the different underlying assumptions made by different researchers when they have analyzed the data.

Consideration of the nondimensionalized equations for catchment evolution presented in this chapter will demonstrate that the varying views of catchment evolution of different researchers are consistent, if somewhat simplistic. In addition, some erroneous ideas regarding classification of landscape types will be corrected by use of similarity analysis.

The nondimensional governing equations of a catchment at deterministic dynamic equilibrium are given by Equation (6.17). To simplify the analysis that follows diffusive processes will be neglected; this is assumed to be a good approximation at the field scale. To further simplify the analysis, it will be assumed that the tectonic uplift is constant with time, but variable in space i.e.

$$c_0(x,t) = \overline{c}_0 \, c_0'(\underline{x})$$

where $\overline{c}_0$ = tectonic uplift rate scale introduced in Equation (6.11)

$c_0'(\underline{x})$ = nondimensional tectonic uplift variable in space

At dynamic equilibrium the channel network has stopped growing so that the channelization equation may be ignored for all intents and purposes. Of prime interest is the balance between the tectonic uplift and erosion. The conditions of dynamic

equilibrium then follow from Equation (6.17a) and are

$$DET_c = -\frac{1}{c_0'(\underline{x})} \sum_i [(Q_i')^{m_1} (S_{ij}')^{n_1} I_{ij} f'(Y_i)] \quad \forall j$$

(6.61)

where DET_c = the Dynamic Equilibrium Tectonic Uplift/Sediment transport Number (see Equation 6.18) defined using the catchment length scales L_d and L_e.

The nondimensional catchment domain is given by

$$\underline{\Omega}' = \frac{1}{\alpha_1^2 L_x^2} \underline{\Omega}$$

where $\underline{\Omega}$ = the original catchment

Equation (6.61) has been arranged so that all the scales of the catchment are on the left–hand side in DET_c, while all the nondimensional catchment characteristics are on the right–hand side. Since all the scales on the left–hand side are independent of time and space, then all differences between catchments at dynamic equilibrium must be parameterized by the individual terms on the right–hand side. From the similarity conditions of Equation (6.20) two similar catchments must, by definition, have Q', S', $f'(Y)$, $c_0'(x)$ equal. Thus DET_c must be equal for two <u>similar</u> catchments. Moreover since Equation (6.61) is the nondimensional catchment downwasting due to fluvial erosion divided by the nondimensional tectonic uplift, and since they must be equal for the nondimensional catchment to be in dynamic equilibrium, then

$$DET_c = 1$$

for _all_ catchments, similar or not.

To examine the question of catchment similarity consider two catchments ((1) and (2)) in which fluvial erosion is the only erosive process. In this case Equation (6.61) is applicable. Because $DET_c = 1$ for all catchments then transport similarity between the two catchments at dynamic equilibrium is automatically satisfied ($DET_{c(1)} = DET_{c(2)}$). However, for two catchments to project onto the same nondimensional catchment (i.e. for two catchments to be similar) then the Catchment Scale Number, α_1, must also be the same (i.e. $\alpha_{1(1)} = \alpha_{1(2)}$) so that the nondimensional drainage densities are the same. If this latter condition is satisfied then the catchments satisfy the physical similarity conditions. In this case the hypsometric curves for the two catchments will be identical, except for some small statistical scatter (see Section 7.2).

The classification of similar landscapes should be made on the nondimensional catchment characteristics contributing to DET_c on the right-hand side of Equation (6.61) and α_1. The scales in DET_c do not form a basis on which to classify the landscape form. This is because all scales, except L_S, result from external sources. That is, L_R results from the runoff and thus the rainfall, L_x the length of the hillslopes and thus the fixed network, $\overline{\beta}_1$ the landscape erodability and thus the regional geology and $\overline{c}_0$ the rate of tectonic uplift. Since DET_c is constant across all catchments, then for two catchments (1) and (2) their slope scales adjust in the following fashion

$$\frac{L_{S(1)}}{L_{S(2)}} = \left[\left(\frac{\overline{c}_0 \ T_R^{m_1} \ \rho_s (1-n) \ L_g^2}{\overline{\beta}_1 \ L_x^{2m_1} \ L_R^{m_1}} \right)_{(1)} \cdot \left(\frac{\overline{\beta}_1 \ L_x^{2m_1} \ L_R^{m_1}}{\overline{c}_0 \ T_R^{m_1} \ \rho_s (1-n) \ L_g^2} \right)_{(2)} \right]^{1/n_1}$$

(6.62)

233

Thus the slopes of the catchment adjust in response to the external inputs. This is the way dynamic equilibrium is maintained; slopes adjust to the uplift and erosional regime. The slopes of two catchments in dynamic equilibrium are related to the other physical inputs.

The central issue of this section is landscape classification, in particular the views of Scheidegger (1987). For the following discussion the similarity conditions in Equation (6.61) are applied and it is shown that Scheidegger's fluvial landscape classification criteria and Equation (6.61) are incompatible. He classifies fluvial landforms on the basis of

1. Tectonic activity – into regions of "high activity," "medium activity," and "low activity."

2. Aridity – into regions of "humid" and "arid" climate.

These classification criteria can be compared with the similarity conditions expressed in Equation (6.61). Logically, those catchments that can be scaled so that they satisfy the similitude conditions (so that nondimensional catchments look the same) should be classified into the same categories. There are 4 characteristics that are spatially distributed and which contribute to DET_c and thus the form of the catchment (Equation 6.61)

1. Channel network form, through the spatial distribution of $f'(Y(\underline{x}))$.

2. Drainage directions, through the spatial distribution of $Q'(\underline{x})$.

3. Slope form, through the spatial distribution $S'(\underline{x})$.

4. Tectonic uplift, through its spatial distribution $c_0'(\underline{x})$

For catchments to be similar all these characteristics must be equal (Equation 6.20). For dissimilar catchments, recalling that $DET_c = 1$, then only 3 of the 4 characteristics are independent. Since $c_0'(\underline{x})$ is externally applied then one of $f'(Y(x))$, $Q'(x)$, and $S'(x)$ must be a dependent characteristic. These characteristics are not the ones used by Scheidegger. Both the tectonic activity and aridity he uses appear as

length scales on the left—hand side of Equation (6.61), as the tectonic uplift scale, $\overline{c}_0$, and the runoff scale, L_R. The importance of this observation is that the scales that Scheidegger uses to classify landscapes are components of the nondimensional number DET_c. Thus the catchment slopes (whose magnitudes are given by L_S) adjust to compensate for the aridity and tectonic activity to ensure that DET_c remains equal to 1 (Equation 6.62).

To exemplify this consider the case where only the runoff length scale, L_R, and the tectonic uplift, c_0, differ between the two catchments (1) and (2). Then the relationship between the slope scales of the catchments is (Equation 6.62)

$$\frac{L_{S(1)}}{L_{S(2)}} = \frac{c_{0(1)}}{c_{0(2)}} \frac{L_{R(2)}^{m_1/n_1}}{L_{R(1)}^{m_1/n_1}}$$

$$= c_{0(1,2)} \, L_{R(1,2)}^{-m_1/n_1} \tag{6.63}$$

For two catchments, (1) being more arid than (2) (i.e $L_{R(1,2)} < 1$) the slopes in catchment (1) will be higher than in catchment (2). A more arid catchment will have higher slopes because of its reduced ability to erode comparative to the tectonic input. If the tectonic uplift in catchment (1) is less than in catchment (2) (i.e. $c_{0(1,2)} < 1$), then slopes will be least in catchment (1), all other processes being unchanged. However, in both cases the form of the catchment will remain unchanged provided the 3 inputs to DET_c remain unchanged. If the spatial pattern of tectonic uplift and runoff is the same in both catchments (1) and (2) then these 4 characteristics will not be changed.

This discussion demonstrates the error of classifying landscapes on the basis of their aridity and tectonic activity.

A complicating issue in the use of Equations (6.61) and (6.62) is that they ignore potential transient effects. In this chapter the distinction between the transient stage of catchment evolution and the dynamic equilibrium stage has been stressed. This distinction follows directly from the governing equations for catchment evolution. This distinction is rarely made in the field because of difficulties distinguishing these two stages. The transient stage is very important, however, since it is at this stage that the channel network is formed and the predominant directions of hillslope drainage fixed. The patterns of erosion are largely fixed once the transient stage is complete. In addition, in Section 8.3 it will be shown that the network growth model adopted in this work produces networks whose branch head positions in space can only be explained by the transient portion of catchment evolution.

Another transient effect arises if slopes are increased by changes in tectonic or runoff input so that active channelization occurs. This factor can be quantified by looking at the scaling of the activator equation (Equation 6.18c) using the activator number.

Finally the assumption of dynamic equilibrium used in this section may be violated. From the nondimensional analysis it is apparent that there is a single degree of freedom at dynamic equilibrium, the vertical length scale of the catchment; the slopes in the catchment adjust to all other physical properties of the catchment, in particular the tectonic uplift. If teconic uplift varies slowly in time compared with erosional processes (i.e. TT < TS) then the catchment may be considered to be in dynamic equilibrium.

Davis (1924) on the other hand thinks in terms of cycles of tectonic events, followed by long periods of erosion. In the governing equations of this work, this corresponds to dirac function inputs of tectonic uplift i.e.

$$c_0(x,t) = \overline{c}_0[\delta(t - \overline{t}_0) + \delta(t - t_1) + \delta(t - t_2) + ...]$$

where $\quad t_0, t_1, t_2$ = times of tectonic uplift events.

This scenario yields a deterministic dynamic equilibrium that is a flat plain. Catchment form can only be sensibly described by use of the transient form of the catchment evolution equations where the catchment is alternately rising and declining with time; dynamic equilibrium is meaningless. This endlessly transient catchment form is one extreme in the geomorphologic evolution of catchments.

Another alternative view, at the other extreme of catchment behavior, is taken by Scheidegger (1987). He thinks in terms of tectonic events occurring concurrently with erosion, with a mean temporal equilibrium being obtained between erosion and tectonic uplift; that is, a landscape permanently in dynamic equilibrium. Neither the view of Davis or Scheidegger is complete. The only complete view is one that explicitly considers transient effects and timescales of adjustment of catchment elevations.

Finally, parenthetically, an observation is made on landscape form in different regions. That many desert areas are quite flat, and many tropical areas are quite mountainous suggests that other processes simultaneously scale to compensate for the aridity effect on catchment slopes so that a scale change in runoff scale cannot be considered independently of other scale changes, such as tectonic uplift. That elevations, and thus orographic effects on rainfall, are quite high in many wet areas (e.g., the Pacific rim archipelago) and low in many desert areas (e.g., Australia) may be an explanation. A first order justification for this assertion, based on Equation (6.61), follows. Eagleson (1970) and Chow (1964) present evidence that rainfall r is approximately proportional to the elevation, E, everything else being equal. In addition, from Equations (6.1) and (6.8) elevation and slope scale in the same fashion,

if the horizontal length scale is constant. Thus a similarity condition relating rainfall, elevation, and slopes may be derived so that

$$\frac{r}{r'} = \frac{E}{E'} = \frac{z}{z'} = L_z = \frac{S}{S'} = L_S \qquad (6.64)$$

If, in addition, it is assumed, to first order, that runoff, $\overline{R}$, is proportional to rainfall, then the runoff and vertical length scales can be related

$$L_R = \frac{\overline{R}}{\overline{R}'} \approx \frac{r}{r'} = L_z \qquad (6.65)$$

It follows from Equations (6.64) and (6.65) that the slope scale and runoff scale are equal (i.e $L_R = L_S$). Substituting this result into Equations (6.22a) and (6.61) and moving β_1 and L_x to the right–hand side (since scaling of sediment transport rate and horizontal lengths is not considered here) then

$$\frac{\overline{c}_0}{L_R^{n_1+m_1}} = \frac{\beta_1 \, L_x^{2m_1}}{\rho_s(1-n) \, L_g^2} \cdot DET_c \qquad (6.66)$$

Since similarity requires that DET_c be constant, then as c_0 (i.e tectonic uplift) is decreased and thus the catchment relief, then L_R (i.e runoff) decreases to compensate. This decrease in runoff follows from the reduced orographic effects on the rainfall (and thus runoff) distribution. To be similar, catchments with low tectonic inputs will tend to be arid and catchments with high tectonic inputs will tend to be wet; this follows from the observed orographic relationship between elevation and rainfall.

CHAPTER 7
RESULTS

7.1 Introduction

This chapter is the first of two dealing with the analysis of catchment simulations over a range of physical parameters. Here the numerous characteristics of the simulations are analyzed. Where possible, statistics are compared with those results obtained by field workers, though in most field studies the appropriate physical parameters are not reported so that the comparisons are largely qualitative. This chapter is not, however, primarily concerned with the comparison of the model with field data; that is the purpose of Chapter 8.

The chapter begins with a detailed consideration of the effects of randomness on statistics that have been used to characterize hillslopes and channel networks. It will be shown that many of the topological measures of channel network form are very sensitive to changes in elevations. A mechanism whereby topologically distinct random networks (Shreve, 1966) arise quite naturally out of the growth process will be proposed.

One of the most important characteristics of the simulation model is its ability to explicitly differentiate between the channels and the hillslopes. The length scales of the exterior links and the hillslope draining to them will be related to the underlying physics of the activator function. Equations will be proposed that relate the drainage density and the rate of network growth to the physical processes acting in the catchment.

Many channel network link properties have been measured by field workers. Unfortunately, there exists a great deal of scatter in their results and many studies are contradictory. Link properties are examined and compared with the typical ranges observed in the field. Some conclusions will be reached regarding the significance, or

otherwise, of proposed correlations. These conclusions will be justified in terms of the governing physics.

The related question of whether there exists a fundamental horizontal scale in the catchment will be addressed. It will be shown that it is possible to relate many of the horizontal length scales, such as mean link length, source area, hillslope length, etc., and that the relationships appear to be relatively independent of the adopted physics.

These correlations together with the drainage density relationships will be used to show that there are only two independent length scales; one vertical and one horizontal. A formulation in terms of the rate of tectonic uplift and total area of the catchment will be presented. All other length scales follow from these. This result will be used to justify the form of the nondimensionalization of the governing equations, presented in Chapter 6.

Hypsometric curves are presented for two different tectonic uplift regimes. The first is an episodic uplift with erosional downwasting, the second is a continuous uplift leading to dynamic equilibrium. Differences in the hypsometric curves which have previously been attributed to age are shown to result from differing uplift histories in catchments.

Finally the effect of hillslope diffusion processes (e.g., creep, landslide, rainsplash) on catchment form is addressed. While the channel network is modified in line with the random growth hypothesis of Section 7.2, it is shown that the overall form of the catchment is relatively unaffected.

7.2 Randomness in Channel Networks

This section is devoted to assessing the randomness that appears to be inherent in channel networks when the physical similarity conditions of Chapter 6 are satisfied. Much work has gone into quantifying this topological randomness (e.g., Shreve, 1966;

Mock, 1971; Mesa, 1986). The link between the channel network growth process and the physics of the runoff and transport within the catchment have largely been ignored. This section will demonstrate that very simple externally applied randomness results in very complicated, and random, network statistics. It will become apparent that those statistics that parameterize the channel network topology exhibit the greatest variability. Those statistics that parameterize more continuous variables such as elevation exhibit less variability.

To demonstrate the random effects, two sets of runs will be examined.

1. The runs CR2–6, CR7–1, CR7–2, and CR7–3. These runs were identical in every respect except that the random field elevation perturbations applied to the initially flat catchments (0.25% of the initial catchment value) were different.

2. The runs CR2–3, CR4–6, CR4–7, CR4–8, CR4–9. These runs were identical except that the vertical scale of the catchment (parameterized by L_e, see Chapter 6) was varied by varying the tectonic uplift, c_0, and that the time step of the numerical solver was varied. These runs display the variability resulting from numerical effects.

The four runs of point 1 satisfy the conditions for transient statistical similarity (i.e. Equation 6.15) so are physically similar. The same is true of the runs in point 2. Both sets of simulations produced networks that were variable in form. The final channel networks are shown in Figure 7.1.

The consideration of randomness will start with the statistics of network topology based on network order. These statistics were summarized in Section 2.2.1.1 where the Horton, Strahler, and Tokunaga ordering models were introduced.

Plots showing the variation with time, as the channel network grows, of Strahlers' bifurcation, length, slope and area ratios are given in Figures 7.2 through
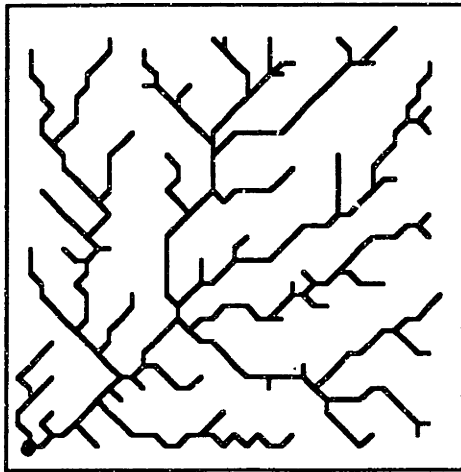
CR2-6
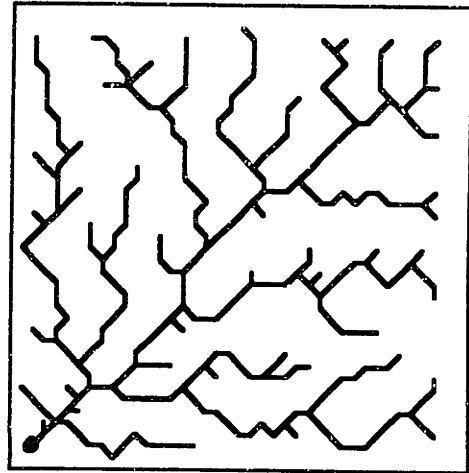
CR7-1
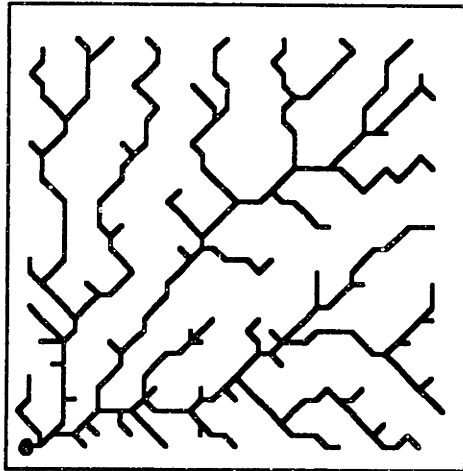
CR7-2
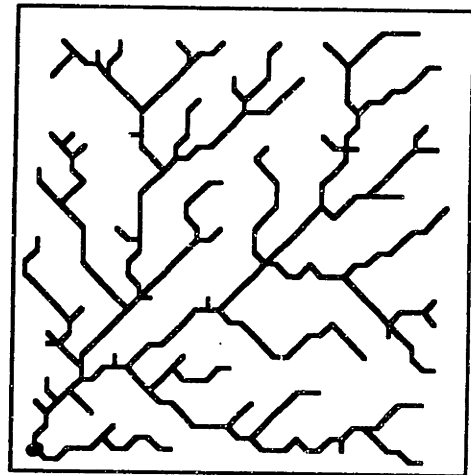
CR7-3

● Catchment outlet

Figure 7.1: (a) random channel networks from
initial elevation perturbations,
simulations CR2-6, CR7-1, CR7-2, CR7-3

242
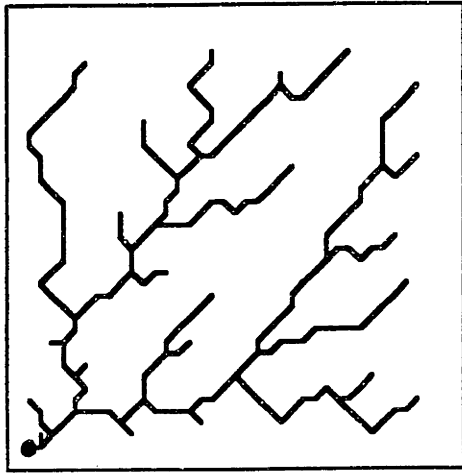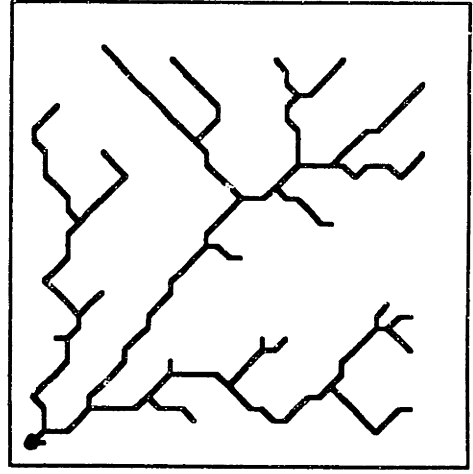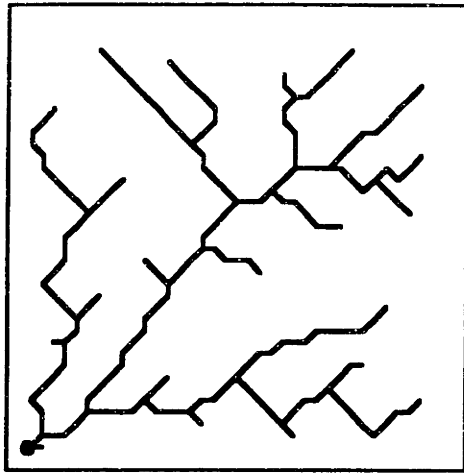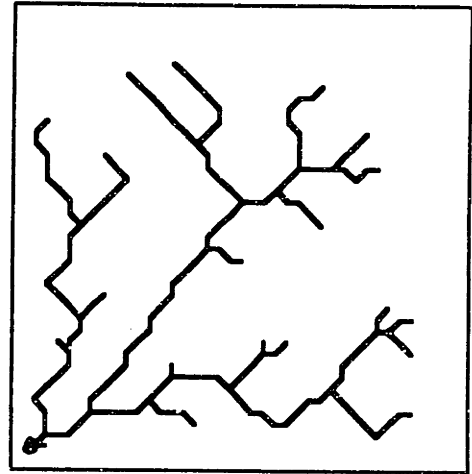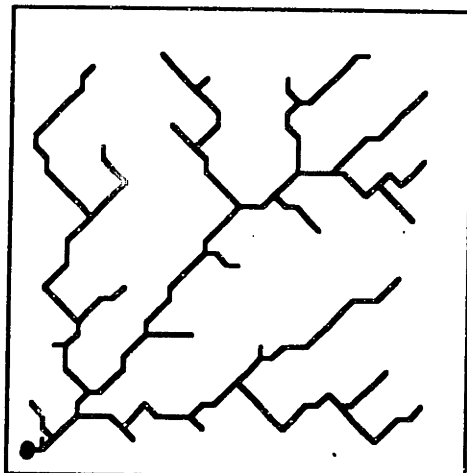
CR2-3

CR4-6

CR4-7

CR4-8

CR4-9

● Catchment outlet

Figure 7.1 (ctd): (b) random channel networks from numerical effects, simulations CR2-3, CR4-6, CR4-7, CR4-8, CR4-9

7.5. The values of the ratios for each order are plotted separately. There is considerable variability in all of the ratios through time, and there is a degree of variability in these statistics even when the network has stopped growing.

One explanation for the variability in the bifurcation ratio, with time, follows. This explanation borrows ideas from the allometric growth hypothesis and the topologically random networks described in Chapter 2. At some time we have a network of magnitude (m). A channel will branch off from the side of an existing channel creating a new network of magnitude (m+1). The location of this branching follows from the chaotic physics controlling network growth. Once that link has branched off the network has grown allometrically. The Strahler bifurcation ratio for these two networks, different by just one link, will vary with each cycle of allometric growth. The very large fluctuations in the bifurcation ratio result from the bifurcation ratios of two networks being significantly different even if the networks differ by just a single link. Kirchener (personal communication) in a study of mapping errors for channel networks looked at the effect of randomly adding a single link to a network. He found that the addition of a single link to a given magnitude 8 network resulted in magnitude 9 networks that were drawn from the population of magnitude 9 topologically distinct random networks (TDRNs) and that the magnitude 9 TDRNs were approximately equally likely. If the TDRNs are equally likely it follows that the bifurcation ratio for magnitude 9 networks will be drawn from the probability distribution for bifurcation ratios exhibited by the TDRNs of Shreve for magnitude 9. Shreve (1966) showed that the range of bifurcation ratios is typically 2.5 to 6.

From the topological viewpoint the final network will be the result of a random allometric growth process. The bifurcation ratio will come from the distribution of bifurcation ratios resulting from Shreve's postulate of equal likelihood of occurrence of each topologically distinct random network.
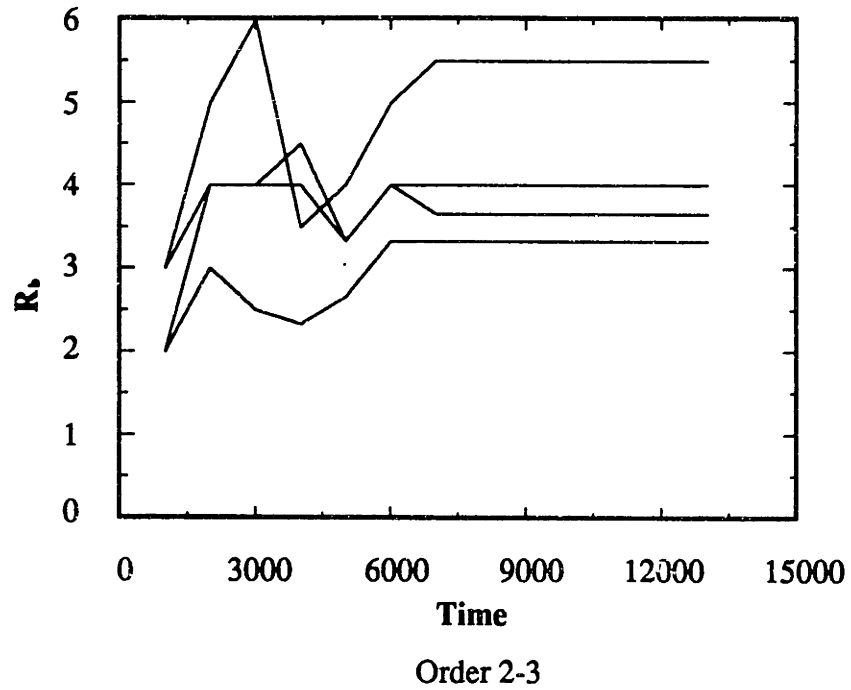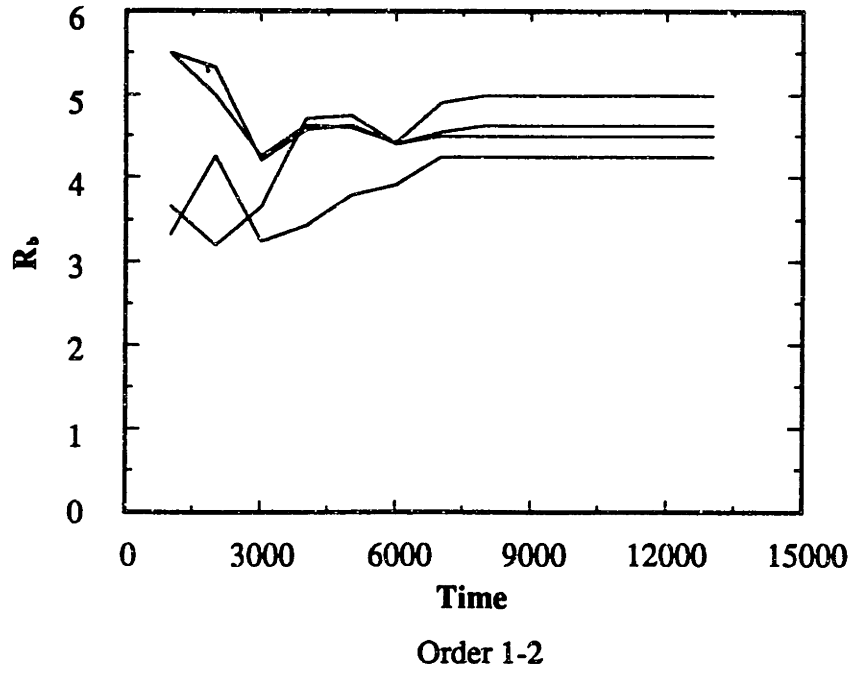
Order 1-2

Order 2-3

Figure 7.2: Strahler bifurcation ratio with time: simulations CR2-6, CR7-1, CR7-2, CR7-3
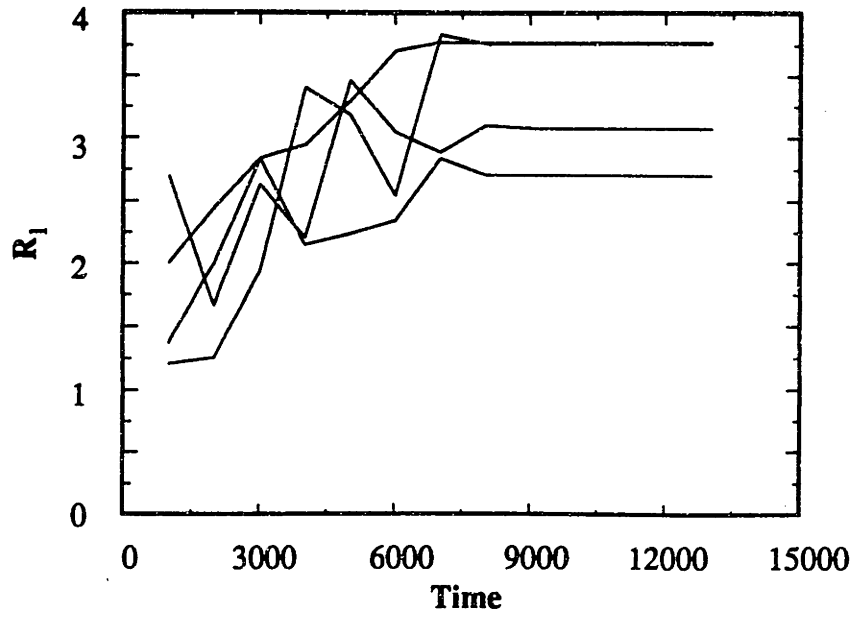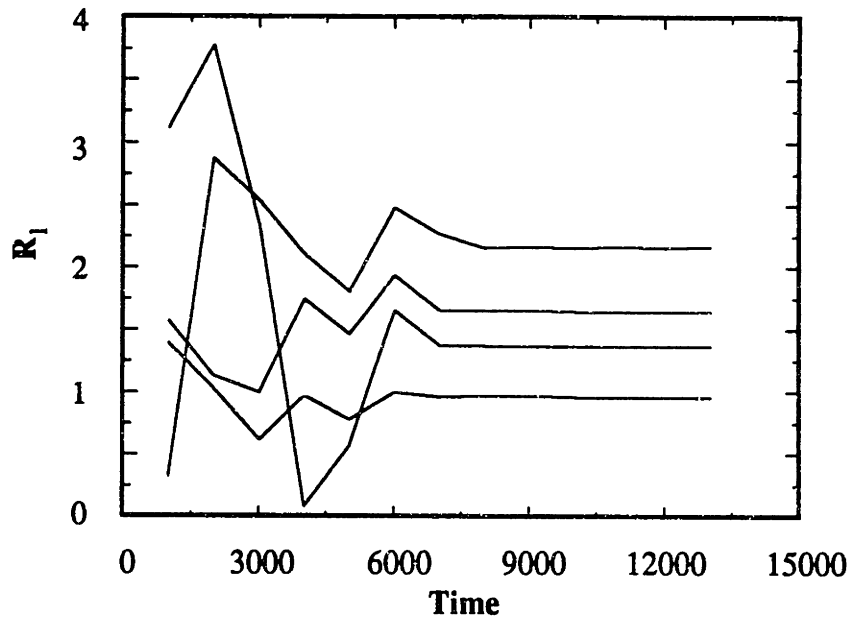
Order 1-2

Order 2-3

Figure 7.3: Strahler length ratio with time:
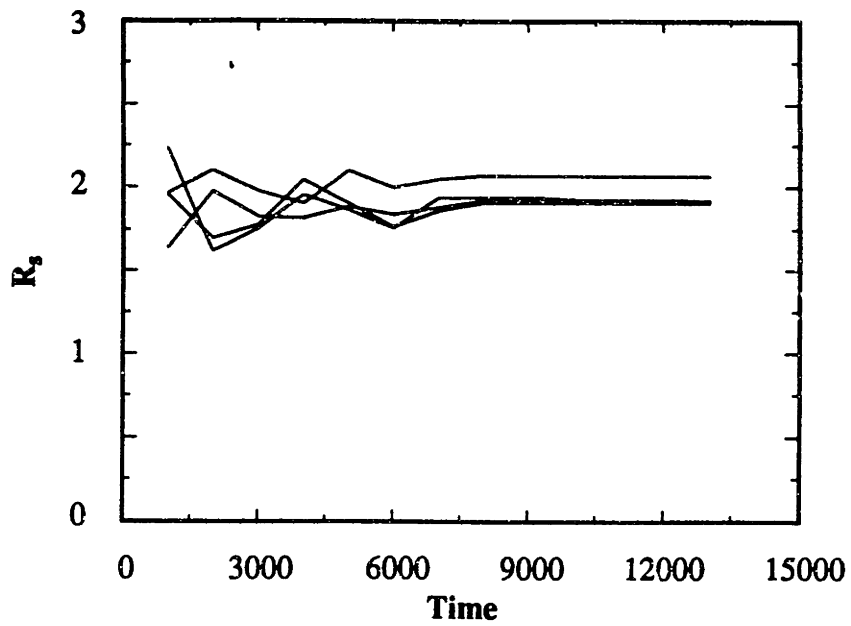
Order 1-2
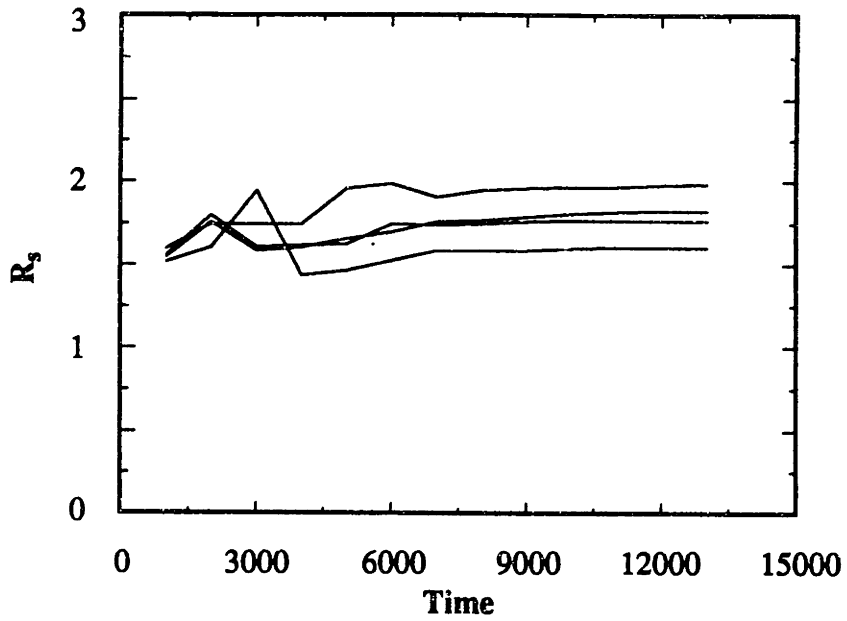
Order 2-3

Figure 7.4: Strahler slope ratio with time: simulations CR2-6, CR7-1, CR7-2, CR7-3
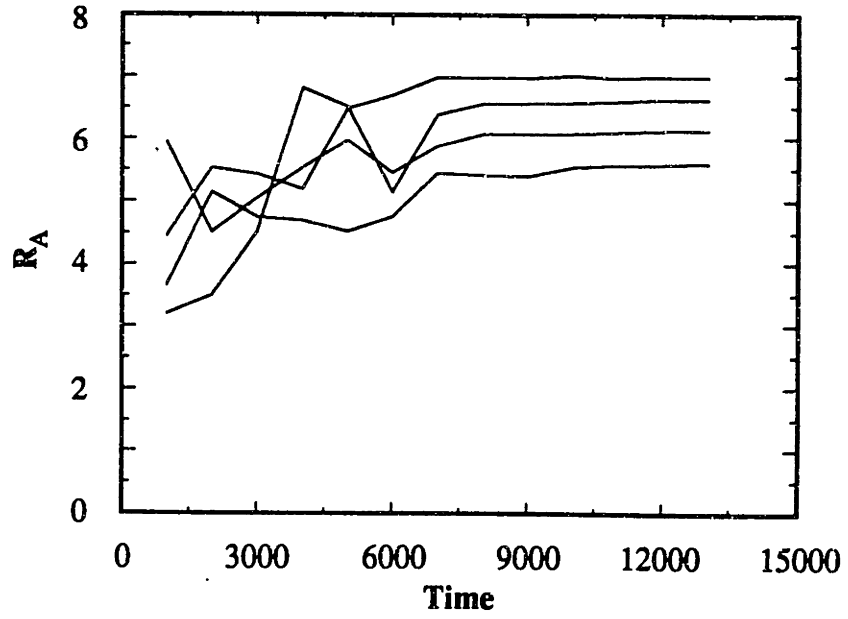
Order 1-2
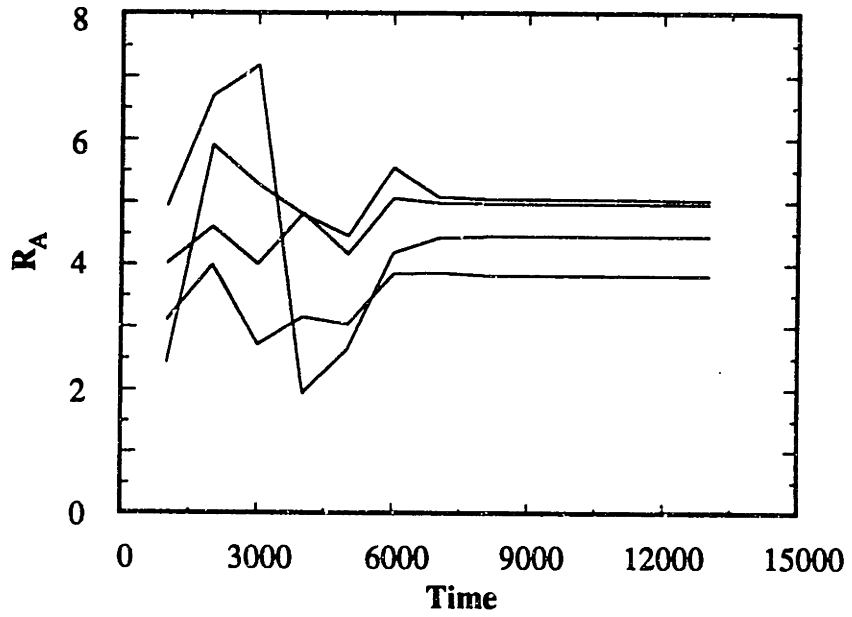
Order 2-3

Figure 7.5: Strahler area ratio with time: simulations CR2-6, CR7-1, CR7-2, CR7-3

This growth hypothesis ignores growth site screening. Screening will effectively bias the growth sites so that growth is more likely at the extremities of the network than the root. It is believed that this systematic process will not invalidate the argument for random allometric growth of the network, it will simply condition the link sampling process during allometric growth.

Parker (1977) simulated the growth of networks using the rainfall erosion experimental facility at Colorado State University. He mapped the planar form of the networks that were formed with time. The bifurcation ratios with time for these networks are given in Figure 7.6. There were two networks; 1 and 2. These networks corresponded to two separate experiments starting from uniform unchannelized initial conditions. Since the runoff/unit time for the applied rainfall was constant with time Parker plotted all his data using cumulative volume of runoff as a surrogate of time; for consistency we do too. For Network 1 this data shows the same degree of scatter in the bifurcation ratio as the simulated results and there does seem to be a reduction of the scatter with time, as the network's magnitude increases. For Network 2 there appears to be a decline in the scatter of bifurcation ratio up to volume 20000. After this time the network was "refreshed" a number of times by increasing the catchment relief and allowing erosion to continue anew until the network stabilized to a new, higher, drainage density. This new network development may account for the increased scatter observed after cumulative runoff volume 20000 in Network 2.

For the simulation results the Strahler ratios for slope, area and length also varied considerably with time, although scatter appears to diminish with time. The major fluctuations apparent in these figures correspond to times when the Strahler order of the network increases. As the network grows a link that may have been part of an order (i) stream becomes part of an order (i+1) stream, even though that link's characteristics have changed little in the meantime. Through the allometric growth process whole sections of the catchment change order, even though the link
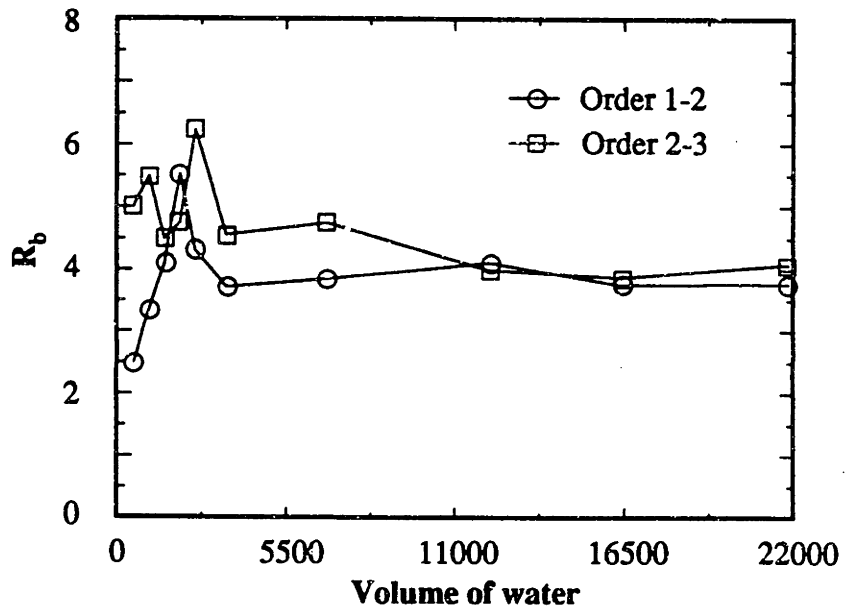
(a) Experiment 1
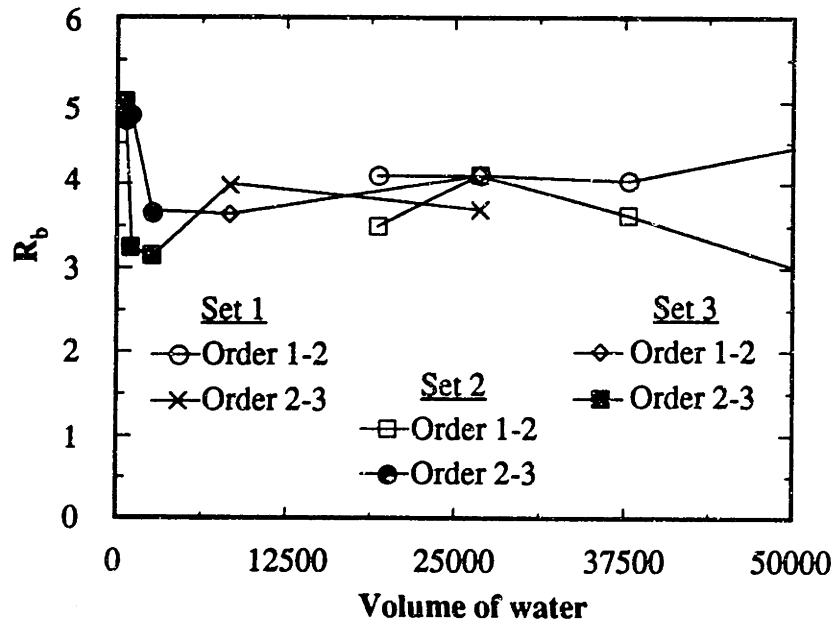
(b) Experiment 2

Figure 7.6: Strahler bifurcation ratio with time:
Parker (1977) experimental data

characteristics have changed only slightly. This is because the order of the network is very dependent upon the topological arrangement of the links. In a similar sense to the case of bifurcation ratio, two catchments that vary only in the position of a single link may have very different Strahler statistics. Thus there is a great deal of scatter inherent in the Strahler statistics.

The bifurcation, length, slope, and area ratios, with time averaged over the entire catchment, are shown in Figures 7.7 to 7.10. Compared to the Strahler ratios calculated for individual orders, the ratios exhibit less scatter, reflecting the larger sample size. However, like the Strahler ratios for the individual orders, there is still substantial scatter in the statistics averaged over the entire catchment both during and after active network development.

The trends with time for the Tokunaga statistics, K and ϵ, based on Strahler ordering, are presented in Figures 7.11 and 7.12. These statistics also show considerable scatter with time. This scatter arises for the same reasons as the scatter in the bifurcation ratio, both statistics are based on the Strahler ordering scheme so that random allometric growth is again important. The Tokunaga parameters provide an alternative way to measure stream numbers. Figure 7.13 shows the trend with time of the bifurcation ratio calculated from Tokunaga's K and ϵ for an infinite catchment ($R_b(\infty)$ in Equation 2.6). The coefficient of variation in R_b from the Tokunaga statistics is similar to that of the bifurcation ratio calculated directly from the stream numbers (Figures 7.2 and 7.7) so that Tokunaga statistics do not provide a more reliable means of assessing the mean stream populations. While the Tokunaga number theory has conceptual advantages over the bifurcation ratio in the mean sense (Section 2.2.2.4) the statistics appear to be no more reliable estimators of the population than is the bifurcation ratio.

Thus there are difficulties with the Strahler and Tokunaga ordering statistics
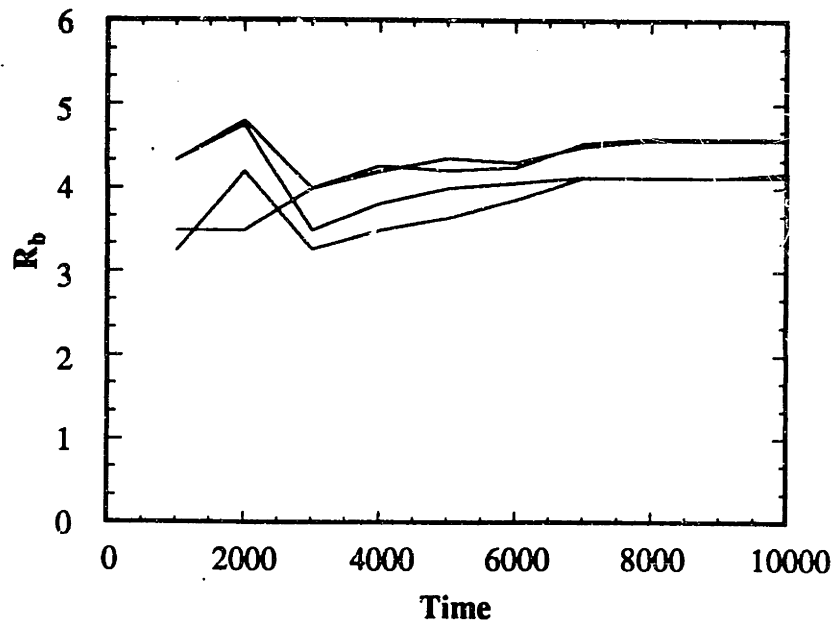
Figure 7.7: Average bifurcation ratio:
simulations CR2-6, CR7-1,CR7-2,CR7-3
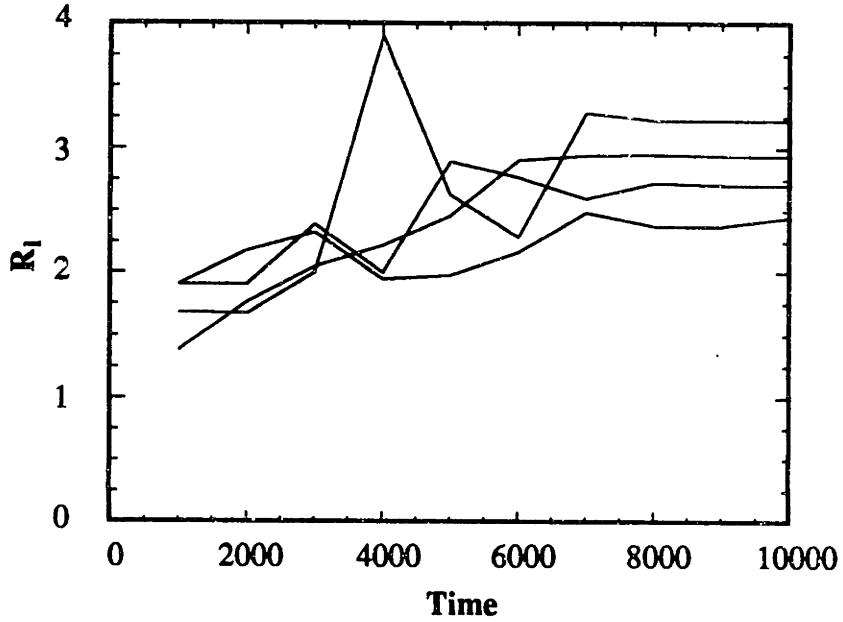
Figure 7.8: Average length ratio:
simulations CR2-6, CR7-1,CR7-2,CR7-3
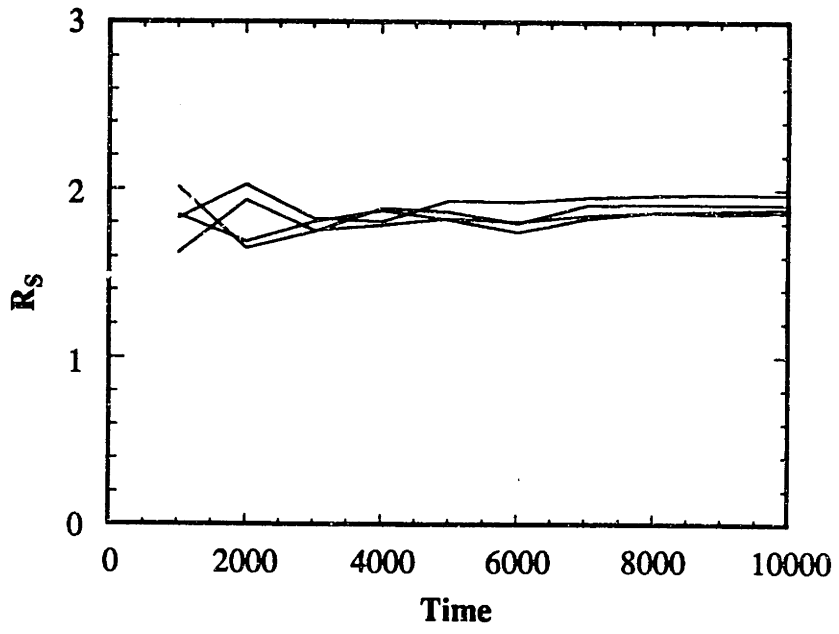
Figure 7.9: Average slope ratio:
simulations CR2-6, CR7-1,CR7-2,CR7-3
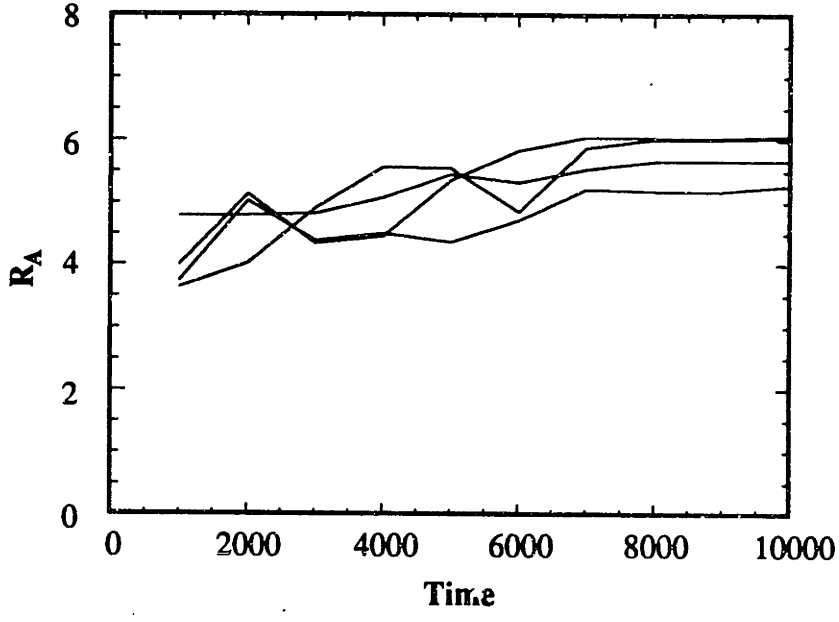
Figure 7.10: Average area ratio:
simulations CR2-6, CR7-1,CR7-2,CR7-3
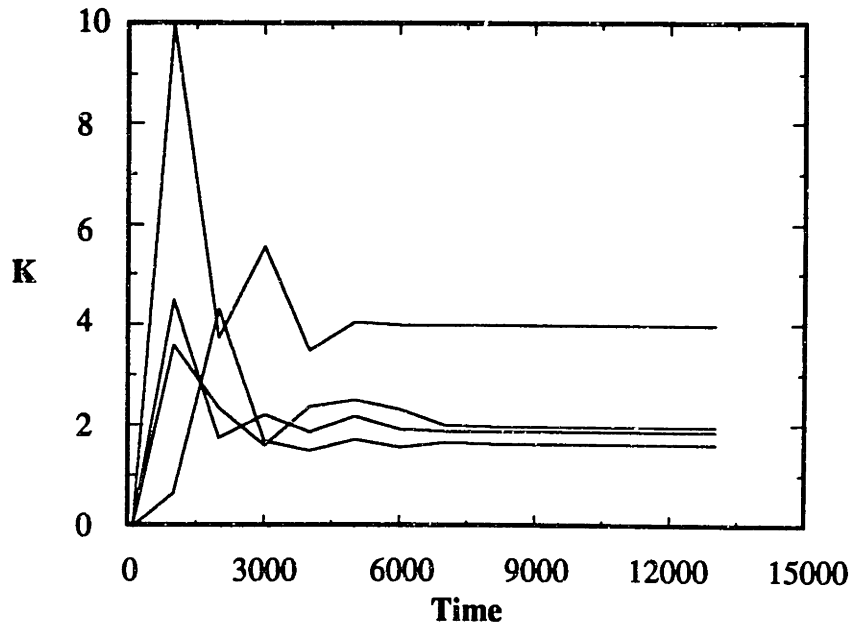
Figure 7.11: Tokunaga K with time:
simulations CR2-6, CR7-1, CR7-2, CR7-3
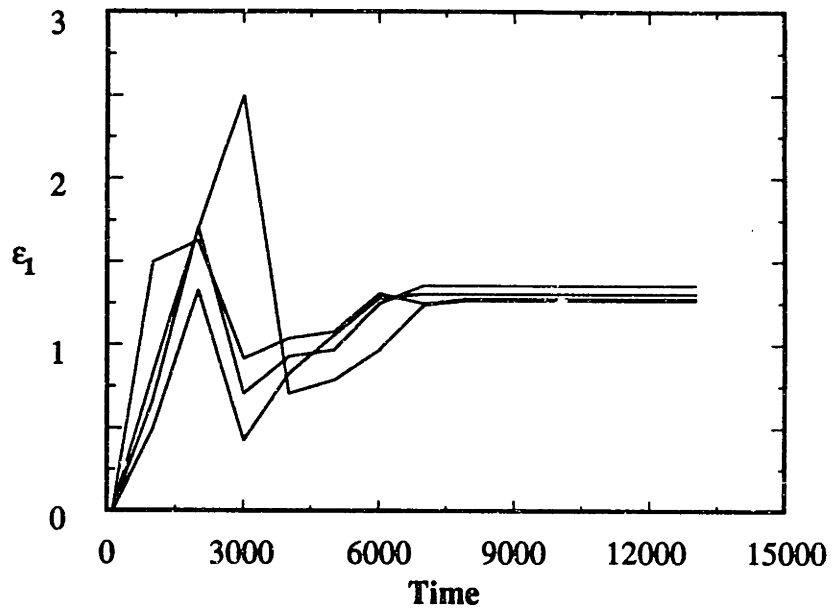
Figure 7.12: Tokunaga ε_1 with time:
simulations CR2-6, CR7-1, CR7-2, CR7-3
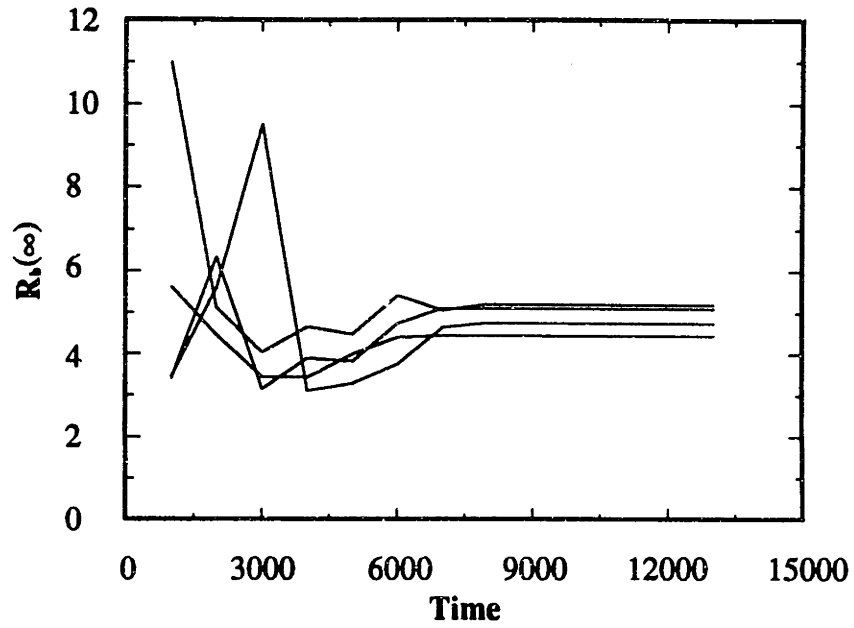
Figure 7.13: $R_b(\infty)$ calculated from Tokunaga's
K and ε with time:
simulations CR2-6, CR7-1, CR7-2, CR7-3

that become apparent when temporal aspects of catchment form are considered. There is considerable inherent variability in the statistics, and this scatter follows directly from the allometric growth of the network. The final channel network reflects this topological randomness. This growth argument is useful for understanding why field catchments exhibit considerable variability in the ordering statistics. The random allometric growth hypothesis also suggests mechanisms whereby the topologically random networks of Shreve (1966) may develop in the field environment.

When a network grows, the growth occurs in two ways:

1.   exterior links extending into the catchment, and

2.   lateral branching from either exterior or interior links.

These two growth mechanisms both induce randomness in the lengths of the links. Figure 7.14 shows the scatter in the mean link lengths for the developing networks. Exterior links appear no more variable in length than interior links and the mean length of both types grows with time.

Figure 7.15 shows the increase of magnitude with time. Figure 7.16 shows the increase in drainage density with time. The variability of drainage density is very small and the variability of magnitude with time is larger. For a fixed catchment area drainage density is a function of only magnitude and mean link length so that variations in mean link length and magnitude are negatively correlated (Equations 6.32 and 6.36). Thus some of the variability in the mean link length can be explained by the randomness of the allometric growth process. It cannot result from any innate variability in the physical length of the network because the drainage density variation is quite small.

Figure 7.17 shows the width functions for the final networks. Despite the scatter of the width functions, it is possible to hypothesize a mean width function that generally increases with distance from the outlet then drops dramatically after it reaches its peak value. The actual shape shown in Figure 7.17 is not believed to be
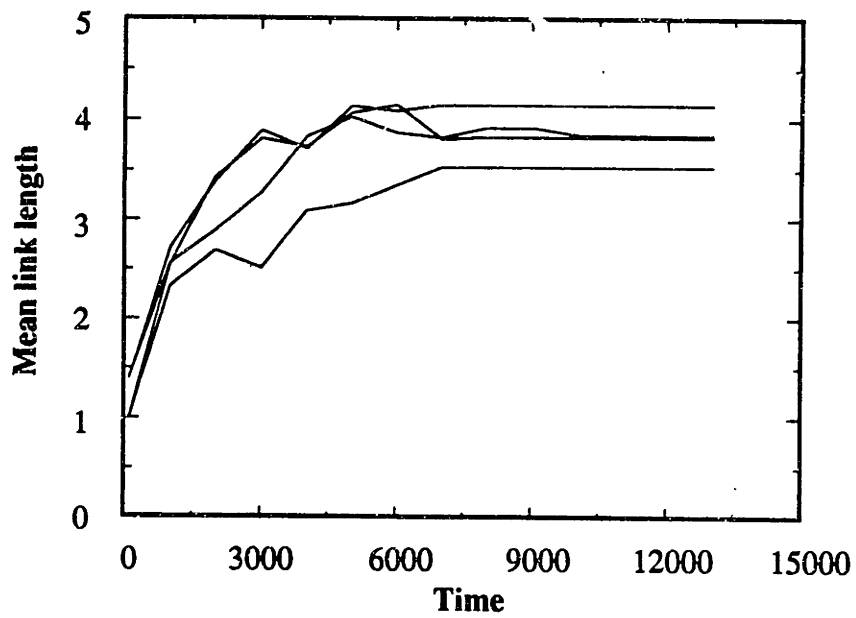
Figure 7.14: Mean link length with time:
simulations CR2-6, CR7-1, CR7-2, CR7 3
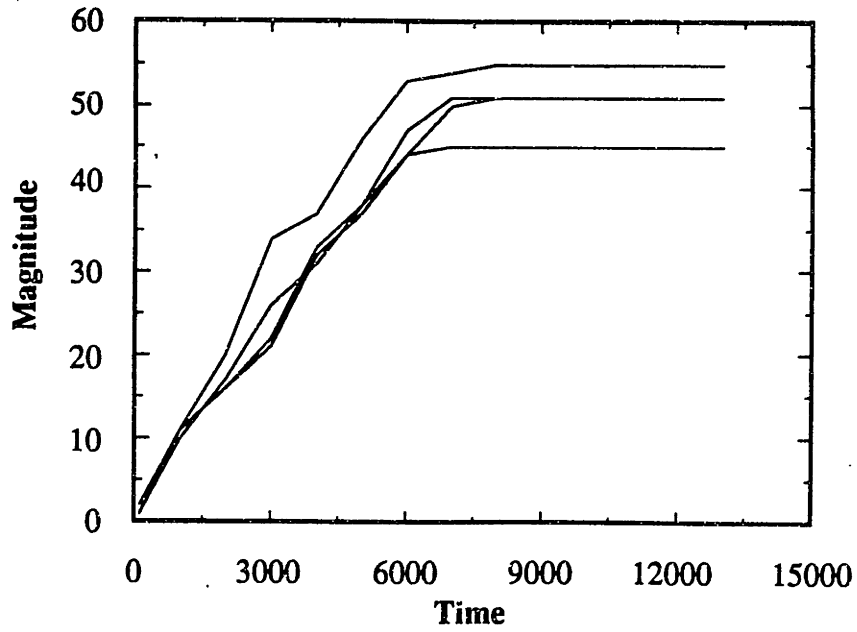
Figure 7.15: Magnitude with time: simulations CR2-6, CR7-1, CR7-2, CR7-3
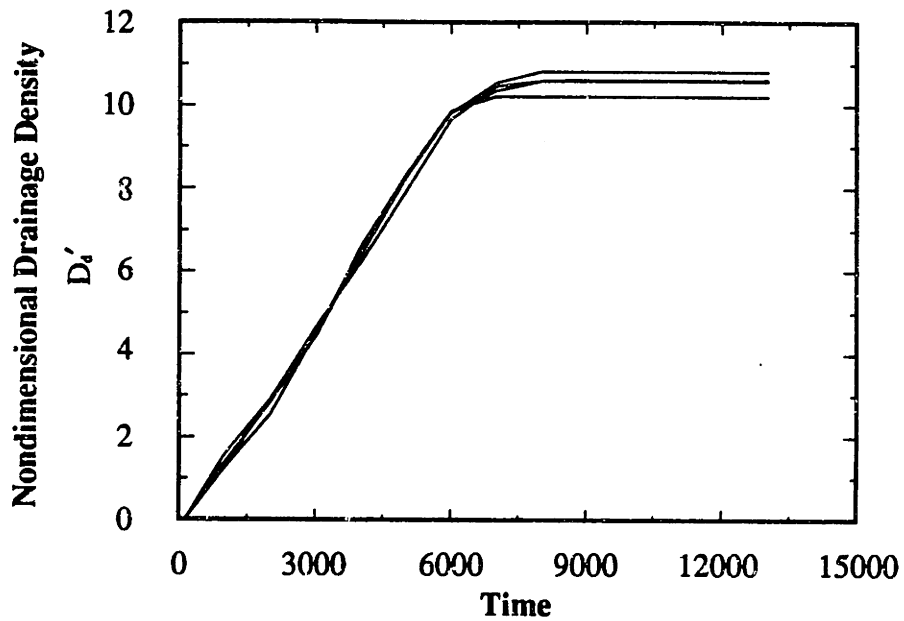
Figure 7.16: Drainage density with time: simulations CR2-6, CR7-1, CR7-2, CR7-3
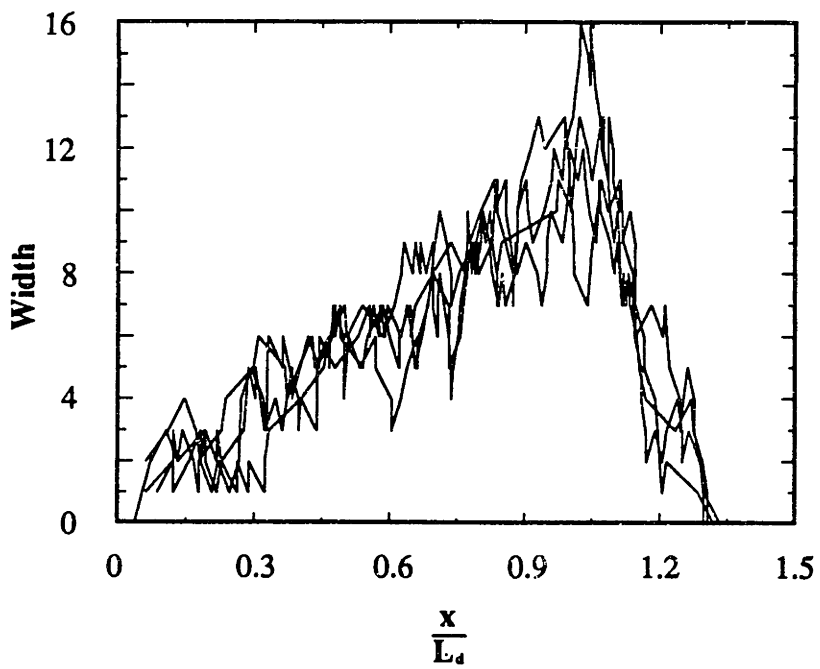
Figure 7.17: Final width functions: simulations CR2-6, CR7-1, CR7-2, CR7-3

generally applicable, but is believed to reflect the shape of the catchment (i.e., a square in this case).

Finally 7.18 shows that the variability of the hypsometric integral is quite small, even with quite variable planar network characteristics. Figure 7.19 shows the sediment yield of the total catchment through time. It may be noted that a change in area under the hypsometric curve (the hypsometric integral) reflects the loss of sediment within the catchment so that the sediment yield and the hypsometric curve/integral are related.

The conclusion drawn from this section is that there is considerable inherent variability in many of the standard statistics for measuring the catchment and its network, particularly those describing topological properties. It is important to note that the catchment simulations in this section satisfy the physical similarity conditions of Chapter 6. All differences between the simulations are due to stochastic effects. The idea of proving the hypothesis that the Strahler ratios are constant seems both hopeless and pointless. It is hopeless because the inherent variability of the network's characteristics means that networks must be impractically large to determine these statistics with the required degree of accuracy. It is pointless because if the Strahler ratios are important in the determination of some derived characteristic (e.g., the GIUH of Section 2.2.1), then the inherent variability of the statistics for realistically sized catchments necessitates the determination of these statistics for each catchment individually rather than using some mean regional values. Of what use is knowing that the underlying Strahler ratio is constant if the sample statistics of networks are not close to the underlying value for practical size problems?

On the other hand there are a number of characteristics like the magnitude, drainage density, and the hypsometric curve that exhibit little variability. This is because they are physically based, rather than topologically based, characteristics.
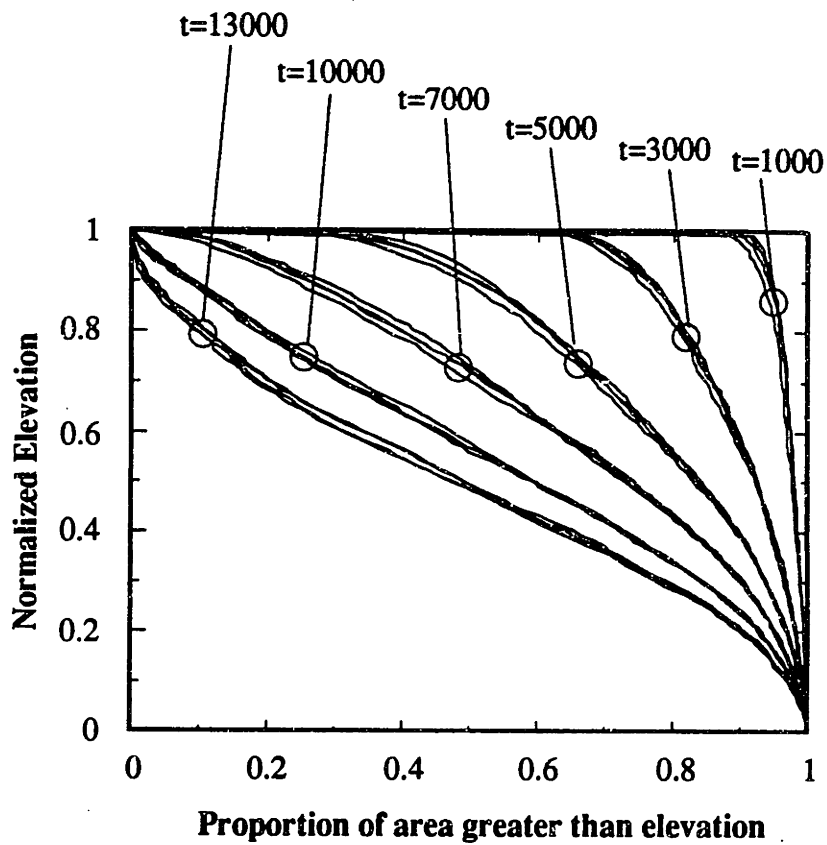
Figure 7.18: Hypsometric curves with time:
simulations CR2-6, CR7-1, CR7-2, CR7-3
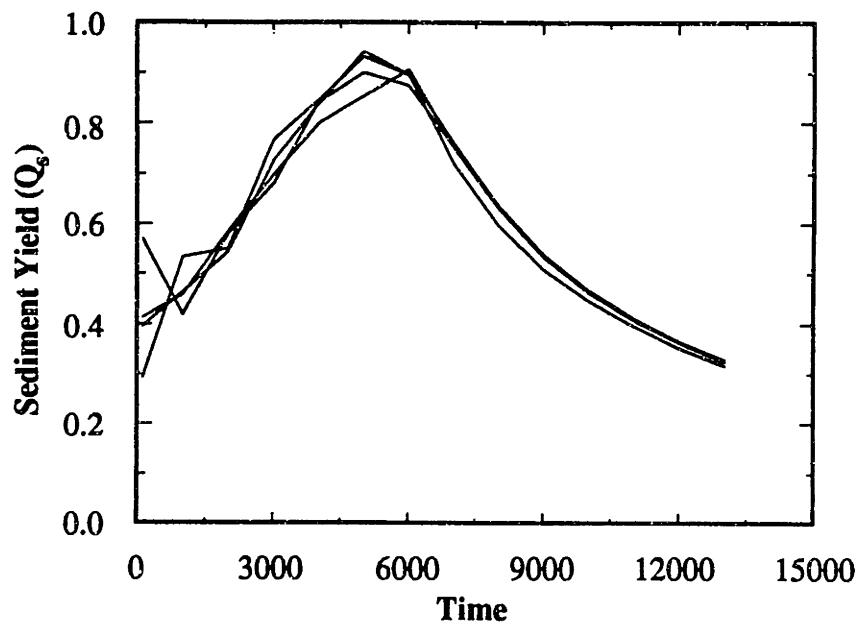
Figure 7.19: Sediment yield from the catchment:
simulations CR2-6, CR7-1, CR7-2, CR7-3

They are also, in some sense, integrated over the catchment. Small variations in the catchments do not result in large variations in the statistics.

Whether derived characteristics of the catchment, such as the instantaneous unit hydrograph (IUH), are highly variable or not, dependent upon the channel network variability, is an open question. Gregory and Walling (1968) and Rodriguez–Iturbe and Valdes (1979) observed an increase in the peak runoff of the IUH with increased drainage density and decreased link length, both of which are low variability properties of the catchment (Figure 7.16). Rodriguez–Iturbe and Valdes (1979), however, hypothesized a dependency of the IUH on Strahler statistics, which are high variability properties of the catchment (Figures 7.2 to 7.10). If the IUH is totally dependent upon network topology then the shape of the IUH might be expected to be highly variable from catchment to catchment, even if the catchments satisfy the physical similarity conditions of Chapter 6. If the IUH is dependent upon low variability properties such as drainage density or the hypsometric curve then the IUH would not be expected to be highly variable from catchment to catchment. The relative importance of high and low variability properties in determining the shape of the IUH is still an open question. The answer to that question will determine how variable is the shape of the IUH from catchment to catchment when the physical similarity conditions of Chapter 6 are satisfied.

7.3 Drainage Density and Its Variation with Time

The main characteristic provided by the catchment simulation model is the channel network. The channel network is, as previously described, the central feature of the model. The catchment characteristics are generated through the long term preferential erosion within the channel network, relative to the hillslopes. Since the channel network is central to the model, then it is important to understand how it varies with the nondimensional parameters of Chapter 6.

This section will look at three important features:

1. parameterization of where the channel starts in terms of the hillslope length scales and other physical characteristics;

2. parameterization of the drainage density in terms of the catchment length scales and other physical characteristics; and

3. parameterization of the rate at which the channel network grows in terms of catchment length scales and other physical characteristics.

The previous section indicates that the inherent variability of drainage density is low. Thus systematic trends with physical parameters can be easily distinguished from stochastic components.

7.3.1 Drainage Density and Hillslope Length Scales

This section aims to show that the fundamental invariant property of the landscape is neither the drainage density nor the mean hillslope length but is the activator number based on hillslope lengths and drops, runoff rates and erodability, TA_h. For a fixed runoff rate, catchment erodability and hillslope slopes, the drainage density is fixed, but variation in runoff rates, erodability or slopes changes the drainage density and thus the mean hillslope lengths.

Consider first the mechanism that controls channel growth. For a channel to grow, the activator threshold must be exceeded; for a channel to stop growing, the activator around the branch head must fall below that threshold. For a node that is on the verge of differentiating from hillslope to channel the activator term in the channelization equation is

$$\frac{\text{activator}}{\text{activator threshold}} = c_1 \beta_5 Q^{m_5} S^{n_5} = 1 \tag{7.1}$$

where $\frac{1}{c_1}$ = the activator threshold.

Thus the channel will stop growing when the contributing area or runoff rate decreases (causing Q to decrease), the local slope at the branch head decreases, the activator threshold $1/c_1$ increases, or the activator coefficient β_5 decreases. If the hillslopes above the channel heads have a consistent nondimensional form then

$$z' = f(x') \tag{7.2}$$

where z' $= z/R_h$ = nondimensional elevation above the channel

R_h = mean hillslope relief

x' $= x/L_x$ = nondimensional distance from the channel

L_x = mean hillslope length

The nondimensionalizations of Equations (6.12) and (6.13) can then be used to give

$$c_1 \beta_5 Q^{m_5} S^{n_5} = \left[\frac{L_z^{n_5} L_x^{2m_5 - n_5} L_R^{m_5} c_1 \beta_5}{T_R^{m_5}} \right] a' = TA_h \, a' \tag{7.3}$$

where TA_h = hillslope activator number, where the length scales, L_x and L_z,
are those for the hillslope above the channel head

a' = nondimensional activator at the channel head dependent on the nondimensional hillslope, $z' = f(x')$ and $a' = \beta_5 Q'^{m_5} S'^{n_5}$.

If the nondimensional hillslopes around the channel head are the same for all channel heads then a' does not vary systematically. Equations (7.1) and (7.3) then yield

$$TA_h = C \tag{7.4}$$

where C = constant

From the definition of TA_h in Equation (7.3) and noting that typically $(2m_5-n_5) > 0$ it can be seen that source areas $(\equiv L_x^2)$ decrease with increased slope $(\equiv \frac{L_z}{L_x})$, increased hillslope drop $(\equiv L_z)$, increased runoff $(\equiv L_R)$ and increased activator threshold $(\equiv c_1)$. A decrease in source area is equivalent to an increase in drainage density

This result is consistent with theoretical results obtained by Kirkby (1986), for catchments dominated by Hortonian style runoff, though not for catchments dominated by subsurface saturation style runoff. This result is also consistent with field data presented by Dietrich, et al. (1986) and Montgomery and Dietrich (1988) (see Chapter 8 for a detailed discussion).

This provides a useful way of parameterizing the hillslopes around the channel head. The inputs to TA_h are simple, hillslope horizontal and vertical scales (e.g. mean hillslope length and mean hillslope drop), runoff length and time scales (e.g. mm/hr), and the coefficients c_1 and β_5. In principle Equation (7.4) should be true irrespective of the physical parameters since the physical parameters are included in the nondimensional number TA_h. The only proviso on Equation (7.4) is that the nondimensional hillslope, and thus a', does not change. A priori only a change in the hillslope sediment transport equation exponents (i.e. m_1 and n_1) will change hillslope form and a' appreciably (Kirkby, 1971 or Figure 2.10).

To verify the validity of Equation (7.4), TA_h was evaluated for the simulations using as the hillslope length scales

L_x = (mean Strahler first order area)$^{1/2}$

L_z = mean drop of hillslopes in the Strahler first order areas

266

The results are presented in Table 7.1. The variation in TA_h is $\pm 20\%$. This is good considering that the length scales used to evaluate TA_h are approximations to the actual length scales that govern channel growth. A similar result was found for experimental data by Dietrich, et al. (1986). The L_x, based on Strahler 1st order area, was satisfactory because the ratio of contributing source area at the channel head to the Strahler 1st order area was remarkably constant across all the physical parameters, except for a trend with the ratio m_5/n_5 (see Section 7.3). The chosen L_z, based on mean hillslope drop, was satisfactory because it is representative of the slope at the channel head; a higher hillslope drop was accompanied by a higher hillslope slope at the channel head.

The fluctuations in the observed value of TA_h are considered to arise from three sources:

1. Differences in the rates of growth of channel networks. Figure 2.10 showed that the equilibrium hillslope profile for hillslopes is concave up for the physical parameters of the simulations. During the growth phase, however, when erosion in the channel proceeds much faster than in the hillslope, the hillslope profile has a tendency to be concave downward, particularly near the channel. The faster the network grows, the more pronounced the effect. Thus the characteristic hillslope profile is dependent upon the timescales of growth of the network and Equation (7.4) is invalidated at times before dynamic equilibrium. This effect was exhibited by runs CR5–2, where channels developed faster, and CR3–4 and CR3–5 where overland erosion proceeded more slowly.

2. Variation in the relative sizes of the channel source area and the Strahler first order area. The source area was a greater percentage of the first order area for an activator function with high slope dependence (e.g.,

TABLE 7.1

TA_h for the Hillslopes

Run	TA_h	Run	TA_h
CR1–3	1.15	CR4–7	1.99
CR2–3	1.22	CR4–9	1.20
CR3–3	1.21	CR5–1	1.20
CR3–4	0.98	CR5–2	1.15
CR3–5	1.11	CR9–6	0.90
CR4–2	1.21	CR9–8	0.88
CR4–5	1.21	CR9–10	0.86
CR4–6	1.22	CR9–11	0.88

(source area)/(Strahler 1st order area) is approximately 78% for CR9 runs, $m_5/n_5 = 0.5$) than for low slope dependence (e.g., approximately 50% for CR1 and CR2 runs, $m_5/n_5 = 1.33$). Thus the use of the Strahler first order areas to define L_x only approximately reflects the governing horizontal length scale, L_x, based on source areas. This dependence accounts for almost all the scatter in TA_h for the CR9 runs in Table 7.1.

3. Since the local slope at the channel head, which controls the activator, and the mean slope, used to define L_z, are different, then TA_h will be dependent on the value of n_5 in the activator equation. None of the simulations in Table 7.1 exhibit this effect since m_5 was varied and n_5 kept constant to vary the value of m_5/n_5.

In summary, it is very easy to determine the point at which channels stop growing in terms of the local hillslope slopes at the channel heads and source areas. This follows directly from the form of the activator and differentiation equations. It is that point on the hillslope where the activator falls below the activator threshold. It is also possible to parameterize this point in terms of the mean properties that are typically measured for hillslope and exterior links, e.g., Strahler first order area, exterior link area, hillslope drops.

An important observation which goes to the heart of the scales of dissection of the landscape will be made here. Equation (7.4) defines the relationship between the fundamental (i.e hillslope) scales of the landscape. Traditionally workers have considered the mean hillslope length, L_x, as the fundamental scale of dissection. However, Equation (7.4), which governs where channels stop growing, defines a relationship between the hillslope horizontal length scale (e.g. mean hillslope length), L_x, and the hillslope vertical length scale (e.g. the mean hillslope drop), L_z. Another independent relationship is required to fix either L_x or L_z uniquely. This is important

since it indicates that the drainage density (the inverse of the mean hillslope length) is variable and is dependent upon the hillslopes; the steeper the hillslopes, the shorter the hillslopes and the higher the drainage density. Strahler (1958) observed that " ... areas of lower drainage density commonly have ... gentler valley side slopes [compared with areas with higher drainage densities] " (p 295).

Solving this problem requires looking at catchment horizontal and vertical length scales, L_d and L_e. A catchment that is steeper overall (i.e. L_d/L_e high) will also have steeper hillslopes. It will be shown later that the mean hillslope and catchment slopes are related through the sediment transport continuity of the model. But catchment scales are determined by such effects as tectonic uplift so that the higher the tectonic uplift, the steeper the catchment, the steeper the hillslope and the higher the drainage density. If the catchment area is less and the tectonic uplift is the same then similar trends are postulated.

It is through a relationship between the catchment length scales and the drainage density that we will be able to determine the hillslope length scales uniquely. Thus it will be shown that the catchment scales and the hillslope length scales are intimately connected. The next section develops the required relationship while the final connection between the scales will be presented in Section 7.5.

7.3.2 Drainage Density and Catchment Length Scales

We now discuss the issue of predicting the drainage density based on catchment properties. The results reported above provide a means of predicting the drainage density, but only if the hillslope drops or slopes are known. We will now discuss the case where only the catchment horizontal and vertical length scale are known (e.g., catchment area, catchment relief or mean catchment slope).

There are two ways of determining the drainage density from the catchment length scales. The first approach is by developing the relationship from first principles

A speculative relationship between the hillslope length scale (defined as mean hillslope length) and catchment length scale (defined as the square root of the catchment area) that was dependent on network magnitude or order was developed in Chapter 6 (Equation 6.37) In a similar fashion a speculative relationship between the vertical length scales of the hillslopes and the catchment through the order of the catchment was also developed (Equation 6.53). These derivations suggest an approach that eliminates magnitude between Equations (6.37) and (6.53) together with Equation (7.3) to derive an expression for drainage density. That is, Equation (6.37) and (6.53) give relationships between the catchment and hillslope length scales of the form

$$L_d = f_1(L_x, m)$$
$$L_e = f_2(L_z, L_x, m) \tag{7.5}$$

while Equation (7.3) gives a relationship between the hillslope length scales of the form

$$L_z = f_3(L_x) \tag{7.6}$$

Drainage density is just a function of hillslope length scale so that

$$L_x = f_4(D_d) \tag{7.7}$$

Substitution of Equations (7.6) and (7.7) into Equation (7.5) yields

$$L_d = f_1(f_4(D_d), m) \tag{7.8a}$$
$$L_e = f_2(f_4(D_d), m) \tag{7.8b}$$

In principle, the magnitude m can be eliminated between the equations (7.8a)

and (7.8b) yielding a relationship for drainage density D_d dependent only on the catchment horizontal and vertical length scales. Analytical difficulties and various approximations inherent in Equations (6.37) and (6.53) that lead to Equations (7.5) preclude this approach.

The second, and more realistic, approach is to perform a multiple regression on the nondimensional drainage density and the nondimensional parameters derived in Chapter 6 using the catchment simulations. A number of different relationships were fitted, but the best, both conceptually and in the mean square sense, was

$$D_d' = \frac{1}{28} \left[\frac{TA_c - TA_h}{TA_h} \right]^{0.53} O_t^{-0.06} \left[2.5 + \frac{n_5}{m_5} \right]^{4.12} \tag{7.9}$$

where D_d' = nondimensional drainage density

$\quad\quad\quad\quad = D_d L_d$

$\quad\quad L_d$ = horizontal length scale of the catchment

$\quad\quad\quad\quad = (\text{catchment area})^{1/2}$

$\quad\quad L_e$ = vertical length scale of the catchment = catchment relief

$\quad\quad TA_h$ = TA defined using the hillslope length scales, L_x and L_z, as defined in Equation (7.3)

$\quad\quad TA_c$ = TA defined using the catchment horizontal and vertical length scales, L_d and L_e.

The explanative power of this equation is quite good, the fit explaining 90% of the variance of the fitted points. Figure 7.20 shows a plot of observed versus estimated D_d' for the multiple correlation above. The observed values plotted are those drainage densities calculated from the simulations, while the estimated values are calculated from Equation (7.9). Some verification points, not used in fitting the regression, are also shown in the figure. The fit of these verification points is quite satisfactory. The
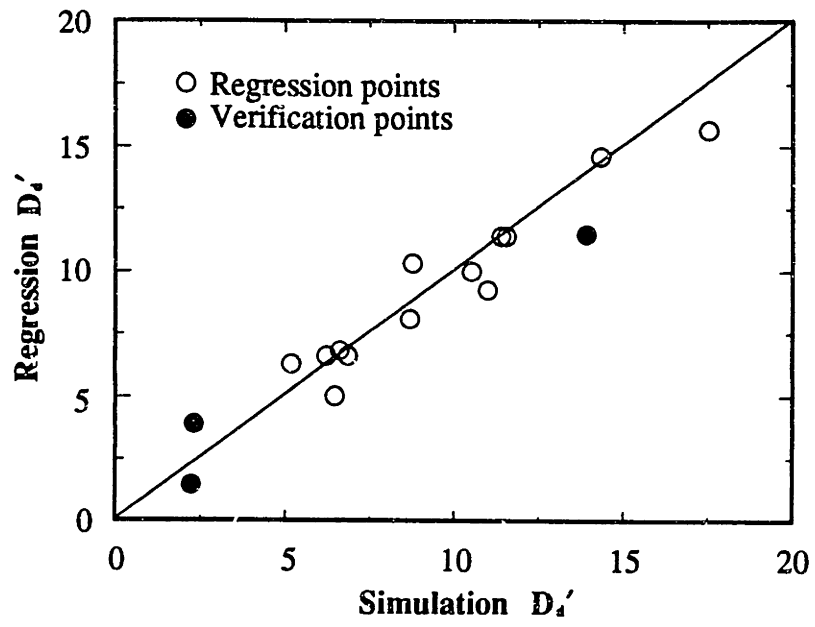
Figure 7.20: Simulation versus regression
drainage density .

simulations used for calibration and verification are given in Appendix D. Figure 7.21 shows 4 simulated networks that vary only in their value of TA_c, slowing the variation in drainage density.

The form of the multiple correlation is consistent with an analytical expression that can be derived for nondimensional drainage density when the activator function is a function solely of area (as in Tarboton, et al., 1988). Let the activator function be of the form

$$a = \beta_5 A$$

where A = area draining to the channel.

This expression for activator is just Equations (5.2a) and (6.5) with the parameters $m_5 = 1$, $n_5 = 0$ and R = 1. This constitutes one extreme of the spectrum of different, and valid, (though not necessarily realistic) activator functions. It would be an advantage if the drainage density expression was consistent at this extreme value for m_5/n_5 since Equation (7.9) does claim to allow for variation in m_5/n_5. Channel growth is initiated when the area draining to a node is higher than the so—called support area or source area (A_s). Then the activator number for the hillslope is, noting the change in the definition of the horizontal length scale compared to Equation (7.1) for convenience

$$TA_h = \beta_5 \, c_1 \, L_s^2 \tag{7.10}$$

where L_s = length scale of the hillslope = $(A_s)^{1/2}$.

Using Equation (2.12) the nondimensional drainage density can be expressed as
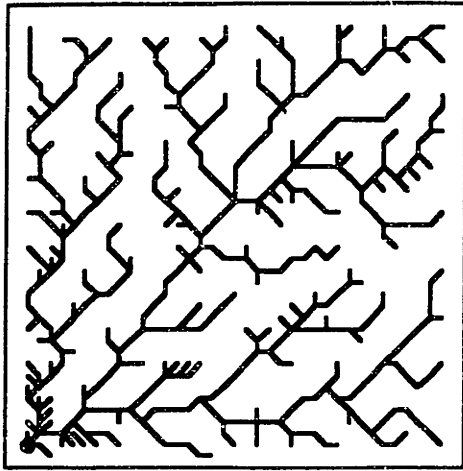
$TA_c = 15.77 \times 10^{-3}$
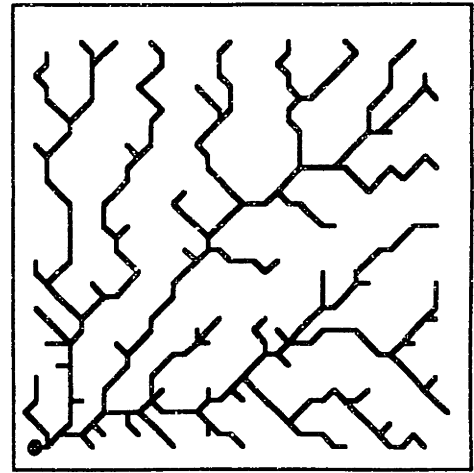
$TA_c = 12.61 \times 10^{-3}$
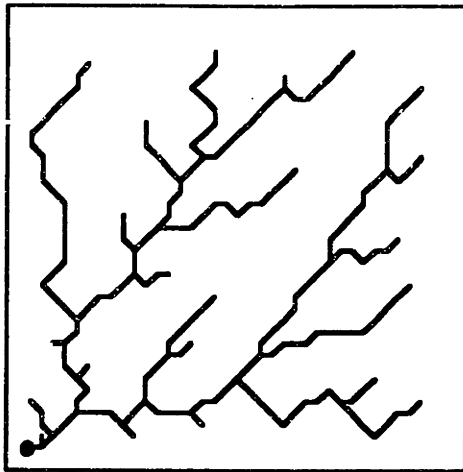
$TA_c = 9.46 \times 10^{-3}$
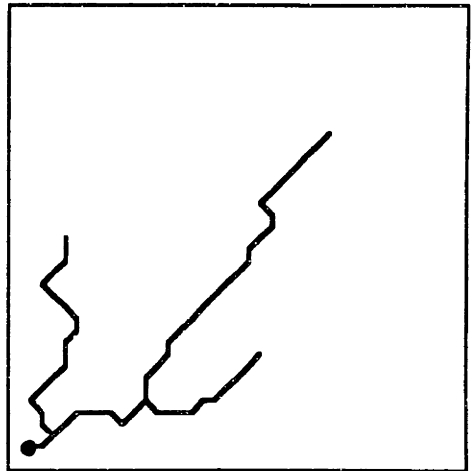
$TA_c = 5.68 \times 10^{-3}$

● Catchment outlet

Figure 7.21: Drainage density variation with TA_c

$$D_d' = L_d\, D_d = \frac{L_d}{2L_x} = \frac{L_d(\beta_5 c_1)^{1/2}}{2\eta_1 (TA_h)^{1/2}} \qquad (7.11)$$

where L_d = length scale of the catchment

                        = (catchment area)$^{1/2}$

           L_x = mean hillslope length = $\eta_1\, L_s$

The catchment scale activator number is

$$TA_c = \beta_5\, c_1\, L_d^2$$

Substituting this result into Equation (7.11) for L_d yields the nondimensional drainage density relationship

$$D_d' = \frac{1}{2\eta_1} \left[\frac{TA_c}{TA_h}\right]^{1/2} \qquad (7.12)$$

The trend of this result with TA_c and TA_h is consistent with Equation (7.9) if $TA_c \gg TA_h$ (i.e., catchment large) which is also required for the definition of drainage density in Equation (2.12) to be correct.

Return now to Equation (7.9). Drainage density is positively correlated with $(TA_c - TA_h)$. When this term is zero, (i.e., $TA_c = TA_h$), the drainage density is also zero. This occurs when either the size of the catchment, L_d, or the relief of the catchment, L_e, is insufficient to trigger channelization and the simulated catchment is totally hillslope. The TA_h term in the denominator of this term provides consistency with Equation (7.12) and slightly reduces the variability in prediction of the regression.

Drainage density is negatively correlated with the ratio of overland sediment

transport to channel sediment transport, O_t. This results from sediment transport continuity in the hillslope and the channel. Sediment transport continuity at the channel head requires that the slope of the hillslopes must be higher than the slope in the channel since at the channel head and for a one-dimensional hillslope

$$\beta_1 \, Q_c^{m_1} \, S_c^{n_1} = O_t \, \beta_1 \, Q_h^{m_1} \, S_h^{n_1} \tag{7.13}$$

so that at the channel head, where $Q_c = Q_h$

$$\frac{S_c}{S_h} = O_t^{\frac{1}{n_1}} \tag{7.14}$$

This result is also true for a two-dimensional hillslope-channel interface (see Section 8.3). As O_t decreases, and hillslope erosion decreases relative to the channel, S_c/S_h decreases. But the total relief of the catchment is unchanged so that S_h increases a little and S_c decreases a little to compensate. Since the hillslope slope has increased to maintain sediment transport continuity, then to maintain the activator balance at the channel head (i.e. TA_h = constant), the source area must decrease. Thus the channel extends into the hillslope region and the drainage density is increased.

7.3.3 The Rate of Increase of Drainage Density

The final property of the channel network to be dealt in this section is the rate of increase of drainage density, or the rate of growth of the channel network. Figure 7.12 shows a typical growth history for a network. It is possible to identify four stages in

the increase of drainage density with time; these are shown in Figure 7.22.

1. A short transition where slopes and drainage directions are beginning to develop around the notch at the catchment outlet.

2. The main growth stage where drainage density grows almost linearly with time until close to the ultimate drainage density.

3. A transition stage where the growth rate decreases as a result of reductions in the contributing area to the channel heads. This reduction could result from applied boundary conditions or competition from adjacent networks.

4. The ultimate drainage density, where growth stops.

The drainage density curve is dominated by the second, linear growth, stage. Figures 7.22 and 7.16 can be compared with the drainage density growth postulated for Howard's headward growth model in Figure 2.6. The curves are quite similar.

The linear growth stage for the network results from two competing processes. The first of these two processes is that as the network grows the magnitude of the network and the number of potential channel growth sites also grows. This produces the concave up portion of the drainage density curve in Figure 7.22 at early times. However, the number of growth sites is not proportional to magnitude since many channel heads do not grow. This results from the screening effect of the surrounding networks near the network root. Growth sites are concentrated in the outer parts of the network. This screening effect is qualitatively similar to that observed in other network growth models, which were described in Chapters 3 and 4. The net effect is that there is modest increase in the number of growth sites. The increase in growth sites produces an upward curvature in the drainage density with time.

The second of the two processes is that as the network extends the relief of hillslopes around the channels diminishes at an ever increasing rate. Since the stream profile   is
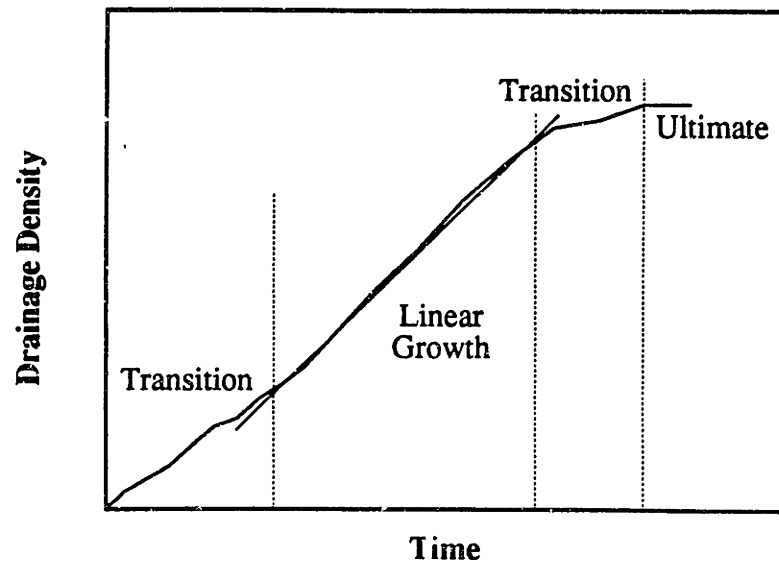
Figure 7.22: Schematic of stages of drainage density growth with time
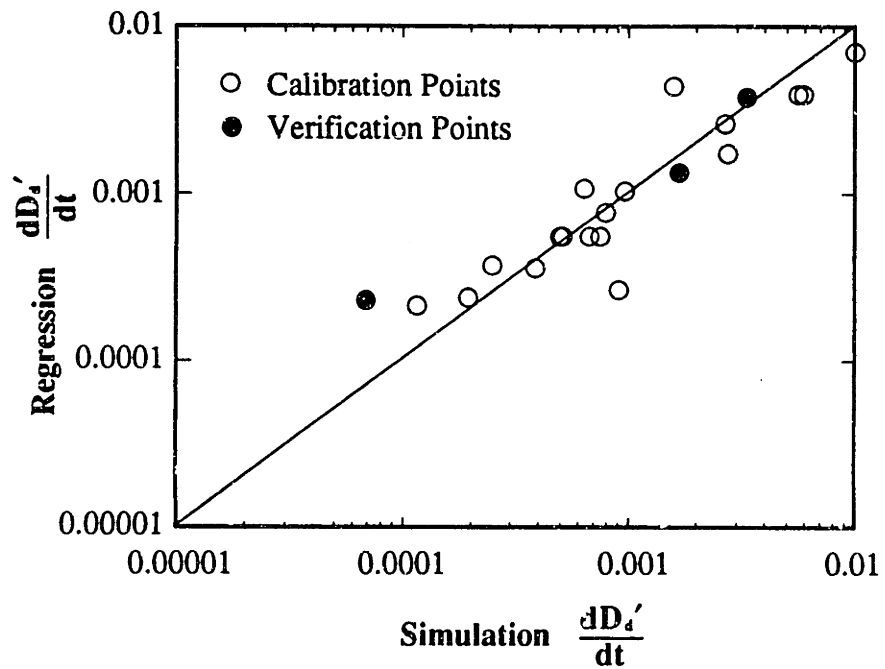
Figure 7.23: Simulation versus regression for drainage density growth with time

concave up, the rate of increase in elevation at the channel head per unit horizontal advance rises (and so the hillslope relief diminishes) at an ever increasing rate as the channel network grows. This reduction in hillslope relief slows the growth of each channel head by reducing the hillslope erosion. The total effect, though not large, is sufficient to offset the overall increase in the number of growth sites producing the linear growth stage. Howard's headward growth model does not model this second process and grows at the same rate everywhere, so it is believed it will not have a linear growth stage.

The slope of the linear part of the drainage density growth curve was fitted to the simulation data using multiple regression to yield

$$\frac{dD'_d}{dt'} = 5920 \ O_t^{0.18} \ TC^{0.15} \ TA_c^{3.09} \ TS^{0.56} \ (\frac{n_5}{m_5})^{5.88} \tag{7.15}$$

An observed versus estimated plot of the data is provided in Figure 7.23. A number of verification points are plotted and their agreement with the predictions is quite satisfactory.

To highlight the physical dependencies, this equation may be reformulated as

$$\frac{dD'_d}{dt'} = 5920 \ (O_t \cdot TS)^{0.18} \ TS^{0.38} \ TC^{0.15} \ TA_c^{3.09} \ (\frac{n_5}{m_5})^{5.88}$$

$$\tag{7.16}$$

The first term of Equation (7.16), $O_t \cdot TS$, is the nondimensional number for the hillslope erosion rate while the second term, TS, is the nondimensional number for the channel erosion rate. The rate of increase of drainage density is positively correlated with both. As channel erosion increases, because of the preferential channel erosion, the slopes in the hillslopes around the channel head are increased. This creates a larger region of activation around the channel head so that the rate of extension of the

channels is faster. A similar correlation between the size of the activated area and the rate of network extension was qualitatively noted during the study of the Meinhardt network model in Chapter 4, though the effect was never quantified.

In a similar sense increased erosion rates in the hillslopes also increase the size of the activated region. It is through the overland erosion term that the elevation changes in the channel are transmitted into the hillslopes. The faster these changes are transmitted, the faster the activated region can propagate into the hillslope and the faster the channel head can grow.

An incremental change in overland erosion has less effect than channel erosion because for hillslope erosion there is a competing process at work. As hillslope erosion increases, the area of influence of the channel increases, but the hillslope slopes also decrease. If hillslope erosion proceeds too quickly compared with the channel erosion then the slopes and thus the activator will be too low to trigger channelization; the hillslope is simply eroded flat too quickly in comparison to the rate of change of the channel elevation. It is thus postulated that for high values of $O_t \cdot TS$ the drainage density growth rate should decrease with increased $O_t \cdot TS$. Such behavior was qualitatively observed in the Meinhardt model. Such behavior was not observed in the physical model, but this possibility was never actively pursued. It may be, and it is so believed, that for realistic values of O_t this behavior is never seen.

The rate of drainage density increase is positively correlated with the rate of growth of channels at a point, TC. If a channel grows faster at a point, then the faster growth in the sediment transport rate in the channel results in higher slopes in the hillslope surrounding the channel head. This term is related to the transient sediment transport number, TS. A lower sediment transport rate in the channel may be partially offset by a faster growth of the channel at that point, so that during the hillslope to channel transition period the sediment transport rates are about the same.

Recall that during transition sediment transport rates are interpolated between that for the hillslope and that for the channel.

The rate of drainage density is positively correlated with the catchment scale activator number, TA_c. This term is the dominating influence on the rate of growth. This catchment scale activator number is also strongly correlated with the drainage density. Thus an increase in the drainage density is paralleled by an increase in the rate of growth of the network. Interestingly, Equations (7.9) and (7.15) suggest increases in TA_c result in increases in both drainage density and drainage density growth rate, but imply a <u>reduction</u> in the nondimensional time to ultimate drainage density.

There is a threshold on the value drainage density may take depending on whether $TA_c \lessgtr TA_h$ in Equation (7.9); there is no threshold on rate of growth in Equation (7.15). This is not an inconsistency. It simply says that there is a minimum network growth rate which occurs when $TA_c = TA_h$ in Equation (7.9). An explanation for this minimum growth rate follows from channel growth being driven by a region of high activator around the channel head. This region of high activator is produced by the steep slopes around the channel head arising from the preferential erosion in the channel eroding down the bottoms of the hillslopes. If the network grows too slowly, then hillslope erosion has sufficient time to flatten the hillslopes around the channel head so that the region of high activator disappears and the network can no longer grow. Thus the network must grow at a minimum rate to ensure that the steep slopes around the channel head are maintained. This minimum rate will be related to hillslope and channel erosion rates, as indicated by Equation (7.16).

The positive correlation of the network growth rate with n_5/m_5 follows from several network properties. The higher the drainage density, the faster the   rk growth rate; both the growth rate and drainage density are positively correlated with

282

n_5/m_5. For large values of n_5/m_5 the effect of network screening is reduced. For high n_5/m_5, channel growth is dominated by the local hillslope slopes around the channel head which are less affected by the surrounding network, than are the contributing areas. The hillslope slopes around the channel head arise primarily from the hillslope–channel erosion interactions, contributing areas primarily from drainage directions and overland erosion. This reduced screening effect for high n_5/m_5 means that there are more growth sites within the network so that overall growth rates are increased.

7.3.4 Conclusions

This section has correlated the nondimensional numbers introduced in Chapter 6 with properties related to drainage density and network growth.

Channel heads stop growing when the activator, dependent upon the contributing area and local slopes, falls below the activator threshold. At a channel head that is on the verge of growing activator is equal to the activator threshold. The activator number based on mean hillslope length scales, TA_h, was also shown to be constant. The importance of this relationship is that it provides the link between the hillslope length and the hillslope drop; it cannot, however, be used for prediction of the drainage density without knowledge of the hillslope drops. It was postulated, and it will be shown in Section 7.5, that catchment slopes resulting from tectonic uplift provide the missing link for determining the drainage density.

The drainage density was dependent upon two activator numbers, one based on hillslope length scales, TA_h, and one based on catchment length scales, TA_c. It was also found to be weakly dependent on the ratio of hillslope to channel erosion rates.

The rate of network growth was found to be strongly related to drainage density. The timescale of network growth of network growth was found to be positively correlated to the erosion and channel growth rates in a form that was

283

dimensionally consistent.

Both the drainage density and rate of network growth exhibited strong dependencies on the form of the activator function. It was possible to parameterize this effect via the ratio n_5/m_5. For large values of this parameter, both drainage density and rate of growth were large.

For the moment we leave the question of the interaction of horizontal and vertical length scales. It will returned to in Section 7.5

7.4 Planar Properties of the Landscape

This section is devoted to the planar properties of the simulated channel networks and the contributing hillslopes. These planar properties are related to the fundamental horizontal length scale of the catchment, hereto loosely referred to as the mean hillslope length. Much research has been devoted to parameterizing channel network planar properties irrespective of the catchment slope properties. These properties will compared with the results of the simulations.

7.4.1 Planar Link Properties

This section discusses the planar properties of the links in the simulated networks. Field workers have observed many relationships, some contradictory. Some examples of relationships include

1. interior and exterior links have different probability distributions (Smart, 1969)

2. link lengths are correlated with the link magnitude or the downstream link magnitude (Smart, 1981; Abrahams, 1984)

3. link length distributions are either exponentially (Smart, 1969) or gamma distributed (Abrahams, 1984; van der Tak, 1988).

284

To test whether mean interior and exterior link lengths were significantly different the mean link lengths for interior and exterior links were calculated for 17 of the catchment simulations. There were no apparent differences in the link distributions for different runs except in the mean. All used the same activator function so that potential variations in the relative lengths of interior links to exterior links from this source were eliminated. To eliminate the variation in the mean of the link lengths from catchment to catchment, the ratio of external link lengths to internal link lengths L_e/L_i was calculated. Because each simulated network had different magnitudes, the estimation variance of the <u>mean</u> link lengths for each network was different and related to the number of links

$$\sigma^2_{\overline{L}_e} = \frac{\sigma^2_{L_e}}{m}$$

where $\qquad \overline{L}_e \qquad$ = mean exterior link length

$\qquad$ m $\qquad$ = magnitude

$\qquad \sigma^2_{L_e}, \sigma^2_{\overline{L}_e} \qquad$ = variance of the link length and the estimate of the mean link

$\qquad$ length, respectively

and likewise for interior links

$$\sigma^2_{\overline{L}_i} = \frac{\sigma^2_{L_i}}{m-1}$$

where $\overline{L}_i$ = mean interior link length.

To eliminate this source of nonstationarity in the variance of the mean, a t—test was performed using a null hypothesis of

$$H_0 : \dfrac{(\frac{L_e}{L_i} - 1)}{\dfrac{\sigma}{L_e/L_i}} = 0 \qquad\qquad (7.17)$$

Values of this variable together with the original mean interior and exterior link lengths are given in Table 7.2. A two–sided t–test indicated that the null hypothesis could not be rejected at the 5% significance level; i.e. the mean exterior and interior link lengths were not significantly different. Not only that but the 95% confidence interval for the test variable was −0.30 to 0.37 with the mean value of the test variable in Equation (7.17) being 0.04. If these values accurately represent the population moments, then approximately 70000 links would be required to reject the null hypothesis, an enormous number of links. This suggests that the rejection of H_0 is quite strong and that for all intents and purposes the interior and exterior links should be considered to have the same mean length.

Smart (1981) suggested that link lengths are positively correlated with the magnitude of the downstream link. To test this hypothesis, link lengths were plotted against the link magnitude and the downstream link magnitude for the simulations CR2–6, CR7–1, CR7–2, and CR7–3 (Figures 7.24 and 7.25). The correlations between the link length and magnitudes are poor ranging from 0.07 to 0.21. None are significant. An analysis by eye of the figures suggests that an envelope curve on the link length distribution might indicate a decreasing maximum link length with link magnitude and no apparent trend with downstream link magnitude. This lack of correlation between the link lengths and magnitude is supported by field data (Abrahams, 1984).

The final link property examined was the probability distribution of the link lengths. Smart (1968) suggested that link lengths should be distributed with an

TABLE 7.2

Mean Exterior and Interior Link Lengths

RUN	$L_e^{(*)}$	$L_i^{(*)}$	$\dfrac{(\frac{L_e}{L_i} - 1)}{\sigma_{L_e}/L_i}$	$L_\ell^{(*)}$	σ_{L_ℓ}	$\dfrac{L_\ell^2}{\sigma_{L_\ell}^2}$
CR1–3	2.34	2.22	0.29	2.27	1.93	1.38
CR2–3	6.58	4.81	0.79	5.72	5.04	1.29
CR2–4	21.50	6.24	1.29	15.40	9.69	2.52
CR2–6	3.47	3.60	−0.18	3.53	2.96	1.42
CR3–3	1.87	1.59	1.00	1.73	1.34	1.67
CR3–4	3.42	4.61	−1.24	4.01	3.53	1.29
CR3–5	3.34	3.31	0.04	3.32	2.94	1.27
CR4–2	3.48	3.03	0.71	3.26	2.45	1.77
CR4–5	3.26	3.07	0.32	3.17	2.37	1.78
CR4–6	5.37	5.89	−0.31	5.62	3.94	2.03
CR4–7	5.88	5.66	0.14	5.77	2.68	2.46
CR4–9	5.29	5.25	0.03	5.27	3.71	2.01
CR5–1	5.60	6.29	−0.39	5.93	4.24	1.95
CR5–2	2.06	2.27	−0.64	2.16	1.89	1.31
CR7–1	3.84	3.88	−0.05	3.86	3.14	1.51
CR7–2	3.85	4.47	−0.64	4.15	3.63	1.30
CR7–3	3.55	4.13	−0.53	3.84	3.08	1.55

(*) L_e, L_i = mean exterior and interior link lengths respectively
L_ℓ = mean link length
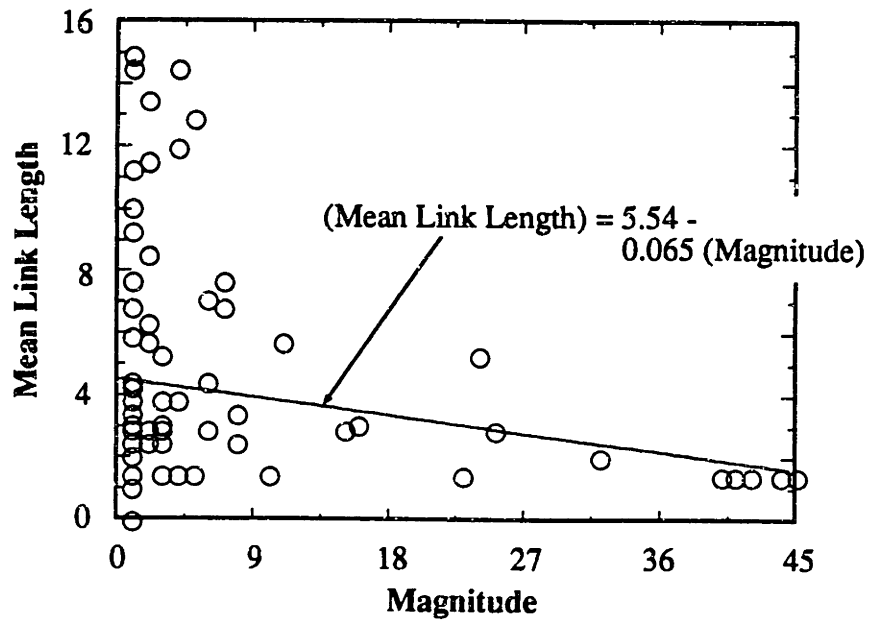
Figure 7.24: Mean link length versus link magnitude: Simulation CR2-6
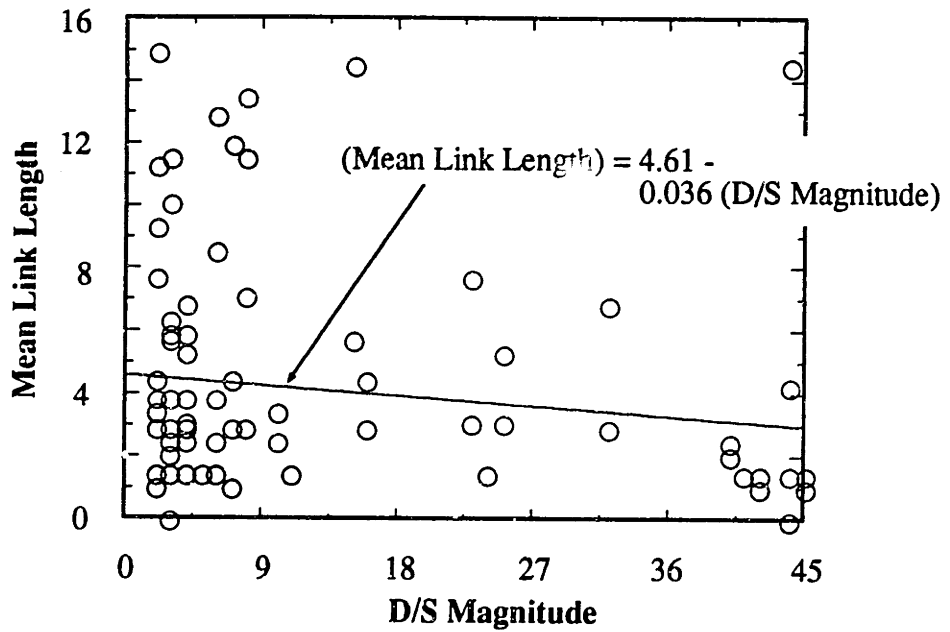
Figure 7.25: Mean link length versus downstream link magnitude: Simulation CR2-6

exponential distribution. More recent evidence (Abrahams, 1984; van der Tak, 1988) suggested that a gamma distribution with a shape parameter, α, around 2 provided a better fit to the observed data.

To test these hypotheses with the simulated data, the moments of the link length distribution were calculated, lumping interior and exterior together. As was previously shown, the mean length of interior and exterior links were shown to be not significantly different, and it is believed that this is also true for higher order moments so that the link distributions for interior and exterior links may be considered jointly. This reduces the estimation error on the moments calculated below. The mean, standard deviation and gamma distribution shape parameter determined by the method of moments are given in Table 7.2. The shape factor, α, for the gamma distribution was calculated from

$$\alpha = (\tfrac{\mu}{\sigma})^2$$

using the moment estimates of mean and variance. The mean shape factor for the simulated catchments was 1.7 ± 0.4. van der Tak (1988) calculated a shape factor of 1.8 for two basins in New Hampshire and one in Arizona, using maximum likelihood estimation. Using the method of moments his results for shape factor are 1.66 ± 0.64. The simulation results are consistent with this result suggesting that link lengths are gamma distributed with a shape parameter of about 1.5 − 2.0.

7.4.2 Fundamental Length Scales of the Catchment

This section will briefly discuss observed relationships between various "fundamental" horizontal length scales of the catchment. They all, in some sense, capture the length scale of dissection of the landscape loosely referred to as the mean hillslope length. It will be demonstrated that while this is a reasonable viewpoint,

there are some subtle differences between the various measures that reflect the physical processes at work in the hillslope.

Horton (1945) proposed that the drainage density and the mean hillslope length be related by

$$D_d = \frac{1}{2L_h} \tag{7.18}$$

and argued that channels form when runoff occurs over longer distances than the mean hillslope length.

Other useful measures of the level of dissection are

1. mean link length, L_ℓ

2. mean interior link length, L_i

3. mean exterior link length, L_e

4. square root of the first order Strahler area, A_1

Recently Montgomery and Dietrich (1988) introduced the concept of source area, A_s, that area observed to contribute to the channel head. The square root of the mean source area, A_s, is another length scale (Equation 7.10).

Another length scale calculated here and related to the source area concept is the mean hillslope length of the lateral contributing areas, $L_{\ell h}$, hereafter called the lateral hillslope length. This is different to the definition of mean hillslope length based on the drainage density since it excludes the source area so that only the lateral flow to the channels is considered. The lateral hillslope length will always be less than the mean hillslope length. It is defined as

$$L_{\ell h} = \frac{A - mA_s}{2\ell} \tag{7.19}$$

where

A = catchment area

ℓ = total length of channel network

m = network magnitude

A_s = mean source area contributing to the channel head

The various length scale definitions are illustrated in Figure 7.26.

These length scales were calculated for 20 of the computer simulations listed in Appendix D. There appeared to be no significant trends with any nondimensional numbers except for a trend with the ratio n_5/m_5. All results presented here are for $n_5/m_5 = 0.75$. The various length scales were then plotted against each other and a simple least squares regression of the following form fitted

$$(\text{length scale}_1) = \alpha \, (\text{length scale}_2)$$

The results of this regression are given in Table 7.3, and the correlation coefficients for each of the fits is given in Table 7.4. In general, the correlation between the various length scales is very high (typically better than r = 0.98).

Some notable relationships are:

1. $L_e = 1.27\, L_i$, yet Section 7.4 showed that L_e and L_i are not significantly different (implying $L_e = L_i$). Note that the correlation coefficient in Table 7.4 is relatively low. This anomalous regression is believed to result from the regression being fitted by simple least squares; i.e., it does not reflect the estimation error on the mean link length which is dependent on network magnitude. The hypothesis test on link lengths in Equation (7.17) is weighted by the estimation error of the means. The correlation between the interior link length, L_i, and the other exterior link length scales, $(A_1)^{1/2}$, and $(A_3)^{1/2}$ is relatively poor.

TABLE 7.3

Regression of Length Scales from simulations: $y = a\,x$

x \ y	$A_1^{1/2}$	L_e	$A_s^{1/2}$	L_ℓ	L_i	L_h
L_e	1.07					
$A_s^{1/2}$	1.37	1.28				
L_ℓ	1.28	1.18	0.928			
L_i	1.42	1.27	1.03	1.16		
L_h	2.25	2.08	1.64	1.74	1.24	
$L_{\ell h}$	3.71	3.39	2.69	2.93	2.24	1.66

TABLE 7.4

Correlation Coefficient of Regression (R^2) on length scales from simulations

x \ y	$A_1^{1/2}$	L_e	$A_s^{1/2}$	L_ℓ	L_i	L_h
L_e	0.994					
$A_s^{1/2}$	0.999	0.993				
L_ℓ	0.987	0.970	0.984			
L_i	0.755	0.710	0.757	0.852		
L_h	0.997	0.984	0.996	0.993	0.803	
$L_{\ell h}$	0.953	0.920	0.947	0.986	0.918	0.970

2. The source area is typically about 50–70% of the Strahler first order area. This regression has a high statistical significance but is only valid for $n_5/m_5 = 0.75$. There is a strong dependence on n_5/m_5; for $n_5/m_5 = 2$ the source area is 70% of the Strahler first order area.

3. $L_e = 1.08 \sqrt{A_l}$, the length of the exterior links are very strongly related to area contributing to that link.

4. There is a significant difference between the length of lateral hillslopes, $L_{\ell h}$, and the mean hillslope determined from drainage density, L_h. This difference reflects systematic differences in the hillslope form for the source area and the lateral hillslope areas. A typical first order area is approximately twice as long as it is wide (Figure 7.26). The area upstream of the branch head (the source area) is 50% of the exterior link area so that there is significant difference between the lateral hillslope length and the mean hillslope length.

This difference between lateral hillslope length and the length scale of the source area is important. We have previously noted that at the channel head (or any potential channel head about to branch off laterally)

$$TA_h = constant$$

If the hillslope length is shorter, then everything else being equal, the slopes must be higher to compensate and maintain the value of TA_h. This effect is clearly seen in Figure 5.10. The source areas contributing to the channel head tend to be long and have relatively low slope, while those hillslopes contributing laterally tend to be shorter and have higher slopes. Valleys at early time tend to be long and flat
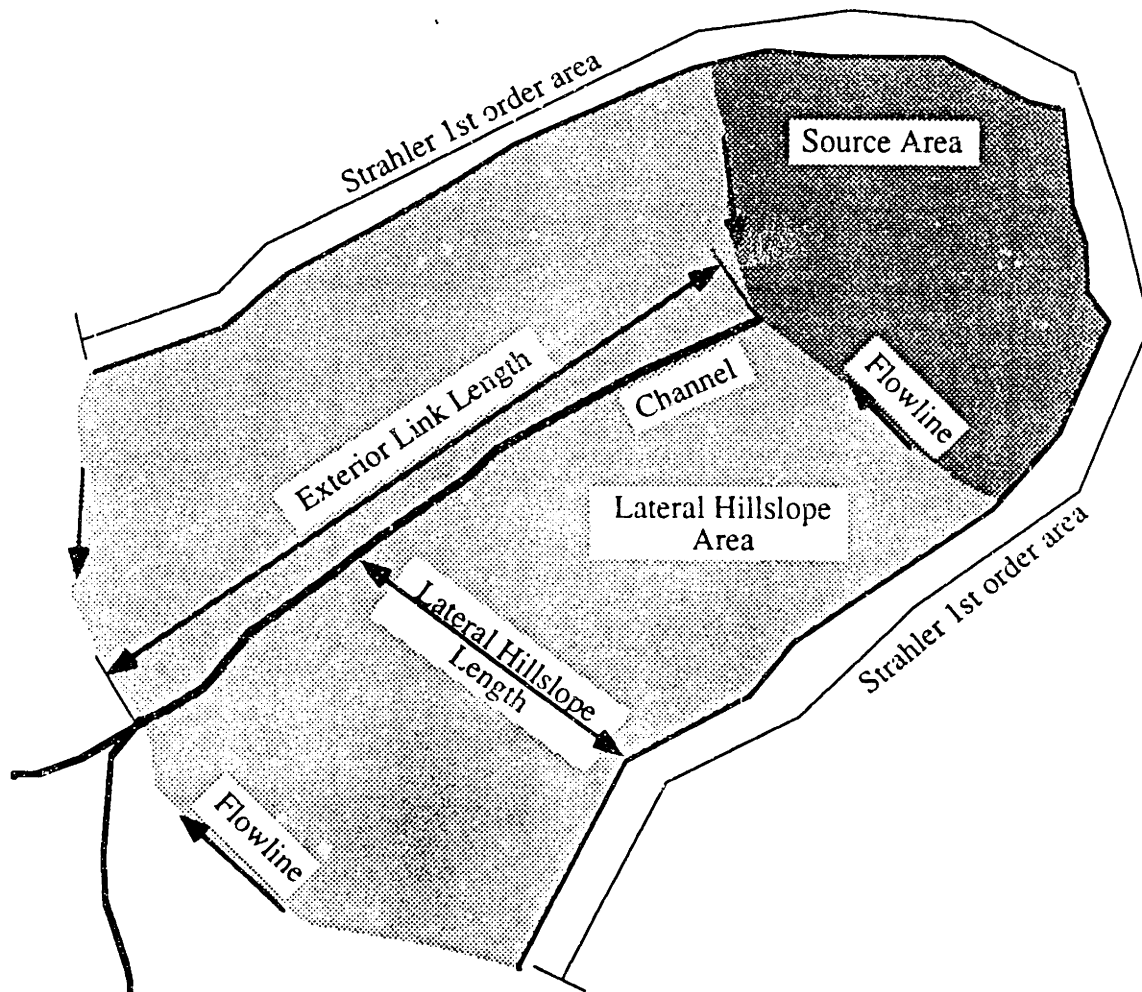
Figure 7.26: Definition of the hillslope
length scales

longitudinally, even above the channel head, and steep laterally. Towards dynamic equilibrium this tendency is reduced

The differences between the lateral hillslope length and the mean hillslope length from drainage density influences the hillslope ruggedness number calculated by Strahler (Equation 2.13) where

$$R_h = \frac{HD_d}{S_g} = \frac{H}{2L_h S_g}$$
$$\approx \frac{H}{3.3 \ L_{\ell h} S_g} \qquad \text{(from Table 7.3)} \qquad (7.20)$$

where   H   = mean hillslope drop

            S_g = some characteristic hillslope slope.

The lateral hillslope length and the mean hillslope lengths relationship is from Table 7.3. If S_g is the mean lateral hillslope slope (i.e. $S_g = H/L_{\ell h}$) this equation becomes

$$R_h = 0.30 \qquad (7.21)$$

where Strahler (1958) proposed the right—hand side value to be 0.5. Strahler found values of R_h ranging from 0.3 to 1.0.

7.4.3 Link Contributing Area Relationship

Smart (1972) suggested that the area draining to an exterior link was different to the area draining laterally to an interior link. He defined a dimensionless variable, α, such that

$$\alpha = \frac{A_{exterior}}{A_{interior}}$$

and found its mean value to be 1.85 ± 1.5 for data from 12 catchments. The area data for the 17 simulations that were used in Section 7.4 were analyzed; the resulting α values are presented in Table 7.5. In Section 7.4 a significance test was performed to test if interior and exterior link lengths were significantly different. In a similar fashion to that done for the test on link lengths (Section 7.4.1) the link areas were transformed using the estimated variances of the mean areas to produce a statistically homogeneous variable. A hypothesis test was performed on this transformed variable with the null hypothesis that the interior and exterior link areas are the same i.e.

$$H_0 : \frac{\alpha - 1}{\sigma_\alpha} = 0$$

At the 5% level a t-test indicated α to be significantly greater than 1, with the 5% confidence limits on α being 2.34 and 3.10. The mean and standard deviation for α from the simulations was 2.53 ± 2.8. Given the large standard deviations in α and the 5% bounds, the simulated results are considered to be consistent with the results of Smart (1972).

Shreve (1974) found for Rockcastle Creek, Kentucky, that the ratio of the mean exterior link length to the mean interior link length was 1.04 and the ratio of the mean exterior link area to mean interior link area was 1.56. These are consistent with the results above.

The difference in the areas contributing to exterior and interior links is due to the source areas contributing to the exterior links. In Section 7.4.2 the mean lateral hillslope length and the mean link lengths were linearly related (there was no significant difference between the lateral hillslope lengths for interior and exterior links). Since exterior and interior link lengths are the same, exterior and interior links have the same lateral contributing area but the exterior links have an additional

TABLE 7.5

Mean Exterior and Interior Link Areas

Run	$\overline{A}_e$	$\overline{A}_i$	α	$\dfrac{A_e/A_i - 1}{\sigma_{A_e/A_i}}$
CR1–3	7.21	3.49	2.07	2.17
CR2–3	6.58	4.81	1.37	0.79
CR2–4	480.0	36.5	13.15	2.34
CR2–6	15.93	9.00	1.77	1.45
CR3–3	4.65	1.69	2.75	3.73
CR3–4	19.7	15.7	1.26	0.69
CR3–5	15.5	8.23	1.89	1.76
CR4–2	14.4	6.14	2.34	2.31
CR4–5	13.8	6.53	2.11	2.31
CR4–6	45.9	24.6	1.86	1.37
CR4–7	47.8	22.0	2.18	1.67
CR4–9	42.6	20.7	2.06	1.66
CR5–1	43.5	26.8	1.62	1.13
CR5–2	6.65	3.63	1.83	2.02
CR7–1	17.0	9.66	1.76	1.64
CR7–2	17.6	13.3	1.33	0.81
CR7–3	16.7	10.5	1.60	1.37

(*) A_e, A_i = lateral area draining to exterior and interior links respectively.

hillslope contribution from the source areas contributing to the branch head. This source area is about 50% of the contributing area to exterior links (Table 7.5) so that exterior link areas are about twice the interior link areas, consistent with the results of the hypothesis test above.

7.5 Relating Hillslope and Catchment Length Scales and Channel Network Magnitude

This section will demonstrate that the results of Sections 7.3 and 7.4 may be used to relate the catchment area and catchment relief to the hillslope length and hillslope relief, and thus to drainage density and magnitude. In this way it will be shown that there are only two independent length scales, the catchment size and the catchment relief. All other vertical and horizontal length scales, as well as the channel network magnitude, follow directly, given the assumed catchment physics. Through the relationship developed below the essential unity of the length scales will be demonstrated; neither the hillslope nor the channel scales can be considered independently of the other. The drainage density is intimately connected to both the slopes of the catchment and the slopes of hillslopes.

In Section 7.3 the hillslope length scale, L_x, was defined as the square root of the Strahler first order area. This definition will be used below. The catchment length scale, L_d, as in Section 7.3, will be defined as the square root of the catchment area. The catchment vertical scale, L_e, will be defined as the catchment relief and the hillslope vertical scale, L_z, as the mean drop of the hillslopes in the Strahler first order areas.

Drainage density and the mean hillslope length are related by (Horton, 1945)

$$D_d = \frac{1}{2L_h} \tag{7.22}$$

and from the definition of drainage density, we may relate drainage density, magnitude, m, mean link length, L_ℓ, and the catchment horizontal scale, L_d, by

$$D_d = \frac{(2m-1)L_\ell}{L_d^2} \qquad (7.23)$$

From the previous section the mean hillslope length, L_h, and the hillslope length scale, L_x, can be related by

$$L_h = \eta_1 \, L_x$$

where $\qquad \eta_1 \qquad = \frac{1}{2.25}$ (from Table 7.3).

The mean link length can be related to the hillslope length scale by

$$L_\ell = \eta_2 \, L_x$$

where $\qquad \eta_2 \qquad = \frac{1}{1.28}$ (from Table 7.3).

Substituting these relationships into Equations (7.22) and (7.23) and eliminating drainage density yields the following relationship between the hillslope, the catchment horizontal length scales and the magnitude

$$L_x = \frac{L_d}{\sqrt{2(2m-1)\eta_1\eta_2}} \qquad (7.24)$$

It was shown in Section 7.3 that the hillslope activator number, TA_h, is constant so that

$$TA_h = C \tag{7.25}$$

where
$$TA_h = \frac{L_x^{2m_5-n_5} \; L_z^{n_5} \; L_R^{m_5} \; c_1 \; \bar{\beta}_5}{T_R}.$$

This equation provides the connection between the hillslope horizontal and vertical length scales and upon rearrangement yields

$$L_z = \left[\frac{CT_R}{c_1 \bar{\beta}_5}\right]^{\frac{1}{n_5}} L_x^{\frac{n_5-2m_5}{n_5}} L_R^{-\frac{m_5}{n_5}} \tag{7.26}$$

Substituting Equation (7.24) into Equation (7.26) provides a relationship between the hillslope vertical length scale, the catchment horizontal length scale and the magnitude.

To connect the magnitude to the catchment vertical and horizontal length scales, it is necessary to use the nondimensional drainage density relationship of Equation (7.9). Simplifying this drainage density relationship yields

$$L_d D_d = \frac{1}{28} \left[\left(\frac{L_d}{L_x}\right)^{2m_5-n_5} \left(\frac{L_e}{L_z}\right)^{n_5} - 1\right]^{0.53} O_t^{-0.06} \left(2.5 + \frac{n_5}{m_5}\right)^{4.12} \tag{7.27}$$

Substituting Equation (7.22) for the drainage density in this equation and Equation (7.26) for the hillslope vertical length scale yields the following expression for the horizontal hillslope length scale

$$L_x = \frac{14L_d}{\eta_1} \left[\left[\frac{c_1 \bar{\beta}_5}{C\,T_R} \right] L_R^{m_5} L_d^{2m_5 - n_5} L_e^{n_5} - 1 \right]^{-0.53} O_t^{0.06} \left[2.5 + \frac{n_5}{m_5} \right]^{-4.12}$$

$$(7.28)$$

This equation can be rearranged to highlight the dependence of the hillslope horizontal scale upon the mean catchment slope scale $S = L_e/L_d$

$$L_x = \frac{14L_d}{\eta_1} \left[\left[\frac{c_1 \bar{\beta}_5}{C\,T_R} \right] L_R^{m_5} L_d^{2m_5} S^{n_5} - 1 \right]^{-0.53} O_t^{0.06} \left[2.5 + \frac{n_5}{m_5} \right]^{-4.12}$$

$$(7.29)$$

and for a large catchment $(L_d >> 1)$ then

$$L_x \approx \frac{14L_d^{(1-m_5)}}{\eta_1} \left[\left[\frac{c_1 \bar{\beta}_5}{C\,T_R} \right] L_R^{m_5} S^{n_5} \right]^{-0.53} O_t^{0.66} \left[2.5 + \frac{n_5}{m_5} \right]^{-4.12}$$

$$(7.30)$$

As the catchment becomes steeper (S increases) the hillslope length, L_x, becomes shorter.

The hillslope slope scale, L_S ($=L_z/L_x$), and the catchment slope scale can be related so that

$$L_S \approx \left[\frac{14L_d}{\eta_1} \right]^{\frac{-2m_5}{n_5}} \left[\frac{c_1 \bar{\beta}_5}{C\,T_R} \right]^{\frac{m_5-1}{n_5}} \cdot$$
$$\left[L_R^{1-m_5} O_t^{-0.1} \right]^{\frac{m_5}{n_5}} \left[2.5 + \frac{n_5}{m_5} \right]^{\frac{8.2\,m_5}{n_5}} S^{1.1\,m_5}$$

$$(7.31)$$

As the mean catchment slope increases the hillslope slope roughly increases with the

square of the mean catchment slope ($m_5 \approx 2$), with a commensurate increase in drainage density.

The importance of these equations is that they are dependent upon only two independent length scales: the catchment horizontal length scale, L_d, and the catchment vertical length scale, L_e. All other catchment scales can be expressed as dependent variables of L_d and L_e. That is,

1. The hillslope horizontal length scale, L_x, follows directly from Equation (7.28)

2. The hillslope vertical length scale, L_z, follows from Equations (7.28) and (7.26).

3. The catchment magnitude, m, follows from Equations (7.28) and (7.24).

4. Other horizontal scales such as source area size, lateral hillslope length and mean link length follow from Equation (7.28) and the regressions of Section 7.5.

That all catchment scales can be expressed in terms of catchment size and catchment relief is important for three reasons. The first, and foremost, reason is that they show the importance of the catchment scale processes in determining the hillslope properties. Through the dependence on catchment slopes it is shown that the hillslope length scales are dependent on catchment scale processes like tectonic uplift. The mechanism for the connection of scales is the channel network. Thus while geomorphologists observe the drainage density, and attribute to it the fundamental scale of the landscape, the hillslopes are governed by larger scale processes as much as by hillslope activator processes. Neither hillslope nor catchment scales are preeminent, both interact to produce the observed landscape.

The second reason these relationships are important is that they establish that there are only two independent length scales: one vertical and one horizontal. The

nondimensionalization of the governing equations presented in Chapter 6 used only two length scales: one vertical and one horizontal. If there were more than two independent length scales then that nondimensionalization would be erroneous.

The third reason this relationship is important is purely pragmatic. Hillslope length scales can now be predicted from catchment properties. In the field setting the two known variables will be the tectonic uplift and the size of the basin. If the tectonic uplift is known, then the catchment relief may be determined.

1. For a single uplift event, the catchment relief is approximately the magnitude of the tectonic uplift.

2. For continuous uplift, the catchment relief is the cumulative amount of tectonic uplift at the time the network stops growing. It is not the catchment relief at dynamic equilibrium, which, as will be shown later, is slightly larger. The catchment relief is determined from equation (7.16) iteratively. A final catchment relief is assumed fixing the nondimensional numbers and the network growth rate is calculated, which determines the total uplift at the time the network attains its ultimate drainage density implied by the assumed relief. By trial and error the initially assumed catchment relief, and that given by the tectonic uplift and the rate of growth of the network can be matched.

The size of the catchment, the other input into the predictive equation, may be constrained by geologic conditions, experimentally applied boundary conditions or the competition between adjacent catchments. If, for instance, tectonic uplift comes from a thrust block uplift then the areal extent of the uplift will effectively govern the maximum possible size of the catchment.

Once the two catchment scales are determined from the governing physical conditions, the hillslope and network scales follow directly. An alternative interpretation of Equations (7.24) and (7.31) is that with a knowledge of the hillslope

length scales these equations give the relationship between catchment area ($\equiv L_d^2$) and catchment relief ($\equiv L_e$) for the catchment draining a downstream point in the channel. At the downstream outlet of the catchment these alternative views of the equations are equivalent.

For the typical physical conditions where $0 < (2m_5 - n_5) < 1$, the following trends in hillslope length can be deduced from Equation (7.28).

1. For constant uplift, as the catchment area increases the hillslope length increases. This is because the catchment, overall, is flatter and less erosive, so that hillslopes are flatter and thus longer so as to maintain $TA_h = C$.

2. As the catchment relief increases, the hillslope length decreases. This is because the catchment, overall, is steeper. Again $TA_h = C$ must be maintained.

3. As the ratio of hillslope erosion to channel erosion rates, O_t, decreases, the hillslope length decreases. This is because hillslopes get steeper and channels flatter. This follows from the sediment transport continuity around the channel head discussed in Section 7.3.2.

4. As the hillslope activator increases, or the activator threshold decreases, the hillslope length decreases. This is because the erosion potential of the hillslope has increased comparative to the hillslope's ability to resist the erosion. The hillslope activator may increase if, for instance, the runoff increases.

For the change in conditions of points 1 and 2 above (i.e. changes in the mean catchment slope) the trends of the hillslope relief, L_z, are the inverse of those for the hillslope length scale, L_x. This follows from $TA_h = C$. For the changes in conditions of point 3 (i.e. increased erosive power) the trends of L_z are similar to those of L_x.

7.6 Hypsometric Relations and Dynamic Equilibrium

The hypsometric curve has long been used by geomorphologists as an indicator of the age of catchments. For instance, Schumm (1956) attributed differences in hypsometric curves for different catchments in the same region to the different ages of the catchments (Figure 2.9). To test this hypothesis, hypsometric curves have been plotted for two different simulations. The first simulation, based on run EB11–9–5, consisted of an event uplift followed by network development and decline of the overall catchment elevations with time. This corresponds to an episodic uplift, erosional history. The uplift history is thus

$$c_0(t) = \bar{c}_0 \, \delta(t - t_0)$$

where $\delta(t)$ = dirac delta function.

The second simulation, based on runs CR2–3 and CR8–1, consisted of an uplift event followed by erosion and network growth. At the time at which the network attained its maximum drainage density a second uplift, this time continuous with time, that just balanced the sediment outflow from the catchment, was introduced. The catchment was then allowed to proceed to dynamic equilibrium where the tectonic uplift just balances the sediment outflow. The uplift history in this case is

$$c_0(t) = \bar{c}_0 \, \delta(t) + \bar{\bar{c}}_0 \, H(t - t_0)$$

where $H(t-t_0)$ = Heavyside step function

$$= \begin{cases} 0 & t < t_0 \\ 1 & t \geq t_0 \end{cases}$$

The hypsometric curves, for various times, for the episodic uplift case are given

in Figure 7.27, while those for the dynamic equilibrium case are given in Figure 7.28. The hypsometric curves in Figure 7.27 were calculated from a simulation using a 20 x 20 grid while those in Figure 7.28 were for a simulation using a 40 x 40 grid. Thus the curves in Figure 7.27 appear rougher reflecting the coarser spatial resolution of this simulation. In addition, there are some slight random differences between the hypsometric curves because the networks were generated from different initial random elevation perturbations. Figure 7.14 indicates that this sampling effect should be small.

The time evolution of the two curves is very different. However, there are some broad similarities between the runs. Both runs are predominantly concave down at early times, and the area under the curve diminishes with time. This is consistent with Schumm (1956). Both curves, at early times, exhibit a kink in them where the slope of the hypsometric curve changes dramatically. This kink corresponds on the horizontal scale to the drainage density of the catchment (Figure 7.27). Nodes in the channel have dramatically reduced elevation, while nodes on the hillslope undergo little elevation change from the initial relief. As the drainage density of the network increases and more of the catchment is channelized, the kink in the hypsometric curve moves to the left, indicating that more and more of the catchment is falling under the influence of the channel network. At times long after the network has stopped growing the notch disappears because hillslope erosion has had sufficient time to come into equilibrium with the new stable channel network, removing the sharp differentiation in elevation between the hillslope and the channel network. A similar kink exists in the general form of the hypsometric curve for young catchments proposed by Strahler (1952, 1964) and observed by Schumm (1956). It seems reasonable to suggest that in the field, analogously with the simulations, that this kink indicates that proportion of the catchment in valleys under the influence of the more quickly eroding channel network.

kink corresponding
to portion of
catchment under
the influence of
the channels

Figure 7.27: Hypsometric curve: Episodic uplift,
elevations normalized by initial relief
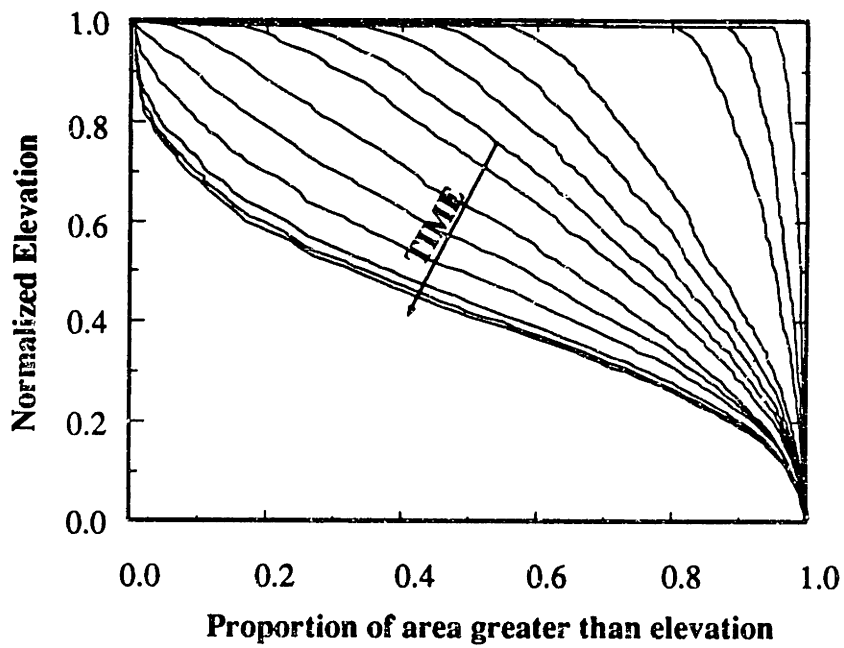
Figure 7.28: Hypsometric curve: Continuous uplift,
elevations normalized by initial relief

307

In Figure 7.27 the hillslope portion of the hypsometric curves declines at a faster rate than in Figure 7.28. The simulation EB11–9–5, on which Figure 7.27 is based, used an overland erosion rate, O_t, that was three times higher than that in the simulation CR2–3/CR8–1, upon which Figure 7.28 is based.

The most interesting aspect of the hypsometric curves is their trend with time long after the network has stopped growing. The hypsometric curve for episodic uplift converges to a curve that is concave up everywhere. The general trend is consistent with that observed for erosional development of catchments in spoil heaps by Schumm (1956) (Figure 2.9). The uplift histories for the simulation data and Schumm's field data are comparable; the initially flat spoil heap of Schumm's study could be conceptualized as having resulted from a single uplift event. If the hypsometric curve for the simulations is replotted so that the elevation axis is now normalized by the original catchment relief (Figure 7.29) another similarity appears. The stage where the curve is concave down overall corresponds to that period when the catchment relief is unchanged from the original catchment relief. The period when the hypsometric curve is concave up corresponds to the period when the catchment relief is declining from the original value. The reason for this behavior is that when the catchment relief is declining the portions with the slowest rate of decline are the low contributing areas, high elevation regions near the catchment watershed. In some sense these low erosion zones are "holding up" the edges of the catchment while everywhere else the catchment is declining more quickly. Then the hypsometric curve is concave upward. This argument parallels that for the existence of a Monadnock stage of catchment development where "isolated rock may form isolated hills rising above a generally subdued surface..." (p. 4–69 Strahler, 1964). It is, apparently, unnecessary to require isolated rocky outcrops with their accompanying source limitation of transport; the low erosion regions around the catchment watersheds suffice to produce a monadnock stage
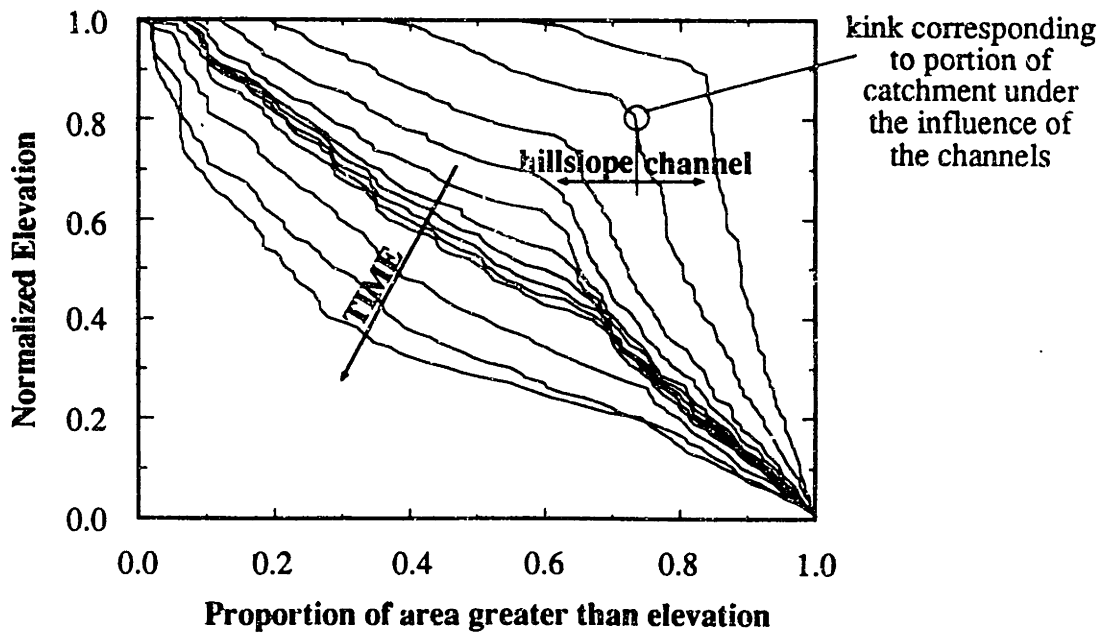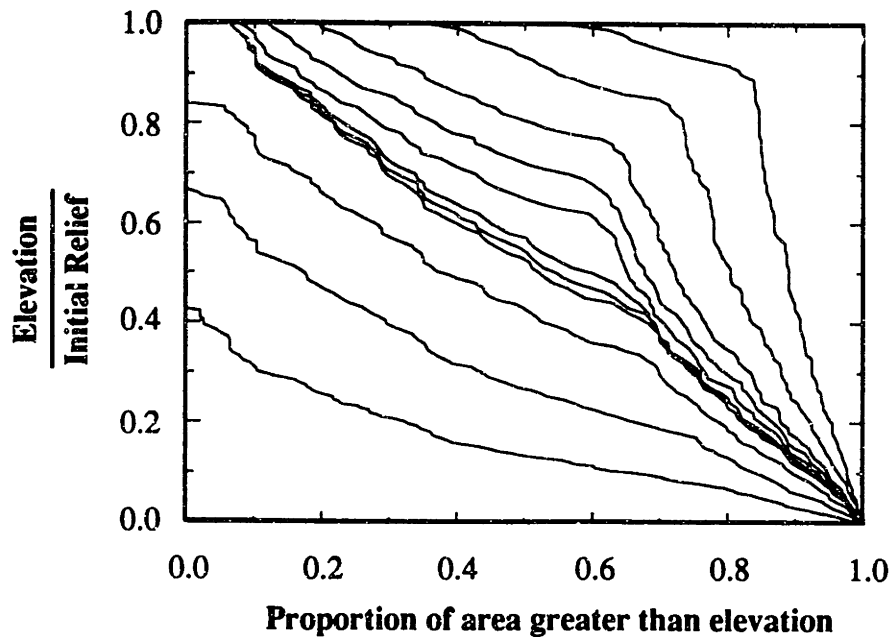
Figure 7.29: Hypsometric curve: Episodic uplift, elevations normalized by current relief
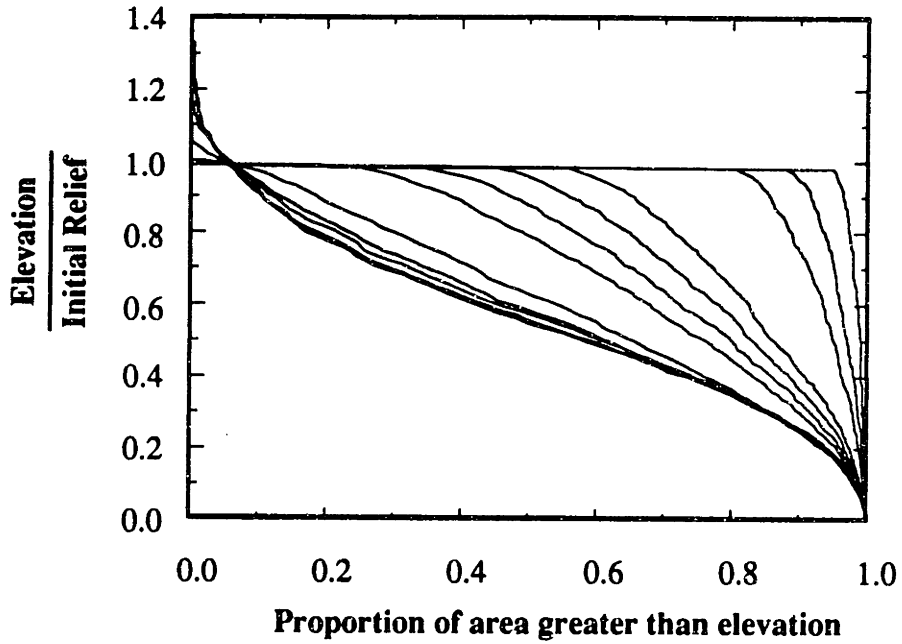
Figure 7.30: Hypsometric curve: Continuous uplift, elevations normalized by current relief

hypsometric curve where there is an overall catchment downwasting.

The hypsometric curve for the dynamic equilibrium case (Figure 7.28) tends with time to an S—shaped curve very similar to the curve observed by Strahler (1964) and which he calls the "equilibrium stage". Contrary to expectations, it appears in Figure 7.28 that even after the continuous tectonic uplift is imposed, that the mean elevation of the catchment is still declining. This impression results from the curve being normalized against the current catchment relief. If instead, the curve is normalized by the original catchment relief, then it becomes clear that catchment relief increases with time, but that the mean elevation, as expected, is constant (Figure 7.30). There is a redistribution of elevation within the catchment which is accompanied by an increase in catchment relief as the catchment approaches dynamic equilibrium. Hillslopes tend to become higher and steeper, the valley portion of the catchment remains much the same. Despite the requirement that $TA_h = C$ (Equation 7.24) the hillslopes can become steeper overall without inducing active network growth. The slopes of the hillslopes immediately upstream of the channel heads do not increase but the slopes of the hillslopes close to the watershed do; the nondimensional shape of the hillslope is changed. This characteristic will be shown more clearly in Section 8.3 and Figures 8.6 and 8.7 where slope—area relationships for channels and hillslopes at dynamic equilibrium will be discussed in detail.

These qualitative differences in the hypsometric curves are corroborated by contours of elevations with time (Figures 7.31 and 7.32). Both figures show the contours of elevation (normalized by the current catchment relief) for the time period after attainment of the maximum drainage density. This time period corresponds to that period when the catchment relief is either increasing or decreasing from its initial value (Figures 7.29 and 7.30). For the episodic uplift case (Figure 7.31) there is a substantial increase with time in the proportion of area at low elevations. For the

(a) time = 10000
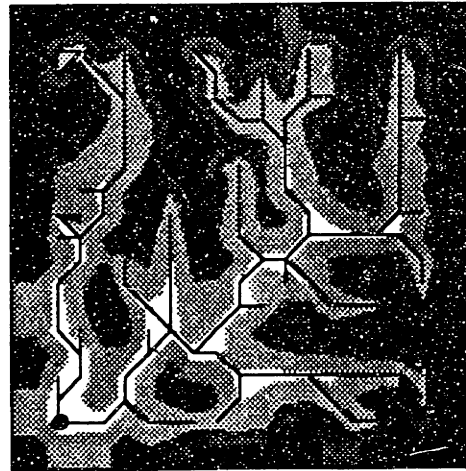
(b) time = 20000
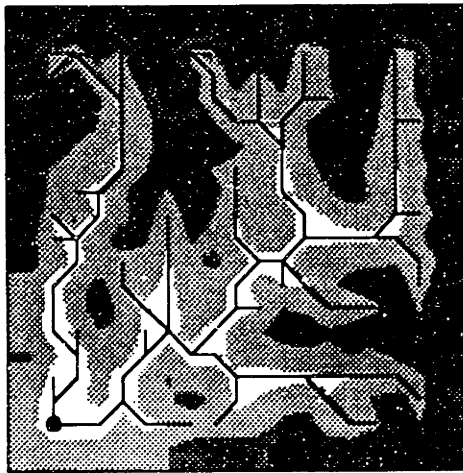
(c) time = 40000

(d) time = 100000

Maximum
Elevation

Minimum
Elevation
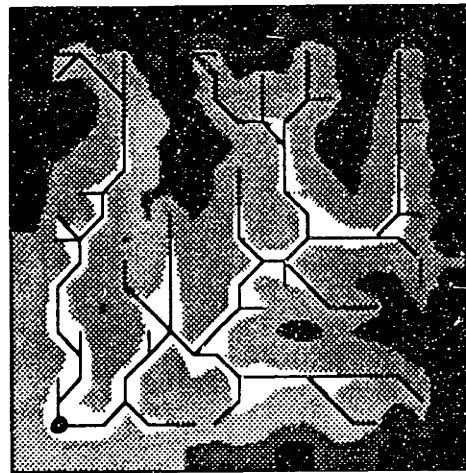
● Catchment outlet

Figure 7.31: Elevation contours for episodic uplift,
simulation EB11-9-5.
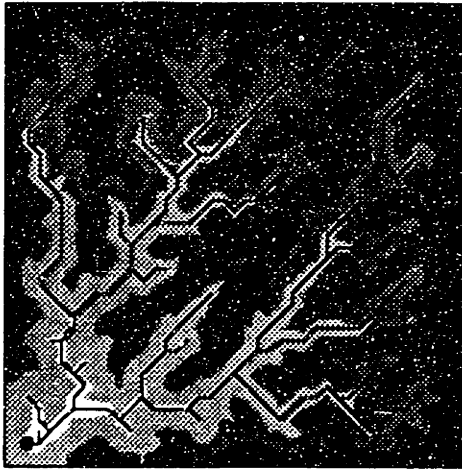
(a) time = 15000
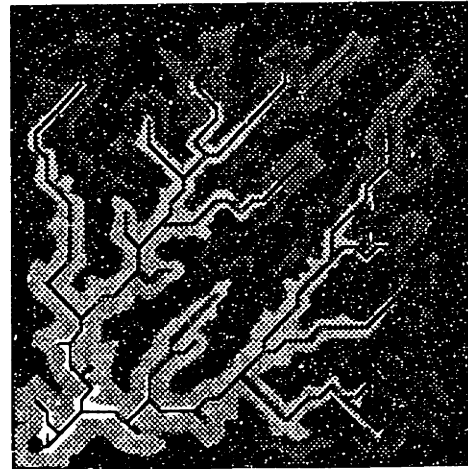
(b) time = 25000

(c) time = 35000

(d) time = 60000

Maximum
Elevation

Minimum
Elevation
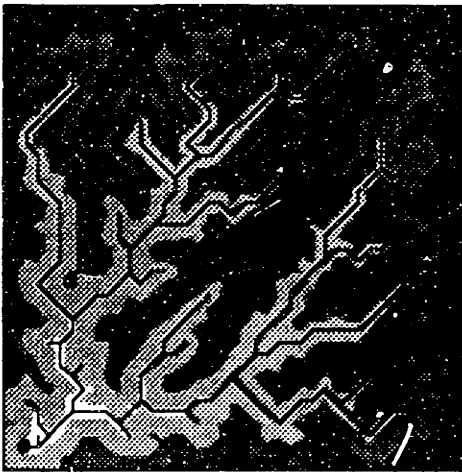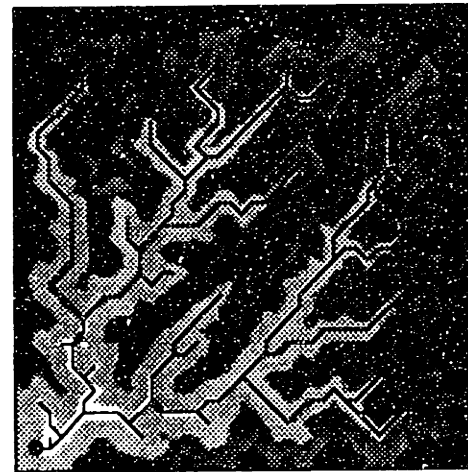
● Catchment outlet

Figure 7.32: Elevation contours for continuous uplift,
simulation CR8-1.

continuous uplift–dynamic equilibrium case (Figure 7.32) changes in elevation distribution are confined to highest elevations. The proportion of area at the highest elevations is declining with time. The trends with time of the elevation contours are consistent with those of the hypsometric curves for both catchments.

On the basis of these the simulation results it is asserted that the differences between the "equilibrium" or "mature" stages and the "Monadnock" stage referred to by Strahler (1964) is not one of age, but one of tectonic uplift history. The mature stage, concave down at low elevations and concave up for high elevations reflects a landscape in dynamic equilibrium with the tectonic uplift regime. The Monadnock stage, concave up for all elevations reflects a declining catchment subjected to episodic uplift events. This latter type of catchment is not in dynamic equilibrium.

This differentiation of catchments into event uplift versus continuous uplift may be complicated by other interactions in the field. In particular the relative rates of change of tectonic uplift and the rate of erosion are important. A catchment with a short timescale of elevation change due to high erosion rates (i.e., high TS, Chapter 6) will go to dynamic equilibrium faster than one with long timescale and low erosion rates (i.e., low TS). The catchment with high TS will react more quickly to changes in the tectonic environment. Thus for the same uplift history two catchments different only in their TS value may exhibit different hypsometric curves. The catchment with high TS will be closer to dynamic equilibrium than the catchment with low TS; the catchment with low TS will have an event uplift hypsometric curve, the catchment with high TS will have one more like the dynamic equilibrium case.

7.7 Sensitivity Study of Diffusive Hillslope Processes

To assess the effect of the diffusive term in Equation (5.1a) on hillslope and network characteristics, a sensitivity test on the parameter D_z was performed. This parameter controls the magnitude of the Fickian diffusion in the elevation evolution

equation and was designed to model the diffusive transport processes described in Section 2.3.3 (e.g. creep, rainsplash, and landslide). As noted in Appendix B fluvial sediment transport cannot be modelled by a Fickian diffusion term so it is modelled separately.

A value was adopted for D_z that resulted in about 10% of the catchment being dominated by diffusive effects. This value of D_z was determined by trial and error, the value adopted is given in Appendix D for the simulation CR10–1. A node in the grid was considered to be dominated by diffusion effects if the rate of elevation change due to the diffusion term was greater than the rate of change of elevation due to fluvial sediment transport. That is, from Equation (5.1a), diffusion domination occurs at node j when

$$\left| D_z \frac{\partial^2 z_j}{\partial x_i^2} \right| > \left| \frac{1}{\rho_s(1-n)} \sum_i [Q_i^{m_1} S_i^{n_1} I_{ij} f(Y_j)] \right|$$

Two comparable runs, different only in the magnitude of the D_z, are shown in Figure 7.33. As expected following the discussions of Section 7.2, the simulated networks are quite different. Any small perturbation in elevations results in self–reinforcing changes in the network realized by the simulation. The diffusion, though a minor component spatially, provides the small perturbations necessary to change the network generated.

Figure (7.33c) shows the spatial region over which diffusion dominates sediment transport when the network stopped growing. These regions mostly lie on or near the watersheds where discharge, and thus fluvial sediment transport, is low. This consistent with the observation that rainsplash dominates at the watersheds (Dunne, 1988b) and that proceeding downslope rainsplash is dominated by fluvial transport (Dunne and Aubrey, 1986).

$$TD = 0 \qquad TD = 2.0 \times 10^{-8}$$
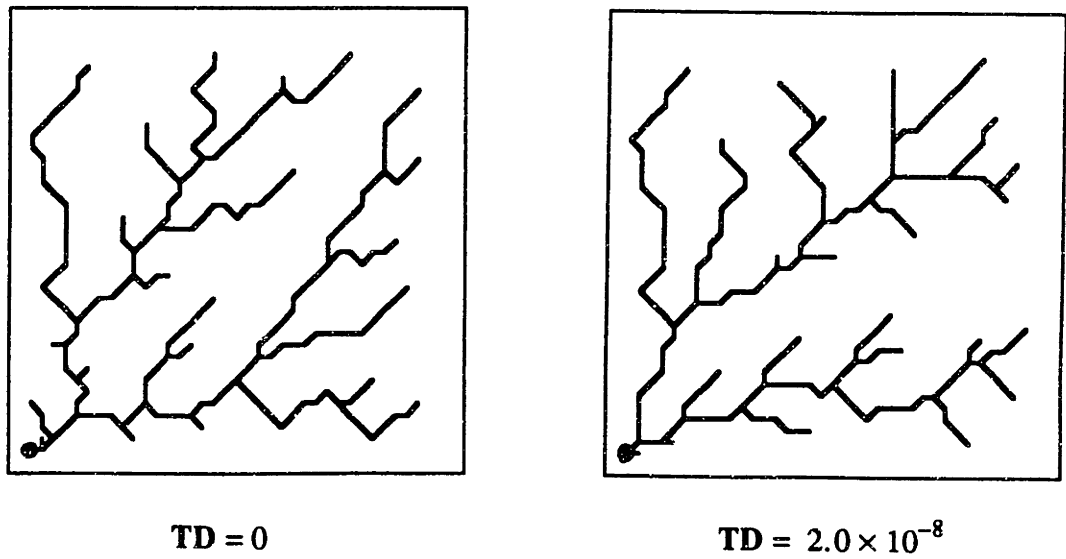
(a) Channel networks with and without diffusion

● Catchment outlet
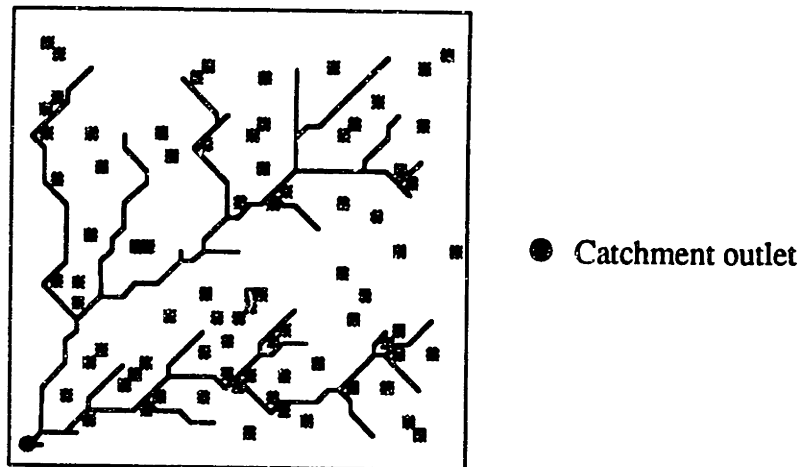
(b) Regions of diffusion domination

Figure 7.33: Sensitivity of the channel network to diffusive transport

Hypsometric curves are plotted for the diffusive and nondiffusive cases in Figure 7.34. There is very little difference between the form of the hypsometric curves reflecting the small aggregate effect of diffusion on the elevations. The horizontal position of the kink in the hypsometric curve, which Section 7.6 showed was indicative of the drainage density of the catchment, varies between the two runs. The diffusive run has a slightly faster rate of network growth than the nondiffusive case so that the kink moves faster from right to left. A plot of the drainage density with time (Figure 7.35) corroborates this. Since the area under the hypsometric curve (the hypsometric integral) is indicative of the mean elevation of the catchment above the outlet, the elevations in the diffusive network declined slightly faster at early times. This is reflected in a slightly higher sediment transport out of the catchment (about 20% higher). At the time when the network stopped growing the hypsometric curves are very similar. The reason for the faster extension of the diffusive network was the increased overland transport of the added diffusive transport compared to fluvial transport alone in the nondiffusive case.

Some network and hillslope statistics are provided in Table 7.6 for both the diffusive and nondiffusive cases, for the time when the network stopped growing. The topological statistics, as expected, are greatly different, due to the differences in the generated networks. Otherwise the effects of the diffusion term are small.

In conclusion, the diffusion term has been shown to have a relatively minor effect on catchment form. As expected the channel network simulated was significantly changed because of the chaotic network growth mechanism. Topological statistics reflected these differences. Other statistics were less affected by the diffusion effects. The spatial distribution of diffusion domination within the catchment was qualitatively consistent with field observations.
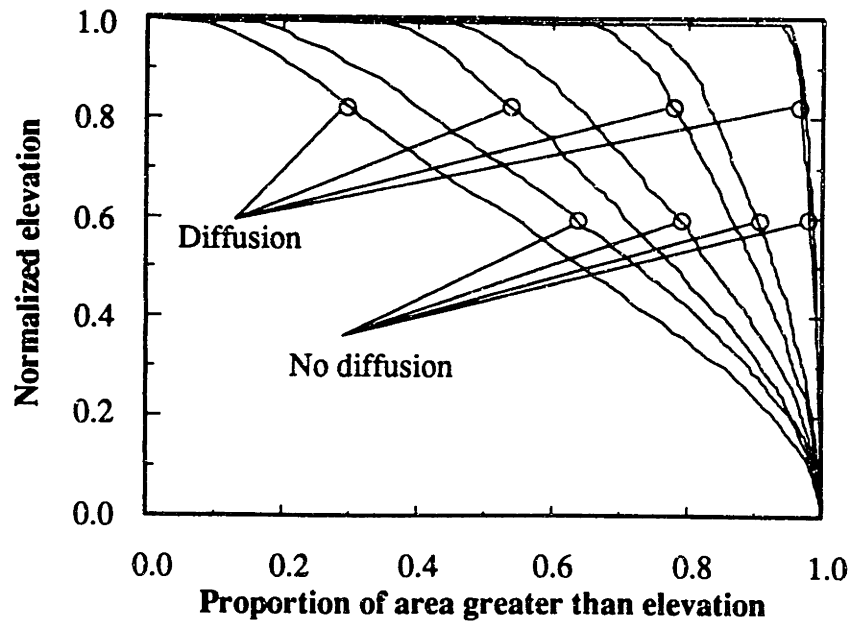
Figure 7.34: Hypsometric Curve: with and without hillslope diffusion, simulations CR2-3, CR10-1
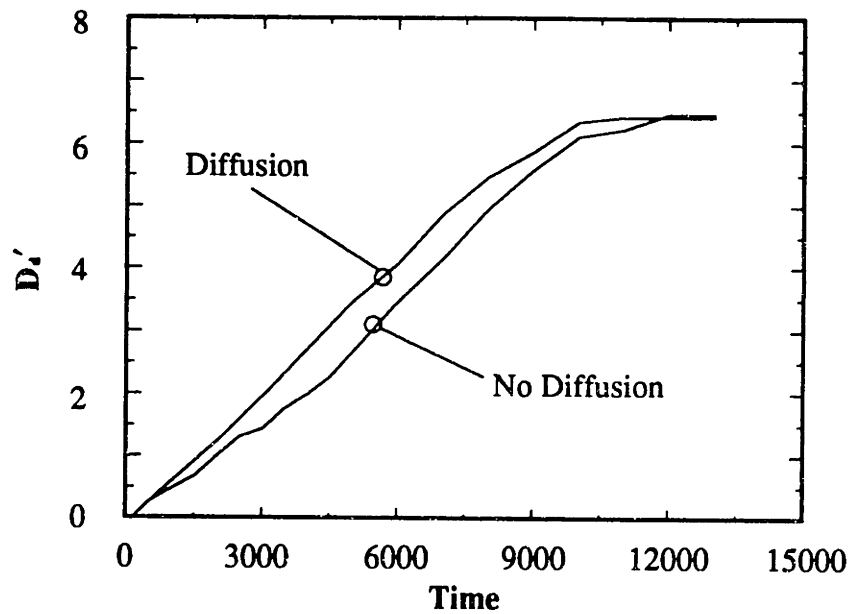
Figure 7.35: Drainage density with time: with and without hillslope diffusion, simulations CR2-3, CR10-1

TABLE 7.6

Summary Statistics for Networks Based on Diffusive and Nondiffusive Hillslope Processes

Statistic	$TD = 0$ [1]	$TD = 2 \times 10^{-8}$ [2]
R_b	5.50	4.60
R_L	3.00	1.43
R_S	1.76	1.60
R_A	7.02	3.28
K	1.78	3.29
ϵ	2.25	0.37
D'_d	6.80	6.44
R_c	9.90	9.90
$R_h^{(3)}$	4.18	4.05
$S_{he}^{(4)}$	1.63	1.63

(1) Results from simulation CR2–3, $t' = 12000$

(2) Results from simulation CR10–1, $t' = 12000$

(3) R_h = mean hillslope drop,
exterior links

(4) S_{he} = mean hillslope slope,
exterior links

CHAPTER 8

EXPERIMENTAL VERIFICATION

8.1 Introduction

This chapter discusses aspects of the experimental verification of the catchment evolution model presented in this work. At the outset it is noted that the catchment evolution model is sufficiently novel that there is no preexisting data set capable of a comprehensive verification of the model. The published data may, however, be used to verify some of the components of the model; the presentation of this chapter is governed by this pragmatic constraint.

If the data does not exist to completely verify the model, it is important to identify, and verify, the more speculative components of the model. Below is a short summary of some of these aspects of the model. This is not to say that this chapter will provide adequate verification of all of these components; it won't. Rather the intention is to highlight the aspects of the model that require experimental verification, whether in this chapter or in later work.

The sediment transport formulas in the elevation equation are as good as can be expected given the state of the art in fluvial sediment transport. The formulation, nonlinearly dependent on discharge and slope, is consistent with accepted instantaneous transport for rivers and hillslopes. The conceptual step from the instantaneous sediment transport to the long–term mean sediment transport involving flood frequencies and the mean annual discharge is more speculative, even though it is commonly used by theoretical geomorphologists. The postulate of different sediment transport rates in the channel and in the hillslopes (viz. the O_t factor in Equation 5.1) is consistent with some data obtained at short timescales. There is little experimental evidence supporting it at long time scales, though it follows theoretically from short

time scale behavior.

The activator concept, and the formulation adopted, nonlinearly dependent on discharge and slope, is particularly novel. Recent data apparently supporting the formulation will be examined in Section 8.4.

The channel head growth mechanisms in the model, apart from the use of the activator function, have been conceptualized into the autocatalytic differentiation equation and the timescale of channelization $1/d_t$. Once the activator threshold is exceeded channelization proceeds irrespective of conditions within the catchment. Very little work has been reported on the long term dynamics of channel growth and dependence of growth rates on physical properties. This lack of knowledge implies that we are unable to say with any certainty, without extensive site specific field work, what are the important physical processes in determining the timescale $1/d_t$.

In the model, channels are conceptualized as distinct entities with well defined channel heads. For gully erosion this seems a reasonable approximation, but others suggest that this viewpoint is not general enough (Kirkby, 1988). Since, in this model, the channel network is only used to differentiate rates of sediment transport, the issue is <u>not</u> whether channels are distinct or not, it <u>is</u> whether there is a distinct point within the catchment (e.g., the channel head) where the mean rate of sediment transport changes abruptly for the timescales of landscape formation.

Finally in both the activator and sediment transport formulas, the definition of discharge adopted is the same. For the long term average equation this assumption needs verification since the discharge adopted is the mean annual discharge. In any experimental verification, the determination of the dominant runoff mechanism is a central problem. Though no results have been presented here, the computer model is able to model both Horton and subsurface saturation runoff, the latter by use of a saturation threshold, (e.g. Beven and Kirkby, 1979; O Loughlin, 1981). For subsurface saturation the important question is then what are the dominant mass movement

mechanisms in that part of the catchment that is unsaturated on the hillslopes?

The list of the components of the model that need verification is daunting. Many, however, go to the very heart of understanding the hillslope–channel interactions that occur over geologic time. This chapter will not answer all the questions, but confine itself to three experimental results that have been reported in the literature.

The first work that will be examined is the experimental small scale catchment simulation studies carried out by Parker and Mosley and reported in Schumm, et al. (1987). These data are the best reported, and best controlled, experiments related to catchment evolution. Scale effects in the sediment transport processes preclude the use of a great deal of the data obtained, and from our perspective many interesting variables were not measured. Nevertheless, it will be shown that the simulation model exhibits both qualitative and quantitative similarities to the experimental data.

The second work that will be examined is work empirically relating slopes and contributing areas (Flint, 1974; Tarboton, et al., 1989). It is shown that the observed relationship can be derived from the simulation model analytically and that this result corresponds well with that determined from networks based on digital elevation data.

The third, and final, work that will be examined is an observed relationship between channel head contributing area and local hillslope slope (Montgomery and Dietrich, 1988). It is shown that the data is consistent with the adopted activator formulation, and that networks simulated using an activator equation fitted to this data are realistic.

8.2 Plot Scale Catchment Evolution Experiments

Schumm, et. al. (1987) presents a comprehensive summary of work performed at the Colorado State University in 1970's and 1980's using their experimental rainfall erosion facility. The range of experimental work carried out was far reaching and

touched upon almost all aspects of fluvial geomorphology. The particular work that is of importance here is that described in Parker (1977). Parker simulated the growth of channel networks using simulated rainfall and uniform sloping initial conditions on elevations. Initial slopes were varied over almost an order of magnitude.

The following data from Parker (1977) will be discussed:

1.   Planar network form with time, e.g., magnitude, Strahler ratios.

2.   Length scales with time, e.g., drainage density, exterior link lengths. .

3.   Sediment yields with time.

4.   Variation of drainage density with catchment slope.

The first of these points, planar network form with time, and in particular Strahler ratios, was discussed in Section 7.2 and will not be further discussed here.

Before proceeding, however, it is important to list the limitations of the data reported by Parker. This will provide a perspective on the verification of the simulation model that follows. There are some problems in interpreting Parker's data.

The first problem concerns the hillslope erosion processes. Schumm, et. al. (1987) notes that the hillslope erosion processes in Parker's experiment were dominated by rainsplash effects. This rainsplash dominance arises from the small horizontal scale of the experimental catchments. Short hillslopes have low discharges and small fluvial erosion terms. As an example of the use of the nondimensional of the simulation model, Chapter 6 showed that for typical catchment scales diffusive transport from rainsplash is about 10^{11} times more important relative to fluvial transport on the hillslopes of Parker's basin compared to typical field scale catchments. In the introduction of this chapter, the ratio of the hillslope and channel erosion rates, the O_t factor, was identified as one component requiring verification. At the field scale, and in the model, it is postulated that the same, or similar erosive processes dominate on both the hillslopes and the channels; with an unconfirmed component being the ratio of the erosion rates. In Parker's experiment the erosive processes

acting on the hillslope and in the channel are different: in the channel it is fluvial erosion; on the hillslopes it is diffusive rainsplash. These two processes scale differently with both discharge and slope so the relative magnitudes of erosion from Parker's experiment are not applicable to the field scale.

The second point concerns the rate of growth of the network. The processes that govern the rate of advance of the channel head are not well understood (Dunne, 1989). In the model these processes have been conceptualized, and the growth rate of a channel at a point parameterized by d_t. If d_t is increased, the channel grows faster. Conversely, if d_t is decreased, the channel grows more slowly. In addition, there are erosion forces acting at the channel head, like undercutting, landsliding, and sediment transport, that are not considered explicitly in the model except through the activator function. The activator function provides a threshold above which channelization occurs, it does not explicitly give the rate of growth of the channels. How the growth rate scales with area, discharge and slope is unknown. It is also probable that the processes that are dominant at the experimental plot scale are not dominant at the field scale. Thus the activator function and parameter d_t for Parker's experiment may not be applicable at the field scale.

A third problem is that Parker's data does not allow verification of the stochastic component of the model because he used a uniform rainfall rate with time. Thus the dependence of the sediment transport rate on mean annual peak discharge cannot be tested.

Despite difficulties in the interpretation of the Parker data, some important qualitative characteristics of his data are observed in the results from the simulation model. One qualitative characteristic showed by both simulation and experimental data is the scatter in the lengths of TS and S type links (see Section 2.2.3 for the definition of link types) during the growth process. Parker observed that the lengths of the S links displayed a greater variance than the length of the TS links and that the TS

323

link length distribution did not vary greatly with time. Similar characteristics were not observed in the simulations (Figure 8.1).

Parker noted that the sediment yield from the experimental catchments initially rose sharply to a peak then declined slowly with time. He found that the decline with time could be fitted with a curve of the form

$$Y = \alpha \, V^\gamma \qquad\qquad\qquad (8.1)$$

where   Y   = sediment yield in mass/time

V   = volume of water, a surrogate for time (V = RT)

α, γ = fitting coefficients

= 850, −0.86 for Experiment 1

= 78, −0.15 for Experiment 2.

The difference between Experiments 1 and 2 was that in Experiment 1 a notch was applied at the outlet so that network growth occurred primarily by notch advance. In Experiment 2 no initial notch was applied so that the network did not grow so fast nor so distinctly, initially. The slower decline of sediment yield in the second experiment was because "initial sediment yields were not high, but continued incision of the channel and an increase in sediment transport efficiency maintained higher sediment yields for a longer period during basin evolution" (Parker, 1977, p. 29).

For the simulated catchments a similar power law decline of sediment yield was observed. Figure 8.2 shows typical plots of sediment yield versus time for the simulated catchments. Power laws have been fitted to the sediment yield decline, and the fits, which are quite good individually are shown in the figures. The power of the sediment yield decline, γ, is in the range of −1.2 to −1.8. These values are about 50% to 100% higher than the single value for Experiment 1 reported by Parker. Despite this the results of the simulations are considered consistent with Parker's for two reasons.
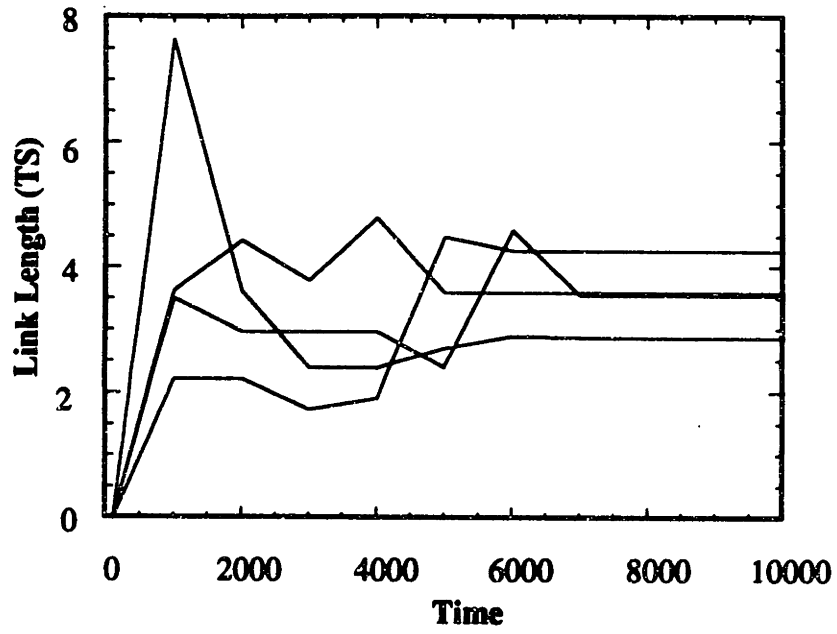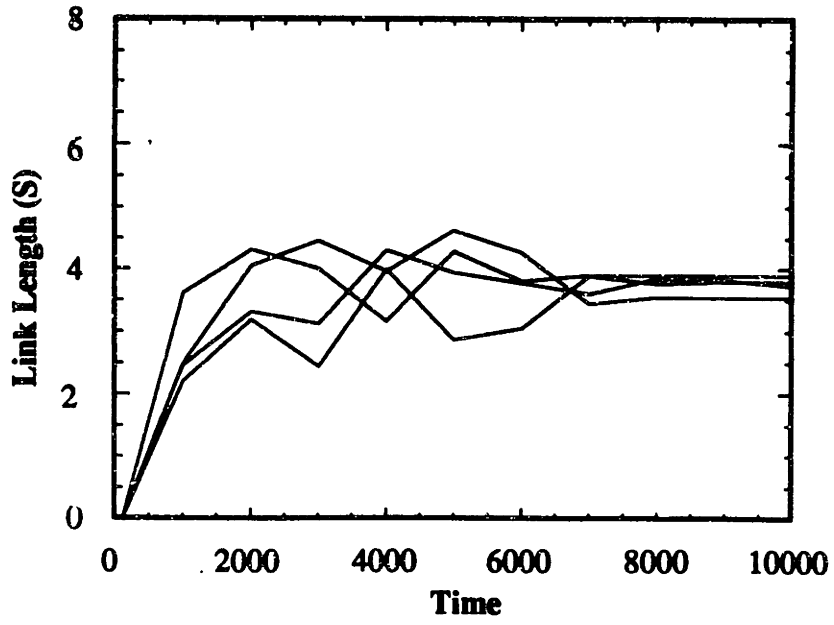
Figure 8.1: Exterior link lengths with time, stratified on topological classification, simulations CR2-6, CR7-1, CR7-2, CR7-3
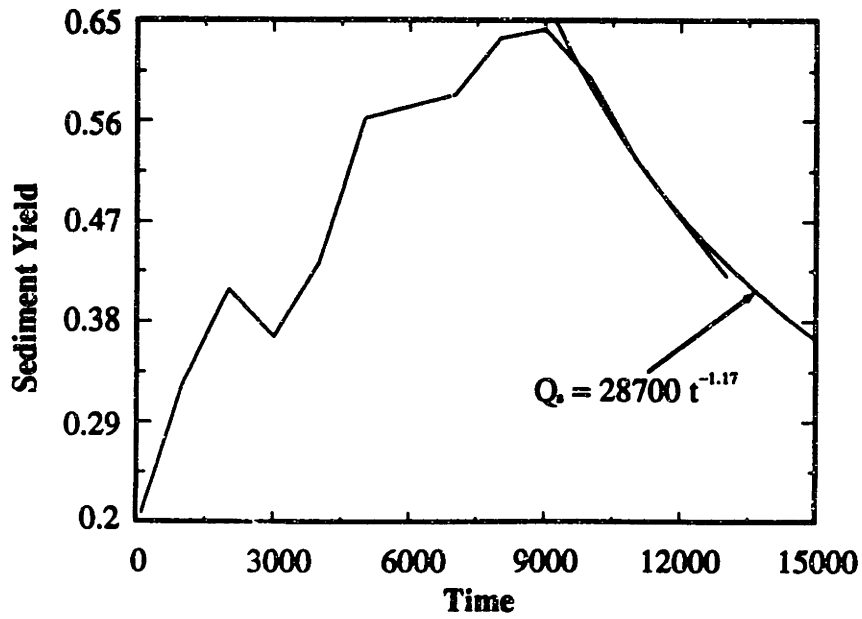
(a) CR2-3
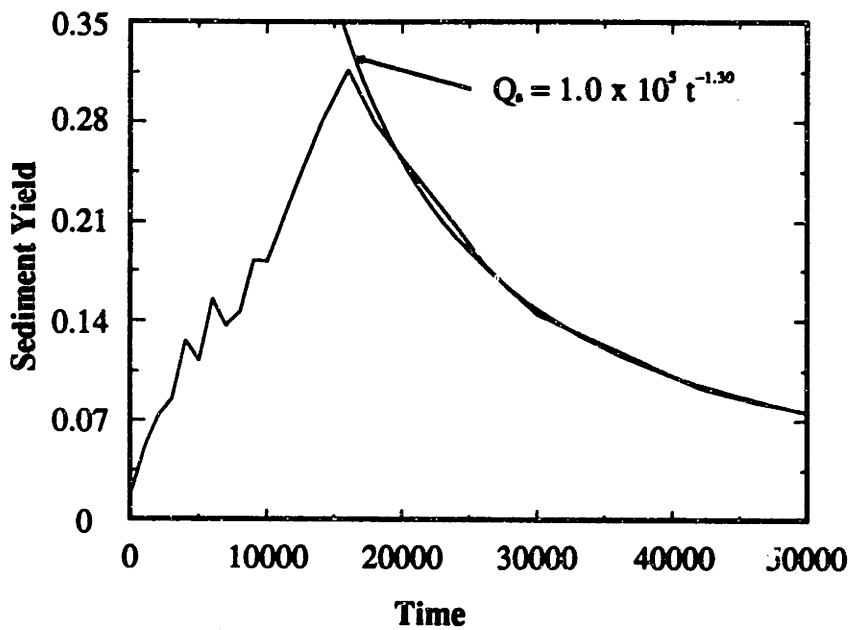
(b) CR3-4

Figure 8.2: Sediment Yield with time, typical results

326

$$Q_s = 1.63 \times 10^7 \, t^{-1.12}$$

Time

(c) CR3-5

$$Q_s = 5180 \, t^{-1.02}$$

Time

(d) CR5-2
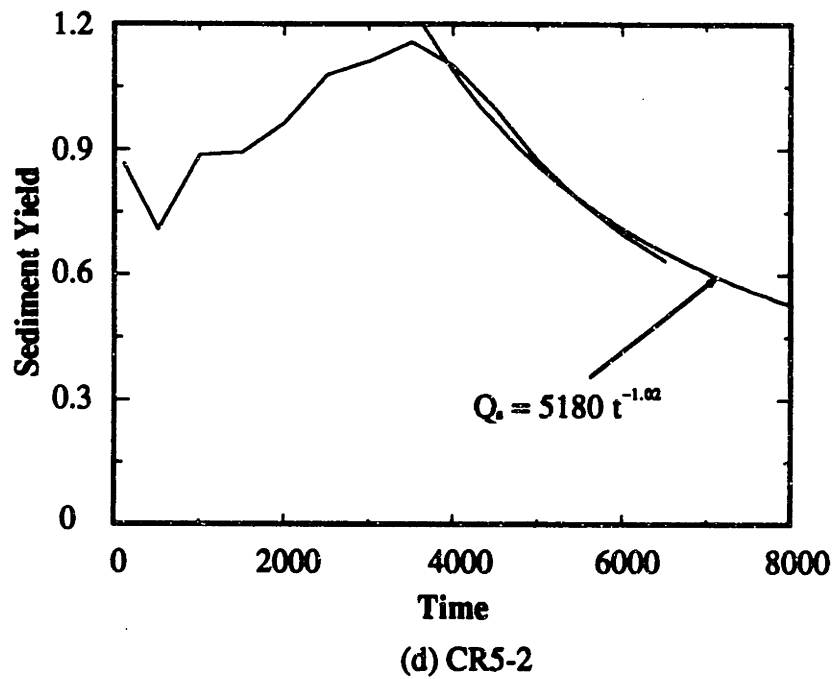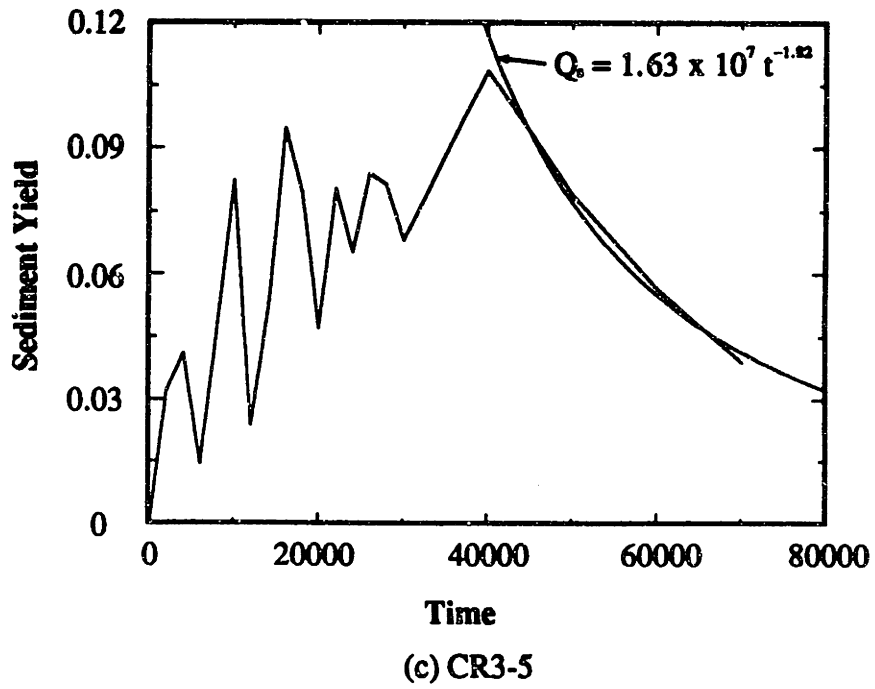
Figure 8.2: (ctd) Sediment yield with time, typical results

1. The values for α and γ, reported by Schumm for Experiment 1 given above, were for a single experiment, with no indication of their repeatability. The variation of γ from -0.86 to -0.15 simply by removing the notch suggests that they are sensitive to experimental fluctuations. This is supported by Schumm who notes that slumps and slope failures result in the yield "fluctuating dramatically."

2. Parker plots the yield versus time for different catchment reliefs (what are effectively 6 related experiments) and it appears that the rate of decline of sediment yield, γ, increases with catchment relief (Figure 6.2 in Parker, 1977). The reported value for γ, of -0.86, is the value for the catchment with the lowest relief. Despite considerable scatter of the data it appears that γ could be up to twice as large (i.e., $\gamma \approx -1.6$) for the higher relief basins (Figure 8.3).

The final result of Schumm, et. al. (1987) that will be compared with the simulation results is a regression between the drainage density and the mean catchment slope obtained from the experimental data. A relationship for drainage density based on catchment horizontal and vertical length scales was presented in Section 7.3. It will be shown that the trend of drainage density with slope for the experimental data is similar to the trend of the drainage density equation based on the simulation data.

Schumm presented the relationship between drainage density and catchment slope as

$$D_d = 0.91 + 22.4\,S \tag{8.2}$$

where S = (initial catchment relief from notch to watershed)/(length of the experimental basin)
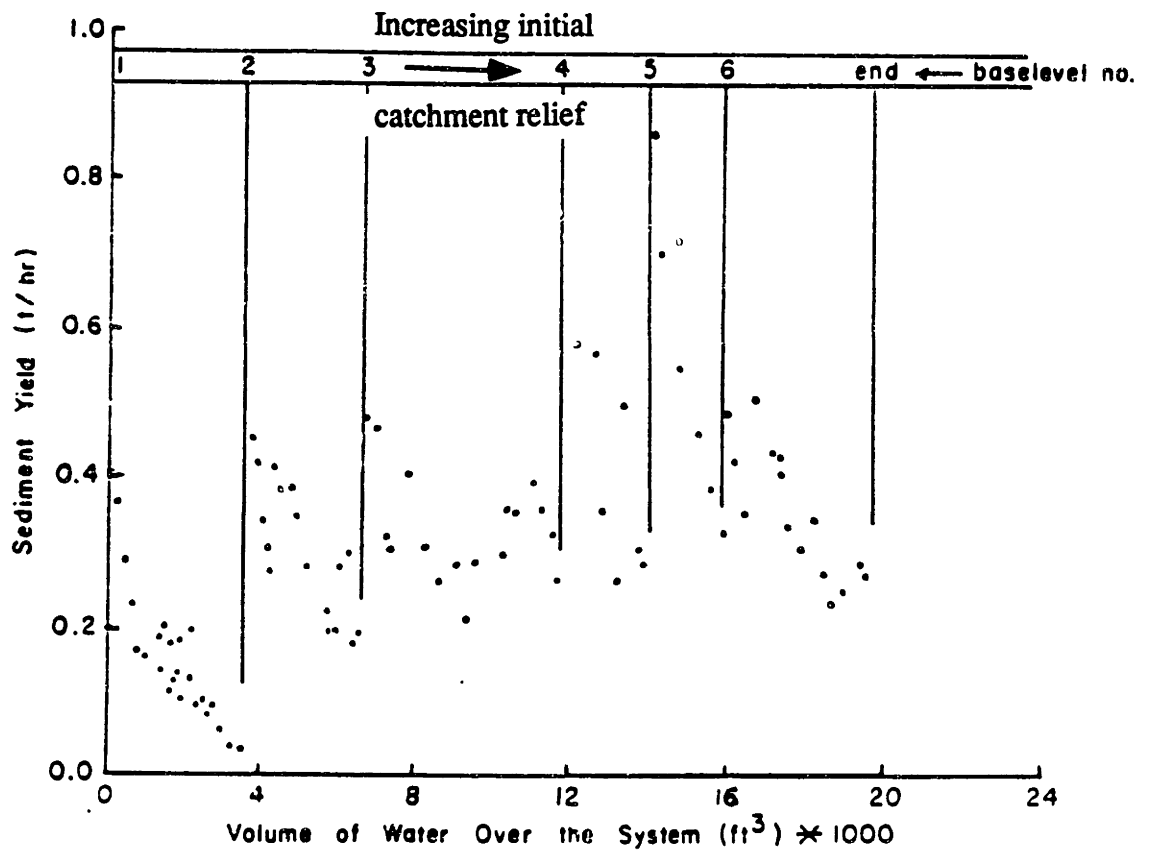
Figure 8.3: Observed sediment yield with time
(from Parker, 1977)

The failing of this expression is that it predicts a nonzero drainage density when the catchment slope is zero. The data was replotted on a log–log scale and a power law expression obtained

$$D_d = 13.1 \ S^{0.62} \tag{8.3}$$

The fit of this latter expression is as good as the original and has the added benefit of predicting zero drainage density for zero slope. The data points and the new regressions are plotted in Figure 8.4

Using the simulation data a general expression for drainage density was obtained (Equation 7.9). The difficulty of this expression is that it has a threshold on TA_h and is not of the form of Equation (8.3) complicating the verification of the simulation data. An alternative expression for the drainage density of the simulation data was developed. This equation is

$$D_d' = 5 \times 10^3 \ O_t^{-0.11} \ TA_c^{1.48} \tag{8.4}$$

This equation is applicable for:

1. $TA_c \gg TA_h$ (i.e., drainage density is high).

2. $m_5 = 0.4$, $n_5 = 0.3$ (i.e., activator is overland flow velocity).

Equation (7.9) is more general than Equation (8.4), but in fact fits the appropriate simulation data little better than the above equation.

The catchment slope can be approximated, for the purpose of the scaling analysis, by $S = L_e/L_d$, using the catchment vertical length scale, L_e, and the catchment horizontal length scale, L_d. In Parker's study the catchment area was fixed, being the area of the experimental basin. Consequently the catchment horizontal length scale, L_d, is fixed so that changes in slopes are equivalent to changes
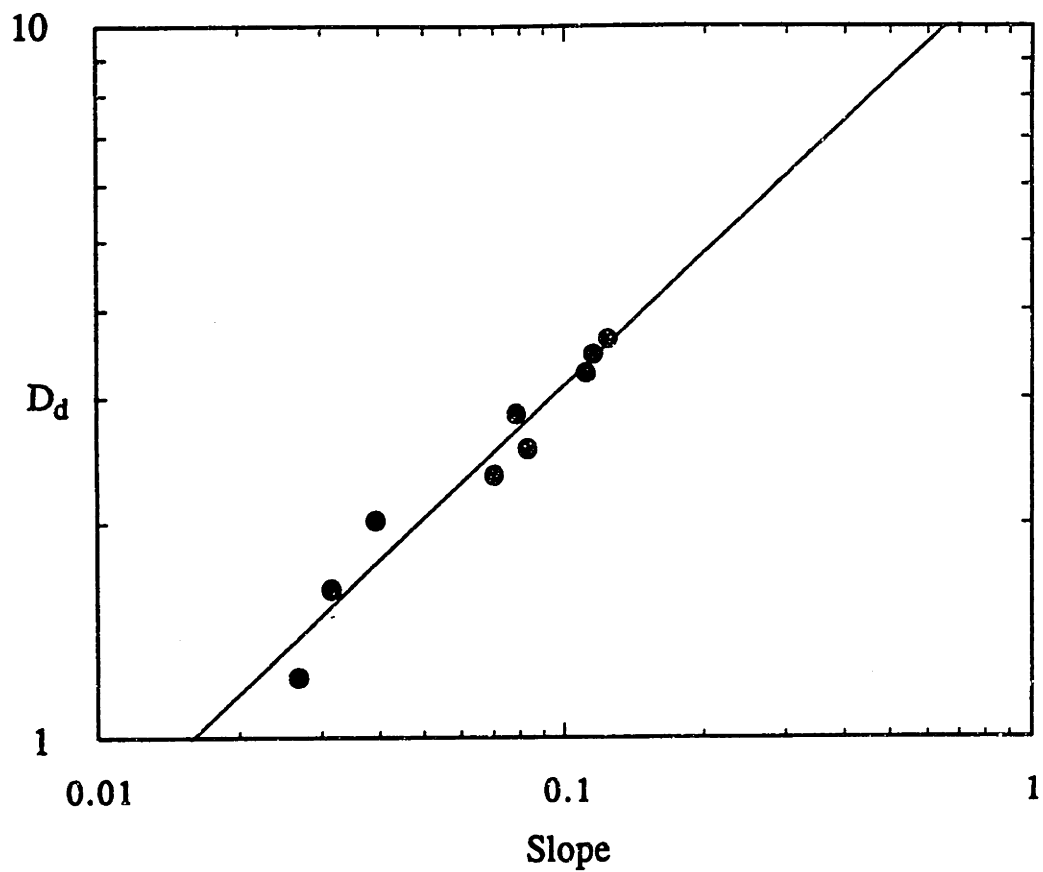
Figure 8.4: Drainage density versus catchment slope:
Schumm (1987) data

in the catchment vertical length scale, L_e.

From Equation (8.3) the trend of the experimental data is

$$D_d \sim S^{0.62} \sim L_e^{0.62} \tag{8.5}$$

Similarly, from Equation (8.4) the trend of the drainage density from simulation data is

$$D_d \sim O_t^{-0.11} TA_c^{1.48} \sim O_t^{-0.11} L_e^{0.44} \tag{8.6}$$

The O_t factor complicates the interpretation of the simulation data. The result in Equation (8.4) is for the case where both hillslope and channel erosion are dominated by fluvial sediment transport and

$$Q_{s,channel} = \beta_1 Q^{m_1} S^{n_1} \tag{8.7}$$

$$Q_{s,hillslope} = \beta_1 O_t Q^{m_1} S^{n_1} \tag{8.8}$$

where $m_1 = 1.8$, $n_1 = 2.1$ for Einstein–Brown sediment transport and

$$O_t = \frac{Q_{s,hillslope}}{Q_{s,channel}} \tag{8.9}$$

The factor O_t is the ratio of the hillslope and channel erosion rates. The experimental catchments of Parker had hillslopes dominated by rainsplash so that the hillslope sediment transport rate is (Equation 2.21)

$$Q_{s,hillslope} = D_z \frac{\partial S}{\partial x} \qquad (8.10)$$

We may interpret O_t in its most general sense as just the ratio of transport rates on the hillslope and the channel so that for the experimental data

$$O_t = \frac{Q_{s,hillslope}}{Q_{s,channel}} = \left[\frac{D_z}{\beta_1 Q^{m_1}} \right] \left[\frac{\frac{\partial S}{\partial x}}{S^{n_1}} \right] \qquad (8.11)$$

If the slope terms in this equation are replaced by the catchment vertical scale, as in Equations (8.5) and (8.6), then

$$O_t \sim \frac{\frac{\partial S}{\partial x}}{S^{n_1}} = L_e^{1-n_1} = L_e^{-1.1} \qquad (8.12)$$

where $n_1 = 2.1$

Thus as the catchment becomes steeper (i.e., L_e increases) the ratio of the sediment transport rate on the hillslope to that in the channels decreases, i.e., the difference between them becomes larger. Variation of hillslope form with slope is not addressed by this analysis though it is not expected that any variations will be dominant.

If Equation (8.12) is substituted into Equation (8.6), then

$$D_d \sim L_e^{0.56} \qquad (8.13)$$

Comparing Equation (8.5), from the experimental data, and Equation (8.13), from the simulation data, the agreement of the trends in drainage density with catchment vertical scale, and thus mean slope of the catchment, is good. The differences of the

powers of slope are within experimental error as can be seen in Figure 8.4 where an equation of the form of Equation (8.13) is plotted.

8.3 Renormalization of Contributing Areas, Magnitudes and Slopes

Tarboton, et al. (1989) have found that channel networks derived from digital elevation data show a relationship between area and mean slope in the channels of the form

$$A^{\alpha}S = \text{constant} \tag{8.14}$$

where   A   = contributing area to the downstream end of the link

S   = mean slope of the link

α = a renormalization coefficient in the range of 0.4 to 0.7 (Tarboton, unpublished data).

Using the catchment evolution model developed herein, it is possible to derive this relationship for a catchment in dynamic equilibrium. This relationship relates the tectonic uplift regime of the catchment with the erosional processes occurring in the channels.

For a catchment in dynamic equilibrium the tectonic uplift for any catchment, or part thereof, equals the average erosional loss rate over that catchment. Thus, by continuity in the channel at the outlet of the catchment, we may say

$$c_0 A = Q_{s,\text{outlet}} = \beta_1 Q^{m_1} S^{n_1} \tag{8.15}$$

where c_0 = tectonic uplift rate

A   = catchment area

$$Q \qquad = \text{channel discharge at the outlet}$$

$$S \qquad = \text{channel slope at the outlet}$$

A commonly observed relationship between discharge and area (e.g. Leopold, et al., 1964) is

$$Q = \alpha_1 A^{\alpha_2} \tag{8.16}$$

Substituting this relationship into Equation (8.15) and rearranging yields

$$A^{\frac{\alpha_2 m_1 - 1}{n_1}} S = \left[\frac{1}{\alpha_1}\right]^{\frac{m_1}{n_1}} \left[\frac{c_0}{\beta_1}\right]^{\frac{1}{n_1}} \tag{8.17}$$

Thus the predicted renormalization coefficient, α_p, is

$$\alpha_p = \frac{\alpha_2 m_1 - 1}{n_1} \tag{8.18}$$

To verify this relationship the renormalization of Equation (8.14) was fitted to the catchment simulation of CR8–1, a simulation which had been allowed to proceed to dynamic equilibrium. Equation (8.17) is applicable to the instantaneous areas and slopes. A regression on these areas and slopes is shown in Figure 8.5. The fit is very satisfactory ($r > 0.99$). The fitted renormalization coefficient is $\alpha = 0.374$. Using the parameters of run CR8–1 ($\alpha_2 = 1.0$, $m_1 = 1.8$, $n_1 = 2.1$), the predicted renormalization coefficient is $\alpha_p = 0.381$, a very good agreement with the fitted value.

Tarboton, et al. found a value of the renormalization coefficient of $\alpha = 0.47$ using mean link slopes and the total contributing area to the bottom of the link. Further data on the sediment transport equation and flood frequencies for their
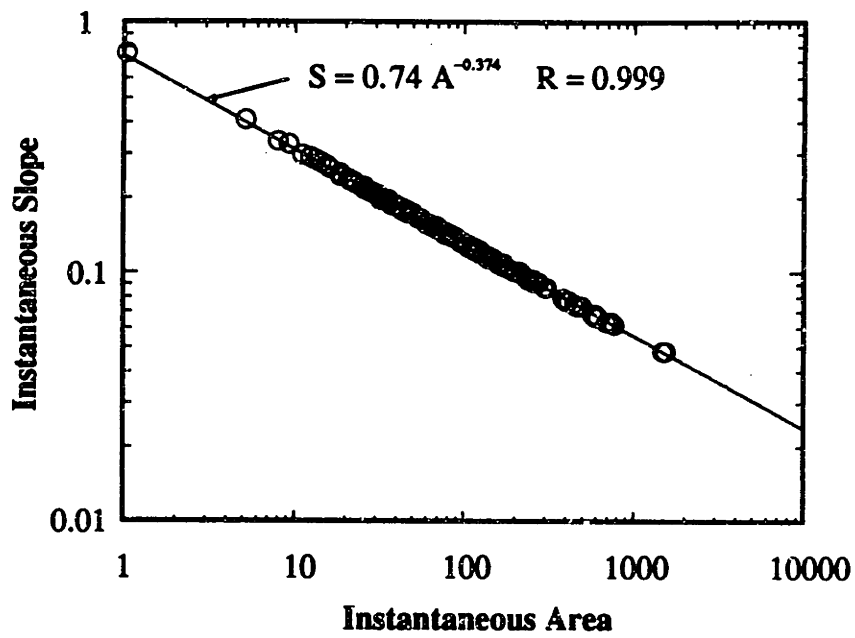
Figure 8.5 : Correlation between slopes and areas
in channels at dynamic equilibrium:
verification of theoretical expression.

catchment is required to verify the fitted renormalization. Nevertheless, the observed renormalization is within the bounds of the predicted renormalization coefficient (Equation 8.8). If, for instance, the same Einstein–Brown sediment transport equation is used, but with p = 3.5 (instead of p = 3), then with $\alpha_2 = 1$, the predicted renormalization coefficient is $\alpha_p = 0.45$. Thus the renormalization coefficient, α, is fairly dependent on the governing sediment transport equations for the catchment. If the flood frequency distribution with area for the catchment is known, so that α_2 may be estimated, then α is potentially an estimator of the exponents in the sediment transport equation.

It should be emphasized that there exists considerable variability in the fitted renormalization coefficient depending on how it is estimated. In addition to the renormalization relationship with area (Equation 8.14) Tarboton, et al. fitted relationships of the form

$$(2m-1)^{\alpha'} S = \text{constant} \tag{8.19}$$

$$m^{\alpha''} S = \text{constant} \tag{8.20}$$

where   m   = magnitude

α', α'' = renormalization coefficients for (2m−1) and m, respectively

Flint (1974) fitted Equation (8.19) to 11 catchments and found α' to be in the range 0.37 to 0.83 with correlation coefficients of around 0.9. Equations (8.14), (8.19) and (8.20) were fitted to the simulation CR8–1, using definitions of area and slope consistent with Tarboton to yield the coefficients $\alpha = 0.33$, $\alpha' = 0.42$, $\alpha'' = 0.53$ with correlation coefficients of r = 0.99, 0.86, 0.85 respectively. Thus there is considerable variability in the renormalization coefficient depending on the variables used in the

renormalization.

A commonly accepted relationship for area and magnitude is

$$E[A] = (2m-1) \, E[A_\ell]$$ (8.21)

where $E[\cdot]$ = expectation operator

A_ℓ = area draining laterally to a link, assumed the same distribution for both exterior and interior links.

It was shown in Chapter 7 that there is a significant difference between the area draining to an exterior link and the area draining to an interior link. More correctly then, the mean area should be expressed as

$$E[A] = m \, E[A_e] + (m-1) \, E[A_i]$$ (8.22)

where A_e, A_i = area draining laterally to exterior and interior links respectively.

Equation (8.19) is equivalent to Equation (8.14) only if Equation (8.21) is true. As noted, however, Equation (8.22) is more satisfactory. The effect of the difference in exterior areas is clear in Figure 8.6 where $(2m-1)$ and area are plotted versus link slope. For the regression of $(2m-1)$ versus slope the difference in the fitted renormalization coefficient is solely in the magnitude 1 links (i.e., exterior links), where slopes are anomalously high. If a regression is fitted to $(2m-1)$ and slope, ignoring the exterior links, then the fitted renormalization coefficient is closer to that based on area, α. This is consistent with Equation (8.22).

The extra area draining to exterior links is believed to be one possible reason

338
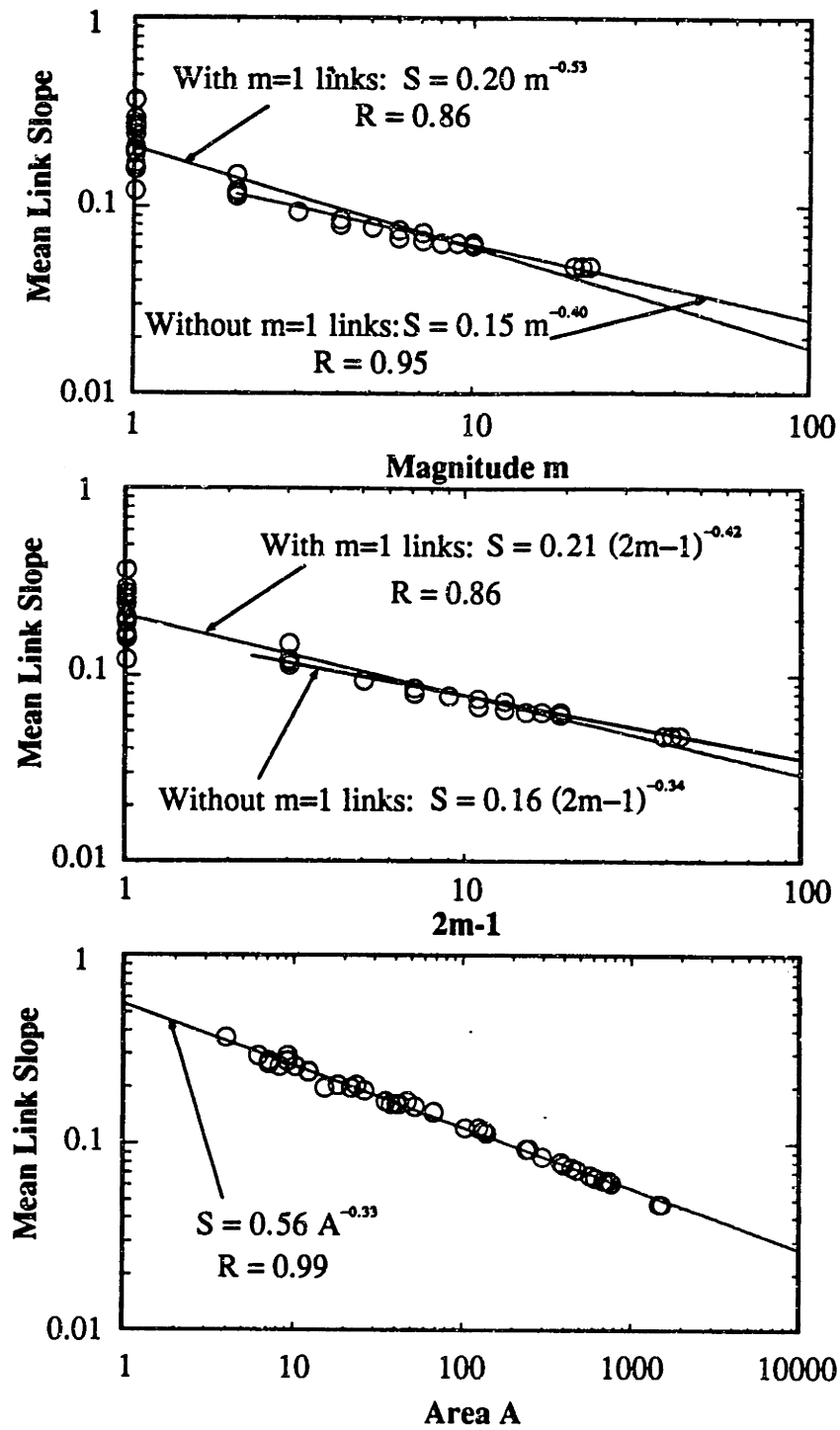
Figure 8.6: Mean link slope renormalization based on link magnitude and link area.

why Tarboton, et al. (1989) observed considerable scatter in the fitted renormalization coefficient depending on whether α, α' or α'' is fitted to the data. The trend of $\alpha < \alpha' < \alpha''$ in the simulated catchments is consistent with that for their field data

It is also possible to fit the renormalization to the hillslopes. Equation (8.15) is satisfied in the channel flowing out of the catchment, irrespective of the erosion processes within the catchment. It can, however, be applied with equal validity to the scales less than the hillslope scale. The sediment outflow is that corresponding to transport on the hillslope, so that

$$c_0 A = \beta_1 \, O_t \, Q^{m_1} S^{n_1} \tag{8.23}$$

where $\quad A \quad =$ contributing area of hillslope

$\quad\quad\quad\quad O_t \quad =$ hillslope channel erosion rate factor

and the renormalization equation for the hillslope becomes

$$A^{\frac{\alpha_2 m_1 - 1}{n_1}} S = \left[\frac{1}{\alpha_1}\right]^{\frac{m_1}{n_1}} \left[\frac{c_0}{\beta_1 O_t}\right]^{\frac{1}{n_1}} \tag{8.24}$$

where this equation is derived in a similar fashion to Equation (8.17). If hillslope erosion is governed by the same physics as the channel, then the renormalization coefficient for the hillslope and the channel will be the same. The only difference will be the value of the constant on the right–hand side.

Figure 8.7 plots the instantaneous slopes and areas for run CR8–1 for the whole catchment, when at dynamic equilibrium. For comparison, two common measures of hillslope scale are indicated in this plot. The first is the mean Strahler first order area. The second is the mean source area. These length scales are common measures of
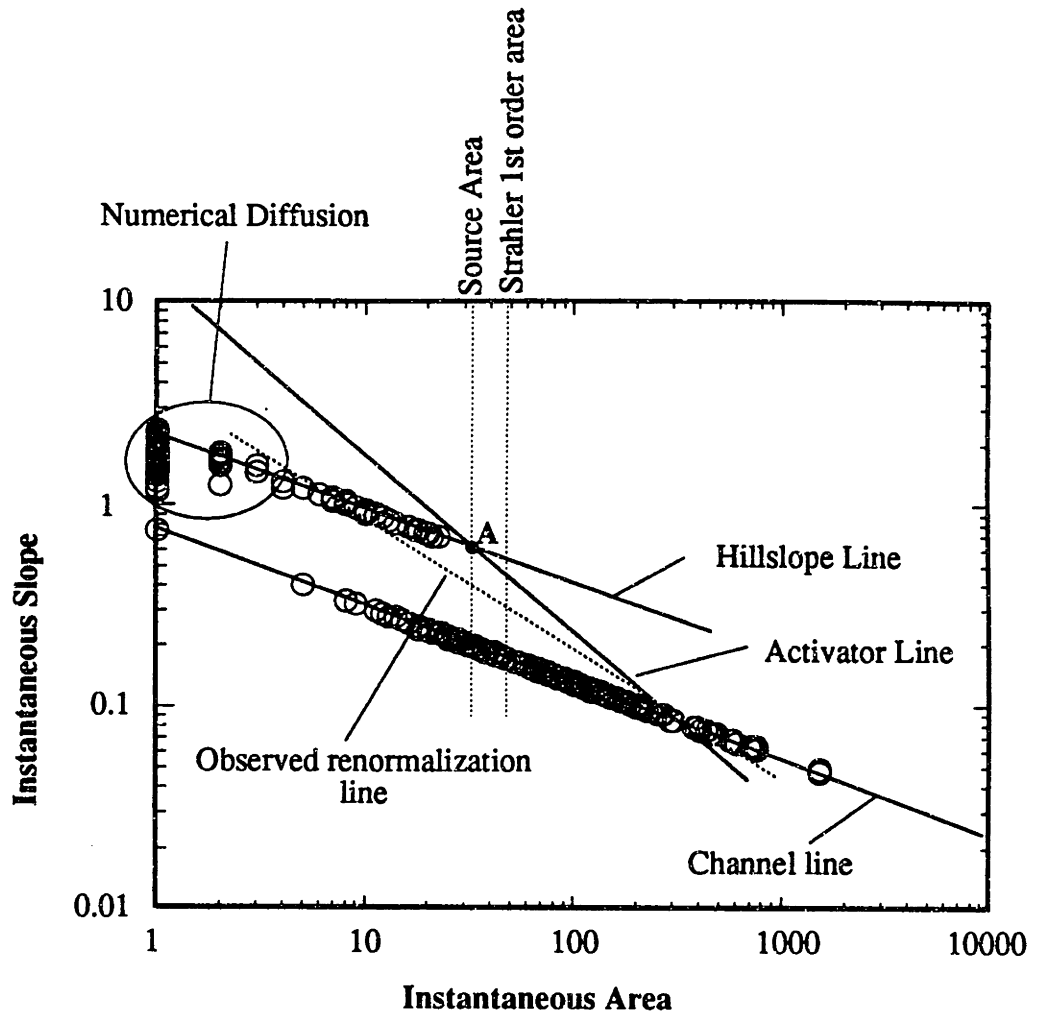
Figure 8.7: Instantaneous areas and slope at
dynamic equilibrium: simulation CR8-1

hillslope scale. As predicted by Equations (8.16) and (8.24), the hillslope line falls higher than the channel line. That is, for the same area the slopes are higher for hillslopes than channels. This slope change is required by sediment transport continuity at the boundary between the hillslope and the channel. This was discussed for the one-dimensional case in Section 7.3. The predicted difference in the slopes, from Equations (8.16) and (8.24) is

$$\frac{S_h}{S_c} = O_t^{-\frac{1}{n_1}} \tag{8.25}$$

and for run CR8-1 ($O_t = 0.1$, $n_1 = 2.1$), this gives $S_h/S_c = 3.0$, which is corroborated by Figure 8.7.

This sharp differentiation between the slopes of the hillslopes and channels is not as pronounced when the catchment is far from dynamic equilibrium. Figure 8.8 shows the catchment of Figure 8.7 at earlier times, when the network is still growing. For areas greater than 10, where timescales of sediment transport are short and elevations proceed to equilibrium quickly, the differentiation of slopes between the two regimes is clear. For smaller areas, where sediment transport is low and elevation change timescales long, there has been insufficient time to develop this differentiation. Note that the scatter of slopes increases as areas decrease, reflecting the progressively longer timescales of elevation adjustment; small areas are further from dynamic equilibrium. The question of timescales of elevation transients and equilibrium are dealt with in detail Appendix A, in the context of the numerical solution of the fluvial sediment transport equations.

The sharp differentiation between the slopes on the hillslopes and in the channels in Figure 8.7 was not observed by Tarboton, et al. Figure 8.7, however, represents the ideal situation and there are many reasons why the differentiation in

slopes may not be observed. First, both Tarboton, et al. and Flint observed scatter in the plotted slopes about the trend of up to half an order of magnitude. In addition, before dynamic equilibrium the differentiation in slopes is not sharp (Figure 8.8). Both of these effects would tend to hide any differences between the hillslopes and the channels.

The activator function used here is perfectly deterministic so that the differentiation between the hillslopes and channel is perfect; i.e. the activator line in Figure 8.7 and 8.8 is a line. In reality both the activator function and the activator threshold will be variable in space so that the activator line will rather tend to a region of transition from hillslope to channel. The hillslope and channel lines in Figure 8.7 will still exist but the horizontal demarcation between hillslope and channel will be blurred so that the region in which both channel and hillslope coexist (left of point A in Figure 8.7) will be larger. If the slopes are averaged for a given area then the mean line will fall through the channel line at high areas and hillslopes at small areas. A possible renormalization that results from this is drawn in Figure 8.7 and is labelled the observed renormalization line. The slope of this line, −0.5 to −0.6, is consistent with slopes calculated by Tarboton, et al.

Tarboton, et al. (1989) noted a deviation from the power law trend of slope for small areas. If their derived network had a high drainage density, they found that with decreasing area mean slopes reached a maximum value and then slopes leveled off or decreased for further decreases in area. This deviation is not totally unexpected. The renormalization of Equation (8.14) predicts infinite slopes at the watershed where the contributing area is zero. This is clearly unrealistic and indicates that other physical mechanisms must dominate hillslope form for small areas. It has been noted several times that the diffusion term of Equation (5.1a) dominates the fluvial term for small areas. Using the full governing Equation (5.2a), Equation (8.15) can be generalized to
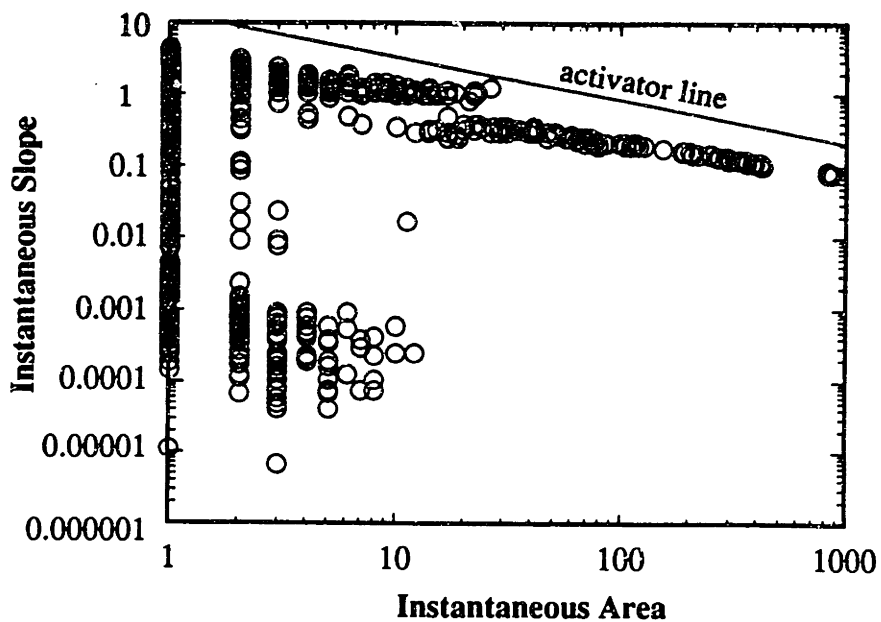
(a) Time = 6000
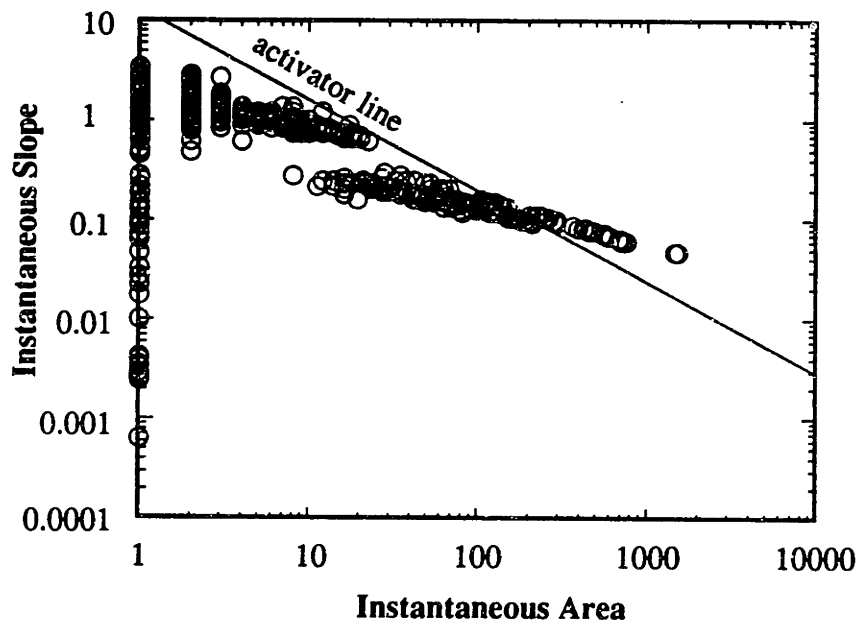
(b) Time = 12000

Figure 8.8: Instantaneous areas and slopes before dynamic equilibrium: simulation CR8-1

$$c_0 A = \beta_1 Q^{m_1} S^{n_1} + D_z L_g S \tag{8.26}$$

where L_g = width of the catchment outlet = 1 (in the simulations)
and substituting for discharge

$$c_0 A = \beta_1 \alpha_1^{m_1} A^{\alpha_2 m_1} S^{n_1} + D_z L_g S \tag{8.27}$$

Clearly as area diminishes the second term dominates and the equation can be approximated by

$$c_0 A = D_z L_g S$$

or

$$A^{-1} S = \frac{c_0}{D_z L_g}$$

or in nondimensional form

$$A'^{-1} S' = \frac{TT}{TD} = \frac{DET}{DED}$$

so that the renormalization coefficient, α_p, for diffusion dominated hillslopes is (Figure 8.9a) $\alpha_p = -1$. The slopes for areas intermediate between the two regimes are indicated in Figure 8.9a and are calculated by solving Equation (8.27) for slopes with area. Simulation CR8–3 was performed to confirm this prediction. The simulation was diffusion dominated and the TA_c was set so that the whole catchment was hillslope (i.e. $TA_c < TA_h$). The areas and slopes near dynamic equilibrium and the
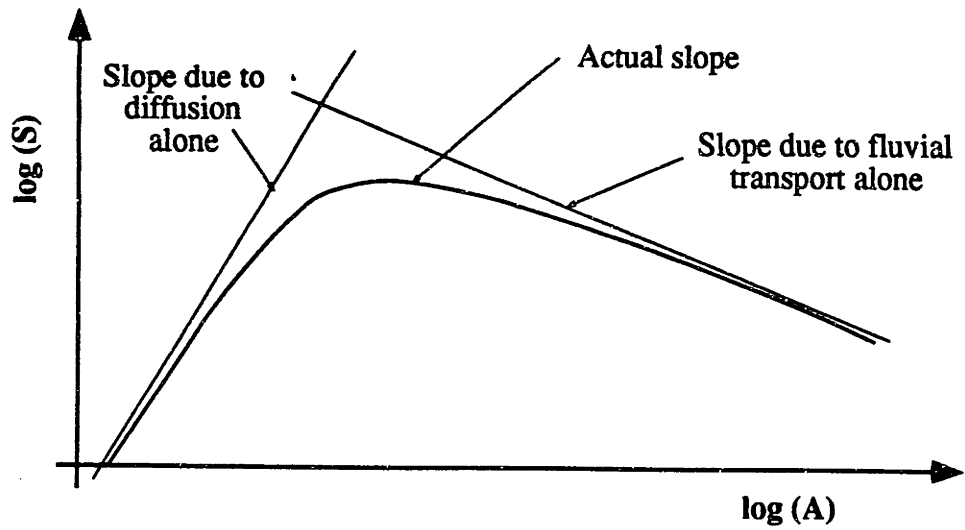
(a)
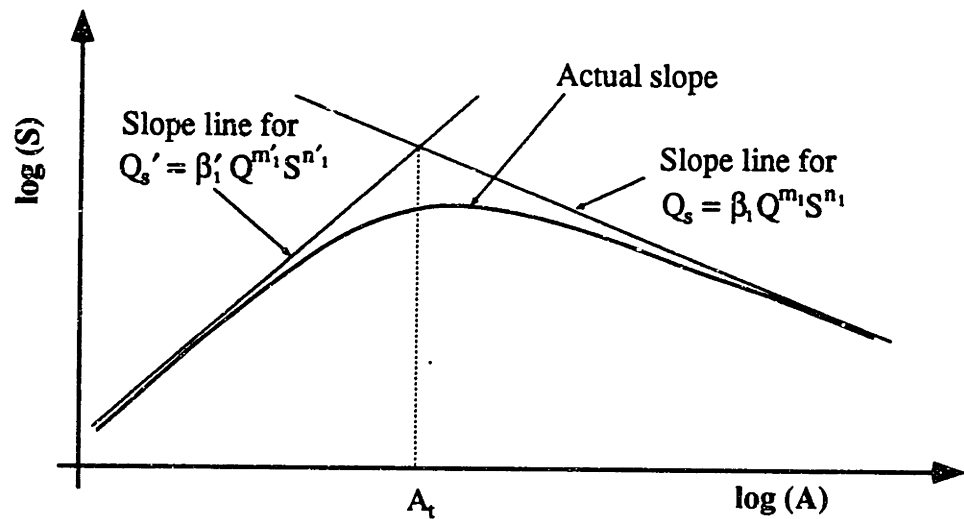
$$A_t = \left[\left[\frac{1}{\alpha_1}\right]^{(m_1 n_1' - m_1' n_1)} c_0^{(n_1' - n_1)}\left[\frac{(\beta_1')^{n_1}}{\beta_1^{n_1'}}\right]\right]^{\frac{1}{\alpha_2 (m_1 n_1' - m_1' n_1) - (n_1' - n_1)}}$$

(b)

Figure 8.9: Diffusion dominated sediment transport and the area-slope renormalization

calculated renormalization line are shown in Figure 8.10. Some nodes with low slopes are apparent, corresponding to portions of the catchment far from dynamic equilibrium (computer time limitations constrained how far the catchment could be allowed to evolve). Note that all slopes fall below the renormalization line because of interactions with the fluvial sediment transport term. This assertion is justified as follows. The problem statement of Equation (8.26) can be generalized with two (or more) different sediment transport mechanisms so that

$$c_0 A = \beta_1 Q^{m_1} S^{n_1} + \beta_1' Q^{m_1'} S^{n_1'} \qquad (8.28)$$

and the slopes at dynamic equilibrium are shown in Figure 8.9b. The main practical problem is to determine the relative magnitudes of the two processes β_1 and β_1'; this determines the area threshold below which one of the processes (the one with smaller α_p) in Equation (8.28) governs the slopes and above which the other process dominates. The scale parameter that controls this is (Figure 8.9b)

$$\overline{\beta}_1 = \left[\frac{\beta_1'^{n_1}}{\beta_1^{n_1'}} \right]^{\frac{1}{\alpha_2(m_1 n_1' - m_1' n_1) - (n_1' - n_1)}} \qquad (8.29)$$

For $\overline{\beta}_1$ large the transition area is large and for $\overline{\beta}_1$ small the transition area is small.

We now leave the issue of the renormalization coefficient and discuss some of the differences in hillslope and channel characteristics highlighted by Figures 8.7 and 8.8. As was shown in Section 7.3, if the channel head is not to grow, then in the hillslope around the channel head,
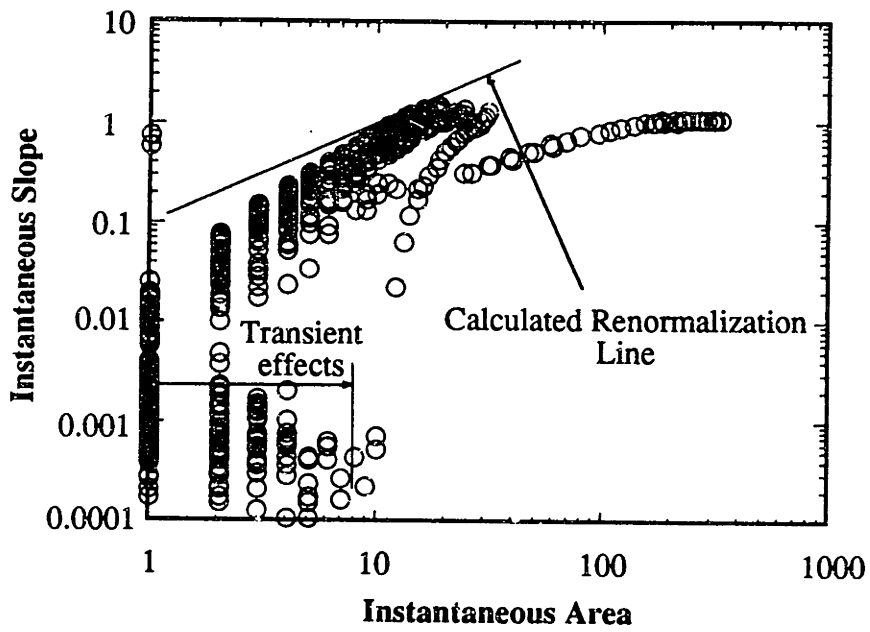
Figure 8.10: Instantaneous area and slopes for diffusion domination: simulation CR8-3

$$\beta_5 \, Q^{m_5} \, S^{n_5} < a_{\text{threshold}} \qquad (8.30)$$

If the channel head is just at the point of growing, then in the hillslope around the channel head

$$\beta_5 \, Q^{m_5} \, S^{n_5} = a_{\text{threshold}} \qquad (8.31)$$

and if the channel head is actively growing, then in the hillslope around the channel head

$$\beta_5 \, Q^{m_5} \, S^{n_5} > a_{\text{threshold}} \qquad (8.32)$$

Equation (8.31) is the dividing line between two regimes of behavior, active channel growth or no–channel growth and will be referred to as the activator line. This line determines the hillslope scales through the hillslope activator number, TA_h. It was shown in Section 7.3.1 that TA_h = constant which is equivalent to Equation (8.31).

All the hillslope nodes in Figure 8.7 must fall within the regime of Equation (8.30); i.e. to the left of the activator line. Some of the hillslopes around the channel heads will fall on the intersection of the activator line and the hillslope line. This situation was illustrated in Figures 8.7 and 8.8 where the activator threshold was plotted. Since not all channel heads lie on the intersection of the hillslope and activator lines, small perturbations in the activator or the threshold will only cause some of the hillslopes around channel heads to channelize, not all of them. This is illustrated in Figure 8.10a where the activator threshold for wetter conditions is plotted.

Another interesting point in Figure 8.7 is that there exists a portion of the channel network that falls to the left of activator line, within the regime of Equation (8.25). This is apparently contradictory.

There is no contradiction if we consider the case where the channelization mechanism is autocatalytic (Figure 8.11a). Once the activator threshold is exceeded channelization occurs irrespective of other catchment characteristics. This is the situation modelled in this work in Equation (5.1). A channel node to the left of the activator line arises when a hillslope node goes into the domain of Equation (8.32) either through attaining a high slope or large contributing area. The hillslope node then differentiates into a channel (i.e. it falls vertically to the channel line in Figure 8.8). Unless the area draining to that node increases dramatically during the channelization process, this node will fall to the left of the activator line after channelization. Channels will be formed at the first point downstream at which the area is sufficient to cause the activator to exceed the activator threshold. Immediately downstream of this point, in the channel, slopes are lower and these points in the channel will fall to the left of the activator line.

However, this does not explain the existence of channels to the left of point A in Figure 8.7. These channel nodes result from high transient slopes in the hillslopes at times before dynamic equilibrium. At early times there is significant scatter in the hillslope slopes (e.g., Figure 8.8). Some hillslope slopes are large enough to trigger channelization, even though their slope at dynamic equilibrium (given by the hillslope line in Figure 8.7) would be too low to do so. This influence of transient slopes is an important point since it indicates that even at dynamic equilibrium there are parts of the channel network that exist only because of transient effects during the catchment evolution process

On the other hand, consider the case where channelization is not an autocatalytic process. In this case for a hillslope node to become a channel, activator
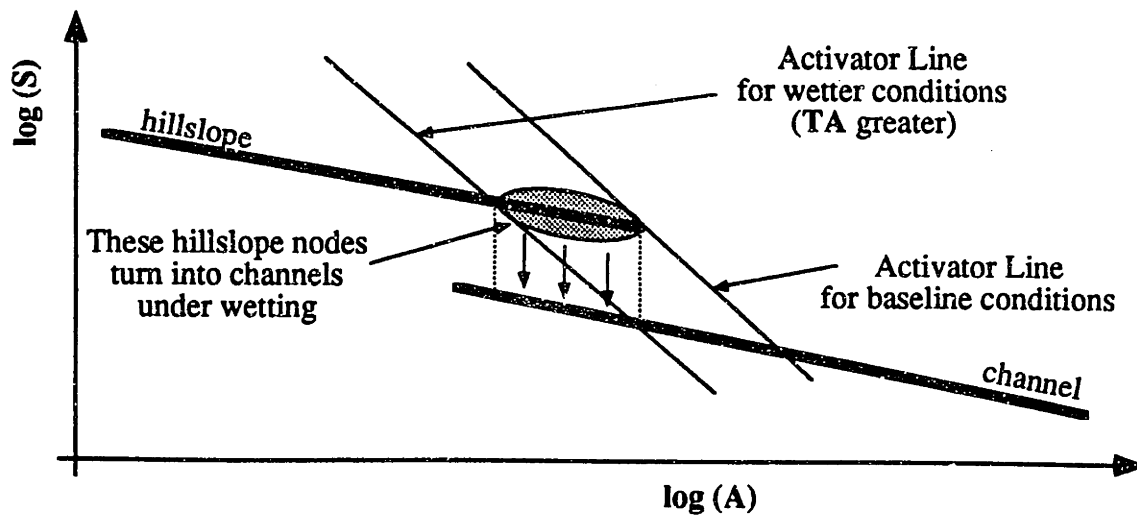
(a) Autocatalytic channelization
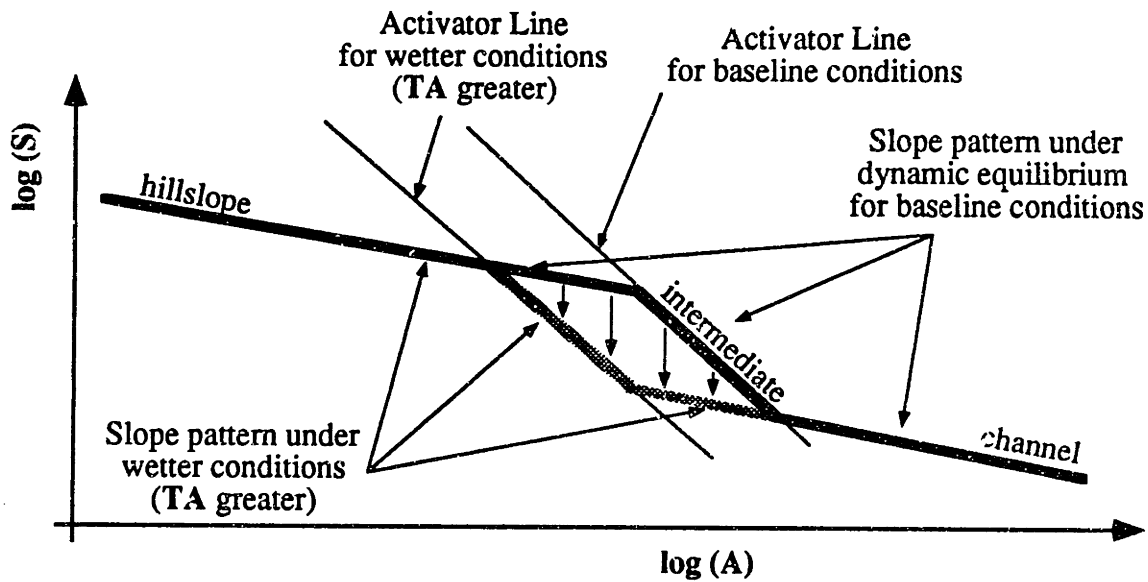
(a) Nonautocatalytic channelization

Figure 8.11: Channelization under wetting conditions

must exceed the activator threshold throughout the transition from hillslope to channel, i.e., during the transition the node must always fall to the right of the activator line. If a hillslope node is just to the right of the activator line then it starts to channelize. It will fall vertically with time in Figure 8.7 (providing its contributing area does not change substantially). At some later time this node will cross the activator line at which time channelization can proceed no further. That node will be stuck in a state intermediate between hillslope and channel. This case is illustrated in Figure 8.11. The points that fall on the activator line are intermediate between hillslope and channel. Since there is no distinct transition from channel to hillslope, as was the case for autocatalytic channelization, then channel heads will no longer be distinct. Rather channel "heads" will have a gradual transition from channel to hillslope as described by Kirkby (1988).

There is one similarity between the case of autocatalytic and nonautocatalytic channelization. For the nonautocatalytic channelization case under wetter conditions transition nodes become more channel—like. Under dryer conditions, however, intermediate nodes do not become more hillslope like; the activator function only triggers hillslopes to become channels in the differentiation equation, not vice versa. For the autocatalytic channelization case under wetter conditions the channel network will extend. Under drier conditions the channel network cannot recede unless a mechanism that explicitly obliterates channels is postulated. Thus both types of channelization postulate a mechanism for channel growth that is one way: from hillslope to channel. Neither explicitly postulates a mechanism for obliterating channels.

8.4 Channel Initiation Mechanisms

Dietrich, et al. (1986) and Montgomery and Dietrich (1988) studied the hillslopes surrounding channel heads in a number of field sites in California and

Oregon. They observed a hyperbolic relationship between the area upstream of the channel head, the source area, and the local slope on the hillslope in the steepest direction immediately upstream of the channel head. In addition, they observed a negative correlation between the prevailing mean annual rainfall and the size of the source area. More rainfall was accompanied by a smaller source area for a given slope.

The data presented in Montgomery and Dietrich (1988) are reproduced in Figure 8.12. The three study sites, Coos Bay, Sierra Nevada, and Marin County have mean annual rainfalls of 1500 mm, 260 mm, and 760 mm, respectively.

The activator formulation of Equation (5.2a) is

$$a = \beta_5 \, Q^{m_5} S^{n_5}$$

and if the Hortonian runoff discharge–area relationship of Equation (6.5) is adopted

$$Q = RA$$

where R = effective runoff rate.

then the activator function may be expressed as

$$a = \beta_5 \, R^{m_5} A^{m_5} S^{n_5} \tag{8.33}$$

If the data of Montgomery and Dietrich represent the activator line of Section 7.3.1 and Equation (8.31), then their data may be fitted by an equation of the form of

$$A^{m_5} S^{n_5} = \frac{a_{threshold}}{\beta_5 \, R^{m_5}} \tag{8.34}$$

353
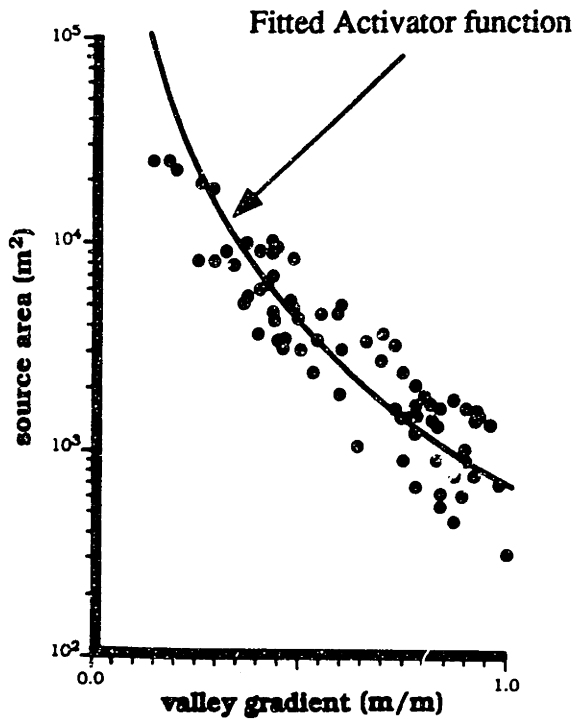
(a) Coos Bay
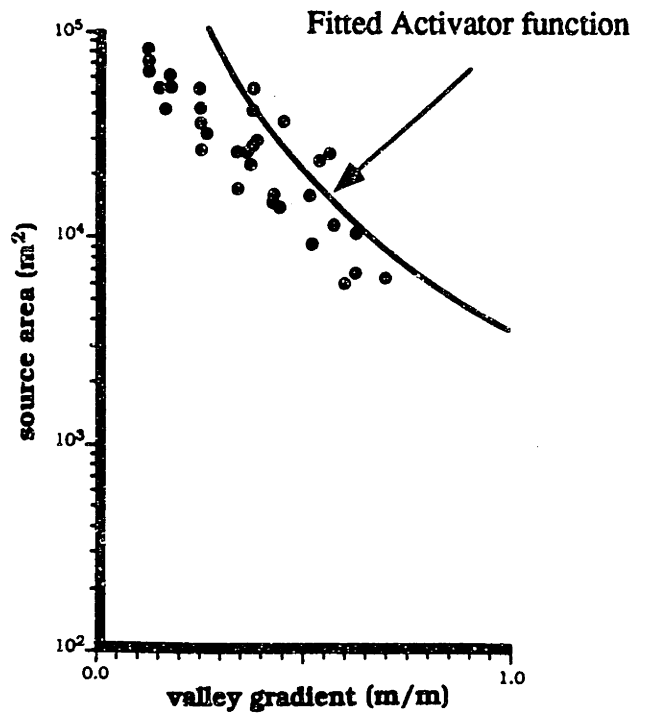
(b) Sierra Nevada
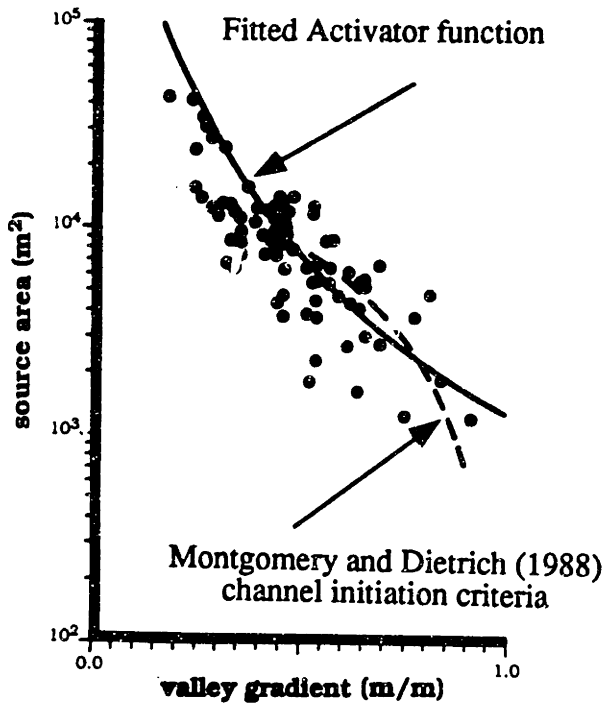
Figure 8.12: Hillslope slope and contributing
area data: Montgomery and Dietrich (1988)
(original drawings from W. Dietrich)

(c) Marin County

354

or

$$A \, S^{n_5/m_5} = \left[\frac{a_{threshold}}{\beta_5} \right]^{\frac{1}{m_5}} \frac{1}{R}$$

$$= a'_{threshold} \tag{8.35}$$

where $a'_{threshold}$ = the activator threshold normalized (but not

nondimensionalized) for prevailing runoff rates, and

other catchment conditions.

Equation (8.35) was fitted by eye to the data of Figure 8.12 Each of the three data sets were fitted individually and the best regressions are listed in Table 8.1. These regressions are also plotted in Figure 8.12. The general form of the fitted activator line was

$$A \, S^{2.5} = \text{constant} \tag{8.36}$$

These regression equations are very different to those fitted by Montgomery and Dietrich (the dotted line in Figure 8.12c). Their mechanism based on groundwater stream sapping and landsliding is concave down, the regression fits are concave up.

In addition, Montgomery and Dietrich note that the data contradicts results obtained by Kirkby (1986) for a subsurface saturation runoff mechanism with channel initiation based on overland flow sediment transport. Kirkby predicts an increase of source area with slope, since as slope increases the saturated area decreases reducing the overland flow erosion potential. If on the other hand the hillslope is dominated by Hortonian overland flow in triangular rills, flow governed by Mannings equation, and if bottom shear stress is assumed to be the process that triggers channelization, then the functional form of the activator equation (Table 5.2) is

TABLE 8.1

Activator Regression;
Montgomery and Dietrich (1988) Data

Site	Regression
Coos Bay	$AS^{2.5} = 710$
Sierra Nevada	$AS^{2.5} = 2020$
Marin County	$AS^{2.5} = 1240$

$$a = \tau_* = (\beta \, Q \, S^{2.2})^{0.375}$$

This overland flow mechanism is approximately consistent with the field data of Figure 8.12. Of course, this argument does not constitute a proof that the channel initiation mechanism is shear stress in rilled flow. Rather it indicates the difficulty of attributing a physical interpretation to the channel initiation mechanic.n without detailed field work to understand both the dominant, or geomorphologically effective, runoff mechanisms and the detailed mechanics of channel extension at the channel head. For instance, preliminary field work suggests that groundwater flow plays an important part in the growth of channels heads in Marin County (Dietrich, personal communication) yet Equation (8.34) does not explicitly involve groundwater processes.

Montgomery and Dietrich also noted a systematic deviation in the source areas measured for the three study sites and qualitatively attributed that deviation to differences in the average rainfall, and thus runoff, at each site. A large amount of the systematic deviation of the constant in Equation (8.36) can be quantitatively explained by the dependence of Equation (8.34) on runoff rate. Consider the simplest case possible where the runoff in Equation (8.35) is proportional to the mean annual rainfall (i.e. $r \, \alpha \, R$). Using the equations fit by eye to the data (Table 8.1) gives a best fit equation of

$$Q \, S^{2.5} \quad \alpha \quad r \, A \, S^{2.5} = 0.87 \qquad\qquad (8.37)$$

where r = rainfall rate.

This regression reduces the coefficient of variation of the residuals from 50% to 19%. The fit is hardly perfect, and only three points were used in the regression so not too much significance should be placed on this result. What it does indicate, however, is that the trend of the field data is consistent with runoff component of the activation

357

criteria fitted to the data. This lends credence to the form of the activator criteria.

Finally it should be noted that the slopes in the channels are irrelevant in Equation (8.37). The relevant slopes are the hillslope slopes. As has been stressed elsewhere in this work, the channel growth process is governed by the hillslope processes; the channels act only to guide the overall form of the hillslopes by the preferential drainage that occurs to the channel.

The fitted activation criteria is consistent with those proposed in Table 5.2, though it might be noted that the fit of the area to slope is at the high end of the range of n_5/m_5 proposed. Most of the simulations discussed in Chapter 7 used a value of $n_5/m_5 = 0.75$, while the fitted value is 2.5. To check that reasonable networks and catchments can still be generated with this value of n_5/m_5 a number of simulations were run with $n_5/m_5 = 2.0$. The networks and the catchment elevations at the time when the channel network stopped growing are given in Figures 8.13 and 8.14. The results are reasonable, but do display a number of interesting characteristics.

The first interesting characteristic is that the networks generated did not always fill the entire grid when the tectonic uplift applied was an initial instantaneous uplift. In some of the runs the network only extended into about half the grid and then stopped. This behavior is explained by examining TA_h for $n_5/m_5 = 2.0$ (Equation 6.13e).

$$TA_h = \frac{L_z^{n_5} \; L_R^{m_5} \; c_1 \; \beta_5}{T_R^{m_5}} \tag{8.38}$$

Note that L_x does not appear here because $2m_5 - n_5 = 0$. Thus $TA_h = \text{constant}$ requires $L_z = \text{constant}$; all other scales in Equation (8.38) are externally applied. At the root of the network, where the channel is at its lowest elevation, the hillslope relief,

358
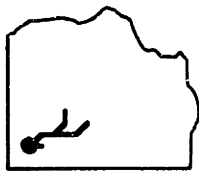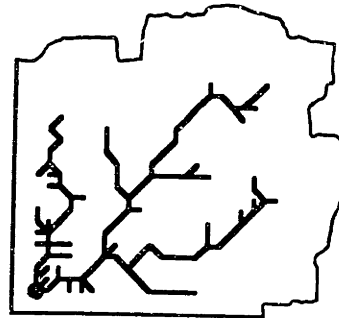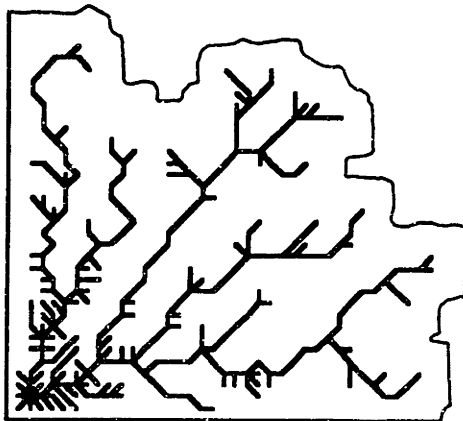
(a) CR9-10
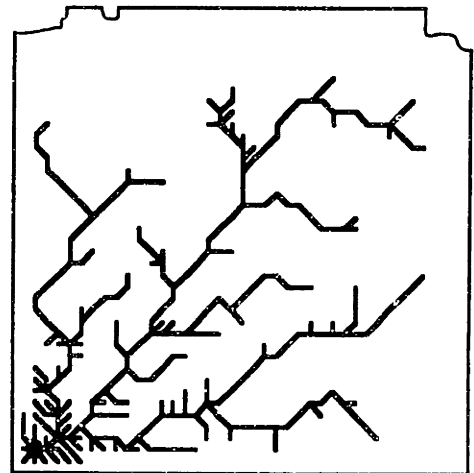
(b) CR9-11

(c) CR9-8

(d) CR9-6

● Catchment outlet

Figure 8.13: Channel networks: Activator $a = \beta_5\, Q\, S^2$
simulations CR9-6, CR9-8, CR9-10, CR9-11

(a) CR9-10

(b) CR9-11

(c) CR9-8

(d) CR9-6

Max.   Min.
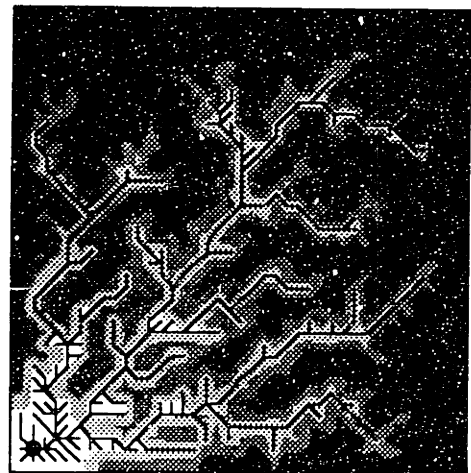
● Catchment outlet

Figure 8.14: Catchment elevations: Activator $a = \beta_5 \, Q \, S^2$
simulations CR9-6, CR9-8, CR9-10, CR9-11

L_z, is highest. Thus channels will grow here. For points upstream, which are at a higher elevation, the hillslope relief is lower and TA_h is lower. At some point TA_h falls below the value above which channels grow, at which stage the channel can no longer grow. For hillslope relief lower than this threshold relief the activator is insufficient to extend the channelization. As $a'_{threshold}$ is increased so that this threshold hillslope relief is increased it is found that the channel network is unable to grow the fill the whole region. This argument is valid for Davisian, event tectonic uplift, regimes. Regimes with continuous tectonic uplift may suffer the same fate if overland erosion rates are very high relative to the rate of extension of the channel network so that hillslopes never attain the necessary high reliefs.

The second interesting characteristic of the simulation is related to the value of n_5/m_5. For the simulations of Chapter 7, with $n_5/m_5 = 0.75$, the relief threshold characteristic noted above does not occur because TA_h involves a term of the form $L_x^{2m_5-n_5}$, where $2m_5-n_5 > 0$. If there is insufficient hillslope relief to induce channelization, then channelization stops until overland erosion has proceeded enough to increase the contributing area to that channel head (i.e L_x is increased). Lower relief just means increased hillslope lengths, L_x, before channelization will occur. For $n_5/m_5 = 2.0$, the contributing area has no effect whatsoever. For $n_5/m_5 > 2.0$, decreasing the contributing area increases TA_h and thus the tendency to channelization. This arises because, in TA_h, as the contributing area increases, the slopes decrease, and this decrease of slope dominates the increase in area. Since Equation (7.15) indicates that TA_c dominates all other nondimensional numbers in setting the rate of growth of the network, channel growth is fastest when the slopes are highest. Fast channel growth will tend to maintain high slopes. At some slope, however, the hillslopes will become unstable and landsliding will become the dominant mechanism for hillslope sediment transport. This follows from the equation for

landslide transport proposed by Andrews and Bucknam (1987) where transport tends to infinity as the slope approaches the angle of repose (Equation 2.23). This is consistent with Montgomery and Dietrich who reported micro–landsliding at some of their field sites, and the fit to their data of $n_5/m_5 = 2.5$ in Table 8.1 puts their data in this steep slope regime.

The third interesting characteristic is the influence of n_5/m_5 on growth site screening. A clear trend in drainage density can be seen in Figure 8.13; drainage density is highest at the outlet and least near the edges of the network. Screening of growth sites occurs in the catchment simulation model by the reduction of contributing areas to nodes near the root of the channel network. Slopes are about the same everywhere. For $n_5/m_5 < 2$, TA_h is highest at the network extremities. As n_5/m_5 increases the effect of screening is reduced because of the relative lesser importance of areas and network growth is less biased towards the network extremity. Drainage densities will tend to be higher inside the network.

Finally the dynamic equilibrium and renormalization ideas of the previous section have some important implications for the interpretation of the data in Figure 8.12. The channel slope–area renormalization of the previous section and the autocatalytic channelization process mean that a number of different, and conflicting, interpretations can be given to the data.

The first interpretation is the case where the catchments are at dynamic equilibrium and the channelization mechanism is autocatalytic. In this case all of the data of Figure 8.12 must be on the hillslope line of Figure 8.8. In this case the Montgomery and Dietrich data is not measuring the activator line of Figure 8.8, but is in fact measuring the renormalization function for the hillslope. Moreover, the value of α for a regression fit of Equation 8.14 to the data of Figure 8.11 is $\alpha \approx 0.33 - 0.4$. This value is within the range of values for α_p (Equation 8.17) using fluvial sediment transport equations (Table 5.1). This is one explanation for their data.

The second interpretation that may be given to the data is that the catchment is at dynamic equilibrium, but that the channelization mechanism is not autocatalytic. This case was illustrated in Figure 8.11b. In this case it is probable that the data of Montgomery and Dietrich measure the activator line, in Figure 8.11b labelled transition. Note that in this case the channel heads would be indistinct and not well defined; Dietrich (personal communication) indicates this to be the case at some of the study sites.

The third interpretation that may be given to the data is that the catchment is not in dynamic equilibrium and that channel head growth is still occurring. In this case the hillslopes around growing channels will lie to the right or above the activator line. Nongrowing channels will be to the left or below the activator line. The observed data would probably be indicative of the activator line. Montgomery (personal communication) has observed that some of the channel heads are actively eroding. This suggests that at short time scales, the channel network is growing into the hillslope; catchments cannot be at dynamic equilibrium. Whether this is true for long time scales is unknown, the current growth may be no more than a short term random fluctuation of the channel head position.

Which of these three interpretations of the presented data is correct cannot be determined without further study. Each interpretation is consistent with some of the observations so that none can be totally eliminated from consideration at this stage. Either of the second cr third interpretations suggest that the data measures the activator line and thus the activator function. This was the interpretation that was adopted in this section.

In conclusion, the data of Dietrich, et al. (1986) and Montgomery and Dietrich (1988) has been examined and reinterpreted using the activator function of this work. After fitting the activator function catchments were simulated and the results were satisfactory. The simulated channel networks did exhibit a trend in drainage density

within a single network and it was postulated that this result followed from the reduced growth site screening for large values of n_5/m_5, the regime of Montgomery and Dietrich's data. Some questions were raised regarding the interpretation of the presented data that can only be resolved by further field work

CHAPTER 9

CONCLUSIONS

9.1 Summary of Results

This work consisted of three sections. The first section was devoted to understanding those processes that create networks, whether they be hydrologic or otherwise. On the basis of extensive simulation with a model from the biological literature a number of necessary conditions were proposed for any physical system that generated networks. Elements of chaotic behavior were discovered in the network growth process of this model. This chaos explained the observed random network generation. Because of similarities in the physics there is reason to believe that these results are applicable to the channel network growth model proposed in this work.

The second section of this work developed the physically based model of channel network growth and hillslope evolution. This model simulates the long term changes in elevations within the catchment and the consequent effect on channel network growth and hillslope form. The changes in elevation are modeled by continuity equations for flow and sediment transport; elevation changes result from local imbalances in the sediment transport. A channelization mechanism, called activator, which is nonlinearly dependent on discharge and local slope is used and it is the spatial distribution of the activator around the channel head that governs where and whether the channel head grows; a channel grows if the activator exceeds a threshold. A central component of the model is that erosion in the channels takes place at a faster rate than in the hillslope. This preferential erosion in the channels results in convergence of the flow on the hillslopes towards the channels. It is the convergence of flow that triggers channel growth. Thus the interaction between the hillslopes and the channels over long timescales is central to the final form of the channel network, and

the hillslopes.

The sediment transport formulation used was related to a generally accepted sediment transport equation; Einstein—Brown. It was shown that the instantaneous sediment transport can be related to the long term mean sediment transport with a simple dependence on the mean annual peak discharge. The activator function was justified, in a similar fashion, using commonly accepted thresholds for channel erosion protection.

A nondimensionalization of the equations was presented and nondimensional numbers that control the physical processes within the catchment were proposed. Using these nondimensional numbers similarity conditions for disparate catchments were presented; catchments that satisfy the similarity conditions have comparable physics. This work allows, for the first time, the comparison of catchments in a physically meaningful fashion. The implications of the nondimensionalization for a proposed classification scheme of fluvial landscapes were discussed and some questions raised regarding the unsuitability of aridity and rate of tectonic uplift as landscape classification criteria.

The third section of the work concentrated on the interpretation of the results of using the proposed catchment evolution model in a numerical simulation mode. It was demonstrated that the network's growth process was sensitively dependent on the elevations. For catchments satisfying the similarity conditions noted, the topological statistics of the generated networks varied greatly both in time and between simulations. This variation was consistent with results from small scale model studies. Statistics of the catchment not based on topological characteristics exhibited much less variability.

One of the nondimensional numbers, the activator number measuring the magnitude of the channel initiation process, was found to be a significant influence on the drainage density within the catchment. If the length scales used for calculating the

activator number were the hillslope length and hillslope drop the activator number was found to approximately constant. This constancy of the hillslope activator number was the fundamental relationship determining the hillslope length scales and thus the drainage density. Importantly the activator number was dependent on both hillslope slope and hillslope length so that steeper hillslopes meant shorter hillslopes and a higher drainage density. Without information on the hillslope slopes the hillslope activator number does not completely determine the drainage density. It was demonstrated that the hillslope slopes were related to the catchment slopes and in turn the catchment slopes were related to the tectonic uplift regime. Steeper catchments had steeper hillslopes and through the hillslope activator number higher drainage densities. A regression between drainage density and the catchment slopes provided the crucial link between catchment scales and hillslope scales that allowed the determination of the hillslope length. Thus neither the hillslope nor the channel network may considered separately, they are only components of a larger system, the catchment, which must be considered in its unity to understand the development of landscape form.

The drainage density and the rate of increase of drainage density were related to the nondimensional numbers and physically based explanations of the trends proposed. The planar properties of the networks and the hillslopes were examined. It was noted that many of the measures of link length and hillslope length are linearly related and these relationships appeared to be independent of the majority of the governing physics. The form of the catchment at dynamic equilibrium was examined. In particular, it was shown that differences in the hypsometric curve that have been previously attributed to catchment age are in fact attributable to the tectonic uplift history of the catchment; i.e. whether tectonic uplift was a single uplift event (i.e. Davisian) or continuous uplift leading to dynamic equilibrium.

The verification of the model against existing experimental data was an

important component of this work. Many characteristics of the network were examined and successfully compared with the reported experimental data. The model was compared with trends observed for small scale experimental catchments generated in erosion facilities. The trends of sediment transport with time and drainage density with slope are consistent with those reported.

Experimental results relating channel slope to link magnitude and contributing area were examined. Similar results were observed in the simulations. This correlation was explained by an equation relating the sediment transport and tectonic uplift within the catchment at dynamic equilibrium. This result, which followed from the proposed catchment evolution model, explained many, apparently anomalous, characteristics of the field data. Deviations in slopes at small areas were demonstrated to result from diffusion domination of transport on the hillslopes. It was shown that the slope—area regression may be a useful means for determining the dominant sediment transport processes both within the channels and on the hillslopes.

Recent data relating hillslope slopes and contributing area draining to channel heads were satisfactorily explained by the activator function adopted in this work. The activator function was also able to explain an negative correlation observed between the channel head source areas and mean annual rainfall at the site. This followed from the activator's functional dependence on discharge, and thus runoff. Catchment's simulated with the activator function fitted to the data appeared realistic. In addition, the nondimensionalization of the activator indicated that the fitted channel head extension process should induce localized landsliding, which was consistent with the field observations.

In conclusion, it is noted that the catchment evolution model proposed in this work was consistent with most of the available field data. Some credence may be placed on the predictions of the model. While the simulation model is not perfect it is strongly believed that the model provides a powerful new tool with which to examine

the physical processes that govern the geomorphology and hydrology of catchments.

9.2 Further Research

Further research based on the work presented here can proceed in two directions. First, since many of the large scale interactions within the model are poorly understood there is a need for further verification of the model against field data. Second, the model, model structure and nondimensionalization of the proposed physics provide opportunities to explore aspects of both hydrology and geomorphology that have been previously inaccessible.

Consider the field verification of the model. Here further work can take two directions. The first direction is to attempt to obtain all the physical parameters required for the model in order to confirm the predictions of the model. As noted in the introduction to Chapter 8 these data requirements are quite daunting and include

1. Channel and hillslope sediment transport equations.

2. Flood frequency distributions related to area.

3. Channel head advance rates and an accurate mapping of the geographic position of channel heads. Determination of which heads are advancing and which are not so that the activator function may be determined.

4. Detailed elevation data for the catchment.

The second direction that experimental verification may take is along the lines of Section 8.3 where the field data were used to verify simulation results from the model rather than the actual governing processes. If theoretical predictions can be made with the model then these can be used to design a well controlled field experiment measuring the physical properties of greatest interest.

For instance, in Section 8.3 a relationship between slope and area is discussed. The comparison of this section was incomplete since flood frequency data and sediment transport relations were not available to verify the fitted relation. The comparison

might be considered something of a parameter estimation exercise since the model was shown to be a plausible explanation for the observed trend not that it was the correct explanation. More work could usefully be carried out here.

Field measurement of slopes and areas in both the channel and the hillslopes at the channel heads may be able to verify the sharp difference in hillslope and channel slopes predicted by the model at dynamic equilibrium. A paired comparison of slopes in the channel and the hillslope at the channel head may be able to eliminate much of the scatter observed when the slopes and areas from digital elevation data are lumped together.

The mean long term sediment transport relation derived from the instantaneous sediment transport relation could be verified from field data. Sediment transport over a number of storms could be examined and the net result integrated.

There are considerable prospects for further work theoretically using the numerical model developed. Ideally much of this work would proceed in parallel with field work to allow verification of predictions made. Some ideas for extension of this work include

1. Incorporation of the subsurface runoff mechanism. The current computer code uses Hortonian overland flow. The code was designed to facilitate the extension to subsurface saturation runoff. Simulated catchment characteristics could then be examined. New nondimensional parameters that result from the new physics could be correlated to catchment characteristics, as was done in Chapter 7. It may be possible to distinguish between those catchments with Hortonian runoff and those with subsurface saturation runoff on the basis of these catchment characteristics.

2. Other sediment transport mechanisms on the hillslopes may be examined. The physically based explanation of the area–slope

renormalization provides a framework within which this may be examined. With sufficiently detailed elevation data it may be possible to infer the dominant transport mechanism on the hillslope remotely, which could then be verified in the field.

3. The constant velocity assumption implicit in many instantaneous unit hydrograph (IUH) models may be examined. The renormalization of slope on area provides a physical basis on which to base estimates of flood wave velocity.

4. An extension of point 3 would be to develop a synthetic IUH by examining the unit response of the synthetic catchments. The shape of the IUH could parameterized in a particularly parsimonious fashion based solely on the nondimensional numbers.

5. The similarity conditions proposed in Chapter 6 could be examined to increase their generality in the classification of landscape. In a practical setting it may be possible to use these similarity conditions in the restoration of land to its natural condition after open—pit mining operations. The numerical model would also be useful in this regard. It might also be a useful design tool for predicting risk of gully erosion from modification of landuse and climate.

Any extension of this work will doubtless require a great deal of numerical simulation. While substantial effort has already been devoted to the solution of the sediment transport equation there is a need for further work. The existing technique is explicit and suffers the stability problems of explicit schemes. The use of the predictor—corrector algorithm with nonlinear extrapolation only serves to defer the stability problems. Currently the model is very computer time intensive, each simulation taking of the order of 1–2 CRAY X–MP CPU hours (approximately 100–200 microVAX hours). For the grid sizes used this is comparable with other

two—dimensional sediment transport models. Nevertheless, it is apparent that some effort will need to be devoted to the numerical solution scheme before significantly larger grids can be solved.

In conclusion, it might be noted that this work has barely scratched the surface of a novel approach for examining the geomorphology and hydrology of the landscape. It is believed that the model of this work provides a useful framework that can be extended and generalized to study many unexplored aspects of catchment geomorphology and hydrology.

REFERENCES

Abrahams A D. 1984. "Channel networks: A geomorphological perspective." *Water Resources Research*, 20(2): pp 161 - 168.

Acton F S. 1970. *Numerical methods that work.* New York: Harper Row.

Ahnert F. 1976. "Brief description of a comprehensive three-dimensional process-response model for landform development." *Zeitschrift fur Geomorphologie N.F. Supplement 25*: pp 29 - 49.

____. 1984. "Local relief and the height limits of mountain ranges." *American Journal of Science*, 204: pp 1035 - 1055.

Andrews D J and Bucknam R C. 1987. "Fitting degradation of shoreline scarps by a nonlinear diffusion model." *Journal of Geophysical Research*, 92(B12): pp 12857 - 12867.

Ball R C. 1986. "DLA in the real world." Stanley H E and Ostrowsky N, *On growth and form. Fractal and non-fractal patterns in physics*, Institut d'Etrudes Scientifiques de Cargèse, Corsica, France, June 1985.

Beven K J and Kirkby M J. 1979. "A physically based variable contributing area model of basin hydrology." *Hydrological Sciences Bulletin*, 24(1): pp 43 - 69.

Beven K J and Wood E F. 1983. " Catchment geomorphology and the dynamics of runoff contributing areas." *Journal of Hydrology*, 65: pp 139 - 158.

Briere C. 1983. " Mathematical analysis of an activator-inhibition system in a discrete field." *Mathematical Bioscences*, 73: pp 51 - 60.

Brunsden D. 1980. "Applicable models of long term landform evolution." *Zeitschrift fur Geomorphologie Supplement 36*: pp 16 - 26.

Calver A. 1978. "Modelling headwater development." *Earth Surface Processes*, 3: pp 233 - 241

Chase C G. 1988. "Fluvial landsculpting and the fractal dimension of topography." *EOS*, 69(44) :p 1207.

Chow V T. 1964. *Handbook of Engineering Hydrology*.

Coelho-Netto A L, Fernandes N F and Moura J R. 1987. "Drainage network expansion by gully advance through RAMPA deposits in SE Brazilian Plateau: Bananal, SP." *EOS*, 68(44) :p 1273.

Coniglio A and Stanley H E. 29 March,1984. "Screening of deeply invaginated clusters and the critical behavior of the random superconducting network." *Physical Review Letters*, 52(13) :pp 1068 - 1071.

Cordova J R, Rodriguez-Iturbe I and Vaca P. 1982. "On the development of drainage networks." Walling D E, Recent developments in the explanation and prediction of erosion and sediment yield, Exeter, UK, 19 - 30 July 1982.

Coventry R J, Moss A J and Verster E. 1988. "Thin surface soil layers attributable to rain-flow transportation on low-angle slopes: An example from semi-arid tropical Queensland, Australia." *Earth Surface Processes and Landforms*, 13 :pp 421-430.

Cuiling W E H. 1963. "Soil creep and the development of hillside slopes." *Journal of Geology*, 71(2) :pp 127-161.

___. 1986. "Highly erratic spatial variability of soil-pH on Iping common, West Sussex." *Catena (Catena Verlag)*, 13 :pp 81-98.

___. 1988. "Dimension and entropy in the soil-covered landscape." *Earth Surface Processes and Landforms*, 13 :pp 619 - 648.

Culling W E H and Datko M. 1987. "The fractal geometry of the soil covered landscape." *Earth Surface Processes and Landforms*, 12 :pp 369 - 385.

Daoud M. 1986. "Branched polymers." Stanley H E and Ostrowsky N, On growth and form. Fractal and non-fractal patterns in physics, Institut d'Etrudes Scientifiques de Cargèse, Corsica, France, June 1985.

Davis F W, Dozier J and Dubayah R O. 1988. "Topographic variation in solar radiation over the Konza Prarie." *EOS*, 69(16) :p 359.

Dietrich W E, Wilson C J and Reneau S L. 1986. "Hollows, colluvium, and landslides in soil-mantled landscapes." *Hillslope processes*, Ed Abrahams A D. pp 361 - 388. Binghampton Series on Geomorphology.

Dunne T. 1969. *Runoff production in a humid area*. Ph.D thesis, Johns Hopkins University, Baltimore.

___. 1980. "Formation and controls of channel networks." *Progress in Physical Geography*, 4 :pp 211-239.

___. 1988. "Sheetwash as a diffusive process stabilizing hillslopes against channel incision." *EOS*, 69(16) :p 345.

___. 1989. "Hydrology, mechanics and geomorphic implications of erosion by subsurface floe." *Groundwater Geomorphology*, Ed Higgins. Geological Society of America.

Dunne T and Aubrey B F. 1986. "Evaluation of Horton's theory of sheetwash and rill erosion on the basis of field experiments." *Hillslope processes*, Ed Abrahams A D. Binghampton Series on Geomorphology.

Eagleson P S. 1970. *Dynamic Hydrology*. NY: McGraw Hill.

Flint J J. 1974. "Stream gradient as a function of order, magnitude and discharge." *Water Resources Research*, 10(5) : pp 969 - 973.

Gardiner V and Gregory K J. 1981. "Drainage density in rainfall - runoff modeiling." Singh V P, Proceedings of the international symposium on rainfall and runoff modelling, Mississippi State University, May 18 - 21 1981.

Gelhar L W and Axness C L. 1983. "Three dimensional stochastic analysis of macrodispersion in aquifers." *Water Resources Research*, 19(1) :pp 161 - 180.

Gierer A. 1981. "Some physical, mathematical and evolutionary aspects of biological pattern formation." Brenner S, Murray J D and Wolpert L, Theories of biological pattern formation : A Royal Society discussion, London, 1981.

Graf W L. 1971. *Hydraulics of sediment transport.* NY: McGraw Hill.

Granero M I, Porati A and Zanacca D. 1977. "A bifurcation analysis of pattern formation in a diffusion governed morphogenetic field." *Journal of Mathematical Biology*, 6 :pp 21-27.

Gregory K J. 1976. "Drainage networks and climate." *Geomorphology and climate*, Ed Derbyshire E. pp 289-318. London: Wiley.

Gregory K J and Walling D E. 1968. "The variation of drainage density within a catchment." *International Association for Scientific Hydrology Bulletin*, 13 :pp 61-68.

Gupta V K, Waymire E and Wang C T. 1980. "A representation of an instanteous unit hydrograph from geomorphology." *Water Resources Research*, 16(5), pp 855 - 862.

Gupta V K and Waymire E. 1983. "On the formulation of an analytical approach to hydrologic response and similarity at the basin scale." *Journal of Hydrology*, 65

Hack J T. 1957. "Studies of longitudinal stream profiles in Virginia and Maryland." USGS.

Haken H and Olbrich H. 1978. "Analytical treatment of pattern formation in the Gierer-Meinhardt model of morphogenesis." *Journal of Mathematical Biology*, 6 :pp 317-331.

Havlena J A and Gross G W. 1988. "Climatic and tectonic controls on a semiarid ephemeral draiange system." *EOS*, 69(16) :p 363.

Henderson F M. 1966. *Open channel flow.* New York: MacMillan.

Hirsch M W and Smale S. 1974. *Differential equations, dynamical systems and linear algebra.* New York: Academic.

Holden A V. 1986. *Chaos.* Princeton: Princeton.

Horton R E. 1945. "Erosional development of steams and their drainage basins; hydrophysical approach to quantitative morphology." *Bulletin of the Geological Society of America*, 56 :pp 275-370.

Howard A D. 1971. "Simulation of stream networks by headward growth and branching." *Geographical Analysis*, 3 :pp 29 - 50.

Huggett R J. 1988. "Dissipative systems: Implications for geomorphology." *Earth Surface Processes and Landforms*, 13 : pp 45 - 49.

Iverson R M and LaHusen R G. 1988. "Dynamic pore-pressure fluctuations in rapidly shearing granular media: theory." *EOS*, 69(44) :p 1195.

Keiley A D, Malin M C and Nielson G M. 1988. "Terrain simulation using a model of stream erosion." *Computer Graphics, ACM SIGGRAPH*, 22(4) :pp 263 - 268.

Kirkby M J. 1971. "Hillslope process-response models based on the continuity equation." in *Slopes: form and process*, pp 15 - 30. Institute of British Geographers special Publications, 3, London: Institute of British Geographers.

___. 17 Nov 1988. "Hillslopes and hollows." *Nature*, 336 :p 201.

Kolb M, Botet R, Jullien R and Herrmann H J. 1986. "Flocculation and gelation in cluster aggregation." Stanley H E and Ostrowsky N, On growth and form. Fractal and non-fractal patterns in physics, Institut d'Etrudes Scientifiques de Cargèse, Corsica, France, June 1985.

Leopold L B and Langbein W B. 1962. *The concept of entropy in landscape evolution.* USGS: U.S. Geological Survey Professional papers 500-A.

Leopold L B, Wolman M G and Miller J P. 1964. *Fluvial processes in geomorphology.* London: Freeman.

Luke J C. 1974. "Special solutions for nonlinear erosion problems." *Journal of Geophysical Research*, 79(26) :pp 4035 - 4040.

Mandelbrot B B. 1983. *The fractal geometry of nature.* New York: Freeman.

McDermott G E and Pilgrim D H. 1983. "A design flood method for arid western New South Wales based on bankfull estimates." *Transactions of Institution of Engineers (Australia)*, CE25(2).

McHugh J and Prestergaard K. 1988. "Subsurface and surface flow in a discontinuous gully system." *EOS*, 69(16) :p 362.

Meakin P. 1986. "Computer simulation of growth and aggregation processes." Stanley H E and Ostrowsky N, On growth and form. Fractal and non-fractal patterns in physics, Institut d'Etrudes Scientifiques de Cargèse, Corsica, France, June 1985.

Meinhardt H A. 1976. "Morphogenesis of Lines and Nets." *Differentiation*, 6 :pp 117-123.

____. 1982. *Models of biological pattern formation*. Berlin: Academic Press.

Melton M A. 1957. *An analysis of the relations amoung elements of climate surface properties and geomorphology*, Office of Naval Research.

Mesa O J. 1986. *Analysis of channel networks parameterized by elevation*. Ph.D thesis, University of Mississippi.

Milly C and Eagleson P S. 1987. "Effects of spatial variability on annual average water balance." *Water Resources Research*, 12(11) :pp 2135 - 2143.

____. 1988. "Effect of storm scale on surface runoff volume." *Water Resources Research*, 24(4) :pp 620 - 624.

Mitchison G J. 1980. "A model for vein formation in higher plants." *Proceedings of the Royal Socity of London: Series B*, 207 :pp 79-109.

Mock S J. 1971. "A classification of channel links in stream networks." *Water Resources Research*, 7(6) :pp 1558-1566.

Montgomery D R and Dietrich W E. 17 November, 1988. "Where do channels begin?" *Nature*, 336 :pp 232 - 234.

Moon F C. 1987. *Chaotic Vibrations: An introduction for applied scientists and engineers*. New York: Wiley.

Moore I D and Burch G J. 1986. "Sediment transport capacity of sheet and rill flow : Application of unit stream power theory." *Water Resources Research*, 22(8) :pp 1350 - 1360.

Morisawa M. 1964. "Development or drainage systems on an upraised lake floor." *American Journal of Science*, 262 :pp 340 - 354.

Nicolis G and Prigogine I. 1977. *Self organization in non-equilibrium systems. From dissipative structures to order through fluctuations.* NY: Wiley.

Nittmann J, Daccord G and Stanley H E. 14 March, 1985. "Fractal growth of viscous fingers: quantitative characterization of a fluid instability phenomenon." *Nature,* 314 :pp 141 - 144.

___. 1986. "When are viscous fingers fractal." Fractals in Physics, Ed Pietronero L and Tosatti E. pp 193 - 202. New York: Elsevier.

O'Loughlin E M. 1981. "Saturation regions in catchments and their relations to soil and topographic properties." *Journal of Hydrology,* 53 :pp 229 - 246.

___. 1986. "Prediction of surface saturation zones in natural catchments by topographic analysis." *Water Resources Research,* 22(5) :pp 794 - 804.

Parker R S. 1977. *Experimental study of drainage basin evolution and its hydrologic implications,* Ph.D thesis, Colorado State University, Fort collins, Colorado.

Pilgrim D H. 1977. "Isochrones of travel time and distribution of flood storage from a tracer study on a small watershed." *Water Resources Research,* 13(3) :pp 587 - 595.

___. 1983. "Some problems in transferring relationships between small and large drainage basins and between regions." *Journal of Hydrology,* 65 :pp 49 - 72.

Priest R F, Bradford J M and Spomer R G. 1975. "Mechanisms of erosion and sediment movement from gullies." Proceedings of the Sediment Yield Workshop, USDA Sedimentation Laboratory, Oxford, Mississippi, 28 - 30 November, 1972.

Price R K. 1975. "Flood Routing Methods." Flood Studies Conference, London, 1975.

Rodriguez-Iturbe I and Escobar L A. 1982. "The dependence of drainage density on climate and geomorphology." *Hydrological Sciences - Journal - des Sciences Hydrologiques,* 27(2) :pp 129-137.

Rodriguez-Iturbe I, Gonzalez-Sanabria M, and Bras R L. 1982. "A geomorphoclimatic theory of the instantaneous unit hydrograph." *Water Resources Research,* 18(4).

Rodriguez-Iturbe I and Valdes J B. 1979. "The geomorphologic structure of hydrologic response." *Water Resources Research*, 15(6) :pp 1409-1420.

Roth G. 1986. Sulla formazione di un reticolo dendritico su superfici erodibili: Un modello analitico a base idrodinamica. Ph.D thesis, Universita di Firenze, Universita di Genova, Universita di Padova, Genoa.

Roth G, Siccardi F and Rosso R. 1989. "Hydrodynamic description of the erosional development of drainage patterns." *Water Resources Research*, 25(2) :pp 319 - 332.

Rothman D H and Gunstensen A K. 1988. "Nonlinear numerical modelling of two phase flow." *EOS*, 69(44) :p 1437.

Scheidegger A E. 1970. *Theoretical geomorphology*. Berlin: Springer Verlag.

____. 1987. *Systematic geomorphology*. NY: Springer-Verlag.

Schumm S A. 1956. "Evolution of drainage systems and slopes in badlands at Perth Amboy, New Jersey." *Bulletin of the Geological Society*, 67 :pp 597-646.

Schumm S A, Mosley M P and Weaver W E. 1987. *Experimental fluvial geomorphology*. New York: Wliey.

Segel L A. 1984. *Modeling dynamic phenomena in molecular and cellular biology*. Cambridge: Cambridge University Press.

Shaw G H and Mooers H D. 1988. "Development of drainage networks: A numerical model." *EOS*, 69(16) :p 362.

Shreve R L. 1966. "Statistical law of stream numbers." *Journal of Geology*, 74 :pp 17-37.

____. 1967. "Infinite topologically random channel networks." *Journal of Geology*, 75 :pp 178-186.

____. 1969. "Stream lengths and basin areas in topologically random channel networks." *Journal of Geology*, 77 :pp 397-414.

____. 1974. "Variation of mainstream length with basin area in river networks." *Water Resources Research*, 10(6) :pp 1167-1177.

Smart J S. 1967. "A comment on Horton's law of stream number." *Water Resources Research*, 3(3) :pp 773 - 776.

___. 1968. "Statistical properties of stream lengths." *Water Resources Research*, 4 :pp 1001 - 1014.

___. 1969. "Distribution of interior link lengths in natural channel networks." *Water Resources Research*, 5 :pp 1337 - 1342.

___. 1972. "Quantitative characterization of channel network structure." *Water Resources Research*, 8(6) :pp 1487-1496.

___. 1981. "Link lengths and channel network topology." *Earth Surface Processes and Landforms*, 6 :pp 77-79.

Smart J S, Surkan A J and Considine J P. 1967. "Digital simulation of channel networks." *IAHS-AISH*, 75 :pp 87-98.

Smith T R and Bretherton F P. 1972. "Stability and the conservation of mass in drainage basin evolution." *Water Resources Research*, 8(6) :pp 1506 - 1529.

Stanley H E. 1986. "Form: An introduction to self-similarity and fractal behavior." Stanley H E and Ostrowsky N, On growth and form. Fractal and non-fractal patterns in physics, Institut d'Etrudes Scientifiques de Cargèse, Corsica, France, June 1985.

Strahler A N. 1952. "Hypsometric (area-altitude) analysis of erosional topography." *Bulletin of the Geological Society of America*, 63 :pp 1117 - 1142.

___. 1958. "Dimensional analysis applied to fluvially eroded landforms." *Bulletin of the Geological Society of America*, 69 :pp 279 - 3000.

___. 1964. "Quantitative geomorphology of drainage basins and channel networks." *Handbook of applied hydrology*, Ed. Chow V T. pp 4-39 - 4-76. New York: McGraw Hill.

Surkan A J. "Synthetic hydrographs: Effects of network geometry", *Water Resources Research*, 5(1), :pp 112 - 128

Tarboton D G, Bras R L and Rodriguez-Iturbe I. 1988. "The fractal nature of river networks." *Water Resources Research*, 24(8) :pp 1317 - 1322.

____. 1988. "Scale versus scaling in river basin geomorphology." *EOS*, 69(16) :p 344.

____. 1989. "Scaling and elevation in river networks." *Water Resources Research*, :submitted.

Thornes J B. 1983. "Evolutionary Geomorphology." *Geography*, :pp 225 - 235.

Toffoli T and Margolus N. 1987. *Cellular Automata Machines*. Boston: MIT.

Tokunaga E. 1978. "Consideration on the composition of drainage networks and their evolution." Tokyo Metropolitan University.

Troutman B M and Karlinger M R. 1984. "On the expected width function for topologically random channel networks." *Journal of Applied Probability*, 22 :pp 836 - 849.

van der Tak L D. 1988. *Part 1: Stream length distributions, hillslope effects and other refinements of the Geomorphologic IUH. Part 2: Topologically random channel networks and Horton's laws: The Howard network simulation model revisited.* Engineers thesis, Massachusetts Institute of Technology, Massachusetts.

Vanoni V A. 1975. *Sedimentation Engineering*. New York: ASCE.

Witten T A and Sander L M. 9 November, 1981. "Diffusion-limited aggregation, a kinetic critical phenomenon." *Physical Review Letters*, 47(19) :pp 1400 - 1403.

Woldenburg M J. 1966. "Horton's laws justified in terms of allometric growth and steady state in open systems." *Bulletin of the Geological Society of America*, 77 :pp 431 - 434.

Wolman M G and Miller J P. 1960. "Magnitude and frequency of forces in geomorphic processes." *Journal of Geology*, 68 :pp 54-74.

Wong T H F and Laurenson E M. 1983. "Wave speed - discharge relations in natural channels." *Water Resources Research*, 19 :pp 701 - 706.

Yang C T. 1971. "Potential energy and stream morphology." *Water Resources Research*, 7(2) :pp 311-322.

Zavoianu I. 1985. *Morphometry of drainage basins*, Developments in Water Science, No 20, Amsterdam: Elsevier.

Zharkov S M. 1987. "Progradation and lateral filling complexes of sedimentational traps." *Transactions of the USSR Academy of Sciences - Earth Sciences section*, 292(1) : pp 61-63.

APPENDIX A

IMPLEMENTATION DETAILS OF
THE "BRANCH" NETWORK SIMULATION CODE AND
THE "FEOUT" NETWORK ANALYSIS CODE

A.1 Introduction

This appendix describes the numerical and algorithmic details of the two computer codes that have been developed for this work. The first and most important computer code is the numerical implementation of the physically based channel network and hillslope evolution model described in Chapter 5. The solution of these equations is very difficult because of their strong nonlinearity and numerical stiffness. A detailed discussion of the adopted solution algorithm is presented below.

The second computer code is one that analyzes the output of the network simulation code. The large amount of data generated by the network simulation model makes it very difficult to satisfactorily analyze the data by hand. The code is sufficiently general to have uses for catchment analysis quite apart from the analysis of the simulation data herein.

A.2 The Branch Network and Catchment Simulation Code

A.2.1 Introduction

The code is written in FORTRAN 77 and the latest version of the code is V6.30. It exists in two forms, one optimized for running on a standard serial computer (e.g., VAX, Macintosh) and one optimized for running on a CRAY X–MP vector supercomputer. Unless otherwise noted, all data in this thesis were generated with the vectorized version V6.28. The differences between versions V6.28 and V6.30 are cosmetic.

The variable identifiers in the code and their correspondence to parameters in

the model detailed in Chapter 5 are given in Table A.1.

TABLE A.1

Model Parameters and Computer Variables

Computer Variable	Model Parameters
z	z
Y	Y
pore	n
Dz	D_z
b1	β_1
m1	m_1
n1	n_1
b5	β_5
m5	m_5
n5	n_5
dtime	d_t
otime	0_t
b3	β_3
n3	n_3
notch	c_0

To solve for elevation and channelization, a catchment shape is chosen and it is subdivided by a square grid, equally spaced in both directions. The node points of the grid are the points at which the elevation and channelization function are determined. The code has some limited capability for modeling irregular boundaries, but none of the data presented in this work have used this capability (Figure A.1).

The code is written so that the timescale of elevation change, T, (Equation 6.4) is unity. All variables involving time dependence are defined relative to this timescale.
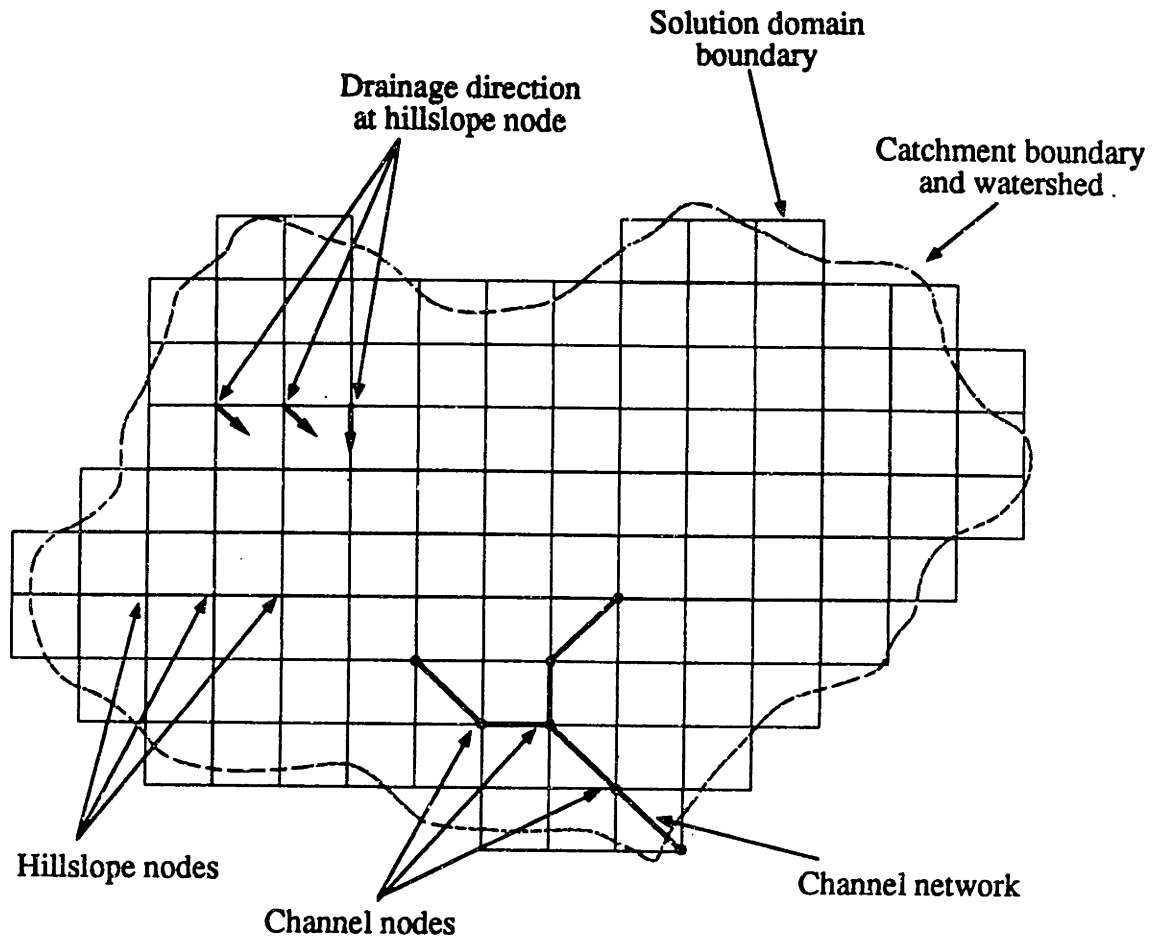
Figure A.1: Definition of terms used in the
solution algorithm

The timestep used in the solution of elevations is defined by an input parameter as a proportion of this timescale. In all results presented here the time indicated is time nondimensionalized by the timescale T; in the code T is referred to, somewhat erroneously, as the timestep. In addition, the code assumes that the grid spacing in each direction is unity so that L_g in Equation (5.1a) is unity for all the simulations. All length scales quoted in this work are, unless otherwise stated, in grid length units.

A number of different tectonic uplift regimes are allowed. All tectonic uplifts are defined relative to the elevation of the catchment outlet. Either continuous with time uplift or no uplift are possible with the spatial distribution being either uniform in space or linearly varying along one of the principle directions of the grid. When combined with the capability of the code to output data files listing all states and parameters at any specified time and the ability to restart from those files using modified parameters (e.g. tectonic uplift) this creates a flexible capability for modeling both time and space varying tectonic uplift. Note that a spatially uniform tectonic uplift with a fixed notch elevation is logically equivalent to a lowered notch elevation with no tectonic uplift; only the datum from which the elevations are defined changes.

A.2.2 Determination of the Discharge and Local Slope

To begin the calculations initial elevations are assigned to the grid and the catchment assigned an initial pattern of channelization. All the simulations of this thesis use as initial conditions a flat surface for elevations with no channelization. A node that is the outlet to the catchment is chosen and its elevation fixed. Tectonic uplift is applied to the remaining nodes within the catchment. Very small random perturbations to the initial elevations are applied with a typical value being 0.25% of the initial tectonic uplift.

Using this elevation information, the computer code determines a drainage direction at each node. A node may only drain into one of the 8 nodes directly

adjacent to it. The drainage direction determines the node to which all the flow drains. Broadly speaking, the node drains in the steepest direction though the exact means of determining this steepest direction varies as follows. If the node is a channel, then the drainage direction is towards the node with steepest downhill slope constrained so that

1. Drainage is only allowed into another channel. Channels cannot drain to hillslope nodes.

2. If a channel drains across a diagonal, as shown in Figure A.2(a), then the value of the slope in the diagonal direction is defined as that slope to the channel rather than that to the node diagonally opposite. This is consistent with the catchment surface being approximated by a triangular finite element with the opposite side having the elevations in the channel and the apex having the elevation at the node. The slope is thus the diagonal slope to the channel. In Figure A.2b it is clear that any drainage to point A ultimately ends up at node labelled B. To avoid the potential for drainage directions that cross, the drainage direction at the node is defined as being directly to node B, rather than diagonally.

If the node is a hillslope node, then the drainage direction is to the adjacent node with the steepest downhill slope. If a channel cuts across the diagonal as in Figure A.2(a), the slope and drainage direction algorithm of point 2 above is used.

These rules for determining the slopes and drainage directions ensure that

1. Crossing channels or crossing drainage directions do not occur.

2. The restriction of channel nodes to draining into other channel nodes reduces the effects of small numerical instabilities.

3. The channel network and hillslope flows represent a good approximation to the flow in the direction of steepest slope.

The contributing area at each point is determined by summing all the areas of
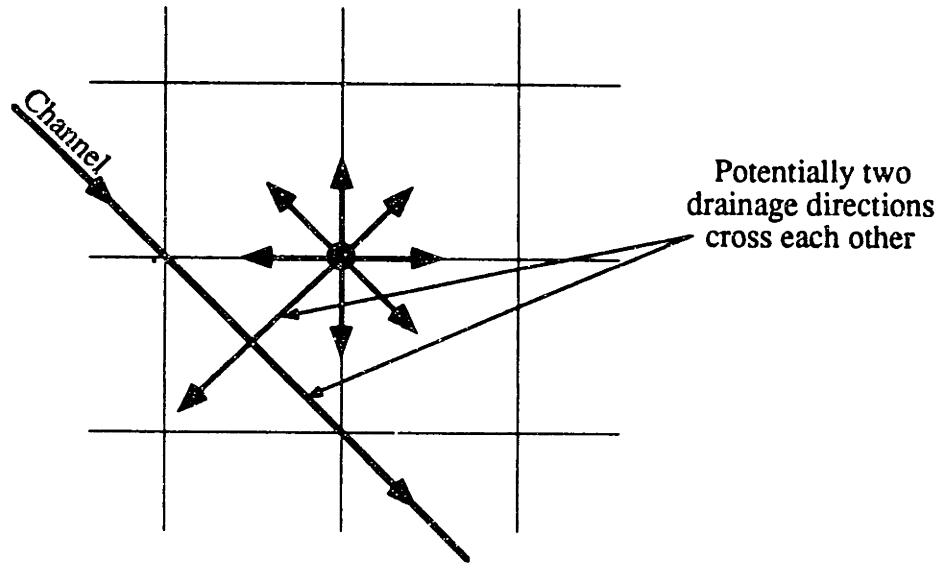
(a) Potential drainage direction
and channel conflict

Potentially two
drainage directions
cross each other

Direction of steepest downhill slope
(direction in which assigned slope is determined)

Assigned direction of flow

A

B

Triangular element interpolation
of elevations
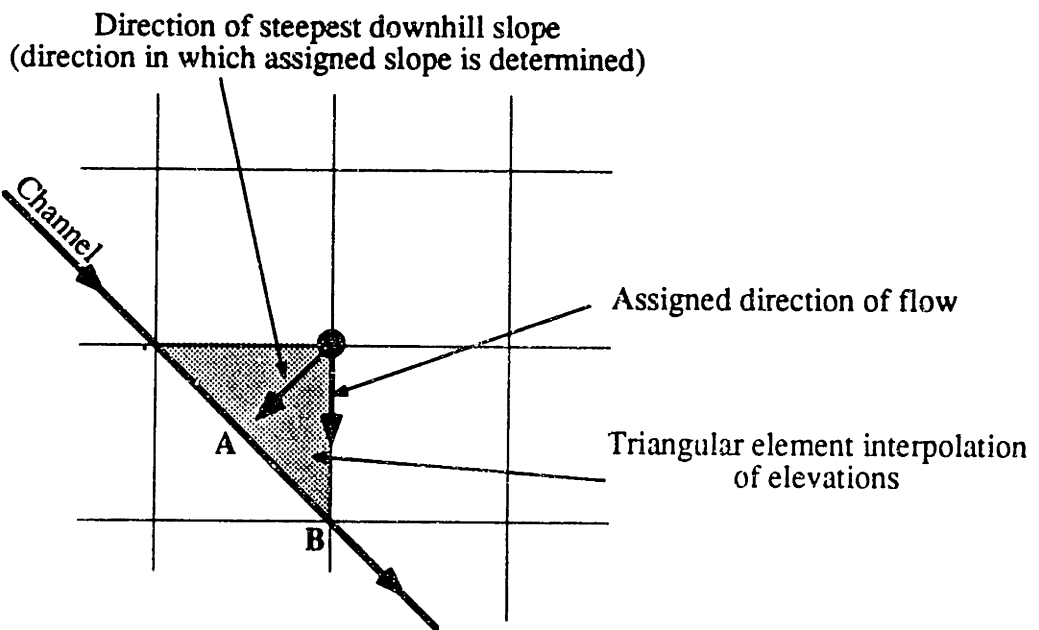
(b) Drainage direction conflict
resolution strategy

Figure A.2: Slope and drainage direction determination
in the presence of diagonal channels

the nodes contributing to that node. The algorithm to do this is based on the equation (see Figure A.3)

$$A_j = \sum_{i=1}^{8} (I_{ij} A_i) + 1$$

where $\quad A_i \quad$ = the area of the adjacent node i

$$I_{ij} = \begin{cases} 1 \text{ if node i drains to node j} \\ 0 \text{ if node i does not drain to node j} \end{cases}$$

This equation is solved iteratively by relaxation. Discharge is related to area by the expression

$$Q = \beta_3 A^{m_3}$$

In this work only a value of $m_3 = 1$ has been used.

A2.3 Solution of the Sediment Transport Continuity Equation for Elevation

Elevation changes result from imbalances in sediment transport. If more sediment enters a node than leaves, then the elevation at that node increases and vice versa. The balance of sediment transport at a node is determined by evaluating the sediment transport into that node and subtracting the sediment transport out of that node. The sediment transport equation, which is given by (variables are defined in Section 5.2)

$$Q_s = f(Y) Q^{m_1} S^{n_1}$$

is evaluated at every node. The discharge used is that discharge based on area

390

Node j for which area
is being determined
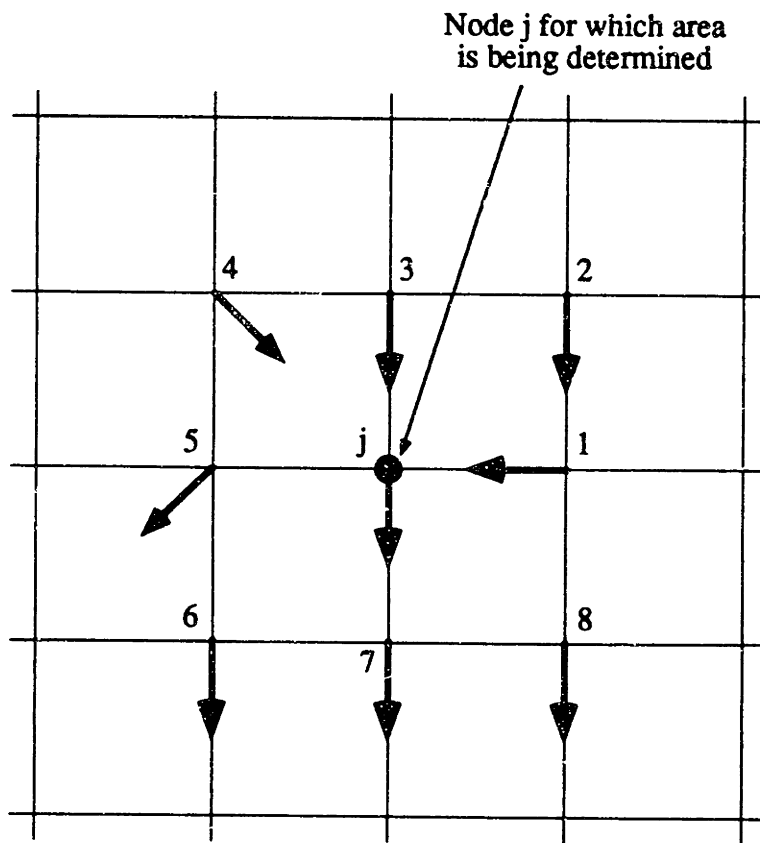
Channelisation Indicator Function Values

$$I_{1j} = I_{3j} = I_{4j} = 1$$

$$I_{2j} = I_{5j} = I_{6j} = I_{7j} = I_{8j} = 0$$

Figure A.3: Area Determination

determined in the previous section and the slope used is the local slope in the steepest direction as determined in the previous section. The rate of change of elevation at a node due to fluvial sediment transport alone is then given by

$$\frac{\partial z_j}{\partial t} = \frac{1}{\rho_s(1-n)\Delta x_1 \Delta x_2} \sum_{i=1}^{N} I_{ij} \, f(Y_i) \, Q_i^{m_1} \, S_i^{n_1} \tag{A.1}$$

where $\quad I_{ij} \quad = \begin{cases} 1 & \text{if node } i \text{ drains into node } j \\ 0 & \text{if node } i \text{ does not drain into node } j \\ -1 & \text{when } i = j \end{cases}$

$\quad\quad\quad\quad N \quad$ = number of nodes in the grid

$\quad\quad\quad\quad n \quad$ = porosity of sediment

$\quad\quad\quad\quad \rho_s \quad$ = density of sediment

$\quad\quad\quad\quad \Delta x_k \quad$ = grid spacing in the k direction = L_g

Diffusive transport is also allowed. The Fickian diffusion term is evaluated in space by a five point centered finite difference approximation so that

$$D_z \frac{\partial^2 z}{\partial x_k^2} \approx D_z \left[\frac{z_{i+1,j} + z_{i,j+1} + z_{i-1,j} + z_{i,j-1} - 4z_{i,j}}{\Delta x_1 \Delta x_2} \right]$$

$$\tag{A.2}$$

where $\quad\quad D_z \quad$ = diffusivity

$\quad\quad\quad\quad z_{i,j} \quad$ = elevation at the node with the (x_1, x_2) coordinates equal to (i,j)

$\quad\quad\quad\quad \dfrac{\partial^2 z}{\partial x_k^2} = \dfrac{\partial^2 z}{\partial x_1^2} + \dfrac{\partial^2 z}{\partial x_2^2} \quad$; Einstein implied summation convention

Equations (A.1) and (A.2) are solved in time by an explicit finite difference scheme. An explicit scheme was adopted over an implicit scheme, despite the poorer stability properties, because of the difficulty of solving for the drainage directions,

area, and discharges implicitly. Comparison of the adopted scheme, described below, with an implicit scheme of Roth (1986), used for hillslope simulation, indicates similar computational burdens for roughly comparable problems. A number of different numerical techniques for solving the elevation equation were tested. The two most successful techniques, described below, are related; the second is an extension of the first. They are both based on a two point predictor–corrector scheme (Acton, 1970).

The first scheme uses linear extrapolation for the predictor step, and parabolic interpolation for the corrector step. In this scheme the predictor equation is given by

$$\underline{z}^P(t_0+\Delta t) = \underline{z}(t_0) + \Delta t \left. \frac{\partial \underline{z}}{\partial t} \right|_{\underline{z}=\underline{z}(t_0), \underline{Y}=\underline{Y}(t_0)}$$

$$\underline{Y}^P(t_0+\Delta t) = \underline{Y}(t_0) + \Delta t \left. \frac{\partial \underline{Y}}{\partial t} \right|_{\underline{z}=\underline{z}(t_0), \underline{Y}=\underline{Y}(t_0)} \tag{A.3}$$

The corrector step which uses the states and derivatives with time at time t_0, and the predicted derivatives with time at time $t_0+\Delta t$ interpolates those with a parabola to give the corrected states of

$$\underline{z}^C(t_0+\Delta t) = \underline{z}(t_0) + \frac{\Delta t}{2} \left[\left. \frac{\partial \underline{z}}{\partial t} \right|_{\underline{z}=\underline{z}^P(t_0+\Delta t), \underline{Y}=\underline{Y}^P(t_0+\Delta t)} \right.$$

$$\left. + \left. \frac{\partial \underline{z}}{\partial t} \right|_{\underline{z}=\underline{z}(t_0), \underline{Y}=\underline{Y}(t_0)} \right] \tag{A.4}$$

$$\underline{Y}^c(t_0+\Delta t) = \underline{Y}(t_0) + \frac{\Delta t}{2} \left[\frac{\partial \underline{Y}}{\partial t} \bigg|_{\underline{z}=\underline{z}^p(t_0+\Delta t), \underline{Y}=\underline{Y}^p(t_0+\Delta t)} \right.$$

$$\left. + \frac{\partial \underline{Y}}{\partial t} \bigg|_{\underline{z}=\underline{z}(t_0), \underline{Y}=\underline{Y}(t_0)} \right]$$

An equivalent way of writing the elevation predictor–corrector equations, that leads to the second solution strategy, is

$$\underline{z}^p(t_0+\Delta t) = \underline{z}(t_0) + \Delta\underline{z}^p \qquad\qquad (A.5)$$

$$\underline{z}^c(t_0+\Delta t) = \underline{z}(t_0) + 0.5 \left(\Delta\underline{z}^p + \Delta\underline{z}^c \right) \qquad\qquad (A.6)$$

where
$$\Delta\underline{z}^p = \Delta t \frac{\partial \underline{z}}{\partial t} \bigg|_{\underline{z}=\underline{z}(t_0), \underline{Y}=\underline{Y}(t_0)}$$

$$\Delta z^c = \Delta t \frac{\partial \underline{z}}{\partial t} \bigg|_{\underline{z}=\underline{z}^p(t_0+\Delta t), \underline{Y}=\underline{Y}^p(t_0+\Delta t)}$$

Similar expressions may be written for the channelization equation. In the reformulated problem $\Delta\underline{z}^p$ and $\Delta\underline{z}^c$ are the predicted and corrected changes in elevation from the original elevations, $z(t_0)$. In the equations above we use linear extrapolation and parabolic interpolation to determine these elevation changes. Practically, if we can extrapolate and interpolate more accurately then the elevation changes from this new method may be used as the elevation changes in Equations (A.5) and (A.6). The derivation that follows will develop a nonlinear extrapolation method that is an approximation to the exact solution of Equation (A.1). This nonlinear solution will be used to predict the changes in elevation on the basis of the physical problem, rather than using simple linear extrapolation. The new solver has significant stability advantages and is particularly good at alleviating the numerical

stiffness problems of the sediment transport equation. This technique was used to produce the simulation results in this work.

An underlying assumption of the following derivations is that the dominant term in the elevation equation is the sediment transport term, with diffusion considered to be a second—order effect. Stiffness in the elevation equation arises because the area, and thus the discharge, at different points in the catchment varies from 1 to NM, where NM is the area of the simulated catchment on an N x M grid. At nodes near the outlet of the catchment (i.e., nodes with large contributing area and thus discharge) small fluctuations in elevations are corrected very quickly because of the large discharge in the sediment transport equation. Near the upper reaches of the catchment fluctuations where areas are less perturbations in elevation are corrected more slowly. Using the scheme of Equations (A.3) and (A.4) requires small time steps both for numerical stability and to capture the detail of these fluctuations in elevations.

Instead of using the linear extrapolation for elevation of Equations (A.3) and (A.4), an approximate analytic solution to Equation (A.1) is used for both the extrapolation of the predictor step and the interpolation of the corrector step.

Consider the sediment continuity at node j. Equation (A.1) can be written as

$$\frac{\partial z_j}{\partial t} = \frac{1}{1-n}\left[\sum_{\substack{i \\ i \neq j}} f(Y_i)\, I_{ij}\, Q_i^{m_1}\, S_i^{n_1} - f(Y_j)\, Q_j^{m_1}\, S_j^{n_1}\right] \qquad (A.7)$$

where the summation is over all the nodes in the grid. The local slope for node i, a node flowing into node j, can be expressed as

$$S_i = \frac{z_i - z_j}{\ell_{ij}}$$

where ℓ_{ij} = distance between nodes i and j

For node j, where flow is out of the node, the local slope can be expressed as

$$S_j = \frac{z_j - z_k}{\ell_{kj}}$$

where k = node that node j drains into

Substituting these results into Equation (A.7) yields

$$\frac{\partial z_j}{\partial t} = \frac{1}{1-n} \left[\sum_{\substack{i \\ i \neq j}} f(Y_i) \, I_{ij} \, Q_i^{m_1} \left[\frac{z_i - z_j}{\ell_{ij}} \right]^{n_1} - f(Y_j) \, Q_j^{m_1} \left[\frac{z_j - z_k}{\ell_{jk}} \right]^{n_1} \right]$$

(A.8)

This is a system of NM first order nonlinear differential equations in z_j. This system of equations is too difficult to solve as it stands. So we look at one node in isolation, node j with elevation z_j at time t_0. We assume that for the period of one timestep (i.e. time from t_0 to $t_0 + \Delta t$) that all the elevations of the surrounding nodes do not change, nor do the flow directions, so that I_{ij} and Q_i do not change. Under these assumptions there exists an equilibrium elevation for node j, z_j^*, towards which $z_j(t)$ will tend with time. So as time tends to infinity, we get

$$\frac{1}{1-n} \left[\sum_{\substack{i \\ i \neq j}} f(Y_i) \, S_{ij} \, Q_i^{m_1} \left[\frac{z_i - z_j^*}{\ell_{ij}} \right]^{n_1} - f(Y_j) \, Q_j^{m_1} \left[\frac{z_j^* - z_k}{\ell_{jk}} \right]^{n_1} \right] = 0$$

(A.9)

We desire to know

1. The value of the equilibrium elevation, z_j^*

2. The rate at which $z_j(t)$ tends to z_j^* so that we may predict $z_j(t_0+\Delta t)$, where Δt is the timestep.

It is shown in Section A2.4 that the differential equation for $z_j(t)$ in Equation (A.8) can be approximated by using the elevations at time $t = t_0$ as the initial conditions

$$\frac{\partial z_j}{\partial t} = \frac{\partial z_j}{\partial t}\bigg|_{t=t_0} \frac{(z_j(t) - z_j^*)^{n_1}}{(z_j(t_0) - z_j^*)^{n_1}} \quad \text{for } t \geq t_0 \qquad (A.10)$$

Solution of this equation yields the solution for $z_j(t_0 + \Delta t)$ as

$$z_j(t_0+\Delta t) = z_j^* + \left[z_j(t_0)-z_j^*\right]^{\frac{n_1}{n_1-1}} \left[z_j(t_0)-z_j^* + \frac{\partial z_j}{\partial t}\bigg|_{t=t_0}(1-m)\Delta t\right]^{\frac{1}{1-n_1}} \qquad (A.11)$$

This equation gives the elevation at a node j, z_j, for time $t \geq t_0$. The solution, however, requires knowledge of the equilibrium elevation z_j^*. This elevation, z_j^*, is estimated from the elevations at time $t=t_0$. To find this equilibrium elevation z_j^*, we desire the solution, in terms of z_j^*, of

397

$$\frac{\partial z_j}{\partial t}\bigg|_{z=z_j^*} = \frac{1}{1-n}\left[\sum_{\substack{i \\ i \neq j}} f(Y_i)\, I_{ij}\, Q_i^{m_1}\left[\frac{z_i - z_j^*}{\ell_{ij}}\right]^{n_1}\right.$$

$$\left. -\, f(Y_j)\, Q_j^{m_1}\left[\frac{z_j^* - z_k}{\ell_{jk}}\right]^{n_1}\right] = 0$$

To determine z_j^* this expression is approximated by a Taylor series (expanded to linear terms), around the elevation at time t_0, $z_j(t_0)$, so that

$$\frac{\partial z_j}{\partial t} \approx \frac{\partial z_j}{\partial t}\bigg|_{z_j = z_j(t_0)} + \left[z_j - z_j(t_0)\right]\left[\frac{\partial}{\partial z_j}\left(\frac{\partial z_j}{\partial t}\right)\right]\bigg|_{z_j = z_j(t_0)} = 0$$

(A.12)

Solving for the equilibrium solution $z_j = z_j^*$ yields the estimate of z_j^* based on conditions at $t = t_0$

$$z_j^* = \frac{z_j(t_0)\left[\frac{\partial}{\partial z_j}\left(\frac{\partial z_j}{\partial t}\right)\right]\bigg|_{z_j = z_j(t_0)} - \frac{\partial z_j}{\partial t}\bigg|_{z_j = z_j(t_0)}}{\left[\frac{\partial}{\partial z_j}\left(\frac{\partial z_j}{\partial t}\right)\right]\bigg|_{z_j = z_j(t_0)}}$$

(A.13)

The derivative with respect to z_j can be determined from Equation (A.8) and is given by

$$\frac{\partial}{\partial z_j}\left(\frac{\partial z_j}{\partial t}\right) = \frac{1}{1-n}\left[\sum_{\substack{i\\i\neq j}}\frac{f(Y_i)I_{ij}Q_i^{m_1}}{\ell_{ij}}\left[\frac{z_i-z_j}{\ell_{ij}}\right]^{n_1-1} + \frac{f(Y_j)Q_j^{m_1}}{\ell_{jk}}\left[\frac{z_j-z_k}{\ell_{jk}}\right]^{n_1-1}\right]$$

$$= \frac{1}{1-n}\left[\sum_{\substack{i\\i\neq j}}\frac{f(Y_i)I_{ij}Q_i^{m_1}S_i^{n_1-1}}{\ell_{ij}} + \frac{f(Y_j)Q_j^{m_1}S_j^{n_1-1}}{\ell_{jk}}\right]$$

$$(A.14)$$

The solution for z_j^* in Equation (A.13) and (A.14) is an approximation. How good the approximation is depends on how good is the linearization of Equation (A.12). If $n_1=1$ the linearization is exact. However, the typical value of n_1 is about 2. A numerical experiment was performed to assess this error by iteration on Equations (A.10) and (A.11) in a loop of the form

1. Evaluate the new z_j^*, for all j, using Equations (A.13) and (A.14).

2. Using these new values of z_j^* as the new values for z_i and $z_j(t_0)$, evaluate for the new value of z_j^*. The drainage directions were not reevaluated.

3. Repeat Step 2 until convergence is observed.

It was found that convergence to a maximum relative error of 0.1% over the entire grid was typically observed within 3 to 5 iterations, for $n_1 = 2.1$. This was considered satisfactory. Note, however, that iterations are not performed within the simulation code when determining z_j^*. This decision was a trade—off between accuracy and added speed of the algorithm. The results from Equations (A.13) and (A.14) are used in the direct solution for the new elevations at each timestep.

The primary advantage of the solution technique of Equations (A.11), (A.13), and (A.14) is that it yields more stable solutions over a broader range of timescales. The Euler predictor—corrector method of Equations (A.3) and (A.4) with its linear

extrapolation has a tendency to exhibit oscillations in the elevations for large values of $\beta_1 Q^{m_1} \Delta t$. If the timestep is too large, then elevation oscillations occur in the downstream regions of the channel network. These elevation oscillations grow until eventually some downstream elevations are higher than upstream elevations. This induces oscillations in the drainage directions and thus the contributing areas and discharges. When these effects feedback into the elevation and sediment transport equation, the solution strategy fails catastrophically. The means by which these oscillations initially develop are illustrated in Figure A.4a.

The new technique based on Equations (A.11), (A.13), and (A.14) is equivalent to a two point predictor—corrector using nonlinear extrapolation. For large $\beta_1 Q^{m_1} \Delta t$, the elevation predicted is the equilibrium elevation given by Equations (A.13) and (A.14), z_j^*. For very small $\beta_1 Q^{m_1} \Delta t$, the solver converges to the two point predictor—corrector method with linear extrapolation (Equations A.3 and A.4). For intermediate values the elevations are intermediate between these two extremes. These results are a good approximation to the exact solution, and are illustrated in Figure A.4b.

Sensitivity studies on a 20 x 20 grid showed that both schemes were mass conservative for small time steps (e.g., nondimensional timestep $\Delta t = 0.001$). The new nonlinear extrapolation maintained mass conservation up to a nondimensional timestep of $\Delta t = 0.1$ for some test cases. The linear extrapolator was typically not mass conservative, nor stable, for nondimensional timesteps above $\Delta t = 0.01$. Considering the added computational burden of the nonlinear solver, it was about five times more efficient on a scalar machine, and it was equally vectorizable so that gains were not lost on the supercomputer through inefficient vectorization.

The stability threshold is dependent upon Q^{m_1} with $Q = O(NM)$ (NM = number of nodes in the grid) and $m_1 \approx 2$ (see Section 5.5). Thus for a grid twice as
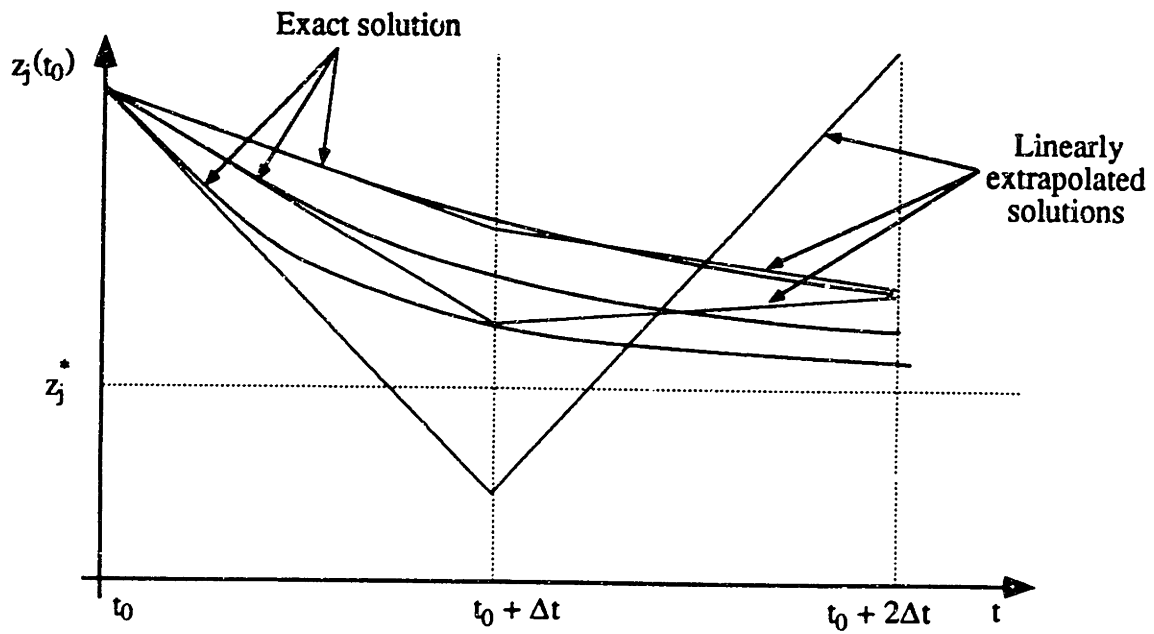
(a) The linear solver
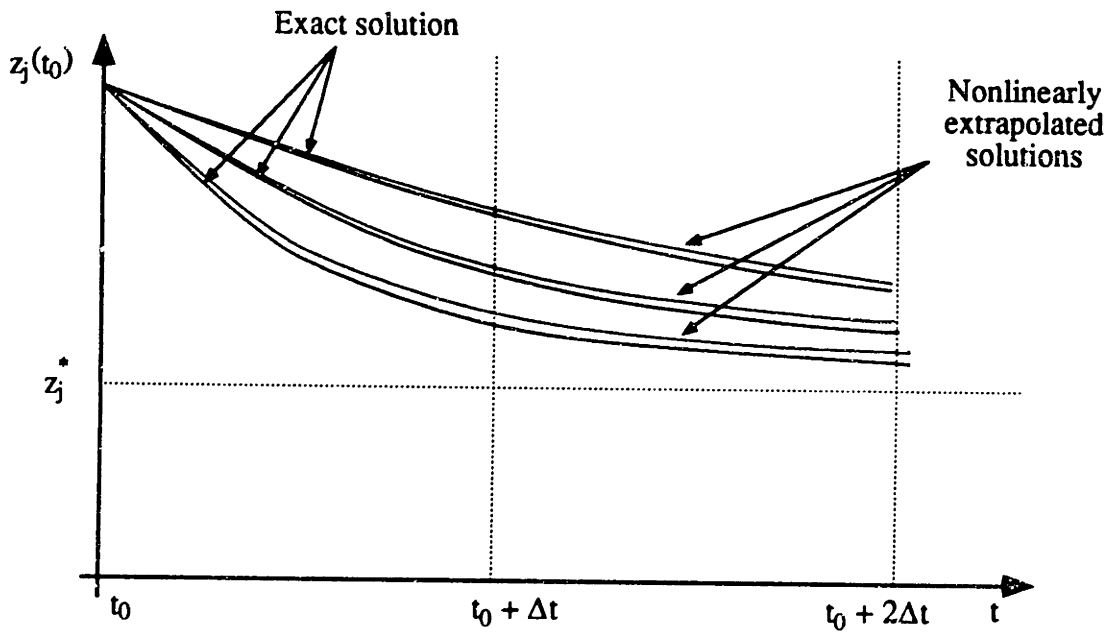
(b) The nonlinear solver

Figure A.4: Elevation with time for the linear and nonlinear solvers
: Oscillation and timescale effects

large (40 x 40), the efficiency differential of the new method is about four times that found for the 20 x 20 grid; the new method is approximately twenty times more efficient for the 40 x 40 case. For the 40 x 40 case with the nonlinear solver, the solution up to ultimate drainage density typically takes about one CPU hour on the CRAY X–MP supercomputer.

A.2.4 Derivation of the Approximation of Elevations with Time

This section will demonstrate under what conditions the approximation of Equation (A.10) is satisfactory.

$$\frac{\partial z_j}{\partial t} = \frac{\partial z_j}{\partial t}\bigg|_{t=t_0} \cdot \frac{(z_j(t)-z_j^*)^{n_1}}{(z_j(t_0)-z_j^*)^{n_1}} \quad \text{for } t \geq t_0$$

First we note that this equation is consistent with the exact solution of the equation when $t=t_0$. It is also consistent as $t \rightarrow \infty$ where $z_j(t) \rightarrow z_j^*$ and $\frac{\partial z_j}{\partial t} \rightarrow 0$. Note that for $(z_j(t)-z_j^*)$ positive, $\frac{\partial z_j}{\partial t}$ is negative and vice versa.

The general solution to the problem is too complex to present here. Rather an error analysis of a simple example will be used where there is only one inflow node and no convergence of flow. The profile of the channel is illustrated in Figure A.5. The sediment transport continuity equation for this case is
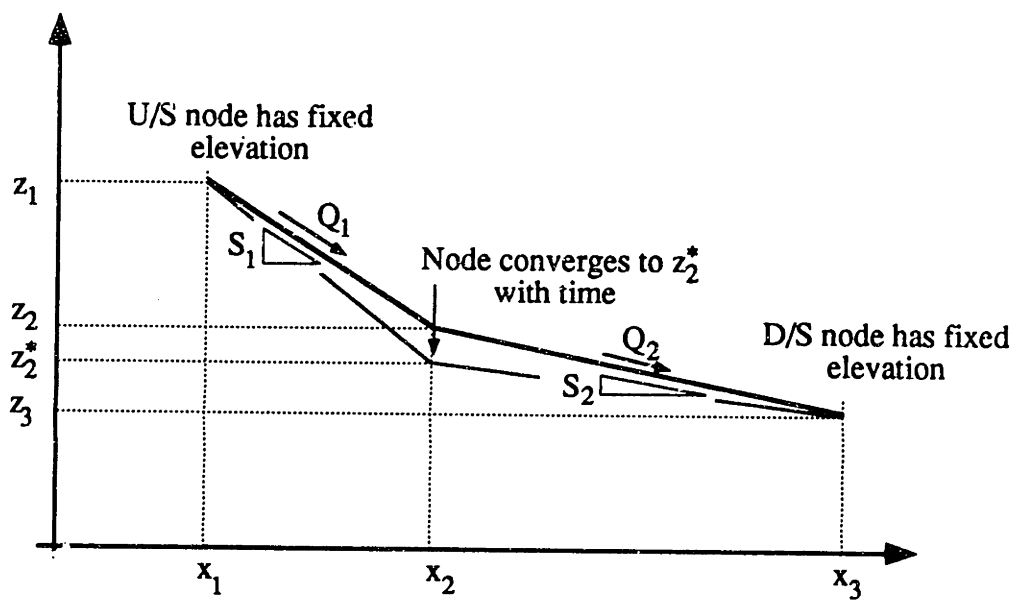
Figure A.5: Elevation changes with time:
numerical solution technique

$$\frac{\partial z_2}{\partial t} = \beta_1 \left(Q_1^{m_1} S_1^{n_1} - Q_2^{m_1} S_2^{n_1} \right) \quad \text{for } t \geq t_0$$

$$= \beta_1 \left[Q_1^{m_1} \left[\frac{(z_1(0)-z_2)^{n_1}}{\ell_1^{n_1}} \right] - Q_2^{m_1} \left[\frac{(z_2-z_3(0))^{n_1}}{\ell_2^{n_1}} \right] \right]$$

(A.15)

where z_2 = elevation of the node being examined

       Q_1, Q_2 = inflow and outflow discharges respectively

       S_1, S_2 = inflow and outflow slopes respectively

$z_1(0), z_3(0)$ = elevations at the upstream and downstream nodes respectively at

         time $t=t_0$.

       ℓ_1, ℓ_2 = distance of the upstream and downstream nodes from the node

         being examined

When $z_2=z^*$, the equilibrium elevation for node 2, then

$$Q_1^{m_1} \left[\frac{(z_1(0)-z_2^*)^{n_1}}{\ell_1^{n_1}} \right] = Q_2^{m_1} \left[\frac{(z_2^*-z_3(0))^{n_1}}{\ell_2^{n_1}} \right]$$

(A.16)

It simplifies notation below if Equations (A.10), (A.15) and (A.16) are expressed in transformed coordinates, $z'=z-z^*$, so that

$$\frac{\partial z_2'}{\partial t} = \beta_1 \left[Q_1^{m_1} \left[\frac{(z_1'(0)-z_2')^{n_1}}{\ell_1^{n_1}} \right] - Q_2^{m_1} \left[\frac{(z_2'-z_3'(0))^{n_1}}{\ell_2^{n_1}} \right] \right]$$

(A.17)

$$\frac{Q_1^{m_1} \, z_1'(0)^{n_1}}{\ell_1^{n_1}} = \frac{Q_2^{m_1} \, (-z_3'(0))^{n_1}}{\ell_2^{n_1}} \qquad (A.18)$$

$$\frac{\partial z_2'}{\partial t} = \frac{\partial z_2'}{\partial t}\bigg|_{t=t_0} \cdot \frac{z_2'^{n_1}}{z_2'(0)^{n_1}} \qquad (A.19)$$

Combining Equation (A.17) and (A.19) yields the error expression for the approximation proposed

$$E = \frac{\partial z_2'}{\partial t} - \frac{\partial z_2'}{\partial t}\bigg|_{t=t_0} \cdot \frac{z_2'^{n_1}}{z_2'(0)^{n_1}}$$

$$= \beta_1 \left[\frac{Q_1^{m_1}}{\ell_1^{n_1}} \left[(z_1'(0)-z_2')^{n_1} - \frac{z_2'^{n_1}}{z_2'(0)^{n_1}} (z_1'(0)-z_2'(0))^{n_1} \right] \right.$$

$$\left. - \frac{Q_2^{m_1}}{\ell_2^{n_1}} \left[(z_2'-z_3'(0))^{n_1} - \frac{z_2'^{n_1}}{z_2'(0)^{n_1}} (z_2'(0)-z_3'(0))^{n_1} \right] \right]$$

$$\qquad (A.20)$$

where E = the error of the approximation

If we express $z_2' = z_2'(0) + \Delta$ where Δ is the change in z_2 with time from the initial value and is considered small then using Taylors Series up to linear terms

$$(z_1'(0) - z_2')^{m_1} - (\frac{z_2'}{z_2'(0)})^{n_1} (z_1'(0) - z_2'(0))^{n_1}$$

$$= [(z_1'(0) - z_2'(0))^{n_1} + n_1 \, \Delta \, (z_1'(0) - z_2'(0))^{n_1 - 1}] \cdot$$

$$\frac{z_2'(0)^{n_1} + n_1 \, \Delta \, z_2'(0)^{n_1 - 1}}{z_2'(0)^{n_1}} [z_1'(0) - z_2'(0)]^{n_1}$$

$$= n_1 \, \Delta \, (z_1'(0) + z_2'(0))^{n_1 - 1} (\frac{z_1'(0)}{z_2'(0)})$$

$$= n_1 \, \Delta \, (\frac{z_1'(0)}{z_2'(0)} + 1)^{n_1 - 1} z_1'(0) \qquad \text{(A.21)}$$

We obtain a similar expression for the second term on the right hand size of Equation (A.20) by substituting for $z_1'(0)$ with $(-z_3'(0))$ in Equation (A.21). Substituting these results into Equation (A.20) yields

$$E = \frac{n_1 \, \Delta \, \beta_1}{(z_2'(0))^{n_1 - 1}} \left[\frac{Q_1^{m_1} z_1'(0)^{n_1}}{\ell_1^{n_1}} (1 - \frac{z_2'(0)}{z_1'(0)})^{n_1 - 1} - \right.$$

$$\left. \frac{Q_2^{m_1} (-z_3'(0))^{n_1}}{\ell_2^{n_1}} (1 - \frac{z_2'(0)}{z_3'(0)})^{n_1 - 1} \right]$$

$$\text{(A.22)}$$

This expression can be further simplified to yield

$$E = \frac{n_1 \, \Delta \, \beta_1}{\left(z_2'(0)\right)^{n_1-1}} \left(1 - \frac{z_2'(0)}{z_1'(0)}\right)^{n_1-1} .$$

$$\left[\frac{Q_1^{m_1} \, z_1'(0)^{n_1}}{\ell_1^{n_1}} - \frac{Q_2^{m_1} \, z_2'(0)^{n_1}}{\ell_2^{n_1}} - \right.$$

$$\left. \frac{Q_2^{m_1} \, (-z_3'(0))^{n_1}}{\ell_2^{n_1}} \left[\frac{z_2'(0)}{z_3'(0)}\right]^{n_1-1} \left[\frac{1 - \frac{z_1'(0)}{z_3'(0)}}{\frac{z_1'(0)}{z_3'(0)} - \frac{z_2'(0)}{z_3'(0)}} \right]^{n_1-1} \right]$$

and substituting the equilibrium conditions of Equation (A.18) the expression for the error is

$$E = \frac{n_1 \, \Delta \, \beta_1 \, z_3'(0)}{\ell_2^{n_1}} \left| \frac{z_3'(0) - z_1'(0)}{z_1'(0)} \right|^{n_1-1}$$

and for a small horizontal discretization then $z_3'(0)$ is small (since slopes are finite then $\Delta \sim S \, \Delta x$) so that the error expression is $E \sim O(\Delta x^2)$. Thus the approximation in Equation (A.10) is good and the error involved is approximately of order Δx^2 where Δx is the spatial discretization.

A.3 The FEOUT Network and Catchment Analysis Code

FEOUT is a general purpose network and catchment data analysis program. It directly uses output files from BRANCH. The latest version of this code is V1.40. The code is written in FORTRAN 77 and uses GKS for graphics output. The program is command line driven and has rudimentary on-line help. All the data analyses of this

work were performed with this code.

Some of the capabilities of the code include:

1. Strahler and Tokunaga analysis by order for stream length, slope, area and drop. Mean hillslope slopes and drops stratified on the Strahler order of the stream they drain into.

2. Link analyses for slope, contributing area and magnitude stratified on Strahler order, exterior versus interior link type, topological classification (after Mock, 1971) and magnitude.

3. Hillslope slopes, drops, mean length and source area.

4. Distributional properties (using the method of moments) of the above characteristics.

5. Catchment, hillslope and channel relief, and drainage density.

6. Fractal characteristics of the channels.

7. Saturation area properties (after O Loughlin, 1981).

8. Plotting of the planar properties of the channel networks and the contributing hillslopes.

9. Plotting of profiles of hillslopes and channels.

10. Plotting of hypsometric curves.

APPENDIX B

ONE–DIMENSIONAL ADVECTION–DIFFUSION FORMULATION
OF SEDIMENT TRANSPORT

B.1 Introduction

This section will demonstrate that it is possible to reformulate the one–dimensional form of the overland flow sediment transport equation, as presented in Chapter 5, into a one–dimensional advection–diffusion equation; that is, the spatial coupling of elevations is the result of advection and diffusion of elevation changes. The diffusion and advection coefficients in the new formulation are nonstationary and nonhomogeneous. It should be noted that the analysis that follows is only applicable to the overland flow runoff mechanism and, in particular, is not applicable to the subsurface saturation mechanism.

The advantage of this new formulation of sediment transport is that it facilitates comparison of the elevation equation developed for this work with previous published work in this area, of the kind described in Chapters 3 and 4. Almost all of these models use a homogeneous, stationary, and isotropic Fickian diffusion term for the spatial coupling of elevations. It will be shown that the subtle differences between the formulations used in previous work and that in this work cause important differences in the form of the catchments generated.

Generalization of this reformulation to two–dimensions has not been performed because of the difficulties caused by the convergence and divergence of the flow. The effect of lateral inflow on the one–dimensional formulation will be examined to demonstrate some of the simpler characteristics of two–dimensional flow.

B.2 The One–Dimensional Advection–Diffusion Formulation

This section will develop the mathematics of the advection–diffusion

formulation of sediment transport. As in Chapter 5, the sediment transport equation is of the form

$$Q_s = \beta \, Q^m \, S^n \tag{B.1}$$

It is recognized that it is highly likely the coefficients will vary systematically in space because of the natural sorting of sediment in the downstream direction (Rose, personal communication). The added complexity of including the variability in the following analysis only clouds the central issue while the simplification of ignoring it does not invalidate the conclusions.

Considering sediment transport continuity on an incremental element (x is defined in the downslope direction and z vertically upwards) then

$$\frac{dz}{dt} \Delta x = Q_s(x) - \left[Q_s(x) + \frac{dQ_s}{dx} \Delta x \right] = -\frac{dQ_s}{dx} \Delta x \tag{B.2}$$

so that the sediment transport continuity equation yields the elevation evolution equation

$$\frac{dz}{dt} = -\frac{dQ_s}{dx} \tag{B.3}$$

To determine the right–hand side of Equation (B.3), we can use the sediment transport equation (B.1) so that

$$\frac{dQ_s}{dx} = \beta \left[m \, Q^{m-1} \frac{dQ}{dx} S^n + n \, S^{n-1} \frac{dS}{dx} Q^m \right] \tag{B.4}$$

and if we collect terms and note that

410

$$S = -\frac{dz}{dx} \quad \text{and} \quad \frac{dS}{dx} = -\frac{d^2z}{dx^2} \tag{B.5}$$

then Equation (B.3) can be expressed as

$$\frac{dz}{dt} - \left[\beta\, m\, Q^{m-1}\, S^{n-1}\, \frac{dQ}{dx}\right] \frac{dz}{dx} = (n\, Q^m\, S^{n-1})\, \frac{d^2z}{dx^2} \tag{B.6}$$

This equation can be compared with the standard formulation of the advection–diffusion equation, i.e.,

$$\frac{dz}{dt} - v\, \frac{dz}{dx} = D\, \frac{d^2z}{dx^2} \tag{B.7}$$

where v = advection wave speed

           D = diffusivity

We can thus conclude that Equation (B.6) is a nonlinear form of the advection–diffusion equation with variable advection speed and variable diffusivity. These rate constants are:

$$v = \beta\, m\, Q^{m-1}\, S^{n-1}\, \frac{dQ}{dx} \quad : \text{Advection wave speed} \tag{B.8}$$

$$D = \beta\, n\, Q^m\, S^{n-1} \quad\quad\quad : \text{Diffusivity} \tag{B.9}$$

B.3 Characteristics of the Advection–Diffusion Formulation

In this section two types of overland flow will be considered: instream flows which might typically correspond to channel flows, and lateral flows which might

411

typically correspond to flows from the hillslope at right angles to the direction of instream flows. The traditional linear advection–diffusion equation with constant coefficients as in Equation (B.7) operates in two modes: advection dominated and diffusion dominated. It will be shown in this section that the advection dominated mode corresponds to conditions where lateral inflows dominate instream flows, while the diffusion dominated mode corresponds to conditions where instream flows dominate lateral inflows.

To examine the conditions under which advection dominates diffusion in Equation (B.6), consider the ratio of the advection velocity and diffusion coefficient

$$
\frac{\text{advection}}{\text{diffusion}} \approx \frac{\beta \; m \; Q^{m-1} \; S^{n-1} \; \frac{dQ}{dx}}{\beta \; n \; Q^{m} \; S^{n-1}} \tag{B.10}
$$

or

$$
\frac{\text{advection}}{\text{diffusion}} \approx \frac{m \; \frac{dQ}{dx}}{n \; Q} \tag{B.11}
$$

Since the n/m ratio is typically of the order of 1, then the advection domination occurs when

$$
\frac{\frac{dQ}{dx}}{Q} \gg 1 \tag{B.12}
$$

Since the numerator of Equation (B.12) represents the lateral inflow and the denominator represents the instream flow, then advection domination occurs when lateral inflow is an important contributor to instream flows.

It shall now be considered where in the channel network and hillslope, lateral

inflows are important in comparison to instream flows. Initially the analysis will be confined to scales larger than the first order basin. If drainage density is constant throughout the catchment, then it is reasonable to assume that the mean hillslope length is independent of the position within the catchment (Horton, 1945). The mean lateral inflow term can then be expressed as

$$\frac{dQ}{ds} = R_{hs}\ell \tag{B.13}$$

where R_{hs} = average runoff excess rate over the hillslope draining to the
position   x in the channel

ℓ = hillslope length

s   = coordinate direction along the hillslope stream line or channel.

If instream discharge is related to catchment area in an analogous fashion to Equation (B.13) then

$$Q = R_{ca}A \tag{B.14}$$

where   A   = contributing area upstream

R_{ca} = average runoff excess rate over the contributing area,

then Equation (B.11) can be expressed as

$$\frac{\text{advection}}{\text{diffusion}} \approx \frac{m}{n}\frac{R_{hs}}{R_{ca}}\frac{\ell}{A} \tag{B.15}$$

Two features are apparent in Equation (B.15):

1.   Rainfall/runoff rates only occur as a correction factor relating the average upstream runoff rate with the local runoff rate.

2.   With typical values of n, m, R_{hs}, and R_{ca}, Equation (B.15) is small for all but the most upstream sections of the first-order basins, so diffusion dominates for all except the hillslope. Other than in the hillslope region by the definitions of the hillslope length and first-order basin, it follows that A is much larger than 1.

Thus, in a one-dimensional sense, large portions of the catchment can be approximated by the equation

$$\frac{dz}{dt} = (n \ Q^m \ S^{n-1}) \frac{d^2z}{ds^2} \tag{B.16}$$

The most important feature of Equation (B.16) is that the diffusivity is not constant in space. Neither is it constant in time because continuing erosion will result in changes in slopes and discharges. This is in contrast to a model recently proposed by Shaw and Mooers (1988) where the proposed model for hillslope and channel evolution was based on the diffusion equation, i.e.,

$$\frac{dz}{dt} = \begin{cases} D_{hs} \dfrac{d^2z}{ds^2} & \text{hillslope} \\[3em] D_{ch} \dfrac{d^2z}{ds^2} & \text{channel} \end{cases} \tag{B.17}$$

where D_{hs}, D_{ch} = diffusivites for elevations in the channel and hillslope respectively

and while the diffusivities did vary from the hillslope regions to the channel regions, they were considered constant within their respective regions.

Consider now the region of the catchment where advection dominates, i.e.,

hillslope scale. The appropriate elevation equation is

$$\frac{dz}{dt} = \left[\beta \, m \, Q^{m-1} \, S^{n-1} \, \frac{dQ}{ds} \right] \frac{dz}{ds} \qquad \text{(B.18)}$$

It can be noted that the advection term in Equation (B.18) goes to zero at the watershed. This condition is true irrespective of the value of slope at the boundary provided only that it remain finite. Thus at the domain boundary

$$\frac{dz}{dt} = 0 \qquad \text{(B.19)}$$

This is not true in hilly terrain without rocky outcrops because of the dominance of soil creep, rainsplash, and rockfall near the watershed (Dunne, 1988).

This is another important difference between the sediment transport model and the model proposed by Shaw and Mooers. The authors enforce a zero slope condition at the boundary. This clearly is inconsistent with the requirements of Equation (B.19). If sediment transport is the only process operational at the boundary, then in fact a slope condition cannot be applied and a first type condition (i.e., specified elevation) is mandatory. On the other hand, the importance of a slope condition will be dependent on the relative importance of the diffusive processes of creep, rainsplash and rockfall near the boundary at the scales of modeling. This scale will be dependent on the grid resolution adopted in the numerical model. The experimental evidence seems to indicate that the scale where these processes are dominant is very small and will only be important for catchments at the hectare scale (Moore and Burch, 1988). Thus at the scales that are typically resolved by modeling, it is reasonable to assume that the appropriate boundary conditions are those for the sediment transport model rather than the diffusion model as modelled by Shaw and Mooers. In other words there are three important scales in real catchments.

1. **Large scale** (channel network scale) — Diffusion dominated (sediment transport)

2. **Medium scale** (hillslope scale) — Advection dominated (sediment transport)

3. **Small scale** (hillslope scale) — Diffusion dominated (rainsplash, creep, rockfall),

but because of the very small scale of the third process and resolution limitations of the computer model, the third process should be ignored totally by the modeling.

Problems with the pattern of catchment elevations around the watershed are apparent in the results of Shaw and Mooers (1988), and it is believed that the inappropriate boundary conditions are the problem.

B.4 Hillslope Retreat Versus Hillslope Degradation

The question of whether hillslope retreat or hillslope degradation is the dominant process in the long—term evolution of hillslope form has vexed geomorphologists for some time (e.g., Zharkov, 1987). The derivations of the previous sections begin to provide some answers to this question.

Hillslope retreat occurs when a hillslope of a constant shape or form retreats away from the valley bottom towards the watershed (Figure B.1a). This movement can be characterized by the retreat speed (m/century, etc.). Hillslope degradation occurs where the hillslope surface declines with time, slopes decline, and the watershed elevation decline. There is no characteristic hillslope form in the latter case (Figure B.1b).

If fluvial sediment transport is considered the dominant hillslope process then, for a one—dimensional hillslope at least, Equation (B.1) is applicable to describing the elevation evolution of the hillslope. The reformulation leading to Equation (B.6) suggests that hillslope evolution may be characterized by two processes.

(a) Retreat

Hillslope profile

aging

Away from the watershed

Elevation decline
at the watershed

Hillslope profile

aging
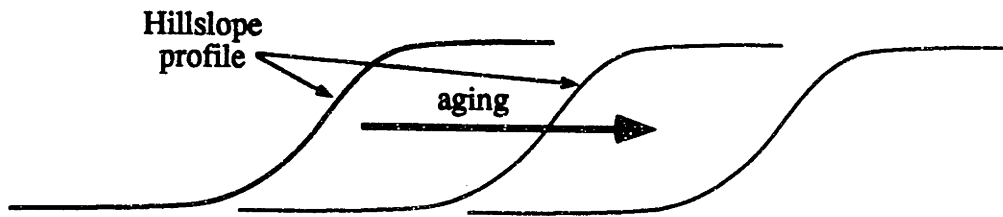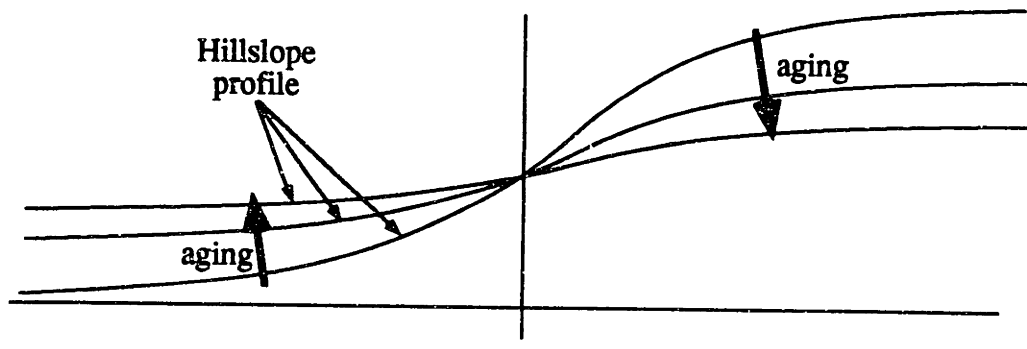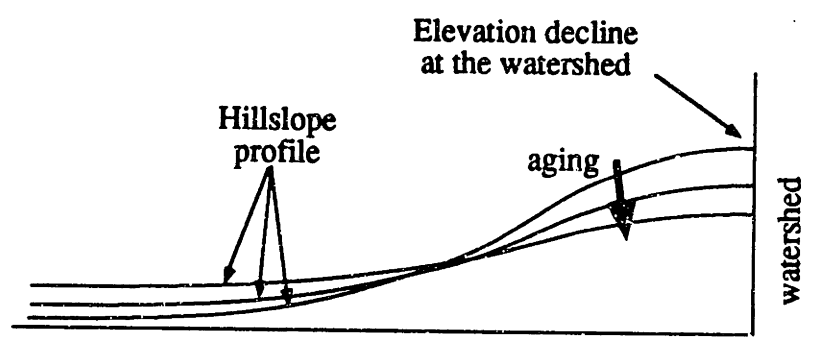
watershed

Near the watershed

(b) Degradation

Figure B.1: Hillslope retreat and hillslope
degradation illustrated

417

1. An advection speed for the constant form hillslope, analogous to hillslope retreat.

2. A diffusivity corresponding to hillslope degradation.

Thus it can be said that

$$\text{Hillslope Retreat Speed} = \beta\, m\, Q^{m-1}\, S^{n-1}\, \frac{dQ}{dx} \qquad (B.20a)$$

$$\text{Hillslope Degradation Diffusivity} = \beta\, n\, Q^{m}\, S^{n-1} \qquad (B.20b)$$

and the retreat process dominates degradation when

$$\frac{m}{n}\, \frac{d\ell nQ}{dx} \gg 1 \qquad (B.21)$$

and vice versa.

If, for the moment, variation within the hillslope is ignored, then hillslope retreat dominates when convergence of flow is strong so that rate of change of discharge is large.

The slope and discharge dependence of Equations (B.20) suggest that neither speed nor diffusivity are constant within a single hillslope, as previously noted. Thus these equations can only give average rates of degradation and retreat for the two processes, over the entire hillslope. Retreat speed, for instance will be minimum at the watershed and the channel, where slopes are low and maximum at a point slightly downstream of the steepest part of the hillslope. This effect will tend to steepen the upper parts of the hillslope and flatten the lower parts of the hillslope. Similar non–linear effects can be noted for the degradation component.

Equation (B.20a) provides a comparison with the hillslope simulation method of Shaw and Mooers (1988), which is linear diffusion, so that the transport law is

$$Q_s = D \frac{dz}{dx} = D\, S$$

Chase (1988) recently proposed an event–based model for elevation evolution based on transport, linear in discharge, so that

$$Q_s = D\, Q \frac{dz}{dx} = D\, Q\, S$$

The retreat speeds and diffusivities are given in Table B.1. It is notable that Shaw and Mooers model cannot model slope retreat at all, and relies entirely on slope degradation for hillslope form. Chase's model removes the slope dependence in the retreat and diffusivities, though the discharge dependence of Equation (B.20) remains. Thus for Chase's model retreat speed will be constant over the hillslope if the flow convergence remains the same on the hillslope. His diffusivity will be maximum at the downstream end of the hillslope rather than in the mid–section of the hillslope.

TABLE B.1

Retreat Speeds and Diffusivities

for Hillslopes

Transport	Retreat Speed	Diffusivity
$\beta Q^n S^m$	$\beta m Q^{n-1} S^{n-1} \frac{dQ}{dx}$	$\beta n\, Q^m S^{n-1}$
$D\, S$	0	D
$D\, Q\, S$	$D \frac{dQ}{dx}$	$D\, Q$

In the context of the area–slope renormalization discussed in Section 8.3 the

Shaw and Mooers transport model will give an entirely different trend to that of this work or Chase. Since their model is diffusion only they will give an increasing slope with area, Chase's model a constant slope (if runoff is proportional to area) while the fluvial transport model of this work will give decreasing slope with area. It is observed that slope decreases with increasing area for large areas and increases with area for very small areas. The reader is referred to Section 8.3 for further details.

In conclusion it has been shown that hillslope retreat rates and degradation rates may be derived, to first order, from the hillslope erosion process. The expressions obtained are simple and it may be easy to calibrate these equations from paleo—landscape data. Alternatively measurement of sediment transport process should enable an answer to be obtained to the question of which hillslope process is dominant: hillslope retreat or hillslope degradation.

APPENDIX C

THE PHYSICAL BASES OF THE SEDIMENT TRANSPORT

AND ACTIVATOR FUNCTIONS

C.1 Physical Basis for the Sediment Transport Equation

C.1.1 Introduction

The generic sediment transport formula used in this work was

$$Q_s = \beta_1 Q^{m_1} S^{n_1} \qquad \qquad (C.1)$$

where

Q_s = sediment transport, in mass/time

Q = discharge

S = channel or hillslope slope

Though a sediment transport equation of this form has been used by geomorphologists in previous work (Smith and Bretherton, 1972), this formulation of the sediment transport equation is unconventional compared with the form used by specialists in fluvial sediment transport (e.g., Vanoni, 1975). The unconventional formulation was adopted because it has a number of significant advantages computationally over the traditional formulation (see Appendix A for details). This section aims to show how the new formulation may be obtained from the Einstein–Brown equation, a commonly accepted fluvial sediment transport formula. It will be shown that a minimal number of simplifying assumptions are required.

In addition, it will be shown how the Einstein–Brown equation, an instantaneous sediment transport relation, can be converted into a mean temporal

sediment transport relation for long timescales. It will be shown that the simple form of Equation (C.1) can be maintained under this temporal averaging, and that only the coefficient β_1 is modified with coefficients m_1 and n_1 unchanged. Thus the temporal averaging results in 'the rate of sediment transport being changed; its form, dependent on m_1 and n_1, is unchanged.

A large number of competing fluvial sediment transport formulas have been described in the literature. The scatter in the predictions from these formulas are quite substantial. Large differences are not unusual both between the equations and the observed data, and between the equations themselves. It is not the intention of this section to discuss and compare the merits of the various equations, good descriptions of the various formulas may be found elsewhere (Graf, 1971; Vanoni, 1975). Rather this section will concentrate on one particular, relatively accepted, equation, the Einstein–Brown equation. This equation captures, at least conceptually, the form of sediment transport; that is, its dependence on depth, velocity, sediment size, and channel geometry.

C.1.2 Deriving the Instanteous Sediment Transport Equation.

The Einstein–Brown equation is expressed in terms of a non–dimensional sediment transport ϕ, and a nondimensional shear stress, $\frac{1}{\psi}$. Vanoni (1975) gives the governing equation as

$$\phi = 40 \left(\frac{1}{\psi}\right)^3 \tag{C.2}$$

where

$$\phi = \frac{q_s}{\gamma_s \, F_1 \, \sqrt{g(s - 1) \, d_s^3}}$$

(C.3)

$$\frac{1}{\psi} = \frac{\tau_0}{\gamma(s - 1) \, d_s}$$

(C.4)

$$F_1 = \sqrt{\frac{2}{3} + \frac{36\nu^2}{gd_s^3(s - 1)}} - \sqrt{\frac{36\nu^2}{gd_s^3(s - 1)}}$$

(C.5)

and the notation used is:

q_s = sediment discharge, mass/time/(unit width)

s = specific gravity of sediment

γ = ρg = specific weight of water

d_s = a representative diameter for the sediment particle. Normally d_{50}, the 50 percentile diameter, is used.

g = acceleration due to gravity

τ_0 = γRS = bottom shear stress

R = hydraulic radius

S = bed slope

ν = kinematic viscosity of water

Throughout this work, it is assumed that sediment characteristics over the catchment are homogeneous so that variations in sediment specific gravity s and representative diameter d_s may be ignored. This assumption is consistent with the objective expressed in Section 5.1; that is, developing the minimal model that will create realistic branched channel networks and realistic hillslopes. It is recognized that sediment characteristics change within the catchment in a systematic fashion,

423

particularly along hillslopes. Such systematic variations should be modifiers of the form of the catchment, rather than a fundamental agent in creating the form of the channel network and hillslopes. Additionally, much of the work on the variation of sediment characteristics down hillslopes is of a preliminary nature (e.g., Coventry, et al., 1988), so that it is difficult to allow for such effects in a reliable, and general, way. For homogeneous sediment characteristics F_1, in Equation (C.5) is constant.

The Einstein–Brown equation may be simplified to

$$q_s = F_2 \ (RS)^3 \qquad\qquad (C.6)$$

where

$$F_2 = 40 \ \gamma_s \ F_1 \sqrt{g(s \ - \ 1) \ d_s^3} \ \left[\frac{1}{(s \ - \ 1) \ d_s}\right]^3 \qquad\qquad (C.7)$$

Under the assumption of homogeneous sediment characteristics, since F_1 is spatially constant so is F_2. However, even if sediment characteristics do vary, the variation is confined to the multiplicative constant F_2. The exponent of 3 on hydraulic radius and slope in Equation (C.6) follows from the original Einstein–Brown work. Recent work indicates that this exponent can vary slightly but falls in the range of 2.5 and 3.5 (Madsen, personal communication). For this reason the derivations that follow will use a more general form of Equation (C.6), one where the exponent is allowed to vary so that

$$q_s = F_2 \ (RS)^p \qquad\qquad (C.8)$$

The objective of the derivations that follow is to remove the dependence of

sediment transport on hydraulic radius, R, in Equation (C.8) and reformulate this equation so that it is dependent on only discharge, Q, and slope, S, and is in the form of generic sediment transport equation (C.1). This is achieved by use of Mannings equation for discharge

$$Q = \frac{R^{5/3} \, S^{1/2} \, P}{n}$$ (C.9)

where   P   = wetted perimeter of flow

         n   = Mannings roughness coefficient

Eliminating the hydraulic radius R from Equations (C.8) and (C.9) yields

$$q_s = [F_2 \, (\tfrac{n}{P})^{\frac{3p}{5}}] \, Q^{\frac{3p}{5}} \, S^{\frac{7p}{10}} = F'(P) \, Q^{\frac{3p}{5}} \, S^{\frac{7p}{10}}$$ (C.10)

It will be noted that the multiplicative constant F′ is dependent upon the wetted perimeter, P. The form, and relevance, of this dependence will be examined by looking at three different types of channel cross–section.

Case 1: Wide Channel or Overland Flow

The simplest case that can be considered is that of a wide channel, or overland flow/unit width, as illustrated in Figure C.1. In these cases the governing sediment equation for the wide channel

$$Q_s \;\; = [F_2 \, (\tfrac{n}{w})^{\frac{3p}{5}}] \, Q^{\frac{3p}{5}} \, S^{\frac{7p}{10}}$$
$$\;\;\;\; = F' \, (Q^{\frac{3p}{5}} \, S^{\frac{7p}{10}})$$ (C.11)
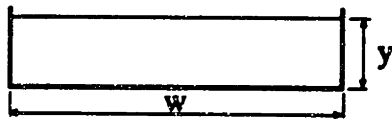
(a) Wide channel

$$R = y$$
$$P = w$$

(b) Overland flow/Flooded rills
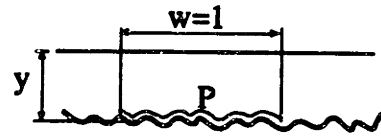
$$R = y$$
$$P = wr \quad r = 1; \text{ flat bottom}$$
$$\quad\quad\quad\quad r > 1; \text{ rilled bottom}$$
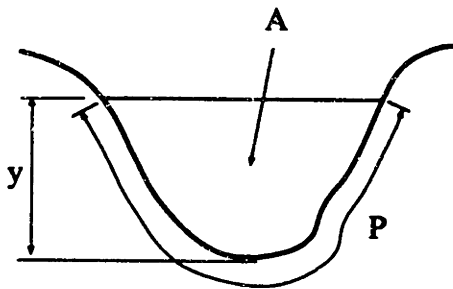
(c) Narrow channel

$$R = \frac{A}{P}$$

(d) Rilled flow

$$R = \frac{A}{P}$$
$$A = \sum_{i=1}^{N} A_i \; ; \qquad P = \sum_{i=1}^{N} P_i$$
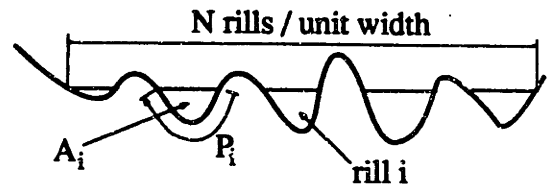
(e) Triangular channel

Figure C.1: Channel and hillslope flow geometries

where w   = width of the channel.

For overland flow/unit width, the governing equation is

$$q_s = [F_2 \, (\tfrac{n}{w_r})^{\tfrac{3p}{5}} \,] \, q^{\tfrac{3p}{5}} \, s^{\tfrac{7p}{10}}$$

$$= F'' q^{\tfrac{3p}{5}} \, s^{\tfrac{7p}{10}} \qquad\qquad (C.12)$$

where

q_s = sediment discharge in units mass/(unit width)/(unit time)

w_r = wetted perimeter/unit width of bottom

q = discharge/unit width.

The multiplicative constants F' and F'' are constant and dependent, in a well defined way, on flow geometry and sediment characteristics. That they are constant results from the wetted perimeter being independent of flow depth. The wetted perimeter in the overland flow case is dependent on the shape of the bottom of the flow. If the bottom is perfectly flat then $w_r = 1$. If the bottom is rilled then $w_r > 1$, the actual value depending on the bottom geometry.

For the specific example of the Einstein–Brown equation (p = 3), Equation (C.11) simplifies to

$$Q_s \quad = F' \, Q^{1.8} \, s^{2.1} \qquad\qquad (C.13)$$

where F' $= [F_2 \, (\tfrac{n}{w})^{1.8}]$.

Using the Chezy equation for discharge instead of Mannings equation, Henderson (1966) obtained, for the Einstein–Brown equation, the expression

$$Q_s = F' Q^2 S^2$$

It may be noted that Equations (C.11) to (C.13) are exact and require no approximation.

Case 2: Triangular Channel

The second case that will be considered is flow in a triangular channel. This case is slightly more complicated than the first because both the hydraulic radius and wetted perimeter change with depth of flow. It will be shown that the sediment transport in a triangular channel can also be formulated in the functional form of the generic sediment transport equation of Equation (C.1).

Consider a triangular channel with shape as in Figure C.1 with the geometry

$$y = a_1 |x| \qquad\qquad\qquad\qquad (C.14)$$

where

y = elevation above the bottom of the channel

x = half width of channel

a_1 = side slope of the channel

An expression will be derived for the wetted perimeter in terms of discharge and slope. This result will be substituted into Equation (C.10). The geometric characteristics of channel cross–sectional area, A, wetted perimeter, P, and hydraulic radius, R, can be expressed in terms of depth, y, as

$$A = a_1 y^2 \qquad\qquad\qquad\qquad (C.15)$$

$$P = 2\sqrt{1 + a_1^2}\; y \qquad\qquad\qquad\qquad (C.16)$$

$$R = \frac{a_1}{2\sqrt{1 + a_1^2}} \, y \qquad\qquad (C.17)$$

These expressions are substituted in Mannings equation for discharge (Equation C.9), and after rearrangement and elimination of the depth y, a relationship between wetted perimeter, discharge, and slope is obtained

$$P = [2.38 \, (\frac{1 + a_1^2}{a_1})^{5/8} \, n^{3/8}] \, Q^{3/8} \, S^{-3/16} \qquad\qquad (C.18)$$

Substituting this result into Equation (C.10) and simplifying yields the sediment transport equation for a triangular channel

$$Q_s = [F_2 \frac{n^{3p/8}}{2.38} \, (\frac{a_1}{1 + a_1^2})^{3p/8}] \, Q^{0.375p} \, S^{0.813p} \qquad\qquad (C.19)$$

For Einstein–Brown's value of p = 3, this becomes

$$Q_s = F' \, Q^{1.13} \, S^{2.44} \qquad\qquad (C.20)$$

where $F' = [F_2 \dfrac{n^{1.13}}{2.38} \, (\dfrac{a_1}{1 + a_1^2})^{1.13}].$

Comparing the result for the wide channel (Equation C.13) and that for the triangular channel (Equation C.20), the main difference is the much reduced dependence on discharge for the triangular channel. This results from the strong dependence of discharge on depth of flow ($Q \approx y^{8/3}$). Large changes in discharge have

429

significantly reduced effects on flow depths, stresses and thus sediment transport in a triangular channel.

We will now discuss a case where wetted perimeter is dependent on discharge, F is dependent on discharge, and where an approximation is necessary to formulate the equations as the generic sediment transport equation.

Case 3: Narrow Channel or Rilled Flows

The third, and potentially most interesting, flow geometry is a channel or rill, where the flow has not exceeded the conveyance capacity of the channel, as illustrated in Figure C.1. In this case the wetted perimeter varies with discharge, typically in a non-linear fashion. It will be necessary in this case to introduce a number of approximations to allow Equation (C.10) to be expressed in the form of Equation (C.1) where the coefficients β_1, m_1, and n_1 are independent of discharge.

To simplify the analysis, and without significant loss of generality, channels and rills with the following geometry will be considered.

$$y = a_1 x^{b_1} \tag{C.21}$$

where

y	= elevation above the bottom of the channel
x	= the half width of the channel
a_1, b_1	= channel geometry coefficients

To be able to express Equation (C.10) in a form dependent only on discharge and slope, it will be necessary to find a relationship between wetted perimeter, discharge and slope in the form

$$P = a_2 \, Q^{b_2} \, S^{b_3} \qquad\qquad (C.22)$$

Mannings equations for discharge gives the discharge as

$$Q = \frac{A^{5/3} \, S^{1/2}}{P^{2/3} \, n} \qquad\qquad (C.23)$$

This relates the cross–sectional area, discharge and wetted perimeter and can be used to eliminate of discharge in Equation (C.22) so that

$$P^{\frac{3+2b_2}{3}} = A^{\frac{5b_2}{3}} \left(\frac{a_2 \, S^{b_3 + \frac{b_2}{2}}}{n^{b_2}} \right) \qquad\qquad (C.24)$$

Both cross–sectional area, A, and wetted perimeter, P, are dependent on depth but independent of the slope, S. This means that

$$b_3 + \frac{b_2}{2} = 0 \qquad\qquad (C.25)$$

so that Equation (C.24) can be rewritten as

$$P = \left[\frac{a_2}{n^{b_2}} \right]^{\frac{3}{3+2b_2}} A^{\frac{5b_2}{3+2b_2}} \qquad\qquad (C.26)$$

Both area, A, and wetted perimeter, P, are dependent upon the channel geometry and depth of the flow in the channel, y. For the assumed channel geometry (Equation

C.21)

$$A = \frac{2a_1 b_1}{(b_1 + 1)} (y')^{\frac{1+b_1}{b_1}}$$ (C.27)

$$P = 2a_1 \int_0^{y'} \left[\left(\frac{1}{a_1 b_1}\right) z'^{\frac{2(1-b_1)}{b_1}} + 1 \right]^{1/2} dz'$$ (C.28)

where $\quad y' \quad = \frac{y}{a} =$ nondimensional depth of flow

$\quad z' \quad = \frac{z}{a} =$ nondimensional elevation above the bottom of the channel.

A simple analytic expression relating the area, A, and wetted perimeter, P, does not exist for this channel geometry. The approach adopted is to obtain an approximate expression in the form of Equation (C.26) by eliminating y' between Equations (C.27) and (C.28). It is convenient to fit the derivative of P with respect to y'. That is, the expression

$$\frac{dP}{dy'} = \left(\frac{5 b_2}{3+2b_2}\right)\left(\frac{a_2}{n^{1/2}}\right)^{\frac{3}{3+2b_2}} \frac{dA}{dy'} A^{\frac{3b_2-3}{3+3b_2}}$$ (C.29)

will be fitted using the expressions

$$\frac{dA}{dy'} = 2a_1 (y')^{\frac{1}{b_1}}$$ (C.30)

$$\frac{dP}{dy'} = 2a_1 \left[\left(\frac{1}{a_1 b_1}\right) y'^{\frac{2(1-b_1)}{b_1}} + 1 \right]^{1/2}$$ (C.31)

Equation (C.24) can be rewritten as

$$\frac{dP}{dy'} = \left(\frac{5b_2}{3+2b_2}\right) \left(\frac{a_2}{b_2}\right)^{\frac{3}{3+2b_2}} \left[\left(\frac{2a_1 b_1}{b_1+1}\right)\left(\frac{1}{2a_1}\right)^{1+b_1}\right]^{\frac{3b_2-3}{2b_2+3}} \cdot$$

$$\left[\frac{dA}{dy'}\right]^{1 + \frac{3(b_1+1)(b_2-1)}{2b_2+3}}$$

(C.32)

In addition, it is necessary to fit Equation (C.32) over a specified range of y'. Leopold and Langbein (1962) point out that the maximum lateral slope of a channel cross–section, at the top of the bank, should be the angle of repose of the material, this being the steepest angle at which the sediment material is stable. Using the channel cross–section geometry (Equation C.21), an expression for the maximum lateral slope and thus the maximum depth of the channel is obtained

$$\frac{dy}{dx}\bigg|_{y'=y'_{max}} = \tan \theta_r$$

(C.33)

where θ_r = angle of repose of sediment material.

The maximum depth of channel follows and is

$$y'_{max} = \left[\frac{\tan \theta_r}{a_1 b_1}\right]^{\frac{b_1}{b_1-1}}$$

(C.34)

A typical value for the angle of respose for cohesionless, granular material is 30°. Values of the maximum depth of the channel area are tabulated in Table C.1 (angle of repose of 30°) for various values of channel geometry parameters a_1 and b_1.

TABLE C.1

Maximum Nondimensional Depth, y', for a Channel with

Variable Geometry, Angle of Repose $= 30^{o}$

	Maximum depth $= y'$ max		
	$a_1 = 1$	$a_1 = 3$	$a_1 = 10$
$b_1 = 2$	0.75	0.083	0.0075
$b_1 = 3$	0.44	0.084	0.0139
$b_1 = 4$	0.33	0.076	0.0152

Curves of $\frac{dP}{dy'}$ versus $\frac{dA}{dy'}$ for these values of a_1 and b_1 are shown in Figure C.2. In this figure, power law curves have been fitted. These fitted curves were used to determine values of a_2, b_2, and b_3 from Equation (C.32) and the resulting coefficients are given in Table C.2. Substituting these values of a_2, b_2, and b_3 into Equations (C.27) and (C.10) yielded the sediment transport formulas given in Table 5.1. For comparison purposes, the exact sediment transport formulas for the wide channel and triangular channel previously derived are also tabulated.

The channel cross–section shape has little effect on the functional dependence of sediment transport to channel slope. The cross–section dependence is most apparent in the power on discharge, with a lesser dependence being displayed by the multiplicative constant. Table 5.1 indicates the range of values that m_1 and n_1 may take in Equation (C.1). That is

$$m_1 \quad \epsilon \ [0.375p, 0.6p]$$
$$n_1 \quad \epsilon \ [0.7p, 0.813p] \quad\quad\quad\quad\quad\quad\quad\quad (C.35)$$
$$\frac{m_1}{n_1} \quad \epsilon \ [0.413, 0.857]$$

In conclusion, it has been shown that a good approximation to accepted sediment transport formulas is given by Equation (C.1), with an appropriate choice of coefficients. Appropriate values of the coefficients β_1, m_1 and n_1 have been derived for a variety of channel geometries, using the Einstein–Brown sediment transport equation as an example.

It has been noted that the coefficients of Equation (C.1), β_1, m_1 and n_1 are dependent on channel geometry and sediment characterization in a well defined way. For uniform overland flow, a wide channel and a triangular channel flow Equation (C.1) is an exact representation of the Einstein–Brown equation. For the case of
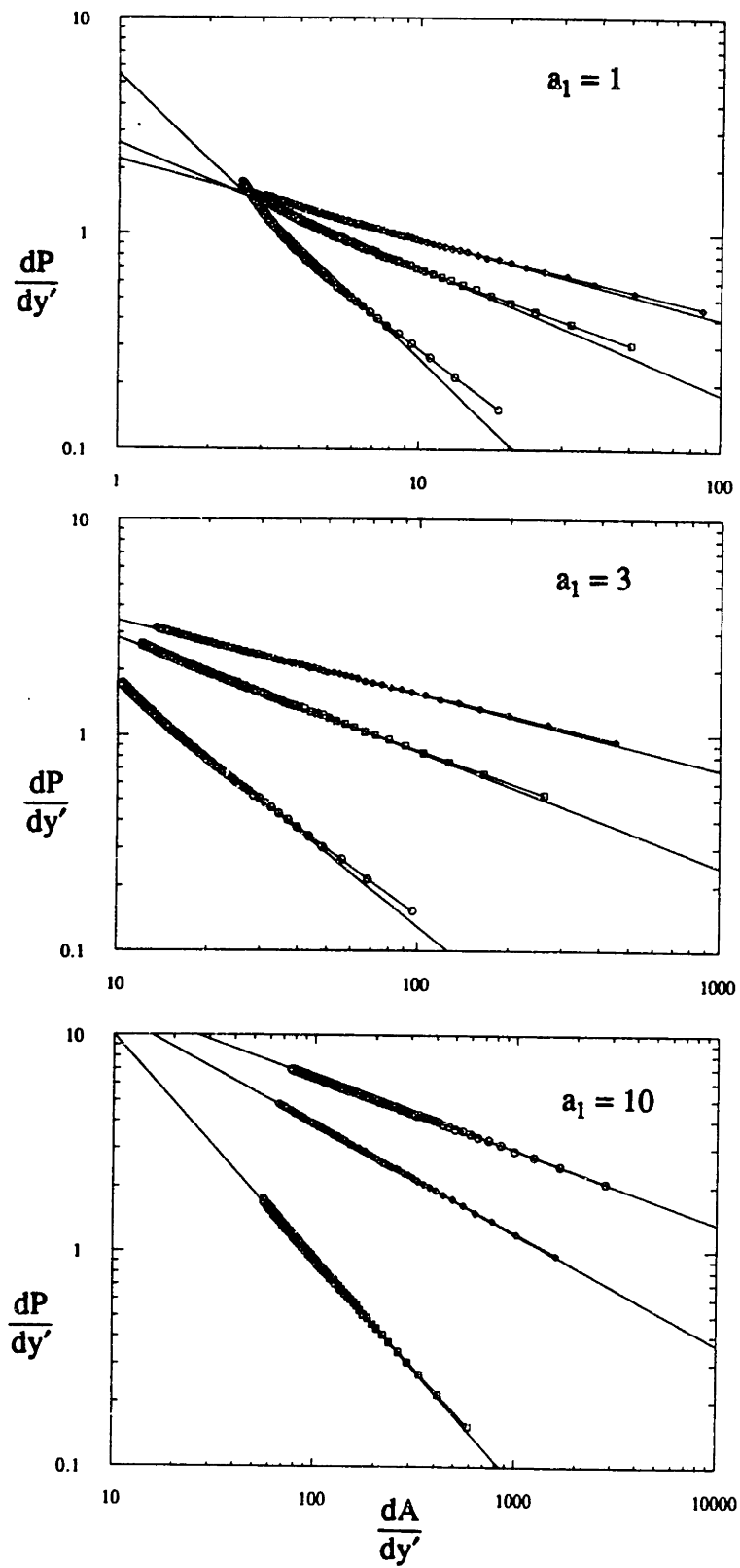
Figure C.2: Regression of $\dfrac{dA}{dy'}$ versus $\dfrac{dP}{dy'}$ for variable channel geometry

TABLE C.2

Fitted Parameters a_2, b_2, and b_3

Channel Geometry		Fitted Wetted Perimeter Parameters		
a_1	b_1	b_2	b_3	b_2
1	2	0.298	−0.149	$3.84\ n^{0.298}$
1	3	0.221	−0.111	$3.98\ n^{0.221}$
1	4	0.175	−0.088	$4.17\ n^{0.175}$
3	2	0.260	−0.130	$4.49\ n^{0.260}$
3	3	0.189	−0.095	$5.19\ n^{0.189}$
3	4	0.149	−0.074	$5.94\ n^{0.147}$
10	2	0.241	−0.121	$5.11\ n^{0.241}$
10	3	0.174	−0.087	$6.97\ n^{0.174}$
10	4	0.136	−0.068	$8.71\ n^{0.136}$

variable channel geometry, the error in approximating Einstein–Brown with constant β_1, m_1, and n_1 is small, particularly in comparison with the errors of prediction of the sediment transport formulas.

Moore and Burch (1987) used unit stream power theory and the experimental data of Mosley (Schumm, 1987) to derive a sediment transport equation for rills on hillslopes. Their equation was

$$
\begin{aligned}
Q_s \quad &\approx Q^{1.6} S^{1.8} \\
&\approx (Q^{0.53} S^{0.6})^p
\end{aligned}
$$

where $p = 3$.

This equation corresponds well with a case intermediate between the wide channel and the variable geometry channel (Table 5.1).

C.1.3 Deriving the Mean Temporal Sediment Transport Equation

The derivations thus far, however, have been limited to the instantaneous sediment discharge. The Einstein–Brown equation has mostly been verified for sediment transport at an instant in time, not averaged over long periods of time. It remains to be shown that while Equation (C.1) is satisfactory for describing the instantaneous discharge, that it is also a satisfactory representation of the mean temporal sediment discharge over long time scales.

The time scales of interest in this work are typically of the order of thousands of years; time enough to average over the distribution of geomorphologically effective runoff events (Section 2.4). It will be shown that by temporal averaging over the distribution of flood hydrographs, where Equation (C.1) describes the instantaneous sediment discharge, a modified version of Equation (C.1) can be obtained for the mean temporal sediment discharge. A new value of the multiplicative constant β will be

obtained that will be dependent on the moments of the distribution of flood events. The interpretation of discharge is also changed.

Hereafter, to avoid confusion of notation we will use the parameters of Equation (C.1), i.e., β_1, m_1 and n_1 to represent those parameters applicable to the mean temporal sediment discharge equation and the equation

$$Q_s = \beta_2 \, Q^{m_2} \, S^{n_2} \tag{C.36}$$

with its parameters β_2, m_2 and n_2 to represent those parameters applicable to the instantaneous sediment discharge equation.

In most unregulated rivers the amount of sediment transported during flood events far exceeds that transported during the long low flow periods that separate these events. This follows from the highly nonlinear dependence of sediment transport on Q in Equation (C.1), where m_2 is typically of the order 2 of more, as previously described. In addition, the flow volume distribution is highly skewed towards high discharges, so that high discharges carry a disproportionately large volume of sediment. It is then a useful, and very good, approximation to ignore all sediment transported during the low flow periods, and consider sediment transport as occurring in discrete events, these events corresponding to the flood hydrographs. The mean temporal sediment transport is then obtained by averaging over the distribution of transport events.

Consider a single flood hydrograph described by the discharge with time Q(t), with a characteristic duration T_p, and a characteristic discharge Q_p. In the following derivation the characteristic discharge will be considered to be the peak discharge (Figure C.3) The total sediment load carried by this single hydrograph, $\hat{Q}_s$, is given by
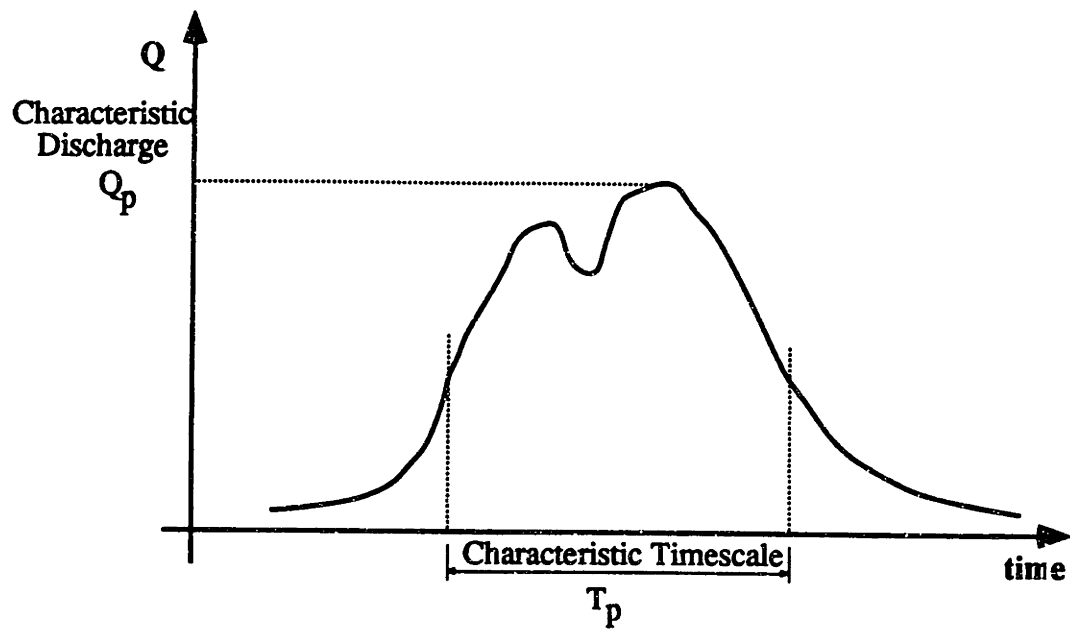
Figure C.3: Definition of hydrograph characteristic length and time scales

$$\hat{Q}_s = \beta_2 \, S^{m_2} \int_{-\infty}^{\infty} Q(t)^{m_2} \, dt \qquad (C.37)$$

where $\hat{Q}_s$ = sediment load of the hydrograph, mass/hydrograph.

The channel slope has been assumed constant over the length of the hydrograph. If we define a nondimensional time and a nondimensional discharge so that

$$t' = \frac{t}{T_p}$$

$$Q'(t) = \frac{Q(t)}{Q_p}$$

then the sediment load/hydrograph (Equation C.37) may be expressed as

$$\hat{Q}_s = [\beta_2 T_p \int_{-\infty}^{\infty} (Q'(t'))^{m_2} dt'] \, Q_p^{m_2} \, S^{n_2} \qquad (C.38)$$

Note that the total sediment load per flood event is in the form of the generic sediment transport equation of Equation (C.1). If flood hydrographs are considered to have a characteristic or average shape, $Q'(t')$ then it is possible to parameterize the total sediment load in a hydrograph by its peak discharge Q_p, and its duration T_p with a constant multiplicative constant β. It is now possible to average this equation over the probability distribution of hydrographs, parameterized by their random peak discharge, Q_p, and duration, T_p. Knowing the frequency in time of these events would permit the expression of the mean temporal sediment transport equation in terms of flood frequency, mean peak discharge, and mean duration. The details follow.

Consider the shape of the flood hydrograph, $Q'(t')$, fixed and independent of Q_p and T_p. The distributions of the peak discharges and times of duration of floods

can be expressed in terms of their means and random fluctuations around those means so that

$$Q_p = \overline{Q}_p + Q_p^*$$

$$T_p = \overline{T}_p + T_p^*$$

$$E[Q_p] = \overline{Q}_p; \quad E[Q_p^*] = 0; \quad E[(Q_p^*)^2) = \sigma_{Q_p}^2$$

$$E[T_p] = \overline{T}_p; \quad E[T_p^*] = 0; \quad E[(T_p^*)^2] = \sigma_{T_p}^2$$

$$E[T_p^* Q_p^*] = \sigma_{Q_p T_p}^2$$

where

$$\overline{Q}_p \quad = \text{mean peak discharge}$$

$$\overline{T}_p \quad = \text{mean time of duration}$$

$$Q_p^*, T_p^* \quad = \text{random fluctuation of peak discharge and time of duration about mean}$$
$$\text{values.}$$

The sediment load in each flood hydrograph can then be considered to be a random variable, derived from the probability distributions of Q_p and T_p. The sediment load per hydrograph can also be expressed as a mean and random fluctuation about that mean so that

$$\hat{Q}_s = \overline{\hat{Q}}_s + \hat{Q}_s^*$$

$$E[\hat{Q}_s] = \overline{\hat{Q}}_s; \quad E[\hat{Q}_s^*] = 0; \quad E[(\hat{Q}_s^*)^2] = \sigma_{\hat{Q}_s}^2$$

where

$\overline{Q}_s$ = mean sediment load/hydrograph

$\hat{Q}_s^*$ = random fluctuation of the sediment load/hydrograph about its mean value

Substituting these expressions into Equation (C.38) yields the perturbation equation for sediment load/flood event

$$\overline{Q}_s + \hat{Q}_s^* = [\beta_2 \int_{-\infty}^{\infty} (Q'(t'))^{m_2} dt'] \, (\overline{T}_p + T_p^*)(\overline{Q}_p + Q_p^*)^{m_2} S^{n_2}$$

(C.39)

We are interested in the mean of this random sediment load over many flood events. Define a rate constant λ which is the average number of flood events N occurring in a given period T as that

$$\lambda = E[\frac{N}{T}]$$

So in any given period T, the sediment transport rate $(\overline{Q}_s + \overline{Q}_s^*)$ averaged over the period T will be

$$(\overline{Q}_s + \overline{Q}_s^*) = \frac{(\overline{Q}_s + \hat{Q}_s^*)N}{T}$$

$$= [\beta_2 \int_{-\infty}^{\infty} (Q'(t')^{m_2} dt'] \frac{NS^{n_2}}{T} (\overline{T}_p + T_p^*)(\overline{Q} + Q_p^*)^{m_2}$$

(C.40)

where $\overline{Q}_s$ = mean sediment transport rate over period T

$\overline{Q}_s^*$ = perturbation of the sediment transport rate from mean in period T

The mean temporal sediment transport rate (load/(unit time)) is obtained by taking the expectation with time of Equation (C.40), so that

$$\overline{Q}_s = E[\overline{Q}_s + \overline{Q}_s^*]$$

$$= \beta_2 S^{n_2} \left[\int_{-\infty}^{\infty} (Q'(t'))^{m_2} dt' \right] \cdot$$

$$E[\frac{N}{T}(\overline{T}_p + T_p^*)(\overline{Q}_p + Q_p^*)^{m_2}] \qquad (C.41)$$

If the rate at which flood events occur, parameterized by λ, is independent of both the duration and peak discharge of the flood events this equation may be simplified to obtain

$$\overline{Q}_s = \beta_2 \lambda S^{n_2} \left[\int_{-\infty}^{\infty} (Q'(t'))^{m_2} dt' \right] \cdot$$

$$E[(T_p + T_p^*)(\overline{Q}_p + Q_p^*)^{m_2}] \qquad (C.42)$$

In order to evaluate the expectation in this equation, it is necessary to expand the nonlinear term in peak discharge by Taylor's series around the mean peak discharge, yielding

$$\overline{Q}_s = \beta_2 \lambda S^{n_2} \left[\int_{-\infty}^{\infty} (Q'(t'))^{m_2} dt' \right] \cdot$$

$$E[(\overline{T}_p + T_p^*)(\overline{Q}_p^{m_2} + m_2 Q_p^* \overline{Q}_p^{m_2-1} + m_2(m_2-1)(Q_p^*)^2(\overline{Q}_p)^{m_2-2} + ...]$$

$$(C.43)$$

444

If the expectation is evaluated and perturbation terms of order 3 or more are neglected, then the following expression is obtained for the mean sediment transport rate

$$\overline{Q}_s = \left[\beta_2 \, \overline{T}_p \, \lambda \int_{-\infty}^{\infty} (Q'(t'))^{m_2} \, dt' \right] \cdot$$

$$\left[1 + m_2(m_2-1) \frac{\sigma^2_{Q_p}}{\overline{Q}^2_p} + m_2 \frac{\sigma^2_{Q_p T_p}}{\overline{Q}_p \, \overline{T}_p} \right] \overline{Q}_p^{m_2} \, S^{n_2}$$

$$(C.44)$$

It should be noted that this equation is in the form of the generic sediment transport equation where the correspondence between coefficients in the instantaneous equation (C.36) and the time averaged equation (C.1) is as follows.

$$m_1 = m_2$$

$$n_1 = n_2$$

$$\beta_1 = \beta_2 \left[\overline{T}_p \, \lambda \int_{-\infty}^{\infty} (Q'(t'))^{m_2} \, dt' \right] \cdot \left[1 + m_2(m_2-1) \frac{\sigma^2_{Q_p}}{\overline{Q}^2_p} + m_2 \frac{\sigma^2_{Q_p T_p}}{\overline{Q}_p \, \overline{T}_p} \right]$$

Thus it is seen that the instantaneous sediment transport equation and the mean temporal sediment transport are of the same form and are simply related, at least in principle.

The major difference between the instantaneous and time averaged equations is the definition of the discharge. The dependence of the mean temporal sediment

transport on the mean peak discharge is important. An important concept in river engineering is that of the "dominant" discharge. The dominant discharge is that discharge which, from a sediment transport viewpoint, is equivalent to the distribution of discharges over many flood events. The channel size and shape is typically asserted to be in balance with the dominant discharge, leading to this discharge sometimes being referred to as the channel forming discharge. In this case the dominant discharge is sometimes considered synonymous with the bankful discharge (Pilgrim and McDermott, 1981). It is generally asserted that the bankful discharge has a recurrence interval of 1.5 to 2.5 years, roughly corresponding to the mean annual peak discharge. Thus the dependence of the mean sediment transport on the mean peak discharge in Equation (C.44) appears in a fashion consistent with current ideas regarding channel cross–section equilibrium. The derivation leading to Equation (C.44) provides some insight into why this might be so.

There is one important difference between the dominant discharge above and the mean peak discharge in Equation (C.44). The dominant discharge is, in practice, equated to the mean annual peak discharge with the mean being determined from the series of annual maximum discharges. However, $\overline{Q}_p$ in Equation (C.44) is the average of the peak discharges of all flood events that carry significant sediment loads. Thus $\overline{Q}_p$ is the mean peak discharge from a flood frequency analysis based on exceedance series, with the lower cutoff being that discharge below which sediment load is insignificant. The relative merits of each of these flood frequency analysis techniques has been widely debated in the hydrology literature; currently there is no prevailing opinion. It is not clear whether the mean peak discharge calculated from each method is necessarily the same; in any case, for Equation (C.44) the exceedance series, not the annual maxima, should be used.

This analysis that started at the instantaneous sediment transport equation

(C.36) and that arrived at the time averaged sediment transport equation (C.44) is a standard second order perturbation analysis (Gelhar and Axness, 1981). By ignoring third order perturbation terms in Q_p in the evaluation of Equation (C.43), it is implicitly assumed that the frequency distribution of Q_p is symmetric. Practical experience in flood frequency analysis suggests that this is unreasonable, since skewed distributions such as log–normal, log–Pearson type III and Gumbel are commonly found to provide satisfactory fits to flood frequencies.

Incorporation of higher order moments, so that skewed distributions of Q_p can be modelled does not invalidate the conclusion reached for the second order analysis. For instance, if the third moment of Q_p were to be modelled (i.e., $E[(Q_p^*)^3]$), so that non–zero skewness may be modelled, then the multiplicative constant β_1, in Equation (C.44) is modified and becomes

$$
\beta_1 = \beta_2 \left[\overline{T}_p \, \lambda \int_{-\infty}^{\infty} (Q'(t'))^{m_1} \, dt' \right] \cdot
$$
$$
\left[1 + m_2(m_2-1) \frac{\sigma_{Q_p}^2}{\overline{Q}_p^2} + m_2(m_2-1)(m_2-2) \frac{\gamma_{Q_p}(\sigma_{Q_p}^2)^{3/2}}{\overline{Q}_p^3} + \right.
$$
$$
\left. m_2 \frac{\sigma_{Q_p T_p}^2}{\overline{Q}_p \, \overline{T}_p} \right] \tag{C.45}
$$

where

γ_{Q_p} = skewness coefficient of Q_p

$$
= \frac{\frac{1}{N} \sum_{i=1}^{N} (Q_p^*)_i^3}{(\sigma_{Q_p}^2)^{3/2}}
$$

N = total number of flood events

Note that the incorporation of skewness only modifies the value of β_1, and not the functional form of the sediment transport equation. Additional, higher order, moments of Q_p, such as kurtosis, may be incorporated. The marginal benefit of including higher order moments diminishes, however, because of the increased difficulty and unreliability of estimating them. The most complicated frequency distribution that is currently used by hydrologists is the log–Pearson type III which uses the mean, variance, and a shape parameter. They may be determined from the first three moments of the distribution of Q_p.

In conclusion, this section has shown two things:

1. A commonly accepted instantaneous total load sediment transport equation, Einstein–Brown, may be reformulated into the functional form used in this work. This reformulation requires some assumptions regarding channel cross–section geometry. In all cases, the approximations implicit in the reformulation are small. It was shown how the coefficients of the Einstein–Brown equation may be used to derive the parameters of the generic sediment transport (Equation C.1). Thus the functional form of sediment transport used in this work may be directly related to experimentally measurable quantities.

2. The instantaneous sediment transport equation adopted in this work can be time averaged to give a mean temporal sediment transport equation. The result is in the functional form used in this work. The discharge used in the mean temporal equation is the mean peak discharge. This mean peak discharge is interpreted as the "dominant" or channel forming discharge. The transport coefficient β_1 is dependent upon the distribution of flood peaks, parameterized by mean peak discharge.

C.2 Physical Basis for the Activator Mechanism

C.2.1 Introduction

This section will provide a linkage between the types of processes that are observed to trigger, or activate, channelization in the field, and Equation (5.2a) which is used to represent the activator mechanism in this work. It should be stated at the outset that the state–of–the–art understanding of these processes is, at best, primitive. A qualitative understanding of the processes is beginning to be developed at the current time. Quantitative understanding of the range of processes is less advanced, and is either the subject of current research, such as groundwater induced stream growth (Dunne, 1989) or has yet to be fully addressed, such as is the case for overland flow induced channel growth.

The generic equation used to represent the channelization activator is

$$a = \beta_5 \, Q^{m_5} \, S^{n_5} \tag{C.46}$$

In Section 5.3 it was noted that the purpose of the activator in the channelization equation (5.1b) is to trigger the one way process modelling the transition from hillslope to channel. This process is triggered at the time when the activator exceeds the channelization threshold at that point. This threshold in Equation (5.1b) is

$$a_{threshold} = \frac{1}{c_1} \tag{C.47}$$

Once this threshold is exceeded, the channelization process proceeds at a rate governed by the timescale for channelization, $1/d_t$. The form of Equation (5.1b) is such that the actual value of activator during the transition stage between hillslope

and channel (i.e., when Y is between 0.1 and 1.0) is relatively insensitive to the actual value of the activator, a, provided that activator is not substantially larger than the threshold value. Formally this can be expressed as

$$a < 5\, a_{threshold} \tag{C.48}$$

Thus the importance of the activator is in the triggering of the channelization, rather than in governing the rate of channelization. Once channelization is triggered, growth occurs independently of the activator level. It was noted in Section 5.3 that during the channelization process sediment transport is considered to be intermediate between the value of transport for a hillslope and the value of transport for a channel. Because the rate of channelization is independent of activator level, and so is independent of drainage directions (which follow from the hillslope form), the form of the transition from hillslope to channel sediment transport is relatively unimportant.

The question that this section addresses is: What physical processes can Equation (C.46) be claimed to simulate given the way the activator threshold is used in the model (Equation 5.1)? Some examples of physical processes will follow.

An important observation must be made before we begin. In Equation (5.1) there is only one definition of discharge and this single definition was used in both the activator and sediment transport equations. It was noted in Section C.1 that this discharge may be interpreted in two different ways, depending on the interpretation to be placed on the elevation evolution equation (5.1a). They are:

1. If Equation (5.1a) is to be interpreted as the instantaneous change of elevation with time, then Q is the instantaneous discharge.

2. If Equation (5.1a) is to be interpreted as the mean change of elevation with time, short term elevation perturbations being averaged out, then Q

is the mean of all peak discharges of all significant flood events, $\overline{Q}_p$.

Activator must be interpreted in an analogous fashion to sediment transport. In the one case, the activator occurs instantaneously in time. In another case, the activator is considered to be the mean over many storm events. This mean activator must be expressed in terms of the mean peak dischage $\overline{Q}_p$. The question of what constitutes the mean temporal activator of channelization, over many flood events, is very poorly understood, and has not been satisfactorily addressed in the literature. Calver (1978) provided a preliminary understanding on the basis of a conceptualization of the instantaneous channel growth mechanism. The lack of theoretical basis for the conceptualization, and the lack of any experimental evidence, make this work somewhat speculative.

C.2.2 Deriving the Activator Equation

In the derivation of the activator formulas that follow a number of different physical mechanisms will be examined for different channel geometries. In the cases where the channel initiation mechanism, or activator, is considered be either overland flow velocity or overland flow bottom shear stress two channel geometries will be examined; a wide channel of uniform depth and a triangular channel. These activator mechanism are based on Hortonian overland flow mechanisms.

Recently a threshold based on a groundwater driven stream sapping mechanism has been proposed (Dunne, 1989). It can be formulated into the form of the activator function used in this work.

Case 1: Overland Flow Velocity

One of the most common criteria for the design of erosion works is overland flow velocity. Many engineering handbooks give tabulations of allowable velocities for

various forms of ground cover or erosion protection (e.g., Henderson, 1966). The overland flow velocity can be expressed in the generic form of activator in Equation (C.46).

If the wide channel assumption is made so that hydraulic radius, R, is equal to the flow depth, y, and the wetted perimeter, P, is independent of discharge, then Mannings equation can be written as

$$v = \frac{R^{2/3} S^{1/2}}{n} \qquad \text{(C.49)}$$

and the discharge for a wide channel of width w can be written as

$$Q = \frac{R^{5/3} S^{1/2} w}{n} \qquad \text{(C.50)}$$

Combining these equations yields

$$v = [\frac{1}{w^{2/5} n^{3/5}}] Q^{0.4} S^{0.3} \qquad \text{(C.51)}$$

or expressing this result in terms of discharge/(unit width), q,

$$v = [\frac{1}{n^{3/5}}] q^{0.4} S^{0.3} \qquad \text{(C.52)}$$

A similar expression for velocity in a triangular channel may also be derived. Consider a triangular channel geometry with sideslope a_1 (Equation C.14) and Mannings equation for velocity and discharge (Equations C.49 and C.50). Using the expression for hydraulic radius and wetted perimeter of Equations (C.16) and (C.17),

we may obtain, after elimination of the flow depth, y, the following expression for average velocity in the channel

$$v = \left[\frac{a_1}{4(1 + a_1^2)n^3}\right]^{0.25} Q^{0.25} S^{0.375} \qquad (C.53)$$

Equations (C.51), (C.52), and (C.53) are expressed in the general form of activator equation given by Equation (C.46). Coefficients β_5, m_5, and n_5 all follow directly from the hydraulics of the environment and the roughness of the surface.

Case 2: Overland Flow Shear Stress

An important concept in modern sediment transport theory is that of a threshold bed shear stress, below which no sediment transport takes place. This concept was first clearly expressed by Shields (1936) (Vanoni, 1975) in his now famous Shields diagram which relates bottom shear stress with sediment transport. Thresholds on shear stress are included in transport formulas due to Duboys (in 1897), Shields (in 1936) and Laursen (in 1958), among others. Without sediment transport the erosion necessary to scour out a channel cannot take place.

For a wide channel, or overland flow region, the bottom shear stress τ_* is given by

$$\tau_* = \gamma R S \qquad (C.54)$$

Using the expression for discharge from Mannings equation (C.50) to substitute for depth y (y = R for the wide channel) in Equation (C.54) gives the bottom shear stress for a wide channel as

$$\tau_* = [\gamma(\tfrac{n}{w})^{3/5}] \, Q^{0.6} \, S^{0.7} \tag{C.55}$$

The equivalent expression in terms of discharge/(unit width), q, is

$$\tau_* = [\gamma \, n^{3/5}] \, q^{0.6} \, S^{0.7} \tag{C.56}$$

A similar expression for shear stress in a triangular channel may also be derived. With the triangular channel geometry with sideslope a_1 (Equation C.14), Mannings equation for discharge (Equation C.8), the shear stress equation (C.54) and the expressions for hydraulic radius and wetted perimeter of Equations (C.16) and (C.17), we may obtain, after elimination of the flow depth, y, the following expression for the average shear stress on the channel bottom.

$$\tau_* = \left[\frac{n \, a_1 \, \gamma^{8/3}}{4 \sqrt{1 + a_1^2}} \right]^{3/8} Q^{0.375} \, S^{0.813} \tag{C.57}$$

As for the wide channel case Equations (C.55), (C.56) and (C.57) are in the form of the generic activator equation (C.46). The coefficients β_5, m_5, and n_5 follow directly from the hydraulics of the environment and the roughness of the surface. It is interesting to note that the dependence of β_5 on Mannings roughness coefficient is qualitatively different from the case of overland velocity.

It might be noted that there are two opposing viewpoints regarding the applicability of a threshold shear stress to channel formation. The first viewpoint is that channel formation cannot occur unless sediment transport occurs, thus the shear stress threshold must be exceeded for channel formation to occur. However, it must be noted that active sediment transport occurs on the hillslopes, albeit at lower rates.

Thus the shear stress threshold must be exceeded on the hillslope as well as in the channel. These contradictions would suggest that the shear stress threshold is not the crucial one for triggering channelization. Priest et al. (1975) supports the view that a shear stress threshold is not the critical one for channel growth. He notes that shear stresses in the channel are insufficent to scour the channel, even though they are sufficent to removed already loosened material.

The second viewpoint is that a second, higher, threshold may conceivably exist that corresponds to the channelization threshold. This threshold may represent the breaking through of the mat of grass and root material in the upper soil zone. It is this effect that should be modelled as the critical threshold on channelization. This was the philosophy of the preceding derivations. This philosophy is consistent with observations that substantial overland erosion may take place on the hillslope without rilling or channelization (Dunne and Aubrey, 1986).

Case 3: Groundwater Stream Sapping

Dunne (1969) proposed a conceptualization of a groundwater process where groundwater streamtubes converged onto a seepage face at a channel head. The groundwater discharged from the seepage face and into the channel, eroding and extending the channel head. This conceptualization has been supported by other experimental evidence (Coelho Netto, et al., 1987) and is illustrated in Figure C.4. Channel advance occurs because of the erosion and piping at the seepage face, followed by head collapse, provided only that the in—channel sediment transport is sufficient to remove the loosened material from the seepage face and carry it downstream. The form and magnitude of the processes that act at the seepage face are still not known exactly, and seem to vary from site to site (Dunne, 1989). As a result there is some uncertainty about the actual physical variables that are responsible for most of the bulk movement at the channel head, e.g., hydraulic gradient, pore pressure, velocity,
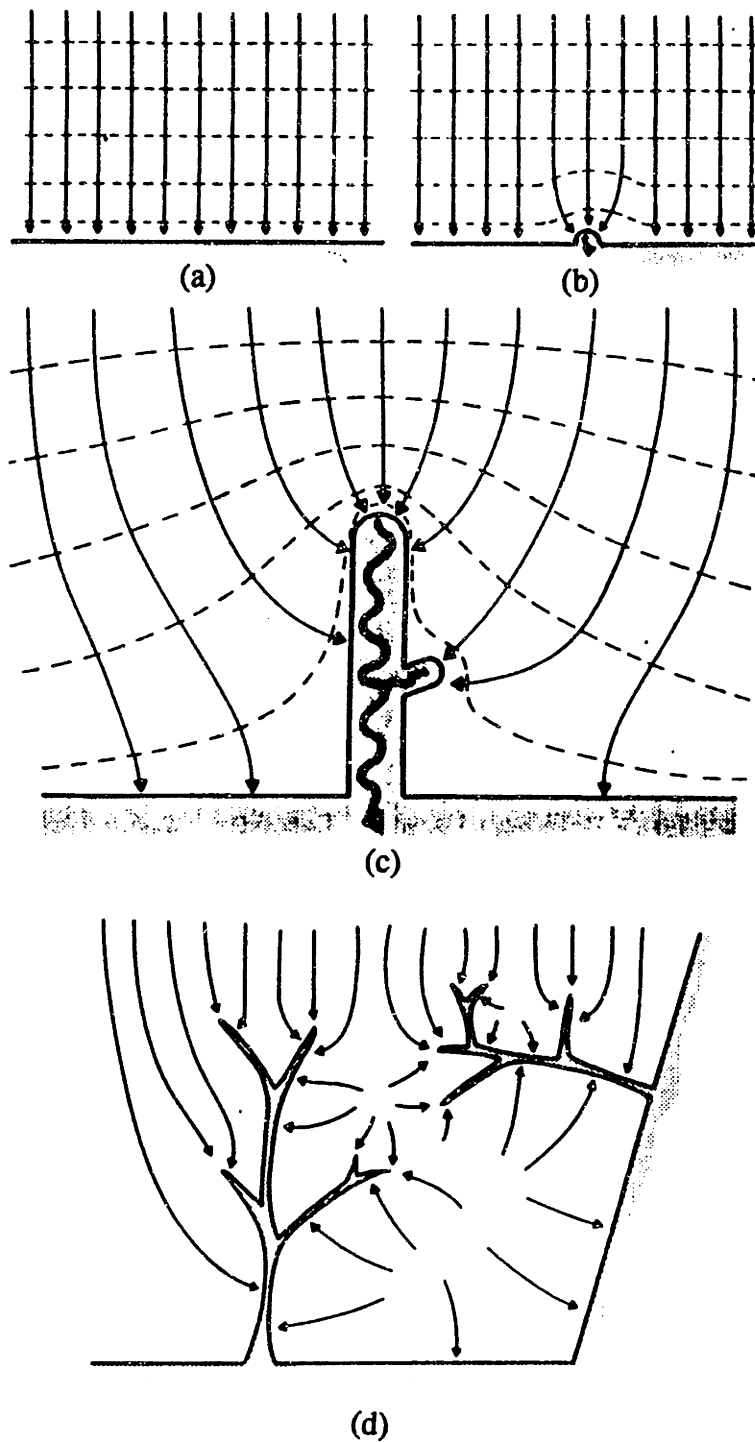
Figure C.4: Dunne's stream sapping hypothesis (from Dunne, 1980)
(a) Uniform initial conditions   (b) Budding of a new channel
(c) Flow convergence around the channel head
(d) Flow patterns around a network of channels

specific discharge or grain shear stress.

Dunne (1989) suggested a threshold on hydraulic gradient above which erosion at the seepage face would occur by piping. He believed velocity effects are not important at the seepage face. This hydraulic gradient threshold is

$$\left[\frac{dH}{dx}\right]_{threshold} = (s-1)(1-n)$$
(5.58)

where

$\frac{dH}{dx}$ = groundwater hydrualic gradient at the seepage face

s = specific gravity of the sediment material

n = porosity

To reformulate Equation (C.58) into the form of the generic activator equation Darcy's law is used. Darcy's law for groundwater flow at the seepage face is

$$Q = K\frac{dH}{dx}h\ w$$
(C.59)

where

K = hydraulic conductivity

h = height of the seepage face

w = width of the seepage face

Formulating this in the form of the generic activator, where head gradient is the activator yields

$$\frac{dH}{dx} = \left[\frac{1}{Khw}\right]Q$$
(C.60)

Again, as for the overland flow cases, the coefficients β_5, m_5, and n_5 in

457

Equation (C.46) follow directly from the hydraulics of the environment. The hydraulic conductivity, K, though heterogeneous in space, is a well defined quantity, at least in its spatial mean. The seepage face area, hw, is perhaps less well defined. It is probably a function of the shape of the channel head and the depth of the soil, these properties being well defined within a catchment.

Note that the groundwater hydraulic gradient $\frac{dH}{dx}$ and the landscape slope S are not necessarily equal. There is no immediate connection between the variations in the groundwater table and variations in the landscape elevations. In Appendix A, where details of the computer model are discussed, it is noted that the discharge is obtained from an analysis of flow directions derived from the land surface topography. Groundwater flow directions may not follow the surface topography. For the groundwater flow activator the discharge used in the sediment transport equation will be different from that in the activator mechanism, the former from surface flow considerations, the latter from groundwater flow considerations. This issue is outside the scope of this work and will not be dealt with here. A slightly simpler case is when the pattern of elevation and groundwater head are the same, and where subsurface saturation is the dominant runoff mechanism. This is the case of O'Loughlin (1986). He assumes that the pattern of elevation change in the landscape and the groundwork phreatic surface are the same. The method for determining discharge used in this work is satisfactory in this case. Note, however, that channel growth on the basis of groundwater flow calls into question the whole idea of hillslope erosion resulting from overland flow that is implicit in Sections A.1 and 5.4.

In summary three different and physically based mechanisms for triggering channel growth have been examined. All three of these mechanisms can be formulated in the form of the generic activator equation of Equation (C.46). In each case the nonlinear dependence of the activator mechanism on discharge and space followed directly from the physics of the activation mechanism. In all three cases the nonlinear

dependence on discharge and slope was different. The derived activator equations are summarized in Table 5.2.

C.2.3 Scaling of the Activator Threshold

The differentiating feature between the different activation mechanisms is the ratio (m_5/n_5). If (m_5/n_5) is equal for two different mechanisms, then they are equivalent in their behavior except for a scaling of the activator number TA (see Chapter 6). A short justification of this assertion follows. Similarity with respect to TA was discussed in Chapter 6 for constant m_5 and n_5. The derivation below relaxes that limitation so that m_5/n_5 is constant but n_5 is variable.

The reader will recall that the region where activator exceeds the activator threshold is used solely to indicate where channelization can occur. Thus two different activator functions, in the same catchment, are equivalent if the regions where the activator exceeds the activator threshold are the same. This is equivalent the saying that the line $\underline{x} = \underline{C}$ along which $a(\underline{x}) = a_{threshold}$ is the same for both of the catchments. Thus for two catchments (1) and (2) we require

$$a_{(1)}(\underline{C}) = a_{(2)}(\underline{C}) = a_{threshold} \ \left(= \frac{1}{c_1} \right) \tag{C.62}$$

or casting this into nondimensional form using the activator number, TA,

$$TA_{(1)} \left[Q'(\underline{C}')^{m_5(1)/n_5(1)} S'(\underline{C}') \right]^{n_5(1)}$$

$$= \ TA_{(2)} \left[Q'(\underline{C}')^{m_5(2)/n_5(2)} S'(\underline{C}') \right]^{n_5(2)} \tag{C.63}$$

Now if

$$m_{5(1)}/n_{5(1)} = m_{5(2)}/n_{5(2)} = m_5/n_5$$

so that

$$\frac{m_{5(1,2)}}{n_{5(1,2)}} = 1$$

where

$$m_{5(1,2)} = \frac{m_{5(1)}}{m_{5(2)}}$$

$$n_{5(1,2)} = \frac{n_{5(1)}}{n_{5(2)}}$$

then Equation (C.63) may be rearranged to give

$$\left[Q'(\underline{C}')^{m_5/n_5} S'(\underline{C}') \right] = \left[\frac{TA_{(1)}}{TA_{(2)}} \right]^{\frac{1}{n_{5(2)} - n_{5(1)}}}$$

$$(C.64)$$

The right hand side is a constant independent of $\underline{x}$. To obtain a similarity condition relating $TA_{(1)}$ and $TA_{(2)}$ we raise Equation (C.64) to the power of $n_{5(1)}$, multiply by $TA_{(1)}$ and substitute for the left hand side with Equation (C.63). This gives

$$\frac{TA_{(2)}}{TA_{(1)}} = (1)^{n_{5(2)} - n_{5(1)}} = 1$$

$$(C.65)$$

Equation (C.65) is the similarity condition for activator for two catchments (1) and (2)

when $m_{5(1,2)}/n_{5(1,2)} = 1$. This similarity condition is identical with that derived in Chapter 6 from the governing equation (Equation 5.1).

The foregoing argument is the justification for the assertion that two activator mechanisms with equal (m_5/n_5) are equivalent except for a scaling coefficient. Table 5.2 contains a list of the activator mechanisms that have been examined in this section, their governing equation, and their m_5/n_5 ratios.

APPENDIX D

SIMULATION DATA

Run	TT 10^{-6}	TS	TC	TA 10^{-3}	TD 10^{-6}	O_t	$\frac{m_5}{n_5}$
CR1–3	1	1.76	1	15.77	0	0.1	1.33
CR1–4	1	0.35	1	15.77	0	0.1	1.33
CR2–3	1	1.76	1	9.46	0	0.1	1.33
CR2–4	1	1.76	1	5.68	0	0.1	1.33
CR2–6	1	1.76	1	12.61	0	0.1	1.33
CR3–3	1	1.76	1	15.77	0	0.01	1.33
CR3–4	1	1.76	1	9.46	0	0.01	1.33
CR3–5	1	1.76	1	9.46	0	0.001	1.33
CR4–1	1	0.43	1	10.72	0	0.1	1.33
CR4–2	1	7.19	1	13.88	0	0.1	1.33
CR4–3	1	0.16	1	10.72	0	0.1	1.33
CR4–4	1	20.1	1	13.88	0	0.1	1.33
CR4–5	1	7.19	1	13.88	0	0.1	1.33
CR4–6	1	1.76	1	9.46	0	0.1	1.33
CR4–7	1	1.76	1	9.46	0	0.1	1.33
CR4–8	1	1.76	1	9.46	0	0.1	1.33
CR4–9	1	1.76	1	9.46	0	0.1	1.33
CR4–10	1	1.76	1	9.46	0	0.1	1.33
CR5–1	1	1.76	10	9.46	0	0.1	1.33
CR5–2	1	1.76	10	15.77	0	0.1	1.33

Run	TT 10^{-6}	TS	TC	TA 10^{-3}	TD 10^{-6}	O_t	$\dfrac{m_5}{n_5}$
CR7–1	1	1.76	1	12.61	0	0.1	1.33
CR7–2	1	1.76	1	12.61	0	0.1	1.33
CR7–3	1	1.76	1	12.61	0	0.1	1.33
CR8–1	31.0	— Run to dynamic equilibrium of CR2–3					
CR8–3	13.9	— Run to dynamic equilibrium of CR10–2					
CR9–6	1	6.68	1	2.59	0	0.01	0.50
CR9–8	1	1.68	1	2.59	0	0.01	0.50
CR9–10	1	0.10	1	2.24	0	0.01	0.50
CR9–11	1	0.73	1	2.44	0	0.01	0.50
CR10–1	1	1.76	1	9.46	0.02	0.1	1.33
CR10–2	1	0.03	1	1.03	1.44	0.1	1.33
CR11–1	1	1.76	1	3.53	0	0.1	0.92
EB11–9	1	0.85	1	169.0	0	0.1	1.33

All runs used $n_5=0.3$, $m_1=1.8$, $n_1=2.1$ except
EB11–9 which used $n_5=0.75$

Regression of Drainage Density with Nondimensional numbers

Simulations used for the calibration of the regression were

CR1–3, CR2–3, CR3–3, CR3–4, CR3–5, CR4–2, CR4–5, CR4–6, CR4–7, CR4–9, CR5–1, CR5–2, CR9–6, CR9–11.

Simulations used for the verification of the regression were

CR2–4, CR2–6, CR11–1

Regression of Drainage Density Growth Rate with Nondimensional numbers

Simulations used for the calibration of the regression were

CR1–3, CR1–4, CR2–3, CR3–3, CR3–4, CR3–5, CR4–1, CR4–2, CR4–3, CR4–4, CR4–5, CR4–6, CR4–7, CR4–8, CR4–9, CR4–10, CR5–1, CR9–6, CR9–11.

Simulations used for the verification of the regression were

CR2–4, CR2–6, CR5–2