Hillslope Form and Process

M.A.CARSON and M.J.KIRKBY

HILLSLOPE FORM AND PROCESS

by M. A. CARSON
Associate Professor of Geography, McGill University

and M. J. KIRKBY
Lecturer in Geography, Bristol University

CAMBRIDGE
AT THE UNIVERSITY PRESS 1972

Published by the Syndics of the Cambridge University Press
Bentley House, 200 Euston Road, London NW1 2DB
American Branch: 32 East 57th Street, New York, N.Y.10022

Library of Congress Catalogue Card Number: 74-163061

ISBN: 0 521 08234 X

Printed in Great Britain by Alden & Mowbray Ltd
at the Alden Press, Oxford

PREFACE

The study of hillslopes has attracted geomorphologists for a long time and it is viewed by many as the most central single theme in the subject. Much research has been directed to the topic over the past two decades, but, to date, no attempt has been made to produce a comprehensive survey of it in the way that Leopold, Wolman and Miller's *Fluvial processes in geomorphology* did for stream channel form and process. A similar survey for hillslopes has been our aim in writing this book. It is directed at the graduate level, but it should also be readily understood by any final-year undergraduate with a basic science background. We have clearly drawn heavily on the work of others, often geomorphologists or pedologists, but equally engineers working in the fields of rock mechanics, soil mechanics or hydraulics. We feel strongly that slope studies and geomorphology generally must be approached from an inter-disciplinary point of view.

Although we have both done our best to give a balanced account of the topics on which we have written, there remain minor differences in viewpoint between the sections which the two of us have written. In a subject where few, if any, conclusions are final, we have felt it honest to retain such differences and leave the reader free to accept either or neither version. At the same time it may assist him to distinguish these intended differences if we list our responsibilities for individual chapters, as follows:

M. A. Carson: Chapters 2, 4, 6, 7, 11, 12, and 15;

M. J. Kirkby: Chapters 1, 3, 5, 8, 9, 10, 13, 14, and 16.

However, we would like to affirm that we consider this book to be a joint venture, with responsibility equally shared between us: our names appear in the title in alphabetical order according to bibliographical practice.

Parts of the manuscript have been critically reviewed by our colleagues, especially J. B. Bird, R. J. Chorley, D. Ingle Smith and Eiju Yatsu, to all of whom we owe a large debt. Similarly we would like to thank the Editorial Board of the Cambridge Geographical Studies for their advice and continued patience in the preparation of the text. Lastly our thanks go to our wives for their constant assistance in reviewing, proof-reading and living with the manuscript during the last two years.

April 1971 M. A. CARSON M. J. KIRKBY

ACKNOWLEDGEMENTS

Thanks are due to the following authors, publishers and learned bodies for permission to reproduce figures and passages from books and journals.

Addison-Wesley (fig. 7.10); Allyn and Bacon (fig. 7.7); American Geophysical Union (figs. 4.1, 8.15); American Journal of Science (figs. 6.5, 6.9, 6.11, 9.4, 9.17, 11.8, 13.6); American Society of Civil Engineers (figs. 7.3, 10.7); The Arctic Institute of North America (fig. 12.3); The Association of American Geographers (fig. 12.1); Biuletyn Peryglacjalny (fig. 12.6); Butterworths (fig. 7.8*c*); Cambridge University Press (figs. 2.4*a*, *b*); R. J. Chorley; Civil Engineering (fig. 3.6); D. R. Currey; Die Erde (fig. 15.5); Dover Publications; Geografiska Annaler (figs. 10.9, 12.4); Geographical Analysis (figs. 7.16, 7.17); The Geological Society of America (fig. 15.1); K. J. Gregory; Institute of British Geographers (figs. 11.7, 16.8); The Council of the Institution of Civil Engineers, London (figs. 4.10, 4.19, 7.12); International Association for Scientific Hydrology (fig. 16.21*a*); H. Jenny; L. C. King; D. Koons; Koninklijke Nederlandse Akademie Van Wetenschappen; W. B. Langbein; Longman (fig. 11.3); Macmillan & Co. Ltd (London) (figs. 2.1, 2.2, 15.3, 15.4); McGraw-Hill Book Co. (fig. 4.12); M. A. Melton; Methuen & Co. Ltd (figs. 3.13, 7.5, 8.9); National Research Council – Highway Research Board, (U.S.A.) (fig. 4.13); Nature Norwegian Geotechnical Institute (figs. 6.8, 7.13; table 4.1); Office of Naval Research, (U.S.A.) (fig. 16.20); J. C. Pugh; A. Rapp; Royal Geographical Society (Geographical Journal) (fig. 12.2); Royal Society of London (fig. 3.8 (inset)); S. A. Schumm; A. N. Strahler; Soil Science (figs. 3.4, 3.10); University of Chicago Press (fig. 15.2); University of Toronto Press (fig. 4.7); U.S. Department of Agriculture (figs. 3.12, 8.13); U.S. Geological Survey (figs. 9.2, 9.3, 16.21*b*); Vierteljahsschrift der Naturforschenden Gesellschaft (figs. 6.13, 6.14, 6.15); D. E. Walling; A. Young; E. G. Youngs; Zeitschrift fur Geomorphologie (figs. 6.10, 12.5).

CONTENTS

INTRODUCTION:

CHAPTER 1

GEOMORPHIC SYSTEMS AND MODELS

Hillslopes have the distinction of being the commonest and, at the same time, least studied geomorphic features, especially in terms of the processes acting on them. In some ways it is their very ubiquity which is responsible for this neglect because researchers have preferred to study unique or restricted features instead of facing the massive sampling problem which is involved in characterizing slopes. In addition, hillslopes present difficult research problems because their forms change either slowly or infrequently, particularly when compared to rivers.

The study of slopes is essential not only to an understanding of natural landscapes, but also as a practical means for controlling erosion and sedimentation which result when man modifies the landscape through agriculture, engineering construction or dumping operations. Cultivation and grazing have long accelerated natural rates of surface wash and have led to the initiation of gullies; and this aspect of soil erosion has been studied intensively particularly since the 1920s. Stripping of vegetation for agriculture has also induced more rapid wind erosion in some areas. Engineering construction often produces artificially steep slopes in cut or fill material, and the aim of design procedures is to reduce the tendency of such slopes to fail suddenly in a landslide, especially in the short term.

Geomorphologists, in common with other environmental scientists, study complex systems which contain within them many interactions and feedbacks. The two main ways of beginning to understand a system are, first inductively, through measuring landscape variables at sample points and analysing the results using statistical, usually multivariate, methods, and second deductively, building models on established physical laws whenever possible. These two approaches are inevitably interlocked, the inductive approach helping to generate some deductive models and testing others. In this book our emphasis is on landforms and the processes which form them, and consequently on deductive process–response models rather than on statistical analysis of data which are necessary to verify them.

Different workers have approached geomorphic problems from different standpoints which have, often unintentionally, limited the range of problems which could be tackled and the types of solution which were acceptable. There is not one correct paradigm of this sort: each is the product of its time and stimulates a period of fruitful research before being rejected. Geomorphic fashions have swung from the qualitative statements of early observers to the highly mathematical treatments of Bakker and Scheidegger, and back again to inductive studies derived from field measurement (Carson, 1969*b*). The approach of this book, with its emphasis on process–response models, is no less a part of current fashion but seems to represent a paradigm which is not yet exhausted. To examine this approach, we must delve a little into systems and model methodology.

SYSTEMS

Harvey (1969, p. 451) has described a system as containing:

'(1) A set of elements identified with some variable attribute of objects.
(2) A set of relationships between the attributes of objects.
(3) A set of relationships between those attributes of objects and the environment.'

In geomorphology, the objects are normally landforms, and their attributes consist of the topographic, soil, vegetation, etc., properties of the landforms. The relationships between these attributes consist of exchanges of energy, debris or water so that there are causal or other links between the attributes. The choice of elements which belong to the system, as opposed to the environment, varies with the problem and with scale, as will be discussed below.

An important distinction which clarifies the different approaches of geomorphologists, is that between open and closed systems. A closed system is one in which there are no relationships between the system elements and the environment. While it is trivially true that no geomorphic system is closed, much of classical physics and chemistry has developed from closed system models and Chorley (1962) has pointed out the extent to which closed system *thinking* has influenced geomorphic thinking. Closed systems typically consist of near-frictionless perpetual motion machines, for example the solar system; or else of systems in which initial variations in mechanical and thermal energy are progressively smoothed out to produce a totally uniform, constant-temperature equilibrium. The former type of equilibrium is somewhat analogous to a geomorphic 'steady state' in which the elements of the system are simply constant through time and unaffected by the environment over the time scale considered. The latter type of equilibrium is analogous to a peneplain type of equilibrium in which the landscape system is progressively running down towards a dead level uniform plain. This is not to suggest that these geomorphic equilibria are supposed to occur within a closed system, only that they are types of equilibrium which arise from closed system *thinking*.

In a closed system, the sum of mechanical and thermal energies is conserved, but only in a frictionless system is each *separately* conserved. Otherwise frictional losses convert mechanical energy into heat energy and the system shows a progressive increase in entropy, which is itself defined in two ways which can be shown to be equivalent. If an amount of heat energy ΔH is associated with material at absolute temperature θ, then the entropy of the system, S is given by:

$$S = \sum_{\theta} \frac{\Delta H}{\theta}. \tag{1.1}$$

Alternatively the molecules of the system can be considered to be in discrete energy states 1, 2, 3 . . ., n, a proportion p_i of the molecules being in state i. On this definition, the entropy of the system,

$$S \propto -\sum_{i=1}^{n} p_i \log p_i. \tag{1.2}$$

It can be shown that at equilibrium, when all mechanical energy has been transformed to heat, the entropy is at a maximum, corresponding to a condition of constant temperature, and, in the unconstrained case, to a condition in which all the p_is of equation 1.2 are equal. This condition is one in which there is a minimum of organization or differentiation between the various states; or a condition for which the *a priori* knowledge of the system is at a minimum.

In an open system, which has links with the external environment, the energy of the system is no longer conserved, and equilibrium positions and criteria for equilibria are more difficult to define. Open systems may run down towards a static equilibrium, but many maintain a dynamic equilibrium (Hack, 1960) which responds to variations in the external environment, and is maintained by negative feedbacks which react against an environmental stimulus so that the system partly accommodates the stimulus and partly arranges itself to resist the stimulus. This self-regulation mechanism has been formalized, initially in chemistry, as Le Châtelier's principle, which states:

> Any system in . . . equilibrium undergoes, as a result of a variation in one of the factors governing the equilibrium, a compensating change in a direction such that, had this change occurred alone it would have produced a variation of the factor considered in the opposite direction. (Prigogine and Defay, 1954, p. 262.)

This statement is almost identical to the definition proposed by Mackin (1948) for 'grade' of a river; but its possible geomorphic implications have not been fully explored.

An alternative criterion for equilibrium in an open system which has been proposed for organisms (Denbigh, 1951) is that the *rate of production* of entropy is a minimum. If entropy is seen as a measure of differentiation, it can be seen that the survival of an organism is dependent on maintaining its internal differentiation, and that a breakdown in differentiation is fatal to it. In a system where temperature differences are unimportant, this condition is equivalent to one of minimum work; a concept which has been widely used as a criterion in otherwise indeterminate mechanical systems, and has been applied to river meanders (Bagnold, 1960). The analogy between a geomorphic system and an organism is an attractive one, because both exhibit a steady progression over a long time span, in conjunction with a near-equilibrium in response to short-term fluctuations. It is therefore reasonable to hypothesize that geomorphic features measured over decades can be treated as being in a true dynamic equilibrium like that of an organism, and not merely as showing a trend which is obscured by large amounts of random noise in short measuring periods.

The concept of minimum entropy production has been used in a somewhat different manner by Leopold and Langbein (1962). An analogy was made between heat energy and mechanical energy, and between absolute temperature and elevation above base level, to define entropy as in equation 1.1. As with other analogies it must be tested within the appropriate system. The thermodynamic argument only shows that a quantity (entropy) which is conserved in a frictionless (reversible) cycle must have the form:

$$\Delta S = \frac{\Delta H}{f(\theta)}, \tag{1.3}$$

where f can be any function of θ; and the particular function, f, must be chosen with reference to the particular system. For a river system where θ is elevation, the function can only be a constant, so that minimum entropy production corresponds to minimum work. In the same paper Leopold and Langbein attempt to estimate geomorphic distributions in terms of most probable states, an approach which is linked to the concept of entropy in its information theory context, and which has since led to the development of a theory of minimum variance (Langbein and Leopold, 1966) for meanders and other features of channel geometry.

The discussion of landscapes as systems leads naturally to a comparison with other systems, which provide a fruitful source of analogies and allow hypotheses to be set up for the geomorphic system. This procedure has been formalized as General Systems Theory (von Bertalanffy, 1951) and has been discussed at length in the geomorphic context by Chorley (1962), but it has not reached a level at which conclusions for *general* systems can be applied to particular systems without re-testing in the particular system.

TIME AND SPACE SCALES

Schumm and Lichty (1965) have considered the conflict between short-term equilibrium and long-term evolution for landscapes, referred to

TABLE 1.1. *Systems for the study of drainage basin variables* (modified from Schumm and Lichty, 1965, table 1)

Variables	Cyclic	Graded (*a*)	Graded (*b*)	Steady
Climate	*E*	*E*	*E*	*E*
Geology	*E*	*E*	*E*	*E*
Regional relief	*S*	*E*	*E*	*E*
Slope forms	*S*	*S*	*E*	*E*
Soil properties	*S*	*S*	*E*	*E*
Vegetation properties	*S*	*S*	*S*	*E*
Drainage density	*S*	*S*	*S*	*E*
Sediment discharge	*S*	*S*	*S*	*S*
Channel geometry and micro-morphology	*I*	*S*	*S*	*E*
Water discharge	*I*	*I*	*S*	*S*

E = environmental variable.
S = element of system studied.
I = irrelevant or meaningless variable in system studied.

above, in terms of the choice of a suitable system. If we wish to examine geomorphic change over a particular time span, then features which change only in much longer periods may be considered as fixed, that is as belonging to the external environment. Landscape features which change in the chosen time period form elements of the system to be studied; and features which change very rapidly within the chosen time span can be considered as having average values which are a part of the system, and random deviations which are irrelevant to the system.

Schumm and Lichty distinguish cyclic, graded and steady time spans.

> *Cyclic time* . . . refers to a time span encompassing an erosion cycle . . . A fluvial system when viewed from this perspective is an open system undergoing continued change . . . (p. 113).
>
> The *graded time* span refers to a short span of cyclic time during which a graded condition or dynamic equilibrium exists . . . . When viewed from this perspective

one sees a continual adjustment between elements of the system, for events occur in which negative feedback (self-regulation) dominates (p. 114).

During a *steady time* span a true steady state may exist in contrast to the dynamic equilibria of graded time. These brief periods of time are referred to as a steady time span because in hydraulics steady flow occurs when none of the variables involved in a section change with time (p. 115).

The periods implied by these definitions depend on the rate at which a particular landscape is developing, but might, under normal conditions, refer to periods of about 10^6, 10^2 and 10^{-2} years respectively. Table 1.1 (adapted from Schumm and Lichty, table 1) indicates systems for study which may be appropriate to the time scale we are concerned with, but close scrutiny shows that almost all the variables listed may be split into components which are relevant to each time span. For example, long-term average climate and climatic change are relevant to *cyclic time*; frequency distributions of present climate are relevant to *graded time*; and instantaneous rainfall and temperature values are relevant to *steady time*. The *graded time* span has been divided into two in table 1.1, because *graded time* spans mean different things in slope and river studies. *Graded time* (*b*) corresponds more closely to Schumm and Lichty's usage, and appears to correspond to a time period in which the channel geometry and channel extension reach equilibrium, say ten- to one-hundred-year periods. *Graded time* (*a*) refers to a longer period, say 1,000 to 100,000 years, in which slope profile forms reach an equilibrium to which the word grade has also been applied.

For the longer time spans of *cyclic time* and *graded time* (*a*), the system is essentially a sediment system, and the system can be defined in terms of sediment distribution and flows, and their variation over time and space. In this system the flows of water are incidental, merely providing an agent for the transportation of the debris. For the short time periods of *graded time* (*b*) and *steady time*, however, the overall topography is fixed and the system describes the flows of water over this surface. Sediment discharge remains a system variable, but only 5 per cent or less of the available flow energy is absorbed in debris transportation. This water system is entirely relevant to hydrologic studies and forms most of the system for considering equilibrium hydraulic geometry and drainage density, but it is of only marginal relevance to slope studies. Since all measurements are made in *steady time* spans and even long periods of record apply only to *graded time* (*b*) for typical systems, slope studies are limited by the problem of extrapolating to long time periods. It is clear that the critical variable must be sediment discharge, since only this is a systems variable in all time spans; but it is also clear that extrapolations must be rather crude over long time spans.

Although *cyclic* and *steady time* spans refer to very long and short time periods respectively, the establishment of *graded time* implies that dynamic equilibria exist, and that the *graded time* span is not merely an artifact of the period of accurate records. If a dynamic equilibrium exists, then a *graded time* system can be examined in which the fluctuations of the system variables need not be referred to events in longer time spans. All that has been said about time scales applies equally to space scales, and we may study systems which refer to a single point or quadrat; to the whole world; or to intermediate areas which show maximum interaction and minimal linkage with other areas. For slope studies, the appropriate spatial systems appear to be the drainage basin and the slope profile. Because slope profiles, defined as lines of greatest slope, may migrate laterally during their evolution, they are less satisfactory than drainage basins in minimizing external links, but their greater simplicity outweighs this disadvantage, particularly if we consider profiles on which the contours remain straight so that lateral migration is minimized.

On a slope profile sediment is transferred from point to point, with little or no transfer across the profile, so that hillslopes can be simplified to two-dimensional spatial systems with a relatively small loss of information and considerable gain in analytical simplicity. In this book, therefore, we have concentrated on the slope profile system over long periods of time; that is, we view it as a sediment system. In chapter 16, we will return to drainage basins, to examine the extent to which the third dimension modifies the slope models considered in this book, but it is as well here to anticipate a little and state that the modifications entailed are relatively slight.

MODELS OF HILLSLOPE SYSTEMS

Many early workers, with the notable exception of Gilbert, were preoccupied with the way in which the external environment changes through cyclic time, but it seems that we can only examine this after we have constructed sound models linking process and form within the system (Carson, 1969*b*). Theory construction proceeds, as always, by an alternation of deductive and inductive steps (fig. 1.1). For example, a study may set out to test a model; and will do this by (1) making deductions which are testable, (2) designing an experiment to collect suitable data, (3) analysing this data to a form which is comparable with the deductions from the model and (4) testing. If the test is satisfactory, then the model is not disproved and may be checked against other areas: otherwise the model must be modified using fresh concepts to refine it.

Davis (1899) is chiefly responsible for the historical bias of geomorphology towards genetic, evolutionary models, but although his general evolutionary framework remains valid (see chapter 2), the neglect of both process and slope geometry by Davis and his successors limits the explanatory power of their deductions. In this book we are seeking to examine processes and profiles in a deductive framework which is compatible with established physical laws. The basic models are causal, although this term is taken to include causal feedback loops.

On the process level, the causal model is of a mechanical equilibrium between *forces* which tend to transport material and *resistances* which

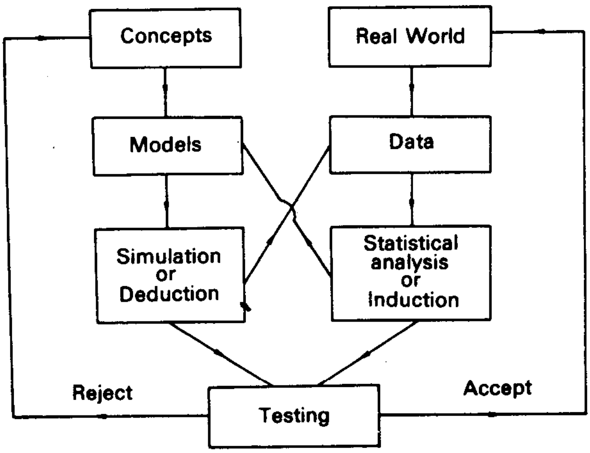

Fig. 1.1. The inductive–deductive model building process.

oppose the movement. Analysis is directed to an understanding of the factors responsible for the movement, and also to an understanding of temporal and spatial variations in its rate of action. On the landform level, the model seeks to link spatial variations in transport rate to sequences of slope profile development; that is, a process–response model. Formally this may be done as a differential equation (see chapters 5 and 15) which relates the temporal process rate to the spatial change in hillslope forms. Despite Hack's (1960) paper on dynamic equilibrium, little has been done on the application of equilibrium models to hillslope forms, although the criteria for open system equilibria appear to be applicable. It might, however, be noted that equilibrium models have been most successful in river studies at the micro- or mesoscale, and least successful in predicting overall long profiles.

At an early stage in an investigation it may be preferable to aim at generating models after collecting and analysing data which are considered relevant to a range of possible models. This inductive approach relies heavily on multivariate statistical methods, notably multiple regression and factor analyses. However, the inherent difficulty in making large numbers of reliable measurements, especially of process rates, so limits sample sizes that there are severe practical problems in verifying or generating models. The best that can be done is to choose sample points which are contiguous and so minimize extraneous differences, but which at the same time retain a full range of internal variation due to the operation of the system – in other words to select sample points from within slope profiles or basins, which thus form suitable sample clusters as well as suitable spatial systems. A more serious problem which cannot usually be overcome is the impossibility of adequate sampling over time spans relevant to slope profile development.

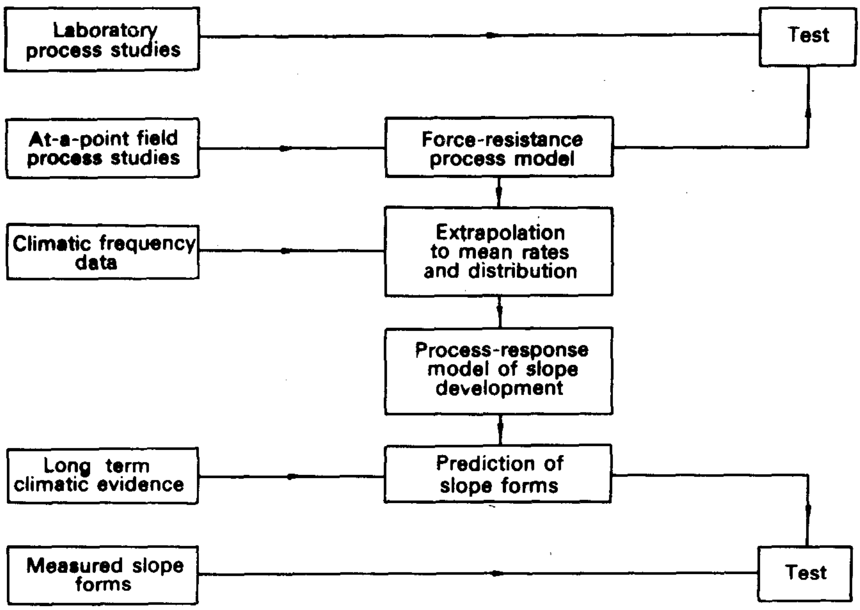

Fig. 1.2. The pattern of hillslope studies.

The most elegant solution to this problem is to find locations where space may be substituted for time (Savigear, 1952); but this is rarely possible in a simple way. Alternatively, we can invoke the ergodic hypothesis (Harvey, 1969, p. 128) and assume that the spatial distribution of a variable is the same as its temporal distribution. On a world scale, this is equivalent to an assumption of strict uniformitarianism – a rather doubtful hypothesis; and within a basin it implies that all profiles follow the same sequence, but at different rates – a statement which appears, on the available evidence, to be false.

The difficulties inherent in slope studies mean that the most effective studies are of slope *process* at a point over short periods, and of slope *forms* over an area. The process studies must be linked to a physical force-resistance model, which can be independently tested in some cases against laboratory experiments, and then extrapolated using climatic frequency data to produce long-term rates, and which then form the basis for a process–response model (fig. 1.2). The process–response model may be combined with long-term climatic data to produce predictions of slope forms which may be compared with measured forms, and this is the only other point at which an independent test can be made. If the test rejects the predicted forms, then the error may be derived from any of the stages, and not only from the original process measurements.

Although there is no unique way to examine slope forms and processes, the relationship between data and models shown in fig. 1.2 typifies most recent work, although it clearly excludes Davis' approach. This book follows the same pattern, although reflecting the preponderance of quantitative work on *process* rather than *form*: an imbalance in recent work which is perhaps a reaction to an earlier over-emphasis on slope forms, interpreted in the context of denudation chronology.