

Catena 46 (2001) 89-102



www.elsevier.com/locate/catena

Modelling the interactions between soil surface properties and water erosion Modelisation des interactions entre caractéristiques superficielles des sols et érosion hydrique

Mike Kirkby

School of Geography, University of Leeds, Leeds, UK

Abstract

Interactions between soil surface properties, water runoff and erosion occur at many time and space scales. The paper focuses primarily on the effects of soil microtopography at a range of spatial scales and considers some new conceptual approaches to modelling its effect on runoff generation, flow concentration and sediment transport. In this paper, the local generation of runoff is considered through the simplifying concept of a discrete runoff threshold. The effect of random roughness is considered independently in the down-slope and cross-slope directions. Roughness elements measured in the cross-slope direction concentrate the flow within depressions, and create greater total sediment transport. Roughness measured in the down-slope direction represents elements which pond or impede the flow, reducing mean flow velocity and reducing sediment transport. This two-component approach to roughness provides a reasonable approximation to behaviour on fractally generated two-dimensional surfaces. If the roughness elements are assumed to be normally distributed, and independently random in the two directions, then explicit integrations can be made over the distributions, providing significant corrections to the relation-ship between mean flow depth and sediments transporting capacity.

A second conceptual model is presented, exploring the implications of spatially random variations in surface flux capacity (storm total infiltration), surface (depression) storage and storm rainfall totals. It is demonstrated that some at-a-point runoff may, in principle, be generated, even where surface capacity appears to greatly exceed storm rainfall. In conclusion, it is argued that these effects, together with others not covered here, should be included in the next generation of soil erosion models. © 2001 Elsevier Science B.V. All rights reserved.

Keywords: Water erosion; Modelling; Time and space scales; Vertical and lateral soil variations

E-mail address: mike@geog.leeds.ac.uk (M. Kirkby).

1. Introduction

Soil erosion models commonly include parameters based on fixed soil properties or moisture-tension curves, but rarely take account of the dynamics linking soil and surface properties with water erosion processes, which are relevant at a range of time and space scales. The four sub-systems of Soil, Vegetation, Atmosphere and Surface interact dynamically with one another within each site, and sites are linked by lateral transfers of water, sediment and solutes in a downslope direction (Fig. 1).

Over short within-storm periods, soils properties influence erosion mainly through their hydrology, aggregate stability, and cation dispersion. Over periods of decades, erosion can lead to profile truncation and armouring of the surface, and these changes then influence subsequent rates and patterns of erosion. Over the still longer periods of soil evolution, the balance between mechanical and chemical denudation determines the entire weathering profile and its variation downslope to create hillslope catenas, which in turn also influence the distribution of erosion rates.

Similarly, erosion interacts with soil properties at a wide range of spatial scales. Macro-pores and crusting are examples of dominant influences at the finest $(10^{-3}-10^{-1} \text{ m})$ scales, and give way in importance to spatial patterning produced by cultivation or natural vegetation location at moderate (1-10 m) scales, to catena differences at the hillslope (10^2-10^3 m) scale, to major soil series at the catchment (10^4-10^5 m) scale and to climatically based major soil groups at continental (10^6-10^8 m) scales. In all cases, there are also strong links with soil via the vegetation cover, so that the direct



Fig. 1. Dynamic relationships between sub-systems at a site and downslope.

importance of the soil is greatest under sparse covers. The direct links with the soil are therefore strongest under arid and semi-arid climates for uncultivated areas. Under conditions of global change, relatively slow changes in soils and erosion rates are giving way to much more rapid modifications under the impact of climate and landuse change. Under these strongly non-equilibrium conditions, many 'natural' inter-relationships between the soils and the associated pattern of erosion will be broken down, demanding a more fundamental assessment of cause and effect in highly disturbed systems and making it difficult to use analogue studies with any confidence.

On natural semi-arid surfaces, microtopography evolves through sediment transfers and through the dynamics associated with mounds around perennial shrubs, with a time scale of decades to hundreds of years. On cultivated fields, microtopography is created by tillage operations and destroyed by sediment transfers over an annual cycle. These two systems have rather different dynamics but a number of important features in common.

This paper explores developments with respect to one set of processes which are believed to be important, and have been relatively neglected—the interactions between overland flow, water erosion and surface microtopography. This theme is being explored through research in a number of related erosion models, which are concerned with the wide range of relevant spatial and temporal scales. Some are primarily related to hydrology, others to soil erosion under global change while others are concerned with the evolution of soils and landforms over geological time spans. These include the MEDALUS and MEDRUSH models (Kirkby et al., 1997, 1998), the Regional Degradation Index (RDI/CSEP: Kirkby and Cox, 1995) and a number of models for long-term landscape evolution (e.g. Kirkby, 1991, 1992, 1994).

2. Overland flow and sediment transport on rough surfaces

Roughness due to random irregularities on a soil surface is usually well represented as a normal distribution of elevations relative to a mean, and is characterized by its standard deviation, or RMS roughness. There may also be various forms of systematic roughness due to tillage operations or, at a coarser scale, to terracing. Their orientation with respect to the slope direction is vitally important, particularly for long linear features like plough ridges.

The scale dependence of roughness is not fully understood, but it is clear that roughness occurs at a number of scales, and natural roughness may well have a fractal or similar structure, with larger RMS roughness as it is measured with a coarser grid of points. At sufficiently coarse scales, roughness of natural surfaces merges into landscape elements such as ephemeral gullies and stream valleys. Cultivated areas also show systematic roughness elements, most strikingly at the scales of plough furrows and terrace systems.

Depressions concentrate ponding, and, if net rainfall persists, progressively join together to form flow paths. After the end of rainfall, infiltration persists longer in depressions as they drain. It may be helpful to distinguish two components of roughness:

cross-slope roughness measured from transects taken across the slope, along the generalised contour direction, which is related to the connectivity of flow downslope along a sequence of depressions; and down-slope roughness (from transects down the generalised line of steepest descent), which tends to encourage ponded storage of stationary water. In some environments, there is a consistent relationship between position in the microtopography and infiltration capacity. It is more common for low points to be less permeable than highs, due to crusting of depressions and/or low bulk density associated with plants on higher areas (plough ridges or vegetation mounds), although the reverse is also possible.

Various natural processes modify surface roughness over time. Microtopography is reduced by rainsplash, which tends to fill depressions at the expense of the higher points, and is more effective in higher frequency microtopography, where height differences are associated with steeper local gradients. This process is very important in cultivated areas, where plough ridges tend to be poorly consolidated and therefore highly vulnerable. Once connected pathways of connected overland flow are established, erosion along them increases the cross-slope roughness, whereas deposition tends to reduce it. This effect is more important in uncultivated areas, where flow paths persist, and are not destroyed by tillage, although frost and wetting–drying cycles may have a similar effect (Schumm, 1964). Vegetation mounds associated with perennial shrubs also tend to increase roughness in semi-arid areas, by protecting the soil beneath from rainsplash and by the concentration of burrowing animals beneath them.

A 'dead-zone' model has been developed for flow over rough hillslopes (Kirkby and McMahon, 1997). Surface Roughness within the flow strip model is represented by cross-slope and down-slope roughness elements. Cross-slope random or regular (plough furrows etc.) roughness elements concentrate flow and sediment transport along depressions, and rougher surfaces generate greater sediment transport. Down-slope roughness elements create closed depressions within random linear flow paths, behind plough ridges and at terrace borders. They are represented as a distributed dead zone of zero flow, using a model which has proved effective in rivers draining areas of up to 10 km² (Beven, 1979). For the part of the flow which takes place at levels above the dead zone, the velocity is considered to be at constant rate, c. This model gives a strongly nonlinear relationship between mean flow depth and discharge (Fig. 2), but with a constant kinematic routing velocity. Although clearly a simplification, this two-parameter model of the surface is able to satisfactorily reproduce many of the features of overland flow hydrographs, and provides a plausible model for the spatial redistribution of sediment transport, and consequent roughness changes. Comparisons with flow simulated for inclined isotropic fractal surfaces show that the two-component model provides an acceptable first approximation to the more general situation, illustrated in Fig. 3 for a surface with an excess dimension of 0.2. For the fractal surface, the reduction in discharge is greater at very low values, but satisfactorily models the break-away from the 1:1 line for no dead zone. The computational advantages gained by using a constant routing velocity are considerable, as is demonstrated by comparable work on the Geomorphological Unit Hydrograph (Rodriguez-Iturbe and Valdez, 1979). For a surface with cross-slope random roughness of z and dead-zone depth of z^* , the total flow at stage y (relative to the mean surface level) may be explicitly integrated over the normal



Fig. 2. Discharge, q, vs. mean depth, y, where z = RMS roughness, $c = routing velocity and <math>z^* = \text{'dead-zone'}$ storage.

distributions, provided that the cross-slope and down-slope distributions are assumed to be independently random. This integration gives:

$$\bar{y}(y) = \frac{1}{z\sqrt{2\pi}} \int_{-\infty}^{z} (y - y') \exp\left(-\frac{{y'}^2}{2z^2}\right) dy'$$

$$q(y) = \frac{c}{z\sqrt{2\pi}} \int_{-\infty}^{y - z^*} (y - z^* - y') \exp\left(-\frac{{y'}^2}{2z^2}\right) dy' = c\bar{y}(y - z^*), \quad (1)$$

where $\overline{y}(y)$ is the mean flow depth at stage y.



Fig. 3. Simulated discharge vs. mean depth for an inclined fractal surface, with excess dimension 0.2.

For simulated fractal surfaces imposed on a uniform overall slope gradient, the progress of storage can be simulated as rainfall is added. Ponded areas grow until they drain or coalesce, and an increasing fraction of additional water runs off. The maximum storage is scaled to the roughness and decreases with gradient (Fig. 4), along a family of curves, each defined by a particular realisation of the random surface. This computational approach, though less general than Eq. (1), allows for the inclusion of infiltration, the development of crusting and erosional evolution of the microtopography, although this work is still at an early stage.

Simulations and theoretical considerations indicate that the dominant roughness elements are those for which the mean flow depth is less than the roughness at that scale. These dominant (emergent) roughness elements sub-divide the flow into threads or channels, whereas smaller elements are submerged and subsequently have little influence on flow or sediment transport, as may be seen from Eq. (1) above which shows that, for large depths ($y \gg z^*$), $\overline{y}(y) \sim y$. Similarly, emergent roughness is able to direct the flow, so that flow directions along plough furrows may be substantially different from the macroscopic downslope gradient vector. Thus, the controlling roughness elements tend to be more widely spaced in larger event, and generally, because discharge tends to increase progressively, downslope. On uncultivated land, there is generally a gradual increase in roughness downslope, even before the formation of discrete channels (Dunne and Aubry, 1986), and these patterns develop over periods of perhaps centuries after the abandonment of agricultural land, through the growth of vegetation mounds (in semi-arid rangelands) and through some differential sediment movement. For systematic roughness elements under cultivation, the same argument about the dominance of emergent roughness elements implies that there is a similar shift in dominance from microtopography to furrow patterns to field systems for larger storms and for larger areas.



Fig. 4. Relationship between log(Storage) and Gradient on a fractally rough surface $(32 \times 32 \text{ cells})$ with an excess dimension of 0.2). Curves show different replicate runs with same parameters.

For sediment transport in rough flows, similar explicit integrations can be made across random roughness elements (Kirkby et al., 1997). If transporting capacity is proportional to discharge squared, a reasonable first approximation in many cases, then the effect of the roughness term is to add a second term to create an excellent approximation:

$$S = 0.85K(q^2 + qq_0), (2)$$

where K is an erodibility term which includes the effects of gradient, q is the width-averaged water discharge, as obtained above in Eq. (1) and q_0 is the 'roughness discharge' = cz.

Thus, it may be seen that the sediment transport at low discharges is greater, but less effective in incising rills and channels on rough surfaces, due to the dominantly linear form of the second term in Eq. (2).

A similar analysis may be applied to the distribution of grain sizes on the surface. Assuming equal mobility in detachment, and differential deposition with a travel distance inversely proportional to grain size, the difference between transported and deposited grain sizes can be simulated from the moments of the grain size distribution. Fig. 5 shows the difference in phi grain sizes calculated for a range of transport stages (indicated by travel distances) and sorting (indicated by the variance of phi grain size of the surface material). Full numerical summations over the distributions can be simply forecast in terms of the phi-variance, as shown by the heavy line in Fig. 5. Similarly, the



Fig. 5. Difference in phi grain size between transported and deposited material. Plotted points from full numerical integration over log-normal grainsize distributions. Line is for simplified model $(T - D) = 0.6\sigma^2$, where *T*, *D* are, respectively, the phi mean sizes for transported and deposited material, and σ is the phi standard deviation of the transported material.

Mean travel distance for deposited material can be well estimated from the travel distance for the D_{84} , that is for the phi mean + 1 standard deviation.

These simulations show that roughness effects can be effectively incorporated within erosion models which retain an acceptable level of simplicity, with the addition of only the parameters for cross-slope and downslope roughness. These theoretical methods still require experimental validation, but are based on well founded physical and mathematical principles.

3. Surface threshold erosion model

Pursuing these concepts at a coarser scale, of fields within a catchment, following the concepts developed by Auzet et al. (1993) which appear to lend support to the theoretical conclusions above, it seems that surface and above-surface properties may be more important in estimating overland flow runoff than detailed consideration of sub-surface properties. From this basis, a simple conceptual model is proposed for the spatial distribution of at-a-point storm runoff generation. The central simplifying parameter discussed here is the threshold of runoff, considered as the depth of water entering the soil before runoff occurs, and closely related to the concepts of total or dead-zone storage above. This threshold clearly has a spatial distribution, which may be characterized by a mean and variance, and might be represented by a normal, gamma or log-normal distribution. The discussion focuses on agricultural land, but can also adapted to uncultivated vegetation.

At a local scale, which may represent a single field (or partition with uniform tillage directions), conditions may be described by a series of layers (Fig. 6). Each layer may be considered as transforming the distribution of rainfall intensity and rainfall energy. The layers are here treated as statistically independent, but may be significantly correlated, for example though the concentration of crusting in depressions. Typical transformations of mean rainfall intensity or energy flux as it passes through a layer will take the form:

$$\Delta S = I - L(I, S, \Xi) - Q(S, \Xi), \tag{3}$$

where S is the storage in the layer, I is the inflow, L is a loss term, Q is the outflow and Ξ represents environment variables.

This expression controls changes in the mean flux, introducing delays and losses as appropriate, and is sufficiently general to cover most transformations of rainfall amount or energy. For example, crown cover may simply attenuate energy as $\exp(-\text{LAI})$, while the leaves also act to intercept and evaporate some of the water.

A second equation is needed to estimate areal re-distribution of the transmitted throughfall as it passes through the canopy and soil. Each layer tends to increase the variance of the distribution by selective interception in some patches (e.g. the vegetation crown canopy or areas of better developed surface crust) and concentration in other patches (e.g. stemflow and leaf drip around the edges of a shrub). For simplicity, this re-distribution term is not further considered here, but it clearly depends on the spatial correlation or morphological structure of the surface and its vegetation.

Referring to Fig. 6, each of the layers has a different dynamic, and it is important to make sure that these have the correct interactions, and that each changes appropriately over time. The vegetation crown and stems respond to crop growth, which should itself be responding to the environment in terms of rainfall, irrigation, etc. Surface stoniness and microtopography are being modified by each cultivation. Between cultivations, roughness and crusting respond strongly to freeze–thaw and rainfall energy, which reduce roughness and increase crusting in susceptible soils. Finally, soil moisture, through a simple infiltration and evapotranspiration model, influences the distribution of runoff thresholds for the water which has reached it. An effective model based on these principles clearly has to be run as a continuous simulation, because storages interact continuously with new rainfall over time.

Fig. 7 illustrates the ways in which the layers can interact, in the simplest possible case, to generate at-a-point overland flow runoff. If is assumed here that the distribution



Fig. 6. Conceptual partitioning and attenuation of water and energy.



of storm rainfall amount, and of the storage characteristics of each layer, can be described by a gamma distribution, and that the layers are statistically independent of each other. In Fig. 7, the distribution of storm rainfall amount is shown as 5 ± 1 mm (mean ± 1 standard deviation), giving the frequency distribution shown. Two layers are shown, representing the surface depression storage capacity $(10 \pm 4 \text{ mm})$ and the sub-surface storage capacity $(3 \pm 2 \text{ mm})$. Their combined effect is obtained by convolution, so that a total capacity of, say 12 mm, can be obtained with 12-mm surface storage plus 0-mm sub-surface storage; 11-mm surface plus 1-mm subsurface; 10-mm plus 2-mm; and so on. This convolution gives the frequency distribution for the total storage of the two layers combined, which is, as expected, roughly equal to the sum of the two capacities, peaking at 13 = 10 + 3 mm, as can be seen in Fig. 7.

Runoff is generated where the total storm rainfall exceeds this combined storage capacity. A comparable convolution is required to estimate how much of the rainfall runs off. It may be seen in Fig. 7 that most runoff is generated where the higher rainfall amounts combine with the lower storage capacities. That is where the rainfall and combined storage curves overlap, in the region between 4 and 10 mm. Although this example shows quite a small total runoff (0.026 mm out of a total storm rainfall of 5 mm), it is clear that only a consideration of the spatial distributions can allow any rational estimation of the runoff total.

The convolution of rainfall with the composite properties of the canopy and soil layers is relevant at a range of scales, from the behaviour within a small plot to the response of whole catchments. It is clear that, in many cases, storm runoff is generated from only small contributing areas within both semi-natural (e.g. Bull et al., 2000) and largely agricultural (e.g. Auzet et al., 1993) catchments. Although the assumption of statistical independence of the sub-surface storage layers and of rainfall intensity, which

have been made here, may not be valid in some cases, it has a simplicity which may outweigh the necessarily greater data requirements of a more inter-correlated model. One critical issue is clearly the structure of interactions within an area of study. Where there are strong preferential flow paths, either within the soil as macro-pore flow; or within catchments as overland flow channels, not only does the flow network and its properties become dominant, but the system becomes highly nonlinear. In this case, issues of connectivity become critical and local runoff generation has to be explicitly linked to the observed flow pathways, whether of soil macro-pores, artificial field drains or dynamically eroding gully networks.

In addition to these spatial distributions of storage over a small area, which lead to more runoff than is indicated by a simple subtraction of storage capacity from rainfall, particularly at low intensities, there are effects due to temporal variations in rainfall, and the interaction between these two components. Fig. 8 shows the result of estimating at-a-point runoff as a function of cumulative rainfall, under the slightly more complex assumption of a Green and Ampt (1911) infiltration equation, in the form:

$$f = A + B/S$$
 or *i*, whichever is less
 $S = \int_0^t (f - A)$ where the integral is taken over positive excursions. (4)

In Fig. 8, steady rainfall intensities are compared with rainfall of random intensity and exponential distribution. As well as the expected variability for random sequences, it is clear that random intensities generate considerably more runoff than the same average rate. This figure also provides some commentary on the use of the assumption of a



Fig. 8. Simulated point runoff from Green–Ampt Equation (A = 5; B = 50), for intensities constant and varying randomly with exponential distribution.

constant storage threshold which has been used above. It is argued that this simplified form is often an adequate description of the process, unless very detailed rainfall intensities are available, although it is of course recognised that the relationship is still dependent on average rainfall intensity.

The temporal variability of rainfall and the spatial variability of storage capacity interact as flows are routed downslope. Runoff generated at a point only reaches the slope base where there is not complete re-infiltration as it crosses intervening areas. Fig. 9 illustrates a simulation of this interaction, for independent random exponential distributions of rainfall and storage. It can be seen that, as expected and widely observed in the field, the runoff coefficient is small, and declines downslope where mean rainfall is less than mean storage capacity, but that some flow is generated from the areas nearest the slope base. When rainfall exceeds average storage, the relationship reverses, and the effect of run-on saturates the soils and produces some increase in the runoff coefficient downslope.

Clearly this distance-discharge relationship has an important impact on the increase of both water and sediment discharge downslope. The overall effect appears to be that, after summing these relationships over the frequency distributions of storm rainfall, discharge and sediment discharge both increase much less than linearly with area drained, and that this effect is stronger in more arid (or more permeable) areas. Discharge and sediment discharge may be taken to increase with area as:

$$q \propto a^n$$

$$S \propto q^2 \propto a^{2n},$$
(5)

where a is the area drained per unit contour width and n is an empirical exponent.



Fig. 9. Simulated downslope runoff coefficients for randomly varying runoff threshold and rainfall intensities over time.

The implication of the simulations reported here is that the exponent *n* depends on the relationship between mean storm rainfall and mean storage capacity, and shows a general increase from arid to humid climates, probably along the range n = 0-5-1.0.

The discussion above is for Hortonian overland flow. If comparable arguments are applied to Saturation overland flow, which is preferentially produced at the base of long hillslopes, their behaviour may be seen as providing n values $\gg 1.0$, so that the slope-length dependence of sediment transport equations is strongly dependent on climate, with consequences for the concavity of the associated slope profiles.

4. Conclusion

It is argued that interactions between surface properties, principally microtopography and crusting, influence the quantity and spatial pattern of surface runoff. These interactions occur at all scales from that of the individual roughness element up to cultivated fields and catchments. Although the parameterisation of roughness has been seen to be scale and storm dependent, the concept of effective storage capacity or runoff threshold appears to be robust across the range of relevant scales, and is, in all cases, strongly dependent on soil surface properties and their distribution.

Random variations in both storm intensity and storage or infiltration characteristics also influence the amount and distribution of runoff, and the interaction between these variabilities produces strong dependencies between mean storm size and slope length for a given area. All of these inter-relationships are beginning to be understood well enough that we can begin to build them into effective erosion and landscape development models. They also provide an important conceptual basis for scaling up from single plot areas to catchments scales, allowing better understanding of experimental results and their application to larger areas.

Acknowledgements

The author acknowledges the financial support and intellectual stimulation provided by the MEDALUS project (EC, DG XII), the GCTE Soil Erosion Network and the RIDES project (PNSE/PNRH).

References

- Auzet, A.V., Boiffin, J., Papy, F., Ludwig, B., Maucorps, J., 1993. Rill erosion as a function of the characteristics of cultivated catchments in the North of France. Catena 20, 41–62.
- Beven, K.J., 1979. On the generalised kinematic routing method. Water Resources Research 15, 1238-1242.

Bull, L.J., Kirkby, M.J., Shannon, J., Hooke, J.M., 2000. The variation in estimated discharge in relation to the location of storm cells in SE Spain. Catena 38, 191–209.

Dunne, T., Aubry, B.F., 1986. Evaluation of Horton's theory of sheetwash and rill erosion on the basis of field experiments. In: Abrahams, A.D. (Ed.), Hillslope Processes. Allen and Unwin, London, pp. 31–53.

- Green, W.H., Ampt, G.A., 1911. Studies on soil physics I: the flow of air and water through soils. J. Agric. Sci. 4, 1–24.
- Kirkby, M.J., 1991. Sediment travel distance as an experimental and model variable in particulate movement. Catena Supplement 19, 111–128.
- Kirkby, M.J., 1992. An erosion-limited hillslope evolution model. Catena Supplement 23, 157-187.
- Kirkby, M.J., 1994. Thresholds and instability in stream head hollows: a model of magnitude and frequency for wash processes. In: Kirkby, M.J. (Ed.), Process Models and Theoretical Geomorphology. BGRG Symposium. Wiley, Chap. 14.
- Kirkby, M.J., 1998. Modelling across scales: the MEDALUS family of models. In: Boardman, J., Favis-Mortlock, D. (Eds.), Modelling Soil Erosion by Water. Springer, Berlin, pp. 161–173.
- Kirkby, M.J., Cox, N.J., 1995. A climatic index for soil erosion potential (CSEP) including seasonal and vegetation factors. Catena 25, 333–352.
- Kirkby, M.J., Mcmahon, M.L., 1997. A dead-zone model for flow over rough hillslopes. Paper Presented atEGS Annual Conference, April.
- Kirkby, M.J., Abrahart, R., McMahon, M.D., Shao, J., Thornes, J.B., 1997. MEDALUS soil erosion models for global change. Geomorphology 24, 35–49.
- Rodriguez-Iturbe, I., Valdez, J.B., 1979. The geomorphological structure of Unit Response. Water Resources Research 15, 1490–1520.
- Schumm, S.A., 1964. Seasonal variations of erosion rates and processes on hillslopes in western Colorado. Zeitschrift fur Geomorphologie, Supplementband 5, 215–238.