Earth Surface Processes and Landforms
Earth Surf. Process. Landforms **33**, 1113–1128 (2008)
Published online 23 January 2008 in Wiley InterScience
(www.interscience.wiley.com) **DOI:** 10.1002/esp.1622

A transport-distance approach to scaling erosion rates: 3. Evaluating scaling characteristics of MAHLERAN

John Wainwright,[1]* Anthony J. Parsons,[1] Eva N. Müller,[2] Richard E. Brazier,[3] D. Mark Powell[4] and Bantigegne Fenti[5]

[1] *Sheffield Centre for International Drylands Research, Department of Geography, University of Sheffield, Sheffield, UK*
[2] *Institut für Geoökologie, Universität Potsdam, Potsdam, Germany*
[3] *Department of Geography, University of Exeter, Exeter, UK*
[4] *Department of Geography, University of Leicester, Leicester, UK*
[5] *Queensland Department of Natural Resources, Mines and Energy, Indooroopilly, Queensland, Australia*

**Correspondence to: John Wainwright, Sheffield Centre for International Drylands Research, Department of Geography, University of Sheffield, Winter Street, Sheffield S10 2TN, UK. E-mail: j.wainwright@sheffield.ac.uk*

Received 21 May 2007; Revised 7 September 2007; Accepted 27 September 2007

Abstract

In the two previous papers of this series, we demonstrated how a novel approach to erosion modelling (MAHLERAN – Model for Assessing Hillslope-Landscape Erosion, Runoff And Nutrients) provided distinct advantages in terms of process representation and explicit scaling characteristics when compared with existing models. A first evaluation furthermore demonstrated the ability of the model to reproduce spatial and temporal patterns of erosion and their particle-size characteristics on a large rainfall-simulation plot. In this paper, we carry out a more detailed evaluation of the model using monitored erosion events on plots of different size. The evaluation uses four plots of 21·01, 115·94, 56·84 and 302·19 m^2, with lengths of 4·12, 14·48, 18·95 and 27·78 m, respectively, on similar soils to the rainfall-simulation plot, for which runoff and erosion were monitored under natural rainfall. Although the model produces the correct ranking of the magnitude of erosion events, it performs less well in reproducing the absolute values and particle-size distributions of the eroded sediment. The implications of these results are evaluated in terms of requirements for process understanding and data for parameterization of improved soil-erosion models. We suggest that there are major weaknesses in the current understanding and data underpinning existing models. Consequently, a more holistic re-evaluation is required that produces functional relationships for different processes that are mutually consistent, and that have appropriate parameterization data to support their use in a wide range of environmental conditions.

Keywords: **erosion; sediment transport; soil-erosion model; scaling; validation; data**

Introduction

In previous papers (Wainwright *et al.*, in press a, b) we presented and tested a soil-erosion model MAHLERAN (Model for Assessing Hillslope-Landscape Erosion, Runoff And Nutrients) that we believe to be based upon a more robust conceptualization of soil-erosion processes than previous models. A key argument underpinning the development of this model is that a model that is conceptually sound will reproduce observed relationships between erosion rate and scale of observation, as has been demonstrated by the analytical approach of Wainwright *et al.* (2001) and Parsons *et al.* (2004). In this paper, we build on the spatial and temporal testing of the MAHLERAN approach using simple rainfall events in rainfall simulation (Wainwright *et al.*, in press b) to provide a more robust test of the approach employing more realistic conditions. The aim of this paper is first to test the performance of the model at the runoff-event scale against data obtained from a range of plot sizes, and second to use the results of this test to discuss issues associated with soil-erosion modelling.

Table I. Summary of events used in model testing

Event date	Abbott		Dud		Laurel		Wise	
	Plot runoff (l)	Sediment yield (kg)	Plot runoff (l)	Sediment yield (kg)	Plot runoff (l)	Sediment yield (kg)	Plot runoff (l)	Sediment yield (kg)
30/7/2000	1272·32	7·81						
10/8/2000	1509·81	7·51						
11/8/2000	63·46	0·41	3·84	0·08	27·99	0·18		
20/8/2000	29·18	0·30	24·11	0·24	59·93	0·45		
5/8/2001			38·80	0·90	14·36	0·57	5·58	2·03
11/8/2001	47·57	0·44	57·24	0·28			31·29	1·68
12/9/2001	53·59	0·86	63·87	0·42			65·04	0·88
19/7/2002			37·43	0·28	47·95	0·44		
26/7/2002	579·94	2·05			265·71	2·00		
4/8/2002							3780·15	27·39

Field Data

To evaluate the performance of the model at a range of scales, detailed testing data are required that meet several criteria. First, they must relate to a variety of scales of measurement, with other environmental variability reduced to a minimum. Secondly, there must be detailed rainfall, runoff and sediment-production data for a range of storm events. These data must be of a high quality so that confidence can be placed in the test results. Thirdly, it is helpful if the focus of the test is on the sediment-transport component, and thus the data should relate to conditions where the hydrological conditions are reasonably well understood. The data presented by Parsons *et al.* (2006; Brazier *et al.*, 2007) meet all of these criteria. Four runoff plots (named Laurel, Abbott, Dud and Wise) were constructed at the Walnut Gulch Experimental Watershed, AZ, USA (31° 44′ 23″ N, 110° 3′ 53″ W) in areas with soils described as being on coarse-loamy, mixed, thermic, Ustochreptic Calciorthids and fine-loamy, mixed, thermic Ustalfic Haplargids (Breckenfield *et al.*, 1995) and vegetation dominated by creosotebush (*Larrea tridentata*), albeit with a number of other shrub species including *Acacia constricta*, *Dasylirion wheeleri*, *Rhus microphylla* and *Yucca elata*, with a ground layer dominated by *Dyssodia acerosa* and *Zinnia pumila*. These plots had areas of 21·01, 115·94, 56·84 and 302·19 m^2, respectively, and lengths of 4·12, 14·48, 18·95 and 27·78 m, respectively. The plots were monitored during natural rainfall events over the summer monsoon seasons of 2000 to 2002. Because of the issues concerning data quality raised by Parsons *et al.* (2006), only those events considered as providing high quality estimates of runoff and sediment production have been used in this model analysis. This selection provides a test data set made up of 22 example events (Table I). However, these events are unfortunately unevenly distributed among the plots, with six events at Laurel, seven at Abbott, five at Dud and four at Wise. Five storm events provide information at three different scales of measurement, but it did not prove possible to obtain measurements with high quality estimates for all four plot sizes simultaneously. However, as each storm is treated as a separate event in subsequent analyses in this paper, this lack of synchronicity (which is common in convective events in dryland regions) does not affect the conclusions drawn.

Digital elevation models (DEMs) were derived for all of the plots using a total station, surveyed on 0·5 m grids. Surface conditions were characterized using quadrats on the same grids as the DEMs to provide estimates of stone pavement and vegetation cover. Distributed particle-size information was then obtained using the same method as in Wainwright *et al.* (in press b), by using information from the fine fractions based on samples obtained by scraping off the surface layer of soil, scaled according to the percentage pavement cover. There were eight samples for the Laurel plot, 15 each from Dud and Abbott and 27 from Wise. Parsons *et al.* (2006) provide more details of the monitored-plot data, demonstrating their intercomparability with the large-plot data used in the model evaluation in Wainwright *et al.* (in press b).

Evaluation Of Mahleran Using Monitored Events

Parameterization

While in principle the parameterization used in our previous testing of Mahleran (Wainwright *et al.*, in press b) should hold for the monitored events, given that the parameters were derived on soils with similar physical characteristics,

DOI: 10.1002/esp

in practice there is some inconsistency in the results obtained with this parameterization. This inconsistency arises for a number of reasons. First, the monitored events have naturally varying rainfall, whereas the final infiltration rates used in the previous testing were derived for steady rainfall. The use of the Smith–Parlange (1978) infiltration model attempts to overcome this limitation (Wainwright and Parsons, 2002), but also presents further problems in that the depth to the wetting front must be known *a priori* and that it introduces a difficult-to-measure parameter (wetting-front suction, estimated using soil characteristics (Clapp and Hornberger, 1978), but not spatially distributed). Secondly, although rainfall intensity is a parameter in the estimation of final infiltration rate, the experiments upon which the relationship in Equation (1) in Wainwright *et al.* (in press a) was based had a much smaller range of intensities than those observed using 1 min interval data. However, there is no empirical evidence that the final infiltration rate would vary on such a short timescale, and so average event intensities have been used (which also overcomes the extrapolation issue). Thirdly, the initial soil-moisture content is a parameter of the Smith–Parlange model, but it could only be estimated in the present case, and is assumed to be spatially constant, although other information (Müller *et al.*, in press) suggests that this is a weak assumption. Fourthly, infiltration rates are known to vary through a rainfall season (Simanton and Renard, 1992 – see also the discussion in Wainwright *et al.*, 2000) and indeed dynamically through a rainfall event due to the formation and destruction of crusts (Luk and Cai, 1990; Parsons *et al.*, 2003). Notwithstanding the importance of these factors for the hydrology-model estimates, because the aim of this paper is to test the erosion model, a simple calibration procedure has been used to ensure the hydrology estimates are the best possible ones for the erosion modelling. By extension of the analysis in the previous paper (Wainwright *et al.*, in press b), we will also assume that these hydrology estimates produce the best available hydraulics estimates. The dominant vegetation on the monitored plots is the same as in the previous paper, so that the parameterization of the effect of vegetation cover on raindrop kinetic energy reaching the surface (Equation (4) in Wainwright *et al.*, in press b) can be used.

Results

For most of the events in Table I outflow hydrographs are available, and so calibration is carried out by maximizing the Nash–Sutcliffe (N–S) efficiency statistic (and minimizing the normalized root-mean square error, NRMSE) for the modelled versus observed hydrographs. Where hydrographs were not available due to equipment failure (Abbott, 11/8/2001; Dud, 11/8/2000; Laurel, 11/8/2000 and 19/7/2002), calibration was simply against total runoff. Calibration was carried out in two steps. First, the initial soil-moisture content (θ_0) was estimated based on knowledge of the pattern of rainfall events through the three seasons measured. If the model produced acceptable outflow hydrographs (in practice, N–S > 0·725), then no further calibration was carried out. Additionally, in cases where it was not possible to obtain better fits even using the more detailed calibration procedure discussed below, the results of this first calibration were used for simplicity, even if N–S < 0·725. Secondly, if it was not possible to produce an acceptable fit in this way, θ_0 was varied to see whether an acceptable fit was possible. Thirdly, only if variations in θ_0 could not produce satisfactory hydrographs were the values of final infiltration rate and wetting-front suction varied to optimize the goodness of fit. These variations were carried out by simple multiplication of the estimated base values. Fourthly, in a limited number of cases, it was also necessary to modify the effective depth to the wetting front. It is thought that this procedure provides the most parsimonious and data-based approach to calibration of the hydrographs. The results of this procedure are presented in Table II, with examples of calibrated hydrographs illustrated in Figure 1. Where the model performs poorly, it is typically in the case of events with multiple flow peaks, for which the timing but not the magnitudes of all the peaks is reproducible, and in some of the smaller sized events.

For each of the calibrated hydrographs on each of the plots, the erosion model was then applied using the optimized sediment-transport parameters obtained from the rainfall-simulation plot for each of the hydraulic relationships (Equations (2) and (3b) of Wainwright *et al.*, in press b). Comparison of modelled and measured sediment yield was only possible for total plot sediment yield (Table III), as time series were only available for suspended load in the supercritical flume at the plot outlet, and only then once a critical depth of flow triggered the auto-sampler. A summary of the results is presented in Figure 2. Abbott gives the best overall results, producing reasonable estimates for total sediment for all of the events except that on 30/7/2000. The Nash–Sutcliffe efficiency for modelled sediment yield relative to observed increases from −0·22 to 0·88 if this event is omitted (using Equation (2) in Wainwright *et al.*, in press b, hydraulics – the respective figures for Equation (3b) in Wainwright *et al.*, in press b, hydraulics are −0·83 and 0·91), although note that there is no obvious reason for omitting this event from the dataset given that the modelled hydrograph has a Nash–Sutcliffe efficiency of 0·93. Results for Dud tend to be underpredicted at low observed yields, but overpredicted for the highest yield (overall Nash–Sutcliffe efficiency is −1·13 for Equation (2) in Wainwright *et al.*, in press b, hydraulics and −2·59 for Equation (3b) in Wainwright *et al.*, in press b, hydraulics). Underprediction occurs throughout the range of values at Laurel (overall Nash–Sutcliffe efficiencies of −8·49 and −10·07 for hydraulic equations (2) and (3b) in Wainwright *et al.*, in press b, respectively), although again the pattern is dominated by a

DOI: 10.1002/esp

Table II. Hydrologic parameters used in the modelling of the runoff events from the monitored plots at Walnut Gulch. Parameters were derived by maximizing the Nash–Sutcliffe efficiency of the observed versus modelled hydrographs. The events for which an asterisk is present in the Nash–Sutcliffe efficiency and NRMSE columns have been fitted against total runoff alone, because equipment failures prevented fitting against the hydrograph. Events marked with a dagger represent the cases where the best possible fits were by calibration of initial soil-moisture content alone, but produced low Nash–Sutcliffe efficiencies

Plot/event	Initial moisture content (θ_0) ($m^3\ m^{-3}$)	Base value of wetting-front suction (ψ) (mm)	Effective depth to wetting front (m)	Final infiltration rate multiplier	Wetting-front suction multiplier	Nash–Sutcliffe	NRMSE (%)
Abbott							
30/7/2000	0·05	46·6	0·3	1	1	0·9309	27·7
10/8/2000	0·10	46·6	0·3	1	1	0·9651	21·7
11/8/2000	0·11	46·6	0·6	1·5	0·8	0·6567	129·5
20/8/2000	0·05	46·6	0·3	1	1	0·9679	29·9
11/8/2001	0·05	46·6	0·3	1	1	*	*
12/9/2001	0·02	46·6	0·3	1·55	1	0·9289	55·9
26/7/2002	0·03	46·6	0·3	1·8	0·5	0·8473	60·8
Dud							
11/8/2000	0·064 25	46·6	0·3	1	1	*	*
20/8/2000	0·06	46·6	0·3	1	1	0·3747†	73·4
5/8/2001	0·02	46·6	0·3	1·3	1	0·8935	137·8
19/7/2002	0·02	46·6	0·3	1	1	0·5900†	327·5
26/7/2002	0·02	46·6	0·3	1·4525	1	0·6992	112·4
Laurel							
11/8/2000	0·033 75	46·6	0·3	1	1	*	*
20/8/2000	0·015	46·6	0·3	0·125	4·0	0·8816	49·5
5/8/2001	0·022 5	46·6	0·3	1	1	–0·6400†	187·4
11/8/2001	0·032 5	46·6	0·3	1	1	0·6354†	41·8
12/9/2001	0·125	46·6	0·3	2	0·5	0·6248	72·4
19/7/2002	0·02	46·6	0·3	3·39	1	*	*
Wise							
5/8/2001	0·02	46·6	0·3	1	1	0·7250	222·6
11/8/2001	0·02	46·6	0·3	1·15	1	0·3096	163·7
12/9/2001	0·02	46·6	0·3	1·5	1	0·6600	93·3
4/8/2002	0·15	46·6	0·3	1	1	0·8014	72·7

single large outlier. In this case – 5/8/2001 – the hydrograph is very poorly fitted (Nash–Sutcliffe = −0·64), although removing this point still maintains negative Nash–Sutcliffe efficiencies (–0·62 and −1·47, respectively). The results for Wise are dominated by the extreme event of 4/8/2002, in which the runoff was nearly two orders of magnitude greater than the next largest event successfully monitored. However, in this case, the hydrograph was well reproduced (Nash–Sutcliffe = 0·80), which fails to explain why the model might produce such a significant overestimation of the sediment yield. All of the results at Wise are problematic, in that the modelled yields for the smaller events are in reverse order compared with the observed ones – the model suggests a positive relationship between total runoff and erosion, whereas the observed data show the reverse pattern. Because of these trends, removing the extreme event actually makes the model efficiency worse in the Equation (2) (in Wainwright *et al.*, in press b) hydraulic case (–14·00 with all data points compared with −18·02 with the extreme point omitted), although there is a slight improvement in the Equation (3b) (in Wainwright *et al.*, in press b) hydraulic case (the respective efficiencies being –25·12 and –19·76).

Thus, for none of the plots does the use of all the data produce Nash–Sutcliffe efficiencies that are better than using the average rate (equivalent to an efficiency value of zero). There is no obvious relationship between size of rainfall event and the squared deviation of modelled erosion from the expected value, except for the previously mentioned inverse case for the smaller events at Wise. Similarly, there is no relationship between the goodness of fit of the flow hydrograph and the squared deviation of modelled erosion from the expected value, although the hydrographs are typically best reproduced at Abbott. However, although the absolute estimated values may be incorrect in relation to the observed values, the results are appropriately ordered. Combining the data from all the plots, Spearman's rank

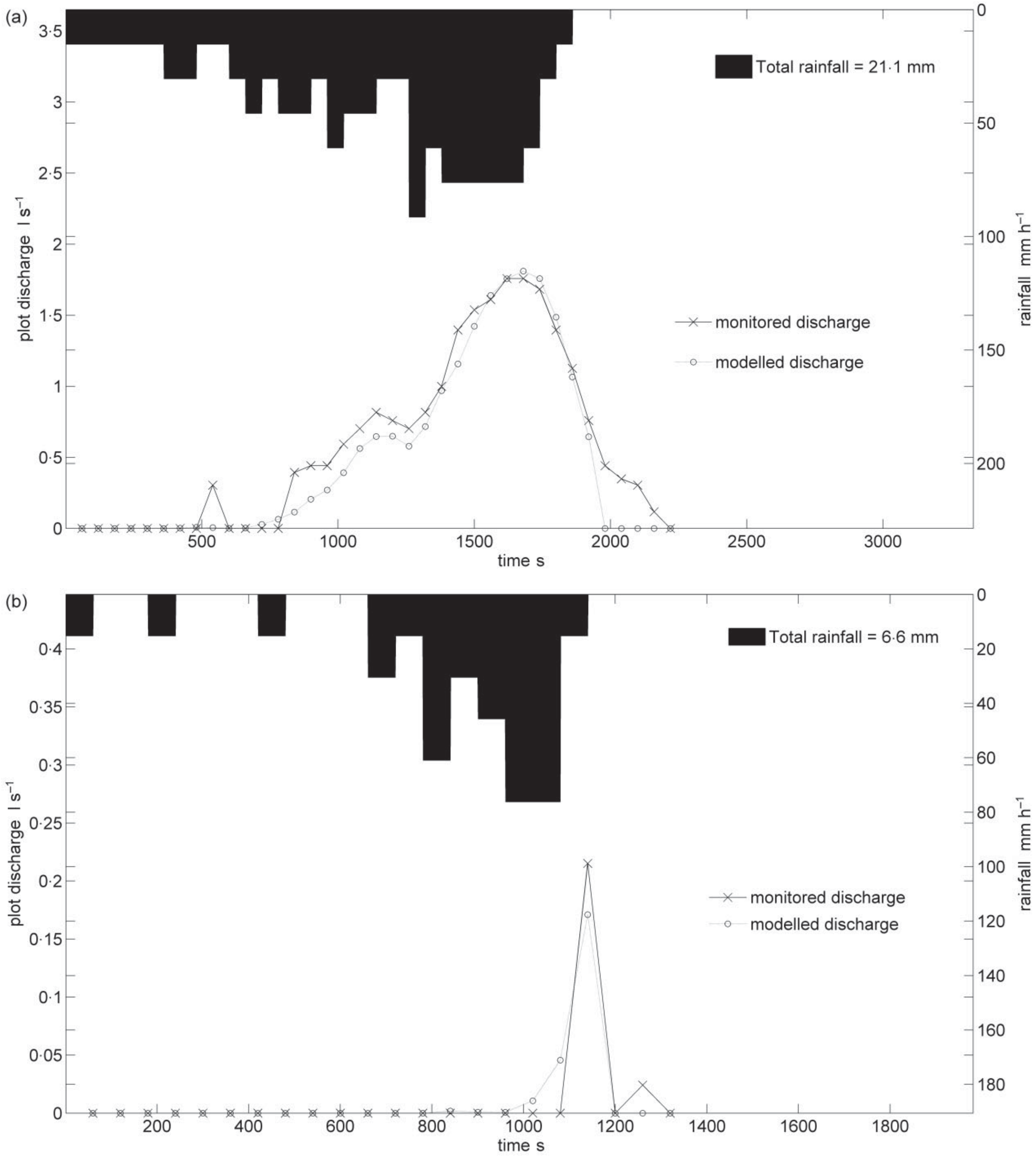

Figure 1. Examples of hydrographs used in the model for a range of event sizes and goodnesses of fit: (a) Abbott event 30/7/2000 – calibration only by modifying initial soil-moisture content; (b) Dud event 5/8/2001 – calibration of initial soil-moisture content and final infiltration rate; (c) Laurel event 5/8/2001 – calibration of initial soil-moisture content alone, as no other conditions were able to reproduce the first small peak or the timing of the main runoff peak; (d) Wise event 11/8/2001 – calibration of initial soil-moisture content and final infiltration rate, with no other conditions being able to reproduce the exact timing of the first hydrograph peak, although the second peak and total runoff are well reproduced.

DOI: 10.1002/esp

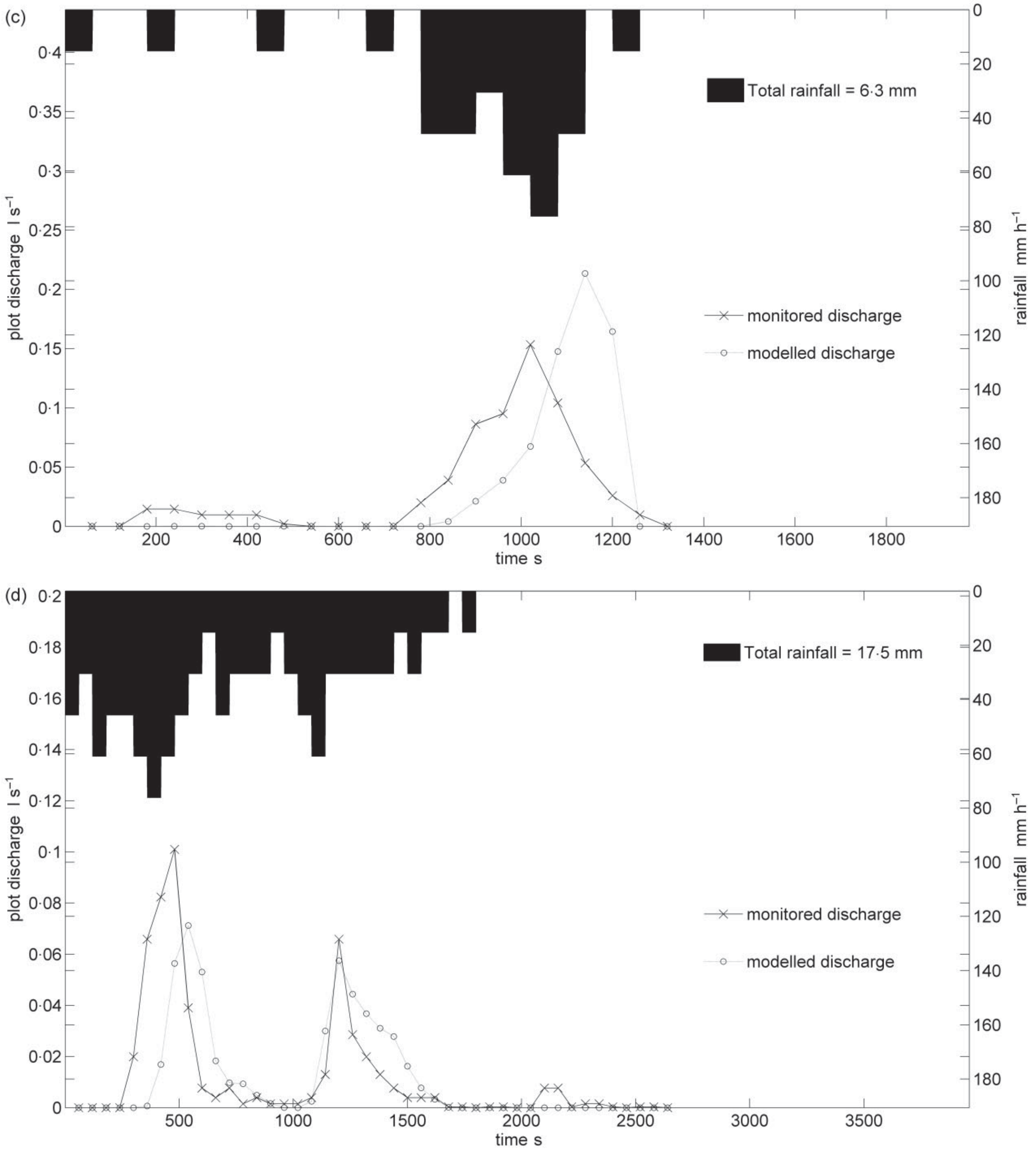

Figure 1. *Continued*

correlation coefficient comparing observed with modelled total sediment is 0·746 using Equation (2) (in Wainwright *et al.*, in press b) hydraulics and 0·715 using Equation (3b) (in Wainwright *et al.*, in press b) hydraulics ($n = 22$, $p < 0{\cdot}0005$ in both cases). This result suggests that the broad pattern of modelled results is correct (cf. Nearing, 2003), and that the cause of the errors in absolute estimates of sediment production is more due to problems with parameterization for specific conditions, rather than due to problems with the model structure.

DOI: 10.1002/esp

Table III. Summary of erosion-simulation results on monitored plots at Walnut Gulch

Plot/event date	Plot runoff (l)	Observed sediment yield (kg)	Proportions by model sediment size class*						Modelled sediment yield Eq. (2) (kg)	Proportions by model sediment size class*						Modelled sediment yield Eq. (3b) (k)g	Proportions by model sediment size class*					
			$\varphi = 1$	$\varphi = 2$	$\varphi = 3$	$\varphi = 4$	$\varphi = 5$	$\varphi = 6$		$\varphi = 1$	$\varphi = 2$	$\varphi = 3$	$\varphi = 4$	$\varphi = 5$	$\varphi = 6$		$\varphi = 1$	$\varphi = 2$	$\varphi = 3$	$\varphi = 4$	$\varphi = 5$	$\varphi = 6$
Abbott																						
30/07/2000	1272·32	7·81	0·05	0·24	0·22	0·24	0·23	0·03	4·19	0·29	0·70	0·01	0·00	0·00	0·00	3·30	0·28	0·71	0·01	0·00	0·00	0·00
10/08/2000	1509·81	7·51	0·50	0·25	0·10	0·15	0·00	0·00	8·09	0·26	0·73	0·01	0·00	0·00	0·00	7·49	0·25	0·74	0·01	0·00	0·00	0·00
11/08/2000	63·46	0·41	0·07	0·29	0·17	0·21	0·23	0·04	0·01	0·90	0·10	0·00	0·00	0·00	0·00	0·01	0·94	0·06	0·00	0·00	0·00	0·00
20/08/2000	29·18	0·30	0·11	0·31	0·17	0·21	0·14	0·06	0·04	0·41	0·59	0·00	0·00	0·00	0·00	0·02	0·46	0·54	0·00	0·00	0·00	0·00
11/08/2001	47·57	0·44	0·08	0·29	0·17	0·17	0·19	0·10	0·38	0·45	0·55	0·00	0·00	0·00	0·00	0·53	0·44	0·55	0·00	0·00	0·00	0·00
12/09/2001	53·59	0·86	0·17	0·24	0·12	0·17	0·21	0·08	0·47	0·31	0·68	0·01	0·00	0·00	0·00	0·58	0·36	0·63	0·01	0·00	0·00	0·00
26/07/2002	579·94	2·05	no data						1·59	0·36	0·63	0·00	0·00	0·00	0·00	1·41	0·34	0·66	0·00	0·00	0·00	0·00
Dud																						
11/08/2000	27·99	0·18	0·13	0·35	0·18	0·24	0·09	0·02	0·00	0·87	0·13	0·00	0·00	0·00	0·00	0·00	0·90	0·10	0·00	0·00	0·00	0·00
20/08/2000	59·93	0·45	0·17	0·34	0·21	0·20	0·08	0·00	0·24	0·28	0·71	0·00	0·00	0·00	0·00	0·29	0·28	0·71	0·00	0·00	0·00	0·00
5/08/01	14·36	0·57	0·06	0·25	0·15	0·19	0·28	0·07	0·28	0·23	0·76	0·01	0·00	0·00	0·00	0·32	0·23	0·76	0·01	0·00	0·00	0·00
19/07/2002	47·95	0·44	no data						0·22	0·44	0·56	0·00	0·00	0·00	0·00	0·19	0·40	0·60	0·00	0·00	0·00	0·00
26/07/2002	265·71	2·00	no data						2·96	0·25	0·74	0·00	0·00	0·00	0·00	3·31	0·26	0·74	0·00	0·00	0·00	0·00
Laurel																						
11/08/2000	3·84	0·08	0·04	0·21	0·16	0·26	0·33	0·00	0·00	0·99	0·01	0·00	0·00	0·00	0·00	0·00	0·99	0·01	0·00	0·00	0·00	0·00
20/08/2000	24·11	0·24	0·12	0·27	0·17	0·23	0·19	0·02	0·03	0·52	0·48	0·00	0·00	0·00	0·00	0·02	0·54	0·46	0·00	0·00	0·00	0·00
5/08/01	38·80	0·90	0·10	0·22	0·14	0·20	0·20	0·14	0·11	0·37	0·62	0·00	0·00	0·00	0·00	0·07	0·35	0·65	0·00	0·00	0·00	0·00
11/08/2001	57·24	0·28	0·11	0·25	0·16	0·22	0·22	0·05	0·12	0·47	0·53	0·00	0·00	0·00	0·00	0·09	0·52	0·48	0·00	0·00	0·00	0·00
12/09/2001	63·87	0·42	0·17	0·20	0·15	0·19	0·21	0·08	0·35	0·38	0·62	0·00	0·00	0·00	0·00	0·19	0·42	0·58	0·00	0·00	0·00	0·00
19/07/2002	37·43	0·28	no data						0·06	0·48	0·52	0·00	0·00	0·00	0·00	0·04	0·49	0·51	0·00	0·00	0·00	0·00
Wise																						
5/08/01	5·58	2·03	0·03	0·10	0·08	0·18	0·38	0·24	0·01	0·51	0·49	0·00	0·00	0·00	0·00	0·00	0·65	0·35	0·00	0·00	0·00	0·00
11/08/2001	31·29	1·68	0·06	0·14	0·10	0·21	0·33	0·16	0·21	0·40	0·60	0·00	0·00	0·00	0·00	0·08	0·42	0·58	0·00	0·00	0·00	0·00
12/09/2001	65·04	0·88	0·14	0·14	0·09	0·20	0·34	0·09	0·25	0·34	0·65	0·00	0·00	0·00	0·00	0·13	0·32	0·67	0·00	0·00	0·00	0·00
4/08/02	3780·15	27·39	no data						77·43	0·18	0·77	0·02	0·03	0·00	0·00	93·46	0·17	0·77	0·02	0·03	0·00	0·00

* $\varphi = 1$ (<63 μm), $\varphi = 2$ (63–125 μm), $\varphi = 3$ (125–1000 μm), $\varphi = 4$ (1–2 mm), $\varphi = 5$ (2–12 mm), $\varphi = 6$ (>12 mm).

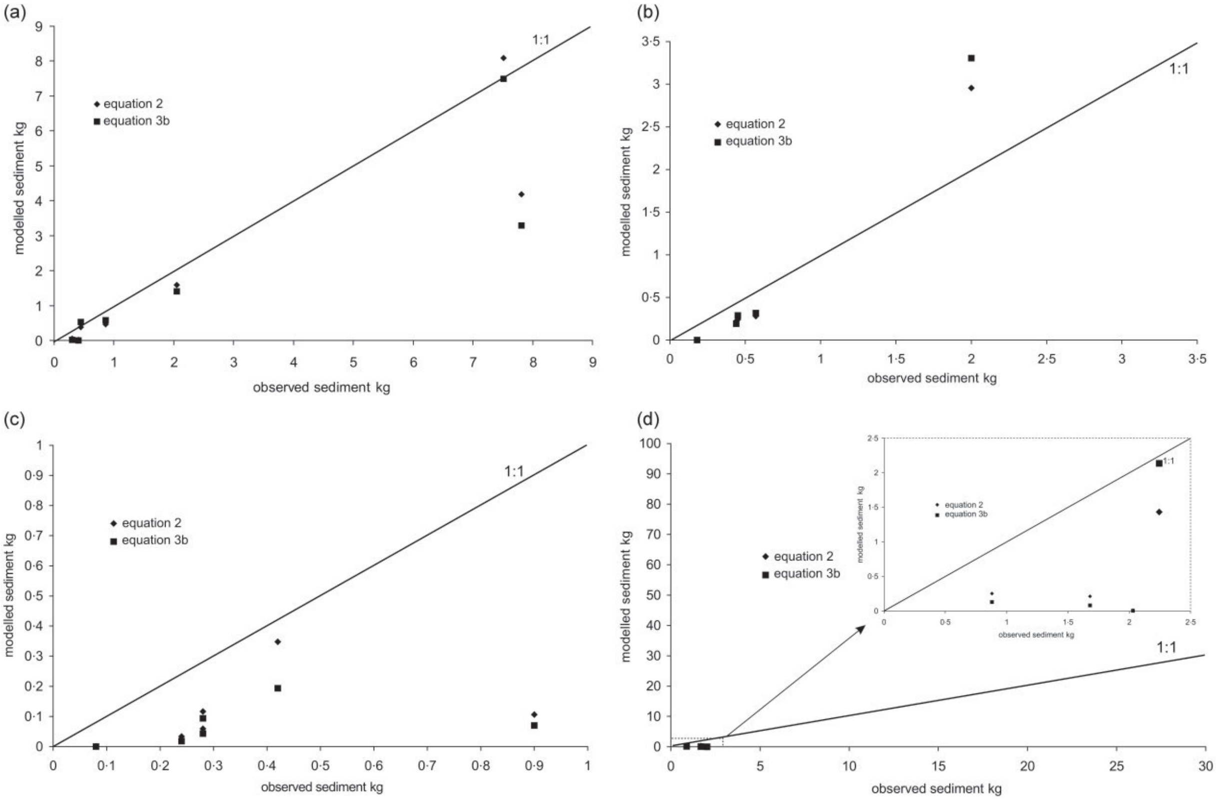

Figure 2. Summary of erosion-model results versus observed sediment yield for the monitored plot events at Walnut Gulch using optimized sediment-transport parameters for hydraulic relationships (2) and (3b) (in Wainwright *et al.*, in press b): (a) Abbott; (b) Dud; (c) Laurel; (d) Wise (with inset showing detail of events with small yields).

A comparison of observed and modelled particle-size distributions of the eroded sediment (Table III and Figure 3) shows that, although in the majority of cases the model correctly predicts the modal size class of eroded sediment, the overall distribution is poorly predicted. Specifically, the model produces a much narrower range of particle-size classes than is observed, and overpredicts the proportion of finer material. This result is worse than, but consistent with, that on the rainfall-simulation plot.

Issues with Model Discretization and Scaling Characteristics

One potential reason for the difference in model performance between that in Wainwright *et al.* (in press b) and the present case is the slight difference in cell size in the two cases. Because of the original sampling design, the former was discretized on a 0·61 m grid, whereas the latter used a 0·5 m grid, meaning that there was no obvious way in which data from one could be interpolated to have the same resolution as the other. It is thus possible that the optimization of the erosion parameters that was carried out contains an implicit scale dependence, which might explain some of the differences in erosion estimates, even where the hydrology is well reproduced. To evaluate the possibility of implicit scale dependence, a further analysis of sensitivity to model cell size was carried out on the 5°, 100 m long × 30 m wide, uniform, planar slope used in the sensitivity analyses in Wainwright *et al.* (in press b). All conditions were kept equal, apart from cell size, which was varied using values of 0·5, 1·0, 1·5, 2·0, 2·5, 5·0, 7·5 and 10·0 m.

The results of this sensitivity analysis are shown in Figure 4. There is a clear cell-size effect in the results, which is particularly marked on the upper part of the slope. The modelled peak event sediment flux is 0·285 kg m^{-1} and occurs at a distance of 7·5 m from the divide for the 0·5 m cell-size run. The locus of the peak flux increases distance downslope with increasing cell size, occurring at 9 m for 1 m cells, 10 m for 2·5 m cells, 15 m for 5 m cells, 15 m for

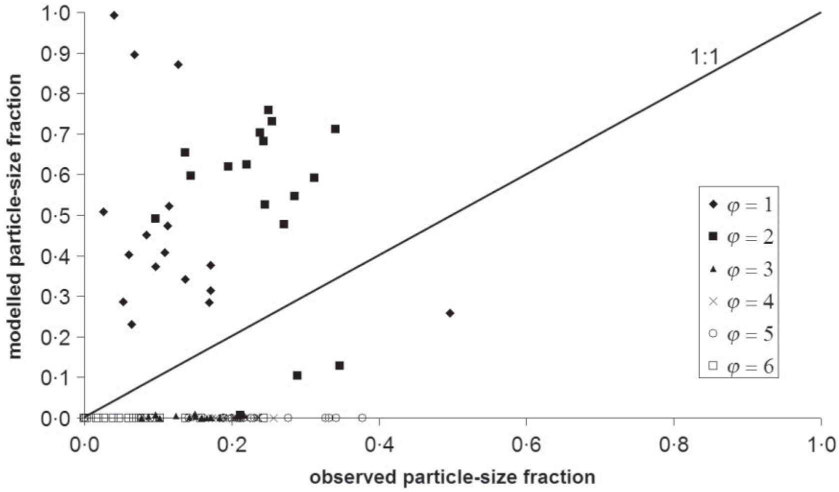

Figure 3. Observed versus predicted particle-size distributions of eroded sediment on the monitored plots.

7·5 m cells and 20 m for 10 m cells. There is also a decrease in the peak sediment flux, with values of 0·236 kg m^{-1} for 1 m cells, 0·143 kg m^{-1} for 2·5 m cells, 0·072 kg m^{-1} for 5 m cells, 0·040 kg m^{-1} for 7·5 m cells and 0·023 kg m^{-1} for 10 m cells. Differences in fluxes are less significant for greater distances downslope, 0·074 kg m^{-1} for 0·5 m cells, 0·067 kg m^{-1} for 1 m cells, 0·047 kg m^{-1} for 2·5 m cells, 0·032 kg m^{-1} for 5 m cells, 0·009 kg m^{-1} for 7·5 m cells and 0·004 kg m^{-1} for 10 m cells. Because of these changes in sediment flux, sediment yield is markedly different at the top of the slope, even accounting for the lower resolution of the larger cell sizes. Although both detachment and deposition follow very similar patterns both spatially and in terms of magnitudes, it can be seen that the net erosion rates are significantly different in the upper 30 m of the slope. Smaller cell sizes produce more rapid declines in net erosion, accounting for the greater curvature in the sediment-flux patterns downslope. It should be noted that these differences according to cell size are solely attributable to the erosion component of the model; in all cases, the hydrological component produces identical results at each point downslope for the different cell sizes.

To understand this cell-size effect, it is necessary to return to the sediment-continuity equation used in MAHLERAN (Equation (6c) in Wainwright *et al*., in press a):

$$\frac{\partial h_{s,\varphi}}{\partial t} + \frac{\partial q_{s,\varphi}}{\partial x} - \varepsilon_\varphi + d_\varphi = 0 \qquad (1)$$

where $h_{s,\varphi}$ is the equivalent depth of sediment in transport [m]
φ is an index relating to a specific size class of sediment
$q_{s,\varphi}$ is the unit discharge of sediment [m^2 s^{-1}]
ε_φ is the rate of erosion of the surface [m s^{-1}]
d_φ is the rate of deposition [m s^{-1}].

Of particular importance is the net erosion term ($\varepsilon_\varphi + d_\varphi$). For the model to scale correctly, it is necessary for the net erosion to have equal values at identical points on the slope using different cell sizes. To illustrate why this condition does not occur, we take the simple case comparing net erosion at a point 1 m from the top of the slope for cell sizes of 0·5 and 1·0 m, using the notation $\varepsilon_{\varphi,\Delta x,i}$ and $d_{\varphi,\Delta x,i}$, where Δx refers to the cell size and i refers to the position along the slope. One consequence of the travel-distance approach, as noted by Parsons *et al.* (2004), is the fact that deposition at any point on the slope can be directly estimated if the erosion upslope and travel distance are known. Thus, for the case with 0·5 m cells, we can note that

$$d_{\varphi,0\cdot5,1} = \varepsilon_{\varphi,0\cdot5,0\cdot5}[(1 - e^{-1L}) - (1 - e^{-0\cdot5L})] + \varepsilon_{\varphi,0\cdot5,1}[(1 - e^{-0\cdot5L}) - (1 - e^{-0\cdot5L})] \qquad (2a)$$

where the bracketed terms give the proportion of material deposited in a cell for a given travel distance L, assuming that travel distances are exponentially distributed. Simplifying, we have

$$d_{\varphi,0\cdot5,1} = \varepsilon_{\varphi,0\cdot5,0\cdot5}[e^{-0\cdot5L} - e^{-1L}] + \varepsilon_{\varphi,0\cdot5,1}[1 - e^{-0\cdot5L}] \qquad (2b).$$

DOI: 10.1002/esp

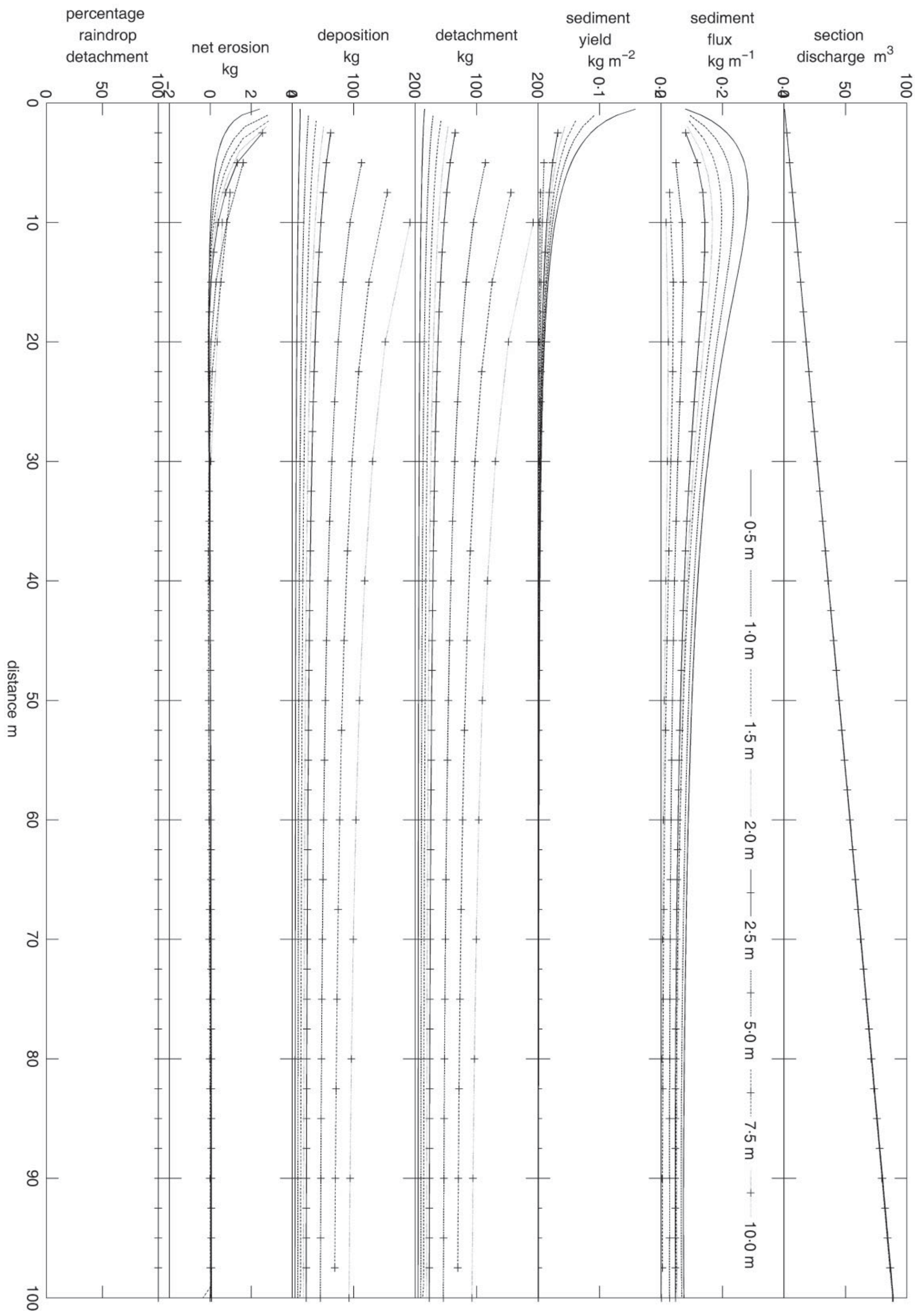

Figure 4. Sensitivity analysis to model cell size for a 5°, uniform, planar 100 m long × 30 m wide slope.

The equivalent expression for 1·0 m cells is

$$d_{\varphi,1,1} = \varepsilon_{\varphi,1,1}[1 - e^{-1L}]. \tag{3}$$

Thus, for process scaling, it follows that

$$\varepsilon_{\varphi,0\cdot5,0\cdot5}[e^{-0\cdot5L} - e^{-1L}] + \varepsilon_{\varphi,0\cdot5,1}[1 - e^{-0\cdot5L}] = \varepsilon_{\varphi,1,1}[1 - e^{-1L}] \tag{4}$$

or, in other words, that detachment must be spatially homogeneous. In the case of splash, this condition holds for a uniform slope as used in these sensitivity analyses. However, once flow occurs, there is a spatial pattern of detachment because of the feedback with flow depth. In the simplest case (Parsons *et al.*, 2004), flow depth will change in proportion to $x^{2/3}$ so detachment in the unconcentrated flow case will be proportional to $e^{-\beta_\varphi x^{4/9}}$, while travel distance will change in proportion to $x^{2/3}$ so that deposition will be proportional to $e^{-x^{2/3}}$. While it would be possible to produce a scaling factor that accounted for this difference in the simple, homogeneous uniform-slope case, its application would be meaningless in realistic conditions, where hydrological heterogeneity further complicates the pattern. More significantly, this lack of balance suggests a limitation of our understanding of different parts of the erosion process. We suggest that this limitation is derived from the empirical basis of the different elements from different sources (and thus sometimes incompatible simplifying assumptions). A more holistic approach to the derivation of the erosion model subcomponents is required to address the problem using consistent, process-based principles.

Discussion

In the review by Wainwright *et al.* (in press a) of the limitations of existing soil-erosion models, it was noted that a number of problems had arisen because untested hypotheses had passed unchecked into the literature. It is thus important to evaluate the extent to which the same case applies to the approach presented here. In some cases, the limitations will relate to the model structure or the empirical basis of process knowledge. However, in other cases, limitations are created by the lack of data for parameterization and model testing, or even by limited model evaluation, so it is necessary to consider these three sources of problems.

Model limitations

The model structure provides limitations in terms of the hydrology sub-model used to characterize spatio-temporal flow fields, and in particular the detail of the flow hydraulics. As noted above (and by Wainwright and Parsons, 1998), the flow hydraulics are critical in controlling a variety of erosion processes and the interactions between them. The current version of MAHLERAN employs a kinematic wave model with dynamic feedback between flow depth and roughness, so that although it is an improvement over most implementations of the kinematic wave approximation, it is still limited in its approach. Such limitations are discussed in detail elsewhere (e.g. Ponce, 1991; Singh, 1996; Moussa and Bocquillon, 1996). However, it is clear that there are major limitations in terms of how the friction characteristics of any flow model are dynamically variable spatially and temporally, and further work is required on this topic before models can be improved. Of particular importance is the transition to concentrated flows, and issues relating to these dynamics (e.g. Rice and Wilson, 1990; Giménez and Govers, 2001) and representation of processes at sub-grid scale. At present, MAHLERAN assumes a linear mixture of concentrated and unconcentrated erosion within a cell, but this approach almost certainly underestimates the amount of concentrated erosion that takes place (Abrahams *et al.*, 1989; Nearing, 1991, 1994; Parsons and Wainwright, 2006). Because of error propagation, no process-based erosion model can be better than the hydrology and hydraulics models upon which it is based. It is clearly necessary to carry out research in developing these fields in tandem (see previous discussions in Wainwright and Parsons, 1998; Brazier *et al.*, 2000).

In terms of the structure of the erosion sub-model, there are a number of issues that need to be resolved. First, as noted above, there is a significant amount of literature that suggests that entrainment by raindrop detachment and by concentrated flows will vary dynamically throughout an event. Notwithstanding limitations regarding the detailed spatial parameterization of these dynamics, there is insufficient evidence in the literature that enables the implementation of the dynamics according to different particle-size ranges. Secondly, there are issues relating to the model structure and the interaction of different elements derived from different empirical sources. The example of model cell size discussed above demonstrates this point, and illustrates the benefits of using a modelling approach to investigate the limits of data from different sources. Although the empirical relationships that specify the feedback with flow

DOI: 10.1002/esp

depth and detachment (Torri *et al.*, 1987) and of exponential distributions of travel distance (Hubbell and Sayre, 1964; Parsons *et al.*, 1993; Van Dijk *et al.*, 2002) may be reasonable assumptions on their own, the model demonstrates that they are mutually incompatible. In both cases, alternatives have been suggested. For example, Kinnell (1993) suggests that there should be a peak in detachment at flow depths of two raindrop diameters with lower detachment rates at lesser and greater flow depths, although the data to support his case are also woefully limited in that there are no data to support the initial rise in detachment, other than fitting a point at the origin, which is an artefact of considering sediment concentration in the flow. Similarly, other distribution functions such as the gamma distribution have been suggested as representing the variability in transport distance (Grigg, 1970; Yang and Sayre, 1971; Hassan *et al.*, 1991; Wainwright and Thornes, 1991). While the exponential distribution provides a special case of the gamma distribution, it cannot describe all observed patterns of observed distribution; but it has the advantage of simplicity of parameterization. There is a need to characterize these distributions with significantly larger datasets than are currently available. There is consequently a need to characterize different parts of the erosion system in an integrated way to avoid these problems. Thirdly, the basis of the parameterization of different parts of the detachment and transport functions is very limited (e.g., the data on particle-size effects on raindrop detachment is largely confined to the work of Quansah, 1981, Savat and Poesen, 1981, and Poesen and Savat, 1981, and contains no information on dynamics such as changes of aggregate stability, whereas travel distances in shallow overland flows are only understood from the work of Parsons *et al.*, 1998, who used a range of relatively coarse particles because of experimental limitations). In part, parameterization weaknesses arise from the move away from appropriate empirical studies in recent years, but also they occur because the conceptual framework has focussed on what have been shown above to be inappropriate models of the erosion process. The negative heuristic has thus directed research away from considering the appropriate details of the process. There is thus a need to develop holistic approaches in the study of, for example, the interacting effects of particle size and the other controlling factors discussed above on rates of sediment detachment and transport. Fourthly, because of limitations of empirical observations, it has been necessary to extrapolate approaches beyond the data upon which they were originally based. For example, the Hassan *et al.* (1992) data on transport distances and virtual velocities were derived from a range of datasets in channels, whereas MAHLERAN applies these relationships to all concentrated flows. It is necessary to evaluate the extent to which such an extrapolation is reasonable, and whether the use of the ordinary regression equations by Hassan *et al.* (1992) to make them compatible with hillslope observations is appropriate. Fifthly, there is an almost complete lack of information on suspended-sediment-transport distances (but see Verhoff *et al.*, 1980). In this case, it has been necessary to extrapolate a process-based model from æolian geomorphology to provide appropriate parameterization information. Although, conceptually, there should be no problem with this extrapolation in that the physical differences in the transporting media are accounted for, in practice there are likely to be some differences, for example in the ways in which the flows interact with roughness elements. Sixthly, there are also likely to be dynamics in the particle-travel distances. For example, Ferguson *et al.* (2002) have demonstrated an apparent deceleration of tagged particle movement through time, which may be due to experimental artefacts, but may also relate to the structure of bed material in channels. At present, there is no way to represent such structural changes within MAHLERAN, and they are not included within the parameterization from either Hassan *et al.* (1992) or Parsons *et al.* (1998 – see Wainwright and Thornes, 1991, for a similar observation to that of Ferguson *et al.* for particles moving on hillslopes). There are also missing dynamics in that transport-distance and virtual-velocity parameterizations used are for a single set of conditions, whereas in reality they are likely to be more complex functions of factors such as surface roughness. Seventhly, the model contains no feedback at present between erosion and surface topography for reasons of computing efficiency and numerical stability. While this approximation may be reasonable at the event scale, especially when considering dominantly unconcentrated flows, it becomes increasingly untenable in conditions that are more rapidly changing in concentrated conditions. Eighthly, it should be recognized that all erosion models will exhibit characteristics of strong spatial autocorrelation. While this property is particularly clear in the transport-distance formulation used here, where deposition can be explicitly written in terms of erosion from nearby locations (see also Parsons *et al.*, 2004; Kirkby, 1991; 1992), it is also the case where the concentration-based formulation following Bennett (1974) is used. This autocorrelation leads to the propagation of large errors because of the net erosion term in the model, which is ill conditioned, as the erosion and deposition terms are almost identical but of opposite signs (Engeln-Müllges and Uhlig, 1996). These conditions will particularly hold near the divide where transport distances are typically very short, and models will become less sensitive further from the divide and in channels.

In addition to limitations that are specific to this model, it must be recognized that overall all models will be constrained by the empirical base upon which their parameterization rests. In particular, there will always be problems related to the sensitivity of the various sediment sub-equations, limited ranges of the dependent variables and boundary conditions. Their performance outside these conditions is uncertain. There are thus fundamental empirical issues that relate to the basis of models and their representation of soil, vegetation, microtopography and other controlling variables.

DOI: 10.1002/esp

Data limitations

In some cases, data limitations have already been discussed above in the sense that the lack of empirical underpinnings of different model equations undermines their broader use, or indeed application to different sets of dynamics. However, there is also a significant dearth of information that would permit the detailed testing of any erosion model (cf. Parsons and Wainwright, 2000). As noted above, it is fundamentally important that all aspects of the system are well constrained, so that issues relating to uncertainty and equifinality can be addressed. In this way, the rainfall-simulation plot data used are the best available in that they provide hydrology and erosion data both spatially and temporally, with independent estimates of parameters. There seem to be few – if any – other data sets available that provide such detailed data, but even so there are significant limitations, relating for example to the uncertainty of observations. It is only when such uncertainties derived from standard field techniques are propagated through the modelling process that it is possible to see the limitations of existing field approaches. Modelling thus provides a key feedback to the understanding of how to investigate geomorphological questions, and provide appropriate datasets. There are key issues raised for example by the monitored data, where there are significant difficulties in obtaining spatial patterns of runoff and erosion data, partly because of technological limitations, but also because of the difficulty of obtaining information without disturbing sites to an extent that they would no longer provide meaningful information.

Any use of monitored data leads to the potential for bias, especially as datasets with the sufficient level of required detail are likely to be very small. For example, the lack of a complete range of events on all sizes of monitored plots in the analysis here leads to potential issues, in that it is often easier to simulate the hydrology of larger and simpler events. It is difficult to evaluate, given the lack of spatial data, whether the lack of success in reproducing the sediment fluxes from the monitored plots is due to problems with parameterizing the erosion sub-component of the model, with parameterizing the hydrology and hydraulics of the slope, with the ability of the infiltration model to represent temporally variable rainfall, or with the erosion model structure itself or the parameterization of the processes in the erosion model. Further work – particularly that which considers spatio-temporal observations of hydraulic and erosion dynamics with an iterative model-development exercise – is thus required to address these different issues.

Limitations of Model Evaluation

Following the work of Oreskes *et al.* (1994), there has been a significant amount of debate about the rôle of model evaluation in demonstrating the applicability or credibility of models in the environmental sciences (see also discussions by Anderson and Bates, 2001; Mulligan and Wainwright, 2003; Demeritt and Wainwright, 2005). Had we conveniently stopped the process of model evaluation at the end of Wainwright *et al.* (in press b), then it would have been perfectly acceptable within the standard peer-review and model-evaluation processes to make broader claims about the general applicability of the MAHLERAN modelling approach. However, we have decided to produce as severe a test as is possible with as high quality spatial and temporal data as is currently available. While this approach brings potential dangers in undermining the credibility of the overall modelling approach, we believe that it is important to present the different model evaluations together, in that they tell us more than the sum of their parts. The different aspects of the evaluation allow us to probe into more detail about the structure, behaviour and parameterization of the model. Furthermore, the problems highlighted in the current paper in no way detract from the critique of existing approaches presented by Wainwright *et al.* (in press a). These problems still exist whatever the result of the detailed evaluation of MAHLERAN.

Despite the fact that Richards (1990) warned the geomorphological community of over-simplistic interpretations of model evaluations using empirical data, the problem of affirming the consequent is still all too common. In the same way that correlation does not demonstrate causation, the lack of fit between a model and a particular set of empirical data does not *necessarily* invalidate the whole model. Beyond the case where the data themselves are at error, it may be that what are essentially subsidiary hypotheses or parameterizations are essentially the parts of the modelling procedure that fail. In the presentation of MAHLERAN, we have been careful to separate what are considered to be the model structure, the ways in which model components are parameterized and the ways in which parameters are generated for field conditions so that the model may be applied in a particular setting. While we have tried to minimize the impact of the latter, the discussion above demonstrates that no simple approach is entirely infallible, and that issues relating to this parameterization do cause problems with the model evaluation. In generating a process-based model (rather than a physically based model: Mulligan and Wainwright, 2003), the model of the physics of the processes needs to be parameterized with observations of how the different model components operate and interact in reality. As there are no physically based derivations of detachment, entrainment and travel distance/deposition that have been demonstrated for the erosion system, and certainly none that are consistent for dynamically changing

DOI: 10.1002/esp

conditions in any but the most simplistic cases, this empirical part of the modelling process is necessary. However, we believe that we have been totally honest about the potential limitations of this approach and its impacts on the modelling process and consequences for model testing. MAHLERAN seems to produce credible patterns and dynamics of erosion for conditions that are dominated by large, relatively slowly changing or steady state events. These are exactly the sorts of condition for which the model parameterizations were developed, due to requirements of simplicity or lack or appropriate resources or technology. The model seems to work less well in dynamically changing conditions, for which our current understanding is far more limited. Thus, the detailed model evaluation supports the need for the collection of more data to constrain uncertainties (Beven, 2006; Hall *et al.*, 2007; Todini and Mantovan, in press; Silberstein, 2006). Furthermore, the assessment of discretization effects in the model shows that some of our existing empirical understanding or simplifications of these systems are mutually incompatible, especially when we construct difficult tests for models across spatial scales of application. Until there is a holistic understanding of these different components, the problems outlined here imply that it is not possible to evaluate MAHLERAN (or indeed any other erosion model) fully.

Conclusions

In this paper we have tested MAHLERAN against field data obtained from plots of different sizes. Although the results are less satisfactory in absolute terms than our previous test (see Wainwright *et al.*, in press b), at least they generally reproduce the relative order of measured fluxes. One reason for this difference may be the unfortunate difference in model cell size between the two sets of test data arising from limitations in the original datasets. As a consequence, it has been shown that the transport-distance model as currently formulated is still not able to account for differences in spatial scale in erosion measurements as originally suggested by Parsons *et al.* (2004), despite the fact that Parsons *et al.* (2006) have shown that it can explain observed patterns of erosion at different scales in the field. It is demonstrated that the problem is more likely to relate to limitations in empirical data used to parameterize different model components than in the model itself.

A modelling exercise such as this one, combining a conceptual basis with empirical data collection, is invaluable for demonstrating the limitations of existing understandings of hydrologic and geomorphic systems. The analyses demonstrate the inadequacy of current empirical support for erosion models, especially the limited information regarding the structure and parameterization of detachment and transport functions, and for the process-based evaluation of simulation results. Consequently, they provide the agenda for future research. It is important to remember that this significant amount of further empirical work must be carried out in parallel to the development of the numerical model(s) in order to produce an erosion model that is founded upon a strong empirical base with a non-reductionist basis. Only once this is accomplished can the model be applied with confidence to identify erosion patterns in a broad range of environments and at different spatial and temporal resolutions.

Acknowledgements

The field data were collected at Walnut Gulch Experimental Watershed over a period from 1988 to 2002. We are grateful to the respective research leaders at Southwest Watershed Research Center, Tucson, for their support, and the field staff at Walnut Gulch Experimental Watershed for their unstinting efforts. We would like to thank two anonymous reviewers for their suggestions, especially with regard to splitting the original paper into three parts. This work was funded by NERC grant GR3/12754 and NSF grant DEB 00-80412.

References

Abrahams AD, Parsons AJ, Luk S-H. 1989. Distribution of depth of overland-flow on desert hillslopes and its implications for modeling soil-erosion. *Journal of Hydrology* **106**: 177–184.

Anderson MG, Bates PD (eds). 2001. *Model Validation: Perspectives in Hydrological Science*. Wiley: Chichester.

Bennett JP. 1974. Concepts of mathematical modelling of sediment yield. *Water Resources Research* **10**: 485–492.

Beven K. 2006. On undermining the science? *Hydrological Processes* **20**: 3141–3146.

Brazier RE, Beven KJ, Freer J, Rowan JS. 2000. Equifinality and uncertainty in physically based soil-erosion models: application of the GLUE methodology to WEPP – the Water Erosion Prediction Project for sites in the UK and US. *Earth Surface Processes and Landforms* **25**: 825–845.

Brazier RE, Parsons AJ, Wainwright J, Powell DM, Schlesinger WH. 2007. Upscaling understanding of nutrient dynamics associated with overland flow in a semi-arid environment. *Biogeochemistry*. DOI: 10.1007/s10533-007-9070-x

Breckenfield DJ, Svetlik WA, McGuire CE. 1995. *Soil Survey of Walnut Gulch Experimental Watershed*. United States Department of Agriculture, Soil Conservation Service.

Clapp RB, Hornberger GM. 1978. Empirical equations for some soil-hydraulic properties. *Water Resources Research* **14**: 601–604.

Demeritt D, Wainwright J. 2005. Models, modelling and geography. In *Questioning Geography*, Castree N, Rogers A, Sherman D (eds). Blackwell: Oxford; 206–225.

Engeln-Müllges G, Uhlig F. 1996. *Numerical Algorithms with FORTRAN*. Springer: Berlin.

Ferguson RI, Bloomer DJ, Hoey TB, Werritty A. 2002. Mobility of river tracer pebbles over different timescales. *Water Resources Research* **38**(5): 1045. DOI: 10.1029/2001WR000254

Giménez R, Govers G. 2001. Interaction between bed roughness and flow hydraulics in eroding rills. *Water Resources Research* **37**: 791–799.

Grigg NS. 1970. Motion of single particles in alluvial channels. *Journal of the Hydraulics Division American Society of Civil Engineers* **96**(HY12): 2501–2518.

Hall J, O'Connell E, Ewen J. 2007. On not undermining the science: coherence, validation and expertise. *Hydrological Processes* **21**: 985–988.

Hassan MA, Church M, Ashworth PJ. 1992. Virtual rate and mean distance of travel of individual clasts in gravel-bed channels. *Earth Surface Processes and Landforms* **17**: 617–627.

Hassan MA, Church M, Schick AP. 1991. Distance of movement of coarse particles in gravel bed streams. *Water Resources Research* **27**: 503–511.

Hubbell DW, Sayre WW. 1964. Sand transport studies with radioactive tracers. *Journal of the Hydraulics Division, ASCE* **90**: 39–68.

Kinnell PIA. 1993. Sediment concentrations resulting from flow depth/drop size interactions in shallow overland flow. *Transactions of the American Society of Agricultural Engineers* **36**: 1099–1013.

Kirkby MJ. 1991. Sediment travel distance as an experimental and model variable in particulate movement. In *Erosion, Transport and Deposition Processes*, Bork H-R, de Ploey J, Schick AP (eds). *Catena Supplement* **19**: 111–128.

Kirkby MJ. 1992. An erosion-limited hillslope evolution model. In *Functional Geomorphology: Landform Analysis and Models*, Schmidt K-H, de Ploey J (eds). *Catena Supplement* **23**: 157–187.

Luk S-H, Cai QG. 1990. Laboratory experiments on crust development and rainsplash erosion of loess soils, China. *Catena* **17**: 261–276.

Moussa R, Bocquillon C. 1996. Criteria for the choice of flood-routing methods in natural channels. *Journal of Hydrology* **186**: 1–30.

Müller EN, Wainwright J, Parsons AJ. In press. Spatial variability of soil and nutrient parameters within grasslands and shrublands of a semi-arid environment, Jornada Basin, New Mexico. *Ecohydrology*.

Mulligan M, Wainwright J. 2003. Modelling and model building. In *Environmental Modelling: Finding Simplicity in Complexity*, Wainwright J, Mulligan M (eds). Wiley: Chichester; 1–73.

Nearing MA. 1991. A probabilistic model of soil detachment by shallow turbulent flow. *Transactions of the American Society of Agricultural Engineers* **34**: 81–85.

Nearing MA. 1994. Detachment of soil by flowing water under turbulent and laminar conditions. *Soil Science Society of America Journal* **58**: 1612–1614.

Nearing MA. 2003. Soil erosion and conversation. In *Environmental Modelling: Finding Simplixity in Complexity*, Wainwright J, Mulligan M (eds). John Wiley and Sons: Chichester; 277–290.

Oreskes N, Shrader-Frechette K, Bellitz K. 1994. Verification, validation and confirmation of numerical models in the Earth Sciences. *Science* **263**: 641–646.

Parsons AJ, Brazier RE, Wainwright J, Powell DM. 2006. Scale relationships in hillslope runoff and erosion. *Earth Surface Processes and Landforms* **31**: 1384–1393.

Parsons AJ, Stromberg SGL, Greener M. 1998. Sediment-transport competence of rain-impacted interrill overland flow. *Earth Surface Processes and Landforms* **23**: 365–375.

Parsons AJ, Wainwright J. 2000. A process-based evaluation of a process-based soil-erosion model. In *Soil Erosion – Application of Physically Based Models*, Schmidt J (ed.). Springer: Berlin; 181–198.

Parsons AJ, Wainwright J. 2006. Depth distribution of interill overland flow and the formation of rills. *Hydrological Processes* **20**: 1511–1523.

Parsons AJ, Wainwright J, Abrahams AD. 1993. Tracing sediment movement on semi-arid grassland using magnetic susceptibility. *Earth Surface Processes and Landforms* **18**: 721–732.

Parsons AJ, Wainwright J, Powell DM, Kaduk J, Brazier RE. 2004. A conceptual model for understanding and predicting erosion by water. *Earth Surface Processes and Landforms* **29**: 1293–1302.

Parsons AJ, Wainwright J, Schlesinger WH, Abrahams AD. 2003. Sediment and nutrient transport by overland flow in mesquite nabkha, southern New Mexico. *Journal of Arid Environments* **53**: 61–71.

Poesen J, Savat J. 1981. Detachment and transportation of sediments by raindrop splash. Part II: Detachability and transport. *Catena* **8**: 19–41.

Ponce VM. 1991. The kinematic wave controversy. *Journal of Hydraulic Engineering American Society of Civil Engineers* **117**: 511–525.

Quansah C. 1981. The effect of soil type, slope, rain intensity and their interactions on splash detachment and transport. *Journal of Soil Science* **32**: 215–224.

Rice CT, Wilson BN. 1990. *Analysis of Dynamic Variables of Rill Flow*, ASAE Paper 902011 (Presentation at the 1990 International Summer Meeting, Columbus, Ohio, June 24–27, 1990). ASAE: St Joseph, MI.

Richards KS. 1990. 'Real' geomorphology. *Earth Surface Processes and Landforms* **15**: 195–197.

Savat J, Poesen J. 1981. Detachment and transportation of loose sediments by raindrop splash. Part I: The calculation of absolute data on detachability and transportability. *Catena* **8**: 1–17.

Silberstein RP. 2006. Hydrological models are so good, do we still need data? *Environmental Modelling and Software* **21**: 1340–1352.
Simanton JR, Renard KG. 1992. Upland erosion research on rangeland. In *Overland Flow: Hydraulics and Erosion Mechanics*, Parsons AJ, Abrahams AD (eds). University College London Press: London; 335–376.
Singh VP. 1996. *Kinematic Wave Modelling in Water Resources*. Wiley: Chichester.
Smith RE, Parlange J-Y. 1978. A parameter-efficient hydrologic infiltration model. *Water Resources Research* **14**: 533–538.
Todini E, Mantovan P. 2007. Comment on: 'On undermining the science?' by Keith Beven. *Hydrological Processes* **21**: 1633–1638.
Torri D, Sfalanga M, del Sette M. 1987. Splash detachment: runoff depth and soil cohesion. *Catena* **14**: 149–155.
Van Dijk AIJM, Meesters AGCA, Bruijnzeel LA. 2002. Exponential distribution theory and the interpretation of splash detachment and transport experiments. *Soil Science Society of America Journal* **66**: 1466–1474.
Verhoff FH, Yaksich SM, Melfi DA. 1980. River nutrient and chemical transport estimation. *Journal of the Environmental Engineering Division ASCE* **106**: 591–608.
Wainwright J, Parsons AJ. 1998. Sensitivity of sediment-transport equations to errors in hydraulic models of overland flow. In *Modelling Soil Erosion by Water*, Boardman J, Favis-Mortlock D (eds). Springer: Berlin; 271–284.
Wainwright J, Parsons AJ. 2002. The effect of temporal variations in rainfall on scale dependency in runoff coefficients. *Water Resources Research* **38**(12): 1271. DOI: 10.1029/2000WR000188
Wainwright J, Parsons AJ, Powell DM, Brazier R. 2001. A new conceptual framework for understanding and predicting erosion by water from hillslopes and catchments. In *Soil Erosion Research for the 21st Century. Proceedings of the International Symposium*, Ascough II JC, Flanagan DC (eds). American Society of Agricultural Engineers: St. Joseph MI; 607–610.
Wainwright J, Parsons AJ, Abrahams AD. 2000. Plot-scale studies of vegetation, overland flow and erosion interactions: case studies from Arizona and New Mexico. *Hydrological Processes* **14**: 2921–2943.
Wainwright J, Parsons AJ, Müller EṆ, Brazier RE, Powell DM, Fenti B. In press a. A transport-distance approach to scaling erosion rates: 1. Background and model development. *Earth Surface Processes and Landforms*. DOI: 10.1002/esp1624
Wainwright J, Parsons AJ, Müller EṆ, Brazier RE, Powell DM, Fenti B. In press b. A transport-distance approach to scaling erosion rates: 2. Sensitivity and evaluation of MAHLERAN. *Earth Surface Processes and Landforms*. DOI: 10.1002/esp1623
Wainwright J, Thornes JB. 1991. Computer and hardware simulations of archæological sediment transport. In *Computer Applications and Quantitative Methods in Archæology 1990*, BAR International Series 565, Lockyear K, Rahtz S (eds). Oxford; 183–194.
Yang CT, Sayre WW. 1971. Stochastic model for sand dispersion. *Proceedings of the American Society of Civil Engineers, Journal of the Hydraulics Division* **97**(HY2): 265–288.