2006

Surface process models and the links between tectonics and topography

Alexandru T. Codilean  
*University of Glasgow, codilean@uow.edu.au*

Paul Bishop  
*University of Glasgow*

Trevor B. Hoey  
*University of Glasgow*

Publication Details  
Surface process models and the links between tectonics and topography

Abstract
Advances in the theoretical understanding of large-scale tectonic and surface processes, along with a rapid growth of computing technology, have stimulated interest in the use of numerical surface process models (SPMs) of long-term landscape evolution, especially in relation to the links between tectonics and topography. Because of these advances and possibilities and because SPMs continue to play an important part in recent geological, geomorphological, thermochronological and other geosciences research, the models warrant review and assessment. This review summarizes and evaluates the important issues concerning SPMs of long-term landscape evolution that have been addressed only in a passing way by previous authors. The issues reviewed here are: (1) the formulation of the 'laws' that represent fluvial and hillslope processes in SPMs; (2) the implementation of the various algorithms on numerical grids; (3) model parameterization and calibration; and (4) model testing.

Keywords
Long-term landscape evolution, model parameterization and calibration, model testing, numerical modelling, process specification, surface process models, tectonics and topography, GeoQuest

Disciplines
Medicine and Health Sciences | Social and Behavioral Sciences

Publication Details

This journal article is available at Research Online: http://ro.uow.edu.au/smhpapers/1514
Surface Process Models and the links between tectonics and topography

Alexandru T. Codilean*, Paul Bishop and Trevor B. Hoey

Department of Geographical and Earth Sciences,
University of Glasgow,
Glasgow, G12 8QQ, United Kingdom,
Telephone: +44 (0) 141 330 4782, Fax: +44 (0) 141 330 4894

* Corresponding author.
E-mail address: tcodilean@ges.gla.ac.uk
Abstract

Advances in the theoretical understanding of large-scale tectonic and surface processes, along with a rapid growth of computing technology, have stimulated interest in the use of numerical surface process models (SPMs) of long-term landscape evolution, especially in relation to the links between tectonics and topography. Because of these advances and possibilities and because SPMs continue to play an important part in recent geological, geomorphological, thermochronological and other geosciences research, the models warrant review and assessment. This review summarises and evaluates the important issues concerning SPMs of long-term landscape evolution that have been addressed only in a passing way by previous authors. The issues reviewed here are: (1) the formulation of the ‘laws’ that represent fluvial and hillslope processes in SPMs; (2) the implementation of the various algorithms on numerical grids; (3) model parameterisation and calibration; and (4) model testing.

Key words: long-term landscape evolution; model parameterisation and calibration; model testing; numerical modelling; process specification; surface process models; tectonics and topography.
I   Introduction

The re-emergence over the last decade or so of research into the links between large-scale tectonic processes\(^1\) and long-term landscape evolution reflects several factors. These include early elaborations of the plate tectonic paradigm so as to generate new questions about the links between tectonic processes and Earth surface topographic evolution (e.g., Holmes 1965; Dewey and Bird, 1970), and the early exploration of these links in the context of the evolution of passive continental margins (e.g., Ollier, 1982, 1985; Bishop, 1986; Lister et al., 1986; Summerfield, 1991). Perhaps equally important was the rapid growth in computing power, speed and memory, thereby enabling the large number of computations required to simulate landscape evolution over the temporal and spatial scales appropriate to exploring the links between tectonics and topography in a plate tectonic framework.

The ‘Tectonics and Topography’ Special Issues of the Journal of Geophysical Research (cf. Merritts and Ellis, 1994) marked a definitive coming-together of these developments, not least in the numerical models (surface process models – SPMs) that were published in these

\(^1\) Our use of the term ‘large-scale’ to mean ‘generalised and covering a large area’ (e.g., a continental margin or a whole orogen) is not strictly correct. In mapping, ‘large-scale’ means detailed (and therefore generally covering a small area). The usage is retained here, however, because it has become common in the literature and in common parlance. The latter means that usage seems to be changing such that it now seems counter-intuitive to think of the numerical models of long-term landscape evolution considered here as ‘small-scale’.  

3
Special Issues, exploring the links between tectonics and long-term landscape evolution. Numerical modelling papers in the Special Issues include models of the evolution of high elevation passive continental margins (Gilchrist et al., 1994; Kooi and Beaumont, 1994; Tucker and Slingerland, 1994) and of other settings, including plateau margins (Masek et al., 1994), of planated plateau surfaces (Gregory and Chase, 1994) and of actively uplifting areas (Anderson, 1994; Rosenbloom and Anderson, 1994). Many of these models explicitly acknowledge a debt to the early numerical models of Koons (1989) and Beaumont et al. (1992) for collisional orogens.

These trends continue with a growing number of SPMs of large-scale, long-term landscape evolution, and the maturing of the plate tectonic paradigm (Beaumont et al., 2000). Ever-increasing computing power and memory mean that sophisticated coupling of climatic and tectonic processes can be attempted in order to test for feedbacks between climate and topography of the types envisaged, for example, by Molnar and England (1990) and Raymo and Ruddiman (1992). It should even be possible now to explore the pattern of long-term evolution of landscapes as climates change during continental drift, in the way envisaged by Bowler (1982), for example, for Australia’s northward drift during the Cenozoic. Likewise, the feedbacks between topography, denudation and climate in an uplifting convergent zone mountain block can be explored numerically in terms of the differences in the patterns of crustal uplift, denudation and sediment flux that result from differences in prevailing wind direction (Beaumont et al., 1992, 2000).

These advances and possibilities prompt questions as to the validity and usefulness of these models. Acceptance of the value of these numerical models is by no means
universal (e.g., Ollier and Pain, 2000), and there is debate within the modelling community as to the appropriate way(s) to formulate, for example, the surface process algorithms or ‘laws’ that lie at the heart of the models (see below). For these reasons, and because the models continue to play an important part in recent geological, geomorphological, thermochronological and other geosciences research (e.g., Kooi and Beaumont, 1996; Tucker and Slingerland, 1996; van der Beek et al., 1995; Whipple and Tucker, 1999; Beaumont et al., 2000; Tucker et al., 2001a; van der Beek et al., 2002; Braun, 2003; Snyder et al., 2003; Gasparini et al., 2004; Tucker, 2004; Braun and van der Beek, in press), the SPMs warrant review and assessment.

Several earlier reviews of SPMs and landscape evolution (e.g., Beaumont et al., 2000; Burbank and Anderson, 2001; Tucker et al., 2001b) either were focused on specific types of models (e.g., Beaumont et al.’s (2000) review of coupled tectonic – surface process models with applications to rifted margins and collisional orogens) or were limited to a brief summary of the different ways numerical modelling of landscape evolution has been attempted (e.g. Tucker et al., 2001b). Coulthard’s (2001) brief comparisons of some of the principal SPMs is mainly a ‘software review’ focused on ‘practical’ aspects, such as, the type of operating system supported by the various models and the input and output file formats.

More recent reviews of numerical modelling of landscape evolution have been provided by Martin and Church (2004) and Willgoose (submitted), the latter focusing on debates related to the processes represented in SPMs. Willgoose (submitted) also provides a very useful ‘feature’ comparison of eight different SPMs, summarising properties, such as the type of transport and erosion mechanism, the number of grading fractions used for
tracking transported/deposited sediment, types of tectonics models and mode of representation of topography. Martin and Church (2004) have examined the theoretical and methodological issues that are involved in numerical and other modelling of landscape evolution, concentrating on the quantitative study of geomorphological processes. Their emphasis on the importance of scale (both spatial and temporal) for the appropriate specification of the system being studied and the relevant processes is noteworthy. The ‘art’ of any type of modelling lies in the capability of the modeller to differentiate between the different processes operating within a system, isolating the relevant processes and ignoring others, so that the model itself is relatively easy to understand, implement and interpret, but still remains a valid representation of reality. Martin and Church (2004) have highlighted how understanding the scaling properties of landscapes can help in the selection of appropriate domains for the study of landscape change and therefore in the proper identification of the relevant processes. Their overview is less detailed on other important issues in the numerical modelling of landscape evolution. Their treatment of the formulation of fluvial process algorithms is brief, despite the fact that landscape evolution models are practically driven by these, hillslope processes having only a secondary role (see below). Other important issues, such as the numerical implementation of the various algorithms or the calibration of parameters are inappropriate for Martin and Church’s (2004) review. Likewise, the need for testing model outcomes, an issue that has received much attention recently (e.g., Hancock, 2003; Hoey et al., 2003; Willgoose et al., 2003) is only briefly addressed by Martin and Church (2004).

The lack, to date, of a thorough critical review of SPMs and landscape evolution is understandable given the wide-ranging nature of the issue and the multitude of directions
such a review could take. This present review aims to fill this gap and to address in a critical manner important issues concerning SPMs and long-term landscape evolution that have been only briefly addressed by previous authors. For the sake of this review, we accept that it is valid to attempt numerical modelling of long-term landscape evolution, if only to use the models heuristically, as a means to explore ‘what-if’ questions and the sensitivity of landscape evolution to various controlling factors (cf. Oreskes et al., 1994). Smaller-scale numerical models focusing on individual catchment evolution over hundreds to thousands of years (e.g., Willgoose et al., 1991a, 1991b; Coulthard et al., 1997; Tucker and Slingerland, 1997; Willgoose and Riley, 1998; Tucker et al., 2001a; Hancock and Willgoose, 2004) are not our primary focus here. These smaller-scale models have many elements in common with the larger-scale models but there are important differences (although some of the smaller scale models (e.g., that of Tucker et al., 2001a) can mimic the behaviour of their larger-scale counterparts if configured appropriately. Process formulation and morphological response may be considerably more detailed in the smaller-scale models, even down to, for example, simulation of detailed sedimentology and depositional stratigraphy (e.g., Coulthard et al., 1997; Tucker et al., 1999, 2001b). Scale differences between the catchment and the mountain range or continental margin also mean that individual hillslopes are not represented in the large-scale SPMs (see below), and that these SPMs treat rainfall only very simply (Tucker et al., 2001a; Tucker, 2004). The larger-scale SPMs, therefore, are associated with a suite of distinctive methodological issues. We begin with a summary of the structure and operation of numerical models of long-term landscape evolution, and then move to review model testing.
II Approaches to the numerical modelling of long-term landscape evolution

SPMs are used to test generic conceptual models of landscape evolution (e.g., Kooi and Beaumont’s (1996) modelling of the classic models of landscape evolution) and/or to model the landform evolution and denudation history of specific regions (e.g., van der Beek and Braun’s (1998) modelling of long-term landscape evolution of the SE Australian passive continental margin). Either way, full numerical modelling of long-term landscape evolution requires: 1. the numerical representation (algorithms) of the process ‘laws’ that determine the rates and pathways of landscape evolution, including both surface processes and tectonic processes; and 2. the implementation of these algorithms in a way that allows the surface and tectonic processes to interact in an integrated and meaningful way. We are more concerned here with surface process models but note in passing that attempts to couple tectonic models and SPMs must take account of the important methodological issue of reconciling the way in which mass fluxes are handled by the two types of models. As Beaumont et al. (2000) have pointed out, SPMs fundamentally operate in planform and it is not always possible to couple this planform operation to two-dimensional (vertical section) tectonic (geodynamic) models of the type, for example, developed by Beaumont et al. (1992, 1996). Other issues include reconciling the different temporal and spatial scales over which surface and tectonic processes operate (Beaumont et al. 2000). The magnitude–frequency characteristics of surface process events, for example, are not always clear (e.g., Wolman and Miller, 1960) and deciding the appropriate temporal resolution (time step) for an SPM is not a trivial issue. Moreover, whatever temporal resolution is chosen for the SPM, it will likely be different from the magnitude–frequency characteristics of the
tectonic model (Beaumont et al. 2000). The only exception to the latter point may be situations in which extreme rock uplift rates are matched by extreme rock incision rates reflecting relatively unusual combinations of very steep channel gradients, extreme discharges (e.g., driven by typhoons) and seismic shaking. Such relatively rare combinations of factors are found in Taiwan, for example (e.g., Hartshorn et al., 2002; Dadson et al., 2003).

1 Landscape evolution processes

Landscape evolution in all SPMs is governed by at least two groups of processes: slope processes (short-range diffusive hillslope transport) and fluvial processes (or long-range advective fluvial transport), both operating in a drainage network through which fluvial discharge is collected and routed (Figure 1). Adequate formulation and coupling of the algorithms to represent these surface processes are not trivial issues. Howard et al. (1994) have commented on inadequate algorithm formulation, especially in terms of the adequacy of representation of the range of fluvial processes and channel types, as well as the great difficulties of representing the broad range of slope processes by one algorithm (see Martin’s (2000) and Martin and Church’s (2004) discussion of the latter issue). It would not be unfair to say, however, that these concerns have not been widely acted on, and the formulation of the algorithms to represent surface processes in SPMs has generally followed the spirit of Kooi and Beaumont (1994):

“As demonstrated in other sciences, a fruitful approach to such problems of [the range of] scale [that must be captured in SPMs] is to set aside (for the time being) the small-scale, short-timescale picture and to explore
simple relationships cast in terms of large-scale, long-timescale quantities”
(Kooi and Beaumont 1994, p.12,192).

The approach, which does not completely satisfy Howard et al.’s (1994) call for “calibratable, mechanistic, transport/erosion laws” (p. 13,971), has been to formulate algorithms that are thought to represent adequately or ‘capture’ the key processes involved in landscape evolution and not necessarily to attempt to formulate completely physically-based algorithms that describe the full physics/mechanics of the particular process. In this spirit, and following the work of Howard and Kerby (1983), fluvial processes of bedrock incision have been widely formulated for numerical modelling purposes as various functions of stream power (the product of channel slope, $S$, and catchment area, $A$, a surrogate for channel discharge). Slope processes have been conceptualised as diffusive processes dependent on slope or topographic curvature.

Figure 1 around here

van der Beek and Bishop (2003) have recently reviewed the various forms that the fluvial incision law may take. In the simplest formulation, fluvial incision is a function of stream power, unit stream power, or basal shear stress; Whipple and Tucker (1999) and Snyder et al. (2000) provide full derivations of all three formulations and show that they reduce to the general form:

$$I = K A^n S^m$$  \hspace{1cm} (1)
where $I$ is fluvial incision, $K$ is a dimensional coefficient of erosion, $A$ is catchment area (a surrogate for channel discharge), $S$ is channel gradient, and $m$ and $n$ are constants. The value of the exponent of slope, $n$, depends on the dominant erosion processes and has been argued to vary between $\sim 0.67$ and $\sim 1.67$ (Whipple et al., 2000). This bedrock incision law has been widely used in numerical models of landscape development (e.g., Anderson, 1994; Rosenbloom and Anderson, 1994; Tucker and Slingerland, 1994; Whipple and Tucker, 1999; Willett, 1999). Some treatments of fluvial processes, such as that of Tucker and Slingerland (1994) and Tucker et al. (2001b), have explicitly recognised the fundamentally different fluvial processes associated with bedrock channels (detachment-limited) and alluvial channels (transport-limited) (cf. Howard et al., 1994), and incorporate algorithms for both processes. In this case, the channel is treated as bedrock when potential fluvial sediment transport exceeds the available sediment, and as alluvial when the amount of sediment available for transport exceeds the transport capacity. In this type of formulation, both alluvial and bedrock fluvial incision are expressed, nonetheless, as functions of stream power. In an attempt to represent the importance of debris flow scour in channel headwaters, Anderson (1994) formulated fluvial incision as a nonlinear function of $S$ (a function of $S^3$).

In a further development of this approach, other fluvial incision algorithms acknowledge that there is a critical stream power or shear stress, $C_0$, that must be exceeded before incision occurs, and formulate an ‘Excess Stream Power’ algorithm:

$$I = (K A^n S^n - C_0)$$  \hspace{1cm} (2)

This form has been used by Densmore et al. (1998), Tucker and Slingerland (1997), Sklar and Dietrich (1998), Lavé and Avouac (2001) and Tucker et al. (1999, 2001b).
The role of sediment transport has also been taken into account in formulating the fluvial incision algorithm by elaborating the stream power model to describe bedrock incision as inversely proportional to a characteristic erosional length scale $L_f$ (a measure of the substrate erodibility) and directly proportional to the degree of disequilibrium in the fluvial sediment flux or ‘undercapacity’:

$$I = \frac{1}{L_f} (q_{eq} - q_s)$$  \hspace{1cm} (3)

$q_{eq}$ is the equilibrium carrying capacity of the river, considered proportional to the linear stream power, and $q_s$ is the local sediment flux (see van der Beek and Bishop (2003) for more detail.) This Undercapacity model was originally formulated by Beaumont et al. (1992) and Kooi and Beaumont (1994), and has been used subsequently by Braun and Sambridge (1997), van der Beek and Braun (1998, 1999) and van der Beek et al. (1999, 2002). The value of $L_f$ varies according to whether the substrate is bedrock (detachment-limited; high $L_f$) or alluvium (transport-limited; low $L_f$) and van der Beek et al. (2002) argued that the ‘undercapacity’ formulation has the advantage of treating transport-limited and detachment-limited behaviour in a single algorithm. Whipple and Tucker (1999) have argued, however, that the physical basis of this formulation is less clear. The numerical modelling literature is certainly dominated by the shear stress formulation of the fluvial incision law, but as Whipple and Tucker (1999) have also noted, useful insights have been gained from the undercapacity formulation.

A fourth formulation of the fluvial incision algorithm, which has not been widely used in SPMs, has a more sophisticated treatment of the role of sediment (Sklar and Dietrich 1998). Sklar and Dietrich’s (1998) approach calculates (i) the fraction of channel bed
composed of exposed bedrock, which is assumed to depend on the excess transport capacity as in (3), (ii) the particle impact rate per unit area, which depends on sediment flux as well as grain size characteristics and saltation length, and (iii) the volume of material removed per particle impact, which is a function of the particle’s kinetic energy.

A simplified version of this model, in which the terms in (ii) other than sediment flux and those in (iii) are assumed constant, is equivalent to Slingerland et al.’s (1997) empirical model which predicts the same macro-scale behaviour. This ‘Tools’ model acknowledges sediment’s role, both as a ‘tool’ for incision and, at other times, as a protector of the bed, and can be formulated as follows:

\[
I = \frac{q_s}{W L_f} \left( 1 - \frac{q_s}{q_{eq}} \right)
\]  

(4)

where \( W \) represents the river width (see van der Beek and Bishop (2003) for more details).

van der Beek and Bishop (2003) have assessed the validity of the various fluvial incision algorithms by comparing actual fluvial long profiles to those generated from initial long profiles provided by the upper surfaces of Early Miocene valley-filling lava flows. They concluded that all these algorithms are capable of predicting the observed amounts of incision reasonably well. However, some of the models (the transport-limited stream power model and the nonlinear versions of the undercapacity and tools models) have parameter combinations that have no physical significance and some of the models’ best fit parameter combinations are such that they mimic the behaviour of other models (e.g., best fit excess stream power models behave as detachment-limited stream power models).
As already noted, the majority of SPMs use forms of equation (1) for their fluvial incision law, with some using equation (3).

3 Hillslope processes

The hillslope processes algorithms encapsulate diffusive, short-range transport and are therefore taken to represent the cumulative effects of soil creep, rain splash, landslides, earth flows and other hillslope processes (Carson and Kirkby, 1972; Dietrich et al., 1995). Howard et al. (1994) have argued that the complexity of these processes means that it is so far impractical to attempt to represent them in SPMs, and they advocated concentrating modelling efforts on the evolution of the fluvial system long profile, while implicitly treating the evolution of hillslopes by having them ‘follow’ long profile evolution.

By and large, however, hillslope processes are included in SPMs, and are represented as slow continuous diffusive processes related to topographic curvature (the second derivative of elevation, z):

\[
\frac{\partial z}{\partial t} = -\nabla ( - k_d \nabla z) = k_d \nabla^2 z
\]  

This formulation precludes slope processes that are dependent on distance from the drainage divide (such as slopewash) and processes that entail a threshold slope angle, such as landsliding (Rosenbloom and Anderson, 1994). Landsliding may be implicitly included in an SPM by setting a high parameterisation for a single diffusivity algorithm (Anderson, 1994), but is usually addressed by using at least two hillslope process algorithms, one to represent landsliding and the second to represent the other diffusive
processes, such as soil creep and rain splash (e.g., Braun and Sambridge, 1997; van der Beek et al., 2002). If explicitly included, landsliding is generally set to occur when a threshold slope angle is exceeded; different threshold slope angles may be used to distinguish regolith landsliding and deeper bedrock mass failure (Tucker and Slingerland, 1994). Such relatively simple formulation (i.e., in terms of exceedance of a threshold slope angle) is made more sophisticated by including a probability of failure when the threshold slope angle is exceeded, perhaps also including the time since last failure (e.g., van der Beek et al., 2002). In Tucker and Slingerland’s (1994) model, three hillslope processes are represented: shallow sediment failure (regolith landsliding); deeper rock failure (bedrock mass failure); and diffusive hillslope transport (creep etc). Tucker and Slingerland (1994) also included an algorithm for bedrock weathering (‘sediment production’) on hillslopes, and followed Ahnert’s (1976) and Anderson and Humphrey’s (1989) formulation of an exponential decline in regolith production rates with increasing regolith thickness (see also Anderson (1994) and Rosenbloom and Anderson (1994)).

Rapid hillslope processes (landsliding) are not explicitly represented in all SPMs, such as in some of the catchment-scale models (e.g., Willgoose et al., 1991a and Tucker et al., 2001a) or some models of passive continental margin evolution (e.g., Kooi and Beaumont, 1994). The latter is probably not unreasonable, given the low average slope angles that generally characterise passive margins, although landsliding must figure in the evolution of high elevation passive margin escarpments. Landsliding and/or bedrock failure is a key process in tectonically active areas (Burbank et al., 1996; Hovius, 2000) and is explicitly included in SPMs of these settings (Allen and Hovius, 1998; Densmore et al., 1998; Allen and Densmore, 2000).
Implementation

The continuous topography of the landscape forming the domain of a SPM is approximated using digital terrain models in the form of regular or irregular grids (Figure 1). The various algorithms encapsulating the process ‘laws’ are implemented on these grids, and the rate of change of elevation is calculated evaluating the sediment mass balance of every node\(^2\) by numerically solving the sediment mass continuity equation in a downstream order. Generally each node hosts the algorithms for both fluvial and slope processes. The grid is therefore not strictly a digital elevation model with only fluvial processes acting at ‘channel’ nodes and slope processes acting only at adjacent ‘hillslope’ nodes: all nodes can experience fluvial and hillslope processes. The gradient between nodes \((S\) in eq. 1) is one component in determining fluvial stream power and this is also the gradient on which the rate of diffusive (hillslope) processes depend. As noted above the latter may be formulated as a function of the rate of change of this gradient (eq. 5), but is also often numerically implemented as a function of gradient only (eq. 6a).

The models work in terms of sediment fluxes: the total sediment derived from the area surrounding a node is determined using the hillslope algorithm(s) and this, plus what is delivered to the node from upstream and whatever sediment may be stored at the node from the previous time-step, is the total sediment available at the node for fluvial

\(^2\) In the case of regular grids, “node” denotes the grid cell itself. For irregular grids, for which the data structure is more complex, “node” denotes the actual points that are connected together by the triangle edges (see Figures 1 C and D).
transport processes, however they are formulated. Untransported sediment remains ‘stored’ at the node as alluvial sediment. Therefore, for an irregular grid, the rate at which the elevation of a given node changes as a result of erosion and/or deposition caused by hillslope and fluvial processes respectively can be calculated as:

\[
\frac{dz_i}{dt}_{\text{hillslope}} = \frac{k_i}{A_i} \sum w_{ij} S_{ij} \]  \hspace{1cm} (6a)

\[
\frac{dz_i}{dt}_{\text{fluvial}} = \frac{\left( \sum Q_{s_{\text{in}}} \right) - Q_{s_{\text{out}}}}{A_i} \]  \hspace{1cm} (6b)

where \( Q_{s_{\text{in}}}, Q_{s_{\text{out}}} \) stand for the volume of sediment entering and leaving cell \( i \), \( A_i \) is the area of the cell (constant for regular grids), \( w_{ij} \) represents the width of the common face of the two adjacent cells, and \( S_{ij} \) is the downslope gradient between the two nodes calculated as: \( S_{ij} = \frac{(z_i - z_j)}{l_{ij}} \) with \( l \) being the length connecting the two points, \( z_i \) and \( z_j \) (Figure 1D). The form of equation (6b) might be different for regular (rectangular) grids if there is a distinction between diagonal and cardinal neighbours. The two equations also depend on the type of the flow routing algorithm employed (see below). The sediment fluxes are ‘handed down’ through the grid mesh and leave the model at its lowest point (or are deposited if the fluvial discharge is ponded at the bottom of the drainage net).

**Figure 2 around here**

After the sediment fluxes have been routed downstream, the total surface elevational change of the model is computed and the corresponding denudational ‘unloading’ in the erosional part of the model (generally the bedrock terrain) and depositional ‘loading’, generally in the depositional part of the terrain, are used to compute the flexural isostatic
response of the surface. This is usually modelled by the vertical deflection of a thin elastic plate overlying an ideal fluid. Ignoring horizontal forces the equation describing this takes the form:

$$D \frac{d^4 \omega}{dx^4} + (\rho_m - \rho_c)g \omega = q(x)$$  \hspace{1cm} (7)

where $D$ is the flexural rigidity of the lithosphere, $\omega$ is the vertical flexural displacement, $\rho_m$ and $\rho_c$ are the density of the mantle and the crust respectively, $g$ is gravitational acceleration and $q(x)$ represents the applied crustal load (Turcotte and Schubert, 2002).

Almost all SPMs route water and sediment downstream using the D8 method (O’Callaghan and Mark, 1984; Jenson and Domingue, 1988), or an adaptation of it (Figures 2 and 3). The main advantage of this method, which passes all the sediment originating from a given node to its lowest natural neighbour, is its simplicity and ease of implementation. However, the D8 method limits the number of possible directions of downstream routing to the number of natural neighbours (eight, in the case of regular grids), contributing to the artificial look of the ‘landscapes’ in SPMs with a regular distribution of nodes (Figures 1 A1, A2 and 4). This non-naturalness is emphasised in coarse spatial resolution grids (1km or even 5km) as well as by the simplistic nature of the D8 flow routing algorithm (Desmet and Govers, 1996; Tarboton, 1997). This artificiality is also present, but less obvious, in irregular grids.

The algorithm used to route water and sediment down-slope plays an important, but generally ignored, role in the simulated landscape evolution and the characteristics of the resulting topography. The flow routing algorithm has a direct influence on the key variable of catchment area ($A$ in equations 1 and 2) (Figure 3). Some catchment scale
models acknowledge this importance and treat flow routing in a more complex manner (e.g., Coulthard et al., 1996, 1997); some even implement several algorithms (e.g., Willgoose, 2004). It might be argued that routing flow in only one direction is reasonable when dealing with coarse resolution grids (e.g., >1km). However, the larger the area represented by the grid cell, the more likely it is for water and sediment originating from that area to drain in more than one direction due to the range of landscape features that can be encompassed by an area of the size of a grid cell.

Figure 3 around here

The use of more complex flow routing algorithms (e.g., Quinn et al., 1991; Freeman, 1991; Lea, 1992; Costa-Cabral and Burges, 1994, Tarboton, 1997) generally produces more naturally looking ‘landscapes’, but with coarse resolutions artificiality remains unavoidable. To address these shortcomings of rectangular grids and to give the modelled landscapes a more ‘realistic’ appearance, Braun and Sambridge (1997) introduced an irregular mesh (see also Tucker et al., 2001b). Tucker et al. (2001b) have also pointed out that the real advantage of using variable mesh grids is that they permit the ‘marrying’ of 2D SPMs and 3D tectonic models, as well as the development of migrating river channels in smaller-scale models of meandering systems. Perhaps most importantly, this irregular grid is self-adapting in that the resolution of the grid automatically varies anywhere on the model as a function of the model’s process rates at that locality. Thus, where rates of landscape change are high (as, say, at a headward-retreating knickpoint or on steep slopes), the grid resolution increases so as to capture increased detail of the landscape evolution at that locality. Where rates of change are
slower, the grid resolution automatically adapts to a lower value. This is computationally more efficient than running the whole model at the high resolution necessary to gain all the necessary information at the localities with high rates of change. At the same time, irregular grids have a much more complex data structure, therefore requiring more complex algorithms and more storage space. Moreover, if they are not configured properly, numerical models using self-adapting irregular grids can be computationally more intensive than their regular grid counterparts.

Figure 4 around here

In the case of a high-resolution grid, it may be that some nodes act only as hillslope nodes (namely, when the fluvial algorithm at a node simply acts as a ‘conveyor belt’, ‘handing on’ to the next node the sediment derived by the diffusive hillslope processes). In this case, the rate of landscape lowering would be determined by the rate of hillslope lowering alone (so long, of course, as there was sufficient stream power to advect away the sediment derived by these hillslope processes). However, denudational rebound means that baselevel lowering and correspondingly increasing ‘fluvial’ gradients between nodes (and therefore increased stream power) are integral parts of the larger-scale models, which is turn means that fluvial incision is an important process in these larger-scale models. Indeed, high resolution models, in which fluvial processes act solely as a conveyor belt for sediment, are such that they cannot currently incorporate denudational rebound that is computed on the basis of the model-derived sediment fluxes and re-arrangement of loads. Such models are of too high a resolution – and therefore computationally too intensive – to be able to encompass the length scales necessary to
accommodate the lithospheric flexural properties for the computation of denudational rebound from loading and unloading. Thus, fluvial incision generally ‘leads’ landscape evolution in numerical models and hillslopes ‘follow’, and once a dynamic equilibrium (steady state) landscape is established, hillslope morphology and process rates remain unaltered unless this equilibrium is perturbed (cf. Anderson, 1994) (see below). In this situation, some models are implemented by not modelling hillslope processes explicitly and by simply setting hillslopes to move up or down at the same rate as the changes in channel elevation, which are explicitly modelled (e.g., Tucker and Slingerland, 1996). This approach follows the procedures advocated by Howard et al. (1994).

5 Parameterisation and Calibration

The values of the governing equations’ parameters (e.g., the values of $K$, $m$ and $n$ in eq 1 and $k_d$ in eq 5) must be set before an SPM can be run. The values that $m$ and $n$ take depend on the form of the process law: for incision linearly proportional to bed shear stress, $m = 0.33$ and $n \approx 0.67$ (Howard and Kerby, 1983; Howard et al., 1994; Whipple et al., 2000), whereas with $m = n = 1$ equation (1) represents stream power per unit length (Seidl and Dietrich, 1992; Seidl et al., 1994; Howard et al., 1994) and with $m \approx 0.5$ and $n \approx 1$, stream power per unit bed area (Whipple and Tucker, 1999). The appropriate values of $m$ and $n$ are set either on the basis of an a priori judgement as to the appropriate form of the fluvial incision algorithm (e.g., $m = 1$; $n = 1$; Tucker and Slingerland, 1994), or by analysis of best fits for channel slope – drainage area relationships for that area (e.g., the Zagros: Tucker and Slingerland, 1996; Siwalik Hills: Kirby and Whipple, 2001; upper Lachlan valley, SE Australia: Stock and Montgomery, 1999; van der Beek and Bishop,
2003), by exploiting ‘Playfair’s Law’ for tributary-trunk confluences (Seidl and Dietrich, 1992) or by using knickpoint celerity (Whipple and Tucker, 1999; Bishop et al., in press). Whipple and Tucker (1999) have shown that the slope exponent, $n$, has a significant influence on the magnitude and timescale of fluvial response to imposed tectonic or climatic forcing, with response times being greater for lower values of $n$ (Whipple 2001). Tucker and Whipple (2002) have noted that nonlinearity in the slope exponent has a fundamental impact on the shape of channel profiles during transient state. Streams respond to baselevel fall through upstream migrating knickpoints, but for $n > 1$ knickpoints are rapidly dissipated due to the faster upstream propagation of their steeper portions. In this situation, profiles evolve towards a characteristic declining form (Tucker and Whipple, 2002; Baldwin et al., 2003). The $m/n$ ratio also provides a measure of the overall convexity of the channel long profile (Whipple and Tucker, 1999).

The value of $K$, the dimensional erosional coefficient in eq 1, is also commonly determined empirically for given $m$ and $n$ values, and may vary by many orders of magnitude depending on lithology and climate (e.g., $K = 10^{-7}$-10^{-3}\text{ m}^{0.2}/\text{yr}$; Stock and Montgomery, 1999) and uplift rate ($K = 1-8 \times 10^{-5}\text{ m}^{0.7}/\text{yr}$; Snyder et al., 2000). Whipple et al. (2000) used river incision into an historic ash flow in Alaska to obtain values for $K$ (and $n$). Tucker and Slingerland (1996) determined the value of $K$ by comparing observed and predicted topographic relief. In this procedure, the numerical model is run several times with only the value of $K$ being varied until the predicted relief approximates the one calculated from a DEM of the area.

There is a several order of magnitude range in the values of the diffusivity coefficient used in the SPMs ($k_d = 10^{-3}$-10^3\text{ m}^2/\text{yr}$; Fernandes and Dietrich, 1997; Martin and Church,
This is despite the fact that measured diffusivities tend to be toward the lower end of this range \( (k_d = 10^{-4}-10^{-2} \text{ m}^2/\text{yr}; \text{Martin, 2000}) \). These higher-than-measured values of diffusivity have been used in many SPMs that do not explicitly model landsliding in order to obtain more realistic landscapes (e.g. Koons, 1989; Kooi and Beaumont, 1994, 1996). SPMs that implicitly model landsliding use values that are much closer to field observations (e.g., \( 1 \times 10^{-2}; \) Tucker and Slingerland, 1997; \( 1 \times 10^{-3}; \) van der Beek et al., 1999); however, the size of individual landslides is determined by the grid resolution of the SPM.

The coarse spatial resolutions employed by some SPMs (1km: e.g., Kooi and Beaumont, 1994; 5km: e.g., van der Beek et al., 1999) mean that topographic elements, such as hillslopes and river channels, are much larger than their real world counterparts on which measurements can be made. Likewise, the temporal resolution commonly encapsulated in an SPM means that model time steps are much longer, and generally seek to encompass (and therefore represent) much larger ‘events’, than those for which field data are available to generate parameters (e.g., Cascade’s model time step that notionally represents 100 years of landscape evolution - Braun and Sambridge, 1997). In other words, SPM parameters often cannot be derived from – or calibrated by – field measurements. As Anderson (1994) has noted: “The effective diffusivities necessary to ensure that the resulting model elevations remain reasonable are commonly many times those measured in the field” (p.20,162). That is, model parameterisation may serve to ensure that the model produces ‘realistic’ output. In fact, the model parameters may have no direct physical significance outside the model context (van der Beek and Braun, 1998), serving, for example, to translate ‘model time’ to ‘real time’ (van der Beek et al., 1999). In other words, rather than having explicit physical meaning, the model
parameters dictate the form and process rates predicted by models, and so the parameters for an SPM for a particular region may be able to be constrained by reference to the landforms of that region (noting the potential for inherent circularity here – van der Beek et al. 1999). The absolute values of process rates in an SPM may therefore be of less interest than spatial and temporal variations in process rates predicted by the model (e.g., van der Beek et al., 1999; Boogart et al., 2003). Tucker et al. (2001b) have made a further point in relation to variable mesh grids that is relevant here: “… the use of a variable spatial resolution complicates the inclusion of ‘scale-dependent physics’ (i.e., equations whose rate constants depend on spatial scale). This may be a blessing in disguise, for although it makes the problem of calibration in engineering applications more difficult it also provides a disincentive to scale-dependent ‘tuning’ of parameters” (p.29).

III Testing

The large literature on the testing of geomorphological numerical models has barely yet touched on testing of SPMs. Anderson (1994) has made the general point that SPMs “should be testable through measurement of geophysical, geologic and geomorphic information” (p.20,161), and this would appear to be most sensibly undertaken with reference to the reason for, or the nature of the problem being addressed in, using the SPM. Numerical modelling of long-term landscape evolution is undertaken for a range of reasons: to test tectonic models for the evolution of particular regions and thereby to gain insight into larger scale (plate) tectonic processes (e.g., Koons, 1989; Anderson, 1994; Gilchrist et al., 1994; Rosenbloom and Anderson, 1994); to test geomorphic models of long-term landscape evolution of particular regions without explicit linkage to larger scale (plate) tectonic processes (e.g., van der Beek et al., 1999); to attempt to understand
the key controls of landscape evolution in large-scale generic landscape types (e.g., high
elevation passive margins: Kooi and Beaumont, 1994; Tucker and Slingerland, 1994;
convergent orogens: Willett, 1999); and even to test classic models of long-term
landscape evolution, such as in Kooi and Beaumont’s (1996) use of an SPM to assess the
Davisian Cycle of Erosion. There are generic and specific issues of testing for each of
these rationales for using an SPM.

1 Internal structure of the model

An essential and powerful test of a model is its internal structure and logic. SPMs
routinely represent up to six processes of landscape formation (not all include as many as
this, however, on the argument that there is no need, for example, to represent
landsliding in some terrains, or even bedrock weathering) and all of these must operate in
the correct logical order. For example, the correct downstream order of all cells has to be
determined before runoff routing and mass balance calculations can commence. In order
to achieve this, variables such as slope have to be calculated at the beginning of each time
step and stored. Baselevel change (e.g., by surface uplift) is applied to the topography
after all cells have been accounted for and have had their values updated. The relatively
small number of processes represented by these numerical models and the relatively
simple interactions between the different modules make logical inconsistencies relatively
easily identifiable since their presence should be obvious in the modelling results. In
more complex numerical models, the effects of logical inconsistencies become subtler
and therefore more difficult to identify. When numerical models are used to provide
insights into the controls of long-term landscape evolution and their interactions (e.g.,
Tucker and Slingerland, 1994; Kooi and Beaumont, 1994), qualitative comparison of
their results with generalized versions of real landscapes is a major goal. In this type of heuristic use of models (cf. Oreskes et al., 1994), the validity of the results can be accepted as valid only if there is an *a priori* assumption of internal consistency and physical meaningfulness of the governing equations. Hoey et al. (2003) have noted that this *a priori* assumption is often based on geomorphological ‘intuition’ rather than direct evidence.

**Figure 5 around here**

2 Qualitative tests

Numerical modelling of landscape evolution does not have to be considered as an attempt to construct an exact replica of an existing landscape. It can be considered more as a “thought experiment” meant to improve our understanding of either the range of processes that form and shape the landscape (e.g., Kooi and Beaumont, 1994; Tucker and Slingerland, 1994; Tucker and Bras, 1998) or of the history (and perhaps the future) of a specific place (e.g., Ellis et al., 1999; van der Beek et al., 1999, 2001). Qualitative assessment of model results is often the primary means of model testing in the case of such experiments. In their modelling of south-eastern Australia, van der Beek and Braun (1999) and van der Beek et al. (1999) considered the replication of particular attributes of the studied site to be the primary factor in accepting or refuting the model results. Likewise, Ellis et al. (1999) used the Zscape SPM (Densmore et al., 1998) to explore the development of mountainous topography in the Basin and Range province, USA. In a related vein, Tucker and Bras (1998) used a numerical model of landscape evolution to
explore the influence of different types of hillslope processes and their process rates and thresholds (e.g., simple competition between creep and runoff erosion or simple threshold landsliding) on the three-dimensional morphology of the landscape. The ‘signature’ of the dominant processes is very clear on the resulting topographies, which have a very strong resemblance to real world landscapes on which the same range of processes operate (Figure 5). Tucker and Bras (1998) were not aiming at testing the different algorithms and were more interested in exploring the influence of these processes on the shape and form of the modelled landscapes. However, the distinct morphologies of the simulated landscapes and their resemblance (in a qualitative way) to real landscapes illustrate the robustness of the different algorithms and their capacity to capture and model the operating processes. Likewise, Kooi and Beaumont’s (1994) modelling of passive margin escarpment evolution under “arid” and “humid” conditions highlights the qualitative differences in the simulated landscapes.

Notwithstanding these successes in generating qualitative differences between landscapes simulated under different controlling conditions, several issues emerge when this style of model testing is employed, notably the issue of equifinality (Oreskes et al., 1994; Beven, 1996). Even in the case of simple numerical models different scenarios can produce topographies of very similar appearances, which are indistinguishable when using simple qualitative methods alone to test the results; Figure 6A presents an example of such an effect. A related issue arises when the synthetic topography does not ‘match’ the natural one. The next ‘logical’ step in such a case would be to abandon the model. However, as Bras et al. (2003) have argued, the inability to confirm results does not imply that the model is flawed. Therefore the only purpose such ‘tests’ have is to reassure us that everything has worked ‘as planned’. 
A universally accepted statistic to capture and quantify landscape for comparison with SPM output is yet to be agreed but several authors have made use of a wide range of topographic / geomorphic descriptors in an attempt to test SPMs. A summary of these statistics is presented in Table 1, the following discussion being limited only to those that have been used more extensively in the literature. Braun and Sambridge (1997), for example, showed that the drainage network of their synthetic landscape obeys Horton’s (1945) and Schumm’s (1956) laws of drainage network composition, but they also noted the near-universality of these relationships (Kirchner, 1993). Failure to obey these laws would indicate serious problems in the SPM, but these statistically inevitable relationships cannot be used in model testing since, like qualitative tests, they are incapable of falsifying models (Willgoose, 1994b). Fractal analyses have been employed to describe the characteristics of landforms (e.g., Mandelbrot, 1982; Huang and Turcotte, 1989; Weissel et al., 1994, 1995). Two widely used measures are fractal dimension and roughness amplitude, defining the relief of a landscape as well as its length-scaling properties (average relief at unit length). Braun and Sambridge (1997) calculated these two measures in order to compare the scaling properties of their synthetic landscapes with those of natural topographies and obtained values for fractal dimension in the range of 2.18 – 2.49, well within the range of values estimated for natural landscapes (2.1-2.7; Mandelbrot, 1982). van der Beek and Braun (1998) subsequently found that fractal dimension in their real landscape DEM data had a random spatial variation and was not
correlated with any other morphometric measure, but showed a predictable behaviour in their synthetic landscapes. They also found that the second fractal measure (roughness amplitude) was quantitatively controlled by the SPM’s input parameters (e.g., initial elevation and tectonic uplift), and they therefore concluded that these measures were inadequate for testing model output.

Willgoose (1994b) has shown that area-slope and area-slope-elevation relations (Flint, 1974; Tarboton et al., 1989; Willgoose et al., 1991d; Willgoose, 1994a) are robust against the spatial and temporal variation of the model inputs, asserting that these statistics are suitable for testing SPMs against field data. Subsequently Willgoose and colleagues have used these two statistics along with a series of other geomorphological descriptors (i.e. hypsometric curve, width function, cumulative area diagram) to test the outputs of their catchment-scale SPMs against DEM data from real catchments (e.g., Hancock et al., 2002) and experimental catchments produced with a landscape simulator (e.g., Hancock and Willgoose, 2002, 2004). Figure 7 provides an example of the use of area-slope and area-elevation relations.

Landscapes are dynamic features that change through time and these topographic statistics are only able to provide a ‘snapshot’ of the landscape at a given time. Therefore, when using such topographic descriptors to test model results one needs to be certain that both the synthetic and natural landscapes are at similar ‘stages’ of their evolution. Hancock et al. (2002) and Hancock and Willgoose (2002, 2003, 2004) achieve this by using catchments considered to be in a state of dynamic equilibrium (steady state) and allowing the models (both the SPM and the experimental landscape simulator [a physical model]) to evolve until the simulated landscapes reach this state (Willgoose et al., 1991d,
Willgoose, 1994b). The use of landscape simulators eliminates the need to test numerical models against ‘snapshots’ of the real world and allows for comparisons of results at different stages of evolution, therefore enabling a more rigorous assessment of model behaviour. However, landscape simulators are themselves models that also need testing (e.g. Hancock and Willgoose, 2003). Therefore, comparing their outputs with those of numerical models of landscape evolution may prove useful but cannot be considered to be an appropriate methodology for validation.

Figure 6 around here

4 Tests incorporating history

Landscapes are dynamic features that evolve through time, and so it is desirable to shift from trying to assess whether SPM-generated synthetic landscapes ‘look and feel’ like natural ones towards assessing whether they function as such and have had the same ‘history’ as a comparable natural landscape. The latter may be achieved by comparing measures such as the known and simulated sediment fluxes from the landscape, which are indicators of the rates of processes and landscape evolution. In their study of the geomorphic evolution of southwestern Africa, Gilchrist et al. (1994) used a SPM to explore the factors that control landscape development and the context in which escarpments may have evolved. They presented four numerical experiments that investigate the roles of initial topography, lithology, climate change, and interior catchment baselevel lowering on the styles of evolution of an initially high elevation margin bordered by an escarpment. Comparing field evidence to the modelled patterns
of denudation for the four experiments showed that, although the use of different parameter values can produce similar results, the model trajectories under these different parameterisations may be distinctly different. van der Beek and Braun (1998) found that estimates of long-term denudation rates impose tighter constraints on SPM parameter values than does the observed present-day topography alone.

Thermochronology (e.g., Cockburn et al., 2000; Persano et al., 2002; Brown et al., 2002) and cosmogenic isotope analyses (e.g., Lal, 1991; Cockburn and Summerfield, 2004) provide potentially useful tests of model performance over long time periods. SPMs calculate process rates at all nodes at every time step, meaning that synthetic thermochronological and cosmogenic isotope concentration data can be generated. In the case of fission track data (van der Beek et al., 1999), assumptions about surface temperature and geothermal gradient allow maps of fission track ages and track lengths to be generated and directly compared to field data. Hoey et al. (2003) presented an example of model result and field data for the Sierra Nevada mountain block, Spain (Figure 6). The SPM’s predictions of apatite fission track ages for the initial uplift function (Scenario I in Figure 6) is satisfactory but the track length predictions underestimate measured apatite fission track lengths. Longer track lengths require a more rapid passage through the fission track partial annealing zone (Gleadow and Brown, 2000) and so the under-estimation of track lengths suggests the probable form of the revised uplift function. A revised uplift function (Scenario II – rapid for 1 Ma and slower for 4 Ma) provides the same prediction of fission track ages and a better prediction of measured track lengths. The strong gradient in modelled fission track ages and the comparative paucity of field data combine to make formal observed vs. predicted fission track age testing difficult. However, the track length information coupled with
other properties of the model landscape (Figure 7) enable alternative uplift scenarios to be differentiated.

The fission track-based testing has the advantages of being direct and quantitative, and of relating model and field properties that operate over directly comparable scales. Sediment fluxes and thermochronological data can be used to compare the temporal evolution of modelled and actual topographies. These data combined with the morphometric statistics of the final topography become powerful tools in testing numerical models of long-term landscape evolution as well as in better constraining the SPM's parameter values. In addition, further work on modelling and directly measuring probability density functions of detrital cosmogenic isotope concentrations (e.g., Hoey et al., in preparation) has the potential to provide tests of the temporal evolution of the model for the shorter (more recent) time scales, while also carrying a signature of the topography that should be consistent between modelled and actual topographies.

**Figure 7 around here**

**IV Discussion**

Needless to say, with rapidly increasing computing power, a convergence between the smaller- and larger-scale models is increasingly possible, but as Tucker et al. (2001b) recently noted, it is difficult to envisage the incorporation of individual storm events in simulations of landscape evolution over 10s to 100s Myrs. In any event, it is not clear whether this ever-increasing ‘accuracy’ or verisimilitude will add anything, given the aim
of numerical modelling. As Merritts and Ellis (1994) expressed it: “like any experiment, they [numerical models of landscape development] enable us to explore the ‘what if’ questions by changing one variable at a time” (p.12139). Kooi and Beaumont (1994) and Gilchrist et al. (1994) have provided excellent examples of this approach, highlighting the use of numerical models to ask the ‘what if’ questions, rather than for the testing of particular landscape histories. van der Beek et al.’s (1999) use of a numerical model to explore the evolution of SE Australian landscapes and drainage systems has highlighted the ways in which these modelling exercises can provide insights as to possible pathways of landscape evolution rather than ‘proving’ one history or another.

Successful prediction of long-term landscape evolution is both of practical importance for geological applications and of intellectual importance as the results permit the evaluation of core concepts in geomorphology. The latter appreciation means that the evaluation of competing generic models of landscape evolution should become increasingly possible. The results of almost any numerical model can be used to address such issues. However, if these results are to contain more than speculation there needs to be rigorous testing of the models, which, in turn, necessitates that the purpose of the models is explicit. Likewise, it must be clear that the outcome of the testing is not pre-ordained by the structure and parameterisation of the SPM and the formulation of its process laws. For example, it would seem inevitable that SPMs built around advective (fluvial) and diffusive (hillslope) process laws that are fundamentally formulated in terms of slope (i.e., the $S$ term in the equations above) will lead to slope decline over time in the absence of active uplift to ‘drive’ river incision. Thus it is not surprising that the landscape evolution simulated by a SPM formulated in this way is similar to that proposed by WM Davis in his Geographical Cycle. The application of the SPM in this
way is not therefore a particularly powerful assessment of the validity of the Davisian Cycle (cf. Kooi and Beaumont 1996).

Models aiming to elucidate the long-term evolution of a particular locality must be able to generate outputs that are directly comparable to available field data for that locality, at the appropriate temporal and spatial resolutions. If the numerical model is being used to test a conceptual model the SPM’s internal consistency, and the ‘process realism’ of its governing equations, are the prime requirements for model validity. When confidence in internal consistency and process realism is high, the numerical model can provide the prototype against which the real world is tested. Of course, continued failure to find any part of the real world that matches the prototype will ultimately lead to dissatisfaction with, and will therefore be a test of, the prototype, but this ‘testing’ of the prototype is not the primary, or at least initial, objective. One of the difficulties in testing SPMs lies in avoiding the problem of model equifinality (Beven, 1996). This can arise either because different model settings produce indistinguishable results, or because the measures used for model testing are insensitive to model settings (eg., Braun and Sambridge, 1997). Testing such models thus requires that the landscape properties used in testing are able to provide critical differentiation between model settings, and that strong statements of model expectations are used (Beven, 1996; van der Beek and Braun, 1999). Creative experiments (Beven, 1996) are needed to produce strong tests of model assumptions and hypotheses. As Beven (1996) further points out, failure in a test can often be avoided by modellers adding or refining their assumptions, and this can easily lead to modeling becoming separated from reality, a shortcoming cautioned against by Klemes (1994) and more generally by philosophers of science, such as Popper (1959), when arguing against ad hoc modifications to hypotheses. Effective testing of geomorphological models
requires strong statements of expectations, which in turn requires that field research and modelling are carried out jointly. If geomorphologists fail to test and verify models at appropriate scales the potential for any more than heuristic use of models will be limited, and it will remain easy for modeling sceptics to dismiss model results as untested speculation. Finally, there needs to be continued recognition that both model outputs and prototype data are often poorly specified and subject to considerable uncertainty. Modelling studies cannot always seek direct, definitive answers but can be directed at constraining the uncertainty introduced by contingent behaviour (Beven, 1996; Kirkby, 1996).

Acknowledgements

Funding to ATC from a Glasgow University Scholarship and a Universities UK - ORS Award are acknowledged.
References

Ahnert, F. 1976: Brief description of a comprehensive three-dimensional process-
response model of landform development. Zeitschrift für Geomorphologie,
Supplementband 25, 29–49.

Research 12, 367-80.


Anderson, R.S. 1994: Evolution of the Santa Cruz Mountains, California, through


Martin, Y. 2000: Modelling hillslope evolution: linear and nonlinear transport relations. *Geomorphology* 34, 1-21


Strahler, A.N. 1952: Hypsometric (area-altitude) analysis of erosional topography.

*Geological Society of America Bulletin* 63, 1117-42.


www.geog.leeds.ac.uk/people/g.willgoose/grwpages/siberia_manual.pdf


Figure Captions

Figure 1. (A1) Tucker and Slingerland's (1994) conceptualisation of the key landscape evolution processes, and (A2) their model's regular grid representation of the topography and drainage (Adapted from Tucker and Slingerland, 1994 (c) The American Geophysical Union). (B) Cartoon of the CASCADE surface processes model and the governing equations (Adapted from van der Beek et al., 1999 (c) Taylor & Francis Ltd - http://www.tandf.co.uk/journals). (C) and (D) The building blocks of regular (rectangular) and irregular grids. The notations are those used in equations [6a] and [6b].

Figure 2. Flow accumulation maps from the NW part of the Rio Torrente, Sierra Nevada, S. Spain. The DEM (A) has a resolution of 10 m, obtained by resampling a 3 m resolution elevation model produced by digital photogrammetry from aerial photographs. The area depicted by the map is dominated by diffusive hillslope processes (rock and debris slides) with channelised flow being almost completely absent on the hillslopes. (B), (C) and (D) show flow accumulation maps calculated using three different flow routing methods: (B) D8 (Deterministic 8) (O’Callaghan and Mark, 1984; Jenson and Domingue, 1988) routes all water originating from a cell to its lowest downhill neighbour; (C) Multiple Direction (Freeman, 1991; Quinn et al., 1991) routes water to all the downhill neighbours, with each of them receiving a certain proportion based on the gradient between the source cell and the neighbours; and (D) D-Infinity (Deterministic Infinity) (Tarboton, 1997) behaves like the D8 method when the direction of flow is a multiple of 45 degrees (function of topographic curvature), and like the Multiple Direction method otherwise. Flow accumulation is defined as the number of DEM cells that drain through a given cell, and it is highly dependent on the flow routing algorithm. Flow accumulation
maps are illustrations of how runoff is being routed through a catchment by different algorithms. It is clear from (B) that D8, in which all flow is channelised, even on the hillslopes, is not capable of modelling the dispersion of water and sediment. Moreover, by routing material to only eight directions, D8 is highly insensitive to the real topography of the landscape because all channels are parallel irrespective of underlaying topography. The methods illustrated by (C) and (D) allow more flexibility in terms of flow directions (no parallel flow lines) and can model dispersion (channelised flow does not occur on the hillslopes). It is clear that any SPM using the D8 flow routing method would fail to model the diffusive hillslope processes that dominate this particular landscape.

**Figure 3.** Comparisons, using the DEM in Figure 2, of catchment areas calculated using three different flow routing methods: (A) D8 vs. Multiple Direction, and (B) D8 vs. D-Infinity. Note the degree of departure from the 1:1 line for catchment areas less than required for fluvial flow. The magnitude of this discrepancy prompts the need for more research into the validity of the various flow routing algorithms and their appropriateness at various DEM resolutions.

**Figure 4.** Comparison of results of numerical modelling of landscape evolution using CASCADE on (A) regular and (B) irregular grids with rose diagrams comparing river segment orientations and lengths for the two experiments (insets).

*Source:* Modified from Braun and Sambridge, 1997; Figures 1, 11 and 13.

**Figure 5.** Examples of simulated landscapes with varying hillslope processes: (A) simultaneous action of creep and runoff erosion, (B) hillslopes dominated by pore
pressure driven landslides and (C) runoff production by saturation-excess overland flow.

*Source:* Modified from Tucker and Bras, 1998; Figures 1, 5 and 7 (c) The American Geophysical Union.

**Figure 6.** Field and modelled data from the Sierra Nevada, Spain: (A) Digital elevation model derived from ~1 km data (global 30-arc-second elevation data, GTOPO30, U.S. Geological Survey, EROS data centre). (B) Simulated topographies generated using the CASCADE SPM (Braun and Sambridge, 1997) on a 1 km grid with model parameter values based on the Sierra Nevada. Scenario I uses a temporally uniform surface uplift rate over 5 Ma, with a maximum surface uplift rate of 1.25 mm.a$^{-1}$ near the central western end of the mountain block (the point of maximum elevation in the DEM), declining sinusoidally to zero at the mountain block’s outer edges. Scenario II employs rapid uplift for 1 Ma (maximum rate of 2.6 mm.a$^{-1}$ at the central western end of the orogen) followed by a reduced uplift (1.2 mm.a$^{-1}$) rate for a further 4 Ma. The (C) apatite fission track thermochronological ages and (D) apatite fission track mean track lengths based on the two model scenarios in (B) show important differences between the two scenarios, with scenario II more consistent with published field data (see Hoey *et al.*, 2003 for more detail).

**Figure 7.** Cumulative frequency distributions of (A) area-slope and (B) area-elevation for the two model runs in Figure 6B and for 1km and 4km grid resolutions in the Scenario 1 model for the area-slope data. Note the differential performance of the two scenarios for area-slope and area-elevation. Note how the topographies (Figure 6B) and the area-elevation plots (B) do not discriminate between the two uplift functions suggesting the attainment of topographic equilibrium by different routes (‘equifinality’ in the sense used
by Oreskes et al. (1994) and Beven (1996)). Note also that assessment of the goodness of fit of the modelled and actual topographies requires comparisons of the modelled and actual data at the same (or at least comparable) resolutions.

List of Tables

Table 1. Summary of topographic / geomorphic statistics employed for testing SPMs.
Table 1. Summary of topographic / geomorphic statistics employed for testing SPMs

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Description</th>
<th>Assumption on which testing is based</th>
<th>Sample references</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Horton and Schumm statistics</strong></td>
<td>A series of measures relating stream order to: the number (law of stream numbers; Horton, 1945), length (law of stream lengths; Horton, 1945) and drainage area (law of drainage areas, Schumm, 1956) of streams of that given order</td>
<td>Synthetic drainage networks should have scaling and branching statistics similar to natural networks</td>
<td>Braun and Sambridge (1997); Willgoose et al. (2003)</td>
</tr>
<tr>
<td><strong>Strahler statistics</strong></td>
<td>Similar to the Horton and Schumm statistics, but uses a different stream ordering system (Strahler, 1952)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Tokunaga statistics</strong></td>
<td>A scale invariant measure of river network branchiness based on the Strahler ordering system (Tokunaga, 1978)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>OCN energy</strong></td>
<td>Measure of total energy expenditure in the drainage network as defined by the Optimal Channel Network concept (Rodrigues-Iturbe et al., 1992b; Rinaldo et al., 1992). It is a function of link length and drainage area (Rigon et al., 1993)</td>
<td>The total energy expenditure in the synthetic drainage network should be similar to that in the natural catchment</td>
<td>Willgoose et al. (2003)</td>
</tr>
<tr>
<td><strong>Catchment convergence</strong></td>
<td>The average number of nodes that drain into a downstream node. Calculated only for grid-based DEMs (Ibbitt et al., 1999)</td>
<td>Synthetic catchments should have a convergence value similar to that derived from the ‘real world’ DEM of the natural catchment</td>
<td>Ibbitt et al. (1999); Willgoose et al. (2003)</td>
</tr>
<tr>
<td><strong>Hypsometric curve</strong></td>
<td>Nondimensional plot of the area of land above a given elevation (Strahler, 1952)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Area-slope and area-slope-elevation relations</strong></td>
<td>Plot of the area draining through a point versus the topographic gradient and elevation at that point (Flint, 1974; Tarboton et al., 1989; Willgoose et al., 1991d; Willgoose, 1994a)</td>
<td>The characteristics of these functions depend on the nature and distribution of the dominant surface processes, catchment geometry and network form (Perera and Willgoose, 1998; Willgoose and Hancock, 1998)</td>
<td>Willgoose (1994b); Willgoose and Hancock (1998); Ibbitt et al. (1999); Willgoose (1999); Hancock and Willgoose (2002); Hancock et al. (2002); Hancock (2003); Hancock and Willgoose (2003); Willgoose et al. (2003); Hancock and Willgoose (2004)</td>
</tr>
<tr>
<td><strong>Width function</strong></td>
<td>Plot of the number of drainage paths at a given distance from the outlet (Surkan, 1968)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Cumulative area distribution (CAD)</strong></td>
<td>Plot of the proportion of the catchment that has a drainage area greater than or equal to a specified drainage area (Rodrigues-Iturbe et al., 1992a, La Barbera and Roth, 1994)</td>
<td>The functions associated with synthetic landscapes should reproduce the characteristic forms of the natural topography</td>
<td></td>
</tr>
<tr>
<td><strong>Slope of CAD</strong></td>
<td>Plot of the slope of the cumulative area (Rodrigues-Iturbe et al., 1992a, Perera and Willgoose, 1998)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>River network fractal dimension</strong></td>
<td>Measure of river network complexity (i.e. how thoroughly do river channels fill the drainage area); it takes values between 1(smooth lines) and 2 (Knighton, 1998)</td>
<td>Studies of natural river networks (Tarboton, et al., 1988; La Barbera and Rosso, 1989) have obtained values for this statistic in the range: 1.5-2. Synthetic networks, therefore, should have a fractal dimension close to this range</td>
<td>Braun and Sambridge (1997)</td>
</tr>
<tr>
<td><strong>Landscape fractal dimension and roughness amplitude</strong></td>
<td>Defines the relief and length scaling properties (average relief at unit length) of the landscape (Mark and Aronson, 1984; Chase, 1992)</td>
<td>Synthetic landscapes should exhibit the same range of fractal properties as their natural counterparts</td>
<td>Braun and Sambridge (1997); van der Beek and Braun (1998)</td>
</tr>
<tr>
<td><strong>Planform geometry statistics</strong></td>
<td>A series of measures (i.e. curvature at a point and in its vicinity; the moments and spectral characteristics of the curvature series; statistical properties of the generalised planform) aimed at defining the planform characteristics of the landscape with the scope of discriminating between different types of morphologies (Howard, 1995)</td>
<td>Different processes yield landscapes with distinct morphologies; therefore different SPM parameterisations should yield synthetic landscapes with distinct planform geometries</td>
<td>Howard (1995)</td>
</tr>
</tbody>
</table>

Note: Nearly all studies use a combination of approaches, and the association of a particular study with one of the groups of topographic / geomorphic statistics implies only that the study in question exemplifies the use of at least one of the statistics contained in that group and not that it excludes other statistics or approaches.