Chapter 3
Contemporary environmental models

3.1 Introduction

Chapter 1 has outlined the aims of the thesis, namely to use or develop a model to study the effects of environmental change on an upland catchment. There are a wide range of models and techniques currently developed for the study of geomorphological problems and catchment evolution that are designed to operate over a wide range of temporal and spatial scales. This chapter will critically review relevant examples of these environmental models and conclude by selecting a methodology from which a model has been developed.

3.2 Fluvial models

Fluvial models have possibly seen the largest amount of research, which is undoubtedly linked to the importance and hazards of rivers to modern life. Consequently engineers need to calculate and predict the effects of river flow for channelisation and flood defence constructions. As this engineering perspective concentrates on comparatively short term predictions, there is a bias towards models aiming for finite, quantitative results, trying to predict what will happen instead of addressing why.

3.2.1 Cross sectional approach

Some of the most successful and widely used hydraulic models used are based around the reconstruction of water surfaces across cross sections. These generally use a 'step-backwater' calculation, using an iterative solution of Manning's or similar roughness function to determine the stage for a given discharge at that cross section. One of the most widely used are the HEC family of models developed in the 1960s by the U.S. Army engineering corps. These have now become quite sophisticated, with various fixes and patches for modelling hydraulically complicated areas such as bridges and weirs. Furthermore, these cross sectional stage calculations are often now integrated within a GIS environment to extrapolate inundation areas over large sections of flood plains. The latest versions (HEC 6) incorporate existing sediment transport relationships using water surfaces calculated by the HEC model. Tingsanchali and Supharatid (1996) tested the performance of this model against results from a 12m test flume. They found that Yang (1972) sediment transport equation was most effective, and that altering the hydraulic parameters (e.g. roughness and resolution) had less effect than the sediment transport law used. Their experiments showed that the model produced good representations of long term patterns of scour and deposition, but accuracy was reduced in complex flow regions with highly non-uniform flow. This is one of the main problem with HEC models, as channel flow is complex and therefore can never be fully represented in one dimension for example secondary circulation is not incorporated. Furthermore, the HEC family of step-backwater models all have difficulties dealing with complicated channel patterns or flow areas, for example at a confluence. However, for forecasting flood levels and engineering purposes in relatively simple single channel sections they have proved invaluable.

Other authors have taken the one-dimensional approach and added terms to account for such cross channel flows. Alabyan (1996) used a series of 1-D sections linked together to simulate the development of meander bends. He simulated secondary circulation according to the radius of curvature which was then used to calculate cross stream flow circulation and hence bank erosion. This allowed meander bends to propagate or expand outwards, with a consequent effect on the downstream bends. Nicholas (1996) used a similar approach to model erosion and incision on a straight channel with a gravel bed and cohesive banks. He found that bank failure was controlled by the cohesivity of the bank material, but the temporal sequence of bank failure is controlled by factors that influence the rates of cross stream sediment transport.

However, all of these one-dimensional schemes only look at the wetted channel cross section. Whilst their simplicity has led to their success, it is also their major limitation as they fail to incorporate slope processes, fine scale changes in channel morphology or secondary circulation.

3.2.2 Finite element modelling

Finite mesh or finite element models set about answering many of the criticisms of one dimensional flow models by solving the Navier-Stokes flow equations for conservation of mass and momentum in two or three dimensions, to produce a more accurate representation of both the water surface and flow velocities. These models have developed a distinct literature, as computational fluid dynamics (CFD).

Bates et al (1997) compared two generalised two-dimensional finite element codes, RMA-2 and TELEMAC-2D on an 11km reach of the River Culm, Devon U.K. Using a mesh of nodes and links to represent the channel and valley floor topography, (Figure 3.1). RMA-2 and TELEMAC-2D both solve second order partial differential equations for depth averaged flow derived from the full three-dimensional Navier-Stokes equations. This gives the depth of flow showing inundation areas and calculates the velocity in two Cartesian directions. Bates *et al.* (1997) then assessed the accuracy and sensitivity to changes in the numerical solution used and the resolution of the finite mesh. They noted problems associated with calculating the free surface at channel margins and discussed the merits of different calculation techniques. Additionally, they measured the calculation times for the passage of a mean annual flood at different resolution meshes, which ranged from 265 minutes for 1200 nodes to 2930 minutes for 5600 nodes on an HP workstation. They concluded that there was little change in accuracy moving from one resolution to another and the model generated good flood plain inundation results. However, the construction of the finite element mesh takes a long time and there were problems obtaining sufficiently accurate topographic and bathymetric data.

Miller (1994) takes a small section of canyon channel and compares the flow at varying stages with and without a restriction over one side of the channel. His work re-creating the cross stream gradients caused by the restriction, throws doubt upon the ability of one dimensional step backwater hydraulic models in areas of complex morphology.

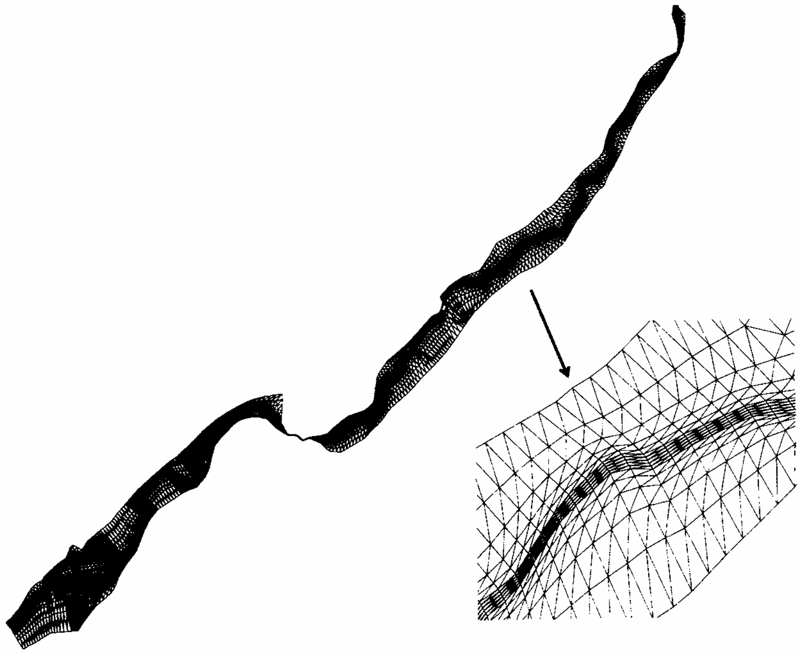

Figure 3.1. Finite element mesh for River Culm, from Bates et al. (1997).

Nicholas and Walling (1997) also carried out a modelling study on a different section of the River Culm. However, contrasting to Bates *et al.* (1997) they use a finite difference grid of fixed cell size coupled with a simple sedimentation model to determine rates of sedimentation on the floodplain. They created a 5m grid for a 600 by 600m area of channel and floodplain (Figure 3.2) and monitored the section for 16 months recording stage and sediment discharge. Instead of solving for water depth at each node for each time step during the passage of the flood, depths were calculated for several levels of stage by measuring the upstream and downstream stages then back calculating the water surface over the entire flood area. This pre-calculation effectively created a hydrograph, stage and thus inundation area for each time step reducing computation time and allowing concentration on the sediment transport. The deposition was calculated by using the measured suspended sediment concentrations and determining deposition according to an empirical settlement relationship for each node over a size range from 8 to >63 um. The modelled floodplain deposition amounts and pattern were then compared to those measured in the field from Astroturf sediment traps and from radio-nucleide dating. From the good fit obtained, the authors concluded that this method adequately integrates the many flows and currents associated with the complexities of floodplain topography. Furthermore, they show that during the early stage of the flood, most deposition is caused by ponding, and that as flood levels rise the ponded areas become interconnected creating a number of flow paths. As stage increases even more, with total floodplain inundation, the topography becomes the dominant factor. However, the work of Nicholas and Walling incorporates several major and questionable simplifications. Their adaptations of the Navier Stokes expressions fail to account for the momentum of flow between nodes and they assume that all flow occurs in the downstream direction of the main valley floor. Deposition is not integrated into the topography, and material cannot be eroded or re-entrained. This technique also requires that the spacing or time-step between each stage/inundation calculation is sufficiently small to capture all dynamic sediment movements in detail. However, this work is one of the few detailed modelling studies that has been successfully validated against field measurements.

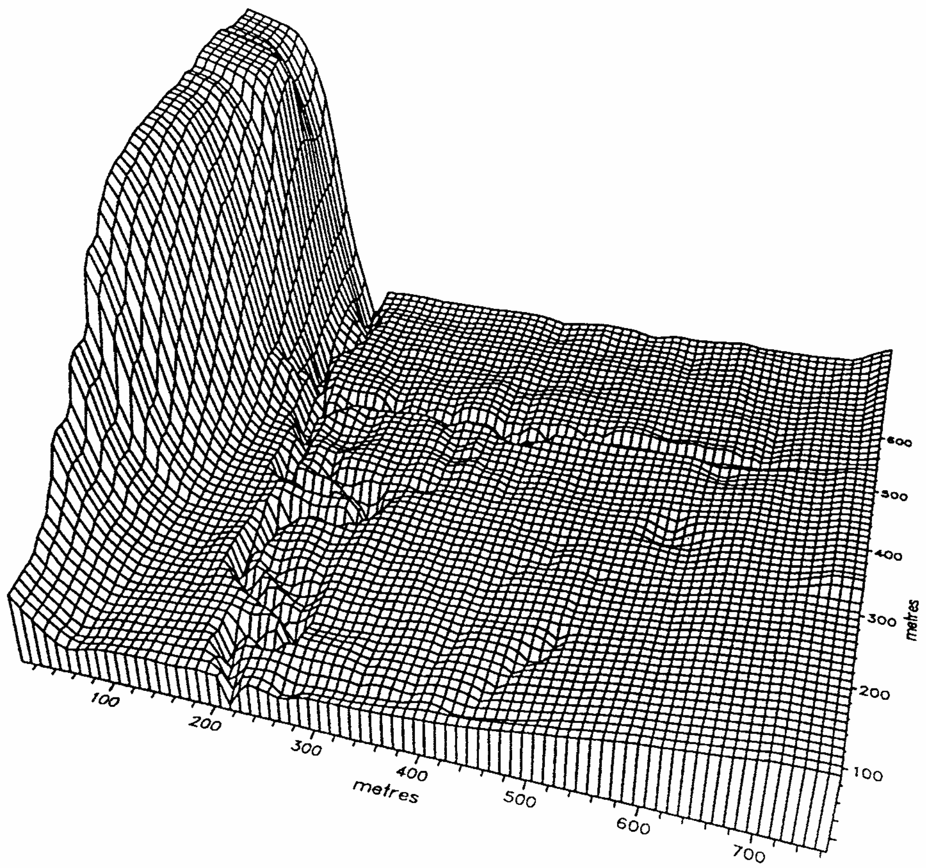

Figure 3.2. Finite element mesh used by Nicholas and Walling. (1997).

Lane (1998), reviewed a selection of such models and discussed some of the problems and difficulties associated with them. He identified three critical issues for all these CFD models. Firstly, the complexity of the Navier Stokes equations ideally requires a grid large enough to cover the area of interest, but with spacing small enough to show the smallest turbulent motion. Lane (1998) shows that to solve the relatively simple problem of steady flow in a pipe, 10^{22} calculations would be required taking an estimated 2 million years to calculate. Thus, to simplify these equations, empirical or semi-empirical expressions are substituted in a technique called Reynolds averaging. which allows the equations to be solved far more quickly. However, there are a number of methods for applying this averaging and a variety of exponents used. Furthermore, by simplifying these complex equations, some facets of fluid flow will be lost for example fine scale turbulence driving secondary circulation.

The second critical issue is the use of depth averaging and secondary circulation where velocities are averaged over depths and horizontally. This leads to problems of how to calculate or represent secondary circulation and bed roughness.

Lane's third issue is the need to make steady flow assumptions. Some authors use a 'Rigid Lid' strategy, assuming that water is in a steady flow with a fixed water surface. They argue that over the time and space of a floodplain, erosion and depositional processes triggered by turbulence will average each other out. However, most geomorphological problems are associated with unsteady flow and complex relationships between discharge and stage, for example when a channel floods over-bank, or during the steady rise and fall of a hydrograph. One-dimensional approaches, such as the previously described HEC series, deal with such non-steady flow but have their limitations. Therefore because of this non-steady flow, two or three dimensional schemes have to reconcile changes in the boundary of the water relative to the grid or topography. This can be achieved by using wet or dry cells, allowing the finite grid to deform or change with the wet boundary or by defining a large number of elements around the boundary to accommodate the change. These techniques all present problems, such as defining wet or dry cells, ensuring that the 'deforming' grid follows the topography correctly and the increase in computational complexity if a large number of extra nodes are required to represent changing boundaries. Furthermore there are issues surrounding the definition of the finite difference grid or finite element surface. This requires a finer resolution grid in zones where there are sudden changes in topography or roughness. Lane warns that whilst it is easy to become excited about these methods which apparently are not hampered by the inaccuracies of field and laboratory study, (CFD's) '*models in themselves are truncations of reality'*. Furthermore, '*their predictions are circumscribed by the accuracy and extent of the initial and boundary conditions, the assumptions made by the modeller and the dependence of the model on parameterisations which may have a weak physical basis.'*

When reviewing the application of CFD models to fluvial geomorphology, only Nicholas and Walling (1997) account for suspended sediment deposition and none of these schemes accounts for bedload transport. This may be due to several reasons. Whilst the flow of water and other liquids is fairly well understood, and can be well represented by equations, the movement of water and sediment is not well studied. The movement of bedload as opposed to suspended sediment again is a fairly poorly understood phenomenon as discussed later in 3.6 and 3.7. The introduction of large concentrations of suspended sediment (as in a flood situation) also changes the physical properties of the water/sediment mix. Additionally, as this mix is not homogenous the spatial and temporal heterogeneity's in density and flow further complicates the problem.

When finite element models are coupled to bedload transport models, the bed topography is continually changing, so that the element mesh has to be re-defined for every iteration. As Bates *et al.* (1997) describe, calculating a 5600 node mesh currently takes 2930 minutes which is unacceptably long. Nicholas and Walling (1997) circumvent this problem by pre calculating stages and not incorporating elevation alterations. But as changes in bed and flood-plain topography influence the flow and inundation area, this is an important feedback. The irony of this omission is highlighted by their conclusions regarding the importance of floodplain topography. Their approximation may be an acceptable when dealing with a few millimetres of deposition from a single moderate flood, but should not be acceptable if simulation periods and flood magnitudes are increased. Finally, as Lane (1998) implied, the results vary depending upon which formulation or package is used, because of the different simplifications and assumptions adopted in each.

3.2.3 Other fluvial methods

Operating over a larger temporal and spatial scale, there are a number of models designed to simulate the alluvial architecture, describing channel change and channel belt development. These models are of great use for oil extraction, as deposits are more readily extracted from within the gravel deposits of buried gravel river beds. As they are simulating the stratigraphy of aggrading flood plains over thousands of years, they are not concerned with the fine detail but only the general position and size of the channel. The first series of models to try and re-construct this stratigraphy were those of Allen (1978) and Bridge & Leeder (1979) often referred to as Leeder, Allen and Bridge models (LAB). Their approach revolves around avulsions, the frequency of which is driven by local sedimentation rates. Avulsion is triggered when the channel aggrades above a threshold level, causing it to switch to another lower path. Bridge and Leeder (1979) also assumed that the avulsion rate is uncoupled from the local sedimentation rate, and there are doubts as to whether channel growth and stacking can be represented in such a simple way.
Examining channel behaviour over similar time-scales, Veldkamp (1992) created a three dimensional model of Quaternary terrace development, modelling terrace stratigraphy and valley asymmetry in the River Allier, France. This represented the response of a channel belt to changes in uplift rates and simulated the development of Quaternary terrace sequences. He used a simple channel capacity erosion scheme, with a set of basic rules governing channel incision or aggradation. He concluded that terrace formation and preservation was determined by tectonics, but that the stratigraphy was controlled by climatic drivers.

3.3 Slope models

Processes acting upon hillslopes are responsible for shaping the profile and form of the landscape. For temperate climates such as the UK, soil creep and mass movements are generally the most important processes. Kirkby (1992) developed an erosion limited hillslope evolution model, that disaggregated hillslope erosional processes into detachment rates and transport distances. For mass movement the detachment rate is proportional to a rate parameter once a slope threshold is exceeded (equation 3.1)

$$D = D_0 \Lambda (\Lambda - \Lambda_0)^m \text{ (3.1)}$$

Where Λ_o is the threshold gradient, D_0 a rate parameter and m an empirical exponent varying from 1-3. This angle is generally less steep than a talus gradient (on which moving material will come to rest. This talus angle is used to determine the transport distance for the eroded material (equation 3.2)

$$h = \frac{h_0}{\Lambda_T - \Lambda} \text{ (3.2)}$$

Here, h is the travel distance, and Λ_t the gradient on which moving material will just come to rest. Where slopes are gentle enough, these approximations allow material to be re-deposited in a realistic manner.

Using a more sophisticated technique, Brookes *et al.*(1993) used a coupled slope hydrology and stability model to investigate Holocene shallow slope failure in Northern Scotland. This detailed study integrated soil development and hydrology over long time periods to investigate the role of different storm magnitudes on soils of varying ages. Furthermore, they compared the model results with a record of landslide events and climate change. They showed that as soils grew older they became more vulnerable to failure down to angles as low as 30 degrees. Despite finding some links between climate and landslide records, they decided that no firm conclusions could be draw, and to concentrate on the processes rather than the causes. Their model showed how slope behaviour was not a simple process and was tightly linked to the local hydrology, soil age and development.

Whilst solifluction or gelifluction is not a common processes today in the British Isles, there is widespread evidence of it during the early Holocene and it therefore has some relevance. Kirkby (1995) developed a regional model for gelifluction rates, using global climate data sets to estimate regional variations in gelifluction under existing and differing climates. This used a two stages, the first estimated the depth of the active layer in response to changes in global temperature and then the mass rate of gelifluction when the active layer of the surface becomes saturated. The process is complex, as when the soil thaws, a small active layer becomes rapidly saturated and may slide. But as temperatures increase the depth of the active layer increases until solifluction never happens. For the UK, this shows a ‘wave’ of gelifluction as the mean annual temperature crosses a 0^0C isotherm, with rapid slope erosion for the first 2-3000 years of the Holocene. Due to the higher latitudes, this period of enhanced slope activity would have occurred in Scotland about 2000 years later.

3.4 Hydrological models

There have been many hydrological models but probably the most widely applied and researched is TOPMODEL (Beven and Kirkby, 1979). This model is based around a simple topographic index $K = a / \tan\beta$ where a is the area draining through a point from up-slope and $\tan\beta$ is the local slope angle. Central to TOPMODEL is the assumption that there is a fundamental link between this topographic index and soil moisture. As gradients fall and the K index increases, the water table rises closer to the surface (Figure 3.3 A). As it rises however water is transmitted via through flow at a rate that increases exponentially. Therefore, as the water table rises, the rate of water loss from a point increases. As a simple parallel, the soil water store may be compared with a bucket with a small hole in the base. As more water is added to the bucket, water is forced through the hole at a greater rate. Two key variables control the behaviour of the model; the transmissivity or rate of movement *(k)* which is in turn scaled by a parameter *m* (3.3 B). This parameter *m* effectively determines the rate of rise or fall in the water table, in other words the 'flashiness' of the hydrograph, and can be calculated from hydrograph recession curves. The simplicity of this approach has been the key to its success. In particular, the derivation of the topographic index from DEM's is a relatively straightforward exercise meaning that TOPMODEL or its concepts can easily and automatically applied to large areas. The technique is not free from problems with sensitivities to the grid cell resolution and *m* parameter, but the wide application and depth of literature is testament to its robustness.

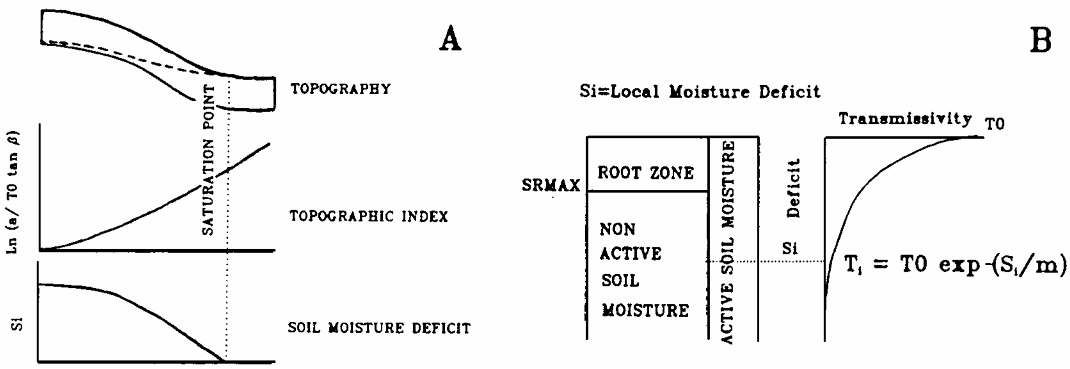

Figure 3.3. Interactions between TOPMODEL parameters, from Quinn et al. (1991).

3.5 Cellular models

3.5.1 Introduction

Cellular models or cellular automata (CA) have become increasingly popular, where a process is represented within a two-dimensional framework or mesh of grid cells. The spatial nature of many geographical processes and the widespread availability of gridded data from remote sensing sources are well suited to such cellular models. Furthermore, a digital computer processes data serially, whereas cellular automata are suited to processing in parallel, making them an ideal application for parallel supercomputers. Wolfram (1984) identifies five key factors which briefly define cellular automata:

1. they consist of a discrete lattice of cells,
2. they evolve in discrete time steps,
3. each cell takes on a finite set of possible states,
4. the state of each cell evolves according to the same deterministic laws,
5. the laws for cell evolution depend only on interactions with immediately neighbouring cells.

Although the concept of CA models is basic, the interaction between the cells can give complex, often non linear behaviour.

" even though the elementary components of a system may follow simple laws, the behaviour of the large collection of components which comprise the whole system may be very complex." Wolfram 1984.

CA models have been applied in many contexts, from ecological modelling to forest fire simulation (Clarke *et al.* 1994). Process based cellular models in geomorphology are reviewed below, with examples of non-linear and emergent behaviour and a discussion of the strengths and weaknesses of this approach. These models form into two clear groups, landscape evolution models, concerned with the long term development, and more specific cellular automaton models, examining aspects of this evolution in finer detail. As the former have external influences, they do not strictly adhere to rule 5 and should thus be classed as finite difference models.

However, as they depend heavily upon interactions between cells, as per rules 1-4, and have much in common with CA, for this review these landscape evolution schemes are termed cellular models as opposed to cellular automaton.

3.5.2 Landscape evolution models

Several authors have used a cellular framework to integrate and explore the effects of slope and channel processes in the evolution of landscapes. As this category of model is concerned with the long term catchment development (Greater than 100 000 years) they concentrate upon processes such as mass movement and creep, and tend to simplify fluvial processes.

In the late 1970s the first series of these landscape evolution models emerged. Whilst providing significant advances in modelling landform development, they were quite basic. Anhert (1976) used a 10 by 10 grid upon which geological patterns were set, forcing the location and density of a drainage pattern, concentrating heavily on slope processes. Armstrong (1976) used a larger 40 by 40 grid with a fixed drainage pattern and separate transport laws for slopes and channels. Cordova *et al.* (1983) used a process based approach with a randomly generated surface and a uniform transport law. However this only accounted for fluvial erosion and therefore ignored diffusive processes such as creep and rainsplash, resulting in the continuous growth of hollows. Furthermore, Cordova *et al.* (1983) only used a grid of 16 by 16, limiting the area modelled.

Kirkby (1987) introduced a more sophisticated approach to investigate how slope processes influence drainage density, especially the evolution of the stream head. He investigated the assumption that channel heads will grow where material is removed more rapidly from a hollow by fluvial processes, than it is in-filled by diffusive processes such as rainsplash and creep (Smith and Bretherton, 1972). He assumed that creep and rainsplash are largely proportional to the slope, and used a term for the estimated daily rainfall required before overland flow, for wash processes. These laws were then applied to a grid of 64 by 64 cells, each representing $10m^2$, which were skewed to represent a hexagonal mesh. This initial surface was given a fractal roughness from which the drainage network was allowed to evolve. Overland flow was routed to the lowest neighbour and he assumed that flow occupied the whole width of the cell. This means that a cell is either dominated by fluvial processes or the slope. To prevent the build up of computational instabilities, a variable time-step was used which prevented the gradient between cells changing by more than 20 %. Over runs of up to 240 000 simulated years, the model showed the development of a drainage network in a variety of situations, including a good example of stream capture. More importantly, it demonstrated how slope and channel processes combine and compete to scale the landscape and the drainage density.

Willgoose *et al.* (1991a, 1991b, 1991c and 1994) developed the computer model SIBERIA to explore links between hillslopes and a growing channel network, and to investigate whether channel growth was governed by hillslope form. They used terms for slope, diffusive and fluvial processes as well as a threshold for channel initiation dependant upon contributing drainage area and slope. As cells are defined either as channels or slopes, according to this threshold, the drainage network can expand and contract with differing magnitudes of rainfall. This interaction is one that Willgoose *et al.* (1994) considered vital in the development of a catchment and their concept of hydrogeomorphology. As they were trying to model the long term changes, mean peak discharge is used, ignoring single flood events. The main implication of this assumption is that the full range of flood events shapes a catchment, Willgoose *et al.* (1994) stated; "*Any single flood and erosion event is only a minor fluctuation in the life of a catchment"*. SIBERIA was then used to investigate dynamic equilibrium in landscape evolution, and Willgoose *et al.* discovered that a catchment can reach a statistical equilibrium where the sediment discharge equals the tectonic uplift, effectively the slope and fluvial erosion balancing uplift. The impacts of a diffusive slope process dominated, and fluvial controlled landscape were compared, showing how both can reach a dynamic equilibrium but with different forms. The area-slope relationships of the simulated landscapes were compared favourably against a field site.

Howard (1994) used a detachment limited approach to investigate the relative roles of diffusional or dispersive (wash and creep) and concentrative or advective processes (channel or stream head erosion). Gilbert (1909) summarised their impacts as *"on the upper slopes, where water currents are weak, soil creep dominates and* *the profiles are convex. On lower slopes water flow dominates and profiles are concave."*

Their model is similar to that of Willgoose *et al.* (1994), but contrasts in several ways. The model is detachment limited, whereas all the models described above are transport limited. It assumes that erosion in many locations, especially headwaters, is detachment limited. Furthermore, Howard's (1994) model allowed a cell both fluvial and slope transport, whereas the Willgoose model uses a threshold to distinguish them.

For a 100 by 100 matrix, Howard's model defines the channel within each relevant grid cell, and channel dimensions are determined by hydraulic geometry relationships. Material is also eroded between cells using detachment limited expressions for rainsplash, creep and mass movement. Bedrock also weathers to generate soil, but soil depth is not expressly modelled, but simply assumed to keep pace with material removed by creep and rainsplash. Mass movement occurs when a slope threshold is exceeded. Fluvial erosion incorporates both bedload and suspended transport, and there is a threshold between alluvial and non alluvial channels, *'a non alluvial channel is defined as one in which the bedload sediment flux is less than a capacity load'*. As this model concentrates far more on the fluvial action than its predecessors, the higher fluvial process rates require a smaller time step. To minimise computer run time, the averages of material added and removed by fluvial action are used, and this aggrades or erodes the whole cell, not just the channel portion within it. Like Willgoose *et al.* (1994), Howard assumes that given the long timescale of the simulation, the multiple creation and dissection of terraces together with flood plain re-working leads to only small alterations in the basin as a whole. He validates this fluvial scheme by comparing long profiles formed with this model to those developed by a traditional finite difference scheme.

Howard finds that a complex drainage pattern forms on an easily erodible regolith giving a high drainage density. When a threshold for fluvial erosion is introduced, there is a sharp transition between channel and slope, with a lower drainage density. This possibly simulates a temperate basin where there is substantial subsurface flow. Furthermore, this threshold shifts the dominant processes from concentrated fluvial action to diffusion on slopes. There is also increased inheritance from initial conditions, with channels developing a shallow drainage network defined by the fractal surface. He concludes that mathematical simulation of the processes is reasonable and that the simulated basins behaviour conforms to our existing ideas of evolution. For the future, he suggests more development on appropriate rates for process laws and validation of the model with reference to a target landscape.

Tucker and Slingerland (1994) examine a more geological issue, investigating the development of erosional escarpments. Their model is designed to find the conditions necessary for the long-term retreat of a rift based escarpment. Previous authors had thought this was a response to uplift, but sedimentation of the basin at the base of the escarpment also leads to a change in uplift rates through isostasy. This model can be considered as a cross between Ahnert (1976) and Howard (1994), with two different types of grid cell, rock and sediment. It has expressions for fluvial transport, shallow sediment landsliding, rock mass failure and diffusive slope processes. Their model allows for the weathering of bedrock by assuming that sediment production rates increase exponentially as the sediment mantle becomes thicker, allowing the development of a more easily erodible regolith. This formulation is used to investigate the effects of a low sediment production rate creating steep sloped arid landforms. They also speculate this may be relevant in steep temperate catchments where sediment transport exceeds supply and are in effect detachment limited. They conclude that the lateral retreat of a channel into an escarpment depends upon a balance between sediment supply and transport efficiency. With low weathering and high fluvial transport rates, bedrock channels cut back rapidly without being clogged by slope derived sediments.

3.5.3 Cellular automaton models

There are a few studies focusing on more specific areas of slope and fluvial geomorphology and they appear as successful as the landscape evolution models at simulating these. Favis-Mortlock (1996) used a CA model to investigate the evolution of rill networks. In direct contrast to Willgoose *et al.*(1994) who averaged his rainfall, Favis-Mortlock models the impact of individual raindrops or ‘run-off packets’ on a semi arid hillslope. He states*; “ A statistical description of a dynamic competitive process may well not capture those essential qualities of the system which emerge as a result of the underlying component-level interactions”.* These packets of rainfall erode the hillslope according to a stream power law, and are routed to the lowest neighbour. He describes the evolution of a rill network as a complex system demonstrating emergent behaviour, with an apparently chaotic sediment discharge. Despite the many shortcomings, which he lists, he considers that his 'evolving cellular automaton system' reproduces many of the larger scale features of a rill network. Simulated planform and rill spacing compare well with field measurements.

One of the most significant papers using cellular automata is that of Murray and Paola (1994). They use a simple cellular model to simulate the processes in a braided river. A section of uniform slope, 22 by 200 cells in size, is given an initial fractal or 'white noise' roughness. 'Water' is then introduced at the top of the slope and moved downstream cell by cell. The elevation of each cell is changed according to the sediment transport law used. The top and bottom ends of this slope are fixed as in a flume. From each cell, water is distributed to any or all of the three immediate downstream cells using the weighting algorithm (equation. 3.3).

$$Q_i = Q_0\left(S_i^n \Big/ \sum S_i^n \right) \quad (3.3)$$

Where Q_i is the discharge, S_i the slope to the downstream neighbour *i* and Q_0 is the total discharge carried. They describe the use of several sediment transport relationships, all driven by the discharge from the cell and the slope. An additional term is included which allows lateral movement of material from cells at right angles to the main flow.

By using a simple routing and sediment transport law, they produce a model which gives a visually realistic simulation of a braided channel. This includes a constantly migrating channel, whose form and magnitude remain similar, displaying a form of dynamic equilibrium. The model also produces a non-linear sediment discharge, giving pulses of sediment even though the discharge is constant. Murray and Paola (1994) suggest that these fluctuations develop due to channel migration into areas that may be un-channelled or shallow. Power series analysis of the total sediment load past a point in the model is compared favourably with that of a laboratory modelled stream. They also compare the geometric statistics for the average number of channel threads in a cross section in relation to channel width.

3.5.4 Examples of non-linear behaviour

Many geomorphic systems can be described as non-linear (Phillips, 1996) and one of the most interesting facets of CA models are their non-linear outputs and self-organisational tendencies. Murray and Paola (1994), whilst not specifically addressing non-linearity or chaotic processes mention the periodic and random switching of channel patterns, as well as the pulsing of sediment load. Favis-Mortlock (1996) describes how his rill model exhibits chaotic sediment discharge. These are both examples of how a simple set of equations is able to give rise to complex system behaviour. Both these CA models also appear to show a degree of self-organisation in that they evolve to a stable pattern. Certainly Murray and Paola's model organises itself from a blank surface into a braided channel pattern, and the rill-grow model similarly develops a channel or rill network.

Self-organisation has also been noted by Phillips (1995), who gives examples of the development of soil horizons and channel geometry. Other authors have examined the role of self-organisation and self organised criticality (SOC) in the development of channel networks and braided channel patterns (Sapozhnikov and Foufoula-Georgiou, 1996). A simple example of SOC is a sand-pile, where the addition of sand to the pile eventually steepens the slopes to the threshold of stability. At this point a local slide or failure is started which may trigger secondary failures below, or undermine the slopes above. These failure events may be small and locally constrained or may cover the whole system. The importance of studying SOC in braided channels was justified by Sapozhnikov and Foufoula-Georgiou (1996) who suggest that it ' *might indicate the presence of universal features in the underlying mechanisms responsible for the formation of braided rivers'*. However, self-organisation may not be an essential component of geomorphological development. Phillips (1995) stated '*While it would be accurate enough to say that geomorphic evolution* ***can*** *be self-organising, it would not be accurate to say that it* ***is*** *self-organising'*.

Phillips (1996) further discusses SOC and the importance of non-linear dynamic systems (NDS). Whilst finding many examples of NDS in geomorphology, he concludes that applying this knowledge can be difficult. For example, demonstrating that a system exhibits non-linear behaviour can be used to provide plausible explanations for unexplained phenomena, but rarely explains how the processes operate or evolve. However, if a system is recognised as being non-linear, different numerical techniques may be needed, for example the use by climate modellers of average values from an ensemble of simulations.

3.5.5 Limitations of cellular models

Cellular models seem to provide a medium for incorporating many processes within a simple frame work, which conventional modelling strategies fail to do. However, the technique is not free from problems.

3.5.5.1 Grid scale problems

Kirkby (1987) describes errors associated with the size of the grid cell and states that cell size will inevitably conceal a distribution of unit areas within the cell. This may have significant effects, especially if flow threads are much smaller than the cell. Howard (1994) negates this by defining channel parameters within the grid cell, but these depend upon empirical hydraulic geometry relationships. Howard then blurs this increase of resolution by averaging out the results of erosion or deposition over the whole cells elevation. The importance of grid-cell size will change with the scale of the problem or landscape studied, but remains an area of uncertainty. As Kirkby (1987) states:

"*The use of small cells plainly increases accuracy, but at the expense of the total area that may realistically be modelled".*

None of the other studies mention the effects of changing grid cell size, however Zhang and Montgomery (1994) computed the TOPMODEL topographic index (A/tanB) using DEM grid sizes of 2,4, 10, 30 and 90m. These were then used to drive TOPMODEL for two basins in the NW USA. They found the finest grid sizes (2 and 4m) were not very different from the 10m size, but the larger cell sizes were significantly different.

Ultimately, as the effects of grid cell size depend upon the processes studied, the methods of parameterisation, relief and topography, the choice of cell size is unique to the context of study.

3.5.5.2 Process representation and parameterisation

Fluvial erosion may well be considered the most important process in the evolution of many landscapes, yet is represented by a variety of methods. In relation to the slope processes operating, fluvial action is rapid, which may explain their typically lumped representation. The erosion laws used are all based upon reasonable assumptions, but none have been calibrated or based upon field data. For example, Howard (1994) splits the model into a fluvial and slope sections, then uses averages to update the cells. Furthermore, whilst Howard (1994) and Tucker and Slingerland (1994) differentiate between bedrock and alluvial channels, neither of their models account for erosion of different grainsizes. In gravel and boulder bed channels, channel armouring is an important process, as well as the development of stratigraphies.

Vegetation growth or change is not accounted for in any of the previously described methods and has important interactions. The empirical relationship between runoff, sediment discharge and vegetation has been summarised by Langbein and Schumm (1958), and vegetation has a large control on slope stability, overland flow and hydrology as described in 2.3. Hydrology is not incorporated in these models and with vegetation is crucial in catchment behaviour. Furthermore, there are vital links between the hydrology and slope processes (section 3.3).

Whilst all models are by definition abstract, none of the cellular models discussed above attempts to use real units. The laws and processes represented may be based upon observed rates, but no actual values are used. Furthermore, several are based upon an initial fractal generated surface. This begs the question as to whether the model results are simply a matter of inheritance from the initial conditions. In some ways, using a fractal surface can be seen as a fair approximation, as any initial surface will have a roughness to some degree. As Phillips (1996) stated *'Minor variations in initial conditions are not just unknown, but also unknowable'.* It might also be argued that the erosion laws used are designed to enlarge hollows in this initial surface and it is therefore important to be aware of any sensitivities to this surface. In addition, many of the processes are grossly simplified. It is sometimes convenient or sometimes only possible to model using a statistical representation of the process but this too is a further abstraction of the landscape processes operating.

3.5.5.3 Steepest flow algorithm

With the exception of Murray and Paola's (1994) model, all the above models use a steepest descent algorithm for determining flow routing where all the flow is routed down the path of steepest descent. Desmet and Govers (1996) compared six different routing algorithms for calculating the drainage area of a test catchment. These algorithms fell into three groups firstly single flow or steepest descent, as described above. Secondly, multiple flow, which route the water to the lower neighbours according to the relative slopes as in equation 3.4, where *Qi* is the amount received by the lower neighbour *i* and *S* is slope between the cell in question and the neighbour *i.*

$$Qi = Q_o \frac{S_i^n}{\sum S_i^n} \qquad (3.4)$$

The third group are 'flux decomposition' where the water is distributed according to the angle of direction of the greatest slope. Desmet and Govers (1996) found that the multiple and flux methods were significantly different from the single flow algorithms, as the latter cannot account for divergent flow. Flux methods were superior by a small margin but more complicated and cannot split flow in more than two directions. The multiple routing method is similar to that used successfully by Quinn *et al*. (1991) and Freeman (1991) where an exponent of 1.1 is used, and both authors account for the difference in slope of diagonal neighbours by dividing the slope to these cells by 1.41 ($\sqrt{2}$). Interestingly, in their cellular model of braided rivers, Murray and Paola (1994) use a similar technique, yet with n = 0.5 - 1. Although Desmet and Govers (1996) work was based on calculating drainage area, this can have large impacts in modelling a catchment. Kirkby (1987) acknowledges that is may be a potential source of error in establishing the size and development of hillslope areas, for example tributaries or stream heads may form in the incorrect places. The lack of divergent flow also suppresses the formation of distributaries

(Kirkby, 1987). Whilst this may not be a problem in large catchments with dendritic networks, it is certainly inappropriate for areas of divergence such as alluvial fans. Furthermore, a single flow direction tends to suppress the migration of channels and will not allow the development of a multi-channelled or braided pattern.

The main reason for the dominance of single flow algorithms, is one of simplicity (Kirkby, 1987), as flow only has to be assigned one flow direction. This is compounded further, because in order to route flow in multiple directions, the model has to work from the highest cell to the lowest sequentially. This means that all the cells have to be sorted. Despite ingenious sorting algorithms, it is still a computationally expensive operation and this may be the main limiting factor, restricting the models to a maximum of 100 by 100 cells (Howard, 1994). Murray and Paola (1994) avoid this sorting by routing the water only to the three downstream cells, in effect pushing the water downstream. This is adequate for modelling a simple reach but cannot be applied to a more complex channel with tributaries or meanders flowing back against the main valley slope. Furthermore, they fail to account for the effect of flow depth, routing all the water around a cell even if it is only fractionally higher. They admit this deficiency, as water and sediment flow uphill in a real stream, yet this approximation is also thought to aid the generation of a braided pattern.

3.5.5.4 Validation of results

Validation of cellular models has proved difficult, and this has been partly linked to their lack of real scales. As a consequence, validation has mainly been qualitative, describing how the landscapes are similar, for example, "*this model captures the main spatial and temporal features of real braided rivers*" Murray and Paola (1994). Secondly, it reflects the lack of statistics available to characterise a landscape. In geomorphological studies, landscapes are generally described but rarely quantified. This shortcoming is clear from several of the previously mentioned studies, as authors look for a comparison statistic. Willgoose *et al.* (1994) mentions using a hypsometric curve but concentrates upon a slope/elevation against area graph. This demonstrates the relative importance of diffusive and fluvial processes and he compares his model outputs to statistics from a field site in Pokolbin. Howard (1994) also uses a hypsometric curve to show different states of catchment evolution, but additionally compares the relationship between drainage area and gradient favourably with that from channel head areas in the Tennessee Valley, California. These methods are all general, describing the overall form of the catchment, which is appropriate when investigating the topography of entire catchments, but more descriptive statistics are required for specific features.

Murray and Paola (1994) attack this issue by using two methods to validate the output of their model. Firstly they used a comparison of averaged geometric statistics, the number of channels, average widest channel width and variance of the number of channels. The plots show a good comparison, the model remaining statistically stable over time with geometric statistics matching those of real braided rivers. Secondly, they used a non-dimensional power spectrum of the simulated sediment load over time. These plots effectively show the pulsing bedload from a real and the model channel. They are non-dimensional, as the real river and model are operating over different time and space scales, but the graph demonstrates that the dynamics of the systems are similar.

However, validation remains a genuine difficulty, as we have little data to validate our models against. Field data are largely based upon point samples, with a very low spatial and temporal resolution. Even the topographic data in the form of a DEM are rarely at a resolution finer than 50m. Ironically, our landscape evolution models may give us more detail about the model landscape than we can presently quantify from the field.

3.5.6 Summary of CA models

CA models appear to offer several advantages over conventional methodologies and a comparison is shown in Table 3.1. However CA models are relatively untried and have only been applied to abstract situations to explore concepts, and not applied to landscapes. As a technique it offers much promise, in particular with respect to the routing method of Murray and Paola (1994) and the ability to incorporate many processes within one framework.

Conventional	Cellular Automaton
Frequently a series of linked cross sections or slope profiles.	Grid format, allowing easy transfer from surveyed topographic data and DTM to CA format
Sometimes complex implementation of several laws and equations	Simple. One set of rules is applied to all cells.
Difficult to integrate different processes acting at different spatial and temporal scales from hillslopes to channel.	Cellular format eases integration of varying processes, as all at same spatial scale.
Often site specific.	Simplicity and ease of DTM integration maintain a generic approach.
1 or 2 dimensional (sometimes linked) cross sections.	2 or 3 dimensions.
Fine scale <1m.	Medium scale >= 1m.
Representation of detailed processes especially within channel.	Poor representation of processes such as secondary circulation, but maintains a generic view.
Tried and tested.	Experimental.
Quantitative results.	Mainly qualitative with some quantitative results.

Table 3.1 Comparison of conventional and cellular automaton modelling techniques.

3.6 Sediment transport laws

In landscape evolution, fluvial transport is an important process. The river channel has been described as the *'Jerky conveyor belt'* (Ferguson *Pers. Comm.*), removing material eroded from the hillslopes. However, this description belies the complexities involved in fluvial sediment transport. Whilst we have a good understanding of the physics and concepts of pebble entrainment and motion, the complexity and diversity of river channels has led to a poor application of this knowledge to prediction. Therefore the selection and use of the correct sediment transport approximation will be vital to the success of a model.

Gomez and Church (1989) categorised sediment transport laws into four groups, discharge, tractive force, stochastic and stream power (Table 3.2).

Formula	Type
Meyer Peter (1934)	Discharge
Schoklitsch (1934)	Discharge
Schoklitsch (1943)	Discharge
Meyer-Peter and Muller (1948)	Tractive force
Parker (1982)	Tractive force
Du Boys-Straub (1935)	Tractive force
Einstein (1950)	Stochastic
Yalin (1963)	Tractive force
Ackers and White (1973)	Tractive force
Bagnold (1980)	Stream power

Table 3.2. Ten contemporary sediment transport equations from Gomez and Church, (1989).

These formulae were either developed from a theoretical background (e.g. Einstein, 1950 and Yalin 1963) or through direct field observation (e.g., Schoklitsch, 1934), but are all to some extent calibrated by comparison with field data. Many of these were compared directly to the Authors' own data, leading to an equation which may be single site specific or, at best, based upon a limited range of data. Furthermore, these formulae are generally robust when compared to their own data sets, but rarely so good when used on other data. Furthermore, as a consequence of the difficulty and expense in measuring bedload accurately, there is a shortage of data sets, and these are outnumbered by transport equations! The net result is that no one or group of formulae have been universally accepted.

Gomez and Church (1989) describe a comprehensive assessment of sediment transport formulae. They tested twelve formulae (Figure 3.4) against seven data sets. They conclude that in general situations Bagnold's formulae is best, but where detailed local hydraulic information is available, the Einstein and Parker methods deserve further attention.

Bathurst *et al.* (1987) examine six sediment transport equations in the context of a steep mountain rivers. This is an especially useful review, as all the data sets used by Gomez and Church (1989) were from lowland rivers with slopes of less than 0.01. Bathurst finds that most of the tractive force formulae break down with slopes greater than 0.1 and where relative submergence (d/D_{50}) is less than 10. He concludes that the Schoklitsch approach performs best, as it uses a discharge slope product instead of depth. However this equation is not well suited to representing a wide range of grain sizes.

With the exception of the Einstein (1950) equation, these formulae are all total load formulations, with bedload transport calculated for all grainsize fractions lumped together. If erosion and deposition for each grainsize is required these have to be derived from a distribution. The Einstein formulae, however, calculates a total bedload by summing the sediment transport for each fraction. This raises the issue of how well these total load formulae might model the transport of bimodal sediment distributions.

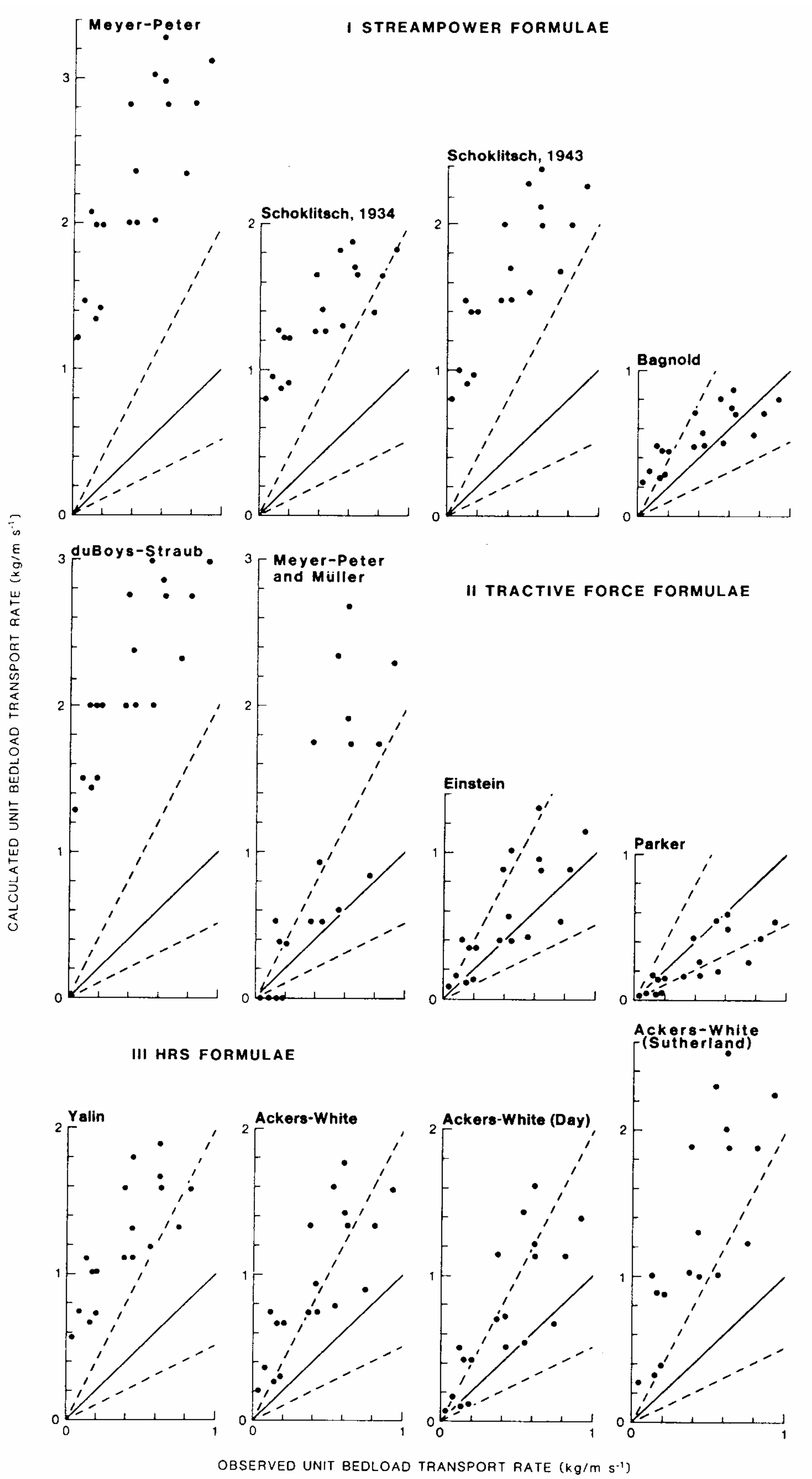

Figure 3.4. Performance of observed vs predicted sediment discharge rates for Elbow river, from Gomez and Church (1989).

The Einstein approach also differs from the others as it has no critical discharge for entrainment. All the other formulae have a threshold for the entrainment of a particle, which at first appears to be a logical conclusion. However, the distribution of sizes in a river-bed is a continuum, but for ease of measurement they are grouped into Phi size classes. If applied to a continuous range of sediment sizes, threshold formulae will have a gradual effect on sediment discharge. But if represented in Phi classes then each size fraction has a threshold below which there is no transport. This is a poor representation as even during low flows there is some sediment transport. For example if there were none how could there be slow incision through an alluvial fan? If the sediment transport and depth/slope product is plotted for the Einstein (1950) equation, a log-log relationship is shown. This closely mimics a threshold, with a sudden increase in erosion, yet there is always a small amount. Whilst in the short term this difference will be negligible, over the course of hundreds or thousands of years of river flow it may be important.

3.7 Grainsize modelling

Vital components in the dynamics of gravel bed rivers are interactions between the bedload and flow. One characterisitic observed in these channels is the development of a coarse armoured layer or pavement, thought to be formed by the selective removal of finer particles leaving the coarse lag behind. This layer is frequently coarser than the material below and consequently protects it from entrainment. This phenomenon however, means that if flow conditions allow this layer to be breached, a reservoir of fine material is made available for transport. This process led to the formulation of theories of equal mobility (Andrews, 1983) where material in the bed of a channel was observed to move (nearly) together, once a threshold flow had been reached. In developing a model with a strong fluvial element, the inclusion of different grainsize fractions is desirable to allow these armouring effects to develop as well as to study down-stream fining and the development of depositional bars.

There are other models which incorporate grainsize. The principal protagonist of this is Gary Parker who in 1990 and 1991 created a series of models to investigate the effects of grainsize on selective transport, equal mobility, downstream fining and abrasion. For this integration, Parker (1990, 1991a, 1991b) introduced the idea of the 'active layer', an area on the stream bed which acts as a transition zone between the river bedload and the underlying material. This approximation is based on field observation of gravel bed rivers, where the surface armour appears as a layer overlying the base material. Within the structure of these models, Parker (1991a, 1991b) and Hoey and Ferguson (1994) use this active layer as the interface for erosion and deposition. The active layer is divided into grainsize fractions according to Phi classes and material is eroded or deposited into this active layer in proportion to the relative power of the flow and the sediment transport laws applied. The thickness of the active layer is the subject of some debate. Hoey and Ferguson (1994) define the thickness as being 2 x D84, although Hassan (1990) and Hassan and Church (1992) suggest it is more a function of transport intensity. The thickness of this layer is therefore difficult to define. Hoey and Ferguson's (1994) approach of 2 x D84 would appear to integrate the effects of increasing transport rates and may be a fair approximation, as a bedload or active layer comprised of coarser particles would be thicker. However, from a modelling perspective, this may create unnecessary complexity as if the active layer is continually changing in composition, the calculation of its thickness becomes iterative.

The transfer of material between the active layer and the underlying bed material is also a point of contention and there is an absence of information on the physics of sediment deposition (Hoey and Ferguson, 1994). During deposition, material will be deposited into the active layer, and so from the active layer moved to the base. Conversely during periods of erosion, material will be integrated into the active layer from the base. This is shown in Figure 3.5.

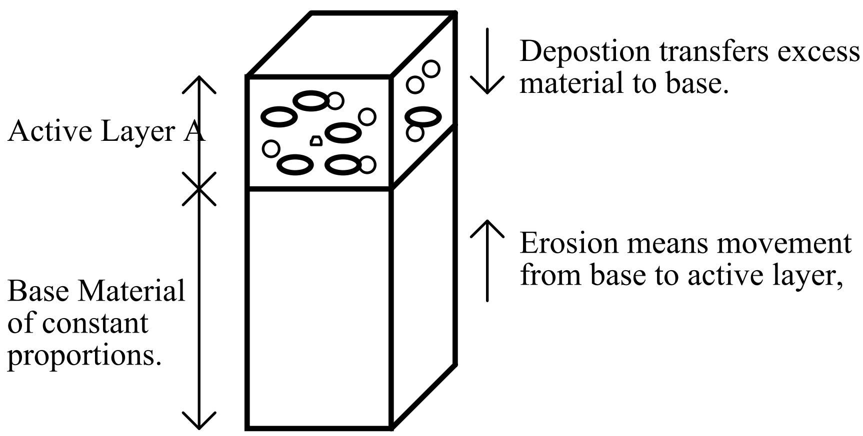

Figure 3.5. Diagram demonstrating the movement of sediment between active layers and the stream bed.

Hoey and Ferguson (1994) state that during degradation there is direct incorporation of material from the subsurface to the active layer, $E_i = f_i$ Where E_i is the proportion of material to be moved to the active layer, and f_i the proportion of sediment in the active layer for $_i$ grainsize class. During aggradation, the situation is less clear. Parker (1991a, 199b) discusses two situations where $E_i = F_i$ and $E_i = p_i$, where E_i is proportion to be removed from the active layer (exchange proportion) F_i the active layer proportion and p_i bedload or depositing proportion. Both these cases seem to correspond to differing scenarios. The former during slow aggradation and the latter during rapid, where for example an avalanche bar migrating downstream. Parker (1991a, 1991b) showed that in periods of continuous aggradation where $E_i = p_i$, this prevented downstream fining occurring. Hoey and Ferguson (1994) overcomes this by introducing an exchange parameter c to use a combination of these approaches (equation 3.5)

$$E_i = cF_i + (1-c)p_i \quad (3.5)$$

When c is set to 0.3, only a fraction is left in the active layer, the rest moved to the substrate. In recent physical and numerical modelling, Toro-Escobar *et al* (1996) supports this proportion. Their coefficient was however calculated from a flume experiment and is the result of several hours of degradation and aggradation over a 20m reach. There is no time scaling in this experiment which is needed to allow filtering through the active layer. For example, if a thick carpet of sediment were deposited upon the base of the channel, 70% of the material would not instantly be moved into the substrate, but would work its way down over time.

The number of active layers required is also uncertain. Cui *et al.* (1996) thought that at least three were required to simulate downstream fining, though others have used more (e.g. De Silvio, 1992). The number of active layers will influence the accuracy of the model, although this depends upon how vertically active the channel is. The sub surface layer or layers must be great enough to encompass the full range of channel incision or aggradation. If deposited material is re-worked then it is vital that this material is stored in its deposited grainsize constitution and not diffused by mixing throughout a single deep active layer.

3.8 Overview

Fluvial action is often the most dynamic and important process in the evolution of a basin. For this and other reasons, fluvial models, operating at a variety of scales, have taken precedence in geomorphology. As described previously in this chapter, these models range from three dimensional simulation of circulation surrounding a confluence, detailed two dimensional finite element grids of water surface profiles (Nicholas and Walling 1997, Bates *et al.* 1997) and the more 'classic' one dimensional approach of calculating over cross sections, such as HEC II. Many appear successful, but due to the complexity of solving the complex Navier-Stokes equations used, are computationally restricted to operating in a confined area. Furthermore, they fail to account for processes outside the study reach, such as mass movement, time varying discharge and changes in upstream sediment supply. Howard (1994), Tucker and Slingerland (1994), Willgoose *et al.* (1994) and Kirkby (1987) use a different approach, placing the emphasis on the slope processes. Howard (1994) and Tucker and Slingerland (1994) simplify channel operations to a sub grid cell process, with values for width and depth calculated using empirical relationships. This approach allows the aggradation and degradation of the channel, in the context of the whole basin, but does not allow the formation of terrace/flood plain stratigraphy, or differing channel patterns, all of which geomorphologists use to interpret past environmental change.

Whilst both river based and slope based approaches are fruitful, the former, hydraulic approach trades basin scale realism for local flood plain accuracy, whereas the latter sacrifices channel accuracy for realism at the basin scale. The split is due in part to the history of numerical flow modelling within engineering and partly to the choice of scale.

When modelling landscape evolution, there are numerous processes interacting through a wide range of time and space scales. These vary from the entrainment of a grain in a split second, to creep on a mountainside over thousands of years. Incorporating these small scale processes in a basin model is troublesome, because of these scale ranges. The computationally intensive nature of finite element methods makes their use impracticable over the long timescale that slope influences require

(>1000 years), and it is similarly impossible for them to provide models for the full spectrum of flood events. Furthermore, over the course of a flood, basins are spatially dynamic. Stream heads may extend and new tributaries and channels may form. For hydraulic modelling this creates numerous problems, as changes in bed/floodplain topography and spatial changes in the network require a frequent re-definition of the mesh of nodes used, which is highly time consuming, especially if a curvilinear approach is used.

Therefore, to meet the aims set out in chapter two, a methodology is needed that is capable of modelling all the relevant processes (slope, fluvial and hydrological) at a scale large enough to encompass an entire catchment yet fine enough to allow the reproduction of alluvial landforms. Furthermore this model should be capable of simulating catchment response to environmental changes over time scales ranging from a single flood event to the Holocene.