Global Sensitivity Analysis of Parameter Uncertainty in Landscape Evolution Models

Christopher J. Skinner[1], Tom J. Coulthard[1], Wolfgang Schwanghart[2], Marco J. Van De Wiel[3], and Greg Hancock[4]

[1]School of Environmental Sciences, University of Hull, Hull, UK

[2]Institute of Earth and Environmental Science, Potsdam University, Potsdam-Golm, Germany

[3]Centre for Agroecology, Water and Resilience, Coventry University, Coventry, UK

[4]University of Newcastle, Callaghan, Australia

Corresponding Author: C. J. Skinner (c.skinner@hull.ac.uk)

Abstract

The evaluation and verification of Landscape Evolution Models (LEMs) has long been limited by a lack of suitable observational data and statistical measures which can fully capture the complexity of landscape changes. This lack of data limits the use of objective function based evaluation prolific in other modelling fields, and restricts the application of sensitivity analyses in the models and consequential the assessment of model uncertainties. To overcome this deficiency, a novel model function approach has been developed, with each model function representing an aspect of model behaviour, which allows for the application of sensitivity analyses. The model function approach is used to assess the relative sensitivity of the CAESAR-Lisflood LEM to a set of model parameters by applying the Morris Method sensitivity analysis for two contrasting catchments. The test revealed that for both catchments the model was most sensitive to the choice of the sediment transport formula, and that each parameter influenced model behaviours differently, with model functions relating to internal geomorphic changes responding in a different way to those relating to the sediment yields from the catchment outlet. The model functions proved useful for providing a way of evaluating the sensitivity of LEMs in the absence of data and methods for an objective function approach.

1. Introduction

Landscape Evolution Models (LEMs) investigate how the Earth's surface evolves over timescales ranging from hundreds to millions of years (Coulthard and Van De Wiel, 2012; Martin and Church, 2004; Pazzaglia, 2003; Tucker and Hancock, 2010; Van De Wiel et al., 2011). They represent the earth's surface with a regular or irregular mesh and simulate how the surface evolves over time as a function of tectonic processes, and erosion and deposition from Earth surface processes. LEMs have proved to be very useful scientific tools to understand how Earth surface processes interact to shape the landscape.

More recently, LEMs have improved considerably in their ability to simulate the physical environment, and this has developed in parallel with improvements in computational efficiency and power. This allows LEMs to go beyond highly simplified models of landform development and to also incorporate increasingly complex processes such as pedogenesis (Vanwalleghem et al., 2013; Welivitiya et al., 2016) and periglacial processes (Andersen et al., 2015; Egholm et al., 2015). Other processes are now being handled in more detail such as hydrodynamic flow models and aeolian processes (Adams et al., 2017; Coulthard et al., 2013; Liu and Coulthard, 2017). These developments led to Coulthard et al. (2013) describing them as 'second generation' LEMs that extend previously explanatory and explorative models to be used for prediction of future changes in landscapes, such as for the mining industry (e.g., Hancock et al., 2017; Saynor et al., 2012).

However, more detailed physical representations of the processes that shape the Earth's surface involve a larger number of parameters that are typically estimated from proxy data or theoretical considerations, or are completely unknown (Oreskes et al., 1994; Petersen, 2012). If LEMs are to be operationally used for prediction or as decision-making tools in the future, their outputs must be evaluated against the uncertainty in input parameters – a task that is increasingly difficult for a large number of parameters. Through sensitivity analysis (SA) investigates how variations in the output of a numerical model can be attributed to its input factors (Pianosi et al., 2016). This is useful for identifying key parameters for later calibration but this has rarely been conducted for LEMs. The aim of this study is thus to conduct a SA of the widely used and highly parameterized LEM CAESAR-Lisflood (Coulthard et al., 2013) - in particular, we wish to be able to detect the parameters that have the greatest influence on the model's simulation output. As model sensitivity may be influenced by different landscapes, we run the SA in two individual and distinct catchments.

1.1 Sensitivity Analysis and Landscape Evolution Models

The application of SA in environmental modelling has a history spanning four decades (Norton, 2008) and forms an important component of using models for decision-making, including model development, calibration and uncertainty analysis (Yang, 2011). SA addresses five key questions (Cariboni et al., 2007; Neumann, 2012; Song et al., 2012, 2015):

1. Which parameters have the greatest influence on the model?
2. If additional data could be used to reduce the uncertainty in a parameter, which would most reduce the model output variance?
3. Are there parameters with such low influence that their values could be fixed without impact on the model outputs?
4. If parameter values emerge as incorrect, how will they influence model outputs?
5. Which parameters influence model outputs in different regions (parameter space)?

Clearly, based on the above, an appraisal of model sensitivity is important to fully understand and apply model results. In a review of applications of SA in environmental models, Yang (2011) identified two common approaches to SA – local and global. Local SA are limited, considering only the impacts of factors on model outputs locally, i.e., within a restricted region of the model's parameter space, whilst global SA typically utilise Monte-Carlo methods to assess the sensitivity of impacts across the whole parameter space (Yang, 2011). For complex models with non-linear behaviours, the use of Local SA can be highly biased as they neglect the non-linear interactions between parameters (Oakley and O'Hagan, 2004; Pappenberger et al., 2006; Yang, 2011). Global SA are more computationally expensive, but as the methods are more reliable, they are attractive to modellers (Yang, 2011).

The use of SA as a routine component of model assessment and calibration is common place in climatic, meteorological, hydrological, hydraulic and many other modelling fields. However, for LEMs there are surprisingly few examples of SA being carried out. This can be explained by three inter-related issues: (i) LEMs typically have a large number of model parameters; (ii) long model run times can make multiple simulations for SA impractical; and (iii) model behaviour can be highly non-linear (e.g., Coulthard and Van De Wiel, 2007; Larsen et al., 2014; Van De Wiel and Coulthard, 2010), leading to potentially complex SA interpretations. Large numbers of model parameters and long run times, in particular, make Monte-Carlo methods extremely time consuming – and therefore often unviable.

There are several studies on how LEMs respond to variable forcing, process changes and model parameters, including changes in climate variability and precipitation resolution (Armitage et al., 2017; Coulthard and Skinner, 2016a; Ijjasz-Vasquez et al., 1992; Tucker and Bras, 2000), channel widths (Attal et al., 2008), vegetation (Collins, 2004; Istanbulluoglu and Bras, 2005), and variations in initial conditions (Hancock, 2006; Hancock et al., 2016; Ijjasz-Vasquez et al., 1992; Willgoose et al., 2003). Campforts et al. (2017) investigated how different numerical solvers affect LEM simulation. Yet few studies explicitly perform SA and most of the applications described above are exploring LEM sensitivity to processes, or changes in environmental conditions, and are more correctly referred to as exploratory tests (Larsen et al., 2014). On the other hand, investigations to ascertain the model's response to potential uncertainties (e.g from model parameterisation) can be deemed as true SA (eg, Armitage et al., 2017; Coulthard and Skinner, 2016a; Hancock et al., 2016).

Hydrological models faced similar issues to LEMs in the past, i.e., model complexity and long processing times when applying SA. To overcome them, hydrologists have used the Morris Method (MM; Morris, 1991). The MM can be regarded as a global SA, although it actually performs multiple local SAs sampled from across the full parameter space – this produces a series of local evaluations, the mean of which is an approximation of the global variance (van Griensven et al., 2006; Norton, 2009; Saltelli et al., 2000). The main strength of the MM is its computational efficiency. Herman et al. (2013) showed that the MM could estimate similar variance in model outputs to the Sobol' Variance-based global SA method (Sobol', 2001), yet required 300 times less evaluations, and significant less data storage for an application to a distributed catchment hydrological model. The robustness of this approach has been further shown by numerous workers (e.g., Brockmann and Morgenroth, 2007; Pappenberger et al., 2008; Yang, 2011). However, the MM cannot provide a full quantitative assessment of parameter sensitivity and is dependent upon the user-defined bounds to the parameter space. It can successfully rank parameters between the least and most influential to model outputs, but cannot determine parameters' exact relative influence (Brockmann and Morgenroth, 2007). These advantages and limitations entail that MM has primarily been used during the pre-screening stage of models, isolating the most influential parameters for further SA with quantitative, yet more computationally expensive, methods (e.g., Ratto et al., 2007; Song et al., 2015; Yang, 2011; Ziliani et al., 2013).

(Ziliani et al., 2013) performed a two-stage SA for the CAESAR LEM, utilising the MM (as adapted by Campolongo et al., 2007). Whilst this study demonstrated the feasibility of applying the MM as a global SA to a reach-scale LEM, it was applied as a pre-screening stage to identify the most relevant parameters for model calibration. In contrast, our study focuses on SA as a tool to investigate parameter influence on model behaviour.

1.2 Metrics for Landscape Evolution Model Assessment

Evaluating LEMs is challenged by the paucity of comprehensive field data against which they can be assessed and the lack of measures for calibration and validation (Hancock et al., 2016; Hancock and Willgoose, 2001; Tucker and Hancock, 2010). Moreover, some LEMs (e.g., CAESAR-Lisflood) simulate short (annual to decadal) and long-term (millennial time scales and longer) landscape changes, necessitating data and methods to assess them across variable time scales. Thus, while SA of environmental models often rely on objective functions (e.g., the Nash-Sutcliffe score between observed and simulated values; Nash and Sutcliffe, 1970), this approach is generally not practical for LEMs. With few exceptions (e.g., Ziliani et al., 2013), results from LEMs are therefore frequently assessed qualitatively, relying on visual interpretation of the simulated landforms or cross-section profiles (e.g., Coulthard and Skinner, 2016b; Hancock et al., 2010, 2015; Hancock and Coulthard, 2012).

Catchment outlet statistics, such as sediment yield time series, allow for comparison between simulations to indicate a catchment's response to perturbations (e.g., Coulthard et al., 2012; Coulthard and Skinner, 2016b; Hancock and Coulthard, 2012). However, sediment yield time series rarely provide a sufficiently complete picture of a catchment's geomorphic response. For example, Coulthard and Skinner (2016b) showed that simulations calibrated to provide equivalent sediment yields produced different landforms. For planning purposes these internal catchment changes are likely to be more useful than catchment sediment yields. Moreover, changing topography potentially instigates a feedback process that leads to complex, often non-linear catchment behaviour (Coulthard and Van De Wiel, 2007, 2013; Hancock et al., 2016; Jerolmack and Paola, 2010; Van De Wiel and Coulthard, 2010).

Finally, the spatially and temporally heterogeneous response of erosion and deposition patterns in LEMs also makes "pixel-to-pixel" comparisons difficult. For example, in a valley reach, gross patterns of erosion and deposition may be identical but with the channel on the other side of the valley – yielding a poor pixel-to-pixel comparison.

Few studies have tested metrics to compare topographic data or physical experiments to simulated elevation changes by LEM (Hancock et al., 2010, 2011; Hancock and Willgoose, 2001; Ibbitt et al., 1999). However, although the metrics often suggested a good agreement, visual analysis of the final DEMs indicated clear differences between the physical models and the simulations (Hancock and Willgoose, 2001). There is, therefore, a clear need for better statistical methods for critically evaluating and comparing landscapes that can also be used for evaluating the accuracy (or otherwise) of LEMs.

The paucity of observational data and the lack of measures that amalgamate the complexity of spatio-temporal landscape change into a single metric have prevented the objective function approach to be common in modelling landscape evolution. Instead, LEMs can be evaluated by observing the changes in model outputs reflective of model behaviour – these model functions can be used in lieu of objective functions to allow the sensitivity of LEMs to be assessed. Model functions would be best used as a set in combination to allow assessment across a range of model behaviours, and would also be transferable across a range of catchments. Such an approach formalises existing methods of evaluating LEM outputs and provides a framework from which multi-criteria objective function approaches can be applied when suitable observation become available.

1.3 A Global SA for a catchment LEM

This study uses MM to assess the sensitivity of CAESAR-Lisflood to a range of user-defined parameters, and therefore demonstrates the first application a global SA to a catchment LEM. We selected 15 model parameters (here we consider the choice of sediment transport formula as parameter)either because of their known importance to the model or because the model's response to the parameter is presently poorly understood. Although not all the 15 model parameters are universal between LEMs, many LEMs have equivalents. Moreover, we developed a set of 15 model functions that reflect core behavioural responses of the model. These will indicate whether the same parameters influence all behaviours, or whether the different behaviours respond to different parameters. The choice of 15 model parameters and 15 model functions is coincidental. We conducted the SA in two catchments with contrasting environmental settings to assess how transferable an individual SA is to different conditions.

It is important to state that this study is an illustration of the potential for using the MM to inform an operator of how model parameter choices can impact the performance and behaviour of their model. It is not an attempt to reproduce or calibrate the CAESAR-Lisflood model to real-world observations, although the model has been applied to each catchment previously.

2. Methods

We apply the MM to perform a global SA on the CAESAR-Lisflood model for two contrasting catchments (more detail in Section 2.3): the Upper Swale, UK (181 km^2, temperate, perennial), and Tin Camp Creek, Australia (0.5 km^2, tropical, ephemeral). Each individual simulation runs for a 30 year period, where the first 10 years are used as a spin-up to reduce the impacts of transient model behaviour and therefore output analysis starts after year 10 of the simulation. The CAESAR-Lisflood model is used in catchment mode, the simulations have no representation of suspended sediments and bed rock, and the dune and soil evolution modules are not used. Form drag is not directly considered within the model but is reflected within the setting of the Manning's n Roughness

Coefficient. For each catchment, we assess the 15 user-defined parameters against a set of 15 model functions. Finally, we also assess the changes in elevations across different sections of the catchments.

For clarity, we here define some terms used frequently throughout this manuscript:

- Parameter – Adjustable value within a model. The value is determined during model set-up and remains constant throughout a given simulation. The value is often based on recorded values or adjusted during calibration.
- Objective function – an error score between model outputs and observations used to evaluate model performance.
- Model function – a measure derived from model outputs used to evaluate model behaviour in lieu of an adequate objective function.
- Elementary effect (EE) – a value used as part of the Morris Method, indicating the change in function value (objective or model) resulting from a change of parameter value during a single repeat.
- Main effect (ME) – the mean of the elementary effects from all repeats, for a specified parameter and a specified function.

2.1 CAESAR-Lisflood

The LEM used is the CAESAR-Lisflood model (Coulthard et al., 2013). CAESAR-Lisflood is a second generation LEM, capable of simulations with greater physical realism than first generation models but also with increased complexity – the model features a large number of fixed, physically-based, or user-defined parameters. This additional complexity may result in an increased non-linearity and sensitivity to model parameters. We used CAESAR-Lisflood v1.8, without any additional modifications to the model's functionality from the version freely available online.

A full description of the CAESAR-Lisflood model can be found in Coulthard et al. (2013), and its core functionality is only summarised here. The model utilises an initial DEM built from a regular grid of cells, and in the catchment mode (as used in this model set up) is driven by a rainfall timeseries, which can be lumped or spatially distributed (Coulthard and Skinner, 2016b). At each timestep the rainfall input is converted to surface runoff using TOPMODEL (Beven and Kirkby, 1979), and distributed across the catchment and routed using the Lisflood-FP component (Bates et al., 2010). The CAESAR component of the model drives the landscape development using sediment transport formulae based on flow depths and velocities derived from the Lisflood-FP component. Bed load is distributed to neighbouring cells proportionally based on relative bed elevations. This study has not used the suspended sediment processes in the model. The model can handle nine different grain sizes, and information is stored in surface and sub-surface layers where only the top surface layer is 'active' for erosion and deposition. A comprehensive description of this process can be found in Van De Wiel et al., 2007).

CAESAR-Lisflood is freely available and since 1996 there have been over 60 published studies using the model over a wide range of temporal and spatial scales (Skinner and Coulthard, 2017). These previous studies provide useful background into model parameter interactions helping to inform the choice of the user-defined parameters used for the SA as described in Section 2.4. Some studies have also investigated the model's sensitivities to external factors - for example, Coulthard and Skinner (2016) investigated the sensitivity of the CAESAR-Lisflood model to the spatial and temporal resolution of precipitation. Other studies have investigated the influence of individual processes or forcings. For example, Coulthard and Van De Wiel (2017) examined how land-use influences the outputs of the model.

2.2 Morris Method

Our study used the MM described in Ziliani et al. (2013), i.e., the original MM of Morris (1991), as extended by Campolongo et al. (2007), and applied the “sensitivity” package in the R Statistical Environment (Pujol, 2009) to generate the parameter sets for the SA.

To set up the MM we selected a number of parameters to be assessed, specifying a minimum and maximum range for each, plus a number of iterative steps. The parameter values are equally spaced based on the range and number of steps – for example, a parameter with a range of 2 to 10 and 5 iterative steps would have available values of 2, 4, 6, 8, and 10. This is done for each parameter and, where possible, the same number of iterative steps was used for each.

The MM samples the global parameter space by performing multiple local SAs referred to as repeats. The first simulation in each repeat is made up of a randomly assigned selection of parameter values from the available values. To set up the second simulation in the repeat a single parameter is randomly selected and its value changed by a random number of iterative steps – if we use the example above, if simulation 1 used the value 4, changing this to 2 or 6 would be one iterative step change (where one step is a change in value of 2), to 8 would be two steps, and using 10 would be three steps. For simulation 3 in the repeat another randomly selected parameter is changed although previously changed parameters are no longer available to be selected. This is continued until no further parameters are available to be changed, therefore in our study each repeat contains 16 tests – 1 starting set of parameters, plus 15 parameter changes. In this study we have used 100 repeats, for a total of 1600 individual simulations – for comparison, the implementation of the MM by Ziliani et al. (2013) used 10 repeats.

The sensitivity of the model to changes in parameter values is evaluated by the changes of objective function values between sequential tests within repeats relative to the number of incremental steps the parameter value has been changed by. The change in objective function score between two sequential tests divided by the number of incremental step changes is an elementary effect (EE) of that objective function and the parameter changed, as shown by Equation 1 -

Equation 1

$$d_{ij} = \left| \frac{y(x_1 x_2 \ldots, x_{i-1}, x_i + \Delta_{i,} x_{i+1}, \ldots, x_k) - y\,(x_1 x_2 \ldots, x_{i-1}, x_{i,} x_{i+1}, \ldots, x_k)}{\Delta_i} \right|$$

Here d_{ij} is the value of the j^{th} EE ($j = 1, \ldots, r$; where *r* is the number of repetitions (here *r* = 100)) of the i^{th} parameter (e.g., *i* =1 refers to sediment transport formula, see Table 1), x_i is the value of the i^{th} parameter, *k* is the number of parameters investigated (here 15), $y(x_1, x_2, \ldots, x_k)$ is the value of the selected objective function, and Δ_i is the change in incremental steps parameter *i* was altered.

After all 1600 tests have been performed, the main effect (ME) for each objective function and parameter is calculated from the mean of the relevant EEs – the higher the ME the greater the model's sensitivity. Alongside the ME, the standard deviation of the EEs is also calculated as this provides an indication of the non-linearity within the model.

2.3 Study Basins

2.3.1. Upper Swale, UK

The Swale catchment, UK, is a medium sized basin (181 km^2) with 500 m of relief (Figure 1). It has been used extensively in previous CAESAR/CAESAR-Lisflood applications (Coulthard et al., 2012; Coulthard and Macklin, 2001; Coulthard and Skinner, 2016a; Coulthard and Van De Wiel, 2013). For this SA, it represents a medium basin in a temperate climate. All simulations on the Swale are use a 50 m resolution DEM based on airborne LiDAR. Precipitation inputs are 10 years of NIMROD composite

RADAR rainfall estimates (Met Office, 2003), applied at a 1 h temporal and 5 km spatial resolution, and repeated three times for a 30 year timeseries.

2.3.2. Tin Camp Creek, Australia

The Tin Camp Creek catchment is a small sub-catchment (0.5 km^2) of the full Tin Camp Creek system (Hancock et al., 2010; Hancock, 2006) (Figure 1). The basin has 45 m of relief and is in the tropical region of the Northern Territory, Australia. In contrast to the Swale, Tin Camp Creek is a small basin and the region has pronounced wet and dry seasons, with short intense rainstorms a feature of wet season precipitation. The DEM is at 10 m grid cell resolution produced from high resolution digital photogrammetry (Hancock, 2012).The rainfall input is taken from observations from a single raingauge at Jabiru Airport, providing a 1 h – lumped (single catchment-average) resolution timeseries for 23 years, with the first 7 years repeated to produce a continuous 30 year timeseries..

2.3.2 Stream Orders

The changes in the mean elevation across different areas of the catchments were assessed as an illustration of spatial differences in geomorphic change. Each basin was sub-divided into regions corresponding to the watersheds of five stream orders based on the proportion of the catchment drained in the initial DEM – 1^{st} ≤1 %; 2^{nd} ≥1 %; 3^{rd} ≥10 %; 4^{th} ≥25 %; 5^{th} ≥50 % (see Figure 1). This method is novel and was developed to provide a consistent method of sub-dividing both catchments independent of factors such as connectivity and DEM resolution.

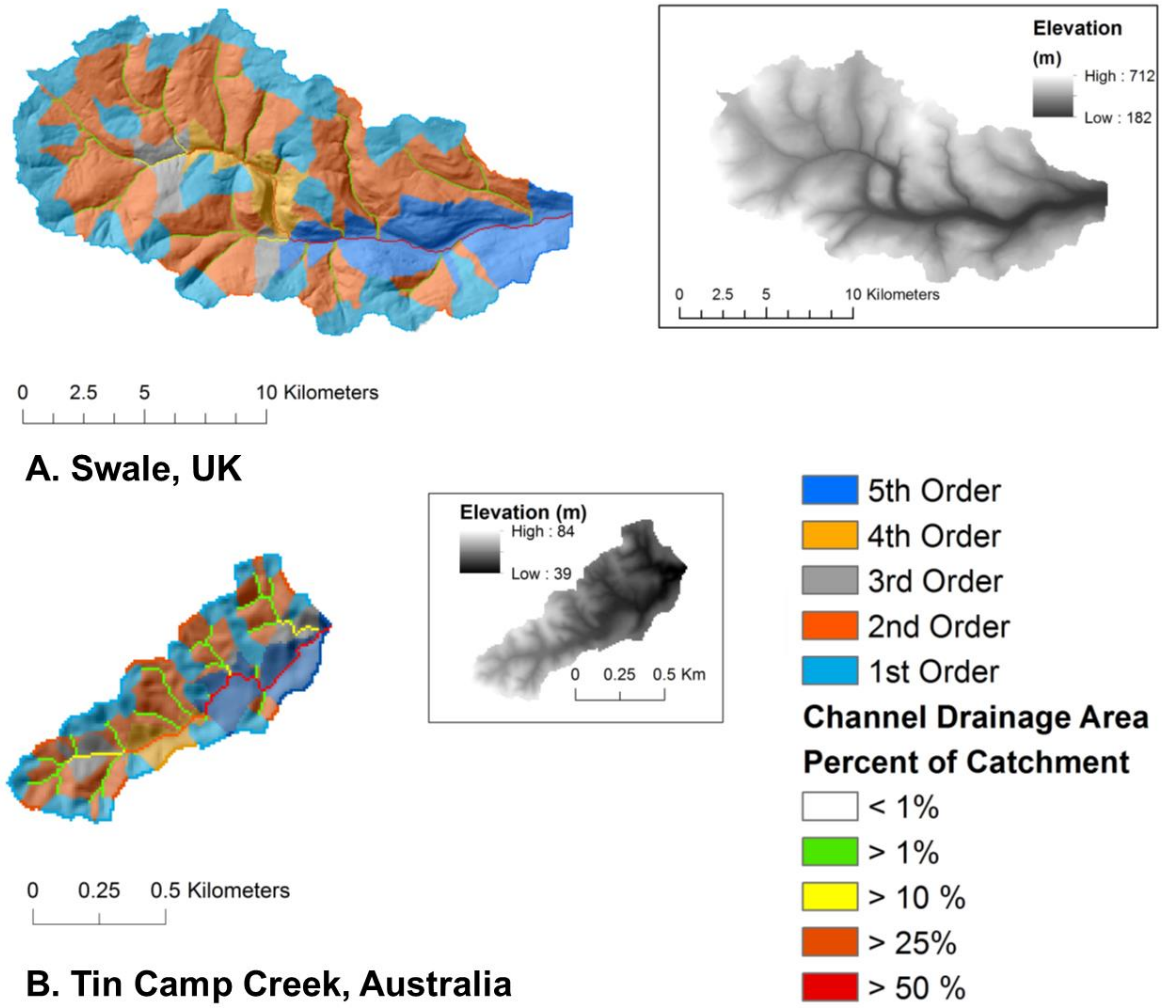


Figure 1 – Elevation map for the Upper Swale catchment, UK (top), and Tin Camp Creek catchment, Australia (bottom) – note the differences in scale between catchments. Each catchment is sub-divided into watersheds of five stream orders based on the proportion of the catchment drained.

2.4 User-Defined Parameters

The MM implemented here used 15 user-defined parameters, each with 5 iterative step values (as described in Section 2.2). The only exception was the choice of sediment transport formula parameter (SED, Table 1) where only two options are available. The parameters, their ranges, and available values are shown in Table 1.

Table 1 – User-defined parameters used and the min-max values for the two study catchments.

Code	Parameter	Steps	Upper Swale	Tin Camp Creek
(1) SED	Sediment Transport Formula	2	1 Wilcock & Crowe / 2 Einstein	1 Wilcock & Crowe / 2 Einstein
(2) MEL	Max Erode Limit (m)	5	0.01; 0.015; 0.02; 0.025; 0.03	0.001; 0.0015; 0.002; 0.0025; 0.003
(3) CLR	In Channel Lateral Erosion Rate	5	10; 15; 20; 25; 30	10; 15; 20; 25; 30
(4) LAT	Lateral Erosion Rate	5	$2.5e^{-6}$; $3.75e^{-6}$; $5e^{-6}$; $6.25e^{-6}$; $7.5e^{-6}$	$1.5e^{-6}$; $2.25e^{-6}$; $3e^{-6}$; $3.75e^{-6}$; $4.5e^{-6}$
(5) VEG	Vegetation Critical Shear Stress (Pa)	5	10; 15; 20; 25; 30	2; 3.25; 4.5; 5.75; 7
(6) MAT	Grass Maturity Rate (yr)	5	0.5; 0.75; 1; 1.25; 1.5	0.5; 0.875; 1.25; 1.625; 2
(7) SCR	Soil Creep Rate (m/yr)	5	0.00125; 0.001875; 0.0025; 0.003125; 0.00375	0.00125; 0.001875; 0.0025; 0.003125; 0.00375
(8) SFT	Slope Failure Threshold (°)	5	40; 42.5; 45; 47.5; 50	40; 42.5; 45; 47.5; 50
(9) IOD	In/Out Difference ($m^3.s^{-1}$)	5	2.5; 3.75; 5; 6.25; 7.5	0.1; 0.175; 0.25; 0.325; 0.4
(10) MinQ	Min Q Value (m)	5	0.25; 0.375; 0.5; 0.625; 0.75	0.025; 0.0375; 0.05; 0.0625; 0.075
(11) MaxQ	Max Q Value (m)	5	2.5; 3.75; 5; 6.25; 7.5	2.5; 3.75; 5; 6.25; 7.5
(12) SEC	Slope for Edge Cells	5	0.0025; 0.00375; 0.005; 0.00625; 0.0075	0.0025; 0.00375; 0.005; 0.00625; 0.0075
(13) EVR	Evaporation Rate (m/d)	5	0.00067; 0.001005; 0.00134; 0.001675; 0.00201	0.0025; 0.004375; 0.00625; 0.008125; 0.01
(14) MNR	Manning's n Roughness	5	0.03; 0.035; 0.04; 0.045; 0.05	0.03; 0.0325; 0.035; 0.0375; 0.04
(15) GSS	Grain Size Set	5	Set 1; Set 2; Set 3; Set 4; Set 5	Set 1; Set 2; Set 3; Set 4; Set 5

The MM varies the value of each parameter tested once per repeat, and here we use 100 repeats. Therefore, careful consideration was required in the selection of parameters as each parameter tested added 100 model runs to the test – there are 49 user-defined parameters in the version of CAESAR-Lisflood model used (v1.8), and even excluding parameters associated with dune and soil development, there are still 35 user-defined parameters. To test each would require 3600 model runs for each catchment, yet the inclusion of some parameters is likely to add little value. Thus this was narrowed to a set of 15 user-defined parameters (Table 1) with the selection based largely on prior knowledge of the importance of these parameters, or due to a lack of previous knowledge of the influence of the parameters on the model – full justification of the selection of parameters, and descriptions of their purpose within the model, can be found in Supplementary Material S1.

The MM is subjective in that the relative sensitivities shown depend on the minimum and maximum range values set by the user. Therefore, it is necessary to set each parameter's range to be broadly equal to the others in order to obtain useful information. To be consistent, where possible we have used a default value taken from past calibrations and varied this by +/- 25 % and +/- 50 %. There are some instances where this was not appropriate and a minimum and maximum bound was set instead, with 5 iterative steps of equal distance determined (for example, the Manning's n Roughness for Tin Camp Creek where +/- 50 % would have resulted in obviously physically unrealistic values – see Table 1 for values used).

Here we have considered the selection of sediment transport formula as a parameter despite doing so is to change the functional form of the model. For clarity, and in line with how the choice is presented within the Graphical User Interface of the model, we will henceforth consider this choice in the same way as a parameter. The sediment transport formulae employed for SED were Einstein (derived for sand-bed rivers) (Einstein, 1950) and Wilcock & Crowe (formulated on sediment ranges between 0.5 and 64 mm) (Wilcock and Crowe, 2003). These were not selected as representing the best fit for the catchments simulated but because they are the formulae available in the unmodified version of CAESAR-Lisflood. The sediment transport formulae parameter was applied as a binary choice, with the model switching from one formula to the other once per repeat (no other parameter values were varied when this occurs, as per the description of the MM in Section 2.2). It was assumed that this change constituted a single iterative step change for calculating related EEs.

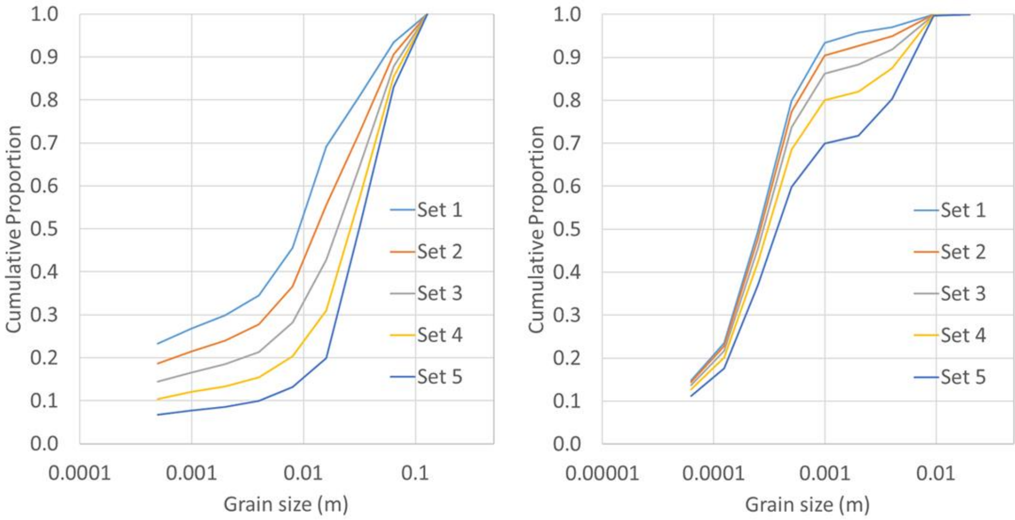


Figure 2 – Sediment grain size distribution sets for the Upper Swale (left) and Tin Camp Creek (right), showing the cumulative proportions.

Grain size distribution has been shown to influence erosion patterns and erosion rate(Hancock and Coulthard, 2012)). It is more difficult to define iterative steps for the sediment grain size sets which include 9 different grain sizes and proportions in each. Instead, these were skewed by altering the proportions of the five smallest grain sizes +/- 25 % and 50 %, and the opposite to the four largest grain sizes, before adjusting the final proportions to equal one based on the relative values. This produces two sets biased for smaller grain sizes (Sets 1 and 2), and two sets biased for larger grain sizes (Sets 4 and 5), as well as the default grain size set (Set 3) (Figure2). Note, that the grain size sets presented in Figure 2 contain non-cohesive silts and this requires an extrapolation of the sediment transport formulae (Van De Wiel et al., 2007).

2.5 Model Functions

The common method of assessing a model's sensitivity to parameters values via SA, and the method employed by the MM, is to observe the variations to objective function measures. However, the difficulties in applying an objective function approach to LEMs were highlighted in Section 1.2, and in order to apply an SA a novel approach is required. The method we have developed eschews the objective function approach and instead assesses the model against a series of model functions designed to reflect some of the core behaviours displayed in the model – these can be seen in Table 2. This represents a philosophical difference to traditional applications of SA – here we are not testing the model against its skill in simulating the physical environment, but rather how the model responds behaviourally to changes in the user-defined parameters detailed in Section 2.4. The 15 model functions (Table 2) are simple, scalable and transferable between different catchment types, and can be applied to simulations of different timeframes. The model functions are based on outputs which are not unique to CAESAR-Lisflood, so can be applied to other LEM and geomorphic models

Table 2 – Model Functions and the associated core behaviours.

Model Function	Core Behaviour
Total Sediment Yield (m^3)	Catchment Sediment Yield
Mean Daily Sediment Yield (m^3)	
Peak Daily Sediment Yield (m^3)	
Time to Peak Sediment Yield (s)	
Days when Sediment Yield > Baseline (d)	
Total Net Erosion (m^3)	Internal Geomorphology
Total Net Deposition (m^3)	
Area with > 0.02 m Erosion (m^2)	
Area with > 0.02 m Deposition (m^2)	
Total Discharge (m^3)	Catchment Discharge
Mean Daily Discharge (m^3)	
Peak Daily Discharge (m^3)	
Time to Peak Discharge (s)	
Days when Discharge > Baseline (d)	
Total Model Iterations (calculations)	Model Efficiency

The model functions were applied to the MM as described in Section 2.2, substituting the model functions in place of the objective functions with no further changes to the method. Model function values were calculated at the end of each simulation.

To summarise the large amount of information produced, the ME of each parameter and model function combination was normalised based on the proportion of the ME for highest ranking parameter for that model function – therefore the highest ranked parameter for each model function always scored 1. The scores for each parameter were aggregated for across all model functions based on the mean of the scores. The model functions were sub-divided into core behaviour groups (Table 2), and the scores aggregated again for each core behaviour. The same was also done, separately, for the standard deviations of each parameter and model function.

3. Results

3.1 All Model Functions

Figure 3 shows the spread of parameter influence for both catchments, where a higher mean of the aggregated MEs indicates greater sensitivity in the model to that parameter, and ahigher standard deviation shows greater non-linearity when interacting with other parameters. Table 3 shows the parameters ranked for both catchments, based on the aggregated mean ME values. The most influential parameter is SED (see Table 1 for full description of parameter abreviations), ranked top for both catchments and also being most influential by a reasonable margin, having an aggregated mean of at least 0.2 higher than the 2nd ranked parameter. Other parameters, such as VEG, IOD, MNR, MinQ and GSS, rank highly or mid-range. There is a visually close correlation between the most influential parameters and those that display the most non-linearity (Figure 3).

Table 3 – Parameters ranked by means for each catchment from the aggregated scores for all Elementary Effects. SED = sediment transport formula; MEL = maximum erode limit; CLR = in channel lateral erosion rate; LAT = lateral erosion rate; VEG = vegetation critical shear stress; MAT = grass maturity rate; SCR = soil creep

rate; SFT = slope failure threshold; IOD in/out difference; MinQ = minimum Q value; MaxQ = maximum Q value; SEC = slope for edge cells; EVR = evaporation rate; MNR = Manning's n roughness coefficient; and GSS = grain size set.

Rank (by mean: 1 = most influential)	**Upper Swale**	**Tin Camp Creek**
1	SED	SED
2	MNR	SEC
3	IOD	VEG
4	GSS	GSS
5	EVR	MinQ
6	VEG	IOD
7	MinQ	MNR
8	LAT	MAT
9	CLR	SCR
10	SCR	MEL
11	SEC	LAT
12	MAT	CLR
13	MEL	MaxQ
14	MaxQ	SFT
15	SFT	EVR

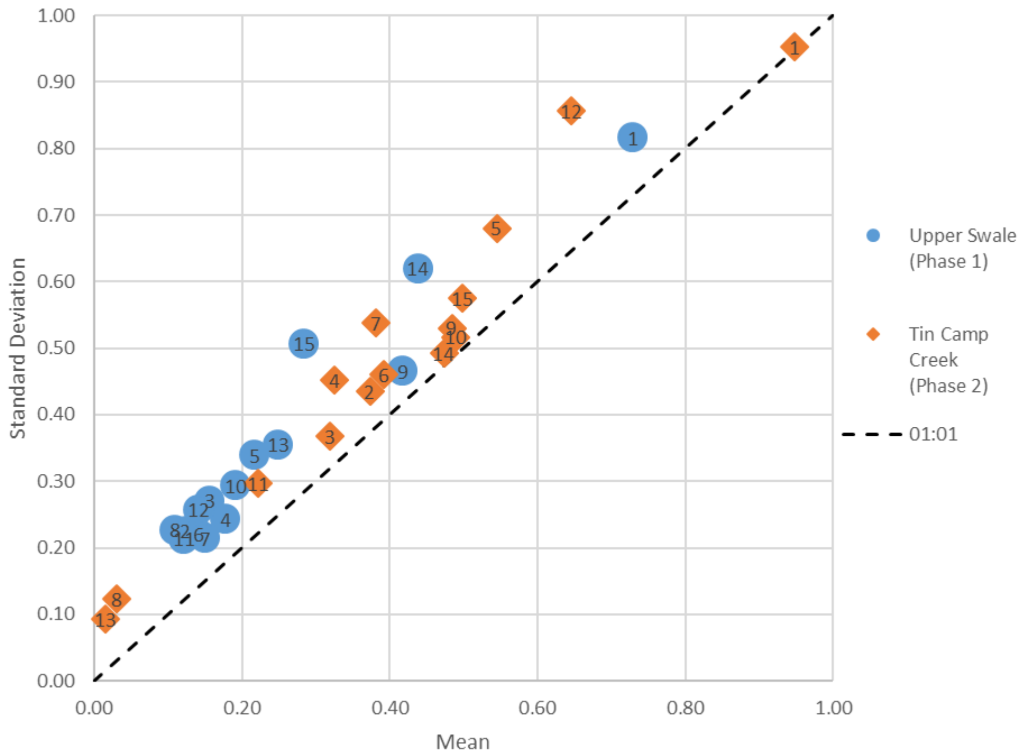


Figure 3 – Aggregated scores for all Elementary Effects where: 1 = sediment transport formula (SED); 2 = maximum erode limit (MEL); 3 = in channel lateral erosion rate (CLR); 4 = lateral erosion rate (LAT); 5 = critical vegetation shear stress (VEG); 6 = grass maturity rate (MAT); 7 = soil creep rate (SCR); 8 = slope failure threshold (SFT); 9 = in/out difference (IOD); 10 = minimum Q value (MinQ); 11 = maximum Q value (MaxQ); 12 = slope for edge cells (SEC); 13 = evaporation rate (EVR); 14 = Manning's n roughness coefficient (MNR); and 15 = grain size set (GSS).

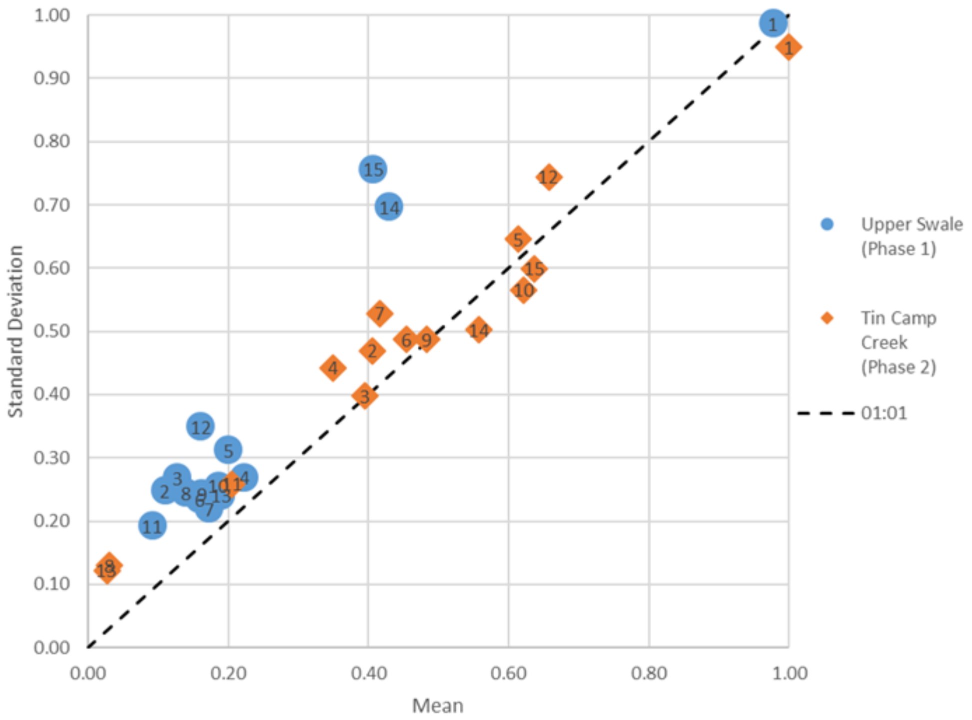
3.2 Catchment Sediment Yield Vs Internal Geomorphology

The core behaviours of catchment sediment yield and internal geomorphology show a different response to the changes in parameter values, as can be seen in Figure 4, and also the rankings in Table 4. For both catchments, SED is ranked as most influential for catchment sediment yields. For influence on the internal geomorphology, SEC ranks higher in the Tin Camp Creek catchment. The Upper Swale catchment displays a similar response with both behaviours, with SED and MNR most influential and by similar amounts, although GSS has less influence on internal geomorphology. The change in response for Tin Camp Creek is more varied – SED is less influential on internal geomorphology, and SEC is the most influential with a higher aggregated mean. GSS is slightly less influential, and MNR slightly more, and VEG is more influential on the internal geomorphology than it is on catchment sediment yield. For both model functions, there again is a strong visually correlation between those parameters showing the most influence and those showing the most non-linear behaviour.

Table 4 – Parameters ranked by means for each catchment from the aggregated scores for catchment sediment yields (SY) and internal geomorphology (IG) elementary effects. SED = sediment transport formula; MEL = maximum erode limit; CLR = in channel lateral erosion rate; LAT = lateral erosion rate; VEG = vegetation critical shear stress; MAT = grass maturity rate; SCR = soil creep rate; SFT = slope failure threshold; IOD in/out difference; MinQ = minimum Q value; MaxQ = maximum Q value; SEC = slope for edge cells; EVR = evaporation rate; MNR = Manning's n roughness coefficient; and GSS = grain size set.

Rank (by mean: 1 = most influential)	**Upper Swale**		**Tin Camp Creek**	
	SY	**IG**	**SY**	**IG**
1	SED	SED	SED	SEC
2	MNR	MNR	SEC	SED
3	GSS	GSS	GSS	VEG
4	LAT	VEG	MinQ	MNR
5	VEG	CLR	VEG	MinQ
6	EVR	LAT	MNR	GSS
7	MinQ	MinQ	IOD	SCR
8	SCR	MaxQ	MAT	MAT
9	IOD	EVR	SCR	IOD
10	SEC	IOD	MEL	LAT

11	MAT	MAT	CLR	MEL
12	SFT	SEC	LAT	CLR
13	CLR	SCR	MaxQ	MaxQ
14	MEL	MEL	SFT	SFT
15	MaxQ	SFT	EVR	EVR



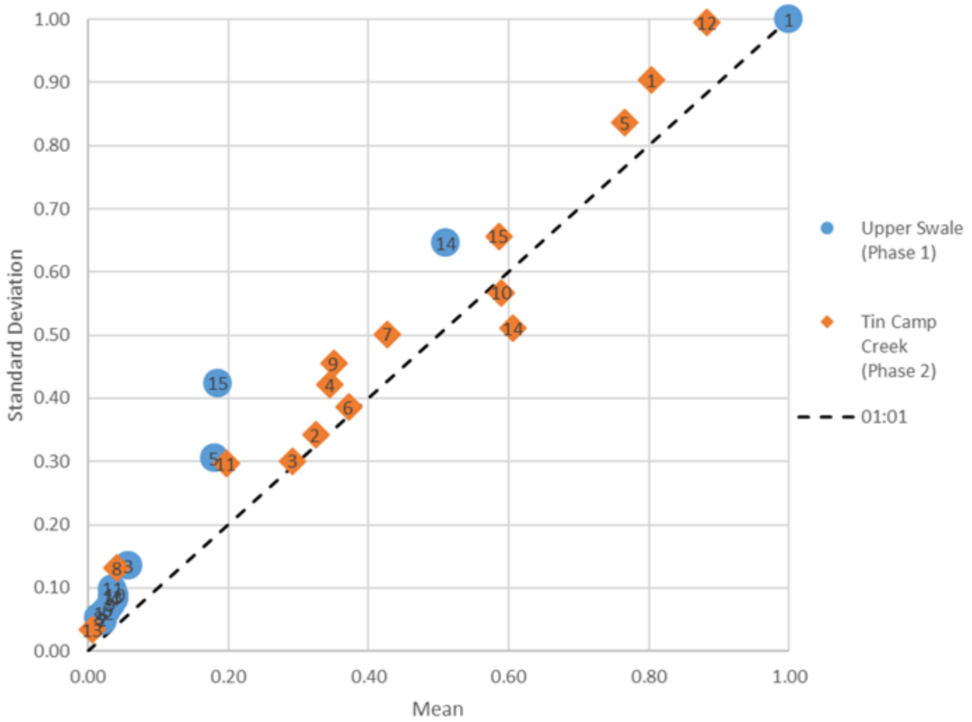


Figure 4 – Aggregated scores for sediment yield (top) and internal geomorphology (bottom) where: 1 = sediment transport formula (SED); 2 = maximum erode limit (MEL); 3 = in channel lateral erosion rate (CLR); 4 = lateral erosion rate (LAT); 5 = critical vegetation shear stress (VEG); 6 = grass maturity rate (MAT); 7 = soil creep rate (SCR); 8 = slope failure threshold (SFT); 9 = in/out difference (IOD); 10 = minimum Q value (MinQ);

11 = maximum Q value (MaxQ); 12 = slope for edge cells (SEC); 13 = evaporation rate (EVR); 14 = Manning's n roughness coefficient (MNR); and 15 = grain size set (GSS).

3.3 Changes in the Mean Elevations

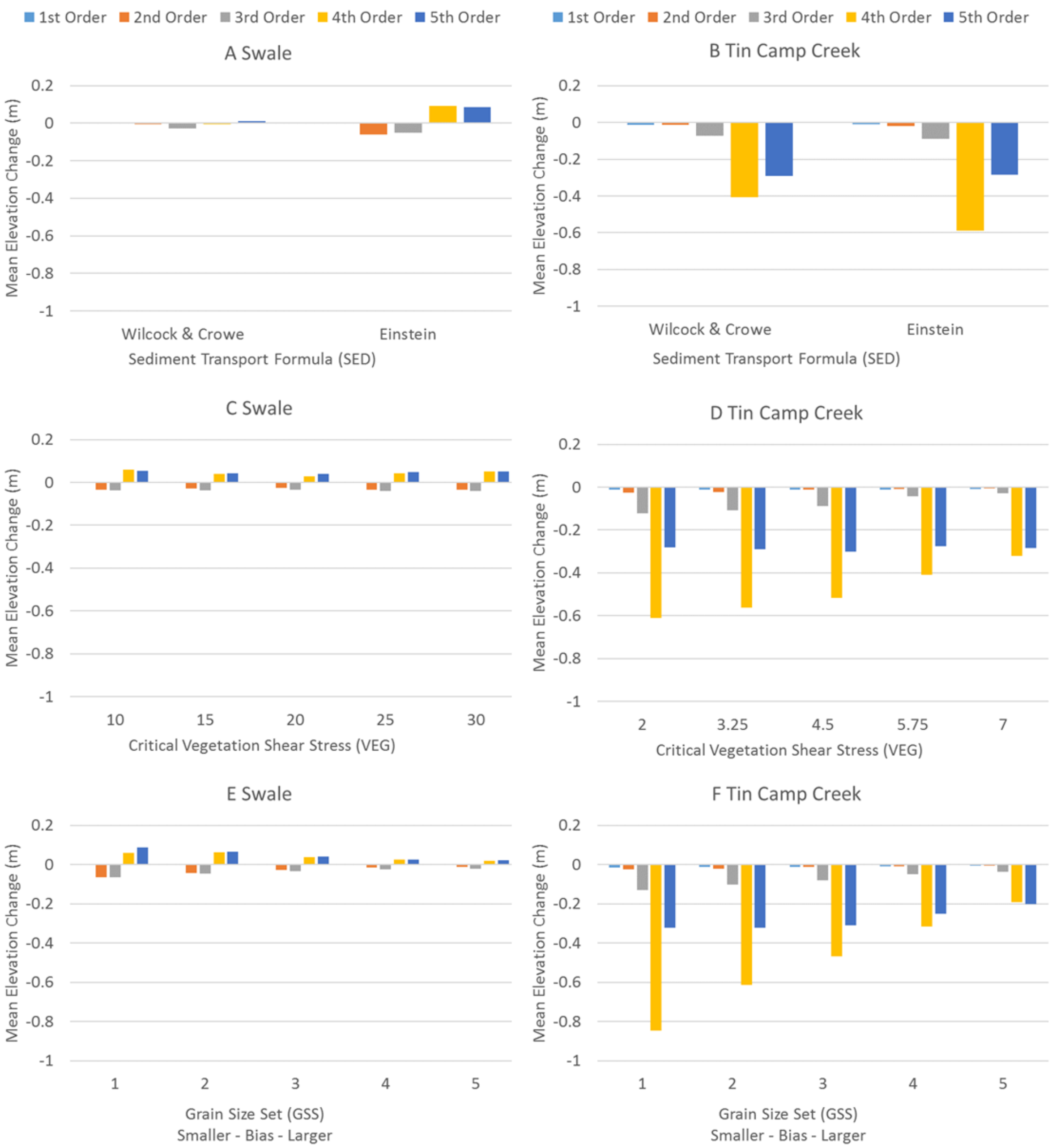


Figure 5 – Illustration of changes in the mean elevations for Upper Swale (A, C and E), and Tin Camp Creek (B, D and F) for the tests split by SED (A and B), VEG (C and D), and GSS (E and F) where 1 and 2 are biased smaller,

and 4 and 5 are biased larger. The catchment is sub-divided into watersheds of five stream orders, based on proportion of catchment drained.

The test results were binned by the parameter values used, and the mean changes in the mean elevations across the 5 stream orders calculated – Figure 5 illustrates how changes in parameter values might influence the spatial patterns of landscape change using SED, VEG and GSS as examples. For SED (Fig 5.A and 5.B), the most obvious difference is the scale of changes seen using each formula with Einstein generally showing greater change. For Tin Camp Creek (Fig 5.B) the spatial changes are similar, but for the larger Swale (Fig 5.A) there are differences in relative rates in 2nd and 4th order areas. In the Swale, VEG (Fig 5.C) appears to have little impact on the patterns and scale of changes, yet in Tin Camp Creek (Fig 5.D) there is reduction in the rates of erosion across the catchment with higher values, except in the 5th order areas which remain at a similar level. Finally, both catchments show a reduction in rates of erosion with a greater proportion of larger grain sizes, yet this is more pronounced 4th order areas in Tin Camp Creek (Fig 5.F).

4. Discussion

The results reveal some important insights into the application of the SA to LEMs generally, and also on specific behaviours of the CAESAR-Lisflood model. Here we discuss model functions (Section 4.1), sediment transport formulae (Section 4.2), implications for calibrating LEMs (Section 4.3), full uncertainty analyses of LEMs (Section 4.4), and limitations of this study (Section 4.5).

4.1 Model Functions

Our findings show that different model functions provide us with different indications of model sensitivity. This has important implications for how to measure LEM performance – and more widely how to quantify and assess geomorphic change within a basin. For example, Figure 4 and Table 4 show how any LEM assessment must depend on the applied metric for comparison. Model functions that quantify sediment yield (derived at the catchment outlet) indicate different sensitivities compared to model functions that quantify the internal landform response (based on spatial measures from within the catchment). Whilst at-a-point sediment yields are straightforward to extract from simulation data and easily related to field measurements (e.g., gauges, although these have their own associated uncertainties), similar or identical yields may conceal very different behaviours within the basin. This highlights an important aspect of LEM calibration: changes in sediment yields from a catchment outlet only provide partial information of what is changing internally. We therefore argue that metrics incorporating *spatial* changes in the basin (as well as bulk figures) are vital for assessing LEM performance. (i.e., time series of high resolution DEM data from LiDAR/photogrammetry) This is especially important as the shape of the landscape – where material has been eroded and deposited – is effectively the basins geomorphic memory and will directly influence subsequent model performance. For other basin scale models (e.g., hydrological models) this aspect is possibly not so important over longer-terms given the limited temporal extent memory of basin antecedence. Some of the challenges of LEM output comparison are similar to those of meteorology/climatology and may require a shift in expectation from end users as to what is possible. For example, predicting detailed patterns of local erosion and deposition is akin to predicting weather (low comparability especially over longer time scales) but more general (spatial and temporal) patterns of basin change are similar to climate predictions (better comparability especially for longer time scales).

4.2. Sediment Transport Formulae

Our SA shows that the choice of sediment transport formula (SED) had a very strong impact on the model functions. As sediment transport formulae are also integrated into other LEMs and geomorphic models they will affect their outcomes too. Looking at sediment transport formulae themselves, Gomez and Church (1989) tested 11 different sediment transport formulae to the same data sets and showed widespread variation in predictions – in some cases over orders of magnitude. The variation in the model performance can be explained by the derivation of the sediment transport formulae themselves, that are often theory-based but fitted to limited laboratory and field data, sometimes representing temporal averages over equilibrium conditions (Gomez and Church, 1989). The formulae do not, and were likely never intended to, represent the full variation of actual flow conditions in natural river. As LEMs commonly amalgamate a set of geomorphic models or transport formulae, their performance hinges on a number of individual model components. Therefore, when applied to different situations, they may not be appropriate (Coulthard et al., 2007a).

4.3 Implications for Calibrating LEMs

This, however, presents a challenge, as it is highly likely that the sediment transport formula to be used was neither designed nor calibrated for a particular model application. The SIBERIA model (Hancock et al., 2010, 2016, 2017; Hancock and Willgoose, 2001; Willgoose et al., 2003) overcomes this issue by having a version of the Einstein sediment transport formula (Einstein, 1950) that is calibrated or tuned to field data on erosion rates. However, even when calibrated, LEMs (and their sediment transport formulae) face another hurdle with the non-stationarity of basin sediment yields. For example, a calibrated LEM will be adjusted to perform for a set of observed sediment outputs or erosion and deposition patterns. If, due to climate change for example, sediment supply, rainfall or channel flows increase outside of the range of the initial calibration then that initial calibration may no longer be valid (Coulthard et al., 2007b). This is similar to issues faced by calibrating hydrological models (e.g., Li et al., 2012) though the non-linear sediment response of LEMs like CAESAR-Lisflood (Coulthard et al., 2012) may make LEMs more sensitive to this. Such a non-linear sediment response to hydrological increases can be traced to the calculation of sediment transport as a square or cubic function of flow velocity. Furthermore, this analysis suggests that detailed justification and calibration of model choices around sediment transport will lead to the most effective gains in model skill.

4.4 Full Uncertainty Analysis

It is important to note that the MM does not provide an absolute value of sensitivity, but ranks each factor based on its relative influence on the model. This means it can be used to assess the main sources of uncertainty on a particular model set up. The next step is then to establish how the uncertainty caused by model parameters (e.g., the choice of sediment transport formula) compares to other identified sources of uncertainty, such as rainfall input uncertainty, DEM observation and resolution uncertainty, and length of spin-up period. For example, it may be that the choice of sediment transport formula may only be a minor source of uncertainty compared to the DEM resolution, or equally, it might be the most significant source of uncertainty in a LEM's ouput.

Importantly, whilst the simulation of long-term development of landscapes may be somewhat resilient to some uncertainties, e.g., initial conditions (Hancock et al., 2016), any attempt to reproduce, predict or forecast physical changes should have the same appreciation of uncertainty and rigorous testing that is applied to models in other fields (e.g., hydrology and hydraulics). There are many methods available, but when discussing CAESAR-Lisflood the applications applied to Lisflood-FP seem a reasonable place to start. Lisflood-FP has been rigorously tested and benchmarked for decision-making purposes (Hunter et al., 2005; Neelz & Pender, 2013), and the use of SA to assess model response and uncertainty is standard practise (Di Baldassarre et al., 2009; Fewtrell et al., 2008, 2011; Hall et al., 2005; Horritt and Bates, 2001, 2002; Hunter et al., 2008; Neal et al., 2011; Sampson et al., 2012), often as a stage of calibration using the GLUE method (Aronica et al., 2002; Bates et al., 2004; Horritt et al., 2006; Hunter et al., 2005; Pappenberger et al., 2007; Wong et al., 2015). Uncertainty in model predictions can be accounted for by utilising probabilistic measures and uncertainty cascades (for example, Pappenberger et al., 2005; Stephens et al., 2012). This is not considered unique to CAESAR-Lisflood, and any application of an LEM or other geomorphic model for operational, decision-making or forecasting applications should make full consideration of all associated uncertainties.

4.5. Limitations

The main limitation of the MM is the subjectivity in selection of parameter values and ranges. Here, this has been mitigated by consistently selected ranges of +/- 50 % of a default value obtained from previous calibrations (where feasible). An issue emerges with categorical parameters, such as SED, where multiple values cannot be placed in spectrum across a range between minimum and maximum values. The MM has no formal method for dealing with such categorical parameters, so here it has been assumed that switching from one formula to another is a single iterative step change, and this would be the same even with more choices available. This reflects the purpose of the MM, which is to inform about the relative importance of choices of parameter values on the performance/behaviour of the model. However, to assess the impact of this single step-change assumption, we performed a further analysis, where it was assumed that switching formula was a change of four iterative steps. This analysis shows that the relative sensitivity of the model to the sediment transport formula choice becomes less important, with other parameters such as Manning's n Roughness and grain size sets increasing in relative influence (see Supplementary Material S2 for full results of this analysis).

An obvious limitation to this exercise is computational resource. This study incorporated 1600 individual model runs to test the behavioural response of the model to 15 parameters, in just two catchments, and this partly influenced the choice to limit simulation periods to 20 years. The bulk of simulations used Intel i7-5960X processors and using Solid State Drives (SSD), yet the run times varied considerably depending on the parameter sets chosen. As an indication, the mean simulation run time for the first repeat in each catchment was 11 hours and 23 minutes for the Swale and 21 minutes for

Tin Camp Creek. We used a batch mode functionality of CAESAR-Lisflood to run simulations of each repeat (16 model runs each) consecutively, and distributed batches across different machines – this is feasible for the model set ups described. However, for long-term simulations for catchments the size of the Upper Swale, individual model runs can take several weeks and running several runs consecutively becomes prohibitive. One solution would be to distribute the jobs on High Performance Computing (HPC) facilities, where the time for a single model run would not significantly decrease, but several hundred, even thousands, of individual model runs can be performed coincidently.

Here, the methodology has only been applied to the CAESAR-Lisflood model, and although some of the findings will be unique to CAESAR-Lisflood and the model set ups presented, they have implications for all LEMs. Importantly, the methodology can serve as a highly useful tool for users to determine the behaviour of any LEM model set up prior to calibration and/or simulation.

5. Conclusions

The feasibility of performing global SA to a highly parameterised catchment LEM has been demonstrated through the application of the MM to the CAESAR-Lisflood model. The analysis was repeated over two different catchments suggesting some model behaviours are universal, and others vary depending on the catchment characteristics providing crucial information to inform future model developments. This analysis confirms that the sediment transport formulae are a significant source of uncertainty in LEMs, and in the CAESAR-Lisflood model the use of one formula over another can result in an order of magnitude differences in sediment yields when all other factors are kept constant. Another finding with relevance to SA and calibration of LEMs was the influence of parameters on each model function, showing that one aspect of model behaviour (e.g., catchment sediment yield) is not fully reflective of other, albeit related, model behaviours (e.g., internal geomorphology).

In addition to the above, the results reveal the parameters in CAESAR-Lisflood which exert the greatest influence, and whilst we can only apply this to the CAESAR-Lisflood model itself, it is likely that LEMs with comparable parameters will display similar behaviours. Some of the most influential parameters, like Manning's n roughness coefficient, grain size distributions, and vegetation critical shear stress are physically-based, so any uncertainty can be reduced by more detailed field measurements. We also show that parameters that determine the numerical efficiency of CAESAR-Lisflood exert a medium influence on the simulation results. Although some parameters exerted less influence on model behaviour relative to others, there were no parameters which did not influence the model in some way.

The application of a global SA should become a vital step in any investigation using LEMs. This paper has demonstrated that the use of the MM is efficient for this purpose and yielded some valuable insights into model behaviour that can ultimately feed back into model set up, as well as future model development.

Model and Data Availability

The data produced by this study is made available on request from the corresponding author. The CAESAR-Lisflood model used in this study is freely available under a GNU licence from http://www.coulthard.org.uk

Competing Interests

The authors declare that they have no conflict of interest.

Acknowledgements

The authors wish to thank the two reviewers, Andy Wickert and Daniel Hobley, for their insightful and helpful comments which have improved this manuscript. The Landscape Evolution Model Sensitivity Investigation Project (LEMSIP) has emerged from the Field and Computer Simulation in Landscape Evolution (FACSIMILE) network. The aims of the project are to collate and generate knowledge pertaining to the sensitivities and uncertainties associated with Landscape Evolution Models, and how these influence the simulation of landscape development. The authors wish to thank the Young Geomorphologists group who donated computational resource. This work was supported by the NERC Flooding from Intense Rainfall (FFIR) project, Susceptibility of Basins to Intense Rainfall and Flooding (SINATRA) NE/K008668/1. The CAESAR-Lisflood model used in this study is freely available under a GNU licence from http://www.coulthard.org.uk

References

Adams, J. M., Gasparini, N. M., Hobley, D. E. J., Tucker, G. E., Hutton, E. W. H., Nudurupati, S. S. and Istanbulluoglu, E.: The Landlab v1.0 OverlandFlow component: a Python tool for computing shallow-water flow across watersheds, Geosci. Model Dev, 10, 1645–1663, doi:10.5194/gmd-10-1645-2017, 2017.

Andersen, J. L., Egholm, D. L., Knudsen, M. F., Jansen, J. D. and Nielsen, S. B.: The periglacial engine of mountain erosion - Part 1: Rates of frost cracking and frost creep, Earth Surf. Dyn., 3(4), 447–462, doi:10.5194/esurf-3-447-2015, 2015.

Armitage, J. J., Whittaker, A. C., Zakari, M. and Campforts, B.: Numerical modelling landscape and sediment flux response to precipitation rate change, Earth Surf. Dyn. Discuss., (May), 1–31, doi:10.5194/esurf-2017-34, 2017.

Aronica, G., Bates, P. D. and Horritt, M. S.: Assessing the uncertainty in distributed model predictions using observed binary pattern information within GLUE, Hydrol. Process., 16(10), 2001–2016, doi:10.1002/hyp.398, 2002.

Attal, M., Tucker, G. E., Whittaker, A. C., Cowie, P. A. and Roberts, G. P.: Modelling fluvial incision and transient landscape evolution: Influence of dynamic Channel adjustment, J. Geophys. Res. Earth Surf., 113(3), 1–16, doi:10.1029/2007JF000893, 2008.

Di Baldassarre, G., Schumann, G. and Bates, P. D.: A technique for the calibration of hydraulic models using uncertain satellite observations of flood extent, J. Hydrol., 367(3), 276–282, doi:10.1016/j.jhydrol.2009.01.020, 2009.

Bates, P. D., Horritt, M. S., Aronica, G. and Beven, K.: Bayesian updating of flood inundation likelihoods conditioned on flood extent data, Hydrol. Process., 18(17), 3347–3370, doi:10.1002/hyp.1499, 2004.

Bates, P. D., Horritt, M. S. and Fewtrell, T. J.: A simple inertial formulation of the shallow water equations for efficient two-dimensional flood inundation modelling, J. Hydrol., 387(1–2), 33–45, doi:10.1016/j.jhydrol.2010.03.027, 2010.

Beven, K. and Kirkby, M.: A physically based, variable contributing area model of basin hydrology/Un modèle à base physique de zone d'appel variable de l'hydrologie du bassin versant, Hydrol. Sci. J., 24(1), 37–41 [online] Available from: http://www.tandfonline.com/doi/abs/10.1080/02626667909491834 (Accessed 8 May 2014), 1979.

Brockmann, D. and Morgenroth, E.: Comparing global sensitivity analysis for a biofilm model for two-step nitrification using the qualitative screening method of Morris or the quantitative variance-based Fourier Amplitude Sensitivity Test (FAST), Water Sci. Technol., 56(8), 85–93, doi:10.2166/wst.2007.600, 2007.

Campforts, B., Schwanghart, W., Govers, G.: Accurate simulation of transient landscape evolution by eliminating numerical diffusion: the TTLEM 1.0 model. Earth Surface Dynamics, 5, 47-66, doi: 10.5194/esurf-5-47-2017, 2017.Campolongo, F., Cariboni, J. and Saltelli, A.: An effective screening design for sensitivity analysis of large models, Environ. Model. Softw., 22(10), 1509–1518, doi:10.1016/j.envsoft.2006.10.004, 2007.

Cariboni, J., Gatelli, D., Liska, R. and Saltelli, A.: The role of sensitivity analysis in ecological modelling, Ecol. Modell., 203(1–2), 167–182, doi:10.1016/j.ecolmodel.2005.10.045, 2007.

Collins, D. B. G.: Modeling the effects of vegetation-erosion coupling on landscape evolution, J. Geophys. Res., 109(F3), 1–11, doi:10.1029/2003JF000028, 2004.

Coulthard, T., Hicks, D. and Wiel, M. Van De: Cellular modelling of river catchments and reaches: Advantages, limitations and prospects, Geomorphology, 90(3–4), 192–207, doi:10.1016/j.geomorph.2006.10.030, 2007a.

Coulthard, T., Neal, J., Bates, P., Ramirez, J., de Almeida, G. and Hancock, G.: Integrating the LISFLOOD-FP 2D hydrodynamic model with the CAESAR model: implications for modelling landscape evolution, Earth Surf. …, 38(15), 1897–1906, doi:10.1002/esp.3478, 2013.

Coulthard, T. J. and Macklin, M. G.: How sensitive are river systems to climate and land-use changes? A model-based evaluation, J. Quat. Sci., 16(4), 347–351, doi:10.1002/jqs.604, 2001.

Coulthard, T. J. and Skinner, C. J.: The sensitivity of landscape evolution models to spatial and temporal rainfall resolution, Earth Surf. Dyn., 4(3), 757–771, doi:10.5194/esurf-4-757-2016, 2016a.

Coulthard, T. J. and Skinner, C. J.: The sensitivity of landscape evolution models to spatial and temporal rainfall resolution, Earth Surf. Dyn. Discuss., 1–28, doi:10.5194/esurf-2016-2, 2016b.

Coulthard, T. J. and Van De Wiel, M. J.: Quantifying fluvial non linearity and finding self organized criticality? Insights from simulations of river basin evolution, Geomorphology, 91(3–4), 216–235, doi:10.1016/j.geomorph.2007.04.011, 2007.

Coulthard, T. J. and Van De Wiel, M. J.: Modelling river history and evolution, Philos. Trans. R. Soc. A Math. Phys. Eng. Sci., 370(1966), 2123–2142, doi:10.1098/rsta.2011.0597, 2012.

Coulthard, T. J. and Van De Wiel, M. J.: Climate, tectonics or morphology: What signals can we see in drainage basin sediment yields?, Earth Surf. Dyn., 1(1), 13–27, doi:10.5194/esurf-1-13-2013, 2013.

Coulthard, T. J. and Van De Wiel, M. J.: Modelling long term basin scale sediment connectivity, driven by spatial land use changes, Geomorphology, 277, 265–281, doi:10.1016/j.geomorph.2016.05.027, 2017.

Coulthard, T. J., Lewin, J. and Macklin, M. G.: 12 Non-stationarity of basin scale sediment delivery in response to climate change, Dev. Earth Surf. Process., 11(07), 315–331, doi:10.1016/S0928-2025(07)11131-7, 2007b.

Coulthard, T. J., Ramirez, J., Fowler, H. J. and Glenis, V.: Using the UKCP09 probabilistic scenarios to model the amplified impact of climate change on drainage basin sediment yield, Hydrol. Earth Syst. Sci., 16(11), 4401–4416, doi:10.5194/hess-16-4401-2012, 2012.

Egholm, D. L., Andersen, J. L., Knudsen, M. F., Jansen, J. D. and Nielsen, S. B.: The periglacial engine of mountain erosion - Part 2: Modelling large-scale landscape evolution, Earth Surf. Dyn., 3(4), 463–482, doi:10.5194/esurf-3-463-2015, 2015.

Einstein, H. A.: The Bed-Load Function for Sediment Transportation in Open Channel Flows, Soil Conserv. Serv., (1026), 1–31 [online] Available from: https://ponce.sdsu.edu/einstein_bedload_function.pdf (Accessed 4 July 2018), 1950.

Fewtrell, T. J., Bates, P. D., Horritt, M. and Hunter, N. M.: Evaluating the effect of scale in flood inundation modelling in urban environments, Hydrol. Process., 22(26), 5107–5118, doi:10.1002/hyp.7148, 2008.

Fewtrell, T. J., Duncan, A., Sampson, C. C., Neal, J. C. and Bates, P. D.: Benchmarking urban flood models of varying complexity and scale using high resolution terrestrial LiDAR data, Phys. Chem. Earth, Parts A/B/C, 36(7), 281–291, doi:10.1016/j.pce.2010.12.011, 2011.

Gomez, B. and Church, M.: An Assessment of Bedload Sediment transport Formulae for Gravel Bed Rivers, Water Resour. Res., 25(6), 1161–1186, 1989.

van Griensven, A., Meixner, T., Grunwald, S., Bishop, T., Diluzio, M. and Srinivasan, R.: A global sensitivity analysis tool for the parameters of multi-variable catchment models, J. Hydrol., 324(1–4), 10–23, doi:10.1016/j.jhydrol.2005.09.008, 2006.

Hall, J. W., Tarantola, S., Bates, P. D. and Horritt, M. S.: Distributed Sensitivity Analysis of Flood Inundation Model Calibration, J. Hydraul. Eng., 131(2), 117–126, doi:10.1061/(ASCE)0733-9429(2005)131:2(117), 2005.

Hancock, G. and Willgoose, G.: Use of a landscape simulator in the validation of the SIBERIA catchment evolution model: Declining equilibrium landforms, Water Resour. Res., 37(7), 1981–1992, doi:10.1029/2001WR900002, 2001.

Hancock, G. R.: The impact of different gridding methods on catchment geomorphology and soil erosion over long timescales using a landscape evolution model, Earth Surf. Process. Landforms, 31(8), 1035–1050, doi:10.1002/esp.1306, 2006.

Hancock, G. R. and Coulthard, T. J.: Channel movement and erosion response to rainfall variability in southeast Australia, Hydrol. Process., 26(5), 663–673, doi:10.1002/hyp.8166, 2012.

Hancock, G. R., Lowry, J. B. C., Coulthard, T. J., Evans, K. G. and Moliere, D. R.: A catchment scale evaluation of the SIBERIA and CAESAR landscape evolution models, Earth Surf. Process. Landforms, 35(8), 863–875, doi:10.1002/esp.1863, 2010.

Hancock, G. R., Coulthard, T. J., Martinez, C. and Kalma, J. D.: An evaluation of landscape evolution models to simulate decadal and centennial scale soil erosion in grassland catchments, J. Hydrol., 398(3–4), 171–183, doi:10.1016/j.jhydrol.2010.12.002, 2011.

Hancock, G. R., Lowry, J. B. C. and Coulthard, T. J.: Catchment reconstruction - erosional stability at millennial time scales using landscape evolution models, Geomorphology, 231, 15–27, doi:10.1016/j.geomorph.2014.10.034, 2015.

Hancock, G. R., Coulthard, T. J. and Lowry, J. B. C.: Predicting uncertainty in sediment transport and landscape evolution - the influence of initial surface conditions, Comput. Geosci., 90, 117–130, doi:10.1016/j.cageo.2015.08.014, 2016.

Hancock, G. R., Verdon-Kidd, D. and Lowry, J. B. C.: Sediment output from a post-mining catchment - Centennial impacts using stochastically generated rainfall, J. Hydrol., 544, 180–194, doi:10.1016/j.jhydrol.2016.11.027, 2017.

Herman, J. D., Kollat, J. B., Reed, P. M. and Wagener, T.: Technical Note: Method of Morris effectively reduces the computational demands of global sensitivity analysis for distributed watershed models, Hydrol. Earth Syst. Sci., 17(7), 2893–2903, doi:10.5194/hess-17-2893-2013, 2013.

Horritt, M., Bates, P. and Mattinson, M.: Effects of mesh resolution and topographic representation in 2D finite volume models of shallow water fluvial flow, J. Hydrol., 329(1–2), 306–314, doi:10.1016/j.jhydrol.2006.02.016, 2006.

Horritt, M. S. and Bates, P. D.: Effects of spatial resolution on a raster based model of flood flow, J. Hydrol., 253(1–4), 239–249, doi:10.1016/S0022-1694(01)00490-5, 2001.

Horritt, M. S. and Bates, P. D.: Evaluation of 1D and 2D numerical models for predicting river flood inundation, J. Hydrol., 268(1), 87–99, doi:10.1016/S0022-1694(02)00121-X, 2002.

Hunter, N. M., Horritt, M. S., Bates, P. D., Wilson, M. D. and Werner, M. G. F.: An adaptive time step solution for raster-based storage cell modelling of floodplain inundation, Adv. Water Resour., 28(9), 975–991, doi:10.1016/j.advwatres.2005.03.007, 2005.

Hunter, N. M., Bates, P. D., Neelz, S., Pender, G., Villanueva, I., Wright, N. G., Liang, D., Falconer, R.

A., Lin, B., Waller, S., Crossley, A. J. and Mason, D. C.: Benchmarking 2D hydraulic models for urban flooding, Proc. Inst. Civ. Eng. - Water Manag., 161(1), 13–30, doi:10.1680/wama.2008.161.1.13, 2008.

Ibbitt, R. P., Willgoose, G. R. and Duncan, M. J.: Channel network simulation models compared with data from the Ashley River, New Zealand, Water Resour. Res., 35(12), 3875–3890, doi:10.1029/1999WR900245, 1999.

Ijjasz-Vasquez, E. J., Bras, R. L. and Moglen, G. E.: Sensitivity of a basin evolution model to the nature of runoff production and to initial conditions, Water Resour. Res., 28(10), 2733–2741, doi:10.1029/92WR01561, 1992.

Istanbulluoglu, E. and Bras, R. L.: Vegetation-modulated landscape evolution: Effects of vegetation on landscape processes, drainage density, and topography, J. Geophys. Res. Earth Surf., 110(2), 1–19, doi:10.1029/2004JF000249, 2005.

Jerolmack, D. J. and Paola, C.: Shredding of environmental signals by sediment transport, Geophys. Res. Lett., 37(19), 1–5, doi:10.1029/2010GL044638, 2010.

Larsen, L., Thomas, C., Eppinga, M. and Coulthard, T.: Exploratory modeling: Extracting causality from complexity, Eos (Washington. DC)., 95(32), 285–286, doi:10.1002/2014EO320001, 2014.

Li, C., Zhang, L., Wang, H., Zhang, Y., Yu, F. and Yan, D.: The transferability of hydrological models under nonstationary climatic conditions., Hydrol. Earth …, 16(4), 1239–1254, doi:10.5194/hess-16-1239-2012, 2012.

Liu, B. and Coulthard, T. J.: Modelling the interaction of aeolian and fluvial processes with a combined cellular model of sand dunes and river systems, Comput. Geosci., 106, 1–9, doi:10.1016/j.cageo.2017.05.003, 2017.

Martin, Y. and Church, M.: Numerical modelling of landscape evolution: geomorphological perspectives, Prog. Phys. Geogr., 28(3), 317–339, doi:10.1191/0309133304pp412ra, 2004.

Met Office: 5km UK Composite Rainfall Data from the Met Office NIMROD System, NCAS Br. Atmos. Data Centre, available at : http://catalogue.ceda.ac.uk/uuid/82adec1f896af6169112d09cc1174499 (last access: 20 September 2016), 2003.

Morris, M. D.: Factorial Sampling Plans for Preliminary Computational Experiments, Technometrics, 33(2), 161–174, doi:10.2307/1269043, 1991.

Nash, J. and Sutcliffe, J.: River flow forecasting through conceptual models part I—A discussion of principles, J. Hydrol., 10, 282–290 [online] Available from: http://www.sciencedirect.com/science/article/pii/0022169470902556 (Accessed 8 May 2014), 1970.

Neal, J., Schumann, G., Fewtrell, T., Budimir, M., Bates, P. and Mason, D.: Evaluating a new LISFLOOD-FP formulation with data from the summer 2007 floods in Tewkesbury, UK, J. Flood Risk Manag., 4(2), 88–95, doi:10.1111/j.1753-318X.2011.01093.x, 2011.

Neelz, S. & Pender, G.: Benchmarking the latest generation of 2D hydraulic modelling packages. [online] Available from: http://evidence.environment-agency.gov.uk/FCERM/Libraries/FCERM_Project_Documents/SC120002_Benchmarking_2D_hydraulic_models_Report.sflb.ashx, 2013.

Neumann, M. B.: Comparison of sensitivity analysis methods for pollutant degradation modelling: A case study from drinking water treatment, Sci. Total Environ., 433(October), 530–537, doi:10.1016/j.scitotenv.2012.06.026, 2012.

Norton, J. P.: Algebraic sensitivity analysis of environmental models, Environ. Model. Softw., 23, 963–972, doi:10.1016/j.envsoft.2007.11.007, 2008.

Norton, J. P.: Selection of Morris trajectories for initial sensitivity analysis, IFAC., 2009.

Oakley, J. E. and O'Hagan, A.: Probabilistic Sensitivity Analysis of Complex Models : A Bayesian Approach Author (s): Jeremy E . Oakley and Anthony O'Hagan Published by : Wiley for the Royal Statistical Society Stable URL : http://www.jstor.org/stable/3647504 Probabilistic sensitiv, , 66(3), 751–769, 2004.

Oreskes, N., Shrader-Frechette, K. and Belitz, K.: Verification, Validation, and Confirmation of Numerical Models in the Earth Sciences, Science (80-.)., 263, 641–646, doi:10.2307/2883078, 1994.

Pappenberger, F., Beven, K. J., Hunter, N. M., Bates, P. D., Gouweleeuw, B. T., Thielen, J. and Roo, A. P. J. De: Cascading model uncertainty from medium range weather forecasts (10 days) through a rainfall-runoff model to flood inundation predictions within the European Flood Forecasting System (EFFS), Hydrol. Earth Syst. Sci. Discuss., 9(4), 381–393, doi:10.5194/hess-9-381-2005, 2005.

Pappenberger, F., Harvey, H., Beven, K., Hall, J. and Meadowcroft, I.: Decision tree for choosing an uncertainty analysis methodology : a wiki experiment, Hydrol. Process., 20, 3793–3798, doi:10.1002/hyp, 2006.

Pappenberger, F., Frodsham, K., Beven, K., Romanowicz, R. and Matgen, P.: Fuzzy set approach to calibrating distributed flood inundation models using remote sensing observations, Hydrol. Earth Syst. Sci. Discuss., 11(2), 739–752 [online] Available from: https://hal.archives-ouvertes.fr/hal-00305049/ (Accessed 24 May 2017), 2007.

Pappenberger, F., Beven, K. J., Ratto, M. and Matgen, P.: Multi-method global sensitivity analysis of flood inundation models, Adv. Water Resour., 31(1), 1–14, doi:10.1016/j.advwatres.2007.04.009, 2008.

Pazzaglia, F. J.: Landscape evolution models, pp. 247–274., 2003.

Petersen, A. C. (Arthur C.: Simulating nature : a philosophical study of computer-simulation uncertainties and their role in climate science and policy advice, CRC Press. [online] Available from: https://books.google.co.uk/books?hl=en&lr=&id=I4GhNkiPv3EC&oi=fnd&pg=PP1&dq=Simulating+Nature:+A+Philosophical+Study+of+Computer-Simulation+Uncertainties+and+Their+Role+in+Climate+Science+and+Policy+Advice&ots=EKmUbPTtVZ&sig=BisleTDNw3E0_EpozyLbxjJHUdg#v=onepage&q=Simulating Nature%3A A Philosophical Study of Computer-Simulation Uncertainties and Their Role in Climate Science and Policy Advice&f=false (Accessed 18 August 2017), 2012.

Pianosi, F., Beven, K., Freer, J., Hall, J. W., Rougier, J., Stephenson, D. B. and Wagener, T.: Sensitivity analysis of environmental models: A systematic review with practical workflow, Environ. Model. Softw., 79, 214–232, doi:10.1016/j.envsoft.2016.02.008, 2016.

Pujol, G.: R Package "sensitivity". Version 1.4-0, 2009.

R Hancock, G.: Modelling stream sediment concentration: An assessment of enhanced rainfall and storm frequency, J. Hydrol., 430–431, 1–12, doi:10.1016/j.jhydrol.2012.01.022, 2012.

Ratto, M., Pagano, A. and Young, P.: State Dependent Parameter metamodelling and sensitivity analysis, Comput. Phys. Commun., 177(11), 863–876, doi:10.1016/j.cpc.2007.07.011, 2007.

Saltelli, A., Chan, K. and Scott, E. M.: Sensitivity Analysis, John Wiley, New York, 2000.

Sampson, C. C., Fewtrell, T. J., Duncan, A., Shaad, K., Horritt, M. S. and Bates, P. D.: Use of terrestrial laser scanning data to drive decimetric resolution urban inundation models, Adv. Water Resour., 41, 1–17, doi:10.1016/j.advwatres.2012.02.010, 2012.

Saynor, M. J., Lowry, J., Erskine, W. D., Coulthard, T. and Hancock, G.: Assessing Erosion and Run-Off Performance of a Trial Rehabilitated, Proc. Life Mine Conf. July 2012, (July), 10–12, 2012.

Skinner, C. and Coulthard, T.: Caesar-Lisflood Existing Applications Parameter Listings - May 2017, , doi:10.5281/ZENODO.800558, 2017.

Sobol', I.: Global Sensitivity Indices for Nonlinear Mathematical Models:Review, Math. Comput. Simul., 55, 271–280, doi:10.1016/S0378-4754(00)00270-6, 2001.

Song, X., Zhan, C., Xia, J. and Kong, F.: An efficient global sensitivity analysis approach for distributed hydrological model, J. Geogr. Sci., 22(2), 209–222, doi:10.1007/s11442-012-0922-5, 2012.

Song, X., Zhang, J., Zhan, C., Xuan, Y., Ye, M. and Xu, C.: Global sensitivity analysis in hydrological modeling: Review of concepts, methods, theoretical framework, and applications, J. Hydrol., 523(225), 739–757, doi:10.1016/j.jhydrol.2015.02.013, 2015.

Stephens, E. M., Bates, P. D., Freer, J. E. and Mason, D. C.: The impact of uncertainty in satellite data on the assessment of flood inundation models, J. Hydrol., 414–415, 162–173, doi:10.1016/j.jhydrol.2011.10.040, 2012.

Tucker, G. E. and Bras, R. L.: A stochastic approach to modelling the role of rainfall variability in drainage basin evolution, Water Resour. Res., 36(7), 1953, doi:10.1029/2000WR900065, 2000.

Tucker, G. E. and Hancock, G. R.: Modelling landscape evolution, Earth Surf. Process. Landforms, 35(1), 28–50, doi:10.1002/esp.1952, 2010.

Vanwalleghem, T., Stockmann, U., Minasny, B. and McBratney, A. B.: A quantitative model for integrating landscape evolution and soil formation, J. Geophys. Res. Earth Surf., 118(2), 331–347, doi:10.1029/2011JF002296, 2013.

Welivitiya, W. D. D. P., Willgoose, G. R., Hancock, G. R. and Cohen, S.: Exploring the sensitivity on a soil area-slope-grading relationship to changes in process parameters using a pedogenesis model, Earth Surf. Dyn., 4(3), 607–625, doi:10.5194/esurf-4-607-2016, 2016.

Van De Wiel, M. J. and Coulthard, T. J.: Self-organized criticality in river basins: Challenging sedimentary records of environmental change, Geology, 38(1), 87–90, doi:10.1130/G30490.1, 2010.

Van De Wiel, M. J., Coulthard, T. J., Macklin, M. G. and Lewin, J.: Embedding reach-scale fluvial dynamics within the CAESAR cellular automaton landscape evolution model, Geomorphology, 90(3–4), 283–301, doi:10.1016/j.geomorph.2006.10.024, 2007.

Van De Wiel, M. J., Coulthard, T. J., Macklin, M. G. and Lewin, J.: Modelling the response of river systems to environmental change: Progress, problems and prospects for palaeo-environmental reconstructions, Earth-Science Rev., 104(1–3), 167–185, doi:10.1016/j.earscirev.2010.10.004, 2011.

Wilcock, P. R. and Crowe, J. C.: Surface-based Transport Model for Mixed-Size Sediment, J. Hydraul. Eng., 129(2), 120–128, doi:10.1061/(ASCE)0733-9429(2003)129:2(120), 2003.

Willgoose, G. R., Hancock, G. R. and Kuczera, G.: A Framework for the Quantitative Testing of Landform Evolution Models, pp. 195–216, American Geophysical Union., 2003.

Wong, J. S., Freer, J. E., Bates, P. D., Sear, D. A. and Stephens, E. M.: Sensitivity of a hydraulic model to channel erosion uncertainty during extreme flooding, Hydrol. Process., 29(2), 261–279, doi:10.1002/hyp.10148, 2015.

Yang, J.: Convergence and uncertainty analyses in Monte-Carlo based sensitivity analysis, Environ. Model. Softw., 26(4), 444–457, doi:10.1016/j.envsoft.2010.10.007, 2011.

Ziliani, L., Surian, N., Coulthard, T. J. and Tarantola, S.: Reduced-complexity modeling of braided rivers: Assessing model performance by sensitivity analysis, calibration, and validation, J. Geophys. Res. Earth Surf., 118(4), 2243–2262, doi:10.1002/jgrf.20154, 2013.