Modelling the decadal dynamics of reach-scale river channel evolution and floodplain turnover in CAESAR-Lisflood

Christopher J. Feeney,[1]* [iD] Richard C. Chiverrell,[1] Hugh G. Smith,[2] Janet M. Hooke[1] and James R. Cooper[1] [iD]
[1] School of Environmental Sciences, University of Liverpool, Liverpool L69 7ZT UK
[2] Landcare Research, Private Bag 1 1052, Palmerston North 4472, New Zealand

*Correspondence to: Christopher J. Feeney, School of Environmental Sciences, University of Liverpool, Liverpool L69 7ZT, UK. E-mail: c.j.feeney@liverpool.ac.uk

Earth Surface Processes and Landforms

ABSTRACT: Sedimentary deposits provide records of environmental change quantifying erosion fluxes conditioned by natural and anthropogenic disturbances. These fluxes are lagged by internal storage, particularly within floodplains, complicating reconstruction of environmental changes. The time sediment remains in storage underpins the interpretation of sedimentary records and accurate monitoring of pollutant fluxes. Turnover time is a measure of the timeframe to erode every floodplain surface. CAESAR-Lisflood is used to simulate fluvial evolution at reach scale, providing a basis for quantifying environmental changes on the timescales of sediment storage. We evaluate the accuracy of CAESAR-Lisflood simulations of channel changes and turnover times for alluvial floodplains using historical channel changes reconstructed for 10 reaches in northern England to quantify model accuracy in replicating mean annual erosion, deposition and channel lateral migration rates, alongside planform morphology. Here, a split-sample testing approach is adopted, whereby five of the reaches were calibrated and the resulting parameter values were applied to the other reaches to evaluate the transferability of parameter settings. The lowest overall integrated error identified the best-fit simulations and showed that modelled process rates were within ~25–50% of rates from historical reconstructions, generally. Calibrated parameters for some reaches are widely transferable, producing accurate geomorphic changes for some uncalibrated sites. However, large errors along some reaches indicate that reach-specific parameterization is recommended. Turnover times are underpinned by the assumption that areas of floodplain previously unvisited by the channel are reworked. This assumption has been challenged by studies that show floodplain (re)occupation rates vary spatially. However, this limitation is less important for the short-duration simulations presented here. The simulations reconstruct floodplain turnover times estimated by mapped rates mostly successfully, demonstrating the potential applicability of calibrated parameters over much longer timescales. Errors in the form of under-predicted erosion rates propagated, resulting in over-predicted turnover times by even greater magnitudes. © 2020 The Authors. Earth Surface Processes and Landforms Published by John Wiley & Sons Ltd.

KEYWORDS: CAESAR-Lisflood; channel change; floodplains; sediment; turnover

Introduction

Sedimentary units preserved in depositional basins provide useful evidence of environmental change, including records of sediment fluxes driven by tectonic, climatic and anthropogenic disturbances (Whittaker et al., 2010). In catchments, sediments are mobilized from upland sources, pass through transfer zones, and ultimately are deposited in long-term or permanent storage (Schumm, 1977). However, signals of sediment flux are typically lagged by internal storage mechanisms (e.g. Castelltort and Van Den Driessche, 2003; Blöthe and Korup, 2013). Additionally, multiple drivers can operate in parallel to initiate erosion, complicating the interpretation of sedimentary records (Coulthard and Van De Wiel, 2013).

Floodplains are significant sedimentary reservoirs in fluvial systems, covering 8×10^5–2×10^6 km² of the Earth's land surface (Leopold et al., 1964; Tockner and Stanford, 2002; Mitsch and Gosselink, 2015). A combination of vertical and lateral accretion processes, limited only by sediment supply, allows floodplains to accumulate rapidly large volumes of sediment over a short period of time (decades to centuries) (Trimble, 2010). However, sediment removal from the floodplain is largely restricted to lateral channel erosion, hence sediments may remain stored for a prolonged period of time – potentially many thousands of years (Trimble, 2010). It is this 'fast in, slow out' principle (*sensu* Trimble, 2010) that contributes to storage processes lagging sediment signals generated upstream.

Floodplain turnover, defined here as the set of processes of sediment removal from the floodplain, provides a useful measure for the longevity of sediment in storage (Bolin and Rodhe, 1973). Aalto et al. (2008) calculated turnover times of

~1000 years for floodplain sediments along 300 km of the Strickland River, Papua New Guinea, by dividing floodplain area by long-term average lateral erosion rates. Beechie et al. (2006) and Konrad (2012) quantified rates of (re)occupation of floodplain areas by river channels for several valley floors in North America.

Sediment storage within floodplains is also important in terrestrial geochemical cycling. Sutfin et al. (2015) estimate that soil organic carbon stored in floodplains accounts for 12–80 Pg C worldwide. Simulations, combining channel–floodplain evolution with biosphere–atmosphere carbon exchanges, show that the residence time (the mean amount of time a mass of particles resides in storage) of organic carbon is strongly controlled by sediment residence times (Torres et al., 2017). River catchments with a legacy of metal mining feature floodplains enriched with potentially harmful contaminants (e.g. Lewin and Macklin, 1987; Macklin and Dowsett, 1989; Bird et al., 2009). Modelling rates of floodplain turnover can help to constrain the timescale for decontamination of polluted catchments. For example, approximately 123 000 tonnes of Pb is stored within the main channel belt of the River Swale, UK, and under present rates of floodplain erosion, will take >5000 years to be removed from the catchment (Dennis et al., 2009).

River channel morphology and processes governing changes over time and space, such as riverbank erosion, are important controls on floodplain sediment storage (Hooke, 1980). As derivatives of their channel systems, any significant environmental change should lead to a transformation of floodplain forms and processes (Nanson and Croke, 1992). In their analysis of different floodplain environments in the US Pacific Northwest, Beechie et al. (2006) determined that the mean ages (defined here as the time since deposition) of floodplain sediments decreased from 85 to 63, 41 and 12 years for straight, meandering, island-braided and braided reaches, respectively. Miller and Friedman (2009) identified a strong positive correlation ($r = 0.95$) between the magnitude of the peak instantaneous flow event and rates of floodplain erosion that occurred in a given measurement interval. Separating out the multiple potential independent variables that control channel changes and the persistence of sediment storage in floodplains remains an important challenge (O'Connor et al., 2003). Therefore, methods that allow for careful control of potentially important variables, such as numerical modelling, need to be employed.

Landscape evolution models (LEMs) are used to simulate the redistribution of sediments in the landscape, and can operate over a range of spatial and temporal scales and environmental conditions (Tucker and Hancock, 2010). CAESAR-Lisflood is a coupled hydrodynamic LEM capable of simulating channel change and floodplain evolution at catchment and reach spatial scales (Coulthard et al., 2013), and has been used to quantify the effects of environmental change on sediment storage and fluxes. Applications of CAESAR-Lisflood include modelling the effects of vegetation cover on catchment-scale sediment delivery (Coulthard and Van De Wiel, 2017), impacts of tectonic uplift and rainfall-regime variability on sediment fluxes (Coulthard and Van De Wiel, 2013) and quantifying the role of self-organized criticality in governing bedload yields (Van De Wiel and Coulthard, 2010). At reach scale, Ziliani and Surian (2012) related phases of channel change along the Tagliamento River, Italy to changes in unit stream power, riverbank protection and a legacy of sediment mining. These studies highlight the importance of sediment storage at both catchment and reach scales and demonstrate the potential for modelling river channel changes and sediment residence times with CAESAR-Lisflood.

Here we test the effectiveness of CAESAR-Lisflood in reproducing historical river channel changes and modelling processes of floodplain turnover. To achieve this, we address several objectives. First, recent historical channel changes were reconstructed for a range of floodplain reaches across northern England. Second, we evaluated model performance by assessing how accurately the simulations replicated historical channel changes using the annual erosion, deposition and lateral migration rates of channels as performance criteria. Further evaluation of model performance included quantifying the spatial match between modelled and mapped patterns of channels and floodplain. From our modelling ensemble, the lowest overall error across these criteria identified the best-fit simulations and hence, the parameterization of variables governing erosion rates and channel morphology. A split-sample testing approach was adopted, whereby calibration focused on half of the selected reaches, with the other half reserved for further testing of model parameter accuracy. The most accurate simulation provided the erosion rates to model the areal extents of floodplain occupation by the river channel and extrapolate erosion rates to predict floodplain turnover times, which were compared with results derived from historical observations. Finally, we discuss the implications of our findings, including suggestions for the direction of future research into modelling channel–floodplain evolution processes.

Methods

Overview and model description

CAESAR-Lisflood is a coupled hydrodynamic LEM capable of simulating channel change and floodplain evolution at reach scale (Coulthard et al., 2013). The model routes water and sediments across a regular grid of cells, including divergent and convergent flows. This allows both single and multi-thread channel patterns to be modelled, meaning a wide range of floodplain evolution processes can be captured (Coulthard et al., 2002). During simulations, each cell records the elevation, sediment grain size distribution, flow depth, vegetation conditions, net erosion and deposition since the start of the simulation, updating these with each timestep (Coulthard et al., 2002). CAESAR-Lisflood operates using a digital elevation model (DEM) of a real-world system or an artificially generated system. The model can be used for hind- or fore-casting purposes and to test experimental 'what-if?' scenarios (Coulthard and Macklin, 2001, 2003; Coulthard et al., 2002, 2005; Coulthard and Van De Wiel, 2007, 2013, 2017).

The scope of this study is twofold. The first stage determines how realistically CAESAR-Lisflood can model river channel changes. The selected reaches have multi-decadal overlapping records of river flows and channel patterns. These reaches provide a test for the effectiveness of CAESAR-Lisflood at modelling channel changes through comparison of mapped historical and hind-cast modelled channel changes. Parameter values representing vegetation cover, sedimentology and hydrology were derived from the existing literature and are described further in a later section. An ensemble of runs, testing different combinations of values for parameters governing erosion rates and channel morphology, were tested to calibrate the model. A key part of this model evaluation process was determining the range and variation in selected parameter values between similar reaches.

The second stage involved applying simulated geomorphic changes to determine the total area of occupation of floodplain by the river channel. These results are compared with geomorphic changes derived from mapped datasets to evaluate further

how effectively CAESAR-Lisflood could be used to model floodplain turnover processes over longer timescales. Evaluation is favoured over validation as it is difficult, or even impossible, to account for all uncertainty that exists in nature and quantitative models (Oreskes, 1998). Previous application of CAESAR-Lisflood to a reach of the River Pellice, Italy demonstrated that validation was inhibited by incomplete understanding of natural phenomena, input data uncertainties, and the inability of a reduced-complexity model to replicate nature completely (Pasculli and Audisio, 2015). Furthermore, data limitations prohibit validation here. Only two channel change survey periods exist for each of the 10 sites over the combined historical map and flow record. As the model requires at least a portion of the first survey period to spin up, calibration focuses on the second survey period. Hence, we evaluate the accuracy of modelled channel changes during the calibration period, rather than a validation approach involving calibration using one time period and application to another.

Study sites and model data inputs

Ten 1 km-length valley floor reaches were chosen from across northern England (Figure 1) and modelled using CAESAR-Lisflood version 1.9b. Historical maps show that the selected reaches incorporate a gradient in the rates of geomorphic change. These reaches also exhibit a variety of channel planform characteristics (e.g. sinuosity, meander wavelength, occurrence of multiple channels) and differences in floodplain morphology (e.g. valley width, number of terraces). Nearby flow gauges provide the multi-decadal time series of discharges required to run CAESAR-Lisflood, and these overlap with a rich historical record of channel changes discerned from maps and aerial imagery (Table I).

DEMs obtained from sources listed in Table I were clipped to reach boundaries, each encompassing 1 km of valley length, and were resampled to a coarser grid resolution to expedite simulations (Table II). The contemporary channel was smoothed out using elevation interpolation tools in RasterEdit software (available at https://sourceforge.net/projects/caesar-lisflood/files/) and the historical channel from the start of the simulation period was burned in using the Raster Calculator tool in ArcGIS (Table II).

Nine sediment grain size classes (mm) in the downloaded model were taken originally from the Swale catchment (see Table 1 in Coulthard and Van De Wiel, 2017), and used as the basis for establishing individual separate grain size distributions for channel and floodplain cells (Figure 2) using the 'grainfilemaker' tool (available at https://sourceforge.net/projects/caesar-lisflood/files/). The proportions of each grain size class (see 'Default' in Figure 2) were rearranged for the nine grain sizes to match closely published data on channel bed
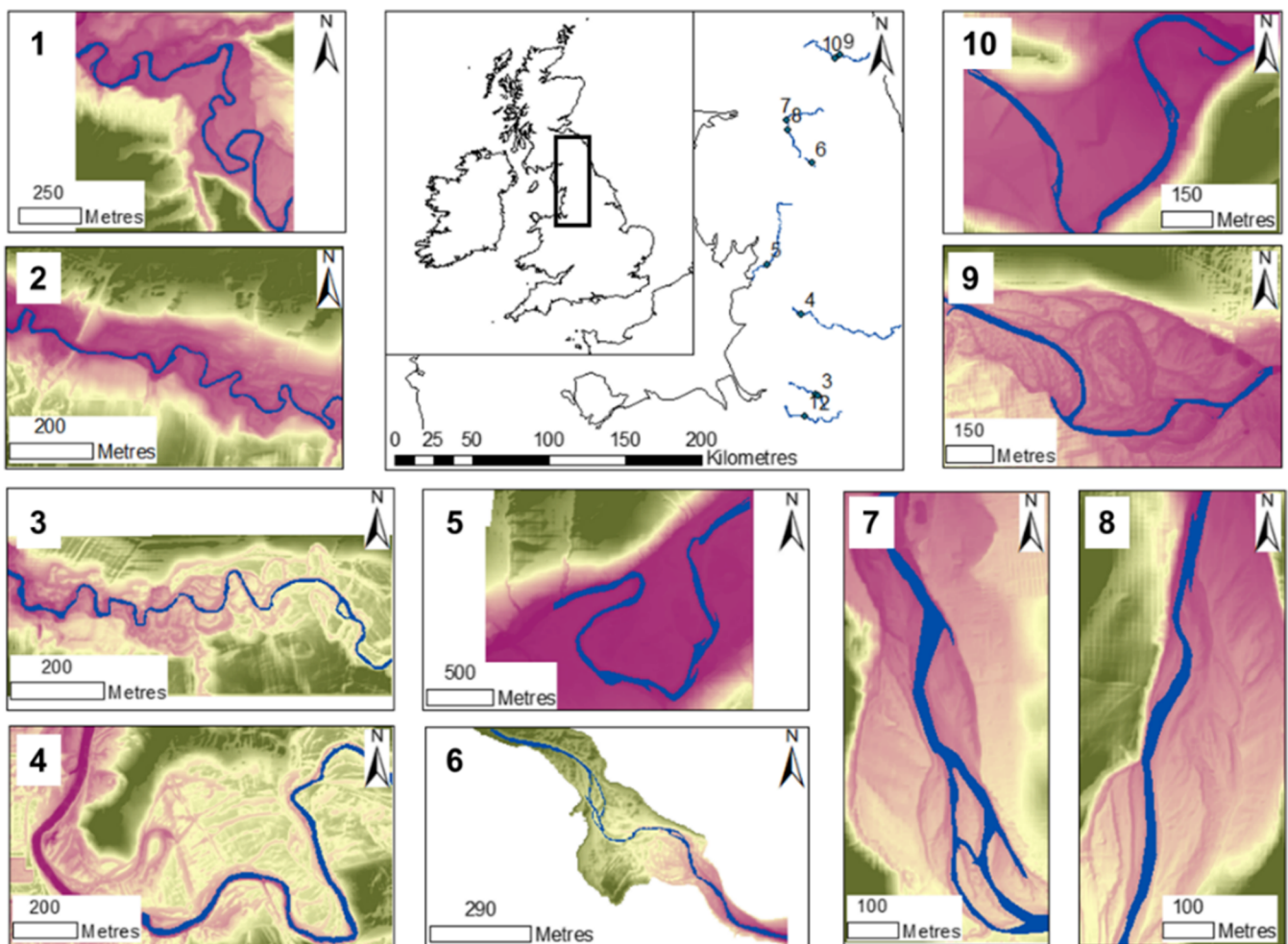
1. Dane, 2. Bollin1, 3. Bollin2, 4. Calder, 5. Lune, 6. Harwood Beck, 7. South Tyne1, 8. South Tyne2, 9. Coquet1, 10. Coquet2
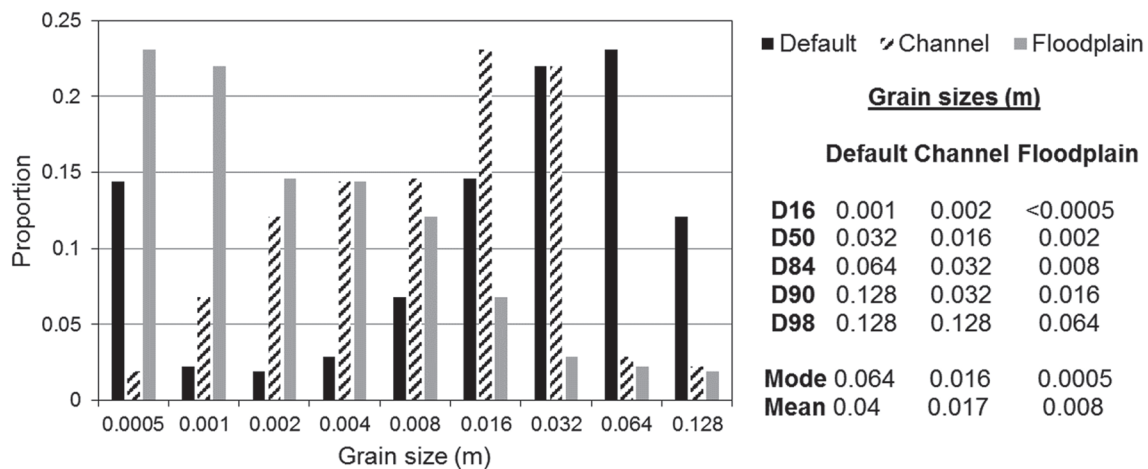
Figure 1. Environment Agency LiDAR 2 m and Ordnance Survey Terrain 5 m elevation models for the study reaches: 1, Dane; 2, Bollin1; 3, Bollin2; 4, Calder; 5, Lune; 6, Harwood Beck; 7, South Tyne1; 8, South Tyne2; 9, Coquet1; 10, Coquet2. © Crown copyright and database rights 2019 Environment Agency and Ordnance Survey. [Colour figure can be viewed at wileyonlinelibrary.com]

Table I. Datasets used in historical channel change reconstruction and model setup

Dataset	Description and purpose	Source
DEMs	DEMs of sites are required inputs for CAESAR-Lisflood. 5–15 m resolution depending on channel width and model run times	Environment Agency (2 m LiDAR) OS Terrain 5 DTM (EDINA Digimap)
Daily flow records	A flow series is required to run CAESAR-Lisflood in reach mode. Daily historical discharge data from the nearest available flow gauge	National River Flow Archive
Historical maps	1:10 000 scale maps. Contemporary OS map and historical map from approximately the start of the daily flow record selected for each site. Where possible, a map from the middle of the flow record period was also acquired	Ordnance Survey (EDINA Digimap)
Aerial imagery	Aerial imagery from the early 2000s chosen to digitize channel morphology from the middle of the available flow record in cases where historical maps were unavailable. 25 cm resolution	Google Earth

Table II. Combined overlapping flow and map/aerial image records for each site, and inputs used to run the model

Site	Flow record used to run the model	Map/image years	DEM resolution (m)
Bollin1	Wilmslow, 01/01/1987–30/09/2014	1987, 2003, 2015	5
Bollin2	Wilmslow, 01/01/1976–30/09/2014	1976, 2003, 2015	5
Calder	Whalley Weir, 01/01/1988–30/09/2015	1988, 2003, 2015	10
Coquet1	Rothbury, 01/01/1976–30/09/2015	1978, 2002, 2015	10
Coquet2	Rothbury, 01/01/1976–30/09/2015	1979, 2002, 2015	10
Dane	Rudheath, 01/01/1969–30/09/2015	1969, 1992, 2015	10
Harwood Beck	Harwood, 01/01/1977–30/09/2015	1977, 2001, 2015	5
Lune	Caton, 01/01/1979–30/09/2015	1974, 2000, 2015	15
South Tyne1	Featherstone, 01/01/1982–30/09/2015	1982, 2003, 2015	15
South Tyne2	Featherstone, 01/01/1982–30/09/2015	1982, 2003, 2015	10

	Default	Channel	Floodplain
D16	0.001	0.002	<0.0005
D50	0.032	0.016	0.002
D84	0.064	0.032	0.008
D90	0.128	0.032	0.016
D98	0.128	0.128	0.064
Mode	0.064	0.016	0.0005
Mean	0.04	0.017	0.008

Figure 2. Grain size distributions for our simulations. The order of size proportions in the Swale catchment distribution was rearranged such that floodplains consisted predominantly of fine sediments (<2 mm) and channels consisted mainly of coarse sands and gravels. If erosion exceeds the combined depth of the surface layer and 10 sub-surface strata, the unaltered 'default' grain size distribution of the Swale will characterize the new surface layer.

and floodplain grain size distributions. This was to account for floodplains and terraces tending to consist of finer-sized sediments and channel beds consisting of coarser material due to bed armouring.

Setting the same grain size distribution for all reaches was intended to provide greater control over drivers of geomorphic change, facilitating the testing of calibrated parameter values on uncalibrated reaches. Published data for reaches of the Coquet, Lune and South Tyne near our chosen sites show comparable grain sizes to each other and to our input grain size distributions (see Fuller et al., 2002; Graham et al., 2005; Wittenberg and Newson, 2005). While descriptions of channel bed and bar grain sizes for the River Dane reflect our channel bed

grain size distribution (Hooke, 2003), local floodplain and terrace material consist mainly of sediments that are finer than the smallest grain size (0.0005 m) in our floodplain grain size distribution (Hooke et al., 1990). This may affect the values selected for parameters governing erosion rates and limit the application of calibrated parameters to experimental modelling.

CAESAR-Lisflood requires time to 'spin up' – a process characterized by the winnowing of fine sediments from the channel bed and an exaggerated rate of geomorphic change and sediment output at the beginning of simulations. To determine the duration of model spin-up, a separate set of simulations for each reach was run using a synthetic hydrograph. Synthetic hydrographs were generated by starting with the maximum
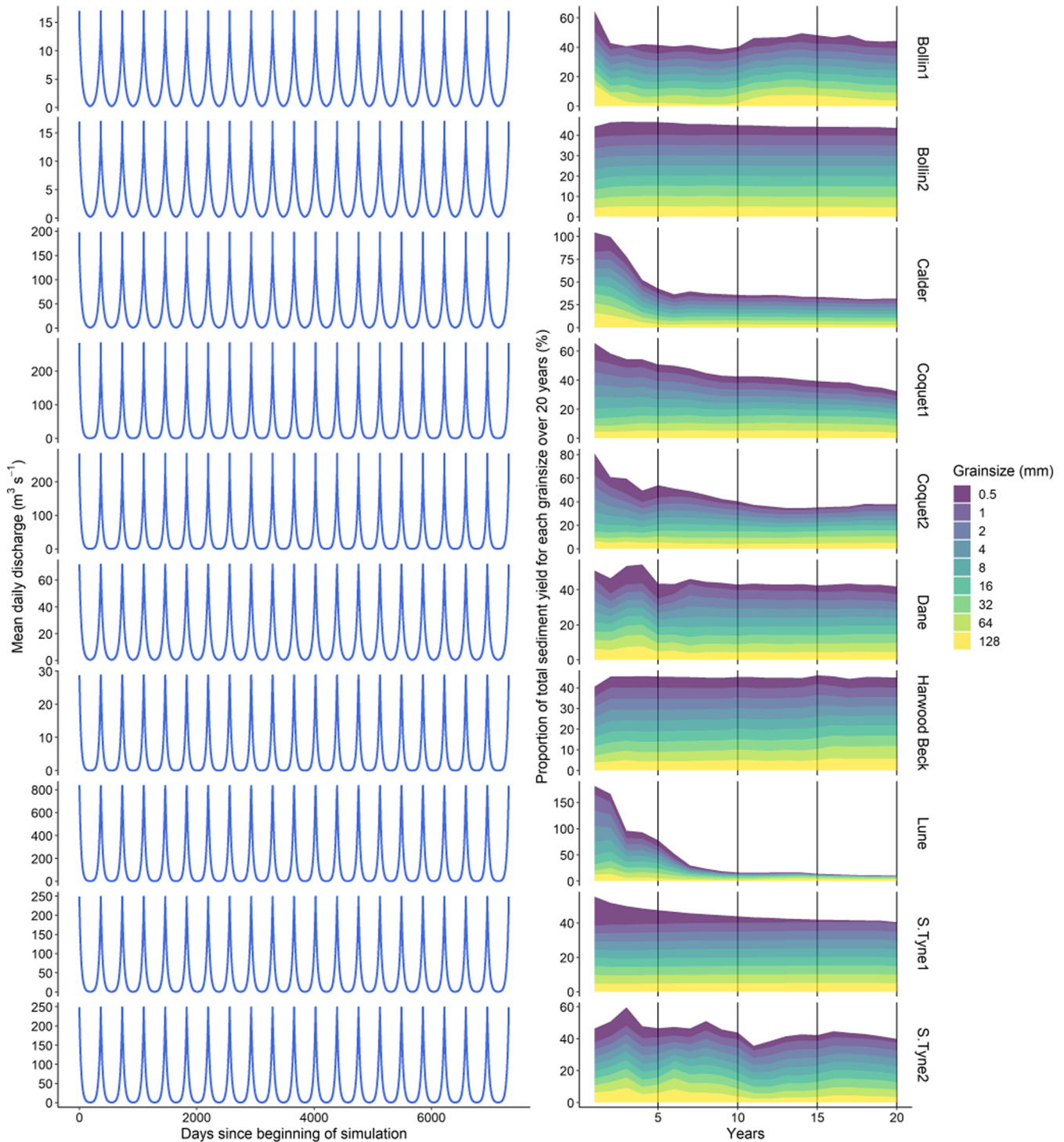
Figure 3. (a) Synthetic hydrographs developed for model spin-up simulations. (b) Modelled annual sediment yields for each grain size from simulations designed to identify the model 'spin-up' time. Annual sediment yields were normalized by the cumulative total sediment yield over 20 years and expressed as percentages. When inter-annual variability in sediment yields stabilizes, spin-up is deemed to have completed. *Note:* Plots for each of the grain sizes are stacked, hence the large upper limits along the y axis. [Colour figure can be viewed at wileyonlinelibrary.com]

discharge in the historic flow series, lowering to the minimum recorded discharge after 182 days, and increasing to the maximum value again at day 365 (see Batz, 2010). This 365-day flow series was repeated to generate a 20-year synthetic flow series with one large flood event per year (Figure 3a). This approach was chosen to ensure that inter-annual variability in sediment yields could not result from any inter-annual variations in discharge. Assessment of inter-annual variability in sediment yields over 20 years (Figure 3b) suggests initially exaggerated rates of sediment flux stabilized within 10 years. This is comparable with results from the Tagliamento (Batz, 2010) and Toutle-Cowlitz (Meadows, 2014) systems. Simulations were therefore run for

the full 25–45 years of flow data in Table II, with evaluation of calibrated model parameters conducted using the latter of the two time intervals, from the 1990s/2000s to present day (see Table II), to ensure model spin-up effects had ceased and were not a component of the results.

Historical daily flow records for the periods listed in Table II were used to drive simulations. Missing data were estimated by linear interpolation between recorded measurements. For the Dane flow series, a gap of more than 8 months (01/06/1978–12/02/1979) was filled using a linear regression between data from both the nearby Ashbrook gauge (River Weaver) and Rudheath gauge (Dane). No historical record of sediment

transport data for the daily flow data taken from the streamflow gauges was available. Therefore, over the course of each simulation, the sediment output recorded at the reach outlet for each model iteration was recirculated into the top of the reach for subsequent iterations to provide upstream sediment fluxes.

Reconstructing historical channel changes

The main river channel within the selected reaches was digitized from the modern Ordnance Survey (OS) map, a historical map dating to the start of the combined map and flow record, and aerial imagery from the middle of the combined record (Table II). Prior to digitizing, aerial imagery was georeferenced to the modern base map. Typically, between 15 and 25 ground control points per image were used, concentrated close to the river channel, and wherever possible, located at 'hard edge' features (*sensu* Hughes et al., 2006), such as the corners of field boundary walls. Photos were georeferenced to the OSGB36 coordinate system using second-order polynomial transformation to reduce the overall error associated with image distortions and then resampled using the nearest-neighbour method. On average, total root mean square error (RMSE) for rectified images was 2.84 m (standard deviation 1.15 m). For the Dane, a second historical map (1992) was used *in lieu* of aerial imagery as the midpoint record (see Supporting Information Part A for further details).

For our CAESAR-Lisflood simulations, DEMs, water depths, grain size information, including daily sediment fluxes, and animation image files were saved annually. For each of the corresponding historical map and aerial image years, DEM and water depth files were first converted from ASCII to raster format. Water depths were subsequently converted to shapefiles and edited to the dimensions of the channel area, based on visual assessment of the DEMs and animation images.

Digitized channels were overlaid to calculate areas of erosion and deposition between years. Areas abandoned by the channel were classified as 'deposition', and areas newly occupied by the channel as 'erosion'. Areas that indicated a combination of erosion and deposition (e.g. floodplain areas between a former and a more recent channel position) were classified as 'erosion and deposition' (Figure 4a). Total eroded and deposited areas (m^2) were converted to rates (m^2 $year^{-1}$) by dividing by the number of years in the interval between the modern and historic channel.

Channel centrelines from each year were extracted from channel polygons and lateral migration rates were calculated from measured migration distances, which were taken along equally spaced transects created perpendicular to the mapped historical migration area centreline (Figure 4b; see also Supporting Information Part B) using the Channel Migration toolbox in ArcGIS (Legg et al., 2014). For each transect that recorded a lateral migration distance (m), values were converted to rates (m $year^{-1}$) in the same way as for erosion and deposition. For all sites, the mean annual lateral migration rate was calculated using data recorded along each transect.

Floodplains and digitized channels from the mid-point and end-years of the record were rasterized to the same DEM grid resolution used in our simulations. For each of the map and model channel–floodplain raster grids, cells were assigned a value depending on whether they were part of the floodplain (2) or channel (1), and these were subtracted from each other
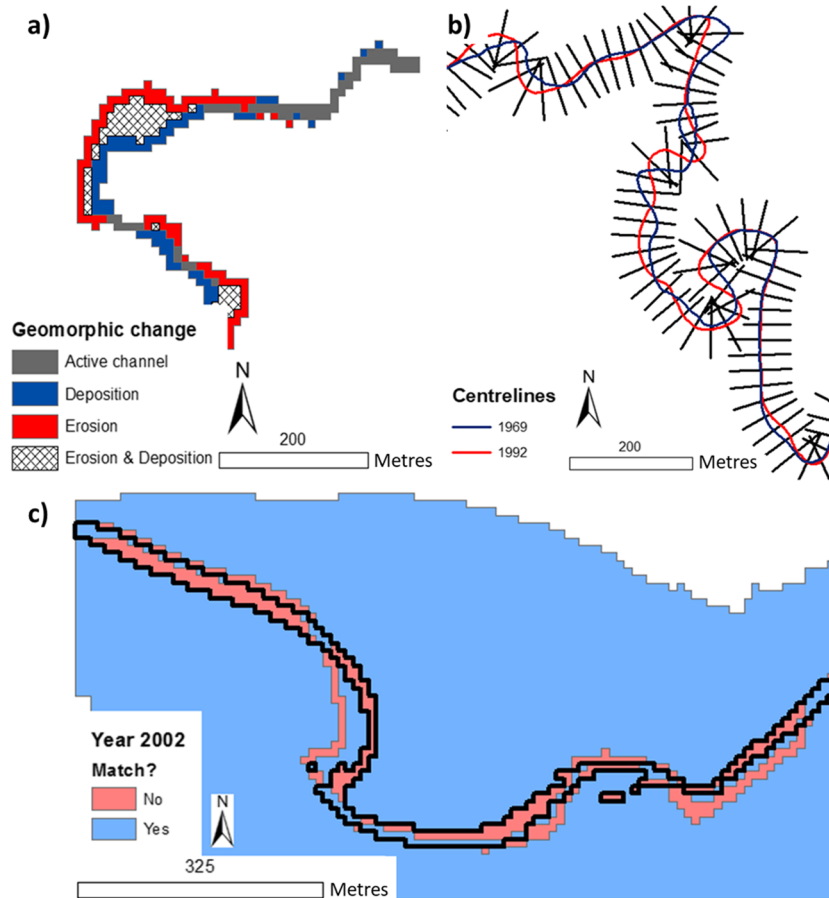

Figure 4. (a) Channel and floodplain geomorphic change classification (example: simulation of Coquet2). (b) Channel centrelines and lateral migration transects (example: two centrelines extracted from historical channels of the River Dane). (c) Map–model raster overlay indicating successful landform cell matches between model and map datasets, with mapped channel boundary highlighted in black (example: Coquet1 reach from 2002). [Colour figure can be viewed at wileyonlinelibrary.com]

in the Raster Calculator. Cells with a value of 0 following sub-traction indicated a match between mapped and modelled landforms (Figure 4c). The number of successful matches of channel and floodplain cells between the model and mapped datasets was divided by reach area and expressed as a percentage.

Model calibration, performance and evaluation

Calibration

The procedure for model calibration, error quantification and selection of parameter values largely follows the framework outlined in a previous CAESAR-Lisflood modelling study by Meadows (2014). Model calibration consisted of two key stages. First, values were selected for the parameters listed in Table III.

Second, different combinations of values were trialled for two equations that make up the lateral erosion algorithm in CAESAR-Lisflood (Coulthard and Van De Wiel, 2006; Van De Wiel et al., 2007). The parameters for these two equations are lateral erosion (θ) and in-channel erosion rate (λ):

$$E_{lat} = \frac{1}{R_{ca}} \theta \tau T \qquad (1)$$

where E_{lat} is the rate at which bank cells are lowered (m timestep^{-1}), R_{ca} is radius of curvature (m), θ is a user-specified lateral erosion parameter (dimensionless), τ is the critical shear stress of the cell next to the channel bank (N m^{-2}) and T is time (s);

$$\Delta Z_{n-1} = V_{n-1} \lambda \frac{(Z_n - Z_{n-1})}{Dx} \qquad (2)$$

where ΔZ is the change in cell elevation (m), V is volume of eroded sediment (m^3), Dx is grid cell size (m), λ is a user-specified in-channel erosion rate (dimensionless), n and $n - 1$ are the donor and recipient cells, respectively, in the context of sediment transfer.

The first equation governs the rate at which the channel laterally migrates. Local radius of bend curvature is determined for each cell by identifying channel edge cells (Coulthard and Van De Wiel, 2006) and determining whether these cells reside along the inside or outside of the meander bend, based on the number of surrounding wet and dry cells (Coulthard et al., 2007). Radius of curvature is then input into Equation (1) to determine the lateral erosion rate. As the accuracy of the radius of curvature depends on the number of passes the edge-smoothing filter makes (Van De Wiel et al., 2007), modelled lateral erosion rates will depend on the choice of values for N_{smooth} and N_{shift} (Table III), as well as θ.

Equation (2) controls channel hydraulic radius and allows sediment to move laterally within the channel, independently of the effects of channel sinuosity resulting from Equation (1) (Coulthard et al., 2013). Higher values of λ allow greater volumes of sediment to transfer between the donor and recipient cells. Effectively, Equation (2) represents sediment cohesion, with lower values of λ indicating boundary materials that are more cohesive, limiting lateral sediment redistribution, and vice versa for higher values of λ (Coulthard et al., 2013).

Some systems exhibit significant mid-channel deposition. A user-defined parameter, C_{max} (Table III), sets the maximum size difference between radius of curvature values for outer bank cells and values for inner bank cells between consecutive smoothing iterations (Van De Wiel et al., 2007). This affects the volume of sediment transferred across the cross-channel radius of curvature gradient over successive model iterations, with higher values facilitating mid-channel deposition and formation of mid-channel bars (Van De Wiel et al., 2007).

During calibration, most sites were run with grass cover, which includes grass, forb and herbaceous plants that are <2 m in height (Lyons et al., 2000), based on observations from aerial imagery. Although the Dane reach consists of a mix of both grass and forest cover, it was judged that forest cover, encompassing trees and shrubs that are ≥2 m in height (Lyons et al., 2000), was more appropriate. Many of the river banks were lined with trees and, based on historical analysis of a 5 km-long river corridor just upstream (Hooke and Chen, 2016), tree cover is shown to be increasing in this area.

Table III describes the three vegetation model parameters and choice of values for forest and grass cover. The highest threshold shear stress values listed in Table 2 of Fischenich (2001) for 'long native grasses' (~80 N m^{-2}) and 'hardwood trees' (~120 N m^{-2}) were chosen for grass and forest, respectively. Literature on the relationship between vegetation maturity and rates of sediment erosion is scant. Some sources state a threshold of about 20 years to distinguish between mature and immature forest (e.g. Trimble, 2004; Fryirs and Brierley, 2013). Given the typical lifecycles of herbaceous plants (Lack and Evans, 2001), we assume that grass reaches full maturity within a year. Thus, maturity times of 1 and 20 years were set for grass and forest, respectively. By default, CAESAR-Lisflood sets the proportion of erosion allowed to occur at full maturity to 0.1. Assuming this parameter is adjustable by increments of 0.1 and mature forest provides the greatest protection from erosion, but allows some erosion to occur, a value of 0.1 was set for forest cover. Based on a relationship of unit riparian biomass to net geomorphic work along river channels (see Figure 9 in Trimble, 2004), grass cover was judged to also protect sediments from erosion by a significant amount, but not quite as much as fully mature, steady-state forest cover. Thus, a value of 0.2 was set here. Manning's n coefficient was adjusted (based on Chow, 1959) to reflect impacts of different vegetation cover types on floodplain hydraulic roughness. Values of 0.035–0.05 and 0.06–0.08 were used for grass and forest, respectively.

No cross-parameterization took place between simulations during calibration (i.e. one parameter at a time was altered, while the others were used as handles). For each reach, a range of values for θ, λ, N_{smooth}, N_{shift} and C_{max} was identified for testing, based on recommendations in the literature (see Table III).

Selection of parameter values

Assessment of model accuracy focused on four key characteristics: surface areal erosion rates, surface areal deposition rates, lateral migration rates, and successful matches between modelled and mapped landforms. Erosion and deposition were judged jointly most important, followed by lateral migration, with successful landform matches least important. Areal erosion and deposition directly quantify the extents of floodplain destruction and creation, respectively. Lateral migration is judged to be the primary control on erosion and deposition. Successful landform matches between modelled and mapped datasets assess how accurately the model reproduces spatial patterns of planform channel change, including active channel width/area. When values for the four criteria for assessing accuracy were calculated for each simulation, they were compared with the corresponding mapped data to calculate absolute relative error values. Absolute error measures ensure both positive (over-estimation) and negative (under-estimation) observations contribute to the overall error and do not cancel each other out (Bennett et al., 2013). For each criterion, error values were compared across calibration runs for a reach and were ranked from 1 to n (smallest to largest error).

Table III. The sediment, vegetation and flow model parameters (modified from Tables 4.3, 4.5 and 4.9 in Meadows, 2014). Guidance for parameter selection comes from the CAESAR-Lisflood website (https://sourceforge.net/p/caesar-lisflood/wiki/Home/) and Coulthard et al. (2013). Details on operation of sediment model parameters are found in Coulthard and Van De Wiel (2006) and Van De Wiel et al. (2007) and for the flow model in Bates et al. (2010) and Coulthard et al. (2013).

Parameter (units)	Operation	Impact and sensitivity	Parameter value selection
Sediment entrainment, transport and redistribution			
Sediment transport equation	Calculates volume of sediment eroded from a cell in a single iteration	Transport rates can vary significantly between the two formulae: Einstein (1950) and Wilcock and Crowe (2003)	Wilcock and Crowe (2003)
v_f (m s^{-1})	Suspended sediment settling velocity	Lower values increase suspended sediment deposition	Estimated using Stokes' law
ΔZ_{max} (m)	Maximum allowable elevation change per cell per iteration	Lower values restrict timestep and enhance model stability	Default = 0.02 used; can be lower for fine-resolution DEMs (10 m)
L_h (m)	Stratum thickness within the active layer system	Low values could affect transport rates through detachment limitation	Must be at least four times ΔZ_{max}. 0.2 used
N_{smooth}	Number of passes made by edge-smoothing filter that is used to calculate radius of curvature	Low values may cause irregular lateral development; high values may lead to over-smoothed channels	Low values for high-sinuosity meandering or braided systems; higher for low-sinuosity systems. Must be an integer. Estimated using meander wavelength (in grid cells)
N_{shift}	Determines number of cells the cross-channel gradient shifts downstream to model downstream bend migration	May lead to unrealistic evolution of meander bends if set incorrectly	One-tenth of N_{smooth} and must be an integer
ΔC_{max}	Calculates cross-channel gradient from radius of curvature which is used to control lateral transfer of eroded sediment from the outer to inner banks	Lower values ensure distribution across the entire channel width, but increase model run time. Higher values encourage mid-channel deposition	Default = 0.0001 used; can be increased or decreased by an order of magnitude depending on channel width or if modelling mid-channel deposition
Vegetation			
τ_{crveg} (N m^{-2})	Threshold bed shear stress for vegetation removal by flows	Higher values relate to greater resistance to removal	Estimated from vegetative strength data[a] For grass cover: 80 For forest cover: 120
T_{veg} (years)	Time required for vegetation to reach full maturity	Higher values negate the role of this over short timescales	Estimated from known vegetation growth rates For grass cover: 1 For forest cover: 20
Proportion of erosion allowable at full maturity	Determines how vegetation maturity affects erosion	0–1 scale: 1 means vegetation has no impact on erosion; 0 means no erosion occurs at full maturity	0–1 (default = 0.1). Estimated by known effects of vegetation on erosion For grass cover: 0.2 For forest cover: 0.1
Hydrology and flow routing			
Q_{min} (m^3 s^{-1})	Run-time optimization. Minimum discharge in a cell necessary for depth to be calculated	Higher values reduce run times but may restrict flow and erosion in peripheral cells if too high	One-tenth of DEM grid resolution (m)
d_{min} (m)	Run-time optimization. Minimum water depth in a cell necessary for erosion to be calculated	Higher values reduce run times but may restrict flow and erosion in peripheral cells if too high	Default = 0.01 used. Can be lowered for grid cells <5 m or increased for cells >50 m
Q_{diff} (m^3 s^{-1})	Run-time optimization. Difference between the input and expected output discharge that can allow the model to shift to steady-state mode	Higher values reduce run times, but if too high can cause too many discharge events to be missed and lead to numerical instabilities	Can be approximated by catchment mean annual flow. User judgement also necessary regarding acceptable difference between input and output discharge and speed of model operation

(Continues)

Table 3. (Continued)

Parameter (units)	Operation	Impact and sensitivity	Parameter value selection
h_{flow} threshold (m)	Run-time optimization. Depth through which water can flow between two cells	Higher values reduce run times but may limit flow when gradient between neighbouring cells is low	Default = 0.00001 used
Courant number	Controls model timestep and affects stability	Higher values increase timestep but increase risk of instabilities such as chequerboarding (rapid flow reversals between cells)	Must be 0.3–0.7. 0.3 chosen as this is recommended for 10 m cells or finer; coarser grids can use larger values
Froude number	Controls model stability and flow rate between cells	Low values reduce speed of flood waves and erosion. High values may induce chequerboarding	Default = 0.8 used. Lower values could be used for very deep, slow flows
Manning's n	Calculation of flow depth and velocity	High values increase flow depths, reduce flow velocities and reduce erosion	Estimated from reach/catchment attributes and relevant literature. Values chosen from reference table (Chow, 1959)
Slope for edge cells	Calculates flow out of the model at the downstream boundary	High values can cause excessive bed scour and upstream-propagating knickpoints; low values can cause excessive deposition at the outlet	Mean bed slope of the channel near the DEM outlet

aRefers to critical shear stress values of 'long native grasses' and 'hardwood trees' presented in Table 2 of Fischenich (2001). These values are recommended permissible thresholds for stream restoration materials and are based on data for natural vegetation.

In order to determine the most accurate model run overall, the analytic hierarchy process (AHP) was used (following Meadows, 2014). The AHP is a decision-making framework whereby multiple criteria of differing importance, derived from user judgement, are compared in a pairwise manner in order to select the best outcome from a range of competing alternatives (Saaty, 1990; Vargas, 1990; Saaty and Vargas, 2012). A goal (i.e. selecting the most accurate simulation from a wider ensemble of calibration runs) and the criteria (see previous paragraph) to meet the goal are defined. A matrix records the judgements in which each criterion is compared with all other criteria. Judgements represent the relative importance of a criterion in the far-left column of the matrix over a criterion in the top row (see example in the Supporting Information) using a 1 to 9 scale (Table IV). All seven possible combinations of 1 to 9 judgement ranks were used whereby erosion and deposition were always judged to be jointly most important, matching landforms least important, and lateral migration only slightly less important (i.e. one rank lower) than erosion and deposition. Assuming this order of importance stays the same, the full range of possible judgement values is considered, eliminating one potential area of user bias.

After all combinations of judgement values were trialled, the resulting weights meant that erosion and deposition each contributed 35.1–39.2%, lateral migration 16.2–18.9%, and landforms matched 5.5–10.9% to the final model score. For each criterion, simulations were ranked $1 - n$ (1 = lowest and n = highest degree of error). The calculated weights were multiplied by the ranks of the corresponding accuracy criteria and summed. The simulation with the smallest cumulative value was ranked best (a rank of 1) in terms of overall model accuracy. The simulation ranked best the most times out of all seven combinations was selected (see Supporting Information Part C for more details of how the AHP method works, including a step-by-step guide using data from the Dane as an example). The AHP was conducted in R using the 'AHP Package' (Glur, 2018).

Evaluating the accuracy of calibrated parameter value ranges
A split-sample testing approach was conducted here. Five of the 10 reaches were reserved for calibration, while the other five reaches were reserved to test the robustness of calibrated parameters more comprehensively. Each of the five reaches reserved for calibration were parameterized individually, regardless of whether some reaches were from the same river system. To evaluate the selection of parameter values more thoroughly, the calibrated lateral erosion rate (θ) and in-channel erosion rate (λ) parameter values from each of the five calibration reaches were tested on the remaining five uncalibrated reaches. Rates of erosion, deposition and lateral migration, and successful landform matches between map and model datasets, were calculated to assess which of the five calibrated parameter combinations was the most accurate, and to what extent, by comparison with the mapped reconstructions.

Modelling floodplain turnover processes
Floodplain turnover can be quantified using a number of measured variables, including the turnover time (the time required for every area of floodplain to be eroded by the channel) and the area of the floodplain that has been occupied by the river channel at least once over time (Figure 5). Quantified erosion rates should be strong predictors of floodplain occupation rates by the channel. However, the 10 sites range widely in a number of physical characteristics, including area (0.14–1.67 km^2), sinuosity (1.04–2.48) and slope (0.002–0.008). Therefore, in order to compare erosion rates and areal extents of floodplain occupation by the channel across sites and between mapped and modelled datasets more appropriately, values for these two

Table IV. The 1–9 scale of importance for pairwise comparison between selection criteria (modified from Saaty, 1986, 2008)

| Rank of importance	Definition	Explanation
1	Equal	Equal contribution of criteria to the objective
2	Weak or slight	—
3	Moderate	Slightly favouring one criterion over another
4	Moderate plus	—
5	Strong	Strongly favouring one criterion over another
6	Strong plus	—
7	Very strong	Very strongly favouring a criterion
8	Very, very strong	—
9	Extreme	Evidence favouring a criterion is of the highest degree of importance over another
Reciprocals of above	If criterion *a* has one of the above rank values assigned to it when compared with criterion *b*, then *b* has the reciprocal value when compared with *a*	—
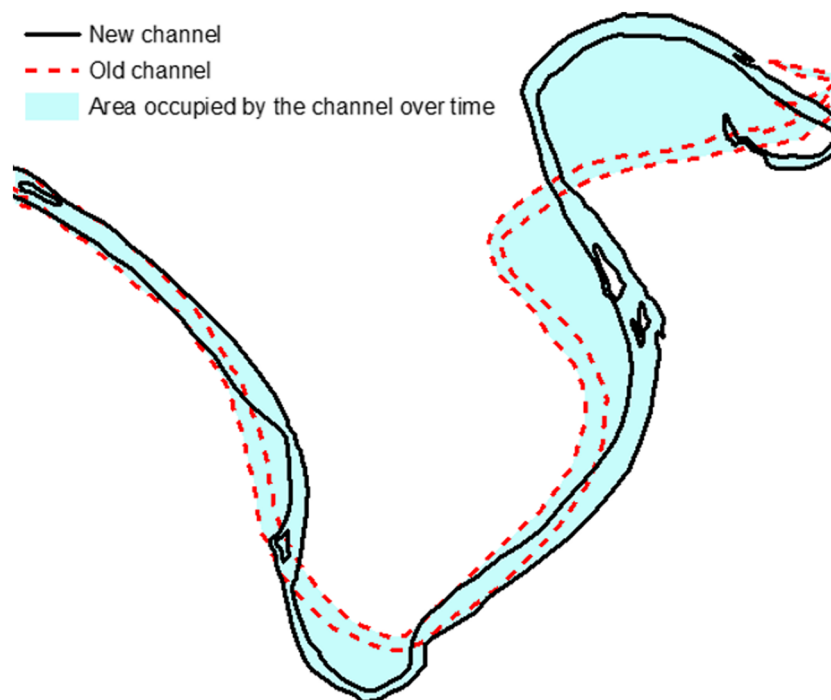
Figure 5. Example of the area occupied by the channel at least once over time. The shaded polygon includes the areas of floodplain occupied by both the old and new channel positions and the area of floodplain that was eroded as the channel migrated from its old to its new position. [Colour figure can be viewed at wileyonlinelibrary.com]

variables were normalized by dividing by the total area of each reach. Then, values derived from mapped reconstructions for these two processes were arranged by decreasing order of size across the 10 reaches to establish gradients. The values derived from modelled reconstructions were subsequently plotted adjacent to these to evaluate how accurately mapped values were reproduced. Relationships between erosion rates and total floodplain area occupied by the channel over time were determined. These relationships were then used to estimate the length of time required for all areas of the floodplain to become occupied by the channel (i.e. the floodplain turnover time) for all the reaches.

Results

Model calibration and performance

On average, 30 simulations were run for each of the five sites reserved for model calibration (total of 154 runs). These included initial test runs to determine the upper and lower bounds for the lateral erosion rate parameter, which for most sites ranged over a single order of magnitude (10^{-5}–10^{-6}). Table V lists selected parameter values for each site from the most accurate simulations. For single channel meandering systems, values of 15–20 for in-channel erosion rate (λ) were found to be most accurate. As the Harwood Beck reach was relatively wide in places (despite being one of the narrowest rivers on average of the 10 sites) and showed evidence of braiding and mid-channel sediment deposition, $\lambda = 23$ was used here.

Results for the four evaluation criteria – erosion rate, deposition rate, lateral migration rate and percentage of matching landforms – demonstrate that simulations replicated historical changes accurately. Erosion rates derived from mapped reconstructions were matched closely by modelled erosion rates across all five sites. Reconstruction accuracy ranged from an over-prediction of 6% along the Lune to an under-prediction of 36% for Coquet2 (Figure 6). Deposition rates along both of

Table V. Summary of calibrated values for the two erosion rate parameters, number of passes for edge-smoothing filter and number of cells the cross-channel gradient shifts downstream for the selected model run (lowest overall error). Test ranges and increments are also given for the two erosion rate parameters (*note:* values in the lateral erosion rate column are multiplied by 10^{-5})

Site		Lateral erosion rate, θ (×10^{-5})	In-channel erosion rate (λ)	No. of passes by edge-smoothing filter (N_{smooth})	No. of cells cross-channel gradient shifts downstream (N_{shift})
Coquet1	Calibrated:	0.1	19		
	Test range:	0.1–1	15–20		
	Increments:	0.1	1	80	8
Coquet2	Calibrated:	0.3	15		
	Test range:	0.05–0.35	15–20		
	Increments:	0.05	1	80	8
Dane	Calibrated:	0.1	14		
	Test range:	0.1–0.12	10–20		
	Increments:	0.01	1	10	1
Harwood Beck	Calibrated:	0.1	23		
	Test range:	—	15–25		
	Increments:	—	1	100	10
Lune	Calibrated:	2	15		
	Test range:	1–5	15–20		
	Increments:	1	1	30	3

the Coquet reaches were reconstructed accurately by the model. However, deposition rates were under-estimated along the Dane and the Lune by as much as 70 and 79%, respectively, and over-estimated for the Harwood Beck by more than 800% (Figure 6). Lateral migration rates were reconstructed accurately overall. Again, the most accurate reconstructions were for the Coquet reaches (over-estimated by 3% for Coquet1 and under-estimated by 11% for Coquet2). The least accurate result was for the Harwood Beck, where mean lateral migration rates were over-estimated by 80% (Figure 6). Successful matches between modelled and mapped landforms are approximately 85% or more on average for all sites (Figure 6).

Evaluation of model parameterization

The five calibrated reaches vary widely in physical characteristics, including reach and channel area, valley slope, number of peak over threshold (POT) events per year, and mean channel sinuosity and thalweg lengths. Reaches reserved for evaluating calibrated parameter value ranges show a similar variability in physical characteristics to the calibrated reaches. Hence, the five reserved reaches should provide an appropriate basis to evaluate the transferability of calibrated parameter values (Table VI).

Figure 7 shows the simulated rates of erosion, deposition, lateral migration and the degree of successful reconstruction of mapped landforms for each of the tested calibrated parameter value combinations when applied to the reserved reaches. Erosion rates of the most accurate simulations (highlighted in red) were over-predicted by ~11% for Bollin2 and under-predicted for the other reaches by ~22–46% (Figure 7). Deposition rates were under-predicted by ~21 and ~23% for the Calder and Bollin1 reaches, respectively. Elsewhere, rates were over-predicted by as much as ~200% (Figure 7). Lateral migration rates were under-predicted by between ~9 and 21% for four reaches and over-predicted by 36% for Bollin2. Mapped landforms were reconstructed to a similar level of accuracy to the calibrated reaches, with successful reconstruction ranging from ~86 to ~93%.

Parameter values calibrated for the two Coquet reaches were found to produce the most accurate simulations for four of the reserved reaches. This would suggest a high potential of transferability in the choice of parameter values for the Coquet reaches. However, the failure to reconstruct mapped data for

untested cases using parameter combinations from other calibrated reaches would suggest the need to calibrate reaches individually. This also applies if the reaches in question come from the same river system.

Floodplain turnover processes

When mapped floodplain erosion rates were normalized by reach area, the upper reaches of the South Tyne (South Tyne2) and Bollin (Bollin2) were shown to be the most actively changing systems over time, with erosion rates of 7.2 and 6.8% of floodplain area per year, respectively (Figure 8). The Dane, Coquet1 and Lune reaches were the least active, with normalized erosion rates approximately four times lower (Figure 8). Modelled erosion rates followed a similar gradient overall, with Bollin2 recording the highest erosion rates and the Dane, Coquet1 and South Tyne1 reaches the lowest rates (Figure 8). Areal extents of floodplain occupation by the channel produced a similar gradient to erosion rates for the mapped dataset. Harwood Beck and the upper reaches of the South Tyne and Bollin had the highest values (>15%), while the Coquet1 reach had the lowest value of <6% (Figure 8). Simulations produced similar results to mapped data, including a similar overall gradient.

Normalized erosion rates were found to correlate positively with the extent of channel occupation of the floodplain for both the mapped and modelled datasets ($R = 0.76$ and 0.74, respectively) (Figure 9a). Results of linear regression indicate that the extent of channel floodplain occupation could be predicted from erosion rates. Linear models could be fitted to both the map and model datasets, with similar gradients and intercepts (Figure 9b).

Assuming that the extent of channel floodplain occupation increases linearly with erosion rates up to 100% floodplain occupation (or turnover), the linear models in Figure 9b could be used to estimate floodplain turnover times (as defined by Bolin and Rodhe, 1973). First, normalized erosion rates (x) are estimated when floodplain occupation by the channel over time is 100% ($y = 100$). The estimated normalized erosion rates when $y = 100$ are then divided by the normalized erosion rates for each site in Figure 8. These values are then multiplied by the measurement interval (years) for the channel change at each reach (e.g. the measurement interval for the Dane, 1992–2015 = 23 years). Estimated floodplain turnover times ranged
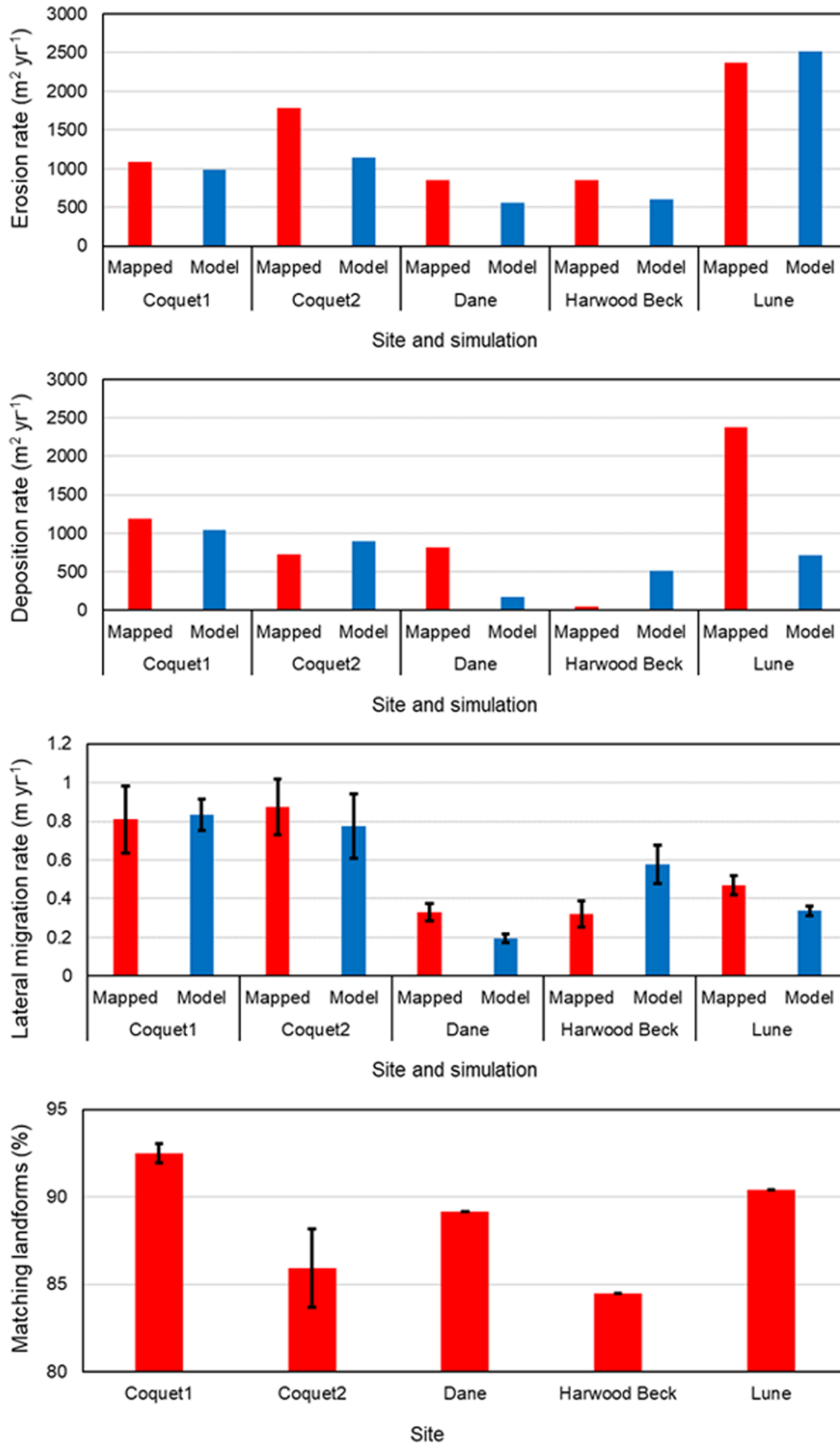
Figure 6. Results of the single most accurate model run for each site compared to data from mapped historical reconstructions for the four evaluation criteria: erosion rate, deposition rate, lateral migration rate and successful landform matches between mapped and modelled datasets. Mean lateral migration rates are calculated from all measurement transects. For matching landforms, the mean is calculated from successful matches at the mid-point and endpoint years. Error bars equal one standard error of the mean. [Colour figure can be viewed at wileyonlinelibrary.com]

widely, between 137 and 1095 years for mapped erosion rates and between 72 and 1080 years for modelled erosion rates. Estimated turnover times were replicated accurately by the model overall. Turnover times for half of the sites were over- or under-

predicted by <20% and under-predicted for four others by <50% (Figure 10). Turnover time for South Tyne1 was over-predicted by the model by nearly 400% (Figure 10). This is un-surprising as modelled erosion rates, using the Coquet1

Table VI. Physical characteristics of the calibrated and reserved reaches. Values relate to mapped datasets. Mean channel area, sinuosity and thalweg length values are calculated for the three mapped channel years listed in Table II. POT stands for peak over threshold events and the mean is calculated based on the number of individual days with a discharge higher than the specified POT flow for the stream gauges listed in Table II

	Reach area (10^3 m^2)	Mean channel area (10^3 m^2)	Valley slope	Mean channel sinuosity	Mean no. POT events per year	Mean thalweg length (m)
Calibrated reaches						
Coquet1	724.9	258.825	0.002	1.33	3	1367.4
Coquet2	423.8	369.65	0.004	1.77	3	1744.4
Dane	474.5	382.591	0.004	2.48	5	2604.8
Harwood Beck	138.65	145.978	0.003	1.15	3	1188.2
Lune	1667.475	2071.74	0.002	2.47	5	4169.6
Reserved reaches						
Bollin1	140.2	120.038	0.002	1.52	1	1507.4
Bollin2	155.775	153.646	0.005	1.84	1	1654.1
Calder	448.2	352.528	0.004	2.01	7	1800
S. Tyne1	440.325	461.276	0.008	1.25	3	1209.1
S. Tyne2	262.9	286.757	0.008	1.04	3	1017

parameter values (Table VII) were eight times lower than rates measured from mapped reconstructions along this reach (see Figure 8). This suggests that despite being the most accurate overall, the Coquet1 reach's calibrated parameters are still a poor fit for South Tyne1, emphasizing further the necessity of calibrating on a reach-specific basis. These results also show how errors in predicted erosion rates by the model can propagate when extrapolated to predict floodplain erosion dynamics over longer timeframes.

Discussion

Model calibration and reconstruction of geomorphic processes

Generally, modelled channel morphology seemed to replicate mapped reconstructions accurately. In some cases, there was a tendency for the size of meander bends to be over-estimated. This is especially true for reaches where channel cut-offs occurred during the studied record. For example, in both Bollin reaches, modelled cut-offs occurred either later than they should have, failed to occur at all along some bends or occurred in the wrong places. However, given the stochastic and self-ordering nature of channel cut-off behaviour (Hooke, 2004; Camporeale et al., 2008), simulated channel cut-offs may be unlikely to occur in exactly the same places and/or at exactly the same time. Hence, it may be satisfactory to settle for similar numbers of cut-offs occurring over the course of the simulation to the compared mapped record. Nevertheless, issues with simulating channel cut-offs have been identified previously for simulations of a reach of the River Teifi (Coulthard and Van De Wiel, 2006). Conversely, for the Coquet2 and Dane reaches, cut-offs formed too easily during calibration trials, limiting the sizes of λ and θ values that could be tested.

In most cases, modelled lateral migration rates were within the error bounds of mapped data. However, there was a pattern of under-estimated modelled mean values for these three rates. Modelled lateral migration distances, recorded along regularly spaced transects, tend to lie within a smaller range of values and are skewed towards lower values (Figure 11). The use of single λ and θ values may restrict the range of lateral migration distances along transects. In particular, the choice of low values for systems prone to forming cut-offs during simulations (e.g. Coquet2 and Dane; Figures 11e and f) would very likely skew

this range towards lower lateral migration rates. It is possible for our simulations that a relatively coarse grid cell resolution (e.g. 10 m) means that a relatively high threshold lateral migration distance (>5 m) is necessary for a floodplain cell to become a channel cell and vice versa. DEM scaling effects have been shown to influence erosion and deposition rates in previous CAESAR-Lisflood applications. For example, soil erosion plot simulations from the Hühnerwasser catchment, Germany revealed a decrease in rill network density with increased grid cell size, as incisions with cross-sections smaller than the cell size could not be initiated (Schneider, 2013). Elsewhere, running simulations using a higher DEM resolution than 2 m was suggested as a way to improve the accuracy of modelled reconstructions of ephemeral gully geometry and spatial dynamics (Hoober et al., 2017).

Channel changes simulated along four out of five of the reaches reserved for evaluating calibrated parameters were reconstructed most accurately using parameters calibrated for the two Coquet reaches. This demonstrates some potential transferability in calibrated parameters from one reach to an untested reach. However, the wide range in calibrated values for parameters listed in Table V, and the results displayed in Figure 7 and Table VII, demonstrate that reach-specific parameterization is necessary with CAESAR-Lisflood. As an example, the Lune required a value for the lateral erosion rate coefficient, θ, that was 20 times larger than the value set for the Dane (Table V), a reach with similar average sinuosity and POT events per year (Table VI) as well as similar channel change behaviour (gradual lateral migration with no avulsions). However, because the Lune had a channel that was ~1.5 times longer and ~5.5 times larger in area than the Dane, much higher values for the lateral erosion rate parameter were needed to generate similar rates in mapped lateral migration.

Channel changes in both Coquet reaches consisted purely of lateral erosion and growth of meander bends. Larger values for the lateral erosion rate parameter for Coquet2 were likely due to a higher mapped annual erosion rate than recorded in Coquet1 (Figure 6). The lower λ setting reflected a balance between maximizing the size of θ to drive higher lateral migration rates, while preventing an avulsion across the neck of the large central meander loop in Coquet2.

Uncertainties in mapped channel changes complicate the evaluation of the accuracy of modelled channel changes. Map projection errors, feature exaggerations and/or distortions, projection errors, georectification errors of aerial imagery and channel digitization errors combine with the outcome that mapped lateral migration, erosion and deposition rates are
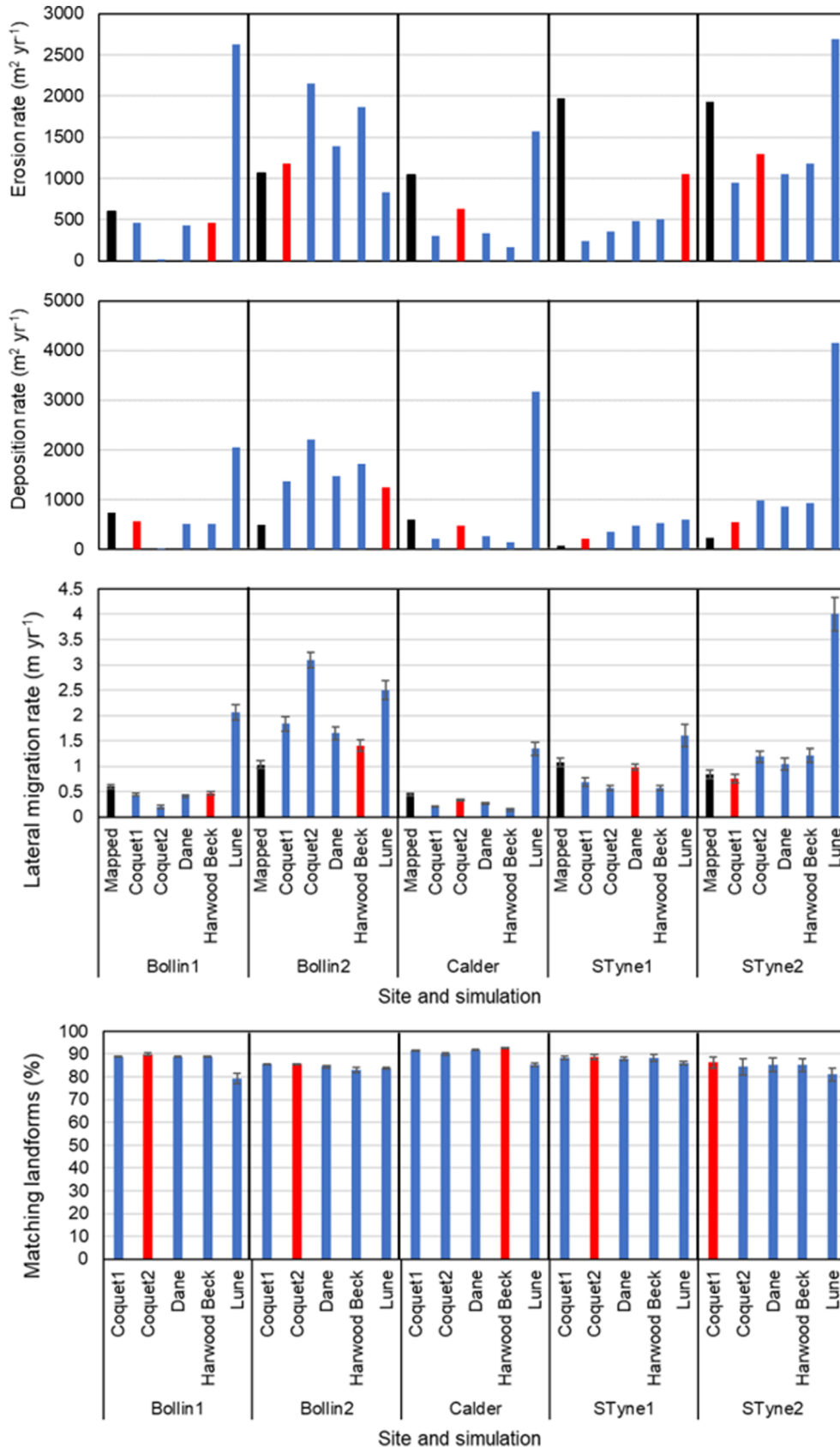
Figure 7. Results of split-sample test, running the five reserved reaches (Bollin1, Bollin2, Calder, S. Tyne1 and S. Tyne2) with parameter values from the calibrated sites (Coquet1, Coquet2, Dane, Harwood Beck and Lune). The most accurate result is highlighted in red. Error bars equal one standard error of the mean. [Colour figure can be viewed at wileyonlinelibrary.com]

over-estimated to some degree (see Downward et al., 1994). However, the degree of over-estimation is difficult to quantify precisely, due in part to the variable nature of historical sources and the accuracy of channel digitization by the analyst (Downward et al., 1994). Most of the 10 sites displayed slight under-

predictions in erosion, deposition and lateral migration rates compared to the mapped reconstructions. Considering the likely over-estimation in mapped geomorphic change rates, modelled geomorphic changes may be more accurate than shown in Figures 6 and 7. However, where modelled changes
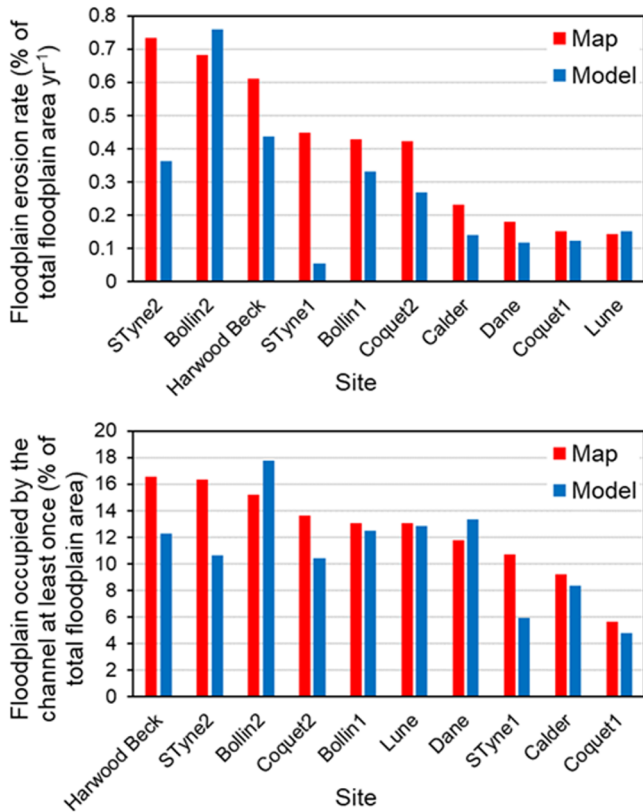
Figure 8. Erosion rates and estimated areal extents of floodplain occupation derived from mapped mean erosion rates. Data for these two dependent variables are normalized by floodplain area. Data in each plot are ordered from largest to smallest based on mapped data (red) to establish a gradient across the 10 reaches, with modelled data (blue) adjacent for comparison. [Colour figure can be viewed at wileyonlinelibrary.com]

Figure 9. (a) Pearson's correlations for mapped and modelled data between erosion rate and floodplain occupation by the channel (after both of these variables have been normalized by reach area). (b) Linear regression results to predict floodplain (re)occupation extents by the channel from erosion rates.

were over-estimated compared to the mapped reconstructions (e.g. Lune erosion rates and Harwood Beck deposition rates), the modelled results may be less accurate than portrayed in Figures 6 and 7.

As sediments are transferred between channels and adjoining floodplain, there is the potential for uneven exchange in terms of sediment volumes and/or grain size distributions. For instance, as a channel migrates laterally, coarse sediment transported as bedload can be deposited as point bars, while comparatively finer material is eroded from the floodplain (Lauer and Parker, 2008). Point bars tend to contain a lower volume of sediment than material eroded from the cutbank, as these develop at lower elevations and because of the effects of channel curvature. Unless this net imbalance is accounted for by other processes (e.g. overbank deposition), the channel–floodplain system could be out of equilibrium, and if this is the case with our simulations, it could call into question our calibrations. CAESAR-Lisflood is designed to conserve the mass of sediment in each size class (Coulthard et al., 2013), meaning that any variability between upstream and downstream grain size distributions will be accounted for by corresponding changes in sediment volumes within the reach. To test this independently, the saved output discharge and sediment flux file from the end of the original Coquet1, Coquet2, Dane and Lune calibration simulations were used as inputs to drive new simulations (with sediment recirculation switched off). Grain size distributions for the input and output sediment transport files were compared for high flow days selected to maximize transport of coarse sediment (discharges > 5% exceedance probability; NRFA, 2019), spaced through the
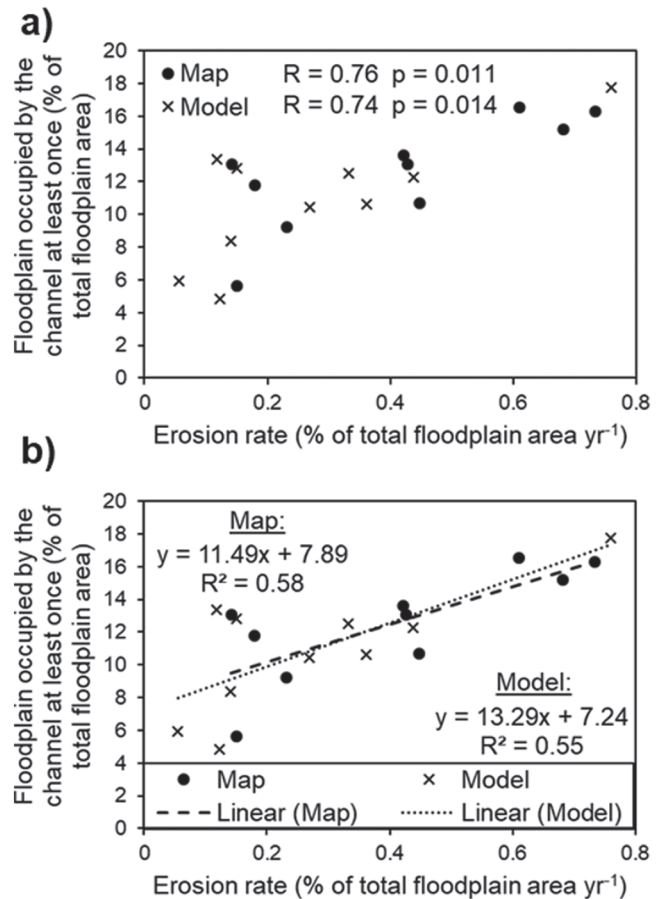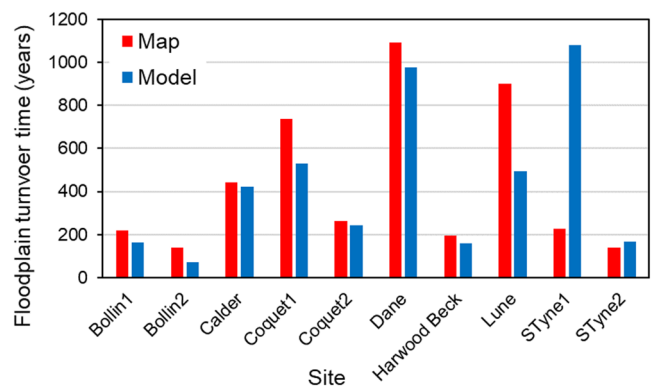

Figure 10. Estimated times to complete floodplain occupation (turnover). Turnover times for the map and model datasets are estimated by solving the linear equations from Figure 9b when $y = 100$ (complete floodplain erosion), dividing by the normalized floodplain erosion rates presented in Figure 8, and multiplying by the number of years in the channel change measurement interval for each site. [Colour figure can be viewed at wileyonlinelibrary.com]

simulations at ~1, 2, 4, 8, 16 and 32 years. In general, output sediment fluxes showed similar distributions to input sediment fluxes, indicating that uneven sediment exchanges were not a feature (or were handled mass-conservatively by CAESAR-Lisflood) during model calibration (see Supporting Information Part D for further details).

Table VII. The most accurate combination of calibrated parameter values for each of the reserved reaches

Reserved reach	Reach with the most accurate calibrated parameters overall when applied to the reserved reach	Lateral erosion rate, θ ($\times 10^{-6}$)	In-channel erosion rate, λ
Bollin1	Harwood Beck	1	23
Bollin2	Coquet1	1	19
Calder	Coquet2	3	15
South Tyne1	Coquet1	1	19
	Coquet1	1	19
South Tyne2	Coquet2	3	15

Floodplain turnover processes

Our simulations captured accurately reach-specific turnover times and variability across reaches for the most part. Modelled estimates of floodplain turnover time corresponded well with mapped estimates, with values ranging over an order of magnitude (from ~100 to ~1000 years). Turnover times for each site, estimated from mapped erosion rates and using the linear relationship of these fitted in Figure 9b, were reproduced accurately in most cases by the modelled erosion rates and the resultant linear model of these data. On average, the margin of error between turnover times estimated from mapped erosion rates and from modelled erosion rates was between ~15 and ~20%. However, under-predicted erosion rates of ~85% resulted in modelled turnover times exceeding mapped turnover times by ~400% along the South Tyne1 reach. These results demonstrate that for the most part, erosion rates predicted from calibrated parameter values over decadal timescales can be extrapolated to simulate accurately longer-term floodplain evolution (over centennial to millennial timescales). This increases confidence in the choice of parameter values for most of the 10 reaches.

Our estimation of the extent of channel floodplain occupation is predicated on two assumptions: (i) all areas of the floodplain will be eroded with equal probability; (ii) rates of geomorphic change will remain constant through time. However, several studies have shown that these assumptions do not necessarily hold. Floodplain vegetation age reconstructions along the Morice River, Canada revealed that reoccupation of abandoned channels occurred much more frequently than the creation of new channels (Gottesfeld and Johnson Gottesfeld, 1990). Analysis of several valley floor systems from North America indicates that the probability of a floodplain area becoming (re)occupied by the river channel can be predicted from floodplain age using a power-law relationship (Konrad, 2012). This is also corroborated by reconstructions of channel–floodplain dynamics in both field (Phillips et al., 2007; Miller and Friedman, 2009) and numerical modelling (Bradley and Tucker, 2013) contexts. If parts of the floodplain were to remain uneroded for longer than the mass balance equation of floodplain turnover time – area/erosion rate (Bolin and Rodhe, 1973) – predicts, we would expect to see a decrease in the proportion of new floodplain eroded with each successive timestep (O'Connor et al., 2003). There would hence be under-quantification for areas of the floodplain unvisited by the channel, and over-quantification in areas where repeated rapid cycling of erosion and deposition occurs (Miller and Friedman, 2009). Due to the limited timeframe of our simulations (only one channel change measurement interval per reach), the reoccupation of abandoned channels by back-and-forth channel movements is not recorded at all. Therefore, while it appears highly likely that repeated reoccupation of younger surfaces would occur in our 10 tested reaches, longer-term simulations, employing methods that captured repeated reworking of floodplain patches, are required to verify this for certain.

Bradley and Tucker (2013) and Torres et al. (2017) demonstrated, through simulating several thousand years of river channel changes, that some parts of the floodplain were reworked many times while the channel would abandon other parts, leaving them unoccupied indefinitely. This is unsurprising for some systems, given the presence of alluvial terraces that may be thousands of years or more in age – including along some of the tested reaches here (e.g. the Dane and the South Tyne). As part of our assessment of CAESAR-Lisflood's accuracy in reconstructing historical channel changes, channel polygons were overlain to determine total areas of floodplain erosion and deposition. This could be expanded to calculate age and storage time (the length of time until sediment is released from storage) values for every floodplain cell and timestep. In fact, CAESAR-Lisflood would offer significant advantages over earlier studies. For instance, it would be possible to model sediment storage behaviour arising from multi-thread channel patterns.

Implications and suggestions for future research

Trimble (2010) characterized sediment storage in alluvial floodplains as 'fast in, slow out', whereby sediments accumulate rapidly (decades to centuries) via vertical and lateral accretion, and sediments are removed via lateral erosion processes over much longer timeframes (centuries to millennia). Based on this, floodplains store significant quantities of sediment, and often for extensive periods of time, delaying their delivery to catchment outlets and introducing lags between erosion signals upstream and their appearance in the stratigraphic record (Hoffmann, 2015). Floodplains, as significant 'shredders' of upstream erosion signals, are diverse in terms of their processes and forms and are liable to change as environmental conditions (e.g. streamflow and vegetation cover) change (Nanson and Croke, 1992). Our findings have demonstrated that floodplain turnover, including erosion rates and areal extents of floodplain occupation by river channels, can vary widely across different sites, encompassing differences in morphology (e.g. area, valley slope, channel pattern). Reach-scale floodplain sediment storage is critical for models of particle trajectories and travel times through valley floor systems (Pizzuto et al., 2017). Modellers simulating sediment transit through valley floor corridors must incorporate the high variability of floodplain sediment storage dynamics into their models. This is particularly important considering some sediments will become incorporated into chronic or permanent storage, and may subsequently be liberated from storage as environmental conditions change (Hoffmann, 2015).

Our model calibration and performance assessment demonstrate the accuracy of parameterized erosion rate equations in driving 'realistic' dynamics of channel changes in CAESAR-Lisflood. These calibrated parameters can be applied using a
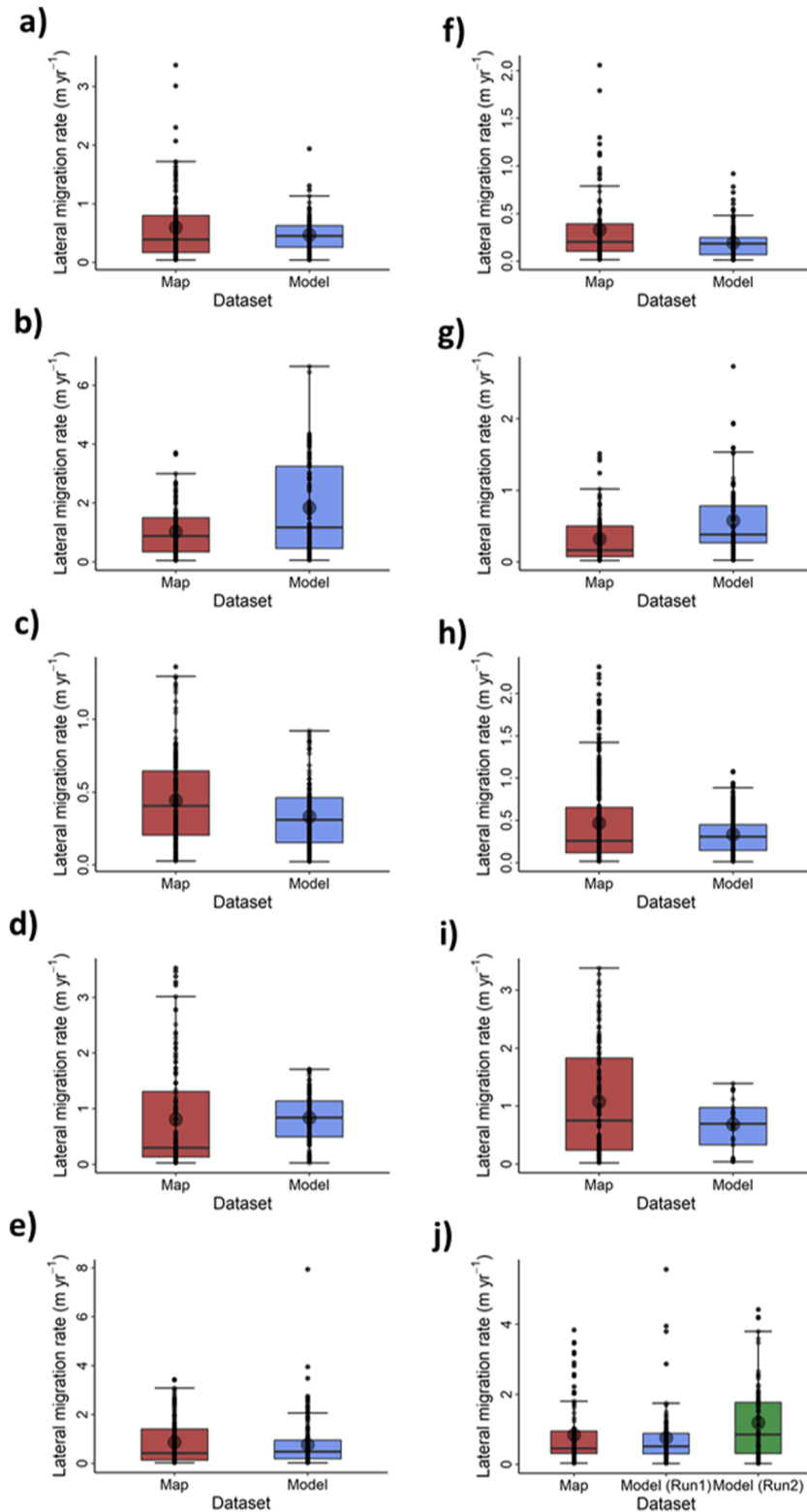
Figure 11. Annual lateral migration rates calculated for map and model channel changes along equally spaced transects (every 10 m along a mapped historical meander belt centreline). Boxplots for both time intervals (big circles represent mean values) for all measured lateral migration rates from top left to bottom right: (a) Bollin1, (b) Bollin2, (c) Calder, (d) Coquet1, (e) Coquet2, (f) Dane, (g) Harwood Beck, (h) Lune, (i) South Tyne1 and (j) South Tyne2. *Note:* Two simulations were found to be jointly most accurate for South Tyne2. [Colour figure can be viewed at wileyonlinelibrary.com]

DEM, grain size distribution and flow series to drive channel changes over much longer timeframes (10^2–10^4 years). Taking this longer perspective could form a series of modelling experiments exploring how different river systems rework their floodplains (e.g. the development of alluvial terraces) and assessing the impacts of environmental changes such as variations in vegetation cover and flood magnitudes on channel–floodplain systems.

Conclusion

CAESAR-Lisflood was applied here to reconstruct river channel changes successfully, including key geomorphic processes of erosion, deposition, lateral migration and landform reconstruction. The robustness of ranges of calibrated parameter values was also demonstrated through successful application of calibrated parameter values to untested reaches via a split-sample

testing approach to reconstruct geomorphic changes accurately. CAESAR-Lisflood was used to predict the extent of channel floodplain occupation from erosion rates, with simulations producing a similar relationship between these two variables to that derived from mapped reconstructions. Key conclusions include the following.

(1) CAESAR-Lisflood has been used to reconstruct geomorphic changes, including channel planform and erosion, deposition and lateral migration rates, of 10 alluvial reaches from across the north of England. This application demonstrates both the feasibility and convenience in parameterizing the model to specific real-world sites, and the utility of such sites and calibrated parameters as templates for virtual flume laboratory settings where experimental modelling can be undertaken. Although our analysis reveals that parameters calibrated for one reach can be applied to model accurately channel changes along a similar reach, we argue that the lack of transferability of parameter values from most of the calibrated reaches shows that reach-specific calibration is needed to produce accurate simulations.

(2) Normalized erosion rates show a positive correlation with the extent of channel floodplain occupation. Floodplain turnover times, estimated using linear models derived from the relationship between erosion rates and floodplain occupation extents, are reconstructed accurately by the model for most sites.

(3) Our results demonstrate that CAESAR-Lisflood has utility in both simulating river channel changes accurately and quantifying longer-term dynamics such as the role of floodplains in lagging signals of sediment supply downstream. Our calibrated parameter values provide a basis for further simulation of channel change and floodplain turnover over much longer time periods. We suggest that further research should focus on quantifying sediment storage timescales by modelling the timing, location and spatial extent of channel occupation, with particular emphasis on quantifying the effects of environmental change on the longevity of sediment storage.

Conflict of Interest Statement

The authors declare no conflict of interest in the work undertaken in this manuscript.

Data Availability Statement

Research data are not shared.

References

Aalto R, Lauer JW, Dietrich WE. 2008. Spatial and temporal dynamics of sediment accumulation and exchange along Strickland River floodplains (Papua New Guinea) over decadal-to-centennial timescales. *Journal of Geophysical Research: Earth Surface* **113**(1): 1–22.

Bates PD, Horritt MS, Fewtrell TJ. 2010. A simple inertial formulation of the shallow water equations for efficient two-dimensional flood inundation modelling. *Journal of Hydrology* **387**: 33–45.

Batz N. 2010. *Predicting ecological diversity of floodplains using a hydromorphic model (CAESAR)*. accessed date December 17, 2019. Available at: http://edepot.wur.nl/165143.

Beechie TJ, Liermann M, Pollock MM, Baker S, Davies J. 2006. Channel pattern and river-floodplain dynamics in forested mountain river systems. *Geomorphology* **78**(1–2): 124–141.

Bennett ND, Croke BFW, Guariso G, Guillaume JHA, Hamilton SH, Jakeman AJ, Marsili-Libelli S, Newham LTH, Norton JP, Perrin C, Pierce SA, Robson B, Seppelt R, Voinov AA, Fath BD, Andreassian V. 2013. Characterising performance of environmental models. *Environmental Modelling and Software* **40**: 1–20.

Bird G, Brewer PA, Macklin MG, Nikolova M, Kotsev T, Mollov M, Swain C. 2009. Dispersal of contaminant metals in the mining-affected Danube and Maritsa drainage basins, Bulgaria, Eastern Europe. *Water, Air, and Soil Pollution* **206**(1–4): 105–127.

Blöthe JH, Korup O. 2013. Millennial lag times in the Himalayan sediment routing system. *Earth and Planetary Science Letters* **382**: 38–46.

Bolin B, Rodhe H. 1973. A note on the concepts of age distribution and transit time in natural reservoirs. *Tellus A* **25**(1): 58–62.

Bradley DN, Tucker GE. 2013. The storage time, age, and erosion hazard of laterally accreted sediment on the floodplain of a simulated meandering river. *Journal of Geophysical Research: Earth Surface* **118**(3): 1308–1319.

Camporeale C, Perucca E, Ridolfi L. 2008. Significance of cutoff in meandering river dynamics. *Journal of Geophysical Research: Earth Surface* **113**: 1–11. F01001.

Castelltort S, Van Den Driessche J. 2003. How plausible are high-frequency sediment supply-driven cycles in the stratigraphic record? *Sedimentary Geology* **157**(1–2): 3–13.

Chow VT. 1959. *Open Channel Hydraulics.* McGraw-Hill: New York.

Coulthard TJ, Macklin MG. 2001. How sensitive are river systems to climate and land-use changes? A model-based evaluation. *Journal of Quaternary Science* **16**(4): 347–351.

Coulthard TJ, Macklin MG. 2003. Modeling long-term contamination in river systems from historical metal mining. *Geology* **31**(5): 451–454.

Coulthard TJ, Van De Wiel MJ. 2006. A cellular model of river meandering. *Earth Surface Processes and Landforms* **31**(1): 123–132.

Coulthard TJ, Van De Wiel MJ. 2007. Quantifying fluvial non linearity and finding self organized criticality? Insights from simulations of river basin evolution. *Geomorphology* **91**(3–4): 216–235.

Coulthard TJ, Van De Wiel MJ. 2013. Climate, tectonics or morphology: what signals can we see in drainage basin sediment yields? *Earth Surface Dynamics* **1**(1): 13–27.

Coulthard TJ, Van De Wiel MJ. 2017. Modelling long term basin scale sediment connectivity, driven by spatial land use changes. *Geomorphology* **277**: 265–281.

Coulthard TJ, Macklin MG, Kirkby MJ. 2002. A cellular model of Holocene upland river basin and alluvial fan evolution. *Earth Surface Processes and Landforms* **27**(3): 269–288.

Coulthard TJ, Lewin J, Macklin MG. 2005. Modelling differential catchment response to environmental change. *Geomorphology* **69**(1–4): 222–241.

Coulthard TJ, Hicks DM, Van De Wiel MJ. 2007. Cellular modelling of river catchments and reaches: advantages, limitations and prospects. *Geomorphology* **90**(3–4): 192–207.

Coulthard TJ, Neal JC, Bates PD, Ramirez J, de Almeida GAM, Hancock GR. 2013. Integrating the LISFLOOD-FP 2D hydrodynamic model with the CAESAR model: implications for modelling landscape evolution. *Earth Surface Processes and Landforms* **38**(15): 1897–1906.

Dennis IA, Coulthard TJ, Brewer P, Macklin MG. 2009. The role of floodplains in attenuating contaminated sediment fluxes in formerly mined drainage basins. *Earth Surface Processes and Landforms* **34**: 453–466.

Downward SR, Gurnell AM, Brookes A. 1994. A methodology for quantifying river channel planform change using GIS. *IAHS Publications* **224**: 449–456.

Einstein HA. 1950. *The Bed-load Function of Sediment Transportation in Open Flows.* US Department of Agriculture: Washington, D.C.

Fischenich C. 2001. *Stability Thresholds for Stream Restoration*. Engineer Research and Development Center: Vicksburg, MS.

Fryirs, K. A. & Brierley, G. J. 2013. Geomorphic Analysis of River Systems An Approach to Reading the Landscape, Chichester: John Wiley & Sons Ltd.

Fuller IC, Passmore D, Heritage D, Brewer P. 2002. Annual sediment budgets in an unstable gravel-bed river: the River Coquet, northern England. *Geological Society of London Special Publications* **191**(1): 115–131.

Glur C. 2018. Package 'ahp'. Available at https://cran.r-project.org/web/packages/ahp/ahp.pdf [].

Gottesfeld AS, Johnson Gottesfeld LM. 1990. Floodplain dynamics of a wandering river, dendrochronology of the Morice River, British Columbia, Canada. *Geomorphology* **3**(2): 159–179.

Graham DJ, Rice SP, Reid I. 2005. A transferable method for the automated grain sizing of river gravels. *Water Resources Research* **41**(7): 1–12. W07020.

Hoffmann T. 2015. Sediment residence time and connectivity in non-equilibrium and transient geomorphic systems. *Earth-Science Reviews* **150**: 609–627.

Hoober D, Svoray T, Cohen S. 2017. Using a landform evolution model to study ephemeral gullying in agricultural fields: the effects of rainfall patterns on ephemeral gully dynamics. *Earth Surface Processes and Landforms* **42**(8): 1213–1226.

Hooke JM. 1980. Magnitude and distribution of rates of river bank erosion. *Earth Surface Processes* **5**(2): 143–157.

Hooke J. 2003. Coarse sediment connectivity in river channel systems: a conceptual framework and methodology. *Geomorphology* **56**(1–2): 79–94.

Hooke JM. 2004. Cutoffs galore!: occurrence and causes of multiple cutoffs on a meandering river. *Geomorphology* **61**(3–4): 225–238.

Hooke J, Chen H. 2016. Evidence of increase in woody vegetation in a river corridor, Northwest England, 1984–2007. *Journal of Maps* **12**(3): 484–491.

Hooke JM, Harvey AM, Miller SY, Redmond CE. 1990. The chronology and stratigraphy of the alluvial terraces of the River Dane Valley, Cheshire, NW England. *Earth Surface Processes and Landforms* **15**(8): 717–737.

Hughes ML, McDowell PF, Marcus WA. 2006. Accuracy assessment of georectified aerial photographs: implications for measuring lateral channel movement in a GIS. *Geomorphology* **74**(1–4): 1–16.

Konrad CP. 2012. Reoccupation of floodplains by rivers and its relation to the age structure of floodplain vegetation. *Journal of Geophysical Research: Biogeosciences* **117**(G4): 1–15.

Lack AJ, Evans DE. 2001. *Plant Biology*. Springer-Verlag: New York.

Lauer JW, Parker G. 2008. Net local removal of floodplain sediment by river meander migration. *Geomorphology* **96**(1–2): 123–149.

Legg NT, Heimburg C, Collins BD, Olson PL. 2014. *The Channel Migration Toolbox: ArcGIS® Tools for Measuring Stream Channel Migration*. accessed date: December 17, 2019. Available at https://fortress.wa.gov/ecy/publications/SummaryPages/1406032.html.

Leopold LB, Wolman MG, Miller JP. 1964. *Fluvial Processes in Geomorphology*. Courier Dover Publications: Mineola, NY.

Lewin J, Macklin MG. 1987. Metal mining and floodplain sedimentation in Britain. In *International Geomorphology 1986, Proceedings of the First International Conference on Geomorphology Part 1*, Gardiner V (ed). Wiley: Chichester; 1008–1027.

Lyons J, Thimble SW, Paine LK. 2000. Grass versus trees: managing riparian areas to benefit streams of central north america. *Journal of the American Water Resources Association* **36**(4): 919–930.

Macklin MG, Dowsett RB. 1989. The chemical and physical speciation of trace metals in fine grained overbank flood sediments in the Tyne basin, north-east England. *Catena* **16**(2): 135–151.

Meadows T. 2014. *Forecasting long-term sediment yield from the Upper North Fork Toutle River, Mount St Helens, USA*. PhD thesis, University of Nottingham.

Miller JR, Friedman JM. 2009. Influence of flow variability on floodplain formation and destruction, Little Missouri River, North Dakota. *Geological Society of America Bulletin* **121**(5–6): 752–759.

Mitsch WJ, Gosselink JG. 2015. *Wetlands*, 5th edn. Wiley: Hoboken, NJ.

Nanson GC, Croke JC. 1992. A genetic classification of floodplains. *Geomorphology* **4**(6): 459–486.

NRFA. 2019. *National River Flow Archive*. Available at https://nrfa.ceh.ac.uk/ [accessed 17 December 2019].

O'Connor JE, Jones MA, Haluska TL. 2003. Flood plain and channel dynamics of the Quinault and Queets Rivers, Washington, USA. *Geomorphology* **51**(1–3): 31–59.

Oreskes N. 1998. Evaluation (not validation) of quantitative models. *Environmental Health Perspectives* **106**(Suppl. 6): 1453–1460.

Pasculli A, Audisio C. 2015. Cellular automata modelling of fluvial evolution: real and parametric numerical results comparison along River Pellice (NW Italy). *Environmental Modeling and Assessment* **20**(5): 425–441.

Phillips JD, Marden M, Gomez B. 2007. Residence time of alluvium in an aggrading fluvial system. *Earth Surface Processes and Landforms* **32**: 307–316.

Pizzuto J, Keeler J, Skalak KJ, Karwan DL. 2017. Storage filters upland suspended sediment signals delivered from watersheds. *Geology* **45**(2): 151–154.

Saaty TL. 1990. How to make a decision: the analytic hierarchy process. *European Journal of Operational Research* **48**(1): 9–26.

Saaty TL, Vargas LG. 2012. *Models, Methods, Concepts & Applications of the Analytic Hierarchy Process*, 2nd edn. Springer Science & Business Media: Berlin.

Schneider A. 2013. *Spatial and temporal development of sediment mass balances during the initial phase of landform evolution in a small catchment*. accessed date December 17, 2019. Available at http://webdoc.sub.gwdg.de/ebook/serien/qa/GeoRS/03.pdf.

Schumm SA. 1977. *The Fluvial System*. Wiley: New York.

Sutfin NA, Wohl EE, Dwire KA. 2015. Banking carbon: a review of organic carbon storage and physical factors influencing retention in floodplains and riparian ecosystems. *Earth Surface Processes and Landforms* **41**(1): 38–60.

Tockner K, Stanford JA. 2002. Riverine flood plains: present state and future trends. *Environmental Conservation* **29**(3): 308–330.

Torres MA, Limaye A, Ganti V, Lamb MP, West AJ, Fischer WW. 2017. Model predictions of long-lived storage of organic carbon in river deposits. *Earth Surface Dynamics Discussions* **5**: 711–730.

Trimble S. 2004. Effects of riparian vegetation on stream channel stability and sediment budgets. In *Riparian Vegetation and Fluvial Geomorphology – Water Science and Application*, Bennett SJ, Simon A (eds). American Geopysical Union: Washington, D.C.; 153–169.

Trimble SW. 2010. Streams, valleys and floodplains in the sediment cascade. In *Sediment Cascades: An Integrated Approach*, Burt T, Allison R (eds). Wiley: Hoboken, NJ; 307–343.

Tucker GE, Hancock GR. 2010. Modelling landscape evolution. *Earth Surface Processes and Landforms* **35**: 28–50.

Van De Wiel MJ, Coulthard TJ. 2010. Self-organized criticality in river basins: challenging sedimentary records of environmental change. *Geology* **38**(1): 87–90.

Van De Wiel MJ, Coulthard TJ, Macklin MG, Lewin J. 2007. Embedding reach-scale fluvial dynamics within the CAESAR cellular automaton landscape evolution model. *Geomorphology* **90**(3–4): 283–301.

Vargas LG. 1990. An overview of the analytical hierarchy process and its applications. *Europeon Journal of Operational Research* **48**: 2–4.

Whittaker AC, Attal M, Allen PA. 2010. Characterising the origin, nature and fate of sediment exported from catchments perturbed by active tectonics. *Basin Research* **22**(6): 809–828.

Wilcock PR, Crowe JC. 2003. Surface-based transport model for mixed-size sediment. *Journal of Hydraulic Engineering* **129**(2): 120–128.

Wittenberg L, Newson MD. 2005. Particle clusters in gravel-bed rivers: an experimental morphological approach to bed material transport and stability concepts. *Earth Surface Processes and Landforms* **30**(11): 1351–1368.

Ziliani L, Surian N. 2012. Evolutionary trajectory of channel morphology and controlling factors in a large gravel-bed river. *Geomorphology* **173–174**: 104–117.

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Data S1. Supporting Information