

Stream power indices correspond poorly with observations of alluvial river channel adjustment

Chris Parker  | James Davey

Department of Geography and Environmental Management, University of the West of England, Bristol, UK

Correspondence

Chris Parker, Department of Geography and Environmental Management, University of the West of England, Bristol BS16 1QY, UK.
Email: chris2.parker@uwe.ac.uk

Abstract

A variety of stream power-based approaches for predicting catchment-scale alluvial channel adjustment have been developed. There is an international interest in applying these to inform river catchment management. However, there is some uncertainty regarding their ability to make consistently accurate predictions. This study evaluates the performance of a range of stream power indices for predicting observed channel adjustment. Remotely sensed data were used to generate 33 different stream power indices every 50 m across the networks of five test catchments. The performances of the indices were evaluated by comparing them against observations of erosion and deposition-dominated channels extracted from the UK's River Habitat Survey (RHS) database. A selection of metrics for evaluating the performance of the indices were calculated. The key finding from this study is that the stream power indices were poorly associated with the observations of alluvial channel adjustment. It is not clear whether this poor association is due to limitations with stream power indices or the suitability of the RHS observations. However, this is not the first study to find a weak association between stream power and observed adjustment. Therefore, caution is recommended to anyone hoping to take advantage of the practicability of stream power indices until further testing is applied using alternative observation datasets. An additional finding from this study is the inconsistency of outcomes between different measures of model performance. It is recommended that future studies also employ multiple model performance measures rather than relying on accuracy alone.

KEYWORDS

adjustment, alluvial, deposition, erosion, river, sediment, stream power

1 | INTRODUCTION

Lane (1955) described alluvial river channels as tending towards a state of balance, using

$$Q.S \propto Q_s . D_{50} \quad (1)$$

where Q is water discharge (m^3/s), S is bed slope, Q_s is sediment supply rate ($\text{kg}/\text{m}/\text{s}$) and D_{50} is the median diameter of sediment supplied (m). The terms on the left represent the sediment transport capacity of the flow, and the terms on the right represent sediment supply. Lane's (1955) balance concept implies that alluvial channel

adjustments are driven by imbalances between the quantity of sediment input to the reach (supply) and the quantity that can be transferred downstream (capacity); channels with excess sediment supply are likely to experience deposition-dominated adjustment while those with excess sediment transport capacity are likely to experience erosion-dominated adjustment. These imbalances can have important implications for the management of flood risk (Naulin et al., 2015; Rinaldi et al., 2009; Stover & Montgomery, 2001), damage to infrastructure in the river corridor (Bowman et al., 2021; Feeney et al., 2022; Li et al., 2021) and ecological status (Ekka et al., 2020; Hauer et al., 2018; Hendry et al., 2003; Lorenz et al., 2004; Soulsby et al., 2001; Wohl et al., 2015).

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial](https://creativecommons.org/licenses/by-nc/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2023 The Authors. *Earth Surface Processes and Landforms* published by John Wiley & Sons Ltd.

There is an international demand for approaches that provide information about likely alluvial channel adjustment at the catchment scale (Owens, 2005); Marcinkowski et al. (2022) described the need to predict geomorphological adjustment across Poland in order to meet requirements of the EU Water Framework Directive; Bowman et al. (2021) reviewed approaches that may be suitable for predicting likely geomorphological activity across catchments in England in order to reduce the expense of managing alluvial erosion and deposition; Papangelakis et al. (2022) described the need for predicting erosion likelihood across urban catchments in Canada; and, as part of its natural flood management strategy, the Scottish Environmental Protection Agency has already produced maps of likely erosion and deposition areas for catchments across Scotland (SEPA, 2013).

Total stream power (Ω , W/m^1) and unit width stream power (ω , W/m^2) are measures of the energy used to drive geomorphological change (Bagnold, 1966), calculated using

$$\Omega = \gamma \cdot Q \cdot S \quad (2)$$

$$\omega = \gamma \cdot Q \cdot S / w \quad (3)$$

where γ is the unit weight of water (9810 N/m^3), Q is an indicative discharge (m^3/s), slope is energy slope (m/m), which is often approximated by bed slope, and w is the width of the flow (m), often approximated by channel bankfull width when using flood flow discharges (Bagnold, 1966; Barker et al., 2009). The ability to measure discharge, slope and channel width using remotely sensed data make Ω and ω practically useful parameters for predicting alluvial channel adjustment at the catchment scale. This, alongside empirical evidence of its control over sediment transport capacity (Bagnold, 1980; Eaton & Church, 2011; Lammers & Bledsoe, 2018b; Martin & Church, 2000; Parker et al., 2011), has led to a number of stream power-based approaches being developed: Bizzi and Lerner (2013) used a combination of both absolute values of Ω and ω and gradients of each across upstream lengths of 3, 5 and 10 km; Parker et al. (2015) developed ST:REAM (Sediment Transport: Reach Equilibrium Assessment Method), which automatically divides river networks into reaches with relatively homogenous ω and then divides the ω of each reach by the ω of its upstream neighbour(s); Lea and Legleiter (2016) used a combination of the ω gradient between adjacent reaches and whether ω was above an entrainment threshold; Soar et al. (2017) developed REAS (River Energy Audit Scheme), which uses stream power integrated across flow duration curves to calculate the annual geomorphic energy (AGE) of reaches, with changes in AGE between reaches indicating the likelihood of erosion and deposition; and Ghunowa et al. (2021) developed the Stream Power Index for Networks (SPIN) to predict morphological change within urban rivers.

There is evidence that these catchment-scale stream power-based approaches can work well in predicting channel adjustment. Bizzi and Lerner (2013) found that absolute and gradient values of Ω and ω could be combined to predict the observations of channel status from River Habitat Survey (RHS) data; ST:REAM correctly predicted the (field observation-based) status of 87.5% of sites within the Taff catchment in Wales (Parker et al., 2015); Yochum et al. (2017) and Sholtes et al. (2018) found that absolute values of Ω and ω and gradient values of ω were significant predictors of geomorphic response to the 2013 Colorado Front Range flood; Marcinkowski

et al. (2022) found that an approach based on ST:REAM correctly predicted the status of >75% of locations across Poland where management of either erosion or deposition issues had been applied; and Papangelakis et al. (2022) found that the frequency of erosion control structures constructed within Etobicoke Creek in Canada was greater with increasing unit width stream power.

However, there is also evidence of limitations in the ability of catchment-scale stream power-based approaches to predict alluvial adjustment. Newson et al. (1998) had limited success using ω to predict the stability status of 484 RHS reaches across England and Wales; Lea and Legleiter (2016) found that the ω gradient was weakly associated with measurements of sediment flux from aerial imagery; Camenen et al. (2016) found that a stream power-based model produced poor representations of sediment budgets along the Middle Loire River; Soar et al. (2017) found that energy budgets produced by REAS corresponded poorly with field observations along the River Kent in England; and Bowman et al. (2021) found that the performance of ST:REAM within three English river catchments was highly variable, with its accuracy being much better on the high-energy River Kent than the low-energy River Stour.

To summarize the above, absolute and relative stream power-based approaches to predicting catchment-scale alluvial channel adjustment have been developed but there is some uncertainty regarding their ability to make consistently accurate predictions. Given the global demand for catchment-scale predictions of alluvial adjustment to inform river management, there is a need to increase empirical evidence of the accuracy of stream power-based techniques. To fill this research gap, this paper aims to evaluate the performance of a range of stream power indices for predicting observed channel adjustment across a range of river catchments. This paper will first introduce the case study river catchments, the remotely sensed datasets used to generate the stream power indices and the secondary dataset used to classify the observed channel status. Next, the paper describes the methods used to generate the observed channel status and the stream power indices, before describing how the performances of the stream power indices are measured. The results are then presented and discussed.

2 | METHODOLOGY AND METHODS

This study uses remotely sensed data to generate a range of different stream power indices every 50 m across the networks of five British river catchments with varying characteristics. The indices being investigated include:

- Absolute values of ω , Ω and T (a new parameter that represents total sediment transport capacity across the channel width).
- Reach-based balances of ω , Ω and T , following an approach similar to ST:REAM (Parker et al., 2015), with a range of five different reach resolutions.
- Point-based balances of ω , Ω and T , following an approach similar to Bizzi and Lerner (2013), with a range of five different lengths used for the upstream average.

The performance of these indices has been evaluated by comparing them against observations of erosion and deposition-dominated

channels extracted from the UK's RHS database. A variety of metrics for evaluating the performance of the indices have been calculated based on the correspondence between the indices and the extracted observations.

2.1 | Data and test catchments

This study used the following data inputs to calculate the 33 stream power indices across each of the test catchments: Environment Agency 2 m LiDAR composite DTM (2020), Ordnance Survey Terrain 5 m DTM (2020), Ordnance Survey OpenRivers river network polyline (2020), Ordnance Survey Mastermap Water theme channel polygons (2020) and Centre for Ecology and Hydrology's hydrological gauging station Q_{med} values (2020). Observations of channel status were extracted from the UK's RHS database, which uses a well-established procedure for surveying physical and habitat features of a river reach (Environment Agency, 2003; Raven et al., 1998). RHS data were collected by field surveys from 1994 to 2016 along 500 m river reaches; morphological and physical habitat features are recorded at 10 evenly spaced 'spot-checks', whereas other features are qualitatively described within a 'sweep-up' across the 500 m reach.

The five selected test catchments are: the Spey in the Scottish Highlands; the Tweed on the Scottish–English border; the Wyre in Northwest England; the Medway in Southeast England; and the Camel in Southwest England (Figure 1). These catchments were selected because they have both a very high coverage of RHS reaches and are varied in their characteristics (Table 1).

2.2 | Classification of observed channel status

The dominant process acting within a river channel can be qualitatively evaluated by the interpretation of field observations (Sear et al., 2010). For instance, the extended presence of unvegetated depositional bars indicates a rich sediment supply from upstream that is partially stored in the reach. Erosion features such as eroding cliffs and vertical or undercut banks indicate processes of bank erosion and are an indication of the amount of sediment mobilized towards downstream. This study classified the observed status of each RHS reach using an approach similar to that of Bizzi and Lerner (2013), who had modified an approach used by Newson et al. (1998). First, alluvial, unconfined channels were isolated by removing any reaches that did not have alluvial material (cobbles, pebbles, gravels, sand, silt, clay) as the dominant material in the majority of their bed and bank spot-checks. Then, the criteria in Table 2 were used to identify any reaches that could be classified as erosion dominated (extensive erosional features and limited depositional features) or deposition dominated (extensive depositional features and limited erosional features).

2.3 | Measurement and calculation of stream power indices

The processes for measuring and calculating the stream power indices for each test catchment are outlined in Figure 2. These are based on those described by Parker et al. (2015), with some updates. All data

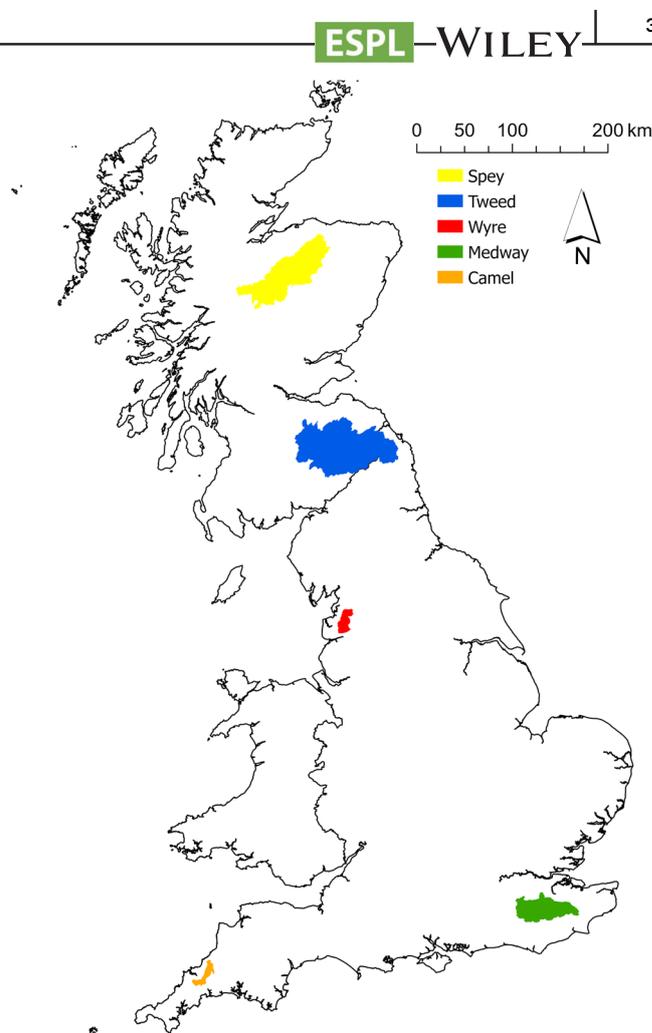


FIGURE 1 Location of the test catchments. The Spey in the Scottish Highlands; the Tweed on the Scottish–English border; the Wyre in Northwest England; the Medway in Southeast England; and the Camel in Southwest England.

processing was performed using a combination of ESRI's ArcGIS Pro and Microsoft Excel. The steps followed for each test catchment were:

1. Using the 'Mosaic' tool to combine the Terrain 5 m DTM with the 2 m LiDAR DTM, using the 5 m DTM to fill in any gaps in the higher-resolution and accuracy LiDAR DTM. This has the benefits of ensuring coverage across the entire catchment whilst maximizing the accuracy.
2. Using the 'Contour' tool to convert the combined DTM raster to 2 m contours. This contour dataset was used later to measure channel slope.
3. Using the 'Topo to Raster' tool to interpolate a 10 m DTM from the 2 m contours and the OpenRivers polyline of the known river channel network.
4. Using the 'Minus' tool to 'burn' the location of the known river channel network into the 10 m DTM, in order to ensure that the modelled river channel location matched its known location.
5. Using the 'Fill' tool to fill any pits (local elevation minima) within the DTM, in order to prevent them obstructing the modelled progress of water flowing downslope across the catchment surface.
6. Using the D8 algorithm within the 'Flow Direction' tool to identify the outgoing flow direction for each raster cell.

TABLE 1 Key information for the test catchments

Catchment	Medway	Camel	Wyre	Spey	Tweed
Catchment outlet location (British National Grid Reference)	Teston (TQ708530)	Denby (SX017681)	St Michaels (SD463411)	Boat o Brig (NJ318517)	Norham (NT898477)
Catchment area (km ²)	1256.1	208.8	275.0	2861.2	4390.0
Mean flow (m ³ /s)	11.10	5.96	6.83	65.39	81.31
Base flow index	0.40	0.63	0.31	0.59	0.52
Q_{med} (m ³ /s)	131.0	57.5	131.0	507.0	825.0
Average annual rainfall (mm)	743	1338	1251	1119	955
Mean drainage path slope (m/m)	0.054	0.088	0.071	0.157	0.136
BFIHOST (soil permeability)	0.44	0.56	0.37	0.48	0.49
Total RHS survey reaches	536	129	139	187	473
Erosion-dominated RHS reaches	94	6	6	9	40
Deposition-dominated RHS reaches	26	35	47	60	106
Stream power (ω , W/m ²): Q_1 (lower quartile boundary)	22.2	68.6	80.8	83.8	81.1
Stream power (ω , W/m ²): Q_2 (median)	44.8	113.8	122.9	141.5	149.1
Stream power (ω , W/m ²): Q_3 (upper quartile boundary)	80.7	171.6	202.0	239.2	266.9

TABLE 2 Criteria used to classify alluvial RHS reaches as either erosion or deposition dominated

Erosion-dominated reach	Deposition-dominated reach
<ul style="list-style-type: none"> EITHER >4 eroding earth cliffs[#] OR (>2 eroding earth cliffs[#] AND vertical/undercut bank profile extended⁺) Sum of all types of unvegetated bars (point, side or mid-channel) NOT >3[#] 	<ul style="list-style-type: none"> Sum of all types of unvegetated bars (point, side or mid-channel) >3[#] NEITHER >4 eroding earth cliffs[#] NOR (>2 eroding earth cliffs[#] AND vertical/undercut bank profile extended⁺)

[#]Within the 10 spot-checks.

⁺Within the sweep-up.

- Using the 'Flow Accumulation' tool to identify the total number of upstream cells that contribute flow into each cell.
- Using the 'Times' tool to calculate the drainage area of each cell by multiplying each cell's flow accumulation value by the area of a 10 m cell (0.0001 km²).
- Using the 'Greater Than Equal' and 'Stream to Feature' tools to create a polyline representation of all cells with a drainage area >0.5 km².
- Using the 'Generate Points Along Lines' tool to create points spaced 50 m apart along all branches that contribute at least 1% of the total catchment drainage area.
- Using the median annual flood (Q_{med}) and drainage areas (A) for flow gauges across the catchment to create a power relationship of the form $Q_{med} = \alpha A^\beta$, where α and β are constants derived for each catchment. This is the approach used by most studies that represent stream power at the catchment scale (Bizzi & Lerner, 2013; Knighton, 1999; Papangelakis et al., 2022; Parker et al., 2015). Table 3 provides details of the regression relationships for each of the test catchments.

- Using the drainage area raster and the derived Q_{med} relationship to predict the Q_{med} for each of the points across the river catchment network.
- Using the elevation difference between the contours up and downstream of each point, and the along-stream distance between those contours, to calculate the channel slope (S) for each point. This 'vertical slice slope measurement' using contours has previously been demonstrated to be less sensitive to stepping artifacts and other sources of noise than 'horizontal slice slope measurement' using DTMs (Reinfelds et al., 2004; Vocal Ferencevic & Ashmore, 2012; Wobus et al., 2006). Using the River Wyre as an example,
- Figure 3 illustrates how vertical slice slope measurements provide a more appropriate representation of channel slope than horizontal slice slope measurements.
- Measuring the channel width (w) of the appropriate MasterMap Water polygon at each of the points across the river catchment network.
- Using the values of Q_{med} , S and w to calculate absolute values of ω , Ω and T at each of the points across the river catchment network (Figure 4a). T provides a representation of total sediment transport capacity across the channel width by accounting for the non-linear relationship between ω and transport capacity (Bagnold, 1980; Martin & Church, 2000), calculated using

$$T = \omega^{3/2} \cdot w \quad (4)$$

- Using the values of ω , Ω and T to calculate reach-based balances of each, following an approach similar to ST:REAM (Parker et al., 2015), with a range of five different reach resolutions (Figure 4b). This involved:
 - Clustering the sequence of point values along each branch into internally homogenous reaches using Gill's (1970) global zonation algorithm in the same manner as Parker et al. (2015). This

FIGURE 2 Flowchart of processes involved in measuring and calculating stream power indices.

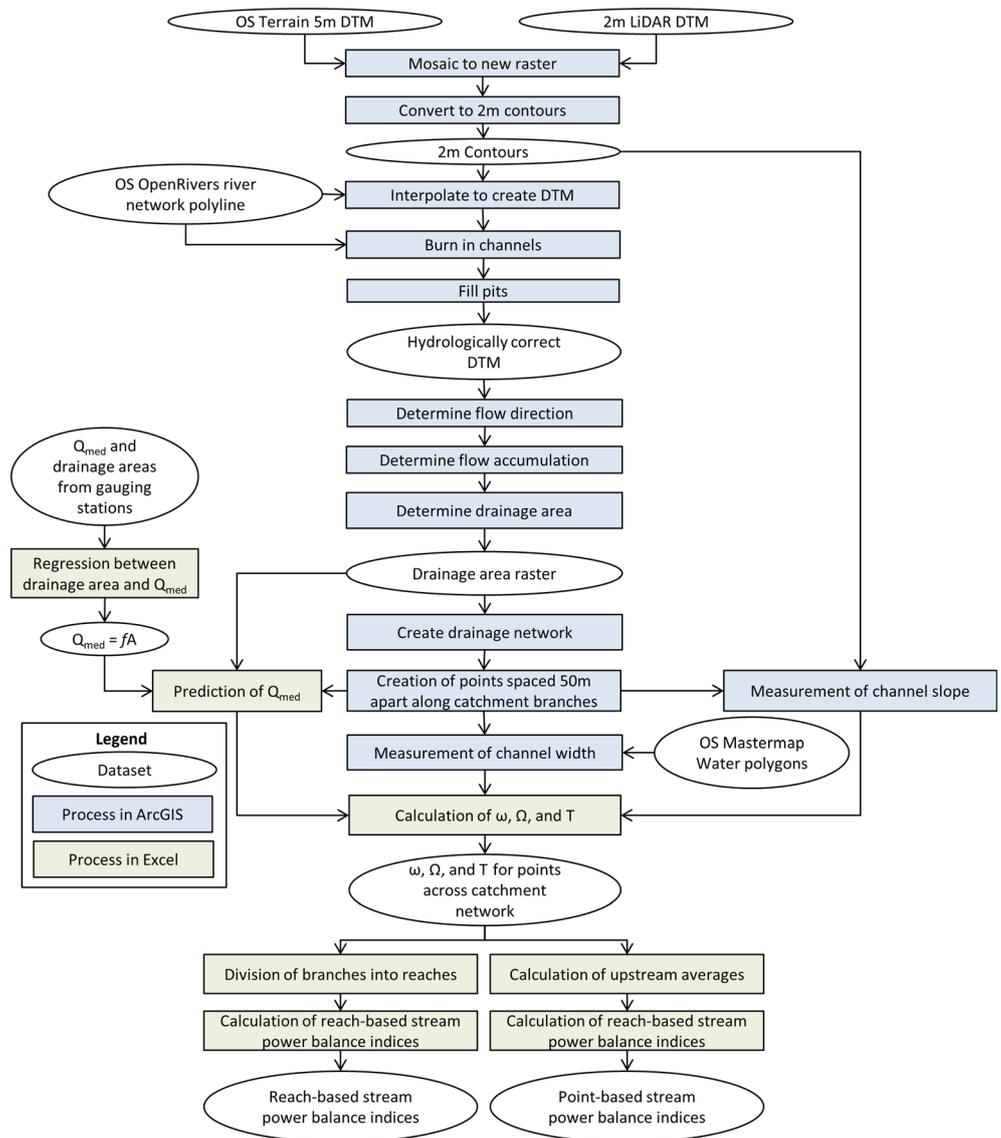


TABLE 3 Details of regression relationships used to predict median annual flood within each of the test catchments, following the form $Q_{med} = \alpha A^\beta$

Catchment	α	β	R^2
Medway	1.209	0.6358	0.9567
Camel	0.6428	0.8466	0.9843
Wyre	0.987	0.9279	0.8478
Spey	1.1284	0.7247	0.8268
Tweed	3.1548	0.6543	0.9891

was performed for five different values of R (0.5, 0.6, 0.7, 0.8 and 0.9), where R represents the proportion of the variability of the point values that are explained by the reaches—so that higher values of R involve the division of branches into smaller reaches.

- b For each reach, calculating the balance between its mean and the mean of its upstream neighbour(s), repeated for each value of R .
- c The calculated reach balance value was then applied to each of the 50 m points within the reach.

- 18. Using the values of ω , Ω and T to calculate point-based balances of each, following an approach similar to Bizzi and Lerner (2013), with a range of five different lengths used for the upstream average (Figure 4c). This involved:
 - a For each point, calculating the mean across points along a set upstream distance. This was performed for five different upstream distances: 1, 3, 5 and 10 km, and all upstream points.
 - b For each point, calculating the balance between its value and the upstream mean, repeated for each distance.

The above steps resulted in values of the following 33 stream power indices for 50 m-spaced points across each of the five test catchment networks: ω , Ω , T , $\omega_{balance-reach-0.5}$, $\omega_{balance-reach-0.6}$, $\omega_{balance-reach-0.7}$, $\omega_{balance-reach-0.8}$, $\omega_{balance-reach-0.9}$, $\Omega_{balance-reach-0.5}$, $\Omega_{balance-reach-0.6}$, $\Omega_{balance-reach-0.7}$, $\Omega_{balance-reach-0.8}$, $\Omega_{balance-reach-0.9}$, $T_{balance-reach-0.5}$, $T_{balance-reach-0.6}$, $T_{balance-reach-0.7}$, $T_{balance-reach-0.8}$, $T_{balance-reach-0.9}$, $\omega_{balance-point-1km}$, $\omega_{balance-point-3km}$, $\omega_{balance-point-5km}$, $\omega_{balance-point-10km}$, $\omega_{balance-point-All}$, $\Omega_{balance-point-1km}$, $\Omega_{balance-point-3km}$, $\Omega_{balance-point-5km}$, $\Omega_{balance-point-10km}$, $\Omega_{balance-point-All}$, $T_{balance-point-1km}$, $T_{balance-point-3km}$, $T_{balance-point-5km}$, $T_{balance-point-10km}$, $T_{balance-point-All}$.

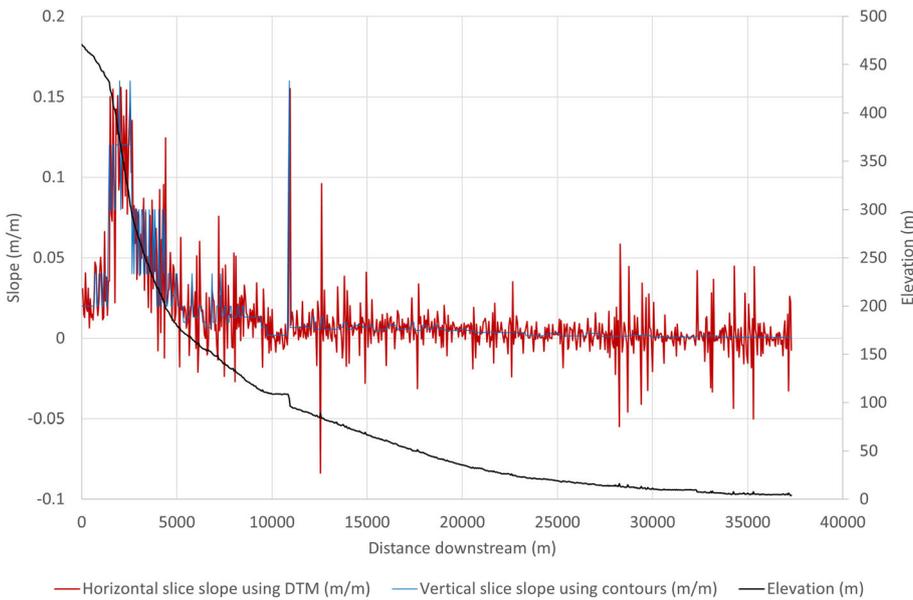


FIGURE 3 Comparison between horizontal slice and vertical slice slope measurement techniques on the main branch of the River Wyre.

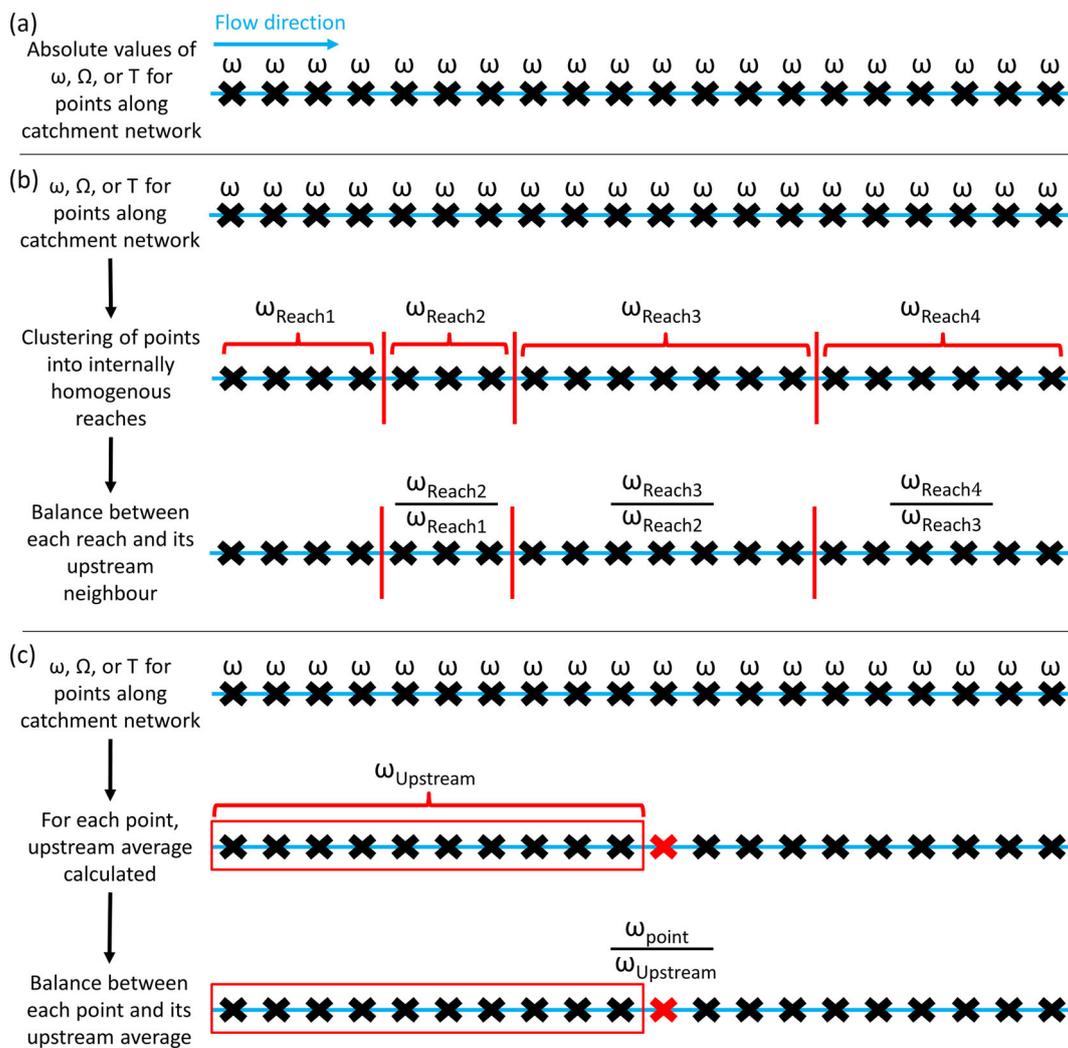


FIGURE 4 Conceptual diagrams representing the differences between the difference stream power parameters: absolute values (a), reach-based balances (b) and point-based balances (c).

2.4 | Assessing performance of stream power indices

The performance of all 33 stream power indices, in each of the five test catchments, has been evaluated based on their correspondence with RHS reaches classified as deposition dominated or erosion dominated. To match up the observations from the RHS reaches with the appropriate stream power index values, the 50 m stream power point closest to the middle of the 500 m RHS reach was used. An alternative approach, applied by Bizzi and Lerner (2013), would have been to average the stream power index values from all 10 of the 50 m points that fall within the 500 m RHS reach.

The confusion matrix for assessing the performance of the 30 stream power balance indices is displayed in Figure 5. Parker et al. (2015) evaluated the performance of ST:REAM by measuring its accuracy, with accuracy defined as the proportion of total observations predicted correctly:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (5)$$

where TP , TN , FP and FN , respectively, refer to the number of true positives, true negatives, false positives and false negatives. However, accuracy does not provide a reliable representation of predictive ability because it is affected by the distribution of observed values (Huang & Ling, 2005). For example, a model that is biased towards predicting that a channel will be erosion dominated will achieve high accuracy in a catchment where most observed points are erosion dominated. In addition, accuracy also requires the *a priori* definition of a threshold value. Whilst this suits the balance indices (for which a default threshold of 1 inherently applies), the absolute values of ω , Ω and T require bespoke thresholds that require calibration and so cannot be tested using accuracy in this study. Accuracy can range from 0 to 1, with a value of 0.5 being no better than a random allocator.

Therefore, to properly evaluate the performance of the stream power indices, some alternative model performance measures have been utilized. The additional performance measures applied are the Matthews correlation coefficient (MCC) and the area under the receiver operating curve (AUC).

		Observed channel status	
		Erosion dominated	Deposition dominated
Stream power balance	>1	True positive	False positive
	<1	False negative	True negative

FIGURE 5 Confusion matrix for assessing the performance of the stream power balance indices.

MCC treats the observed outcomes and predicted outcomes as two binary variables and computes their correlation coefficient using

$$\text{MCC} = \frac{(TP \cdot TN) - (FP \cdot FN)}{\sqrt{(TP + FP) \cdot (TP + FN) \cdot (TN + FP) \cdot (TN + FN)}} \quad (6)$$

MCC equally takes into account all four elements of the confusion matrix, with a high value meaning that both positive and negative outcomes are predicted well, even if there is imbalance in the distribution of observed outcomes (Chicco et al., 2021). However, like accuracy, MCC also requires the *a priori* definition of a threshold value and so cannot be used to test the performance of the absolute values of ω , Ω and T . MCC can range from -1 to 1 , with a value of 0 being no better than a random allocator.

AUC is the area under a curve that plots the true positive rate (TPR – the proportion of all observed positives that are predicted as being positive) against the false positive rate (FPR – the proportion of all observed negatives that are incorrectly predicted as being positive) as the threshold value changes, where TPR and FPR are measured using

$$\text{TPR} = \frac{TP}{TP + FN} \quad (7)$$

$$\text{FPR} = \frac{FP}{FP + TN} \quad (8)$$

AUC provides the probability that the model ranks a random positive observation more highly than a random negative observation. As well as being insensitive to imbalances in the distribution of observed outcomes (Huang & Ling, 2005), AUC also does not need a threshold to be set (as it measures cumulative performance across the full range of possible thresholds) – so that, unlike the other performance measures, it can be applied to the absolute values of ω , Ω and T . AUC can range from 0 to 1 , with a value of 0.5 being no better than a random allocator.

3 | RESULTS

Figure 6 displays the observed channel locations classified as either erosion or deposition dominated using the information from the RHS database and the criteria set out in Table 2. The number of RHS reaches classified as erosion and deposition-dominated status for each catchment is provided in Table 1. In each catchment there is an imbalance between the number of erosion and deposition-dominated reaches, with four of the catchments having more deposition-dominated reaches and the Medway, the catchment with the lowest energy, counterintuitively having more erosion-dominated reaches. In addition, within some of the catchments there is clustering of erosion and deposition reaches. The Tweed has deposition-dominated reaches across most of its catchment but has a cluster of erosion-dominated reaches within its eastern branch; the Medway has lots of erosion-dominated reaches in its western headwaters and a cluster of deposition-dominated reaches in the headwaters of one of its southern branches; the Wyre has deposition-dominated reaches across the upper and middle portions of its northern branch, with a small cluster

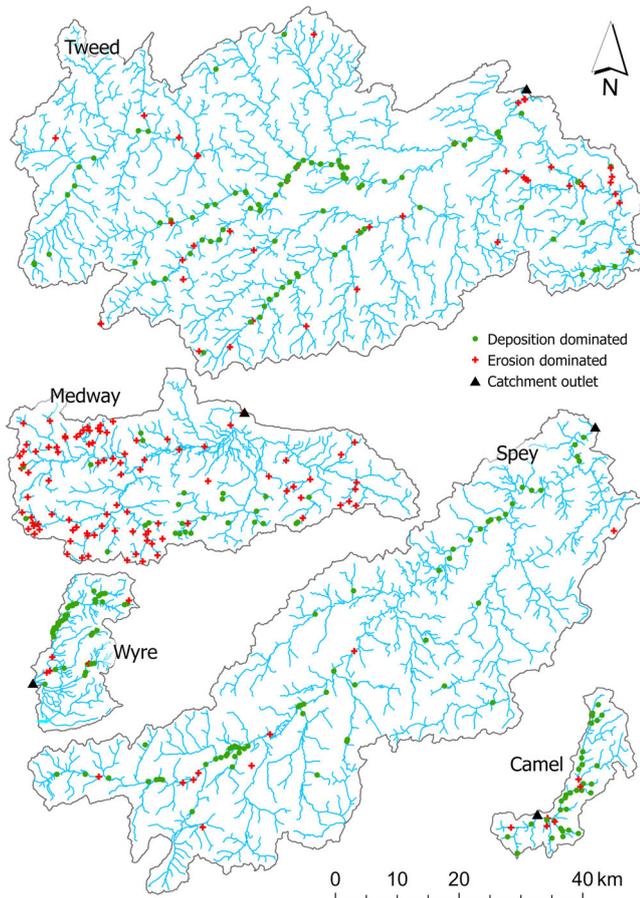


FIGURE 6 Channel locations across each of the five test catchments classified as either erosion or deposition dominated based on observations from the RHS.

of erosion-dominated reaches near the mouth; and the Spey has deposition-dominated reaches along its main branch, interrupted by a cluster of erosion-dominated reaches towards its upstream extent.

To illustrate the variation in stream power between and within the five test catchments, Figure 7 displays the calculated unit width stream power values (ω) for points spaced every 50 m along each catchment network. Summary statistics of ω for each catchment are also provided in Table 1. The Medway has considerably lower values of ω than the other four catchments. There is substantial variation in ω within each of the test catchments but the nature of the spatial distribution varies between the catchments: the Tweed has high ω values across most of the catchment, with low values along the higher-order channels near the outlet and very low values at some headwater reservoirs; the Medway has low ω values across most of the catchment, with higher ω values in the headwaters of some of the tributaries; the Wyre has high ω values across most of the lower and middle-order channels, but low ω values along the higher-order channels near the outlet; the tributaries of the Spey have high ω values that are interrupted by low ω values at some headwater reservoirs, and the main branch of the Spey starts with low ω values and gradually increases to high near the mouth; the Camel has high ω values across most of the catchment, with low ω values along some of the branches feeding down from the northeast. Visually, the variation in ω values within and between the catchments (Figure 7) does not appear to reflect the locations of reaches observed to be erosion and deposition dominated

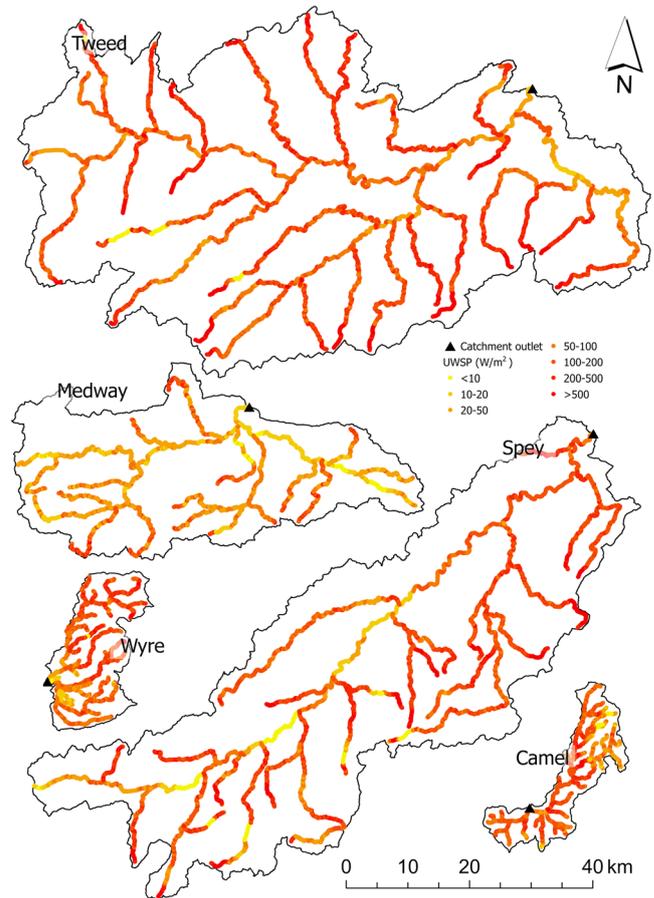


FIGURE 7 Calculated unit width stream power (ω) values for points spaced every 50 m across the catchment network of each of the five test catchments.

(Figure 6). The remainder of this section describes whether this visual assessment is supported statistically.

Figures 8–10 display the performance metrics for the various stream power indices within each of the five test catchments and across all of the catchments combined. The three key observations to make from the results are outlined below.

Firstly, the overall correspondence between each of the stream power indices and the observed channel status was poor. None of the indices achieved a result better than a random allocation consistently across all test catchments and performance metrics. Whilst most of the indices achieved moderately successful levels of accuracy for the Spey and Tweed catchments, this is undermined by both poor accuracy performance for the other catchments and poor performance in those catchments for the other metrics.

Secondly, there is a substantial difference between the outputs of the different performance metrics. The results for the accuracy metric vary considerably across the five test catchments, whilst the MCC and AUC metrics are more consistently poor across all catchments.

Thirdly, there appears to be some moderate variation in performance between the different stream power indices: the majority of point-based balance metrics achieve higher MCC and AUC scores than the majority of reach-based metrics; the absolute values of ω , Ω and T achieve lower AUC scores than the majority of balance indices; the point-based balances using shorter upstream average distances

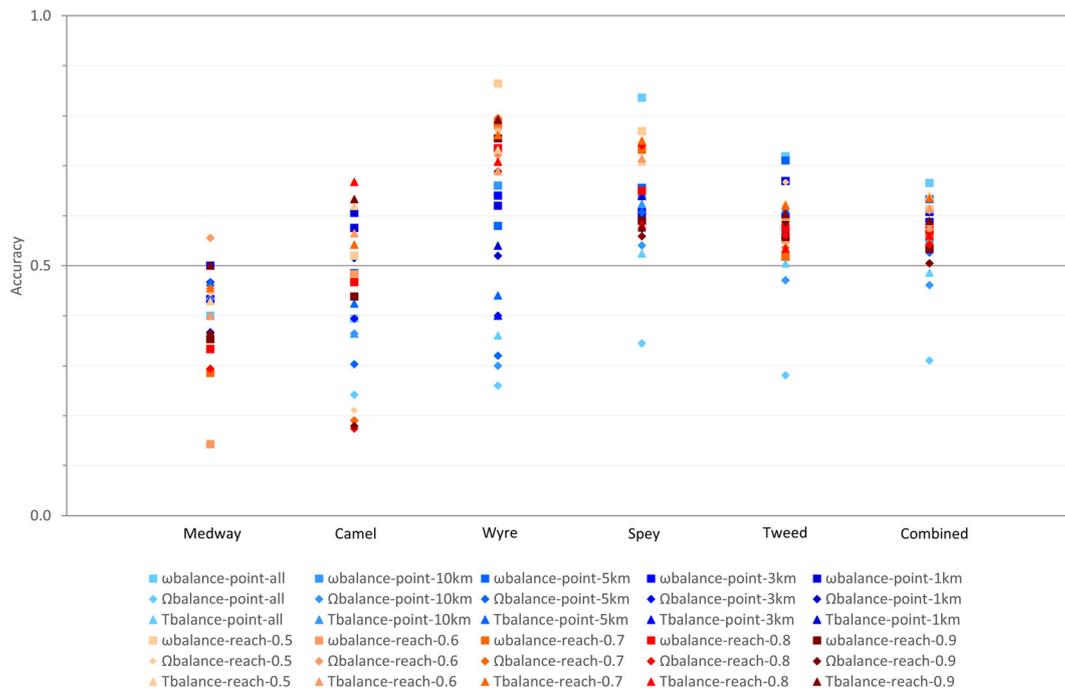


FIGURE 8 Accuracy performance metric for the 30 balance indices in each of the five test catchments, and for all catchments combined. Accuracy can range from 0 to 1, with a value of 0.5 being equivalent to a random allocator.

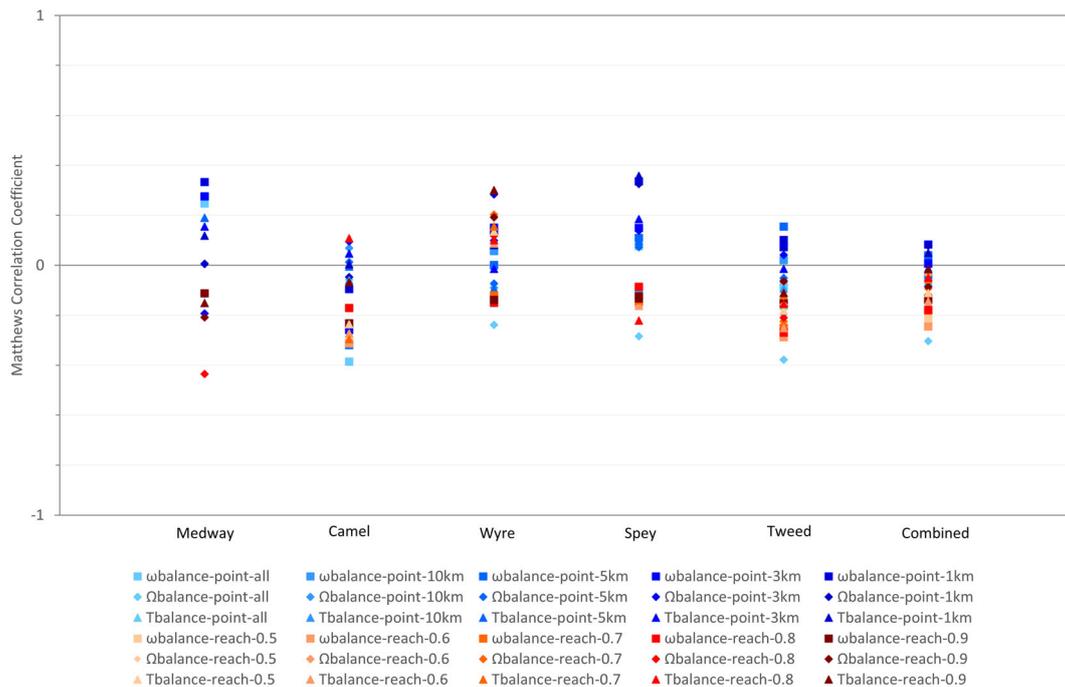


FIGURE 9 Matthews correlation coefficient performance metric for the 30 balance indices in each of the five test catchments, and for all catchments combined. MCC can range from -1 to 1 , with a value of 0 being equivalent to a random allocator.

achieve higher MCC and AUC than those using longer upstream distances. However, any such patterns are overshadowed by the fact that all of the stream power indices performed poorly and so are not considered further.

As described in Section 2.4, instead of matching each RHS reach to the stream power index of the point that is closest to its middle (as was applied in this study), it is possible to use the mean of the 10 stream power index values that fall within the 500 m RHS reach. To test the impact of choosing one approach over

another, Figure 11 compares the AUC scores from the Wyre catchment achieved by using reach-mean stream power indices with those achieved using the mid-point stream power indices. The results show that averaging the stream power indices across the points within the RHS reach did not improve their measured performance. In fact, a paired *t*-test found that the AUC scores using the reach-averaged approach were slightly, but statistically significantly, lower than those using the mid-point approach (mean difference = 0.1089 , *t*-value = 5.73 , *p*-value = 0.000).

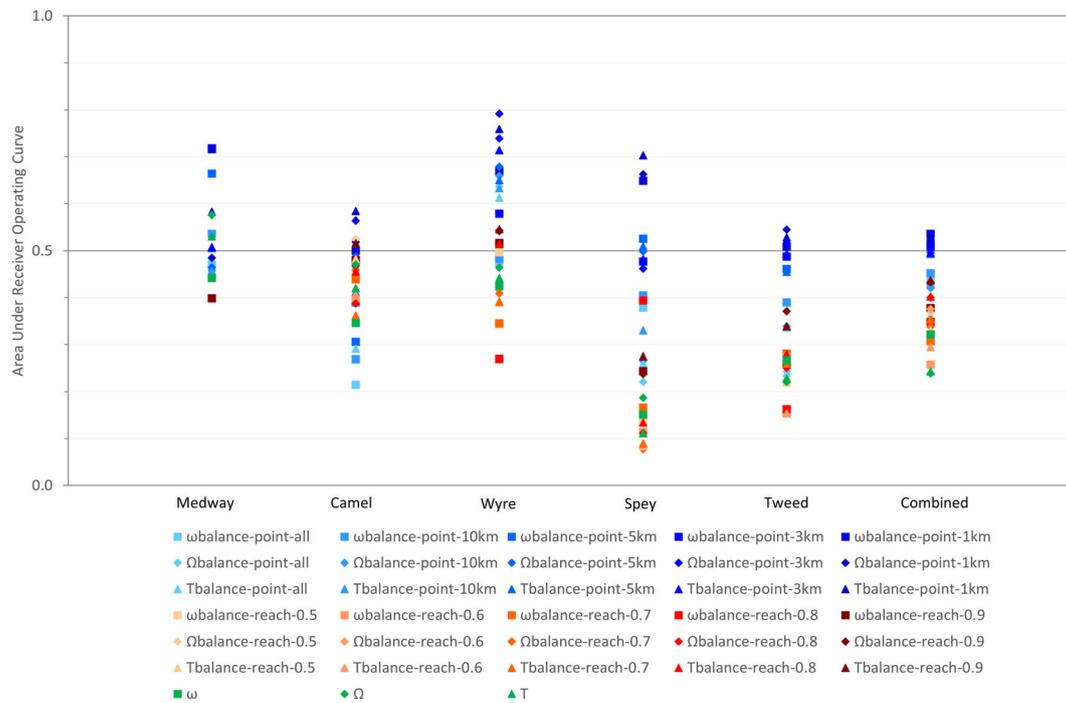


FIGURE 10 AUC performance metric for the 30 balance indices and the three absolute indices in each of the five test catchments, and for all catchments combined. AUC can range from 0 to 1, with a value of 0.5 being equivalent to a random allocator.

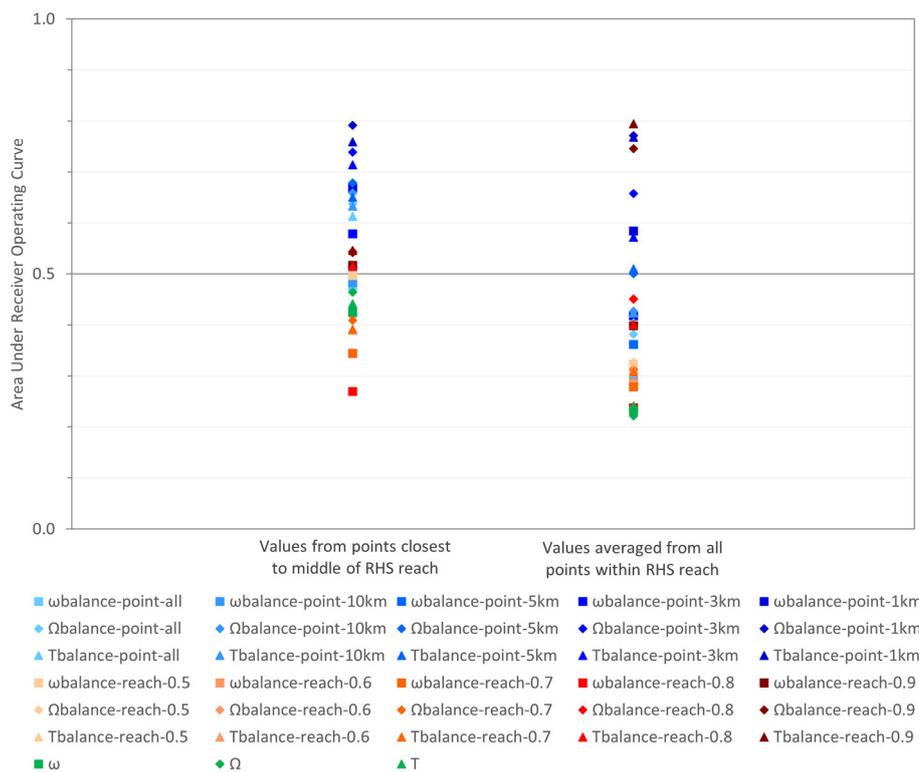


FIGURE 11 Comparison of AUC performance metric (for the 30 balance indices and the three absolute indices) between using the index value from the point in the middle of the RHS reach and using the mean of all points within the 500 m RHS reach. AUC can range from 0 to 1, with a value of 0.5 being equivalent to a random allocator.

4 | DISCUSSION

The poor correspondence between the stream power indices and the observations of alluvial channel adjustment are surprising given the contrasting findings of numerous previous studies (Biron et al., 2013; Bizzi & Lerner, 2013; Marcinkowski et al., 2022; Papangelakis et al., 2022; Parker et al., 2015; Sholtes et al., 2018; Vocal Ferencevic & Ashmore, 2012; Yochum et al., 2017). The poor performance is even more surprising given the observational dataset was of

only erosion and deposition-dominated RHS reaches—making differentiation between them an easier task than if balanced RHS reaches were also included. There are two possible explanations for this study's surprising findings: either the calculated stream power indices are not reliable predictors of alluvial adjustment; or there are limitations with the observational dataset used in this study. The following two sections tackle these possibilities in turn, with a third section considering the suitability of accuracy as a measure of model performance.

4.1 | Limitations of stream power indices for predicting alluvial adjustment

Using simple stream power indices to predict alluvial channel adjustment involves many simplifications of what are complex, non-linear and dynamic systems. In particular: the stream power indices in this study do not comprehensively represent the factors affecting bed material transport; they cannot properly represent sediment supply; they are limited by their temporally and spatially rigid representations of the fluvial system; they do not account for variations in a channel's degrees of freedom; and they are limited by the accuracy of remotely sensed data available at the catchment scale.

Whilst there is much evidence of stream power's influence over sediment transport capacity (Bagnold, 1980; Eaton & Church, 2011; Lammers & Bledsoe, 2018b; Martin & Church, 2000; Parker et al., 2011), a single stream power value cannot comprehensively represent the sediment output from a river channel. Both bed material size (Bagnold, 1980) and sorting (Wilcock & Crowe, 2003) influence sediment transport rate and so any variation in these along the river channel network will cause variations in sediment output that cannot be predicted using stream power alone. In addition, stream power indices that are based on Q_{med} alone do not necessarily represent the most geomorphically effective flow impacting a river channel (Doyle & Shields, 2008). Further, it is not just the magnitude of peak flows that is important for channel adjustment during a flood event. The duration (Costa & O'Connor, 2011; Gervasi et al., 2021) and sequencing (Eagle et al., 2021; Major et al., 2021) of flood events are also of critical importance. Not accounting for the impact of either bed material characteristics or the full flow regime limits the ability of stream power indices to predict sediment output.

On their own, stream power indices cannot represent all of the controls over sediment supply to a river channel. Whilst stream power balance indices use the stream power of upstream channels as an indicator of likely sediment supply, this incorrectly assumes that sediment supplied from upstream is itself not supply limited. As bed material transport is known to be limited by the availability of sediment for transport (Wolman et al., 1997), the stream power of upstream channels alone is not sufficient to predict the bed material transported from upstream. Additionally, stream power indices do not account for sediment supplied from sources other than upstream channels. Sediment delivered from hillslopes can locally alter the morphology of streams and increase the supply, transport and storage of coarse sediment within the channel network (Benda & Dunne, 1997). There have been recent advances in our ability to represent sediment supply from river catchments and their networks (Czuba et al., 2017; Schmitt et al., 2016) that could be used to improve predictions of channel adjustment in combination with stream power indices.

Stream power indices are static representations of what are inherently dynamic systems. From a temporal perspective, providing a single snapshot of the balances in energy within the catchment network ignores the form-process feedbacks that are important in how alluvial channels adjust over time (Ashworth & Ferguson, 1986). Additionally, the rigid representations of the spatial scales involved in the balance indices do not properly represent the dynamic nature of fluvial processes. Whilst the point-based balance indices require a defined length of upstream reach to represent the sediment supply, in

reality the length of upstream channel supplying sediment will vary with the nature of the particular channel and with the magnitude of each flow event. Similarly, whilst the reach-based balance indices divide the channel network into discrete homogenous reaches, real channels will both have significant variations within reaches and more subtle transitions between reaches.

On their own, stream power indices do not account for how restrictions on a river channel's degrees of freedom limit its adjustment (Phillips, 2007). A channel's degrees of freedom can be limited by either natural factors, like valley confinement, or artificial structures such as bridges and weirs. The importance of valley confinement for alluvial adjustment has been demonstrated by several studies, including Righini et al. (2017), which found that confinement had a strong influence on the degree of widening experienced by alluvial reaches of two Mediterranean rivers during an extreme flood. The impact of artificial structures on channel adjustment was demonstrated by Camenen et al. (2016), who found that a stream power-based approach performed particularly poorly in predicting alluvial adjustment along the Middle Loire River within complex reaches that were interrupted by bridges and weirs. It may be possible to improve the performance of stream power indices by combining them with representations of valley confinement like that of O'Brien et al. (2019).

The application of stream power indices at the catchment scale necessitated using techniques for measuring discharge, slope and width that may not be completely accurate: the discharge measurements assumed that there is a consistent relationship between Q_{med} and drainage area across the entire catchment (which can be disrupted by variations in geology, land use and topography); the slope measurements do not capture any variations in slope that may fall within the 2 m contours; and the width measurements are based on OS Mastermap polygons that do not consistently represent the width of the flow at Q_{med} .

Some, or all, of the limitations explored above may be responsible for the poor correspondence between stream power indices and observed channel status in this study and others (Bowman et al., 2021; Camenen et al., 2016; Lea & Legleiter, 2016; Soar et al., 2017). There are a variety of alternative, more complex approaches to predicting alluvial channel adjustment at the catchment scale that resolve one or more of these limitations. These include: connectivity models like CASCADE (Catchment Sediment Connectivity And Delivery), which can predict the sinks for specified sediment sources across the river channel network (Schmitt et al., 2016) or predict sediment supply from across the catchment (Bizzi et al., 2021; Schmitt et al., 2018); channel evolution models like REM (River Erosion Model), which can simulate decadal channel evolution dynamics to predict the location of bed and bank adjustment across river channel networks (Lammers & Bledsoe, 2018a); and cellular models like CAESAR (Cellular Automaton Evolutionary Slope And River), which can simulate dynamic evolution across hillslopes and channels over thousands of years to predict the location of net erosion and deposition across catchments (van de Wiel et al., 2007). Whilst more complex approaches like these require data inputs beyond those used for stream power indices, rapid improvements to remote sensing technology are enabling increased data collection for catchment-scale modelling purposes (Piégay et al., 2020).

4.2 | Limitations of river habitat survey data for defining observed channel status

The RHS database was used as the basis for defining observed channel status because it provided a large pre-existing dataset and because previous studies used it for similar purposes (Bizzi & Lerner, 2013; Newson et al., 1998). However, there are multiple reasons why it may not be a reliable observational set for evaluating predictions of alluvial adjustment.

Firstly, the RHS was designed to survey physical habitats rather than provide geomorphological assessment (Newson et al., 1998). As a result, it has a limited capability to detect erosion and deposition processes. The extent of depositional and erosional features is not directly quantified and there are no evaluations of sediment longitudinal continuity, channel-floodplain connectivity or bed armouring. This makes quantitative comparisons of erosion and deposition features difficult when using the RHS (Bizzi & Lerner, 2013).

Secondly, the features exhibited by adjusting channels can confuse simple classification schemes like that applied in Table 2. Erosion-dominated channels can have depositional features (e.g. bars) composed of recently eroded material and deposition-dominated channels can have erosional features due to flow being diverted around depositional forms. For example, in their investigation into the impact of stream power on alluvial adjustment during the 2013 Colorado Front Range flood event, Sholtes et al. (2018) found that the highest levels of bank erosion occurred within reaches that had lower stream power than upstream, due to depositional processes like braiding and channel avulsion.

Thirdly, the RHS data have been collected over three decades (starting in 1994). Whilst it is unlikely that many of the surveyed reaches have changed their morphological status in that period, it is possible that some will have. As a result, it may not be appropriate to use the observations of channel adjustment from RHS data as a means of testing the performance of stream power indices derived from spatial data collected in 2020.

There are a variety of alternative sources of observed morphological status that could be used for evaluating predictions of alluvial adjustment. One option is field surveys, either those primarily collected for the purpose of evaluating the performance of model predictions (e.g. Parker et al., 2015) or repurposing of those that have been collected previously (e.g. Bowman et al., 2021). A second option is to use databases of applied management activities in a manner similar to both Marcinkowski et al. (2022), who evaluated a stream power index against locations where either erosion (backfilling) or deposition (dredging) management actions had been applied, and Papangelakis et al. (2022), who compared stream power indices against the density of erosion control structures. A third option is interpretation from aerial photography like that applied by both Mazgareanu et al. (2020) and Lea and Legleiter (2016). Finally, it is possible to derive remotely sensed elevation changes, using either ground-based, airborne or spaceborne techniques like those used by Sholtes et al. (2018), Conesa-García et al. (2022) and others described by Piégay et al. (2020).

4.3 | Limitations of accuracy as a model performance measure

An additional finding from this study is the inconsistency of outcomes from the different measures of model performance. Our results support the arguments of Huang and Ling (2005) that accuracy is not a reliable representation of a model's predictive ability. Our results found that it was possible to achieve accuracy values as high as 0.87 ($\omega_{\text{balance-reach-0.5}}$ on the Wyre catchment) when the more robust MCC and AUC performance measures indicated that the respective model performed no better than a random allocator. This is considered to be due to the accuracy metric being affected by the imbalanced distribution of observed values in the test catchments, whilst both MCC and AUC are relatively insensitive to the distribution of observed outcomes (Chicco et al., 2021; Huang & Ling, 2005).

5 | CONCLUSIONS

The key finding from this study is that the 33 stream power indices tested across five catchments were poorly associated with the observations of alluvial channel adjustment from the RHS database. It is not clear whether this poor association is due to limitations with stream power indices or the suitability of the RHS observations. However, this is not the first study to find a weak association between stream power and observed adjustment (Bowman et al., 2021; Camenen et al., 2016; Lea & Legleiter, 2016; Soar et al., 2017). Therefore, caution is recommended to anyone hoping to take advantage of the practicability of stream power indices for predicting alluvial adjustment, at least until further testing is applied using alternative observation datasets. This is an important finding given the international interest in applying stream power indices within river catchment management.

Despite the poor associations between stream power indices and observations of alluvial channel adjustment in this study, numerous studies have previously demonstrated a correlation between stream power and fluvial adjustment (Biron et al., 2013; Bizzi & Lerner, 2013; Marcinkowski et al., 2022; Papangelakis et al., 2022; Parker et al., 2015; Sholtes et al., 2018; Vocal Ferencevic & Ashmore, 2012; Yochum et al., 2017), and so stream power may still be a useful parameter for predicting river channel adjustment when used appropriately and cautiously. It is unclear why those studies found an association between stream power and fluvial adjustment when this study did not, particularly in the case of Bizzi and Lerner (2013), who used a similar source for their observed status. Therefore, it is recommended that further research is performed into why stream power indices perform better in some circumstances than others, and whether stream power can be effectively combined with representations of confinement (like that of O'Brien et al., 2019) and sediment supply (like that of Schmitt et al., 2016) to more accurately predict channel adjustment.

An additional finding from this study is the inconsistency of outcomes from different measures of model performance. It is recommended that future studies also employ multiple model performance measures in order to get a comprehensive understanding of the performance of their models.

DATA AVAILABILITY STATEMENT

Data available on request from lead author.

ORCID

Chris Parker  <https://orcid.org/0000-0002-7631-3092>

REFERENCES

- Ashworth, P.J. & Ferguson, R.I. (1986) Interrelationships of channel processes, changes and sediments in a proglacial braided river. *Geografiska Annaler, Series A: Physical Geography*, 68(4), 361–371. Available from: <https://doi.org/10.2307/521527>
- Bagnold, R.A. (1966) An approach to the sediment transport problem from general physics. In: Thorne, C.R., MacArthur, R.C. & Bradley, J.B. (Eds.) *The Physics of Sediment Transport by Wind and Water*. New York: American Society of Civil Engineers, pp. 231–291.
- Bagnold, R.A. (1980) An empirical correlation of bedload transport rates in flumes and natural rivers. In: Thorne, C.R., MacArthur, R.C. & Bradley, J.B. (Eds.) *The Physics of Sediment Transport by Wind and Water* (pp. 323–345). New York: American Society of Civil Engineers.
- Barker, D.M., Lawler, D.M., Knight, D.W., Morris, D.G., Davies, H.N. & Stewart, E.J. (2009) Longitudinal distributions of river flood power: The combined automated flood, elevation and stream power (CAFES) methodology. *Earth Surface Processes and Landforms*, 34(2), 280–290. Available from: <https://doi.org/10.1002/esp.1723>
- Benda, L. & Dunne, T. (1997) Stochastic forcing of sediment supply to channel networks from landsliding and debris flow. *Water Resources Research*, 33(12), 2849–2863. Available from: <https://doi.org/10.1029/97WR02388>
- Biron, P.M., Choné, G., Buffin-Bélanger, T., Demers, S. & Olsen, T. (2013) Improvement of streams hydro-geomorphological assessment using LiDAR DEMs. *Earth Surface Processes and Landforms*, 38(15), 1808–1821. Available from: <https://doi.org/10.1002/esp.3425>
- Bizzi, S. & Lerner, D.N. (2013) The use of stream power as an indicator of channel sensitivity to erosion and deposition processes. *River Research and Applications*, 31(1), 16–27. Available from: <https://doi.org/10.1002/rra.2717>
- Bizzi, S., Tangi, M., Schmitt, R.J.P., Pitlick, J., Piégay, H. & Castelletti, A.F. (2021) Sediment transport at the network scale and its link to channel morphology in the braided Vjosa River system. *Earth Surface Processes and Landforms*, 46(14), 2946–2962. Available from: <https://doi.org/10.1002/esp.5225>
- Bowman H, Jeffries R, Ing R, Hemsworth M, Todd-Burley N, Hankin B, Soar P, Thorne C. (2021). *Investigating ways to predict channel changes to inform flood risk management now and in the future*. Budapest. Paper presented at the 4th European Conference on Flood Risk Management – FLOODrisk 2020. Available from: <https://doi.org/10.3311/FloodRisk2020.2.26>
- Camenen, B., Grabowski, R.C., Latapie, A., Paquier, A., Solari, L. & Rodrigues, S. (2016) On the estimation of the bed-material transport and budget along a river segment: Application to the middle Loire River, France. *Aquatic Sciences*, 78(1), 71–81. Available from: <https://doi.org/10.1007/s00027-015-0442-3>
- Chicco, D., Tötsch, N. & Jurman, G. (2021) The Matthews correlation coefficient (MCC) is more reliable than balanced accuracy, bookmaker informedness, and markedness in two-class confusion matrix evaluation. *BioData Mining*, 14(1), 13. Available from: <https://doi.org/10.1186/s13040-021-00244-z>
- Conesa-García, C., Puig-Mengual, C., Riquelme, A., Tomás, R., Martínez-Capel, F., García-Lorenzo, R., et al. (2022) Changes in stream power and morphological adjustments at the event-scale and high spatial resolution along an ephemeral gravel-bed channel. *Geomorphology*, 398, 108053. Available from: <https://doi.org/10.1016/j.geomorph.2021.108053>
- Costa JE, O'Connor JE. 2011. Geomorphically effective floods. In *Natural and Anthropogenic Influences in Fluvial Geomorphology*, Costa JE, Miller AJ, Potter KW, Wilcock PR (eds). American Geophysical Union: Washington, DC; 45–56.
- Czuba, J.A., Foufoula-Georgiou, E., Gran, K.B., Belmont, P. & Wilcock, P.R. (2017) Interplay between spatially explicit sediment sourcing, hierarchical river-network structure, and in-channel bed material sediment transport and storage dynamics. *Journal of Geophysical Research: Earth Surface*, 122(5), 1090–1120. Available from: <https://doi.org/10.1002/2016JF003965>
- Doyle, M.W. & Shields, C.A. (2008) An alternative measure of discharge effectiveness. *Earth Surface Processes and Landforms*, 33(2), 308–316. Available from: <https://doi.org/10.1002/esp.1543>
- Environment Agency. (2003) *River Habitat Survey in Britain and Ireland Field Survey Guidance Manual: 2003 Version*. Bristol: Environment Agency.
- Eagle, L.J.B., Carrivick, J.L., Milner, A.M., Brown, L.E. & Klaar, M.J. (2021) Repeated high flows drive morphological change in rivers in recently deglaciated catchments. *Earth Surface Processes and Landforms*, 46(7), 1294–1310. Available from: <https://doi.org/10.1002/esp.5098>
- Eaton, B.C. & Church, M. (2011) A rational sediment transport scaling relation based on dimensionless stream power. *Earth Surface Processes and Landforms*, 36(7), 901–910. Available from: <https://doi.org/10.1002/esp.2120>
- Ekka, A., Pande, S., Jiang, Y. & van der Zaag, P. (2020) Anthropogenic modifications and river ecosystem services: A landscape perspective. *Water*, 12(10), 2706. Available from: <https://doi.org/10.3390/w12102706>
- Feeney, C.J., Godfrey, S., Cooper, J.R., Plater, A.J. & Dodds, D. (2022) Forecasting riverine erosion hazards to electricity transmission towers under increasing flow magnitudes. *Climate Risk Management*, 36, 100439. Available from: <https://doi.org/10.1016/j.crm.2022.100439>
- Gervasi, A.A., Pasternack, G.B. & East, A.E. (2021) Flooding duration and volume more important than peak discharge in explaining 18 years of gravel–cobble river change. *Earth Surface Processes and Landforms*, 46(15), 3194–3212. Available from: <https://doi.org/10.1002/esp.5230>
- Ghunowa, K., MacVicar, B.J. & Ashmore, P. (2021) Stream power index for networks (SPIN) toolbox for decision support in urbanizing watersheds. *Environmental Modelling & Software*, 144, 105185. Available from: <https://doi.org/10.1016/j.envsoft.2021.105185>
- Gill, D. (1970) Application of a statistical zonation method to reservoir evaluation and digitized-log analysis. *Bulletin of the American Association of Petroleum Geologists*, 54, 719–729. Available from: <https://doi.org/10.1306/5D25CA35-16C1-11D7-8645000102C1865D>
- Hauer C, Leitner P, Unfer G, Pulg U, Habersack H, Graf W. 2018. The role of sediment and sediment dynamics in the aquatic environment. In *Riverine Ecosystem Management*, Schmutz S, Sendzimir J (eds). Springer International: Cham; 151–169.
- Hendry, K., Cragg-Hine, D., O'Grady, M., Sambrook-Smith, H. & Stephen, A. (2003) Management of habitat for rehabilitation and enhancement of salmonid stocks. *Fisheries Research*, 62(2), 171–192. Available from: [https://doi.org/10.1016/S0165-7836\(02\)00161-3](https://doi.org/10.1016/S0165-7836(02)00161-3)
- Huang, J. & Ling, C.X. (2005) Using AUC and accuracy in evaluating learning algorithms. *IEEE Transactions on Knowledge and Data Engineering*, 17(3), 299–310. Available from: <https://doi.org/10.1109/TKDE.2005.50>
- Knighton, A.D. (1999) Downstream variation in stream power. *Geomorphology*, 29(3–4), 293–306. Available from: [https://doi.org/10.1016/S0169-555X\(99\)00015-X](https://doi.org/10.1016/S0169-555X(99)00015-X)
- Lammers, R.W. & Bledsoe, B.P. (2018a) A network scale, intermediate complexity model for simulating channel evolution over years to decades. *Journal of Hydrology*, 566, 886–900. Available from: <https://doi.org/10.1016/j.jhydrol.2018.09.036>
- Lammers, R.W. & Bledsoe, B.P. (2018b) Parsimonious sediment transport equations based on Bagnold's stream power approach. *Earth Surface Processes and Landforms*, 43(1), 242–258. Available from: <https://doi.org/10.1002/esp.4237>
- Lane, E.W. (1955) Design of stable channels. *Transactions of the American Society of Civil Engineers*, 120(1), 1234–1279. Available from: <https://doi.org/10.1061/TACEAT.0007188>

- Lea, D.M. & Legleiter, C.J. (2016) Mapping spatial patterns of stream power and channel change along a gravel-bed river in northern Yellowstone. *Geomorphology*, 252, 66–79. Available from: <https://doi.org/10.1016/j.geomorph.2015.05.033>
- Li, X., Cooper, J.R. & Plater, A.J. (2021) Quantifying erosion hazards and economic damage to critical infrastructure in river catchments: Impact of a warming climate. *Climate Risk Management*, 32, 100287. Available from: <https://doi.org/10.1016/j.crm.2021.100287>
- Lorenz, A., Hering, D., Feld, C. & Rolauffs, P. (2004) A new method for assessing the impact of hydromorphological degradation on the macroinvertebrate fauna of five German stream types. *Hydrobiologia*, 516(1–3), 107–127. Available from: <https://doi.org/10.1023/B%3AHYDR.0000025261.79761.b3>
- Major, J.J., Spicer, K.R. & Mosbrucker, A.R. (2021) Effective hydrological events in an evolving mid-latitude mountain river system following cataclysmic disturbance—a saga of multiple influences. *Water Resources Research*, 57(2), 1–28. Available from: <https://doi.org/10.1029/2019WR026851>
- Marcinkowski, P., Kiczko, A. & Kardel, I. (2022) Large-scale assessment of spatial variability in stream power distribution: An indicator of channel erosion and deposition potential. *Journal of Hydrology*, 605, 127319. Available from: <https://doi.org/10.1016/j.jhydrol.2021.127319>
- Martin, Y. & Church, M. (2000) Re-examination of Bagnold's empirical bedload formulae. *Earth Surface Processes and Landforms*, 25(9), 1011–1024. Available from: [https://doi.org/10.1002/1096-9837\(200008\)25:9<1011::AID-ESP114>3.0.CO;2-H](https://doi.org/10.1002/1096-9837(200008)25:9<1011::AID-ESP114>3.0.CO;2-H)
- Mazgareanu, I., Biron, P.M. & Buffin-Bélanger, T. (2020) A fuzzy GIS model to determine confluence morphological sensitivity to tributary inputs at the watershed scale. *Geomorphology*, 357, 107095. Available from: <https://doi.org/10.1016/j.geomorph.2020.107095>
- Naulin, M., Kortenhuis, A. & Oumeraci, H. (2015) Reliability-based flood defense analysis in an integrated risk assessment. *Coastal Engineering Journal*, 57(1), 1540005-1–1540005-35. Available from: <https://doi.org/10.1142/S0578563415400057>
- Newton, M.D., Clark, M.J., Sear, D.A. & Brookes, A. (1998) The geomorphological basis for classifying rivers. *Aquatic Conservation-Marine and Freshwater Ecosystems*, 8, 415–430.
- O'Brien, G.R., Wheaton, J.M., Fryirs, K., Macfarlane, W.W., Brierley, G., Whitehead, K., et al. (2019) Mapping valley bottom confinement at the network scale. *Earth Surface Processes and Landforms*, 44(9), 1828–1845. Available from: <https://doi.org/10.1002/esp.4615>
- Owens, P. (2005) Conceptual models and budgets for sediment management at the river basin scale. *Journal of Soils and Sediments*, 5(4), 201–212. Available from: <https://doi.org/10.1065/jss2005.05.133>
- Papangelakis, E., MacVicar, B., Ashmore, P., Gingerich, D. & Bright, C. (2022) Testing a watershed-scale stream power index tool for erosion risk assessment in an urban river. *Journal of Sustainable Water in the Built Environment*, 8(3). Available from: <https://doi.org/10.1061/JSWBAY.0000989>
- Parker, C., Clifford, N.J. & Thorne, C.R. (2011) Understanding the influence of slope on the threshold of coarse grain motion: Revisiting critical stream power. *Geomorphology*, 126(1–2), 51–65. Available from: <https://doi.org/10.1016/j.geomorph.2010.10.027>
- Parker, C., Thorne, C.R. & Clifford, N.J. (2015) Development of ST:REAM: A reach-based stream power balance approach for predicting alluvial river channel adjustment. *Earth Surface Processes and Landforms*, 40(3), 403–413. Available from: <https://doi.org/10.1002/esp.3641>
- Phillips, J.D. (2007) The perfect landscape. *Geomorphology*, 84(3–4), 159–169. Available from: <https://doi.org/10.1016/j.geomorph.2006.01.039>
- Piégay, H., Arnaud, F., Belletti, B., Bertrand, M., Bizzi, S., Carbonneau, P., et al. (2020) Remotely sensed rivers in the Anthropocene: State of the art and prospects. *Earth Surface Processes and Landforms*, 45(1), 157–188. Available from: <https://doi.org/10.1002/esp.4787>
- Raven, P.J., Holmes, N.T.H., Dawson, F.H., Fox, P.J.A., Everard, M., Fozzard, I.R., et al. (1998) *River Habitat Quality: The Physical Character of Rivers and Streams in the UK and Isle of Man*. Bristol: Environment Agency.
- Reinfelds, I., Cohen, T., Batten, P. & Brierley, G. (2004) Assessment of downstream trends in channel gradient, total and specific stream power: A GIS approach. *Geomorphology*, 60(3–4), 403–416. Available from: <https://doi.org/10.1016/j.geomorph.2003.10.003>
- Righini, M., Surian, N., Wohl, E., Marchi, L., Comiti, F., Amponsah, W., et al. (2017) Geomorphic response to an extreme flood in two Mediterranean rivers (northeastern Sardinia, Italy): Analysis of controlling factors. *Geomorphology*, 290, 184–199. Available from: <https://doi.org/10.1016/j.geomorph.2017.04.014>
- Rinaldi, M., Simoncini, C. & Piégay, H. (2009) Scientific design strategy for promoting sustainable sediment management: The case of the Magra River (central-northern Italy). *River Research and Applications*, 25(5), 607–625. Available from: <https://doi.org/10.1002/rra.1243>
- Schmitt, R.J.P., Bizzi, S. & Castelletti, A. (2016) Tracking multiple sediment cascades at the river network scale identifies controls and emerging patterns of sediment connectivity. *Water Resources Research*, 52(5), 3941–3965. Available from: <https://doi.org/10.1002/2015WR018097>
- Schmitt, R.J.P., Bizzi, S., Castelletti, A.F. & Kondolf, G.M. (2018) Stochastic modeling of sediment connectivity for reconstructing sand fluxes and origins in the unmonitored Se Kong, Se San, and Sre Pok tributaries of the Mekong River. *Journal of Geophysical Research: Earth Surface*, 123(1), 2–25. Available from: <https://doi.org/10.1002/2016JF004105>
- Sear, D., Newson, M.D. & Thorne, C.R. (2010) *Guidebook of Applied Fluvial Geomorphology*. London: Thomas Telford.
- SEPA. (2013). *Identifying Opportunities for Natural Flood Management*. Holytown: Scottish Environment Protection Agency.
- Sholtes, J.S., Yochum, S.E., Scott, J.A. & Bledsoe, B.P. (2018) Longitudinal variability of geomorphic response to floods. *Earth Surface Processes and Landforms*, 43(15), 3099–3113. Available from: <https://doi.org/10.1002/esp.4472>
- Soar, P.J., Wallerstein, N.P. & Thorne, C.R. (2017) Quantifying river channel stability at the basin scale. *Water*, 9(2), 133. Available from: <https://doi.org/10.3390/w9020133>
- Soulsby, C., Youngson, A.F., Moir, H.J. & Malcom, I.A. (2001) Fine sediment influence on salmonid spawning habitat in a lowland agricultural stream: A preliminary assessment. *The Science of the Total Environment*, 265(1–3), 295–307. Available from: [https://doi.org/10.1016/S0048-9697\(00\)00672-0](https://doi.org/10.1016/S0048-9697(00)00672-0)
- Stover, S.C. & Montgomery, D.R. (2001) Channel change and flooding, Skokomish River, Washington. *Journal of Hydrology*, 243(3–4), 272–286. Available from: [https://doi.org/10.1016/S0022-1694\(00\)00421-2](https://doi.org/10.1016/S0022-1694(00)00421-2)
- van de Wiel, M.J., Coulthard, T.J., Macklin, M.G. & Lewin, J. (2007) Embedding reach-scale fluvial dynamics within the CAESAR cellular automaton landscape evolution model. *Geomorphology*, 90(3–4), 283–301. Available from: <https://doi.org/10.1016/j.geomorph.2006.10.024>
- Vocal Ferencevic, M. & Ashmore, P.E. (2012) Creating and evaluating digital elevation model-based stream-power map as a stream assessment tool. *River Research and Applications*, 28(9), 1394–1416. Available from: <https://doi.org/10.1002/rra.1523>
- Wilcock, P.R. & Crowe, J.C. (2003) Surface-based transport model for mixed-size sediment. *Journal of Hydraulic Engineering - ASCE*, 129(2), 120–128. Available from: [https://doi.org/10.1061/\(ASCE\)0733-9429\(2003\)129:2\(120\)](https://doi.org/10.1061/(ASCE)0733-9429(2003)129:2(120))
- Wobus, C., Whipple, K.X., Kirby, E., Snyder, N., Johnson, J., Spyropoulou, K., et al. (2006) Tectonics from topography: Procedures, promise, and pitfalls. *Geological Society of America Special Paper*, 398, 55–74. Available from: [https://doi.org/10.1130/2006.2398\(04\)](https://doi.org/10.1130/2006.2398(04))
- Wohl, E., Bledsoe, B.P., Jacobson, R.B., Poff, N.L., Rathburn, S.L., Walters, D.M., et al. (2015) The natural sediment regime in rivers: Broadening the foundation for ecosystem management. *Bioscience*, 65(4), 358–371. Available from: <https://doi.org/10.1093/biosci/biv002>

- Wolman MG, Emmett WW, Whiting PJ, Thomas R, King J. 1997. United States' Expert Report Disclosing Methodologies for Quantification of Organic Act Claims. Consolidated Subcase No. 63-25243. Prepared for the Court.
- Yochum, S.E., Sholtes, J.S., Scott, J.A. & Bledsoe, B.P. (2017) Stream power framework for predicting geomorphic change: The 2013 Colorado Front Range flood. *Geomorphology*, 292, 178-192. Available from: <https://doi.org/10.1016/j.geomorph.2017.03.004>

How to cite this article: Parker, C. & Davey, J. (2023) Stream power indices correspond poorly with observations of alluvial river channel adjustment. *Earth Surface Processes and Landforms*, 1-15. Available from: <https://doi.org/10.1002/esp.5550>