

AUTOMATIC DELINEATION OF FUNCTIONAL RIVER REACH BOUNDARIES FOR RIVER RESEARCH AND APPLICATIONS.

CHRIS PARKER*, NICHOLAS J. CLIFFORD**, AND COLIN R. THORNE***

**Corresponding author. Department of Geography and Environmental Management,
University of the west of England, Bristol, BS16 1QY. 0117 3283116.*

Chris2.Parker@UWE.ac.uk.

*** Department of Geography, King's College London, London, WC2R 2LS.*

****School of Geography, University of Nottingham, University Park, Nottingham, NG7 2RD.*

ABSTRACT

The river reach is a pervasive term within contemporary river research and applications. Yet, despite its prevalence, there is a notable lack of consistency in its definition. This paper identifies the prevalence of two broad types of reach definition within the academic literature: operational, and functional, and argues that a functional definition is more suitable for applications within river research and management. A range of statistical sequence zonation algorithms that were originally derived for geological well log analysis are compared for their ability to automatically identify functional reach boundaries. An analysis of variance -based global boundary hunting algorithm is identified as the most suitable. Two potential practical applications of automatic reach delineation are demonstrated: identification of functional reach boundaries first, in a sequence of predicted sediment transport capacity for reach-based sediment transport models; and second, in a sequence of RHS Habitat Quality Assessment scores for identification of areas in a catchment in need of habitat restoration efforts. Finally, the paper discusses how this type of functional reach identification procedure might be applied in many other areas of river research and applications, and how a multivariate version of a statistical zonation algorithm might prove useful in facilitating integrated catchment management by identifying reach boundaries common across all variables of interest in the system.

KEY WORDS: river management; reach; zonation; boundary hunting; integrated catchment management; sediment transport; ST:REAM; RHS

INTRODUCTION

The pervasive yet obscure nature of the river reach

The “river reach”, as a term of reference, plays an important role within both the academic and applied riverine communities. Its use is ubiquitous across sub-disciplines and across geographical regions, as evidenced from the following series of quotations:

“The bed load formulae examined here are all one-dimensional equations parameterized by *reach*-average hydrologic and sedimentologic variables” (Barry *et al.*, 2004: 18); “For all sites the flow cross sectional data were averaged from measurements made at several cross sections along a *reach*...” (Bathurst, 2002: 18); “...framework for modelling size selective transport and sorting, capable of being implemented in an RCM at the *reach* scale” (Brasington and Richards, 2007: 174); “For the goodness of fit, data were compared first, for the entire *reach*, and subsequently for riffle and pool *sub-reaches*.” (Clifford *et al.*, 2005: 3635); “PHABSIM assesses the habitat ‘performance’ of a *reach* by defining its ‘usable area’ for a particular (target) species, based on a function of discharge and channel structure.” (Emery *et al.*, 2003: 534); “Information on the magnitude and variability of flow regimes at the river *reach* scale is central to aspects of water resources and water quality management.” (Holmes *et al.*, 2002: 721); “The overall effect on catchment-scale flood generation will be a function of the spatial location and extent of the landscape areas and river channel *reaches* affected”(O’Connell *et al.*, 2005: 14).

This prevalence results from the need for a manageable reference scale within natural systems where the interrelations between forms and processes are inherently scale-dependent, and where this scale dependency has some form of functional attribute. For example, in order to collect information on UK river habitat status, the Environment Agency’s River Habitat Survey involves the inspection of the physical structure of thousands of individual reaches across the UK (Raven *et al.*, 1998a). Here, the collection of habitat information in reaches is used to facilitate national coverage of catchments in manageable sample cases. Similarly, hydrologists often utilise the river reach as a reference-scale within broad-scale hydrological models. In order to operate at the broad-scale, these models simplify the catchment network into a series of interconnected reach-based “building blocks” (Paz and Collischonn, 2007). For geomorphologists, the reach represents a means of simplifying forms and processes that vary and interact over a continuum of spatial scales. This is evident most explicitly within reach-based sediment transport models such as the Riverine Accounting and Transport model (RAT: Graf, 1996), the Sediment Impact Assessment Method (SIAM: Gibson and Little,

2006), the River Energy Audit Scheme (REAS: Wallerstein *et al.*, 2006) and the Sediment Transport Reach Equilibrium Assessment Method (ST:REAM: Parker, 2010), where sediment transport processes which vary across a range of spatial scales are simplified into interactions between discrete sections of the river channel.

Yet, despite the prevalence of the term, definition of the river reach is far from consistent. The majority of studies fail to define their concept of a reach, and of those that do, there is considerable variation. The Environment Agency's RHS reaches are a standard 500m in length throughout the UK, yet for those hydrologists interested in broad-scale modelling a reach is often just the stretch of river between two junctions in a hydrological system (Hellweger and Maidment, 1999). This divergence of definition and use is not only true between sub-disciplines within river science, but also exists internally within each sub-discipline. For example, within the field of geomorphology, a reach has been defined as: a minimum of 10-20 channel widths in a classification of mountain channel morphologies (Montgomery and Buffington, 1997); falling between tributary junctions and grid cell boundaries in a catchment-scale sediment routing model (Benda and Dunne, 1997); and as a geomorphologically homogeneous stretch of river, the boundaries of which are defined by observed changes in channel morphology (Eyquem, 2007). In fact, the only definition that can completely encompass all observed usages of the term reach is "a length of river"!

This range of definition and use, may, in fact, hinder research and application in river science. As a simplified example, within a single project one fluvial geomorphologist may perform a reach-scale channel morphology survey over ten channel widths that is to be used by another fluvial geomorphologist within a piece of modelling software whose reach length is 1km. Contemporary river science has seen calls for a shift towards a fully integrated multidisciplinary approach to catchment management driven in particular by the European Water Framework Directive (Harper and Ferguson, 1995; Newson, 2002; Raven *et al.*, 2002; Eyquem, 2007; Orr *et al.*, 2008). Yet, the inconsistent usage of a term that is applied so widely across all river science disciplines could constitute a major obstruction to this idealised goal and form of river management.

Operational and functional definitions of the river reach

Existing definitions of river reaches currently fall into two major types: operational definitions that describe reach lengths in terms of a set spatial distances, for example 500m or 10 channel widths; and functional definitions that describe reach lengths based on the distance over which a certain form or process occurs, for example between hydrological

channel junctions, changes in channel morphology, or shifts in habitat structure. When looking for consistency in terminology, an operational definition of constant length initially seems the obvious choice – a reach is a stretch of channel 1km in length, for example. But what they offer in consistency, these types of operational definitions lack in flexibility, and may map onto channel forms or processes purely serendipitously. Not only may a uniform reach length definition of 1km be unsuitable for applications where the topic of interest varies over scales of 100m or 10km, but also, the assigned reach boundaries are unlikely to occur at natural breaks in the form or process of interest. As a result, operationally defined reach-averaged representations are potentially unreliable, and inter-reach comparisons potentially meaningless.

On the other hand, functional reach definitions are necessarily inconsistent in terms of reach length, even within a single study. Within fluvial geomorphology, a channel reach is often functionally defined as a stretch of river composed of largely homogeneous geomorphological units, the boundaries of which are defined by significant changes in morphology (Eyquem, 2007). For example, when Graf (1996) identified 11 functional reaches along the ~20km long Los Alamos Canyon in New Mexico, they ranged from 61m to 4568m in length. Yet, despite their irregular lengths, these reaches were consistent in terms of their definition: they each represented specific and definable segments with internally constant processes and forms that differed from those of neighbouring segments. Functional definitions such as this therefore offer a consistent (in terms of definition) yet flexible (in terms of length) means of describing reaches. As a generic methodology, defining reaches as lengths of channel within which forms and/or processes are more, rather than less, similar has the potential to be utilised across all river science disciplines. For example, a river biologist may define river reaches in terms of stretches of river inhabited by relatively consistent species types, the boundaries of which represent significant changes in those qualities. Similarly this type of definition could equally be applied to water quality scientists, ecologists and hydrologists. Implicitly, such definitions are suggested in various river classifications and typologies, most notably those which couple channel plan and cross-sectional form with relative position and ecological function within a catchment such as the River Styles Framework (Brierley *et al.*, 2002). By explicitly assigning functional reaches based on minimising within reach variation in the form or process of interest, river researchers and practitioners could more clearly justify the assumption that, in catchment-models, and for catchment research and management, reaches are lengths of channel with sequential and scale-dependent homogeneous properties.

One major obstacle that prevents the practical application of functional definitions of river reaches is that they generally require detailed *a priori* knowledge of the system in question in order to identify where the functional reach boundaries lie. In Graf's (1996) research, this was possible since he was dealing with a relatively small, and intensively studied research catchment. In many river management applications, this is not the case – functional reach boundaries need to be identified without the resources available to explore the catchment in detail. Therefore, if functional definitions of reach boundaries are to be widely applied within river science, a means of identification that does not require detailed *a priori* knowledge is required. With the continued increase in the quantity and quality of data of interest available to river scientists at the catchment scale, this is now possible using statistical techniques from data available in digital form and derived from remotely-sensed and map sources, as well as field survey. For example, in their study of planform dynamics of the Lower Mississippi River, Harmar and Clifford (2006) divided the river into a series of reaches of similar planform characteristics, on the basis of a statistical zonation algorithm to a data series based on lateral direction changes digitised from historic maps. This approach represented a novel means of objectively identifying functional reach boundaries without in depth *a priori* knowledge of the processes acting at the study site.

The purpose of this paper is to build upon the work of Harmar and Clifford (2006; 2007) and explore the potential of using statistical methods to define functional reach boundaries based on the concept that the reach characteristics of interest are internally homogenous and comparatively distinct. First, a number of alternative statistical methods for identifying reach boundaries based on techniques currently used within geology for identification of stratigraphic units are introduced. Second, the suitability of each of these alternative methods is assessed using a univariate test dataset of sediment transport capacity along the River Taff catchment in South Wales. Based on this assessment, a preferred method for delineating river reach boundaries is identified. Third, two potential practical applications of the chosen reach delineation method are illustrated: identification of functional reach boundaries in sequences of predicted sediment transport capacity for reach-based sediment transport models; and identification of functional reach boundaries in a sequence of RHS Habitat Quality Assessment scores for identification of areas in a catchment in need of habitat restoration efforts. Finally, the utility of automatic reach delineation methods, including how they may assist in integrated catchment management, is discussed.

STUDY SITE, DATA SERIES AND METHODS

Study site and data sequence: sediment transport capacity along the River Taff, South Wales

The data sequence used to test the algorithms within this paper is the predicted bed material transport capacity for the Mean Annual Flow ($Q_{s,MAF}$) along the main stem of the River Taff in South Wales. (Predicted bed material transport capacity was selected as the test variable as it is the parameter of interest within ST:REAM, a reach-based sediment balance model used to demonstrate a potential application of automatic functional reach boundary hunting algorithms used later in the paper.)

The Taff is a large river in South Wales (Figure 1). Its main stem rises in the Brecon Beacons south-west of Pen Y Fan and flows over 60km south to enter the Severn Estuary at Cardiff. Several major tributaries join the river: the Nant Ffrwd, Taff Fechan, Nant Morlais, Nant Rhydyar, Taff Bargoed, Cynon and Rhondda. Combined, these channels drain approximately 500km². Annual rainfall across the catchment ranges from 950mm/year at Cardiff to 2400mm/year in the Brecon Beacons.

Figure 1. Location map of the River Taff catchment in South Wales.

In order to generate a sequence of $Q_{s,MAF}$ values along the Taff, main stem channel slope, width and annual flow distribution values were obtained every 50m along the channel: channel slope was derived from a Environment Agency LiDAR digital elevation model of the catchment (Figure 2A); Ordnance Survey MasterMap data were used to measure channel widths (Figure 2B); and annual flow distribution curves were derived by applying an ungauged catchment flow estimation methodology similar to Low Flows 2000 (Young *et al.*, 2003) (Figure 2C). These datasets were used to calculate $Q_{s,MAF}$ at each point in the sequence by assuming a consistent bed material type of cobble and applying a general stream power-based bed material transport function that relates a dimensionless unit width sediment transport rate (q_s^*) to a dimensionless unit width stream power (ω^*)

$$q_s^* = \begin{cases} 100 \cdot \omega^{*6} & \text{for } \omega^* < 0.25 \\ 0.2 \cdot \omega^{*1.5} & \text{for } \omega^* \geq 0.25 \end{cases} \quad (1)$$

where $q_s^* = \frac{Q_s/w}{g \cdot (\rho_s - \rho_w) \cdot \sqrt{\frac{\rho_s - \rho_w}{\rho_w} \cdot g \cdot D_i^3}}$, $\omega^* = \frac{\omega}{g \cdot (\rho_s - \rho_w) \cdot \sqrt{\frac{\rho_s - \rho_w}{\rho_w} \cdot g \cdot D_i^3}}$, Q_s is the predicted rate of bed material transport at the discharge of interest (Q_{MAF}) in N^1s^{-1} (submerged weight), w is the active channel width in m, ω is the unit width stream power at the discharge of interest (Q_{MAF}) in $N^1m^{-1}s^{-1}$ calculated using $\omega = (Q \cdot S \cdot g \cdot \rho_w)/w$, Q is the discharge of interest (Q_{MAF}) in m^3s^{-1} , S is the energy slope (approximated by channel slope), g is gravitational acceleration in m^2s^{-1} (assumed to be 9.81), ρ_s is the density of sediment in kg^1m^{-3} (assumed to be 2650), ρ_w is the density of water in kg^1m^{-3} (assumed to be 1000), and D_i is the assumed diameter of the bed material being transported in m^1 (assumed to be 0.1 throughout the Taff). Further details of this bed material transport function are given by Parker (2010). Figure 2D shows the resultant sequence of $Q_{s,MAF}$ values along the main stem of the River Taff.

Figure 2. Data used for the River Taff main stem, values every 50m. (A) LiDAR channel elevation and slope values; (B) Mastermap width channel values; (C) Q_{MAF} values derived from catchment variables; (D) Q_{MAF} bed material transport capacity values.

Sequence zonation algorithms

Various sub-disciplines within geology have had a long interest in dividing data sequences into relatively uniform segments that are distinctive from other adjacent segments (Davis, 2002). Well logs need to be subdivided into relatively uniform sections that represent zones of consistent lithology and which therefore correspond to stratigraphic units. Paleontologists zone stratigraphic sequences on the basis of consistent abundance of microfossils. Airborne radiometric traverses may be subdivided into zones that can be interpreted as belts of uniform rock composition or mineralisation. Given their expertise in identifying relatively homogenous stratigraphic units, when attempting to identify relatively homogenous river units (reaches) it seems prudent to make use of progress already made within geology.

There are essentially two contrasting approaches to zonation: “local boundary hunting” and “global zonation” (Davis, 2002). Both methods are applied to data series which propagate in either time or space. Local boundary hunting procedures begin at one end of a sequence and progressively move to the other end, identifying abrupt changes in average

value. Webster (1973) proposed one of the original versions of this type of procedure. His method involved a sampling “window” of a specified width passing through each of the data points (Figure 3A). The window is divided into two parts either side of the point in question. The technique involves comparing the difference between the points located in either half of the window and plotting these differences as the window moves through the data sequence (Figure 3B). The principle is that the difference between the two sides of the window will be largest where the most significant discontinuities occur in the data. Various statistics may be used to quantify the difference between the two sides of the window. One of the more commonly applied is the generalised distance (D^2)

$$D^2 = \frac{(\bar{X}_1 - \bar{X}_2)^2}{s_1^2 + s_2^2} \quad (2)$$

where \bar{X}_1 and \bar{X}_2 are the mean values from either side of the window, and s_1^2 and s_2^2 are the variance for the data either side of the window. The zone (reach) boundaries are then identified based on the points in the series with the maximum intra-window generalised distances (Figure 3C).

Webster (1973) noted that the performance of the procedure depends upon the width of the moving window (see Figure 3B and 3C vs. Figure 3D and 3E). A wide window will average across small zones, subduing any erratic variation but also masking any local variability. A narrow window is more sensitive and will identify local changes in the sequence, but may also pick up noise from the original sequence.

****Figure 3. Demonstration of Webster’s local boundary hunting method on a downstream sequence of bed material transport capacity for the main stem of the River Taff, South Wales.*

*(A) Downstream plot of bed material transport capacity with sampling “window”; (B) Downstream plot of generalised distances in bed material transport capacity using a window width of 500m; (C) Downstream plot of reach-averaged bed material transport capacities using reach boundaries identified using a window width of 500m; (D) Downstream plot of generalised distances in bed material transport capacity using a window width of 2km; (E) Downstream plot of reach-averaged bed material transport capacities using reach boundaries identified using a window width of 2km****

By contrast, global zonation procedures break the sequence into segments which are as internally homogenous as possible and as distinct as possible from their adjacent segments. Unlike the local boundary hunting methods, global procedures consider the entire sequence at once, rather than just the portion within a window (Davis, 2002). Three different global zonation methods are considered here. The first was originally devised by Gill (1970) in order to analyse well-logs and applies an iterative analysis of variance approach. The data sequence begins as one long zone (reach) and is temporarily divided into two zones, with the provisional partition falling between the first and second points in the sequence. At this stage, the sum of squares within the two temporary zones (SS_w) is calculated using

$$SS_w = \sum_{j=1}^m \sum_{i=1}^{n_j} (x_{ij} - \bar{X}_{*j})^2 / \sum_{j=1}^m n_j - m \quad (3)$$

where x_{ij} is the i th point within zone j , \bar{X}_{*j} is the mean of the j th zone, n_j is the number of points in the j th zone, and m is the number of zones. Once this has been calculated, the partition between the two zones is moved along the sequence to successive positions and SS_w is calculated for every possible position of the partition. The partition which results in the lowest SS_w is selected as the first zonal boundary, forming two zones (Figure 4). Then, the procedure starts again, with the SS_w calculated for every possible position of the second partition, the minimum of which is used to divide the sequence into three zones. In this manner, Gill's (1970) method seeks to identify the zonation that minimises the variance within each zone (reach) and maximises the difference between the zones (reaches). The zonation procedure continues until the proportion of total variance explained by the zonation (R) increases beyond a specified level.

An alternative, global zonation procedure was published by Hawkins and Merriam (1973). They found that with a non-recursive procedure such as Gill's (1970), it is possible that the location chosen as the optimal partition at one stage in the iteration may no longer be optimal when looking to insert the next partition. As a solution, Hawkins and Merriam (1973) proposed a procedure that is similar to Gill's (1970), but is recursive and takes advantage of Bellman's principle of optimality to ensure that the final set of zone boundaries is the best possible combination. Like Gill's (1970) method, Hawkins and Merriam's (1973) recursive

procedure begins with the data sequence as one continuous zone. Then, whereas Gill's method makes its initial division based on the minimum SS_w from all potential locations of the first partition, Hawkins and Merriam's procedure calculates the SS_w value for every possible combination of the first *two* partitions. Once the combination of two partitions that results in the minimum SS_w is found, the procedure applies just the first of these partitions, dividing the sequence into two zones. In order to divide the sequence into three zones, the procedure considers which combination of the second and third partitions will result in the lowest SS_w . Again, the process continues until the proportion of total variance explained by the zonation (R) increases beyond a specified level. Because of the recursive nature of Hawkins and Merriam's method, for a given number of zones, it is guaranteed to have the smallest within zone variance out of all of the possible combinations. However, this optimality is achieved at higher computational cost.

****Figure 4. Demonstration of Gill's analysis-of-variance global zonation method on a downstream sequence of bed material transport capacity for the main stem of the River Taff, South Wales. The numbers represent the order in which the partitions are made in the sequence.****

Bohling *et al.* (1998) developed a zonation procedure based on a hierarchical cluster analysis in order to analyse well-logs. This method differs from that of Gill (1970) and Hawkins and Merriam (1973) in that, rather than starting with the data sequence as one contiguous zone, it begins with the data sequence divided into as many zones (reaches) as there are points in the sequence. The first iteration involves calculating the distance between the value of every zone i and its neighbour, $i + 1$. The pair of zones with the smallest difference is combined into one zone. In the next iteration, this new composite zone is treated as a single object defined by the mean value of the points within the zone. The process continues, with more and more zones being joined together based on their similarity to each other (Figure 5). Unlike the global zonation methods which reduce SS_w with every iteration, the cluster method begins with zero within-zone variance (SS_w) and each iteration results in an increase in SS_w until the proportion of total variance explained by the zonation (R) falls below a specified level.

*** Figure 5. Demonstration of Bohling's hierarchical cluster global zonation method on a downstream sequence of bed material transport capacity for the main stem of the River Taff, South Wales.***

For comparative purposes, all four of the sequence zonation algorithms described above were applied to the River Taff predicted sediment transport capacity data sequence and their relative performances were analysed.

RESULTS

Figure 6 compares the performance of each algorithm in defining internally homogenous and comparatively distinct reaches. Performance is based on how the proportion of variability explained by the reach boundaries (R) increases with the number of reaches identified, where $R = SS_B / (SS_B + SS_W)$, and SS_B is a measure of the variance of the reach means about the grand mean of the whole sequence (\bar{X}_{**})

$$SS_B = \sum_{j=1}^m (\bar{X}_{*j} - \bar{X}_{**})^2 / m - 1 \quad (4)$$

SS_W , j , m and \bar{X}_{*j} are as defined in Equation 3. This assumes that the higher the proportion of variability explained by the reach boundaries (R) for a given number of reaches, the stronger the performance, and hence the better suited the zonation algorithm is to defining functional river reaches.

The three global zonation algorithms all performed better than Webster's local boundary hunting method when dividing the initial sequence into reaches. The principal weakness of local boundary hunting procedures like Webster's (1973) within this type of application is that they are concerned with finding local breaks in the sequence, with little reference to the importance of these breaks in the sequence as a whole. Therefore, they do not necessarily prioritise the break points that are of most importance in minimising intra-reach and maximising inter-reach differences. Further, as identified by Davis (2002) and illustrated in both Figure 3 and Figure 6, the performance of local boundary hunting procedures is

dependent on the width of the moving window, a variable which is selected somewhat arbitrarily. This makes the effective application of a local boundary hunting method difficult without detailed *a priori* knowledge to inform the choice of window width.

The three global zonation algorithms explain similar proportions of variation when the initial data sequence is divided into multiple reaches. However, after the sequence has been divided into approximately 25 reaches, the proportion of variance explained by the Bohling algorithm falls below that explained when the same number of reaches identified by the Gill, and Hawkins and Merriam methods. A detailed comparison of the actual reach boundaries identified by the Gill and Bohling methods reveals an important difference in how they prioritise reach boundary placement. Figure 7 shows the reach extents identified by each of the two algorithms at two equivalent numbers of reach boundaries. The Bohling, clustering-based, method tends to identify boundaries at large, local inconsistencies in the data sequence. This is because it is at these points that the clustering method avoids grouping points either side into reaches. By contrast, the Gill, analysis of variance, method identifies boundaries based on broader-scale differences in the data sequence, rather than individual local/temporary inconsistencies. This means that individual large local variations are tolerated within a reach, as long as overall the total variation within all reaches is kept to a minimum. This is preferred since large local changes, such as steps in a step-pool reach, or weirs in a low gradient reach with a number of grade control structures, do not necessarily equate to functional reach boundaries when considered within the context of an entire catchment. This also means that the analysis of variance approaches are less sensitive to local discrepancies in data. On this basis, the analysis of variance approaches adopted by Gill (1970) and Hawkins and Merriam (1973) are better suited to the identification of river reach boundaries. Not only do they statistically minimise within reach variation and maximise between reach differences, but they also identify reach boundaries based on broad-scale functional changes since they are less influenced by local inconsistencies in the data sequence.

****Figure 6. Change in R (proportion of variability explained by reach boundaries) with the number of reach boundaries identified for each of the zonation algorithms considered. ****

****Figure 7. Comparison of reach boundaries selected by the Gill and Bohling zonation algorithms along the River Taff main stem. (A) Downstream plot of bed material transport*

*capacity; (B) Reaches identified by the two zonation algorithms after 5 boundaries; (C) identified by the two zonation algorithms after 17 boundaries.****

Unsurprisingly given their related structure, the Gill (1970) and Hawkins and Merriam (1973) methods explain almost exactly the same proportion of variation for every number of reach boundaries. In fact, the reach boundaries identified by these two algorithms are also nearly identical. As mentioned earlier, the recursive nature of Hawkins and Merriam's method comes at a high computational cost, especially when dealing with a large number of data points – a standard modern desktop computer (Dual Core 2Ghz processor, 2GB RAM) took more than seven days to place 100 reach boundaries in the Taff data sequence. Since, in this application at least, there is little difference between the reach boundaries identified by Gill's and Hawkins and Merriam's method, the faster, non-recursive algorithm would be preferred for practical reasons. Future research may yet, however, develop a new, more advanced algorithm that is even better suited to the identification of functional river reach boundaries. The remainder of this paper addresses how statistical identification of functional reach boundaries in general can be useful to river research and management using two examples, one of which builds upon the River Taff sediment transport capacity example.

PRACTICAL APPLICATIONS OF SEQUENCE ZONATION ALGORITHMS

Catchment-scale reach-based sediment transport modelling – River Taff, South Wales

An important concern in contemporary river science and management is the assessment of broad-scale sediment dynamics. To simplify the complex nature of sediment transfer through catchments, it is common practice to divide channels into relatively homogenous reaches, as demonstrated by reach-based sediment transport models such as SIAM (Gibson and Little, 2006), and RAT (Graf, 1996). SIAM is model that is designed to determine annual reach-based sediment budgets at the catchment scale. SIAM performs sediment transport computations by grain size class on user-defined reaches, and integrates the computed transport rates with flow duration information to compute average annual sediment transport capacities. These are compared with the annual quantity of sediment predicted to enter the reach to evaluate sediment continuity for all the reaches in the system (Gibson and Little, 2006). Similarly, RAT also defines sediment transport in a series of

linked user-defined reaches. The RAT model replicates fluvial processes across the system over a simulated time period by mathematically routing water and sediment through the sequence of reaches, calculating the average force available for transport within each reach, the average resistance of each reach, and the resulting erosion or deposition throughout each reach. However, both these reach-based methodologies currently require the user to manually identify river reach boundaries based on user knowledge of the catchment (e.g. Graf, 1996).

As identified earlier, this type of approach to reach delineation has its limitations, not simply because of its subjective nature, but also because it relies on the user being very familiar with the catchment - substantial investment into catchment reconnaissance may, for example, need to be made, which is often not feasible. Yet, without this knowledge of the catchment, user-designated reaches may be unrepresentative and consequently impact detrimentally on model outputs. Clearly, a means of automatically delineating reach boundaries like that demonstrated here would be an extremely useful addition to catchment-scale reach-based sediment models.

In order to demonstrate the potential usefulness of statistical sequence zonation algorithms for reach-based sediment models, the reaches defined by Gill's (1970) algorithm on the River Taff earlier were applied within a new reach-based sediment transport balance model, ST:REAM (Parker, 2010). In a manner similar to that of SIAM (Gibson and Little, 2006), ST:REAM treats a catchment network as a series of reaches and predicts reach sediment status based on the balance of annual sediment transport capacity between adjacent reaches, calculated using reach-averaged properties. ST:REAM differs from previous approaches as not only does it use a statistical zonation algorithm to automatically identify its own functional reach boundary locations, but it also explicitly restricts data inputs to those generally already available at the catchment scale: channel width, slope and annual flow distribution, without a requirement for reach-specific grain size information.

Having divided each of the branches in the catchment into discrete reaches based on the branches' longitudinal sequences of $Q_{s,MAF}$, for each reach the predicted total annual bed material transport capacity ($Q_{s,Annual}$) is calculated using

$$Q_{s,Annual} = \sum_{k=1}^p T_k \cdot Q_{s,k} \quad (5)$$

where T_k is the total time-span in s^1 , of discharge class k , p is the total number of discharge classes, and Q_{s_k} is the predicted bed material transport capacity of the reach at discharge class k calculated using a relationship equivalent to that described by Equation 1 and reach-averaged values of S , Q_k and w . Next, for each reach, the predicted total annual bed material transport capacity is compared with that of its immediate upstream neighbour (or neighbours if the reach in question is downstream of a confluence). Based on the assumption that the annual bed material supply delivered to a reach is equivalent to the predicted total annual bed material transport capacity of its upstream neighbour(s), a predicted capacity supply ratio (CSR) can be calculated. The CSR values give the ratio between the annual transport capacity of the reach, and the annual transport capacity of the reach(es) immediately upstream. Reaches that are identified as completely non-alluvial, such as weir crests and bedrock gorges, can be designated as such within the model and their maximum transport capacity is then limited to that of their upstream neighbour(s).

Figure 8 illustrates the outputs of the ST:REAM approach when applied to the Taff catchment using reach boundaries that explain 1% of the total variation in bed material transport capacity ($R=0.01$). The images labelled A to F were obtained as part of a reconnaissance of the Taff catchment performed to validate the outputs of the ST:REAM approach. In general, as shown by Figure 8, the outputs from ST:REAM correspond well with observations made during the catchment reconnaissance. Reaches predicted to have high $CSRs$ were generally observed to be starved of sediment compared with their upstream neighbours, whilst reaches with low $CSRs$ were generally observed to have excess sediment compared with their upstream neighbours.

This first example application demonstrates the efficacy of using a statistical zonation algorithm to identify functional reach boundaries within reach-based sediment transport models. Without a means to automatically define the reaches used within ST:REAM it would have been necessary to either make a significant investment in a detailed catchment reconnaissance, or use a constant operational reach length to define boundaries that did not align with shifts in sediment transport capacity. Incorporating a means of automatically identifying functional reach boundaries into other, existing reach-based sediment transport models such as SIAM and RAT would significantly increase their utility. Most recently, applications of REAS, another model which treats a catchment as a series of reaches and compares annual integrated stream power between adjacent reaches to predict reach sediment

status (Wallerstein *et al.*, 2006), have shown that by using a zonation algorithm to find reach boundaries, model outputs are less erratic and provide a clearer representation of broad-scale sediment dynamics (Soar and Wallerstein, in review).

An important consideration when applying a sequence zonation algorithm to distinguish reach boundaries within a reach-based sediment balance approach like ST:REAM is the proportion of variation explained by the reach boundaries (R). A higher value of R results in more reach boundaries and therefore smaller reach lengths. Therefore, R must be selected carefully with respect to the scale of examination required by the application. For example, an R value of 0.01 was deemed appropriate when applying ST:REAM to the Taff. With a higher R values the zonation algorithm began to identify individual riffle and pool units as separate reaches, whilst lower R values provided a representation of the catchment that was too general than that required for management purposes.

****Figure 8. Outputs from the ST:REAM representation of the Taff catchment based on functional reaches identified using Gill's (1970) approach to sequence zonation. Photos A-F were obtained during a catchment reconnaissance.****

Broad-scale reach habitat quality restoration prioritisation – River Frome, Somerset

Driven in part by the European Water Framework Directive (Harper and Ferguson, 1995; Newson, 2002; Raven *et al.*, 2002; Eyquem, 2007; Orr *et al.*, 2008), there has been increased interest in the role that river habitat restoration plays in river management. In order to achieve 'good ecological status' throughout a river catchment, river managers need to identify, and prioritise, those parts of a catchment that are in most need of habitat restoration. The most widely available data source that gives information on river habitat quality is the Environment Agency's River Habitat Survey (RHS), a system for assessing the character and habitat quality of rivers based on their physical structure (Raven *et al.*, 1998a). It has four distinct components: (i) a standard field survey method; (ii) a computer database, for entering results from survey sites and comparing them with information from other sites throughout the UK and Isle of Man; (iii) a suite of methods for assessing habitat quality; and (iv) a system for describing the extent of artificial channel modification. The RHS field method itself is a systematic collection of data associated with the physical structure of watercourses. As identified earlier, data collection is based on operationally defined reaches that are a constant 500m in length. Map information is collected for each 500m reach and includes grid

reference, altitude, slope, geology, height of source and distance from source. During the field survey, features of the channel (both in-stream and banks) and adjacent river corridor are recorded. Channel substrate, habitat features, aquatic vegetation types, the complexity of bank vegetation structure and the type of artificial modification to the channel and banks are all recorded at each of 10 'spot-checks' located at 50 m intervals. A 'sweep-up' checklist is also completed to ensure that features and modifications not occurring at the spot checks are recorded. Cross-section measurements of water and bankfull width, bank height and water depth are made at one representative location to provide information about geomorphological processes acting on the channel. The numbers of riffles, pools and point bars found in the site are also recorded. In all, more than 200 compulsory data entries are made at each site, collectively building a comprehensive picture of habitat diversity and character.

The Habitat Quality Assessment (HQA) scoring system is a broad measure of the diversity and 'naturalness' of physical (habitat) structure in the channel and river corridor based on data contained within the RHS. The HQA score of a site is determined by the presence and extent of habitat features of known wildlife interest recorded during the field survey according to a points tariff. Rare features nationally such as waterfalls more than 5 m high and extensive fallen trees attract additional points. Point scoring for the HQA system is based on informed professional judgment. Features that score within the HQA are consistent with those included in the 'System for Evaluating Rivers for Conservation' (SERCON), for which a panel of ecological experts identified the attributes of most value to riverine wildlife (Boon *et al.*, 1997).

Whilst Raven *et al.* (1998a) warn that comparison of HQA scores across different river types is not meaningful, by comparing the HQA scores at a reach with scores from elsewhere in the same catchment, it is possible to identify whether the location has a relatively high or low habitat quality status. Ideally, this information could be used by river managers to identify locations within a catchment where river habitat restoration efforts would be most effective. However, as demonstrated by Figure 9A, which shows the HQA scores for a series of RHS reaches along the River Frome in Somerset, it is not clear where sites of high or low habitat quality are located within the catchment. This is because the boundaries of the operationally defined, uniform 500m RHS reaches are not aligned with the forms and processes that influence habitat quality, but are instead selected arbitrarily. This means that variations in river habitat quality at a spatial scale finer than 500m would not be apparent from the RHS based HQA scores. Similarly, variations in river habitat quality at a spatial scale coarser than 500m would not be initially apparent based on the HQA scores

from the operationally defined 500m reaches; nor would habitat quality variations corresponding to a 500m spatial scale, because these would generally be only partially represented, unless the 500 m chosen were entirely coincident, by chance, with the habitat unit or pattern. Therefore, since the reach boundaries applied and the forms and processes of interest do not necessarily align, difficulty arises in identifying the location and spatial extent of the sections of channel over which habitat quality is low/high, and therefore the sections of channel over which management action is the highest/lowest priority.

Although not relevant to any functional reaches that may occur at a scale smaller than the 500m RHS reach length, an automatic reach delineation algorithm can be applied to a longitudinal sequence of HQA scores to identify functional reaches that are associated with broad-scale variations in river habitat quality. Figure 9B demonstrates how the sequence of HQA scores for the River Frome in Figure 9A can be automatically discretised into a series of functional reaches using Gill's (1970) analysis of variance global boundary hunting algorithm ($R = 0.05$). The detection of reaches based on statistically defined boundaries in the HQA data sequence, rather than on an arbitrarily selected uniform length, not only identifies meaningful changes in habitat quality but also results in a clear spatial structure to the data (Figure 9C). These outputs are therefore useful to river managers looking to prioritise parts of a catchment that are in need of river habitat restoration. Further, from a researcher's perspective, the broad-scale variations in habitat quality illustrated in Figure 9C could also be linked to other catchment variables in order to determine broad-scale drivers of habitat quality status.

As was recognised previously with reference to ST:REAM, selecting the proportion of variance explained by the reach boundaries (R) was an important decision when applying Gill's (1970) algorithm to the River Frome HQA sequence. Using values of R significantly higher than 0.05 resulted in a large number of reaches that were at a scale too fine for catchment management purposes. Conversely, using values of R significantly lower than 0.05 did not distinguish enough separate reaches to be useful for management purposes.

****Figure 9. Demonstration of automatic delineation of reach boundaries based on RHS Habitat Quality Assessment (HQA) scores on the River Frome, Somerset. (A) Downstream plot of HQA scores; (B) Mean HQA scores for reaches identified using Gill's (1970) approach to sequence zonation; (C) Map of mean HQA scores for reaches identified using Gill's (1970) approach to sequence zonation.****

DISCUSSION AND CONCLUSIONS

Given both the frequency of occurrence of the term river reach within river research and applications, and the current ambiguity in its meaning, there is much to be gained from seeking a reproducible and precise functional definition of the term. Based on an analysis of various statistical sequence zonation algorithms originally derived for geological well log analysis, Gill's (1970) global method has been shown to be the most suitable for the automatic identification of functional reach boundaries. This arises from its ability to minimise within reach variance / maximise inter-reach differences, and to consider broad-scale trends in data sequences without requiring excessive computational time. This statistical methodology ensures a consistency in definition, while allowing flexibility in actual reach length dependent on the nature of the form or process of interest.

Automatic functional reach delineation can be applied within all sub-disciplines of river research and management. Here, reach boundaries automatically identified using a method based on Gill's (1970) zonation algorithm have been successfully applied within a catchment-scale reach-based sediment balance model (ST:REAM) of the Taff catchment in Wales. This represents an improvement over existing reach-based sediment transport models where reach boundaries are identified based on the user's *a priori* knowledge of the catchment, which is often not very detailed. Gill's zonation algorithm was also used to split a sequence of Habitat Quality Assessment scores along the River Frome in Somerset into discrete reaches. This resulted in a clearer spatial representation of the way in which habitat quality varies in the catchment and could help river practitioners identify the location and extent of reaches that should be prioritised for habitat restoration.

As well as being valuable to river managers, being able to objectively identify functional reaches is also useful for researchers. As introduced earlier, Harmar and Clifford (2006) applied a statistical zonation algorithm to a sequence of lateral direction change values along the Lower Mississippi River in order to investigate the causes of spatial patterning in planform change. As with the analysis performed herein, Harmar and Clifford (2006) used a method similar to Gill's (1970) zonation algorithm to identify their reaches. By dividing the downstream sequence into distinct reaches of relatively homogenous planform type they could attempt to identify potential spatially distributed controls over river channel planform. Based on their findings Harmar and Clifford (2006) concluded that the division of their

downstream sequences into reaches helped reveal the time and spatial scales that characterised planform dynamics of meander trains within the Lower Mississippi River.

This type of automatic functional reach identification procedure could be applied to numerous other aspects of river management. For example, applications interested in how water quality varies spatially within the catchment could apply an automatic functional reach delineation procedure to a longitudinal series of Water Quality Index (WQI) values. The resultant discretised reaches and their averaged WQI values could help simplify downstream spatial variation and emphasise meaningful spatial structure in the data. Similarly, this can be applied to sequences of bed material size, invertebrate richness, channel width, or pollutant concentration.

The statistical basis of the analytical methods demonstrated here is common to all potential applications, but the nature of the data series used will vary. All of the applications considered so far have been based upon univariate data series. But these univariate series can differ in two ways. First, there are differences in terms of whether they either represent a form or status attribute (e.g. habitat quality, channel width or lead concentration) or a process attribute (e.g. predicted sediment transport rate, planform change or flow velocity). Second, there are differences in terms of whether the variable is either truly univariate (e.g. channel width, bed material size or flow velocity) or is formulated through combination of a number of different variables, e.g. predicted sediment transport rate, HQA score, or WQI value. Nevertheless, despite the variety of data types used, the suitability of global zonation methods of the kind applied here is unaffected: they do not depend on the distribution of the data, the sampling theory, nor any other underlying assumptions. In fact, the zonation methods are not strictly 'statistical'; they do not draw inferences about a population that is represented by a sample (Davis, 2002). It is the sequence itself that is of interest, and questions of sampling and probability are not relevant.

Within the applications of Gill's (1970) algorithm demonstrated here it has been identified that a key decision when attempting to automatically delineate reach boundaries is the selection of an appropriate R value. Choosing the proportion of variation to be explained by the automatically identified reach boundaries represents the primary subjective decision within application of the zonation algorithm. Unfortunately there is no ideal value of R that can be recommended for all applications: its selection depends upon the scale of examination required by the problem at hand.

While it has been demonstrated that global zonation procedures can be of important practical use to many aspects of river research and management by identifying functional river reaches based on univariate data series, a potentially more important application remains unexplored. Since the mid-1990s, the previous tradition of reporting on different aspects of rivers in isolation has been superseded by a more multidisciplinary approach based on the principles of integrated river basin management (Harper and Ferguson, 1995). Raven *et al.* (1998b), for example, suggested that, in order to facilitate such an integrated approach, all management “viewpoints” need to be based upon a consistent yet flexible approach. Yet, as identified at the outset, whilst the term “reach” is used extensively within all aspects of river research and applications there is little consistency in its definition. Clearly, when geomorphologists are referring to reaches that differ in extent to those reaches that are of interest to hydrologists or ecologists, it creates confusion that can hinder successful integrated catchment management.

Given the above, a further feature of interest in data sequence zonation algorithms is their potential extension to multivariate data sequences (Davis, 2002). To facilitate this, data sequences must be standardised so that they each have equal influence on the zonation, and the resultant boundaries represent divisions that are dominant across all of the contributing variables. There is consequently the potential to utilise an algorithm similar to that proposed by Gill (1970), but across several data sequences that each represent a variable of interest within an integrated catchment management approach. For example, functional reach boundaries could be identified based on sequences representing sediment transport capacity, habitat quality, water quality, invertebrate abundance, and low flow discharge. The resultant combined functional reaches could then act as a consistent spatial framework within which all of these aspects of riverine management could be considered.

Following from the above, further research into advancing the utility of automatic reach boundary hunting algorithms, not only into achieving the optimum statistical procedure for reach designation, but also more importantly, to explore how this type of technique can be deployed in developing practical strategies underpinning integrated catchment management, is recommended.

ACKNOWLEDGEMENTS

This work was carried out as part of a Ph.D. thesis funded by EPSRC doctoral training account *EP/P502632* as part of the Flood Risk Management Research Consortium. The method for the automatic identification of reach boundaries was developed based on earlier work carried out by Philip Soar (University of Portsmouth) and Nick Wallerstein (Heriot Watt University). LiDAR data for the Taff catchment was provided by the Environment Agency, Mastermap data was provided by Edina Digimap on behalf of Ordnance Survey, catchment hydrology data was obtained from the Centre for Ecology and Hydrology.

REFERENCES

- Barry, J.J., Buffington, J.M., King, J., 2004. A general power equation for predicting bed load transport rates in gravel bed rivers. *Water Resources Research*, doi:10.1029/2004WR003190.
- Bathurst, J.C., 2002. At-a-site variation and minimum flow resistance for mountain rivers. *Journal of Hydrology* 269, 11-26.
- Benda, L., Dunne, T., 1997. Stochastic forcing of sediment routing and storage in channel networks. *Water Resources Research* 33(12), 2865-2880.
- Bohling, G., Doveton, J., Guy, B., Watney, L., Bhattacharya, S., 1998. *PfEFFER 2.0 Manual*. Kansas Geological Survey, Lawrence, 164 pp.
- Boon, P.J., Holmes, N.T.H., Maitland, P.S., Rowell, T.A., Davies, J., 1997. A system for evaluating rivers for conservation (SERCON): development, structure and function. In: Boon, P.J. and Howell, D.L. (Eds.), *Freshwater Quality: Defining the Indefinable?* The Stationary Office, Edinburgh, pp. 299-326.
- Brasington, J., Richards, K., 2007. Reduced-complexity, physically-based geomorphological modelling for catchment and river management. *Geomorphology* 90, 171-177.
- Brierley, G., Fryirs, K., Outhet, D., Massey, C., 2002. Application of the River Styles framework as a basis for river management in New South Wales, Australia. *Applied Geography* 22(1), 91-122.
- Clifford, N.J., Soar, P.J., Harmar, O.P., Gurnell, A.M., Petts, G.E., Emery, J.C., 2005. Assessment of hydrodynamic simulation results for eco-hydraulic and eco-hydrological applications: a spatial semivariance approach. *Hydrological Processes* 19(18), 3631-3648.
- Davis, J.C., 2002. *Statistics and Data Analysis in Geology*. John Wiley & Sons, New York, 638 pp.
- Emery, J.C., Gurnell, A.M., Clifford, N.J., Petts, G.E., Morrissey, I.P., Soar, P.J., 2003. Classifying the hydraulic performance of riffle-pool bedforms for habitat assessment and river rehabilitation design. *River Research and Applications* 19(5-6), 533-549.

Eyquem, J., 2007. Using fluvial geomorphology to inform integrated river basin management. *Water and Environment Journal* 21(1), 54-60.

Gibson, S.A., Little, C.D., Jr., 2006. *Implementation of the Sediment Impact Assessment Model (SIAM) in HEC-RAS*. In: Bernard, J.M. (Ed.), Eighth Federal Interagency Sedimentation Conference. Federal Interagency Subcommittee on Sedimentation, Reno, Nevada, pp. 65-72.

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(5), 719-729.

Graf, W.L., 1996. Transport and deposition of plutonium-contaminated sediments by fluvial processes, Los Alamos Canyon, New Mexico. *Geological Society of America Bulletin* 108, 1342-1355.

Harmar, O.P., Clifford, N.J., 2006. Planform dynamics of the Lower Mississippi River. *Earth Surface Processes and Landforms* 31, 825-843.

Harmar, O.P., Clifford, N.J., 2007. Geomorphological explanation of the long profile of the Lower Mississippi River. *Geomorphology* 84, 222-240.

Harper, D., Ferguson, A.J.D., 1995. *The Ecological Basis for River Management*. John Wiley, Chichester, 630 pp.

Hawkins, D.M., Merriam, D.F., 1973. Optimal Zonation of Digitized Sequential Data. *Mathematical Geology* 5(4), 389-395.

Hellweger, F.L., Maidment, D.R., 1999. Definition and connection of hydrologic elements using geographic data. *Journal of Hydrologic Engineering* 4(1), 10-18.

Holmes, M.G.R., Young, A.R., Gustard, A., Grew, R., 2002. A region of influence approach to predicting flow duration curves within ungauged catchments. *Hydrology and Earth System Sciences* 6(4), 721-731.

Montgomery, D.R., Buffington, J.M., 1997. Channel-reach morphology in mountain drainage basins. *Geological Society of America Bulletin* 109(5), 596-611.

Newson, M.D., 2002. Geomorphological concepts and tools for sustainable river ecosystem management. *Aquatic Conservation-Marine and Freshwater Ecosystems* 12(4), 365-379.

O'Connell, P.E., Beven, K.J., Carney, J.N., Clements, R.O., Ewen, J., Fowler, H., Harris, G.L., Hollis, J., Morris, J., O'Donnell, G.M., Packman, J.C., Parkin, A., Quinn, P.F., Rose, S.C., Shepher, M., Tellier, S., 2005. *Review of impacts of rural land use and management on flood generation: Impact study report*. FD2114/TR, Department of the Environment, Farming and Rural Affairs R&D.

Orr, H.G., Large, A.R.G., Newson, M.D., Walsh, C.L., 2008. A predictive typology for characterising hydromorphology. *Geomorphology* 100, 32-40.

Parker, C., 2010. *Quantifying catchment-scale coarse sediment dynamics in British rivers*. Thesis[Ph.D. Dissertation], University of Nottingham, Nottingham.

Paz, A.R., Collischonn, W., 2007. River reach length and slope estimates for large-scale hydrological models based on a relatively high-resolution digital elevation model. *Journal of Hydrology* 343, 127-139.

Raven, P.J., Holmes, N.T.H., Dawson, F.H., Everard, M., 1998a. Quality assessment using River Habitat Survey data. *Aquatic Conservation-Marine and Freshwater Ecosystems* 8(4), 477-499.

Raven, P.J., Boon, P.J., Dawson, F.H., Ferguson, A.J.D., 1998b. Towards an integrated approach to classifying and evaluating rivers in the UK. *Aquatic Conservation-Marine and Freshwater Ecosystems* 8(4), 383-393.

Raven, P.J., Holmes, N.T.H., Charrier, P., Dawson, F.H., Naura, M., Boon, P.J., 2002. Towards a harmonized approach for hydromorphological assessment of rivers in Europe: a qualitative comparison of three survey methods. *Aquatic Conservation-Marine and Freshwater Ecosystems* 12(4), 405-424.

Soar, P., Wallerstein, N.P., in review. Characterising sediment transfer in river channels using stream power. *River Research and Applications*.

Wallerstein, N.P., Soar, P., Thorne, C.R., 2006. *River Energy Auditing Scheme (REAS) for catchment flood management planning*. In: Ferreira, R.M.L., Alves, E.C.T.L., Leal, J.G.A.B. and Cardoso, A.H. (Eds.), *River Flow 2006*. Taylor & Francis, Lisbon, Portugal.

Webster, R., 1973. Automatic soil-boundary location from transect data. *Mathematical Geology* 5(1), 27-37.

Young, A.R., Grew, R., Holmes, M.G.R., 2003. Low flows 2000: a national water resources assessment and decision support tool. *Water Science and Technology* 48(10), 119-126.

Figure Captions

Figure 1. Location map of the River Taff catchment in South Wales.

Figure 2. Data used for the River Taff main stem, values every 50m. (A) LiDAR channel elevation and slope values; (B) Mastermap width channel values; (C) Q_{MAF} values derived from catchment variables; (D) Q_{MAF} bed material transport capacity values.

Figure 3. Demonstration of Webster's local boundary hunting method on a downstream sequence of bed material transport capacity for the main stem of the River Taff, South Wales. (A) Downstream plot of bed material transport capacity with sampling "window"; (B) Downstream plot of generalised distances in bed material transport capacity using a window width of 500m; (C) Downstream plot of reach-averaged bed material transport capacities using reach boundaries identified using a window width of 500m; (D) Downstream plot of generalised distances in bed material transport capacity using a window width of 2km; (E) Downstream plot of reach-averaged bed material transport capacities using reach boundaries identified using a window width of 2km.

Figure 4. Demonstration of Gill's analysis of variance global zonation method on a downstream sequence of bed material transport capacity for the main stem of the River Taff, South Wales. The numbers represent the order in which the partitions are made in the sequence.

Figure 5. Demonstration of Bohling's hierarchical cluster global zonation method on a downstream sequence of bed material transport capacity for the main stem of the River Taff, South Wales.

Figure 6. Change in R (proportion of variability explained by reach boundaries) with the number of reach boundaries identified for each of the zonation algorithms considered.

Figure 7. Comparison of reach boundaries selected by the Gill and Bohling zonation algorithms along the River Taff main stem. (A) Downstream plot of bed material transport

capacity; (B) Reaches identified by the two zonation algorithms after 5 boundaries; (C) Reaches identified by the two zonation algorithms after 17 boundaries.

Figure 8. Outputs from the ST:REAM representation of the Taff catchment based on functional reaches identified using Gill's (1970) approach to sequence zonation. Photos A-F were obtained during a catchment reconnaissance.

Figure 9. Demonstration of automatic delineation of reach boundaries based on RHS Habitat Quality Assessment (HQA) scores on the River Frome, Somerset. (A) Downstream plot of HQA scores; (B) Mean HQA scores for reaches identified using Gill's (1970) approach to sequence zonation; (C) Map of mean HQA scores for reaches identified using Gill's (1970) approach to sequence zonation.