Geomorphology 194 (2013) 121–131

Contents lists available at SciVerse ScienceDirect

Geomorphology

journal homepage: www.elsevier.com/locate/geomorph

Successful predictions of river characteristics across England and Wales based on ordination

Ian P. Vaughan [a,*], Faye L. Merrix-Jones [a], José Antonio Constantine [b]

[a] Cardiff School of Biosciences, Cardiff University, Cardiff CF10 3AX, UK
[b] School of Earth & Ocean Sciences, Cardiff University, Cardiff CF10 3YE, UK

ARTICLE INFO

Article history:
Received 13 July 2012
Received in revised form 12 March 2013
Accepted 14 March 2013
Available online 22 April 2013

Keywords:
River Habitat Survey
Principal component analysis
Hydromorphology
Prediction
River classification

ABSTRACT

Predictions of river channel form under current conditions or in response to environmental or management changes and the rapid comparison of different channel reaches are important tasks in river management. River classification is a common and valuable framework to address these aims but may suffer from the necessity to force a continuum of channel morphology into discrete groups. More generally, the scope to test the ability of predictive tools has been limited because of a shortage of field data. In this study, we used principal component analysis (PCA) to identify the main sources of variation among river reaches in England and Wales based on a set of 20 variables expected to correlate with channel morphology. The PCA scores were then used to predict the distributions of a wide range of hydromorphic features based on >4000 reaches surveyed in the River Habitat Survey baseline. For comparison, the predictive ability of three pairs of variables (channel slope–discharge, slope–catchment area, and specific power–catchment area) was also tested. The PCA identified specific stream power, channel size and groundwater input as the main sources of variation among reaches. Regression models using PCA scores or paired variables were effective predictors of a range of channel characteristics, including predominant substrate, flow biotopes, and channel vegetation. Channel cross sections and anthropogenic modifications were less predictable. All of the approaches permitted simple plots of river reaches and quantitative comparisons of the (dis)similarity among individual reaches, whilst the paired variables also minimised the data requirements. Our work reiterates the value of simple, paired variables as a basis for rapidly comparing river reaches and, for the first time, quantifies the predictive ability of these approaches across a wide range of channel characteristics at a national scale. Principal component analysis provides a valuable exploratory tool for identifying the main sources of variation in complex, multivariate data from which a simplified version (e.g., specific power and area) could be adopted.

1. Introduction

When field data are not available, predictions of river channel form may be valuable in a range of contexts. Predictions may be made of the outcomes of river restoration schemes (e.g., Reichert et al., 2007; Steel et al., 2008) or of reference conditions, against which the observed channel form can be compared in site assessments (e.g., Davies et al., 2000; Frappier and Eckert, 2007). Predictive models can estimate the extent and distribution of different channel types or features in a catchment (e.g., Jeffers, 1998; Mugodo et al., 2006). More generally, implicit predictions are used widely in river research and management, identifying reaches that are predicted to respond to environmental or anthropogenic changes in similar ways, that require similar management interventions, or that support a similar flora and fauna (Orr et al., 2008). Such predictions underlie river classifications, providing important tools for river management, conservation, and communication among disciplines (Kondolf et al., 2003; Vaughan et al., 2009).

Recent decades have seen the development of a diverse range of fluvial classifications, with scales ranging from individual channel features (e.g., riffles and point bars) to entire catchments or regions (e.g., Frissell et al., 1986; Brierley and Fryirs, 2000), data collected in the field or from GIS and remote sensing and from reaches grouped according to empirical, taxonomic criteria, or by process-based typologies (Newson et al., 1998a; Kondolf et al., 2003; Orr et al., 2008). The fundamental assumption that unites all of these approaches is that the study entities can be separated into discrete categories. Support for this is found where geomorphic thresholds have been identified among channel forms, such as between bedrock and alluvial channels (e.g., Montgomery et al., 1996) or single-threaded and braided planforms (e.g., van den Berg, 1995). Conversely, even where classifications are developed, river environments are acknowledged to represent continua of

* Corresponding author. Tel.: +44 29 20874545.
E-mail address: vaughanip@cardiff.ac.uk (I.P. Vaughan).

0169-555X/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.geomorph.2013.03.036

channel forms (e.g., Nanson and Croke, 1992; Montgomery and Buffington, 1997; Kleinhans, 2010). In such situations, a more natural framework may be to quantify the (dis)similarity among individual reaches, rather than placing reaches into discrete categories.

Ordination methods, such as principal component analysis (PCA) and multidimensional scaling, identify the major sources of variation in multivariate data and attempt to project such data onto a small number of ordination axes that will be simple to interpret and the differences among sites readily quantified (Hastie et al., 2001). In the majority of cases, two axes are used so that the data can be readily plotted; indeed plotting is often the main aim of ordination (Ter Braak, 1995). If based on controls on channel morphology, ordination scores could underpin simple, yet powerful predictions of channel character (Jeffers, 1998). For rivers, ordinations are related to the powerful two-dimensional plots that are common in fluvial geomorphology, such as slope and catchment area (see Montgomery and Buffington, 1997) or slope and discharge (see Church, 2002), but they allow a greater number of variables to be considered simultaneously and the main sources of inter-reach differences from amongst those variables to be identified. Compared to classification approaches, few riverine studies have considered ordination in spite of the potential advantages (e.g., Harvey et al., 2008a, b; Gurnell et al., 2012), notably real promise in terms of simplicity and efficacy at separating channel characteristics, such as predominant substratum (e.g., Jeffers, 1998).

In the current study, we ordinate rivers from across England and Wales using GIS-based variables and use the ordination axes to predict river channel form and a range of hydromorphic features, quantifying their predictive ability in detail. We base this on the recent River Habitat Survey (RHS) baseline of England and Wales, which involved a stratified random sample of nearly 5000 river reaches (Seager et al., 2012). Whilst there are limitations to the geomorphic information recorded by RHS (Newson, 2002), they are offset by the wide range of information recorded and the ability to carry out large-scale ($>100{,}000$ km^2) studies of river morphology (e.g., Newson et al., 1998a; Emery et al., 2004; Harvey et al., 2008b). Our primary aims were to (i) use ordination to identify the principal sources of variation amongst English and Welsh rivers from a range of potential controls on channel form; (ii) quantify the predictive ability of the ordination scores across a wide range of hydromorphic 'features'; and (iii) compare the predictive ability of the ordination to established bivariate axes (e.g., slope and discharge) and an earlier ordination of British rivers (Jeffers, 1998). We also illustrate how simple ad hoc comparisons of river reaches can be made using ordination techniques without recourse to formal classifications.

2. Methods

2.1. RHS baseline survey

The RHS was developed by the National Rivers Authority (now the Environment Agency) in the mid-1990s as a simple way of characterising the physical structure of rivers (Raven et al., 1997). The survey involves a description of a 500-m length of channel, with channel and bank features recorded at 10 equidistant locations known as 'spot-checks' and general site characteristics (e.g., valley cross section, land use within 50 m of the channel) recorded along the complete 500-m length in a second 'sweep-up' stage of the survey (Environment Agency, 2003). The survey records predominant bed and bank materials; flow biotopes; the presence of features such as bars, riffles and pools; bank profiles; artificial features and modifications to the channel; and vegetation in the channel, on the banks and in the riparian zone (Raven et al., 1997; Environment Agency, 2003). All of the feature categories have clear definitions, and surveyors are trained to ensure consistency of recording (Environment Agency, 2003). In 2007 and 2008, a second RHS baseline survey of England and Wales was completed, with 4849 reaches surveyed under a stratified random survey design (Seager et al., 2012). Three randomly located sites were surveyed in every 10-km national grid square, with two sites on rivers present on 1:250,000 scale maps and the third site on streams only present on the 1:50,000 scale map, providing better coverage of low order streams (Seager et al., 2012).

2.2. GIS-derived variables

Potential predictive variables were chosen to reflect some of the key drivers of channel form, whilst restricting the list to variables that could be derived easily from a GIS (e.g., Newson et al., 1998a; Reinfelds et al., 2004; Harvey et al., 2008b; Orr et al., 2008). Previous work has shown that a number of catchment and reach-scale variables can successfully predict aspects of hydrology and morphology (e.g., NERC, 1975; Jeffers, 1998; Institute of Hydrology, 1999). From a geomorphological perspective, the form of alluvial channels reflects prevailing flow and sediment transport conditions, which are functions of regional climate, topography, underlying geology, and land use. For example, upland channels are better connected to hillslope sediment supplies than their lowland counterparts, but experience more varied flow conditions as they respond rapidly to individual precipitation events and lack significant base flow that moderates discharge between individual storms. In our analysis, variables were identified to capture measures of topography (e.g., channel gradient and channel confinement), regional climate (e.g., annual precipitation), local flow dynamics and underlying geology (Table 1). Data were derived from 4080 RHS sites within catchments of >15 ha (Fig. 1), which represented 84% of the 2007–2008 RHS baseline survey. Sites were excluded where it was not possible to derive all of the catchment level variables — these were mainly in low-lying, artificially drained areas (e.g., in eastern England), usually reflecting a modified (nondendritic) drainage network and/or low relief, making it difficult to determine the source or delineate the catchment area above a reach. All data manipulations were carried out in ArcInfo (v. 10; ESRI, Redlands, USA), and data summaries within catchment polygons were calculated with the geospatial modelling environment (v. 0.5; Beyer, 2011).

The baseflow index (BFI; Table 1), which represents the long-term ratio between baseflow and total discharge, was modelled using linear regression with catchment geology as predictors (Bloomfield et al., 2009). The model was calibrated using data from 837 English and Welsh gauging stations included in the UK hydrometric register (Marsh and Hannaford, 2008) for which we calculated catchment areas using the same method as for RHS sites (Table 1). Three variables represented the percentage of the catchment with drift geology divided into high, mixed, and low permeability, following the category definitions in Marsh and Hannaford (2008); and a further six variables described the percentage of the catchment underlain by the six main aquifers in the UK (Chalk, Lower Greensand, Jurassic limestone, Permo-Triassic sandstones, Magnesian limestone, and Carboniferous limestone; see Allen et al., 1997). Drift and solid geology were extracted from 1:625,000 British Geological Survey maps. All variables were significant at $p < 0.001$, except low permeability drift geology ($p = 0.02$), and the adjusted R^2 was 0.58. The model was then used to predict BFI values for all RHS sites.

The rate of potential energy expenditure per unit channel length, or stream power, has been used by many studies in stream classification (e.g., Nanson and Croke, 1992; van den Berg, 1995; Bizzi and Lerner, 2012) and in assessing the behaviour of streams (e.g., Whipple and Tucker, 1999). When normalised by channel width, this variable is denoted as specific stream power (ω), which has been used widely in studies of sediment transport and controls of channel morphology (e.g., Newson et al., 1998a; Knighton, 1999; Reinfelds et al., 2004; Parker et al., 2011b). Most previous studies have focused on ω during near-bankfull flows, such as the mean annual flood (van den Berg, 1995; Knighton, 1999). Following Gurnell et al. (2010) and Bizzi and Lerner (2012), we used the median annual flood (Q_{MED}; m^3 s^{-1}) to calculate ω given that it approximates bankfull discharge (Wharton, 1995) but can be more reliably estimated from short time series as it is less