Geomorphology 253 (2016) 38–47

Contents lists available at ScienceDirect

Geomorphology

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

Automatic procedures for river reach delineation: Univariate and multivariate approaches in a fluvial context

V. Martínez-Fernández *, J. Solana-Gutiérrez, M. González del Tánago, D. García de Jalón

E.T.S. Ingeniería de Montes, Forestal y del Medio Natural, Universidad Politécnica de Madrid, Spain

ARTICLE INFO

Article history:
Received 10 July 2015
Received in revised form 29 September 2015
Accepted 30 September 2015
Available online 3 October 2015

Keywords:
River segmentation
River reach
Permutation method
GIS
Randomization

ABSTRACT

Segmenting the continuum of rivers into homogeneous reaches is an important issue in river research and management. Automatic procedures provide significance, objectivity, and repeatability. Although univariate techniques are frequently used to identify river reaches, multivariate approaches offer a more integrative context. Three nonparametric methods (multi-response permutation procedures (MRPP) with an advance in the significance level estimation, the Pettitt and Mann–Kendall tests) are applied for segmenting the river based on three geomorphic variables (valley width, active channel width, and channel slope) systematically measured in a GIS environment. The cited techniques have been applied to the Curueño River (NW Spain) to illustrate the methods, we analyse reach distribution along the river longitudinal profile.

The methods successfully characterize the evident transitions along fluvial systems and also others less noticeable. The three methods provide more reaches according to valley width and less reaches according to channel slope (18.0 and 3.7 reaches on average, respectively). In contrast to the Mann–Kendall test, MRPP and Pettitt tests provide more stable segmentations when significance level varies. However, the Pettitt test provides irregular segmentations for regular patterns. The MRPP both univariate and multivariate applications enables a wider scope for the segmentation issue, which is useful in diverse aspects of fluvial domain.

1. Introduction

Dividing a river into homogeneous reaches is a determinant aspect of river research and management (Brenden et al., 2008), as natural system studies require manageable reference scales to assess relationships between forms and processes (Parker et al., 2012). Also, considering different reaches along the river improves the interpretation of geomorphologic processes and forms, which is essential for developing sustainable restoration approaches and decision making (Gurnell et al., 2014).

The expert criteria approach and graphical methods for delineating river reaches have been widely used for many purposes (Frissell et al., 1986; Schumm et al., 1994; Habersack, 2000; Fausch et al., 2002; Brierley and Fryirs, 2005; Thorp et al., 2006; Beechie et al., 2010; Merovich et al., 2013; McCluney et al., 2014). This approach offers advantages such as the consideration of existing knowledge about the relationships between biological and geomorphic characteristics (Brenden et al., 2008). However, in the resulting segmentation the repeatability of the division is difficult because of its subjective basis, and it may not be statistically significant being that the degree of assurance is unknown (Bizzi and Lerner, 2012).

An alternative to the expert-criteria approach is to use automatic techniques to find statistically significant discontinuities for segmenting rivers. The primary challenge associated with these methods is detecting boundaries for the system characteristics, which is advantageous because of the efficiency, repeatability, and objectivity as opposed to an expert opinion approach (Alber and Piégay, 2011). Recently, new research perspectives have been opened to detect longitudinal discontinuities along the fluvial system. Several authors have detected discontinuities within the river in a univariate way, in relation to valley or channel width (Alber and Piégay, 2011), predicted sediment transport (Parker et al., 2012) or floodplain width (Notebaert and Piégay, 2013). Leviandier et al. (2012) suggested that segmenting a river based on only one variable could be useful as a first step in the geomorphic characterization, but depending on the purpose, this variable should be sufficiently integrative to provide the geomorphic context of the river. Then multivariate approaches that consider some drivers on fluvial processes could provide a greater potential for a better comprehension of the system (Bizzi and Lerner, 2012). Few multivariate approaches have been considered in river research and management. Brenden et al. (2008) proposed a spatially constrained clustering programme for river valley segmentation and applied it to seven physicochemical attributes determinant of fish distribution. Each reach between confluences in the network was taken as the sampling unit; therefore with this method detecting discontinuities between confluences is not possible. Bizzi and Lerner (2012) applied neural network

* Corresponding author at: E.T.S. Ingeniería de Montes, Forestal y del Medio Natural, Universidad Politécnica de Madrid, Ciudad Universitaria, s/n, 28040 Madrid, Spain.
E-mail address: vanesa.mfernandez@upm.es (V. Martínez-Fernández).

http://dx.doi.org/10.1016/j.geomorph.2015.09.029

models (Kohonen, 1982) over geomorphic drivers, previously inferred as grid cell resolution. They generated a nonlinear classification of visually interpretable clusters, followed by cluster techniques (Vesanto et al., 2000) to identify different channel types.

Thus, the automatic detection of homogeneous reaches along rivers is by now a challenging perspective. These procedures have a great potential to improve the delineation of river reaches in a quantitative way instead of a descriptive way, more objectively, and less susceptible to operator subjectivity. In particular, multivariate approaches have been less developed in fluvial geomorphology from a scientific point of view. In this regard, multi-response permutation procedures (Mielke, 1991) could be an alternative for segmenting the river, generating reaches that are internally homogeneous and significantly different from the adjacent reaches. Its nonparametric condition together with its univariate and multivariate application offer advantages in a fluvial context, which is a domain commonly lacking the numerous assumptions required by parametric techniques and frequently multivariate. It was successfully applied by Orlowski et al. (1995) to classify the lower Mississippi River into geomorphically distinct reaches, evaluating the significance of a previous classification proposed by Schumm et al. (1994). The method was effectively applied over geomorphic variables taken from the 1880 and 1915 hydrographic surveys, which were measured at not necessarily regular intervals.

The recent advances in geographic information system technology (GIS), analysis tools, and data mining techniques and the current availability of medium to high resolution digital data (digital elevation models, DEM, with a 5-m or less horizontal grid and a vertical accuracy of 0.5 m, are already available in many countries) offer new opportunities to explore the potentiality of statistical techniques, particularly multi-response permutation procedures, in assessing objectively geomorphic properties at different scales, which should be exploited by river managers and researchers.

Our research aims to compare different automatic procedures for segmenting a river into homogeneous units based on geomorphic variables, considering univariate and multivariate algorithms and discussing their advantages and limitations. To achieve this purpose, first we describe the extraction procedure of geomorphic variables (valley width, active channel width and channel slope). Secondly, we introduce the statistical methods: (i) multi-response permutation procedures (Mielke, 1991), with an advance in the significance level approximation using a randomization test (Manly, 1997); (ii) the Pettitt test (Pettitt, 1979), a univariate method previously used in the fluvial context (Zhang et al., 2008; Alber and Piégay, 2011); and (iii) the Mann–Kendall test (Mann, 1945; Kendall, 1975), which is also univariate and is less frequently used in spatial context (Alibert et al., 2011). Finally, we present their applications over a gravel-bed river in NW Spain (the Curueño River). What is learned from the procedure for extracting the variables in a GIS environment and the application of the methods in the fluvial domain is discussed and possible future applications are addressed.

2. Materials and methods

The methodology is structured in two phases: first, the data production process based on the extraction of variables in a GIS environment and, second, the application of the three methods considered for segmentation.

To explore the performance of algorithms the Curueño River (42°51′ N, 5°24′ W), in NW Spain, was chosen (Fig. 1A). The Curueño is a gravel-bed river that rises in the Cantabrian Mountains (NW Spain) and flows over 45 km south to enter the Porma River with a mean slope of 0.96%. This river drains ~293 km^2, and the elevation ranges between 850 and 2150 m above sea level. Annual precipitation in the catchment ranges between 800 and 1300 mm (http://aemet.es), presenting the river with a mean annual discharge of 5.53 m^3/s under a perennial flashy hydrological regime..

2.1. Data production

To test the statistical techniques we selected three geomorphic variables related to fluvial processes commonly used for spatial analysis of stream networks (Kondolf et al., 2003; Fryirs and Brierley, 2013): the valley bottom width, the channel slope, and the active channel width (combining unvegetated bars and low-flow channel width).

First, the required geographic elements for extracting these variables were delineated: the valley bottom area, the streamline, and the active channel area. They were digitized manually combining information from a DEM with a 5-m spatial resolution, dated 2010 (www.ign.es), and orthophotographs with a 0.25–0.5 m spatial resolution, dated 2011, following the recommendations by Alber and Piégay (2011) and Gurnell (1997). The GIS analysis was supported by ESRI ArcMap version 9.3, with the ArcHydrotools, 3DAnalysis, Spatial Analyst, and Xtoolspro extensions.

Secondly, the reference axis of each element was required for locating each measurement along the element. The streamline was considered as the reference axis for channel variables. The axis of the valley bottom was defined by using a semiautomatic procedure based on Thiessen polygonalization, which extracts the skeleton of every polygon and ramified polygon (Alber and Piégay, 2011).

Finally, the measurements of the variables were undertaken. Valley bottom widths were systematically measured orthogonally to their reference axis every 200 m (Fig. 1C). Average active channel width was calculated as the active channel area divided by the streamline length (Manners et al., 2014) within each patch encompassed by valley width lines. Channel slope was calculated in each patch by dividing the difference between the upstream and downstream elevation values by the streamline length within the patch (Fig. 1C).

2.2. Methods for segmentation

Three nonparametric methods are investigated in this paper: (i) multi-response permutation procedures, which can be applied as a univariate or multivariate technique, (ii) the Pettitt test, and (iii) the Mann–Kendall test, with the latter two being univariante. The methods used in this paper are nonparametric techniques to avoid the numerous assumptions required by parametric techniques, which are almost never obtained from observations of natural systems.

2.2.1. Multi-response permutation procedures

Multi-response permutation procedures are nonparametric techniques that allow the system to be classified into homogeneous and significant groups and the multivariate dimension of the river morphology to be taken into consideration. A complete mathematical description is contained in the work by Mielke (1991) and Orlowski et al. (1993).

The MRPP evaluates the uniqueness of previously defined groups relative to all other possible permutations of the objects. The null hypothesis states that equal probabilities are assigned to each of the possible allocations of the objects into the groups. Hypothesis testing is based upon the MRPP statistic that quantifies the separation between groups by considering the objects in a Euclidian data space. It is calculated as the weighted average of the within-group between-point Euclidean distance average, indicating in the case of small values a tendency for clustering.

To estimate the significance level of the partition, we evaluate the *p*-value defined as the proportion of MRPP statistic values, calculated for all possible partitions, that are less than or equal to the observed MRPP statistic.

Generally, possible partitions or combinations are so many that the procedure is extremely time-intensive (Mielke, 1991) and the need for approximation procedures becomes essential. Mielke (1991) proposed using a Pearson type III distribution for *p*-value approximation. In our case, we suggest using a randomization test (Manly, 1997) to evaluate the candidate partition *p*-value. A randomization test is a