AGU PUBLICATIONS

Water Resources Research

RESEARCH ARTICLE

10.1002/2017WR020390

Mapping the temporary and perennial character of whole river networks

A. M. González-Ferreras[1] and **J. Barquín**[1]

[1]Environmental Hydraulics Institute, Universidad de Cantabria-Avda. Isabel Torres, 15, Parque Científico y Tecnológico de Cantabria, Santander, Spain

Key Points:

- Mapping flow permanence facilitates a better understanding of the interactions between terrestrial and aquatic systems at large scales
- The temporary/perennial character of river segments for a whole river network is usually not available, incomplete, or not very precise
- We present an easy methodology for estimating the occurrence and extent of temporary and perennial river segments in a whole river network

Supporting Information:

- Supporting Information S1

Correspondence to:

A. M. González-Ferreras,
gferrerasam@unican.es

Citation:

González-Ferreras, A. M., and J. Barquín (2017), Mapping the temporary and perennial character of whole river networks, *Water Resour. Res.*, *53*, 6709–6724, doi:10.1002/2017WR020390.

Received 9 JAN 2017

Accepted 13 JUL 2017

Abstract Knowledge of the spatial distribution of temporary and perennial river channels in a whole catchment is important for effective integrated basin management and river biodiversity conservation. However, this information is usually not available or is incomplete. In this study, we present a statistically based methodology to classify river segments from a whole river network (Deva-Cares catchment, Northern Spain) as temporary or perennial. This method is based on an a priori classification of a subset of river segments as temporary or perennial, using field surveys and aerial images, and then running Random Forest models to predict classification membership for the rest of the river network. The independent variables and the river network were derived following a computer-based geospatial simulation of riverine landscapes. The model results show high values of overall accuracy, sensitivity, and specificity for the evaluation of the fitted model to the training and testing data set ($\geq$0.9). The most important independent variables were catchment area, area occupied by broadleaf forest, minimum monthly precipitation in August, and average catchment elevation. The final map shows 7525 temporary river segments (1012.5 km) and 3731 perennial river segments (662.5 km). A subsequent validation of the mapping results using River Habitat Survey data and expert knowledge supported the validity of the proposed maps. We conclude that the proposed methodology is a valid method for mapping the limits of flow permanence that could substantially increase our understanding of the spatial links between terrestrial and aquatic interfaces, improving the research, management, and conservation of river biodiversity and functioning.

1. Introduction

Streams may be classified as temporary or perennial according to the permanence of their surface flow. Temporary streams are waterways that cease flowing at some point in space and time along their course [*Acuña et al.*, 2014]. Flow cessation may be caused by transmission loss, evapotranspiration, downward shifts in groundwater tables, hillslope runoff recession, or freeze-up [*Larned et al.*, 2010] and is part of the natural hydrology for streams and rivers globally [*Acuña et al.*, 2014]. Temporary streams and rivers have been defined using several terms (i.e., interrupted, intermittent, temporary, ephemeral, episodic, seasonal) according to different classifications in terms of flow, drying, and periodicity [e.g., *Uys and O'Keeffe*, 1997]. For simplicity, in this paper we refer to all of these types of systems as temporary.

The number of studies focusing on temporary streams and rivers has increased exponentially since the 1990s [*Datry et al.*, 2011], and there is a growing scientific interest in the ecology of temporary waterways due to their role in the water and carbon cycles. These river segments contain important links between water stored in soils, aquifers, snowpacks, glaciers, and the atmosphere, and they are also important for the provision of a wide range of ecosystem services [*Larned et al.*, 2010]. Temporary streams are not only naturally widespread in dry climate areas, but they also comprise many of the first-order streams in most drainages in wetter climates [*Nikolaidis et al.*, 2013], accounting for a significant proportion of the total number, length, and discharge volume of the world's rivers [*Larned et al.*, 2010; *Tooth*, 2000]. Moreover, in the coming years, the number and length of temporary river segments and the duration and magnitude of temporary flows may increase in areas that experience drying trends due to climate change, land use alteration, and water abstraction [*Datry et al.*, 2014], which could have important consequences for river biodiversity and functioning at a catchment scale. For example, the simplification of river networks and the alteration of water fluxes have been shown to reduce the capacity of fluvial systems to recover from natural disturbances [*Sabater and Tockner*, 2010]. The loss of perennial streams and rivers or the reduction of their lengths has large social, economic, and ecological consequences, so managers seek better ways to track and monitor the status of these systems [*Turner and Richter*, 2011].

The dynamic characteristics of temporary rivers present significant challenges for the assessment of ecological conditions and potentially affect the accuracy of monitoring results [*Arthington et al*., 2014]. These challenges have not been properly addressed in some legislation, such as the Water Framework Directive [*European Commission*, 2000; *Nikolaidis et al*., 2013]. Standard methods for monitoring perennial and temporary streams typically collect measurements from too few locations and do not effectively characterize the spatial extent of these terrestrial-aquatic systems.

Although a global inventory of temporary streams has not yet been compiled, several estimates exist and collectively underscore their abundance (see *McDonough et al*. [2011] for more information). For specific areas, the methods most used to map the spatial distribution of temporary streams are topographic maps, aerial images, and field surveys [e.g., *Robinson et al*., 2016]. However, these methods are intrinsically labor-intensive and subjective and are not generally applicable for mapping over large areas. In recent years, these methods have been combined with various modeling techniques [e.g., *Sando and Blasch*, 2015] to objectively estimate temporary streams and rivers for whole river networks. Most of the previous studies that use statistical modeling to estimate temporary and perennial rivers have been performed at large regional scales where many flow gauging stations exist. These approaches are based on long flow gauging records that allow the definition of different types of temporary and perennial river flows, which are then predicted to the whole region based on specific catchment attributes [e.g., *Shortridge et al*., 2016; *Snelder et al*., 2013]. However, obtaining comprehensive flow gauging records (i.e., from a sufficient number of gauges and/or time series lengths) can be very difficult or even impossible for most catchments, where flow gauges are very sparse and flow records for temporary rivers are underrepresented. This situation calls for the development of other types of approaches and data sets that facilitate the accurate estimation of the spatial distribution of temporary and perennial rivers.

Recognizing the limitations of the current digital resources and maps in representing channel extent and the degree of flow permanence (e.g., the difficulty of differentiating between temporary and perennial rivers in some forested areas using remote sensing) is important for many reasons. For example, temporary streams and rivers represent a dominant interface between terrestrial and freshwater ecosystems, and their hydrology is a critical factor influencing patterns and processes in river networks. Moreover, accurate hydrography is a key tool for monitoring, modeling, and decision making [*Fritz et al*., 2013]. In this regard, the extent of the temporary and perennial segments in a whole river network is a basic information need for formulating appropriate strategies for biodiversity conservation. As an example, there are some recent initiatives to compile databases that integrate all the available information on the environmental characteristics and biodiversity of temporary rivers (IRBAS database) [*Leigh et al*., 2017]. For some organisms (e.g., fish), the available habitat is determined by the connectivity and extent of the perennial network. For other organisms (e.g., amphibians), the connectivity and extent of the temporary network is crucial for their survival. Temporary streams may, for example, serve as important amphibian nursery areas, because they support fewer predators than perennial streams [*Reid and Ziemer*, 1994]. For conservation strategies and planning, it is therefore important to know the length of both perennial and temporary networks and the delineation of the boundaries between them.

Accordingly, the main goals of this paper are to develop a mapping strategy to: (1) estimate the occurrence and extent of perennial and temporary segments in a whole river network at a local catchment scale and (2) determine the main variables which play a fundamental role in determining their spatial distribution.

2. Materials and Methods

2.1. Study Area

The study area entails the Deva-Cares catchment, located in Northern Spain and covering an area of 1200 km^2, which drains into the Cantabric Sea (Figure 1). A large part of the catchment is located within the Picos de Europa National Park, which is part of the Cordillera Cantabrica mountain range. The average altitude of the catchment is 1100 m and the mean slope is 50.36% [*Grupo de Emisarios Submarinos e Hidraulica Ambiental (GESHA)*, 2005]. The study area is geologically diverse (including sandstone, dolomite, and marl) but is mainly dominated by limestone karst formations and by conglomerates and slates [*IGME*, 2015]. The

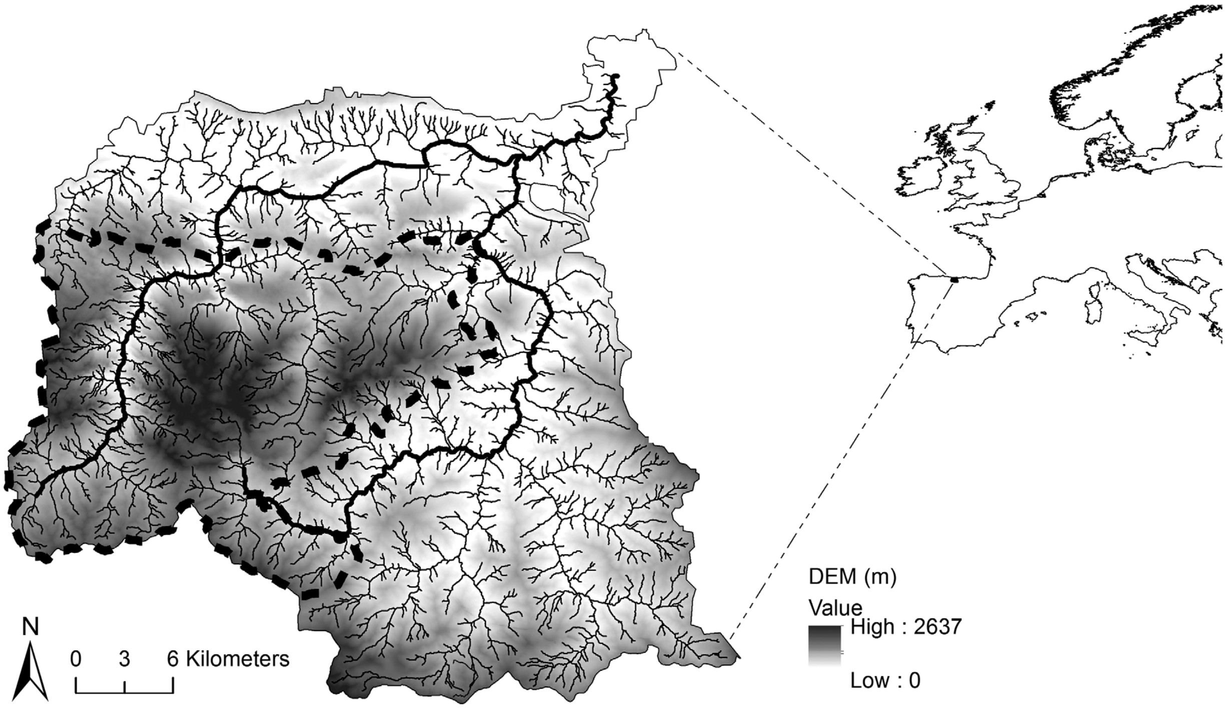

Figure 1. Map of the study area and representation of the Deva-Cares river network. Dashed line represents the limits of Picos de Europa National Park and black lines represent the two main axes of the river network, Deva and Cares rivers.

hydrogeology and geomorphology of this catchment is highly influenced by the presence of the karst massif of Picos de Europa. This karst contains 13% of the world's shafts known to be deeper than 1.000 m [*Ballesteros et al*., 2011] and some 3648 documented cavities encompassing 355 km of conduits [*Ballesteros et al*., 2015]. At altitudes above 1700 m there is practically no vegetation, and evapotranspiration is below 200 mm. These upper karst zones experience very quick infiltration from snow and rainfall, making this zone one of the main karst aquifer recharge areas [*Fernández-Giber et al*., 2000]. Subsequently, very huge vadose zones develop (well over 1000 m in some areas) before reaching the water table which is very close to the fluvial valleys. The main aquifer discharge occurs through the many springs on the valley bottoms of the main tributaries and even through the river beds, although there are also springs with significantly lower flows in the upper parts of the karst discharging small perched aquifers [*Adrados et al*., 2012; *Fernández-Giber et al*., 2000]. The rivers are characterized by high channel gradients and short channel lengths, with canyons up to 2000 m deep which evidence the significant fluvial incision in the area [*Ballesteros et al*., 2011]. The southeastern part of the catchment, which includes the two main Deva tributaries (Figure 1) and the uppermost part of the Cares tributary catchment, has a different hydrological character dominated by surface runoff flows imposed by old mature forests over shales and conglomerates.

The climate is temperate [*Rivas-Martínez et al*., 2004], but the climatic conditions are highly variable seasonally and spatially, mainly driven by two factors: its proximity to the sea (less than 50 km to the coast in a straight line) and orographic effects (high mountain tops, rising up to 2600 m above sea level). Snow is common during the winter and accounts for nearly 20% of the annual precipitation, which exceeds 2000 mm above 1000 m [*Fernández-Giber et al*., 2000]. The study area is located primarily in the Eurosiberian biogeographic region, but also spans the Mediterranean region, which is expected to experience significant hydrological impacts because of climate change [e.g., *Sánchez de Dios et al*., 2009]. The natural deciduous forest below 400 m is dominated by *Fraxinus excelsior*, *Tilia* sp., *Corylus avellana*, *Acer* spp., and *Quercus* spp. while *Populus* spp. *Quercus robur*, *Quercus petraea*, *Fagus sylvatica*, and *Ilex aquifolium* dominate between 400 and 1100 m. From 1100 to 1800 m *Betula* sp. is dominant, while alpine mountain grasslands and denuded rocks are the dominant features at higher altitudes. The Mediterranean influence produces a native vegetation community dominated by holm *Quercus ilex* and *Quercus pyrenaica*.

2.2. Mapping Approach

The mapping approach presented in this study comprises four steps. First, we created a Virtual Watershed with the aim of obtaining a river network (a digital representation of the surface water drainage network) that incorporates all the environmental information needed to generate the independent variables. Second, within this digital platform, we included information about the a priori classification of a subset of river segments (reaches of the river network) as being either temporary or perennial. This information was gathered from specially designed field surveys and aerial image data. Third, we selected several independent variables which are significant for determining the perennial/temporary character of a river segment. Fourth, we used Random Forest models to predict the temporary/perennial character of those river segments for which there was no empirical information (i.e., no field surveys or aerial image data). The final map integrates the empirical observations with the modeled ones. All these different steps are described in detail below.

2.2.1. Virtual Watershed Approach

For this study, a Virtual Watershed was built using the Bldgrds and Netrace software packages which are contained in the "NetMap" platform (www.terrainworks.org) [*Benda et al*., 2015; *Miller*, 2002]. Virtual watersheds are computer-based geospatial simulations of riverine landscapes that include digital elevation models (DEM), synthetic hydrography, and their coupling, using a data structure to support the required analytical capabilities (for more information, see *Barquin et al*. [2015] and *Benda et al*. [2015]). The river network (see Figure 1) was delineated using flow directions inferred from a 25 m DEM. To estimate the location of channel heads, we employed two criteria, one for low-gradient areas and the other applied to high-gradient areas. In the first case, channel expansion occurs primarily through fluvial processes and in the second case, channel expansion may occur via mass wasting processes. Both cases employ a slope-dependent drainage area threshold [*Dietrich et al*., 1993; *Montgomery and Dietrich*, 1992] following the equation $a_{cr}S^{\alpha} = C$, where a_{cr} is a critical specific drainage area required for channel initiation, S is the surface gradient, α is an exponent (which varies between 1 and 2), and C is a constant. The values used were $a_{cr} = 40\ m^2$ (for low-gradient areas) and $300\ m^2$ (for high-gradient areas), $\alpha = 2$, $S = 0.2$ (low gradient threshold), and 0.35 (high-gradient threshold). Values of S separate channel initiation into two process domains; mass wasting and fluvial erosion of surface material. In addition to drainage-area-dependent thresholds, we required a minimum topographic convergence at the channel heads, indicated by plan-curvature values of 0.00025 or greater in low-gradient areas, and 0.01 or greater in high-gradient areas over a minimum flow length of 40 m. Physically, the C value reflects regional properties of the soil, bedrock, and climate. To set threshold values that reproduce appropriate channel densities, we followed the process described in *Miller* [2002] and in previous studies in the region [e.g., *Benda et al*., 2011]. The final river network comprised 11,256 river segments and set the spatial network for the integration of all the following information.

2.2.2. Dependent Variables

One important aspect within this study is the definition of what we considered to be temporary river segments. This definition is completely constrained by the lack of flow gauging records for the temporary river segments in the studied catchment. This lack of data prevents a clear-cut definition of the different types of temporary river segments from being made based on the frequency and duration of zero flows [e.g., *Snelder et al*., 2013]. Instead, in this study we defined temporary river segments as those with zero flow (i.e., cessation of surface water flow, although water may be present but only in disconnected pools) during the summer (the low-flow season) in average hydrological years (see below). In contrast, we define perennial river segments to be those with perennial flow in average hydrological years. Information about the location of temporary and perennial river segments was collected from different sources and then integrated into the river network following a series of steps:

1. *Field data collection*. A specific field survey campaign was designed to map temporary river segments during the 2011 and 2014 summer seasons. We visited 74 and 75 different river segments, respectively (a total of 149 independent observations) over the two seasons. We then classified each segment as being temporary or perennial according to the existence of surface flow. The summer of 2011 was considered to be a normal year according to the standardized precipitation index (SPI) developed by the Spanish Meteorological Office in its hydro-meteorological annual report [*Spanish National Meteorological Agency (AEMET)*, 2016] (http://www.aemet.es/es/serviciosclimaticos) while the summer of 2014 was regarded as moderately wet. Thus, we consider our mapping exercise to be a composite image of the low-flow season in average hydrological years.

2. *Aerial image data*. To complement the obtained field data, we visually identified temporary river segments within the rocky highlands of the Picos de Europa Karst's central massif (upstream of our surveyed tributaries and in nonsurveyed ones) using aerial images. The streams in this area do not flow during the low-flow season because of the numerous karstic ducts and the low amount of water retained in the upper catchment (see the study area description). In this area, it is relatively easy to visually identify flowing water as there is hardly any vegetation. We also complemented the field data by looking for perennial river segments in the lower parts of the catchment. In these areas, the width of the river's channels is greater and the bank vegetation does not cover the segments entirely, allowing uninterrupted water flows to be identified. Aerial imagery was obtained using the Spanish National Geographic Institute's (IGN) web service of maps (WMS), and aerial images from 2011 and 2014 (with resolutions of 0.25 and 0.5 m depending on the flight) from the PNOA (National Plan of Aerial Orthophotography) project (http://www.ign.es/wms/pnoa-historico). PNOA aims to obtain digital aerial orthophotos of the entire Spanish territory, with an update period of 2–3 years and performing flights during the spring to summer period.
3. *Integration of data*. This is a crucial step in our methodology, as important assumptions have been made that need to be carefully considered when applying this approach to other catchments. First of all, river segments that were identified as perennial or temporary were located and labeled in the digital river network. Then we made two important assumptions: (1) we extended the temporary network upstream by assuming that all river segments upstream of a long temporary segment ($>$500 m) would also be temporary and (2) we extended the perennial network downstream by assuming that all river segments downstream of a perennial segment would also be perennial. These assumptions were made on the basis of the main hydro-geological functioning of this catchment as described in the study area section. Most of the tributaries in the upper parts of the karst only have water during the snowmelt season or during heavy rainfall, while most river segments in the lower valley have perennial flow maintained by a wide network of perennial springs that discharge the karst aquifer [*Adrados et al*., 2012; *Fernández-Giber et al*., 2000]. Finally, two other limitations were also imposed before the modeling stages. First, this study does not include very small perennial river channels that are maintained by small spring sources in the upper part of the karst and which run dry before entering a perennial flow channel (most of them run for less than 50 m [*Adrados et al*., 2012]). Second, this study also does not include perennial segments that run for more than 0.5 km and then go dry, as there is only one such case recorded in the catchment, associated with a karstic sink (Liordes Polje) according to *Adrados et al*. [2012].

2.2.3. Independent Variables

We selected a range of independent variables (Table 1) describing several environmental attributes that could possibly be important for determining the perennial/temporary character of the river segments in the studied area. These include topography ($n = 5$), climate ($n = 6$), land cover ($n = 2$), and geology ($n = 4$). The assignment of stream attributes to individual segments of the river network was performed using NetMap tools, and the digital information (topography, climate, land cover, and geology) was summarized across a range of spatial scales, from entire catchments (drainage areas for each river segment), to adjacent hillslopes draining into individual river segments (drainage wings for each segment, referred to as segment wings).

In other studies, the different variables used as model predictors (Table 1) have been derived from a wide range of digital resources. For brevity, we refer only to those works in which specific details are given on the methodological aspects used to derive them (for more information about the description of the independent variables, see *Álvarez-Cabria et al*. [2016, 2017], *González-Ferreras et al*. [2016], and *Peñas et al*. [2014]).

Although multicollinearity has no influence on the predictive performance of the model we used in this study (see below), variable importance measurements can be affected [*Boulesteix et al*., 2012] and the partial plots representation (see below) of the predictor-response relationship is more reliable when the predictors have low correlation [*Friedman and Meulman*, 2003]. For this reason, and to avoid potential problems, we developed a correlation matrix (Spearman rank correlation) for the segments characterized in the previous phase, and when pairs of variables had a correlation $>|0.7|$ only one was retained for modeling (see Table 1 and supporting information Figure S1). The variables retained were those that most reduced the total number of variables in the model (i.e., those that were correlated with a larger number of other predictors).

Table 1. Initial Set of Independent Variables Attributed to the River Network[a]

Type	Code	Definition	Units	Correlated With:
Topographic	**AREA_SQKM**	**Total catchment area**	**km^2**	DRAIN_DEN
	MN_ELEV	**Average catchment elevation from the considered river segment to the upper most river segment in the river network**	**m**	MN_DEN, MN_EP
	MN_GRAD	**Average catchment gradient from the considered river segment to the upper most river segment in the river network**	**%**	
	VAL_FLOOR	**Width of the valley floor at 2 × bankfull depth elevations above the channel**	**m**	
	DRAIN_DEN	Drainage density. Number of segments divided by the catchment area	N° of rivers confluences by catchment area	**AREA_SQKM**
Climatic	**MN_TEMP**	**Mean annual catchment temperature**	**°C**	MN_maxT08
	MN_PREC	**Mean annual catchment precipitation**	**mm**	
	MN_EP	Mean annual catchment potential evapotranspiration	mm	**MN_ELEV, MN_maxE08**
	MN_minP08	**Accumulated value (average variable value from the consider river segment to the upper most river segment in the river network) for this variable: minimum value within the monthly list (1980–2006) of mean precipitation values in August**	**mm**	
	MN_maxE08	**Accumulated value (average variable value from the consider river segment to the upper most river segment in the river network) for this variable: maximum value within the monthly list (1980–2006) of mean potential evapotranspiration values in August**	**mm**	MN_EP
	MN_maxT08	Accumulated value (average variable value from the consider river segment to the upper most river segment in the river network) for this variable: maximum value within the monthly list (1980–2006) of mean temperature values in August	°C	**MN_TEMP**
Land cover	**MN_BLF**	**Area occupied by broadleaf forest from the considered segment to the most upper catchment point in the catchment**	**%**	
	MN_DEN	Area occupied by denuded areas from the considered segment to the most upper catchment point in the catchment	%	**MN_ELEV**
Geological	**LC_HARD**	**Average rock hardness within the segment wings**	**1–5**	
	MN_HARD	**Average rock hardness from the considered segment to the most upper catchment point in the catchment**	**1–5**	
	MN_COND	Average rock conductivity from the considered segment to the most upper catchment point in the catchment	1–5	**MN_PERM**
	MN_PERM	**Average rock permeability from the considered segment to the most upper catchment point in the catchment**	**1–5**	MN_COND

[a] Bold variables are uncorrelated variables (Spearman rank correlation $\leq|0.7|$) comprised in the final set of independent variables (supporting information Figure S1 shows the Spearman rank correlation matrix).

2.2.4. Modeling

We decided to use the Random Forest (RF) classification model [*Breiman*, 2001] with two classes (temporary and perennial). RF is a nonparametric method developed by *Breiman* [2001] that comprises an ensemble of individual Classification and Regression Trees (CART) [*Breiman et al.*, 1984] based on the aggregation of a large number of decision trees (a forest) from which a final prediction is averaged for all trees. RF presents a random variation by growing each tree with a bootstrap sample from the training data and using only a small random sample of the predictors to define the split at each node, where the predictions for the trees are performed using a voting system. The advantages of RF include very high classification accuracy, determination of variable importance, and the ability to model complex interactions among predictor variables [*Cutler et al.*, 2007]. Moreover, recent studies have shown that RF models predict spatial patterns in river characteristics better than other more conventional methods [e.g., *Booker and Snelder*, 2012]. The RF technique has been previously applied in water resource studies to predict spatial patterns of different ecosystem components such as river bed surface grain size [*Snelder et al.*, 2011], biotic indices [*Álvarez-Cabria et al.*, 2017], and lake trophic state [*Hollister et al.*, 2016], among others.

RFs were developed using the R statistical language with the “caret” package, version 6.0–41 [*Kuhn*, 2008]. We used the additional feature-selection model in caret that uses the ”randomForest” [*Liaw and Wiener*, 2002] and “Boruta” packages [*Kursa and Rudnicki*, 2010]. Implementing a RF model with the Boruta algorithm assists with the selection of the most relevant independent variables to include in the RF model. Boruta is a feature selection wrapper algorithm that iteratively removes the features which proved to be less relevant than random probes [*Kursa and Rudnicki*, 2010]. Dependent-variable data was randomly partitioned into training (75%) and testing (25%) data sets, preserving the overall class distribution of perennial and temporary river segments. To fit the model, we used the cross-validation resampling method in the training set. To minimize any bias resulting from the random data splitting, a tenfold cross-validation was repeated 5 times for each of the models. This step was performed twice; first, applying the feature-selection method (Random Forest with Boruta), and then, with the final independent variables selected to be included in the modeling (the selected model). Although model performance can be optimized for the number of trees and the number of predictors used at each split, we used the recommended default values. These values were the square root of the number of predictor variables used to define the number of variables available for splitting at each tree node, and 500 as the maximum number of trees.

The average of the overall accuracy (proportion of the total number of segments that are correctly identified), sensitivity (proportion of temporary segments that are correctly identified as such), and specificity (proportion of perennial segments that are correctly identified as such) statistics was calculated for the resampling results with the optimal variables selected. These statistics were estimated according to the following equations: overall accuracy = (TP + TN)/(TP + FP + TN + FN), sensitivity = TP/(TP + FN), and specificity = TN/(TN + FP), where TP are true positives, TN are true negatives, FP are false positives, and FN are false negatives. We considered the temporary class as positive and the perennial class as negative. Exact binomial 95% confidence intervals (CIs) were also calculated for overall accuracy, sensitivity, and specificity (see *Collett* [2002] for details). Then, the averages of the overall accuracy, sensitivity, and specificity for the resampling results were chosen as model performance measures and the entire training set was used to fit the final model. We evaluated the fitted model on the test data set and calculated the same three statistics in order to compare the results.

The importance of the independent variables was calculated according to the results of the Mean Decrease Gini Index that measures the total decrease in node impurity, averaged over all trees using the Gini Index (the purer a node is, the smaller the Gini Index is, indicating that a node contains observations which are predominantly from a single class). The Gini Index is defined as $i(t)=\sum_{i\neq j} p(i|t)p(j|t)$ where $p(i|t)$ is the probability that a case is in class i given that it is node t and $p(j|t)$ is the probability that a case is in class j given that it is node t [*Breiman et al.*, 1984]. We also used partial dependence plots to show the marginal contribution of the most important variables to the response. These plots are not a perfect representation of the effects of each variable, but they provide useful information for illustration and may be used to graphically characterize relationships between individual predictor variables and predicted probabilities of a class presence [*Cutler et al.*, 2007; *Friedman and Meulman*, 2003]. We used the fitted model to predict the temporary or perennial membership of segments of the river network without any empirical information on the dependent variable (i.e., without field or aerial image data). Finally, we integrated the predicted classification values with the empirical information on class membership to achieve a final map showing the temporary/perennial character of the whole river network. The maps were created with ArcGIS [*ESRI*, 2014]. All models were developed using R 3.1.3 software [*R Core Team*, 2015] and the RStudio editor [*RStudio*, 2015].

2.3. Validation of Mapping Results

For an alternative validation of our final maps (which integrate the dependent variables based on empirical information and the RF model predictions), we used two external data sources. First, we used field data from River Habitat Surveys (RHS) [*Environment Agency*, 2003] carried out in the area and, second we also garnered expert knowledge from actual forest guards in the area.

Field data were obtained from an existing database with RHS data (www.rhs.ihcantabria.com). Data were obtained from 64 500 m long river segments during the summer seasons of 2008 (n = 42), 2009 (n = 8), 2010 (n = 1), and 2011 (n = 13). We used the information from section E of the RHS field form, where up to 10 different flow types are recorded every 50 m (dry flow is also recorded). Because a river segment could only be classified as temporary or perennial, we considered river segments to be temporary if at least half of the spot checks were dry, otherwise they were classified as perennial. Hydrologically, the summer seasons were considered "normal" for 2008, "normal" for 2009, "moderately wet" for 2010, and "normal" for 2011, according to the standardized precipitation index developed by the Spanish Meteorological Office for the relevant months [*AEMET*, 2016].

To design the forest guards' validation of our mapping results, we considered the administrative organization of the Deva-Cares catchment. This catchment is divided into three different administrative regions: Castilla y León (≈150 km^2), Principado de Asturias (≈415 km^2), and Cantabria (≈640 km^2). The forest guards spent most of their days in the field up and down along the catchment because of the many duties they perform (e.g., biodiversity inventories and monitoring, enforcement of environmental regulations, issues relating to fishing and game, providing assistance to local farmers, etc.). As an example, pursuant to regional fish management policies, they annually electro-fish those parts of the river network that dry out in the summer to rescue trout that get trapped in drying pools. This means that they know the river sections that dry out each year very well. Because of their experience and knowledge of the area, we consider their opinions to be a very valuable source of data and, thus, an appropriate validation approach for the final map. In order to take advantage of their knowledge and experience and use it for our purposes, the final map with the spatial distribution of perennial and temporary segments was presented to them at meetings which were held at the headquarters of the Asturias and Cantabrian regions. During those meetings, booklets with the maps covering the headquarters' domain were handed out. These booklets were then recollected after 2–3 weeks, containing annotations made by the headquarters' personnel on the limits of the perennial/temporary character of the river network within their domain.

These results were then assigned to individual segments of the river network to perform a comparative assessment of their expert opinions, the RHS data, and the temporary/perennial characterization from our mapping results. To achieve this, we created a confusion matrix with the mapping results, the expert opinions from the forest guards and the RHS observations as reference data. Overall accuracy (proportion of the total number of river segments that are correctly identified), sensitivity (proportion of temporary segments that are correctly identified as such), and specificity (proportion of perennial segments that are correctly identified as such) statistics were calculated for both with 95% CIs.

3. Results

3.1. Mapping Approach

A total of 2701 river segments of the river network were included in the a priori classification, with 1282 classified as perennial (308.5 km) and 1419 as temporary (194.5 km). Twelve independent variables were included in the final RF model with Boruta (Table 1), all of which were selected as relevant. Therefore, all the variables were included in the RF model to determine the temporary/perennial character of the 8555 segments without empirical information. The average of the overall accuracy (95% CIs), sensitivity (95% CIs), and specificity (95% CIs) for the resampling results were 1 (0.98–1), 1 (0.96–1), and 0.99 (0.95–1), respectively. The values of these three statistics on the test data set were 0.99 (0.98–1), 1 (0.98–1), and 0.99 (0.97–1) for overall accuracy (95% CIs), sensitivity (95% CIs), and specificity (95% CIs), respectively.

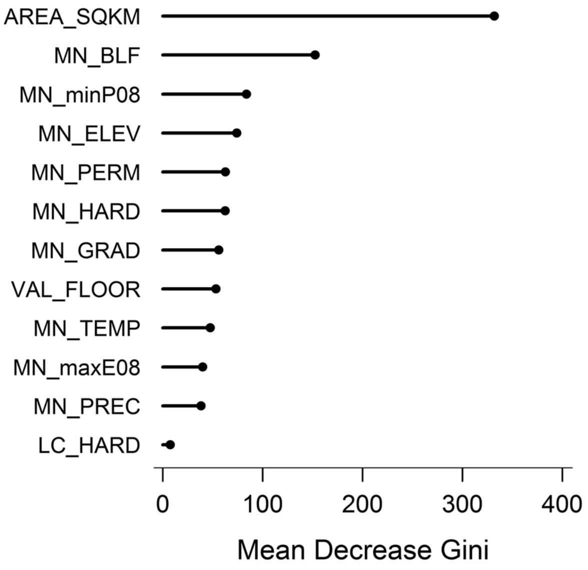

Figure 2. Importance of the independent variables (see variable code description in Table 1) in the fitted model in relation to the Mean Decrease Gini Index.

The most important independent variables (Mean Decrease Gini Index; Figure 2) were catchment area (AREA_SKQM), area occupied by broadleaf forest in the upstream catchment (MN_BLF), minimum monthly value of precipitation in August in the upstream catchment (MN_minP08), and average catchment elevation of the upstream catchment (MN_ELEV). Conversely, the variable with the least importance in the model was the average rock hardness within the segment wings (LC_HARD). Partial plots of the most influential variables (Figure 3) indicate that as watershed area decreases, there is a higher probability of a temporary classification (Figure 3a). Conversely, as watershed area increases, there is a higher probability of a perennial classification with an important threshold at approximately 20 km^2. In the case of MN_BLF, the probability of temporary classification decreases as the percentage of

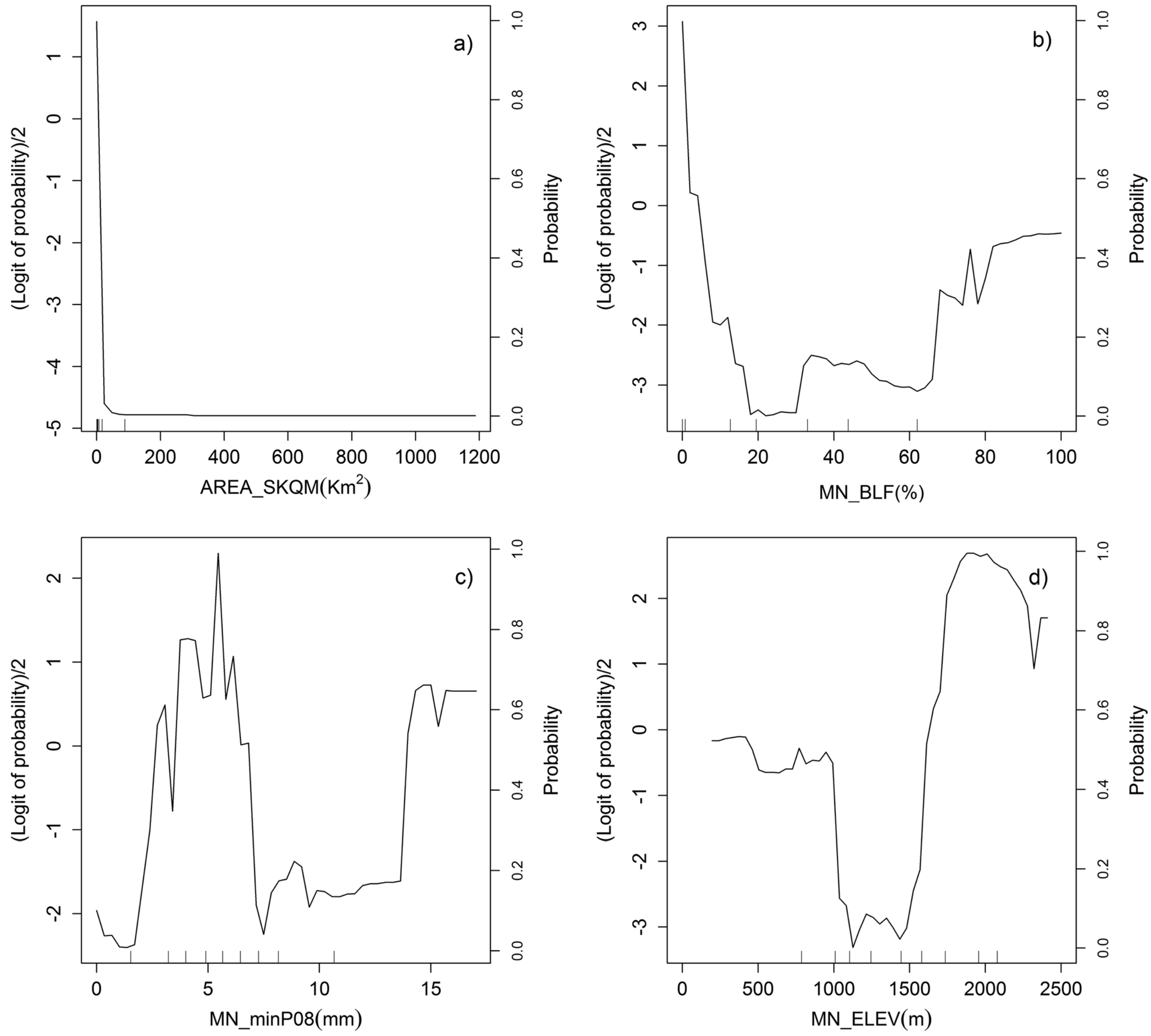

Figure 3. Partial dependence plots for the four most influential variables in the model for the temporary class ((a) AREA_SQKM; (b) MN_BLF; (c) MN_minP08; (d) MN_ELEV: see variable code description in Table 1). The "rug" at the bottom show the deciles of the distribution of sites across that independent variable. In the case of the perennial class, partial dependent plot is the mirror image of these partial dependence plot, and only one class was used for interpretation. The values of the left *y* axis (logit of probability/2) are also represented in probabilities on the right *y* axis.

forest in the catchment increases (until ≈20%). For MN_BLF values between ≈20 and 70%, the probability of temporary segments is low and more or less constant, but doubles for MN_BLF values above 70% (Figure 3b). For the other two most important variables (MN_minP08 and MN_ELEV; Figures 3c and 3d), the probability of perennial classification is higher for their median values, and in the case of MN_minP08, when these variable approaches zero (see Figure 4 for the distribution of temporary/perennial character recorded in the empirical observations).

The fitted model predicted 6106 (818 km) river segments to be temporary, while 2449 (354 km) were regarded as perennial (Figure 5b). The final map of the Deva-Cares catchment in low-flow conditions (Figure 5c) comprises the predicted river segments (Figure 5b) and the river segments with initial a priori classification data (Figure 5a). This final map shows 7525 temporary river segments (1012.5 km) and 3731 perennial river segments (662.5 km).

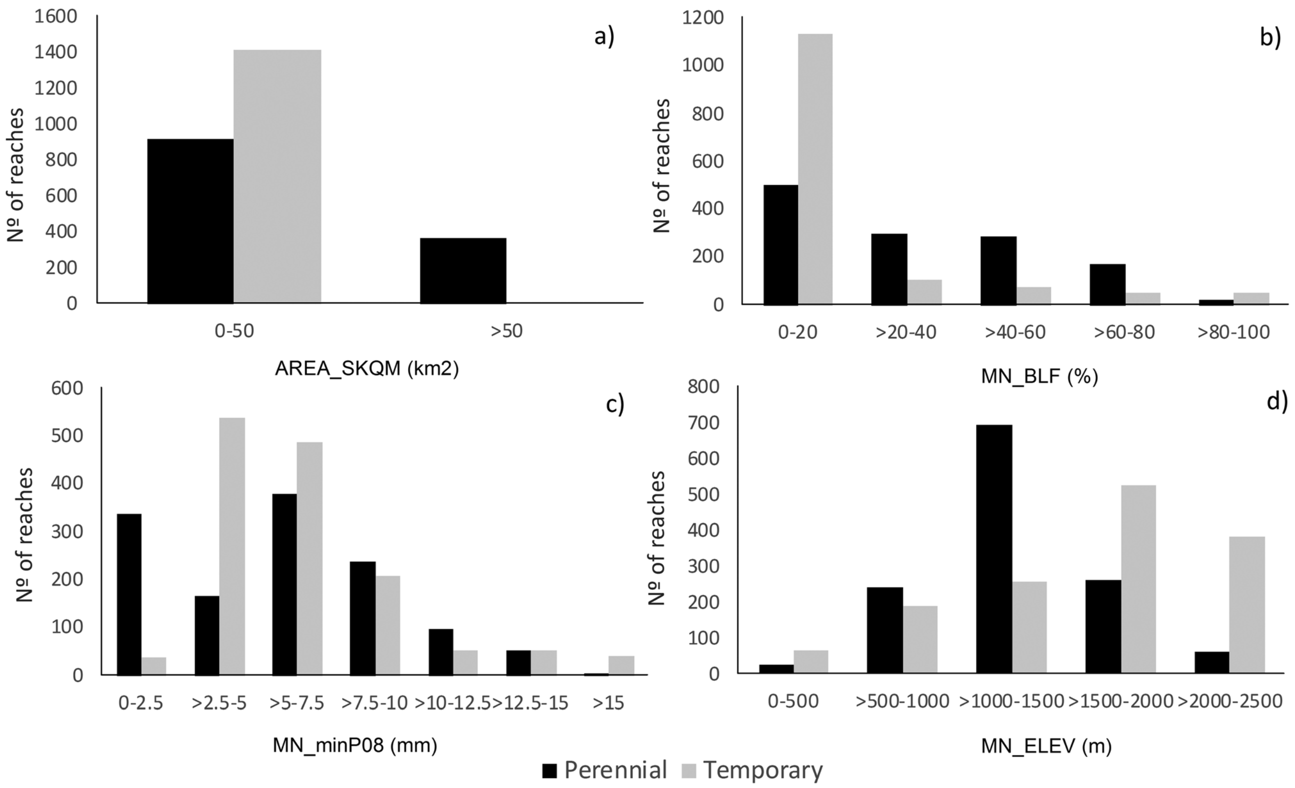

Figure 4. Distribution of the temporary and perennial character in the 2701 segments used as dependent variables according to different ranges of the four most important independent variables in the Random Forest model ((a) AREA_SQKM; (b) MN_BLF; (c) MN_minP08; (d) MN_ELEV: see variable code description in Table 1).

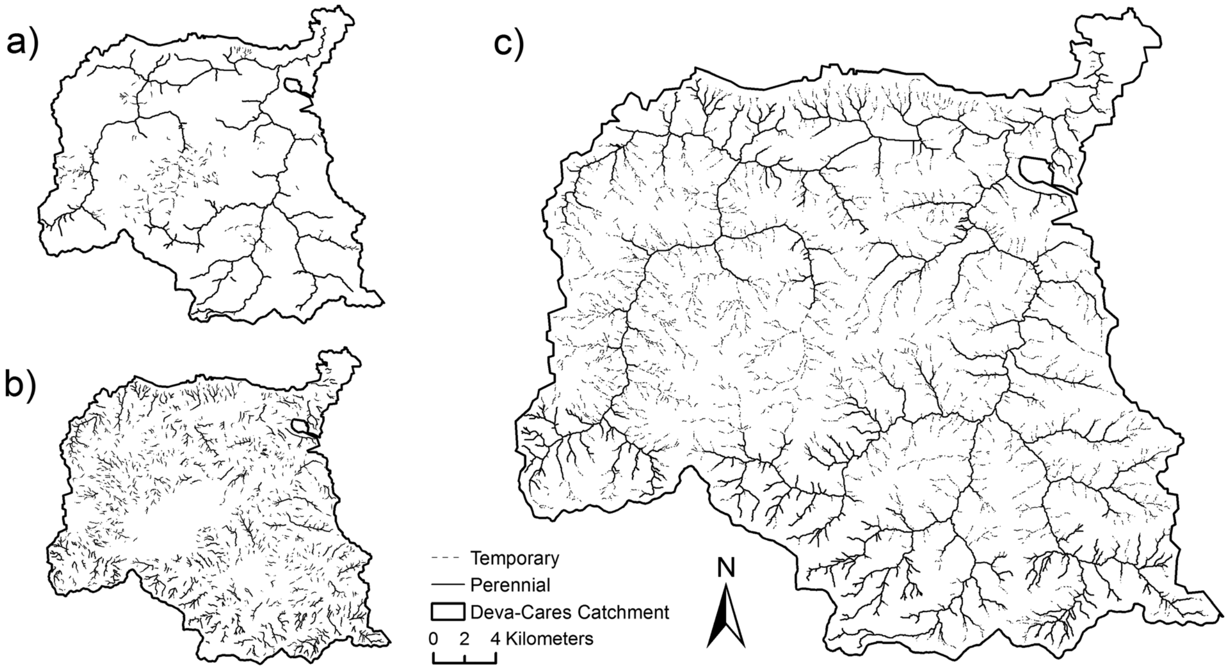

Figure 5. (a) Spatial distribution of the segments whose class membership was assigned by aerial images or field measurements. (b) Spatial distribution of the segments whose class membership was predicted with Random Forest model. (c) Representation of the temporary and perennial segments in the entire river network comprising the segments predicted (modeled with Random Forest) and segments used as dependent variables (assigned by aerial images or field measurements).

3.2. Validation of Mapping Results

The validation of the perennial/temporary characterization of the mapping results (Figure 5c) with RHS data (64 segments with RHS surveys data) shows the following values of overall accuracy = 0.91 (0.81–0.96), sensitivity = 0.67 (0.09–0.99), and specificity = 0.92 (0.82–0.97).

Regarding the evaluation of the expert opinions of the forest guards, the evaluation from Principado de Asturias could not be quantitatively assessed, because they did not provide geographical indications on the map. However, their general comments were that they agreed with the map, although suggesting that the highest perennial tributaries located in the Picos de Europa National Park should be slightly trimmed down. Conversely, the Cantabrian forest guards provided geographical indications for their entire respective area (5623 river segments). The evaluation of the Cantabrian forest guards shows high values of overall accuracy = 0.9 (0.9–0.91), sensitivity = 0.9 (0.89–0.91), and specificity = 0.91 (0.9–0.92).

4. Discussion

The strategy presented in this study to map the perennial/temporary character of river segments for a whole river network has been shown to be a plausible approach from a statistical point of view, and from the perspective of experts in the field. The length of the temporary river segments (1012.5 km) in our study area represented around 60% of the total channel length (1675 km) of the river network, so temporary streams should not be neglected when conducting catchment scale studies. Similar percentages have been found in other studies. For example, *Nadeau and Rains* [2007] reports that 59% of the total stream length in the USA, excluding Alaska, are temporary streams. *Tzoraki et al*. [2007] reports that 43% of the total area of Greece is drained by temporary rivers. We believe that the maps generated in this study could be a key digital resource for future research and management strategies in the selected catchment, with numerous potential applications.

4.1. Mapping Approach

Field surveys and aerial images provide information that is relatively easy to collect in order to classify rivers as temporal or perennial together with a modeling approach. Other studies have used other sources of data, such as the involvement of citizens [e.g., *Datry et al*., 2016; *Turner and Richter*, 2011] and gauging stations [e.g., *Snelder et al*., 2013]. Gauging stations provide information about the frequency and duration of temporary flow using flow time series, and facilitate a more specific definition of flow intermittency. However, most catchments typically have few gauging stations, and these are usually located in perennial river segments. This was the case in our study area, in which there were few active official gauging stations (three in river channels and four in reservoirs) all of which were situated in perennial river segments.

In this study, modeling the temporary and perennial character of river segments yielded high values of overall accuracy, sensitivity and specificity in the evaluation of the fitted model on the training and testing data set. These high values demonstrate the model's capacity to learn from the training data set and to identify the main relationships between the dependent and independent variables, and indicates the prediction capacity of the model. Conversely, *Snelder et al*. [2013] reported that RF performed poorly at classifying flow regimes and suggested that the performance results were due to the fact that intermittence is also controlled by processes acting at smaller scales. We believe that the good performance of the model in our case might be related to a number of different issues. First, the hydrogeological functioning of the Deva-Cares catchment allowed us to make some assumptions (section 2.2.2; i.e., considering any segment upstream of a segment empirically classified as temporary to also be temporary, and any segment downstream of a segment empirically classified as perennial to also be perennial) that may have improved the performance of some of the selected predictor variables. In this regard, the catchment area played an important role in differentiating between temporal and perennial river segments, with 20 km^2 being an important catchment size threshold. However, these assumptions might not hold in other catchments where different hydrological processes dominate, for example in large rivers where alluvial aquifers generate temporary flows [e.g., *Larned et al*., 2011], or where anthropogenic water regulation changes the temporary/perennial character of a river segment (e.g., water abstraction or diversion [*Datry et al*., 2014]). In these cases, it might also be necessary to include other variables related to groundwater dynamics, such as riverbed permeability or aquifer structure [*Snelder et al*., 2013] or anthrophic variables such us the distance to the point of abstraction. Second, our modeling approach was developed in a local catchment in which very small tributaries were considered (employing a slope-dependent drainage area threshold for estimating the location of channel heads in the delineation of the river network), but limitations to identifying temporary/perennial segments had been previously identified using different criteria (e.g., permanent channel length criteria). These assumptions might also be inappropriate when working at very large regional scales where most of the headwaters have been trimmed off from the river network (e.g., establishing a minimum area or order) and local knowledge regarding the abundance and location of special cases (e.g., catchments with numerous sinkholes) is not available [e.g., *Snelder et al*., 2013]. Third, the use of aerial images in the current study was possible because of the lack of woody vegetation at high altitudes and the large channel width of the lower segments. These characteristics allowed us to use the aerial images to identify temporary and perennial segments, respectively, in the river network. In other locations with different characteristics, the use of unmanned aerial vehicles (UAVs) might be an appropriate approach to cover large parts of the river network that could be used as segments in the training data set. While the use of UAVs for hydrological processes and modeling is still experimental and in continuous development, some studies have already shown satisfactory results [*Spence and Mengistu*, 2016]. Finally, our training data set included a balanced training data set (47.5% perennial and 52.5% temporary), which has been shown to be very relevant for increasing the accuracy of RF models [*González-Ferreras et al*., 2016]. Conversely, extremely imbalanced classes result in poor accuracy for the minority class because RF tends to focus more on the prediction accuracy of the majority class [*Chen et al*., 2004].

In relation to the independent variables, the catchment area and broadleaf-forest-percentage cover in the upstream catchment were the most important variables for classifying a segment as temporary or perennial, together with the minimum monthly value of precipitation in August, and the average catchment elevation. Catchment area has been shown to be an important determinant of perennial and temporary channels in other studies [*Snelder et al*., 2013; *Svec et al*., 2005]. The high frequency of temporary river segments in small catchment areas is supported by the temporary/perennial character distribution obtained from the empirical observations (Figure 4a).

The second most important variable, forest cover in the catchment, showed similar responses at either extreme of its continuum. It is known that different types of forest can have an important role in hydrology [*Cui et al*., 2012]. However, within this study we only considered broadleaf forest, because other types of forests were only present in 1% of the river segments with very low cover values. Minimum forest cover values tend to correlate with temporary character occurring more frequently (Figure 3b). These results match with the high occurrence of temporary river segments in the alpine area (Figure 4b). In the Deva-Cares catchment, temporary segments are common in the alpine karst areas where mountain grasslands and denuded rocks dominate and, consequently, a low percentage of forest is found. At the other extreme, the probability of temporary river segments also increased with higher values of forest cover. This result is consistent with studies that find increased water yields following timber harvesting [e.g., *Smerdon et al*., 2009]. In our case, this result corresponds to small catchments ($\leq$5 km^2) where broadleaf forests have been preserved (i.e., a high percentage of cover). Catchment forest coverage has also been shown to play a key role in determining hydrological spatial patterns in other studies [e.g., *Cui et al*., 2012]. Moreover, other variables related to forest presence could also play important roles in determining the temporary or perennial character of small river segments. For example, *Belmar et al*. [2016] has shown that mature forests in Cantabrian catchments may provide higher base flows during the summer months compared to young forests.

4.2. Validation of Mapping Results

The validation of our results using RHS survey data and expert knowledge from forest guards supports the methodology proposed in this work for mapping the temporal and perennial character of river segments. Both validations have shown high values of overall accuracy ($\geq$0.9). In the specific case of RHS data, the proportion of perennial segments correctly identified was high (specificity = 0.92), but the proportion of temporary segments correctly identified presented lower values (sensitivity = 0.67). This sensitivity value could be related to the lower proportion of temporary segments (n = 3) within the RHS data set (n = 64). This could negatively influence the validation process due to an unbalanced class data set or a dearth of information caused by a small sample size [*Ali et al*., 2015]. For this reason, including a larger number of temporary segments within the RHS data set would provide for a more robust validation process. When using the expert opinion data, the quantitative evaluation showed high values of sensitivity and specificity (0.9 and 0.91, respectively) in the Cantabrian region, which represents more than 50% of the study area. The qualitative assessment of the other administrative area (Principado de Asturias) was also positive. Regarding the model and the independent validations, we consider that the final map results represent a major improvement of the digital cartography of temporary and perennial river channels in the study area. The length of the segments using our mapping approach is greater than in existing maps. Our map has 1675 km of river segments, entailing 1012.5 km of temporary rivers and 662.5 km of perennial rivers, while the most detailed existing map of the area [*National Geographic Institute (IGN)*, 2016] (National Topographical Base of Spain at scale 1: 25000, BTN25) has 1126 km of river channels, listing 845 km of temporary rivers and 280 km of perennial rivers. Other studies have shown that this kind of topographic map also underestimates the total length of river channels [e.g., *Hansen*, 2001]. Moreover, comparing this map with our results, a great part of the length of the temporary rivers in the BTN25 (≈40%) are actually perennial according to our map, while less than 15% of the total length of perennial rivers in the BTN25 are temporary according to our mapping approach. Thus, our results extend and improve the existing information, highlighting the need to update the current information.

The knowledge gathered from the forest guards from the Cantabrian region allowed us to identify sinkholes in the area, which were not taken into account in the a priori classification which was based on our assumptions. This situation comprised six zones, where perennial flows exist upstream of temporary segments. The total length of temporary river segments downstream of perennial flows is 3.4 km, and the total length of perennial river segments upstream of temporary segments is 16.3 km. These values support our initial assumptions, because their length in relation to the whole river network length (temporary and perennial) in the Cantabrian part of the catchment (850.5 km) is minimal. However, the identification of these areas, especially in catchments where such areas are significantly represented, is important for the continuity of river processes and has a major impact on river biodiversity and functioning patterns. Accordingly, we stress the need to gather this type of punctual information in future studies whenever possible.

4.3. Potential Applications

Mapping the temporary or perennial character of a river network can provide a better understanding of hydrologic systems and the interaction between terrestrial and aquatic interfaces at large spatial scales (i.e., landscapes or watersheds). The current lack of data on how temporary and perennial river segments are spatially organized in river networks hinders the development of regional applications. To compensate, upscaling methods are needed to extrapolate information from river segments to whole catchments [*Blöschl*, 2006]. In our study, the combined use of field data, aerial images, and modeling provides a simple and replicable methodology to upscale information about perennial and temporary river channel character to the whole river network. This simple representation of the limits between temporary and perennial river segments for whole river networks at a catchment scale constitutes a first step in various applications. This exercise estimates the extent and magnitude of each class in the catchment and can stimulate research and management for a variety of different objectives. Below, we describe the importance of some of these applications to illustrate the relevance of this exercise:

1. Monitoring of areas likely to experience changes due to global change. Some evidence suggests that climate-driven temporary flow has increased and that it will continue to increase in the future [*Larned et al.*, 2010]. It is therefore important to delineate the potentially most sensitive zones that might suffer from the effects of global change. RF models can provide class probability results to help to delineate these zones by mapping the probability of a segment's being temporary or perennial instead of using the hard binary classification (temporary/perennial). In our study area, for example, we ran RF probability class and around 130 km of channels have a temporary class probability between 0.45 and 0.55. All of them are located in catchments with areas less than 8.5 km^2 and could be identified as potentially most sensitive or transition zones.
2. Delimit terrestrial and aquatic systems and locate their interfaces. Because temporary streams are hydrologically dynamic, providing both terrestrial and aquatic habitats, their characterization and delimitation is important for determining different interactions, such as exchanges between terrestrial and aquatic organic matter and biotic interactions between terrestrial and aquatic organisms [*Datry et al.*, 2011]. Moreover, temporary channels function differently than perennial ones with respect to biogeochemical fluxes [*Datry et al.*, 2014] and may have different impacts on carbon and nutrient fluxes. Therefore, it is important to identify the transitions between aquatic and terrestrial phases, referred to as hot spots and hot moments for biogeochemistry [*McClain et al.*, 2003].

3. Habitat delineation for different groups of organisms. In temporary streams that desiccate entirely, fish must seek refuge in perennial segments, whereas in streams that dry partially, some fish can survive in disconnected pools [e.g., *Pires et al*., 1999]. Considering that stream drying is stressful for fish and causes high mortality [*Davey and Kelly*, 2007], fishes are generally restricted to perennially flowing segments. In the case of amphibians, temporary streams are important, because the differential vulnerability of larvae to aquatic predators may exclude some species from perennial streams [*Woodward*, 1983]. The temporary/perennial character of a channel has been shown to be important for invertebrates [e.g., *Leigh et al*., 2016] and algae [e.g., *Robson and Matthews*, 2004]. The spatial distribution of temporary and perennial channels has important implications for the distribution patterns of aquatic species and is thus a key feature for aquatic biodiversity management and conservation.

5. Conclusions

This study proposes a relatively easy method for estimating the occurrence and extent of perennial and temporary segments in whole river networks at catchment scale. This strategy allows the determination of which variables are the most important based on catchment characteristics which play a fundamental role in determining the spatial distribution of flow permanence. Our approach can be applied anywhere in the world using minimal data resources, although the applicability of our assumptions to other locations should be carefully considered and be based on a working knowledge of the fundamental hydrological processes of the studied catchment. Thus, with a few field surveys taken in the low-flow season, access to aerial images and a virtual watershed approach [*Benda et al*., 2015], it is possible to obtain dependent and independent variables to build a temporal and perennial river segment classification model. Incorporating available knowledge from locals and experts in the area can also enhance the mapping approach and improve the final digital map representing the spatial distribution of temporary and perennial segments in the particular river network.

Acknowledgments

We would like to thank the Journal Editor and the three referees for their comments and suggestions, which have greatly improved the manuscript. This study was partly funded by the Spanish Ministry of Economy and Competitiveness as part of the RIVERLANDS (Ref: BIA-2012–33572) and HYDRA (Ref: BIA-2015–71197) projects. Alexia María González-Ferreras is supported by a predoctoral research grant (Ref: BES-2013–065770) from the Spanish Ministry of Economy and Competitiveness, and José Barquín was supported by a Ramon y Cajal grant (Ref: RYC-2011–08313) from the Spanish Ministry of Economy and Competitiveness. We would like to thank the Government of Cantabria, the Principado de Asturias and the forest guards of the study areas for providing useful information. We would also like to acknowledge the Interautonomic Consortium of the Picos de Europa National Park and the Biodiversity Foundation from the Ministry of Agriculture, Food and Environment, for their advice and project support. Finally, we would also like to thank all the people involved in the field data collection, and those who read an early draft of the manuscript and suggested several improvements. The data and the data sources used in this study are cited and explained in the text. Readers can obtain further information about the data supporting the analysis and conclusions by contacting the corresponding author.

References

Acuña, V., T. Datry, J. Marshall, D. Barceló, C. N. Dahm, A. Ginebreda, G. McGregor, S. Sabater, K. Tockner, and M. A. Palmer (2014), Why should we care about temporary waterways?, *Science*, *343*(6175), 1080–1081, doi:10.1126/science.1246666.

Adrados, L., et al. (2012), *Parque Nacional de los Picos de Europa. Guía Geológica (Picos de Europa National Park. Geological guide)*, 2nd ed., Adrados Ediciones, Oviedo, Spain.

Ali, A., S. M. Shamsuddin, and A. L. Ralescu (2015), Classification with class imbalance problem: A review, *Int. J. Adv. Soft Comput. Appl.*, *7*(3), 176–204.

Álvarez-Cabria, M., J. Barquín, and F. J. Peñas (2016), Modelling the spatial and seasonal variability of water quality for entire river networks: Relationships with natural and anthropogenic factors, *Sci. Total Environ.*, *545–546*, 152–162, doi:10.1016/j.scitotenv.2015.12.109.

Álvarez-Cabria, M., A. M. González-Ferreras, F. J. Peñas, and J. Barquín (2017), Modelling macroinvertebrate and fish biotic indices: From reaches to entire river networks, *Sci. Total Environ.*, *577*, 308–318, doi:10.1016/j.scitotenv.2016.10.186.

Arthington, A. H., J. M. Bernardo, and M. Ilheu (2014), Temporary rivers: Linking ecohydrology, ecological quality and reconciliation ecology, *River Res. Appl.*, *30*(10), 1209–1215, doi:10.1002/rra.2831.

Ballesteros, D., M. Jiménez-Sánchez, J. García-Sansegundo, and S. Giralt (2011), Geological methods applied to speleogenetical research in vertical caves: The example of Torca Teyera shaft (Picos de Europa, northern Spain), *Carbonates Evaporites*, *26*(1), 29–40, doi:10.1007/s13146-011-0052-7.

Ballesteros, D., M. Jiménez-Sánchez, M. J. Domínguez-Cuesta, J. García-Sansegundo, and M. Meléndez-Asensio (2015), Geoheritage and geodiversity evaluation of Endokarst landscapes: The Picos de Europa National Park, North Spain, in *Hydrogeological and Environmental Investigations in Karst Systems*, edited by B. Andreo et al., pp. 619–627, Springer, Berlin Heidelberg.

Barquin, J., et al. (2015), Coupling virtual watersheds with ecosystem services assessment: A 21st century platform to support river research and management, *WIREs Water*, *2*(6), 609–621, doi:10.1002/wat2.1106.

Belmar, O., J. Barquín, J. M. Álvarez-Martínez, F. J. Peñas, and M. Del Jesus (2016), The role of forest maturity on catchment hydrologic stability, *Hydrol. Earth Syst. Sci. Discuss.*, *2016*, 1–17, doi:10.5194/hess-2016-471.

Benda, L., D. Miller, and J. Barquín (2011), Creating a catchment scale perspective for river restoration, *Hydrol. Earth Syst. Sci.*, *15*(9), 2995–3015, doi:10.5194/hess-15-2995-2011.

Benda, L., D. Miller, J. Barquín, R. McCleary, T. Cai, and Y. Ji (2015), Building virtual watersheds: A global opportunity to strengthen resource management and conservation, *Environ. Manage.*, *57*(3), 722–739, doi:10.1007/s00267-015-0634-6.

Blöschl, G. (2006), Statistical upscaling and downscaling in hydrology, in *Encyclopedia of Hydrological Sciences*, edited by M. G. Anderson, pp. 135–154, John Wiley, Chichester, U. K.

Booker, D. J., and T. H. Snelder (2012), Comparing methods for estimating flow duration curves at ungauged sites, *J. Hydrol.*, *434–435*, 78–94, doi:10.1016/j.jhydrol.2012.02.031.

Boulesteix, A.-L., S. Janitza, J. Kruppa, and I. R. König (2012), Overview of random forest methodology and practical guidance with emphasis on computational biology and bioinformatics, *WIREs Data Mining Knowl. Discov.*, *2*(6), 493–507, doi:10.1002/widm.1072.

Breiman, L. (2001), Random Forests, *Mach. Learn.*, *45*(1), 5–32, doi:10.1023/a:1010933404324.

Breiman, L., J. Friedman, R. A. Olshen, and C. J. Stone (1984), Classification and Regression Trees, Wadsworth Int. Group, Belmont, Calif.

Chen, C., A. Liaw, and L. Breiman (2004), Using random forest to learn imbalanced data, Univ. of Calif., Berkeley.

Collett, D. (2002), *Modelling Binary Data*, 2nd ed., Taylor and Francis, Boca Raton, Fla.

Cui, X., S. Liu, and X. Wei (2012), Impacts of forest changes on hydrology: A case study of large watersheds in the upper reaches of Minjiang River watershed in China, *Hydrol. Earth Syst. Sci.*, *16*(11), 4279–4290, doi:10.5194/hess-16-4279-2012.

Cutler, D. R., T. C. Edwards Jr., K. H. Beard, A. Cutler, and K. T. Hess (2007), Random forests for classification in ecology, *Ecology*, *88*(11), 2783–2792, doi:10.1890/07-0539.1.

Datry, T., D. B. Arscott, and S. Sabater (2011), Recent perspectives on temporary river ecology, *Aquat. Sci.*, *73*(4), 453–457, doi:10.1007/s00027-011-0236-1.

Datry, T., S. T. Larned, and K. Tockner (2014), Intermittent rivers: A challenge for freshwater ecology, *Bioscience*, *64*(3), 229–235, doi:10.1093/biosci/bit027.

Datry, T., H. Pella, C. Leigh, N. Bonada, and B. Hugueny (2016), A landscape approach to advance intermittent river ecology, *Freshwater Biol.*, *61*(8), 1200–1213, doi:10.1111/fwb.12645.

Davey, A. J. H., and D. J. Kelly (2007), Fish community responses to drying disturbances in an intermittent stream: A landscape perspective, *Freshwater Biol.*, *52*(9), 1719–1733, doi:10.1111/j.1365-2427.2007.01800.x.

Dietrich, W. E., C. J. Wilson, D. R. Montgomery, and J. McKean (1993), Analysis of erosion thresholds, channel networks, and landscape morphology using a digital terrain model, *J. Geol.*, *101*(2), 259–278, doi:10.1086/648220.

Environment Agency (2003), *River Habitat Survey in Britain and Ireland, Field Survey Guidance Manual: 2003*, Bristol, U. K.

ESRI (2014), ArcGIS Desktop 10.2.1 (GIS software), Redlands, Calif.

European Commission (2000), *Directive 2000/60/EC of the European Parliament and of the Council—Establishing a Framework for Community Action in the Field of Water Policy*, European Union, Brussels, Belgium.

Fernández-Giber, E., J. M. Calaforra, and C. Rossi (2000), Speleogenesis in the Picos de Europa Massif, Northern Spain, in *Speleogenesis: Evolution of Karst Aquifers*, edited by A. B. Klimchouk et al., pp. 352–357, Natl. Speleol. Soc., Huntsville, Alabama.

Friedman, J. H., and J. J. Meulman (2003), Multiple additive regression trees with application in epidemiology, *Stat. Med.*, *22*(9), 1365–1381, doi:10.1002/sim.1501.

Fritz, K. M., E. Hagenbuch, E. D'Amico, M. Reif, P. J. Wigington, S. G. Leibowitz, R. L. Comeleo, J. L. Ebersole, and T.-L. Nadeau (2013), Comparing the extent and permanence of headwater streams from two field surveys to values from hydrographic databases and maps, *J. Am. Water Resour. Assoc.*, *49*(4), 867–882, doi:10.1111/jawr.12040.

Grupo de Emisarios Submarinos e Hidraulica Ambiental (GESHA) (2005), Estudio de los Recursos Hídricos de los ríos de la vertiente norte de Cantabria (Study of the Water Resources of the rivers of the north of Cantabria), Univ. of Cantabria, Santander, Spain.

González-Ferreras, A. M., J. Barquín, and F. J. Peñas (2016), Integration of habitat models to predict fish distributions in several watersheds of Northern Spain, *J. Appl. Ichthyol.*, *32*(1), 204–216, doi:10.1111/jai.13024.

Hansen, W. F. (2001), Identifying stream types and management implications, *For. Ecol. Manage.*, *143*(1–3), 39–46, doi:10.1016/S0378-1127(00)00503-X.

Hollister, J. W., W. B. Milstead, and B. J. Kreakie (2016), Modeling lake trophic state: A random forest approach, *Ecosphere*, *7*(3), e01321, doi:10.1002/ecs2.1321.

IGME (2015), Mapa de Litologias de España 1/1.000.000 (Spain map lithologies 1/1.000.000), Madrid, Spain. [Available at http://mapas.igme.es/Servicios/default.aspx#IGME_Litologias_1M.]

Kuhn, M. (2008), Building predictive models in R using the caret package, *J. Stat. Software*, *28*(5), 1–26.

Kursa, M. B., and W. R. Rudnicki (2010), Feature selection with the Boruta package, *J. Stat. Software*, *36*(11), 1–13.

Larned, S. T., T. Datry, D. B. Arscott, and K. Tockner (2010), Emerging concepts in temporary-river ecology, *Freshwater Biol.*, *55*(4), 717–738, doi:10.1111/j.1365-2427.2009.02322.x.

Larned, S. T., J. Schmidt, T. Datry, C. P. Konrad, J. K. Dumas, and J. C. Diettrich (2011), Longitudinal river ecohydrology: Flow variation down the lengths of alluvial rivers, *Ecohydrology*, *4*(4), 532–548, doi:10.1002/eco.126.

Leigh, C., N. Bonada, A. J. Boulton, B. Hugueny, S. T. Larned, R. Vander Vorste, and T. Datry (2016), Invertebrate assemblage responses and the dual roles of resistance and resilience to drying in intermittent rivers, *Aquat. Sci.*, *78*(2), 291–301, doi:10.1007/s00027-015-0427-2.

Leigh, C., B. Laporte, N. Bonada, K. Fritz, H. Pella, E. Sauquet, K. Tockner, and T. Datry (2017), IRBAS: An online database to collate, analyze, and synthesize data on the biodiversity and ecology of intermittent rivers worldwide, *Ecol. Evol.*, *7*(3), 815–823, doi:10.1002/ece3.2679.

Liaw, A., and M. Wiener (2002), Classification and regression by randomForest, *R News*, *2*(3), 18–22.

McClain, E. M., et al. (2003), Biogeochemical hot spots and hot moments at the interface of terrestrial and aquatic ecosystems, *Ecosystems*, *6*(4), 301–312, doi:10.1007/s10021-003-0161-9.

McDonough, O. T., J. D. Hosen, and M. A. Palmer (2011), Temporary streams: The hydrology, geography, and ecology of non-perennially flowing waters, in *River Ecosystems: Dynamics, Management and Conservation*, edited by H. S. Elliot and L. E. Martin, pp. 64–68, Nova Sci., Hauppauge, New York.

Miller, D. (2002), Program for DEM analysis, in *Landscape Dynamics and Forest Management*, *Gen. Tech. Rep. RMRS-GTR-101CD*, USDA For. Serv., Rocky Mountain Res. Sta., Fort Collins, Colo.

Montgomery, D. R., and W. E. Dietrich (1992), Channel Initiation and the problem of landscape scale, *Science*, *255*(5046), 826–830, doi:10.1126/science.255.5046.826.

Nadeau, T.-L., and M. C. Rains (2007), Hydrological connectivity between headwater streams and downstream waters: How science can inform policy, *J. Am. Water Resour. Assoc.*, *43*(1), 118–133, doi:10.1111/j.1752-1688.2007.00010.x.

National Geographic Institute (IGN) (2016), Base Topográfica Nacional de España 1:25.000 - BTN25 (National Topographical Base of Spain 1:25.000 - BTN25). [Available at http://centrodedescargas.cnig.es/CentroDescargas/.]

Nikolaidis, N. P., et al. (2013), Towards sustainable management of Mediterranean river basins: Policy recommendations on management aspects of temporary streams, *Water Policy*, *15*(5), 830–849, doi:10.2166/wp.2013.158.

Peñas, F. J., J. Barquin, T. H. Snelder, D. J. Booker, and C. Alvarez (2014), The influence of methodological procedures on hydrological classification performance, *Hydrol. Earth Syst. Sci.*, *18*(9), 3393–3409, doi:10.5194/hess-18-3393-2014.

Pires, A. M., I. G. Cowx, and M. M. Coelho (1999), Seasonal changes in fish community structure of intermittent streams in the middle reaches of the Guadiana basin, Portugal, *J. Fish Biol.*, *54*(2), 235–249, doi:10.1111/j.1095-8649.1999.tb00827.x.

R Core Team (2015), R: A language and environment for statistical computing, Found. for Stat. Comput., Vienna.

Reid, L. M., and R. R. Ziemer (1994), Evaluating the biological significance of intermittent streams, in *Issues in Watershed Analysis. Discussions at Interdisciplinary and Interagency Workshops*, Humboldt Interagency Watershed Anal. Cent., edited by USDA Forest Service Pacific Southwest Research Station, McKinleyville, Calif.

Rivas-Martínez, S., A. Penas, and T. E. Díaz (2004), Bioclimatic map of Europe, Bioclimates, Cartogr. Serv., Univ. of León, León, Spain.

Robinson, C. T., D. Tonolla, B. Imhof, R. Vukelic, and U. Uehlinger (2016), Flow intermittency, physico-chemistry and function of headwater streams in an Alpine glacial catchment, *Aquat. Sci.*, *78*(2), 327–341, doi:10.1007/s00027-015-0434-3.

Robson, B. J., and T. G. Matthews (2004), Drought refuges affect algal recolonization in intermittent streams, *River Res. Appl.*, *20*(7), 753–763, doi:10.1002/rra.789.

RStudio (2015), RStudio: Integrated development environment for R (Version 0.98.1103), Boston, Mass.

Sabater, S., and K. Tockner (2010), Effects of hydrologic alterations on the ecological quality of river ecosystems, in *Water Scarcity in the Mediterranean: Perspectives Under Global Change*, edited by S. Sabater and D. Barceló, pp. 15–39, Springer, Berlin Heidelberg.

Sánchez de Dios, R., M. Benito-Garzón, and H. Sainz-Ollero (2009), Present and future extension of the Iberian submediterranean territories as determined from the distribution of marcescent oaks, *Plant Ecol.*, *204*(2), 189–205, doi:10.1007/s11258-009-9584-5.

Sando, R., and K. W. Blasch (2015), Predicting alpine headwater stream intermittency: A case study in the northern Rocky Mountains, *Ecohydrol. Hydrobiol.*, *15*(2), 68–80, doi:10.1016/j.ecohyd.2015.04.002.

Shortridge, J. E., S. D. Guikema, and B. F. Zaitchik (2016), Machine learning methods for empirical streamflow simulation: A comparison of model accuracy, interpretability, and uncertainty in seasonal watersheds, *Hydrol. Earth Syst. Sci.*, *20*(7), 2611–2628, doi:10.5194/hess-20-2611-2016.

Smerdon, B. D., T. Redding, and J. Beckers (2009), An overview of the effects of forest management on groundwater hydrology, *BC J. Ecosyst. Manage.*, *10*(1), 22–44.

Snelder, T. H., N. Lamouroux, and H. Pella (2011), Empirical modelling of large scale patterns in river bed surface grain size, *Geomorphology*, *127*(3–4), 189–197, doi:10.1016/j.geomorph.2010.12.015.

Snelder, T. H., T. Datry, N. Lamouroux, S. T. Larned, E. Sauquet, H. Pella, and C. Catalogne (2013), Regionalization of patterns of flow intermittence from gauging station records, *Hydrol. Earth Syst. Sci.*, *17*(7), 2685–2699, doi:10.5194/hess-17-2685-2013.

Spanish National Meteorological Agency (AEMET) (2016), Resumen climatológico nacional sobre el estado del clima y la evolución de las principales variables climáticas (National climatological summary on the state of the climate and the evolution of the main climatic variables), Madrid, Spain. [Available at http://www.aemet.es/es/serviciosclimaticos/vigilancia_clima/resumenes.]

Spence, C., and S. Mengistu (2016), Deployment of an unmanned aerial system to assist in mapping an intermittent stream, *Hydrol. Processes*, *30*(3), 493–500, doi:10.1002/hyp.10597.

Svec, J. R., R. K. Kolka, and J. W. Stringer (2005), Defining perennial, intermittent, and ephemeral channels in Eastern Kentucky: Application to forestry best management practices, *For. Ecol. Manage.*, *214*(1–3), 170–182, doi:10.1016/j.foreco.2005.04.008.

Tooth, S. (2000), Process, form and change in dryland rivers: A review of recent research, *Earth Sci. Rev.*, *51*(1–4), 67–107, doi:10.1016/S0012-8252(00)00014-3.

Turner, D. S., and H. E. Richter (2011), Wet/dry mapping: Using citizen scientists to monitor the extent of perennial surface flow in dryland regions, *Environ. Manage.*, *47*(3), 497–505, doi:10.1007/s00267-010-9607-y.

Tzoraki, O., N. P. Nikolaidis, Y. Amaxidis, and N. T. Skoulikidis (2007), In-stream biogeochemical processes of a temporary river, *Environ. Sci. Technol.*, *41*(4), 1225–1231, doi:10.1021/es062193h.

Uys, C. M., and H. J. O'Keeffe (1997), Simple words and fuzzy zones: Early directions for temporary river research in South Africa, *Environ. Manage.*, *21*(4), 517–531, doi:10.1007/s002679900047.

Woodward, B. D. (1983), Predator-prey interactions and breeding-pond use of temporary-pond species in a desert anuran community, *Ecology*, *64*(6), 1549–1555, doi:10.2307/1937509.