Limnetica, 37 (1): 145-158(2018). DOI: 10.23818/limn.37.12
 ISSN: 0213-8409

A comparison of modeling techniques to predict hydrological indices in ungauged rivers

Francisco J. Peñas[1,2,*], Jose Barquín[3] and César Álvarez[3]

[1] Department of Civil Engineering. Universidad Católica de la Santísima Concepción, Alonso de Ribera 2850, Concepción, Chile.
[2] Centro de Investigación en Biodiversidad y Ambientes Sustentables. "CIBAS". Universidad Católica de la Santísima Concepción, Alonso de Ribera 2850, Concepción, Chile.
[3] Environmental Hydraulics Institute. "IH Cantabria". Universidad de Cantabria. Parque Científico y Tecnológico de Cantabria. C/Isabel Torres, 15, 39011, Santander, Spain.

[*] Corresponding author: fpenas@ucsc.cl

Received: 23/06/17 Accepted: 11/12/17

ABSTRACT

A comparison of modeling techniques to predict hydrological indices in ungauged rivers

Predicting the natural flow regime in ungauged rivers is an important challenge in water resource management and ecological research. We developed models to predict 16 hydrological indices in a river network covering the northern third of the Iberian Peninsula. Multiple Linear Regression (MLR), Generalized Additive Models (GAMs), Random Forest (RF) and Adaptive Neuro Fuzzy Inference System (ANFIS) were used and compared according to their prediction accuracy. The results showed that predictive performance varied greatly depending on the modeled hydrological attribute. The magnitude and frequency indices were predicted with excellent accuracy. In contrast, no technique was capable of developing precise models for hydrological indices of timing, duration and rate of change. This is mainly related to the lack of proper environmental databases on the scales on which these flow regime patterns are influenced. In addition, complex modeling techniques did not always outperform linear models and no single approach was optimal for all indices. ANFIS and GAMs provided the best results; however, other issues such as computational cost and the level of knowledge required to apply the method and interpret the results should be taken into account.

Key words: natural flow regime, prediction, linear regression, generalized additive models, machine learning

RESUMEN

Comparación de técnicas de modelado para predecir índices hidrológicos en ríos no aforados

La predicción del régimen natural de caudales en ríos no aforados representa un problema esencial para superar los nuevos retos a los que se enfrenta la gestión de los recursos hídricos y la ecología de los sistemas de agua dulce. En este trabajo hemos desarrollado modelos para predecir 16 índices hidrológicos en la red fluvial que cubre el tercio norte de la Península Ibérica. En concreto se han desarrollado y comparado Regresiones Lineales Múltiples (RLM), Modelos Aditivos Generalizados (MAG), Bosques Aleatorios (BA) y Sistemas Adaptativos de Inferencia de Lógica Difusa (SAILD). Los resultados han puesto de manifiesto que la capacidad predictiva varía significativamente dependiendo del tipo de índice hidrológico modelado. Los modelos de los índices de magnitud y frecuencia mostraron una capacidad predictiva excelente. Por el contrario, los modelos de los índices hidrológicos relacionados con la temporalidad, la duración de periodos de caudales altos o bajos y la tasa de cambio mostraron una capacidad de predicción limitada. Esto se relaciona, en gran medida, con la falta de bases de datos de variables predictoras con escalas espacio-temporal adecuadas. Por otro lado, las técnicas estadísticas más complejas no siempre mostraron capacidades predictivas mayores que los RLM y, además, no se encontró un método que ofreciese resultados óptimos para todos los índices. SAILD y MAG obtuvieron, por norma general, los mejores resultados, sin embrago, consideramos que otros elementos, tales como los recursos computacionales requeridos o la experiencia

necesaria para aplicar la técnica e interpretar los resultados, deben tenerse en muy en cuenta a la hora de seleccionar el método más adecuado.

Palabras clave: *Régimen natural de caudales, predicción, regresión múltiple, modelos aditivos generalizados, aprendizaje automático*

INTRODUCTION

River flow regime is a key element that structures freshwater ecosystems (Poff *et al.*, 1997). Indeed, the understanding of the bio-physical associations between hydrological variability and stream biological communities is a critical scientific and management challenge (Alvarez-Cabria *et al.*, 2017). However, it is frequently the case that streamflow data are not available at a site of interest such as where biomonitoring is carried out (Poff & Zimmerman, 2010; Sanborn & Bledsoe, 2006). This hinders the exploration of the flow regime influence on stream ecology and ultimately the management of these systems.

Natural flow regime can be described through a collection of ecologically relevant hydrological indices (Olden & Poff, 2003). Hence, interest in the prediction of these hydrological indices in ungauged streams has grown rapidly in recent years (Carlisle *et al.*, 2010; Kennen *et al.*, 2008). Most of the work has been aimed at addressing water yield and flooding issues. Thus, models to predict average flows, flood quantiles, flow duration curves or low-flow parameters dominate the literature (Sanborn & Bledsoe, 2006). In contrast, prediction of ecologically relevant hydrological indices has received limited attention (Carlisle *et al.*, 2010; Knight *et al.*, 2011; Sanborn & Bledsoe, 2006).

Multiple linear regression has been the most commonly used statistical technique to predict hydrological indices in ungauged sites (Knight *et al.*, 2011). However, the potential improvement in model performance when using other modeling procedures that do not assume specific distribu-

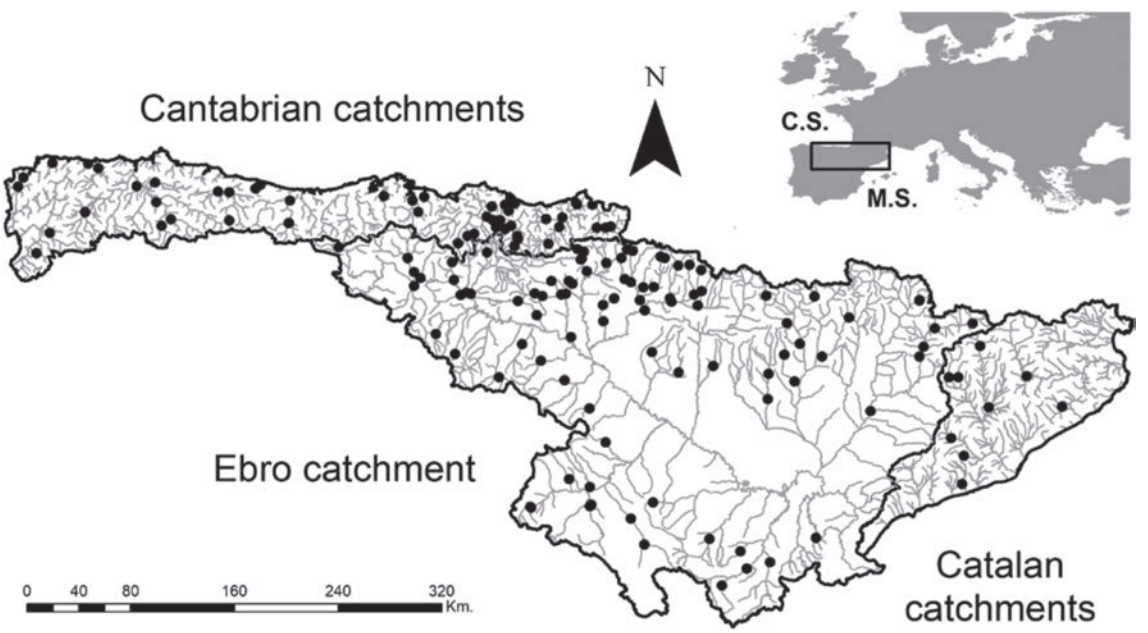

Figure 1. Map of unregulated gauges (●; n=156) in the study area. Black lines divide the Cantabrian, Ebro and Catalan catchments. (CS: Cantabrian Sea; MS: Mediterranean Sea). *Mapa de aforos no regulados en el área de estudio (●; n=156). Las líneas negras dividen las cuencas del Cantábrico, del Ebro y de Cataluña (CS: Mar Cántabro; MS: Mar Mediterráneo).*