Analysing the Capability of the Catchment's Spectral Signature for the Regionalization of Hydrological Parameters

Laura Fragoso-Campón [1], Pablo Durán-Barroso [2] and Elia Quirós [3]

[1] Department of Graphic Expression, Universidad de Extremadura, Cáceres, 10003, Spain

[2] Department of Construction, Universidad de Extremadura, Cáceres, 10003, Spain

[3] Department of Graphic Expression, Universidad de Extremadura, Cáceres, 10003, Spain

Corresponding Author: Laura Fragoso-Campón. Universidad de Extremadura, Escuela Politécnica, Avda. de la Universidad s/n, Cáceres, 10003, Spain. E-mail: laurafragoso@unex.es

Abstract

Water resource management in ungauged catchments is complex due to the uncertainties around the hydrological parameters that dominate the streamflow behaviour. These parameters are usually defined by regionalization approaches in which hydrological response patterns are transferred from gauged to ungauged basins. Regression-based methods using physical properties derived from cartographic data sources are widely used. The current remote sensing techniques offer us new standpoints in regionalisation processing since the hydrological response depends on the physical attributes related to the spectral responses of the territory. Moreover, machine learning approaches have not been specifically applied to the regionalization of hydrologic parameters. This work studies the capability of a catchment's spectral response based on Sentinel-1 and Sentinel-2 data to address a regression-based regionalization of hydrological parameters using a machine learning approach. Hydrological modelling was conducted by the HBV-light model. We tested the random forest algorithm in several regionalization scenarios: the new approach using the catchments' spectral signature, the traditional method using physical properties and a fusion of them. The calibration results were excellent (median KGE = 0.83), and the regionalized parameters obtained with the random forest algorithm achieved good performance in which the three scenarios showed almost the same goodness of fit (median KGE = 0.45 to 0.50). We found that the effectiveness depends on the climatic environment and that predictions in humid catchments exhibited better performance than those in the driest catchments. The physical approach (median KGE= 0.71) exhibited better performance than the spectral approach (median KGE= 0.64) in humid catchments, whereas spectral regionalization (median KGE= 0.33) outperformed the physical scenario in the driest catchments (median KGE= 0.25). Herein, our results confirm that regionalization is still challenging in Mediterranean climate variants where the new spectral approach showed promising results and time series of satellite data could improve seasonal regionalization methodologies.

Keywords: HBV, hydrology, random forest, runoff, SAR, Sentinel.

1. INTRODUCTION

Water resource management requires accurate estimations of all hydrological processes involved in water supply, flood and drought evaluations and eco-hydrology (Cui et al., 2020; Hrachowitz et al., 2013). The complexity of hydrological modelling relies on the availability of input data and on the uncertainties of the parameters that must be calibrated to obtain good accuracy in streamflow simulations (Beck et al., 2020). Moreover, calibration can lead to non-unique combinations of best parameters (Bárdossy, 2007). Therefore, the uncertainty is even greater in ungauged catchments where calibration of model parameters is not feasible, thus becoming a challenge for hydrologists.

For ungauged catchments, the parameter regionalization approach is used for transferring hydrological response patterns from gauged to ungauged (Hrachowitz et al., 2013), and there are several regionalization methods (spatial proximity, similarity-based, regression-based, and hydrological signatures-based) that have been proven to be effective, alone or in combination, in previous studies (see reviews by Hrachowitz et al. (2013), by J Parajka et al. (2013) and updated review by Guo, Zhang, Zhang, and Wang (2021). Although there are several worldwide studies, there is no optimal regionalization model; furthermore, the effectiveness of the regionalization will depend on the environment and the particular hydroclimatic attributes of an area (J Parajka et al., 2013).

Specifically, for regression-based parameter regionalization, the widest models used are linear or multiple regression (Merz & Blöschl, 2004; Jan Seibert, 1999) or transfer functions (Beck et al., 2020; Götzinger & Bárdossy, 2007; Hundecha & Bárdossy, 2004). Nevertheless, the transferability of parameters does not always lead to satisfactory simulation results due to nonlinear functions between the heterogeneous catchments and their flow regimen (Bárdossy, 2007). In the last decade, machine learning methods have been proven to be effective in analysing nonlinear relationships in many geosciences areas, particularly in water science areas (see review by Tyralis, Papacharalampous, and Langousis (2019)), but only a few studies have used machine learning for the regionalization approach of streamflow simulations (Buzacott, Tran, van Ogtrop, & Vervoort, 2019; Kult, Fry, Gronewold, & Choi, 2014; Prieto Sierra, Le Vine, Kavetski, García Alonso, & Medina Santamaría, 2019; Snelder et al., 2009; Zhang, Chiew, Li, & Post, 2018). Consequently, to the best of our knowledge, there is a lack of studies applying machine learning algorithms specifically to the regionalization of hydrologic parameters.

Related to the physical properties, topographic, land use, soil, and geological data from cartographic sources are traditionally the most commonly used for catchment characterization for parameter regionalization (Booij, 2005; Hundecha & Bárdossy, 2004; Kult, Fry, Gronewold, & Choi, 2014; Merz & Blöschl, 2004; Parajuli, Jayakody, & Ouyang, 2018). Recent studies have incorporated remote sensing data as predictors, such as those of Choubin et al. (2019), who used Moderate Resolution Imaging Spectroradiometer (MODIS) images and some derived products, such as vegetation and biophysical indices, to estimate streamflow in ungauged catchments. Landsat-8-derived indices have been used for flood susceptibility mapping (Bui et al., 2020), and Beck et al. (2020) used the Mean Normalized Difference Vegetation Index (NDVI) from the SPOT vegetation program as a predictor in their global regionalization of hydrologic parameters. However, apart from the vegetation indices, remote sensing data can also add a valuable characterization of the catchments because the spectral response will depend on the vegetation (type and the cover factor) and on the geological and lithological characteristics of the study area. Moreover, the vegetation types are also related to the nature of the soil lithotypes (Costa, Santos, Melo, & Santos, 2017). Hence, the reflectance captured by the optical sensor is related to the surface mineralogy depending on the composition of the soil of the geological formations of the area (Rajendran & Nasir, 2021). Moreover, the response of synthetic aperture radar (SAR) backscatter intensity is affected by surface roughness, soil moisture content (Purinton & Bookhagen, 2020), dielectric constant and grain size (Lu, Yang, & Meng, 2021). Consequently, considering that the abovementioned characteristics are directly involved in the hydrological response of the territory, the spectral response of a catchment can be related to its hydrological response. However, to our knowledge, the capability of the spectral signature to study catchment properties and to be used as predictors in regression-based regionalization of hydrological parameters has not yet been documented.

Therefore, the aim of our study was to explore the capability of a catchment's spectral response for a regression-based regionalization of hydrological parameters using a machine learning approach. The specific goals were i) to obtain the spectral signatures of the catchments based on the Sentinel satellites, ii) to characterize the physical soil properties based on the cartographic information of the European Soil Data Centre, iii) to test the random forest algorithm for the regression of the hydrological parameters, and iv) to assess the accuracy of the regionalization from two standpoints: spectral and physical properties.

2. DATA AND METHODS

2.1 Study area

The study area corresponds to 18 gauged watersheds throughout the Extremadura region in Spain (Figure 1). The criteria for the selection of the watersheds were based on the availability of gain station data measuring an upstream area below 1000 km² of mainly forested cover and not being controlled by the regulation of any dam or reservoir to study the natural regimen of the rivers. Table 1 shows the main climatic characteristics of the watersheds (Rivas-Martinez & Rivas-Saenz, 1996-2019), where the annual mean temperature and the mean annual precipitation range from 10 to 17 °C and from 446 to 1323 mm, respectively. Precipitation occurs mainly from October to April, while June, July and August suffer a significant drought with no or close to zero precipitation amount.

[Insert Figure 1]

[Insert Table 1]

2.2 Data Preprocessing

2.2.1 River discharge

The river discharge data were supplied by the Automatic Hydrological Information Systems (SAIH) of the hydrographic regions of *Tajo* and *Guadiana* to which the watersheds belong and which have had data recorded since they were put into operation in 2008. Therefore, the daily discharge from January 2008 to December 2019 was analysed. To ensure the quality of discharge observations, the hydrograph was analysed by visual inspection to guarantee the lack of gaps and outliers. Since the natural regimen of the rivers studied is intermittent, as recommended in Crochemore et al. (2020), it is important to distinguish between high-flow peaks instead of numerical outliers.

2.2.2 Precipitation and temperature

In this work, the daily gridded dataset of temperatures (maximum and minimum) and 24 h accumulated precipitation developed by the Spanish Meteorological Agency (AEMET) available in AEMET (2019) was used. These grided datasets were obtained through a statistical interpolation analysis of ground observation stations, with a spatial resolution of 0.05° in a rotated grid (CORDEX compliant) based on HIRLAM-AEMET Numerical Weather Prediction operational analyses (AEMET, 2017). To obtain the daily series in each watershed, we first reprojected the watershed into the rotated grid, and then, the average daily values in each watershed were processed. Considering that the interpolation of these data refers to the location of the ground observation stations, to analyse variations in temperature and precipitation with height, the differences in the elevation to the reference ground data and the mean distance to the nearest stations were analysed for each catchment.

2.2.3 Potential Evapotranspiration

The daily potential evapotranspiration was calculated using the 1985 Hargreaves ET_o equation since it only requires measured temperature data (see equation 8 in Hargreaves, ASCE, and Allen (2003)). The extraterrestrial radiation (R_a) values for the different latitudes of the watersheds were estimated following the equations in Allen, Pereira, Raes, and Smith (1998). Finally, the long-term monthly mean values were obtained from the daily values of potential evapotranspiration from January 2008 to December 2019.

2.2.4 Land Cover and Topographic data

The land cover map used in this work (Figure 1) was developed in a previous study where hydrological land cover categories were mapped based on the runoff generation capability of the land cover types using remote sensing techniques (Fragoso-Campón, Quirós, & Gutiérrez Gallego, 2021) and can be divided into three groups: forested, agricultural and impervious cover. Forested land cover categories comprise evergreen forest, deciduous forest, *dehesas*, shrubs, and herbaceous vegetation. The *dehesa* is a typical cover from Extremadura (Fragoso-Campón, Quirós, & Gutiérrez Gallego, 2020) defined by Devesa Alcaraz (1995) as "*pasturelands populated by holm and/or cork oaks, with an understorey of open grassland, cereal crops, or Mediterranean scrub*". The agricultural land cover categories are rainfed crops (mainly olive trees, vineyards and cold-season annual crops and cherry trees) and irrigated crops (seasonal warm crops). Finally, impervious surfaces such as rocky outcrops, bare soil, roads, and urban areas.

Digital elevation models (DEMs) at a spatial resolution of 25 m developed by the Spanish National Geographic Institute (IGN, 2021) were used for topographic data. Table 5 shows the properties considered as predictors in this study.

2.2.5 Soil Data

The spatial distribution of soil characteristics was obtained from the European Soil Data Centre (Joint Research Centre, 2020). Here, two different sources were used: we considered the information in the European Soil Database v2.0 (Panagos, 2006) for the characterization of the deeper groundwater response, whereas the updated topsoil and the physical properties for Europe developed in Ballabio, Panagos, and Monatanarella (2016) for the shallow groundwater response. Table 5 shows the physical soil properties considered a predictor in this study.

2.2.6 Sentinel Data

The spectral response of the catchments was studied using images collected from the Sentinel-1 (S1) and Sentinel-2 (S2) missions of the Copernicus Program. S1 is a C-band SAR sensor with dual VV and VH polarization, and S2 is a multispectral sensor working in the visible (VIS), near-infrared (NIR) and shortwave infrared (SWIR) bands. The analysis was addressed in multidate format: one S2 image acquired in summer when the soil and vegetation shows the best spectral separability in the Mediterranean environment (Fragoso-Campón, Quirós, Mora, Gutiérrez Gallego, & Durán-Barroso, 2020; Godinho, Guiomar, & Gil, 2017) and two S1 images, one acquired in summer close to the S2 acquisition date and another S1 image in winter to capture seasonal phenological differences. The satellite images were preprocessed using the Sentinel Application Platform (SNAP) software developed by the European Space Agency (ESA). S1 scene preprocessing included calibration (radiometric normalization), terrain correction and speckle filtering. S2 preprocessing included resampling, reprojection and mosaic building (further information described in Fragoso-Campón et al. (2021)). In addition, for a better characterization of the watersheds, spectral-derived metrics (vegetation, water and soil indices) and texture metrics derived from the grey-level cooccurrence matrix (GLCM) were also used (Haralick & Shanmugam, 1973). These metrics have been proven to be very useful in previous works related to land cover

(Fragoso-Campón et al., 2021) and lithological analysis (Lu et al., 2021; Radford, Cracknell, Roach, & Cumming, 2018). A complete description of these metrics is shown in Table 5.

2.3 Hydrological model

2.3.1 Model description

Herein, hydrological modelling was conducted by a conceptual-continuous rainfall-runoff model at the catchment scale using the HBV-light version of the HBV model (Bergström, 1995) developed in J. Seibert and Vis (2012). The model includes different routines (snow, soil, groundwater and routing) and simulates catchment discharge based on time series of precipitation, temperature and potential evaporation data (detailed description of the software and formulation are included in J. Seibert and Vis (2012)). We studied the discharge at a daily time step for 11 years, dividing the time series into two periods: warm-up (January 2008 to September 2014) and simulation (October 2014 to December 2019).

The watersheds were analysed with a semidistributed approach by considering four elevation zones (counting quantile distribution of the DEM) and three vegetation classes. The vegetation categories were grouped according to their runoff capability (low-medium-high): vegetation type 1 (low) comprises evergreen forest, deciduous forest, *dehesas*, and shrubs; vegetation type 2 (medium) comprises herbaceous vegetation and agricultural land cover; and vegetation type 3 (high) comprises impervious surfaces.

We used a model structure of three groundwater (GW) boxes: storage in the soil top zone (STZ), storage in the soil upper zone (SUZ) and storage in the soil lower zone (SLZ). Both the STZ and SUZ boxes are distributed using a box for each elevation-vegetation unit (Figure 2). The parameters involved in the model are related to each routine, as shown in Table 2.

[Insert Figure 2]

[Insert Table 2]

2.3.2 Model Accuracy Evaluation

The accuracy evaluation of the models was made following the goodness of fit (GoF) functions implemented in HBV light (Table 3).

[Insert Table 3]

2.3.3 Model Calibration

The calibration of the model was addressed in two stages: the individual catchment approach and the cluster of catchments approach.

In the first stage, each catchment was examined individually to obtain the best combination of parameters in each catchment using the tools available in the HBV-light software: Monte Carlo simulations and genetic algorithm and Powell optimization (GAP). First, a Monte Carlo simulation of 50 000 runs was calculated. In each run, the parameter values were randomly chosen within the given range (Table 2), ensuring the lack of bias in the calibration procedure (Jan Seibert, 1999). Then, the results were analysed to set the optimum range of PCALT and TCALT parameters to minimize the water balance volume error. Then, another 50 000 Monte Carlo simulations were run again, and the simulation results were analysed considering the goodness of fit function results for each parameter, defining the optimum range to maximize the accuracy of the model. The last step consisted of the application of the GAP algorithm within the optimal boundaries of the parameters obtained in the previous steps, and as a result, the parameters were fine-tuned using Powell's quadratically convergent method as described in J. Seibert and Vis (2012). The best parameter value combination was calculated for an objective function using Reff, KGE and Reff Peak with weight values of 0.2, 0.6, and 0.2, respectively.

In the second stage, the best values were analysed from the cluster criteria standpoint. As mentioned in previous studies, calibration could lead to a nonunique combination of best parameters (Bárdossy, 2007; Hundecha & Bárdossy, 2004); then, if necessary, parameters out of the cluster trend were revaluated using the GAP algorithm. Cluster analysis was performed using the Ward hierarchical clustering method (Ward Jr, 1963) with Euclidean distances implemented in R Stats Package R Software (R-Core-Team, 2018). This cluster analysis grouped the watersheds that were maximally similar with respect to their characteristics considering the physical and spectral information proposed in this work. Formerly, for the cluster analysis, the parameters were scaled with a centering approach (subtracting the mean and dividing by the standard deviation).

As a result, the best combination of the parameters in each catchment was established (Best-Par). In addition, an evaluation of the parameter sensitivity was assessed in each catchment using Monte Carlo simulation of 10 000 runs, one simulation per parameter, using Best-Par combination, varying one parameter at a simulation and measuring the decrease in accuracy when changing the parameter value within the original range (Table 2). Hence, to analyse how well the models fit the best value of the calibration, the objective GoF values were calculated in each run and compared to the target objective value. For all the runs, the minimum, mean, median and maximum values were calculated in each simulation for each parameter.

2.4 Regionalization process

The regionalization process relies on the hypothesis that the hydrological response in watersheds with similar characteristics is meant to be similar. Therefore, once the Best-Par combination was obtained in each catchment, the next step was to establish the relationship of those parameters to the catchment characterization. In this work, we addressed the regionalization analysis considering two standpoints: on the one hand, the variables related to precipitation, temperature and snow routine, which will be named climatic regionalization (CR), based on the topographic situation and climatology, and, on the other hand, the parameters involved in soil, groundwater and routing routines, which will be named ground regionalization (GR), depending on the catchment's properties in terms of spectral profile, morphology, land cover and soil characteristics. In addition, for the GR, we considered three scenarios to characterize the catchment properties, as shown in Table 4. The complete description of the predictors used for the regression is shown in Table 5.

[Insert Table 4]

The regression analysis between parameters and characteristics in each scenario was conducted with the random forest (RF) algorithm (Breiman, 2001), which is a nonparametric machine learning method that has been proven effective working with mixed-origin-input predictors such as those used in this work. Here, the regressions were performed using the RandomForest R package (Liaw & Wiener, 2002) with a number of decision trees (Ntree) of 5 000. The number of variables to be selected when growing the trees (Mtry) was automatically trained by the algorithm for each scenario depending on the number of predictors in each scenario.

Due to the limited number of watersheds, we conducted an iterative regression adjustment with a one catchment out of bag approach, which consists of leaving out of the algorithm the characteristics of one of the catchments and running the RF algorithm to calculate the regionalized parameters (Reg-Par) considering the data of the others. After the 18 RF simulations, all the Reg-Pars were compared to the Best-Par, and the RMSE was calculated for each parameter. Finally, the accuracy assessment of the regionalization in the different scenarios was addressed by running the HBV model using the Reg-Par obtained for each catchment analysing the variations in the GoF functions when compared to the results of the Best-Par.

[Insert Table 5]

3. RESULTS

3.1 Model calibration

After the calibration process, the Best-Par combination was obtained for each catchment, and the main statistics are shown for in each cluster in Table 6. Figure 3 shows the spectral and physical density profiles of the catchments categorized into groups that are maximally similar with respect to their properties based on the dendrogram of the clustering analysis.

[Insert Figure 3]

[Insert Table 6]

The accuracy of the models achieved excellent results, as shown in Table 7, where the global median objective GoF function value was 0.73. The best performing clusters are numbers 2, 3 and 4, which perform better than the global mean value, while clusters 1 and 5 are slightly lower but still outstanding.

[Insert Table 7]

The sensitivity analysis of the parameters measured how well the models fitted to the objective target value in each catchment for the Monte Carlo simulation of 10000 runs. Therefore, values of adjustment of approximately 1 imply that the GoF is as good as the results achieved by Best-Par, so that the model is less sensitive to the parameter value.

Evaluating all the catchments together, the results showed that the model accuracy was most sensitive to variations in FC, UZL and K0 (Figure 4). In addition, differences in the sensitivity among the catchment clusters were observed. In this sense, cluster 1 and cluster 5 showed the highest sensitivity to BETA, FC and UZL, while cluster 3 showed high sensitivity to FC and PCALT. In contrast, the lowest sensitivities appeared for cluster 2 and cluster 4, which were sensitive to K0, PCALT and TCALT and to FC, UZL and K0, respectively.

[Insert Figure 4]

3.2 Regionalization of parameters

3.2.1 Random forest regressions

The regression analysis was conducted with the RF algorithm in an iterative regression adjustment with one catchment out of bag at a time. Table 8 shows the RMSE between Reg-Par and Best-Par, and Figure 5 shows the correlation in the most sensitive parameters of the model. The CR Reg-Par fit well to the Best-Par values, achieving good correlations and reasonable RMSE. The GR achieved almost the same performance in the three scenarios, where PGR achieved slightly better results in terms of both RMSE (Table 8) and Pearson correlation coefficient (Figure 5).

[Insert Table 8]

[Insert Figure 5]

The importance of each predictor in the regression algorithms was analysed in terms of the increase in the mean squared error when a predictor is randomly permuted (%IncMSE), and higher values suggest a more important role of the predictor in the regression (Figure 6). In the CR, the most influential predictor is a topographic measurement referring to differences in the height of the catchments to the ground-observation digital elevation model (Delta_sd) and the climatic variables (T_mean and P_mean). In the SGR, the more valuable predictors are mainly related to texture metrics (for both SAR and optical bands) and the blue and NIR bands, and VIs such as MSAVI2 and SAVI seem to have a high influence on the prediction. In the PGR, the topsoil properties are the most valuable predictors together with vegetation coverage. When all the GR predictors are combined in the GGR, the topsoil attributes remain valuable predictors, together with the texture-derived metrics from the SAR and NIR bands.

[Insert Figure 6]

3.2.2 Accuracy Assessment

Figure 7 shows the accuracy assessment of models in the different regionalization scenarios when the Reg-Par were used compared to results for Best-Par in each catchment where the value for a perfect fit is 1 and values below 0 indicate poor fit. Figure 8 represents the boxplot graphic of the variations in the GoF measurements in the scenarios for all the catchments and the behaviours grouped by clusters. The general trend was that the decrease in the GoF functions was higher for the GR than for the CR, and the three scenarios in GR achieved almost the same performance. The CR&GGR scenario achieved slightly lower accuracies. Against the general trend, the efficiency for log(Q) was highly improved in all GR scenarios. Regarding the differences among the groups, clusters 2 and 4 showed the best performance in the regionalization validation, whereas cluster 5 achieved the lowest accuracies. The SGR results were better in clusters 1, 3 and 5, followed by the PGR scenario. The GoF function with more differences is ReffPeak, mostly in clusters 1 and 5.

[Insert Figure 7]

[Insert Figure 8].

Finally, an example of observed and simulated streamflows in the scenarios are graphically presented in Figure 9 at the Angeles catchment.

[Insert Figure 9].

4. DISCUSSION

In this study, the capability of the catchment's spectral signature for the regionalization of hydrological parameters was analysed using a regression-based machine learning approach.

For that purpose, the spectral profiles of the catchments based on the Sentinel-1 and Sentinel-2 satellites were processed, as well as the physical characterization of land cover and morphological and soil properties. Herein, we test the random forest algorithm for addressing the regression of the hydrological parameters from different standpoints: on the one hand, the CR where the variables related to precipitation, temperature and snow routine were analysed, based on the topographic situation and climatology, and, on the other hand, the GR were the hydrological parameters involved in soil, groundwater and routing routines were studied considering the new approach using the catchments' spectral signature (SGR) exclusively, the classical method using physical properties (PGR), and the mixed using all the spectral and physical approach (GGR). The evaluation of the overall efficiency using the most common GoF functions showed very good performance in the calibration stage (Table 7) and was still outstanding using the regionalization results (Figure 9). The general trend was that the decrease in the GoF functions was higher for the GR than for the CR, and the three GR scenarios achieved almost the same performance.

It is worth noting that the lack of previous studies at similar latitudes and Mediterranean climate variants hinders the comparison of the effectiveness of our work; consequently, we compared it with previous studies using any regionalization approach.

Concerning our calibration results, referred to the validation period, which showed a median overall KGE value of 0.83 (ranging from 0.78 to 0.86 in the clusters) outperformed previous works, such as those reported by Beck et al. (2020) with a median KGE of approximately 0.3 referred for the arid climate group (visual interpretation of Figure 5 in their work), and our results are similar to those reported by Alfieri et al. (2020) for Europe (visual interpretation of Figure 6 in their work). From another common GoF measurement standpoint, Reff, also referred to as Nash-Sutcliffe efficiency (NSE), our results showed a median overall value of 0.68 (ranging from 0.63 to 0.79 in the clusters), in agreement with those reported by Merz and Blöschl (2004) (NSE = 0.63) in Austria and with those reported by Jin, Xu, Zhang, and Chen (2009) (Reff =0.78). In addition, our results are better than those reported by Götzinger and Bárdossy (2007) (NSE = 0.53) in Germany and similar to those reported by Juraj Parajka, Blöschl, and Merz (2007) (NSE=0.66-0.69) again in Austria. It is worth noting that our calibration was carried out based on the optimization of an objective function where Reff and KGE were weighted 0.20 and 0.60, respectively. We also considered Reff Peak weighted 0.20.

Regarding our regionalized results, the GGR scenario, obtained by the combination of the spectral and physical attributes, did not improve the regionalization accuracies in any cluster, perhaps because the spectral response of the terrain is conditioned by the physical characteristics of the territory and the sum of the predictors does not provide new information but rather redundant information. Specifically, comparing the effectiveness of SGR *versus* PGR, the median Reff value obtained in the SGR (0.54) was slightly lower than the results of the PGR (0.57). These values are in agreement with those reported by Merz and Blöschl (2004) (NSE = 0.56) in their best-regression-based scenario and with those reported by Götzinger and Bárdossy (2007) (NSE = 0.50) using a combined method (using the Lipschitz condition and a monotony condition) and better than those reported by Masih, Uhlenbrook, Maskey, and Ahmad (2010) with a median NSE = 0.47 (estimated from Table 7 in their study using the 7^{th} best ranked) in their flow-duration-curve-based regionalization approach. In contrast, our global results are worse than the results reported by Jin et al. (2009) (Reff = 0.72) in their proxy basin-based regionalization in a subtropical climate catchment. One of the main reasons could be the differences in climatic properties discussed below.

The clustering analysis showed different efficiencies depending on the climatic standpoint. In this sense, the prediction catchment classified as the Tocsm bioclimatic variant (clusters 2 to 4) exhibited better performance than that classified in the Mpc bioclimatic variant (clusters 1 and 5) in both calibration and regionalization stages (Figure 10). This is in agreement with the findings in the review by J Parajka et al. (2013), who indicated that the performance of runoff predictions tends to be lower in arid than in cold and humid regions, and with the finding by In Goswami, O'connor, and Bhattarai (2007), where the worst results were obtained in the driest catchment of their study. In the Mpc catchments, the median Reff value was slightly better for the SGR (0.42) than for the PGR (0.35), whereas in the Tocsm catchments, the median Reff values were similar for both scenarios (0.7). Considering this, the efficiencies of the SGR and PGR in the Tocsm catchments are in agreement with those abovementioned by Jin et al. (2009). Concerning the CR, we found a greater decrease in the model efficiency in the Tocsm catchments (median Reff=0.73 to 0.64; median KGE= 0.84 to 0.67, median Obj=0.79 to 0.62) than in the Mpc catchments (median Reff=0.64 to 0.63; median KGE= 0.78 to 0.77, median Obj= 0.69 to 0.69) and showed how the uncertainty in precipitation has more influence in wet catchments than in dry catchments, in agreement with the analysis by Pianosi and Wagener (2016).

[Insert Figure 10]

We compared our regionalization effectiveness with the performance of the global high-resolution regionalized parameters (GloH2O) dataset developed by Beck et al. (2020). For that, we processed the ten cross-validation folds of available model parameters (BETA, FC, K0, K1, K2, LP, PERC, UZL, TT, CFMAX, CFR, and CWH) in the GloH2O dataset, and the ensemble-mean value was averaged in each catchment. For the parameters not mentioned above (PCALT, TCALT, SFCF, MAXBAS and Cet), we assumed the Best-Par obtained in our calibration, and after that, the HBV models were run again for each catchment. Figure 11a shows that our regional approach outperformed the global GloH2O achievements in all GoFs analysed, and specifically the GoF reported in their study, our results (in terms of median values) SGR (KGE = 0.45), PGR and GGR (both KGE = 0.50) outperformed the global GloH2O achievements (KGE = 0.20). Nevertheless, the latter value is in agreement with their own results for the arid climate class (where it is classified our study area in Spain as shown in Figure 5 in Beck et al. (2020)) for which they informed a median KGE of approximately -0.05 for the calibration results (ranging from -0.20 in the 25th percentile to 0.15 in the 75th percentile). Therefore, given that in their global study, it appears that there are no study catchments in the Mediterranean area (see Figure 2 in Beck et al. (2020)), this finding would confirm that the GloH2O parameters in our study area achieved an effectiveness of approximately the same order of magnitude as those reported in the original study for arid climate classes. In addition, GloH2O achieved acceptable results in Tocsm and showed a median KGE value of approximately 0.28, whereas the results in Mpc were worse at approximately -0.14. The main reason for these differences in the model efficiency can be due to the differences in the estimation of the most sensitive parameters in the catchments under study. It was observed that GloH2O overestimated the FC and UZL values and underestimated the K0 values when compared to the Best-Par and Reg-Par obtained in our work (Figure 11b).

[Insert Figure 11]

Regarding the PGR approach, we used the soil information of the European Soil Data Centre that exhibited excellent performance. The results showed that soil information has a strong influence on the regression of parameters since the topsoil properties are the most valuable predictors in the PGR and remain in GGR as valuable predictors, together with the texture-derived metrics from SAR and NIR bands. The SGR showed better results than the PGR in the Mpc catchments, which are driest and characterized by medium to low vegetation coverage. This might be because the more valuable predictors in SGR are mainly related to texture metrics for both SAR and optical bands. The SAR texture metrics are influenced by the grain size of the surface (Lu et al., 2021), and the radar capability for lithological analysis is better in sparse coverage than in dense vegetation areas (Radford et al., 2018). In addition, the optical information of the blue and NIR bands resulted as valuable predictors, as these latter bands were proven to be useful for geological applications in arid regions (Rajendran & Nasir, 2021).

Therefore, it is worth noting that effectiveness in hydrological models highly depends on the climatic environment, and when comparing results with other studies, it seems to be highly recommended to analyse the findings from the climatic rather than from the methodological standpoint. Herein, our results confirm that regionalization is still challenging in Mediterranean bioclimate variants where the new spectral approach SGR showed promising results. One of the key issues in the regionalization of hydrological parameters is the availability of attributes to be considered predictors (Merz & Blöschl, 2004) and its variability in nomenclature across regions or countries. Therefore, having continuous soil data throughout Europe, together with Sentinel information, offers us new opportunities in the regionalization of parameters at the European scale. As mentioned above, the previously studied sites are mainly concentrated in central and northern Europe: Austria (Merz & Blöschl, 2004; Juraj Parajka et al., 2007), Germany (Bárdossy, 2007; Götzinger & Bárdossy, 2007), and Iran (Masih et al., 2010), and there is an absolute lack of studies on catchments in the Mediterranean environment. Future work could focus on regionalization in the Mediterranean area to complete the range of catchment typologies. Moreover, Sentinel data offer continuous coverage throughout almost all over the world, offering new possibilities in areas where cartographic information is not available. In addition, time series of satellite data could improve seasonal regionalization approaches characterizing the hydrologic response of the catchments and the seasonal variability that characterizes Mediterranean catchment flood events.

5. CONCLUSIONS

In this study, the capability of the catchment's spectral signature for the regionalization of hydrological parameters was studied using a regression-based machine learning approach. The calibration results were excellent (median KGE = 0.83), and regionalized parameters with the random forest algorithm achieved good performance. The general trend showed that the decrease of the efficiency was lower for the climate regionalization (variables related to precipitation, temperature and snow routine) (median KGE = 0.74) than for the ground regionalization (hydrological parameters involved in soil, groundwater and routing routines), where the three scenarios achieved almost the same performance: the new approach using the catchments' spectral signature from the Sentinel-1 and Sentinel-2 satellites (median KGE = 0.45), the traditional method using physical properties (data provided by the European Soil Data Centre) (median KGE = 0.50) and a fusion of spectral and physical properties (median KGE = 0.50). We found that effectiveness in the hydrological models highly depends on the climatic environment, and the prediction in catchments classified as temperate oceanic sub-Mediterranean (Tocsm) bioclimatic variants exhibited better performance than that classified in the Mediterranean pluviseasonal-continental (Mpc) variant. The physical approach, using the soil information of the European Soil Data Centre, exhibited a brilliant performance, and the spectral approach showed better results in the driest catchments in the Mpc climatic variant. Herein, our results confirm that regionalization is still challenging in Mediterranean bioclimate variants where the new spectral approach showed promising results. Therefore, having continuous soil data throughout Europe, together with Sentinel information, offers us new opportunities in the regionalization of parameters at the European scale, especially to fill the gap of regionalization studies in the Mediterranean environment. Moreover, the continuous coverage of Sentinel data worldwide offers new possibilities in areas where cartographic information is unavailable. In addition, time series of satellite data can improve seasonal regionalization approaches characterizing the hydrologic response of the catchments and its seasonal variability.

ACKNOWLEDGEMENTS

This research was funded by the *Consejería de Economía, Ciencia y Agenda Digital* of *Junta de Extremadura* and the European Social Fund: A way of doing Europe, through the "Financing of Predoctoral Contracts for the Training of Doctors in Public Research and Development Centers belonging to the Extremadura System of Science, Technology, and Innovation [file PD16018]."

This work was also supported by the *Consejería de Economía, Ciencia y Agenda Digital of Junta de Extremadura* (Spain) and co-funded by the European Regional Development Fund under Grants GR18052 (DESOSTE), GR18028 (KRAKEN) and, also supported by *Universidad de Extremadura* under Grant 2021-1852X2-CONVENIO 025Ñ16.

We thank *Confederación Hidrográfica del Guadiana* and *Tajo* for providing the river discharge series. We thank the AEMET for providing the climatic information. We thank the European Soil Data Centre (ESDAC) for providing the data about the topsoil physical properties for Europe and the European Soil Database v2.0. We thank the University of Zurich (Department of Geography) for providing us with a copy of the HBV-light software.

DATA AVAILABILITY

The random-forest regression models developed in this study are available from the corresponding author upon reasonable request.

REFERENCES

AEMET. (2017). Serie de precipitación diaria en rejilla con fines climáticos. Nota técnica 24 de AEMET. Retrieved June, 2019, from https://www.aemet.es/es/conocermas/recursos_en_linea/publicaciones_y_estudios/publicaciones/detalles/NT_24_AEMET

AEMET. (2019). Datos observacionales Precipitación y Temperatura. Retrieved December, 2020, from http://www.aemet.es/es/serviciosclimaticos/cambio_climat/datos_diarios?w=2

Alfieri, Lorenzo, Lorini, Valerio, Hirpa, Feyera A, Harrigan, Shaun, Zsoter, Ervin, Prudhomme, Christel, & Salamon, Peter. (2020). A global streamflow reanalysis for 1980–2018. Journal of Hydrology X, 6, 100049.

Allen, Richard G, Pereira, Luis S., Raes, Dirk , & Smith, Martin (1998). Chapter 3 - Meteorological data. In F.-F. a. A. O. o. t. U. Nations (Ed.), Crop evapotranspiration -

Guidelines for computing crop water requirements - FAO Irrigation and drainage paper 56. Rome: FAO - Food and Agriculture Organization of the United Nations.

Ballabio, Cristiano, Panagos, Panos, & Monatanarella, Luca. (2016). Mapping topsoil physical properties at European scale using the LUCAS database. Geoderma, 261, 110-123. doi: 10.1016/j.geoderma.2015.07.006

Bárdossy, A. (2007). Calibration of hydrological model parameters for ungauged catchments. Hydrology and Earth System Sciences, 11(2), 703-710. doi: 10.5194/hess-11-703-2007

Beck, Hylke E, Pan, Ming, Lin, Peirong, Seibert, Jan, van Dijk, Albert IJM, & Wood, Eric F. (2020). Global fully distributed parameter regionalization based on observed streamflow from 4,229 headwater catchments. Journal of Geophysical Research: Atmospheres, 125(17), e2019JD031485. doi: 10.1029/2019JD031485

Bergström, Sten. (1995). The HBV model. Computer models of watershed hydrology., 443-476.

Breiman, Leo. (2001). Random forests. Machine learning, 45(1), 5-32. doi: 10.1023/A:1010933404324

Bui, Quang-Thanh, Nguyen, Quoc-Huy, Nguyen, Xuan Linh, Pham, Vu Dong, Nguyen, Huu Duy, & Pham, Van-Manh. (2020). Verification of novel integrations of swarm intelligence algorithms into deep learning neural network for flood susceptibility mapping. Journal of Hydrology, 581, 124379. doi: 10.1016/j.jhydrol.2019.124379

Buzacott, Alexander JV, Tran, Bruce, van Ogtrop, Floris F, & Vervoort, R Willem. (2019). Conceptual Models and Calibration Performance—Investigating Catchment Bias. Water, 11(11), 2424. doi: 10.3390/w11112424

Choubin, Bahram, Solaimani, Karim, Rezanezhad, Fereidoun, Roshan, Mahmoud Habibnejad , Malekian, Arash, & Shamshirband, Shahaboddin. (2019). Streamflow regionalization using a similarity approach in ungauged basins: Application of the geo-environmental signatures in the Karkheh River Basin, Iran. Catena, 182, 104128. doi: j.catena.2019.104128

Costa, Silas, Santos, Vagner, Melo, Danilo, & Santos, Pablo. (2017). Evaluation of Landsat 8 and Sentinel-2A data on the correlation between geological mapping and NDVI. Paper presented at the 2017 First IEEE International Symposium of Geoscience and Remote Sensing (GRSS-CHILE).

Crochemore, Louise, Isberg, Kristina, Pimentel, Rafael, Pineda, Luis, Hasan, Abdulghani, & Arheimer, Berit. (2020). Lessons learnt from checking the quality of openly accessible river flow data worldwide. Hydrological Sciences Journal, 65(5), 699-711. doi: 10.1080/02626667.2019.1659509

Cui, Zhen, Huang, Jiacong, Tian, Feng, Gao, Junfeng, Wang, Xuesong, & Li, Jun. (2020). Quantifying the impacts of climate change and land use on hydrological processes: A comparison between mountain and lowland agricultural watersheds. Hydrological Processes. doi: 10.1002/hyp.13950

Devesa Alcaraz, Juan Antonio. (1995). Vegetación y flora de Extremadura. Badajoz: Universitas Editorial.

Fragoso-Campón, Laura, Quirós, Elia, & Gutiérrez Gallego, José Antonio (2020). Dehesa environment mapping with transference of a Random Forest classifier to neighboring ultra-high spatial resolution imagery at class and macro-class land cover levels. Stochastic Environmental Research and Risk Assessment, 1-32. doi: 10.1007/s00477-020-01880-3

Fragoso-Campón, Laura, Quirós, Elia, & Gutiérrez Gallego, José Antonio (2021). Optimization of land cover mapping through improvements in Sentinel-1 and Sentinel-2 image dimensionality and data mining feature selection for hydrological modeling. Stochastic Environmental Research and Risk Assessment. doi: 10.1007/s00477-021-02014-z

Fragoso-Campón, Laura, Quirós, Elia, Mora, Julián, Gutiérrez Gallego, José Antonio, & Durán-Barroso, Pablo. (2020). Overstory-understory land cover mapping at the watershed scale: accuracy enhancement by multitemporal remote sensing analysis and LiDAR. Environmental Science and Pollution Research, 27(1), 75-88. doi: 10.1007/s11356-019-04520-8

Godinho, Sérgio, Guiomar, Nuno, & Gil, Artur. (2017). Estimating tree canopy cover percentage in a mediterranean silvopastoral systems using Sentinel-2A imagery and the stochastic gradient boosting algorithm. International Journal of Remote Sensing, 1-23. doi: 10.1080/01431161.2017.1399480

Goswami, M, O'connor, KM, & Bhattarai, KP. (2007). Development of regionalization procedures using a multi-model approach for flow simulation in an ungauged catchment. Journal of Hydrology, 333(2-4), 517-531.

Götzinger, J, & Bárdossy, A. (2007). Comparison of four regionalization methods for a distributed hydrological model. Journal of Hydrology, 333(2-4), 374-384. doi: 10.1016/j.jhydrol.2006.09.008

Guo, Yuhan, Zhang, Yongqiang, Zhang, Lu, & Wang, Zhonggen. (2021). Regionalization of hydrological modeling for predicting streamflow in ungauged catchments: A comprehensive review. Wiley Interdisciplinary Reviews: Water, e1487. doi: 10.1002/wat2.1487

Haralick, Robert M, & Shanmugam, Karthikeyan. (1973). Textural features for image classification. IEEE Transactions on systems, man, and cybernetics(6), 610-621. doi: 10.1109/TSMC.1973.4309314

Hargreaves, George H, ASCE, F, & Allen, Richard G. (2003). History and Evaluation of Hargreaves Evapotranspiration equation. Journal of Irrigation and Drainage Engineering, 129(1), 53-63. doi: 10.1061/(ASCE)0733-9437(2003)129:1(53)

Hrachowitz, Markus, Savenije, HHG, Blöschl, G, McDonnell, JJ, Sivapalan, M, Pomeroy, JW, . . . Cudennec, C. (2013). A decade of Predictions in Ungauged Basins (PUB)—a review. Hydrological sciences journal, 58(6), 1198-1255. doi: 10.1080/02626667.2013.803183

Hundecha, Yeshewatesfa, & Bárdossy, András. (2004). Modeling of the effect of land use changes on the runoff generation of a river basin through parameter regionalization of a watershed model. Journal of hydrology, 292(1-4), 281-295. doi: j.jhydrol.2004.01.002

IGN. (2021). Centro de Descargas.

Centro Nacional de Información Geográfica. Retrieved April, 2017, from http://centrodedescargas.cnig.es/CentroDescargas/index.jsp

Jin, Xiaoli, Xu, Chong-yu, Zhang, Qi, & Chen, Yongqin David. (2009). Regionalization study of a conceptual hydrological model in Dongjiang basin, south China. Quaternary International, 208(1-2), 129-137. doi: 10.1016/j.quaint.2008.08.006

Joint Research Centre. (2020). European Soil Data Centre (ESDAC). Retrieved 20/07/2020, from https://esdac.jrc.ec.europa.eu/

Kult, Jonathan M, Fry, Lauren M, Gronewold, Andrew D, & Choi, Woonsup. (2014). Regionalization of hydrologic response in the Great Lakes basin: Considerations of temporal scales of analysis. Journal of Hydrology, 519, 2224-2237. doi: 10.1016/j.jhydrol.2014.09.083

Liaw, Andy, & Wiener, Matthew. (2002). Classification and regression by randomForest. R News, 2/3.

Lu, Yi, Yang, Changbao, & Meng, Zhiguo. (2021). Lithology Discrimination Using Sentinel-1 Dual-Pol Data and SRTM Data. Remote Sensing, 13(7), 1280. doi: 10.3390/rs13071280

Masih, Ilyas, Uhlenbrook, Stefan, Maskey, Shreedhar, & Ahmad, MD. (2010). Regionalization of a conceptual rainfall–runoff model based on similarity of the flow duration curve: A case study from the semi-arid Karkheh basin, Iran. Journal of hydrology, 391(1-2), 188-201. doi: 10.1016/j.jhydrol.2010.07.018

Merz, Ralf, & Blöschl, Günter. (2004). Regionalisation of catchment model parameters. Journal of hydrology, 287(1-4), 95-123. doi: 10.1016/j.jhydrol.2003.09.028

Panagos, Panos. (2006). The European soil database. GEO: connexion, 5(7), 32-33.

Parajka, J, Viglione, A, Rogger, M, Salinas, JL, Sivapalan, M, & Blöschl, G. (2013). Comparative assessment of predictions in ungauged basins–Part 1: Runoff-hydrograph studies. Hydrology and Earth System Sciences, 17(5), 1783-1795. doi: 10.5194/hess-17-1783-2013

Parajka, Juraj, Blöschl, Günter, & Merz, Ralf. (2007). Regional calibration of catchment models: Potential for ungauged catchments. Water Resources Research, 43(6). doi: 10.1029/2006WR005271

Pianosi, Francesca, & Wagener, Thorsten. (2016). Understanding the time−varying importance of different uncertainty sources in hydrological modelling using global sensitivity analysis. Hydrological Processes, 30(22), 3991-4003. doi: 10.1002/hyp.10968

Prieto Sierra, Cristina, Le Vine, Nataliya, Kavetski, Dmitri, García Alonso, Eduardo, & Medina Santamaría, Raúl. (2019). Flow prediction in ungauged catchments using probabilistic random forests regionalization and new statistical adequacy tests. doi: 10.1029/2018WR023254

Purinton, Benjamin, & Bookhagen, Bodo. (2020). Multiband (X, C, L) radar amplitude analysis for a mixed sand-and gravel-bed river in the eastern Central Andes. Remote Sensing of Environment, 246, 111799. doi: 10.1016/j.rse.2020.111799

R-Core-Team. (2018). R: A language and environment for statistical computing. Vienna (Austria): R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Radford, Declan DG, Cracknell, Matthew J, Roach, Michael J, & Cumming, Grace V. (2018). Geological mapping in western Tasmania using radar and random forests. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 11(9), 3075-3087. doi: 10.1109/JSTARS.2018.2855207

Rajendran, Sankaran, & Nasir, Sobhi. (2021). ASTER mapping of gypsum deposits of Thumrait region of southern Oman. Resource Geology. doi: 10.1111/rge.12245

Rivas-Martinez, S, & Rivas-Saenz, S. (1996-2019). Worldwide Bioclimatic Classification System. Retrieved July, 2019, from http://www.globalbioclimatics.org/form/tb_map/index.htm

Seibert, J., & Vis, M. J. P. (2012). Teaching hydrological modeling with a user-friendly catchment-runoff-model software package. Hydrol. Earth Syst. Sci., 16(9), 3315-3325. doi: 10.5194/hess-16-3315-2012

Seibert, Jan. (1999). Regionalization of parameters for a conceptual rainfall-runoff model. Agricultural and forest meteorology, 98, 279-293. doi: 10.1016/S0168-1923(99)00105-7

Snelder, Ton H, Lamouroux, Nicolas, Leathwick, John R, Pella, Hervé, Sauquet, Eric, & Shankar, Ude. (2009). Predictive mapping of the natural flow regimes of France. Journal of Hydrology, 373(1-2), 57-67. doi: 10.1016/j.jhydrol.2009.04.011

Tyralis, Hristos, Papacharalampous, Georgia, & Langousis, Andreas. (2019). A brief review of random forests for water scientists and practitioners and their recent history in water resources. Water, 11(5), 910. doi: 10.3390/w11050910

Ward Jr, Joe H. (1963). Hierarchical grouping to optimize an objective function. Journal of the American statistical association, 58(301), 236-244. doi: 10.1080/01621459.1963.10500845

Zhang, Yongqiang, Chiew, Francis HS, Li, Ming, & Post, David. (2018). Predicting runoff signatures using regression and hydrological modeling approaches. Water Resources Research, 54(10), 7859-7878. doi: 10.1029/2018WR023325