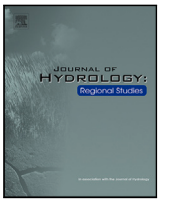
Journal of Hydrology: Regional Studies 47 (2023) 101427

Contents lists available at ScienceDirect

Journal of Hydrology: Regional Studies

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



Regional streamflow prediction in northwest Spain: A comparative analysis of regionalisation schemes

Juan F. Farfán *, Luis Cea

Universidade da Coruña, Water and Environmental Engineering Group, Center for Technological Innovation in Construction and Civil Engineering (CITEEC), Elviña, 15071 A Coruña, Spain

ARTICLE INFO

Dataset link: https://www.meteogalicia.gal/, https://augasdegalicia.xunta.gal/

Keywords:
Ungauged basin
Regionalisation
Spatial proximity
Physical similarity
Artificial Neural Networks
Hydrological model

ABSTRACT

Study Region: The present study was conducted in 24 watersheds located in the region of Galicia, in the northwest of Spain, covering an extension of approximately 13,000 km^2.

Study focus: This study is focused on the application and evaluation of different schemes for streamflow Prediction in Ungauged Basins (PUB). The MHIA model (Spanish acronym for *Modelo HIdrológico Agregado*), is first used to reproduce the observed time series of discharge in several gauged basins. Then, six different regionalisation schemes are applied to transfer the hydrological model parameters to ungauged catchments. For that purpose, we explore and compare two physical similarity, two spatial proximity and two regression-based regionalisation schemes. Output averaging (also known as ensemble modelling) as well as parameter averaging implementations of the physical similarity and spatial proximity methods are analysed.

New hydrological insights: The most efficient methods are those based on output averaging, with acceptable success rates (SR) in 88% of the cases. On the other hand, the parameter averaging-based methods have the lowest SR. The methods based on spatial proximity output averaging provide the best performance when the receptor basin has a sufficient number of nearby donor basins. On the other hand, the methods based on physical similarity output averaging show a better performance in areas where there is a low density of donor catchments. The regression-based methods showed the lowest performance in all cases. The existence of correlations between the performance of the regionalisation schemes and the area of the receptor catchments was observed, with higher performances in large basins than in small basins.

1. Introduction

The successful application of hydrological models as water resources management tools requires accurate and comprehensive input data, as well as observed discharge series to calibrate the model. However, the availability of streamflow data in some catchments is very scarce, precluding the calibration of the models and leading to the development of methods for streamflow prediction in ungauged basins (PUB) (Guo et al., 2021; Tarek et al., 2021).

Different methodologies have been developed over the last two decades to transfer information from gauged (donor) to ungauged (receptor) basins, which are often referred to as regionalisation methods (Sivapalan et al., 2003; Prieto et al., 2019; Kratzert et al., 2019; Prieto et al., 2022). Despite the efforts made in previous studies, there is not a unique generally accepted method for streamflow simulation in ungauged catchments, and several factors are known to affect the performance of the regionalisation

* Corresponding author.
E-mail address: j.farfan@udc.es (J.F. Farfán).

https://doi.org/10.1016/j.ejrh.2023.101427
Received 6 December 2022; Received in revised form 30 March 2023; Accepted 15 May 2023
Available online 2 June 2023
2214-5818/© 2023 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

methods, such as regional climate, the physical characteristics of the basins and the existence of human interventions (Oudin et al., 2008; Swain and Patra, 2017; Arsenault and Brissette, 2014; Parajka et al., 2013; Razavi and Coulibaly, 2013).

Guo et al. (2021) emphasise that there are different ways of categorising methods for predicting streamflow in ungauged catchments, with parameter transfer-based and signature transfer-based methods being among these classifications. The methods based on parameter transfer implement parameter regionalisation using various catchment similarity measures, which can be categorised into three groups: physical similarity methods, spatial proximity methods, and regression-based methods.

Physical similarity methods (Burn and Boorman, 1993; Kay et al., 2007; He et al., 2011) focus on the similarity of catchment properties and characteristics, while spatial proximity methods (Cislaghi et al., 2020; Merz and Blöschl, 2004) consider the spatial relationships between catchments. Regression-based methods (Heuvelmans et al., 2006; Arsenault and Brissette, 2014) use statistical regression techniques to establish relationships between catchment attributes and model parameters.

Hydrological signature-based methods are commonly used to evaluate the hydrological characteristics of catchments at different time scales, and they are categorised into two groups: direct regionalisation of hydrological signatures and constraining model parameter sets using these signatures (Guo et al., 2021). Typical flow indices used in these methods include monthly and average flows, runoff coefficients, quantiles, slope of flow duration curves, and the baseflow index, as noted in studies by Guo et al. (2021), McMillan et al. (2017). While this work focuses on parameter transfer-based methods, readers are referred to the works of Kavetski et al. (2018), Westerberg et al. (2016), McMillan et al. (2017), Addor et al. (2018), Prieto et al. (2022) for a complete and detailed description of signature-based approaches.

Physical similarity methods assume that similar climatic and geomorphological characteristics result in similar hydrological responses (Merz and Blöschl, 2004). Spatial proximity methods assume that catchments in close proximity have similar hydrological behaviours (Cislaghi et al., 2020; Merz and Blöschl, 2004). Regression-based methods attempt to establish mathematical relationships between hydrological model parameters and catchment characteristics (Heuvelmans et al., 2006; Almeida et al., 2016; Guo et al., 2021).

Both physical similarity and regression-based methods require rigorous selection of the most relevant Catchment Descriptors (CD) - catchment geology, soil, land use, topography and climate (Prieto et al., 2022; Almeida et al., 2016; Tarek et al., 2021). These descriptors are typically defined based on expert knowledge of the study basin and subsequent correlation analysis (Heuvelmans et al., 2006; Mwakalila, 2003). In contrast, spatial proximity methods do not require CDs since model parameters are transferred from nearby basins using spatial interpolation techniques based on basin centroids (Cislaghi et al., 2020). This approach has shown to provide better results than physical similarity methods in areas with a sufficient number of donor catchments (Oudin et al., 2008; Cislaghi et al., 2020). However, the main limitation of spatial proximity methods is that they require a large number of gauged catchments surrounding the target basin Oudin et al. (2008), Razavi and Coulibaly (2013), Merz and Blöschl (2004), Shepard (1968), Oudin et al. (2008), Merz and Blöschl (2004), Ssegane et al. (2012), Samuel et al. (2011).

Another approach based on physical similarity or spatial proximity involves calculating the average hydrographs obtained from running the hydrological model with parameter sets of donor catchments (Oudin et al., 2008; Zhang and Chiew, 2009; Samuel et al., 2011; Arsenault and Brissette, 2014; Poissant et al., 2017). This output-averaging method is commonly referred to as ensemble modelling and has shown remarkable results in different studies (Viney et al., 2009; Farfán and Cea, 2022a; Najafi and Moradkhani, 2016; Arsenault and Brissette, 2014). Ensemble modelling has outperformed parameter-averaging methods in previous studies (Yang et al., 2020; Arsenault and Brissette, 2014; Poissant et al., 2017). Additionally, combinations of physical similarity and spatial proximity methods can be performed by incorporating information about the spatial location of the catchments as another Catchment Descriptor (CD) within a physical similarity scheme (e.g., the centroid coordinates of the donor catchments), which in some cases can enhance the results obtained with output-averaging methods (Arsenault and Brissette, 2014; Poissant et al., 2017).

In regression-based methods, a regression model is derived between the value of each model parameter and the CDs (Almeida et al., 2016; Heuvelmans et al., 2006; Arsenault and Brissette, 2014; Tarek et al., 2021; Swain and Patra, 2017). The regression is used to predict the model parameters in the ungauged basin, which are then used to run the hydrological model in the ungauged basin Tarek et al. (2021), Wang et al. (2014). To apply these methods, it is recommended to consider as many donor basins as possible to establish a robust regression (He et al., 2011). The most commonly applied techniques in this methodology are multiple linear regression and artificial neural networks (ANNs), with ANNs having the best results (Heuvelmans et al., 2006; Kokkonen et al., 2003; Post, 2009). It is important to note that the main assumption of physical similarity-based and regression-based methods, that catchments with similar CDs exhibit similar hydrological responses, may be compromised due to the effects of equifinality. Equifinality means that there are different parameter sets that produce equally acceptable results during model calibration processes (Beven, 2006; Arsenault and Brissette, 2014).

Given the challenge of characterising the influence of equifinality on the predictions obtained with regionalisation schemes, employing a bootstrapping method provides a valuable approach for evaluating the uncertainty related to this particular problem (Heuvelmans et al., 2006; Arsenault and Brissette, 2014; Poissant et al., 2017). Bootstrapping allows the evaluation of regionalisation schemes using a sample of behavioural parameter sets, aiding in the quantification of the range of possible outcomes and their associated uncertainties, thereby enhancing the reliability of the regionalisation process. Behavioural parameter sets are typically identified as those that produce model performance where the likelihood function value exceeds a pre-established acceptable threshold (Beven and Freer, 2001; Yin et al., 2020). For streamflow predictions in ungauged catchments, behavioural parameter sets are also obtained by selecting a fixed number of top-performing simulations from the entire sample of parameter sets of each donor basin Arsenault and Brissette (2014), Poissant et al. (2017).

This study aims to apply, evaluate, and compare different schemes for streamflow prediction in ungauged basins in northwest Spain using 24 watersheds. Climatic and geomorphological information from various sources is used to determine the Catchment

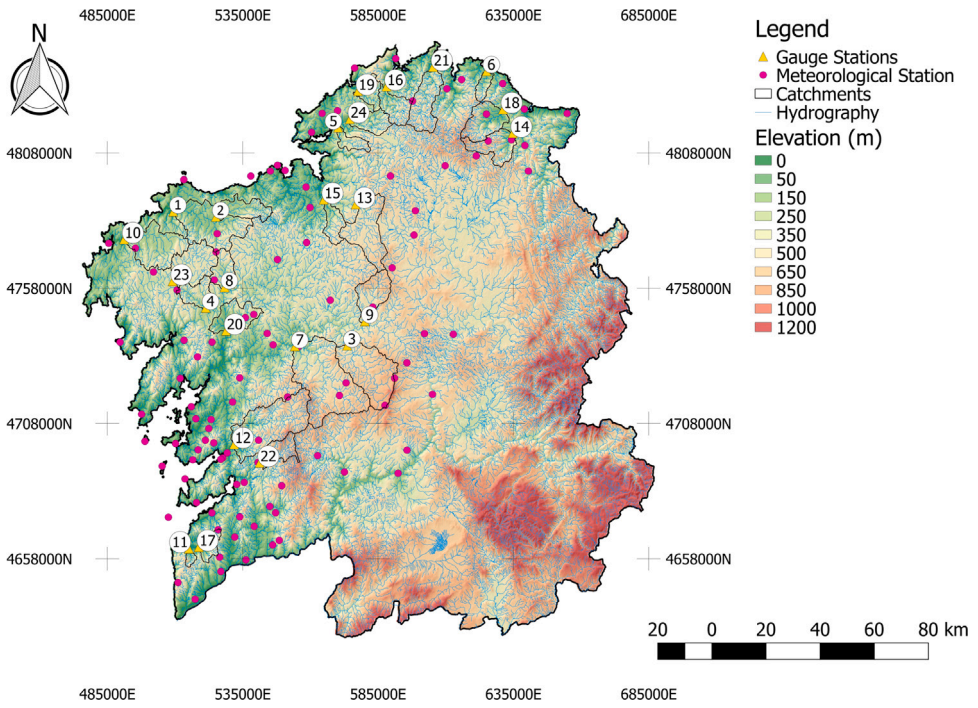


Fig. 1. Location of the 24 study catchments.

Descriptors (CDs) that are subsequently used in the regionalisation schemes. To this end, the methodologies that have provided the best results in previous studies are selected and applied to the regionalisation of streamflow series in the 24 basins of the study region (Arsenault and Brissette, 2014; Poissant et al., 2017; Cislaghi et al., 2020; Heuvelmans et al., 2006). A total of six regionalisation schemes are evaluated and organised as follows: two schemes based on spatial proximity, with an average of model parameters and an average of model output, two schemes based on physical similarity with an average of model parameters and an average of model output, and two regression-based methods, Multiple Linear Regression and Artificial Neural Networks. In order to characterise the uncertainty in the streamflow predictions, the transfer of parameters is carried out using a bootstrapping technique over a sample of model behavioural parameters. The validation of the schemes is developed using a leave-one-out process. The results are analysed in terms of the probability distribution of the goodness-of-fit, the Success Rate (SR) of the regionalisation schemes, and the spatial distribution of the results over the study region.

2. Study area and data

This section describes the data sources and the characteristics of the 24 basins included in this study. The catchments are located in Galicia, in the northwest of Spain. The available information was used to evaluate 25 potential CDs for each watershed, which were classified as: (1) morphological, (2) hydro-climatic and (3) lithological (Heuvelmans et al., 2006). The meteorological and discharge data used in this study were obtained from the agencies Meteo Galicia (https://www.meteogalicia.gal/) and Augas de Galicia (https://augasdegalicia.xunta.gal/), respectively. The data are pre-processed and undergoes filtering procedures by these agencies to ensure reliability before being made publicly available.

The 24 catchments cover a total area of approximately 13,000 km^2 (Fig. 1). The area of each individual catchment varies between 17 and 542 km^2, with elevations ranges from 16 to 406 meters above sea level (m.a.s.l). The minimum average slope of the catchments is 7% and the maximum 28.3%. The length of the main river channel varies between 9 and 62 km, and the drainage density from 0.57 to 1.72 km/km^2 (Table 1).

The coordinates of the basin centroids have also been incorporated as basin descriptors since, as reported in previous studies, they can enhance the application of methods based on physical similarity (Oudin et al., 2008; Arsenault and Brissette, 2014; Poissant et al., 2017).

There are 101 meteorological stations distributed over the study area that were used to calculate basin-averaged time series of precipitation and air temperature at 10-minute resolution. The time series of temperature and precipitation were aggregated at hourly scale, from 2008 to 2018 (10 years). The mean annual precipitation in the region varies from 1224 to 2115 mm, while the mean annual temperature ranks between 10.9 and 13.6 ∘C. The catchments with the highest mean annual precipitation are located in the western region, with values around 2000 mm, while in those located towards the east the annual precipitation varies between

Table 1
Catchment descriptors.

	Catchment descriptors	Abbreviation	Unit	Min	Max
1	Longitude of the basin centroid (UTM H29)	X_c	m	502921	628069
2	Latitude of the basin centroid (UTM H29)	Y_c	m	4658071	4832820
3	Area	A	km^2	17.05	542.39
4	Perimeter	Pr	km	24.70	157.93
5	Main stream length	SL	km	9.10	62.20
6	Drainage network length	DNL	km	24.89	613.68
7	Drainage density	DD	km/km^2	0.57	1.72
8	Minimum elevation	MinE	m	16.52	404.95
9	Maximum elevation	MaxE	m	395.15	943.05
10	Average slope	Slp	%	7.70	28.30
11	Curve Number	CN	–	60.00	78.00
12	Clay percentage at depth 1 (0–5 cm)	Cl_1	%	12.77	18.73
13	Silt percentage at depth 1 (0–5 cm)	Si_1	%	26.59	44.77
14	Sand percentage at depth 1 (0–5 cm)	Sa_1	%	37.98	59.62
15	Clay percentage at depth 2 (30–60 cm)	Cl_2	%	16.21	27.09
16	Silt percentage at depth 2 (30–60 cm)	Si_2	%	23.58	40.34
17	Sand percentage at depth 2 (30–60 cm)	Sa_2	%	38.89	52.82
18	Clay percentage at depth 3 (100–200 cm)	Cl_3	%	17.12	25.11
19	Silt percentage at depth 3 (100–200 cm)	Si_3	%	22.06	37.94
20	Sand percentage at depth 3 (100–200 cm)	Sa_3	%	39.49	53.78
21	Agriculture land use	AgrUse	%	8.00	48.00
22	Non economic land use	NAgrUse	%	47.00	89.00
23	Remaining land uses	RLandUse	%	1.00	18.00
24	Mean annual precipitation	P_{mean}	mm	1224.46	2115.03
25	Mean temperature	T_{mean}	C	10.86	13.57

1200 and 1500 mm. The highest mean annual temperature occurs in the Southwest of the study region, and it gradually decreases towards the Northeast.

The Curve Number (CN) is commonly employed in hydrological modelling to represent soil infiltration capacity and its impact on runoff generation processes (Boughton, 1989). Despite not being a direct physical property of the basin, it is derived from three physical conditions: soil type, land cover, and terrain slope, which are the most influential factors on soil infiltration capacity. Therefore, it was decided to include it as a potential CD as it can synthesise this information for application in regionalisation schemes. The CN values are calculated using the methodology developed by Ferrer (2003) and described in the technical report by Álvarez (2010). The pixel resolution is 500 m x 500 m. The CN values were averaged within each catchment and ranged from 60 to 78, with the highest values located in the central part of the study area where agricultural land is concentrated (Fig. 2).

Land use data was obtained from the Spanish Land Occupation Information System (SIOSE), and land use categories were assigned based on the Hierarchical INSPIRE Land Use Classification System (HILUCS) developed by the Infrastructure for Spatial Information in Europe (INSPIRE) (https://inspire.ec.europa.eu/). The land use data layer is stored in GPKG format and contains 191,047 polygon spatial objects, each representing a specific land use type. The areas of these polygons were summed within each catchment to determine the percentages of each land use type.

The majority of land uses in the study catchments are classified as *Agriculture* and *Not in Economic Uses* (natural areas not dedicated to any economic use), which can occupy up to 48% and 89% of the land, respectively. Other HILUCS categories, such as *Residential Use* or *Transportation Networks*, only occupy small areas of the catchments. Therefore, for simplicity, the sum of these land uses was categorised as *Remaining Uses* (Fig. 2).

Soil types were obtained from the SoilGrids platform, which provides global predictions of soil classes and properties. Readers can refer to the website https://soilgrids.org/ and to Hengl et al. (2017) for further details on these data. The SoilGrids platform includes soil type information at three different depths: depth 1 (0–5 cm), depth 2 (30–60 cm), and depth 3 (100–200 cm). The downloaded data consist of GeoTIFF files with a pixel resolution of approximately 200 m x 200 m, which were then integrated into each basin.

The soil texture diagram shown in Fig. 3 was used to characterise the soil types. The diagram indicates that most of the soils are classified as loam and sandy loam, without significant variations in depth.

3. Methods

3.1. Hydrological model

The hydrological model used in this study is the MHIA (acronym for Lumped Hydrological Model in Spanish). The model performs a balance of the water volume in the soil, considering processes such as precipitation, infiltration, percolation, evapotranspiration, and exfiltration, to estimate the hydrograph at the basin outlet. The calculation and concepts of these processes are explained in detail in Farfán and Cea (2022b).

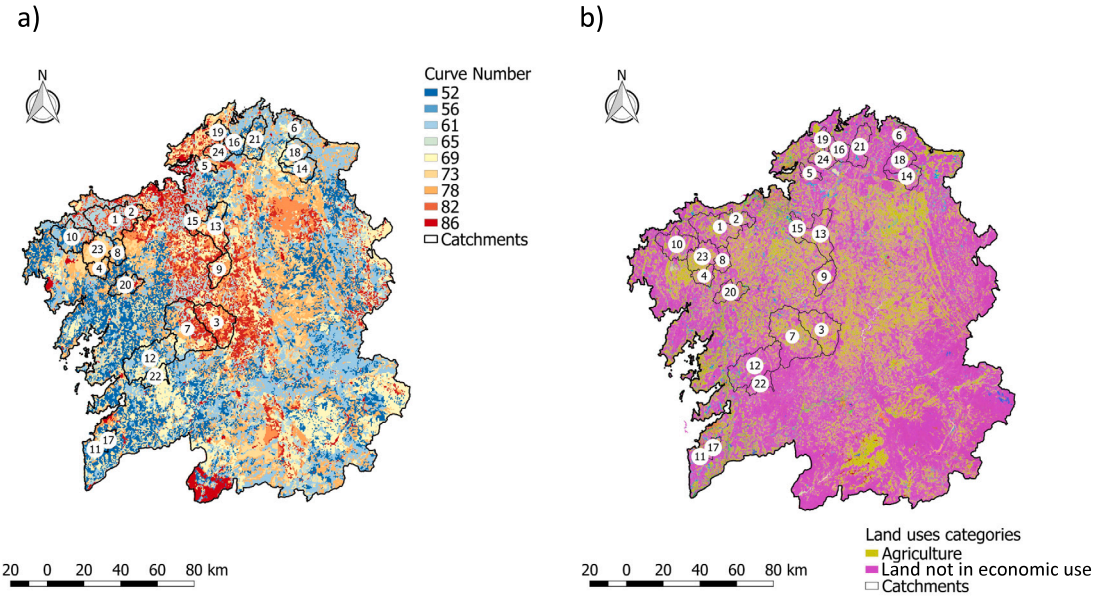


Fig. 2. (a) Curve Number distribution in the study region (b) Land uses in the study region. The map legend indicates uniquely *Agriculture* and *Not in economic use*. The non-listened uses are categorised in Table 1 as *Remaining uses*.

Table 2
Calibration parameters of the MHIA model (Farfán and Cea, 2022b).

Parameter	Symbol	Unit	Lower limit	Upper limit
Ratio between CN_{dry} and CN	ϕ		0.1	1
Exponent of percolation	m_1		5	60
Percolation permeability	K_s	mm/h	0.1	20
Exponent of exfiltration	m_2		15	75
Exfiltration permeability	K_b	mm/h	0.1	20
Coefficient lag-time relationship for surface runoff	k_1		0.1	6.5
Parameter for scaling gamma function for surface runoff	n_1		1	10
Coefficient lag-time relationship for groundwater flow	k_2		0.1	6.5
Parameter for scaling gamma function for groundwater flow	n_2		1	10
Parameter for potential evapotranspiration	b		0.4	2
Initial abstraction coefficient	α		0.0	0.2
Decaying coefficient for P_{acum}	d		0.01	1
Correction coefficient for S	a		1	4

The model requires input data such as time series of precipitation and temperature (averaged over the entire basin), as well as basin-specific parameters that need to be calibrated from observed discharge time series (Table 2). The study adopted a time resolution of 1 h.

The MHIA model represents the basin as a reservoir with a variable volume of water and a maximum storage capacity, and computes the hydrological processes at each computational time step. The code for the MHIA model is provided in Farfán and Cea (2022b), and the model parameters are summarised in Table 2. The MHIA model has been previously used in the study region and produced satisfactory results as demonstrated in Farfán and Cea (2022a).

The available precipitation, temperature, and discharge time series, aggregated at an hourly time scale, were split into two non-overlapping periods for model calibration and validation, spanning from 2008 to 2013 (calibration) and from 2013 to 2018 (validation). The model calibration was done independently for each catchment using a Monte Carlo approach with 5000 parameter sets. The parameters were sampled in their feasible space using a quasi-random method. The 25 best parameter sets (behavioural parameters) in each catchment, in terms of goodness of fit, were used to validate the model and perform the transfer to the receptor catchments.

In this study, the Nash–Sutcliffe Efficiency (NSE) (Nash and Sutcliffe, 1970) is used as a metric to evaluate the goodness of fit in each catchment. NSE is a commonly used metric in similar research and facilitates comparisons with other studies (e.g., Arsenault and Brissette, Cislaghi et al., Razavi and Coulibaly, Guo et al.). However, variations in the number of catchments and research periods among comparable studies can introduce bias when comparing NSE results (Guo et al., 2021). Therefore, it is important to consider this potential bias when interpreting and comparing the results. The NSE coefficient is calculated as follows:

$$\mathrm{NSE} = 1 - \frac{\sum_{i=1}^{n} (Q_{obs,i} - Q_{sim,i})^2}{\sum_{i=1}^{n} (Q_{obs,i} - \overline{Q_{obs}})^2} \tag{1}$$

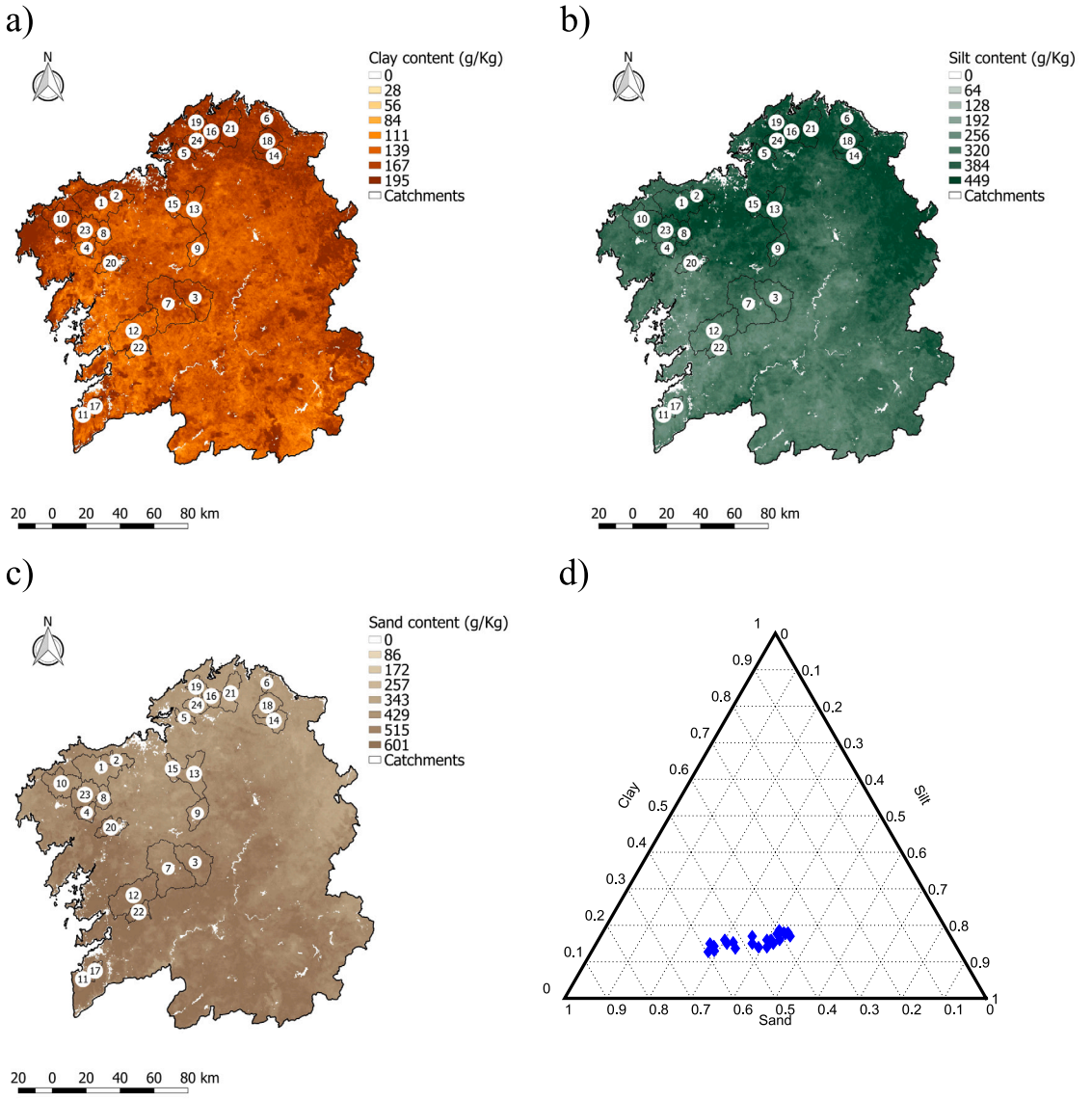


Fig. 3. Soil types in the study region at depth 1. (a) Clay content, (b) Silt content, (c) Sand content, (d) Texture Diagram for soil classification (Science and Administration, 1975).

where $Q_{obs,i}$ and $Q_{sim,i}$ are the ith observed and simulated discharge value, respectively, and $\overline{Q_{obs}}$ is the mean value of the observed discharge series.

3.2. Regionalisation schemes

The regionalisation schemes analysed are summarised in Table 5, and their mathematical formulation is provided in the Appendix.

3.2.1. Physical similarity methods

Physical similarity methods assume that model parameters from one basin can be transferred to another basin with similar physical and climatic attributes (Oudin et al., 2008; Cislaghi et al., 2020). In this study, we applied a physical similarity method using two different approaches. The first approach used Inverse Distance Weight (IDW) interpolation (Shepard, 1968) to estimate a set of model parameters in the receptor basin based on those of the donor basins. The second approach used IDW interpolation on the outlet hydrographs computed in the receptor basin using the parameter sets of the donor basins. This second approach, also known as ensemble modelling, has been applied in different studies with acceptable results (Oudin et al., 2008; Merz and Blöschl,

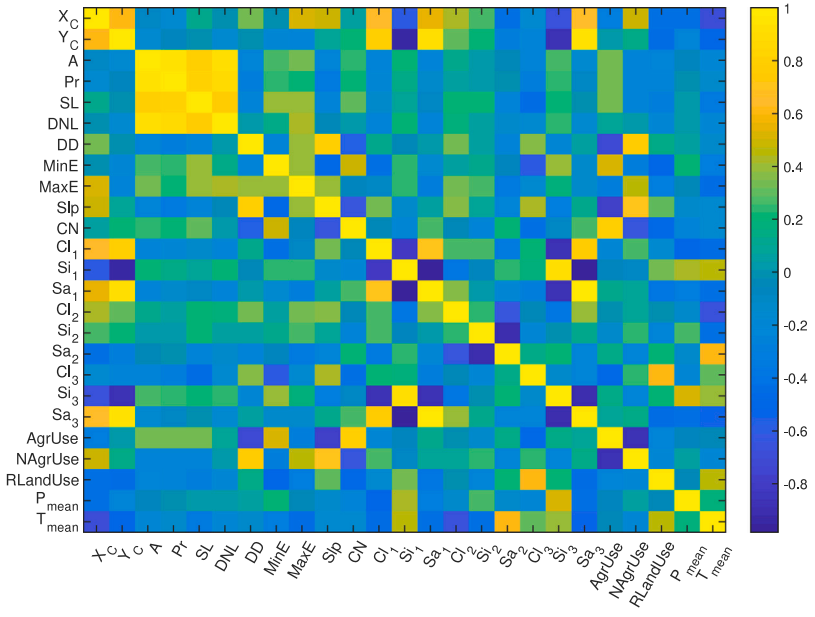


Fig. 4. Correlations between catchment descriptors.

Table 3
Groups of catchment descriptors. The descriptor selected as representative of each correlated group is indicated with the superscript *.

	Group	Catchment descriptors
Correlated CDs	1	Y_c^*, Cl_1, Si_1, Sa_1, Si_3, Sa_3
	2	A*, Pr, SL, DNL
	3	DD*, SLP, AgrUse, NagrUse, CN
	4	Si_2^*, Sa_2
Non-correlated CDs	5	X_c, MinE, MaxE , Cl_2, Cl_3, RLandUse, P_{mean}, T_{mean}

2004; Razavi and Coulibaly, 2016; Arsenault and Brissette, 2014). Between 5 and 10 donor basins were evaluated to determine the number of donors that provide the most reliable results, and after initial evaluation, the number of donor catchments was set to 10, following the procedures of previous studies (Arsenault et al., 2019; Poissant et al., 2017; Arsenault and Brissette, 2014). Further details on the mathematical formulations used in each approach can be found in Appendix.

A preliminary analysis of the correlations between the 25 catchment descriptors (CD) included in Table 1 was done in order to identify strong correlations. Four groups of strongly correlated CDs ($|r| > 0.7$) (Cislaghi et al., 2020; Goswami et al., 2007; Wagener et al., 2007) were identified, while those CDs that are not correlated with any other were included in a fifth group (Table 3). For each correlated group only one CD was used as representative of the whole group in order to avoid redundancies (Cislaghi et al., 2020; Wagener et al., 2007).

Group 1 contains the following CDs: Y_c, Cl_1, Si_1, Sa_1, Si_3, Sa_3. These CDs show positive correlations between 0.84 and 0.96, and negative correlations between −0.95 and −0.96. Considering the strong correlations between these CDs, we have only used the Latitude of the basin centroid (Y_c) to characterise (Table 3) this group (Cislaghi et al., 2020).

The second group includes the following CDs: A, Pr, SL and DNL. These CDs show strong positive correlations ranging from 0.78 to 0.95 (Fig. 4). The area of the catchment (A) was selected as the representative descriptor of this group.

Group 3 includes: AgrUse, Slp, DD, NAgrUse and CN, with positive correlations ranging from 0.78 to 0.80 and negative correlations ranging from −0.79 to −0.89. The representative CD selected in this group was the Drainage Density (DD).

In the case of group 4, which includes only the soil types for depth 2 (30–60 cm), the selected CD was Si_2.

Finally, group 5 includes the CDs that do not have any significant correlation with any other CD and therefore, all of them have been considered for the application of the regionalisation methods. The values of the considered CDs are shown in Table 4

3.2.2. Spatial proximity method

Spatial proximity methods assume that catchments that are close to each other have a similar hydrological behaviour, assuming that climatic, lithological and morphological conditions do not vary significantly in space (Merz and Blöschl, 2004; Cislaghi et al., 2020).

The current methodology employs an approach similar to the physical similarity scheme, utilising both parameter-averaging and output-averaging techniques. The difference lies in the use of IDW interpolation, which in this case calculates weights based on the

Table 4
Values of the selected catchment descriptors for each basin in the study area.

ID	Basin	X_c m	Y_c m	A km^2	DD km/km^2	MinE m	MaxE m	Si_2 %	Cl_2 %	Cl_3 %	RLandUse %	P mm	T C°
1	Anllons	526420	4783068	433.53	1.03	59.43	473.43	41	20	22	7	1403	12.5
2	Anllons Carballo	535304	4787157	120.17	1.09	114.27	470.62	44	21	21	11	1224	12.9
3	Arnego Ulla	581745	4726991	330.09	0.82	404.95	943.05	53	22	20	3	1343	10.9
4	Barcala Tambre	517816	4756248	91.35	1.06	167.18	472.77	42	19	20	8	2115	12.6
5	Belelle	575228	4812210	55.42	1.02	57.47	561.86	45	16	20	3	1358	13.6
6	Covo	623969	4832821	42.35	1.44	44.29	623.8	49	19	22	1	1226	11.5
7	Deza	565856	4723404	542.39	0.98	205.53	831.99	39	21	20	5	1289	12.3
8	Dubra Tambre	527832	4764897	92.67	0.99	197.03	494.79	50	21	17	3	1909	12.4
9	Furelos	583169	4755852	151.02	0.97	365.02	747.25	40	20	20	2	1399	12.1
10	Grande Camarinas	502921	4773477	251.74	0.85	80.56	453.46	46	19	23	5	1639	12.7
11	Grova	515605	4658071	17.05	1.46	34.72	626.61	39	21	25	17	1297	13.2
12	Lerez	545408	4707565	412.07	1.49	70.84	861.62	49	21	21	6	1818	12.9
13	Mandeo	581149	4779205	247.67	1.04	328.62	709.75	41	22	17	4	1394	11.8
14	Masma	628069	4810607	144.46	1.29	75.31	888.87	52	22	20	7	1554	11.5
15	Mendo	568519	4782179	83.34	1.1	71.31	538.22	39	21	20	6	1393	12.3
16	Mera	591187	4825020	109.38	1.52	77.81	612.79	49	27	22	11	1472	12
17	Minor	522746	4663003	62.55	1.44	16.52	578.37	42	19	25	18	1592	13.5
18	Ouro	624219	4819497	161.99	1.42	45.96	861.28	48	27	22	5	1591	10.9
19	Rego das Mestas	582296	4830510	71.55	1.19	74.93	478.48	43	20	23	6	1406	12.5
20	Sar	532072	4747463	138.94	0.57	33.64	395.15	46	16	21	14	1730	13.2
21	Sor Baixo	602577	4826713	182.63	1.72	159.97	683.24	45	22	21	7	1538	10.9
22	Verdugo	548414	4697721	102.18	1.7	324.25	939.3	45	18	21	4	2081	13
23	Xallas1	516840	4766828	202.58	0.98	313.25	472.95	49	22	20	5	2028	12.1
24	Xubia	582604	4819666	99.75	1.22	47.32	515.1	40	22	22	7	1489	12.3

Table 5
Schemes used for the regionalisation of the streamflow time series.

Regionalisation scheme	Approach	Description	Abbreviation
Physical similarity	Parameter estimation	IDW from catchment descriptors	$PhysicalPar$
	Ensemble modelling	Weighted average with IDW from catchment descriptors	$PhysicalEns$
Spatial proximity	Parameter estimation	IDW from centroids of the catchments	$SpatialPar$
	Ensemble modelling	Weighted average with IDW from centroids	$SpatialEns$
Regression-based	Parameter estimation	Multiple linear regression	$LinReg$
	Parameter estimation	Artificial Neural Networks	ANN

distances between the centroids of the catchments. In this type of scheme, the IDW interpolation aims to enhance the influence of nearby basins while reducing the impact of more distant ones on the final averaging process. The detailed mathematical formulations of each method can be found in Appendix.

3.2.3. Regression-based methods

Regression-based methods relate a dependent variable (in our case each model parameter) to a set of independent variables (in our case the CDs) (Oudin et al., 2008; Razavi and Coulibaly, 2013). The CDs used in the regression must be selected taking into account their influence on the hydrological response of the catchment. Different combinations of CDs have been used in previous studies (Merz and Blöschl, 2004; Cislaghi et al., 2020; Post, 2009; Heuvelmans et al., 2006), including meteorological and geomorphological variables as those included in Table 1.

We applied a feed forward back propagation Artificial Neural Network (ANN) of three layers (an input layer, a hidden layer and an output layer) and a Multiple Linear Regression ($LinReg$). The predictors used for each model parameter were the CDs that have a Spearman's correlation coefficient with that model parameter higher than 0.25 (Fig. 5). In the same way as in the physical similarity schemes, only one CD from each highly correlated group (Table 3) was used in the regression, in order to avoid redundancies.

Both regression-based methods were trained using a leave-one-out-cross-validation scheme (Oudin et al., 2008; Merz and Blöschl, 2004).

3.3. Characterisation of uncertainty on the streamflow predictions and success rate (SR)

A bootstrapping technique was used to estimate the uncertainty on the streamflow predictions obtained from the different regionalisation schemes (Efron, 1992; Heuvelmans et al., 2006; Arsenault and Brissette, 2014; Poissant et al., 2017). As mentioned in Section 3.1, the hydrological model was calibrated in each basin following a Monte Carlo approach, and the parameters of the 25 best simulations were retained as the behavioural parameter sets for each donor basin. In order to obtain the streamflow prediction in a receptor catchment, the regionalisation schemes were applied 250 times, each time using a random behavioural parameter set from each donor basin. In this manner, 250 streamflow predictions were obtained for each regionalisation scheme in each receptor basin.

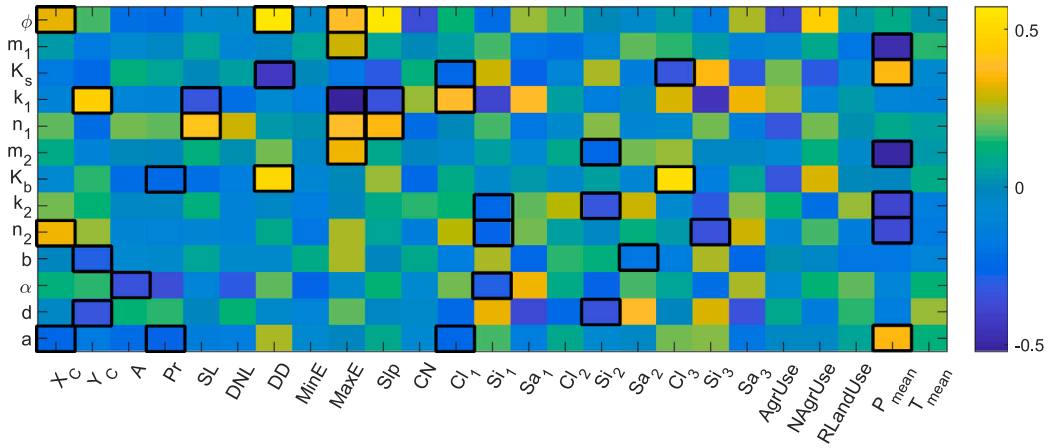


Fig. 5. Spearman's correlation coefficient between CDs (horizontal axis) and model parameters (vertical axis). The black squares indicate the CDs used as independent variables to predict the respective model parameters in the vertical axis.

From these, the performance of each regionalisation scheme was evaluated in terms of the Success Rate (SR) proposed by Arsenault and Brissette (2014). The SR is defined as the total number of acceptable predictions divided by the total number of predictions and is calculated using the following equation:

$$\mathrm{SR} = \frac{\text{Number of acceptable predictions}}{\text{Total number of predictions}} \tag{2}$$

A SR of 0.6 indicates that the regionalisation scheme was successful on 60% of the 250 bootstrapped predictions. The threshold established to consider a prediction as successful was that its NSE was larger than 75% of the average NSE derived from the 25 behavioural parameter sets obtained in calibration (Section 3.1). This threshold is slightly lower than the one used by Arsenault and Brissette (2014) because the number of basins included in our study is smaller and thus, the performance of the regionalisation scheme is expected to be to some degree lower.

This allows the estimation of the SR of each of the methods, the distribution of the expected regionalised NSE values and the calculation of the mean and standard deviation of the distribution.

4. Results

4.1. Calibration and validation of the hydrological model

The MHIA hydrological model was independently calibrated in each basin, using 5000 quasi-random parameter sets and retaining the 25 best simulations in terms of NSE during the calibration period (2008–2013), as described in Section 3.1. After identify the 25 behavioural parameter sets for each basin, these sets were used to run the model during the validation period (2013–2018) and the entire period (2008–2018). In the appendix section, we provide plots related to the calibrated model for some of the study catchments using the behavioural sets and the Generalised Likelihood Uncertainty Estimation methodology (GLUE) (Beven and Freer, 2001; Beven, 2006).

Fig. 6 shows the average Nash–Sutcliffe Efficiency (NSE) coefficient computed at an hourly scale for the 25 behavioural simulations in each basin. The results are presented for the calibration period (2008–2013), the validation period (2013–2018), and the entire period (2008–2018). The Grova basin (basin 11 in Fig. 1 and Table 4) has the lowest NSE value, around 0.5. This catchment is very small, covering only 17 km^2, and is located in the southwestern limit of the study region. About one-third of the basins exhibit mean NSE values between 0.50 and 0.65, while the other two-thirds show mean NSE values between 0.65 and 0.80.

In the following sections, the regionalisation schemes are evaluated only in the entire period (2008–2018) and at an hourly scale. The NSE values corresponding to the entire period in Fig. 6 can be considered as a reference value for the evaluation of the regionalisation schemes.

4.2. Physical similarity methods

Fig. 7 shows the probability distribution of the NSE coefficients obtained from the 250 bootstrapped parameter sets (Section 3.3) in each basin for the period 2008–2018, using the two physical similarity methods ($Physical\,Ens$ and $Physical\,Par$ in Table 5). The histograms of the NSE coefficients corresponding to the 25 behavioural simulations in each basin (Section 4.1) are also shown in red for reference and comparison purposes.

The global average of the NSE obtained with the $Physical\,Ens$ and $Physical\,Par$ methods is 0.68 and 0.56 respectively, with a standard deviation of 0.13 in both cases. In overall terms, the distributions of the results present similar amplitudes with the

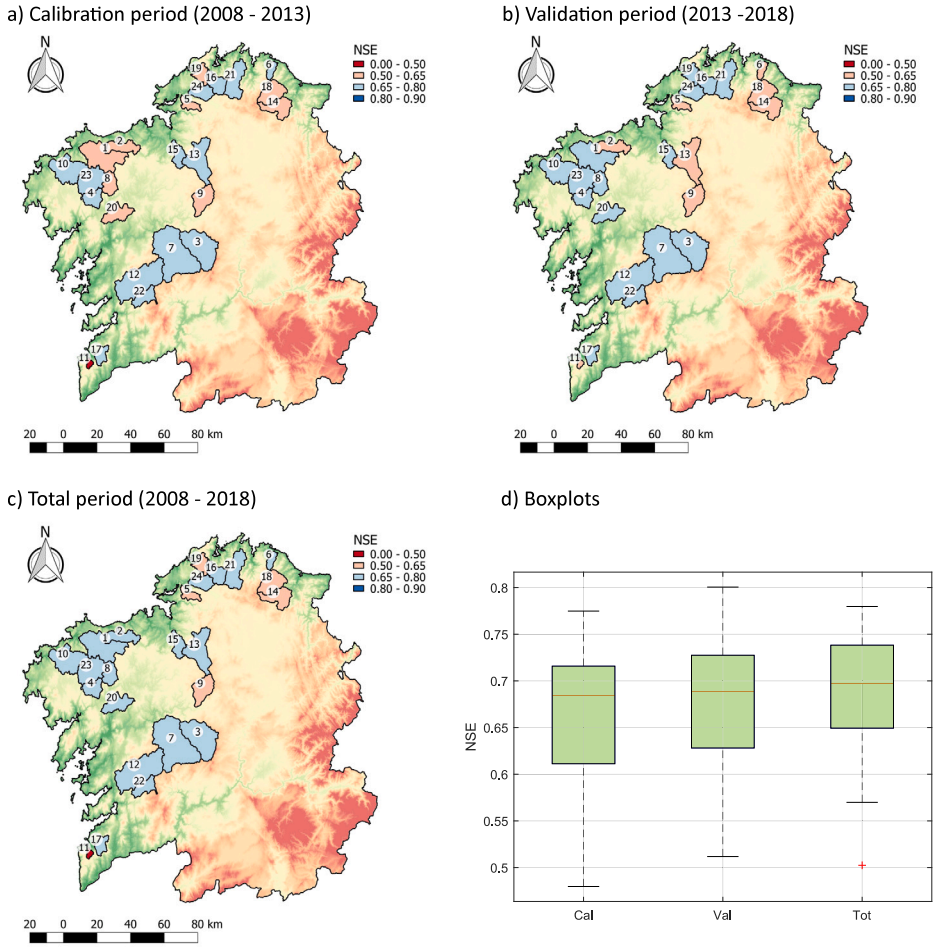


Fig. 6. NSE coefficients in the calibration, validation and total periods. The values shown in figures (a), (b) and (c) correspond to the average NSE of the 25 behavioural simulations.

exception of the Miñor (17) and Ouro (18) rivers in the $PhysicalPar$ method, where a wider dispersion is observed. The results of the ensemble-based physical similarity method ($PhysicalEns$) far outperformed the $PhysicalPar$ in most cases. Exceptions are the basins Belelle (5), Grova (11) and Verdugo (22). In some basins the $PhysicalEns$ method reached higher NSE coefficients than those obtained in the calibration of the hydrological model (Fig. 7). This is a well-known characteristic of ensemble modelling methods, and it is explained by the error compensation between simulations, leading to predictions with better fits than those obtained with a single parameter set. The reader is referred to Kumar et al. (2015), Farfán et al. (2020), Farfán and Cea (2022a) for more details on ensemble modelling.

4.3. Spatial proximity methods

Fig. 8 shows the probability distribution of the NSE coefficients obtained with the $SpatialEns$ and $SpatialPar$ methods in the 24 basins. Overall, the results are similar to those obtained with the physical similarity methods. The average NSE obtained with $SpatialEns$ (0.65) is significantly better than the average NSE obtained with $SpatialPar$ (0.56), with standard deviations of the NSE values of 0.13 and 0.15 respectively. In no case the $SpatialPar$ scheme outperformed the $SpatialEns$ method. There are also a few catchments (e.g. Ouro and Miñor basins) in which there is a high dispersion in the NSE distribution of the $SpatialEns$ $SpatialPar$ results, which indicates a high uncertainty in the results given by this type regionalisation method (Fig. 7). The possible causes of this behaviour are analysed in Section 4.5 in terms of the location of the watersheds in the study area and the density of nearby watersheds.

4.4. Regression-based methods

The results obtained with the regression-based schemes are shown in Fig. 9. The global average NSE obtained with the ANN scheme was 0.59 with a standard deviation of 0.10, while the global average NSE obtained with the $LinReg$ scheme was 0.57, with

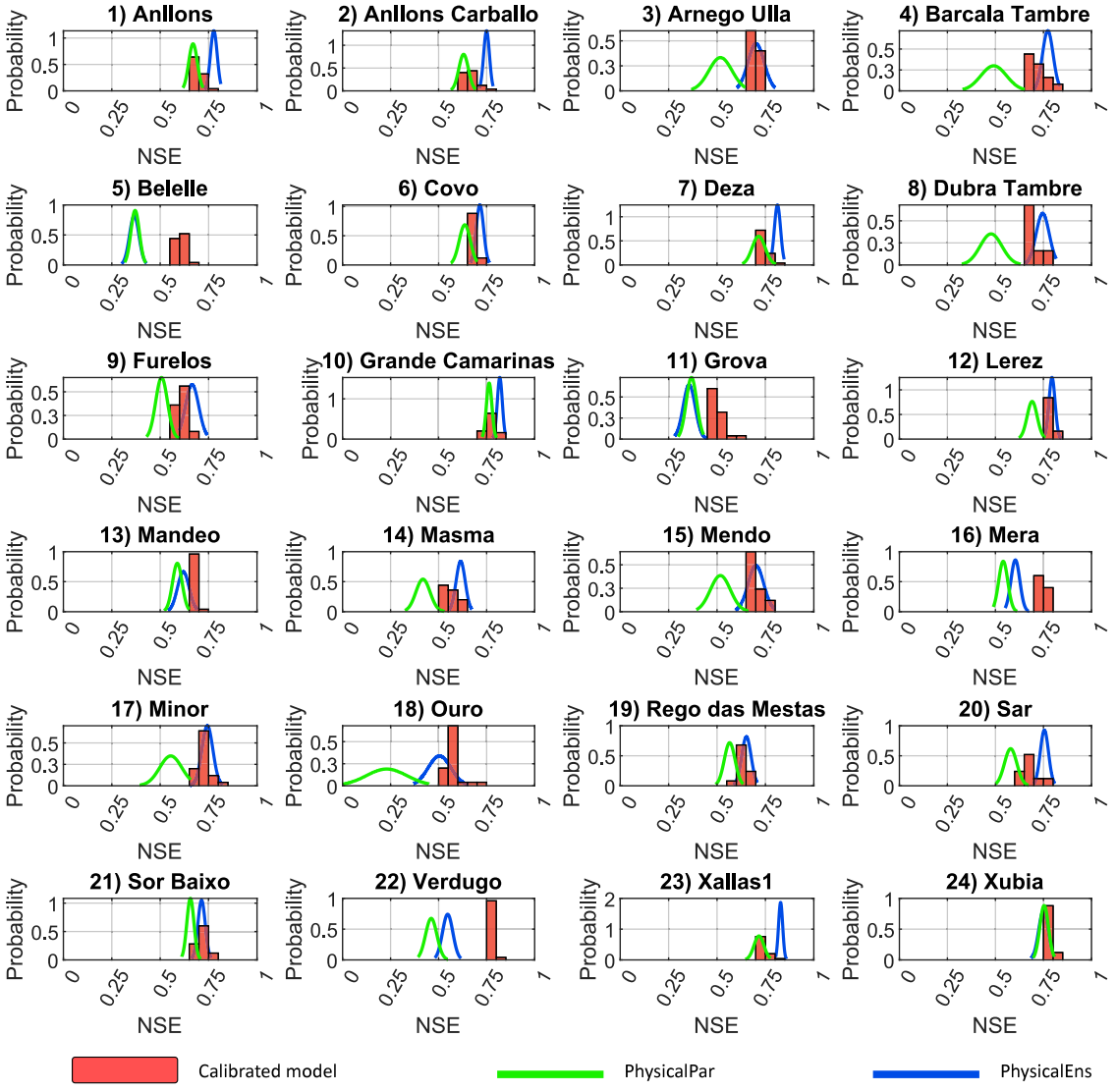


Fig. 7. Probability distribution of the NSE coefficients for the physical similarity regionalisation schemes.

Table 6
Global performance of the regionalisation schemes.

Ranking	Scheme	$\overline{NSE}$	σ_{NSE}	SR > 0.75
1	PhysicalEns	0.68	0.13	21/24
2	SpatialEns	0.65	0.13	21/24
3	ANN	0.58	0.12	16/24
4	LinReg	0.57	0.12	17/24
5	SpatialPar	0.56	0.15	16/24
6	PhysicalPar	0.56	0.13	13/24

a standard deviation of 0.11. The ANN scheme has shown marginally better results than the $LinReg$ scheme, but the difference in performance is not relevant. The distribution of the NSE obtained shows that there is a large uncertainty related to the performance of these methods in the study basins.

4.5. Discussion

The success rate (SR) obtained in each catchment with each regionalisation scheme is shown in Fig. 11.

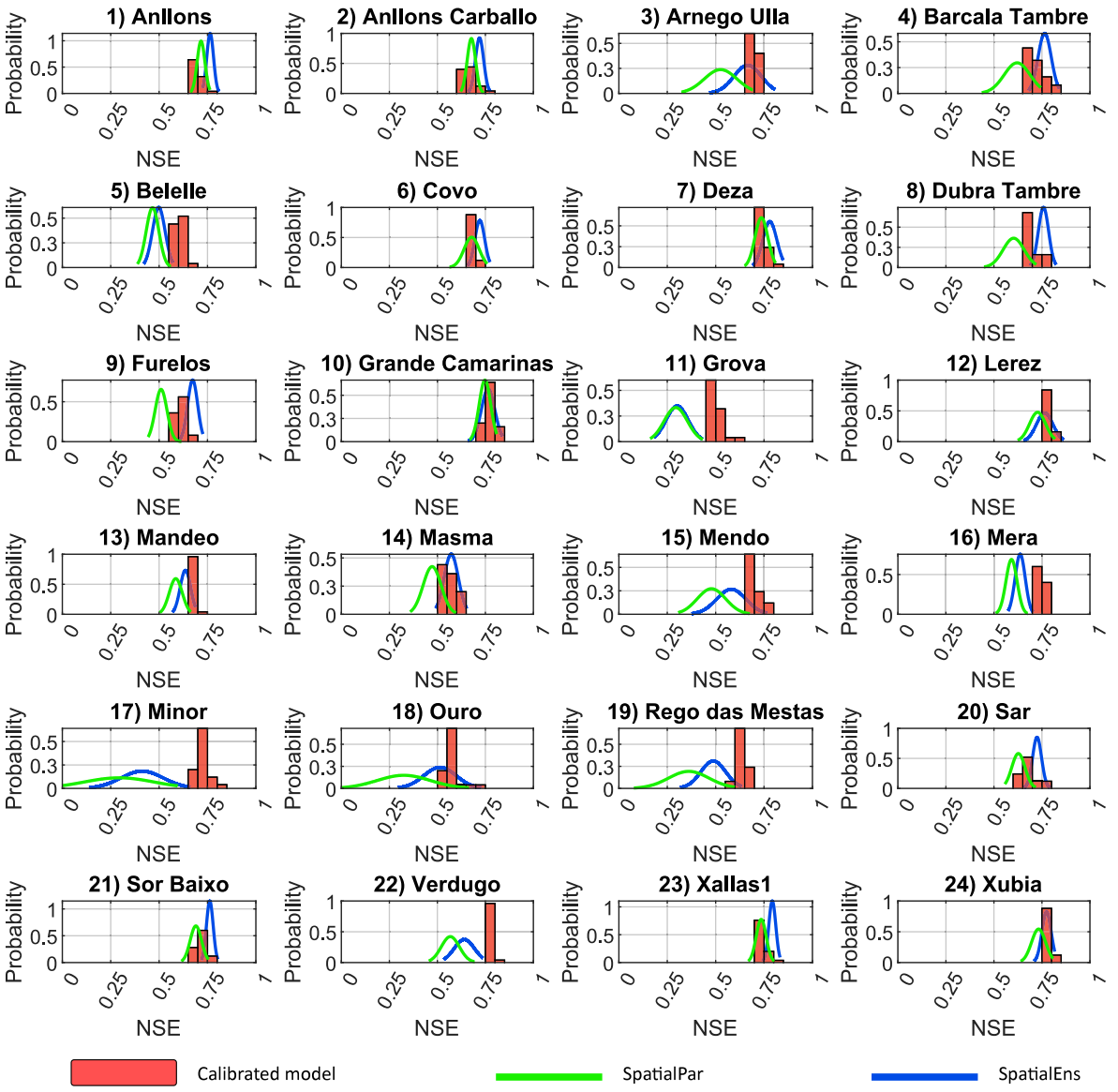


Fig. 8. Probability distribution of the NSE coefficients for the spatial proximity regionalisation schemes.

The two schemes based on ensemble modelling are clearly those with the best performance. In the case of catchments 4, 8, 14, 15, 16 and 18, the $Physical Par$ scheme has a SR close to 0, while the $Physical Ens$ scheme has SRs greater than 0.75 in the same catchments. Something similar occurs with catchments 15, 18 and 19, where the $Spatial Par$ scheme has low SR, while the $Spatial Ens$ method produces satisfactory SR in all of them. Additionally, there are no cases where the $Physical Par$ or $Spatial Par$ methods produce higher SRs than the $Physical Ens$ or $Spatial Ens$ schemes.

Regarding regression-based schemes, poor streamflow reproduction (SR) persists for catchments 4, 8, 14, and 18. The reason why ensemble-based methods provide the best results is that these schemes transfer the full set of parameters from the donor to the receptor catchments, preserving the interaction and potential equifinality between hydrologic model parameters. These interactions are lost when parameters are averaged or estimated individually, as occurs in parameter interpolation and regression-based methods. This is in line with the works of Oudin et al. (2008), Arsenault and Brissette (2014), Poissant et al. (2017), which show the importance of transferring a complete set of model parameters rather than transferring parameters individually from one catchment

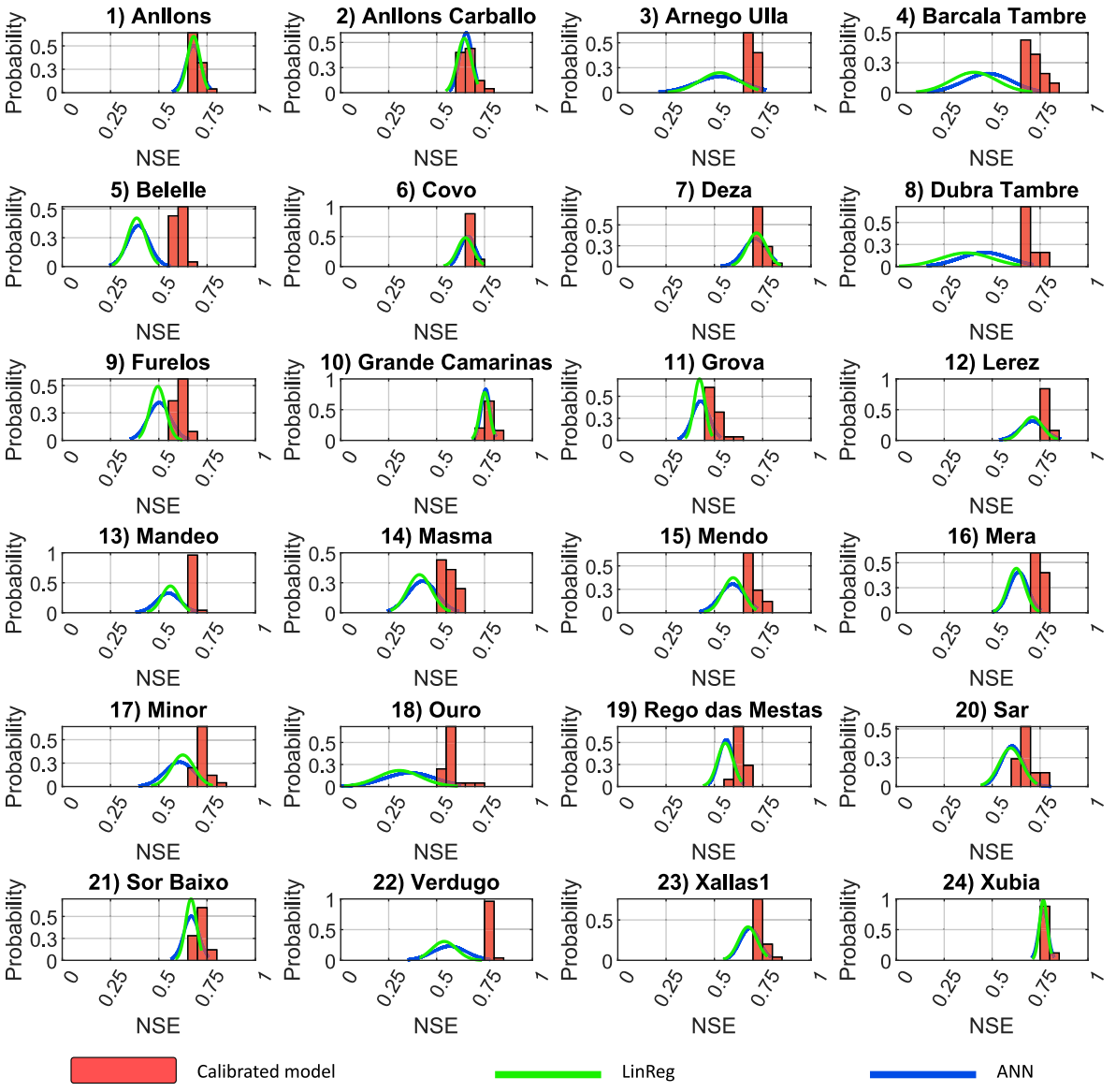


Fig. 9. Probability distribution of the NSE coefficients for the ANN and $LinReg$ regionalisation schemes.

to another. This drawback may be enhanced due to the hourly time scale of the modelling process in the present study, as parameter uncertainty tends to dominate at finer temporal scales (Li et al., 2016; Farfán and Cea, 2022a).

Table 6 shows that, although the $PhysicalEns$ scheme produces the best results in terms of average NSE, its results are similar to those of the $SpatialEns$ scheme in terms of SR. In both schemes, 21 of the 24 catchments present a SR greater than 0.75. The difference between these schemes is in the catchments in which the low SR occurs. In this regard, an advantage of the spatial proximity-based methods over the other schemes applied in the present study is that by observing the spatial distribution of the results, it is possible to understand in part the cases with low SR, if these correspond to a low concentration of nearby donor catchments. This explains the SR close to 0 in catchments Miñor (17) and Grova (11), which are too distant from their donor catchments (3, 4, 7, 8, 9, 10, 12, 20, 22, 23) in comparison with the rest of the rivers that show satisfactory results. The lack of nearby donor catchments has a clear effect on the performance of spatial proximity methods, as pointed out by Ali et al. (2012), Beck et al. (2016).

When the $PhysicialPar$ and $PhysicalEns$ schemes were applied, this limitation has been partially overcome since, as shown in Fig. 11, the SR in the Miñor river is greater than 0.75 for the $PhysicalEns$ scheme, and greater than 0.5 in $PhysicalPar$ schemes. This is because in physical similarity schemes the lack of nearby catchments is not crucial, since it is compensated by other donor

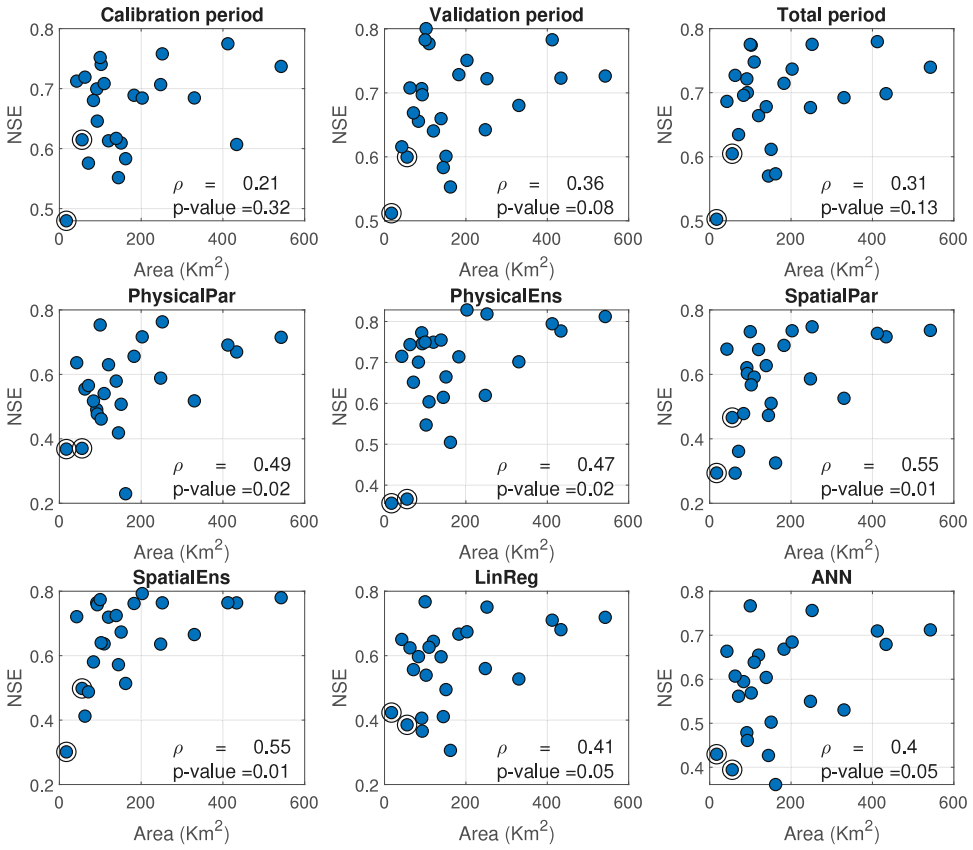


Fig. 10. Area of the basin vs. average NSE for the behavioural simulations (upper row) and different regionalisation schemes (middle and lower rows). ρ is the Spearman's coefficient. The Grova and Belelle basins are marked within the black circle.

catchments with similar physical characteristics. However, this does not explain the persistently poor results in the Grova (11), Belelle (5) and Verdugo (22) catchments, the two last ones with SRs close to 0.

The $LinReg$ and ANN methods applied in the catchments Grova (11) and Belelle (5) present SRs greater than 0.75, which can be attributed to the fact that in the regression-based schemes, the catchment gap is not a drawback, since in these schemes all the catchments are incorporated in the regression of the parameters in the remaining receiving catchment. Nevertheless, and despite the fact that these methods present the best results in catchments 11 and 17, catchments with unsatisfactory SR are also distributed within the study area, with slightly better results in the case of the ANN scheme, probably due to its ability to identify non-linear relationships between the CDs and the model parameters. The catchments with low SRs are mostly the same as those in which the $PhysicalPar$ and $SpatialPar$ methods also fail to produce good predictions. These low SRs can be attributed to the loss of the interaction between the model parameters. However, the black box nature of the regression-based methods does not allow detailed analysis of the cases with unsatisfactory results and the processes that may be involved in the relationship between input and output variables, being this a major drawback in the application of this type of methodologies (Heuvelmans et al., 2006; Razavi and Coulibaly, 2017; Farfán et al., 2020).

In addition, it is important to take into account that as explained by He et al. (2011) and Guo et al. (2021), when applied in an area with a reduced number of catchments, regression-based methods may not reach their maximum efficiency. This is because the main assumption of regression-based schemes states that there is an underlying relationship between CDs and parameters. However, in practice, this neglects the principle of equifinality (Oudin et al., 2008; Poissant et al., 2017).

Our results are in agreement with studies such as those by Kokkonen et al. (2003), Young (2006) in which regression-based methods performed better than parameter averaging methods (in our case $PhysicalPar$ and $SpatialPar$). Nevertheless, these results seem to be due to the large uncertainty of these methods resulting from the high variability in the regionalised parameter sets. A possible approach to overcome these shortcomings could be to calibrate the model by using regionalised flow indices. These types

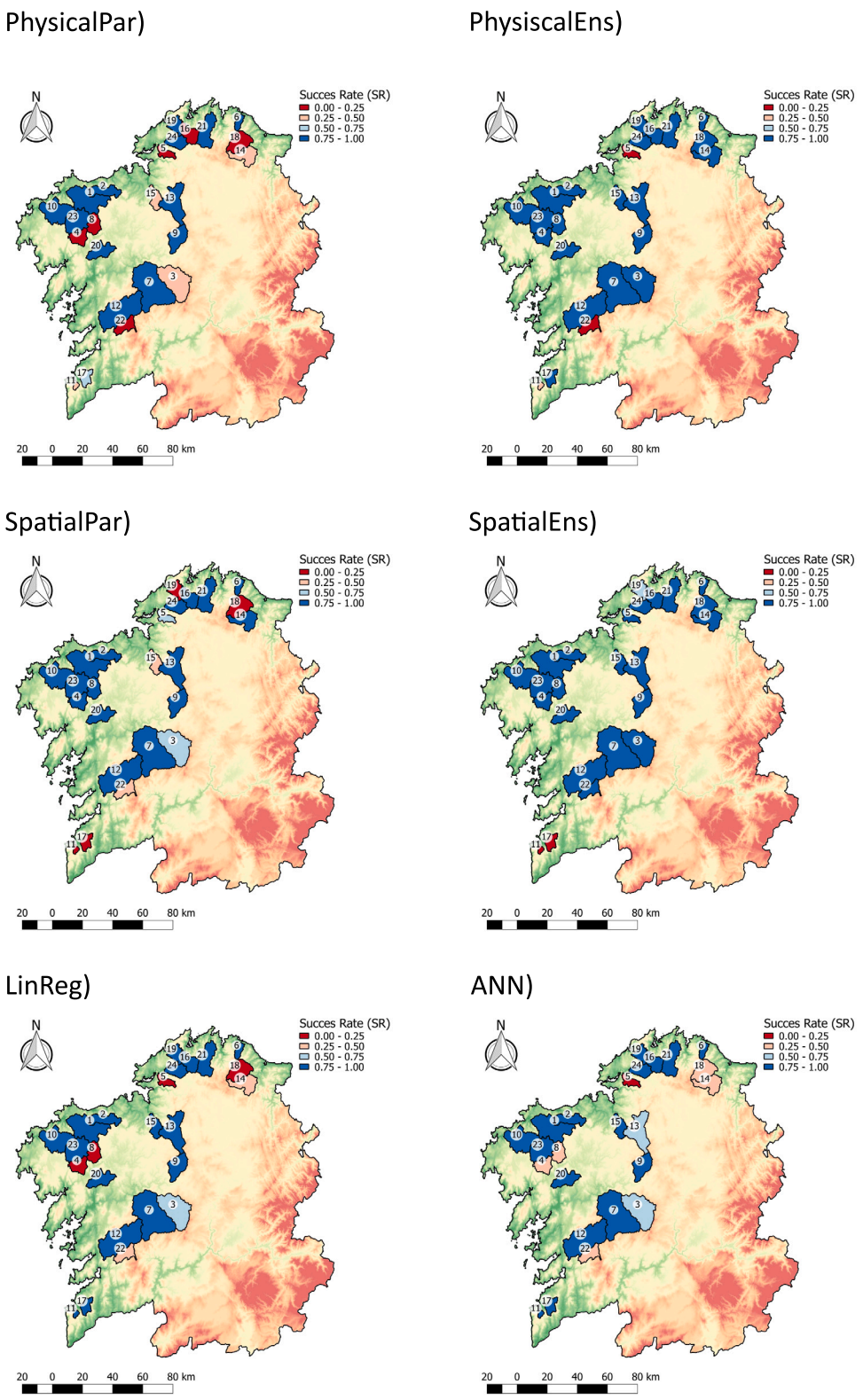


Fig. 11. Spatial distribution of the mean SR for the regionalisation schemes.

of schemes are constructed by establishing approximations between flow indices (such as mean annual flow) and CDs from donor catchments. These relationships are then extrapolated to the ungauged catchments. These kinds of schemes have been implemented in previous studies using machine learning techniques such as random forests or Bayesian frameworks (Almeida et al., 2016; Prieto et al., 2019, 2022; Westerberg et al., 2016; Kavetski et al., 2018).

Fig. 10 shows that in the calibration, validation and total periods, the basins with smaller area tend to present lower NSE values than those of larger area. This effect seems to be enhanced in the case of regionalisation schemes, where moderate correlations between mean NSE bootstrapped values and the catchment area have been detected. This correlation is more pronounced in methods based on spatial proximity and physical similarity. Based on this, the poor results in the Belelle (5) and Grova (11) basins are explained by a limitation in the structure of the hydrological model that makes it less accurate in small basins which causes that the regionalisation schemes tend to lose accuracy in basins with smaller area. This explains the low SR obtained in the Grova (11) river basin, which has an area of 17.5 km^2, being the smallest in our study area and partly those of the Belelle (5) river basin, which has an area of 55 km^2, being the third smallest.

5. Conclusions

The objectives of this study were focused on apply and evaluate different schemes for the regionalisation of streamflow series in Galicia, northwestern Spain. To this end, six regionalisation schemes were evaluated and organised as follows, two schemes based on physical similarity, with average of parameters and average of outputs, two schemes based on spatial proximity with average of parameters and average of outputs and two regression based methods, multiple linear regression and Artificial Neural Networks. To account for the uncertainty inherent to this type of study a bootstrapping technique was used. The results were analysed in terms of the probability distribution of the goodness-of-fit and the Success Rate of the regionalisation schemes.

The output-averaging methods provided the best results because they preserve the interaction between the parameters of the hydrological model. These interactions are lost when the parameters are averaged or estimated individually, as it is the case in the regression-based methods and in the methods based on parameter interpolation. This is in agreement with the works of Oudin et al. (2008), Arsenault and Brissette (2014), Poissant et al. (2017), which show the importance of transferring a complete set of model parameters instead of transferring the parameters individually.

The comparison of the performance of the different regionalisation schemes shows that, When a receptor basin has a sufficient number of nearby donor basins, the $SpatialEns$ methods provides the best performance. On the other hand, the $PhysicalEns$ method shows a better performance in areas where there is a low density of donor catchments.

The existence of correlations between the average NSE of the regionalisation schemes and the area of the basins shows that the applied regionalisation schemes tend to lose accuracy in basins with a small area (NSE > 0.5 when A > 100 km^2). This is probably related to a limitation in the structure of the hydrological model, which causes low NSE values in the basins with smaller areas. Therefore, in future research, other model structures could be compared and evaluated in terms of their ability to capture the dominant hydrological processes in different catchments. This would help to identify models that can adequately represent the "uniqueness-of-the-place" (Prieto et al., 2022).

Regression-based methods can offer better results in watersheds where the other schemes are inefficient as is the case of Grova and Miñor rivers, although with a wide distribution of the results, which can lead to poor performance in watersheds where the other schemes perform well, so their application is not recommended in areas with a number of watersheds similar to that of the present study.

CRediT authorship contribution statement

Juan F. Farfán: Conceptualisation, Methodology, Code, Writing – original draft . **Luis Cea:** Visualisation, Supervision, Writing – review & editing.

Data availability

The meteorological data was obtained from the agency MeteoGalicia at https://www.meteogalicia.gal/. The discharge data has been provided by the regional water administration Augas de Galicia at https://augasdegalicia.xunta.gal/.

Funding

This study has received financial support from the Galician government (*Xunta de Galicia*) as part of its pre-doctoral fellowship program (*Axudas de apoio á etapa predoutoral 2019*) Register Nº ED481A-2019/014. Funding for open access charge: Xunta de Galicia (project Nº ED431C 2022/010) and Universidade da Coruña/CISUG

Appendix

Calibration and validation of the MHIA model

Spatial proximity methods

The weights for the average of the parameters are calculated by means of the IDW interpolation (Shepard, 1968)applied as follows:

$$\mathbf{U}j^{\text{receptor}} = \sum_{i=1}^{M-1} \mathbf{W}_{j,i} \cdot \theta i^{\text{donor}} \tag{3}$$

$$\mathbf{W}_{j,i} = \left(\frac{d_{j,i}^{-2}}{\sum_{i=1}^{M-1} d_{j,i}^{-2}} \right) \tag{4}$$

where $\mathbf{U}j^{\text{receptor}}$ is the transferred parameter for the jth ungauged catchment, M is the number of nearby catchments used in the interpolation, θi^{donor} are the model parameters of the ith donor catchment, $\mathbf{W}_{j,i}$ is the vector of weights computed with the IDW interpolation. Then, $d_{j,i}$ is the distance between the centroids of the receptor (j) and donor (i) catchments (see Fig. 12). In the case of the output-averaging, the discharge series in the receptor catchments are calculated as:

$$\mathbf{Q}_j^{receptor} = \sum_{i=1}^{M} \mathbf{Q}(\theta_i^{donor}) \cdot \mathbf{W}_{j,i} \tag{5}$$

where $\mathbf{Q}_j^{receptor}$ is the hydrograph for the receptor jth catchment, $Q(P_i^{donor})$ is the hydrograph obtained by averaging the hydrographs generated by running the model with the parameters of the donor catchments $i = 1...M$ and $W_{i=1..M}$ is the vector of weights obtained from Eq. (4).

Regression-based methods

In the case of ANN, we have used a feed forward back propagation neural network consisting of three layers: an input layer, a hidden layer and an output layer, which can be represented as:

$$\mathbf{U}_j^{receptor} = \mathbf{W}_1 \times tanh\left(\mathbf{W}_2 \times \mathbf{CD} + \beta\right) \tag{6}$$

where the catchment descriptors (**CD**) are the input of the ANN, $\mathbf{U}_j^{receptor}$ is the output of the ANN (the 13 parameters of the hydrological model), β is the bias, $\mathbf{W}_1$ is the matrix of the weights between the hidden and the output layer, and $\mathbf{W}_2$ is the matrix of the weights between the input and the hidden layer.

The Multiple Linear Regression (MLR) model was implemented as:

$$\mathbf{U}_j^{receptor} = \left(\mathbf{CD} \times \mathbf{W} + \boldsymbol{\beta}\right) \tag{7}$$

where $\mathbf{U}_j^{receptor}$ is the output of the MLR model (the 13 parameters of the hydrological model), **CD** are the catchment descriptors and **W** is the matrix of linear regression coefficients and β is a bias vector.

Physical similarity methods

In the present work we have applied a physical similarity method using two different approaches. In the first approach a set of parameters is estimated using the IDW interpolation as:

$$\mathbf{U}_j^{receptor} = \sum_{i=1}^{M-1} \mathbf{W}_{j,i} \cdot \boldsymbol{\theta}_i^{donor} \tag{8}$$

$$\mathbf{W}_{j,i} = \left(\frac{\mathbf{pd}_{j,i}^{-2}}{\sum_{i=1}^{M-1} \mathbf{pd}_{j,i}^{-2}} \right) \tag{9}$$

where $U_j^{receptor}$ is the predicted parameter for the jth catchment, M is the number of catchments, θ_i^{donor} are the model parameters of the ith donor catchment, $W_{j,i}$ is the vector of weights to transfer the model parameters from the basin i to the basin j, and $pd_{j,i}$ is the euclidean distance between the CDs of the receptor and donor catchments, calculated as:

$$\mathbf{pd}_{j,i} = \sqrt{\sum_{k=1}^{N} \cdot (\mathbf{CD}_{k,j} - \mathbf{CD}_{k,i})^2} \tag{10}$$

where $\mathbf{CD}_{k,j}$ is the kth catchment descriptor of the jth catchment, N is the number of catchment descriptors. In order to apply Eq. (10), the descriptors are normalised to $[-1; +1]$ (Garambois et al., 2015; Cislaghi et al., 2020; Zhang and Chiew, 2009; Oudin et al., 2008).

The output-averaging is calculated as follows:

$$\mathbf{Q}_j = \sum_{i=1}^{M} \mathbf{Q}(\theta_i) \cdot \mathbf{W}_{j,i} \tag{11}$$

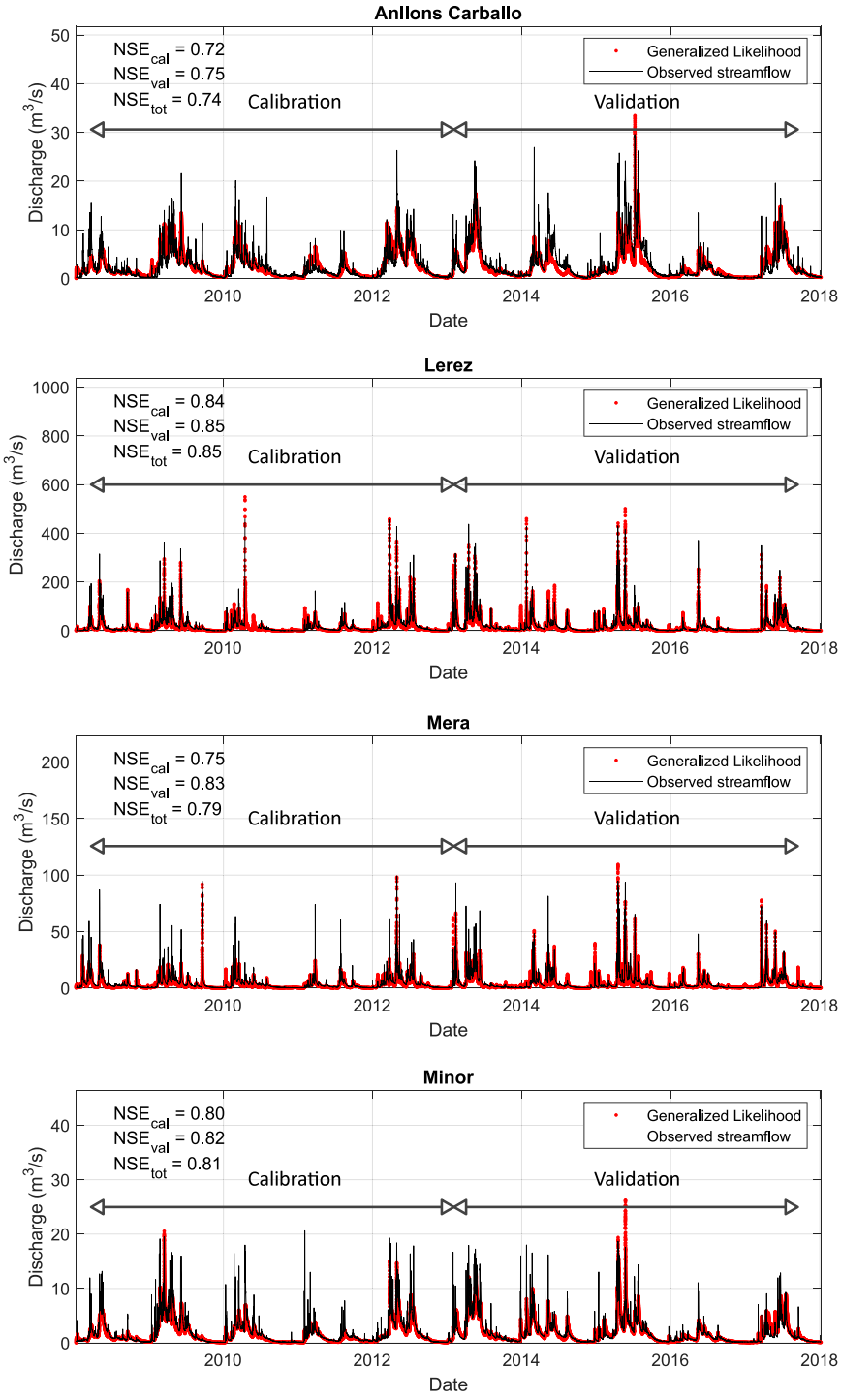


Fig. 12. Calibration and validation of the MHIA model using the GLUE methodology for some basins in the study region (Beven and Freer, 2001; Beven, 2006).

where $\mathbf{Q}_j$ is the hydrograph for the receptor jth catchment, $\mathbf{Q}(\theta_{i=1..M})$ are the hydrographs obtained form the model run with the parameters of the donor catchments $i = 1...M$ and $\mathbf{W}_{i=1..M}$ is the vector of weights obtained from Eq. (9) used for the averaging of the hydrographs.

References

Addor, N., Nearing, G., Prieto, C., Newman, A.J., Le Vine, N., Clark, M.P., 2018. A ranking of hydrological signatures based on their predictability in space. Water Resour. Res. 54 (11), 8792–8812.

Ali, G., Tetzlaff, D., Soulsby, C., McDonnell, J.J., Capell, R., 2012. A comparison of similarity indices for catchment classification using a cross-regional dataset. Adv. Water Resour. 40, 11–22.

Almeida, S., Le Vine, N., McIntyre, N., Wagener, T., Buytaert, W., 2016. Accounting for dependencies in regionalized signatures for predictions in ungauged catchments. Hydrol. Earth Syst. Sci. 20 (2), 887–901.

Álvarez, A., 2010. Mapa de caudales máximos Memoria Técnica. Madrid: Centro de Estudios HidrograÁficos Del CEDEX.

Arsenault, R., Breton-Dufour, M., Poulin, A., Dallaire, G., Romero-Lopez, R., 2019. Streamflow prediction in ungauged basins: analysis of regionalization methods in a hydrologically heterogeneous region of Mexico. Hydrol. Sci. J. 64 (11), 1297–1311.

Arsenault, R., Brissette, F.P., 2014. Continuous streamflow prediction in ungauged basins: The effects of equifinality and parameter set selection on uncertainty in regionalization approaches. Water Resour. Res. 50 (7), 6135–6153.

Beck, H.E., van Dijk, A.I., De Roo, A., Miralles, D.G., McVicar, T.R., Schellekens, J., Bruijnzeel, L.A., 2016. Global-scale regionalization of hydrologic model parameters. Water Resour. Res. 52 (5), 3599–3622.

Beven, K., 2006. A manifesto for the equifinality thesis. J. Hydrol. 320 (1–2), 18–36.

Beven, K., Freer, J., 2001. Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the glue methodology. J. Hydrol. 249 (1–4), 11–29.

Boughton, W., 1989. A review of the USDA SCS curve number method. Soil Res. 27 (3), 511–523.

Burn, D.H., Boorman, D.B., 1993. Estimation of hydrological parameters at ungauged catchments. J. Hydrol. 143 (3–4), 429–454.

Cislaghi, A., Masseroni, D., Massari, C., Camici, S., Brocca, L., 2020. Combining a rainfall–runoff model and a regionalization approach for flood and water resource assessment in the western Po Valley, Italy. Hydrol. Sci. J. 65 (3), 348–370.

Efron, B., 1992. Bootstrap methods: another look at the jackknife. In: Breakthroughs in Statistics. Springer, pp. 569–593.

Farfán, J.F., Cea, L., 2022a. Improving the predictive skills of hydrological models using a combinatorial optimization algorithm and artificial neural networks. Model. Earth Syst. Environ. 1–16.

Farfán, J.F., Cea, L., 2022b. MHIA model (modelo hidrológico agregado). http://dx.doi.org/10.4211/hs.d98161b9f3fb4d03a4358f6c8b5f2c04, https://www.hydroshare.org/resource/d98161b9f3fb4d03a4358f6c8b5f2c04/.

Farfán, J.F., Palacios, K., Ulloa, J., Avilés, A., 2020. A hybrid neural network-based technique to improve the flow forecasting of physical and data-driven models: Methodology and case studies in andean watersheds. J. Hydrol. Regional Stud. 27, 100652.

Ferrer, M., 2003. Análisis de Nuevas Fuentes de Datos para la Estimación del Parámetro Número de Curva del Modelo Hidrológico del SCS. Serie Cuadernos de InvestigaciÓN CEDEX C, 48.

Garambois, P.-A., Roux, H., Larnier, K., Labat, D., Dartus, D., 2015. Parameter regionalization for a process-oriented distributed model dedicated to flash floods. J. Hydrol. 525, 383–399.

Goswami, M., O'connor, K., Bhattarai, K., 2007. Development of regionalisation procedures using a multi-model approach for flow simulation in an ungauged catchment. J. Hydrol. 333 (2–4), 517–531.

Guo, Y., Zhang, Y., Zhang, L., Wang, Z., 2021. Regionalization of hydrological modeling for predicting streamflow in ungauged catchments: A comprehensive review. Wiley Interdiscip. Rev. Water 8 (1), e1487.

He, Y., Bárdossy, A., Zehe, E., 2011. A review of regionalisation for continuous streamflow simulation. Hydrol. Earth Syst. Sci. 15 (11), 3539–3553.

Hengl, T., Mendes de Jesus, J., Heuvelink, G.B., Ruiperez Gonzalez, M., Kilibarda, M., Blagotić, A., Shangguan, W., Wright, M.N., Geng, X., Bauer-Marschallinger, B., et al., 2017. SoilGrids250m: Global gridded soil information based on machine learning. PLoS One 12 (2), e0169748.

Heuvelmans, G., Muys, B., Feyen, J., 2006. Regionalisation of the parameters of a hydrological model: Comparison of linear regression models with artificial neural nets. J. Hydrol. 319 (1–4), 245–265.

Kavetski, D., Fenicia, F., Reichert, P., Albert, C., 2018. Signature-domain calibration of hydrological models using approximate Bayesian computation: Theory and comparison to existing applications. Water Resour. Res. 54 (6), 4059–4083.

Kay, A., Jones, D., Crooks, S., Kjeldsen, T., Fung, C., 2007. An investigation of site-similarity approaches to generalisation of a rainfall–runoff model. Hydrol. Earth Syst. Sci. 11 (1), 500–515.

Kokkonen, T.S., Jakeman, A.J., Young, P.C., Koivusalo, H.J., 2003. Predicting daily flows in ungauged catchments: model regionalization from catchment descriptors at the Coweeta Hydrologic Laboratory, North Carolina. Hydrol. Process. 17 (11), 2219–2238.

Kratzert, F., Klotz, D., Herrnegger, M., Sampson, A.K., Hochreiter, S., Nearing, G.S., 2019. Toward improved predictions in ungauged basins: Exploiting the power of machine learning. Water Resour. Res. 55 (12), 11344–11354.

Kumar, A., Singh, R., Jena, P.P., Chatterjee, C., Mishra, A., 2015. Identification of the best multi-model combination for simulating river discharge. J. Hydrol. 525, 313–325.

Li, W., Sankarasubramanian, A., Ranjithan, R., Sinha, T., 2016. Role of multimodel combination and data assimilation in improving streamflow prediction over multiple time scales. Stochastic Environ Res. Risk Assess. 30, 2255–2269.

McMillan, H., Westerberg, I., Branger, F., 2017. Five guidelines for selecting hydrological signatures. Hydrol. Process. 31 (26), 4757–4761.

Merz, R., Blöschl, G., 2004. Regionalisation of catchment model parameters. J. Hydrol. 287 (1–4), 95–123.

Mwakalila, S., 2003. Estimation of stream flows of ungauged catchments for river basin management. Phys. Chem. Earth Parts A/B/C 28 (20–27), 935–942.

Najafi, M.R., Moradkhani, H., 2016. Ensemble combination of seasonal streamflow forecasts. J. Hydrol. Eng. 21 (1), 04015043.

Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models part I—A discussion of principles. J. Hydrol. 10 (3), 282–290.

Oudin, L., Andréassian, V., Perrin, C., Michel, C., Le Moine, N., 2008. Spatial proximity, physical similarity, regression and ungaged catchments: A comparison of regionalization approaches based on 913 French catchments. Water Resour. Res. 44 (3).

Parajka, J., Viglione, A., Rogger, M., Salinas, J., Sivapalan, M., Blöschl, G., 2013. Comparative assessment of predictions in ungauged basins–Part 1: Runoff-hydrograph studies. Hydrol. Earth Syst. Sci. 17 (5), 1783–1795.

Poissant, D., Arsenault, R., Brissette, F., 2017. Impact of parameter set dimensionality and calibration procedures on streamflow prediction at ungauged catchments. J. Hydrol. Regional Stud. 12, 220–237.

Post, D.A., 2009. Regionalizing rainfall–runoff model parameters to predict the daily streamflow of ungauged catchments in the dry tropics. Hydrol. Res. 40 (5), 433–444.

Prieto, C., Le Vine, N., Kavetski, D., Fenicia, F., Scheidegger, A., Vitolo, C., 2022. An exploration of Bayesian identification of dominant hydrological mechanisms in ungauged catchments. Water Resour. Res. 58 (3), e2021WR030705.

Prieto, C., Le Vine, N., Kavetski, D., García, E., Medina, R., 2019. Flow prediction in ungauged catchments using probabilistic random forests regionalization and new statistical adequacy tests. Water Resour. Res. 55 (5), 4364–4392.

Razavi, T., Coulibaly, P., 2013. Streamflow prediction in ungauged basins: review of regionalization methods. J. Hydrol. Eng. 18 (8), 958–975.

Razavi, T., Coulibaly, P., 2016. Improving streamflow estimation in ungauged basins using a multi-modelling approach. Hydrol. Sci. J. 61 (15), 2668–2679.

Razavi, T., Coulibaly, P., 2017. An evaluation of regionalization and watershed classification schemes for continuous daily streamflow prediction in ungauged watersheds. Can. Water Resour. J. Rev. Can. Ressour. Hydriques 42 (1), 2–20.
Samuel, J., Coulibaly, P., Metcalfe, R.A., 2011. Estimation of continuous streamflow in ontario ungauged basins: comparison of regionalization methods. J. Hydrol. Eng. 16 (5), 447–459.
Science, U.S., Administration, E., 1975. Soil Taxonomy: A Basic System of Soil Classification for Making and Interpreting Soil Surveys. In: 1, (436), US Department of Agriculture.
Shepard, D., 1968. A two-dimensional interpolation function for irregularly-spaced data. In: Proceedings of the 1968 23rd ACM National Conference. pp. 517–524.
Sivapalan, M., Takeuchi, K., Franks, S., Gupta, V., Karambiri, H., Lakshmi, V., Liang, X., McDonnell, J., Mendiondo, E., O'connell, P., et al., 2003. IAHS decade on predictions in ungauged basins (PUB), 2003–2012: Shaping an exciting future for the hydrological sciences. Hydrol. Sci. J. 48 (6), 857–880.
Ssegane, H., Tollner, E., Mohamoud, Y., Rasmussen, T., Dowd, J., 2012. Advances in variable selection methods I: Causal selection methods versus stepwise regression and principal component analysis on data of known and unknown functional relationships. J. Hydrol. 438, 16–25.
Swain, J.B., Patra, K.C., 2017. Streamflow estimation in ungauged catchments using regionalization techniques. J. Hydrol. 554, 420–433.
Tarek, M., Arsenault, R., Brissette, F., Martel, J.-L., 2021. Daily streamflow prediction in ungauged basins: an analysis of common regionalization methods over the African continent. Hydrol. Sci. J. 1–17.
Viney, N.R., Bormann, H., Breuer, L., Bronstert, A., Croke, B.F., Frede, H., Gräff, T., Hubrechts, L., Huisman, J.A., Jakeman, A.J., et al., 2009. Assessing the impact of land use change on hydrology by ensemble modelling (LUCHEM) II: Ensemble combinations and predictions. Adv. Water Resour. 32 (2), 147–158.
Wagener, T., Sivapalan, M., Troch, P., Woods, R., 2007. Catchment classification and hydrologic similarity. Geogr. Compass 1 (4), 901–931.
Wang, G., Zhang, J., Jin, J., Liu, Y., He, R., Bao, Z., Liu, C., Li, Y., 2014. Regional calibration of a water balance model for estimating stream flow in ungauged areas of the Yellow River Basin. Quat. Int. 336, 65–72.
Westerberg, I.K., Wagener, T., Coxon, G., McMillan, H.K., Castellarin, A., Montanari, A., Freer, J., 2016. Uncertainty in hydrological signatures for gauged and ungauged catchments. Water Resour. Res. 52 (3), 1847–1865.
Yang, X., Magnusson, J., Huang, S., Beldring, S., Xu, C.-Y., 2020. Dependence of regionalization methods on the complexity of hydrological models in multiple climatic regions. J. Hydrol. 582, 124357.
Yin, Z., Liao, W., Lei, X., Wang, H., 2020. Parallel hydrological model parameter uncertainty analysis based on message-passing interface. Water 12 (10), 2667.
Young, A.R., 2006. Stream flow simulation within UK ungauged catchments using a daily rainfall-runoff model. J. Hydrol. 320 (1–2), 155–172.
Zhang, Y., Chiew, F., 2009. Evaluation of regionalisation methods for predicting runoff in ungauged catchments in southeast Australia. In: 18th World IMACS Congress and MODSIM09 International Congress on Modelling and Simulation. Modelling and Simulation Society of Australia and New Zealand and . . . , pp. 3442–3448.