Hydrology Research

Hydrology Research Vol 54 No 7, 885 doi: 10.2166/nh.2023.027

Regionalization methods for PUB: a comprehensive review of progress after the PUB decade

Xue Yang [a], Fengnian Li[a], Wenyan Qi[b,*], Mengyuan Zhang[a], Chengxi Yu[a] and Chong-Yu Xu[c]

[a] State Key Laboratory of Eco-Hydraulics in Northwest Arid Region, Xi'an University of Technology, Xi'an, Shaanxi 710048, China
[b] School of Civil Engineering, Lanzhou Jiao Tong University, Lanzhou 730070, China
[c] Department of Geosciences, University of Oslo, P.O. Box 1047 Blindern, Oslo N-0316, Norway
*Corresponding author. E-mail: wenyan.qi@lzjtu.edu.cn

XY, 0000-0003-1853-8573

ABSTRACT

This paper presents an updated review of model-dependent regionalization methods in hydrology since the PUB decade, incorporating new regions and methodological advancements. Two categories of regionalization methods are discussed: distance-based and regression-based, with various modification approaches. Several factors affecting the accuracy of regionalization performance are identified, including hydrological models, climate characteristics, data availability, and regionalization techniques. The review concludes that distance-based regionalization methods with an output averaging option from multiple donor catchments are the most statistically reliable, and a threshold of 0.5–0.8 for donor selection is optimal for performance. The parsimonious hydrological model is also recommended, particularly in data-limited contexts. Other insights include the effectiveness of the ensemble concept and limited impact of prior classification. Additionally, it is found that the general performance difference between climatic classes is larger than between methods, and regression-based methods may have large uncertainties in tropical regions. Overall, this study provides practical guidance for improving regionalization studies and advancing the field of hydrology.

Key words: hydrological model, prediction in ungauged basins (PUB), regionalization methods, runoff prediction

HIGHLIGHTS

- A deep analysis based on last decade's regionalization studies over the world.
- The parsimonious hydrological model seems more robust with limited prior knowledge of regionalization.
- Too high threshold (larger than 0.8) to donor selection results in performance deterioration.
- Ensemble concept could help to modify the performance, rather better than the prior classification.
- Climate plays an important role in regionalization.

1. INTRODUCTION

Streamflow is a fundamental component of the water cycle, and achieving reliable and continuous streamflow simulation is considered an essential objective within the hydrological community (Boughton & Droop 2003; Brocca *et al.* 2011; Cislaghi *et al.* 2020). The scientific and practical importance of reliable and continuous streamflow simulation cannot be overstated as it helps solve various engineering and water resource problems, such as reservoir design and management, predicting climatic and anthropogenic impacts on water resources, and risk assessment for extreme hydrological events (Parajka *et al.* 2013). Over decades of development and improvement, hydrological models, particularly rainfall-runoff models, have become the most common and useful tools for streamflow/runoff prediction and the estimation of water balance (Beven 2012). These models provide a visual representation of water systems using a set of equations that estimate the amount of runoff from rainfall based on various parameters (Devi *et al.* 2015). For practical applications, these models are calibrated to achieve optimal parameter values through the comparison of observations and simulations.

However, many catchments around the world are ungauged or poorly gauged, which presents a significant challenge (Sivapalan *et al.* 2003; Gorgoglione *et al.* 2019; Narbondo *et al.* 2020), especially in developing countries (Mulligan 2012).

Even more concerning is that global efforts to collect streamflow data are decreasing, particularly in headwater basins (Kirchner 2006; Davids *et al.* 2019). Without the necessary information to drive hydrological models, the calibration process cannot be performed (Oudin *et al.* 2008; He *et al.* 2011). As a result, simulating runoff in ungauged or poorly gauged basins becomes exceedingly difficult. To address this fundamental challenge in hydrological science that has persisted for decades, the International Association of Hydrological Sciences (IAHS) launched the 'Decade on Predictions in Ungauged Basins (PUB): 2003–2012' (Cislaghi *et al.* 2020). The PUB decade aimed to improve the capacity for reliable predictions in basins lacking observations (Sivapalan *et al.* 2003).

Regionalization is the most commonly used method for predicting ungauged streams in PUB studies (Hrachowitz *et al.* 2013; Arsenault *et al.* 2021). Blöschl & Sivapalan (1995) define regionalization as the process of transferring model parameters (MP) from similar catchments to the ungauged one of interest. Egbuniwe & Todd (1976) tested the transferability of MP between neighboring Kontagora and Malendo watersheds even before this definition was established. In recent years, with increased attention on PUB studies, the definition of regionalization has expanded. Tarek *et al.* (2021) describe how regionalization can be used to estimate MP for generating continuous streamflow in ungauged basins. Xu (1999) estimates MP in ungauged basins by relating them to catchment descriptors (CDs). In PUB research, various approaches have been proposed to estimate MP in ungauged basins, which fall into two categories: distance-based methods and regression-based methods (e.g., He *et al.* 2011; Pool *et al.* 2021). Distance-based methods transfer information, including MP, from hydrologically similar catchments using a similarity metric based on spatial location or catchment attributes (e.g., Parajka *et al.* 2005; Oudin *et al.* 2008; Yang *et al.* 2018). Regression-based methods relate individual MP to selected catchment properties (e.g., Skaugen *et al.* 2015; Yang *et al.* 2018) and use this relationship to estimate the MP in ungauged basins.

In order to determine the most suitable regionalization method for PUB research applications, a regional evaluation of various methods has been conducted. Arsenault *et al.* (2019) evaluated both physical catchment descriptor-based and spatial-proximity (distance-based) regionalization methods on 30 catchments in Mexico's semi-arid and humid regions. They concluded that the distance-based method provides more robust results compared to methods relying on physical CDs. Yang *et al.* (2020) compared both types of regionalization methods using four hydrological models on 86 basins in different climate sub-regions in Norway. They found that the physical similarity (distance-based) method consistently outperformed other methods across all models and sub-regions. A study conducted on 264 catchments in Canada by Arsenault & Brissette (2016) revealed that the preference for distance-based regionalization method depends on the number of donor catchments applied. The regression-based method was found to be more effective than the distance-based method with only one similar donor, but the distance-based method with multiple donors performed better.

The PUB decade has seen significant progress in improving regionalization methods while also identifying their limitations. For example, McIntyre *et al.* (2005) suggested using the entire model parameter set for distance-based regionalization methods, which preserves the integrity of the parameters as a set (e.g., Oudin *et al.* 2010; He *et al.* 2011). Livneh & Lettenmaier (2013) proposed replacing local parameters with zonal parameters when building a regression model with local CDs. Their approach was validated in experiments over 220 catchments in the United States.

In studies conducted in regions with high diversity and heterogeneity, there are significant differences in performance among different regionalization methods. Petheram *et al.* (2012) investigated the performance of regionalization methods in tropical and arid regions in Australia and found that the regionalization methods performed better in the tropical region than in the arid region. Bao *et al.* (2012) evaluated the performance of regionalization methods in three different types of hydro-climatic regions (i.e., humid, semi-humid, and arid regions). The results showed that the accuracy of the regression-based method decreased more significantly than the similarity-based method from humid to arid regions. Yang *et al.* (2020) assessed the performance of regionalization methods using 86 catchments over three climatic regions in Norway and found that regionalization performance declined from oceanic to continental and polar tundra regions. Qi *et al.* (2021) explored the performance dependence of different regionalization methods on climatic regions by using thousands of catchments under varying hydro-climatic conditions in North America. The results displayed that the regionalization performance was worst in the arid region, and the performance differences among methods were larger in the arid region than in others. Additionally, some studies investigated the relationship between the regionalization performance and the aridity index (i.e., an index that presents the degree of dryness of the climate of a site) and concluded that the regionalization performance decreases as the aridity index increases for all regionalization methods (Viglione *et al.* 2013; Guo *et al.* 2020).

Parajka *et al.* (2013) carried out an excellent comparative assessment of 3,874 catchments from 34 previous studies during the PUB decade (2003–2012) to summarize and explore the differences between various studies. They proposed a two-level comparison assessment and presented the first global perspective, which provides valuable reference and guidance for regionalization studies. It was observed that there has been an increasing number of regionalization studies since the PUB decade (2003–2012) (see Figure 1), with many of them being conducted in new regions (e.g., Althoff *et al.* 2022). Moreover, recent studies have included more catchments on average, located in relatively heterogeneous regions. From a global and statistical viewpoint, it is helpful to establish a more robust consensus from recent studies. Furthermore, many improvements have been made recently (e.g., Xu *et al.* 2022), making it timely to review the current state of knowledge and discuss the challenges that we are currently facing. Therefore, the objective of this study is to summarize what we have learned from recent studies and provide conclusions that can guide our future endeavors both in research and practical application. It is hoped that new research will be based systematically on previous research and fill gaps rather than merely reaffirming the consensus formed by prior work in the field (Xu 2021). To assess the new achievements of regionalization studies during the post-PUB decade (2013–2022) compared to those reviewed by Parajka *et al.* (2013), this study employs the same comparison

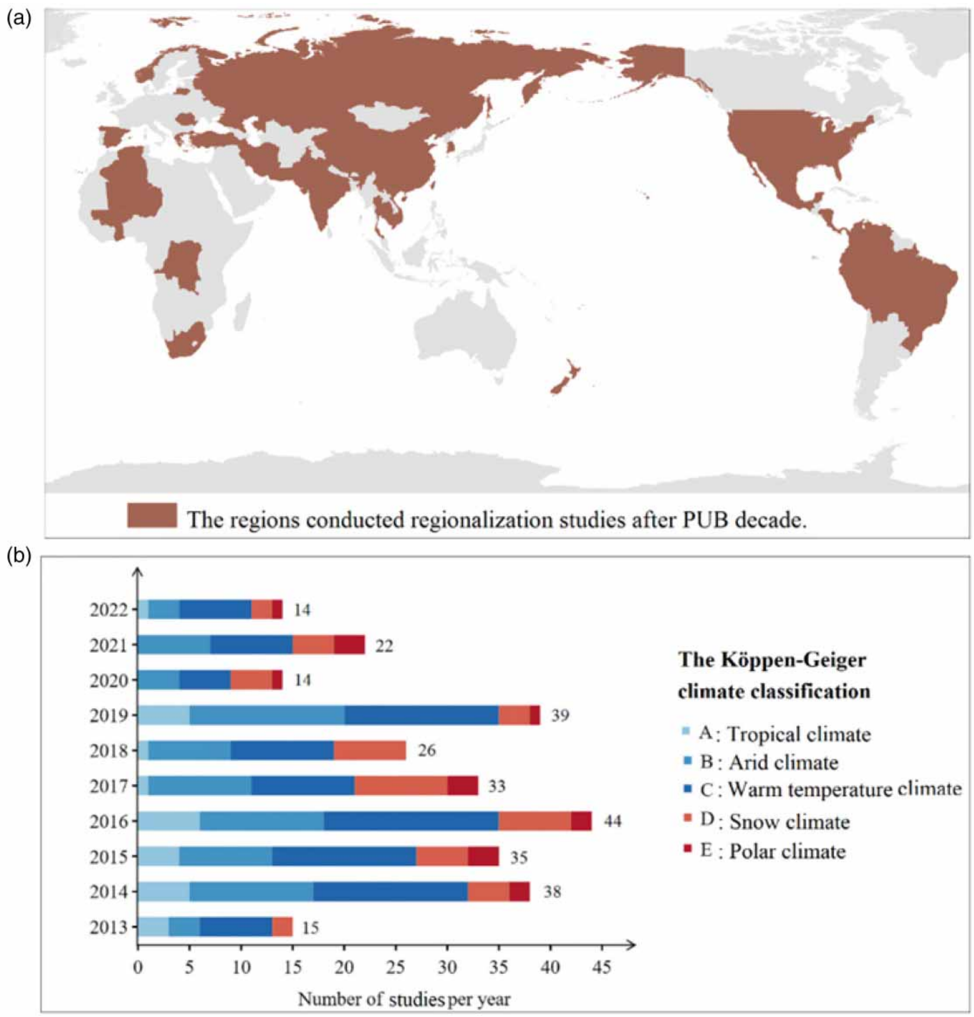

Figure 1 | The information of regionalization studies after PUB decade. (a) The region information where the first regionalization study was conducted after PUB decade (Note: The studies are searched from the Institute for Scientific Information (ISI) Web of Knowledge and Google Scholar including the terms 'hydrology', 'regionalization', and 'prediction in ungauged basins'. All the studies are then stratified at the country level based on their region(s)); (b) the annual number of regionalization studies over different Köppen–Geiger climate classes (Note: If there is more than one Köppen–Geiger climate class in a study, the number is increased to all these classes).

assessment method proposed by Parajka *et al.* (2013). The list of studies used for the statistical analysis is provided in the appendix table.

2. REGIONALIZATION METHODS

As mentioned previously, regionalization involves deriving model parameters in ungauged catchments from hydrological MP calibrated in gauged catchments. There are several regionalization methods, which can be broadly categorized into two classes: (1) distance-based methods (including both spatial-proximity and physical similarity (PS) methods); and (2) regression-based methods (which involve two steps). A more detailed description of each regionalization method is provided below.

2.1. Distance-based methods

The concept of catchment similarity forms the basis for applying distance-based regionalization methods in PUB (He *et al.* 2011). Distance-based regionalization methods can be classified into two categories based on the catchment properties used in catchment similarity evaluation: (a) spatial-proximity methods (SP); and (b) PS methods. The SP method utilizes geographical location information in catchment similarity evaluation, with the assumption that neighboring catchments exhibit similar behavior (e.g., Vandewiele *et al.* 1991). On the other hand, the PS method uses catchment attribute information to assess catchment similarity, such as mean elevation, forest cover types, and soil types (e.g., Kokkonen *et al.* 2003; Parajka *et al.* 2005; Samuel *et al.* 2011, 2012). The PS method assumes that catchments with similar physical characteristics exhibit similar hydrological responses (e.g., Kokkonen *et al.* 2003; McIntyre *et al.* 2005). In PUB research, both methods follow the same process, as depicted in Figure 2.

The first step in the regionalization process is to determine the 'donor catchment(s).' These are assessed based on hydrological MP that can be achieved through the calibration process and then transferred to ungauged basins for regionalization applications (Blöschl & Merz 2005). Different catchment attributes are considered in this step using both SP and PS methods. Geographical location information of catchments is used for the SP method, which is consistent across all studies, but the measure of distance can vary depending on whether the centroid or outlet location is used (Cislaghi *et al.* 2020). For the PS method, more catchment attributes are included, but these can differ between studies due to data availability and subjective experience (He *et al.* 2011). Land cover, area, mean annual precipitation, and mean elevation are the four most frequently applied catchment attributes identified through reviewed studies. However, similarity indices used to assess the

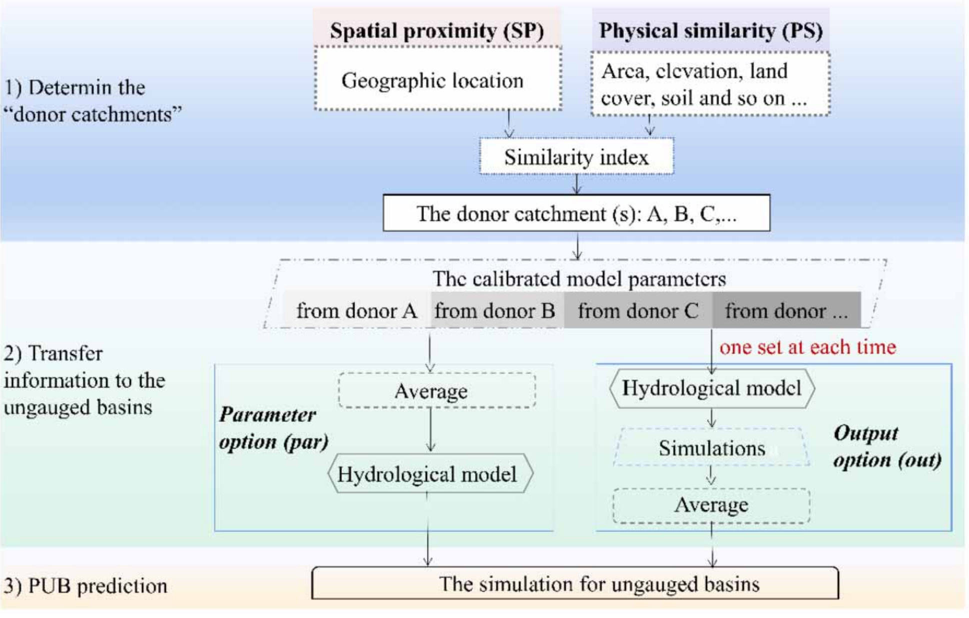

Figure 2 | The flow chart of distance-based regionalization methods.

similarity between gauged and ungauged basins, like Euclidean distance, can be the same for both SP and PS methods. Catchment similarity is crucial for distance-based regionalization methods, with smaller distances indicating higher similarity. Various similarity indicators for PS methods are presented by Guo *et al.* (2020). Based on similarity evaluation, gauged catchments with higher similarity are preferred as donor catchments, but the number of donor catchments selected is highly dependent on researchers' experience, which plays an essential role in performance during PUB regionalization (Guo *et al.* 2020; Qi *et al.* 2021). For instance, Yang *et al.* (2020) observed that the NSE (Nash–Sutcliffe Efficiency) difference between different numbers of donor catchments could be larger than 0.1.

The second step in the regionalization process involves transferring information from gauged basins to ungauged basins (see Figure 2). Based on previous studies and Figure 2, using multiple donor catchments is preferred for achieving good performance in PUB. When there are multiple donor catchments, two averaging options can be used to obtain final simulations for the target catchment (McIntyre *et al.* 2005). The first option is to average parameter values from the donor catchments, referred to as the parameter option (par). The second option is to average discharges simulated using an individual parameter set from each donor catchment, called the output option (out). Thus, four approaches are obtained by combining distance-based regionalization methods (SP and PS) with the two averaging options. After completing these two steps, streamflow simulation can be performed for ungauged basins.

2.2. Regression-based methods

The regression-based method is one of the most commonly used regionalization methods in hydrology (Beck *et al.* 2016; Guo *et al.* 2020), particularly two-step regression. This study focuses on discussing only two-step regression methods that utilize the established relationship between CDs and MP to derive parameter values for ungauged basins (Xu 1999, 2003; Timilsena & Piechota 2007; Beck *et al.* 2016). Many hydrological models have been found to be suitable for application across various climatic and physiographic conditions (Strömqvist *et al.* 2012; Hrachowitz *et al.* 2013) and have been widely used in PUB studies.

The regression-based regionalization method assumes that (1) a well-defined relationship exists between observable CDs and MP, and (2) the CDs used in regression provide information relevant to hydrological behavior at ungauged sites (Merz *et al.* 2006; Yang *et al.* 2018). Figure 3 illustrates the process of applying the regression-based method in PUB. The

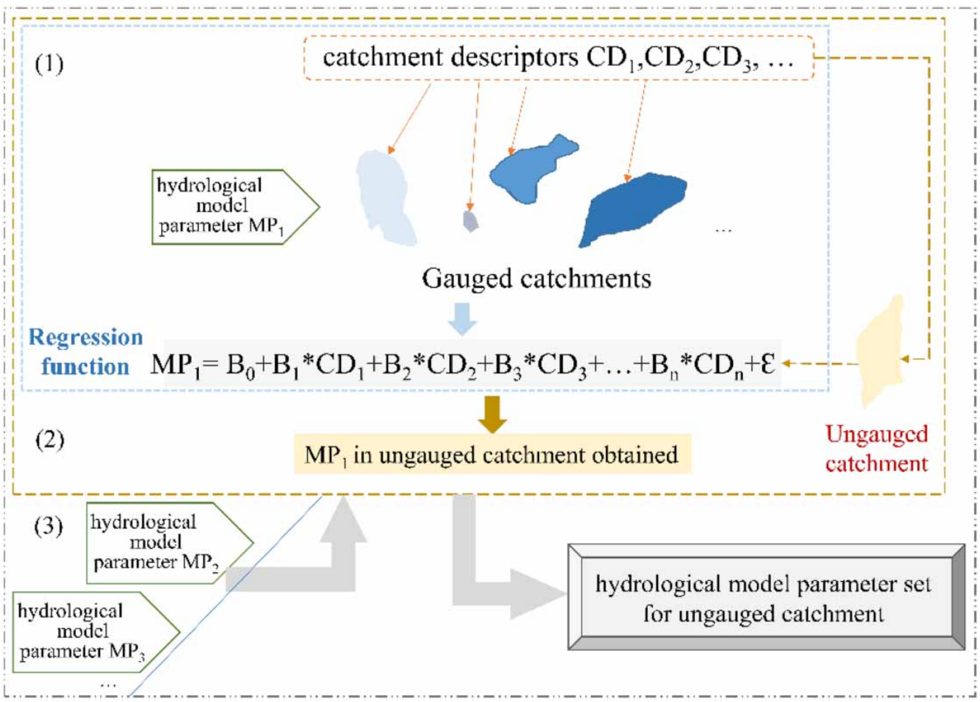

Figure 3 | The flow chart of regression-based regionalization methods.

first step involves constructing a regression relationship (function) for each model parameter using gauged catchments. Both linear and non-linear functions are possible and may differ depending on the regression method used and the CDs included in the function. Additionally, the types of CDs utilized in the regression function may vary among MP. The second step involves obtaining and applying CDs from ungauged catchments to the regression function to predict model parameter values. The final step is to repeat these processes for other MP until the entire set of MP for the ungauged catchment is obtained. Hydrological models can then be applied to simulate and predict streamflow.

2.3. Recent developments

To maximize the benefits and reduce the drawbacks of traditional methods in regionalization, researchers have proposed the concept of ensemble prediction as an alternative option for ungauged basins. For example, Waseem *et al.* (2015) introduced ensemble hydrological prediction (EHP) as an improved regionalization method to predict hydrological variables in ungauged basins. The basic idea of EHP is to give weights to traditional regionalization simulations based on their performance, which has been proven to be more robust in certain Pakistan catchments. Similarly, Razavi & Coulibaly (2016) proposed a multi-modelling approach to improve streamflow estimation in ungauged basins. Based on evaluations conducted on 90 Canadian watersheds, ensemble regionalization was found to perform better.

In regionalization research, similar hydrological responses are assumed within homogeneous groups (e.g., Hailegeorgis *et al.* 2015; Swain & Patra 2019). Recently, some studies have pre-classified catchments into groups to assess the impact. For example, Yang *et al.* (2018) grouped 118 Norwegian catchments, which were widely distributed, into five groups based on climatic indices but found no distinct effect. In contrast, Swain & Patra (2019) observed an improvement in regionalization performance in India when catchments were pre-classified, and they recommended this practice for future PUB studies. However, Pool *et al.* (2021) found that prior classification led to lower performance based on estimations across 671 U.S. catchments. This difference may be due to differences in classification methods or factors used in prior classification, and further exploration is needed to determine the reason behind this discrepancy.

Hydrological signatures, which describe the statistical or dynamical properties of hydrologic characteristics of catchments on different time scales, have recently received increased attention for regionalizing these signatures from donor catchments to ungauged basins. Atieh *et al.* (2017) proposed gene expression programming (GEP) and artificial neural networks (ANN) models to predict flow duration curves (FDCs) at ungauged basins across North America. Choubin *et al.* (2019) applied a comprehensive set of signatures to estimate streamflow in ungauged basins, which proved to be an efficient method.

To investigate whether poorly calibrated donor catchments could impair the performance of PUB simulations or not, some studies introduced threshold values of model efficiency during the calibration stage to filter donor selection (e.g., Arsenault & Brissette 2014; Qi *et al.* 2021). Arsenault & Brissette (2014) used a threshold of calibrated NSE values larger than 0.7 for donor selection over 268 Canadian catchments and found that well-calibrated donors would not guarantee good performance in PUB. Garambois *et al.* (2015) observed that poorly modeled catchments could also produce good performances when transferred to the ungauged catchment. Tarek *et al.* (2021) compared two thresholds (KGE (Kling Gupta Efficiency) values above 0.6 and 0.8 in the calibration, KGE is a widely used metric for evaluating the goodness of-fit of model simulations and corresponding observations proposed by Gupta *et al.* (2009)) as the threshold for donor catchments over 314 catchments in Africa. They detected decreased regionalization performance when applying the threshold 0.8, but no clear change when applying the threshold 0.6. Qi *et al.* (2021) set five thresholds for KGE (all, 0.6, 0.7, 0.8, 0.9) when evaluating the influence on the regionalization performance over 3,444 catchments in Canada and the US. The result showed deteriorated regionalization performance for higher thresholds (0.8 and 0.9) and no clear difference in regionalization performance for lower thresholds (0.6 and 0.7). In summary, setting too high a threshold value results in the use of too few donor catchments, which in turn causes a clear deterioration in regionalization performance.

3. THE COMPARISON ASSESSMENT

3.1. The donor influences

The main concept behind regionalization involves utilizing information from donor catchments, which are selected from a ‘donor pool’. Thus, the analysis of donor influence is conducted from two perspectives: (1) the size of the ‘donor pool’ (which equals the total number of basins included minus one), defined as data availability by Parajka *et al.* (2013), and (2) the quantity and quality aspects of the applied donor basins.

3.1.1. Results of the distance-based method

Figure 4 displays the performance of distance-based regionalization with respect to the size of the donor pool and the number of applied donor catchments. The performance is measured by the mean NSE value from globally distributed studies. Subplot (a) presents the NSE values from the reviewed studies, indicating that the use of multiple donor catchments reduces random errors compared to using a single donor (e.g., Arsenault *et al.* 2015; Guo *et al.* 2020). This is evident from the spread-out NSE values observed for the single donor case. The other subplots show the averaged result over the studies. The results indicate that employing multiple donors tends to yield higher accuracy than using a single donor for streamflow predictions in ungauged basins (see subplot (b)), consistent with many regional studies (e.g., Livneh & Lettenmaier 2013; Ergen & Kentel 2016; Li & Zhang 2017; Yang *et al.* 2018, 2020). Regarding the influence of the 'donor pool' size (see subplot (c)), the best and worst performances occur in the smallest and moderate groups, with mean NSE values of 0.66 and 0.53, respectively. The reasons for the best performance in the smallest group could be the close distance between catchments or the subjective selection of donors (e.g., Parajka *et al.* 2013). Relatively better performance for groups larger than 100 could be attributed to high streamflow network density (e.g., Oudin *et al.* 2008; Lebecherel *et al.* 2016). Regarding donor influences in both aspects (see subplot (d)), using two to three donors is generally suggested, since the NSE value is higher than 0.6

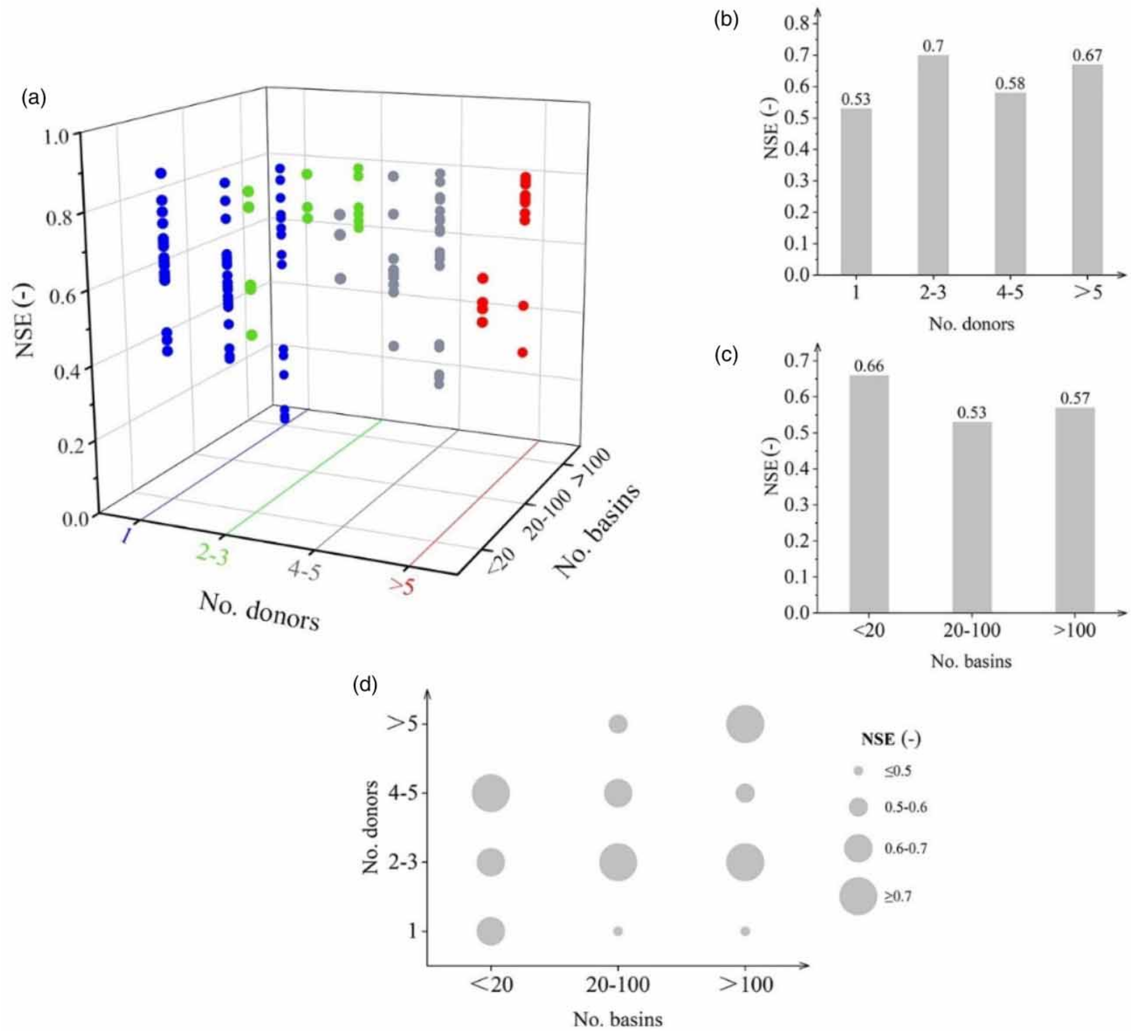

Figure 4 | The influence from the numbers of applied donor catchments (No. of donors) and the size of 'donor pool' (No. of basins) on the performance of distance-based regionalization methods. The scatter plot (a) shows the median NSE values from the studies; the bar chart and bubble plot (b–d) presents the averaged NSE values over the studies.

regardless of the number of basins included. This may be because a large number of donors imply that the used catchments are less and less similar to the target catchment, leading to a deterioration in performance (Neri *et al.* 2020). On the other hand, using a single donor catchment does not seem to be a good option if the donor pool size is larger than 20.

When regionalizing with multiple donors, there are four approaches that can be chosen. Figure 5 compares the mean NSE values from the reviewed studies and the following conclusions can be drawn. Firstly, the output option outperforms the parameter option with higher average NSE values, which is consistent with findings from many regionalization studies (e.g., Heng & Suetsugi 2014; Yang *et al.* 2018; Arsenault *et al.* 2019; Qi *et al.* 2021). Secondly, the performance difference between the SP and PS methods is marginal, supporting Xu's statement (2021), but the difference between studies in each group is significant, as found by Parajka *et al.* (2013). When considering the impact of MP, hydrological models with fewer parameters have higher median and mean NSE values, suggesting that parsimonious models are recommended to estimate streamflow in ungauged basins (e.g., Xu 1999; Oudin *et al.* 2008; Skaugen *et al.* 2015). The group of hydrological models with more than 12 parameters has a significantly larger performance difference. Therefore, greater uncertainty may arise if limited prior knowledge is available, which supports Arsenault *et al.*'s claim (2019). It should be noted that these findings are not conclusive due to the limited number of studies included in each group, particularly for the small parameter option groups.

In addition to the number of donor catchments, regionalization performance is also influenced by the 'quality' of these donors. Qi *et al.* (2022) investigated this relationship by setting seven different thresholds for donor selection and analyzing the impact on over 2,277 global catchments. They included four hydrological models and suggested five donor basins. Figure 6 illustrates how distance-based regionalization performance changes with the thresholds, using the average value across all hydrological models. The results show that regionalization performance (assessed by KGE, which is free of the influence of unhelpful interactions among components) deteriorated for higher thresholds (0.8 and 0.9), while no clear difference was observed for lower thresholds. This is primarily because a high threshold value results in a small 'donor pool' and low streamflow network density, as previously noted by Tarek *et al.* (2021) and Xu (2021). To further explore the influence of streamflow network density on regionalization performance, Lebecherel *et al.* (2016) conducted a study in France and observed that the SP method's performance decreased as the size of the 'donor pool' decreased. This finding was confirmed by Neri *et al.* (2020) in their study conducted in Austria.

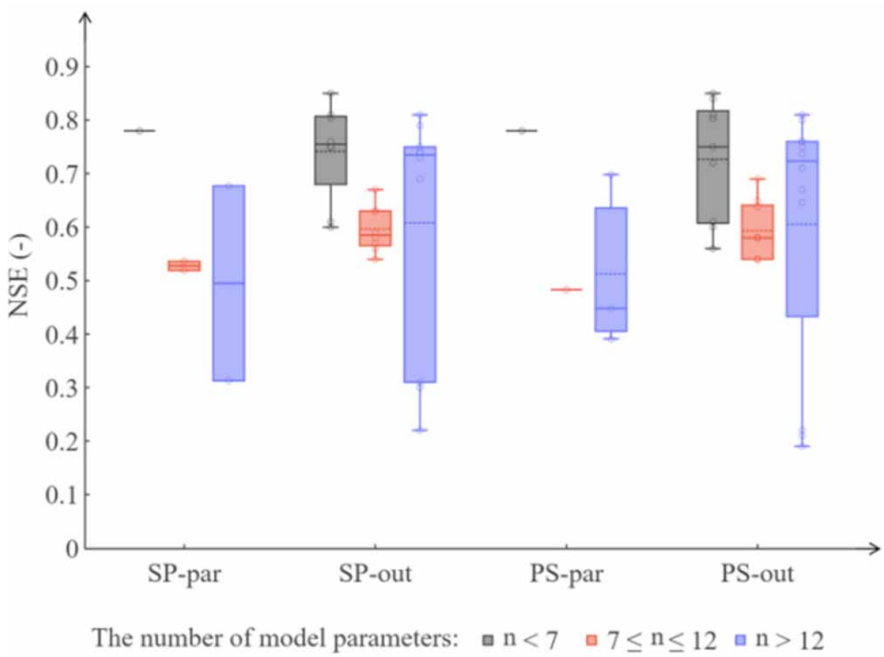

Figure 5 | The regionalization performance comparison between distance-based methods, with the influence from the number of hydrological MP. The solid and dash line in the box presents the median and mean value, respectively.

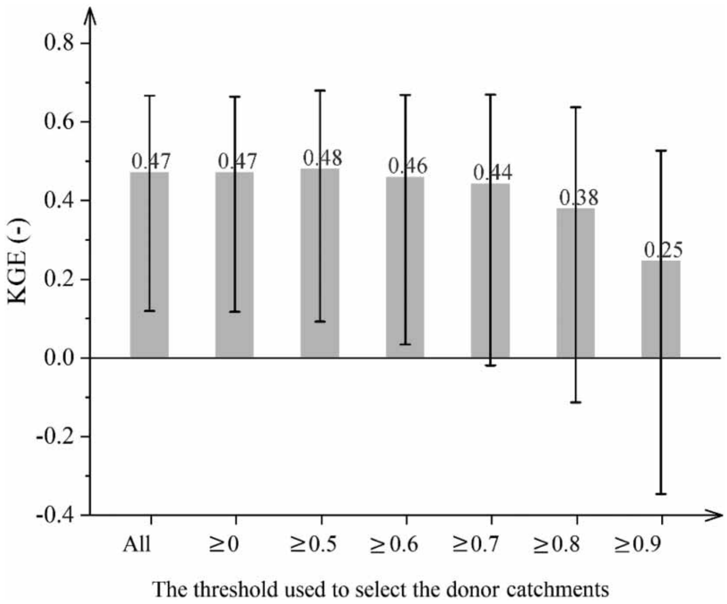

Figure 6 | The regionalization performance comparison with the influence of the threshold used to select the donor catchments. (The figure is reproduced from Qi *et al.* (2022) with permission).

3.1.2. The regression-based method results

For regression-based regionalization methods, the regression function is built on gauged basins, and the correlation between MP is ignored. The number of included catchments, or the size of the 'donor pool,' and the number of MP are considered as two important factors affecting regionalization performance. Figure 7 compares the performance of three groups with different numbers of MP and varied sizes of 'donor pool.'

Based on the scatter plot, it appears that many studies apply hydrological models with fewer parameters for regression methods, and most of them perform relatively well (with NSE values higher than 0.5). Along with the subplot (b), this result suggests that models with fewer parameters provide better performance based on the higher mean NSE value. This finding is in line with the distance-based method and some previous studies (e.g., Xu 2003; Clark *et al.* 2017). However, there are still 7 studies that yield NSE values higher than 0.6 using models with more than 12 parameters. Thus, we must

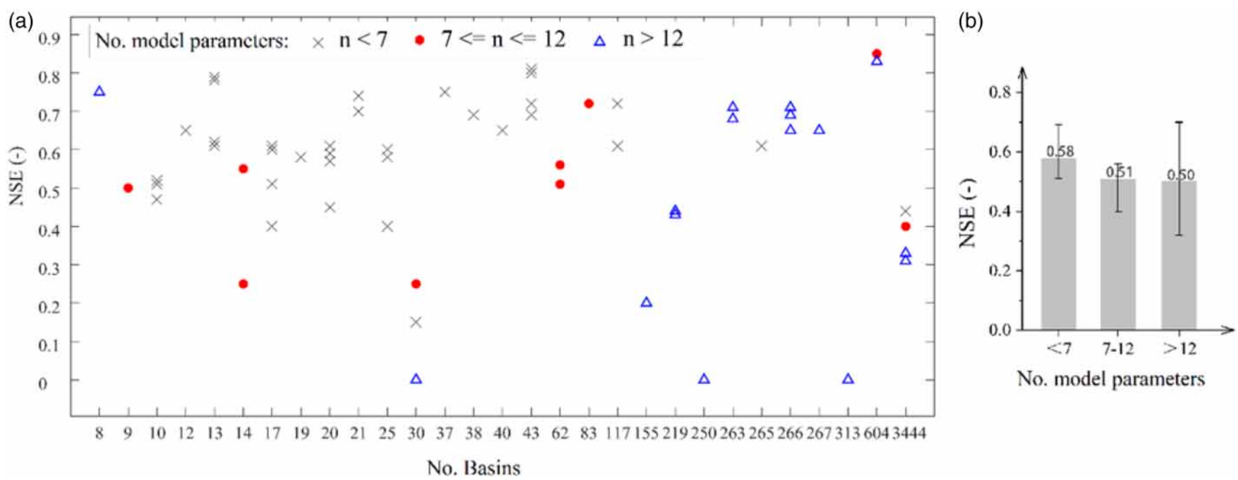

Figure 7 | The regression-based regionalization performance comparison between groups with different numbers of MP, with the consideration of the 'donor pool' size effect.

accept the notion that there is no inherent incentive to prefer a parsimonious hydrological model over a model with adequate complexity for regionalization studies (Arsenault *et al.* 2015, 2019; Poissant *et al.* 2017). Concerning the difference between studies, hydrological models with more than 12 parameters yield the largest statistically significant uncertainty. This result is consistent with the finding of Arsenault *et al.* (2019), which suggests that using a more complicated model may result in larger uncertainty. Therefore, simple hydrological models are suggested for application when there is a lack of pre-knowledge. However, models with adequate complexity are encouraged for application to improve regionalization performance in PUB.

The *x*-axis of the scatter plot displays the size of the 'donor pool.' There is no clear pattern observed in general, indicating that the performance is not dependent on the number of catchments, which is consistent with the findings of Yang *et al.* (2018) and Parajka *et al.* (2013). However, slightly higher NSE values appear when the 'donor pool' size is greater than 30 and parameters are less than 7. This improved performance with a larger 'donor pool' is related to a stronger regression relationship.

3.1.3. Comparison of distance-based and regression-based methods

Generally, most distance-based methods borrow information from less than 6 donors, while regression-based methods statistically require more than 20 donors to build the function. Therefore, only the size of the 'donor pool' is considered for comparison between the two methods (see Figure 8). At first glance, a clear performance difference occurs between the two regionalization methods over all three groups. For the group of total catchments less than 20, the distance-based method outperforms the regression-based method to a relatively large extent, with a median NSE value of about 0.65 and 0.52, respectively. However, the reverse finding is true for the middle group, where the median value of the distance-based method is only about 0.51, 0.08 lower than the regression-based method. Two factors are involved: one is that a clear lower performance was detected for the distance-based method with a larger dataset (Parajka *et al.* 2013); another is that many recent studies applied the regression-argument method or a non-linear regression approach to improve the stability of the relationship (e.g., Zhang *et al.* 2018). When considering the performance difference between studies, both regionalization methods display an increasing trend with the increased size of the donor pool.

3.2. Which method performs best?

This question has motivated numerous hydrology studies, and Figure 9 displays the statistical comparison between two regionalization methods: using one versus multiple donor catchments via distance-based methods. We are considering these methods separately here due to their large differences. In the distance-based method, almost 70% of the reviewed studies have mean NSE values lower than 0.6 when applying one donor catchment. In contrast, approximately 60% of the reviewed studies obtained high mean NSE values (greater than 0.6) when utilizing multiple donor catchments. For the regression-based method, about 46% of studies reviewed had a mean NSE value larger than 0.6, with generally moderate performance.

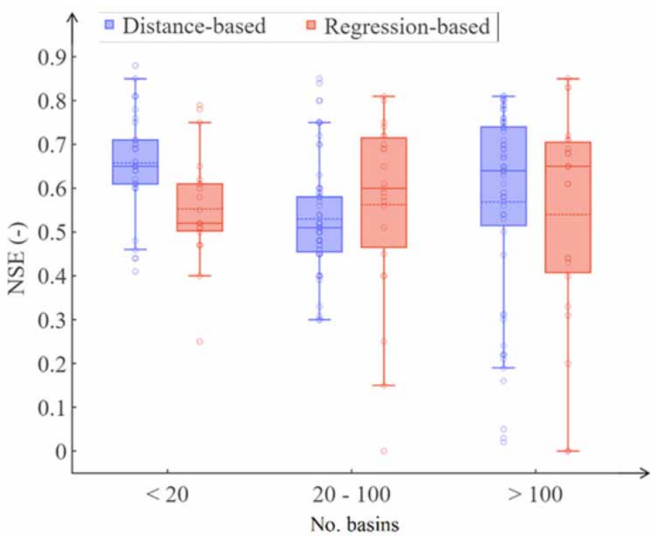

Figure 8 | The performance comparison between two methods with different 'donor pool' sizes. The solid and dash lines display the median and mean values, respectively.

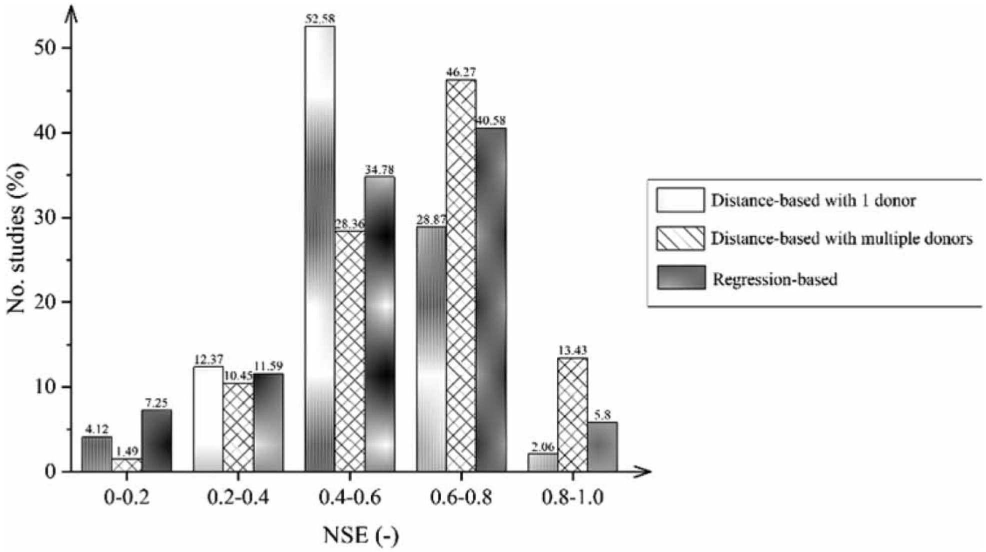

Figure 9 | The performance assessment between kinds of regionalization methods.

Therefore, we recommend using the distance-based method with multiple donor catchments, which is consistent with many regionalization studies (e.g., Qi *et al.* 2021; Tarek *et al.* 2021). The regression-based method may perform better than the distance-based method using one donor catchment due to recent modifications in regression-based methods (e.g., Livneh & Lettenmaier 2013; Hamel *et al.* 2017; Zhang *et al.* 2018).

With the advancement of PUB science and methodology, certain modifications have been proposed, classified mainly into three categories. Figure 10 presents the corresponding performance changes. Generally, the ensemble methods have shown a substantial improvement in performance. Including hydrological signatures showed relatively minor differences. Pre-classification, on the other hand, presented limited influence. Two notices should be taken into consideration: Firstly, pre-classification might result in worse performance as the 25% quantile value is lower than 0. Secondly, the considerable performance increase was mainly due to the very low performance before the modification made by Razavi & Coulibaly (2016). Thus, it is recommended to include hydrological signatures and use ensemble methods to enhance general regionalization performance for future PUB applications.

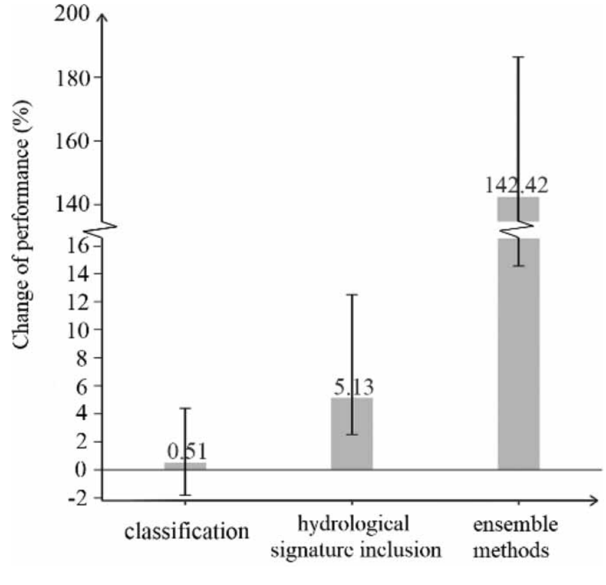

Figure 10 | The performance assessment between modified regionalization methods.

3.3. How does the climate affect the performance?

We conducted an assessment over different Köppen–Geiger climate classes to explore climatic influences, which are increasingly applied in PUB studies since the climate plays a significant role in hydrological response characteristics (Gao *et al.* 2016; Shrestha *et al.* 2021). Figure 11 shows the comparison results between climate classes from various aspects.

Regarding subplot (a), models with parameters higher than 12 for climatic classes A, B, and C yielded all the low NSE values, whereas averaged NSE values were visibly higher for the group of models with parameters less than 7 in classes D and E. This result tends to suggest using parsimonious models for relatively humid regions. When hydrological models include 7–12 parameters, the performance in class A is generally better than in other classes, whose NSE values are mostly higher than 0.5. Connecting the Köppen–Geiger climate classification with the map of FAO (Food and Agriculture Organization of the United Nations) aridity index (Spinoni *et al.* 2014) suggests using models with adequate complexity in relatively arid regions. In subplot (b), there are only five values for the regression-based method in Class A and E, which are statistically meaningless, and these results are excluded from the following comparison. According to the general performance assessed by mean NSE values (the dashed lines), the distance-based method is considered to outperform the regression-based method across all climate groups, especially for class C. This is consistent with global results and many regional evaluation studies (Yang *et al.* 2018; Zhu *et al.* 2021) but to a more pronounced extent. Comparing the difference in general performance between climate classes and between two regionalization methods, the former is larger, which confirms the statement that climatic influence is stronger than regionalization methods for PUB performance (Parajka *et al.* 2013; Yang *et al.* 2020). When assessing the distributions of NSE values over studies across all climate classes, the regression-based method displays a relatively larger spread. This means that the performance is more unpredictable or the prediction is in larger uncertainty if applying the regression-based method in these regions. Taking the interquartile range as the uncertainty measure, distinct differences exist between climate groups for both methods. Therefore, climatic impact is not ignorable in regionalization performance (Haaf *et al.* 2020; Qi *et al.* 2021; Xu 2021).

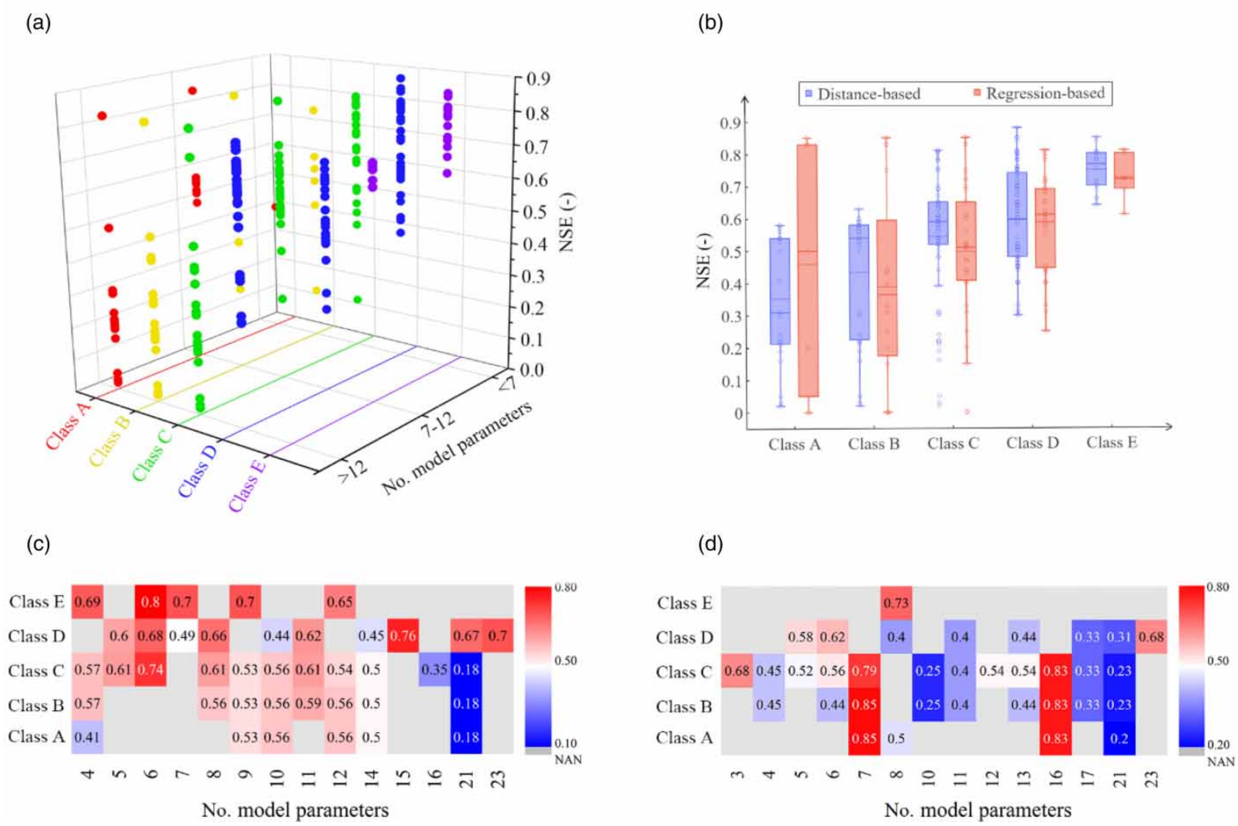

Figure 11 | The performance comparison over different Köppen–Geiger climate classes. (a) The original NSE values from reviewed studies; (b) boxplots for different regionalization methods, the solid and dashed lines present the median and mean values, respectively; (c,d) present the averaged NSE value over corresponding studies for distance-based method and regression-based method, respectively.

The subplots (c) and (d) present a comparison of mean NSE values with the corresponding number of MP used to represent model complexity (e.g., Parajka *et al.* 2013; Jehn *et al.* 2019). For distance-based methods, the performance decreases from class E to class A, with a corresponding decrease in aridity. This supports the finding that higher accuracy is generally achieved in humid rather than arid regions (e.g., Guo *et al.* 2020; Xu 2021), and that hydrological model capacity is considered a negative factor (e.g., Parajka *et al.* 2013; Gao *et al.* 2016). Regarding performance differences between studies across climatic classes, class C displays the largest difference, and very low performance appears for models with more than 15 parameters. Taking this difference as uncertainty, it means that randomly selecting a hydrological model for runoff simulation might produce a larger difference in runoff simulation for class C than for other classes. For regression-based methods, classes A and E are excluded from the discussion due to the few values, while more studies are encouraged in these classes. For the remaining classes, there is no distinct dependence between climatic classes and MP, but the chance of low accuracy for models with more than seven parameters is greater. However, several studies have obtained high NSE values (possibly over 0.83) when applying models with 16 parameters, which may be due to stronger physically relevant applied hydrological models that are clearly identifiable (Xu 2021). In comparing the performance between methods, model complexity seems to affect regression-based methods more than distance-based methods. This is considered mainly due to the influence of model parameter interaction on regression-based methods (e.g., Zhang & Chiew 2009; Yang *et al.* 2018).

4. CONCLUSIONS AND CHALLENGES

This paper compares the performance of two classes of regionalization methods in predicting runoff hydrographs in ungauged basins using the assessment method proposed by Parajka *et al.* (2013). However, the study is based on a completely different set of regionalization studies published after 2013. The review revealed that the accuracy of the regionalized PUB performance is affected by various factors, including hydrological models, climate characteristics, data availability, and regionalization approaches (e.g., Pool *et al.* 2021; Xu 2021). Nevertheless, based on comparative assessments and previous regionalization studies, some meaningful, consensus knowledge has been obtained.

- The distance-based regionalization method, with an output averaging option from multiple donor catchments, is statistically the most effective method. If the donor pool is larger than 20, it is not recommended to use one donor catchment.
- If prior knowledge is limited, the parsimonious hydrological model is a robust choice for future regionalization studies.
- The ensemble concept is the most attractive modification, and prior classification does not necessarily improve the performance of the study.
- Adding a low threshold (less than 0.5) to the donor selection has a limited impact, but using a threshold that is too high (larger than 0.8) results in performance deterioration.
- The total number of included catchments could affect the ranking of the two regionalization methods, and a larger performance difference may be observed with an increased number of total basins.
- The general performance difference between climatic classes is larger than between the two methods. Large uncertainties might exist if applying the regression-based method in tropical climatic regions.

Despite over two decades of development and numerous regionalization studies in public urban basin (PUB) research, there remain several challenges that must be addressed for future studies of regionalization in PUB:

(1) There are too few studies on the impact of station density on regionalization methods to summarize its global pattern. Until now, this objective has only been explored in some European regions located in relatively humid areas. As such, it is imperative to conduct studies in different regions to provide new insights and assess the differences among varying climatic and geographic regions.
(2) A comprehensive understanding of hydrological processes and regionalization methods is essential, especially in relatively arid areas. Many studies have found worse performance and high uncertainties in these regions. However, not enough research has been conducted to identify the factors that influence the final result. Answering this question is central to further improving regionalization methods, and it should be considered a meaningful future direction of research.
(3) Techniques that widen the availability of data for PUB science should be prioritized. The development of remote sensing has led to the emergence of various new data sets, and understanding how to apply these data to improve regionalization performance is an area that needs further research. Additionally, as big data and machine learning science continue to grow, it becomes more attractive to explore how to extract better value from limited observation data.

(4) Climate change has made it essential to predict streamflow in ungauged basins under non-stationary climates, and this requires more investigation. In recent years, a large number of hydrologic variable products, such as evaporation, soil moisture, and terrestrial water storage change, have been generated. These products provide diverse hydrologic data sets at finer spatial and temporal resolutions. Therefore, exploring how to use these data sets to carry out real-time correction on hydrological process variables to overcome the shortcomings of traditional regionalization methods under changing environmental conditions needs further investigation.

ACKNOWLEDGEMENT

The National Natural Science Foundation of China (Grant No. 52209035), the Education Department of Shaanxi Provincial Government (Project No. 21JT032), the Science and Technology Program of Gansu Province (22JR11RA143), and the Research Council of Norway (FRINATEK Project 274310) supported this work.

AUTHORSHIP CONTRIBUTION STATEMENT

Xue Yang conceptualized the whole article, developed the methodology, conducted funding acquisition, and wrote the original draft. Fengnian Li and Wenyan Qi rendered support in data curation, analysis, investigation, and wrote the review and edited the article. Mengyuan Zhang and Chengxi Yu brought resources, rendered support in data curation, analysis, and reviewed and edited the article. Chong-Yu Xu rendered support in the initiation of the study, funding acquisition, and review and editing.

DATA AVAILABILITY STATEMENT

All relevant data are included in the paper or its Supplementary Information.

CONFLICT OF INTEREST

The authors declare there is no conflict.

REFERENCES

Althoff, D., Rodrigues, L. N. & David, D. 2022 Predicting runoff series in ungauged basins of the Brazilian Cerrado biome. *Environ. Modell. Software* **149**, 105315.

Arsenault, R. & Brissette, P. F. 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.

Arsenault, R. & Brissette, P. F. 2016 Comparison of climate datasets for lumped hydrological modeling over the continental United States. *J. Hydrol.* **537**, 334–345.

Arsenault, R., Poissant, D. & Brissette, F. 2015 Parameter dimensionality reduction of a conceptual model for streamflow prediction in Canadian, snowmelt dominated ungauged basins. *Adv. Water Resour.* **85**, 27–44.

Arsenault, R., Dufour, M. B., Poulin, A., Dallaire, G. & Lopez, R. 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., Tarek, M., Brissette, F. & Martel, J. 2021 Daily streamflow prediction in ungauged basins: an analysis of common regionalization methods over the African continent. *Hydrol. Sci. J.* **66** (11), 1695–1711.

Atieh, M., Taylor, G., Sattar, A. M. A. & Gharabaghi, B. 2017 Prediction of flow duration curves for ungauged basins. *J. Hydrol.* **545**, 383–394.

Bao, Z., Zhang, J., Liu, J., Fu, G., Wang, G., He, R., Yan, X., Jin, J. & Liu, H. 2012 Comparison of regionalization approaches based on regression and similarity for predictions in ungauged catchments under multiple hydro-climatic conditions. *J. Hydrol.* **466–467**, 37–46.

Beck, H. E., Dijk, A. I., Roo, A. D., 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. J. 2012 *Rainfall-Runoff Modelling: The Primer*, 2nd edn. Wiley-Blackwell, Chichester, West Sussex.

Blöschl, G. & Merz, R. 2005 Flood frequency regionalization – spatial proximity vs. catchment attributes. *J. Hydrol.* **302** (1–4), 283–306.

Blöschl, G. & Sivapalan, M. 1995 Scale issues in hydrological modelling: a review. *Hydrol. Processes* **9** (3–4), 251–290.

Boughton, W. & Droop, O. 2003 Continuous simulation for design flood estimation – a review. *Environ. Modell. Software* **18** (4), 309–318.

Brocca, L., Melone, F. & Moramarco, T. 2011 Distributed rainfall-runoff modelling for flood frequency estimation and flood forecasting. *Hydrol. Processes* **25** (18), 2801–2813.

Choubin, B., Solaimani, K., Rezanezhad, F., Roshan, M. H., Malekian, A. & Shamshirband, S. 2019 Streamflow regionalization using a similarity approach in ungauged basins: application of the geo-environmental signatures in the Karkheh River Basin, Iran. *Catena* **182**, 15.

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.

Clark, G. E., Ahn, K. H. & Palmer, R. N. 2017 Assessing a regression-based regionalization approach to ungauged sites with various hydrologic models in a forested catchment in the northeastern United States. *J. Hydrol. Eng.* **22** (12), 05017027.

Davids, J. C., Rutten, M. M., Pandey, A., Devkota, N., van Oyen, W. D., Prajapati, R. & van de Giesen, N. 2019 Citizen science flow – an assessment of simple streamflow measurement methods. *Hydrol. Earth Syst. Sci.* **23** (2), 1045–1065.

Devi, K. G., Ganasri, P. B. & Dwarakish, S. G. 2015 A review on hydrological models. *Aquatic Procedia* **4**, 1001–1007.

Egbuniwe, N. & Todd, D. K. 1976 Application of the Stanford watershed model to Nigerian watersheds. *J. Am. Water Resour. Assoc.* **12** (3), 449–460.

Ergen, K. & Kentel, E. 2016 An integrated map correlation method and multiple-source sites drainage-area ratio method for estimating streamflows at ungauged catchments: a case study of the Western Black Sea Region, Turkey. *J. Environ. Manage.* **166**, 309–320.

Gao, G., Fu, B., Wang, S., Liang, W. & Jiang, X. 2016 Determining the hydrological responses to climate variability and land use/cover change in the Loess Plateau with the Budyko framework. *Sci. Total Environ.* **557–558**, 331–342.

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.

Gorgoglione, A., Gioia, A. & Iacobellis, V. 2019 A framework for assessing modeling performance and effects of rainfall-catchment-drainage characteristics on nutrient urban runoff in poorly gauged watersheds. *Sustainability* **11** (18), 4933.

Guo, Y., Zhang, Y., Zhang, L. & Wang, Z. 2020 Regionalization of hydrological modeling for predicting streamflow in ungauged catchments: a comprehensive review. *Wiley Interdiscip. Rev. Water* **8** (5), 32.

Gupta, H. V., Kling, H., Yilmaz, K. K. & Martinez, G. F. 2009 Decomposition of the mean squared error and NSE performance criteria: implications for improving hydrological modelling. *J. Hydrol.* **377** (1), 80–91.

Haaf, E., Giese, M., Heudorfer, B., Stahl, K. & Barthel, R. 2020 Physiographic and climatic controls on regional groundwater dynamics. *Water Resour. Res.* **56** (10), e2019WR026545.

Hailegeorgis, T. T., Abdella, Y. S., Alfredsen, K. & Kolberg, S. 2015 Evaluation of regionalization methods for hourly continuous streamflow simulation using distributed models in boreal catchments. *J. Hydrol. Eng.* **20** (11), 04015028.

Hamel, P., Guswa, A. J., Sahl, J. & Zhang, L. 2017 Predicting dry-season flows with a monthly rainfall-runoff model: performance for gauged and ungauged catchments. *Hydrol. Processes* **31** (22), 3844–3858.

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

Heng, S. & Suetsugi, T. 2014 Comparison of regionalization approaches in parameterizing sediment rating curve in ungauged catchments for subsequent instantaneous sediment yield prediction. *J. Hydrol.* **512**, 240–253.

Hrachowitz, M., Savenije, H. H. G., Blöschl, G., McDonnell, J. J., Sivapalan, M., Pomeroy, J. W., Arheimer, B., Blume, T., Clark, M. P., Ehret, U., Fenicia, F., Freer, J. E., Gelfan, A., Troch, P. A., Uhlenbrook, S., Wagener, T., Winsemius, H. C., Wood, R. S., Zehe, E. & Cudennec, C. 2013 A decade of predictions in ungauged basins (PUB) – a review. *Hydrolog. Sci. J.* **58** (6), 1198–1255.

Jehn, F. U., Chamorro, A., Houska, T. & Breuer, L. 2019 Trade-offs between parameter constraints and model realism: a case study. *Sci. Rep.* **9** (1), 1–12.

Kirchner, J. W. 2006 Getting the right answers for the right reasons: linking measurements, analyses, and models to advance the science of hydrology. *Water Resour. Res.* **42** (3), W03S04.

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. Processes* **17** (11), 2219–2238.

Lebecherel, L., Andréassian, V. & Perrin, C. 2016 On evaluating the robustness of spatial-proximity-based regionalization methods. *J. Hydrol.* **539**, 196–203.

Li, H. & Zhang, Y. 2017 Regionalising rainfall-runoff modelling for predicting daily runoff: comparing gridded spatial proximity and gridded integrated similarity approaches against their lumped counterparts. *J. Hydrol.* **550**, 279–293.

Livneh, B. & Lettenmaier, D. P. 2013 Regional parameter estimation for the unified land model. *Water Resour. Res.* **49** (1), 100–114.

McIntyre, N., Lee, H., Wheater, H., Young, A. & Wagener, T. 2005 Ensemble predictions of runoff in ungauged catchments. *Water Resour. Res.* **41** (12), 1–14.

Merz, R., Blöschl, G. & Parajka, J. 2006 Regionalisation methods in rainfall-runoff modelling using large samples. *Lr. Model Param. Exp. IAHS Publ.* **307**, 117–125.

Mulligan, M. 2012 Waterworld: a self-parameterising, physically based model for application in data-poor but problem-rich environments globally. *Hydrol. Res.* **44** (5), 748–769.

Narbondo, S., Gorgoglione, A., Crisci, M. & Chreties, C. 2020 Enhancing physical similarity approach to predict runoff in ungauged watersheds in sub-tropical regions. *Water* **12** (2), 528.

Neri, M., Parajka, J. & Toth, E. 2020 Importance of the informative content in the study area when regionalising rainfall-runoff model parameters: the role of nested catchments and gauging station density. *Hydrol. Earth Syst. Sci.* **24**, 5149–5171.

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), 1–15.

Oudin, L., Kay, A., Andréassian, V. & Perrin, C. 2010 Are seemingly physically similar catchments truly hydrologically similar? *Water Resour. Res.* **46** (11), 1–15.

Parajka, J., Merz, R. & Blöschl, G. 2005 A comparison of regionalisation methods for catchment model parameters. *Hydrol. Earth Syst. Sci.* **9** (3), 157–171.

Parajka, J., Viglione, A., Rogger, M., Salinas, J. L., 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.

Petheram, C., Rustomji, P., Chiew, F. H. S. & Vleeshouwer, J. 2012 Rainfall-runoff modelling in northern Australia: a guide to modelling strategies in the tropics. *J. Hydrol* **462–463**, 28–41.
Poissant, D., Arsenault, R. & Brissette, F. 2017 Impact of parameter set dimensionality and calibration procedures on streamflow prediction at ungauged catchments. *J. Hydrol.: Reg. Stud.* **12**, 220–237.
Pool, S., Vis, M. & Seibert, J. 2021 Regionalization for ungauged catchments-lessons learned from a comparative large-sample study. *Water Resour. Res.* **57** (10), e2021WR030437.
Qi, W., Chen, J., Li, L., Xu, C., Xiang, Y., Zhang, S. & Wang, H. 2021 Impact of the number of donor catchments and the efficiency threshold on regionalization performance of hydrological models. *J. Hydrol.* **601**, 126680.
Qi, W., Chen, J., Li, L., Xu, C., Li, J., Xiang, Y. & Zhang, S. 2022 Regionalization of catchment hydrological model parameters for global water resources simulations. *Hydrol. Res.* **53** (3), 441–466.
Razavi, T. & Coulibaly, P. 2016 Improving streamflow estimation in ungauged basins using a multi-modelling approach. *Hydrol. Sci. J.* **61** (15), 2668–2679.
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.
Samuel, J., Coulibaly, P. & Metcalfe, R. A. 2012 Evaluation of future flow variability in ungauged basins: validation of combined methods. *Adv. Water Resour.* **35**, 121–140.
Shrestha, S., Bae, D. H., Hok, P., Ghimire, S. & Pokhrel, Y. 2021 Future hydrology and hydrological extremes under climate change in Asian river basins. *Sci. Rep.* **11** (1), 17089.
Sivapalan, M., Takeuchi, K., Franks, S. W., Gupta, V. K., Karambiri, H., Lakshmi, V., Liang, X., McDonnell, J. J., Mendionndo, E. M., O'Connell, P. E., Oki, T., Pomeroy, J. W., Schertzer, D., Uhlenbrook, S. & Zehe, E. 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.
Skaugen, T., Peerebom, I. O. & Nilsson, A. 2015 Use of a parsimonious rainfall-runoff model for predicting hydrological response in ungauged basins. *Hydrol. Processes* **29** (8), 1999–2013.
Spinoni, J., Naumann, G., Carrao, H., Barbosa, P. & Vogt, J. 2014 World drought frequency, duration, and severity for 1951–2010. *Int. J. Climatol.* **34** (8), 2792–2804.
Strömqvist, J., Arheimer, B., Dahné, J., Donnelly, C. & Lindström, G. 2012 Water and nutrient predictions in ungauged basins: set-up and evaluation of a model at the national scale. *Hydrolog. Sci. J.* **57** (2), 229–247.
Swain, J. B. & Patra, K. C. 2019 Impact assessment of land use/land cover and climate change on streamflow regionalization in an ungauged catchment. *J. Water Clim. Change* **10** (3), 554–568.
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.* **66**, 1695–1711.
Timilsena, J. & Piechota, T. 2007 Regionalization and reconstruction of snow water equivalent in the upper Colorado River basin. *J. Hydrol.* **352** (1–2), 94–106.
Vandewiele, G. L., Xu, C. Y. & Huybrechts, W. 1991 Regionalisation of physically-based water balance models in Belgium. Application to ungauged catchments. *Water Resour. Res.* **5** (3–4), 199–208.
Viglione, A., Parajka, J., Rogger, M., Salinas, J. L., Laaha, G., Sivapalan, M. & Blöschl, G. 2013 Comparative assessment of predictions in ungauged basins – part 3: runoff signatures in Austria. *Hydrol. Earth. Syst. Sci.* **17** (6), 2263–2279.
Waseem, M., Ajmal, M. & Kim, T.-W. 2015 Ensemble hydrological prediction of streamflow percentile at ungauged basins in Pakistan. *J. Hydrol.* **525**, 130–137.
Xu, C. Y. 1999 Estimation of parameters of a conceptual water balance model for ungauged catchments. *Water Resour. Manage.* **13** (5), 353–368.
Xu, C. Y. 2003 Testing the transferability of regression equations derived from small sub-catchments to a large area in central Sweden. *Hydrol. Earth Syst. Sci.* **7** (3), 317–324.
Xu, C. Y. 2021 Issues influencing accuracy of hydrological modeling in a changing environment. *Water Sci. Eng.* **14** (2), 167–170.
Xu, Z., Wu, Z., Shao, Q., He, H. & Guo, X. 2022 A two-step calibration framework for hydrological parameter regionalization based on streamflow and remote sensing evapotranspiration. *J. Hydrol.* **613**, 128320.
Yang, X., Magnusson, J., Rizzi, J. & Xu, C. Y. 2018 Runoff prediction in ungauged catchments in Norway: comparison of regionalization approaches. *Hydrol. Res.* **49** (2), 487–505.
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.
Zhang, Y. & Chiew, F. 2009 Evaluation of regionalization methods for predicting runoff in ungauged catchments in Southeast Australia. In: *18th World IMACS/MODSIM Congress*, Cairns, Australia, pp. 3442–3448.
Zhang, R., Liu, J., Gao, H. & Mao, G. 2018 Can multi-objective calibration of streamflow guarantee better hydrological model accuracy? *J. Hydroinf.* **20** (3), 687–698.
Zhu, Y., Liu, L., Qin, F., Zhou, L., Zhang, X., Chen, T., Li, X. & Ao, T. 2021 Application of the regression-augmented regionalization approach for BTOP model in ungauged basins. *Water* **13** (16), 2294.

First received 2 March 2023; accepted in revised form 12 June 2023. Available online 13 July 2023