Regionalization of hydrological model parameters using gradient boosting machine

Zhihong Song[1,2], Jun Xia[1,2,3], Gangsheng Wang[1,2,4], Dunxian She[1,2], Chen Hu[1,2], Si Hong[1,2]

[1]State Key Laboratory of Water Resources and Hydropower Engineering Science, Wuhan University, Wuhan, 430072, China

[2]Hubei Key Laboratory of Water System Science for Sponge City Construction, Wuhan University, Wuhan, 430072, China

[3]Key Laboratory of Water Cycle and Related Land Surface Processes, Chinese Academy of Sciences, Beijing, 10010, China

[4]Institute for Water-Carbon Cycles and Carbon Neutrality, Wuhan University, Wuhan, 430072, China

Correspondence to: Jun Xia (xiajun666@whu.edu.cn) and Gangsheng Wang (wanggs@whu.edu.cn)

Abstract. Regionalization of hydrological model parameters is key to hydrological predictions in ungauged basins. The commonly used multiple linear regression (MLR) method may not be applicable in complex and nonlinear relationships between model parameters and watershed properties. Moreover, most regionalization methods assume lumped parameters for each catchment without considering within-catchment heterogeneity. Here we incorporated the Penman-Monteith-Leuning (PML) equation into the Distributed Time-Variant Gain Model (DTVGM) to improve the mechanistic representation of the evapotranspiration (ET) process. We calibrated six key model parameters grid-by-grid across China using a multivariable calibration strategy, which incorporates spatiotemporal runoff and ET datasets (0.25°, monthly) as reference. In addition, we used the gradient boosting machine (GBM), a machine learning technique, to portray the dependence of model parameters on soil and terrain attributes in four distinct climatic zones across China. We show that the modified DTVGM could reasonably estimate the runoff and ET over China using the calibrated parameters, but performed better in humid than arid regions for the validation period. The regionalized parameters by the GBM method exhibited better spatial coherence relative to the calibrated grid-by-grid parameters. In addition, GBM outperformed the stepwise MLR method in both parameter regionalization and gridded runoff simulations at national scale, though the improvement is not significant pertaining to watershed streamflow validation due to most of the watersheds being located in humid regions. We also revealed that the slope, saturated soil moisture content, and elevation are the most important explanatory variables to inform model parameters based on the GBM approach. The machine-learning-based regionalization approach provides an effective alternative to deriving hydrological model parameters from watershed properties, particularly in ungauged regions.

Keywords: Regionalization; Gradient Boosting Machine; Distributed hydrological model; Soil; Terrain;

1 Introduction

Hydrological modeling can provide quantitative extrapolation or prediction of runoff and water balance (Beven, 2011; He et al., 2011), which serves as the basis for water management for human livelihood, agriculture, industry, and environment (Hobeichi et al., 2018; Montanari et al., 2013; Parajka et al., 2013; Zhang et al., 2020). Hydrological models often require streamflow and/or other observations to calibrate parameters (Beck et al., 2020). However, it is difficult to parameterize a hydrological model at large scales (e.g., from national to global) or remote regions due to the sparse or the lack of observing stations. Under such circumstances, attempts have been made to use reanalysis datasets, not the observations, for model calibration and validation (Bai et al., 2018; Dembélé et al., 2020; Huang et al., 2020; Immerzeel and Droogers, 2008; Zhang et al., 2020). For example, Dembélé et al., (2020) developed a novel multivariate calibration framework on spatial patterns by combining streamflow with satellite datasets, including evapotranspiration (ET), soil moisture, and terrestrial water storage to parameterize a distributed hydrological model. Zhang et al., (2020) and Huang et al., (2020) demonstrated encouraging potential in the calibration of hydrological models solely against remote sensed ET data (or bias-corrected remote sensed data) without the need for observed streamflow data (i.e., the runoff-free calibration approach) to predict runoff in ungauged basins. The current study attempted to use spatiotemporal ET and runoff data for grid-by-grid calibration of a distributed hydrological model.

Hydrological models generally rely on regionalization methods to tackle the Predictions in Ungauged Basins (PUBs) by transferring information from gauged to ungauged catchments (He et al., 2011; Parajka et al., 2013; Razavi and Coulibaly, 2013). Regionalization of hydrological parameters generally includes three categories, similarity-based, regression-based, and hydrological signatures-based (Guo et al., 2021). The similarity-based regionalization presumes that the catchments with similar characteristics have the same hydrological response, such as the spatial proximity (Oudin et al., 2008; Parajka et al., 2005; Samuel et al., 2011; Vandewiele and Elias, 1995), and the physical similarity (Beck et al., 2016; Oudin et al., 2010; Yang et al., 2018; Zhang and Chiew, 2009). The regression-based method aims to establish the regression relationship between hydrological parameters and catchment characteristics (e.g., soil, topography, and climate variables), which helps to estimate model parameters in ungauged regions (Hundecha and Bárdossy, 2004; Livneh and Lettenmaier, 2013; Xu, 1999; Young, 2006). In addition,

some researchers have managed to transplant hydrological signatures (runoff depth, runoff ratio, flow percentile, flood frequency, baseflow index, flow change rate, etc.) from gauged catchments to ungauged basins (Castiglioni et al., 2010; Oubeidillah et al., 2014; Yang et al., 2019).

Among the most utilized regionalization techniques are probably regressions between the model parameters and physiographic catchment attributes as they are simple, fast, and intuitive (Bao et al., 2012; Heuvelmans et al., 2006; Oudin et al., 2008; Pagliero et al., 2019; Parajka et al., 2005; Razavi and Coulibaly, 2013; Young, 2006). Typically, the multiple linear regression (MLR) is widely used to estimate model parameters (Pagliero et al., 2019; Parajka et al., 2005; Sefton and Howarth, 1998). However, there are several limitations for regression approach, including fewer representative results of linear regression due to multicollinearity in catchment attributes, a high correlation between explanatory variables, complex and nonlinear relationships with high nonstationarity between physical catchment descriptors and model parameters (Blöschl, 2005; Guo et al., 2021; Kuczera and Mroczkowski, 1998; Pagliero et al., 2019; Yang et al., 2020; Zhang et al., 2018).

Machine learning techniques provide an alternative to overcome these issues for the linear regression approaches. For example, Sun et al., (2014) found that the Gaussian process regression is superior to traditional linear regression and artificial neural network models in most cases for probabilistic streamflow forecasting in 438 catchments across the United States. Zhang et al., (2018) assessed the regression tree ensemble approach compared with MLR, log-transformed MLR, and hydrological modeling in 605 catchments across Australia, which outperforms the two linear regressions in predicting signatures of flow dynamics. Prieto et al., (2019) implemented a machine learning technique, i.e., random forests, combined with a Bayesian inference formulated for regionalized principal components of a set of flow indices in 92 catchments in northern Spain. Heuvelmans et al., (2006) demonstrated that artificial neural networks could provide a useful alternative in some cases compared with linear regression, especially when we could physically explain the nonlinear relationship between parameters and catchment descriptors.

Here we attempted to estimate parameters as a function of watershed features in ungauged areas using a machine learning method, i.e., the gradient boosting machine (GBM) (Friedman, 2001). GBMs are a family of powerful machine-learning techniques that have achieved considerable success across many

domains, such as image classification (Lawrence, 2004), text classification (Natekin and Knoll, 2013), pattern recognition (Schütz et al., 2019), motion detection (Bouwman et al., 2020), and ecological and environmental issues (Bouwman et al., 2020). The principal idea behind GBM is to consecutively construct the new base-learners, which optimally improves prediction in combination with the already existing ensemble leading to a more accurate estimate of the response variable (Natekin and Knoll, 2013; Schütz et al., 2019). We used the MLR to evaluate and compare the effectiveness of GBM for parameter regionalization.

We incorporated the Penman-Monteith-Leuning (PML) equation (Leuning et al., 2008) into the Distributed Time-Variant Gain Model (DTVGM) (Wang et al., 2009; Xia et al., 2005) forced with state-of-the-art meteorological data to predict runoff and ET over China. Our specific objectives were to (i) explore the capability of GBM for parameter regionalization relative to the traditional MLR method; (ii) develop hydrological parameters that are linked to soil and terrain properties; and (iii) identify the critical factors for these hydrological parameters in distinct climatic zones.

2 Materials and Methods

We ran a hydrological model, i.e., Distributed Time-Variant Gain Model with the Penman-Monteith-Leuning equation (DTVGM-PML) developed in this study, for country of China in a spatially distributed manner to account for within-catchment heterogeneity, the discrepancy in scale, and thus rainfall-runoff behavior between catchments and grid cells (Beck et al., 2020). Figure 1 presents an overview of this study. We first conducted a multivariable calibration to derive an ensemble of grid-by-grid parameter maps for regionalization. The multivariable datasets for model calibration and validation included a runoff product (Zhang et al., 2014) and an evapotranspiration (ET) dataset (Martens et al., 2017). The runoff product ($0.25° \times 0.25°$) was derived from the Variable Infiltration Capacity (VIC) model where the simulated monthly streamflow matches well with the measurements at the major river basins in China (Zhang et al., 2014). We obtained the total runoff depths from the dataset for model calibration and validation. The ET product used in this study was the 0.25°-resolution Global Land Evaporation Amsterdam Model (GLEAM) V3.3a dataset (Martens et al., 2017), which has been demonstrated to estimate actual ET with reasonable accuracy in China (Yang et al., 2017).

We regionalized model parameters using the MLR and GBM methods in terms of the explanatory
120 variables, including topographical and soil characteristics. Finally, we compared the model performance
with parameters from grid-by-grid calibration and regionalization.
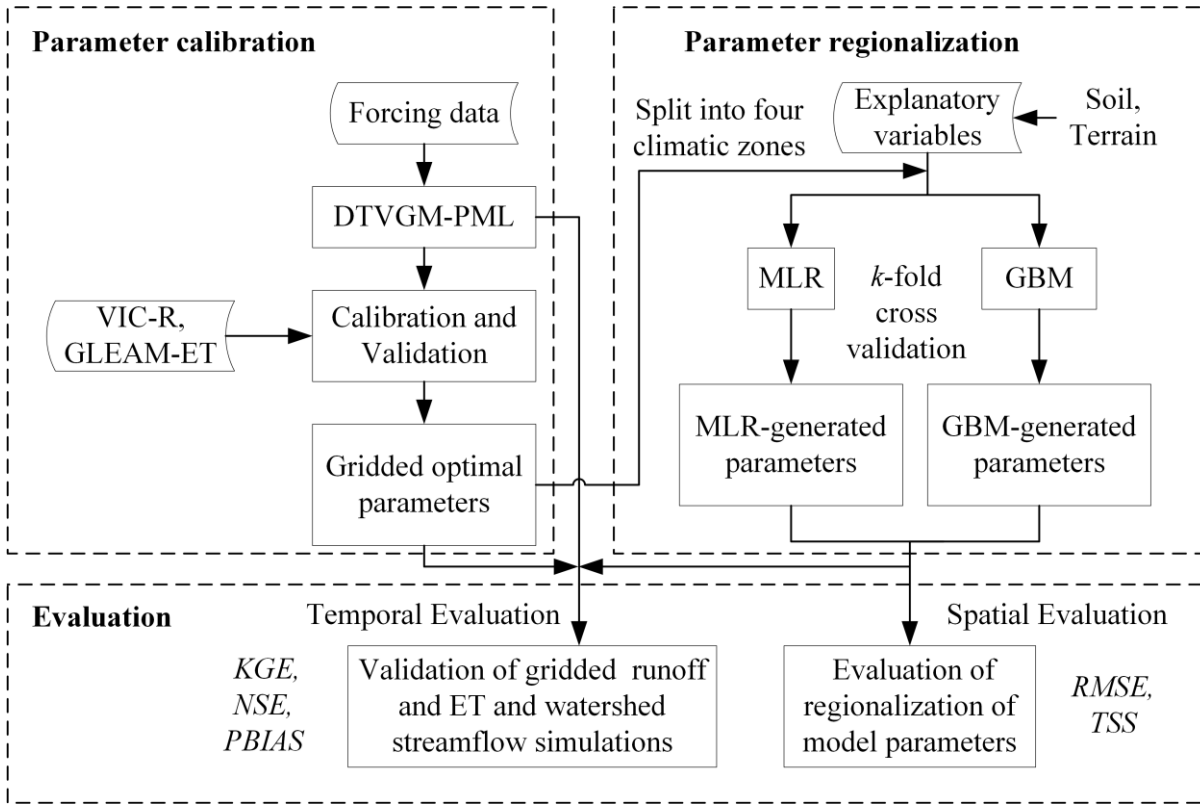
**Figure 1. Flowchart of parameter calibration, validation, and regionalization. DTVGM-PML: the distributed
time-variant gain model with the penman-monteith-leuning equation; VIC-R: runoff simulated by the Variable
125 Infiltration Capacity (VIC) model; GLEAM-ET: evapotranspiration dataset from the Global Land Evaporation
Amsterdam Model (GLEAM) V3.3a; MLR: stepwise multiple linear regression; GBM: gradient boost machine;
KGE: Kling-Gupta efficiency; NSE: Nash-Sutcliffe efficiency; PBIAS: percent bias; RMSE: root-mean-square
error; TSS: Taylor skill score.**

2.1 Application of the hydrological model across China

130 We developed the DTVGM-PML model (Figure 2) in this study to implement hydrological modelling
across China. The original DTVGM model has been successfully applied to many river basins in China
(Cai et al., 2014; Ning et al., 2016; Wang et al., 2009; Xia et al., 2005; Zeng et al., 2020; Zhan et al.,
2013; Zou et al., 2017). We selected the model for its parsimonious model structure with limited free
parameters due to the lack of ground-based measurements (e.g., observed streamflow) for model

calibration in many regions of China. A great deal of previous studies have highlighted the importance of incorporating the vegetation change information into hydrological models to achieve better performance in hydrological simulations (Donohue et al., 2007, 2010; Gerten, 2013; Ivanov et al., 2008; Lei et al., 2014; Thompson et al., 2011). Additionally, it has been demonstrated that coupling the PML equation into hydrological models can improve the hydrological simulations under vegetation greening conditions (Bai et al., 2018; Li et al., 2009; Zhang et al., 2009; Zhou et al., 2013). Thus, we replaced the empirical evaporation module in the DTVGM with the PML equation to improve the mechanistic representation of the ET process. We also coupled a snow routine from the HBV (Hydrologiska Byråns Vattenavdelning) model (Seibert and Vis, 2012) and the Gash rainfall interception model (van Dijk and Bruijnzeel, 2001) to improve relevant processes in DTVGM. The DTVGM-PML model ran on a grid-scale, with a spatial resolution of $0.25°×0.25°$. The gridded runoff simulated by the DTVGM-PML was routed by the Lohmann routing model for specific watersheds (Lohmann et al., 1996).

In the PML equation, we estimated the ratio of soil evaporation to the equilibrium rate (f) using the relative soil water storage, W/W_M, simulated in the DTVGM. Another key parameter, the maximum stomatal conductance (g_{sx}), was assigned for each land cover type recommended by Zhang et al., (2017). Other insensitive parameters were held constant since ET simulations have no significant accuracy loss (Bai et al., 2018; Leuning et al., 2008; Zhang et al., 2008).

In this study, we calibrated six parameters in the runoff generation process of the DTVGM-PML: two parameters (g_1, g_2) that control the nonlinear surface runoff generation, the subsurface runoff generation coefficient (k_s), the groundwater recharge coefficient (k_r) and recession coefficient (k_g), and the soil moisture storage capacity (W_M) (Figure 2). We provided detailed descriptions of the DTVGM-PML in the supplementary materials.

Figure 2. The model structure of the DTVGM-PML.

2.2 Parameter calibration strategy

160   In this study, we performed the 0.25° grid-by-grid calibration by fitting gridded monthly runoff and ET data at a national scale (a total of 15640 grid-cells) owing to the limited long-term observed streamflow data. The modeling period spanned from 1982–2012 and consisted of 15 years (1998–2012) of the calibration period and 16 years (1982–1997) of the validation period. We used the Shuffled Complex Evolution (SCE-UA) algorithm (Duan et al., 1992, 1994) for model calibration by minimizing a multi-

165   variable function (see Eq.(1)). The objective function is expressed as the Euclidean distance (denoted by F) that combines the Kling-Gupta efficiency (KGE, see Eq. (2)) (Gupta et al., 2009; Kling et al., 2012) of monthly runoff (KGE_R) and ET (KGE_{ET}). The KGE is a comprehensive criterion to measure the agreement between observed and simulated values ranging from $-\infty$ to 1, with an optimal value of 1.

$$F = \sqrt{w_1(1 - KGE_R)^2 + w_2(1 - KGE_{ET})^2},$$ (1)

170    $$KGE = 1 - \sqrt{(r - 1)^2 + (\beta - 1)^2 + (\gamma - 1)^2},$$ (2)

where w_1, w_2 are the weights assigned to runoff and ET evaluation, respectively. In this study, w_1, w_2 were both equal to 1. r is the Pearson correlation coefficient, β denotes the bias term (i.e., a ratio of means), and γ is the variability term (i.e., a ratio of coefficients of variation).

2.3 Parameter regionalization strategy

Regionalization techniques generally include two types, such as distance-based (spatial proximity, physical similarity) and regression-based methods (He et al., 2011). In this study, we obtained the 0.25° gridded parameters of the DTVGM-PML based on a multi-variable calibration, which were then divided into four climatic zones over China (see Figure 3, such as humid, semi-humid, semi-arid, and arid region). For each grid-cell, we estimated the relationship between the calibrated parameters (respond variable) and physical properties (explanatory variables, e.g., topographical attributes and soil characteristics, see Table 1) by a machine learning technique, the gradient boosting machine (GBM) (Friedman, 2001). For comparison of the performance of GBM, we also examined the traditional regression method, the multiple linear regression method (MLR) (Heuvelmans et al., 2006; Waseem et al., 2016), for parameter regionalization as the benchmark.

The GBM is a powerful machine learning technique to train decision trees in a gradual, additive, and sequential manner (Friedman et al., 2000; Friedman, 2001; Natekin and Knoll, 2013). The main idea of the GBM is to add new models with respect to the error of the whole ensemble learned so far to the ensemble sequentially to boosts its performance iteratively. The final GBM model is a stagewise additive model of previous individual trees. The GBM has been proven successful across many domains, including classification problems (Lawrence, 2004; Xia et al., 2020) and regression problems (Liao et al., 2020; Xenochristou et al., 2020; Yan et al., 2019) which is the case for this study.

The MLR approach is a standard multiple linear regression to relate response variables to the explanatory variables in a simple, fast and straightforward manner (Lima et al., 2015; Zhang et al., 2018). And the stepwise selection of predictors is applied to minimize the possible errors resulting in the best performing model and then identify the most influenced physiographical variable (Lima et al., 2015; Shu and Ouarda, 2012; Waseem et al., 2016). Unlike the GBM method, the MLR can explicitly quantify the relationship between explanatory variables and response variables through a regression equation.

The GBM and MLR modeling were conducted using the 'gbm' and 'lmStepAIC' methods, respectively, with the k-fold cross validation in the R package 'caret' (Kuhn et al., 2020). The performing of the k-fold cross validation (k=10 in this study) can help to reduce the chances of overfitting, leading to less prediction variability and, therefore, improved accuracy (Natekin and Knoll, 2013). First, we considered two categories of representative explanatory variables for regression modeling in each grid-cell: (i) topographical variables, e.g., elevation (m), and slope (°); (ii) soil variables, e.g., sand content (g/kg), silt content (g/kg), clay content (g/kg), field capacity, wilting point, residual moisture content, saturated moisture content; saturated hydraulic conductivity (cm d^{-1}). Second, we eliminated the grid-cells with either KGE_R or KGE_{ET} less than zero, which perform poorly in simulating runoff or ET, and then split the remaining grid-cells into four subsets according to the climatic zones. Finally, we trained and evaluated the GBM and MLR modeling using the model parameters in each subset and the relevant explanatory variables.

2.4 Evaluation criteria

We used the Root-mean-square error (RMSE) to evaluate the performance of the parameter prediction based on the GBM and MLR modeling. A lower RMSE indicates better performance than a higher one. We also calculated the Taylor skill score (TSS) (Taylor, 2001) to express a synthetic measure of the prediction skill of the MLR and GBM modeling for model parameters. The TSS is a numerical summary of the Taylor diagram, varying from zero (least skillful) to one (most skillful). The TSS, as a comprehensive metric of correlation coefficient, standard deviation, and root mean square error, has been widely used in model evaluation (Mohan and Bhaskaran, 2019; Taylor, 2001). As defined in Eq., the TSS increases monotonically with increasing correlation ($r{\rightarrow}r_0$) for any given variance, and increases as the modeled variance approaches the observed variance (SDR$\rightarrow$1) for any given correlation. The Kling-Gupta efficiency (KGE) (Gupta et al., 2009; Kling et al., 2012), Percent Bias (PBIAS) (Gupta et al., 1999), and Nash-Sutcliffe efficiency (NSE) (Nash and Sutcliffe, 1970) were used for the evaluation of model simulations based on three parameter sets. The PBIAS varying from $-\infty$ to $+\infty$ measures the extent to which the simulated values are overestimated (a positive value) or underestimated (a negative value) relative to the observed values. The NSE is a widely used evaluation index to assess the predictive skill

of hydrological models, which ranges from $-\infty$ to its perfect score, that is, 1. In addition to the KGE shown in Eq. (2), other evaluation criteria are formulated as follows:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i - Y_i)^2}, \tag{3}$$

$$TSS = \frac{4(1+r)^4}{\left(SDR+\frac{1}{SDR}\right)^2 (1+r_0)^4}, \tag{4}$$

$$r = \frac{\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sigma_X \sigma_Y}, \tag{5}$$

$$SDR = \frac{\sigma_X}{\sigma_Y}, \tag{6}$$

$$\sigma_X = \sqrt{\frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n}}, \sigma_Y = \sqrt{\frac{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}{n}}, \tag{7}$$

where X, Y are the calibrated parameter and regionalized parameter in DTVGM-PML (e.g., the runoff generation parameters, g_1 and g_2), respectively; the subscript i denotes the ith sample of the gridded parameter; $\bar{X}$, $\bar{Y}$ are the mean values of X and Y, respectively; σ_X, σ_Y are the spatial standard deviation of calibrated parameter and regionalized parameter, respectively; r represents the spatial correlation coefficient between X and Y; r_0 is the maximum correlation attainable and usually set to 0.999; SDR is the ratio of σ_X to σ_Y; and n is the total number of values for X (and Y).

$$PBIAS = \frac{\sum_{i=1}^{n}(Y_i - X_i)}{\sum_{i=1}^{n}X_i} \times 100\%, \tag{8}$$

$$NSE = 1 - \frac{\sum_{i=1}^{n}(X_i - Y_i)^2}{\sum_{i=1}^{n}(X_i - \bar{X})^2}, \tag{9}$$

where X, Y are the observed and simulated values (e.g., observed streamflow and DTVGM-PML simulated streamflow), respectively; the subscript i denotes the ith time step (day or month) of hydrological variables (i.e., runoff, ET, and streamflow); $\bar{X}$ is the mean value of X; and n is the total number of values for X (and Y).

2.5 Data sources and processing

245 Data used in this study consisted of three categories, such as forcing data, attribute data, and evaluation data (see Table 1).

(1) Model forcing data, including the climate forcing data and land surface data, were used for DTVGM-PML simulation. We used the China meteorological forcing dataset (CMFD) provided by National Tibetan Plateau Data Center (He et al., 2020; Yang et al., 2010; Yang and He, 2019), which contains

250 seven daily variables (see Table 1) with a spatial resolution of 0.1°. The land surface variables used here included leaf area index (LAI) and albedo obtained from the 8-day composite $0.05° \times 0.05°$ GLASS product in National Earth System Science Data Center. The 8-day composite data were interpolated into the daily data using a piecewise cubic Hermite polynomial and then smoothed by the Savitzky-Golay filtering method (Fang et al., 2008; Li et al., 2009; Ruffin et al., 2008).

255 (2) Ten attribute variables were considered, including topographical attributes and soil characteristics. The elevation and slope were derived from 90m digital elevation model. The soil data provided by Liu et al., (2020) included eight variables summarized in Table 1 at multiple depths 0–5, 5–15, 15–30, 30–60, 60–100, and 100–200 cm. We transformed multiple-layer soil data into single-layer data using a weighted average method.

260 (3) In addition to the total runoff data from the VIC simulation for China and ET data from the GLEAM v3.3a product used for parameter calibration (section 2.2), the evaluation data also included observed daily streamflow from 31 representative watersheds (see Figure 3) for streamflow validation over China. Table S1 lists the basic information of the 31 hydrological stations.

Since the DTVGM-PML ran at a daily time step for 1982–2012 with a spatial resolution of 0.25°, all
265 spatial data used in this study were resampled to 0.25° consistently.

Table 1. Model Input for Model Simulation, Training, and Evaluation

Data type	Name	Unit	Sources
Forcing data	2 m air temperature	K	http://data.tpdc.ac.cn
	Surface air pressure	Pa	Idem
	Specific humidity	kg kg^{-1}	Idem
	10 m wind speed	m s^{-1}	Idem
	Downward shortwave radiation	W m^{-2}	Idem
	Downward longwave radiation	W m^{-2}	Idem
	Precipitation rate	mm hr^{-1}	Idem
	LAI	m^2 m^{-2}	http://www.geodata.cn
	Albedo	Unitless	Idem
Attribute data	Elevation (elev)	m	http://srtm.csi.cgiar.org/
	Slope (slp)	°	Estimated using elevation
	Sand content (snd)	g kg^{-1}	(Liu et al., 2020)
	Silt content (slt)	g kg^{-1}	Idem
	Clay content (cly)	g kg^{-1}	Idem
	Field capacity (fc)	cm^3 cm^{-3}	Idem
	Wilting point (pw)	cm^3 cm^{-3}	Idem
	Residual moisture content (thr)	cm^3 cm^{-3}	Idem
	Saturated moisture content (ths)	cm^3 cm^{-3}	Idem
	Saturated hydraulic conductivity (ksat)	cm d^{-1}	Idem
Evaluation data	Runoff	kg m^{-2}	VIC simulation for China (Zhang et al., 2014)
	ET	mm month^{-1}	GLEAM v3.3a product (https://www.gleam.eu/)
	Streamflow	m^3 s^{-1}	China's Hydrological Year Book

Note. The attribute variable abbreviation is shown in the parenthesis after the variable name.
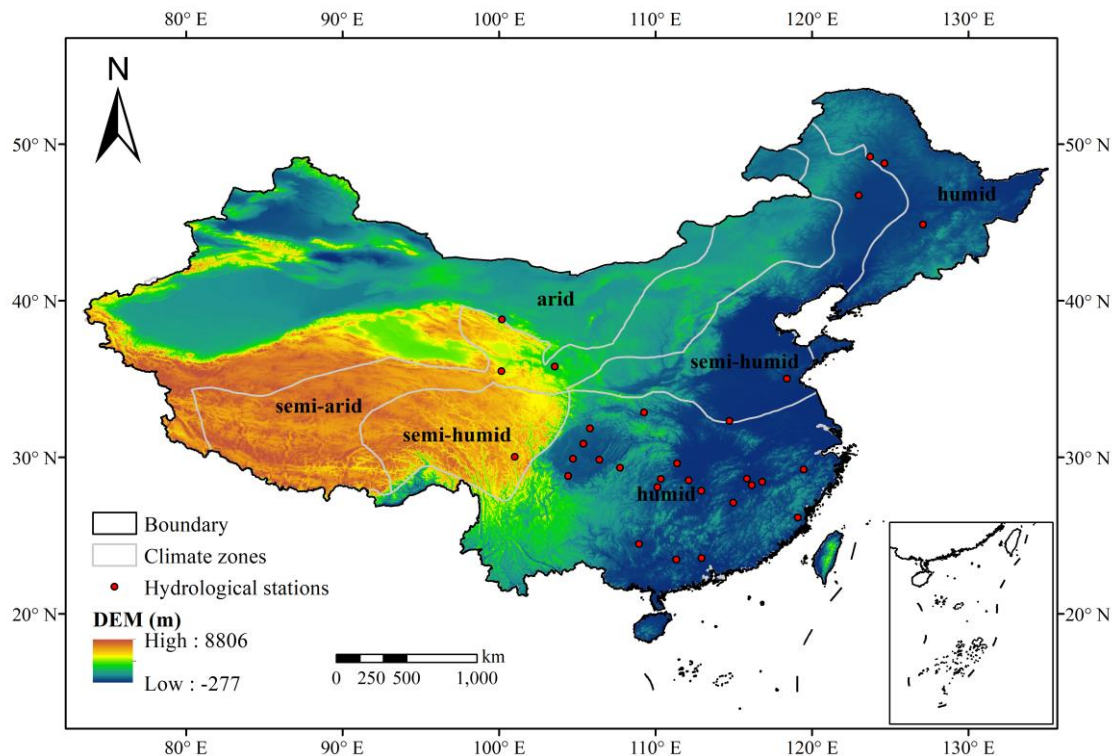
Figure 3. Location of climatic zones (humid, semi-humid, semi-arid, and arid) and hydrological stations in China.

3 Results

3.1 Simulation of DTVGM-PML

The model calibration set out using reference data (gridded monthly runoff and ET) from the former 16 years (1982–1997) or the latter 15 years (1998–2012) with the remaining period as the validation period. The model shows slightly better results when using data from the latter period than the first period for calibration (Fig. S1), which was probably due to the better data quality in the latter period than in the former period. Several previous studies have adopted this strategy, i.e., using data from the latter period for model calibration (Mizukami et al., 2017; Newman et al., 2017; Yang et al., 2018). Thus, the present study used the calibrated model parameters from the latter period for regionalization.

We also performed the comparison of model performance in hydrological simulation between the original DTVGM (without PML) and DTVGM-PML. As shown in Fig. S2, the KGE and PBIAS values of runoff

simulation (Fig. S2a, b) by DTVGM were lower than those from DTVGM-PML. The median KGE and PBIAS values of ET simulation (Fig. S2c, d) were comparable between the two models. In summary, DTVGM-PML can help to improve hydrological simulations relative to DTVGM. Additionally, the consideration of vegetation dynamics by the PML equation in DTVGM-PML would improve the mechanistic understanding of hydrological response under vegetation greening, which is lacking in DTVGM.

Figure 4 presents the spatial patterns of mean annual runoff and ET simulations derived from the DTVGM-PML during 1982–2012. Both the mean annual runoff and ET simulations show a decreasing trend from southeast to northwest China, with the highest values in humid tropical and subtropical regions, intermediate values in temperate regions, and lowest values in cold and arid regions. Figure 5 shows the model performance of runoff and ET simulations in calibration and validation periods. The median KGE and PBIAS values for runoff simulation were 0.78 and 0.8%, respectively, in the calibration period, and 0.69 and −14.2%, respectively, in the validation period. The corresponding statistical values for ET simulation were 0.70, −8.2% and 0.68, −13.1%, respectively. Overall, the DTVGM-PML could well simulate the monthly runoff and ET over China.
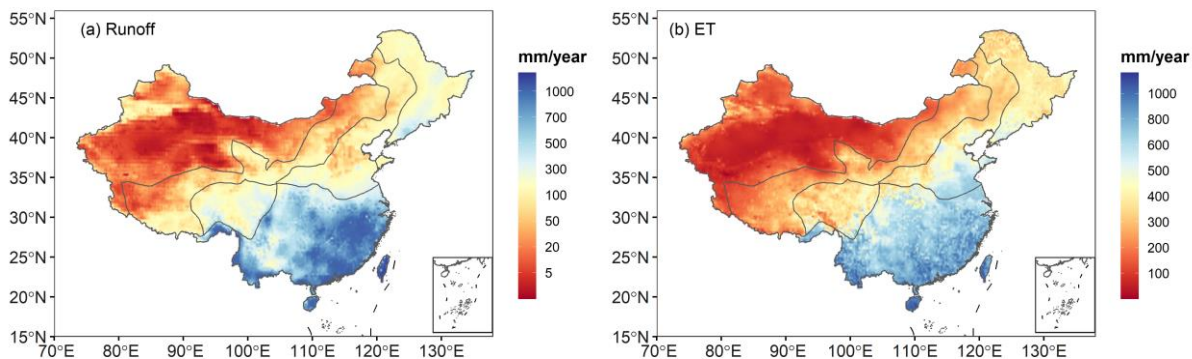
Figure 4. Spatial patterns of mean annual runoff (a) and evapotranspiration (ET) (b) simulations by the DTVGM-PML during 1982–2012.
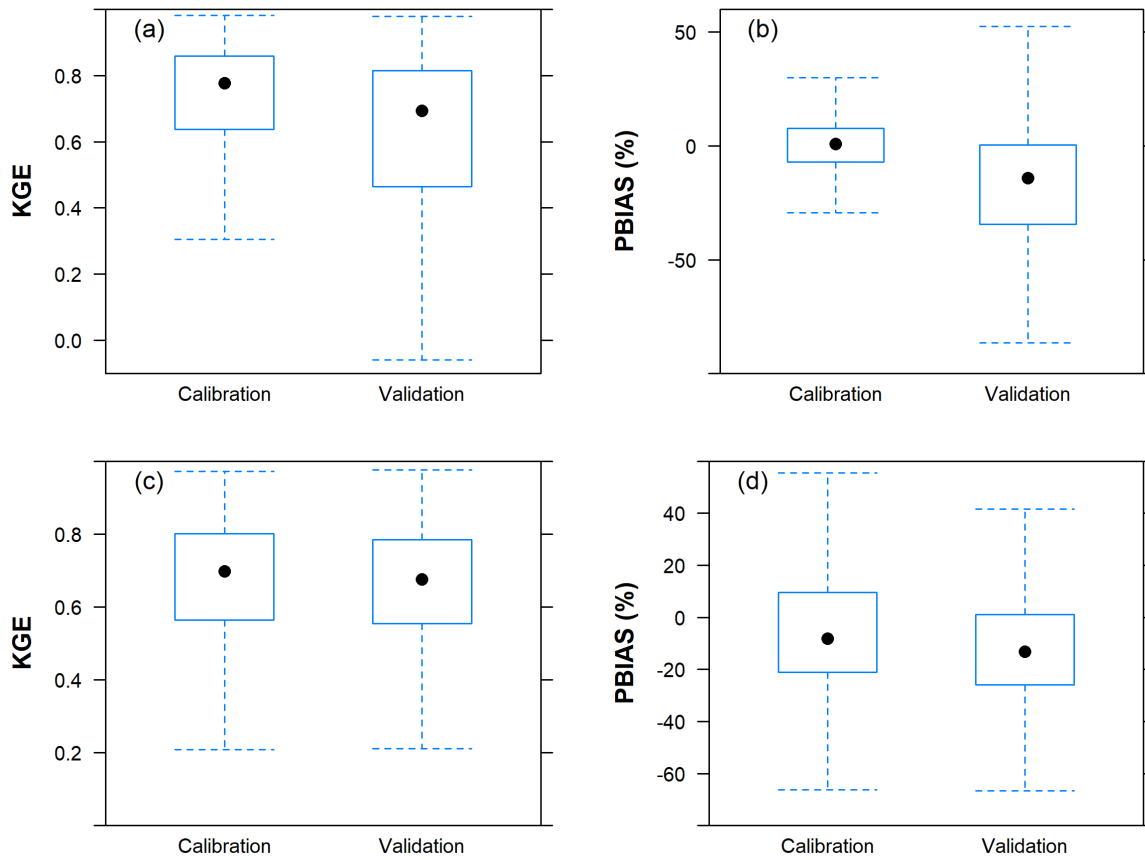
Figure 5. Model performance of runoff (KGE: a, and PBIAS: b) and evapotranspiration (ET) (KGE: c, and PBIAS: d) simulations in calibration and validation periods. The boxplot was generated using data from a total of 15640 grid cells over China. KGE denotes the Kling-Gupta efficiency. PBIAS denotes the percent bias.

3.2 Regionalization of model parameters

We evaluated the regionalization model performance in terms of RMSE for six parameters in four climatic zones. As shown in Figure 6, GBM appears better in predicting model parameters than MLR because of the lower RMSE for all parameters in humid regions. We found consistently better accuracy of GBM in semi-humid, semi-arid, and arid areas (Fig. S3 to Fig. S5). Additionally, the difference in model performance between GBM and MLR was significant (p-value < 0.05) as per the Kruskal-Wallis test (Hollander et al., 2013). Overall, these results suggest that the performance of GBM is significantly better than that of MLR for six parameters in four climatic zones. We also calculated the Taylor skill scores

(TSS) with the grid-scale calibrated parameters as the reference parameters to evaluate the regionalization model performance for estimating each parameter at each grid. As shown in Fig. S6, GBM obviously outperformed MLR with a higher TSS, suggesting that the GBM-regionalized parameters presented a
315   higher spatial agreement with reference parameter values than the MLR-generated parameters.
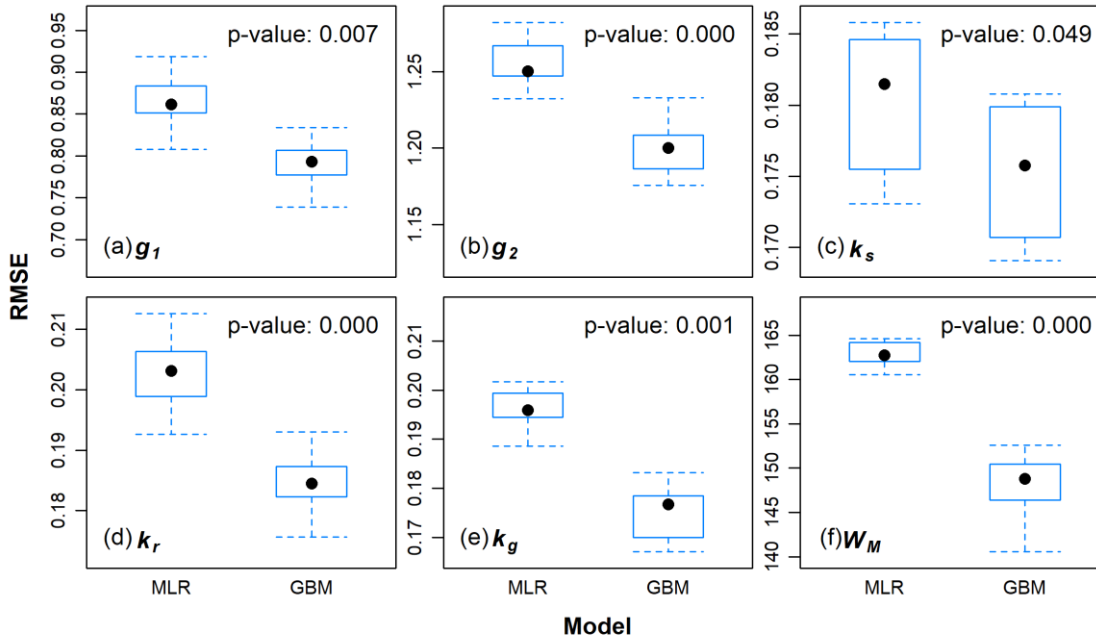
Figure 6. Performance evaluation of multiple linear regression (MLR) and gradient boosting machine (GBM) for six parameters, (a) g_1, (b) g_2, (c) k_s, (d) k_r, (e) k_g, (f) W_M, in humid region. MLR and GBM denote the multiple linear regression with stepwise selection and the gradient boosting machine model. The boxplot is generated from
320   **the 10 samples in k-fold cross validation. We use the non-parametric Kruskal-Wallis (KW) test to determine the significance of difference in the performance between MLR and GBM at a significance level of 0.05.**

Figure 7 shows the spatial patterns of the three parameter sets derived from calibration and regionalization (MLR and GBM). Generally, both MLR and GBM derived parameters exhibited good agreement spatially
325   with the calibrated parameters. As the model parameters were related to topography and soil properties, the parameters generated by MLR and GBM show exquisite spatial patterns and a much better spatial coherence than the calibrated parameters. Compared with the MLR generated parameters, the GBM generated parameters presented more consistency with the calibrated parameters in space. For example (Figure 7 e1, e2, and e3), the MLR underestimated the parameter k_g in part of Western China (nearly

330   0.25–0.5) relative to the calibrated parameters (about 0.5–0.8), while the GBM-derived parameters (0.5–0.75) were more consistent with calibrated values. In summary, the regionalized parameters generated by the regionalization methods (MLR and GBM) exhibited better spatial coherence relative to the calibrated parameters with spatial discontinuities. The GBM derived more accordant parameters with the calibration than the MLR.

335
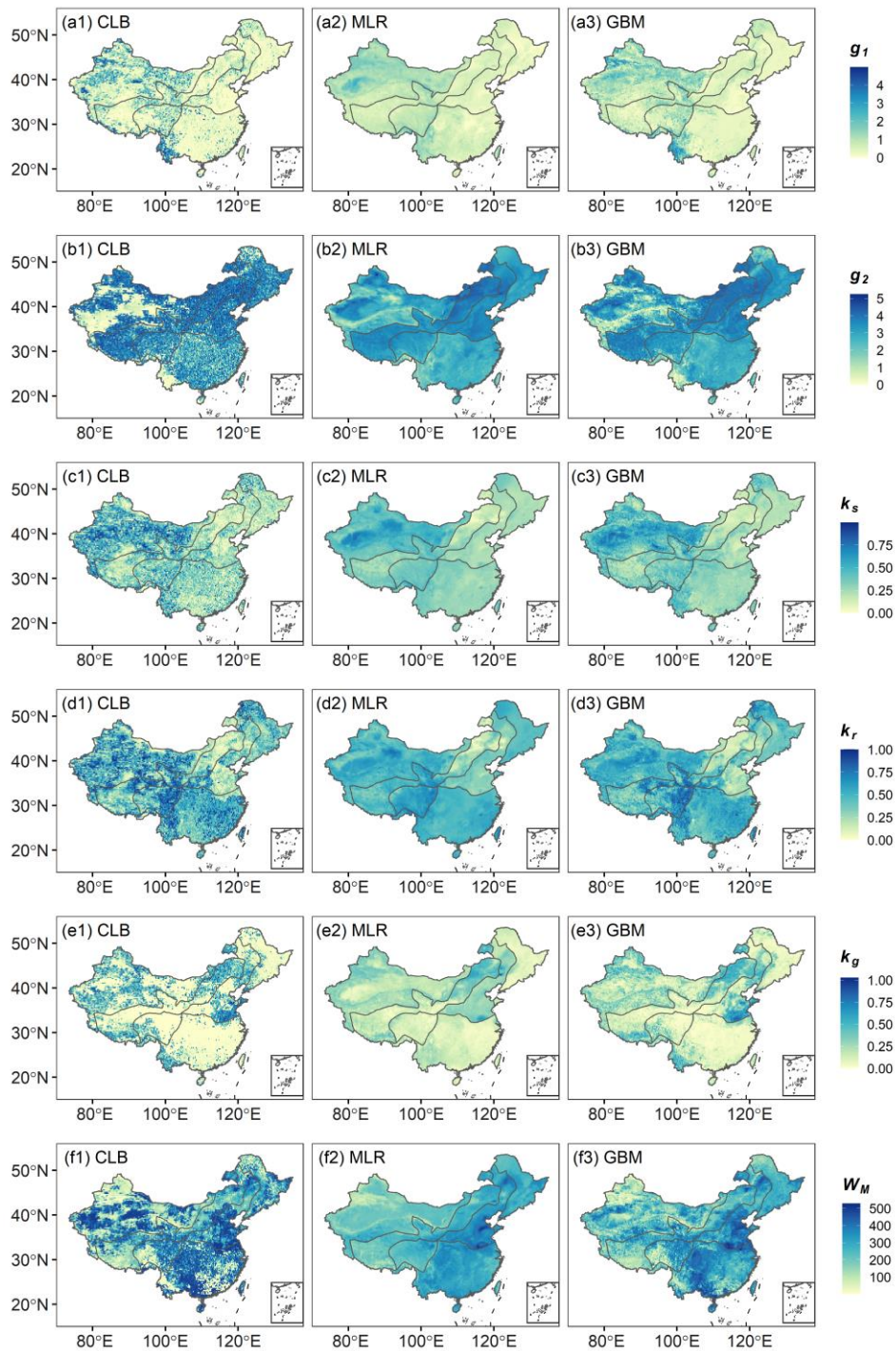
Figure 7. Spatial patterns of model parameters: (a) g_1, (b) g_2, (c) k_s, (d) k_r, (e) k_g, (f) W_M, derived from (1) calibration (CLB), (2) multiple linear regression (MLR) and (3) gradient boosting machine (GBM).

3.3 Validation of gridded runoff and ET simulations based on parameter regionalization

To assess the effectiveness of parameter regionalization, we compared the model performance of runoff and ET simulations with the regionalized parameter sets (MLR and GBM) to that with the calibrated parameters. Figure 8 and Figure 9 present cumulative density function (CDF) plots of KGE values for runoff and ET simulations in both the calibration (solid lines) and validation (dashed lines) periods over four climatic zones. The KGE values were computed based on the DTVGM-PML simulations using three parameter sets: (1) grid-by-grid calibration (black lines), (2) MLR generation (blue lines), and (3) GBM generation (red lines).

Regarding the runoff simulation (Figure 8), median KGE values produced by calibrated parameters were 0.783 in humid regions, 0.755 in semi-humid regions, 0.704 in semi-arid regions, and 0.442 in arid regions for the validation period. The median KGE score based on the MLR method was worse by 0.139 averaged in four climatic zones than that from simulation using calibrated parameters for the validation period. While as shown in Figure 8, all the KGE values in four regions from GBM parameters were superior to the MLR parameters. The corresponding difference of KGE was 0.023 relative to the calibration for the validation period.

In contrast to runoff simulation, however, the distributions of KGE for ET simulation from the regionalized parameters were significantly close to that based on the calibrated parameters in each region as shown in Fig. S7. The median KGE values from three parameter sets were around 0.68 in humid regions, 0.74 in semi-humid regions, 0.72 in semi-arid regions, and 0.53 in arid regions for the validation period. Overall, the performance of ET simulation from regionalization was comparable to that from calibration satisfactorily.
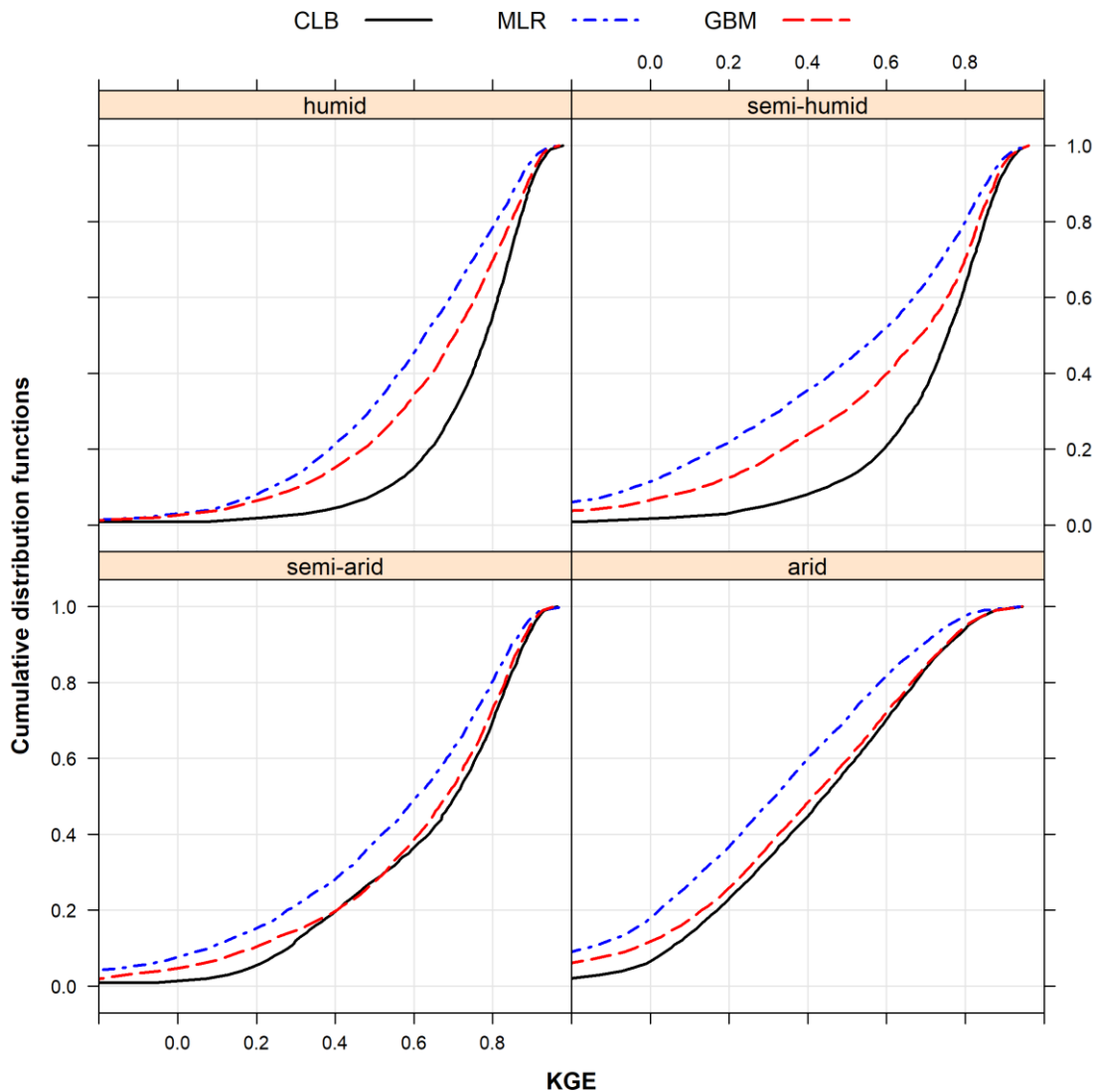
Figure 8. Cumulative distribution functions (CDFs) of *KGE* **for runoff simulation based on three parameter sets (black lines for calibration (CLB), blue lines for multiple linear regression (MLR), red lines for gradient boosting machine (GBM)) in the validation period over four climatic zones. KGE denotes the Kling-Gupta efficiency.**

3.4 Validation of watershed streamflow simulations based on parameter regionalization

To give insight into the performance of streamflow simulations based on parameter regionalization, we calculated NSE, KGE, and PBIAS at 31 representative watersheds with both daily and monthly

streamflow validation (Figure 9). According to the Kruskal-Wallis test, there were insignificant differences (p-values > 0.05) in five criteria based on the three parameter sets. Both MLR and GBM performed as well as the calibration with high median scores of NSE and KGE at both daily (NSE was nearly 0.64 and KGE was nearly 0.67) and monthly (NSE was nearly 0.84 and KGE was nearly 0.78) scales and median PBIAS close to zero (around −4.1%). Since the stations were almost located in humid and semi-humid regions (28 of 31 in Figure 3), we could expect that the parameters derived from the regionalization can reasonably generate monthly streamflow with good agreements to the observations (in line with results from Figure 8). We also obtained encouraging results of daily streamflow simulation with satisfying accuracy. In terms of the three stations in arid and semi-arid regions, the performance of MLR was slightly poorer than that of calibration and GBM, as the daily and monthly NSE values, for instance, at Tangnaihai station, were 0.494 and 0.586 compared with the corresponding values of 0.675 and 0.751 for calibration, 0.631 and 0.719 for GBM, respectively.
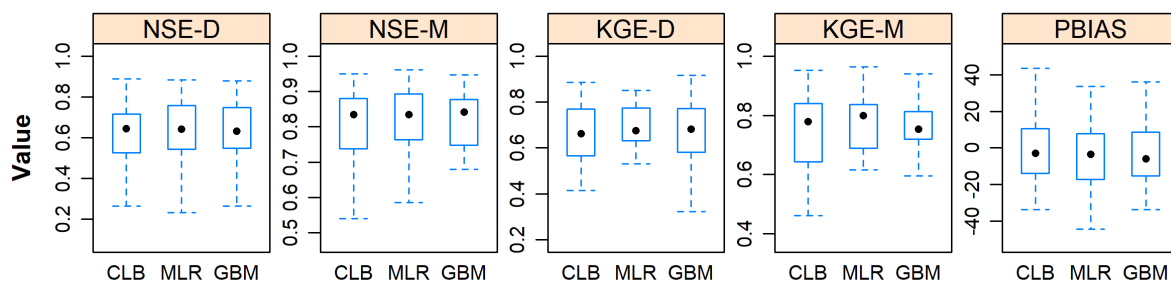
Figure 9. The model performance statistics of streamflow simulations on 31 hydrological stations based on three parameter sets. 'D' and 'M' denote daily and monthly evaluation, respectively.

3.5 Identification of important factors for model parameters

We further estimated the relative importance of each explanatory variable based on the GBM model, which was determined by averaging the improvement (decrease) in the squared error at each split over all trees made by each variable, with a range from 0 (least important) and 100 (most important) (Natekin and Knoll, 2013; Xia et al., 2020). Figure 10 presents the relative importance of explanatory variables from the GBM model for all six parameters in four climatic zones and the margin plots containing the mean relative importance for each parameter or each climatic zone.

390   Generally, we found the slope (slp), saturated moisture content (ths), and elevation (elev) were the most critical explanatory variables to inform the model parameters (Figure 10g5). Whereas there were several differences among parameters or regions. In humid and semi-humid regions, the terrain properties, including slope and elevation, were likely to determine most model parameters (Figure 10g1 and g2). As for semi-arid and arid regions, most parameters primarily depended on the saturated moisture content that

395   becomes a constraining factor for runoff generation in dry areas. For the three parameters (g_1, g_2, and k_s) that control surface and subsurface runoff generation, the dominant factors were slope and saturated moisture content (Figure 10a5, b5, and c5). However, the parameters for groundwater recharge (k_r) and recession (k_g) were mainly controlled by the saturated moisture content, followed by elevation (Figure 10d5 and e5). In terms of the parameter W_M, which represents the soil moisture storage capacity, the slope,

400   elevation followed by the saturated hydraulic conductivity were more essential predictors than the saturated moisture content (Figure 10f5).
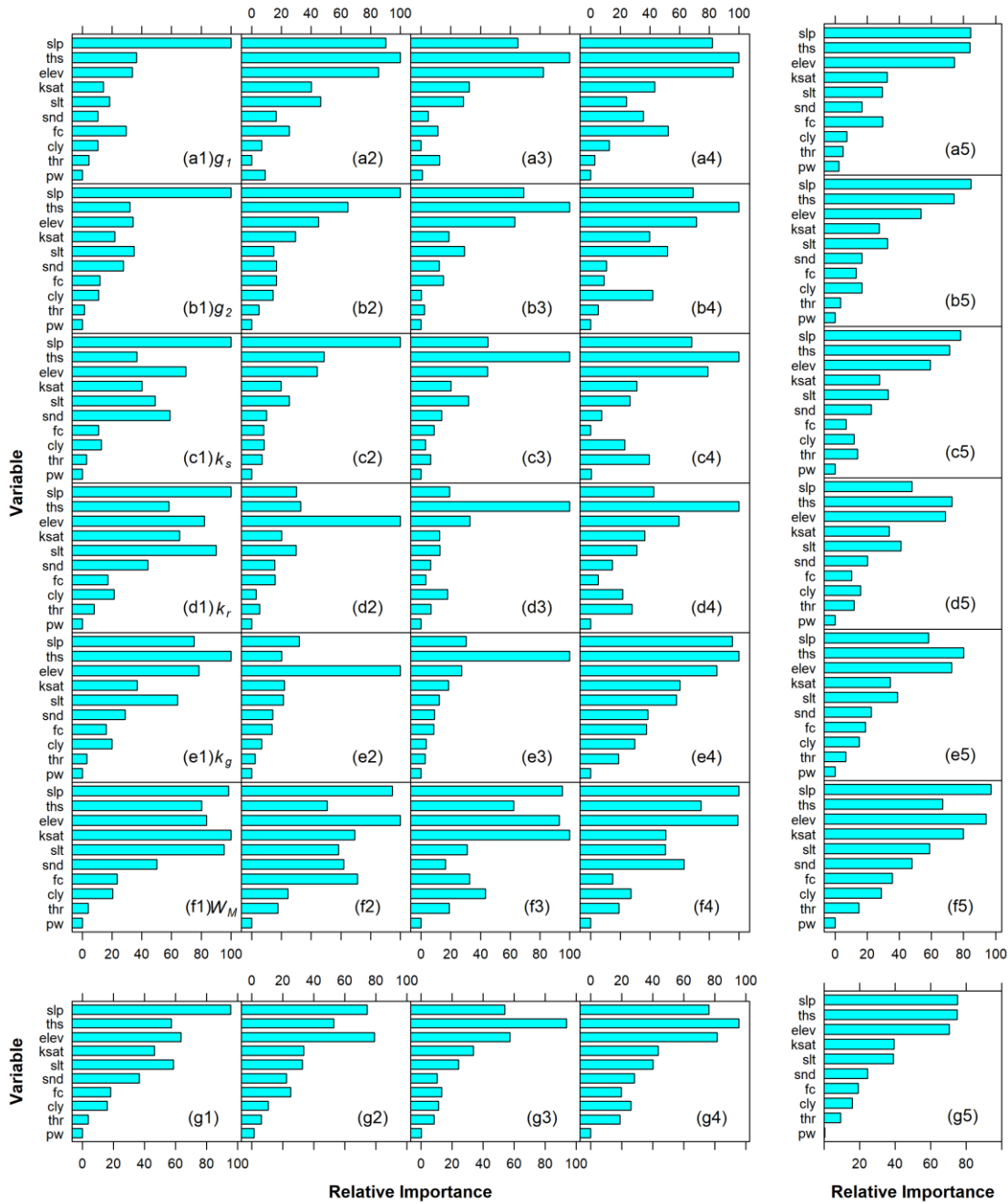
Figure 10. Relative importance of explanatory variables from the gradient boosting machine (GBM) model for (a) g_1, (b) g_2, (c) k_s, (d) k_r, (e) k_g, (f) W_M in four climatic zones (1, humid; 2, semi-humid; 3, semi-arid; 4, arid). The mean relative importance over four regions for each parameter is shown in the right column (a5-g5). The mean relative importance over six parameters for each climatic zone is shown in the bottom row (g1-g5). Explanatory variables include elev (Elevation), slp (Slope), snd (Sand content), slt (Silt content), cly (Clay content), fc (Field

405

capacity), pw (Wilting point), thr (Residual moisture content), ths (Saturated moisture content), ksat (Saturated hydraulic conductivity).

410 **4 Discussion**

4.1 Performance of calibration and regionalization for DTVGM-PML

Previous studies have established the effectiveness of the multivariate calibration framework for hydrological models (Bai et al., 2018; Dembélé et al., 2020; Demirel et al., 2018; Finger et al., 2015; Nijzink et al., 2018; Xie et al., 2020). The current study performed a multiple variable calibration strategy

415 to calibrate model parameters in each grid against the reference datasets (VIC runoff and GLEAM ET) during the 15-year calibration period, followed by an independent model validation against the reference gridded runoff and ET during the 16-year validation period. We also validated the model using observed streamflow at 31 representative hydrological stations in diverse climatic zones. The gridded runoff in each watershed was routed to the corresponding hydrological station using consistent routing parameters.

420 Despite that the streamflow data are commonly used in the calibration of hydrological models (Dembélé et al., (2020), previous studies have explored the potential in model calibration solely against remotely sensed ET data (without the need for gauged streamflow data) and achieved encouraging results in streamflow simulation (Huang et al., 2020; Zhang et al., 2020). The satisfactory performance in model validation of streamflow suggests the high reliability of the multivariate calibration strategy used in this

425 study.

Hydrological simulation in arid and semi-arid regions is still challenging (Huang et al., 2016; Wheater et al., 2007; Yang et al., 2019). There appears better performance in runoff simulations in (semi-) humid regions than (semi-) arid regions (Figure 8). Regionalization methods tend to perform more poorly in drier regions, which is expected in common modeling practices (Guo et al., 2021; Parajka et al., 2013).

430 The results support the general knowledge of runoff prediction in different climatic zones, i.e., it is more challenging to achieve good performance in (semi-) arid regions than humid regions. The relatively poor performance in (semi-) arid regions may be attributed to model structural deficiencies, forcing errors, as well as uncertainties in the reference data for calibration. For example, the large-scale hydrological models may ignore complex processes, like surface-groundwater interactions and channel losses

(Oubeidillah et al., 2014). The quality of forcing data (e.g., precipitation) also influences the performance of hydrological models in runoff simulation (Mizukami et al., 2017). Wang et al., (2016) found systematic overestimates of CMFD precipitation over the Qinghai-Tibetan Plateau. Furthermore, unlike the observed streamflow, the reference data, such as runoff and ET data, are not the standard actual observed data. This might also explain the better performance in our model validation against observed streamflow at watershed scale with regionalized parameters than calibrated parameters. Zhang et al., (2014) suggested that the data should be used with caution in western China with significant potential uncertainties in hydrologic simulation due to the lack of meteorological observations. And issues in arid regions have always been a challenge for the VIC model (Oubeidillah et al., 2014; Yang et al., 2019). Yang et al., (2017) indicated that the GLEAM ET data showed significantly systematic bias and overestimated the eddy covariance ET measurements at forest sites. We argue that these reanalysis datasets are precious for large-scale calibration of hydrological models in terms of both spatial and temporal dynamics, though it will inevitably introduce uncertainties to a certain degree.

4.2 Prediction of model parameters by machine learning methods

Given the complex and nonlinear relationships with high nonstationarity between physical catchment descriptors and model parameters, it is a daunting challenge to develop a robust method for the regionalization of hydrological models (Guo et al., 2021). Machine learning approaches provide a promising tool, as an alternative to conventional linear regression models, to regionalize hydrological model parameters. Machine learning techniques are focused on specific tasks, like classification, regression, and have been widely used in many hydrological issues (Adnan et al., 2019; Huntingford et al., 2019; Lima et al., 2015; Rajaee et al., 2019; Shen, 2018; Yaseen et al., 2015; Zhang et al., 2018). In this study, we built a predictive GBM model for parameter regionalization of DTVGM-PML compared with an MLR model using the terrain and soil properties as explanatory variables. Overall, the GBM outperformed the MLR based on evaluation in three aspects: (i) higher accuracy in reproducing the calibrated parameters as indicated by significantly lower RMSE with GBM against MLR (Figure 6) and more consistent spatial pattern with calibrated parameter values (Figure 7); (ii) better performance in runoff simulations based on parameters generated from GBM than MLR (Figure 8); (iii) comparable

results of streamflow validation in 31 watersheds based on regionalization and better performance in several stations in (semi-) arid region by GBM than MLR (Figure 9). Note that the parameters in ET estimation (PML method in DTVGM-PML) were not involved in model calibration. Consequently, the performance of ET simulation from regionalization was comparable to that from calibration (Fig. S7). Taken together, the GBM method, as an ensemble technique, can achieve higher accuracy in parameter regionalization than MLR (Natekin and Knoll, 2013).

We noticed that the GBM outperformed the MLR in grid-scale runoff simulations but showed insignificant difference in watershed streamflow validation. It is likely due to that (i) the observed watershed streamflow data are independent of the girded runoff data in this study, which could lead to differential modelling performance in these two datasets; (ii) most of the watersheds (i.e., 28 out of 31) for streamflow validation are within humid regions, where the difference in performance of grid-scale runoff simulations was relatively small, compared with non-humid regions, between regionalized and calibrated parameters; and (iii) the flow routing of grid-scale runoff within a watershed may smooth the heterogeneity in runoff from multiple spatially distributed grid. We suggest that streamflow from diverse watersheds, especially in arid regions, are needed for model validation with parameters derived from different regionalization approaches. The possible inconsistent results between the evaluation of parameter regionalization and the validation of streamflow also imply the necessity of watershed-scale streamflow validation following parameter regionalization.

The proposed GBM method explored the relationship between model parameters and the terrain and soil attributes, offering a helpful approach to estimate model parameters for hydrological simulations. It can also achieve satisfactory accuracy, especially in (semi-) humid regions. This study also motivates further investigations including, but not limited to (i) improvement in model structure to better represent hydrological mechanisms in complicated underlying surface conditions and changing environment, (ii) selection of more physical attributes such as vegetation or climatic factors for regionalization.

Although the GBM cannot provide an explicit formula that intuitively links the response variable with explanatory variables like the MLR, it can also estimate response variables based on explanatory variables. More importantly, the machine learning-based regionalization methods can identify essential

driving factors, which develop a primary appraisal of how important terrain and soil properties are for
490   parameters of hydrological models.

4.3 Important attributes dominating model parameters

The runoff process is primarily controlled by regional climatic regime, vegetation, land use, topography,
and soils (Dunne and Leopold, 1978; Freeze, 1974; Mizukami et al., 2017; Tarboton, 2003). We used the
GBM method to predict parameter values by available topographic and soil properties, such as slope,
495   elevation, saturated moisture content, saturated hydraulic conductivity, field capacity, soil texture. Our
findings of variable importance using the GBM model quantitatively indicate that the runoff generation
parameters of DTVGM-PML are majorly controlled by slope, saturated moisture content, and elevation.
Moreover, the results in different climatic zones show that terrain attributes significantly influence the
runoff process in relatively humid regions. At the same time, the saturated moisture content becomes a
500   limiting factor in drier areas.

Prior studies have noted the incredible impact of slope on runoff generation (Akbarimehr and Naghdi,
2012; Chaplot and Le Bissonnais, 2003; Garg et al., 2013; Tarboton, 2003). Steeper slopes lead to faster
drainage in aquifers (Beck et al., 2020; Post and Jakeman, 1996; Zecharias and Brutsaert, 1988). The
slope seems to be the most critical factor for parameters g_1, g_2, and k_s that control surface and subsurface
505   runoff generation. Garg et al., (2013) investigated that the slope affects surface runoff estimation
significantly for Solani watershed using the modified NRCS-CN (natural resources conservation service
curve number) method. Several reports have shown that storm runoff by subsurface flow requires steep,
convex slopes and high saturated hydraulic conductivities (Freeze, 1972, 1974; Montgomery and
Dietrich, 2002). The parameter g_1 presents the surface runoff coefficient when the soil water storage
510   reaches its maximum value, i.e., $W/W_M=1$, and the parameter g_2 is the power of the relative soil water
storage (W/W_M). Steeper slope results in an increase in the value of g_1, a decrease in g_2 (Fig. S8 a1 and
b1), and thus leading to an increase in the amount of surface runoff, which supports the previous findings
that runoff amount increases with increasing slope (Chaplot and Le Bissonnais, 2003; Huang, 1995).

Regarding the groundwater recharge and baseflow, the corresponding parameters (k_r and k_s) strongly
515   depend on the saturated moisture content. The results are consistent with Chiew and Siriwardena, (2005)

who found that the groundwater recharge parameter and baseflow recession parameter of SIMHYD (a simplified version of the HYDROLOG) model are highly correlated with plant available water holding capacity, a proxy of the soil water storage capacity (Mckenzie et al., 2000). While concerning the soil moisture storage capacity, W_M, in DTVGM-PML, the saturated moisture content is likely to be less important than the slope, elevation, and saturated hydrologic conductivity. Note that W_M is different from the saturated moisture content in soil. The latter is equivalent to effective porosity and is simplified as the one-layer value in this study. Whereas the soil moisture state variables in many conceptual hydrological models do not act in the same way as in the real world (Zhuo and Han, 2016). The surface slope is correlated with soil depth (Tesfa et al., 2009), which strongly influences W_M. If the soil depth data, more likely to be associated with W_M, are available, they should be incorporated in parameter regionalization to obtain more reliable and intelligible results.

4.4 Necessity of parameter regionalization

Hydrologic models often rely on regionalization approaches to transfer information from small to large spatial scale (e.g., from gridcell to subbasin, watershed, and regional scale) (Beck et al., 2020; Mizukami et al., 2017), and from gauged to ungauged catchments (He et al., 2011; Hrachowitz et al., 2013; Pagliero et al., 2019; Parajka et al., 2013). In this study, though parameters were calibrated and thus available at each gridcell, the parameter values at around 450 gridcells were not reliable owning to poor model performance (i.e., KEG < 0) (Knoben et al., 2018; Koskinen et al., 2017; Sutanudjaja et al., 2018). Therefore, we only used the calibrated parameters with KGE ≥ 0 (i.e., representing better model performance) for regionalization of parameters. The model performance for 53% of these gridcells (with KGE < 0 prior to regionalization) were improved when we re-ran the model with regionalized parameters. Particularly, the KGE values in 37% of the gridcells (with KGE < 0 prior to regionalization) became positive, indicating a substantive improvement of the modeling performance.

Even though the parameters were well calibrated and available at each gridcell, one might think whether and which topographic and edaphic properties mediate these hydrological parameters. Our machine learning (i.e., GBM) based regionalization of parameters enables to estimate six key hydrological parameters using site-specific characteristics. Following the regionalization of parameters, our results of

variable importance quantitatively indicate that the runoff generation parameters are majorly controlled by slope, saturated soil moisture content, and elevation. Moreover, the terrain attributes significantly regulate the runoff processes in relatively humid regions, while the saturated soil moisture content becomes a limiting factor in arid areas. The regionalization of parameters will improve our mechanistic understanding of the runoff generation processes and associated key hydrological parameters under different topographic and edaphic conditions.

5 Conclusion

We conducted parameter regionalization for the DTVGM-PML model using a machine learning technique, the gradient boost machine, compared with the traditional multiple linear regression method. We show that the GBM model is superior to the MLR in predicting model parameters as a function of topographical and soil characteristics due to its significantly lower biases and higher spatial agreement for almost all parameters in four distinct climatic zones. The regionalized parameters also exhibited better spatial coherence relative to the grid-by-grid calibrated parameters. Regarding the model validation of streamflow simulations in 31 hydrological stations, MLR- and GBM-generated parameters could simulate streamflow as accurately as the results with grid-by-grid calibrated parameters (median daily and monthly NSE are 0.65 and 0.84, respectively), with the GBM being preferable to MLR in arid regions. This study suggests the watershed-scale streamflow validation following parameter regionalization is necessary due to potential inconsistent results between parameter regionalization evaluation and the streamflow validation. Based on the GBM regionalization results, we found that the slope, saturated moisture content, and elevation are the most important explanatory variables to inform model parameters. Our results indicate that machine learning techniques can be a useful alternative to the conventional regression approach to better predict hydrological model parameters. This is particularly significant for hydrological predictions in ungauged basins. The methods developed and insights gained from this study can also improve the interpretation and prediction of parameters in other large-scale hydrological and environmental models.

Code and data availability

The model code used in this study is available via https://github.com/Songzh101/GBM_HydroModel_Regionalization. The climatic forcing data, land surface data, elevation data, and evapotranspiration data are available online, as described in Table 1. The streamflow in the 31 watersheds were obtained from China's Hydrological Year Book. These data are not publicly available because of governmental restrictions, but can be accessed by contacting the corresponding author. The soil properties are provided by Liu et al., (2020). The VIC simulations for China are obtained from Zhang et al., (2014).

Author contribution

ZS and GW designed the study. DS, GW and ZS developed the model code and performed the simulations. ZS prepared the manuscript with contributions from all co-authors. JX, GW, DS, CH, SH contributed to results discussion and critical review of the manuscript.

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgments

This work was funded by the National Natural Science Foundation of China (No. 41890823), the National Key R&D Program of China (Grant Number: 2017YFA0603702), and the Excellent Young Scientists Fund. The climatic forcing data set are provided by National Tibetan Plateau Data Center (http://data.tpdc.ac.cn). The land surface data are provided by National Earth System Science Data Center, National Science & Technology Infrastructure of China. (http://www.geodata.cn). We also appreciate the data support from Zhang et al., (2014) and Liu et al., (2020).

References

Adnan, R.M., Liang, Z., Trajkovic, S., Zounemat-Kermani, M., Li, B., and Kisi, O.: Daily streamflow prediction using optimally pruned extreme learning machine, J. Hydrol., 577, 123981, https://doi.org/10.1016/j.jhydrol.2019.123981, 2019

595 Akbarimehr, M., and Naghdi, R.: Assessing the relationship of slope and runoff volume on skid trails (Case study: Nav 3 district), Journal of forest science (Praha), 58(No. 8), 357-362, 10.17221/26/2012-JFS, 2012

Bai, P., Liu, X., Zhang, Y., and Liu, C.: Incorporating vegetation dynamics noticeably improved performance of hydrological model under vegetation greening, Sci. Total Environ., 643, 610-622, 10.1016/j.scitotenv.2018.06.233, 2018

Bai, P., Liu, X., and Liu, C.: Improving hydrological simulations by incorporating GRACE data for model calibration, J.

600 Hydrol., 557, 291-304, 10.1016/j.jhydrol.2017.12.025, 2018

Bao, Z., Zhang, J., Liu, J., Fu, G., Wang, G., He, R., Yan, X., Jin, J., and Liu, H.: 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, 10.1016/j.jhydrol.2012.07.048, 2012

Beck, H.E., Pan, M., Lin, P., Seibert, J., Dijk, A.I.J.M., and Wood, E.F.: Global Fully Distributed Parameter Regionalization

605 Based on Observed Streamflow From 4,229 Headwater Catchments, Journal of Geophysical Research: Atmospheres, 125(17), 10.1029/2019JD031485, 2020

Beck, H.E., van Dijk, A.I.J.M., de Roo, A., Miralles, D.G., Mcvicar, T.R., Schellekens, J., and Bruijnzeel, L.A.: Global-scale regionalization of hydrologic model parameters, Water Resour. Res., 52(5), 3599-3622, 10.1002/2015WR018247, 2016

Beven, K.J.: Rainfall-runoff modelling: the primer, John Wiley & Sons, 2011.

610 Blöschl, G.: Rainfall-Runoff Modeling of Ungauged Catchments, in: Encyclopedia of Hydrological Sciences, edited, 1-19, https://doi.org/10.1002/0470848944.hsa140, 2005.

Bouwman, A., Savchuk, A., Abbaspourghomi, A., and Visser, B.: Automated Step Detection in Inertial Measurement Unit Data From Turkeys, Front Genet, 11, 207, 2020

Cai, M., Yang, S., Zeng, H., Zhao, C., and Wang, S.: A Distributed Hydrological Model Driven by Multi-Source Spatial Data

615 and Its Application in the Ili River Basin of Central Asia, Water Resour. Manag., 28(10), 2851-2866, 10.1007/s11269-014-0641-z, 2014

Castiglioni, S., Lombardi, L., Toth, E., Castellarin, A., and Montanari, A.: Calibration of rainfall-runoff models in ungauged basins: A regional maximum likelihood approach, Adv. Water Resour., 33(10), 2010

Chaplot, V.A.M., and Le Bissonnais, Y.: Runoff Features for Interrill Erosion at Different Rainfall Intensities, Slope Lengths,

620 and Gradients in an Agricultural Loessial Hillslope, Soil Sci. Soc. Am. J., 67(3), 844-851, https://doi.org/10.2136/sssaj2003.8440, 2003

Chiew, F.H.S., and Siriwardena, L.: Estimation of SIMHYD Parameter Values for Application in Ungauged Catchments, edited, Melbourne, AUSTRALIA. p.pp. 2883-2889, 2005.

Dembélé, M., Hrachowitz, M., Savenije, H.H.G., Mariéthoz, G., and Schaefli, B.: Improving the Predictive Skill of a

625 Distributed Hydrological Model by Calibration on Spatial Patterns With Multiple Satellite Data Sets, Water Resour. Res., 56(1), 10.1029/2019WR026085, 2020

Demirel, M.C., Mai, J., Mendiguren, G., Koch, J., Samaniego, L., and Stisen, S.: Combining satellite data and appropriate objective functions for improved spatial pattern performance of a distributed hydrologic model, Hydrol. Earth Syst. Sc., 22(2), 1299-1315, 10.5194/hess-22-1299-2018, 2018

630 Donohue, R.J., Roderick, M.L., and Mcvicar, T.R.: On the importance of including vegetation dynamics in Budyko's hydrological model, Hydrol. Earth Syst. Sci., 11(2), 983-995, 10.5194/hess-11-983-2007, 2007

Donohue, R.J., Roderick, M.L., and Mcvicar, T.R.: Can dynamic vegetation information improve the accuracy of Budyko's hydrological model? J. Hydrol., 390(1-2), 23-34, 10.1016/j.jhydrol.2010.06.025, 2010

Duan, Q., Sorooshian, S., and Gupta, V.K.: Effective and efficient global optimization for conceptual rainfall‐runoff models,

635 Water Resour. Res., 28(4), 1015-1031, 10.1029/91WR02985, 1992

Duan, Q., Sorooshian, S., and Gupta, V.K.: Optimal use of the SCE-UA global optimization method for calibrating watershed models, J. Hydrol., 158(3), 265-284, https://doi.org/10.1016/0022-1694(94)90057-4, 1994

Dunne, T., and Leopold, L.B.: Water in environmental planning, Macmillan, 1978.

Fan, C., Song, C., Liu, K., Ke, L., Xue, B., Chen, T., Fu, C., and Cheng, J.: Century‐Scale Reconstruction of Water Storage

640 Changes of the Largest Lake in the Inner Mongolia Plateau Using a Machine Learning Approach, Water Resour. Res., 57(2), 10.1029/2020WR028831, 2021

Fang, H., Liang, S., Townshend, J.R., and Dickinson, R.E.: Spatially and temporally continuous LAI data sets based on an integrated filtering method: Examples from North America, Remote Sens. Environ., 112(1), 75-93, https://doi.org/10.1016/j.rse.2006.07.026, 2008

645 Finger, D., Vis, M., Huss, M., and Seibert, J.: The value of multiple data set calibration versus model complexity for improving the performance of hydrological models in mountain catchments, Water Resour. Res., 51(4), 1939-1958, 10.1002/2014WR015712, 2015

Freeze, R.A.: Role of subsurface flow in generating surface runoff: 2. Upstream source areas, Water Resour. Res., 8(5), 1272-1283, https://doi.org/10.1029/WR008i005p01272, 1972

650 Freeze, R.A.: Streamflow generation, Rev. Geophys., 12(4), 627-647, https://doi.org/10.1029/RG012i004p00627, 1974

Friedman, J., Hastie, T., and Tibshirani, R.: Additive logistic regression: a statistical view of boosting (With discussion and a rejoinder by the authors), The Annals of Statistics, 28(2), 337-407, 10.1214/aos/1016218223, 2000

Friedman, J.H.: Greedy Function Approximation: A Gradient Boosting Machine, The Annals of Statistics, 29(5), 1189-1232, 10.1214/aos/1013203451, 2001

655 Garg, V., Nikam, B.R., Thakur, P.K., and Aggarwal, S.P.: Assessment of the effect of slope on runoff potential of a watershed using NRCS-CN method, Int. J. of Hydrology Science and Technology, 3(2), 141-159, https://doi.org/10.1504/IJHST.2013.057626, 2013

Gerten, D.: A vital link: water and vegetation in the Anthropocene, Hydrol. Earth Syst. Sc., 17(10), 3841-3852, 10.5194/hess-17-3841-2013, 2013

660   Guo, Y., Zhang, Y., Zhang, L., and Wang, Z.: Regionalization of hydrological modeling for predicting streamflow in ungauged catchments: A comprehensive review, Wires Water, 8(1), 10.1002/wat2.1487, 2021

Gupta, H.V., Kling, H., Yilmaz, K.K., and Martinez, G.F.: Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling, J. Hydrol., 377(1-2), 80-91, 10.1016/j.jhydrol.2009.08.003, 2009

Gupta, H.V., Sorooshian, S., and Yapo, P.O.: Status of Automatic Calibration for Hydrologic Models: Comparison with

665   Multilevel Expert Calibration, J. Hydrol. Eng., 4(2), 10.1061/(ASCE)1084-0699(1999)4:2(135), 1999

He, J., Yang, K., Tang, W., Lu, H., Qin, J., Chen, Y., and Li, X.: The first high-resolution meteorological forcing dataset for land process studies over China, Sci Data, 7(1), 10.1038/s41597-020-0369-y, 2020

He, Y., Bárdossy, A., and Zehe, E.: A review of regionalisation for continuous streamflow simulation, Hydrol. Earth Syst. Sc., 15(11), 3539-3553, 10.5194/hess-15-3539-2011, 2011

670   Heuvelmans, G., Muys, B., and Feyen, J.: Regionalisation of the parameters of a hydrological model: Comparison of linear regression models with artificial neural nets, J. Hydrol., 319(1-4), 245-265, 10.1016/j.jhydrol.2005.07.030, 2006

Hobeichi, S., Abramowitz, G., Evans, J., and Beck, H.E.: Linear Optimal Runoff Aggregate (LORA): A global gridded synthesis runoff product, Hydrology and earth system sciences discussions, 1-26, 10.5194/hess-2018-386, 2018

Hollander, M., Wolfe, D.A., and Chicken, E.: Nonparametric statistical methods, John Wiley & Sons, New York, 2013.

675   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., Gupta, H.V., Hughes, D.A., Hut, R.W., Montanari, A., Pande, S., Tetzlaff, D., Troch, P.A., Uhlenbrook, S., Wagener, T., Winsemius, H.C., Woods, R.A., Zehe, E., and Cudennec, C.: A decade of Predictions in Ungauged Basins (PUB) — a review, Hydrological Sciences Journal, 58(6), 1198-1255, 10.1080/02626667.2013.803183, 2013

680   Huang, C.: Empirical Analysis of Slope and Runoff For Sediment Delivery from Interrill Areas, Soil Sci. Soc. Am. J., 59(4), 982-990, https://doi.org/10.2136/sssaj1995.03615995005900040004x, 1995

Huang, P., Li, Z., Chen, J., Li, Q., and Yao, C.: Event-based hydrological modeling for detecting dominant hydrological process and suitable model strategy for semi-arid catchments, J. Hydrol., 542, 292-303, 10.1016/j.jhydrol.2016.09.001, 2016

Huang, Q., Qin, G., Zhang, Y., Tang, Q., Liu, C., Xia, J., Chiew, F.H.S., and Post, D.: Using Remote Sensing Data‐Based

685   Hydrological Model Calibrations for Predicting Runoff in Ungauged or Poorly Gauged Catchments, Water Resour. Res., 56(8), 10.1029/2020WR028205, 2020

Hundecha, Y., and Bárdossy, A.: Modeling of the effect of land use changes on the runoff generation of a river basin through parameter regionalization of a watershed model, J. Hydrol., 292(1), 281-295, https://doi.org/10.1016/j.jhydrol.2004.01.002, 2004

690   Huntingford, C., Jeffers, E.S., Bonsall, M.B., Christensen, H.M., Lees, T., and Yang, H.: Machine learning and artificial intelligence to aid climate change research and preparedness, Environ Res Lett, 14(12), 124007, 10.1088/1748-9326/ab4e55,

2019

Immerzeel, W.W., and Droogers, P.: Calibration of a distributed hydrological model based on satellite evapotranspiration, J. Hydrol., 349(3-4), 411-424, 10.1016/j.jhydrol.2007.11.017, 2008

695   Ivanov, V.Y., Bras, R.L., and Vivoni, E.R.: Vegetation-hydrology dynamics in complex terrain of semiarid areas: 1. A mechanistic approach to modeling dynamic feedbacks, Water Resour. Res., 44(3), 10.1029/2006WR005588, 2008

Kling, H., Fuchs, M., and Paulin, M.: Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios, J. Hydrol., 424-425, 264-277, https://doi.org/10.1016/j.jhydrol.2012.01.011, 2012

Knoben, W.J.M., Woods, R.A., and Freer, J.E.: A Quantitative Hydrological Climate Classification Evaluated With

700   Independent Streamflow Data, Water Resour. Res., 54(7), 5088-5109, https://doi.org/10.1029/2018WR022913, 2018

Koskinen, M., Tahvanainen, T., Sarkkola, S., Menberu, M.W., Laurén, A., Sallantaus, T., Marttila, H., Ronkanen, A., Parviainen, M., Tolvanen, A., Koivusalo, H., and Nieminen, M.: Restoration of nutrient-rich forestry-drained peatlands poses a risk for high exports of dissolved organic carbon, nitrogen, and phosphorus, Sci. Total Environ., 586, 858-869, https://doi.org/10.1016/j.scitotenv.2017.02.065, 2017

705   Kuczera, G., and Mroczkowski, M.: Assessment of hydrologic parameter uncertainty and the worth of multiresponse data, Water Resour. Res., 34(6), 1481-1489, https://doi.org/10.1029/98WR00496, 1998

Kuhn, M., Wing, J., Weston, S., Williams, A., Keefer, C., Engelhardt, A., Cooper, T., Mayer, Z., Kenkel, B., and Benesty, M.: R Package 'caret': classification and Regression Training, https://github.com/topepo/caret/, 2020

Lawrence, R.: Classification of remotely sensed imagery using stochastic gradient boosting as a refinement of classification

710   tree analysis, Remote Sens. Environ., 90(3), 331-336, 10.1016/j.rse.2004.01.007, 2004

Lei, H., Yang, D., and Huang, M.: Impacts of climate change and vegetation dynamics on runoff in the mountainous region of the Haihe River basin in the past five decades, J. Hydrol., 511, 786-799, 10.1016/j.jhydrol.2014.02.029, 2014

Leuning, R., Zhang, Y., Rajaud, A., Cleugh, H., and Tu, K.: A simple surface conductance model to estimate regional evaporation using MODIS leaf area index and the Penman-Monteith equation, Water Resour. Res., 44(10),

715   https://doi.org/10.1029/2007WR006562, 2008

Li, H., Zhang, Y., Chiew, F.H.S., and Xu, S.: Predicting runoff in ungauged catchments by using Xinanjiang model with MODIS leaf area index, J. Hydrol., 370(1-4), 155-162, 10.1016/j.jhydrol.2009.03.003, 2009

Liao, S., Liu, Z., Liu, B., Cheng, C., Jin, X., and Zhao, Z.: Multistep-ahead daily inflow forecasting using the ERA-Interim reanalysis data set based on gradient-boosting regression trees, Hydrol. Earth Syst. Sc., 24(5), 2343-2363, 10.5194/hess-24-

720   2343-2020, 2020

Lima, A.R., Cannon, A.J., and Hsieh, W.W.: Nonlinear regression in environmental sciences using extreme learning machines: A comparative evaluation, Environ. Modell. Softw., 73, 175-188, 10.1016/j.envsoft.2015.08.002, 2015

Liu, F., Zhang, G., Song, X., Li, D., Zhao, Y., Yang, J., Wu, H., and Yang, F.: High-resolution and three-dimensional mapping of soil texture of China, Geoderma, 361, 114061, 10.1016/j.geoderma.2019.114061, 2020

725   Livneh, B., and Lettenmaier, D.P.: Regional parameter estimation for the unified land model, Water Resour. Res., 49(1), 100-

114, https://doi.org/10.1029/2012WR012220, 2013

Lohmann, D., Nolte-Holube, R., and Raschke, E.: A large-scale horizontal routing model to be coupled to land surface parametrization schemes, Tellus A: Dynamic Meteorology and Oceanography, 48(5), 708-721, 10.3402/tellusa.v48i5.12200, 1996

730   Martens, B., Miralles, D.G., Lievens, H., van der Schalie, R., de Jeu, R.A.M., Fernández-Prieto, D., Beck, H.E., Dorigo, W.A., and Verhoest, N.E.C.: GLEAM v3: satellite-based land evaporation and root-zone soil moisture, Geosci Model Dev, 10(5), 1903-1925, 10.5194/gmd-10-1903-2017, 2017

Mckenzie, N.J., Jacquier, D.W., Ashton, L.J., and Cresswell, H.P.: Estimation of soil properties using the Atlas of Australian Soils, CSIRO Land and Water Technical Report, 11(00), 1-12, 2000

735   Mizukami, N., Clark, M.P., Newman, A.J., Wood, A.W., Gutmann, E.D., Nijssen, B., Rakovec, O., and Samaniego, L.: Towards seamless large-domain parameter estimation for hydrologic models, Water Resour. Res., 53(9), 8020-8040, 10.1002/2017WR020401, 2017

Mohan, S., and Bhaskaran, P.K.: Evaluation of CMIP5 climate model projections for surface wind speed over the Indian Ocean region, Clim. Dynam., 53(9-10), 5415-5435, 10.1007/s00382-019-04874-2, 2019

740   Montanari, A., Young, G., Savenije, H.H.G., Hughes, D., Wagener, T., Ren, L.L., Koutsoyiannis, D., Cudennec, C., Toth, E., Grimaldi, S., Blöschl, G., Sivapalan, M., Beven, K., Gupta, H., Hipsey, M., Schaefli, B., Arheimer, B., Boegh, E., Schymanski, S.J., Di Baldassarre, G., Yu, B., Hubert, P., Huang, Y., Schumann, A., Post, D.A., Srinivasan, V., Harman, C., Thompson, S., Rogger, M., Viglione, A., Mcmillan, H., Characklis, G., Pang, Z., and Belyaev, V.: "Panta Rhei—Everything Flows": Change in hydrology and society—The IAHS Scientific Decade 2013–2022, Hydrological Sciences Journal, 58(6), 1256-

745   1275, 10.1080/02626667.2013.809088, 2013

Montgomery, D.R., and Dietrich, W.E.: Runoff generation in a steep, soil-mantled landscape, Water Resour. Res., 38(9), 1-7, 10.1029/2001WR000822, 2002

Nash, J.E., and Sutcliffe, J.V.: River flow forecasting through conceptual models part I — A discussion of principles, J. Hydrol., 10(3), 282-290, https://doi.org/10.1016/0022-1694(70)90255-6, 1970

750   Natekin, A., and Knoll, A.: Gradient boosting machines, a tutorial, Front Neurorobotics, 7, 10.3389/fnbot.2013.00021, 2013

Newman, A.J., Mizukami, N., Clark, M.P., Wood, A.W., Nijssen, B., and Nearing, G.: Benchmarking of a Physically Based Hydrologic Model, J. Hydrometeorol., 18(8), 2215-2225, 10.1175/JHM-D-16-0284.1, 2017

Nijzink, R.C., Almeida, S., Pechlivanidis, I.G., Capell, R., Gustafssons, D., Arheimer, B., Parajka, J., Freer, J., Han, D., Wagener, T., Nooijen, R.R.P., Savenije, H.H.G., and Hrachowitz, M.: Constraining Conceptual Hydrological Models With

755   Multiple Information Sources, Water Resour. Res., 54(10), 8332-8362, 10.1029/2017WR021895, 2018

Ning, L., Xia, J., Zhan, C., and Zhang, Y.: Runoff of arid and semi-arid regions simulated and projected by CLM-DTVGM and its multi-scale fluctuations as revealed by EEMD analysis, J. Arid Land, 8(04), 506-520, 10.1007/s40333-016-0126-4, 2016

Oubeidillah, A.A., Kao, S.C., Ashfaq, M., Naz, B.S., and Tootle, G.: A large-scale, high-resolution hydrological model parameter data set for climate change impact assessment for the conterminous US, Hydrol. Earth Syst. Sc., 18(1), 67-84, 10.5194/hess-18-67-2014, 2014

Oudin, L., Andréassian, V., Perrin, C., Michel, C., and Le Moine, N.: Spatial proximity, physical similarity, regression and ungaged catchments: A comparison of regionalization approaches based on 913 French catchments, Water Resour. Res., 44(3), https://doi.org/10.1029/2007WR006240, 2008

Oudin, L., Kay, A., Andréassian, V., and Perrin, C.: Are seemingly physically similar catchments truly hydrologically similar? Water Resour. Res., 46(11), https://doi.org/10.1029/2009WR008887, 2010

Pagliero, L., Bouraoui, F., Diels, J., Willems, P., and Mcintyre, N.: Investigating regionalization techniques for large-scale hydrological modelling, J. Hydrol., 570, 220-235, 10.1016/j.jhydrol.2018.12.071, 2019

Parajka, J., Andreassian, V., Archfield, S., B A Rdossy, A., Chiew, F., Duan, Q., Gelfan, A., Hlavcova, K., Merz, R., Mcintyre, N., Oudin, L., Perrin, C., Rogger, M., Salinas, J.L., Savenije, H., Sk O Ien, J.O., Wagener, T., Zehe, E., and Zhang, Y.: Predictions of runoff hydrographs in ungauged basins, in: Runoff prediction in ungauged basins: Synthesis across processes, places and scales, edited, Cambridge university press, 227-269, 2013.

Parajka, J., Merz, R., and Blöschl, G.: A comparison of regionalisation methods for catchment model parameters, Hydrol. Earth Syst. Sc., 9(3), 157-171, 10.5194/hess-9-157-2005, 2005

Parajka, J., Viglione, A., Rogger, M., Salinas, J.L., Sivapalan, M., and Blöschl, G.: Comparative assessment of predictions in ungauged basins – Part 1: Runoff-hydrograph studies, Hydrol. Earth Syst. Sc., 17(5), 1783-1795, 10.5194/hess-17-1783-2013, 2013

Post, D.A., and Jakeman, A.J.: RELATIONSHIPS BETWEEN CATCHMENT ATTRIBUTES AND HYDROLOGICAL RESPONSE CHARACTERISTICS IN SMALL AUSTRALIAN MOUNTAIN ASH CATCHMENTS, Hydrol. Process., 10(6), 877-892, https://doi.org/10.1002/(SICI)1099-1085(199606)10:6<877::AID-HYP377>3.0.CO;2-T, 1996

Prieto, C., Le Vine, N., Kavetski, D., García, E., and Medina, R.: Flow Prediction in Ungauged Catchments Using Probabilistic Random Forests Regionalization and New Statistical Adequacy Tests, Water Resour. Res., 55(5), 4364-4392, 10.1029/2018WR023254, 2019

Rajaee, T., Ebrahimi, H., and Nourani, V.: A review of the artificial intelligence methods in groundwater level modeling, J. Hydrol., 572, 336-351, https://doi.org/10.1016/j.jhydrol.2018.12.037, 2019

Razavi, T., and Coulibaly, P.: Streamflow Prediction in Ungauged Basins: Review of Regionalization Methods, J. Hydrol. Eng., 18(8), 958-975, 10.1061/(ASCE)HE.1943-5584.0000690, 2013

Ruffin, C., King, R.L., and Younan, N.H.: A Combined Derivative Spectroscopy and Savitzky-Golay Filtering Method for the Analysis of Hyperspectral Data, Gisci Remote Sens, 45(1), 1-15, 10.2747/1548-1603.45.1.1, 2008

Samuel, J., Coulibaly, P., and Metcalfe, R.A.: Estimation of Continuous Streamflow in Ontario Ungauged Basins: Comparison of Regionalization Methods, J. Hydrol. Eng., 16(5), 447-459, 10.1061/(ASCE)HE.1943-5584.0000338, 2011

Schütz, N., Leichtle, A.B., and Riesen, K.: A comparative study of pattern recognition algorithms for predicting the inpatient

mortality risk using routine laboratory measurements, Artif. Intell. Rev., 52(4), 2559-2573, 10.1007/s10462-018-9625-3, 2019

Sefton, C.E.M., and Howarth, S.M.: Relationships between dynamic response characteristics and physical descriptors of catchments in England and Wales, J. Hydrol., 211(1), 1-16, https://doi.org/10.1016/S0022-1694(98)00163-2, 1998

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

Shen, C.: A Transdisciplinary Review of Deep Learning Research and Its Relevance for Water Resources Scientists, Water Resour. Res., 54(11), 8558-8593, 10.1029/2018WR022643, 2018

Shu, C., and Ouarda, T.B.M.J.: Improved methods for daily streamflow estimates at ungauged sites, Water Resour. Res., 48(2), 10.1029/2011WR011501, 2012

Sun, A.Y., Wang, D., and Xu, X.: Monthly streamflow forecasting using Gaussian Process Regression, J. Hydrol., 511, 72-81, 10.1016/j.jhydrol.2014.01.023, 2014

Sutanudjaja, E.H., van Beek, R., Wanders, N., Wada, Y., Bosmans, J.H.C., Drost, N., van der Ent, R.J., de Graaf, I.E.M., Hoch, J.M., de Jong, K., Karssenberg, D., López López, P., Peßenteiner, S., Schmitz, O., Straatsma, M.W., Vannametee, E., Wisser, D., and Bierkens, M.F.P.: PCR-GLOBWB 2: a 5 arcmin global hydrological and water resources model, Geosci Model Dev, 11(6), 2429-2453, 10.5194/gmd-11-2429-2018, 2018

Tarboton, D.G.: Rainfall-runoff processes, Utah State University, 1(2), 2003

Taylor, K.E.: Summarizing multiple aspects of model performance in a single diagram, Journal of Geophysical Research: Atmospheres, 106(D7), 7183-7192, https://doi.org/10.1029/2000JD900719, 2001

Tesfa, T.K., Tarboton, D.G., Chandler, D.G., and Mcnamara, J.P.: Modeling soil depth from topographic and land cover attributes, Water Resour. Res., 45(10), https://doi.org/10.1029/2008WR007474, 2009

Thompson, S.E., Harman, C.J., Konings, A.G., Sivapalan, M., Neal, A., and Troch, P.A.: Comparative hydrology across AmeriFlux sites: The variable roles of climate, vegetation, and groundwater, Water Resour. Res., 47(10), 10.1029/2010WR009797, 2011

van Dijk, A.I.J.M., and Bruijnzeel, L.A.: Modelling rainfall interception by vegetation of variable density using an adapted analytical model. Part 1. Model description, Journal of hydrology (Amsterdam), 247(3-4), 230-238, 10.1016/s0022-1694(01)00392-4, 2001

Vandewiele, G.L., and Elias, A.: Monthly water balance of ungauged catchments obtained by geographical regionalization, J. Hydrol., 170(1), 277-291, https://doi.org/10.1016/0022-1694(95)02681-E, 1995

Wang, G., Xia, J., and Chen, J.: Quantification of effects of climate variations and human activities on runoff by a monthly water balance model: A case study of the Chaobai River basin in northern China, Water Resour. Res., 45(7), 10.1029/2007WR006768, 2009

Wang, Y., Nan, Z., Chen, H., and Wu, X.: Correction of daily precipitation data of ITPCAS dataset over the Qinghai-Tibetan Plateau with KNN model, in: 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 593-596, 2016.

Waseem, M., Ajmal, M., and Kim, T.: Improving the flow duration curve predictability at ungauged sites using a constrained hydrologic regression technique, Ksce J. Civ Eng, 20(7), 3012-3021, 10.1007/s12205-016-0038-z, 2016

Wei, Z., Meng, Y., Zhang, W., Peng, J., and Meng, L.: Downscaling SMAP soil moisture estimation with gradient boosting decision tree regression over the Tibetan Plateau, Remote Sens. Environ., 225, 30-44, 10.1016/j.rse.2019.02.022, 2019

Wheater, H., Sorooshian, S., and Sharma, K.D.: Hydrological modelling in arid and semi-arid areas, Cambridge University Press, 2007.

Xenochristou, M., Hutton, C., Hofman, J., and Kapelan, Z.: Water Demand Forecasting Accuracy and Influencing Factors at Different Spatial Scales Using a Gradient Boosting Machine, Water Resour. Res., 56(8), 10.1029/2019WR026304, 2020

Xia, J., Wang, G., Tan, G., Ye, A., and Huang, G.H.: Development of distributed time-variant gain model for nonlinear hydrological systems, Science in China Series D: Earth Sciences, 48(6), 713-723, 10.1360/03yd0183, 2005

Xia, R., Wang, G., Zhang, Y., Yang, P., Yang, Z., Ding, S., Jia, X., Yang, C., Liu, C., Ma, S., Lin, J., Wang, X., Hou, X., Zhang, K., Gao, X., Duan, P., and Qian, C.: River algal blooms are well predicted by antecedent environmental conditions, Water Res., 185, 116221, 10.1016/j.watres.2020.116221, 2020

Xie, K., Liu, P., Zhang, J., Wang, G., Zhang, X., and Zhou, L.: Identification of spatially distributed parameters of hydrological models using the dimension-adaptive key grid calibration strategy, J. Hydrol., 125772, 10.1016/j.jhydrol.2020.125772, 2020

Xu, C.Y.: Estimation of Parameters of a Conceptual Water Balance Model for Ungauged Catchments, Water Resour. Manag., 13(5), 353-368, 10.1023/A:1008191517801, 1999

Yan, J., Jia, S., Lv, A., and Zhu, W.: Water Resources Assessment of China's Transboundary River Basins Using a Machine Learning Approach, Water Resour. Res., 55(1), 632-655, 10.1029/2018WR023044, 2019

Yang, K., He, J., Tang, W., Qin, J., and Cheng, C.C.K.: On downward shortwave and longwave radiations over high altitude regions: Observation and modeling in the Tibetan Plateau, Agr. Forest Meteorol., 150(1), 38-46, 10.1016/j.agrformet.2009.08.004, 2010

Yang, K., and He, J.: China meteorological forcing dataset (1979-2018), National Tibetan Plateau Data Center, 10.11888/AtmosphericPhysics.tpe.249369.file, 2019

Yang, X., Magnusson, J., Huang, S., Beldring, S., and Xu, C.: Dependence of regionalization methods on the complexity of hydrological models in multiple climatic regions, J. Hydrol., 582(C), 124357, https://doi.org/10.1016/j.jhydrol.2019.124357, 2020

Yang, X., Magnusson, J., Rizzi, J., and Xu, C.: Runoff prediction in ungauged catchments in Norway: comparison of regionalization approaches, Hydrol Res, 49(2), 487-505, 10.2166/nh.2017.071, 2018

Yang, X., Yong, B., Ren, L., Zhang, Y., and Long, D.: Multi-scale validation of GLEAM evapotranspiration products over China via ChinaFLUX ET measurements, Int. J. Remote Sens., 38(20), 5688-5709, 10.1080/01431161.2017.1346400, 2017

Yang, Y., Pan, M., Beck, H.E., Fisher, C.K., Beighley, R.E., Kao, S.C., Hong, Y., and Wood, E.F.: In Quest of Calibration Density and Consistency in Hydrologic Modeling: Distributed Parameter Calibration against Streamflow Characteristics, Water Resour. Res., 55(9), 7784-7803, 10.1029/2018WR024178, 2019

Yaseen, Z.M., El-Shafie, A., Jaafar, O., Afan, H.A., and Sayl, K.N.: Artificial intelligence based models for stream-flow forecasting: 2000 – 2015, J. Hydrol., 530, 829-844, 10.1016/j.jhydrol.2015.10.038, 2015

Young, A.R.: Stream flow simulation within UK ungauged catchments using a daily rainfall-runoff model, J. Hydrol., 320(1), 155-172, https://doi.org/10.1016/j.jhydrol.2005.07.017, 2006

865 Zecharias, Y.B., and Brutsaert, W.: Recession characteristics of groundwater outflow and base flow from mountainous watersheds, Water Resour. Res., 24(10), 1651-1658, https://doi.org/10.1029/WR024i010p01651, 1988

Zeng, S., Xia, J., Chen, X., Zou, L., Du, H., and She, D.: Integrated land-surface hydrological and biogeochemical processes in simulating water, energy and carbon fluxes over two different ecosystems, J. Hydrol., 582, 124390, 10.1016/j.jhydrol.2019.124390, 2020

870 Zhan, C., Song, X., Xia, J., and Tong, C.: An efficient integrated approach for global sensitivity analysis of hydrological model parameters, Environ. Modell. Softw., 41, 39-52, 10.1016/j.envsoft.2012.10.009, 2013

Zhang, X., Tang, Q., Pan, M., and Tang, Y.: A Long-Term Land Surface Hydrologic Fluxes and States Dataset for China, J. Hydrometeorol., 15(5), 2067-2084, 10.1175/JHM-D-13-0170.1, 2014

Zhang, Y., Chiew, F.H.S., Li, M., and Post, D.: Predicting Runoff Signatures Using Regression and Hydrological Modeling
875 Approaches, Water Resour. Res., 54(10), 7859-7878, 10.1029/2018WR023325, 2018

Zhang, Y., Chiew, F.H.S., Liu, C., Tang, Q., Xia, J., Tian, J., Kong, D., and Li, C.: Can Remotely Sensed Actual Evapotranspiration Facilitate Hydrological Prediction in Ungauged Regions Without Runoff Calibration? Water Resour. Res., 56(1), 10.1029/2019WR026236, 2020

Zhang, Y., Chiew, F.H.S., Peña-Arancibia, J., Sun, F., Li, H., and Leuning, R.: Global variation of transpiration and soil
880 evaporation and the role of their major climate drivers, Journal of Geophysical Research: Atmospheres, 122(13), 6868-6881, 10.1002/2017JD027025, 2017

Zhang, Y., Chiew, F.H.S., Zhang, L., Leuning, R., and Cleugh, H.A.: Estimating catchment evaporation and runoff using MODIS leaf area index and the Penman-Monteith equation, Water Resour. Res., 44(10), 10.1029/2007WR006563, 2008

Zhang, Y., Chiew, F.H.S., Zhang, L., and Li, H.: Use of Remotely Sensed Actual Evapotranspiration to Improve Rainfall –
885 Runoff Modeling in Southeast Australia, J. Hydrometeorol., 10(4), 969-980, 10.1175/2009JHM1061.1, 2009

Zhang, Y.Q., and Chiew, F.: Evaluation of regionalisation methods for predicting runoff in ungauged catchments in southeast Australia, in: 18th World IMACS/MODSIM Congress, Cairns, Australia, 3442-3448, 2009.

Zhang, Z., Zhang, Q., and Singh, V.P.: Univariate streamflow forecasting using commonly used data-driven models: literature review and case study, Hydrological Sciences Journal, 63(7), 1091-1111, 10.1080/02626667.2018.1469756, 2018

890 Zhou, Y., Zhang, Y., Vaze, J., Lane, P., and Xu, S.: Improving runoff estimates using remote sensing vegetation data for bushfire impacted catchments, Agr. Forest Meteorol., 182-183, 332-341, 10.1016/j.agrformet.2013.04.018, 2013

Zhuo, L., and Han, D.: Could operational hydrological models be made compatible with satellite soil moisture observations? Hydrol. Process., 30(10), 1637-1648, 10.1002/hyp.10804, 2016

Zou, L., Zhan, C., Xia, J., Wang, T., and Gippel, C.J.: Implementation of evapotranspiration data assimilation with catchment

895   scale distributed hydrological model via an ensemble Kalman Filter, J. Hydrol., 549, 685-702, 10.1016/j.jhydrol.2017.04.036, 2017