CNRD v1.0

A High-Quality Natural Runoff Dataset for Hydrological and Climate Studies in China

Jiaojiao Gou, Chiyuan Miao, Luis Samaniego, Mu Xiao, Jingwen Wu, and Xiaoying Guo

ABSTRACT: Reliable, spatiotemporally continuous runoff records are necessary for identifying climate change impacts and planning effective water management strategies. Existing Chinese runoff data to date have been produced from sparse, poor-quality gauge measurements at different time scales. We have developed a new, quality-controlled gridded runoff dataset, the China Natural Runoff Dataset version 1.0 (CNRD v1.0), which provides daily, monthly, and annual 0.25° runoff estimates for the period 1961–2018 over China. CNRD v1.0 was generated using the Variable Infiltration Capacity (VIC) model. A comprehensive parameter uncertainty analysis framework incorporating parameter sensitivity analysis, optimization, and regionalization with 200 natural or near-natural gauge catchments was used to train the VIC model. Overall, the results show well-calibrated parameters for most gauged catchments except arid and semiarid areas, and the skill scores present high values for all catchments. For the pseudo-/test-ungauged catchments, the model parameters estimated by the multiscale parameter regionalization technique offer the best regionalization solution. CNRD v1.0 is the first free public dataset of gridded natural runoff estimated using a comprehensive model parameter uncertainty analysis framework for China. These results indicate that CNRD v1.0 has high potential for application to long-term hydrological and climate studies in China and to improve international runoff databases for global-scale studies.

KEYWORDS: Runoff; Databases; Hydrologic cycle; Hydrologic models; Land surface model

AFFILIATIONS: Gou, Miao, Wu, and Guo—State Key Laboratory of Earth Surface Processes and Resource Ecology, Faculty of Geographical Science, Beijing Normal University, Beijing, China; Samaniego—UFZ-Helmholtz Centre for Environmental Research, Leipzig, Germany; Xiao—Department of Geography, University of California, Los Angeles, Los Angeles, California

Water is a key natural resource on which all life on Earth depends, and it is essential for societies to flourish (Gudmundsson et al. 2017). Although it is believed that the amount of water will not diminish on shorter than geological time scales (Oki and Kanae 2006), the extremely low fraction of available freshwater and the increasing water demand from a continuously growing human population have made water scarcity one of the major problems of the twenty-first century. Meanwhile, changes in atmospheric thermodynamics in the context of ongoing global warming—seen in, for example, increasingly extreme precipitation events (Papalexiou and Montanari 2019; Miao et al. 2019), higher potential evapotranspiration, and earlier snowmelt seasons—have intensified the global water cycle (Huntington 2006; Miao et al. 2016; Oki and Kanae 2006), further aggravating global water stress. In addition, water is also one of the most direct mediums through which people can perceive the effects of climate change. Changes in frequency and duration of hydrological extremes, such as droughts (Samaniego et al. 2018), floods (Ficchì and Stephens 2019), and glacial melting (Gao et al. 2019), are prime ways to experience climatic change impacts on the environment and the socio-economic conditions of a region. In this context, better tools, such as improved hydrological flow databases, are needed for managing terrestrial water resources and detecting the effect of climate change in space and time at resolutions suitable for hydrometeorology studies (e.g., Berghuijs et al. 2017; Gudmundsson et al. 2017).

Generally, ground-based hydrological gauge networks are the main sources of streamflow data. Observed river flow data have been widely used in various hydroclimate studies for applications such as the design of water distribution systems and irrigation networks (Hu et al. 2010; Tetzlaff et al. 2017) and studies of climate change impacts on water resources (Tang et al. 2019). Although gauges can directly measure streamflow from river channels and the hillslope, such point measurements are sparse; also, the number of hydrometric stations has declined in many parts of the world (Mishra and Coulibaly 2009). Previous studies have also highlighted that current data collection networks are inadequate for providing the information required to understand and explain changes in natural systems (Mishra and Coulibaly 2009; Mitchell and Shrubsole 1994). For example, there is an extreme lack of gauged streamflow data in the Arctic and the Tibetan Plateau owing to their harsh environments. In addition, flow measurements usually have data gaps due to technical or maintenance issues; for instance, hydrological stations can be damaged during flood events (Gao et al. 2018; Tencaliec et al. 2015). These missing records in the datasets can cause erroneous summary data interpretation or unreliable scientific analysis (Tencaliec et al. 2015). At the same time, anthropogenic influences on natural processes are large and widespread. The global area of irrigated agriculture constitutes 40% of the total area used for agricultural production (Fig. 1a; Meier et al. 2018), and the unprecedented numbers of dams and reservoirs constructed have resulted in about 7,320 large dams (capacity ≥ 0.1 km^3) being in operation worldwide in 2019 (Fig. 1b; the Global Reservoir and Dam Database v1.3, http://globaldamwatch .org/data/). Thus, natural hydrological cycles have been dramatically modified by human activities such as domestic water withdrawal, irrigation, reservoir regulation, and river diversions (Oki and Kanae 2006), and so gauge measurements cannot solely represent natural hydrological processes. Overall, hydrometeorologists face a great challenge in using gauged flow data to capture the variability signal and assess long-term trends of natural hydrological cycles.
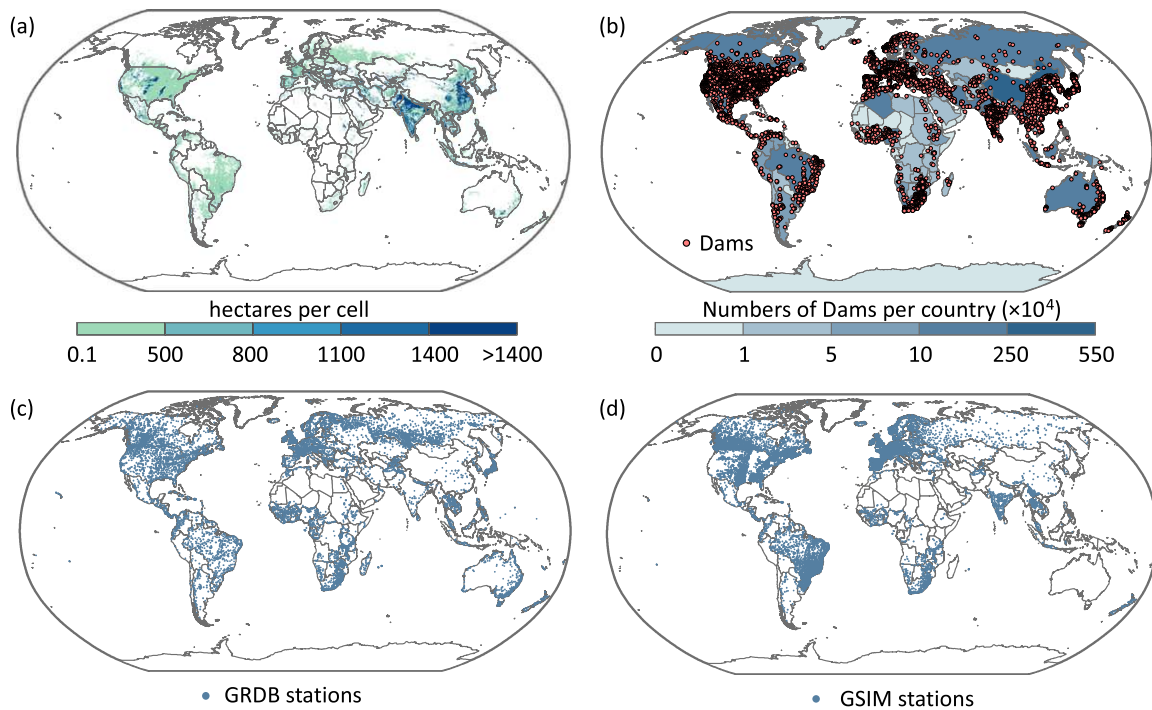
Fig. 1. Distribution of (a) global irrigated areas, (b) constructed large dams (capacity ≥ 0.1 km³; according to GRanD), (c) Global Runoff Data Base (GRDB) hydrological stations, and (d) Global Streamflow Indices and Metadata archive (GSIM) hydrological stations.

Reconstruction of continuous natural flow data records has been an emerging research area in the past three decades. Since the 1980s, the Global Runoff Data Base (GRDB) that operates under the auspices of the World Meteorological Organization has been steadily updated based on submissions from national authorities (Do et al. 2018). To date, the GRDB comprises discharge data from more than 9,900 gauging stations worldwide and is the primary dataset used in large-scale hydroclimate studies (e.g., Ficchì and Stephens 2019; Markonis et al. 2018). The Global Streamflow Indices and Metadata (GSIM) archive is a worldwide collection of streamflow metadata and indices, which is derived from a total of 35,002 hydrological stations used in 12 global or regional streamflow databases (Do et al. 2018). The Dai and Trenberth Global River Flow and Continental Discharge Dataset has recorded the historical monthly streamflow into the oceans at the farthest downstream stations for the world's 925 largest ocean-reaching rivers (Dai 2017). The Global Streamflow Characteristics Dataset (Beck et al. 2015) and the Global Runoff Reconstruction dataset (GRUN; Ghiggi et al. 2019) are two other global flow datasets developed by a neural network and a machine learning algorithm, respectively. These global flow datasets undoubtedly provide fundamental records for water resources management and climate change monitoring around the world. However, construction of these global datasets remains resource intensive, and they contain notable data gaps in some regions, especially in China (Figs. 1c–d and Fig. ES1; GRDB and GSIM are two examples).

China is climate vulnerable due to its remarkable topographic gradients, monsoon climate, and rapid economic development (Miao et al. 2016). Climate change has also increased the urgency of understanding, regulating, and forecasting China's freshwater flows (Piao et al. 2010). Such work requires reliable, spatiotemporally continuous runoff records; however, current flow datasets across China are inadequate. As with the global datasets, high-quality national-scale datasets, such as the Long-Term Land Surface Hydrologic Fluxes and States Dataset for China produced by Zhang et al. (2014), also use data from relatively few stations. Thus, this study presents a new long-term, high-quality runoff dataset, the China natural runoff dataset version 1.0 (CNRD v1.0), which spans 1961–2018 with daily

and monthly temporal and 0.25° spatial resolution. The gridded runoff record in CNRD v1.0 was reconstructed based on a distributed land surface model combined with a comprehensive parameter uncertainty analysis framework, i.e., one that includes parameter sensitivity analysis, optimization, and regionalization. The 200 natural or near-natural gauge stations with the lowest fractions of missing data were used to train the model—a larger number than in previous studies. All of these characteristics make CNRD v1.0 a useful data product for driving hydrological and climate studies over China, especially for ungauged or poorly gauged areas, and the data will also contribute to improving global runoff databases.

Methods and data sources

Hydrological modeling and data sources. The Variable Infiltration Capacity (VIC) macroscale hydrological model (Liang et al. 1994) was used to produce CNRD v1.0. VIC is an offline land surface model that is able to capture transient basin discharge (Gou et al. 2020) and project the terrestrial water cycle (X. G. He et al. 2020; Sheffield et al. 2014). In this study, we ran the VIC model version 4.2 in water balance mode in two stages: model training and data production. The former refers to the stage that creates a high-reliability land surface model to replicate "natural" hydrology processes based on the routing structure for the basin with natural or near-natural streamflow records. Given that China has experienced rapid economic development since the 1980s, the period 1960–79 was chosen for this model training stage to reduce anthropogenic influences, such as dam construction, urbanization, and land–use changes on natural hydrological processes (Gou et al. 2020). The data production stage forms a long-term natural runoff series covering the period 1961–2018. We ran the trained model with a 6-hourly time step at 0.25° spatial resolution, spanning a total of 15,775 grid cells across mainland China (hereinafter "China").

To drive the model, we used 0.25° gridded daily precipitation, maximum temperature, minimum temperature, and wind speed from two gridded datasets for the period 1961–2018. The first group of meteorological datasets covers most of these years (1961–2014) and was constructed using data from ~2,400 weather stations (Fig. ES2) acquired from the China Meteorological Administration. The second dataset, the China Meteorological Forcing Dataset (J. He et al. 2020), which supplied the rest of the years (2015–18), was made through fusion of remote sensing products, reanalysis datasets, and in situ station data. These two meteorological datasets passed the consistency test during the overlap period (Fig. ES3). Other climate forcing variables, such as downward shortwave radiation, longwave radiation, vapor pressure, and air pressure were simulated offline by the Mountain Microclimate Simulation module of VIC (Bohn et al. 2013). The required lower boundary conditions for the model include soil, vegetation, and topographical data and are detailed in our previous model work (Gou et al. 2020).

For the model training stage, we used 200 natural or near-natural gauged catchments over 10 river basins across China to optimize the simulated natural runoff (Fig. 2). The gauged monthly streamflow was obtained from the hydrological yearbook of China and local water resources departments. The gauges were categorized into three groups: 1) naturalized gauges where the influences of human activities are removed, 2) near-natural gauges without dams or reservoirs upstream, and 3) a few gauges with a low level of dam influence (Fig. 2). Naturalized gauges without water management effects (e.g., irrigation and reservoir regulation) were developed by the Bureau of Hydrology of the Chinese Ministry of Water Resources based on the water balance principle (see details in appendix A). Gauges without dams or reservoirs upstream before 1980 may be seen as near-natural catchments, and could therefore be used directly with naturalized gauges for model training. We were particularly cautious with the third catchment group, i.e., those that were gauged and had influenced by dams. Although the flow regimes are well balanced, with no discernible abrupt changes or shifts in the third
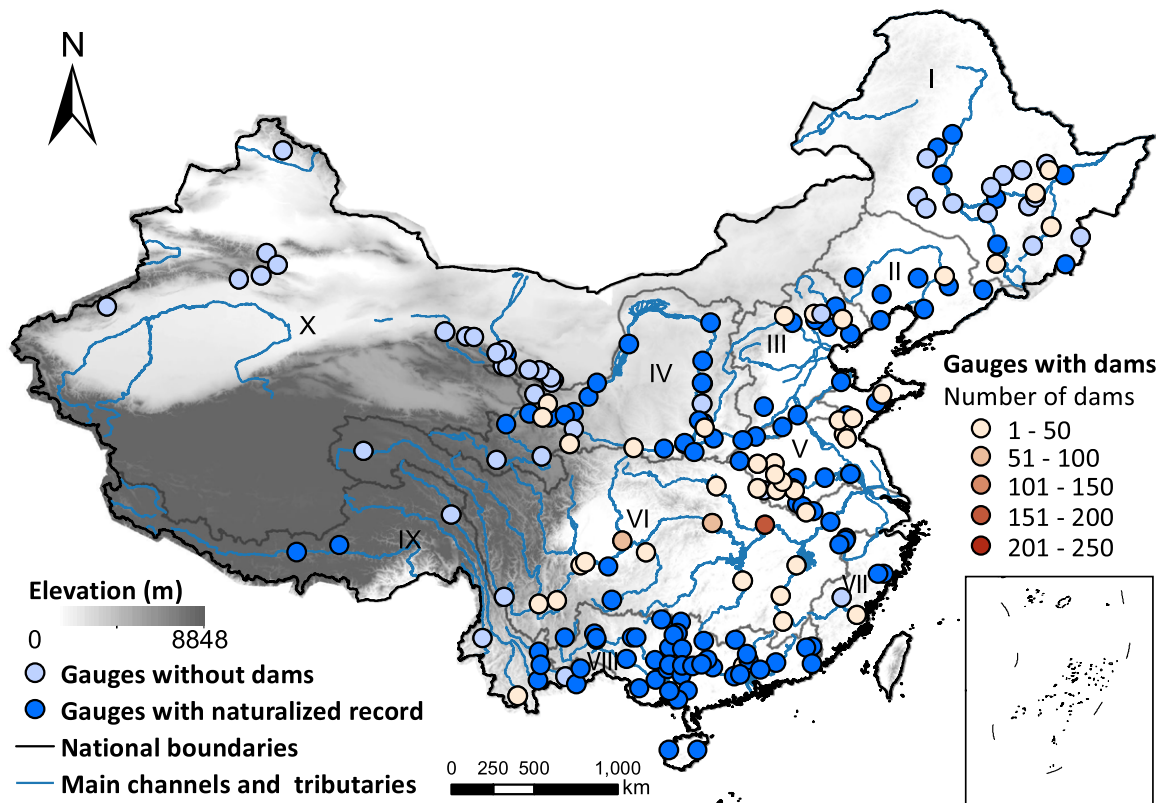
Fig. 2. Gauge records classified according to whether they are 1) naturalized (dark blue), 2) observed but without dam influence (light blue), or 3) influenced by dams (yellow to red). For those influenced by dams, the color gradation indicates the number of upstream dams. The number of dams before 1979 was extracted from a dam point dataset (GRanD v1.3, http://globaldamwatch.org/data/). The gray boundaries indicate the 10 major river basin areas over China: I—Songhua River; II—Liao River; III—Hai River; IV—Yellow River; V—Huai River; VI—Yangtze River; VII—Southeast River drainage system; VIII—Pearl River; IX—Southwest River drainage system; and X—Northwest River drainage system.

catchment group (Fig. ES4), parameters from this type of gauged catchment were not applied to the ungauged catchments. For the data production stage, we evaluated the performance of CNRD v1.0 and, for the sake of comparison, an ensemble of 18 simulations from the second phase of the Inter-Sectoral Impact Model Intercomparison Project (ISIMIP2a; https://esg.pik-potsdam.de/search/isimip/), and a recently published dataset of gridded runoff data (GRUN; Ghiggi et al. 2019). The evaluation was performed at a monthly temporal scale for the over-lapping time periods between 1971 and 2010 from these datasets. To be consistent with the resolution of two global products, we aggregated CNRD v1.0 to 0.5° × 0.5° grid to calculate the cell-to-cell accordance. Further details on ISIMIP2a and GRUN can be found in appendix B.

Overview of parameter uncertainty analysis framework. In this study, the parameter uncertainty analysis framework is used to train the VIC land surface model using the 200 in situ streamflow observation sites shown in Fig. 2. The framework for this model training is shown in Fig. 3 and includes parameter sensitivity analysis, parameter optimization, and parameter regionalization. First, we screened the important parameters for runoff simulation from a set of runoff-related parameters (Fig. 3a). The parameter sensitivity analysis is needed to reduce parameter dimensionality because having a large number of tunable parameters creates a heavy burden for parameter optimization (Bennett et al. 2018; Cuntz et al. 2015). Parameter optimization followed the parameter sensitivity analysis. For each of the 200 naturalized or near-natural catchments, an adaptive surrogate modeling-based optimization (ASMO)
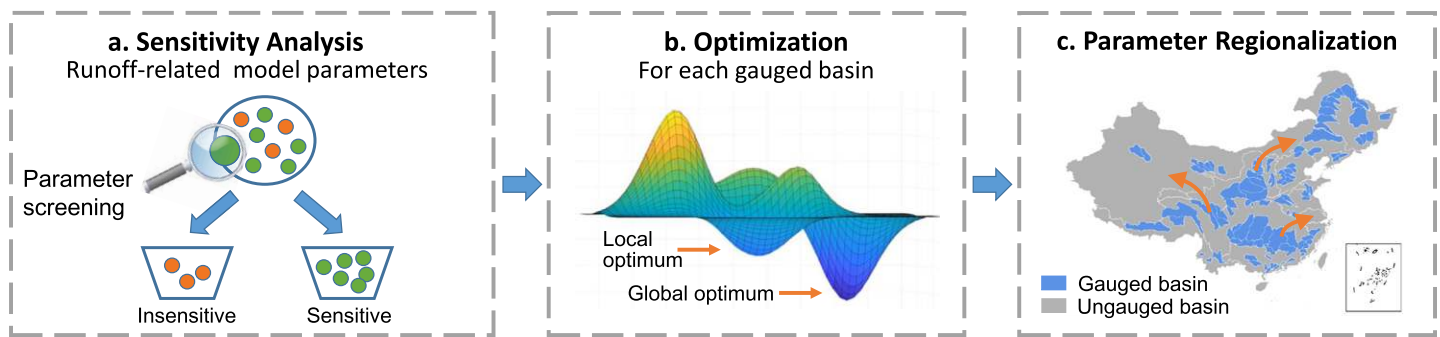
Fig. 3. Parameter uncertainty analysis framework, including (a) sensitivity analysis, (b) optimization, and (c) regionalization.

algorithm was adopted for finding the optimal parameter solution (Fig. 3b; see details in the "Parameter sensitivity analysis and optimization for gauged catchments" section). A certain amount of parameter tuning is necessary to produce model predictions that can match corresponding observations (Gupta et al. 1999). However, even though all hydrologic models can to some degree benefit from parameter calibration to improve their runoff simulations in gauged catchments, the large area of unavailable streamflow observations make regionalization approaches extremely important in transferring information from gauged (donor) to ungauged (receptor) catchments (Beck et al. 2016; Parajka et al. 2013). Therefore, we designed a series of experiments to investigate parameter uncertainties involved with predicting runoff for the ungauged catchments (Fig. 3c; see details in the "Parameter regionalization evaluation for ungauged catchments" section). To assess the parameter performance of the regionalization methods for runoff estimation in ungauged catchments, each of the 200 catchments was used in turn as if it were ungauged, following a jackknife cross-validation procedure (i.e., pseudo/test ungauged).

Parameter sensitivity analysis and optimization for gauged catchments. The parameter sensitivity analysis is based on three qualitative global sensitivity analysis (GSA) methods—the sum-of-trees model, the multivariate adaptive regression splines technique, and the delta test—and one quantitative GSA method, the metamodel-based Sobol′ method. For each large river basin, 6,000 training simulations were run for the period from 1960 to 1979 based on samples of parameter combinations obtained from the Sobol′ sequence for selected catchments. Then the sensitivity scores for all streamflow-related parameters were computed by combining both qualitative and quantitative GSA methods based on the training simulations. Once we ascertained the important parameters, then we used an automatic optimization algorithm to estimate the optimal parameter sequences for 200 natural or near-natural gauged catchments. The ASMO algorithm developed by our team members, Wang et al. (2014), was used in this study to optimize the catchment-specific sensitive parameters of the VIC model. The initial sampling was conducted using the Sobol′ sequence—one of the quasi-Monte Carlo sampling methods—and the sample size was set equal to 20 times the number of sensitive parameters. Gaussian processes were used to construct an error response surface (i.e., the surrogate model) by using the initial sample points. Then parameter optimization of the surrogate model and adaptive sampling of the existing response surface were repeated until the convergence criteria for parameter optimization of the real physical model were met. A global optimization algorithm—shuffled complex evolution (Duan et al. 1992)—and the minimum interpolating surface method were used as the core optimization algorithm and adaptive sampling strategy, respectively. We define the performance of streamflow simulation (the output variable of interest), by calculating the Nash–Sutcliffe model efficiency coefficient (NSE) of monthly streamflow for the parameter sensitivity analysis and optimization. We tuned

the catchment-specific sensitive parameters in the VIC model during the calibration period (1961–69); then the tuned parameters were validated for the validation period (1970–79). More information about the parameter sensitivity analysis and optimization methods can be seen in the online supplemental material.

Parameter regionalization evaluation for ungauged catchments. Parameter estimation for ungauged catchments for which no streamflow data are available and hence no direct parameter calibration is possible remains one of the biggest challenges (Mizukami et al. 2017; Oudin et al. 2008; Samaniego et al. 2010). Therefore, hydrologists have developed a series of parameter transfer strategies, such as regression (Magette et al. 1976), catchment grouping (similarity based; Burn and Boorman 1993), and simultaneous regionalization (Hundecha and Bárdossy 2004), which can be used to estimate model parameter values on any ungauged catchment in a definable region of consistent hydrological response. Three parameter transfer approaches were considered in this study: regionalization based on spatial distance similarity, regionalization based on physical similarity, and multiscale parameter regionalization (MPR). For spatial distance regionalization (DSR) similarity and physical similarity regionalization (PSR), five donors (gauged catchments) were used, because Bao et al. (2012) examined the optimal number of donors and achieved good results using five donors in China. Seven catchment descriptors—annual precipitation, temperature, dryness index (the ratio of the potential evapotranspiration to precipitation), normalized difference vegetation index, elevation, slope, and soil depth—were considered in the PSR method to calculate the rank of the of the donor catchment. For more information about the similarity regionalization, see appendix C.

The MPR technique was proposed by Samaniego et al. (2010), and this regionalization approach focuses on using transfer functions to relate geophysical features at the finest scale with model parameters at the finest scale, and then upscale them to the selected modeling spatial scale (normally much coarser) (Mizukami et al. 2017). Although a simultaneous regionalization method, MPR differs from the traditional standard regionalization methods that define catchment predictors at the modeling unit scale, because it accounts for the subgrid variability of catchment predictors (Samaniego et al. 2010). In this study, we coupled the ASMO algorithm to the MPR technique to conduct a simultaneous parameter estimation for both gauged and pseudo-ungauged catchments. As shown in Livneh et al. (2015), the use of finer soil texture properties in MPR among other physiographical predictors plays a significant role not only on the potential for ease the transferability of a hydrological model across scales and locations but also for reducing the predictive uncertainty of the estimated fluxes. On the contrary to those regionalization methods that use coarse or lumped predictor counterparts (e.g., DSR, or PSR), MPR delivers seamless parameters fields as shown in Mizukami et al. (2017) and Samaniego et al. (2010). As the workflow diagram for MPR–ASMO shows in Fig. ES5, this method involves three steps: 1) parameter transfer, 2) parameter upscaling, and 3) simultaneous parameter optimization. The first and second steps identified a form of the transfer function and the appropriate scaling operator for each model parameter, respectively. The model parameters and their transfer functions and upscaling operators are shown in Table 1. Eight model global parameters, including an infiltration parameter (B), three baseflow parameters (Ds, Dm, and Ws), the second soil layer drainage parameter (E_2) and soil depth parameters (D_1, D_2, and D_3), are involved in simultaneous parameter optimization with different combinations of sensitive parameters. Detailed parameter descriptions are given in the online supplement. The soil depth parameters were chosen in the model calibration process for their high sensitivity as indices for streamflow in most river basins across China (Table ES1; Gou et al. 2020). Previous empirical transfer functions with geophysical features for each model tunable parameter were used in this study (Table 1). Detailed information on all variables in Table 1 can been seen in Table 2.

Table 1. Regionalization transfer functions and upscaling operators used in VIC model.

Parameters (λ)	Transfer functions	β	Lower bound	Upper bound	Default values	Upscaling operators
B	$\dfrac{\ln\sigma_{ele} - \beta_1}{\ln\sigma_{ele} + 10\beta_2}$ [a]	β_1	−2	1	0	Arithmetic mean
		β_2	0.8	1.2	1	
Ds	$\beta_3 \times \left\{ A_0 + \displaystyle\int_{A_0}^{1} \left[1 - (1-A)^{1/B}\right]^{2\mu+3} dA \right\}$ [b]	β_3	0.3	1.3	1	Harmonic mean
Dm	$\dfrac{1}{10^{\beta_4}} K_s$ [b]	β_4	1.75	2.5	2	Harmonic mean
Ws	$\beta_5 \times \left(\dfrac{W_f}{W_m}\right)$ [b]	β_5	0.3	1.3	1	Harmonic mean
E_2	$\beta_6 \times (3 + 2/\lambda)$ [c]	β_6	0.3	1.3	1	Harmonic mean
D_1	$\beta_7 \times h_1$	β_7	0.3	1.3	1	Harmonic mean
D_2	$\beta_8 \times h_2$	β_8	0.3	1.3	1	Harmonic mean
D_3	$\beta_9 \times h_3$	β_9	0.3	1.3	1	Harmonic mean

[a] Source: Mizukami et al. (2017).
[b] Source: Bao et al. (2012).
[c] Source: https://vic.readthedocs.io/en/master/.

The third (key) step comprised parameter estimation by adjusting transfer function parameters (β) rather than adjusting the model global parameter (λ) themselves; each model iteration in an automated calibration algorithm goes through all the MPR processes. The objective function (O) of the MPR–ASMO method is defined for the aggregation of all catchments but is weighted toward the catchment in which the model performance is the worst (Hundecha and Bárdossy 2004):

$$O = \sum_{i=1}^{M} \varepsilon_i^2 + M \min_{i=1,\dots,M} \varepsilon_i^2, \tag{1}$$

where M is the number of calibration catchments, and the ε_i refers to the Nash–Sutcliffe model efficiency coefficient (NSE) of monthly streamflow in each calibration catchment.

To evaluate the performance of the different regionalization approaches, we designed five experiments (Table 3). In the MPR, PSR, and DSR experiments, the parameters in pseudo-ungauged catchments were obtained from the corresponding regionalization methods. Two reference experiments were also considered: one for which the optimal solution obtained in parameter optimization was used (Ref1) and another for which a single median parameter set was used across the corresponding major basins (Ref2). These two reference experiments can be interpreted as providing the upper (practically unreachable) limit and the lower limit of the regionalization schemes, respectively. The results obtained with the three regionalization schemes tested were expected to lie between the two extremes.

Results and discussion
Parameter optimization for the gauged catchments.
Based on the sensitivity analysis results of our previous work by Gou et al. (2020), a list of VIC tunable parameters for each large river basin (Table ES1) was identified that are important for streamflow simulation. Subsequently, the ASMO algorithm was used to optimize those parameters for each of the 200 gauged catchments;

Table 2. Variables list.

Variable symbol	Brief description	Unit	Sources
λ	Global parameter	—	Calculated by transfer functions and upscaling operators
$\beta_1 - \beta_9$	Transfer function parameters	—	Found by optimization algorithm
σ_{ele}	Standard deviation of elevation	m	Global Multi-resolution Terrain Elevation Data 2010 dataset (www.usgs.gov/land-resources/eros/coastal-changes-and-impacts/gmted2010)
W_f	Subgrid field capacity	mm	China Dataset of Soil Hydraulic Parameters (Dai et al. 2013)
W_m	Saturated soil moisture	mm	China Dataset of Soil Hydraulic Parameters (Dai et al. 2013)
A_0	Saturated fraction of the subgrid area, in which the average soil moisture is W_f	—	Calculated from the variable soil moisture capacity curve (Bao et al. 2012)
μ	Brooks–Corey equation parameter	—	Empirical parameter from Cosby et al. (1984)
K_s	Saturated hydraulic conductivity	mm day^{-1}	China Dataset of Soil Hydraulic Parameters by Dai et al. (2013)
γ	Soil pore size distribution parameter	—	Calculated by the sand and clay soil particle-size fraction
$h_1 - h_3$	Three soil layer thickness parameters	m	China Dataset of Soil Hydraulic Parameters by Dai et al. (2013)

Table 3. Model parameter regionalization experimental design in this study.

Experiments	Sources for parameters for pseudo-ungauged catchments
Ref1	Using the optimal solution obtained in parameter optimization
Ref2	Using a single median parameter set across the corresponding major basin
MPR	Estimated with the multiscale parameter regionalization technique
PSR	Estimated with the physical similarity regionalization technique
DSR	Estimated with the spatial distance similarity regionalization technique

the skill scores (NSE) of simulated monthly streamflow for those catchments are shown in Fig. 4. The NSE between the naturalized and simulated streamflow varied between 0.39 and 0.98 under the calibration period, and between 0.28 and 0.98 under the validation periods. The NSE showed high values for all catchments, with an average of 0.83 and 0.80 for calibration and validation modes, respectively. Moreover, most catchments show good performance for peak flow (July–September) except the northwestern catchments (catchments 181–200, Fig. ES6). The poor performance of peak flow in northwest drainage system perhaps owing to the model structure uncertainties over snow-dominated and arid
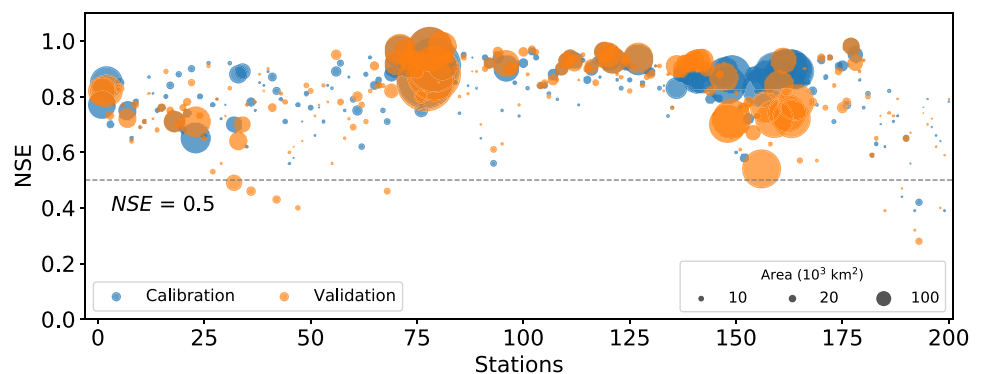

Fig. 4. Parameter optimization results (expressed as NSE) of VIC-simulated monthly streamflows for 200 gauged catchments during the calibration (blue circles) and validation periods (orange circles). The size of each circle indicates the drainage area of corresponding catchment (10³ km²).

regions (X. He et al. 2020). The physical parameterization scheme related to snowmelt are not well understood (Pan et al. 2003), and other key hydrological processes (e.g., glacier dynamics) are missing in the model. Therefore, peak flow information compiled in the CNRD v1.0 should be interpreted with caution over northwestern regions and specific model optimization may need to be considered before using the data. Figure 5 shows the spatial distribution
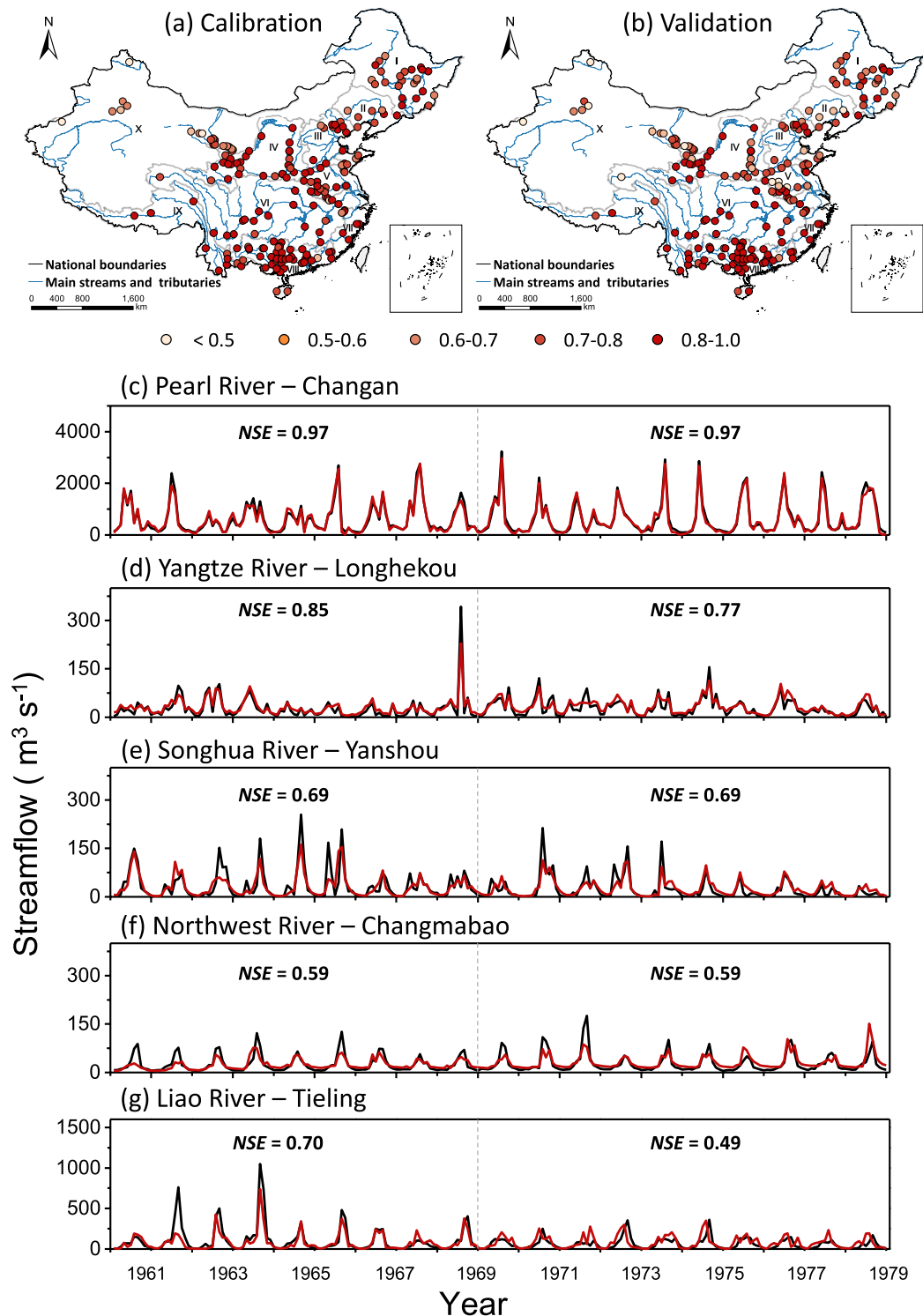

Fig. 5. Spatial distribution of the model performance of the calibrated catchments. (a),(b) The spatial distribution of optimization results (expressed as NSE) of VIC-simulated monthly streamflow for all 200 gauged stations during the calibration period (1961–69) and the validation period (1970–79). (c)–(g) Time series of monthly streamflow for the selected gauged catchments. The black and red lines indicate naturalized and the simulated streamflows.

of the model performance of the calibrated catchments (expressed as the NSE), and the comparison between the naturalized streamflow and modeled monthly streamflow for five selected catchments. In both time periods, the NSE values across China are generally high in most of the southern basins and low in the northern basins, particularly in the northwestern basins (Figs. 5a,b). This uneven spatial pattern of model performance can be attributed the inadequate precipitation gauge densities (Fig. ES2) and the arid and semiarid hydroclimatic regimes in the northern region. The precipitation data provide an important upper boundary condition of runoff simulation, and the accuracy of these data to a large extent determines the representation of surface hydrological processes, while also affecting parameter values and model performance (Xu et al. 2013). In addition, the mechanism of rainfall–runoff generation is more complex in semiarid and semihumid regions than that in humid regions because the infiltration excess runoff and saturation excess runoff always interact to varying degrees in those regions (Atkinson et al. 2002; Hu et al. 2005). Therefore, all models, simple or complex, produce more accurate results with more confidence in humid regions. And caution is needed when applying the model simulations in northwestern river drainages, for the number of gauge stations used in the model training process is not sufficient there. In the five selected catchments, the temporal evolution of river flow is in general well captured; both the timing and amount of simulated streamflow peaks and valleys closely match the naturalized streamflow (Figs. 5c–g). High flows were slightly underestimated for arid or semiarid basins in the Songhua, Northwest, and Liao River basins (Figs. 5e–g). Even if the model does not perform well in some arid and semiarid areas, the results overall show well-calibrated parameters for most gauged catchments.

Parameter regionalization performance for the pseudo-ungauged catchments. Five regionalization experiments (Table 3) were designed to evaluate the performance of the different regionalization schemes in transferring calibrated parameter sets from the gauged (donor) catchments to pseudo-ungauged catchments. Figure 6 summarizes the results by showing the distributions of model efficiencies given by the five regionalization scheme experiments over China as a whole. Unsurprisingly, the Ref1 and Ref2 experiments produced the best and worst (in terms of median NSE) regionalization results, respectively, with values of the median NSE equal to 0.85 (Ref1) and 0.68 (Ref2) in the calibration period and 0.82 (Ref1) and 0.67 (Ref2) in the validation period (Fig. 6a); the cumulative frequency distributions lines for the NSE values obtained from the Ref1 and Ref2 experiments are located on the far right and left of the clustered lines (Figs. 6b,c). The differences found between the three regionalization methods were discernible but not very large. MPR offered the best regionalization solution (median NSE = 0.76 for the calibration period and 0.72 for the validation period; Fig. 6a). PSR was the worst of three regionalization schemes (median NSE = 0.71 for the calibration period and 0.68 for the validation period), with results close to the Ref2 median parameter solution (Fig. 6). The DSR experiment produced similar median NSE values to the MPR experiment (0.75 and 0.71 for the calibration and validation periods, respectively; Fig. 6a); the cumulative frequency distributions of the DSR and MPR experiments are very close, although MPR performed better in the calibration period (Figs. 6b,c). At basin scale, MPR exhibited the leading performance of the three regionalization methods in humid areas, including the Yangtze River basin, Southeast River drainage system, and Pearl River basin (Fig. 7), which indicates that MPR produced more satisfactory results in humid regions than in arid regions. Overall, MPR generated more accurate parameter results than those obtained with the PSR and DSR methods for estimating the parameters of pseudo-ungauged catchments. The parameters generated by MPR are spatially continuous (i.e., seamless; see Samaniego et al. 2010) in relation to geophysical attributes of each river basin (Fig. ES7). Therefore, the model parameters for real-ungauged catchments were
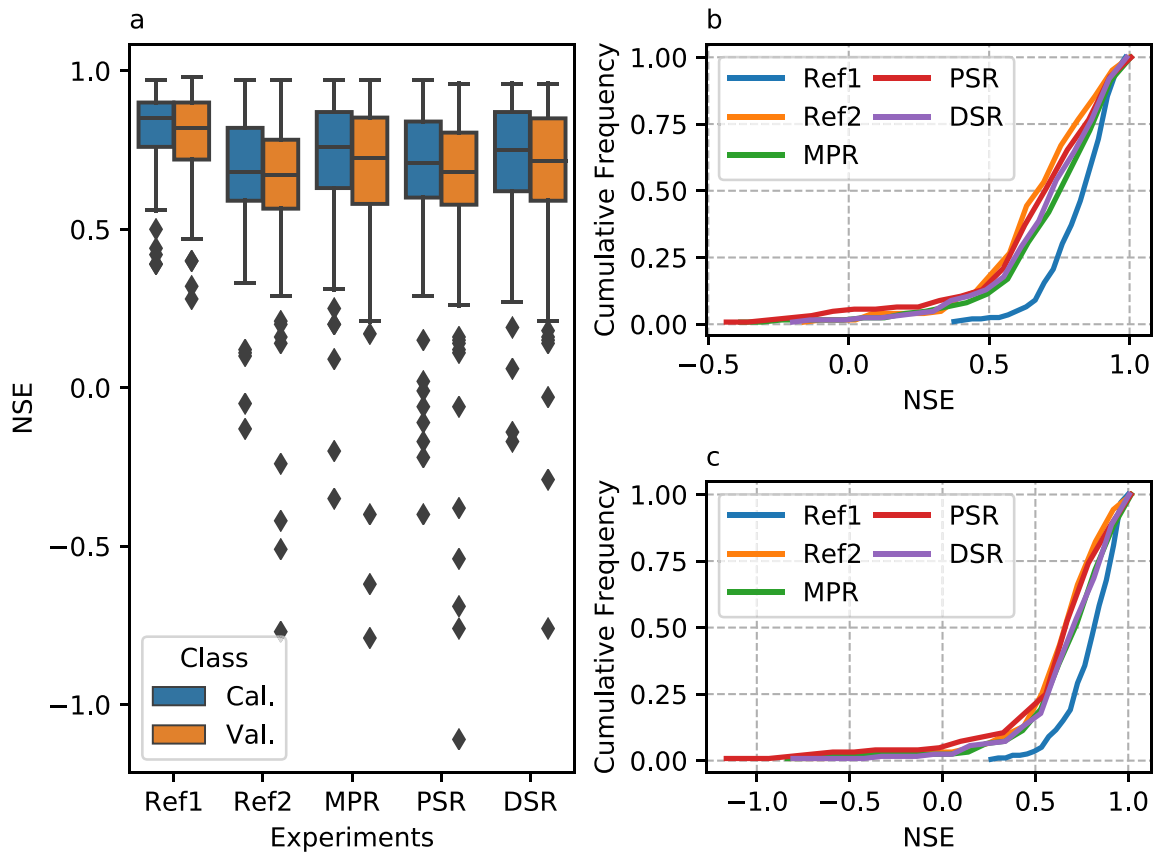
Fig. 6. Comparison of the performance of different parameter regionalization experiments across all pseudo-ungauged catchments. (a) Boxplot illustrating the skill-score distributions (expressed as NSE) for all five regionalization experiments for the calibration periods (blue boxes) and validation periods (orange boxes). (b),(c) Cumulative frequency distributions of the skill scores for all five regionalization experiments for the calibration period and the validation period, respectively, and the colors of the curves distinguish the different regionalization experiments.

obtained based on the MPR technique by means of simultaneously calibrating the gauged catchments within a river basin.

CNRD v1.0 product evaluation. Using the well trained VIC land surface model from the results of the parameter uncertainty analysis framework, we created the CNRD v1.0 dataset. CNRD v1.0 is a 0.25° daily mean and monthly mean natural runoff reconstruction product that spans the period 1 January 1961 to 31 December 2018, over mainland China. To investigate whether the CNRD v1.0 gridded runoff product was properly reconstructed, two global gridded runoff datasets, ISIMIP2a and GRUN (appendix B) were used as references for product evaluation. Figures 8a–c compare the spatial distributions of the long-term-mean (1971–2010) annual total runoff for ISIMIP, CNRD v1.0, and GRUN. The overall pattern of the reconstructed runoff record in CNRD v1.0 is similar to the runoff record independently derived from the multimodel ensemble mean of ISIMIP from GRUN (Figs. 8a–c). Runoff maps from CNRD v1.0 show more continuous transitions in runoff distribution compared to ISIMIP and GRUN. Possible reasons for the roughness of ISIMIP and GRUN runoff data relate to the coarse resolution of those two datasets (0.5°, compared to the 0.25° resolution of CNRD v1.0) and the multimodel-mean technique used to them. CNRD v1.0 performs better than the two global runoff datasets in representing the geographic distribution of China's water resources across complex terrain and climate regions. For example, CNRD v1.0 is better at representing the high-value runoff center produced by the monsoon climate over southeast China and the snowmelt runoff produced by the special
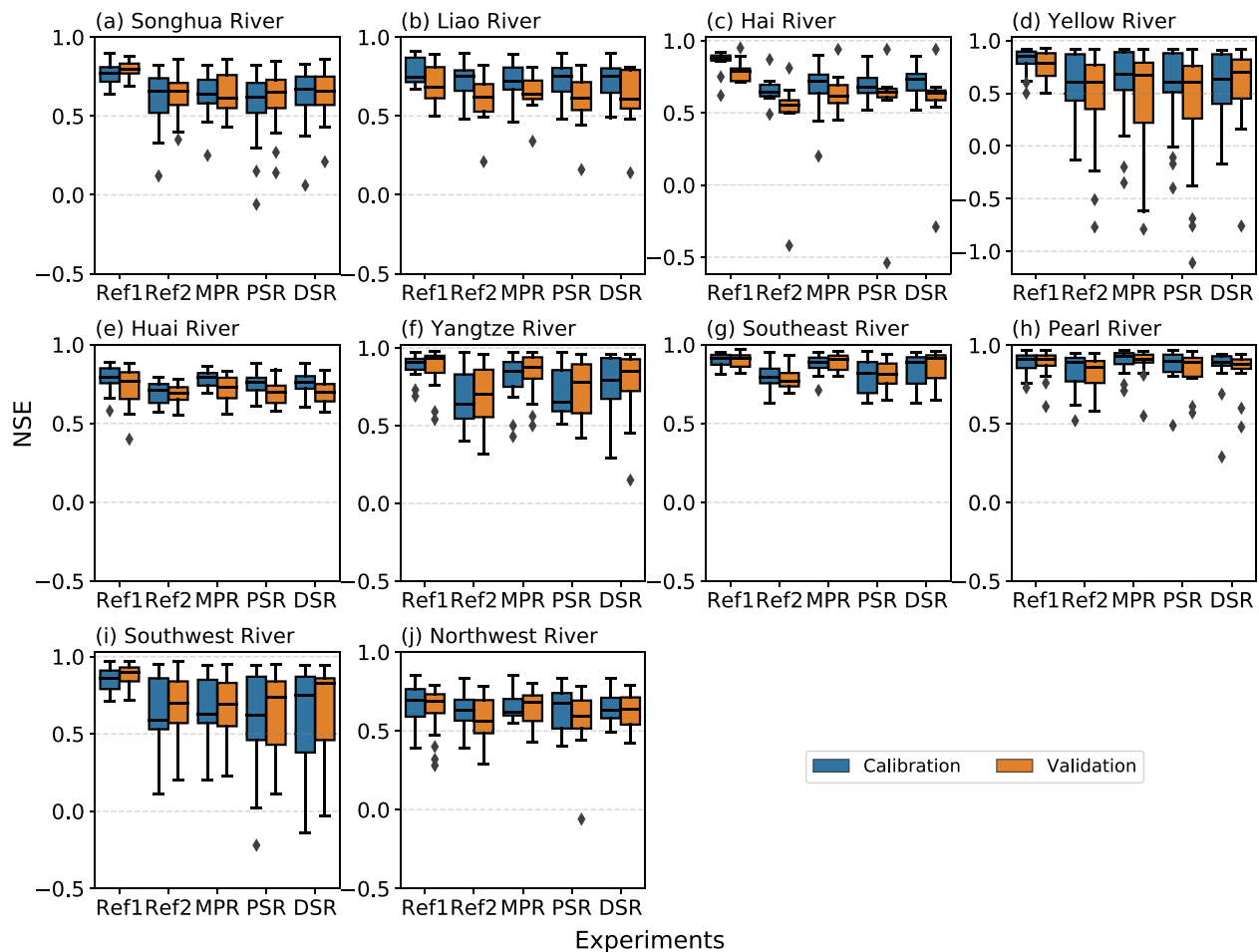
Fig. 7. Comparison of the performance (expressed as NSE) of different parameter regionalization experiments for monthly streamflow in pseudo-ungauged catchments across 10 river basins: (a) Songhua River basin, (b) Liao River basin, (c) Hai River basin, (d) Yellow River basin, (e) Huai River basin, (f) Yangtze River basin, (g) Southeast River drainage system, (h) Pearl River basin, (i) Southwest River drainage system, and (j) Northwest River drainage system.

geographic environment of northwest China (where two basins are sandwiched between three mountains); those features are not well captured by either the ISIMIP or the GRUN maps (Figs. 8a–c). This is mainly because runoff recorded by ISIMIP and GRUN is not well calibrated in China, since the streamflow gauge stations used in the China region during the data production process were sparse. Cell-to-cell comparisons between the CNRD v1.0 runoff map and the ISIMIP and GRUN runoff maps are shown in Figs. 8d and 8e, respectively. These comparisons show that CNRD v1.0 gives results in accordance with ISIMIP and GRUN, the respective levels of agreement (expressed by the coefficient of determination R^2) being 0.92 and 0.72.

CNRD v1.0 was also compared with ISIMIP and GRUN for annual and monthly temporal scales over China; mean annual and multiyear monthly mean runoff time series for the three datasets were calculated and are shown in Fig. 9. The results show an overall agreement between the reconstructed CNRD v1.0 and ISIMIP and GRUN runoff for both interannual and annual scales. The timings of dry and wet years in CNRD v1.0 closely match the two global runoff datasets, although CNRD v1.0 always generates a larger magnitude runoff value (Fig. 9a). The annual cycles of monthly runoff show similar performance to the interannual cycles, and the timing of high-flow and low-flow months is well captured by CNRD v1.0. Whereas the CNRD v1.0 values invariably have largest magnitudes throughout the spring and winter seasons; the summer and autumn season runoff in CNRD v1.0 agrees with the
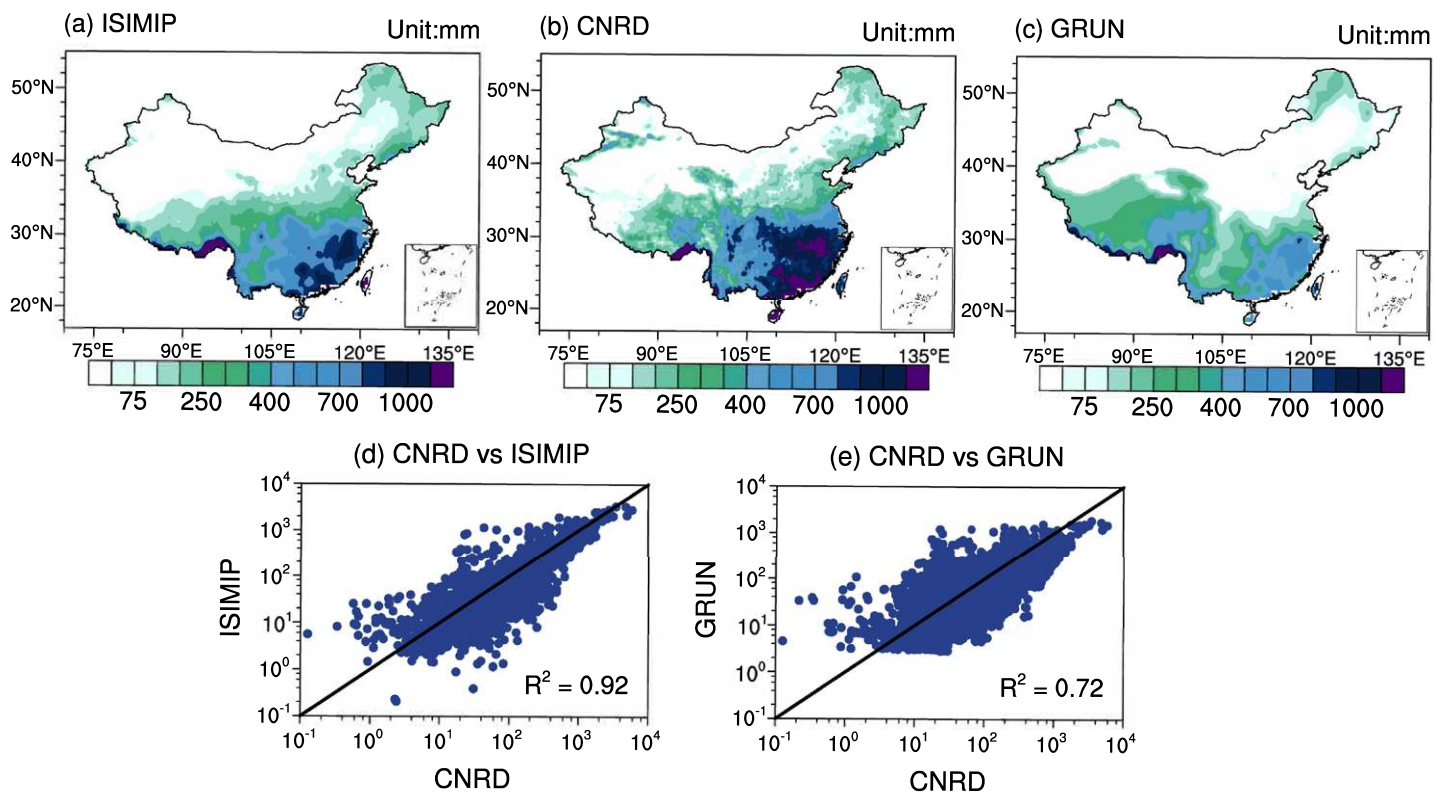
Fig. 8. CNRD v1.0 product evaluation. (top) Spatial distribution of the long-term mean annual (1961–2010) total runoff for (a) the ISIMIP multimodel ensemble mean, (b) the CNRD v1.0 dataset, and (c) the GRUN dataset. (bottom) Spatial correlation (d) between CNRD v1.0 and ISIMIP and (e) between CNRD v1.0 and GRUN.

higher range of the ISIMIP2a simulations. Overall, the temporal dynamics of runoff are well reproduced by CNRD v1.0 in comparison to the global runoff data.

Potential applications of CNRD v1.0. This section briefly discusses three potential applications of the CNRD v1.0 dataset in hydrological and climate studies, including water resources management, climate change assessment for terrestrial water availability, and cross-validation of satellite-observed runoff data. China has serious water scarcity problems due to the tremendous impact of topography gradients and a monsoon climate (Piao et al. 2010). The uneven distribution of water resources and growing water demand will cause water competition across China between irrigators, domestic water users, and the energy sector (Zhu et al. 2017). The CNRD v1.0 dataset has been developed to provide spatiotemporally continuous natural runoff estimates at the national scale, which has the potential to support water resources management and allocation. Moreover, there is also a need for tools to identify possible physical mechanisms and processes of the terrestrial water cycle and how it is responding to a warming climate (Huntington 2006; Sun et al. 2018, 2019). As a reliable spatiotemporally continuous runoff dataset, CNRD v1.0 could potentially be used for detecting large-scale climate features such as El Niño–Southern Oscillation impacts on hydrologic processes. Thus, it could provide additional decision-making support for water managers who are developing plans for their local communities to adapt to climate change. In addition, recent advances in satellite-based optical remote sensors (RS) offer promising alternatives for monitoring global river discharge from space (Huang et al. 2018; Lin et al. 2019). However, using RS observations for surface runoff estimation is subject to large uncertainties, such as weather and vegetation cover impacts on optical sensors. Further corrections to these RS observations might be essential to make them accurate enough for hydrometeorology applications. The CNRD v1.0 dataset we have introduced, with its

multiple levels of quality control in the data production process, could potentially be used to cross validate remotely sensed data or in other scientific applications requiring spatiotemporally continuous discharge estimates.

Conclusions and outlook

We present here CNRD v1.0, a gridded runoff dataset reconstructed to span 1961–2018 at daily and monthly temporal resolution and 0.25° spatial resolution. As a long-term spatiotemporally continuous natural runoff record, CNRD v1.0 is the first free public dataset constructed using a comprehensive model parameter uncertainty analysis framework across China. CNRD v1.0 is generated using a land surface model, which was used to enrich the relatively shorter and discontinuous spatiotemporal distribution of gauged streamflow data, thus filling in gaps or constructing time series of comparable length. Sensitivity analysis methods were used to identify the important parameters for streamflow simulation and an ASMO algorithm was used to tune those important parameters in the VIC model based on data from 200 gauged catchments. Another important quality control used in producing this dataset is the use of the MPR technique to estimate parameters in the ungauged catchments based on physical characteristics of corresponding catchments.
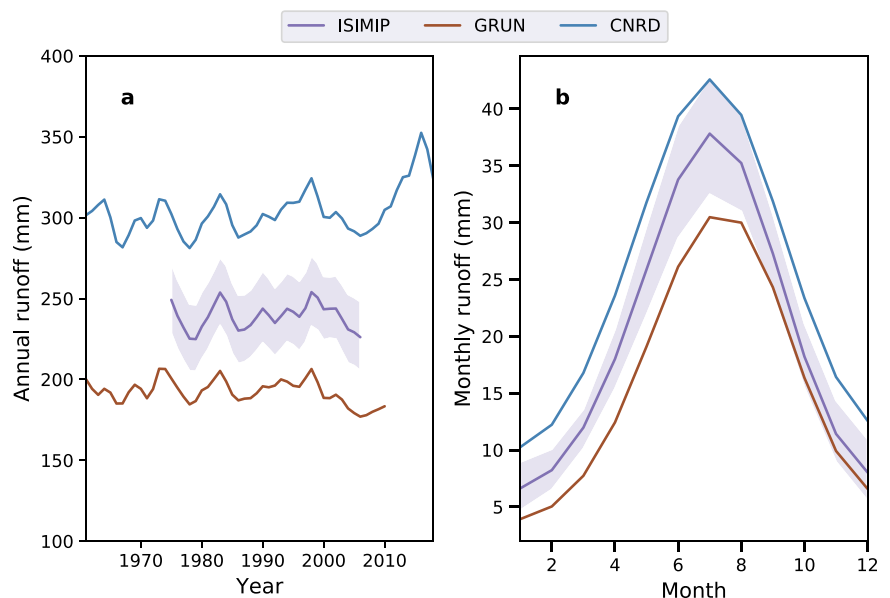

Fig. 9. Smoothed CNRD v1.0 data over China compared with the ISIMIP multimodel ensemble mean and the GRUN dataset for (a) annual runoff and (b) annual cycles of monthly mean runoff. Shading denotes the 95% confidence interval of the ISIMIP global hydrological models. The colored lines indicate the different runoff statistics for different datasets during the period 1961–2018. Note that the statistical period used in (b) is limited to 1961–2010 to ensure consistency of the datasets for comparison. We applied Gaussian smoothing to each data record (sigma = 1).

In the model training stage, the results overall show well-calibrated parameters for most gauged catchments except in arid and semiarid areas; NSE values were high for all catchments, averaging 0.83 under calibration mode and 0.80 under validation mode. For the pseudo-ungauged catchment, MPR offered the best regionalization solution to estimate the model parameters, with values of the median NSE equaling 0.76 and 0.72 for the calibration and validation periods, respectively. In the data production stage, we evaluated CNRD v1.0 against two other runoff datasets, ISIMIP and GRUN, as references at a monthly time scale during the period 1971–2010. CNRD v1.0 performs better than the two global runoff datasets in representing China's water resources distribution under complex terrain and climate conditions because of the sparseness of streamflow gauge stations across China that were used in producing these global datasets. Cell-to-cell comparisons between the CNRD v1.0 runoff map and the ISIMIP and GRUN runoff maps show overall agreement ($R^2 = 0.92$ and 0.72, respectively). In addition, the temporal dynamics of runoff are well reproduced by CNRD v1.0, because the timing of dry versus wet years and high-flow versus low-flow months of CNRD v1.0 closely match the two global runoff datasets.

We have demonstrated the potential of CNRD v1.0 as a new dataset for hydroclimate studies. CNRD v1.0 is publicly available at https://doi.org/10.6084/m9.figshare.13185410. In practice, several considerations must be taken into account before using this dataset. The CNRD v1.0

runoff estimates are constructed from two climate forcing datasets that partially overlap in time, which may introduce uncertainty between the pre- and post-2014 parts of the runoff data, although these two forcing datasets have passed a consistency check in most regions of China (Fig. ES3). CNRD v1.0 is also limited by insufficient training stations in the Northwest River drainage system, which is inevitable due to the inadequate gauge network density in this area. Therefore, the CNRD v1.0 should be applied with caution for the northwestern regions if water amount instead of relative changes is to be assessed at those regions. The current version of the CNRD product only provides a gridded runoff record; but work is in progress to complete additional flow products for the catchment (gauges) scale and the river (reaches) scale. Next steps within the development of this runoff dataset should focus on multimodel ensemble that reduce the uncertainty of runoff simulation and improve the simulation of arid/semiarid regions runoff across China. A previous study observed the overestimation of active evapotranspiration (AET) values within the VIC model, resulting in underestimation of the runoff simulation (Rakovec et al. 2019). It is intended that multi-objective optimization involving other water fluxes (e.g., AET and soil moisture) and multiple models with different parameterization schemes will be considered for our future work. For runoff estimation in ungauged catchments, it is crucially important to improve the understanding of parameter transfer functions by considering more geophysical features, such as soil porosity, wilting point, and land cover.

Appendix A: Naturalized-gauge streamflow reconstruction method

Naturalized-gauge streamflow reconstruction refers to removing the anthropogenic influences affecting observed streamflow data to approximate the natural condition. Naturalized streamflow is a widely used data product for understanding natural hydrological processes (Tu et al. 2015) and the impact of human activities (Yuan et al. 2017) and for calibrating hydrological models (Gou et al. 2020; Nijssen et al. 1997). The naturalized streamflow was calculated as the sum of the observed streamflow and the direct water abstraction based on abundant data from different sectors (Gou et al. 2020; Yuan et al. 2017):

$$W_n = W_m + W_{ir} + W_{iw} + W_{dw} \pm W_d \pm W_f \pm W_r, \tag{A1}$$

where W_n is the naturalized streamflow; W_m is the measured streamflow at gauge stations; W_{ir} is the volume of water diverted from the river for irrigation, i.e., irrigated surface water transpired by crops, evaporated from bare ground and open water surfaces, absorbed into soil (through infiltration, percolation, and recharge to shallow groundwater), or leaked from river or channel beds to groundwater during transportation; W_{iw} is the volume of water extracted for industrial consumption; W_{dw} is the volume of water removed for domestic consumption; W_d is the volume of water diverted from the river to basins; W_f is the volume diverted for flood protection; and W_r is the consumption of water stored in reservoirs.

Appendix B: ISIMIP2a and GRUN

Two global gridded runoff datasets, the runoff simulations from the second phase of the

Inter-Sectoral Impact Model Intercomparison Project (ISIMIP2a, 0.5° × 0.5°), and an observation-based global gridded Global Runoff Reconstruction dataset (GRUN, 0.5° × 0.5°) by Ghiggi et al. (2019), were used in this study as references to evaluate the accuracy of CNRD v1.0. ISIMIP2a is a community-driven climate-impact modeling initiative that provides a global hydrological model simulation protocol and defines a set of common simulation scenarios (https://esg.pik-potsdam .de/search/isimip/). In this study, we use an ensemble of 18 runoff simulations participating in ISIMIP2a from six global hydrological models (GHM) under three global meteorological forcing products to represent the general condition of global terrestrial water resources. The six GHMs are the Distributed Biosphere–Hydrological (DBH) model, H08, the Lund–Potsdam–Jena managed Land (LPJmL) model, the Minimal Advanced Treatments of Surface Interaction and Runoff (MATSIRO) model, the PCRaster Global Water Balance (PCR-GLOBWB) model, and the Water-Global Analysis and Prognosis (WaterGAP2) model. Each GHM was forced by three selected global meteorological forcing products, namely, the PGMFD v.2, GSWP3, and WFDEI datasets under the "nosoc" socio-economic simulation scenario (without anthropogenic influences) for the period from 1971 to 2010. GRUN is an observation-based global gridded runoff dataset covering the period from 1902 to 2014 (https://doi.org/10.6084/m9.figshare.9228176). The gridded monthly runoff time series in GRUN is produced by a trained machine-learning algorithm (forced by GSWP3) from in situ streamflow observations. The ensemble means of 50 reconstructions obtained from different subsets of training data were used in this study.

Appendix C: Similarity regionalization method

In similarity-based regionalization, the parameters in ungauged catchments are directly transferred from geographic neighbors (spatial and/or physical), the rationale for this being that the climatic, geological, or topographic characteristics are likely to be relatively homogeneous within a region, so neighbor flow regimes should behave similarly (Bao et al. 2012; Oudin et al. 2008). For the spatial similarity regionalization, the Euclidean distances between pairs of catchments were used to define the rank of the of the donor catchment:

$$S_d = \sqrt{\left(X_t - X_d\right)^2 + \left(Y_t - Y_d\right)^2}, \tag{A2}$$

where (X_t, Y_t) and (X_d, Y_d) are the geographical coordinates of the centroids of the receptor and donor catchments, respectively.

For the physical similarity regionalization, we selected seven catchment descriptors—annual precipitation, temperature, dryness index, normalized difference vegetation index (NDVI), elevation, slope, and soil depth—to represent the physical condition of the donor catchment. The annual precipitation and the multiyear mean temperature were computed using the 0.25° gridded daily precipitation and mean temperature from the China Meteorological Administration, constructed from ~2,400 stations during the period from 1961 to 1979. The dryness index is the ratio of the potential evapotranspiration to precipitation; the potential evapotranspiration was calculated by the method of Hargreaves and Samani (1985). NDVI values were extracted from the Advanced Very High Resolution Radiometer (AVHRR)-based Global Inventory Modeling and Mapping Studies (GIMMS) NDVI dataset during the period from 1961 to 1979. Elevation and slope information were extracted based on the 1-km digital elevation model dataset from the Cold and Arid Regions Sciences Data Center at Lanzhou (http://westdc.westgis.ac.cn). Soil depths were obtained from the 30-arc-s soil database of China, which was produced by Dai et al. (2013). In this study, each descriptor had the same weight in the proximity computation, and the mean ranks were computed using the ranks of the donor catchment for each descriptor. The mean rank was then used to rank the donor catchments by decreasing proximity in the physical similarity regionalization.

References

Atkinson, S. E., R. A. Woods, and M. Sivapalan, 2002: Climate and landscape controls on water balance model complexity over changing timescales. *Water Resour. Res.*, **38**, 1314, https://doi.org/10.1029/2002WR001487.

Bao, Z. X., and Coauthors, 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, https://doi.org/10.1016/j.jhydrol.2012.07.048.

Beck, H. E., A. De Roo, and A. I. J. M. Van Dijk, 2015: Global maps of streamflow characteristics based on observations from several thousand catchments. *J. Hydrometeor.*, **16**, 1478–1501, https://doi.org/10.1175/JHM-D-14-0155.1.

——, A. I. J. M. Van Dijk, A. De Roo, D. G. Miralles, T. R. Mcvicar, J. Schellekens, and L. A. Bruijnzeel, 2016: Global-scale regionalization of hydrologic model parameters. *Water Resour. Res.*, **52**, 3599–3622, https://doi.org/10.1002/2015WR018247.

Bennett, K. E., J. R. U. Blanco, A. L. Atchley, N. M. Urban, A. Jonko, and R. S. Middleton, 2018: Global sensitivity of simulated water balance indicators under future climate change in the Colorado basin. *Water Resour. Res.*, **54**, 132–149, https://doi.org/10.1002/2017WR020471.

Berghuijs, W. R., J. R. Larsen, T. H. Van Emmerik, and R. A. Woods, 2017: A global assessment of runoff sensitivity to changes in precipitation, potential evaporation, and other factors. *Water Resour. Res.*, **53**, 8475–8486, https://doi.org/10.1002/2017WR021593.

Bohn, T. J., B. Livneh, J. W. Oyler, S. W. Running, B. Nijssen, and D. P. Lettenmaier, 2013: Global evaluation of MTCLIM and related algorithms for forcing of ecological and hydrological models. *Agric. Meteor.*, **176**, 38–49, https://doi.org/10.1016/j.agrformet.2013.03.003.

Burn, D. H., and D. B. Boorman, 1993: Estimation of hydrological parameters at ungauged catchments. *J. Hydrol.*, **143**, 429–454, https://doi.org/10.1016/0022-1694(93)90203-L.

Cosby, B. J., G. M. Hornberger, R. B. Clapp, and T. R. Ginn, 1984: A statistical exploration of the relationships of soil moisture characteristics to the physical properties of soils. *Water Resour. Res.*, **20**, 682–690, https://doi.org/10.1029/WR020i006p00682.

Cuntz, M., and Coauthors, 2015: Computationally inexpensive identification of noninformative model parameters by sequential screening. *Water Resour. Res.*, **51**, 6417–6441, https://doi.org/10.1002/2015WR016907.

Dai, A. G., 2017: Dai and Trenberth global river flow and continental discharge dataset. Research Data Archive at NCAR CISL, accessed 26 April 2020, https://doi.org/10.5065/D6V69H1T.

Dai, Y. J., W. Shangguan, Q. Y. Duan, B. Y. Liu, S. H. Fu, and G. Y. Niu, 2013: Development of a China dataset of soil hydraulic parameters using pedotransfer functions for land surface modeling. *J. Hydrometeor.*, **14**, 869–887, https://doi.org/10.1175/JHM-D-12-0149.1.

Do, H. X., L. Gudmundsson, M. Leonard, and S. Westra, 2018: The Global Streamflow Indices and Metadata Archive (GSIM) – Part 1: The production of a daily streamflow archive and metadata. *Earth Syst. Sci. Data*, **10**, 765–785, https://doi.org/10.5194/essd-10-765-2018.

Duan, Q. Y., S. Sorooshian, and V. Gupta, 1992: Effective and efficient global optimization for conceptual rainfall-runoff models. *Water Resour. Res.*, **28**, 1015–1031, https://doi.org/10.1029/91WR02985.

Ficchì, A., and L. Stephens, 2019: Climate variability alters flood timing across Africa. *Geophys. Res. Lett.*, **46**, 8809–8819, https://doi.org/10.1029/2019GL081988.

Gao, Y. B., C. Merz, G. Lischeid, and M. Schneider, 2018: A review on missing hydrological data processing. *Environ. Earth Sci.*, **77**, 47, https://doi.org/10.1007/s12665-018-7228-6.

Gao, J., T. Yao, V. Masson-Delmotte, H. C. Steen-Larsen, and W. Wang, 2019: Collapsing glaciers threaten Asia's water supplies. *Nature*, **565**, 19–21, https://doi.org/10.1038/d41586-018-07838-4.

Ghiggi, G., V. Humphrey, S. I. Seneviratne, and L. Gudmundsson, 2019: GRUN: An observation-based global gridded runoff dataset from 1902 to 2014. *Earth Syst. Sci. Data*, **11**, 1655–1674, https://doi.org/10.5194/essd-11-1655-2019.

Gou, J. J., C. Y. Miao, Q. Y. Duan, Q. H. Tang, Z. H. Di, W. H. Liao, J. W. Wu, and R. Zhou, 2020: Sensitivity analysis-based automatic parameter calibration of the variable infiltration capacity (VIC) model for streamflow simulations over China. *Water Resour. Res.*, **56**, e2019WR025968, https://doi.org/10.1029/2019WR025968.

Gudmundsson, L., S. I. Seneviratne, and X. B. Zhang, 2017: Anthropogenic climate change detected in European renewable freshwater resources. *Nat. Climate Change*, **7**, 813–816, https://doi.org/10.1038/nclimate3416.

Gupta, H. V., S. Sorooshian, and O. Yapo Patrice, 1999: Status of automatic calibration for hydrologic models: Comparison with multilevel expert calibration. *J. Hydrol. Eng.*, **4**, 135–143, https://doi.org/10.1061/(ASCE)1084-0699(1999)4:2(135).

Hargreaves, G. H., and Z. A. Samani, 1985: Reference crop evapotranspiration from temperature. *Appl. Eng. Agric.*, **1**, 96–99, https://doi.org/10.13031/2013.26773.

He, J., K. Yang, W. J. Tang, H. Lu, J. Qin, Y. Y. Chen, and X. Li, 2020: The first high-resolution meteorological forcing dataset for land process studies over China. *Sci. Data*, **7**, 25, https://doi.org/10.1038/s41597-020-0369-y.

He, X. G., M. Pan, Z. W. Wei, E. F. Wood, and J. Sheffield, 2020: A global drought and flood catalogue from 1950 to 2016. *Bull. Amer. Meteor. Soc.*, **101**, E508–E535, https://doi.org/10.1175/BAMS-D-18-0269.1.

Hu, C. H., S. L. Guo, L. H. Xiong, and D. Z. Peng, 2005: A modified Xinanjiang model and its application in northern China. *Nord. Hydrol.*, **36**, 175–192, https://doi.org/10.2166/nh.2005.0013.

Hu, Y., J. P. Moiwo, Y. Yang, S. Han, and Y. Yang, 2010: Agricultural water-saving and sustainable groundwater management in Shijiazhuang Irrigation District, North China Plain. *J. Hydrol.*, **393**, 219–232, https://doi.org/10.1016/j.jhydrol.2010.08.017.

Huang, C., Y. Chen, S. Zhang, and J. Wu, 2018: Detecting, extracting, and monitoring surface water from space using optical sensors: A review. *Rev. Geophys.*, **56**, 333–360, https://doi.org/10.1029/2018RG000598.

Hundecha, Y., and A. Bárdossy, 2004: 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**, 281–295, https://doi.org/10.1016/j.jhydrol.2004.01.002.

Huntington, T. G., 2006: Evidence for intensification of the global water cycle: Review and synthesis. *J. Hydrol.*, **319**, 83–95, https://doi.org/10.1016/j.jhydrol.2005.07.003.

Liang, X., D. P. Lettenmaier, E. F. Wood, and S. J. Burges, 1994: A simple hydrologically based model of land surface water and energy fluxes for general circulation models. *J. Geophys. Res.*, **99**, 14 415–14 428, https://doi.org/10.1029/94JD00483.

Lin, P. R., and Coauthors, 2019: Global reconstruction of naturalized river flows at 2.94 million reaches. *Water Resour. Res.*, **55**, 6499–6516, https://doi.org/10.1029/2019WR025287.

Livneh, B., R. Kumar, and L. Samaniego, 2015: Influence of soil textural properties on hydrologic fluxes in the Mississippi River basin. *Hydrol. Processes*, **29**, 4638–4655, https://doi.org/10.1002/hyp.10601.

Magette, W., V. Shanholtz, and J. Carr, 1976: Estimating selected parameters for the Kentucky watershed model from watershed characteristics. *Water Resour. Res.*, **12**, 472–476, https://doi.org/10.1029/WR012i003p00472.

Markonis, Y., Y. Moustakis, C. Nasika, P. Sychova, P. Dimitriadis, M. Hanel, P. Máca, and S. Papalexiou, 2018: Global estimation of long-term persistence in annual river runoff. *Adv. Water Resour.*, **113**, 1–12, https://doi.org/10.1016/j.advwatres.2018.01.003.

Meier, J., F. Zabel, and W. Mauser, 2018: A global approach to estimate irrigated areas – A comparison between different data and statistics. *Hydrol. Earth Syst. Sci.*, **22**, 1119–1133, https://doi.org/10.5194/hess-22-1119-2018.

Miao, C. Y., Q. H. Sun, A. G. L. Borthwick, and Q. Y. Duan, 2016: Linkage between hourly precipitation events and atmospheric temperature changes over China during the warm season. *Sci. Rep.*, **6**, 22543, https://doi.org/10.1038/srep22543.

——, Q. Y. Duan, Q. H. Sun, X. H. Lei, and H. Li, 2019: Non-uniform changes in different categories of precipitation intensity across China and the associated large-scale circulations. *Environ. Res. Lett.*, **14**, 025004, https://doi.org/10.1088/1748-9326/aaf306.

Mishra, A. K., and P. Coulibaly, 2009: Developments in hydrometric network design: A review. *Rev. Geophys.*, **47**, RG2001, https://doi.org/10.1029/2007RG000243.

Mitchell, B., and D. Shrubsole, 1994: Canadian water management: Visions for sustainability. Canadian Water Resources Association, 76 pp.

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

Nijssen, B., D. P. Lettenmaier, X. Liang, S. W. Wetzel, and E. F. Wood, 1997: Streamflow simulation for continental-scale river basins. *Water Resour. Res.*, **33**, 711–724, https://doi.org/10.1029/96WR03517.

Oki, T., and S. Kanae, 2006: Global hydrological cycles and world water resources. *Science*, **313**, 1068–1072, https://doi.org/10.1126/science.1128845.

Oudin, L., V. Andréassian, C. Perrin, C. Michel, and N. Le Moine, 2008: Spatial proximity, physical similarity, regression and ungaged catchments: A comparison of regionalization approaches based on 913 French catchments. *Water Resour. Res.*, **44**, W03413, https://doi.org/10.1029/2007WR006240.

Pan, M., and Coauthors, 2003: Snow process modeling in the North American Land Data Assimilation System (NLDAS): 2. Evaluation of model simulated snow water equivalent. *J. Geophys. Res.*, **108**, 8850, https://doi.org/10.1029/2003JD003994.

Papalexiou, S. M., and A. Montanari, 2019: Global and regional increase of precipitation extremes under global warming. *Water Resour. Res.*, **55**, 4901–4914, https://doi.org/10.1029/2018WR024067.

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

Piao, S. L., and Coauthors, 2010: The impacts of climate change on water resources and agriculture in China. *Nature*, **467**, 43–51, https://doi.org/10.1038/nature09364.

Rakovec, O., N. Mizukami, R. Kumar, A. J. Newman, S. Thober, A. W. Wood, M. P. Clark, and L. Samaniego, 2019: Diagnostic evaluation of large-domain hydrologic models calibrated across the contiguous United States. *J. Geophys. Res. Atmos.*, **124**, 13 991–14 007, https://doi.org/10.1029/2019JD030767.

Samaniego, L., and Coauthors, 2018: Anthropogenic warming exacerbates European soil moisture droughts. *Nat. Climate Change*, **8**, 421–426, https://doi.org/10.1038/s41558-018-0138-5.

——, R. Kumar, and S. Attinger, 2010: Multiscale parameter regionalization of a grid-based hydrologic model at the mesoscale. *Water Resour. Res.*, **46**, W05523, https://doi.org/10.1029/2008WR007327.

Sheffield, J., and Coauthors, 2014: A drought monitoring and forecasting system for sub-Sahara African water resources and food security. *Bull. Amer. Meteor. Soc.*, **95**, 861–882, https://doi.org/10.1175/BAMS-D-12-00124.1.

Sun, Q. H., C. Y. Miao, Q. Y. Duan, H. Ashouri, S. Sorooshian, and K. L. Hsu, 2018: A review of global precipitation data sets: Data sources, estimation, and intercomparisons. *Rev. Geophys.*, **56**, 79–107, https://doi.org/10.1002/2017RG000574.

——, ——, M. Hanel, A. G. L. Borthwick, Q. Y. Duan, D. Y. Ji, and H. Li, 2019: Global heat stress on health, wildfires, and agricultural crops under different levels of climate warming. *Environ. Int.*, **128**, 125–136, https://doi.org/10.1016/j.envint.2019.04.025.

Tang, Y., Q. H. Tang, Z. G. Wang, F. H. S. Chiew, X. J. Zhang, and H. Xiao, 2019: Different precipitation elasticity of runoff for precipitation increase and decrease at watershed scale. *J. Geophys. Res. Atmos.*, **124**, 11 932–11 943, https://doi.org/10.1029/2018JD030129.

Tencaliec, P., A.-C. Favre, C. Prieur, and T. Mathevet, 2015: Reconstruction of missing daily streamflow data using dynamic regression models. *Water Resour. Res.*, **51**, 9447–9463, https://doi.org/10.1002/2015WR017399.

Tetzlaff, D., S. K. Carey, J. P. Mcnamara, H. Laudon, and C. Soulsby, 2017: The essential value of long-term experimental data for hydrology and water management. *Water Resour. Res.*, **53**, 2598–2604, https://doi.org/10.1002/2017WR020838.

Tu, X. J., V. P. Singh, X. H. Chen, L. Chen, Q. Zhang, and Y. Zhao, 2015: Intra-annual distribution of streamflow and individual impacts of climate change and human activities in the Dongjiang River Basin, China. *Water Resour. Manage.*, **29**, 2677–2695, https://doi.org/10.1007/s11269-015-0963-5.

Wang, C., Q. Y. Duan, W. Gong, A. Z. Ye, Z. H. Di, and C. Y. Miao, 2014: An evaluation of adaptive surrogate modeling based optimization with two benchmark problems. *Environ. Modell. Software*, **60**, 167–179, https://doi.org/10.1016/j.envsoft.2014.05.026.

Xu, H. L., C.-Y. Xu, H. Chen, Z. X. Zhang, and L. Li, 2013: Assessing the influence of rain gauge density and distribution on hydrological model performance in a humid region of China. *J. Hydrol.*, **505**, 1–12, https://doi.org/10.1016/j.jhydrol.2013.09.004.

Yuan, X., M. Zhang, L. Y. Wang, and T. Zhou, 2017: Understanding and seasonal forecasting of hydrological drought in the Anthropocene. *Hydrol. Earth Syst. Sci.*, **21**, 5477–5492, https://doi.org/10.5194/hess-21-5477-2017.

Zhang, X. J., Q. H. Tang, M. Pan, and Y. Tang, 2014: A long-term land surface hydrologic fluxes and states dataset for China. *J. Hydrometeor.*, **15**, 2067–2084, https://doi.org/10.1175/JHM-D-13-0170.1.

Zhu, Y. N., Y. Zhao, H. H. Li, L. Z. Wang, L. Li, and S. Jiang, 2017: Quantitative analysis of the water-energy-climate nexus in Shanxi Province, China. *Energy Procedia*, **142**, 2341–2347, https://doi.org/10.1016/j.egypro.2017.12.164.