



# A hydrological emulator for global applications – HE v1.0.0

Yaling Liu<sup>1,2</sup>, Mohamad Hejazi<sup>1</sup>, Hongyi Li<sup>3</sup>, Xuesong Zhang<sup>1</sup>, and Guoyong Leng<sup>1</sup>

<sup>1</sup>Joint Global Change Research Institute, Pacific Northwest National Laboratory, 5825 University Research Court, College Park, MD 20740, USA

<sup>2</sup>Department of Earth and Environmental Engineering, Columbia University, New York, NY 10027, USA

<sup>3</sup>Department of Land Resources and Environmental Sciences, Montana State University, Bozeman, MT 59717, USA

**Correspondence:** Yaling Liu (cauliuyaling@gmail.com)

Received: 9 May 2017 – Discussion started: 16 June 2017

Revised: 31 January 2018 – Accepted: 19 February 2018 – Published: 23 March 2018

**Abstract.** While global hydrological models (GHMs) are very useful in exploring water resources and interactions between the Earth and human systems, their use often requires numerous model inputs, complex model calibration, and high computation costs. To overcome these challenges, we construct an efficient open-source and ready-to-use hydrological emulator (HE) that can mimic complex GHMs at a range of spatial scales (e.g., basin, region, globe). More specifically, we construct both a lumped and a distributed scheme of the HE based on the monthly *abcd* model to explore the trade-off between computational cost and model fidelity. Model predictability and computational efficiency are evaluated in simulating global runoff from 1971 to 2010 with both the lumped and distributed schemes. The results are compared against the runoff product from the widely used Variable Infiltration Capacity (VIC) model. Our evaluation indicates that the lumped and distributed schemes present comparable results regarding annual total quantity, spatial pattern, and temporal variation of the major water fluxes (e.g., total runoff, evapotranspiration) across the global 235 basins (e.g., correlation coefficient  $r$  between the annual total runoff from either of these two schemes and the VIC is  $> 0.96$ ), except for several cold (e.g., Arctic, interior Tibet), dry (e.g., North Africa) and mountainous (e.g., Argentina) regions. Compared against the monthly total runoff product from the VIC (aggregated from daily runoff), the global mean Kling–Gupta efficiencies are 0.75 and 0.79 for the lumped and distributed schemes, respectively, with the distributed scheme better capturing spatial heterogeneity. Notably, the computation efficiency of the lumped scheme is 2 orders of magnitude higher than the distributed one and 7 orders more efficient than the VIC model. A case study of uncertainty analysis

for the world's 16 basins with top annual streamflow is conducted using 100 000 model simulations, and it demonstrates the lumped scheme's extraordinary advantage in computational efficiency. Our results suggest that the revised lumped *abcd* model can serve as an efficient and reasonable HE for complex GHMs and is suitable for broad practical use, and the distributed scheme is also an efficient alternative if spatial heterogeneity is of more interest.

## 1 Introduction

A global hydrological model (GHM) is an effective tool to understand how water moves between soil, plants, and the atmosphere. In terms of spatial discretization, hydrological models can be classified into (1) lumped models treating one basin as a homogeneous whole and disregarding spatial variations, such as the Sacramento Soil Moisture Accounting Model (Burnash et al., 1973), and (2) distributed models, where the entire basin is divided into small spatial units (e.g., square cells or triangulated irregular network) to capture spatial variability, such as the PCRaster Global Water Balance (Van Beek and Bierkens, 2009) and the WASMOD-M (Water And Snow balance MODELing system for macro-scale; Widén-Nilsson et al., 2007). For simplicity, models with a division of one basin into separate areas or subbasins are also categorized as distributed ones here. The corresponding predictability and computational efficiency of GHMs may vary from model to model, due to differences in complexity and structure. Recent years have seen rapid progress in GHMs. They are widely used in assessing the impacts of climate change and land surface changes on the water cy-

cle (Alcamo and Henrichs, 2002; Arnell and Gosling, 2013; Liu et al., 2013, 2014; Nijssen et al., 2001a), exploring the spatial and temporal distribution of water resources (Abdulla et al., 1996; Alkama et al., 2010; Bierkens and Van Beek, 2009; Gerten et al., 2005; Tang et al., 2010), examining how human activities alter water demand and water resources (De Graaf et al., 2014; Döll et al., 2009; Hanasaki et al., 2008; Liu et al., 2015; Rost et al., 2008; Vörösmarty et al., 2000), and investigating the interactions between human activities and water availability by incorporating GHM with integrated assessment models (Kim et al., 2016).

Applying GHMs usually requires miscellaneous inputs, high computational costs, and a complex calibration process. These challenges stand out in practical situations, especially when the computational resources are limited. For instance, sensitivity analysis and uncertainty quantification are often needed for decision making, but the users usually cannot afford to run a large number of simulations with many GHMs like the Variable Infiltration Capacity (VIC) (also categorized as a land surface model, LSM) due to their high computational expense (Oubeidillah et al., 2014). Another situation is when the users seek reasonable estimates of water resources with minimal efforts rather than acquiring highly accurate estimates through expensive inputs of time and efforts. For example, when users seek to explore the hydroclimatology of a region and its long-term water balance (Sankarasubramanian and Vogel, 2002), then GHMs with fine spatial (e.g.,  $1/8^\circ$ ) and temporal resolution (e.g., hourly) are not necessarily needed. In this case, simple models that possess reasonable predictability and are computationally efficient tend to be more suitable. In addition, some studies have shown that GHMs or LSMs are sometimes outperformed by simple empirical statistical models (Abramowitz, 2005; Abramowitz et al., 2008; Best et al., 2015), suggesting that some GHMs or LSMs may underutilize the information in their climate inputs and that model complexity may undermine accurate prediction. This also indicates the potential advantages of a simple model over complex GHMs or LSMs. Thus, constructing simple models that can emulate the dynamics of more complex and computational expensive models (e.g., GHMs or LSMs) is warranted.

The motivation of this work arises from the need to construct a hydrological emulator (HE) that can efficiently mimic the complex GHMs to address the abovementioned issues for practical use, which provides the opportunity of speeding up simulations at the cost of introducing some simplification. We develop a HE that is easy to use and efficient for any interested groups or individuals to assess water cycle at basin/regional/global scales. This HE possesses the following features: (1) minimum number of parameters, (2) minimal climate input that is easy to acquire, (3) simple model structure, (4) reasonable model fidelity that captures both the spatial and temporal variability, (5) high computational efficiency, (6) applicable in a range of spatial scales, and (7) open-source and well-documented.

To achieve our goal of identifying a suitable HE, we have explored many hydrological models to find one that may meet our needs. We start with a simple baseline model characterized by mean seasonal cycle, i.e., the interannual mean value for every calendar day (Schaeffli and Gupta, 2007). Among others, we also explore the *abcd* model because (1) it is widely used and proven to have reasonable predictability (Fernandez et al., 2000; Martinez and Gupta, 2010; Sankarasubramanian and Vogel, 2002, 2003; Thomas, 1981; Vandewiele and Xu, 1992; Vogel and Sankarasubramanian, 2003), (2) it uses a monthly time step and requires less computational cost than daily or hourly models, (3) it has a solid physical basis and hence has potential to be extended to other temporal scales (Wang and Tang, 2014), (4) it requires minimal and easily available inputs, (5) it only involves 4–7 parameters, and (6) it can simulate variables of interest such as recharge, direct runoff, and baseflow that many other simple models cannot simulate (Vörösmarty et al., 1998). This study marks the first time that the *abcd*-based model is applied globally and also the first time the predictability and computational efficiency for both the lumped and distributed schemes are evaluated. Below we describe the baseline and the *abcd* models and data in Sect. 2; we present the evaluation of the two models and discuss their appropriateness of serving as a HE in Sect. 3; finally, in Sect. 4 we summarize this work with concluding remarks.

## 2 Methods and data

### 2.1 Model description

We examine two simple models – the baseline and the *abcd* model (both lumped and distributed scheme) in order to identify a suitable one for serving as a HE.

#### 2.1.1 Baseline model

Following the work of Schaeffli and Gupta (2007), we explore a baseline model characterized by the interannual mean value for every calendar day, i.e., climatology. In this study, the baseline model is based on monthly climatology runoff, which comes from a model simulation product – i.e., the runoff product from the Variable Infiltration Capacity (VIC) model (Leng et al., 2015). Specifically, we first calculate the grid-level interannual mean value for each of the 365 calendar days from the daily runoff of the benchmark product during 1971–2010 (see Sect. 2.3.2) and then aggregate daily climatology runoff to monthly climatology runoff at grid level. The baseline model here uses monthly climatology runoff for prediction. For example, if the climatology runoff for July in one grid cell is  $100 \text{ mm mon}^{-1}$ , then the prediction of total runoff for July of every year in that specific grid cell is  $100 \text{ mm mon}^{-1}$ .

2.1.2 The *abcd* model

The monthly *abcd* model was first introduced by Thomas (1981) to improve the national water assessment for the US, with a simple analytical framework using only a few descriptive parameters. It has been widely used across the world, especially for the US (Martinez and Gupta, 2010; Sankarasubramanian and Vogel, 2002, 2003). The model uses potential evapotranspiration (PET) and precipitation ( $P$ ) as input. The model defines four parameters  $a$ ,  $b$ ,  $c$ , and  $d$  that reflect regime characteristics (Sankarasubramanian and Vogel, 2002; Thomas, 1981) to simulate water fluxes (e.g., evapotranspiration, runoff, groundwater recharge) and pools (e.g., soil moisture, groundwater). The parameters  $a$  and  $b$  pertain to runoff characteristics, and  $c$  and  $d$  relate to groundwater. Specifically, the parameter  $a$  reflects the propensity of runoff to occur before the soil is fully saturated. The parameter  $b$  is an upper limit on the sum of evapotranspiration (ET) and soil moisture storage. The parameter  $c$  indicates the degree of recharge to groundwater and is related to the fraction of mean runoff that arises from groundwater discharge. The parameter  $d$  is the release rate of groundwater to baseflow, and thus the reciprocal of  $d$  is the groundwater residence time. Snow is not part of the original *abcd* model, which may result in poor performance of the model in cold regions where snow significantly affects the hydrological cycle. The work of Martinez and Gupta (2010) has added snow processes into the original *abcd* model, where the snowpack accumulation and snowmelt are estimated based on air temperature. Their work indicated that the incorporation of the snow processes in the monthly *abcd* model has significantly improved model performance in snow-covered area in the conterminous United States (see Fig. 4 in Martinez and Gupta, 2010).

In this study, we adopt the *abcd* framework from Martinez and Gupta (2010) (Fig. 1); meanwhile, we make three modifications to suit the needs of a HE for global applications. First, in order to enhance the model efficiency with as least necessary parameters as possible, instead of involving three tunable snow-related parameters in the calibration process, we set the values for two of the parameters (i.e., the temperature threshold above or below which all precipitation falls as rainfall or snow) from the literature (Wen et al., 2013) and only keep one tunable parameter  $m$  – the snowmelt coefficient ( $0 < m < 1$ ). Second, we introduce the baseflow index (BFI) into the calibration process to improve the partition of total runoff between the direct runoff and baseflow (see Sect. 2.4). Third, unlike previous studies, which only used the lumped scheme, we first explore the values of model application in a distributed scheme with a grid resolution of  $0.5^\circ$ . The detailed model descriptions and equations are presented in Appendix A, and the descriptions and ranges of model parameters are listed in Table 1.

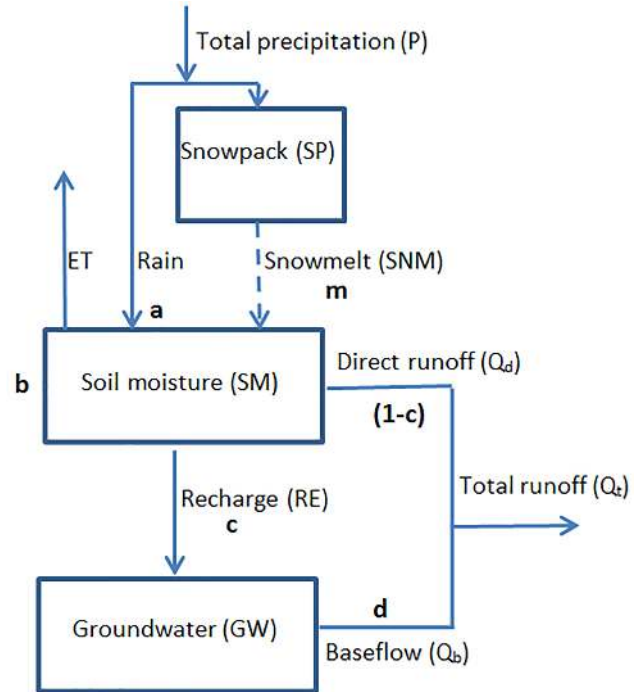


Figure 1. Schematic diagram of the *abcd* model, with enhancements of snow and partition of total runoff between direct runoff and baseflow.

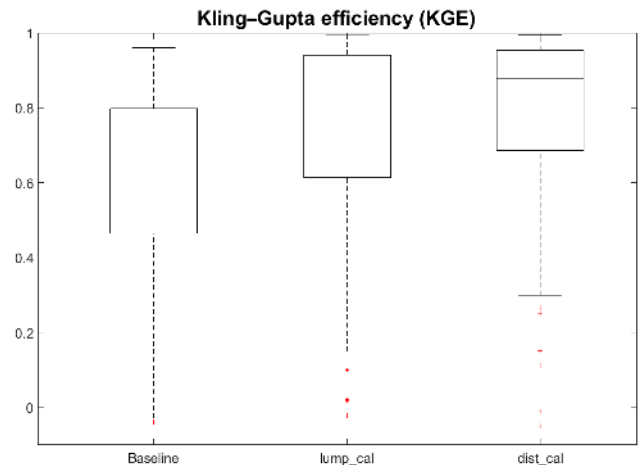


Figure 2. Kling–Gupta efficiency of the simulated basin-level total runoff across the global 235 basins (lump: lumped; dist: distributed; cal: calibration; the  $x$  axis labels of “lump\_cal” or “dist\_cal” represent the lumped/distributed scheme during the calibration period).

2.2 Model structure

In terms of the *abcd* model, we evaluate both the lumped and distributed model schemes, although most previous applications of the model are conducted in a lumped scheme (Bai et al., 2015; Fernandez et al., 2000; Martinez and Gupta, 2010; Sankarasubramanian and Vogel, 2002, 2003;

**Table 1.** Parameter description and ranges for the *abcd* model (the parameters *a*, *c*, *d* and *m* are dimensionless, and the unit for parameter *b* is mm).

Parameter	Description	Range	References
<i>a</i>	Propensity of runoff to occur before the soil is fully saturated	0–1	Alley (1984), Martinez and
<i>b</i>	Upper limit on the sum of evapotranspiration and soil moisture storage	0–4000	Gupta (2010), Sankarasubramanian
<i>c</i>	Degree of recharge to groundwater	0–1	and Vogel (2002), Vandewiele
<i>d</i>	Release rate of groundwater to baseflow	0–1	and Xu (1992)
<i>m</i>	Snowmelt coefficient	0–1	Wen et al. (2013), Martinez and Gupta (2010)

Vandewiele and Xu, 1992; Vogel and Sankarasubramanian, 2003). In the lumped scheme, each of the 235 river basins is lumped as a single unit, and each of the data input (see Sect. 2.3.1) are the lumped average across the entire basin, and thus all the model outputs are lumped as well. In terms of the distributed scheme, however, each  $0.5^\circ$  grid cell has its own data inputs, and likewise, the model outputs are simulated at the grid level. Although the two schemes differ in the spatial resolution of their inputs and outputs, their within-basin parameters are uniform. We use basin-uniform rather than grid-specific parameters for the distributed scheme for two reasons: (1) to enhance computational efficiency and (2) to avoid drastically different parameters for neighboring grid cells that may be unrealistic. Note that lateral flows between grid cells and basins are not included at this stage for the *abcd* model. For the baseline model, as documented in Sect. 2.1.1, every  $0.5^\circ$  grid cell of each basin has its own monthly climatology runoff estimates for each of the 12 calendar months.

## 2.3 Data

### 2.3.1 Climate data

The climate data needed for the *abcd* model only involve monthly total precipitation and monthly mean, maximum, and minimum air temperature. The data we use are obtained from WATer and global CHange (WATCH; Weedon et al., 2011), spanning the period of 1971–2010, and they are  $0.5^\circ$  gridded global monthly data. The climate data are used for model simulation over the global 235 major river basins (Kim et al., 2016). Additionally, we use the Hargreaves–Samani method (Hargreaves and Samani, 1982) to estimate potential evapotranspiration (PET), which is a required input for the *abcd* model, and it needs climate data of mean, maximum, and minimum temperatures for the calculation.

### 2.3.2 Benchmark runoff product

In this study, the *abcd* model is tested for its ability to emulate the naturalized hydrological processes of a reference model since the “true” naturalized hydrological processes are unknown. The “perfect model” approach is well adopted in climate modeling studies where one model is treated as

“observations” while the others are tested for their ability to reproduce observations (Murphy et al., 2004; Tebaldi and Knutti, 2007). Here, we use the process-based VIC model as the perfect model, which was also driven by the WATCH climate forcing.

The VIC runoff product here is a global simulation with a daily time step and spatial resolution of  $0.5^\circ$  for the period of 1971–2010, and the VIC daily runoff is aggregated to monthly data to be consistent with the temporal scale of the *abcd* model. The VIC model settings used in this study are based on the University of Washington VIC Global applications (<http://www.hydro.washington.edu/Lettenmaier/Models/VIC/Datasets/Datasets.shtml>, last access: 8 October 2016). The sub-grid variability in soil, vegetation, and terrain characteristics is represented in sub-grid area-specific parameter classifications. Soil texture and bulk densities are derived by combining the World Inventory of Soil Emission Potentials database (Batjes, 1995) and the 5 min digital soil map of the world from the Food and Agricultural Organization (FAO, 1998). Based on the work of Cosby et al. (1984), the remaining soil properties (e.g., porosity, saturated hydraulic conductivity, and unsaturated hydraulic conductivity) are derived. Vegetation type data are obtained from the global land classification of Hansen et al. (2000). Parameters including the infiltration parameter, soil layer depths, and those governing the baseflow function were calibrated for major global river basins and transferred to the global domain as documented in Nijssen et al. (2001b), based on which Zhang et al. (2014) and Leng et al. (2015) conducted additional calibrations in the domain of China. In this study, the VIC model was forced by WATCH climate forcing at the daily time step (Weedon et al., 2011), based on the calibrated parameters from Nijssen et al. (2001b), Zhang et al. (2014), and Leng et al. (2015). The simulated runoff used in this study has recently been validated globally within the framework of the Inter-Sectoral Impact Model Intercomparison Project and shows reasonable performance compared to other hydrological models (Hattermann et al., 2017; Krysanova and Hattermann, 2017).

The VIC runoff product (Hattermann et al., 2017; Leng et al., 2015) is then used as a benchmark for calibrating and validating the *abcd* model due to two reasons. First, VIC

runoff has been evaluated across many regions of the globe and is proven to work reasonably well (Abdulla et al., 1996; Hattermann et al., 2017; Maurer et al., 2001; Nijssen et al., 1997, 2001b). Second, the simulated monthly runoff by the *abcd* model is more representative of “natural conditions” because human activities (e.g., reservoir regulations and upstream water withdrawals) are currently not represented in the model. Thus, it tends to be more reasonable to compare the simulated runoff against the VIC natural runoff product rather than comparing against observed streamflow data from stream gauges (Dai et al., 2009; Wilkinson et al., 2014). Despite potential bias in the VIC runoff product, using it as a benchmark here is to demonstrate the capability of the HE developed in this work to mimic complex GHMs. Furthermore, the application of the HE is not tied to the VIC model and should be able to emulate other GHMs.

The VIC runoff product compares well to other products (see Figs. S1 and S2 in the Supplement), including the University of New Hampshire/Global Runoff Data Centre (UNH/GRDC) runoff product (Fekete and Vorosmarty, 2011; Fekete et al., 2002) and the global streamflow product (Dai et al., 2009). The scatterplot pattern of the VIC long-term annual runoff product vs. the GRDC product (GRDC, 2017) matches well with that of the UNH/GRDC runoff vs. the GRDC product (streamflow is transferred to the same unit as runoff by dividing by the basin area), which means the behavior of the VIC runoff product is similar to that of the UNH/GRDC product. Further, the correlation coefficient of the VIC and the UNH/GRDC long-term annual runoff is as high as 0.83 across the global 235 basins (Fig. S2). This suggests the reasonableness of the VIC runoff product because the UNH/GRDC runoff is calibrated with the GRDC observations. At the same time, the discrepancies between the VIC runoff products and the streamflow products (Fig. S2) may be attributed to human activities, such as reservoir regulations and upstream water withdrawals, which are not embedded in the runoff but reflected in the streamflow. This is because the VIC model simulates runoff in natural conditions, and then a stand-alone routing model can be used to route these flows downstream (Nijssen et al., 2001b). The routing model may account for human activities such as water extractions and reservoir operations (Haddeland et al., 2014). However, here we use the VIC runoff under natural conditions as the benchmark product, which explains the discrepancies between the VIC runoff and observed streamflow products.

Uncertainties arising from the runoff process in the VIC model should be acknowledged. The implementation of different runoff generation schemes (e.g., TOPMODEL – Topography based hydrological MODEL) within the same modeling framework is an alternative that can be adopted in the future to explore the uncertainty range. A recent inter-model comparison study shows that the VIC model falls within the range of large model ensembles (Hattermann et al., 2017). Notably, groundwater and its interaction with river and land surface are not represented in the model.

Thus, the model may not be able to fully capture the hydrologic responses in areas where lateral flow and the three-way streamflow–aquifer–land interactions are important. Further, vegetation dynamics and water management that may affect runoff are not considered in the model simulations. Nonetheless, the use of the HE documented here is not tied to the VIC, and it could be used to emulate other GHMs of interest.

## 2.4 Model calibration

Typically, most applications of the *abcd* model utilize single-objective optimization for total runoff (or streamflow) during the calibration process to minimize the difference between measured and simulated streamflow (Bai et al., 2015; Martinez and Gupta, 2010; Sankarasubramanian and Vogel, 2002). While this may lead to a good fit for simulated total runoff, however, it may result in inappropriate partition of total runoff between direct runoff and baseflow. To improve the accuracy of the simulated total runoff and the partition between direct runoff and baseflow, we introduce the baseflow index (BFI) into the objective function.

Unlike the baseline model, the *abcd* model requires a calibration step for reasonable parameterization so as to enable good prediction. As mentioned above, we incorporate BFI into the objective function during the calibration process. On one side, we maximize the Kling–Gupta efficiency (KGE) (Gupta et al., 2009), which is used as a metric to measure the accuracy of the simulated total runoff relative to the VIC benchmark runoff. The KGE is defined as the difference of unity and the Euclidian distance (ED) from the ideal point; thus, we maximize KGE through minimizing the ED. The KGE and ED are calculated as follows (Gupta et al., 2009):

$$\text{KGE} = 1 - \text{ED}, \quad (1)$$

$$\text{ED} = \sqrt{(r - 1)^2 + (\alpha - 1)^2 + (\beta - 1)^2}, \quad (2)$$

$$r = \frac{\text{Cov}_{so}}{\sigma_s \cdot \sigma_o}, \quad (3)$$

$$\alpha = \sigma_s / \sigma_o, \quad (4)$$

$$\beta = \mu_s / \mu_o, \quad (5)$$

where  $r$ ,  $\alpha$ ,  $\beta$ , and  $\text{Cov}_{so}$  are relative variability, bias, correlation coefficient, and covariance between the simulated and observed values (here we treat the VIC runoff as the observed), respectively;  $\mu$  and  $\sigma$  represent the mean and SD (subscripts “s” and “o” stand for simulated and observed values). On the other side, we also nudge the simulated BFI towards the benchmark BFI (here we treat the benchmark BFI as the observed) – the mean BFI of the four products from (Beck et al., 2013). Then, the objective function is as follows:

$$\min (\text{ED} + \text{abs}(\text{BFI}_{\text{obs}} - \text{BFI}_{\text{sim}})), \quad (6)$$

where min stands for minimizing the value in the parenthesis, abs represents absolute value, ED is the Euclidian distance between the simulated and observed total runoff (Gupta

et al., 2009), and  $BFI_{obs}$  and  $BFI_{sim}$  are the observed and simulated BFI, respectively. Here we treat the benchmark runoff from the VIC and BFI from Beck et al. (2013) as observed values. We then minimize the objective function for parameter optimization by utilizing a genetic algorithm (GA) routine (Deb et al., 2002). Note that for the distributed model scheme, we aggregate the grid-level total runoff estimates to basin level and then nudge it toward the basin-level benchmark total runoff during the calibration process.

## 2.5 Model simulations

To evaluate the predictability and efficiency of the baseline and the *abcd* model so as to identify a suitable one to serve as a HE, we have conducted a series of simulations. Specifically, for the baseline model, no simulations are needed as it uses the interannual mean value for each month – 12 monthly values – as prediction, so we just replicate the 12 monthly runoff values for 1971–2010 and for each of the global 235 basins and then compare them against the benchmark runoff product. For the *abcd* model, two sets of model simulations across the global 235 basins are conducted, with one set for calibration and the other one for validation, for both the lumped and distributed model schemes. For the first set, we run the model for each basin for the period of 1971–1990 to get basin-specific parameters by using the GA approach (see Sect. 2.4). For the second set, using the parameters identified in the first set of simulation, we run the model for the period of 1991–2010 to validate the model predictability and also evaluate the computational efficiency. Model inputs and outputs in the distributed scheme are at a spatial resolution of  $0.5^\circ$ , whereas those in the lumped scheme are all in lumped single units for each basin. All model simulations are conducted at a monthly time step. Note that users can run the identified HE for the global 235 basins, or for as many basins as they want for either scheme, as all the related basin-specific input data and calibrated parameters for both schemes are open-source.

## 3 Results and discussions

### 3.1 Comparison of performances between the baseline and the *abcd* model

Generally, we find that the baseline model performs worse than the *abcd* model (Fig. 2). The baseline model exhibits a lower global mean KGE value (0.61) than the lumped and distributed schemes of the *abcd* model (0.75 and 0.79, respectively). In addition, our analysis indicates that the incorporation of BFI into the objective function leads to a significant improvement in the partition of total runoff between direct runoff and baseflow (Figs. 3 and S4), without compromising predictability for total runoff, i.e., the global mean KGE values for modeled total runoff with or without the incorporation of BFI are almost the same (0.75 vs. 0.76).

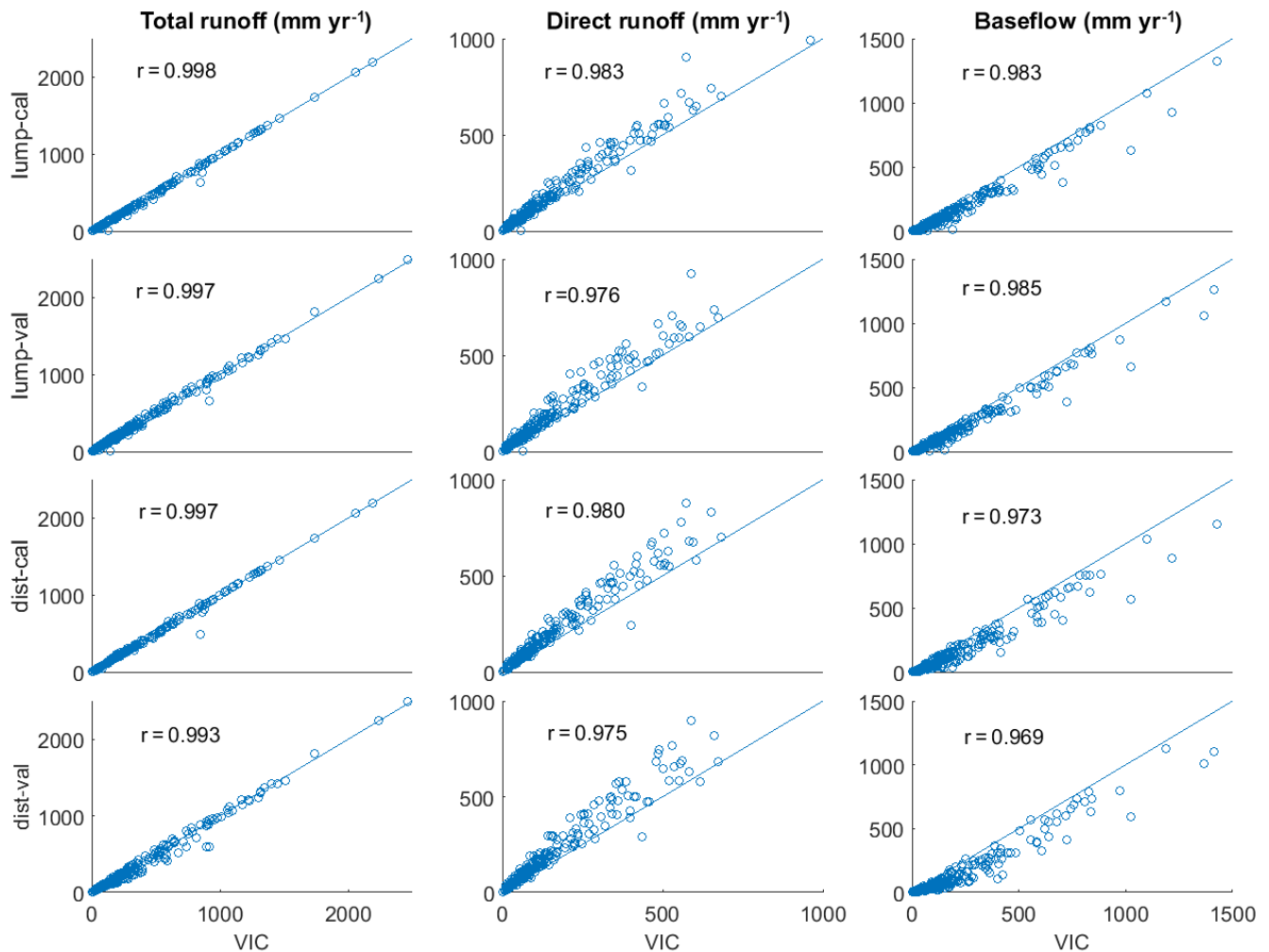
Specifically, for the case of involving both the total runoff and BFI in the objective function, the correlation efficiencies ( $r$ ) between the long-term annual benchmark and modeled direct runoff and between benchmark and modeled baseflow from the lumped scheme across global basins are both 0.98 (Fig. 3) and are much higher than those of 0.86 and 0.72 in the case of only involving the total runoff in the objective function (Fig. S4). Given the superiority of the *abcd* model over the baseline model, we focus in the following sections on evaluating the predictability and computational efficiency of the *abcd* model and its potential to serve as a HE.

### 3.2 Evaluation of model predictability

In terms of total runoff, we find that the lumped and distributed schemes are comparably capable of simulating long-term mean annual quantity, temporal variations, and spatial patterns for the vast majority of river basins globally (Figs. 3–5). Estimates of long-term mean annual total runoff from both the lumped and distributed schemes match very well with that of VIC total runoff across the 235 basins, with a correlation coefficient ( $r$ ) higher than 0.96, for both the calibration and validation period (Fig. 3). Similarly, the basin-level estimates of long-term mean annual direct runoff and baseflow also match well with those of the VIC across the globe, for both schemes and both periods (Fig. 3). This suggests that both schemes possess the capability of partitioning total runoff.

Furthermore, both schemes display good capability of capturing the seasonal variations in the total runoff for both the calibration and validation period (Figs. 4 and S5). Meanwhile, although the spatial patterns of annual total runoff from the lumped scheme present a general match with that of the VIC, this does not reflect the spatial variations inside a basin, which are, however, captured by the distributed scheme (Fig. 5). Likewise, overall much lower percentage differences between the modeled runoff from the distributed scheme and the VIC runoff product than those between the VIC and the lumped scheme further corroborate the significantly better performance of the distributed scheme (Fig. S6). Both schemes still show large percentage differences in some dry (e.g., North Africa) or cold regions (e.g., Tibetan Plateau). This is because the runoff there is at a low magnitude, and thus small changes in runoff will lead to large percentage differences. Therefore, the distributed scheme provides overall slightly higher KGE (Fig. 6), with a global mean KGE value of 0.79 as compared to 0.75 for the lumped scheme (Fig. 2).

To ensure good model predictability for the major water fluxes, we also evaluate the modeled ET estimates. The modeled ET compares reasonably well with the VIC ET product as well as with the mean synthesis of the LandFlux-EVAL ET product (Mueller et al., 2013), displaying similar spatial variations (Fig. S7). Likewise, the distributed *abcd* scheme tends to have better capability in presenting spatial hetero-

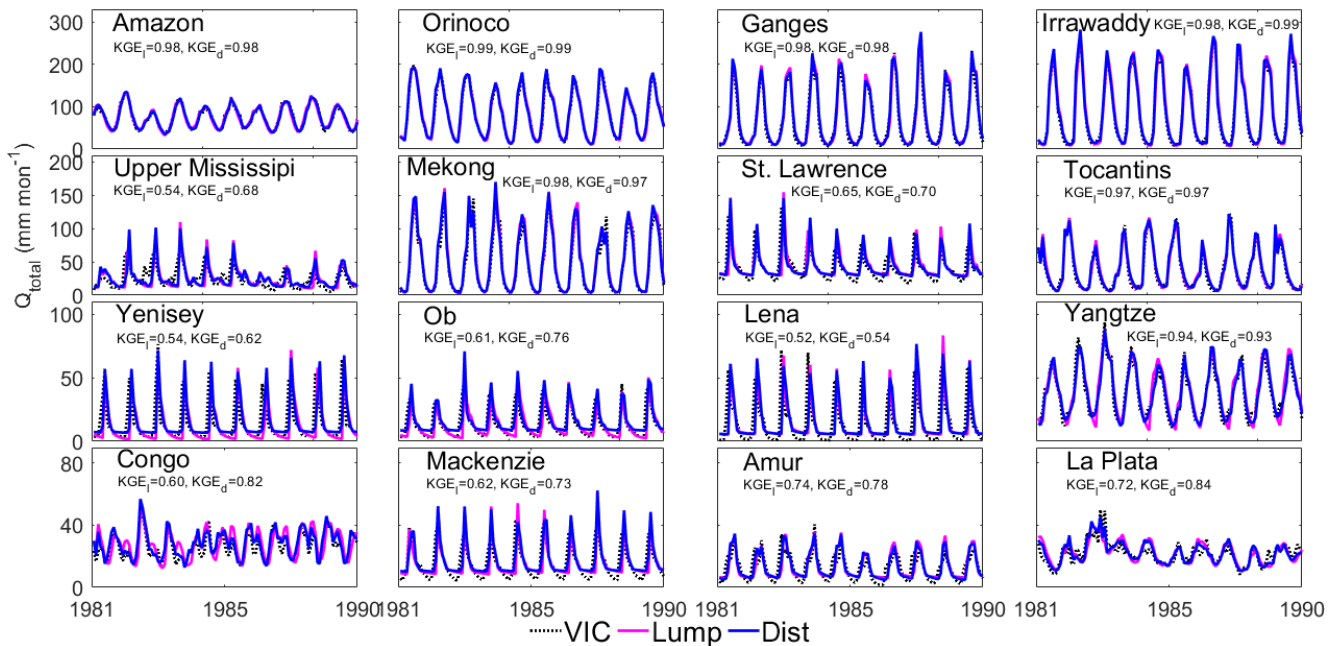


**Figure 3.** Comparison of basin-specific long-term annual total runoff, direct runoff, and baseflow estimates from both the lumped and distributed *abcd* model schemes against VIC products, across the global 235 basins and for the calibration period of 1971–1990 and validation period of 1991–2010. The labels are denoted as a combination of model scheme and period, where lump and dist stand for lumped and distributed model scheme and cal and val represent the calibration and validation period, respectively. These denotations remain the same for all figures in this work. Note that the basin-level VIC baseflow is derived by multiplying the gridded VIC long-term annual total runoff and the mean of the four gridded baseflow index products from Beck et al. (2013) and then aggregating from grid level to basin level. The basin-level VIC direct runoff is then calculated by subtracting baseflow from the total runoff.

generality than the lumped one. In addition, the percentage differences between our modeled ET and the VIC ET product further confirm that the distributed scheme significantly outperforms the lumped one (Fig. S8), with much lower differences from the VIC ET product, although discrepancies still exist in some extremely cold (e.g., Greenland) or dry regions (e.g., North Africa), which is because small differences in ET will lead to a large percentage difference in those regions with low ET. Further, given that the changes in basin-scale monthly soil moisture are relatively small, precipitation should approximate to the sum of ET and runoff according to the water mass balance; the good predictability of seasonality in runoff as illustrated in Fig. 4 also reflects similar performance for ET.

The distributed scheme appears to outperform the lumped scheme in terms of goodness of fit, especially in some cold (e.g., Arctic, northern European, interior Tibet) and in some dry (e.g., North Africa) regions (Fig. 6). This is possibly because distributed inputs can reflect basin-level heterogeneity, and thus better capture the characteristic of the hydrological conditions in those regions. However, neither scheme performs well at the southern end of the Andes Mountains (Fig. 6). This may be attributed to the complex land surface characteristics in that mountainous area, which cannot be resolved due to the coarse spatial resolution. Moreover, the distributed scheme seems not to perform very well in some cold regions (Fig. 6), which is possibly due to a lack of representation of permafrost in the model.





**Figure 4.** Time series of basin-specific total runoff ( $Q_{\text{total}}$ ) from the VIC product and the lumped and distributed *abcd* schemes for the world's 16 river basins with top annual flow (Dai et al., 2009) during 1981–1990 (part of the calibration period 1971–1990).  $KGE_l$  and  $KGE_d$  stand for the KGE value for the lumped and distributed scheme, respectively.

Previous studies investigating the credibility of lumped and distributed hydrological models indicate that, in many cases, lumped models perform comparably to or just as well as distributed models (Asadi, 2013; Brihnet and Benaabdiate, 2016; Ghavidelfar et al., 2011; Michaud and Sorooshian, 1994; Obled et al., 1994; Reed et al., 2004; Refsgaard and Knudsen, 1996; Yao et al., 1998). However, distributed models may have advantages for predicting runoff in ungauged watersheds (Reed et al., 2004; Refsgaard and Knudsen, 1996) and for capturing the spatial distribution of runoff due to heterogeneity in rainfall patterns or in land surface (Downer et al., 2002; Paudel et al., 2011; Yao et al., 1998). Our results on the predictability of the lumped and distributed *abcd* model are in line with previous findings in the literature.

The good agreement between our modeled water fluxes, including total runoff, direct runoff, baseflow, and ET, and the benchmark products provides confidence in the capability of both the lumped and distributed schemes in estimating temporal and spatial variations in major water fluxes across the globe. In addition, to identify a suitable HE, the required computation cost is another key factor as detailed below.

### 3.3 Evaluation of computational efficiency

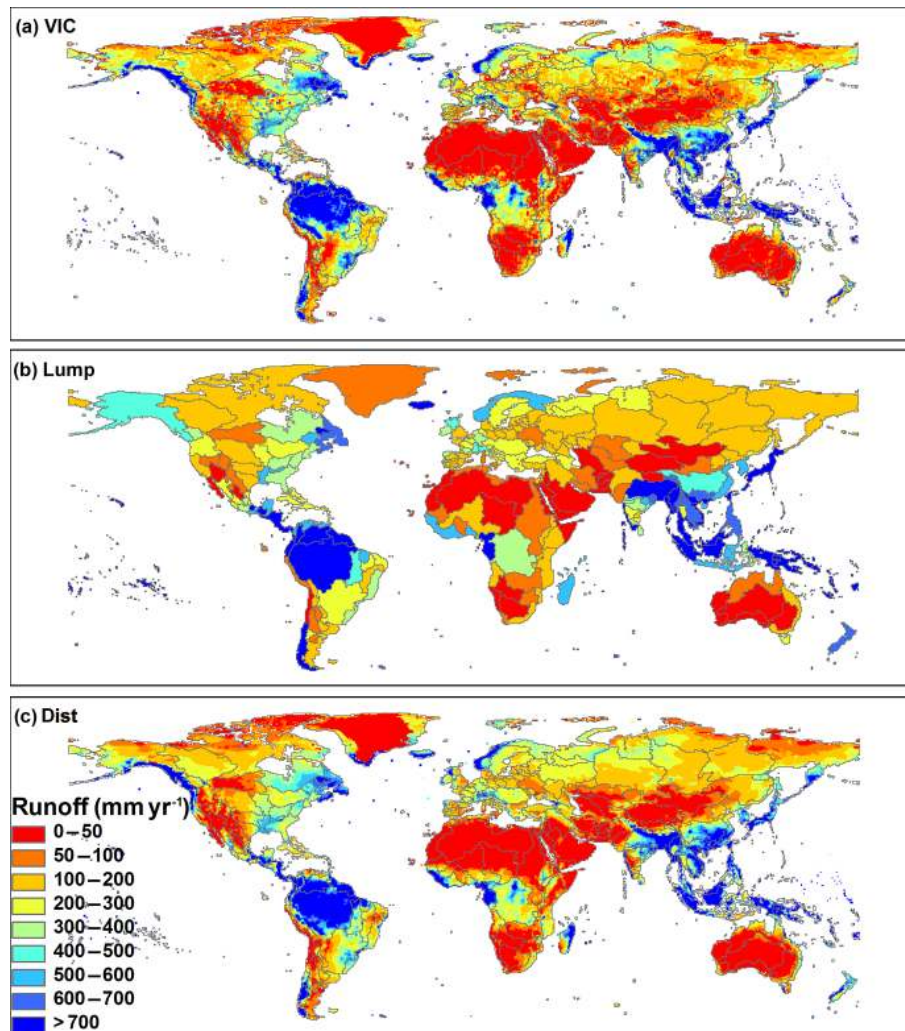
While the performance of model predictability is comparable for the lumped and distributed schemes as elucidated above, great disparities still exist for the runtime of the two schemes and the VIC model (Table S1 in the Supplement).

Take the Amazon Basin that covers a total number of 2002  $0.5^\circ$  grid cells as an example: it takes 11.05 min for model calibration via the GA method for the distributed scheme but only 0.16 min for the lumped one. A similar disparity is also found for model simulation with calibrated parameters, with a runtime of 0.03 and 3.20 s for a 1000-year simulation of the Amazon Basin for the lumped and distributed schemes, respectively. However, according to the authors' experience, it will take  $\sim 1$  week for the VIC model to accomplish the same job, which is far more computationally expensive. In general, the computational efficiency of the lumped scheme is 2 orders of magnitude higher than the distributed one, although that of the distributed one is still much higher than the VIC ( $\sim 5$  orders of magnitude) and many other GHMs and LSMs. Note that all of the simulations here are conducted on the Pacific Northwest National Laboratory (PNNL)'s Institutional Computing (PIC) Constance cluster using one core (Intel Xeon 2.3 GHz CPU) with the same configuration.

### 3.4 Potential application of the *abcd* model as a hydrological emulator

The good predictability and computational efficiency of both the distributed or lumped schemes as elucidated in Sects. 3.2 and 3.3 suggest its suitability for serving as HEs that can efficiently emulate complex GHMs (e.g., the VIC or others). The source codes, input data, and basin-specific parameters across the globe for both the lumped and distributed schemes are open-source and well-documented, which will make the



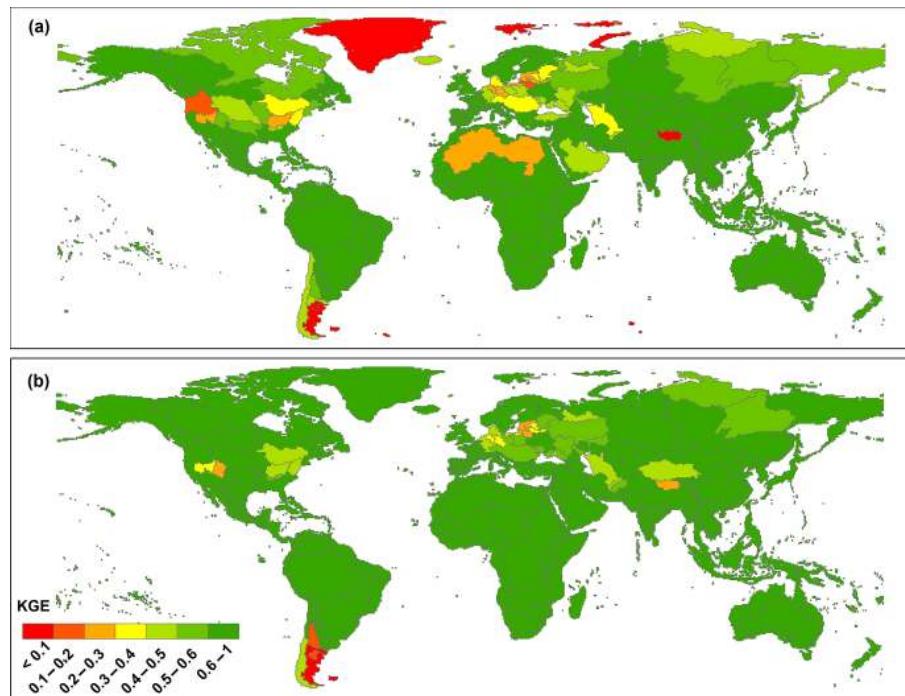


**Figure 5.** Spatial patterns of long-term annual total runoff ( $\text{mm yr}^{-1}$ ) during 1971–1990 across the global 235 basins: (a) VIC runoff product, (b) total runoff estimates from the lumped *abcd* scheme (Lump: lumped), and (c) total runoff estimates from the distributed *abcd* scheme (Dist: distributed).

HE ready to use and facilitate the schemes' wide and easy use with minimal efforts.

The choice of either the distributed or lumped scheme as HE depends on the user's specific needs. There is a trade-off between the model predictability and computational efficiency. While the distributed scheme tends to better capture the spatial heterogeneity of water fluxes and can produce grid-level outputs that the lumped scheme cannot, it incurs a higher computational cost than the lumped scheme. For applications that aim to strike a balance between predictability and computation cost, such as a practical assessment of water resources or an estimation of water supply for integrated assessment models (IAMs) or a quantification of uncertainty and sensitivity analyses, it would be reasonable to employ the lumped scheme as a HE. The lumped scheme is especially advantageous due to its minimal calibration and computational cost, parsimonious efforts for model imple-

mentation, and reasonable fidelity in estimating major water fluxes (e.g., runoff, ET). For users from the IAM community, the lumped scheme might be sufficiently suitable for their needs since (1) the lumped scheme can operate at the same spatial resolution at which IAMs typically balance water demands and supplies (Edmonds et al., 1997; Kim et al., 2006, 2016) and (2) the inherent uncertainty of the lumped scheme is likely comparable to or even overshadowed by the intrinsic uncertainty of IAMs (Kraucunas et al., 2015; O'Neill et al., 2014). Similarly, for users who aim to conduct uncertainty and sensitivity analyses, the high computational efficiency of the lumped scheme allows the users to emulate the hydrological model of interest (e.g., GHMs, LSMs) and then run a large number of simulations to conduct the model's uncertainty and sensitivity analysis (Scott et al., 2016). Therefore, the high computational efficiency makes the lumped scheme more appealing as a HE in these cases. However, if the re-



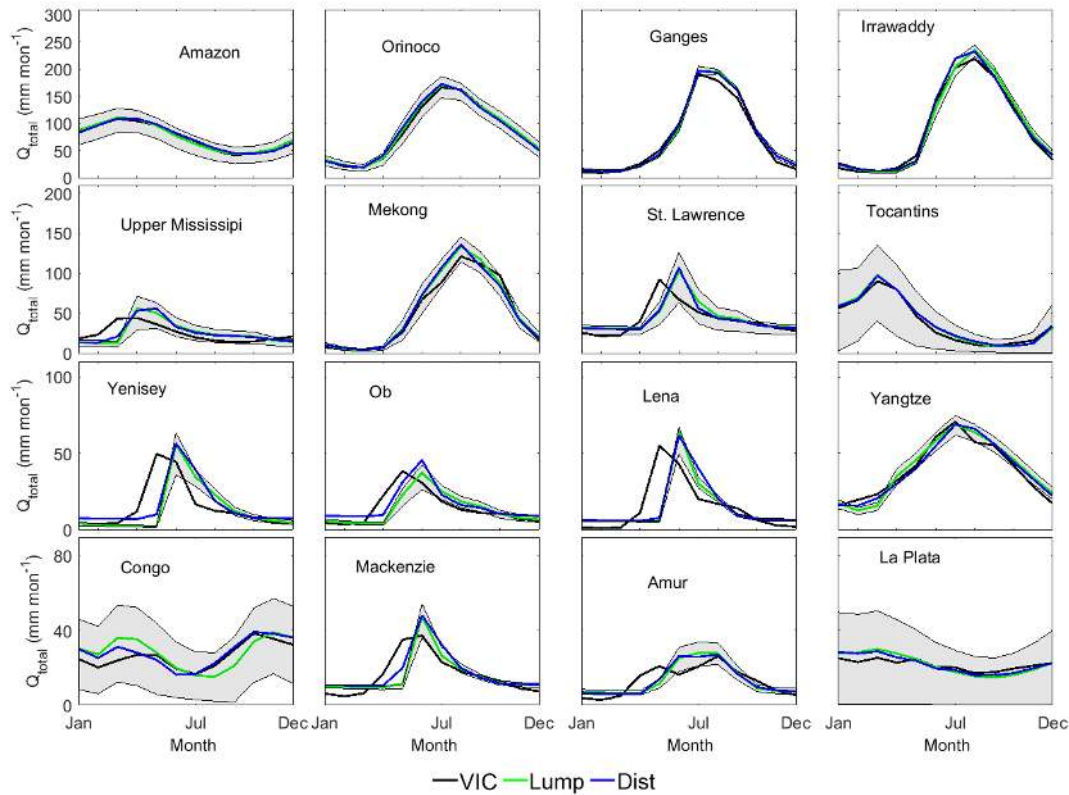
**Figure 6.** The spatial pattern of the Kling–Gupta efficiency (KGE) for the total runoff estimates of the global 235 basins for the calibration period of 1971–1990: (a) the lumped *abcd* scheme; and (b) the distributed *abcd* scheme.

search questions hinge on the gridded estimates or emphasize the spatial heterogeneity of the water fluxes or pools, it would be more desirable to deploy the distributed scheme as a HE instead. For example, a follow-up work is coupling the distributed scheme of the HE with a widely used IAM, the Global Change Assessment Model (GCAM; Edmonds et al., 1997) and then using the coupled model to investigate the impacts of a variety of land use policies on global water scarcity, where the HE is used to estimate grid-level runoff globally under different land use policies.

While many studies indicate that basin runoff generation is sensitive to factors such as physical characteristics, spatiotemporal variability in storage distribution, and forcing input, evidence also shows that basin response can be captured using a handful of parameters (Hsu et al., 1995; Young and Parkinson, 2002). In this study, the lumped scheme of the HE ignores the spatiotemporal variability in basin characteristics by averaging the input forcing data; consequently, the associated responses in within-basin runoff or ET variations cannot be captured. In contrast, the distributed scheme presents a better performance in capturing spatiotemporal variability in runoff and ET with the use of the same input data, and without increasing the number of parameters. Thus, the use of the distributed scheme is preferred when the tradeoff in the computational efficiency is not a constraining factor.

Moreover, a combination of a top–down approach (Sivapalan et al., 2003) and a multi-objective approach to model evaluation (Gupta et al., 1998) could be used to explore in-

ternal basin behavior, wherein the top–down approach would start from a simple structure and then progressively expand based on its caveats in reproducing overall basin behavior (e.g., Jothityangkoon et al., 2001). In this study we adopt a similar framework, by starting from a baseline model and then expanding to the *abcd* model with snow representation, also by incorporating the baseflow index into the objective function to exert a multi-objective approach. Our assessment indicates that a baseline model characterized by mean seasonal cycle still holds promise in predicting runoff at basins with small variability in basin characteristics, such as the basins of the Ob’, Lena, Yenisey, Siberia and Mackenzie in the Arctic area, where the baseline model yields KGE values of greater than 0.90 from our evaluation. Further, while Martinez and Gupta (2010) indicated that the incorporation of the snow component and an additional snow parameter into the original *abcd* model has greatly improved model performance in snow-dominated regions, areas without prevailing snow (e.g., the tropical zone) could still utilize the original version of the *abcd* model to keep the model as parsimonious as possible without compromising model predictability. In addition, although our results reveal that the incorporation of the baseflow index into the objective function generally improves the model performance in the partitioning of runoff between direct runoff and baseflow, simply employing a single-objective approach (i.e., only involving total runoff) also works well for some basins such as northern interior Africa and interior Australia. Thus, the single-objective



**Figure 7.** Parameter-induced uncertainty in total runoff estimate of the lumped *abcd* scheme for the world's 16 river basins with top annual flow. The lines of VIC, Lump and Dist stand for the VIC benchmark runoff product and total runoff estimates from the lumped and distributed *abcd* scheme with the calibrated parameters, respectively, and the gray area represents the spread derived from variations in parameters of the lumped scheme.

approach is also acceptable for those basins with the advantage of simplicity without compromise in performance. In short, according to specific basin characteristics and the research needs, suitable model complexity and the number of parameters could be identified by following the above-mentioned scenarios, such that either the baseline model or a reduced format of the HE (e.g., without snow representation or single-objective) could be potentially utilized with the merits of simplicity, reasonable predictability, and computational efficiency rather than adopting the full format of the HE. This HE could be used to emulate a wide range of models with different spatial and temporal complexities, and its performance may vary from model to model. Thus, examining and comparing the extent to which the HE could mimic the behaviors of different GHMs and LSMs is of future research interest to us. In addition, future research can extend this work by systematically investigating the role of different levels of inputs and parameters on model performance in different basins across the globe.

Based upon our open-source HE and the validated basin-specific parameters across the globe, researchers can easily investigate the variations in water budgets at the basin, regional, or global scale of interest, with minimum require-

ments regarding input data, efficient computation performance, and reasonable model fidelity. Likewise, researchers can utilize the framework of the HE with any alternative input data or recalibrate the HE to emulate other complex GHMs or LSMs of interest to meet their own needs.

### 3.5 Case study for uncertainty analysis

To demonstrate the capability of the examined *abcd* model serving as a HE, we use the lumped scheme to conduct parameter-induced uncertainty analysis for the runoff simulation at the world's 16 river basins with top annual flow (Dai et al., 2009). Specifically, for each of the 16 basins, we first apply  $\pm 10\%$  change to each of the five calibrated parameters ( $a$ ,  $b$ ,  $c$ ,  $d$ ,  $m$ ) to compose varying ranges; note that we just truncate the range to those ranges listed in Table 1 if the  $\pm 10\%$  change exceeds the valid range. Then we randomly sample the five parameters from corresponding ranges for 100 000 times (i.e., 100 000 combinations of parameters). After that, we run the lumped scheme 100 000 times for each basin with the 100 000 combinations of parameters to examine the parameter-induced uncertainty in total runoff. The uncertainty analysis indicates that most basins are robust to

changes in parameters, except for the Tocantins, Congo and La Plata (Fig. 7). In other words, for the basins Congo and La Plata, slight changes in parameters may lead to large changes in runoff estimates. Then the uncertainty in the calibrated parameters for the two basins may lead to large bias in the simulated runoff, which may more or less explain why modeled runoff for the two basins tends to have higher biases than that for other basins (Fig. 4). Notably, the 100 000 simulations only take  $\sim 80$  s on a Dell Workstation T5810 with one Intel Xeon 3.5 GHz CPU, which demonstrates the extraordinary computational efficiency of the lumped scheme and its advantage for serving as a HE.

#### 4 Conclusions

Toward addressing the issue that many global hydrological models (GHMs) are computationally expensive and thus users cannot afford to conduct a large number of simulations for various tasks, we firstly construct a hydrological emulator (HE) that possesses both reasonable predictability and computation efficiency for global applications in this work. Built upon the widely used *abcd* model, we have adopted two snow-related parameters from literature rather than tuning them for parameter parsimony and also have improved the partition of total runoff between the direct runoff and baseflow by introducing the baseflow index into the objective function of the parameter optimization. We then evaluate the appropriateness of the model serving as an emulator for a complex GHM – the VIC – for both the lumped and distributed model schemes by examining their predictability and computational efficiency.

In general, both distributed and lumped schemes have a comparably good capability of simulating spatial and temporal variations in the water balance components (i.e., total runoff, direct runoff, baseflow, evapotranspiration). Meanwhile, the distributed scheme has slightly better performance than the lumped one (e.g., capturing spatial heterogeneity), with a mean Kling–Gupta efficiency of 0.79 vs. 0.75 across the global 235 basins, and it also provides grid-level estimates that the lumped scheme is incapable of. Additionally, the distributed scheme performs better in extreme climate regimes (e.g., Arctic, North Africa) and Europe. However, the distributed scheme incurs 2 more orders of magnitude of computational cost than the lumped one. A case study of an uncertainty analysis with 100 000 simulations for each of the world's 16 basins with top annual streamflow further demonstrates the lumped scheme's extraordinary advantage in terms of computational efficiency. Therefore, the lumped scheme could be an appropriate HE: it has reasonable predictability and high computational efficiency. At the same time, the distributed scheme could be a suitable alternative for research questions that hinge on grid-level spatial heterogeneity. Finally, upon open-sourcing and documenting it well, the HE is ready to use and it provides researchers with an easy way to

investigate the variations in water budgets at a variety of spatial scales of interest (e.g., basin, region, or globe), with minimum requirements of efforts, reasonable model predictability, and extraordinary computational efficiency.

*Code and data availability.* The hydrological emulator (HE) is freely available on the open-source software site GitHub (<https://github.com/JGCRI/hydro-emulator/>). We have released the version of the specific HE v1.0.0 referenced in this paper at <https://github.com/JGCRI/hydro-emulator/releases/tag/v1.0.0>, where the source code (written in Matlab), all related inputs, calibrated parameters and outputs for each of the global 235 basins as well as the user manual are available. In addition, the HE documented here has been translated into Python and is being incorporated into Xanthos (Li et al., 2017), which is an open-source global hydrologic model that allows users to run different combinations of evapotranspiration, runoff, and routing models. The HE will be the default runoff model used in Xanthos 2.0 and will be available on GitHub (<https://github.com/JGCRI/xanthos>).

**Appendix: Descriptions and equations of the *abcd* model**

The *abcd* model was first introduced by Thomas (1981), and Martinez and Gupta (2010) added snow processes into the model. In this work, we adopted the snow scheme in Martinez and Gupta (2010):

$$\text{Snow}_i = \begin{cases} 0 & T^{\text{rain}} < T_i^{\text{min}} \\ P_i \cdot \frac{T^{\text{rain}} - T_i^{\text{min}}}{T^{\text{rain}} - T^{\text{snow}}} & T^{\text{snow}} < T_i^{\text{min}} < T^{\text{rain}} \\ P_i & T_i^{\text{min}} < T^{\text{snow}}, \end{cases} \quad (1)$$

$$\text{SP}_i = \text{SP}_{i-1} - \text{SNM}_i + \text{Snow}_i, \quad (2)$$

$$\text{SNM}_i = \begin{cases} 0 & T_i^{\text{min}} < T^{\text{snow}} \\ (\text{SP}_{i-1} + \text{Snow}_i) \cdot m & T^{\text{snow}} < T_i^{\text{min}} < T^{\text{rain}} \\ \frac{T^{\text{rain}} - T_i^{\text{min}}}{T^{\text{rain}} - T^{\text{snow}}} \cdot (\text{SP}_{i-1} + \text{Snow}_i) \cdot m & T^{\text{rain}} < T_i^{\text{min}}, \end{cases} \quad (3)$$

where  $P_i$ ,  $\text{SP}_i$ ,  $\text{SNM}_i$ , and  $\text{Snow}_i$  are total precipitation, snowpack storage, snowmelt, and the precipitation as snowfall at time step  $i$ , respectively;  $T^{\text{rain}}$  (or  $T^{\text{snow}}$ ) stands for the temperature threshold above (or below) which all precipitation falls as rainfall (or snow), and  $T_i^{\text{min}}$  is the minimum temperature at time step  $i$ , and the parameter  $m$  is the snowmelt coefficient. Rather than keeping the three parameters  $T^{\text{rain}}$ ,  $T^{\text{snow}}$ , and  $m$ , we adopt the  $T^{\text{rain}}$  value of 2.5 °C and  $T^{\text{snow}}$  value of 0.6 °C (Wen et al., 2013) and thus only keep one snowmelt-related parameter  $m$  in the model, in order to alleviate the computation load during the parameter optimization process.

The model defines two state variables “available water” and “evapotranspiration opportunity”, denoted as  $W_i$  and  $Y_i$ , respectively. The  $W_i$  is defined as

$$W_i = \text{SM}_{i-1} + \text{Rain}_i + \text{SNM}_i, \quad (4)$$

where  $\text{SM}_{i-1}$  is soil moisture at the beginning of time step  $i$ ;  $\text{Rain}_i$  and  $\text{SNM}_i$  are rainfall and snowmelt during period  $i$ .

$Y_i$  stands for the maximum water that can leave the soil as evapotranspiration (ET) at period  $i$ , and it is defined as below:

$$Y_i = \text{ET}_i + \text{SM}_i, \quad (5)$$

where  $\text{ET}_i$  is the actual ET at time period  $i$  and  $\text{SM}_i$  is soil moisture at the end of time step  $i$ . Further,  $Y_i$  has a nonlinear relationship with  $W_i$  as

$$Y_i = \frac{W_i - b}{2a} - \sqrt{\left(\frac{W_i - b}{2a}\right)^2 - W_i \cdot b/a}, \quad (6)$$

where  $a$  and  $b$  are parameters detailed in Sect. 2.1.2.

The allocation of  $W_i$  between  $\text{ET}_i$  and  $\text{SM}_i$  is estimated by assuming that the loss of soil moisture by ET will be proportional to potential evapotranspiration (PET) as

$$\frac{d\text{SM}}{dt} = -\text{PET} \cdot \frac{\text{SM}}{b}, \quad (7)$$

where PET is calculated by using the Hargreaves–Samani method (Hargreaves and Samani, 1982).

After integrating the above differential equation and assuming  $\text{SM}_{i-1} = Y_i$ ,  $\text{SM}_i$  can be derived as

$$\text{SM}_i = Y_i \cdot \exp(-\text{PET}_i/b). \quad (8)$$

Then,  $\text{ET}_i$  can be calculated through Eq. (5).

In the model framework,  $W_i - Y_i$  is the sum of the groundwater recharge (RE) and direct runoff ( $Q_d$ ), and the allocation is determined by the parameter  $c$ :

$$\text{RE}_i = c \cdot (W_i - Y_i), \quad (9)$$

$$Q_d = (1 - c) \cdot (W_i - Y_i). \quad (10)$$

The baseflow from the groundwater (GW) pool is modeled as

$$Q_b = d \cdot \text{GW}_i, \quad (11)$$

where  $d$  is a parameter reflecting the release rate of groundwater to baseflow. Then the total runoff ( $Q_t$ ) is the sum of the direct runoff and baseflow:

$$Q_t = Q_d + Q_b. \quad (12)$$

The  $\text{GW}_i$  is the sum of groundwater storage at the end of the last time step and the groundwater recharge minus the baseflow, and  $\text{GW}_i$  is derived as

$$\text{GW}_i = \frac{\text{GW}_{i-1} + \text{RE}_i}{1 + d}. \quad (13)$$

Then, all the water fluxes and pools are solved.



The Supplement related to this article is available online at <https://doi.org/10.5194/gmd-11-1077-2018-supplement>.

*Author contributions.* YL and MH designed this work, and all co-authors offered help through discussions. YL developed the hydrological emulator and conducted the simulations and evaluations. YL wrote the manuscript, and all co-authors contributed to the revision.

*Competing interests.* The authors declare that they have no conflict of interests.

*Acknowledgements.* This research was supported by the US Department of Energy, Office of Science, as part of research in the Multi-Sector Dynamics, Earth and Environmental System Modeling Program. The Pacific Northwest National Laboratory is operated for DOE by Battelle Memorial Institute under contract DE-AC05-76RL01830. The views and opinions expressed in this paper are those of the authors alone. We thank Chris Vernon for his help in maintaining the Github repository for the hydrological emulator (<https://github.com/JGCRI/hydro-emulator/>).

Edited by: Bethanna Jackson

Reviewed by: Tirthankar Roy, Shusen Wang, and two anonymous referees

## References

- Abdulla, F. A., Lettenmaier, D. P., Wood, E. F., and Smith, J. A.: Application of a macroscale hydrologic model to estimate the water balance of the Arkansas-Red River Basin, *J. Geophys. Res.-Atmos.*, 101, 7449–7459, 1996.
- Abramowitz, G.: Towards a benchmark for land surface models, *Geophys. Res. Lett.*, 32, L22702, <https://doi.org/10.1029/2005GL024419>, 2005.
- Abramowitz, G., Leuning, R., Clark, M., and Pitman, A.: Evaluating the performance of land surface models, *J. Climate*, 21, 5468–5481, 2008.
- Alcamo, J. and Henrichs, T.: Critical regions: A model-based estimation of world water resources sensitive to global changes, *Aquat. Sci.*, 64, 352–362, 2002.
- Alkama, R., Decharme, B., Douville, H., Becker, M., Cazenave, A., Sheffield, J., Voltaire, A., Tyteca, S., and Le Moigne, P.: Global evaluation of the ISBA-TRIP continental hydrological system. Part I: Comparison to GRACE terrestrial water storage estimates and in situ river discharges, *J. Hydrometeorol.*, 11, 583–600, 2010.
- Alley, W. M.: On the treatment of evapotranspiration, soil moisture accounting and aquifer recharge in monthly water balance models, *Water Resour. Res.*, 20, 1137–1149, 1984.
- Arnell, N. W. and Gosling, S. N.: The impacts of climate change on river flow regimes at the global scale, *J. Hydrol.*, 486, 351–364, 2013.
- Asadi, A.: The Comparison of Lumped and Distributed Models for Estimating Flood Hydrograph (Study Area: Kabkian Basin), *Journal of Electronics and Communication Engineering Research*, 1, 7–13, 2013.
- Bai, P., Liu, X., Liang, K., and Liu, C.: Comparison of performance of twelve monthly water balance models in different climatic catchments of China, *J. Hydrol.*, 529, 1030–1040, 2015.
- Batjes, N. H.: A homogenized soil data file for global environmental research: A subset of FAO, ISRIC and NRCS profiles (Version 1.0), ISRIC, Wageningen, The Netherlands, 1995.
- Beck, H. E., Dijk, A. I., Miralles, D. G., Jeu, R. A., McVicar, T. R., and Schellekens, J.: Global patterns in base flow index and recession based on streamflow observations from 3394 catchments, *Water Resour. Res.*, 49, 7843–7863, 2013.
- Best, M. J., Abramowitz, G., Johnson, H. R., Pitman, A. J., Balsamo, G., Boone, A., Cuntz, M., Decharme, B., Dirmeyer, P. A., Dong, J., Ek, M., Guo, Z., Haverd, V., van den Hurk, B. J., Nearing, G. S., Pak, B., Peters-Lidard, C., Santanello, J. A., Stevens, L., and Vuichard, N.: The plumbing of land surface models: benchmarking model performance, *J. Hydrometeorol.*, 16, 1425–1442, 2015.
- Bierkens, M. and Van Beek, L.: Seasonal predictability of European discharge: NAO and hydrological response time, *J. Hydrometeorol.*, 10, 953–968, 2009.
- Brirhet, H. and Benaabidate, L.: Comparison Of Two Hydrological Models (Lumped And Distributed) Over A Pilot Area Of The Issen Watershed In The Souss Basin, Morocco, *Eur. Sci. J.*, 12, 18, <https://doi.org/10.19044/esj.2016.v12n18p347>, 2016.
- Burnash, R. J., Ferral, R. L., and McGuire, R. A.: A generalized streamflow simulation system, conceptual modeling for digital computers, US Department of Commerce, National Weather Service, and State of California, Department of Water Resources, Sacramento, CA, 1973.
- Cosby, B., Hornberger, G., Clapp, R., and Ginn, T.: A statistical exploration of the relationships of soil moisture characteristics to the physical properties of soils, *Water Resour. Res.*, 20, 682–690, 1984.
- Dai, A., Qian, T., Trenberth, K. E., and Milliman, J. D.: Changes in continental freshwater discharge from 1948 to 2004, *J. Climate*, 22, 2773–2792, 2009.
- De Graaf, I., Van Beek, L., Wada, Y., and Bierkens, M.: Dynamic attribution of global water demand to surface water and groundwater resources: Effects of abstractions and return flows on river discharges, *Adv. Water Resour.*, 64, 21–33, 2014.
- Deb, K., Pratap, A., Agarwal, S., and Meyarivan, T.: A fast and elitist multiobjective genetic algorithm: NSGA-II, *IEEE T. Evolut. Comput.*, 6, 182–197, 2002.
- Döll, P., Fiedler, K., and Zhang, J.: Global-scale analysis of river flow alterations due to water withdrawals and reservoirs, *Hydrol. Earth Syst. Sci.*, 13, 2413–2432, <https://doi.org/10.5194/hess-13-2413-2009>, 2009.
- Downer, C. W., Ogdin, F. L., Martin, W. D., and Harmon, R. S.: Theory, development, and applicability of the surface water hydrologic model CASC2D, *Hydrol. Process.*, 16, 255–275, 2002.
- Edmonds, J., Wise, M., Pitcher, H., Richels, R., Wigley, T., and Maccracken, C.: An integrated assessment of climate change and the accelerated introduction of advanced energy technologies-an application of MiniCAM 1.0, *Mitig. Adapt. Strat. Gl.*, 1, 311–339, 1997.



- FAO: Digital soil map of the world and derived soil properties, in: Organization, F. A. (Ed.). Land and 410 Water Digital Media Series 1, FAO, Rome, Italy, 1998.
- Fekete, B. and Vorosmarty, C.: ISLSCP II UNH/GRDC Composite Monthly Runoff. ISLSCP Initiative II Collection, edited by: Hall, FG, Collatz, G., Meeson, B., Los, S., Brown de Colstoun, E., and Landis, D., Data Set, available at: <http://daac.ornl.gov/>, from Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridge, Tennessee, USA, 2011.
- Fekete, B. M., Vörösmarty, C. J., and Grabs, W.: High-resolution fields of global runoff combining observed river discharge and simulated water balances, *Global Biogeochem. Cy.*, 16, 15–1–15–10, 2002.
- Fernandez, W., Vogel, R., and Sankarasubramanian, A.: Regional calibration of a watershed model, *Hydrolog. Sci. J.*, 45, 689–707, 2000.
- Gerten, D., Hoff, H., Bondeau, A., Lucht, W., Smith, P., and Zaehle, S.: Contemporary “green” water flows: Simulations with a dynamic global vegetation and water balance model, *Phys. Chem. Earth, Parts A/B/C*, 30, 334–338, 2005.
- Ghavidelfar, S., Alvankar, S. R., and Razmkhah, A.: Comparison of the lumped and quasi-distributed Clark runoff models in simulating flood hydrographs on a semi-arid watershed, *Water Resour. Manage.*, 25, 1775–1790, 2011.
- GRDC: BfG The GRDC – Global Runoff Database, available at: [http://www.bafg.de/GRDC/EN/01\\_GRDC/13\\_dtbse/database\\_node.html](http://www.bafg.de/GRDC/EN/01_GRDC/13_dtbse/database_node.html), last access: 13 September 2017.
- 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, 80–91, 2009.
- Gupta, H. V., Sorooshian, S., and Yapo, P. O.: Toward improved calibration of hydrologic models: Multiple and noncommensurable measures of information, *Water Resour. Res.*, 34, 751–763, 1998.
- Haddeland, I., Heinke, J., Biemans, H., Eisner, S., Flörke, M., Hanasaki, N., Konzmann, M., Ludwig, F., Masaki, Y., Schewe, J., and Stacke, T.: Global water resources affected by human interventions and climate change, *P. Natl. Acad. Sci. USA*, 111, 3251–3256, 2014.
- Hanasaki, N., Kanae, S., Oki, T., Masuda, K., Motoya, K., Shirakawa, N., Shen, Y., and Tanaka, K.: An integrated model for the assessment of global water resources – Part 2: Applications and assessments, *Hydrol. Earth Syst. Sci.*, 12, 1027–1037, <https://doi.org/10.5194/hess-12-1027-2008>, 2008.
- Hansen, M., DeFries, R., Townshend, J. R., and Sohlberg, R.: Global land cover classification at 1 km spatial resolution using a classification tree approach, *Int. J. Remote Sens.*, 21, 1331–1364, 2000.
- Hargreaves, G. H. and Samani, Z. A.: Estimating potential evapotranspiration, *J. Irrig. Drain. Div.*, 108, 225–230, 1982.
- Hattermann, F. F., Krysanova, V., Gosling, S. N., Dankers, R., Dagupati, P., Donnelly, C., Flörke, M., Huang, S., Motovilov, Y., Buda, S., and Yang, T.: Cross-scale intercomparison of climate change impacts simulated by regional and global hydrological models in eleven large river basins, *Climatic Change*, 141, 561–576, 2017.
- Hsu, K., Gupta, H. V., and Sorooshian, S.: Artificial neural network modeling of the rainfall–runoff process, *Water Resour. Res.*, 31, 2517–2530, 1995.
- Jothityangkoon, C., Sivapalan, M., and Farmer, D.: Process controls of water balance variability in a large semi-arid catchment: Downward approach to hydrological model development, *J. Hydrol.*, 254, 174–198, 2001.
- Kim, S. H., Edmonds, J., Lurz, J., Smith, S. J., and Wise, M.: The O bj ECTS framework for integrated assessment: Hybrid modeling of transportation, *Energ. J.*, 27, 63–91, 2006.
- Kim, S. H., Hejazi, M., Liu, L., Calvin, K., Clarke, L., Edmonds, J., Kyle, P., Patel, P., Wise, M., and Davies, E.: Balancing global water availability and use at basin scale in an integrated assessment model, *Climatic Change*, 136, 217–231, 2016.
- Kraucunas, I., Clarke, L., Dirks, J., Hathaway, J., Hejazi, M., Hibbard, K., Huang, M., Jin, C., Kintner-Meyer, M., Kleese van Dam, K., Leung, R., Li, H.-Y., Moss, R., Peterson, M., Rice, J., Scott, M., Thomson, A., Voisin, N., and West, T.: Investigating the nexus of climate, energy, water, and land at decision-relevant scales: the Platform for Regional Integrated Modeling and Analysis (PRIMA), *Clim. Change*, 129, 573–588, 2015.
- Krysanova, V. and Hattermann, F. F.: Intercomparison of climate change impacts in 12 large river basins: overview of methods and summary of results, *Clim. Change*, 141, 363–379, 2017.
- Leng, G., Tang, Q., and Rayburg, S.: Climate change impacts on meteorological, agricultural and hydrological droughts in China, *Global Planet. Change*, 126, 23–34, 2015.
- Li, X., Vernon, C. R., Hejazi, M. I., Link, R. P., Feng, L., Liu, Y., and Rauchenstein, L. T.: Xanthos – A Global Hydrologic Model, *Journal of Open Research Software*, 5, 1–7, <https://doi.org/10.5334/jors.181>, 2017.
- Liu, Y., Pan, Z., Zhuang, Q., Miralles, D. G., Teuling, A. J., Zhang, T., An, P., Dong, Z., Zhang, J., He, D., Wang, L., Pan, X., Bai, W., and Niyogi, D.: Agriculture intensifies soil moisture decline in Northern China, *Sci. Rep.*, 5, 11261, <https://doi.org/10.1038/srep11261>, 2015.
- Liu, Y., Zhuang, Q., Chen, M., Pan, Z., Tchebakova, N., Sokolov, A., Kicklighter, D., Melillo, J., Sirin, A., Zhou, G., He, Y., Chen, J., Bowling, L., Miralles, D., and Parfenova, E.: Response of evapotranspiration and water availability to changing climate and land cover on the Mongolian Plateau during the 21st century, *Global Planet. Change*, 108, 85–99, 2013.
- Liu, Y., Zhuang, Q., Pan, Z., Miralles, D., Tchebakova, N., Kicklighter, D., Chen, J., Sirin, A., He, Y., Zhou, G., and Melillo, J.: Response of evapotranspiration and water availability to the changing climate in Northern Eurasia, *Climatic Change*, 126, 413–427, 2014.
- Martinez, G. F. and Gupta, H. V.: Toward improved identification of hydrological models: A diagnostic evaluation of the “abcd” monthly water balance model for the conterminous United States, *Water Resour. Res.*, 46, W08507, <https://doi.org/10.1029/2009WR008294>, 2010.
- Maurer, E. P., O’Donnell, G. M., Lettenmaier, D. P., and Roads, J. O.: Evaluation of the land surface water budget in NCEP/NCAR and NCEP/DOE reanalyses using an off-line hydrologic model, *J. Geophys. Res.-Atmos.*, 106, 17841–17862, 2001.
- Michaud, J. and Sorooshian, S.: Comparison of simple versus complex distributed runoff models on a midsized semiarid watershed, *Water Resour. Res.*, 30, 593–605, 1994.
- Mueller, B., Hirschi, M., Jimenez, C., Ciais, P., Dirmeyer, P. A., Dolman, A. J., Fisher, J. B., Jung, M., Ludwig, F., Maignan,

- F., Miralles, D. G., McCabe, M. F., Reichstein, M., Sheffield, J., Wang, K., Wood, E. F., Zhang, Y., and Seneviratne, S. I.: Benchmark products for land evapotranspiration: LandFlux-EVAL multi-data set synthesis, *Hydrol. Earth Syst. Sci.*, 17, 3707–3720, <https://doi.org/10.5194/hess-17-3707-2013>, 2013.
- Murphy, J. M., Sexton, D. M., Barnett, D. N., Jones, G. S., Webb, M. J., Collins, M., and Stainforth, D. A.: Quantification of modelling uncertainties in a large ensemble of climate change simulations, *Nature*, 430, 768–772, 2004.
- Nijssen, B., Lettenmaier, D. P., Liang, X., Wetzel, S. W., and Wood, E. F.: Streamflow simulation for continental-scale river basins, *Water Resour. Res.*, 33, 711–724, 1997.
- Nijssen, B., O'Donnell, G. M., Hamlet, A. F., and Lettenmaier, D. P.: Hydrologic sensitivity of global rivers to climate change, *Climatic Change*, 50, 143–175, 2001a.
- Nijssen, B., O'Donnell, G. M., Lettenmaier, D. P., Lohmann, D., and Wood, E. F.: Predicting the discharge of global rivers, *J. Climate*, 14, 3307–3323, 2001b.
- O'Neill, B. C., Krieger, E., Riahi, K., Ebi, K. L., Hallegatte, S., Carter, T. R., Mathur, R., and van Vuuren, D. P.: A new scenario framework for climate change research: the concept of shared socioeconomic pathways, *Climatic Change*, 122, 387–400, 2014.
- Obled, C., Wendling, J., and Beven, K.: The sensitivity of hydrological models to spatial rainfall patterns: an evaluation using observed data, *J. Hydrol.*, 159, 305–333, 1994.
- 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. Sci.*, 18, 67–84, <https://doi.org/10.5194/hess-18-67-2014>, 2014.
- Paudel, M., Nelson, E. J., Downer, C. W., and Hotchkiss, R.: Comparing the capability of distributed and lumped hydrologic models for analyzing the effects of land use change, *J. Hydroinform.*, 13, 461–473, 2011.
- Reed, S., Koren, V., Smith, M., Zhang, Z., Moreda, F., Seo, D. J., and Participants, D. M. I. P.: Overall distributed model intercomparison project results, *J. Hydrol.*, 298, 27–60, 2004.
- Refsgaard, J. C. and Knudsen, J.: Operational validation and intercomparison of different types of hydrological models, *Water Resour. Res.*, 32, 2189–2202, 1996.
- Rost, S., Gerten, D., and Heyder, U.: Human alterations of the terrestrial water cycle through land management, *Adv. Geosci.*, 18, 43–50, <https://doi.org/10.5194/adgeo-18-43-2008>, 2008.
- Sankarasubramanian, A. and Vogel, R. M.: Annual hydroclimatology of the United States, *Water Resour. Res.*, 38, 19-1–19-12, <https://doi.org/doi:10.1029/2001WR000619>, 2002.
- Sankarasubramanian, A. and Vogel, R. M.: Hydroclimatology of the continental United States, *Geophys. Res. Lett.*, 30, 1363, <https://doi.org/10.1029/2002GL015937>, 2003.
- Schaefli, B. and Gupta, H. V.: Do Nash values have value?, *Hydrol. Process.*, 21, 2075–2080, 2007.
- Scott, M. J., Daly, D. S., Hejazi, M. I., Kyle, G. P., Liu, L., McJeon, H. C., Mundra, A., Patel, P. L., Rice, J. S., and Voisin, N.: Sensitivity of future U.S. Water shortages to socioeconomic and climate drivers: a case study in Georgia using an integrated human-earth system modeling framework, *Climatic Change*, 136, 233–246, 2016.
- Sivapalan, M., Blöschl, G., Zhang, L., and Vertessy, R.: Downward approach to hydrological prediction, *Hydrol. Process.*, 17, 2101–2111, 2003.
- Tang, Q., Gao, H., Yeh, P., Oki, T., Su, F., and Lettenmaier, D. P.: Dynamics of terrestrial water storage change from satellite and surface observations and modeling, *J. Hydrometeorol.*, 11, 156–170, 2010.
- Tebaldi, C. and Knutti, R.: The use of the multi-model ensemble in probabilistic climate projections, *Philos. T. R. Soc. A*, 365, 2053–2075, 2007.
- Thomas, H.: Improved methods for national water assessment, Report WR15249270, US Water Resource Council, Washington, DC, 1981.
- Van Beek, L. and Bierkens, M. F.: The global hydrological model PCR-GLOBWB: conceptualization, parameterization and verification, Utrecht University, Utrecht, the Netherlands, 2009.
- Vandewiele, G. and Xu, C.-Y.: Methodology and comparative study of monthly water balance models in Belgium, China and Burma, *J. Hydrol.*, 134, 315–347, 1992.
- Vogel, R. M. and Sankarasubramanian, A.: Validation of a watershed model without calibration, *Water Resour. Res.*, 39, 1292, <https://doi.org/10.1029/2002WR001940>, 2003.
- Vörösmarty, C. J., Federer, C. A., and Schloss, A. L.: Potential evaporation functions compared on US watersheds: Possible implications for global-scale water balance and terrestrial ecosystem modeling, *J. Hydrol.*, 207, 147–169, 1998.
- Vörösmarty, C. J., Green, P., Salisbury, J., and Lammers, R. B.: Global water resources: vulnerability from climate change and population growth, *Science*, 289, 284–288, 2000.
- Wang, D. and Tang, Y.: A one-parameter Budyko model for water balance captures emergent behavior in darwinian hydrologic models, *Geophys. Res. Lett.*, 41, 4569–4577, 2014.
- Weedon, G. P., Gomes, S., Viterbo, P., Shuttleworth, W. J., Blyth, E., Österle, H., Adam, J. C., Bellouin, N., Boucher, O., and Best, M.: Creation of the WATCH forcing data and its use to assess global and regional reference crop evaporation over land during the twentieth century, *J. Hydrometeorol.*, 12, 823–848, 2011.
- Wen, L., Nagabhatla, N., Lü, S., and Wang, S.-Y.: Impact of rain snow threshold temperature on snow depth simulation in land surface and regional atmospheric models, *Adv. Atmos. Sci.*, 30, 1449–1460, 2013.
- Widén-Nilsson, E., Halldin, S., and Xu, C.-Y.: Global water-balance modelling with WASMOD-M: Parameter estimation and regionalisation, *J. Hydrol.*, 340, 105–118, 2007.
- Wilkinson, K., von Zubern, M., and Scherzer, J.: Global Freshwater Fluxes into the World Oceans, Tech. Report prepared for the GRDC, Koblenz, Federal Institute of Hydrology (BfG), GRDC Report No. 44, [https://doi.org/10.5675/GRDC\\_Report\\_44](https://doi.org/10.5675/GRDC_Report_44), 23 pp., available at: [http://www.bafg.de/GRDC/EN/02\\_srvcs/24\\_rprttrs/report\\_44.pdf](http://www.bafg.de/GRDC/EN/02_srvcs/24_rprttrs/report_44.pdf) (last access: 10 October 2016), 2014.
- Young, P. C. and Parkinson, S.: Simplicity out of complexity, in *Environmental Foresight and Models: A Manifesto*, edited by: Beck, M. B., Elsevier Science, the Netherlands, 251–294, 2002.
- Yao, H., Hashino, M., Terakawa, A., and Suzuki, T.: Comparison of distributed and lumped hydrological models, *Proceedings of Hydraulic Engineering*, 42, 163–168, 1998.
- Zhang, X.-J., Tang, Q., Pan, M., and Tang, Y.: A long-term land surface hydrologic fluxes and states dataset for China, *J. Hydrometeorol.*, 15, 2067–2084, 2014.