# Calibration of the US Geological Survey National Hydrologic Model in Ungauged Basins Using Statistical At-Site Streamflow Simulations

William H. Farmer[1]; Jacob H. LaFontaine[2]; and Lauren E. Hay[3]

**Abstract:** In the absence of measured streamflow, statistically simulated daily streamflow can be used to support the ability of physical models to represent hydrologic processes at ungauged locations. This study determined the feasibility of using statistical simulations in place of measured streamflow to calibrate physical models in ungauged basins. Daily streamflow was simulated at each of the 1,410 gauged watersheds using a cross-validated implementation of pooled ordinary kriging (POK). In this manner, the streamflow at each gauge was simulated as if no at-site streamflow information were available. The National Hydrologic Model application of the Precipitation-Runoff Modeling System was then calibrated through two separate procedures: (1) with measured streamflow, and (2) with statistically simulated streamflow in lieu of measured streamflow. Calibrating with statistically simulated streamflow produced performance within 23% of the performance of applications with knowledge of at-site measurements. Furthermore, statistically generated streamflow produced accurate timing information, which, when combined with alternative data sets (e.g., evapotranspiration, recharge, and so forth), can be used to improve representation of hydrologic processes at ungauged locations. **DOI:** 10.1061/(ASCE)HE.1943-5584.0001854. *© 2019 American Society of Civil Engineers.*

**Author keywords:** Prediction in ungauged basins; Physical hydrologic models; Statistical hydrologic models.

## Introduction

A reliable understanding of hydrologic processes, especially streamflow, is essential for resource managers. However, the characterization of hydrologic processes in areas without long-term monitoring remains one of the fundamental challenges of hydrologic sciences. When it comes to the consideration of streamflow, continental-scale data networks such as the USGS National Streamflow Information Program (NSIP) and National Water Information System (NWIS) (USGS 2017) provide the backbone of data for water resources management in the United States. Unfortunately, not every stream has a stream gauge and decisions about water resources may need to be made in a watershed in which no streamflow information is available. Approximately 74% of the land in the conterminous United States (CONUS) drains to an active stream gauge, with 13% of these areas being unaffected by anthropogenic effects (Kiang et al. 2013). This leaves large tracts of the United States without data with which to characterize hydrologic processes, a problem not unfamiliar across the globe. The International Association of Hydrologic Sciences committed a decade to advancing knowledge in this field under the heading of prediction in ungauged basins (PUB) (Sivapalan 2003). Throughout this decade, as in the previous decades of work, the problem of hydrologic characterization has been largely addressed with either statistical or physical modeling (Farmer 2016).

The ungauged problem considers how to provide hydrologic information to stakeholders when the particular point in space and time that is of interest is not monitored. In reference to streamflow, a number of methods have been proposed; however, no one method has been universally accepted or demonstrated to work in all environments (Blöschl et al. 2013, Tables A7–A10). Two approaches, which are best conceived as ends of a continuum, have been proposed to tackle PUB: physical and statistical modeling. The former relies on a physical understanding of hydrologic processes, whereas the latter leverages historical correlations.

On one end of this continuum, physically based hydrologic models, commonly called rainfall-runoff, process-based, or deterministic models, are simplified representations of the physical processes driving hydrology (Bicknell et al. 2005; Markstrom et al. 2015; Neitsch et al. 2005; Kendy et al. 2011). The appeal of physical models is that they provide some degree of methodological understanding of hydrologic processes that can be leveraged for scenario analysis. However, in the absence of measured streamflow information for calibration, physical models are difficult to calibrate to at-site information. Without at-site information, various suggestions have been made: regionalizing parameter values (Sawicz et al. 2011; Bock et al. 2016), linking parameters to basin physiography (Merz and Blöschl 2004; Zhang et al. 2008; Seibert 1999; Peel et al. 2000; Oudin et al. 2008; Oudin et al. 2010; Samuel et al. 2011), defining measures of basin similarity (Tegegne and Kim 2018), and building models that do not require calibration (Bergstrom 2006; Shi et al. 2008; Rozalis et al. 2010). All have been met with mixed reviews, and have advantages and disadvantages depending on the motivation of the application, but most remain unsatisfying. Transferring information from gauged to ungauged locations remains challenging. Issues of data scarcity mean that, for the use of physical models, the hydrologic community needs to develop a better understanding of the value of various types of data available

[1]Water Mission Area, US Geological Survey, Denver, CO 80226 (corresponding author). ORCID: https://orcid.org/0000-0002-2865-2196. Email: wfarmer@usgs.gov

[2]South Atlantic Water Science Center, US Geological Survey, Norcross, GA 30093. ORCID: https://orcid.org/0000-0003-4923-2630

[3]Water Mission Area, US Geological Survey, Denver, CO 80226.

Note. This manuscript was submitted on September 26, 2018; approved on July 10, 2019; published online on September 14, 2019. Discussion period open until February 14, 2020; separate discussions must be submitted for individual papers. This paper is part of the ***Journal of Hydrologic Engineering***, © ASCE, ISSN 1084-0699.

and how these data can be pooled into large sample hydrology (Andréassian et al. 2006; Gupta et al. 2012; Hrachowitz et al. 2013).

On the other end of this continuum of hydrologic representation, statistical hydrologic models leverage regional information and historical correlations to simulate hydrologic signatures. Although the results can be accurate, particularly when specific hydrologically relevant metrics are targeted without concern for other hydrologic processes, statistical models do not always provide a robust understanding of physical processes. Examples of such models are those that simulate streamflow characteristics (Worland et al. 2018; Archfield et al. 2013a; Bassiouni et al. 2016), streamflow signatures such as complete streamflow duration curves (Pugliese et al. 2016; Blum et al. 2017; Farmer et al. 2014b; Atieh et al. 2015; Atieh et al. 2017; Reichl and Hack 2017), and hydrographs (Farmer 2016; Archfield et al. 2013b; Farmer et al. 2014a, 2015, 2018; Farmer and Koltun 2017). Because statistical methods lack a methodological foundation of physical processes, it can be difficult to use statistical models for extrapolation beyond the space-time region for which they were developed.

It is perhaps overly simplistic to classify hydrologic characterization into physical and statistical methods. There are, of course, many statistical assumptions that go into the development of any physical model, from the mathematical representation of processes to the calibration of parameters. Similarly, the best statistical approaches rely on physically meaningful correlations rather than arbitrary or hydrologically meaningless correlations. All methods of hydrologic characterization, especially at ungauged locations, rely on some fusion of physical and statistical methods.

This work proposes a novel technique combining the advantages of both approaches to produce well-calibrated physical models suitable for historical and future long-range (decadal) simulation at ungauged locations. This work hypothesizes that statistical methods of daily hydrograph simulation can be used to inform at-site calibrations of physical models to allow for process exploration and daily hydrograph simulation at ungauged locations. Because large regions of the globe remain unmonitored, it is difficult to develop well-calibrated physical models with at-site information that is unavailable without monitoring. Statistical methods are easier to regionalize to ungauged locations, but lack strong process representations. By using statistical methods as a first-pass simulation of historical hydrographs upon which to calibrate a physical model, this work provides a novel combination of physical and statistical methods for historical, daily hydrograph simulation. The objective of this study is to quantify the performance that can be expected when calibrating a physically based hydrologic model to a statistically simulated hydrograph.

The USGS National Hydrologic Model application of the Precipitation-Runoff Modeling System (NHM-PRMS) (Regan et al. 2018) was used to simulate streamflow in 1,410 gauged watersheds across the CONUS to test the feasibility of improving streamflow simulations by linking statistical and physically based hydrologic models. The NHM-PRMS was calibrated using measured streamflow (PRMS-M) and daily hydrographs from pooled ordinary kriging (PRMS-POK) (Farmer 2016) at 1,410 gauged watersheds across the CONUS. The intent was not to justify these specific models, but rather to show how the approach of using simulated at-site calibration information can be used across the country, particularly in ungauged basins. Further research could explore the sensitivity of choosing other statistical or process-based models.

This approach is similar to, but fundamentally different from, the historical practice of calibrating a physical model to reconstructed reservoir inflows. In the case of reservoir inflows, proxies exist for accurately reconstructing inflows (e.g., reservoir levels, reservoir releases, and reservoir gauges). In the case presented here, there are no at-site observations on which to rely. Furthermore, this work considers a continental-extent, automated calibration procedure. This is not the basin-by-basin procedure adopted for reservoir reconstructions. For these reasons, and others presented throughout (Sivapalan 2003; Blöschl et al. 2013), this work represents a novel and meaningful contribution to the literature. Through this demonstration, it is possible to begin to ask questions about how ungauged calibration can be improved (e.g., by including uncertainty in simulations and climate or other elements of the hydrologic cycle).

## Study Area

This study consisted of 1,410 gauged watersheds distributed across the CONUS, with the majority of watersheds located in the east and near the west coast (Fig. 1). The gauged watersheds were selected from the 1,575 watersheds chosen by Bock et al. (2016), and this set was further restricted by considering only sites for which the monthly water balance model produced positive Nash–Sutcliffe model efficiencies (Nash and Sutcliffe 1970; Bock et al. 2016) and sites considered reference quality, as defined by Falcone (2011) as having limited hydrologic alteration and management. Stream gauge coverage is less dense in the central United States, with stream gauges generally having larger drainage areas compared with the those in the rest of the CONUS. Study watersheds are present in every major hydrologic unit across the CONUS and range across hydroclimatic regimes and basin size. Drainage areas for the study watersheds ranged from 4 to 7,680 $km^2$ (median size 469 $km^2$). Elevations for the outlets of the study watersheds ranged from −0.8 to 3,170 m (median elevation 192 m) relative to the North American Vertical Datum of 1988. Annual precipitation accumulations for the study watersheds ranged from 200 to 4,000 mm (median 1,200 mm). Daily measured streamflow for the 1,410 stream gauges was retrieved from the NWIS (USGS 2017) using the dataRetrieval R script (Hirsch and DeCicco 2015) for the period from October 1, 2000 through September 30, 2010. All data used in this work are available in the associated data release (Farmer et al. 2019).

## Methodology

The ability of a physically based hydrologic model to simulate ungauged streamflow when provided with statistically simulated streamflow was evaluated through two calibration methodologies. First, the physically based hydrologic model was calibrated to the measured streamflow records described previously using an automated procedure. This is an idealized case in which complete knowledge of at-site information is available. Second, daily hydrographs were simulated using purely statistical methods at each stream gauge as if no streamflow information were available at that location. These statistically simulated hydrographs were then used to calibrate the physically based hydrologic model. This second methodology represents an at-site calibration without any at-site measurements of streamflow. Contrasting the performance of these two methodologies (complete measured streamflow versus complete lack of measured streamflow) allowed for an assessment of the potential fusion of statistical and physical streamflow modeling for ungauged locations. Although the methods used for simulation here are specific to POK and PRMS, as outlined in the following sections, the general principle could, with some discretion, be expanded for the use of any combination of models. The POK and PRMS methodologies are not presented as novel contributions and are not advocated as the major contribution of this work. For this reason, this section outlines the methodologies with heavy

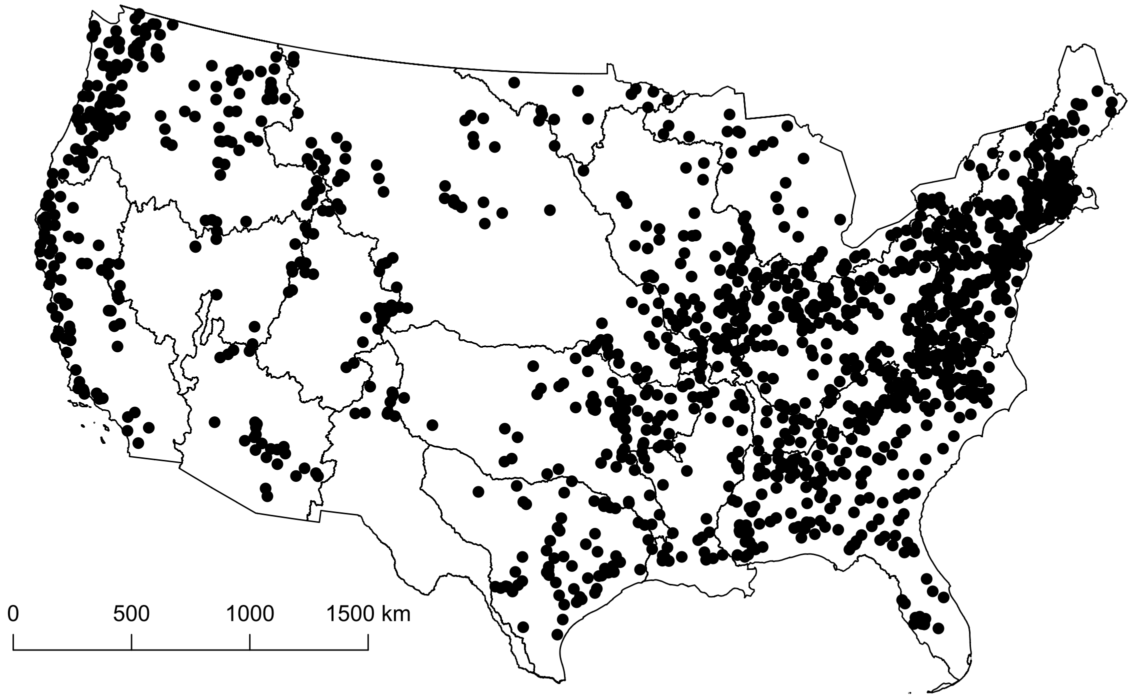


**Fig. 1.** Locations of 1,410 stream gauges used in this simulation and analysis.

references for the interested user. The novelty in the methodology presented here is the combination of physical and statistical simulations through calibration.

### Pooled Ordinary Kriging

Following the methodology of Farmer (2016), daily hydrographs were statistically simulated at each stream gauge using a cross-validated implementation of pooled ordinary kriging (POK). The streamflow at each stream gauge was simulated as if no at-site streamflow information were available. The POK procedure then simulated daily streamflow as a weighted sum of concurrent daily streamflow within the same two-digit hydrologic unit. The two-digit hydrologic units are large regions that roughly align with the major river basins of the United States and are commonly used for classification (Seaber et al. 1987).

The POK method proceeds by looking at the semivariance between concurrent daily streamflow measurements within a two-digit hydrologic unit. These semivariances are then pooled across all days of the streamflow records to estimate a single, time-invariant semivariogram linking the correlation structure between stream gauges to a function of separation distance. This approximation of the regional correlation structure is then used to, for a particular target estimation location, assign weights to each active stream gauge on the target date. These weights inform the weighted summation of concurrent streamflow used to simulate streamflow at each target location. To improve the distribution of simulated values and limit the impact of streamflow ranging over multiple orders of magnitude, all streamflow was converted by dividing by the drainage area and taking the natural logarithm. Zero-valued streamflow was augmented by 0.142 L/s (0.005 cfs) to avoid undefined logarithms in this analysis.

### Precipitation-Runoff Modeling System

For this study, the Precipitation-Runoff Modeling System was used to simulate streamflow for 1,410 watersheds across the CONUS. The 1,410 PRMS models were extracted from the USGS NHM-PRMS, which was developed to support coordinated, comprehensive, and consistent hydrologic model development and application across the CONUS (Regan et al. 2018, 2019). The PRMS (Markstrom et al. 2015) is a modular, deterministic, distributed-parameter; physical process–based hydrologic simulation code. It was developed to evaluate impacts of various combinations of climate, physical characteristics, and simulation options on hydrologic response and water distribution at the watershed scale. The PRMS computes water flow and storage from and to the atmosphere, plant canopy, land surface, snowpack, surface depressions, shallow subsurface zone, deep aquifers, stream segments, and lakes. Physical characteristics including topography, soils, vegetation, geology, and land use are used to characterize and derive parameters required in simulation algorithms, spatial discretization, and topological connectivity. Computations of the hydrologic processes are forced by historical, current, or potential future climate and land-cover data. Daily precipitation and minimum and maximum air temperature data are required for a PRMS simulation. Simulations operate on a daily time step with periods from days to centuries.

The PRMS modeling units [hydrologic response units (HRUs)], stream segments, and default parameters, for the 1,410 study watersheds were extracted from the Geospatial Fabric (GF) database (Viger and Bock 2014). The GF, created for the USGS NHM infrastructure, is a framework for CONUS-scale modeling derived from the National Hydrography Dataset Plus, version 1 (NHDPlus) (Bondelid et al. 2010). Catchments and flowlines from the NHDPlus were aggregated into HRUs and stream segments, respectively. The NHM-PRMS default parameters were documented by Regan et al. (2018) and archived by Driscoll et al. (2017).

The Daymet climate forcings (Thornton et al. 1997; Thornton and Running 1999; Thornton et al. 2000) were used to force the 1,410 subsetted PRMS models for this study. Daymet is a 1 × 1-km gridded, daily-time-step climate product for the CONUS and parts of Canada and Mexico for the period 1980–2015. These data were summarized for the 1,410 study watersheds using the USGS

GeoData Portal (Blodgett et al. 2011). Preprocessed climate data sets allow for CONUS-scale assessments to be completed with consistently developed climate forcings. In particular, daily simulations with the PRMS are driven by precipitation, minimum temperature, and maximum temperature.

The PRMS is a physically based model that requires calibration. Based on work by Markstrom et al. (2016), Table 1 lists the parameters that were calibrated as described subsequently. The PRMS work presented here used a combination of two approaches to estimate model parameters: (1) estimating initial values based on measurable characteristics when available, and (2) traditional model calibration with an objective function on streamflow. A key step in hydrologic model calibration is to have good initial estimates of model parameters. Where possible, the initial values for parameters listed in Table 1 were estimated using techniques that relate model parameters to land surface characteristics including climate, soils, vegetation and topography, as described by Regan et al. (2018). The initial model parameters that were estimated remain idealized representations; they require further calibration to ensure that their ultimate values reflect the landscape at the scale of the modeling units. Other parameters in Table 1 have an initial value for calibration that is constant for every HRU because there was no observable information to inform the initial parameter value. These parameters that start with nondistributed values are noted in Table 1.

The PRMS models for each of the 1,410 watersheds were calibrated with two separate streamflow data sets: (1) measured streamflow from USGS stream gauges, and (2) simulated streamflow developed with POK. The PRMS models were calibrated using Luca, a software package developed to carry out multiple-objective, stepwise, automated calibration for the PRMS (Hay et al. 2006; Hay and Umemoto 2006) using the shuffled complex evolution global search algorithm (Duan et al. 1992, 1993). The Luca software was set up to match simulated streamflow with either measured or POK-generated streamflow. Objective functions targeting the minimization of normalized RMSE (NRMSE) were used in the calibration procedure, as recommended by Hay et al. (2006), Hay and Umemoto (2006), and Hay et al. (2018). The first calibration uses measured streamflow for simulation optimization, congruent with how hydrologic models previously have been developed. The second calibration scenario uses statistically generated streamflow (POK), providing an avenue for the use of streamflow simulation as calibration targets in ungauged basins. By comparing PRMS simulations calibrated to measured streamflow and statistically generated streamflow, estimates of simulation confidence in areas without measured streamflow can be provided.

Both PRMS model calibrations use a four-step methodology that calibrates parameters to components of the daily hydrograph and other elements of the water balance. Both calibration procedures start with the same initial parameter file containing all PRMS parameters as extracted from the USGS NHM infrastructure. A sensitivity analysis based on the work by Markstrom et al. (2016) was used to identify which model parameters would be calibrated in each step (Table 1). These calibrated parameter values replace the respective parameter values in the parameter file; the updated parameter file becomes the initial parameter file for the next calibration step. Completion of the four calibration steps constitutes a round. Once a parameter is calibrated, it does not change for the remainder of that calibration round. This process is repeated until no further improvement in model performance is obtained—four rounds for each of the two calibration procedures, based on previous experience (Hay et al. 2006; Hay and Umemoto 2006; Hay et al. 2018).

The first calibration step targeted streamflow volume in PRMS, examining annual and monthly streamflow volumes. The second and third steps calibrated the parameters to indicators of hydrologic alteration environmental flow components (EFC) from The Nature

**Table 1.** List of parameters calibrated for this application of PRMS, including information on calibration procedures and results

| Parameter | Calibration step | Distributed | Correlation in changes |
|---|---|---|---|
| Temperature when precipitation is assumed to be rain [tmax_allrain] | Streamflow timing | No | 0.32 |
| Maximum possible percentage area contributing to surface runoff [carea_max] | Streamflow timing | No | 0.29 |
| Transmission coefficient for short-wave radiation through winter vegetation canopy [rad_trncf] | Streamflow timing | Yes | 0.63 |
| Exponent in nonlinear contributing area algorithm [smidx_exp] | High streamflow | No | 0.27 |
| Summer rain interception storage capacity for major vegetation type [srain_intcp] | High streamflow | Yes | 0.50 |
| Monthly adjustment factor to measured precipitation [snow_cbh_adj] | Water balance | No | 0.39 |
| Linear coefficient in equation to compute groundwater (GW) discharge [gwflow_coef] | Low streamflow | Yes | 0.48 |
| Fraction of potential evapotranspiration that is sublimated from snow in canopy and snowpack [potet_sublim] | Streamflow timing | No | 0.03 |
| Maximum amount of capillary reservoir excess routed directly to GW reservoir [soil2gw_max] | Low streamflow | Yes | 0.57 |
| Monthly adjustment factor to measured precipitation [rain_cbh_adj] | Water balance | No | 0.34 |
| Monthly maximum air temperature when precipitation is assumed to be snow [tmax_allsnow] | Streamflow timing | No | 0.40 |
| Fraction of soil zone in which preferential flow occurs [pref_flow_den] | Streamflow timing | No | 0.74 |
| Coefficient in nonlinear contributing area algorithm [smidx_coef] | High streamflow | No | 0.23 |
| Monthly minimum air temperature adjustment factor [tmin_cbh_adj] | Streamflow timing | No | 0.55 |
| Nonlinear preferential-flow routing coefficient [fastcoef_lin] | Streamflow timing | Yes | 0.38 |
| Average emissivity of air on days without precipitation [emis_noppt] | Streamflow timing | No | 0.15 |
| Maximum storage for soil recharge zone [soil_rechr_max] | Streamflow timing | Yes | 0.46 |
| Monthly maximum air temperature adjustment factor [tmax_cbh_adj] | Streamflow timing | No | 0.58 |
| Nonlinear coefficient in equation to route gravity-reservoir storage down slope [slowcoef_sq] | Streamflow timing | No | 0.42 |
| Free-water holding capacity of snowpack [freeh2o_cap] | Streamflow timing | No | 0.27 |
| Monthly convection condensation energy coefficient [cecn_coef] | Streamflow timing | No | 0.07 |
| Maximum available water holding capacity of capillary reservoir [soil_moist_max] | Streamflow timing | Yes | 0.56 |
| Temperature index to determine specific date of start of transpiration period [transp_tmax] | Streamflow timing | No | 0.03 |
| Maximum threshold snowpack water equivalent below which snow covered–area curve is applied. [snarea_thresh] | Streamflow timing | Yes | 0.44 |
| Maximum fraction of potential solar radiation that may reach ground [radmax] | Streamflow timing | No | 0.15 |

Conservancy (TNC 2009). The daily streamflow was split into high- and low-flow components based on the five EFC categories. In the second step, daily streamflow that was considered part of the high-flow regime, categorized as a high-flow pulse, small flood or large flood, was calibrated. In the third step, streamflow considered part of the low-flow regime, categorized as low flow or extreme low flow, was calibrated. The fourth step calibrated parameters to the complete daily streamflow record.

### Analysis and Evaluation

Accuracy of each method of hydrograph simulation was assessed using the normalized root mean squared error (NRMSE)

$$\mathrm{NRMSE} = \sqrt{\frac{\sum_{i=1}^{n}(Q_{o,i} - Q_{s,i})^2}{\sum_{i=1}^{n}(Q_{o,i} - \bar{Q}_o)^2}} \tag{1}$$

where $Q_o$ = measured streamflow; $Q_s$ = simulated streamflow; $\bar{Q}_o$ = mean measured streamflow; and $n$ = number of time steps (days) in the record. In addition to the NRMSE, the NRMSE of the logarithms of streamflow, i.e., $\ln Q_o$ rather than $Q_o$, was used; this is denoted NRMSEL. When logarithms were taken of streamflow, a small value [0.028 L/s (0.001 cfs)] was added to each streamflow value to avoid the logarithm of zero values. As discussed subsequently, the use of logarithms weights the high and low values differently.

To compare the difference in overall performance of methods across sites, the NRMSE can be converted to the Nash–Sutcliffe model efficiency (NSE) (Nash and Sutcliffe 1970)

$$\mathrm{NSE} = 1 - \mathrm{NRMSE}^2 \tag{2}$$

In the same form that NSE is a function of NRMSE, the NSE of the logarithms of streamflow (NSEL) is a function of NRMSEL.

The agreement of any two methods was considered by calculating the percentage difference (PD) in NRMSE as

$$\mathrm{PD} = 200\frac{\mathrm{NRMSE}_1 - \mathrm{NRMSE}_2}{\mathrm{NRMSE}_1 + \mathrm{NRMSE}_2} \tag{3}$$

where the subscripts define the methods being compared. This same metric can be computed for the NRMSEL. The effect of one method over another, taking the latter to be an initial value, is the percentage change (PC) in the NRMSE or NRMSEL. The percentage change in the NRMSE is given as

$$\mathrm{PC} = 100\frac{\mathrm{NRMSE}_N - \mathrm{NRMSE}_R}{\mathrm{NRMSE}_R} \tag{4}$$

where the subscript $R$ denotes the NRMSE of the reference method and the subscript $N$ denotes the NRMSE of the novel method being considered. Again, the percentage change can be computed using the NRMSEL as well.

All data, including the streamflow observations and simulations used and the analysis conducted in this work, are available in the associated data release (Farmer et al. 2019).

## Results

The following sections compare streamflow from (1) POK, (2) PRMS models calibrated using measurements (PRMS-M), and (3) PRMS models calibrated using POK (PRMS-POK) with reference to streamflow measurements at 1,410 watersheds across the CONUS. Favorable comparisons between PRMS-M and PRMS-POK may determine where POK can be used to provide streamflow information to calibrate the PRMS in ungauged basins. The differences indicate a gain or loss of accuracy in simulating a particular hydrograph.

The PRMS-POK represents a fully ungauged application of the PRMS. The resulting hydrographs from 1,410 stream gauges yielded a median NSE of 0.56. The PRMS-POK performance was less than the median NSE of 0.73 produced by PRMS-M, but not substantially different than the median NSE of 0.58 produced by POK. Fig. 2 presents the distribution of performance of all three methods in terms of NSE. Boxplot whiskers in Fig. 2(d) extend to the 5th and 95th percentiles. The edges of the box identify the 25th and 75th percentiles. The solid line indicates the median value, and the asterisk indicates the mean. The PRMS-M methodology produced NSEs above 0.50 at more than 91% of sites [Fig. 2(d)]. This is not overly surprising, because the PRMS-M methodology has access to measured at-site information. Without any at-site measurements, the POK methodology was able to produce NSEs greater than 0.50 at more than 62% of sites across the country. When the POK methodology was used as a calibration target for PRMS, through the PRMS-POK methodology, more than 60% of sites yielded NSEs greater than 0.50. Figs. 2(a–c) show the correspondence of performance in these three methods. The NSEs produced by PRMS-M demonstrated a 0.21 Spearman rank correlation with the NSEs of POK; although the correlation was relatively weak, Fig. 2(a) shows that the methods tended to perform similarly well in some places. The Spearman rank correlation of PRMS-POK and POK was 0.80, showing that the PRMS-POK methodology strongly reproduced the underlying calibration target.

When logarithms were taken of streamflow, with a small additive adjustment to retain zero-valued measurements, the PRMS-POK produced a median NSEL of 0.50. This was less than the median performance of POK (0.67) and the PRMS-M (0.58). Fig. 3 captures the distribution of this performance for all three methods in terms of NSEL. Boxplot whiskers in Fig. 3(d) extend to the 5th and 95th percentiles. The edges of the box identify the 25th and 75th percentiles. The solid line indicates the median value, and the asterisk indicates the mean. Whereas the NSE calculation was affected more by errors in large streamflow, the NSEL was affected more by errors in low streamflow. Although the points in the scatterplot are much more dispersed, the Spearman rank correlation between the NSELs of PRMS-M and POK (0.35) was greater than the same correlation between NSEs.

Considering the overall NRMSE metric, the PRMS-M performed better than POK at 71% of sites, although the median percentage difference [Eq. (3)] was 18%, with 46% of the sites having an absolute percentage difference of less than 25%. When the errors in low streamflow were weighted more heavily, the NRMSEL tipped in favor of POK, with a median difference of −17%, with POK performing better at 74% of sites. The PRMS-M methodology, which used the measured hydrograph for calibration, performed better than POK, which had no knowledge of the measured hydrograph, when calculating performance without any transformations.

Removing all knowledge of at-site measurements, PRMS-POK produced an overall NRMSE within 23% (median value across all sites) of PRMS-M. PRMS-POK produced better NRMSEs than PRMS-M at 5.7% of sites. In terms of NRMSEL, the performance of PRMS-POK was within 5.2% (median value across all sites) of PRMS-M. The PRMS-POK methodology produced better NRMSELs than the PRMS-M methodology at 37% of sites. The correspondence of overall NRMSEs is shown in Fig. 4(d), whereas Fig. 5(d) shows the correspondence of overall NRMSELs. In both figures, each subplot shows the metric computed over a different

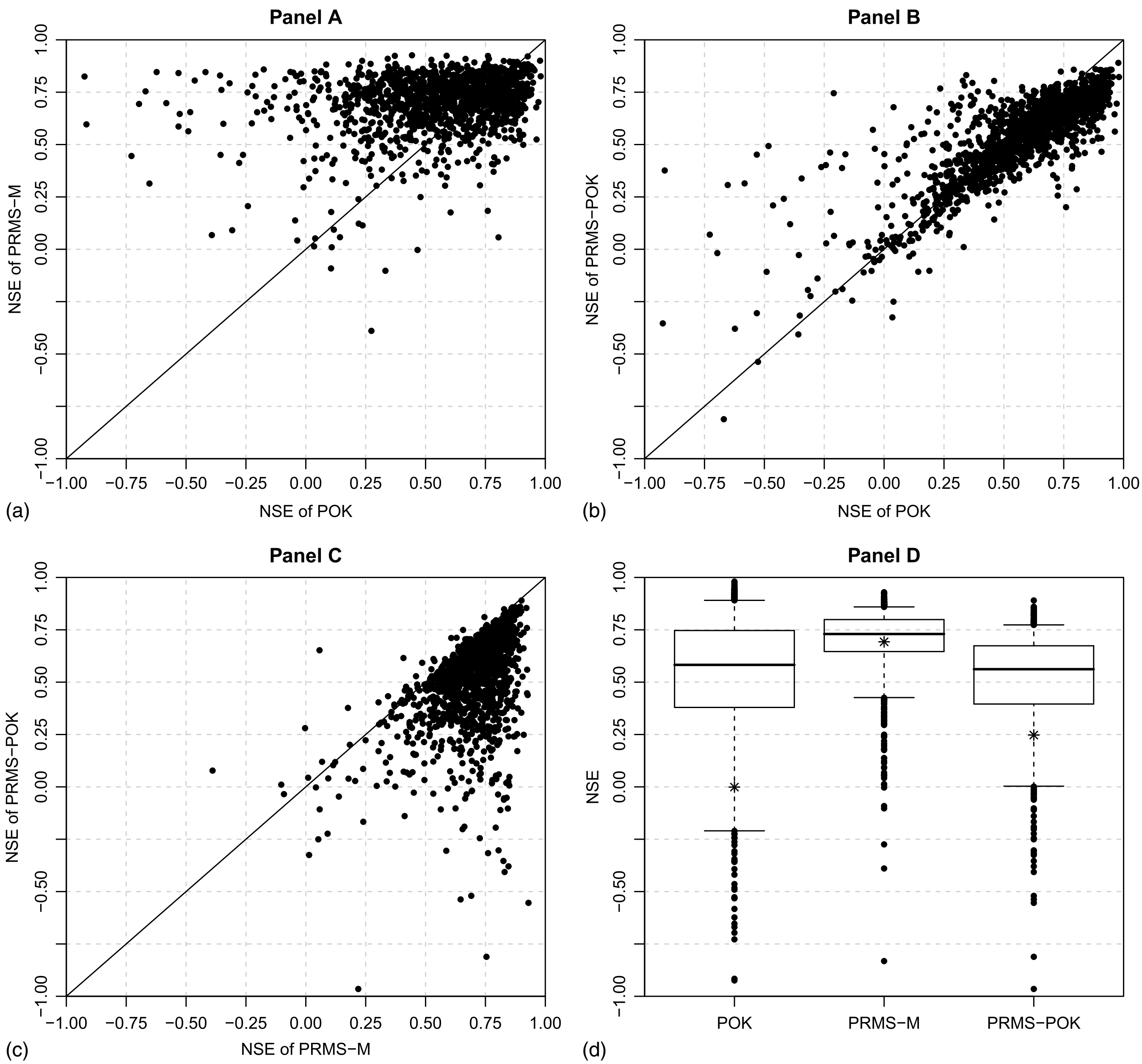


**Fig. 2.** Correspondence and distribution of the Nash–Sutcliffe efficiency of streamflow (NSE) at each of 1,410 stream gauges for each simulation method.

subset or averaging of streamflow values. Boxplot whiskers extend to the 5th and 95th percentiles. The edges of the box identify the 25th and 75th percentiles. The solid line indicates the median value and the asterisk indicates the mean.

The correspondence of NRMSEs and NRMSELs between PRMS-M and PRMS-POK is shown in Figs. 4 and 5, respectively, for monthly streamflow [Figs. 4(a) and 5(a)], low streamflow [Figs. 4(b) and 5(b)] and high streamflow [Figs. 4(c) and 5(c)]. The NRMSEs on monthly streamflow simulations produced by PRMS-POK were within 32% (median value across all sites) of those produced by PRMS-M, and the NRMSEs on high streamflow simulations were within 26% (median value across all sites). With a median effect of −9.6%, the PRMS-POK methodology improved the NRMSE of low streamflow simulations compared with PRMS-M. The NRMSELs (Fig. 5) show that the performance corresponded well, despite the wide scatter: Taking a median effect across all sites, the NRMSELs of PRMS-POK were within 5.3%, 2.0%, and 22% for monthly, low, and high streamflow, respectively, relative to PRMS-M.

The difference between the NRMSE and NRMSEL metrics is their balancing of errors in low and high streamflow. The logarithmic transformation will more equitably weight low and high streamflow. The large difference between the NRMSE and NRMSEL suggests that the NRMSE is sensitive to errors in high streamflow. This sensitivity is likely to affect the calibration procedure used in this study, allowing calibration to be dominated by high streamflow despite the fact that, on average, only 26% of daily streamflow was classified as high streamflow. However, the target of calibration should be taken under deep consideration, because those same 26% of days account for an average of 73% of the total volume of streamflow across sites. Although the merits of any particular calibration scheme can be debated at length, the sensitivity is controlled here because both methods, i.e., PRMS-M and PRMS-POK, use the same calibration scheme.

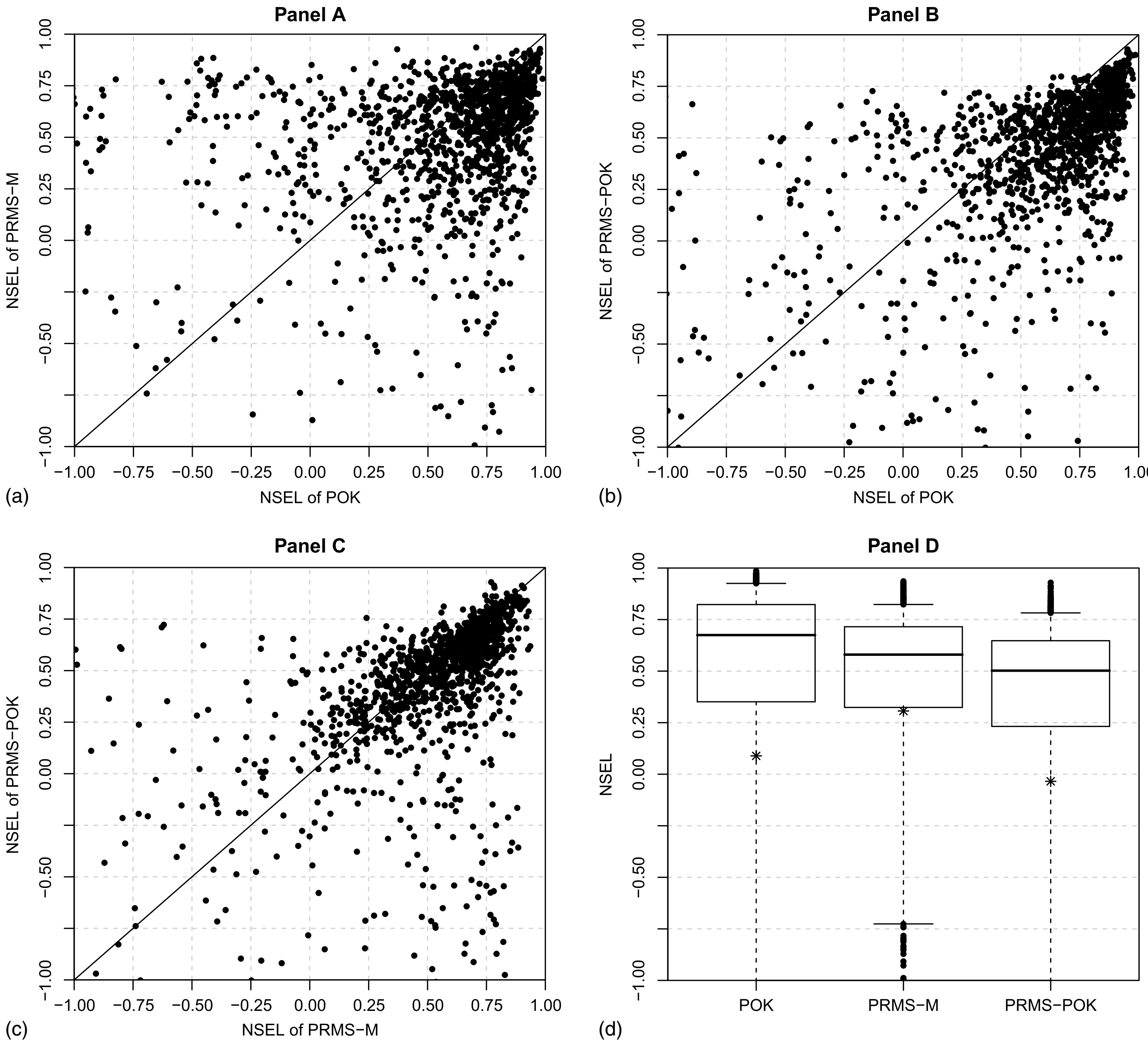


**Fig. 3.** Correspondence and distribution of the Nash–Sutcliffe efficiency of the natural logarithms streamflow (NSEL) at each of 1,410 stream gauges for each simulation method.

PRMS-POK improved the simulation of low streamflow compared with PRMS-M simulations [Fig. 4(b)]. Whereas few sites saw an improvement in overall NRMSE (5.7%) and NRMSEL (37%), a greater proportion of sites demonstrated improvements in the NRMSE (64%) [Fig. 4(b)] and NRMSEL (45%) [Fig. 5(b)] of low streamflow. Recent work (Andréassian et al. 2012; Poncelet et al. 2016) has shown that, in the presence of high streamflow, it is difficult to simulate both high and low streamflow equally well. The statistical time series, however, is likely to have a dampened representation of the high streamflow (Farmer 2016; Farmer and Vogel 2016). Because the PRMS calibration algorithm is not being as severely stretched in opposite directions when targeting statistically generated streamflow (PRMS-POK) than when targeting measured streamflow (PRMS-M), it may be possible to more accurately represent low streamflow using PRMS-POK.

The hypothesis that more-extreme differences between high and low streamflow limit the ability of a process-based model to reproduce both low and high streamflow magnitudes, although not fully explored, is supported by the initial exploration presented subsequently. The ratio of standard deviation (SD) of all streamflow in the measured record to the standard deviation of all streamflow in the statistically generated record indicates the differences in variability. If the magnitude of high streamflow were reduced relative to low streamflow, this ratio would be greater than one. Here, the median ratio was 1.35, with nearly 80% of sites having smaller standard deviations in the statistically generated time series. This ratio can be contrasted with the ratio of the low-streamflow NRMSE of PRMS-M to the low-streamflow NRMSE of PRMS-POK. This NRMSE ratio would be greater than 1 if the PRMS-POK methods were better able to reproduce measured low streamflow, as was true



J. Hydrol. Eng., 2019, 24(11): 04019049

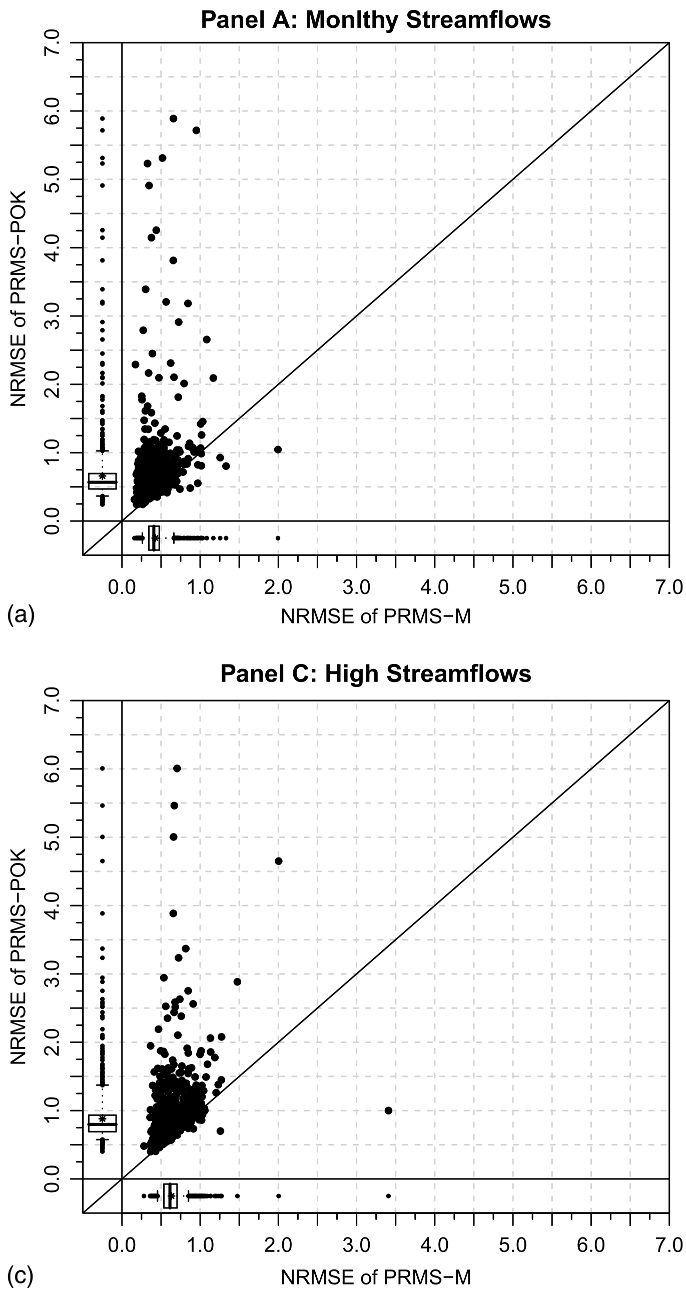


**Fig. 4.** Correspondence and distribution of the normalized RMSE of streamflow (NRMSE) at each of 1,410 stream gauges for the calibration of the Precipitation-Runoff Modeling System to measured streamflow (PRMS-M) and calibration to POK streamflow (PRMS-POK).

for 64% of sites used in this study (if there were little difference, one would expect a random 50% split). Fig. 6 plots these values against each other, showing a 0.44 Pearson correlation coefficient between the two ratios: because the statistically generated time series was less variable, the ability of PRMS-POK to reproduce measured low streamflow generally increased. Although not an exhaustive exploration, this provides evidence to motivate further research into parameter estimation and optimization algorithms and targets. The horizontal axis in Fig. 6 shows the ratio of the standard deviation of all measured streamflow to the standard deviation of streamflow simulated by pooled ordinary kriging. The vertical axis shows the ratio of the NRMSE of low, daily streamflow simulated by calibrating the Precipitation-Runoff Modeling System to measurements (numerator) and the same metric from calibrations with statistically generated streamflow (denominator).

## Discussion

The NHM-PRMS was calibrated using measured streamflow (PRMS-M) and daily hydrographs from pooled ordinary kriging (PRMS-POK) at 1,410 gauged watersheds across the CONUS. Both statistically and physically based methods can provide accurate estimates of historical streamflow. Statistically based methods such as POK use measured streamflow to generate synthetic hydrographs, are easy to regionalize to ungauged locations, but lack strong process representations because they are targeted toward producing the best streamflow. Statistical estimates of streamflow depend on the availability and quality of measured streamflow data in proximity to the basin location of interest, without consideration of hydrologic setting or complex hydrologic processes that may be present in a given basin. Additionally, methods such as POK can

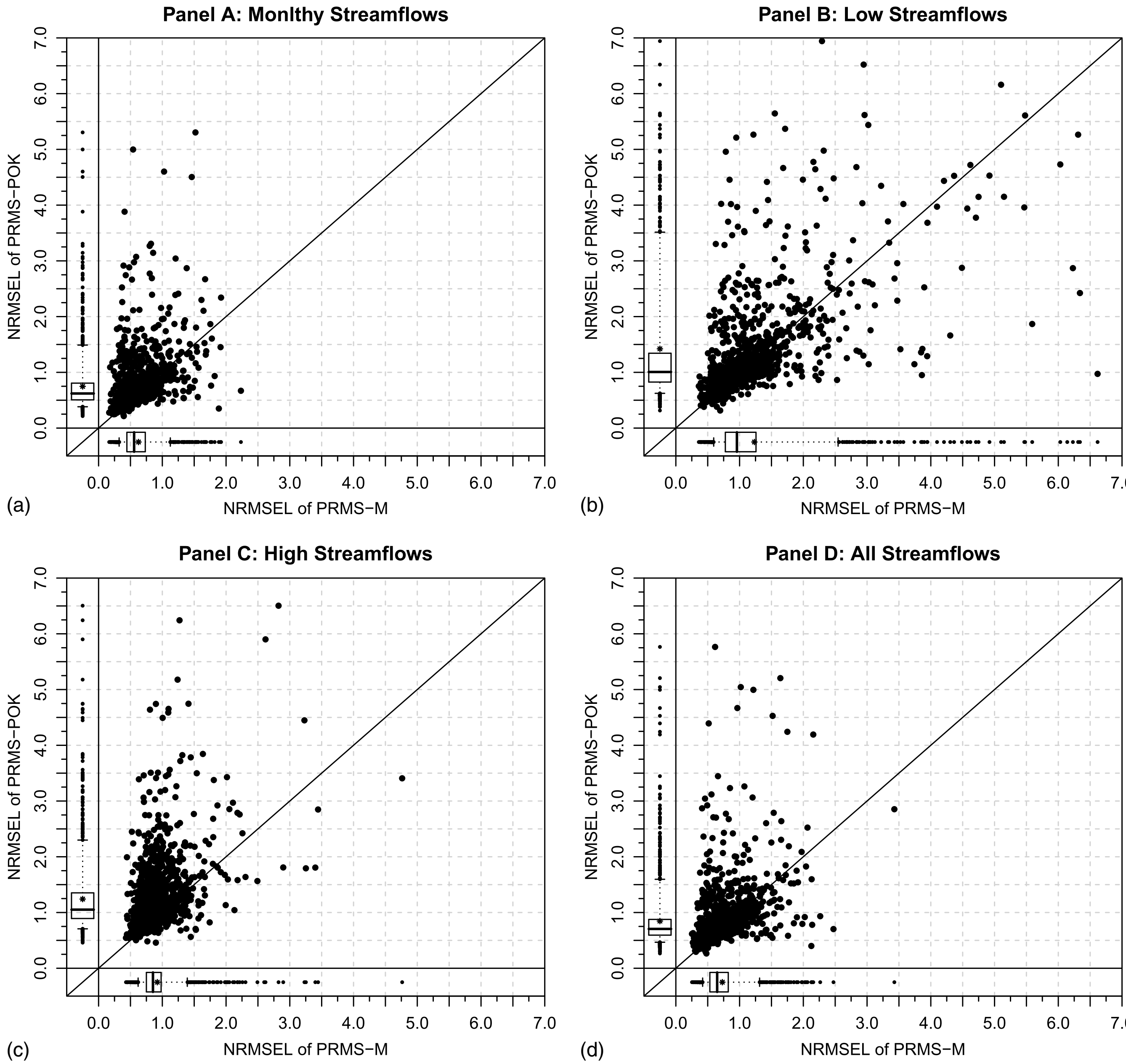


**Fig. 5.** Correspondence and distribution of the normalized RMSE of the natural logarithms of streamflow (NRMSEL) at each of 1,410 stream gauges for the calibration of the Precipitation-Runoff Modeling System to measured streamflow (PRMS-M) and calibration to POK streamflow (PRMS-POK).

provide relatively accurate estimates of streamflow for historical or current conditions, but are limited for applications considering potential changes in the hydroclimatic regime or land cover of a basin of interest.

Physical hydrologic models use surrogate basin characteristics of climate and land cover to simulate streamflow. In general, they are much more data-intensive, may be less accurate for historical or current periods compared with statistical hydrologic models, but are able to include the effects of potential climate and landscape change when such data are available. In addition, physically based methods, such as PRMS, can go beyond simulating the best streamflow, providing a suite of modules that represent each part of the water cycle. Simulations include streamflow with the added benefit of producing a closed water budget as part of the simulation.

Integrating statistical and physical models (PRMS-POK) joins the strengths of both methods. Without any knowledge of at-site streamflow measurements, the PRMS-POK process-based simulation was able to reproduce the daily hydrograph with an accuracy within 23% of the application with complete knowledge of measurements (PRMS-M). This provides invaluable information in a completely ungauged application, but we may be able to achieve additional improvements by considering alternative data sets. The approach, showing little difference between PRMS-POK and POK, demonstrates that little streamflow accuracy is lost with the inclusion of a complete physical water balance.

Measured streamflow data are collected using consistent methodologies by the USGS for the CONUS, but the effects of land cover disturbance, water use, and complex channel behavior hinder the use of much of the measured streamflow record for model

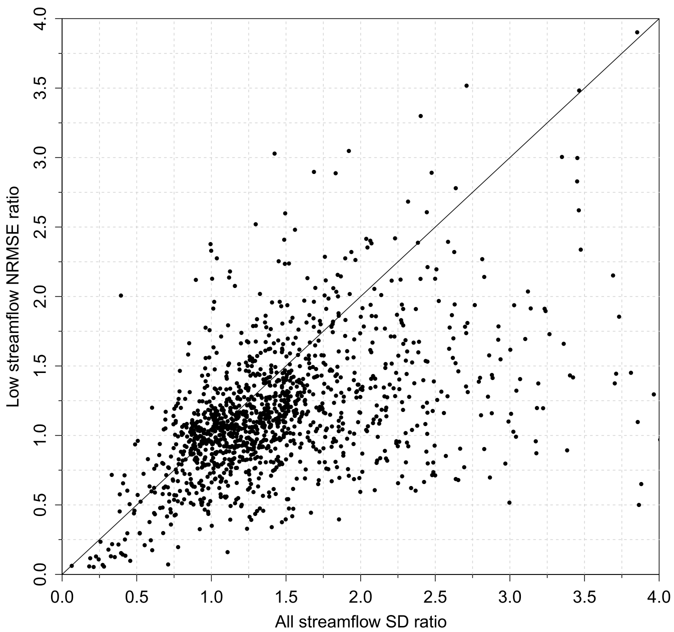


**Fig. 6.** Correspondence between a measure of calibration variability and the accuracy of simulations of low streamflow.

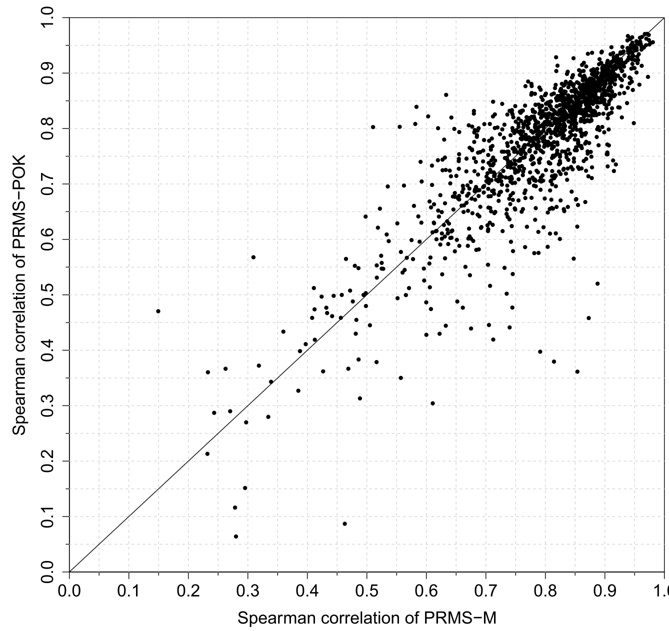


**Fig. 7.** Correspondence of Spearman rank correlations of two methodologies for simulating streamflow relative to measurements.

development. An estimated 87% of the CONUS (Kiang et al. 2013) does not have streamflow information free of anthropogenic influences for the calibration of physical hydrologic models. The PRMS-POK methodology can be effectively used to increase the spatial coverage and resolution of the streamflow calibration information and allow for unimpaired streamflow to be estimated as a point of comparison with impaired flows.

Calibration of the parameters of the PRMS is essential to developing a physically realistic hydrologic model, although adjusting those parameters through a traditional objective function (e.g., sum of squared errors in streamflow) alone is insufficient. Table 1 lists the parameters that were subjected to the four-step calibration procedure defined by Regan et al. (2018), as recommended by sensitivity analyses conducted by Markstrom et al. (2016). Table 1 indicates which parameters started from universal default values across the country and which initial conditions were modified by at-site observations of characteristics such as land cover, the former being undistributed initial conditions and the latter being distributed initial conditions. Considering the percentage change in each parameter with each calibration data set, the distributed parameters had a much stronger Pearson correlation in percentage changes. That is, for parameters that started from distributed initial values, both calibrations produced similar parameters. When initial conditions were undistributed, the percentage changes were less strongly correlated, suggesting concerns of equifinality (Beven and Freer 2001).

Providing streamflow as the sole calibration target of a physically based model is not enough—when physically based models are optimized only for aspects of streamflow (as was done here), simulations of other parts of the water budget may provide inaccurate or nonrealistic results (Refsgaard 1997; Kuppel et al. 2018; Hay et al. 2018). The intermediate process variables computed by the physically based model could be characterized by parameter values that do not replicate those hydrological processes in the physical system (Hay and Umemoto 2006). Alternative data sources (evapotranspiration, snow water equivalent, soil moisture, and so on) in addition to streamflow should be used to constrain parameter values (Fang et al. 2013; Rajib et al. 2016; Thorstensen et al. 2016; Hay et al. 2018). This provides more information to optimize the simulation of the water budget and reduce the effects of equifinality (reduce parameter ranges that provide equally good simulations) (Beven and Freer 2001).

In general, the PRMS-POK method did not simulate the high flows as accurately as PRMS-M, indicating that overall streamflow volumes will not be as accurate (high flows constitute approximately 73% of total volume on average). Although the volume estimates might be degraded, the daily timing estimates using PRMS-POK were similar to those from PRMS-M. Fig. 7 contrasts the Spearman rank correlation between daily flows simulated with PRMS-M relative to measurements against the same measure for PRMS-POK. The horizontal axis shows the Spearman rank correlation between streamflow measurements and streamflow simulated by PRMS-M. The vertical axis shows the Spearman rank correlation between streamflow measurements and streamflow simulated by PRMS-POK. There was a 0.84 Spearman rank correlation between these rank correlations, indicating that the application with and without knowledge of streamflow measurements produces the timing (relative ranking) of daily streamflow similarly. Combining alternative data sources, as described previously, with statistically generated streamflow in a calibration procedure should enhance the simulation of the entire water budget and provide additional information on streamflow timing.

The varying performance, both strengths and weaknesses, of the PRMS and the POK methodologies is likely a product of separate information sources. The statistical approach uses information from regional streamflow measurements, whereas the physical model uses information coming largely from meteorological data sets. For this reason, the combination of the two approaches has great potential. Combining these information sources could be what causes improvements in certain hydrograph characteristics. The sequential calibration presented here will limit the impact of meteorological forcing because it attempts, through calibration, to reproduce the statistical streamflow. If the statistical simulation is inherently flawed, the result will not prove useful. However, this may also be true of limited accuracy in measured data from hydrologically complex regions. In both cases, with statistical simulation and with measured streamflow, the methodology could be improved by incorporating measurement and simulation uncertainty in calibration.

A limitation of the PRMS-POK strategy is the uncertainty associated with POK simulations, or any statistical simulations. However, this is not a flaw exclusive to statistical simulation, nor is it wise to completely ignore measurements where they are available. All calibration data sets, including measured streamflow, will contain errors and uncertainty. The key, and an essential part of future research, will be to understand how to use this information correctly. The results presented here show that, despite possibly greater uncertainty in the statistical simulations, there is some information to support modeling in ungauged basins. Further research might explore the development of a probabilistic framework to link measurement or simulation uncertainty into the coupling of physical and statistical models. Indeed, the POK method does produce estimates of uncertainty that might prove beneficial if used in this way.

Several model intercomparison studies in recent years have shown that no one hydrologic simulation method is best for all scenarios (Farmer et al. 2014a; Caldwell et al. 2015). Based on those findings, and acknowledging the strengths and limitations of various hydrologic simulation methodologies, the PRMS-POK method combines multiple modeling methods to provide hydrologic simulations that are more consistent for both gauged and ungauged areas of a geographic region of interest than either method individually. The ability to simulate streamflow response in ungauged basins with similar accuracy as those with measured streamflow data has been a goal of the hydrologic research community for some time (Hrachowitz et al. 2013). Improved streamflow simulations combined with accurately estimating all other parts of the water budget will provide managers with improved information to make decisions.

## Summary and Conclusions

Characterizing hydrology at ungauged locations is one of the most frequent challenges confronting the field of hydrology today. Improving our ability to reproduce hydrographs at ungauged locations will vastly improve our management of water resources. Considering two extremes of a continuum, two approaches have been used to address this problem. One proposes the use of statistical tools to simulate measured hydrology, but these tools rely on the density of nearby measured information and rarely allow for robust extrapolation. The other proposes the use of process-based models, but it is often difficult to calibrate these models without at-site information. This work proposes a novel solution, combining the accuracy and ease-of-implementation of statistical tools with the process-based possibilities of more-physical models.

The National Hydrologic Model application of the Precipitation-Runoff Modeling System was used to simulated streamflow in 1,410 gauged watersheds across the conterminous United States to test the feasibility of improving streamflow simulations by linking statistical and physical hydrologic models. This work showed how a process-based model could be directly calibrated to statistically generated at-site information with no measured values and achieve a general performance within 23% of applications with complete knowledge of measurements. This performance varied widely by site, with some sites actually showing improvements when disregarding measured streamflow. For low streamflow, 64% of sites saw improvements in the ability to reproduce low streamflow. Indeed, initial evidence demonstrated a connection with the reduction in peakiness of streamflow as driving this improvement, although further work is needed. The greatest strength of statistically generated streamflow may be its ability to reproduce the timing of daily streamflow for calibration in physical models at ungauged locations. Although the use of statistically generated streamflow may allow for the application of process-based models to unmonitored locations, it is hypothesized that the inclusion of additional data sets of other processes may further constrain models and result in improvements of model application.

## Acknowledgments

We are indebted to the anonymous reviewers who took the time to consider this manuscript. All data, including streamflow observations and simulations, used and analysis conducted in this work are available in the associated data release (Farmer et al. 2019). This work was funded through the USGS Water Availability and Use Science Program.

## References

Andréassian, V., A. Hall, N. Chahinian, and J. Schaake. 2006. *Large sample basin experiments for hydrological model parameterization: Results of the model parameter experiment: MOPEX*. Number 307. Paris: International Association of Hydrological Sciences.

Andréassian, V., N. Le Moine, C. Perrin, M.-H. Ramos, L. Oudin, T. Mathevet, J. Lerat, and L. Berthet. 2012. "All that glitters is not gold: The case of calibrating hydrological models." *Hydrol. Processes* 26 (14): 2206–2210. https://doi.org/10.1002/hyp.9264.

Archfield, S. A., A. Pugliese, A. Castellarin, J. O. Skøien, and J. E. Kiang. 2013a. "Topological and canonical kriging for design flood prediction in ungauged catchments: An improvement over a traditional regional regression approach?" *Hydrol. Earth Syst. Sci.* 17 (4): 1575–1588. https://doi.org/10.5194/hess-17-1575-2013.

Archfield, S. A., P. A. Steeves, J. D. Guthrie, and K. G. Ries, III. 2013b. "Towards a publicly available, map-based regional software tool to estimate unregulated daily streamflow at ungauged rivers." *Geosci. Model Dev.* 6 (1): 101–115. https://doi.org/10.5194/gmd-6-101-2013.

Atieh, M., B. Gharabaghi, and R. Rudra. 2015. "Entropy-based neural networks model for flow duration curves at ungauged sites." *J. Hydrol.* 529 (3): 1007–1020. https://doi.org/10.1016/j.jhydrol.2015.08.068.

Atieh, M., G. Taylor, A. M. Sattar, and B. Gharabaghi. 2017. "Prediction of flow duration curves for ungauged basins." *J. Hydrol.* 545 (Feb): 383–394. https://doi.org/10.1016/j.jhydrol.2016.12.048.

Bassiouni, M., R. M. Vogel, and S. A. Archfield. 2016. "Panel regressions to estimate low-flow response to rainfall variability in ungaged basins." *Water Resour. Res.* 52 (12): 9470–9494. https://doi.org/10.1002/2016WR018718.

Bergstrom, S. 2006. *Large sample basin experiments for hydrological model parameterization: Results of the model parameter experiment: MOPEX*. Number 307. Paris: International Association of Hydrological Sciences.

Beven, K., and J. Freer. 2001. "Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the glue methodology." *J. Hydrol.* 249 (1): 11–29. https://doi.org/10.1016/S0022-1694(01)00421-8.

Bicknell, B., J. Imhoff, J. Kittle, T. Jobes, and A. Donigian. 2005. *HSPF version 12.2 users manual*. Athens, GA: USEPA, National Exposure Research Laboratory.

Blodgett, D. L., N. L. Booth, T. C. Kunicki, J. I. Walker, and R. J. Viger. 2011. *Description and testing of the geo data portal: Data integration framework and web processing services for environmental science collaboration*. Open-File Rep. 2011-1157. Reston, VA: USGS.

Blöschl, G., M. Sivapalan, T. Wagener, A. Viglione, and H. Savenije. 2013. *Runoff prediction in ungauged basins: Synthesis across processes, places and scales*. Cambridge, UK: Cambridge University Press.

Blum, A. G., S. A. Archfield, and R. M. Vogel. 2017. "On the probability distribution of daily streamflow in the united states." *Hydrol. Earth Syst. Sci.* 21 (6): 3093–3103. https://doi.org/10.5194/hess-21-3093-2017.

Bock, A. R., L. E. Hay, G. J. McCabe, S. L. Markstrom, and R. D. Atkinson. 2016. "Parameter regionalization of a monthly water balance model for the conterminous united states." *Hydrol. Earth Syst. Sci.* 20 (7): 2861–2876. https://doi.org/10.5194/hess-20-2861-2016.

Bondelid, T., C. Johnston, C. McKay, R. Moore, and A. Rea. 2010. "NHDPlus version 1." Accessed March 14, 2019. http://www.horizon-systems.com/NHDPlus/NHDPlusV1_home.php.

Caldwell, P. V., et al. 2015. "A comparison of hydrologic models for ecological flows and water availability." *Ecohydrology* 8 (8): 1525–1546. https://doi.org/10.1002/eco.1602.

Driscoll, J., S. Markstrom, R. Regan, L. Hay, and R. Viger. 2017. "Data release." In *National Hydrologic Model Parameter Database*. Reston, VA: US Geological Survey.

Duan, Q., S. Sorooshian, and V. Gupta. 1992. "Effective and efficient global optimization for conceptual rainfall-runoff models." *Water Resour. Res.* 28 (4): 1015–1031. https://doi.org/10.1029/91WR02985.

Duan, Q. Y., V. K. Gupta, and S. Sorooshian. 1993. "Shuffled complex evolution approach for effective and efficient global minimization." *J. Optim. Theory Appl.* 76 (3): 501–521. https://doi.org/10.1007/BF00939380.

Falcone, J. 2011. "Gages-II: Geospatial attributes of gages for evaluating streamflow." In *Dataset*. Reston, VA: USGS.

Fang, X., J. W. Pomeroy, C. R. Ellis, M. K. MacDonald, C. M. DeBeer, and T. Brown. 2013. "Multi-variable evaluation of hydrological model predictions for a headwater basin in the Canadian Rocky Mountains." *Hydrol. Earth Syst. Sci.* 17 (4): 1635–1659. https://doi.org/10.5194/hess-17-1635-2013.

Farmer, W., S. Archfield, T. Over, L. Hay, J. LaFontaine, and J. Kiang. 2014a. *A comparison of methods to predict historical daily streamflow time series in the southeastern United States*. Scientific Investigations Rep. 2014-5231. Reston, VA: USGS.

Farmer, W., R. Knight, D. Eash, K. Hutchinson, S. Linhart, D. Christiansen, S. Archfield, T. Over, and J. Kiang. 2015. *Evaluation of statistical and rainfall-runoff models for predicting historical daily streamflow time series in the Des Moines and Iowa River watersheds*. Scientific Investigations Rep. 2015-5089. Reston, VA: USGS.

Farmer, W. H. 2016. "Ordinary kriging as a tool to estimate historical daily streamflow records." *Hydrol. Earth Syst. Sci.* 20 (7): 2721–2735. https://doi.org/10.5194/hess-20-2721-2016.

Farmer, W. H., and G. Koltun. 2017. "Geospatial tools effectively estimate nonexceedance probabilities of daily streamflow at ungauged and intermittently gauged locations in ohio." *J. Hydrol.: Reg. Stud.* 13 (Oct): 208–221. https://doi.org/10.1016/j.ejrh.2017.08.006.

Farmer, W. H., J. H. LaFontaine, and L. E. Hay. 2019. "Physical and statistical simulations of daily streamflow (2000-2010) across the continental United States for an analysis of blended simulation methods." In *US Geological Survey data release*. Reston, VA: USGS.

Farmer, W. H., T. M. Over, and J. E. Kiang. 2018. "Bias correction of simulated historical daily streamflow at ungauged locations by using independently estimated flow duration curves." *Hydrol. Earth Syst. Sci.* 22 (11): 5741–5758. https://doi.org/10.5194/hess-22-5741-2018.

Farmer, W. H., T. M. Over, and R. M. Vogel. 2014b. "Multiple regression and inverse moments improve the characterization of the spatial scaling behavior of daily streamflows in the Southeast United States." *Water Resour. Res.* 51 (3): 1775–1796. https://doi.org/10.1002/2014WR015924.

Farmer, W. H., and R. M. Vogel. 2016. "On the deterministic and stochastic use of hydrologic models." *Water Resour. Res.* 52 (7): 5619–5633. https://doi.org/10.1002/2016WR019129.

Gupta, H. V., M. P. Clark, J. A. Vrugt, G. Abramowitz, and M. Ye. 2012. "Towards a comprehensive assessment of model structural adequacy." *Water Resour. Res.* 48 (8): W08301. https://doi.org/10.1029/2011WR011044.

Hay, L., P. Norton, R. Viger, S. Markstrom, R. Steven Regan, and M. Vanderhoof. 2018. "Modelling surface-water depression storage in a prairie pothole region." *Hydrol. Processes* 32 (4): 462–479. https://doi.org/10.1002/hyp.11416.

Hay, L., and M. Umemoto. 2006. *Multiple-objective stepwise calibration using Luca*. Open-File Rep. 2006-1323. Reston, VA: USGS.

Hay, L. E., G. H. Leavesley, M. P. Clark, S. L. Markstrom, R. J. Viger, and M. Umemoto. 2006. "Step wise multiple objective calibration of a hydrologic model for a snowmelt dominated basin." *J. Am. Water Resour. Assoc.* 42 (4): 877–890. https://doi.org/10.1111/j.1752-1688.2006.tb04501.x.

Hirsch, R., and L. DeCicco. 2015. "User guide to exploration and graphics for river trends (egret) and dataretrieval: R packages for hydrologic data." In *Techniques and methods Book 4, Chapter A10*: Reston, VA: USGS.

Hrachowitz, M., et al. 2013. "A decade of Predictions in Ungauged Basins (PUB)—A review." *Hydrol. Sci. J.* 58 (6): 1198–1255. https://doi.org/10.1080/02626667.2013.803183.

Kendy, E., M. Cutrofello, and M. Fekadu. 2011. "Unified hydrologic model for assessing human and climate impacts on streamflows at multiple geographic scales." Accessed September 8, 2017. https://www.sciencebase.gov/catalog/item/595e9001e4b0d1f9f0570425.

Kiang, J., D. Stewart, S. Archfield, E. Osborne, and K. Eng. 2013. *A national streamflow network gap analysis*. Scientific Investigations Rep. 2013-5013. Reston, VA: USGS.

Kuppel, S., D. Tetzlaff, M. P. Maneta, and C. Soulsby. 2018. "EcH$_2$O-iso 1.0: Water isotopes and age tracking in a process-based, distributed ecohydrological model." *Geosci. Model Dev.* 11 (7): 3045–3069. https://doi.org/10.5194/gmd-11-3045-2018.

Markstrom, S., R. Regan, L. Hay, R. Viger, R. Webb, R. Payn, and J. LaFontaine. 2015. "*Prms-IV, the precipitation-runoff modeling system, version 4*." In *Techniques and methods book 6, chapter B7*. Reston, VA: USGS.

Markstrom, S. L., L. E. Hay, and M. P. Clark. 2016. "Towards simplification of hydrologic modeling: Identification of dominant processes." *Hydrol. Earth Syst. Sci.* 20 (11): 4655–4671. https://doi.org/10.5194/hess-20-4655-2016.

Merz, R., and G. Blöschl. 2004. "Regionalisation of catchment model parameters." *J. Hydrol.* 287 (1): 95–123. https://doi.org/10.1016/j.jhydrol.2003.09.028.

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

Neitsch, S., J. Arnold, and J. Kiniry. 2005. *Soil and water assessment tool, theoretical documentation*. Rep. No. TR-406 Washington, DC: US Dept. of Agriculture, Agricultural Research Service.

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 (3): W03413. https://doi.org/10.1029/2007WR006240.

Oudin, L., A. Kay, V. Andréassian, and C. Perrin. 2010. "Are seemingly physically similar catchments truly hydrologically similar?" *Water Resour. Res.* 46 (11): W11558. https://doi.org/10.1029/2009WR008887.

Peel, M. C., F. H. S. Chiew, A. W. Western, and T. A. McMahon. 2000. *Extension of unimpaired monthly streamflow data and regionalization of parameter values to estimate streamflow in ungauged catchments*. Report to National Land and Water Resources Audit. Parkville, Australia: Univ. of Melbourne.

Poncelet, C., R. Merz, B. Merz, J. Parajka, L. Oudin, V. Andréassian, and C. Perrin. 2016. "Process-based interpretation of conceptual hydrological model performance using a multinational catchment set." *Water Resour. Res.* 53 (8): 7247–7268. https://doi.org/10.1002/2016WR019991.

Pugliese, A., W. H. Farmer, A. Castellarin, S. A. Archfield, and R. M. Vogel. 2016. "Regional flow duration curves: Geostatistical techniques versus multivariate regression." *Adv. Water Resour.* 96 (Oct): 11–22. https://doi.org/10.1016/j.advwatres.2016.06.008.

Rajib, M. A., V. Merwade, and Z. Yu. 2016. "Multi-objective calibration of a hydrologic model using spatially distributed remotely sensed/in-situ soil moisture." *J. Hydrol.* 536 (May): 192–207. https://doi.org/10.1016/j.jhydrol.2016.02.037.

Refsgaard, J. C. 1997. "Parameterisation, calibration and validation of distributed hydrological models." *J. Hydrol.* 198 (1): 69–97. https://doi.org/10.1016/S0022-1694(96)03329-X.

Regan, R., K. Juracek, L. Hay, S. Markstrom, R. Viger, J. Driscoll, J. LaFontaine, and P. Norton. 2019. "The US Geological Survey National Hydrologic Model infrastructure: Rationale, description, and application of a watershed-scale model for the conterminous United States." *Environ. Modell. Software* 111 (Jan): 192–203. https://doi.org/10.1016/j.envsoft.2018.09.023.

Regan, R., S. Markstrom, L. Hay, R. Viger, P. Norton, J. Driscoll, and J. LaFontaine. 2018. "Description of the national hydrologic model for use with the precipitation-runoff modeling system (PRMS)." In *Techniques and methods book 6, chapter B9*. Reston, VA: USGS.

Reichl, F., and J. Hack. 2017. "Derivation of flow duration curves to estimate hydropower generation potential in data-scarce regions." *Water* 9 (8). https://doi.org/10.3390/w9080572.

Rozalis, S., E. Morin, Y. Yair, and C. Price. 2010. "Flash flood prediction using an uncalibrated hydrological model and radar rainfall data in a Mediterranean watershed under changing hydrological conditions." *J. Hydrol.* 394 (1): 245–255. https://doi.org/10.1016/j.jhydrol.2010.03.021.

Samuel, J., P. Coulibaly, and R. A. Metcalfe. 2011. "Estimation of continuous streamflow in Ontario ungauged basins: Comparison of regionalization methods." *J. Hydrol. Eng.* 16 (5): 447–459. https://doi.org/10.1061/(ASCE)HE.1943-5584.0000338.

Sawicz, K., T. Wagener, M. Sivapalan, P. A. Troch, and G. Carrillo. 2011. "Catchment classification: Empirical analysis of hydrologic similarity based on catchment function in the eastern USA." *Hydrol. Earth Syst. Sci.* 15 (9): 2895–2911. https://doi.org/10.5194/hess-15-2895-2011.

Seaber, P. R., F. P. Kapinos, and G. L. Knapp. 1987. "*Hydrologic unit maps*." In *Water supply paper 2294*. Reston, VA: USGS.

Seibert, J. 1999. "Regionalisation of parameters for a conceptual rainfall-runoff model." *Agric. For. Meteorol.* 98–99 (Dec): 279–293. https://doi.org/10.1016/S0168-1923(99)00105-7.

Shi, X., A. W. Wood, and D. P. Lettenmaier. 2008. "How essential is hydrologic model calibration to seasonal streamflow forecasting?" *J. Hydrometeorol.* 9 (6): 1350–1363. https://doi.org/10.1175/2008JHM1001.1.

Sivapalan, M. 2003. "Prediction in ungauged basins: A grand challenge for theoretical hydrology." *Hydrol. Processes* 17 (15): 3163–3170. https://doi.org/10.1002/hyp.5155.

Tegegne, G., and Y.-O. Kim. 2018. "Modelling ungauged catchments using the catchment runoff response similarity." *J. Hydrol.* 564 (Sep): 452–466. https://doi.org/10.1016/j.jhydrol.2018.07.042.

Thornton, P. E., H. Hasenauer, and M. A. White. 2000. "Simultaneous estimation of daily solar radiation and humidity from observed temperature and precipitation: An application over complex terrain in Austria." *Agric. For. Meteorol.* 104 (4): 255–271. https://doi.org/10.1016/S0168-1923(00)00170-2.

Thornton, P. E., and S. W. Running. 1999. "An improved algorithm for estimating incident daily solar radiation from measurements of temperature, humidity, and precipitation." *Agric. For. Meteorol.* 93 (4): 211–228. https://doi.org/10.1016/S0168-1923(98)00126-9.

Thornton, P. E., S. W. Running, and M. A. White. 1997. "Generating surfaces of daily meteorological variables over large regions of complex terrain." *J. Hydrol.* 190 (3): 214–251. https://doi.org/10.1016/S0022-1694(96)03128-9.

Thorstensen, A., P. Nguyen, K. Hsu, and S. Sorooshian. 2016. "Using densely distributed soil moisture observations for calibration of a hydrologic model." *J. Hydrometeorol.* 17 (2): 571–590. https://doi.org/10.1175/JHM-D-15-0071.1.

TNC (The Nature Conservancy). 2009. *Indicators of hydrologic alteration version 7.1 user's manual*. Arlington, VA: TNC.

USGS. 2017. "USGS water data for the nation." In *National Water Information System Database*. Reston, VA: USGS.

Viger, R., and A. Bock. 2014. "GIS features of the geospatial fabric for national hydrologic modeling." In *Metadata*. Reston, VA: USGS.

Worland, S. C., W. H. Farmer, and J. E. Kiang. 2018. "Improving predictions of hydrological low-flow indices in ungaged basins using machine learning." *Environ. Modell. Software* 101 (Mar): 169–182. https://doi.org/10.1016/j.envsoft.2017.12.021.

Zhang, X., R. Srinivasan, and M. V. Liew. 2008. "Multi-site calibration of the swat model for hydrologic modeling." *Trans. ASABE* 51 (6): 2039–2049. https://doi.org/10.13031/2013.25407.