Comparison of spatial interpolation methods for the estimation of precipitation patterns at different time scales to improve the accuracy of discharge simulations
Dedi Liu, Qin Zhao, Dezhi Fu, Shenglian Guo, Pan Liu and Yujie Zeng

ABSTRACT

Interpolating precipitation data is of prime importance to hydrological design, modeling, and water resource management. Various models have been developed that estimate spatial precipitation patterns. The purpose of this study is to analyze different precipitation interpolation schemes at different time scales in order to improve the accuracy of discharge simulations. The study was carried out in the upstream area of the Changjiang River basin. The performance of all selected methods was assessed using cross-validation schemes, with the mixed methods ultimately displaying the best performance at all three time scales. However, the differences in performance between the spatial interpolation methods decreased with increasing time scales. The unifying catchment Soil and Water Assessment Tool (SWAT), ‘abcd’, and the Budyko equation were employed at the daily, monthly, and annual scales, respectively, to simulate discharge. The performance of the discharge simulation at the monthly and annual time scales was consistent with their ranks of spatial precipitation estimation. For coarse, or long period, precipitation, there were no significant differences. However, the mixed methods performed better than the single model for the daily, or short, time scale with respect to the accuracy of the discharge simulation.

Key words | ‘abcd’, Budyko equation, precipitation, spatial interpolation, SWAT

HIGHLIGHTS

- The performances of different precipitation interpolation schemes at different time scales have been figured out.
- The mixed spatial interpolation schemes for precipitation are found to be the best at all the time scales due to its own parameters and structure.
- The accuracy of discharge simulation is found to be sensitive to the spatial interpolation schemes for precipitation at short time scale while there are no significant differences at monthly and annual time scales.
- The importance of time scales in the precipitation interpolation are recommended to be considered before determining the areal precipitation.
- The proposed framework of the selecting the methods can also be applied to another spatial precipitation interpolation methods.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Licence (CC BY 4.0), which permits copying, adaptation and redistribution, provided the original work is properly cited (http://creativecommons.org/licenses/by/4.0/).

doi: 10.2166/nh.2020.146
Precipitation is one of the most important variables in the water cycle. Accurate continuous spatial precipitation data plays a significant role in the hydrological planning, modeling, and management of water resource systems (Parker et al. 2019). As precipitation is often only recorded at individual gauge sites and is very dispersed at larger scales (Vicente-Serrano et al. 2003), continuous spatial data over a region cannot be directly collected from measurements.

In order to solve this problem, various spatial interpolation methods for the estimation of precipitation have been developed based on point measurements over the past several years (Chaubey et al. 1999; Zhang & Srinivasan 2009). Spatial interpolation methods range from conceptually simple weighting methods, such as Thiessen polygons (Thiessen 1911) or Inverse Distance Weighting schemes, to more complex approaches using artificial intelligence and covariates (Burrough & Mc Donnell 1998; Vicente-Serrano et al. 2003). In addition to elevation data, other parameters extracted from digital elevation models (DEMs) are used as covariates for interpolating precipitation. Grid data, such as Radar (NEXRAD) (Teegavarapu et al. 2012) and satellite data acquired by Tropical Rainfall Measuring Mission (TRMM) products (Wagner et al. 2012), are increasingly used in spatial interpolation as covariates. The mixed, or multi-method ensemble, approach has been increasingly used to improve model estimations (Hagedorn et al. 2005). The contribution of each individual model in the mixed, or multi-method ensemble, approach is represented by its weight. The use of simple weighting methods is still very common, as there are always insufficient or low-resolution covariate data and requires only low intensity computation. As both simple weighting methods and complex approaches have their own advantages and disadvantages based on specific case studies, it is necessary to select an appropriate spatial interpolation method for a given input dataset (Burrough & Mc Donnell 1998). Many comparative studies on precipitation interpolation techniques have been conducted (Vicente-Serrano et al. 2003; Li & Heap 2011; Bennet et al. 2013; Plouffe et al. 2015; Ding et al. 2018); however, there is little consensus on the optimal interpolator for precipitation (Dirks et al. 1998; Zimmerman et al. 1999; Price et al. 2000; Plouffe 2015). Although some studies have compared the interpolation performance for different station densities (Krajewski 1987; Goudenhoofdt & Delobbe 2009; Yoon et al. 2012; Nanding et al. 2015), very few have considered different time scales (Dirks et al. 1998; Bárdossy & Pegram 2014; Plouffe et al. 2015). As precipitation data properties differ with their time scales, the performance of a spatial interpolation method is affected. Therefore, it is important to develop a process for determining which method of spatial interpolation is best suited for different time scales.

The aims of this study are to provide a framework and useful suggestions on how to determine the most effective spatial interpolation methods for precipitation at different time scales, and to improve the accuracy of discharge simulations. The accuracy assessment for the precipitation interpolation results is done by applying cross-validation techniques (Hattermann et al. 2005) at different time scales, and then simulating the corresponding runoff driven by the precipitation interpolation results using a unifying hydrological model at the different time scales. The propagation of the uncertainties from the precipitation spatial interpolation method at each time scale is also discussed.

METHODOLOGY

In order to evaluate the performance of various precipitation interpolation methods with regard to their suitability at different time scales, daily, monthly, and annual precipitation distributions were selected for testing via cross-validation schemes. The results from each precipitation interpolation method were also used to drive hydrological models in order to analyze the effects of the different precipitation interpolation methods on the discharge simulation. A framework of the entire workflow process employed in this study is depicted in Figure 1, with associated abbreviations shown in Table 1.

Spatial interpolation methods for estimating precipitation distribution

Spatial interpolation methods can be grouped into four categories: local methods, global methods, geostatistical...
methods, and mixed methods (Burrough & Mc Donnell 1998; Plouffe et al. 2015). Six different methods (Table 1) were applied and compared in our case study. These methods were evaluated on daily, monthly, and annual scales.

Nearest Neighbor (NN) and Inverse Distance Weighting (IDW) are local spatial interpolation methods that utilize the concept of Tobler’s first law to estimate precipitation by distance (Chen & Liu 2014). The NN (which is also called the Thiessen polygons interpolation method) method assumes that point $p_u$ at the unsampled spatial location is equal to the closest available rainfall, $p_i$. If $\lambda_{u,i}$ is the distance between locations $u$ and $i$, the weight of $\lambda_{u,i}$ is 1 for the location $i$ corresponding to the minimum distance, $d_{u,i}$, and zero for all other locations (Thiessen 1911; Nikolopoulos et al. 2015). The IDW method, recommended by the Handbook of Hydrology by ASCE (1996), is easy to implement (Teegavarapu & Elshorbagy 2008). IDW (Shepard 1968) assumes that weights, $\lambda_{u,i}$, are also related to the distance, $d_{u,i}$, from the sampled point, $p_i$, to the estimated point, $p_u$, as given by Equation (1).

$$p_u = \sum_{i=1}^{N} \lambda_i p_i = \sum_{i=1}^{N} \left( \frac{d_{u,i}^{-\alpha}}{\sum_{j=1}^{N} d_{u,j}^{-\alpha}} \right) p_i$$  

where $N$ is the number of sample points, and $\alpha$ is the power parameter in the interpolation. The variable $\alpha$ is the friction distance (Vieux 2001), ranging from 1.0 to 6.0, and is the main factor affecting the estimation. It also controls the decay of the weights with distance. The higher the value of $\alpha$, the greater the influence of the points closest to the interpolated location (Nikolopoulos et al. 2015).

Different from the basis of Tobler’s first law, where the precipitation is estimated by distance for the two local spatial interpolation methods (NN and IDW), trend surface analysis (TSA) is a global method that treats each value in a geographic area as a centroid in a grid pattern (Wang 2006).
If $U$ and $V$ are geographic coordinates, $p_{ij}$ is a single observation of the mapped variable at coordinates $(U_i, V_j)$. The $p_u$ at an unmeasured spatial location can be estimated by Equation (2) (Wong et al. 2005).

$$p_u = \sum_{i,j \leq N} \lambda_{ij} U_i V_j$$

where $\lambda_{ij}$ is the polynomial coefficient and $N$ is the degree of the polynomial that represents the best surface of the precipitation pattern. $\lambda_{ij}$ and $N$ are determined by using the least-squares approximation (Krumbein 1959) or minimum Akaike Information Criterion (AIC) (Bumhan & Anderson 2002; Soares et al. 2015). The major issue with TSA is selection of the appropriate functional form to model the trend (Sullivan & Unwin 2005).

As TSA, NN, and IDW do not exploit the statistical properties of the observational samples (Foehn et al. 2018), the weights of the Ordinary Kriging (OK) and Co-kriging (CK) methods are optimized based on information that is inherent in the measured data (Wagner et al. 2012). The weights in OK are obtained by solving Equations (3)–(5).

$$\begin{align*}
\sum_{i=1}^{N} \lambda_i \gamma(x_i - x_u) + \phi &= \gamma(x_i - x_u), \text{ for all } j \\
\sum_{i=1}^{N} \lambda_i &= 1
\end{align*}$$

where $\gamma(x_i - x_j)$ represents the value of the semivariogram function for the distance between points $x_i$ and $x_j$. $\gamma(x_i - x_u)$ is the value of the distance between $x_i$ and the estimated location $x_u$, and $\phi$ is the Lagrange parameter, which is equal to the mean distance. On the basis of the second-order stationary hypothesis, the semivariogram is derived by fitting a semivariogram model to the empirical one, and can be determined for all distances, $h$, by solving Equation (4).

$$\hat{\gamma}(h) = \frac{1}{2N} \sum_{i=1}^{N} [p(x_i) - p(x_i + h)]^2$$

Gaussian (Equation (5a)), spherical (Equation (5b)), and exponential (Equation (5c)) semivariogram models are employed to fit the empirical model shown in Equation (4).

$$\begin{align*}
\gamma(h) &= \begin{cases} 
0 & h = 0 \\
C_0 + C \left(1 - e^{-\frac{h}{\lambda}}\right) & h > 0
\end{cases}
\quad (5a) \\
\gamma(h) &= \begin{cases} 
0 & h = 0 \\
C_0 + C \left(\frac{3h}{2a^2} - \frac{1}{2a^2}\right) & 0 < h \leq a \\
C_0 + C & h > a
\end{cases}
\quad (5b) \\
\gamma(h) &= \begin{cases} 
0 & h = 0 \\
C_0 + C \left(1 - e^{-\frac{h}{2a}}\right) & h > 0
\end{cases}
\quad (5c)
\end{align*}$$

where $C_0$, $C_0 + C$, and $a$ are the parameters known as the nugget, the sill, and the range, respectively. CK offers the possibility of considering more than one variable in the kriging interpolation, especially elevation (Pardo-Iguzquiza et al. 2011), and is expressed as following:

$$p^*(x_u) = \sum_{i=1}^{N_1} \lambda_{1i} p_1(x_{1i}) + \sum_{j=1}^{N_2} \lambda_{2j} p_2(x_{2j})$$

where the subscripts ‘1’ and ‘2’ refer to the primary and secondary networks, respectively, and $N_1$ and $N_2$ are the number of available rain gauges in the primary and secondary networks over the period, respectively. Weights $\lambda_1$ and $\lambda_2$ are obtained by solving Equation (7).

$$\begin{align*}
\sum_{i=1}^{N_1} \lambda_{1i} \gamma_{11}(x_{1i} - x_i) + \sum_{j=1}^{N_2} \lambda_{2j} \gamma_{21}(x_{2j} - x_i) + \mu_1 &= \gamma_{21}(x_u - x_i), I = 1, \ldots, N_1 \\
\sum_{i=1}^{N_1} \lambda_{1i} \gamma_{21}(x_{1i} - x_i) + \sum_{j=1}^{N_2} \lambda_{2j} \gamma_{22}(x_{2j} - x_i) + \mu_2 &= \gamma_{22}(x_u - x_i), J = 1, \ldots, N_2 \\
\sum_{i=1}^{N_1} \lambda_{1i} &= 1, \sum_{j=1}^{N_2} \lambda_{2j} = 0
\end{align*}$$
In order to combine the physical relationships between climatic data and geographic or topographic variables, as done in the global TSA method, and the spatial correlation between the information recorded at weather stations, as done in OK, a mixed method, called TSA-OK, has been proposed to improve the accuracy of spatial interpolation methods. The TSA is an inexact interpolator, and has residual errors (as shown in Equation (8)). The residuals from the TSA method are then interpolated by the OK method (Equation (9)). The exact climatic variable for the final prediction is obtained through Equation (10).

\[
p(x) = p^*(x) + E^*(x)
\]  

Although all four categories of spatial interpolation methods have been widely used in the estimation of precipitation distribution, none have been able to select the correct method with certainty a priori. The uncertainties associated with selection of the method should be taken into account in order to reduce prediction error. Bayesian Model Averaging (BMA) addresses this issue by its inherent incorporation of method uncertainty into predictors via considering all competing methods (including variations in their parameters), and making predictions using a weighted average of all of them, with weights representing the posterior model probabilities (Schneider & Yaşar 2016; Zhang & Taflanidis 2019). Consider a predicted or interpolated precipitation, \( p \), the corresponding target, or training data, \( \text{p}_T \), and an ensemble of all spatial interpolation method (i.e., TSA, NN, IDW, OK, and CK) simulations \( \{f_1, f_2, f_3, f_4, f_5\} \) of the precipitation. According to the law of total probability (Bulmer 1979), the probability density function of \( p \) can be expressed as:

\[
\Pr(p|f_k) = \frac{\text{Pr}(p\mid f_k) \Pr(f_k)}{\text{Pr}(p)}
\]

where \( \Pr(p|f_k) \) is the probability prediction given by the simulation of \( f_k \) alone, and \( \Pr(f_k|p_T) \) is the posterior probability of the method prediction, given the data set, \( p_T \). \( \Pr(f_k|p_T) \) is also identified as a fractional statistical weight, \( w_k \), of the \( k \)th spatial interpolation method, whose magnitude reflects how well \( f_k \) matches \( p_T \), and \( \sum_{k=1}^{5} w_k = 1 \). In order to determine the prior \( \Pr(p|f_k) \) for the parameters and the prior probability of each method, it is computationally convenient to assume that \( \Pr(p|f_k) \) is a Gaussian distribution defined by a mean, \( \mu_k \), and variance, \( \sigma^2_k \) (Raftery et al. 2005; Duan & Phillips 2010).

\[
\Pr(p|\theta_k) = g(p|\theta_k)
\]

where \( g(\cdot) \) is the Gaussian probability density function, and \( \theta_k = [\mu_k, \sigma^2_k] \) is the parameter vector.

Substituting (12) into (11),

\[
\Pr(p|f_1, f_2, f_3, f_4, f_5) = \sum_{k=1}^{5} g(p|\theta_k)w_k
\]

The log-likelihood function, \( l \), was maximized to obtain both the Bayesian weight, \( w_k \), and the parameter vector, \( \theta_k \), for its convenient computation. The log-likelihood function, \( l \), is approximated as:

\[
l(\theta_1, \theta_2, \theta_3, \theta_4, \theta_5) = \sum_{k=1}^{5} \log \left( \sum_{k=1}^{5} g(p_T|\theta_k)w_k \right)
\]

In order to eliminate the overreach of biasing inferences on or the bias from only one of several methods considered, simple linear regression is often used, and both the target and simulations are pre-processed using the Box-Cox transformation to align them more closely with the Gaussian distribution (Chen et al. 2015). The parameters \( \theta \) and \( w \) can be estimated by the maximum the logarithm of the likelihood function shown in Equation (14). As the function cannot be maximized analytically, the Expectation-Maximization (EM) algorithm is used to numerically solve the problem (Wu 1985; McLachlan & Krishnan 1997; Raftery et al. 2005).

**Hydrological models**

Wang and Zhao (2014) and Zhao et al. (2016) derived a unifying catchment water balance model for different time
scales through the maximum entropy production principle. The Budyko-type model at the annual scale, the ‘abcd’ model at the monthly scale, and the Soil Conservation Service (SCS) curve number model at the event scale are special cases of unifying catchment water balance models. In order to minimize the uncertainties of the hydrological model structures, the effects of the various interpolation inputs on the water balance and runoff dynamics at different time scales were assessed using the Soil & Water Assessment Tool (SWAT) based on the SCS model at the daily scale, ‘abcd’ at the monthly scale, and Budyko-type at the annual scale.

**SWAT model**

Since the SWAT is a continuous, semi-distributed, process-based basin scale model (Arnold et al. 1998), the basin was divided into multiple sub-basins which were then further subdivided into hydrologic response units (HRUs) (Wu et al. 2019). Water entering each sub-basin stream via direct runoff, soil lateral flow, or groundwater flow was then routed through the watershed stream network (Aliyari et al. 2019). Direct runoff was estimated using the Soil Conservation Service Curve Number (SCS-CN) model (Hjelmfelt 1991). Precipitation was partitioned into direct runoff and soil wetting, where soil wetting included initial abstraction and continuing abstraction. The direct runoff was assumed to equal the product of the remaining water after initial abstraction and the ratio of continuing abstraction to its potential continuing abstraction. Lateral flow, as a function of the interflow-recession coefficient, was calculated using the kinematic storage model (Neitsch et al. 2005; Chen et al. 2019). Groundwater flow in shallow (active) and deep (inactive) aquifers was simulated using a conceptual linear one-reservoir (shallow aquifer storage) approach (Fu et al. 2015; Pan et al. 2017).

**‘Abcd’ model**

The ‘abcd’ model is a nonlinear monthly hydrological model that was originally proposed by Thomas (1981) for national water assessment. The key assumption of the ‘abcd’ model is that the evapotranspiration opportunity is nonlinearly related to the available water in a manner such that the evapotranspiration opportunity increases quickly with the available water for water-limited conditions, but asymptotically approaches a maximum constant value for energy-limited conditions (Martinez & Gupa 2010). The structure of the model can be found in Thomas (1981), Alley (1984), Fernandez et al. (2000), and Sankarasubramanian & Vogel (2002).

**Budyko model**

At the watershed scale, if the water storage change in the mean annual water balance was negligible, the mean annual precipitation was portioned into runoff and evaporation (Budyko 1958). The partitioning of the precipitation was determined by the competition between the available water (i.e., the annual precipitation) and available energy as measured by potential evaporation. The evaporative index, expressed by the ratio between the actual evaporation and precipitation, is a non-linear function of the climate dryness index, which was determined by the ratio of potential evaporation to precipitation. Then, the runoff was simply the difference between the precipitation and actual evaporation (Yang et al. 2008).

**Statistical criteria for assessing the agreements of the spatial interpolation methods and hydrological models**

**Statistical criteria for the spatial interpolation used to estimate precipitation**

The performance of every spatial interpolation used to estimate precipitation was assessed and compared using the leave-one-out, or cross-validation, approach (Goovaerts 2000). This approach estimated precipitation at a gauge location using all observations except the one corresponding to the interpolation location (Isaaks & Srivastava 1989; Foehn et al. 2008). The comparison criterions were the Mean Absolute Error (MAE), Pearson Product moment correlation measures (R), and Root-Mean-Square Error (RMSE), which measure overall accuracy (MAE and R) or reflect the extreme outliers (RMSE) (Neil et al. 2013; Berndt & Haberlandt 2018). The procedure was undertaken on different time scales (daily, monthly, and annually in this
study) for each of the spatial interpolation methods.

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^{n} |\hat{p}_i - o_i|$$  \hspace{2cm} (15)

$$R = \frac{\sum_{i=1}^{n} (\hat{p}_i - p)(o_i - \bar{o})}{\sqrt{\sum_{i=1}^{n} (\hat{p}_i - p)^2 \sqrt{\sum_{i=1}^{n} (o_i - \bar{o})^2}}$$  \hspace{2cm} (16)

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (\hat{p}_i - o_i)^2}$$  \hspace{2cm} (17)

where $O_i$ and $\hat{p}_i$ are the observation and estimation at the $i$th gauge station, respectively.

**Performance assessment of hydrological models**

To evaluate the fitness of the hydrological models, Nash-Sutcliffe Efficiency (NSE) (Nash & Sutcliffe 1970) and Percentage Bias (PBIAS) (Moriasi et al. 2007) were used to measure the simulation accuracy of the discharge and are given by Equations (18) and (19).

$$\text{NSE} = 1 - \frac{\sum_{i=1}^{n} (Q_{oi} - Q_{si})^2}{\sum_{i=1}^{n} (Q_{oi} - Q_{s0})^2}$$  \hspace{2cm} (18)

$$\text{PBIAS} = 100 \times \frac{\sum_{i=1}^{n} (Q_{oi} - Q_{si})}{\sum_{i=1}^{n} Q_{oi}}$$  \hspace{2cm} (19)

**STUDY AREA AND DATA SET**

The study area is located upstream of Yichang city in Hubei and is defined as the upper reach of the Changjiang River (24.5° N–27.8° N; 90.6° E–111.5° E), which has a large topographic gradient ranging from 200 m to 6,500 m. The river is 4,500 km long with a basin area of $1.0 \times 10^6$ km$^2$ (Figure 2). The upper reach of the Changjiang River has a subtropical, humid monsoon climate, which is characterized by seasonal rainfall from June to October. Owing to its

![Figure 2](https://iwaponline.com/hr/article-pdf/doi/10.2166/nh.2020.146/706769/nh2020146.pdf)
extensive geographic area and various sources of precipitation (i.e., the South Asia summer monsoon (SASM), Plateau summer monsoon, and elevation), the precipitation spatial distribution is irregular at different time scales. As such, the precipitation spatial interpolations cannot ignore the impacts of time scales in the study area. The mean annual precipitation amount is 1,232 mm, and the annual average temperature is 15–18 °C. The dominant soil types are purple, yellow, and paddy soils. The Three Gorges Dam, the world’s largest hydropower project to date, is located 44 km upstream of the Yichang hydrological gauge station.

Daily precipitation measurements at 53 gauge stations, with continuous data collected from 1,960 to 2008, were collected from the website of the China Meteorological Bureau (http://data.cma.cn/). These precipitation stations displayed a relatively uniform distribution over the study area as shown in Figure 2. The daily discharge data at the Yichang gauge station from 1960 to 2008 were provided by the Bureau of Hydrology of the Changjiang Water Resources Commission of China.

RESULTS AND DISCUSSION

Parameters of precipitation spatial interpolation calibration

The critical problem associated with the TSA is selection of the optimal polynomial order (Chayes 1970). In this study, the degree of fitting (AIC) was used to determine the polynomial order of the TSA. The lower the AIC value a particular method produced, the better fitness the observed data provided. As presented in Table 2, a third order polynomial was assumed to be sufficient to capture all the daily, monthly, and annual precipitation variations.

To identify the optimal parameters of the IDW method at the daily, monthly, and annual time scales, the α value, number of sample points, N, the research radius, the minimum MAE and RMSE, and the maximum R should be determined. The values of MAE, RMSE, and R, shown in Figure 3, revealed that the optimal values for parameters N and α can be determined as shown in Table 3.

Three semivariogram (Gaussian, spherical, and exponential) models were adopted for fitting the empirical model in the OK and CK spatial interpolation methods. The cross-validation technique was used to calibrate the parameters of the semivariogram models. The best-fitting method was determined by selecting the model with the lowest MAE and RMSE, or the highest R. There were no significant differences among the three semivariogram models or between the OK and CK methods at the daily (Figure 4(a)), monthly (Figure 4(b)), and annual (Figure 4(c)) time scales. However, the spherical semivariogram model was the best, as determined by the average statistical criteria values (shown in Table 4) and was selected by the OK and CK spatial interpolation methods.

The global optimal BMA weights for the log-likelihood function (13) were obtained by using the EM algorithm with daily, monthly, and annual time steps (shown in Figure 5). As there were five models, the average weights were approximately 0.2. The uncertainty of the daily weights was the highest (Figure 5(a)), while the annual was the lowest (Figure 5(c)). Only the expected BMA estimation was adopted in this study.

Performance of the spatial interpolation methods for estimating precipitation at different time scales

After the parameters of the five interpolation methods (TSA, OK, CK, TSA-OK, and BMA) were calibrated, their performance, including NN and IDW, without calibration parameters was evaluated using the goodness-of-fit indicators (MAE, R, and RMSE) based on the cross-validation scheme. Considering the observed data from the entire basin, the performance of all seven spatial interpolation methods is shown in boxplots in Figure 6. The mixed methods (TSA-OK and BMA) performed best at the daily,
monthly, and annual time scales. BMA encompassed five individual methods, while TSA-OK encompassed only two single methods. Their performance was similar considering their MAE, R, and RMSE values at the monthly and annual time scales. The BMA weights at the daily scale (shown in Figure 5) were more dispersive than at the monthly and annual scales. It can be seen that the performance of BMA combined with seven methods was better than TSA-OK at the daily scale. The more dispersive the BMA weights, the more accurate the results from the methods that combined one or more approaches. If the weights of every single method in BMA were equal or similar, there was no significant difference between the mixed methods that were combined with various numbers of single methods. As the spatial precipitation correlations for short periods (daily) varied more than those for longer periods (monthly and annually), it was determined that a mixed

![Figure 3](image_url)  
Figure 3 | The values of MAE, R and RMSE with the variations of the numbers of rainfall stations N and the α values at (a) daily, (b) monthly and (c) annual time scales.

| Table 3 | The determined parameters for the IDW spatial interpolation methods |
| Parameter | daily | monthly | annual |
| N | 4 | 6 | 6 |
| α | 1.6 | 1.8 | 1.6 |
Figure 4 | The performances boxplots of the three semivariogram models for OK and CK methods at 53 precipitation gauge sites: (a) daily, (b) monthly and (c) annual. The whiskers of the boxplots are drawn to the 90th and 10th percentiles of very metrics. The center line (red line in the rectangle) is the median and the dot is the average values. The rectangle is defined by the 75th and 25th percentiles as upper and lower hinges (same meanings in the following boxplots Figures 5 and 6).
method with more single models was more accurate than a mixed method with fewer ones. Although the best performance of the mixed method might be partly due to the different number of parameters, the input for every method was just the precipitation series, and recommendation of a mixed method will still be useful in practice.

Of the single methods evaluated, the best performance was from IDW. OK was better than CK, TSA was better than NN and CK, and OK was better than TSA at the monthly and annual time scales, while the former was worse than the latter at the daily scale. The spatial variation of the monthly precipitation was higher than that of the annual precipitation. The values of $R$ increased from the daily, monthly, and annual time scales, while their variation ranges decreased, which was different from the values of $R$ at each site. The performance of each single method was not significantly different at the monthly and annual scales from the perspective of the entire study area. The annual data compensated for the possible differences in daily precipitation amounts, as in the case of moving storms, which was one reason to improve the interpolation methods. Therefore, while any method could be recommended to estimate the monthly or annual spatial precipitation distribution, IDW was the primary recommendation.

The performance of all seven methods was also examined at 53 gauge stations. At a given station, the larger MAE, $R$, and RMSE values are shown in darker red, while the smaller values are shown in darker blue, as seen in Figure 7. According to the gauge station distribution shown in Figure 2, the better performance locations were

### Table 4
The average statistical criteria values of the semivariogram models for OK and CK spatial interpolation methods

| Semivariogram models | Statistical criteria | OK | CK |
|----------------------|----------------------|----|----|
|                      | daily                | monthly | annual | daily | monthly | annual |
| Gaussian             | MAE (mm)            | 2.76 | 27.32 | 164.55 | 2.95 | 28.33 | 173.39 |
|                      | $R$                 | 0.44 | 0.67 | 0.77 | 0.40 | 0.66 | 0.70 |
|                      | RMSE (mm)           | 4.26 | 37.37 | 212.56 | 4.48 | 38.38 | 242.65 |
| Spherical            | MAE (mm)            | 2.74 | 27.28 | 164.48 | 2.90 | 28.12 | 173.27 |
|                      | $R$                 | 0.45 | 0.68 | 0.78 | 0.42 | 0.67 | 0.72 |
|                      | RMSE (mm)           | 4.25 | 37.34 | 212.24 | 4.41 | 38.32 | 242.55 |
| Exponential          | MAE (mm)            | 2.77 | 27.31 | 165.48 | 2.97 | 28.23 | 173.31 |
|                      | $R$                 | 0.45 | 0.67 | 0.75 | 0.41 | 0.67 | 0.71 |
|                      | RMSE (mm)           | 4.27 | 37.39 | 212.59 | 4.46 | 38.40 | 243.08 |

| Figure 5 | BMA weights boxplots reflecting the matching degrees for NN, IDW, ISA, OK, CK spatial interpolation methods: (a) daily; (b) monthly; (c) annual. |
Figure 6 | The performances boxplots of selected seven precipitation interpolation methods for study area at (a) daily; (b) monthly; (c) annual scales.
found downstream of the study area where the elevation was lower. Estimations for the upstream sites from the 3rd to the 30th gauge stations were not as good as those downstream. As the precipitation gauge station density was low at the border of the study area (see Figure 3), the poor performance at these sites can be partly attributed to the lack of precipitation gauges with similar characteristics. Considering the $R$ values, the performance of each method at the monthly and annual time scales improved compared to the daily time scale (e.g., No. 15 gauge site). However, the $R$ values at the monthly scale were better than those at the annual scale for most sites. As precipitation in the study area comes mainly from monsoons, the spatial correlations between sampling sites were strong and beneficial to spatial interpolation. The performance of each method at some sites was poor for every time scale (e.g., No. 17 gauge site).

In order to directly compare the performance of each method at a given gauge site, the worst method was given a value of ‘one’ according to the MAE, $R$, and RMSE values at a particular time scale, while the best method was given a value of ‘seven’. Applying these ‘points’ to each gauge site, Figure 7 is transformed into Figure 8. It can be easily seen that the performance of each method was different for different time scales. The mixed methods (TSA-OK and BMA) performed best at the daily scale, while CK, OK, and NN performed worst among the methods. As the time scales increased to the monthly and annual scales, no method with significant strengths emerged. However, some models demonstrated good performance at some sites, but poor performance at other sites. For example, BMA for monthly precipitation produced the lowest MAE and RMSE values at the second gauge site, but produced higher values at the 17th site. The rank score color as determined by the $MAE$, $R$, and $RMSE$ values were not always the same at a given gauge site for a particular time scale. It should be noted that the performance of the NN method was high at the monthly and annual time scales, while it demonstrated poor performance at the daily time scale.

**Performance of hydrological models driven by precipitation at different time scales**

The performance of the SWAT, ‘abcd’, and Budyko hydrological models driven by the interpolated precipitation were evaluated using the observed discharge data at the Yichang gauge station from 1960 to 2008 at the daily, monthly, and annual time scales. The ranks of the NSE values for the discharge simulation at the daily, monthly, and annual time scales were the same as that of the $R$ values for the spatial interpolation methods at the 53 gauge stations (shown in Table 5 and Figure 7). However, the ranks of $PBIAS$ were consistent with those of the $R$ values with respect to the entire study area (shown in Table 5 and Figure 6). The highest NSE values were obtained from the ‘abcd’ model at the
monthly scale for the calibration and validation periods. The lowest PBIAS values were obtained from the Budyko model at the annual scale for the calibration and validation periods. Since the distribution of precipitation at the monthly scale impacts the discharge at the Yichang gauge station, the annual discharge was primarily determined by the accuracy of the interpolated precipitation over the entire study area. A more favorable NSE value indicated that the BMA spatial interpolation method provided better results except for the modeled runoff at the daily time scale for the calibration period. Additionally, the ranks of the performance of the discharge simulation at the monthly and annual time scales were consistent with the ranks of precipitation estimation with respect to the NSE values for the calibration and validation periods. The differences in the discharge simulation at the daily scale, driven by the spatially interpolated precipitation, were significant with respect to the NSE or PBIAS values (shown in Table 5). The range of their differences was 0.05–0.06 for NSE (which accounts for nearly 10% of the average NSE value) and 5.50%–5.76% for PBIAS (which accounts for nearly 50% of the average PBIAS value). However, the differences at the monthly and annual time scales were smaller. If the metrics NSE and PBIAS were determined for the total period from 1960 to 2008, BMA would be found to have the best performance for the discharge simulation at the three time scales. However, the ranks of the discharge performance driven by another spatial precipitation interpolation scheme were different from those of spatial precipitation based on cross-validation techniques. It can also be inferred from the discharge simulation shown in Table 5 that the sample splitting for calibration and validation led to uncertainties in the discharge simulation.

 Scatterplots depicting the observed discharge versus the simulated discharge at different time scales for low discharges exhibited a closer linear relationship than at high discharges. In addition, the simulated discharge was lower than the observed at the daily and monthly time scales for the calibration and validation periods. The observed vs. simulated discharge plots driven by the NN spatially interpolated precipitation at the daily scale for the validation period in Figure 9(a2) showed more concordance between the simulated and observed values than in the calibration period (Figure 9(a1)). The spatial precipitation relationships between each gauge station at the daily scale were not as strong as at the monthly and annual time scales. The land surface conditions, which consider such things as water storage in the soil, contributed to the flow generation and routing. Therefore, the performance of the discharge simulation at the daily scale was not exactly the same as the corresponding precipitation estimate. The difference in large discharge values for the calibration period was more than in the validation period, while the NSE and PBIAS values for the calibration period were less than in the validation period (Table 5).
Some of the error associated with the results of the areal precipitation estimation and discharge simulation can be accounted for when looking at the entire study area. Given that the upstream portion of the study area had significantly fewer available sampling points than the downstream portion (Figure 2), the precipitation discrepancies become problematic, as precipitation for large regions that cover many different elevations was estimated with little input data. Finding accurate spatial relationships for precipitation between sparse and relatively dense gauge stations for different time scales that could improve spatial precipitation estimation and discharge simulation was the aim of this study.

It can be concluded that mixed methods, such as TSA-OK and BMA, performed best at the daily, monthly, and annual time scales. If the weights of every single method in BMA were equal or similar, there was not much difference between the mixed methods, which were combined with various single methods. As a single method, IDW produced low values for the MAE and RMSE metrics for all three time scales. However, it did not perform the same at all sites. As the time scale increased to monthly and annual, there were no methods with significant strengths. However, several models showed good performance at some sites, but poor performance at others. There was no method that performed well at every site in all three time scales. As the time scale increased, the differences in interpolated precipitation between each method shrunk. The selection of spatial interpolating methods for areal precipitation at small time scales should be given extra consideration.

The adopted hydrological models for the discharge simulation should be different from their time scales. SWAT based on SCS, abcd, and the Budyko hydrological models have been proven as unifying catchment water balance models for different time scales through the maximum entropy production principle (Wang et al. 2014; Zhao et al. 2015). Propagation of the uncertainty of spatial precipitation interpolation methods were comparable at different time scales. It was found that the uncertainty of spatial precipitation interpolation methods caused by the different spatial precipitation interpolation methods were comparable at different time scales. It was found that the uncertainty of spatial precipitation interpolation methods based on SCS, abcd, and the Budyko hydrological models have been proven as unifying catchment water balance models for different time scales (Wang et al. 2014; Zhao et al. 2016). Propagation of the uncertainty of spatial precipitation interpolation methods demonstrated a significant advantage in spatially interpolating precipitation at small time scales.

The uncertainty of spatial precipitation interpolation methods should be given extra consideration. The mixed methods demonstrated a significant advantage in spatially interpolating precipitation at small time scales. Table 5 shows the performances of hydrological models driven by differently interpolated precipitation inputs.

| Spatial precipitation interpolation scheme | SWAT (Daily) | abcd (Monthly) | Budyko (Annual) |
|-------------------------------------------|-------------|---------------|---------------|
|                                           | NSE         | PBIAS(%)      | NSE           | PBIAS(%)      | NSE           | PBIAS(%)      |
| NN                                        | 0.702       | 0.744         | 0.699         | 0.916         | 0.908         | 0.914         |
| IDW                                       | 0.749       | 0.729         | 0.828         | 0.936         | 0.916         | 0.930         |
| TSA                                       | 0.732       | 0.723         | 0.728         | 0.933         | 0.909         | 0.926         |
| OK                                        | 0.724       | 0.718         | 0.723         | 0.930         | 0.908         | 0.923         |
| CK                                        | 0.701       | 0.721         | 0.696         | 0.920         | 0.906         | 0.916         |
| TSA-OK                                    | 0.736       | 0.692         | 0.732         | 0.936         | 0.908         | 0.930         |
| BMA                                       | 0.753       | 0.684         | 0.851         | 0.939         | 0.916         | 0.932         |

* I is for the calibration period (1960–1992); II is for validation period (1993–2008) and III is for whole period (1960–2008).
with respect to the discharge simulation. The performance of the daily discharge simulation was not consistent with the performance of the daily precipitation interpolation, because factors such as water storage and initial water conditions influence the discharge simulation in addition to precipitation. It was typical for the discharge simulation to underestimate high discharge. Ultimately, there was no dominant spatial precipitation interpolation for the discharge simulation at small time scales (i.e., daily). Although the selection of spatial interpolation methods for areal precipitation estimation and discharge simulation at different scales have been proposed based on our case study, more case studies should be done for more general conclusion and guideline.

ACKNOWLEDGEMENT

The authors gratefully acknowledge the financial support from the National Natural Science Foundation of China (Nos. 51579183, 91647106, 51879194, and 51525902). This work is also partly funded by the Ministry of Foreign Affairs of Denmark and administered by Danida Fellowship Centre (File number: 18-M01-DTU).
REFERENCES

Aliyari, F., Bailey, R. T., Tasdighi, A., Dozier, A., Arabi, M. & Zeiler, K. 2019 Coupled SWAT-MODFLOW model for large-scale mixed agro-urban river basins. Environ. Modell. Software 115, 200–210. doi:10.1016/j.envsoft.2019.02.014.

Alley, W. M. 1984 On the treatment of evapotranspiration, soil moisture accounting, and aquifer recharge in monthly water balance models. Water Resour. Res. 20 (8), 1137–1149.

Arnold, J. G., Srinivasan, R., Muttiah, R. S. & Williams, J. R. 1998 Large area hydrologic modeling and assessment Part I: model development. J. Am. Water Resour. Assoc. 34 (1), 73–89. doi:10.1111/j.1752-1688.1998.tb05961.

ASCE 1996 Hydrology of Handbook, Task Committee on Handbook of Hydrology (Management Group D of ASCE), 2nd edn. American Society of Civil Engineers, ASCE, New York.

Bárdossy, A. & Pegram, G. 2014 Infilling missing precipitation records—a comparison of a new copula-based method with other techniques. J. Hydrol. 519, 1162–1170.

Bennet, N. D., Croke, B. F. W., Guariso, G., Guillaume, J. H. A., Hamilton, S. H., Jakeman, A. J., Marsili-Libelli, S., Newham, L. T. H., Norton, J. P., Perrin, C., Pierce, S. A., Robson, B., Seppelt, R., Voinov, A. A., Fath, B. D. & Andreassian, V. 2015 Characterising performance of environmental models. Environ. Modell. Software 40, 1–20.

Berndt, C. & Haberlandt, U. 2018 Spatial interpolation of climate variables in Northern Germany – influence of temporal resolution and network density. J. Hydrol.: Reg. Stud. 15, 184–202.

Blumer, M. G. 1979 Principles of Statistics. Dover, New York, p. 252.

Budyko, M. I. 1958 The Heat Balance of the earth's surface. translated from Russian by N.A. Stepanova. U.S. Dep. of Commer., Washington, DC, p. 259.

Bunham, K. P. & Anderson, D. R. 2002 Model Selection and Multi-Model Inference: A Practical Information-Theoretic Approach. Springer, Heidelberg.

Burrough, P. & Mc Donnell, R. 1998 Principles of Geographical Information Systems. Oxford University Press, New York.

Chayes, F. 1970 On deciding whether trend surfaces of progressively higher order are meaningful. Geol. Soc. Am. Bull. 81, 1273–1278.

Chen, F. W. & Liu, C. W. 2012 Estimation of the spatial rainfall distribution using inverse distance weighting (IDW) in the middle of Taiwan. Paddy Water Environ. 10 (3), 209–222.

Chen, Y., Yuan, W. P., Xia, J. Z., Fisher, J. B., Dong, W. J., Zhang, X. T., Liang, S. L., Ye, A. Z., Cai, W. W. & Feng, J. M. 2015 Using Bayesian model averaging to estimate terrestrial evapotranspiration in China. J Hydrol. 528, 537–549.

Chen, Y., Xu, C. Y., Chen, X. W., Xu, Y. P., Yin, Y. X., Gao, L. & Liu, M. B. 2019 Uncertainty in simulation of land-use change impacts on catchment runoff with multi-timescales based on the comparison of the HSPF and SWAT models. J Hydrol. 573, 486–500.

Ding, Q., Wang, Y. & Zhuang, D. F. 2018 Comparison of the common spatial interpolation methods used to analyze potentially toxic elements surrounding mining regions. J. Environ. Manage. 212, 23–31.

Dirks, K., Hay, J., Stow, C. & Harris, D. 1998 High-resolution studies of rainfall on Norfolk Island: part II: interpolation of rainfall data. J. Hydrol. 208 (3–4), 187–193.

Duan, Q. & Phillips, T. J. 2010 Bayesian estimation of local signal and noise in multimodel simulations of climate change. J. Geophys. Res. 115 (D18), D18123. http://dx.doi.org/10.1029/2009JD013654.

Fernandez, W., Vogel, R. M. & Sankarasubramanian. 2000 A. Regional calibration of a watershed model. Hydrol. Sci. J. 45 (5), 689–707.

Foehn, A., Hernández, J. G., Schaefl, B. & De Cesare, G. 2018 Spatial interpolation of precipitation from multiple rain gauge networks and weather radar data for operational applications in Alpine catchments. J Hydrol. 563, 1092–1110.

Fu, C. S., James, A. L. & Yao, H. X. 2015 Investigations of uncertainty in SWAT hydrologic simulations: a case study of a Canadian Shield catchment. Hydrol Process. 29 (18), 4000–4017.

Goovaerts, P. 2000 Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall. J Hydrol. 228, 113–129.

Goudenhoofd, E. & Delobbe, L. 2009 Evaluation of radar-gauge merging methods for quantitative precipitation estimates. Hydrol. Earth Syst. Sci. 13, 195–203.

Hagedorn, R., Doblas-reyes, F. J. & Palmer, T. N. 2005 The rational behind the success of multi-model ensembles in seasonal forecasting-I. Basic concept. Tellus 57A, 219–233.

Hattermann, F., Krysanova, V., Wechsung, F. & Wattenbach, M. 2005 Runoff simulation on the macroscale with the ecohydrological model SWIM in the Elbe catchment—validation and uncertainty analysis. Hydrol. Process. 19, 693–714.

Hjelmfelt, A. T. 1991 Investigation of curve number procedure. J. Hydrol. Eng. 17 (6), 725–735.

Isaaks, E. H. & Srivastava, R. M. 1989 An Introduction to Applied Geostatistics. Oxford University Press, New York, pp. 351–368.

Krumbein, W. C. 1959 Trend surface analysis of contour-type maps with irregular control-point spacing. J. Geophys. Res. 64 (7), 823–834.

Li, J. & Heap, A. D. 2011 A review of comparative studies of spatial interpolation methods in environmental sciences: Performance and impact factors. Ecol. Inform. 6 (3–4), 228–241.

Martinez, G. F. & Gupa, H. 2010 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. doi:10.1029/2009WR008294.
McLachlan, G. J. & Krishnan, T. 1997 The EM Algorithm and Extensions. Wiley, New York.

Moriase, D. N., Arnold, J. G., Van Liew, M. W., Bingner, R. L., Harmel, R. D. & Veith, T. L. 2007 Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Trans. ASABE 50 (3), 885–900.

Nash, J. E. & Sutcliffe, J. V. 1970 River flow forecasting through conceptual models part I – a discussion of principles. J. Hydrol. 10, 282–290.

Neitsch, S. L., Arnold, J. G., Kiniry, J. R. & Williams, J. R. 2005 Soil and Water Assessment Tool Theoretical Documentation, Version 2005.

Nikolopoulos, E. I., Borga, M., Creutin, J. D. & Marra, F. 2015 Estimation of debris flow triggering rainfall: influence of rain gauge density and interpolation methods. Geomorphology 243, 40–50.

Pan, S. H., Liu, D. D., Wang, Z. L., Zhao, Q., Zou, H., Hou, Y. K., Liu, P. & Xiong, L. H. 2017 Runoff responses to climate and land use/cover changes under future scenarios. Water 9 (7), 475. https://doi.org/10.3390/w9070475.

Pardo-Iguzquiza, E., Rodríguez-Galiano, V. F., Chica-Olmo, M. & Atkinson, P. M. 2011 Image fusion by spatially adaptive filtering using downsampling cokriging. ISPRS. J. Photogramm. 66, 337–346.

Parker, S. R., Adams, S. K., Lammers, R. W., Stein, E. D. & Bledsoe, B. P. 2009 Targeted hydrologic model calibration to improve prediction of ecologically-relevant flow metrics. J. Hydrol. 573, 546–556.

Plouffe, C. C., Robertson, C. & Chandrapala, L. 2015 Comparing interpolation techniques for monthly rainfall mapping using multiple evaluation criteria and auxiliary data sources: a case study of Sri Lanka. Environ. Modell. Software 67, 57–71.

Price, D. T., McKenney, D. W., Nalder, I. A., Hutchinson, M. F. & Kesteven, J. L. 2000 A comparison of two statistical methods for spatial interpolation of Canadian monthly mean climate data. Agric. Meteorol. 101 (2–3), 81–94.

Rafferty, A. E., Gneiting, T., Balabdaoui, F. & Polakowski, M. 2005 Using Bayesian models averaging to calibrate forecast ensembles. Mon. Weather Rev. 133 (5), 1155–1174. http://dx.doi.org/10.1175/MWR2906.1.

Sankarasubramanian, A. & Vogel, R. M. 2002 Annual hydroclimatology of the United States. Water Resour. Res. 38 (6), 1083. doi:10.1029/2001WR000619.

Schneider, M. P. A. & Yaşar, Y. 2016 Is inequality deadly and for whom? a Bayesian model averaging analysis. The Social Science J. 53, 357–370.

Soares, T. N., Diniz-Filho, J. A. F., Nabout, J. C., Telles, M. P. C., Terribile, L. C. & Chaves, L. J. 2015 Patterns of genetic variability in central and peripheral populations of Dipteryx alata (Fabaceae) in the Brazilian Cerrado. Plant Syst. Evol. 301 (5), 1315–1324.

Sullivan, D. O. & Unwin, D. J. 2003 Geographical Information Analysis. John Wiley & Sons, Inc., N.J.

Teegavarapu, R. S. V. & Elshorbagy, A. 2005 Fuzzy-set-based error measure for hydrologic model evaluation. J. Hydrol. 7, 199–208.

Teegavarapu, R. S. V., Meskele, T. & Pathak, C. S. 2012 Geo-spatial grid-based transformations of precipitation estimates using spatial interpolation methods. Comput. Geosci. 40, 28–39.

Thiessen, A. H. 1911 Precipitation averages for large areas. Mon. Weather Rev. 39 (7), 1082–1084.

Thomas, H. A. 1981 Improved Methods for National Water Assessment: Final Report. U.S. Geol. Surv. Water Resour. Contract WR15249270, p. 44.

Vicente-Serrano, S. M., Saz-Sánchez, M. A. & Cuadrat, J. M. 2003 Comparative analysis of interpolation methods in the middle Ebro Valley (Spain): application to annual precipitation and temperature. Clim. Res. 24, 161–180.

Vieux, B. E. 2001 Distributed Hydrologic Modeling Using GIS, Water Science and Technology Library. Kluwer Academic Publishers.

Wagner, P. D., Fiener, P., Wilken, F., Kumar, S. & Schneider, K. 2012 Comparison and evaluation of spatial interpolation schemes for daily rainfall in data scarce regions. J Hydrol. 464–465, 388–400.

Wang, F. 2006 Quantitative Methods and Applications in GIS. CRC Press, Boca Raton, Florida, USA.

Wang, D. & Tang, Y. 2014 A one-parameter Budyko model for water balance captures emergent behavior in Darwinian hydrologic models. Geophys. Res. Lett. 41, 4569–4577.

Wong, K. W., Wong, P. M., Gedeon, T. D. & Fung, C. C. 2003 Rainfall prediction model using soft computing technique. Soft. Comput. 7, 434–438.

Wu, C. F. J. 1985 On the convergence properties of the EM algorithm. Ann. Stat. 11, 95–103.

Wu, D., Cui, Y. L., Xie, X. H. & Luo, Y. F. 2019 Improvement and testing of SWAT for multi-source irrigation systems with paddy rice. J Hydrol. 568, 1031–1041. https://doi.org/10.1016/j.jhydrol.2018.11.057.

Yang, H. B., Yang, D. W., Lei, Z. D. & Sun, F. B. 2008 New analytical derivation of the mean annual water-energy balance equation. Water Resour. Res. 44 (3). http://dx.doi.org/10.1029/2007WR006135.

Zhang, X. & Srinivasan, R. 2009 GIS-based spatial precipitation estimation: a comparison of geostatistical approaches. J. Am. Water Resour. Assoc. 45 (4), 894–906.

Zhang, J. & Taflanidis, A. A. 2019 Bayesian model averaging for Krige regression structure selection. Probabilist. Eng. Mech. 56, 58–70. https://doi.org/10.1016/j.probengmech.2019.02.002.

Zhao, J., Wang, D., Yang, H. & Sivapalan, M. 2016 Unifying catchment water balance models for different time scales through the maximum entropy production principle. Water Resour. Res. 52. https://doi.org/10.1002/2016WR018977.

Zimmerman, D., Pavlik, C., Ruggles, A. & Armstrong, M. P. 1999 An experimental comparison of ordinary and universal kriging and inverse distance weighting. Math. Geol. 31 (4), 375–390.

First received 17 September 2019; accepted in revised form 4 March 2020. Available online 29 June 2020.