Hydro-stochastic interpolation coupling with Budyko approach for prediction of mean annual runoff

Ning Qiu[a,b], Xi Chen[d,a,b]*, Qi Hu[c], Jintao Liu[a,b], Richao Huang[a,b], Man Gao[a,b]

[a] *State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering Hohai University, Nanjing 210098, China*
[b] *College of Hydrology and Water Resources, Hohai University, Nanjing 210098, China*
[c] *School of Natural Resources, University of Nebraska-Lincoln, Lincoln NE 68583, U.S.*
[d] *Institute of Surface-Earth System Science, Tianjin University, Tianjin China*
**Corresponding author E-mail: xichen@hhu.edu.cn*

Abstract

Hydro-stochastic interpolation method based on traditional block-kriging has often been used to predict mean annual runoff in river basins. A caveat in this method is that the statistic technique provides little physical insight on relationships between the runoff and the external forcings, such as climate and landscape. In this study, the spatial runoff is decomposed into a deterministic trend and stochastic fluctuations describing deviations from it. The former is described by the Budyko method (Fu's equation) and the latter by hydro-stochastic interpolation. The coupled method of stochastic interpolation and the Budyko method is applied to spatially interpolate runoff in the Huaihe River Basin of China, after dividing it into 40 sub-basins. Results show that the coupled method significantly improves the accuracy of predicted mean annual runoff. The error of the predicted runoff from the coupled method is much smaller than that from the Budyko method and the hydro-stochastic interpolation method. The determination coefficient for cross-validation, R_{cv}^2, from the coupled method is 0.93, much larger than 0.81 from the Budyko method and 0.54 from the hydro-stochastic interpolation. Further comparisons indicate that the coupled method also has improved the problem of overestimating low runoff and underestimating high runoff suffered by the other two methods. These results support that the coupled method offers an effective and accurate way to predict mean annual runoff in river basins.

Keywords: Coupled Budyko method and hydro-stochastic interpolation; mean annual runoff; prediction accuracy; Huaihe River Basin

1. Introduction

Runoff observed at the outlet of a basin is a crucial element for investigating the hydrological cycle of the basin. The runoff is influenced by both deterministic and stochastic processes. Estimating the spatial patterns of runoff and associated distribution of water resources in ungauged basins has been one of the key problems in hydrology (Sivapalan et al., 2003) and a thorny issue in water management and planning (Imbach, 2010; Greenwood et al., 2011).

In estimating and predicting regional water resources availability, regional or global runoff mapping using geostatistical interpolation method has often been applied. In such geostatistical approaches, the value of a regional variable at a given location is estimated as a weighted average of observed values at surrounding locations. The spatial interpolation of runoff, assumed as an auto-correlated generalized stochastic field (Jones, 2009), uses secondary information often referring to more than one variable (Li and Heap, 2008). Spatial autocorrelation is measured by the covariance or semi-variance between pairs of points as a function of their Euclidian distance (such as in the ordinary kriging). The values obtained by geostatistical or kriging interpolation methods are the best linear unbiased estimate in the sense that the expected bias is zero and the kriging mean squared error is minimized (Skøien et al., 2006). The ordinary kriging (OK) estimates the local mean as constant, and corresponding residuals are considered as random. Because the spatial mean could also be used as a trend or nonstationary variation in space, OK has been further developed into various geostatistical interpolation methods, such as kriging with a trend by incorporating local trend within the neighborhood search window as a smoothly varying function of the coordinates. Block kriging (BK) is an extension of OK for estimating a block value instead of a point value by replacing the point-to-point covariance with the point-to-block covariance (Wackernagel, 1995).

Unlike precipitation or evaporation which we often interpolate to find its values at specific points in space, runoff is an integrated spatially continuous process in basins (Lenton and RodriguezIturbe, 1977; Creutin and Obled, 1982; Tabios and Salas, 1985; Dingman et al., 1988; Barancourt et al., 1992; Bloschl, 2005). Streamflow shows some degrees of natural organization and connection of water basins (Dooge, 1986; Sivapalan, 2005), e.g., rivers that connect sub-basins. The river network constraints water balance between upstream and downstream in a basin. The hierarchically organized river network requires that the sum of the interpolated discharge from the sub-basins equals to the observed runoff at the outlet of the entire basin. Previous studies have indicated that runoff interpolation may overestimate the actual runoff without adequate spatial variation information of runoff (Arnell, 1995), e.g., neglecting the river network in connecting sub-basins or processing basin runoff behavior as “points" in space (Villeneuve et al, 1979; Hisdal and Tveito, 1993). Given the obvious nested structure of basins, Gottschalk (1993a and b) developed a hydro-stochastic approach for runoff interpolation. It takes a full account of the concept that runoff is an integrated course in the hierarchical structure of river network. Distance between a pair of basins is measured by geostatistical distance instead of the Euclidian distance. The covariogram among points in conventional spatial interpolation is replaced by covariogram between basins.

In this concept, runoff is considered spatially homogeneous over basins, i.e., the expected value of runoff is a constant in space (Sauquet, 2006).

The observed patterns of runoff reveal systematic deviations from homogeneity due to influences, such as heterogeneous rainfall. We can describe the hydrological variables of interest in deterministic forms of functions, curves or distributions, and construct conceptual and mathematical models to predict hydro-climate variability (Wagener et al, 2007). For example, Qiao (1982), Arnell (1992) and Gao et al. (2017) used such approach and derived empirical relationships between runoff and its controlling factors of climate, land-cover and topography in various basins. However, the deterministic method in describing complex runoff patterns suffers inevitable loss of information (Wagener et al, 2007) because of existence of uncertainty in many hydrological processes and especially in observations. Thus, hydrological variables also contain information of stochastic nature and should be treated as outcomes of both deterministic and stochastic processes. Recently, the method of kriging with an external drift (KED) was introduced (Goovaerts, 1997; Li and Heap, 2008; Laaha et al., 2013). It accounts for deterministic patterns of spatial variable and also incorporates the local trend within the neighborhood search window as a linear function of a smoothly varying secondary variable, instead of a function of the spatial coordinates.

The inclusion of deterministic terms in the original geostatistical methods has been shown to increase interpolation accuracy of basin variables, such as mean annual runoff (Sauquet, 2006), stream temperature (Laaha et al., 2013) and groundwater table (Holman et al., 2009). Those deterministic terms are often described by empirical formulae linking spatial features, e.g., variability of mean annual runoff in elevation (Sauquet, 2006), and relationship between the mean annual stream temperature and the altitude of gauges (Laaha et al., 2013). As a semi-empirical approach for modelling the deterministic process of runoff, the Budyko framework has been popularly used to analyze relationship between mean annual runoff and climatic factors, e.g., aridity index (Milly, 1994; Koster and Suarez, 1999; Zhang et al., 2001; Donohue et al., 2007; Li et al., 2013; Greve et al., 2014). Many efforts have also been devoted to improving the Budyko method by including effects of other external forcing factors, such as land-use and land-cover (Donohue et al., 2007; Li et al., 2013; Han et al., 2011; Yang et al., 2007), soil properties (Porporato et al., 2004; Donohue et al., 2012), topography (Shao et al., 2012; Xu et al., 2013; Gao et al., 2017), hydro-climatic variations of seasonality (Milly, 1994; Gentine et al., 2012; Berghuijs et al., 2014) and groundwater levels (Istanbulluoglu et al., 2012). However, it has been found that use of the deterministic equation in the Budyko method alone still comes with large errors in prediction of runoff in many basins/areas (e.g., Potter and Zhang, 2009; Jiang et al., 2015).

The aim of this study is to combine the stochastic interpolation with semi-empirical Budyko method to improve spatial interpolation/prediction of mean annual runoff in the Huaihe River Basin (HRB), China. In this study, the spatial runoff from sub-basins in the HRB is separated into the deterministic trend and its residuals both of which are estimated by the Budyko framework and interpolation method. The residuals that are calculated as difference between the observed and the estimated runoff from the Budyko method, are used in the hydro-stochastic interpolation as described in Gottschalk (1993a, 1993b, and 2000). After that, the runoff of any sub-basin is predicted as the sum of the interpolated residual and the Budyko estimated value. The improved method is tested in the HRB. For comparison, the leave-one-out cross-validation approach was applied to evaluate performance of the three interpolation methods: the Budyko method, hydro-stochastic interpolation, and our coupled Budyko and hydro-stochastic interpolation method.

2. Methodologies

2.1 Spatial estimation of mean annual runoff by Budyko method

The Budyko method explains the variability of mean annual water balance on a regional or global scale. It describes dependence of actual evapotranspiration (*E*) on precipitation (*P*) and potential evapotranspiration (E_0) (Williams et al., 2012). The original relationship ($E/P \sim E_0/P$) derived by Budyko (1974) is deterministic and nonparametric. It was later developed into parametric forms (Fu, 1981; Choudhury, 1999; Yang et al., 2008; Gerrits et al., 2009; Wang and Tang, 2014). Among all the parametric forms of Budyko curves, the one-parameter equation derived by Fu (Fu, 1981, Zhang et al. 2004) has been used frequently. This equation is written as

$$\frac{E}{P} = 1 + \frac{E_0}{P} - \left(1 + \left(\frac{E_0}{P}\right)^{\omega}\right)^{\frac{1}{\omega}} \quad (1)$$

or

$$R = P \cdot \left(1 + \left(\frac{E_0}{P}\right)^{\omega}\right)^{\frac{1}{\omega}} - E_0 \quad (2)$$

where, P, E, E_0, and R are mean annual precipitation, actual evapotranspiration, potential evapotranspiration, and runoff (units: mm), respectively, and ω is a dimensionless model parameter within the range of $(1, \infty)$. In these formulae, the larger the ω is, the smaller the partition of precipitation into runoff.

The parameter ω can be calculated using observed P, E_0 and R in gauged sub-basins. The mean value of ω of a basin can be obtained by averaging ω of sub-basins, or by best fitting the curve in Eq. (1) with $E/P \sim E_0/P$ ($E = P - R$) from sub-basin observations by minimizing the mean absolute error (*MAE*) (Legates and McCabe, 1999). Using the mean value of ω, Eq. (2) can be used to predict ungauged basin runoff or to interpolate spatial variation of runoff, using meteorological data in targeted sub-basins (Parajka and Szolgay, 1998).

2.2 Hydro-stochastic interpolation method

Gottschalk (1993a) described the hydro-stochastic interpolation method for spatial prediction of runoff based on the kriging method. Gottschalk's method redefines a relevant distance between basins and identifies the river network and supplemental water balance constraints as follows.

As a spatially integrated continuous process, the predicted runoff of a specific unit of an area A_0 in a basin can be expressed as

$$r^*(A_0) = \sum_{i=1}^{n} \lambda_i r(A_i) = \Lambda^T R \quad (3)$$

where, $r^*(A_0)$ is the predicted runoff of that unit, $r(A_i)$ is the observed runoff in a gauged basin i with an area A_i ($i = 1, \ldots n$, n is the total number of gauged basins), λ_i is the weights of a gauged basin i, and Λ is the transposed column vector of the weights, and R is the column vector of runoff $r(A_i)$.

Because $r^*(A_0)$ is an estimate of the true value $r(A_0)$, the best linear unbiased estimate should satisfy $E[\,r^*(A_0) - r(A_0)\,] = 0$, in which E is the expected value. If the runoff is taken as a point process at the location of interest u_0, to achieve the goal of minimizing the error of estimation for a point process, the following set of equations has been developed to solve for the optimal weights under the second order stationary assumption for hydrologic variables (Ripley, 1976),

$$\begin{cases} \sum_{j=1}^{n} \lambda_i Cov(u_i, u_j) + \mu = Cov(u_i, u_0), \quad i, j = 1,2,\ldots,n \\ \sum_{i=1}^{n} \lambda_i = 1\,. \end{cases} \tag{4}$$

In the above, $Cov(u_i, u_j)$ is the theoretical covariance function between each pair of gauged stations (i=1,…, n, j=1,2…, n), and $Cov(u_i, u_0)$ is the theoretical covariance of runoff between the location of interest u_0 and each of the gauged stations u_i, μ is the Lagrange multiplier. After calculating the weights, λ_i, and substituting them into Eq. (3), we can solve for $r^*(A_0)$.

According to Sauquet et al. (2000), a basin consisting of n sub-basins with areas A_j (j = 1, ..., n) and observations of runoff can be further divided into M non-overlapping sub-basins with areas ΔA_i. Those M sub-basins can be used as the fundamental units in hydro-stochastic interpolation. The sum of the interpolated runoff for each non-overlapping sub-basin should be equal to the observed runoff at the river outlet.

This constraint can be written as

$$R_T = \sum_{i=1}^{M} \Delta A_i \, r(\Delta A_i) \tag{5}$$

where, R_T is the streamflow observed at the outlet of the basin, ΔA_i is the non-overlapping area of sub-basin i, $r(\Delta A_i)$ is the runoff depth for sub-basin i ($i = 1, \ldots, M$). The runoff prediction for each ΔA_i is a linear combination of weights and runoff observations in the *n* sub-basins, i.e., $r(\Delta A_i) = \sum_{j=1}^{n} \lambda_j^i r(A_j)$. Substituting it in Eq. (5) we can get

$$R_T = \sum_{i=1}^{M} \Delta A_i \left(\sum_{j=1}^{n} \lambda_j^i r(A_j)\right) = \sum_{i=1}^{M} n_i a \left(\sum_{j=1}^{n} \lambda_j^i r(A_j)\right) \quad (6)$$

In Eq. (6), $r(A_j)$ is the runoff depth for sub-basins *j* (*j* = 1, ..., *n*) with discharge observations, and λ_j^i is the weight (*i*= 1, ..., *M*; *j*= 1, ..., *n*).

Sauquet et al. (2000) divided the whole basin into n_T grids with equal area a. The discharge data are converted into runoff depth under the assumption that the runoff distribution across each basin is uniform. Thus, $R_T = n_T a r_T$ for the runoff depth r_T at the outlet of the basin in Eq. (6).

Based on the constraint of Eqs. (5) and (6) and considering basin areas of the river network, Sauquet et al. (2000) derived the weight matrices and described a hydro-stochastic method to optimize the weights λ_j^i (*i*= 1, ..., *M*; *j*= 1, ..., *n*) in Eq. (6). Their interpolation is calculated on multiple *M* non-overlapping sub-basins simultaneously.

To develop the theoretical covariance function and weight matrices, the key step is to define the distance between pairs of sub-basins from the identified runoff hierarchical structure in the river network. The appropriate geostatistical distance between sub-basins *A* and *B* defined by Gottschalk (1993b) is expressed as the expectation of distances of all the possible pairs of points inside *A* and *B*, i.e.,

$$d(A, B) = \frac{1}{AB} \int\int_{AB} ||u_1 - u_2|| du_1 du_2 \quad (7)$$

where, A and B are the areas of sub-basins A and B, respectively, u_1 and u_2 are the locations of pairs of points inside basins A and B, du_1 and du_2 are the differential symbol of u_1 and u_2, respectively.

The theoretical covariogram, $Cov(A, B)$, is derived in a similar way as geostatistical distance by averaging the point process covariance function Cov_p

$$Cov(A, B) = \frac{1}{AB} \int \int_{AB} Cov_p(||u_1 - u_2||) du_1 du_2 \qquad (8)$$

where, $Cov_p(||u_1 - u_2||)$ is the theoretical covariance function value of pairs of points in basins A and B and with distance $||u_1 - u_2||$.

In Eq. (8), the geostatistical distance $d(A, B)$ between A and B is calculated based on grid division in each of the sub-basins (Sauquet et al., 2000). We can obtain the mean distance d between all possible pairs of points (point at the center of a grid) in sub-basins A and B. For n sub-basins with observations, there are $n(n+1)/2$ pairs of the sub-basins with the mean distance d_i (i=1, …, $n(n+1)/2$).

Corresponding to the mean distance d_i between pairs of sub-basins, the empirical covariogram $Cov_e(d_i)$ can be calculated using the runoff depth of pairs of the sub-basins. The geostatistical distances d_i are then divided into fixed intervals (50 km in this study) to calculate the mean of $Cov_e(d_i)$ within each of the distance interval. Finally, the mean of $Cov_e(d_i)$ vs. the geostatistical distances d_i is used to draw a scatter diagram of the empirical covariogram $Cov_e(d_i)$.

The trial-and-error fitting method is used to calibrate Cov_p(d) in Eq. (8) aiming to best fit Cov_e(d). Only independent sub-basins are used to calculate the covariance function in order to avoid spatial correlation of nested sub-basins.

2.3 Coupling hydro-stochastic interpolation with Budyko method

The above stochastic interpolation procedure assumes a stationary stochastic variation of runoff among sub-basins or spatial homogeneity in runoff (Sauquet, 2006), despite variations in river network. For nonstationary variation of runoff resulting from spatial heterogeneity in a river network, the spatial runoff can be decomposed into nonstationary deterministic and stochastic components, i.e.,

$$R(x) = R_d(x) + R_s(x). \tag{9}$$

In (9), $R(x)$ is runoff at location x, $R_d(x)$ is the deterministic component of the spatial trend or the external drift (Wackernagel, 1995) that results in nonstationary variability in space. $R_s(x)$ is the stochastic component considered to be stationary.

In this study, R in Eq. (2) is used as an external drift function in estimating the deterministic component $R_d(x)$ in all sub-basins, i.e., $R_d(x)$ in Eq. (9) is substituted in Eq. (2) by setting $R_d(x) = R$. The residuals between $R_d(x)$ and observed runoff are calculated for all gauged sub-basins. Furthermore, these residuals are interpolated for all ungauged sub-basins and set as the stochastic component $R_s(x)$ in Eq. (9) using the "residual kriging" method (Sauquet, 2006). In particular, $R_s(x)$ in Eq. (9) is replaced by $r^*(A_0)$ in Eq. (3) by after setting $r^*(A_0) = R_s(x)$ for the hydro-stochastic interpolation scheme described in section 2.2. The superposition of these estimates of both components on the right-hand side in Eq. (9) yields the prediction of $R(x)$.

2.4 Cross validation

To validate the prediction procedure described above, we use the leave-one-out cross-validation method (Kearns, 1999). In addition, we examine and compare quantitatively the performances of our coupled model with the Budyko and the hydro-stochastic interpolation method. The performance of each method is evaluated by the following metrics (Laaha and Bloschl, 2006):

$$MAE = \frac{1}{n}\sum_{j=1}^{n}[R(x_i) - R^*(x_i)\ ] \quad (10)$$

$$MSE = \frac{1}{n}\sum_{j=1}^{n}[R(x_i) - R^*(x_i)\ ]^2 \quad (11)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{j=1}^{n}[R(x_i) - R^*(x_i)\ ]^2} \quad (12)$$

where, $R^*(x)$ and $R(x)$ are the predicted and observed runoff, respectively, *MAE* is mean absolute error, *MSE* is mean square error, and *RMSE* is the root-mean-square error. The determination coefficient for cross-validation is

$$\mathrm{R}_{cv}^2 = 1 - \frac{V_{cv}}{V_{NK}} \quad (13)$$

where, V_{cv} is the mean square error (*MSE*), and V_{NK} is the spatial variance ($V_{NK} = \frac{\sum_{j=1}^{n}[R(x_i) - \bar{R}\ ]^2}{n-1}$, in which $\bar{R}$ is the mean $R(x)$) of the runoff over all the tested sub-basins. In addition to these evaluation metrics, the prediction result is further evaluated by regression analysis of the observation vs. prediction.

3. Study catchment and data

The Huaihe River Basin (HRB), which is the sixth largest river basin in China, is used in evaluation of our coupled model and in comparison of it to the other two methods. HRB has a strong precipitation gradient from humid climate in the east and semi-humid in the west (Hu, 2008), and it is one of the major agricultural areas in China with the highest human population density in the country. Each year, millions of tons of water are consumed to meet the needs of the population and agriculture production. Water resources per capita and per unit area is less than one-fifth of the national average.

Moreover, more than 50% of the water resources is exploited, much higher than the recommended rate for international inland river basins (30%) (Yan et al, 2011). Intense precipitation occurring in a few very rainy months makes the region highly vulnerable to severe floods as well as droughts (Zhang et al., 2015). Thus, having the knowledge of spatial distribution of runoff is vital for water resources planning and management for the region.

Our study area is located upstream of the Bengbu Sluice in HRB and has a size of 121,000 km^2 (Fig. 1). The river network in the area is derived from data packages of the National Fundamental Geographic Information System, developed by National Geomatics Center of China. The study area is divided into 40 sub-basins, according to available hydrological stations with records from 1961-2000 (Fig. 2). Areas of the sub-basins vary from the smallest of 17.9 km^2 to the largest of 3,0630 km^2. Among the 40 sub-basins, there are 27 independent sub-basins and 13 nested sub-basins.

Annual precipitation data used in this study are from 1961-2000 and are obtained from a monthly mean climatological dataset at 0.5-degree spatial resolution. Those data were developed at China Meteorological Administration and are accessible at: http://data.cma.cn/data/detail/dataCode/SURF_CLI_CHN_PRE_MON_GRID_0.5.html. The dataset was derived from observations at 2472 stations in China, using Thin Plate Spline (TPS) interpolation method and the ANUSPLIN software. Pan evaporation data at 21 meteorological stations in HRB are used to interpolate E_0 using the ordinary kriging interpolation method and ArcGIS. The interpolated E_0 are used to derive the annual potential evapotranspiration in the sub-basins of HRB. The statistical features of mean annual precipitation (*P*), E_0 and runoff depth (*R*) in the period from 1961-2000 are summarized in Table 1. They show that over that period *P* varied from 638-1,629 mm, annual temperature varied from 11-16℃, and the mean annual E_0 varied between 900-1,200 mm. The sub-basins in the north are relatively dry with the dryness index (E_0/P) above 1.3 in, for example, the sub-basins ZM, ZQ, XY and ZK (Fig. 2). The sub-basins in the south are wetter with dryness index below 0.8 in sub-basins of MS, HBT and HC. The average mean annual *R* is about 400 mm, and fluctuating from a minimum of 90 mm in the northern region of the area to a maximum of 1000 mm in the south. The temporal and spatial variation in runoff is relatively small in the south and large in the north.

4 Results

4.1 Prediction of runoff by Budyko method

Actual evapotranspiration *E* (Table 1) is estimated using long-term mean annual water balance ($E=P-R$) from 1961–2000 at the 40 sub-basins (Table 1). Also shown in Table 1 is the calculated ω values for the sub-basins. They vary from 1.43 in the sub-basin HWH to 3.16 in JJJ. The average of ω is 2.32 for the 40 sub-basins. The actual vs. potential evapotranspiration in terms of E/P vs. E_0/P is shown in Fig. 3. The best fit (curve) for E/P vs. E_0/P, or *R* vs. E_0/P distribution by Eq. (1) or (2) is also shown in Fig, 3, and gives an alternative for average ω of the sub-basin. The fitted value of ω for the 40 sub-basins determined from this process is 2.213, very close to that calculated directly from the 40 individual sub-basins.

Using $\omega=2.213$ in our study basin, Fu's equation in Eq. (2) can be written as

$$R = P \cdot \left(1 + \left(\frac{E_0}{P}\right)^{2.213}\right)^{\frac{1}{2.213}} - E_0. \quad (14)$$

Eq. (14) and Fig. 3 clearly show the deterministic trend of runoff in the study basin. According to the water limit line of the arid edge at which $E = P$ and the energy limit line of the wet edge at which $E = E_0$ shown in Fig. 3a, the smaller the index $\frac{E_0}{P}$ is, the smaller the $\frac{E}{P}$ is (Fig. 3a) or the larger the runoff is (Fig. 3b) from the sub-basins in HRB. In Figs. 3b and 3c, the lower R in the north sub-basins indicates drier conditions (E_0/P>1.4) in those sub-basins, while the higher R in the south sub-basins indicates wetter conditions (E_0/P<0.8).

Using P and E_0 in the 40 sub-basins given in Table 1, the predicted runoff R by Eq. (14), the Budyko method, and the deviation of the prediction from the observation are calculated. The results are summarized in Tables 1 and 2. The *MAE* of predicted R is 94 mm, and *RMSE* is 112 mm. The largest absolute error is in sub-basin HWH (328 mm) and the smallest in sub-basin XX (24 mm). The largest relative error is 81.6% of the observed runoff in sub-basin XZ and the smallest is 5.0% of the observed runoff in XHD. They represent absolute errors of 91 and 37 mm in those two sub-basins, respectively.

4.2 Runoff from the hydro-stochastic interpolation method

For comparison, the observed runoff was used in the hydro-stochastic interpolation following the procedure detailed in section 2.2. In order to obtain the distance d between the sub-basin pairs, the study area is divided into 40 row × 50 column. According to Eq. (7), the geostatistical distance between any two sub-basins, A and B, is calculated by averaging the distances between all pairs of grid points in sub-basins A and B (all the possible sub-basin pairs are 40×41/2 for the 40 sub-basins in this study). According to the estimated distance for pairs of sub-basins and the observed runoff at the 40 sub-basins (Table 1), the empirical covariance $Cov_e(d)$ is estimated for each pair of the sub-basins. From plots of the mean $Cov_e(d)$ of the independent sub-basin pairs vs. the corresponding distances *d* with an interval of 50 km, we get an empirical covariogram that is shown in Fig. 4. The theoretical covariance function $Cov_p(d)$ fitting to the empirical covariogram is determined

$$Cov_p(d) = 6 \times 10^5 \exp(-d/28.62). \quad (15)$$

This function is further used in calculation of the average theoretical covariances $Cov(A,B)$ in Eq. (8). Subsequently, the weight matrices are determined using our program in MATLAB.

The interpolation results (R) over the 40 sub-basins along with the deviations from the observation are shown in Table 1. The *MAE* and *RMSE* of R are 134 mm and 176mm, respectively. The largest absolute error is in the sub-basin HWH (448 mm) and the smallest in XHD (3 mm) (Table 2). The largest relative error is 85.1% of the observed runoff in the sub-basin ZK, and the smallest is 0.4% of the observed runoff in XHD. They represent absolute error of 105 and 3 mm, respectively. These results indicate that the errors from this interpolation method are in general larger than those from the Budyko method, suggesting that the observed runoff is more influenced by the deterministic trend in the basins.

4.3 Hydro-stochastic interpolation with Fu's equation (our coupled method)

We use Fu's equation, Eq. (2), to evaluate the deterministic trend or the external drift function $R_d{}^*(x)$, and the deviation from the trend from the observation, $R_s{}^*(x)$, assuming a spatially auto-correlated process. The $R_s{}^*(x)$ is then used in the hydro-stochastic interpolation. The results are shown in Table 1.

The empirical covariogram of $R_s{}^*(x)$ for each pair of sub-basins versus sub-basin distances is shown in Fig. 5. From Fig. 5, we obtain the following exponential function for $Cov_p(d)$

$$Cov_p(d) = 3000 \exp(-d/48.34). \quad (16)$$

From Eq. (16), weight matrices of runoff deviation are determined using our program in MATLAB, and used to predict runoff deviation. Because this interpolation scheme represents the spatial runoff deviation, the sum of the interpolated runoff deviation and the simulated runoff by Fu's equation is the total interpolated runoff in the sub-basins.

The predicted runoff using this procedure is given in Table 1, with the *MAE* at 47 mm and *RMSE* at 69 mm over the 40 sub-basins. The largest absolute error is at the sub-basin HWH (236 mm) and the smallest at JJJ (2 mm) (Table 2). The largest relative error is 42.1% of the observed runoff at BB, and the smallest is 0.3% of the observed runoff from the sub-basin JJJ. They represent the absolute errors of 90 and 2 mm, respectively.

4.4 Comparisons of predicted runoff by the three methods

As shown in Table 2, our coupled method of the deterministic and stochastic processes described in this study significantly reduces the runoff prediction error in our study region. The *MAE* and *RMSE* of the runoff from our coupled method are much smaller than those from the Budyko and the hydro-stochastic interpolation methods. The maximum error of runoff at the sub-basin HWH is significantly reduced; the error is 236 mm from the coupled method compared to 328 mm from the Budyko method and 448 mm from the hydro-stochastic interpolation. In cross-validation (Table 2), our coupled method has R_{cv}^2=0.93, much larger than 0.81 and 0.54 from the Budyko method and the hydro-stochastic interpolation, respectively.

Our correlation analysis between predicted and observed R is shown in Fig. 6. The predicted runoff from our coupled method is highly correlated with the observed (R^2=0.95). In contrast, R^2=0.82 and 0.58 for the Budyko method and the hydro-stochastic interpolation, respectively. Our analysis indicates that the latter two methods overestimate low runoff and underestimate high runoff, as shown by large departures from 1:1 line in Fig. 6. Similar large deviation of the runoff predicted by the hydro-stochastic interpolation has also been reported in the previous work by Sauquet et al. (2000), Laaha and Bloschl (2006) and Yan et al. (2011).

Spatial distributions of runoff in the HRB calculated from the three methods are shown in Fig. 7. They again show significant differences. Compared to the result from our coupled method (Fig. 7c), the Budyko method (Fig. 7a) and hydro-stochastic interpolation (Fig. 7b) considerably overestimate sub-basin runoff in the north of the basin, where climate is relatively dry and runoff is small (ranging from 140-280 mm). Among the predicted runoff in the largest non-overlapping area upstream of BB in the basin, the one made by our coupled method is 125 mm, and the one made by the Budyko method and the hydro-stochastic interpolation is 264 and 179 mm, respectively. The results from our coupled method describe most closely the observed distribution of runoff in the HRB (Fig. 7d).

5. Discussions and conclusions

In this study, we use the Budyko's deterministic method to describe mean annual runoff as an integrated spatially continuous process determined by both the hydro-climatic elements and the hierarchical river network. A deviation from the Budyko estimated runoff is used by the hydro-stochastic interpolation that assumes spatially auto-correlated error. The deterministic aspects of runoff described by Budyko method reflect regional trends at locations (sub-basins), and their deviations caused by stochastic processes are considered by the weights as a function of autocorrelation and distance. Weights are larger for near points/basins and smaller for distant points/basins. Information from both the Budyko method and the hydro-stochastic interpolation are taken into account in our coupled model to predict the runoff.

We have tested this coupled method and compared its results to the Budyko method and the hydro-stochastic interpolation in the Huaihe River basin (HRB) in China. Our comparison results show that the deterministic process strongly affects spatial variations in runoff over the 40 sub-basins in HRB. The error of predicted runoff in terms of *MAE* and *RMSE* from the Budyko method is smaller than that from the hydro-stochastic interpolation method. In addition, the cross-validation result shows that the deterministic coefficient R_{cv}^2 from the Budyko method is larger than that from the traditional hydro-stochastic interpolation. These results suggest that estimation of runoff determined by the Budyko method in conjunction with the random deviations described by the hydro-stochastic interpolation method can improve the accuracy of predicted runoff. Our coupled method takes this approach, and its results show that it outperforms both the Budyko method and the stochastic interpolation by significantly increasing the runoff prediction accuracy. The interpolation errors described by *MAE* and *RMSE* from our coupled method are reduced to 47 and 69 mm, respectively, over the 40 sub-basins in the HRB. The largest error in predicted runoff in the HRB, at the sub-basin HWH, is also significantly reduced. That error is reduced to be 236 mm in our coupled method from being 328 mm in the Budyko method and 448 mm in the hydro-stochastic interpolation method. The cross-validation results show that the deterministic coefficient R_{cv}^2 in our coupled method is 0.93, much larger than 0.81 and 0.54 in the Budyko and the hydro-stochastic interpolation method, respectively. Furthermore, prediction from our coupled method describes the high and low runoff in sub-basins of the HRB more accurately than by the other two methods.

While substantial progress has been made by our coupled method, its results show that more effort is needed to further improve the accuracy of runoff prediction. There remain large runoff prediction errors from our coupled method at some sub-basins, e.g., the large sub-basins ZK and BB where the relative error of predicted runoff is larger than 40% of the observed runoff. Such large errors could result partially from insufficient number of observation stations in the large sub-basins (see Fig. 1). Other possible causes could be from additional external factors influencing the runoff, such as land-cover, soil properties, hydro-climatic variations of seasonality and groundwater levels. Including some or all these effects to improve the Budyko method will aid our understanding of the deterministic processes and help increase runoff prediction accuracy by our coupled method.

Acknowledgement

We thank the editor Erwin Zehe, the reviewers M. Mälicke and J.O. Skøien for their comments and suggestions that helped improve this manuscript substantially. The research was supported by the National Natural Science Foundation of China (No. 51190091 and 41571130071).

References

Arnell., N. W.: Factors controlling the effects of climate change on river flow regimes in a humid temperate environment, Journal of hydrology, 132(1-4), 321-342, 1992.

Arnell, N. W.: Grid mapping of river discharge. J. Hydrol., 167, 39-56, 1995.

Barancourt, C., Creutin, J. D., and Rivoirard, J.: A method for delineating and estimating rainfall fields, Wat. Resour. Res., 28, 1133-1144, 1992.

Berghuijs, W. R., Woods, R. A., and Hrachowitz, M.: A precipitation shift from snow towards rain leads to a decrease in streamflow, Nature Clim. Change, 4(7), 583–586, 2014.

Bishop, G. D., Church, M. R., Aber, J. D., Neilson, R. P., Ollinger, S. V., and Daly, C.: A comparison of mapped estimates of long term runoff in the northeast United States, Journal of Hydrology, 206: 176-190, 1998.

Bloschl, G.: Rainfall-runoff modelling of ungauged catchments, Article 133, in: Encyclopedia of Hydrological Sciences, edited by: Anderson, M. G., pp. 2061–2080, Wiley, Chicester, 2005.

Bloschl, G., Sivapalan, M., and Wagener T.: Runoff Prediction in Ungauged Basins: Synthesis Across Processes, Places and Scales, Cambridge Univ. Press, Cambridge, U. K, 2013.

Budyko, M. I.: Climate and Life, Academic, New York, 1974.

Choudhury, B.: Evaluation of an empirical equation for annual evaporation using field observations and results from a biophysical model, J. Hydrol., 216(1–2), 99–110, 1999.

Creutin, J. D. and Obled, C.: Objective analysis and mapping techniques for rainfall fields an objective comparison, Wat. Resour. Res., 18, 413-431, 1982.

Degaetano, A. T. and Belcher, B. N.: Spatial interpolation of daily maximum and

minimum air temperature based on meteorological model analyses and independent observations, Journal of Applied Meteorology & Climatology, 46(11), 1981-1992, 2006.
Dingman, S. L., Seely-Reynolds, D. M. and Reynolds, R. C.: Application of kriging to estimating mean annual precipitation in a region of orographic influence, Wat. Resour. Bull., 24, 329-339, 1988.
Donohue, R. J., Roderick, M. L., and McVicar, T. R.: On the importance of including vegetation dynamics in Budyko's hydrological model, Hydrol. Earth Syst. Sci., 11(2), 983–995, 2007.
Donohue, R. J., Roderick, M. L., and McVicar, T. R.: Roots, storms and soil pores: Incorporating key ecohydrological processes into Budyko's hydrological model, 436-437, 35–50, 2012.
Dooge, J. C. I.: Looking for hydrologic laws. Water Resources Research 22 (9), 46S–58S, (2003). Linear theory of hydrologic systems. EGU Reprint Series (Originally published in 1965), Katlenburg-Lindau, Germany, 1986.
Fu, B.: On the calculation of the evaporation from land surface (in Chinese), Sci. Atmos. Sin., 1(5), 23–31, 1981.
Gao, M., Chen, X., Liu, J., and Zhang, Z. Regionalization of annual runoff characteristics and its indication of co-dependence among hydro-climate–landscape factors in Jinghe River Basin, China. Stoch Env Res Risk A, 1-18.
Gentine, P., D'Odorico, P., Lintner, B. R., Sivandran, G., and Salvucci, G.: Interdependence of climate, soil, and vegetation as constrained by the Budyko curve, Geophys. Res. Lett., 39(19), L19404, 2012.
Gerrits, A. M. J., Savenije, H. H. G., Veling, E. J. M. and Pfister, L.: Analytical derivation of the Budyko curve based on rainfall characteristics and a simple evaporation model, Water Resour. Res., 45, W04403, 2009.
Gottschalk, L.: Correlation and covariance of runoff, Stochas. Hydrol. Hydraul., 7, 85-101, 1993a.
Gottschalk, L.: Interpolation of runoff applying objective methods, Stochas. Hydrol. Hydraul., 7, 269-281, 1993b.
Gottschalk, L., Krasovskaia, I., Leblois, E., and Sauquet, E.: Mapping mean and variance of runoff in a river basin, Hydrology and Earth System Sciences Discussions, 3(2), 299-333, 2006.
Goovaerts, P.: Geostatistics for natural resources evaluation, Oxford University Press on Demand, 1997.
Greenwood, A. J. B., Benyon, R. G., and Lane. P. N. J.: A method for assessing the hydrological impact of afforestation using regional mean annual data and empirical rainfall–runoff curves, Journal of Hydrology, 411(1–2), 49-65, 2011.
Greve, P., Orlowsky, B., Mueller, B., Sheffield, J., Reichstein, M., and Seneviratne, S. I.: Global assessment of trends in wetting and drying over land, Nat. Geosci., 7(10), 716–721, 2014.
Han, S., Hu, H., Yang, D., and Liu, Q.: Irrigation impact on annual water balance of the oases in Tarim Basin, Northwest China, Hydrol. Process, 25, 167–174, 2011.
Hisdal, H., Tveito, O. E.: Generation of runoff series at ungauged locations using

empirical orthogonal functions in combination with kriging, Stochas Hydrol. Hydraul., 6, 255-269, 1993.
Hollingsworth, A., Lönnberg, P.: The verification of objective analyses: diagnostics of analysis system performance, Meteorology & Atmospheric Physics, 40(1-3), 3-27, 1989.
Holman, I. P., Tascone, D., and Hess, T. M.: A comparison of stochastic and deterministic downscaling methods for modelling potential groundwater recharge under climate change in East Anglia, UK: implications for groundwater resource management, Hydrogeology Journal, 17(7), 1629-1641, 2009.
Hu, W. W., Wang, G. X., Deng, W., and Li, S. N.: The influence of dams on eco hydrological conditions in the Huaihe River basin, China, Ecological Engineering, 33(3), 233-241, 2008.
Imbach, P. L., Molina, L. G., Locatelli, B., Roupsard, O., Ciais, P., Corrales, L., and Mahé, G.: Climatology-based regional modelling of potential vegetation and average annual long-term runoff for Mesoamerica, Hydrology Earth System Sciences, 14(10), 1801-1817, 2010.
Istanbulluoglu, E., Wang, T., Wright, O. M., and Lenters, J. D.: Interpretation of hydrologic trends from a water balance perspective: The role of groundwater storage in the Budyko hypothesis, Water Resour. Res., 48, W00H16, 2012.
Jakeman, A. J. and Hornberger, G. M.: How much complexity is warranted in a rainfall-runoff model? Water Resources Research, 29(8), 2637-2649, 2010.
Jiang, C., Xiong, L., Wang, D., Liu, P., Guo, S., and Xu, C. Y.: Separating the impacts of climate change and human activities on runoff using the Budyko-type equations with time-varying parameters, Journal of Hydrology, 522, 326-338, 2015.
Jin, X., Xu, C. Y., Zhang, Q., and Chen, Y. D.: Regionalization study of a conceptual hydrological model in Dongjiang basin, South China, Quaternary International, 208(1–2), 129-137. 2009.
Jones, O. D.: A stochastic runoff model incorporating spatial variability. 18th world IMACS CONGRESS AND MODSIM09 International congress on modelling and simulation: interfacing modelling and simulation with mathematical and computational sciences, 157(1), 1865-1871, 2009.
Jutman, T.: Runoff, Climate, Lakes and Rivers: National Atlas of Sweden. Stockholm: SNA Publishing, 106-111, 1995.
Kearns, M. and Ron, D.: Algorithmic stability and sanity-check bounds for leave-one-out cross-validation, Neural computation, 11(6), 1427-1453, 1999.
Koster, R. D. and Suarez M. J.: A simple framework for examining the inter annual variability of land surface moisture fluxes, J. Clim., 12(7), 1911–1917, 1999.
Laaha, G. and Bloschl., G.: Seasonality indices for regionalizing low flows, Hydrological Processes, 20(18), 3851-3878, 2006.
Laaha, G., Skøien, J. O., Nobilis, F., and Blöschl, G.: Spatial prediction of stream temperatures using Top-kriging with an external drift, Environmental Modeling & Assessment, 18(6), 671-683, 2013.
Legates, D. R. and McCabe. G. J.: Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation, Water resources research, 35(1), 233-

241, 1999.
Lenton, R. L. and Rodriguez-Iturbe, I.: Rainfall network system analysis: the optimal esimation of total areal storm depth, Wat. Resour. Res., 13, 825-836, 1977.
Li, D., Pan, M., Cong, Z., Zhang, L., and Wood, E.: Vegetation control on water and energy balance within the Budyko framework, Water Resour. Res., 49(2), 969–976, 2013.
Li, J. and Heap. A. D.: A review of spatial interpolation methods for environmental scientists, 137-145, 2008.
Luo, W., Taylor, M. C. and Parker., S. R.: A comparison of spatial interpolation methods to estimate continuous wind speed surfaces using irregularly distributed data from England and Wales, International Journal of Climatology, 28(7), 947–959, 2008.
Milly, P. C. D.: Climate, soil water storage, and the average annual water balance, Water Resour. Res., 30(7), 2143–2156, 1994.
Niehoff, D. Fritsch, U., and Bronstert, A.: Land-use impacts on storm-runoff generation: scenarios of land-use change and simulation of hydrological response in a meso-scale catchment in SW-Germany, Journal of Hydrology, 267(1–2), 80-93, 2002.
Parajka, J. and Szolgay, J. Grid-based mapping of long-term mean annual potential and actual evapotranspiration in Slovakia, IAHS Publications-Series of Proceedings and Reports-Intern Assoc Hydrological Sciences, 248, 123-130, 1998.
Porporato, A., Daly, E., and Rodriguez-Iturbe, I.: Soil water balance and ecosystem response to climate change, Am. Nat., 164(5), 625–632, 2004.
Potter, N. J. and Zhang, L.: Inter annual variability of catchment water balance in Australia, Journal of Hydrology, 369(1), 120-129, 2009.
Qiao., C. F.: Mapping runoff isocline of Hai, Luan River basin. Hydrology, (s1), 63-66, 1982.
Ripley., B. D.: The second-order analysis of stationary point processes, Journal of applied probability, 13(2), 255-266, 1976.
Sauquet, E. Mapping mean annual river discharges: Geostatistical developments for incorporating river network dependencies, Journal of Hydrology 331, 300– 314, 2006.
Sauquet, E., Gottschalk, L., and Leblois. E.: Mapping average annual runoff: a hierarchical approach applying a stochastic interpolation scheme, Hydrological Sciences Journal, 45(6), 799-815, 2000.
Shao, Q., Traylen, A., and Zhang, L.: Nonparametric method for estimating the effects of climatic and catchment characteristics on mean annual evapotranspiration, Water Resour. Res., 48, W03517, 2012.
Sivapalan, M.: Pattern, processes and function: elements of a unified theory of hydrology at the catchment scale. In: Anderson, M. (ed.) Encyclopedia of hydrological sciences, London: John Wiley, pp. 193–219, 2005.
Sivapalan, M., Takeuchi, K., Franks, S. W., Gupta, V. K., Karambiri, H., Lakshmi, V., ... and Oki, T.: Iahs decade on predictions in ungauged basins (pub), 2003–2012: shaping an exciting future for the hydrological sciences, Hydrological Sciences Journal, 48(6), 857-880, 2003.
Skøien, J. O., Merz, R., and Bloschl., G.: Top-kriging geostatistics on stream networks, Hydrology and Earth System Sciences Discussions, 2(6), 2253-2286, 2005.

Tabios, G. Q. and Salas, J. D.: A comparative analysis of techniques for spatial interpolation of precipitation, Wat. Resour. Bull., 21, 365-380, 1985.
Villeneuve, J. P., Morin, G., Bobée, B., Leblanc, D., and Delhomme, J. P.: Kriging in the design of streamflow sampling networks, Wat. Resour. Res., 15, 1833-184, 1979.
Wackernagel. H.: Multivariate geostatistics, Berlin: Springer, 1995.
Wagener, T., Sivapalan, M., Troch, P., and Woods, R.: Catchment classification and hydrologic similarity, Geography compass, 1(4), 901-931, 2007.
Wang, D. and Tang Y.: A one-parameter Budyko model for water balance captures emergent behavior in darwinian hydrologic models, Geophys. Res. Lett., 41, 4569–4577, 2014.
Williams, C. A., Reichstein, M., Buchmann, N., Baldocchi, D., Beer, C., Schwalm, C. ... and Papale, D.: Climate and vegetation controls on the surface water balance: Synthesis of evapotranspiration measured across a global network of flux towers, Water Resources Research, 48(6), 2012.
Xu, X., Liu, W., Scanlon, B. R., Zhang, L., and Pan, M. Local and global factors controlling water-energy balances within the Budyko framework, Geophys. Res. Lett., 40, 6123–6129, 2013.
Yan, Z., Xia, J., and Gottschalk, L.: Mapping runoff based on hydro-stochastic approach for the Huaihe River Basin, China, Journal of Geographical Sciences, 21(3), 441-457, 2011.
Yang, D., Sun, F., Liu, Z., Cong, Z., Ni, G., and Lei, Z. Analyzing spatial and temporal variability of annual water-energy balance in non-humid regions of China using the Budyko hypothesis, Water Resour. Res, 43, W04426, 2007.
Yang, H., Yang, D. Z. Lei, and Sun, F.: New analytical derivation of the mean annual water-energy balance equation, Water Resour. Res., 44, W03410, 2008.
Zhang, L., Dawes, W. R. G., and Walker, R.: Response of mean annual evapotranspiration to vegetation changes at catchment scale, Water Resour. Res., 37(3), 701–708, 2001.
Zhang, L., Hickel, K., Dawes, W. R., Chiew, F. H., Western, A. W., and Briggs, P. R.: A rational function approach for estimating mean annual evapotranspiration, Water Resources Research, 40(2), 2004.
Zhang, R., Chen, X. Zhang, Z. and Shi, P.: Evolution of hydrological drought under the regulation of two reservoirs in the headwater basin of the Huaihe River, China, Stochastic environmental research and risk assessment, 29(2), 487-499, 2015.

Captions of figures:

1. Figure 1 Topography and river network of study area
2. Figure 2 Sub-basins and hydrological stations of study area.
3. Figure 3 $E/P \sim E_0/P$, (b) $R \sim E_0/P$ for the 40 sub-basins (the solid line is the best fit function), and (c) sub-basins in the north and south of the study basin. Note: in (b) and (c), blue color indicates wetter climate in the south and yellow color indicates drier climate in the north.
4. Figure 4 Empirical covariogram ($Cov_e(d)$) from sub-basin runoff data and theoretical covariogram by fitted covariance function $Cov_p(d)$ of study area.
5. Figure 5 Empirical covariogram ($Cov_e(d)$) from the residual $R_s(x)$ and theoretical covariogram by fitted covariance function $Cov_p(d)$ of study area.
6. Figure 6 Cross validation of predicted runoff vs. observation by (a) Budyko method, (b) hydro-stochastic interpolation, and (c) our coupled method. The dashed-line is 1:1.
7. Figure 7 Spatial distribution of mean annul runoff estimated from (a) Budyko method, (b) hydro-stochastic interpolation, (c) our coupled method, and (d) observation.

Table 1 Summary of hydro-meteorological data and predicted runoff of sub-basins in study area

No.	Stations	Basin area (km^2)	P (mm)	R (mm)	E_0 (mm)	E_0/P	E (mm)	Budyko method			Hydro-stochastic interpolation		Coupled method	
								ω	Predicted R (mm)	Error (mm)	Predicted R (mm)	Error (mm)	Predicted R (mm)	Error (mm)
1	CTG	3090	1012	366	932	0.92	646	2.41	399	32.85	371	4.90	348	17.84
2	XHD	1431	1517	740	974	0.64	776	2.41	777	36.94	737	2.70	692	47.82
3	SQ	3094	822	168	1024	1.25	653	2.83	248	79.29	285	116.77	178	10.10
4	MS	1970	1517	672	957	0.63	845	3.06	786	114.28	584	88.45	662	10.13
5	BGS	2730	877	225	1029	1.17	651	2.57	279	53.93	247	22.39	181	44.01
6	XC	4110	945	225	997	1.06	720	3.02	332	106.82	272	46.77	212	13.11
7	BT	11280	910	223	993	1.09	687	2.85	310	86.94	275	52.25	219	3.74
8	ZK	25800	678	123	1061	1.56	555	2.54	163	39.96	228	104.65	61	61.70
9	JJJ	5930	1347	513	969	0.72	834	3.16	640	127.27	520	7.49	512	1.49
10	HB	16005	1092	335	937	0.86	757	3.15	455	120.48	334	1.02	360	25.01
11	ZQ	3410	739	118	1083	1.47	621	2.83	190	71.71	219	101.07	141	23.40
12	HPT	4370	1629	764	984	0.60	865	2.92	868	103.53	755	9.22	712	51.64
13	XX	10190	987	367	1053	1.07	620	2.10	343	23.77	381	13.73	424	56.96
14	BB	121330	850	215	1024	1.20	635	2.54	264	49.48	394	179.16	125	90.46
15	WJB	30630	1003	294	957	0.95	709	2.85	384	90.29	304	9.65	287	6.90
16	LZ	390	963	345	1078	1.12	618	2.09	320	24.96	320	25.08	399	53.75
17	NLD	1500	1019	439	1101	1.08	581	1.86	351	88.30	309	129.64	401	37.56
18	ZMD	109	690	212	1093	1.58	478	1.94	163	48.65	281	68.78	235	22.53
19	BLY	737	1504	868	1126	0.75	635	1.69	695	173.27	639	229.05	794	74.23
20	HWH	292	1560	1068	1127	0.72	492	1.43	740	328.03	619	448.83	832	236.16
21	ZC	493	1512	838	1112	0.74	674	1.79	708	130.23	695	142.77	777	61.19
22	BQY	284	1268	693	1094	0.86	575	1.68	527	166.21	349	344.06	604	89.35
23	QL	178	1559	970	1090	0.70	589	1.60	756	214.17	646	324.06	840	130.17
24	HNZ	805	1480	640	1114	0.75	840	2.41	681	41.37	577	63.05	585	55.20
25	TJH	152	1305	699	1090	0.84	605	1.74	556	143.66	262	437.02	589	110.18
26	LX	77.8	1025	484	1079	1.05	540	1.75	361	123.77	241	242.88	436	48.01
27	ZLS	1880	755	253	1104	1.46	502	1.91	194	58.45	169	84.28	233	19.94
28	ZT	501	1021	437	1101	1.08	583	1.87	351	85.87	242	195.10	411	26.08

29	XGS	375	830	302	1088	1.31	528	1.91	238	63.74	243	58.60	297	5.46
30	JZ	46	1103	583	1107	1.00	520	1.63	404	178.81	200	382.51	455	127.50
31	GC	620	638	111	1055	1.65	528	2.51	145	34.18	139	28.42	103	8.08
32	ZM	2106	645	97	1039	1.61	548	2.72	150	53.48	141	43.80	105	7.58
33	YZ	814	979	235	1083	1.11	743	2.85	329	94.07	277	42.13	246	11.24
34	XZ	1120	746	111	1040	1.39	636	3.06	202	90.66	167	56.30	152	40.95
35	GZ	1030	855	342	1098	1.28	513	1.81	250	92.10	255	86.54	307	35.14
36	DPL	1770	1067	331	1066	1.00	736	2.57	393	61.62	339	8.02	342	11.39
37	XX2	256	1301	606	1092	0.84	695	2.00	552	53.68	705	99.36	552	53.82
38	PH	17.9	1248	708	1094	0.88	540	1.61	512	196.04	604	104.35	512	195.90
39	HC	2050	1255	454	1095	0.87	802	2.54	517	63.36	363	91.02	409	44.52
40	HK	2141	871	227	1077	1.24	644	2.44	264	37.28	309	82.40	186	41.22

Table 2 Interpolation cross-validation errors between the predicted and observed runoff at 40 sub-basins for the three methods

Evaluation indicators	Budyko method	Hydro-stochastic interpolation	Coupling method
MAE (mm)	94	134	47
MSE (mm^2)	12561	31024	4798
RMSE (mm)	112	176	69
Max absolute error (mm)	328	448	236
Min absolute error (mm)	24	3	2
Max relative error (%)	82	86	50
Min relative error (%)	5	0.3	0.3
R^2_{cv}	0.81	0.54	0.93

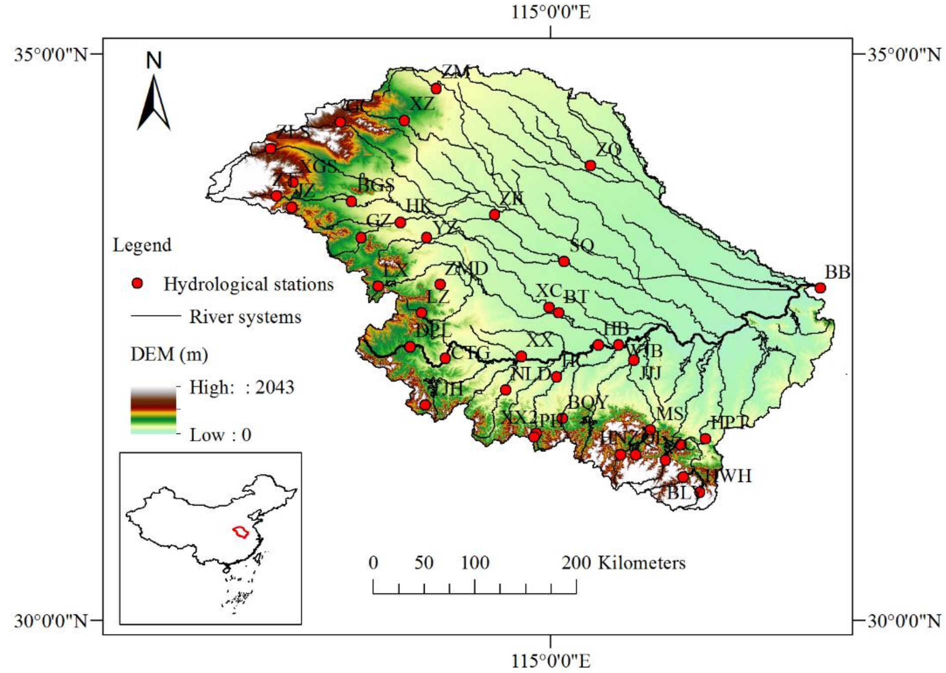

Figure 1: Topography and river network of study area.

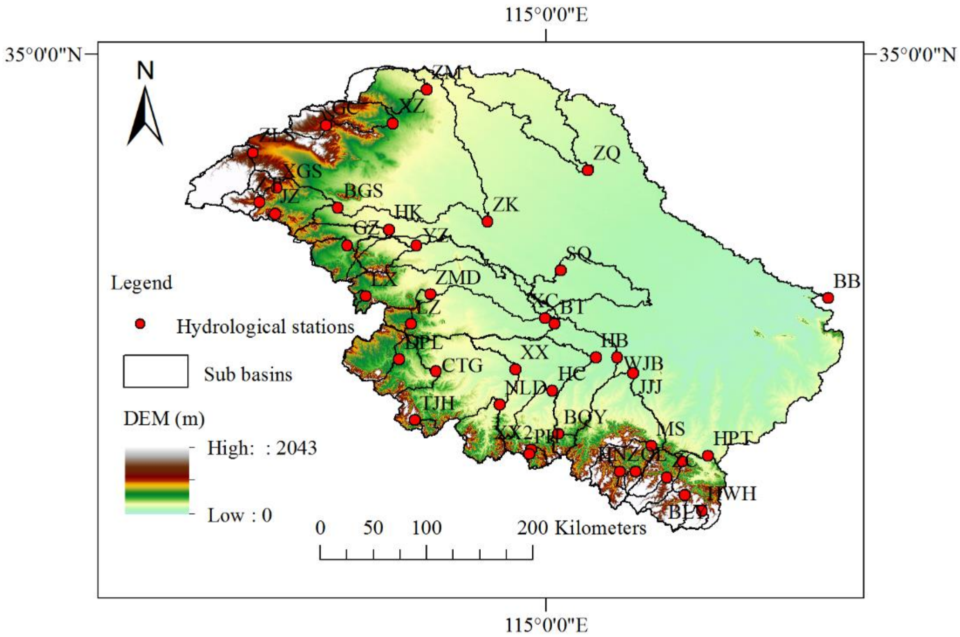

Figure 2: Sub-basins and hydrological stations of study area.

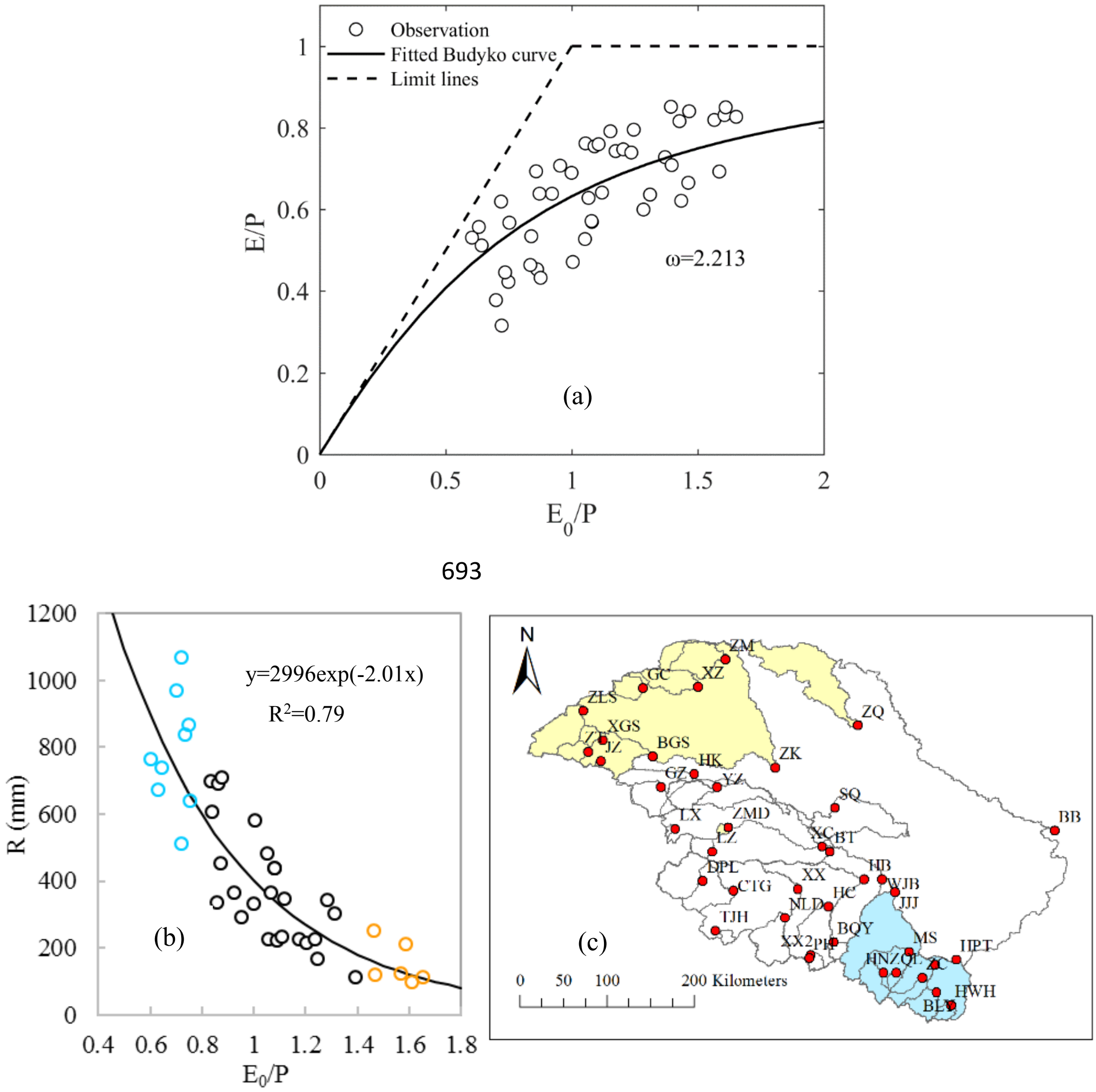

Figure 3: (a) $E/P \sim E_0/P$, (b) $R \sim E_0/P$ for the 40 sub-basins (the solid line is the best fit function), and (c) sub-basins in the north and south of the study basin. Note: in (b) and (c), blue color indicates wetter climate in the south and yellow color indicates drier climate in the north.

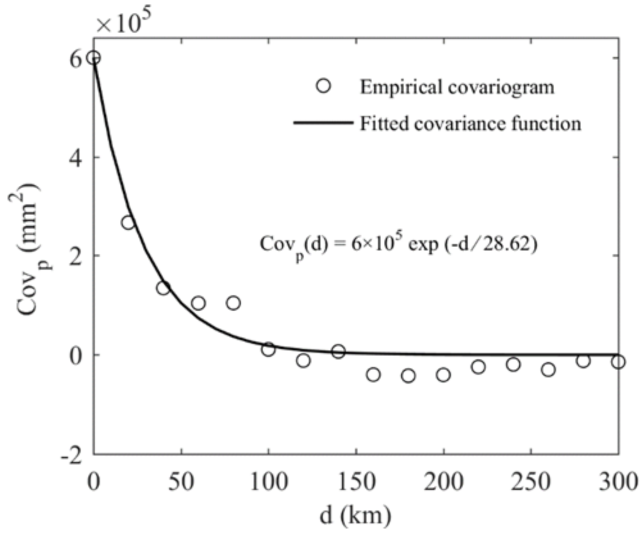

Figure 4: Empirical covariogram ($Cov_e\,(d)$) from sub-basin runoff data and theoretical covariogram by fitted covariance function $Cov_p\,(d)$ of study area

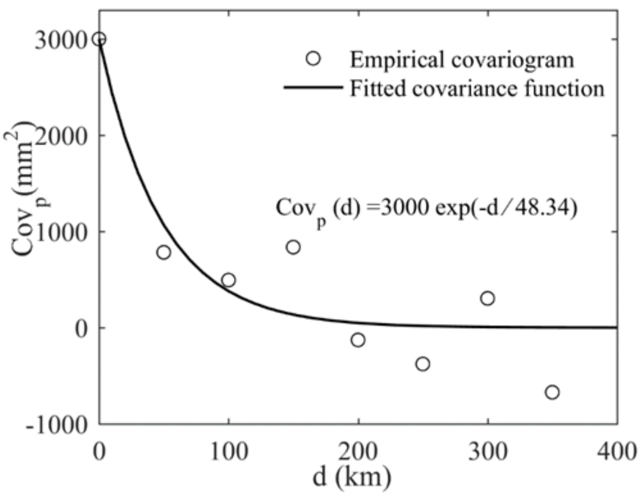

Figure 5: Empirical covariogram ($Cov_e(d)$) from the residual $R_s(x)$ and theoretical covariogram by fitted covariance function $Cov_p(d)$ of study area.

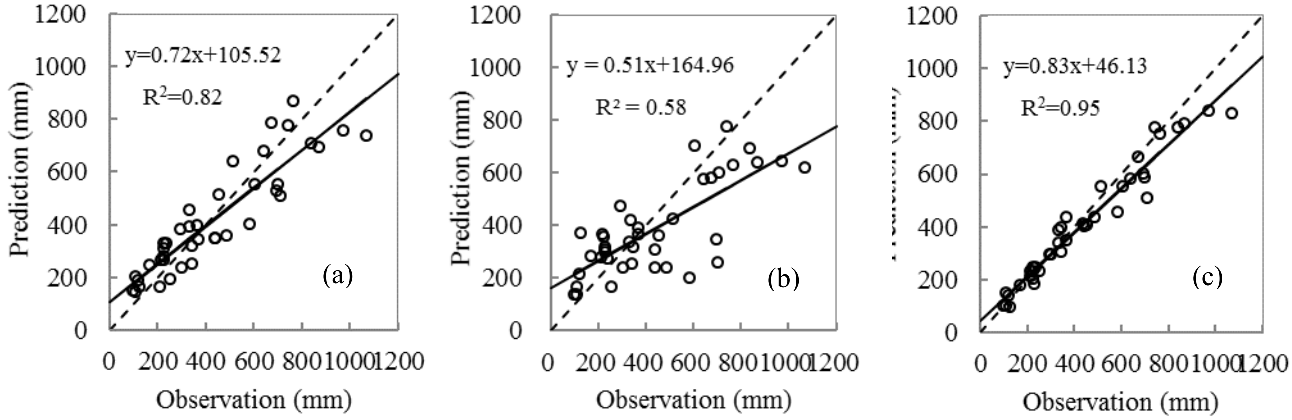

Figure 6: Cross validation of predicted runoff vs. observation by (a) Budyko method, (b) hydro-stochastic interpolation, and (c) our coupled method. The dashed-line is 1:1.

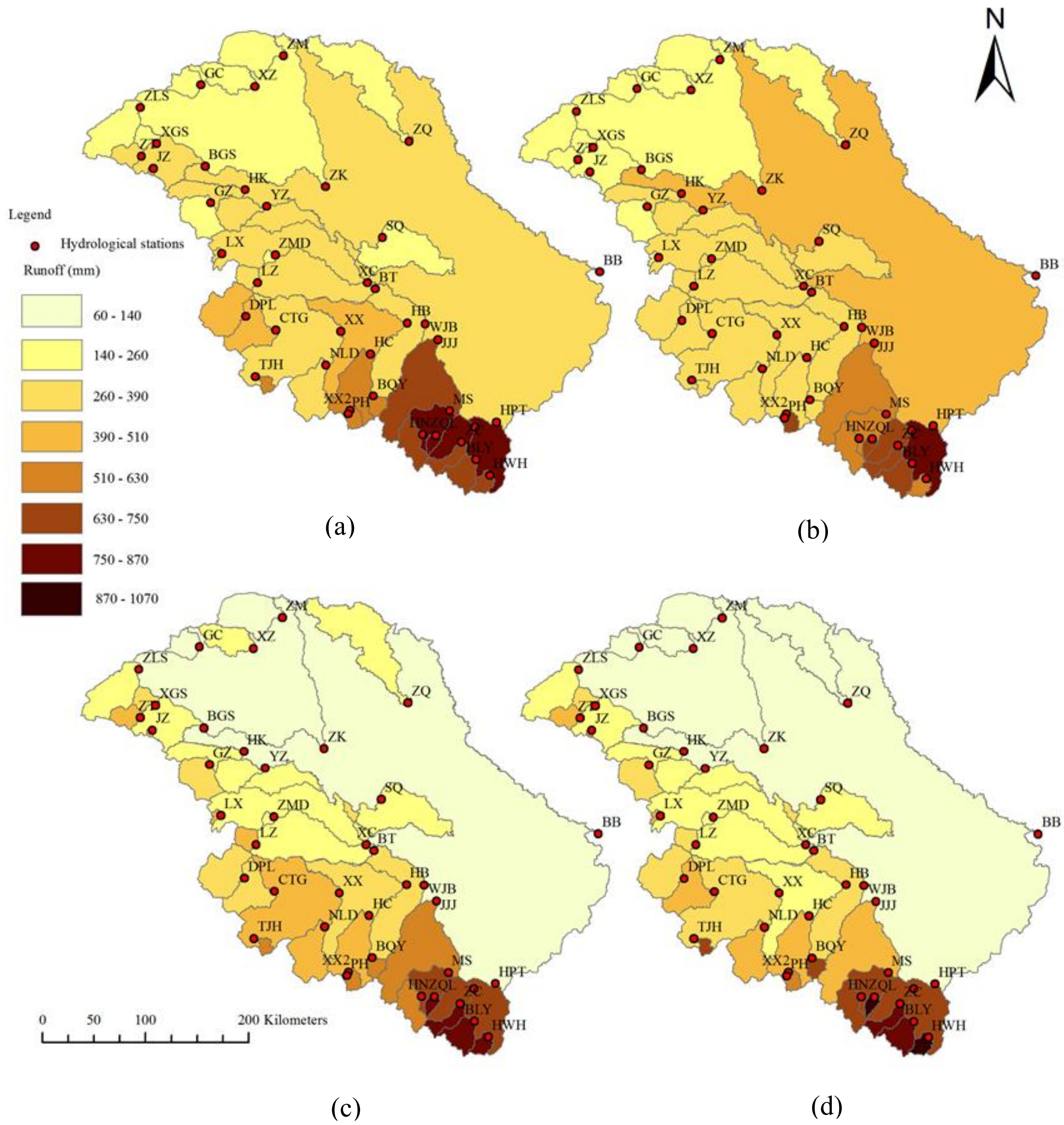

Figure 7: Spatial distribution of mean annul runoff estimated from (a) Budyko method, (b) hydro-stochastic interpolation, (c) our coupled method, and (d) observation.