Dear Editor and Reviewer,

We sincerely thank the editor and the reviewer for your reading our previous submission and for your valuable feedbacks that have helped us in improving this manuscript, entitled "*Hydro-stochastic interpolation coupling with Budyko approach for spatial prediction of mean annual runoff*" (ID: hess-2017-472). We have carefully studied the reviewer's constructive comments and made extensive modifications in our revised manuscript. Our point-to-point responses to the reviewer are listed below. The reviewer's comments are laid out in italicized blue font and have been numbered for each comment. Our responses are given in black. In our revised manuscript, all changes are highlighted using blue-colored text. Please contact me if you may have any questions about this revision.

Thank you again for your editorial work that has really helped us a lot.

Sincerely yours,

Xi Chen
On behalf of all co-authors

Major points

1. l.63 – 71: This paragraph consists of only two sentences, which are way too long and thus, were confusing for me. In the first sentence the authors make two different points. First, streamflow is a combined landscape information and second, that climate-landscape variability leads to non-stationary runoff observations. I kindly ask the authors to separate these points and reword the following statements in order to foster the structure. Especially the term "deterministic term" (l. 65) needs more and clearer introduction. This is in the following work also referred to as "deterministic trend" and is of fundamental importance for the proposed method. Introducing this term in more detail will significantly increase my text comprehension for the entire work.
The second sentence in this paragraph (l. 68 – 71) does in my opinion not connect to the first one and it was not obvious what this sentence shall emphasize. What trend does the "the spatially nonstationary trend of runoff" (l. 68) refer to? And how is a runoff trend interpretable as "hydrological regionalization in terms of hydro-climate and landscape data" (l. 68 – 69)? What I read out of this sentence is that non-stationary runoff is caused by heterogeneity in hydro-climate and the landscape and can be described by empirical relationships as done in the presented studies (l. 71). But this is not exactly what is written down in this paragraph.
In my opinion, the authors shall rewrite the whole paragraph in shorter, non-nested sentences.
Answer: According to the reviewer's suggestion, the paragraph has been rewritten in the revised manuscript (refer to *l.61~80).* The section of *line.72~82* in the previous manuscript was moved to *line.104~111* and *line.114~117*.

The whole paragraph has been rewritten as:

Unlike precipitation or evaporation which we often interpolate to find its values at specific points in space, runoff is an integrated spatial 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 exhibits some degrees of natural organization or connection of water basins (Dooge, 1986; Sivapalan, 2005), e.g., rivers that connect sub-basins. The river network structure constraints water balance between upstream and downstream in a basin. The hierarchically organized river structure requires that the sum of interpolated discharge from each of the sub-basins equals to the observed runoff in 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 1993b) developed a hydro-stochastic approach for runoff interpolation. It takes full account of the concept that runoff is an integrated course in the hierarchical structure of river basin systems. Distance between a pair of basins is measured along the river network by geostatistical distance instead of Euclidian distance. The covariogram among points in conventional spatial interpolation is replaced by covariogram between basins. However, in this concept, spatial runoff is considered spatially homogeneous over basins (Sauquet, 2006).

The observed patterns of runoff reveal systematic deviations from a purely homogeneous field due to the influence of some deterministic process acting in the basin, such as higher or lower runoff corresponding to larger or smaller rainfall over space. 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). Qiao (1982), Arnell (1992) and Gao et al. (2017) used such methods and derived empirical relationships between runoff and its controlling factors of climate, land use and topography in various basins. However, the deterministic method in describing complex patterns suffers inevitable loss of information (Wagener et al, 2007) because of existence of uncertainty in many hydrological processes and 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 the basin variables, e.g., mean annual runoff (Sauquet, 2006), stream temperature (Laaha et al., 2013) and groundwater table (Holman et al., 2009). Nevertheless, the deterministic term is mostly described by an empirical formula linking the spatial features, e.g. variability of mean annual runoff with elevation (Sauquet, 2006) and relationship between the mean annual stream temperature and altitude of the gauge (Laaha et al., 2013). As a simple 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) over space (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 been devoted to improving the Budyko method by deriving parameters with other external driving factors, such as land use/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 found that use of the deterministic equation of the Budyko method still came with large errors when it was used in prediction of runoff in any specific basin/area (Potter and Zhang, 2009; Jiang et al., 2015).

2. I would strongly recommend to completely rework the whole section 5 from line 408 to 451, due to many factors. Above all, this whole section is neither a discussion nor a conclusion in my opinion.

The first paragraph (l.409-424) basically lays the framework for coupling "deterministic and statistical models"(l.420), which is used as a justification for the proposed method. The paragraph itself seems to be helpful and relevant but should thus be moved to the introduction, somewhere located (and linked) to the paragraph l.105-122.

This paragraph is followed by two paragraphs that summarize major parts of the publication. l.425-434 summarizes the proposed method; while l.435 - 447 summarizes the reported result.

The only conclusions drawn can be found in the last paragraph (l.448-451). In my opinion, these conclusion are way too general. Furthermore the authors presented a new interpolation method, while long-term climate change impacts are modeled into the future, which would require an extrapolation. Thus, the proposed method is not appropriate to predict climate change impacts.

As the authors presented some interesting results in this publication, it should be easy to draw some more immediate and definite conclusions.

<u>Answer:</u> Discussions and conclusions have been revised in the manuscript. The first paragraph in the previous version of this manuscript (*l.409~424*) was moved to introduction in the revised manuscript (refer to *l.64~66, l.83~86 and l.88~93* in the Answer of *Major point 1*). We also revised the conclusions in more details. The last paragraph on long-term climate change impacts was replaced by discussions on reduction of errors from our method.

The section of Discussions and Conclusions from *l.460~487* in the revised manuscript is listed below.

We have tested our coupled method and compared the results with the Budyko method and the traditional hydro-stochastic interpolation in the Huaihe River basin (HRB) in China. Our results show that the deterministic process controls runoff variation in space instead of the stochastic process over the 40 sub-basins in HRB. The interpolation errors in terms of *MAE* and *RMSE* from the Budyko are smaller than those from the hydro-stochastic interpolation, and the cross-validation outcome of the deterministic coefficient R_{cv}^2 from the Budyko method is much larger than that from the traditional hydro-stochastic interpolation. Nevertheless, use of the deviation of runoff from the Budyko method estimation, instead of the observations, as the random error in Gottschalk's interpolation method can further improve interpolation accuracy of spatial runoff. The coupled method outperforms both the Budyko method and the stochastic interpolation by significantly increasing the spatial interpolation and prediction accuracy. The interpolation errors in terms of *MAE* and *RMSE* from our coupled method reduce to 47 and 69mm, respectively, over the 40 sub-basins. The maximum error at HWH is significantly reduced as well. It is 236 mm from our coupled method much smaller than 328 mm from the Budyko method and 448 mm from the hydro-stochastic interpolation. The cross-validation outcome of the deterministic coefficient R_{cv}^2 from our coupled method is 0.93, much larger than 0.81 and 0.54 from the Budyko method and the hydro-stochastic interpolation, respectively. The prediction from the coupled method captures most accurately the regional high and low runoff in the HRB among the three methods.

Our results show that there are still large interpolation errors from our coupled method at some of sub-basins, e.g., in larger sub-basins of ZK and BB where the relative errors are larger than 40%. Such large errors could be resulted from insufficient number of observation stations in such large sub-basins (see Fig. 1). Other possible reasons may be from additionally external factors that drive runoff in space, such as land use/land cover, soil properties, topography, hydro-climatic variations of seasonality and groundwater levels. Use of such data in improving Budyko method will enhance our understanding the deterministic process and thus increase the interpolation accuracy.

3. l.136 – 141: To me it is not clear why the authors have chosen Fu's equation. In the introduction to Budyko approaches (l. 129 – 136) the authors introduced a number of adjustments and improvements to the original approach suggested by other studies and highlighted their importance. Fu's equation does not incorporate any of these, but rather a "dimensionless model parameter"(l. 144), which does only control the "partitioning of precipitation into runoff"(l. 145). The authors are kindly asked to give more insights on this decision. Additionally, the calibration of this parameter is just mentioned in l. 146, but not further described.

Answer: In these adjustments, they only changed the parameter of Budyko equations, e.g., establishing relationship of ω in Fu's equation with land surface data. We agree that the improved Budyko approaches in consideration of other driving factors in addition to the aridity index could improve the prediction accuracy of runoff. However, they need plenty of basin characteristics that are often not available or inaccurately quantified in many areas. In our revision, we adopted the parametric Budyko curves of $E/P{\sim}E_0/P$. In the revision, we discussed how this possible way could increase accuracy of spatial runoff interpolation.

The calculation of parameter ω is given in *l.147~154* in the revised manuscript.

The parameter ω can be calculated by observed P, E_0 and R at each of gauged sub-basins. The mean value of ω can be obtained by averaging ω of sub-basins or least square method, i.e. to find the best fitting ω in Eqs. (1) or (2) by minimizing the mean absolute error (MAE) between predicted and estimated annual evapotranspiration E (=P-R) from the long-term water balance over sub-basins (Legates and McCabe, 1999). With known ω, Eq. (2) can be used for prediction of ungauged basin runoff or interpolation of spatial variation of runoff by using meteorological data in the target sub-basins (Parajka and Szolgay, 1998).

4. l.240 - 244: For my understanding, this is the key paragraph of the methodology as it describes the actual coupling of Budyko with hydro-stochastic interpolation. I would summarize this as: 1.: Rd(x) in equation (18) is substituted with equation (2) by setting Rd(x) = R. and calculated for all basins. 2.: The residuals between Rd(x) and observed R is calculated for all gauged basins. Further, these residuals are interpolated for all ungauged basins by "residual kriging"(l.243). and set as Rs(x) 3.: Equation (18) applies as the final result of this study.
Following the cited "residual kriging" from Sauquet (2006) it was not clear to me, how exactly the "residual kriging" is performed on the ungauged basins. The residuals from this study would be described by a first order polynomial ("Accounting for spatial heterogenity," last paragraph, in Sauquet (2006)), and be combined with ξq, the error in residuals. But, for me, it is not clear how this ξq or the g from Sauquet (2006) were calculated. From my point of view, the interpolation scheme described in Sauquet (2006) seems to be closely related to the general approach presented by the authors. Then, the delimitation between the two studies was not clear to me from the introduction. In any case a clarification of how Rs(x) is calculated, how section 2.2 sets in and is linked here would be highly appreciated.

Answer: We completely revised the related paragraphs. *l.240~244* in the previous version of the manuscript was revised as you suggested (*l.275~282*). In our work, the spatial heterogeneity was described by the Budyko equation, and a deviation from the Budyko method estimation is taken as the residual at all observation stations. Then the original hydro-stochastic interpolation approach was used for interpolation of the residual. The superposition of the estimates of the residual and the Budyko equation yields the prediction of runoff R.

The "residual kriging" is performed on the ungauged basins (e.g., non-overlapping sub-basins) by simultaneously optimizing the weights λ_j^i (i= 1, ..., *M*; j= 1, ..., *n*) according to Eqs. (11)-(15) (see *l.217~235* and *l.275~282* in the revised manuscript).

Our approach is similar to that of Sauquet (2006). A major difference is that we applied a semi-empirical approach of the Budyko, while Sauquet (2006) used an empirical formula (average annual runoff with mean elevation in Fig. 6 of Sauquet (2006)) in his description of spatial heterogeneity over basins (the calculation process of $R_s(x)$ please refer to *l.404-412* in our revised manuscript).

The content of *l.217~222* beneath Eq.(11) is as follows:

where, L_i is the weights column vector with respect to the $i-th$ non-overlapping sub-basins, each of which relates to the weights λ_j^i (*i*= 1,..., *M*; *j*=1,…, *n*) of sample observations, and μ^* and μ^i are all the Lagrange coefficients. To simultaneously optimize the weights λ_j^i (*i*= 1, ..., *M*; *j*= 1, ..., *n*), the covariance function values $Cov(u_i, u_j)$ of pair points in Eqs. (6) and (7) are represented as covariance function values of the pair sub-basins.

The content of *l.229~235* beneath Eqs.(12)-(15) is as follows.

where, A_i and A_j are the available basins (*i*=1,2…, *n*; *j*=1,2…, *n*, *n* is the number of gauged sub-basins), ΔA_i is the non-overlapping area for sub-basin i $(i = 1, \dots M)$, $Cov(A_i, A_j)$ is theoretical covariance function value of pairs of available basins, $Cov(A_n, \Delta A_i)$ is theoretical covariance function value of an available basin and the predicting basin, n_i is the number of grids of ΔA_i, n_T is the number of grids of the whole basin, $r(A_j)$ is runoff depth for basins j (j = 1, ..., n) with discharge observations, V_i^T is the transposed column vector of V_i, μ^i is the Lagrange coefficients.

The content of *l.275~282* is as follows:

In this study, Eq. (2) is used as an external drift function for estimation of the deterministic component $R_d(x)$ in all basins, i.e., $R_d(x)$ in Eq. (18) is substituted for Eq. (2) by setting $R_d(x) = R$. The residuals between $R_d(x)$ and observed R is calculated for all gauged basins. Furthermore, these residuals are interpolated for all ungauged basins by the "residual kriging" (Sauquet, 2006), i.e., $R_s(x)$ in Eq. (18) is replaced by $r^*(A_0)$ in Eq. (3) by setting $r^*(A_0)$=$R_s(x)$ for the hydro-stochastic interpolation method in section 2.2. The superposition of the estimates of both components according to Eq. (18) yields the prediction of R.

The content of *l.404~412* is as follows:

The empirical covariogram of $R_s{}^*(x)$ for each pair of sub-basins versus sub-basin distances is plotted in Fig. 5. The following exponential function is obtained from best fitting the empirical covariogram

$$Cov_p(d) = 3000\,\exp(-d/48.34). \qquad (25)$$

From (25), matrices C, C_0, K, V and G in Eqs. (9)-(15) are calculated using MATLAB, and the weight coefficient matrix of runoff deviation is subsequently calculated to predict runoff deviation. Since this interpolation scheme represents the spatial runoff deviation, the sum of the interpolated runoff deviation and the simulated runoff by Fu's equation is regarded as the total interpolated runoff in the sub-basins.

5. l.219 - 220: Which scatter diagram are you referring to, here? Furthermore I can hardly imagine how such a diagram would look like. For my understanding, an empirical covariogram relates the separating distances of lag classes the inner-class covariance observed in the data. Please describe how a diagram like this shall be scattered over the distances between all sub-basin combinations. Furthermore, equation (17) presented in line 222 is used to derive a theoretical covariogram. From my understanding (and in fact I am not sure what u1; u2; du1; du2 are referring to here, see minor point below) this will yield a single value Cov(A; B) for sub-basins A and B. Does the theoretical covariogram then relate Cov(A; B) to d(A; B) defined in (16) (l.217)? If so, a more descriptive and clear explanation in the respective paragraph would be highly appreciated.
Additionally, do Cov(A; B) (l.220) and Cov(ui; un) (l.178-180) describe the same thing?
Answer: We completely revised Section 2.2 for clear description of the empirical and theoretical covariograms (*l.*245~264 in the revised manuscript).

The content of *l.245~264* is presented below:

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

$$Cov(A,B) = \frac{1}{AB}\int\int_{AB} Cov_p(||u_1 - u_2||)du_1 du_2 \quad (17)$$

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

In the above, the geostatistical distance $d(A,B)$ between *A* and *B* is calculated based on grid division in each of the sub-basin (Sauquet and Gottschalk, 2000). We can obtain the mean distance *d* between all possible pairs of points (center points of grids) within the two sub-basins (*A* and *B*). For *n* sub-basins with observations, there are $n(n+1)/2$ pairs of the sub-basins, correspondingly with the mean distance d_i (i=1,…, $n(n+1)/2$).

Corresponding to the mean distance d_i between pairs of the 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 ranged with an interval (50 km in this study) to obtain the mean of $Cov_e(d_i)$ for the distance interval. Finally, the mean of $Cov_e(d_i)$ vs. the geostatistical distances d_i within each of the range intervals 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. (17) aiming to best fit Cov_e(d). Only independent sub-basins are used to calculate the covariance function to avoid spatial correlation of the nested sub-basins.

6. The authors should consider to report their result more consistently and comprehensive. Beside a cross-validation, the authors compare the three different interpolation approaches by comparing the errors each method yielded. This error reporting in line 377-379; 355-356 and 328-331 shall be harmonized and report the same numbers.
I would suggest reporting the overall minimum, maximum and mean error found in a single sub-basin, along with the minimum, maximum and mean relative error (as share of basin-specific runoff) found in any sub-basin. Both kind of errors can be reported as a absolute (in mm) and relative (in %) number. In my opinion this makes sense as, for example, the sub-basin yielding the biggest absolute error in equation 2 (which is HWH), does not show the biggest relative error (as eg. SQ shows a bigger relative error).
Beside reporting these important numbers, the authors should consider to report these numbers in table 2, as well.
Answer: We revised discussions of the results from the three methods in a way more consistent to that suggested by the reviewer, and the error reporting in *l.377~379; 355~356* and *328~331*(in the previous manuscript) have been harmonized in same numbers (refer to *l.413~415; 390~392* and *366~370*).
We modified Table 2 by adding those numbers as the reviewer suggested.

The contents of *l.413~417* is presented below:

Prediction outcome of runoff is listed in Table 1, with the *MAE* of 47 mm and *RMSE* of 69mm over the 40 sub-basins. The largest absolute error is at HWH (236 mm) and the smallest at JJJ (2 mm) (Table 2). The largest relative error is about 42.1% of the observed runoff at BB station, and the smallest is 0.3% of the observed runoff at JJJ station, corresponding to the absolute errors of 90 and 2 mm, respectively.

The contents of *390~394* is presented below:

The *MAE* and *RMSE* are 134 mm and 176mm, respectively. The largest absolute error is at HWH (448 mm) and the smallest at XHD (3 mm) (Table 2). The largest relative error is about 85.1% of the observed runoff at ZK station and the smallest is 0.4% of the observed runoff at XHD, corresponding to the absolute errors of 105 and 3 mm, respectively.

The contents of *366~370* is presented below:

The *MAE* of Budyko runoff prediction is 94 mm, and the *RMSE* is 112 mm. The largest absolute error is at HWH (328 mm) and the smallest at XX (24 mm) (Table 1 and 2). The largest relative error is about 81.6% of the observed runoff at XZ station and the smallest is 5.0% of the observed runoff at XHD, corresponding to the absolute errors of 91 and 37 mm at those two stations, respectively.

The modified Table 2 is as below:

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

Minor points

1. l.322 - 323: This observation is not supported by fig. 3. From my point of view it is not possible to derive the location of a sub-basin from this figure.

Answer: We added locations of the sub-basins in the north and the south section in the revised Figs. 3(b) and 3(c) (also shown below), showing runoff distribution lower in the north and higher in the south.

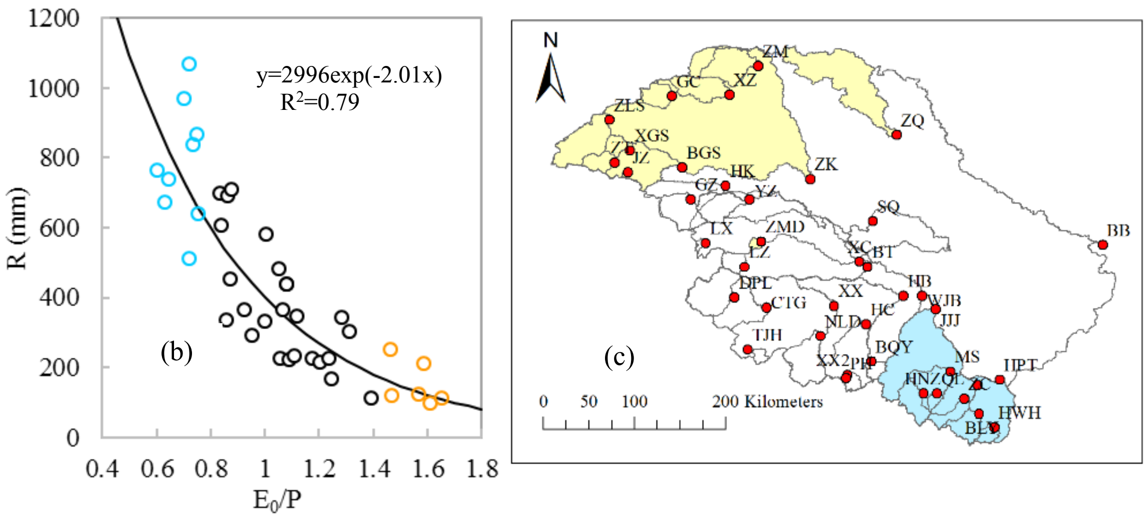

Figure 3 Plot of (b) *R ~E₀/P* for 40 sub-basins of HRB (c) sub-basins in the north and south of HRB. Note: in Fig 3(b) and (c), blue color indicates wetter climate in south and yellow color indicates the drier climate in north.

2. l.83 – 88: The authors make different points here within one long confusing sentence. They are kindly asked to break this sentence down to the core statements of: 1.: runoff is an integrated spatial continuous process, not a field like precipitation; 2.: runoff interpolation must take the stream network into account; 3.: the stream network constraints the water balance up- and downstream. Furthermore, please clarify the connection between a water balance constraint and assumed runoff properties that can be traced back to field properties.

Answer: Taking the Reviewer's suggestion, we broke this long confusing sentence into several short ones to make it easy to read and understand (refer to *l.61~67* in the Answer of *Major point 1*).

3. l.90 – 91: Please explain "lateral streamflow" (l. 90). What is that and how is it connected to the topic? None of the two presented studies, that shall explain the link between runoff overestimation and "neglecting lateral streamflow" contain the term "lateral streamflow." Please clarify what the two studies actually indicate.

Answer: We changed this expression as the river network in connecting sub-basins (the content is included in the Answer of *Major point 1*).

4. l.92 – 96. For my understanding, this part is not linked to the other parts of this paragraph or the introduction as far as I read it at that point. Why is this important? Additionally, "hydro-stochastic interpolation" (l. 92) was not clear to me at that point and the authors might consider some more explanation. Furthermore, the difference between "Euclidean distances" (l. 94) used in "conventional stochastic methods" and the "spatial distance" (l. 95) is too vague for me. Consider adding an explanation.

Answer: We deleted this sentence. In the revised manuscript, we added discussion on how to obtain spatial distance between a pair of basins (*l.238~244 and l.250~255* in the revised manuscript).
The contents of *l.250~255* is presented in the Answer of *Major point 5*.

The contents of *l.238~244* is presented below:

runoff hierarchical structure in the river system. 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*:

$$\mathrm{d}(A,B)=\frac{1}{AB}\int\int_{AB}||u_1-u_2||du_1du_2 \qquad (16)$$

where, A and B are the areas of sub-basins A and B, 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.

5. Please clarify what "samples" refers to in line 171.

<u>Answer:</u> We revised it as gauged stations in *line 177*. The content of *l.174~178* is presented below:

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

where, $Cov(u_i, u_j)$ is the theoretical covariance function value between each pair of gauged stations (i=1,…n, j=1,2…n), and $Cov(u_i, u_0)$ is the theoretical covariance value between the location of interest u_0 and each of the gauged stations u_i, μ is the Lagrange multiplier.

6. l. 362-364: The authors are asked to clarify what "trend removal" refers to here, as no kind of trend removal was reported in the methods. From that, what kind of assumptions do you "justify" from applying a trend removal? Do you assume the residuals to be spatially autocorrelated or do you assume an existing spatial autocorrelated random error underlying the residuals themselves, as the "hydro-stochastic interpolation" is performed on the residuals?
Consider extending the corresponding methods part.
<u>Answer:</u> We deleted this sentence. Here, we assumed the residuals to be spatially auto correlated, which is the basic condition for the stochastic interpolation method.

7. Is the du1; du2 used in (17) (l.222) and (18) (l.236) the same thing, or does the d from (17) refer to the d(A; B) calculated in (16) (l. 217)? If not, what is (16) then used for? If yes, please clarify the difference of the two used du1; du2.
<u>Answer:</u> We clarified the explanations of these items in the revised manuscript (refer to *l.242~244, l.250~255, and l.272~274*).
The content of *l.242~244* is presented in the Answer of *Minor point* 4.
The content of *l.250~255* is presented in the Answer of *Major point* 5.
The content of *l.272~274* beneath Eq.(18) is presented below:

In (18), $R(x)$ is runoff at location x, $R_d(x)$ is the deterministic component of the spatial trend and the external drift (Wackernagel, 1995) that results in nonstationary variability.

8. Please describe what "spatial variance" (l. 259) exactly means here and how it is defined.

Answer: Spatial variance is the variance of observed runoff data, and is calculated from this formula: $V_{NK} = \frac{\sum(R(x_i)-\bar{R})^2}{n-1}$, where $\bar{R}$ is mean $R(x)$. We have added this formula in the revised manuscript (*l.296~297)*.

9. The used precipitation data is described to be a "climatological dataset" (l.287). What kind of data product is this? An interpolated and aggregated map from a observation network? A radar product?

Answer: The used precipitation data are from the monthly precipitation dataset of China at 0.5° spatial resolution constructed by China Meteorological Administration**.** The dataset is on the basis of 2472 observational stations of China reorganized by National Meteorological Information Center, using Thin Plate Spline (TPS) interpolation method in ANUSPLIN software to obtain. It offers the monthly precipitation grid data of China started from 1961.The website from where we download the data is: http://data.cma.cn/data/detail/dataCode/SURF_CLI_CHN_PRE_MON_GRID_0.5.html. *(refer to l.326~328).*

10. l.290 How was this interpolation conducted? "ArcGIS" is capable of more than one interpolation method. Please name the method, not the tool.

Answer: According to the reviewer's advice, this sentence in lines 289 to 290 was modified as:

"Pan evaporation data at 21 meteorological stations in HRB are used to interpolate spatial potential evapotranspiration using ordinary kriging interpolation method via ArcGIS." *(refer to l.328~330).*

11. What is the "relative error [of] 91 mm"? Is this the absolute error at XZ station, where the relative error is the largest observed of 81.6%?

Answer: It is the absolute error. We have corrected it and pointed out the stations in the revised manuscript *(refer to l.368~370)*.

The content of *l.368~370* is in the Answer of *Major points 6.*

12. l. 340 - 348: Why was HRB divided into a grid? The corresponding methodological description of these results (l. 212 - 217) did not mention this step. Furthermore, for me the link between equation 24 (l.337), equation 25 (l. 349) and figure 4 is not clear. Both equations describe a empirical covariance C(d), while figure 4 shows a "covarinance function" along with an "empirical covariogram." Which one does refer to what here? The authors are kindly asked to make this clearer and the notation more distinct.

Answer: We revised the descriptions in text and Fig. 4. We first calculate empirical covariogram ($Cov_e(d)$) using sub-basin runoff data, and then use it to fit covariance function Cov_p (d) and obtain theoretical covariogram by the integration of Eq. (17). This part of modifications in the revised manuscript are shown in *l.245 ~264* which is listed in the Answer of *Major point 5*.

13. l.351 How shall equation 25 be used to "calculate the theoretical covarinace matrix Cov(A; B)*"? In line 220* Cov(A; B*) was described as a "theoretical covariogram," not a matrix. Is* Covp *in equation 17 than the same as* C(d) *in equation 25? Is the* d *in equation*

25 then derived from equation 16 for each sub-basin pair A,B? Are u*1*; u*2 in equation 16 then the grid points mentioned in l.340 - 348 or the "samples" mentioned in line 171? Clarifying this specific step in the methods wherever appropriate would be highly appreciated.*

Answer: We have clarified it on *l.250 ~261* in the revised manuscript.
The content of *l.250 ~261* is presented in the Answer of *Major point 5*.

14. *l.403-404 Did you mean that figure 7 (a) and (b) overestimate runoff, instead of underestimate, as stated? Because (a) ranges from 145mm - 280mm in the north and (b) ranges from 140mm - 280mm, in contrast (c) ranges from 60mm to 250mm in the north. Adding another sub-figure to figure 7 showing the measured runoff values can make figure 7 even more meaningful. Additionally, I would strongly recommend using the same value ranges for the color codes in figure 7, this will make the sub-figures more comparable and consistent.*

Answer: We revised it according to the reviewer's suggestion *(refer to l.440~443 and Fig.7(d))*.
The revised *Fig.7* is shown at the end of this Response letter.

The content of *l.440~443* is presented as follows:

interpolation methods. Compared with our coupled method (Fig. 7c), the Budyko method (Fig. 7a) and hydro-stochastic interpolation (Fig. 7b) markedly overestimate sub-basin runoff in the north where climate is relatively dry and runoff is small (ranging from 140mm - 280mm).

Technical points

1. In my opinion all the figures should be revised. The figure captions shall be extended and describe all figure elements. This is especially true for figures 3,4,5 and 6. Consider adding legends to figures 3 and 6.

Answer: According to the reviewer's suggestion, we have re-drawn all the figures, revised figure captions, and added legend to Fig. 3(a).
All the revised figures with their captions are shown at the end of this Response letter.

2. The authors are kindly asked to revise all their equations. Please make sure, that all used symbols are explained beneath the equation. This is especially true for μ_* and μ_i (l. 189); The sub- or superscripted T used in e.g. in l 192; the undefined symbols u_1; u_2; du_1; du_2 (l.217); $Cov(u_i; u_n)$ in l.178-179, 194-197) .
Wherever possible the symbol description shall also include the used unit. The unit was only given in a single case.

Answer: Thank you for your kind reminder.
We checked all the equations in the manuscript and made sure all the symbols are given their meanings beneath the equations.

3. l.129 – 136: This part is in fact a literature review on Budyko approaches and should thus be *moved from the methodology part into the introduction*.

Answer: We moved them to the Introduction (*l.108 ~114*) which is shown in the Answer of *Major point 1*.

4. l.334 - 339: In my opinion, these are methods and should be moved to the correct section.

Answer: We described it in the Introduction which is shown in the Answer of *Major point 1*.

5. l.314 - 316: Consider moving this to the methods (l.147-148), where the "calibration" is not further described.

Answer: We have already moved the contents of lines *314~316* (in the previous manuscript) to lines *148~151* which is presented in the Answer of *Major point 3*, and explained in more details the "calibration" of parameter ω.

6. What exactly is meant by "drainage basin" in line 224? In the preceding text the authors referred to basins and sub-basins.

Answer: "drainage basin" in this manuscript is changed to "basins" or "sub-basins" (*l.263* and *l.264*) through this revision.

7. Consider replacing "method with semi-empirical approach" (l. 112) with "method with semi-empirical Budyko approach," in order to be even more clear here.

Answer: We have replaced this term (refer to *l. 118 ~119*).

8. l.405 The authors should consider replacing "area above BB" with "area upstream of BB" or "area south of BB," to be more precise here.

Answer: We have revised this term in *l.444* in the revised manuscript. We also changed the wording in the caption of Figs. 1 and 2.

9. In line 388, I would not state that "0.93 [is] much larger than 0.81 and 0.54," as 0.93 - 0.81 is in fact smaller than 0.81 - 0.54. I would rather sayr "cross-validation outcome R^2_{cv} performed best for the coupled method (0.93)..." or something similar.

Answer: We changed the expressions to "In terms of the cross-validation outcome in Table 2, our coupled method performed best with R_{cv}^2 as large as 0.93, much larger than 0.81 and 0.54 from the Budyko method and the hydro-stochastic interpolation, respectively" (refer to *l.424~427*). It does not refer to the difference value between them (0.93 - 0.81 and 0.81 - 0.54).

10. The authors are asked to consider adding an overview map locating HRB in China. This could be added to figure 1 or as a fourth sub-figure to figure 7

Answer: We have added it in Fig.1. Fig.1 is shown at the end of this Response letter.

Other modifications:

Following errors have been corrected in the revision.

1. Line 141: Equation (2)

The original equation (2) was: $R = \left(1 + \left(\frac{E_0}{P}\right)^{\omega}\right)^{\frac{1}{\omega}} - E_0$, in which the symbol P was missed and we added it, that is: $R = P \cdot \left(1 + \left(\frac{E_0}{P}\right)^{\omega}\right)^{\frac{1}{\omega}} - E_0$.

2. Line 320: Equation (23)

The original equation (23) was $R = \left(1 + \left(\frac{E_0}{P}\right)^{2.213}\right)^{\frac{1}{2.213}} - E_0$. Similarly, it was modified as $R = P \cdot \left(1 + \left(\frac{E_0}{P}\right)^{2.213}\right)^{\frac{1}{2.213}} - E_0$.

3. Line 149: the words "sub basins"

"Sub basins" was changed to "sub-basins through the revision.

Other modifications not listed here were highlighted using blue-colored text in the revised manuscript.

All the revised figures are listed below:

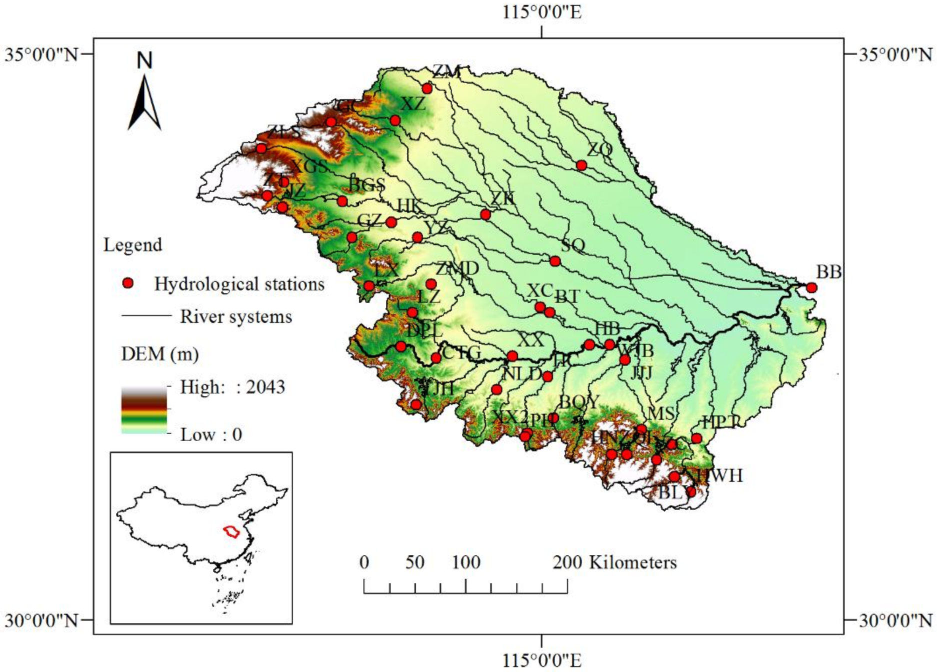

Figure 1 Topography and river systems of HRB upstream of BB.

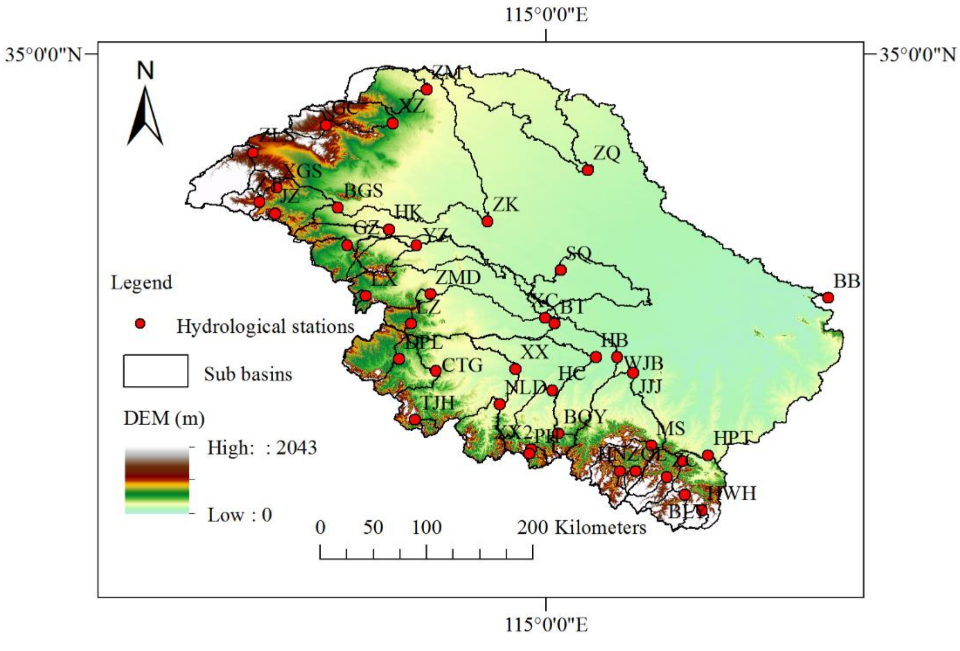

Figure 2 Sub-basins and hydrological stations of HRB upstream of BB.

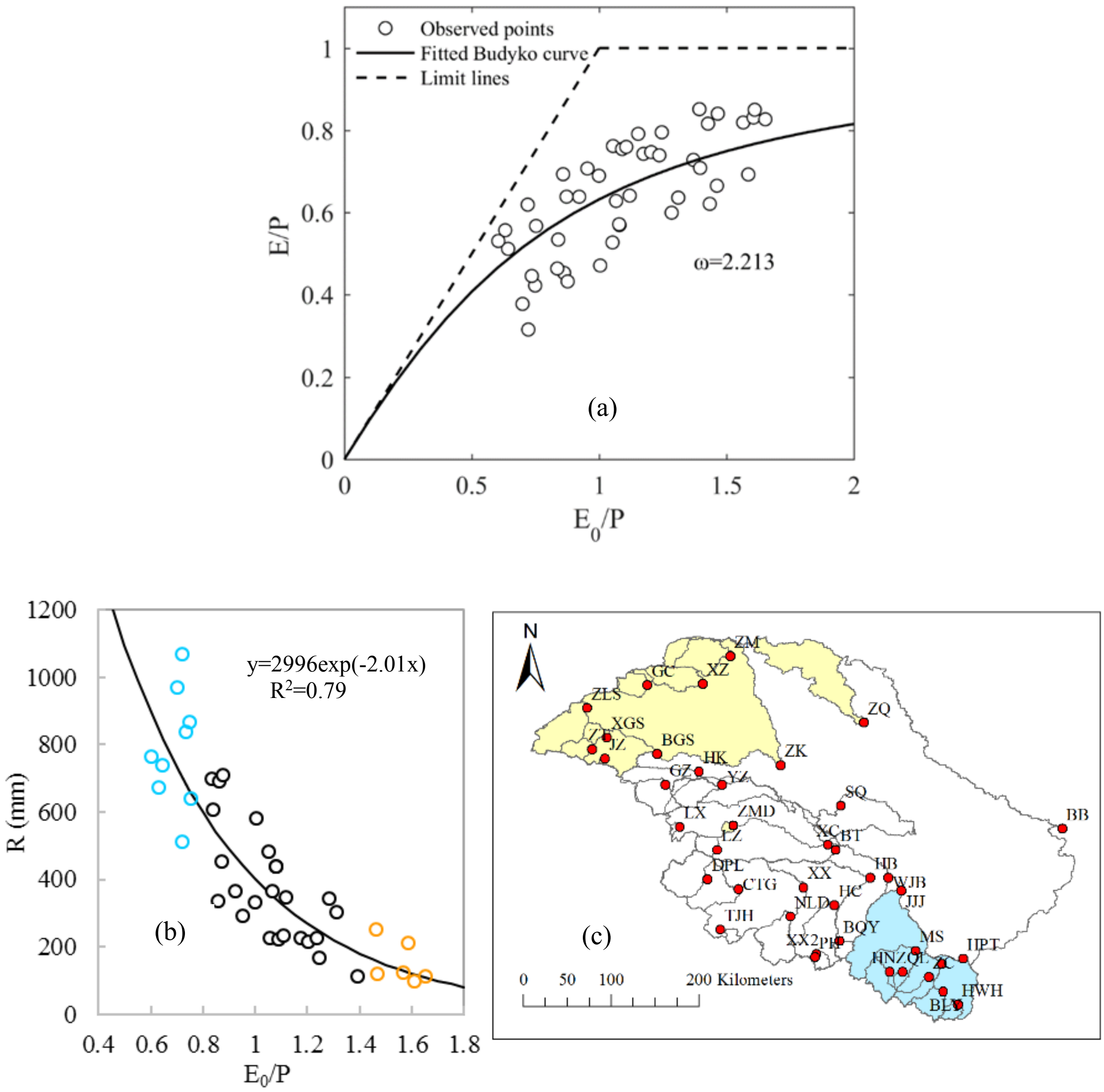

Figure 3 Plot of (a) $E/P \sim E_0/P$ and (b) $R \sim E_0/P$ for 40 sub-basins of HRB (c) sub-basins in the north and south of HRB. Note: in Fig 3(b) and (c), blue color indicates wetter climate in south and yellow color indicates the drier climate in north.

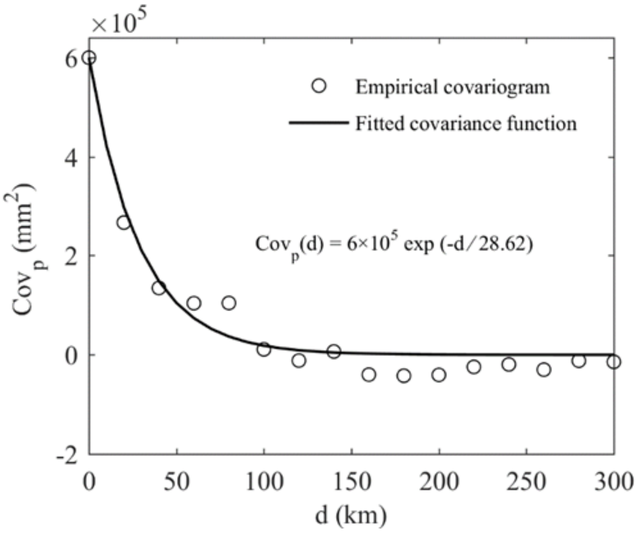

Figure 4 Empirical covariogram ($Cov_e(d)$) from sub-basin runoff data and theoretical covariogram by fitted covariance function $Cov_p(d)$ of HRB.

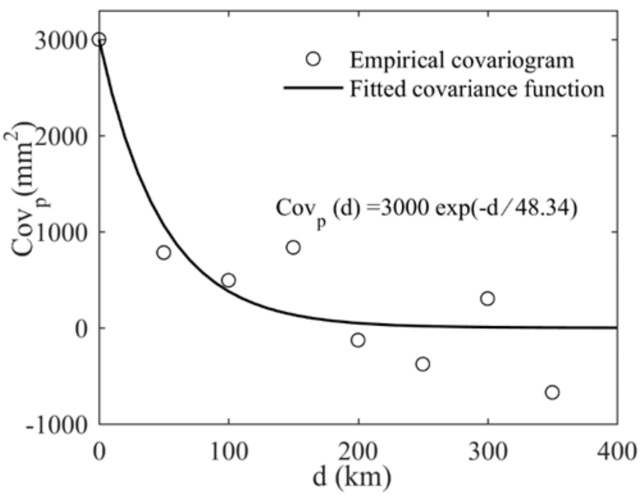

Figure 5 Empirical covariogram ($Cov_e(d)$) from the residual $R_s^*(x)$ and theoretical covariogram by fitted covariance function $Cov_p(d)$ of HRB.

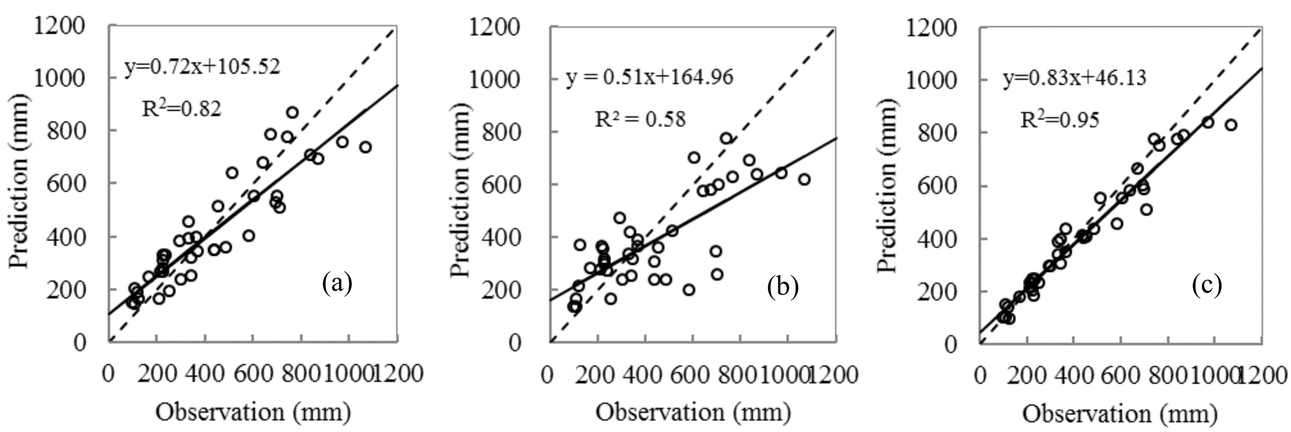

Figure 6 Cross validation of runoff prediction vs. observation by (a) Budyko method, (b) hydro-stochastic interpolation, and (c) coupled method. The dash line is 1:1.

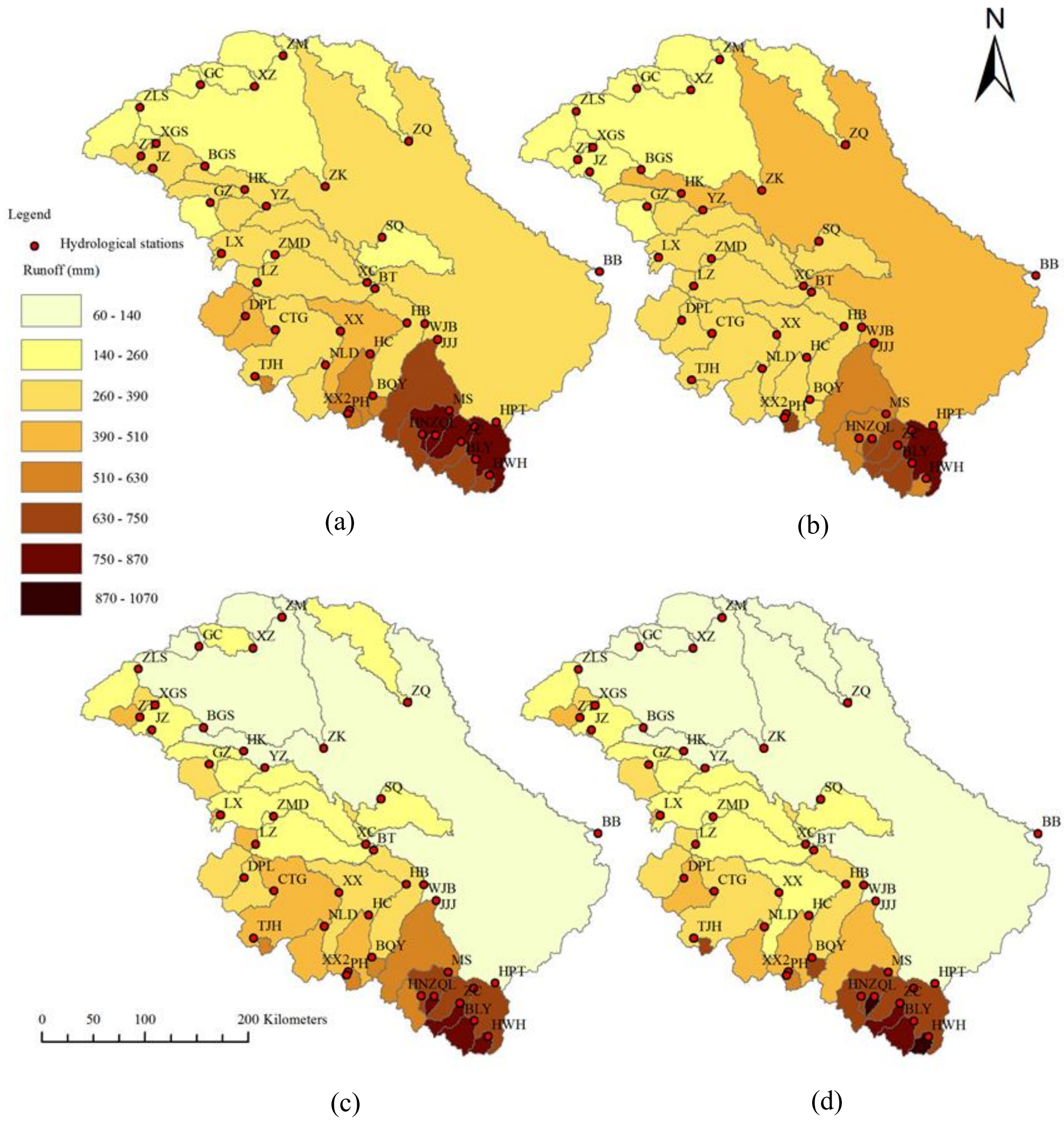

Figure 7 Spatial distribution of mean annul runoff estimated from (a) Budyko method, (b) hydro-stochastic interpolation, (c) coupled method, and (d) observed data.