




A Recursive Approach to Long-Term Prediction of Monthly Precipitation Using Genetic Programming

Suning Liu^{1,2,3} · Haiyun Shi^{1,2,4} 

Received: 8 August 2018 / Accepted: 11 December 2018 /

Published online: 15 December 2018

© The Author(s) 2018, corrected publication May/2019

Abstract

Precipitation is regarded as the basic component of the global hydrological cycle. This study develops a recursive approach to long-term prediction of monthly precipitation using genetic programming (GP), taking the Three-River Headwaters Region (TRHR) in China as the study area. The daily precipitation data recorded at 29 meteorological stations during 1961–2014 are collected, among which the data during 1961–2000 are for calibration and the remaining data are for validation. To develop this approach, first, the preliminary estimations of annual precipitation are computed based on a statistical method. Second, the percentage of the monthly precipitation for each month of a year is calculated as the mean monthly precipitation divided by the mean annual precipitation during the study period, and then the preliminary estimation of monthly precipitation for each month of a year is obtained. Third, since GP can be used to improve the prediction results through establishing the relationship of the observations with the preliminary estimations at the past and current times, it is adopted to improve the preliminary estimations. The calibration and validation results reveal that the recursive approach involving GP can provide the more accurate predictions of monthly precipitation. Finally, this approach is used to predict the monthly precipitation over the TRHR till 2050. Overall, the proposed method and the obtained results will enhance our understanding and facilitate future studies regarding the long-term prediction of precipitation in such regions.

Keywords Monthly precipitation · Recursive approach · Long-term prediction · Genetic programming · Three-River Headwaters Region

1 Introduction

As a basic meteorological variable, precipitation is of great importance to a variety of fields such as water resources management, agriculture, and ecological environment assessment (Wei et al. 2005; Kisi and Cimen 2011; Shi et al. 2016; Liu and Chui 2018). In the past several decades, global warming has greatly influenced the water circulation and hydrological processes (Garbrecht et al. 2004; IPCC 2013), and thus, has attracted the attentions of numerous researchers (Groisman et al. 2005; Westra et al. 2013; Cao and Pan

The original version of this article was revised due to a retrospective Open Access order.

✉ Haiyun Shi
shihy@sustc.edu.cn

2014). There are difficulties in the accurate prediction of precipitation because of the complexity of physical processes (Chow et al. 1988; Kulligowski and Barros 1998), especially for long-term prediction. As a result, many efforts have been made to develop appropriate methods to predict precipitation, which can be classified into the following types, e.g., dynamical methods (Claußnitzer and Névir 2009; Landman et al. 2014), statistical methods (Barnston and Smith 1996; Lee and Ouarda 2010; Kim et al. 2017; Chardon et al. 2018), soft computing methods (Silverman and Dracup 2000; Patal and Cigizoglu 2009; Ortiz-García et al. 2014), and numerical weather prediction methods (Richardson 2005; Park et al. 2008). Overall, the above-mentioned methods have been widely used and can be efficient in precipitation prediction; however, most of them rely on utilizing other climatic indices (e.g., El Niño-Southern Oscillation, ENSO) and climatic variables (e.g., sea surface temperature, SST). Moreover, precipitation can be influenced by a variety of related meteorological variables, e.g., air pressure, air temperature and wind speed (Singh 1988; Benestad 2013). However, due to the variable weather conditions and complex terrain orography, there is still limited amount of available meteorological data over the regions such as the Three-River Headwaters Region (TRHR), which may lead to great uncertainties in precipitation prediction (Xue et al. 2017). To this end, it would be valuable to develop a new approach to precipitation prediction for the designated regions (e.g., the TRHR), independent of other climatic indices and variables.

The TRHR is a plateau mountainous region in the western China. It is particularly sensitive to climate change, bringing serious disturbances to local ecosystem (Immerzeel et al. 2010; Tong et al. 2014; Shi et al. 2016, 2017; Xi et al. 2018). With reference to precipitation, several studies have reported an increasing trend in the annual precipitation (Liang et al. 2013; Tong et al. 2014; Shi et al. 2016) and precipitation extremes (Cao and Pan 2014) over this region. Moreover, it is worth noting that, for regions with large elevation variations (e.g., the TRHR), precipitation can be characterized by significant spatial variation, and the relationship between precipitation and elevation can be expressed by various functional forms (Naoum and Tsanis 2004; Chu 2012). Since precipitation is one of the basic component of hydrological cycle, to make accurate long-term prediction of precipitation for the TRHR has both scientific and practical significances (Xue et al. 2017). Therefore, for the designated regions such as the TRHR, it is important to develop an appropriate approach to precipitation prediction, especially for long-term prediction at the monthly scale.

To fill these gaps, the objective of this study is to develop a recursive approach to long-term prediction of monthly precipitation using genetic programming (GP), without the utilization of other climatic indices and variables. Compared to previous studies on the topic of precipitation prediction, the significances of this study can be summarized as follows: first, no other climatic indices and variables are introduced to predict precipitation. Second, a statistical method is proposed to compute the preliminary estimations of annual and monthly precipitation, and then GP is adopted as an optimization method to improve the preliminary estimations. Third, the prediction results of the monthly precipitation over the TRHR till 2050 are obtained, which cannot be found in existing literature. Overall, the proposed approach and the obtained results will enhance our understanding and facilitate future studies regarding the long-term prediction of precipitation in such regions. Following this Introduction, this paper has been divided into three further sections as follows: Section 2 introduces the data and methods, Section 3 presents the results and related discussion, and Section 4 draws several conclusions.

2 Data and Methods

2.1 Study Area and Research Data

The TRHR, well-known as the sources of the Yangtze River, the Yellow River and the Lantsang River, is located in the Qinghai Province in the western China (89°45' - 102°23' E, 31°39' - 36°12' N) (see Fig. 1). With an area of 0.3 million km², it accounts for 43% of the total area of the Qinghai Province (Cao and Pan 2014). The elevation of this region varies between 2000 m and 6600 m, and the mean value is over 4000 m. This region lies in the temperate zone, mainly dominated by a plateau monsoon climate (Liu and Yin 2002; Duan et al. 2013). Normally, the annual precipitation ranges from 262 mm to 773 mm in this region (Yi et al. 2013), and more than 80% of the annual precipitation occurs in the wet season from May to October (Liang et al. 2013; Shi et al. 2016).

There are 37 meteorological stations available inside or around the TRHR. The daily observed data can be downloaded for free on the official website of China Meteorological Administration (2016). As most of these stations were built in the late 1950s, only the data from 1961 are selected in this study to ensure that the lengths of the datasets from different stations are consistent. For the designated station, missing data in a certain year are interpolated

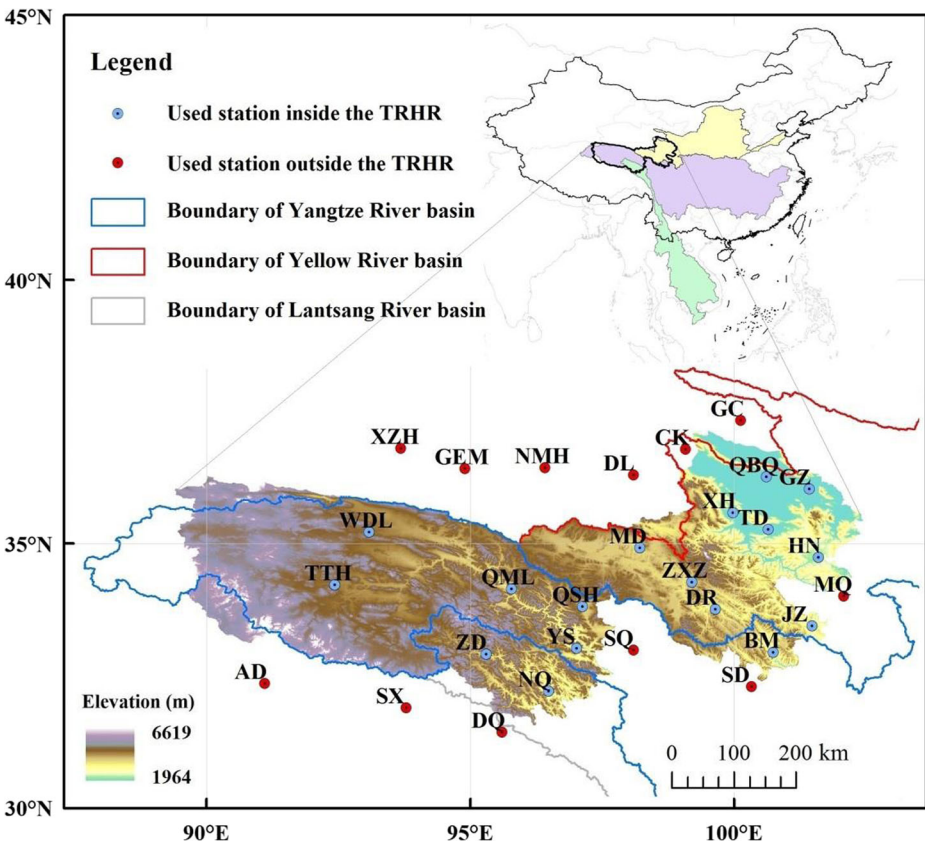


Fig. 1 The locations of meteorological stations over the TRHR

using the data of the neighboring stations in the same year; however, stations with severe lack of data for more than twenty years are excluded directly (Shi et al. 2017). After this elimination, there are 29 meteorological stations left with complete daily observations from 1961 to 2014. Among them, 17 stations are located inside the TRHR (the blue points in Fig. 1) and the others are located outside the TRHR (the red points in Fig. 1 and the stations in italic format in Table 1). For each station, the annual and monthly observed precipitation can be obtained from the daily observed data. Moreover, the mean annual and monthly precipitation over the TRHR can be computed using the Thiessen polygon method (Thiessen and Alter 1911; Brassel and Reif 1979). In this study, the observations during 1961–2000 were used for calibration of the proposed approach, while the remaining data during 2001–2014 were used for validation.

2.2 Statistical Method for Computing the Preliminary Estimations of Precipitation

Both the observed precipitation and the preliminary estimations of precipitation are the essential inputs of the recursive approach. In this study, a statistical method is proposed to compute the preliminary estimations of precipitation, firstly at the annual scale and then at the monthly scale.

Table 1 General information of the meteorological stations over the TRHR

Station name	Abbreviation	Longitude (°E)	Latitude (°N)	Elevation (m)	Years with observed data
Qiaboqia	QBQ	100.62	36.27	2835	1953–2014
Guizhou	GZ	101.43	36.03	2237	1956–2014
Wudaoliang	WDL	93.08	35.22	4612	1956–2014
Xinghai	XH	99.98	35.58	3323	1960–2014
Tongde	TD	100.65	35.27	3289	1954–1998
Tuotuohe	TTH	92.43	34.22	4533	1956–2014
Zaduo	ZD	95.30	32.90	4066	1956–2014
Qumalai	QML	95.78	34.13	4175	1956–2014
Yushu	YS	97.02	33.02	3681	1951–2014
Maduo	MD	98.22	34.92	4272	1953–2014
Qingshuihe	QSH	97.13	33.80	4415	1956–2014
Zhongxinzhan	ZXZ	99.20	34.27	4211	1959–1997
Dari	DR	99.65	33.75	3968	1956–2014
Henan	HN	101.60	34.73	3519	1959–2014
Jiuzhi	JZ	101.48	33.43	3629	1958–2014
Nangqian	NQ	96.48	32.20	3644	1956–2014
Banma	BM	100.75	32.93	3503	1960–2014
<i>Seda</i>	<i>SD</i>	<i>100.33</i>	<i>32.28</i>	<i>3929</i>	<i>1961–2014</i>
<i>Suoxian</i>	<i>SX</i>	<i>93.78</i>	<i>31.88</i>	<i>4023</i>	<i>1956–2014</i>
<i>Dingqing</i>	<i>DQ</i>	<i>95.60</i>	<i>31.42</i>	<i>3873</i>	<i>1954–2014</i>
<i>Maqu</i>	<i>MQ</i>	<i>102.08</i>	<i>34.00</i>	<i>3471</i>	<i>1967–2014</i>
<i>Shiqu</i>	<i>SQ</i>	<i>98.10</i>	<i>32.98</i>	<i>4086</i>	<i>1960–2014</i>
<i>Anduo</i>	<i>AD</i>	<i>91.10</i>	<i>32.35</i>	<i>4800</i>	<i>1965–2014</i>
<i>Xiaozhaohuo</i>	<i>XZH</i>	<i>93.68</i>	<i>36.80</i>	<i>2767</i>	<i>1960–2014</i>
<i>Gangcha</i>	<i>GC</i>	<i>100.13</i>	<i>37.33</i>	<i>3321</i>	<i>1957–2014</i>
<i>Geermu</i>	<i>GEM</i>	<i>94.9</i>	<i>36.42</i>	<i>2807.6</i>	<i>1955–2014</i>
<i>Nuomuhong</i>	<i>NMH</i>	<i>96.42</i>	<i>36.43</i>	<i>2790</i>	<i>1956–2014</i>
<i>Dulan</i>	<i>DL</i>	<i>98.10</i>	<i>36.30</i>	<i>3191</i>	<i>1954–2014</i>
<i>Chaka</i>	<i>CK</i>	<i>99.08</i>	<i>36.78</i>	<i>3088</i>	<i>1955–2000</i>

Stations in italic format are located outside the TRHR

2.2.1 Annual Precipitation

Known from our previous studies (Shi et al. 2016, 2017), the annual precipitation over the TRHR has experienced a significant increasing trend during 1961–2014. Thus, it is supposed that, for the designated year, the preliminary estimation of annual precipitation, $P_{\text{ann, est}}$, can be expressed as follows:

$$P_{\text{ann, est}} = P_{\text{ann, trend}} + P_{\text{ann, var}} \quad (1)$$

where $P_{\text{ann, trend}}$ and $P_{\text{ann, var}}$ denote the tendency and variation parts of the preliminary estimation of annual precipitation, respectively. The tendency part can be easily obtained using the linear regression method based on the observed annual precipitation during the study period:

$$P_{\text{ann, trend}} = a \cdot Y + b \quad (2)$$

where Y denotes the year number (e.g., 1961, 1962, ..., 2014 in this study), and a and b are the coefficients of the regression equation.

For the variation part of the preliminary estimation of annual precipitation, the following method is proposed. First, for each year during the study period, the difference between the observed annual precipitation and the tendency part ($P_{\text{ann, trend}}$) is calculated, and a set of differences (S_{diff}) can be obtained. Second, assuming that S_{diff} obeys the normal distribution (see Subsection 3.1 for details), the variation part ($P_{\text{ann, var}}$) for each year during the study period can be calculated through generating random numbers from the normal distribution function with mean and standard deviation of S_{diff} . Then, together with the tendency part ($P_{\text{ann, trend}}$), the preliminary estimation of annual precipitation for each year ($P_{\text{ann, est}}$) can be computed as the sum of $P_{\text{ann, trend}}$ and $P_{\text{ann, var}}$ (see Eq. (1)).

2.2.2 Monthly Precipitation

In order to compute the preliminary estimation of monthly precipitation for each month of a year ($P_{\text{mon, est, } i}$, $i = 1, 2, \dots, 12$), the following method is proposed. First, the mean annual precipitation ($P_{\text{ann, mean}}$) and the mean monthly precipitation for each month of a year during the study period ($P_{\text{mon, mean, } i}$, $i = 1, 2, \dots, 12$) are calculated with the observed precipitation. Second, the percentage of the mean monthly precipitation for each month of a year ($PCT_{\text{mon, mean, } i}$, $i = 1, 2, \dots, 12$) is computed as $P_{\text{mon, mean, } i} / P_{\text{ann, mean}}$. Third, using the preliminary estimation of annual precipitation for each year ($P_{\text{ann, est}}$) which has been obtained in the previous subsection, the preliminary estimation of monthly precipitation for each month of a year ($P_{\text{mon, est, } i}$) can then be computed as the product of $PCT_{\text{mon, mean, } i}$ and $P_{\text{ann, est}}$. The mathematical formula to calculate $P_{\text{mon, est, } i}$ can be expressed as follows:

$$P_{\text{mon, est, } i} = \frac{P_{\text{mon, mean, } i}}{P_{\text{ann, mean}}} \cdot P_{\text{ann, est}}, \quad i = 1, 2, \dots, 12 \quad (3)$$

2.3 Optimization Method for Improving the Preliminary Estimations

In this study, GP is adopted as an error updating scheme to improve the preliminary estimations of precipitation, which are obtained from the proposed statistical method. The details are introduced in the following.

2.3.1 Brief Introduction of GP

GP, which was firstly proposed by Cramer (1985) and later greatly expanded by Koza (1992, 1994), is a type of evolutionary algorithm, a subset of machine learning. Normally, GP implements an algorithm that uses random crossover, mutation, a fitness function, and multiple generations of evolution to resolve a user-defined task. GP is applicable to automatically discovering a functional relationship among features in data (i.e., symbolic regression, instead of traditional numerical regression). GP has been successfully used in various fields with complex optimization and search problems (e.g., Nordin and Banzhaf 1997; Muttill and Lee 2005; Wedge et al. 2009; Xu et al. 2016), including hydrology (e.g., Babovic and Keijzer 2000; Liong et al. 2002; Chadalawada et al. 2017).

Generally, GP gives each solution in a tree structure, with an operator function (e.g., four rules of arithmetic, trigonometric function, exponential function, logarithmic function, and so on) in every tree node and an operand (e.g., variable and number) in every terminal node, necessitating the evaluation of mathematical and logical expressions (Fig. 2). Crossover and mutation are the two major processes of producing new individuals in GP. In the crossover, two individuals (i.e., parents) are selected and sub-tree crossover randomly selects a crossover tree node in each parent tree (Fig. 2a). Then, two new individuals (i.e., children) are produced by exchanging the two selected crossover tree nodes, as illustrated in Fig. 2a. By contrast, in the mutation, each tree node is randomly considered and exchanged by another operator function with a certain probability (Fig. 2b). Therefore, it does not usually change the tree structure but the information content in the parse tree, which may serve to free the search from the possibility of being trapped in local optima (Khu et al. 2001; Fallah-Mehdipour et al. 2012).

In this study, Genetic Symbolic Regression (GSR), which is a special application of GP in the field of symbolic regression (Khu et al. 2001), will be used as the optimization method for improving the preliminary estimations of precipitation. In symbolic regression, both the most suitable functional form and the coefficients should be determined. This is quite different from traditional numerical regression, in which the functional form is pre-determined and only the coefficients should be determined. Thus, based on the observed precipitation and the preliminary estimations of precipitation, GSR will be conducted to explore a mathematical expression in symbolic form in this study.

2.3.2 Application of GP in the Recursive Approach

It is worth noting that an important application of GP is real-time prediction (e.g., Khu et al. 2001; Muttill and Lee 2005; Gaur and Deo 2008; Fallah-Mehdipour et al. 2012; Xu et al. 2016). In this study, such real-time prediction indicates that the observed monthly precipitation of the current year can be used as the input to predict the monthly precipitation of the next year. As a result, once the observed monthly precipitation in the initial year and the preliminary estimations of monthly precipitation in the initial year and the next year are available, GP can be applied to improve the preliminary estimations of monthly precipitation in the next year, and thus, the more accurate predictions of monthly precipitation in the next year can be obtained.

Mathematically, for the designated month, the observed monthly precipitation in the year t , $P_{\text{mon, obs, } t}$ can be expressed as follows:

$$P_{\text{mon, obs, } t} = P_{\text{mon, est, } t} + \varepsilon_t \quad (4)$$

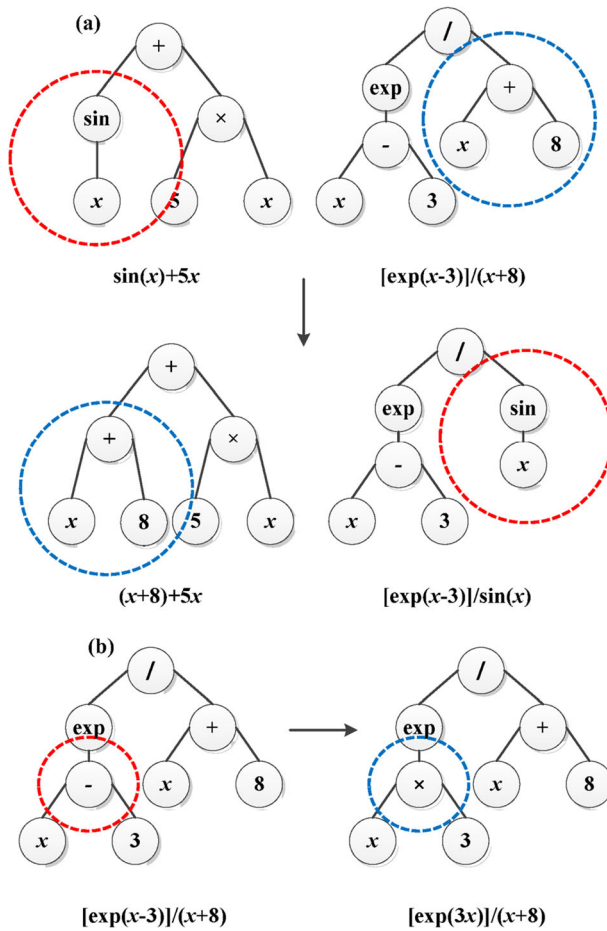


Fig. 2 (a) Crossover and (b) mutation in GP

where $P_{\text{mon, est, } t}$ is the preliminary estimation of monthly precipitation in the year t , and ε_t is the estimation error of monthly precipitation in the year t . This study assumes that the estimation error of monthly precipitation in a year will only be affected by the preliminary estimation of monthly precipitation and the estimation error in the previous year. Therefore, GSR is used to establish the functional relationship between the estimation error in the year t_{next} ($\varepsilon_{t_{\text{next}}}$) and other variables, including the preliminary estimation of monthly precipitation in the year t_{next} ($P_{\text{mon, est, } t_{\text{next}}}$), the preliminary estimation of monthly precipitation in the year t ($P_{\text{mon, est, } t}$), and the estimation error in the year t (ε_t). As the form of functional relationship generated from GSR can be various, this study uses a universal symbol, i.e., $f()$, to denote the functional relationship. The mathematical formula to calculate $\varepsilon_{t_{\text{next}}}$ can be expressed as follows:

$$\varepsilon_{t_{\text{next}}} = f(P_{\text{mon, est, } t_{\text{next}}}, P_{\text{mon, est, } t}, \varepsilon_t) \tag{5}$$

Then, the improved estimation of monthly precipitation in the year t_{next} , $P_{\text{mon, imp, } t_{\text{next}}}$, can be obtained:

$$P_{\text{mon,imp,t,next}} = P_{\text{mon,est,t,next}} + \varepsilon_{t_next} \tag{6}$$

Moreover, it is worth noting that two types of ε_t values may be used in Eq. (5). The first type is the actual estimation error in the case that the observed monthly precipitation is available; in this study, this type of ε_t values is used for calibration and validation (see Subsection 3.2 for details). However, for future precipitation prediction, the observed monthly precipitation is unavailable; in that case, the improved estimation of monthly precipitation can be regarded as the substitution of the observed monthly precipitation. As a result, the second type is the estimation error between the improved estimation of monthly precipitation and the preliminary estimation of monthly precipitation; this type of ε_t values is only used for prediction (see Subsection 3.3 for details).

Then, the procedures of the recursive approach to long-term prediction of monthly precipitation using GP can be summarized as follows:

1. For each year during the study period, the statistical method proposed in Subsection 2.2 is applied to compute the preliminary estimations of monthly precipitation for each month of that year (i.e., $P_{\text{mon, est, } i}$, $i = 1, 2, \dots, 12$);
2. For the designated month in the year t , the estimation error between $P_{\text{mon, est, } t}$ and $P_{\text{mon, obs, } t}$ (or $P_{\text{mon, imp, } t}$), i.e., ε_t , is computed using Eq. (4);
3. For the corresponding month in the year t_next , GSR is used to derive the functional relationship of the estimation error (i.e., ε_{t_next}) with the preliminary estimation of monthly precipitation in the year t_next ($P_{\text{mon, est, } t_next}$), the preliminary estimation of monthly precipitation in the year t ($P_{\text{mon, est, } t}$), and the estimation error in the year t (i.e., ε_t), as shown in Eq. (5);
4. Finally, the improved estimation of monthly precipitation in the year t_next (i.e., $P_{\text{mon, imp, } t_next}$) is calculated as the sum of $P_{\text{mon, est, } t_next}$ and ε_{t_next} , as shown in Eq. (6).

2.4 Assessment Criterion

To evaluate the performances of the prediction results, the *NSCE* (Nash-Sutcliffe Coefficient of Efficiency) (Nash and Sutcliffe 1970) is selected as assessment criterion, and the relevant equation is given as follows:

$$NSCE = 1 - \frac{\sum_{j=1}^N (X_{j,\text{obs}} - X_{j,\text{pre}})^2}{\sum_{j=1}^N (X_{j,\text{obs}} - \bar{X}_{\text{obs}})^2} \tag{7}$$

where $X_{i,\text{obs}}$ and $X_{i,\text{pre}}$ are the j -th observation and prediction, respectively; \bar{X}_{obs} is the mean value of the observations; and N is the sample size. The *NSCE* can measure the goodness of fit, and its value will approach 1.0 if the prediction results are close to the observations.

In addition, to further evaluate the performances of the prediction results, residual analysis is implemented (Shi et al. 2014), and the equation for computing the standardized residual is given as follows:

$$SR_j = e_j/s \tag{8}$$

where SR_j is the j -th standardized residual; e_j is the j -th residual; s is the standard deviation of the residuals. e_j and s can be computed as follows:

$$e_j = X_{j,obs} - X_{j,pre} \quad (9)$$

$$s = \sqrt{\frac{1}{N-2} \sum_{j=1}^N e_j^2} \quad (10)$$

3 Results and Discussion

3.1 Preliminary Estimations of Precipitation

Based on the observed precipitation, the mean annual precipitation ($P_{ann, mean}$) was 423.0 mm over the TRHR during 1961–2014, and the annual precipitation showed a significant increasing trend with the change rate of 7.11 mm/decade at the significance level of $p < 0.1$. The maximum annual precipitation (i.e., 514.8 mm) occurred in 1989, while the minimum one (i.e., 362.9 mm) occurred in 1969 (Fig. 3a). Moreover, the annual precipitation showed a decreasing trend with the change rate of -2.6 mm/decade during 1961–2002 but a significant increasing trend with the change rate of 19.3 mm/decade during 2003–2014 (Shi et al. 2016). Therefore, it was infeasible to calibrate and validate the statistical method for computing the preliminary estimations of precipitation; instead, to present the overall trend during 1961–2014 which would be regarded as the trend during 2015–2050, all the data during 1961–2014 were used to make the preliminary estimations of precipitation be close to the observed precipitation as much as possible. Following the method proposed in Subsection 2.2.1, the preliminary estimations of annual precipitation were computed. Based on the linear regression method, the tendency part of annual precipitation (i.e., $P_{ann, trend}$) could be preliminarily estimated as follows:

$$P_{ann, trend} = 0.711 \cdot Y - 990.13 \quad (11)$$

Then, for each year during 1961–2014, the difference between the observed annual precipitation and the tendency part was calculated, and the result (i.e., S_{diff}) is shown in Fig. 3b. It is observed that the differences varied between -61.32 mm and 90.74 mm, with the mean of 0.0085 mm and the standard deviation of 36.69 mm. Moreover, the distribution of S_{diff} was checked by Kolmogorov-Smirnov test at the significance level of $p < 0.05$, which confirmed that S_{diff} obeyed the normal distribution. Thus, the variation part of annual precipitation (i.e., $P_{ann, var}$) for each year during 1961–2014 could be preliminarily estimated, and the preliminary estimation of annual precipitation ($P_{ann, est}$) for each year during 1961–2014 was computed as the sum of $P_{ann, trend}$ and $P_{ann, var}$.

It is worth noting that random numbers were generated for $P_{ann, var}$, which might bring uncertainty in estimation; thus, one set of $P_{ann, var}$ would not be enough to represent the performance of the statistical method for computing the preliminary estimations of precipitation. In this study, ten sets of $P_{ann, var}$ were generated, and thus ten sets of $P_{ann, est}$ were obtained. For each year during 1961–2014, there were ten $P_{ann, est}$ values, among which, the mean, maximum and minimum $P_{ann, est}$ values were selected, respectively. The left part of Fig. 4

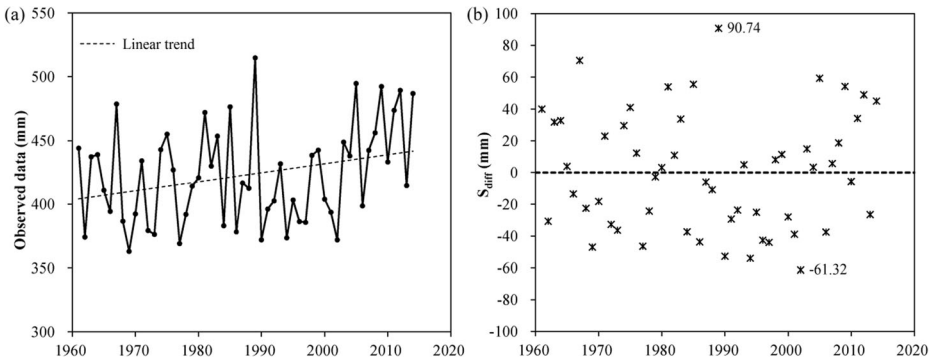


Fig. 3 (a) The annual precipitation over the TRHR during 1961–2014 and the linear trend; and (b) The differences between the observed annual precipitation and the tendency part of the annual precipitation over the TRHR during 1961–2014

shows the preliminary estimations of annual precipitation during 1961–2014, which would be used for calibration and validation of the proposed recursive approach in the next subsection. The observed annual precipitation during 1961–2014 (marked by “+”) could basically be covered by the range between the maximum and minimum $P_{ann, est}$ values (i.e., the solid lines), and the mean $P_{ann, est}$ values (marked by “o”) generally showed the increasing trend.

In this study, the mean $P_{ann, est}$ values were used to compute the preliminary estimation of monthly precipitation for each month of a year ($P_{mon, est, i}$, $i = 1, 2, \dots, 12$). According to the method proposed in Subsection 2.2.2, the percentage of the mean monthly precipitation for each month of a year ($PCT_{mon, mean, i}$, $i = 1, 2, \dots, 12$) should be firstly computed. Based on the observed precipitation, the mean monthly precipitation for each month of a year ($P_{mon, mean, i}$, $i = 1, 2, \dots, 12$) during 1961–2014 was obtained (Table 2). It is observed that the maximum value of the mean monthly precipitation (i.e., 92.7 mm) occurred in July, followed by the mean monthly precipitation in June (i.e., 81.3 mm) and August (i.e., 78.4 mm), and the minimum value (i.e., 1.9 mm) occurred in December. Because the mean annual precipitation ($P_{ann, mean}$) was 423.0 mm during 1961–2014, the percentage of the mean monthly precipitation for each month of a year ($PCT_{mon, mean, i}$, $i = 1, 2, \dots, 12$) could be computed (Table 2). Then, for each year during 1961–2014, the preliminary estimation of monthly precipitation for each month ($P_{mon, est, i}$, $i = 1, 2, \dots, 12$) was computed using Eq. (3).

Figure 5 shows the comparison of the preliminary estimations of monthly precipitation (marked by “□”) against the observations for each month of a year during 1961–2014. It is

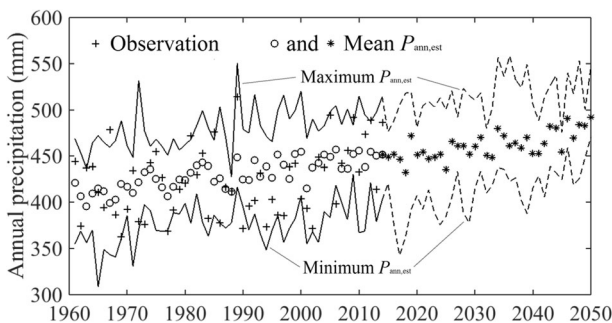


Fig. 4 The preliminary estimations of annual precipitation during 1961–2050. Note: the period of 1961–2014 is used for calibration and validation, and the period of 2015–2050 is used for prediction

Table 2 The mean monthly precipitation and the percentage accounted for by the mean monthly precipitation for each month of a year during 1961–2014

Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
$P_{\text{mon, mean, } i}(\text{mm})$	3.2	4.3	8.1	15.9	44.6	81.3	92.7	78.4	64.6	23.6	3.8	1.9
$PCT_{\text{mon, mean, } i}(\%)$	0.8	1.0	1.9	3.8	10.6	19.3	21.9	18.6	15.3	5.6	0.9	0.5

observed that the preliminary estimations were basically close to the observations and uniformly distributed on both sides of the red dash line, and the *NSCE* value was relatively high (i.e., 0.9195). However, there were more underestimated preliminary estimations in the case of large observations. The result of residual analysis (marked by “□” in Fig. 6) showed that the standardized residuals of the preliminary estimations were generally within ± 2 when the preliminary estimations were smaller than 60 mm, but a number of points were beyond the range of ± 2 when the preliminary estimations were larger than 60 mm (Fig. 6). This could also be represented by the significant deviations of the points from the red dash line in Fig. 5. It is worth noting that the preliminary estimations larger than 60 mm mainly appeared in the wet season, especially from June to September. Therefore, the preliminary estimations of monthly precipitation should be further improved, especially for those in the wet season.

3.2 Improved Estimations of Precipitation

In this study, GP was considered as the optimization method for improving the preliminary estimations of monthly precipitation, and the main GP relevant parameters included size of population (100), number of generation (100), tournament size (6), and operator functions (times, minus, plus, sqrt, square). As mentioned above, the observations and preliminary estimations during 1961–2000 were used for calibration of the proposed method, while the remaining data during

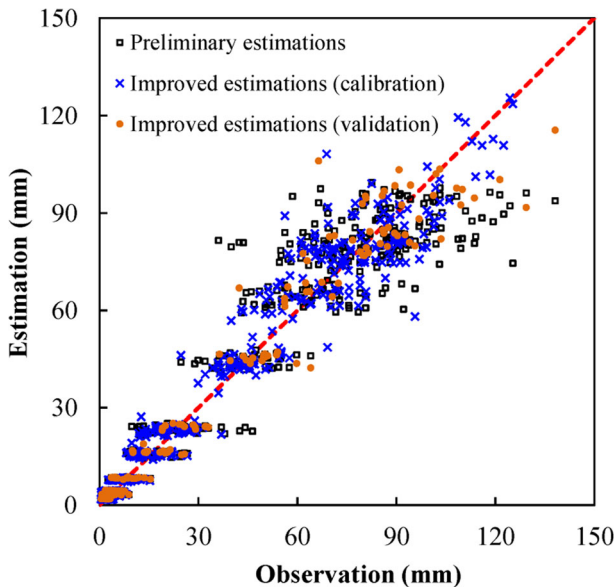


Fig. 5 Comparison of the preliminary/improved estimations of monthly precipitation against the observations for each month of a year during 1961–2014

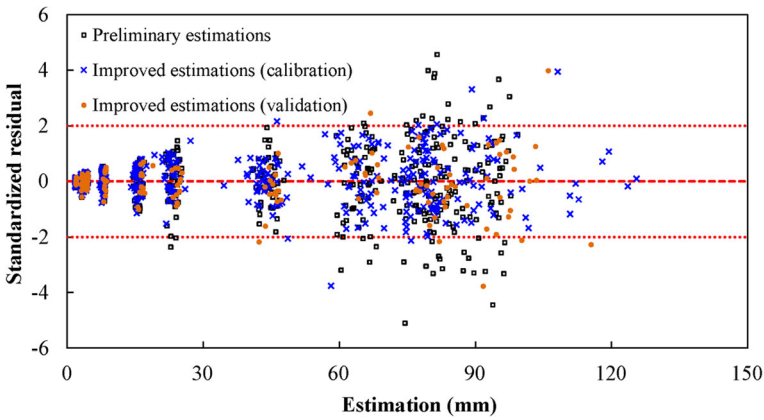


Fig. 6 Residual analysis based on the preliminary/improved estimations of monthly precipitation for each month of a year during 1961–2014

2001–2014 were used for validation. Before optimization, the *NSCE* values derived from the observations and preliminary estimations (i.e., Case I) were 0.9190 and 0.9205 for the two periods of 1961–2000 and 2001–2014, respectively (Table 3). Following the procedures in Subsection 2.3.2, the optimization results were obtained and presented as follows.

3.2.1 Calibration

In this study, GP was applied to derive the functional relationship as given in Eq. (5), and the improved estimations of monthly precipitation were calculated using Eq. (6). It is worth noting that two cases of the improved estimations of monthly precipitation (i.e., Case II and Case III) were considered during the calibration period of 1961–2000. Case II denoted that the improved estimations, which were computed considering all the preliminary estimations as a whole, were compared against the observations. Case III denoted that the improved estimations, which were computed considering the preliminary estimations with the standardized residuals beyond ± 2 separately, were compared against the observations. It is observed from Table 3 that the *NSCE* value derived from Case II (i.e., 0.9222) was only a little higher than that derived from Case I (i.e., 0.9190), while the *NSCE* value derived from Case III (i.e., 0.9520) was much higher than those derived from Case I and Case II. As a result, this study regarded the improved estimations derived from Case III as the better ones and the relevant functional relationships derived from GP were expressed as follows:

$$\begin{aligned}
 & P_{\text{mon,est,t,next}} \leq 60 \text{ mm} : \\
 & P_{\text{mon,imp,t,next}} = 1.023 \cdot P_{\text{mon,est,t,next}} - 1.023 \cdot |P_{\text{mon,est,t,next}} - P_{\text{mon,est,t}}| \\
 & \quad - 0.00004854 \cdot P_{\text{mon,est,t,next}}^2 \cdot [\varepsilon_t + (P_{\text{mon,est,t,next}} - P_{\text{mon,est,t}})] \\
 & \quad + 0.00141 \cdot \varepsilon_t \cdot P_{\text{mon,est,t}} \cdot |P_{\text{mon,est,t,next}} - P_{\text{mon,est,t}}| + 0.0001584 \\
 & \quad P_{\text{mon,est,t,next}} > 60 \text{ mm} : \\
 & P_{\text{mon,imp,t,next}} = P_{\text{mon,est,t,next}} + \left[12.6166 \cdot \varepsilon_t + 71.9757 \cdot (P_{\text{mon,est,t,next}} - P_{\text{mon,est,t}}) \right. \\
 & \quad - 0.3897 \cdot \varepsilon_t^2 - 10.4948 \cdot (P_{\text{mon,est,t,next}} - P_{\text{mon,est,t}})^2 \\
 & \quad \left. - 3.9947 \cdot \varepsilon_t \cdot (P_{\text{mon,est,t,next}} - P_{\text{mon,est,t}}) - 51.3250 \right] \\
 & \quad / \left[1 - 0.3686 \cdot \varepsilon_t - 1.9015 \cdot (P_{\text{mon,est,t,next}} - P_{\text{mon,est,t}}) \right. \\
 & \quad \left. + 0.04282 \cdot \varepsilon_t^2 + 1.1741 \cdot (P_{\text{mon,est,t,next}} - P_{\text{mon,est,t}})^2 \right. \\
 & \quad \left. + 0.4571 \cdot \varepsilon_t \cdot (P_{\text{mon,est,t,next}} - P_{\text{mon,est,t}}) \right] \tag{12}
 \end{aligned}$$

Figure 5 shows the comparison of the improved estimations of monthly precipitation derived from Case III (marked by blue “×”) against the observations for each month of a year during the calibration period of 1961–2000. The improved estimations were closer to the observations than the preliminary estimations (marked by “□”), especially for those larger than 60 mm. The result of residual analysis on the improved estimations (marked by blue “×” in Fig. 6) showed that most of the standardized residuals were within ± 2 even if the improved estimations were larger than 60 mm (Fig. 6). Moreover, it is worth noting that there were more improved estimations than preliminary estimations which were distributed between 90 mm and 120 mm, especially for those around 120 mm. This indicated that the larger observations could be better estimated by the improved estimations obtained from the proposed method.

3.2.2 Validation

Based on Eq. (12), the improved estimations of monthly precipitation for each month of a year during the period of 2001–2014 could be computed, and then, the validation of the proposed method was conducted. Figure 5 also shows the comparison of the improved estimations of monthly precipitation (marked by orange “.”) against the observations for each month of a year during the validation period of 2001–2014. The result indicated that the improved estimations were closer to the observations than the preliminary estimations (marked by “□”), and the distribution of the improved estimations during the validation period was similar to that during the calibration period. Moreover, Fig. 6 shows that most of the standardized residuals of the improved estimations during the validation period (marked by orange “.”) were within ± 2 , similar to those during the calibration period. Table 3 lists the *NSCE* values of the three cases (i.e., Case I, Case II and Case III), and the results derived from Case II (i.e., 0.9243) was only a little higher than that derived from Case I (i.e., 0.9205), while the *NSCE* value derived from Case III (i.e., 0.9517) was much higher than those derived from Case I and Case II.

To further evaluate the performance of the proposed recursive approach in precipitation prediction, additional validation was conducted as follows. For a designated year during the validation period of 2001–2014, the observations in the previous year were substituted by the improved estimations of monthly precipitation in the previous year, and then, the improved estimations of monthly precipitation in the designated year were computed for Case II and Case III, respectively. The relevant *NSCE* values were 0.9237 for Case II and 0.9524 for Case III, respectively. It indicated that, even when the observations were unavailable and the improved estimations were used as the substitutions of the observations for prediction, the performance of the proposed recursive approach could be as good as that using the observations directly. Therefore, the recursive approach proposed in this study could be adopted for

Table 3 The *NSCE* values of the three cases

Period	Case I	Case II	Case III
Calibration (1961–2000)	0.9190	0.9222	0.9520
Validation (2001–2014)	0.9205	0.9243	0.9517

Case I denotes that the preliminary estimations are compared against the observations. Case II denotes that the improved estimations, which are computed considering all the preliminary estimations as a whole, are compared against the observations. Case III denotes that the improved estimations, which are computed considering the preliminary estimations with the standardized residuals beyond ± 2 separately, are compared against the observations

the long-term prediction of monthly precipitation over the TRHR (see the next subsection for details). Nevertheless, the estimations of monthly precipitation could be further updated if the observations became available for subsequent predictions (Khu et al. 2001).

3.3 Long-Term Prediction of Monthly Precipitation

After calibration and validation, the proposed recursive approach has been proved to be applicable to long-term prediction of monthly precipitation over the TRHR. In this study, the predictions of monthly precipitation over the TRHR till 2050 were computed based on this recursive approach.

First, the preliminary estimations of annual precipitation were computed. For each year during 2015–2050, the tendency part ($P_{\text{ann, trend}}$) and the variation part ($P_{\text{ann, var}}$) of annual precipitation were estimated, and the preliminary estimation of annual precipitation ($P_{\text{ann, est}}$) was computed as the sum of $P_{\text{ann, trend}}$ and $P_{\text{ann, var}}$. Similar to that in Subsection 3.1, ten sets of $P_{\text{ann, var}}$ (as well as $P_{\text{ann, est}}$) were generated, and the mean, maximum and minimum $P_{\text{ann, est}}$ values for each year during 2015–2050 were selected, respectively. The right part of Fig. 4 shows the relevant results.

Second, the mean $P_{\text{ann, est}}$ values were used to compute the preliminary estimations of monthly precipitation for each month of a year during 2015–2050 ($P_{\text{mon, est, } i}$, $i = 1, 2, \dots, 12$), following the method proposed in Subsection 2.2.2.

Third, the improved estimations of monthly precipitation were calculated using Eq. (12). Since the observations during 2015–2050 were unavailable, the improved estimations of monthly precipitation in the previous year were regarded as the substitution of the observations in the previous year when applying GP.

Figure 7a shows the predictions (both the preliminary estimations and the improved estimations) of monthly precipitation for each month of a year during 2015–2050. It is observed that the series of the improved estimations had the more significant variation than the series of the preliminary estimations. In addition, some months in the wet season (e.g., July 2029 and July 2041) would have much higher values of precipitation than other months. Figure 7b shows the predictions of annual precipitation derived from the improved estimations for each year during 2015–2050. It is observed that the annual precipitation would present an overall increasing trend with significant inter-annual variation. This is reasonable because the same variation of the annual precipitation was found over the TRHR during 1961–2014 (Shi et al. 2016) (see Fig. 3a).

3.4 Discussion

It is worth noting that, in this study, ten sets of $P_{\text{ann, est}}$ were generated to address the problem that one set of $P_{\text{ann, est}}$ would not be enough to represent the performance of the statistical method; therefore, the impact of the number of sets should be discussed. In this section, comparative analysis is conducted between two results from ten sets and twenty sets of $P_{\text{ann, est}}$, respectively. For the result from ten sets of $P_{\text{ann, est}}$, the variation ranges of the mean, maximum and minimum values were 381–469 mm, 406–549 mm, and 320–426 mm, respectively; the averages of the mean, maximum and minimum values were 425 mm, 472 mm, and 375 mm, respectively. For the result from twenty sets of $P_{\text{ann, est}}$, the variation ranges of the mean, maximum and minimum values were 383–468 mm, 445–567 mm, and 294–413 mm, respectively; the averages of the mean, maximum and minimum values were 424 mm, 490 mm, and

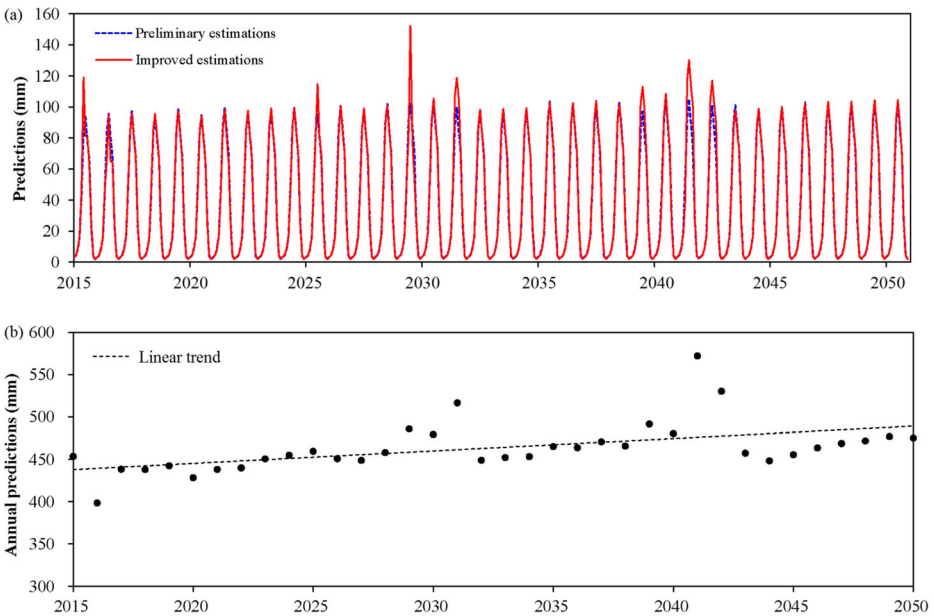


Fig. 7 Predictions of (a) monthly precipitation for each month of a year during 2015–2050 and (b) annual precipitation for each year during 2015–2050

358 mm, respectively. It could be concluded that the maximum values would generally increase with larger number of sets, while the minimum values would generally decrease with larger number of sets. In contrast, the number of sets would not significantly influence the mean values. Since the mean values were used to compute the preliminary estimations of monthly precipitation in this study, the impact of the number of sets could be neglected.

Applying the proposed recursive approach for predicting monthly precipitation, we also need fully aware of the following four limitations. First, to some extent, the proposed recursive approach is region-dependent. The study area is located in a semi-arid region, and this approach is applicable. For different regions (e.g., humid regions), the variation characteristics of precipitation may be different, and thus the equations (e.g., Eqs. (11) and (12)) should be reestablished based on the observed data and the estimations. Second, this study adopted the mean $P_{\text{ann, est}}$ values for calculating the preliminary estimations of monthly precipitation, and thus, the results were in an average sense. If the maximum $P_{\text{ann, est}}$, the minimum $P_{\text{ann, est}}$, or any one of the ten sets of $P_{\text{ann, est}}$ values were used, the results might be different. Therefore, how to determine the most appropriate one is an important issue that is worth investigating in future works. Third, due to the quite different trends of annual precipitation between the two periods of 1961–2002 and 2003–2014, the statistical method for computing the preliminary estimations of precipitation was not calibrated and validated in this study, making the proposed recursive approach not a complete prediction scheme. However, for other regions with a consistent trend of annual precipitation, the proposed recursive approach could be further improved. Fourth, when predicting monthly precipitation over the TRHR during 2015–2050, the variation characteristics of precipitation were assumed to be the same as those during 1961–2014. Therefore, whether the variation characteristics of precipitation would change in the future is another issue that needs further study.

Nevertheless, with the awareness of the above limitations, the recursive approach proposed in this study can provide a new avenue of long-term prediction of monthly precipitation

independent of other climatic indices and variables, which would be valuable for hydrology research around the world.

4 Conclusions

This study proposes a recursive approach to long-term prediction of monthly precipitation using GP, taking the TRHR as the study area. The major contributions of this study can be summarized as follows:

First, a statistical method for computing the preliminary estimations of monthly precipitation was proposed. For the TRHR during 1961–2014, the preliminary estimations were basically close to the observations except for those large values appeared in the wet season, which should be further improved.

Second, an optimization method with application of GP was proposed to improve the preliminary estimations. For the TRHR during 1961–2014, the improved estimations of monthly precipitation showed much better performance than the preliminary estimations.

Third, based on the proposed recursive approach, the monthly precipitation for each month of a year during 2015–2050 was predicted. The results indicated that there would be some very wet months in the near future. From a realistic perspective, more precipitation, especially for short-duration, high-intensity rains in the wet season, might lead to an increased risk of floods; therefore, it would be important to formulate related policies for reducing the losses caused by possible floods.

Overall, the proposed methods and the related conclusions can potentially benefit the utilization of some hydrological models at the monthly scale (Wang et al. 2009; Gosling and Arnell 2011), evaluating climate change (Wang et al. 2009), and exploring water resources sustainability (Barros et al. 2011; Chen et al. 2016; Li et al. 2018).

Acknowledgments This study was supported by the Natural Science Foundation of Qinghai Province funded project (No. 2017-ZJ-911), Guangdong Provincial Key Laboratory of Soil and Groundwater Pollution Control (No. 2017B030301012), and State Environmental Protection Key Laboratory of Integrated Surface Water-Groundwater Pollution Control. The authors are also grateful to Editor, Associate Editor, and the two anonymous reviewers who offered the insightful comments leading to improvement of this paper.

Compliance with Ethical Standards

Conflict of Interest Statement None.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, duplication, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

- Babovic V, Keijzer M (2000) Genetic programming as a model induction engine. *J Hydroinf* 2(1):35–60
- Barnston AG, Smith TM (1996) Specification and prediction of global surface temperature and precipitation from global SST using CCA. *J Clim* 9:2660–2697

- Barros R, Isidoro D, Aragüés R (2011) Long-term water balances in La Violada irrigation district (Spain): I. sequential assessment and minimization of closing errors. *Agric Water Manag* 102:35–45
- Benestad RE (2013) Association between trends in daily rainfall percentiles and the global mean temperature. *J Geophys Res-Atmos* 118(19):10802–10810
- Brassel KE, Reif D (1979) A procedure to generate Thiessen polygons. *Geogr Anal* 11:289–303
- Cao LG, Pan SM (2014) Changes in precipitation extremes over the “Three-River Headwaters” region, hinterland of the Tibetan plateau, during 1960–2012. *Quat Int* 321:105–115
- Chadalawada J, Havlicek V, Babovic V (2017) A genetic programming approach to system identification of rainfall-runoff models. *Water Resour Manag* 31(12):3975–3992
- Chardon J, Hingray B, Favre A-C (2018) An adaptive two-stage analog/regression model for probabilistic prediction of small-scale precipitation in France. *Hydrol Earth Syst Sci* 22:265–286
- Chen J, Shi HY, Sivakumar B, Peart MR (2016) Population, water, food, energy and dams. *Renew Sust Energ Rev* 56:18–28
- China Meteorological Administration (2016) Daily meteorological observation data sets of China. http://data.cma.gov.cn/data/detail/dataCode/SURF_CLI_CHN_MUL_DAY_V3.0.html
- Chow VT, Maidment DR, Mays LW (1988) *Applied hydrology*. McGraw-Hill, New York
- Chu HJ (2012) Assessing the relationships between elevation and extreme precipitation with various durations in southern Taiwan using spatial regression models. *Hydrol Process* 26(21):3174–3181
- Claußnitzer A, Névír P (2009) Analysis of quantitative precipitation forecasts using the dynamic state index. *Atmos Res* 94(4):694–703
- Cramer NL (1985) A representation for the adaptive generation of simple sequential programs. In: John J (ed) *Proc of an inter Conf on genetic algorithms and the applications Grefenstette*. Carnegie Mellon University, Pittsburgh
- Duan AM, Hu J, Xiao ZX (2013) The Tibetan plateau summer monsoon in the CMIP5 simulations. *J Clim* 26(19):7747–7766
- Fallah-Mehdipour E, Haddad OB, Mariño MA (2012) Real-time operation of reservoir system by genetic programming. *Water Resour Manag* 26(14):4091–4103
- Garbrecht J, Van Liew M, Brown GO (2004) Trends in precipitation, streamflow, and evapotranspiration in the Great Plains of the United States. *J Hydrol Eng* 9(5):360–367
- Gaur S, Deo MC (2008) Real-time wave forecasting using genetic programming. *Ocean Eng* 35:1166–1172
- Gosling SN, Amell NW (2011) Simulating current global river runoff with a global hydrological model: model revisions, validation, and sensitivity analysis. *Hydrol Process* 25:1129–1145
- Groisman PY, Knight RW, Easterling DR, Karl TR, Hegerl GC, Razuvaev VN (2005) Trends in intense precipitation in the climate record. *J Clim* 18:1326–1350
- Immerzeel WW, van Beek LPH, Bierkens MFP (2010) Climate change will affect the Asian water towers. *Science* 328:1382–1385
- Intergovernmental Panel on Climate Change (IPCC) (2013) Summary for policymakers. In: Stocker TF et al (eds) *Climate change 2013: the physical science basis. Contribution of working group I to the fifth assessment report of the intergovernmental panel on climate change*. Cambridge University Press, Cambridge
- Khu ST, Liong SY, Babovic V, Madsen H, Muttill N (2001) Genetic programming and its application in real-time runoff forecasting. *J Am Water Resour Assoc* 37(2):439–451
- Kim J, Oh H-S, Lim Y, Kang H-S (2017) Seasonal precipitation prediction via data-adaptive principal component regression. *Int J Climatol* 37:75–86
- Kisi O, Cimen M (2011) A wavelet-support vector machine conjunction model for monthly streamflow forecasting. *J Hydrol* 399:132–140
- Koza JR (1992) *Genetic programming: on the programming of computers by means of natural selection*. MIT Press, Cambridge
- Koza JR (1994) *Genetic programming II: automatic discovery of reusable programs*. MIT Press, Cambridge
- Kulligowski RJ, Barros AP (1998) Localized precipitation from a numerical weather prediction model using artificial neural networks. *Weather Forecast* 13:1195–1205
- Landman WA, Beraki A, DeWitt D, Lötter D (2014) SST prediction methodologies and verification considerations for dynamical mid-summer rainfall forecasts for South Africa. *Water SA* 40:615–622
- Lee T, Ouarda TBMJ (2010) Long-term prediction of precipitation and hydrologic extremes with nonstationary oscillation processes. *J Geophys Res Atmos* 115:D13107. <https://doi.org/10.1029/2009JD012801>

- Li JY, Li TJ, Liu SN, Shi HY (2018) An efficient method for mapping high-resolution global river discharge based on the algorithms of drainage network extraction. *Water* 10(4):533. <https://doi.org/10.3390/w10040533>
- Liang LQ, Li LJ, Liu CM, Cuo L (2013) Climate change in the Tibetan plateau three Rivers source region: 1960–2009. *Int J Climatol* 33:2900–2916
- Liong SY, Gautam TR, Khu ST et al (2002) Genetic programming: a new paradigm in rainfall runoff modeling. *J Am Water Resour Assoc* 38(3):705–718
- Liu SN, Chui TFM (2018) Impacts of different rainfall patterns on hyporheic zone under transient conditions. *J Hydrol* 561:598–608
- Liu XD, Yin ZY (2002) Sensitivity of east Asian monsoon climate to the uplift of the Tibetan plateau. *Palaeogeogr Palaeoclimatol Palaeoecol* 183(3–4):223–245
- Muttill N, Lee JHW (2005) Genetic programming for analysis and real-time prediction of coastal algal blooms. *Ecol Model* 189(3–4):363–376
- Naoum S, Tsanis IK (2004) Orographic precipitation modeling with multiple linear regression. *J Hydrol Eng* 9(2):79–102
- Nash JE, Sutcliffe JV (1970) River flow forecasting through conceptual models part 1 - a discussion of principles. *J Hydrol* 10(3):282–290
- Nordin P, Banzhaf W (1997) Real time control of a Khepera robot using genetic programming. *Control Cybern* 26(3):533–561
- Ortiz-García EG, Salcedo-Sanz S, Casanova-Mateo C (2014) Accurate precipitation prediction with support vector classifiers: a study including novel predictive variables and observational data. *Atmos Res* 139:128–136
- Park Y-Y, Buizza R, Leutbecher M (2008) TIGGE: preliminary results on comparing and combining ensembles. European Centre for Medium-Range Weather Forecasts, Reading
- Partal T, Cigizoglu HK (2009) Prediction of daily precipitation using wavelet-neural networks. *Hydrol Sci J* 54: 234–246
- Richardson D (2005) The THORPEX Interactive Grand Global Ensemble (TIGGE). *Geophysical Research Abstracts* 7: Abstract EGU05-A-02815
- Shi HY, Fu XD, Chen J et al (2014) Spatial distribution of monthly potential evaporation over mountainous regions: case of the Lhasa River basin, China. *Hydrol Sci J* 59(10):1856–1871
- Shi HY, Li TJ, Wei JH et al (2016) Spatial and temporal characteristics of precipitation over the Three-River Headwaters region during 1961–2014. *J Hydrol Reg Stud* 6:52–65
- Shi HY, Li TJ, Wei JH (2017) Evaluation of the gridded CRU TS precipitation dataset with the point raingauge records over the Three-River Headwaters region. *J Hydrol* 548:322–332
- Silverman D, Dracup JA (2000) Artificial neural networks and long-range precipitation prediction in California. *J Appl Meteorol* 39:57–66
- Singh VP (1988) *Hydrologic systems: watershed modeling*. Prentice Hall, New Jersey
- Thiessen AJ, Alter JC (1911) Precipitation averages for large areas. *Mon Weather Rev* 39:1082–1084
- Tong LG, Xu XL, Fu Y, Li S (2014) Wetland changes and their responses to climate change in the “Three-River Headwaters” Region of China since the 1990s. *Energies* 7:2515–2534
- Wang GS, Xia J, Chen J (2009) Quantification of effects of climate variations and human activities on runoff by a monthly water balance model: a case study of the Chaobai River basin in northern China. *Water Resour Res* 45(7). <https://doi.org/10.1029/2007WR006768>
- Wedge DC, Das A, Dost R et al (2009) Real-time vapour sensing using an OFET-based electronic nose and genetic programming. *Sensors Actuators B Chem* 143(1):365–372
- Wei H, Li JL, Liang TG (2005) Study on the estimation of precipitation resources for rainwater harvesting agriculture in semi-arid land of China. *Agric Water Manag* 71:33–45
- Westra S, Alexander LV, Zwiers FW (2013) Global increasing trends in annual maximum daily precipitation. *J Clim* 26(11):3904–3918
- Xi Y, Miao CY, Wu JW et al (2018) Spatiotemporal changes in extreme temperature and precipitation events in the Three-Rivers Headwater Region, China. *J Geophys Res-Atmos* 123. <https://doi.org/10.1029/2017JD028226>
- Xu CC, Liu P, Wang W, Zhang Y (2016) Real-time identification of traffic conditions prone to injury and non-injury crashes on freeways using genetic programming. *J Adv Transp* 50(5):701–716
- Xue T, Xu J, Guan Z et al (2017) An assessment of the impact of ATMS and CrIS data assimilation on precipitation prediction over the Tibetan plateau. *Atmos Meas Tech* 10:2517–2531
- Yi XS, Li GS, Yin YY (2013) Spatio-temporal variation of precipitation in the Three-River Headwater Region from 1961 to 2010. *J Geogr Sci* 23(3):447–464

Affiliations

Suning Liu^{1,2,3} · Haiyun Shi^{1,2,4}

- ¹ State Environmental Protection Key Laboratory of Integrated Surface Water-Groundwater Pollution Control, School of Environmental Science and Engineering, Southern University of Science and Technology, Shenzhen, China
- ² Guangdong Provincial Key Laboratory of Soil and Groundwater Pollution Control, School of Environmental Science and Engineering, Southern University of Science and Technology, Shenzhen, China
- ³ Department of Civil Engineering, The University of Hong Kong, Hong Kong, China
- ⁴ State Key Laboratory of Plateau Ecology and Agriculture, Qinghai University, Xining, Qinghai, China