Revisiting the hydrological basis of the Budyko framework with the principle of hydrologically similar groups

Yuchan Chen¹, Xiuzhi Chen¹, Meimei Xue¹, Chuanxun Yang²,³, Wei Zheng¹, Jun Cao⁴, Wenting Yan¹, and Wenping Yuan¹

¹Guangdong Province Data Center of Terrestrial and Marine Ecosystems Carbon Cycle, Guangdong Province Key Laboratory for Climate Change and Natural Disaster Studies, School of Atmospheric Sciences, Sun Yat-sen University & Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai), Zhuhai, 519082, China
²Guangzhou Institute of Geochemistry, Chinese Academy of Sciences, Guangzhou, 510640, China
³College of Earth and Planetary Sciences, University of Chinese Academy of Sciences, Beijing, 100049, China
⁴Institute of Ecological Civilization and Green Development, Guangdong Provincial Academy of Environmental Science, Guangzhou, 510635, China

Correspondence: Xiuzhi Chen (chenxzh73@mail.sysu.edu.cn)

Received: 8 August 2022 – Discussion started: 2 September 2022
Revised: 14 April 2023 – Accepted: 21 April 2023 – Published: 17 May 2023

Abstract. The Budyko framework is a simple and effective tool for estimating the water balance of watersheds. Quantification of the watershed-characteristic-related parameter (Pw) is critical for accurate water balance simulations with the Budyko framework. However, there is no universal method for calculating Pw as the interactions between hydrologic, climatic, and watershed characteristic factors differ greatly across watersheds. To fill this research gap, this study introduced the principle of hydrologically similar groups into the Budyko framework for quantifying the Pw of watersheds in similar environments. We first classified the 366 selected watersheds worldwide into six hydrologically similar groups based on watershed attributes, including climate, soil, and vegetation. Results show that soil moisture (SM) and fractional vegetation cover (FVC) are two controlling factors of the Pw in each group. The SM exhibits a power-law relationship with the Pw values, with increasing SM leading to higher Pw values in dry watersheds (SM ≤ 20 mm) and lower Pw values in humid watersheds (SM > 20 mm). Additionally, the FVC shows to be linearly correlated with the Pw values in most hydrologically similar groups, except in that group with moist soil and no strong rainfall seasonality (SM > 20 mm and seasonal index (SI) ≤ 0.4). Multiple non-linear regression models between Pw and the controlling factors (SM and FVC) were developed to individually estimate the Pw of six hydrologically similar groups. Cross-validations using the bootstrap sampling method ($R^2 = 0.63$) and validations of time-series Global Runoff Data Centre (GRDC) data ($R^2 = 0.89$) both indicate that the proposed models perform satisfactorily in estimating the Pw parameter in the Budyko framework. Overall, this study is a new attempt to quantify the unknown Pw in the Budyko framework using the method for hydrologically similar groups. The results will be helpful in improving the applicability of the Budyko framework for estimating the annual runoff of watersheds in diverse climates and with different characteristics.

1 Introduction

There has been an increasing interest in estimating the water balance of watersheds with a simple and effective tool – the Budyko framework. Unlike process-based models that typically require a large number of parameters as inputs for accurate simulations (Caracciolo et al., 2018; Lei et al., 2014), the Budyko framework is a top-down approach that is rooted on a firm physical basis, relating a catchment’s long-term evaporative ratio (ratio between actual evapotranspiration and precipitation) to its aridity index (ratio between potential evapotranspiration and precipitation) (Vora and Singh, 2021; Sivapalan, 2003; Wang and Tang, 2014). Recently, the

Published by Copernicus Publications on behalf of the European Geosciences Union.
Budyko framework has been widely used for assessing linkages and feedbacks between climate forcing and land surface characteristics on water and energy cycles (Zhang et al., 2001; Milly and Shmakin, 2002; Li et al., 2013; Xu et al., 2013), prompting a great deal of empirical-, theoretical-, and process-based studies (Chen and Sivapalan, 2020; Roderick and Farquhar, 2011; Rau et al., 2018; Goswami and Goyal, 2022).

The original Budyko equation assumes that evapotranspiration is mainly controlled by precipitation (representing the availability of water) and potential evapotranspiration (representing the availability of energy) (Budyko, 1974; Wang et al., 2022). Despite its solid performance, the original Budyko equation still produces a bias between modeled and measured evapotranspiration or runoff because it does not consider the effects of watershed characteristics other than mean annual climatic conditions on water balance (Kim and Chun, 2021; Zhang et al., 2001). As a result, hydrologists have invested considerable efforts to improve model performance by introducing parameters related to watershed characteristics (watershed-characteristic-related parameter, Pw) into the original Budyko equation. Popular parametric equations of the Budyko framework are presented in Table 1.

From a hydrological point of view, Pw controls the fraction of precipitation diverted into runoff for a given aridity index (Caracciolo et al., 2018). Watersheds with higher Pw values partition more precipitation to evapotranspiration and consequently less to runoff than those with lower Pw values; some studies defined Pw as the water retention capacity of a watershed (Fu, 1981; Zhou et al., 2015). Overall, Pw denotes the adjustment of water–energy partitioning by various watershed characteristics (Yao et al., 2017; Li et al., 2013).

During the past decades, researchers have done lots of work to quantify Pw for the accurate simulation of evapotranspiration or runoff using the Budyko framework (Wang et al., 2022; Yao et al., 2017; Guo et al., 2019; Yu et al., 2021) and considerably improved the estimation of Pw by taking into account the influence of watershed characteristics (Fu, 1981; Liu and Liang, 2015; Guan et al., 2022; Yang et al., 2008). Although there is agreement that Pw represents the integrated effects of various environmental factors (Wang et al., 2022; S. Liu et al., 2022; Yu et al., 2021; Gan et al., 2021), studies still differed greatly as to what factors and effects should relate to Pw and failed to give a general framework for quantifying it. For instance, whether the Pw in the Budyko framework is controlled by vegetation or not has been much debated. Ning et al. (2017) found that Pw generally correlated positively with vegetation cover. Zhang et al. (2018) obtained the sensitivity of Pw to changes in leaf area index (LAI) by taking a derivative of the Pw function with respect to LAI, implying a crucial role of vegetation cover in impacting Pw. However, other studies indicated that most regions or watersheds show no significant influences of vegetation indices or cover on Pw (Li et al., 2013; Liu et al., 2021). For example, Li et al. (2013) noted that the variations in the Pw values are not entirely controlled by vegetation cover in small catchments. Another study by Liu et al. (2021) also found a weak correlation between the vegetation leaf area index and Pw. Therefore, more in-depth studies are needed for revisiting the hydrological basis of Pw in the Budyko framework.

Here, we hypothesize that watersheds with similar climatic, hydrologic, and watershed characteristics have consistent controlling factors of Pw in the Budyko framework. Classifying watersheds into groups that are hydrologically similar may help us identify how Pw responds to different watershed characteristic factors. However, to date, few studies have been conducted on classifying watersheds based on the highly variable hydro-climate–Pw relationships in the Budyko framework. This may be an important reason why researchers disagree about the factors and extent of the influence on Pw.

This study proposes a new approach to address the research gap in accurately estimating the Pw parameter in the Budyko framework by classifying watersheds into hydrologically similar groups and developing a framework for estimating Pw (PwM) in each group to simulate global runoff. More specifically, we collected 726 hydrological records in 366 watersheds from published literature for analyses. These 726 samples were classified into six hydrologically similar groups according to the hydrologically homogenous attributes of watersheds using the “decision tree regressor” method. Then, we identified the controlling factors of Pw from various environmental factors in each hydrologically similar group and developed multiple non-linear regression models for estimating Pw in the Budyko framework. We expect that classifying watersheds into hydrologically similar groups can help explore the effect of watershed characteristics on their water balance and interpret the physical meaning of the Pw in the Budyko framework. This study highlights the need to account for the interactions among hydrologic, climatic, and watershed characteristic factors for explaining Pw in the Budyko framework.

2 Fu’s formula

This study employed Fu’s formula (Zhou et al., 2015) to analyze Pw in the Budyko framework. Fu’s equation is a commonly used parametric equation in Budyko-type formulas due to its versatility and adaptability (Zhou et al., 2015). The formula is expressed as

$$\frac{R}{P} = \left(1 + \left(\frac{P}{PET}\right)^{-Pw}\right)^{\frac{1}{Pw}} - \left(\frac{P}{PET}\right)^{-1},$$

where $R/P$ is a dimensionless annual water yield coefficient, $P/PET$ is an aridity index, and Pw is a dimensionless constant varying from 1 to infinity and representing water retention capacity for evapotranspiration. When Pw = 1, all the precipitation becomes flow, and the residence time is 0.
Table 1. Parametric formulations of the Budyko framework (Pw – watershed-characteristic-related parameter; ET – actual evaporation, R – runoff, P – precipitation, PET – potential evapotranspiration, all in mm yr\(^{-1}\)).

<table>
<thead>
<tr>
<th>Reference</th>
<th>Formulation</th>
<th>Pw (theoretical range)</th>
<th>Reference values of Pw</th>
</tr>
</thead>
</table>
| Budyko (1974)                 | \[
|                               | \left[ \frac{\text{PET} \cdot \tanh\left(\frac{\text{PET}}{P}\right)}{\left(1 - \exp\left(-\frac{\text{PET}}{P}\right)\right)}\right]^{0.5} 
|                               | w                                                                           | Trees – 2.0, Plants – 0.5 |
| Zhang et al. (2001)           | \[
|                               | \frac{1 + w \cdot \text{PET}}{1 + w \cdot \text{PET} + \left(\frac{\text{PET}}{P}\right)^w} 
|                               | w                                                                           | Trees – 2.0, Plants – 0.5 |
| Turc (1954), Mezentsev (1955),| \[
|                               | \frac{1}{1 + \left(P \text{PET}\right)^n} 
|                               | n                                                                           | Field – 2.6, River basins – 1.8 |
| Choudhury (1999), Yang et al. (2008) | \[
|                               | \left[1 + \left(P \text{PET}\right)^m\right]^{\frac{1}{m}} - \left(P \text{PET}\right)^{-1} 
|                               | m                                                                           | Forest – 2.83, Shrub – 2.33, Grassland or cropland – 2.28, Mixed land – 2.12 |
| Wang and Tang (2014)          | \[
|                               | \frac{1 + \text{PET} - \sqrt{\frac{(1 + \text{PET})^2 - 4\varepsilon (2 - \varepsilon) P \text{PET}}{2\varepsilon (2 - \varepsilon)}}}{\varepsilon} 
|                               | \varepsilon                                                                 | 0.55–0.58 |
|                               | \left[1 + \left(P \text{PET}\right)^{-m}\right]^{\frac{2}{m}} - \left(P \text{PET}\right)^{-1} 
|                               | m                                                                           | Forest – 2.83, Shrub – 2.33, Grassland or cropland – 2.28, Mixed land – 2.12 |

When Pw tends to infinity, the runoff approaches the difference between precipitation and potential evapotranspiration. In this scenario, all precipitation remains in the watershed, and all available water is lost through evapotranspiration. The duration of water residence equals the time for converting all precipitation to evapotranspiration. However, in natural watersheds, it may be difficult to observe Pw approaching infinity since it is nearly impossible for all precipitation to be retained in the watershed. The natural watersheds with a high Pw value may be “non-conservative” (i.e., precipitation is not the sum of streamflow and evapotranspiration), as a portion of the water that remains in the watershed may not be solely from precipitation but may include groundwater flow and other difficult to measure flows. As a result, it may be challenging to accurately estimate the water balance, especially in regions with complex hydrological systems (De Lavenne and Andréassian, 2018; Goswami and O’Connor, 2010). As a precautionary measure, this study sets an empirical upper limit of 10 for Pw to ensure that the watersheds in question remain conservative.

3 Data

3.1 Hydrological data

Hydrological data for modeling, including runoff and corresponding precipitation data, were collected from published literature (726 samples listed in Supplement 1, Fig. 1). Potential evapotranspiration data were downloaded from version 4.05 of the CRU TS (Climatic Research Unit gridded Time Series) climate dataset (https://doi.org/10.6084/m9.figshare.11980500.v1), which is produced by the CRU at the University of East Anglia. For consistency, we used potential evapotranspiration values extracted from the CRU TS dataset of all watersheds listed in Supplement 1, even for studies with potential evapotranspiration values reported. The potential evapotranspiration values were extracted based on the coordinate points of watersheds. Using collected and extracted the annual average runoff, precipitation and potential evapotranspiration data for the observation period, we calculated the annual water yield coefficient (R/P) and aridity index (P/PET) for each sample. Then, we derive the annual average Pw value of each sample for the corresponding period according to Eq. (1).

Observed river discharge data for validation were obtained from the Global Runoff Data Centre (GRDC, https://www.bafg.de/GRDC/EN/02_srvcs/21_tmsrs/riverdischarge_node.html, last access: 16 April 2021). Only the GRDC stations meeting the following criteria were selected for further analysis: (1) The sites with continuous time-series runoff observations during the period 2000–2016 and corresponding surface soil moisture (SM), fractional vegetation cover (FVC) and seasonal index (SI) data were also available during such a period. (2) The drainage area
3.2 Watershed characteristic-related data

The watershed characteristic-related factors mainly include SM (0–10 cm underground), FVC and SI of Walsh and Lawler (1981). For the collected watersheds from published literature without boundary files, these three datasets were extracted from grid data according to the coordinate points of these watersheds. For the GRDC watersheds, records of these three fields were extracted from grid data based on the boundary files provided by GRDC Watershed Boundaries (2011). The sources of datasets are summarized in Table 2.
Table 2. Data sources for watershed characteristic factors.

<table>
<thead>
<tr>
<th>Watershed characteristic factors</th>
<th>Data source/version</th>
<th>Units</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surface soil moisture (0–10 cm underground, SM)</td>
<td>GLDAS Noah Land Surface Model L4</td>
<td>mm</td>
<td>Rodell et al. (2004)</td>
</tr>
<tr>
<td>Fractional vegetation cover (FVC)</td>
<td>GLASS FVC V4</td>
<td>m² m⁻²</td>
<td>Liang et al. (2021)</td>
</tr>
<tr>
<td>Seasonal index (SI)</td>
<td>CRU TS dataset version 4.03, global maps of seasonality indices</td>
<td>dimensionless</td>
<td>Walsh and Lawler (1981); Feng (2019)</td>
</tr>
</tbody>
</table>

Table 3. Classification of watersheds.

<table>
<thead>
<tr>
<th>Soil moisture classifier</th>
<th>Water soil regime</th>
<th>Seasonality index classifier</th>
<th>Seasonality precipitation regime</th>
<th>Fractional vegetation cover classifier</th>
<th>Vegetation cover regime</th>
<th>Name of the group</th>
</tr>
</thead>
<tbody>
<tr>
<td>SM ≤ 20</td>
<td>Dry soil</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>IN_D</td>
</tr>
<tr>
<td>SM &gt; 20</td>
<td>Wet soil</td>
<td>SI ≤ 0.4</td>
<td>Seasonless</td>
<td>FVC ≤ 0.2</td>
<td>Low density</td>
<td>IN_WP</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.4 &lt; SI ≤ 0.8</td>
<td>Marked seasonality</td>
<td>0.2 &lt; FVC ≤ 0.5</td>
<td>Middle density</td>
<td>IN_WM_S</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>FVC &gt; 0.5</td>
<td>High density</td>
<td>IN_WM_M</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SI &gt; 0.8</td>
<td>Extreme seasonality</td>
<td></td>
<td></td>
<td>IN_WE</td>
</tr>
</tbody>
</table>

in Supplement 2. Based on the criteria used by Walsh and Lawler (1981), we divided the SI into three parts (SI ≤ 0.4, 0.4 < SI ≤ 0.8, SI > 0.8) to represent three hydroclimatic seasonality (precipitation spread throughout the year, marked seasonality with a short drier season, extreme seasonality with a long drier season). Finally, six hydrologically similar groups were classified (Table 3).

4.2 Setup of proposed Pw simulation model (PwM)

4.2.1 PwM with the classification of hydrologically similar groups

We performed regression analysis between the Pw and watershed characteristic variables to determine the input variables of the PwM. The variables whose $R^2$ of the regression model was greater than 0.1 were selected as input variables. We used a polynomial as the basic model form. Each term of the polynomial depends on the regression model of the corresponding variable and the Pw. For each hydrological group, the Pw value is modeled as the following function:

$$Pw = \sum Coef_n \times f(Var_n),$$  \hspace{1cm} (2)

where $Pw$ represents the value of Pw, $Var_n$ represents the input variable that passes the regression test, $f$ corresponds to the function derived from the regression of $Pw$ on $Var_n$, and $Coef_n$ represents the empirical coefficient fitted by multiple non-linear regression (MNR).

4.2.2 PwM without classification of hydrologically similar groups

For comparison, we estimated Pw without the hydrologically similar groups, defined as non_PwM. The non_PwM was defined as follows:

$$non_Pw = a_1 \times SM^2 + a_2 \times SM + b_1 \times FVC^2 + b_2 \times FVC,$$  \hspace{1cm} (3)

where non_Pw is the annual value of Pw simulated by non_PwM; SM is the annual average value of surface soil moisture (0–10 cm underground); FVC is the annual average value of fractional vegetation cover; and $a_1$, $a_2$, $b_1$, and $b_2$ represent the empirical coefficients fitted by the least-squares method.

4.3 Model validation

4.3.1 Performance metrics

Three performance metrics were used to assess the accuracy of PwM. The variable $N$ is the number of observations, $i$ is the $i$th value to be simulated, and $y_s$ and $y_o$ are the simulated and observed series, respectively.

The relative bias (RelBIAS) represents systematic errors. A positive value indicates a general overestimation, while a negative one indicates an underestimation. The perfect agreement is achieved when RelBIAS equals zero. RelBIAS is defined as

$$RelBIAS = \frac{mean(y_s - y_o)}{mean(y_o)}.$$  \hspace{1cm} (4)
Figure 2. Regression between Pw in Fu’s formula and (a) SM (SM \( \leq 20 \) mm), (b) SM (SM >20 mm), (c) FVC (IN\(_D\)), (d) FVC (IN\(_{WP}\)), (e) FVC (IN\(_{WMS}\)), (f) FVC (IN\(_{WMM}\)), (g) FVC (IN\(_{WML}\)), and (h) FVC (IN\(_{WE}\)). Symbol shapes indicate SM (dots) and FVC (squares).

The coefficient of determination (\( R^2 \)) assesses the linear relationship between the simulated and observed time series data and is defined as

\[
R^2 = \frac{\sum_{i=1}^{N} (y_i^s - \bar{y}_o) (y_i^o - \bar{y}_o)}{\left( \sum_{i=1}^{N} (y_i^o - \bar{y}_o)^2 \right)^{0.5}}.
\] (5)

The Nash–Sutcliffe efficiency (NSE) (Nash and Sutcliffe, 1970), a goodness-of-fit index, is usually used to assess the accuracy of the model. When NSE = 1, the model predictions perfectly match the observed data. A value higher than 0 indicates that the modeled mean is a good predictor compared to the observed value. It is defined as

\[
\text{NSE} = 1 - \frac{\sum_{i=1}^{N} (y_i^s - y_i^o)^2}{\sum_{i=1}^{N} (y_i^o - \bar{y}_o)^2}.
\] (6)

4.3.2 Cross-validations using the bootstrap sampling method

We used cross-validation to test the stability of the proposed PwM using the bootstrap sampling method. The collected public data were split into two parts, one for model training and the other for model validation. A subset of 60% of the data was randomly selected using the bootstrap sampling method for training PwM. The remaining 40% of the data was used to evaluate the model performance using the validation metrics in Sect. 4.3.1. For each metric, the variable \( N \) is the number of test sets, \( i \) is the \( i \)th value to be simulated by the trained PwM, and \( y_s \) and \( y_o \) are the simulated and observed series of test sets, respectively. The process was repeated randomly 10,000 times. We documented the cross-validation result of each bootstrapping and showed them in the violin plot (Fig. 3).

4.3.3 Validations of GRDC time-series runoff reconstruction results

To further assess the model performance, we applied the proposed PwM to Fu’s model to reconstruct the time-series runoff data of GRDC from 2000 to 2016. Finally, the time-series runoff data from 545 GRDC stations, which were selected by Sect. 3.1, were used to evaluate the model performance using the validation metrics in Sect. 4.3.1. For each metric, the terms \( y_s \) and \( y_o \) represent the simulated and observed time-series runoff data, respectively.

5 Results and discussion

5.1 The new proposed model for estimating Pw in Fu’s formula

The regressions between Pw in Fu’s formula and watershed characteristic variables collected from globally published datasets are shown in Fig. 2.

As shown in Fig. 2a, b, the relationship between Pw and SM conforms to a power function, consistent with prior findings reported by Chen and Sivapalan (2020). The important finding here is that there is a critical soil moisture threshold at 20 mm that separates watersheds with two different water balances. In watersheds characterized by arid conditions (SM \( \leq 20 \) mm), as shown in Fig. 2a, the Pw values have an upward trend as SM values increase. On the other side, in watersheds characterized by humid conditions (SM >20 mm), as shown in Fig. 2b, the Pw values exhibit a decreasing trend as SM values increase. This is likely because transpiration usually increases as soil water increases in relatively dry conditions (Jiao et al., 2019; Bierhuizen, 1958; Wang et al., 2012; Yao...
et al., 2016; Schwarzel et al., 2020). However, once the soil moisture exceeds the threshold (20 mm in this study), the acceleration of transpiration from soil moisture slows down quickly (Havranek and Benecke, 1978; Verhoef and Egea, 2014; Metselaar and De Jong Van Lier, 2007). These findings are very in line with previous studies (Havranek and Benecke, 1978; Jiao et al., 2019; Cavanaugh et al., 2011; Ducharne et al., 1998), although the threshold of soil moisture varies slightly in these studies (e.g., 0.25 m³ m⁻³ in Ducharne et al., 1998, 0.10 m³ m⁻³ in Cavanaugh et al., 2011 and 0.20 m³ m⁻³ in Jiao et al., 2019).

As shown in Fig. 2c–h, the FVC is linearly correlated with the Pw values of watersheds in most hydrologically similar groups but differs greatly between different groups. In dry watersheds (INWP), the relationship between Pw and FVC followed a positive linear function (Fig. 2c). This finding is consistent with the major view that vegetation transpiration increases (reflected by the increased Pw) with increasing vegetation cover in regions with insufficient soil moisture (Wang et al., 2012; Yao et al., 2016; Schwarzel et al., 2020). For those small and wet watersheds, vegetation-related factors are considered to be weakly correlated with Pw (Liu et al., 2021; Padrón et al., 2017; Yang et al., 2014). However, our study reveals a positive linear correlation between Pw and FVC in the INWMS (Fig. 2e) and INWE groups (Fig. 2h), whereas a negative linear correlation is observed in the INWM (Fig. 2f) and INWML groups (Fig. 2g). Only in the INWP group is the relationship between Pw and FVC not significant. These results indicate that the relationship between Pw and FVC may be stronger than what was previously believed, and this relationship varies across different groups characterized by specific combinations of FVC and SI. This confirms that climate, soil moisture, and vegetation cover are not independent factors affecting the water balance (Gan et al., 2021; Yang et al., 2009). Coupling vegetation with other catchment properties resulted in greater Pw variations (Gan et al., 2021).

Based on the results of the regression analysis illustrated in Fig. 2, the proposed PwM employs SM and FVC as input variables (i.e., Var_n) for all groups, except for the INWP group, for which FVC was not chosen. The formula in PwM for calculating the Pw is modeled as a sum of a power function of SM and a linear function of FVC, given by Eq. (7):

\[
Pw = \begin{cases} 
0.91 \times SM^{0.38} + 1.48 \times FVC \\
(\text{IN}_D, SM \leq 20) \\
28.72 \times SM^{-0.76} \\
(\text{IN}_W, SM > 20, SI \leq 0.4) \\
39.03 \times SM^{-0.96} + 11.82 \times FVC \\
(\text{IN}_WMS, SM > 20, 0.4 < SI \leq 0.8, FVC \leq 0.2) \\
33.76 \times SM^{-0.71} - 4.221 \times FVC \\
(\text{IN}_WML, SM > 20, 0.4 < SI \leq 0.8, FVC > 0.5) \\
3078 \times SM^{-2.43} + 3.53 \times FVC \\
(\text{IN}_WE, SM > 20, SI > 0.8), 
\end{cases}
\]

where Pw is the annual value of Pw, SM is the annual average value of soil moisture (0–10 cm underground), and FVC is the annual average value of fractional vegetation cover.

5.2 Cross-validations based on data collected from globally published literature

The performances of PwM and non_PwM were cross-validated based on the data collected from globally published literature using the bootstrap sampling method (Fig. 3). On average, the ensemble RelBIAS of Pw simulated by PwM is slightly negative (Fig. 3a), indicating a weak tendency to underestimate the values of Pw but with a maximum relative bias less than 0.1. The interquartile range of R² for PwM is from 0.35 to 0.40, with a median of 0.37. The scores of R² are higher than 0.3 in more than 95 % of the bootstrap sampling events. The NSE skill scores show that in most bootstrap samplings, the estimation-error-estimated variance for PwM is less than the variance of the observations (NSE > 0), with an interquartile range from 0.33 to 0.39. In comparison, the maximum relative bias of the Pw simulated by the
Figure 4. Cross-validation results of PwM for (a) IND, (b) INWP, (c) INWMS, (d) INWMM, (e) INWML, and (f) INWE.

Figure 5. Simulated R/P using PwM in comparison with the observations collected from published literature. (a) Scatter plots between R/P (yellow: simulation; green: observations) and P/PET. (b) Difference between simulated R/P from PmM and observations from the published datasets.

non_PwM is 0.12, the median of $R^2$ is 0.13, and the median of NSE is 0.13. Overall, cross-validations show that the performance of the PwM with the hydrologically similar groups is better and more stable than that of the non_PwM.

Grouping watersheds based on their hydrological similarities ensures that watersheds within the same category exhibit similar behaviors in settings with comparable climate, soil, and vegetation characteristics (Kanishka and Eldho, 2017; Sinha et al., 2019). The model developed based on the principle of hydrologically similar groups considers the unique hydrological characteristics of different watersheds and can more accurately simulate the hydrological response in complex watershed systems (Santra et al., 2011; Jin et al., 2017; Kouwen et al., 1993; Gao et al., 2018; Kanishka and Eldho, 2017). As a comparison, in the non_PwM, all watersheds were lumped into a single category and showed a similar hydrological response to changes in watershed characteristics. That non_PwM, as the similar model used in previous studies (Zhang et al., 2018; Liang et al., 2015; Xu et al., 2013), may overlook and oversimplify the intricate interplay between climate, watershed characteristics, and hydrology, thereby potentially resulting in less precise predictions of Pw across diverse watersheds.

The skill scores of cross-validations for the six groups are shown in Fig. 4. Though its overall RelBIAS is negative, PwM tends to overestimate values of Pw in the INWP group (the median of RelBIAS is positive). The INWMS group scores highest in $R^2$, with a median of 0.73, while the INWP group scores the lowest, with a median of 0.16. The grouped NSE scores show more uncertainty than the overall, especially in the INWMS: the lower adjacent value (LAV) larger than zero indicates more skill than the mean of observations; however, the outliers are far below zero. The low NSE value may be due to the low number of watersheds sampled in this interval, which increased the inconclusive results.
Figure 5 shows the simulated $R/P$ by PwM in comparison to site observations. The $R^2$ between the observed and the simulated values is 0.63 (Fig. 5a). The model performs well in humid regions with $P/PET \geq 1$ in southeast America, Europe, central China, and southeast Australia. However, PwM likely underestimated the runoff in the arid ($P/PET < 0.2$) and semi-arid regions ($0.2 \leq P/PET < 0.5$), mainly in western America and northwest China (Fig. 5b).

5.3 Validations of reconstructing the time-series GRDC runoff

For the selected 545 GRDC watersheds, the annual runoff estimated by PwM ranges from 229.84 to 320.34 mm, which is slightly lower than the observed range of GRDC (265.82–345.50 mm yr$^{-1}$) (Fig. 6a). Overall, the temporal evolution of runoff is captured well in the period 2000–2010. However, since 2011, the consistency between reconstructed runoff and
GRDC runoff has decreased, and the reconstruction results have consistently been lower than the GRDC observations. The scatter plot between simulated and observed \( R/P \) also shows a slight underestimation of reconstructed global long-term mean runoff (Fig. 6b). The spatial patterns of long-term mean runoff reconstruction are shown in Fig. 6c–f. The estimated time-series runoff shows lower values in the west of the United States and south of Africa and shows higher values in the northeastern United States and the European Mediterranean area, in comparison with the GRDC time series.

Figure 7 displays the skill scores of the reconstructed runoff by PwM in comparison with the GRDC ensemble from 2000–2016. It can be seen that, generally, the result of reconstruction by PwM is satisfactory, as indicated by the RelBIAS close to 0. The underestimation of runoff mainly occurs in the high mountains of the western United States (Fig. 7a), where the runoff is much smaller. Humid regions such as the northeastern United States and the European Mediterranean area have quite high \( R^2 \) values, while lower values are observed in the semi-arid (0.2 \( \leq P/PET < 0.5 \)) and the dry sub-humid (0.5 \( \leq P/PET < 0.65 \)) regions, which are mainly located in the western and midwestern United States (Fig. 7e–h). There are low NSE scores in the watersheds where runoff is unusually underestimated or overestimated (Fig. 7i–l), especially in the western United States.

We classified the GRDC data into nine geographic regions (Fig. 1) and further evaluated the performance of PwM in each sub-region individually. In general, the simulated time-series runoff is consistent with the time-series observations (Figs. 8, 9), except in the western United States, where runoff was consistently underestimated (Fig. 8a). Spatially, there is an underestimation of runoff in sub-regions like the western United States (Fig. 8a) and high latitudes in North America (Fig. 8f). The runoff underestimation is more severe in the arid areas of the western United States (Fig. 9a) than in the relatively wet areas of northwest North America (Fig. 9f). The reconstructed time-series runoff levels in the Milk River watershed (GRDC station number: 4220501) and Near Lethbridge watershed (GRDC station number: 4213111) both show an underestimation of annual runoff in arid areas. The Milk River and Near Lethbridge are two adjacent watersheds with similar drainage areas located on the border of the United States and Canada. However, the underestimation is more serious in the Milk River watershed (RelBIAS = -0.32, annual mean \( P/PET = 0.52 \)) than in the Near Lethbridge watershed (RelBIAS = -0.27, annual mean \( P/PET = 0.55 \)). Interestingly, the spatial pattern of runoff underestimation almost coincides with that of the glaciers. Therefore, we considered that glacial meltwater might be the probable cause of runoff underestimation in glacier-covered areas (Li et al., 2021), where glacial snowmelt plays a more important role as a water input in arid regions than in wet ones. Therefore, the underestimation of runoff in the western United States is greater than in northwest North America. Temporally, the runoff was mostly underestimated by

Figure 8. Observed time-series runoff versus reconstructed time-series runoff. Nine geographic sub-regions were in Fig. 1: North America (a west, b southwest, c midwest, d northeast, e southeast, f except for the USA), (g) South America, (h) Africa, and (i) Europe.
PwM in the year 2011, when the world experienced abnormally high temperatures (Frölicher et al., 2018; NOAA-CEI, 2011), and glacier melting was thus accelerated and increased runoff (Du et al., 2022; J. Liu et al., 2022).

In this paper, we selected Fu’s new equation and developed a universal framework for estimating Pw. Our results show that, to a large extent, the Pw in Budyko equation can be well estimated by the PwM using only soil moisture and fractional vegetation cover parameters. This indicates that soil moisture and fractional vegetation cover strongly control the water balance of watersheds (Gan et al., 2021; Chen and Sivapalan, 2020; Yang et al., 2009; Wang et al., 2021). The better performance of PwM than non_PwM supports our hypothesis that watersheds with similar climatic, hydrologic, and watershed-related characteristics have consistent controlling factors of Pw in the Budyko framework and suggest that the classification of watersheds can reduce uncertainty and improve the accuracy of Pw and runoff predictions.

6 Conclusions

This study developed a new framework for estimating the Pw in the Budyko framework for watersheds in similar environments based on the principle of hydrologically similar groups. The proposed method not only represented runoff observations in 366 watersheds from published literature but also reconstructed the time-series runoff in 545 GRDC stations. The findings indicated that Pw is closely related to SM and FVC, and the relationship varies across specific hydrologically similar groups. However, due to the complexity of hydrological processes, the new framework could not fully account for the impacts of all other factors, which might result in an underestimation of runoff in regions with glaciers or under climates with temperature anomalies. Overall, our findings lay a sound basis for estimating Pw in the Budyko framework, providing references for calibrating hydrological models, and will be helpful in improving global runoff estimations.

Code availability. The pieces of code that were used for all analyses are available from the authors upon request.
Data availability. All data used in this study are publicly available. Potential evapotranspiration data are available from CRU TS (https://doi.org/10.6084/m9.figshare.11980500.v1), Scientific Data Curation Team, 2020), precipitation data used to model validation are available from GPCC (https://doi.org/10.5676/DWD_GPCC/FP_M_V2018_050), Schneider et al., 2018), observed river discharge data are available from GRDC (https://www.bafg.de/GRDC/EN/02_srrcs/21_tmrs/riverdischarge_node.html, GRDC, 2020), SM data are available from GLDAS (https://doi.org/10.5067/SXAVCZFAQLNO, Beaudoin and Rodell, 2020), FVC data are available from GLASS (http://www.glass.umd.edu/05D/FVCI/, National Earth System Science Data Center, 2020), and SI data are available from HydroShare (http://www.hydroshare.org/resource/ff287e90c9e947a78e351c8d07d9d3f3, Feng, 2019).

Supplement. The supplement related to this article is available online at: https://doi.org/10.5194/hess-27-1929-2023-supplement.

Author contributions. XC designed the study and proposed the scientific hypothesis. YC implemented the experiments, conducted the analysis, and wrote the paper. MX helped with data collection and checked the technical adequacy of the experiments. CY and WZ helped with data processing. WPY provided guidance on the SI. CY, CJ, WTY, and WPY reviewed and edited the manuscript. XC oversaw the study and conducted manuscript revision as a mentor.

Competing interests. The contact author has declared that none of the authors has any competing interests.

Disclaimer. Publisher’s note: Copernicus Publications remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Acknowledgements. We would like to express our gratitude to Roger Moussa for handling the manuscript, as well as the two anonymous referees for their constructive comments which have greatly contributed to improving the quality of our manuscript.

Financial support. This study was financed by the Guangdong Major Project of Basic and Applied Basic Research (grant no. 2020B0301030004), the National Natural Science Foundation of China (grant nos. 31971458 and 41971275), the Innovation Group Project of Southern Marine Science and Engineering Guangdong Laboratory (Zhuha; grant no. 311021009), and the Special High-level Plan Project of Guangdong Province (grant no. 2016TQ03Z354).

Review statement. This paper was edited by Roger Moussa and reviewed by two anonymous referees.

References


Feng, X.: Global maps of seasonality indices, HydroShare [data set], http://www.hydroshare.org/resource/ff287e90c9e947a78e351c8d07d9d3f3 (last access: 11 April 2022), 2019.


GRDC: GRDC discharge data, GRDC [data set], https://www.bafg.de/GRDC/EN/02_srvcs/21_tmsrs/riverdischarge_node.html (last access: 16 April 2021), 2021.


