Using the Budyko framework for calibrating a global hydrological model in ungauged catchments of the world

Peter Burek, Peter Greve, Yusuke Satoh, and Yoshihide Wada
IIASA, Water, Laxenburg, Austria (burek@iiasa.ac.at)
1 Problem

Calibration without discharge data

Global Hydrological Model
CWATM
http://www.iiasa.ac.at/cwatm
https://cwatm.github.io/

1: River: Rhine, Station: Lobith, No of runs: 1296

(a) Streamflow time series for calibration period

KGE = 0.84, NSE = 0.66, R = 0.85, B = 3.71%

Dense network and actual data
Sparse network of non actual data

Global Runoff Data Centre
2 Idea

Using the empirical relation

Budyko function

for calibration

Hypotheses:
Budyko calibration results will be not as good fitting simulated to the observed discharge as if it is calibrated for discharge itself, but it will be an improvement against an unfitted a priori parameter run.

Advantage:
Precipitation, and evaporation is available everywhere.

Mikhail Budyko

Budyko function
(Budyko, 1958, 1974)
3 Method Calibration

Instead: Finding a parameter set which represents discharge data

Finding a parameter set which represents the Budyko function

Budyko curve: \( \frac{E}{P} = 1 + \frac{E_{pot}}{P} - (1 + \left( \frac{E_{pot}}{P} \right)^{\omega})^{1/\omega} \) with: \( \omega = 2.6 \)
“Budyko” Calibration For River Rhine

Objective functions:
KGE: Kling Gupta Efficiency
NS: Nash-Sutcliffe Efficiency

<table>
<thead>
<tr>
<th></th>
<th>Obs</th>
<th>Sim0</th>
</tr>
</thead>
<tbody>
<tr>
<td>KGE</td>
<td>0.548</td>
<td></td>
</tr>
<tr>
<td>NS</td>
<td>0.643</td>
<td></td>
</tr>
</tbody>
</table>
"Budyko" Calibration For River Rhine

Objective functions:

<table>
<thead>
<tr>
<th></th>
<th>Obs</th>
<th>Sim0</th>
<th>SimBud</th>
<th>SimDis</th>
</tr>
</thead>
<tbody>
<tr>
<td>KGE</td>
<td>0.548</td>
<td>0.809</td>
<td>0.917</td>
<td></td>
</tr>
<tr>
<td>NS</td>
<td>0.643</td>
<td>0.633</td>
<td>0.836</td>
<td></td>
</tr>
</tbody>
</table>

KGE: Kling Gupta Efficiency
NS: Nash-Sutcliffe Efficiency
Global discharge data

The monitoring network of discharge data is sparse in large part of the globe, and there is **no** mechanism in place to collect and distribute river discharge data **globally** on a real-time base.

Global Runoff Data Centre (GRDC) (2017)

Do et al. (2018): see also  
EGU2018-5994: Wed, 11 Apr, 15:30–15:45, Room 2.31  

Different ways to overcome the problem of having no discharge time series

- Regionalization of discharge data
e.g. Barbarossa et al. 2018

- Regionalization of model parameter
e.g. Beck et al. 2016

- Calibration with discharge from satellite derived data
e.g. Revilla-Romero et al. (2015)

Community Water Model (CWATM)

Development of a community driven global water model by WAT Program, IIASA

Model design

- CWATM represents one of the new key elements of IIASA’s Water program to assess water supply, water demand and environmental needs at global and regional level
- The hydrologic model is open source and flexible to link in different aspects of the water energy food nexus

Global discharge demo

- Vision
  - Our vision for the short to medium term work is to introduce water quality and to consider qualitative and quantitative measures of transboundary river and groundwater governance into an integrated modelling framework.

Contact

- www.iiasa.ac.at/cwatm
- wfas.info@iiasa.ac.at
Calibration of river discharge

River: Rhine Station: Lobith

River: Klamath, Station: USGS 11523000 - Orleans, CA

River: Murray River station: Wakool Junction

**Calibration:**
- Daily run of 12 to 20 years
- Compared to daily or monthly observed discharge
- Objective function: KGE’
  - KGE’: modified Kling-Gupta efficiency
  - NSE: Nash-Sutcliffe Efficiency
  - R2: Correlation coefficient
  - B: Bias
Calibration is using an evolutionary computation framework in Python called DEAP (Fortin et al., 2012). DEAP implemented the evolutionary algorithm NSGA-II (Deb et al., 2002) which is used here as single objective optimization.

Evolution of parameter space

Parameter space for 8 parameter
Discharge:
Daily (or monthly) pairs of observed and simulated discharge at gauging stations

Objective function:
Modified version of the Kling-Gupta Efficiency (Kling et al., 2012),

\[ KGE' = 1 - \sqrt{(r - 1)^2 + (\beta - 1)^2 + (\gamma - 1)^2} \]

where: \( \beta = \frac{\mu_s}{\mu_o} \) and \( \gamma = \frac{CV_s}{CV_o} = \frac{\sigma_s/\mu_s}{\sigma_o/\mu_o} \)

Where:
r as the correlation coefficient between simulated and observed discharge (dimensionless), \( \beta \) as the bias ratio (dimensionless) and \( \gamma \) as the variability ratio.
CV is the coefficient of variation, \( \mu \) is the mean streamflow \( [m^3 s^{-1}] \) and \( \sigma \) is the standard deviation of the streamflow \( [m^3 s^{-1}] \). KGE’, \( r \), \( \beta \) and \( \gamma \) have their optimum at unity.

For discharge calibration 12 parameters are calibrated. For each important hydrological process - snow, evaporation, soil, groundwater, routing, lakes up to 3 parameters are used.

Because the Budyko curve looks at runoff generation (and evaporation) at grid cell level the runoff concentration and the routing processes are not sensitive to the objective function of the Budyko calibration. Therefore only 5 parameters are calibrated.

For each grid cell the sum of daily precipitation ($P$), potential evaporation ($ETP$) and actual evapotranspiration ($ETA$) is calculated. From these three sums the coordinate in the “Budyko space” are calculated:

$$x = \frac{ETP}{P}; \quad y = \frac{ETA}{P}$$

Depending on the period of calibration the sum is calculated for 10 to 15 years. The “Budyko space” spanned by $x,y$ for each grid cell should be close to the Budyko curve:

$$y = 1 + x - (1 + x^\omega)^{1/\omega} \quad \text{with fixed } \omega = 2.6.$$  

Here the distance of Kolmogorov-Smirnov (maximum distance of a point to the function) is used as objective function and the calibration algorithm is minimizing this distance.
Improvements

- Using another test than KS for Budyko e.g. min distance of all points to a function, or other statistical test e.g. Anderson-Darling

- A fixed $\omega = 2.6$ is used for all station. Could be variable depending on the climate zone.

- At the moment only the water balance of a grid cell without incoming discharge and evaporation from rivers and lakes are estimated. Precipitation = Runoff + Evaporation
  The storage term is not used:
  Precipitation = Runoff + Evaporation + $\Delta S$
Rhine (Lobith, Germany)
The “Budyko” run gives a good improve compared to the a priori parameter run (Sim0).
The a priori parameter run is overestimating observed discharge by far (84%) while the Budyko run is even underestimating observed discharge. Overall Budyko cal. is a major improvement.
Upper Nile – Lake Vitoria (Jinja, Uganda)

The a priori parameter run is overestimating (36%) observed discharge. Discharge calibrated discharge fit very well (KGE = 0.92, NSE = 0.85)

Budyko cal. is half way from uncalibrated to discharge calibrated. Overall it is an improvement

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Obs</th>
<th>Sim0</th>
<th>SimBudyko</th>
<th>SimDis</th>
<th>SimDisBudyko</th>
</tr>
</thead>
<tbody>
<tr>
<td>KGE</td>
<td>0.571</td>
<td>0.772</td>
<td>0.922</td>
<td>0.846</td>
<td>0.798</td>
</tr>
<tr>
<td>NS</td>
<td>-1.877</td>
<td>0.002</td>
<td>-1.681</td>
<td>0.180</td>
<td>0.798</td>
</tr>
<tr>
<td>NSlog</td>
<td>-0.852</td>
<td>0.970</td>
<td>0.940</td>
<td>0.990</td>
<td>-2.75%</td>
</tr>
<tr>
<td>BCE</td>
<td>0.354</td>
<td>0.354</td>
<td>0.354</td>
<td>0.354</td>
<td>0.354</td>
</tr>
<tr>
<td>RMSE</td>
<td>377</td>
<td>222</td>
<td>87</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MAE</td>
<td>354</td>
<td>354</td>
<td>354</td>
<td>354</td>
<td>72</td>
</tr>
<tr>
<td>Mins</td>
<td>992</td>
<td>1346</td>
<td>1165</td>
<td>965</td>
<td></td>
</tr>
<tr>
<td>Mins</td>
<td>474</td>
<td>820</td>
<td>594</td>
<td>529</td>
<td></td>
</tr>
<tr>
<td>HS</td>
<td>652</td>
<td>956</td>
<td>763</td>
<td>649</td>
<td></td>
</tr>
<tr>
<td>HSlog</td>
<td>977</td>
<td>1212</td>
<td>1128</td>
<td>914</td>
<td></td>
</tr>
<tr>
<td>HS95%</td>
<td>1384</td>
<td>1774</td>
<td>1774</td>
<td>1380</td>
<td></td>
</tr>
<tr>
<td>HS99%</td>
<td>1688</td>
<td>1785</td>
<td>1774</td>
<td>1595</td>
<td></td>
</tr>
<tr>
<td>HSmax</td>
<td>1661</td>
<td>1801</td>
<td>1809</td>
<td>1646</td>
<td></td>
</tr>
</tbody>
</table>
The a priori parameter run is overestimating observed discharge. Budyko cal. is a reasonable improvement towards discharge calibration.
Danube - Kienstock, Austria  
Zimnicea, Romania

Danube - Kienstock, Austria  
Catchment area: 96,000km²

Zimnicea, Romania  
Catchment area: 648,400km²

River: Danube Station: Kienstock, AT

(a) Streamflow time series for calibration period

Streamflow [m³/s] with:
KGE = 0.60, NSE = 0.60, R = 0.78, B = 2.44 %

(b) Budyko plot

Budyko curve: \( \frac{E}{P} = 1 + \frac{E_{pot}}{P} \times \left(1 + \left(\frac{E_{pot}}{P}\right)^{\omega}\right)^{-1/\omega} \) with: \( \omega = 2.6 \)

(c) Budyko obj. function evolution

<table>
<thead>
<tr>
<th>Obs</th>
<th>Sim0</th>
<th>Sim/P0</th>
<th>Sim/P0c</th>
</tr>
</thead>
<tbody>
<tr>
<td>KGE</td>
<td>0.605</td>
<td>0.605</td>
<td>0.605</td>
</tr>
<tr>
<td>NS</td>
<td>0.802</td>
<td>0.802</td>
<td>0.802</td>
</tr>
<tr>
<td>NSE</td>
<td>0.800</td>
<td>0.800</td>
<td>0.800</td>
</tr>
<tr>
<td>R</td>
<td>1.072</td>
<td>1.072</td>
<td>1.072</td>
</tr>
<tr>
<td>Bias</td>
<td>0.05 %</td>
<td>0.05 %</td>
<td>0.05 %</td>
</tr>
<tr>
<td>RMSE</td>
<td>423</td>
<td>535</td>
<td>529</td>
</tr>
<tr>
<td>MAE</td>
<td>411</td>
<td>372</td>
<td>350</td>
</tr>
</tbody>
</table>

(d) Monthly Q climatology cal. period

Streamflow [m³/s] with:
95 % 1994
90 % 1996
85 % 1998
75 % 1999
65 % 2001
50 % 2002
30 % 2004
95 % 2005
Max 10.073

River: Danube Station: Zimnicea, RO

(a) Streamflow time series for calibration period

Streamflow [m³/s] with:
KGE = 0.56, NSE = 0.40, R = 0.85, B = 21.73 %

(b) Budyko plot

Budyko curve: \( \frac{E}{P} = 1 + \frac{E_{pot}}{P} \times \left(1 + \left(\frac{E_{pot}}{P}\right)^{\omega}\right)^{-1/\omega} \) with: \( \omega = 2.6 \)

(c) Budyko obj. function evolution

<table>
<thead>
<tr>
<th>Obs</th>
<th>Sim0</th>
<th>Sim/P0</th>
<th>Sim/P0c</th>
</tr>
</thead>
<tbody>
<tr>
<td>KGE</td>
<td>0.524</td>
<td>0.565</td>
<td>0.537</td>
</tr>
<tr>
<td>NS</td>
<td>0.404</td>
<td>0.404</td>
<td>0.404</td>
</tr>
<tr>
<td>NSE</td>
<td>0.403</td>
<td>0.403</td>
<td>0.403</td>
</tr>
<tr>
<td>R</td>
<td>2.375</td>
<td>2.375</td>
<td>2.375</td>
</tr>
<tr>
<td>Bias</td>
<td>27.12 %</td>
<td>27.73 %</td>
<td>27.73 %</td>
</tr>
<tr>
<td>RMSE</td>
<td>213</td>
<td>309</td>
<td>148</td>
</tr>
<tr>
<td>MAE</td>
<td>1819</td>
<td>1363</td>
<td>1070</td>
</tr>
</tbody>
</table>

(d) Monthly Q climatology cal. period

Streamflow [m³/s] with:
95 % 2003
90 % 2007
85 % 2009
75 % 2011
65 % 2013
50 % 2015
30 % 2017
95 % 2019
Max 31.840

The catchment area of this basin is 4.7 Mio. km². The average observed discharge is 170,000 m³/s. Discharge at this station depends mostly on the timing, that means mostly on the routing and lake parameters.

Therefore Budyko cal. does not significantly improve the a priori parameter run.
Murray-Darling - Wakool Junction, Australia

Murray river is running through a semi-arid region. Most of the discharge is lost during this transfer. As the Budyko cal. is only looking at the grid-cell balance, it cannot be expected to be effective.

The a priori parameter run is overestimating observed discharge by 600%. Transmission lost is calibrated by the routing process. Discharge calibration gives reasonable good results and Budyko improves the results a little bit, but still not sufficient.