Enhancing Hydrological Modeling with Bias-Corrected Satellite Weather Data in Data-Scarce Catchments: A Comparative Analysis of SWAT and GR4J Models

Youness Hrour[1*], Zahra Thomas[1], Pauline Rousseau-Gueutin[2], Yassine Ait Brahim[3], Ophélie Fovet[1]

[1]Institut Agro, INRAE, SAS, F-35000 Rennes, France
[2]Univ Rennes, EHESP, Rennes, France
[3]International Water Research Institute (IWRI), Mohammed VI Polytechnic University (UM6P), Hay My Rachid, Ben Guerir 43150, Morocco

*** Correspondence:**
Youness Hrour
Hrour.iav@gmail.com

Appendix 1: Soil data used in the SWAT model

The Harmonized World Soil Database (Fao et al., 2012), was used in this study. This database contains 16,000 mapped units with two different soil layers (0-30 cm and 30-100 cm depth) (Nachtergaele et al., 2010). Table A1 lists all soil properties required by SWAT, their sources, references and calculation methods.

Table S1 : SWAT model input parameters for each soil type.

SWAT code	Description	Source
SOL_ZMX	Maximum rooting depth of soil profile (mm)	HWSD
SOL_Z	Depth from soil surface to bottom of layer (mm)	HWSD
SOL_BD	Moist bulk density (Mg/m^3 or g/cm^3)	HWSD
SOL_AWC	Available water capacity of the soil layer (mm H2O/mm soil)	HWSD
SOL_K	Saturated hydraulic conductivity (mm/hr)	Jabro's equation (Jabro ,1992)
HYDGRP	Soil hydrologic group (A, B, C, or D)	HWSD
SOL_CBN	Organic carbon content (% soil weight)	HWSD
CLAY	Clay content (% soil weight)	HWSD
SILT	Silt content (% soil weight)	HWSD
SAND	Sand content (% soil weight)	HWSD
ROCK	Rock fragment content (% total weight)	HWSD
SOL_ALB	Moist soil albedo	--
USLE_K	USLE equation soil erodibility (K) factor (units: 0.013 (metric ton m2 hr)/(m3-metric ton cm))	Wiliams' equation (Neitsch et al., 2011; Williams, 1995)

- **Jabro's equation (Jabro ,1992) :**

$$SOL\_K = 1000\,.\,exp^{(11{,}86-0{,}81\times\log(\mathrm{SILT})-1{,}09\times\log(\mathrm{CLAY})-4{,}64\times SOL_{BD})} \qquad \text{Eq. A1.1}$$

- **Wiliams' equation (Williams, 1995) :**

$$\mathrm{USLE_K} = f_{\mathrm{csand}} \times f_{cl-si} \times f_{orgc} \times f_{\mathrm{hisand}} \qquad \text{Eq. A1.2}$$

With:

$$f_{csand} = 0{,}2 + 0{,}3 \times \exp^{-0{,}256\times\mathrm{SAND}\times\left(\frac{1-\mathrm{SILT}}{100}\right)} \qquad Eq. A1.3$$

And:

$$f_{cl-si} = \left(\frac{\mathrm{SILT}}{\mathrm{SAND}+\mathrm{SILT}}\right)^{0{,}3} \qquad Eq. A1.4$$

$$f_{orgc} = (1 - 0{,}25 \times \mathrm{Orgc} + exp^{3{,}72-2{,}95\times\mathrm{orgc}}) \qquad Eq. A1.5$$

$$f_{\mathrm{hisand}} = 1 - \frac{0{,}7 \times \frac{1-\mathrm{SAND}}{100}}{\frac{1-\mathrm{SAND}}{100}} + exp^{-5{,}51\times 22{,}9\times\left(\frac{1-\mathrm{SAND}}{100}\right)} \qquad Eq. A1.6$$

Where Orgc is the percentage (%) of organic matter. For Sol_K, SILT, SOL_BD, USLE_K, CLAY, and SAND see Table A1.

Appendix 2: SWAT model parameters taken into account for the sensitivity analysis.

Table S2 : SWAT model parameters considered for the sensitivity analysis. The prefix 'v_' denotes a replacement, while 'r_' denotes a relative change (the existing parameter is multiplied by (1 + the specified value)".

Variable name	Definition	Unit	Spatial level	Default value	Initial inetrval
v_SURLAG (.bsn)	Surface runoff lag coefficient	days	Sub-basin	4	[1 , 31]
r_CN2 (.mgt)	Condition II curve number	-	HRU	Changes for HRU	[-0.75 , 0.75]*
r_OV_N (.hru)	Manning's "n" value for overland flow	-	HRU	Changes for HRU	[0.01 , 30]
r_SOL_AWC (.sol)	Available water capacity of the soil layer	mm H_20/mm sol	HRU	Changes for HRU	[-0.5 , 0.5]*
r_SOL_K (.sol)	Saturated hydraulic conductivity	mm/hr	HRU	Changes for HRU	[-0.5 , 0.5]*
r_SOL_BD (.sol)	Moist bulk density	Mg/m^3 or g/cm^3	HRU	Changes for HRU	[-0.5 , 0.5]*
v_ESCO (.hru ; .bsn)	Soil evaporation compensation factor	-	Sub-basin	0.95	[0,1]
v_EPCO (.hru ; .bsn)	Plant uptake compensation factor	-	Sub-basin	1	[0,1]
v_GW_delay (.gw)	Groundwater delay	days	HRU	31	[0 , 450]
Alpha_BF (.gw)	Baseflow alpha factor	1/days	HRU	0.048	[0 , 1]
v_GW_Revap (.gw)	Groundwater "revap" coefficient	-	HRU	0.02	[0.02 , 0.2]
v_GWQMN (.gw)	Threshold depth of water in the shallow aquifer required for return flow to occur	mm H2O	HRU	1000	[0 , 5000]
v_RCHRG_DP (.gw)	Deep aquifer percolation fraction	-	HRU	0.05	[0 , 1]
v_REVAPMN (.gw)	Threshold depth of water in the shallow aquifer for “revap” or percolation to the deep aquifer to occur	mm H2O	HRU	0	[0 , 500]
v_Lat_Time (.hru)	Lateral flow travel time	days	HRU	1	[0 , 180]
r_CH_N1 (.sub)	Manning’s “n” value for the tributary channels	-	Sub-basin	0.014	[0 , 0.3]
r_CH_N2 (.rte)	Manning's "n" value for the main channel	-	Sub-basin	1.014	[0 , 0.3]
r_HRU_SLP (.hru)	Average slope steepness	m/m	HRU	Changes for HRU	[-0.25 , 0.25]*
r_SLSUBBSN (.hru)	Average slope length	m	HRU	Changes for HRU	[-0.2 , 0.2]*

Appendix 3:

The performance of the models was assessed by calculating the efficiency criteria:

- **Nash-Sutcliffe efficiency coefficient (NSE):** is a commonly used measure of the accuracy of hydrological models (Nash and Sutcliffe, 1970). The NSE ranges from -∞ (no better than the mean of the observations) to 1 (perfect prediction).

$$NSE = 1 - \frac{\sum_{i=1}^{n}\left(Q_{sim}^{i} - Q_{obs}^{i}\right)^2}{\sum_{i=1}^{n}\left(Q_{obs}^{i} - \overline{Q_{obs}}\right)^2} \qquad Eq. A3.1$$

- **Coefficient of determination (R^2):** expresses the degree of linear correlation between measured and simulated discharge values. An R^2 of 1 indicates that the model predicts discharge perfectly, while an R^2 of 0 indicates that the model predicts discharge very poorly.

$$R^2 = \frac{\left[\sum_{i=1}^{n}\left(Q_{obs}^{i} - \overline{Q_{obs}}\right)\left(Q_{sim}^{i} - \overline{Q_{sim}}\right)\right]^2}{\sum_{i=1}^{n}\left(Q_{obs}^{i} - \overline{Q_{obs}}\right)^2 \sum_{i=1}^{n}\left(Q_{sim}^{i} - \overline{Q_{sim}}\right)^2} \qquad \mathrm{Eq.}\, A3.2$$

- **Root mean square error (RMSE):** is the standard deviation of the residuals, which are the discrepancies between the observations and the model simulations. A lower RMSE indicates closer agreement between observations and model simulations.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}\left(Q_{sim}^{i} - Q_{obs}^{i}\right)^2}{n}} \qquad \mathrm{Eq.}\, A3.3$$

- **Percentage bias (Pbias):** quantifies the average tendency of simulated discharge rates to deviate from their measured equivalents. A PBIAS value close to 0 indicates that the model is able to accurately reproduce the observed discharge, while a value significantly different from 0 indicates that the model has a systematic bias in its simulations. In general, a PBIAS value of less than 25% is considered satisfactory (Moriasi et al., 2007). Where Q_{obs}, Q_{sim} and $\overline{Q_{ob}}$, and $\overline{Q_{sim}}$ are the observed discharge, simulated discharge, the mean of the observed and simulated discharge at time step i, respectively, and n is the number of observations.