remote sensing

# Supplementary Materials: Assessment of GPM and TRMM Multi-satellite Precipitation Products in Streamflow Simulations in a Data-sparse Mountainous Watershed in Myanmar

**Fei Yuan, Limin Zhang, Khin Wah Wah Win, Liliang Ren, Chongxu Zhao, Yonghua Zhu, Shanhu Jiang and Yi Liu**

## S1. Detailed Descriptions of Streamflow Simulation Schemes

To investigate the feasibility of the IMERG and 3B42V7 precipitation products in streamflow simulations in the Chindwin River basin, five precipitation data sets (the gauge-based precipitation data, the original and corrected 3B42V7 data sets, and the original and corrected IMERG data sets on a 0.25° resolution) were, respectively, used to drive the XAJ model to perform historical daily hydrological simulations in the Chindwin River basin from 1 April 2014 to 31 December 2015.

For these five simulation runs, the XAJ model parameters were independently calibrated by fitting the calculated historical daily streamflow against the observed data at the five streamflow stations, with different precipitation inputs. Model calibration was achieved with the aid of the Shuffled Complex Evolution (SCE-UA) automatic optimization method [1-2]. It should be noted that the hydrological model parameters are supposed to be ideally calibrated using the high-quality observed precipitation data measured in a dense gauge network. In this way, the calibrated parameters may rationally characterize watershed hydrological features, and they should be adopted for all simulation runs with different satellite precipitation inputs. However, the gauge-based precipitation data in this study were derived from a sparsely-distributed rain gauge network (five weather stations in a controlled area of 110,350 km$^2$) and were not adequate to represent the real rainfall regimes for model calibration. Therefore the simulation runs using the 3B42V7 and IMERG products adopted the precipitation product-specific model parameters, instead of the rain gage-calibrated model parameters. This model calibration strategy is likely to produce some parameter values that are beyond their physical ranges. To refrain from this problem, the searching space was strictly defined within the physical ranges of the XAJ parameters in the SCE-UA algorithm.

Although the maxima of the Nash–Sutcliffe model efficiency coefficient (NSE) is a commonly-used objective function for hydrological model optimization, it tends to give higher importance to high flows, and the optimized model may not accurately capture low flow processes. To derive the model parameters that can comprehensively characterize both high- and low-flow processes, we added the maxima of the log-transformed NSE (LogNSE) to the objective function, which is given by

$$f = max(NSE) + max(LogNSE) \tag{1}$$

$$NSE = 1 - \frac{\sum_{i=1}^{n}(Q_i^s - Q_i^o)^2}{\sum_{i=1}^{n}(Q_i^o - \overline{Q^o})^2} \tag{2}$$

$$LogNSE = 1 - \frac{\sum_{i=1}^{n}[log(Q_i^s) - log(Q_i^o)]^2}{\sum_{i=1}^{n}[log(Q_i^o) - \overline{log(Q^o)}]^2} \tag{3}$$

where $n$ is the sample size of the observed or calculated streamflow time series; $Q_i^o$ and $Q_i^s$ denote the observed and calculated daily streamflow at the $i$th time step (m$^3$/s); $\overline{Q^o}$ and $\overline{log(Q^o)}$ represent the mean observed streamflow and mean log-transformed observed streamflow(m$^3$/s), respectively.

Although the SCE-UA algorithm has proved to be consistent, effective, and efficient in locating the parameter values of a hydrologic model that optimizes a given objective function, it has a critical deficiency of population degeneration, which supposes the global optimum is located on a linearly-operated subspace of the parameter space [3]. This assumption is likely to lead to the situation that SCE-UA misses the global optimum or misconverges. To retrieve hydrological parameter sets that best possibly approximate the global optimum, 20 independently initialized runs of the SCE-UA optimizations were executed. For each SCE-UA run, optimization was terminated when iteration reached 10,000 times or population converged within 0.1% of the parameter space. Finally, the parameter set in the optimization run with the highest objective function value was adopted for historical streamflow simulation.

In this study, the entire Chindwin River was separated into 204 grid cells on a 0.25° resolution, and the calibration and validation periods were defined as 1 April 2014 – 31 December 2014 and 1 January 2015 – 31 December 2015, respectively. Since there are five streamflow stations located in the watershed (Figure 1), the entire basin was divided into five parameter zones, which are the area controlled by the Hkamti station, and the four interval regions between the upstream and downstream stations. Within a parameter zone, the parameter values for each grid cell are identical.

After precipitation product-specific calibrations, the XAJ model driven by the five different precipitation data sets was adopted for historical daily hydrological simulations. The five sets of simulated streamflow time series using different precipitation inputs were compared against the observed streamflow to assess the feasibility of 3B42V7 and IMERG precipitation products in streamflow simulations in the study area. The accuracy of streamflow simulations were evaluated using four statistical indices, namely, BIAS, CC, NSE, and LogNSE.

## References

1. Duan, Q.; Soorooshian, S.; Gupta, V. Effective and efficient global optimization for conceptual rainfall-runoff models. *Water Resour. Res.* **1992**. *28*, 1015–1031, doi: 10.1029/91WR02985.
2. Duan, Q.; Gupta, V.K.; Sorooshian, S. A shuffled complex evolution approach for effective and efficient global minimization. *J. Optimiz. Method Appl.* **1993**, *76*, 501–521, doi: 10.1007/BF00939380.
3. Chu, W.; Gao, X.; Sorrooshian, S. Improving the shuffled complex evolution scheme for optimization of complex nonlinear hydrological systems: Application to the calibration of the Sacramento soil-moisture accounting model. *Water Resour. Res.* **2010**. *46*, W09530, doi: 10.1029/2010WR00924.

**Table S1.** Physical meaning and sensitivity of the XAJ model parameters and the calibrated parameter values in the area controlled by the Hkamti station for the simulation runs using the gauge-based, original 3B42V7 and original IMERG precipitation data sets.

| Model Parameters | Physical Meaning and Sensitivity | Precipitation Inputs | | |
|---|---|---|---|---|
| | | **Gauge** | **Original 3B42V7** | **Original IMERG** |
| **K** | Coefficient of potential evapotransipiration; very sensitive | 1.379 | 0.864 | 0.838 |
| **WUM** | Mean tention water capacity in the upper soil layer (mm); insensitive | 10.215 | 18.685 | 19.642 |
| **WLM** | Mean tention water capacity in the lower soil layer (mm); insensitive | 67.156 | 85.231 | 87.253 |
| **WDM** | Mean tention water capacity in the deep soil layer (mm); insensitive | 42.356 | 56.235 | 53.156 |
| **C** | Coefficient of deep evapotranspiration; very insensitive | 0.198 | 0.178 | 0.182 |
| **B** | Exponent of the tention water storage curve; insensitive | 0.486 | 0.401 | 0.396 |
| **EX** | Exponent of the free water storage curve; insensitive | 1.35 | 1.13 | 1.21 |
| **SM** | Areal mean free water storage (mm); very sensitive | 86.153 | 12.103 | 11.856 |
| **KI** | Outflow coefficient of interflow; sensitive | 0.611 | 0.498 | 0.398 |
| **KG** | Outflow coefficient of groudwater runoff; sensitive | 0.089 | 0.192 | 0.302 |
| **CS** | Recession constant of surface runoff; very sensitive | 0.459 | 0.589 | 0.598 |
| **CI** | Recsssion constant of interflow; very sensitive | 0.912 | 0.935 | 0.942 |
| **CG** | Recession constant of groundwater runoff; very sensitive | 0.999 | 0.977 | 0.977 |
| **Kmus** | Storage-time constant for the sub-reach in the Muskingum routing method (h); very sensitive | 28.353 | 29.956 | 29.451 |
| **Xmus** | Proportionality constant in the Muskingum method; very sensitive | 0.015 | 0.297 | 0.307 |

**Table S2.** Statistical indices of the simulated monthly streamflow processes at the five weather stations in the Chindwin River basin (1 April 2014 – 31 December 2015).

| Streamflow Stations | Precipitation Inputs | BIAS (%) | CC | NSE |
|---|---|---|---|---|
| **Hkamti** | Gauge | −2.9 | 0.965 | 0.928 |
| | Original 3B42V7 | −16.6 | 0.951 | 0.842 |
| | Corrected 3B42V7 | −1.9 | 0.978 | 0.952 |
| | Original IMERG | −28.7 | 0.947 | 0.720 |
| | Corrected IMERG | −1.2 | 0.955 | 0.908 |
| **Homalin** | Gauge | −0.7 | 0.973 | 0.933 |
| | Original 3B42V7 | −16.9 | 0.948 | 0.825 |
| | Corrected 3B42V7 | −5.3 | 0.978 | 0.946 |
| | Original IMERG | −31.2 | 0.944 | 0.660 |
| | Corrected IMERG | 0.6 | 0.948 | 0.883 |
| **Mawlaik** | Gauge | −1.8 | 0.978 | 0.956 |
| | Original 3B42V7 | −10.9 | 0.976 | 0.904 |
| | Corrected 3B42V7 | −0.9 | 0.983 | 0.966 |
| | Original IMERG | −23.5 | 0.977 | 0.807 |
| | Corrected IMERG | 2.2 | 0.916 | 0.839 |
| **Kalewa** | Gauge | −5.2 | 0.966 | 0.930 |
| | Original 3B42V7 | −13.1 | 0.968 | 0.877 |
| | Corrected 3B42V7 | −4.7 | 0.971 | 0.939 |
| | Original IMERG | −25.9 | 0.964 | 0.764 |
| | Corrected IMERG | −1.5 | 0.909 | 0.818 |
| **Monywa** | Gauge | −4.7 | 0.939 | 0.948 |
| | Original 3B42V7 | −14.4 | 0.909 | 0.870 |
| | Corrected 3B42V7 | −5.6 | 0.944 | 0.960 |
| | Original IMERG | −24.8 | 0.887 | 0.778 |
| | Corrected IMERG | −1.3 | 0.913 | 0.836 |

Sample size for monthly streamflow time series is 20.

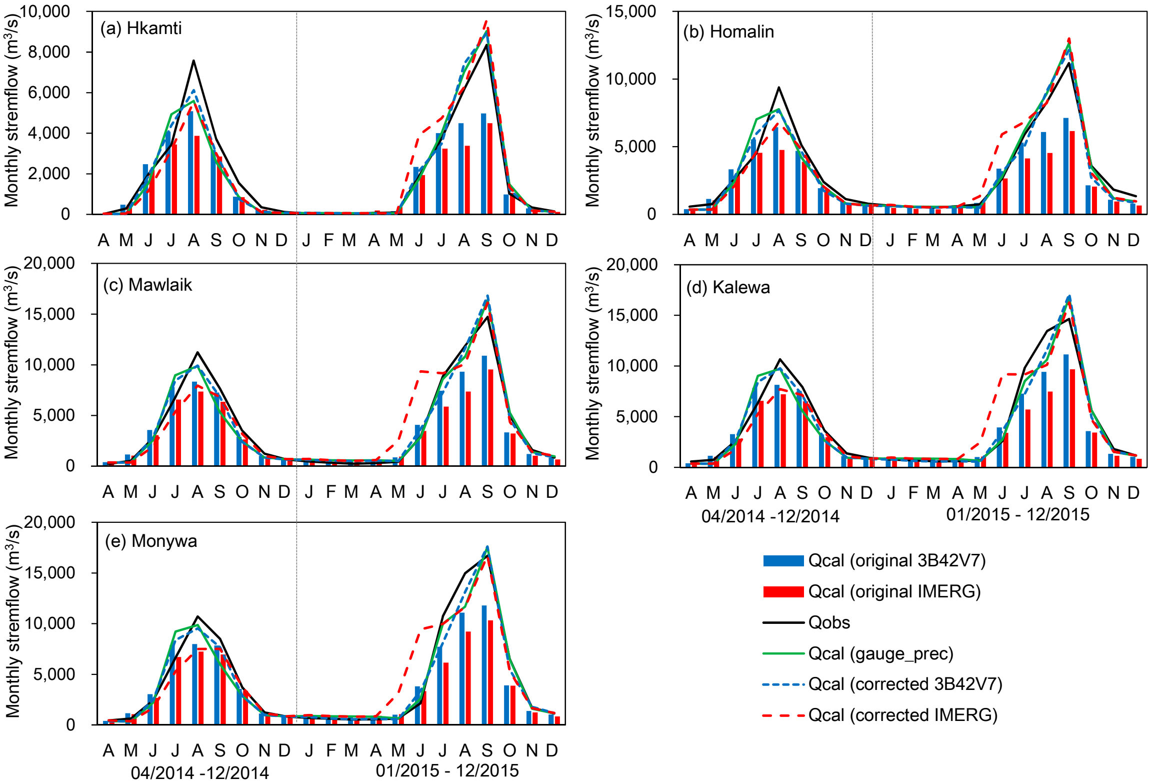

**Figure S1.** Simulated monthly hydrographs using the gauge-based precipitation data, the original 3B42V7 and IMERG precipitation estimates, and their corrected data sets at the fiver streamflow stations. The terms in the figure are the same as in Figure 8 in main text of the article.