

HYDROLOGICAL MODELING AT THE DAILY LEVEL OF THE BOLIVIAN AMAZON FOR HYDROELECTRIC USE PURPOSES, COCHABAMBA-BOLIVIA.

First A. Author Roxana Guillen Salvador
Universidad Mayor de San Simón
e-mail: roxana.guillen@fcyt.umss.edu.bo

Co-authors: Mauricio F. Villazón Gómez¹, Santiago Mendoza Paz², Andres Gonzales Amaya¹, William Usher³
mauriciovillazon@fcyt.umss.edu.bo, santiago.mendoza@ird.fr,
andresgonzales.a@fcyt.umss.edu.bo, wusher@kth.se

¹Hydraulic Laboratory, Universidad Mayor de San Simón (Bolivia)

²IRD, French National Research Institute for Sustainable Development (France)

³KTH, Royal Institute of Technology (Sweden)

ABSTRACT

Bolivia has diverse topographical features, including areas with high potential for the installation of hydroelectric plants. The country has two major water systems: the macro basin of the Bolivian Amazon, with a potential of 34,208 MW, and the macro basin of La Plata with 5,359 MW (Zárate, S., Villazón, M. F., Balderrama, S., & Quoilin, S., 2021), values that exceed the current national demand. However, Bolivia is undergoing a transition and diversification of its energy matrix by incorporating renewable energy sources, including solar and wind power (ENDE, 2024).

The objective of this study is to implement a continuous daily hydrological modeling system for the macro basin of the Bolivian Amazon. This tool aims to support the development of hydroelectric projects by providing the necessary information regarding water availability in the area, while taking into account environmental considerations and ecosystem preservation.

The WEAP (Water Evaluation and Planning) software was applied using the Soil Moisture Method (Precipitation–Runoff) to simulate the hydrological processes during the period 1980–2020.

Secondary information on land use and land cover was employed for model parameterization, while climatological inputs were obtained from the Gridded Meteorological Ensemble Tool (GMET) (NCAR, 2022). The hydrological model was subsequently rescaled both spatially and temporally to represent four priority areas, corresponding to hydroelectric and multipurpose systems.

The hydrological assessment was carried out in the Bolivian Amazon basin, specifically in its headwater catchments, through the hydrometric stations coded as Q_AM_01 (El Kaka area with 21,374.86 km²), Q_AM_02 (La Paz, 31,626.85 km²), Q_AM_03 (Rurrenabaque, 48,566.91 km²), and Q_AM_04 (Gundonovia, 20,419.48 km²).

Q_AM_04 sub-basins were utilized. The statistical evaluation of the hydrological modeling in these sub-basins showed Nash-Sutcliffe Efficiency (NSE) values ranging from 0.56 to 0.72, Kling-Gupta Efficiency (KGE) values between 0.74 and 0.85, percentage bias (PBIAS) from –4.8 % to 0.095 %, the ratio of the root mean square error to the standard deviation of the measured data (RSR) from 0.53 to 0.67, and LogNSE values between 0.59 and 0.74. These

results indicate that the model performed satisfactorily across all stations, with the highest accuracy observed at Q_AM_01, thereby demonstrating the robustness and reliability of the calibration process (Moriassi, et al., 2007) (Moriassi, D. N., et al., 2015) (Gupta, H. V., Kling, H., Yilmaz, K. K., & Martinez, G. F., Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling, 2009) (Kling, H., Fuchs, M., & Paulin, M., 2012).

KEYWORDS

Hydrological modeling, Rainfall–runoff modeling, Bolivian Amazon, WEAP, GMET, Hydropower, Calibration.

INTRODUCTION

Using hydrometric records from the Bolivian Amazon, daily-scale data were employed for the hydrological model, which was implemented through the WEAP software.

WEAP (Water Evaluation and Planning) is a hydrological planning and management software developed by the Stockholm Environment Institute (SEI). It is based on a continuous and aggregated modeling approach within a deterministic mathematical framework, applying the fundamental principle of water balance (SEI, 2015). The platform enables the definition of temporal and spatial domains, system components, and alternative scenarios for evaluation (Sieber, J., Purkey, D., & Huber-Lee, A., 2005). In this study, the Soil Moisture Method was implemented, a semi-distributed conceptual model that represents the soil as two reservoirs: the root zone (surface zone) and the deep zone (aquifer), allowing the simulation of runoff generation and percolation processes.

The information provided to the WEAP software includes topography from a Digital Elevation Model (DEM) such as HydroSHEDS-SRTM, which is used for the delineation of hydrological basins and the disaggregation of hydrological units (land cover). Climatological data are defined using NetCDF products from ESA-CCI-LC (European Space Agency Climate Change Initiative – Land Cover), GMET-2020 (Gridded Meteorological Dataset 2020), and SEI(Stockholm Environment Institute), respectively.

The basic unit used for the representation of areas is the basin, referred to as a 'catchment,' to which modelling parameters are assigned. In the calibration process, the six most representative parameters are detailed in Table 1.

Table 1. Modelling Parameters

Parameter	Description
Kc	Crop coefficient
SWC	Soil Water Capacity [mm].
DWC	Deep Water Capacity [mm]
RRF	Runoff Resistance Factor.
RZC (Ks)	Root Zone Conductivity [mm/day]
DC (Kd)	Deep Conductivity [mm/day].

Source: WEAP modeling parameters, 2025

The root zone, located in the upper reservoir, is influenced by the parameters Kc, RRF, SWC, RZC, and the preferential flow direction governed by the slope. The lower tank, or deep zone, depends on the parameters DC, DWC, and Z2, which refer to the percentage level of the tank; these latter parameters assign a single value per catchment.

The efficiency of the model is evaluated using statistical criteria applied to the series of measured in situ flows versus the modeled ones (Moriasi, D. N., et al., 2015). The representative statistical parameters taken as performance evaluation guidelines during the model calibration stage are Percent Bias (PBIAS), is calculated with the equation 1.

$$PBIAS = \left[\frac{\sum_{i=1}^n (Y_i^{obs} - Y_i^{sim}) * 100}{\sum_{i=1}^n (Y_i^{obs})} \right] \quad (1)$$

Nash-Sutcliffe Efficiency (NSE), is calculated with the equation 2, and the ratio of the root mean square error to the standard deviation of the measured data (RSR) [5], calculated with the equation 3.

$$NSE = 1 - \left[\frac{\sum_{i=1}^n (Y_i^{obs} - Y_i^{sim})^2}{\sum_{i=1}^n (Y_i^{obs} - Y^{mean})^2} \right] \quad (2)$$

$$RSR = \frac{RMSE}{STDEV_{obs}} = \frac{\left[\sqrt{\sum_{i=1}^n (Y_i^{obs} - Y_i^{sim})^2} \right]}{\left[\sqrt{\sum_{i=1}^n (Y_i^{obs} - Y^{mean})^2} \right]} \quad (3)$$

The Kling-Gupta Efficiency (KGE) metric was also applied, supported by studies that have evaluated its suitability for hydrological modeling, particularly in exploring its behavior and relationship with other performance metrics such as the Nash–Sutcliffe Efficiency (NSE) (Gupta, H. V., Kling, H., Yilmaz, K. K., & Martinez, G. F., 2009), (Knoben, W. J. M., Freer, J. E., & Woods, R. A., 2019), (Kling, H., Fuchs, M., & Paulin, M., 2012). The KGE was calculated according to Equation 4.

$$KGE = 1 - \sqrt{\left(\frac{cov(sim, obs)}{\sigma_{sim}\sigma_{obs}} - 1 \right)^2 + \left(\frac{\sigma_{sim}}{\sigma_{obs}} - 1 \right)^2 + \left(\frac{\mu_{sim}}{\mu_{obs}} - 1 \right)^2} \quad (4)$$

Where σ is the standard deviation and μ is the mean, therefore the KGE uses correlation, variability, and volume for model evaluation.

In this study, we also analyzed the Log-Nash–Sutcliffe Efficiency (Log-NSE), which is calculated by applying the NSE metric to logarithmically transformed flow data. This variant increases the relative weight of low flows and is particularly valuable when accurate simulation of base-flow conditions is critical to hydrological model evaluation. Its use has been widely supported in the hydrological modeling literature, where Log-NSE has been recommended as a complementary performance metric to address the limitations of the traditional NSE, especially in representing both high- and low-flow regimes (Krause, P., Boyle, D. P., & Bäse, F., 2005), (Santos, L., Thirel, G., & Perrin, C., 2018), (U.S. Geological Survey, 2021). For instance, in a machine learning streamflow modeling study, Log-NSE values were comparable or slightly higher than traditional NSE, indicating robust model skill across flow magnitudes (U.S. Geological Survey, 2021). In the present research, the inclusion of Log-NSE complements the evaluation through NSE and KGE, providing a more nuanced insight into model accuracy under low-flow conditions.

$$LogNSE = 1 - \left[\frac{\sum_{t=1}^n (\log(Q_{obs,t} + 1) - \log(Q_{sim,t} + 1))^2}{\sum_{t=1}^n (\log(Q_{obs,t} + 1) - \log(Q_{obs} + 1))^2} \right] \quad (5)$$

Where:

- $Q_{obs,t}$ = observed discharge at time t

- $Q_{sim,t}$ = simulated discharge at time t
- n = number of time steps
- $\overline{\log(Q_{obs} + 1)}$ = mean of the log-transformed observed discharges

Adding +1 inside the log ensures that zero flow values do not cause undefined results.

This research builds upon the framework and baseline information of the 2017 Bolivia Surface Water Balance (BHSB), developed at the monthly scale by the Stockholm Environment Institute (SEI) in collaboration with the National Meteorology and Hydrology Service (SENMHI), the Institute of Hydraulics at Universidad Mayor de San Andrés (IHH-UMSA), the Hydraulics Laboratory at Universidad Mayor de San Simón (LH-UMSS), the National Center for Atmospheric Research (NCAR, USA), and the Ministry of Environment and Water (MMAyA) (Ministry of Environment and Water Bolivia, 2018), (NCAR, 2022), (SEI, 2015). The present study adapts this framework to a hydrological model focused on four prioritized sub-basins, using updated and corrected meteorological data from GMET through 2020, and calibrating the model specifically for these basins to improve daily streamflow representation.

METHOD

For the implementation of the hydrological model (Figure 1), the WEAP software was used, employing the Soil Moisture Method (precipitation–runoff). The study area exhibits strongly contrasting wet and dry periods, with the rainy season generally occurring from November to March across the region, and the dry season prevailing during the remaining months (Ministry of Environment and Water Bolivia, 2018). The model focuses on four prioritized sub-basins, each monitored by an existing hydrometric station with labels Q_AM_01, Q_AM_02, Q_AM_03, Q_AM_04 (Q_EL KAKA, Q_LA PAZ, Q_RURRENABAQUE, and Q_GUNDONOVIA respectively.), which were used for calibration and validation. The modeling period spans from 1980 to 2020, using updated and corrected meteorological data from the Gridded Meteorological Ensemble Tool (GMET) (NCAR, 2022). Calibration was performed primarily manually, complemented by a Python-based tool developed for this study that allows semi-automatic parameter adjustment within defined ranges, enhancing efficiency and reproducibility.

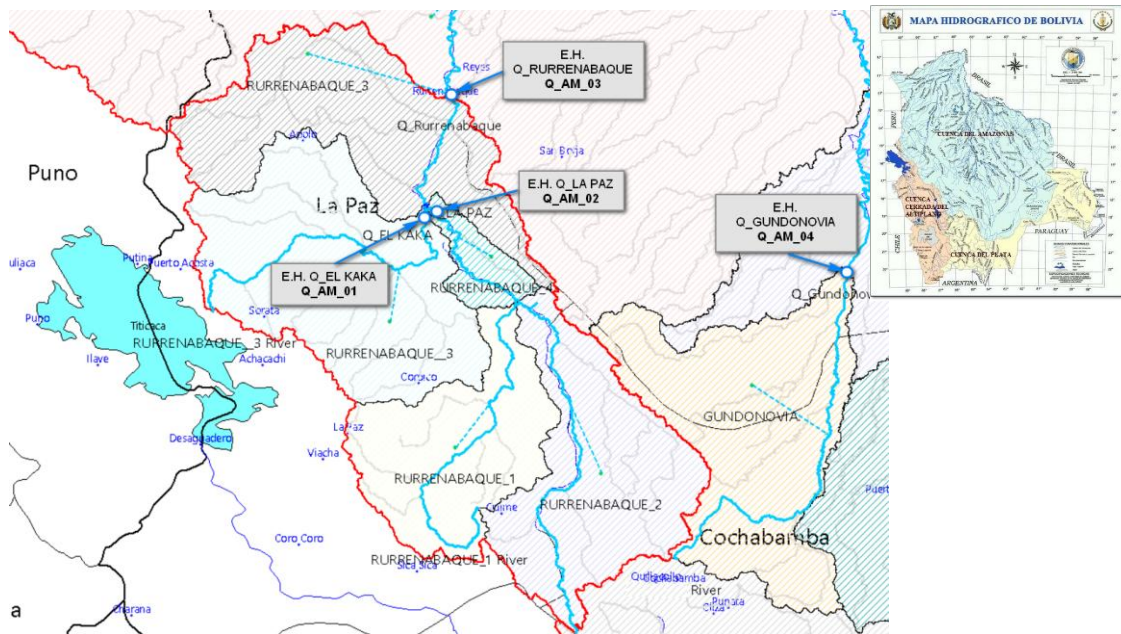


Figure 1. WEAP layout of the four modeled sub-basins with hydrometric stations used for calibration in four hydrometric stations Q_AM_01, Q_AM_02, Q_AM_03, Q_AM_04.

The hydrological model utilizes daily precipitation and temperature data from the Gridded Meteorological Ensemble Tool (GMET), which employs station data to create continuous, gridded analyses. Unlike other tools, GMET incorporates historical data and estimates potential prediction errors in its assessments, allowing for the quantification of uncertainty through interpolation that estimates the probability of occurrence and distribution of precipitation volumes (NCAR, 2022).

Previous studies, such as the Bolivia Surface Water Balance (BHSB), used hydrological information at a monthly scale (Ministry of Environment and Water Bolivia, 2018). In the present model, hydrological data are used at a daily scale

Previous studies, such as the Bolivia Surface Water Balance (BHSB), used hydrological information at a monthly scale (Ministry of Environment and Water Bolivia, 2018). In the present model, hydrological data are used at a daily scale, along with spatial refinement focused on four prioritized sub-basins. Daily precipitation and temperature inputs were obtained from the GMET (Gridded Meteorological Ensemble Tool), which provides spatially gridded meteorological data that account for variability in precipitation distribution across the study area (NCAR, 2022). In the Amazonian portion of the model, annual precipitation varies approximately from 500 mm to over 6000 mm, reflecting both topographic and regional climatic gradients.

The downscaling procedure was conducted in three stages: first, the hydrological model was adapted from a monthly to a daily temporal resolution; second, runoff coefficients were verified using observed flow data from the region's hydrometric stations; and third, the model was calibrated and validated at the daily scale to ensure reliable projections under the same conditions.

The adaptation of the continuous hydrological model in WEAP incorporated data on vegetation cover and soil types, along with climatic variables such as relative humidity, wind speed, and sunshine hours, projected on an annual cyclical basis. Daily precipitation and temperature inputs were obtained from the Gridded Meteorological Ensemble Tool (GMET), which was also employed in the Bolivia Surface Water Balance (BHSB) (Ministry of Environment and Water Bolivia, 2018), (NCAR, 2022).

The verification of runoff coefficients highlighted uncertainties in the GMET 2020 precipitation and temperature data, primarily associated with the infilling process, as well as with the density and spatial distribution of available meteorological stations. These factors influence the accuracy of gridded climate datasets and, consequently, hydrological simulations. Despite these limitations, GMET provides one of the most consistent and updated meteorological products for Bolivia, making it suitable for hydrological modeling applications at the daily scale (NCAR, 2022).

These sources of uncertainty affect the simulated runoff and its relationship with precipitation at the basin scale. In contrast to the national-level BHSB, which applied precipitation correction factors (PCF) in several basins to address underestimation issues, the present study focused on four priority basins where daily calibration was conducted using available hydrometric stations. Although sub-basins such as Zongo, Miguillas, Misicuni, Corani, and Ivirizu are not individually delineated in the current WEAP configuration, they are encompassed within the modeled macro-basins. Consequently, the calibration performed here provides a robust foundation for future refinements, including the explicit incorporation and calibration of these sub-basins in subsequent stages of model development.

Runoff was validated at each hydrometric station by comparing the accumulated observed flows with precipitation data from the GMET-2020 meteorological grid. These precipitation values were multiplied by the basin area and the average daily rainfall. The results are summarized in Table 2, which presents the Runoff Coefficient (CR) and the Precipitation Correction Factor (PCF), both derived from the GMET-2020 dataset. These values were calculated using a code, script that considered only the periods with available observed flow data, within the complete model covering the entire national territory.

Table 2: Runoff coefficients (CR)

NRO.	CACHMENT	CR GMET 2020
1	Rurrenabaque	0.62
2	Penhas Amarillas	0.49
3	Portachuelo 2 estrellas	0.48
4	El Sena	0.58
5	Riberalta	0.55
6	Cachuela Esperanza	0.49
7	Paraiso	0.29
8	Abapo	0.29
9	Pailas	0.31
10	Angostura	0.23
11	La Belgica	0.25
12	Puente Eisenhower	0.13
13	El Carmen	0.23
14	Puerto Villarroel	1.66
15	Santa Rosa de Chapare	0.87
16	Camiaco	0.49
17	Gundonovia	0.37
18	Los Puentes	0.60
19	Puerto Almacen	0.60
20	San Borjita	0.22
21	Puerto Junin	0.41
22	Puerto Siles	0.43
23	Matto Grosso	0.19
24	Pimenteiras	0.21
25	Pedras Negras	0.18
26	Principe	0.17
27	Campamento More	0.17
28	Guayanamerin	0.28
29	Abuna	0.38
30	Porto Velho	0.38
31	Brasilea	0.29
32	Rio Branco	0.27

For the calibration of the hydrological model, the performance was evaluated using several widely recognized statistical criteria, including the Nash-Sutcliffe Efficiency (NSE), Kling-Gupta Efficiency (KGE), Percent Bias (PBIAS), the Ratio of the Root Mean Square Error to the Standard Deviation of Measured Data (RSR), and the Logarithmic Nash-Sutcliffe Efficiency (Log-NSE). PBIAS measures the average tendency of simulated flows to overestimate or underestimate observed flows, providing a clear indication of model bias (Gupta, H. V., Kling, H., Yilmaz, K. K., & Martinez, G. F., 2009). NSE and Log-NSE assess

the goodness-of-fit between observed and simulated data, with Log-NSE giving additional weight to low flows for improved evaluation of base-flow conditions (Santos, L., Thirel, G., & Perrin, C., 2018) (U.S. Geological Survey, 2021). KGE evaluates correlation, bias, and variability, offering a comprehensive assessment of model performance (Gupta, H. V., Kling, H., Yilmaz, K. K., & Martinez, G. F., 2009) (Knoben, W. J. M., Freer, J. E., & Woods, R. A., 2019) (Kling, H., Fuchs, M., & Paulin, M., 2012). Finally, RSR provides a normalized error metric derived from the RMSE, serving as a complementary indicator for model evaluation (Moriassi, et al., 2007) (Table 3, 4 and 5).

Table 3: General statistical indices of watershed simulation accuracy (Moriassi, D. N., et al., 2015)

Performance Rating	RSR	NSE	PBIAS (%) Streamflow
Very good	$0.00 \leq \text{RSR} \leq 0.50$	$0.75 < \text{NSE} \leq 1.00$	$\text{PBIAS} < \pm 10$
Good	$0.50 < \text{RSR} \leq 0.60$	$0.65 < \text{NSE} \leq 0.75$	$\pm 10 \leq \text{PBIAS} < \pm 15$
Satisfactory	$0.60 < \text{RSR} \leq 0.70$	$0.50 < \text{NSE} \leq 0.65$	$\pm 15 \leq \text{PBIAS} < \pm 25$
Unsatisfactory	$\text{RSR} > 0.70$	$\text{NSE} \leq 0.50$	$\text{PBIAS} \geq \pm 25$

The Kling-Gupta Efficiency (KGE) was employed as an additional performance metric to complement NSE and Log-NSE, providing a comprehensive assessment of the hydrological model's ability to reproduce observed streamflow dynamics. KGE considers the correlation, variability, and bias between simulated and observed flows, offering a balanced evaluation of model performance across different flow regimes. Table 4 presents the classification of model performance based on KGE values, adapted from widely recognized hydrological modeling studies (Kling et al., 2012; Gupta et al., 2009; Moriassi et al., 2007; Andréassian et al., 2015; Bennett et al., 2013; Fowler et al., 2018; McCuen et al., 2006; Beven, 2018; Medici et al., 2019). This classification enables the identification of areas where the model performs very well, well, satisfactorily, or unsatisfactorily, thus supporting robust model evaluation and guiding future calibration efforts.

Table 4: Hydrological model performance classification based on KGE

Performance Rating	KGE	References
Very good	$0.90 \leq \text{KGE} \leq 1.00$	(Kling, Gupta, & Andres, 2012) (Gupta, H. V., Kling, H., Yilmaz, K. K., & Martinez, G. F., 2009) (Moriassi, et al., 2007)
Good	$0.75 \leq \text{KGE} < 0.90$	(Andréassian, V., Perrin, C., Michel, C., Usart-Sanchez, I., & Lavabre, J., 2015)
Satisfactory	$0.50 \leq \text{KGE} < 0.75$	(Bennett, N. D., et al., 2013) (Fowler, K. J., Peel, M. C., Western, A. W., & Zhang, L., 2018) (McCuen, R. H., Knight, Z., & Cutter, A. G., 2006)
Unsatisfactory	$\text{KGE} < 0.50$	(Beven, 2018) (Medici, G., Abrahão, R., Siqueira, J., & Marco, D. A., 2019)

Source: Author elaboration based on the cited literature.

In addition to traditional performance metrics such as NSE and KGE, the Log-Nash–Sutcliffe Efficiency (Log-NSE) was employed to evaluate the model's capability in simulating low-flow conditions. Log-NSE applies the NSE criterion to logarithmically transformed streamflow data, thereby increasing the sensitivity to low flows and providing a more balanced assessment across the full range of observed discharges. The classification of model performance according to Log-NSE is presented in Table 5, based on established hydrological modeling studies (Pushpalatha et al., 2012; Perrin et al., 2001; Krause et al., 2005; Fowler et al., 2018; Beven, 2018; Medici et al., 2019). This metric complements NSE and KGE by offering nuanced insight into the model's performance under low-flow conditions, which are critical for hydropower planning and water resource management.

Table 5. Classification of hydrological model performance based on Log-NSE

Performance Rating	Log-NSE	References
Very good	$0.75 \leq \text{Log-NSE} \leq 1.00$	(Pushpalatha, R., Perrin, C., Le Moine, N., & Andréassian, V., 2012) (Fowler, K. J., Peel, M. C., Western, A. W., & Zhang, L., 2018)
Good	$0.65 \leq \text{Log-NSE} < 0.75$	(Perrin, C., Michel, C., & Andréassian, V., 2001) (Fowler, K. J., Peel, M. C., Western, A. W., & Zhang, L., 2018)
Satisfactory	$0.50 \leq \text{Log-NSE} < 0.65$	(Krause, P., Boyle, D. P., & Bäse, F., 2005) (McCuen, R. H., Knight, Z., & Cutter, A. G., 2006)
Unsatisfactory	$\text{Log-NSE} < 0.50$	(Beven, 2018) (Medici, G., Abrahão, R., Siqueira, J., & Marco, D. A., 2019)

Source: Author elaboration based on the cited literature.

RESULTS

The hydrological model was calibrated and validated using observed flow data from four hydrometric stations within the Rurrenabaque and Gundonovia basins. Model performance was assessed using multiple statistical indicators: Nash–Sutcliffe Efficiency (NSE), Kling–Gupta Efficiency (KGE), Percent Bias (PBIAS), the ratio of the root means square error to the standard deviation of measured data (RSR), and Log-Nash–Sutcliffe Efficiency (Log-NSE). These metrics provide complementary insights into the model's ability to reproduce both high and low flows, capturing temporal variability and base-flow conditions relevant for hydropower and water management applications.

Table 6 summarizes the model performance for each hydrometric station. The results indicate good to very good agreement between observed and simulated flows.

Hydrometric Station Code	Station	NSE	KGE	PBIAS (%)	RSR	Log-NSE
Q_AM_01	Q_EL KAKA	0.72	0.85	-0.24	0.53	0.73
Q_AM_02	Q_LA PAZ	0.64	0.74	0.095	0.60	0.71
Q_AM_03	Q_RURRENABAQUE	0.62	0.74	-0.16	0.62	0.74
Q_AM_04	Q_GUNDONOVIA	0.56	0.74	1.8	0.67	0.59

The highest performance is observed at Q_AM_01, with a KGE of 0.85 and Log-NSE of 0.73, classified as very good according to the criteria of Moriasi et al. (2007) and Kling et al. (2012). Q_AM_02 and Q_AM_03 display good performance, while Q_AM_04 is classified as satisfactory, reflecting slightly lower NSE and Log-NSE values but still adequate for hydrological applications.

Figures 2–5 present the observed versus simulated daily flows for each station. These figures illustrate the model's ability to reproduce the magnitude and temporal distribution of streamflow, including low-flow periods, which are particularly important for assessing hydropower potential. The Log-NSE values confirm that low flows are well captured, complementing the overall performance evaluation provided by NSE and KGE.

Overall, the combined evaluation using NSE, KGE, RSR, PBIAS, and Log-NSE indicates that the model performs well in the study area. These results provide a robust basis for hydrological assessments and future water resource management and hydropower planning in these basins.

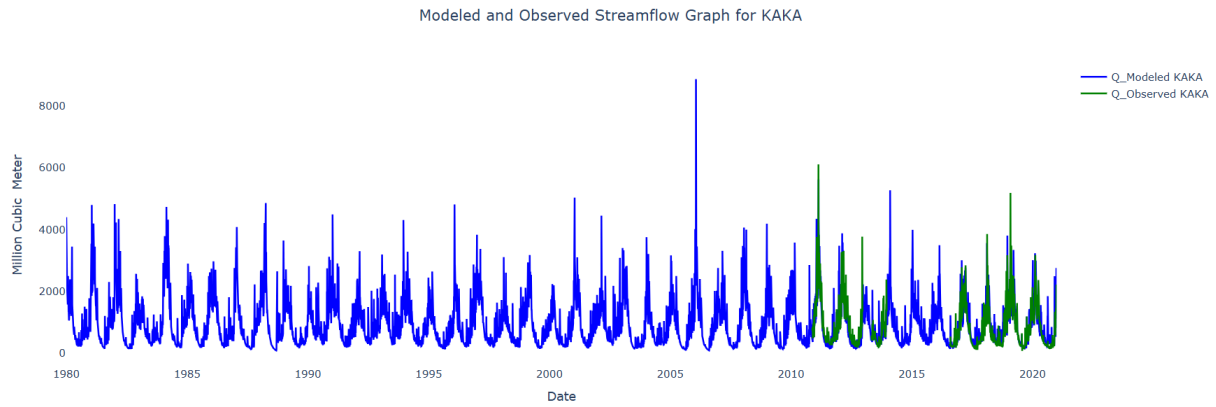


Figure 2: Modeled and observed streamflow at Q_AM_01 hydrometric station.

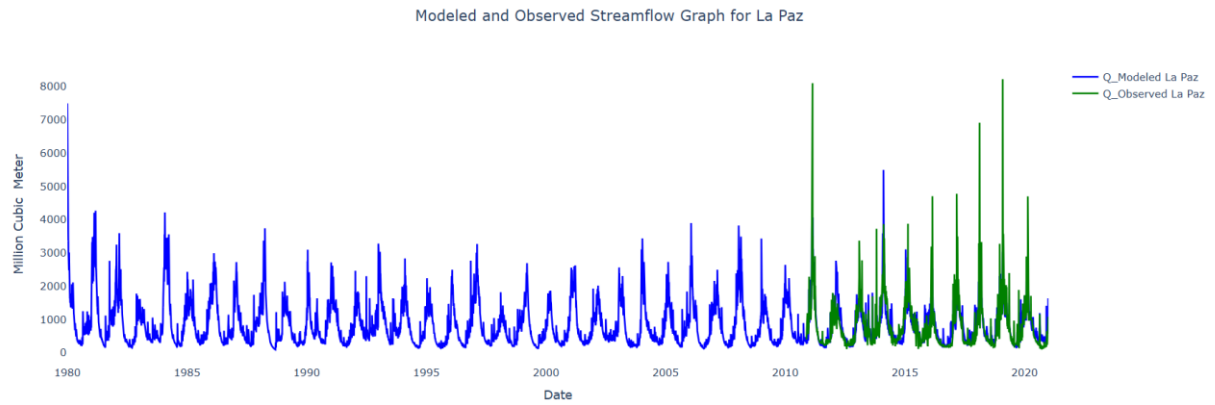


Figure 3: Modeled and observed streamflow at Q_AM_02 hydrometric station.

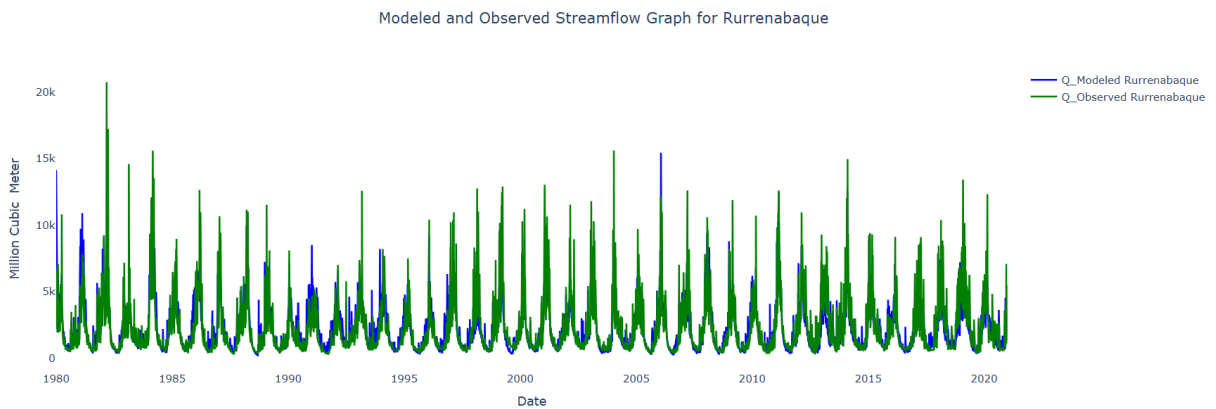


Figure 4: Modeled and observed streamflow at Q_AM_03 hydrometric station.

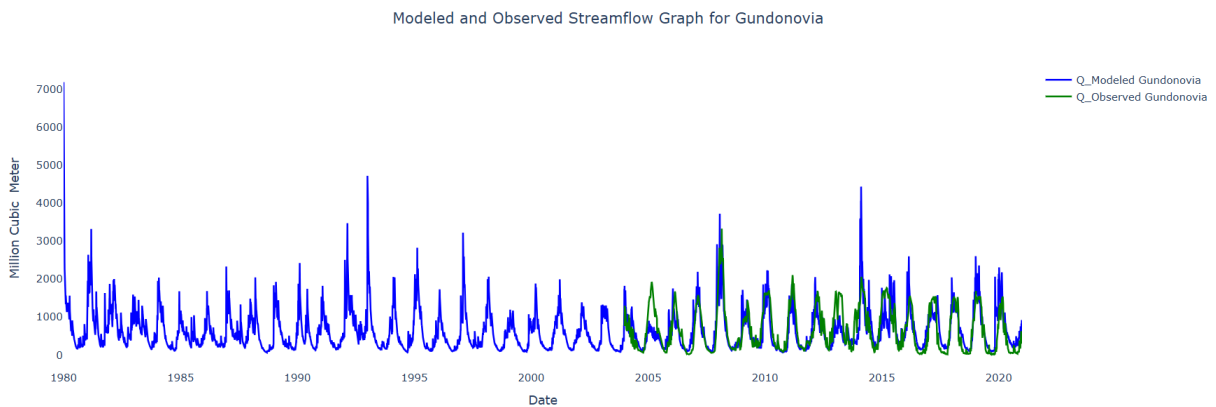


Figure 5: Modeled and observed streamflow at Q_AM_04 hydrometric station.

DISCUSSION

The evaluation of the WEAP hydrological model across the four hydrometric stations in the Q_AM_01, Q_AM_02, Q_AM_03, Q_AM_04 basins demonstrates that the model achieves satisfactory to very good performance, depending on the site and the selected statistical indicator. The Kling–Gupta Efficiency (KGE) and Log-Nash–Sutcliffe Efficiency (Log-NSE) proved to be the most informative metrics, as they integrate correlation, variability, and bias, while also providing sensitivity to low flows, which are critical for water resources management in the Bolivian Amazon. According to the classification criteria of Moriasi et al. (2007), Kling et al. (2012), and Pushpalatha et al. (2012), the model showed very good performance at Q_AM_01 (KGE = 0.85; Log-NSE = 0.73), good performance at Q_AM_02 and Q_AM_03, and satisfactory but acceptable performance at Q_AM_04. This gradient of performance across stations reflects both the spatial heterogeneity of the basins and the challenges of representing local hydrological processes with limited data.

The satisfactory performance of the model, even under data-scarce conditions, confirms the robustness of WEAP as a decision-support tool. Nevertheless, the relatively lower NSE and RSR values at Q_AM_04 indicate that improvements are needed in representing runoff generation and flow routing in this sub-basin. Possible sources of uncertainty include errors in precipitation forcing from the GMET dataset, generalized soil and land-use parameterization, and the absence of explicit groundwater dynamics in the current model configuration. Previous studies (e.g., Beven, 2018; Medici et al., 2019) have emphasized that such structural and data-related uncertainties can limit the predictive accuracy of semi-distributed hydrological models, particularly in regions with strong rainfall seasonality and complex geomorphology such as the Amazonian piedmont.

An important finding is the ability of the model to capture low-flow conditions, as evidenced by the Log-NSE values above 0.70 at three of the four stations. This is particularly relevant for hydropower and water supply planning, where reliable simulation of baseflows is as important as capturing peak discharges. The performance of Log-NSE complements NSE and KGE, confirming that the model not only reproduces mean flow dynamics but also provides confidence in assessing water availability during dry periods. Such capability strengthens the model's applicability for evaluating hydropower potential, ecosystem flow requirements, and drought resilience.

Despite these encouraging results, the model still exhibits limitations in capturing extreme high-flow events, as suggested by the moderate NSE and RSR values. This limitation is common in hydrological models calibrated with daily data (Fowler et al., 2018), where short-term rainfall–runoff dynamics are smoothed out. Incorporating higher-resolution climate inputs, improved representation of soil infiltration, and sub-basin-scale parameter calibration could help address these challenges. Additionally, the integration of automated multi-objective calibration methods (e.g., SUFI-2, DREAM, or Bayesian approaches) would enable more systematic exploration of parameter space, reducing uncertainty and enhancing predictive skill.

In summary, the hydrological modeling of the Rurrenabaque and Gundonovia basins demonstrates that WEAP, when evaluated with complementary statistical metrics (NSE, KGE, PBIAS, RSR, and Log-NSE), provides reliable results for hydrological assessments in data-scarce contexts. While the performance ranges from satisfactory to very good across stations, the model successfully reproduces both seasonal variability and low-flow regimes, offering a sound foundation for hydropower planning, water allocation, and adaptation strategies under

future climate variability scenarios. Continued refinement of input data and calibration strategies will further enhance its applicability as a strategic water management tool in Bolivia.

CONCLUSIONS AND RECOMMENDATIONS

The calibration and validation of the WEAP hydrological model in the Rurrenabaque and Gundonovia basins demonstrate that the model provides satisfactory to very good performance, depending on the station and the evaluation metric. The combined use of NSE, KGE, PBIAS, RSR, and Log-NSE proved essential for a comprehensive assessment, capturing both average flow dynamics and low-flow regimes that are critical for water resources management in the Bolivian Amazon. The model achieved very good performance at Q_AM_01, good performance at Q_AM_02 and Q_AM_03, and satisfactory but acceptable performance at Q_AM_04, reflecting both the robustness of WEAP and the spatial heterogeneity of hydrological processes in the study area.

A key conclusion is that the model effectively reproduces low-flow conditions, as confirmed by Log-NSE values above 0.70 in most stations. This strengthens confidence in its applicability for hydropower planning, drought management, and ecological flow assessments, where reliable representation of baseflows is essential. However, limitations remain in the simulation of extreme high-flow events, highlighting the need for improved climate inputs, refined soil–vegetation parameterization, and more detailed representation of runoff generation processes.

Based on these findings, the following recommendations are proposed:

1. *Data improvement:* Enhance the spatial and temporal resolution of precipitation and climate forcing data, particularly in poorly monitored sub-basins, to reduce input uncertainty.
2. *Parameter refinement:* Implement sub-basin scale calibration of soil and land-use parameters to better capture spatial heterogeneity.
3. *Groundwater dynamics:* Incorporate explicit representation of groundwater processes to improve flow routing and baseflow simulation.
4. *High-flow representation:* Use higher-resolution hydrometeorological data and infiltration parameters to improve the simulation of extreme flow events.
5. *Advanced calibration:* Apply automated multi-objective calibration techniques (e.g., SUFI-2, DREAM, Bayesian approaches) to systematically explore parameter space and reduce model uncertainty.
6. *Future applications:* Extend the modeling framework to assess *climate change impacts, hydropower potential, and water allocation scenarios*, ensuring that future management strategies are based on robust and reliable hydrological simulations.

In conclusion, despite existing limitations, WEAP has proven to be a reliable and practical tool for hydrological assessments in data-scarce regions such as the Bolivian Amazon. With continued refinement of input datasets and calibration strategies, its applicability as a decision-support tool for integrated water resources management and climate adaptation in Bolivia can be further strengthened.

ACKNOWLEDGMENTS

We would like to thank the Royal Institute of Technology in Stockholm KTH for the research stay period that helped enhance the calibration of the model.

Also, to Hydraulic Laboratory of the Universidad Mayor de San Simón (LHUMSS) for providing observed flow data from the hydrometric stations near the study area.

Our gratitude extends to the Deutscher Akademischer Austauschdienst (DAAD) and the University of Bonn for their support of the MIBIO project components through the Universidad Mayor de San Simón and the Universidad Católica Boliviana of Cochabamba.

We also acknowledge the Integral Biodiversity Monitoring Project in Response to the Impact of Hydroelectric Megaprojects (MIBIO) for their efforts in data collection and overall project logistics.

REFERENCES

- Andréassian, V., Perrin, C., Michel, C., Usart-Sanchez, I., & Lavabre, J. (2015). Impact of imperfect rainfall knowledge on the efficiency and the parameters of watershed models. *Water Resources Research*, 37(6), 1757-1770. doi:<https://doi.org/10.1029/2000WR900365>
- Bennett, N. D., Croke, B. F., Guariso, G., Guillaume, J. H. A., Hamilton, S. H., Jakeman, A. J., . . . Newham, L. T. H. (2013). Characterising performance of environmental models. *Environmental Modelling & Software*, 40, 1-20.
- Beven, K. J. (2018). *Rainfall-Runoff Modelling: The Primer (3rd ed.)*. Wiley-Blackwell.
- ENDE. (2024). *ENDE continuará apostando por la generación de energía renovable para el cambio de la matriz energética nacional*. Retrieved from <https://www.ende.bo>
- Fowler, K. J., Peel, M. C., Western, A. W., & Zhang, L. (2018). Improved rainfall-runoff calibration for drying climates using a proxy-catchment approach. *Journal of Hydrology*, 564, 327-341.
- Gupta, H. V., Kling, H., Yilmaz, K. K., & Martinez, G. F. (2009). Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of Hydrology*, 80-91.
- Kling, H., Fuchs, M., & Paulin, M. (2012). Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios. *Journal of Hydrology*, 424-425, 264-277. doi:<https://doi.org/10.1016/j.jhydrol.2012.01.011>
- Kling, H., Gupta, H. V., & Andres, G. T. (2012). On the assessment of uncertainty in hydrological models. *Water Resources Research*, W03509., 48(3). doi:<https://doi.org/10.1029/2011WR011462>
- Knoben, W. J. M., Freer, J. E., & Woods, R. A. (2019). Inherent benchmark or not? Comparing Nash–Sutcliffe and Kling–Gupta efficiency scores. *Hydrology and Earth System Sciences*(23(10)), 4323–4331. doi:<https://doi.org/10.5194/hess-23-4323-2019>
- Krause, P., Boyle, D. P., & Bäse, F. (2005). Comparison of different efficiency criteria for hydrological model assessment. *Advances in Geosciences*, 5, 89–97. doi:<https://doi.org/10.5194/adgeo-5-89-2005>
- McCuen, R. H., Knight, Z., & Cutter, A. G. (2006). Evaluation of the Nash-Sutcliffe efficiency index. *Journal of Hydrologic Engineering*, 11(6), 597-602.
- Medici, G., Abrahão, R., Siqueira, J., & Marco, D. A. (2019). Improvements in hydrological modeling through parameter sensitivity analysis and model calibration. *Water*, 11(6), 1250.

- Ministry of Environment and Water Bolivia. (2018). *Balance Hídrico Superficial de Bolivia (BHSB), periodo 1980–2016*. doi:<https://doi.org/10.5281/zenodo.10850934>
- Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Bingner, R. L., Harmel, R. D., & Veith, T. L. (2015). Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Journal of Hydrologic Engineering*, A4014001(20(3)).
- Moriasi, D., Arnold, J. G., Van Liew, M. W., Bingner, R. L., Harmel, R. D., & Veith, T. L. (2007). Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transactions of the ASABE*, 50(3), 885–900.
- NCAR. (2022). Gridded Meteorological Ensemble Tool (GMET). *National Center for Atmospheric Research*. Retrieved from <https://ral.ucar.edu/projects/gmet>
- Perrin, C., Michel, C., & Andréassian, V. (2001). Does a large number of parameters improve model performance? *Hydrological Processes*(15(1)), 1–15.
- Pushpalatha, R., Perrin, C., Le Moine, N., & Andréassian, V. (2012). Using the performance of hydrological models: a focus on low flows. *Hydrology and Earth System Sciences*, 16(8), 3125–3140.
- Santos, L., Thirel, G., & Perrin, C. (2018). Technical note: Pitfalls in using log-transformed flows within the KGE criterion. *Hydrology and Earth System Sciences*(22), 4583–4591. doi:<https://doi.org/10.5194/hess-22-4583-2018>
- SEI. (2015). *Stockholm Environment Institute, Water Evaluation And Planning, User Guide*. Stockholm Environment Institute.
- SEI. (2022). *Gridded Meteorological Ensemble Tool (GMET)*. Stockholm Environment Institute. Retrieved from <https://www.sei.org/projects/gridded-meteorological-ensemble-tool-gmet/>
- Sieber, J., Purkey, D., & Huber-Lee, A. (2005). WEAP21—A demand-, priority-, and preference-driven water planning model: Part 1: Model characteristics. *Water International*, 30(4), 487–500.
- U.S. Geological Survey. (2021). Evaluation of Predictive Performance and Variable Importance—Machine Learning Approach to Modeling Streamflow with Sparse Data in Ungaged Watersheds on the Wyoming Range (USGS Scientific Investigations Report 2021-5093). *U.S. Department of the Interior*. doi:<https://doi.org/10.3133/sir20215093>
- Zárate, S., Villazón, M. F., Balderrama, S., & Quoilin, S. (2021). Modeling hydropower to assess its contribution to flexibility services in the Bolivian power system. In Conference on Sustainable Development of Energy. *Water and Environment Systems (SDEWES)*.