Estimation of River Discharge from Satellite Observations

1Dhanendra Bahekar, 2Isthiyaq Ahmad, 3A. D. Prasad

1PG Scholar, Department of Civil Engineering, National Institute of Technology, Raipur, Chhattisgarh, India
Email - dhanendrabhekar1996@gmail.com
2Assistant Professor, Department of Civil Engineering, National Institute of Technology, Raipur, Chhattisgarh, India
Email - iahmad.ce@nitrr.ac.in
3Assistant Professor, Department of Civil Engineering, National Institute of Technology, Raipur, Chhattisgarh, India
Email - adprasadiiit@gmail.com

Article History: Received:11 January 2021; Accepted: 27 February 2021; Published online: 5 April 2021

Abstract – The role of streamflow is very important in any type of hydrologic. For very effective flood routing and hydraulic structure design, it is important to have a large dataset of past years. We now have a conceptual rainfall-runoff model that can predict streamflow based on pre-existing datasets. Because there is no or very little observed data in un-gauged basins, calibrating these models to predict daily streamflow becomes difficult. Nowadays, parameters for example river width can be observed using satellite images, and some studies show a promising associated relation between discharge and river width. The suggested study demonstrates a method for calculating streamflow from river width extracted with the help of satellite imagery. To predict streamflow, hydrological models are calibrated using river width instead of in situ observed streamflow, and for estimating uncertainty Generalized Likelihood Uncertainty Estimation (GLUE) is used. For validation, the suggested method is implemented in the Kharun river basin situated in the Chhattisgarh state of India. The obtained Nash-Sutcliffe efficiency is 92.6 % for simulated river discharge in 2019-2020 at the 50% quantile, which is promising.

Keywords – Hydrological model, Discharge estimation, RivWidth, Un-gauged basin, GLUE

I. INTRODUCTION

Streamflow (river discharge) is a key parameter there in the hydrological cycle. Previous information of river flows is needed for any hydraulic analysis or hydrometeorology. Prehistoric river discharge datasets are needed for flood routing, hydraulic structure design, and forecasting. Historically, Gauge-Discharge (GD) stations have been utilized in order to measure streamflow in major rivers. There are a lot of ungauged basins all over the world. In India, for example, the Garra and Kulsi river basins are un-gauged. It is difficult for hydrologic modelers to simulate flooding in un-gauged catchments [1].

The use of remote sensing and GIS software can be utilized to calculate streamflow in any un-gauged basin. The previous study has shown that there is a close association between stage and discharge as well as width and discharge. Since river stream width can be derived from satellite data very easily, therefore, the width-discharge relationship is favored for prediction purposes. The following paragraph addresses several research on the subject.

Some research on calculating river flow from satellite data has been undertaken in the past. These researches are classified as hydrological model-based and direct equation-based. The strong linear relationship between flow velocity - discharge, stage - discharge and width – discharge exists and was defined by Leopold et al (1953) [2]. This research by Leopold et al. (1953) was updated and improved by several authors in different study areas as an at-a-station relationship. Ashmore P. et al (2005) [3] suggested a somewhat different approach for calculating running river surface water width since he used ground-based photographs and acquired orthophotos of river site and stage datasets. They found the average $R^2$ value for a linear function 0.70, for a power function 0.76, and for a polynomial function 0.80 comparing width and discharge. In place of linear regression as done by Leopold et al. (1953) [2], multiple regression analyses were utilized by Bjerklie et al. (2007) [4,5]. He developed five different models to determine the streamflow and width relationship, and from those 5 models, the most reliable model was $Q = 7.22 W^{1.02} Y^{1.74} S^{0.35}$. In this equation $Y$ depicts the depth of the river and ‘S’ depicts the slope of the river. The research was undertaken in 108 separate rivers of various forms, such as large rivers, narrow rivers, braided rivers, and full-width rivers. Similarly, some researchers utilized bed material characteristics to model streamflow [6]. Irrespective of the direct relationship, some researchers indicated a quantile approach, while some authors also recommend a regionalization approach [7, 8, 9].
There are several significant barriers to streamflow estimation in an ungauged basin that must be addressed. The first problem is parameterizing width-discharge relationships. These parameters must be carefully parameterized since they represent cross-sectional details and hydrological conditions of corresponding river stations. Rainfall-runoff models are appropriate for parameterization since they are dependent on mathematical and physical relationships. For promising results, the models should be calibrated. Second, there are questions about satellite imagery. These two concerns are discussed using the Generalized Probability Uncertainty Calculation (GLUE) method [10, 11]. Calibration of the model by GLUE is conducted using the width of the river water surface derived from satellite observation, as well as uncertainty estimate for both width extraction and discharge prediction. Streamflow prediction for any river basin can be performed using the calibrated model.

The rest portion of this paper is divided into the following sections: Section II describes the proposed methodology, HYMOD, and GLUE. Section III includes a case study in which the suggested methodology is performed. Section IV discusses the case study's findings and results. Section V describes the discussion of the study's potential reach in the future. Finally, in section VI, the study's conclusion is addressed.

II. PROPOSED METHODOLOGY

2.1 Modified objective function for calibrating Hydrological Model –

Since this article focuses on the methods that would be used to derive river discharge from satellite measurements, such as water surface width, the objective function for the the general rainfall-runoff model will be modified to take water surface width into account.

Any hydrological model is developed with the help of a basic rainfall-runoff relationship that can simulate streamflow for the respective river basin. General relation describing a hydrological model is as follows –

\[
Q = f \left( \frac{I}{\eta} \right) \tag{1}
\]

Where I depict precipitation, a forced model input parameter. Q depicts model output streamflow in m³/sec, η depicts model parameter vector.

In order to consider satellite measurements for model calibration, a power equation comparing stream flow and river width was established on the basis of the hydraulic relationship, which is represented as

\[
W = g \left( \frac{Q}{\lambda} \right) \tag{2}
\]

In this equation ‘W’ denotes the river stream width, the parameter vector represented by ‘λ’, and ‘g’ denotes the function relating hydraulic conditions. After coupling the two models above and utilizing equation (1) output river discharge as the input in equation (2). The following will have an interactive model as:

\[
W = F \left( \frac{I}{\theta} \right) \tag{3}
\]

In this equation, the terms W and I are the river width and rainfall respectively mentioned in equations (1) and (2), and θ depicts the model parameters vector which comprises every parameter of η and λ. A combined function is depicted by F, which is developed by incorporating the above given 2 models (eq. 1 and 2).

By elements tuning of θ with the aim of minimization of the difference between satellite observed river width and HYMOD simulated river width was used to calibrate this integrated model. “In this process, it is presumed that the calibrated model parameters represent the hydrological characteristics of the river basin as well as the situation at the river outlet” [12].

Leopold et al (1953) [2] investigated the efficiency of station cross-section dependent hydraulic geometry, which describes the relationship between river top-water surface width and streamflow as a basic power function; this equation can be used as a hydraulic model, such as:

\[
W = a Q^b \tag{4}
\]
The terms ‘a’ and ‘b’ represent mathematical equation parameters, while the term ‘W’ represents river depth. The parameter vector $\lambda$ takes into account two parameters, ‘a’ and ‘b,’ these two are calibrated as $\lambda$.

2.2. Concept of HYMOD Hydrological Model

Boyle D. P. (2001) [11] created a Hydrological Model abbreviated as HYMOD, a daily times series based model that is a type of parsimonious model that takes into account conceptual elements of the hydrological model. Sun W. et al (2015) [12-15] has successfully applied HYMOD model in various basins, including the Irrawaddy basin in Myanmar. The basic theory behind HYMOD is runoff generation from excess infiltration. In a proposed study by Moore (1985) [16] gave the principle of probability-distribution, and stream flow generation with the help of this model is explained using this principle. The function given below depicts spatial variability of the basin as:

$$F(C) = 1 - (1 - \frac{C}{C_{\text{max}}})^{B_{\text{exp}}}$$

(5)

In this equation,
F - Storage capacity at any point of the river basin in cumulative order
C - Storage capacity of water;
$C_{\text{max}}$ - The river basin’s maximum storage capacity of water (mm)
$B_{\text{exp}}$ - Within the basin the degree of spatial variability of soil moisture.

In the hydrological model HYMOD, streamflow is conceptually represented with the help of three quick-flow tanks, while groundwater is represented with the help of a single slow-release reservoir parallel to the quick flow tanks.
It is known that spatial variability exists in every river system, so the basin is split into sub-basins to account for such variability. Cmax, Bexp, and Alpha (α) are 3 identical parameters in HYMOD that account for runoff generation. Kq and Ks are 2 routing parameters. Table 1 and Figure 1 provide summary detail about HYMOD and its parameters.

<table>
<thead>
<tr>
<th>Indication of HYMOD parameters</th>
<th>Unit</th>
<th>Variation of parameter value (Range)</th>
<th>Illustration of parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>C&lt;sub&gt;max&lt;/sub&gt;</td>
<td>millimeters</td>
<td>1–100</td>
<td>Basin’s Maximum water storage capacity</td>
</tr>
<tr>
<td>B&lt;sub&gt;exp&lt;/sub&gt;</td>
<td>..</td>
<td>0–2</td>
<td>Within the basin the degree of spatial variability of soil moisture.</td>
</tr>
<tr>
<td>α</td>
<td>..</td>
<td>0.2–0.99</td>
<td>Factor distributing the flow between slow and quick release reservoirs</td>
</tr>
<tr>
<td>K&lt;sub&gt;q&lt;/sub&gt;</td>
<td>In days</td>
<td>0.5–1.2</td>
<td>Slow release reservoir’s residence time</td>
</tr>
<tr>
<td>K&lt;sub&gt;s&lt;/sub&gt;</td>
<td>In days</td>
<td>0.01–0.5</td>
<td>Quick release reservoir’s residence time</td>
</tr>
</tbody>
</table>

2.3. GLUE (Generalized Likelihood Uncertainty Estimation) –

In order to estimate variability and calibrating hydrological models, GLUE proves to be a good tool. It's a form of Monte Carlo method. Its working aim is to create a behavioral model, or a collection of parameters, from a range of parameter set combinations [17]. The term "behavioral model" refers to an acceptable model as assessed by historical data and prior information. In order to perform GLUE, initially, parameter distribution was selected at random, and a huge number of model parameter combinations from that distribution were used for running the model for obtaining behavioral sets based on a probability threshold value that denotes the goodness of fit of the respective parameter combination that is simulating the system. As the likelihood function value increases, it indicates that the predicted discharge best correlates to the observed discharge [18, 19]. “The threshold value which is used to separate behavioral parameter as well as non-behavioral parameter combinations can be obtained either as a permissible deviation from the maximum likelihood value in the sample or in the form of a fixed percentage of the total simulations count” [12, 13]. Output estimation is derived using these probability values Cumulative Distribution Function (CDF). With CDF, variability in simulation is calculated at 50 percent and 95 percent confidence levels, as recommended by authors in several GLUE-based research.

When dealing with GLUE, a collection of parameter sets are developed; to identify the wellness for simulation, a Likelihood L[θ] is given to those sets. In the context of L[θ], any indices which can define the model's performance can be used. The threshold value for L[θ] is fixed for achieving behavioral parameter sets (suitable for successful simulation). It is therefore guaranteed that when the rescaling is completed parameter sets that are behavioral, their sum of L[θ] will be equal to unity. The likelihood can be given as:

\[ L_p[θ] = C * L[θ] \]  \hspace{2cm} (6)

In equation

- L[θ] - The parameter sets likelihood value.
- L_p[θ] - The likelihood value after rescaling.
- C - Rescaling constant.

Then, calculation of uncertainty quantile is done by simulating the cumulative distribution at time series in step ‘t’.

\[ P_t(W_t < w) = \sum_{i=1}^{m} L_p[θ_i | W_t < w] \]  \hspace{2cm} (7)

In this equation -
- P_t(W_t < w) – It represents cumulative probability of river width Wt less than ‘w’
Lp[θi] – represents parameter set θ likelihood value after rescaling
m - number of those parameter sets which are satisfying the given condition Wt < w.

In same way, river discharge simulation is done by using the similar cumulative distribution and set of parameters. Many experiments have shown that in the case of streamflow simulation, 50 percent quantile simulation produces positive results [12, 13].

III. A CASE STUDY: ON KHARUN RIVER (PATHARIDIH STATION)

3.1 Study Area –
This method is applied to a 5-kilometer stretch of the Kharun River, first point from the ‘PATHARDIH’ G-D station to end point Bana. The investigation site is located in Raipur city in Chhattisgarh state of INDIA. The selected Kharun River contributes its flow to the Seonath River. Seonath is a major tributary of Mahanadi basin. The Kharun river originates in Chhattisgarh province, in Petechua village in tehsil Balod, in the district of Durg. The Kharun River stretches for around 164 kilometres. The Kharun River joins Seonath near the place named Somnath. The Kharun River Basin has a catchment area of approximately 4191 km². Kharun is positioned between 20° 33’ 30” and 21° 33’ 38” N latitude and 81° 17’ 51” and 81° 55’ 25” E longitude in terms of its geographical coordinates.” Modak S. (2012) [20]. The Bay of Bengal monsoon From July to September brings heavy rain in the Kharun river basin.

![Kharun river basin map](image)

Figure 3. Kharun river basin map

▲ represents Patharidih station.
Patharidih G-D station is located at latitude 21° 20' 26.88" N and longitude 81° 35' 42" E. Fig. - 3 depicts location of the Patharidih station within Kharun basin. Central Water Commission INDIA monitors and maintains this GD station.

3.2 Extracting width of Kharun river at Patharidih using LANDSAT-8 images –

Satellite imageries are extremely helpful in order to obtain data from any location, as Landsat 1 is one of the old satellites that assisted scientists in obtaining massive amounts of data from 1970-1990. LANDSAT 8 images with a resolution of 30 m are being used for this study. The launching of Landsat 8 has been done on 11 February 2013. It is equipped with the Thermal Infrared Sensor (TIRS) and the Operational Land Imager (OLI). Landsat 8 images of the Patharidih site were obtained for the calibration period of 2014-2018. For every year, three images are downloaded: one for the pre-monsoon season, one for the monsoon season, and one for the post-monsoon season. A total of fifteen images having very less cloud cover around 10%, were obtained for extracting river width from the USGS website (https://earthexplorer.usgs.gov/).

The RivWidth script in ENVI-IDL (Environment for Visualizing Images – Interactive Data Language) was used to extract river width. Pavelsky introduced RivWidth in 2008 (Pavelsky, 2008) [21]. For extracting river water width from satellite imageries numerous researchers recommended RivWidth in place of the manual measurement by dividing river surface area by length of considered river portion. To reduce error, the mean river water surface width for a given stretch is measured against the water surface width at each station. The channel mask (fig. 4 b) and river mask (fig. 4 c) prepared of every Landsat imagery are used as input for the ENVI-IDL based RivWidth algorithm.

![Figure 4. a. 5 Km stretch of Kharun River](image1)

![Figure 4. b. Input Channel Mask](image2)

![Figure 4. c. Input River Mask](image3)

From studies it is well established that only fifteen satellite imagery data are insufficient for stream flow thus, daily mean rainfall record evidence was used to provide help for the inference. Precipitation data for the Patharidih area is collected from the WRIS stands for Water Resources Information System developed jointly by Central Water Commission India and ISRO. Because of the spatial variation in precipitation, it is evident that the Patharidih site has distinct rainy and dry seasons. According to Sun W. (2015) [12], "Based on the experiences at the Mekong basin in Pukse, we can conclude that at least 2 satellite images, one during the rainy season and another during the dry season, are adequate."

**Table -2 Satellite image extracted effective width of river**
3.3 Setting up model and calibration with GLUE –

Data sets from 2014 to 2018 were used to calibrate 5 HYMOD model parameters and 2 width-discharge equation parameters. In HYMOD, two forced input parameters are daily average rainfall and potential evapotranspiration. WRIS system was used to obtain datasets for years 2014 to 2018. (https://indiawris.gov.in/wris/#/RiverMonitoring). Previously parameter ranges must be set in needed to execute GLUE, since a significant number of combination of parameter sets would be created in GLUE through applying Latin Hypercube Sampling method. Previous research information is used in Latin Hypercube Sampling to generate parameter sets, for parameters distribution, and determine parameter range. When the parameter distribution is unknown because of insufficient data, a basic uniform distribution is utilized. Refer above given Table 1 for model parameter range which is suggested by previous researchers. Width-Discharge equation parameter ‘a’ varies from 100-1600 and ‘b’ varies from 0.03-0.17 in our study.

For defining a range, the width-discharge equation parameters ‘a’ and ‘b’ seem to be more complicated. According to previous research in the equation $W = a*Q^b$, the defined parameter ‘a’ is affected by the shape of the cross-section of river, size of the river, the slope of energy line of the river, and hydraulic conductance. The value of equation parameter ‘a’ ranges theoretically from zero to infinity, however it’s not practically possible; according to Dingman (2007) [23], the value of parameter an is less than 80% of the bankfull river width. The shape of the cross-section determines parameter b. The b value for an approximately rectangle-shaped cross-section is 0. According to research, b varies from 0.015 - 0.28. (W Sun, 2015).

The next step prior to running GLUE is to define a threshold likelihood ($L[θ]$) value for differentiating between behavioral parameter sets as well as non-behavioral parameter sets. For a smaller set of observations (like 15, in this study), we used the Relative Mean Square Error’s inverse, abbreviated RMSE, to calculate $L[θ]$ invers RMSE is give by –

$$L[θ] = \frac{1}{\sqrt{\left(\frac{1}{n}\right) * \Sigma(W_{obs,t} - W_{sim,t})^2}}$$

(8)
In equation (8) the term $W_{\text{obs},i}$ refers to ‘i’ th satellite imagery based observed width and $W_{\text{sim},i}$ refers to ‘i’ th model simulated width. By applying trial and error method we obtained threshold value as 0.0075. From generated 50,000 sets, we obtained 1134 behavioral parameter sets. The stream flow simulation was then performed using 1134 behavioral parameter sets. “Similarly, for estimation of stream discharge, 50 percent quantile is best suited for uncertainty estimation.” [13] (W Sun, 2010a).

IV. RESULTS

Upon completion of parameter calibration, the final width discharge power equation $W = 125.31 * Q^{0.0382}$ with $R^2 = 0.9158$ was obtained. A high $R^2$ value indicates that a and b are well calibrated.

After calibration streamflow simulation for years 2014-2018 was done using a 50 percent quantile uncertainty band. (Refer Figure 5).

![Figure 5. Rainfall and corresponding simulated discharge at 50 percent quantile for years 2014-2018](image)

![Figure 6. Discharge simulation for years 2019-2020 at 50 percent quantile](image)
To validate the developed approach, we collected river discharge data for the Patharidih site with the help of WRIS for the years 2019-2020, and we ran the model over the same time period of 2019-2020 with the obtained 1134 behavioral parameter sets. Figure 6 and 7 depicts a daily time series simulation of river discharge for 2019-2020 and simulation uncertainty limit band for 2019-2020 respectively. The monthly mean observed discharge was compared to the mean monthly mean simulated discharge at the 5 percent, 50 percent, and 95 percent quantiles (Refer figure 8). The Nash-Sutcliffe Efficiency for simulated monthly river discharge at the 50 percent quantile was 92.6 %. The model performed admirably during the high flow season but not so well during the low flow season. The variation in hydrological characteristics between the dry and monsoon seasons, as well as the low streamflow forecast during the summer season, resulted in slightly poor model performance mostly during summer (dry) season.

Figure 8. Monthly average simulated river discharge comparison with observed discharge between years 2019-2020

V. Discussion: Future scope of methodology
The technique for estimating river discharge with the help of satellite data is described here; this technique can be used for gauged and un-gauged basins during unfavorable circumstances arise, such as a GD station failure.

The width of a river's water surface derived from satellite imagery is an important parameter since it is easily retrieved. To improve accuracy, the correct width-discharge relationship for the research region must be established after proper calibration. As extraction of river top water width obtained through satellite imagery is easy and it also increase accuracy. After complete calibration, the width-discharge equation for the study area will be developed. Since there would be no substantial difference in water-surface width in case of rectangular river cross section, this approach can’t be used when the river is rectangular. The satellite imagery spatial resolution and the calibration scheme assess the precision of this technique. When fine resolution images for Eg. Quickbird (0.61 m) are available, more accurate width extraction will be possible, yielding more promising width-streamflow relationship.

Another factor of consideration is the river width measurement process from satellite images. In case of fine-resolution images, more accurate width measurement can be achieved, but for coarse-resolution imageries such as Landsat (30 m), the RivWidth algorithm defined by Pavelsky (2008) can be used to minimize error, and we utilized that in this case study.

Yang (2019) [22] used a new process for extracting river width from Google Earth Engine using RivWidthCloud algorithm. GLUE is a very effective technique for calibration and uncertainty calculation that has been defined by a number of researchers.

For calibration and validation of presented technique in any un-gauged basin, minimum of one year of observed regular streamflow at the site is needed. Just a few satellite measurements were used in previous studies [12], such as W Sun, (2015) utilized six observations only. In this analysis, three satellite image observation were selected for every year, totaling fifteen records, resulting in more precise width discharge relationships. If there are further images available, they must be taken into account in the analysis.

For the correctly calibrated model, this approach may lead to real-time streamflow observation in the future. There is a fantastic tool available. Google Earth Engine (GEE) supports both Java and Python scripting languages. It is simple to import satellite images from GEE, and direct extraction of river width is possible in seconds using RivWidthCloud [22]. Real-time streamflow tracking would be feasible if we can build a more accurate model with less uncertainties. In future fine-resolution satellite images will become available over time, and if anyone have enough resources, must utilize those imageries to extract width.

VI. CONCLUSION

With the value of streamflow in mind, this suggested approach will reliably forecast discharge in both gauged and un-gauged basins. To begin forecasting, some specific information about general cross section, minimum of 2 satellite image derived river width (each for dry and monsoon season) for every year and minimum of 1 year observed river discharge dataset is necessary.

With these datasets, GLUE's automated calibration technique successfully overcomes uncertainty in the streamflow simulation after calibration for years 2014-2018, yielding an NSE of 92.6 percent for simulated discharge for 2-year 2019-2020 estimation. As seen in this case study, three photos are taken per year for the not-so-wide river Kharun, and calibration was completed is an accurate approximation.

This technology was originally designed for large river basins such as the Pyay, Irrawaddy, and others, but this is the first time it has been applied to a small river basin, the Kharun. The RivWidth algorithm was used to solve the problem of extracting width from low-resolution imagery in this case. Just one parameter set, ‘a’, was influenced by lower width values.

The above-mentioned methodology can currently be used to predict stream discharge in a flow, but with some tuning, we can use it in combination with a genetic algorithm to minimize uncertainty and get closer to real-time river discharge prediction. With advancements in resolution of satellite images and satellite sensor, this approach will provide more precise and consistent data.

This approach will often yield more precise and consistent data, potentially leading to real-time streamflow estimation with RivWidthCloud.
REFERENCES


