Applying gridded rainfall data to calculate inflow discharge into Ban Chat hydropower reservoir basin

Chinh Kien Nguyen  
Hydrodynamic and Disaster Mitigation  
Institute of Mechanics - VAST  
Hanoi, Vietnam  
nckien@imech.vast.vn

Thanh Huong Duong Thi  
Hydrodynamic and Disaster Mitigation  
Institute of Mechanics - VAST  
Hanoi, Vietnam  
dthuong imech@gmail.com

Hang Nguyen Thi  
Hydrodynamic and Disaster Mitigation  
Institute of Mechanics - VAST  
Hanoi, Vietnam  
hangnguyen imech@gmail.com

Thanh Hang Do  
National Centre for  
Hydrometeorological Forecasting  
Hanoi, Vietnam  
dothanhhang234@gmail.com

Tuan Dung Nguyen  
University of Engineering and Technology - VNU  
Hanoi, Vietnam  
tuandungvnu206@gmail.com

Abstract— Because of complex terrain, especially in the northern mountainous region, it is very difficult to collect rain data as input for hydrological models to calculate flow. The monitoring network system is small and unevenly distributed, the data frequency is not high, and the locations of monitoring stations do not fully reflect the basin’s rainfall regime. Therefore, other rainfall data sources are being sought to supplement and combine with measured rain to improve these limitations. This article presents the results of using gridded rainfall (calculated from the radar reflectivity and the product of the global numerical model) as input to the HEC-HMS hydrological model to calculate discharge into Ban Chat hydropower reservoir basin. The calculation results show the good ability of using gridded rainfall for hydrological modeling to calculate runoff in this basin as well as other basins.

Keywords— Gridded rainfall, radar reflectivity, HEC-HMS.

I. INTRODUCTION

Rainfall data are important input data in streamflow simulation. With limited funding and complex terrain, the rain monitoring network across the country is too small and uneven, and the locations of monitoring stations do not fully reflect the rain regime across the basin. Currently, in the world and Vietnam, there have been many studies on the use of gridded rain in computational hydrological models for river basins without or lacking ground measuring stations. Gridded rain data are calculated from remote sensing images or from numerical models with the advantages of wide and uniform distribution in space, continuous frequency over time, and have been combined with actual measured rain data to as input for computational models.

In the world, there are many researchers on gridded rainfall data, such as Andrew D. King [1], Kamal Ahmed [2] have provided comparisons and analyzes of using gridded rainfall from satellite images with rainfall observed from measuring stations. The resulting gridded datasets can be used to investigate trends and changes in extreme rainfall once examined. C. R. Tozer [3] did statistical analysis with three monthly rainfall datasets in Australia. The results show clearly different runoff associated with each different rainfall data source.

In Vietnam, rainfall forecasts from satellite images and radar images combined with rainfall measured at meteorological stations have also been researched by many authors in flood forecasting with many topics, projects as Nguyen Van Thang [9], Thai Be Van [10]. Recently, Pham Thanh Hung [6], Nguyen Thi Lien [8] used gridded rain data and provided quality assessments by comparing with a number of heavy rains at meteorological stations.

In this article, the authors use gridded rainfall data to calculate the inflow discharge into Ban Chat hydropower reservoir basin.

II. RESEARCH METHODS

In this study, two gridded rain data are used: gridded rain calculated from radar reflectivity and rain extracted from the US global meteorological model (GFS).

A. Gridded rain data from radar returns

In Vietnam, there are 10 weather radar stations spread across the country, providing data every 10 minutes. Radar data was collected in 2021-2022 from the National Hydrometeorological Network Center, Ministry of Natural Resources and Environment [13].

From the collected data, the authors used two methods to calculate gridded rain from radar reflectivity, which are the empirical formula (EF) and artificial neural network (ANN).
Chinh Kien Nguyen et al.

1. Empirical formula

The most widely used method of calculating rain by radar is based on the relationship between radar reflectivity and rainfall rate using a formula of the Marshall-Palmer form: \( Z = a \cdot R^b \).

Change to form \( R = c \cdot 10^{dZ'} \) (*)

\[
Z = 10^{7b} = a \cdot R^b \rightarrow R = \frac{1}{\sqrt{a}} \cdot 10^{0.7b} \\
\rightarrow R = c \cdot 10^{dZ'} \text{ with } c = \frac{1}{\sqrt{a}}; d = \frac{1}{10. b}
\]

In which:

- \( Z \) is the radio reflectivity, in which \( Z = 10^{Z/10} \) (\( Z \) in \( \text{dbZ} \) unit);
- \( R \) is rainfall rate;
- \( a, b, c, d \) are experimental coefficients.

To detect these experimental coefficients, the authors used the Shuffled Complex Evolution (SCE) method [5]. The SCE method is initialized by choosing the parameters \( p \) and \( m \) where \( p \) is the number of complexes, \( m \) is the number of points in each complex. Sample space \( s \) is a random sample in the feasible space of parameters using a uniform probability distribution and calculating the objective function value at each point. Then, the points in \( s \) are arranged in order of increasing value of the objective function. These points will be divided into \( p \) complexes, each complex consisting of \( m \) points. Each complex will evolve independently according to the Downhill Simplex method. The next step, shuffling, combines the points in the developed complexes into a new sample set based on the information of the original sample set. This development and disturbance will be repeated until the convergence criteria are satisfied.

2. Artificial neural network method

An artificial neural network is built from basic components that are artificial neurons with many inputs and one output “Fig. 2”. Each neuron simulates a biological neuron, including an activation threshold (bias) and an activation function (or transfer function), characterizing the properties of the neuron [7].

\[
\text{Fig. 2. The structure of an artificial neuron}
\]

In which:

- \( x_i \) (\( i = 1, m \)): The set of input signals of the neuron;
- \( w_{jk} \): Set of links, each link between the \( j \)th input signal and \( k \) neuron is represented by a weight, randomly initialized at the time of network initialization and continuously updated during the learning process;
- \( \sum \): The totalizer is used to calculate the sum of the product of the inputs with their associated weights.
- \( b_k \): Bias, included as a component of the transfer function;
- \( \phi(.) \): Transfer function is also called Activation function. Input data of this function are the result of the totalizer and the given bias;
- \( y_k \): The output signal of a neuron, with each neuron having at most one output.

Although each individual neuron can perform certain information processing functions, the power of neural computing is largely achieved by combining neurons in a unified architecture. “Fig. 3” shows the most widely used multi-layer feedforward network in ANN models.

\[
\text{Fig. 3. Straight propagation neural network}
\]

The advantage of an artificial neural network is that it allows building a computational model with a very high ability to learn from data. The neural network training process is based on the regression error between the calculated value and the actual measured value. The training algorithm will adjust the connection weights of the neural network to minimize the regression error on the training samples. After the network is successfully trained, the weight matrix will be updated for use in the forecasting process.
Applying gridded rain data to calculate discharge flow of Ban Chat hydropower reservoir basin

B. Rain from GFS model (Global Forecasting System)

The Global Forecast System (GFS) is a weather forecast model from the US National Centers for Environmental Prediction (NCEP), generating data for dozens of atmospheric and land-land variables, including temperature, wind, precipitation, soil moisture and atmospheric ozone concentration with a horizontal resolution of 27 km between gridded points. Interim resolution includes analysis and forecasts for up to 16 days. The system incorporates four separate models (atmosphere, ocean, land/land and sea ice model) that work together to accurately describe weather conditions [17].

![Rain from GFS model](image)

For outside the US, the model results have a resolution of 27 km x 27 km, which the authors use as input for hydrological models.

III. SETTING THE MODEL

A. HEC-HMS model

HEC-HMS model [11, 12] (Hydrologic Engineering Center - Hydrologic Modeling System): is a rain-runoff hydrological model of the American Association of Military Engineers. The model was built to simulate the flood process from rain of the basin system based on geographic information system (GIS). The model converts the rain process into flow in each basin, then calculates it in natural rivers, etc. The model uses average parameters in time and space to simulate the flow process. Depending on the geographical and topographical characteristics of each sub-basin, rainfall data, and underground flow to apply appropriate loss calculation and calculation methods.

![Model diagram](image)

We can represent the model with the following diagram:

\[ Y = X - Z \]

B. Data collection

Ban Chat hydropower reservoir basin is located in the Nam Mu river basin in Lai Chau province in Northern Vietnam. Near Vietnam-China border and abuts Dien Bien, Son La and Lao Cai province. The reservoir has an area of 60.5 km² and a capacity of 21,377 billion m³, stretching over 8 communes of Than Uyen district and 4 communes of Tan Uyen district, Lai Chau province. The total basin area is 2,050 km², its altitudes from 340 m to 3,100 m above sea level. Nearly 70% of the basin area is below 700 m altitude. The study area has a typical tropical climate with two distinct seasons: the rainy season lasts from May to September, and the dry season lasts from November to March of the following year.

![Map of the research basin](image)

Input data of the model includes:

**Terrain data:**
- Digital elevation model DEM taken from JAXA’s Global ALOS 3D World, the Japanese aerospace exploration agency with a resolution of 30 x 30 m in 2022 [15];
- Land cover and land use map has a resolution of 10 x 10 m with 20 detailed layers, a product of cooperation between ESRI and Microsoft in 2020 [16];
- Soil map with 250 m x 250 m resolution taken from the United States Department of Agriculture (USDA) in 2020.

**Meteorological, hydrological and radar data:**
- Measured rain data in the Ban Chat hydropower reservoir basin in 2021 and 2022 at every 6 hours: Mu Cang Chai, Tam Duong, Than Uyen and Ta Gia stations from the data source of the Hydrometeorological Forecasting Center National, General Department of Hydrometeorology;
- The inflow of Ban Chat hydropower reservoir basin in 2021 and 2022 is from the online hydroelectric reservoirs operating system of Vietnam Electricity Group - EVN [14];
- Radar reflectivity data is from the free website of the High Air Meteorological Station - Ministry of Natural Resources and Environment [13].
The product is extracted from the GFS global meteorological model with a resolution of 27 km x 27 km [17].

C. Setup model for the Ban Chat hydropower reservoir basin

Based on the digital elevation model map DEM, the Ban Chat hydrological basin is divided into 7 sub-basins.

In this study, the authors used the Clark synthetic unit process path method [9] to calculate the flow, the Green and Ampt seepage method and used the Lag method to calculate floods in open channels.

Using rain options in the HEC-HMS model.

- Rain station: Specified Station (SS) and Inverse Distance (ID)
- The 500 m x 500 m gridded rain is calculated from the product of the GFS global model and from the radar returns (ANN and EF).

IV. RESULTS AND DISCUSSIONS

In this article, the authors chose two floods in 2021 and 2022 to calculate and determine the set of parameters to simulate the flow of Ban Chat hydropower reservoir basin.

<table>
<thead>
<tr>
<th>Plan</th>
<th>Start</th>
<th>End</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calibration</td>
<td>07/8/2021</td>
<td>22/8/2021</td>
</tr>
<tr>
<td>Verification</td>
<td>05/6/2022</td>
<td>10/6/2022</td>
</tr>
</tbody>
</table>

A. Calibration

The 2021 calculation results of HEC-HMS model are calibrated, input is measured rain data at the stations and gridded rain products. Fig. 8 describes the results between the calculation of the Specified Station method (SS), the Inverse Distance ratio (ID) and the measured inflow discharge of Ban Chat hydropower reservoir basin. Although the calculated peak discharge is lower than the measured value, the SS option has NSE = 0.68 and the ID option has NSE = 0.69. The results are good and show the model’s well ability to simulate the flow in the calibration.

The calibration results using gridded rain show that the ANN and EF methods are better than rain station. Only the gridded rain from the GFS source is not satisfactory and only be used for qualitative reference in extreme weather situations.
B. Verification

The set of calibrated parameters is used in the verification with measured rain data at the stations and gridded rain products in 2022. Fig. 10 shows the calculated inflows of Ban Chat hydropower reservoir basin in SS, ID rain station options and measured rain data. In which the NSE index of the SS is 0.6 and ID is 0.46. These indexes are at the average level.

![Verification results by SS and ID methods in rain station option.](image1)

“Fig. 11” describes the calculated inflows of Ban Chat hydropower reservoir basin in ANN, EF, GFS gridded rain options with measured rain data. The results show that ANN option is the best (NSE = 0.67) followed by EF option (NSE = 0.64) but GFS method is not qualified (NSE = 0.16).

![Verification results by ANN, EF and GFS methods in gridded rain option.](image2)

---

### TABLE 2. NSE INDEX TABLE OF CALIBRATION STEP

<table>
<thead>
<tr>
<th>NSE</th>
<th>SS</th>
<th>ID</th>
<th>ANN</th>
<th>EF</th>
<th>GFS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calibration</td>
<td>0.68</td>
<td>0.69</td>
<td>0.77</td>
<td>0.76</td>
<td>0.18</td>
</tr>
</tbody>
</table>

---

### TABLE 3. NSE INDEX TABLE OF VERIFICATION STEP

<table>
<thead>
<tr>
<th>NSE INDEX</th>
</tr>
</thead>
<tbody>
<tr>
<td>Verification</td>
</tr>
<tr>
<td>-------------</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

---

**V. CONCLUSION**

Through implementing a number of inflow calculations to Ban Chat hydropower reservoir basin, the authors came to the following conclusions:

- The results show that the calculation method, with gridded rain data obtains by calculating rainfall from radar reflectivity, the results of artificial neural network method is better than results of the empirical formula;

- As for the gridded rain data of meteorological models as data from the GFS model with sparse resolution (27 km x 27 km) cannot be used to calculate for this basin. So, a better resolution or reference to the products of other meteorological models are needed. The meteorological model of the National Center for Hydro-Meteorological Forecasting, that has resolution of 1 km x 1 km and data assimilation with ground measuring stations, is a typical one.

In the next studies, it is necessary to improve the accuracy of gridded rain from radar reflectivity (by detecting experimental coefficients or training artificial neural networks with longer data series) and combine assimilating data with ground measuring station data for better calculation results.

---

**ACKNOWLEDGMENT**

This work is funded by the research project from the Institute of Mechanics, Vietnam Academy of Science and Technology (VAST), code VCH.TX.03/2023.

---

**REFERENCES**


