Science on Natural Resources and Environment **54** (2024) 180-188

Science on Natural Resources and Environment
Journal homepage: tapchikhtnmt.hunre.edu.vn

CALCULATING THE INFLOW TO TUYEN QUANG RESERVOIR BY HEC-HMS MODEL USING DIFFERENT LULC MAPS AND BASIN DIVISIONS

Nguyen Tuan Anh*, Nguyen Chinh Kien, Duong Thi Thanh Huong, Do Thanh Hang, Nguyen Thi Hang, Tran Thi Thanh Huyen
Institute of Mechanics, Vietnam Academy of Science and Technology

Received 04 November 2024; Accepted 23 December 2024

Abstract

For mathematical models in general and hydrological models in particular, the accuracy of input data and the detail in using that data greatly affect the calculation results. For each specific requirement, how detailed the input data needed to ensure the accuracy while still meeting the requirements of real-time tasks is an issue that needs to be surveyed and researched. In this paper, the authors use HEC-HMS model with two types of LULC (Land use-Land cover) maps and different ratios of basin division to calculate Tuyen Quang reservoir inflow for August 2023 and July 2024 floods. The results show that the Jaxa LULC map gives better results than the Microsoft and Esri LULC map and divides Tuyen Quang basin into 29 sub-basins is more suitable than 09 and 49 sub-basins. From there, the influence of LULC map as well as the ratio of basin division on the accuracy and calculation time are commented.

Keywords: LULC map; Tuyen Quang reservoir; HEC-HMS.

***Corresponding author, Email:** ntanh@imech.vast.vn

DOI: http://doi.org/10.63064/khtnmt.2024.651

1. Introduction

In distributed or semi-distributed hydrological models such as Marine (France), WetSpa (Belgium), TopModel (UK), HEC-HMS (USA) [1], detailed data on basin characteristics (Topography data, Land use-Land cover map, Soil map,...) greatly affect the calculation results of basin inflow. However, these data are difficult to accurately measure on a large scale because of the very high cost. Thanks to remote sensing technology and artificial intelligence, several types of LULC maps are built and provided for free with a good accuracy, synchronized over time on large scale and continuously updated annually. Two types of LULC maps are free and the best today.

The LULC map has resolution of 10 × 10 m including 12 layers built by analyzing images taken from the advanced ground observation satellite ALOS-4 of the Japanese aerospace agency (Jaxa). The collected data were validated by using random sampling stratification, following the procedure for assessing the accuracy of LULC changes proposed by Olofsson et al. Accordingly, 600 points are randomly distributed on 12 layers corresponding to 12 layers of LULC map. The error matrix of the samples is created based on the actual validation data. Then, the error matrix of the area ratio is calculated by multiplying each row by the area ratio of that row [3].

LULC map has resolution of 10 × 10 m from Microsoft & Esri including 9 layers built by analyzing Sentinel-2 satellite images based on artificial intelligence models trained with billions of pixels labeled images. This dataset, produced by Impact Observatory, (Microsoft & Esri) displays a global map of LULC for the years 2017 - 2023. Each map is a composite of LULC predictions for 9 classes throughout the year in order to generate a representative snapshot of each year. This dataset used billions of human-labeled pixels (curated by the US National Geographic Society) to train a deep learning model for land classification. Each global map was produced by applying this model to the Sentinel-2 annual scene collections from the Microsoft Planetary Computer. Each of the maps has an assessed average accuracy of over 75 %.

These maps have been improved from Impact Observatory's previous release and provide a relative reduction in the amount of anomalous change between classes, particularly between "Bare" and any of the vegetative classes "Trees," "Crops," "Flooded Vegetation," and "Rangeland". This updated time series of annual global maps is also re-aligned to match the ESA UTM tiling grid for Sentinel-2 imagery [4].

High resolution and detailed maps will help the model reflect the basin characteristics of the study area better. In distributed hydrological models the more detailed the data, the larger the calculation volume and requiring more time for processing. It also leads to the phenomenon of errors accumulation and poor results. For the semi-distributed model, if the area of the sub-basin is too large, it will not be able to describe all the characteristics of the region but if it is too small it will take a long time to adjust and calculate the model. It is also encountered the phenomenon of error accumulation if the model is not established well.

In Vietnam, the author Bui Dinh Lap [2] has studied how to divide and index the basin with the aim of mitigating the spatial impact of input factors such as soil, LULC map as well as the influence of spatial rainfall distribution,...

In this paper, the authors calculate Tuyen Quang reservoir inflow using two types of free and best available LULC maps from the Japan Aerospace Agency (Jaxa) and Microsoft & Esri. Furthermore, in order to study the influence of detail in basin division on the results and calculation time the authors divided the Tuyen Quang basin into sub-basins with different numbers.

2. Data collection

The authors collected the following data for calculating the Tuyen Quang reservoir inflow (The rainfall runoff throughout Tuyen Quang basin called Tuyen Quang reservoir inflow).

Digital Topographic map (DEM) has 30 × 30 m resolution data, provided free of charge by the Japan Aerospace Agency (Jaxa) in 2022 [5].

Soil map has 250 × 250 m resolution provided by FAO in 2020.

LULC map has 10 × 10 m resolution from the Japan Aerospace Agency and Microsoft & Esri in 2023.

Measured rainfall data in Tuyen Quang basin at the rain stations: Bac Me, Bao Lac, Chiem Hoa, Cho Ra, Dau Dang, Dong Van, Ha Giang, Ham Yen, Na Hang in 2023, 2024 from the National Center for Hydrometeorological Forecasting, General Department of Hydrometeorology (Figure 2, 3). Tuyen Quang reservoir inflows in 2023 and 2024 from Vietnam Electricity (EVN) [6].

Figure 1: LULC map

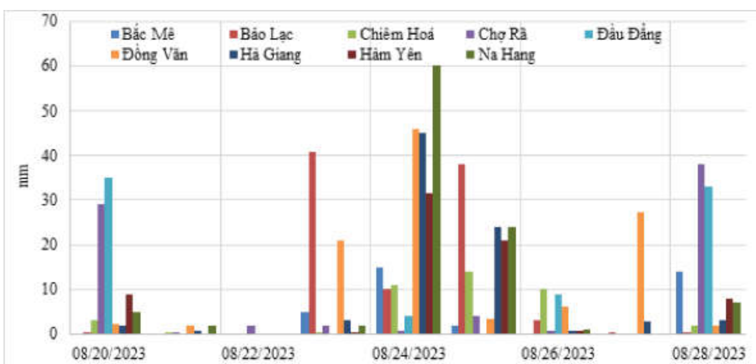

Figure 2: Rain data on Tuyen Quang basin in 2023

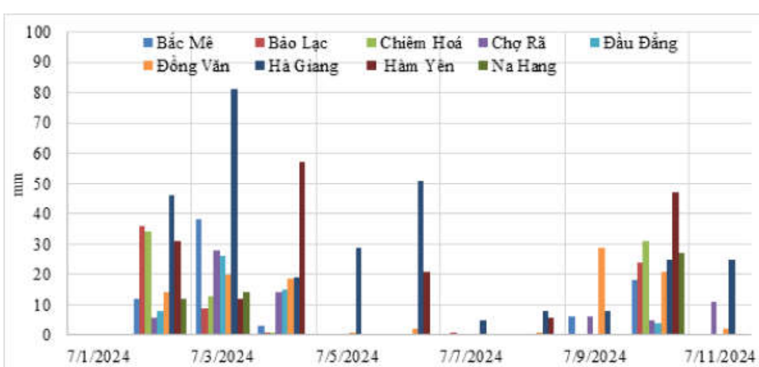

Figure 3: Rain data on Tuyen Quang basin in 2024

3. Model setup

The HEC-HMS hydrological model [7] stands for Hydrologic Engineering Center - Hydrologic Modeling System, built by the American Society of Military Engineers to calculate the process of converting rain into runoff on each sub-basin, then routing the flow on the river system [8]. In this model, the process of converting runoff from rain is calculated according to the following processes: Rain on the sub-basins will be collected on each sub-basin after infiltration process into the soil (Loss), then these flows are routed along the river branches concentrated at the outlet point of the basin. In this entire process, the loss greatly affects the final result. These processes are modeled by many different methods through average parameters of time and space. These parameters are adjusted to suit the characteristics of each studied basin.

3.1. Loss method

In HEC-HMS, loss can be calculated by many methods, to use the characteristics of the LULC map the model offers two methods: "Grid Curve Number" or "Curve Number" (CN). The CN method is more often used because it is simpler, more stable, easier to use and more useful, especially in river basins with little or no measuring stations. The CN method was introduced by the United States Land Conservation Agency (USDA) to calculate surface runoff from a certain amount of rainfall in river basins. It is a method constructed from a set of experimental equations and parameters. Its core is to determine the characteristics that hold or store the maximum amount of water in the soil according to its depth. These characteristics depend on factors: LULC, Soil types, Initial moisture conditions,...

The CN method [9] is built on the water equilibrium equation with two basic hypotheses:

- Hypothesis of balance between total rainfall and total surface water and permeability (1).

- Hypothesis of the relationship between initial permeability and maximum retention or penetration (2).

Therefore, the CN method will include:

- Water balance equation:

$$P = I_a + F + Q \qquad (1)$$

- Proportional equilibrium hypothesis:

$$\frac{F}{S} = \frac{Q}{P - Ia} \qquad (2)$$

in which: P is the total rainfall; I_a is the initial permeability (mm); F is cumulative permeability; Q is the Surface Water Level (mm); S is the maximum retention or penetration capacity (mm).

Combining equations (1) and (2) would have:

$$Q = \frac{(P - Ia)^2}{P - Ia + S} \qquad (3)$$

I_a is taken as 0.2S based on experimental research, S is calculated from soil and LULC map.

- The formula to calculate CN is as follows:

$$\text{CN} = \frac{25400}{254 + S} \qquad (4)$$

The "Curve Number set" (CNs) is calculated and used in the HEC-HMS model according to the following steps:

- The author counted the LULC classification from the two types of collected maps. Because these LULC maps have different ways in dividing data layers according to different land cover types, a full classification system is needed to serve as a standard for calculation. The authors chose the TR-55 Technical Release, the most widely used hydrological approach in the US since 1975, using the "Soil Conservation Service" flow equation so that the flow rate, the total volume, the hydrological flows is maximum. On the TR-55 Technical Release to refer to the estimated value of the CNs for different land use situations on hydrological land groups for each map, the authors have grouped and classified according to the common standard table (Table 1) on each grid cell.

- Use DEM data, soil map, LULC maps and sub-basins to calculate the CNs for each sub-basin using GIS software.

- At last, the CNs is used in the HEC-HMS model in the form of losses including: CN; Initial Loss (%); Impervious (%).

Table 1. Curve index reference table of maps with Technical Release TR-55

No.	MICROSOFT & ESRI	JAXA	SCS TR-55	A	B	C	D
1	Water, Flooded vegetation, Snow/Ice	Water - Aquaculture	-	0	0	0	0
2	Crops	Rice/Other Crops	Row crops	64	75	82	85
3	Tree	Woody, Crops/Orchards, Evergreen Forest, Deciduous Forest, Plantation Forest, Mangrove Forest	Woods	30	55	70	77
4	Built Area	Urban/Built-up	Residential districts by average lot size	54	70	80	85
5	Clouds	-	-	100	100	100	100
6	Bare Ground	Barren	Fallow: Bare soil	77	86	91	94
7	Rangeland	Grass/Shrub	Pasture, grassland, or range-continuous forage for graving	39	61	74	80

3.2. Sub-basin division

The authors divided the Tuyen Quang basin into different numbers of sub-basins (09, 29 and 49) to study the impact of the basin division on the Tuyen Quang reservoir inflow (Figure 4).

Figure 4: Basin division into 09, 29, 49 sub-basins

In addition to the loss method according to the CNs mentioned above, which are calculated with fixed values, other parameters of each sub-basin and reach need to be optimized such as time of water concentration, flow transmission parameters,… Based on the error between the calculated basin outflow and the observed value, these parameters need to be optimized so that the error is minimum.

Table 2. Number of optimization parameters

Number of sub-basin	Parameters
09	22
29	65
49	122

In the HEC-HMS model, each sub-basin and reach has its own parameters that need to be optimized. Therefore, the number of sub-basins divided will directly affect the accumulated error as well as time of parameter optimization.

3.3. Parameter optimization method

In the HEC-HMS model, there are two automatic parameter optimization methods. Simplex and Differential Evolution methods:

- The Simplex method is a parameters optimization method that uses a Simplex (a set of parameter values) to move to a peak (a position in multidimensional space) at which the value of the objective function (usually the root mean square error) is minimal. This is a direct optimization method (based on geometric operations such as reflection, expansion, contraction) and is often applied to non-linear optimization problems where the derivative may not be known.

- The Differential Evolution method by maintaining a population of the problem's solutions and creating new solutions by adding weighted distribution components randomly to the components of the existing solutions in order to create new solutions to the population. After each generation, the best fit solution (the objective function is the smallest) is kept.

In the previous publication of the authors, the parameter optimal tool of HEC-HMS is used to calculate Ban Chat reservoir inflow [10]. The Differential Evolution method has shown its superiority and was chosen to optimize time of water concentration, flow transmission,... for Tuyen Quang basin. Using the ending condition of iterations (tens of thousands) and a small tolerance (0.01 m^3/s) to ensure the convergence of the method.

- The Differential Evolution method by maintaining a population of the problem's solutions and creating new solutions by adding weighted distribution components randomly to the components of the existing solutions in order to create new solutions to the population. After each generation, the best fit solution (the objective function is the smallest) is kept.

In the previous publication of the authors, the parameter optimal tool of HEC-HMS is used to calculate Ban Chat reservoir inflow [10]. The Differential Evolution method has shown its superiority and was chosen to optimize time of water concentration, flow transmission,... for Tuyen Quang basin. Using the ending condition of iterations (tens of thousands) and a small tolerance (0.01 m^3/s) to ensure the convergence of the method.

4. Calculation and results

4.1. Assessment of LULC map influence on calculation of Tuyen Quang reservoir inflow

Using 2 types of collected LULC map and dividing Tuyen Quang basin into 9 sub-basins the authors calibrating and verifying using CNs according to empirical values that references to equivalent values of other basins (User), LULC map from Jaxa (Jaxa LULC) and LULC map from Microsoft & Esri (ME LULC) for the same division plan. Tuyen Quang basin is divided into 9 sub-basins, the authors calibrated and verified the HEC-HMS model to calculate the Tuyen Quang reservoir inflow with floods in August 2023 and July 2024. The calculation results are shown in Figure 5, 6 and evaluated by the Nash-Sutcliffe Efficiency indexes (NSE).

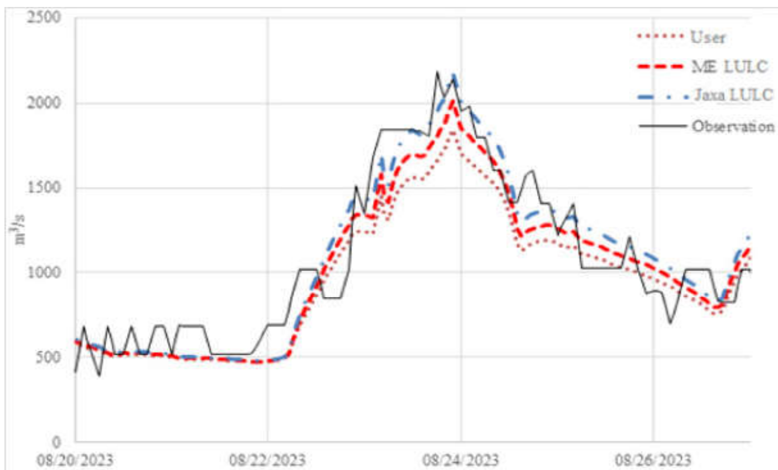

Figure 5: Calibration of Tuyen Quang reservoir inflow in 07/2024

Figure 6: Verification of Tuyen Quang reservoir inflow in 07/2024

Table 3 is the assessment table of the results according to the NSE index. The results show that Jaxa's LULC map gives the best results with the NSE index reaching 0.82 for the calibration and 0.66 for the verification.

Table 3. NSE index

Plan	NSE	
	Calibration	Verification
KN	0.61	0.58
ME	0.75	0.62
JX	0.82	0.66

4.2. Assessment of basin division influence on calculation of Tuyen Quang reservoir inflow

In this article, the authors have not yet attempted to find the optimal number of sub-basins in Tuyen Quang, Just only tested and evaluated 03 ways to divide the basins gradually increasing in number. Starting from the default division of the HEC-HMS model of 09 (JX-09) sub-basins, the authors increase the number of sub-basins (corresponding to the increase in parameters and calculation time) to 29 (JX-29) and 49 (JX-49) sub-basins (To suit existing research conditions as computer system and research time).

The results of calibration and verification in 8/2023 and 7/2024 are shown in Figure 7, 8.

The calculation time and NSE index are shown in Table 4.

Table 4. NSE index and calculation time

Plan	NSE		Cal Time (s)
	Calibration	Verification	
JX-09	0.82	0.66	163
JX-29	0.85	0.69	455
JX-49	0.75	0.63	1130

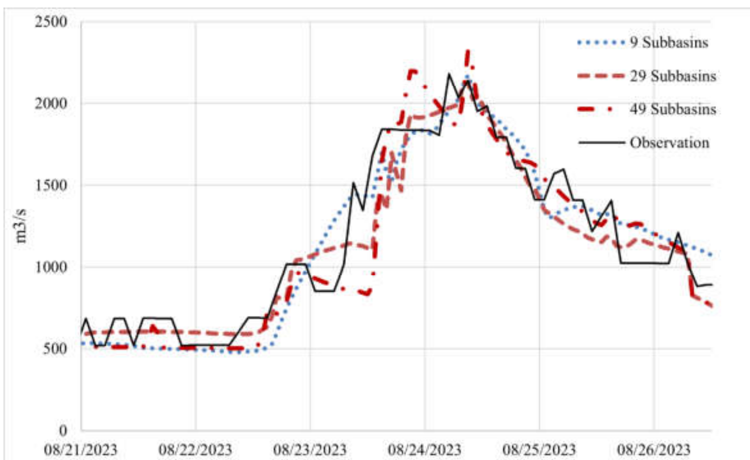

Figure 7: Calibration of Tuyen Quang reservoir inflow in 08/2023

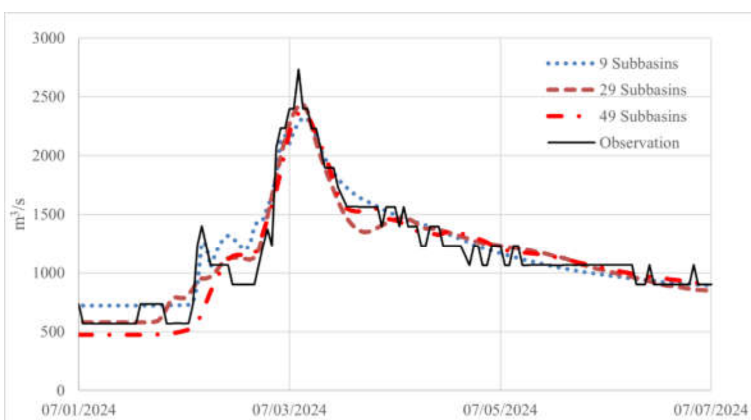

Figure 8: Verification of Tuyen Quang reservoir inflow in 07/2024

Through this table, we can see that the plan to divide Tuyen Quang basin into 29 sub-basins gives the best NSE index. In addition, the calculated flow and the measured data of this plan is also the most similar in both phase and amplitude. About calculation time, the JX-29 plan does not increase too much compared to JX-09, but it is faster than JX-49, and has a much better result.

5. Conclusion

Through the calculation results of Tuyen Quang reservoir inflow using different LULC map and different basin division the authors draw the following conclusions:

For Tuyen Quang basin, using Jaxa's LULC map to calculate loss gives better results than using Microsoft & Esri LULC map;

The division of the basin into 29 sub-basins is more suitable than the division of the basin into 09 or 49 sub-basins. It not only gives best results (NSE Calibration: 0.85 and NSE verification 0.69) but also reduces the calculation time compared with 49 sub-basins plans. That is effective in real-time forecasting Tuyen Quang reservoir inflow.

However, these are initial comments on the use of the available LULC maps and reasonable basin division. It is necessary to continue experimenting with more accurate LULC maps such as the map of the General Statistics Office of Vietnam, Vietnam Department of Survey, Mapping and Geographic Information with field-verified data information, in order to obtain a better model. This study may be used to support managers in regional planning and development (afforestation) as well as in optimally operating hydro-electric reservoirs for power generation and natural disasters mitigation.

Acknowledgments: The research team would like to sincerely thank the support of the Institute of Mechanic's project *"Research on selection of LULC maps and detail in basin division to calculate Ban Chat reservoir inflow"*, code VCH.TX.07/2024 in implementing and publishing this study.

REFERENCES

[1]. Chinh Kien Nguyen (2020). *Research on building a concentrated hydrology model for flood forecast in Vietnam river basins*. Institute of Mechanic' s project, Hanoi, p. 5 - 6.

[2]. Dinh Lap Bui (2016). *Research on application of basin number in hydrology models*. Journal of Hydro-Meteorology, 665, p. 46 - 51.

[3]. https://www.eorc.jaxa.jp/ALOS/en/dataset/LULC/LULC_vnm_v2309_e.htm.

[4]. https://livingatlas.arcgis.com/landcover/.

[5]. https://www.eorc.jaxa.jp/ALOS/en/dataset/fnf_e.htm.

[6]. https://www.evn.com.vn/c3/thong-tin-ho-thuy-dien.aspx.

[7]. US Army Corps of Engineers (2023). *Hydrological Model systems HEC-HMS User's Manual Version 4.12*. November 2023.

[8]. Tuan Anh Nguyen, Thanh Huong Duong Thi, Thanh Huyen Tran Thi (2020). *Setup HEC - HMS model to forecast inflow to Tuyen Quang reservoir in 2020*. Journal of Science on Natural Resources and Environment, 0866-7608, (34), p. 100 - 108.

[9]. Surendra Kumar Mishra & Vijay P. Singh (2023). *SCS-CN methods* in Soil Conservation Service Curve Number (SCS-CN) Methodology.

[10]. Tuan Anh Nguyen, Thi Hang Nguyen, Chinh Kien Nguyen, Thi Thanh Huong Duong, Viet Duc Nguyen, Thanh Hang Do (2023). *Optimized parameters for Tuyen Quang reservoir basin using HEC-HMS model*. In The 7th International Conference on Engineering Mechanics and Automation, Hanoi.