The 7th International Conference on
Engineering Mechanics and Automation (ICEMA - 2023)
Hanoi, November 11, 2023

DOI: 10.15625/vap.2023.0137

Optimized parameters for Tuyen Quang reservoir basin using HEC-HMS model

Tuan Anh Nguyen
Hydrodynamic and Disaster Mitigation Institute of Mechanics - VAST
Hanoi, Vietnam
ntanhvc@gmail.com

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

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

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

Viet Duc Nguyen
University of Engineering and Technology - VNU
Hanoi, Vietnam
vietduc1721@gmail.com

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

***Abstract*— In the calculation of discharge into hydroelectric reservoirs, the calibration and verification of the hydrological model HEC–HMS by trial process takes a long time and the found parameters is not optimal. This paper presents the results of using the Trial Optimal feature in the HEC-HMS to automatically detect the optimal set of parameters, calculate the inflow to Tuyen Quang reservoir. At first, the values of parameters are automatically assigned to the model. Then, HEC-HMS uses the Trial Optimal feature to detect the optimal parameters through iterative and selective calculation based on the observed data. The optimal parameters are applied to simulate the inlet discharge from May 23, 2022 to May 31, 2022 and verified the discharge from June 10, 2022 to June 17, 2022. The calculation results, that are estimated the accuracy by NSE index, show that the Trial Optimal feature has made the process of detecting the optimal parameters more accurate.**

Keywords*— *HEC-HMS, Optimal, Tuyen Quang reservoir.

I. INTRODUCTION OF HEC - HMS

HEC-HMS is a product of the Hydrologic Engineering Center within the U.S. Army Corps of Engineers. It is designed to simulate the precipitation-runoff processes of drainage basins. That can be applicable in a wide range of geographic areas for solving the widest possible range of problems. Including flood hydrology, natural watershed runoff ,...

The program is capable of representing many different watersheds. A model of the watershed is constructed by separating the water cycle into manageable pieces and constructing boundaries around the watershed of interest. Any mass or energy flux in the cycle can then be represented with a mathematical model. In most cases, several model choices are available for representing each flux. Each mathematical model included in the program is suitable in different environments and under different conditions. Making the correct choice requires knowledge of the watershed, the goals of the hydrologic study, and engineering judgement.

In HEC-HMS observed data can be used to optimize model's performance by automatically estimating parameters. In this paper the authors used two different approaches in HEC- HMS optimization: Simplex and Differential Evolution (DE). Simplex method starts with given initial parameters and adjusts them so that the simulation results match the observed data as closely as possible. Differential Evolution generates set of parameters and deals with uncertain parameters. Differential Evolution uses mutation as a search mechanism and selection to direct the search toward the prospective regions in the feasible region. It will result in different parameter sets with each experiment.

A. Simplex optimal

Simplex searches for the optimal parameter value without using derivatives of the objective function to guide the search. Instead this algorithm relies on a simpler direct search. In this search, parameter estimates are selected with a strategy that uses knowledge gained in prior iterations to identify good estimates, to reject bad estimates, and to generate better estimates from the pattern established by the good.

The simplex search uses a simplex - a set of alternative parameter values. For a model with n parameters, the simplex has n+1 different sets of parameters.

For example, if the model has two parameters, a set of three estimates of each of the two parameters is included in the simplex. Geometrically, the n model parameters can be visualized as dimensions in space, the simplex as a polyhedron in the n-dimensional space, and each set of parameters as one of the n+1 vertices of the polyhedron. In the case of the two-parameter model, then, the simplex is a triangle in two-dimensional space, as illustrated in "Fig. 1".

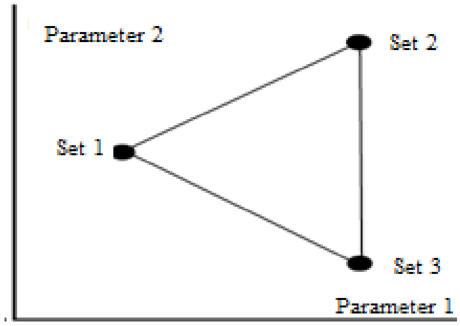

Fig. 1. Initial simplex for a 2-parameter model

This method is to find a vertex at which the value of the objective function is a minimum. To do so, it uses the following operations:

- Comparison: The first step in the evolution is to find the vertex of the simplex that yields the worst (greatest) value of the objective function and the vertex that yields the best (least) value of the objective function. In “Fig. 2”, these are labeled W and B, respectively.
- Reflection: The next step is to find the centroid of all vertices, excluding vertex W; this centroid is labeled C in “Fig. 2”. The algorithm then defines a line from W, through the centroid, and reflects a distance WC along the line to define a new vertex R, as illustrated “Fig. 2”.

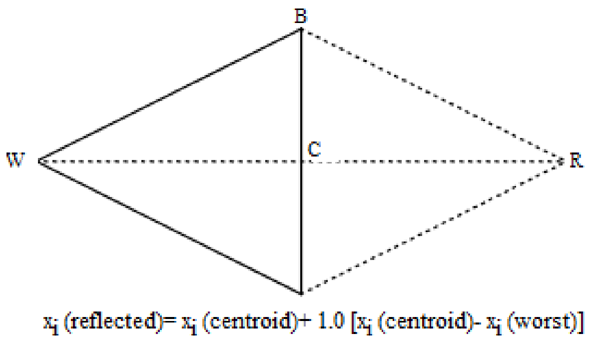

Fig. 2. Reflection of a simplex

- Expansion: If the parameter set represented by vertex R is better than, or as good as, the best vertex, the algorithm further expands the simplex in the same direction, as illustrated in “Fig. 3”. This defines an expanded vertex, labeled E in the figure. If the expanded vertex is better than the best, the worst vertex of the simplex is replaced with the expanded vertex. If the expanded vertex is not better than the best, the worst vertex is replaced with the reflected vertex.

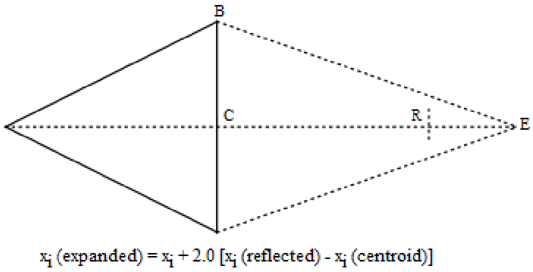

Fig. 3. Expansion of a simplex

- Contraction: If the reflected vertex is worse than the best vertex, but better than some other vertex (excluding the worst), the simplex is contracted by replacing the worst vertex with the reflected vertex. If the reflected vertex is not better than any other, excluding the worst, the simplex is contracted. This is illustrated in “Fig. 4”. To do so, the worst vertex is shifted along the line toward the centroid. If the objective function for this contracted vertex is better, the worst vertex is replaced with this vertex.

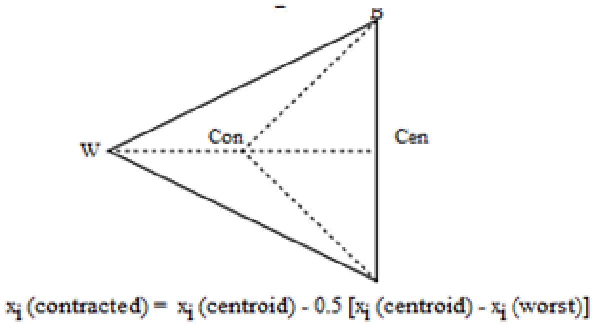

Fig. 4. Contraction of a simplex

- Reduction: If the contracted vertex is not an improvement, the simplex is reduced by moving all vertices toward the best vertex. This yields new vertices R_1 and R_2, as shown in “Fig. 5”.

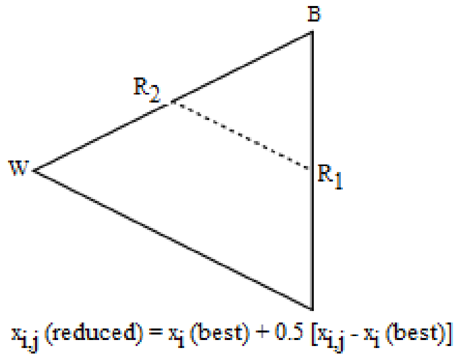

Fig. 5. Reduction of a simplex

The search terminates when either of the following criterion is satisfied:

$$\sqrt{\sum_{j=1, j \| worst}^{n} \frac{(z_j - z_c)^2}{n-1}} < tolerance$$

In which: n = number of parameters; j = index of a vertex, c = index of centroid vertex; and z_j and z_C = objective function values for vertices j and c, respectively.

Or the number of iterations reaches 50 times the number of parameters.

The parameters represented by the best vertex when the search terminates are reported as the optimal parameter values [1].

B. Differential evolution optimal

DE is a population based search technique which utilizes NP variables as population of D dimensional parameter vectors for each generation. The initial population is chosen randomly if no information is available about the problem. In the case of the available preliminary solution, the initial population is often generated by adding normally distributed random deviations to the preliminary solution. The basic idea behind DE is a new scheme for generating trial parameter vectors. DE generates new parameter vectors by adding the weighted difference vector between two population members to a third member. If the resulting vector yields a lower objective function value than a predetermined population member, the newly generated vector replaces the vector with which it was compared. In addition, the best parameter vector is evaluated for every generation in order to keep track of the progress that is made during the optimization process. Extracting the distance and the direction information from the population to generate random deviations result in an adaptive scheme with excellent convergence properties.

DE maintains two arrays, each of which holds a population size NP and D dimensional, real-valued vectors. The primary array holds the current vector population, while the secondary array accumulates vectors that are selected for the next generation. In each generation, NP competitions are held to determine the composition of the next generation.

Every pair of vectors (X_a, X_b) defines a vector differential: (X_a-X_b). When X_a and X_b are chosen randomly, their weighted differential is used to perturb another randomly chosen vector X_c. This process can be mathematically expressed as:

$$X'_c = X_c + F (X_a - X_b)$$

The weighting, or scaling, factor F is a user supplied constant in the optimal range between 0.5 and 1.0. In every generation, each primary array vector X_i is targeted for crossover with a vector like X'_c to produce a trial vector X_t. Thus, the trial vector is the child of two parents, a noisy random vector and the target vector against which it must compete. Uniform crossover (that can take child vector parameters from one parent more often than it does from others) is used with a crossover constant (CR), in the optimal range of 0.5 to 1.0 which actually represents the probability that the child vector inherits the parameter values from the noisy random vector.

When CR = 1, for example, every trial vector parameter is certain to come from X'_c. On the other hand, if CR = 0, all but one trial vector parameter comes from the target vector. To ensure that X_t differs from X_i by at least one parameter, the final trial vector parameter always comes from the noisy random vector even when CR = 0. Then the objective function corresponding to the trial vector is compared with that of the target vector, and the vector that has the lower objective function value (for minimization) of the two would survive for the next generation. This process is continued until the termination criterion of a preset maximum number of generations (MAXGEN) is met, and difference in objective function values between two consecutive generations reaches a small value [2].

II. HEC-HMS SETUP

The authors collected data from different reliable sources:

- 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 [4];
- 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 [5];
- Soil map with 250 m x 250 m resolution taken from the United States Department of Agriculture (USDA) in 2020;
- Measured rain data at Bac Me, Bao Lac, Chiem Hoa, Cho Ra, Dau Dang, Dong Van, Na Hang stations in 2022 at every 6 hours are collected from National Centre for Hydro-Meteorological Forecasting of Vietnam.
- Data of inflow discharge into Tuyen Quang reservoir in 2022 from Electricity of Vietnam [6].

From DEM data, the authors divided the Tuyen Quang reservoir basin into 15 sub-basins “Fig. 6”, inputted the coordinates of rain stations, rain data, and inflow discharge for calculation.

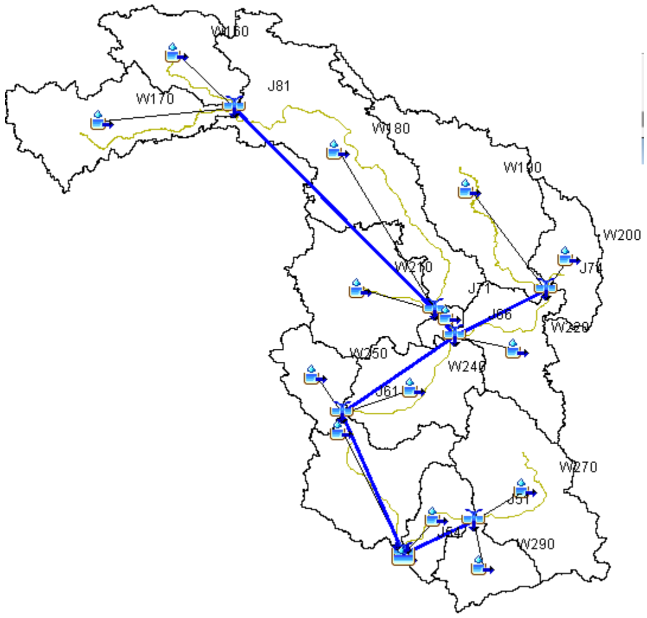

Fig. 6. Sub-basins of Tuyen Quang reservoir basin

The main input parameters include: Green and Amp permeability, flow transform and routing. That includes:

- Green and Amp Loss: Calculate losses using the Green and Ampt method.
- Flow Transform: SCS Unit Hydrograph method determines the unit hydrograph curve.
- Routing: Using Lag method, this is a simple method with the assumption that the flow from the upstream flows intact to the downstream after a delay period.

All these above parameters are setup in the model and used to optimize in order to get the best result of model calibration. These parameters are showed in “Table I”.

TABLE I. MAIN PARAMETERS

Parameter	Unit	Value
Green and Ampt - Conductivity	mm/hr	0.4
Green and Ampt - Suction Scale Factor		0.6
Green and Ampt - Initial Content Scale Factor		0.8
Green and Ampt - Saturated Content Scale Factor		0.9
SCS Unit Hydrograph	minute	570
Routing - Lag - Initial Discharge	m3/ s	5
Routing - Lag	minute	200

The basin is divided into 15 basins and 7 reaches, the total number of parameters that used to be adjusted is 82. The initial values of the parameters are estimated as follow:

- Water concentration time in the basin (T_c) and water storage coefficient in the basin (R) are calculated according to the length of the main river from the source to the outlet of the basin, the length from the outlet to the center of the basin and the average slope of the flow path;
- Green and Amp parameters are estimated from land cover and land use maps of the basin;
- Transmission time depends on the length and slope of the river branch.

The maximum and minimum value range of each parameter is selected corresponding to these initial values.

III. SCENARIOS AND CALCULATION

Had the model established, the authors have simulated and calibrated the inflow discharge into Tuyen Quang reservoir with two different rain station assignment methods: Specified Station (SS) and Inverse Distance (ID) as shown in "Table II":

TABLE II. SCENARIOS AND CALCULATION

Scenarios		SS	ID
Calibration 23/5/2022 30/5/2022	Calibration	PA1	PA2
	Simplex Optimal	PA3	PA4
	DE Optimal	PA5	PA6
Verification 10/6/2022 - 17/6/2022		PA7	PA8

The results of SS calibration and optimization are shown as "Fig. 7":

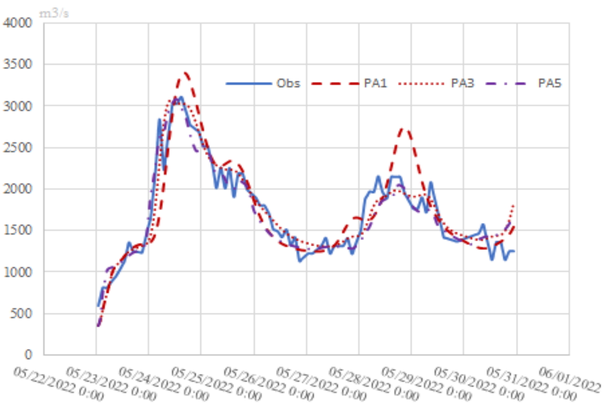

Fig. 7. SS calibration and optimization

The results of ID calibration and optimization are shown as "Fig. 8":

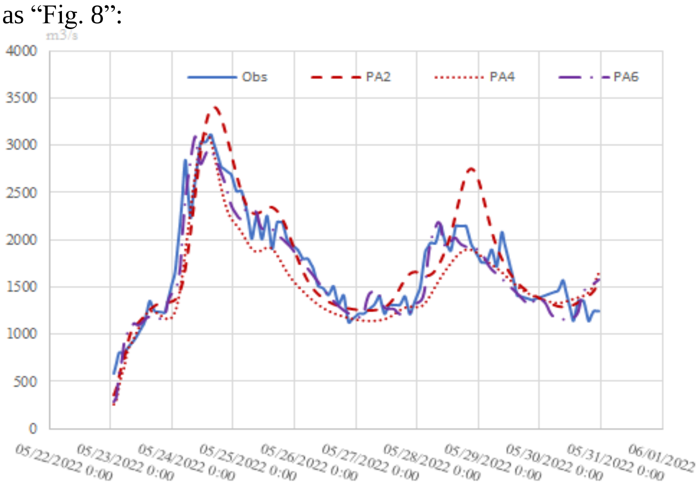

Fig. 8. ID calibration and optimization

The calculation results are evaluated using the NSE index (Table III).

TABLE III. NSE INDEX OF THE RESULTS

Scenarios	SS	ID
Calibration	0.80	0.77
Simplex Optimal	0.91	0.92
DE Optimal	0.91	0.94

Although calculation time of each calibration is short, but achieving the best final result requires trial and error many times and depends on the experience of the forecaster. In addition to reducing time of calculation, Optimal also provides outstanding results as the NSE index above. Notice that the calculation time of the Simplex algorithm only reaches 1/3 calculation time of the DE algorithm, but the NSE index results are not as good as.

The NSE index shows that DE optimal yield the best results. Using the set of DE optimal parameters that obtained after calibration, the authors have verified the model (PA7, PA8) as "Fig. 9".

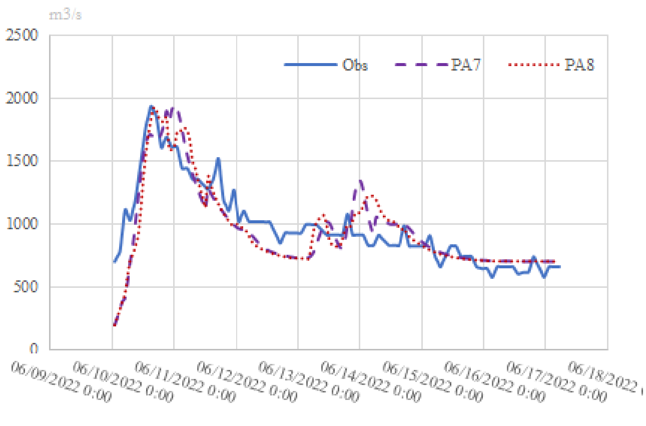

Fig. 9. Verification results

Verification results are evaluated by the NSE index, the SS verification NSE = 0.66, DI verification NSE = 0.67. This shows that the parameters obtained in the calibration accurately and reflects the characteristics of the basin. The NSE value of the verification are good.

For a deeper evaluation of the DE optimal method, the authors calculated additional options with different parameters. Then apply the optimized value of the parameters to validate the model. “Tables IV, V” are the results of DE optimal with different parameters.

TABLE IV. DE OPTIMAL WITH DIFFERENT ITERATIONS

Population size 30								
Ite	20	50	100	200	300	400	500	900
NSE	0.166	0.753	0.937	0.928	0.927	0.820	0.813	0.801

TABLE V. DE OPTIMAL WITH DIFFERENT POLULATION SIZES

Iteration 100								
Size	10	20	30	40	50	60	70	80
NSE	-1.10	0.896	0.937	0.945	0.939	0.930	0.935	0.922

It is noticed that when the population size is 30 or more, the NSE value has reached around the maximum level (NSE ≈ 0.94). The difference in NSE value is due to the random selection of initial values. If these random values are closer to the maximum value (easier to reach the optimal solution), NSE will achieve better results.

The result shows that, corresponding 82 parameters that need to be adjusted, the Tuyen Quang basin only needs a population size of 40 and 100 iterations to get the best result of DE optimal. This number of parameters is reasonable for Tuyen Quang basin as well as a reference for other basins.

IV. CONCLUSION

Calculation results show that the parameter optimization feature in the HEC-HMS hydrological model greatly helps in finding optimal parameters for model calibration.

For the Tuyen Quang basin, the DE algorithm has higher accuracy but the running time is longer than the Simplex algorithm. DE algorithm, for the Tuyen Quang basin devided into 15 sub-basins), gives the best result when population size is 40 in 100 iterations.

Calculating scenarios using the algorithm to automatically detect optimal parameters brings many benefits and helps forecasters gain more experience to apply to the Tuyen Quang basin as well as other basins.

ACKNOWLEDGMENT

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

REFERENCES

[1]. USACE, “Hydrologic Modeling System HEC–HMS Technical Reference Manual”, 2000, 102-104.

[2]. Vasan Arunachalam,”Optimization Using Differential Evolution”, Water Resources Research Report, July 2008, 3-4.

[3]. Nguyen Chinh Kien, “Testing the application of artificial neural networks in hydrological and hydraulic forecasting”, The 10th National Conference on Mechanics, Ha Noi, 12/2017.

[4]. https://www.eorc.jaxa.jp/ALOS/en/dataset/fnf_e.htm (01/11/2023)

[5]. https://livingatlas.arcgis.com/landcover/ (01/11/2023)

[6]. https://www.evn.com.vn/c3/thong-tin-ho-thuy-dien.aspx (01/11/2023)