Research article

Examination of Optimal Search Method of Unknown Parameters in Tank Model by Monte Carlo Method

KOJI MIWA
Extension Center, Institute of Environmental Rehabilitation and Conservation, Japan

KOSUKE MUKAE
Hitonomori Corporation, Kanagawa, Japan

HIROMU OKAZAWA*
Faculty of Regional Environment Science, Tokyo University of Agriculture, Tokyo, Japan
Email: h1okazaw@nodai.ac.jp

TADAO YAMAMOTO
Research Faculty of Agriculture, Hokkaido University, Hokkaido, Japan

TAKASHI INOUE
Research Faculty of Agriculture, Hokkaido University, Hokkaido, Japan

EMMANUEL OKIRIA
Graduate School of Agriculture, Tokyo University of Agriculture, Tokyo, Japan

Received 9 April 2018 Accepted 5 March 2020 (*Corresponding Author)

Abstract Recent years, flood disasters have become rampant due to climate change. For the risk mitigation, building up a hydrological run-off model which can be used accurately and quickly helps communities prepare effective disaster prevention measures. Tank Model is a hydrological run-off model proposed in Japan. On one hand, its calculation is relatively simple. On the other hand, many unknown parameters have to be identified. Thus, the more random number it generates at a time, the more it takes time for calibration. This research examines the optimal method of identifying the unknown parameters by Monte Carlo method, considering the improvement of efficiency for practical use. Montel Carlo method can generate the massive number of random samples and help us select the best combinations of unknown parameters. Among the generated random numbers, the optimal parameters are searched which can fit to the actual measured values with the minimal number of samples. This research examines how many random numbers need to be generated at minimum to obtain the optimal parameters. The number (*N*) of random samples were seven kinds of 100, 1,000, 10,000, 100,000, 1,000,000, 10,000,000 and 100,000,000. Every time random samples generated by Montel Carlo method, the maximum and minimum numbers of each unknown parameter among the five best combinations were applied for the retrieval range of the next simulation. In this way, the simulation was repeated ten times for each kind of sample number. As the result, the watershed with simple land-use could obtain the optimal parameter with fewer samples (N = 1,000) than the watershed with complex land-use type (N = 1,000,000). With these sample numbers, the prediction accuracy for both watersheds were high. It is considered that the complex land-use watershed had lower accuracy rate because the water runoff was influenced by the underdrainage facilities.

Keywords hydrological runoff model, Tank Model, Monte Carlo simulation, watershed

INTRODUCTION

Hydrological disasters such as storms, flood and drought have been rampant in many parts of the world in the recent years. Especially floods have caused huge economic damages and human casualties. To prevent or mitigate the damages, the improvement of the prediction accuracy and speed on water-runoff by a hydrological model is one of the measures. Generally, the prediction failure of a hydrological runoff model is attributed to errors in input data, model structure and parameter setting.

A hydrological run-off model, Tank Model is a lumped conceptual model to predict rainfall-runoff, developed by Sugawara (1985). It is considered a simple model but requires model parameters to be identified for calibration. To identify optimal parameters, various methods have been studied. Tanakamaru (1995) identified the accuracy of the Stuffled Complex Evolution method and the multistart Powell method for the parameter estimation. Tada (2007) examined the Particle Swarm Optimization algorithm and discussed its effectiveness under certain conditions. Fujihara et al. (2003) investigated the selection of objective functions to be used for identification of the parameters. For other examples, Kalman Filter (Ichihara et al., 2000) and Monte Carlo method (Mukae et al., 2017) have been applied. Monte Carlo is a simple simulation which generates a huge number of numeric random samples. However, the optimal number of samples to be generated has not been examined.

OBJECTIVE

This research aims to clarify the optimal search method by Monte Carlo method to identify the parameters of the Tank Model quickly and accurately. In this study, the simulation of Monte Carlo was implemented by a software, MATLAB developed by MathWorks Inc. The fewer samples, the less time it takes to simulate. Therefore, it is important to find the minimal number of samples to determine the optimal parameters promptly and accurately. Through this examination, it is expected to suggest a minimal number of samples to identify the parameters with expedition and accuracy in the Tank Model.

METHODOLOGY

Study Area and Data Collection

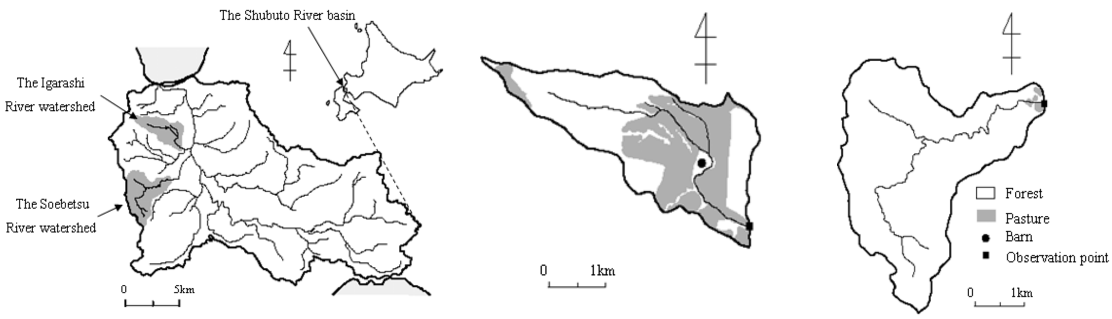

Fig. 1 The Shubuto River Basin, Hokkaido, Japan (left), land-use of Igarashi River Watershed (center) and Soebetsu River Watershed (right) (*Okazawa, et al. 2002*)

The study sites of this research are watersheds of the Igarashi River and the Soebetsu River. Both of them, in fact, are the size of the stream rather than rivers and tributaries of the Shubuto River, located in southwestern Hokkaido, Japan (Fig. 1). The Igarashi River watershed has the area of 6.9 km^2 and the river length is 7.3 km. This watershed consists of a complex land-use with 2.7 km^2 of pastureland in the downstream basin and 4.2 km^2 of forestland in the upper and middle basin, which covers 31% and 69% of the watershed respectively. This pastureland is mainly used for livestock and there is cropland in a part of the upper basin. Therefore, this watershed contains the land-use of agricultural and forest lands. On the other hand, the watershed area of the Soebetsu is 11.5 km^2, and the river length is 11.1 km. It is a simplex land-use only consisted of forest land. Since these watersheds are close each other, the meteo-hydrological data, such as precipitation amount and evapotranspiration for both areas have the similarities.

This research used data of the daily record of river discharge and precipitation from June 1, 1998 to October 31, 1998 (153 days) observed by Okazawa, et al. (2002). The daily data of temperature, wind speed and hours of sunshine which are required to calculate daily evapotranspiration, are obtained from AMeDAS in Kuromatsunai Town managed by Japan Meteorological Agency.

Hydrological Model: Tank Model

The tank model sets four tanks vertically and sequentially as shown in Fig. 2. In the top tank, precipitation amount is entered and evapotranspiration is subtracted. When there is no water in the top tank, evapotranspiration is subtracted from the lower tanks in order. The water flows from the upper to lower tanks through the bottom outlet as percolation and infiltration or flows out of the side outlets as runoffs which are set at a certain height in each tank. The runoff from the side outlets is considered surface flow in the top tank, intermediate runoff in the second tank, sub-base runoff in the third tank, and base flow in the fourth tank. The total of these runoff becomes the river runoff. Until the water level reaches side outlets, the runoff does not start and the water keeps flowing downward.

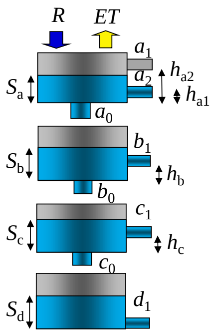

Fig. 2 Concept of Tank Model

The water level of the top tank of simulation is calculated by Eq. (1).

$$S_{a\,i+1} = R - ET + S_{a\,i} - Q_{a1\,i} - Q_{a2\,i} - Q_{a0\,i} \qquad (i = 1 \cdots n) \tag{1}$$

Where R is daily rainfall (mm), ET is daily evapotranspiration calculated by FAO Penman-Monteith method. Q_{a1} is water volume discharged by the upper side outlet, Q_{a2} is water volume discharged by the lower side outlet and Q_{a0} is water volume discharged to the second tank from the bottom, of the top tank. i is elapsed day(s), the maximum of n is 153 (days). The side and bottom outlets are calculated as Eqs. (2) to (6).

$$\text{if } S_{a\,i} \geq h_{a1\,i},\ \ Q_{a1\,i} = a_1(S_{a\,i} - h_{a1\,i}) \tag{2}$$

$$\text{if } S_{a\,i} < h_{a1\,i},\ \ Q_{a1\,i} = 0 \tag{3}$$

$$\text{if } S_{a\,i} \geq h_{a2\,i},\ \ Q_{a2\,i} = a_2(S_{a\,i} - h_{a2\,i}) \tag{4}$$

$$\text{if } S_{a\,i} < h_{a2\,i},\ \ Q_{a2\,i} = 0 \tag{5}$$

$$Q_{a0\,i} = a_0\, S_{a\,i} \tag{6}$$

Where h_{a1} and h_{a2} are the height of side outlet, a_1, a_2 and a_0 are coefficient of discharge in each outlet in the top tank. The second and third tanks are calculated in the same way as Eqs. (7) to (14).
The second tank:

$$S_{b\,i+1} = Q_{a0\,i} + S_{b\,i} - Q_{b1\,i} - Q_{b0\,i} \tag{7}$$

$$\text{if } S_{b\,i} \geq h_{b1\,i},\ \ Q_{b1\,i} = b_1(S_{b\,i} - h_{b\,i}) \tag{8}$$

$$\text{if } S_{b\,i} \geq h_{b1\,i}, \;\; Q_{b1\,i} = 0 \tag{9}$$

$$Q_{b0\,i} = b_0\, S_{b\,i} \tag{10}$$

The third tank:

$$S_{c\,i+1} = Q_{b0\,i} + S_{c\,i} - Q_{c1\,i} - Q_{c0\,i} \tag{11}$$

$$\text{if } S_{c\,i} \geq h_{c1\,i}, \;\; Q_{c1\,i} = c_1(S_{c\,i} - h_{c\,i}) \tag{12}$$

$$\text{if } S_{c\,i} \geq h_{c1\,i}, \;\; Q_{c1\,i} = 0 \tag{13}$$

$$Q_{c0\,i} = c_0\, S_{c\,i} \tag{14}$$

As the forth tank is considered the bottom, it is calculated as Eqs. (11) to (13).

$$S_{d\,i+1} = Q_{c0\,i} + S_{d\,i} - Q_{d1\,i} \tag{11}$$

$$\text{if } S_{d\,i} \geq 0, \;\; Q_{d1\,i} = d_1\, S_{d\,i} \tag{12}$$

$$\text{if } S_{d\,i} = 0, \;\; Q_{d1\,i} = 0 \tag{13}$$

Calibration of water runoff in the tank model requires to calculate and identify the sixteen parameters, which are "a_1", "a_2", "a_0", "b_1", "b_0", "c_1", "c_0", "d_1", "$S_{a\,1}$", "$S_{b\,1}$", "$S_{c\,1}$", "$S_{d\,1}$", "h_{a1}", "h_{a2}", "h_b" and "h_c" as seen in Fig. 2. *a, b* and *c* are coefficient of discharge from the side and bottom outlets in the top, second and third tanks respectively. The bottom tank has only one side outlet, *d*. *S* is water level of each tank. *h* is the height of the side outlets from the bottom of each tank. Only the top tank has two side outlets.

As an evaluation function, the compatibility of the actual value of river discharge and the estimated value is evaluated by Nash-Sutcliffe efficiency (*NSE*) as Eq. (14). The closer the *NSE* value is to 1.0, the higher the accuracy of the simulation is. Moriasi et al. (2007) divided *NSE* into ≤ 0.50, 0.50 to 0.65, 0.65 to 0.75 and > 0.75 and evaluated unsatisfactory, satisfactory, good and very good, respectively. N stands for the total number of calculation time, Q_{obs}(i) is the actual river discharge at any time step of i, Q_{sim}(i) is the estimated river discharge of at any time step of i, and Q_{av} is the mean value of the actual river discharge. *NSE* value is calculated as Eq. (14).

$$NS = 1 - \frac{\sum_{i=1}^{N} \left[Q_{obs}(i) - Q_{sim}(i)\right]^2}{\sum_{i=1}^{N} \left[Q_{obs}(i) - Q_{av}\right]^2} \tag{14}$$

Calibration Approach

The flow of this research method is as shown in Fig. 3. The initial retrieval range for each parameter was decided as shown in Table 1. The random numbers were created between the minimum and maximum value for each parameter by Monte Carlo method. The numbers of random samples were 1.0×10^2, 1.0×10^3, 1.0×10^4, 1.0×10^5, 1.0×10^6, 1.0×10^7 and 1.0×10^8. These samples were calculated in the tank model to obtain the *NSE* value for each sample. From the top five of highest *NSE* value, the maximum and minimum values of each parameter were confirmed and used as the next retrieval range of maximum and minimum. This process was repeated nine times. In this way, the retrieval range was narrowed down. The value which has the highest *NSE* value was selected as the parameter obtained from the 10th simulation.

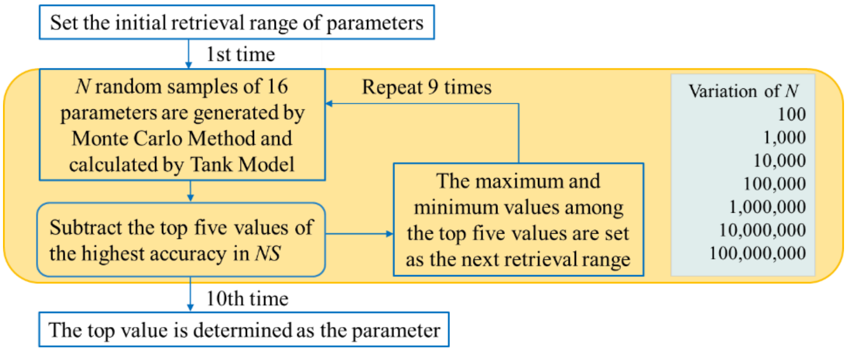

Fig. 3 Flow of calculation method with Monte Carlo method

Table 1 Initial retrieval range of value in each parameter

	a_1	a_2	a_0	b_1	b_0	c_1	c_0	d_1
Min. value	0	0	0	0	0	0	0	0
Max. value	0.3	0.3	0.1	0.1	0.1	0.1	0.1	0.1

	S_{a1}	S_{b1}	S_{c1}	S_{d1}	h_{a1}	h_{a2}	h_b	h_c
Min. value	0	0	0	0	15	0	20	20
Max. value	30	50	200	300	50	20	50	100

RESULTS AND DISCUSSION

The results of the simulation for Igarashi River and Soebetsu River watersheds are shown in Table 2. In general, the higher the sample number (*N*) is, the higher *NSE* value is obtained. For the Igrarashi River watershed where land-use is the mix of forest and agriculture, the increase of *NSE* value slowed down after $N = 1.0\text{x}10^7$ (*NSE* value = 0.801). However, since the increase from $N = 1.0\text{x}10^6$ (*NSE* value = 0.799) to $N = 1.0\text{x}10^7$ is very slight and the value was rather decreased at $N = 1.0\text{x}10^8$ (*NSE* value = 0.798), the minimal and optimal sample number was considered *N* = 1,000,000. The *NSE* value was insufficient in $N = 1.0\text{x}10^2$ as the value was lower than 7.0. On the other hand, for the Soebetsu River watershed where land-use is solely forest, the *NSE* value was high enough even for *N* = 100 (*NSE* value = 0.848). The increase from $N = 1.0\text{x}10^3$ onward was limited. Thus, $N = 1.0\text{x}10^3$ is considered optimal.

Comparing both watersheds, the forest watershed could obtain the optimal parameter with much fewer samples than the mixed-land use watershed. From these, it is considered that the simple land-use watershed has the simple water movement, thus could obtain the optimal parameter with fewer sample numbers.

Table 2 NSE value in Igarashi and Soebetsu watersheds after the tenth simulation

No. of random sample numbers	1.0×10^2	1.0×10^3	1.0×10^4	1.0×10^5	1.0×10^6	1.0×10^7	1.0×10^8
NSE Value in Igarashi watershed	0.690	0.766	0.765	0.776	0.799	0.801	0.798
NSE Value in Soebetsu watershed	0.848	0.872	0.872	0.875	0.878	0.880	0.882

Hyetograph and hydrograph for each watershed are shown in Fig. 4 and 5. For the Igarashi River, the optimal parameters were obtained from $N = 1.0\text{x}10^6$ and for Soebetsu $N=1.0\text{x}10^3$. Comparing to the actual river discharge, both simulation had high accuracy. However, in the Igarashi, although the gap was little at the peak of discharge at the rainfall time, it was large when water is calm with little rain (base flow condition of river). In addition, the value of simulation is higher than the actual value in June when rainfall is low. On the other hand, the value of simulation was lower than the actual value in July when there is almost no rainfall. Such occurrence is considered because of drainage facilities such as underdrain in agricultural lands. Therefore, water movement becomes more complicated in agricultural watershed than forest watershed. For this reason, the prediction accuracy becomes lower in agricultural watersheds than forest watershed by four layers of tank model.

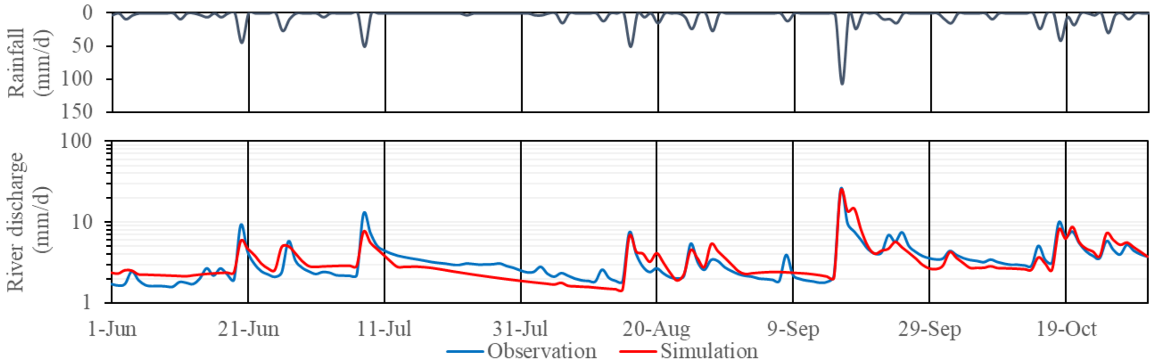

Fig. 4 Hyetograph and hydrograph with Tank model simulation in the Igarashi

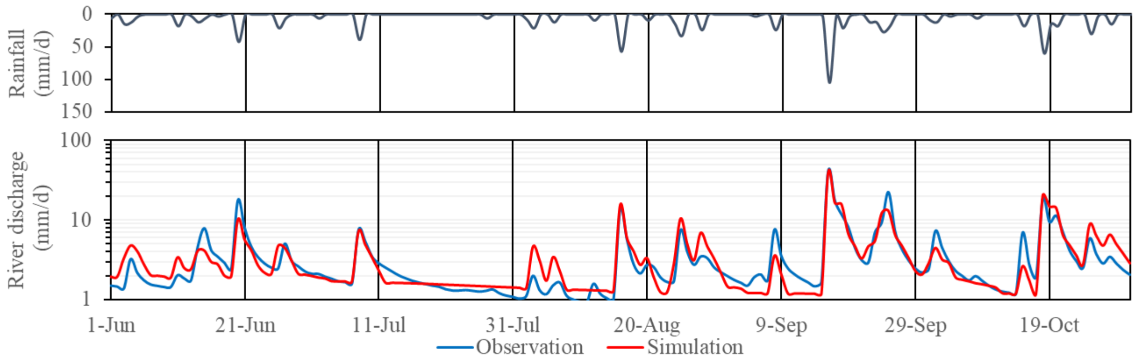

Fig. 5 Hyetograph and hydrograph with Tank Model simulation in the Soebetsu

The above results showed that the larger number of sample is, the higher the *NSE* value tends to be obtained up to a certain point. When the land-use is simple, the high accuracy is obtained with fewer samples than the case of complex land-use. Although there are several different methods to identify parameters, by using Monte Carlo method with an appropriate software, it is able to predict the rainfall-runoff for simple-land use area with a limited time and effort. In case of complex land-use, it is suggested to conduct the prediction considering the time required to conduct the calibration a number of times.

CONCLUSION

This research showed the predictability of rainfall-runoff with the minimal number of samples to obtain the results as swiftly and accurately as possible. The results showed that the larger number of sample numbers is, the higher the *NSE* value is obtained. However, the increasing rate is slowed down at certain sample numbers and the further attempts seemed unnecessary for both watersheds. In the research site for this study, the simple forest watershed required fewer number of random samples, which was $N = 1.0\text{x}10^3$, compared with the complex land-use watershed with forest and agricultural lands, which was $N = 1.0\text{x}10^6$. It is assumed that agricultural watersheds have lower accuracy in four layers of tank model as they require additional calculation in parameter setting as the runoff is influenced by drainage facilities. In the future study, it is important to continue testing in different cases to generalize the optimal number of random samples to confirm the parameters by Monte Carlo method in the tank model. For the case of agricultural watersheds, it needs to consider the water movement influenced by drainage facilities in the calibration process.

ACKNOWLEDGEMENTS

This research was supported by the grant-aid from Tokyo University of Agriculture in Japan, and by the joint usage/research program of the Institute of Materials and Systems for Sustainability (IMaSS), Nagoya University.

REFERENCES

Fujihara, Y., Tanakamaru, H., Hata, T. and Tada, A. 2003. Study on selection of objective functions for calibration of rainfall-runoff models. Transactions of the Japanese Society of Irrigation, Drainage and Reclamation Engineering, 225, 423-435.

Ichihara, K., Toyokawa, K., Sawaguchi, I. and Kawashima, H. 2000. Real time runoff forecasting with tank models whose parameter values are identified by Kalman filter. Journal of the Japanese Forestry Society, 82 (2), 125-131.

Moriasi, N.D., Arnold, J.G., Van Liew, M.W., Bingner, R.L., Harmel, R.D. and Veith, T.L. 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Transactions of the American Society of Agricultural and Biological Engineers, 50 (3), 885-900.

Mukae, K., Miwa, K., Okazawa, H. and Fujikawa, T. 2017. Ecosystem service assessment in agricultural watershed by using TOPMODEL. AGROFOR International Jounal, 2 (3), 27-36.

Sugawara, M. 1985. Tank model, For the derivation of river discharge from rainfall. Journal of Geography, 94 (4), 1-13.

Tada, T. 2007. Parameter optimazation of hydrological model using the PSO algorithm. 20 (5), 450-461.

Tanakamaru, H. 1995. Parameter estimation for the tank model using global optimization. 178 (63-4), 103-112.