An Estimation of Rainfall using Fuzzy Set-Genetic Algorithms Model

Teerawat Thongwan, Anongrit Kangrang and Sahalaph Homwuttiwong
Department of Civil Engineering, Faculty of Engineering, Mahasarakham University,
Khamriang, Kantharawichai, Maha sarakham, 44150, Thailand

Abstract: Problem statement: Damaged by floods are natural disasters that have violence cause significant damage and economic and social. If we can prevent disasters that may occur in advance is important. So an estimated rainfall data is important information for prevention disasters. **Approach:** The objective of this study is to apply a fuzzy set theory to estimate rainfall. The genetic algorithm was applied to calibrate the fuzzy set model. The proposed model considered only a few basic hydrological parameters including temperature, humidity, wind speed and solar radiation. The proposed model was applied to estimate the rainfall in the Chi River Basin (in the northeast region of Thailand) using 5-minute historic data. **Results:** The results have shown that the obtained rainfalls of the improved model are close to the rainfall of the actual rainfall record. Furthermore, the results presented that the genetic algorithm calibration provided the optimal condition of membership function. **Conclusions/Recommendations:** The proposed fuzzy-GA model can be used to estimate the rainfall, given only the basic hydrological parameters; temperature, humidity, wind speed and solar radiation. The fuzzy set model considering 4 variables using rainfall duration data is more effective than the model using the continuous rainfall data.

Key words: Fuzzy set, genetic algorithm, rainfall forecast, hydrological parameters

INTRODUCTION

Damaged by floods are natural disasters that have violence cause significant damage and economic and social. If we can prevent disasters that may occur in advance is important. So an estimated rainfall data is important information for prevention disasters. Measures to prevent and mitigate damage caused by flooding to effectively and promptly. Nowadays, rainfall forecast was studied by using mathematical models (Frencha *et al*., 1992; Hormwichian *et al*., 2009; Otok and Suhartono, 2009; Andrieu *et al*., 1996). These models required a lot of hydrological and meteorological data. Hence, there are limited for forecasting rainfall lacking data.

A Fuzzy set is mathematical theory for describing the interested variables from uncertain factors or variables like seasonal inflows. The relationship between input and output variables is defined from fuzzy rule, according to human processes in thinking and decision. In addition, fuzzy rules are relatively easy to explain and understand. Recently, the fuzzy model was accepted to describe the relationship of the uncertain variables (Shrestha *et al*., 1996; Panigrahi and Mujumda, 2000; Alsabbah *et al*., 2008; Hassan *et al*., 2010). Often, the calibration processes of the fuzzy

model were performed by manual adjusting (trial and error) the membership functions and rule bases. However, depending on the result of the adjustment, it does not guarantee to yield the optimal solution.

Genetic algorithm is search and optimization techniques based on the principles of national selection and genetics. GA is a robust method for searching for the optimum solution of a complex problem. It can provide the near global optimal solution. The GA was applied to solve the optimal solution of water resource problems (Goldberg, 1989; Chang *et al*., 2005; Kangrang *et al*., 2008; Chen, 2003; Jairaj and Vedula, 2000). The best part of GA is that they can handle any type of objective function. Recently, the genetic algorithms technique is applied to calibrate the membership of the fuzzy model in water resources problem (Kangrang and Chaleeraktrakoon, 2007a; Sivanpheng *et al*., 2009). They had shown that the highest coefficient of determination was provided in the calibration process.

Therefore, this study brings out fuzzy set theory applied to the estimation of forecasting rainfall and GA technique for calibration in the process of modeling fuzzy sets. The improved model was applied to estimate rainfall in the Chi River Basin (in the northeast region of Thailand).

Corresponding Author: Teerawat Thongwan, Department of Civil Engineering, Faculty of Engineering,
Mahasarakham University, Mahasarakham, Thailand

MATERIALS AND METHODS

Model formulation: System inputs include the temperature, humidity, wind speed and solar radiation. Output is the rainfall data. There are four steps for developing fuzzy model as the following.

The first step is to transform the crisp inputs into fuzzy variable through the membership function, called Fuzzification process. The number and type of membership functions are constructed based on statistical data and experience of engineers, generally upon the considering problem. This study used the triangular membership function for describing the input and output variables.

The second step, the fuzzy rule bases are created using the 5-min historical data of all parameters and fuzzy operator. The historical data were collected from the Kosumphisai Meteorological Station as shown in Fig. 1. These fuzzy operators AND and OR are applied to combine the input variables.

The third step is to apply the input membership functions and the rule bases to obtain the output membership functions. This step is done by the implication method which obtaining a fuzzy set of output when given a single number of each inputs. Then the output membership functions of each rule are jointed to one output fuzzy set, called aggregation process.

Finally, the process is defuzzification that a fuzzy set of output is converted into a single crisp value. The most common defuzzification method is the "centroid" evaluation, which returns the center of area under the curve.

The adequacy of the fuzzy model is evaluated by considering the coefficient of determination (R^2) which defined based on the irrigation efficiency estimation errors as:

$$R^2 = \frac{\left(\sum \phi_j \hat{\phi}_j - m\overline{\phi_j}\overline{\hat{\phi}_j}\right)^2}{\left(\sum \phi_j{}^2 - m\overline{\phi_j}^2\right)\left(\sum \hat{\phi}_j{}^2 - m\overline{\hat{\phi}_j}^2\right)} \tag{1}$$

where, ϕ_j is the estimated rainfall of the 5 min duration record j which calculated using fuzzy model, $\hat{\phi}_j$ is the actual rainfall of 5 min duration record j which collected from the Meteorological Station, $\overline{\phi_j}$ and $\overline{\hat{\phi}_j}$ are respectively the average of above mentions and m is the number of data record. The fuzzy model is calibrated by adjusting the membership functions and rule bases using the genetic algorithms technique, these performances will be stopped when the results obtained the highest coefficient of determination (closed to 1.0).

Fig. 1: Location map of the meteorological station

The calibration processes using GA are described as follows. GA requires encoding schemes that transform the decision variables into chromosome. Then, the genetic operations (reproduction, crossover and mutation) are performed. These genetic operations will generate new sets of chromosomes. In this study, each decision variable represents a parameter of membership function. The objective function of the search is to maximize the coefficient of determination (R^2). This study used population size = 80, crossover probability = 0.9and mutation probability = 0.01.

There are 4 cases that were considered in the study including 2-input parameter, 3-input parameter, 4-input parameter and 4-input parameter with duration of rainfall. Illustrative application will be shown in the results section.

RESULTS

2-input parameter consideration: The input parameters were applied to create the fuzzy model only 2 variables. However, all couples of input variables were used to create and test with their historical data record. Table 1 shows coefficient of determination of each couple. The couple of relative humidity and solar radiation provide the highest coefficient of determination as 0.2968.

Table 2 showed the lag time providing the highest coefficient of determination for 2-input consideration. Figure 2 present the coefficient of determination of lag time consideration for 2-input parameter (relative humidity-solar radiation).

Table 1: The coefficient of determination of 2-input consideration

Input variables	R^2
Wind speed-temperature	0.2811
Wind speed-relative humidity	0.2536
Wind speed-solar radiation	0.2449
Temperature-relative humidity	0.2551
Temperature-solar radiation	0.2735
Relative humidity-solar radiation	0.2968

Table 2: The Lag time providing the highest coefficient of determination of 2-input consideration

Input variables	Lag time	R^2
Wind speed-temperature	1 h	0.4072
Wind speed-relative humidity	5 min	0.3916
Wind speed-solar radiation	30 min	0.4233
Temperature-relative humidity	2 h	0.4334
Temperature-solar radiation	30 min	0.4227
Relative humidity-solar radiation	10 min	0.5264

Table 3: The coefficient of determination of 3-input consideration

Input variables	R^2
Wind speed-temperature-relative humidity	0.3321
Wind speed-temperature-solar radiation	0.3062
Wind speed-relative humidity-solar radiation	0.3133
Temperature-relative humidity-solar radiation	0.3359

Table 4: The coefficient of determination of 3-input consideration in case of lag time

Input variables	Lag time	R^2
Wind speed-temperature-relative humidity	12 h	0.7048
Wind speed-temperature-solar radiation	9 h	0.7376
Wind speed-relative humidity-solar radiation	5 min	0.7333
Temperature-relative humidity-solar radiation	2 h	0.8001

Fig. 2: The coefficient of determination of lag time consideration for 2-input parameter

Fig. 3: The coefficient of determination of lag time consideration for 3-input parameter
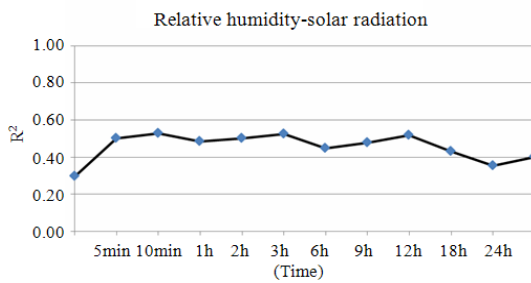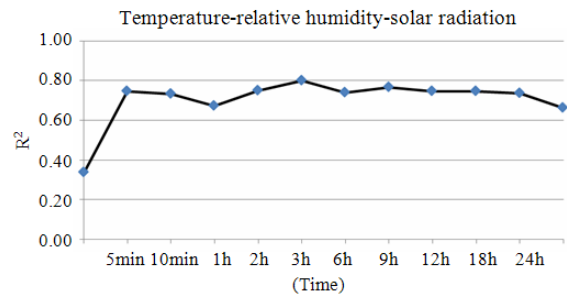

Fig. 4: The coefficient of determination of lag time consideration for 4-input parameter

3-input parameter consideration: Table 3 shows the coefficient of determination of 3-input parameter. The input of temperature-relative humidity-solar radiation provided the highest coefficient of determination as 0.3359. Lag time conditions were performed; the coefficients of determination for all cases were present in Table 4.

Figure 3 present the coefficient of determination of lag time consideration for 3-input parameter (Temperature-Relative Humidity-Solar Radiation). The coefficient of determination of lag time 2 h is the highest.

4-input parameter consideration: The 4-input parameters were applied to create the fuzzy model using the present time of data record. The highest coefficient of determination is 0.3562. When considered by any lag time, it found that the coefficient of determination is increase to 0.8383 at lag time 1 h. Figure 4 presents the coefficient of determination of lag time consideration for 4-input parameter.
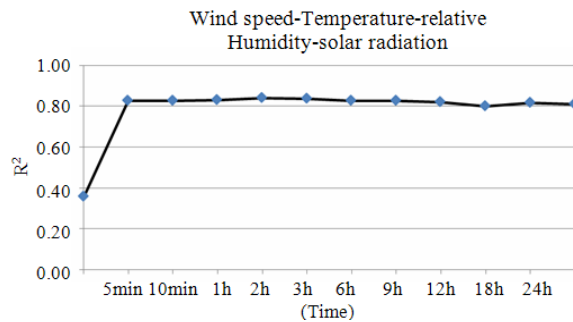
Duration of rainfall consideration: The 4-input parameters were applied to create the fuzzy model using rainfall duration only. There are 3 rain durations were considered in this study; short duration for lower than 15 minutes continuous raining, medium for 15-60 minutes continuous raining and long rainfall duration for over than 60 min continuous raining. They found that the coefficients of determination of all cases are higher than the coefficients of determination of 4-input parameter using continuous rainfall record.

Figure 5-7 show estimated rainfall and actual rainfall for short medium and long rainfall duration respectively. They indicated that the estimated rainfalls are close to the actual rainfall for all cases with the coefficients of determination 0.8732, 0.8622 and 0.8672 respectively.
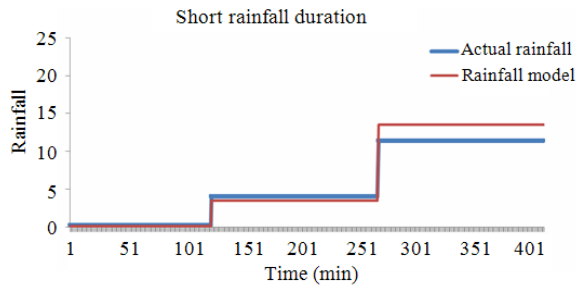
Fig. 5: Estimated rainfall and actual rainfall for short rainfall duration
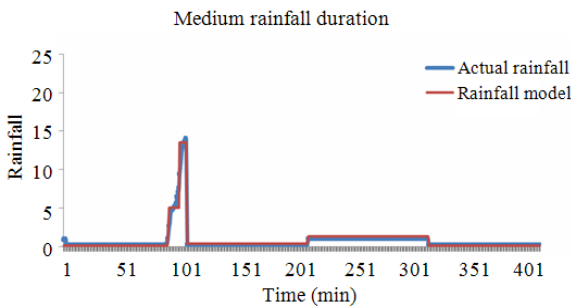

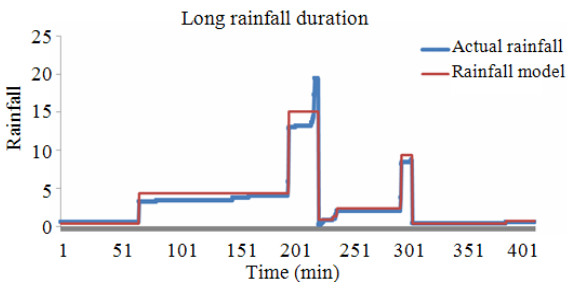Fig. 6: Estimated rainfall and actual rainfall for medium rainfall duration

Fig. 7: Estimated rainfall and actual rainfall for long rainfall duration

DISCUSSION

2-input parameter consideration: As shown in Table 1, the couple of relative humidity and solar radiation provided the highest coefficient of determination because rainfall event may affect from climate and hydrologic variables in the past according to the previous study (Frencha *et al*., 1992), hence the lag time by 5, 10, 30 min, 1-3, 9, 12, 18 and 24 h were performed. They indicated that the highest coefficients of determination of the lag time consideration are higher than the coefficient of determination of the present time in Table 1 for all couples. The coefficient

of determination of lag time consideration for 2-input parameter (relative humidity-solar radiation) were shown in Fig. 2. They present that the coefficient of determination of lag time from 5 min-9 h are not significantly deferent. It can conclude that rainfall event is affected by the previous situation more than the present situation (Frencha *et al*., 1992).

3-input parameter consideration: The 3-input parameters were applied to create the fuzzy model considering the present input data record. It indicated that input of temperature-relative humidity-solar radiation provided the highest coefficient. When considered by any lag time, they found that the coefficients of determination are increase for all cases (Table 4). As shown in Fig. 3, the coefficient of determination of lag time 2 h is the highest. However, the coefficients of determination of lag time over 2 h are slightly decrease.

4-input parameter consideration: The results presented that considering all variables to create fuzzy set model appropriate to estimate rainfall. They also indicated that the coefficients of determination of lag time from 5 minute to 24 h are similarly. Fore this reason, because all parameters are together effect to rainfall. According to the applied fuzzy set to find the interested variable from factors or variables that effected to the interested variable (Shrestha *et al*., 1996; Panigrahi and Mujumda, 2000; Bouhouche *et al*., 2007; Kangrang and Chaleeraktrakoon, 2007b; Alsabbah *et al*., 2008; Hassan *et al*., 2010).

Duration of rainfall consideration: The coefficients of determination of all cases are higher than the coefficients of determination of 4-input parameter using continuous rainfall record as shown in Fig. 5-7. They can conclude that the fuzzy set model considering 4 variables using rainfall duration data is more effective than the model using the continuous rainfall data.

CONCLUSION

This study applied a fuzzy model for estimating the rainfall. Genetic algorithm technique was used to calibrate membership function condition of fuzzy sets model. The improved model considered only a few basic hydrological parameters including temperature, humidity wind speed and solar radiation. The model was applied to estimate the rainfall in the Chi River Basin using 5-minute historic data. Results showed that the fuzzy-GA model can be used to estimate the rainfall, given only the basic hydrological parameters;

temperature, humidity wind speed and solar radiation. The obtained rainfalls of the improved model are close to the rainfall of the actual rainfall record. The fuzzy set model considering 4 variables using rainfall duration data is more effective than the model using the continuous rainfall data.

REFERENCES

Alsabbah, S., T. Mughrabi, S. Alsabbah and T. Mughrabi, 2008. Speed estimation of adaptive fuzzy-controlled piezo-electric motor using MLP-neural network. Am. J. Applied Sci., 5: 1518-1521. DOI: 10.3844/ajassp.2008.1518.1521

Andrieu, H., M.N. Andrieu, V. Thauvin and W.F. Krajewski, 1996. Adaptation and application of a quantitative rainfall forecasting model in a mountainous region. J. Hydrol., 184: 243-259. DOI: 10.1016/0022-1694(95)02977-X

Bouhouche, S., M. Lahreche, A. Moussaoui and J. Bast, 2007. Quality monitoring using principal component analysis and fuzzy logic application in continuous casting process. Am. J. Applied Sci., 4: 637-644. DOI: 10.3844/ajassp.2007.637.644

Chang, F.J., L. Chen and L.C. Chang, 2005. Optimizing the reservoir operating rule curves by genetic algorithms. Hydrol. Proc., 19: 2277-2289. DOI: 10.1002/hyp.5674

Chen, L., 2003. Real coded genetic algorithm optimization of long term reservoir operation. J. Am. Water Res. Assoc., 39: 1157-1165. DOI: 10.1111/j.1752-1688.2003.tb03699.x

Frencha, M.N., W.F. Krajewskia and R.R. Cuykendall, 1992. Rainfall forecasting in space and time using a neural network. J. Hydrol., 137: 1-31. DOI: 10.1016/0022-1694(92)90046-X

Goldberg, D.E., 1989. Genetic Algorithms in Search, Optimizationand Machine Learning. 1st Edn., Addison-Wesley Publishing Company Inc., London. pp: 412.

Hassan, L.H., M. Moghavvemi and H.A.F. Mohamed, 2010. Takagi-Sugeno fuzzy gains scheduled pi controller for enhancement of power system stability. Am. J. Applied Sci., 7: 145-152. DOI: 10.3844/ajassp.2010.145.152

Hormwichian, R, Kangrang, A. And Lamom, A., 2009. A conditional genetic algorithm model for searching optimal reservoir rule curves. Am. J. Applied Sci., 9: 3375-3380. DOI: 10.3923/jas.2009.3575.3580

Jairaj, P.G. and S. Vedula, 2000. Multireservoir system optimization using fuzzy mathematical programming. Water Res. Manage., 14: 457-472. DOI: 10.1023/A:1011117918943

Kangrang, A. and C. Chaleeraktrakoon, 2007a. A Fuzzy-GAs model for determining varied irrigation efficiency. Am. J. Applied Sci., 4: 339-345. DOI: 10.3844/ajassp.2007.339.345

Kangrang, A. and C. Chaleeraktrakoon, 2007b. An estimation of irrigation efficiency of limited water resource area. Am. J. Applied Sci., 7: 2744-2749. DOI: 10.3923/jas.2007.2744.2749

Kangrang, A., Phumphan, A and Chaiyapoom, W, 2008. stochastic inflow simulation for searching rule curves. Am. J. Applied Sci., 5: 221-226. DOI: 10.3844/ajassp.2008.221.226

Otok, B.W. and Suhartono, 2009. Development of rainfall forecasting model in Indonesia by using astar, transfer function and ARIMA methods. Eur. J. Sci. Res., 38: 386-395 http://www.eurojournals.com/ejsr_38_3_05.pdf

Panigrahi, D.P. and P.P. Mujumda, 2000. Research operation modeling with fuzzy logic. Water Res. Manage., 14: 89-109. http://civil.iisc.ernet.in/~pradeep/WRM-panigrahi.pdf

Shrestha, B.P., L. Duckstein and E.Z. Stakhiv, 1996. Fuzzy rule-based modeling of reservoir operation. J. Water Res. Plan. Manage., 122: 262-269. DOI: 10.1061/(ASCE)0733-9496(1996)122:4(262)

Sivanpheng, O., A. Kangrang and A. Lamom, 2009. A varied-utilized soil type in linear programming model for irrigation planning. Am. J. Eng. Applied Sci., 2: 133-138. DOI: 10.3844/ajeassp.2009.133.138