J. Agr. Sci. Tech. (2016) Vol. 18: 1571-1581

Fuzzy-GA Approach for Estimating Rainfall over Upper Chi-Mun Basins of Thailand

A. Kangrang[1*], and W. Jiwlong[1]

ABSTRACT

The present study examines the fuzzy sets model for computing rainfall over the Upper Chi-Mun basins in the Northeastern region of Thailand based on historical weather data from five stations' rain gauges under the radar umbrella, temperature, relative humidity, and radar reflectivity. Data were collected during June 2009 to August 2009 of the rainfall reflectivity record from the Royal Rainmaking Research Centre at Pimai, Nakhon Ratchasima Province, and for the surface rainfall, automatic rain gauges were used. The results showed that the Fuzzy-GAs model could be used effectively to estimate rainfall given only three parameters: temperature, relative humidity and radar reflectivity. Furthermore, the results show that the genetic algorithm calibration provided the optimal conditions of the membership function. The simulation results indicated that the results of the Fuzzy-GA model were close to the observed rainfall data more than the results of a multiple linear regression model for both calibration and validation processes. Consequently, we are confident that a Fuzzy-GA model is a useful tool for estimating rainfall.

Keywords: Fuzzy set, Genetic algorithm, Optimization technique, Radar reflectivity.

INTRODUCTION

Rainfall data is essential in hydrological studies as it constitutes the base of all the computations in water resources management and decision making problems. The obtained rainfall data from a rain gauge is only the point rainfall that has been recorded in a particular time scale. Therefore, due to a lack of the necessary climatological data, it is difficult to obtain a reliable rainfall measurement. In addition, some parameters are difficult to measure and record in small meteorological stations. However, other parameters, such as temperature, relative humidity, and radar reflectivity are easily measured and collected by small meteorological stations. The advantage of using radar information is to obtain rainfall information for the ungauged areas, such as mountainous remote areas where there are problems of rain gauge scarcity and unavailability of power supply. However, there are radar stations distributed all over the country, and the Thai Meteorological Department provides the radar information of each station via its website. It seems to be more reliable to use only these data to estimate the rainfall by a new technique.

A fuzzy set is a mathematical theory for describing the variables of interest from uncertain factors, or variables, such as temperature, relative humidity and radar reflectivity from a radar image. The relationship between input and output variables is defined from a fuzzy rule, according to human processes in thinking and decision making. In addition, fuzzy rules are relatively easy to explain and understand. Recently, a fuzzy model was accepted to describe the relationship between uncertain variables (Zadeh, 1998;

[1] Department of Civil Engineering, Faculty of Engineering, Mahasarakham University, 44150, Thailand.
* Corresponding author; E-mail: anongrit.k@msu.ac.th

[DOR: 20.1001.1.16807073.2016.18.6.6.4]

Jairaj and Vedula, 2000). Often, the calibration processes of the fuzzy model were performed by manual adjustment (trial and error) of the membership functions and rule bases. Therefore, depending on the result of the human adjustment, there is no guarantee to yield the optimal solution. Moreover, the membership function in the fuzzy set theory (Zadeh, 1965) is a new concept rather than a traditional mathematical way to describe a problem.

Genetic Algorithms (GA) have advantages over classical optimization methods (Holland, 1975; Goldberg, 1989), and they have become one of the most widely used techniques for solving a number of hydrological and water resource problems (Wang, 1991; Franchini, 1996; Kangrang and Chaleeraktrakoon, 2007). The best part of GAs is that they can handle any type of objective function describing decision variables. The fuzzy-GA model can be used to estimate the rainfall, given only the basic hydrological parameters. (Thongwan *et al*., 2011)

In this study, the aim was to employ a Fuzzy-GA and multiple linear regression model collectively to estimate the rainfall, with the focus on the two different models. First, the rainfall estimation using the GA technique was applied to calibrate the membership of the fuzzy model. Second, the rainfall estimation using the multiple linear regression model was adopted to evaluate the rainfall.

MATERIALS AND METHODS

Model Formulation

The fuzzy sets model and its rule-based system were applied to estimate the rainfall. The input parameters were temperature, relative humidity and radar reflectivity. The output was the rainfall data period. The steps for working were as follows:

Firstly, the input variables were transformed to the fuzzy variable through the membership function. The number and type of membership functions were constructed based on statistical data and the experience of the engineers, as normal upon considering the problem (Saruwatari and Yomota, 1995; Jang *et al.*, 1997). The fuzzy sets with trapezoidal membership functions were used to describe all parameters due to their easy computation. Secondly, the fuzzy rule bases were created using 3-hourly climatological data and fuzzy operators. These fuzzy operators including AND and OR were applied to combine the input variables. In addition, the input membership functions and the rule bases were applied to obtain the output membership functions. This step was done by the implication method that obtains a fuzzy set of outputs when given a single number for each input. Then the output membership functions of each rule were joined to one output fuzzy set. Finally, a fuzzy set of the output was converted into a single crisp value using the centroid method.

The calculated rainfall from the fuzzy and multiple linear regression models were used to evaluate the obtained rainfall. The adequacy of the fuzzy model was evaluated by considering the coefficient of determination (R^2), which is defined based on the estimated rainfall as:

$$R^2 = \frac{\left[\sum Rm_i Ra_i - n\overline{Rm_i}\,\overline{Ra_i}\right]^2}{\left[\sum Rm_i^2 - n\overline{Rm_i}^2\right]\left[\sum Ra_i^2 - n\overline{Ra_i}^2\right]} \quad (1)$$

Where, Rm_i is the estimated rainfall using the fuzzy model of day *i*, Ra_i is the observed rainfall of day i, $\overline{Rm_i}$ and $\overline{Ra_i}$ are, respectively, the average of those as mentioned above, and n is the number of daily data points. The fuzzy model was calibrated by adjusting the membership functions and rule bases using the GA technique, which will be stopped when the results obtain the highest coefficient of determination (close to 1.00).

The calibration processes using the GA were as follows. The GA technique requires encoding schemes that transform the decision variables into chromosomes. The decision variables were the typical membership function of each type. Equations (2) and (3) show the typical membership function of the trapezoidal type. The trapezoidal curve is a function of a vector x, and depends on four scalar parameters a, b, c, and d, as given by:

$$\mu_A(x,a,b,c,d)=\begin{Bmatrix} 0, & x\le a \\ \frac{x-a}{b-a}, & a\le x\le b \\ 1, & b\le x\le c \\ \frac{d-x}{d-c}, & c\le x\le d \\ 0, & d\le x \end{Bmatrix} \quad (2)$$

or, more compactly, by :

$$\mu_A(x,a,b,c,d)=\max\left(\min\left(\frac{x-a}{b-a},1,\frac{d-x}{d-c}\right),0\right) \quad (3)$$

Where, $\mu_A(x)$ is the membership function of value x for the input or output variable (i.e., temperature, relative humidity, and radar reflectivity). The parameters a and d locate the "feet" of the trapezoid and the parameters b and c locate the "shoulders." They show that the decision variables of each membership function for one group are a, b, c and d. These variables were transferred into the chromosome for searching in the process of the GA. Then, the genetic operations (reproduction, crossover, and mutation) were performed. These genetic operations generated new sets of chromosomes. The objective function of the search was to maximize the R^2. This study used a population size of 80, Crossover probability= 0.9, and Mutation probability= 0.01 (Goldberg, 1989). The search was stopped when the highest coefficient of determination was obtained, hence the optimal values of a, b, c and d were determined. This study considered the number of membership functions of each parameter from two to four groups based on the distribution of historical data.

Data and Methodology

These 3-hourly climatological data, such as temperature and relative humidity, of five rain gauges from Thai Meteorological Department Stations and the radar reflectivity from the Department of Royal Rainmaking and Agricultural Aviation at the Pimai site, Nakorn Ratchasima Province (RRM) were used in the study. The meteorological stations were sta.387401 Mahasarakham, sta.403201 Chaiyapum, sta.405201 Roi Et, sta.431201 Nakorn Ratchasima and sta.436201 Burirum, most of which are located in the Northeastern region of Thailand. The locations are presented in Figure 1.

Radar reflectivity was obtained by an image analysis GIS extension tool from rain events that occurred in the Northeastern region of Thailand during June to August 2009 for long rainfall. The reflectivity was recorded from the Department of Royal Rainmaking and Agricultural Aviation at the Pimai site, Nakorn Ratchasima Province, which corresponds to 2.5 km of CAPPI radar products at the Pimai site from a *S*-band polarmetric radar that transmits radiation with a wavelength of 10.7 cm and produces a beam width of 1.2 degrees with a maximum range of 480 km, as illustrated in Table 1. The CAPPI products have six minutes and 1 sq km temporal and spatial resolutions, respectively. This study assumes that there is no bias caused by the bright band effect and different observation altitudes at 1.5 km (Compliew and Khuanyuen, 2003). The CAPPI data lies within 240 km of the radar, and reflectivity values less than 10 dBz and greater than 50 dBz were excluded from the analysis to avoid the effect of noise in the measured radar reflectivity.

The rainfall from the Thai Meteorological Department (TMD) rain gauges was measured every three hours, whereas the reflectivity from the radar was measured every six minutes. Thus, a normalized process for radar reflectivity should be done prior to the Buffer Probability Technique (BPT) process and be averaged into three hourly data blocks with the

[DOR: 20.1001.1.16807073.2016.18.6.6.4]

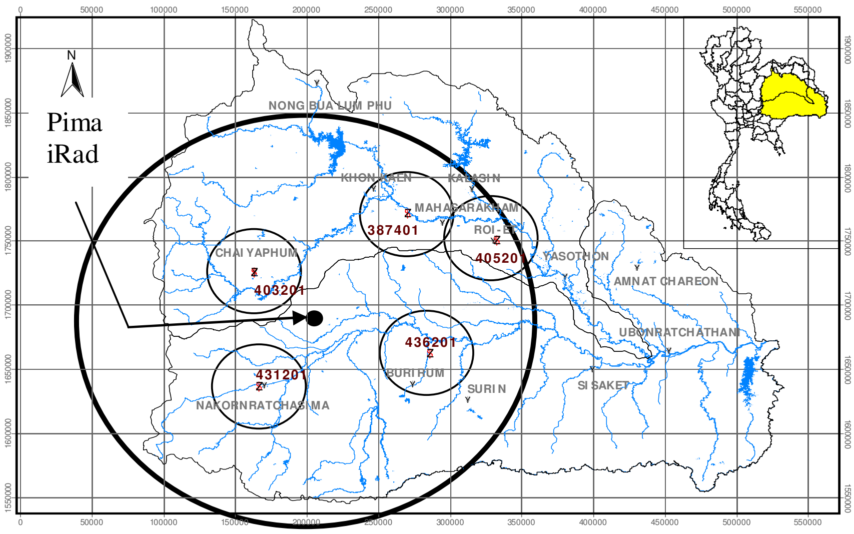

Figure 1. Locations of the five meteorological stations and radar umbrella.

Table 1. Characteristics of radar at Pimai site.

Details of radar	Characteristics
Type of radar	Doppler weather surveillance Radar model DWSR-8500 S, S band
Wave length (cm)	10.7
Beam width (Degree)	1.2
Pulse length (Microsecond)	0.8
Resolution of record data	1 degree×1 degree×1 km
Maximum transmission power (Kw)	850
Maximum range (km)	480
Sequence of elevation angles	Operation A: 0.8, 1.7, 2.5 Operation B: 3.4, 4.2, 5.1, 6.0, 7.4, 9.2, 11.6, 14.8, 18.4, 22.0

same time interval as the rainfall measurement. The BPT was developed with the assumption that raindrops may not fall vertically into the rain gauge because of the wind effect. To decrease this error, the dBz value, which corresponds to the R value at time t, can be calculated from the arithmetic mean of the dBz values detected at time t within a one-km buffer area over the rain gauge (Piman *et al.*, 2007; Tantanee *et al.*, 2008).

These data were used to compute the rainfall by the Fuzzy-GA and the multiple linear regression models: Only temperature, relative humidity, and radar reflectivity were used in the proposed model. The relationship between rainfall and temperature as well as relative humidity and radar reflectivity for five rain gauge meteorological stations during 1 June to 31 August, 2009, indicated that the temperature (22–30°C) affects the occurrence of rainfall, as shown in Figure 2. A low temperature reflected high rainfall. The relative humidity (80–100%) affected the rainfall as shown in Figure 3; high relative humidity indicated high rainfall. The radar reflectivity values were varied (10–50 dBz) as shown in Figure 4, and high radar reflectivity reflected high rainfall. These data were used to set the initial membership functions and rule base of the fuzzy model.

Table 2 shows the fuzzy rule bases using AND and OR operators, such as “If the medium temperature AND the low relative

[DOR: 20.1001.1.16807073.2016.18.6.4]

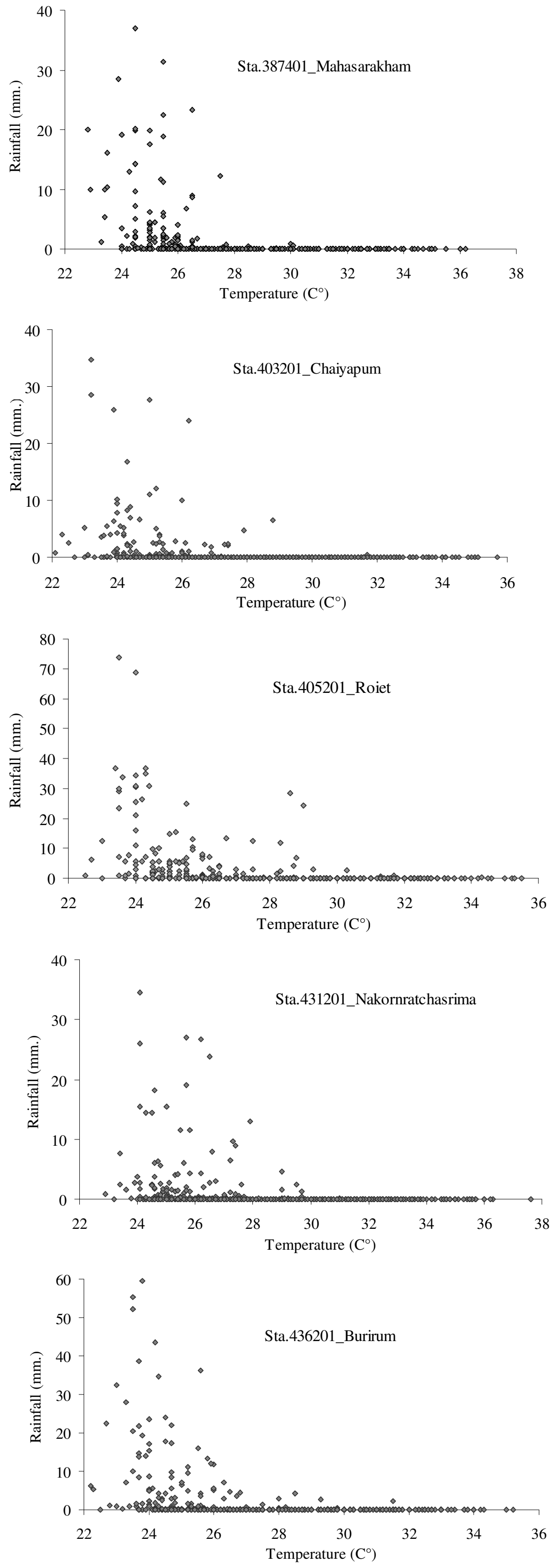

Figure 2. Temperature range effect on occurrence of rainfall for five rain gauge stations.

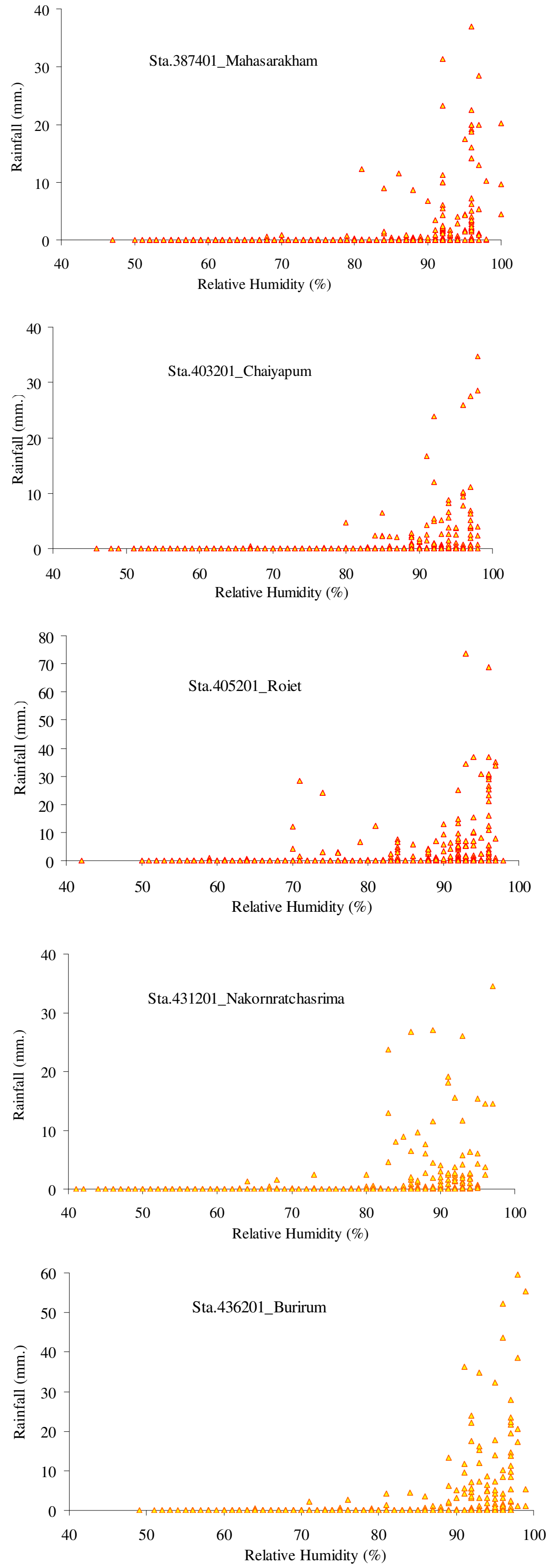

Figure 3. Relative humidity range effect on occurrence of rainfall for five rain gauge stations.

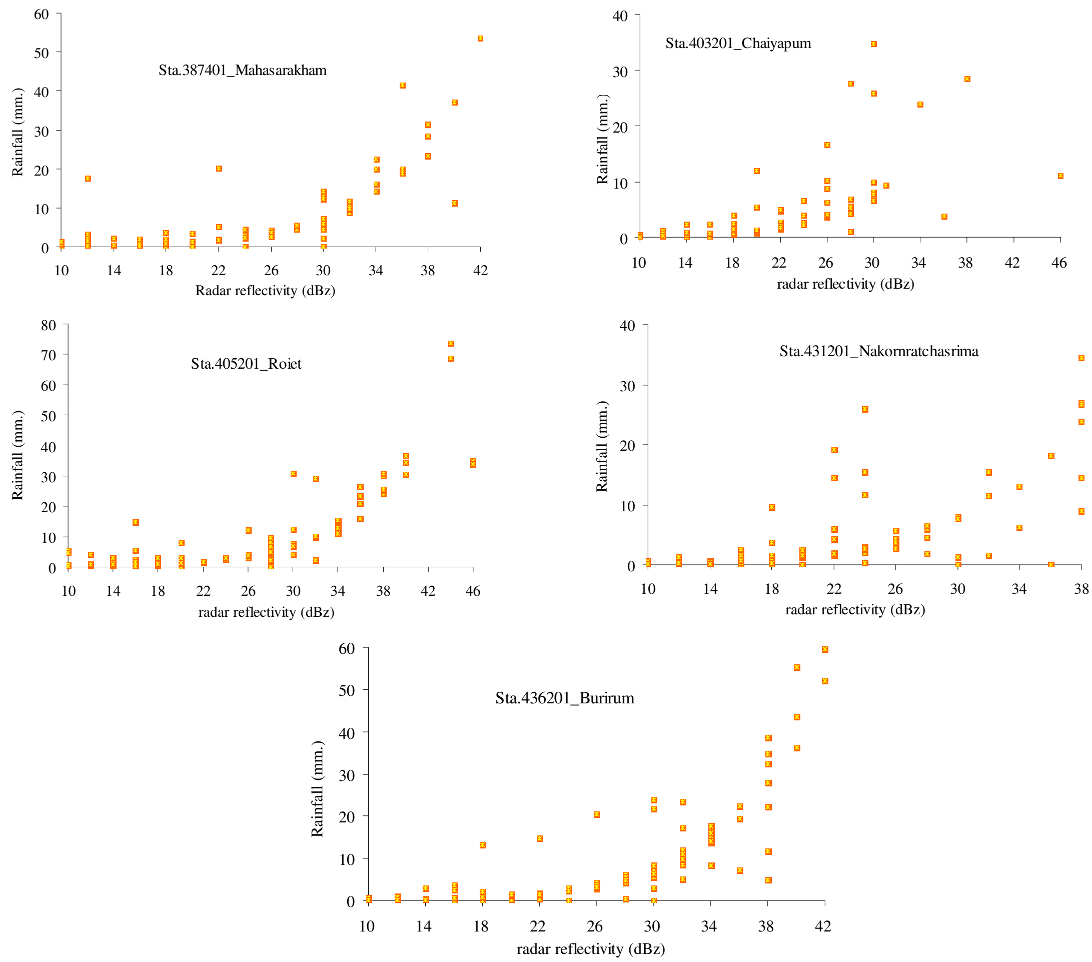

Figure4. Radar reflectivity range effect on occurrence of rainfall for five rain gauge stations.

Table 2. Example of fuzzy rule bases for estimating rainfall.

No	IF Temperature	AND Humidity	AND Reflectivity	THEN Rainfall
1	Low	Low	Low	Low
2	Low	Low	Medium	Medium
3	Low	Low	High	Medium
4	Low	Medium	Low	Low
5	Low	Medium	Medium	Low
6	Low	Medium	High	Medium
7	Low	High	Low	Low
8	Low	High	Medium	Low
9	Low	High	High	Low
10	Medium	Low	Low	Low
11	Medium	Low	Medium	Medium
12	Medium	Low	High	High
13	Medium	Medium	Low	Low
14	Medium	Medium	Medium	Medium
15	Medium	Medium	High	High
16	Medium	High	Low	Low
17	Medium	High	Medium	Low
18	Medium	High	High	Medium

humidity AND high reflectivity THEN the rainfall is high", or "If the medium temperature AND the low relative humidity AND low reflectivity THEN the rainfall is low." These eighteen rule bases were used to construct the relationship between the input and output parameters.

RESULTS AND DISCUSSION

Model Calibration

To determine the optimal model parameters, the Fuzzy-GA model was calibrated with the 3-hourly climatological data that presents the R^2 of several membership functions using the GA calibration with the number of optimization groups for each parameter. The period of the calibration process was from 1 June to 31 August, 2009.

Table 3 presents the input parameters for optimal number of membership functions and the R^2 for each station. Seven sets of four input parameters were applied to create the fuzzy model as rainfall (rain), temperature (temp), Relative humidity (Rh) and radar reflectivity (dBz). The results indicated that the R^2 of the number 3-3-3-3 of the trapezoidal membership function were the highest values of 0.877, 0.895, 0.912, 0.904 and 0.938 for sta.387401, sta.403201, sta.405201, sta.431201 and sta.436201, respectively. It was found from these results that the optimal number and shape of membership functions give the highest R^2. Hence, this condition was accepted to calculate the rainfall. However, the accepted model needed to be verified with other data for evaluating the performance of the model.

Model Verification

The calibrated model parameters from the previous section were further validated using another climatological data set recorded from 1 June to 31 October, 2011, which was not considered during the calibration process. These data from the five stations were used to check the rainfall from the multiple linear regression and the Fuzzy-GA models.

Figure 5 shows the forecasted rainfall by Fuzzy-GA and Multiple linear regression models and the observed rainfall at Mahasarakham, Chaiyapum, Roiet, Nakhon Ratchasima, and Burirum stations. It is apparent that the curves for all station are similarly distributed. In addition, Figure 6 shows the forecasted accumulated rainfall by Fuzzy-GA and Multiple linear regression models and the observed accumulated rainfall at Mahasarakham, Chaiyapum, Roi Et, Nakorn Ratchasima and Burirum stations, where the accumulated rainfall for each station is 1,070.1, 952.9, 1,268.2, 876.0

Table 3. Optimal number of membership function and coefficient of determination (R^2) for each station.[a]

Parameter	Sta.387401		Sta.403201		Sta.405201		Sta.431201		Sta.436201	
	MF	R^2	MF	R^2	MF	R^2	MF	R^2	MF	R^2
Rain-Temp	2-3	0.410	2-3	0.561	2-3	0.550	2-3	0.423	2-3	0.524
Rain-dBz	3-3	0.404	3-3	0.604	2-3	0.524	2-3	0.321	3-3	0.384
Rain-Rh	3-3	0.873	3-3	0.870	3-3	0.805	3-3	0.840	3-3	0.855
Rain- Temp-dBz	2-2-3	0.407	2-2-3	0.591	2-2-3	0.585	2-2-2	0.420	2-2-3	0.502
Rain-Temp-Rh	2-3-3	0.869	2-3-3	0.876	2-3-3	0.818	2-3-2	0.775	2-3-2	0.709
Rain-dBz-Rh	2-3-3	0.870	2-3-3	0.878	3-3-3	0.890	3-2-3	0.872	3-3-3	0.865
Rain-Temp-dBz-Rh	3-3-3-3	0.877	3-3-3-3	0.895	3-3-3-3	0.912	3-3-3-3	0.904	3-3-3-3	0.938

[a] MF= Number of Membership Function; Temp= Temperature; dBz= Reflectivity, Rh= Relative humidity.

[DOR: 20.1001.1.16807073.2016.18.6.4]

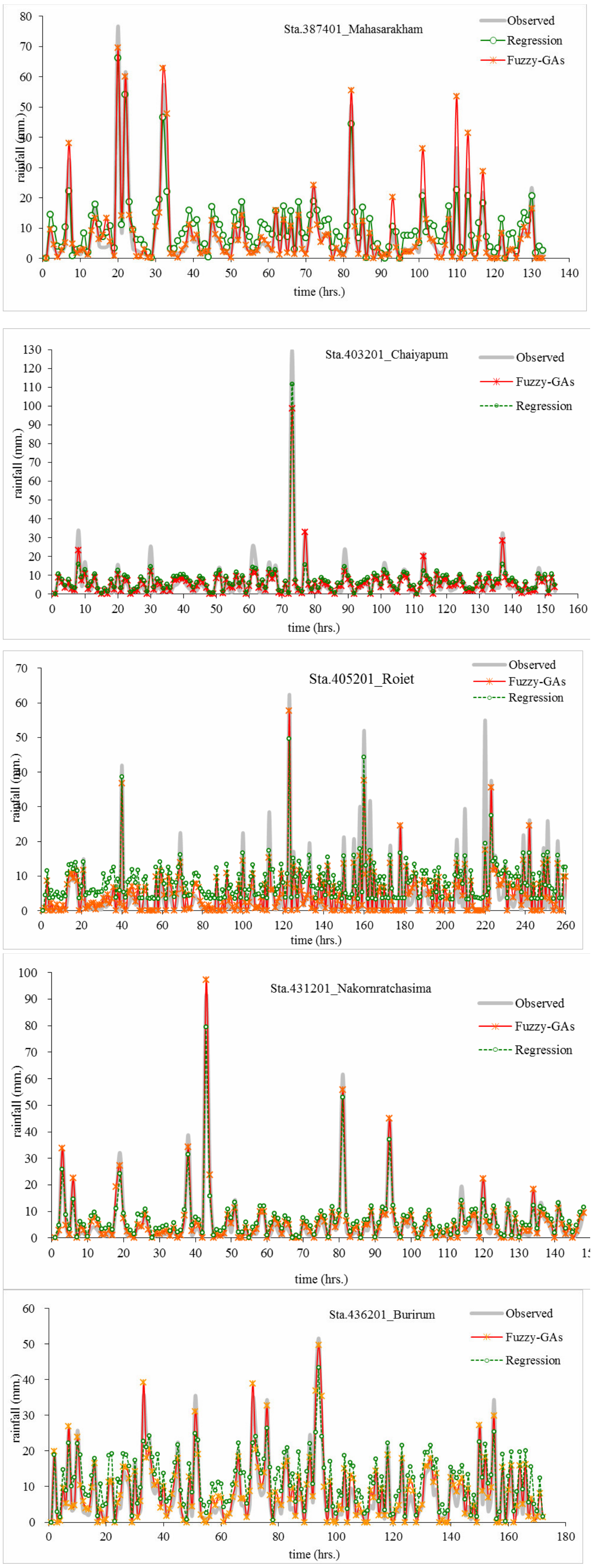

Figure 5. Observed and computed rainfall from Fuzzy-GA and multiple linear regression models for each station.

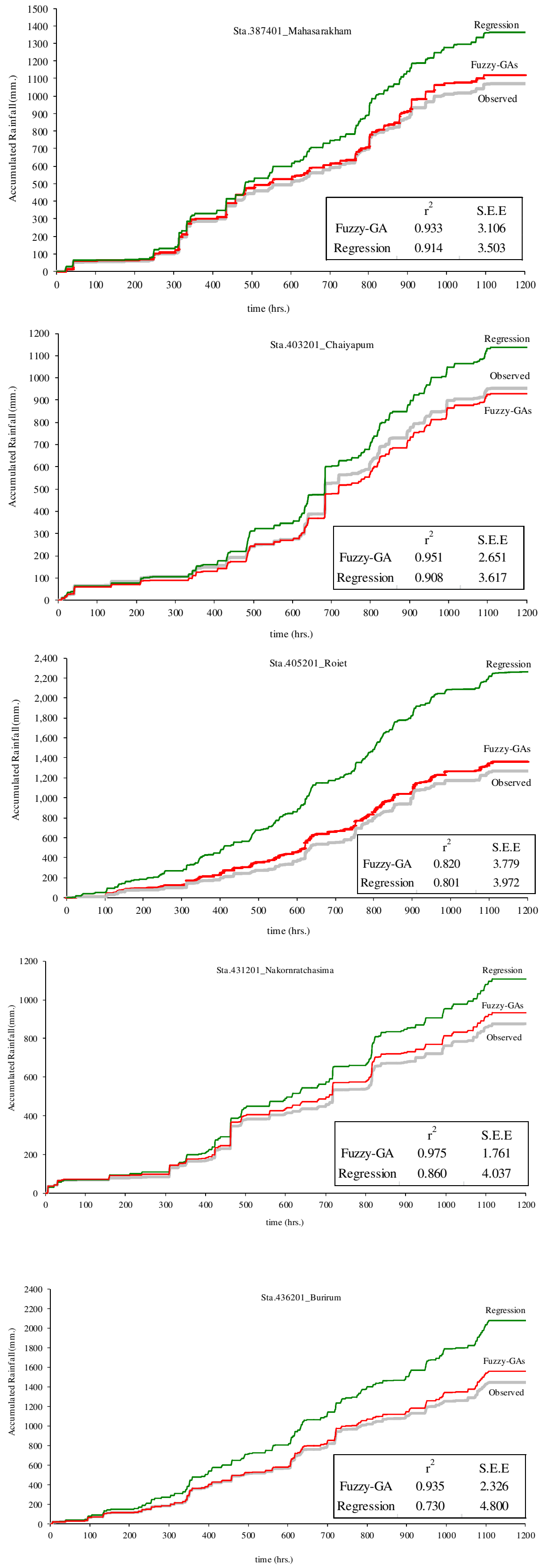

Figure 6. Observed and computed accumulated rainfall from Fuzzy-GA and multiple linear regression models for each station.

and 1,446.3 mm, respectively.

It can be seen that the computed rainfalls using the Fuzzy-GA and multiple linear regression models for the five stations were very close to the observed rainfall based on statistical performance indicators: Standard Error of Estimate (SEE) and R^2.

The S.E.Es for sta.387401, sta.403201, sta.405201, sta.431201 and sta.436201 computed by the multiple linear regression model (Table 4) were 3.503, 3.617, 3.972, 4.037 and 4.800, while when computed by the Fuzzy-GA, they were 3.106, 2.651, 3.779, 1.761 and 2.326, respectively.

The R^2 values for sta.387401, sta.403201, sta.405201, sta.431201 and sta.436201 computed by the multiple linear regression model were 0.914, 0.908, 0.801, 0.960 and 0.730, while the corresponding values for the Fuzzy-GA were 0.933, 0.951, 0.820, 0.975 and 0.935, respectively.

This indicated that the calculations by the two methods were close to each other, with the R^2 between 0.801–0.975. Values of R^2 were higher than 0.800 for all stations. From the verification results, it can be confirmed that the calibration of the model using the Fuzzy-GA technique is suitable to be used to calculate the rainfall. The rainfall calculations by the Fuzzy-GA model were closer to the observed rainfall than with the multiple linear regression model. There was an apparent excellent modelling case at sta. 431201 Nakorn Ratchasima, as shown by the statistical results, which are in agreement with the graphical results. The R^2 values and *SEE* for the Fuzzy-GA model validation were 0.975 and 1.761, respectively. In addition, poor modelling at sta. 405201 Roi Et, as shown by the statistical results, is in agreement with the graphical results. The R^2 values and *SEE* for Fuzzy-GA model validation were 0.820 and 3.779, respectively.

CONCLUSIONS

To estimate rainfall using a Fuzzy-GA model, the input parameters consisting of 3-hourly climatological data including the temperature, relative humidity, and radar reflectivity from five meteorological stations located in the Northeastern region of Thailand were used in this study.

Results show that the rainfall data from the Fuzzy-GA model were closer to the observed data than the results of the multiple linear regression model. Comparing these two models of rainfall estimation, the obtained R^2 from both processes were slightly different, with the Fuzzy-GA model providing a lower standard error of estimation than the multiple linear regression model. Therefore, we are confident that a Fuzzy-GA model is a useful tool for estimating rainfall.

Finally, we propose that further research should be conducted to study the application of fuzzy operators and optimal rule base size. Moreover, more data sets should be tested to predict the suitable membership function in the prediction of the time series data.

Table 4. Comparison of regression and Fuzzy-GAs model for each station in verification process.

Station code	Station name	Regression model		Fuzzy-GA model	
		R^2	Standard Error of Estimate (SEE)	R^2	Standard Error of Estimate (SEE)
387401	Mahasarakham	0.914	3.503	0.933	3.106
403201	Chaiyapum	0.908	3.617	0.951	2.651
405201	Roi Et	0.801	3.972	0.820	3.779
431201	Nakorn Ratchasima	0.860	4.037	0.975	1.761
436201	Burirum	0.730	4.800	0.935	2.326

[DOR: 20.1001.1.16807073.2016.18.6.6.4]

ACKNOWLEGEMENTS

The authors would like to express their appreciation to the Thai Meteorological Department and the Bureau of Royal Rainmaking and Agricultural Aviation for supplying some hydrologic data used in this research.

REFERENCES

1. Compliew, S. and Khuanyuen, B. 2003. The Relation of Radar Reflectivity and Ground Rainfall in the Northeast of Thailand. *Agricultural Engineering 4th Proceeding*, KU Home Kasetsart University, Bangkok, Thailand.
2. Franchini, M. 1996. Using a Genetic Algorithm Combined with a Local Search Method for the Automatic Calibration of Conceptual Rainfall-Runoff Models. *Hydrolog. Sci. J.*, **41(1):** 21-40.
3. Goldberg, D. E. 1989. *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison-Wesley Publishing Company Inc., London.
4. Holland, J. H. 1975. *Adaptation in Natural and Artificial Systems*. University of Michigan Press, Ann Arbor, MI.
5. Jairaj, P. G. and Vedula, S. 2000. Multireservoir System Optimization Using Fuzzy Mathematical Programming. *Water Resour. Manag.,* **14(6):** 457-472.
6. Jang, S. R., Sun, C. T. and Mizutani, E. 1997. *Neuro-Fuzzy and Soft Computing*. Prentice-Hall Inc., USA.
7. Kangrang, A. and Chaleeraktrakoon, C. 2007. Genetic Algorithms Connected Simulation with Smoothing Function for Searching Rule Curves. *Am. J. Appl. Sci.,* **4(2):** 73-79.
8. Piman, T., Babel, M. S., Gupta, A. D. and Weesakul, S. 2007. Development of a Window Correlation Matching Method for Improved Radar Rainfall Estimation. *Hydrol. Earth Syst. Sc.,* **11:** 1361-1372.
9. Saruwatari, N. and Yomota, A. 1995. Forecasting System of Irrigation Water on Paddy Field by Fuzzy Theory. *Agric. Water Manage.,* **28:** 163-167.
10. Tantanee, S., Prakarnrat, S., Polsan, P. and Weesakul, U. 2008. Estimation of Rainfall-Radar Reflectivity Relationship Using Buffer Probability Technique (BPT). *Proceeding 4th IASME/WSEAS International Conference on Energy, Environmental, Ecosystems and Sustainable Development (EEESD'08)*, PP. 460-466.
11. Thongwan, T., Kangrang, A. and Homwuttiwong, S. 2011. An Estimation of Rainfall Using Fuzzy Set-genetic Algorithms Models. *Am. J. Engg. Appl. Sci.,* **4(1):** 77-81.
12. Wang, Q. J. 1991. The Genetic Algorithm and Its Application to Calibrating Conceptual Rainfall-runoff Models. *Water Resour. Res.,* **27(9):** 2467-2471.
13. Zadeh, L. A. 1965. Fuzzy Sets. *Information Control.*, **8:** 338-353.
14. Zadeh, L. A. 1998. Roles of Soft Computing and Fuzzy Logic in Conception Design and Deployment of Information/Intelligent Systems. *Adv. Study Inst. Soft Comput. Appl. Antalya*, **162:** 1-9.

روش فازی-الگوریتم ژنتیکی (Fuzzy-GA) برای برآورد باران در بالا دست حوضه های چی مون در تایلند

ا. کانگرانگ، و. و. جیولنگ

چکیده

هدف این پژوهش بررسی مدل فازی برای محاسبه بارندگی بر مبنای آمار چند ساله باران سنج های تحت پوشش رادار، درجه حرارت، و نم نسبی و ضریب انعکاس رادار(radar reflectivity) از ۵ ایستگاه هواشناسی در بالادست حوضه های چی مون در شمال شرقی تایلند بود. آمار ضریب انعکاس رادار در طی ماه ژوئن ۲۰۰۹ تا اوت ۲۰۰۹ از مرکز سلطنتی تحقیقات باران سازی در منطقه Pimai در استان Nakhon Ratchasimaبرداشت شد و برای باران های سطحی از باران سنج های خودکار استفاده شد. نتایج نشان داد که با کار برد مدل های Fuzzy-GA و در دست داشتن فقط سه پارامتر درجه حرارت ، نم نسبی، و ضریب انعکاس رادار می توان مقدار باران را به طور موثری برآورد کرد. همچنین، نتایج حاکی از آن بود که برای تابع عضویت (membership function)، واسنجی الگوریتم ژنتیکی شرایط بهینه را فراهم می کند. یافته های این بررسی نمایانگر آن بود که برای هر دو فرایند واسنجی و راستی آزمایی، نتایج شبیه سازی با مدلFuzzy-GA در مقایسه با نتایج مدل رگرسیون خطی چند گانه، به اعداد اندازه گیری شده نزدیک تر بود. در نتیجه، اطمینان داریم که مدل Fuzzy-GA ابزار مفیدی برای برآورد باران است.