

Parametric Studies of ANFIS Family Capability for Thunderstorm Prediction

Wayan Suparta and Wahyu Sasongko Putro

Abstract Thunderstorms are an unpredictable natural hazard. They affect daily human life especially for space launcher development program in the near future. The variations in meteorological parameters were used to capture thunderstorm activity. In this study, six parameters, namely pressure, temperature, relative humidity, cloud, rainfall and precipitable water vapor were analyzed in order to develop a thunderstorm prediction system. To realize the development of this thunderstorm prediction system, we developed a thunderstorm prediction model based on the Adaptive Neuro-fuzzy Inference System (ANFIS) family (ANFIS FCM, ANFIS FSC, and ANFIS Human Expert). Three models from the ANFIS family were assessed to ascertain their capability for thunderstorm prediction. Input and output variables were taken from the Tawau meteorology station. The results showed that the thunderstorm prediction model based on ANFIS Human Expert showed a good efficiency with an estimated error prediction of <2% with root mean square error (RMSE) and percentage error (PE) values of 3.028% and 23.545% respectively compared to RMSE and PE of ANFIS FCM and ANFIS FSC.

1 Introduction

Unpredicted natural hazards pose a risk to human life. High thunderstorms activity is composed of pressure, temperature, relative humidity, and cloud (oktas) parameters, etc. Fluctuations in atmospheric variability will particularly affect the environmental thunderstorm level in region such as in Tawau, Sabah. The thunderstorm level in this area increases during the intermonsoon period (from summer

W. Suparta (✉)

Dept of Electrical Engineering, Sanata Dharma University, 55282 Yogyakarta, Indonesia

e-mail: drwaynesparta@gmail.com; wayan@ukm.edu.my

W. Suparta · W.S. Putro

Space Science Centre (ANGKASA), Institute of Climate Change,

Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor Darul Ehsan, Malaysia

to winter) [1]. The frequency of rainfall has increased since September 2014) [2] and this has resulted in flash flood events. To anticipate the thunderstorm damage in this area, many researchers have studied the atmospheric conditions in order to characterize and develop a thunderstorm prediction model.

A preliminary study of atmospheric characterization over the Tawau area was successfully conducted by Suparta and Putro [1] to classify thunderstorm activity. The thunderstorm activity increased during the intermonsoon periods (from winter to summer and summer to winter). However, in winter and summer the thunderstorm activity decreased by 10~20% due to climate change effects on the seasonal monsoon in Asia [3]. Loo et al. [4] studied the seasonal monsoon in Asia throughout the twentieth century and discovered a connection between monsoon rainfall and global warming. The increasing rainfall frequencies were caused by Typhoon Saola and then Typhoon Haikui [5]. Due to dangerous effect of thunderstorms, Albar et al. [6] successfully estimated thunderstorm activities by using rainfall radar data from Jeddah, Saudi Arabia. Empirical relationship methods (reflectivity and rainfall rate) were used to assess the rainfall depth and thunderstorm status. However, the empirical relationship methods have a disadvantage due to the lack of area coverage and a low level of detectable signal. In order to improve this shortcoming, Velden et al. [7] used measurements from satellite images (the Dvorak technique) to find thunderstorm status based on cloud evolution. Spiridonov and Čurić also successfully developed a thunderstorm forecasting model to evaluate the status of thunderstorm using a numerical method [8]. However, the numerical method has a disadvantage when large amounts of data are used. Litta et al. [9] successfully developed a thunderstorm forecasting model using Artificial Neural Network (ANN) with Levenberg Marquardt and Delta-Bar-Delta logarithms to improve the training process using the numerical model. Suykens et al. [10] obtains the limitations of the ANN method when it reached a state of convergence (steady).

The capability of ANN decreased at a relatively slow pace before they reached a steady position. However, Rajasekaran and Pai [11] successfully improved the capability of ANN with an added Fuzzy algorithm, namely the Adaptive Neuro-Fuzzy Inference System (ANFIS). Suparta and Alhasa [12] studied two types of ANFIS, namely ANFIS Fuzzy C-Means (FCM) and ANFIS Fuzzy Subtractive clustering (FSC) in atmospheric applications by using surface meteorological data. The estimation result showed that ANFIS FCM is comparable to ANFIS FSC. The disadvantage of ANFIS FCM depends on the situation in question and the application. Cao et al. [13] successfully designed the ANFIS New Fuzzy Reasoning Model (NFRM), called the ANFIS Human Expert (HE), and this has greater capability than ANFIS FCM and FSC. As a result of this problem, ANFIS HE was proposed to estimate and measure the thunderstorm model. In this study, we compare the capability of ANFIS HE with ANFIS FCM and FSC to predict thunderstorm events in Tawau, Sabah, Malaysia.

2 Methodology

2.1 Data and Location

As thunderstorm activity increases over Tawau area during the intermonsoon periods, this region was selected to be an area of study. Stohlgren [14] obtained four meteorological parameters (temperature, wind flow, precipitation and relative humidity) as the source of thunderstorm activity in the Rocky Mountains area (USA) during spring and summer. Based on the different geographical conditions between Tawau and the Rocky Mountains, six meteorological parameters, namely pressure (P), temperature (T), relative humidity (H), cloud intensity (C), precipitation (Pr), and precipitable water vapor from radiosonde (RSPWV) were analyzed to reveal thunderstorm activity. Data from 1 January 2013 to 31 December 2013 taken from the Malaysian Meteorological Department (MetMalaysia), weather underground, NASA, and Wyoming University over Tawau ground-based station (located at 4.32°N, 118.12°E and at an elevation of 17 m) were used to develop a thunderstorm prediction model. The data were processed using Matlab and an error measurement was filtered from the ground-based sensor by replacing it with Not a Number (NaN) data.

2.2 Adaptive Neuro-Fuzzy Inference System (ANFIS)

Based on the conceptual study to improve training process using ANN, the Fuzzy Inference System (FIS) was constructed to develop a degree of membership functions (MFs) of both input and output data. Here, input and output data were clustered to create fuzzification process over FIS. The two types of FIS, Sugeno and Mamdani were applied with a linear technique. In this study, Sugeno FIS were applied to create an ANFIS family model (ANFIS FCM, ANFIS FSC, and ANFIS Human Expert). The main concept of membership function Sugeno FIS has two input rules (A and B) and two consequent parts (f_1 and f_2) of the output rule, where $f_1(p_1, q_1, r_1)$ and $f_2(p_2, q_2, r_2)$ are a premise part of the linear parameter for the Takagi-Sugeno fuzzy model. The sum product of f_1 and f_2 for the final weight is expressed as

$$f = \overline{w_1}f_1 + \overline{w_2}f_2 \quad (1)$$

where f , $\overline{w_1}$ and $\overline{w_2}$ are linear parameter average for the output target, and weight average rule one and two respectively. Furthermore, the premise and consequent parameter of the Takagi-Sugeno concept of the FIS and fuzzy reasoning mechanism is described in Fig. 1. Here, the premise and consequent parameter of the Takagi-Sugeno concept had two input rules and output rules [12].

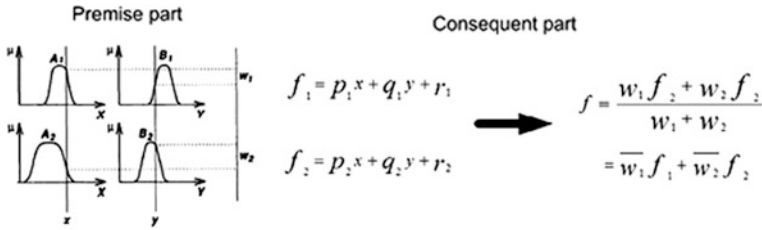


Fig. 1 Fuzzy reasoning mechanism and Takagi-Sugeno FIS

Thus, according to Eq. (1), the Takagi-sugeno ANFIS model as described in Fig. 2 used two rules x and y which is A_1 and A_2 , B_1 and B_2 as MFs of input parameter, respectively [12].

Rule 1 : if x is A_1 and y is B_1 then $f_1 = p_1x + q_1y + r_1$

Rule 2 : if x is A_2 and y is B_2 then $f_2 = p_2x + q_2y + r_2$

Furthermore, the architecture of the ANFIS model consists of five layers. Each layer has its own function to process input data. In the first layer, we generated the MFs for every node with an output parameter and generalized using a Gaussian function. In layer two, we found the output result using an AND operator (T-Norm operator) for each node. In layer three, we calculated the ratio of firing strength to find the best output result. In layer four, we calculate the total output based on the parameter input of all firing strength nodes. In the last layer, the calculation of average signal input from the input parameter was performed. In this study, we used a Gaussian function and three MFs from the selected parameters. The selected parameters were analyzed using Minitab software in order to find the best configuration parameter of input and output.

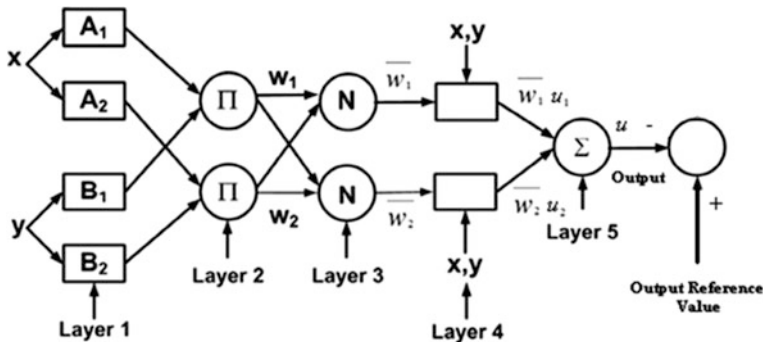


Fig. 2 Adaptive Neuro-Fuzzy Inference System structure

2.3 Capability of ANFIS Family to Predict Thunderstorm Events

As thunderstorms are categorized as a deadly natural hazard, the capabilities of the ANFIS family—FCM, FSC and Human Expert (HE)—were compared. The different concepts of three types of ANFIS are located in the MFs design over FIS. Figure 3 shows the comparison of MFs designed over FIS for input and output parameters. The MFs of (3a) the ANFIS FCM and (3b) FSC models were designed using a cluster generated by a Matlab program. However, the MFs of (3c) the ANFIS HE model were designed using an author expert and a Gaussian function to improve training, testing, and validation results.

The results from training data using ANFIS designed with MFs over FIS were processed to obtain an estimation model. The best estimation model was chosen on the basis of iteration limit and training time. However, in other cases iteration and training will not guarantee a good estimation result. Thus, the RMSE and Percent True were calculated in order to obtain the best estimation result for the testing and

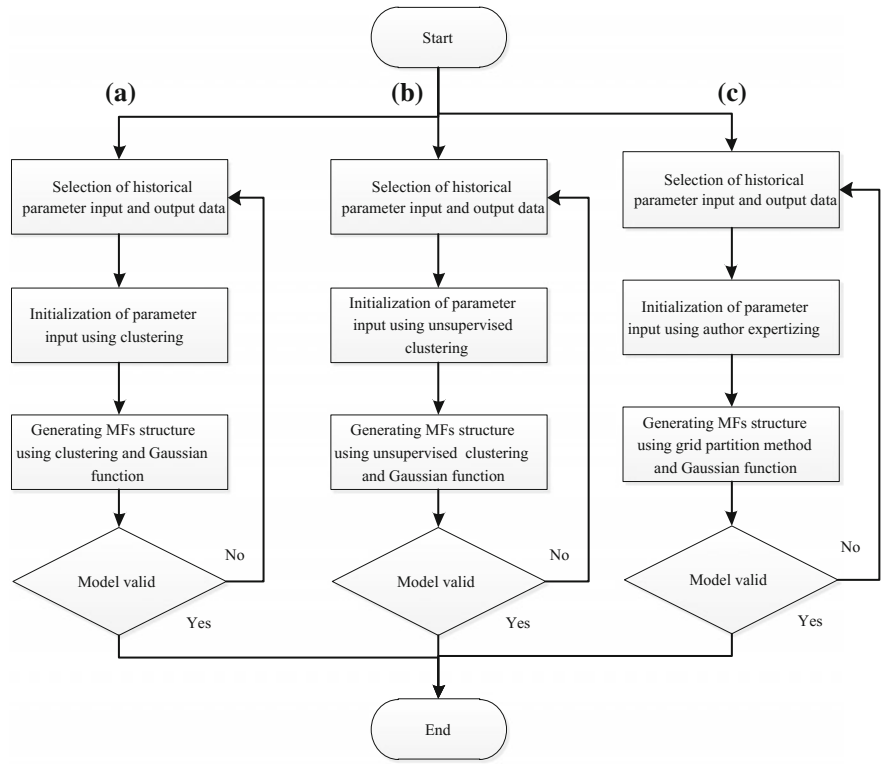


Fig. 3 Flowchart of MFs design of ANFIS families for a FCM, b FSC, and c HE

validation process. In this study, the prediction of thunderstorm events was calculated using an exponent Lyapunov chaotic time series [15] as expressed in Eq. 2

$$\vec{X}(t) = [(x(t), x(t - \tau), \dots, x(t - (m - 1)\tau)] \quad (2)$$

where $\vec{X}(t)$ is the prediction result, t is the scalar index of the estimation result (from the ANFIS model), τ is the time delay and m is the collapse dimension.

3 Results and Discussion

Figure 4 shows the variation of six meteorological parameters over Tawau station used as input parameters to find thunderstorm events. Data from six input parameters due to a lack of measurement data were processed using a Matlab software. Here, blank data is Not a Number (NaN) captured from lack of measurement after processing using Matlab software. In addition, hourly measurement data from the meteorological sensor were averaged on a daily basis and with a target to estimate thunderstorm data using a Matlab software.

During intermonsoon season I (winter to summer) in March, April, and May the RSPWV decreased when the temperature (T) was normal and relative humidity (H) increased by less than 90% respectively during mid-April 2013. However, during intermonsoon season II (summer to winter) in September, October, and

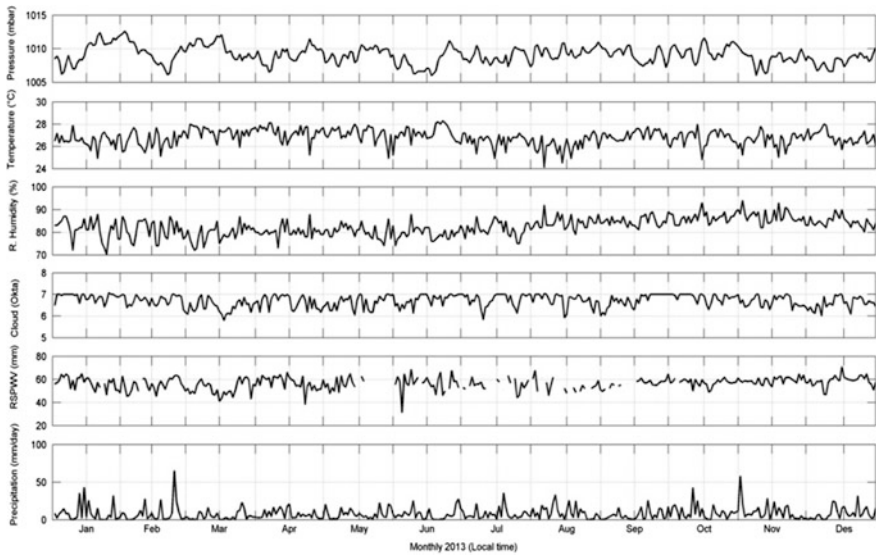


Fig. 4 The variation of meteorological parameters used to find thunderstorm events over Tawau station

November, the T parameter increased when the H parameter decreased and RSPWV data was normal respectively during mid-October 2013. The different situation between intermonsoon seasons I and II was due to a tropical storm event in the south Philippines [1]. Thus, to capture the tropical storm event the configuration method was proposed to calculate the standard error (S-value). Furthermore, six parameter inputs and one output parameter (thunderstorm) were paired into 148 configuration inputs and outputs to calculate the S-value using Minitab software. The best result of configuration parameter input and output are provided in Table 1. Here, the best configuration meteorological parameters are H and RSPWV.

The best configuration parameters (H and RSPWV) were selected to predict thunderstorm activity. During the development of the new model, the six configuration parameters were used to design the MFs using a Gaussian function. Figure 5 presents the degree of MFs surface meteorological data for the ANFIS model. As shown in the figure, the three Gaussian MFs were chosen to design the FIS model on the basis of accuracy, training time, and case study. The FIS was added to layer one of the ANFIS model for the thunderstorm estimation model. During training, testing, and validating of the observation parameters in the ANFIS model, we selected the data into 70%, 10%, and 20%, respectively for input parameter and with the target of thunderstorm activity. Furthermore, the comparison of results between ANFIS FCM, FSC, and HE used to capture thunderstorm activity over 2013 are described in Fig. 6. Here, the ANFIS family models were successful in predicting thunderstorm activity using six configuration input parameters.

Figure 6 shows the results for estimation of thunderstorm activity using ANFIS HE, FCM, and FSC. Generally, all the ANFIS models follow the observation pattern except for May 2013 due to there being no data of RSPWV. A statistical method was applied to obtain the capability of ANFIS HE, FCM, and FSC models (see Table 2).

Table 1 The best configuration parameter for predicting thunderstorm events

Configuration input parameter	Output parameter	Standard error (S-value)
P , T , and H	Thunderstorm	0.431
...	...	...
...	...	...
P , T , H , and C	Thunderstorm	0.438
P , T , H , C , and RSPWV	Thunderstorm	0.439
P , T , H , C , RSPWV, and Pr	Thunderstorm	0.429
H and RSPWV	Thunderstorm	0.427
T , H , and RSPWV	Thunderstorm	0.432
T and H	Thunderstorm	0.431
...	...	...
T , H , and C	Thunderstorm	0.431

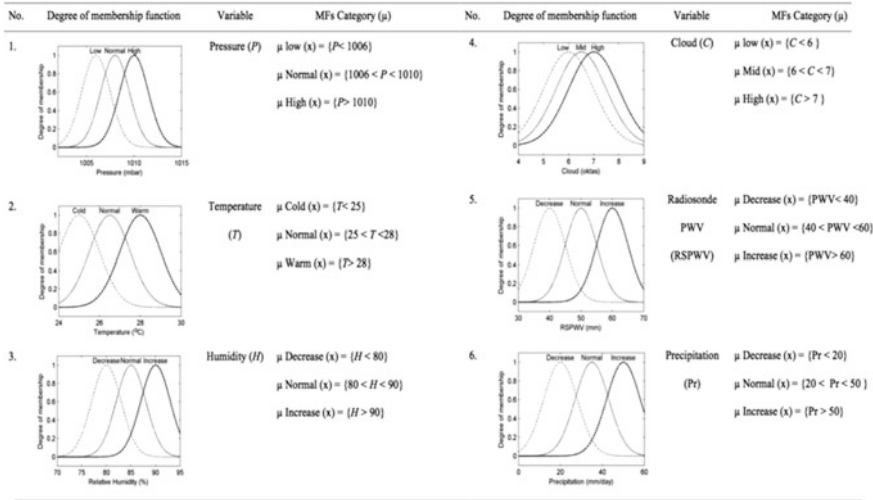


Fig. 5 The degree of MFs meteorological data for the ANFIS model

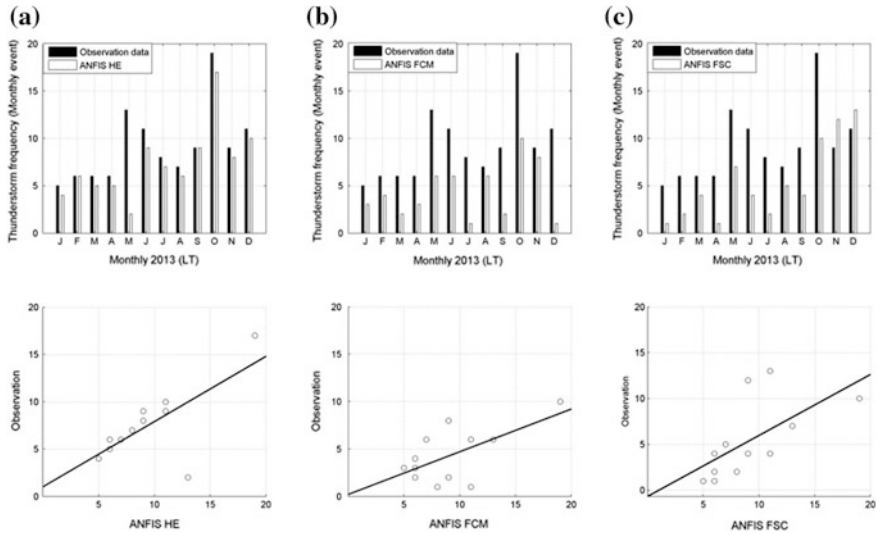


Fig. 6 Comparison of observation data and their relationship in terms of thunderstorm frequency with **a** ANFIS HE, **b** ANFIS FCM, and **c** ANFIS FSC

In order to reach a thunderstorm prediction model, the chaotic time series were applied in the ANFIS model by using five reference data. The comparison of estimation error from the thunderstorm prediction models is compiled in Fig. 7. The figure shows that the thunderstorm value increased every day/event. The trends of estimation error for each step in thunderstorm prediction increased up to eight steps

Table 2 Statistical comparison of each ANFIS family model’s ability to estimation thunderstorm events

ANFIS Model	Correlation (R^2)	RMSE (%)	MAE (%)	PE (%)
HE	0.868	3.028	2.333	23.545
FCM	0.739	4.173	3.250	36.471
FSC	0.582	5.091	4.417	49.276

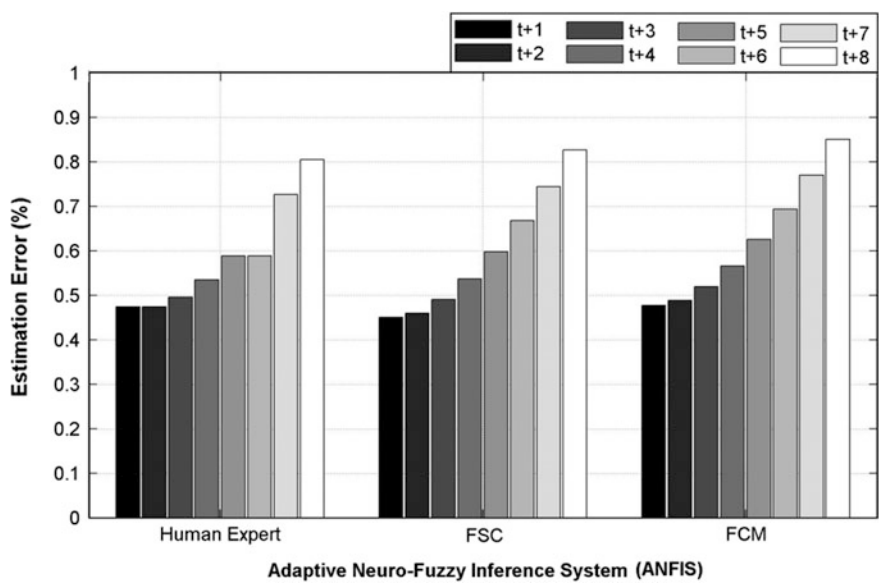


Fig. 7 The comparison of estimation error in thunderstorm prediction using three models from one step to eight steps ahead

ahead (max). The increase in the estimation error was caused by data resolution on a daily basis. Finally, ANFIS HE was used to establish a prediction model with the greatest capability compared with ANFIS FCM and FSC.

4 Conclusion

A comparison of the ANFIS families conducted using meteorological data as the input has the capability to predict thunderstorm activity. The input used was six configurations of meteorological parameters and one output of thunderstorm frequency, and were applied to design MFs over FIS. The estimation result showed that ANFIS HE had a favorable result compared to ANFIS FCM and FSC. Furthermore, the estimation error of the thunderstorm prediction results indicates

that ANFIS HE also shows a good capability to detect thunderstorm events over the Tawau area. However, the ANFIS model looks roughly to follow the trend of observation data in May 2013 due to a lack of RSPWV data. To improve the RSPWV data, it is suggested that it must be constructed using PWV data from GPS data in the near future.

Acknowledgements This research is supported by Ministry of Science, Technology and Innovation Malaysia (MOSTI) under Science Fund 01-01-02-SF1100 grant and partly funded by Flagship Program: ZF-2014-016 grant. The authors would like thank MetMalaysia, weather underground, NASA and Wyoming University for providing the meteorological data used in this study.

References

1. Suparta W, Putro WS, Singh MSJ, Asillam MF (2015) Characterization of GPS and meteorological parameters for mesoscale convective systems model over Tawau, Malaysia. *Adv Sci Lett* 21:203–206
2. Suparta W, Putro WS, Singh MSJ, Asillam MF (2015) The estimation of rainfall and precipitation variation during 2011 convective system using an artificial neural network over Tawau, Sabah. Paper presented at the 2015 IEEE international conference on space science and communication, space science centre (ANGKASA), Universiti Kebangsaan Malaysia, Langkawi, 10–12 August 2015
3. Chen TC, Wang SY, Yen MC (2007) Enhancement of afternoon thunderstorm activity by urbanization in a valley: Taipei. *J Appl Meteorol Climatol* 46:1324–1340
4. Loo YY, Billa L, Singh A (2015) Effect of climate change on seasonal monsoon in Asia and its impact on the variability of monsoon rainfall in Southeast Asia. *Geosci Front* 6:817–823
5. Heistermann M, Crisologo I, Abon CC, Racoma BA, Jacobi S, Servando NT, David CPC, Bronstert A (2013) Using the new Philippine radar network to reconstruct the Habagat of August 2012 monsoon event around Metropolitan Manila. *Nat Hazards Earth Syst Sci* 13:653–657
6. Albar AMH, AL-Khalaf AK, Mohamed HAB (2015) Radar rainfall estimation of a severe thunderstorm over Jeddah. *Atmos Clim Sci* 5:302
7. Velden C, Harper B, Wells F, Beven JL II (2006) The Dvorak tropical cyclone intensity estimation technique. *Bull Am Meteorol Soc* 87:1195
8. Spiridonov V, Curic M (2015) A storm modeling system as an advanced tool in prediction of well organized slowly moving convective cloud system and early warning of severe weather risk. *Asia Pac J Atmos Sci* 51:61–75
9. Litta AJ, Mary Idicula S, Mohanty UC (2013) Artificial neural network model in prediction of meteorological parameters during premonsoon thunderstorms. *Int J Atmos Sci* 1:1–14
10. Suykens JAK (2001) Nonlinear modeling and support vector machine. Paper presented at the 2001 IEEE instrumentation and measurement technology conference, Budapest convention centre, Budapest, Hungary, 21–23 May 2001
11. Rajasekaran S, Pai GV (2004) Neural networks. *Fuzzy Logic and Genetic Algorithms*. PEU Learning Pvt. Ltd, New Delhi, pp 17–456
12. Suparta W, Alhasa KM (2015) Modeling of zenith path delay over Antarctica using an adaptive neuro fuzzy inference system technique. *Expert Syst Appl* 42:1050–1064
13. Cao Z, Kandel A, Li L (1990) A new model of fuzzy reasoning. *Fuzzy Sets Syst* 36:311–325

14. Stohlgren TJ, Chase TN, Pielke RA, Kittel TG, Baron J (1998) Evidence that local land use practices influence regional climate, vegetation, and stream flow patterns in adjacent natural areas. *Glob Change Biol* 4:495–504
15. Xu X, Tang Q, Xia H, Zhang Y, Li W, Huo X (2016) Chaotic time series prediction for prenatal exposure to polychlorinated biphenyls in umbilical cord blood using the least squares SEATR model. *Sci Rep* 6:25005

Space Science and Communication for Sustainability

Suparta, W.; Abdullah, M.; Ismail, M. (Eds.)

2018, IX, 348 p. 223 illus., 181 illus. in color., Hardcover

ISBN: 978-981-10-6573-6