Short Term Weather Forecasting Using Fusion of Fuzzy-Artificial Neural Network

B. Putra
Department of Engineering Physics
Institut Teknologi Sepuluh
Nopember (ITS)
Surabaya, Indonesia
budiman.putra@len.co.id

B.T. Atmaja
Department of Engineering Physics
Institut Teknologi Sepuluh
Nopember (ITS)
Surabaya, Indonesia
btatmaja@gmail.com

S. Hidayat
Department of Electronic and
Informatics Engineering
Universitas Gajah Mada (UGM)
Yogyakarta, Indonesia
syahroni_ee11@mail.ugm.ac.id

Abstract— **Weather forecasting is important need in daily modern life. Until now, the weather forecasting is only shown in long term whereas the short term forecasting is also important for some cases. In this research, the system of short period weather forecasting was designed based on the current weather parameters consisted of temperature, humidity, air pressure, the direction of the wind, wind speed and present weather condition. This system used fusion of Fuzzy Logic, Artificial Neural Network (ANN) combined with *Lavendberg-Marquadt* method and Fuzzy C-Mean (FCM) as analysis and prediction method. Based on the result, this complex method can predict the weather continuously despite the change of unpredictable patterns. This improvement is mainly caused by updating response system based on continuous learning of various cases given. Furthermore, this system has clear reasoning logic on the fuzzy logic instead of its neural network method. The performance of proposed system has accuracy up to 78% for validity among three possible weathers *i.e* shiny, cloudy and rainy.**

Keywords - short term, weather forecasting, fuzzy logic, Fuzzy-C-Means, Artificial Neural Network, Lavendberg-Marquadt.

I. INTRODUCTION

Weather is a part of people daily life. Many people such as farmers, fishermen depend on the weather to do their daily activity. For extreme case, a pilot of an aircraft must take into account the weather for the flight, especially during *take-off* and *landing* because it will affect to the visibility and *take-off/landing* process. So, weather forecasting provides critical information for some cases. In severe weather situations, short-term forecasts and warnings can help to save lives of people and protect property. It is important thing to make accurate weather prediction. Moreover, recently unpredictable of daily weather is a problem and it needs to predict the short-term weather with high degree accuracy.

Many methods are created for weather forecasting. The latest applicable are five following methods: Persistence Method, Trends Method, Climatology Method, Analogue Method and Numerical Weather Prediction Method. Those methods are taken under attention in to approach a model creation. It ensures optimal input data for the system for short-term weather forecasting. Other researchers are work in soft computing methods to create blind prediction based on input data.

It has been reported, the work of following people give contribution in weather forecasting research. Ingsrisawang [9] has investigated implementation of machine learning for rain forecasting. Then, Shahi [1] has developed the using fusion of Fuzzy C-Mean Clustering and Type-2 Fuzzy Logic for short term weather forecasting. The implementation of neural network for building the efficient weather forecasting also has done by Baboo [10] and the probabilistic method has been proposed by Fraley [11]. In this paper, another different system is proposed to improve the performance and accuracy of weather prediction built from the fusion of Neural Network and Fuzzy Inference and C-Means System. Hypothetically, this fusion will be a good method due to the characteristics of each method. Neural Networks are good for recognizing patterns by training way but it is not good in explaining how it reaches its decisions. In the other side, Fuzzy logic system is good for explaining its decisions but it cannot automatically acquire the rules. Thus, by combining both methods, the forecasting system can be built to give more valid result.

II. THEORITICAL REVIEW

A. Weather Prediction

Rain is one form of precipitation that forms when water droplets fall to the earth's surface from clouds. Phase change experienced by the water into water vapor and water back in the form of precipitation can be explained by the Psychometric chart.

Many methods have been made to forecast rain, one of which is the analogue method. This method is based on the fact that information on a *weather chart* at the current time can be the same as the information on the weather chart in the past. This can be used to predict the weather based on past experience it. For the *forecaster*, this method is almost the same as pattern recognition.

Variables observed as meteorological data to predict rainfall are listed below.

1) Temperature

Temperatures on earth are influenced by solar heating. When daytime heating will occur in the evening air, the opposite will be cooling. The air is a worse heat conductor than the mainland, so that during the daytime land heat faster than air. The highest intensity of heating by the sun during the day occurred right. However, the highest surface air temperature is occurred just a few hours later, this is caused by mainland which still store heat energy for two to four hours afterward.

In the night, the land quickly release the heat radiation into the air, so air in the hotter surface becomes cooler due to heat transfer from air to the ground. The heat from the air at higher altitudes then flow into the layer of air near the earth's surface, causing then so called the radiation inversion.

2) Humidity

Humidity or moisture is used to indicate the amount of water vapor in the air. There are many ways to express humidity; absolute humidity can be expressed by showing the ratio of the mass of water vapor in air volume. Specific humidity indicates the mass ratio of water vapor with air mass. However, generally humidity is indicated by the relative humidity (RH). Relative humidity is the ratio of the actual amount of water vapor in the air compared to the maximum amount of water vapor required for saturation at a certain temperature and pressure.

3) Sea Level Pressure

Pressure means the amount of force or pressure that affects the surface area. Because air mass and gravity, it also has the air pressure.

The pressure can vary with altitude. Atmospheric pressure will drop as much as 10 millibars in increasing altitude of 100 meters. Pressure monitoring station will be corrected by any measurement results with the height.

4) Visibility

Visibility can be defined as the distance of view, namely the extent to which the human eye can see clearly. Rain, snow and fog can be a barrier to the view that makes visibility low.

5) Wind

Since the resultant force, the air will move from high pressure to low pressure. Moving air then is called as wind. The amount of wind speed is influenced by how large the difference pressure between two places.

In addition to having the speed, the wind also has a direction. The instrument used to measure wind speeds in the stratosphere is called an anemometer.

6) Cloud

Clouds are formed by water vapor which is condensed because of enormous pressure and low temperature. Water vapor can be formed from seawater and river water that evaporates due to heating from the sun.

7) Present Weather (Current weather)

How to predict the weather in simplest is method of persistence. This method says that what happens today will happen tomorrow. Just as if today at 10 noon the air temperature is 30 ° C, then tomorrow at 10 the air temperature is also 30 ° C. However, this method can not be used if weather conditions in a region highly complex and changeable. However, the weather forecasters will always consider the weather conditions at this time to predict the weather several days ahead by looking weather trends that occur from time to time.

B. Artificial Neural Network (ANN)

Artificial Neural Networks (ANN) is the artificial representation of the human brain. ANN tried to simulate the human brain by the learning process.

There are several types of ANN, however, almost all have the same components. As human brain, ANN also consists of several neurons, and among these neurons there are connectors. These neurons will present information received through the connection leading to the other neurons. In the ANN, this relationship is known as weights. The information is stored on a specific value on the weight.

1) Neurons

It is a simple processing unit. In these neurons, there are mechanisms for data processing, such as multiplication with the input weights and activation functions.

2) Weight

The weight is a weight value of an input into the ANN. In ANN, the weight will be adjusted by process of adaptation in order to have a function ANN is as expected.

3) Activation Function Neural Networks

The sum of each input that have been multiplied by weighting matrix will be compared with a threshold value (threshold) specified by each neuron activation function.

4) Layer

Layer is a set of neurons that perform similar function. In practice, ANN consists of several layers, including input layer, hidden layer (hidden layer) and output layer. Figure 1 below shows the structure of neuron ANN.

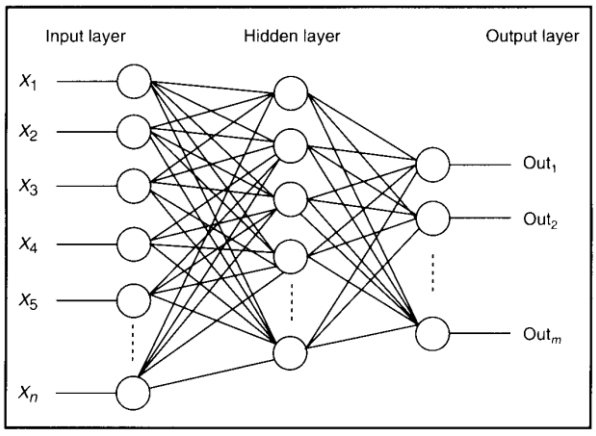


Figure 1. Structure of Artificial Neural Network With Single Hidden Layer

C. Lavendberg Marquadt Algorithm

The *Levenberg-Marquardt* algorithm was designed to approach second-order training speed without having to compute the Hessian matrix H. When the performance function has the form of a sum of squares (as is typical in training feed forward networks), then the Hessian matrix can be approximated as

$$H = J^T J \tag{1}$$

Then, the gradient can be computed as

$$g = J^T e \tag{2}$$

J is the *Jacobian* matrix which contains first derivatives of the network errors with respect to weight and bias. Then, e is a vector of network errors. The *Jacobian* matrix can be computed through a standard back propagation technique which less complex than computing the Hessian matrix.

The *Levenberg-Marquardt* algorithm uses this approximation to the Hessian matrix in the following,

$$X_{k+1} = X_k - \left[J^T J + \mu I\right]^{-1} J^T e \tag{3}$$

When the scalar μ is zero, this is turn to Newton's method using the approximate Hessian matrix. When μ is large, this becomes gradient descent with a small step size. Thus, μ is decreased after each successful step (reduction in performance function) and is increased only when a tentative step would increase the performance function. In this way, the performance function will always be reduced at each iteration of the algorithms.

D. Fuzzy Logic

Facing the problem of uncertainty, generally expert statisticians will finish with the theory of probability. Fuzzy logic will resolve the uncertainty with a better approach through a possibility. Review the case of temperature in an area. Temperature is not only defined by the heat and not hot (cold), with the numbers 1 and 0. But at a certain temperature, humans would define it as very hot, hot, half-hot, half-cold and others. Fuzzy logic gives a more human-like in these conditions with a value of 0 to 1. For example with a very hot, heat it to 0.75, and half-hot half-cold is shown to 0.5.

The basics of fuzzy logic are consisted of:

1) Fuzzy sets theory
In a universe, each member will be grouped into appropriate sets. As in the set of meteorological data for rainfall, pressure, RH and surface wind will be categorized in the set.

2) Membership function
A Membership Function (MF) is a curve that defines how each point in space (space) is mapped in the input membership value or degree of membership between 0 and 1. Membership function is expressed in μ.

3) Logical operation
In fuzzy logic, will use standard Boolean logic, such as AND, OR, and NOT. In stating the truth as 1 and 0,

In fuzzy logic, AND operation will be denoted by min (A, B), whereas the OR operation with max (A, B). Complement or NOT will be denoted by 1-A.

The results of Boolean logic operation that is composed of two values, 0 and 1, will also yield a value of 0 and 1. While fuzzy logic which has a value between 0 to 1 or multi-valued, then the value of the operating result is 0 to 1.

4) If-then rule
If-then rule is used to express statements in fuzzy logic. This rule is formulated as follows.

$$\text{If x } \textit{is} \text{ A } \textit{Then} \text{ y } \textit{is} \text{ B} \tag{4}$$

x is the input while y is the output. A and B are linguistic values of x and y. Sentence "x is A" is called antecedent or premise, while the phrase "y is B" is the consequent or conclusion

E. Fuzzy C-Means

Fuzzy C-Means (FCM) is a method of clustering which allows one piece of data to belong to two or more clusters. This method is frequently used in pattern recognition. It is based on minimization of the following objective function:

$$J_m = \sum_{i=j}^{N}\sum_{j=1}^{C} u_{ij}^{m}\left\|x_i - c_j\right\|^2, 1 \le m < \infty \tag{5}$$

where m is any real number greater than 1, u_{ij} is the degree of membership of x_i in the cluster j, x_i is the *i-th* of d-dimensional measured data, c_j is the d-dimension center of the cluster, and $\|*\|$ is any norm expressing the similarity between any measured data and the center.

Fuzzy partitioning is carried out through an iterative optimization of the objective function shown above. The update of membership u_{ij} and the cluster centers c_j are given by:

$$u_{ij} = \frac{1}{\sum_{k=1}^{C}\left(\frac{\left\|x_i - c_j\right\|}{\left\|x_i - c_k\right\|}\right)^{\frac{2}{m-1}}}, \quad c_j = \frac{\sum_{i=1}^{N} u_{ij}^{m} . x_i}{\sum_{i=1}^{N} u_{ij}^{m}} \tag{6}$$

This iteration will stop with $\max_{ij}\left\{\left|u_{ij}^{(k+1)}-u_{ij}^{k}\right|\right\}<\varepsilon$, where ε is a termination criterion between 0 and 1, whereas k is the iteration steps. This procedure converges to a local minimum or a saddle point of J_m.

III. SIMULATION AND RESULT

To get an accurate prediction system in weather prediction, it is required two sub-stages of the formulation and training of design, and the validation system in which both stages are done in simulation. Thus, the validity of system after formulation and training process will be tested using the testing data. Below is step-by-step to build Neuro-Fuzzy System.

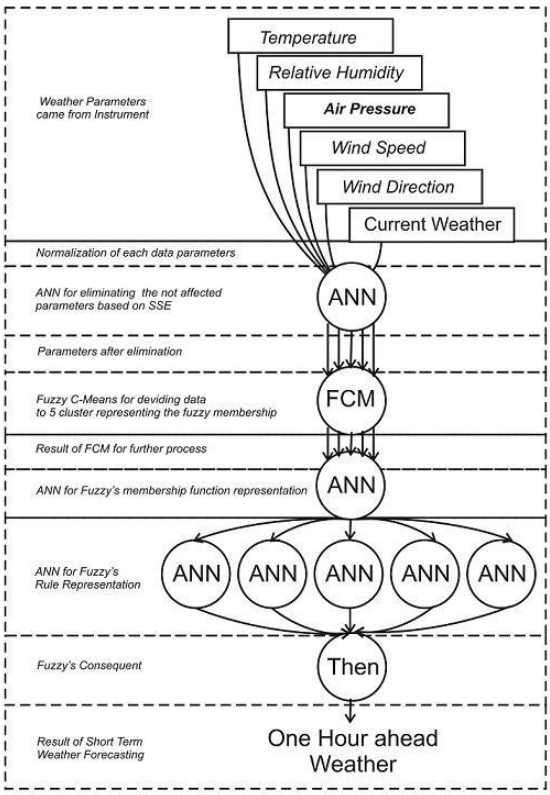


Figure 2. Step of Training System and Learning Prediction

1) *Design and Training of Forecasting System*

The design of prediction system consists of several determination stages and learning system with training data. Learning and validation data for each stage consists of 92 randomly observational data in date 26 to 28 October 2008 and 4 to 6 November 2008. Design and learning process for this system can be described in Figure 2.

The following are the stages of formation and learning process of proposed system,

a) *Casting: Selection of weather parameters*

Preliminary data for the input observational weather prediction system consists of temperature (x1), humidity (x2), air pressure (x3), the direction of the wind speed (x4), wind speed (x5) and current weather conditions (x6). The target of the system is weather next one hour. The weather parameters are represented in numerical forms to be processed further. For this research, the weather is divided into three type *i.e* sunny, cloudy and rainy. For this simulation, the state is represented by the numerical constants such as 1, 2 and 3.

The learning for selecting parameter affecting the system uses *Lavendberg-Marquadt* algorithm with a hidden layer consisted of 6 neurons. Network parameter of the maximum epoch is 500; rate of learning is 1; tolerance error MSE is 10^{-3}; and the initialization of neuron weight is set randomly.

The learning process uses all inputs, eliminating the variables x1, x2, x2, x3, x4, x5, x6. Below is table of the learning process result,

TABLE I. CLUSTER CENTER VALUE DATA FOR EACH VARIABLE

No	Varible eliminated	MSE	SSE
1	All Variable	0.01559	1.4341
2	x1	0.01564	1.4393
3	x2	0.01564	1.4392
4	x3	0.01370	1.2606
5	x4	0.01519	1.3983
6	x5	0.01549	1.4257
7	x6	0.01602	1.4743

Based on result above, variable x3 is eliminated because it has the smallest SSE among other circumstances. In other word, variable x3 is not very affected to weather patterns for prediction purposes. On the next stage variable x3 is ignored. Input is renamed and variables now consist of five from x1 to x5.

b) *Clustering: Grouping Variable Input*

To initialize the data as the fuzzy antecedent interpretation, the input variables are grouped into some groups. The data-casting has been grouped by Fuzzy C-means. Fuzzy C-means is constructed with a number of class as many as five classes, a factor loading of 2 units. With these parameters derived from the cluster center value of each input parameter in each class,

TABLE II. CLUSTER CENTER VALUE DATA FOR EACH VARIABLE

No	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
x1	27.5799	26.425	27.9496	26.0741	26.2046
x2	84.1585	88.8543	82.8204	89.5039	88.827
x3	79.6288	127.7022	73.0261	191.7688	212.4769
x4	4.4607	5.2099	5.622	4.4576	4.8088
x5	1.9877	2.0219	1.9854	1.9685	1.9601
x6	0.6537	0.6731	0.6522	0.6800	0.6749

c) *Training: Training ANN for Fuzzy's Antecedent*

Using input data and target data, these data are trained by artificial neural network interpreted as Fuzzy Antecedent to

obtain membership function. This training uses *Lavendberg-Marquadt* method with hidden layers consisted of 6 neurons. Network parameters of the maximum epoch is 1000; learning rate = 0.8; tolerance error MSE = $10^{-3;}$ and the initialization of neuron weight is set randomly. Figure 3 below shows learning process in simulation,

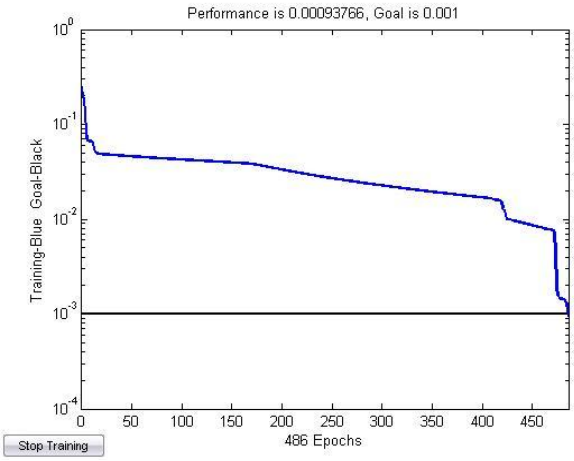
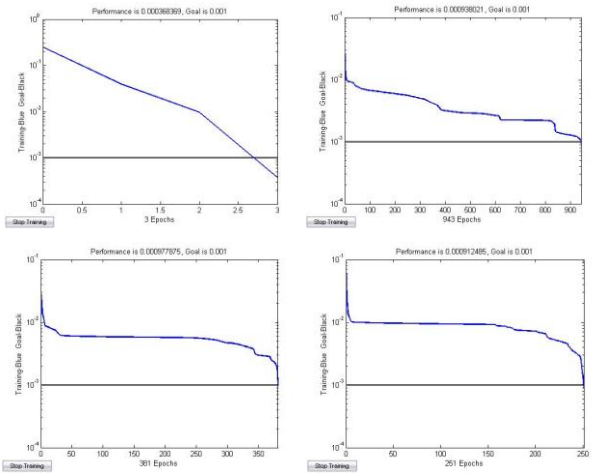


Figure 3. Learning Process Making Membership Functions

Based on above figure, the convergence is reached after iteration in around 500 iterations.

d) Training: Training ANN for Fuzzy's Consequent

For each cluster/rule, it was built an ANN with a target output of each cluster. For this system, because there were five classes/rules then there are five ANN formed. ANN used the *Lavendberg-Marquadt* method with a hidden layer (8 neurons). Network parameters of the maximum epoch = 1000; learning rate = 0.9; and tolerance error MSE = 10^{-3}



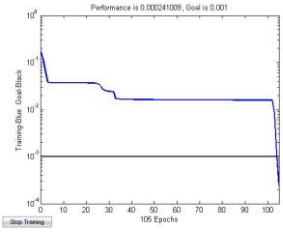


Figure 4. Learning Process of Each Rule (cluster)

Based on those above figures, it can be seen that all of the training stages reach the convergence state before the maximum epoch.

e) Calculation Output System

After the system has been built in stages as above, the system then is tested with training data. The output is obtained through the equation in the previous section. From the test results, it was obtained the correct prediction as many as 68 data out of 92 data for validation.

2) Formulation and Testing

From the design process and learning the system, it can be obtained formulation weather prediction system for the city of Surabaya on the basis of the concept Neural Fuzzy System (NFS). The figure below is the final of weather forecasting using the fusion of Fuzzy logic and artificial neural network. In this system consists of five input nodes, five neurons in the hidden layer interpreted as the Fuzzy Logic Antecedent and one node output interpreted as the Fuzzy Logic Consequent.

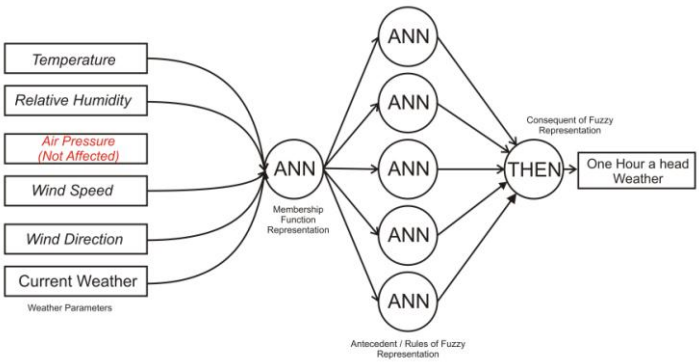


Figure 5. Final Design of Short Term Weather Forecasting

For testing the validity, system was tested by validation data which got as much as 68 of 92 given data or about 74% were correct.

The system has been designed to be adaptive, or in other words it can be trained again when error value increase because there are changes in weather patterns. By using some recent data, the system will be able to perform the update the data for readjusting network weight which is the main part of the prediction.

TABLE III. THE RESULTS OF VALIDATION TESTING OF SYSTEMS DESIGN AND LEARNING USING FINAL RESULT OF FORECASTING

Training Data	Weather	Accuracy
	Shiny	83%
	Cloudy	89%
	Rainy	38%
Validation Data	Weather	Accuracy
	Shiny	20%
	Cloudy	78%
	Rainy	17%

From the data table above, prediction for the cloudy weather is quite high compared to others. This is caused by data for learning was dominated by the cloudy data so the network is more sensitive to weather patterns during cloudy. To improve the system performance on the other weather condition, there is necessary to update data with other condition.

Computation process in the learning stage is big enough. However, in the implementation, it is fairly small. The number of clusters formed by ANN is determined based on technical considerations such as error result (SSE) and computation process. So for further research, the number of cluster can be readjusted to search the best configuration to make the high accurate prediction.

IV. SUMMARY AND CONCLUSION

The method of short term weather forecasting was developed using combined of Fuzzy Logic, Fuzzy C-Means, Artificial Neural Network used *Lavendberg-Marquadt* algorithm. This system gives reliable prediction with around 78% accurate when cloudy, 20% in shiny and 15% in rainy. From these testing results, the method needs more training data for more accurate prediction especially for shiny and rainy condition. By adding the training data, this system should result more accurate hourly prediction.

ACKNOWLEDGMENT

The authors gratefully acknowledge to Bureau of Meteorology, Climatology and Geophysics Republic of Indonesia (BMKG) Juanda, Surabaya for providing data for the realization of this research.

REFERENCES

[1] Imam Maqsood, M Riaz Khan, Ajith Abraham, "An Ensemble of Neural Network for Weather Forecasting ", in Neural Compt & Application (2004) 13: pp 112-122.

[2] Paras, Sanjay Mathur, Avinash Kumar, and Mahesh Chandra, "A Feature Based Neural Network Model for Weather Forecasting", World Academy of Science, Engineering and Technology 34, 2007.

[3] James W. Taylor and Roberto Buizza," Neural Network Load Forecasting with Weather Ensemble Predictions", unpublished.

[4] N. Q. Hung, M. S. Babel, S. Weesakul, and N. K. Tripathi," An artificial neural network model for rainfall forecasting in Bangkok, Thailand", in Hydrology and Earth System Science ed 2009, 13 pp 1413-1425

[5] Marius Matreata," Overview of the artificial neural networks and fuzzy logic applications in operational hydrological forecasting systems ".unpublished.

[6] P. Pinson a n d G. N. Kariniotakis, "Wind Power Forecasting using Fuzzy Neural Networks Enhanced with On-line Prediction Risk Assessment". In IEEE Bologna Power Tech Conference June 23-26 2003

[7] Y. Tulunay, "A Fuzzy Neural Network Model To Forecast The Percent Cloud Coverage And Cloud Top Temperature Maps " in Annels Geophysicae 2008, 26, pp 3945-3954.

[8] Shahi, Ahmad, "An Effective Fuzzy C-Mean And Type-2 Fuzzy Logic For Weather Forecasting", Journal of Theoretical and Applied Information Technology 6 Vol 5 No 5, 2009.

[9] Ingsrisawang, Lily ," Machine Learning Techniques for Short-Term Rain Forecasting System in the Northeastern Part of Thailand", World Academic Scioence and Engineering 41, 2008

[10] Baboo,Dr. S. Santhosh," An Efficient Weather Forecasting System using Artificial Neural Network", International Journal of Environmental Science and Development, Vol. 1, No. 4, October 2010 ISSN: 2010-0264

[11] Fraley, Chris," ProbabilisticWeather ForecastinginR",unpublished.

[12] Simeonov, Ivan," Embedded system for short-term weather forecasting", International Conference on Computer Systems and Technologies - CompSysTech'06, 2006.

[13] Tektaş, Mehmet," Weather Forecasting Using ANFIS and ARIMA MODELS. A Case Study for Istanbul, Journal of Environmental Research, Engineering and Management, 2010. No. 1(51), P. 5 – 10, ISSN 1392-1649