

# Short-Term Localized Weather Forecasting By Using Different Artificial Neural Network Algorithm In Tropical Climate

Noor Zuraidin Mohd-Safar<sup>1</sup>, David Ndzi<sup>1</sup>, Ioannis Kagalidis<sup>1</sup>, Yanyang Yang<sup>1</sup>, Ammar Zakaria<sup>2</sup>

<sup>1</sup>School of Engineering, University of Portsmouth, United Kingdom

<sup>2</sup>School of Mechatronic Engineering, Universiti Malaysia Perlis, Malaysia

{noorzuraidin.mohdsafar, david.ndzi, ioannis.kagalidis, linda.yang}@port.ac.uk

ammarzakaria@unimap.edu.my

**Abstract-** This paper evaluates the performance of localized weather forecasting model using Artificial Neural Network (ANN) with different ANN algorithms in a tropical climate. Three ANN algorithms namely, Levenberg-Marquardt, Bayesian Regularization and Scaled Conjugate Gradient are used in the short-term weather forecasting model. The study focuses on the data from North-West Malaysia (Chuping). Meteorological data such as atmospheric pressure, temperature, dew point, humidity and wind speed are used as input parameters. One hour ahead forecasted results for atmospheric pressure, temperature and humidity were compared and analyzed and they show that ANN with Levenberg-Marquardt algorithm performs best.

**Keywords-** artificial neural network; ANN; short-term weather forecasting; neural network; soft computing; tropics; tropical climate

## I. INTRODUCTION

Meteorological processes are highly non-linear and complicated to predict at high spatial resolutions. Weather forecasting provide critical information about future weather that is important for flooding disaster prediction system and disaster management. This information is also important to businesses, industry, agricultural sector [1], government and local authorities for a wide range of reasons. Soft Computing (SC) techniques such as Artificial Neural Network (ANN) can be used to predict the behaviour of such non-linear conditions [2]. Since the weather processes are non-linear and follow an irregular trend, ANN is a better technique for analysing and identifying the structural relationship between the meteorological parameters [3]. Large data from satellites, radar, weather stations and sensors are processed continuously on a daily basis. This data is transformed into useful information that are used to forecast the weather in the next hours or days. Weather forecasting systems use complex computer algorithm that demand high performance computers and require high resolution spatial data [4]. However, this study will take advantage of localized meteorological to forecast localized weather conditions. Meteorological data from Chuping in North-West Malaysia is selected for this study. There are many techniques that are used in weather forecasting ranging from simple observations to highly complex computerized mathematical models. Numerical Weather Prediction (NWP) method is commonly used for weather forecasting in Malaysia

[5]. NWP forecasting is suitable for large areas and not for localized forecasting [4]. Furthermore, NWP model is adapted from non-tropical regions such as Europe and Japan [5]. This study uses atmospheric pressure, temperature, dew point, humidity and wind speed data from a single weather station. Different ANN algorithms such Levenberg-Marquardt (LM), Bayesian Regularization (BR) and Scaled Conjugate Gradient (SCG) algorithms have been used to forecast atmospheric pressure, temperature and humidity for one hour ahead prediction.

The performance of each ANN algorithms is determined by evaluating the magnitude of the error and the correlation coefficient value between observed and forecasted values. Results show that LM yields better results compared to BR and SCG. This paper is organized into six sections. Section II presents the related works in weather forecasting using ANN. Section III presents the information of study area and data availability. Methodology will be discussed in Section IV. Result and discussion of the finding are presented in Section V. Finally, Section VI discusses the conclusion and future works.

## II. LITERATURE

### A. Related works

Weather forecasting is one of the challenging problems especially for tropical climate where the meteorological condition is dynamic and constantly changing. Many studies have been carried out in weather forecasting using ANN [5-8]. This section summarizes some practical applications of ANN weather forecasting. [6] presents the use of ANN to provide daily forecasts for temperature, wind speed, and humidity for southern Saskatchewan, Canada. The results show that empirical statistical modelling is outperformed by a proposed ANN with radial basis function. [7] obtained temperature and relative humidity weekly forecasts using ANN time series analysis. The network model used is a Multilayer perceptron (MPL) feed forward ANN model with back propagation learning (BPL). The error is less than 3% for 15 weeks temperature and humidity forecast. [7] proposed that statistical parameters can be used as input parameter in ANN model for weeks ahead weather forecasting. Daily temperature forecasting model is presented in [8] using ANN with an additional input.

Ten years of meteorological data from Kermanshah in Iran have been used in the study. The results show that the best performance of MLP-BPL ANN architecture uses sigmoid transfer function in the hidden layer, linear function in the output layer and SCG algorithm [8]. Recently, [9] used deep learning ANN method to forecast air temperature for short-term period prediction in North-western Nevada, United States. One year of data from 2012 to 2013 was used in the study. Hourly meteorological data such as atmospheric pressure, temperature, humidity, precipitation and wind speed were also used. The results show that deep learning ANN by reconstructing the input parameters and combining related meteorological parameters such as barometric pressure, humidity and wind speed data achieved 97% accuracy whilst basic ANN implementation yields 94% accuracy. The proposed model in [9] was developed for temperate climate. However this study is applied in tropical climate where the behaviour and pattern of the meteorological parameters is different. Most of the weather forecasting methods in literature are geo-location dependant. The current advances in ANN methodology for modelling non-linear and dynamical phenomena are the motivation to investigate the application of different ANN algorithms for hourly weather forecasting.

### B. Artificial Neural Network

ANN Structure has a number of interconnected artificial neurons. It is a mathematical model that functions like a neuron of a human brain. A supervised learning ANN must be trained using a training dataset enabling it to create by itself the patterns and the rules governing the network. Feed forward ANN was introduced and is known as perceptron. This model uses a single input perceptron layer with a single output. The disadvantage of this approach is that a perceptron is not able to train and recognize many types of patterns. A single neuron consists of two parts; a weighted coefficient and a transfer function. Fig. 1 shows a single neuron that consists of a weighted input, summing function and transfer function.

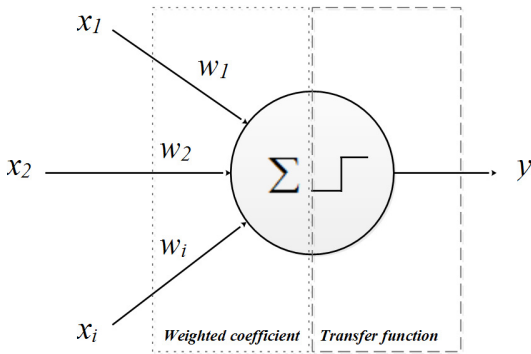


Fig. 1. Single neuron

For a neuron receiving  $n$  inputs, each input  $x_i$  (where  $i=1 \dots n$ ) is weighted by multiplying it by a weight,  $w_i$ . The sum of the  $w_i x_i$  products gives the net activation of the neuron. This activation value is subjected to a transfer function ( $f$ ) to produce the neuron's output  $y$ :

$$y = f \left( \sum_{i=1}^n w_i x_i \right) \quad (1)$$

Adding more perceptron layers in a network topology will increase the ANN's ability to recognize more classes of patterns. This additional layer is known as MLP. MLP works by adding more layers of nodes between the input and output nodes. The neurons are arranged into an input layer, an output layer and one or more hidden layers in between. The learning process for MLP is known as BPL [2][10][11]. The process repetitively calculates an error function and adjust the weights to achieve a minimum error. The transformation of weights depends on the following steps:

1. Each learning step starts with an input signal from training dataset.
2. Then the process will determine the output values for each neuron in each network layer.
3. Finally, based on the above step, weight is selected to map the input with the output.

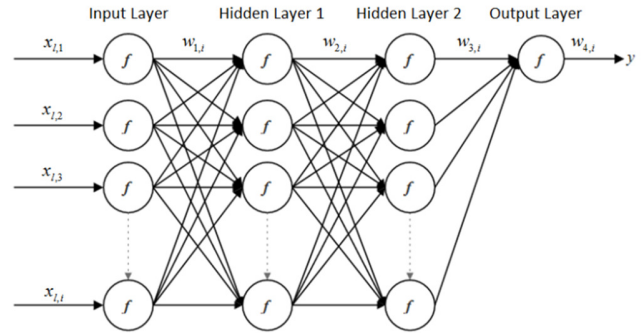


Fig. 2. ANN layer representation

Fig. 2. illustrates the propagation of signal traversing through each neuron from input layer, hidden layer and finally the target. Supposed that the layer representation has  $L$  number of layers, including input, hidden and output layer, while  $l$  is representing the input layer and hidden layer that has  $N$  number of nodes in the form of  $N(l)$ , where  $l=(0, 1, \dots, L)$ ;  $l=0$  is the input parameter) and  $i=1, \dots, N(l)$  is the node that has an output from a previous layer and it will be the input for the next layer,  $y_{l,i}$  is the output that depends on incoming signals  $x_{l,i}$  and parameters  $\alpha, \beta, \gamma$ . Thus the following equation is the generalization of the output from each node:

$$y_{l,i} = f_{l,i}(x_{l-1,1} w_{l-1,1}, \dots, x_{l-1,N(l-1)}, \alpha, \beta, \gamma, \dots) \quad (2)$$

After the propagation of signals is completed, the next step is to compare the output signal with the desired output ( $z$ ) from the training dataset. In general, the difference is the error signal  $\delta_{l,i}$ .

$$\delta_{l,i} = z - y_{l,i} \quad (3)$$

Assuming that the training dataset has  $P$  entries, the error measured for the  $P^{\text{th}}$  entry of the training dataset is the sum of the squared error:

$$E_p = \sum_{k=1}^{N(L)} (z - y_{l,i})^2 \quad (4)$$

where  $k$  is the number of components of  $z$  (desired output) and  $y_{l,i}$  (predictive output). The weights for a particular node are adjusted in direct proportion to the error by propagating it back through all neurons using Gradient descent algorithm. Gradient descent algorithm finds the weight that will deliver the minimized error. The algorithm can be simplified into two steps as follows:

1. Obtain gradient vector.
2. Calculate the error signal  $\varepsilon_{l,i}$  as the derivative of the error measure  $E_p$  with respect to the output node  $i$  in layer  $l$  in both direct and indirect paths. The ordered derivative can be express as in the following equation:

$$\varepsilon_{l,i} = \frac{\partial^+ E_p}{\partial y_{l,i}} \quad (5)$$

Every input parameter that is used to train the network is associated with the output pattern. The ANN forecasting models use standard supervised learning MLP trained with BPL algorithm [2][10][11]. The entire process of training in MPL-BPL implementation is summarized in the following steps:

1. Initialize the weights using the training dataset by mapping the input data to the desired output data.
2. Initialize bias by randomly selecting data from training dataset.
3. Compute the output of neurons, error and update the weights.
4. Update all weights and bias and repeat step 3 for all training data
5. Repeat step 3 and step 4 until the error is minimized.

In this study, three BPL training algorithm were applied to evaluate their performances, each of them having a variety of different computation and storage requirements. Table I summarizes the characteristics of the algorithms.

TABLE I. ANN TRAINING ALGORITHM.

Algorithm	Description
<b>LM</b>	Faster training algorithm for networks with moderate size, with memory reduction capability when the training data set is large. Using approximation method to update network weights and biases. LM algorithm for network training is used in [12] and [13].
<b>BR</b>	BR training algorithm improves generalization and reduces the difficulty of the determination of the optimum network architecture. BR uses object function such as MSE to improve generalization regularization technique requires expensive computation to reach optimal level. Detail discussion of BR is described in [14], [15] and [16].
<b>SCG</b>	SCG is designed based on optimization technique in numerical analysis named Conjugate Gradient Method [17]. Unlike other conjugate gradient approach SCG was designed to avoid time consuming line search, it may required more iteration to converge but the number of computation is reduced because line searched is avoided. This technique does not require any user specified parameters and its computation is faster and inexpensive. Detailed description of the algorithm can be found in [18].

### III. STUDY AREA

Chuping is a small town in Perlis, Malaysia. It has 22,000 hectares of agricultural and plantation land area. The climate is tropical characterized by sunny days, high temperature and high humidity. In a whole year, the average day period is from 7 a.m. to 7 p.m. The average temperature is 27.5°C. Hourly meteorological data of atmospheric pressure, dry bulb temperature, dew point, humidity, wind speed, wind direction, rainfall amount and rainfall rate are measured and have been used in this study. Three years of data from January 2012 to December 2014 that consists of 26304 dataset have been used. Elementary meteorological parameter characteristic for Chuping are summarized in Table II.

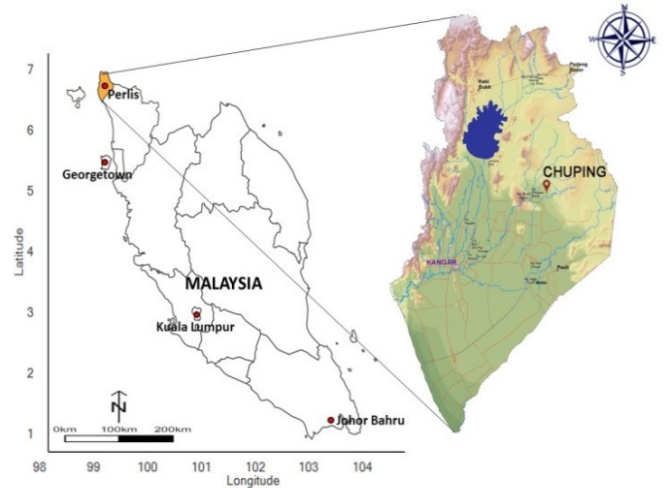


Fig. 3. Map of Malaysia and Chuping

TABLE II. PARAMETER ANALYSIS FOR CHUPING

Parameter	Uni	Max	Mea	Min	Media	$\sigma$	$\sigma^2$
Pressure	HPa	1017.	1009.	1001.	1009.6	2.0	4.0
Temperatur	$^{\circ}C$	$\hat{3}8.1$	$\hat{2}7.5$	$\hat{1}6.1$	26.4	3.2	10.2
Dew Point	$^{\circ}C$	29.4	24.5	15.3	24.5	1.5	2.2
Humidity	%	100	83.7	29	88	12.	148.
Wind Speed	$ms^{-1}$	5.1	1.1	0	1.1	$\hat{0}.9$	$\hat{0}.7$

$\sigma$  =Standard Deviation,  $\sigma^2$ =variance

#### IV. METHODOLOGY

##### A. Data cleaning

A common problem in meteorological study is that of missing data due to insufficient sampling, faulty data acquisition or instrument measurement error. Data cleaning methods have been used in this study to reduce noisy data. Due to high dimensionality of the parameters, Principal Component Analysis (PCA) algorithm is a suitable model to determine the missing data [19]. Imputation, a process of determining the value of missing data, is used [20]. In one dimensional data, basic statistical analysis such as minimum, maximum, median, standard deviation and variance (Table II) is one of the important analysis to interpret parameter trends and changes [21].

##### B. ANN Design

Designing an ANN model will follow the standard steps from input data to network trained and finally, use the network [22].

###### 1) ANN Preliminary implementation:

In the design stage, several models were tested in order to find the optimum ANN model for each algorithm. Data was divided into training, validating and testing dataset. In order to avoid overfitting and increase generalization, an appropriate amount of the training dataset was selected and used for cross validation. During the training process, the errors in training and validation were monitored. When the error in the validation set increases, the training should be stopped because the point of best generalization has been reached. The cross validation approach with split-sample training was adopted for the training of ANN models in this study. Three years of data with hourly observed parameters were randomly divided into training, validation and test datasets with 70% for training, 15% for validation and 15% for testing.

###### 2) Data Normalization

ANNs learn faster and give better performance if the input variables are pre-processed before it is used to train the network [10]. A normalization method is used to pre-process input and target data. Data normalization is the process of scaling the data to fall within a smaller range. The advantage of scaling data is to make all weighted neurons to remain within a small and predictable range. Normalization scaling between -1.0 to 1.0 is adopted and normalized data was generated using the following equation:

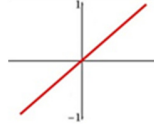
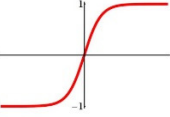
$$x'_{scale} = 2 \left( \frac{x - x_{min}}{x_{max} - x_{min}} \right) - 1 \quad (6)$$

where  $x$  is the value before normalization,  $x_{max}$  is the maximum value and  $x_{min}$  is the minimum value in the dataset.

###### 3) Input layer, hidden layer and output layer

In the input layer, each neuron received a single input which originated from meteorological data. However, the hidden layer and the output layer can accept an arbitrary number of inputs based on the type of chosen interconnection of the neurons. Two hidden layers with 20 and 10 neurons have been used in ANN training. Table III shows Linear (LT) and Hyperbolic Tangent Sigmoid (HT) transfer functions (or activation functions) that have been used in this study. HT has been used in the first hidden layer while LT has been used in the second hidden layer.

TABLE III. TRANSFER FUNCTION

Transfer function	Graphical illustration	Mathematical equation
Linear		$y = x$
Hyperbolic Tangent Sigmoid		$y = \frac{e^x - e^{-x}}{e^x + e^{-x}}$

##### C. Experimental setup

For the purposes of this study, the input vectors of hourly atmospheric pressure, temperature, dew point, humidity and wind speed have been used for training. One hour lag of target vector is used as additional input for each forecast parameter. The target vectors are the observed metrological parameter (pressure, temperature and humidity). Each of the experimental setup determine the optimum outcome for one hour ahead forecast.

In this paper, weather forecasting performance for pressure ( $P$ ), temperature ( $T$ ) and humidity ( $H$ ) are evaluated. Other parameters such as dew point ( $DP$ ) and wind speed ( $WS$ ) were used as additional input parameters. LM, BR and SCG algorithms have been used in the same ANN architecture. The accuracy of each algorithm has been analysed. In general the input parameters,  $P_{input}$ , and output parameters,  $P_{output}$ , for the ANN architecture can be described as follows:

$$P_{input} = [P_{t-n}, T_{t-n}, DP_{t-n}, H_{t-n}, WS_{t-n}, PF_{t-n}] \quad (7)$$

$$P_{output} = [PF_t] \quad (8)$$

where  $PF$  is the metrological parameter observed at time  $t$  and  $n$  is the hour(s) before the forecast. In this setup,  $n = 1$  and  $t = 2$ .

The ANN used in the experimental setup was trained using BPL-MLP training algorithm as shown in Fig. 4.

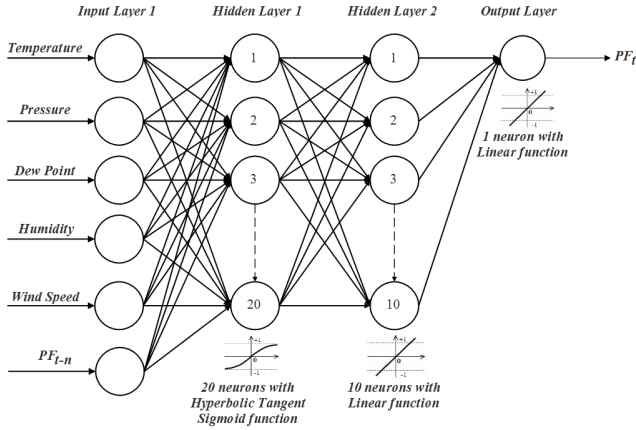


Fig. 4. Weather forecasting ANN architecture.

The structure of ANN forecast model can be summarized as follows:

- Input layer:  $N$  number of neurons where  $N > 0$
- Hidden layer: Two hidden layer with 20 and 10 neurons.
- Output layer: One output layer, where the output for the next one hour parameter forecast is obtained.
- Training functions: LM, BR and SCG algorithm.
- Transfer functions: HT and LT.
- Training set: 22360 datasets.
- Test set: 3944 datasets.
- Training iterations: 1000 epochs.
- Performance function:  $MSE=0.0001$ .

#### D. Accuracy Performance Indices

The performance of all networks was measured using Mean Absolute Error ( $MAE$ ), Root Mean Square Error ( $MSE$ ) and correlation coefficient between the observed and forecasted value. In (9), (10) and (11),  $y_t$  is the observed value,  $\hat{y}_t$  is forecasted value and  $n$  is the number of observation.

$$MAE = \frac{1}{n} \sum_{t=1}^n |y_t - \hat{y}_t| \quad (9)$$

$$MSE = \frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_t)^2 \quad (10)$$

The correlation coefficient is a measure of the linear dependency between two variables. If each variable has  $n$  scalar observations, then the Pearson correlation coefficient [23] is defined as:

$$R = \frac{n \sum y_t \hat{y}_t - (\sum y_t)(\sum \hat{y}_t)}{\sqrt{n(\sum y_t^2) - (\sum y_t)^2} \sqrt{n(\sum \hat{y}_t^2) - (\sum \hat{y}_t)^2}} \quad (11)$$

The  $R$  value is an indication of the relationship between the forecasted and observed value. If  $R=1$ , this indicates that there is an exact linear relationship. If  $R$  is close to zero, then there is non-linear relationship between them.

## V. RESULTS AND DISCUSSIONS

The purpose of this study was to evaluate the accuracy of LM, BR and SCG algorithms use in weather forecasting with focus on atmospheric pressure, temperature and humidity parameters. The input parameters are from past meteorological data. The results show that LM is the best forecasting algorithm. Results and comparison are summarized in Table IV. LM yields low MAE and MSE compared to BR and SCG algorithms. The correlation coefficient value produced by LM is greater than 0.95 for pressure, temperature and humidity forecast. Fig. 5 shows the training, validation, testing and overall regression plots for temperature forecast using LM algorithm. LM offered the best accuracy, followed by BR and SCG. In terms of processing time, SCG is faster [18] compared to LM and BR but it does not produce good convergence. BR algorithm takes more time compared to LM and SCG during training but it converges faster. Fig. 6, Fig. 7 and Fig. 8 show the comparison between forecast and observed results for pressure, temperature and humidity using LM algorithm.

TABLE IV. ACCURACY PERFORMANCE RESULT

	ALGORITHM	MAE	RMSE	R
Pressure	LM	0.4410	0.5475	0.9618
	BR	0.4449	0.5664	0.9614
	SCG	0.4772	0.5851	0.9563
Temperature	LM	0.6369	0.9168	0.9577
	BR	0.6492	0.9270	0.9568
	SCG	0.6759	0.9541	0.9542
Humidity	LM	2.3684	3.4442	0.9592
	BR	2.4086	3.4858	0.9582
	SCG	2.6817	3.7069	0.9525

BR=Bayesian Regularization, LM= Levenberg-Marquardt, SCG=Scaled Conjugate Gradient

## VI. CONCLUSION AND FUTURE WORKS

The proposed ANN forecast model has been trained using meteorological data from a tropical area, North Malaysia. Results from this study has shown that ANN forecast model with LM algorithm offers significant potential for weather condition forecasting for localized tropical region. From the overall performance, an ANN weather forecasting model is capable of capturing the dynamic behaviour of atmospheric pressure, temperature and humidity for one hour forecasting. Modelling for longer time window forecasting should be implemented and tested. Wind speed and dew point forecasting from the same dataset will be carried out for an effective weather forecast system. This study offers a potentially new approach for weather forecasting system for localized tropical climate. Forecast results from this study are very useful in environmental condition prediction such as rain occurrences and rain intensity. Furthermore, when rain forecasting is reliable, it will contribute to an effective water resources management, flood prediction, drought mitigation, ecological studies and climate change impact assessment.

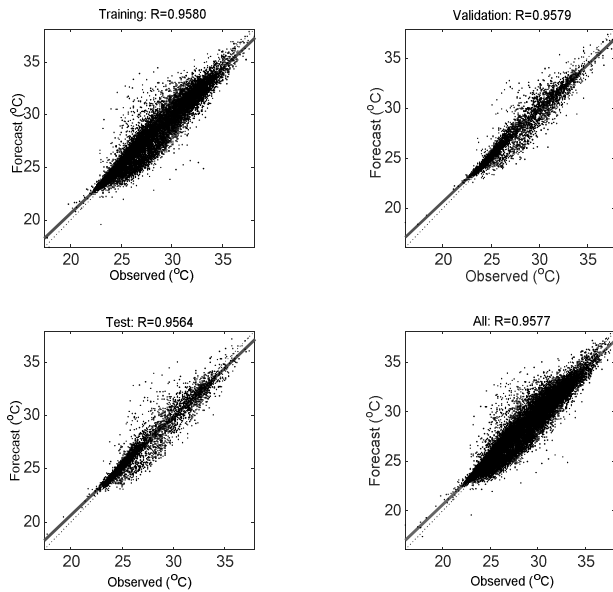


Fig. 5. Regression plot for temperature forecast using ANN with LM algorithm.

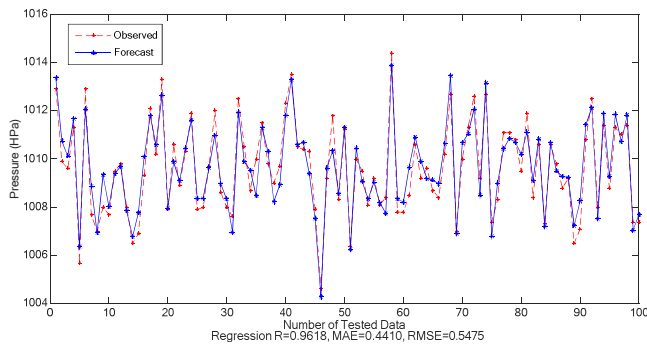


Fig. 6. Simulation result for Pressure using LM algorithm

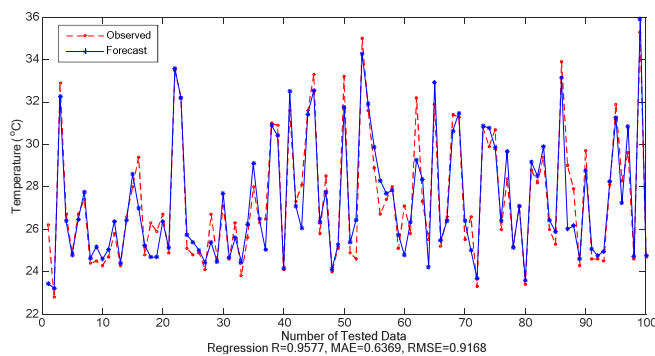


Fig. 7. Simulation result for Temperature using LM algorithm

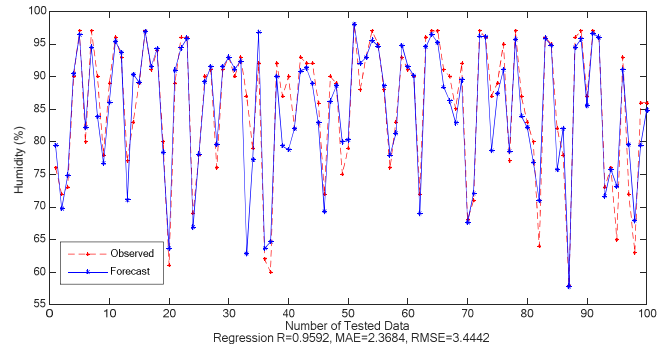


Fig. 8. Simulation result for Humidity using LM algorithm

## REFERENCES

- [1] D. L. Ndzi, A. Harun, F. M. Ramli, M. L. Kamarudin, A. Zakaria, A. Y. M. Shakaff, M. N. Jaafar, S. Zhou, and R. S. Farook, 'Wireless sensor network coverage measurement and planning in mixed crop farming', *Computers and Electronics in Agriculture*, vol. 105, pp. 83–94, 2014.
- [2] G. K. Rahul and M. Khurana, 'A Comparative Study Review of Soft Computing Approach in Weather Forecasting', no. 5, pp. 295–299, 2012.
- [3] B. Cheng and D. M. Titterton, 'Neural Networks: A Review from a Statistical Perspective', *Statistical Sciences*, vol. 9, no. 1, pp. 2–30, 1994.
- [4] V. B. Nikam and B. B. Meshram, 'Modeling Rainfall Prediction Using Data Mining Method: A Bayesian Approach', *2013 Fifth International Conference on Computational Intelligence, Modelling and Simulation*, pp. 132–136, 2013.
- [5] A. Shahi, 'An Effective Fuzzy C-Mean and Type-2 Fuzzy', *Journal of Theoretical and Applied Information Technology*, pp. 556–567, 2009.
- [6] I. Maqsood, M. Khan, and A. Abraham, 'An ensemble of neural networks for weather forecasting', *Neural Computing and Applications*, vol. 13, pp. 112–122, 2004.
- [7] S. M. Paras, A. Kumar, and M. Chandra, 'A feature based neural network model for weather forecasting', *International Journal of Computational Intelligence*, vol. 4, no. 3, 2009.
- [8] M. Hayati and Z. Mohebi, 'Application of artificial neural networks for temperature forecasting', *World Academy of Science, Engineering and Technology*, vol. 28, no. 2, pp. 275–279, 2007.
- [9] M. Hossain, B. Rekabdar, S. J. Louis, and S. Dascalu, 'Forecasting the weather of Nevada: A deep learning approach', *Proceedings of the International Joint Conference on Neural Networks*, vol. 2015-Sept, pp. 2–7, 2015.
- [10] Y. A. LeCun, L. Bottou, G. B. Orr, and K. R. Møller, 'Efficient backprop', *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 7700 LECTURE NO, pp. 9–48, 2012.
- [11] A. Abraham, N. S. Philip, and K. B. Joseph, 'Will We Have a Wet Summer? Soft Computing Models for Long-term Rainfall Forecasting', 1992.
- [12] F. D. Foresee and M. T. Hagan, 'Gauss-Newton approximation to Bayesian learning', in *Proceedings of International Conference on Neural Networks (ICNN)*, pp. 97–102, 1997.
- [13] V. Pellakuri, D. Rajeswara Rao, P. Lakshmi Prasanna, and M. V. B. T. Santhi, 'A conceptual framework for approaching predictive modeling using multivariate regression analysis vs artificial neural network', *Journal of Theoretical and Applied Information Technology*, vol. 77, no. 2, pp. 287–290, 2015.
- [14] J. J. Moré, 'The Levenberg-Marquardt algorithm: Implementation and theory', in *Lecture Notes in Mathematics*, Springer Science + Business Media, pp. 105–116.
- [15] D. J. C. MacKay, 'A Practical Bayesian Framework for Backpropagation Networks', *Neural Computation*, vol. 4, no. 3, pp. 448–472, May 1992.
- [16] D. J. C. MacKay, 'Bayesian Interpolation', *Neural Computation*, vol. 4, no. 3, pp. 415–447, May 1992.
- [17] J. R. Shewchuk, 'An Introduction to the Conjugate Gradient Method Without the Agonizing Pain', *Science*, vol. 49, no. CS-94–125, p. 64, 1994.
- [18] M. F. Møller, 'A scaled conjugate gradient algorithm for fast supervised

- learning', *Neural Networks*, vol. 6, no. 4, pp. 525–533, Jan. 1993.
- [19] T. Schneider, 'Analysis of incomplete climate data: estimation of mean values and covariance matrices and imputation of missing values', *J. Climate*, vol. 14, pp. 853–871, 2001.
- [20] J. Josse, J. Pages, and F. Husson, 'Multiple imputation in principal component analysis', *Advances in Data Analysis and Classification*, vol. 5, no. 3, pp. 231–246, 2011.
- [21] B. Manly, *Statistics for Environmental Science and Management*. Chapman and Hall/{CRC}, 2000.
- [22] M. B. Howard Demuth and M. Hagan, 'Neural network toolbox user's guide', 2012.
- [23] J. Hauke and T. Kossowski, 'Comparison of Values of Pearson's and Spearman's Correlation Coefficients on the Same Sets of Data', *Quaestiones Geographicae*, vol. 30, no. 2, pp. 87–93, 2011.