Forecasting Rainfall Distribution Using Artificial Neural Networks for Johor Rivers

Nur Farhana Hordri[a*], Siti Sophiayati Yuhaniz[b] , Kamilia Kamardin[b]

[a]*Soft Computing Research Group, Faculty of Computing, Universiti Teknologi Malaysia and UTM Big Data Centre*

[b]*Advanced Informatics School, Universiti Teknologi Malaysia and UTM Big Data Centre*

Abstract

The study is conducted to forecast the rainfall distribution in the areas around Johor, Malaysia. Although there are many other factors, we will be using the rainfall distribution factor only. The forecasting method that is going to be used in this study is the Artificial Neural Networks (ANN) which will be trained using back propagation learning algorithm. To produce the best model, several propagation models will be constructed in the algorithm. The value of learning rate parameter and momentum parameter will also be used and constantly changed based on the number of hidden nodes. The data is prepared and filtered using data pre-processing. Data pre-processing includes data cleaning, normalisation, transformation, feature extraction and selection. The product of data pre-processing is the final training set. At the end of the experiment, the best model was selected and the strength of the relationship of each model based on their activation functions that have been used was compared. The result of the model produces the minimum error value and has a stronger relationship between the actual data value and forecast data value is the best model among the best.

Keywords: data pre-processing, forecasting, rainfall distribution, artificial neural network, back propagation

1. Introduction

Data pre-processing is any process that is done manually or automated via computer program involving these steps; input of data, summarising, analysing and converting data into useful information. The process can be programmed to be run in automated mode, and executed by machine at own convenience. Data pre-processing incorporates steps such as recording, analysing, organising, summarising, calculating, distributing and

* Corresponding author. E-mail address: nfarhana64@live.utm.my

storing the data, among others [1, 3, 6]. Since the process resulted in data to be more beneficial and informative when it is organised, this leads to data pre-processing system, often identified as information system, in general.

Contradictory to this belief, the correct term for data pre-processing is to manipulate raw data to be more informative and this involves using raw data as an input to produce useful information. Data pre-processing differs from data conversion, where both processes have similar step i.e. conversion of raw data, but data pre-processing takes the conversion to another extent; by manipulating the data.

The raw data for this study were collected from Water Resources and Hydrology Management for 2006 until June 2010 that covers three areas in Johor which is Senai, Kg. Seelong and Kg. Separa.

For the purpose of this study, a computer program named Microsoft Excel (MS Excel) is used for data pre-processing. The steps are performed in order as follows; data were obtained in the form of a notepad file (.txt). Microsoft Office Excel Worksheet file (.xlsx) was then used to open the text file. The data saved in text file is in delimited format, hence conversion of each character such as comma or tab, were then separated into each field in MS Excel. Once the conversion had been done, the data appears to be organised and informative.

2. Empty Data

After the data had been converted into Excel file (.xlsx), it was noted that there are empty data within these data, which could affect forecasting for this study. These empty data may result from technical problems during shipping process. Problems include data may not be in its original condition, damaged data or obsolete data due to data storage problem or the data itself had been gathered for quite some time ago, which may not reflect the current environment.

Therefore, in this study, we decided to use dummy data to replace the empty data. The dummy data that were used are based on previous data trend within the same month in previous year. Considering there is no significant climate change, assumptions were made as follows, December 2006 is considered as heavy rainfall season; therefore December 2007 is also considered as heavy rainfall season.

3. Data Normalisation

Once the empty data step had been completed, data normalisation was then performed. Data normalisation is a process of grouping elements of data into the form of tables. Within normalisation process, test runs are performed on data with several conditions to resolve conflicts encountered during adding, deleting, modifying or reading the database. When hiccups are found or problem arises during test, data is then divided into several tables again to acquire optimal database.

The purpose in normalising the data is to ensure that the automated output generated is more convergent [4]. In this process, the data shall be converted into the form of value, between 0 to 1, with the intention to facilitate learning process and network calculation process. Concurrent to this process, transformation equation is used as linear transformation to enable large value of the original data to be scaled into smaller range. Therefore, it is expected that the learning phase to be accelerated and concentrated.

In this study, normalisation is performed to rescale the data in the range of 0 to 1.

$$x_n = \frac{x_o - x_{min}}{x_{max} - x_{min}}$$

Where,
x_n = normalised data
x_o = *i*th input/output data
x_{min} = minimum input/output data
x_{max} = maximum input/output data
Results from data normalisation, generated data in the range of 0 to 1.

4. Analyse Real Data for Each Year

Following are the data analysis for the three places studied for five consecutive years.

4.1. Senai, Johor

Figure 1 (a)-(e) refer to rainfall level's graphs from 2006 till June 2010 in Senai, Johor. These figures represent the amount of rainfall level recorded for each year on monthly basis.

For Figure 1(a), it was found that the highest amount of rainfall recorded during the year is in December with reading of 418 mm and least amount of total rainfall recorded is in March, with reading of 31.8 mm. Total rainfall recorded during the year is 1,890.8 mm. This exhibits that the last quarter of 2006 in this area is a rainy season due to heavy rainfall where total rainfall in November is 221 mm. In mid-2006, minimum rainfall were recorded since most of the rainfall level are less than 150 mm.

In Figure 1(b), it was identified that the highest amount of rainfall recorded during the year is also in December, with reading of 405 mm and minimum amount of rainfall recorded is in May, 54 mm. Total rainfall recorded for 2006 is 1743 mm. The month identified with the highest rainfall after December is January which amounted to 380 mm. This is due to the month of January is a subsequent month after previous year's December. In mid-2007, the rainfall level recorded is low since most rainfall level recorded are less than 105 mm.

In Figure 1(c), the highest amount of rainfall in 2008 was recorded in December, with reading of 216 mm and minimum total rainfall recorded was in March with 24.5 mm. Total rainfall recorded for the year is 1041 mm. The figure exhibited irregular rainfall level during the months of July, August and September with an average of 60.5 mm. By mid-2008, the rainfall level recorded is low since most months during this period, recorded rainfall rate of less than 70 mm. Rainfall level recorded in 2008 displays similar pattern with rainfall level recorded in 2007.

In Figure 1(d), it was found this figure fluctuates from the previous figure whereby, the highest amount of rainfall level recorded during 2009 was recorded in May with 447 mm and the lowest rainfall level recorded is in January, with 4 mm. Total rainfall during the year is 2489.5 mm. Within this year, rainfall level showed irregular pattern since there exists few months with high rainfall level and another few months with low rainfall level. The five months with rainfall level recorded exceeding 300 mm are March, April, May, August and November. In contrast, the three months with rainfall level recorded less than 100 mm are January, May and September.

The final figure is Figure 1(e), which displays only the first six months of 2010. It was noted that the highest rainfall level recorded within the six months is in April, with reading of 431.5 mm and the minimum total rainfall level recorded is in February; 84 mm. Total rainfall level recorded during the first six months is 1229.8 mm. In the month of March and June, rainfall rate recorded is almost similar, with reading of 219.3 mm and 227 mm, respectively.

As a conclusion, for the past five years, Senai exhibits high rate of rainfall during early and end of the year. The lowest rate of rainfall is recorded in the middle of the year. Thus, the mid-year is considered as dry season, while early and end of the year are considered as rainy season.

4.2. Kg. Seelong

Figure 2 (a)-(e) refers to rainfall level's graphs from 2006 till June 2010 in Kg. Seelong, Johor. These figures represent the amount of rainfall recorded for each year on monthly basis.

As exhibited in Figure 2(a), it was found that the maximum amount of rainfall recorded during the year is in December, with reading of 881.5 mm and the least amount of rainfall recorded is in May; 37.5 mm. Total rainfall level recorded in 2006 is 2566.5 mm. It is worth to note that from January till November 2006, the rainfall rate is less than 300 mm. In first quarter of 2006, minimum rainfall level was recorded since the first three months exhibit levels of rainfall less than 100 mm.

In Figure 2(b), it was identified that the highest amount of rainfall level recorded during the year is in December, with reading of 484.5 mm, similar with previous year, and the least amount of rainfall recorded is in February, with reading of 0 mm, which indicates no rainfall. Total rainfall recorded for 2007 is 2624 mm.

There are three months with rainfall rates approaching to 0 mm which is in February, March and April, with readings of 0 mm, 5 mm and 3.5 mm respectively. It was noted that there are three months with similar average rainfall distribution rate of 322.33 mm. The months that recorded the similar reading are January, May and June.

In Figure 2(c), it was found that the figure fluctuates from previous figure, where the highest level of rainfall recorded during the year is in March, with reading of 457 mm and this was tailed by April and January, respectively, with reading of 386 mm and 342 mm. The least amount of rainfall recorded is in September, with reading of 23.5 mm. There six months which do not exceed 150 mm rainfall rate are February, May, June, September, October and December. The readings are 123.5 mm, 64 mm, 104.5 mm, 23.5 mm, 135.5 mm and 126.5 mm respectively. Total rainfall level recorded for 2008 is 2436 mm.

In Figure 2(d), the highest amount of rainfall recorded during the year is recorded in July, with reading of 189.5 mm and the least amount of rainfall recorded is in February with reading of 5.5 mm. The seven months, which recorded rainfall rates less than 80 mm are as follows; January, February, March, April, June, August and September. Rainfall rates recorded for the months are 8.5 mm, 5.5 mm, 31 mm, 48 mm, 78.5 mm, 73 mm and 56 mm, respectively. The three months with rainfall rates recorded exceeding 150 mm are July, October and November, and the distribution is as follows; 189.5 mm, 169.5 mm and 160 mm. Total rainfall recorded for 2009 is 1018 mm.

The final figure is Figure 2(e), which only exhibits the first six months of 2010. The highest rainfall level recorded within six months is in January with reading stands at 143.5 mm and the minimum rainfall rate recorded is in April with reading of 1 mm. Total rainfall for the first six months of 2010 is 407 mm. For the month of May and June, rainfall level does not exceed 60 mm. The rainfall levels recorded are 24.5 mm and 51.5 mm, respectively.

As a conclusion, the level of rainfall recorded for the past five years in Kg. Seelong is erratic. It seemed that rainy seasons occur twice; during mid-year and at the end of the year. In this area, even monthly minimal rainfall level recorded stands at 0 mm. The figures displayed that during dry seasons, the rain distribution rate is low, and this differs with monsoon seasons, where rainfall distribution rate is very high.

4.3. Kg. Separa

Figure 3 (a)-(e) exhibited five graphs of rainfall level recorded from 2006 till June 2010 in Kg. Separa, Skudai Johor. These figures represent the amount of rainfall for each year on monthly basis.

Figure 3(a) exhibited the highest amount of rainfall recorded in 2006 is in March, with reading of 997 mm and minimum amount of rainfall level recorded is in May, with reading of 144 mm. Total rainfall during the year is 6137.9 mm. Besides March, the three months with rainfall rates exceeding 800 mm are January, June and December. The months recorded rainfall rate at 808 mm, 997 mm and 889.3 mm, correspondingly. Apart from October, the other five months with rainfall level recorded less than 400 mm are February, April, August, September and November. The months recorded rainfall level as follows; 244.9 mm, 211 mm, 280 mm, 336.7 mm and 191 mm, respectively.

It was identified in Figure 3(b) that the highest amount of rainfall recorded during the year is in June, with reading of 1000 mm and this is closely followed by the month of March, with reading of 997 mm. The least amount of rainfall for the year is in November of 25 mm. Total rainfall for 2007 is 4144 mm. Besides November, the five months that experienced rainfall level less than 200 mm are as follows; February, July, August, September and October. The recorded rainfall level are as follows, 84 mm, 104 mm, 176 mm, 195 mm and 115 mm, respectively. The month of May and December recorded average rate of rainfall level at 494 mm.

For Figure 3(c), the highest amount of rainfall during the year is in January, with reading of 403 mm and this is followed by the month of March with rainfall level of 356 mm. The least amount of rainfall recorded is in September, with reading of 23.5 mm. From July till December 2009, conclusion can be made that the level of rainfall is almost similar. The reading from July till December are as follows; 246 mm, 269 mm, 248 mm, 299 mm, 221 mm and 183 mm, correspondingly. For February and May, both months seemed to have similar rainfall level of 60.5 mm. Total rainfall recorded for the year is 2572 mm.

Figure 3(d) showed that the highest amount of rainfall during the year is in July, with reading of 313 mm and the least total rainfall recorded is in April, with reading of 47 mm. From August till November 2009, it was observed that the rainfall rate has similar pattern. The rainfall rate for the months are as follows, 177.9 mm, 173 mm, 196 mm and 173 mm, respectively. The month of January and February displayed even pattern, where the monthly average rainfall level recorded is 51.1 mm. Total rainfall recorded during the year is 1863.8 mm.

As per previous figures, Figure 3(e) is only limited to the first six months of 2010. It was found the highest rainfall recorded within the six months is in April with reading of 499 mm, and minimum total rainfall recorded is in January with reading of 6 mm. Total rainfall recorded during the year is 1044.2 mm.

In conclusion, the level of rainfall recorded for the past five years in Kg. Separa, Skudai shows predictable pattern. The figures indicated that balance exists between the previous years' level of the highest rainfall with subsequent year's high rainfall level. The level was calculated using average rate for areas within close proximity.

5. Network Learning

Neural networks are used in pattern recognition and forecasting applications. Based on [8], Artificial Neural Network (ANN) has the advantages of working well in solving non-linear problems. However, it needs to be trained to learn the examples to perform the specific task. To date, ANN are being used to study application of previous existing data. Any study using previous existing data is known as training process and need to be done prior to sharpening of neural network's learning ability to simulate the behaviour of the system in different conditions.

Back Propagation Network training is a method of training artificial neural networks used in conjunction with the gradient system, which tries to minimise the Root Mean Square (RMS) [2]. The capability of artificial neural networks depends on trainings conducted. In ensuring favourable results are achieved from the training, availability of previous data is crucial. When more data is made available, it widens the scope of training and ensure betterment of the neural network. In this study, historical data were used approximately as per following proportion: 70% for training purpose, 20% for validation purpose and 10% for testing purpose.

In the phase of training, a volume of training for the model shall be performed. With the tests conducted, errors are able to be identified. Errors from initial training needs to be rectified before subsequent training to ensure better result. Hence, it is expected that the errors will decline in subsequent tests performed. To minimise training volume during this phase, training will stop once error reach its minimum point.

5.1. Distribution and Compilation Data

Subsequent to data normalisation, data were then organised, and divided to be used for three purposes; training, validation and testing. For the purpose of this study, rainfall data were collected from each area, from 2006 till June 2010 on daily basis. It should be noted that 2008 is a leap year, and has 366 days. In ensuring smoothness of training process later on, 2006, 2007 and 2009 had been standardised by adding an extra day in February, hence this leads to assumption that each year has 366 days, similar to 2008.

From the assumption made, data from each area totals to 1647. The data, were then divided by ratio 7:2:1 as described above. Based on the ratio, 1098 data were used for training purpose, 366 data were used for validation purpose and remaining 183 data were used for testing purpose.

Subsequent to the above process, data were then, compiled on weekly basis, and were grouped in the form of matrix (7 x *n*) where *n* equals the number of weeks and 7 stands for seven days in a week. From these groupings, the data structure for training is formed in a matrix (7 x 26). Table 1 below shows the parameter setting for the study.

5.2. Architecture Neural Network Model

The average rainfall for three different areas that were expected to able to predict are Senai, Kg. Seelong and Skudai. Each area is projected to have 27 models with learning algorithms *traingdx*. Each model is distinguished by:

i. The number of hidden nodes;
ii. The combination of constant learning rate and momentum

The network model is built up with 7 input nodes that each node represents data for a day of the week, one level of hidden node and 7 output nodes. Hidden nodes that were used are 7, 14 and 15. The number of hidden nodes was determined using the Kolmogorov theorem; $2n + 1$, where n is the number of input [7]. For the parameter of learning rate and constant momentum rate, the combination of 0.1, 0.5 and 0.9 are used. Table 2 shows a learning rate and constant momentum built for each of the hidden nodes.

Traingdx algorithm is a learning algorithm that is used to design this model. *Traingdx* algorithm's function is to supervise the learning process by updating the values of the weights and the bias of the network. The network is affected by the slope of the momentum factor and learning rate adjustment.

6. Experiment Result

Before conducting the experiment, the data obtained need to be normalised first. Then, the data are divided into three segments of different places. Set A represents Senai, Set B represents Kg. Seelong and Set C represents Kg. Separa. Table 3 indicates the details of the experiment data.

6.1. Experiment 1: Set A Senai

In Senai, the data needs a lot of training since the training performance targets do not always reach the target and cause the value become infinite. During the training, the maximum number of runs has been reached before the target is reached.

Thus, a total of 27 models were produced, five of the model that has the smallest MSE will be selected. Then, the correlation value for each of the five models will be calculated to determine the correlation between the forecast data and actual data. Table 4 shows the MSE and the correlation of the five selected models.

Model ($A_{7,1,1}$) is selected to be the best model for forecasting rainfall distribution for Senai. The Figure 4 below shows the comparison between actual data, forecast data and moving average data.

From the graph , the forecast data (forecast) are not in accordance with the pattern of the actual data (actual_forecast) . However, it can be seen on day 81 until day 117 for the prediction of rainfall rate higher than any other day that is equal to the real data rate of rainfall . In day 158 up to day 182, the patterns for forecast data and the pattern for actual rate is similar. Therefore, it can be seen in April and June is the rainy season rainfall rates based on forecasts and actual rainfall rates.

6.2. Experiment 2: Set B Kg. Seelong

For Kg. Seelong, a total of 27 models were built and five models that have small MSE value were selected. Next, the correlation of these five models will be calculated to find the relationship between the actual data and forecast data. Table 5 shows the selected five models have been selected and the correlation value for each model.

Hence, model ($B_{7,1,1}$) is selected because its correlation value is more significant compare with other models. The Figure 5 below indicates the comparison between actual data, forecast data and moving average data.

From the graph we can see the data of moving average (mvgavg_forecast) pattern of rainfall distribution is very similar to the actual data (actual_forecast). The data also shows that the moving average in January, February, March and June is the rainy season for Kg. Seelong. However, the forecast data (forecast) is more focused and closer to the actual data.

6.3. Experiment 3: Set C Kg. Separa

Among 27 models that has been built for Kg. Separa, there are five models by the smallest number of MSE selected. Table 6 below shows five models together with its number of hidden nodes and parameter learning rate and momentum parameters.

After that, the correlation of the models was calculated to find the relationship of actual data and forecast data. If the correlation value is bigger and closer to one, hence the relationship is stronger. The results of the correlation for the five models can be referred at the table below.

Model ($C_{7,5,1}$) is the best model among five models above since it has smallest MSE and bigger correlation value. The Figure 6 below indicates the comparison between actual data, forecast data and moving average data.

Based on the graph, the moving average data (mvgavg_forecast) is almost equal to the pattern with the actual data (actual_forecast) but for a more accurate, model ($C_{7,5,1}$) is equal to the actual data .

7. Discussion

From the results obtained from three experiments above, we can see that to get the best forecasting model, we should take into an account the correlation of each model besides looking at their MSE alone. This is due to the significance of these forecast data with model selected. Furthermore, in order to see the significance of the model, the structure of hidden nodes also can be changes or increases the number of input nodes and hidden nodes, [5]. However for this study, we only calculate the correlation of the relationship between the actual data and forecasts data and also the relationship of the two data is either inversely proportional or not.

Besides comparing the actual data and the forecast data using Artificial Neural Network, we also compare both data with moving average data. The data is also engaged in the implementation of the model in this study, but only for a minor use because Artificial Neural Networks technique implementation is a major comparison with the actual data.

For experiment 1, the model of moving average data has a pattern that is very similar to the actual data than the forecast data. This can be considered that the data in experiment 1 may inadequate to produce the best model for ANN technique. This is because ANN technique is requires a lot of data to be trained before reaching the target. Moreover, the probability of the structure of input nodes that inappropriate may cause it cannot capture the dynamic networking, [5].

For experiment 2 and 3, since the ANN technique model can generate forecast data similar to the actual data, the forecast data can be used as a network model for Kg. Seelong and Kg. Separa.

8. Limitation

There are number of limitations in this review. The study is focusing in three areas in Johor which is Senai, Kg. Seelong and Kg. Separa. The rainfall distribution data obtained from year 2006 until June 2010.

9. Conclusion

In summary, this study basically presents the results from experiments conducted using different data sets according to the study area. Each study area has 27 models that have been built and the model will be selected by the smallest MSE and the correlation is closer to one. Next, the selected model was then compared to the value of the data with the actual data and the moving average data.

Artificial Neural Network technique used to forecast rainfall in the study area in Johor have been applied and examined in this study. To get the accuracy rate in predicting rainfall is a complex challenge because the rainfall rate is constantly changing in proportion to time. Therefore, some experiments were carried out to produce some of the best model was chosen based on the smallest error. However, the correlations for some models that have small error values were calculated to see the significance and the strength of the relationship between the actual data and the forecasts data. Finally, the model has been able to produce a network that is closest to the actual data rate and the pattern of rainfall is also almost the same. Hence, the implementation model of Artificial Neural Network technique was suitable for rainfall forecasting model.

10. Figures

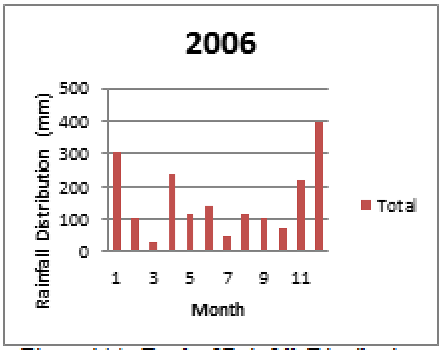

Figure 1(a): Total of Rainfall Distribution for Year 2006

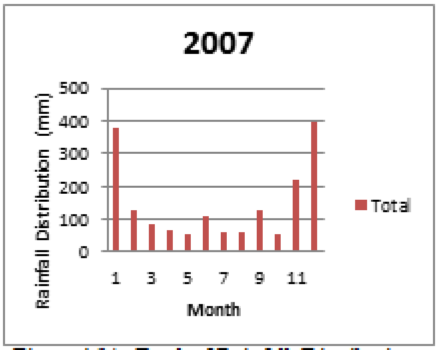

Figure 1(b): Total of Rainfall Distribution for Year 2007

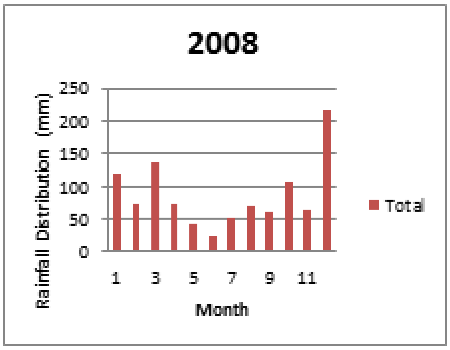

Figure 1(c): Total of Rainfall Distribution for Year 2008

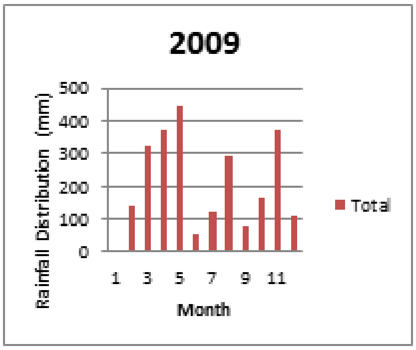

Figure 1(d): Total of Rainfall Distribution for Year 2009

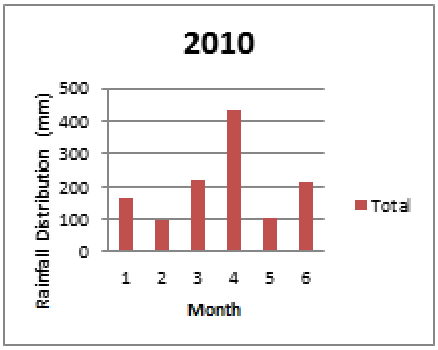

Figure 1(e): Total of Rainfall Distribution for Year 2010

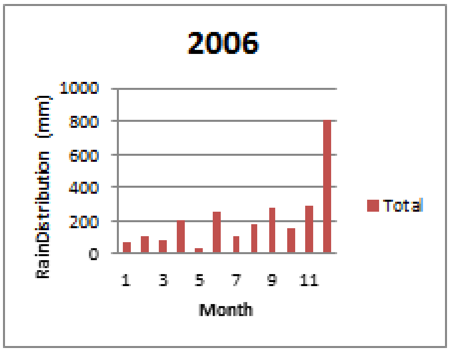

Figure 2(a): Total of Rainfall Distribution for Year 2006

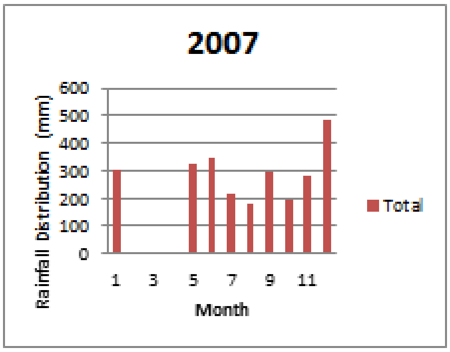

Figure 2(b): Total of Rainfall Distribution for Year 2007

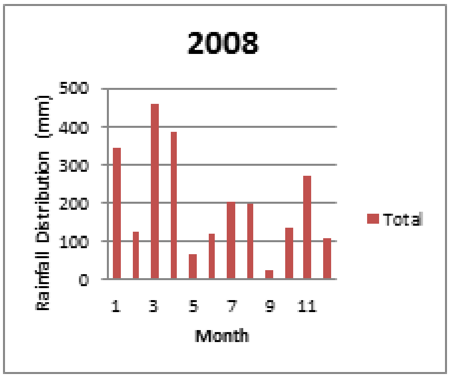

Figure 2(c): Total of Rainfall Distribution for Year 2008

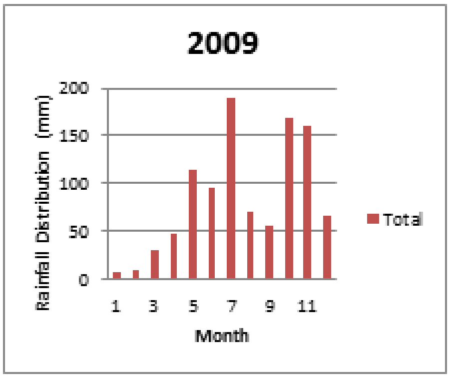

Figure 2(d): Total of Rainfall Distribution for Year 2009

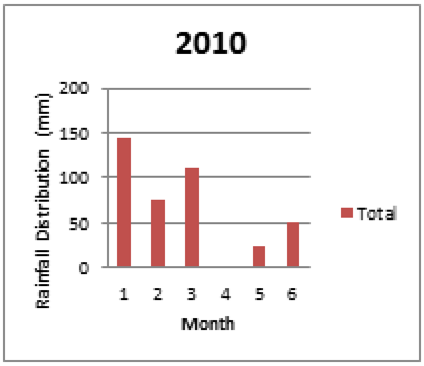

Figure 2(e): Total of Rainfall Distribution for Year 2010

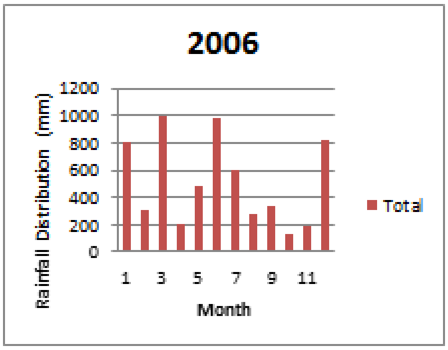

Figure 3(a): Total of Rainfall Distribution for Year 2006

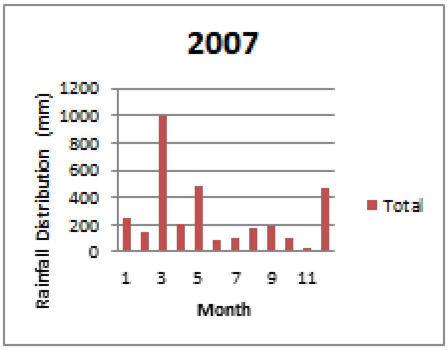

Figure 3(b): Total of Rainfall Distribution for Year 2007

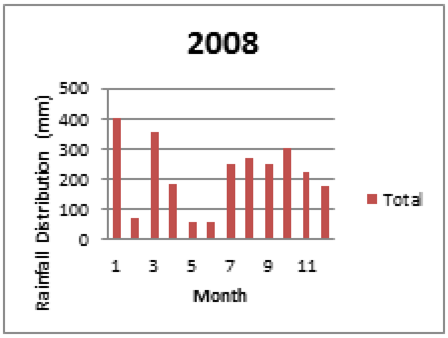

Figure 3(c): Total of Rainfall Distribution for Year 2008

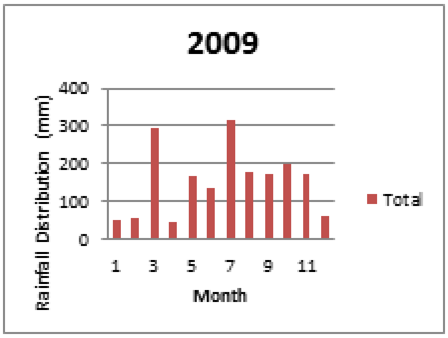

Figure 3(d): Total of Rainfall Distribution for Year 2009

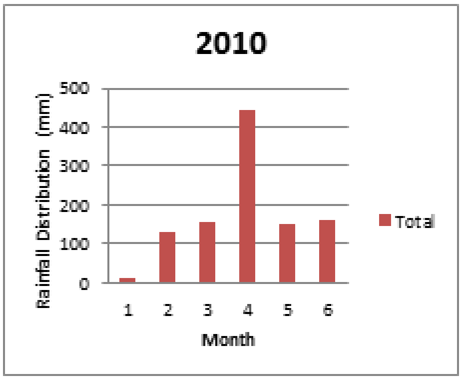

Figure 3(e): Total of Rainfall Distribution for Year 2010

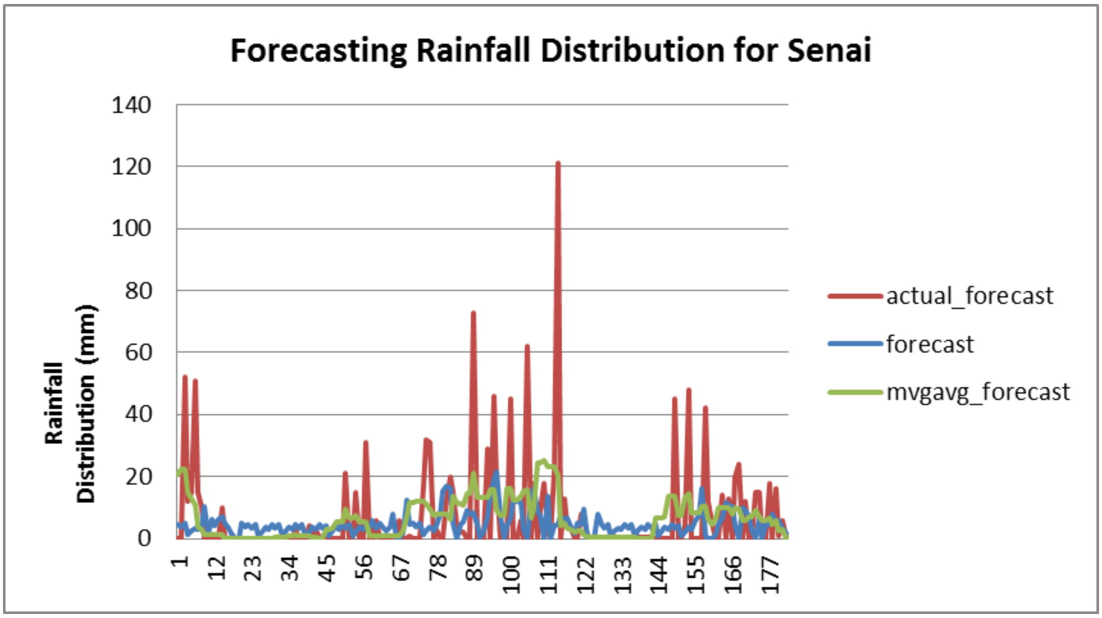

Figure 4. Comparison Graph for Senai

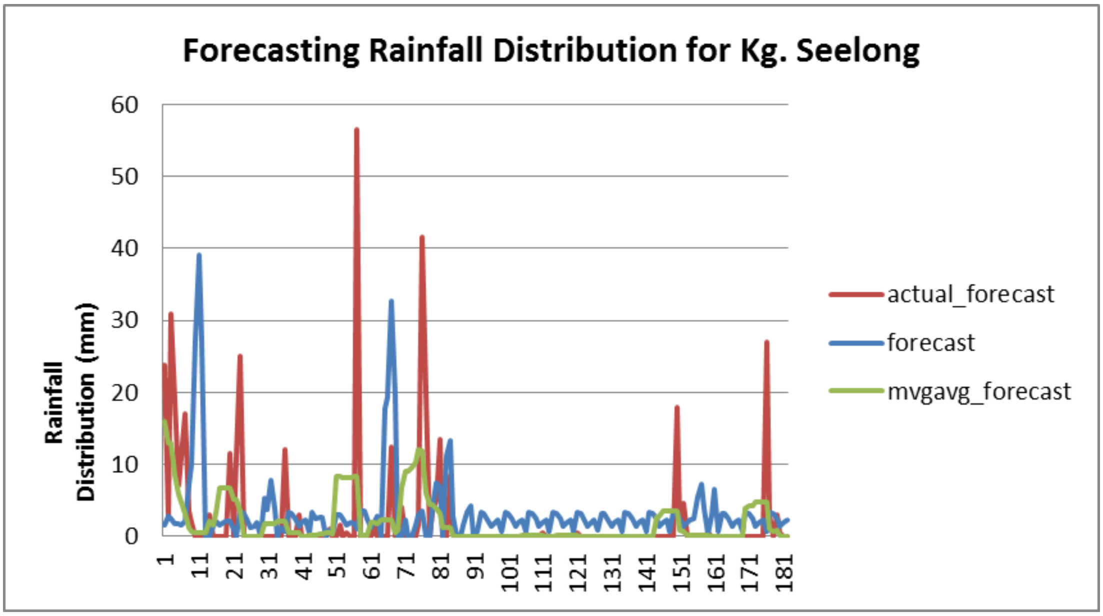

Figure 5. Comparison Graph for Kg. Seelong

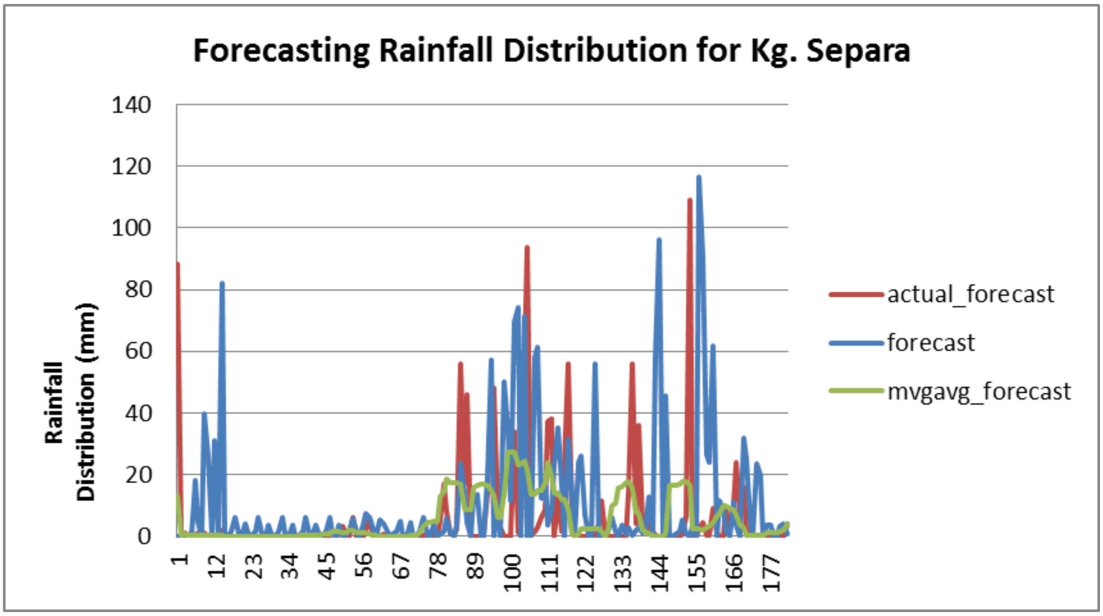

Figure 6. Comparison Graph for Kg. Separa

11. Tables

Table 1. Parameter Setting

Criterion	Settings
Input Data	Rainfall Distributions
Output Data	Rainfall Distributions
% Training Set	70
% Validation Set	20
% Testing Set	10
Performance Evaluation	Root Mean Squared Error

Table 2. Combination Learning rate and Momentum

Learning rate	Momentum
0.1	0.1
	0.5
	0.9
0.5	0.1
	0.5
	0.9
0.9	0.1
	0.5
	0.9

Table 3.Experiment Data

Set	Case Study	Year
A	Kg. Separa, Skudai	2006 – Jun 2010
B	Senai	2006 – Jun 2010
C	Kg. Seelong	2006 – Jun 2010

Table 4.Five Models Selected for Senai

Hidden Nodes	Learning Rate	Momentum Rate	MSE	Correlation (r)
7	0.1	0.1 ($B_{7,1,1}$)	0.0188	0.0752
		0.5 ($B_{7,1,5}$)	0.0184	-0.0383
		0.9 ($B_{7,1,9}$)	0.0173	0.0241
	0.5	0.9 ($B_{7,5,5}$)	0.0175	0.0016
14	0.5	0.5 ($B_{14,5,5}$)	0.0188	0

Table 5.Five Models Selected for Kg. Seelong

Hidden Nodes	Learning Rate	Momentum Rate	MSE	Correlation (r)
7	0.1	0.1 ($C_{7,1,1}$)	0.0426	0.0036
		0.9 ($C_{7,1,9}$)	0.0323	-0.0994
	0.5	0.5 ($C_{7,5,5}$)	0.0407	-0.0994
14	0.9	0.5 ($C_{14,9,5}$)	0.0450	-0.0352
15	0.1	0.5 ($C_{15,1,5}$)	0.0408	-0.0449

Table 6.Five Models Selected for Kg. Separa

Hidden Nodes	Learning Rate	Momentum Rate	MSE	Correlation (r)
7	0.1	0.1 ($A_{7,1,1}$)	0.0852	0.047122
		0.9 ($A_{7,1,9}$)	0.0610	-0.03951
	0.5	0.1 ($A_{7,5,1}$)	0.0688	0.086856
		0.5 ($A_{7,5,5}$)	0.0781	0.036115
15	0.5	0.5 ($A_{15,5,5}$)	0.0864	0.081029

Acknowledgement

This study is supported by Research University Grant (GUP) entitled “Image Super-Resolution Using Iterative Back Projection and Hybridised Biogeography Based Optimisation Algorithm” (08H23).

References

[1]Gutierrez-Osuna, R. and H. T. Nagle (1999). "A Method for Evaluating Data-Preprocessing Techniques for Odour Classification with an Array of Gas Sensors." Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on 29(5): 626-632.

[2]Hecht-Nielsen, R. 1989. "Theory of the Backpropagation Neural Network". In *Proceedings Of The International Joint Conference On Neural Networks* (Vol. 1, pp. 593–605). http://doi.org/10.1109/IJCNN.1989.118638.

[3]Herrero, J., R. Díaz-Uriarte and J. Dopazo (2003). "Gene Expression Data Preprocessing." Bioinformatics 19(5): 655-656.

[4]Juszczak, P., Tax, D. M. J., & Duin, R. P. W. 2002. "Feature scaling in support vector data description". *Proc. 8th Annu. Conf. Adv. School Comput. Imaging*, 95–10.

[5]Khalili N., Khodashenas S.R., Davary K. Dan Karimaldini F. (2011). *Daily Rainfall Forecasting for Mashhad Synoptic Station using Artificial Neural Network.* Universiti Mashhad. 2011 International Conference on Environmental and Computer Science, IPCBEE Vol. 19 (2011)

[6]Kotsiantis, S., D. Kanellopoulos and P. Pintelas (2006). "Data Preprocessing for Supervised Learning." International Journal of Computer Science 1(2): 111-117.

[7]Shamsuddin S. M. (2004). Lecture Note Advanced Artificial Intelligence: Number of Hidden Neurons. Unpublished note, *University Teknologi Malaysia*, Skudai.

[8]Stergiou, C., & Siganos, D. (1999) Neural Networks. *Webpage*.