Investigations

Backpropagation Vs. Radial Basis Function Neural Model: Rainfall Intensity Classification For Flood Prediction Using Meteorology Data

[1]Soo See Chai, [1]Wei Keat Wong and [2]Kok Luong Goh

[1]*Department of Software Engineering and Computing, Faculty of Computer Science and Information Technology, University of Malaysia Sarawak (UNIMAS), 94300, Kota Samarahan, Sarawak, Malaysia*
[2]*Faculty of Cognitive Science and Human Development, UNIMAS, 94300, Kota Samarahan, Sarawak, Malaysia*

Abstract: Rainfall is one of the important weather variables that vary in space and time. High mean daily rainfall (>30 mm) has a high possibility of resulting in flood. Accurate prediction of this variable would save human lives and properties. Soft computing methods have been widely applied in this field. Among the various soft computing methods, Artificial Neural Network (ANN) is the most commonly used methodology. While numerous ANN algorithms were applied, the most commonly applied are the Backpropagation (BPN) and Radial Basis Function (RFN) models. However, there was no research conducted to verify which model among these two produces a superior result. Therefore, this study will fill this gap. In this study, using the meteorology data, the two ANN models were trained to classify the rainfall intensity based on four different classes: Light (<10 mm), moderate (11-30 mm), heavy (31-50 mm) and very heavy (>51 mm). The architecture of the neural networks models based on the different combination of inputs and number of hidden neurons to obtain the optimum classification were verified in this study. The influence of the number of training data on the classification results was also analyzed. Results obtained showed, in term of classification accuracy, BPN model performed better than the RFN model. However, in term of consistency, the RFN model outperformed BPN model.

Keywords: Precipitation, Classification, Backpropagation, Radial Basis Function, Neural Networks

Introduction

Weather forecasting is a complicated procedure yet the most essential and vital process for the mankind nowadays, because it severely affect human activities. Highly accurate weather forecast could help to prevent casualties and damages. Amongst all the weather happenings, floods are the leading cause of natural disaster death world-wide and were responsible for 6.8 million death in the 20th century (Doocy *et al*., 2013). Rainfall intensity is important for flood warning system. Considering the alert system involved in heavy rain management, it would be useful to classify the rainfall intensity based on different threshold. The depth or intensity of the rainfall and its distribution in the temporal and spatial dimensions depend on many

variables, such as pressure, temperature, wind speed and direction (Luk *et al*., 2001). Understanding the complex physical processes that create the rainfall is very challenging. Large number of attempts has been made by researchers to accurately predict rainfall. However, the accuracy obtained by these techniques is still below satisfactory level due to the nonlinear nature of rainfall (Nayak *et al*., 2013). Artificial Neural Network (ANN), which has the ability in handling complex and non-linear problems, has drawn the attention of researchers in the field of weather forecasting. Among the different architectures of ANN, the Backpropagation Network (BPN) and Radial Basis Function (RBF) network are the two main models that are sufficiently suitable for precipitation prediction (Shrivastava *et al*., 2012). Although a lot of works have been done using these two

architectures, the superiority of one architecture over another is not being discussed. Therefore, this paper aims to compare and analyze the performance of these two architectures for rainfall classification.

Data

The study area selected for this study is Kuching city, the capital city of Sarawak located in the east Malaysia. Since Malaysia is located on top of equator, Kuching city has a tropical type of climate with average of five to six hours of sunshine, high temperature and high humidity. The city, located at the southwest of the Sarawak state with latitude 1.6019N and longitude 110.3244E, covers area of 895.09 km^2 and a population of 681,901 (Wikipedia 2015). A collection of historical meteorology parameters of daily measurement was obtained from the Malaysian Meteorological Department. These daily meteorology data from year 2009 to 2013 consisted of seven elements: Minimum temperature (°C), maximum temperature (°C), mean temperature (°C), mean relative humidity (%), mean wind speed (m/s), mean sea level pressure (hPa) and mean precipitation (mm).

Data Pre-Processing

Noise and missing data would affect the performance of ANN models (Sola and Sevilla, 1997). Thus, before training and testing the ANN, it is important to perform data checking and cleansing to maximize the performance of ANN forecasting. In this research, missing and incomplete data which was confirmed by the Sarawak Meteorology Station was deleted from the database.

Data Normalization

The pre-processed database is next gone through normalization. Normalization aims to produce good result and prevent numerical difficulties occurs when performing calculation (Chen *et al.*, 2013). Moreover, according to (Chai *et al.*, 2009), normalization speeds up the training process of the ANN and reduces the likelihood of the ANN getting stuck in local minima. Adequate data normalization before applying it into the ANN can reduce the estimation error generated by the ANN in a factor between 5 and 10 (Sola and Sevilla, 1997). In this study, the input data was normalized so that the minimum and maximum values for each input row are between +1 and -1.

Rainfall Intensity Classification

The rainfall intensity for year 2009 to 2013 is classified into four classes: Light precipitation (<10 mm), moderate precipitation (11-30 mm), heavy precipitation (31-50 mm) and very heavy precipitation (> 51 mm). In the research to estimate rainfall using radar for the Klang River Basin in Selangor, Malaysia (Ramli *et al.*, 2011), three classes namely low (<10

mm), moderate (>10, <30 mm) and heavy (>30 mm) were used. However, in this study, the heavy precipitation class (>30 mm) was sub-divided into 2 classes: Heavy precipitation (31-50 mm) and very heavy precipitation (>51 mm) as rainfall of more than 50 mm could be termed as "hazard precipitation" (Szalińska *et al.*, 2014). Table 1 summarized the different classes used in this study.

Methodology

Input and Output of ANN Models

The input of the ANN models included 6 meteorological data obtained from Department of Irrigation and Drainage (DID) Kuching Division of Sarawak, Malaysia. These six parameters were: Daily minimum temperature (°C), maximum temperature (°C), mean temperature (°C), mean relative humidity (%), mean wind speed (m/s), mean sea level pressure (hPa). Each input node of the ANN models consists of an array of different parameter values at a different time period. The output of the ANN models would be the different class of the rainfall intensity as shown in Table 1.

Data Discretization for Training and Testing Process

The ANN models will be trained by providing "examples" for the models to learn. The "well-learned" ANN will next be tested with some unseen data. In order to accomplish this, the meteorology data obtained were divided into training and testing data according to Table 2. Table 2 showed that, in order to make a fair comparison of the ANN models, the same set of data, i.e., 1 month data ranging from 1 to 31 Dec. 2013, was used for testing. The training of the data was divided into 5 groups of different amount of training data.

BPN Architecture

One single hidden layer feed forward network can approximate any measureable function arbitrarily well regardless of the activation function, the dimension of the input space and the input space environment (Hornik *et al.*, 1989). Therefore, in this study, one hidden layer Backpropagation neural Network (BPN) with Levenberg-Marquardt learning algorithm is used.

RFN Architecture

Radial functions are class of functions which could be applied in any sort of model (linear or non-linear) and any sort of network (single- or multi-layer) (Orr, 1996). Moreover, the single-weight layer network (the input component is feed-forward to the basis functions whose outputs are linearly combined with weights into the net-work outputs) is associated with the traditional RFN model (Broomhead and Lowe, 1988). Consequently, in this study, the single-weight layer RFN model with Gaussian learning algorithm is used.

Table 1. Rainfall intensity classification

Rainfall event Classification	Precipitation (mm)	Class
Light Precipitation	<10	1
Moderate Precipitation	11-30	2
Heavy Precipitation	31-50	3
Very Heavy Precipitation	>51	4

Table 2. Data discretization for training and testing of ANN models

Groups	Training data size	Testing data size
1	12 months data (1 Jan 2009 to 31 Dec 2009)	One month data (1 Dec 2013 to 31 Dec 2013)
2	24 months data (1 Jan 2009 to 31 Dec 2010)	One month data (1 Dec 2013 to 31 Dec 2013)
3	36 months data (1 Jan 2009 to 31 Dec 2011)	One month data (1 Dec 2013 to 31 Dec 2013)
4	48 months data (1 Jan 2009 to 31 Dec 2012)	One month data (1 Dec 2013 to 31 Dec 2013)
5	59 months data (1 Jan 2009 to 31 Dec 2013)	One month data (1 Dec 2013 to 31 Dec 2013)

Experiments Setup

In order to obtain the optimal ANN architecture, the following experiments were carried out:

- Number of hidden neurons
- Number of training data
- Different combination of input data

Number of Hidden Neurons

The number of hidden neurons will influence the error on the nodes to which their output is connected (Sheela and Deepa, 2013). With too many hidden neurons, the system will overestimate and incapable of generalization. On the other hand, having too few hidden neurons will prevent the network from fitting the input data properly and therefore, the robustness of the network would be reduced. According to Sheela and Deepa (2013) who reviewed the methods to fix hidden neuron in neural network for the past 20 years, the existing methods to determine number of hidden neurons are all trial-and-error rule, i.e., experimenting. Although a lot of efforts were taken by the researchers in developing approaches to estimate the number of hidden neurons in the hidden layer, the approximation was also dependable on the type of the database samples which the network is designed for, the number of training samples and the complexity of the target problem. In order to have a rough approximation of the lower and upper bound of the number of hidden neuron, theorem by Paugam-Moisy and Helene (1997) with the formulas below were applied in this research.

Lower bound of hidden neurons:

$$\frac{N_P N_S}{N_I + N_S} \tag{1}$$

Upper bound of hidden neurons:

$$2\left[\frac{N_P}{N_I + N_S}\right] N_S \tag{2}$$

Where:
N_P = The number of learning sets
N_I = The number of inputs
N_S = The number of output

Using these formulas, the estimation for lower and upper bound of hidden neurons for 12 months data would be 52 and 104 respectively. However, according to the review, the researchers in the field of rainfall prediction had utilized less than 20 hidden neurons. Therefore, an average of 10 hidden neurons was selected to be the lower bound as comparing to the calculated figure.

In order to determine the optimal number of hidden neurons for both the BPN and RBFN models, a series of experiments with the number of hidden neurons ranging from 10, 50 and 100 were used. The models were each trained with 12 months of meteorology data and test with 1 month of unseen data (Group 1 in Table 2). Table 3 shows the MSE and R of each of the network model using 10, 50 and 100 hidden neurons. From the graphs in Fig. 1, it can be clearly seen that, the performance of both network models decreases when the number of hidden neuron increases. In term of variance, it could be seen that, the MSE and R for BPN model has higher values (0.004 for both) as comparing to the RFN model (0.002 and 0.003 respectively). The optimum number of hidden neurons for both models was found to be 10 using the 6 inputs.

Number of Training Data

The general belief of training a neural network model is that, with more training data, the network model would be able to generalize better as comparing to training a network model with small training data.

According to Zhu *et al.* (2012), for a given model, one would expect performance to generally increase with the amount of data, but eventually saturate. In their study, they found that additional training data decrease the performance of the network. Therefore, to investigate the minimum amount of data required to

train the network models, the network model with 10 hidden neurons (obtained from previous experiment), 6 inputs and 1 output were trained with different amount of training data. The discretization of the training and testing data is shown in Table 2. Note that, the testing data used was the same for all the groups.

Table 3. BPN model and RFN model performance comparison by using different hidden layer neurons

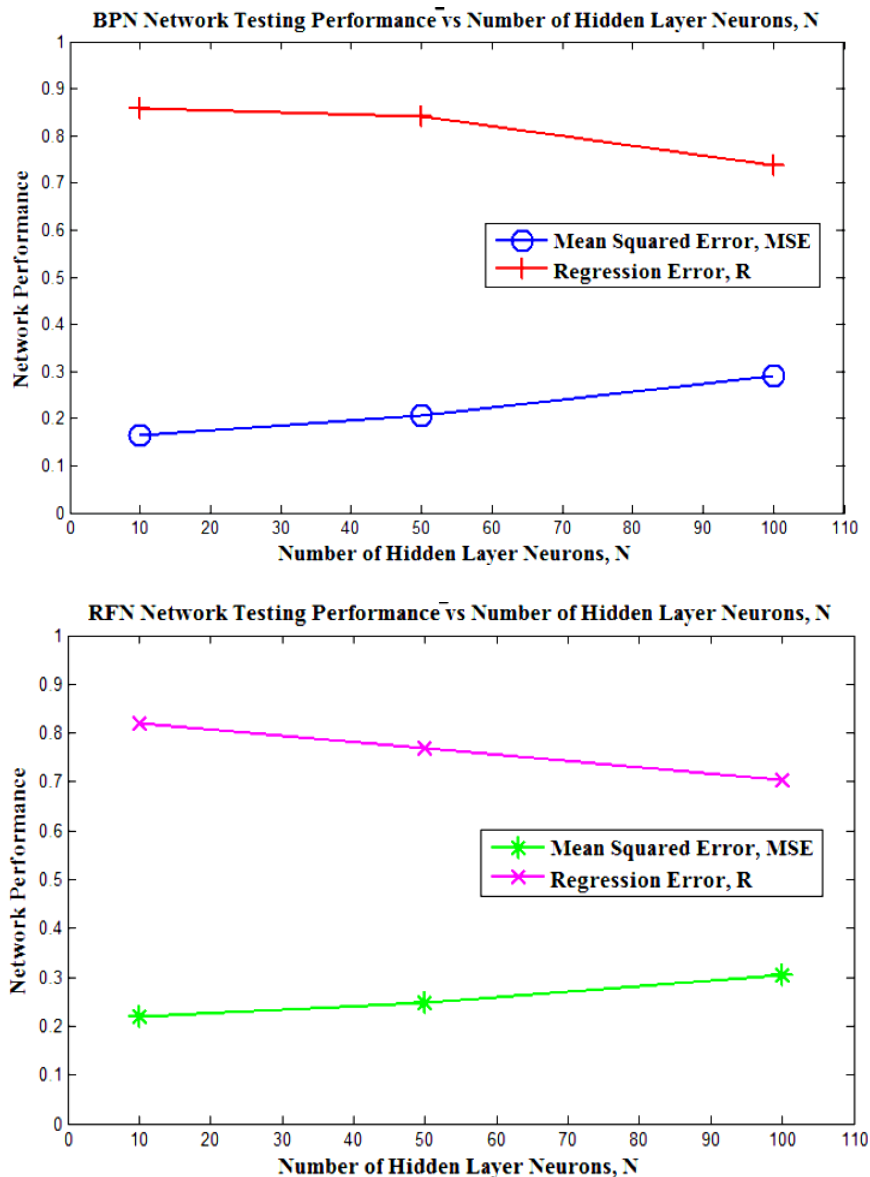| No. of hidden neurons	BPN		RFN	
	MSE	R	MSE	R
10	0.1647	0.8575	0.2206	0.8191
50	0.2065	0.8399	0.2473	0.7763
100	0.2898	0.7366	0.3035	0.7048

Fig. 1. The performance of BPN (top) and RFN (bottom) models in term of MSE and R using 10, 50 and 100 hidden neurons

BPN Model Testing Performance vs Size of Training Data (Months)

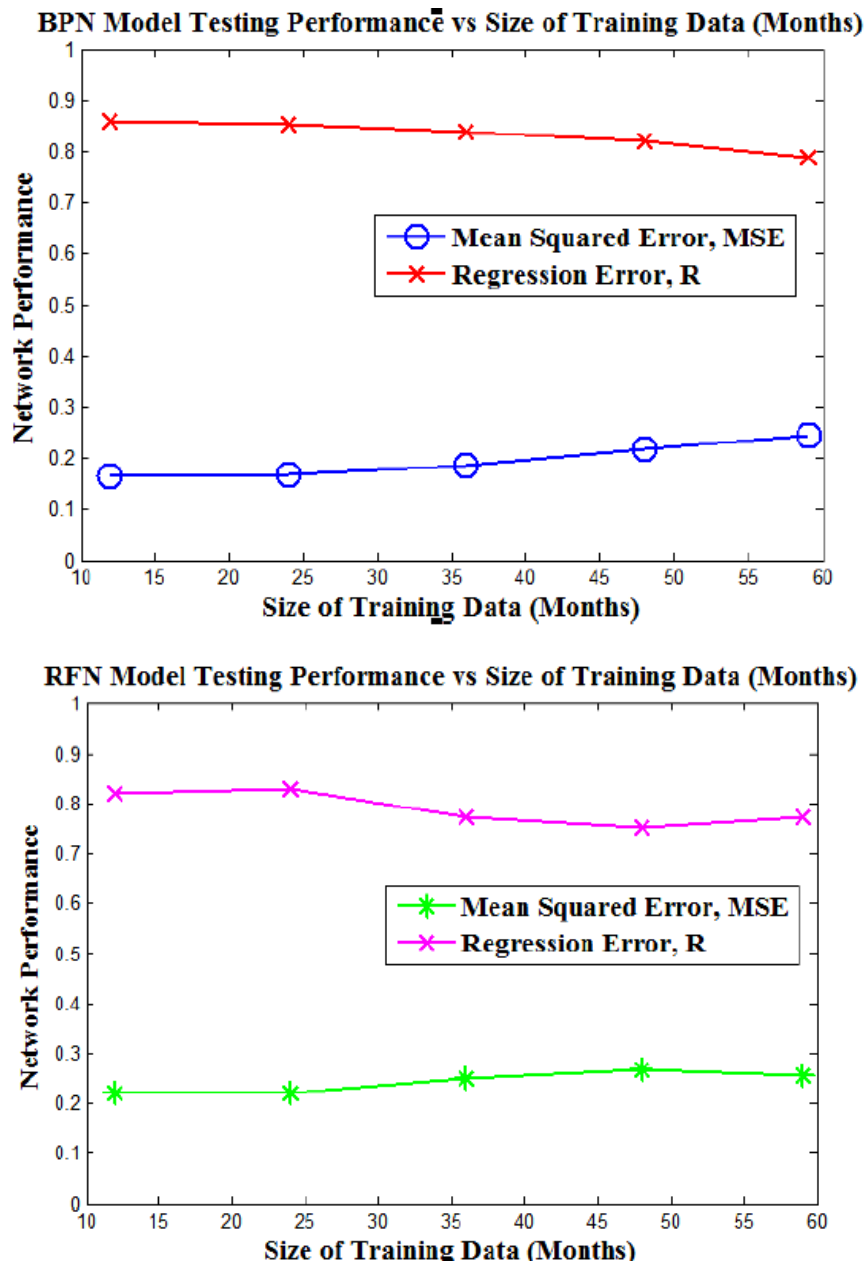**RFN Model Testing Performance vs Size of Training Data (Months)**

Fig. 2. Performance of BPN (top) and RFN (bottom) models using different training size

From the graphs in Fig. 2, it could be seen that, when the training size increases from 12 to 59 months, the performance of the network models decreases. RFN model was found to produce more consistent accuracy with variance of MSE around 0.0005 as comparing to BPN with variance of MSE around 0.001.

Different Combination of Input Data

Selecting the best subset of the input variables is the critical issue in forecasting (Utans *et al.*, 1995). For the case of data driven model like neural networks, the in-put variables are selected from available data and there is no prior assumption of the functional form of the model based on some physical interpretation of the underlying system or process being modeled (Suzuki, 2011). Unrequired input data can significantly increase the learning complexity. Therefore, in order to select the best input variables for rainfall intensity classification using the meteorology data, different combination of input variables were used to find the optimal inputs for the network models. A series of 15 experiments were conducted for this purpose. The inputs of these 15

experiments were listed out in Table 5. The network architecture used was: *m*:10:1 whereby *m* stands for the different number of combination of inputs used.

From the previous experiment, the number of training data used in this experiment is 12 months of training data (Group 1 of Table 4) as it produced the best accuracy result.

From Table 6, it could be seen that, each of the network model performs differently with the different input combination. The BPN model performs the best when the sea level pressure is omitted (MSE = 0.1523,

R = 0.8704) while the RFN model performs best when the wind speed is omitted (MSE = 0.1885, R = 0.8467). For BPN model, the top five best performance of data combinations included: The data combination of all in-put variable except the sea level pressure, the data combination of relative humidity and sea level pressure, the data combination of temperature and relative humidity, the data combination of all input variables except the wind speed and lastly the data combinations of all six elements of meteorology data.

Table 4. BPN and RFN network performance using different training size

	BPN		RFN	
Data training size	MSE	R	MSE	R
12 months (Group 1)	0.1647	0.8575	0.2206	0.8191
24 months (Group 2)	0.1672	0.8511	0.2199	0.8282
36 months (Group 3)	0.1855	0.8375	0.2513	0.7745
48 months (Group 4)	0.2171	0.8204	0.2667	0.7539
59 months (Group 5)	0.2428	0.7891	0.2547	0.7726

Table 5. Abbreviation for different meteorology data combinations

Abbreviation	Data combination (Daily meteorology data)
RH	Relative Humidity
SLP	Sea Level Pressure
Temp	Minimum temperature, maximum temperature, mean temperature
WS	Wind Speed
RHSLP	Relative Humidity and Sea Level Pressure
RHWS	Relative Humidity and Wind Speed
TempRH	Temperature and Relative Humidity
TempSLP	Temperature and Sea Level Pressure
TempWS	Temperature and Wind Speed
WSSLP xRH	Wind Speed and Sea Level Pressure
(without RH) xSLP	Sea level pressure, minimum temperature, maximum temperature, mean temperature and wind speed
(without SLP) xTemp	Relative humidity, minimum temperature, maximum temperature, mean temperature and wind speed
(without Temp) xWS	Relative humidity, sea level pressure and wind speed
(without WS)	Sea level pressure, sea level pressure, minimum temperature, maximum temperature and mean temperature
ALL	Relative humidity, sea level pressure, minimum temperature, maximum temperature, mean temperature and wind speed

Table 6. Performance of BPN model and RFN models using different combination of meteorology data sorted by ascending order of MSE values

BPN			RFN		
Data combination	MSE	R	Data combination	MSE	R
xSLP	0.1523	0.8704	xWS	0.1885	0.8467
RHSLP	0.1524	0.8702	RHSLP	0.2090	0.8345
TempRH	0.1559	0.8656	ALL	0.2206	0.8191
xWS	0.1647	0.8575	xSLP	0.2207	0.8328
ALL	0.1647	0.8575	RHWS	0.2282	0.8011
RH	0.1796	0.8680	RH	0.2338	0.8599
RHWS	0.2060	0.8262	xTemp	0.2575	0.8067
TempSLP	0.2017	0.8368	Temp	0.2580	0.7485
xTemp	0.2107	0.8368	TempRH	0.2598	0.7611
TempWS	0.2210	0.7885	TempSLP	0.2831	0.7440
xRH	0.2412	0.7833	xRH	0.3416	0.6536
Temp	0.2555	0.7581	TempWS	0.3450	0.6403
WSSLP	0.4534	0.5494	WSSLP	0.5139	0.5784
SLP	0.4659	0.8680	SLP	0.6312	0.1501
WS	0.6230	0.1782	WS	0.6453	0.1997

On the other hand, the RFN model performs the best with the following 5 types of data combination: Data combination of all input variables except the wind speed, the data combination of relative humidity and sea level pressure, the data combination of all meteorology data and the data combination of all input variable except the sea level pressure. Table 6 also clearly shows that, the use of single input variable for both network models did not produce descent output. It could be further concluded that in order to produce decent weather forecasting result, either the relative humidity data or temperature data is needed to be included in weather forecast model training.

Discussion

One primary observation from the experiments is that, the RFN model produced a more consistent result in term of accuracy as comparing to the BPN model. This is clearly shown by the smaller variance in the MSE and R values for the experiments of using different hidden neuron and the different number of training data.

In order to further illustrate this, for the experiment of using different hidden neuron of the 59 months data, each of the model was validated with a subset of the training data that was unseen by the network models during the training. A total of 15 training and validations were done for each network model. Figure 3 shows the MSE values obtained for both the BPN and RFN models. In Fig. 3, the MSE values for RFN model during validation process shows consistent value as comparing to the BPN model which shows high fluctuation in the MSE values obtained. This observation is due to the fact that during the BPN training and testing processes, the network generates random weight to adjust or to improve the training and learning process. Due to the random weight provided into the training, the results that are generated vary in training and testing process. On the other hand, RFN model produces a consistent result as comparing to BPN model due to pre-adjustment of 'spread' in the specification of network properties.

The second observation obtained from the series of the experiment is that, with the fixed number of inputs and output, using the same amount of training data, the increase of hidden neuron did not improve the accuracy of the networks. This could be explained by that, the 10 hidden neurons contain the adequate degree of freedom to generalize the target problem and therefore, the increase of the number of hidden neurons worsens the performance of the network models.

In term of number of training data, the general belief is that the accuracy of the neural network could be improved better as the number of data grows. However, from the experiment under the section *Number of Training Data*, it was observed that, this might not be true as the classification accuracy for both network models deteriorate as more training data was provided for the network models to learn. Note that, the network was trained with increasing number of training data but tested with the same data set. In the research carried out by (Kavzoglu 2009) for image classification, in order to improve the accuracy of the neural network, besides increasing the number of training data, the quality of the training samples is also crucial. For neural network to match the target problem, the samples provided for training the neural network must be representative. In our case, when more data was provided for training, the training sample might be more representative for a certain class of the classification. A detailed analysis of the target data for the different group of training data is shown in Table 7. In this table, the number of data for the different rainfall intensity of Class 1 (Light Precipitation) to 4 (Very Heavy Precipitation) is shown clearly. From Table 7, it could be seen that, the data provided for training included all the classes needed for testing. However, when the number of training data increased from 12 months to 59 months, the difference between the numbers of sample data provided for each class became larger. Therefore, when the number of training data increases from 12 months to 59 months, the accuracy of the classification deteriorates as the data provided for training is not representative for each class. This affects both the neural network models but the effects of this is clearer in BPN as comparing to RFN model due to the pre-adjustment of 'spread' in the specification of network properties in RFN model.

Table 7. The distribution of the target data for rainfall intensity classification of Class 1 to 4 when the training data used range from 12, 24, 36, 48 and 59 months

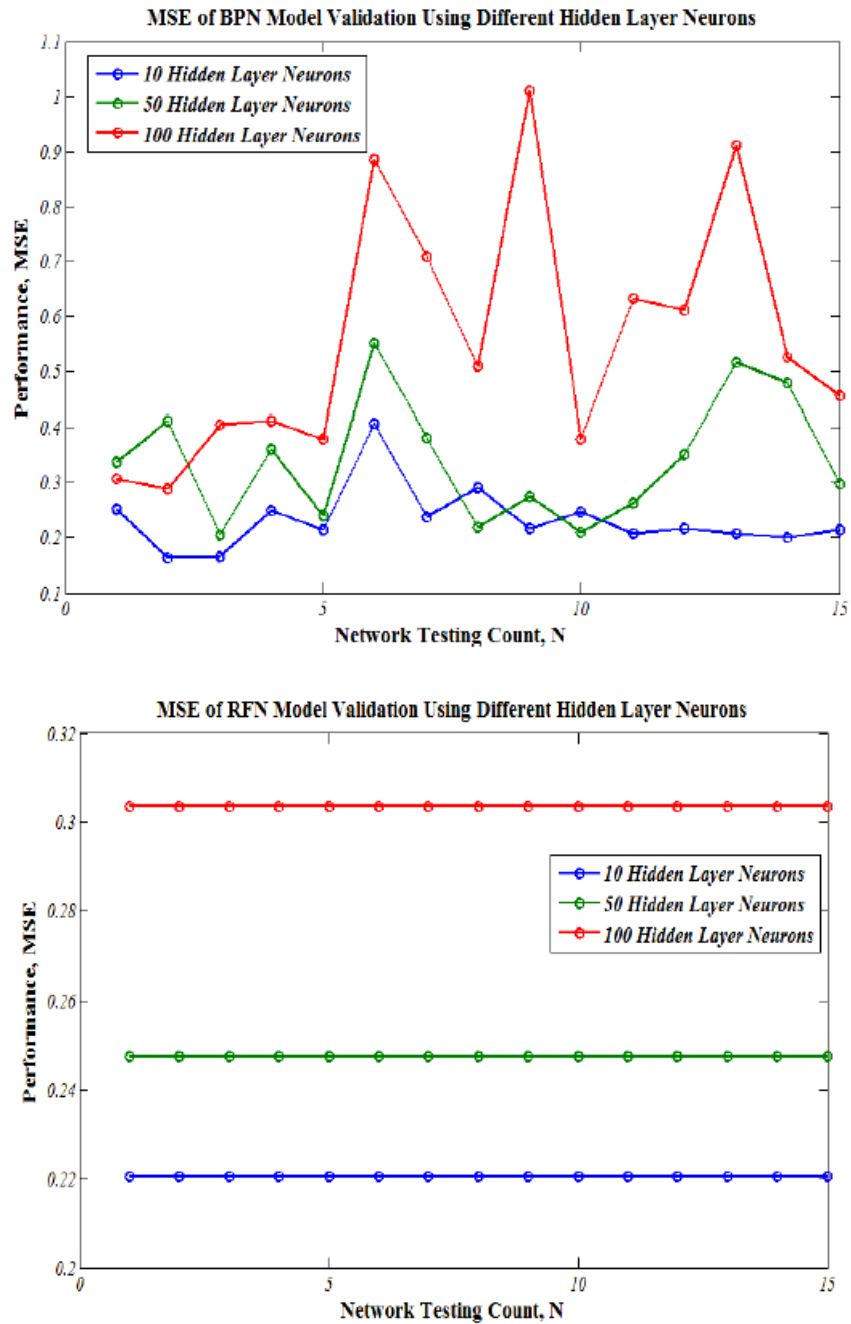| Class | No. of target data | | | | | |
	12 months	24 months	36 months	48 months	59 months	Testing data
1	255	507	762	1024	1268	16
2	64	126	191	251	299	8
3	22	52	69	95	122	1
4	24	45	69	85	95	6

Fig. 3. MSE values for the validation of the BPN (top) and RFN (bottom) model using 10, 50 and 100 hidden neurons

The use of different inputs for the neural network models influenced the accuracy of the rainfall intensity classification. The experiments in *Different Combination of Input Data* showed that, the use of wind speed data and sea level pressure data will deteriorate the classification accuracy of rainfall intensity for both models. For BPN, the model performs the best without the sea level pressure while the RFN achieves the best accuracy result without the wind speed values. On the other hand, the relative humidity data and temperature data are the compulsory meteorology data that needed to be included in weather forecast model training in order for the ANN models to achieve descent accuracy. Another valuable observation from the experiments is that, although there exist a parameter that could deteriorate the accuracy result, the ANN models were able to tolerate by generalizing well when all the input data were used.

Conclusion

From the study, it could be clearly seen that, in order for ANN models to perform well in rainfall intensity classification, there is a need to consider the architecture, training data and the input of the ANN models. Among the two common neural network models used for rainfall classification, based on the data set used, the BPN performs better in term of achieving better classification results as comparing to the RFN model. However, in term of consistency, it could be concluded that, RFN is able to perform better. Although the accuracy results varied when different hidden neurons and training data sets were used, the variances of the accuracy obtained were not as far as the variance obtained for the BPN model. Therefore, depending on the norm of the research, BPN is a useful data driven model when the data for training could be verified of being representative as this will promise descent accuracy. On the other hand, RFN is a good choice in order to obtain a consistent result.

Acknowledgement

Funding Information

Author's Contributions

Soo See Chai: Supervised and wrote majority of the manuscript.

Wei Keat Wong: Collected the data and worked on the experiments.

Kok Luong Goh: Contributed ideas, insights and discussion for the research work.

References

Broomhead, D.S. and D. Lowe, 1988. Radial basis functions, multi-variable functional interpolation and adaptive networks. DTIC Document.

Chai, S.S., J.P. Walker, O. Makarynskyy, M. Kuhn and B. Veenendaal et al., 2009. Use of soil moisture variability in artificial neural network retrieval of soil moisture. Remote Sens., 2: 166-190. DOI: 10.3390/rs2010166

Chen, S., Y. Wang and I. Tsou, 2013. Using artificial neural network approach for modelling rainfall-runoff due to typhoon. J. Earth Syst. Sci., 122: 399-405. DOI: 10.1007/s12040-013-0289-8

Doocy, S., A. Daniels, C. Packer, A. Dick and T.D. Kirsch, 2013. The human impact of earthquakes: A historical review of events 1980-2009 and systematic literature review. PLOS Currents. DOI: 10.1371/currents.dis.67bd14fe457f1db0b5433a8ee20fb833

Hornik, K., M. Stinchcombe and H. White, 1989. Multilayer feedforward networks are universal approximators. Neural Netw., 2: 359-366. DOI: 10.1016/0893-6080(89)90020-8

Kavzoglu, T., 2009. Increasing the accuracy of neural network classification using refined training data. Environ. Modell. Software, 24: 850-858. DOI: 10.1016/j.envsoft.2008.11.012

Luk, K.C., J.E. Ball and A. Sharma, 2001. An application of artificial neural networks for rainfall forecasting. Math. Comput. Modell., 33: 683-693. DOI: 10.1016/S0895-7177(00)00272-7

Nayak, D.R., A. Mahapatra and P. Mishra, 2013. A survey on rainfall prediction using artificial neural network. Int. J. Comput. Applic., 72: 32-40.

Orr, M.J., 1996. Introduction to radial basis function networks. Technical Report, Center for Cognitive Science, University of Edinburgh.

Paugam-Moisy, H. and A.E. Helene, 1997. Size of multilayer networks for exact learning: Analytic approach. Proceedings of the Conference on Advances in Neural Information Processing Systems, (IPS' 97), MIT Press.

Ramli, S., S.H.A. Bakar and W. Tahir, 2011. Radar hydrology: New Z/R relationships for Klang River Basin, Malaysia based on rainfall classification. Proceedings of the IEEE Colloquium on Humanities, Science and Engineering, Dec. 5-6, IEEE Xplore Press, Penang, pp: 537-541. DOI: 10.1109/CHUSER.2011.6163790

Sheela, K.G. and S. Deepa, 2013. Review on methods to fix number of hidden neurons in neural networks. Math. Prob. Eng., 2013: 425740-425750. DOI: 10.1155/2013/425740

Shrivastava, G., S. Karmakar, M.K. Kowar and P. Guhathakurta, 2012. Application of artificial neural networks I n weather forecasting: A comprehensive literature review. Int. J. Comput. Applic., 51: 0975-8887. DOI: 10.5120/8142-1867

Sola, J. and J. Sevilla, 1997. Importance of input data normalization for the application of neural networks to complex industrial problems. IEEE Trans. Nuclear Sci., 44: 1464-1468. DOI: 10.1109/23.589532

Suzuki, K., 2011. Artificial Neural Networks-Methodological Advances and Biomedical Applications. 1st Edn., InTech, Rijeka, ISBN-10: 9533072431, pp: 362.

Szalińska, W., I. Otop and T. Tokarczyk, 2014. Precipitation extremes during flooding in the Odra River Basin in May-June 2010. Meteorol. Hydrol. Water Manage., 2: 13-21.

Utans, J., J. Moody, S. Rehfuss and H. Siegelmann, 1995. Input variable selection for neural networks: Application to predicting the U.S. business cycle. Proceedings of the IEEE/IAFE Computational Intelligence for Financial Engineering, Apr. 9-11, IEEE Xplore Press, New York, pp: 118-122. DOI: 10.1109/CIFER.1995.495263

Zhu, X., C. Vondrick, D. Ramanan and C. Fowlkes, 2012. Do we need more training data or better models for object detection? Proceedings of the British Machine Vision Conference, (MVC' 12), BMVA Press, pp: 80.1-80.11. DOI: 10.5244/C.26.80

Wikipedia, 2015. Demographics of Sarawak. Wikipedia, The Free Encyclopedia.