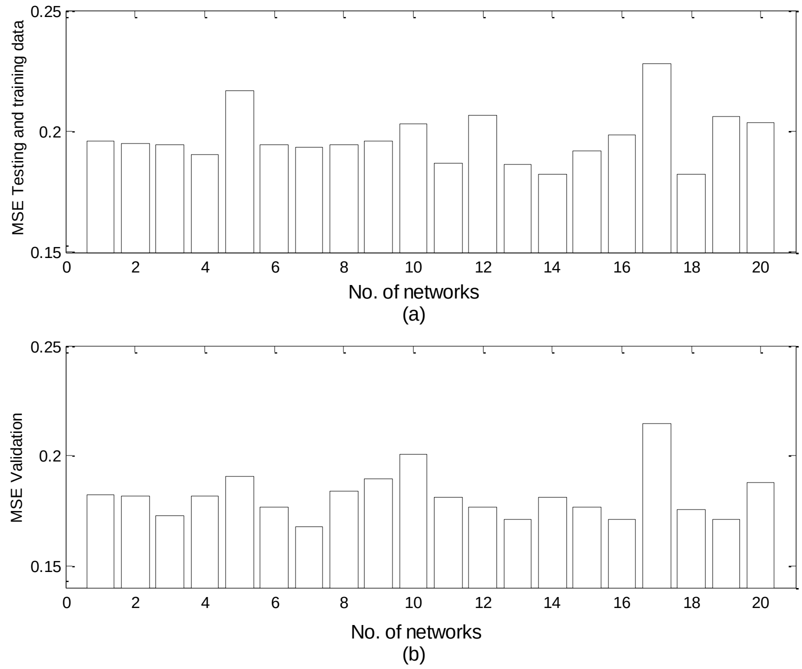

Fig. 7: MSE for single FANN.
(a) MSE for Training and Testing data, (b) MSE for Validation data.

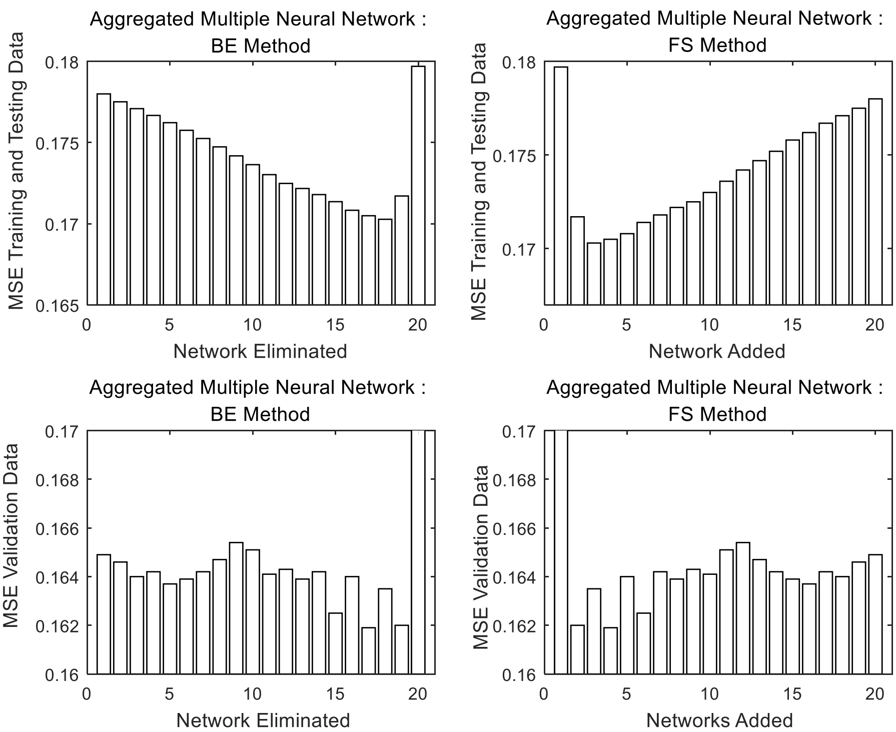

Fig. 8: MSE for aggregated multiple neural networks on the unseen validation data for BE and FS approaches.

Figure 8 shows the multiple neural network performance using selective combinations with BE and FS methods. The performance of aggregated networks on training and testing data is consistent with the performance on the unseen validation data for both selective combination methods. The reduction of MSE in training and testing data for BE and FS combinations are consistent with the reduction of MSE in the unseen/validation data. It shows the robustness of the proposed modelling techniques as compared to the single FANN where the best network performance in training and testing data will not guarantee the best performance in the unseen validation data. The numbers of networks for the final combination are reduced to 3 networks for both methods which show the minimum MSE in Training and testing data that also correspond to MSE in validation data. The final result analysis is shown in Table 2. In this particular case, the FS and BE approaches led to the same individual networks being combined. Even though the number of networks combined was quite small for both selective methods, the most important thing is that both combination approaches perform better than the single FANN.

Table 2: Statistical Analysis of MNN performance on the unseen validation data

	Number of networks combined	MSE	R^2
Single FANN	1	0.1856	0.7950
Combined all MNN	20	0.1649	0.8170
FS Aggregated MNN	3 (12,15,20)	0.1635	0.8200
BE Aggregated MNN	3 (12,15,20)	0.1635	0.8200

As for comparison, Azid et al. [22, 23] did carry out API modelling for the Southern region of Malaysia with 2 different sets of data containing 202,050 and 232,505 observations respectively. In [22] the input was reduced to 10 from 12 possible inputs with the R^2 and RMSE of 0.724 and 7.562 for unseen validation data respectively. On the other hand, in [23], the input was reduced to 5 from a possible 8 with the R^2 and RMSE of 0.618 and 10.017 for unseen data respectively. Therefore, the MNN did perform better than the [22] and [23] API modelling for Malaysia as shown in Table 2 with the R^2 and RMSE of 0.8200 and 0.160 for unseen validation data respectively. This performance was obtained with fewer sample data (1388 observations) as compared in [22] and [23].

5. CONCLUSION

This study proposes single FANN and multiple neural networks to model API based on the environmental monitoring data to get reliable and fast API predictions in order to mitigate the problems related to API. The single FANN does model the API quite well with relatively small MSE and high R^2 values on the unseen data. However, in order to overcome the non-robust nature of single FANN, multiple neural networks are proposed with two selective combination methods. Both selective combination methods further improve the model prediction as compared to single FANN and combining all networks. This clearly shows that it is possible to reduce the number of networks combined for the API prediction without losses in performance.

ACKNOWLEDGEMENT

The authors would like to acknowledge the support from the Universiti Sains Malaysia (USM) and Newcastle University, United Kingdom, and special gratitude to Department of Environment (DOE) Malaysia for providing and giving permission to utilize their air quality data for this study.

REFERENCES

[1] ASMA. “Air Pollutant Index (API),“ Retrieved on July, 2012. Available from http://www.doe.gov.my/portalv1/en/info-umum/english-air-pollutant-index-api/100

[2] Akkoyunlu A, Yetilmezsoy K, Erturk F, Oztemel E. (2010) A neural network-based approach for the prediction of urban SO_2 concentrations in the Istanbul metropolitan area. International Journal of Environment and Pollution, 40(4):301-315.

[3] Wang W, Lu W, Wang X , Leung YT. (2003) Prediction of maximum daily ozone level using combined neural network and statistical characteristics. Environmental International, 29(5):555–562.

[4] Viotti P, Liuti G, Di Genova P. (2002) Atmospheric urban pollution: applications of an artificial neural network (ANN) to the city of Perugia. Ecol Modell., 148(1):27-46.

[5] Sabri G, Tarek KM. (2012) Combination of artificial neural network models for air quality predictions for the region of Annaba, Algeria. Int. J. Environ. Stud., 69(1):79-89.

[6] Amodio M, Andriani E, Cafagna I, Caselli M, Daresta BE, de Gennaro G, Tutino M. (2010) A statistical investigation about sources of PM in South Italy. Atmos. Res., 98:207-218.

[7] Rodriguez S, Querol X, Alastuey A, Kallos G, Kakaliagou O. (2001) Saharan dust contributions to PM10 and TSP levels in Southern and Eastern Spain. Atmos. Environ., 35:2433-2447.

[8] PeyJ, Pérez N, Querol X, Alastuey A, Cusack M, Reche C. (2010) Intense winter atmospheric pollution episodes affecting the Western Mediterranean. Sci. Total Environ., 408(8):1951–1959.

[9] Pohjola MA, Rantamäki M, Kukkonen J, Karppinen A, Berge E. (2004) Meteorological evaluation of a severe air pollution episode in Helsinki on 27-29 December 1995. Boreal Environ. Res., 9(1):75–87.

[10] De Gennaro G, Trizio L, Di Gilio A, Pey J, Pérez N, Cusack M, Querol X. (2013) Neural network model for the prediction of PM10 daily concentrations in two sites in the Western Mediterranean. Sci. Total Environ., 463-464:875–883.

[11] Perez P, Trier A, Reyes J. (2000) Prediction of PM2.5 concentrations several hours in advance using neural networks in Santiago, Chile. Atmos. Environ., 34:1189–-196.

[12] Yan CK, Jian L.(2013) Identification of significant factors for air pollution levels using a neural network based knowledge discovery system. Neurocomputing, 99: 564-569.

[13] Gardner MW, Dorling SR. (1998) Artificial neural networks (the multilayer perceptron) - A review of applications in the atmospheric sciences. Atmos. Environ., 32(14-15):2627-2636.

[14] Perez P, Trier A. (2001) Prediction of NO and NO_2 concentrations near a street with heavy traffic in Santiago, Chile. Atmos. Environ., 35:1783-1789.

[15] Sousa S, Martins F, Alvimferraz M, Pereira M. (2007) Multiple linear regression and artificial neural networks based on principal components to predict ozone concentrations. Environ. Model Softw., 22(1):97-103.

[16] Ul-Saufie AZ, Yahaya AS, Ramli NA, Rosaida N, Hamid HA. (2013) Future daily PM10 concentrations prediction by combining regression models and feedforward backpropagation models with principle component analysis (PCA). Atmos. Environ., 77:621-630.

[17] Cigizoglu KH, Kisi Ö. (2006) Methods to improve the neural network performance in suspended sediment estimation. J. Hydrol., 317:221-238.

[18] Chelani AB, Chalapati RC, Phadke K, Hasan M. (2002) Prediction of sulfur dioxide concentration using artificial neural networks. Environ. Model Softw., 17:161-168.

[19] Gardner MW, Dorling SR. (1999) Neural network modelling and prediction of hourly NO and NO concentrations in urban air in London. Atmos. Environ., 33:709-719.

[20] Zhang J. (1999) “Developing Robust Non-linear Models Through Bootstrap Aggregated Neural Networks. Neurocomputing, 25:93-113.

[21] Ahmad Z, Zhang J. (2009) Selective combination of multiple neural networks for improving model prediction in nonlinear systems modelling through forward selection and backward elimination. Neurocomputing, 72:1198-1204.

[22] Azid A, Juahir H, Latif MT, Mohd Zain S, Osman MR. (2003) Feed-Forward artificial neural network model for Air Pollutant Index prediction in the southern region of Peninsular Malaysia. Journal of Environmental Protection, 4:1-10.

[23] Azid A, Juahir H, Toriman ME, Kamarudin MKA, Mohd Saudi AS, ,Che Hasnam CN, Abdul Aziz NA, Zaman F, Latif MT, Mohamed Zainuddin SF, Osman MR, Yamin M. (2014) Prediction of the Level of Air Pollution Using Principal Component Analysis and Artificial Neural Network Techniques: a Case Study in Malaysia. Water Air Soil Pollut., 225:2063-2077.

NOMENCLATURE

ANN	Artificial Neural Network	-
API	Air Pollution Index	-
BE	Backward Elimination	-
CO	carbon monoxide	mg/l
FANN	Feed-forward Artificial Neural Network	-
FS	Forward Selection	-
MLP	Multi-Layer Perceptron	-
MNN	Multiple Neural Networks	-
MSE	Mean sum square error	-
n	Number of networks combined	-
NO	Nitrogen monoxide	mg/l
NO_2	Nitrogen Oxide	mg/l
O_3	Ozone	mg/l
PCA	Principle Component Analysis	-
PM_{10}	Concentration of particulate matter with a size less than 10 microns	mg/l
$PM_{2.5}$	Concentration of particulate matter with a size less than 2.5 microns	mg/l
R^2	Coefficient determination	g/mol
X	Input data	-
$\hat{X}$	Input data after resampling	-
Y	Output data	-
$\hat{Y}$	Network Prediction Output	-

Subscript

i	Number of network