CHEMICAL ENGINEERING TRANSACTIONS

VOL. 56, 2017

A publication of AIDIC

The Italian Association of Chemical Engineering
Online at www.aidic.it/cet

Guest Editors: Jiří Jaromír Klemeš, Peng Yen Liew, Wai Shin Ho, Jeng Shiun Lim

ISBN 978-88-95608-47-1; **ISSN** 2283-9216

DOI: 10.3303/CET1756015

Prediction of Combustion Efficiency using Multiple Neural Networks

Zainal Ahmad*[,a], Alireza Bahadori[b], Jie Zhang[c]

[a]School of Chemical Engineering, Engineering Campus, Universiti Sains Malaysia, Seri Ampangan, 14300, Nibong Tebal Penang, Malaysia.
[b]School of Environment, Science and Engineering, Southern Cross University, Lismore NSW, Australia.
[c]School of Chemical Engineering and Advanced Materials, Newcastle University , Newcastle upon Tyne NE1 7RU, UK.
chzahmad@usm.my

In order to improve the generalisation capability of neural network based models, combining multiple neural networks (MNN) is proposed in this paper with the application of predicting the combustion efficiency from the boiler. This is due to the fact that single feed forward artificial neural networks (FANN) lack of the robustness due to the overfitting of the models. Combination of MNN was introduced and researchers concentrate on how overfitting can be avoided by combining the single FANN. In this study, the individual FANN are trained using different training data sets and /or from different initial weights, then combined. Instead of choosing the best single FANN model among the networks, all the neural networks are combined. It can also be described as architecture of network consisting of several sub-models and a mechanism which combines the outputs of these sub-models. In this study bootstrap application or bootstrap technique were apply to replicate the initial raw data or to create different training and testing data sets. Bootstrap basically relate or deals with the sampling to create random data sets for training and testing. By creating an equal number of bad and good data sampling, it actually improves the generalisation capability of FANN. The simple averaging method was applied in combining the MNN. The data for modelling was taken from energy management handbook with total of 66 data points where the training and testing consist of 39 samples data while unseen data consist of 27 samples data. The result shows that the MNN combination with simple averaging method did slightly improved the model prediction of the combustion efficiency by using two inputs which are stack temperature rise, dT and also the excess air. The coefficient of determination, r^2, and root mean squared error (RMSE) for unseen data for MNN are 0.9999 and 0.0105.

1. Introduction

Artificial neural network (ANN) had emerged as an attractive tool for non-linear multivariate modelling in the last two decades as Desai et al. (2008) apply it as model for the optimisation as compare to response surface methodology (RSM) and further improvement of the robusness of the single ANN in Zainal et al. (2010). It had typically been used as a “black-box” tool, which was, no prior knowledge about the process was assumed but the goal was to develop a process model based only on observations of its input-output behaviour (Psichogios and Ungar, 1992). ANN was a mathematical system that simulated biological neural networks and was often described as a massively interconnected network structure consisting of many simple processing elements (neurons) with the ability to perform parallel computation for data processing (Agatonovic-Kustrin et al., 1998) and also as tools for optimisation (Baş and Boyacı, 2007). ANN was capable of handling multiple independent and dependent variables simultaneously and to do this prior knowledge on the functional relationship did not need to be known. Each neuron received information through input connections, processed the information and produced the output, which was distributed through output connections. A neural network in its basic form was usually composed of several layers of neurons, there being one input layer, one output layer and at least one hidden layer.
There were various types of ANNs and the most common and popular network was the feedforward artificial neural network (FANN). In FANN, the information from various sets of inputs was fed forward through the

network to optimise the weight between neurons, or to train it. As described by Agatonovic-Kustrin et al. (1998), the error or bias in prediction was then propagated through the system and the inter-unit connections were changed to minimise the error in prediction. A typical neural networks structure can be depicted in Figure 1.

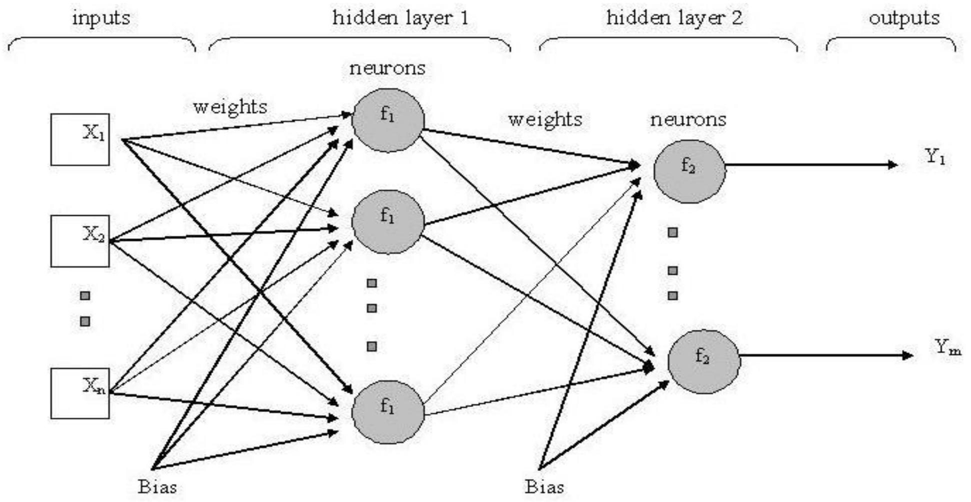

Figure 1: A typical structure of neural networks

A major disadvantage of single FANN was the difficulty in explaining the relation between independent and response variables resulting from the ambiguously defined weight, as a black box (Lou and Nakai, 2001).

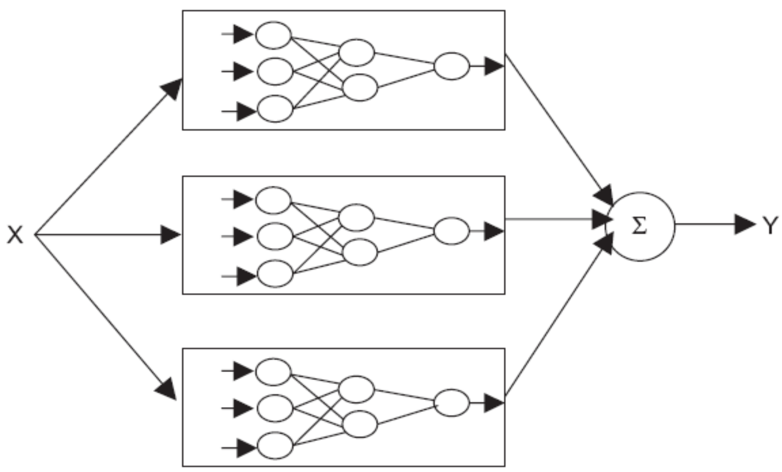

Figure 2: Combining multiple neural networks

Since a good process control performance was greatly dependent on the accuracy of model representation of the process, single FANN models must be robust or stable when they were applied to unseen data (Zainal et al., 2010) and it has been further developed by Ahmad et al. (2008) to improve the accuracy of the single ANN by using Bayesian Combination Predictor (BCP) as a method for final combination. Single neural networks sometimes lack robustness when the data is insufficient especially when dealing with real world data due to the fact that the robustness of the network is related to the representativeness of the training data (Bishop, 1995). Single FANN sometimes suffer badly when applied to unseen data where some neural network might fail to deliver the correct result due to the network training converged to undesired local minima (Hashem,1997), overfitting or noise in the data (McLoone and Irwin, 2001). The combination of multiple neural networks is introduced in this paper with the aim of enhancing the single FANN robustness. MNN basically consist of individual neural networks that are trained using different training data sets and/or from different initial weights, then combined. Instead of choosing the best neural network model among the networks, all the neural networks are combined as shown in Figure 2. The idea of multiple neural networks came up from Wolpert (1992), where he described about stacked generalisation which is a technique for combining different representations of single FANN to improve the overall prediction performance.

2. Single Feedforward Artificial Neural Network (FANN) and Multiple Neural Network (MNN) Model Development

The combustion efficiency for fire heater system of boiler is chosen in this study for the evaluation and analysis of the multiple neural networks (MNN) performance is taken from Turner and Doty (2007) and additional information about the single FANN modelling was taken from Bahadori et al. (2016) In this case study, 20 single FANN were developed from bootstrap re-samples of the original training and testing data. In re-sampling the training and testing data using bootstrap re-sampling techniques, the training and testing data was already in the discrete time function as shown in Eq(1). By re-sampling discrete time function, it does not affect the sequence of input-output mapping of the prediction. Figure 3 illustrates the analogy of bootstrap re-sampling technique. The numbers in the box represent the sample number (Zhang, 1999). Then the single FANN were trained by the Levenberg-Marquardt optimisation algorithm with regularisation and "early stopping". All weights and biases were randomly initialised in the range from –0.1 to 0.1.

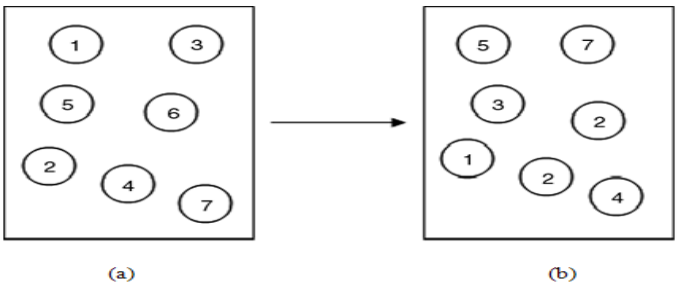

Figure 3: Bootstrap re-sampling: (a) Data samples in the original data set; (b) Data samples in the re-sampled data set (Zhang, 1999)

The individual FANN are single hidden layer feed forward neural networks. The hidden neurons use the logarithmic sigmoid activation function whereas output layer neurons use the linear activation function. In single FANN, the number of hidden neurons was determined using cross validation technique. The numbers of hidden nodes are increased from 5 to 15 and the SSE and r^2 value for the training and testing data are calculated for each node. The SSE and r^2 is plotted against the number of nodes. Different number of hidden neurons will give the different SSE and r^2 value in training and testing data. The network with the lowest SSE on the training and testing data was considered as having the best network topology for this prediction. In addition, in assessing the developed models, SSE on the unseen validation data is used as the performance criterion as well as the new unseen data from the literature. For this case study, the single FANN is developed based on the discrete time of the process as the prediction output at time (t), y(t), is predicted based on the process input at time t, u(t), as follows:

$$\hat{y}(t) = f[u_1(t), u_2(t)] \tag{1}$$

Where u(t) is the process input at time (t), where for this case study is stack temperature and excess air at time t, $\hat{y}(t)$ is the predicted process output at time t, which is the combustion efficiency. The MNN combined all the single FANN using simple averaging method instead of selecting a single neural network model, this approach will improve the accuracy and robustness of the prediction output. The final MNN model prediction is a weighted combination based on the simple averaging of the individual neural network outputs. The simple averaging method is the most common method in combining several model outputs with the weights fixed as shown below:

$$\hat{Y} = w_1\hat{y}_1 + w_2\hat{y}_2 + \cdots + w_n\hat{y}_n \tag{2}$$

where $\hat{y}_i$ is the network prediction from the ith network, where in this case is the efficiency of the combustion, n is the number of networks to be combined, in this case is 20, $\hat{Y}$ is the final prediction output, and w_i = 1/n is the weight for combining the i^{th} network.

3. Result and discussions

The input data of the single FANN model was selected and divided accordingly for training and testing data while the selection of the validation data was selected from the literatures. Single FANN with one hidden layer neural network architecture was utilised because it was sufficient to perform most of nonlinear process mapping. The numbers of data generated were divided into 3 data sets of training, testing and validation data comprising of 27 (80 % from 39 samples data), 12 (20 % from 39 samples data) and 27 samples. The data for training and testing was divided using MATLAB divideint command (MATLAB, 2016). The unseen validation data is the same as in Bahadori and Vuthaluru (2010).

Since there was no theoretical principle in choosing the proper network topology, the number of hidden neurons was determined using cross validation technique to obtain the best one. The numbers of neurons in the hidden layer were varied from 5 to 15 and the network was trained, tested and validated after each addition of neuron. The lowest SSE for trained and tested obtained was 0.0070 with 8 neurons in the hidden layer with SSE in validation of 0.1101 as shown in Table 1. The final single FANN structure or configuration for the combustion efficiency in the boiler was defined as 2-8-1 with the coefficient determination, r^2, of 0.9940.

Table 1: Sum Square Error (SSE) in single FANN with different numbers of hidden neurons

No. of hidden nodes	SSE (Train + Test)	SSE (Unseen/Validation)
5	0.0179	0.1426
6	0.0385	0.2032
7	0.0093	0.0567
8	0.0070	0.1101
9	0.0105	0.0471
10	0.0178	0.0674
11	0.0296	0.1382
12	0.0093	0.0561
13	0.0215	0.0635
14	0.0130	0.0511
15	0.0109	0.0695

The log-sigmoid transfer function was used as the activation function for hidden layer, and linear transfer function was applied for the output layer. The predicted data for training and testing sets are illustrated in Figure 4 for single FANN model with the final structure of 2-8-1. It was obvious that the actual and predicted values were more or less the same and in agreement with each other for the model.

Evaluation of the established single FANN models with the validation or unseen data indicated that generated single FANN models were able to present the combustion efficiency processes quite accurately as shown in Figure 5 in which implied an excellent generalisation capacity of the network.

The MNN combination was applied after replication of the “best” single FANN with 2-8-1 structure. 20 replication of single FANN were made using bootstrap method then afterwards combined it with the simple averaging combination approach.

Comparison of the predicted and actual values for single FANN and MNN are presented in Table 2. It was found that the MNN model was capable in defining the true behaviour of combustion efficiency system with slightly more accurate model as compared to single FANN with the RMSE and coefficient determination, r^2, of 0.0105 and 0.9999.

Table 2: Coefficient determination and root mean squared error for unseen validation data for single FANN and MNN models

	Single FANN		MNN (combination of 20 networks)	
Data set	Root mean squared error (RMSE)	Coefficient of determination (r^2)	Root mean square error (RMSE)	Coefficient of determination (r^2)
Unseen/Validation	0.0855	0.9940	0.0105	0.9999

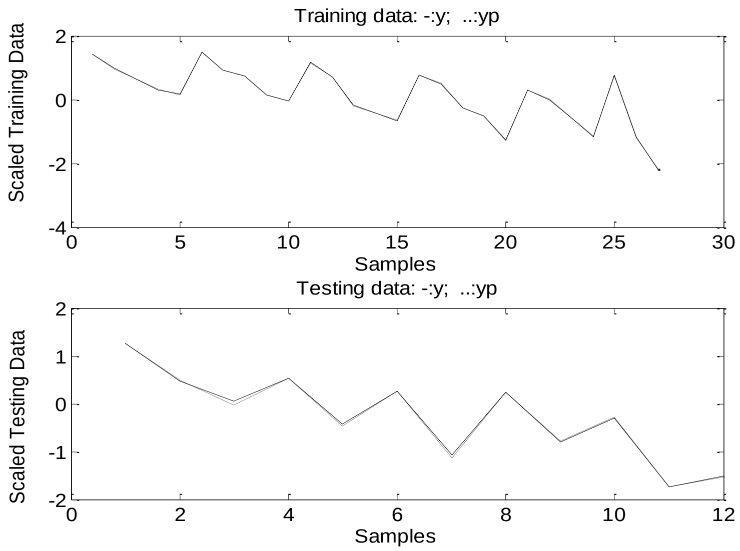

Figure 4: Actual and predicted for training and testing data for single FANN

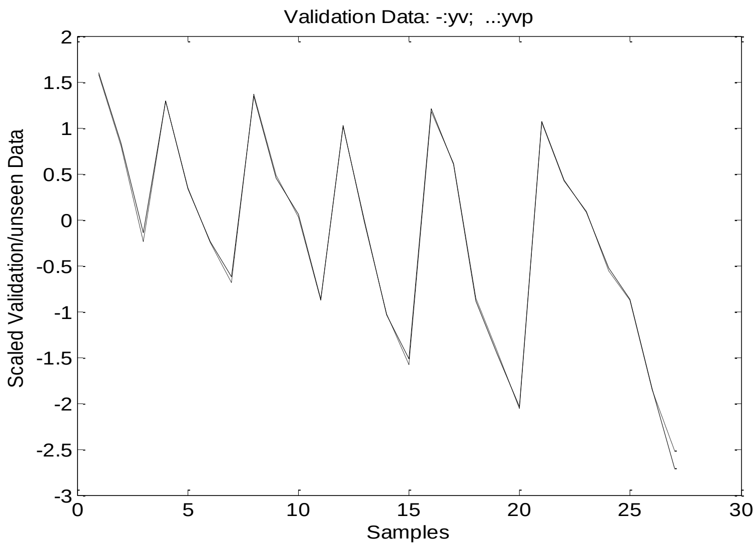

Figure 5: Actual and predicted for unseen/validation data for single FANN

4. Conclusions

In multiple neural networks, the generalisation capabilities of individual networks are not the same and, therefore, different networks generate different errors. Combining these networks can improve the robustness of the neural networks model by sharing and averaging out these errors.

Simple averaging or linear combinations method have been applied in this paper for MNN combination. MNN combination with simple averaging performed quite well with low RMSE and high coefficient determination which is closed to 1 for unseen validation data. The accumulation of prediction errors some time can make the predictions chaotic especially with dealt with the real data. The MNN model has shown its superiority as compared to single FANN in term of validation RMSE and coefficient determination. Based on the results and discussions presented in this paper, MNN combination seems to be a promising modelling method where it improved the robustness of the single FANN in predicting the combustion efficiency and the combination method is a quite simple and practical approach in many applications.

The main finding of this study demonstrates that the combination of individual networks did improve the robustness of the model as compared to single FANN models and its shows it’s superiorly in predicting the combustion efficiency in the boiler. The simple averaging method did improve the performance of the combined MNN model.

Acknowledgments

This work was supported by the Universiti Sains Malaysia (USM) through Graduate on Time (GOT) grant 8023006.

Reference

Agatonovic-Kustrin S., Zecevic M., Zivanovic L., Tucker I.G., 1998, Application of artificial neural networks in HPLC method development, Journal of Pharmaceutical and Biomedical Analysis 17, 69-76.

Ahmad Z., Tang P.H., Mat Noor R.A., 2008, Improving nonlinear process modeling using multiple neural network combination through Bayesian Model Averaging (BMA), IIUM Engineering Journal 9, 19-36.

Bahadori A., Baghban A., Bahadori M., Lee M., Ahmad Z., Zare M., Abdollahi E., 2016, Computational intelligent strategies to predict energy conservation benefits in excess air controlled gas-fired systems, Applied Thermal Engineering 102, 432–446.

Bahadori A., Vuthaluru H.B., 2010, Estimation of energy conservation benefits in excess air controlled gas-fired systems, Fuel Process Technology 91, 1198–1203.

Baş D., Boyacı İ.H., 2007, Modeling and optimization II: Comparison of estimation capabilities of response surface methodology with artificial neural networks in a biochemical reaction, Journal of Food Engineering 78, 846-854.

Bishop C., 1995, Neural Networks for Pattern Recognition, Clarendon Press, Oxford, UK.

Desai K.M., Survase S.A., Saudagar P.S., Lele S.S., Singhal R.S., 2008, Comparison of artificial neural network (ANN) and response surface methodology (RSM) in fermentation media optimization: Case study of fermentative production of scleroglucan, Biochemical Engineering Journal 41, 266–273.

Hashem S., 1997, Optimal Linear Combination, Neural Networks 10 (4), 599-614.

Lou W., Nakai S., 2001, Application of artificial neural networks for predicting the thermal inactivation of bacteria: a combined effect of temperature, pH and water activity, Food Research International 34, 573–579.

Mathlabd, 2016, MathWorks, Boston, United States.

McLoone S., Irwin G., 2001, Improving Neural Networks Training Solution Using Regularisation, Neurocomputing 37, 71-90.

Psichogios D.C., Ungar L.H., 1992, A hybrid neural network-first principles approach to process modeling, AIChE Journal 38, 1499-1511.

Turner W.C., Doty S., 2007, Energy Management Handbook, The Fairmont Press Inc., Lilburn, G, United States.

Wolpert D., 1992, Stacked generalization, Neural Networks 5, 241–259.

Zainal A., Mashitah M.D., Siti Hatijah M., Rabiatul Adawiah M.N., 2010, Nonlinear process modeling of fructosyltransferase (FTase) using bootstrap re-sampling neural network model, Bioprocess and Biosystem Engineering 33, 599-606.

Zhang J., 1999, Developing Robust Non-linear Models Through Bootstrap Aggregated Neural Networks, Neurocomputing 25, 93-113.