Modelling of a Post-combustion CO_2 Capture Process Using Neural Networks

Fei Li[1], Jie Zhang[1], Eni Oko[2], Meihong Wang[2]

[1]School of Chemical Engineering and Advanced Materials,
Newcastle University, Newcastle upon Tyne NE1 7RU, U.K.
e-mail: jie.zhang@newcastle.ac.uk
[2]School of Engineering, University of Hull, Hull HU6 7RX, U.K.

Abstract

This paper presents a study of modelling post-combustion CO_2 capture process using bootstrap aggregated neural networks. The neural network models predict CO_2 production rate and CO_2 capture level using the following variables as model inputs: inlet flue gas flow rate, CO_2 concentration in inlet flue gas, pressure of flue gas, temperature of flue gas, lean solvent flow rate, MEA concentration and temperature of lean solvent. In order to enhance model accuracy and reliability, multiple neural network models are developed from bootstrap re-sampling replications of the original training data and are combined. Bootstrap aggregated model can offer more accurate predictions than a single neural network, as well as provide model prediction confidence bounds. The developed neural network models can then be used in the optimisation of the CO_2 capture process.

Keywords: CO_2 capture; chemical absorption; neural networks; data-driven modelling; reliability.

1   Introduction

Post-combustion Carbon dioxide (CO_2) capture in coal fired power plants is attracting more attention as a result of the large amounts of existing coal fired power plants and none significant changes to equipment configurations required [1]. For the efficient design and operation of a post-combustion CO_2 capture plant, process optimisation is required. Process optimisation requires reliable and efficient process models. Different modelling technologies, such as mechanistic models and statistical models, have been studied to investigate the post-combustion carbon capture process efficiency. Previous studies showed that the establishment of mechanistic models is very time consuming and requires extensive knowledge of the underlying physics of the process. Numerical optimisation typically

required thousands of function evaluations. Evaluation of a detailed mechanistic model is typically computationally very demanding. To overcome this problem, neural network models can be developed from process operational data and used in plant optimization. Neural network models can be developed very quickly from process data and their evaluation is much less computationally demanding than a mechanistic model. Conventional neural networks sometimes suffer from poor generalisation performance due to the limitations in training data and training algorithms. More advanced neural network modelling methods should be utilised. This paper uses bootstrap aggregated neural networks to build data-driven models for a CO_2 capture process with chemical absorption.

The paper is organised as follows. Section 2 presents an overview of CO_2 capture processes. Section 3 presents bootstrap aggregated neural networks. Modelling of a CO_2 capture process using bootstrap aggregated neural networks is detailed in Section 4. Both static and dynamic models are developed. Section 5 draws some concluding remarks.

2   An overview of CO_2 capture processes

2.1   Methods for post-combustion CO_2 capture

Post-combustion CO_2 capture process removes CO_2 emission after the combustion of fossil fuel in a combustor. It is a technology which can be applied to retrofit the most existing coal-fired power plants for NO_x, SO_x and CO_2 capture. Several separation technologies can be employed in this process and they include adsorption, physical absorption, chemical absorption, cryogenics separation and membranes [1].

Specifically, adsorption is a physical process that adopts adsorbents to attach CO_2 to its surface. However, absorption process cannot be applied to large scale power plant flue gas treatment because of the low adsorption capacity of most available absorbents. Physical absorption is a physical process based on Henry's Law. The challenge is the high cost to treat the flue gas streams with CO_2 partial pressures lower than 15 vol%. Cryogenics separation condenses CO_2 in flue gas stream at -56.6 °C in atmospheric pressure and then removes it. As with the physical absorption, because of the high cost of operation, especially refrigeration, the technology can only be applied with high CO_2 concentration. Membrane absorption applies membrane only as a contact device between gas and liquid, while the membrane cannot provide supererogatory selectivity. The outstanding point of this process is the

membranes are more compact so that they will not be influenced by flooding, entrainment, channelling or foaming. Nevertheless, the pressures of gas and liquid phases are required to be equal to enable CO_2 cross the membrane easily. In this case, membrane absorption is only suitable for high concentration of CO_2 in flue gas, such as the flue gas coming from oxyfuel process.

Lastly, the chemical absorption is a chemical reaction process in which the chemical solvent is used to react with CO_2 to form a new intermediate compound with weak bound, and then CO_2 is regenerated in the circumstance of high temperature. The high selectivity and final pure CO_2 stream make chemical absorption being widely used for CO_2 capture of industrial flue gas [1].

2.2   Post-combustion capture via chemical absorption

As shown in Figure 1, CO_2 capture with chemical absorption is mainly consisted of two parts: the absorber and stripper columns, which are both packed columns. The flue gas from power plant is fed into the bottom of absorber and contacted counter-currently with lean amine solution from the top side. The lean amine solution chemically reacts and absorbs CO_2 in flue gas. Then the treated gas stream containing much lower CO_2 contents is generated and leaves from the top of absorber column. The amine solution of much more CO_2 (now rich amine), coming from the bottom of absorber, is pumped to the stripper unit after preheating in cross heat exchanger. In the stripper, the absorber amine solution is regenerated by heating rich amine in a reboiler. The low-pressure steam from power plant is used in reboiler to maintain the operating condition, resulting in large energy consumption. In details, the heat supplied in the reboiler is used for increasing the rich solution coming from the absorber, desorption heat required for separating CO_2 in rich amine, and vaporization of gas in stripper [2]. After that, the vapour is cooled in a condenser and returned to the stripper, while CO_2 leaves the condenser and is compressed for storage. In addition, the amine solution coming from the stripper (now lean amine) is cooled in cross heat exchanger by exchanging heat with rich amine and pumped back to absorber for absorption.

A significant portion of operation cost of CO_2 capture with chemical absorption is the energy requirement. In order to make CO_2 capture economically viable, the process operation should be optimised to identify the best process operating conditions such as temperature and pressure in absorber, stripper, reboiler and condenser. A reliable model is essential in carrying out the optimisation task.
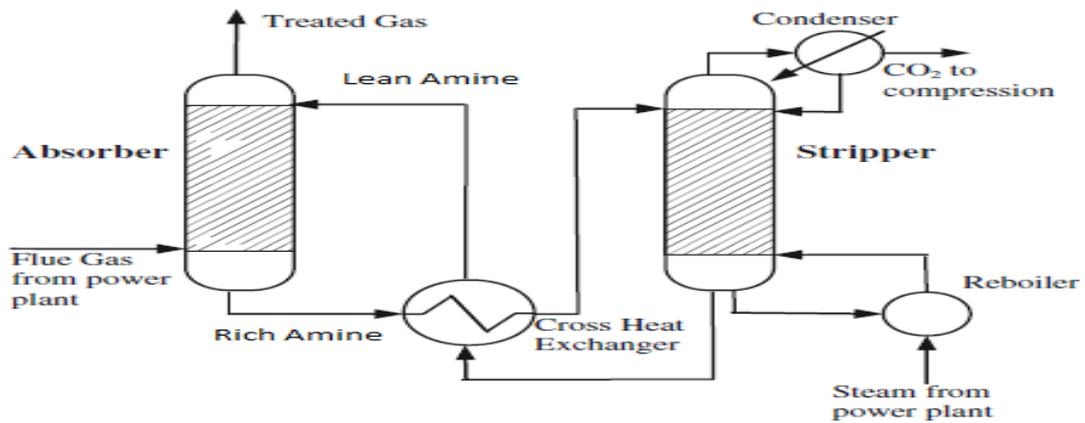
Figure 1. Simplified process flow diagram of chemical absorption process for post-combustion capture [3]

2.3   Overview of previous post-combustion CO_2 capture modelling studies

Post-combustion CO_2 capture with chemical solvent is a reactive absorption, due to two simultaneous phenomena in the process. One is mass transfer of CO_2 from the bulk vapour to the liquid solvent and the other one is chemical reaction between CO_2 and the solvent. As stated in [4], mass transfer rate contributes a lot to reactive absorption design. The relationship between transport and reaction rate will determine where the species can react, such as in the bulk phase, or in the bulk and interfacial regions, or purely in the interfacial layers.

Two-film theory and penetration theory are employed to create rate-based models [1]. In details, for two-film theory, there is an assumption that the liquid and vapour phases are both consisted of two regions: bulk and film. The effects of heat of mass transfer resistances are taken into account only in the laminar film regions. They also explained penetration theory, which assumes the exposure time between every element on surface of liquid and the vapour phase is same. The exposure time affects mass transfer coefficient significantly, because it can imply the effects of hydrodynamic properties of the system.

In the study [5], a steady-state model was developed for the absorber of packed column. This assumes a rate-based mass transfer involving an enhancement factor to estimate the actual absorption rate. The study emphasized that the variation of enhancement along the absorber column is important for model prediction. The evaporation and condensation of water, the

variations in physical properties and heat of chemical reaction all play a vital role to provide a reliable model prediction. On the other hand, most CO_2 absorption took place in the bottom of absorber. In the study [6], a further steady-state model for complete recycling process, including both absorber and stripper, was developed. The rate-based model is presented, which involves different enhancement factors to estimate absorption and desorption rates. In the study [7], a steady-state model was implemented in Aspen Plus based on RADFRAC, to study the effects of chemical reaction and mass transfer on the absorption process.

However, steady-state models are not particularly helpful to understand the impact of post-combustion capture on the operability of the power plant. For instance, what is the response of post-combustion capture plant when the power plant is operating with a varying load? Will modifications (flooding and higher pressure drop) occur during transient conditions, such as start-up and shutdown procedures? Therefore, a dynamic model is necessary to explain and solve all these questions. In the study [3], a dynamic model of absorber was developed with assumptions of equilibrium-based approach in Aspen Plus and rate-based mass transfer in gPROMS. They showed that the rated-based approach gave better prediction than the equilibrium-based approach. Kvamsdal *et al.* (2009) Reference [8] has also developed a dynamic model for an absorption column in gPROMS, with an assumption of rate-based mass transfer. In this study, the enhancement factor is taken into account to count the impact of chemical reaction. In [9], a dynamic model of stripper was created in ACM in Matlab by using rate-based approach. Two operating strategies were carried out in this study: reducing reboiler steam rate with and without adjusting the rich solvent rate. By implementing the ratio of rich solvent rate to steam rate control, the lean loading and temperature remained constant, as well as less response time for the system. The rate-based dynamic model of the amine regeneration unit was also developed in [10], with an enhancement factor to represent the influence of the reactions on the CO_2 mass transfer.

However, these researches so far were only looking at the individual unit (absorber or stripper). Due to the coupling of two columns linked together with a recycle loop in factories, it is inaccurate to analyse the stand-alone columns. Therefore, further researches were concluded to investigate the performance of the complete recycling process by dynamic models. Reference [11] has carried out a study to compare the accuracy of dynamic models for stand-alone columns and integrated columns. The rate-based models assumed all reactions

attained equilibrium, implemented in gPROMS. The absorber and stripper units were linked together with heat exchanger. The results showed that the integrated model predicted the temperature profile better than stand-alone models. In [12], a rate-based model was developed to analyse two dynamic cases, including reducing power plant loading and increasing capture level set point to 95%. They summarized that the CO_2 capture plant had a slower response than power plant. They further explored how capture level affects the power plant loading and difficulties to achieve a steady power plant output quickly.

All these simulation models, relating to chemical-, fluid mechanic- and thermodynamic laws, require extensive knowledge and underlying physics of the process. Even though they can provide advanced features such as customizing component models for the application in hand, there is still a limitation to carry out complicated simulations. For instance, it is difficult to identify which underlying theory and assumption result in the rising uncertainties of the simulation model. In addition, the solution of these simulators is very complex and time consuming. Thus, data-driven "black-boxes" models should be employed as an alternative the mechanism models. In [13], a model of the relationship between critical parameters in post-combustion carbon capture was developed by applying multiple regression. However, it is unable to represent the non-linear relationships among the parameters and the selection of input variables strongly relies on the experts' knowledge. Reference [14] compared three modelling approaches, statistical model, artificial neural network (ANN) model combined with sensitivity analysis (SA) and neuro-fuzzy model. Reference [15] has developed ANN model with sensitivity analysis for a chemical absorption process, by exploring the relationships between inputs and outputs from data set of complete recycling process. However, some previous studies pointed out the disadvantages of single ANN model, such as over-fitting of the training data and poor generalisation performance [16]. The combination of different neural network models would overcome the mentioned shortcomings, thereby increasing the prediction accuracy [17,18]. Bootstrap aggregated neural network [18] is used in this study to model the post-combustion CO_2 capture with chemical absorption.

3   Bootstrap aggregated neural networks

Due to the limitations in training data and training algorithms, it is generally not possible to obtain a perfect neural network model. For example, neural network training might be

trapped in a poor local minimum or the trained network might over fit noise in the training data. Several techniques have been developed to improve neural network generalisation capability, such as regularisation [19], early stopping [20], Bayesian learning [21], training with both dynamic and static process data [22], and combining multiple networks [23-25]. In training with regularisation, the magnitude of network weight is introduced as a penalty term in the neural network training objective function with the purpose of avoiding unnecessarily large network weights which usually leads to poor generalisation. In training with early stopping, neural network performance on the testing data is continuously monitored during the training process and the training process stops when the neural network prediction errors on the testing data start to increase. Among these techniques, combining multiple networks has been shown to be a very promising approach to improving model predictions on unseen data.

Figure 2 shows a bootstrap aggregated neural network model, where several neural network models are developed to model the same relationship. These individual networks are trained on bootstrap replications of the original training data. Instead of selecting a "best" single neural network model, these individual neural networks are combined together to improve model accuracy and robustness. The overall output of the aggregated neural network is a weighted combination of the individual neural network outputs:

$$f(X) = \sum_{i=1}^{n} w_i f_i(X) \tag{1}$$

where $f(X)$ is the aggregated neural network predictor, $f_i(X)$ is the ith neural network, w_i is the aggregating weight for combining the ith neural network, n is the number of neural networks to be combined, and X is a vector of neural network inputs. Since the individual neural networks are highly correlated, appropriate aggregating weights could be obtained through principal component regression [25]. Instead of using constant aggregating weights, the aggregating weights can also dynamically change with the model inputs [26,27]. Another advantage of bootstrap aggregated neural network is that model prediction confidence bounds can be calculated from individual network predictions [18]. The standard error of the ith predicted value is estimated as

$$\sigma_e = \{\frac{1}{n-1}\sum_{b=1}^{n}[y(x_i;W^b) - y(x_i;\cdot)]^2\}^{1/2} \qquad (2)$$

where $y(x_i;\cdot) = \sum_{b=1}^{n} y(x_i;W^b)/n$ and n is the number of neural networks in an aggregated neural network. Assuming that the individual network prediction errors are normally distributed, the 95% prediction confidence bounds can be calculated as $y(x_i;\cdot) \pm 1.96\sigma_e$. A narrower confidence bound, i.e. smaller σ_e, indicates that the associated model prediction is more reliable. Thus, model prediction associated with a narrow prediction confidence bounds is preferred and is considered to be reliable. On the other hand, model prediction with a wide confidence bound is unreliable and should not be trusted.
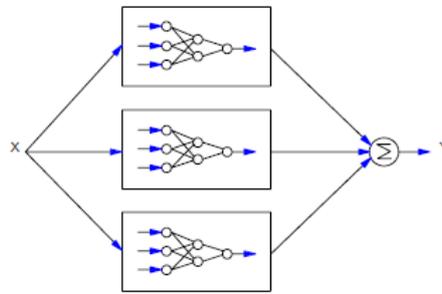
Figure 2. A bootstrap aggregated neural network

4   Modelling of CO$_2$ capture Process

The CO$_2$ capture Process considered here is through chemical absorption.   Detailed mechanistic model for this process was developed in [3] and a simulator based on the mechanistic model was developed in gPROMS at the University of Hull. Simulated static and dynamic process operation data were generated using the simulator.

4.1   Static model

As to static model, only the absorber is modelled. Simulated static process operation data are shown in Figure 3. The process variables that are selected as model input variables are: inlet flue gas flow rate, CO$_2$ concentration in inlet flue gas, pressure of flue gas, temperature of flue gas, lean solvent flow rate, MEA concentration and temperature of lean solvent. They are shown in plots (a) to (g) respectively in Figure 3. CO$_2$ capture level, shown in plot (h) in Figure 3, is taken as the model output variable. Considering that static data is usually not abundant in practice as a process is usually operated in just a few steady states, a small number of data samples are produced as shown in Figure 3.
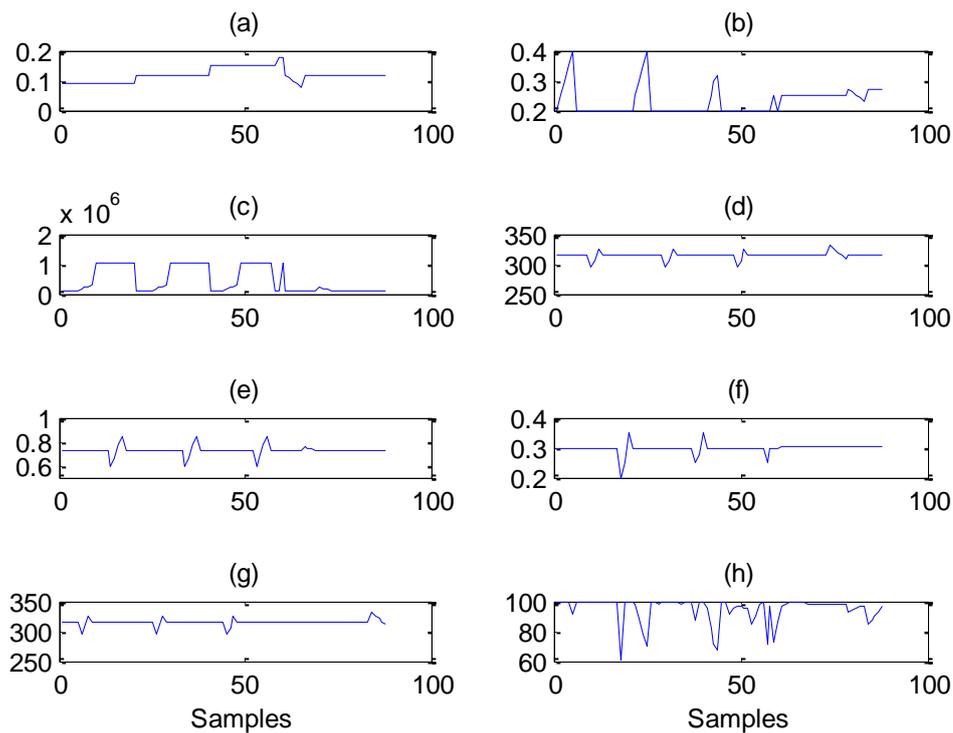
Figure 3. Static process operation data

The generated static data is split into training data (56%), testing data (24%), and unseen validation data (20%). The data is scaled to zero mean and unit variance before they are used for network training. A bootstrap aggregated neural network consists of 30 individual networks is developed. For the development of an individual network, a replication of the training and testing datasets is generated through bootstrap re-sampling with replacement [28] and the network is developed on each bootstrap replication. Each single hidden neural network is a single hidden layer feedforward neural network. The number of hidden neurons in each neural network is determined through cross validation. A number of neural networks with different numbers of hidden neurons are trained on the training data and tested on the testing data. The network with the lowest mean squared errors (MSE) on the testing data is considered to have the appropriate number of hidden neurons. Each network was trained using the Levenberg-Marquardt optimisation algorithm [29] with regularisation and cross-validation based "early-stopping". Figure 4 shows the number of hidden neurons in the individual neural networks. It can be seen that number of hidden neurons vary a lot with different training and testing data sets. The individual networks are then combined through averaging.
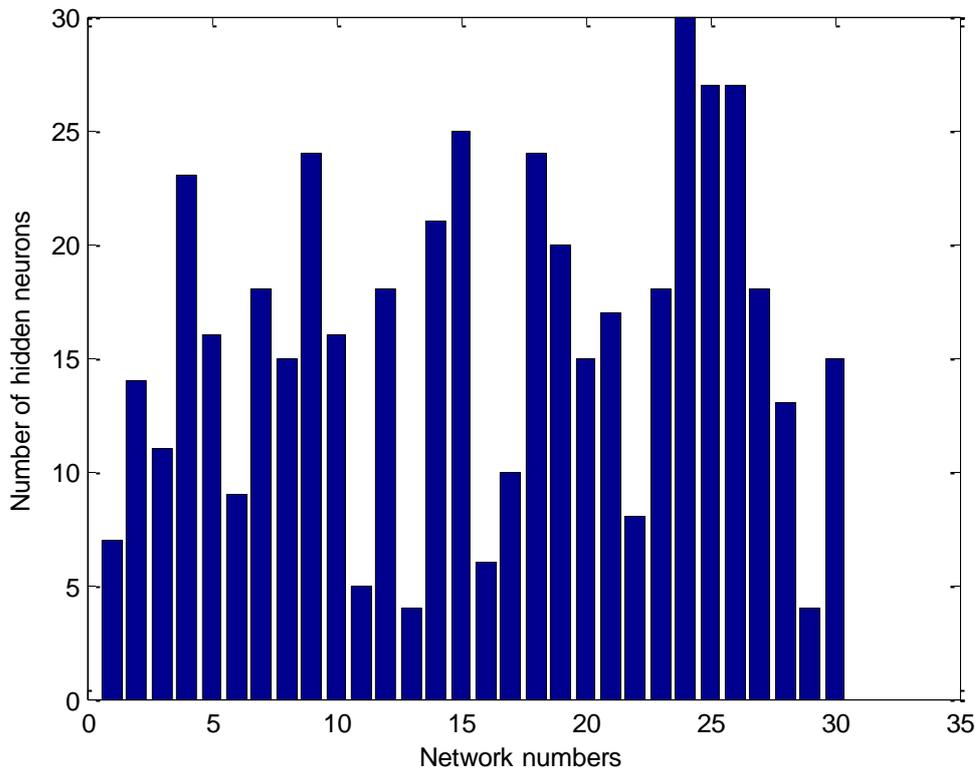
Figure 4. Number of hidden neurons in individual neural networks

Figure 5 shows the mean squared errors (MSE) on training and testing data (top) and on unseen validation data (bottom) from the 30 different single neural networks. Figure 5 shows these from aggregated neural networks with different numbers of constituent networks. It is clearly seen that single neural networks give inconsistent performance on the model building data (training and testing data) and the unseen validation data. For instance, the 14th and 17th networks are among the few best networks in terms of performance on the model building data, but their performance on the unseen validation data is not among the best. The non-robustness of single neural networks is clearly indicated by the difference in performance of individual neural networks on model building data and unseen validation data. Figure 6 clearly indicates that the bootstrap aggregated neural networks give consistent performance on the model building data and on the unseen validation data. In Figure 6, the first bar in each plot represents the first single neural network shown in Figure 5, the second bar represents combining the first two single neural networks, and the last (30th) bar represents combining all the 30 networks. It can be seen from Figure 6 that as more networks are combined, the MSE values on both model building data and unseen validation data decrease and converge to

stable values. Furthermore, bootstrap aggregated neural networks give much more accurate prediction performance than most of the individual networks. This demonstrates that bootstrap aggregated neural networks reliable and accurate prediction performance than single neural networks.
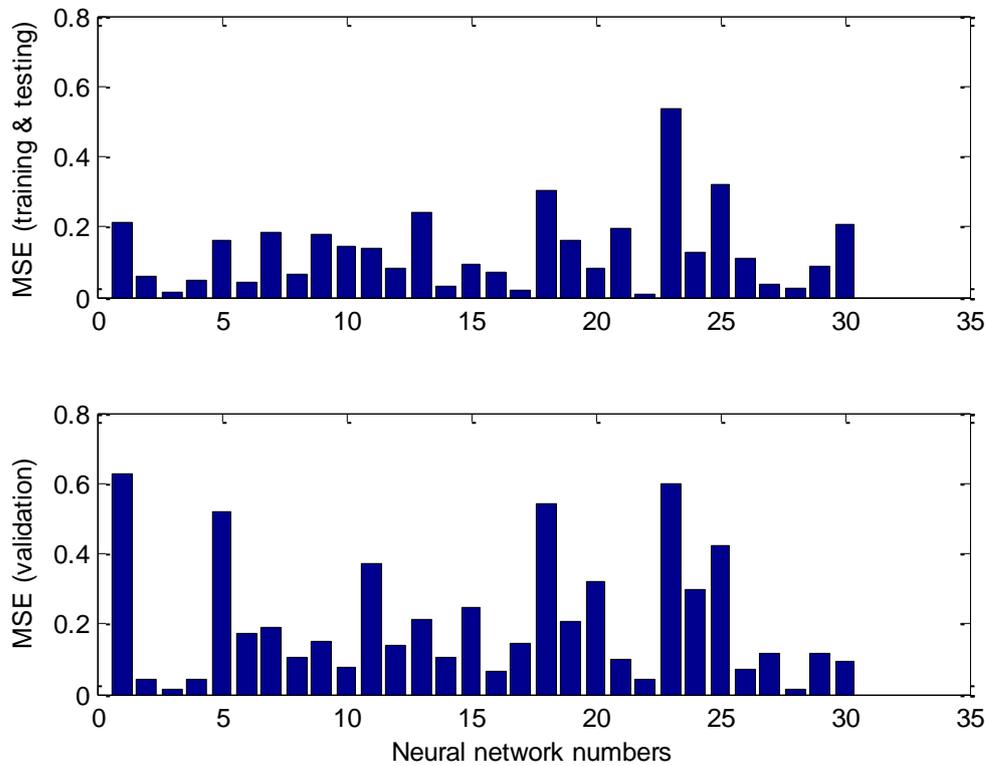

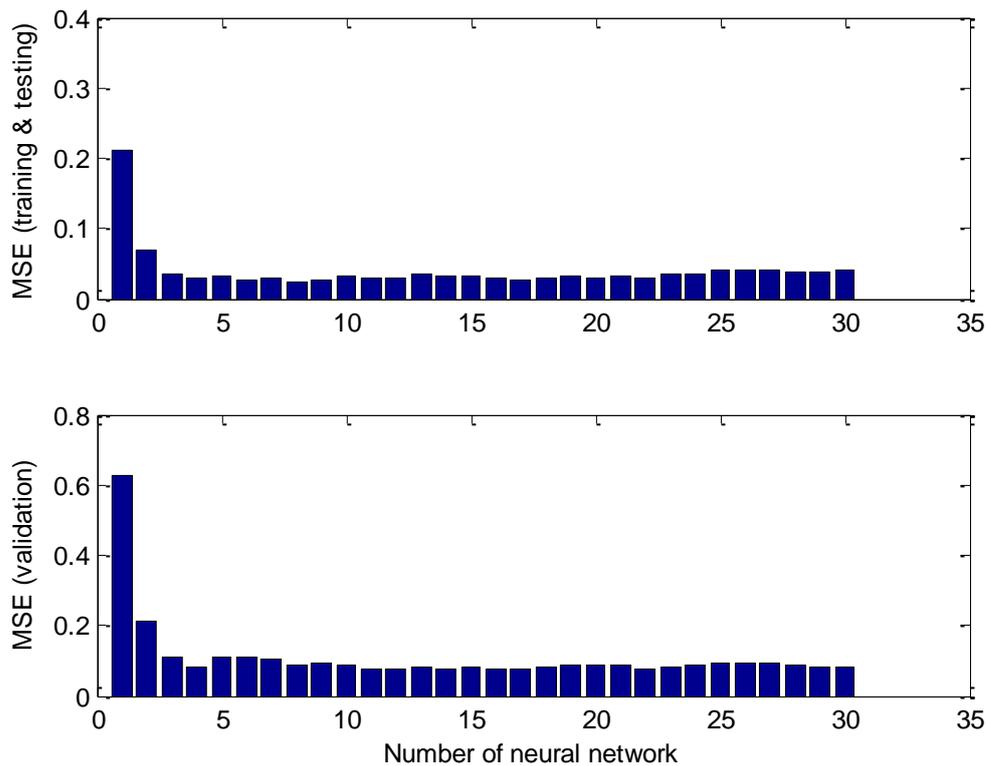Figure 5. MSE of CO_2 capture level for individual neural networks

Figure 6. MSE of CO_2 capture level for aggregated neural networks

Figure 7 shows the actual values, predictions, and 95% confidence bounds of CO_2 capture level on the unseen validation data. Clearly, the predictions by using aggregated neural networks are close to the actual values. The prediction confidence bounds offer extra information to the process operators, such as rejection or acceptation of a particular prediction from the stacked neural network model. The confidence bounds are quite narrow for almost all samples, except for 2nd, 10th, 11th, and 12th samples. Therefore, extra care needs to be taken when using predictions for these samples.
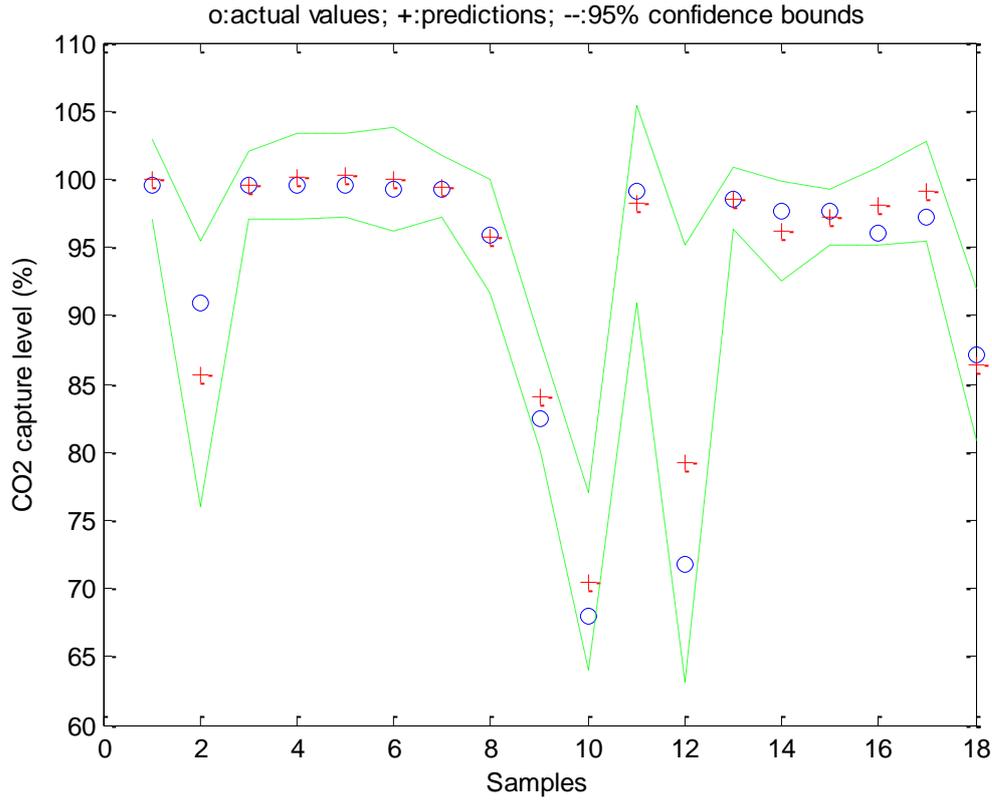
Figure 7. Static model predictions for CO2 capture level on unseen validation data

4.2 Dynamic model

The dynamic simulated process operation data were sampled using a sampling time of 5 s. The generated data were split into training data (56%), testing data (24%), and unseen validation data (20%). The data were scaled to zero mean and unit variance before they were used for neural network training. Two multi-inputs single output (MISO) first order dynamic nonlinear models were developed for CO_2 capture level and CO_2 production rate using bootstrap aggregated neural networks. The developed dynamic model is of the following form:

$$y(t) = f(y(t-1), u_1(t-1), u_2(t-1), \cdots, u_8(t-1)) \tag{3}$$

where y represents CO_2 capture level or CO_2 production rate, u_1 to u_8 are, respectively, inlet gas flow rate, CO_2 concentration in inlet gas flue, inlet gas temperature, inlet gas pressure, MEA circulation rate, lean loading, lean solution temperature, and reboiler temperature.

Each of the nonlinear dynamic models is developed using a bootstrap aggregated neural network consisting of 30 individual neural networks. These individual neural networks are single hidden layer feedforward neural networks. The number of hidden neurons in each network was determined through cross validation. Each network was trained using the Levenberg-Marquardt optimisation algorithm [29] with regularisation and cross-validation based "early-stopping".

Figure 8 shows the MSE values on model building (training and testing) data and unseen validation data from individual neural networks. It can be seen from Figure 8 that the individual networks give various prediction performance. Furthermore, their performance on the training and testing data is not consistent with that on the unseen testing data. For example, network 15 is among the worst performing networks on the training and testing data. However, it offers the best performance on the unseen data. This clearly demonstrates the non-robust nature of single neural networks. Figure 9 shows the MSE values on model building data and unseen validation data from different aggregated neural networks. In Figure 9, the horizontal axes represent the number of individual networks contained in an aggregated neural network. The first bar in Figure 9 represents the first individual neural network shown in Figure 8 and second bar in Figure 9 represents combining the first two individual networks shown in Figure 8. The last bar in Figure 9 represents combining all the 30 neural networks. It can be seen from Figure 9 that bootstrap aggregated neural networks give much more consistent performance on model building data and unseen validation data. The MSE values of aggregated neural networks generally decrease as more networks are combined and converge to a stable level. This occurs in both the model building and unseen data sets. In addition to robustness, Figure 9 also indicates that aggregated neural networks give more accurate performance than individual neural networks. Figure 10 shows the one-step-ahead predictions and multi-step-ahead predictions of CO_2 production rate using aggregated neural networks. It is clearly seen that the predictions are very close to the actual values, except for a few samples where the CO_2 production rates are very high or very low. The slightly larger prediction errors at these samples are likely due to the fact that training data is scare at these extreme operating points. The accurate multi-step-ahead predictions are very encouraging indicating that the model has captured the underlying dynamics of the process. The long range predictions are very accurate till about 90 step-ahead predictions. Such accurate long

range predictions are more than sufficient for model predictive control and real-time optimisation applications.
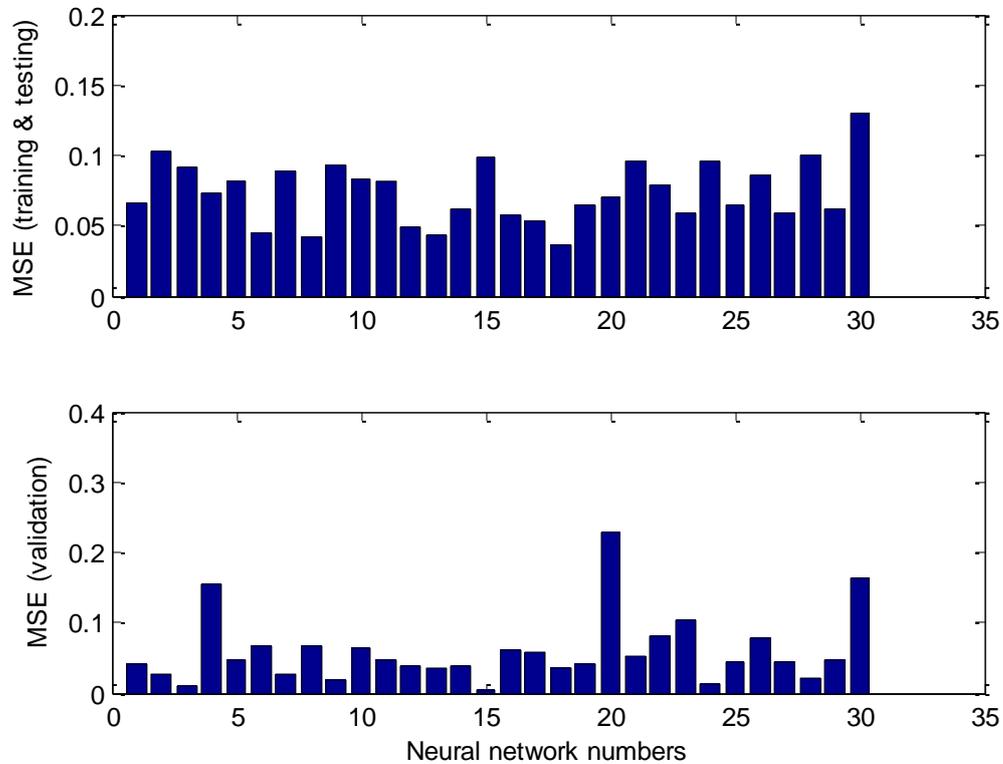

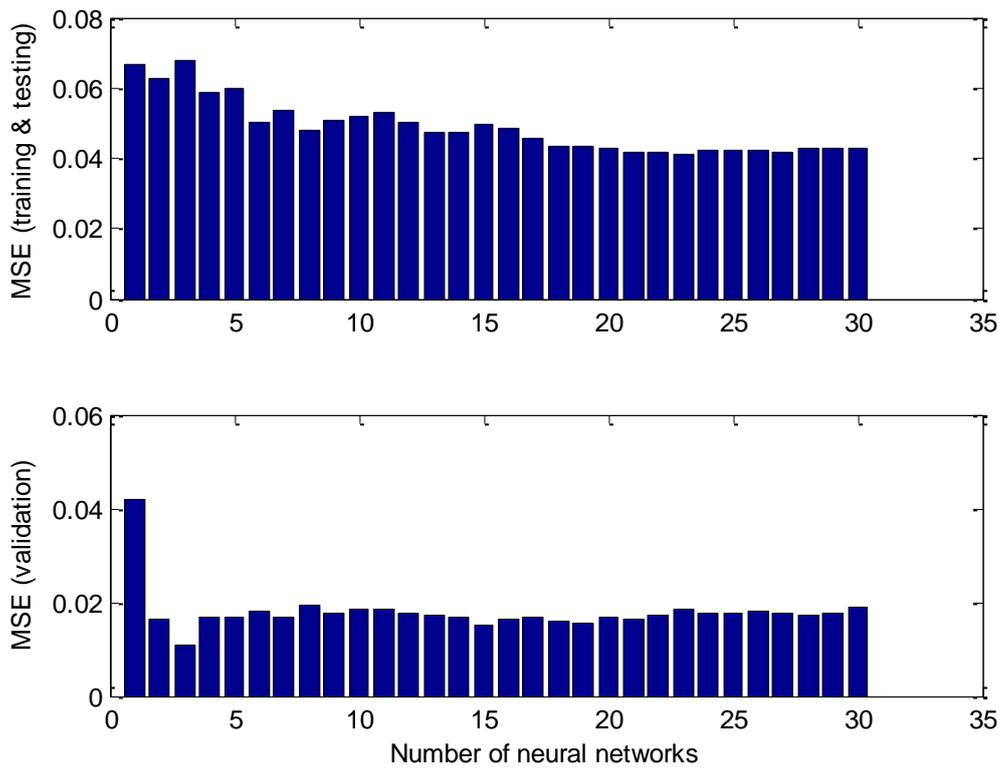Figure 8. MSE of CO_2 production rate for individual neural networks

Figure 9. MSE of CO_2 production rate for aggregated bootstrap neural networks

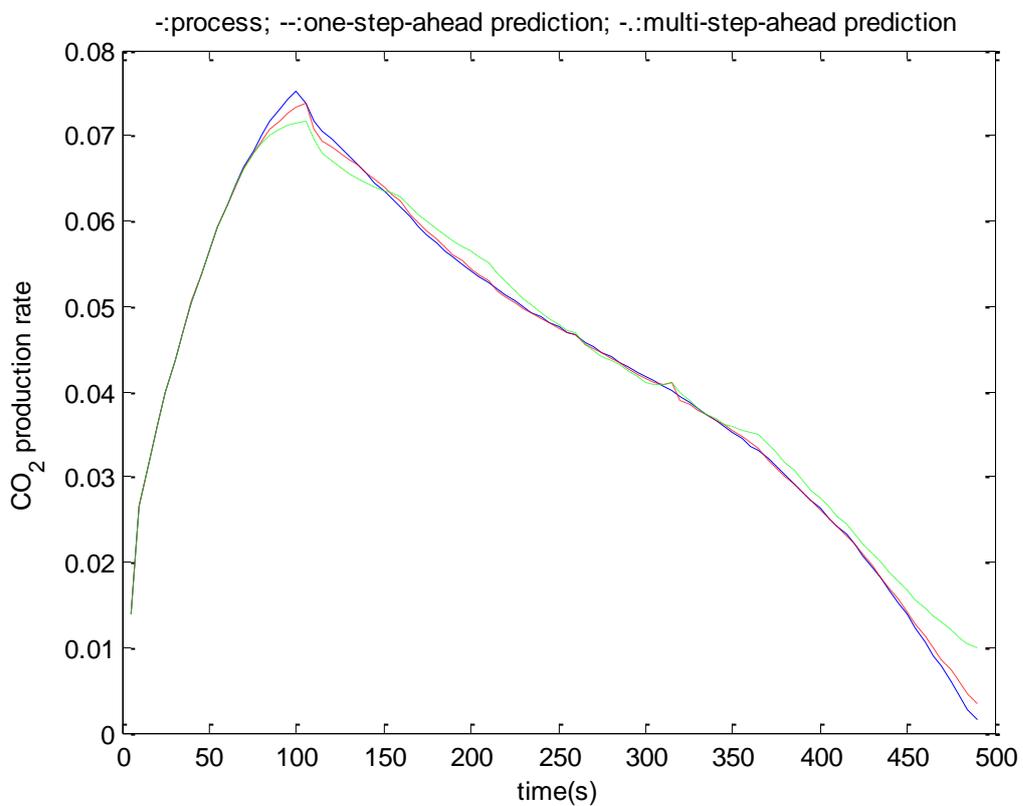-:process; --:one-step-ahead prediction; -.:multi-step-ahead prediction

Figure 10. Dynamic model prediction of CO_2 production rate

The neural network dynamic model for CO_2 capture level is also very accurate as shown in Figure 11. It can be seen from Figure 11 that the long range predictions are accurate until 82-steps-ahead predictions. Again such long prediction horizon is generally adequate for many applications such as model predictive control and real-time optimisations.
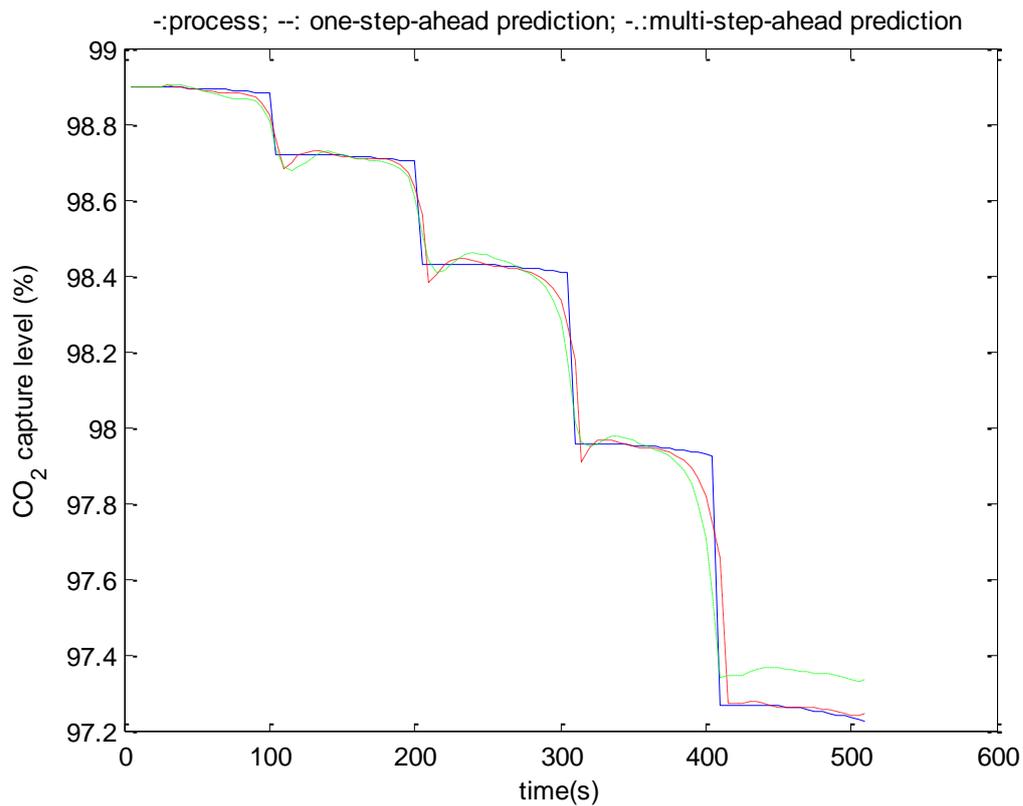
-:process; --: one-step-ahead prediction; -.:multi-step-ahead prediction

Figure 11. Dynamic model prediction of CO_2 capture level

5. Conclusions

The neural network static and dynamic models of CO_2 production rate and CO_2 capture level are developed and they are shown to be able to give accurate predictions. The aggregated neural networks model is found to be the useful tool to predict the post-combustion CO_2 capture process, which is more accuracy and reliable than the traditional neural network models. Bootstrap aggregated neural networks give consistent performance on the model building data and unseen validation data. Furthermore, bootstrap aggregated neural networks can give model prediction confidence bounds, which are a very useful measure on the prediction reliability and can be incorporated in the optimisation framework to give reliable

optimisation results [30]. Reliable optimisation of the CO_2 capture process using the developed neural network models will be studied in the future.

Acknowledgement

References

[1] Wang, M., Lawal, A., Stephenson, P., Sidders, J. and Ramshaw, C. (2011) 'Post-combustion CO2 capture with chemical absorption: A state-of-the-art review', Chemical Engineering Research & Design, 89(9), pp. 1609-1624.

[2] Mores, P., Scenna, N. and Mussati, S. (2012) 'CO2 capture using monoethanolamine (MEA) aqueous solution: Modeling and optimization of the solvent regeneration and CO2 desorption process', Energy, 45(1), pp. 1042-1058.

[3] Lawal, A., Wang, M., Stephenson, P. and Yeung, H. (2009) 'Dynamic modelling of CO2 absorption for post combustion capture in coal-fired power plants', Fuel, 88(12), pp. 2455-2462.

[4] Kenig, E.Y., Schneider, R. and Gorak, A. (2001) 'Reactive absorption: Optimal process design via optimal modelling', Chemical Engineering Science, 56(2), pp. 343-350.

[5] Pintola, T.T., P. Meisen, A. (1993) 'Simulation of pilot plant and industrial CO2-MEA absorbers ', Gas Seperation and Purification, 7(1), pp. 47-52.

[6] Alatiqi, I., Sabri, M.F., Bouhamra, W. and Alper, E. (1994) 'Steady-State Rate-Based Modeling for Co2/Amine Absorption Desorption Systems', Gas Separation & Purification, 8(1), pp. 3-11.

[7] Abu-Zahra, M.R.M., Schneiders, L.H.J., Niederer, J.P.M., Feron, P.H.M. and Versteeg, G.F. (2007) 'CO(2) capture from power plants. Part I. A parametric study of the technical-performance based on monoethanolamine', International Journal of Greenhouse Gas Control, 1(1), pp. 37-46.

[8] Kvamsdal, H.M., Jakobsen, J.P. and Hoff, K.A. (2009) 'Dynamic modeling and simulation of a CO2 absorber column for post-combustion CO2 capture', Chemical Engineering and Processing, 48(1), pp. 135-144.

[9] Ziaii, S., Rochelle, G.T. and Edgar, T.F. (2009) 'Dynamic Modeling to Minimize Energy Use for CO_2 Capture in Power Plants by Aqueous Monoethanolamine', Industrial & Engineering Chemistry Research, 48(13), pp. 6105-6111.

[10] Mores, P., Scenna, N. and Mussati, S. (2012) 'CO_2 capture using monoethanolamine (MEA) aqueous solution: Modeling and optimization of the solvent regeneration and CO_2 desorption process', Energy, 45(1), pp. 1042-1058.

[11] Lawal, A., Wang, M., Stephenson, P., Koumpouras, G. and Yeung, H. (2010) 'Dynamic modelling and analysis of post-combustion CO_2 chemical absorption process for coal-fired power plants', Fuel, 89(10), pp. 2791-2801.

[12] Lawal, A., Wang, M.H., Stephenson, P. and Obi, O. (2012) 'Demonstrating full-scale post-combustion CO_2 capture for coal-fired power plants through dynamic modelling and simulation', Fuel, 101, pp. 115-128.

[13] Zhou, Q., Chan, C.W., Tontiwachiwuthikul, P., Idem, R. and Gelowitz, D. (2009) 'A statistical analysis of the carbon dioxide capture process', International Journal of Greenhouse Gas Control, 3(5), pp. 535-544.

[14] Zhou, Q., Wu, Y.X., Chan, C.W. and Tontiwachwuthikul, P. (2010) 'Applications of Three Data Analysis Techniques for Modeling the Carbon Dioxide Capture Process', 2010 23rd Canadian Conference on Electrical and Computer Engineering (Ccece).

[15] Sipocz, N., Tobiesen, F.A. and Assadi, M. (2011) 'The use of Artificial Neural Network models for CO_2 capture plants', Applied Energy, 88(7), pp. 2368-2376.

[16] Caruana, R., Lawrence, S., &Giles, L. (2000) 'Overrating in neural networks: Backpropagation, conjugate gradient and early stopping', Neural Information Processing System, 13, pp. 402-408.

[17] Clemen, R.T. (1989) 'Combining Forecasts - a Review and Annotated-Bibliography', International Journal of Forecasting, 5(4), pp. 559-583.

[18] Zhang, J. (1999). Developing robust non-linear models through bootstrap aggregated neural networks. Neurocomputing, 25, 93-113.

[19] Bishop, C. (1991). Improving the generalisation properties of radial basis function neural networks, Neural Computation, 13, 579-588.

[20] Bishop, C., Neural Networks for Pattern Recognition. Oxford University Press: Oxford, 1995.

[21] MacKay, D. J. C., 1992. Bayesian interpolation. Neural Computation 4, 415-447.

[22] Zhang, J., 2001. Developing robust neural network models by using both dynamic and static process operating data. Ind. Eng. Chem. Res. 40, 234-241.

[23] Wolpert, D. H., 1992. Stacked generalization. Neural Networks 5, 241-259.

[24] Sridhar, D. V., Seagrave, R. C., Bartlett, E. B., 1996. Process modelling using stacked neural networks. AIChE Journal 42, 2529-2539.

[25] Zhang, J., Morris, A. J., Martin, E. B., Kiparissides, C., 1997. Inferential estimation of polymer quality using stacked neural networks. Computers & Chemical Engineering 21, s1025-s1030.

[26] Ahmad, Z. and Zhang, J. (2005) 'Combination of multiple neural networks using data fusion techniques for enhanced nonlinear process modelling', Computers & Chemical Engineering, 30(2), pp. 295-308.

[27] Ahmad, Z. and J. Zhang, (2006). Combination of multiple neural networks using data fusion techniques for enhanced nonlinear process modelling. Computers & Chemical Engineering, 30, 295-308.

[28] Efron, B. (1982). The Jacknife, the Bootstrap and Other Resampling Plans. Philadelphia: Society for Industrial and Applied Mathematics.

[29] Marquardt, D., 1963. An algorithm for least squares estimation of nonlinear parameters. SIAM J. Appl. Math. 11, 431-441.

[30] Zhang, J. (2004). A reliable neural network model based optimal control strategy for a batch polymerization reactor. *Ind. Eng. Chem. Res*., **43**(4), 1030-1038.