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 capture 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 feedforward 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. Simulated CO_2 capture process operation data from gPROMS simulation are used to build and verify neural network models. Both neural network static and dynamic models are developed and they offer accurate predictions on unseen validation data. 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.

1 Introduction

Post-combustion CO_2 capture for fossil fuel-fired power plants is attracting more attention as a result of the large amounts of existing fossil fuel-fired power plants and no 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 first principle models and statistical models, have been studied to investigate the post-combustion

* Corresponding author

carbon capture process efficiency. Previous studies showed that the establishment of first principle models is very time consuming and requires extensive knowledge of the underlying physics of the process. Numerical optimisation typically requires thousands of function evaluations. Evaluation of a detailed first principle 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 [2]. Neural network models can be developed very quickly from process data and their evaluation is much less computationally demanding compared with a first principle 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 [3-6]. This paper uses bootstrap aggregated neural networks to build data-driven models for a CO_2 capture chemical absorption process with solvents.

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 and results and discussions are detailed in Section 4. Both steady state and dynamic models are developed. Section 5 draws some concluding remarks.

2 An overview of CO_2 capture processes

2.1 Post-combustion CO_2 capture via chemical absorption

Post-combustion CO_2 capture process removes CO_2 emission after the combustion of fossil fuel in a combustor. It can be retrofitted to existing fossil fuel-fired power plants for CO_2 capture. Several separation technologies can be employed in this process and they include adsorption, physical absorption, chemical absorption, cryogenics separation and membrane absorption [1]. Among these methods, chemical absorption is found to be most suitable for CO_2 capture from fossil fuel-fired power plants industrial flue gas due to the high selectivity and final pure CO_2 stream [1].

As shown in Figure 1, CO_2 capture plant with chemical absorption mainly consists of two packed columns, namely absorber and stripper columns. The flue gas from power plant is fed into the bottom of absorber and contacted counter-currently with lean amine solvent from the top side. The lean amine solvent chemically reacts and absorbs CO_2 in the flue gas. Then the

treated gas stream with lower CO_2 concentration leaves from the top of absorber. The amine solution with more CO_2 (now rich amine), coming from the bottom of absorber, is pumped to the stripper after preheating in cross heat exchanger. In the stripper, the rich amine solution is regenerated by heating via a reboiler [7]. The reboiler energy is often provided by low-pressure steam from the steam cycle of the power plant causes large energy consumption. The vapour that results from heating the rich amine is richer in CO_2 than the left over liquid phase (lean amine). The vapour is cooled in a condenser followed by flash separation. The vapour from the flash separator, mostly CO_2 (up to 99%) is compressed and transported to storage sites while the liquid phase from the flash separator is refluxed back to the stripper. Finally, the lean amine is cooled in cross heat exchanger by exchanging heat with rich amine and then pumped back to the absorber.

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 for carrying out the optimisation tasks.

2.2 Review of previous studies in modelling post-combustion CO_2 capture

Post-combustion CO_2 capture with chemical solvent is a reactive absorption involving simultaneous phenomena. One is mass transfer of CO_2 from the bulk vapour to the liquid solvent and the other is chemical reaction between CO_2 and the solvent. As stated in [8], mass transfer rate contributes a lot to reactive absorption design. The relationship between mass 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.

Mass transfer rates between the vapour and liquid phase can be described using different theories, namely two-film theory and penetration theory among others [1]. Two-film theory is however more commonly used due to its simplicity and ease of application. In two-film theory, the liquid and vapour phases are both assumed to consist of two regions; bulk and film region. The effects of heat and mass transfer resistances are taken into account only in the laminar film regions. In penetration theory on the other hand, it is assumed that the

exposure time between every element on surface of the liquid and vapour phase is the same. The exposure time affects mass transfer coefficient significantly because it can imply the effects of hydrodynamic properties of the system.

In Pintola and Meisen [9], a steady-state model was developed for the absorber. This assumes a rate-based mass transfer with 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 reliable model prediction. On the other hand, most CO_2 absorption took place in the bottom of the absorber. In the work of Alatiqi *et al*. [10], a further steady-state model involving integrated absorber and stripper columns was developed. The model is similarly rate-based with different enhancement factors to estimate absorption and desorption rates. In the work of Abu-Zahra *et al*. [11], a steady-state model was implemented in Aspen Plus® using RADFRAC columns, 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 the post-combustion capture plant 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 flooding occur during transient conditions, such as start-up and shutdown procedures? Therefore, a dynamic model is necessary to answer these questions. In Lawal *et al*. [12], a dynamic model of absorber was developed using equilibrium-based approach in Aspen Plus® and rate-based mass transfer in gPROMS®. They showed that the rated-based approach gives better prediction than the equilibrium-based approach. Kvamsdal *et al*. [13] also developed a dynamic model for an absorption column in gPROMS using rate-based mass transfer approach. In this study, an enhancement factor is used to account for the impact of chemical reactions. In Ziaii *et al*. [14], a dynamic model of stripper was created in Aspen Custom Modeller (ACM) 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 remains constant, as well as less response time for the system. The rate-based dynamic model of the amine regeneration unit was also developed in Mores *et al*.

[7], with an enhancement factor to represent the influence of the reactions on the CO_2 mass transfer.

However, these studies only looked at the individual unit (either absorber or stripper). Due to the coupling effect between two columns linked together with a recycle loop, analysis of the stand-alone columns is insufficient to understand the dynamics of the complete post-combustion CO_2 capture process. Therefore, further studies are necessary to investigate the performance of the complete recycling process through dynamic models. Lawal *et al*. [15] carried out a study to compare the accuracy of dynamic models for stand-alone columns and integrated columns using gPROMS®. The rate-based models assumed all reactions attained equilibrium. 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 Lawal *et al*. [16], 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 of the 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-box" models should be employed as an alternative to first principle models. In Zhou *et al*. [17], 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. Zhou *et al*. [18] compared three modelling approaches: statistical model, artificial neural network (ANN) model combined with sensitivity analysis (SA), and neuro-fuzzy model. Sipocz *et al*. [19] has developed ANN model with sensitivity analysis for a chemical absorption process, by exploring the relationships between inputs and outputs from data set of integrated post-combustion CO_2 capture 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 [20]. The combination of different neural network models would overcome the mentioned shortcomings, thereby increasing the prediction accuracy [21, 22]. Bootstrap aggregated neural network [22] 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 [23], early stopping [24], Bayesian learning [25], training with both dynamic and static process data [26], and combining multiple networks [27-29]. In training with regularisation, the magnitude of network weights 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 is terminated 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 [22], where several neural network models are developed to model the same relationship and are then combined. These individual networks can be multilayer feedforward neural networks, radial basis function networks, or recurrent neural networks. In this study, the individual networks are single hidden layer feedforward networks. In order for the combined network to give accurate predictions, the individual networks should be different. These individual networks are trained on bootstrap replications of the original training data [30]. These individual networks can have different number of hidden neurons and their training start from different sets of initial weights. Instead of selecting a "best" single neural network model among these individual networks, these individual neural networks are combined together to improve

model accuracy and robustness. The terminology "bootstrap aggregation" is originated from "bagging predictors" introduced by Breiman in [31]. 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 [29]. Instead of using constant aggregating weights, the aggregating weights can also dynamically change with the model inputs [32, 33].

The reason that combing multiple models can improve model prediction accuracy is illustrated by Perrone and Cooper [34]. Suppose that N independent predictors are combined, then the combined predictor can reduce the mean squared prediction errors by a factor of N. Let the error of the ith model $f_i(x)$ be $\varepsilon_i = f(x) - f_i(x)$, where $f(x)$ is the true function to be approximated, the mean square error can be written as:

$$MSE[f_i] = E[\varepsilon_i^2] \tag{2}$$

The average mean square error is therefore

$$\overline{MSE} = \frac{1}{N}\sum_{i=1}^{N} E[\varepsilon_i^2] \tag{3}$$

If the N models are combined through simple averaging, then

$$f_{AV}(x) = \frac{1}{N}\sum_{i=1}^{N} f_i(x) = f(x) - \frac{1}{N}\sum_{i=1}^{N} \varepsilon_i(x) \tag{4}$$

If we assume that the $\varepsilon_i(x)$ are mutually independent with zero mean, then the mean square error for $f_{AV}(x)$ can be calculated as:

$$\begin{aligned} MSE[f_{AV}] &= E[(\frac{1}{N}\sum_{i=1}^{N} \varepsilon_i)^2] \\ &= \frac{1}{N^2} E[\sum_{i=1}^{N} \varepsilon_i^2] + \frac{1}{N^2} E[\sum_{i \neq j} \varepsilon_i \varepsilon_j] \\ &= \frac{1}{N} \overline{MSE} + \frac{1}{N^2} \sum_{i \neq j} E[\varepsilon_i] E[\varepsilon_j] \end{aligned}$$

$$= \frac{1}{N}\overline{MSE} \quad (5)$$

It can be seen from Eq(5) that combining N independent model can reduce the mean square error by N times. An implication of this result is that significant improvement in model prediction can be obtained if dissimilar models are combined. Zhang et al. [29] propose using principal component regression to combine neural networks where independent contributions from individual network predictions are combined.

Another advantage of bootstrap aggregated neural network is that model prediction confidence bounds can be calculated from individual network predictions [22]. The standard error of the ith predicted value is estimated as

$$\sigma_e = \{\frac{1}{n-1}\sum_{b=1}^{n}[y(x_i;W^b) - y(x_i;\cdot)]^2\}^{1/2} \quad (6)$$

where $y(x_i;\ .) = \sum_{b=1}^{n} y(x_i;W^b)/n$ 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 $y(x_i;\ .) \pm 1.96\sigma_e$. 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.

4 Modelling of CO_2 capture Process

The CO_2 capture process considered here is through chemical absorption. A detailed first principle gPROMS dynamic model for absorber was developed in [12] and a gPROMS dynamic model for the whole process was developed in [15]. Simulators for the absorber and the whole process based on the first principle models were developed in gPROMS®. Simulated steady state and dynamic process operation data were generated using the simulators.

4.1 Steady state model for absorber

For the steady state model, only the absorber is modelled. Simulated steady state absorber operation data using first principle model developed in [12] are shown in Figure 3. The process variables 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 steady state 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.

The generated steady state data was 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 network training. A bootstrap aggregated neural network consisting of 30 individual single hidden layer feedforward networks was developed. For the development of an individual network, a replication of the training and testing datasets was generated through bootstrap re-sampling with replacement [30] and the network was developed on the bootstrap replication. The number of hidden neurons in each neural network was determined through cross validation. A number of neural networks with different numbers of hidden neurons (between 3 and 30) were trained on the training data and tested on the testing data. The network with the lowest mean squared errors (MSE) on the testing data was considered to have the appropriate number of hidden neurons. Each network was trained using the Levenberg-Marquardt optimisation algorithm [35] 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 the number of hidden neurons vary a lot with different training and testing data sets. This indicate that the "best" neural network structure depends on the model building data and slight variation in the model building data can lead to different neural network structure. The individual networks are then combined through averaging in this study.

Figure 5 shows the MSE values on training and testing data (top) and on unseen validation data (bottom) from the 30 different individual neural networks. It is clearly seen that single neural networks give inconsistent performance on the model building data (training and

297 testing data) and on the unseen validation data. For instance, the 14^{th} and 17^{th} networks are
298 among the few best networks in terms of performance on the model building data, but their
299 performance on the unseen validation data is not among the best. The non-robustness of
300 single neural networks is clearly indicated by the difference in performance of individual
301 neural networks on model building data and unseen validation data.
302
303 Figure 6 shows the MSE values on training and testing data (top) and on unseen validation
304 data (bottom) from aggregated neural networks with different numbers of constituent
305 networks. In Figure 6, the first bar in each plot represents the first single neural network
306 shown in Figure 5, the second bar represents combining the first two single neural networks,
307 and the last (30^{th}) bar represents combining all the 30 networks. Figure 6 clearly indicates
308 that the bootstrap aggregated neural networks give consistent performance on the model
309 building data and on the unseen validation data. It can be seen from Figure 6 that as more
310 networks are combined, the MSE values on both model building data and unseen validation
311 data decrease and converge to stable values. Furthermore, bootstrap aggregated neural
312 networks give much more accurate prediction performance than most of the individual
313 networks. This demonstrates that bootstrap aggregated neural networks can give reliable and
314 accurate prediction performance than single neural networks.
315
316 Figure 7 shows the actual values, neural network model predictions, and 95% confidence
317 bounds of CO_2 capture level on the unseen validation data. Clearly, the predictions by using
318 aggregated neural networks are close to the actual values. The prediction confidence bounds
319 offer extra information to the process operators on the prediction reliability. A prediction with
320 narrow prediction confidence bounds is considered to be reliable while, on the other hand, a
321 prediction with wide prediction confidence bounds is considered to be unreliable. Figure 7
322 shows that the model prediction confidence bounds are quite narrow for almost all samples,
323 except for 2^{nd}, 10^{th}, 11^{th}, and 12^{th} samples. Therefore, extra care needs to be taken when using
324 predictions for these samples.
325

326 ## 4.2 Dynamic model for the whole process

327
328 Dynamic simulation of the whole process was carried out using the first principle model
329 developed in [15] and the simulated process operation data were sampled using a sampling

time of 5s. 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 capture rate using bootstrap aggregated neural networks. The developed dynamic model is of the following form:

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

where y represents CO_2 capture level or CO_2 capture rate, u_1 to u_8 are, respectively, inlet gas flow rate, CO_2 concentration in inlet flue gas, 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 [35] with regularisation and cross-validation based "early-stopping".

Figure 8 shows the MSE values on model building (training and testing) data (top) and unseen validation data (bottom) 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 validation 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 (top) and unseen validation data (bottom) from different aggregated neural networks with different number of constituent 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 stable levels. This occurs in both the model building and unseen validation data sets. In addition to robustness, Figure 9 also indicates that aggregated neural networks give more accurate prediction performance than individual neural networks.

Figure 10 shows the one-step-ahead predictions and multi-step-ahead predictions of CO_2 capture rate using aggregated neural networks. In Figure 10, the solid curve is the actual CO_2 capture rate, the dashed curve is the one-step-ahead prediction, and the dash dotted curve is multi-step-ahead prediction. In calculating one-step-ahead predictions, the measured process output (CO_2 capture level or CO_2 capture rate) at sampling time t is used as neural network model input to predict the process output at sampling time t+1. In calculating multi-step-ahead predictions, the neural network model predicted process output at sampling time t is used as neural network model input to predict the process output at sampling time t+1. Accurate multi-step-ahead predictions are hence much more difficult to achieve than accurate one-step-ahead predictions because prediction errors at the previous sampling times are introduced in the model inputs and this could further increase the model prediction errors at future sampling times. It can be clearly seen from Figure 10 that the predictions are very close to the actual values, except for a few samples where the CO_2 capture 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 scarce 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 usually more than sufficient for model predictive control and real-time optimisation applications.

The neural network dynamic model for CO_2 capture level is also very accurate as shown in Figure 11. In Figure 11, the solid curve is the actual CO_2 capture level, the dashed curve is the one-step-ahead prediction, and the dash dotted curve is multi-step-ahead prediction. 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, which will be investigated in future studies.

5. Conclusions

The neural network steady state and dynamic models of CO_2 capture 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 a useful tool to model the post-combustion CO_2 capture process and is more accurate 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 very useful measures on the prediction reliability and can be incorporated in the optimisation framework to give reliable optimisation results [36]. Reliable optimisation of the CO_2 capture process using the developed neural network models will be studied in the future.

Acknowledgement

The work was financially supported by the EU FP7 Programme through the project "Research and Development in Coal-fired Supercritical Power Plant with Post-combustion Carbon Capture using Process Systems Engineering techniques" (Ref: PIRSES-GA-2013-612230).

References

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

[2] Bulsari, A. B., (Ed), (1995). Computer-Aided Chemical Engineering, Volume 6: Neural Networks for Chemical Engineers. Amsterdam: Elsevier.

[3] Zhang, Z. Y., Wang, T., Liu, X. G. Melt index prediction by aggregated RBF neural networks trained with chaotic theory. Neurocomputing 2014; 131: 368-376

[4] Li, J. B., Liu, X. G., Jiang, H. Q., Xiao, Y. D. Melt index prediction by adaptively aggregated RBF neural networks trained with novel ACO algorithm. J Appl Polym Sci. 2012;125: 943-951

[5] Liu, X. G., Zhao, C. Y. Melt index prediction based on fuzzy neural networks and PSO algorithm with online correction strategy. AIChE J 2012;58: 1194-1202

[6] Zhang, J. and Morris, A. J. Recurrent neuro-fuzzy networks for nonlinear process modelling. IEEE Transactions on Neural Networks, 1999; 10: 313-326.

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

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

[9] Pintola, T.T., P. Meisen, A. Simulation of pilot plant and industrial CO2-MEA absorbers. Gas Separation and Purification 1993; 7(1): 47-52.

[10] Alatiqi, I., Sabri, M.F., Bouhamra, W. and Alper, E. Steady-State Rate-Based Modeling for CO2/Amine Absorption Desorption Systems. Gas Separation & Purification 1994, 8(1):3-11.

[11] Abu-Zahra, M.R.M., Schneiders, L.H.J., Niederer, J.P.M., Feron, P.H.M. and Versteeg, G.F. CO2 capture from power plants. Part I. A parametric study of the technical-performance based on monoethanolamine. International Journal of Greenhouse Gas Control 2007; 1(1):37-46.

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

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

[14] Ziaii, S., Rochelle, G.T. and Edgar, T.F. Dynamic Modeling to Minimize Energy Use for CO2 Capture in Power Plants by Aqueous Monoethanolamine. Industrial & Engineering Chemistry Research 2009; 48(13):6105-6111.

[15] Lawal, A., Wang, M., Stephenson, P., Koumpouras, G. and Yeung, H. Dynamic modelling and analysis of post-combustion CO2 chemical absorption process for coal-fired power plants. Fuel 2010, 89(10):2791-2801.

[16] Lawal, A., Wang, M.H., Stephenson, P. and Obi, O. Demonstrating full-scale post-combustion CO2 capture for coal-fired power plants through dynamic modelling and simulation. Fuel 2012; 101:115-128.

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

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

[19] Sipocz, N., Tobiesen, F.A. and Assadi, M. The use of Artificial Neural Network models for CO2 capture plants. Applied Energy 2011; 88(7):2368-2376.

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

[21] Clemen, R.T. Combining Forecasts - a Review and Annotated-Bibliography. International Journal of Forecasting 1989, 5(4):559-583.

[22] Zhang, J. Developing robust non-linear models through bootstrap aggregated neural networks. Neurocomputing 1999; 25:93-113.

[23] Bishop, C. Improving the generalisation properties of radial basis function neural networks. Neural Computation 1991; 13:579-588.

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

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

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

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

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

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

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

[31] Breiman, L. Bagging predictors. Machine Learning, 1996;24: 123–140.

[32] Ahmad, Z., Zhang, J. Bayesian selective combination of multiple neural networks for improving long range predictions in nonlinear process modelling. Neural Computing & Applications 2005; 14, 78-87.

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

[34] Perrone, M. P. and Cooper, L. N. (1993). When networks disagree: ensemble methods for hybrid neural networks. In R. J. Mammone (Ed), Artificial Neural Networks for Speech and Vision. London: Chapman and Hall, 1993; 126-142.

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

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

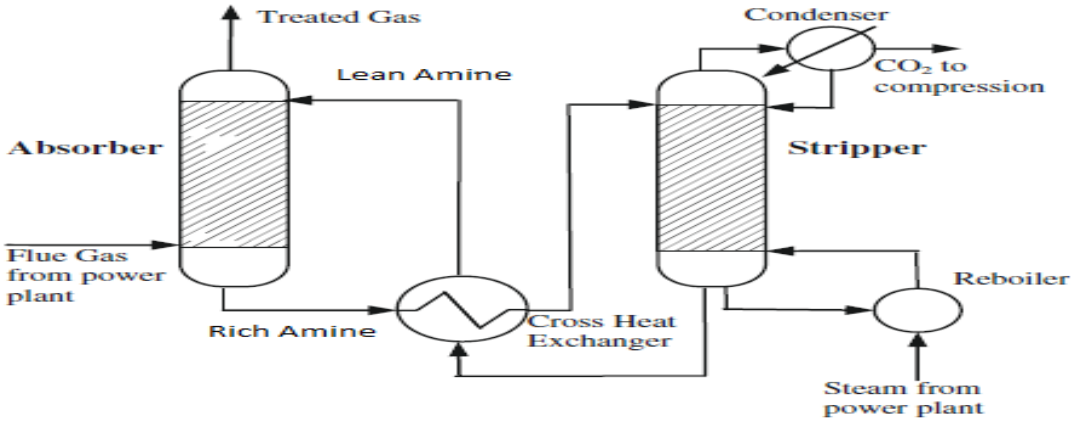


Figure 1. Simplified process flow diagram of chemical absorption process for post-combustion capture [8]

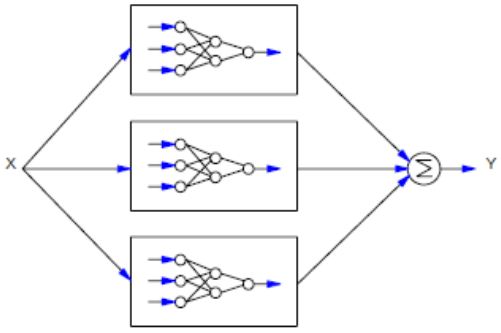
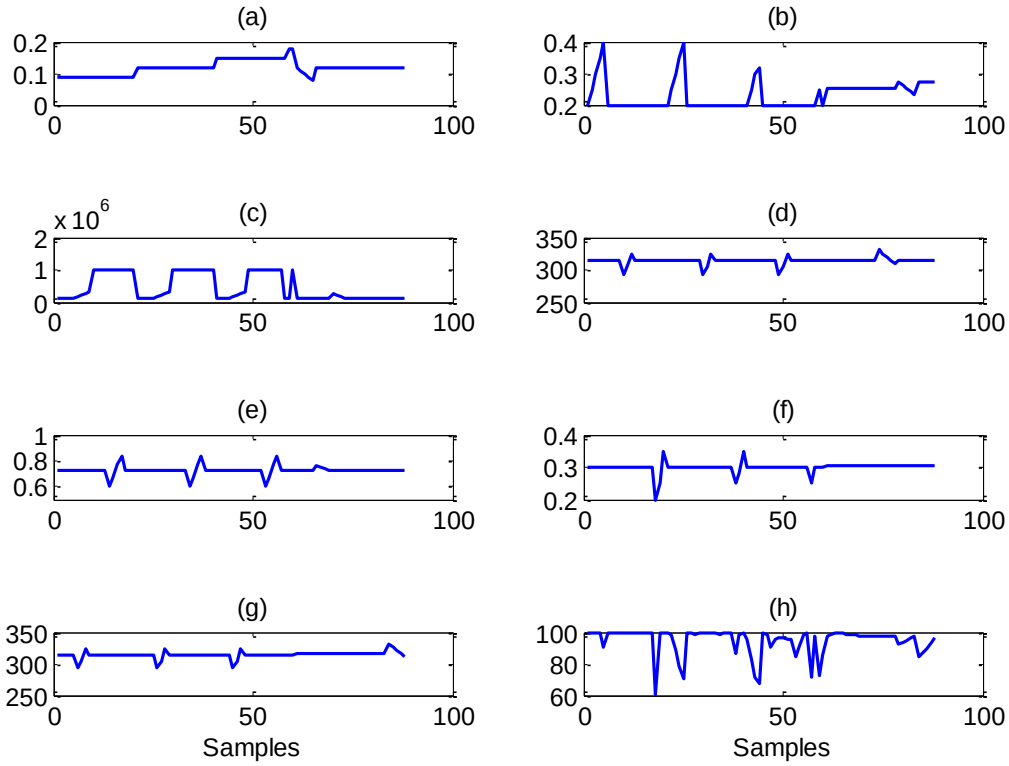


Figure 2. A bootstrap aggregated neural network



527

528

529

530 Figure 3. Steady state absorber operation data: (a). inlet flue gas flow rate (kg/s); (b). CO_2
531 concentration in inlet flue gas (mass fraction); (c). pressure of flue gas (Pa); (d). temperature
532 of flue gas (K); (e). lean solvent flow rate (kg/s); (f). MEA concentration (mass fraction); (g).
533 temperature of lean solvent (K); (h). CO_2 capture level (%)

534

535

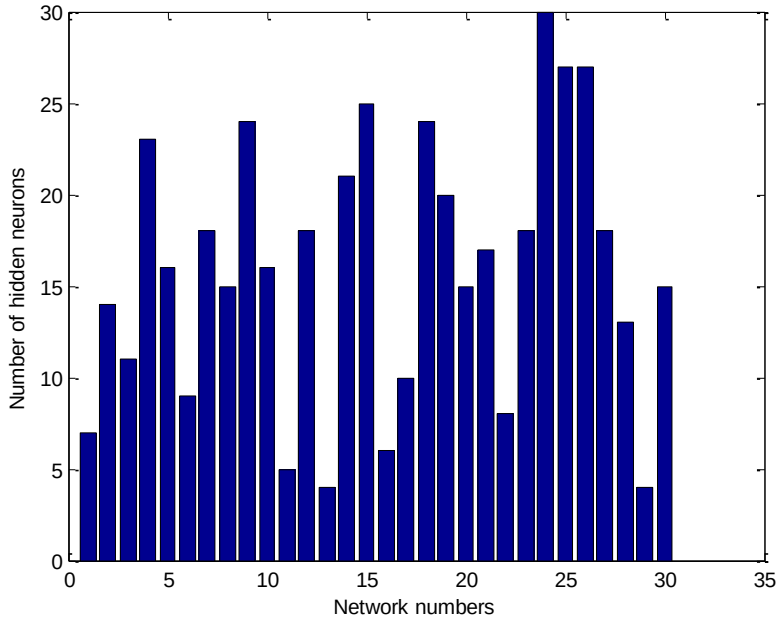


Figure 4. Number of hidden neurons in individual neural networks

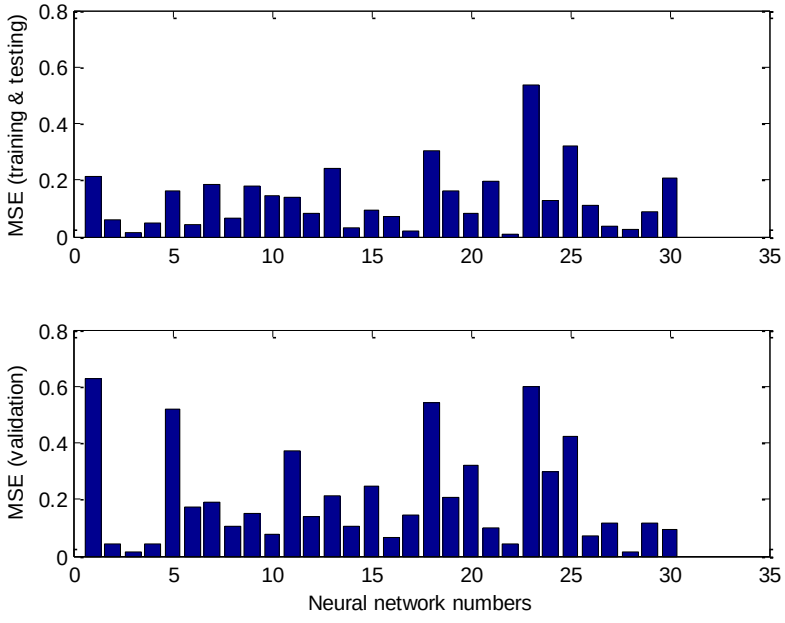


Figure 5. MSE of CO_2 capture level for individual neural networks

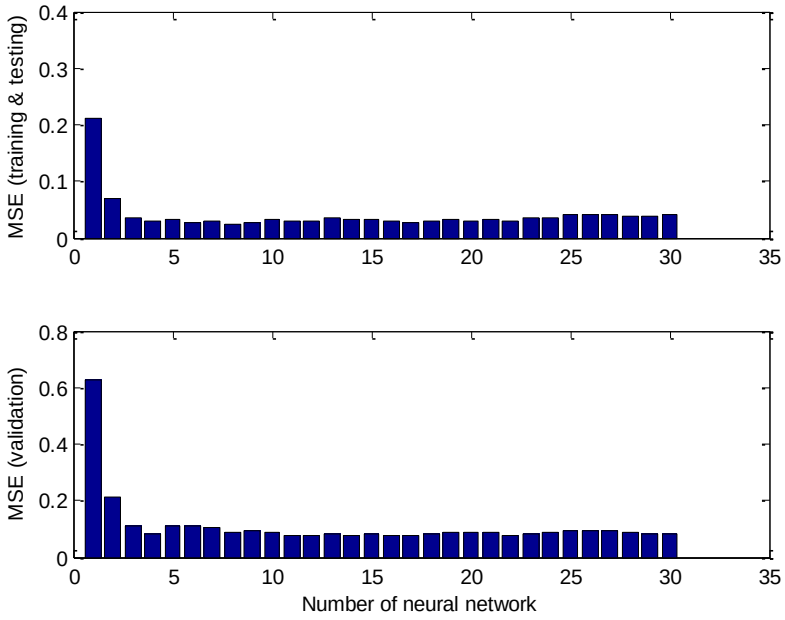


Figure 6. MSE of CO_2 capture level for aggregated neural networks

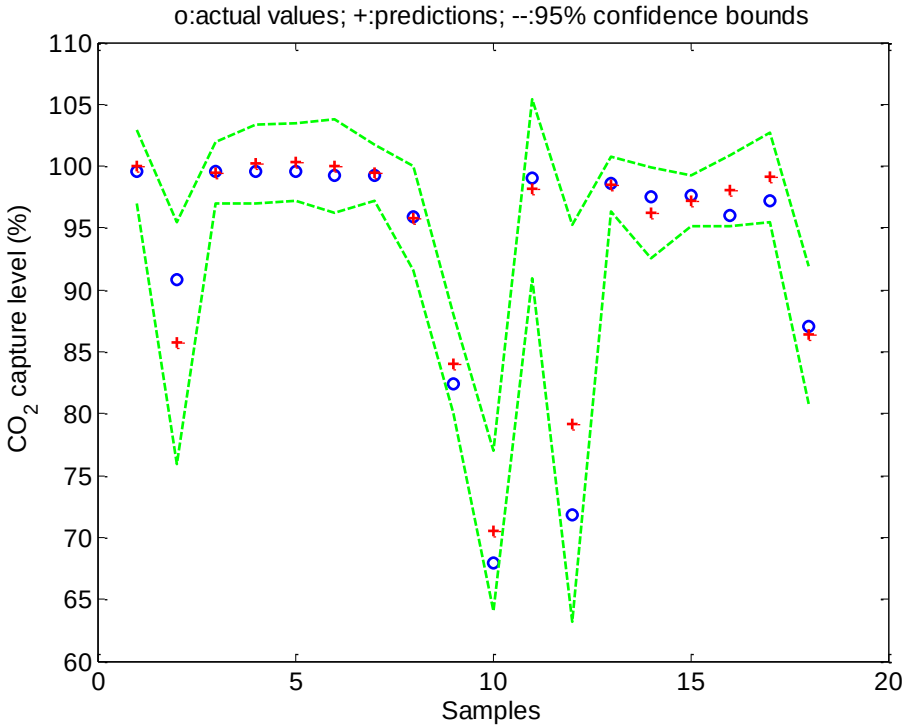
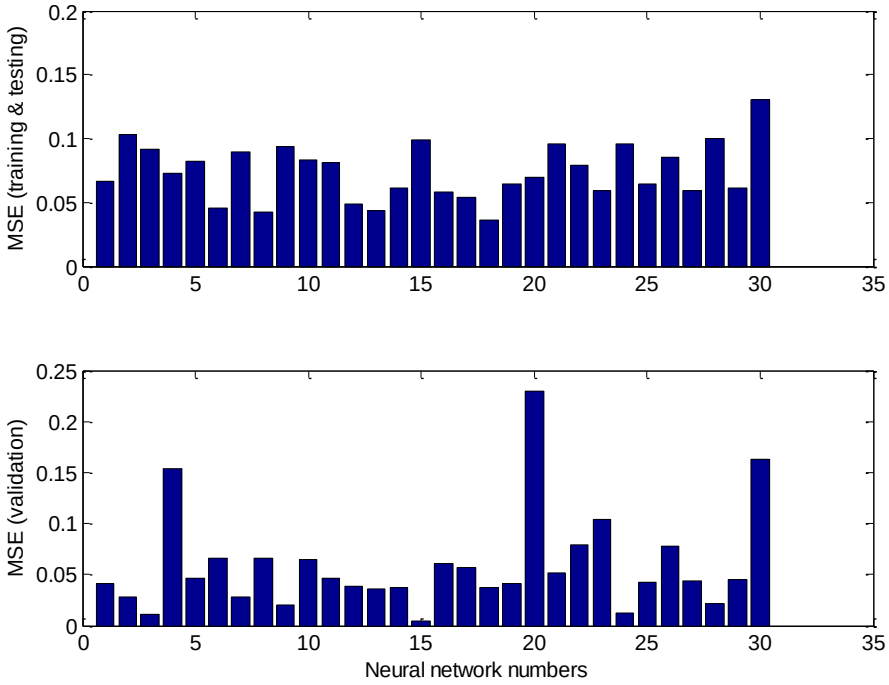
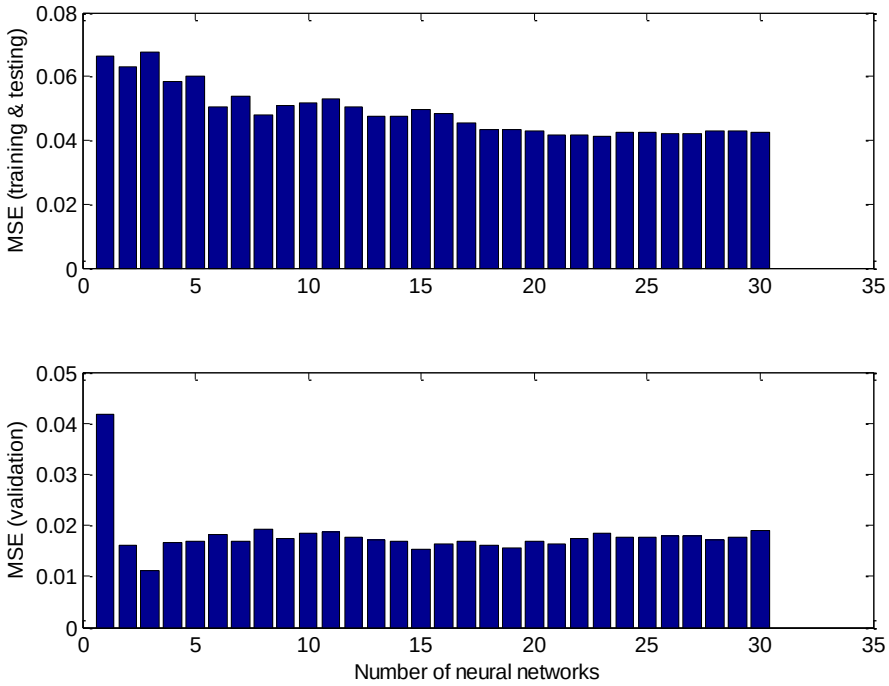


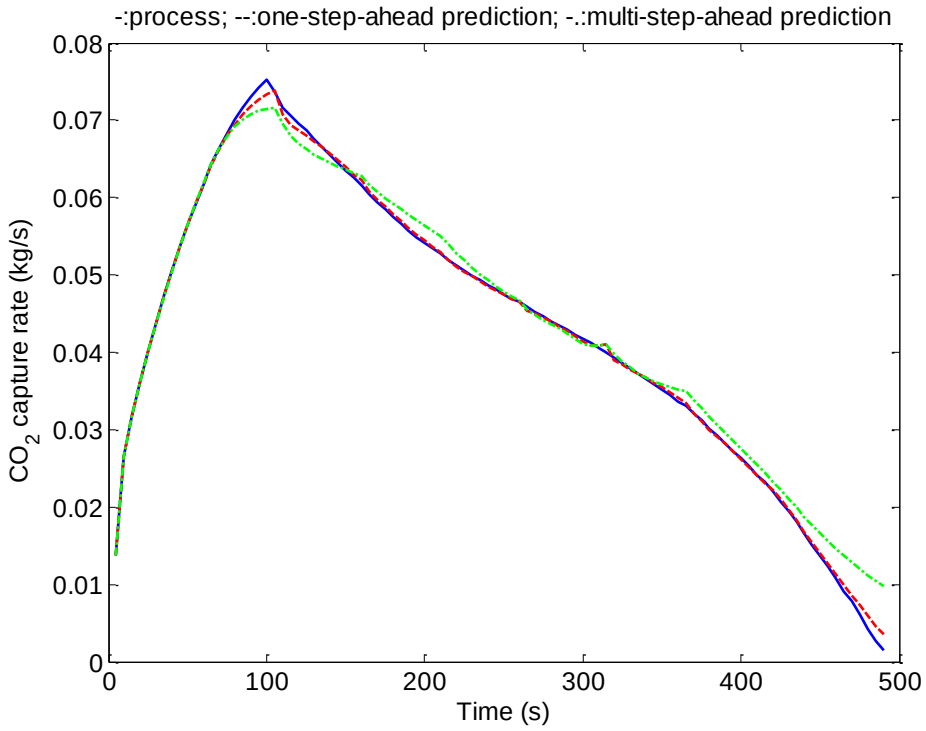
Figure 7. Steady state model predictions for CO_2 capture level on unseen validation data



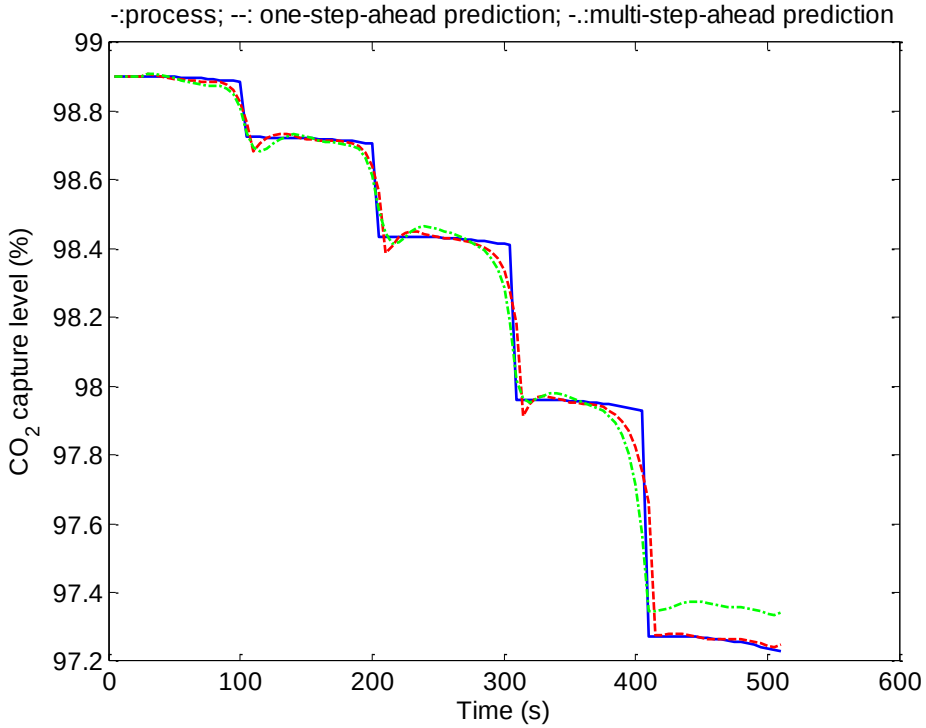
555
556
557 Figure 8. MSE of CO_2 capture rate for individual neural networks
558
559



560
561
562 Figure 9. MSE of CO_2 capture rate for aggregated bootstrap neural networks



563
564
565 Figure 10. Dynamic model prediction of CO_2 capture rate
566
567



568
569
570 Figure 11. Dynamic model prediction of CO_2 capture level