Forecasting greenhouse gas emissions from coal-based resource in power plant using a nonsupervisory artificial neural network

Mohd Fauzi Zanil 1*, Kiat Moon Lee 1, Radin Diana R. Ahmad 2, Sazalina Zakaria 2, and Raja Shazrin Shah RES 3,4

1 Department of Chemical & Petroleum Engineering, Faculty of Engineering, Technology and Built Environment, UCSI University, 56000 Kuala Lumpur, Malaysia.
2 Built Environment and Climate Change Unit, Generation and Environment Department, TNB Research Sdn. Bhd., No.1, Lorong Ayer Itam, Kawasan Institusi Penyelidikan, 43000 Kajang, Selangor.
3 Galaxy Tech Solutions Suite 2.3, Innovation Incubator Complex, University of Malaya, 50603 Kuala Lumpur.
4 Department of Chemical Engineering, Faculty of Engineering, University of Malaya, 50603 Kuala Lumpur.

*E-mail: mohdfauzi@ucsiuniversity.edu.my

Abstract. Machine learning can be a game-changer for a global warming prediction. About 75% global greenhouse gas (GHG) emissions cause by energy sector and this indicate a major concern to global warming community. In this study, non-supervisory machine learning technique has been used to predict the GHG effect relate to net calorific value based on intergovernmental panel on climate change (IPCC) standard. The study focuses on the characteristic of coal that is used in power generation sector and its chemical effluent that obtained from ultimate analysis (dried basis; Carbon, Hydrogen, Oxygen, Nitrogen, Sulphur and Ash) as gas emissions is concerned. The dataset shows, coal from different origin and type produce GHG emissions range approximately between 86.95 and 108.23 k-tonne CO$_2$/TJ with the net calorific value of 19.77 to 27.17 MJ/kg-coal. While, for ultimate analysis, the percentage of Carbon, Hydrogen, Oxygen, Nitrogen, Sulphur and Ash are in the range of [65.05 – 73.3], [1.46 – 5.49], [1.2 – 19.06], [0.3 – 1.20] and [4.82 – 15.96] respectively. In this study, principal component analysis is used to screen the training dataset and feed forward structure from artificial neural networks are used which allows the trained model to determine the GHG emission factor based on the given input data. The network relative errors of year 2017 dataset were used to adjust the weight value and as a result, the networks give r-square of 0.91678, which subsequently the trained networks are simulated for GHG emissions prediction for year 2018 at accuracy of r-square 0.82191. Furthermore, the study also shows, they are significant effect from coal characteristic towards GHS emissions and study proposed an optimal solution to simultaneously maximise power generation (in net calorific value per consumption weight) and reducing GHG value (k-tonne CO$_2$/TJ) of coal plant.

Keywords: Non-supervise ANN, Greenhouse Gas Forecasting, Principal Component Analysis.
1. Introduction
According to the Intergovernmental Panel on Climate Change (IPCC), the global temperature (as in 2017) has increased approximately 1°C above pre-industrial levels, implying 0.2°C increase per decade (“Global Warming of 1.5°C —,” n.d.). It is corresponded to the Fifth Assessment Report by IPCC which reported that the annual global GHG emissions have continued to grow and reached 49.5 Gt of carbon dioxide equivalents (CO₂ eq) in the year 2010, higher than any prior years [1].

GHG emissions is largely contributed by the generation of electricity with approximately 35% of the anthropogenic GHG emissions in 2010 [1]. Electricity generation from fossil fuels (coal, natural gas, and oil) make up 64% of the global energy generated (26,700 TWh) while the remaining are from renewable sources such as solar PV, wind, nuclear, hydro and other renewable sources [2]. The increasing demand of energy for human productivity has led to the increasing demand of fossil fuels which ultimately increases the emission of greenhouse gases (GHG) into the atmosphere. Thus, the indication of emission of GHG is critical aspect to mitigate the GHG effect to environment and this can only be carried out by understanding the correlation of input factor towards GHG emission.

In empirical model, heuristic data from real system are required. The data hold the information regarding the characteristic of input-output of the system process [3]. Artificial neural networks (ANN) are normally and extensively used in empirical modelling design and already establish at the open literature [4-6]. The empirical model able to give estimation of interested factor without going through the physical phenomena of input and output correlation. Hence, for environmental study and analysis, empirical model has a potential to provide an indicator to emission factor for specific power plant.

In this study, the objective is to develop a data driven model to predict the CO₂ GHG emission factor for year 2018 by using artificial ANN machine learning technique. The scope of study involves on utilising the database from 9 coal power plants for financial year of 2017 (for developing an ANN model) and 2018 (for testing ANN model performance). The study is limited to dataset from the ultimate analysis of coal fuel and the emission factor value was used as from IPCC 2006 guideline.

2. Greenhouse Gas Emission
The calculation of GHG mitigation for clean coal technology was done based on an estimation of how much GHG emissions reduction are possible through penetration of advanced clean coal technologies (CCTs) such as ultra-supercritical technologies (USC) for coal-based thermal power plants [7]. While the emission mitigation was calculated using the default carbon sink (0.03 tCO₂/year/tree) as stipulated under the Low Carbon Cities Framework and Assessment System [8].

The emissions intensity assessment conducted to derive estimates of GHG emissions based on the amount of fuel combusted. The intensity of CO₂ emissions was calculated based on emissions from the stationary combustions divided by the net generation. The ultimate analysis of each chemical component in the coal and calorific value has been used to calculate emission factor. In this study, ultimate analysis is used as development dataset to design artificial neural network in next section.

3. Methodology of Non-Supervisory Artificial Neural Networks
Data collection activities has been carried out to make dataset available. This will be an iterative process throughout year 2017 where details of data are taken from fuel information and gas emission composition. This screening process is slow and requires many judgements before it can be made usefulness in data bank inventory. The study pre-processes the input dataset by using non-supervisory technique (principle component analysis) to identify the nature of dataset in data bank inventory. Figure 1 show, two components has been used in case. Figure 1 also indicate nine colours which represent different coal brand used in the power plant. The outcome of this result shows plant 8 and 9 proportionally linear towards each component and separated from other coal brand. As the result, this dataset was removed from the training dataset later in the modelling phase.
In this study, artificial neural networks (ANN) with several hidden layers structure is used. The networks can be formulated as in Eq. 1 [9] and the output is to determine the GHG emission factor based on the given input data from ultimate analysis.

\[
O_m = \text{purelin} \left( \sum_{j=1}^{\text{neuron}} \left[ rw_{m,j} \left( \text{tansig} \left( \sum_{i=1}^{4} [hw_{j,i} \cdot x_i]_m + B_j \right) \right) \right] + C_m \right) ; \quad m = 1, 2, 3 \tag{1}
\]

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure1.png}
\caption{Principal component analysis for input dataset}
\end{figure}

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure2.png}
\caption{ANN configuration with 1 layer for input, hidden and output [9].}
\end{figure}
In this study, six inputs have been selected from ultimate analysis while GHG emission factor is the model output. The configuration of ANN for this study can be written as $\text{ANN}[6:H_n:1]$ where $H$ is number of node in hidden layer while $n$ represent the number of hidden layer in the model. The ANN with Levenberg Marquardt is used as training algorithm while sigmoid is used as activation function at every node in the hidden network.

4. Result and Discussion
The study is to predict the GHG emission factor by using artificial neural networks with PCA assisted result. The study also indirectly indicates the feasibility of the proposed work to model environmental system. The model is validated by using the trained ANN structure where every weight in each node has been optimised. The effectiveness the model is measured by residual analysis (regression) and above 0.8 is desired to have good of fitness. For this reason, many configuration and ways can be implemented to improve the ANN performance. As refer to Figure 3(a), initial performance is below 0.8 and the study increase the number of layer as an alternative solution to improve the fitness. This can be realised as in Figure 3(b) and Figure 3(c). At this stage the generalization of ANN model is assumed, and the trained model is tested to another dataset (for year 2018). For this dataset, ANN model does not have any knowledge about the coal characteristic in year 2018 whereas only 6 inputs are provided, and prediction output is compared with GHG emission factor for each set in the data bank of 2018 (as refers to Figure 4(a) and 4(b)).

![Graphs](a) ANN [6:10:1] (b) ANN [6:10:10:1]
5. Conclusion
In conclusion, results obtained from this study indicate that using the Artificial Neural Networks (ANN) model scheme is a feasible and it shows good prediction of the greenhouse gas (GHG) emission factor. ANN give r-square of 0.8 (above) for Training Phase and subsequently the GHG emissions prediction for year 2018 at accuracy of r-square 0.7 (above). The study also shows, there are significant effect between origin and characteristic. The study can be used to propose a mitigation plan to reduce GHG emission factor. This study has potential to maximise power generation (in net calorific value per consumption weight) and reducing GHG emission factor value (k-tonne CO$_2$/TJ) of coal plant.

Acknowledgement
This work was supported by the AAIBE PROJECT and CERVIE [CONF-INT-FETBE-230] by UCSI University.
References
[1] IPCC, Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Field, C.B., V.R. Barros, D.J. Dokken, K.J. Mach, M.D. Mastrandrea, T.E. Bilir, M. Chatterjee, K.L. Ebi, Y.O. Estrada, R.C. Genova, B. Girma, E.S. Kissel, A.N. Levy, S. MacCracken, P.R. Mastrandrea, and L.L. White (eds.)]. 2014, Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press. 1132.

[2] Malaysia Energy Statistics Handbook, S. Tenaga, Editor. 2018, Suruhanjaya Tenaga.

[3] Antanasijević, D.Z., et al., Forecasting GHG emissions using an optimized artificial neural network model based on correlation and principal component analysis. International Journal of Greenhouse Gas Control, 2014. 20: p. 244-253.

[4] Nabavi-Pelesarai, A., R. Abdi, and S. Rafiee, Neural network modeling of energy use and greenhouse gas emissions of watermelon production systems. Journal of the Saudi Society of Agricultural Sciences, 2016. 15(1): p. 38-47.

[5] Antanasijević, D., et al., Modeling of energy consumption and related GHG (greenhouse gas) intensity and emissions in Europe using general regression neural networks. Energy, 2015. 84: p. 816-824.

[6] Khoshnevisan, B., et al., Application of artificial neural networks for prediction of output energy and GHG emissions in potato production in Iran. Agricultural Systems, 2014. 123: p. 120-127.

[7] ADB. Guidelines for Estimating Greenhouse Gas Emissions of Asian Development Bank Projects: Additional Guidance for Clean Energy Projects. 2017 [cited 2019; Available from: https://www.adb.org/sites/default/files/institutional-document/296466/guidelines-estimating-ghg.pdf.

[8] KeTTHA. Low Carbon Cities Framework and Assessment System. 2011 [cited 2019; Available from: http://lccftrack.greentownship.my/files/LCCF-Book.pdf.

[9] Abnisa, F.A.S., S.D.; bin Zanil, M.F.; Wan Daud, W.M.A.; Indra Mahlia, T.M., The Yield Prediction of Synthetic Fuel Production from Pyrolysis of Plastic Waste by Levenberg–Marquardt Approach in Feedforward Neural Networks Model. Polymers, 2019. 11: p. 1853.