Jurnal Ilmu Alam dan Lingkungan 17 (1), (2026). 35 - 46

Jurnal
Ilmu Alam dan Lingkungan

Open Access https://journal.unhas.ac.id/index.php/jai2

Application of Adaptive Neuro-Fuzzy Inference System Model for Predicting CO_2 Emission Based on Production Rate in Sugarcane Agroindustry

Sri Mursidah[1*], Taufik Djatna[2]

[1]*Department of Environmental Engineering, UIN Walisongo Semarang, Semarang 50185, Jawa Tengah, Indonesia*
[2]*Agro-industrial Engineering, Bogor Agricultural University, Bogor 16680, Jawa Barat, Indonesia*

[*]*Email: sri.mursidah@walisongo.ac.id*

Abstract

Carbon dioxide (CO_2) emissions from industrial processes, particularly in sugar production, have become an important environmental concern that requires accurate prediction and monitoring. This study aims to predict carbon dioxide (CO_2) emissions using the Adaptive Neuro-Fuzzy Inference System (ANFIS) model based on production rate and the fuel used for operating the boiler machine. The model was developed using hypothetical data obtained from the literature, which were divided into 80% training data and 20% testing data. A trial-and-error approach was employed to determine the optimal parameters of the ANFIS model. Various membership function (MF) types and numbers were evaluated, and the optimal configuration was found to be three MFs with a triangular MF type. Model performance was assessed using Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), and the coefficient of determination (R^2). The results show that the ANFIS model provides good predictive performance, with an RMSE value of 17.083, a MAPE of 14.338, and an R^2 of 0.8686. These findings indicate that the proposed ANFIS model has a high capability in estimating CO_2 emissions based on production rates and fuel usage.

Keywords: *ANFIS, CO_2 emission, fuel usage, production rate, sugarcane agroindustry*

INTRODUCTION

The sugar production process generates various environmental impacts that may threaten the sustainability of the industry, particularly in terms of air quality. Among these impacts, carbon dioxide (CO_2) emissions have become a major concern due to their significant contribution to greenhouse gas (GHG) accumulation and global warming (Stocker *et al.,* 2013; Zahedi *et al*., 2018). CO_2 is projected to continuously increase, relying on historical data. Globally, 80% of CO_2 emission is from the combustion of fossil fuel mainly in industry (Kunda & Phiri, 2017). Sugar agroindutry become one of the largest contributors to CO_2 emissions (Saleh *et al.,* 2015). The great development of the sugarcane

P ISSN: 2086 - 4604
E ISSN: 2549 - 8819

production industries has created a need for recalculating the emission of pollutants (Davis & Caldeira, 2010). In the sugar industry, CO_2 emissions mainly arise from fuel combustion in production processes, especially in boilers, lime kilns, diesel generators, electrical equipment in factories, and sugarcane burning for the elimination of waste (Davis & Caldeira, 2010). The combustion of the boiler emits CO_2, which will potentially produce air pollutants (Saleh *et al.,* 2015). Accurate prediction of CO_2 emissions, therefore, requires the identification of key driving factors. Emission levels are primarily influenced by sugar production, fuel type, and fuel consumption. Higher levels of sugar production increase energy demand, leading to greater fuel consumption and consequently higher CO_2 emissions. In addition, variations in fuel type used in boiler operations significantly affect the magnitude of emissions. This study focuses on analyzing these variables to better understand and predict CO_2 emissions in the sugar production process.

There have been many different methods of forecasting due to the different locations and factors of CO_2 emissions (Abdullah & Pauzi, 2015). Accurate prediction of CO_2 emissions in such processes remains challenging due to the complex, nonlinear, and dynamic nature of industrial systems. Conventional mathematical models are often limited in representing these characteristics, especially under conditions involving uncertainty and variability. Therefore, more adaptive and flexible approaches are required to improve prediction performance. Soft computing techniques, such as artificial neural networks and fuzzy logic, have been widely used to model complex and nonlinear systems. Artificial neural networks are capable of learning patterns from data and representing those in synapse weights (Huang *et al.,* 2018), while fuzzy systems can handle uncertainty and linguistic variables. The integration of these methods in the Adaptive Neuro-Fuzzy Inference System (ANFIS) provides a robust approach for modeling nonlinear relationships in complex environments. Adaptive Neuro-Fuzzy Inference Systems (ANFIS) is one of the techniques in neuro-fuzzy systems that provides accurate and reliable predictions (Blanes-Vidal *et al.,* 2016; Zeinalnezhad *et al,* 2020). ANFIS can make input-output mapping based on human knowledge and assign input-output data pairs using hybrid learning procedures. The ANFIS architecture is used in simulations to model nonlinear functions, predict chaotic time series, and identify online nonlinear components in the control system (Zeinalnezhad *et al.,* 2020).

Various studies have explored the application of intelligent techniques in environmental modeling, particularly for air pollution forecasting. Methods such as fuzzy logic and artificial neural networks (ANN) have been widely used due to their ability to handle complex and nonlinear relationships. As mentioned by Woldt *et al*., (2003), fuzzy techniques can be used to prevent industrial pollution. Fuzzy techniques are effective in representing uncertainty and linguistic variables for air quality assessment (Woldt *et al*., 2003; Tokhmehchi & Makui, 2015), while ANN has demonstrated strong performance in learning patterns and providing reliable predictions of air pollution and CO_2 emissions (Saleh *et al*., 2015; Rahman *et al*., 2017; Maleki *et al*., 2019). Furthermore, hybrid approaches integrating ANN with other methods, such as life cycle assessment (Kaab *et al*., 2019) and fuzzy-based systems (Glinsky *et al*., 2016), have been developed to improve prediction accuracy. Despite these advancements, each approach has its limitations. Fuzzy systems rely heavily on expert-defined rules, which may introduce subjectivity, while ANN models often require large datasets and may lack interpretability. Hybrid models, although more accurate, tend to be more complex and computationally demanding. To address these limitations, the Adaptive Neuro-Fuzzy Inference System (ANFIS), which combines the learning capability of ANN and the reasoning mechanism of fuzzy logic, has been proposed as a promising alternative (Asklany *et al*., 2016; Huang *et al*., 2018). Previous studies have shown that ANFIS can provide higher accuracy compared to conventional regression models in predicting air pollution (Zeinalnezhad *et al*., 2020) and effectively model nonlinear systems.

P ISSN: 2086 - 4604
E ISSN: 2549 - 8819

However, existing studies have largely focused on general air pollution forecasting, with limited application of ANFIS for predicting CO_2 emissions in specific industrial contexts. In particular, the integration of key production-related variables, such as sugar output, fuel type, and fuel consumption, in ANFIS-based models remains underexplored in the sugar industry. This gap highlights the need for developing a more tailored and adaptive prediction model that can capture the complex relationships between production activities and CO_2 emissions. Therefore, this study aims to develop a CO_2 emission forecasting model using ANFIS, providing a more accurate and reliable tool for emission prediction in the sugar production process. Accurate emission prediction plays a crucial role in improving air quality (Liu *et al.*, 2019) and can support both industry practitioners and policymakers in making more informed decisions.

METHODOLOGY

This study applies ANFIS modelling to predict CO_2 emissions. The research flow of the present study related to ANFIS modelling is shown in Figure 1. Research flowcharts consist of 4 steps. Step 1 is analysis of data, step 2 is data preprocessing, step 3 is running the model, and step 4 is performance evaluation of the prediction model.

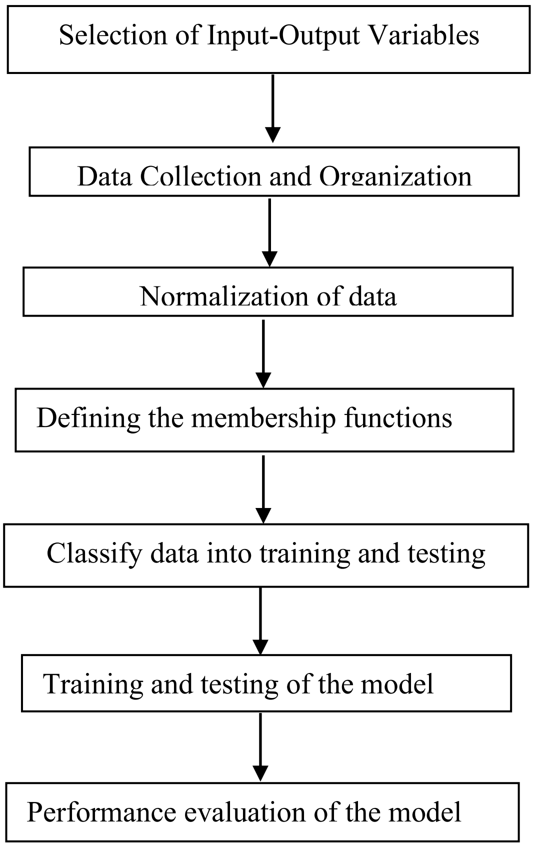

Figure 1. Research Flowchart.

1. Selection of Input and Output Variables

The input used consists of many variables, namely the factors that affect CO_2 emissions released by the sugar industry. Input variables used are sugar production, gasoline consumption, solar consumption, residue consumption, oil consumption, and electricity. Meanwhile, the output variable is the resulting CO_2 emissions.

P ISSN: 2086 - 4604
E ISSN: 2549 - 8819

2. Data Collection and Organization

Data was obtained from a survey of literature and then analyzed using the proposed model. Six activities in the sugar industry have the potential to emit CO_2 emissions, such as cane transport activity, plastic sacks transportation, the use of sugar cane trucks, combustion in the boiler machine, the use of electricity, and the distribution of sugar to consumers. Several variables that affect CO_2 expenditure, such as sugar cane milling machine, boiler, and engine efficiency, lorry mileage, and sugar has a significant impact. CO_2 emissions were measured using both direct and indirect methods. Direct measurement involves activities controlled by the company, such as combustion processes, transport units, refrigeration, and cooling systems. While indirect measurement of the CO_2 emissions of the activity performed outside the control of management companies, such as the purchase of energy from the company or the source of energy. As exemplified using electricity (Saleh *et al.,* 2015). In this research measurement focused on direct measurement in the combustion chamber and indirect measurement in electricity usage. The number of canes processed is also considered because it varies every day, so it can cause variations in resources and energy used.

Table 1. Variables to Develop A Prediction Model (García-Bustamante *et al*., 2018)

Variable	Description	Range
X1	Sugar production (t/a)	2000-222.22
X2	Gasoline consumption (Gg/year)	1-1.5
X3	Solar consumption (Gg/year)	31-47
X4	Residue consumption (Gg/year)	53-93
X5	Oil consumption (Gg/year)	13-23
X6	Baggase Production (Gg/year)	2943603-4930251
X7	Electricity (MJ/MJ fuel)	2-2.5
Y	CO2 Emission (Gg/year)	5732-6157

3. Normalization of Data

All data points for both variables were normalized. This normalization process brings the values of variables between -1 and 1. This normalization of data was done using the following formula (Prasad *et al.,* 2016):

$$X = \left[2\frac{a - a_{min}}{a_{max} - a_{min}}\right] - 1$$

where X is a normalized value, 'a' is the observed value of the variable, a_{max} is the maximum observed value, and a_{min} is the minimum observed value.

4. Defining The Membership Function

The initial step in developing the model is to build a set of primary memberships using ANFIS. Four types of input membership functions (triangular, gaussian, trapezoidal, and gaussian bell) and two types of output membership. The membership functions can be constructed from several basic functions, such as piecewise linear functions, the gaussian distribution function, the sigmoid curve, and quadratic and cubic polynomial curves (Prasad *et al.,* 2016). The MATLAB software package is used for ANFIS modeling.

P ISSN: 2086 - 4604
E ISSN: 2549 - 8819

5. Model Training and Testing

After data loading, membership number and type for each input and output variable were assigned to generate the FIS system, the models are trained for each input combination as described above, and the performance of the model are tested using the testing data sets. The collected data were divided into training (80%), and testing (20%). Both the training data sets and testing data sets are loaded in the ANFIS toolbox of MATLAB software.

6. Performance Evaluation of The Model

Performance of the model can be seen in RMSE (Root Mean Squared Error), MAPE (Mean Absolute Percentage Error), and R (coefficient of determination). They are an indicator that is widely used in literature to measure the correctness of the expanded models. These performance indicators are described as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(P_i - A_i)^2}{\sum_{i=1}^{n} {A_i}^2}$$

$$SE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(P_i - A_i)^2}$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{|P_i - A_i|}{A_i}$$

Where P_i and A_i are the predicted and the actual amounts for the ith record, and n is the number of total records. The estimation accuracy of a model is confirmed by RMSE. R^2 demonstrates the percentage variance illustrated by the model. If R^2 approaches 1, it represents an advanced correlation.

RESULTS AND DISCUSSION

1. ANFIS Model Analysis

Predicting CO_2 emissions in the sugar agroindustry is done using ANFIS modeling. The ANFIS learning process includes training, testing (validation), and checking the data set. The dataset used is 200 hypotetically set of data. Approximately 80% of the dataset is allocated for training data, while 20% of the dataset is allocated for validation and checking. Figure 3 shows the distribution of data training, testing, checking, and the FIS-generated outputs for the ANFIS structure. Three membership functions were used for each input, because MFs number “3” and “triangular” for MF type is appropriate for use after testing many other types of membership functions. Figure 4 shows the ANFIS architecture. The ANFIS information is shown in Table 2.

Table 2. ANFIS Information.

Parameter	Value
Number of nodes	4426
Number of linear parameters	2187
Number of nonlinear parameters	63
Number of training data pairs	159

P ISSN: 2086 - 4604
E ISSN: 2549 - 8819

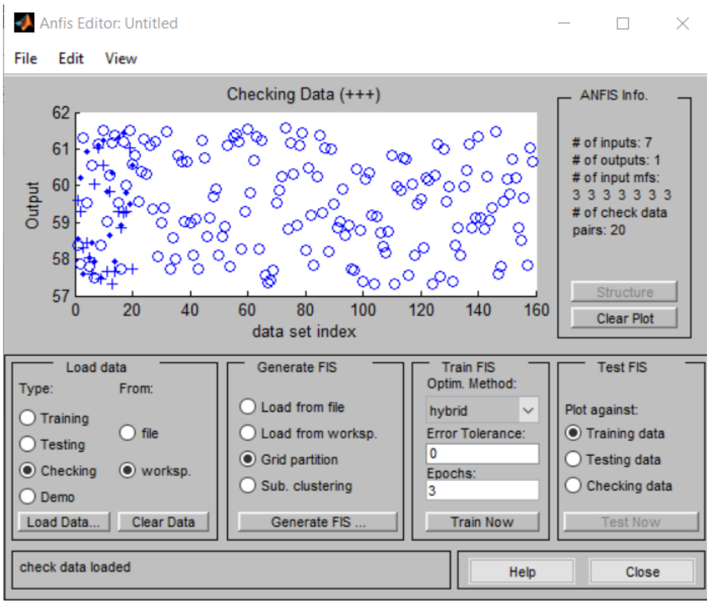

Figure 2. Data Distribution.

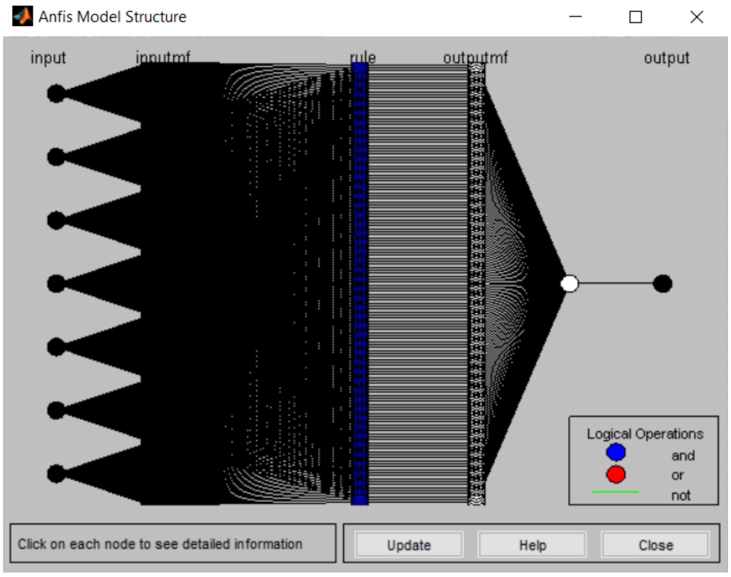

Figure 3. ANFIS Architecture.

The ANFIS model consists of an input layer with seven input variables, one hidden layer consisting of 21 input membership functions, and an output layer with a single output variable. Figure 4-6 provides ANFIS network results for training, validation (testing), and checking of CO_2 emission. In the ANFIS analysis, the inputs and output variables were correlated using three input membership functions and a set of rules. These membership functions and rules produced fuzzy outputs. Finally, the fuzzy outputs were defuzzified to obtain a scalar output.

P ISSN: 2086 - 4604
E ISSN: 2549 - 8819

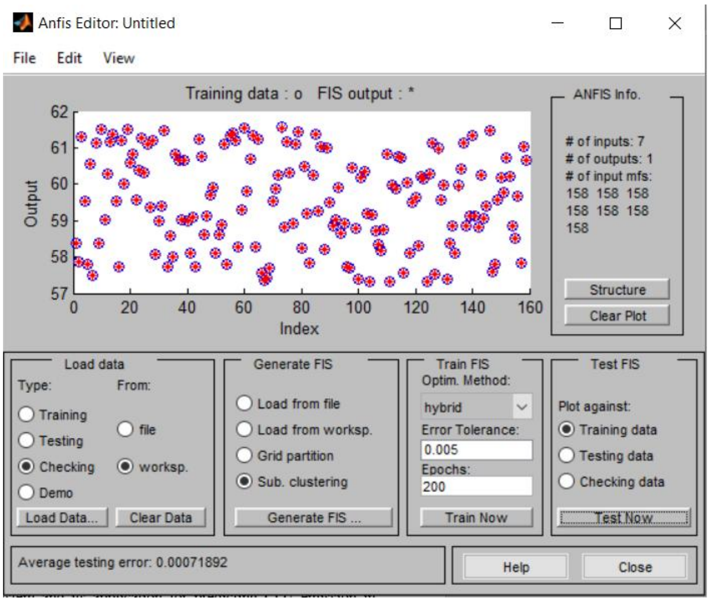

Figure 4. ANFIS Results for Training Data.

Figure 4 is associated with the training data of CO_2 emissions. The FIS outputs are shown with small red stars in these figures. The values related to the training data are marked with small blue circles. The adjacent blue small circles show that the calculated and experimental values at that point are close together, and the total error calculation is small. The larger divergence between stars and circles shows computational errors.

Figure 5. ANFIS Results for Testing Data.

Figure 5 illustrates the results associated with the testing data of CO_2 emissions. The demarcation of the FIS outputs is represented by small red stars, while the testing data are depicted using blue dots. The proximity of the blue dots to the red stars indicates a close agreement between the experimental and calculated values at those points, resulting in a low overall error. The larger deviation between dots and stars shows total error over the experimental range.

P ISSN: 2086 - 4604
E ISSN: 2549 - 8819

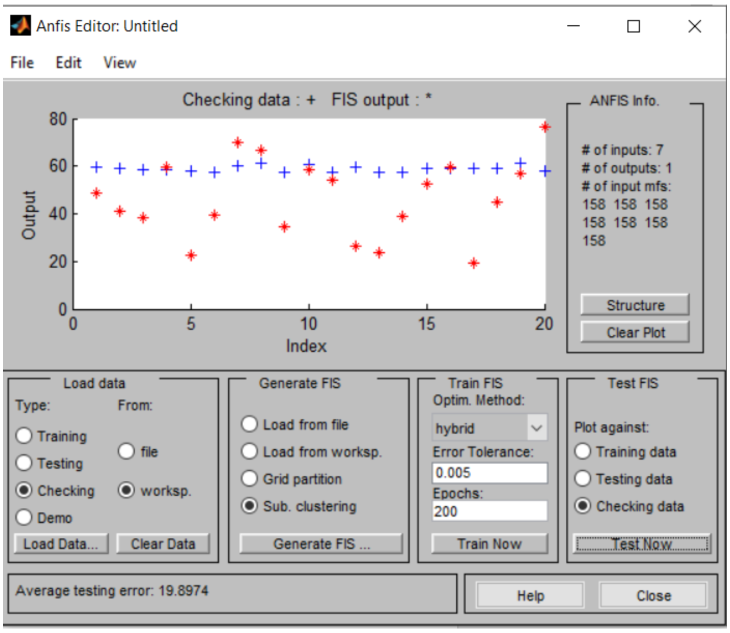

Figure 6. ANFIS Results for Checking Data.

The FIS outputs are marked with small red stars in these figures, while the validation data are represented by small blue crosses. The proximity of the blue crosses to the red stars indicates that the empirical and calculated values at those points are in good agreement, resulting in a low overall error. This pattern suggests that the ANFIS model is able to capture the underlying nonlinear relationships between input variables and CO_2 emissions with a high degree of accuracy. Conversely, larger deviations between the red stars and blue crosses indicate higher prediction errors, which may arise from data variability, noise, or the presence of operating conditions that are not sufficiently represented in the training dataset. The fuzzy surfaces of the developed model, presented in Figure 7, illustrate how CO_2 emissions respond to variations in key input variables.

These surfaces reveal smooth nonlinear trends, indicating that the model effectively generalizes the interaction effects among inputs such as fuel characteristics and production levels. In some regions, steeper gradients can be observed, suggesting that small changes in certain inputs can lead to significant increases in emissions. This behavior is consistent with the combustion process, where emission levels are highly sensitive to energy input and fuel properties. Additionally, relatively flatter regions on the surface indicate ranges where changes in input variables have a less pronounced effect on emissions, implying more stable operating conditions. The contrast between steep and flat regions reflects the varying sensitivity of CO_2 emissions to different operational factors. For example, in regions where emissions respond sharply to fuel consumption, the system is highly sensitive, suggesting that careful monitoring and control of fuel usage could significantly reduce emissions. Conversely, in flatter regions, the system shows robustness to input variations, indicating operational settings where emission levels remain relatively stable even if small fluctuations in input occur. This analysis demonstrates that the ANFIS model not only predicts CO_2 emissions accurately but also provides insights into which input variables most strongly influence emission behavior.

P ISSN: 2086 - 4604
E ISSN: 2549 - 8819

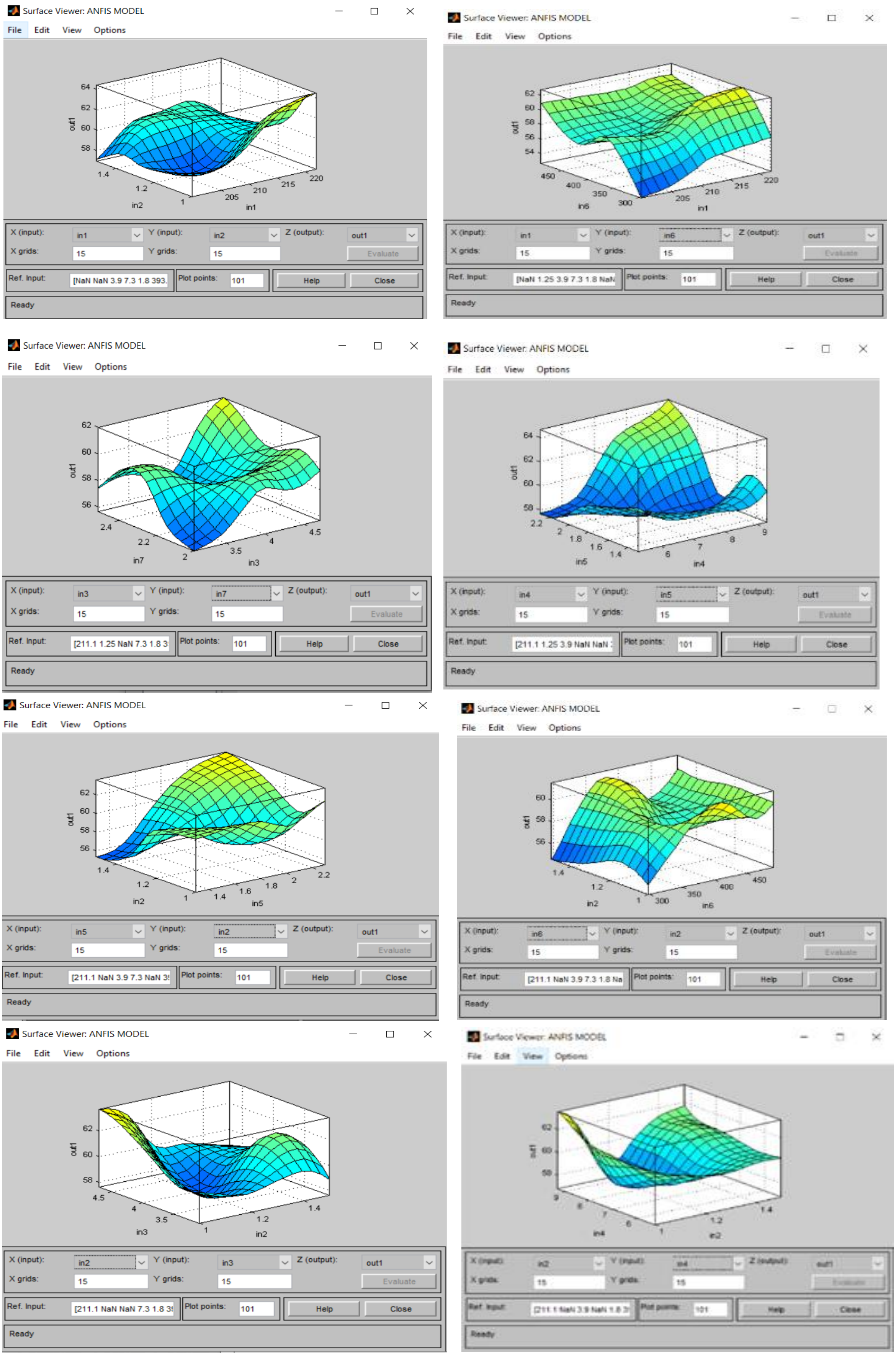

Figure 7. Surface Diagrams Presenting the Dependence of CO_2 Emission on the Input Parameter.

P ISSN: 2086 - 4604
E ISSN: 2549 - 8819

2. Performance Model Analysis

Performance of the model is analyzed by comparing actual and predicted values. Figure 8 shows the scatterplot of the actual versus the predicted CO_2 emission using ANFIS modeling. The performance indicators of the proposed ANFIS model are summarised in Table 3.

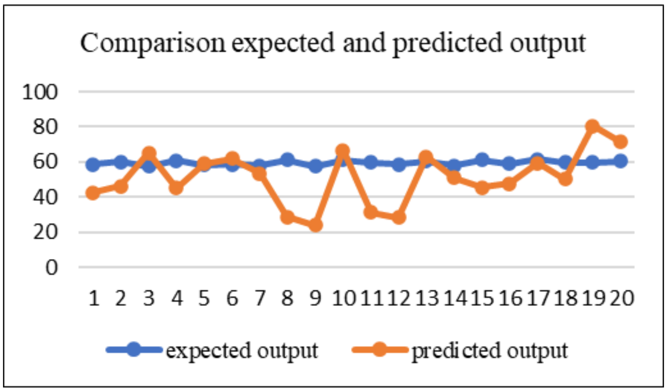

Figure 8. Comparison of Expected and Predicted CO_2 Emission.

Table 3. Performance indicators of ANFIS Model

Emission	**ANFIS model's performance**		
	RMSE	**MAPE**	**R^2**
CO_2	17.083	14.338	0.868

According to the calculated evaluation criteria, the ANFIS model exhibits relatively low errors, indicating good predictive performance. The root mean squared error (RMSE) of the ANFIS model is 17.083. The RMSE is the square root of the variance of the residuals. It indicates the absolute fit of the model to the data–how close the observed data points are to the model's predicted values. Lower values of RMSE indicate a better fit. RMSE is a good measure of how accurately the model predicts the response, and it is the most important criterion for fit if the main purpose of the model is prediction. In this case, the relatively low RMSE demonstrates that the ANFIS model can accurately capture the variation in CO_2 emissions and provide reliable predictions. According to Table 3, the values of Mean Absolute Percentage Errors (MAPE) for the ANFIS model is 14.338. Based on Lewis (1982), the results of a prediction method have excellent forecasting ability if the MAPE value is <10% and have good predictive ability if the MAPE value is between 10% and 20%. The coefficient of determination (R^2) for the ANFIS model is 0.8686, indicating a high proportion of variance in the observed CO_2 emissions explained by the model. This confirms that the model has strong estimation power and can reliably represent the underlying emission dynamics. The combination of RMSE, MAPE, and R^2 values indicates that the ANFIS model is capable of modeling both the central trend and variability in CO_2emissions effectively. Comparisons with other studies support the robustness of the ANFIS approach. For instance, Nassef *et al.*, (2023) applied ANFIS for CO_2 emission prediction and obtained RMSE = 20.89505 and R^2 = 0.98875. Similarly, Khan & Khan (2019) reported that ANFIS outperformed conventional techniques, with RMSE values of 0.1157 for ANFIS and 0.1915 for conventional methods. These results suggest that ANFIS not only provides accurate predictions but also surpasses traditional modeling approaches in capturing nonlinear behaviors in emission data.

P ISSN: 2086 - 4604
E ISSN: 2549 - 8819

CONCLUSION

This paper presents the development of a neuro-fuzzy system and its application for predicting CO_2 emission in the sugarcane agroindustry. The Adaptive Neuro-Fuzzy Inference System algorithm was successfully implemented to model carbon dioxide emissions based on relevant production and energy-related input variables. The model achieved strong predictive performance, as indicated by low RMSE and MAPE values and a high R^2, confirming its reliability in estimating emissions under varying operational conditions. In addition to its predictive capability, the ANFIS model produces interpretable fuzzy rules that reflect the influence of key input variables, providing practical insight for emission monitoring and control. These characteristics make the model suitable for supporting data-driven decision-making in industrial emission management. However, the model performance is dependent on the quality and variability of the input data, and its generalization may be limited to conditions like those used in this study. Future research should therefore emphasize enhancing model robustness through hybrid optimization approaches, such as integrating ANFIS with genetic algorithms and conducting validation using datasets from diverse industrial contexts to improve its broader applicability.

DAFTAR PUSTAKA

Abdullah, L., and Pauzi, H. M., 2015. *Methods In Forecasting Carbon Dioxide Emissions: A Decade Review*. Jurnal Teknologi. 75(1): 67-82.

Asklany, S. A., Elhelow, K., and Abd El-Wahab, M., 2016. *On using Adaptive Hybrid Intelligent Systems in PM10 Prediction*," International Journal Soft Computing and Engineering. 4: 2231-2307.

Blanes-Vidal, V., Cantuaria, M. L., and Nadimi, E. S., 2017. *A Novel Approach for Exposure Assessment in Air Pollution Epidemiological Studies Using Neuro-Fuzzy Inference Systems: Comparison of Exposure Estimates and Exposure-Health Associations*. Environmental Research. 154-196.

Davis, S. J., and Caldeira, K., 2010. *Consumption-Based Accounting of CO2 Emissions*. Proceedings of the National Academy of Sciences. 107(12): 5687-5692.

García-Bustamante, C. A., Aguilar-Rivera, N., Zepeda-Pirrón, M., and Armendáriz-Arnez, C., 2018. *Development of Indicators for the Sustainability of the Sugar Industry*. Environmental & Socio-Economic Studies. 6(4): 22–38.

Glinskiy, V., Serga, L., Khvan, M., and Zaykov, K., 2016. *Fuzzy Neural Networks in the Assessment of Environmental Safety*, Procedia CIRP. 40: 614–618. DOI: https://doi.org/10.1016/j.procir.2016.01.143.

Huang, M., Tian, D., Liu, H., Zhang, C., Yi, X., Cai, J., Ruan, J., Zhang, T., Kong, S., and Ying, G., 2018. *A Hybrid Fuzzy Wavelet Neural Network Model with Self-Adapted Fuzzy C-Means Clustering and Genetic Algorithm for Water Quality Prediction in Rivers*. Complexity. DOI: https://doi.org/10.1155/2018/8241342.

Kaab, A., Sharifi, M., Mobli, H., Nabavi-Pelesaraei, A. and Wing Chau, K., 2019. *Combined Life Cycle Assessment and Artificial Intelligence for Prediction of Output Energy and Environmental Impacts of Sugarcane Production*. Total Environment. 664: 1005–1019.

Khan, M. Z., and Khan, M. F., 2019. *Application of ANFIS, ANN and Fuzzy Time Series Models to CO_2 Emission from the Energy Sector and Global Temperature Increase*. International Journal of Climate Change Strategies and Management. 11(5): 622–642. DOI: https://doi.org/10.1108/IJCCSM-01-2019-0001.

Kunda, D and Phiri, H., 2017. *An Approach for Predicting CO_2 Emissions using Data Mining Techniques,* International Journal of Computers and Applications. 172(3): 7–10.

Liu, Z., Wang, F., Tang, Z. and Tang, J.,2020. *Predictions and Driving Factors of Production-Based CO_2 Emissions in Beijing, China.* Sustainable Cities and Society. 53.

Maleki, H., Sorooshian, A., Goudarzi, G., Baboli, Z., Tahmasebi Birgani, Y., and Rahmati, M., 2019. *Air Pollution Prediction by Using an Artificial Neural Network Model*. Clean Technologies and Environmental Policy. 21(6): 1341–1352.

Nassef, A.M., Olabi, A.G., Rezk, H., and Abdelkareem, M. A. 2023. *Application of Artificial Intelligence to Predict CO_2 Emissions: Critical Step towards Sustainable Environment*. Sustainability. 15(9). DOI: https://doi.org/10.3390/su15097648.

Prasad, K., Gorai, A. K. and Goyal, P., 2016. *Development of ANFIS Models for Air Quality Forecasting and Input Optimization for Reducing the Computational Cost and Time*. Atmospheric Environment. 128: 246–262.

Rahman, P. A., Panchenko, A. A., and Safarov, A. M., 2017. *Using Neural Networks for Prediction of Air Pollution Index in Industrial City*. IOP Conference. Series Earth Environment. 87(4). DOI: https://doi.org/10.1088/1755-1315/87/4/042016.

Saleh, C., Leuveano, R. A. C., Rahman, M. N. A., Deros, B. M., and Dzakiyullah, N. R., 2015. *Prediction of CO_2 Emissions Using an Artificial Neural Network: The Case of The Sugar Industry*. Advanced Science Letters. 21(10): 3079-3083.

Stocker, T.F., D. Qin, G., K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley. 2013. *Climate change 2013 the Physical Science Basis: Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, UK, New York.

Tokhmehchi, N., and Makui, A., 2015. *Developing A Fuzzy Model for Tehran's Air Quality.* Management. 5(1): 109-116.

Woldt, W. E., Dvorak, B. I. and Dahab, M. F., 2003. *Application of Fuzzy Set Theory to Industrial Pollution Prevention: Production System Modeling and Life Cycle Assessment*. Soft Computing. 7(6): 419-433.

Zahedi, A., Ghodrati, S., Ahmadimoghaddam, M., and Jaafarzadeh, N., 2018. *Estimating Greenhouse Gas Emissions Using Emission Factors from the Sugarcane Development Company, Ahvaz, Iran*. Environmental Health Engineering and Management Journal. 5(1): 9-14.

Zeinalnezhad, M., Chofreh, A. G., Goni, F. A and Klemeš, J. J., 2020. *Air Pollution Prediction Using Semi-Experimental Regression Model and Adaptive Neuro-Fuzzy Inference System*. Journal Cleaner Production. 261. DOI: https://doi.org/10.1016/j.jclepro.2020.121218.

P ISSN: 2086 - 4604
E ISSN: 2549 - 8819