BioTech SI

https://doi.org/10.17113/ftb.63.02.25.8792 original scientific paper

Predictive Modelling of H_2S Removal from Biogas Generated from Palm Oil Mill Effluent (POME) Using Biological Scrubber in an Industrial Biogas Plant: Integration of Artificial Neural Network (ANN) and Process Simulation

Running title: Modelling of H_2S Removal from Palm Oil Mil Effluent Biogas

Joanna Lisa Clifford[1], Yi Jing Chan[1]*, Mohd Amran Bin Mohd Yusof[1], Timm Joyce Tiong[1], Siew Shee Lim[1], Chai Siah Lee[2] and Woei-Yenn Tong[3]*

[1]Department of Chemical and Environmental Engineering, University of Nottingham Malaysia, Broga Road, 43500, Semenyih, Selangor Darul Ehsan, Malaysia
[2]Advanced Materials Research Group, Faculty of Engineering, University of Nottingham, NG7 2RD, UK
[3]Universiti Kuala Lumpur, Institute of Medical Science Technology, A1-1, Jalan TKS 1, Taman Kajang Sentral, 43000 Kajang, Selangor, Malaysia

Received: 24 July 2024
Accepted: 31 March 2025

SUMMARY

Research background. Biogas production from Palm Oil Mill Effluent (POME) is inherently unstable due to variations in feedstock composition and operating conditions. These fluctuations can result in inconsistent biogas quality, variable methane content, and fluctuating levels of hydrogen sulphide (H_2S), posing significant challenges for bioscrubbers in removing H_2S to meet the quality standards for gas engines used in electricity generation. This research aims to address these challenges by

*Corresponding authors:
E-mail: yi-jing.chan@nottingham.edu.my (Y.J. Chan); wytong@unikl.edu.my (W.-Y. Tong)

integrating simulation models using a computer programme and Artificial Neural Network (ANN) to predict the performance of a bioscrubber at a POME treatment plant in Johor, Malaysia.

Experimental approach. Initially, the process flowsheet model was simulated using a computer programme. The prediction of H_2S removal was then conducted using a machine learning algorithm, specifically ANN, based on two years of historical data obtained from the biogas plant. Furthermore, a detailed techno-economic analysis was conducted to determine the economic feasibility of the process.

Results and conclusions. Simulation results revealed a biogas yield of 26.12 Nm^3 of biogas per m^3 of POME, aligning with industry data with less than 1 % deviation. The ANN model achieved a high coefficient of determination (R^2) of 0.9 and a low mean squared error (MSE), with the bioscrubber reaching approximately 96 % H_2S removal efficiency. The techno-economic analysis indicated that the process is feasible, with a net present value of $131,000 and a payback period of 7 years.

Novelty and scientific contribution. The integration of ANN and the computer programme provides a robust framework for predicting bioscrubber performance and ensuring stable bioscrubber operation due to their complementary strengths. ANN accurately predicts H_2S removal using daily recorded data, while the computer programme estimates parameters not monitored daily, such as chemical oxygen demand (COD), biological oxygen demand (BOD), and total suspended solids (TSS). This research provides valuable insights into sustainable biogas production practices, offering opportunities to improve energy efficiency and environmental sustainability in the palm oil industry.

Keywords: palm oil mill effluent; biogas; simulation; Artificial Neural Network; bioscrubber

INTRODUCTION

Palm oil is a popular edible vegetable oil in tropical areas, that is derived from the mesocarp of the fruit of oil palm trees (*1*). Malaysia has become one of the top producers and exporters of the palm oil industry in the world. The palm oil industry has been a stable contributor towards the growth of Malaysia's economy. According to the Malaysian Palm Oil Board, the total land usage for palm oil cultivation in 2022 was recorded at approximately 5.67 million hectares. In 2022, crude palm oil (CPO) production in Malaysia was 18.45 million tonnes and is projected to rise by 3 % in 2023 (*2*). However, the potential environmental impacts due to high production rates are becoming more significant. With the number of palm oil mills increasing each year, the increase in by-products such as fresh fruit bunch waste and effluent discharge is inevitable (*3*).

Palm oil mill effluent (POME) is the wastewater produced from palm oil milling activities, mainly from the oil extraction, washing and cleaning processes. Due to its highly polluting properties, POME needs to be properly treated before being discharged into the environment (*4*). POME is a viscous, brown liquid with a pH of 4 to 5. POME is 100 times more polluted than municipal sewage which has high chemical oxygen demand (COD) and biological oxygen demand (BOD) (*3*). POME is one of the major problems in the production industry due to the large amounts of waste produced annually and its disposal issues (*5*). Therefore, POME treatment deserves more focus and emphasis to promote sustainability and circular economy in the industry.

There are several common effluent treatment systems used in the palm oil industry such as ponding systems, covered lagoon systems, closed anaerobic digesters and land application systems (*6*). Among these techniques, biological treatment is the most popular treatment method in many palm oil mills to treat POME (*7*). Most palm oil mills have implemented anaerobic digestion (AD) as the primary treatment for POME. The AD process normally operates at a temperature range of 30-65 °C (*8*). With AD, the biodegradation of POME releases a combustible methane-rich gas, which is captured to convert to renewable energy (*9*).

Biogas is the renewable energy source produced from the AD of POME in the absence of oxygen. Anaerobic digestion is a multi-step biochemical process of converting organic material into biogas. It is a natural fermentation process in which bacteria break down the organic matter into its components until all that is left are gases and a residue called digestate (*10*). Biogas can be used to produce heat and electricity with minimum impact on the climate and is a sustainable organic waste management solution. It is a promising alternative to fossil fuels (*11*). The chemical composition of the biogas is influenced by the characteristics of raw materials and the conditions under which AD is carried out. Biogas production sources encompass a diverse array of inputs such as wastewater, sewage sludge, animal or agricultural waste and landfill materials (*12*).

Biogas consists of mainly methane (CH_4) and carbon dioxide and with trace amounts of nitrogen, oxygen, hydrogen sulphide (H_2S) and water. The biogas produced in palm oil industries is mostly used for electricity generation. The major drawback associated with biogas is the presence of H_2S (*13*). Maintaining low H_2S concentrations is essential for broader environmental, health, and safety considerations. H_2S concentrations in unprocessed biogas range from 50 to 10,000 ppm, depending on the characteristics of POME. H_2S has been proven to corrode various steels, including stainless steel and other copper and nickel alloys. High concentrations of H_2S cause corrosion on equipment and increase the maintenance costs. It is also particularly corrosive on gas engines, shortening the engine life. This, in turn, increases the service and maintenance costs and decreases the conversion of biogas to electricity. Therefore, the recommended level of H_2S in biogas is around the range of 200 to 500 ppm, depending on engine specifications (*14*). Besides, high levels of H_2S in the atmosphere can cause environmental pollution and pose serious health risks, as H_2S is toxic and in high concentrations can lead to respiratory and neurological issues in humans. Additionally, unregulated H_2S emissions contribute to a pungent odor that affects nearby communities and the surrounding environment, potentially violating air quality standards.

Thus, the efficient removal of H_2S from biogas is crucial for ensuring the proper functioning of the gas engine and meeting regulatory standards. Bioscrubbers are used to remove H_2S from biogas in most palm oil mills. In bioscrubbers, microbial activity is observed to degrade contaminants in the biogas such as H_2S before it enters the gas engine (*15*). Due to their lower operating costs compared to physical and chemical methods, biological approaches are preferred. Besides that, they do not generate secondary streams. Moreover, the minimal use of chemicals makes biological approach more cost-effective and environmentally friendly than physicochemical and chemobiological methods.

Despite the potential of bioscrubbers, the lack of comprehensive simulation data poses significant challenges for biogas plants, such as the one in Johor, Malaysia, which serves as a case study in this research. Biogas production from POME is inherently unstable due to variations in feedstock composition, operating conditions, and other factors (*16*). These fluctuations can result in inconsistent biogas quality, variable methane content and fluctuating levels of H_2S, which subsequently affect the efficiency of the bioscrubber. The plant in Johor has reported difficulties in achieving consistent H_2S removal with their current bioscrubber due to these fluctuations and the inherent variability of biological processes. Factors such as temperature, pH, microbial community dynamics, and the concentration of H_2S and other compounds can change unpredictably, leading to inconsistent performance. The complexity of microbial interactions and their sensitivity to environmental changes make it challenging to maintain optimal H_2S removal. These necessitate advanced solutions to manage and mitigate these fluctuations.

Simulation plays a crucial role in addressing this instability by providing means to predict and analyze the behaviour of bioscrubbers under different operating conditions at minimum operating costs. By simulating bioscrubber performance, researchers and engineers can assess how variations in parameters such as inlet gas composition, flow rate, and temperature impact the efficiency of H_2S removal. This enables the identification of optimal operating conditions that maximises H_2S removal efficiency while minimising energy consumption and operational costs. Furthermore, simulation allows for the evaluation of bioscrubber performance over time and under varying conditions, providing insights into its long-term effectiveness and reliability. This is particularly important in the context of biogas plants, where consistent and reliable performance is essential for meeting regulatory requirements and ensuring the sustainable operation of the facility.

Yap and Hasanah (*17*) conducted a simulation study on a water scrubber process using ChemCAD simulation, demonstrating the feasibility of purifying POME-based biogas to produce high-purity methane. However, further optimisation and economic analysis of the process have not been conducted. This highlights the need for comprehensive studies that integrate both technical and economic aspects to assess the feasibility and viability of biogas purification processes. SuperPro Designer (*18*) is a versatile tool widely employed for simulating complex industrial processes, including those related to anaerobic digestion and biogas treatment systems (*19*). It enables detailed modelling of process streams, providing essential data such as COD values, while facilitating energy and mass balance calculations and incorporating critical reactions involved in anaerobic digestion. For POME treatment, this software has been employed to analyze biogas generation, treatment, and wastewater stabilization. Chong *et al.* (*20*) utilized it to assess the performance of an integrated anaerobic-aerobic bioreactor (IAAB), optimizing parameters for improved biogas yield and COD removal. Similarly, Kan *et al.* (*21*) demonstrated its accuracy in predicting treatment outcomes, with less than a 3 % error between simulated and experimental data. Despite its versatility, previous studies have not fully explored the role of bioscrubber systems and H_2S removal in meeting gas engine requirements. This highlights the potential to use SuperPro Designer (*18*) for simulating industrial-scale POME treatment processes, particularly focusing on bioscrubber integration.

Machine learning (ML) has advanced throughout the years, and it can be implemented as a prediction tool in the studies of microbial ecology and system biology (*22*). ML tools were employed in a few studies in analysing and predicting the performance of AD (*23,24*). Furthermore, the performance of biogas purification and H_2S removal has been reported to be affected by the inlet temperature, operation pressure, and biogas flow rate. Thus, to further understand the behaviour of biogas purification, ML tools can be implemented to forecast the performance of H_2S removal from biogas. The artificial neural network (ANN) approach has been reported to be among the most widely utilized ML approaches (*25*). However, applying ANN to bioscrubbers presents challenges, including the need for high-quality data due to the fluctuating conditions under which bioscrubbers operate, such as varying inlet gas composition, temperature, and pressure. This variability can lead to prediction inaccuracies if not properly accounted for. Additionally, non-linear interactions between bioscrubber parameters, like pH and microbial activity, demand advanced ANN architectures and precise model tuning. Overcoming these challenges requires enhanced data collection, model refinement, and real-time operational feedback to improve the accuracy and reliability of ANN predictions. Most importantly, there has been limited research on ANN-based prediction models for H_2S removal in bioscrubbers operating in real biogas plants.

Therefore, this research aims to address this research gap by integrating simulation models using SuperPro (*18*) and ANN to predict the performance of a bioscrubber based on two years of data recorded at a POME treatment plant in Johor. The simulation results were compared with industry data to validate the model's accuracy. Additionally, a techno-economic analysis was conducted to evaluate the feasibility and economic viability of the treatment plant, incorporating both POME treatment and biogas purification process. Through comprehensive simulation and analysis, this study aims to provide valuable insights into sustainable biogas purification practices, contributing to the advancement of energy efficiency and environmental sustainability in the palm oil industry.

MATERIALS AND METHODS

This study investigated the removal of H_2S from raw biogas using a bioscrubber at a POME treatment plant. A simulation of the POME treatment process and biogas purification at an existing plant in Johor was conducted to assess the feasibility of a bioscrubber. Furthermore, an optimisation approach using artificial neural networks (ANN) was applied to analyse the biogas purification process. The research framework outlined in Fig. S1 was developed based on the actual plant data.

Data collection

After identifying the problem and setting the study objectives, the next crucial step is data collection. Most of the data used in this research was obtained through collaboration with a palm oil mill in Johor, Malaysia. Input parameters, such as POME characteristics including COD, total solids (TS), and total volatile solids (TVS), were collected from the mill on a monthly basis. Additionally, data on temperature, biogas flow rate, H_2S concentrations, and biogas composition were collected from their supervisory control and data acquisition (SCADA) system. The data collection for the machine learning model was conducted over a period of two years, which provides detailed operational data for accurate model development and validation.

Data pre-processing

A total of 90 industrial datasets were collected from the palm oil mill in Johor. To generate an effective model for performance analysis, it is essential to pre-process the data to have small variations of the output data. Theoretically, output data usually display a wide range of data. However, if the output domain is too large, the ANN model could appear to be less stable and reliable (*26*). The pre-processing of the data was carried out by normalisation technique whereby the data was scaled to fit within a range of 0-1 using the following equation (*27*):

$$x_{norm} = \frac{x - x_{min}}{x_{max} - x_{min}} \quad /1/$$

where x is the variable, x_{max} is the maximum value and x_{min} is the minimum value.

Overfitting is a common issue that arises during the training of neural networks. To address this problem, the data was divided into three segments: training, testing and validation sets. With this approach, the validation error is constantly monitored and if there is any indication of an increase in it, the training is stopped (*27*). It was configured where 70 % of the datasets were tasked for training the model. Meanwhile, 15 % was set for validation and testing, which proves the accuracy of the machine learning model. The software used to run this ANN model was MATLAB R2022a (*28*). The neural fitting network in the software was the programming language application used to generate the ANN model.

Process simulation

The simulation of the POME treatment and biogas purification processes was executed utilising SuperPro Designer v9.0 (*18*). The input parameters and compositions for the simulation were obtained from the palm oil mill situated in Johor, Malaysia. The composition of POME in the feed stream is shown in Table 1. The feed flowrate was obtained from the data collected from the plant. The flow rate of POME from the cooling pond that enters the mixing tank was recorded weekly. The process flow sheet and equipment selection were then simulated using the information provided by the plant. The sizing calculations performed on SuperPro Designer v9.0 (*18*) were completed after mass balance calculations had zero errors. The simulated process flow sheet is presented in Fig. S2.

As shown in Fig. S2, following exiting a cooling pond, the POME feed enters a mixing tank (P-1) at a flow rate of 14,366 kg/h. The operating conditions of the mixing tank were set at a temperature of 57.3 °C and a pH of 6.89. The mixture then undergoes anaerobic digestion in a covered anaerobic lagoon (P-2). The purpose of a covered digester is to contain the produced CH_4. Loss of these components could lead to a significant profit loss due to a lower purity and emission of greenhouse gases. The biogas generated then requires further purification due to undesired components like H_2S. The biogas purification is carried out via H_2S removal in the bioscrubber (P-9).

In this biogas plant, a low-pressure biological scrubber (LPBS) is used for H_2S removal. The decomposition of H_2S occurs through oxidation, producing sulphur and water (*13*).

A LPBS usually consists of one or more gas-liquid contact columns, connected to a recirculation tank comprised of water with nutrients. Each column is partly filled with packing media that is available in a range of shapes (*29*). Packing media are normally built from non-corrosive materials including glass, metal, ceramics or plastic, and microorganisms are immobilised on it. The shape of the packing media plays an important role in reducing gas pressure drop by providing a small contact area per unit volume and increasing the void fraction for gas passage (*29*). Moreover, the packing media allows microorganisms to form a biofilm (*30*). The biogas stream is routed through the media, while the nutrient water is sprayed down over it. The nutrient water is sourced from the settling pond post-AD, containing minimal solids (below 1 %). Contaminants like H_2S is absorbed from the raw biogas into the liquid in which the microorganisms are growing on the media, where biological oxidation occurs. Normally, the chosen bacteria are Thiobacillus genus for the removal of H_2S from biogas.

Next, the moisture content of this product stream is reduced using a chiller (P-10) (*31*). Meanwhile, the digestate from the anaerobic digester enters a settling pond (P-4) for stabilisation of wastewater sludge before the bottom sludge is recirculated to the original mixing tank. The treated effluent then enters aerobic treatment through a facultative pond (P-6) followed by an algae pond (P-7) for further breakdown of biodegradable matters. The pond effluent then enters a discharge pond (P-17) for further wastewater separation from suspended particles due to settlement. The sludge is then passed to a dumping pond and is sent to a composting plant (*32*).

Neural net fitting

Machine learning (ML) is a discipline of artificial intelligence (AI) that trains machines or software to identify patterns and make predictions using historical data from past experiences with minimal human supervision. Over the years, the advancements of ML in the industry have shown great potential as a prediction tool in many studies. In the POME industry, it can be observed that the integration of ML in recent years is becoming more prevalent for quality estimation (*33*).

ANN consists of interconnected nodes or neurons, linked by weights. Each node in parallel layers receives data from preceding nodes, processes it through a nonlinear function and transmits the result to subsequent nodes. The modelling process of an ANN involves several steps. The data in this study were recorded weekly and compiled each month from July 2021 to June 2023. A total of 90 datasets were used in the development of the machine learning algorithm to obtain the prediction model. With the MATLAB (*28*) settings for neural net fitting of 70 % to validation, 15 % each for training and testing, the training datasets were used to train the ANN model for future prediction.

Hyperparameters define the structure and learning process of a neural network and play a critical role in optimizing the performance of an ANN model. Unlike model parameters, which are learned during training, hyperparameters such as learning rate, the number of hidden layers, and the number of neurons must be determined prior to the training process. In this study, random search was employed for hyperparameter tuning, as it is more efficient and practical in identifying optimal hyperparameter combinations compared to grid search. This approach helps to prevent overfitting or underfitting and ensures that the model generalizes well to unseen data, leading to improved prediction accuracy.

Parameters affecting the bioscrubber's performance, such as the type of packing media, bacteria species, empty bed retention time, and recirculation ratio of liquid media (*34*), play a crucial role. Additionally, other factors including pressure, biogas flow rate, operating pH, and the shape of the packing media can influence the efficiency of the bioscrubber.

The inlet temperature and biogas flow rate were chosen as input data for the model because these parameters are routinely monitored and recorded by operators in biogas plants daily. Specifically, inlet temperature can influence microbial activity within the bioscrubber, while biogas flow rate affects the contact time between biogas and liquid media, impacting H_2S removal efficiency (*35*). The outlet H_2S concentration was chosen as the output data and integrated into the output layer of the model. This is because it serves as a direct metric of bioscrubber performance, reflecting the effectiveness of H_2S removal in the biogas stream. This concentration is a critical parameter for compliance with environmental regulations and is essential for ensuring the biogas meets quality standards for downstream applications.

In this study, statistical analysis of mean squared error (MSE) is used as an indicator of the model’s fit. MSE is determined during the neural network programming using the following equation:

$$MSE = \frac{1}{n}\sum\nolimits_{i=1}^{n}(o_i - p_i)^2 \quad /2/$$

While root mean squared error (RMSE) is evaluated in the same unit, making it a more interpretable measure, MSE is more effective in assessing the distribution of data and overall measure of accuracy. This is because outliers have a greater impact, as all the differences are squared and all errors are positive. Hence, minimising the effect of negative and positive differences cancelling each other out.

The coefficient of determination (R^2), is computed during the ANN modelling using the following equation:

$$R^2 = \frac{\sum_{i=1}^{n}(p_i - o_i)^2(p_i - \bar{p})}{\sum_{i=1}^{n}(o_i - \bar{o})^2\sum_{i=1}^{n}(p_1 - \bar{p})^2} \quad /3/$$

The model fit improves as R approaches 1.

Economic evaluation

Techno-economic analysis (TEA) has emerged as a crucial tool for assessing the economic potential of industrial processes, particularly in the context of the Industrial Revolution (IR) 4.0. As industries evolve, TEA has increasingly incorporated data-driven technologies such as artificial intelligence (AI) and blockchain to optimise processes and economic factors (*36*). This study combines process simulation with TEA to assess the economic feasibility of the biogas plant using the bioscrubber. The analysis involves stages such as process design, equipment sizing, cost estimation, and cash flow analysis. Data on plant operation, design, transportation, and market behaviour were collected from the biogas plant. Market behaviour refers to the trends in pricing and demand for biogas and its by-products, including the feed-in tariff (FiT) policies in Malaysia, which influence the economic viability of the project. Following successful simulation and equipment sizing, economic calculations were conducted using updated prices from 2022. The economic evaluation feature of the SuperPro software (*18*) facilitated the generation of a comprehensive economic evaluation report.

RESULTS AND DISCUSSION

Simulation model

The simulated results including the composition of biogas from AD and its purified form are tabulated in

Table . The simulation results show the final H_2S concentrations in the biogas product to be 43 ppm and the H_2S removal efficiency by the bioscrubber was calculated to be approximately 95.5 %. This indicates that the operational conditions of the current palm oil mill are satisfactory, as the concentration of H_2S in the biogas remains well below the recommended threshold of 500 ppm. Overall, the percentage difference of most parameters in Table 2 is less than 10 %, indicating that the simulation model accurately predicts the performance of the bioscrubber system for these parameters.

Results of artificial neural network (ANN) model

ANN adopts a multi-layer feed-forward neural network using the Levenberg-Marquardt training algorithm. The concept of feed-forward is to take measurements and corrective actions before the process is disturbed. The settings and properties of the prediction model are shown in Table 3. The four graphs in Fig. 1 display the regression analysis of the neural network model with desirable outcomes. This is observed with the values of R^2 of close to 0.9. This suggests that the neural network model establishes strong linearity with the target values. It is highly potential to be used in the palm oil industry for prediction purposes that can greatly enhance the processes.

The graph in Fig. 2a illustrates the MSE for the training, validation and testing of the generated ANN model with respect to the epoch. An epoch in machine learning can be defined as the total number of iterations of all the training data used in one cycle through the algorithm to train the model. The epoch number is a crucial hyperparameter for the machine learning algorithm. Typical values for the number of epochs can range from 10 to 1000 and can keep continuing even larger until the model error is sufficiently minimised (*37*). After running the neural network program several times, the results with the best R^2 were used. According to the results of the ANN model, the best validation performance was 0.0067655 which was observed at epoch 11. This means that during the training process, the neural network achieved its optimal performance on the validation dataset at the 11th iteration of training. After epoch 11, the performance of the model on the validation data might not improve further, leading to overfitting. The ideal ANN model is then developed by focusing on the MSE performance function and the coefficient of determination (R^2). The MSE serves as an estimate for the average squared difference between the predicted and actual output values. A lower MSE value signifies smaller errors in predicting the actual outcomes of the dataset.

Fig. 2b depicts an error histogram bar chart using 20 bins from the ANN model, which illustrates the quantitative difference between the targeted values and the output values following the development of the ANN model. In an error histogram, bins are used to group these differences (errors) into specific ranges, represented as vertical bars. Each bin shows how many predictions fall within a particular error range, allowing for a clearer view of the model's overall accuracy. In this histogram, the model has a notable bin with an error value of -0.00765, containing around 20 instances where predictions fall within this range. This error value is close to zero, indicating that the model is making predictions with a small margin of error, which suggests a relatively high accuracy. Additionally, the zero-error line on the x-axis represents a perfect match between predicted and target values, highlighting the distribution of errors around this ideal point.

Overall, the findings of this study highlight the critical importance of integrating both ANN and SuperPro (*18*) simulation for effective biogas plant management due to their complementary strengths. ANN plays a vital role in accurately estimating specific parameters such as H_2S removal efficiency, using historical data and complex algorithms based on just two input variables, *i.e.* inlet temperature and biogas flow rate. In terms of performance, the ANN model in this study demonstrated high predictive accuracy, further supported by the comparison with the work of Tan *et al.* (*38*), who reported an R^2 value approaching 1 and a MSE of 0.0002 for H_2S concentration prediction. This capability is particularly valuable for optimising process performance and ensuring compliance with gas engine specifications, such as H_2S levels in biogas for combustion.

On the other hand, SuperPro (*18*) simulation extends the predictive power by estimating parameters that are not monitored daily in the plant, such as COD, BOD, TSS, methane content, methane yield, *etc*. and other operational parameters including HRT, OLR, and recirculation flow rate. This broader scope allows for a comprehensive understanding of overall process efficiency, environmental impact, and operational optimisation. By integrating both ANN and SuperPro (*18*), these tools provide a robust framework for enhancing biogas plant management, enabling biogas plant engineers to anticipate performance variations, address potential bottlenecks, and implement proactive measures to ensure stable and efficient bioscrubber operation. This integrated approach not only improves operational decision-making but also contributes to the long-term economic viability of the palm oil industry. This aspect will be further evaluated in the next section.

Techno-economic analysis (TEA)

Techno-economic analysis (TEA) is a method to assess the economic potential of processes in industries (*36*). The economic analysis for this POME treatment plant was carried out using the economic evaluation feature on the SuperPro Designer v9.0 software (*18*). The revenues within the processing plant are mainly generated from the sales of the treated biogas, dried sludge and digested sludge to the composting plant. The reliability of this economic evaluation is up to date, as this report uses prices from 2022. Table 4 shows the brief economic evaluation report generated from the simulation model. According to the sizing calculations performed with the software, the bioscrubber has a volume of 99.39 m^3. Equipment of this size is estimated to have a unit cost of $ 250,000. The payback period is 7 years, which is aligned with those reported by Loh *et al.* (*39*). This shows that the bioscrubber is economically feasible and offers a reasonable return on investment.

In contrast, Tan *et al.* (*38*) reported a shorter payback period of 5.34 years for a local palm oil mill in Pahang, Malaysia. This reduced payback period was achieved following an optimisation study, highlighting the critical role of optimisation in enhancing economic viability. This comparison highlights the potential for improving the bioscrubber's effectiveness, particularly in enhancing H_2S removal efficiency. Additionally, optimising the anaerobic digester's conditions can lead to the production of biogas with lower H_2S concentrations. By focusing on parameters such as feedstock composition and operational conditions, future research can enhance both H_2S management and overall biogas quality. Consequently, conducting optimisation studies will be essential for identifying and implementing strategies that minimise costs while maximising biogas production, ultimately contributing to more sustainable practices in the palm oil industry.

Furthermore, excessive H_2S can reduce equipment lifespan and significantly increasing maintenance expenses (*35*). Implementing an optimised bioscrubber with both ANN and SuperPro (*18*) will mitigate these effects by lowering H_2S levels before the biogas enters critical equipment. This approach would reduce repair costs and extend the lifespan of the biogas plant’s infrastructure, promoting more efficient and cost-effective operations. By improving H_2S management and optimising biogas quality, the integrated strategy ensures the economic benefits of biogas production are fully realised.

Limitations and future works

While this study provides valuable insights into the integration of machine learning and SuperPro (*18*) simulation for biogas purification, it is important to acknowledge its limitations. A key limitation is the reliance on a dataset of 90 records, collected from a single palm oil mill, which may not capture the full range of variability in POME characteristics across different facilities. Although the dataset was gathered under varying operating conditions, its relatively small size may limit the generalizability of the model. To mitigate this, future studies could incorporate techniques such as k-fold cross-validation, which would help validate model reliability across different data subsets and reduce overfitting.

The current simulation was limited to certain controllable parameters (*e.g.* inlet temperature, biogas flow rate), assuming that other operational aspects, like the biofilm stability or packing media configuration, remain consistent. This exclusion of certain bio-physical interactions and operational variances may underrepresent how complex, interdependent factors influence bioscrubber efficiency. In future work, incorporating a broader array of parameters or conducting a sensitivity analysis could help identify the primary variables that impact H_2S removal efficiency, better aligning predictions with real-world conditions.

The performance of both the machine learning model and the SuperPro (*18*) simulation may also be affected by data quality, plant-specific operational conditions, and the limited range of parameters included in this analysis. Future research should consider expanding the dataset to include data from multiple palm oil mills to enhance robustness and the potential to generalize findings. Additionally, conducting a feature importance analysis would help clarify the influence of individual parameters on model predictions, providing a clearer understanding of factors affecting the integrated system's outputs.

To further enhance prediction accuracy, future research could investigate the application of other machine learning models, such as support vector machines or deep learning techniques. Long-term performance assessments across varying seasonal or operational conditions would also be beneficial to evaluate the integrated approach's viability in real-world settings.

CONCLUSIONS

Simulation and prediction of the bioscrubber performance to treat biogas are crucial for meeting the quality standards required for internal combustion gas engines used in electricity generation. This study presents a novel integration of machine learning and process simulation in POME-based biogas plants, offering a data-driven approach to optimize biogas purification and enhance operational efficiency. Given the non-linear nature of many output responses in these processes, selecting an appropriate machine learning algorithm to model the prediction of a unit is crucial. Consequently, ANN was employed to investigate the performance of the bioscrubber in removing H_2S from biogas. Statistical parameters, including the coefficient of determination (R^2) and mean squared error (MSE), were used to evaluate the ANN model's accuracy. The model achieved an R^2 of 0.90 and an MSE of 0.0068, confirming its high predictive accuracy and reliability without the need for further correlations or experimental measures. The simulation of the POME treatment process and biogas purification using the bioscrubber was conducted using SuperPro. Several parameters were observed and compared with industry data, showing deviations of less than 10 %. These results demonstrated the high accuracy and prospects of employing machine learning tools coupled with SuperPro simulation to predict the performance of bioscrubbers in the palm oil industry. Techno-economic analysis also confirms the bioscrubber's economic feasibility, offering a reasonable return on investment with a 7-year payback period. These findings provide a scalable framework for improving biogas treatment processes, reducing reliance on trial-and-error approaches, and supporting the broader adoption of AI-driven decision-making in industrial wastewater treatment. This integrated approach also enhances operational decision-making and supports the long-term economic sustainability of the palm oil sector. Future research should investigate the application of other machine learning models, such as support vector machines or deep learning techniques, to further enhance prediction accuracy.

ACKNOWLEDGEMENTS

The authors acknowledge the support provided by the University of Nottingham Malaysia in this research paper.

CONFLICT OF INTEREST

There is no conflict of interest in this research.

AUTHORS' CONTRIBUTION

Joanna Lisa Clifford conceived and designed the work. Mohd Amran Bin Mohd Yusof conducted the data collection. Joanna Lisa Clifford carried out the data analysis and interpretation, who also performed the analysis and drafted the article. Yi Jing Chan and Timm Joyce Tiong supervised the work. Yi Jing Chan, Timm Joyce Tiong, Siew Shee Lim, and Chai Siah Lee undertook the critical revision. Woei-Yenn Tong gave the final approval of the version to be published.

SUPPLEMENTARY MATERIALS

Supplementary materials are available at www.ftb.com.hr.

ORCID ID

Y.J. Chan https://orcid.org/0000-0003-2123-2356

T.J. Tiong https://orcid.org/0000-0003-1137-3606

S.S. Lim https://orcid.org/0000-0002-1798-2683

C.S. Lee https://orcid.org/0000-0002-7688-4469

REFERENCES

1. Akhbari A, Kutty *PK.*, Chuen *OC.*, Ibrahim *S.* A study of palm oil mill processing and environmental assessment of palm oil mill effluent treatment. Environmental Engineering Research. 2019; 25(2):212–21.
https://doi.org/10.4491/eer.2018.452

2. Chew A. Malaysia palm oil production, exports expected to rise this year - Nikkei Asia. Accessed: Dec. 06, 2023. [Online]. https://asia.nikkei.com/Business/Markets/Commodities/Malaysia-palm-oil-production-exports-expected-to-rise-this-year.

3. Kamyab H, Lee SC, Shahabuddin S, Hashim H, Luu DH. Palm Oil Mill Effluent as an Environmental Pollutant. Palm Oil. Nov. 2018.
https://doi.org/10.5772/INTECHOPEN.75811

4. Poh PE, Yong WJ, Chong MF. Palm oil mill effluent (POME) characteristic in high crop season and the applicability of high-rate anaerobic bioreactors for the treatment of pome. Ind Eng Chem Res. 2010;49(22):11732–40.
https://doi.org/10.1021/IE101486W

5. Chen W-Y, Ho CC, Tan HT, Lee KL, Ling HT, Li H. Artificial neural network (ANN) modelling for biogas production in pre-commercialized integrated anaerobic-aerobic bioreactors (IAAB). Water (Basel). 2022;14(9):1410.
https://doi.org/10.3390/w14091410

6. Chong DJS, Chan YJ, Arumugasamy SK, Yazdi SK, Lim JW. Optimisation and performance evaluation of response surface methodology (RSM), artificial neural network (ANN) and adaptive neuro-fuzzy inference system (ANFIS) in the prediction of biogas production from palm oil mill effluent (POME). Energy. 2023;266:126449.
https://doi.org/10.1016/J.ENERGY.2022.126449

7. Chan YJ, Lee HW, Selvarajoo A. Comparative study of the synergistic effect of decanter cake (DC) and empty fruit bunch (EFB) as the co-substrates in the anaerobic co-digestion (ACD) of palm oil mill effluent (POME). Environ Challenges. 2021;5:100257.
https://doi.org/10.1016/j.envc.2021.100257

8. Muhammad RAM, Torii S. Removal of hydrogen sulfide (H_2S) from biogas using zero-valent iron. J Clean Energy Technol. 2015;3(6):511-5.
https://doi.org/10.7763/JOCET.2015.V3.236

9. Poh PE, Chong MF. Development of anaerobic digestion methods for palm oil mill effluent (POME) treatment. Bioresour Technol. 2009;100(1):1–9.

https://doi.org/10.1016/J.BIORTECH.2008.06.022

10. Bušić A, Kundas S, Zelić B, Krika T, Novak S, Komes D, Babić J, Jokić S, Aladić K, Redovniković IR. Recent trends in biodiesel and biogas production. Food Technol Biotechnol. 2018;56(2):156-63. https://doi.org/10.17113/ftb.56.02.18.5547

11. Zupančič GD, Panjičko M, Zelić B. Biogas production from brewery yeast using an anaerobic sequencing batch reactor (ASBR). Food Technol Biotechnol. 2017;55(2):200-9. https://doi.org/10.17113/ftb.55.02.17.5080

12. Andlar M, Krajacic M, Nestic I, Presecki AV, Šantek B, Novak S, Babić J, Aladić K, Žutinić P, Komes D. Biogas production systems and upgrading technologies: a review. Food Technol Biotechnol. 2021;59(4):387–412. https://doi.org/10.17113/ftb.59.04.21.7300

13. Huynh Nhut H, Le Thi Thanh V, Tran Le L. Removal of h2s in biogas using biotrickling filter: recent development. Process Saf Environ Prot. 2020;144:297–309. https://doi.org/10.1016/J.PSEP.2020.07.011

14. Khoshnevisan B, Tsapekos P, Alvarado-Morales M, Angelidaki I, Tabatabaei M. A review on prospects and challenges of biological H2S removal from biogas with focus on biotrickling filtration and microaerobic desulfurization. Biofuel Res J. 2017;16:741–750. https://doi.org/10.18331/BRJ2016.4.4.6

15. San-Valero P, Penya-roja JM, Álvarez-Hornos FJ, Buitrón G, Gabaldón C, Quijano G. Fully aerobic bioscrubber for the desulfurization of H2S-rich biogas. Fuel. 2019;241:884–891. https://doi.org/10.1016/j.fuel.2018.12.098

16. Yap CC, Ho WS, Yeo HT, Ling HT, Chua YW, Bong CP, Klemeš JJ. Pilot-scale investigation of the integrated anaerobic–aerobic bioreactor (IAAB) treating palm oil mill effluent (POME): Startup and Performance Evaluation. Ind Eng Chem Res. 2021;60(10):3839–3859. https://doi.org/10.1021/acs.iecr.0c05878

17. Yap AC, Hasanah MNUK. Computer aided simulation pome biogas purification system. J Kejuruteraan. 2021;33(2):293–316.

18. SuperPro Designer, v.11, Intelligen, Inc., NJ, USA; 2021. Available from: https://www.intelligen.com/

19. Lok X, Chan YJ, Foo DCY. Simulation and optimisation of full-scale palm oil mill effluent (POME) treatment plant with biogas production. Journal of Water Process Engineering. 2020;38:101558. https://doi.org/10.1016/j.jwpe.2020.101558

20. Chong JWR, Chan YJ, Chong S, Ho YC, Mohamad M, Tan WN, *et al.* Simulation and optimisation of integrated anaerobic-aerobic bioreactor (IAAB) for the treatment of palm oil mill effluent. Processes. 2021; 1;9(7).
https://doi: 10.3390/pr9071124

21. Kan KW, Chan YJ, Tiong TJ, Lim JW. Maximizing biogas yield from palm oil mill effluent (POME) through advanced simulation and optimisation techniques on an industrial scale. Chem Eng Sci. 2024;285:119644.
https://doi.org/10.1016/j.ces.2023.119644.

22. Witten IH, Frank E, Hall MA. Data mining: practical machine learning tools and techniques, 3rd ed. Morgan Kaufmann; 2011.

23. Barik D, Murugan S. An artificial neural network and genetic algorithm optimized model for biogas production from co-digestion of seed cake of karanja and cattle dung. Waste Biomass Valor. 2015;6(6):1015–1027.
https://doi.org/10.1007/S12649-015-9392-1

24. Mata-Alvarez J, Dosta J, Romero-Güiza MS, Fonoll X, Peces M, Astals S. A critical review on anaerobic co-digestion achievements between 2010 and 2013. Renew Sustain Energy Rev. 2014;36:412–427.
https://doi.org/10.1016/J.RSER.2014.04.039

25. Manan W, Abdullah NR. An artificial neural network model for forecasting air pollution. In: IOP Conference Series: Materials Science and Engineering, IOP Publishing; 2021, p. 12032.

26. Sunil A, Aneesh K. Modelling studies by application of artificial neural network using matlab. Journal of Engineering Science and Technology. 2015;10(11):1477-86. Available from:
https://www.researchgate.net/publication/284921018_Modelling_studies_by_application_of_artificial_neural_network_using_matlab

27. Sulaiman NS, Yusof KM. Artificial neural network-based model for quality estimation of refined palm oil. 2015 15th International Conference on Control, Automation and Systems (ICCAS). 2015b.
http://dx.doi.org/10.1109/ICCAS.2015.7364843

28 MATLAB, v. R2023b, The MathWorks, Inc., Natick, MA, USA; 2023. Available from: https://www.mathworks.com/products/matlab/.

29 Reutlinger LE. Design and operation of a biological fixed-film scrubber for removal of hydrogen sulfide gas. State University of New York College of Environmental Science and Forestry; 1994.

30. Ottengraf SPP, Meesters JJP, Van Den Oever AHC, Rozema HR. Biological elimination of volatile xenobiotic compounds in biofilters. Bioprocess Eng. 1986;1(2):61–69. https://doi.org/10.1007/BF00387497

31. Joyce J, Sorensen H. Bioscrubber design. Water Environ Technol. 1999;11(2):37–43.

32 Tan VWG, Chan YJ, Arumugasamy SK, Lim JW. Optimizing biogas production from palm oil mill effluent utilizing integrated machine learning and response surface methodology framework. J Clean Prod. 2023;414:137575. https://doi.org/10.1016/J.JCLEPRO.2023.137575

33. Sulaiman NS, Yusof KM. Artificial neural network-based model for quality estimation of refined palm oil. 2015 15th International Conference on Control, Automation and Systems (ICCAS); 2015.

34. Koe LCC, Yang F. A bioscrubber for hydrogen sulphide removal. Water Sci Technol. 2000;41(6):141–145. https://doi.org/10.2166/wst.2000.0103

35. Mutegoa E, Sahini MG. Approaches to mitigation of hydrogen sulfide during anaerobic digestion process – A review. Heliyon [Internet]. 2023 Sep 1;9(9). https://doi.org/10.1016/j.heliyon.2023.e19768

36. Chai SYW, Phang FJF, Yeo LS, Ngu LH, How BS. Future era of techno-economic analysis: Insights from review. Front Sustain. 2022;3:924047. https://doi.org/10.3389/FRSUS.2022.924047/BIBTEX

37. UNext. What is Epoch in Machine Learning? [Internet]. 2022 [cited 2023 Dec 10]. https://u-next.com/blogs/machine-learning/epoch-in-machine-learning/

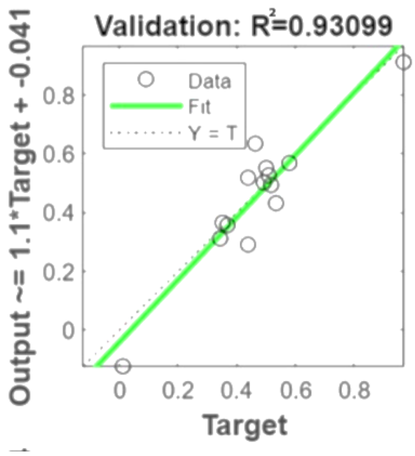
38. Tan VWG, Chan YJ, Arumugasamy SK, Lim JW. Optimizing biogas production from palm oil mill effluent utilizing integrated machine learning and response surface methodology framework. J Clean Prod, 2023;414:137575. https://doi.org/10.1016/j.jclepro.2023.137575

39. Loh SK, Nasrin AB, Mohamad Azri S, Nurul Adela B, Muzzammil N, Daryl Jay T, *et al.* First Report on Malaysia's experiences and development in biogas capture and utilization from palm oil mill effluent under the Economic Transformation Programme: Current and future perspectives. Renewable and Sustainable Energy Reviews [Internet]. 2017;74:1257–74. https://doi.org/10.1016/j.rser.2017.02.066

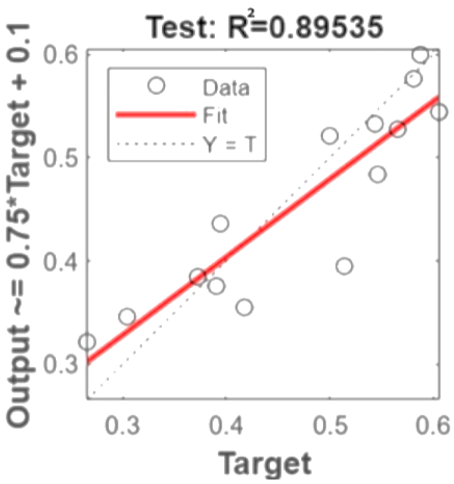
a)


Training: R²=0.8745
Output ~= 0.75*Target + 0.12
Target
Data
Fit
Y = T
0
0.2
0.4
0.6
0.8
1
0.5

b)


Validation: R²=0.93099
Output ~= 1.1*Target + -0.041
Target
Data
Fit
Y = T
0
0.2
0.4
0.6
0.8

c)


Test: R²=0.89535
Output ~= 0.75*Target + 0.1
Target
Data
Fit
Y = T
0.3
0.4
0.5
0.6

d)

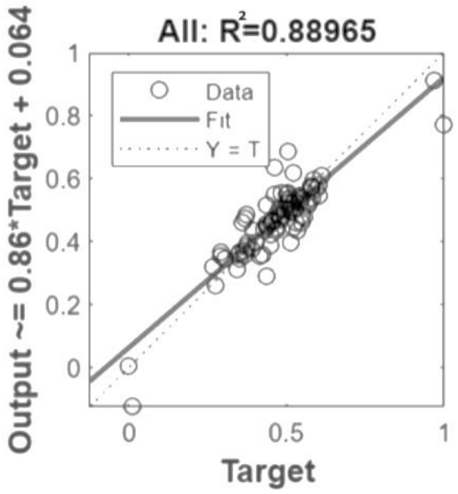
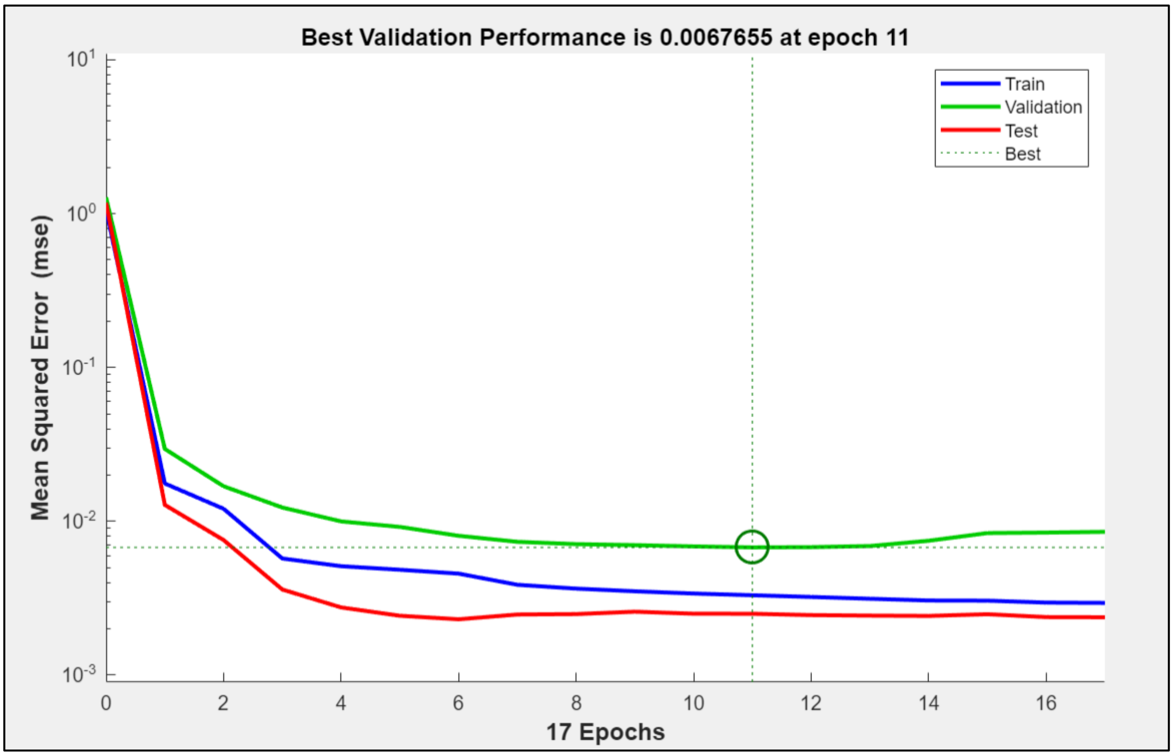


Fig. 1. Regression analysis of the ANN model during: a) training, b) validation, c) testing and d) overall phase

a)



b)

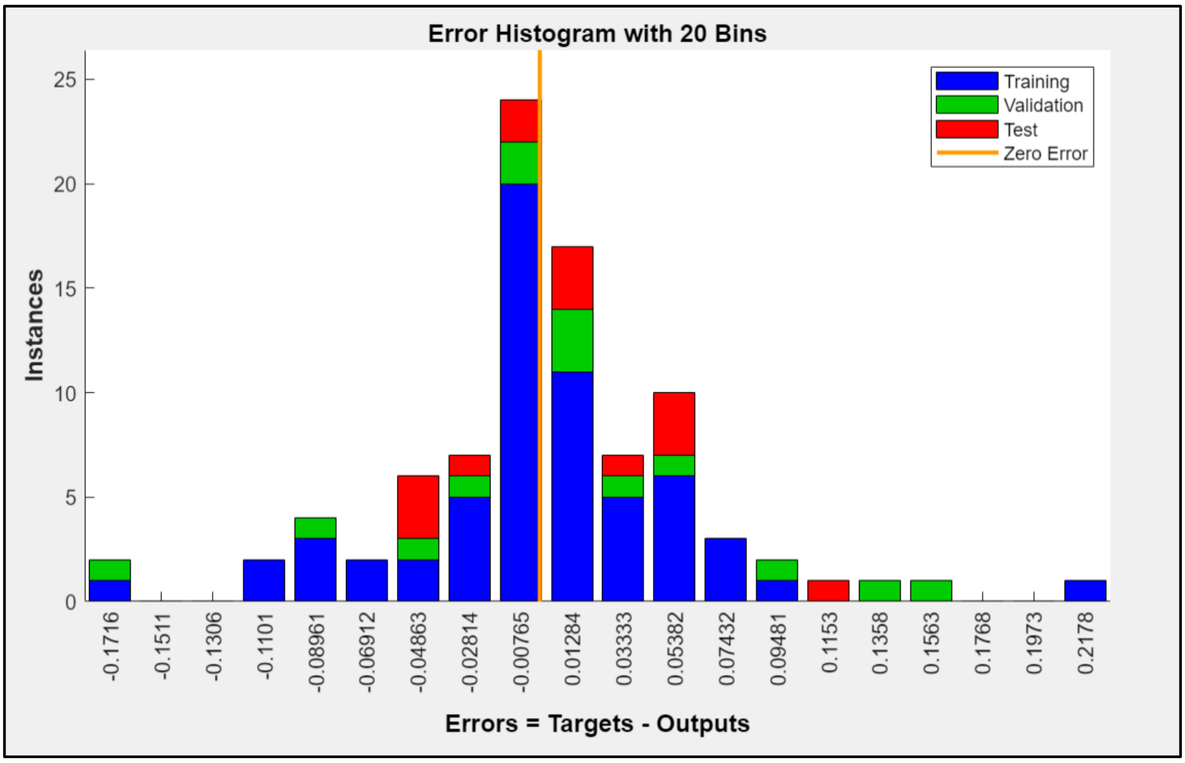


Fig. 2. Artificial neural network (ANN) model: a) performance in terms of mean squared error, b) error histogram with 20 bins

Table 1. Palm oil mill effluent (POME) mass flow rate and feed composition used in the SuperPro (*18*) simulation

Component	Mass flow rate/(kg/h)	Mass composition/%
Biomass	285.376	1.9865
Carbohydrates	285.878	1.9900
Fiber	7.743	0.0539
Nitrate	13.547	0.0943
Oil and Grease	61.040	0.4249
Phosphate	1.839	0.0128
Proteins	61.040	0.4249
Sulfate	1.839	0.0128
Water	13647.457	95.0000

Table 2. Composition of biogas from anaerobic digestion (AD) and purified biogas

Parameter	Compound	Industry data	Simulated data	Relative difference/%
Raw biogas from AD	CH_4	(60.47±0.97) %	62 %	2.5
	H_2S	(943±264) mg/L	957 mg/L	1.5
	CO_2	(34.82±1.09) %	33 %	5.2
	O_2	(0.55±0.11) %	0.51 %	7.3
Purified biogas to gas engine	H_2S	(37±38) mg/L	43 mg/L	16.2
Methane yield	CH_4	(25.68± 6.65) Nm^3 Biogas / m^3 POME	26.12 Nm^3 Biogas / m^3 POME	1.7
Scrubber efficiency (H_2S removal)	H_2S	(96.08±3.86) %	95.51 %	0.6

Table 3. Parameters of neural network model

Setting	Value
Network type	Feed forward
Neurons in input layer	2
Number of hidden layer	2
Neurons in hidden layer	20
Neurons in output layer	1
Transfer function	Tangent Sigmoid
Training function	Levenberg-Marquardt
Performance function	MSE

Table 4. Economic evaluation on the palm oil mill effluent (POME) treatment plant with bioscrubber for the removal of H_2S

Parameter	Value
Total capital investment	$ 9,897,000
Capital investment charged to this project	$ 9,897,000
Operating cost	$ 2,849,000/year
Total revenue	$ 3,734,000/year
Gross margin	23.69 %
Return on investment	14.29 %
Payback time	7.00 years
IRR (after taxes)	7.27 %
NPV (at 7.0 % interest)	$ 131,000

SUPPLEMENTARY METERIAL

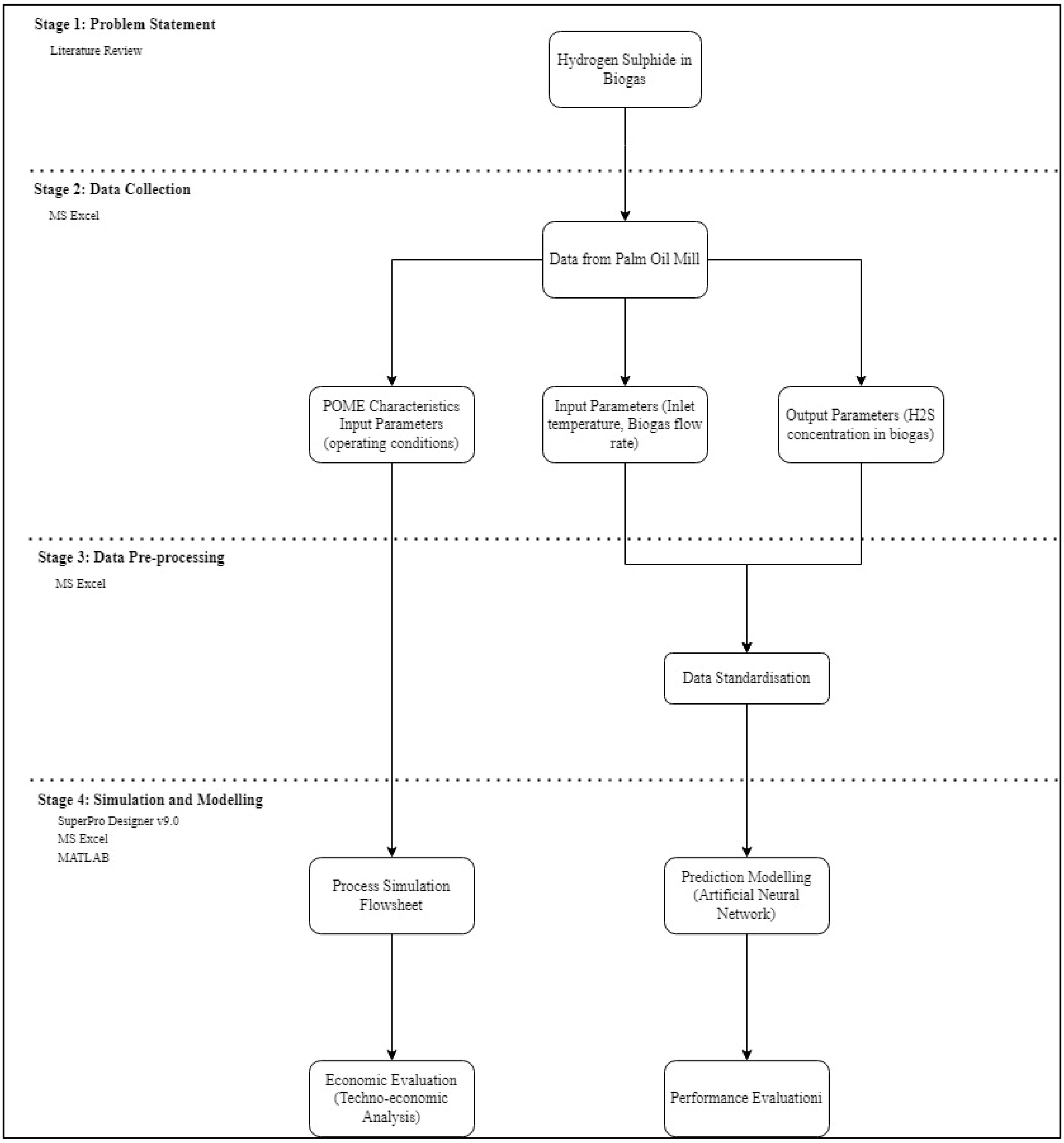


Fig. S1. Framework for simulation and prediction modelling using machine learning and economic analysis

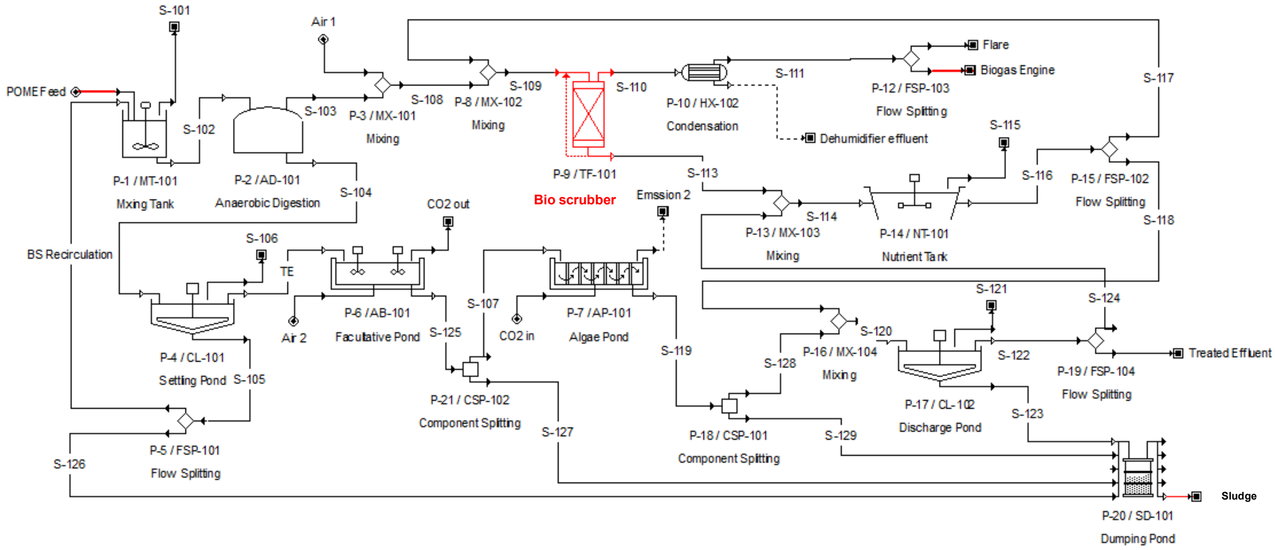


Fig. S2. Simulated process flow sheet for biogas production from palm oil mill effluent (POME)