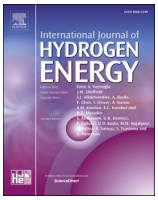
International Journal of Hydrogen Energy 190 (2025) 152153

Contents lists available at ScienceDirect

International Journal of Hydrogen Energy

journal homepage: www.elsevier.com/locate/he



Global sensitivity analysis of blue hydrogen production: a comparative study using machine learning

William George Davies, Paulina Quintanilla, Yang Yang, Salman Masoudi Soltani *

Department of Chemical Engineering, Brunel University of London, Uxbridge, UB8 3PH, United Kingdom

ARTICLE INFO

Keywords:
Hydrogen
Carbon capture
Machine learning
Global sensitivity analysis

ABSTRACT

Data-driven modelling utilising machine learning (ML) techniques offers a powerful alternative to first-principles simulations of chemical processes. In this work, artificial neural networks and random forests were developed as surrogate models, trained on data from a first-principles model of sorption-enhanced steam methane reforming with chemical-looping combustion. These ML-based surrogates were integrated with global sensitivity analysis (GSA) approaches to identify key process drivers and evaluate the comparative performance of different GSA methods in chemical process modelling. The surrogate models achieved an approximately 99 % reduction in computational time compared to first-principles simulations, while maintaining predictive accuracy. Sensitivity analysis demonstrated that the CaO/natural gas (CaO/NG) ratio is a dominant parameter, strongly influencing carbon capture efficiency and hydrogen production performance (cold-gas efficiency and H_2 purity). In-situ CO_2 removal from the reformer was shown to shift equilibrium towards higher hydrogen yields while simultaneously enabling CO_2 capture. Ratios of CaO/NG $\geq$ 1.00 ensured high capture efficiency, while improvements in cold-gas efficiency were observed from ratios $\geq$0.5. Among GSA methods, the Sobol approach delivered high computational efficiency (0.5 s) with first- and second-order sensitivities, whereas Shapley additive explanations provided greater interpretability but at significantly higher computational cost (384 s).

Nomenclature

Air reactor	AR
Artificial neural networks	ANN
Calcium looping	CaL
Calcium oxide/natural gas	CaO/NG
Carbon capture and storage	CCS
design of experiment	DoE
Explainable artificial intelligence	ExAI
Feedforward backpropagation neural network	FBNN
Fuel reactor	FR
Full factorial design	FFD
Global sensitivity analysis	GSA
Local sensitivity analysis	LSA
Machine learning	ML
million tonnes	mT
Natural gas	NG
Physics-informed neural network	PINN
Pressure-swing adsorption	PSA
Principal component analysis	PCA
Random forest	RF
Sensitivity analysis	SA
Shapley additive explanation	SHAP
Sorption-enhanced steam methane reforming	SE-SMR
Sorption-enhanced steam methane reforming with chemical-looping combustion	SE-SMR-CLC
Steam methane reforming	SMR
Steam methane reforming with carbon capture	SMR-CC
Steam/Natural gas	S/NG

(*continued on next column*)

(*continued*)

1. Introduction

1.1. Research background

The development of low-carbon technologies is key to ensuring net zero is reached by 2050 so that the most devasting effects of climate change are mitigated [1]. Hydrogen has been identified as an energy carrier capable of decarbonising a variety of industries such as energy [2], transportation [3] and heat for industrial processes [4]. Although the end use of hydrogen is low-carbon, the production of hydrogen

* Corresponding author.
E-mail address: Salman.MasoudiSoltani@brunel.ac.uk (S. Masoudi Soltani).

https://doi.org/10.1016/j.ijhydene.2025.152153
Received 7 July 2025; Received in revised form 9 October 2025; Accepted 20 October 2025

currently produces large amounts of CO_2. Currently, over 900 million tonnes (Mt) of CO_2 is produced as a result of hydrogen production [5]. In 2022, 95 Mt of hydrogen was produced [5]; however, the majority of that comes from natural gas reforming (i.e. grey hydrogen), coal gasification (i.e. black hydrogen) and as a by-product from refineries and the petrochemical industry [5]. Natural-gas reforming with carbon capture and storage (CCS) (i.e. blue hydrogen) and electrolysis (i.e. green hydrogen) accounted for ~0.7 % of hydrogen produced in 2022 [5]. The distribution between grey, blue and green hydrogen is expected to change drastically by the middle of this century as shown in Fig. 1.

There will be a projected increase of low-carbon hydrogen production by 2030 with 10 Mt of hydrogen being produced via blue hydrogen production routes [5]. Conventionally, blue hydrogen is produced through steam methane reforming (SMR) with carbon capture (CC) technologies (SMR-CC) such as amine scrubbing, as shown in Fig. 2a. Although SMR-CC can capture CO_2 from the process, it has a significant energy penalty and increased cost of hydrogen as a result. Prior works have investigated the intensification of this process by integrating other capture technologies that are able to capture the CO_2 at a lower energy penalty. Sorption-enhanced steam methane reforming (SE-SMR) integrates calcium-looping (CaL) technology with the reformer, as shown in Fig. 2b. Utilising a high-temperature sorbent, such as CaO, allows for in-situ CO_2 capture as H_2 is being produced within the reactor. This process allows for a reduced capital expenditure and produces more hydrogen as a result of Le Chatelier's principle [8]. The reactions for this are shown in R1-4. In SE-SMR, the heat required for the calciner is often supplied via a burner/combustor. CO_2 generated due to combustion can be captured via conventional carbon capture technologies such as amine scrubbing. This however, results in high energy demand due to amine regeneration and the associated process units. Recent advancements have utilised chemical-looping combustion (CLC) to provide the heat for the calciner as shown in Fig. 2c. The benefit of this approach is that the combustion products theoretically comprise only CO_2 and H_2O, significantly reducing the energy penalty of separating out the CO_2 [9]. The reactions utilising iron-oxide with reducing agents is shown in R5-8.

$$CH_{4(g)} + H_2O_{(g)} \leftrightarrow CO_{(g)} + 3H_{2(g)}\ \Delta H_{298K} = 206.2\ \text{kJ.mol}^{-1} \quad \text{R1}$$

$$CO_{(g)} + H_2O_{(g)} \leftrightarrow CO_{2(g)} + H_{2(g)}\ \Delta H_{298K} = -41.2\ \text{kJ.mol}^{-1} \quad \text{R2}$$

$$CH_{4(g)} + 2H_2O_{(g)} \leftrightarrow CO_{2(g)} + 4H_{2(g)}\ \Delta H_{298K} = 165.2\ \text{kJ.mol}^{-1} \quad \text{R3}$$

$$CaO_{(s)} + CO_{2(g)} \leftrightarrow CaCO_{3(s)}\ \Delta H_{298K} = -178.8\ \text{kJ.mol} - 1 \quad \text{R4}$$

$$CH_4 + 12Fe_2O_3 \leftrightarrow 8Fe_3O_4 + CO_2 + 2H_2O\ \Delta H^{O}_{298} = 126.38\ \text{kJ mol}^{-1} \quad \text{R5}$$

$$H_2 + 3Fe_2O_3 \leftrightarrow 2Fe_3O_4 + H_2O\ \Delta H^{O}_{298} = 16.10\ \text{kJ mol}^{-1} \quad \text{R6}$$

$$CO + 3Fe_2O_3 \leftrightarrow 2Fe_3O_4 + CO_2\ \Delta H^{O}_{298} = -25.10\ \text{kJ mol}^{-1} \quad \text{R7}$$

$$4Fe_3O_4 + O_2 \leftrightarrow 6Fe_2O_3\ \Delta H^{o}_{298} = -534.54\ \text{kJ mol}^{-1} \quad \text{R8}$$

Detailed first-principle process modelling and simulation of these blue hydrogen production processes provides a method to analyse these technologies that is economical and reliable [10]. First-principle modelling is built on fundamental 'ab-initio' physio-chemical phenomena such as mass and heat transfer [10]. The use of expert software such as GPROMS and ASPEN allows for processes to be modelled and simulated via a first-principle approach [11]. A first-principle approach provides an accurate and robust method in which the behaviour of processes in steady-state or dynamic mode can be modelled accurately across a vast range of operating parameters. However, the use of ab-initio methods can be time-consuming especially when assessing secondary impacts of the process, e.g. economic or environmental impacts [11].

Machine learning (ML) has provided an alternative method of modelling processes in chemical engineering. For example, artificial neural network (ANN) and random forest (RF) can be trained on data that come from literature (experimental or first-principle modelling) or industry (a hydrogen production plant) [11]. Once this surrogate model is trained, the model can then be employed to predict outputs for a range of input values. This approach has been shown to be able to model processes quickly and accurately although the accuracy does is reduced when predicting outputs outside the range of the training data [12]. The use of ML to develop surrogate models is referred to as a black-box approach or data-driven [10]. Development of an interpretable, data-driven approach that can predict key performance indicators at increased speed is crucial for the chemical engineering discipline. It provides a method to determine how the reactants impact the products; this could potentially be used to make informed decisions in changing certain process parameters to ensure a high purity value product in real-time [13].

Multiple approaches can be used to develop an interpretable data-driven approach. One such method is by the integration of sensitivity analysis (SA) and design of experiment (DoE) methods into ML models in order to assess input parameters and their effect on the output. SA methods aim to measure uncertainty in the output based on the change in the input [14], it can be applied to ML surrogate models in order to provide interpretability to these models [14]. SA methods can be divided into two types: local sensitivity analysis (LSA) and global sensitivity analysis (GSA). Local-sensitivity analysis (LSA) looks at local variation within a local point, hence it is often applied to a single

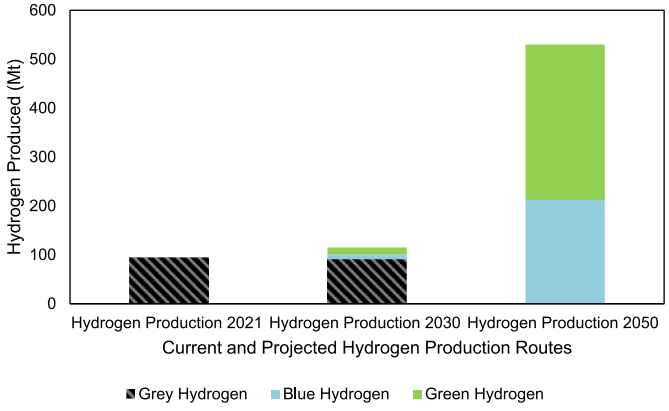


Fig. 1. Current and projected global hydrogen production (Data extracted from Refs. [6,7]).

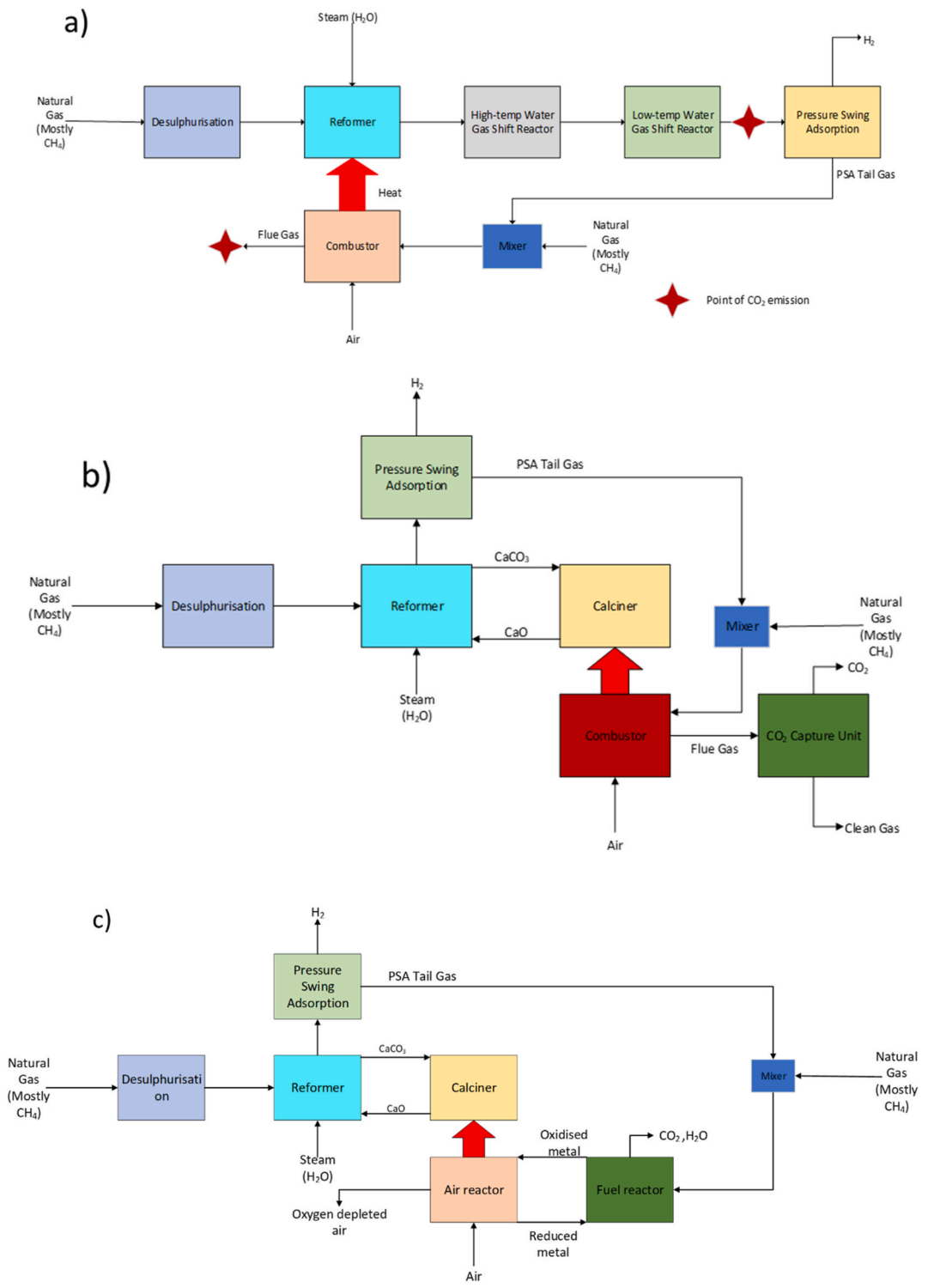


Fig. 2. Fig. 2a) Block-flow diagram of SMR with potential CO_2 capture points. Fig. 2b) SE-SMR block-flow diagram. Fig. 2c) SE-SMR-CLC block-flow diagram.

prediction and a single input. Consequently, it is inadequate to explore large design spaces [15]. However, GSA assesses the variation of all inputs onto the variability of model outputs. This provides an overall analysis of the importance of each feature and provides a way to explain a model in its entirety [15]. SA can be implemented via four distinct methods: variance-based methods, derivative-based methods, density-based methods and model-based methods. Section 2 provides an in-depth description of each approach used in the study.

1.2. Literature review

ML has already been employed in the surrogate modelling of blue hydrogen production processes. Comprehensive reviews of prior works can be found in the literature [11,16]. Nkulikiyinka *et al* (2020) [17] developed soft-sensors for the reformer and calciner utilising both ANNs and RF. The surrogate models developed were able to predict the output of both reactors with high accuracy, with average R^2 values of 0.98–1 for both RF and ANN. They utilised the feature importance for RF and a principal component analysis (PCA) for ANNs to determine the impact of

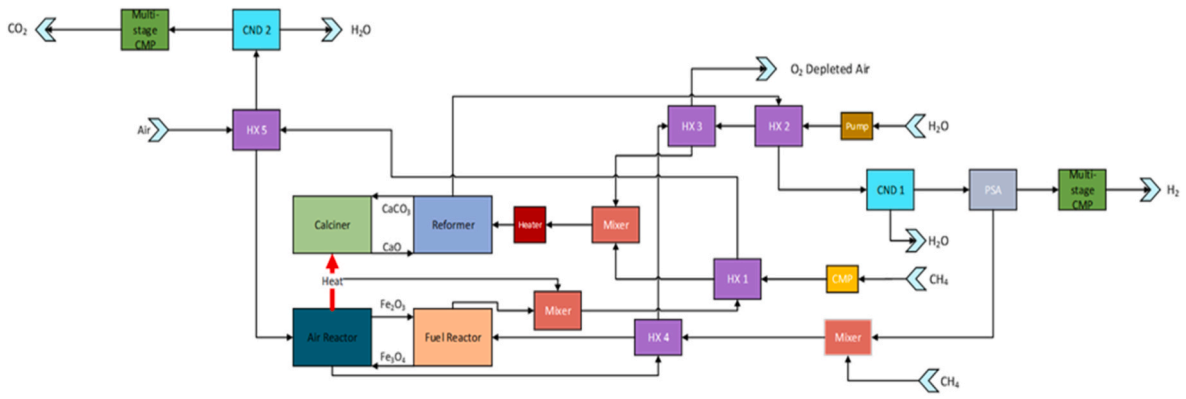


Fig. 3. Simple block-flow diagram of sorption-enhanced steam methane reforming process with chemical-looping combustion, highlighting the inputs that were manipulated in this work.

the inputs on the outputs. Vo *et al* (2022) [18] developed a ML-based optimisation for the SE-SMR process. They utilised a Pearson correlation coefficient to assess the impact of input variables on the outputs, where the optimisation approach was able to produce hydrogen at low-cost ($1.70 per kg of H_2) whilst maintaining a high CO_2 capture rate (90.3 %). The use of screening approaches such as PCA is useful as it allows to discriminate influential inputs from non-influential ones. Careful consideration must be taken when employing this approach since this method mainly aims to reduce the dimensions of the input space by transforming and combining the inputs into a new set of variables that aims to capture the largest variance [19]. This transformation means that the relative contribution of the original dataset to the outputs is lost and requires further steps to analyse this feature's importance.

Zhao *et al* (2021) developed a surrogate model for water gasification of biomass for hydrogen production. They compared different ML approaches to determine the best approach in predicting hydrogen yield [20]. Evaluation of the model variables via feature importance and partial dependence analysis, determined that a high oxygen content in biomass leads to maximum hydrogen reaction efficiency and exergy efficiency. Recent work by Huang *et al* (2025) developed a ML model of a diesel autothermal reforming process by varying certain operating parameters. In this work, a H_2 purity of 68.79 % was achieved [21]. The authors used the Shapley Additive Explanation (SHAP) tool to provide insight into how the operating parameters influence each output [21]. Within these works, the focus is often on the optimisation speed or the accuracy of the model. There is much less effort made on investigating data quantity required to enable the development of an accurate robust interpretable model. This work, therefore, also aims to evaluate what is required from datasets within the investigated area in order to develop an interpretable computational efficient approach to ML-based modelling.

1.3. Objectives and paper motivation

Developing explainable artificial intelligence (ExAI) has become of increasing importance, especially within science and engineering, where the model must be subject to the underlying physical laws. Within data-driven modelling, machine learning-based surrogate models, incorporating global sensitivity analysis (GSA), can provide context to the predictions made by the surrogate model. Incorporating GSA into ML surrogate models has recently been exercised. Stein et al. (2022) looked at assessing different GSA approaches. They determined that Morris approach was extremely robust, performing well in large dimensions [15]. Lucay (2022) implemented a surrogate model combined with GSA approach in a case study of mineral processing [22]. They found that by implementing the surrogate model with GSA, they were able to significantly increase the speed at which they were able to predict the outputs, while still quantifying the influence of the inputs on each output [22]. Zhao et al. (2021) implemented active learning and sensitivity analysis with the ML surrogate model to improve the computational efficiency of a non-linear system a chemical process system [23]. Here, active learning allowed for the model to be further generalised, and the surrogate model developed based on the sensitivity analysis, significantly improved the computational efficiency whilst maintaining high accuracy. Objectives and paper motivation.

The literature review highlights how ML has currently been used within the modelling of blue hydrogen production, serving as soft sensors to predict reactor outputs and used as a surrogate model to optimising blue-hydrogen production in regard to cost and CO_2 capture rate. However, in existing literature, ML is often employed to accelerate optimisation rather than improve interpretability. In process engineering, achieving both computational efficiency and interpretability is crucial when integrating GSA as demonstrated by Zhao *et al* (2021) [23]. The aim of this work is to integrate different GSA and LSA approaches with ML surrogate models, trained on first-principle data (SE-SMR-CLC process model), to assess how these approaches provide interpretability to the model without impacting computational efficiency and recommend strategies that ensure both interpretability and computational efficiency. By providing a comparison among the approaches, the aim is to highlight what GSA method is most appropriate to select for a specific purpose dependant on the dataset and the expected outcomes (justification of the surrogate model and computational efficiency).

Table 1
Breakdown of components and mol fraction within the natural gas.

Natural gas components	mol (%)
CH_4	92.50
N_2	3.00
C_2H_6	3.00
C_3H_8	0.50
C_4H_{10}	0.310
CO_2	0.78

Table 2
Input parameters and values at which they were varied.

	Reformer Temperature (°C)	S/NG Ratio (mol)	Pressure Reformer (bara)	Temperature Calciner (°C)	CaO/NG ratio (mol)
1	400	2	5	900	0.25
2	450	2.5	10	950	0.5
3	500	3	15	1000	0.75
4	550	3.5	20	1050	1.00
5	600	4	25	1100	1.25
6	650	4.5	30	1150	1.50
7	700	5	35	1200	1.75
8	750	5.5	40	1250	2.00

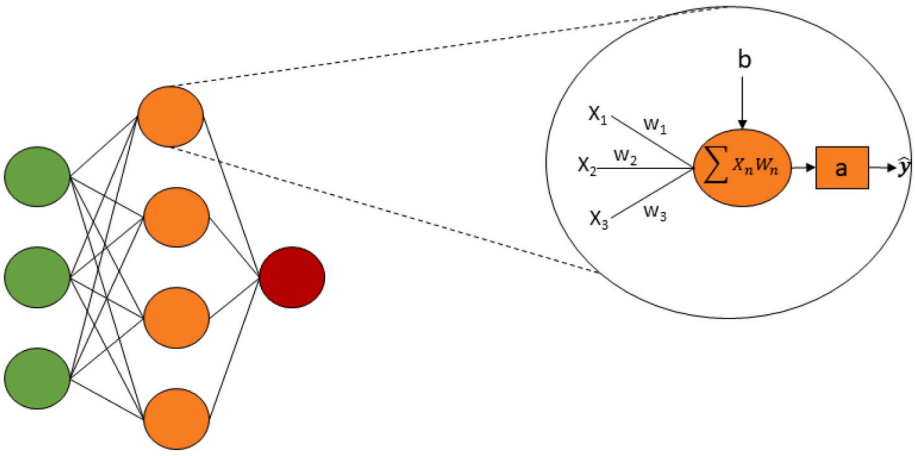


Fig. 4. ANN structure and visualisation of the calculation of a model output.

Furthermore, this study highlights how the amount of data provided is key to ensure robust, accurate ML models, and to development of interpretable ML model for application within science and engineering fields.

The paper is structured as follows. In Section 2, the methodology is discussed, detailing the data generation process, including the variables selected for both the inputs and outputs, the theory behind each ML method, as well as a description of the GSA approaches used in the study. Section 3 discusses the results associated with the training and testing of each model, highlighting the validity of the models when predicting an output. Within section 3, we comprehensively assess the different GSA approaches when applied to the models; we compare the feature importance indices for each approach and the respective computational efficiency in assessing the features importance, as well as discussing which approaches are best for a specific purpose. Finally, within section 4 we discuss the limitations of this approach and provide recommendations on future research.

2. Methodology

To develop a surrogate model and apply GSA techniques to determine the effect of each model input on the output, the following steps were followed:

1. Design of experiments using a full factorial design (FFD);
2. Collection of data from the process model in ASPEN Plus;
3. Training and testing ML surrogate models (artificial neural network and random forest);
4. Integration of GSA techniques with the surrogate model;
5. Comparison of GSA approaches with regard to interpretability, robustness and computational efficiency.

2.1. Process configuration, simulation and data generation

The data was collected based on a process model previously developed in our earlier work using ASPEN Plus (V 12.1) [24]. The process simulated is a steady-state thermodynamic model of a sorption-enhanced steam methane reforming process with chemical-looping combustion (SE-SMR-CLC). The Peng-Robinson Boston-Mathias property package was used to simulate the process [25]. The model was developed and validated in our previous work [24]. Fig. 3 shows a block-flow diagram of the process. The reformer, calciner, air-reactor (AR) and fuel reactor (FR) were modelled using the RGIBBS block in Aspen. In this process, natural gas (NG) is compressed (5–40 bara) and heated (400–750 °C). Liquid H_2O is pressurised (5–40 bara) and then heated (400–750 °C), before being mixed with natural gas and introduced into the reformer, where they also come in contact with fresh CaO. Here R1-4 take place. The solid and syngas are separated after which the $CaCO_3$ is returned to the calciner, where the reverse of R4 takes place. The regenerated CaO is recycled back into the reformer with the CO_2 stream leaving the calciner to be then compressed and stored. The natural gas (NG) composition used in this work is shown in Table 1.

The syngas enters a condenser where it is cooled, and the H_2O is condensed out. The syngas then enters a pressure-swing adsorption system (PSA) that assumes an H_2 recovery of 90 %. The hydrogen then leaves the column to be compressed and stored. The PSA off-gas is recycled and mixed with NG and is heated before entering the fuel reactor (FR), where R5-7 take place. The reduced iron oxide is recycled to be used in the air reactor (AR) where an air stream is heated and enters the FR. Here, R8 takes places with the oxidised iron oxide transported to the FR. The oxygen depleted air is cooled and leaves the system. The CO_2 from the calciner and the gas stream from the fuel reactor are mixed and the mixture is subsequently cooled. After this, the liquid H_2O is condensed out so that a high-purity stream of CO_2 is generated.

The data collected from the model focuses on the hydrogen production system. The other key units within the process (i.e. PSA and CLC systems) remain constant. Five inputs (reformer temperature, steam/NG ratio (S/NG), pressure of reformer, temperature calciner and CaO/NG

Table 3
Initial hyperparameters values selected for neural network developed in this study. Hyperparameters in bold are varied.

Hyperparameters	Hyperparameter description	Initial Value
Hidden layers	Number of layers between the input and output layers	1–3
Neurons in hidden layer	Number of neurons within the hidden layer	50–150
Activation Function	Compute the weighted sums of inputs and biases, decides whether a neuron is activated or not.	ReLU
Learning Rate	Step size at each epoch when minimizing loss.	0.001
Training/Validation/ Test Data Split (%)	The split of the data into the training, validation and test set.	60/20/20
Batch Size	Number of training examples in one pass through the neural network.	32
Epochs	Number of passes the training set is run through the neural network.	50–100
Gradient Descent Method	Updates the weights and biases so that loss is minimised.	Adam
Loss Function	Used for regression-based outputs, calculates the loss between the predicted and actual results.	Mean squared error

ratio) were varied across eight levels with a full-factorial design (FFD). An FFD was selected to enable capturing a comprehensive picture of all possible combinations of the investigated variable in this work. Overall, 32,069 simulation runs were required to complete the design space, the simulation was run on 11th Gen Intel® Core™ I7-11700 @ 2.50 GHZ with eight cores and four threads. The time taken to collect all data points was 8,136 s. These inputs were selected as the operating parameters because these factors often play a role in the performance of the blue hydrogen production plant. Table 2 shows the manipulated variables together with their corresponding values, tested in this work. These input parameters were selected as they have shown to significantly affect the performance of the investigated process i.e. hydrogen production and CO_2 capture. These independent (adjustable) variables have been shown to impact both the hydrogen production and the CO_2 capture efficiency, hence their inclusion within this model ([26,27]).

Five factors were considered as the output (i.e. Key Performance Indicator - KPI) (H_2 purity, CH_4 conversion, cold gas efficiency, CO_2 capture efficiency and CO_2 purity). The formulae used to calculate these outputs (KPIs) are shown below, through Eqs. (1)–(5).

$$H_2\ Purity\ (\%) = \frac{n_{H_2 ref}}{n_{H_2\ ref} + n_{N_2\ ref} + n_{CH_4 ref} + n_{CO\ ref} + n_{CO_2\ ref} + n_{C_2H_6 ref} + n_{C_3H_8 ref} + n_{C_4H_{10} ref}} \times 100 \quad \text{Eq. 1}$$

$$CH_4\ Conversion\ (\%) = \frac{n_{CH_4\ in} - n_{CH_4\ ref}}{n_{CH_4\ in}} \times 100 \quad \text{Eq. 2}$$

$$Cold\ gas\ efficiency\ (\%) = \frac{m_{H_2 out} \times LHV_{H_2}}{m_{NG\ total} \times LHV_{NG}} \times 100 \quad \text{Eq. 3}$$

$$CO_2\ capture\ efficiency\ (\%) = \frac{m_{CO_2\ ref.cap}}{m_{NG\ in}} \times 100 \quad \text{Eq. 4}$$

$$CO_2\ Purity\ (\%) = \frac{n_{CO_2 cap}}{n_{CO_2\ cap} + n_{CH_4\ cap} + n_{CO\ cap} + n_{H_2\ cap} + n_{H_2 O cap} + n_{N_2\ cap} + n_{C_2H_6 cap} + n_{C_3H_8 cap} + n_{C_4H_{10} cap}} \times 100 \quad \text{Eq. 5}$$

Where $n_{x\ ref}$ denotes the molar flowrate of a component leaving the reformer, $n_{CH_4\ in}$ is the molar flow rate of CH_4 entering the reformer and $n_{CH_4\ ref}$ is the molar flow rate of CH_4 leaving the reformer. $m_{H_2 out}$ is the mass flow rate of hydrogen leaving the system, with LHV_{H_2} being the lower heating value of H_2 (120 MJ/kg). $m_{NG\ total}$ is the total mass flow rate of natural gas entering the system and LHV_{NG} is the lower heating value of natural gas (47.10 MJ/kg). $m_{CO_2\ ref.cap}$ is the mass of CO_2 captured in the reformer, with $m_{NG\ in}$ being the mass flow rate of natural gas entering the reformer. n_{xcap} is the molar flow rate of a component leaving the system through the captured CO_2 stream.

All KPIs were individually calculated for each simulation run. The simulation was run under the assumptions of: steady state operation, no temperature gradient in the reactors and negligible pressure drop. These assumptions were made to enhance the model-complexity with minimum impact on the results. Once the data was collected, the data was normalised to ensure a quick convergence rate when training the surrogate models. The first-principle modelling results are provided in the supplementary information. Outliers were determined through an isolation forest as it is a high-dimension dataset and so the isolation forest is highly effective for detecting anomalies without assuming a specific data distribution [28]. Isolation forest was implemented through the scikit-learn package in which the contamination was set to 0.05. Once the outliers were detected, a sample was analysed to characterise these outliers, and if values were physical impossibilities (e.g. CO_2 capture efficiency >1.00), they were removed from the dataset [29].

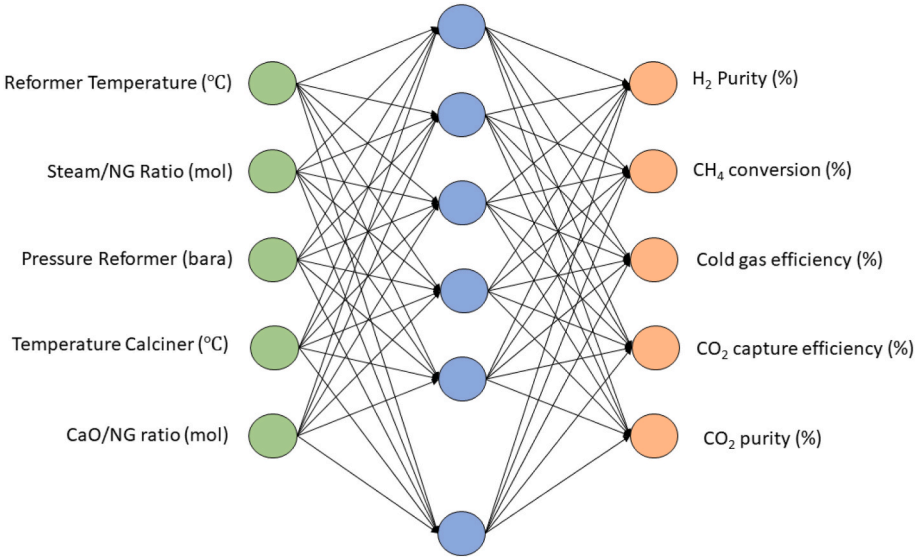


Fig. 5. Neural network structure used to develop the surrogate model in this study.

Table 4
Values at which selected hyperaparameters were varied for ANN.

	Lower	Middle	Upper
Number of neurons	50	100	150
Number of epochs	50	75	100
Number of hidden layers	1	2	3

2.2. Machine learning surrogate model development

Developing the ML surrogate models was done in python v3.6.1 in the Jupyter notebook environment using TensorFlow package [30] for the artificial neural network (ANN) and Scikit-learn [29] for the random forest (RF). The following sections describe the working principle for ANNs and RFs as well as the hyperparameters selected for each surrogate model.

2.2.1. Artificial neural network

ANNs is a machine learning method that connects inputs and outputs via a series of interconnected nodes and layers [31]. They are extremely flexible and able to solve complex problems that are non-linear in nature [32]. The structure of a neural network consists of an input layer, hidden layers (can be multiple hidden layers) and an output layer, as shown in Fig. 4. ANN is based on the principle of information processing in biological systems [31]. The data is passed through each layer, and an output value is produced. The nodes within each layer are known as neurons. The basic structure of a neuron is shown in Fig. 4. To calculate the output of a neuron, Eq. (6) is used.

$$\hat{\boldsymbol{y}} = a\left(\sum_{i=1}^{n} x_i w_i + b\right), \quad \text{Eq. 6}$$

where x_i is an input, w_i is a weight, b is the bias, a is the activation function and $\hat{\boldsymbol{y}}$ is the vector of the model output. Once all the data is passed through (known as an epoch) the loss is calculated (mean squared error), which is the difference between the actual output values and the predicted values predicted by the surrogate model. Back-propagation is then used in which the chain rule is used to calculate the rate at which loss changes in response to any change to a specific weight (or bias) in the network [33]. Once this partial derivative is calculated, the gradient descent is utilised which is where the weights are updated to ensure the loss function is minimised [33]. This process is repeated

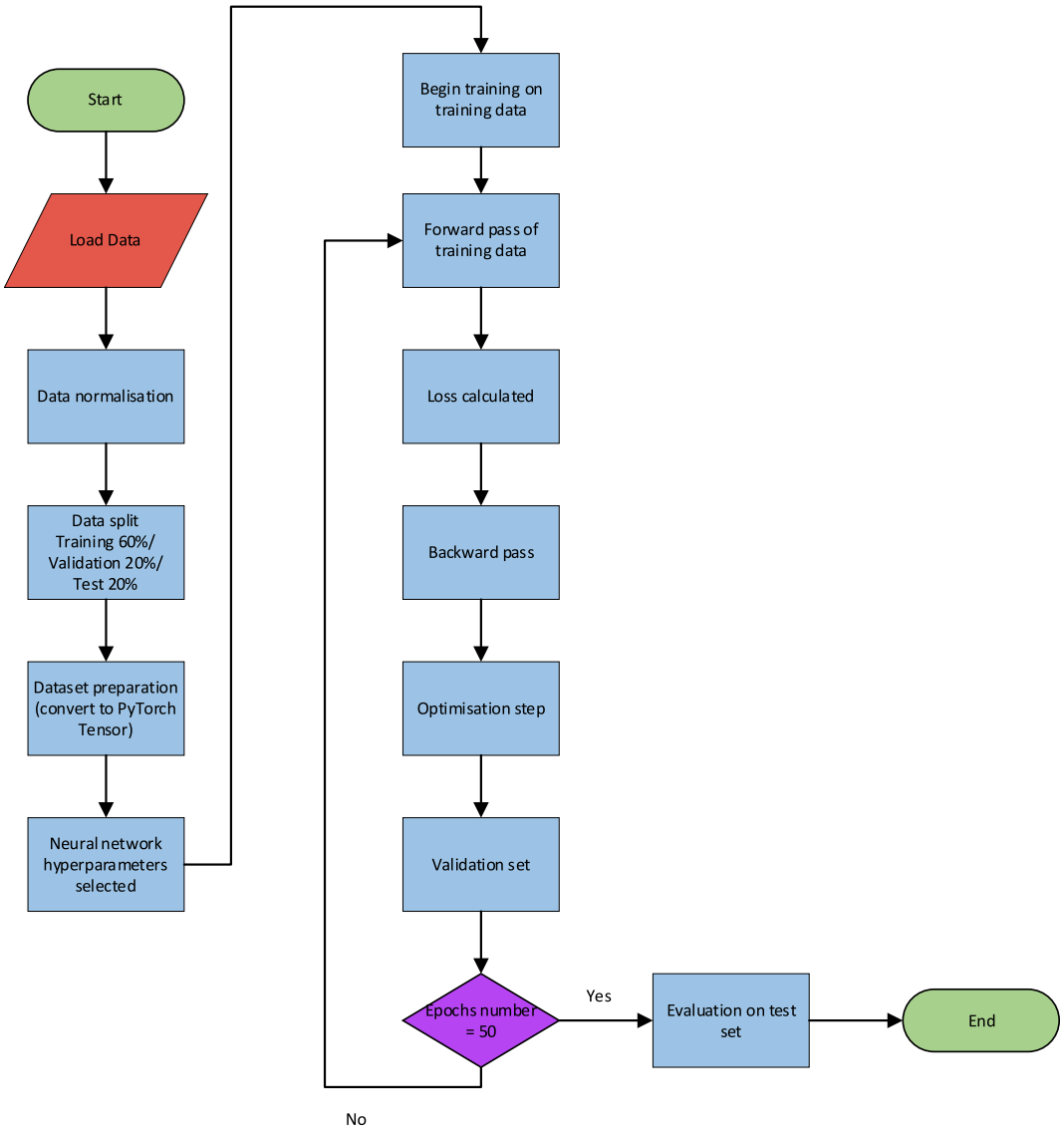


Fig. 6. Flowchart for learning procedure of ANN.

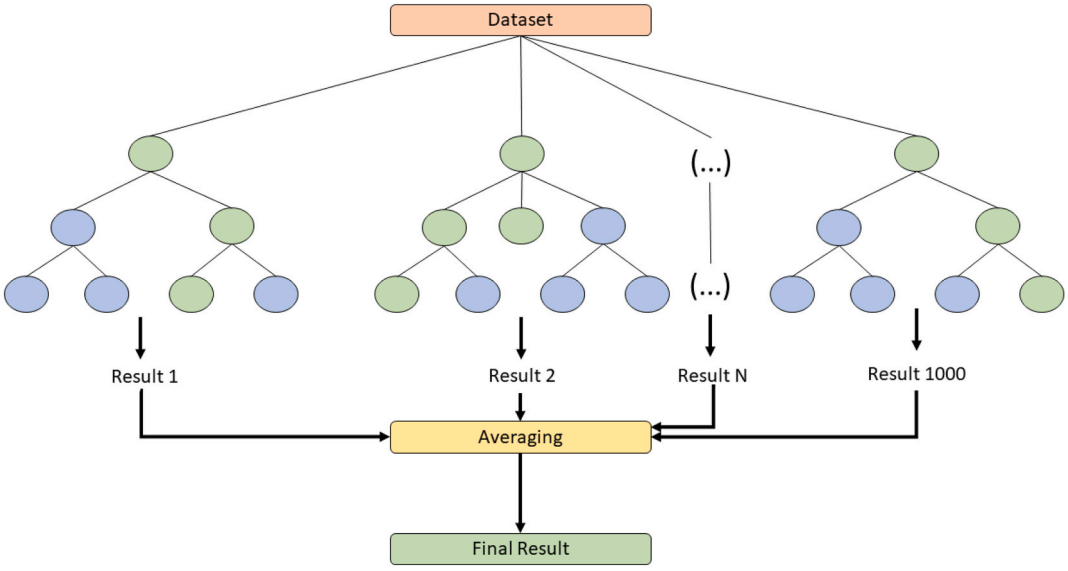


Fig. 7. Structure of initial random forest surrogate model.

Table 5
Initial hyperparameters for the RF surrogate model.

Hyperparameters	Values
Number of estimators	50–150
Max depth	10, 50, None
Training/Test split	60/40
Leaf nodes	None
Minimum samples split	2

Table 6
Values at which selected hyperaparameters were varied for RF.

	Lower	Middle	Upper
Number of Trees	50	100	150
Max Depth	10	50	None

until the loss is minimised.

In this study, a feedforward with backpropagation neural network (FBNN) was selected and developed in python version 3.6.1 using the TensorFlow package [30]. An FBNN was selected due to its ease of use, especially when calculating the backpropagation. Also, since the data is not time sequential, a recurrent neural network is not needed. The initial hyperparameters values selected are shown in Table 3, with Fig. 5 showing the structure of the ANN used. The accuracy of the surrogate model was determined by calculating the R^2 and the mean squared error (MSE), calculated by Eq. (7), respectively.

$$R^2 = 1 - \frac{sum\ squared\ regression}{total\ sum\ of\ squares} = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2}, \quad \text{Eq. 7}$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \quad \text{Eq. 8}$$

where $\hat{y}_i$ is the predicted output, y_i is the actual (synthetic) data, $\bar{y}_i$ is the mean of all values and n is the number of samples. The following hyperparameters were varied through an FFD from the values in Table 4 (number of neurons, number of epochs and number of layers). These parameters were selected as they play a key role in the computational efficiency and the accuracy of the deployed model [34]. This was done to assess the trade-off between average accuracy (R^2, MSE) and training time per epoch. The training and validation sets were used for all the models whilst the test set was used on the selected optimised model. Fig. 6 shows the flowchart detailing how the ANN learns from the dataset.

2.2.2. Random forest

RF is an ensemble-learning approach in which multiple decision trees are "bagged" together to predict an output [35]. Ensemble learning has been shown to have high predictive capability in comparison to other ML approaches and can be used for both regression and classification tasks [11]. Ensemble learning methods combine multiple machine learning methods and aggregate them together to find the average result. RF combines the bagging method with feature randomness to create an uncorrelated mixture of decision trees [36]. Initially developed by Brieman [36], each tree utilises a random part of the input dataset to calculate a random part of the features in each partition, as shown in Eq. (9) [35]. This randomness produces variability between the trees, which reduces the risk and improves the overall prediction [35]. Splits within trees are selected based on the minimisation of the MSE. The tree will then grow until the stopping conditions are met. The output of the whole RF is calculated by Eq. (10).

$$\hat{y}_{dt} = \sum_{m=1}^{M} c_m . \mathbb{I}(x \in R_m), \text{ where } c_m = \frac{1}{|R_m|}\sum_{i \in R_m} y_i \quad \text{Eq. 9}$$

$$\hat{y}_{RF} = \frac{1}{B}\sum_{b=1}^{B}\hat{y}_{dt} \quad \text{Eq. 10}$$

Where R_m is the region within the data y_i. $\hat{y}_{RF}$ is the predicted output of the random forest, and B is the bootstrap samples of the dataset. $\hat{y}_{dt}$ is the predicted output of the decision tree. The structure of the initial RF surrogate model is shown in Fig. 7 and the initial hyperparameters are summarised in Table 5.

The following hyperparameters were varied *via* an FFD from the values in Table 6 (number of trees and max tree depth). This was done to assess the trade-off between average accuracy (R^2, MSE) and training time per epoch. Both parameters were selected as for the RF, both are influential hyperparameters that determine the computational efficiency and the accuracy of the deployed ML model [37]. Fig. 8 provides a flowchart detailing how the RF learns from the dataset.

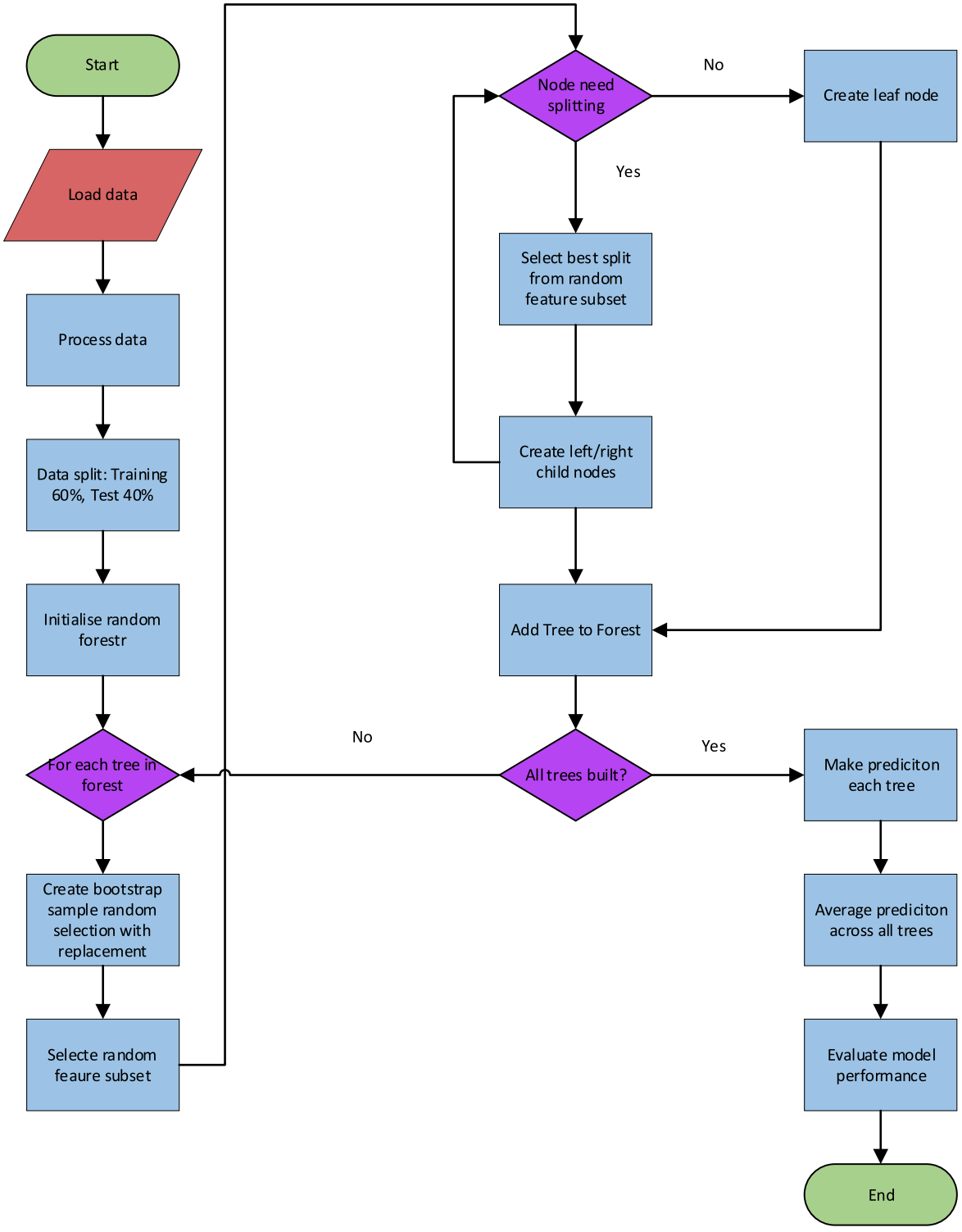


Fig. 8. Flowchart for training and testing of random forest.

2.3. Global sensitivity analysis techniques

Once the ML surrogate models were trained and tested to ensure high accuracy, they were subsequently used to generate sample data to enable GSA. The GSA was conducted using the SAlib package in Python v3.6.1 [38]. The exception to that was Shapley which was conducted using the SHAP package [39]. The following sections describe how each GSA approach used in this study, calculates the sensitivity indices for each input, or a combination of inputs. For each approach, initially a sample of 1024 datapoints was used before the sampling was varied to determine the robustness of each approach (described in Tection 2.4.).

2.3.1. Variance-based methods

Variance-based methods assume that variance is enough to describe the output uncertainty [40]. They calculate a sensitivity of the input variables via an ANOVA-like decomposition of the function [41]

Table 7
Hyperparameters of selected best-performing ANN model.

Hyperparameter	Number of Neurons	Number of hidden layers	Epochs
Value	50	2	100

Table 8
Comparison of the best-performing ANN model with average result of all developed models across R^2, MSE and training time.

	Training R^2 (%)	Training MSE	Validation R^2 (%)	Validation MSE	Training time all epochs (seconds)	Training time per epoch (seconds)
Optimal model	99.99	0.0002	99.97	0.0001	164.36	1.64
Average across all models	99.70	0.000167	99.66	0.000185	139.62	1.87

(Sobol), or through a Fourier transformation (Fourier Amplitude Sensitivity Test).

2.3.1.1. Sobol. The Sobol approach decomposes the variance of the output of a model into fractions that are attributed to an input variable via an ANOVA-like decomposition [42]. The Sobol approach can be used to calculate first-order sensitivity (a single input effect on the variance of the output), second-order sensitivities (a combination of multiple inputs effects on the variance) and total-order Sobol index (measures the overall contribution of an input including both direct (main) and interaction effects). In order to calculate the Sobol indices from model ($f(X)$) and inputs ($X = (X_1, X_2, \ldots, X_k)$). Two independent sampling matrices **A** and **B** are generated of size $N \times k$, with N being the sample size and k being the number of input variables. For each of the inputs, a third matrix is generated, known as $\mathbf{C_i}$, which is a hybrid matrix where columns are from matrix **A** with the i-th column being from matrix **B**. From this the model, outputs are computed as shown in Eq. (11)–(13).

$$y_A = f(\boldsymbol{A}) \quad \text{Eq. 11}$$

$$y_b = f(\boldsymbol{B}) \quad \text{Eq. 12}$$

$$y_{C_i} = f(\boldsymbol{C}_i) \quad \text{Eq. 13}$$

Where y_A is the model outputs evaluated using samples from matrix A. y_b the model outputs evaluated using samples from matrix B. y_{C_i} is the model outputs evaluated using samples from matrix C_i. From this, the total variance (V) can be calculated via Eq. (14).

$$V = \frac{1}{N}\sum_{j=1}^{N} y_A^{(j)2} - \left(\frac{1}{N}\sum_{j=1}^{N} y_A^{(j)}\right)^2 \quad \text{Eq. 14}$$

Where $y_A^{(j)}$ is the j-th output from y_A and $\left(\frac{1}{N}\sum_{j=1}^{N} y_A^{(j)}\right)^2$ is the square of the mean output, which is done to centre the variance. To calculate the first-order effect (V_i) of an input X_i, Eq. (15) is used. In order to calculate the Sobol first-order index (S_i) of an input, Eq. (16) is used.

$$V_i = \frac{1}{N}\sum_{j=1}^{N} y_A^{(j)} y_{C_i}^{(j)} - \left(\frac{1}{N}\sum_{j=1}^{N} y_A^{(j)}\right)^2 \quad \text{Eq. 15}$$

$$S_i = \frac{V_i}{V} \quad \text{Eq. 16}$$

Where $y_{C_i}^{(j)}$ is the j-th output of matrix C_i, to then calculate the total order Sobol index, the variance (V_{-i}) contributed by all inputs except input X_i is calculated by Eq. (17). The total order Sobol index of X_i (S_{T_i}) is calculated by Eq. (18).

Table 9
R^2 of best-performing ANN model across training, validation and test set.

Optimised Model	Training Set	Validation Set	Test Set
R^2	99.99	99.97	99.94

Table 10
Optimised hyperparameters of selected RF model.

Hyperparameter	Number of Tress	Max Tree Depth
Value	100	None

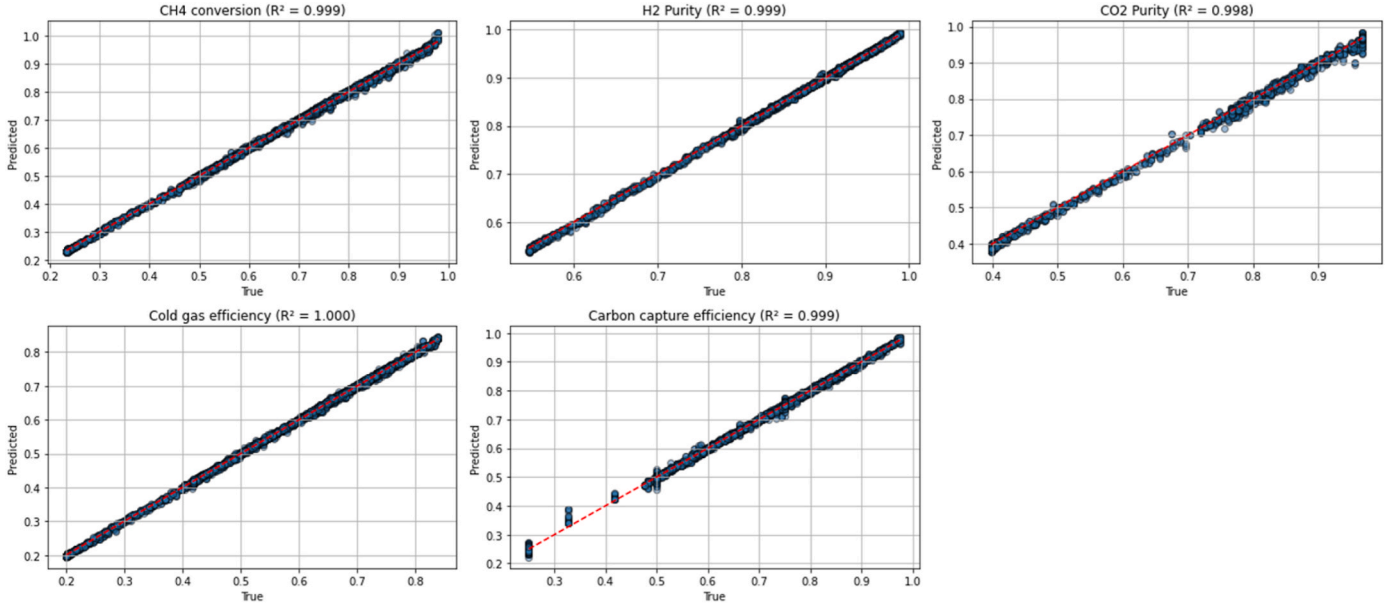


Fig. 9. Average R^2 Plots for each output from the ANN for the training set.

Table 11
RF optimal model R^2, MAE and training time.

	Training R^2	Training MSE	Test R^2	Test MSE	Time taken for training to complete (seconds)
Best performing model	99.99	0.0021	99.92	0.0025	4.38
Average across training	99.71	0.004144	99.63	0.004585	3.84

Table 12
R^2 of optimised RF model across training and test set.

	Training	Test
R^2 (%)	99.99	99.92

$$V_{-i} = \frac{1}{N}\sum_{j=1}^{N} y_B^{(j)} y_{C_i}^{(j)} - \left(\frac{1}{N}\sum_{j=1}^{N} y_A^{(j)}\right)^2 \quad \text{Eq. 17}$$

$$S_{T_i} = 1 - \frac{V_{-i}}{V} \quad \text{Eq. 18}$$

Where $y_B^{(j)}$ is the j-th output value from y_b. The second order effects (S_{ij}) are calculated via Eq. (19).

$$S_{ij} = \frac{V_{ij}}{V} - S_i - S_j \quad \text{Eq. 19}$$

Where S_{ij} is the second order Sobol index of input variables Xi and Xj and V_{ij} is the variance due to the interaction between Xi and Xj.

2.3.1.2. Random-balance design Fourier amplitude sensitivity test.

Random-balance design - Fourier amplitude sensitivity test (RBD-FAST) is another variance-based approach. Here, sensitivity values calculated are independent from the sampling scheme. A Fourier transformation is then used to decompose the variance of the model into partial variances [43]. To determine sensitivities indices for a model with k input factors $(X_1, X_2, \ldots, X_k)$ and N sample points, each input factor is sampled using a random permutation. This ensures that each input is uniformly distributed whilst maintaining orthogonality. The model is then run to obtain outputs (Eq. 20

$$y_j = f\left(x_1^{(j)}, x_2^{(j)}, \ldots, x_k^{(j)}\right), j = 1, 2, \ldots, N \quad \text{Eq. 20}$$

Where y_j is the output of a model. A Fourier transform is applied to compute the Fourier coefficient as shown in Eq. (21).

$$Y_m^{(i)} = \frac{1}{N}\sum_{j=1}^{N} y_j^{(i)} e^{-i\frac{2\pi mj}{N}}, m = 0, 1, \ldots, N-1 \quad \text{Eq. 21}$$

Where $Y_M^{(i)}$ is the m-th Fourier coefficient for an input factor, representing the amplitude and phase at frequency m. The power spectrum at frequency m is determined by Eq. (22).

$$A_m^{(i)} = \left|Y_m^{(i)}\right|^2 \quad \text{Eq. 22}$$

From this, the total variance (V) of an output is determined via Eq. (23). The sensitivity of an input (Xi) is then calculated via Eq. (24).

$$V = 2\sum_{m=1}^{N/2-1} A_m^{(i)} \quad \text{Eq. 23}$$

$$S_i = \frac{2\sum_{p=1}^{M} A_{p\omega_i}^{(i)}}{V} \quad \text{Eq. 24}$$

Where ω_i is the characteristic frequency for Xi, and M is the number of harmonics considered.

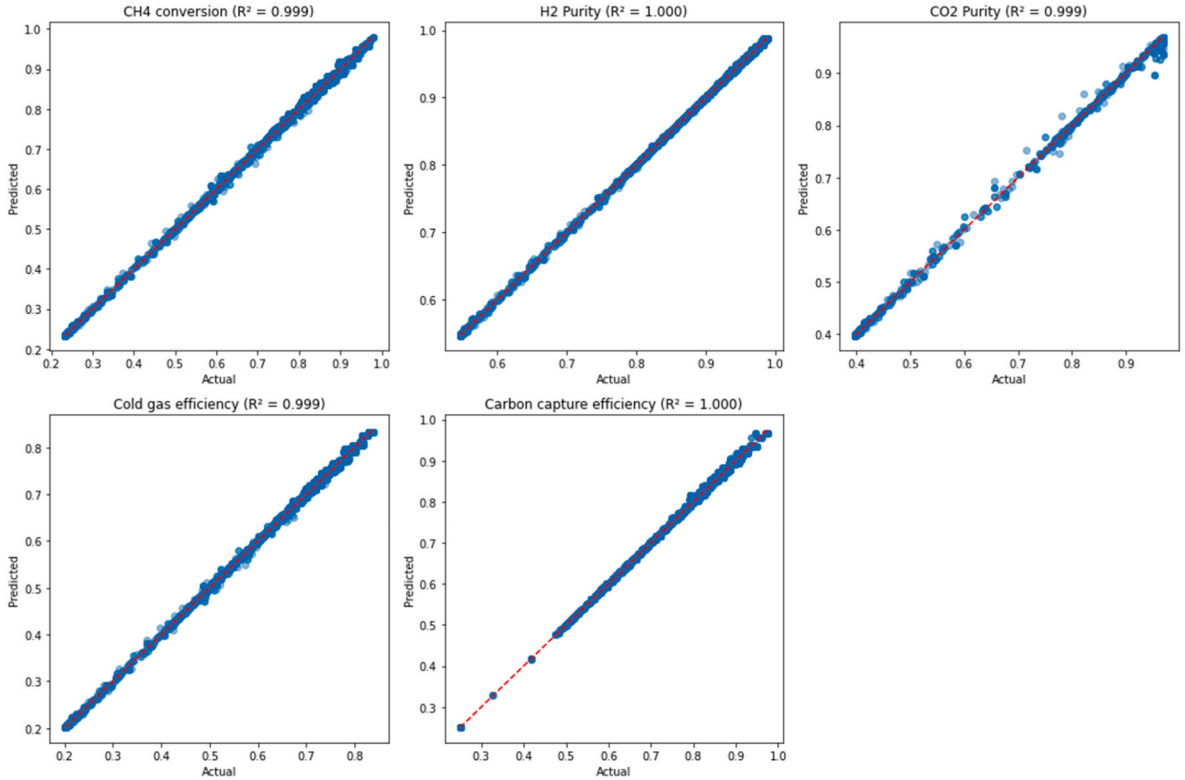


Fig. 10. Average R^2 plots for each output from the RF for the training set.

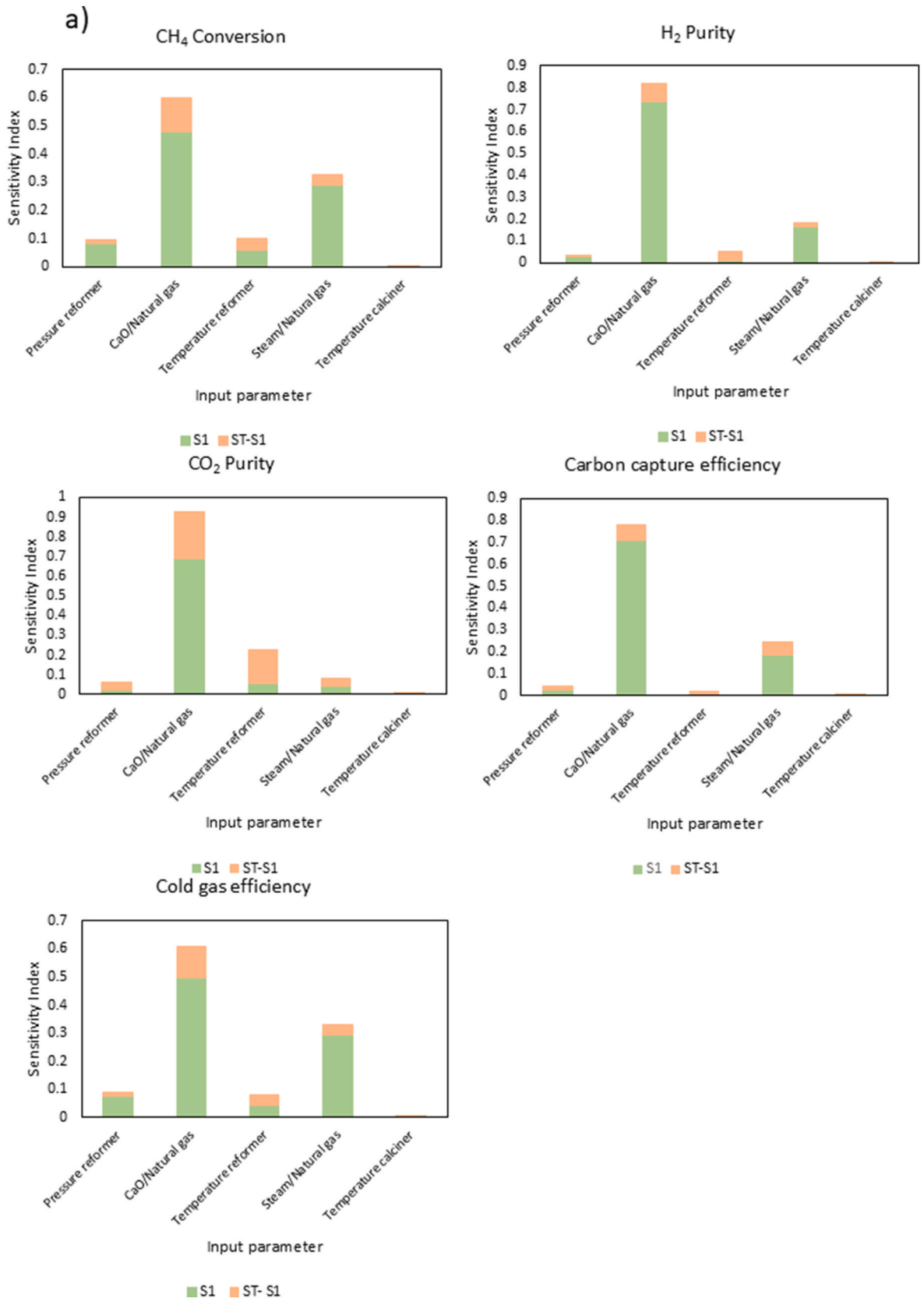


Fig. 11. Total-order sensitivities for each inputs influence on each output. a) ANN. b) RF.

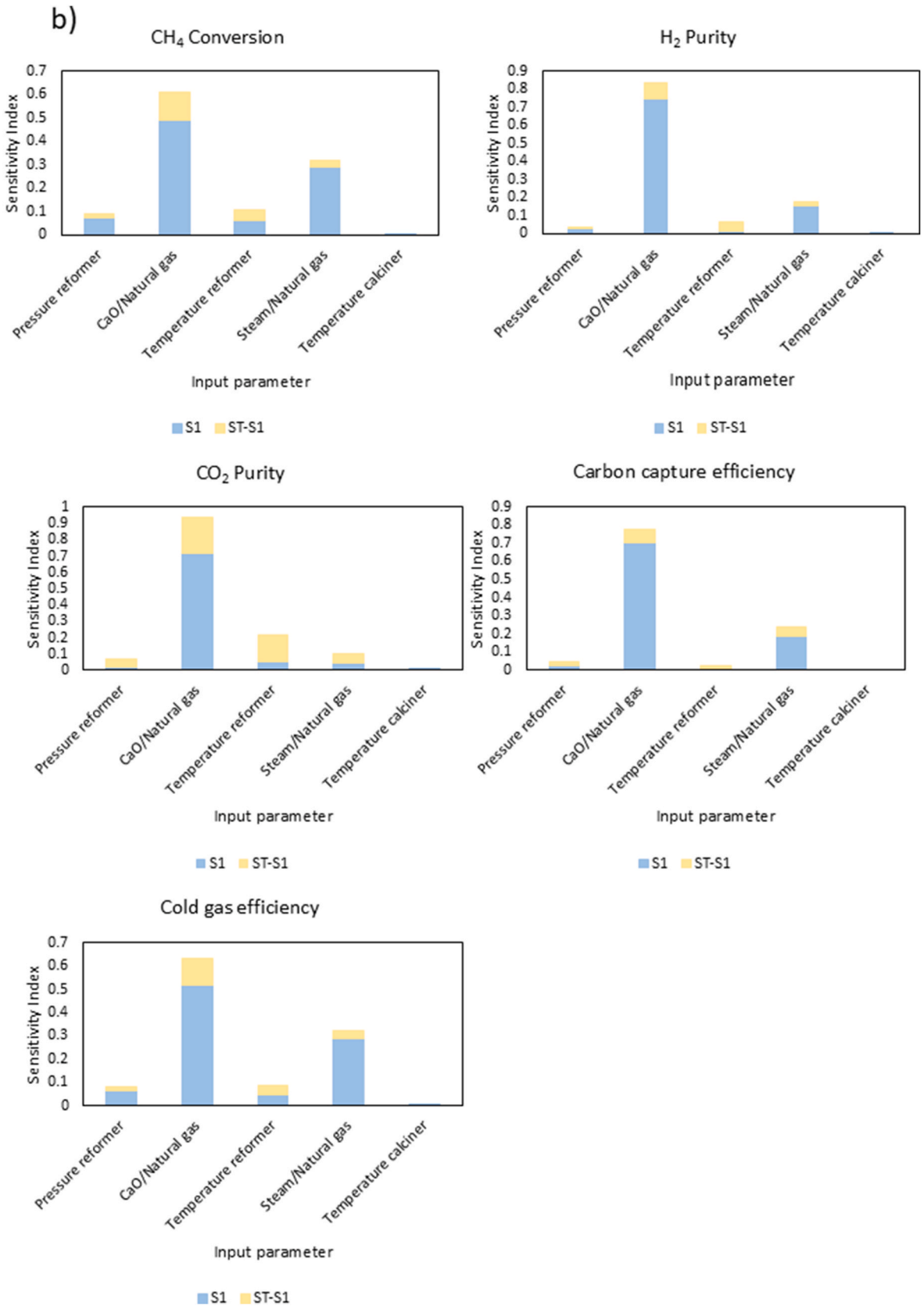


Fig. 11. (*continued*).

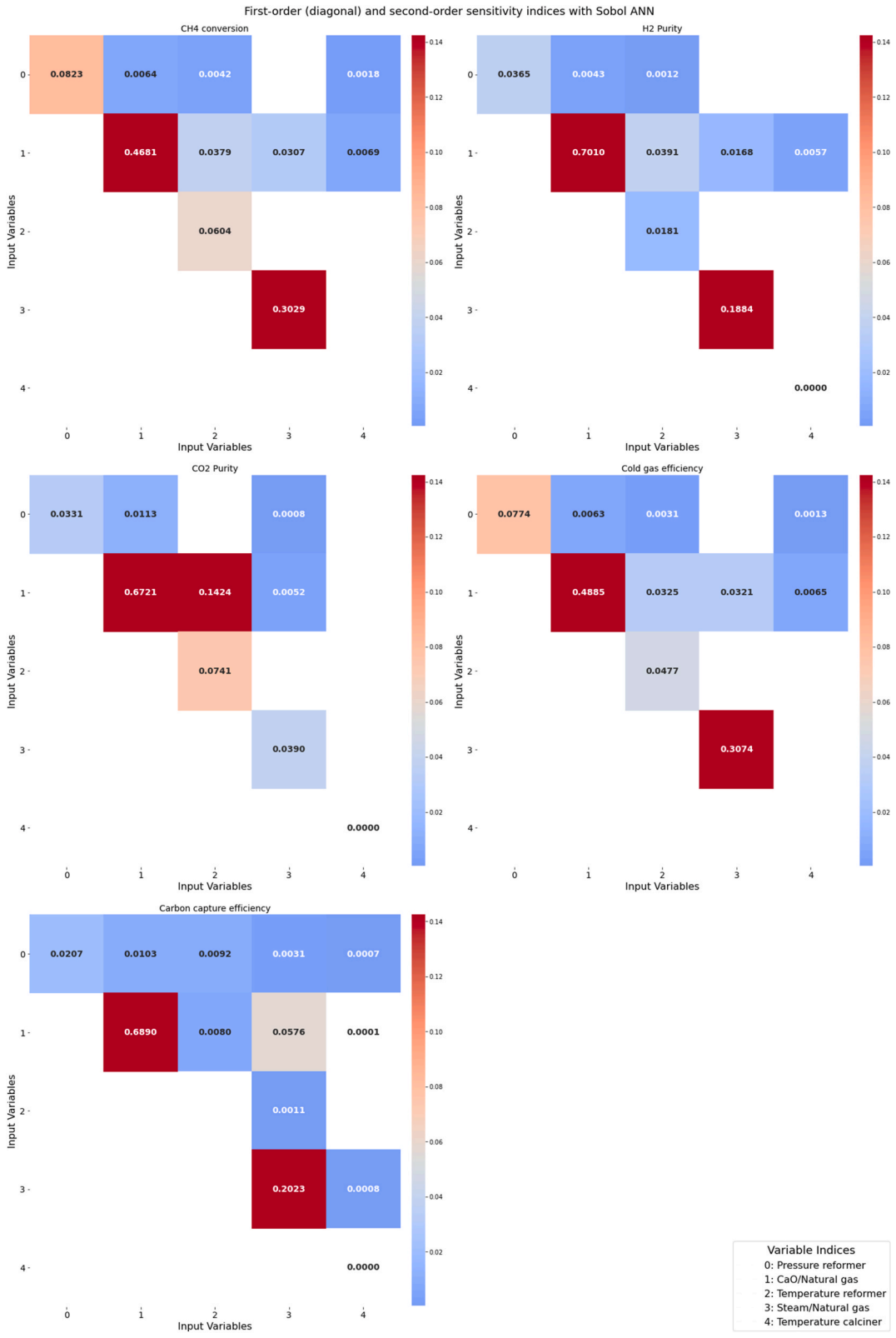


Fig. 12. a) ANN second order Sobol sensitivities. b) RF Sobol second order sensitivities.

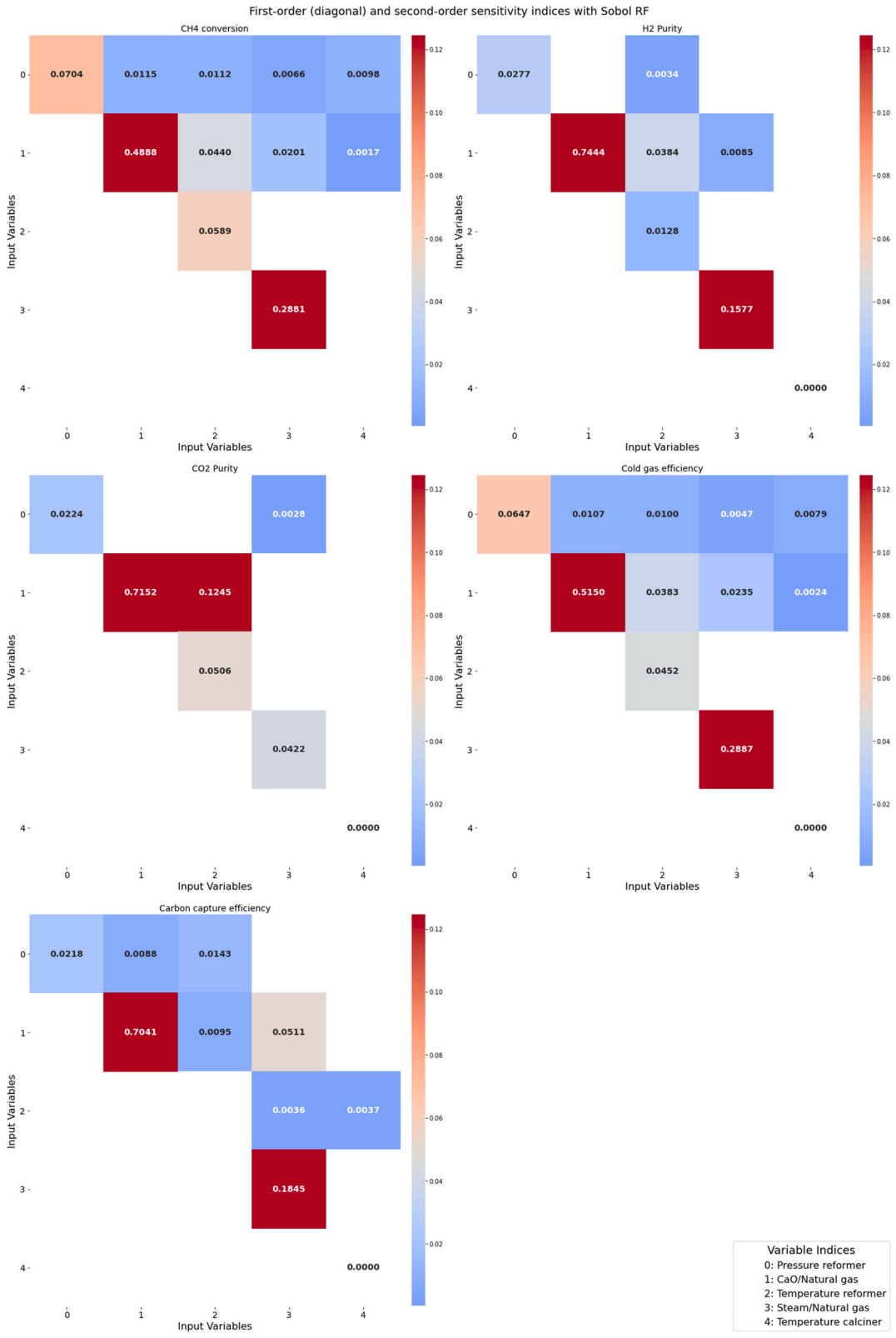


Fig. 12. (*continued*).

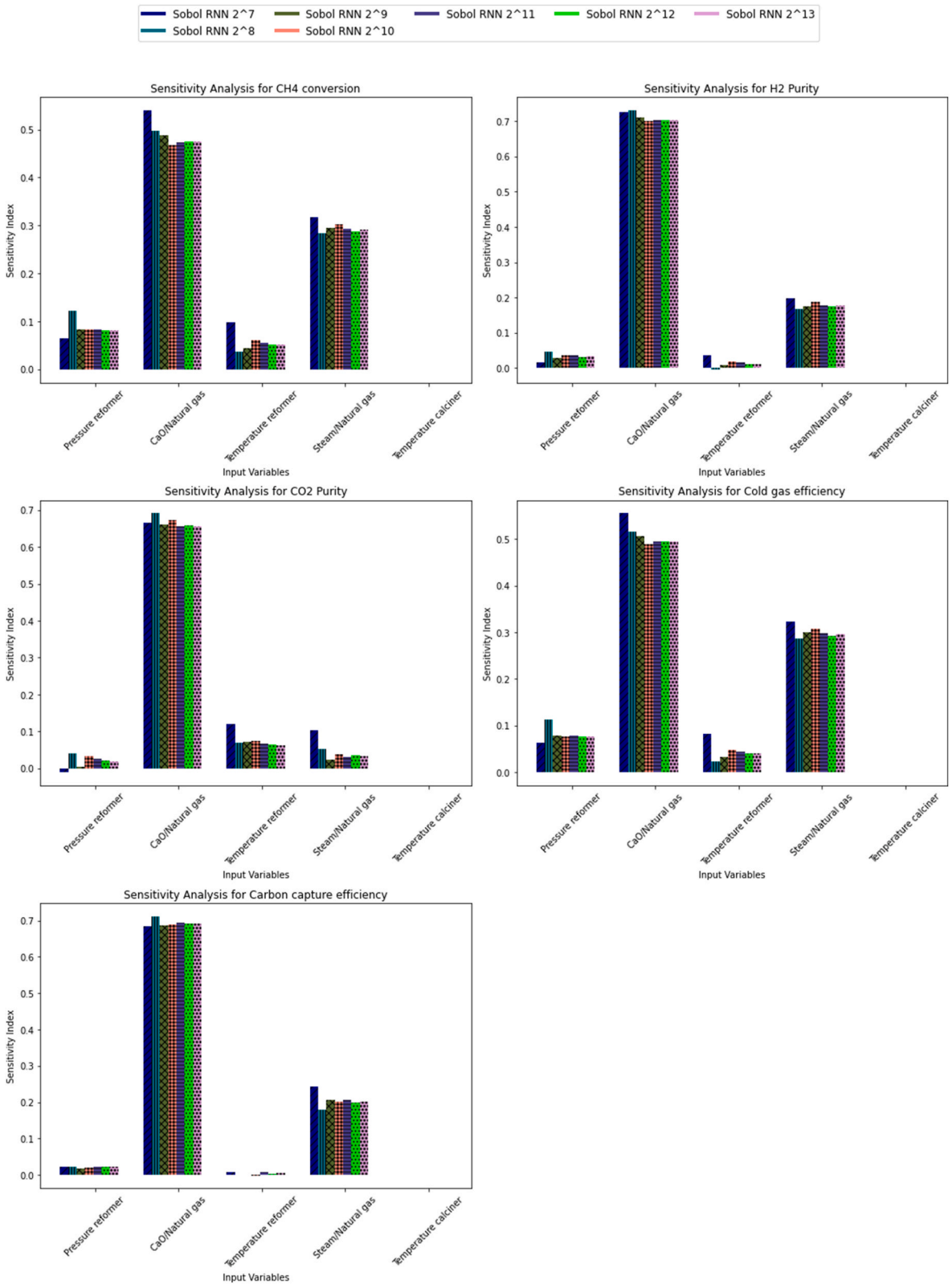


Fig. 13. a) Sobol first-order sensitivity for ANN across a varied sample size. b) Sobol first order sensitivity for RF across a varied sample size.

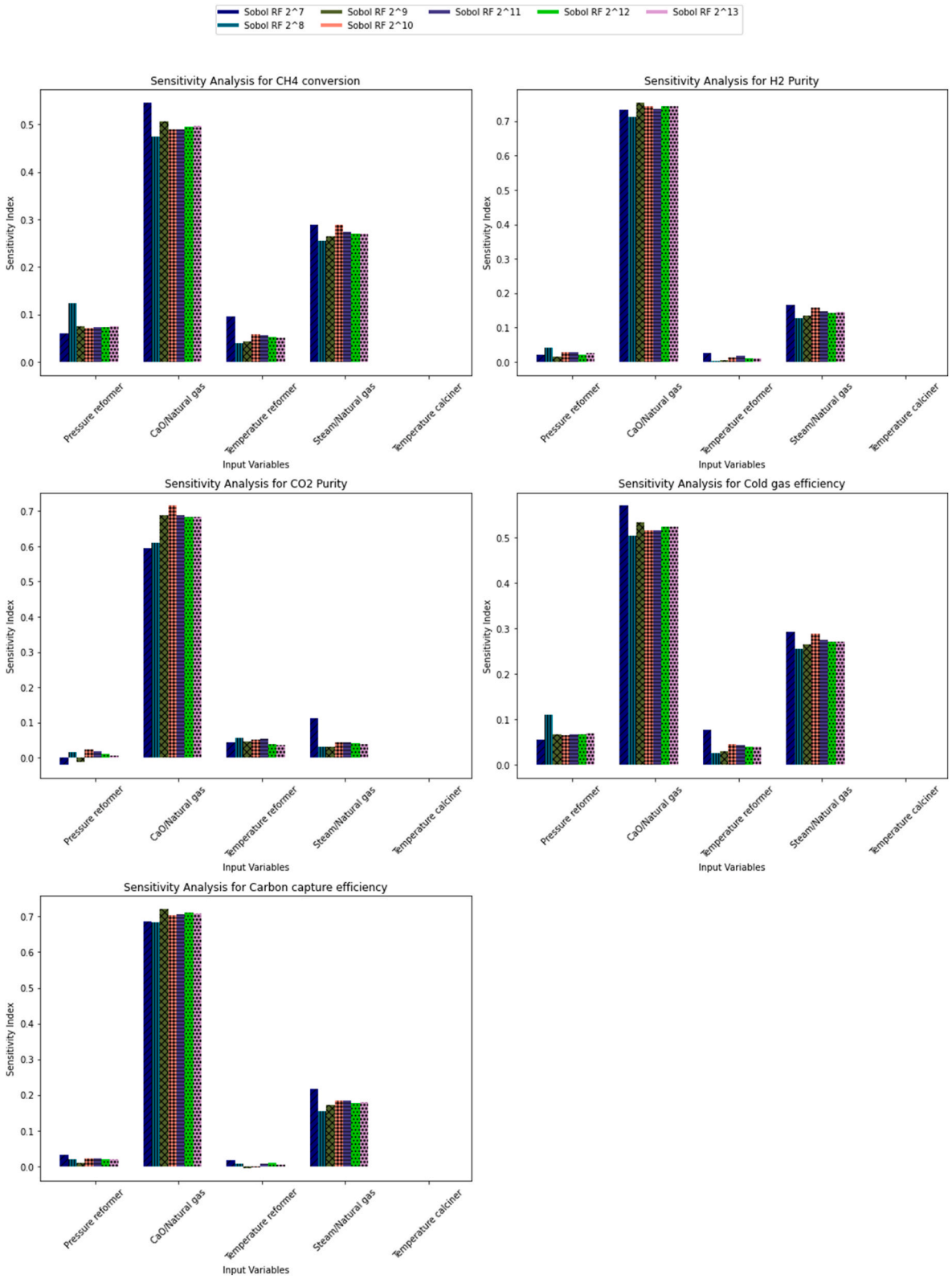


Fig. 13. (*continued*).

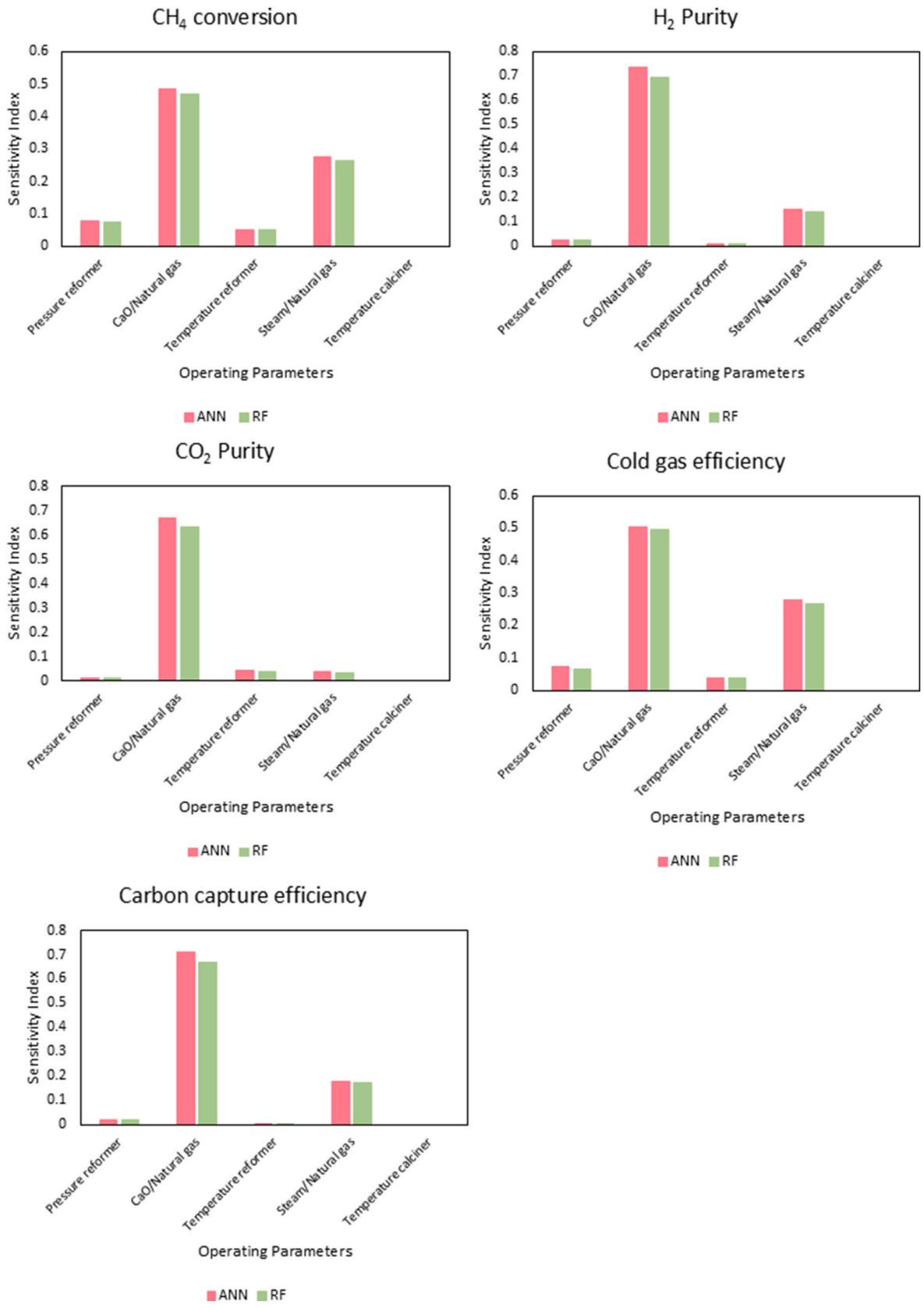


Fig. 14. Sensitivity of an input to each output for both models calculated by the RBD-FAST approach.

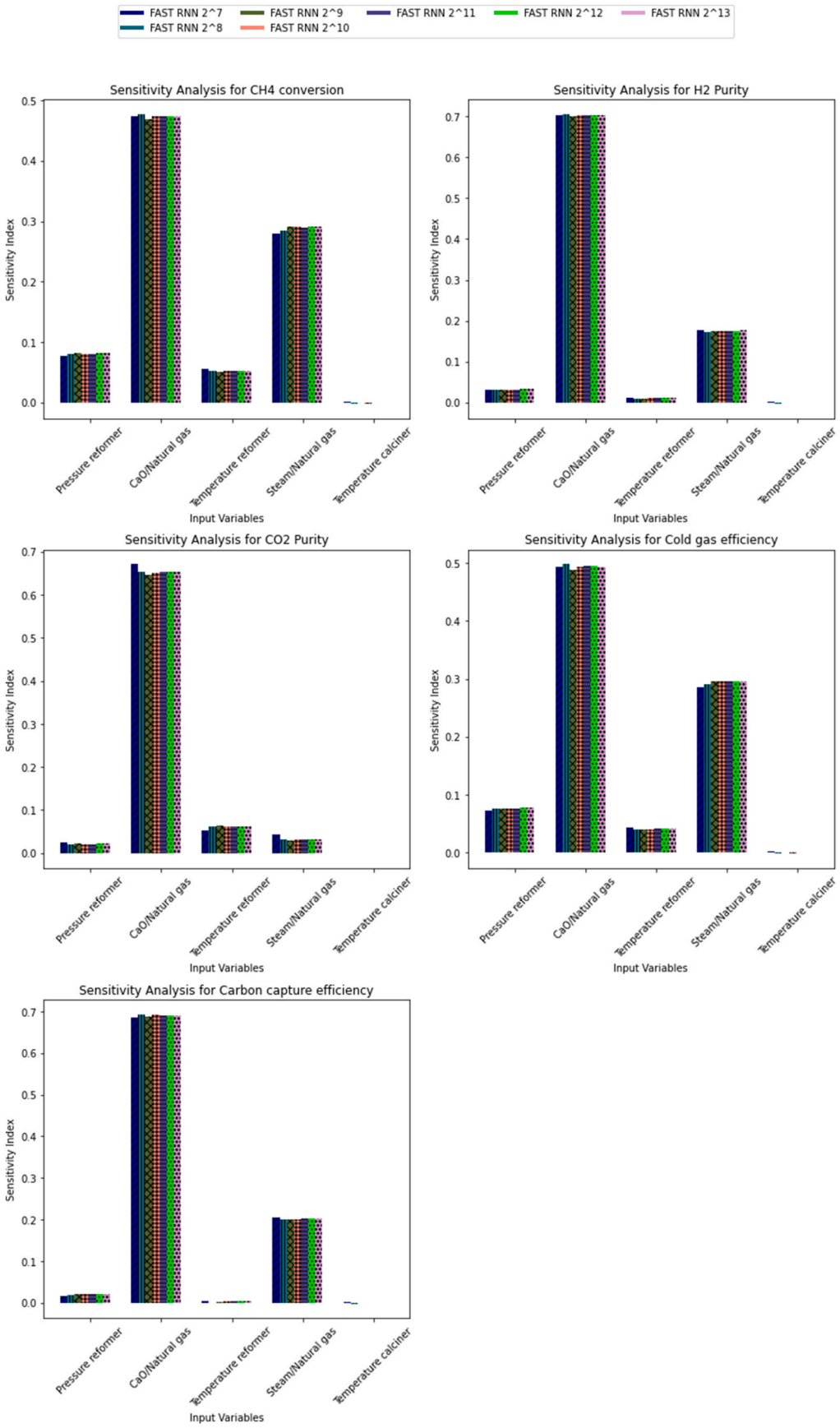


Fig. 15. a) ANN robustness plots for RBD-FAST. b) RF Robustness plots for RBD-FAST.

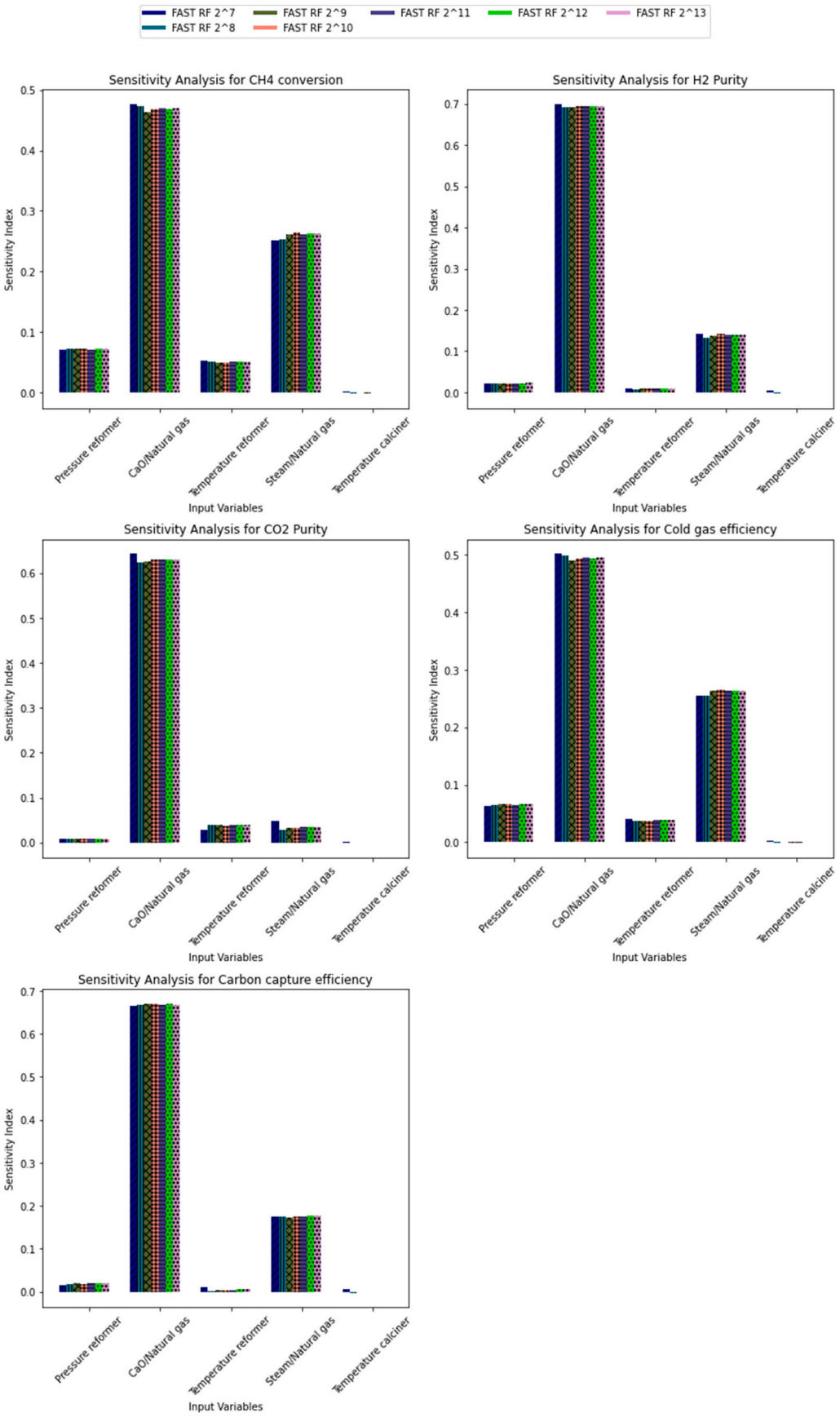


Fig. 15. (*continued*).

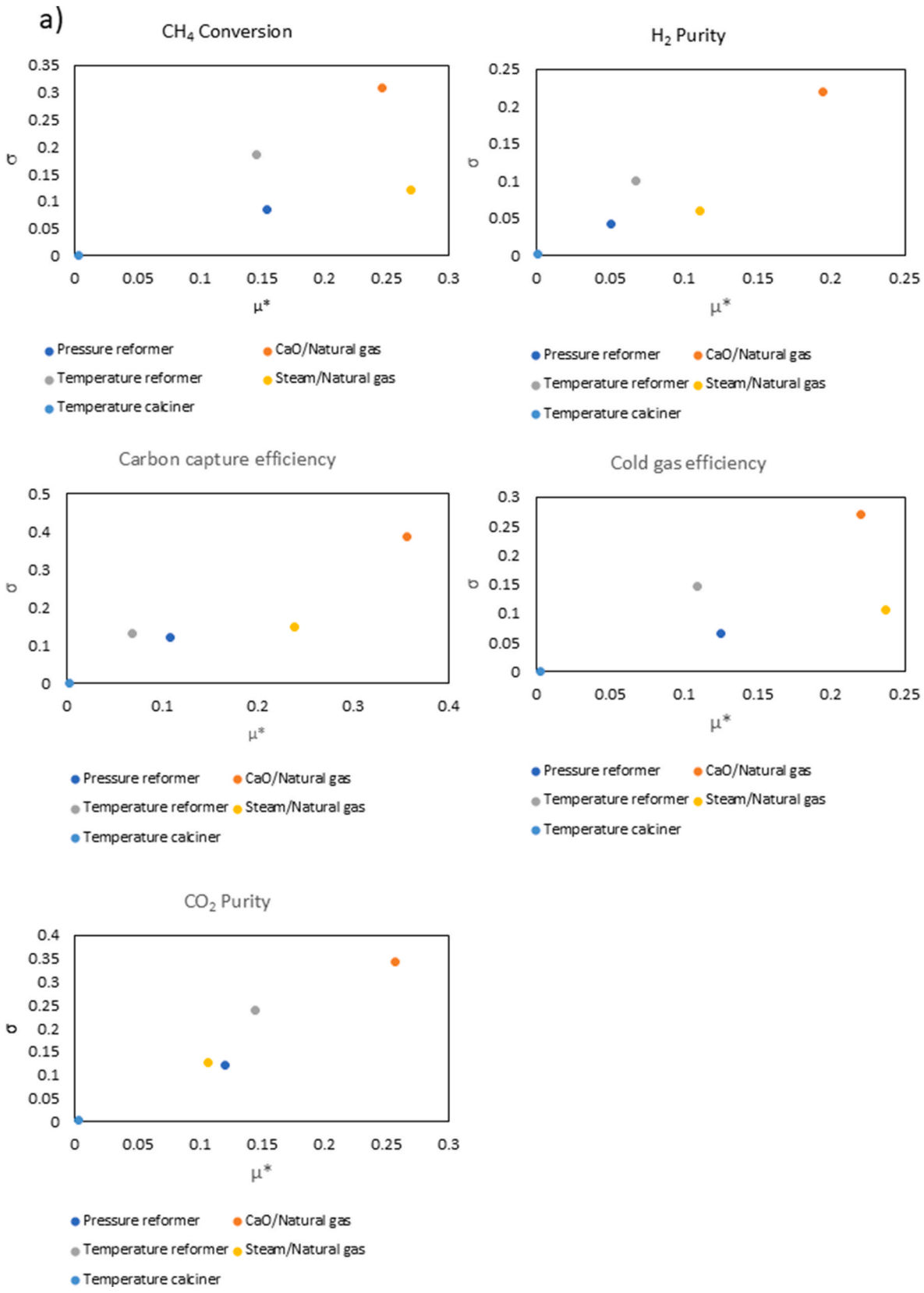


Fig. 16. μ_i^* *vs* σ_i plots calculated by the Morris approach a) ANN b) RF.

b)

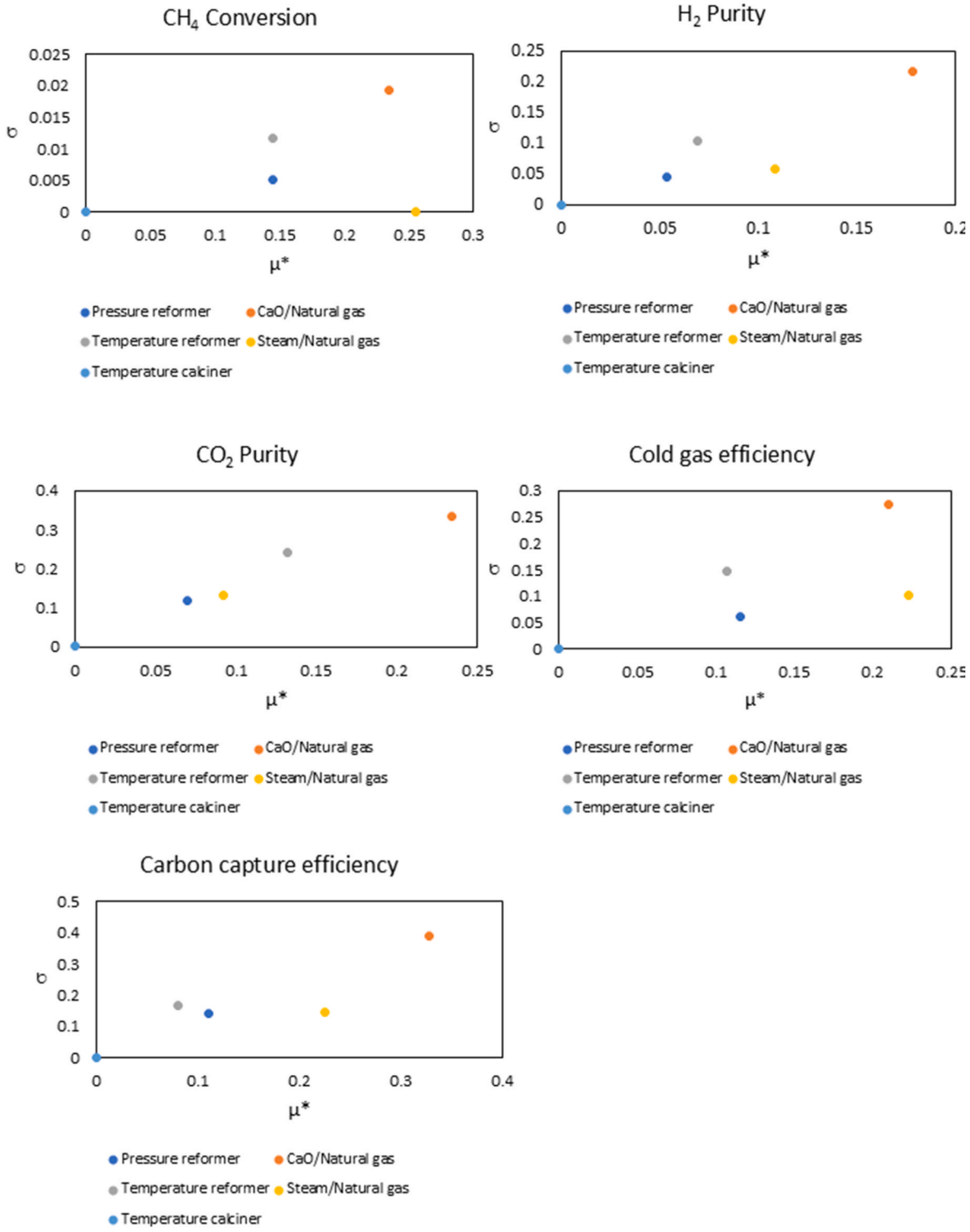


Fig. 16. (*continued*).

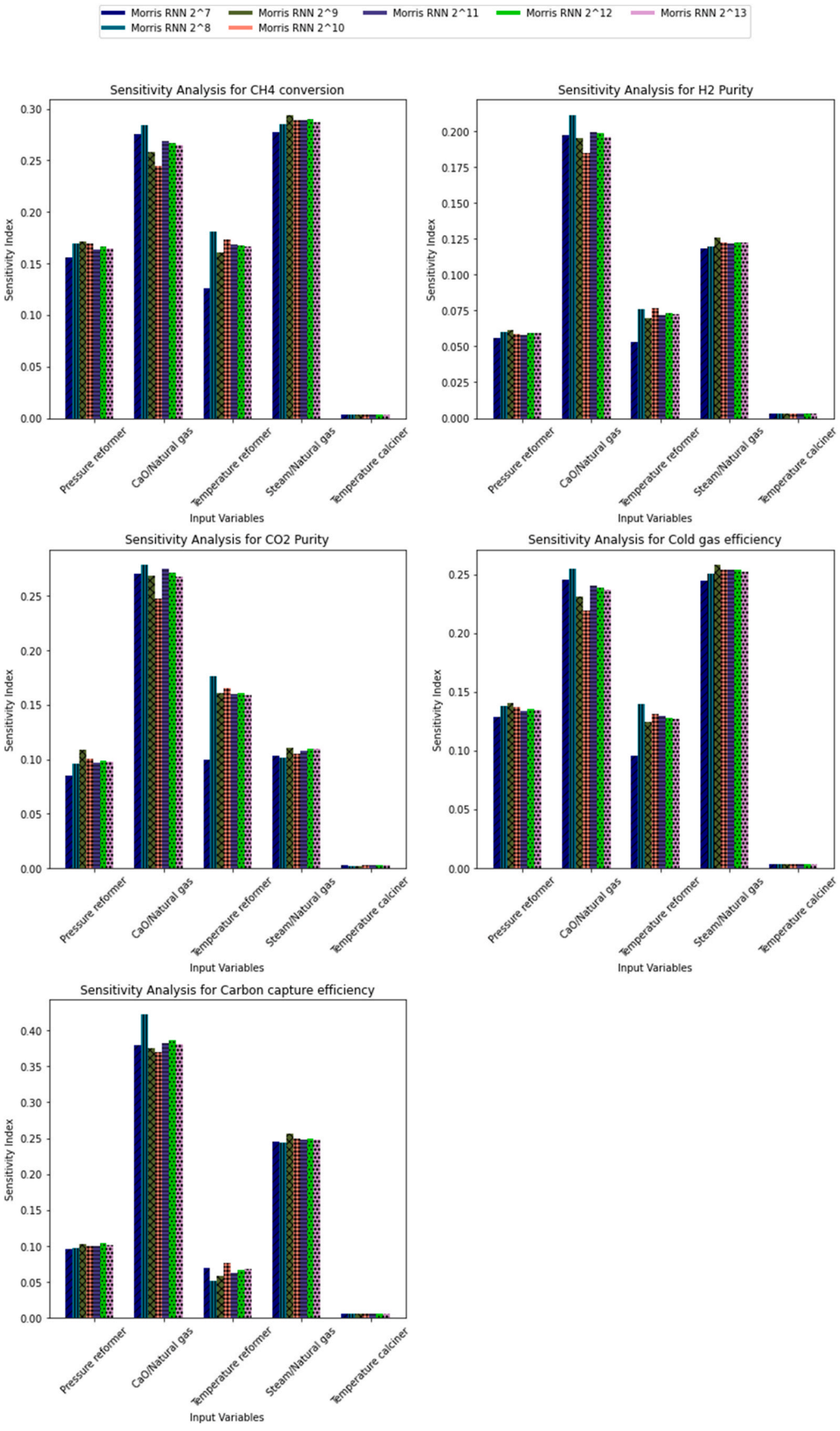


Fig. 17. Robustness for the Morris approach. a) ANN. b) RF.

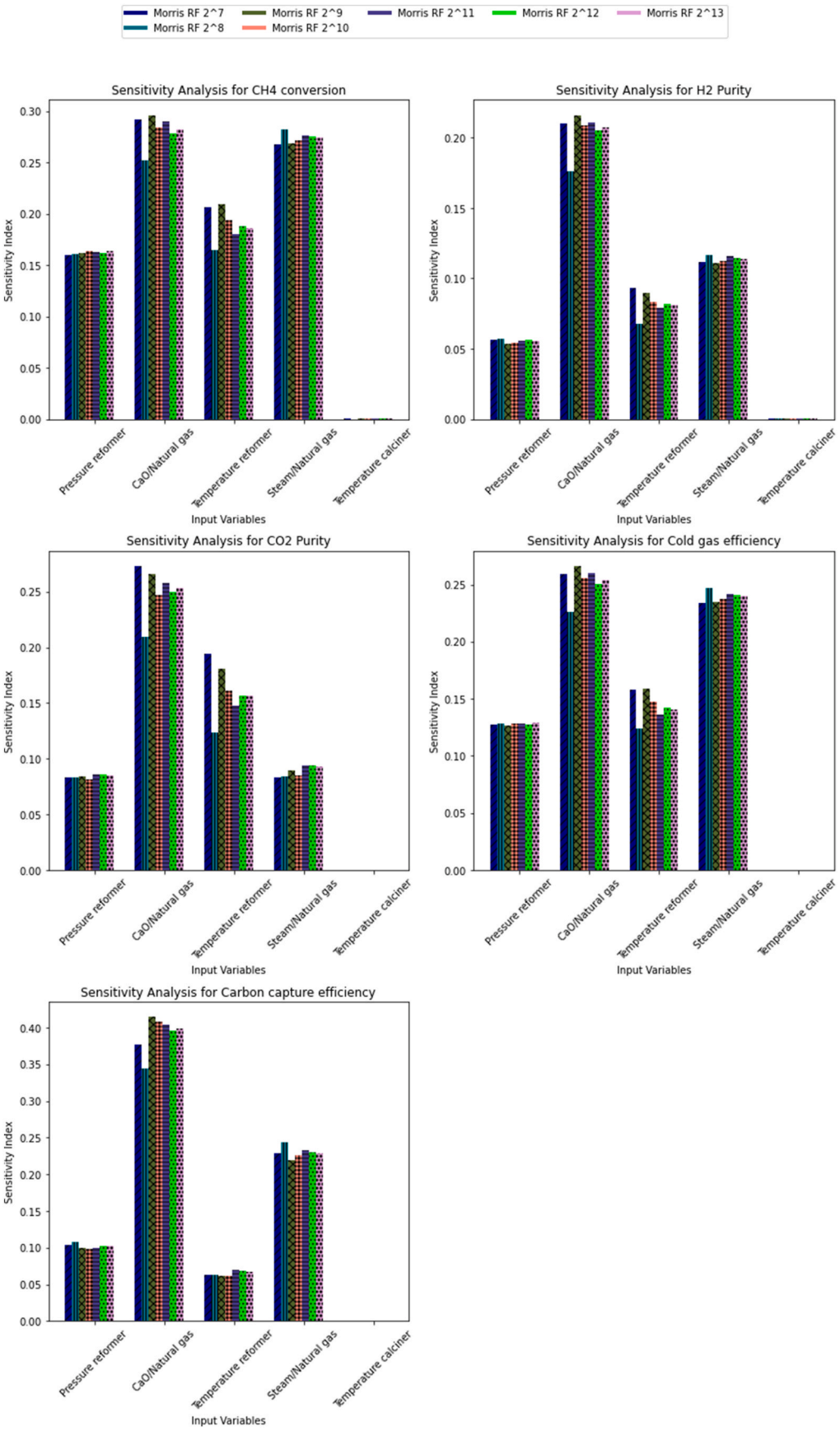


Fig. 17. (*continued*).

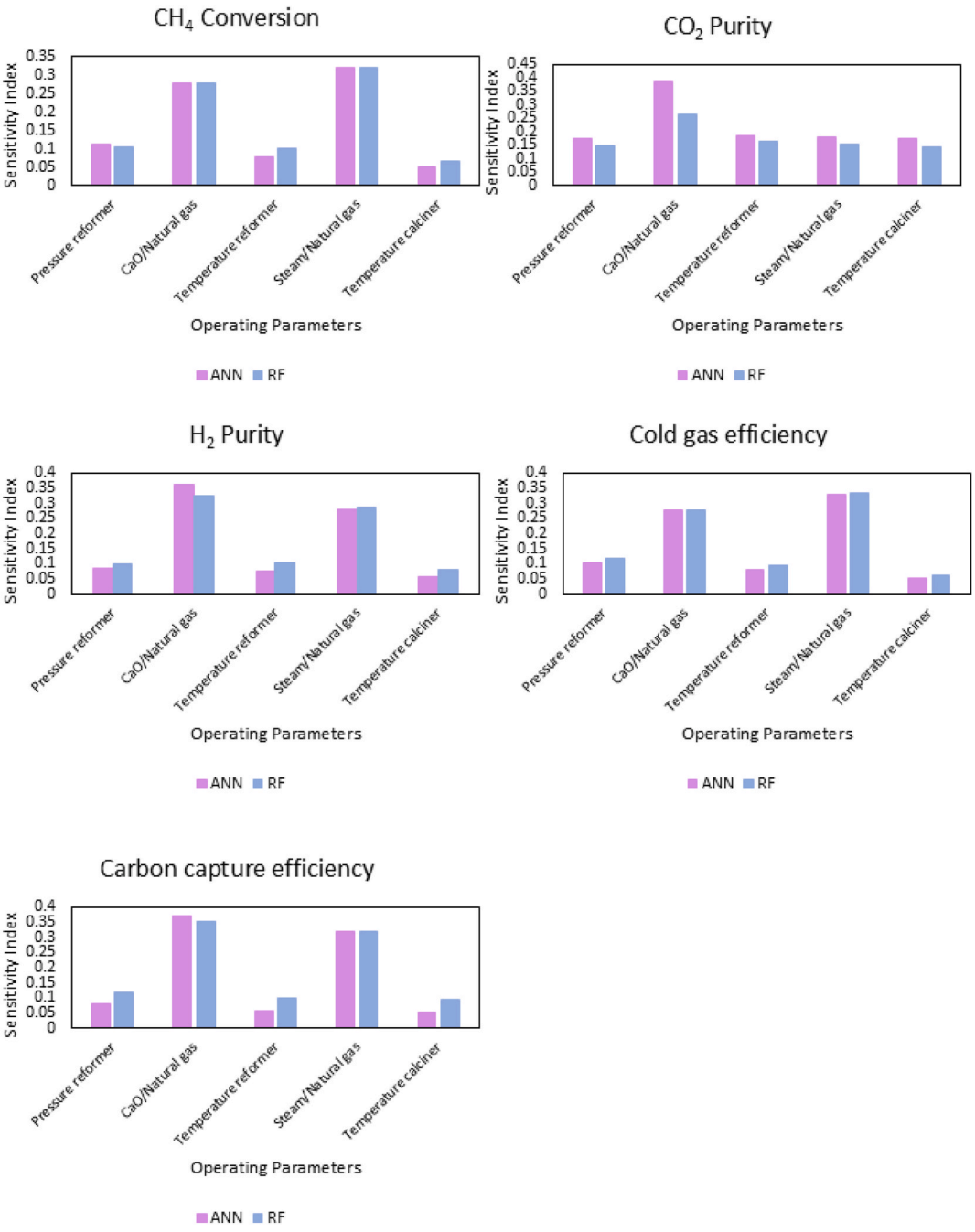


Fig. 18. First-order sensitivities calculated by the Delta approach for both ANN and RF.

2.3.2. *Morris method*

The Morris method [44], is a GSA approach. It is based on the concept of elementary effects (EE_i) which is calculated by Eq. (25). The distribution of the elementary effect is then obtained by sampling a point from the design space [45].

$$EE_i = \frac{[f(x_1, \ldots..., x_i \pm \Delta, \ldots, x_k) - f(x_1, \ldots, x_i, \ldots, x_k)\,]}{\Delta} \quad \text{Eq. 25}$$

Where f is the model output, x_i is the input parameter and Δ is the step size. The first sensitivity measure calculated by the Morris approach is the mean of absolute elementary effects μ_i^* which is an estimate of the mean of the distribution of EE_i, calculated via Eq. (26) [45].

$$\mu_i^* = \frac{1}{r}\sum_{j=1}^{r}\left|EE_i^{(j)}\right| \quad \text{Eq. 26}$$

Where $EE_i^{(j)}$ is the elementary effect of parameter i in the j-th trajectory, with r being the total number of trajectories. The other sensitivity measure calculated by the Morris approach is the standard deviation of elementary effects (σ_i) - an estimate calculated via Eq. (27) [45].

$$\sigma_i = \sqrt{\frac{1}{r-1}\sum_{j=1}^{r}\left(EE_i^{(j)} - \mu_i\right)^2} \quad \text{Eq. 27}$$

Where μ_i is the mean of EE_i (non-absolute).

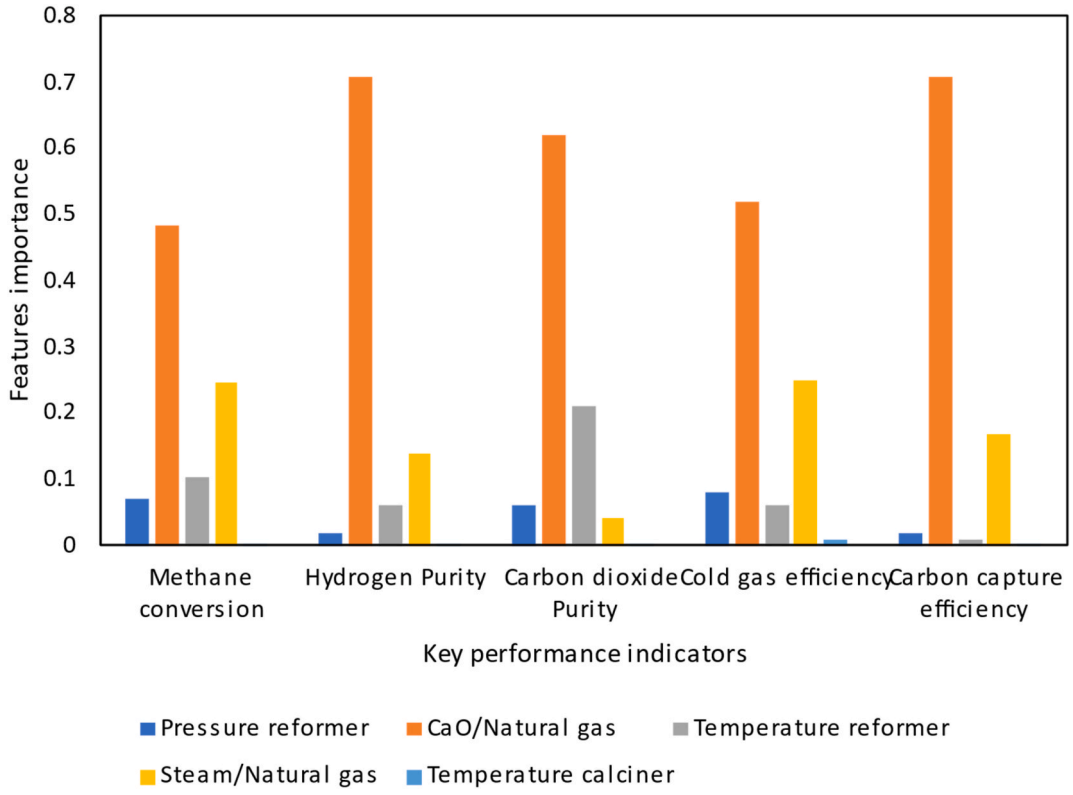


Fig. 19. Features importance for each input across the outputs from the RF model.

2.3.3. Delta

The Delta approach determines sensitivities that are independent of the sampling method used [46]. The main aim is to assess the influence of the entire input on the entire output without a reference to a particular moment within the output [46]. To calculate this sensitivity, Eq. (28) is used.

$$\delta_i = \frac{1}{2}\int f_{X_i}(x)\left[\int \left|f_Y(y) - f_{Y|Xi=x}(y)\right| dy\right] dx \quad \text{Eq. 28}$$

Where $f_{X_i}(x)$ is the marginal probability density of the input factor, and X_i, $f_Y(y)$ is the cumulative probability density function of the model output y. $f_{Y|Xi=x}(y)$ is the conditional density of Y given one of the parameters is a fixed value.

2.3.4. Model-based methods

These approaches all require an existing ML model in order to generate outputs to determine the impact of certain inputs on the outputs.

2.3.4.1. Shapley additive explanations. Shapley additive explanation (SHAP) utilises a game theory approach to explain the output of any ML-based model. It does this by using fair allocation results from cooperative game theory to allocate credit to a model output among its input features [39]. Shapley values are often used for an LSA. GSA can be performed using the SHAP approach by averaging the Shapley value over all instances of the dataset the model was trained on [39]. To calculate a Shapley value Eq. (29) is used.

$$g(z^i) = \phi_0 + \sum_{j=1}^{M} \phi_j z_j' \quad \text{Eq. 29}$$

Where g is the explanation model z' is the coalition vector/Shapley value, M is the maximum coalition size ϕ_j is the feature attribution.

2.3.4.2. Features importance of random forest. The RF ML approach allows for the direct calculation of the feature's importance without further integration with GSA approaches; it does this by utilising a function known as feature importance. The importance of this feature is the mean of the trees within the random forest, calculated as the decreasing out-of-bag (OOB) error of a tree when the values of the variable are randomly permuted in the OOB samples. For regression problems, the OOB error is measured by the mean squared error [35] (Eq. (30)).

$$\widehat{I}(X^J) = \frac{1}{ntree}\sum_{k=1}^{ntree}(y_i - \mu)^2 \quad \text{Eq. 30}$$

Where $\widehat{I}(X^J)$ is the feature importance of feature X^J, and *ntree* is the number of trees. μ is the mean.

2.3.5. Assessment of global sensitivity analysis approaches

The following approaches were assessed on the parameters of interpretability, robustness, and computational efficiency. The interpretability was determined by assessing how much information can be extracted from each approach, concerning the impact of each input on each output. Robustness was assessed by varying the sample size from 2^7 to 2^{13}. This was done to determine the robustness of the GSA approach when dealing with a small range of samples. The computational efficiency was calculated by determining the time taken for the GSA to calculate the importance of the features across the sample from 2^7 to 2^{13}. The cell was executed on an Intel Core i5-7300U CPU at 2.60 GHz with two cores and four threads.

3. Results and discussions

3.1. Machine learning surrogate model development

3.1.1. Neural network training and testing of the surrogate model

During the training of the ANN surrogate model, the hyperparameters were varied throughout the training. Table 7 presents the

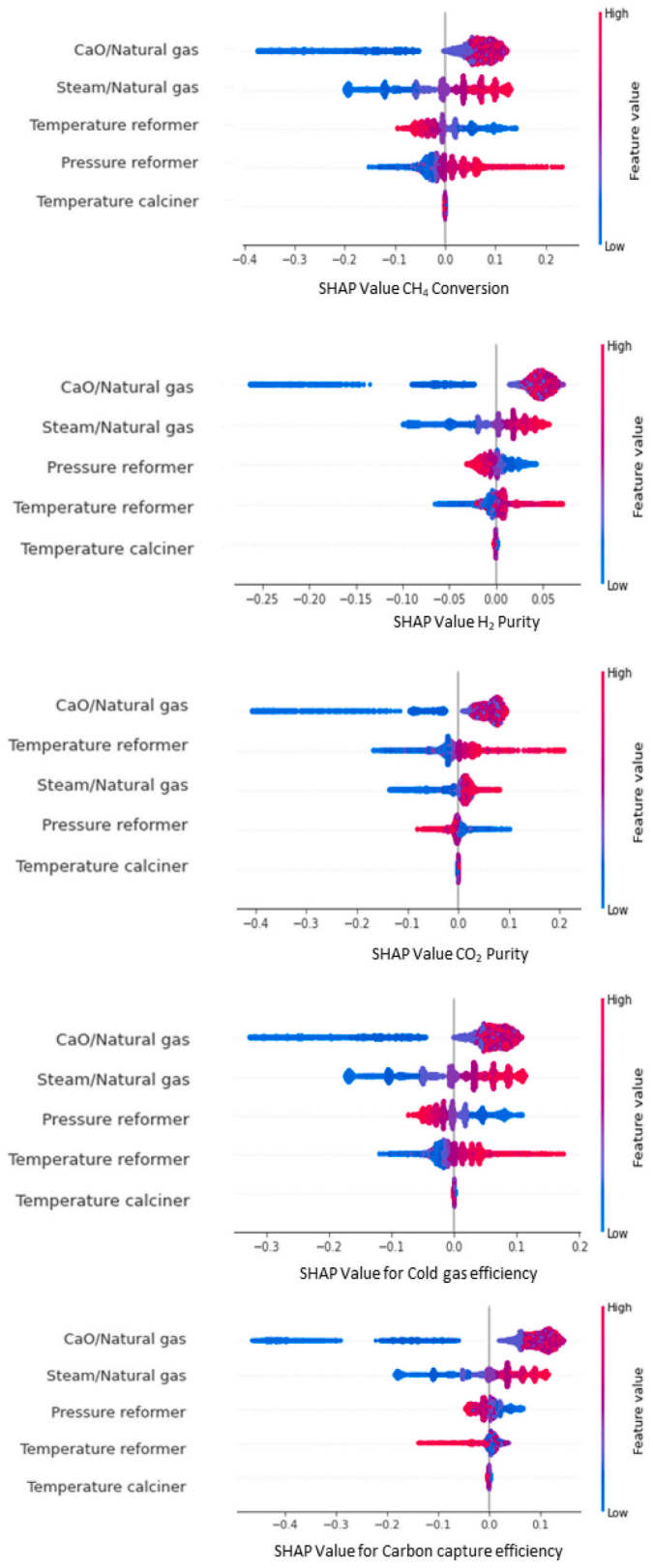


Fig. 20. Beeswarm plots for each output using the SHAP approach.

hyperparameter values for the best-performing model, specifically in terms of time taken per epoch and accuracy. Table 8 presents the optimal model results in terms of accuracy and time taken per epoch, as well as the average result obtained across different models studied.

As shown in Table 8, across the training and validation sets for the optimal model, there is high accuracy while maintaining good computational speed. The best performance model demonstrated high performance in both training and validation, exhibiting a low drop-off in the R^2 value, which highlights its robustness even when dealing with unseen data, suggesting that it does not overfit. Further evaluation metrics to compare the models are shown in the supplementary information. The best performing model although shows better performance than the average across all models, it is not statistically significant. Regardless, this was taken forward to evaluate on the test set. The best-performing model was then tested on the test set to ensure the model's robustness and accuracy when dealing with unseen data as shown in Table 9, which displays the R^2 values across the training, validation, and test datasets. As shown in Table 9, the test R^2 value remains high at 99.94 %, which highlights the model's generalisability in predicting outputs outside of the training set. For all outputs, the R^2 was calculated and plotted (Fig. 9). As shown in Fig. 9, across all 5 outputs, R^2 is high with no significant deviations.

3.1.2. *Random forest validation of the surrogate model*

The optimised hyperparameters in the RF surrogate model are shown in Table 10, together with the results of the accuracy and speed summarised in Table 11. These results show a slightly higher R^2 but lower MAE across the training and validation sets. The time taken for the training to be completed was significantly shorter because the training data went through the RF once. However, when comparing the training per epoch, the neural network is seen to be quicker. The training time helps to understand the training process; however, when deployed, the surrogate model is often used to predict single outputs and is unlikely to be used to predict as many samples as are provided within the training set. Further analysis of the computational efficiency is provided in section 3.3.

The RF and ANN show a similar R^2 across the training and validation sets. The optimised RF model based on the hyperparameters, shown in Table 10, was then evaluated on the test data set to ensure the models' robustness and accuracy when dealing with unseen data. Table 12 shows the R^2 of the test set. As shown in Table 12, the test R^2 is still high at 99.92 %, highlighting the model generalisability. As shown in Tables 9 and 12, the RF approach has a lower R^2 value than the ANN across the test set. This, combined with the increased speed of the ANN, comparing the training time of RF and ANN per Epoch, shows the ANN to have superior performance, although the RF is still highly accurate and can predict outputs at a high speed. Overall, for both approaches, the accuracy is high and the prediction is relatively quick, allowing for a large number of outputs to be generated during the training process. This is partially linked to the large volume of data collected (>30,000 data points). This ensures that the model is robust enough so that when it is tested on unseen data the surrogate model is highly accurate, leading to only a slight drop off with respect to the R^2. For all outputs, the R^2 was calculated and plotted (Fig. 10). As shown in Fig. 10, across all 5 outputs, R^2 is high with no significant deviations.

3.2. *Global sensitivity analysis evaluation of each approach*

3.2.1. *Sobol approach*

For the Sobol approach with an initial sample size of 1,024, the first-order sensitivities (Fig. 11) show good agreement between both the ANN and RF, achieving similar total order sensitivities across each output. As shown in Fig. 11, the first-order sensitivities show that the CaO/NG ratio significantly influences all the outputs, particularly influencing the purities of CO_2 and H_2 as well as the carbon capture efficiency. This is because of the role of CaO within the reformer, which is the removal of CO_2 through R4. This improves the purity of hydrogen, subsequently enhancing the production efficiency of hydrogen by shifting R3 to the right, which is confirmed by considering the impact the CaO/NG ratio on the system. The reformer temperature and pressure have small contribution to the outputs of the system with the calciner temperature

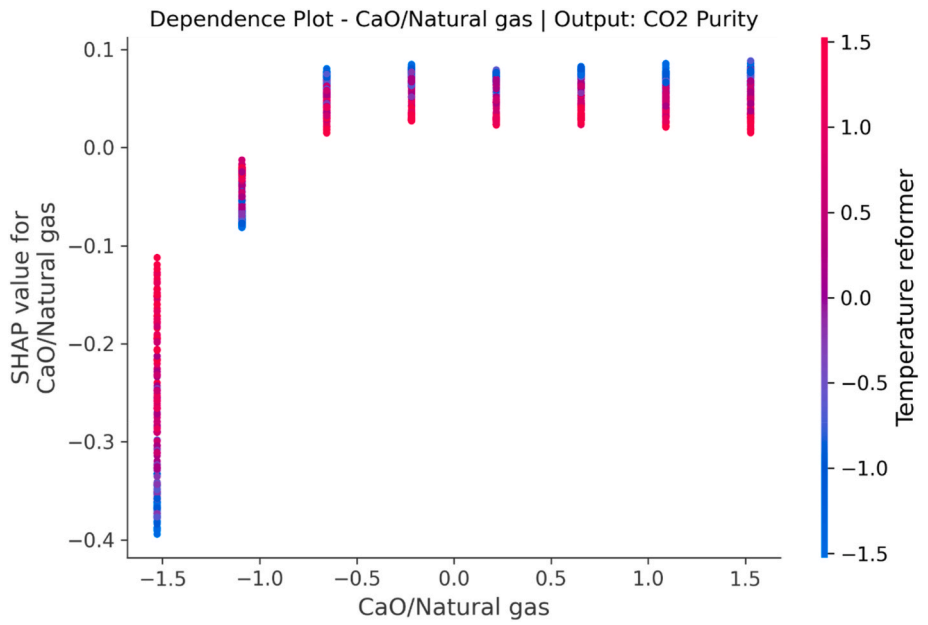


Fig. 21. SHAP dependence plot for the CaO/NG ratio and reformer temperature.

Fig. 22. Local sensitivity analysis of the ANN using the SHAP approach.

having little to no impact on the outputs. Through the Sobol approach, the second order sensitivities are calculated to assess interaction effects between parameters, these are shown in Fig. 12a–b.

As seen in Fig. 12a–b, there is interaction between different inputs. For example, there is a significant interaction between the reformer temperature and the CaO/NG ratio regarding the CO_2 purity. The Sobol approach provides both first- and second-order interaction effects. Although this provides a valuable method to assess input parameters, it still does not explain how an input impacts an output, only that interaction occurs between an input and an output; for example, the CaO/NG and reformer temperature show interactions that impact the CO_2 purity; however, there is no assessment of the value of these inputs and its impact on the output.

The sample size was varied for both models to assess the variation of the Sobol indices, that is calculated based on the sample size (Fig. 13a–b). Variation of the sample size shows that a lower sample size results in deviations in the calculated first-order sensitivity compared to the results from larger sample sizes, which maintain relatively consistent results once a sample size of 1,024 is reached. The smaller sample sizes do show an improved speed in calculating the Sobol first-order sensitivities. It is, however, important to ensure that the ranking of the input features in order of importance is not significantly altered. As shown in Fig. 13a, the operating parameters: reformer temperature and pressure, when at lower sample sizes are shown to switch in the order of importance on the CH_4 conversion KPI. However, both factors have no significant impact on the CH_4 conversion, and therefore, they are not considered significant factors. Overall, the Sobol approach provides a valuable method for assessing the input's first- and second-order sensitivities on each output; however, further analysis is required to determine how a value impacts the output.

3.2.2. *RBD-FAST approach*

Similarly to the Sobol approach, RBD-FAST approach assesses the contribution of an input to an output by the variance. As shown in Fig. 14, the sensitivity values for each input's impact on each output are displayed across both models. Fig. 14 shows a similar result to the first-order sensitivities calculated by the Sobol approach, highlighting the key influence of the CaO/NG across each output. The S/NG ratio also plays a key role in the CH_4 conversion, H_2 purity, and cold gas efficiencies. The robustness of the RBD-FAST was also assessed (Fig. 15a–b). Across both models, the RBD-FAST approach exhibits a similar sensitivity index for each input's impact on the output across all sampling sizes, thereby highlighting the robustness of this approach. However, the RBD-FAST approach, unlike the Sobol approach, does not provide a breakdown of the second-order sensitivities. Hence, there is only an understanding that interaction effects occur through the total order sensitivities, but not how these inputs interact with one another. This means it is helpful for an initial assessment of the system; however, further analysis of the system is required to gain an understanding of interaction and non-linear effects.

3.2.3. *Morris approach*

Morris approach provides two key outputs when assessing the model, which indicate the importance of a model input and its impact on predicting an output. The other output from the Morris method is the variance, which provides insight into how a parameter's effect varies over time. A high value indicates that the parameter's effect varies significantly, suggesting that the output is non-linear, or there is an interaction between them. Fig. 16a–b shows the μ_i^* vs σ_i across both the ANN (Fig. 16a) and the RF (Fig. 16b). As shown in Fig. 16a–b, the CaO/NG ratio is highly influential, suggesting that there are non-linear

ANN Surrogate Model Ranking of Inputs	Pressure	CaO/Natural gas	Temperature Reformer	Steam/Natural gas	Temperature Calciner
H_2 Purity Sobol	3	1	4	2	5
H_2 Purity RBD-FAST	3	1	4	2	5
H_2 Purity Delta	4	1	3	2	5
H_2 Purity Morris	4	1	3	2	5
CH_4 Conversion Sobol	3	1	4	2	5
CH_4 Conversion RBD-FAST	3	1	4	2	5
CH_4 Conversion Delta	3	2	4	1	5
CH_4 Conversion Morris	3	2	4	1	5
Cold gas efficiency Sobol	3	1	4	2	5
Cold gas efficiency RBD-FAST	3	1	4	2	5
Cold gas efficiency Delta	3	2	4	1	5
Cold gas efficiency Morris	3	2	4	1	5
CO_2 capture efficiency Sobol	3	1	4	2	5
CO_2 capture efficiency RBD-FAST	3	1	4	2	5
CO_2 capture efficiency Delta	3	1	4	2	5
CO_2 Capture Rate Morris	3	1	4	2	5
CO_2 Purity Sobol	4	1	2	3	5
CO_2 Purity RBD-FAST	4	1	2	3	5
CO_2Purity Delta	4	1	2	3	5
CO2 Purity Morris	4	1	2	3	5

Fig. 23. Ranking of first-order sensitivity across each approach. a) ANN. b) RF.

impacts on the outputs or the presence of interaction between the other parameters.

Furthermore, it is interesting to note that regarding CH_4 conversion and Cold gas efficiency, the Morris approach shows that the S/NG ratio has a higher μ_i^* than the CaO/NG ratio, suggesting that the S/NG has a greater influence on these outputs (and not the CaO/NG). This differs from the variance-based approaches (Sobol and RBD-FAST) which suggests the opposite based on the calculated first-order sensitivities shown in Figs. 11 and 14 for the Sobol and RBD-FAST approaches respectively. This variation in the ranking of feature importance across different GSA approaches is important to consider, as incorrect statements about the influence of input factors can potentially be made based on this analysis. Although Morris approach, through the σ_i, accounts for both non-linear and interaction effects between different inputs, it is unable to differentiate between the two, indicating the limitation of this approach. Even though it helps to screen variables and initial assessments, a more detailed analysis (e.g. Sobol) is required to quantify the second-order sensitivities.

Based on the variability of the feature's importance, the robustness of this approach is extremely important to assess the impact of sample size on μ_i^*. This is illustrated in Fig. 17a–b for the ANN (Fig. 17a) and RF (Fig. 17b). Variation is seen across both models as the sample size is varied. This is more pronounced for smaller sample sizes. This can suggest that the ranking of the features may vary across different sample sizes.

3.2.4. Delta approach

The delta approach reveals deviations in the importance of features compared to other approaches, as illustrated in Fig. 18. The steam-to-natural gas ratio has the most significant impact on CH_4 conversion and cold-gas efficiency. In contrast, Sobol and RBD-FAST suggest that the CaO-to-NG ratio has the most significant impact. These deviations are dependent on various factors and will be discussed in greater detail in Section 3.3.

3.2.5. Random forest feature importance

The use of features importance within the RF model provides a useful initial evaluation of RF inputs and the associated impact on the output. As shown in Fig. 19, the importance is not too dissimilar from the other approaches, indicating that CaO/NG has a greater impact on all the outputs within this process. However, similar to RBD-FAST and Delta, it only provides first-order effects. Like Morris, it is a valuable screening tool, but beyond that, it is unable to extract further information from the system.

3.2.6. SHAP approach

The SHAP approach provides a comprehensive method to assessing the inputs' features and the impacts on the outputs. As shown in Fig. 20, the Beeswarm plot not only shows the impact of an input (the range of each input) on the output but also how the value of an input (the colour of each point) affects the output. As illustrated in Fig. 20, the CaO/NG

RF Surrogate Model Ranking of Inputs	Pressure	CaO/Natural gas	Temperature Reformer	Steam/Natural gas	Temperature Calciner
H_2 Purity Sobol	4	1	3	2	5
H_2 Purity RBD-FAST	4	1	3	2	5
H_2 Purity Delta	4	1	3	2	5
H_2 Purity Morris	4	1	3	2	5
H_2 Purity Features Importance	4	1	3	2	5
CH_4 Conversion Sobol	3	1	4	2	5
CH_4 Conversion RBD-FAST	3	1	4	2	5
CH_4 Conversion Delta	3	2	4	1	5
CH_4 Conversion Morris	3	2	4	1	5
CH_4 Conversion Features Importance	3	1	4	2	5
Cold gas efficiency Sobol	3	1	4	2	5
Cold gas efficiency RBD-FAST	3	1	4	2	5
Cold gas efficiency Delta	3	1	4	2	5
Cold gas efficiency Morris	3	1	4	2	5
Cold gas efficiency Features Importance	4	1	3	2	5
Carbon capture efficiency Sobol	3	1	4	2	5
Carbon capture efficiency RBD-FAST	3	1	4	2	5
Carbon capture efficiency Delta	3	1	4	2	5
Carbon capture efficiency Morris	3	1	4	2	5
Carbon capture efficiency Features Importance	3	1	4	2	5
CO_2 Purity Sobol	3	1	4	2	5
CO_2 Purity RBD-FAST	3	1	4	2	5
CO_2 Purity Delta	3	1	4	2	5
CO2 Purity Morris	3	1	4	2	5
CO2 Purity Features Importance	4	1	3	2	5

Fig. 23. (*continued*).

ratio has the greatest impact on each of the output. Specifically, when the CaO/NG is lower, the output value is significantly reduced. Whereas at higher CaO/NG ratios, other factors also play a role in the output value. The SHAP approach provides an increased understanding of the system through this added dimension of the value of the input and the impact on the output; however, this comes with an increased computational cost; using a sample size of 500, the time taken to calculate the SHAP values was 384 s, which is significantly higher than the other approaches discussed in this work.

The SHAP analysis provides further insights into the ANN model through the SHAP dependence plots. Based on the previous analysis using the Sobol and Morris methods, there are interaction effects between inputs throughout the process. As shown in Fig. 12a, there is second-order interactions effects between the reformer temperature and the CaO/NG ratio that influence CO_2 purity, utilising the SHAP dependence plots, allows for further evaluation of these interaction effects as demonstrated in Fig. 21. As shown in Fig. 21, at lower CaO/NG ratios, the reformer temperature has a significant impact on the SHAP value. Increasing the reformer temperature significantly increases the SHAP value at lower CaO/NG ratios; however, as the CaO/NG ratio increases, this effect levels off, and varying the reformer temperature has a minimal impact on the SHAP value. Although at higher CaO/NG ratios, reducing the reformer temperature does increase the SHAP value of the CaO/NG, this effect is not significant, as it has a far greater impact at lower CaO/NG ratios. This is due to R4 being an exothermic reaction; once there is enough CaO to remove the CO_2 from the system, lower temperatures are preferred.

For local sensitivity analysis, the SHAP approach provides a breakdown of how a value of the input contributes to a specific output (Fig. 22). Fig. 22 provides a breakdown of each input contribution to the value, showing that the temperature of the reformer and the CaO/CH_4 ratio have a significant positive influence. The SHAP approach can provide a comprehensive assessment of the ML model by determining the SHAP value. It can determine how an input impacts the output

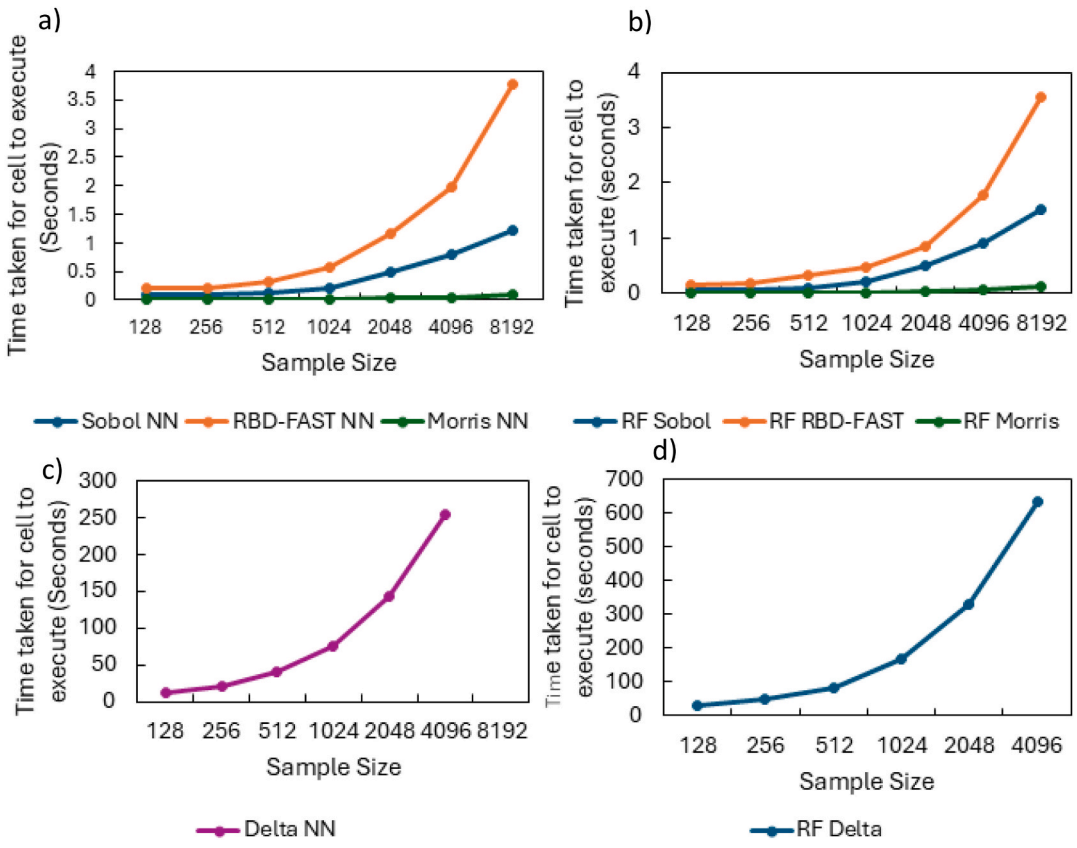


Fig. 24. Time taken for the GSA approaches to calculate sensitivity indices for a) ANN model with Sobol, RBD-FAST and Morris. b) RF model with Sobol, RBD-FAST and Morris. c) ANN model with Delta. d) RF model with delta.

within this ML model development. This approach can provide a valuable means for predictive maintenance by quantifying the impact of each input on the output value. The key drawback of this approach for calculating SHAP values is its computational intensity, especially with large datasets, as calculating the SHAP value is computationally slow.

3.3. *Comparative assessment of the global-sensitivity analysis approaches*

The rankings of each input parameter, indicating their importance in calculating the output, are shown in Fig. 23a–b for both the ANN and RF models, respectively. The variance-based methods demonstrate a similar importance of each feature across all outputs. In contrast, the Morris and Delta approaches exhibit some variation between specific outputs, typically when the impacts of each input are of comparable values. This variation depends on several factors; for example, the Sobol approach is a variance-based method, whereas the Morris method is derivative-based. These methods will effectively ascertain different aspects of sensitivity, potentially leading to different rankings. Furthermore, the sampling method used in a high-dimensional space may influence the rankings of these approaches. The Morris approach employs a trajectory-based sampling method, which may lead to under-sampling in high-dimensional spaces [45].

As shown in Figs. 13, 15 and 17 varying the sample size reveals variation in the calculated sensitivity index. However, with a sample size of more than 1,024 across all approaches, the sensitivity value remains consistent. The sample size is a crucial factor to consider when determining the time required to calculate the sensitivity index. The sample size used to calculate the sensitivity index was varied from 2^7 to 2^{13}, and the time required for the sensitivity index to execute was evaluated for both surrogate models. Fig. 24 shows the results for both the ANN and RF. As shown in Fig. 24, it appears that Delta is computationally extremely slow for both approaches; however, when combined with RF, it becomes especially sluggish. Across both surrogate models, the Morris approach is relatively rapid across all sample sizes, with the Sobol also demonstrating good computational efficiency across all sample sizes. For the Sobol and Morris methods with RF and ANN, there is little difference between the two, with the ANN being slightly faster. This is most likely because the optimised ANN has 50 neurons. In contrast, the optimised RF has 100 trees with no maximum tree depth, which explains why across these sample sizes, the execution speed of the model was slightly slower for the RF approach. Both the Sobol and Morris approaches exhibit excellent computational speed. When Sobol is applied with a large sample size (>1,024), the robustness for the first-order sensitivities remains consistent.

Overall, this is a clear indication that the SHAP approach provides increased interpretability, showcasing not only whether an input has a significant impact on the output value but also how the value of an input affects the output, providing a comprehensive assessment of the surrogate model regarding how an input influences the output. This comes at the cost of speed, taking a considerable amount of time to fully execute. Reducing the sample size can improve speed; however, ensuring that the developed surrogate model maintains high accuracy is essential to guarantee that the SHAP values are reliable. Further methods can be implemented to improve the efficiency. For example, recent work by

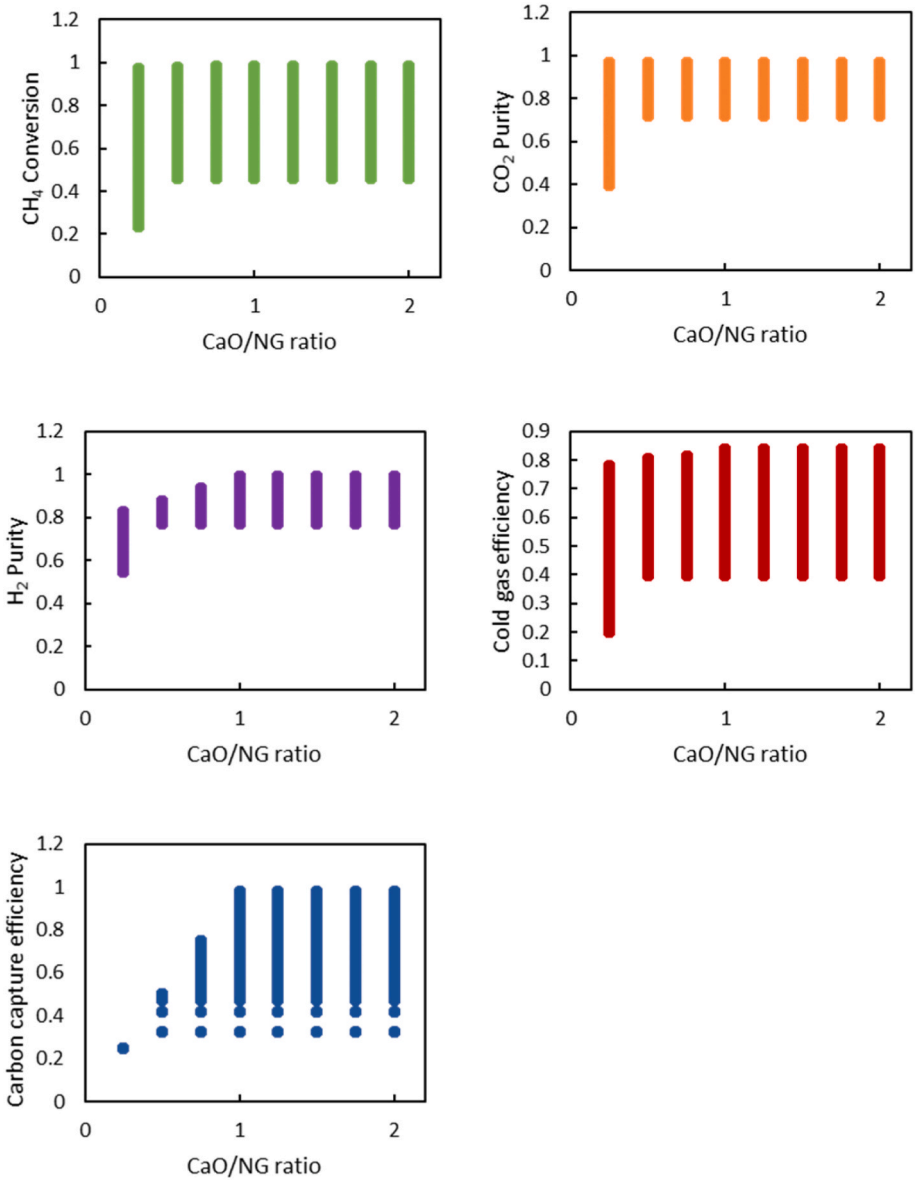


Fig. 25. Univariate analysis of CaO/NG ratio on each KPI.

Ranjbaran et al. (2025) combined K-means clustering with the SHAP approach which reduced the computational time to execute from 400 s to 0.29 s [47]. This approach can provide interpretability whilst ensuring it is still computationally efficient. To ensure computational speed while still providing interpretability, the Sobol and Morris approaches are capable of calculating both first- and second-order effects (although the Morris cannot differentiate between second-order and non-linear interactions). While the RBD-FAST method is computationally efficient, it cannot extend beyond first-order effects, and the Delta approach is also slow and does not progress past first-order effects. Consequently, the combined robustness is not particularly strong. Often, within the literature, feature importance from the RF package is used due to its simplicity and ease in determining the importance of each feature for each output. Although it offers a straightforward method, similar to the Delta approach, it is unable to surpass first-order effects.

Ensuring a large sample size (>1,024) guarantees that the sensitivity values will not vary significantly, with both the RBD-FAST and Morris approaches demonstrating excellent consistency in the sensitivity index value. The Morris approach calculates the μ* and σ indices at significantly faster speeds than the other approaches. The Sobol approach, similar to the Morris approach, can also compute these values swiftly at large sample sizes. The SHAP approach enhances interpretability by illustrating how each input value affects the output. This approach is beneficial for local-sensitivity analysis, as shown in Fig. 22 through the force-plot which for individual output values, it is able to determine how each input contributes to predicting the output. This can be particularly useful when assessing outliers to establish how the inputs affect the output and which factors within those inputs lead to the output value. This application is highly relevant in research areas such as predictive maintenance [48].

As shown across each approach the CaO/NG ratio influences each of the KPIs significantly by removing the CO_2 in-situ it provides two key advantages: the removal of CO_2 ensuring higher CCE and CO_2 purity, and removing CO_2 means that the reforming reaction shifts to the right

to produce more hydrogen improving the CH_4 conversion, cold-gas efficiency and H_2 purity. higher CaO/NG ratios ensures a higher CCE as well as CGE and H_2 purity, as it is such a significant influence across all KPIs a univariate analysis was done to assess each output metric individually which is shown in Fig. 25.Fig. 25 further highlights what the SHAP analysis shows extremely low values of the CaO/NG ratio (<0.5) significantly lowers the performance of the process across all KPIs above this ratio the performance of these KPIs is improved although the variation within these values still shows that other factors within the process such as the S/C and unit operators operating parameters such as temperature and pressure of the reformer play a role within ensuring the KPIs are optimised. To ensure a CCE above 75 % a CaO/NG ratio of 1.00 is required but again other factors play a role to maintain that high CCE otherwise it can be significantly reduced.

4. Challenges and limitations

Using RF and ANN as a preliminary step for global sensitivity analysis can be effective, provided it is done with careful attention to model validation, interpretability, and integration with traditional GSA methods. In this work, a comprehensive assessment of various GSA approaches has been carried out. Although significant progress has been made in developing a thorough evaluation of these GSA methods and integration with ML techniques, key challenges persist in this area of research, particularly regarding the enhanced interpretability of surrogate ML models for blue hydrogen production. Further work must be undertaken on the entire system, such as the CLC system and PSA, to examine how these systems impact the KPIs described. This, in turn, increases the complexity of surrogate modelling, which may affect both the computational efficiency and the accuracy of the model.

The expansion of the dataset to include variations in the feed-gas composition and activity of the catalyst is key to further assess these parameters and their influence on the output, e.g. by incorporating kinetic data such as catalyst activity and its impact on the KPIs within this work. In this work, we have highlighted how developing a surrogate ML model can enhance the speed of the modelling. Future work should focus on the introduction of additional operating parameters such as feed gas composition and kinetic parameters, in order to paint a more comprehensive picture of the process.

Limitation of this current work is that this data-driven model is trained on first-principle datasets, which is based on a steady state thermodynamic process model; it does not include kinetic parameters or dynamic-specific elements. This absence of additional parameters in the process means that when developing a data-driven approach, trained on a first-principle model, the full complexity of the system is not captured. The recent advancement of neural network architectures, such as transformers, allows for dynamic processes to be modelled with great accuracy and has proven effective for optimisation in dynamic processes, including adsorption-based processes such as PSA for H_2 purification ([49–51]). If this approach would to be extended for such dynamic scenarios (e.g. reduction in performance of the CaO over multiple cycles), further operational parameter must be included to assess how the behaviour of the system shapes over time.

The development of neural network architecture has led to the creation of physics-informed neural networks (PINNs) for modelling unit operations and processes in environments with low data availability, with increasing popularity ([52,53]). Integrating with PINNs using GSA approaches would help in understanding the key input factors driving this process, potentially reducing the dimensions of the data-driven model output while still capturing the key factors that drive the process.

The enhanced use of SHAP can be an important tool in science and engineering. It provides increased knowledge of the surrogate model system by offering a comprehensive toolkit capable of assessing the model, determining how an input value impacts the output, as well as examining interaction factors through the SHAP dependence plot and individual inputs via the LSA function. However, in this work, this has solely been employed to evaluate feature impacts since the SHAP-based approach can break down how an input value affects the output. Once the SHAP values are calculated, they could be integrated with optimisation methods to guide the solver towards solutions. Although the computational efficiency of the SHAP approach is a key concern, incorporating the SHAP approach with efficient sampling methods has shown to improve the computational efficiency. Further investigation into the effects of sampling and its impact on accuracy is necessary to determine the applicability of the SHAP approach and its suitability for applications within process control, predictive maintenance, and real-time optimisation.

5. Conclusions

In this work, various GSA and LSA approaches were integrated with ML surrogate models trained on data collected from a first-principle SE-SMR-CLC process model. The development of surrogate ML models for predicting key performance indicators in the SE-SMR-CLC process demonstrates that, with a large number of data points, both the ANN and RF can exhibit high predictive accuracy - even within the test set. This highlights the model's generalisability. A comparison of different GSA approaches reveals that the CaO/NG ratio has the most significant impact on the key performance indicators (i.e. CH_4 conversion, CCE, CGE, CO_2 purity and H_2 purity), emphasising that this intensification method not only provides a means to capture CO_2 but also enhances the performance of the hydrogen production system. This impact is due to the removal of CO_2 in-situ within the reformer; this shifts the thermodynamic equilibrium so that more hydrogen is produced whilst simultaneously capturing CO_2 from the system. This combined approach harnesses the benefits of both ML and classical sensitivity analysis, offering a robust foundation for understanding and enhancing complex systems such as blue hydrogen production. Integrating GSA with ML provides interpretability to surrogate models developed through ML techniques, such as ANN and RF. Depending on the GSA approach, they can ascertain a range of impacts on the system, including both first-order and second-order effects. In particular, the Sobol approach can compute both first- and second-order effects with a large sample size in less than 1 s. The SHAP approach to assessing feature importance provides a method that enables not only the determination of whether an input affects an output, but also the value of an input and its influence on the output.

Future work will further expand implement this approach on systems with greater complexity (kinetic and dynamic considerations), as well as implementing strategies to improve the computational efficiency of the SHAP approach. This work has shown that the integration of surrogate models with GSA in the modelling of chemical processes can enhance computational efficiency, provided that the appropriate GSA approach is selected. As more data is generated, effectively utilising this to improve the modelling and development of low-carbon processes is crucial to accelerating the development and deployment speed of these processes, enabling a net-zero future.

CRediT authorship contribution statement

William George Davies: Writing – original draft, Visualization, Validation, Investigation, Formal analysis. **Paulina Quintanilla:** Methodology, Investigation, Formal analysis, Conceptualization. **Yang Yang:** Visualization, Supervision, Resources. **Salman Masoudi Soltani:** Writing – review & editing, Supervision, Resources, Project administration, Funding acquisition, Conceptualization.

Data availability statement

The data generated in this work is made available at Brunel Figshare database at https://doi.org/10.17633/rd.brunel.29478566.v1.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

The research presented in this work has received financial support from the UK Engineering and Physical Sciences Research Council (EPSRC) through the EPSRC Doctoral Training Partnerships (DTP) award, EP/T518116/1 (project reference: 2688399).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijhydene.2025.152153.

References

[1] Calvin K, et al. "IPCC, 2023: climate change 2023: synthesis report. In: Lee H, Romero J, editors. Contribution of working groups I, II and III to the sixth assessment report of the intergovernmental panel on climate change [core writing team. IPCC; Jul. 2023. https://doi.org/10.59327/IPCC/AR6-9789291691647. Geneva, Switzerland.
[2] Husarek D, Schmugge J, Niessen S. Hydrogen supply chain scenarios for the decarbonisation of a German multi-modal energy system. Int J Hydrogen Energy Nov. 2021;46(76):38008–25. https://doi.org/10.1016/J.IJHYDENE.2021.09.041.
[3] Anandarajah G, McDowall W, Ekins P. Decarbonising road transport with hydrogen and electricity: long term global technology learning scenarios. Int J Hydrogen Energy Mar. 2013;38(8):3419–32. https://doi.org/10.1016/J.IJHYDENE.2012.12.110.
[4] Neuwirth M, Fleiter T, Manz P, Hofmann R. The future potential hydrogen demand in energy-intensive industries - a site-specific approach applied to Germany. Energy Convers Manag Jan. 2022;252:115052. https://doi.org/10.1016/J.ENCONMAN.2021.115052.
[5] I. - International Energy Agency. Global hydrogen review 2023 [Online]. Available: www.iea.org. [Accessed 18 September 2024].
[6] IEA. "Net Zero by 2050," Paris [Online]. Available: https://www.iea.org/reports/net-zero-by-2050. [Accessed 12 January 2025].
[7] IEA. Global hydrogen review 2024 [Online]. Available: https://www.iea.org/reports/global-hydrogen-review-2024. [Accessed 13 January 2025].
[8] Masoudi Soltani S, Lahiri A, Bahzad H, Clough P, Gorbounov M, Yan Y. Sorption-enhanced steam methane reforming for combined CO2 capture and hydrogen production: a state-of-the-art review. Carbon Capture Sci Technol Dec. 2021;1:100003. https://doi.org/10.1016/j.ccst.2021.100003.
[9] Bahzad H, Katayama K, Boot-Handford ME, Mac Dowell N, Shah N, Fennell PS. Iron-based chemical-looping technology for decarbonising iron and steel production. 2019. https://doi.org/10.1016/j.ijggc.2019.06.017.
[10] Pan I, Mason LR, Matar OK. Data-centric engineering: integrating simulation, machine learning and statistics. Challenges and opportunities. Chem Eng Sci Feb. 2022;249:117271. https://doi.org/10.1016/J.CES.2021.117271.
[11] Davies WG, Babamohammadi S, Yang Y, Soltani SM. The rise of the machines: a state-of-the-art technical review on process modelling and machine learning within hydrogen production with carbon capture. 2023. https://doi.org/10.1016/j.jgsce.2023.205104.
[12] Dobbelaere MR, Plehiers PP, Van de Vijver R, Stevens CV, Van Geem KM. Machine learning in chemical engineering: strengths, weaknesses, opportunities, and threats. Engineering Sep. 2021;7(9):1201–11. https://doi.org/10.1016/j.eng.2021.03.019.
[13] Sansana J, et al. Recent trends on hybrid modeling for industry 4.0. Elsevier Ltd.; Aug. 01, 2021. https://doi.org/10.1016/j.compchemeng.2021.107365.
[14] Lee D, Chen W. Tuba dolar data-driven global sensitivity analysis of variable groups for understanding complex physical interactions in engineering design. 2024. https://doi.org/10.1115/1.4064633.
[15] Van Stein B, Raponi E, Sadeghi Z, Bouman N, Van Ham RCHJ, Back T. A comparison of global sensitivity analysis methods for explainable AI with an application in genomic prediction. IEEE Access 2022;10:103364–81. https://doi.org/10.1109/ACCESS.2022.3210175.
[16] Quintanilla P, et al. Artificial intelligence and robotics in the hydrogen lifecycle: a systematic review. Int J Hydrogen Energy Mar. 2025;113:801–17. https://doi.org/10.1016/J.IJHYDENE.2025.03.016.
[17] Nkulikiyinka P, Yan Y, Güleç F, Manovic V, Clough PT. Prediction of sorption enhanced steam methane reforming products from machine learning based soft-sensor models. Energy and AI 2020;2(Nov). https://doi.org/10.1016/j.egyai.2020.100037.
[18] Vo ND, Kang JH, Oh DH, Jung MY, Chung K, Lee CH. Sensitivity analysis and artificial neural network-based optimization for low-carbon H2 production via a sorption-enhanced steam methane reforming (SESMR) process integrated with separation process. Int J Hydrogen Energy Jan. 2022;47(2):820–47. https://doi.org/10.1016/j.ijhydene.2021.10.053.
[19] Jollife IT, Cadima J. Principal component analysis: a review and recent developments. Philos Trans R Soc A Math Phys Eng Sci Apr. 2016;374(2065). https://doi.org/10.1098/RSTA.2015.0202.
[20] Zhao S, Li J, Chen C, Yan B, Tao J, Chen G. Interpretable machine learning for predicting and evaluating hydrogen production via supercritical water gasification of biomass. J Clean Prod 2021;316(Sep). https://doi.org/10.1016/j.jclepro.2021.128244.
[21] Huang J, Liang Z, Liu Y. Smart reforming for hydrogen production via machine learning. Chem Eng Sci Feb. 2025;304:120959. https://doi.org/10.1016/J.CES.2024.120959.
[22] Lucay FA. Accelerating global sensitivity analysis via supervised machine learning tools: case studies for mineral processing models. Minerals Jun. 2022;12(6). https://doi.org/10.3390/min12060750.
[23] Zhao T, Zheng Y, Wu Z. Improving computational efficiency of machine learning modeling of nonlinear processes using sensitivity analysis and active learning. Digital Chemical Engineering Jun. 2022;3:100027. https://doi.org/10.1016/J.DCHE.2022.100027.
[24] Davies WG, Babamohammadi S, Yan Y, Clough PT, Soltani SM. Exergy analysis in intensification of sorption-enhanced steam methane reforming for clean hydrogen production: comparative study and efficiency optimisation. Carbon Capture Sci Technol 2024;12:100202. https://doi.org/10.1016/j.ccst.2024.100202.
[25] Lopez-Echeverry JS, Reif-Acherman S, Araujo-Lopez E. Peng-Robinson equation of state: 40 years through cubics. Fluid Phase Equilib Sep. 2017;447:39–71. https://doi.org/10.1016/J.FLUID.2017.05.007.
[26] Babamohammadi S, Davies WG, Masoudi Soltani S. Probing into the interactions among operating variables in blue hydrogen production: a new approach via design of experiments (DoE). Gas Sci Eng Sep. 2023;117:205071. https://doi.org/10.1016/J.JGSCE.2023.205071.
[27] Yan Y, et al. Process simulations of blue hydrogen production by upgraded sorption enhanced steam methane reforming (SE-SMR) processes. Energy Convers Manag Oct. 2020;222:113144. https://doi.org/10.1016/J.ENCONMAN.2020.113144.
[28] Isolation forest guide: explanation and python implementation | DataCamp [Online]. Available: https://www.datacamp.com/tutorial/isolation-forest?dc_referrer=https%3A%2F%2Fwww.google.com%2F. [Accessed 21 August 2025].
[29] Pedregosa F, et al. Scikit-learn: machine learning in python. J Mach Learn Res 2011;12(85):2825–30 [Online]. Available: http://jmlr.org/papers/v12/pedregosa11a.html. [Accessed 12 November 2024].
[30] Abadi M, et al. TensorFlow: large-scale machine learning on heterogeneous distributed systems [Online]. Available: https://arxiv.org/abs/1603.04467v2. [Accessed 12 November 2024].
[31] Janiesch C, Zschech P, Heinrich K. Machine learning and deep learning. 2021. https://doi.org/10.1007/s12525-021-00475-2/Published.
[32] Yan Y, et al. Harnessing the power of machine learning for carbon capture, utilisation, and storage (CCUS)-a state-of-the-art review. Royal Society of Chemistry Dec. 01, 2021. https://doi.org/10.1039/d1ee02395k.
[33] Sarker IH. Deep learning: a comprehensive overview on techniques, taxonomy, applications and research directions. SN Comput Sci Nov. 2021;2(6):1–20. https://doi.org/10.1007/S42979-021-00815-1/FIGURES/11.
[34] Tune hyperparameters and layers of neural networks. https://www.analyticsvidhya.com/blog/2021/05/tuning-the-hyperparameters-and-layers-of-neural-network-deep-learning/. [Accessed 22 August 2025].
[35] Antoniadis A, Lambert-Lacroix S, Poggi JM. Random forests for global sensitivity analysis: a selective review. Reliab Eng Syst Saf Feb. 2021;206:107312. https://doi.org/10.1016/J.RESS.2020.107312.
[36] Breiman L. Random forests. Mach Learn Oct. 2001;45(1):5–32. https://doi.org/10.1023/A:1010933404324/METRICS.
[37] Hyperparameter tuning the random forest in python | towards data science. https://towardsdatascience.com/hyperparameter-tuning-the-random-forest-in-python-using-scikit-learn-28d2aa77dd74/. [Accessed 22 August 2025].
[38] Herman J, Usher W. SALib: an open-source python library for sensitivity analysis. J Open Source Softw Jan. 2017;2(9):97. https://doi.org/10.21105/JOSS.00097.
[39] Lundberg SM, Lee SI. A unified approach to interpreting model predictions. Adv Neural Inf Process Syst 2017-December:4766–75 [Online]. Available: https://arxiv.org/abs/1705.07874v2. [Accessed 1 November 2024].
[40] Saltelli A, et al. Global sensitivity analysis. The primer," *global sensitivity analysis*. The Primer; Jan. 2008. p. 1–292. https://doi.org/10.1002/9780470725184.
[41] Van Stein B, Raponi E, Sadeghi Z, Bouman N, Van Ham RCHJ, Back T. A comparison of global sensitivity analysis methods for explainable AI with an application in genomic prediction. IEEE Access 2022;10:103364–81. https://doi.org/10.1109/ACCESS.2022.3210175.
[42] Sobol IM. Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. Math Comput Simul Feb. 2001;55(1–3):271–80. https://doi.org/10.1016/S0378-4754(00)00270-6.
[43] Tarantola S, Gatelli D, Mara TA. Random balance designs for the estimation of first order global sensitivity indices. Reliab Eng Syst Saf Jun. 2006;91(6):717–27. https://doi.org/10.1016/J.RESS.2005.06.003.
[44] Morris MD. Factorial sampling plans for preliminary computational experiments. Technometrics May 1991;33(2):161. https://doi.org/10.2307/1269043.
[45] Kucherenko S, Iooss B. Derivative based global sensitivity measures. Handbook of uncertainty quantification. Dec. 2014. p. 1–24. https://doi.org/10.1007/978-3-319-11259-6_36-1.
[46] Plischke E, Borgonovo E, Smith CL. Global sensitivity measures from given data. Eur J Oper Res May 2013;226(3):536–50. https://doi.org/10.1016/J.EJOR.2012.11.047.

[47] Ranjbaran G, Recupero DR, Roy CK, Schneider KA. C-SHAP: a hybrid method for fast and efficient interpretability. Appl Sci Jan. 2025;15(2):672. https://doi.org/10.3390/APP15020672.
[48] Steurtewagen B, Van den Poel D. Adding interpretability to predictive maintenance by machine learning on sensor data. Comput Chem Eng 2021;152(Sep). https://doi.org/10.1016/j.compchemeng.2021.107381.
[49] Yu X, Shen Y, Guan Z, Zhang D, Tang Z, Li W. Multi-objective optimization of ANN-based PSA model for hydrogen purification from steam-methane reforming gas. Int J Hydrogen Energy Mar. 2021;46(21):11740–55. https://doi.org/10.1016/j.ijhydene.2021.01.107.
[50] Yang Y, Zhang Q, Feng X. An innovative gas management methodology based on PSA for efficient gas allocation and utilization in hybrid hydrogen network: integrating process simulation, modeling, and machine learning. Int J Hydrogen Energy 2024;57:224–39. https://doi.org/10.1016/j.ijhydene.2024.01.009.
[51] Subraveti SG, Li Z, Prasad V, Rajendran A. Machine learning-based multiobjective optimization of pressure swing adsorption. Ind Eng Chem Res Nov. 2019;58(44):20412–22. https://doi.org/10.1021/acs.iecr.9b04173.
[52] Kharazmi E, Zhang Z, Karniadakis GE. Variational physics-informed neural networks for solving partial differential equations [Online]. Available: http://arxiv.org/abs/1912.00873. [Accessed 19 August 2024].
[53] G. E. Karniadakis, I. G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang, and L. Yang, "Physics-informed machine learning nature reviews | Physics", doi: 10.1038/s42254-021-00314-5.