Chemical Engineering Journal 426 (2021) 131639

Contents lists available at ScienceDirect

Chemical Engineering Journal

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

Machine learning based predictive model for methanol steam reforming with technical, environmental, and economic perspectives

Manhee Byun [a], Hyunjun Lee [a], Changgwon Choe [a], Seunghyun Cheon [a], Hankwon Lim [a,b,*]

[a] School of Energy and Chemical Engineering, Ulsan National Institute of Science and Technology, Ulsan 44919, Republic of Korea
[b] Department of Energy Engineering, Ulsan National Institute of Science and Technology, 50 UNIST-gil, Eonyang-eup, Ulju-gun, Ulsan 44919, Republic of Korea

ARTICLE INFO

Keywords:
H_2 production
Methanol steam reforming
Membrane filter
Feasibility study
Techno, economic, and environmental analysis
Machine learning based predictive model

ABSTRACT

To overcome limitations of conventional H_2 production approaches such as steam methane reforming (SMR) in a membrane reactor (MR) such as large CO_2 emission and deactivation of catalyst and membrane, a promising alternative H_2 production system of methanol steam reforming (MSR) in serial reactors and membrane filters is reported here, affording its high product yield and a compact design. In this study, technical, environmental, and economic feasibility according to 12 techno-economic parameters and detailed effects of each parameter for this H_2 production system are comprehensively investigated with a machine learning (ML) based predictive model in the following steps: (1) process simulation using Aspen Plus® with detailed thermodynamic phenomena and environmental performance; (2) numerical model using MATLAB® based on technical and environmental performance from the process simulation results; (3) ML-based predictive model having outputs of H_2 production rate, CO_2 emission, and unit H_2 production cost feasibility trained by 12,000 data sets from a numerical model. It is well noted from this study that # of reactors and operating temperature for technical performance, # of reactors and S/C ratio for environmental performance, and operating temperature, # of reactors, reactant, and labor for economic performance are reported as most influential factors.

1. Introduction

As various policies of transition to H_2 society where H_2 is mainly used as fuels continuously increase, there have been many researches for H_2 production methods [1,2]. Reforming of various hydrocarbons and coal gasification were proposed as commercialized H_2 production methods accounting for a very-high portion of the total amount of H_2 produced [3–6], and researches for systematic designs such as membrane reactor also have been followed [7–9]. Especially, methanol steam reforming (MSR) has received much attention due to various advantages: (a) transportation of reactant is relatively easy with existing infrastructure; (b) it doesn't require additional sulfur removal process; (c) it is operated in relatively low temperature compared to steam methane reforming (SMR), widely used in H_2 production, leading to lower heat required and carbon emissions [10–12].

In addition, the H_2 production system of reforming with H_2 permeable membrane filters, serial packed-bed reactors and membrane filters, using *Le Chatelier*'s principle has been introduced as an alternative to a membrane reactor (MR) to overcome its own disadvantages such as deactivation of catalyst and membrane caused by coke formation, less mass transfer leading to higher membrane area required, reduced permeance by liquids, and difficult replacement of inside materials [13–16].

H_2 is treated as a clean energy alternative to fossil fuels based on the following characteristics: (a) it is the most abundant element in earth meaning that it can't be depleted; (b) it has a very high energy density of 143 MJ kg^{-1}, which is double of other conventional fuels; (c) it can be produced by various renewable energy sources such as solar, wind, hydro, and biomass, etc. as electricity storable form leading to reduction of environmental concerns; (d) it produces only water as a byproduct during the generation of energy through combustion [17–22].

However, technologies related with H_2 energy is premature now, expressed as low technology readiness level (TRL) [23], so it is crucial to investigate technical, environmental, and economic feasibility of a certain premature technology to accomplish its commercialization and sustainability: Moura et al. [24] reported possible H_2 production of 40.3 $kgH_2\ d^{-1}$ with 1101.91 kW solar-wind-biogas hybrid H_2 production system and its economic and environmental feasibility based on H_2 production cost range of 0.96–0.16 \$ kWh^{-1} and ecological efficiency of 0.87; Acar and Dincer [25] ranked various for H_2 production source/system and storage options in aspects of technical, environmental,

* Corresponding author.
E-mail address: hklim@unist.ac.kr (H. Lim).

https://doi.org/10.1016/j.cej.2021.131639
Received 4 February 2021; Received in revised form 25 May 2021; Accepted 26 July 2021
Available online 10 August 2021
1385-8947/

Nomenclature

MSR	Methanol steam reforming
SMR	Steam methane reforming
MR	Membrane reactor
TRL	Technology readiness level
PV	Photovoltaic
COGH	Coke oven gas-based hydrogen
CGH	Coal gasification-based hydrogen
WE	Water electrolysis
ESR	Ethane steam reforming
AI	Artificial intelligence
IoT	Internet of things
ML	Machine learning
ANN	Artificial neural network
SVR	Support vector regression
BFD	Block flow diagram
PSA	Pressure swing adsorption
CEPCI	Chemical engineering plant cost index
CRF	Capital recovery factor
RMSE	Root mean square error
MSE	Mean square error
MAE	Mean absolute error
GPR	Gaussian process regression
SVM	Support vector machine

economic, social, and reliable with other options and reported solar energy source and electrical storage as universally suitable options; Acar et al. [26] also investigated H_2 production systems of electrolysis using electricity from fossil fuels (grid electrolysis), wind, and photovoltaic (PV), thermochemical water splitting cycles using nuclear and solar, and photoelectrochemical cell considering various technical, environmental, and economic parameters by hesitant analytic hierarchy process, and suggested grid electrolysis as most sustainable H_2 production method in the near future; Li and Cheng [27] compared technical, environmental, and economic feasibility of coke oven gas-based hydrogen (COGH) and coal gasification-based hydrogen (CGH) and reported enhanced performance of COGH in terms of energy consumption of 34.6%, carbon emissions of 36.7%, capital cost of 27.4%, and operating cost of 8.7% than ones for CGH; Kannah et al. [28] reviewed various H_2 production methods of SMR, pyrolysis, gasification, dark fermentation, photobiolysis, water electrolysis (WE), and renewable liquid reforming and suggested that SMR is most efficient one with its high efficiency of 70–85%,very cheap reactant cost of 0.3 \$ kgH_2^{-1}, and followed H_2 production cost of 1.25–3.50 \$ kgH_2^{-1}; Shah [29] conducted feasibility study for renewable energy sources of wind, solar, biomass, municipal solid waste, geothermal, and micro-hydro and suggested wind as most favorable source in H_2 production and micro-hydro and solar as possible alternatives; El-Emam and Özcan [30] reviewed economic and environmental aspects of various H_2 production methods and systems such as WE, thermochemical water splitting, renewable energy-based H_2 production, and grid-connected one and reported very high close connection between intermittent characteristics of solar and wind and unit H_2 production cost and economic benefit of mature technologies connected by electrolysis; Salkuyeh et al. [31] investigated technical, environmental, and economic feasibility of biomass-based gasification with fluidized bed and entrained flow and reported that biomass cost of 100 \$ ton^{-1}, minimum cost of 115 \$ $tonCO_2^{-1}$, and minimum natural gas cost of 5 \$ GJ^{-1} are necessary to accomplish same minimum H_2 selling price with ones from SMR; El-Chen et al. [32] conducted life cycle assessment, cradle-to-cradle environmental analysis technique, for H_2 production system with microbial electrolysis cell and reported environmentally favorable effect of cathodic gas recovery and H_2 production rate on reduction of carbon emissions in operation and construction phases, especially, 18.8 $kgCO_2$-kgH_2^{-1} with electricity transformation efficiency of 90%, cathodic gas recovery of 90%, and applied voltage of 0.5 V; Temiz and Javani [33] designed and simulated floating PV system where stored H_2 is used to produce electricity through fuel cell and reported that 99.4% of the electricity demand for grid connection (60.30 MWh y^{-1} electric consumption) or fossil fuel (211.94 MWh y^{-1}) cases can be satisfied with its own system; Sadeghi et al. [34] investigated comprehensive economic and environmental feasibilities of four pathways - SMR, coal gasification, PV, and solar thermal electrolysis – for providing H_2 to oil and gas industries with and reported greenhouse gas abatement costs of 0.77 \$ $kgCO_2^{-1}$ and 1.37 \$ $kgCO_2^{-1}$ for PV and solar thermal electrolysis; Heo et al. [35] conducted techno-economic analysis to evaluate the technical, environmental, and economic feasibility of ethane steam reforming (ESR) in a MR and reported that unit H_2 production cost of 2.93 \$ kgH_2^{-1} and CO_2 emission rate of 298.98 $tonCO_2$ y^{-1} for ESR in an MR, 20.4% and 13.3% lower compared to conventional reactor, respectively.

With increasing global trends for a transition to 4th industrial revolution first stated at the 2016 World Economic Forum [36,37], researches for related technologies such as artificial intelligence (AI), internet of things (IoT), big data, and cyber-physical system, etc. have been actively conducted based on its versatility for many fields [38–43]. Especially, due to its high-accurate predictability based on reference data, machine learning (ML) has been widely applied in many fields such as recognition of words, content identification, image processing, spam filtering, fluid dynamics, catalysis, and molecular dynamics adopting its own multi-functionality [44–48]: Gao et al. [49] developed neural-network model trained by 10 million examples to predict chemical conditions of catalyst, solvent, reagent, and temperature showing predictive accuracy of 69.6% for catalyst, solvent, and reagent, comprehensively, 80–90% for individual parameter, and 60–70% for temperature; Yan et al. [50] constructed supervised artificial neural network (ANN) models trained by experimental data for 19 manganese ores showing coefficient of determination of 0.94 and mean absolute error of 0.057 to predict reactivity of the oxygen carriers and oxygen transfer capacity using operating temperature, composition, and mechanical properties as inputs; Coley et al. [51] used 15,000 reference reaction records to train reaction outcome predictable model and its recorded products were ranked as 1, 1–3, and 1–5 for 71.8%, 86.7%, and 90.8% of cases; Singh et al. [52] tried to overcome conventional method to predict and calculate catalytic performance using density functional theory with ML-based prediction model and compare several predictive ways of linear/non-linear regression, random forest, and gaussian process, etc. In addition, features such as coordination # of metal atom, characteristic of the adsorbate, # of bonds broken, binding energy of the adsorbate, and polynomial combinations were used, then surface bond energy turned out to be key factor to determine transition state energies; Hough et al. [53] developed detailed lignin pyrolysis kinetic model using ANN and decision trees showing 99.9% predictive accuracy on new data and possibility to reduce computational expense for prediction of kinetic; Chew and Cocco [54] adopted ML methods of ANN and random forest to determine influence of each parameter in circulating fluidized bed process and obtain predictive capability without first principle related with its process; Zhu et al. [55] investigated effects of temperature and catalysts on n-pentane cracking process producing ethylene and propene using ANN-based ML model with 76% of n-pentane conversion, 88% ethylene yield, and 84% of propene yield deviated less than 2%, 1 wt%, and 1 wt%, respectively; Golkarnarenji et al. [56] constructed support vector regression (SVR) and ANN based model to predict mechanical properties of oxidized polyacrylonitrile fibers during stabilization reaction with average prediction errors of 2.4% (SVR) and 2.7% (ANN) and 3.7% (SVR) and 4.1% (ANN) for Young's modulus and tensile strength.

Even though many researches for technical, environmental, and economic analysis for H_2 production methods and applications of ML-based approaches in chemical engineering have been performed, there are very few papers adopting ML-based prediction model reflecting simulated performance to investigate a technical, environmental, and economic feasibility. So, in this study, comprehensive technical, environmental, and economic feasibility analysis for MSR in serial reactors and membrane filters is conducted in the following steps: (1) Process simulation using Aspen Plus® is carried out to obtain trend of technical and environmental performance according to 6 different technical parameters; (2) For wide range investigation for effects of techno-economic parameters, reaction kinetics used in process simulation are put into MATLAB® and economic analysis with 6 economic parameters is performed. In other words, numerical model, where 12 techno-economic parameters are used in evaluating unit H_2 production rate, CO_2 emission, and unit H_2 production cost is constructed; (3) ML-based technical, environmental, and economic predictive model (Fig. 1) is developed by training of 10,000 reference data sets pairing 12 techno-economic parameters as inputs and outputs of H_2 production rate, CO_2 emission, and unit H_2 production cost generated from a numerical model leading to much time- and cost-efficient prediction (Fig. 2).

2. Methods

2.1. Process simulation and numerical model

To precisely simulate thermodynamic units used in MSR with membrane filters in a numerical model, Aspen Plus®-based process simulation was performed (Fig. 3).

In process simulation model, reactant consisting of liquid methanol and water with a specific ratio (technical parameter 6) was fed into Gibbs reactors (technical parameter 1) at a specified temperature (technical parameter 2) and constant pressure to produce hydrogen. Product streams passed through membrane filters expressed by separators and some amount of H_2 in a product stream was permeated through H_2 permeable membrane in which amount of permeated H_2 can be determined by H_2 permeance (technical parameter 3), differences of H_2 partial pressures, and membrane area (technical parameter 4). Remained stream (retentate side) entered additional reactors to produce additional H_2 or compressor and pressure swing adsorption (PSA) unit, which is one of commercial component separation units [57], to purify H_2. Permeated H_2 was merged with sweep gas (technical parameter 5) flowing through permeate side and entered flash drum to knock-out sweep gas to get pure H_2. In addition, a boiler unit where natural gas is used as fuel for combustion reaction with 20% excess oxygen was constructed to supply the required heat, related to environmental performance, for each reactor where endothermic reactions occur.

However, Aspen Plus®-based process simulation model has its limits to handle comprehensively various technical parameters and connect economic parameters for economic analysis based one technical performance. Therefore, numerical model which simulates technical and environmental performance based on certain units constructed in Aspen Plus®-based process simulation model and analyses economic performance was constructed by MATLAB® and validations for units of reactor (Fig. 4a) and compressor (Fig. 4b), which are highly influential on technical, environmental, and economic feasibility.

2.2. Economic analysis

Unit H_2 production costs ($ kgH_2^{-1}) under various technical and economic parameters were estimated based on itemized cost estimation

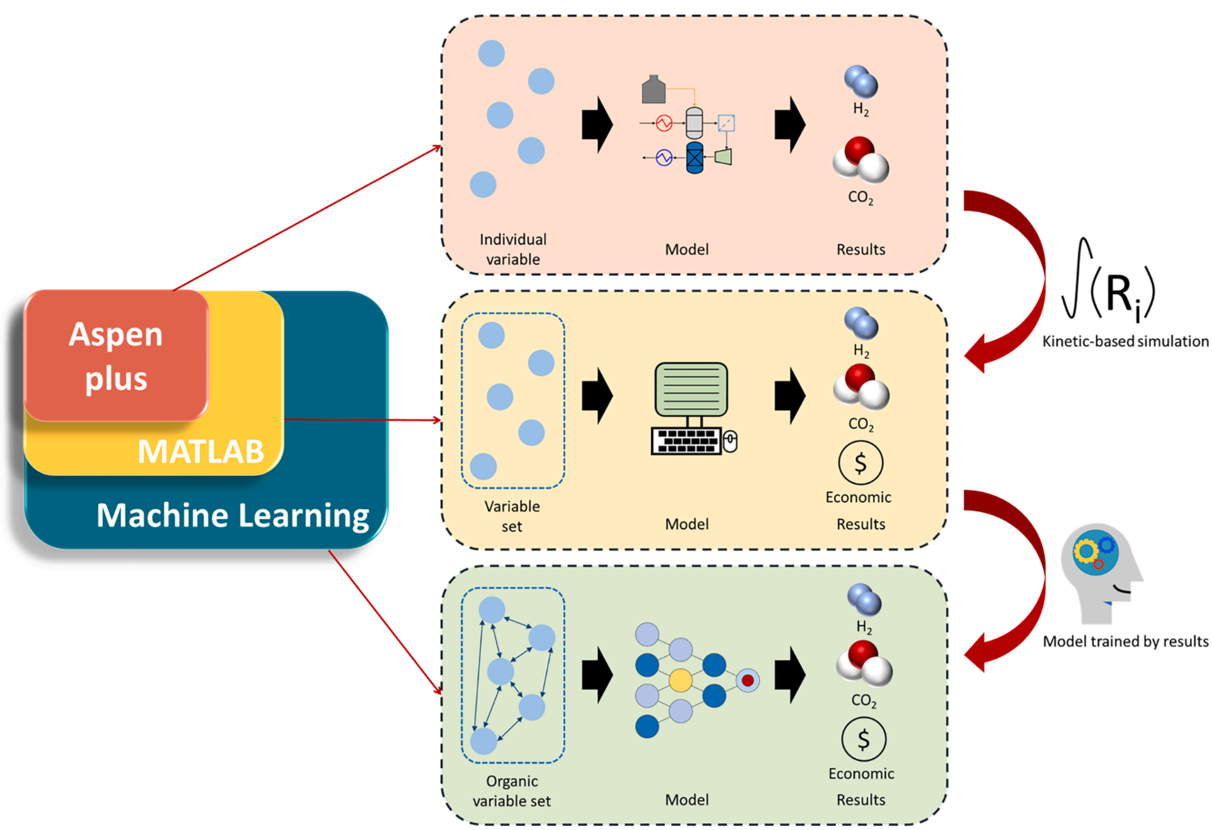

Fig 1. Schematic diagram of development procedure from Aspen Plus® based process simulation followed by numerical model to machine learning (ML) based predictive model to predict technical, environmental, and economic feasibility of the H_2 production system of methanol steam reforming (MSR) in serial reactors and membrane filters.

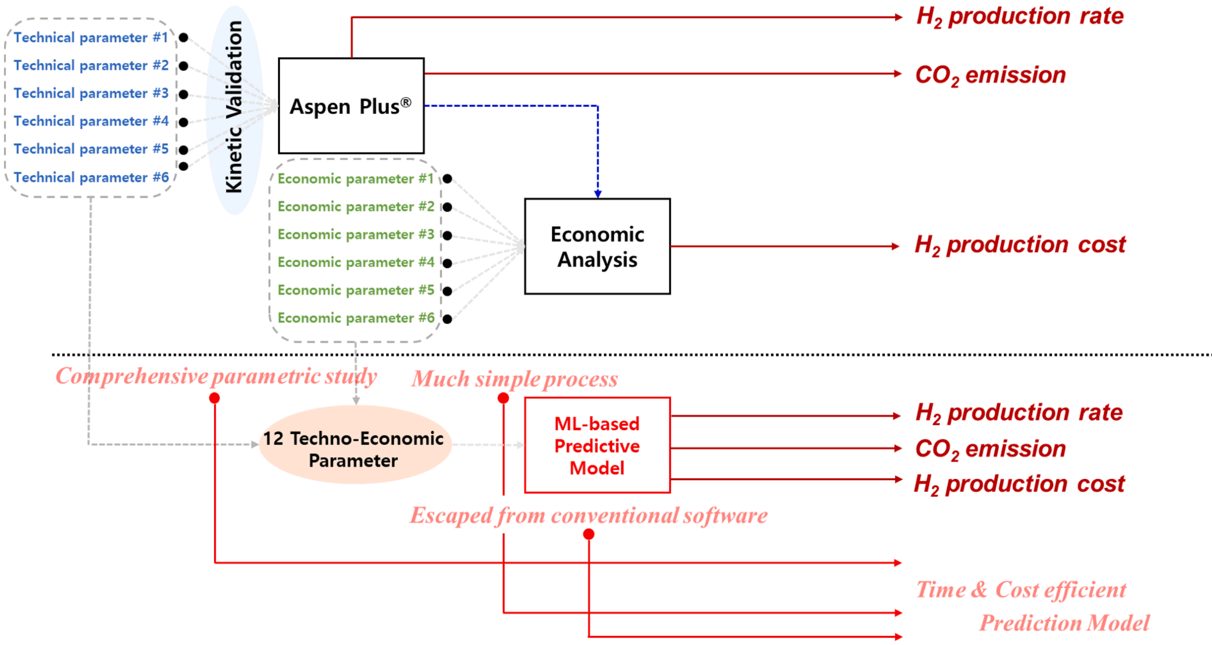

Fig 2. Schematic diagram for comparison of feasibility study between using (a) conventional chemical process simulator and (b) machine learning (ML) based predictive model.

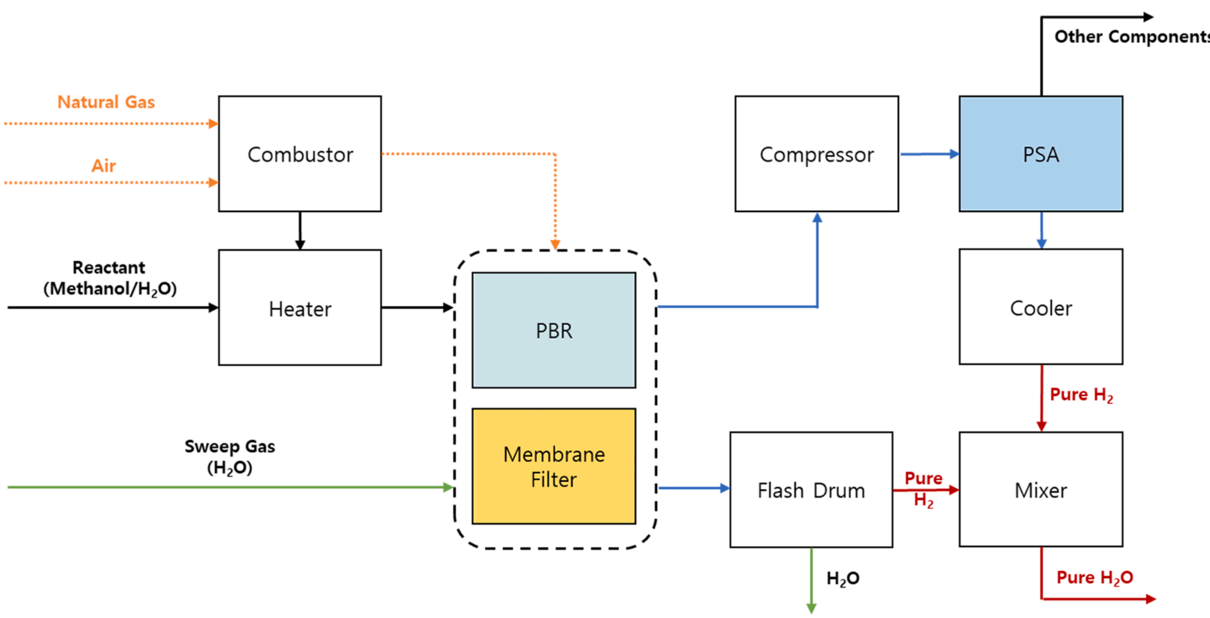

Fig 3. Block flow diagram (BFD) for methanol steam reforming (MSR) in serial reactors and membrane filters.

method, which defines total cost as the sum of capital cost ($) and operating cost ($ y^{-1}), suggested by Turton et al. [58]. Reactor (economic parameter 1), compressor (economic parameter 2), membrane module, PSA, and supplement for capital cost and labor (economic parameter 3), reactant (economic parameter 4), natural gas (economic parameter 5), electricity (economic parameter 6), sweep gas, membrane replacement, PSA operating cost, maintenance, and other costs for operating cost were considered as shown in Table 1. To consider variations of each capital cost over time, chemical engineering plant cost index (CEPCI) was applied and each cost was re-estimated based on Equation (1).

$$C_2 = C_1\left(\frac{I_2}{I_1}\right) \quad (1)$$

where C is an equipment cost and I is a CEPCI.

Also, annualized cost ($ y^{-1}) of each capital cost was obtained by applying capital recovery factor (CRF) as shown in Equation (2).

$$CRF = \frac{i(1+i)^N}{(1+i)^N - 1} \quad (2)$$

where i is a discount rate and N is an economic analysis period.

2.3. Machine-learning model

Based on training data sets consisting of 12 techno-economic parameters previously defined as inputs and technical and environmental performance (H_2 production rates and CO_2 emissions) from a numerical

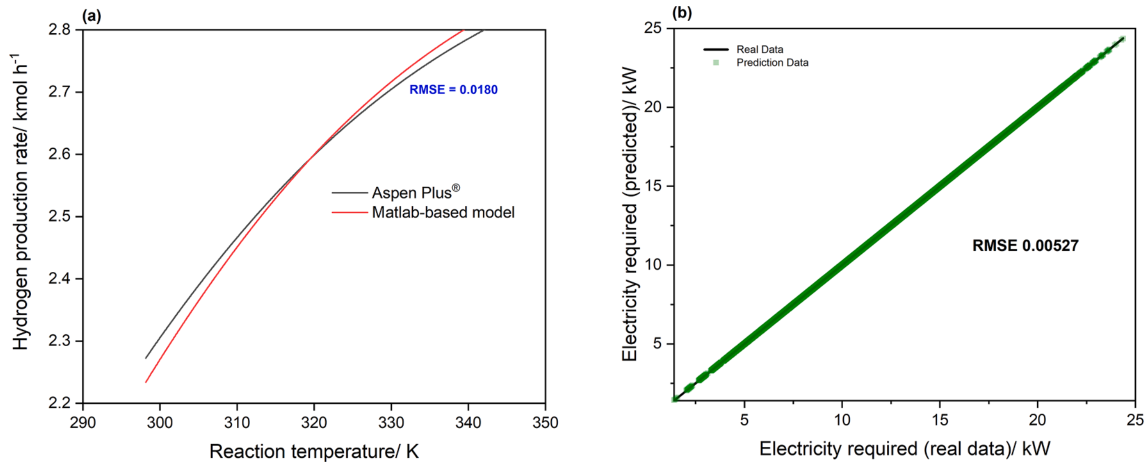

Fig 4. Model validation of numerical model results with process simulation for (a) reactor and (b) compressor.

Table 1
List of economic specifications and assumptions used in economic analysis.

Economic specifications			
Reactor [59] (Economic parameter 1)	28,746 $	Membrane module [65]	5,382 $ m^{-2}
Compressor [60] (Economic parameter 2)	15,000 $	Membrane replacement [66]	20% of membrane module ($ y^{-1})
Labor [61] (Economic parameter 3)	34,167 $ y^{-1} (full-time) 16,667 $ y^{-1} (temporary)	PSA [67]	$\frac{CEPCI}{392.6}\times 1,510,000 \times (\frac{inletflowrate}{500})^{0.6}$ ($)
Reactant [62] (Economic parameter 4)	305 $ ton^{-1}	PSA operating cost [67]	$6.11 \times 100 \times inletflowrateexceptH_2$
Natural gas [63] (Economic parameter 5)	0.0111 $ MJ^{-1}	Supplement	20% of (reactor + compressor +membrane module + PSA) ($)
Electricity [64] (Economic parameter 6)	0.06 $ kWh^{-1}	Maintenance [68]	2% of (reactor + compressor +membrane module + PSA) ($ y^{-1})
Sweep gas (water in this study) [58]	0.067 $ ton^{-1}	Other cost [68]	1% of (reactor + compressor +membrane module + PSA) ($ y^{-1})
Economic assumptions			
CEPCI	603.1	N	10 for compressor, PSA, and membrane module20 for reactor and supplement
i	0.045	Stream factor	0.93

model for a H_2 production system of MSR in serial reactors and membrane filters and unit H_2 production costs obtained by economic analysis, ML-based technical, environmental, and economic feasibility predictive model was constructed using MATLAB® toolbox. In this study, mass 10,000 data sets as training data obtained from numerical model and 5-fold cross validation for its training data were used. In addition, only key 12 techno-economic parameters in training data sets were considered. Furthermore, for much accurate predictive ability of our model, results of various ML-based regression methods were compared in terms of root mean square error (RMSE), mean squared error (MSE), and mean absolute error (MAE) (Table 2) and SVR (Fig. 5a), decision tree regression (Fig. 5b), and gaussian process regression (GPR) (Fig. 5c) were selected for fundamental approaches to a predictive model showing high accuracies and system fitting effect with very less RMSEs of 0.0184, 0.00132, and 0.0443, respectively. Especially, for the prediction of unit H_2 production cost, regression algorithm using GPR shows higher performance with lower RMSE than one from SVR, where statistical approaches were applied to identify certain relations between inputs and outputs, due to its probabilistic approach to figure out the rule to predict output. Among various types of regression model, Cubic support vector machine (SVM) where cubic polynomial kernel function was used with maximum error or specified margin of 0.0124 and C of 0.124, medium tree with minimum number of parent and leaf nodes of 24 and 12, and Matern 5/2 GPR where Matern 5/2 kernel function with σ of 0.0528 and constant basis function were applied were used as regression algorithms to investigate technical, environmental, and economic feasibility, respectively.

2.3.1. Support vector regression (SVR)

The concept of SVM, proposed by Cortes and Vapnik [69] as current form, has been widely applied to both linear and non-linear sophisticated binary data classification problems showing its high empirical performance [70,71]. SVM establishes a decision boundary by mapping input training data, expressed by vectors, to higher dimensional feature space, then linear regression between input vectors and output to predict future output occurs. SVR adopts extended principles of SVM as follows [72,73].

Cortes and Vapnik [69] introduced hyper linear plane that classifies binary data using support vector and the hyperplane can be expressed as:

$$f(x) = wx + b \tag{3}$$

where $f(x)$ is predicted value, w is weight vector, and b is an offset of the regression line or bias.

To classify binary data, ideal hyperplane with optimal w and b is constructed resulting a maximum margin ($\frac{2}{||w||}$), shortest distance between hyperplane and classified data sets (class), subject to:

$$\min w^T w$$

$$s.t. y_i\left(w^T \bullet \mathrm{x_i} + \mathrm{b}\right) \geq 1, \mathrm{i} = 1, \cdots, \mathrm{l} \tag{4}$$

In addition, this model can be modified with empirical error (ξ) to prevent overlap of classified data as follows:

$$\min w^T w + C\sum_i^l \xi_i$$

$$s.t. y_i\left(w^T \bullet \mathrm{x_i} + \mathrm{b}\right) \geq 1, \mathrm{i} = 1 - \xi_\mathrm{i}, \forall \mathrm{i}$$

$$\xi_i \geq 0, \forall i \tag{5}$$

Table 2
Comparison results for different regression approaches in terms of root-mean square error (RMSE), and mean square error (MSE), and mean absolute error (MAE).

	H_2 production rate			CO_2 emission rate			Unit H_2 production cost		
	RMSE	MSE	MAE	RMSE	MSE	MAE	RMSE	MSE	MAE
	Linear								
Linear	0.075	0.006	0.056	0.011	0.000	0.009	0.075	0.006	0.056
Interactions Linear	0.047	0.002	0.038	0.011	0.000	0.009	0.016	0.002	0.037
Robust Linear	0.077	0.006	0.055	0.011	0.000	0.009	0.077	0.006	0.055
Stepwise Linear	0.047	0.002	0.037	0.011	0.000	0.009	0.047	0.002	0.037
	Tree								
Fine Tree	0.038	0.001	0.027	0.002	0.000	0.002	0.041	0.002	0.028
Medium Tree	0.038	0.001	0.028	0.002	0.000	0.001	0.040	0.002	0.029
Coarse Tree	0.046	0.002	0.034	0.002	0.000	0.002	0.045	0.002	0.035
	Support Vector Machine (SVM)								
Linear SVM	0.077	0.006	0.055	0.013	0.000	0.009	0.078	0.006	0.055
Quadratic SVM	0.051	0.003	0.037	0.008	0.000	0.006	0.053	0.003	0.038
Cubic SVM	0.036	0.001	0.026	0.008	0.000	0.005	0.045	0.002	0.033
Fine Gaussian SVM	0.054	0.003	0.039	0.009	0.000	0.006	0.100	0.010	0.080
Medium Gaussian SVM	0.045	0.002	0.032	0.008	0.000	0.006	0.048	0.002	0.035
Coarse Gaussian SVM	0.076	0.006	0.056	0.011	0.000	0.009	0.081	0.007	0.061
	Ensemble								
Boosted Trees	0.134	0.018	0.123	0.007	0.000	0.006	0.134	0.018	0.123
Bagged Trees	0.054	0.003	0.034	0.004	0.000	0.003	0.045	0.002	0.031
	Gaussian Process Regression (GPR)								
Squared Exponential GPR	0.039	0.002	0.029	0.008	0.000	0.006	0.045	0.002	0.033
Matern 5/2 GPR	0.038	0.001	0.028	0.007	0.000	0.005	0.045	0.002	0.033
Exponential GPR	0.040	0.002	0.028	0.007	0.000	0.005	0.046	0.002	0.033
Rational Quadratic GPR	0.039	0.001	0.028	0.008	0.000	0.006	0.045	0.002	0.033

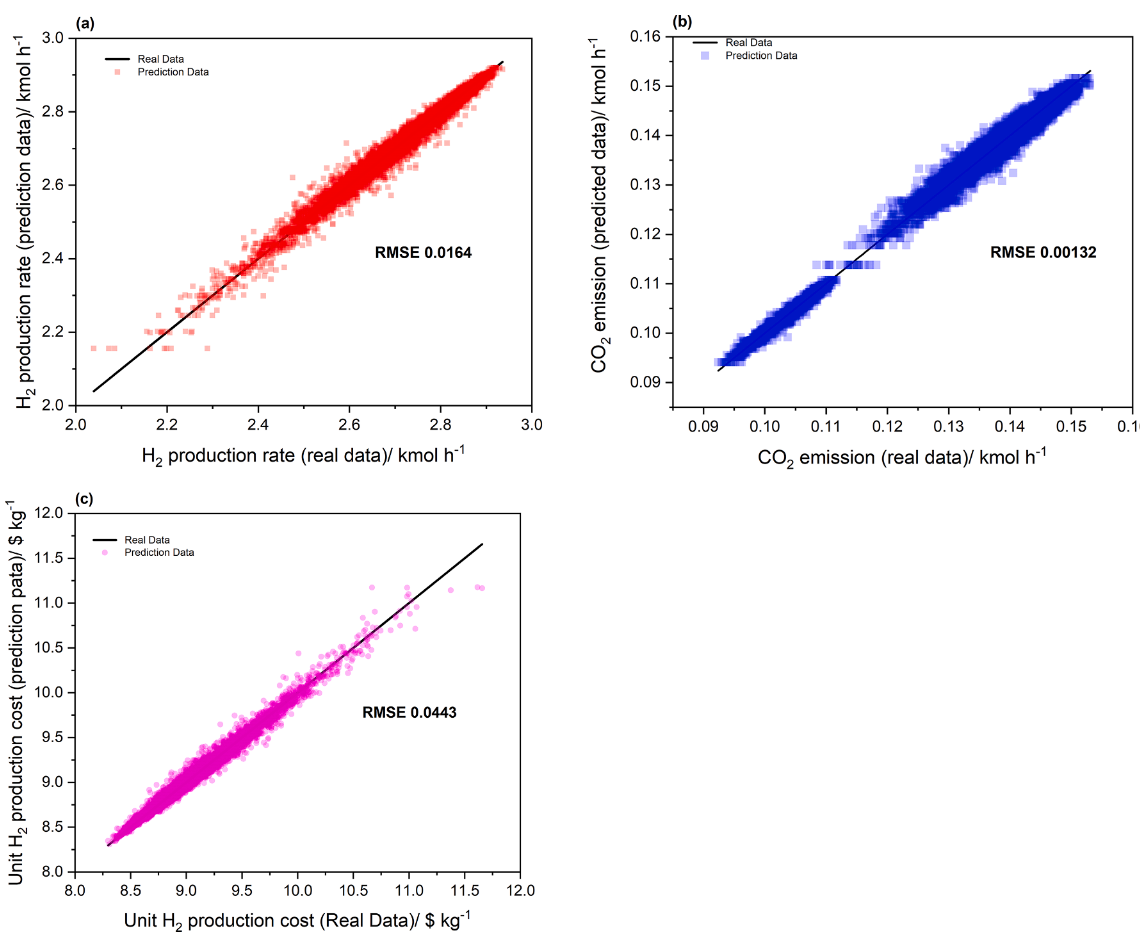

Fig 5. Model validation of prediction results for (a) H_2 production rate, (b) CO_2 emission, and (c) unit H_2 production cost with results from numerical model.

$$\max e^T\alpha - \frac{1}{2}\alpha^T Q\alpha \qquad s.t.\ y^T\alpha = 0, \quad 0 \le \alpha_i \le C_i \forall i \tag{6}$$

where C is a relative weight between margin and ξ, $e = (1, \cdots, 1)^T \in R^i$, and $Q_{i,j} = y_i y_j (x^i)^T x_j$.

Minimizing of $w^T w$ ($||w||^2$), defined as reciprocal of margin, expressed in first term in Equation (4) indicates that optimal linear hyperplane with maximized margin can be obtained. Also, by considering additional constant (C), this model can produce optimal hyperplane with maximized margin and minimized empirical error. w and linear hyperplane of $f(x)$ can be obtained as follows:

$$w = \sum_i \alpha_i^* y_i x^i$$

$$f(x) = \sum_i \alpha_i^* y_i (x^i)^T x + b \tag{7}$$

where α^* is an optimal solution and $x^i(\alpha_i^* > 0)$ is a support vector.

Meanwhile, to cover non-linear data used in various filed, training data set (x_i) can be mapped to new data points (Φ_i) in high-dimensional feature space using kernel function, then new hyperplane ($f_\Phi(x)$), which has non-linearity in original low-dimensional space, is constructed and re-converted to space.

Based on principles of SVM, ε-insensitive loss function ($L_\varepsilon(x, y, f)$) is used instead of terms related with empirical error ($C\sum_i^l \xi_i$) to predict future real value (Equation (5)).

$$L_\varepsilon(x, y, f) = \begin{cases} \left|y - (w^T x + b)\right| - \varepsilon \left|y - (w^T x + b)\right| \geq \varepsilon \\ 0 \left|y - (w^T x + b)\right| < \varepsilon \end{cases} \tag{8}$$

With ε-insensitive loss function, a model developed in SVM is modified maximizing margin and minimizing difference between y_i and $w^T x + b$ as follows:

$$\min w^T w + C \sum_i^l (\xi^+ + \xi^-)$$

$$s.t. \left(w^T x^i + b\right) - y_i \leq \varepsilon + \xi^+, \forall i$$

$$y_i - \left(w^T x^i + b\right) \leq \varepsilon + \xi^+, \forall i$$

$$\xi^+, \xi^- \geq 0, \forall i \tag{9}$$

$$max \left(\sum_i y_i (\alpha_i^- - \alpha_i^+) \right) - \varepsilon \sum_i (\alpha_i^- + \alpha_i^+) - \frac{1}{2} \sum_{i,j} (\alpha_i^- - \alpha_i^+) \left(\alpha_j^- - \alpha_j^+\right) Q_{i,j}$$

$$s.t. \sum_i (\alpha_i^- - \alpha_i^+) = 0$$

$$0 \leq \alpha_i^+, \alpha_i^- \leq C_i, \forall i \tag{10}$$

2.3.2. Decision tree regression

Due to (a) its applicability to both categorical and continuous variable for classification and regression, respectively, (b) insensitivities for several conditions such as outliers, multi-colinearity, heteroskedasticity, (c) good readability, (d) fast training time, and (e) higher training performance for large data set, decision tree-based classification (categorial variable)/ regression (continuous variable) is widely used in various sectors such as analyses for medical signal, image, and network security, etc [74–77]. In decision tree classification/regression, data processing based on if-then rules to classify/regress each data is performed with following procedures: deciding the criteria, selecting splits, and determining when splitting stopped and optimal tree [77].

Predicted values can be evaluated by following three criteria of re-substitution error, cross-validation error, and test sample error as shown in Equation 11–13, respectively.

$$Err(p) = \frac{1}{N} \sum_{i=1}^N (u_i - p(v_i))^2 \tag{11}$$

$$Err(p) = \frac{1}{N_k} \Sigma_k \Sigma_{(u_i, v_i)} \left(v_i - p^{(k)}(u_i)\right)^2 \tag{12}$$

$$Err(p) = \frac{1}{N_2} \Sigma_{(u_i, v_i) \in X_k} \left(v_i - p^{(k)}(u_i)\right)^2 \tag{13}$$

where N = a size of sample X, (u_i, v_i) = learning samples, i = 1, 2, …, N, k = the number of sub samples ($X_1, X_2, \cdots, X_k$) (Equation (12)), and X_2 = subsample of sizes N_2, which is not used in input-predictor correlation.

Based on certain criteria determined in the first step, splits which can be measured by node impurity indicating homogeneity are selected, and then stopped when an optimum tree with right size is obtained by tree pruning.

2.3.3. Gaussian process regression (GPR)

For an input training data set of functional covariates ($\boldsymbol{x}(t) = (x_1(t), \cdots, x_Q(t))^T$) and output training data set of functional response ($y(\boldsymbol{x})$), nonlinear regression model, defined as mapping $f : R^Q \rightarrow R$ can be expressed as follows [78,79]:

$$y(\boldsymbol{x}) = f(\boldsymbol{x}) + b \tag{14}$$

where b is an offset of the regression model or bias.

There have been various approaches to estimate function f such as local polynomial model ($f(\boldsymbol{x}) = \sum_{k=1}^K \alpha_k(x) x^k$) or spline smoothing ($f(\boldsymbol{x}) = \sum_{k=1}^K \alpha_k \Phi_k(x)$), however, most of them have its limitations applied in multidimensional covariates. To overcome this limitation, the Bayesian-based approach, where f is defined as a random function, was introduced to define the Gaussian process with mean function (μ) and covariance function (k) (Equation (15)).

$$Cov(f(\boldsymbol{x}), f(\boldsymbol{x}^{'})) = k(\boldsymbol{x}, \boldsymbol{x}^{'}; \boldsymbol{\theta}) \tag{15}$$

where θ is hyper-parameters to be trained by using training data.

With parameters defined above, the gaussian process regression (GPR) model is expressed as follows:

$$f(x) = GPR[\mu, \mathrm{k}(\theta) | \boldsymbol{x}] \tag{16}$$

Equation (16) implied that converted values of $p(f(x_1), \cdots, f(x_n))$ from $\{t_1, \cdots, t_n\}$ follows the distribution of $N(0, K)$, where K is defined by evaluating covariance function as $[K]_{i,j} = k(t_i, t_j, \theta)$ with mean and variance expressed in Equation (17) and Equation (18).

$$\mu(\mathrm{t}_*) = k_{*f} \left(K + \sigma_{noise}^2 I\right)^{-1} y \tag{17}$$

$$\sigma^2(\mathrm{t}_*) = k(t_*, t_*) - k_{*f} \left(K + \sigma_{noise}^2 I\right)^{-1} k_{*f}^T \tag{18}$$

where t_* is test input, the element i of the row vector k_{*f} is converted covariance between $f(t_*)$ and $f(t_i)$ as $\left[k_{*f}\right] = k(t_*, t_i, \theta)$.

3. Results and discussion

3.1. Techno, environmental, and economic feasibility

To investigate technical, environmental, and economic feasibility of a H_2 production system of MSR in serial reactors and membrane filters, 12,000 data sets for H_2 production rates, CO_2 emissions, and unit H_2 production costs were obtained from ML-based predictive model (Fig. 6). For comprehensive feasibility studies, 6 technical parameters for predictions of H_2 production rates and CO_2 emissions and all 12 techno-economic parameters for unit H_2 production costs considered in this study as inputs for ML-based predictive model were randomly varied within a boundary differently defined for each parameter as shown in Table 3.

Fig. 6a shows the 12,000 data sets of H_2 production rate with maximum of 2.9953 kmol h^{-1}, minimum of 2.2598 kmol h^{-1}, mean of

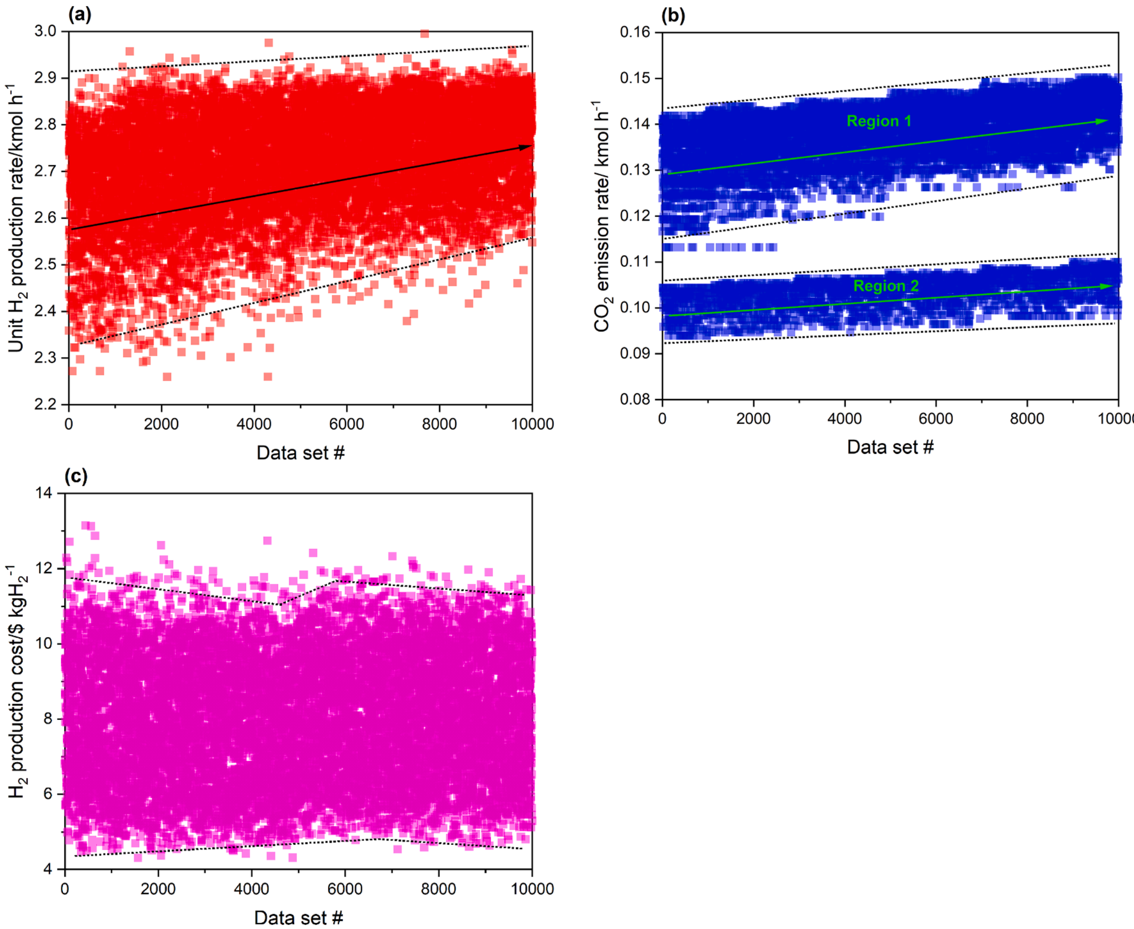

Fig 6. Results of comprehensive (a) technical, (b) environmental, and (c) economic feasibility studies from machine learning (ML) based predictive model according to varied 12 techno-economic parameters in total range.

Table 3
List of variation ranges of 12 techno-economic parameters used in individual technical, economic, and environmental feasibility studies.

Technical parameter	Range	Unit	Economic parameter	Range	Unit
# of Gibbs reactors	2–8		Reactor	14,373–43,119	\$
Operating temperature	25–75	°C	Compressor	7,500–22,500	\$
H_2 permeance	5×10^{-6}-5×10^{-5}	mol m^{-2} s^{-1} Pa^{-1}	Labor	17,083–51,250	\$ y^{-1}
				8,333–25,000	
Membrane area	0.01 – 0.035	m^2	Reactant	152.5 – 457.5	\$ ton^{-1}
Sweep gas flow rate	1 – 10	kmol h^{-1}	Natural gas	0.0056 – 0.0167	\$ MJ^{-1}
S/C ratio	1 – 1.5		Electricity	0.03 – 0.09	\$ kWh^{-1}

2.7250 kmol h^{-1}, median of 2.7395 kmol h^{-1}, mode of 2.6776 kmol h^{-1}, and standard deviation of 0.1131. As the # of reactors increased (expressed as the number of data sets increased), the range of H_2 production rates goes narrower meaning that higher probability for higher H_2 production rates can be observed with increased # of reactors and membrane filters and the centers of data set goes higher indicating an overall increasing tendency of H_2 production rates.

Fig. 6b shows a result of environmental feasibility study divided by region 1 with 9,277 data and region 2 with 2,707 data caused by latent heat of each component with maximum of 0.1506 and 0.1107 kmol h^{-1}, minimum of 0.1167 and 0.0939 kmol h^{-1}, mean of 0.1388 and 0.1036 kmol h^{-1}, median of 0.1395 and 0.1041 kmol h^{-1}, mode of 0.1429 and 0.1088 kmol h^{-1}, and standard deviation of 0.0063 and 0.0036, respectively. Similar to a study for technical performance in terms of H_2 production rate, the centers of data set for both regions go higher, however, the range of CO_2 emissions is relatively maintained constant indicating uniform or proportional probability of CO_2 emissions as the # of reactors increased.

Fig. 6c shows unit H_2 production cost distribution with maximum of 13.1487 \$ kgH_2^{-1}, minimum of 4.3122 \$ kgH_2^{-1}, mean of 8.1045 \$ kgH_2^{-1}, median of 8.0866 \$ kgH_2^{-1}, mode of 8.9912 \$ kgH_2^{-1}, and standard deviation of 1.6856. Compared to previous ones, there is a very wide distribution of unit H_2 production cost showing a high standard deviation (1.6856) representing complex relation between techno-economic parameters and unit H_2 production cost. In addition, unit H_2 production cost increased with the # of reactors increased upon some points, but the trend was reversed at some points implying a possibility of optimized conditions of techno-economic input parameters.

From comprehensive technical, environmental, and economic feasibility studies, trends for H_2 production rates, CO_2 emissions, and unit H_2 production costs were obtained showing proportional trends along with increased # of reactors for unit H_2 production rates and CO_2 emissions and high uncertainty of unit H_2 production cost.

3.2. Effects of technical parameters on H_2 production rate

To investigate effects of technical parameters, listed in Table 3, on technical performance in terms of H_2 production rate, technical analyses for each technical parameter were conducted with 10,000 data of H_2 production rates. At each analysis, only one parameter was varied in low, medium, and high ranges (1/3 of the total range for each parameter) with other 5 parameters varied in total range. Fig. 7 and Table 4 show the results of each technical analysis.

Fig. 7a shows distributions of H_2 production rates with a set of distinct layers for each variation range of # of reactors. For each variation range, medians and maximums for H_2 production rates of 1.7314, 1.9902, and 2.3913 kmol h^{-1} and 2.3868, 2.4789, and 3.000 kmol h^{-1} were obtained showing increases of 13.0% and 27.6%, and 3.7% and 20.4% compared to ones for low variation range. This set of distinct layers was also observed for a case of operating temperature (Fig. 7b), with medians and maximums of 1.7494, 2.0936 (+16.4%), and 2.2233 kmol h^{-1} (+21.3%) and 2.7940, 2.9001 (+3.7%), and 3.000 kmol h^{-1} (+6.9%). In contrast to cases for # of reactors and operating temperature, for H_2 permeance (Fig. 7c), membrane area (Fig. 7d), sweep gas flow rate (Fig. 7e), and S/C ratio (Fig. 7f), no distinguishable trends were observed with very low variations of 0.91%, 3.60%, 2.22%, and 5.28% for medians and 0.62%, 5.38%, 5.62%, and 2.32% for minimums, respectively.

In summary, with several statistical indicators, individual analyses for effects of technical parameters on technical performance were conducted proving # of reactors and operating temperature are very influential in the H_2 production system of MSR in serial reactors and membrane filters.

3.3. Effects of technical parameters on CO_2 emission rate

Based on the same method to divide a range of investigated parameters, distributions of CO_2 emissions and statistical indicators for each parameter varied in low, medium, and high ranges were obtained as shown in Fig. 8 and Table 5. As expected from comprehensive trends of CO_2 emissions expressed in Fig. 6b, divided regions of CO_2 emission data were also observed due to the thermodynamic properties of each component used in MSR. In Table 5, trends of CO_2 emissions for each variation range (low, medium, and high) were classified by frequency of data in each CO_2 emission level (0.09 to 0.11 kmol h^{-1} (low emission level), 0.11 to 0.13 kmol h^{-1} (medium emission level), and 0.13 to 0.15 kmol h^{-1} (high emission level)).

Fig. 8a shows distributions of CO_2 emissions for # of reactors with distinct layers for each variation range with frequencies of data in high emission level of 67.65%, 70.36%, and 74.23% indicating environmental disadvantage of adding reactors. Similar to the increasing trend of the frequency of data in higher emission level caused by larger # of reactors, increased frequencies of data in high CO_2 emission level for medium and high variation range of 78.01% and 77.67% compared to one of 54.56% were observed for S/C ratio (Fig. 8f). In contrast to cases for # of reactors and S/C ratio, there were no significant differences of frequencies of data in high CO_2 emission level between each variation level with fluctuated in a range of 69.72% to 71.05%. Interestingly, the frequency of data in high CO_2 emission level significantly decreased in high variation range suggesting advantage of high operating temperature.

Based on these results, environmental disadvantages caused by higher # of reactors and S/C ratio and lower operating temperature and less influences of H_2 permeance, membrane area, and sweep gas flow rate on environmental feasibility were confirmed.

3.4. Effects of technical parameters on unit H_2 production cost

As technical parameters in chemical engineering system have considerable impacts on the economic feasibility of that [80,81], like economic parameter's ones, economic analyses in different variation ranges for each technical parameter were conducted (Fig. 9 and Table 6).

For effects of technical parameters on unit H_2 production cost, no distinct layers were observed proving its less impact compared to ones on technical performance and environmental feasibility expressed by

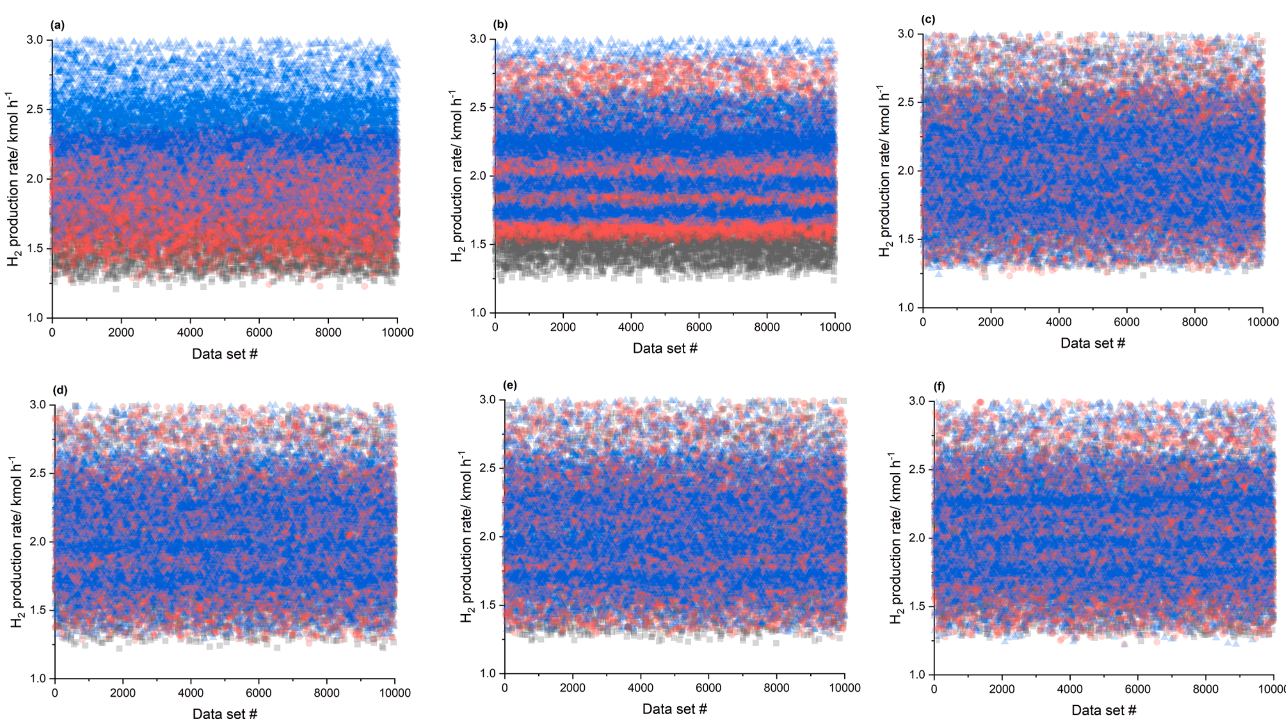

Fig 7. Results of individual technical analysis for technical parameters of (a) # of reactors, (b) operating temperature, (c) H_2 permeance, (d) membrane area, (e) sweep gas flow rate, and (f) S/C ratio varied in low, medium, and high range.

Table 4
Results of individual technical analysis with statistical indicators for technical parameters of (a) # of reactors, (b) operating temperature, (c) H_2 permeance, (d) membrane area, (e) sweep gas flow rate, and (f) S/C ratio varied in low, medium, and high range.

	(a) # of Reactors			**(b) Operating temperature**			**(c) H_2 permeance**		
	Low	Medium	High	Low	Medium	High	Low	Medium	High
Maximum	2.3868	2.4789	3.0000	2.7940	2.9001	3.0000	3.0000	3.0000	3.0000
Minimum	1.2082	1.2284	1.3821	1.2343	1.5069	1.6247	1.2238	1.2332	1.2420
Mean	1.7527	1.9830	2.3694	1.7955	2.1002	2.2021	2.0428	2.0301	2.0309
Median	1.7314	1.9902	2.3913	1.7494	2.0936	2.2233	2.0188	2.0004	2.0013
Standard deviation	0.2312	0.2520	0.3215	0.3225	0.3276	0.3316	0.3706	0.3712	0.3712
	(d) Membrane area			**(e) Sweep gas flow rate**			**(f) S/C ratio**		
	Low	Medium	High	Low	Medium	High	Low	Medium	High
Maximum	3.0000	3.0000	3.0000	3.0000	3.0000	3.0000	3.0000	3.0000	3.0000
Minimum	1.2215	1.2579	1.2910	1.2232	1.2716	1.2961	1.2485	1.2242	1.2195
Mean	1.9939	2.0428	2.0662	2.0073	2.0389	2.0577	1.9837	2.0377	2.0808
Median	1.9592	2.0121	2.0323	1.9765	2.0088	2.0213	1.9480	2.0058	2.0566
Standard deviation	0.3595	0.3696	0.3748	0.3703	0.3681	0.3730	0.3537	0.3718	0.3819

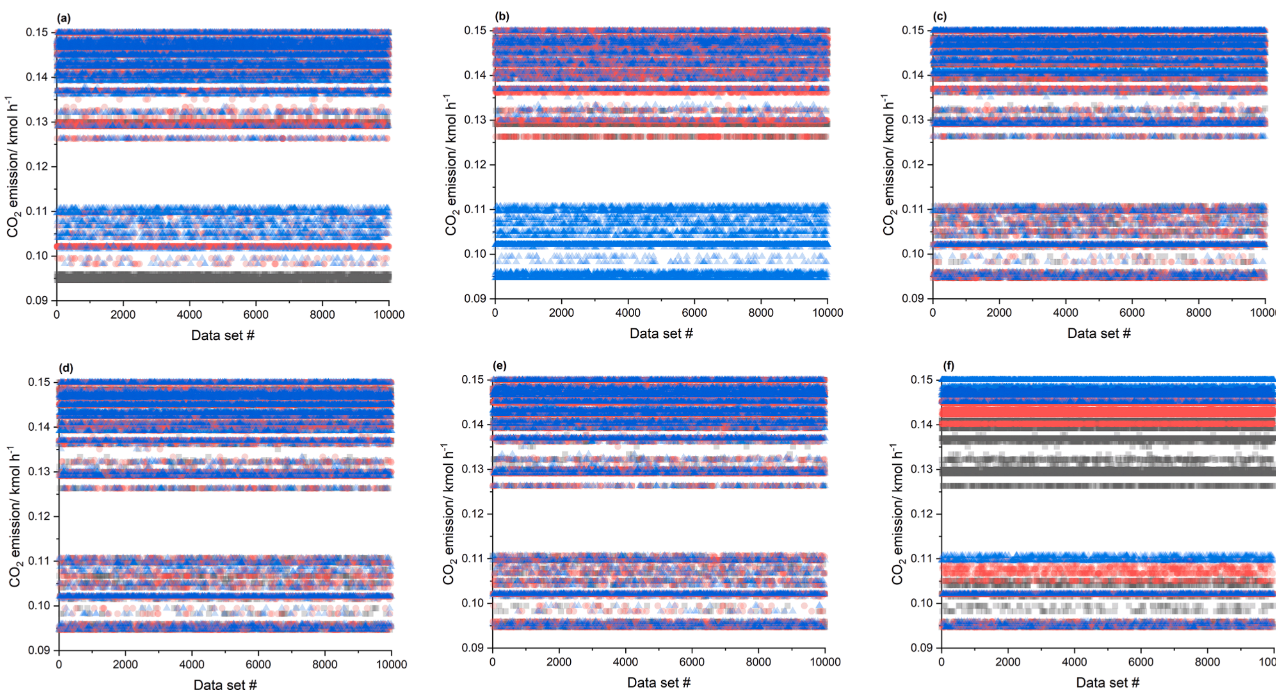

Fig 8. Results of individual environmental analysis for technical parameters of (a) # of reactors, (b) operating temperature, (c) H_2 permeance, (d) membrane area, (e) sweep gas flow rate, and (f) S/C ratio varied in low, medium, and high range.

Table 5
Results of individual environmental analysis with statistical indicators for technical parameters of (a) # of reactors, (b) operating temperature, (c) H_2 permeance, (d) membrane area, (e) sweep gas flow rate, and (f) S/C ratio varied in low, medium, and high range.

	(a) # of Reactors			**(b) Operating temperature**			**(c) H_2 permeance**		
	Low	Medium	High	Low	Medium	High	Low	Medium	High
0.09–0.11	21.76%	21.82%	18.56%	0.00%	0.00%	62.44%	20.77%	21.37%	20.21%
0.11–0.13	10.59%	7.81%	7.21%	13.05%	9.23%	4.70%	8.85%	8.91%	9.31%
0.13–0.15	67.65%	70.37%	74.23%	86.95%	90.77%	32.86%	70.38%	69.72%	70.48%
	(d) Membrane area			**(e) Sweep gas flow rate**			**(f) S/C ratio**		
	Low	Medium	High	Low	Medium	High	Low	Medium	High
0.09–0.11	21.02%	20.84%	21.10%	20.90%	20.95%	20.38%	22.05%	21.99%	18.50%
0.11–0.13	8.89%	8.98%	9.07%	8.73%	9.24%	8.57%	23.39%	0.00%	3.83%
0.13–0.15	70.09%	70.18%	69.83%	70.37%	69.81%	71.05%	54.56%	78.01%	77.67%

unit H_2 production cost and CO_2 emission rate, but there are still certain noticeable trends of data. For operating temperature (Fig. 9b) and # of reactors (Fig. 9a), decreased medians for unit H_2 production cost of 10.0009, 9.7913, and 9.3929 \$ kgH$_2^{-1}$ and 10.0009, 9.7913, and 9.3929 \$ kgH$_2^{-1}$ were obtained showing reductions of 8.7% and 12.9% and 2.1% and 6.1% between low-medium/low–high variation regions, respectively. Especially, maximum and minimum unit H_2 production costs for operating temperature continuously decreased about 8.5% and 11.4% indicating its high relevance to economic feasibility. In contrast to these trends, no significant reductions of unit H_2 production costs in terms of all statistical indicators of maximum, minimum, mean, and median were observed for H_2 permeance (Fig. 9c), membrane area

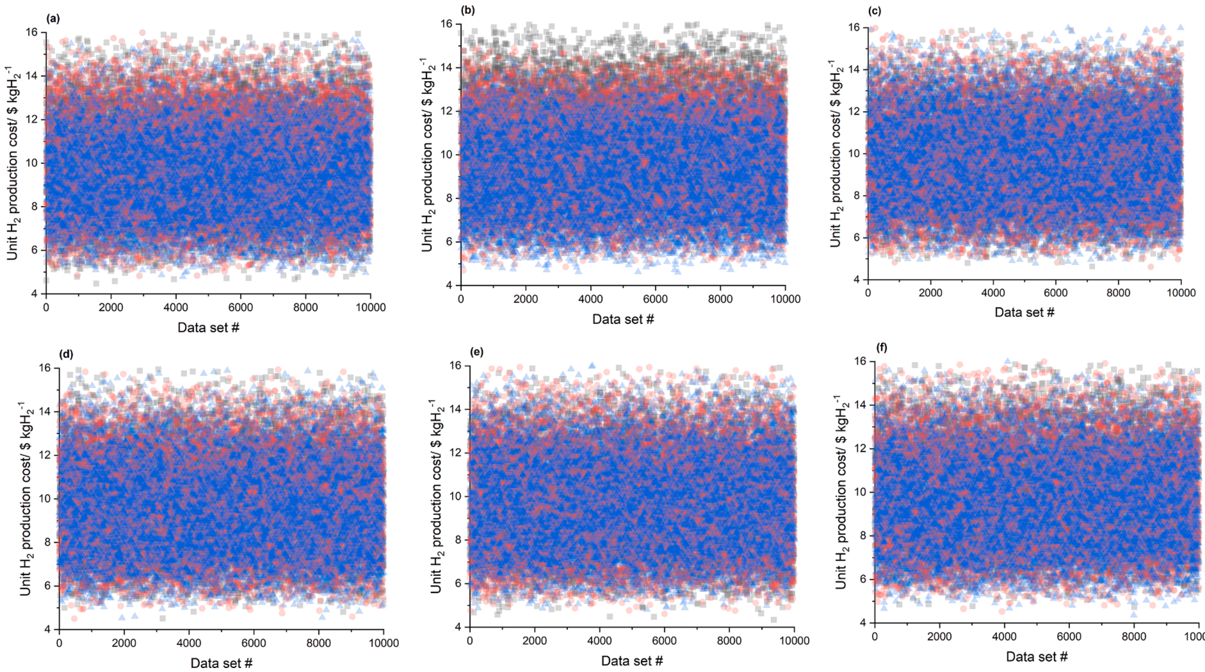

Fig 9. Results of individual economic analysis for technical parameters of (a) # of reactors, (b) operating temperature, (c) H_2 permeance, (d) membrane area, (e) sweep gas flow rate, and (f) S/C ratio varied in low, medium, and high range.

Table 6
Results of individual economic analysis with statistical indicators for technical parameters of (a) # of reactors, (b) operating temperature, (c) H_2 permeance, (d) membrane area, (e) sweep gas flow rate, and (f) S/C ratio varied in low, medium, and high range.

	(a) # of Reactors			**(b) Operating temperature**			**(c) H_2 permeance**		
	Low	Medium	High	Low	Medium	High	Low	Medium	High
Maximum	16.3000	16.3484	16.3666	16.4355	15.3936	15.0404	16.3697	16.1031	16.1195
Minimum	4.4748	4.8284	4.7872	5.2009	4.6937	4.6094	4.6375	4.6157	4.7832
Mean	10.0265	9.8321	9.4639	10.5522	9.6329	9.2251	9.7749	9.7701	9.7828
Median	10.0009	9.7913	9.3929	10.5342	9.6212	9.1725	9.7178	9.7033	9.7189
Standard deviation	2.3054	2.1704	2.0197	2.2660	2.0805	2.0094	2.1644	2.1660	2.1927
	(d) Membrane area			**(e) Sweep gas flow rate**			**(f) S/C ratio**		
	Low	Medium	High	Low	Medium	High	Low	Medium	High
Maximum	16.5291	16.1654	16.4167	16.1392	16.1966	16.2162	16.6669	16.4249	16.1298
Minimum	4.5147	4.5013	4.5348	4.3413	4.6131	4.8175	4.7321	4.5312	4.3563
Mean	9.9049	9.7545	9.6127	9.8276	9.7810	9.7492	10.0051	9.7260	9.5718
Median	9.8557	9.6937	9.5298	9.7837	9.7008	9.6941	9.9615	9.6471	9.5215
Standard deviation	2.2036	2.1703	2.1558	2.2113	2.2065	2.1633	2.2152	2.1989	2.1387

(Fig. 9d), sweep gas flow rate (Fig. 9e), and S/C ratio (Fig. 9f) showing very low variations of median of 0.16%, 3.31%, 0.92%, and 4.42%, respectively.

In short, effects of technical parameters on economic feasibility were investigated, especially, operating temperature and # of reactors were found out as influential factors for unit H_2 production cost.

3.5. Effects of economic parameters on unit H_2 production cost

To investigate effects of economic parameters on the economic feasibility of the system of MSR in serial reactors and membrane filters, distributions of unit H_2 production costs were analyzed in low/medium/high variation ranges of each parameter (Fig. 10 and Table 7).

Similar to distinct layers observed in technical analysis with technical parameters of # of reactors and operating temperature (Fig. 7a and Fig. 7b), there were very distinguishable layers for different high/medium/low variation ranges for reactant showing medians of 11.9160, 9.7259, and 7.5011 $ kgH_2^{-1}, maximums of 16.5043, 14.0570, and 11.1871 $ kgH_2^{-1}, and minimums of 8.0174, 6.4955, and 4.5586 $ kgH_2^{-1}, and standard deviations of 1.3111, 1.2078, and 1.0984. These decreasing trends of median (-37.1%), maximums (–32.2%), and minimums (-43.1%) strongly prove that reactant is most influential factor in determining economic feasibility. Labor also exhibited strong economic impacts with medians of 10.5023, 9.7752 (-6.9%), and 9.0062 $ kgH_2^{-1} (-14.2%), maximums of 16.4641, 15.6844 (-4.7%), and 14.8981 $ kgH_2^{-1} (-9.5%), and minimums 5.7619, 4.9387 (-14.3%), and 4.2473 $ kgH_2^{-1} (-26.3%), even larger standard deviation than one for reactant. In contrast to previous ones, very low variations of median of 0.66%, 0.77%, 1.38%, and 1.79% in each variation range of economic parameters were obtained for reactor (Fig. 10a), compressor (Fig. 10b), natural gas (Fig. 10e), and electricity (Fig. 10f) showing its weak impacts on unit H_2 production cost.

In short, very high impacts of economic parameters of reactant and labor on economic feasibility were confirmed showing very high reduced unit H_2 production cost distributions.

4. Conclusions

Comprehensive technical, environmental, and economic feasibility studies for promising alternative H_2 production system of methanol

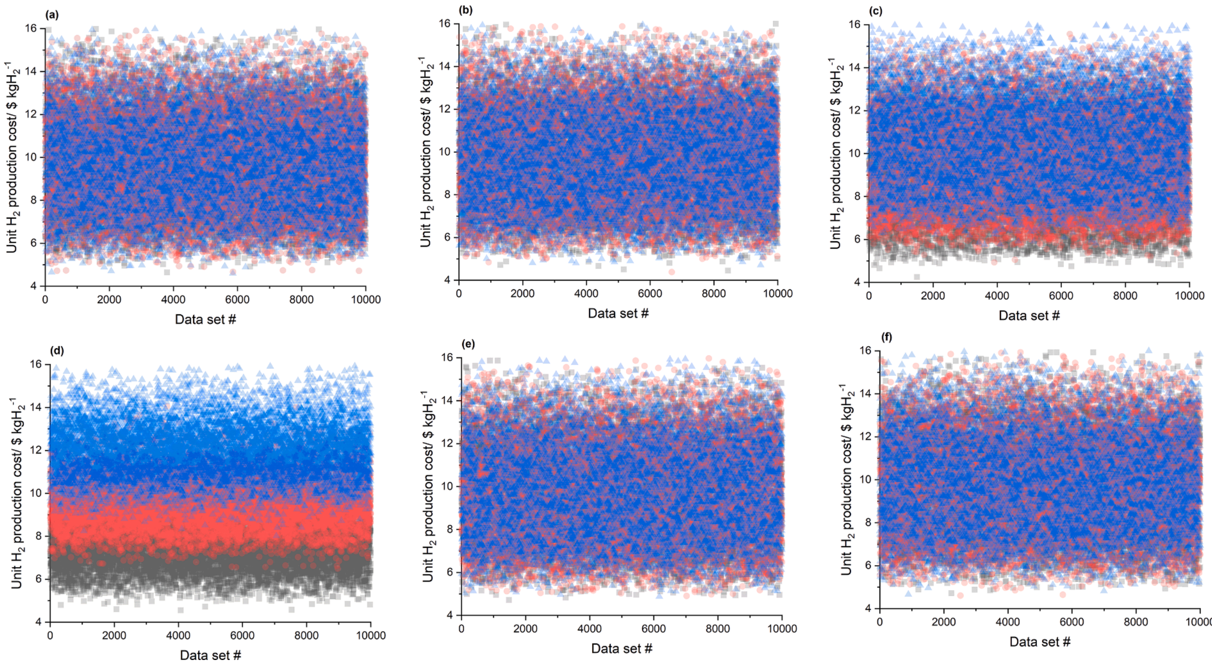

Fig 10. Results of individual economic analysis for economic parameters of (a) reactor, (b) compressor, (c) labor, (d) reactant, (e) natural gas, and (f) electricity varied in low, medium, and high range.

Table 7
Results of individual economic analysis with statistical indicators for economic parameters of (a) # of reactors, (b) operating temperature, (c) H_2 permeance, (d) membrane area, (e) sweep gas flow rate, and (f) S/C ratio varied in low, medium, and high range.

	(a) Reactor			**(b) Compressor**			**(c) Labor**		
	Low	Medium	High	Low	Medium	High	Low	Medium	High
Maximum	16.1010	16.2556	16.3081	16.1037	16.7248	16.1678	14.8981	15.6844	16.4641
Minimum	4.6437	4.6705	4.5964	4.4529	4.3843	4.7103	4.2473	4.9387	5.7619
Mean	9.7391	9.7736	9.7733	9.7418	9.7677	9.8031	9.0355	9.7926	10.5381
Median	9.6542	9.6952	9.7179	9.6728	9.6677	9.7431	9.0062	9.7752	10.5023
Standard deviation	2.1922	2.1935	2.1863	2.1923	2.1908	2.1854	2.0867	2.1199	2.1005
	(d) Reactant			**(e) Natural gas**			**(f) Electricity**		
	Low	Medium	High	Low	Medium	High	Low	Medium	High
Maximum	11.1871	14.0570	16.5043	15.8683	16.2509	16.7044	16.4504	16.4909	16.5678
Minimum	4.5586	6.4955	8.0174	4.7191	4.9063	4.8804	4.7168	4.6090	4.6640
Mean	7.5525	9.7790	11.9944	9.7392	9.7513	9.8597	9.7076	9.8180	9.8565
Median	7.5011	9.7259	11.9160	9.6615	9.6854	9.7967	9.6487	9.7459	9.8243
Standard deviation	1.0984	1.2078	1.3111	2.1869	2.1721	2.1921	2.1980	2.1886	2.1791

steam reforming (MSR) in serial reactors and membrane filters were conducted, and detailed effects of involved techno-economic parameters on H_2 production rate, CO_2 emission, and unit H_2 production cost were investigated with machine learning (ML) based predictive model, following process simulation and numerical model.

For technical performance, the H_2 production system of MSR in serial reactors and membrane filters shows H_2 production rate distribution having maximum of 2.9953 kmol h^{-1}, minimum of 2.2598 kmol h^{-1}, mean of 2.7250 kmol h^{-1}, median of 2.7395 kmol h^{-1}, mode of 2.6776 kmol h^{-1}, and standard deviation of 0.1131 showing high dependence on # of reactors. Interestingly, divided regions of CO_2 emission distribution were observed in a study for environmental feasibility due to its difference of latent heat showing maximum of 0.1506 and 0.1107 kmol h^{-1}, minimum of 0.1167 and 0.0939 kmol h^{-1}, mean of 0.1388 and 0.1036 kmol h^{-1}, median of 0.1395 and 0.1041 kmol h^{-1}, mode of 0.1429 and 0.1088 kmol h^{-1}, and standard deviation of 0.0063 and 0.0036, respectively. In addition, very wide distribution for unit H_2 production cost with maximum of 13.1487 \$ kgH_2^{-1}, minimum of 4.3122 \$ kgH_2^{-1}, mean of 8.1045 \$ kgH_2^{-1}, median of 8.0866 \$ kgH_2^{-1}, mode of 8.9912 \$ kgH_2^{-1}, and standard deviation of 1.6856 were obtained indicating lower effect of # of reactors than on other two feasibility studies.

To investigate more detailed effects of 12 techno-economic parameters on each feasibility, individual technical, environmental, and economic analyses where only parameter was varied in low, medium, and high ranges with other parameters varied in total range were conducted. For effects of technical parameters on technical performance, # of reactors and operating temperature were reported as most influential factors showing +27.6% and +21.3% increased medians in high variation range, respectively. For effects of technical parameters on environmental performance, # of reactors was also figured out as most influential factor showing frequency of 74.23% in high emission level and high variation range. Similar to # of reactors, S/C ratio also exhibits frequencies of 78.01% (medium variation range) and 77.67% (high variation range) in high emission level. For economic feasibility, the fact that technical parameters of operating temperature and # of reactors and economic parameters of reactant and labor have strong impacts on unit H_2 production cost was reported with −12.9%, −6.1%, −37.1%, and −14.2% decreased medians, respectively.

In short, based on ML-based model for prediction of technical, environmental, and economic feasibility of H_2 production system MSR in serial reactors and membrane filters, # of reactors and operating temperature for technical performance, # of reactors and S/C ratio for environmental performance, and operating temperature, # of reactors, reactant, and labor for economic performance were confirmed as most influential factors.

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

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (NRF-2019M1A2A2065614) and Hydrogen Energy Innovation Technology Development Program of the National Research Foundation of Korea (NRF) funded by the Korean government (Ministry of Science and ICT (MSIT)) (NRF-2019M3E6A1064290).

References

[1] H. Nam, H. Nam, S. konishi, Techno-economic analysis of hydrogen production from the nuclear fusion-biomass hybrid system, International Journal of Energy Research (2020).
[2] Irfan Ahmad Gondal, Syed Athar Masood, Rafiullah Khan, Green hydrogen production potential for developing a hydrogen economy in Pakistan, International Journal of Hydrogen Energy 43 (12) (2018) 6011–6039.
[3] Mari Voldsund, Kristin Jordal, Rahul Anantharaman, Hydrogen production with CO_2 capture, International Journal of Hydrogen Energy 41 (9) (2016) 4969–4992.
[4] Ibrahim Dincer, Canan Acar, Innovation in hydrogen production, International Journal of Hydrogen Energy 42 (22) (2017) 14843–14864.
[5] Sivaprakash Sengodan, Rong Lan, John Humphreys, Dongwei Du, Wei Xu, Huanting Wang, Shanwen Tao, Advances in reforming and partial oxidation of hydrocarbons for hydrogen production and fuel cell applications, Renewable and Sustainable Energy Reviews 82 (2018) 761–780.
[6] S.S. Seyitoglu, I. Dincer, A. Kilicarslan, Energy and exergy analyses of hydrogen production by coal gasification, International Journal of Hydrogen Energy 42 (4) (2017) 2592–2600.
[7] Wataru Yonamine, Sivasakthivel Thangavel, Hidenori Ohashi, Chihiro Fushimi, Performance analysis of a water–gas shift membrane reactor for integrated coal gasification combined cycle plant, Energy Conversion and Management 174 (2018) 552–564.
[8] Yukio Hayakawa, Tomonori Miura, Kota Shizuya, Shintaro Wakazono, Kenya Tokunaga, Shinji Kambara, Hydrogen production system combined with a catalytic reactor and a plasma membrane reactor from ammonia, International Journal of Hydrogen Energy 44 (20) (2019) 9987–9993.
[9] Guozhao Ji, Ming Zhao, Geoff Wang, Computational fluid dynamic simulation of a sorption-enhanced palladium membrane reactor for enhancing hydrogen production from methane steam reforming, Energy 147 (2018) 884–895.
[10] Shahid Ali, Kim Sørensen, Mads P. Nielsen, Modeling a novel combined solid oxide electrolysis cell (SOEC) - Biomass gasification renewable methanol production system, Renewable Energy 154 (2020) 1025–1034.
[11] Jiaqi Zhao, Run Shi, Zhenhua Li, Chao Zhou, Tierui Zhang, How to make use of methanol in green catalytic hydrogen production? Nano Select 1 (1) (2020) 12–29.
[12] Manhee Byun, Boreum Lee, Hyunjun Lee, Seungkyo Jung, Hyunjin Ji, Hankwon Lim, Techno-economic and environmental assessment of methanol steam reforming for H_2 production at various scales, International Journal of Hydrogen Energy 45 (46) (2020) 24146–24158.
[13] A. Helmi, E. Fernandez, J. Melendez, D.A.P. Tanaka, F. Gallucci, M.V.S. Annaland, Fluidized bed membrane reactors for ultra pure $_{H}2$ production—A step forward towards commercialization, Molecules 21 (2016) 376–393.
[14] Seung-Won Choi, David S. Sholl, Sankar Nair, Jason S. Moore, Yujun Liu, Ravindra S. Dixit, John G. Pendergast, Modeling and process simulation of hollow fiber membrane reactor systems for propane dehydrogenation, AIChE Journal 63 (10) (2017) 4519–4531.
[15] A. Wunsch, M. Mohr, P. Pfeifer, Intensified LOHC-dehydrogenation using multi-stage microstructures and Pd-Based membranes, Membranes 8 (2018) 112–124.
[16] Moshe Sheintuch, Olga Nekhamkina, Architecture alternatives for propane dehydrogenation in a membrane reactor, Chemical Engineering Journal 347 (2018) 900–912.
[17] J.O. Abe, A.P.I. Popoola, E. Ajenifuja, O.M. Popoola, Hydrogen energy, economy and storage: Review and recommendation, International Journal of Hydrogen Energy 44 (2019) 15072–15086.
[18] P.P. Edwards, V.L. Kuznetsov, W.I.F. David, N.P. Brandon, Hydrogen and fuel cells: Towards a sustainable energy future, Energy Policy 36 (2008) 4356–4362.
[19] L. BARRETO, A. MAKIHIRA, K. RIAHI, The hydrogen economy in the 21st century: a sustainable development scenario, International Journal of Hydrogen Energy 28 (3) (2003) 267–284.
[20] A. MIDILLI, I. DINCER, Hydrogen as a renewable and sustainable solution in reducing global fossil fuel consumption, International Journal of Hydrogen Energy 33 (16) (2008) 4209–4222.
[21] C.C. Elam, C.E.G. Padro, G. Sandrock, A. Luzzi, P. Lindblad, E.F. Hagen, Realizing the hydrogen future: the International Energy Agency's efforts to advance hydrogen energy technologies, International Journal of Hydrogen Energy 28 (2003) 601–607.
[22] Kaveh Mazloomi, Chandima Gomes, Hydrogen as an energy carrier: Prospects and challenges, Renewable and Sustainable Energy Reviews 16 (5) (2012) 3024–3033.
[23] Stefan Schneider, Siegfried Bajohr, Frank Graf, Thomas Kolb, State of the art of hydrogen production via pyrolysis of natural gas, ChemBioEng Reviews 7 (5) (2020) 150–158.
[24] C.H.S. Moura, J.L. Silveira, W. de Q. Lamas, Dynamic production, storage, and use of renewable hydrogen: A technical-economic-environmental analysis in the public transport system in São Paulo state, Brazil, International Journal of Hydrogen Energy 45 (2020) 31585–31598.
[25] Canan Acar, Ibrahim Dincer, Review and evaluation of hydrogen production options for better environment, Journal of Cleaner Production 218 (2019) 835–849.
[26] Canan Acar, Ahmet Beskese, Gül Tekin Temur, Sustainability analysis of different hydrogen production options using hesitant fuzzy AHP, International Journal of Hydrogen Energy 43 (39) (2018) 18059–18076.
[27] Junjie Li, Wanjing Cheng, Comparative life cycle energy consumption, carbon emissions and economic costs of hydrogen production from coke oven gas and coal gasification, International Journal of Hydrogen Energy 45 (51) (2020) 27979–27993.
[28] R.Y. Kannah, S. Kavitha, O.P. Peethi, G. Karthikeyan, N.V. Kumar, R. Banu Dai-Viet, Techno-economic assessment of various hydrogen production methods – A review, Bioresource Technology 319 (2021), 124175.
[29] Syed Ahsan Ali Shah, Feasibility study of renewable energy sources for developing the hydrogen economy in Pakistan, International Journal of Hydrogen Energy 45 (32) (2020) 15841–15854.
[30] Rami S. El-Emam, Hasan Özcan, Comprehensive review on the techno-economics of sustainable large-scale clean hydrogen production, Journal of Cleaner Production 220 (2019) 593–609.
[31] Yaser Khojasteh Salkuyeh, Bradley A. Saville, Heather L. MacLean, Techno-economic analysis and life cycle assessment of hydrogen production from different biomass gasification processes, International Journal of Hydrogen Energy 43 (20) (2018) 9514–9528.
[32] Jingwei Chen, Wenwen Xu, Xiaomin Wu, Jiaqiang E, Na Lu, Taosheng Wang, Hongyan Zuo, System development and environmental performance analysis of a pilot scale microbial electrolysis cell for hydrogen production using urban wastewater, Energy Conversion and Management 193 (2019) 52–63.
[33] Mert Temiz, Nader Javani, Design and analysis of a combined floating photovoltaic system for electricity and hydrogen production, International Journal of Hydrogen Energy 45 (5) (2020) 3457–3469.
[34] S. Sadeghi, S. Ghandehariun, M.A. Rose, Comparative economic and life cycle assessment of solar-based hydrogen production for oil and gas industries, Energy 208 (2020), 118347.
[35] Juheon Heo, Boreum Lee, Hyunjun Lee, Hankwon Lim, Integrative Technical, economic, and environmental feasibility analysis for ethane steam reforming in a membrane reactor for H_2 production, ACS Sustainable Chemistry Engineering 8 (18) (2020) 7011–7019.
[36] World Economic Forum. (2016, January). The future of jobs: Employment, skills and workforce strategy for the fourth industrial revolution. In Global challenge insight report. Geneva: World Economic Forum.
[37] A.-D. Pham, H.-J. Ahn, High precision reducers for industrial robots driving 4th industrial revolution: state of arts, analysis, design, performance evaluation and perspective, International Journal of Precision Engineering and Manufacturing-Green Technology 5 (2018) 519–533.
[38] Manfred Krafft, Laszlo Sajtos, Michael Haenlein, Challenges and opportunities for marketing scholars in times of the fourth industrial revolution, Journal of Interactive Marketing 51 (2020) 1–8.
[39] M.B. Rapanyane, F.R. Sethole, The rise of artificial intelligence and robots in the 4th Industrial Revolution: implications for future South African job creation, Contemporary Social Science 15 (2020) 489–501.
[40] Emil Blixt Hansen, Simon Bøgh, Artificial intelligence and internet of things in small and medium-sized enterprises: A survey, Journal of Manufacturing Systems 58 (2021) 362–372.
[41] Siddharth Sridhar, Adam Hahn, Manimaran Govindarasu, Cyber–physical system security for the electric power grid, Proceedings of the IEEE 100 (1) (2012) 210–224.
[42] Shuang Huang, Chun-Jie Zhou, Shuang-Hua Yang, Yuan-Qing Qin, Cyber-physical system security for networked industrial processes, International Journal of Automation and Computing 12 (6) (2015) 567–578.
[43] D. Gollmann, P. Gurikov, A. Isakov, M. Krotofil, J. Larsen, A. Winnicki, Cyber-physical systems security: Experimental analysis of a vinyl acetate monomer plant, in: Proceedings of the 1st ACM Workshop on Cyber-Physical System Security, 2015, pp. 1–12.
[44] J.H. Lee, J. Shin, M.J. Realff, Machine learning: Overview of the recent progresses and implications for the process system engineering field, Computers and Chemical Engineering 114 (2018) 111–121.
[45] Steven L. Brunton, Bernd R. Noack, Petros Koumoutsakos, Machine Learning for Fluid Mechanics, Annual Review of Fluid Mechanics 52 (1) (2020) 477–508.
[46] Grégoire Montavon, Matthias Rupp, Vivekanand Gobre, Alvaro Vazquez-Mayagoitia, Katja Hansen, Alexandre Tkatchenko, Klaus-Robert Müller, O. Anatole

von Lilienfeld, Machine learning of molecular electronic properties in chemical compound space, New Journal of Physics 15 (9) (2013) 095003, https://doi.org/10.1088/1367-2630/15/9/095003.
[47] S. Hwangbo, R. Al, G. Sin, An integrated framework for plant data-driven process modeling using deep-learning with Monte-Carlo simulation, Computers and Chemical Engineering 143 (2020) 107071–107084.
[48] V. Venkatasubramanian, The promise of artificial intelligence in chemical engineering: Is it here, finally? AIChE Journal 65 (2019) 466–478.
[49] Hanyu Gao, Thomas J. Struble, Connor W. Coley, Yuran Wang, William H. Green, Klavs F. Jensen, Using Machine Learning To Predict Suitable Conditions for Organic Reactions, *ACS Central*, Science 4 (11) (2018) 1465–1476.
[50] Yongliang Yan, Tobias Mattisson, Patrick Moldenhauer, Edward J. Anthony, Peter T. Clough, Applying machine learning algorithms in estimating the performance of heterogeneous, multi-component materials as oxygen carriers for chemical-looping processes, Chemical Engineering Journal 387 (2020) 124072, https://doi.org/10.1016/j.cej.2020.124072.
[51] Connor W. Coley, Regina Barzilay, Tommi S. Jaakkola, William H. Green, Klavs F. Jensen, Prediction of Organic Reaction Outcomes Using Machine Learning, *ACS Central*, Science 3 (5) (2017) 434–443.
[52] Aayush R. Singh, Brian A. Rohr, Joseph A. Gauthier, Jens K. Nørskov, Predicting Chemical Reaction Barriers with a Machine Learning Model, Catalysis Letters 149 (9) (2019) 2347–2354.
[53] B.R. Hough, D.A.C. Beck, D.T. Schwartz, J. Pfaendtner, Application of machine learning to pyrolysis reaction networks: Reducing model solution time to enable process optimization, Computers and Chemical Engineering 104 (2017) 56–63.
[54] J.W. Chew, R.A. Cocco, Application of machine learning methods to understand and predict circulating fluidized bed riser flow characteristics, Chemical Engineering Science 217 (2020) 115503–115512.
[55] Weijun Zhu, Xingwang Liu, Xu Hou, Jiayao Hu, Zhenheng Diao, Application of machine learning to process simulation of n-pentane cracking to produce ethylene and propene, Chinese Journal of Chemical Engineering 28 (7) (2020) 1832–1839.
[56] Gelayol Golkarnarenji, Minoo Naebe, Khashayar Badii, Abbas S. Milani, Reza N. Jazar, Hamid Khayyam, A machine learning case study with limited data for prediction of carbon fiber mechanical properties, Computers in Industry 105 (2019) 123–132.
[57] S. Siracar, T.C. Golden, Purification of Hydrogen by Pressure Swing Adsorption, Separation Science and Technology 35 (2000) 667–687.
[58] R. Turton, R.C. Bailie, W.B. Whiting, J.A. Shaeiwitz, D. Bhattacharyya, Analysis, synthesis, and design of chemical processes, fourth ed., Pearson, New Jersey, 2013.
[59] E.W. Feedstock Equipment design and cost estimation for small modular biomass systems, synthesis gas cleanup, and oxygen separation equipment Task 1: cost Estimates of Small Modular Systems, National Renewable Energy Laboratory 2006.
[60] C. Yang, J. Ogden, Determining the lowest-cost hydrogen delivery mode, International Journal of Hydrogen Energy 32 (2007) 268–286.
[61] Boreum Lee, Heetae Chae, Nak Heon Choi, Changhwan Moon, Sangbong Moon, Hankwon Lim, Economic evaluation with sensitivity and profitability analysis for hydrogen production from water electrolysis in Korea, International Journal of Hydrogen Energy 42 (10) (2017) 6462–6471.
[62] https://www.methanex.com/our-business/pricing [Last Accessed 25 December 2020].
[63] http://www.citygas.or.kr/info/charge.jsp [Last Accessed 25 December 2020].
[64] Chunfeng Song, Qingling Liu, Na Ji, Yasuki Kansha, Atsushi Tsutsumi, Optimization of steam methane reforming coupled with pressure swing adsorption hydrogen production process by heat integration, Applied Energy 154 (2015) 392–401.
[65] Leonardo Roses, Giampaolo Manzolini, Stefano Campanari, Ellart De Wit, Michael Walter, Techno-economic assessment of membrane reactor technologies for pure hydrogen production for fuel cell vehicle fleets, Energy & Fuels 27 (8) (2013) 4423–4431.
[66] Dana Havas, Haiqing Lin, Optimal membranes for biogas upgrade by removing CO_2: High permeance or high selectivity? Separation Science and Technology 52 (2) (2017) 186–196.
[67] Z. Hoffman, Simulation and economic evaluation of coal gasification with SETS reforming process for power production, LSU Master's Theses (2005) 2269.
[68] B.J. Gim, J.W. Kim, H.M. Ko, Economic evaluation of domestic low-temperature water electrolysis hydrogen production, Transactions of the Korean Hydrogen and New Energy Society 22 (2011) 559–567.
[69] C. Cortes, V. Vapnik, Support-vector networks, Machine Learning 20 (1995) 273–297.
[70] I. Ahmad, M. Basheri, M.J. Iqbal, A. Rahim, Performance comparison of support vector machine, random forest, and extreme learning machine for intrusion detection, IEEE Access 6 (2018) 33789.
[71] D.P. Lewis, T. Jebara, W.S. Noble, Support vector machine learning from heterogeneous data: an empirical analysis using protein sequence and structure, Bioinformatics 22 (2006) 2753–2760.
[72] Mathieu Wauters, Mario Vanhoucke, Support vector machine regression for project control forecasting, Automation in Construction 47 (2014) 92–106.
[73] U. Thissen, M. Pepers, B. Üstün, W.J. Melssen, L.M.C. Buydens, Comparing support vector machines to PLS for spectral regression applications, Chemometrics and Intelligent Laboratory Systems 73 (2) (2004) 169–179.
[74] James J. Rothwell, Martyn N. Futter, Nancy B. Dise, A classification and regression tree model of controls on dissolved inorganic nitrogen leaching from European forests, Environmental Pollution 156 (2) (2008) 544–552.
[75] P. TITTONELL, K. SHEPHERD, B. VANLAUWE, K. GILLER, Unravelling the effects of soil and crop management on maize productivity in smallholder agricultural systems of western Kenya—An application of classification and regression tree analysis, Agriculture, Ecosystems & Environment 123 (1-3) (2008) 137–150.
[76] Guangxin Ren, Yujie Wang, Jingming Ning, Zhengzhu Zhang, Using near-infrared hyperspectral imaging with multiple decision tree methods to delineate black tea quality, Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy 237 (2020) 118407, https://doi.org/10.1016/j.saa.2020.118407.
[77] Aleena Swetapadma, Anamika Yadav, A novel decision tree regression-based fault distance estimation scheme for transmission lines, IEEE Transactions on Power Delivery 32 (1) (2017) 234–245.
[78] J.Q. Shi, T. Choi, Gaussian process regression analysis for functional data, CRC Press, 2011.
[79] Ananth Ranganathan, Ming-Hsuan Yang, Jeffrey Ho, Online sparse Gaussian process regression and its applications, IEEE Transactions on Image Processing 20 (2) (2011) 391–404.
[80] H. Lee, B. Lee, M. Byun, H. Lim, Economic and environmental analysis for PEM water electrolysis based on replacement moment and renewable electricity resources, Energy Conversion and Management 224 (2020), 113477.
[81] M. Byun, H. Kim, C. Choe, H. Lim, Conceptual feasibility studies for cost-efficient and bi-functional methylcyclohexane dehydrogenation in a membrane reactor for H_2 storage and production, Energy Conversion and Management 227 (2021), 113576.