Journal of Membrane Science 736 (2025) 124708

Contents lists available at ScienceDirect

Journal of Membrane Science

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

Membrane preparation assisted by integration of machine learning and response surface methodology for CO_2 separation

Lucía Carrillo-Sánchez [a,b], Carlos Téllez [a,b,*], Joaquín Coronas [a,b,**]

[a] Instituto de Nanociencia y Materiales de Aragón (INMA), CSIC-Universidad de Zaragoza, Zaragoza, 50018, Spain
[b] Chemical and Environmental Engineering Department, Universidad de Zaragoza, Zaragoza, 50018, Spain

ARTICLE INFO

Keywords:
CO_2 separation
Polymer Pebax® 1657
Membrane optimization
Machine learning
Response surface methodology
Predictive modeling

ABSTRACT

The separation of carbon dioxide (CO_2) is presented as a current challenge in the environment and energy sector. The primary reason for this is to control the emissions of this gas into the atmosphere and the upgrading of biomethane. In this context, the membrane separation technology seems to be a very sustainable promising tool for such tasks. This work presents a machine learning (ML) study, based on a database created from membrane preparation conditions and gas separation records from the literature, achieved for the CO_2/N_2 and CO_2/CH_4 mixtures using dense membranes of thermoplastic elastomer Pebax® 1657. A comparative analysis of three different ML models was carried out: multiple linear regression, decision tree and random forest. This last algorithm demonstrates the best performance in statistics terms of coefficient of determination and root mean square error. In addition, the combination of the ML random forest with a method based on the design of experiments with response surface methodology (RSM) allowed to identify the favorable conditions for the membrane synthesis, with the objective of enhancing the CO_2 separation performance. This resulted in prepared membranes in the laboratory considering the proposed conditions by RSM with CO_2 permeability and CO_2/X selectivity values of 115 Barrer and 43.5 and 132 Barrer and 16.4 for the CO_2/N_2 and CO_2/CH_4 mixtures, respectively, at 35 °C.

1. Introduction

The carbon dioxide (CO_2) molecule, a principal contributor to global warming due to its role in greenhouse gas emissions, is a product associated with various processes, including the combustion of fossil fuels [1]. The removal, capture and storage of this acid gas can be conducted with the use of conventional technologies such as adsorption, absorption or cryogenic distillation. These present certain disadvantages, comprising elevated energy consumption and substantial environmental impact. Therefore, new improved separation processes are necessary with a greater efficiency, simplicity and smaller footprint [2].

The utilization of membranes for the separation of CO_2 has been proposed as a novel alternative for the capture of CO_2 in post-combustion processes as well as in other processes related to sustainability such as biogas upgrading. This separation proposal offers several advantages, including cost savings in processing, energy efficiency, simplified mechanics and compact designs [3]. Among the various types of membranes (organic, metallic or inorganic), polymeric membranes are particularly promising, due to their flexibility, cost-effectiveness and ease of preparation [4]. Moreover, the acquisition of membranes that exhibit enhanced selectivity for CO_2 over gases such as N_2 and CH_4 is possible, based on the differences in solubility and diffusivity of these gases within the membrane matrix [5]. In these cases, a variety of polymer types are employed, including the PEBA (polyether block amide) series, with the Pebax® commercial brand being the most prominent, which has demonstrated efficacy in the treatment of these gases [6,7]. PEBA is a thermoplastic elastomer that is formed by block copolymers: a “hard” block of aliphatic polyamide (PA), that contributes to mechanical stability, and a “soft or flexible” block of amorphous polyether (PE), that enables high permeation flow and selective CO_2 transport. In fact, PEBA has demonstrated a high degree of affinity for CO_2, due to its polar ether group. The variation in the proportion of each block, as well as their intrinsic composition, causes differences in separation and physical properties [8]. Specifically, the polymer called Pebax® 1657, with suitable properties for CO_2 separation [9], illustrated in Fig. 1 and composed of polyethylene oxide (PEO) and PA in a mass

* Corresponding author. Instituto de Nanociencia y Materiales de Aragón (INMA), CSIC-Universidad de Zaragoza, Zaragoza, 50018, Spain.
** Corresponding author. Instituto de Nanociencia y Materiales de Aragón (INMA), CSIC-Universidad de Zaragoza, Zaragoza, 50018, Spain.
E-mail addresses: ctellez@unizar.es (C. Téllez), coronas@unizar.es (J. Coronas).

https://doi.org/10.1016/j.memsci.2025.124708
Received 15 July 2025; Received in revised form 12 September 2025; Accepted 13 September 2025
Available online 15 September 2025

ratio of 60:40, is employed in this study.

In these polymeric membranes, the efficiency of the gas separation process is determined by the selectivity and the permeability of the membrane material. In addition, according to selectivity and flux density, they can be classified as: porous and non-porous or dense, among others. The most common preparation methods for dense membranes are melt extrusion and solution casting [10]. In this study, the second mentioned method is used, which consists of the preparation of the membrane from a polymer solution with a specific solvent. Subsequently, the deposition or casting of the solution on a surface is carried out, in such a way that the membrane will be formed after the complete evaporation of the solvent. During membrane fabrication, it is important to consider the parameters that influence this process, since they can affect the structure of the membrane and its properties. Some of these parameters or variables may be the type of polymer, concentration of the solution, type of solvent and its volatility, the material of the casting surface, operating times (casting and solution) and the casting temperature [11].

Achieving the highest possible performance from the use of these membranes as well as their optimization requires a detailed study of the experimental variables that are part of the synthesis and preparation of the membranes, in combination with the gas separation conditions. This is a complex task that involves an important investment of resources, including high cost and significant effort in terms of time and experimental work, usually carried out through a trial-and-error strategy.

Consequently, the use of artificial intelligence (AI), due to its capacity to emulate human learning and reasoning abilities, is growing exponentially among society and its application and potency is increasing in different fields, such as medicine or materials technology, among others. Specifically, within the domain of AI, there are tools known as “machine learning (ML)”, which can generate models that, upon undergoing training, exhibit the capacity to discern between classes or predict a desired property. This capability arises from their ability to capture nonlinear relationships between the input variables and the output variable, producing the desired result [12,13]. Therefore, ML techniques facilitate a rapid examination of the variables and the optimization of processes, allowing the evaluation of the object problem. This results in a reduction of required resources and an enhancement in efficiency. To achieve this objective, a ML model can be integrated with other resources, such as the design of experiments with a response surface method [14], molecular design [15] and molecular simulation [16]. Moreover, ML has been previously used in the field of gas separation membranes [17–19] to gain insight into the prediction of the polymer characteristics and discover new ones [20,21], predict the separation properties of a given polymer (e.g. polyimide) [22], find a rapid assessment of MOF based mixed matrix membranes (MMMs) [16], estimate missing values of permeabilities of gases using exiting values for certain gases [23], and aid to choose the best ionic liquid@MOF composite to produce a MMM for CO_2 separation [24], among others. Some works try to use the knowledge obtained with ML to prepare MOF-based MMMs that separate CO_2 [11]. However, most of these approaches lack experimental exploitation and optimization of the knowledge gained from the ML.

The main objective of this work is to use ML to identify the key parameters affecting the preparation of dense membranes of polymer Pebax® 1657 and then apply the information gathered to improve their separation performance. Thus, the work seeks to design a model for the selectivity of membranes of Pebax® 1657, based on experimental variables and considering their respective effects on these membranes. This implies the configuration of a specific database related to Pebax® 1657 membranes from the literature including relevant information on their synthesis and separation behavior, the development of a model able to predict selectivity values based on previous experiments and, therefore, the estimation of the selectivity performance of newly synthesized membranes. Using the AI-generated model, design of experiments (DOE) with the response surface method (RSM) is used to obtain optimal membrane preparation conditions to maximize the separation selectivity. Membranes were therefore subsequently prepared in the laboratory under different conditions, including optimal conditions, and tested for CO_2 separation, validating the methodology. This is the first time that AI has been integrated with DOE-RSM in the tangible membrane preparation for gas separation. In addition, the database was revamped with the additional experiments performed. In short, this approach (see Fig. 1) accelerates the identification of the most influential membrane preparation conditions by optimizing the parameters that lead to optimal selectivity values.

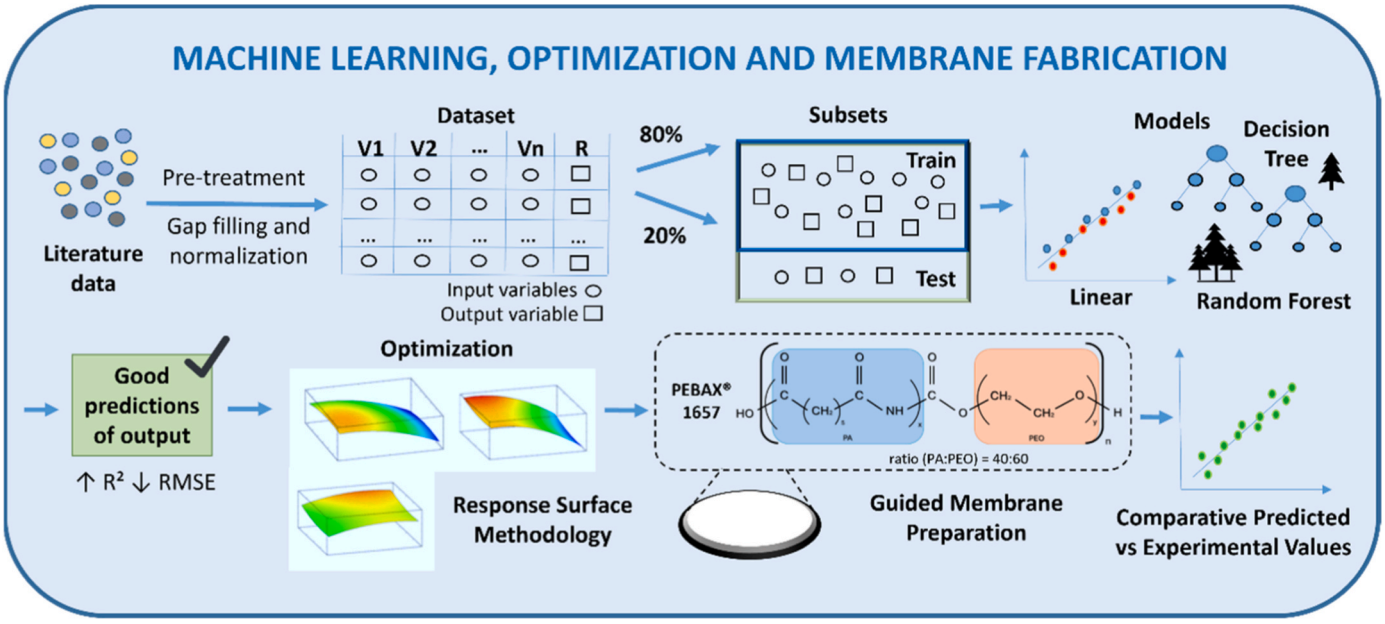

Fig. 1. Workflow integrating ML, optimization and membrane fabrication. Literature data are preprocessed and used to train several predictive models. After model validation, a response surface methodology is applied to optimize the preparation conditions at the laboratory, which are experimentally validated.

2. Materials and methods

2.1. Materials

Pebax® MH 1657 copolymer (60 wt% polyethylene oxide (PEO) and 40 wt% aliphatic polyamide 6 (PA6)) in pellet shape was kindly supplied by Arkema, France. Regarding the solvent, absolute ethanol and deionized water were purchased from Gilca, Spain. Additionally, gases used for gas permeation measurements (i.e., CO_2, N_2, CH_4, He) were supplied by Linde S.A, Spain. These gases were of a purity level greater than 99.9 %, a standard that ensures their quality and reliability for scientific applications.

2.2. Preparation of the membranes

A specific quantity of Pebax® MH 1657 pellets (0.36 g) was dissolved in a mixture of ethanol and deionized water (70:30 wt%) by stirring at several temperatures in the 70–140 °C range for 3 h to obtain solutions of this polymer at different polymer concentrations (in 1.8–10 wt% range). The amount of polymer added was constant to ensure a similar thickness (approximately 60–80 μm) in all membranes under comparison. However, the amount of solvent varied. After cooling, solutions were poured into a Petri dish and dried for one or two days at room temperature or a higher temperature (always less than 60 °C) and under controlled evaporation, until the solvent was almost completely dissipated. Subsequently, membranes were placed in a vacuum oven as a conditioning step from 30 to 60 °C for 4 h–6 h to ensure the complete elimination of any residual solvents that may be present in them. The membrane thickness is measured using a digital disk micrometer (Digimatic Micrometer Quickmike, Mitutoyo Corp.), and five measurements are taken at different points on each membrane, with the thickness being calculated as the average of those measurements. The synthesized membranes are designated as P1657DM_CX, where X denotes specific working conditions, as indicated in Table S1 of the Supplementary Information.

2.3. Gas permeation measurements

Membranes were cut with an area of 12.6 cm^2 and placed in a module consisting of two pieces of stainless steel and a porous support disk (Mott Co.) with a nominal pore size of 20 μm, being gripped inside with Viton o-rings. Gas separation measurements were conducted using feed mixtures of CO_2/N_2 (15:85 cm^3 (STP)·min^{-1}) and CO_2/CH_4 (50:50 cm^3 (STP)·min^{-1}) which were controlled by two Alicat Scientific mass flow controllers (MC-100CCM-D for CO_2 and MC-200CCM-D for N_2 and CH_4) at a pressure of 3 bar on the feed side and different temperatures (25, 35 and 50 °C). The temperature of experiments is controlled by a UNE 200 Memmert oven in which the module is placed. Moreover, the permeate side of the membranes was swept with 10 cm^3 (STP)·min^{-1} of He at a pressure of 1 bar, which was also controlled by a mass flow controller (MC-200CCM-D, Alicat Scientific). Regarding the concentrations of CO_2, N_2 and CH_4 in the output stream, they were measured online using an Agilent 990 micro-gas chromatograph. The permeability value was calculated in Barrer (10^{-10} cm^3 (STP)·cm·cm^{-2}·s^{-1}·$cmHg^{-1}$) and the selectivity was determined as the ratio of the permeabilities of the gases in the mixture.

2.4. Data preparation and machine learning models

2.4.1. Data collection from literature

The aim of this section is to analyze and compile the information available in the literature that is related to the study of dense membranes for separation of CO_2 from other gases (N_2 or CH_4) with the selected Pebax® 1657 polymer. This will allow the constitution of a database to be exploited through a machine learning (ML) model. It is important to examine the conditions and methodologies during the fabrication and measurement processes, as these factors will define the behavior of each membrane, and it will allow future research to improve the results. For this purpose, academic and scientific collections such as Web of Science and Google Scholar were used through keywords such as "CO_2 separation", "Pebax® 1657" and "Membrane". A total of 55 articles [1,5,25–77] were collected, representing 109 samples or entries into our database and in Supplementary Information, Fig. S1 and Table S2 present a summary about the distribution of the variables and some data related with this dataset collection, respectively. However, it is important to note that a greater volume of data was collected; nevertheless, these data are not directly used to construct the model, as some samples contain a high number of incomplete variables and outliers or they do not appear to a high degree to be representative, which can harm the model by making it difficult to learn. Initially, the total collected data consisted of 1846 entries corresponding to 139 samples, of which 30 samples were eliminated, and 218 entries were presented as missing values, representing 22 % of samples and 12 % of data entries, respectively. The main variables where gaps existed were the percentage of CO_2 in the gas mixture, the membrane thickness and temperature and time of the conditioning step. From a future perspective, a valuable improvement would be to collect additional data to increase the representativeness of a broader dataset, thereby allowing the study of other

Table 1
Descriptive statistics (range, mean and standard deviation (sd)) of input variables and output variable that participated in the machine learning model.

Variable (units)	Range	Mean ± sd	Description
Input variables			
Solvent	0.0–2.0	1.5 ± 0.8	"Butanol", "DMF" and "Ethanol/water" were normalized to 0, 1 and 2, respectively.
Polymer conc. (wt%)	0.6–10.0	4.0 ± 1.7	–
Thickness (μm)	30.0–110.0	60.2 ± 17.5	–
Solution temp. (°C)	70.0–140.0	92.3 ± 17.5	–
Solution time (h)	2.0–24.0	8.0 ± 8.3	–
Casting temp. (°C)	25.0–70.0	33.0 ± 13.0	–
Casting time (h)	10.0–72.0	34.1 ± 14.2	–
Conditioning temp. (°C)	30.0–100.0	54.7 ± 11.4	–
Conditioning time (h)	0.0–72.0	17.9 ± 13.7	–
Separation temp. (°C)	25.0–55.0	30.0 ± 6.5	–
Pressure (bar)	1.0–10.0	3.3 ± 2.1	–
CO_2 percentage (%)	10.0–100.0	54.0 ± 35.2	–
Gas mixture	0.0–1.0	0.4 ± 0.5	"CO_2/CH_4" and "CO_2/N_2" were normalized to 0 and 1.
Output variable			
CO_2/N_2 selectivity	20.0–69.6	46.4 ± 11.2	–
CO_2/CH_4 selectivity	4.0–36.0	16.5 ± 5.0	–

conditions. Note that, even if the availability of data was limited, Pebax® 1657 was chosen due to the fact that, besides showing a relevant performance to separate CO_2 containing mixtures, it is one of the most studied membrane polymers, so that the study carried out here could be focused on an important set of preparation variables excluding the polymer type. In any case, data were subjected to several preprocessing steps, which included the imputation of missing values using the mean of the corresponding variable and the implementation of normalization to enhance the interpretation of the data by the model. Consequently, Table 1 shows the statistical data of each input (categorical and numeric) and output variable for the models used in this study.

2.4.2. Machine learning algorithms and models

Most ML algorithms are classified into supervised and unsupervised learning based on the examples provided during training. Supervised learning involves providing a set of input variables (x) along with its corresponding label or output variable (y), while in unsupervised learning only the input variables (x) are provided. Supervised learning tasks can be divided into classification tasks, which classify each sample into a certain group or category, and regression tasks, in which the algorithm establishes relationships between the input variables and the output variable [78]. Typically, the dataset is divided into different subsets, such as training and testing sets.

These algorithms were applied to various models, through the development of a Python code (see the supporting information) using the Google Colab tool. In this study, supervised learning algorithms, of growing complexity, were used for regression tasks, specifically, with certain types of models. *Multiple linear regression* seeks the connection between dependent and independent variables through their respective definitions as input and output characteristics since they are potentially interrelated, a phenomenon known as multicollinearity. In addition, the *decision tree* combines simple numerical tests. Each node contains a test that classifies a feature, and the branches extend from the parent nodes to the child nodes. Thus, they will be divided into new nodes in such a way that the iterations continue. Finally, *random forest* is an ensemble of decision trees that solves the overfitting problem typically associated with decision tree models [79].

In this study, 13 input variables were selected to estimate the CO_2 selectivity (i.e. CO_2/N_2 and CO_2/CH_4 separation selectivities, called CO_2/X throughout the study, X being either N_2 or CH_4) for dense membranes obtained from polymer Pebax® 1657. Selectivity was the output variable of the models. Note that, due to the limited amount of data available, both CO_2/N_2 and CO_2/CH_4 selectivities are integrated in the output of the model which can therefore differentiate between the two studied mixtures. The input characteristics are listed in Table 1. Regarding the types of solvents found in the literature, they are ethanol/water, propanol/water, dimethylformamide (DMF), dimethylacetamide (DMAc), propanol/butanol and butanol. Only the first three solvents from the above list were selected for the study, because they were representative of the dataset. The selected gas mixtures were CO_2/N_2 and CO_2/CH_4 because they are the most studied mixtures in the literature and are suitable for copolymer Pebax® 1657 with proven specific interaction with the CO_2 molecule given its CO_2-philic polar blocks or flexible polyether (complimented with the rigid polyamide blocks) [3].

2.4.3. Performance metrics

To evaluate the usefulness and operation of the models as well as to compare them, the metrics used to measure the quality were the root mean square error (RMSE) and the coefficient of determination (R^2) [80]. The RMSE is the square root of the mean of the errors, so it has the same units as the response or output variable. Its value indicates how far the predictions of the model are, on average, from the real value. The R^2 is dimensionless and describes the proportion of variance of the response or output variable explained by the model relative to the total variance. These metrics can be estimated using Eqs. (1) and (2):

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(P_i - y_i)^2}{n}} \quad (1)$$

$$R^2 = \frac{\sum_{i=1}^{n}(P_i - \bar{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2} \quad (2)$$

where P_i indicates the predicted output by the model, while y_i is the real value of the output, $\bar{y}_i$ represents the sample data mean, and n denotes the number of data entries.

2.4.4. Data normalization

Normalization is a data preprocessing technique that adjusts the characteristics or attributes of data within a specific range to facilitate the results interpretation. Thus, normalization is an essential part of data science and ML, increasing consistency and improving the performance, efficiency and precision of algorithms and models [81]. It is also possible to reduce differences in the scale and magnitude of data features, which is relevant because an attribute with larger values will have a dominant influence on the machine learning model. Many algorithms, such as regression or decision trees, require normalized data for its correct function.

There are numerous normalization techniques, but in this case min-max scaler normalization for numerical input was employed. This technique involves transforming the values of a characteristic into a range of specific values between zero and one. This objective is achieved by subtracting the minimum value of the data from each value and dividing the result by the difference between maximum and minimum values, as indicated in Eq. (3):

$$N_i = \frac{X_i - X_{min}}{X_{max} - X_{min}} \quad (3)$$

where N_i is the normalized data, X_i represents the original data and X_{min} and X_{max} denote the lowest and highest values of each attribute. This approach enables the comparison of attributes with different units and ranges of values within the same dataset, while ensuring a simple implementation. In the case of categorical inputs, the transformation of the data into numerical values was achieved through the implementation of the label encoding technique. Furthermore, a normalized data exhibits an enhanced robustness to potential outliers when compared with alternative normalization methods.

2.5. Study of variables that affect synthesis and measurement processes

The use of a ML model allows the prediction of various experiments; however, the integration of these models with other tools, such as design of experiments (DOE), enables the optimization of variables to achieve the best possible process and, consequently, maximizes the selectivity value of the membranes. DOE is a statistical technique that is used to optimize a process. This technique involves a simultaneous study of different factors to obtain optimal values. A specific range of each variable studied is defined in such a way that the DOE proposes a set of experiments. This optimization is achieved through a multivariate analysis using a Box-Behnken design (BDD) with Statgraphics Centurion software [82], which was used to investigate the relationship between variables that influence the process. The variables of polymer concentration (2–6 wt%), thickness (30–90 μm), polymer solution temperature (70–140 °C), casting temperature (25–75 °C) and conditioning temperature (30–100 °C) were analyzed using the response surface methodology (RSM) [83]. The polymer concentration was used as variable for the DOE because it affects the membrane processing with direct influence on the thickness and homogeneity of the dense film and also with implications in future developments dealing with thin film supported membranes. In any event, 46 combinations of experimental conditions were obtained, which selectivity values were calculated with the

optimum machine learning model.

For the optimization, the remaining variables were maintained constant during the optimization study. The numerical variables were set to specific values, solution time (3 h), casting time (48 h), conditioning time (4 h), gas separation temperature (35 °C), pressure (3 bar) and CO_2 percentage in the feed (15 %). In the case of categorical variables, ethanol/water (70:30 wt%) like polymer solvent and CO_2/N_2 mixture of gas separation were selected.

3. Results and discussion

As depicted in Figs. 1 and 2, a specific database was first built related to Pebax® 1657 membranes from the literature including relevant information on their synthesis and CO_2/X separation performance. Then, a Python code was written and executed using Google Colab platform to apply the dataset in the development of a ML model able to predict the CO_2 separation selectivity values based on experiments reported in the literature and, therefore, the estimation of the selectivity performance of newly synthesized membranes. Next, a methodology of DOE was combined with the ML model and the obtained output was applied to carry out specific experimental work preparing and testing Pebax® 1657 dense membranes for the separation of CO_2/N_2 and CO_2/CH_4 mixtures.

3.1. Correlations between input variables

It is necessary to know the correlation between each pair of input variables that are within the ML model, as a high coefficient between these variables can lead to overfitting. The values representing the Pearson correlation coefficients between the variables of the model for selectivity are presented in Table 2 and illustrated with the corresponding matrix of correlation given in Fig. S2. In this case, the coefficients have low values (less than 0.7), indicating an insufficient correlation between the input variables proposed in the models. It is notable that there are high values that are close to this limit, such as the relationship between the solvent and the solution preparation conditions. This is logical, since it is well established that these conditions are related to the boiling point of each solvent and its ability to dissolve the polymer. Another case is the inverse relationship between the casting conditions, since if a higher temperature is used to evaporate the solvent, less time will be required to obtain the membrane. Therefore, these variables are interconnected and, although they exhibit a certain relationship, it is important that they participate in the model. As a result, the 13 variables enumerated in section 2.4.1 have been selected for the estimation of the output variable, which is selectivity.

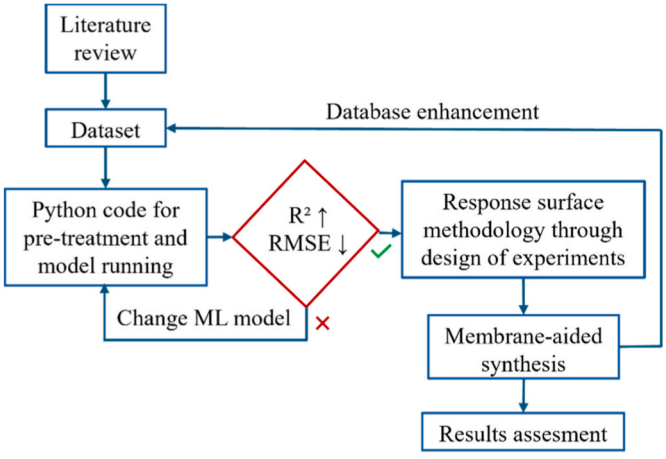

Fig. 2. Flow chart of the proposed procedure.

3.2. Comparison of selected algorithms

After the preparation of the database and subsequent analysis of variables, three distinct models (multiple linear regression, decision tree and random forest) were constructed for the purpose of evaluating their efficacy identifying the most suitable model for the given separation application. To this end, the dataset containing the input variables and the output or response variable was divided into two subsets: one subset designated as the "training set", which serves to train the model, comprising 80 % of the data, and another subset as the "test set", which is not known to the model and is, therefore, used for model evaluation, comprising 20 % of the data (see Table 3).

The values obtained for the multiple linear regression model are $R^2 = 0.833$ and RMSE = 7.10, for the training set, and $R^2 = 0.824$ and RMSE = 6.89, for the testing set. The decision tree model demonstrates R^2 and RMSE values of 0.999 and 0.56 and 0.655 and 9.12 for the training and test sets, respectively. Finally, the metrics for the random forest model are $R^2 = 0.973$ and RMSE = 2.85 and $R^2 = 0.962$ and RMSE = 3.20 for the training and test sets, respectively. The ideal result is a high R^2 value and a low RMSE value, indicating that the model can explain the response variable better and that there is a less difference between the model predictions and the real values from the database. Furthermore, it is important to highlight that the R^2 values for the training and test sets should be similar [84]. A high R^2 value in the training set and a low R^2 value in the test set typically indicates overfitting, whereas a low R^2 value in both sets may suggest a poor learning efficacy in the model. This phenomenon can be detrimental, as it indicates that the model either exhibits a limited capacity for the effective generalization to novel data or fails to accurately capture the underlying patterns in the training data. The application of artificial intelligence to gas separation with membranes is very varied as seen in the introduction. In any case, the R^2 values obtained here are in the range of what can be found in the literature according to the recent study in the subject by Abdollahi et al. [18], R^2 being in the 0.77–0.9999 range. More specifically, Guan et al. [13], studying mixed matrix membranes based on MOF for CO_2 separation using a random forest model (the best model here), with R^2 values in the 0.89–0.91 and 0.70–0.77 ranges for CO_2/CH_4 separation on the training and test datasets, respectively, which corroborates the goodness of our values.

Consequently, the random forest model demonstrates better performance metrics in comparison to the other two alternative models studied here. In fact, Fig. 3 illustrates the comparison between predicted and real CO_2/X selectivity values, revealing a better agreement between them for the random forest model. This is due to the proximity of the points to the red line, which signifies a complete alignment between the values. This result is supported by the capacity of the random forest algorithm to manage both continuous and categorical data, along with its high prediction accuracy [85]. This is because this algorithm is based on ensembles that capture complex and non-linear relationships [86], which is important in the context of membrane preparation, where small variations in variables can interact in a non-linear way affecting the gas separation performance. Conversely, the multiple linear model is characterized by its simplicity, as it does not explain non-linear relationships between input and output variables often leading to prediction errors [87]. Furthermore, the decision tree is affected by overfitting, since the model memorized the training data, but when it encounters unseen test data it lacks the ability to make accurate predictions, as small changes in the data can have a significant effect on the structure of the algorithm [88], resulting in a limited generalization capacity. For this reason, this model performs nearly perfectly in training and poorly in the test set. In contrast, the random forest model generalizes more appropriately to unknown data and combats overfitting by averaging the predictions of several decision trees trained with random subsets, resulting in a more balanced performance [89]. The above can be critical in the preparation of membranes where there are several variables that with small changes influence the membranes properties and therefore the separation results.

Table 2
Pearson correlation coefficients among input variables used to assess multicollinearity and validate their suitability for inclusion in predictive models.

	Solvent	Polym. conc.	Thickness	Solution temp.	Solution time	Cast temp.	Cast time	Cond. temp.	Cond. time	Sep. temp.	Pressure	% CO_2	Gas mixture
Solvent	1												
Polym. conc.	0.16	1											
Thickness	0.07	0.30	1										
Solution temp.	0.68	−0.23	0.04	1									
Solution time	0.62	−0.20	−0.13	0.36	1								
Cast temp.	0.23	0.26	0.16	0.26	−0.14	1							
Cast time	0.07	−0.31	−0.06	0.08	0.08	−0.46	1						
Cond. temp.	0.31	0.08	−0.07	0.20	0.17	0.12	−0.20	1					
Cond. time	0.02	0.03	0.28	−0.12	0.01	−0.22	−0.08	−0.07	1				
Sep. temp.	0.18	0.02	−0.09	0.06	0.13	−0.16	0.27	0.08	−0.05	1			
Pressure	0.14	0.28	0.14	0.13	0.07	0.18	0.14	0.28	−0.01	0.03	1		
% CO_2	0.14	0.06	0.04	0.08	0.07	0.25	−0.11	0.02	−0.14	−0.04	0.09	1	
Gas mixture	0.00	0.15	0.07	0.16	−0.09	0.16	−0.21	0.09	0.04	−0.01	0.01	0.31	1

Table 3
R^2 and RMSE metrics obtained in the training and test sets for the multiple linear regression (MLR), decision tree (DT) and random forest (RF) models.

		MLR	DT	RF
R^2	Train	0.833	0.999	0.973
	Test	0.824	0.655	0.962
RMSE	Train	7.10	0.56	2.85
	Test	6.89	9.12	3.20

Finally, for the purpose of comparison, Fig. S3 shows analogous information to that in Fig. 3 for all the three models (MLR, DT and RF).

Furthermore, Fig. 3 (and similarly the others showing this type of plotting along the article) depicts values of selectivity below ca. 30 corresponding to the CO_2/CH_4, while those above 30 are related to the CO_2/N_2 one, approximately, in agreement with the fact that the type of polymer studied here shows a more selective transport for this last mixture [3], even if, in case of mixtures, working at lower CO_2 concentration values in the feed (15 % vs. 50 %). The above is explained based on the differences in solubility, being greater for methane than for nitrogen, since in terms of diffusion it would be expected that N_2 with a smaller kinetic diameter (0.364 nm) would diffuse faster than methane (0.38 nm) [3,9,90]. In any case, as explained, the selectivity to CO_2 (kinetic diameter 0.33 nm) over both gases is due to its greater solubility and diffusivity in the Pebax® polymer.

As illustrated in Fig. 2, an alternative approach involves extending the database to enhance the model. Therefore, an experiment was conducted in which the results of the experimental membrane tests obtained throughout the present study were incorporated into the entire database. Regarding the training, the model in question exhibited metrics of $R^2 = 0.973$ and RMSE = 2.85, as previously referenced. After the incorporation of the additional data (24 new entries, as shown in Table S1 discussed below), the model demonstrated values of $R^2 = 0.975$ and RMSE = 2.64. These results indicate the consistency of the added experiments, since they do not compromise the performance of the model. Furthermore, it is possible to demonstrate that the extension of the database is a fundamental aspect, since it can allow a decrease in prediction error. Currently, the experimental work done (see Sections 3.4 and 3.5) was mainly designed to the validation of the best membrane preparation conditions aided by the ML model.

Selectivity is the most significant parameter in the CO_2 separation, consequently, the model developed was oriented towards this value. Conversely, permeability constitutes an additional key parameter used for the estimation of selectivity. In this study, an investigation was conducted into a model focused on this parameter. In this instance, for the permeability, the metrics obtained were $R^2 = 0.907$ and RMSE = 12.1 for the training set, and $R^2 = 0.688$ and RMSE = 16.9 for the testing set. The model demonstrates learning capabilities; however, there are cases where it shows difficulty in identifying novel data. This result may be attributable to the necessity for a more substantial dataset that is more representative of a parameter, such as permeability, that can exhibit a wide range of values.

3.3. Importance analysis of input features

Once the random forest model was selected, some detailed analysis was done to gain insight into the application of ML in the prediction of the membrane performance. This analysis gives the degree of the importance of input variables of the random forest model for the estimation of CO_2/X selectivity as a result. Additionally, the sensitivity of input variables increases with the higher absolute value of the relative importance index. Fig. 4 illustrates the degree of importance of each input variable for the estimation of selectivity. The results indicate that the order of importance of input variables for selectivity is as follows: 1. Gas mixture, 2. Separation temperature, 3. Solution temperature, 4. Solvent, 5. CO_2 percentage, 6. Thickness, 7. Pressure, 8. Solution time, 9. Conditioning time, 10. Polymer concentration, 11. Casting time, 12. Conditioning temperature and 13. Casting temperature. Therefore, the variable with the greatest weight corresponds to the gas mixture composition, which is important because gases (in this case, N_2 and CH_4) do not have the same diffusivity and solubility capacity through the membrane. In this case, molecular interactions and kinetic diameters play a key role, thus resulting in different permeability values. Consequently, a specific range for the ratio of gas permeabilities is defined, which in turn defines the separation selectivity. Another significant variable is the measurement or separation temperature, as it affects the permeability with an increase in temperature resulting in a decrease in selectivity due to a decrease in CO_2 solubility. The solution temperature for the preparation of the polymer solution is also important, as it influences the complete dissolution of the polymer, the viscosity of the solution and the structure and morphology of the membrane. The capacity of a given solvent to dissolve the polymer influences the synthesis of membranes and their characteristics. Variables identified with a lower importance are the conditions of the casting step, this is possibly due to the membrane homogeneity and the absence of significant structural changes under sufficiently controlled conditions; conditioning temperature, a step to ensure measurement stability, does not significantly modify the already established membrane characteristics; and polymer concentration may have a significant impact on the processability and thickness of the membrane, but its effect on the gas separation selectivity and structure is not evident. This occurs under controlled conditions and in the absence of defects.

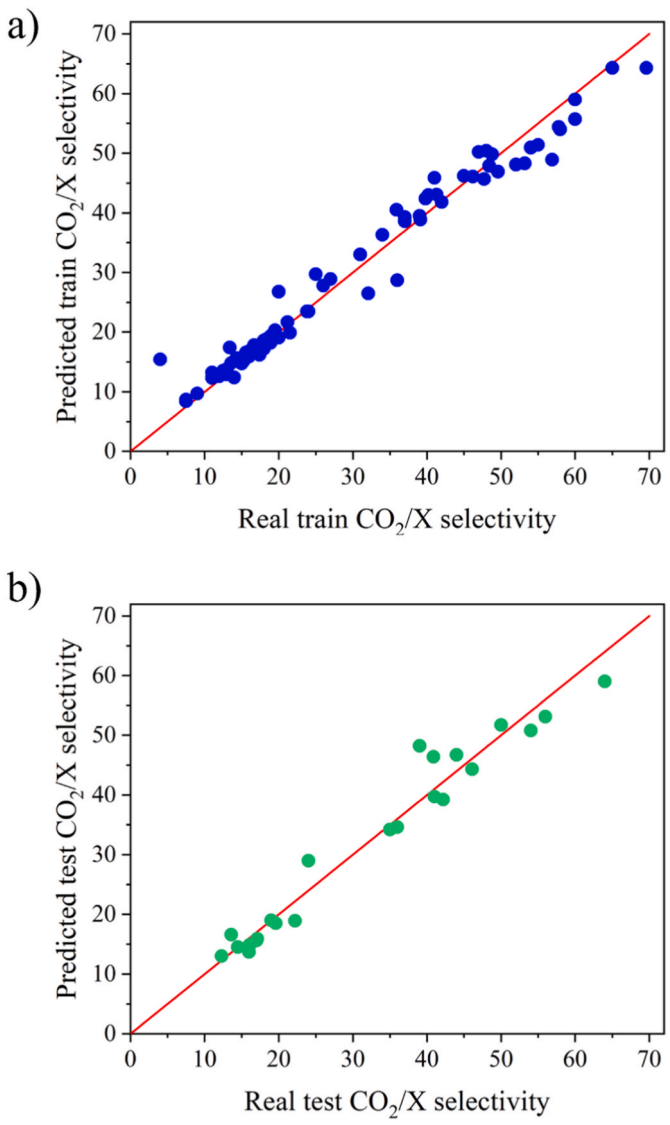

Fig. 3. Comparison between real and predicted values (CO_2/X selectivity) using the random forest model: a) training subset and b) testing subset. The red line in the figure represents the bisector y = x, which indicates the ideal line where the predictions exactly match the real values. The performance of the model is directly proportional to the proximity of the points to the mentioned line. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

3.4. Study and optimization of the procedure for the best conditions

Following the determination of the variables that are more or less important for the model, as well as the establishment of the experimental criteria, only some variables were studied to optimize the process, the study being conducted on a subset of the variables to optimize the process. The remaining variables are fixed at a constant value to simplify the study. Therefore, the variables studied in this section were previously defined in Section 2.5.

As previously discussed, variables such as the polymer concentration, membrane thickness or temperatures involved in the procedure influence the preparation of the membranes and, consequently, the selectivity achieved in gas separation. These variables were studied by a multivariate approach through the design of experiments. For this purpose, a defined range is established for each variable. The minimum, intermediate and maximum values within these ranges are systematically combined, resulting in a set of experimental conditions that

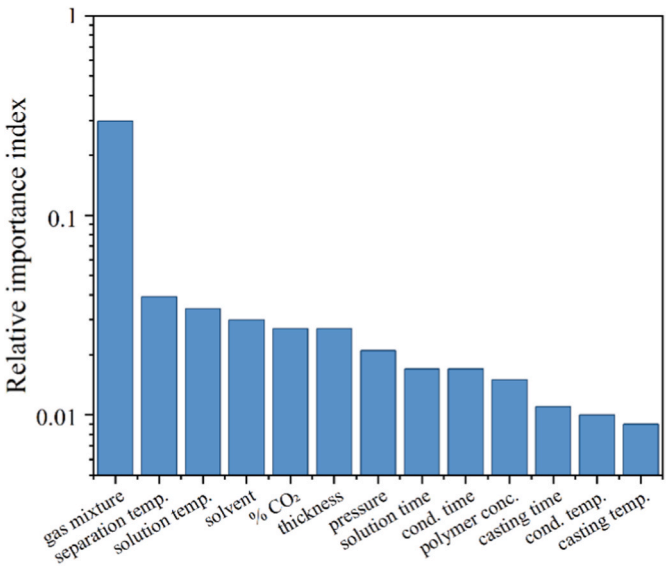

Fig. 4. Ranking for relative importance index of input variables in the random forest model for CO_2/X selectivity prediction. Feature importance was determined by the internal scores of the model, and log-scaling was applied to enhance the visibility of variables with lower relative contributions.

comprehensively cover the entire domain of study. It is important to note that the selectivity values introduced in all the proposed experiments were obtained through the utilization of the optimum machine learning model (i.e., the random forest one) developed in this work.

The analysis of the results was conducted using a response surface methodology, which facilitates the identification of the optimal conditions within the analyzed range. Fig. 5 illustrates the relationship among the three most influential factors studied, revealing that the selectivity increases at low polymer concentrations, with intermediate-to-high membrane thicknesses and low-to-intermediate solution temperatures. Moreover, an evident correlation between these variables and their interdependence across the surface can be discerned. In contrast, combinations involving casting and conditioning temperatures yielded a nearly uniform response, indicating their limited impact on the membrane CO_2/X selectivity. This observation is consistent with the variable importance analysis discussed in Section 3.3, as the order of importance of the variables subjected to the optimization process corresponds to the effects observed in the surface plots. Furthermore, Eq. (4), corresponding to the fitting equation for selectivity, demonstrates the weights of the variables quantitatively through the effects estimated in the design of experiments:

$$\begin{aligned} Selectivity = {} & 41.7 - 2.32A + 3.25B - 8.44C + 0.61D - 0.27E + 0.1AB \\ & - 0.50AC + 0.05AD + 0.25AE - 0.20BC - 0.50BD + 0.60BE \\ & + 0.40CD - 0.35CE + 0.00DE + 2.26A^2 - 3.97B^2 - 5.92C^2 - 0.35D^2 \\ & + 0.29E^2 \end{aligned} \quad (4)$$

where A, B, C, D and E represent polymer concentration (wt%), membrane thickness (μm) and solution, casting and conditioning temperatures (°C), respectively. Moreover, if the coefficients are accompanied by a positive sign, it means that if that factor increases, so will the response, while a negative sign indicates that an increase in the factor has a negative effect on the response. In addition, the order of importance through coefficients is solution temperature > thickness > polymer concentration > casting temperature > conditioning temperature, which is the same order obtained in the importance analysis in the model. Specifically, the solution temperature shows the most significant influence on the response, followed by membrane thickness and

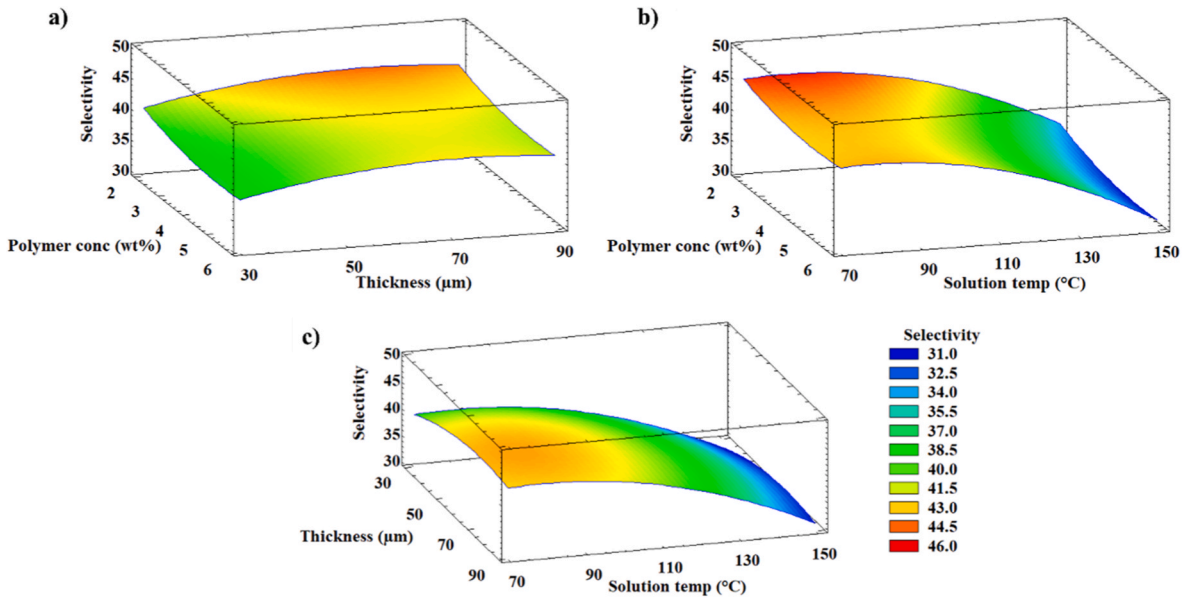

Fig. 5. Response surface plots of CO_2/X selectivity values as a function of the interaction between variables: a) polymer concentration and membrane thickness, b) polymer concentration and solution temperature, and c) membrane thickness and solution temperature. In the plots presented here, the remaining variables not represented in each plot are maintained at their respective intermediate values within the range established during the study. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

polymer concentration, which shows a less variable surface and a reduced influence. Finally, casting and conditioning temperatures were found to be of minimal importance and have no significant effect as shown by their coefficients in the linear terms of Eq. (4), resulting in flat surfaces (not shown). The optimal values derived from the multivariate analysis included a polymer concentration of 2.3 wt%, a membrane thickness of 70 μm, a solution temperature of 80 °C, a casting temperature of 45 °C, and a conditioning temperature of 30 °C. These conditions were selected for the purpose of future studies.

3.5. Experimental measurements and comparison with model estimation

Finally, several membranes were prepared under the optimal conditions suggested by the RSM in the design of experiments and in other experimental conditions, with the aim of verifying the validity of the approach. To verify the importance provided and to determine the capacity of the model to perceive the influence of the change of these factors, some experimental conditions were modified. Table S1 presents the experimental conditions, and a comparison of the experimental values obtained together with the predicted values from the model. Fig. 6a illustrates the training, test and experimental datasets, showing that this last set exhibits a comparable correspondence between the predicted and experimental values, resulting in a favorable predictive performance, with metrics of $R^2 = 0.945$ and RMSE = 3.08. Optimal conditions were also evaluated, revealing that the selectivity values proposed, 45.8 and 44.5 b y both the design of experiments and the model, respectively, are close to the experimentally obtained value of 43.5. As shown in Fig. 6b, a comparison of experiments with a better gas separation performance at 35 °C is provided, indicating that the proposed conditions by the response surface methodology (P1657DM_C6, see Table S1) enable the highest selectivity to be achieved. C5, C7 and C8 are other membranes that were further studied in other conditions. Furthermore, it is possible to observe a close correspondence between the experimental values with the predicted values by the model. This allows us to infer that the ability of the model to predict the dataset values with a normalized error percentage of 10.4 %, is a strong indication of its effectiveness. In addition, the accuracy of the predictions in comparison to the experimental selectivity measurements was determined to be 3.5 % on average, as indicated by the relative standard deviation. Finally, in the case of experimental gas separation, the membranes were measured for approximately 2 h, considering the average of the values of the last 3 measurements performed in each experiment, each of them being conducted in duplicate (i.e., with a different membrane sample). Experimental errors for selectivity, represented as standard deviation, are in the range of 0.2–3.4, where factors such as analytical errors, slight temperature fluctuations, possible membrane instability during experiments and membrane reproducibility itself may contribute to variation between replicates.

3.6. Comparative with other membranes in literature

In this study, analysis and optimization of selectivity were performed within a specific range and under the control of only a subset of variables. Therefore, a comparison with previously published gas separation studies related to the polymer Pebax® 1657 being worked with here is necessary to assess whether a real improvement was achieved. Fig. 7 presents the permeability-selectivity plots for the CO_2/N_2 and CO_2/CH_4 mixtures, constructed using data collected from literature achieved under similar conditions (mainly using the same ethanol/water solvent and similar separation temperatures) and those in the experiments of this work with the optimized membranes P1657DM_C6 (see Table S1). The proposed conditions by RSM, resulted in prepared membranes in the laboratory, with CO_2 permeability and CO_2/X selectivity values of 115 Barrer and 43.5 and 132 Barrer and 16.4 for the CO_2/N_2 and CO_2/CH_4 mixtures, respectively, at 35 °C; and CO_2 permeability and CO_2/X selectivity values of 89 Barrer and 47.8 and 97 Barrer and 18.9 for these gas mixtures, respectively, at 25 °C. Moreover, Table S3 compares these optimized values with some values reported in the literature under similar conditions as mentioned above. To quantify the improvement, the performance of the optimized membranes was ranked among 34 literature-reported data points for the CO_2/N_2 mixture and 39 for the CO_2/CH_4 mixture (shown in Fig. 7), considering both permeability and selectivity simultaneously. The position in plots of each result within these ranked datasets allowed estimating the proportion of literature results that were outperformed. For the CO_2/N_2 and CO_2/CH_4 mixtures, the optimized membraned exceeded approximately 79 % and 92 % of the comparable literature data. On average, the membranes developed in this work outperformed the performance of 85 % of previously

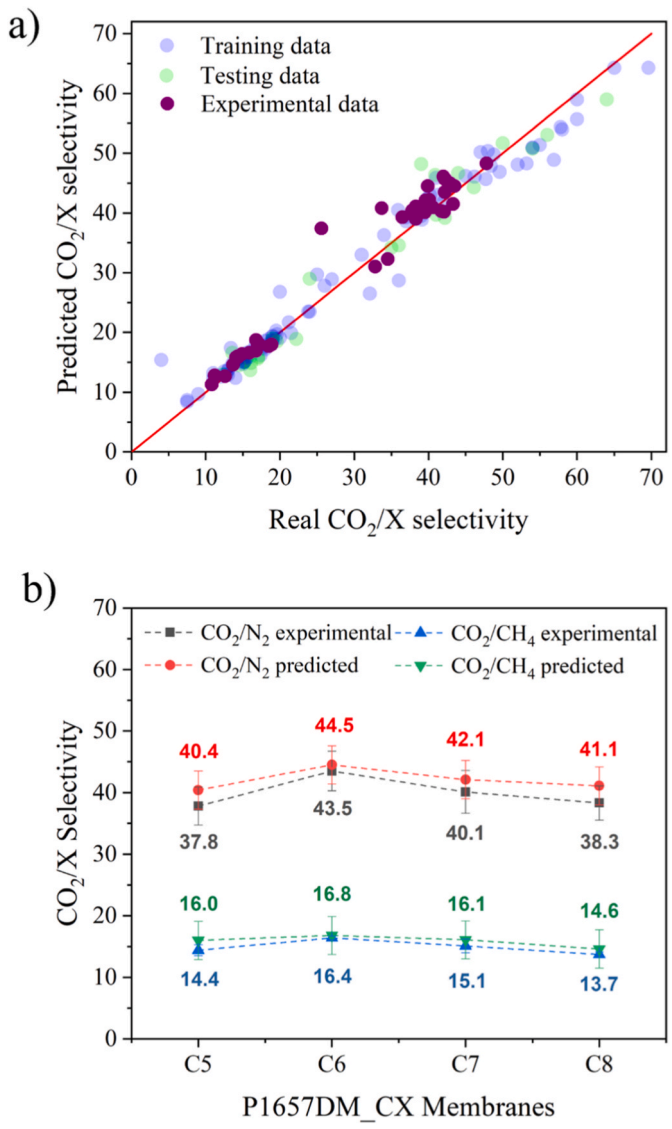

Fig. 6. a) Plot comparing the predicted CO_2/X selectivity versus the real values of selectivity for the training, test and experimental datasets using the developed ML model. b) Comparison of predicted and experimental selectivity values under several conditions for CO_2/N_2 and CO_2/CH_4 gas mixtures. The highest selectivity values were achieved under optimal conditions identified through the response surface methodology. In all cases, the separation temperature was 35 °C. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

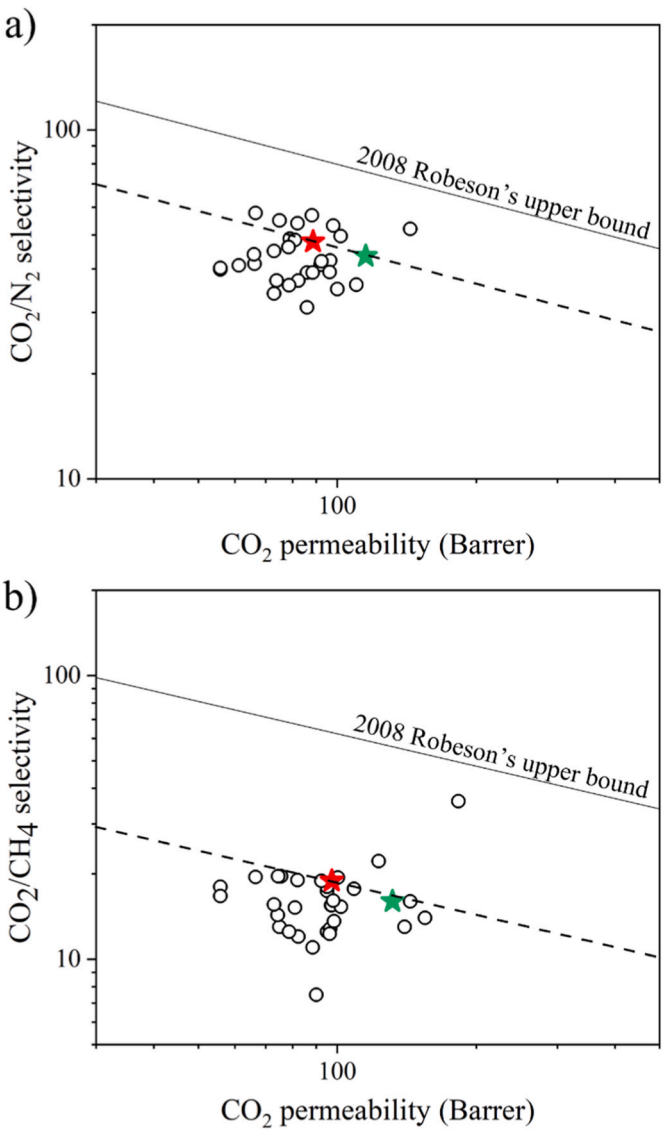

Fig. 7. Selectivities of a) CO_2/N_2 and b) CO_2/CH_4 gas pairs as a function of CO_2 permeability with polymer Pebax® 1657. The improved data from this study (green and red stars at 25 °C and 35 °C, respectively) are compared with literature values (circles) to assess the performance of the prepared dense membranes. The dashed lines are to facilitate the visualization of the data, while the continuous ones correspond to the 2008 Robeson's upper bound [91]. The sample size of the comparison is 73 points (34 and 39 data for CO_2/N_2 and CO_2/CH_4 mixtures, respectively). Some of the values obtained in the literature can be seen in Table S3. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

published results in terms of the permeability-selectivity trade-off (this percentage would be 80 % considering only the selectivity values). As can be observed, this fact confirms that the optimization guided by machine learning and the DOE-RSM methodology led to a significant enhancement in membrane performance.

4. Conclusions

This work describes an approach for the development of a machine learning (ML) model, which can be used to predict membrane separation selectivity values from the preparation conditions of polymer membranes. The purpose of this ML model is to determine the most suitable values of an experiment of membrane preparation for a given gas separation application, before conducting it. The employment of variable analysis tools facilitates the determination of the optimal experimental conditions leading to the improvement of membrane preparation procedures. The method was applied to dense Pebax® 1657 membranes for the gas separation of CO_2/N_2 and CO_2/CH_4 mixtures. Among the three ML models applied (multiple linear regression, decision tree and random forest), the use of the random forest model elucidated the significance of the variables that influence the process, as well as the capacity to predict CO_2/N_2 and CO_2/CH_4 selectivity values from the data collected in the literature and new experimental data more accurately (higher R^2 and lower RMSE values). Moreover, the implementation of a multivariate analysis through the design of experiments with a response surface methodology (DOE-RSM) resulted in the determination of the values of specific experimental variables that enhanced the membrane selectivities obtained. In other words, the integration of the random forest ML model with the RSM suggested a set of experimental conditions for the membrane preparation and successive application to gas

separation mixtures which led to good theoretical-experimental matching and high reproducibility.

Beyond technical accuracy, the integration of machine learning into experimental design also provides clear economic benefits. Conventional trial-and-error methodologies typically employed in the field of membrane development are time-consuming and require of significant expenses and substantial resource requirements. Conversely, data-driven models allow researchers to focus experimental efforts on the most promising conditions, simplifying the experimental approach by reducing both the number of required tests and associated costs. As the pool of known membranes expands, the cost of discovering new, high-performing materials through purely empirical methods increases significantly. Consequently, approaches like the one proposed in this study not only boost prediction and performance but also enhance the efficiency and sustainability of membrane research.

In addition, from a future-oriented standpoint, it is important to emphasize on the accuracy of the predictions, as this will enhance the effectiveness of the ML methodology. Consequently, given the nature of this data-driven tool, key aspects that should be further explored include data acquisition and quality assurance. This includes, for instance, the application of more robust statistical methods for outlier detection, which could be more effectively implemented with a more extensive database. In the future, it would be ideal to conduct studies with a larger dataset to improve the representativeness and reliability of the models, along with the exploration of different algorithms and the optimization of the overall process. An essential approach, which was also implemented in this study, is the validation of the model through experimental verification.

CRediT authorship contribution statement

Lucía Carrillo-Sánchez: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation. **Carlos Téllez:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Joaquín Coronas:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization.

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.

Acknowledgements

Grant PID2022-138582OB-I00 funded by MICIU/AEI/10.13039/501100011033/and by "ERDF A way of making Europe" is gratefully acknowledged. Also, grant T68_23R funded by Aragón Government, IASOMM24004 funded by European Union-NextGenerationEU/PRTR (C16.R1) and grant CEX2023-001286-S funded by MICIU/AEI/10.13039/501100011033/are thanked. In addition, "Investigo Program" of "Servicio Público de Empleo Estatal" which is funded by European Union-NextGenerationEU/PRTR is gratefully acknowledged. We thank Dr. Jesús Asín, associate professor from University of Zaragoza, for the fruitful discussion on data treatment. We thank Dr. Javier Lacasta, associate professor from University of Zaragoza, for the constructive discussion on machine learning applications.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.memsci.2025.124708.

Data availability

Data will be made available on request.

References

[1] A. Car, C. Stropnik, W. Yave, K. Peinemann, PEG modified poly(amide-b-ethylene oxide) membranes for CO_2 separation, J. Membr. Sci. 307 (1) (2008) 88–95, https://doi.org/10.1016/j.memsci.2007.09.023.

[2] A.A. Shamsabadi, F. Seidi, E. Salehi, M. Nozari, A. Rahimpour, M. Soroush, Efficient CO_2-removal using novel mixed-matrix membranes with modified TiO_2 nanoparticles, J. Mater. Chem. A 5 (8) (2017) 4011–4025, https://doi.org/10.1039/c6ta09990d.

[3] A. Embaye, L. Martínez-Izquierdo, M. Malankowska, C. Téllez, J. Coronas, Poly (ether-block-amide) copolymer membranes in CO_2 separation applications, Energy Fuel. 35 (21) (2021) 17085–17102, https://doi.org/10.1021/acs.energyfuels.1c01638.

[4] F. Karamouz, H. Maghsoudi, R. Yegani, Synthesis and characterization of high permeable PEBA membranes for CO_2/CH_4 separation, J. Nat. Gas Sci. Eng. 35 (2016) 980–985, https://doi.org/10.1016/j.jngse.2016.09.036.

[5] S.R. Reijerkerk, M.H. Knoef, K. Nijmeijer, M. Wessling, Poly(ethylene glycol) and poly(dimethyl siloxane): combining their advantages into efficient CO_2 gas separation membranes, J. Membr. Sci. 352 (1-2) (2010) 126–135, https://doi.org/10.1016/j.memsci.2010.02.008.

[6] G. Chatterjee, A. Houde, S. Stern, Poly(ether urethane) and poly(ether urethane urea) membranes with high H_2S/CH_4 selectivity, J. Membr. Sci. 135 (1) (1997) 99–106, https://doi.org/10.1016/S0376-7388(97)00134-8.

[7] R. Zoppi, S. das Neves, S. Nunes, Hybrid films of poly(ethylene oxide-b-amide-6) containing sol-gel silicon or titanium oxide as inorganic fillers: effect of morphology and mechanical properties on gas permeability, Polymer 41 (14) (2000) 5461–5470, https://doi.org/10.1016/S0032-3861(99)00751-X.

[8] X.L. Ren, J.Z. Ren, H. Li, S.C. Feng, M.C. Deng, Poly (amide-6-b-ethylene oxide) multilayer composite membrane for carbon dioxide separation, Int. J. Greenh. Gas Control 8 (2012) 111–120, https://doi.org/10.1016/j.ijggc.2012.01.017.

[9] L. Martínez-Izquierdo, A. Perea-Cachero, M. Malankowska, C. Téllez, J. Coronas, A comparative study between single gas and mixed gas permeation of polyether-block-amide type copolymer membranes, J. Environ. Chem. Eng. 10 (5) (2022), https://doi.org/10.1016/j.jece.2022.108324.

[10] D. Zhao, J. Ren, Y. Qiu, H. Li, K. Hua, X. Li, M. Deng, Effect of graphene oxide on the behavior of poly(amide-6-b-ethylene oxide)/graphene oxide mixed-matrix membranes in the permeation process, J. Appl. Polym. Sci. 132 (41) (2015), https://doi.org/10.1002/app.42624.

[11] G. Guillen, Y. Pan, M. Li, E. Hoek, Preparation and characterization of membranes formed by nonsolvent induced phase separation: a review, Ind. Eng. Chem. Res. 50 (7) (2011) 3798–3817, https://doi.org/10.1021/ie101928r.

[12] K. Brown, S. Brittman, N. Maccaferri, D. Jariwala, U. Ceano, Machine learning in nanoscience: big data at small scales, Nano Lett. 20 (1) (2020) 2–10, https://doi.org/10.1021/acs.nanolett.9b04090.

[13] J. Guan, T. Huang, W. Liu, F. Feng, S. Japip, J. Li, X. Wang, S. Zhang, Design and prediction of metal organic framework-based mixed matrix membranes for CO_2 capture via machine learning, Cell Rep. Phys. Sci. 3 (5) (2022), https://doi.org/10.1016/j.xcrp.2022.100864.

[14] B. Li, R. Yue, L. Shen, C. Chen, R. Li, Y. Xu, M. Zhang, H. Hong, H. Lin, A novel method integrating response surface method with artificial neural network to optimize membrane fabrication for wastewater treatment, J. Clean. Prod. 376 (2022), https://doi.org/10.1016/j.jclepro.2022.134236.

[15] L. Chen, G. Liu, Z. Zhang, Y. Wang, Y. Yang, J. Li, Machine learning and molecular design algorithm assisted discovery of gas separation membranes exceeding the CO_2/CH_4 and CO_2/N_2 upper bounds, Chem. Eng. Sci. 291 (2024), https://doi.org/10.1016/j.ces.2024.119952.

[16] H. Daglar, S. Keskin, Combining machine learning and molecular simulations to unlock gas separation potentials of MOF membranes and MOF/polymer MMMs, ACS Appl. Mater. Interfaces 14 (28) (2022) 32134–32148, https://doi.org/10.1021/acsami.2c08977.

[17] A. Li, J. Chu, S. Huang, Y. Liu, M. He, X. Liu, Machine learning-assisted development of gas separation membranes: a review, Carbon Capture Sci. Technol. 14 (2025), https://doi.org/10.1016/j.ccst.2025.100374.

[18] F. Abdollahi, A. Khosravi, S. Karagöz, A. Keshavarz, A systematic review of recent advances in the application of machine learning in membrane-based gas separation technologies, Appl. Energy 381 (2025), https://doi.org/10.1016/j.apenergy.2024.125203.

[19] J. Wang, K. Tian, D. Li, M. Chen, X. Feng, Y. Zhang, Y. Wang, B. van der Bruggen, Machine learning in gas separation membrane developing: ready for prime time, Sep. Purif. Technol. 313 (2023), https://doi.org/10.1016/j.seppur.2023.123493.

[20] A. Tayyebi, A. Alshami, X. Yu, E. Kolodka, Can machine learning methods guide gas separation membranes fabrication? J. Membr. Sci. Lett. 2 (2) (2022) https://doi.org/10.1016/j.memlet.2022.100033.

[21] J. Barnett, C. Bilchak, Y. Wang, B. Benicewicz, L. Murdock, T. Bereau, S. Kumar, Designing exceptional gas-separation polymer membranes using machine learning, Sci. Adv. 6 (20) (2020), https://doi.org/10.1126/sciadv.aaz4301.

[22] M. Zhao, C. Zhang, Y. Weng, Improved artificial neural networks (ANNs) for predicting the gas separation performance of polyimides, J. Membr. Sci. 681 (2023), https://doi.org/10.1016/j.memsci.2023.121765.

[23] Q. Yuan, M. Longo, A. Thornton, N. McKeown, B. Comesaña-Gándara, J. Jansen, K. Jelfs, Imputation of missing gas permeability data for polymer membranes using machine learning, J. Membr. Sci. 627 (2021), https://doi.org/10.1016/j.memsci.2021.119207.

[24] Z. Zhang, X. Cao, C. Geng, Y. Sun, Y. He, Z. Qiao, C. Zhong, Machine learning aided high-throughput prediction of ionic liquid@MOF composites for membrane-based CO_2 capture, J. Membr. Sci. 650 (2022), https://doi.org/10.1016/j.memsci.2022.120399.

[25] L. Martínez-Izquierdo, M. Malankowska, J. Sánchez-Laínez, C. Téllez, J. Coronas, Poly(ether-block-amide) copolymer membrane for CO_2/N_2 separation: the influence of the casting solution concentration on its morphology, thermal properties and gas separation performance, R. Soc. Open Sci. 6 (9) (2019), https://doi.org/10.1098/rsos.190866.

[26] M. Isanejad, N. Azizi, T. Mohammadi, Pebax membrane for CO_2/CH_4 separation: effects of various solvents on morphology and performance, J. Appl. Polym. Sci. 134 (9) (2017), https://doi.org/10.1002/app.44531.

[27] E. Ahmadpour, A. Shamsabadi, R. Behbahani, M. Aghajani, A. Kargari, Study of CO_2 separation with PVC/pebax composite membrane, J. Nat. Gas Sci. Eng. 21 (2014) 518–523, https://doi.org/10.1016/j.jngse.2014.09.021.

[28] Q. Zhang, M. Zhou, X. Liu, B. Zhang, Pebax/two-dimensional MFI nanosheets mixed-matrix membranes for enhanced CO_2 separation, J. Membr. Sci. 636 (2021), https://doi.org/10.1016/j.memsci.2021.119612.

[29] R. Ding, W. Zheng, K. Yang, Y. Dai, X. Ruan, X. Yan, G. He, Amino-functional ZIF-8 nanocrystals by microemulsion based mixed linker strategy and the enhanced CO_2/N_2 separation, Sep. Purif. Technol. 236 (2020), https://doi.org/10.1016/j.seppur.2019.116209.

[30] X. Zhao, W. Liu, X. Liu, B. Zhang, Mixed matrix membranes incorporated with aminosilane-functionalized SAPO-34 for upgrading CO_2/CH_4 separation performances, Ind. Eng. Chem. Res. 60 (38) (2021) 13927–13937, https://doi.org/10.1021/acs.iecr.1c01522.

[31] R. Ebadi, H. Maghsoudi, A. Babaluo, Fabrication and characterization of Pebax-1657 mixed matrix membrane loaded with Si-CHA zeolite for CO_2 separation from CH_4, J. Nat. Gas Sci. Eng. 90 (2021), https://doi.org/10.1016/j.jngse.2021.103947.

[32] L. Li, J. Wang, H. Li, X. Guan, K. Li, B. Zhang, X. Li, J. Zhang, Constructing multi-dimensional transport pathways by mixed-dimensional fillers in membranes for efficient CO_2 separation, J. Environ. Chem. Eng. 11 (1) (2023), https://doi.org/10.1016/j.jece.2022.109178.

[33] Y. Wang, N. Zhang, H. Wu, Y. Ren, L. Yang, X. Wang, Y. Wu, Y. Liu, R. Zhao, Z. Jiang, Exfoliation-free layered double hydroxides laminates intercalated with amino acids for enhanced CO_2 separation of mixed matrix membrane, J. Membr. Sci. 618 (2021), https://doi.org/10.1016/j.memsci.2020.118691.

[34] F. Guo, B. Li, R. Ding, D. Li, X. Jiang, G. He, W. Xiao, A novel composite material UiO-66@HNT/Pebax mixed matrix membranes for enhanced CO_2/N_2 separation, Membranes 11 (9) (2021), https://doi.org/10.3390/membranes11090693.

[35] C. Wu, H. Guo, X. Liu, B. Zhang, Mixed matrix membrane comprising glycine grafted CuBTC for enhanced CO_2 separation performances with excellent stability under humid atmosphere, Sep. Purif. Technol. 295 (2022), https://doi.org/10.1016/j.seppur.2022.121287.

[36] Q. Xin, W. Shao, Q. Ma, X. Ye, Z. Huang, B. Li, S. Wang, H. Li, Y. Zhang, Efficient CO_2 separation of multi-permselective mixed matrix membranes with a unique interfacial structure regulated by mesoporous nanosheets, ACS Appl. Mater. Interfaces 12 (42) (2020) 48067–48076, https://doi.org/10.1021/acsami.0c10895.

[37] Q. Zhang, H. Guo, G. Muradi, B. Zhang, Tuning the multi-scale structure of mixed-matrix membranes for upgrading CO_2 separation performances, Sep. Purif. Technol. 293 (2022), https://doi.org/10.1016/j.seppur.2022.121118.

[38] T. Khosravi, M. Omidkhah, S. Kaliaguine, D. Rodrigue, Amine-functionalized CUBTC/poly(ether-b-amide-6) (Pebax® MH 1657) mixed matrix membranes for CO_2/CH_4 separation, Can. J. Chem. Eng. 95 (10) (2017) 2024–2033, https://doi.org/10.1002/cjce.22857.

[39] N. Tara, Z. Shamair, N. Habib, M. Craven, M. Bilad, M. Usman, X. Tu, A. Khan, Simultaneous increase in CO_2 permeability and selectivity by BIT-72 and modified BIT-72 based mixed matrix membranes, Chem. Eng. Res. Des. 178 (2022) 136–147, https://doi.org/10.1016/j.cherd.2021.12.007.

[40] X. Jiang, C. Chuah, K. Goh, R. Wang, A facile direct spray-coating of pebax® 1657: towards large-scale thin-film composite membranes for efficient CO_2/N_2 separation, J. Membr. Sci. 638 (2021), https://doi.org/10.1016/j.memsci.2021.119708.

[41] K. Shahrezaei, R. Abedini, M. Lashkarbolooki, A. Rahimpour, A preferential CO_2 separation using binary phases membrane consisting of Pebax®1657 and [Omim][PF_6] ionic liquid, Kor. J. Chem. Eng. 36 (12) (2019) 2085–2094, https://doi.org/10.1007/s11814-019-0402-z.

[42] D. Nobakht, R. Abedini, Improved gas separation performance of Pebax®1657 membrane modified by poly-alcoholic compounds, J. Environ. Chem. Eng. 10 (3) (2022), https://doi.org/10.1016/j.jece.2022.107568.

[43] T. Hassan, N. Jusoh, Y. Yeong, Synthesis and characterization of PEBAX 1657 and hierarchical linde Type-T (h-LTT) zeolite for the fabrication of hybrid membranes, Mater. Today Proc. (2021) 1263–1268, https://doi.org/10.1016/j.matpr.2021.02.792.

[44] P. Bernardo, G. Clarizia, Enhancing gas permeation properties of pebax® 1657 membranes via polysorbate nonionic surfactants doping, Polymers 12 (2) (2020), https://doi.org/10.3390/polym12020253.

[45] R. Habibi, O. Bakhtiari, Preparation and characterization of modified halloysite nanotubes-Pebax nanocomposite membranes for CO_2/CH_4 separation, Chem. Eng. Res. Des. 174 (2021) 199–212, https://doi.org/10.1016/j.cherd.2021.08.010.

[46] N. Soltani, O. Bakhtiari, Preparation of alumina nanotubes for incorporation into CO_2 permselective Pebax-based nanocomposite membranes, Kor. J. Chem. Eng. 38 (7) (2021) 1469–1486, https://doi.org/10.1007/s11814-021-0777-5.

[47] L. Xu, L. Xiang, C. Wang, J. Yu, L. Zhang, Y. Pan, Enhanced permeation performance of polyether-polyamide block copolymer membranes through incorporating ZIF-8 nanocrystals, Chin. J. Chem. Eng. 25 (7) (2017) 882–891, https://doi.org/10.1016/j.cjche.2016.11.007.

[48] L. Fan, Z. Kang, Y. Shen, S. Wang, H. Zhao, H. Sun, X. Hu, H. Sun, R. Wang, D. Sun, Mixed matrix membranes based on metal-organic frameworks with tunable pore size for CO_2 separation, Cryst. Growth Des. 18 (8) (2018) 4365–4371, https://doi.org/10.1021/acs.cgd.8b00307.

[49] C. Li, C. Wu, B. Zhang, Enhanced CO_2/CH_4 separation performances of mixed matrix membranes incorporated with two-dimensional Ni-Based MOF nanosheets, ACS Sustainable Chem. Eng. 8 (1) (2020) 642–648, https://doi.org/10.1021/acssuschemeng.9b06370.

[50] D. Wang, D. Yao, Y. Wang, F. Wang, Y. Xin, S. Song, Z. Zhang, F. Su, Y. Zheng, Carbon nanotubes and graphene oxide-based solvent-free hybrid nanofluids functionalized mixed-matrix membranes for efficient CO_2/N_2 separation, Sep. Purif. Technol. 221 (2019) 421–432, https://doi.org/10.1016/j.seppur.2019.04.005.

[51] M. Gou, W. Zhu, Y. Sun, R. Guo, Introducing two-dimensional metal-organic frameworks with axial coordination anion into Pebax for CO_2/CH_4 separation, Sep. Purif. Technol. 259 (2021), https://doi.org/10.1016/j.seppur.2020.118107.

[52] D. Wang, Y. Zheng, D. Yao, Z. Yang, Y. Xin, F. Wang, Y. Wang, H. Ning, H. Wu, H. Wang, Liquid-like CNT/SiO_2 nanoparticle organic hybrid materials as fillers in mixed matrix composite membranes for enhanced CO_2-selective separation, New J. Chem. 43 (30) (2019) 11949–11958, https://doi.org/10.1039/c9nj02789k.

[53] J. Zhang, Q. Xin, X. Li, M. Yun, R. Xu, S. Wang, Y. Li, L. Lin, X. Ding, H. Ye, Y. Zhang, Mixed matrix membranes comprising aminosilane-functionalized graphene oxide for enhanced CO_2 separation, J. Membr. Sci. 570 (2019) 343–354, https://doi.org/10.1016/j.memsci.2018.10.075.

[54] S. Lian, R. Li, Z. Zhang, Q. Liu, C. Song, S. Lu, Improved CO_2 separation performance and interfacial affinity of composite membranes by incorporating amino acid-based deep eutectic solvents, sep purif, Tech 272 (2021), https://doi.org/10.1016/j.seppur.2021.118953.

[55] G. Huang, A. Isfahani, A. Muchtar, K. Sakurai, B. Shrestha, D. Qin, D. Yamaguch, E. Sivaniah, B. Ghalei, Pebax/ionic liquid modified graphene oxide mixed matrix membranes for enhanced CO_2 capture, J. Membr. Sci. 565 (2018) 370–379, https://doi.org/10.1016/j.memsci.2018.08.026.

[56] H. Yang, S. Liang, P. Zhang, X. Zhang, P. Lu, Y. Liu, X. Cao, Y. Li, Q. Wang, Improved CO_2 separation performance of mixed matrix membranes via expanded layer double hydroxides and methanol post-treatment, J. Membr. Sci. 670 (2023), https://doi.org/10.1016/j.memsci.2023.121345.

[57] M. Longo, M. De Santo, E. Esposito, A. Fuoco, M. Monteleone, L. Giorno, J. Jansen, Force spectroscopy determination of Young's modulus in mixed matrix membranes, Polymer 156 (2018) 22–29, https://doi.org/10.1016/j.polymer.2018.09.043.

[58] H. Wu, X. Li, Y. Li, S. Wang, R. Guo, Z. Jiang, C. Wu, Q. Xin, X. Lu, Facilitated transport mixed matrix membranes incorporated with amine functionalized MCM-41 for enhanced gas separation properties, J. Membr. Sci. 465 (2014) 78–90, https://doi.org/10.1016/j.memsci.2014.04.023.

[59] P. Sharma, Y. Kim, M. Kim, S. Alam, C. Cho, A stable polymeric chain configuration producing high performance PEBAX-1657 membranes for CO_2 separation, Nanoscale Adv. 1 (7) (2019) 2633–2644, https://doi.org/10.1039/c9na00170k.

[60] Y. Li, X. Li, H. Wu, Q. Xin, S. Wang, Y. Liu, Z. Tian, T. Zhou, Z. Jiang, H. Tian, X. Cao, B. Wang, Anionic surfactant-doped Pebax membrane with optimal free volume characteristics for efficient CO_2 separation, J. Membr. Sci. 493 (2015) 460–469, https://doi.org/10.1016/j.memsci.2015.06.046.

[61] N. Habib, Z. Shamair, N. Tara, A. Nizami, F. Akhtar, N. Ahmad, M. Gilani, M. Bilad, A. Khan, Development of highly permeable and selective mixed matrix membranes based on Pebax®1657 and NOTT-300 for CO_2 capture, Sep. Purif. Technol. 234 (2020), https://doi.org/10.1016/j.seppur.2019.116101.

[62] X. Li, Z. Jiang, Y. Wu, H. Zhang, Y. Cheng, R. Guo, H. Wu, High-performance composite membranes incorporated with carboxylic acid nanogels for CO_2 separation, J. Membr. Sci. 495 (2015) 72–80, https://doi.org/10.1016/j.memsci.2015.07.065.

[63] V. Mozaffari, M. Sadeghi, A. Fakhar, G. Khanbabaei, A. Ismail, Gas separation properties of polyurethane/poly(ether-block-amide) (PU/PEBA) blend membranes, Sep. Purif. Technol. 185 (2017) 202–214, https://doi.org/10.1016/j.seppur.2017.05.028.

[64] E. Estahbanati, M. Omidkhah, A. Arnooghint, Interfacial design of ternary mixed matrix membranes containing pebax 1657/Silver-Nanopowder/[BMIM][BF_4] for improved CO_2 separation performance, ACS Appl. Mater. Interfaces 9 (11) (2017) 10094–10105, https://doi.org/10.1021/acsami.6b16539.

[65] H. Wang, Y. Ding, M. Ning, M. Yu, W. Zheng, X. Ruan, Y. Xi, Y. Dai, H. Liu, G. He, Amino-functional CPL-1 with abundant CO_2-philic groups to enhance MMM-based CO_2 separation, Sep. Purif. Technol. 322 (2023), https://doi.org/10.1016/j.seppur.2023.124227.

[66] R. Ding, Q. Wang, X. Ruan, Y. Dai, X. Li, W. Zheng, G. He, Novel and versatile PEI modified ZIF-8 hollow nanotubes to construct CO_2 facilitated transport pathway in MMMs, Sep. Purif. Technol. 289 (2022), https://doi.org/10.1016/j.seppur.2022.120768.

[67] A. Nadeali, M. Pedram, M. Omidkhah, M. Yarmohammadi, Promising performance for efficient CO_2 separation with the p-tert-Butylcalix[4]arene/Pebax-1657 mixed matrix membrane, ACS Sustainable Chem. Eng. 7 (23) (2019) 19015–19026, https://doi.org/10.1021/acssuschemeng.9b04641.
[68] S. Miri, M. Omidkhah, A. Amooghin, T. Matsuura, Membrane-based gas separation accelerated by quaternary mixed matrix membranes, J. Nat. Gas Sci. Eng. 84 (2020), https://doi.org/10.1016/j.jngse.2020.103655.
[69] A. Nadeali, S. Kalantari, M. Yarmohammadi, M. Omidkhah, A. Amooghin, M. Pedram, CO_2 separation properties of a ternary mixed-matrix membrane using ultraselective synthesized macrocyclic organic compounds, ACS Sustainable Chem. Eng. 8 (34) (2020) 12775–12787, https://doi.org/10.1021/acssuschemeng.0c01895.
[70] N. Azizi, M. Isanejad, T. Mohammadi, R. Behbahani, Effect of TiO_2 loading on the morphology and CO_2/CH_4 separation performance of PEBAX-based membranes, Front. Chem. Sci. Eng. 13 (3) (2019) 517–530, https://doi.org/10.1007/s11705-018-1781-0.
[71] Z. Farashi, S. Azizi, M. Arzhandi, Z. Noroozi, N. Azizi, Improving CO_2/CH_4 separation efficiency of Pebax-1657 membrane by adding Al_2O_3 nanoparticles in its matrix, J. Nat. Gas Sci. Eng. 72 (2019), https://doi.org/10.1016/j.jngse.2019.103019.
[72] N. Azizi, M. Arzani, H. Mahdavi, T. Mohammadi, Synthesis and characterization of poly(ether-block-amide) copolymers/multi-walled carbon nanotube nanocomposite membranes for CO_2/CH_4 separation, Kor. J. Chem. Eng. 34 (9) (2017) 2459–2470, https://doi.org/10.1007/s11814-017-0152-8.
[73] S. Ahmadi, T. Mohammadi, N. Azizi, Effect of halloysite nanotubes incorporation on morphology and CO_2/CH_4 separation performance of Pebax-based membranes, Kor. J. Chem. Eng. 38 (1) (2021) 104–113, https://doi.org/10.1007/s11814-020-0650-y.
[74] N. Azizi, H. Mahdavi, M. Isanejad, T. Mohammadi, Effects of low and high molecular mass PEG incorporation into different types of poly(ether-b-amide) copolymers on the permeation properties of CO_2 and CH_4, J. Polym. Res. 24 (9) (2017), https://doi.org/10.1007/s10965-017-1297-1.
[75] E. Estahbanati, M. Omidkh, A. Amooghin, Preparation and characterization of novel Ionic liquid/Pebax membranes for efficient CO_2/light gases separation, J. Ind. Eng. Chem. 51 (2017) 77–89, https://doi.org/10.1016/j.jiec.2017.02.017.
[76] S. Ahmadi, T. Mohammadi, N. Azizi, Superior Pebax-1657/amine-modified halloysite nanotubes mixed-matrix membranes to improve the CO_2/CH_4 separation efficiency, J. Appl. Polym. Sci. 138 (31) (2021), https://doi.org/10.1002/app.50749.
[77] X. Yang, W. Zheng, Y. Xi, W. Guan, X. Yan, X. Ruan, C. Ma, Y. Dai, G. He, Constructing low-resistance and high-selectivity transport multi-channels in mixed matrix membranes for efficient CO_2 separation, J. Membr. Sci. 624 (2021), https://doi.org/10.1016/j.memsci.2020.119046.
[78] H. Yin, M. Xu, Z. Luo, X. Bi, J. Li, S. Zhang, X. Wang, Machine learning for membrane design and discovery, Green Energy Environ. 9 (1) (2024) 54–70, https://doi.org/10.1016/j.gee.2022.12.001.
[79] P. Dansawad, Y. Li, Y. Li, J. Zhang, S. You, W. Li, S. Yi, Machine learning toward improving the performance of membrane-based wastewater treatment: a review, Advanced Membranes 3 (2023), https://doi.org/10.1016/j.advmem.2023.100072.
[80] D. Chicco, M. Warrens, G. Jurman, The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation, PeerJ Comput. Sci. 623 (2021), https://doi.org/10.7717/peerj-cs.623.
[81] K. Rao, G. Saikrishna, K. Supriya, Data preprocessing techniques: emergence and selection towards machine learning models-a practical review using HPA dataset, Multimed. Tool. Appl. 82 (2023) 37177–37196, https://doi.org/10.1007/s11042-023-15087-5.
[82] M. Yolmeh, S. Jafari, Applications of response surface methodology in the food industry processes, Food Bioprocess Technol. 10 (3) (2017) 413–433, https://doi.org/10.1007/s11947-016-1855-2.
[83] L. Pereira, T. Milan, D. Tapia-Blácido, Using response surface methodology (RSM) to optimize 2G bioethanol production: a review, Biomass Bioenergy 151 (2021), https://doi.org/10.1016/j.biombioe.2021.106166.
[84] M. Wang, G. Hui, Y. Pang, S. Wang, S. Chen, Optimization of machine learning approaches for shale gas production forecast, Geoenergy Sci. Eng. 226 (2023), https://doi.org/10.1016/j.geoen.2023.211719.
[85] J. Magidi, L. Nhamo, S. Mpandeli, T. Mabhaudhi, Application of the random forest classifier to map irrigated areas using google Earth engine, Remote Sens. 13 (5) (2021), https://doi.org/10.3390/rs13050876.
[86] L. Breiman, Random forests, Mach. Learn. 45 (1) (2001) 5–32, https://doi.org/10.1023/A:1010933404324.
[87] A. Hosseinzadeh, M. Baziar, H. Alidadi, J. Zhou, A. Altaee, A. Najafpoor, S. Jafarpour, Application of artificial neural network and multiple linear regression in modeling nutrient recovery in vermicompost under different conditions, Bioresour. Technol. 303 (2020), https://doi.org/10.1016/j.biortech.2020.122926.
[88] H. Zhou, J. Zhang, Y. Zhou, X. Guo, Y. Ma, A feature selection algorithm of decision tree based on feature weight, Expert Syst. Appl. 164 (2021), https://doi.org/10.1016/j.eswa.2020.113842.
[89] T. Dietterich, Ensemble methods in machine learning, Mult. Classifier Syst. 1857 (2000) 1–15, https://doi.org/10.1007/3-540-45014-9_1.
[90] H. Ferrari, F. Bernard, L. dos Santos, G. Dias, C. Le Roux, P. Micoud, F. Martin, S. Einloft, Enhancing CO_2/N_2 and CO_2/CH_4 separation in mixed matrix membrane: a comprehensive study on Pebax®1657 with SSMMP/IL for improved efficiency, Polym. Eng. Sci. 64 (6) (2024) 2875–2893, https://doi.org/10.1002/pen.26732.
[91] L. Robeson, The upper bound revisited, J. Membr. Sci. 320 (1-2) (2008) 390–400, https://doi.org/10.1016/j.memsci.2008.04.030.