

Gaussian process surrogate modelling and multi-objective optimization for energy-efficient CO₂ capture using functionalized oil palm ash

Fakhrony Sholahudin Rohman · Muhammad Yusri Ahmad · Dinie Muhammad ·
Dipesh Shikchand Patle · Ashraf Azmi · Syamima Nasrin Mohamed Saleh 

Received: 4 July 2026 / Accepted: 3 September 2026
© The Author(s) 2026

Abstract

Agricultural waste-derived adsorbents offer a sustainable approach for carbon capture while supporting biomass valorization. However, identifying operating conditions that simultaneously maximize CO₂ adsorption and minimize energy consumption remains challenging because adsorption performance depends on multiple interacting variables. This study presents a machine learning-assisted optimization framework for CO₂ capture using KOH-functionalized oil palm ash (OPA). A dataset comprising 346 fixed-bed adsorption experiments was used to develop Gaussian Process Regression (GPR), Support Vector Regression (SVR), and Regression Tree models, employing the OPA-to-KOH ratio, gas flow rate, adsorption temperature, inlet CO₂ concentration, and contact time as input variables. Among the models evaluated, GPR achieved the highest predictive performance, with an R^2 of 0.9984, a root mean square error of 0.0262 mmol g⁻¹, and a mean absolute error of 0.0119 mmol g⁻¹. The validated GPR model was integrated with Genetic Algorithm (GA) and Particle Swarm Optimization (PSO) to maximize adsorption capacity, and with Multi-Objective Particle Swarm Optimization (MOPSO) and Non-dominated Sorting Genetic Algorithm II (NSGA-II) to optimize both adsorption capacity and electrical energy consumption. Both GA and PSO predicted a maximum adsorption capacity of 2.954 mmol g⁻¹, while the multi-objective algorithms identified Pareto-optimal operating conditions that balanced adsorption performance with energy demand. The proposed framework demonstrates the potential of integrating machine learning with multi-objective optimization to support the design of energy-efficient and sustainable CO₂ adsorption processes.

Keywords CO₂ adsorption · KOH-functionalized OPA · Machine learning · Gaussian process regression · Genetic algorithm · Multi-objective optimization · Particle swarm optimization

Abbreviations

AC	Adsorption capacity
ANN	Artificial neural network
BET	Brunauer–Emmett–Teller
BJH	Barrett–Joyner–Halenda

This Article in Press is shared early to give you faster access to new research. It is citable and carries a permanent DOI. The final edited version will replace it automatically.



CCUS	Carbon capture, utilization, and storage
CO ₂	Carbon dioxide
DT	Decision tree
FESEM	Field-emission scanning electron microscopy
FTIR	Fourier-transform infrared spectroscopy
GA	Genetic algorithm
GPR	Gaussian process regression
HCl	Hydrochloric acid
HV	Hypervolume
KOH	Potassium hydroxide
LHS	Latin hypercube sampling
MAE	Mean absolute error
MBE	Mean bias error
MFC	Mass-flow controller
ML	Machine learning
MOO	Multi-objective optimization
MOPSO	Multi-objective particle swarm optimization
N ₂	Nitrogen
NSGA-II	Non-dominated Sorting Genetic Algorithm II
OPA	Oil palm ash
PD	Pure Diversity
PF	Pareto front
PSO	Particle swarm optimization
RMSE	Root mean squared error
SOO	Single-objective optimization
SVM	Support vector machine
SVR	Support vector regression

Roman symbols

A	Effective heat-transfer area used in the optional heat-loss model (m ²)
AC	CO ₂ adsorption capacity (mmol/g)
$\widehat{AC}$	Adsorption capacity predicted by the surrogate model (mmol/g)
$AC_{\min}$	Minimum acceptable predicted adsorption capacity (mmol/g)
AC_n	Normalized adsorption capacity (–)
a_0	Base heater-duty coefficient (–)
a_1	Temperature-dependent heater-duty coefficient (1/°C)
C	Initial CO ₂ concentration used as an optimization variable (vol%)
C_0	CO ₂ concentration at the column inlet (vol%)
C_t	CO ₂ concentration at the column outlet at time t (vol%)
c_1	Cognitive acceleration coefficient in PSO or MOPSO (–)
c_2	Social acceleration coefficient in PSO or MOPSO (–)
D_p	L_p Dissimilarity used in the pure diversity metric (–)
d_i	Nearest-neighbour distance of Pareto solution i (–)
$\bar{d}$	Mean nearest-neighbour distance (–)
E	Electrical energy consumption (kWh)
E_n	Normalized electrical-energy objective (–)
F	Total gas flow rate (mL/min)

F_0	Inlet CO ₂ volumetric flow rate used in the capacity calculation (mL/min)
	Multi-objective function vector (–)
f_{GPR}	Trained GPR mapping from operating variables to adsorption capacity (–)
f_{SOO}	Single-objective function minimized by GA or PSO (–)
$g(\mathbf{x})$	Nonlinear inequality constraint function (–)
$\mathbf{g}$	Global-best or archive-leader position in PSO-based optimization (–)
HV	Hypervolume of the Pareto set (–)
HV_n	Normalized hypervolume (–)
$k(\mathbf{x}, \mathbf{x}')$	GPR covariance or kernel function (–)
m	Mass of adsorbent loaded into the fixed-bed column (g)
N	Number of non-dominated Pareto solutions (–)
N_{obs}	Number of experimental observations (–)
n	Number of samples used in a regression metric (–)
$\dot{n}$	Molar gas-flow rate (mol/s)
p	Exponent used in the pure diversity dissimilarity measure (–)
P_1	Compressor inlet pressure (bar)
P_2	Compressor discharge pressure (bar)
P_{analyser}	Electrical power of the CO ₂ analyser (W)
P_{comp}	Gas-compression power (W)
P_{fixed}	Combined fixed electrical load (W)
P_{heater}	Average heater electrical power (W)
P_{loss}	Optional heat-loss power (W)
P_{MFC}	Electrical power of the mass-flow controller (W)
P_{PC}	Electrical power of the computer and data-acquisition system (W)
P_{rated}	Rated heater power (W)
P_{total}	Total average electrical power (W)
$\mathbf{p}_i$	Personal-best position of particle i (–)
q	Experimentally calculated adsorption capacity (mmol/g)
R	OPA-to-KOH mass ratio (–)
R_g	Universal gas constant (J/(mol K))
R^2	Coefficient of determination (–)
$\mathbf{r}$	Hypervolume reference point (–)
$\mathbf{r}_1, \mathbf{r}_2$	Uniform random vectors used in PSO velocity updates (–)
S	Spacing metric (–)
t	Contact or adsorption time (min)
t_e	End or equilibrium time of an adsorption experiment (min)
t_s	Operating time expressed in seconds (s)
T	Adsorption or set-point temperature (°C)
T_{amb}	Ambient temperature (°C)
T_g	Absolute gas temperature used in compression calculations (K)
U	Overall heat-transfer coefficient (W/(m ² K))
$\mathbf{v}_i^k$	Velocity of particle i at iteration k (–)
$\dot{V}$	Volumetric gas-flow rate (m ³ /s)
w	Inertia weight in PSO or MOPSO (–)
$\mathbf{x}$	Predictor matrix used for machine-learning model development (–)
$\mathbf{x}$	Optimization decision vector, $[R, F, T, C, t]$ (–)
$\mathbf{x}_i^k$	Position of particle i at iteration k (–)

$x_{\min}, x_{\max}$	Lower and upper bounds of the decision variables (–)
x_j	j Th component of an input vector (–)
y_i	Experimental adsorption capacity of observation i (mmol/g)
$\hat{y}_i$	Cross-validated adsorption-capacity estimate for observation i (mmol/g)
$\bar{y}$	Mean experimental adsorption capacity (mmol/g)
z_i	Normalized minimization-objective vector for Pareto solution i (–)

Greek symbols

η	Compressor efficiency (–)
ℓ_j	Characteristic length scale of the j th GPR input (variable-dependent)
$\lambda(\cdot)$	Lebesgue measure used to define hypervolume (–)
σ_f^2	Signal variance of the GPR kernel (response squared)

Subscripts and superscripts

amb	Ambient condition
comp	Compression contribution
fixed	Fixed electrical load
heater	Heater contribution
i, j	Solution, particle, observation, or variable indices
k	Optimization iteration index
loss	Heat-loss contribution
max	Maximum value
min	Minimum value
opt	Optimized value
rated	Rated equipment value
SOO	Single-objective optimization
MOO	Multi-objective optimization
$\hat{(\cdot)}$	Model-predicted quantity
$(\cdot)^*$	Optimal solution or optimal value
$(\cdot)^k$	Quantity evaluated at iteration k

1 Introduction

Carbon dioxide (CO₂) emissions from energy-intensive industries remain one of the greatest challenges in achieving global decarbonization targets. Although advances in renewable energy and improvements in energy efficiency have contributed to emission reduction, many industrial sectors, including power generation, cement production, steel manufacturing, and petrochemical processing, continue to rely on carbon-based fuels and feedstocks. Consequently, carbon capture, utilization, and storage (CCUS) have become an indispensable strategy for mitigating industrial CO₂ emissions while supporting the transition towards a low-carbon economy [1–3]. Among the available carbon capture technologies, adsorption has attracted considerable interest because of its relatively low energy requirement, operational flexibility, straightforward regeneration, and potential for cyclic operation under mild conditions compared with conventional solvent-based absorption processes [2, 3].

The performance and economic viability of adsorption processes are primarily governed by the characteristics of the adsorbent. Various porous materials, including zeolites, metal–organic frameworks (MOFs), porous polymers, silica-based adsorbents, and activated carbons, have been extensively investigated for CO₂ separation. Activated carbon remains particularly attractive because of its excellent thermal stability, chemical resistance,

moisture tolerance, relatively low production cost, and the possibility of being synthesized from renewable biomass resources and agricultural residues [2, 4]. Its adsorption behaviour depends not only on the total surface area but also on pore-size distribution, pore connectivity, surface functional groups, and the accessibility of active adsorption sites. These physicochemical properties can be tailored through appropriate activation and carbonization strategies to enhance adsorption performance for specific applications.

Among the available activation methods, potassium hydroxide (KOH) has been widely employed because of its ability to promote pore development and modify the surface chemistry of biomass-derived carbon materials. During high-temperature activation, KOH facilitates devolatilization, carbon etching, and the formation of interconnected micro- and mesoporous structures, thereby increasing the availability of adsorption sites and improving gas diffusion within the adsorbent matrix [4, 5]. Nevertheless, excessive activation may enlarge pore structures beyond the optimum range for CO₂ adsorption or partially damage the carbon framework, resulting in reduced adsorption performance. Consequently, optimizing the amount of activating agent is essential to achieve an appropriate balance between pore development, structural stability, and adsorption efficiency [4, 6].

The conversion of agricultural waste into functional adsorbents provides an additional environmental benefit by transforming low-value biomass residues into high-value carbon materials, thereby supporting waste minimization and circular resource utilization [5, 7]. This approach is particularly attractive in palm oil-producing countries, where large quantities of oil palm ash (OPA) are generated during the combustion of palm kernel shells and fibres in palm oil mills. OPA is generally regarded as a low-value solid residue despite its considerable carbon content and porous structure. Previous studies have demonstrated that acid washing, carbonization, and KOH functionalization can significantly improve the adsorption properties of OPA for CO₂ capture [1, 8]. In particular, OPA-KOH prepared at an OPA-to-KOH ratio of 1:2 exhibited a BET surface area of 30.95 m² g⁻¹, an average pore diameter of 72.71 Å, and a maximum CO₂ adsorption capacity of approximately 2.9 mmol g⁻¹ [1]. These findings indicate that the adsorption behaviour of OPA is governed by the combined effects of adsorbent preparation, pore structure, surface chemistry, gas composition, operating temperature, and gas–solid contact time rather than by any individual factor.

Although experimental investigations have provided valuable insights into the influence of these operating variables, identifying the optimum adsorption conditions remains a significant challenge because adsorption performance is controlled by complex nonlinear interactions among multiple parameters. The OPA-to-KOH ratio determines the physicochemical characteristics of the adsorbent, whereas gas flow rate, adsorption temperature, inlet CO₂ concentration, and contact time influence mass transfer, adsorption kinetics, and the overall driving force for CO₂ uptake. Conventional one-factor-at-a-time experimentation is useful for examining the influence of individual variables but is often unable to capture their coupled effects. Moreover, exploring a multidimensional operating space exclusively through experiments requires considerable time, materials, and operating costs, making comprehensive optimization impractical for many adsorption systems.

Recent advances in machine learning (ML) have provided an efficient alternative for analysing complex adsorption systems that exhibit nonlinear behaviour and strong interactions among operating variables. Unlike conventional statistical approaches, ML models are capable of learning complex relationships directly from experimental data without requiring explicit mechanistic formulations. Once adequately trained and validated, these surrogate models can rapidly predict adsorption performance for new operating conditions within the experimental domain. This capability significantly reduces the computational and experimental effort required for process evaluation and enables the rapid screening of numerous operating scenarios before experimental verification. Consequently, ML has been increasingly applied in adsorption processes, carbon capture technologies, catalyst development, and broader chemical engineering applications to accelerate process design and optimization.

Despite these advances, three important research gaps remain. First, most experimental studies on biomass-derived adsorbents continue to rely primarily on conventional parametric investigations, where optimization is performed by varying one or two operating variables at a time. Such approaches are often unable to capture the complex interactions among adsorbent preparation parameters and operating conditions that collectively determine adsorption performance. Second, although several studies have employed machine learning models to

predict adsorption behaviour, their primary focus has been on improving predictive accuracy rather than integrating prediction with systematic process optimization. Third, optimization studies frequently prioritize maximum adsorption capacity while overlooking the associated electrical energy required to achieve those operating conditions. Since energy consumption directly influences process economics and environmental sustainability, considering adsorption capacity alone may not necessarily lead to the most practical operating strategy.

Addressing these challenges requires an integrated framework that combines accurate predictive modelling with multi-objective optimization to simultaneously evaluate process performance and energy consumption. Such an approach enables decision-making based not only on maximum CO₂ uptake but also on the trade-offs between adsorption efficiency and operational energy demand. Unlike single-objective optimization, multi-objective optimization simultaneously considers multiple competing objectives without assigning predefined weights, allowing a set of non-dominated solutions to be generated for engineering evaluation. These solutions collectively form the Pareto front, where improvement in one objective can only be achieved by sacrificing at least one other objective. Consequently, the Pareto front provides a systematic basis for selecting operating conditions according to practical design priorities rather than relying on a single optimum solution. Similar decision-making frameworks have been successfully applied in environmental and chemical engineering to balance process performance with operating cost and energy consumption, demonstrating their value for sustainable process development [9]. From a chemical engineering perspective, this integration supports the development of more sustainable adsorption processes by reducing experimental effort while identifying operating conditions that are technically and economically favourable.

Building upon our previous experimental investigation of KOH-functionalized OPA [1], this study develops an integrated surrogate modelling and optimization framework for CO₂ adsorption. The experimental dataset was used to develop and compare Gaussian Process Regression (GPR), Support Vector Regression (SVR), and Regression Tree models for predicting adsorption capacity based on the OPA-to-KOH ratio, gas flow rate, adsorption temperature, inlet CO₂ concentration, and contact time [1, 10, 11]. Among these models, the best-performing surrogate was subsequently coupled with Genetic Algorithm (GA) and Particle Swarm Optimization (PSO) to identify operating conditions that maximize adsorption capacity. Multi-Objective Particle Swarm Optimization (MOPSO) and the Non-dominated Sorting Genetic Algorithm II (NSGA-II) were further employed to establish Pareto-optimal solutions that simultaneously maximize adsorption performance and minimize electrical energy consumption.

The specific objectives of this study were to (1) develop and compare machine learning models for predicting CO₂ adsorption capacity using experimentally measured operating variables; (2) determine the operating conditions that maximize adsorption capacity through single-objective optimization; (3) establish the trade-off between adsorption capacity and electrical energy consumption using two independent multi-objective optimization algorithms; and (4) identify representative minimum-energy, maximum-capacity, and knee-point operating conditions to support engineering decision-making.

Unlike previous studies that primarily focus on adsorption prediction or single-objective optimization, this work integrates experimental adsorption data, surrogate modelling, and evolutionary multi-objective optimization within a unified framework that explicitly incorporates electrical energy consumption into the optimization process. The proposed methodology provides a systematic and reproducible approach for identifying energy-efficient operating conditions while maintaining high CO₂ adsorption performance. Beyond the present case study, the framework can be readily adapted to other adsorption systems and biomass-derived adsorbents, demonstrating the broader potential of artificial intelligence to support sustainable process development in chemical engineering.

2 Materials and experimental methods

2.1 Study basis and data provenance

The experimental dataset analysed in this study was obtained from our previously reported investigation on CO₂ adsorption using KOH-functionalized oil palm ash (OPA) [1]. Rather than repeating the experimental investigation, the present work utilizes the published adsorption data to develop and validate machine learning models, followed by single- and multi-objective optimization of the adsorption process. The experimental procedures are summarized here to provide sufficient methodological context for the development of the computational framework, while detailed descriptions of adsorbent characterization, adsorption kinetics, equilibrium analysis, and thermodynamic evaluation are available in the original publication [1]. The measured adsorption data were first pre-processed, subsequently used for surrogate model development, and finally integrated with evolutionary optimization algorithms to identify optimum operating conditions.

2.2 Biomass source, chemicals, and gases

Oil palm ash was obtained from the United Oil Palm Mill, Pulau Pinang, Malaysia. Hydrochloric acid solution (HCl, 37%) and potassium hydroxide pellets (KOH, ≥98%) were supplied by Merck Sdn. Bhd., Selangor, Malaysia. High-purity carbon dioxide and nitrogen gases, each with a stated purity of 99.9%, were obtained from Well-gas (M) Sdn. Bhd. Distilled water was used for washing, solution preparation, and pH adjustment. The chemicals employed in this study were of analytical grade and were used without further purification (Table 1).

2.3 Preparation of KOH-functionalized oil palm ash

Pristine OPA was initially washed with 1.0 M hydrochloric acid to remove residual ash, inorganic impurities, and loosely bound contaminants. The material was subsequently rinsed with warm distilled water until a neutral pH was achieved, followed by filtration and drying at 105 °C overnight. For chemical activation, KOH pellets were dissolved in distilled water before the dried OPA was impregnated at the desired OPA-to-KOH ratios. The suspension was stirred for 5 h, refluxed at 90 °C for 8 h, and further treated in a water-bath shaker at 50 °C and 150 rpm for 1 h to improve impregnation. The activated samples were then dried overnight at 60 °C. Carbonization was carried out in a horizontal tubular furnace at 700 °C for 1 h under a continuous nitrogen flow of 150 mL min⁻¹ with a heating rate of 10 °C min⁻¹. The inert nitrogen atmosphere minimized oxidation during thermal treatment and promoted the development of a porous carbon structure. The selected carbonization temperature was intended to promote pore development and removal of volatile components while avoiding the excessive structural degradation that can occur at substantially higher temperatures [5, 8]. Activated samples were designated as OPA-KOH(x), where x denotes the OPA-to-KOH ratio (1:0.5, 1:1, 1:2, and 1:3). These preparation conditions were adopted based on our previous optimization study, which demonstrated their suitability for producing porous adsorbents with favourable CO₂ adsorption characteristics [1].

Table 1 Materials used for adsorbent preparation and adsorption experiments

Material	Specification	Purpose
Oil palm ash	Palm-mill combustion residue	Carbon precursor
Hydrochloric acid	37%; diluted to 1.0 M	Removal of ash and inorganic contaminants
Potassium hydroxide	≥98%	Chemical activation and surface functionalization
Carbon dioxide	99.9%	Adsorbate gas
Nitrogen	99.9%	Carrier, dilution, and regeneration gas
Distilled water	Laboratory grade	Washing and KOH-solution preparation

2.4 Adsorbent characterization

Nitrogen physisorption analysis was performed using a Micromeritics ASAP 2020 instrument. Pristine OPA, acid-washed OPA, and OPA-KOH(1:2) were degassed for 6 h at 300 °C under a pressure below 10^{-1} Pa before analysis. The Brunauer–Emmett–Teller method was used to calculate specific surface area, and the Barrett–Joyner–Halenda method was used to estimate pore-size distribution from the desorption branch. Surface morphology was examined using a Quanta FEG 450 field-emission scanning electron microscope. The instrument was operated in high-vacuum mode over an accelerating-voltage range of 1–30 kV. Micrographs were used to compare the dense structure of pristine OPA with the pore development obtained after acid washing, carbonization, and KOH activation. Functional groups were characterized by Fourier-transform infrared spectroscopy using a Perkin-Elmer 2000 spectrometer. Approximately 0.9 mg of sample was mixed with 63 mg of KBr and ground using an agate mortar before being pressed into a pellet. Spectra were recorded over 4000–400 cm^{-1} at room temperature. These analyses were used to verify changes in oxygen-containing groups and other surface functionalities following KOH treatment. Comprehensive characterization results and their relationship with adsorption performance have been reported previously [1]. Therefore, only the essential characterization procedures are summarized here to establish the experimental basis of the machine learning dataset.

2.5 Fixed-bed CO₂ adsorption experiments

Dynamic adsorption experiments were conducted in a fixed-bed column with a radius of 0.55 cm and a height of 14 cm. For each run, 0.5 g of OPA-KOH(x) was loaded into the column. Before adsorption, the packed bed was heated to 110 °C and maintained at that temperature for 1 h under a nitrogen flow of 90 mL min^{-1} to remove moisture and weakly adsorbed species. The bed was then cooled to the required adsorption temperature.

A controlled CO₂/N₂ gas mixture was introduced at the selected total flow rate. The baseline gas composition corresponded to 10 vol% CO₂ in nitrogen, while additional experiments covered inlet CO₂ concentrations from 7 to 15 vol%. The total flow rate was varied between 30 and 60 mL min^{-1} , and the adsorption temperature was varied between 30 and 60 °C. The OPA-to-KOH ratio was also varied to represent the different activated samples. The principal operating variables were therefore the OPA-to-KOH ratio, total flow rate, adsorption temperature, inlet CO₂ concentration, and adsorption time.

The outlet CO₂ concentration was monitored continuously using a Quantek 906 CO₂ analyser with a detection limit of 20 ppm. The analyser was connected to a computer through a HOBO data logger, and the concentration–time data were recorded using HOBOWare Pro software. Each run was continued until the outlet concentration approached the inlet concentration, indicating saturation of the bed. Every operating condition was tested in triplicate to assess repeatability.

The normalized breakthrough response was expressed as

$$\theta(t) = \frac{C_t}{C_0} \quad (1)$$

where C_0 is the inlet CO₂ concentration and C_t is the measured outlet concentration at time t . Values of θ close to zero represent effective capture at the start of adsorption, while θ approaching unity indicates bed saturation.

2.6 Adsorption-capacity calculation

The cumulative adsorption capacity was calculated from the area above the normalized breakthrough curve:

$$q = \frac{C_0 F_0}{m} \int_0^{t_f} \left(1 - \frac{C_t}{C_0}\right) dt \quad (2)$$

where q is the CO_2 adsorption capacity, F_0 is the inlet volumetric flow rate of CO_2 , m is the mass of adsorbent, C_0 and C_t are the inlet and outlet CO_2 concentrations, respectively, and t_f is the final adsorption time. The appropriate gas-volume-to-mole conversion was applied during data processing so that adsorption capacity was reported in mmol g^{-1} . Numerical integration of the recorded breakthrough profile was used to evaluate the time integral.

2.7 Experimental design and dataset construction

The experimental programme produced 346 observations covering changes in adsorbent preparation and operating conditions. For machine-learning analysis, the predictor vector was arranged as

$$\mathbf{x} = [R \quad F \quad T \quad C \quad t]^T \quad (3)$$

and the corresponding response was

$$y = AC \quad (4)$$

where R is the numerical OPA-to-KOH ratio, F is the total flow rate, T is the adsorption temperature, C is the inlet CO_2 concentration, t is the contact time, and AC is the measured adsorption capacity. The numerical ranges retained for model development are summarized in Table 2.

Each row of the dataset corresponded to one recorded operating state and its experimentally calculated adsorption capacity. The time-resolved measurements enabled the learning algorithms to reproduce the development of adsorption capacity during a run rather than predicting only a single equilibrium value. Data were checked for missing and non-finite entries before model training. All predictor variables were retained because each represented either a material-preparation factor or a directly controlled adsorption condition.

2.8 Quality assurance and methodological scope

Repeat measurements, continuous concentration logging, and fixed instrument settings were used to reduce experimental variability. The same adsorbent mass, column geometry, pre-treatment procedure, and saturation criterion were applied throughout the experimental campaign. The machine-learning and optimization stages were restricted to the experimentally observed variable ranges. This restriction was adopted to avoid unsupported extrapolation from the surrogate models.

The experimental method generated the adsorption-capacity dataset used in the subsequent machine-learning comparison and optimization. The present study focused on operating electricity consumption during adsorption, including fixed instrumentation, heater demand, and gas-compression power. Energy associated with OPA collection, acid washing, KOH production, drying, carbonization, and adsorbent replacement was outside the optimization boundary and should be considered separately in a complete life-cycle or techno-economic assessment.

Table 2 Input and output variables used for machine-learning model development

Variable type	Parameter	Symbol	Unit	Minimum	Maximum
Input	OPA-to-KOH ratio	R	Dimensionless	0.33	2.00
Input	Total flow rate	F	mL min^{-1}	30	60
Input	Adsorption temperature	T	$^{\circ}\text{C}$	30	60
Input	Initial CO_2 concentration	C	vol%	7	15
Input	Contact time	t	Min	0	50
Output	Adsorption capacity	AC	mmol g^{-1}	0	3

3 Optimization methodology

3.1 Overall optimization framework

The workflow integrates experimental adsorption data, machine learning, and evolutionary optimization into a unified decision-support framework for identifying energy-efficient CO₂ adsorption conditions. The methodology consisted of three sequential stages: (1) development and validation of machine learning surrogate models, (2) estimation of electrical energy consumption associated with laboratory-scale adsorption, and (3) single- and multi-objective optimization of adsorption performance.

Initially, the experimental dataset comprising 346 adsorption observations was pre-processed and used to train three machine learning models, namely Gaussian Process Regression (GPR), Support Vector Regression (SVR), and Regression Tree (Tree). The predictive performance of GPR, SVR, and Regression Tree models was evaluated using fivefold cross-validation. The same cross-validation partition was applied to all three models to ensure a consistent comparison. In each iteration, fourfolds were used for model training and the remaining fold was used for validation. Predictions from the five validation folds were combined to calculate the overall model-performance indicators. The most accurate model was subsequently selected as the surrogate model for optimization.

The validated surrogate model was then coupled with Genetic Algorithm (GA) and Particle Swarm Optimization (PSO) to determine the operating conditions that maximize CO₂ adsorption capacity. Finally, Multi-Objective Particle Swarm Optimization (MOPSO) and Non-dominated Sorting Genetic Algorithm II (NSGA-II) were employed to simultaneously optimize adsorption capacity and electrical energy consumption, thereby generating Pareto-optimal operating conditions representing different engineering trade-offs. This integrated framework enables rapid evaluation of numerous operating scenarios without requiring additional laboratory experiments, thereby reducing computational effort and experimental cost while supporting informed engineering decision-making.

3.2 Development of machine learning

Machine learning models were developed to establish predictive relationships between the experimental operating variables and CO₂ adsorption capacity. Rather than replacing experimental observations, these models served as computational surrogates that enabled rapid prediction of adsorption performance across the investigated operating domain. The predictor matrix consisted of the OPA-to-KOH ratio, gas flow rate, adsorption temperature, inlet CO₂ concentration, and contact time, while adsorption capacity was used as the response variable. Three regression algorithms with different modelling characteristics were evaluated to determine the most suitable surrogate model. Gaussian Process Regression (GPR) was selected because of its ability to model highly nonlinear relationships while providing smooth interpolation across sparse experimental data. Support Vector Regression (SVR) and Regression Tree models were included as benchmark methods to evaluate prediction robustness and model generalization.

Fivefold cross-validation was used to evaluate the predictive performance of the three surrogate models. The dataset was divided into five mutually exclusive folds using a fixed random seed. For each iteration, the model was fitted using four training folds and evaluated using the remaining validation fold. Each observation was therefore predicted by a model that had not been trained using that observation.

Predictor standardization for GPR and SVR was performed internally during model fitting using only the training observations within each fold. Model parameters were similarly estimated from the corresponding training folds. The validation observations were used only to generate out-of-fold predictions and were not included during model fitting.

The same fold assignments were used for GPR, SVR, and Regression Tree models. The out-of-fold predictions from all five iterations were combined to calculate R^2 , RMSE, MAE, mean bias error, and the regression relationship between experimental and predicted adsorption capacities. No separate independent holdout test set

was used; therefore, the reported results represent fivefold cross-validated performance rather than blind-test performance.

A variance-based global sensitivity analysis was conducted using the Sobol method to quantify the influence of the five operating variables on predicted adsorption capacity and electrical energy consumption. The OPA-to-KOH ratio, flow rate, adsorption temperature, inlet CO₂ concentration, and contact time were sampled independently using uniform distributions over their experimentally investigated ranges. The selected surrogate was used to evaluate adsorption capacity, whereas energy consumption was determined from the explicit laboratory energy model. First-order Sobol indices were calculated to quantify the direct contribution of each variable, while total-order indices included both individual and interaction effects. The analysis used 10,000 base samples, and 95% confidence intervals were estimated using 500 bootstrap replicates.

3.2.1 GPR surrogate model

GPR was employed to model the nonlinear relationship between adsorption capacity and the five experimental operating variables. GPR was selected because it provides highly flexible nonlinear regression while maintaining excellent predictive capability for relatively small experimental datasets. Furthermore, the probabilistic nature of GPR reduces the likelihood of overfitting and provides smooth interpolation between experimental observations. The experimental predictor matrix was arranged as $\mathbf{X} = [R, F, T, C, t]$, while adsorption capacity was used as the response variable. A GPR model with an automatic relevance determination squared-exponential kernel was fitted to represent the nonlinear relationship between the five operating variables and adsorption capacity. The predictors were standardized internally during model training.

The covariance function was expressed as

$$k(\mathbf{x}, \mathbf{x}') = \sigma_f^2 \exp \left[-\frac{1}{2} \sum_{j=1}^5 \frac{(x_j - x'_j)^2}{\ell_j^2} \right]$$

where σ_f^2 is the signal variance and ℓ_j is the characteristic length scale of the j th input variable. The trained surrogate provided the predicted adsorption capacity

$$\widehat{AC} = f_{\text{GPR}}(R, F, T, C, t)$$

for every candidate solution generated during optimization. The use of the surrogate substantially reduced the computational cost because the optimization algorithms did not require repeated experiments [11]. The resulting surrogate model was subsequently embedded within the optimization algorithms, allowing adsorption capacity to be estimated rapidly without repeated experimental evaluation.

The ARD squared-exponential kernel was selected based on the characteristics of the predicted response. Although the outlet CO₂ concentration may exhibit a relatively sharp transition near breakthrough, the response modelled in this study was cumulative adsorption capacity, which was calculated by integrating the breakthrough curve over time. The integration reduces local fluctuations and produces a smooth, continuous response with respect to contact time and the other operating variables. The squared-exponential covariance function was therefore suitable for representing the gradual variation of cumulative adsorption capacity within the investigated experimental domain. Its automatic relevance determination formulation assigns an independent characteristic length scale to each predictor, allowing the model to represent different rates of variation associated with the OPA-to-KOH ratio, flow rate, adsorption temperature, inlet CO₂ concentration, and contact time. Smaller length scales accommodate relatively rapid nonlinear changes, while larger length scales represent more gradual effects.

A Matérn 5/2 kernel, which assumes lower differentiability and permits less smooth functions, may be more appropriate when directly modelling the sharp transition of the outlet concentration breakthrough curve. However,

the smoother squared-exponential kernel was retained because the present response variable was the integrated cumulative adsorption capacity rather than the instantaneous breakthrough concentration.

3.2.2 Support vector machine (SVM)

SVR was developed as an alternative nonlinear regression model for comparison with GPR. SVR seeks to determine a regression function that minimizes prediction error while maintaining model generalization through structural risk minimization. A SVM regression model with an ε -insensitive loss function was developed as an alternative surrogate. The SVM model sought to find a function $f(x)$ that deviates from the actual target values by at most ε for all training data points, while maintaining maximum flatness. The SVM regression function was expressed as:

$$f(x) = \sum_{i=1} (\alpha_i - \alpha_i^*) K(x_i, x) + b$$

where α_i and α_i^* are Lagrange multipliers, $K(x_i, x)$ is the kernel function, and b is the bias term. A Gaussian radial basis function (RBF) kernel was employed:

$$K(x_i, x_j) = \exp\left(-\gamma \|x_i - x_j\|^2\right)$$

where γ is the kernel width parameter. The SVM hyperparameters, including the penalty parameter C_{SVM} (box constraint), ε (epsilon), and γ (kernel scale), were optimized using Bayesian optimization with fivefold cross-validation to minimize the cross-validation mean squared error. A Gaussian radial basis function kernel was employed. The kernel scale was determined automatically by MATLAB using the training observations within each cross-validation fold. Predictor standardization was also estimated exclusively from the corresponding training observations. The validation fold was used only to calculate the out-of-fold predictions. The search ranges were $C_{SVM} \in [0.1, 100]$, $\varepsilon \in [0.001, 1]$, and $\gamma \in [10^{-4}, 10]$. The final optimized SVM model provided the predicted adsorption capacity:

$$AC^\wedge = f_{SVM}(R, F, T, C, t)$$

for each candidate solution [12]. A Gaussian radial basis function kernel was employed because of its ability to capture nonlinear relationships commonly encountered in adsorption systems.

3.2.3 Regression tree (tree)

Tree modelling was included as a third surrogate modelling approach to provide a simple and interpretable benchmark against the more sophisticated nonlinear regression models. Unlike GPR and SVR, Regression Tree recursively partitions the predictor space into homogeneous regions, allowing prediction rules to be represented in a hierarchical structure.

A regression tree model was developed as a third surrogate modelling approach. The tree model partitioned the predictor space into distinct regions and fitted a constant prediction within each region. The tree was grown using recursive binary splitting, where the best split at each node was selected to minimize the sum of squared errors:

$$\min_{j,s} \left[\min_{c_1} \sum_{x_i \in R_1(j,s)} (y_i - c_1)^2 + \min_{c_2} \sum_{x_i \in R_2(j,s)} (y_i - c_2)^2 \right]$$

where j is the splitting variable, s is the split point, and R_1 and R_2 are the resulting regions with predictions c_1 and c_2 (the mean response in each region).

To prevent overfitting, the tree was pruned using cost-complexity pruning, where the following criterion was minimized:

$$\sum_{m=1}^{|T|} \sum_{x_i \in R_m} (y_i - \hat{y}_{R_m})^2 + \alpha |T|$$

where $|T|$ is the number of terminal nodes and α is the complexity parameter. The optimal tree complexity was determined through cross-validation. The trained regression tree provided the predicted adsorption capacity:

$$AC^{\wedge} = f_{Tree}(R, F, T, C, t)$$

for each candidate solution generated during optimization [13].

The three surrogate models (GPR, SVM, and Tree) were compared using R^2 , RMSE, MAE, and mean bias error calculated from the combined out-of-fold predictions generated through fivefold cross-validation. No independent test set was used in the present study. The model with the best predictive performance was selected as the primary surrogate for optimization. GPR was selected as the surrogate model because it achieved the highest cross-validated predictive accuracy and the lowest prediction errors among the evaluated models. Although GPR can provide predictive uncertainty estimates, uncertainty quantification was not investigated in this study and is therefore not used as a basis for model selection.

3.3 Single-objective optimization using GA and PSO

Single-objective optimization (SOO) was performed to determine the operating condition that maximized the adsorption capacity predicted by the GPR model. Genetic Algorithm (GA) and Particle Swarm Optimization (PSO) were evaluated independently using the same decision variables, surrogate model, and experimental bounds.

The SOO problem was formulated as

$$\max_{\mathbf{x}} \widehat{AC}(\mathbf{x})$$

Because the MATLAB solvers minimize their objective functions, adsorption-capacity maximization was implemented as

$$\min_{\mathbf{x}} f_{\text{SOO}}(\mathbf{x}) = -\widehat{AC}(\mathbf{x})$$

subject to

$$\mathbf{x}_{\min} \leq \mathbf{x} \leq \mathbf{x}_{\max}$$

Energy consumption was not included in the SOO objective. Instead, the laboratory energy model was evaluated after optimization at the maximum-adsorption-capacity solution:

$$E_{\text{SOO,opt}} = E(\mathbf{x}_{\text{SOO,opt}})$$

This enabled comparison of the electrical demand associated with the best adsorption condition obtained by each optimizer.

3.3.1 Laboratory electrical energy model

The total average electrical power of the laboratory rig was calculated as

$$P_{\text{total}} = P_{\text{fixed}} + P_{\text{heater}} + P_{\text{comp}} + P_{\text{loss}}$$

The fixed equipment load included a mass flow controller, gas analyser, and computer:

$$P_{\text{fixed}} = P_{\text{MFC}} + P_{\text{analyser}} + P_{\text{PC}} = 5 + 12 + 45 = 62 \text{ W}$$

The heater power was represented by a bounded temperature-dependent duty model:

$$P_{\text{heater}} = P_{\text{rated}} \min [1, \max (0, a_0 + a_1 (T - T_{\text{amb}}))]$$

where $P_{\text{rated}} = 300 \text{ W}$, $T_{\text{amb}} = 27 \text{ }^{\circ}\text{C}$, $a_0 = 0.15$, and $a_1 = 0.02 \text{ per } ^{\circ}\text{C}$ [1]. The gas-compression power was estimated from an isothermal compression relationship. The inlet molar flow rate was calculated as

$$\dot{n} = \frac{P_1 \dot{V}}{R_g T_g}$$

and the compression power was estimated by

$$P_{\text{comp}} = \frac{\dot{n} R_g T_g \ln (P_2 / P_1)}{\eta}$$

where $P_1 = 1.0 \text{ bar}$, $P_2 = 2.0 \text{ bar}$, $T_g = 298.15 \text{ K}$, and $\eta = 0.75$ [1].

But this term was disabled in the reported calculations because the empirical heater-duty model was assumed to include the average power needed to compensate for heat losses. Gas-compression power was included in all reported energy calculations. Only the separate heat-loss term was disabled because the empirical heater-duty model was assumed to represent the average heater demand, including thermal losses. This avoided double counting of the heat-loss contribution.

The total electrical energy was calculated as

$$E(\mathbf{x}) = \frac{P_{\text{total}}(60t)}{3.6 \times 10^6}$$

where t is expressed in minutes and E is obtained in kWh. In the adopted formulation, flow rate, temperature, and contact time directly affected energy consumption. The OPA-to-KOH ratio and initial concentration influenced energy indirectly through their influence on the GPR-predicted adsorption capacity [1]. No surrogate model was developed for the energy vector. Energy consumption was calculated explicitly for every candidate operating point using the equations above. The multi-objective optimization mapping was therefore

$$\mathbf{f}(\mathbf{x}) = \begin{bmatrix} E(\mathbf{x}) \\ -\widehat{AC}(\mathbf{x}) \end{bmatrix}$$

where $\widehat{AC}(\mathbf{x})$ is the adsorption capacity predicted by the trained GPR model. Thus, the first objective was obtained directly from the empirical and thermodynamic energy model, whereas only the adsorption-capacity objective was evaluated using a machine-learning surrogate.

The energy-model parameters were obtained from the equipment ratings, operating specifications, and calibration measurements of the laboratory fixed-bed system. The fixed load comprised the rated or measured power consumption of the mass flow controller, CO_2 analyser, and computer. The heater rated power and coefficients a_0 and a_1 were determined from the observed heater-power response over the experimental temperature range. The ambient temperature, gas inlet conditions, pressure ratio, and compression efficiency corresponded to the operating conditions and specifications of the experimental equipment. These parameters were obtained from our experimental set up. Consequently, the calculated energy values represent laboratory-scale estimates for comparing operating conditions within the present system.

3.3.2 Genetic algorithm

The GA represented each candidate solution as a real-valued vector containing the five operating variables. At each generation, the GPR model predicted adsorption capacity for every individual. Individuals with lower values of $-\widehat{AC}$ were preferentially selected to produce the next generation.

Tournament selection was used to choose parent solutions. Scattered crossover combined the decision variables of the selected parents, while adaptive feasible mutation introduced additional variation and maintained the decision variables within their permitted bounds. Elitism preserved the best solutions between consecutive generations. The GA terminated when the maximum number of generations was reached, the change in objective value fell below the specified tolerance, or the maximum number of stall generations was exceeded [16] (Table 3).

The final GA solution was defined as

$$\mathbf{x}_{GA,opt} = \arg \max_{\mathbf{x}} \widehat{AC}(\mathbf{x})$$

3.3.3 Particle swarm optimization

In PSO, each particle represented one operating condition. The particles moved through the five-dimensional search space according to their individual best positions and the best position identified by the swarm. The velocity and position updates were represented by

$$\mathbf{v}_i^{k+1} = w\mathbf{v}_i^k + c_1\mathbf{r}_1(\mathbf{p}_i - \mathbf{x}_i^k) + c_2\mathbf{r}_2(\mathbf{g} - \mathbf{x}_i^k)$$

and

$$\mathbf{x}_i^{k+1} = \mathbf{x}_i^k + \mathbf{v}_i^{k+1}$$

where $\mathbf{x}_i^k$ and $\mathbf{v}_i^k$ are the position and velocity of particle i at iteration k , $\mathbf{p}_i$ is its personal-best position, $\mathbf{g}$ is the global-best position, w is the inertia weight, c_1 and c_2 are acceleration coefficients, and $\mathbf{r}_1$ and $\mathbf{r}_2$ are uniformly distributed random vectors [14].

The MATLAB particleswarm solver was used with a swarm size of 100, a maximum of 50 iterations, and a function tolerance of 1.0×10^{-6} . The solver defaults were retained for the remaining PSO parameters [15]. The final solution was (Table 4)

$$\mathbf{x}_{PSO,opt} = \arg \max_{\mathbf{x}} \widehat{AC}(\mathbf{x})$$

The GA and PSO solutions were compared using the maximum predicted adsorption capacity, optimized operating conditions, total average power, and calculated electrical energy.

Table 3 Settings used for the single-objective GA

Parameter	Setting
Number of decision variables	5
Population size	100
Maximum generations	100
Elite count	5
Selection method	Tournament selection
Crossover method	Scattered crossover
Crossover fraction	0.80
Mutation method	Adaptive feasible mutation
Function tolerance	1.0×10^{-6}
Maximum stall generations	30
Search bounds	Experimental minimum and maximum

Table 4 Settings used for the single-objective PSO

Parameter	Setting
Number of decision variables	5
Swarm size	100
Maximum iterations	50
Function tolerance	1.0×10^{-6}
Search bounds	Experimental minimum and maximum
Remaining PSO coefficients	MATLAB solver defaults

3.4 Multi-objective optimization using MOPSO and NSGA-II

Multi-objective optimization (MOO) was used to determine the trade-off between adsorption performance and energy consumption. Multi-Objective Particle Swarm Optimization (MOPSO) and the Non-dominated Sorting Genetic Algorithm II (NSGA-II) were applied independently using the same GPR surrogate, laboratory energy model, decision-variable bounds, and adsorption-capacity constraint.

The two-objective problem was formulated as

$$\min_{\mathbf{x}} \mathbf{f}(\mathbf{x}) = [E(\mathbf{x}), -\widehat{AC}(\mathbf{x})]$$

A minimum adsorption-capacity constraint was imposed to prevent selection of low-energy conditions with inadequate capture performance:

$$\widehat{AC}(\mathbf{x}) \geq AC_{\min}$$

or equivalently

$$g(\mathbf{x}) = AC_{\min} - \widehat{AC}(\mathbf{x}) \leq 0$$

Both algorithms generated non-dominated solutions representing different compromises between energy use and adsorption capacity.

3.4.1 Multi-objective particle swarm optimization

The MOPSO population was initialized by Latin hypercube sampling to provide broad coverage of the feasible search domain. Each particle retained a personal-best solution, while feasible non-dominated solutions were stored in an external archive. The velocity and position updates followed the PSO equations given in Sect. 3.1.2, but the global leader was replaced by a leader selected from the Pareto archive.

The inertia weight decreased linearly from $w_{\max}$ to $w_{\min}$ according to

$$w_k = w_{\max} - (w_{\max} - w_{\min}) \frac{k-1}{N_{\text{iter}}-1}$$

Crowding distance was used in binary tournament selection to favour archive members located in less densely populated regions. Constraint dominance was applied when updating the personal best and the archive. A feasible solution was preferred over an infeasible one; when both solutions were infeasible, the one with the lower constraint violation was retained; and conventional Pareto dominance was used when both solutions were feasible [16].

A decaying Gaussian mutation operator was included to improve exploration during early iterations. Particle velocities were limited to 20% of the range of each decision variable. Positions outside the bounds were returned to the nearest boundary and the corresponding velocity components were set to zero (Table 5).

Table 5 Settings used for MOPSO

Parameter	Setting
Number of decision variables	5
Swarm size	250
Maximum iterations	250
External archive capacity	300
Cognitive coefficient, c_1	1.8
Social coefficient, c_2	1.8
Maximum inertia weight, $w_{\max}$	0.90
Minimum inertia weight, $w_{\min}$	0.40
Velocity limit	20% of each variable range
Initial mutation probability	0.15
Initial mutation scale	0.10
Population initialization	Latin hypercube sampling
Leader selection	Crowding-distance binary tournament
Constraint handling	Constrained dominance

3.4.2 Non-dominated sorting genetic algorithm II

NSGA-II was implemented using MATLAB's controlled-elitist gamultiobj solver. The initial population was generated using Latin hypercube sampling. Candidate solutions were ranked using non-dominated sorting, while diversity was preserved by favouring solutions in less crowded regions of the objective space. Crossover and adaptive feasible mutation were used to generate offspring, and elitism retained high-quality non-dominated solutions across generations [17].

A population size of 400 and a maximum of 250 generations were used. The Pareto fraction and crossover fraction were set to 0.85. The function tolerance was 1.0×10^{-10} , and the maximum number of stall generations was 80. Non-vectorized evaluation was retained because the nonlinear adsorption-capacity constraint was evaluated separately for each candidate solution [18]. A dedicated summary table listing the execution parameters for both metaheuristic algorithms is listed in Appendix 2.

Latin hypercube sampling was used only to initialise the MOPSO and NSGA-II populations and was not used to determine the operating conditions of the original fixed-bed experiments. Each normalised LHS point was mapped to the experimentally investigated bounds of the five decision variables, namely an OPA-to-KOH ratio of 0.33–2.00, a total flow rate of 30–60 mL min⁻¹, an adsorption temperature of 30–60 °C, an inlet CO₂ concentration of 7–15 vol%, and a contact time of 0–50 min. Candidate solutions containing non-finite values or lying outside these bounds were not accepted. In MOPSO, positions exceeding the permitted range were returned to the nearest boundary and the corresponding velocity components were set to zero. In NSGA-II, the specified lower and upper bounds and the adaptive feasible mutation operator maintained the solutions within the permitted domain. A minimum predicted adsorption-capacity constraint was additionally imposed to remove operating points with inadequate capture performance. No separate dynamic stability filter was applied because the optimisation was based on a surrogate model of cumulative adsorption capacity rather than a dynamic fixed-bed model. Consequently, the feasibility constraints restricted the search to the experimentally demonstrated operating domain but did not constitute a formal dynamic stability assessment (Table 6).

3.4.3 Pareto-front processing and compromise selection

After optimization, non-finite and near-duplicate objective vectors were removed. The remaining solutions were checked for Pareto dominance and sorted in ascending order of energy consumption. Three representative solutions were identified: the minimum-energy point, the maximum-adsorption-capacity point, and the knee point.

Energy and adsorption capacity were normalized as:

Table 6 Settings used for NSGA-II

Parameter	Setting
Number of decision variables	5
Population size	400
Maximum generations	250
Pareto fraction	0.85
Crossover fraction	0.85
Mutation method	Adaptive feasible mutation
Initial population	Latin hypercube sampling
Function tolerance	1.0×10^{-10}
Maximum stall generations	80
Vectorized evaluation	Disabled
Constraint	Minimum predicted adsorption capacity

$$E_n = \frac{E - E_{\min}}{E_{\max} - E_{\min}}$$

and

$$AC_n = \frac{AC - AC_{\min}}{AC_{\max} - AC_{\min}}$$

The normalized minimization vector was:

$$\mathbf{z} = [E_n, 1 - AC_n]$$

The knee solution was selected as the Pareto point closest to the ideal point [0, 0]:

$$i_{\text{knee}} = \arg \min_i \sqrt{E_{n,i}^2 + (1 - AC_{n,i})^2}$$

This procedure provided a balanced solution without assigning subjective weights to energy and adsorption capacity.

3.4.4 Pareto-front quality metrics

The Pareto fronts generated by MOPSO and NSGA-II were assessed using spacing, hypervolume, and Pure Diversity [19, 20]. The metrics were calculated in the normalized minimization objective space $\mathbf{z} = [E_n, 1 - AC_n]$.

For each Pareto solution, the nearest-neighbour Manhattan distance was calculated as:

$$d_i = \min_{j \neq i} \sum_{m=1}^2 |z_{i,m} - z_{j,m}|$$

The spacing metric was then evaluated by:

$$S = \sqrt{\frac{\sum_{i=1}^N (d_i - \bar{d})^2}{N - 1}}$$

where a lower value indicates more uniform spacing.

Hypervolume was calculated as the region dominated by the Pareto front and bounded by the common normalized reference point

$$\mathbf{r} = [1.10, 1.10]$$

The normalized hypervolume was calculated as:

$$HV_n = \frac{HV}{r_1 r_2}$$

where a higher value indicates improved convergence and objective-space coverage.

Common minimum and maximum values obtained from the combined MOPSO and NSGA-II nondominated solutions were used to normalize the objectives. Energy consumption was normalized as

$$z_E = \frac{E - E_{\min}^{\text{common}}}{E_{\max}^{\text{common}} - E_{\min}^{\text{common}}},$$

whereas adsorption capacity was transformed into a minimization objective as

$$z_{AC} = 1 - \frac{AC - AC_{\min}^{\text{common}}}{AC_{\max}^{\text{common}} - AC_{\min}^{\text{common}}}.$$

Both normalized objectives were dimensionless and bounded between 0 and 1, with the ideal point at [0,0]. The reference point [1.10, 1.10] was outside and inferior to all normalized Pareto solutions. Applying the same normalization limits and reference point to both algorithms ensured that the Hypervolume comparison reflected convergence and objective-space coverage rather than differences in objective scale.

Pure Diversity used the L_p dissimilarity:

$$D_p(z_i, z_j) = \left[\sum_{m=1}^2 |z_{i,m} - z_{j,m}|^p \right]^{1/p}$$

with $p = 0.1$. The Pure Diversity value was obtained by accumulating the dissimilarity contributions selected through the greedy diversity procedure. Higher Pure Diversity indicates a broader and more dissimilar solution set [21, 22].

For a valid comparison between MOPSO and NSGA-II, the same normalization limits, hypervolume reference point, and Pure Diversity exponent were applied to both algorithms. The algorithms were compared using their extreme and knee solutions together with spacing, hypervolume, normalized hypervolume, and Pure Diversity.

4 Results and discussion

This section evaluates the predictive performance of the machine-learning models and discusses the operating conditions obtained from single-objective and multi-objective optimization. Gaussian process regression (GPR), regression tree, and support vector machine (SVM) models were first compared using cross-validated predictions. The best-performing surrogate was then coupled with genetic algorithm (GA), particle swarm optimization (PSO), multi-objective particle swarm optimization (MOPSO), and the non-dominated sorting genetic algorithm II (NSGA-II). The discussion focuses on prediction accuracy, the physical interpretation of the response surfaces, the energy requirement of the optimized conditions, and the trade-off between adsorption capacity and electrical energy consumption.

4.1 Machine-learning model development and validation

4.1.1 Comparative predictive performance

The predictive capability of the three surrogate models was evaluated using the coefficient of determination (R^2), root mean squared error (RMSE), and mean absolute error (MAE), which collectively assess model accuracy,

prediction error, and generalization performance. Evaluating multiple performance metrics provides a more comprehensive assessment than relying on a single statistical indicator, particularly for nonlinear adsorption systems where local prediction errors may influence subsequent optimization. These indicators were calculated from the cross-validated predictions as

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2}$$

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2}$$

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i|$$

where y_i is the experimental adsorption capacity, $\hat{y}_i$ is the cross-validated estimate, $\bar{y}$ is the mean experimental value, and N is the number of observations.

Table 7 summarizes the predictive performance of the evaluated machine learning models.

GPR achieved the highest and most consistent predictive performance, with an average R^2 of 0.9985 ± 0.0006 , RMSE of 0.0246 ± 0.0044 mmol g⁻¹, and MAE of 0.0115 ± 0.0018 mmol g⁻¹. Its near-zero MBE also indicates negligible systematic prediction bias. SVM showed intermediate performance, whereas the Regression Tree produced the largest errors and the greatest fold-to-fold variation. These results confirm that GPR provided the most accurate and stable predictions across the five validation folds. These results indicate that GPR successfully captured the highly nonlinear relationship between adsorption capacity and the combined effects of adsorbent preparation and operating conditions. Furthermore, the regression equation $\hat{y} = 0.994y + 0.0079$ exhibited a slope approaching unity and an intercept close to 0, confirming excellent agreement between predicted and experimental values.

In contrast, the Regression Tree model produced the lowest predictive performance, with an R^2 of 0.8098 ± 0.0833 , indicating limited capability to represent the complex nonlinear behaviour of the adsorption system. Although tree-based models are computationally efficient and easily interpretable, their piecewise partitioning strategy is often insufficient for describing smooth nonlinear relationships involving multiple interacting variables. The SVM model demonstrated intermediate predictive capability ($R^2 = 0.9294 \pm 0.0402$), suggesting that although nonlinear regression improved prediction accuracy compared with the Regression Tree model, it was less effective than GPR for the present dataset.

The superior performance of GPR can be attributed to its probabilistic framework and covariance-based learning strategy, which enable smooth interpolation between experimental observations while maintaining excellent predictive accuracy for relatively small datasets. Similar observations have been reported for adsorption and carbon capture applications, where GPR consistently demonstrated superior predictive capability over conventional regression methods owing to its ability to capture highly nonlinear process behaviour while maintaining robust generalization performance [23]. Similar observations have been reported for adsorption and other chemical engineering systems, where GPR has consistently outperformed conventional regression techniques because of its ability to capture highly nonlinear process behaviour while avoiding excessive model complexity.

Table 7 Predictive performance of the machine-learning models based on combined out-of-fold predictions from fivefold cross-validation

Model	R^2	RMSE (mmol g ⁻¹)	MAE (mmol g ⁻¹)	Regression equation
GPR	0.9985 ± 0.0006	0.0246 ± 0.0044	0.0115 ± 0.0018	$\hat{y} = 0.994y + 0.009$
Regression Tree	0.8098 ± 0.0833	0.2762 ± 0.0572	0.1807 ± 0.0289	$\hat{y} = 0.836y + 0.271$
SVM	0.9294 ± 0.0402	0.1685 ± 0.0559	0.0981 ± 0.0222	$\hat{y} = 0.814y + 0.315$

Fig. 1 Cross-validated regression plots for GPR, regression tree, and SVM. The dashed line represents the ideal relationship, whereas the solid line is the fitted regression line

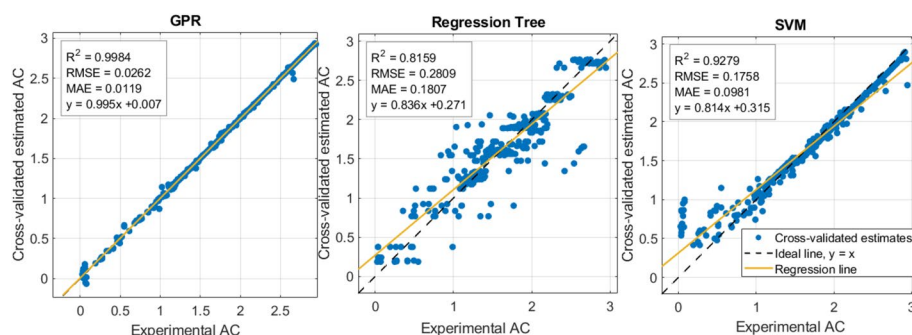
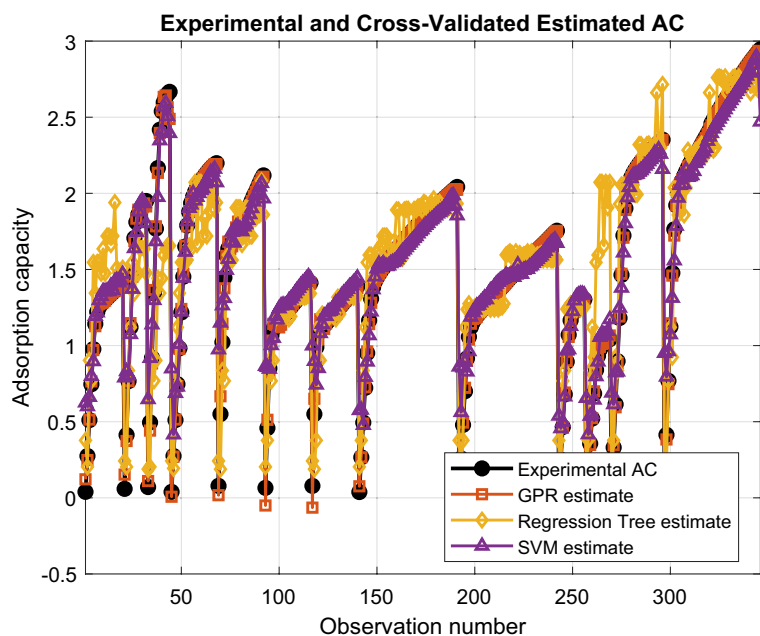


Fig. 2 Experimental adsorption capacity and cross-validated estimates generated by GPR, regression tree, and SVM



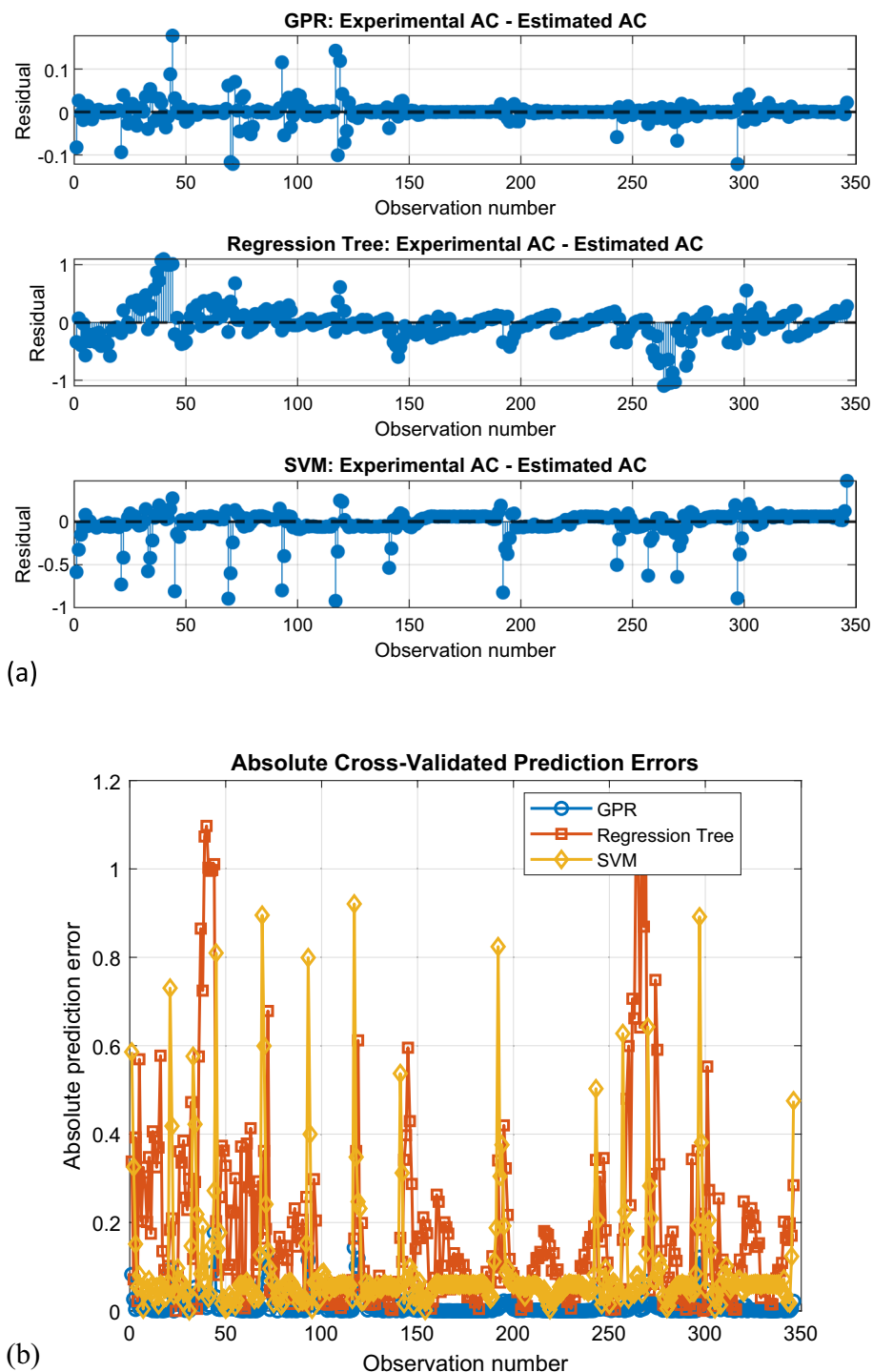
From an engineering perspective, the high predictive accuracy achieved by GPR is particularly important because optimization algorithms repeatedly evaluate the surrogate model during the search process. Prediction errors within the surrogate may propagate throughout optimization, potentially leading to unrealistic operating conditions. Consequently, selecting the most reliable predictive model is essential for ensuring that the optimized operating conditions remain physically meaningful and experimentally achievable (Fig. 1).

4.1.2 Cross-validated response and residual behaviour

Figure 2 compares the experimental adsorption capacities with the cross-validated estimates in observation order. GPR followed the rapid rises, declines, and repeated experimental patterns with little visible delay or amplitude distortion. The tree model reproduced the broad trend but missed several local maxima and minima. SVM tracked the overall pattern more effectively than the tree, although larger deviations occurred in several low-capacity and transition regions.

The residual and absolute-error plots in Fig. 3 provide a clearer assessment of local prediction behaviour. GPR residuals were distributed narrowly around zero, with most errors remaining within approximately ± 0.05 mmol g⁻¹ and only a small number of larger deviations. The absence of a strong trend in these residuals indicates that the GPR model captured the dominant nonlinear structure of the data. The regression tree exhibited the widest residual band and several systematic clusters, confirming that the model underfit portions of the response. SVM

Fig. 3 **a** Cross-validated residuals for GPR, regression tree, and SVM; and **b** absolute cross-validated prediction errors



gave lower residuals than the tree but retained several large negative errors, indicating underprediction at specific observations.

A few GPR estimates close to zero were slightly negative. Although these deviations were small, negative adsorption capacity is not physically meaningful. For future implementation, physical consistency could be improved by constraining the surrogate output to be non-negative, applying a suitable response transformation, or clipping only very small negative predictions after model validation. This limitation did not affect the optimization of maximum adsorption capacity because the optimizer operated in the high-response region.

GPR was selected as the surrogate model for optimization because it combined the highest accuracy with the most stable residual behaviour. Its near-unity regression slope and low cross-validated errors also reduced the risk that the optimization algorithms would exploit large surrogate-model bias. Nevertheless, optimized operating points located close to the experimental boundaries should still be confirmed experimentally because cross-validation assesses interpolation performance and does not guarantee reliable extrapolation outside the sampled domain.

Cross-validation was performed at the observation level because unique identifiers for individual experimental runs were not retained in the compiled dataset. Consequently, correlated time-resolved or replicate observations may have been distributed across different folds. The reported metrics should therefore be interpreted as interpolation performance within the sampled operating domain and may overestimate performance for entirely unseen fixed-bed runs. Future studies should retain run-level identifiers and apply grouped cross-validation to provide a stricter assessment of model generalization.

Figure 8 in Appendix 1 presents the first-order and total-order Sobol indices for adsorption capacity and energy consumption. The variable with the highest total-order index contributed the greatest proportion of output variance. A substantial difference between the total-order and first-order indices indicates that the variable also participated in important nonlinear interactions with the other operating conditions. For energy consumption, contact time, temperature, and flow rate are expected to contribute directly because they appear explicitly in the energy model. In contrast, the OPA-to-KOH ratio and inlet CO₂ concentration do not directly enter the energy equation, and their calculated energy sensitivity indices should therefore be close to zero. Their importance is instead reflected through their effects on the GPR-predicted adsorption capacity and the resulting multi-objective trade-off.

4.2 Single-objective optimization using GA and PSO

4.2.1 Optimum adsorption capacity and operating conditions

GA and PSO were independently applied to maximize the GPR-predicted adsorption capacity. Electrical energy was not part of the single-objective function; it was calculated after optimization at the maximum-capacity condition. The two algorithms converged to practically identical solutions, as summarized in Table 8. Both predicted a maximum adsorption capacity of 2.953786 mmol g⁻¹. The optimum occurred at an OPA-to-KOH ratio of approximately 0.50, a flow rate of 30 mL min⁻¹, a temperature of 30.44 °C, an initial concentration of 14.71 vol%, and a reaction time of 49.13 min.

The relative energy difference between the two solutions was calculated as

$$\Delta E = \frac{E_{\text{GA}} - E_{\text{PSO}}}{E_{\text{GA}}} \times 100 = 0.0032\%$$

Table 8 Comparison of the single-objective GA and PSO results for maximizing adsorption capacity

Performance indicator	GA	PSO	Preferred result
OPA-to-KOH ratio, <i>R</i>	0.499859	0.500000	Equivalent
Flow rate (mL min ⁻¹)	30.000703	30.000069	Equivalent
Temperature (°C)	30.436855	30.436269	Equivalent
Initial concentration (vol%)	14.707558	14.707691	Equivalent
Reaction time (min)	49.125268	49.125046	Equivalent
Maximum predicted adsorption capacity (mmol g ⁻¹)	2.953786	2.953786	Equivalent
Fixed equipment power (W)	62.000000	62.000000	Equivalent
Heater power (W)	65.621130	65.617617	PSO
Compression power (W)	0.046211	0.046210	PSO
Total average power (W)	127.667340	127.663826	PSO
Calculated energy consumption (kWh)	0.104528	0.104525	PSO

This difference is negligible and arises from small differences in flow rate, temperature, and reaction time. The close agreement between two distinct population-based algorithms indicates that the GPR response contains a well-defined maximum within the specified search domain. The convergence of independent optimization algorithms towards nearly identical operating conditions has similarly been reported for adsorption process optimization, indicating that the identified solutions are robust and unlikely to be artefacts of a particular optimization strategy [24]. The result also provides confidence that the solution was not strongly dependent on a particular optimization mechanism or random search trajectory.

4.2.2 Response-surface interpretation

Figures 4 and 5 show the GPR response surfaces evaluated around the GA and PSO solutions. The surfaces produced by both optimizers are effectively identical because their optimum coordinates are almost the same. For the OPA-to-KOH ratio and flow-rate surface, the highest adsorption capacity occurred at the lower end of both variables. Within the modelled range, increasing the OPA-to-KOH ratio beyond approximately 0.5 caused a pronounced reduction in predicted capacity. This response suggests that excessive KOH relative to the ash did not improve the adsorption response represented by the dataset. A lower flow rate was also favourable, which is consistent with a longer gas–solid contact time and a lower likelihood of premature breakthrough in the laboratory bed.

The response predicted by the GPR model is consistent with the physicochemical changes induced by KOH activation of the adsorbent. At relatively low KOH loading, chemical activation promotes pore development and introduces additional oxygen-containing surface functional groups, resulting in a larger accessible surface area and more favourable adsorption sites for CO₂. Consequently, adsorption capacity increases as the pore structure becomes more accessible to the gas molecules. However, further increases in the KOH loading do not necessarily improve adsorption performance. Excessive impregnation may widen or partially collapse the microporous structure, while residual activating species can obstruct micropore entrances, reducing the effective BET surface area available for adsorption. These competing effects explain the reduction in adsorption capacity predicted at higher OPA-to-KOH ratios. Therefore, the nonlinear trend captured by the GPR model is consistent with the experimentally observed relationship between chemical activation, pore development, and CO₂ adsorption performance reported in our previous study.

The temperature-concentration surface showed that high initial CO₂ concentration and low temperature favoured adsorption. The positive effect of concentration can be attributed to an increased gas-phase driving force, whereas the preference for lower temperature is consistent with adsorption that becomes less favourable as thermal energy increases. This behaviour agrees with adsorption theory, where increasing the inlet CO₂ concentration enhances the concentration gradient between the bulk gas phase and the adsorbent surface, thereby increasing the driving force for mass transfer and adsorption until adsorption sites become progressively occupied [25]. This interpretation should be treated as model-based evidence rather than a complete mechanistic conclusion because the current analysis did not independently estimate thermodynamic parameters.

The time-response curves increased rapidly during the early stage, followed by a slower rise toward the predicted maximum. This behaviour is consistent with rapid occupation of readily accessible surface sites followed by slower transport to less accessible pores or progressive site saturation. The optimizer selected a reaction time close to the upper experimental limit because the single-objective formulation rewarded every additional increase in predicted capacity, regardless of the associated energy and processing-time penalty.

Although the single-objective result defines the highest predicted adsorption capacity, it required approximately 0.1045 kWh. This value is substantially higher than the energy levels observed along the low-energy Pareto fronts discussed in Sect. 4.3. The SOO solution is therefore most appropriate when maximum adsorption capacity is the only decision criterion. For practical operation, the high reaction time and corresponding energy requirement justify the use of a multi-objective formulation.

Fig. 4 GPR response surfaces and time response at the GA optimum: **a** OPA-to-KOH ratio and flow rate; **b** temperature and initial concentration; and **c** reaction time

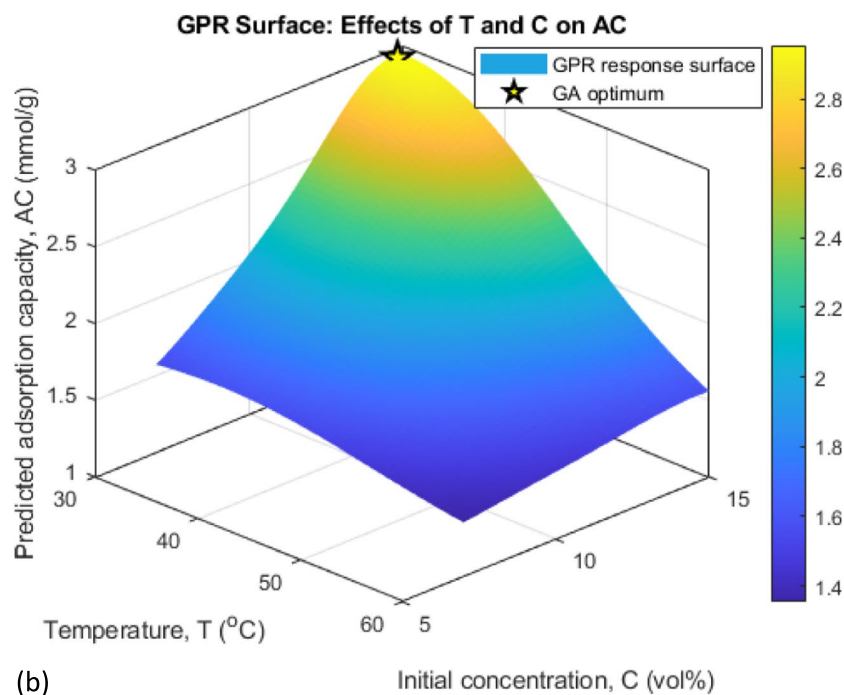
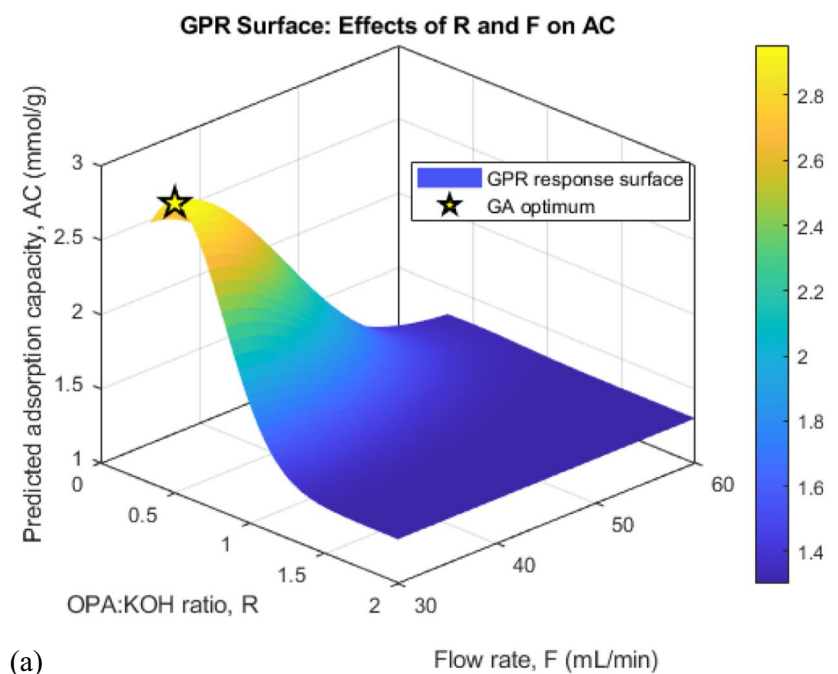
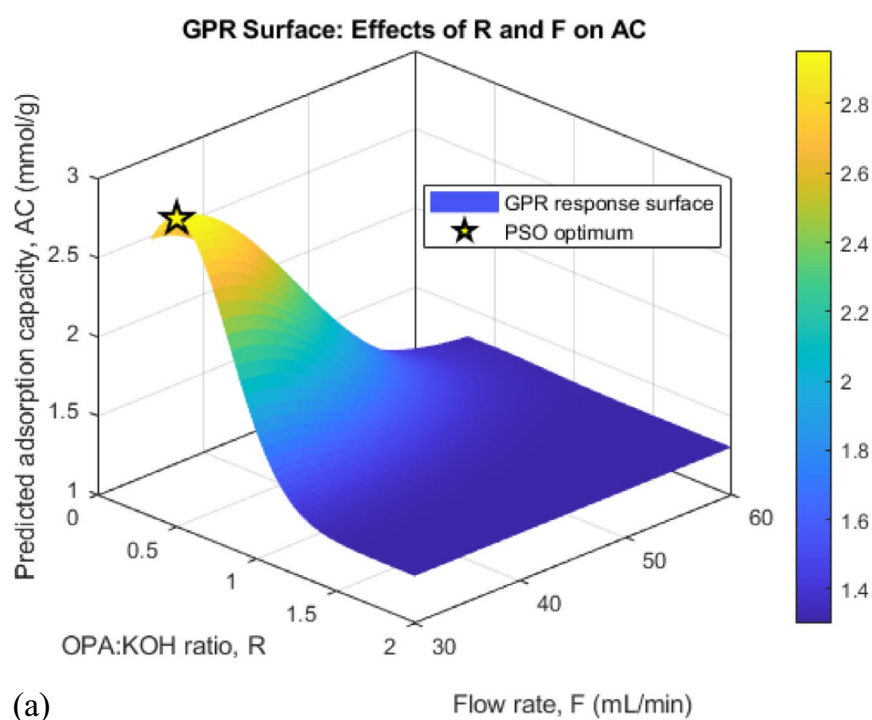
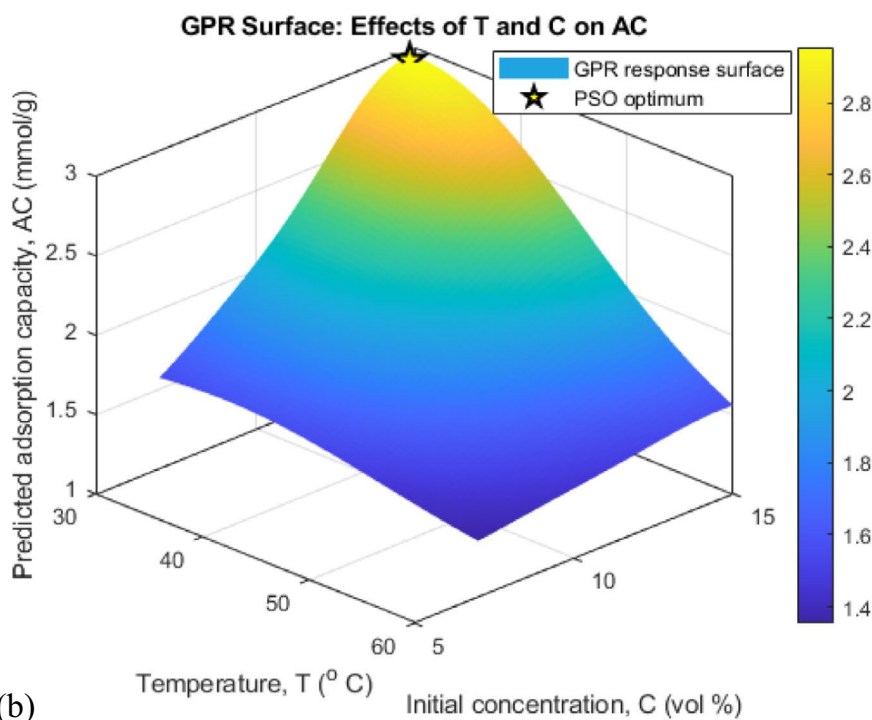


Fig. 5 GPR response surfaces and time response at the PSO optimum: **a** OPA-to-KOH ratio and flow rate; **b** temperature and initial concentration; and **c** reaction time



(a)



(b)

4.3 Multi-objective optimization using NSGA-II and MOPSO

4.3.1 Pareto-front characteristics

NSGA-II and MOPSO were used to minimize electrical energy while maximizing predicted adsorption capacity. Both algorithms generated smooth Pareto fronts, as shown in Fig. 6. The fronts demonstrate a clear conflict between the objectives: increasing adsorption capacity required longer operating time and, consequently, greater energy consumption. The initial region of each front was relatively steep, indicating that substantial improvements in adsorption capacity could be obtained with moderate increases in energy. At higher adsorption capacities, the curves gradually flattened, showing diminishing returns as the systems approached the high-capacity region.

The Pareto front demonstrates the inherent trade-off between maximizing CO₂ adsorption capacity and minimizing electrical energy consumption. No single operating condition simultaneously optimizes both objectives; therefore, the Pareto front provides a set of non-dominated solutions that represent different engineering compromises. Similar behaviour has been reported in multi-objective optimization studies of environmental and chemical engineering systems, where Pareto analysis supports systematic decision-making without requiring predefined weighting factors. Consequently, the final operating point can be selected according to practical priorities such as maximizing adsorption performance, minimizing energy consumption, or achieving a balanced compromise represented by the knee point. This interpretation is consistent with previous studies that employed Pareto-based optimization for sustainable process design and environmental management [26].

For the selection of a compromise solution, the objectives were represented in normalized minimization form as:

$$z_i = \left[\frac{E_i - E_{\min}}{E_{\max} - E_{\min}}, 1 - \frac{AC_i - AC_{\min}}{AC_{\max} - AC_{\min}} \right]$$

and the knee solution was obtained from:

$$i_{\text{knee}} = \underset{i}{\operatorname{argmin}} \sqrt{z_{i,1}^2 + z_{i,2}^2}$$

This criterion selected the point closest to the ideal combination of minimum energy and maximum adsorption capacity without assigning subjective weights to the objectives.

4.3.2 Representative Pareto solutions

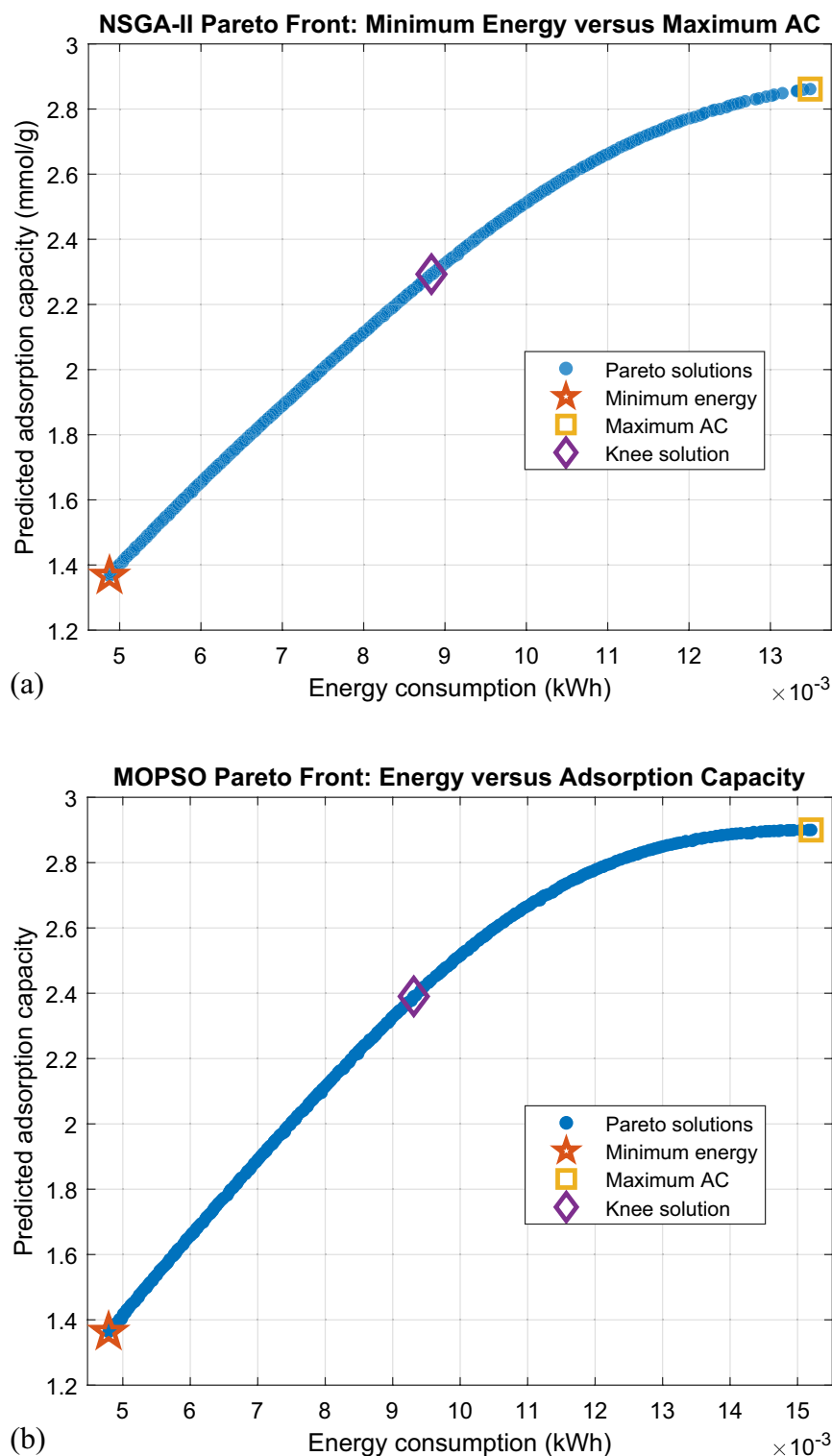
Table 9 compares the minimum-energy, maximum-capacity, and knee solutions. MOPSO achieved the lowest minimum energy, 0.004790 kWh, which was 1.76% below the NSGA-II value. At this point, however, NSGA-II produced a marginally higher adsorption capacity, 1.3663 compared with 1.3627 mmol g⁻¹. This difference was only approximately 0.27%, indicating that the low-energy endpoints were practically similar.

At the maximum-capacity endpoint, MOPSO predicted 2.9001 mmol g⁻¹, approximately 1.36% higher than NSGA-II. This improvement required 0.015202 kWh, which was 12.70% greater than the NSGA-II requirement. The choice between these endpoints therefore depends on whether the additional adsorption capacity justifies the higher energy demand.

The knee solutions provided a more balanced comparison. MOPSO achieved 2.3905 mmol g⁻¹, 4.25% above the NSGA-II knee capacity, while requiring 5.42% more energy. The NSGA-II knee point is preferable when energy minimization is given greater priority, whereas the MOPSO knee point offers a higher capture response for a modest energy increase.

Relative to the PSO single-objective maximum, the MOPSO maximum-capacity solution reduced energy consumption by approximately 85.46% while sacrificing only 1.82% of adsorption capacity. The NSGA-II

Fig. 6 **a** NSGA-II and **b** MOPSO Pareto fronts for energy consumption and predicted adsorption capacity



maximum-capacity point reduced energy by approximately 87.09% with a 3.13% reduction in capacity. These results demonstrate the practical value of multi-objective optimization: a large reduction in electrical energy was achieved at the expense of only a small loss in maximum predicted capacity.

The energy-saving percentage was calculated using:

Table 9 Representative Pareto solutions obtained using NSGA-II and MOPSO

Solution	Indicator	NSGA-II	MOPSO	Preferred result
Minimum energy	Energy consumption (kWh)	0.004876	0.004790	MOPSO
	Predicted adsorption capacity (mmol g ⁻¹)	1.3663	1.3627	NSGA-II
Maximum AC	Energy consumption (kWh)	0.013490	0.015202	NSGA-II
	Predicted adsorption capacity (mmol g ⁻¹)	2.8613	2.9001	MOPSO
Knee	Energy consumption (kWh)	0.008834	0.009312	NSGA-II
	Predicted adsorption capacity (mmol g ⁻¹)	2.2930	2.3905	MOPSO
Knee	OPA-to-KOH ratio, R	0.6821	0.6732	–
	Flow rate (mL min ⁻¹)	30.5169	30.0000	–
	Temperature (°C)	30.0049	30.0004	–
	Initial concentration (vol%)	9.8700	9.8338	–
	Reaction time (min)	4.2375	4.4683	NSGA-II

$$\eta_E = \frac{E_{\text{SOO}} - E_{\text{MOO}}}{E_{\text{SOO}}} \times 100$$

4.3.3 Effects of decision variables along the Pareto front

Figure 7 illustrates the decision-variable distributions along the MOPSO Pareto set. Reaction time showed the clearest relationship with both objectives. Energy increased almost linearly with time because the energy model multiplies average power by operating duration. Adsorption capacity also increased monotonically with time but gradually approached a plateau. The combined behaviour explains the curvature of the Pareto front and confirms that reaction time was the dominant variable governing the energy-capacity trade-off. Trade-offs between adsorption performance and energy consumption are characteristic of adsorption process design because improvements in separation efficiency frequently require longer operating times or increased energy input. Consequently, Pareto-based optimization provides a practical framework for identifying operating conditions that balance competing engineering objectives without introducing subjective weighting factors [27].

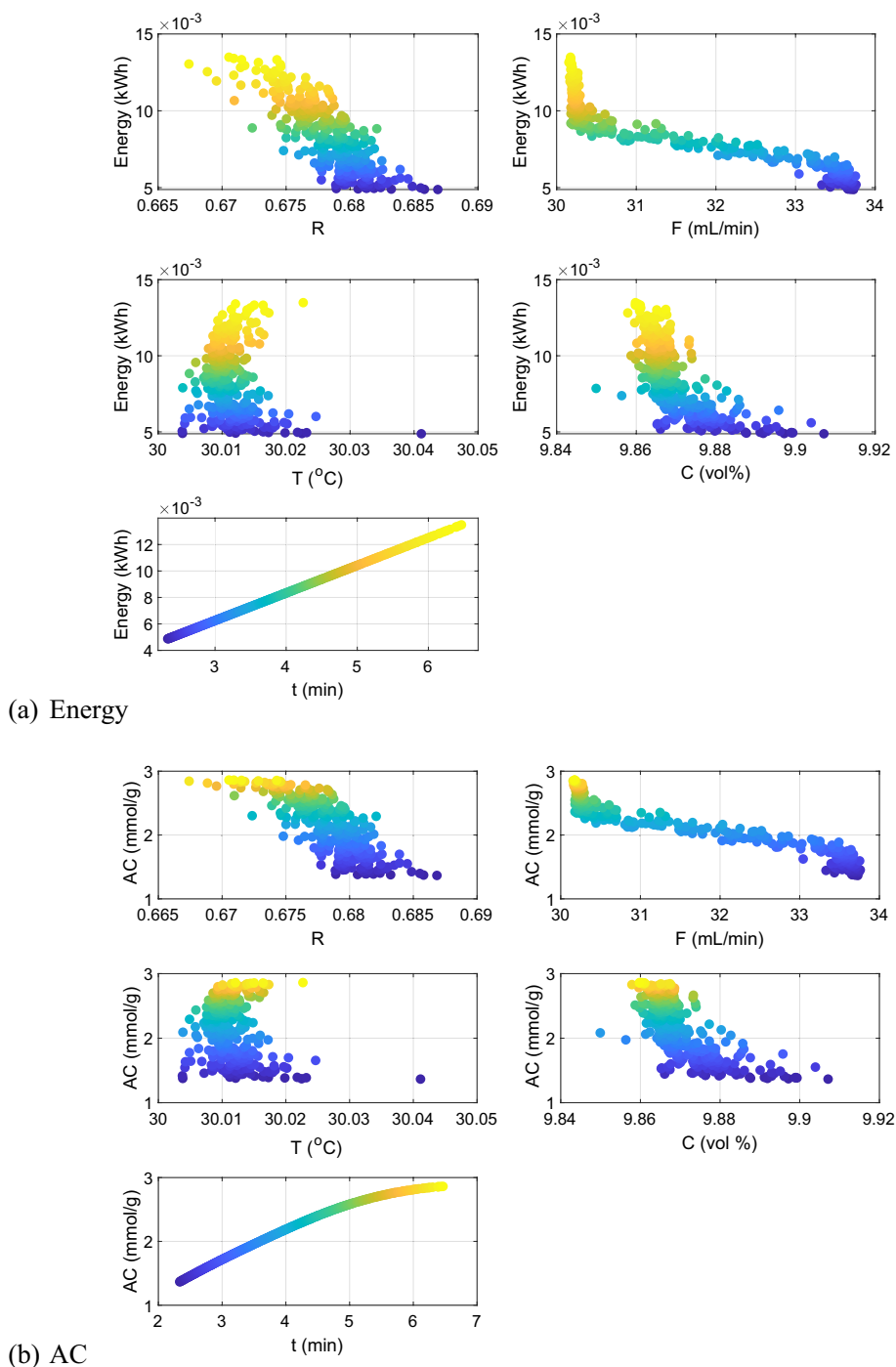
Temperature remained close to its lower bound for most Pareto solutions because a low set point reduced heater power and was also favourable for predicted adsorption. Flow rate also remained near the lower part of its range at the higher-capacity end, consistent with improved contact between the gas and adsorbent. The ratio and initial concentration varied across the Pareto set because they affected the GPR-predicted capacity but did not enter the electrical energy model directly.

The scatter plots in Fig. 7 should not be interpreted as isolated one-factor sensitivity curves. Every point represents a simultaneous combination of five decision variables selected by the optimizer. Apparent relationships between energy and variables such as R or C may therefore arise from correlations among Pareto-optimal operating conditions rather than a direct contribution to the energy equation. Reaction time and temperature have direct effects on energy, whereas R and C influence energy only indirectly through their interaction with the adsorption-capacity objective.

The decline in predicted adsorption capacity at higher KOH loading can be explained by the competing effects of chemical activation and excessive impregnation. At low to moderate loading, KOH promotes carbon etching, removal of volatile components, and development of accessible micropores. It may also increase the availability of basic oxygen-containing surface groups and, where present, nitrogen-containing functionalities that enhance interactions with acidic CO₂ molecules. These effects improve both the number and accessibility of adsorption sites [28].

Beyond the optimum impregnation level, however, excessive KOH can over-activate the carbon structure. Severe carbon etching may widen or collapse micropores, while residual potassium-containing species may block pore entrances and restrict gas diffusion. The resulting reduction in accessible micropore volume and effective BET surface area decreases the number of adsorption sites available to CO₂. Therefore, the lower adsorption capacity predicted by the GPR model at high KOH loading is consistent with pore blockage and structural

Fig. 7 **a** Relationships between the decision variables and energy consumption; and **b** relationships between the decision variables and predicted adsorption capacity along the MOPSO Pareto front



deterioration. This interpretation should be regarded as a physically plausible explanation of the modelled trend because detailed BET and surface-chemistry measurements were not available for every OPA-to-KOH ratio [1, 29].

4.3.4 Pareto-front quality and algorithm comparison

The Pareto-front quality indicators are summarized in Table 10. MOPSO produced a lower spacing value than NSGA-II, indicating a more uniform distribution of neighbouring Pareto solutions. Its hypervolume and

Table 10 Comparison of the multi-objective quality metrics

Metric	Desired direction	NSGA-II	MOPSO	Preferred result
Spacing	Lower	0.002080	0.001883	MOPSO
Hypervolume	Higher	0.819267	0.867420	MOPSO
Normalized hypervolume	Higher	0.677081	0.716876	MOPSO
Pure diversity	Higher	1631.9219	2821.1378	MOPSO
Mean pure diversity contribution	Higher	4.8139	9.4352	MOPSO
Hypervolume reference point	Same setting	[1.10, 1.10]	[1.10, 1.10]	Same
Pure diversity exponent, p	Same setting	0.10	0.10	Same

normalized hypervolume were also higher by approximately 5.88%, showing better combined coverage and convergence within the evaluated normalized objective space. Pure Diversity was 72.87% higher for MOPSO, while its mean Pure Diversity contribution was approximately 96.0% higher. These indicators collectively suggest that MOPSO explored the trade-off surface more broadly and uniformly in the reported runs.

The metric comparison requires careful interpretation. Hypervolume values are directly comparable only when both Pareto fronts are normalized using the same fixed energy and adsorption-capacity limits and evaluated against the same reference point. If each front is normalized using its own minimum and maximum values, the reported values describe within-run front quality but do not provide a fully controlled cross-algorithm comparison. Pure Diversity can also increase with the number of retained solutions; therefore, the mean contribution and a common number of uniformly resampled Pareto points should be considered when making final statistical comparisons.

A second point concerns the consistency of the Pareto endpoints. The single-objective algorithms identified a maximum predicted capacity of $2.9538 \text{ mmol g}^{-1}$ at approximately 0.1045 kWh , whereas the reported MOO fronts terminated at $2.8613\text{--}2.9001 \text{ mmol g}^{-1}$ and $0.0135\text{--}0.0152 \text{ kWh}$. Thus, the MOO results represent a highly energy-efficient segment of the trade-off rather than the full high-energy endpoint identified by SOO. Before final publication, endpoint completeness should be verified by using identical bounds, seeding the SOO maximum-capacity solution into the initial population or archive, and confirming whether it remains non-dominated. This check would distinguish an intentionally restricted practical energy region from incomplete exploration of the full Pareto front.

Within the reported low-energy region, MOPSO provided the stronger overall front quality and the higher-capacity knee point. NSGA-II nevertheless identified lower-energy solutions at the knee and maximum-capacity representative points. Consequently, MOPSO is preferable when broad Pareto coverage and adsorption performance are emphasized, whereas NSGA-II offers attractive alternatives when a stricter energy preference is imposed. In practical implementation, the final operating point should be selected from the Pareto set according to an explicit energy budget or minimum adsorption-capacity requirement and subsequently verified experimentally. The integration of machine learning with surrogate-based optimization is increasingly recognised as an effective strategy for accelerating process development, reducing computational cost, and supporting sustainable process design across a wide range of carbon capture technologies [30–33].

5 Conclusion

This study established an integrated machine learning and optimization framework for identifying effective operating conditions for CO_2 capture using KOH-functionalized oil palm ash. Among the evaluated surrogate models, GPR showed the strongest predictive capability. Its cross-validated R^2 of 0.9984, RMSE of $0.0262 \text{ mmol g}^{-1}$, and MAE of $0.0119 \text{ mmol g}^{-1}$ demonstrated that it could accurately reproduce the nonlinear relationship between adsorption capacity and the OPA-to-KOH ratio, flow rate, temperature, inlet concentration, and contact time. The substantially larger prediction errors obtained with the regression tree and SVM confirmed that model selection was important before the surrogate was used for optimization.

The single-objective results showed that GA and PSO converged to essentially the same maximum-capacity condition. Both algorithms predicted a maximum adsorption capacity of $2.953786 \text{ mmol g}^{-1}$ at an OPA-to-KOH ratio of approximately 0.50, a flow rate of 30 mL min^{-1} , a temperature of about 30.44°C , an inlet concentration of approximately 14.71 vol%, and a contact time of about 49.13 min. The corresponding energy demand was approximately 0.1045 kWh. The agreement between the two algorithms increased confidence that the global region associated with maximum predicted adsorption capacity had been identified. Nevertheless, the long contact time required by this solution also showed that maximizing uptake alone does not necessarily produce an energy-efficient operating policy.

The multi-objective analysis provided a more informative representation of process performance. NSGA-II generated solutions ranging from 0.004876 kWh at $1.3663 \text{ mmol g}^{-1}$ to 0.013490 kWh at $2.8613 \text{ mmol g}^{-1}$. MOPSO extended the attainable range from 0.004790 kWh at $1.3627 \text{ mmol g}^{-1}$ to 0.015202 kWh at $2.9001 \text{ mmol g}^{-1}$. The NSGA-II knee solution required 0.008834 kWh and achieved $2.2930 \text{ mmol g}^{-1}$, whereas the MOPSO knee solution achieved a higher adsorption capacity of $2.3905 \text{ mmol g}^{-1}$ at a moderately higher energy demand of 0.009312 kWh. Thus, NSGA-II provided the lower-energy compromise, while MOPSO favored a higher-capacity compromise.

MOPSO also produced the stronger overall Pareto-front quality, with a lower spacing value of 0.001883 and higher hypervolume, normalized hypervolume, and Pure Diversity values of 0.867420, 0.716876, and 2821.1378, respectively. The corresponding NSGA-II values were 0.002080, 0.819267, 0.677081, and 1631.9219. These indicators suggest that MOPSO generated a more uniformly distributed and diverse set of trade-off solutions with broader objective-space coverage. However, the selection of an operating point should not be based on algorithmic quality metrics alone. The preferred solution depends on whether the main priority is minimum energy, maximum adsorption capacity, or a balanced compromise between them.

The proposed framework demonstrates that experimental adsorption data can be converted into a reliable decision-support tool by combining GPR with complementary single- and multi-objective optimizers. The analysis also confirms that contact time is a major contributor to electrical energy consumption, while low temperature and low flow rate are generally favorable for CO_2 uptake. Before practical implementation, the selected optimum and knee solutions should be verified experimentally. Future work should also incorporate uncertainty bounds, adsorbent regeneration, pressure drop, material-preparation energy, and repeated stochastic runs using common normalization limits for direct algorithm comparison. These extensions would support a more complete assessment of the technical and energy performance of KOH-functionalized OPA for scalable CO_2 capture.

Fig. 8 First-order and total-order Sobol sensitivity indices for the GPR-predicted adsorption capacity and calculated electrical energy consumption. Error bars represent approximate 95% bootstrap confidence intervals. Higher total-order indices indicate a greater contribution to output variance, including interaction effects

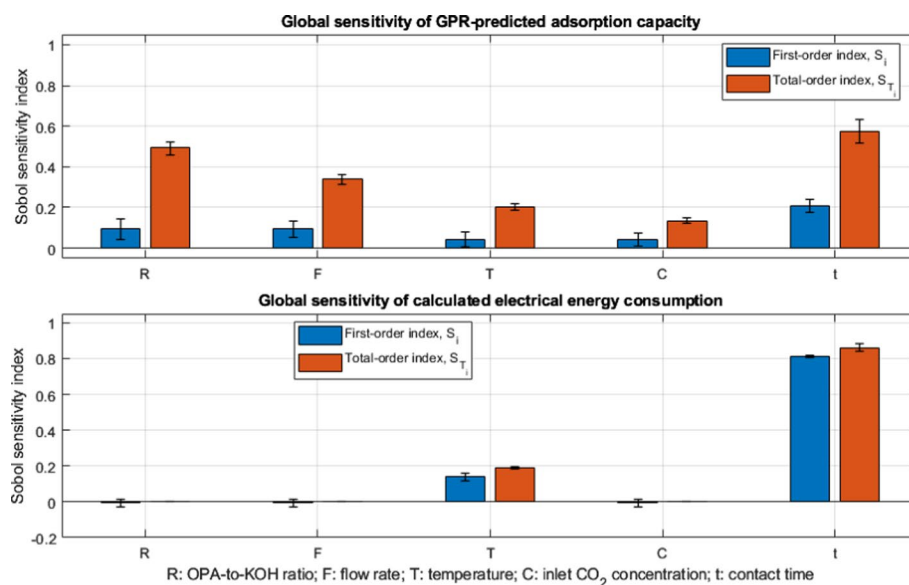


Table 11 Execution parameters used for the MOPSO and NSGA-II algorithms

Execution parameter	MOPSO	NSGA-II
MATLAB implementation	Custom MOPSO algorithm	MATLAB gamultiobj controlled-elitist multiobjective genetic algorithm
Population or swarm size	250 particles	400 individuals
Maximum iterations or generations	250 iterations	250 generations
Number of decision variables	5	5
Random-number seed	rng(1, 'twister')	rng(1, 'twister')
Initialisation method	Latin hypercube sampling	Latin hypercube sampling, with uniform random initialisation as a fallback
External archive size	300 nondominated solutions	Not applicable
Inertia weight, w	Linearly decreased from 0.90 to 0.40	Not applicable
Cognitive learning factor, c1	1.8	Not applicable
Social learning factor, c2	1.8	Not applicable
Velocity limit	20% of each decision-variable range	Not applicable
Crossover fraction	Not applicable	0.85
Crossover operator	Not applicable	MATLAB default gamultiobj crossover operator; not explicitly assigned in the code
Initial mutation probability	0.15	Not explicitly defined
Mutation-probability schedule	Linearly decreased from 0.15 to 0	Controlled internally by the adaptive feasible mutation operator
Initial mutation scale	0.10 of the selected variable range	Not explicitly defined
Mutation-scale schedule	Linearly decreased from 0.10 to 0	Adaptively determined by mutationadaptfeasible
Mutation operator	Gaussian perturbation of one randomly selected decision variable	Adaptive feasible mutation, mutationadaptfeasible
Leader or parent selection	Binary tournament selection based on crowding distance	MATLAB controlled-elitist selection based on nondominated ranking and diversity preservation; selection function not explicitly assigned
Personal-best selection	Constraint-dominance comparison; random retention when neither solution dominates	Not applicable
Pareto-retention fraction	Not applicable	0.85
Diversity preservation	Crowding-distance leader selection and archive truncation	Nondominated ranking and diversity preservation
Constraint handling	Constraint-dominance method	Nonlinear inequality constraint handled by gamultiobj
Minimum adsorption-capacity constraint	40th percentile of the experimental adsorption-capacity data	40th percentile of the experimental adsorption-capacity data
Function tolerance	Not explicitly applied	1×10^{-10}
Maximum stall generations	Not applicable	80 generations
Vectorised evaluation	No	No
Primary stopping condition	Completion of 250 iterations	Completion of 250 generations, attainment of function tolerance, or 80 stall generations

Appendix 1

See Fig. 8.

Appendix 2

See Table 11.

Parameters reported as “not explicitly defined” were not directly assigned in the corresponding MATLAB code and were managed internally by the solver. Inertia weight and learning factors are specific to the particle-swarm formulation and are therefore not applicable to NSGA-II.

Author contributions Fakhrony Sholahudin Rohman: Writing—original draft, Methodology, Formal Analysis; Muhammad Yusri Ahmad: Writing- review & editing, Conceptualization; Dinie Muhammad: Writing- review & editing, Formal Analysis; Dipesh Shikchand Patle: Writing- review & editing, Formal Analysis; Ashraf Azmi: Writing- review & editing, Conceptualization; Syamima Nasrin Mohamed Saleh: Writing- review & editing, Conceptualization.

Funding No funding.

Data availability Data was provided within the manuscript.

Generative AI and AI-assisted technologies in the manuscript preparation process During the preparation of this manuscript, the authors used ChatGPT to improve language clarity, refine sentence structure, and enhance overall readability. All generated content was carefully reviewed, revised, and validated by the authors, who assume full responsibility for the accuracy and integrity of the final published work.

Declarations

Ethics approval and consent to participate Not applicable. This study did not involve human participants or animals and therefore did not require ethical approval. Not applicable. This study did not involve human participants.

Consent or publication Not applicable. The manuscript does not contain any individual person’s data in any form requiring consent for publication.

Competing interests The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

1. Saleh SNM, Rohman FS, Muhammad D, Hussin SM, Hameed BH, Leng CT, et al. Enhanced CO₂ capture using KOH-functionalized oil palm ash adsorbent: experimental and applied machine learning approach. *Carbon Res.* 2025;4:60. <https://doi.org/10.1007/s44246-025-00227-3>.
2. Abuelnoor N, AlHajaj A, Khaleel M, Vega LF, Abu-Zahra MRM. Activated carbons from biomass-based sources for CO₂ capture applications. *Chemosphere.* 2021;282:131111. <https://doi.org/10.1016/j.chemosphere.2021.131111>.
3. Naseer MN, Zaidi AA, Dutta K, Wahab YA, Jaafar J, Nusrat R, et al. Past, present and future of materials’ applications for CO₂ capture: a bibliometric analysis. *Energy Rep.* 2022;8:4252–64. <https://doi.org/10.1016/j.egyr.2022.02.301>.
4. Alabadi A, Razzaque S, Yang Y, Chen S, Tan B. Highly porous activated carbon materials from carbonized biomass with high CO₂ capturing capacity. *Chem Eng J.* 2015;281:606–12. <https://doi.org/10.1016/j.cej.2015.06.032>.
5. Gayathiri M, Pulingam T, Lee KT, Sudesh K. Activated carbon from biomass waste precursors: factors affecting production and adsorption mechanism. *Chemosphere.* 2022;294:133764. <https://doi.org/10.1016/j.chemosphere.2022.133764>.

6. Ghosh S, Barron AR. The effect of KOH concentration on chemical activation of porous carbon sorbents for carbon dioxide uptake and carbon dioxide–methane selectivity: the relative formation of micro- (<2 nm) versus meso- (>2 nm) porosity. *Sustain Energy Fuels*. 2017;1:806–13. <https://doi.org/10.1039/C6SE00102E>.
7. Aisheh YIA. Palm oil fuel ash as a sustainable supplementary cementitious material for concrete: a state-of-the-art review. *Case Stud Constr Mater*. 2023;18:e01770. <https://doi.org/10.1016/j.cscm.2022.e01770>.
8. Zainudin NF, Lee KT, Kamaruddin AH, Bhatia S, Mohamed AR. Study of adsorbent prepared from oil palm ash for flue gas desulfurization. *Sep Purif Technol*. 2005;45:50–60. <https://doi.org/10.1016/j.seppur.2005.02.008>.
9. Choi AES, Park HS. Fuzzy multi-objective optimization case study based on an anaerobic co-digestion process of food waste leachate and piggy wastewater. *J Environ Manag*. 2018;223:314–23. <https://doi.org/10.1016/j.jenvman.2018.06.009>.
10. Zhang J, Zhang X, Li X, Song Z, Shao J, Zhang S, et al. Prediction of CO₂ adsorption of biochar under KOH activation via machine learning. *Carbon Capture Sci Technol*. 2024;13:100309. <https://doi.org/10.1016/j.ccst.2024.100309>.
11. Rohman FS, Muhammad D, Alwi SRW, Kassim NFA. Application bootstrap machine learnings for modeling of sonocatalytic degradation of caffeine in wastewater treatment. In: Nasr M, Negm A, Peng L, editors. *Artificial intelligence applications for a sustainable environment. Green chemistry and sustainable technology*. Cham: Springer; 2025. https://doi.org/10.1007/978-3-031-91199-6_6.
12. Rohman FS, Muhammad D, Azudin NY, Murat MN, Lee CS, Termizi SNAA, et al. Machine learning assisted-barnacles mating optimization of non-enzymatic isoamyl acetate in a solvent-free system using miniaturized intensified reactor. *Discov Chem Eng*. 2026;6:22. <https://doi.org/10.1007/s43938-026-00123-1>.
13. Rohman FS, Muhammad D, Azudin NY, Azmi A, Sheng LC, Termizi SNAA, et al. Surrogate-particle swarm optimization of uncatalyzed esterification of acetic anhydride with isoamyl alcohol in a microreactor system in reducing energy. *Discov Chem Eng*. 2026;6:20. <https://doi.org/10.1007/s43938-026-00124-0>.
14. Rohman FS, Alwi SRW, Termizi SNAA, Muhammad D, Er HA, Azmi A. Real-time accelerated particle swarm optimization for minimizing energy costs with changing demands and fuel price in industrial cogeneration system. *Digit Chem Eng*. 2026;19:100309. <https://doi.org/10.1016/j.dche.2026.100309>.
15. Okwu MO, Tartibu LK. Particle swarm optimisation. In: *Metaheuristic optimization: nature-inspired algorithms swarm and computational intelligence, theory and applications*. Studies in computational intelligence, vol. 927. Cham: Springer; 2021. https://doi.org/10.1007/978-3-030-61111-8_2.
16. Nshimirimana R, Abraham A, Nothnagel G. A multi-objective particle swarm for constraint and unconstrained problems. *Neural Comput Appl*. 2021;33:11355–85. <https://doi.org/10.1007/s00521-020-05555-6>.
17. Rohman FS, Idris I, Muhammad D, Alwi SRW, Zahan KA, Murat MN, et al. Non-dominated sorting-based strategy for optimizing the mixture of initiators in polyethylene reactor. *Process Integr Optim Sustain*. 2023;7:1131–48. <https://doi.org/10.1007/s41660-023-00332-z>.
18. Vu NP, Vu DB, Dinh VT, Vu D, Huy TQ. Optimizing two-stage gear design using NSGA-II with MATLAB: multi-objective approach on mass and efficiency trade-off. *Eng Technol Appl Sci Res*. 2025;15(3):23586–91.
19. Rohman FS, Muhammad D, Zahan KA, Murat MN. Operation and design optimisation of industrial low-density polyethylene tubular reactor for multiple objectives using an evolutionary algorithm-based strategy. *Process Integr Optim Sustain*. 2023;7:655–72. <https://doi.org/10.1007/s41660-023-00308-z>.
20. Rohman FS, Aziz N. Performance metrics analysis of dynamic multi-objective optimization for energy consumption and productivity improvement in batch electrodialysis. *Chem Eng Commun*. 2021;208(4):517–29. <https://doi.org/10.1080/00986445.2019.1674817>.
21. Areola RI, Adebisi AA, Reddy DJ. Comparative benchmarking of multi-objective algorithms for renewable energy system design using Pareto front quality metrics. *Appl Sci*. 2026;16(8):3775. <https://doi.org/10.3390/app16083775>.
22. Kang S, Li K, Wang R. A survey on pareto front learning for multi-objective optimization. *J Membr Comput*. 2024;7:128–34. <https://doi.org/10.1007/s41965-024-00170-z>.
23. Du X, Khan MN, Thakur GC. Machine learning in carbon capture, utilization, storage, and transportation: a review of applications in greenhouse gas emissions reduction. *Processes*. 2025;13(4):1160. <https://doi.org/10.3390/pr13041160>.
24. Takura Y, Ravutla S, Kim J, Ikeda K, Kajiro H, Yajima T, et al. Surrogate model optimization of vacuum pressure swing adsorption using a flexible metal–organic framework with hysteretic sigmoidal isotherms. *Int J Greenh Gas Control*. 2024;138:104260. <https://doi.org/10.1016/j.ijggc.2024.104260>.
25. Henrtin A, Heymans N, Duprez ME, Mouchaham G, Serre C, Wong D, et al. Lab-scale pilot for CO₂ capture vacuum pressure swing adsorption using metal–organic frameworks. *Carbon Capture Sci Technol*. 2024;12:100224. <https://doi.org/10.1016/j.ccst.2024.100224>.
26. Ortenero JR, Choi ES. Electrocoagulation treatment of wastewater: a Pareto frontier identification based on the total dissolved solids and cost. *Chem Eng Trans*. 2022;94:781–6. <https://doi.org/10.3303/CET2294130>.
27. Beck J, Fiedrich D, Brandani S, Fraga S. Multi-objective optimisation using surrogate models for the design of VPSA systems. *Comput Chem Eng*. 2015;82:318–29. <https://doi.org/10.1016/j.compchemeng.2015.07.009>.
28. Serafin J, Dziejarski B, Orlando FCJ, Sreńscek-Nazzal J. Design of highly microporous activated carbons based on walnut shell biomass for H₂ and CO₂ storage. *Carbon*. 2022;201:633–47. <https://doi.org/10.1016/j.carbon.2022.09.013>.

29. Hong SH, Chung K, Bang G, Kim KM, Lee CH. Adsorption equilibria and kinetics of CO₂, CH₄, CO, N₂, and H₂ on KOH-treated activated carbon pellets up to 1000 kPa. *Chem Eng J*. 2021;431:133396. <https://doi.org/10.1016/j.cej.2021.133396>.
30. Wang Z, Yuan P, Yu H, Ma Q, Xu B, Zhao D. Carbon capture, utilization and storage: technology, application, and policy. *Processes*. 2025;13(11):3414. <https://doi.org/10.3390/pr13113414>.
31. He Z, Yen G, Ding J. Knee-based decision making and visualization in many-objective optimization. *IEEE Trans Evol Comput*. 2021;25:292–306. <https://doi.org/10.1109/tevc.2020.3027620>.
32. Sukpancharoen S, Prasartkaew B. Combined heat and power plant using a multi-objective Henry gas solubility optimization algorithm: a thermodynamic investigation of energy, exergy, and economic (3E) analysis. *Heliyon*. 2021;7(9):e08003.
33. Sukpancharoen S, Udol P, Supsomboon P, Srinophakun TR. Integrated multi-objective MINLP optimization for cost-effectiveness and safety in multi-floor process layout for manufacturing industries. *Comput Ind Eng*. 2025;209:111442. <https://doi.org/10.1016/j.cie.2025.111442>.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

**Fakhrony Sholahudin Rohman¹ · Muhammad Yusri Ahmad² · Dinie Muhammad³ ·
Dipesh Shikchand Patle⁴ · Ashraf Azmi⁵ · Syamima Nasrin Mohamed Saleh² **

✉ Syamima Nasrin Mohamed Saleh
mssnasrin@unimas.my

¹ Faculty of Engineering, Universiti Malaysia Sabah, 88400 Kota Kinabalu, Malaysia

² Department of Chemical Engineering and Energy Sustainability, Faculty of Engineering, Universiti Malaysia Sarawak, 94300 Kota Samarahan, Sarawak, Malaysia

³ Faculty of Chemical and Energy Engineering, Universiti Teknologi Malaysia (UTM), 81310 Johor Baharu, Johor, Malaysia

⁴ Department of Chemical Engineering, Motilal Nehru National Institute of Technology Allahabad, Prayagraj 211004, India

⁵ Faculty of Chemical Engineering, Universiti Teknologi MARA, Universiti Teknologi MARA, 40450 Shah Alam, Selangor, Malaysia