Chemical Engineering & Processing: Process Intensification 219 (2026) 110638

Contents lists available at ScienceDirect

Chemical Engineering and Processing - Process Intensification

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

A hybrid surrogate and simulation-based framework for efficient CapEx/OpEx optimization in complex chemical plants☆

Luis Felipe Sánchez[1], Marcello Maria Bozzini[1], Mattia Vallerio*, Flavio Manenti

Dipartimento di Chimica, Materiali e Ingegneria Chimica "Giulio Natta" Politecnico di Milano, Piazza Leonardo da Vinci, 32, Milan, 20133, Italy

ARTICLE INFO

Keywords:
CapEx
OpEx
Derivative-free optimization
Surrogate model

ABSTRACT

Process Intensification aims at the economic and operational efficiency of chemical processes by emphasizing energy integration, unit size reduction, and cost minimization. The optimality of intensified solutions is typically assessed using Process Simulators, especially for complex chemical processes. These tools offer limited reliability and flexibility for the optimization of Capital Expenditures (CapEx), thus restricting their scope to Operational Expenditures (OpEx). As an alternative, external software is commonly required for simulation-based (SIM-OPT) and surrogate-based (SUR-OPT) CapEx/OpEx optimization. This work introduces a framework to select the most efficient optimization methodology based on simulation computational complexity. In addition, it presents a novel methodology (MIX-OPT) providing an efficient trade-off between optimization speed and accuracy. These three approaches were employed to optimize a complex biogas-to-methanol plant. Results showed that SIM-OPT achieved the greatest reduction in the Payback Period (PBP) of the plant (9.28%) with highest computational demand (984 min), SUR-OPT had the shortest computational time (717 min) with moderate PBP reduction (7.89%), and MIX-OPT reached a compromise with a PBP reduction of 8.24% in 884 min. The proposed framework demonstrated that simple simulations benefit from SIM-OPT, complex ones from SUR-OPT, and a wide range of simulations in the middle from the novel MIX-OPT approach.

1. Introduction

Process simulators (PS) are key tools in the chemical engineering practice. They are powerful software tools that solve equation systems based on first principles and established scientific laws to mimic the behavior of real-scale processes [1]. Their accuracy and precision enable the extraction of high-quality data without appealing to real-life experimental campaigns [2]. This is particularly useful in the process design phase to identify operating conditions and plant configurations that can enhance economic performance. For instance, different process alternatives may be tested to improve the process profitability [3]. In the context of Process Intensification (PI), this may include aspects such as energy integration, unit size reduction, and cost minimization [4], which should be assessed in conjunction with economic indicators to ensure their suitability [5]

The profitability of a process depends on its Capital and Operational Expenditures (CapEx and OpEx). While the OpEx depends on well-defined operational variables obtained from material and energy balances, typically obtained from PS, the CapEx depends on several different factors [6]. The equipment cost is strongly related to its particular type, process conditions, and dimensions, which are user-defined on a case-by-case basis [7]. Consequently, PS often offers tools to perform OpEx cost minimization, while CapEx must be evaluated using more manual approaches, such as cost correlations or rules of thumb [8]. For this reason, the literature approaches dealing with simultaneous CapEx and OpEx optimization have typically been supported by external CapEx estimation and optimization methodologies.

Simulation-based process economic optimization (SIM-OPT) may be performed with derivative-based or derivative-free approaches. Although derivative-based methodologies are typically more efficient, derivative information is not usually available in most commercial simulation tools [9]. In addition, the iterative nature of numerical solutions in process simulators, or the presence of unit operations that solve Boundary Value Problems with a determined tolerance may be translated into noise, which affects the accurate estimation of the system's derivatives [10]. For this reason, SIM-OPT typically falls within the derivative-free optimization domain.

Considering the complexities associated with CapEx estimation in simulation environments, some authors have preferred to avoid their calculation or to address them separately. For example, Ramos et al.

☆ This article is part of a Special issue entitled: 'AI for PI' published in Chemical Engineering and Processing - Process Intensification.

* Corresponding author.
E-mail address: mattia.vallerio@polimi.it (M. Vallerio).

[1] All authors contributed equally.

https://doi.org/10.1016/j.cep.2025.110638
Received 7 April 2025; Received in revised form 15 November 2025; Accepted 23 November 2025
Available online 29 November 2025

explicitly avoided the evaluation of CapEx at each iteration in a Fisher–Tropsch synthesis process multiobjective optimization in Aspen Plus due to computational limitations [11]. Nevertheless, their study demonstrated the plant's profitability for a fixed biogas feed flow rate. In contrast, Frosi et al. [12] and Variny et al. [13] proposed energetic optimization and posterior CapEx evaluation using Aspen Plus for the optimization of an ethanol-to-ethylene and a cyclopentyl methyl ether plant, respectively. The analysis allowed the authors to assess the feasibility of constructing each of the evaluated plants. However, only minimizing OpEx does not necessarily lead to the economic optimum over the plant's lifetime. Other authors have implemented external software to evaluate and optimize the CapEx. For instance, Duanmu et al. [14] optimized a series of case studies analyzing general quaternary reactive systems in three configurations, including reactive distillation, pressure swing distillation, and pervaporation in gPROMS Process Simulator. The authors implemented an algorithm in C# to solve the MINLP problem, since they also calculated integer values such as stages of distillation. In addition, they calculated OpEx in the C# script and optimized both OpEx and CapEx with a genetic algorithm. Each optimization problem solved required up to 3 h on a dedicated high-end computer. The study determined that dividing wall systems and hybrid distillation-pervaporation operations saved up to 24% of energy and presented a reduced total cost compared to conventional reactive distillation alternatives. Nevertheless, the computational time needed to perform the optimization is not feasible with more complex process layouts and variables. Kazemi et al. [15] used a multiobjective genetic algorithm to minimize environmental impact and total cost for a bio-methanol production process integrated with a Rankine Cycle for waste heat recovery. They reduced the Global Warming Potential (GWP) by 71% compared to the initial case, while increasing the Internal Rate of Return (IRR) from 20 to 35%. The authors did not present the complexity of the approach and the computational effort to achieve such results.

Despite their accuracy, derivative-free optimization approaches require multiple iterations to map the system's response. This typically leads to high computational costs when solving the optimization problem for complex simulation systems with intricate thermodynamics, kinetics, or interactions between unit operations. In fact, these complex systems and related simulations are common in chemical engineering practice [2]. As an alternative to reduce computational times, surrogate models have been widely used in recent years to approximate the evaluation of computationally expensive systems arising in this context [2,16].

Surrogate-based optimization (SUR-OPT) techniques to solve bound-constrained mixed-integer derivative-free problems present an alternative to quickly generate near-optimum solutions without evaluating the original function or its derivatives [16]. These methods have recently been used to optimize the economic performance of processes such as methanol synthesis [17] and CO_2 capture [18]. For instance, Kum et al. [18] implemented a Deep Neural Network trained with data from 486 Aspen Plus simulations to minimize the total annualized cost of a CO_2 capture process, using 5 optimization variables [18]. The high accuracy of the models, with relative errors around 2%, allowed the optimization and sensitivity analysis of the process. The authors compared the surrogate-based simulation results with the original simulation, obtaining an optimum cost with a relative error of 0.5%. However, considering the large number of simulations used, a direct derivative-free SIM-OPT may have been used instead, with potentially better results. Andreasen and van Baten trained a Kriging model using a dataset with data from 80 experiments of a methanol synthesis process simulated in COCO [17]. The authors determined the optimum pressure, temperature, catalyst volume, and purge ratio to minimize the Levelized Cost of Methanol (LCOM). The highly accurate surrogate, with relative errors lower than 1%, was used to optimize and compare the process with other techno-economic analyses available in the literature. However, the authors did not validate their optimum point found using the original simulation. Moreover, the high accuracy of the surrogate model may be related to the intrinsic simplicity of the simulation, which provided smooth and noiseless data. In addition, the approach excluded steam reforming for syngas production, a key step in methanol synthesis, that can significantly influence the overall economic performance of the process.

The following work presents a framework for the selection of the most efficient CapEx/OpEx optimization methodology among three approaches: (i) derivative-free simulation-based optimization (SIM-OPT), (ii) derivative-free surrogate-based optimization (SUR-OPT), and (iii) a novel mixed methodology (MIX-OPT) combining the advantages of SIM-OPT and surrogate-based sensitivity analysis. The principle behind MIX-OPT is the reduction of the optimization region using information from a surrogate model. This is expected to diminish the transient periods between successive simulations [19], increasing the optimization speed. The three optimization methodologies are tested in a complex process simulation developed in the design phase of a highly integrated methanol synthesis process, including the reforming section, a double-reactor configuration, several recycles, and the methanol distillation zone. The objective functions for optimization are widely used metrics such as the Payback Period (PBP), the Discounted Net Present Value (DNPV), and the Internal Rate of Return (IRR). The degrees of freedom for optimization considered design and operational parameters, such as reactor volumes and reactant flow rates.

2. Methodology

Fig. 1 presents the framework proposed in this work to select the most efficient methodology for the simultaneous CapEx/OpEx optimization in complex chemical simulations. The first step in the framework is the Simulation Sensitivity Analysis. It consists of a preliminary study to determine simulation noise. Although simulations are typically deterministic, simulation noise may be quantified as the difference in the simulation response obtained with the same input data. Simulation noise is typically caused by the required iterative calculations in process simulators with their associated error tolerances [10].

The computational time may be improved by relaxing the tolerances, while the accuracy may be increased by tightening them. Although no general rule exists regarding noise allowance, a typical rule of thumb is to consider 1% as the upper boundary, since higher allowances could become a significant problem for optimization and surrogate modeling [20,21].

The next step is to estimate the Average Simulation Time, which is directly related to the simulation complexity. Simpler simulations may benefit from simulation-based (SIM-OPT) approaches, while more complex ones may require the training of a Surrogate Model to handle the optimization. Indeed, the time threshold to define a simulation as complex is studied in this work. This has been addressed by assuming different simulation times for the same case study to determine which optimization approach would prove more efficient.

Surrogate model training will be further discussed in Section 2.5. In case the trained surrogate model accurately represents the original simulation, it may be directly used for optimization in a SUR-OPT approach, similar to the work of Andreasen et al. [17]. The threshold for defining a surrogate model as accurate has been discussed by different authors [21]. Typically, a rule of thumb is setting an upper limit of 5% relative error near the optimum value for practical surrogate-guided optimization.

Complex simulations, on which a surrogate model has been trained but resulted in a limited accuracy, present a challenge for both SIM-OPT and SUR-OPT. For these particular cases, the MIX-OPT methodology has been proposed. It combines SIM-OPT with the trained surrogate model to provide an efficient trade-off between computational expense and optimization accuracy. In particular, the surrogate model is used in a sensitivity analysis to narrow the optimization region in a zone near the optimum. Then, the SIM-OPT approach is used

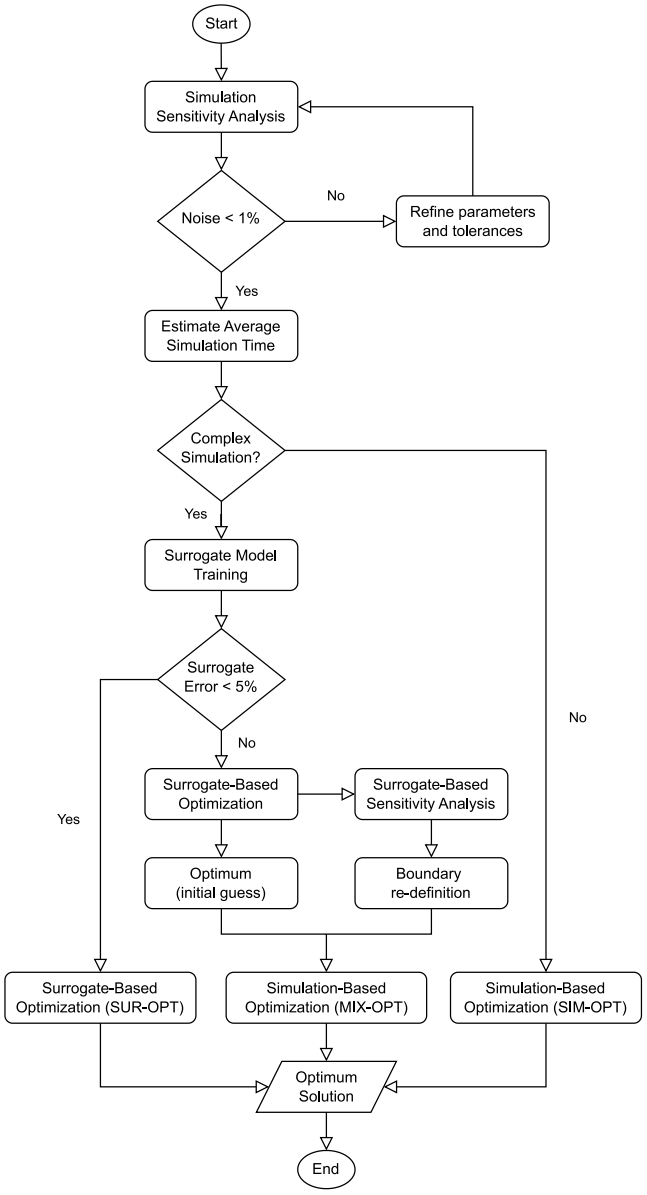


Fig. 1. Proposed framework to select the most efficient methodology for the simultaneous CapEx/OpEx optimization in complex chemical simulations.

in the reduced region, which is expected to shorten the computational time due to smaller steps in the optimization algorithm. This approach leverages a key characteristic of surrogate models: their ability to predict untested points through interpolation [2]. Unlike conventional hybrid surrogate-based optimization frameworks, where the surrogate replaces the process simulator within the optimization loop [22], the MIX-OPT approach applies the surrogate only to redefine the optimization domain before launching the rigorous simulation-based optimization. Consequently, this approach consists of a two-step sequential framework. This sequential strategy confines the search space around promising regions, leading to a quantifiable computational time advantage while fully preserving the accuracy of the rigorous process simulations.

The details of the methodology followed for each approach are presented in this section. Although computational time may not serve as a definitive performance or comparison benchmark, it is reported to provide a general sense of the computational effort involved. The different approaches were implemented using a computer equipped with an 11th Gen Intel Core i9-11900 processor at 2.50 GHz and 32 GB of RAM.

2.1. Case study

The case study used in this work corresponds to a methanol production process from biogas, given its relevance in recent literature due to its pivotal role in achieving net-zero targets [23]. The study is set in the design phase of a plant with a desired transforming capacity of 3600 kg/h of upgraded biogas (74/26 mol% methane/CO_2) into market-grade methanol. As a key consideration, the biogas feedstock must be pretreated to remove impurities and a portion of the carbon dioxide to prevent reformer catalyst poisoning. Moreover, the syngas produced in the reformer unit should ensure a Stoichiometric Number (SN, Eq. (1)) around 2, which is a standard Key Performance Indicator (KPI) of its quality [24].

$$SN = \frac{x_{H_2} - x_{CO_2}}{x_{CO_2} + x_{CO}} \tag{1}$$

The Block Flow Diagram (BFD) of the process is shown in Fig. 2. It presents the four key operations in the process: reforming of biogas to produce syngas, methanol synthesis from syngas, methanol purification, and the combustion of a mixture of methane and off-gases in a furnace to sustain the biogas reforming.

2.2. Process simulation

The process described in Section 2.1 was simulated using Aspen HYSYS V.14, a well-known commercial process simulation software for petrochemical operations [20]. The thermodynamic package chosen was the Soave–Redlich–Kwong equation of state since it is recommended for gas-processing and petrochemical systems [25]. The methanol synthesis was modeled using the isothermal Plug Flow Reactor tool in Aspen HYSYS, accounting for the pressure drops with the Ergun equation. Regarding the methanol reaction kinetics, the Vanden-Bussche and Froment scheme was implemented to describe the reactive stage within the tube bundle [26]. Side reactions were not considered in this approach, as the reactive mixture, derived from upgraded and reformed biogas, contains high concentrations of CO_2 (12 m/m%) [27]. As a consequence, the downstream separation was limited to methanol, water, and non-condensable gases. In contrast, the reformer and the furnace were modeled at thermodynamic equilibrium using the Gibbs Reactor, hence minimizing the Gibbs free energy of the mixture at the selected outlet temperature [28]. Furthermore, the pressure drop considered at each heat exchanger of the process was 10 kPa.

A detailed representation of the simulated reforming section is presented in Fig. 3 as a Process Flow Diagram (PFD). In this scheme, the upgraded biogas feedstock (Stream 1) is preheated using hot flue gases from the furnace. Then, the biogas stream is mixed with steam that has been previously produced in a boiler utilizing the hot flue gases exiting from the furnace. After the reforming operation at 950 °C, the produced syngas is cooled down to 40 °C before a water removal operation. The dry syngas is then sent to the methanol synthesis section, while the condensed water enters a flash operation to remove the solute gases. The water is mixed with a make-up (Stream 18) from the methanol purification section to ensure a hydrogen-to-carbon ratio of 3 in the reformer section, increasing syngas production [29]. However, before entering the reformer, water must be preheated (Stream 15) and vaporized (Stream 3).

Fig. 4 illustrates a detailed scheme of the synthesis section. The syngas from the reforming section (Stream 10) is compressed to 61 bar and mixed with the recycled unreacted gases (Stream 38). The reactive mixture (Stream 21) is preheated with the product stream from the first methanol reactor and further heated up to reach the reactor synthesis temperature, set at 250 °C. The product stream (Stream 25) is cooled down to 40 °C to condense methanol and water (Stream 27), which are then sent to the purification section. The gases are sent to a second

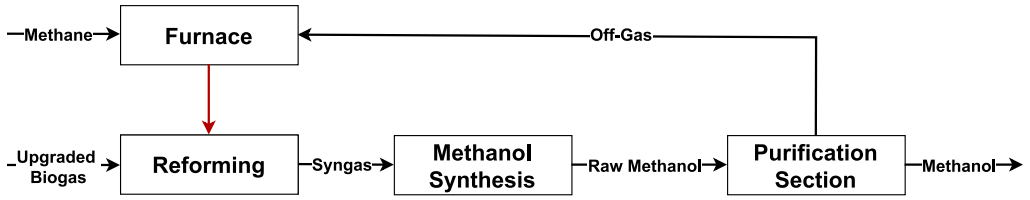


Fig. 2. Block Flow Diagram of the case study.

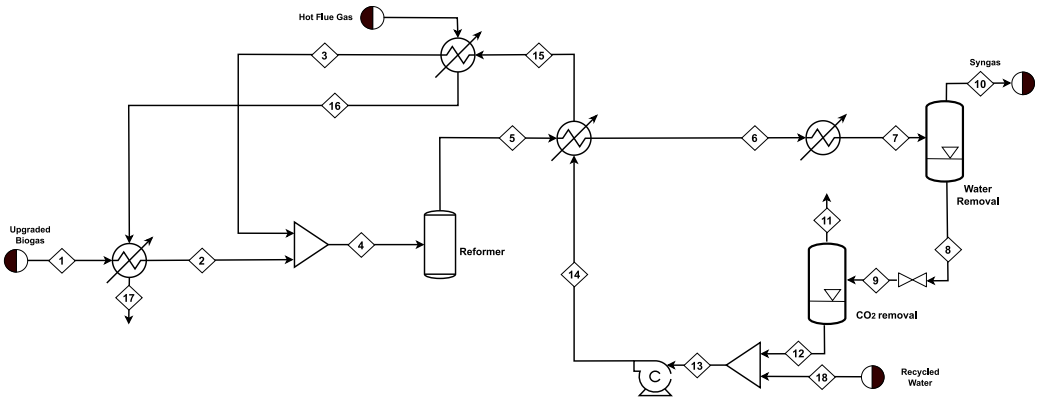


Fig. 3. Process Flow Diagram of the reforming section.

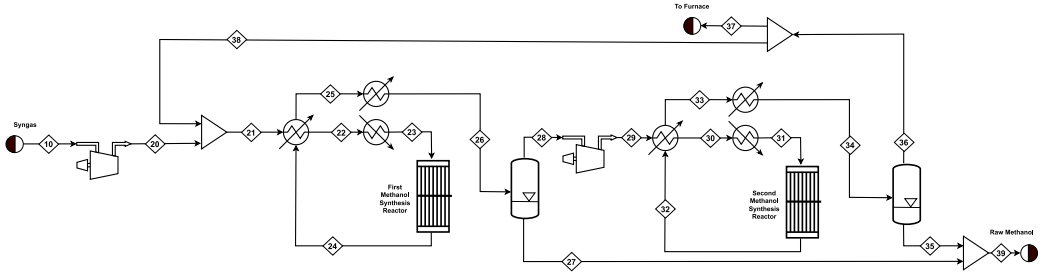


Fig. 4. Process Flow Diagram of the synthesis section.

analogous reactive stage. The unreacted gases are split into two: a 95% fraction recycled to the first reactor (Stream 38) and a 5% fraction sent to the furnace (Stream 37) for combustion [30].

The last two sections of the plant regarding methanol purification and furnace are shown in detail in Fig. 5. Raw methanol undergoes a two-stage separation in two dedicated distillation columns. The first column separates the remaining off-gases used for combustion (Stream 40). Conversely, the second column separates the methanol–water mixture into AA grade methanol (99.85 w/w% purity) in Stream 43 and water in Stream 18. The first column ensures the absence of gases and a 99% recovery of methanol at the bottom. The second column specifications are the methanol fraction and recovery at the top, both equal to 99.9%. This water stream is recycled to the reforming section. Regarding the furnace, part of the unreacted gases from the reaction and separation sections (Streams 37 and 40) are mixed with methane and burned to sustain the reformer reactor's energetic demand.

2.3. CapEx OpEx estimation

The economic assessment of the process was performed using the CapEx OpEx Robust Optimization (CORO) software [31]. CORO is a framework developed in C++ and VBA to estimate the capital and operational expenditures of a plant based on its Aspen HYSYS process simulation and the cost correlations of Turton et al. [8]. The software includes the customization of cost-related variables such as the cost of raw materials, utilities, and product streams for reliable economic estimation.

In this work, the methanol synthesis reactors were modeled as shell-and-tube heat exchangers with a reactive section within the tube bundle. With this configuration, the cost correlation of fixed tube heat exchangers was used [8]. Moreover, the materials chosen for the units were stainless steel in the presence of hydrogen, and carbon steel otherwise.

The utilities used for the economic analysis and their cost are presented in Table 1 [8]. In addition, the table presents the cost of methane

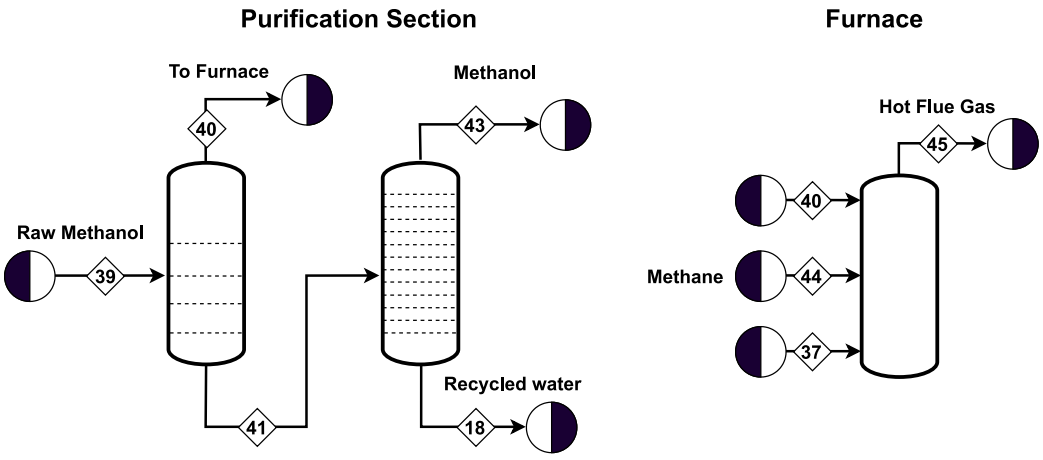


Fig. 5. Process Flow Diagram of the purification section and furnace unit.

Table 1
Raw material and utility cost.

Raw Material	Cost	UoM	Reference
Methane	0.7	€/kg	[25]
Upgraded biogas	0.2	€/kg	[25]
Utility	**Cost**	**UoM**	**Reference**
Electricity	0.1064	€/kWh	[8]
Cooling water	0.0148	USD/m3	[8]
Low Pressure Steam	0.024	USD/kg	[8]
High Pressure Steam	0.038	USD/kg	[8]

Table 2
Output variables.

Name	Description	UoM
PBP	Payback Period	y
DNPV	Discounted Net Present Value	USD
IRR	Internal Rate of Return	–

and upgraded biogas [25]. In this work, the methanol selling price was fixed at 0.5 USD/kg [30]. Low-pressure steam (5 bar) has been used for the reboilers of both distillation columns, while superheated high-pressure steam (40 bar, 350 °C inlet) is applied to heat the feed before the synthesis reactor. Cooling water is used in every cooler of the process simulation, assuming 30 °C as inlet and 40 °C as outlet temperatures, respectively. For simplicity, the heat generated from the methanol reaction has not been considered as a profit or cost. Heat recovery has been deliberately neglected in order to represent a worst-case scenario, avoiding assumptions on the possible use or economic value of the recovered heat.

The key economic indicators of process profitability considered in this study were the Payback Period (PBP), the Discounted Net Present Value (DNPV), and the Internal Rate of Return (IRR), shown in Table 2. The PBP is a financial indicator that takes into account both capital and operational expenditures, as well as the time value of money. The PBP of the process is defined as the exact period when the cumulative cash flow of the project becomes positive, hence when the project starts to be profitable. The DNPV is the sum of all discounted yearly cash flows over the project's lifetime. The IRR is the interest rate value that makes the DNPV equal to zero [32]. These particular metrics were implemented following the economic parameters described in Table 3.

2.4. Process optimization

The degrees of freedom (DOF) for the process optimization were selected by screening the variables expected to influence the economic performance of the plant. These variables were:

Table 3
Parameters for economic assessment.

Parameters	Discount rate	Taxation rate	Timespan	Residual value
Value	6%	24%	25 years	0 USD

- **Reforming temperature (Ref.T)**: Affects methane conversion and syngas production, as well as the operational expenditures of the process. Higher reforming temperatures increase the methane burning requirement but improve the syngas production.
- **Split ratio in synthesis (Split)**: Defined as the ratio between the recycled flow to the synthesis section and the unreacted gas mass flow after the second methanol synthesis reactor (Eq. (2)). It affects the overall syngas conversion and methanol production. Moreover, it influences the dimensions of the units and the total methane consumption.

$$\text{Split Ratio} = \frac{\text{Recycled gas}}{\text{Unreacted gas}} \tag{2}$$

- **Reactor temperatures (R1.T and R2.T)**: Related to the syngas conversion and methanol productivity.
- **Reactor volumes (R1.V and R2.V)**: Related to the syngas conversion and methanol productivity.

The process layout depends on the methanol synthesis pressure. Hence, it was fixed to 60 bar in this study to keep the described process layout. The initial, lower, and upper limits for the different degrees of freedom were set considering their economic and physical constraints. Their values are shown in Table 4. The reforming temperature is initially set to 950 °C to maximize methane conversion while remaining within the operational limits of the material [33]. A minimum value of 850 °C was considered to ensure sufficient methane conversion.

The reactor volumes were set to 4 m^3, which is the value to sustain a Gas Hourly Space Velocity of 10,000 h^{-1}, typical for methanol synthesis reactors [34]. Moreover, typical methanol synthesis temperatures and split ratios were assumed for the initial condition [35].

The DOF were used to optimize the economic performance metrics reported in Table 2. The optimization problem was studied in three separate case studies considering two, four, or all the reported DOF. This division was done to assess their impact on the optimization and the effectiveness of the SUR-OPT and SIM-OPT approaches. The case with two degrees of freedom considered the reforming temperature (Ref.T) and the split ratio (Split), which are variables not related to the synthesis section. These variables are highly correlated. The reforming temperature influences the thermal energy required to reach thermodynamic equilibrium (i.e., methane conversion), while the split ratio

Table 4
Degrees of freedom for the optimization problem.

Number	Name	Description	UoM	Initial	Min	Max
1	Split	Purge split ratio	–	5%	2%	15%
2	Ref.T	Reformer temperature	°C	950	850	950
3	R1.T	First reactor temperature	°C	250	230	270
4	R1.V	First reactor volume	m^3	4	1	4
5	R2.T	Second reactor temperature	°C	250	230	270
6	R2.V	Second reactor volume	m^3	4	1	4

changes the amount of gas sustaining the reforming unit. In contrast, the case study with 4 variables considered synthesis-related variables only, such as the reactors' temperatures and volumes. These variables are correlated with the methanol production and their associated revenues.

The derivative-free SIM-OPT problem was addressed using the BzzMath numerical package *BzzMinimizationRobust*, included in the CORO. It is a well-established mathematical software widely used for both academic and industrial purposes [36]. BzzMinimization Robust couples the derivative-free, valley-oriented OPTNOV algorithm with short Nelder–Mead sweeps, giving it exceptional stability on noisy, discontinuous, and computationally expensive objective functions common in chemical-process optimization. OPTNOV alternates coordinate line searches with simplex restarts along inferred valley directions, preventing simplex degeneration while efficiently traversing rugged response surfaces [37]. Prifti et al. (2022) described the power and reliability of the tool and tested its performance on several custom-made process simulations [38]. During the optimization, any simulation that fails to converge is automatically discarded, and the last successfully converged solution is reloaded and used as the starting point for the subsequent iteration. Convergence tolerances were set to 1×10^{-3} for the objective function and 1×10^{-4} for the degrees of freedom, ensuring both numerical stability and solution consistency.

2.5. Surrogate modeling

The complex and computationally expensive simulation of the methanol synthesis process described in Section 2.2 motivated a surrogate-modeling approach. The surrogate model was trained using simulation data extracted following a one-shot space-filling Design of Experiments methodology, to map the response surface of the different objective functions evenly. The surrogate was used to map the input/output correlation of the optimization variables (inputs, Table 2 and the objective functions (outputs, Table 4).

2.5.1. Design of experiments

The efficiency of one-shot approaches depends on their space-filling characteristics [39], which possibly guarantees an even representation or global approximation of the system's response. For this reason, the selected DoE strategy was the maximin-optimized Latin Hypercube Sampling (maximin LHS). This methodology has been widely implemented in chemical engineering applications [2], as it generates a design with the smallest maximum distance between sampled points, guaranteeing even spatial coverage of the domain. The methodology was applied using the Python package skopt [40].

The number of samples for the training dataset was selected based on Loeppky et al. guidelines [41], who determined that the number of samples should be around 10 times the number of input variables for an efficient computational experiment. For this reason, a total of 70 samples, equal to 10 times the maximum number of degrees of freedom (6 from Table 4) plus 10, were extracted. The additional simulations were proposed to account for potentially non-convergent simulations.

The simulations were executed sequentially, using a Python script linked to the Aspen HYSYS process simulation. If the simulator fails to converge, the next simulation in the experiment list is performed.

2.5.2. Data management and data pre-processing

The data management of this work considered an 80/20 proportion for the training and test sets. The training and test sets were designed independently using the same DoE approach (maximin-optimized LHS, Section 2.5.1) to reach similar sample distributions within the domain. Hence, the test dataset contained 18 samples. As a suggestion, the test set may be designed and simulated during the initial sensitivity analysis described at the beginning of this Section, which guarantees an unbiased evaluation of the space and reduces the total required computational expense. A summary of the data management process is presented in Fig. 6.

The designed samples were introduced to the simulation. The recovered data was validated and normalized in a pre-processing activity, excluding non-convergent simulations using the Pandas Python package [42]. The training set was used for model training, validation, and selection. With 5 folds, the cross-validation technique of the sklearn Python package [40] divided the training set in a proportion of 80/20 for training and validation. This means that each k-fold contained 16% of the total available data (14 points).

2.5.3. Mathematical models

The mathematical models chosen for surrogate modeling were some of the most widely used in the literature, including linear and polynomial regressions, regression forests, boosted trees, support vector regression, and the Kriging Gaussian Process with Radial Basis Function (RBF) kernel [2,43]. These models were available in the sklearn, Pytorch, and GPytorch Python packages [40,44].

Ten mathematical models were trained for each objective function. Model selection was based on the lowest average Mean Absolute Error (MAE) after the 5-fold cross-validation. Since the MAE was calculated over normalized data, it is not only an intuitive, uniform, and accurate metric to evaluate the error but also a representation of the relative error of the model.

2.6. Surrogate-based optimization

Several mathematical models were evaluated for surrogate model training, as discussed in Section 2.5.3. Each model required the use of different Python packages, presented their individual functional form, thus requiring the use of different Python packages, resulting in varied functional forms and different implementation characteristics. Moreover, since derivative information would only be available for a limited number of these models (such as linear and polynomial regressions), a derivative-based approach would not be possible for all of them, even though it would certainly generate accurate results efficiently. These inconsistencies motivated the adoption of a systematic Derivative-Free Optimization approach that could be used independently of the final selected model.

NOMAD (Nonlinear Optimization using the MADS algorithm) was chosen for derivative-free SUR-OPT. It is a well-known open-access black-box derivative-free optimization software that implements the Mesh Adaptive Direct Search (MADS) algorithm [45]. The MADS algorithm defines a mesh in which the surrogate model is evaluated. The mesh size is iteratively adjusted until no further improvement in the optimal value is observed. The algorithm has been designed for noisy functions [45], but has shown accurate results in smooth functions,

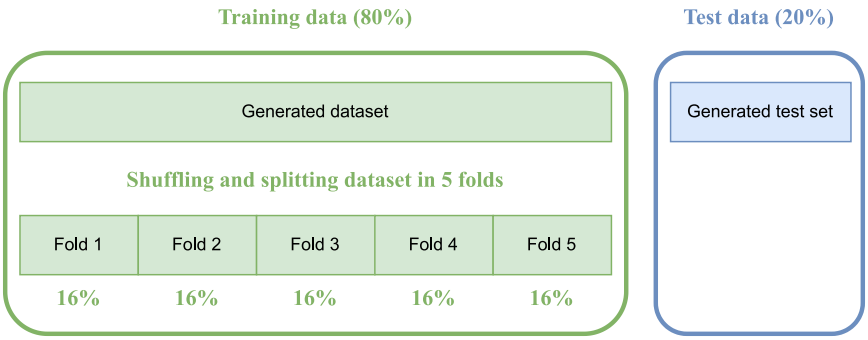


Fig. 6. Data division into training (70 samples), validation (14 samples), and test (18 samples) sets. Training and test data were generated using maximin-optimized LHS.

which are typically derived from surrogate models [16]. Moreover, NOMAD incorporates a variable neighborhood search metaheuristic to escape potential local optima. This may increase its computational cost but also enhances the likelihood of reaching a global optimum [16]. However, since NOMAD was used to optimize the trained surrogate model, each evaluation is inexpensive, allowing to perform a larger number of experiments. These characteristics motivated the selection of NOMAD over the OPTNOV algorithm used for SIM-OPT (Section 2.4).

This work implemented NOMAD using the PyNomad 4.4.0 Python package [45]. The NOMAD parameters were kept at their default values, except for VNS_MADS_SEARCH, which was activated to increase the probability of reaching global optimum values. Moreover, the maximum number of model evaluations (MAX_BB_EVAL) was set to 500 to guarantee an exhaustive domain exploration.

3. Results and analysis

3.1. Base case process simulation

The initial degrees of freedom shown in Table 4 were used to establish the base case scenario of the process simulation described in Section 2.2. The simulation presented a highly interconnected process with complexities associated with the required iterative calculations, typical for sequential solving approaches [9,10]. The overall computational time was affected by the distillation columns solving Boundary Value Problems, and the adjusting and recycling tools iterating to a feasible solution. The three adjusting items regulated the water make-up (Stream 18), the methane requirement (Stream 44), and the methanol synthesis closed-loop pressure (Stream 29), whereas the six recycle operations were required to account for recirculating streams due to the sequential nature of Aspen HYSYS.

Following the methodology in Fig. 1, a preliminary study was conducted to determine the stability and computational expense of the simulation [10]. For this purpose, the test set for the surrogate model, containing 18 simulations, was used (Section 2.5.2). After each simulation, the degrees of freedom were returned to the initial values shown in Table 4. The values of the objective functions after returning to the initial simulation state were recorded and used to quantify the simulation noise. The average and standard deviation of the noise found were 0.17% and 0.11%, respectively. These values are smaller than the 1% threshold mentioned in Section 2 for an accurate simulation, and set the limit for the accuracy of the optimization in both surrogate and simulation-based approaches. The average simulation time during the sensitivity analysis was 8 min.

At the initial simulation conditions, the 3600 kg/h of biogas feedstock (Stream 1) and 2159 kg/h of steam (Stream 16) transformed into 5759 kg/h of syngas with an SN of 1.93 (Stream 5) in the reforming section. The water make-up required in Stream 18 was 1467 kg/h. The furnace required 809 kg/h of methane (Stream 44) plus 680 kg/h of unreacted gases (Streams 37 and 40) to supply the 9.52 MW required by the reformer. The produced syngas entered the methanol synthesis section (Stream 20), where it was mixed with 6022 kg/h of recycled gas (Stream 38), increasing its SN value to 2.80 before the first reactor (Stream 21). The raw methanol produced by the two synthesis stages, with a molar fraction around 0.75, was 5443 kg/h in the first stage (Stream 27) and 4389 kg/h in the second (Stream 35). A total of 534 kg/h of unreacted gases (Stream 37) are purged and burned in the furnace.

Table 5
Base case utility consumption.

Utility	Electricity	CW	LPS	HPS
Consumption	753 kW	208 kg/h	2.05 kg/h	0.291 kg/h

Table 6
Economic assessment of the base case scenario.

KPI	PBP	DNPV	IRR
Value	8.62	24.2	14.7
UoM	y	USD M	%

At the base case scenario conditions, the utility consumption of the process is presented in Table 5, while the economic assessment is reported in Table 6. In particular, it was observed that the reforming and compression units presented the highest bare module costs at 5.7M and 2.6M USD, respectively. In addition, the main operational expenditures were shown by methane and biogas consumption (10.3M and 2.5M USD/y, respectively). Revenues only come from selling methanol.

3.2. CapEx OpEx robust optimization (SIM-OPT)

The derivative-free SIM-OPT problem described in Section 2.4 was solved considering two, four, and six degrees of freedom. Given the high computational expense of the simulation, only the PBP was used as the objective function, although all economic parameters were calculated at the optimal point. The optimal degrees of freedom for each case are presented in Table 7, while the economic key performance indicators are shown in Table 8. All the degrees of freedom chosen demonstrated an influence on the process's economic performance. In particular, the economic performance of the reactors improved at reduced volumes, demonstrating the potential for process intensification from standard volume determination heuristics (see Section 2.4). Moreover, the optimal value for the reformer temperature (Ref.T) exhibited a marked shift from its initial value towards its lower bound, demonstrating that the operating expenditures of sustaining the reformer had a more critical impact on the economics than the conversion of the syngas produced, contrary to the initial hypothesis described on Section 2.4.

Table 8 also presents the total number of iterations required by the CORO derivative-free optimizer to reach the described optimum. For these iterations, the previously estimated average of 8 min per

Table 7
Optimal parameters for the three optimized scenarios.

DOF	Objective function	Ref.T	R1.V	R2.V	R1.T	R2.T	Split
2	PBP	859	**4**	**4**	**250**	**250**	4.7%
4	PBP	**950**	2.10	2.11	253	244	**5.00%**
6	PBP	853	2.13	2.02	253	255	6.64%

Table 8
Key performance indicators for optimized scenarios.

DOF	Objective function	PBP	DNPV	IRR	n° Iteration
Base	–	8.62	24.2	14.7	–
2	PBP	8.34	24.6	15.1	82
4	PBP	8.25	24.3	15.3	64
6	PBP	7.82	25.0	16.0	105

Table 9
Normalized MAE (equivalent to percentage relative error) in the test set for the two best-performing trained mathematical models for the output variables of Table 2.

Name	UoM	Best model	%MAE	2nd Best model	%MAE
PBP	y	Kriging	4.66%	2 Order Polynomial	5.51%
NPV	USD	Kriging	4.79%	2 Order Polynomial	5.57%
IRR	–	Kriging	4.85%	2 Order Polynomial	5.42%

simulation persisted. These iterations showed a dependency on the distance of the initial guess to the final optimization point and the steepness of the response. For instance, similar values were observed when the temperatures and volumes of the reactors changed, which may be related to the presence of a valley and thus to the quicker convergence of the optimizer compared to the case with two degrees of freedom. Regarding the optimization with six degrees of freedom, an improvement of 9.3% in the economic performance was observed. However, the large number of iterations leads to extensive computational time, considering the estimated 8 min per simulation (see Section 3.1).

3.3. Surrogate modeling

The ten mathematical models described in Section 2.5 were trained using the dataset designed with the DoE strategy reported in Section 2.5. The required training time was 104 s for the 9 regression models and 116 s for Kriging. The model accuracy results for the two best mathematical models are presented in Table 9. Moreover, Fig. 7 displays the parity plots for Kriging, the model with the highest prediction accuracy. The final surrogate model of the simulation was the ensemble of the best mathematical models for each output variable prediction. The Kriging Gaussian Process with RBF kernel presented the most satisfactory results. However, simpler models such as linear or polynomial regression also showed high prediction accuracies. This suggests that, although the predicted variables present a marked nonlinear behavior, linear or quadratic functions may still be accurate for quickly mapping their response.

The average MAE for the output variables was 4.6%, which is higher than the surrogates trained by Andreasen and van Baten [17] or Kum et al. [18] in similar applications. This error may be explained by the complexity and high integration of the simulation case used in this work, which complicated its representation with a surrogate. Moreover, the dataset used in this work is smaller than the one reported by these authors, since the objective was to efficiently map the response surface by reducing the computational time. Despite the error being smaller than the threshold defined in Fig. 1, both SUR-OPT and MIX-OPT approaches were evaluated with this surrogate.

3.4. Surrogate-based optimization (SUR-OPT)

The surrogate model of the simulation supported a derivative-free optimization using NOMAD (Section 2.6). The optimization problem was solved using two, four, and six degrees of freedom, as described in the same Section. Due to the fast evaluation time of the surrogate models, the average optimization time was less than one minute for all objective functions defined in Section 2.4. However, the optimization problem was solved nine times (for a total of 9 min) to determine the best solution. The best results for the objective functions are presented in Table 10, while the best values for the optimization variables are presented in Table 11. All the results were validated with a rigorous process simulation to assess the feasibility of the solution.

Table 10 shows that the proposed objective functions reached highly accurate solutions when two, four, or six degrees of freedom were considered. Overall, the accuracy was higher for the PBP and IRR objective functions. The different case studies reached an optimum at the lower boundary of the reformer temperature (Ref.T), at 850 °C. The remaining degrees of freedom showed wide variations in the different cases studied. The presence of an optimum value close to the boundaries of the input space used for training may explain the higher relative error in the result validation with the simulation, since the point density close to the boundaries in algorithms based on LHS decreases, increasing the risk of extrapolation.

3.5. Surrogate-based sensitivity analysis

The trend to reach the lower boundary of the reformer temperature (Ref.T) may suggest a steep behavior of the system response to this variable. To further study the behavior of the variables, a sensitivity analysis was conducted. First, the optimization with 2 DoF was further studied in Figs. 8 and 9. The response surface is described by a steep surface with a marked optimum at a reformer temperature (Ref.T) of 850 °C. In contrast, the reactors' temperatures and volumes presented extensive valleys with similar economic performances, as shown in Figs. 10, 11, and 12. However, the location and extension of the level lines surrounding the smallest PBP suggest the most likely location of the actual optimum of the system. In particular, Fig. 10 shows a lower economic sensitivity at higher reactor temperatures and lower volumes. Moreover, Figs. 11 and 12 suggest that the first reactor should have a higher temperature and volume than the second. The average time to construct each plot using the surrogate models was 3 min due to the evaluation of the 400 points required by the generated mesh (20x20). The smoothness of the response surfaces arises from the surrogate models, which capture the overall trend of the experimental response rather than closely fitting each experiment.

The sensitivity analysis results agree with the initial hypothesis made in Section 3.4 regarding the behavior of the response surface. This information may be further used to define appropriate first guesses and variable boundaries in an eventual derivative-free MIX-OPT. These narrower boundaries could improve the optimization times in simulation-based approaches. For example, defining the boundaries considering a PBP of 7.85 in the sensitivity analysis, the possible initial values and boundaries could be those presented in Table 12.

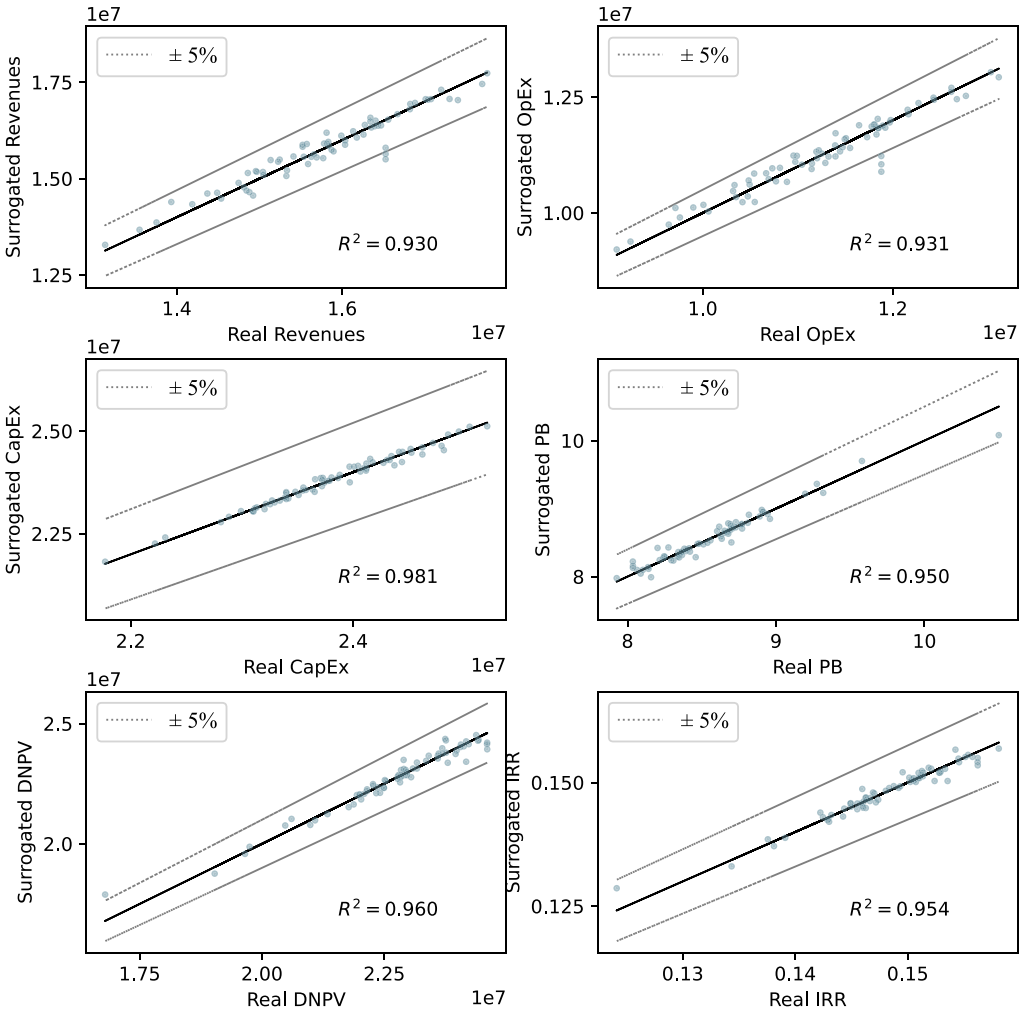


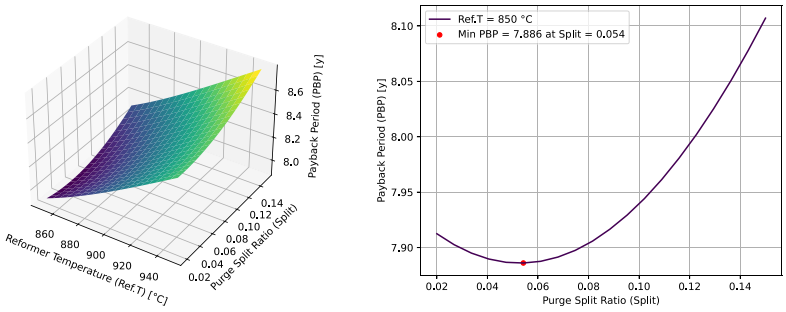
Fig. 7. Parity Plots for the output variables of Table 2 using the best mathematical models for prediction described in Table 9.

Table 10
Optimum values for the objective functions in the different SUR-OPT cases. The optimum values were validated and compared with the simulation to determine their relative error.

DOF	Objective function	Optimum	HYSYS validation	Error
Base	–	8.62	8.62	–
2	PBP	7.89	7.94	0.63%
2	DNPV	$ 24,912,910.71	$ 23,172,132.89	7.51%
2	IRR	15.81%	15.80%	0.06%
4	PBP	7.97	8.11	1.79%
4	DNPV	$ 24,729,745.00	$ 24,539,568.67	0.77%
4	IRR	15.68%	15.46%	1.45%
6	PBP	7.78	8.25	5.70%
6	DNPV	$ 25,740,777.00	$ 23,497,975.03	9.54%
6	IRR	15.97%	15.22%	4.93%

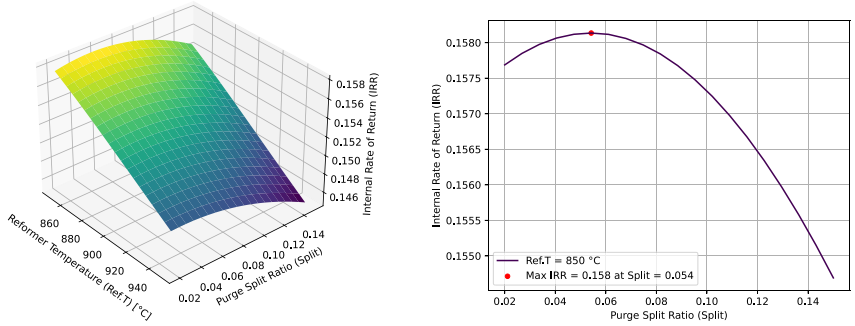
Table 11
Optimum values for the optimization variables in the different SUR-OPT cases. The values in bold were kept constant in the specified optimization case.

DOF	Objective function	Ref.T	R1.V	R2.V	R1.T	R2.T	Split
Base	–	**950.0**	**4.00**	**4.00**	**250.0**	**250.0**	**5.0%**
2	PBP	850.0	**4.00**	**4.00**	**250.0**	**250.0**	5.4%
2	DNPV	850.0	**4.00**	**4.00**	**250.0**	**250.0**	2.0%
2	IRR	850.0	**4.00**	**4.00**	**250.0**	**250.0**	5.4%
4	PBP	**950.0**	1.78	1.38	270.0	251.5	**5.0%**
4	DPV	**950.0**	2.32	2.34	263.7	250.2	**5.0%**
4	IRR	**950.0**	1.72	1.25	270.0	252.6	**5.0%**
6	PBP	850.0	2.38	2.06	263.0	246.0	2.5%
6	DNPV	850.0	3.41	2.99	246.8	245.6	2.0%
6	IRR	850.0	2.11	2.03	267.0	246.5	3.2%



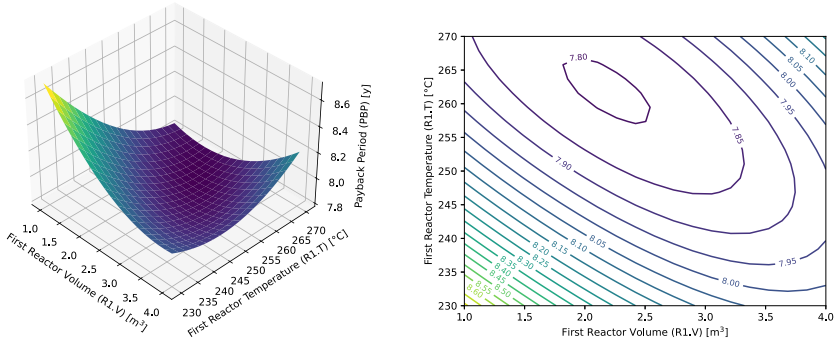
(a) Response surface (b) Response at fixed Ref.T = 850 °C

Fig. 8. PBP as a function of split ratio (Split) and reformer temperature (Ref.T).



(a) Response surface (b) Response at fixed Ref.T = 850 °C

Fig. 9. IRR as a function of split ratio (Split) and reformer temperature (Ref.T).



(a) Response surface (b) Level Lines

Fig. 10. PBP as a function of first reactor temperature (R1.T) and volume (R1.V).

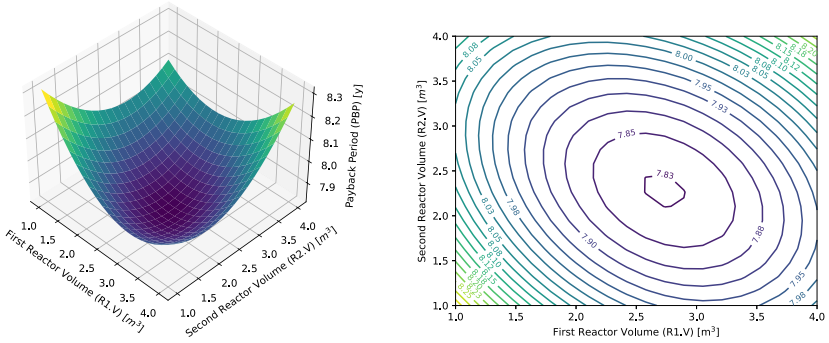
Table 12
Potential initial guess and boundaries for an eventual rigorous SIM-OPT defined from SUR-OPT and sensitivity analysis.

Name	UoM	Start	Min	Max
Split	–	5%	2%	8%
Ref.T	°C	855	850	860
R1.T	°C	265	255	270
R1.V	m^3	2.7	2.2	3.2
R2.T	°C	247	240	255
R2.V	m^3	2.2	1.7	2.7

3.6. Simulation-based optimization with narrower boundaries (MIX-OPT)

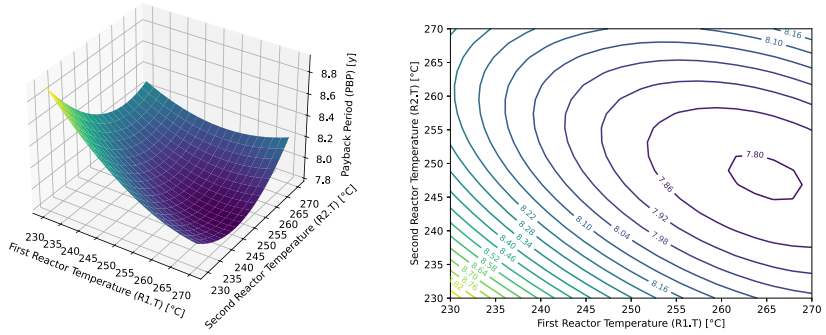
The results for the optimization problem described in Table 12 are presented in Table 13. The narrower boundaries were expected to reduce the transient periods between successive simulations due to the proximity of the points [19]. Indeed, the computational time of the simulation decreased to 2 min. This is explained since convergence in sequential simulators resembles an optimization, which converges more quickly with an appropriate initial guess. In this case, the proximity of the simulations acted as improved initial guesses to reach a new steady-state.

With the MIX-OPT methodology, the optimal value was improved compared to the SUR-OPT approach. However, it was not possible



(a) Response surface (b) Level Lines

Fig. 11. PBP as a function of first and second reactor's volume (R1.V, R2.V).



(a) Response surface (b) Level Lines

Fig. 12. PBP as a function of first and second reactor's temperature (R1.T, R2.T).

Table 13
SIM-OPT results supported on the surrogate-based initial guess and boundaries defined in Table 12.

PBP	DNPV	IRR	n° Iteration
7.91	24.9	15.8	76

to reach the same minimum obtained in the SIM-OPT since the defined boundaries did not include the optimal points of two degrees of freedom (R1.V and R1.T). These values may be included by defining broader optimization boundaries, which would improve the optimum accuracy but would also increase the total computational time.

Table 14 summarizes the results for the three optimization approaches for PBP optimization: SIM-OPT, SUR-OPT, and MIX-OPT. These results could have been affected by the simulation *noise*, which has been estimated to affect the PBP improvement by up to 0.3% (See Section 3.1). The computational times required by simulation, model training, sensitivity analysis, and process optimization are presented as follows:

- **Simulation time:** Refers to the simulation runs used to build the dataset for surrogate model training (70 simulations, 8 min per simulation, for a total of 560 min).
- **Training time:** Required to train the 10 mathematical models evaluated in this work (around 4 min, see Section 3.3).
- **Sensitivity analysis time:** Includes the initial sensitivity analysis time required to assess the simulation complexity (18 simulations, 8 min per simulation, for a total of 144 min) and the surrogate-based sensitivity analysis described in Section 3.5 (15 min). The initial sensitivity test is also used as the test set for the surrogate model (See Section 2.5.2).

Table 14
Summary table of the three optimization approaches followed. The total time is reported as simulation time/model training time/optimization time/sensitivity analysis time/total time.

Approach	Best PBP	PBP improvement	Time (min)
Base	8.62	0.00%	–
SUR-OPT	7.94	7.89%	560/4/9/144/**717**
SIM-OPT	7.82	9.28%	0/0/840/144/**984**
MIX-OPT	7.91	8.24%	560/4/161/159/**884**

- **Optimization time:** Refers to the time required by the NOMAD optimization in the SUR-OPT and MIX-OPT approaches (1 min, repeated 9 times for a total of 9 min) and to the simulation-based approaches' time. In the MIX-OPT, the average time per simulation was considered 2 min instead of 8, since the computational expense was reduced with narrower boundaries, as mentioned above.

Each optimization approach studied in Table 14 presented particular advantages. On the one hand, SUR-OPT demonstrated an overall reduction in the computational time of 20%. The surrogate model trained showed high flexibility and fast calculation speed, ideal to solve different optimization problems, or to study the system behavior via sensitivity analysis. On the other hand, SIM-OPT presented the highest accuracy, but the largest computational times.

The results suggest that the combined use of both optimization approaches (MIX-OPT) offers an efficient trade-off between computational time and optimization accuracy, reducing the total computational expense by 10% while achieving an optimum within 1% of that obtained using SIM-OPT. In case the accuracy of the surrogate model of the simulation was higher, the boundary redefinition could have been tighter,

Table 15
Hypothetical computational time for the three optimization approaches considering an average time per simulation of 1.75 min instead of the original 8. The total time is reported as simulation time/model training time/optimization time/sensitivity analysis time/total time.

Approach	Time (min)
Base	–
SUR-OPT	122/4/9/32/**167**
SIM-OPT	0/0/183/32/**215**
MIX-OPT	122/4/42/47/**215**

further reducing computational expense and/or increasing optimization accuracy.

The MIX-OPT approach may be unnecessary for less complex simulations with shorter computational times, where the SIM-OPT approach should be used. Determining the boundary computational time for this purpose depends on several factors, in particular, the system dimensionality. For a dimensionality of six, as in this case study, the boundary time may be estimated by assuming constant model training, an equal number of simulation runs, and surrogate-based optimization time. Moreover, it is assumed that the simulation time after the sensitivity analysis keeps the ratio observed in this study (i.e., 25% of the original simulation time). In this order of ideas, a simulation with an average convergence time of 1.75 min would have presented the results shown in Table 15, suggesting the SIM-OPT approach would be ideal for simulations up to this computational time.

4. Conclusions

In this work, three derivative-free optimization methodologies were evaluated and compared for the simultaneous CapEx and OpEx optimization during the design of an industrial-scale methanol production plant from biogas: (i) simulation-based (SIM-OPT), (ii) surrogate-based (SUR-OPT), (iii) and the novel proposed method (MIX-OPT). The case study was a highly integrated simulation with long computational times. The SIM-OPT made use of CORO, a robust economic tool for economic assessment, while the SUR-OPT used NOMAD, an open-access efficient optimization software.

SUR-OPT presented the highest calculation speed with the lowest accuracy, the opposite of the SIM-OPT. This suggests that simulations with a short convergence time may benefit from direct simulation-based optimization. The rigorous SIM-OPT was complemented with the surrogate model in a novel MIX-OPT approach by implementing the results of the SUR-OPT and surrogate-based sensitivity analysis to set the initial point and the boundaries of the simulation-based rigorous approach. The MIX-OPT provided an efficient trade-off between the computational time required and the optimization accuracy, even with a surrogate model with an average prediction error of around 5% concerning the simulation.

Although this work presented a comprehensive and efficient framework for process optimization, there are still research questions to address. For instance, the specific optimizers, tolerances and solving algorithms chosen for each optimization strategy (SIM, SUR, or MIX-OPT) may further influence the trade-off between model accuracy and computational time. Moreover, the feasibility of the MIX-OPT approach strongly depends on the average convergence time of the simulation. A preliminary study suggested that problems with similar dimensionality would benefit from the MIX-OPT approach at average times per simulation larger than 1.75 min. However, further research should generalize the optimization approach selection considering the process dimensionality. Moreover, a more rigorous sensitivity analysis for boundary redefinition in the MIX-OPT approach should be investigated, considering how the surrogate model's precision affects its redefinition. Ultimately, the presence of large valleys in the optimization problem could be better addressed through a weighted objective function, which would emphasize capital expenditure (CapEx), favoring solutions with comparable objective values but higher CapEx.

CRediT authorship contribution statement

Luis Felipe Sánchez: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Marcello Maria Bozzini:** Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Mattia Vallerio:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Flavio Manenti:** Writing – review & editing, Validation, Supervision, Resources, Project administration, Funding acquisition, Conceptualization.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the author(s) used ChatGPT and Grammarly for proofreading and support in English grammar and sentence structure on around 5% of the text. After using these tools, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

Declaration of competing interest

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

References

[1] O.T. Kajero, T. Chen, Y. Yao, Y.-C. Chuang, D.S.H. Wong, Meta-modelling in chemical process system engineering, J. Taiwan Inst. Chem. Eng. 73 (2017) 135–145, http://dx.doi.org/10.1016/j.jtice.2016.10.042.
[2] K. McBride, K. Sundmacher, Overview of Surrogate Modeling in Chemical Process Engineering, Chem. Ing. Tech. 91 (3) (2019) 228–239, http://dx.doi.org/10.1002/cite.201800091.
[3] S. Becht, R. Franke, A. Geißelmann, H. Hahn, An industrial view of process intensification, Chem. Eng. Process.: Process. Intensif. 48 (1) (2009) 329–332, http://dx.doi.org/10.1016/j.cep.2008.04.012.
[4] B. Muster-Slawitsch, Process intensification for circular economy, Chem. Eng. Process. - Process. Intensif. 208 (2025) 110095, http://dx.doi.org/10.1016/j.cep.2024.110095.
[5] D.C. Boffito, Scaling process intensification technologies: what does it take to deploy? Chem. Eng. Process. - Process. Intensif. 212 (2025) 110275, http://dx.doi.org/10.1016/j.cep.2025.110275.
[6] M.K. Pasha, L. Dai, D. Liu, M. Guo, W. Du, An overview to process design, simulation and sustainability evaluation of biodiesel production, Biotechnol. Biofuels 14 (1) (2021) 129, http://dx.doi.org/10.1186/s13068-021-01977-z.
[7] K.N. Sun, S.R. Wan Alwi, Z.A. Manan, Heat exchanger network cost optimization considering multiple utilities and different types of heat exchangers, Comput. Chem. Eng. 49 (2013) 194–204, http://dx.doi.org/10.1016/j.compchemeng.2012.10.017.
[8] R. Turton, J. Shaeiwitz, D. Bhattacharyya, W. Whiting, Analysis, Synthesis, and Design of Chemical Processes, fifth ed., Pearson, Boston, 2018.
[9] T. Janus, S. Engell, Iterative Process Design with Surrogate-Assisted Global Flowsheet Optimization, Chem. Ing. Tech. 93 (12) (2021) 2019–2028, http://dx.doi.org/10.1002/cite.202100095.
[10] J.A. Caballero, I.E. Grossmann, An algorithm for the use of surrogate models in modular flowsheet optimization, AIChE J. 54 (10) (2008) 2633–2650, http://dx.doi.org/10.1002/aic.11579.
[11] N.M.V. Ramos, O. Del-Mazo-Alvarado, A. Bonilla-Petriciolet, L.F. de Lima Luz Jr., M.L. Corazza, Multi-objective optimization of syngas production for Fischer-Tropsch synthesis based on biogas catalytic reforming and upgrading, Chem. Eng. Process. - Process. Intensif. 199 (2024) 109758, http://dx.doi.org/10.1016/j.cep.2024.109758.
[12] M. Frosi, A. Tripodi, F. Conte, G. Ramis, N. Mahinpey, I. Rossetti, Ethylene from renewable ethanol: Process optimization and economic feasibility assessment, J. Ind. Eng. Chem. 104 (2021) 272–285, http://dx.doi.org/10.1016/j.jiec.2021.08.026.
[13] M. Variny, L. Hlavatý, Z. Silná, T. Soták, Design and economic evaluation of separation process for novel production of cyclopentyl methyl ether, J. Chem. Technol. Biotechnol. n/a n/a (2025) http://dx.doi.org/10.1002/jctb.7839.

[14] F. Duanmu, D.N. Chia, A. Tsatse, E. Sorensen, Optimal design and operation of reactive distillation systems and reactive dividing wall systems with pressure swing distillation and hybrid distillation-pervaporation, Chem. Eng. Process. - Process. Intensif. 201 (2024) 109832, http://dx.doi.org/10.1016/j.cep.2024.109832.

[15] A. Kazemi, Z. Kazemi, M.A. Pourmohammadi, Combined bio-methanol and power production from bio-oil: Proposing a clean bio-process through waste heat recovery and environmental optimization, Chem. Eng. Process. - Process. Intensif. 201 (2024) 109807, http://dx.doi.org/10.1016/j.cep.2024.109807.

[16] N. Ploskas, N.V. Sahinidis, Review and comparison of algorithms and software for mixed-integer derivative-free optimization, J. Global Optim. 82 (3) (2022) 433–462, http://dx.doi.org/10.1007/s10898-021-01085-0.

[17] A. Andreasen, J. van Baten, Techno-Economic Analysis and Optimization of Gray and Green Methanol Synthesis Using Flowsheet Automation and Surrogate Modeling, Energy Fuels 39 (6) (2025) 3359–3374, http://dx.doi.org/10.1021/acs.energyfuels.4c05806, publisher: American Chemical Society.

[18] J. Kum, H.-T. Oh, J. Park, J.-H. Kang, C.-H. Lee, Techno-economic analysis and optimization of a CO2 absorption process with a solvent looping system at the absorber using an MDEA/PZ blended solvent for steam methane reforming, Chem. Eng. J. 455 (2023) 140685, http://dx.doi.org/10.1016/j.cej.2022.140685.

[19] D. McNickle, G.C. Ewing, K. Pawlikowski, Some effects of transient deletion on sequential steady-state simulation, Simul. Model. Pr. Theory 18 (2) (2010) 177–189, http://dx.doi.org/10.1016/j.simpat.2009.10.004.

[20] AspenTech, AspenTech HYSYS® 2004.2 user guide, 2023.

[21] A. Forrester, A. Sobester, A. Keane, Constraints, In: Engineering Design Via Surrogate Modelling, John Wiley & Sons Ltd, 2008, pp. 117–139.

[22] S. Ravutla, A. Bai, M.J. Realff, F. Boukouvala, Effects of Surrogate Hybridization and Adaptive Sampling for Simulation-Based Optimization, Ind. Eng. Chem. Res. 64 (18) (2025) 9228–9251, http://dx.doi.org/10.1021/acs.iecr.4c03303.

[23] L. Salano, M.M. Bozzini, S. Caspani, G. Bozzano, F. Manenti, Integration of an Autothermal Outer Electrified Reformer Technology for Methanol Production from Biogas: Enhanced Syngas Quality Production and CO2 Capture and Utilization Assessment, Processes 12 (8) (2024) 1598, http://dx.doi.org/10.3390/pr12081598.

[24] G. Bozzano, F. Manenti, Efficient methanol synthesis: Perspectives, technologies and optimization strategies, Prog. Energy Combust. Sci. 56 (2016) 71–105, http://dx.doi.org/10.1016/j.pecs.2016.06.001.

[25] L. Salano, M.M. Bozzini, F. Manenti, Techno-Economic Analysis for Biogas Reforming using PSWA: Case Study on Methanol Synthesis, in: F. Manenti, G.V. Reklaitis (Eds.), Computer Aided Chemical Engineering, in: 34European Symposium on Computer Aided Process Engineering / 15 International Symposium on Process Systems Engineering, 53, Elsevier, 2024, pp. 1045–1050.

[26] K.M.V. Bussche, G.F. Froment, A Steady-State Kinetic Model for Methanol Synthesis and the Water Gas Shift Reaction on a Commercial Cu/ZnO/Al2O3Catalyst, J. Catalysis 161 (1) (1996) 1–10, http://dx.doi.org/10.1006/jcat.1996.0156.

[27] F. Pontzen, W. Liebner, V. Gronemann, M. Rothaemel, B. Ahlers, CO2-based methanol and DME – Efficient technologies for industrial scale production, Catal. Today 171 (1) (2011) 242–250, http://dx.doi.org/10.1016/j.cattod.2011.04.049.

[28] D.V. Demidov, I.V. Mishin, M.N. Mikhailov, Gibbs free energy minimization as a way to optimize the combined steam and carbon dioxide reforming of methane, Int. J. Hydrog. Energy 36 (10) (2011) 5941–5950, http://dx.doi.org/10.1016/j.ijhydene.2011.02.053.

[29] L. Yue, G. Li, G. He, Y. Guo, L. Xu, W. Fang, Impacts of hydrogen to carbon ratio (h/c) on fundamental properties and supercritical cracking performance of hydrocarbon fuels, Chem. Eng. J. 283 (2016) 1216–1223, http://dx.doi.org/10.1016/j.cej.2015.08.081.

[30] M.M. Bozzini, L. Salano, C. Pirola, F. Manenti, Economic Optimization of the Synthesis Section of a Small-Scale Biogas-to-Methanol Plant, in: F. Manenti, G.V. Reklaitis (Eds.), Computer Aided Chemical Engineering, in: 34 European Symposium on Computer Aided Process Engineering / 15 International Symposium on Process Systems Engineering, vol. 53, Elsevier, 2024, pp. 1063–1068.

[31] M. Bozzini, K. Prifti, A. Galeazzi, F. Manenti, Capex Opex Robust Optimization: a Software for Well-founded Economic Estimations and Process Optimization, Chem. Eng. Trans. 99 (2023) 619–624, http://dx.doi.org/10.3303/CET2399104.

[32] D. Bellotti, M. Rivarolo, L. Magistri, A.F. Massardo, Feasibility study of methanol production plant from hydrogen and captured carbon dioxide, J. CO2 Util. 21 (2017) 132–138, http://dx.doi.org/10.1016/j.jcou.2017.07.001.

[33] R.D. Alli, P.A.L. de Souza, M. Mohamedali, L.D. Virla, N. Mahinpey, Tri-reforming of methane for syngas production using Ni catalysts: Current status and future outlook, Catal. Today 407 (2023) 107–124, http://dx.doi.org/10.1016/j.cattod.2022.02.006.

[34] A. Bansode, A. Urakawa, Towards full one-pass conversion of carbon dioxide to methanol and methanol-derived products, J. Catalysis 309 (2014) 66–70, http://dx.doi.org/10.1016/j.jcat.2013.09.005.

[35] E.S. Van-Dal, C. Bouallou, Design and simulation of a methanol production plant from CO2 hydrogenation, J. Clean. Prod. 57 (2013) 38–45, http://dx.doi.org/10.1016/j.jclepro.2013.06.008.

[36] G. Buzzi-Ferraris, F. Manenti, BzzMath: Library Overview and Recent Advances in Numerical Methods, in: I.D.L. Bogle, M. Fairweather (Eds.), Computer Aided Chemical Engineering, in: 22 European Symposium on Computer Aided Process Engineering, vol. 30, Elsevier, 2012, pp. 1312–1316.

[37] G. Buzzi-Ferraris, F. Manenti, Nonlinear Systems and Optimization for the Chemical Engineer, John Wiley & Sons Ltd, 2013.

[38] K. Prifti, A. Galeazzi, M. Barbieri, F. Manenti, A Capex Opex Simultaneous Robust Optimizer: Process Simulation-based Generalized Framework for Reliable Economic Estimations, in: L. Montastruc, S. Negny (Eds.), Computer Aided Chemical Engineering, in: 32 European Symposium on Computer Aided Process Engineering, vol. 51, Elsevier, 2022, pp. 1321–1326.

[39] R. Yondo, E. Andrés, E. Valero, A review on design of experiments and surrogate models in aircraft real-time and many-query aerodynamic analyses, Prog. Aerosp. Sci. 96 (2018) 23–61, http://dx.doi.org/10.1016/j.paerosci.2017.11.003.

[40] O. Grisel, A. Mueller, F. Pedregosa, Lars, A. Gramfort, G. Louppe, P. Prettenhofer, M. Blondel, V. Niculae, A. Joly, J. Nothman, J. Vanderplas, M. Kumar, R. Layton, N. Varoquaux, N. Dawe, J. Schönberger, D.A. Engemann, W. Li, R.R. V, C. Woolam, K. Eren, Eustache, A. Fabisch, A. Passos, bthirion, V. Fritsch, D. Sullivan, H. Alsalhi, M. Wijewardena, Scikit-learn: 0.17.1 release tag for DOI, 2016, http://dx.doi.org/10.5281/zenodo.49911.

[41] J.L. Loeppky, J. Sacks, W.J. Welch, Choosing the Sample Size of a Computer Experiment: A Practical Guide, Technometrics 51 (4) (2009) 366–376, http://dx.doi.org/10.1198/TECH.2009.08040.

[42] Pandas development team, Pandas-dev/pandas: Pandas, 2024, http://dx.doi.org/10.5281/zenodo.13819579.

[43] L.F.S. Martínez, A. Galeazzi, F. Manenti, Surrogate-Based Optimization of the OPEX of a Modular Plant for Biogas Conversion to Methanol Using the MADS Algorithm, in: F. Manenti, G.V. Reklaitis (Eds.), Computer Aided Chemical Engineering, in: 34 European Symposium on Computer Aided Process Engineering / 15 International Symposium on Process Systems Engineering, vol. 53, Elsevier, 2024, pp. 3103–3108, http://dx.doi.org/10.1016/B978-0-443-28824-1.50518-4.

[44] GPyTorch, https://gpytorch.ai/.

[45] S. Le Digabel, Algorithm 909: NOMAD: Nonlinear Optimization with the MADS Algorithm, ACM Trans. Math. Software 37 (4) (2011) 44:1–44:15, http://dx.doi.org/10.1145/1916461.1916468.