

## Surrogate-Based Operability Analysis: Employing Physics-Based Simulation and Machine Learning for Process Systems Engineering Applications

Rafael Ramos Heilbuth<sup>a</sup>, Hannel Ferreira Sarmiento de Freitas<sup>b</sup>, Sérgio M. S. Neiro<sup>a</sup>, Nicolas Spogis<sup>c</sup>, Fernando V. Lima<sup>d\*</sup>

<sup>a</sup>Faculty of Chemical Engineering, Federal University of Uberlândia, Uberlândia-MG, Brazil

<sup>b</sup>Federal Institute of Education, Science and Technology of RN, Nova Cruz-RN, Brazil

<sup>c</sup>School of Chemical Engineering, University of Campinas, Campinas-SP, Brazil

<sup>d</sup>Department of Chemical and Biomedical Engineering, West Virginia University, Morgantown-WV, USA

\*fernando.lima@mail.wvu.edu

### ABSTRACT

Process operability analysis is a powerful framework for evaluating the achievability and feasibility of chemical processes, quantifying the relationship between available inputs and achievable outputs. However, traditional operability tools rely on computationally expensive first-principles models, limiting their practical application in industrial settings. This work presents a comprehensive methodology comparing physics-based process simulation with machine learning surrogates to enable rapid and accurate operability analysis. Three industrially relevant case studies are presented: a heat exchanger with bypass control, an ethanol-water distillation column, and a direct methane aromatization membrane reactor (DMA-MR). Multi-Layer Perceptron (MLP) and Gaussian Process (GP) surrogate models are trained on data generated from rigorous simulations in DWSIM and custom first-principles models. The surrogates demonstrate excellent predictive accuracy ( $R^2 > 0.999$ ) while achieving computational speedups from 4× to over 20,000× when compared to the original first-principles models. Forward mapping, operability index calculations, and inverse mapping procedures are successfully performed using the Opyrability package in Python, revealing feasible operating envelopes and design spaces. Results demonstrate that validated surrogates can reliably replace high-fidelity models for operability analysis, enabling rapid exploration of design alternatives, process optimization, and decision-making under uncertainty. This integrated approach bridges the gap between rigorous process modeling and computationally efficient engineering workflows.

**Keywords:** process operability, surrogate modeling, DWSIM, neural networks, Gaussian processes

### 1. Introduction

Process operability, pioneered by Vinson and Georgakis (2000), is a fundamental concept in chemical engineering that addresses the relationship between process inputs and outputs, quantifying the ability of a process to achieve desired operating conditions given the available manipulated variables' spaces and inherent constraints, providing a systematic methodology for evaluating process achievability, identifying feasible operating regions, and supporting design decisions under uncertainty. However, despite its theoretical elegance and practical value, operability analysis faces significant computational challenges when applied to industrial-scale processes. Traditional approaches require numerous evaluations of high-fidelity process models, particularly for inverse mapping procedures that involve solving multiple constrained optimization problems, which can lead to prohibitive computational times.

In this sense, surrogate modeling (also known as metamodeling) provides a computationally feasible alternative. Recent efforts have successfully integrated machine learning into the operability framework; for instance, Alves et al. (2022) demonstrated that Gaussian Process Regression (also known as Kriging) can substitute first-principles models, reducing computational times by up to four orders of magnitude while maintaining high accuracy. Building upon these advances, particularly in neural networks and Gaussian processes, data-driven surrogates can reliably approximate complex, high-dimensional process behavior.

This work proposes a Python-based framework for surrogate-assisted operability analysis, validated through three industrial case studies: a heat exchanger, an ethanol-distillation column, and a direct methane aromatization membrane reactor (DMA-MR). Training datasets were generated via two distinct approaches:



Realização:



rigorous simulation in DWSIM (for the heat exchanger and distillation cases) and numerical integration of an ODE-based model for the DMA-MR. The methodology integrates scikit-learn for surrogate training and the Opyrability package (Alves et al., 2024) for calculating the Operability Index (OI), focusing on the trade-off between computational speedup and mapping accuracy.

## 2. Methodology

The Python-based methodology follows a four-step pipeline, which includes the steps: (i) Latin Hypercube Sampling (LHS) for uniform input space coverage; (ii) data generation using DWSIM (heat exchanger and distillation) and JAX - a high-performance numerical library enabling automatic differentiation via computational graphs - for the DMA-MR ODE model; (iii) surrogate training using scikit-learn; and (iv) operability analysis via the Opyrability package. The following subsections outline the operability framework, process models, surrogate validation protocols, and summarize the computational configurations for all case studies.

### 2.1 Operability framework

The process operability concept evaluates the mapping between two sets: the Available Input Set (AIS), which defines the feasible region for manipulated variables, typically bounded by physical constraints; and the Achievable Output Set (AOS), which represents all outputs reachable through the process model given the AIS. A process is considered operable if the Desired Output Set (DOS), which specifies target operating conditions, is contained within the AOS. The Operability Index (OI) is defined as the percentage of the desired output region that can be reached by the process, represented by the ratio of their respective hyper-volumes ( $\mu$ ) or Lebesgue measure of the sets within the output domain (Eq.1).

$$OI = [\mu(AOS \cap DOS)/\mu(DOS)] \times 100\% \quad (1)$$

This mapping is characterized by using polytope theory, following the framework established by Vinson and Georgakis (2000). The Available Input Set (AIS) is typically defined in V-representation (V-rep) as the convex hull of the operating limits, while the Desired Output Set (DOS) often incorporates process constraints defined in H-representation (H-rep) as a system of linear inequalities ( $Ax \leq b$ ). As detailed by Carrasco and Lima (2017), the conversion between these representations is a critical step for computing the intersection between the achievable and desired regions. In this work, the Opyrability package (Alves et al., 2024) is utilized to perform these geometric operations and calculate the Operability Index (OI) using the Lebesgue measure ( $\mu$ ).

Finally, inverse mapping identifies the Design Input Set (DIS) required to reach the DOS. This is formulated as a constrained optimization problem, where the surrogate model's efficiency allows for rapid convergence using gradient-based solvers or optimization algorithms for non-convex spaces, ensuring the recovery of the feasible design space.

### 2.2 Process models and data generation

Training data was generated through two distinct modeling and simulation approaches, reflecting different levels of computational complexity. For the heat exchanger and distillation column cases, rigorous steady-state models were developed in DWSIM. The data acquisition was automated via a Python script using the COM interface (Automation3), allowing for seamless communication between the simulator and the sampling algorithm. For the DMA-MR case, a first-principles model consisting of mass and energy balances was implemented. To ensure high-performance execution and enable automatic differentiation, the system of ODEs was solved using the JAX library.

The process data was generated from three benchmark processes, namely: (i) Heat exchanger case: a shell-and-tube heat exchanger with bypass control was modeled in DWSIM to simulate the cooling of a hot process stream using a cooling-water utility. The model uses the Gnielinski correlation for heat transfer coefficients and was automated via the COM interface; (ii) Distillation column case, an ethanol-water separation unit, consisting of a 20-stage column, simulated in DWSIM using the NRTL thermodynamic property package, following the specifications described by Araújo Neto et al. (2020); and (iii) Direct Methane Aromatization Membrane Reactor (DMA-MR) case, implemented as a system of ODEs in JAX, based on the non-isothermal, non-adiabatic model and kinetic parameters provided by Alves et al. (2022).

In all cases, the input space was sampled using Latin Hypercube Sampling (LHS). This strategy ensured a uniform distribution of points across the operating range. The sampling bounds and the specific number of points (N) for each case study were chosen to balance the representation of process non-linearities with the computational time required for the initial data generation. The respective information about the size of the LHS will be presented in the final subsection of the methodology, where all the configurations employed for the three study cases are summarized.

### 2.3 Surrogate modeling and validation

The surrogate modeling workflow was implemented in Python using the scikit-learn library. To ensure numerical stability and improve convergence, both input and output datasets were standardized to zero mean and unit variance using StandardScaler. The data was randomly partitioned into training (80%) and testing (20%) sets to evaluate the models' generalization capability on unseen data.

Two regression strategies were employed based on the problem nature: for the Heat Exchanger and Distillation cases, Multi-Layer Perceptron (MLP) regressors were developed. The network architecture (number of hidden layers and neurons) and training hyperparameters (learning rate, alpha) were optimized using Optuna, a Bayesian optimization framework. The optimization process maximized the validation  $R^2$  score, resulting in architectures typically using ReLU activation functions and the Adam solver for robust convergence. For the DMA-MR case, Gaussian Process (GP) regressors were utilized to capture the non-linear relationship between reactor geometry and performance. A Radial Basis Function (RBF) kernel was employed, with hyperparameters (length scale) optimized during fitting. The trained GP models were subsequently exported as JAX-compatible functions to enable automatic differentiation. Model fidelity was assessed using the Coefficient of Determination ( $R^2$ ) and Root Mean Squared Error (RMSE) on the held-out test set. Only models achieving  $R^2 > 0.98$  were integrated into the operability analysis workflow.

### 2.4 Summary of computational configurations

The specific configurations for the three case studies are summarized in Table 1 in terms of each case, its AIS and AOS and the reference model. It is worth noting that the DMA-MR case addresses a *design operability* problem (optimizing reactor geometry), whereas the Heat Exchanger and Distillation cases focus on a *control operability* problem (optimizing process variables). To demonstrate the applicability of the proposed methodology, different surrogate models were employed: MLP for the heat exchanger and distillation column cases, and GP for the DMA-MR case. All surrogate models were trained using standardized data (zero mean, unit variance). The specific architectures and hyperparameters for the surrogate models are detailed in Table 2.

**Table 1.** Description of case studies, simulation sources, and operability variables.

Case study	Original model	Input variables (AIS)	Output variables (AOS)
Heat exchanger	DWSIM (Rigorous)	Steam flow [kg/h]	Outlet temperature [°C]
		Bypass fraction [0-1]	Thermal efficiency [%]
Distillation	DWSIM (Rigorous)	Distillate purity [mol frac]	Distillate flow [mol/s]
		Bottoms flow [mol/s]	Reboiler duty [kW]
DMA-MR	Python (ODE System)	Tube length [cm]	Benzene production [mg/h]
		Tube diameter [cm]	Methane conversion [%]

**Table 2.** Optimized surrogate model architectures and hyperparameters.

Case study	Surrogate type	Architecture / Kernel	Solver / Optimizer	Activation
Heat exchanger	MLP	4 layers x 256 neurons	Adam	ReLU
Distillation	MLP	4 layers x 256 neurons	Adam	ReLU
DMA-MR	Gaussian process	Kernel: RBF	Max. likelihood	N/A

### 3. Results and Discussion

#### 3.1 Surrogate validation and computational performance

Model validation was performed on the held-out test sets, revealing excellent agreement between surrogate predictions and ground-truth values. Quantitatively, all models achieved  $R^2$  scores exceeding 0.999 and low values of RMSE, confirming that the optimization strategies - Bayesian (Optuna) for MLPs and Maximum Likelihood Estimation for GPs - effectively represented the underlying process mappings.

Computational benchmarks indicate a reduction in evaluation time varying by model complexity. For the rigorous DWSIM simulations (heat exchanger and distillation), the surrogate approach reduced the evaluation time per point from approximately 1.5 seconds (via COM interface) to less than 100 microseconds. For the custom DMA-MR ODE model, which was already computationally efficient due to its JAX implementation, the GP surrogate provided a 4x acceleration when compared to the numerical integration of the ODEs. This performance facilitates the operability workflows, particularly the inverse mapping step, where thousands of iterative function calls required for convergence would be computationally intractable using rigorous simulations. The summary of the metrics obtained through the surrogate models and the subsequent speedup when compared to the first-principles models' runs are presented in Table 3.

**Table 3.** Statistical validation metrics (test set) and computational speedup factors.

Case study	Target output	Test $R^2$	Test RMSE	Computational Speedup
Heat exchanger	Outlet temperature	0.9996	0.165 °C	~ 20,000x (DWSIM→MLP)
	Thermal efficiency	0.9991	0.51 %	
Distillation	Distillate flow	0.9999	0.107 mol/s	~ 20,000x (DWSIM→MLP)
	Reboiler duty	0.9999	0.023 MW	
DMA-MR	Benzene production	> 0.9999	0.0047 mg/h	~ 4x (ODE→GP)
	Methane conversion	> 0.9999	0.0080 %	

#### 3.2 Operability analysis and design spaces

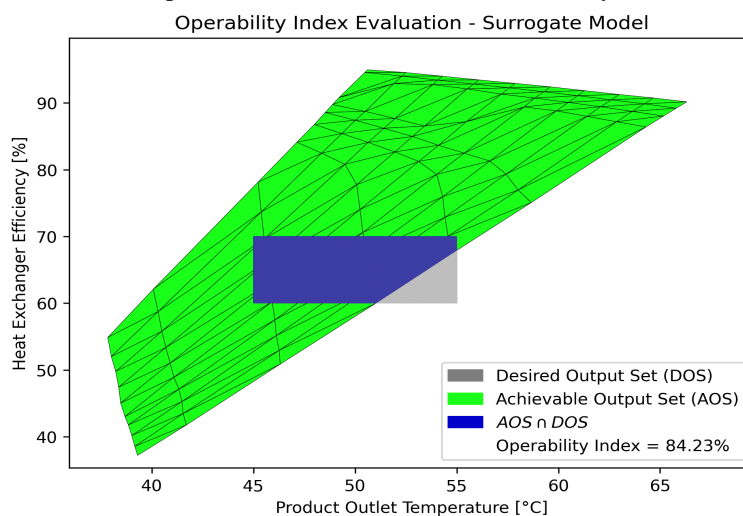
The validated surrogates were integrated into the Operability framework to compute the Achievable Output Sets (AOS) and perform inverse mapping. For the Heat Exchanger case, the analysis quantified the trade-off between Thermal Efficiency and Outlet Temperature. The generated AOS (Figure 1) indicates that while increasing the bypass fraction expands the achievable temperature control range, it reduces thermal efficiency due to lower heat transfer area utilization. The operability index calculation determined the specific steam flow range required to keep the outlet temperature within the setpoint target under input disturbances.

In the Distillation Column case, the surrogate mapped product specifications to operating conditions. The resulting operating envelope (Figure 2) captures the non-linear behavior caused by the ethanol-water azeotrope. As the distillate purity target approaches the azeotropic composition, the required reboiler duty increases

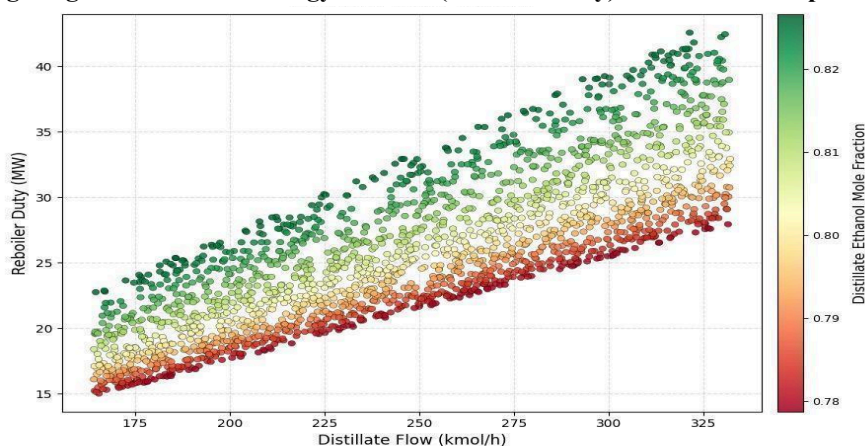
asymptotically. The inverse mapping identified the minimum energy input and reflux ratio necessary to sustain distillate purities above 0.82 (molar fraction), establishing the feasible region for high-purity operation.

Finally, the DMA-MR case applied the framework to design operability. The inverse mapping algorithm determined the reactor geometry (tube length and diameter) required to meet specific performance targets. The calculation recovered the Desired Input Set (DIS) corresponding to a benzene production greater than 10 mg/h and methane conversion above 15%. The resulting design space (Figure 3) delimits the valid geometric combinations that satisfy these production constraints.

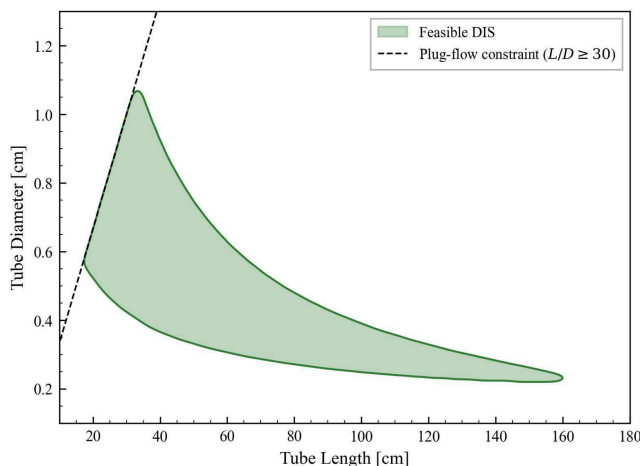
**Figure 1. Achievable Output Set (AOS) for the heat exchanger, illustrating the trade-off between outlet temperature control and thermal efficiency.**



**Figure 2. Operating envelope for the distillation column. The color gradient represents distillate purity, highlighting the increase in energy demand (Reboiler Duty) near the azeotropic region.**



**Figure 3. Inverse Design Mapping for the Membrane Reactor. The plot shows the valid Desired Input Set (DIS) of tube lengths and diameters required to meet production targets, delimiting the valid geometric combinations satisfying the production constraints, including the plug-flow flux constraint.**



#### 4. Conclusions

This study established a computational workflow integrating open-source rigorous simulation (DWSIM) with data-driven surrogates to solve process operability problems. The replacement of phenomenological models by Bayesian-optimized MLPs and Gaussian Processes reduced the evaluation time of steady-state points from the order of seconds to microseconds. This acceleration rendered the iterative inverse mapping algorithms mathematically tractable, allowing for the exhaustive exploration of design and control spaces that is impractical with direct simulation. This aspect is especially important for high-dimensional chemical engineering flowsheets.

In terms of process engineering, the analysis quantified the operational constraints for three distinct unit operations. For the heat exchanger, the Achievable Output Sets (AOS) revealed the thermal efficiency penalty required to expand the temperature control range. In the distillation column, the inverse mapping determined the asymptotic energy costs associated with high-purity specifications near the azeotropic region. Finally, the membrane reactor case validated the method for inverse design, recovering the specific tube geometry (L,D) required to meet production targets with negligible computational cost compared to the original ODE system. In the future, the proposed approach will be applied to high dimensional flowsheets exploring data sets of industrial systems.

#### Data Availability

The supporting code and data for this article can be found online at the following GitHub site: <<https://github.com/Spogis/PSEBR2026>>.

#### References

- Alves, V., Gazzaneo, V., Lima, F. V. A machine learning-based process operability framework using Gaussian processes. *Computers & Chemical Engineering*, 163, 107835, 2022.
- Alves, V., Dinh, S., Kitchin, J. R., Gazzaneo, V., Carrasco, J. C., Lima, F. V. (2024). Opyrability: A Python package for process operability analysis. *Journal of Open Source Software*, 9(94), 5966.
- Carrasco, J. C., Lima, F. V. (2017). Novel operability-based approach for process design and intensification: Application to a membrane reactor for direct methane aromatization. *AIChE Journal*, 63(3), 975-983.
- Vinson, D. R., Georgakis, C. (2000). A new measure of process output controllability. *Journal of Process Control*, 10(2-3), 185-194.
- Araújo Neto, A.P., Farias Neto, G.W., Neves, T.G (2020). Changing product specification in extractive distillation process using intelligent control system. *Neural Comput & Applic*, 32, 13255–13266.