



A machine learning-based process operability framework using Gaussian processes

Victor Alves, Vitor Gazzaneo, Fernando V. Lima*

Department of Chemical and Biomedical Engineering, West Virginia University, Morgantown, WV 26506, United States

ARTICLE INFO

Article history:

Received 9 September 2021

Revised 26 March 2022

Accepted 2 May 2022

Keywords:

Process operability

Machine learning

Kriging

Gaussian process regression

Surrogate modeling

ABSTRACT

The objective in this work is to develop a machine learning-based framework for process operability using surrogate responses based on Kriging (also known as Gaussian Process Regression). Currently, the available operability approaches for nonlinear systems are limited by the problem dimensionality that they can address, not being computationally tractable for high-dimensional systems. The proposed approach will use Kriging-based models to substitute the developed first-principles or process simulation-based models. The built surrogate models can generate responses that are comparable to the first-principles nonlinear models in terms of accuracy, while reducing the computational effort. To achieve this goal, a framework for the systematic analysis of highly nonlinear, large-dimensional systems at steady state is developed. The proposed approach is benchmarked against current operability methods and provides a new direction in the process operability field employing Kriging models. Two case studies associated with natural/shale gas conversion are addressed to illustrate the effectiveness of the proposed methods, namely a membrane reactor for direct methane conversion to fuels and chemicals and a natural gas combined cycle power plant. It is shown that the computational time for operability calculations is significantly decreased when using the developed approach, with reductions of up to four orders of magnitude, while the relative errors with respect to the output responses is below 0.3% for the worst-case scenario considering all cases. This work thus contributes to machine learning formulations and algorithms for process operability to enable the improved design, operations and manufacturing of chemical and energy systems.

© 2022 Elsevier Ltd. All rights reserved.

1. Introduction

Process Operability has been developed in the last two decades as a valuable tool for qualitatively and quantitatively assessing the design and control interface of industrial processes, subjected to expected disturbances and process constraints. Process operability has been extensively applied to steady-state systems and later extended to dynamic processes (Lima and Georgakis, 2010; Gazzaneo et al., 2020).

Since the inception of process operability concepts (Vinson and Georgakis, 2000; Georgakis et al., 2003), several contributions have been made towards addressing the inherent challenges that emerged with the input-output operability mapping of the studied processes. Such challenges include nonlinearity, high-dimensionality and input-output multiplicity of process models that are derived to represent chemical/energy processes. Particu-

larly in the field of process operability, response surface modeling (RSM) was proposed for reducing the complexity of operability calculations for high-dimensional systems (Georgakis and Li, 2010). Additionally, the operability concepts were extended for the analysis of plantwide systems by selecting production rate and product purity as the key variables of focus (Subramanian and Georgakis, 2005). More recently, a series of nonlinear programming (NLP)-based approaches were developed to evaluate the feasibility of achieving desired outputs and calculate what should be the respective inputs to accomplish this goal (Carrasco and Lima, 2015, 2017a, 2017b). In addition, these same studies (Carrasco and Lima, 2015, 2017a, 2017b) extended the operability framework to consider the concepts of process intensification and modularization, as a step forward towards using the operability tools for enabling modular manufacturing. Moreover, mixed-integer linear programming (MILP)-based methods were introduced (Gazzaneo and Lima, 2019; Gazzaneo et al., 2020) employing computational geometry concepts for evaluating the operability regions for process design and intensification. Finally, the main process operability algorithms developed for intensification and modularization were compiled

* Corresponding author.

E-mail addresses: vc00004@mix.wvu.edu (V. Alves), viggazzaneo@mix.wvu.edu (V. Gazzaneo), fernando.lima@mail.wvu.edu (F.V. Lima).

into an open-source Operability App in MATLAB (Gazzaneo et al., 2020) with a user-friendly interface for easy dissemination of the process operability algorithms.

Despite of these past contributions to the process operability field, the challenge regarding tackling nonlinear problems with high-dimensionality using first-principles models (Carrasco and Lima, 2017b, 2018) still remains. This task becomes computationally intractable as it grows in complexity with problem dimensionality, creating the need of recurring to parallel computing (Carrasco and Lima, 2018), an approach that is not always readily available and highly dependent on the computational infrastructure and modeling platform/numerical package used by practitioners or academic researchers. For such challenging high-dimensional applications, the idea of substituting the nonlinear first-principles process model by a surrogate model can be appealing to perform the operability computations in a more efficient manner. In particular, the NLP-based operability approaches (Carrasco and Lima, 2017a, 2017b, 2018) for process design and intensification could benefit of computational time reductions enabled by machine learning-based methods. Moreover, communication challenges have also been reported between process simulation platforms (e.g., Aspen Plus) and numerical packages that are required to perform process operability calculations (e.g., in MATLAB) (Carrasco and Lima, 2018).

To address these challenges, there is a critical need to develop a systematic approach for assessing process operability that has the following features: (i) surrogate modeling of the processes studied to be able to solve high-dimensional and nonlinear models, while maintaining accuracy; (ii) synergy or integration with current nonlinear programming (NLP)-based (Carrasco and Lima, 2017b) and multimodel-based approaches for process operability (Gazzaneo and Lima, 2019); and (iii) facilitated communication between surrogate models developed based on process simulators and numerical packages, using the same computing platform for both (e.g., MATLAB or Python). The use of supervised machine learning-based algorithms known as Kriging (or Gaussian Process Regression) is proposed here as a surrogate to the first-principles models in process operability. Although the implementation of Kriging in the field of Chemical Engineering is not new, having applications ranging from chemical reaction engineering (Maceiczky and deMello, 2014), feasibility analysis (Boukouvala and Ierapetritou, 2012), process optimization (Davis and Ierapetritou, 2007), pharmaceutical processes (Boukouvala et al., 2010), modular flowsheet optimization (Quirante et al., 2015; Caballero and Grossmann, 2008), to Self-Optimizing Control (Alves et al., 2018), this proposed direction for process operability has not yet been reported. This work is structured as follows: Section 2 discusses process operability and Kriging main concepts; Section 3 addresses the proposed approach, and Section 4 shows the application of the proposed approach to two case studies of increased complexity and dimensionality; lastly, Section 5 contains conclusions and suggestions for future developments.

2. Previous work: process operability and Kriging

2.1. Process operability

Process operability has emerged as a viable alternative to the sequential tasks of assessing process design and control, by integrating both tasks in the early design phase of industrial processes (Lima and Georgakis, 2010; Gazzaneo et al., 2020). To perform this task, process operability tools were developed to quantify achievability of process control objectives, given the available limits imposed on the input variables, while considering process constraints and expected disturbances that may occur dur-

ing process operations (Vinson and Georgakis, 2000). In this section, process operability definitions for the case where the disturbances are kept at their nominal values are considered, which corresponds to the main scope of this work. For a complete and in-depth discussion on the previous operability concepts, including the presence of disturbances acting on the process, one must refer mainly to Vinson and Georgakis (2000), Georgakis et al. (2003), Lima et al. (2010), Gazzaneo et al. (2020).

To perform the operability analysis, the main requirement is that a process model that describes the relationship between the input (manipulated and/or disturbance) and output variables should be available (Georgakis et al., 2003). A process model M with m inputs, p outputs, q disturbances and n states, can be defined as in Eq. (1):

$$M = \begin{cases} \dot{x}_s = f(x_s, u, d) \\ y = g(x_s, u, d) \\ h_1(\dot{x}_s, x_s, y, \dot{u}, u, d) = 0 \\ h_2(\dot{x}_s, x_s, y, \dot{u}, u, d) \geq 0 \end{cases} \quad (1)$$

In which $u \in R^m$ are the inputs, $y \in R^p$ are the outputs, $d \in R^q$ are the disturbances and $x_s \in R^n$ are the state variables. Also, f and g are nonlinear maps and h_1 and h_2 correspond to equality and inequality process constraints, respectively. Based on the process model and associated variables, operating spaces/sets were established for the operability calculations and quantification of the Operability Index (OI) (Georgakis et al., 2003). These sets are summarized with proper descriptions and mathematical formulations in Table 1. Refer to Refs. Georgakis et al. (2003), Vinson and Georgakis (2000) for further details on these sets and other operability definitions. Fig. 1 illustrates the main process operability spaces and the definitions that arise. In Fig. 1 (A), the lower and upper bounds for the manipulated/design variables are defined within the AIS. By evaluating these inputs through the process model (M), the outputs are calculated generating the AOS (B). From the desired operation with respect to such outputs, the DOS can be defined (C). The feasible portion of the DOS that is contained within the AOS is specified as the DOS* (C). Lastly, the DIS can be obtained via an inverse mapping (M^{-1}), in which its intersection with the AIS is defined as the DIS* (D).

Especially for process design and intensification purposes, it is important to note that the task of obtaining the inverse model/mapping (M^{-1}) is of paramount importance. As stated above, high-dimensionality, non-linearities, input-output multiplicities, and additional challenges such as infeasibility of the desired targets (Gazzaneo et al., 2020) are the most common issues to be faced when performing this task. In order to circumvent these issues, linear and nonlinear programming-based operability methods were successfully developed (Gazzaneo and Lima, 2019; Carrasco and Lima, 2017b), respectively. Focusing on the latter, denominated NLP-based approach, further definitions regarding the feasibility of the desired inputs and output sets were created. Instead of analytically calculating (M^{-1}), Carrasco and Lima (2017b) proposed an NLP-based optimization problem to calculate the discretized inputs from a specified DOS, using an objective function defined by relative error minimization. More specifically, the approach consists of finding a Feasible Desired Input Set (DIS*) that results in a Feasible Desired Output Set (DOS*), with the latter being optimized to be as close as possible from the original DOS. For the optimization problem of the NLP-based approach, the error minimization function can be assembled considering the process required constraints, intensification/efficiency targets and solved sequentially or using a bilevel programming approach, as showed by Carrasco and Lima (2017b). The possible optimization problems of this approach are adapted in this work for the use of Kriging responses and are discussed in depth in Section 3.

Table 1
Steady-state operability sets: definitions and mathematical formulations.

Operability Set	Description	Mathematical Formulation
Available Input Set (AIS)	Manipulated inputs ($u \in R^m$) based on the design of the process that is limited by the process constraints (Vinson and Georgakis, 2000).	$AIS = \{u \mid u_i^{\min} \leq u_i \leq u_i^{\max}, 1 \leq i \leq m\}$
Expected Disturbance Set (EDS)	Disturbance variables ($d \in R^q$) that can represent process uncertainties and variabilities.	$EDS = \{d \mid d_i^{\min} \leq d_i \leq d_i^{\max}, 1 \leq i \leq q\}$
Achievable Output Set (AOS)	Range of the outputs ($y \in R^n$) that can be achieved using the inputs inside the AIS.	$AOS(d) = \{y \mid y = M(u, d); u \in AIS, d \text{ is fixed}\}$
Desired Output Set (DOS)	Production/target/efficiency requirements for the outputs that do not necessarily meet the ranges of the AOS.	$DOS = \{y \mid y_i^{\min} \leq y_i \leq y_i^{\max}, 1 \leq i \leq n\}$
Desired Input Set (DIS)	Set of inputs required to reach the entire DOS, given a disturbance vector d .	$DIS(d) = \{u \mid u = M^{-1}(y, d); y \in DOS, d \text{ is fixed}\}$
Feasible Desired Output Set (DOS*)	Feasible set of desired outputs calculated via relative error minimization from the DOS using for example the NLP-based approach (Carrasco and Lima, 2017b).	$DOS^* = \{y^* \mid y^* = M(u^*); u^* \in DIS^*\}$
Feasible Desired Input Set (DIS*)	Optimal set of inputs that are required to obtain the Feasible Desired Output Set (DOS*), calculated for example via the NLP-based approach.	$DIS^* = \{u^* \mid u^* = (u_1^*, u_2^*, u_3^* \dots u_l^*)\}$ $i \stackrel{\text{def}}{=} \text{Discretized DOS grid size}$

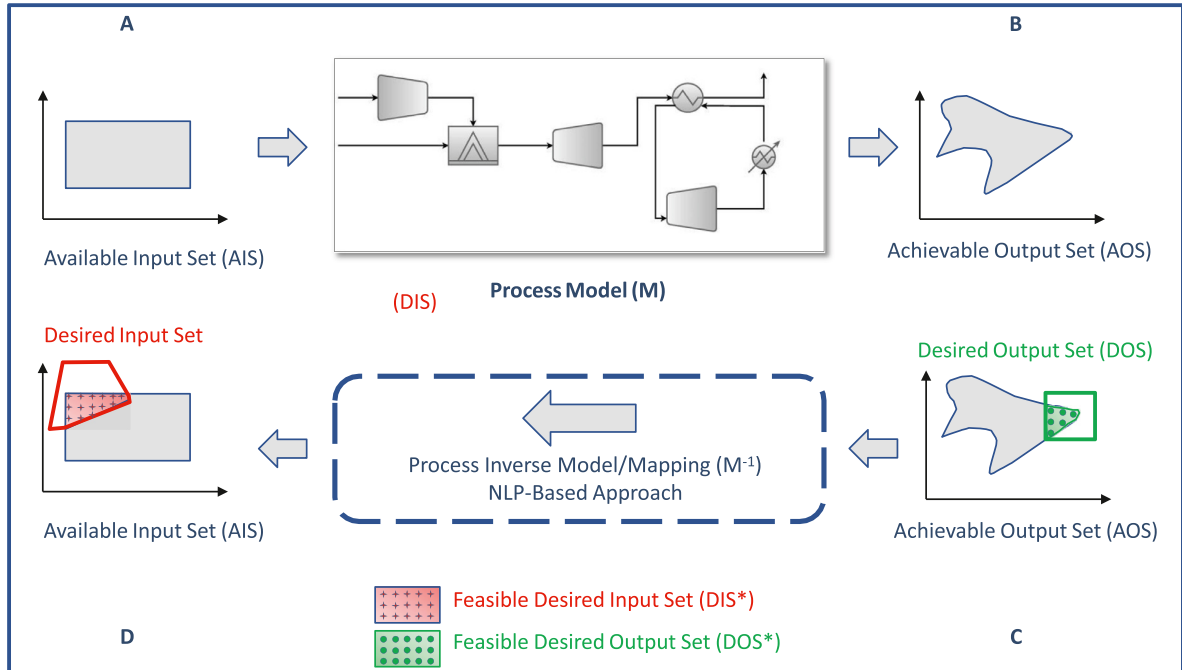


Fig. 1. Visual exploration of main Process Operability sets and definitions.

With the input-output sets defined above, the Operability Index (OI) can be calculated as shown in Eq. (2). A process is considered fully operable when the OI is 1 and if it is less than 1, some regions of the DOS are not achievable (Lima and Georgakis, 2010).

$$OI = \frac{\mu(AOS \cap DOS)}{\mu(DOS)} \quad (2)$$

In which μ indicates the measure of regions, varying depending on the dimensionality of the considered sets, for example length for 1D systems, area for 2D systems, volume for 3D systems and hypervolumes for systems of higher dimensionality (Gazzaneo et al., 2020).

2.2. Surrogate modeling: Kriging (Gaussian process regression)

Kriging models are employed in this work as a surrogate to the process first-principles model or simulation-based models. The Kriging implementation in MATLAB used in this article is based on the work of Lophaven et al. (2002a, 2002b) and similar approaches can be found in Caballero and Grossmann (2008); Quirante et al. (2015). For historical purposes, the

works by Sacks et al. (1989) and Forrester et al. (2008) are also recommended. For a thorough discussion on the subject, refer to Rasmussen and Nickisch (2010), Jones (2001), Forrester et al. (2008).

Considering a set of g experiments $S = [u_1 \dots u_g]^T$, with $u_i \in R^m$ (input variables), and an output vector $Y = [y_1 \dots y_g]^T$, a Kriging metamodel is capable of representing a nonlinear function $Y(x) \in R^p$ with the aid of two terms: a regression ($\mathcal{F}$) and a stochastic function (z) as shown in Eq. (3).

$$\hat{Y}_l(u) = \mathcal{F}(u) + z_l(u), \quad l = 1, \dots, p \quad (3)$$

The regression model is considered as a linear combination of functions $f_j : R^m \rightarrow R$, and typically the regression functions used are polynomials of orders zero to two (Lophaven et al., 2002b; Forrester et al., 2008). Additionally, it is assumed that each z_l has zero mean with covariance between any two given points, u and u' as in Eq. (4).

$$\text{Cov}[z_l(u), z_l(u')] = \sigma_l^2 \mathcal{R}(\theta, u, u'), \quad l = 1, \dots, p \quad (4)$$

In which σ^2 is the process variance for the l^{th} output and $\mathcal{R}(\theta, u, u')$ is a correlation function. There are several correlation

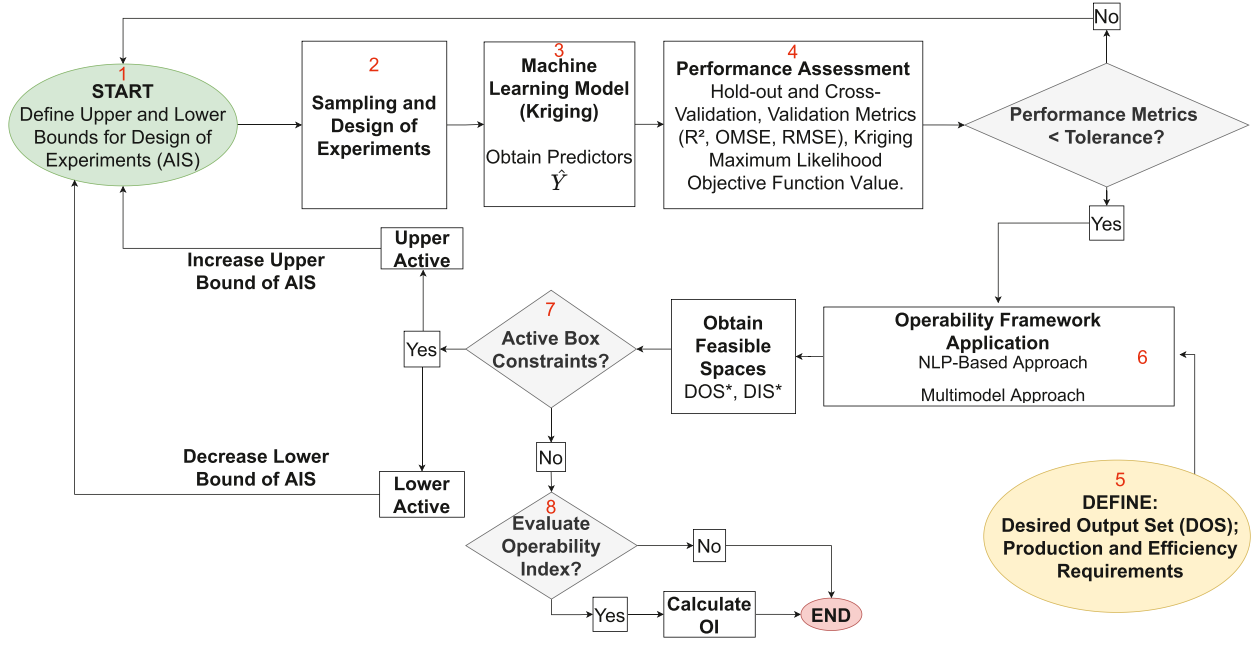


Fig. 2. Flowchart with the proposed method steps.

functions developed for Kriging models in the literature (Lophaven et al., 2002a; 2002b; Rasmussen and Williams, 2008). For the sake of simplicity and given the particular characteristic of being continuously differentiable, in this work, the correlation function assumed is of the form known as squared-exponential or Gaussian form ($p_j = 2$) as shown in Eq. (5).

$$\begin{aligned} \mathcal{R}(\theta, u, u') &= \prod_{j=1}^m \mathcal{R}_j(\theta, u_j - u'_j), (\theta \geq 0) \\ \mathcal{R}_j(\theta, u_j - u'_j) &= \exp\left(-\theta_j(u_j - u'_j)^{p_j}\right), (p_j = 2) \end{aligned} \quad (5)$$

In which θ are defined as hyperparameters and their values can indicate if the inputs are highly correlated or not (Alves et al., 2018) and also how fast the correlation goes to zero as the process moves in the j^{th} coordinate direction (Caballero and Grossmann, 2008). The parameter p_j represents the smoothness of the correlation (Forrester et al., 2008), and reducing its value increases the rate at which the correlation initially drops as the distance between two given points $(u_j - u'_j)$ increases (Forrester et al., 2008).

When $p_j \approx 0$, there is a discontinuity between $\hat{Y}(u_j)$ and $\hat{Y}(u'_j)$ (Lima et al., 2020), excluding the possibility of correlation between two given arbitrary points. Finally, the determination of the hyperparameters θ is a result of an optimization problem with the optimal solution corresponding to the maximum likelihood estimation (Lophaven et al., 2002b), where $|R|$ is the determinant of the correlation matrix R , as in Eq. (6).

$$\min_{\theta} \left\{ \psi(\theta) \equiv |R|^{\frac{1}{2}} \sigma^2 \right\} \quad (6)$$

3. Proposed approach: Kriging-based process operability

The proposed method is depicted in Fig. 2 and will be discussed step-by-step in this section. The main objective is to develop a systematic and generic method capable of performing prior process operability calculations (Carrasco and Lima, 2017b; Gazzaneo and Lima, 2019) while maintaining accuracy and reducing computational effort. In addition, the proposed method should be able to tackle systems of any dimensionality that present strongly nonlinear behavior, challenges that can be addressed with the help of the

Kriging models. Moreover, this method should facilitate the employment of the process operability algorithms, avoiding the need of coupling process modeling tools with numerical packages, as the surrogate model can be used instead of a first-principles/process simulator model.

The following steps outlined in Fig. 2 are detailed below for the developed method:

- 1. Definition of upper and lower bounds of input variables (Available Input Set):** In this initial step, one must define the bounds of the input variables (design and/or manipulated variables). These limits should be as representative to the process as possible, as the Kriging responses created will generate the output data according to the points within the limits defined. Thus, careful selection and inspection of the AIS bounds must be done, in order to create a sampling space that represents the operating space of the model/process and avoid potential infeasibilities (e.g., lack of convergence of the process model). The AIS bounds selection should be based on knowledge about the process/model, such as design limitations and/or manipulated variable operating regions;
- 2. Sampling:** Using the limits defined in the previous step, a sampling technique is employed in order to generate the input data. Space-filling techniques such as Latin Hypercube Sampling (used in this work) are recommended as they have been widely tested in the literature for different applications (Quirante et al., 2015; Alves et al., 2018; Lima et al., 2020). In addition, outliers must be investigated here and deleted from the sampling set if they are not meaningful (e.g., infeasible cases from process simulator runs), in order to ensure that the Kriging responses are generated in the next step with only meaningful data;
- 3. Surrogate model (Kriging) generation:** In this step, the Kriging responses for the p outputs are generated using the input-output data from the previous step. In this work, the DACE (Design and Analysis of Computer Experiments) MATLAB toolbox is used (Lophaven et al., 2002a). However, different implementations may also be employed such as the "scikit-learn" Python package (Pedregosa et al., 2011), which is based mainly on the work of Rasmussen and Williams (2008). Other implementations that could be considered are the MATLAB *fitrgp* func-

tion and the *Gaussian Processes for Machine Learning* toolbox (Rasmussen and Nickisch, 2010). This step might have the highest computational expense since the generation of Kriging surrogates are known to have times up to $O(n^3)$, in which n is the sample size, due to matrix inversion operations needed. Therefore, the required time to generate the Kriging surrogates can increase quickly with the sample sizes. However, the expectation is that this time is still significantly less than using the corresponding nonlinear, first-principles model for process operability calculations;

4. **Performance Assessment:** After generating the surrogate responses, performance assessment must be carried out in order to ensure the robustness and precision of the metamodels. Several metrics can be employed in this step, such as: R^2 fitness, root mean-squared error (RMSE), overall mean-squared error (OMSE), hold-out and cross validation. Hold-out validation is useful for large datasets and cross-validation can be used for smaller ones without being time consuming. Cross-validation assessment is a more effective approach (Caballero and Grossmann, 2008), as the training set is split into k folds and different responses are generated and validated, being able to inspect outliers more effectively. A special case of cross validation is when $k = 1$ and the number of folds corresponds to the size of the sample: “leave-one-out” cross validation, which can be employed when the size of the sampling is not extremely large. In addition, for this particular study, the maximum likelihood function that is optimized at the DACE toolbox can also be used as a metric. Following Lima et al. (2020), the threshold for an acceptable surrogate representation should be below 10^{-5} for the maximum likelihood estimation function value. If the chosen performance assessment metrics are below the defined tolerances, the user can go to the next step and employ the surrogate models in the operability framework.
5. and 6. **Definition of Desired Output Set and operability framework calculations:** At this step, the desired operation is defined (i.e., Desired Output Set (DOS)) and which problem type is to be solved, following Carrasco and Lima (2017a). Depending on the process, it is of particular interest finding an operable region regarding the manipulated variables that ensure that the DOS (or part of it) is satisfied, finding an optimal design using design variables towards process intensification, or even both cases. This problem definition is split into layers of target-problems (Carrasco and Lima, 2017a). In this work these problems are reconsidered, but now using Kriging-based models: **P1 - Relative error minimization between desired (DOS) and feasible (DOS*) output sets:** In this problem, the target is to calculate the DIS^* (Desired Feasible Input Set) that minimizes the relative error between the what is feasible (DOS*) and what is desired (DOS) in terms of output spaces:

$$\begin{aligned} \theta_k &= \min_{u_k} \sum_{j=1}^n ((\hat{y}_{j,k} - y_{j,k}^*) / \hat{y}_{j,k})^2 \\ \text{s.t. : Surrogate responses} \\ u_k^{\min} &\leq u_k \leq u_k^{\max} \\ \hat{c}_1(u_k^*) &\leq 0 \end{aligned} \quad (7)$$

One must choose this problem when it is of particular interest defining a feasible input set region for the process studied for optimal design purposes. This problem can give insights on what should be the best operable region for a particular target or the best design for a particular process (Carrasco and Lima, 2017a). In Eq. (7), $\hat{y}$ represents a Kriging surrogate prediction of the actual variable y , which was previously calculated using the nonlinear model of the process. In addition, if there are any nonlinear constraints, these can be also substituted by surrogates $\hat{c}$. Lastly, u_k represents the input variables for the k^{th}

element of the DOS selected points according to a specified grid (Carrasco and Lima, 2015).

P2 - P1 + Process intensification and/or efficiency target maximization: This problem adds at the final step of P1 the evaluation of a process intensification or efficiency target. This target can be set as footprint reduction, efficiency maximization, or even pollutant emission reduction, for instance. The mathematical representation of this problem is as follows:

$$\begin{aligned} \Omega &= \min_{u_{PI}^*} [PI_{\text{target}}] \\ \text{s.t. : } u_{PI}^* &\in DIS^* \\ \hat{y}_{PI} &\in DOS^*(\hat{y}) \end{aligned} \quad (8)$$

In which Ω represents the function value of the inputs that will intensify the process while ensuring the level of performance desired by the DOS^* .

P1 and P2 can also be solved simultaneously in a bilevel programming approach (Carrasco and Lima, 2018), in conjunction with parallel programming techniques seeking reduction of computational time. However, due to the employment of the Kriging-based models as surrogates to the first-principles nonlinear models, this work did not need to use any parallel or bilevel programming techniques, simplifying the software infrastructure. Moreover, note that all target problems developed (Carrasco and Lima, 2017a) are reformulated to use Kriging models instead of nonlinear, first-principles models. Lastly, it should be noted that the NLP-based approach can only yield one set of solutions for the DIS^* , but more than one set of such solutions may exist. When multiple inputs lead to the same output (i.e., for processes with input-output multiplicities), the solution of the inverse mapping in the NLP-based approach may be influenced by the choice of optimization solver as well as the solver's initial guess. However, as long as the inverse mapping of the DIS^* belong to a bijective sub-region of the input space, the proposed approach does not encounter this challenge. For the case studies performed in the paper, a modified version of the Nelder-Mead Simplex algorithm to accept nonlinear constraints and bounds is used as a derivative-free solver in MATLAB.

- 7 **Active constraint handling:** At this step, one must inspect if there are any lower/upper bound constraints (i.e., inverse-mapped input variables at the lower/upper bound of the AIS). If there are any upper bound active constraints, the upper bounds of the AIS can be increased if possible to further explore more regions, in case these do not correspond to hard/physical constraints (e.g., equipment/safety/operation bounds). The same holds if there are any lower bound active constraints. The limits can then be changed if needed and the user taken back to Step 1; Otherwise, the user can move on the next step;
- 8 **Evaluate OI (Operability Index):** After obtaining all the relevant operability sets (DIS^*/DOS^*) leading to insights on operability regions, the Operability Index (OI) is evaluated using the regions obtained. This evaluation was discussed thoroughly in the past by several authors such as Georgakis et al. (2003), Gazzaneo and Lima (2019), Gazzaneo et al. (2020). The objective in this step is to rank competing designs and/or control structures, depending on the application of interest, using the OI as a metric.

4. Case studies

The case studies below are performed to test the proposed method. The classical operability approach, using first-principles nonlinear process models is run and the Kriging-based approach is validated against it. The objective is to show that the computational time employing the developed approach is significantly reduced while the accuracy of the calculations is maintained. All case

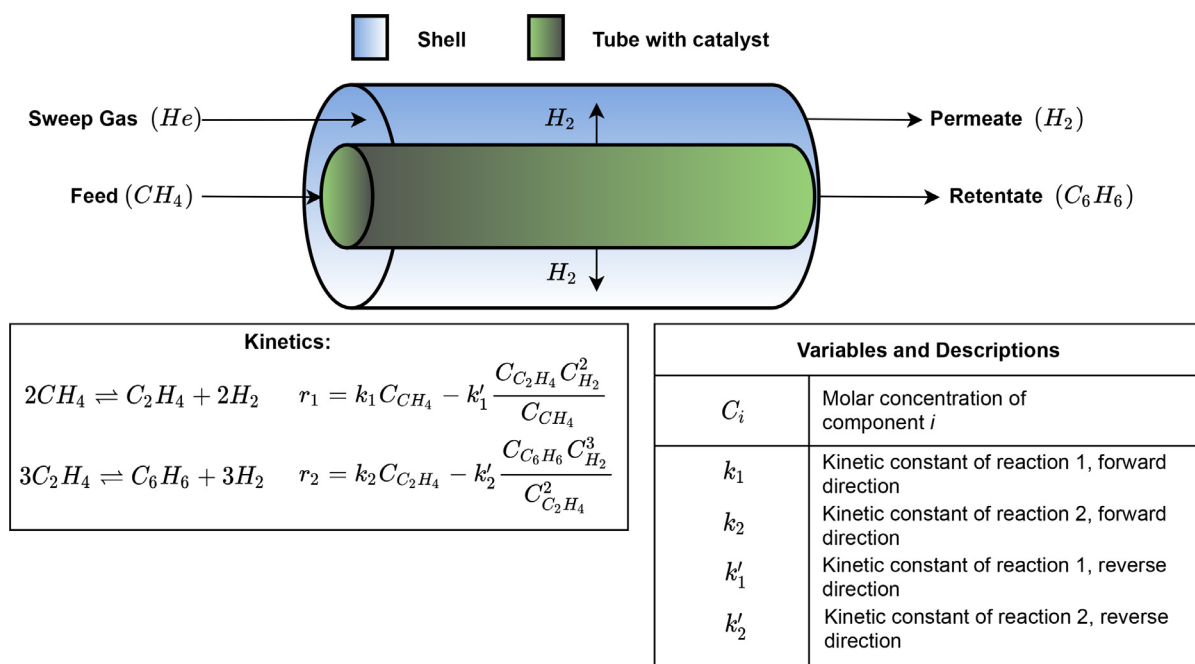


Fig. 3. DMA-MR schematic.

studies were performed on a PC with the following configuration: 16GB RAM, Intel @Core i7 4770 processor. To accurately compare the proposed method with the previous calculations employing nonlinear first-principles models or process simulators, such comparison was made under the same conditions for each problem, including NLP-solver tolerances, initial estimates and for the exact same DOS grid generated using the Latin Hypercube algorithm.

4.1. Direct Methane Aromatization Membrane Reactor (DMA-MR)

The first example consists of the steady-state model of the Direct Methane Aromatization Membrane Reactor (DMA-MR) for hydrogen and benzene production. This process has been studied as an application of process intensification and modularization of natural/shale gas processing (Carrasco and Lima, 2015; 2017a).

The process schematic is depicted in Fig. 3, in which a shell and tube membrane reactor is used for the methane conversion to hydrogen and ethylene, followed by ethylene conversion to benzene and additional hydrogen in a two-step reaction mechanism. Refer to Carrasco and Lima (2015, 2017a) for the full set of differential equations and kinetic parameters that represent the developed first-principles nonlinear process model in MATLAB.

The first application example consists of a 2 x 2 (outputs x inputs) subsystem associated with this DMA-MR process. Despite of the low system dimensionality in this first example, this process model is highly nonlinear due to the presence of the nonlinear reaction kinetics and the membrane flux in the intensified unit. Such nonlinearities make the inverse operability mapping task challenging, and thus this system serves as a good candidate for benchmarking the computational time and accuracy of the proposed methods, when compared to the results of the nonlinear operability methods that are readily available (Carrasco and Lima, 2015; 2017a).

For this application, a Kriging model is created using as input variables the tube diameter and tube length, corresponding to the variables in the AIS. Also, for the DOS, the variables considered are the methane conversion (X_{CH_4}), and the production rate of benzene ($F_{C_6H_6}$). These variables are selected similarly to previous literature for benchmark purposes (Carrasco and Lima, 2017a). The assumed

Table 2

AIS sampling limits for DMA-MR.

	Lower bound	Upper bound
Tube length [cm]	10	300
Tube diameter [cm]	0.1	2

Table 3

Likelihood objective function values.

Variable	ψ
$F_{C_6H_6}$	1.1984×10^{-6}
X_{CH_4}	3.3259×10^{-6}

limits for the AIS are also the same from literature (Carrasco and Lima, 2017a), except for the upper limit for the tube length, which is chosen based on the maximum experimental tube length possible of 300 cm. The limits used for the Kriging sampling procedure can be seen in Table 2.

With the input limits for the design of experiments being defined, 2000 samples are generated using the Latin Hypercube Sampling (LHS) technique and each one of the generated cases are run employing the first-principles nonlinear model for calculating the outputs. With the obtained input-output data, 200 cases were left out for validation purposes and thus the training set consisted of 1800 cases. The chosen number of samples was based on successive performance assessments of the Kriging maximum likelihood objective function and by also using hold-out validation until precise responses for both output variables were achieved (i.e., in this case, the value of the maximum likelihood objective function for the Kriging predictor was below 10^{-5} and $R^2 \geq 0.999$). The final hold-out validation predictions for this case study are depicted in Fig. 4 and the value for the optimized DACE objective function (ψ) likelihood that gives the optimal hyperparameters is given in Table 3.

After performance assessment, the metamodels are then employed to obtain the inverse mapping for operability and subsequently, the DIS*-DOS* calculations. The same DOS from previous work (Carrasco and Lima, 2017a) was employed in this task,

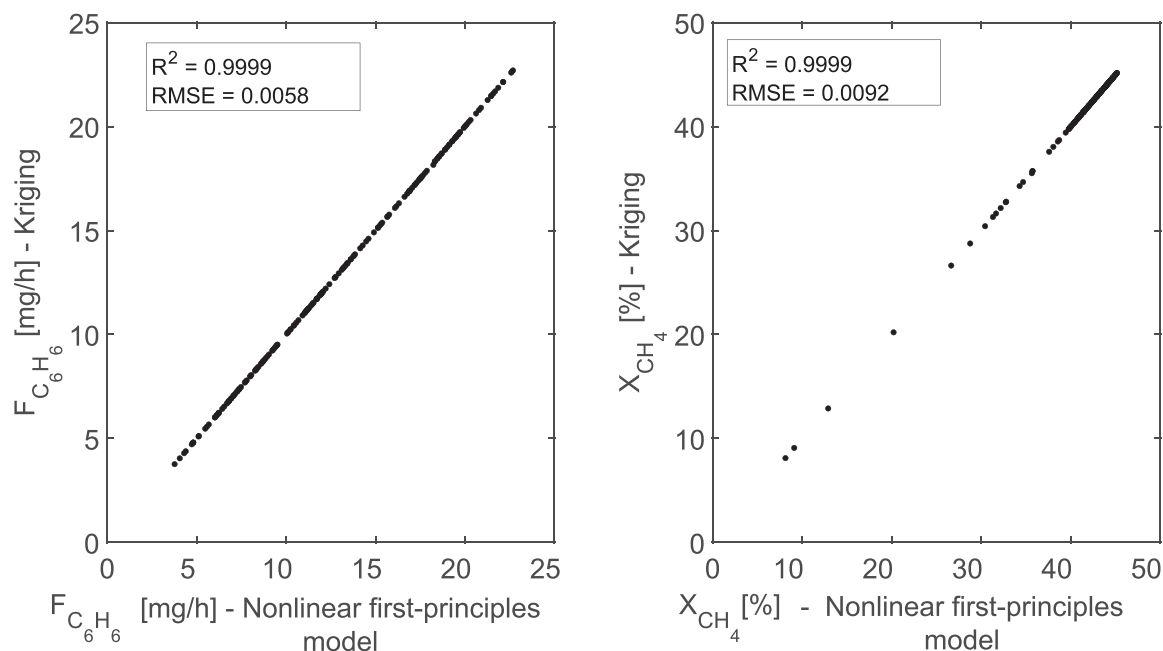


Fig. 4. Hold-out validation for DMA-MR Kriging models.

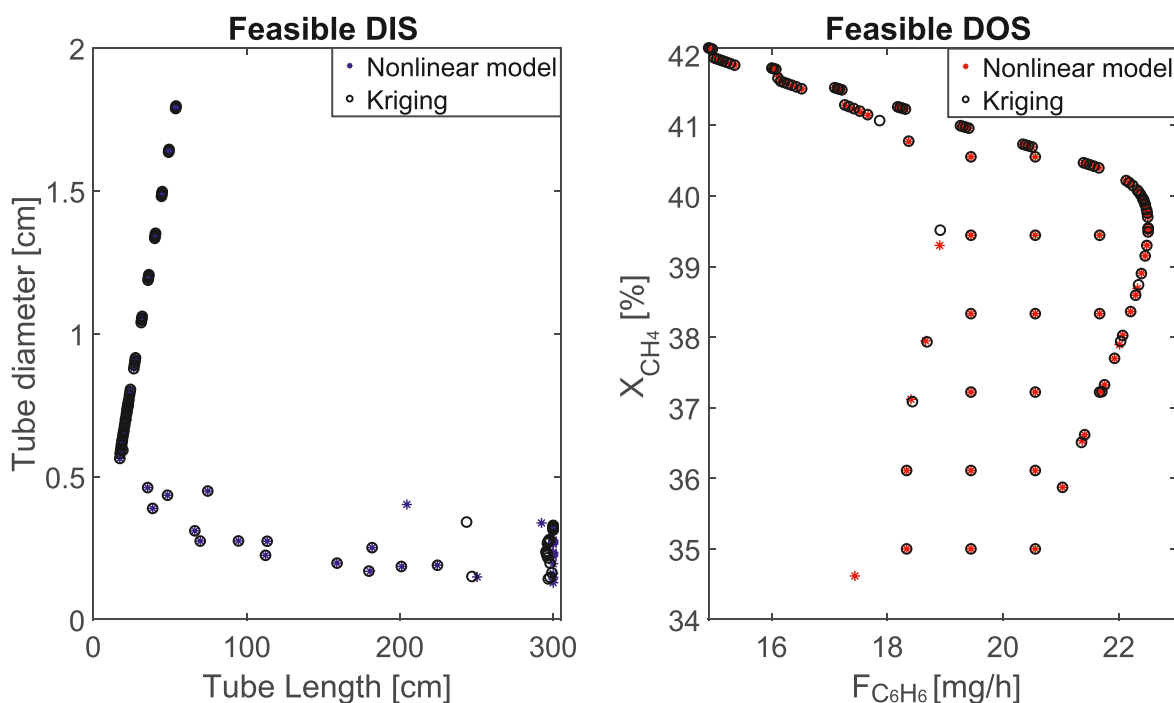


Fig. 5. DIS* and DOS* comparisons: calculations employing Kriging metamodel versus original first-principles nonlinear model.

with limits of 15–25 [mg/h] and 35–45 [%] for the production rate of benzene and conversion of methane, respectively. An additional constraint was imposed to the inverse mapping optimization problem to ensure reactor plug flow operation (e.g., $L/D \geq 30$). For comparison purposes, both the NLP-based approach for inverse mapping and the multimodel approach for OI calculation are employed using the nonlinear first-principles and the Kriging-based models, as shown in Figs. 5 and 6, in which both tasks were performed with accuracy. Note from these figures the small error obtained when comparing the original nonlinear first-principles model against the Kriging-based model responses for operability. Table 4 shows the relative errors between these model responses

Table 4

Relative errors between nonlinear-based and Kriging-based NLP operability approaches for calculating intensified design variables.

Intensified design variable	Nonlinear model	Kriging	Relative error [%]
Tube length [cm]	17.2283	17.2205	0.0454
Tube diameter [cm]	0.5634	0.5636	0.0388

for the membrane reactor modular design, confirming the agreement between the responses. Lastly, Table 5 depicts the average relative error of all DOS* points calculated using the Kriging-based approach, when compared to the first-principles model. Again,

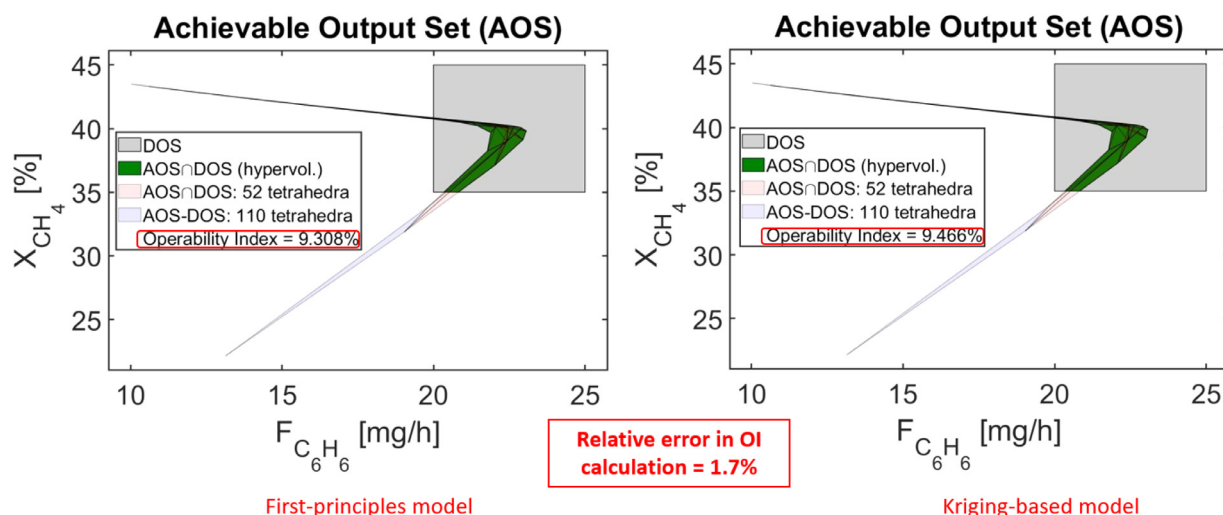


Fig. 6. OI calculations using first-principles model versus Kriging-based model. Small relative error of 1.7% is obtained between calculations.

Table 5

Relative average error for all DOS* grid points: 2x2 DMA-MR.

	Relative average error [%]
X_{CH_4}	0.0367
$F_{C_6H_6}$	0.2013

Table 6

AIS sampling limits for NGCC plant.

	Lower bound	Upper bound
Natural gas flowrate [tonne/h]	0.5	1.5
Steam flowrate [tonne/h]	0.4	4

note that the relative errors are small, indicating the accuracy of the developed method.

The overall computational time for obtaining the results in Tables 4 and 5 when using the first principles nonlinear model for this 2x2 system was of 5 min 38s, and it has been shown previously that the computational time for this system is expected to grow exponentially with problem dimensionality (Carrasco and Lima, 2018). However, the overall computational time associated with the Kriging-based method (generation of 2000 points for the input-output mapping, fitting the Kriging responses and running the optimization problem) was of 58 s. This represents a decrease of 5.8 times in computational time. If this computational time evaluation is done considering only the process operability calculations when using the Kriging-based approach against its nonlinear model-based counterpart, this difference becomes even larger. The required time to obtain the operability sets using the Kriging-based approach is of only 0.0477s against the 5 min 38s previously mentioned for the nonlinear model-based method, thus reducing the computational time by about four orders of magnitude. For comparison purposes, a previous algorithm developed employing Mixed-Integer Linear Programming (MILP) for process operability calculations (Gazzaneo et al., 2018) achieved reduction of computational time of 3.051 times and relative error of 1.28% with respect to the membrane area calculated using the tube length and diameter as input variables, considering the same case study with the same dimensionality as in this work. Hence, the proposed approach in this work has the potential to enable operability calculations for high-dimensional systems that would not be possible otherwise due to intractable computational times. Also, the results above show that the accuracy of the proposed method for generating the input-output mappings for operability analysis was not compromised, while the complexity of calculating such inverse mappings was reduced.

4.2. Natural Gas Combined Cycle (NGCC) plant

For the second case study, the NGCC plant (Carrasco and Lima, 2017b) depicted in Fig. 7 is considered as an example of employing process simulators (in this case, Aspen Plus) as the source of input-output data for the proposed method. This case can be of relevance when the user wants to take advantage of the large thermodynamic database and built-in models from the simulators that can be linked to the proposed framework. All of the equipment and stream conditions were based on the literature (Carrasco and Lima, 2017b), except when mentioned otherwise. Moreover, the NGCC plant is also an interesting example of applications that are capable of processing natural gas in the US with the potential for modular manufacturing (Carrasco and Lima, 2017b). The 2x2 and 3x8 (outputs x inputs) cases are addressed for the NGCC process in order to test the proposed method.

4.2.1. NGCC plant (2x2 case)

The considered input variables for the first NGCC plant case study are the natural gas flowrate and the steam flowrate used in the steam turbine. The air flowrate is considered based on a ratio of the natural gas flowrate following the literature (Carrasco and Lima, 2017b). The Net Plant Power [MW] and Plant Efficiency [%] are considered (Carrasco and Lima, 2017b) as output variables as the analysis of these variables may lead to process intensification and maximization of efficiency targets. The limits for the aforementioned variables are given in Table 6.

With the limits considered for the input variables, 100 samples are generated using the Latin Hypercube sampling technique. Additionally, a COM/OLE communication between MATLAB and Aspen Plus is employed to run the cases. As the number of points is not prohibitive for this case, a “leave-one-out” validation is performed. The results of the leave-one-out validation are depicted in Fig. 8. Note from this figure the good fit of the Kriging model, with relative errors of orders of 10^{-5} and 10^{-3} with respect to the Kriging outputs. Table 7 also indicates that the optimal hyperparameters for the Kriging responses built give an accurate prediction, as the

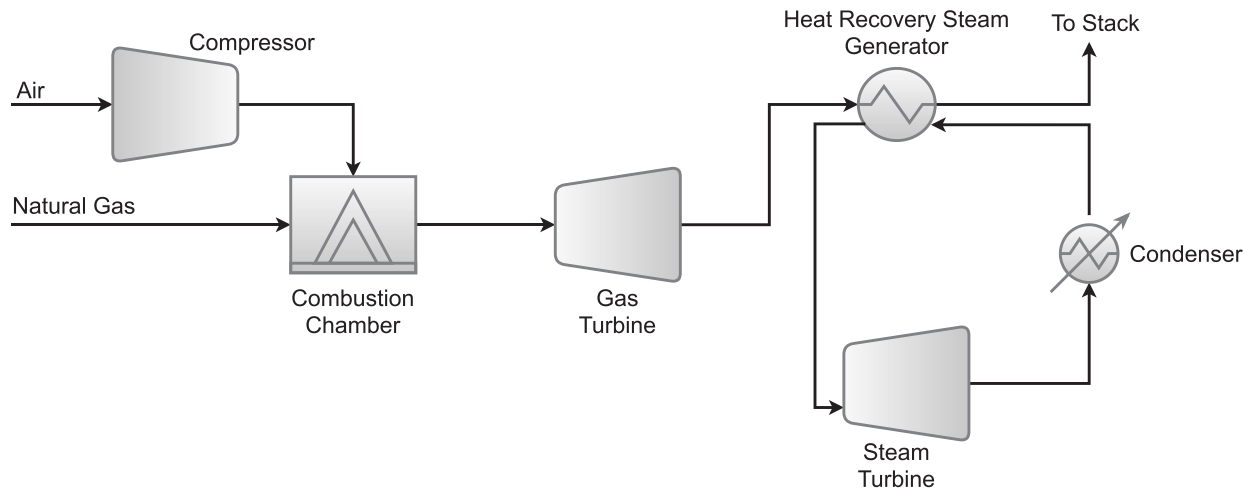


Fig. 7. NGCC simplified process flow diagram based on (Carrasco and Lima, 2017b).

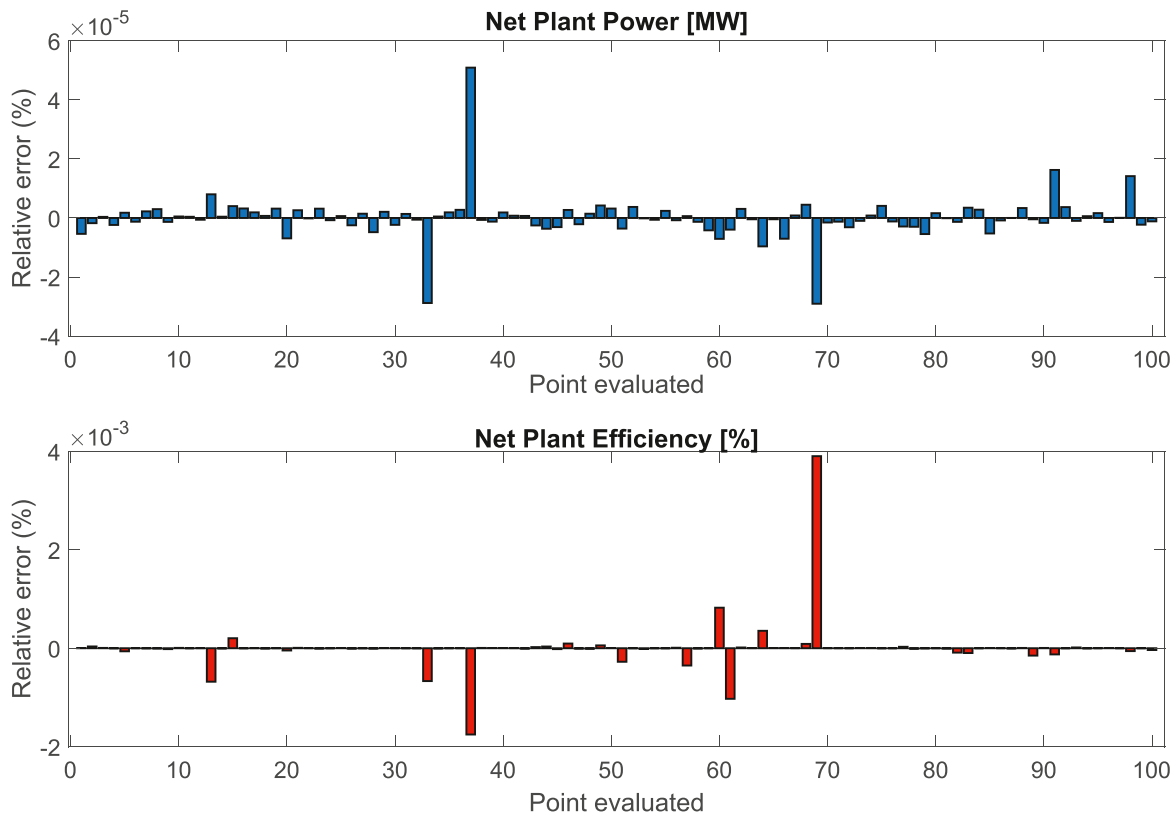


Fig. 8. NGCC leave-one-out cross-validation for output variables.

Table 7
Likelihood objective function value for 2x2 NGCC Plant.

Variable	ψ
Net Plant Power [MW]	9.7533×10^{-12}
Net Plant Efficiency [%]	2.1676×10^{-7}

Table 8
DOS bounds for 2x2 NGCC plant.

	Lower Bound	Upper Bound
Net Plant Power [MW]	0.5	5
Net Plant Efficiency [%]	40	60

values of the likelihood objective function optimized are below the suggested threshold from Section 3.

With the Kriging models properly assessed, they can now be employed in the operability calculations. The DOS considered in the calculations is given in Table 8. Fig. 9 shows that the machine

learning-based operability approach is capable of successfully obtaining the region of operation satisfying the DOS*. Note that, similarly to the DMA-MR example, the DOS* obtained for both process simulator-based (Aspen Plus) and Kriging-based approaches are overlapping.

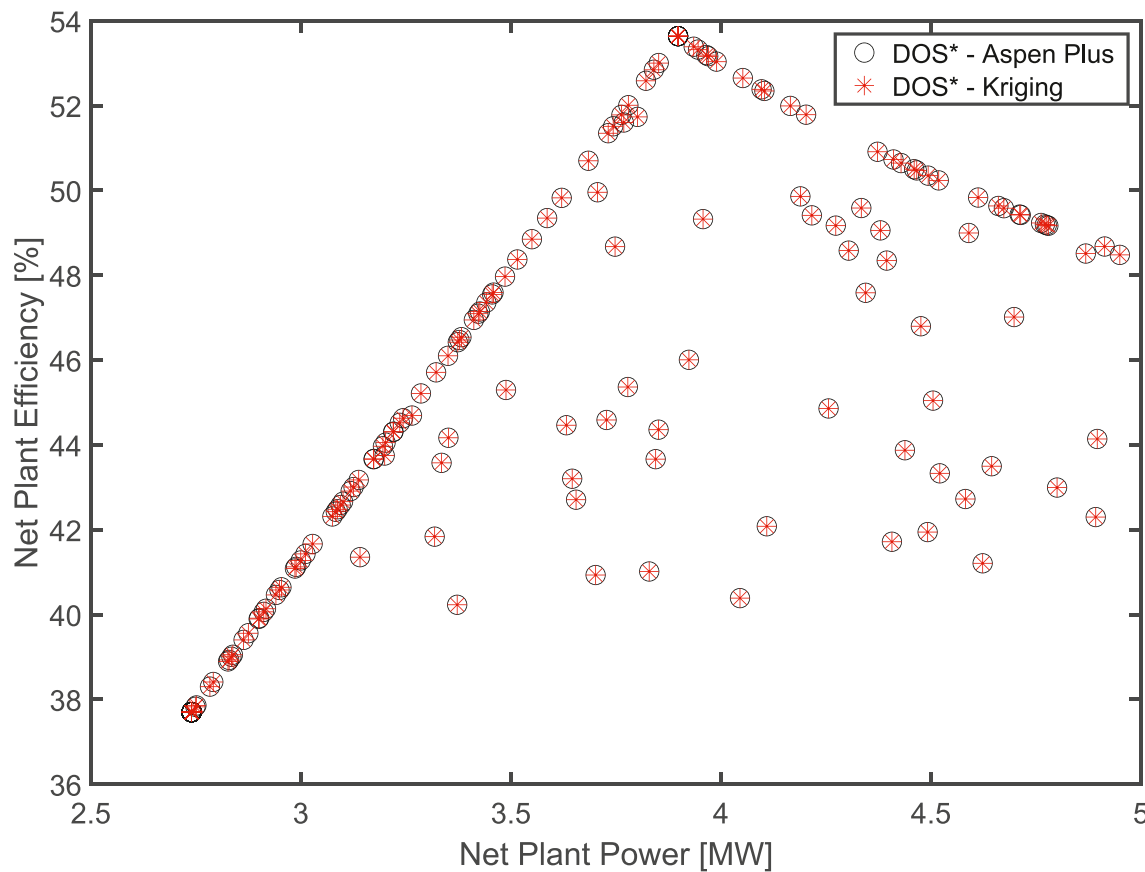


Fig. 9. DOS* comparison: Aspen Plus vs Kriging metamodels for the 2x2 NGCC plant case study.

Table 9

Error analysis of obtained intensified output responses for 2x2 NGCC plant case study.

Output variable	Relative error [%]
Net Plant Power [MW]	1.5070×10^{-7}
Net Plant Efficiency [%]	3.5400×10^{-5}

Table 10

Relative average error for all DOS* grid points: 2x2 NGCC plant case study.

	Relative average error [%]
Net Plant Power [MW]	0.0006
Net Plant Efficiency [%]	0.0019

Table 11

Computational time for 2x2 NGCC plant: Aspen Plus-based vs Kriging-based NLP approach.

NLP Model/Time	Time [h]	Decrease [times]
Aspen Plus	2.2212	-
Kriging	1.00×10^{-3}	2221.1

In order to evaluate the feasibility of the obtained inverse mapping, the optimal point that yielded the highest Net Plant Power and Efficiency is compared for both approaches. It can be seen in Table 9 that the optimal value obtained with the operability framework employing the Kriging model is accurate when compared to the nonlinear model in Aspen Plus. In addition, from the average relative error for all DOS* points shown in Table 10, it can be seen that the errors when using the proposed Kriging-based approach are small for both analyzed variables. Thus, the complexity of employing the process simulator directly with the operability framework (Carrasco and Lima, 2017a) can be eliminated, with the process simulator being used only for the input-output data generation. This can be particularly useful when evaluating the operability of highly nonlinear and complex problems that were already modeled using chemical process simulators.

The Kriging-based NLP approach is also compared against its use with the full process simulator (Aspen Plus) model in terms of computational time. It can be seen in Table 11 that there is a significant decrease in computational effort when using the Kriging-based method. This is an expected result as the Kriging models only predict the output responses that they were trained for, not needing to calculate all system states/equations for every nonlinear solver iteration as needed for simulator models. This fact significantly reduces the computational effort for evaluating each point of the desired operating region.

Fig. 10 shows that the OI evaluation of the Kriging-based approach yields the same value when compared against the operability calculations performed employing the process simulator.

4.2.2. NGCC plant 3x8 case

The same NGCC plant is now considered with 8 input variables and 3 output variables as a high-dimensional case study, following the same variables chosen as in Carrasco and Lima (2018). These variables are shown in Table 12 with their respective AIS bounds considered. The natural gas feed and HRSG steam feed were adjusted to avoid infeasibilities in the AIS regions.

As output variables, the Net Plant Power [MW], Net Plant Efficiency [%] and Capital Cost [\$million] are considered for the

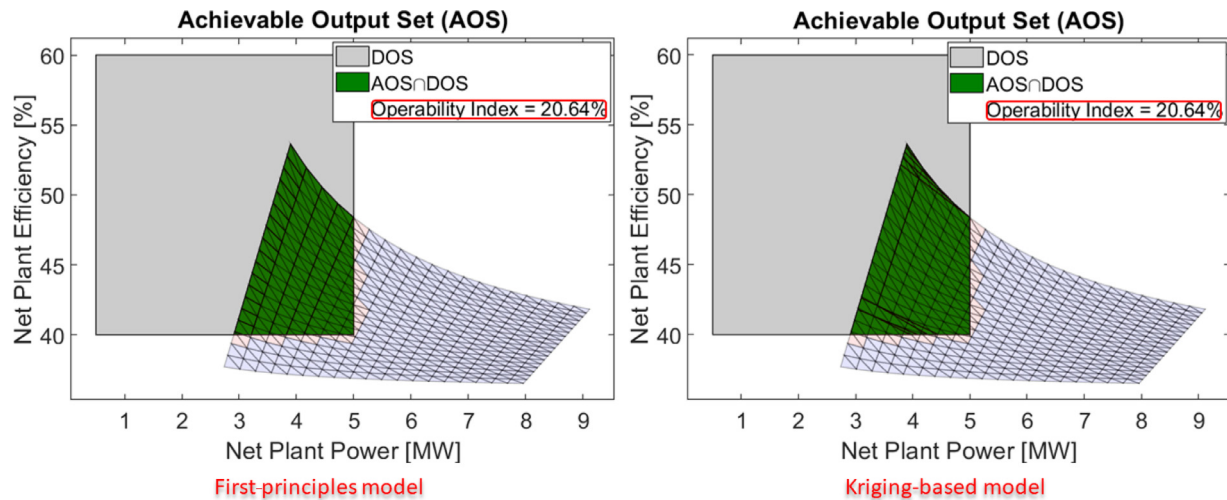


Fig. 10. OI calculations using Aspen Plus simulation versus Kriging-based model.

Table 12
AIS sampling limits for NGCC plant 3x8 case.

	Lower Bound	Upper Bound
Natural gas feed [tonne/h]	0.5	1.5
Heat Recovery Steam Generator steam feed [tonne/h]	0.4	4
Compressor outlet pressure [atm]	10	30
Air feed temperature [K]	273	303
Steam cycle pressure [atm]	100	200
Gas turbine efficiency [%]	0.70	0.85
Air compressor efficiency [%]	0.70	0.85
Steam turbine efficiency [%]	0.70	0.85

Table 13
Root mean-squared errors (RMSE) for NGCC plant 3x8 case outputs.

	RMSE
Net Plant Power [MW]	0.0023
Net Plant Efficiency [%]	0.0299
Capital Cost [\$million]	0.0038

AOS/DOS definitions, following Carrasco and Lima (2018). Considering the AIS bounds from Table 12, 900 samples are generated using the LHS and the outputs are obtained running the process simulator model. With the generated input-output data, a 5-fold cross-validation is performed using the training data and the RMSE results are shown in Table 13. Note from this table the RMSE values indicating that the Kriging models are accurate for operability calculations. Lastly, Table 14 shows the value of the DACE maximum likelihood objective function, also indicating the accuracy of the surrogate model responses.

Table 14
Likelihood objective function values for NGCC plant 3x8 case.

Variable	ψ
Net Plant Power [MW]	1.8632×10^{-6}
Net Plant Efficiency [%]	9.9704×10^{-6}
Capital Cost [\$million]	3.6557×10^{-6}

Table 15
DOS bounds for 3x8 NGCC plant case.

	Lower Bound	Upper Bound
Net Plant Power [MW]	0.5	5
Net Plant Efficiency [%]	40	60
Capital Cost [\$million]	0.1	10

Table 16
Relative error for DOS* intensified design for 3x8 NGCC plant case.

	Aspen Plus	Kriging	Relative Error [%]
Net plant power [MW]	4.1171	4.1159	0.0307
Net plant efficiency [%]	56.6644	56.6327	0.0559
Capital cost [\$million]	8.7403	8.7425	0.0254

After ensuring the accuracy of the generated Kriging responses, a DOS is selected based on target operation, considering as basis the work of Carrasco and Lima (2018). The desired operating range selected for the output variables is shown in Table 15.

The inverse-mapping NLP-based algorithm was run using Kriging-based models for a 500-point grid generated for the DOS, with the bounds defined in Table 15 and LHS to generate the grid. A similar approach was followed for the input data gener-

Table 17
Relative error for DIS* intensified design for 3x8 NGCC plant case.

Input Variable	Aspen Plus	Kriging	Relative error [%]
Natural gas flowrate [kg/h]	500	500.0001	1.9159×10^{-5}
HRSG flowrate [kg/h]	400	400	4.2300×10^{-7}
Air compressor pressure [atm]	18.3676	19.2354	4.7248
Atmospheric air inlet temperature [K]	273	273	1.3900×10^{-7}
HRSG cycle pressure [atm]	184.0897	184.6187	0.2873
Gas turbine efficiency	0.8500	0.8500	2.3100×10^{-8}
Air compressor efficiency	0.8500	0.8500	1.7400×10^{-6}
Steam turbine efficiency	0.8500	0.8500	4.5900×10^{-6}

Table 18

Computational time for 3x8 NGCC Plant case: Aspen Plus-based vs Kriging-based NLP approach.

NLP Model/Time	Time [h]	Decrease [times]
Aspen Plus	31.2800	-
Kriging	0.0815	383.8

Table 19

Average relative error for all DOS* grid points: 8x3 NGCC Plant.

	Average relative error [%]
Net plant power [MW]	0.1281
Net plant efficiency [%]	0.1732
Capital cost [\$ million]	0.2742

ation. Problem P2 is considered in this case study, in which the most efficient modular design is chosen based on maximization of the Net Plant Efficiency. In order to compare the accuracy and show the gain in terms of computational time, the same problem is run using the full process simulator model for the NLP-Based approach. In Tables 16 and 17, the values obtained for the intensified NGCC design for both process simulator-based and Kriging-based NLP calculations are compared. Note from these tables that the relative errors between the two runs are small, confirming the accuracy of the proposed approach for high-dimensional systems. In Table 18, the computational time needed to run the NLP-based approach using the process simulator and the Kriging models are shown. Again one can see that there is a significant decrease in computational time for the proposed approach, enabling the calculation of process operability metrics for high-dimensional systems. Table 19 shows the average relative error for all 500 points of the DOS* grid calculated, indicating that the mapping of the operability regions is done accurately for the desired region using surrogate models. Lastly, it is important to note that the calculations performed with the Kriging-based approach are done using only input-output data from the process simulator model. Thus, problems of model convergence and communication issues between software platforms (process simulator and numerical packages) are therefore mitigated/eliminated.

5. Conclusions

In this work, a systematic method was proposed using Kriging-based models for process operability. It has been shown that the proposed method can generate accurate results (compared to the generated results with the nonlinear first-principles model solutions previously developed) while reducing the complexity and computational effort of operability calculations. The results indicated that a unified approach that depends only on the sampled data and on the accuracy of the surrogate model fitness is produced, allowing the evaluation of complex systems in terms of dimensionality and nonlinearity. Two case studies, using a numerical package for a single unit operation and a process simulator for a complete flowsheet as modeling platform were performed to show the effectiveness of the proposed approach for operability calculations regardless of the data source. Future work is expected to expand the proposed methods to dynamic systems, as a way to reduce the computational effort for enabling online operability calculations, as well as towards using new Kriging formulations for larger datasets and dimensionalities.

Funding

This work was supported by the National Science Foundation CAREER Award 1653098.

Authors' contributions

VA, VG and FVL conceived the research. VA conceived the proposed method. VA conceived the use of Gaussian Process Regression for Process Operability calculations. VA set up the surrogate models for both case studies and applied the proposed method to both case studies. VG provided the first-principles model for the membrane reactor and the NGCC Aspen Plus simulation. VG also provided the MILP formulation for Operability Index calculations. FVL and VG supervised the study. VA wrote the manuscript. FVL and VA supervised the manuscript reviews until it was considered acceptable for submission. All authors read and approved the final manuscript.

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. The authors declare that they have no competing interests.

Acknowledgment

The authors acknowledge the National Science Foundation CAREER Award 1653098 for supporting this research and San Dinh for discussions on input-output multiplicities in process operability.

References

- Alves, V.M.C., Lima, F.S., Silva, S.K., Araujo, A.C.B., 2018. Metamodel-based numerical techniques for self-optimizing control. In: Eng. Chem. Res. 57, 16817–16840. doi:10.1021/acs.iecr.8b04337.
- Boukouvala, F., Ierapetritou, M.G., 2012. Feasibility analysis of black-box processes using an adaptive sampling Kriging-based method. Comput. Chem. Eng. 36, 358–368.
- Boukouvala, F., Muzzio, F.J., Ierapetritou, M.G., 2010. Design space of pharmaceutical processes using data-driven-based methods. J. Pharm. Innov. 5, 119–137.
- Caballero, J.A., Grossmann, I.E., 2008. An algorithm for the use of surrogate models in modular flowsheet optimization. AIChE J. 54, 2633–2650. <https://aiche.onlinelibrary.wiley.com/doi/pdf/10.1002/aic.11579>
- Carrasco, J.C., Lima, F.V., 2015. Nonlinear operability of a membrane reactor for direct methane aromatization. IFAC-PapersOnLine 48, 728–733. 9th IFAC Symposium on Advanced Control of Chemical Processes ADCHEM 2015
- 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 J. 63, 975–983. <https://aiche.onlinelibrary.wiley.com/doi/pdf/10.1002/aic.15439>
- Carrasco, J.C., Lima, F.V., 2017. An optimization-based operability framework for process design and intensification of modular natural gas utilization systems. Comput. Chem. Eng. 105, 246–258.
- Carrasco, J.C., Lima, F.V., 2018. Bilevel and parallel programming-based operability approaches for process intensification and modularity. AIChE J. 64, 3042–3054. <https://aiche.onlinelibrary.wiley.com/doi/pdf/10.1002/aic.16113>
- Davis, E., Ierapetritou, M., 2007. A Kriging method for the solution of nonlinear programs with black-box functions. AIChE J. 53, 2001–2012. <https://aiche.onlinelibrary.wiley.com/doi/pdf/10.1002/aic.11228>
- Forrester, A., Sobester, A., Keane, A., 2008. Engineering Design via Surrogate Modelling: A Practical Guide. John Wiley & Sons.
- Gazzaneo, V., Carrasco, J.C., Lima, F.V., 2018. An MILP-based operability approach for process intensification and design of modular energy systems. In: Eden, M.R., Ierapetritou, M.G., Towler, G.P. (Eds.), 13th International Symposium on Process Systems Engineering (PSE 2018). In: Computer Aided Chemical Engineering, Vol. 44. Elsevier, pp. 2371–2376.
- Gazzaneo, V., Carrasco, J.C., Vinson, D.R., Lima, F.V., 2020. Process operability algorithms: past, present, and future developments. Ind. Eng. Chem. Res. 59, 2457–2470. doi:10.1021/acs.iecr.9b05181.
- Gazzaneo, V., Lima, F.V., 2019. Multilayer operability framework for process design, intensification, and modularization of nonlinear energy systems. Ind. Eng. Chem. Res. 58, 6069–6079. doi:10.1021/acs.iecr.8b05482.
- Georgakis, C., Li, L., 2010. On the calculation of operability sets of nonlinear high-dimensional processes. Ind. Eng. Chem. Res. 49, 8035–8047. doi:10.1021/ie1009316.
- Georgakis, C., Uztürk, D., Subramanian, S., Vinson, D.R., 2003. On the operability of continuous processes. Control Eng Pract 11, 859–869. Process Dynamics and Control
- Jones, D.R., 2001. A taxonomy of global optimization methods based on response surfaces. J. Glob. Optim. 21, 345–383.

- Lima, F.S., Alves, V.M.C., de Araujo, A.C.B., 2020. Metacontrol: a Python based application for self-optimizing control using metamodels. *Comput. Chem. Eng.* 140, 106979.
- Lima, F.V., Georgakis, C., 2010. Input-output operability of control systems: the steady-state case. *J. Process Control* 20, 769–776.
- Lima, F.V., Jia, Z., Ierapetritou, M., Georgakis, C., 2010. Similarities and differences between the concepts of operability and flexibility: the steady-state case. *AIChE J.* 56, 702–716.
- Lophaven, S., Nielsen, H., Søndergaard, J., 2002a. DACE - A Matlab Kriging Toolbox, version 2.0.
- Lophaven, S. N., Nielsen, H. B., Søndergaard, J., 2002b. Aspects of the matlab toolbox DACE. Technical Report, Richard Petersens Plads, Building 321, DK-2800 Kgs. Lyngby, Informatics and Mathematical Modelling, Technical University of Denmark, DTU.
- Maceiczky, R.M., deMello, A.J., 2014. Fast and reliable metamodeling of complex reaction spaces using universal Kriging. *J. Phys. Chem. C* 118, 20026–20033. doi:10.1021/jp506259k.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E., 2011. Scikit-learn: machine learning in Python. *J. Mach. Learn. Res.* 12, 2825–2830.
- Quirante, N., Javaloyes, J., Ruiz-Femenia, R., Caballero, J.A., 2015. Optimization of chemical processes using surrogate models based on a Kriging interpolation. In: Gernaey, K.V., Huusom, J.K., Gani, R. (Eds.), 12th International Symposium on Process Systems Engineering and 25th European Symposium on Computer Aided Process Engineering. In: *Computer Aided Chemical Engineering*, Vol. 37. Elsevier, pp. 179–184.
- Rasmussen, C.E., Nickisch, H., 2010. Gaussian processes for machine learning (gpml) toolbox. *J. Mach. Learn. Res.* 11, 3011–3015.
- Rasmussen, C.E., Williams, C.K.I., 2008. *Gaussian Processes for Machine Learning*. MIT Press.
- Sacks, J., Welch, W.J., Mitchell, T.J., Wynn, H.P., 1989. Design and analysis of computer experiments. *Stat. Sci.* 4 (4), 409–423.
- Subramanian, S., Georgakis, C., 2005. Methodology for the steady-state operability analysis of plantwide systems. *Ind. Eng. Chem. Res.* 44, 7770–7786. doi:10.1021/ie0490076.
- Vinson, D.R., Georgakis, C., 2000. A new measure of process output controllability. *J. Process Control* 10, 185–194.