Data-driven self-optimization of processes in the presence of the model-plant mismatch

Alireza Memarian, * **Santhosh Kumar Varanasi,** *
Biao Huang *,1

* *Department of Chemical and Materials Engineering, University of Alberta, Edmonton, T6G 1H9, Canada.*

Abstract: In this paper, a self-optimization algorithm is developed to find both the optimal operating point and the path from the current condition to the optimal point. Being a model-based strategy, a generalized locally weighted probabilistic principal component regression (PPCR) model that is robust to outliers and can handle missing data, is developed to model the plant. To account for the model-plant mismatch, a penalty term in the form of a robust Gaussian process regression is incorporated into the optimization process. A non-linearity index is utilized to control the accuracy of the local model. Finally, the exploration in optimization for diversity through the acquisition functions is studied. The performance of the proposed algorithm is demonstrated on a simulation case study of a deethanizer column.

Keywords: Process optimization, Generalized probabilistic principal component regression, Gaussian process regression, Model-plant mismatch, Local modeling, Acquisition functions.

1. INTRODUCTION

Increasing productivity, safety, and efficiency have always been the main goals of industrial plants. The objective of plant optimization is to reduce resource wastage and remove bottlenecks while accomplishing the objective of the plant and meeting all plant constraints, including operational, economic, and safety constraints. Due to the reduction in the availability of the raw materials (Manhart et al., 2019), the increase in the market demand for the products because of escalation in the world's population (Mehta et al., 2020), and the environmental concerns like global warming as a result of the emission of the greenhouse gases (GHG) (Zhang et al., 2021), plant optimization has gained more popularity. One of the approaches to do plant optimization is based on the model.

A model is generally obtained through two different approaches, i) first principles and ii) data-driven (Wiebe et al., 2018; Chen et al., 2013). In the first principle model-based approach, the plant is modeled by deriving the governing equations from the fundamental laws, which needs an in-depth understanding of the plant (Pani and Mohanta, 2011). On the other hand, in data-driven model-based approach, a model is built based on the historical data. The closer the developed model is to the plant, the more accurate results can be obtained by solving the optimization problem. However, due to the differences between the model and plant (model-plant mismatch) and the disturbances that may occur during the data collection, the optimal point obtained by solving the optimization problem will be different from the true optimal point (de Avila Ferreira et al., 2018).

To account for the model-plant mismatch in process optimization, the scheme of modifier adaptation was proposed in Oliveira-Silva et al. (2021). In this scheme, the error between the developed model and the plant is incorporated in the objective function while performing the optimization by using the information and measurements collected from the plant. Marchetti et al. (2009) provided a theorem that demonstrates the equivalence of KKT conditions between the plant and the model with the inclusion of the modifier adapters. They suggested the use of gradients of the objective function and constraints calculated from plant measurements as a functional form of modifier adapter. Although the calculation of gradients from noisy plant measurements can be challenging, it is demonstrated to be a reasonably reliable and effective approach. To overcome the challenges with the calculation of gradients, several methods such as nested modifier adapters (Navia et al., 2015), recursive modifier adapters (Marchetti et al., 2010), and derivative-free modifier adapters (Gao et al., 2016) are proposed. Recently, de Avila Ferreira et al. (2018) proposed using Gaussian process regression (GPR) as a modifier adapter. In this work, the historical data and real measurements obtained from the plant are used to train the GP; thereby a nonlinear model that accounts for the model-plant mismatch is obtained. del Rio Chanona et al. (2019) proposed a trust-region framework and the Gaussian process modifier adapters to control the optimization region and to avoid the possibility of violation of constraints. However, the convergence to a local optimal is still a problem in all the aforementioned methods. One of the approaches to overcome this challenge is considering

[1] This work was supported by Natural Sciences and Engineering Research Council of Canada. Corresponding author: Biao Huang (biao.huang@ualberta.ca)

uncertainty when solving the optimization problem (del Rio Chanona et al., 2021).

With the development of reinforcement learning, self-reflective objective is gaining popularity (Kiran et al., 2021). In this approach, the accuracy and reliability of the optimization are improved by consideration of uncertainty. Although many studies have focused on increasing the accuracy of the modifier adaptation, the potential of reinforcement learning has not been studied extensively in the modifier adaptation and optimization problems in general. One of the concepts that can help to increase the accuracy of the optimization is acquisition functions that are used in Bayesian optimization and provide the balance between exploration(trying something new) and exploitation(keep doing what has been done) (del Rio Chanona et al., 2021). In all the aforementioned studies, the modifier adaptation scheme is used along with the first principle models, which essentially requires an in-depth understanding of the process and hence, is not always feasible. In addition to finding the optimal point, an efficient way to steer the process to the optimal point is of paramount importance. Trust-region-based real-time optimization (RTO) is one of the solutions finding an efficient path to the optimal point (Liu and Chen, 2004). However, the application of the data-driven RTO has not been well studied (Powell et al., 2020).

In view of the aforementioned points, a novel self-optimization algorithm is developed in this work that can find both the plant optimal point and the efficient way to shift the current operating condition to the optimal one. The proposed algorithm considers a generalized weighted PPCR model due to its ability to deal with missing and outlier data in both input and output variables, (Memarian et al., 2021; Yuan et al., 2017). Since weighted PPCR is a linear model, and the plant is nonlinear in general, a non-linearity index is used to help the local data-driven model to determine its accuracy. The non-linearity index measures the mismatch between the locally weighted PPCR model and the nonlinear GPR model. Then, this non-linearity index is used to determine the trust range of the generalized locally weighted PPCR model; thereby, the accuracy measure of the model is obtained. In addition, the GPR is used as a modifier adapter to account for the model-plant mismatch. Finally, an acquisition function is adopted to study the exploration during the optimization process.

The remainder of this paper is organized as follows: Section 2 provides the data-driven self-optimization in the presence of model-plant mismatch and the study of acquisition functions for exploration. The efficiency of the algorithm is illustrated through a simulation on a deethanizer column to demonstrate its applicability and feasibility in section 3, and the conclusions are drawn in section 4.

2. DATA-DRIVEN SELF-OPTIMIZATION OF PROCESSES IN THE PRESENCE OF THE MODEL-PLANT MISMATCH

In this section, the data-driven self-optimization of processes in the presence of the model-plant mismatch is presented. The proposed approach utilizes a generalized locally weighted PPCR model that can handle the missing data in both input and output variables along with outliers. Further, due to its weighted local model property, it can efficiently handle the nonlinearity and/or multi-modal nature of plants (Memarian et al., 2021; Yuan et al., 2017). A robust Gaussian process regression model is used to determine the model-plant mismatch between the weighted PPCR model and the plant. To balance the exploitation and exploration in the optimization, the lower confidence bound, described in del Rio Chanona et al. (2021), is used as an acquisition function both in the objective and constraint functions. The details are provided in the rest of this section.

2.1 Generalized weighted PPCR model formulation

One of the important steps while solving an optimization problem is to build a suitable model that can describe the plant with sufficient accuracy. Data-driven modeling is one of the approaches that can help achieving this objective. In the proposed self-optimization algorithm, a generalized weighted PPCR model is used as a data-driven model to describe the plant (Memarian et al., 2021; Yuan et al., 2017). The generalized weighted PPCR model is one of the simplest models in dealing with uncertainties in the plant's datasets.

The generative equation for the generalized locally weighted PPCR model is presented in Eq. (1).

$$\begin{aligned} \boldsymbol{x}_i &= \boldsymbol{P}\boldsymbol{t}_i + \boldsymbol{e}_i + \boldsymbol{\mu}_x, \quad i = 1, 2, \cdots, n \\ \boldsymbol{y}_j &= \boldsymbol{C}\boldsymbol{t}_j + \boldsymbol{f}_j + \boldsymbol{\mu}_y, \quad j = 1, 2, \cdots, n_1 \end{aligned} \tag{1}$$

where, $\boldsymbol{x}_i \in \mathbb{R}^{m\times 1}$ and $\boldsymbol{y}_j \in \mathbb{R}^{r\times 1}$ denote the input and output data, respectively. $\boldsymbol{P} \in \mathbb{R}^{m\times q}$ and $\boldsymbol{C} \in \mathbb{R}^{r\times q}$ are the weighting matrices, and $\boldsymbol{t}_i \in \mathbb{R}^{q\times 1}$ is a vector of latent variables defined in Eq. (2). The variables $\boldsymbol{e}_i \in \mathbb{R}^{m\times 1}$ and $\boldsymbol{f}_j \in \mathbb{R}^{r\times 1}$ denote the noise measurement in input and output, respectively, which are assumed to follow a mixture of two Gaussian distributions given in Eqs. (3) and (4) to account for both outliers and regular noises. The mean values of the input and output variables are denoted as $\boldsymbol{\mu}_x$ and $\boldsymbol{\mu}_y$, respectively. n is the total number of observations where n_1 denotes the number of labeled ones.

$$\boldsymbol{t}_i \sim \mathcal{N}(0, \boldsymbol{I}) \tag{2}$$

$$\boldsymbol{e}_i \sim (1-\delta_x)\mathcal{N}(0, \sigma_x^2\boldsymbol{I}) + \delta_x\mathcal{N}(0, \boldsymbol{\rho}_x^{-1}\sigma_x^2\boldsymbol{I}) \tag{3}$$

$$\boldsymbol{f}_i \sim (1-\delta_y)\mathcal{N}(0, \sigma_y^2\boldsymbol{I}) + \delta_y\mathcal{N}(0, \boldsymbol{\rho}_y^{-1}\sigma_y^2\boldsymbol{I}) \tag{4}$$

Due to the nonlinear and/or multi-modal nature of the plants, developing a single PPCR model to capture the entire plant is not suitable. Thus, to improve the accuracy in modeling, exponential weights are calculated based on Euclidean distance to select the most relevant data points for building the model. The weights are calculated based on Eq. (5).

$$w_i = \exp(\frac{-d_i^2}{\phi}), \quad i = 1, 2, \cdots, n \tag{5}$$

where ϕ is a tuning parameter that defines how the weights are spread across the neighborhood of the testing data to develop the locally weighted PPCR model, and d_i is the Euclidean distance. Further details of the locally weighted PPCR model can be found in Yuan et al. (2017). The model is developed under the framework of expectation

maximization (EM) algorithm. In the $E-step$, the expectation of the log-likelihood function, $Q-function$, presented in Eq. (6) is derived.

$$Q = E_{\boldsymbol{X}_{\boldsymbol{\lambda}},\boldsymbol{T},Q_x,Q_y|\boldsymbol{X}_{o_{\boldsymbol{\lambda}}},\boldsymbol{Y},\boldsymbol{\theta}^{old}} \left[\sum_{i\in\boldsymbol{O}} w_i \log p(\boldsymbol{x}_{i_{\boldsymbol{\lambda}}}, \boldsymbol{y}_i, \boldsymbol{t}_i, Q_x, Q_y \mid \boldsymbol{\theta}) + \sum_{i\in\boldsymbol{M}} w_i \log p(\boldsymbol{x}_{i_{\boldsymbol{\lambda}}}, \boldsymbol{t}_i, Q_x, Q_y \mid \boldsymbol{\theta})\right] \tag{6}$$

where $\theta = \{\boldsymbol{P}, \boldsymbol{C}, \boldsymbol{\mu}_x, \boldsymbol{\mu}_y, \sigma_x, \sigma_y, \delta_x, \delta_y, \boldsymbol{\rho}_x, \boldsymbol{\rho}_y\}$ is the set of model parameters which need to be estimated, defined in Memarian et al. (2021). $\boldsymbol{O}$ and $\boldsymbol{M}$ are the set of observed and missing data, respectively. The $Q-function$ presented in Eq. (6) is calculated by incorporating Eqs. (2)-(5) and following the Bayes rule.

The parameters are estimated through the solution of partial derivatives of the $Q-function$ given in Eq.(6) in the M-step of the algorithm. EM algorithm is an iterative procedure that the $E-step$ and $M-step$ are iterated until the convergence. In the rest of this paper, the generalized locally weighted PPCR model is denoted as G^{PPCR}.

2.2 Data-driven self-optimization algorithm formulation

Since the plant conditions change over time, the historical data that is used to build the model may not be able to accurately describe the current condition of the plant. Therefore, a model-plant mismatch exists between the weighted PPCR model built from the historical data and the current condition of the plant. To account for this model-plant mismatch, de Avila Ferreira et al. (2018) proposed to use the Gaussian process regression (GPR). In our problem, the objective of this GPR model is to build a model by considering the difference between the values of the objective function that are calculated from the plant data (real-time measurements) and the estimation from the locally weighted PPCR model. A similar approach is also used for the variables that have constraint, and the resultant set of equations are shown in Eq. (7).

$$\delta G_i = G_i^P - G_i^{PPCR} \sim \mathcal{GP}(\boldsymbol{\mu}_{\delta G_i}, \boldsymbol{\sigma}^2_{\delta G_i}), \qquad i = 0, \cdots, n_g \tag{7}$$

where n_g is the total number of constraints, where, in Eq. (7), the difference between the variable measurements and their predictions from the locally weighted PPCR is calculated and the mean error is determined through the Gaussian process regression (GPR) model.

Hence, the following optimization problem should be solved.

$$\boldsymbol{u}^{k+1^*} \in \arg\min_{u\in\mathcal{U}} \; [G_0^{PPCR} + \boldsymbol{\mu}^k_{\delta G_0}](\boldsymbol{u}^k) \quad \text{s.t. } [G_i^{PPCR} + \boldsymbol{\mu}^k_{\delta G_i}](\boldsymbol{u}^k) \le 0, \qquad i = 1, \cdots, n_g \tag{8}$$

where $\boldsymbol{\mu}^k_{\delta G_i}$ is the estimated mean of the GP regression that accounts for the term of model-plant mismatch in iteration k. The mean values used in Eq. (8) are those values that are estimated from Eq. (7) which models the model-plant mismatch and correct the bias in both the objective function and constraints in the optimization. $\boldsymbol{u}$ is the manipulated variable that can be defined based on the process.

As discussed in section 2.1, the effective amount of data points relevant to the current operating condition is determined by tuning the parameter ϕ i.e., by decreasing ϕ, fewer data points will effectively contribute to the model construction. If the current operating point is in a highly nonlinear region, building the locally weighted PPCR (linear) model for a given ϕ might not be valid. Thus, by decreasing the parameter ϕ, fewer data points that are closer to the current data point will receive more significant weights; hence, a generalized locally weighted PPCR model in a smaller region will be built. On the other hand, when the weighted PPCR model approximates the nonlinear plant well, we can increase area and have more data point receiving higher weight. Hence, a non-linearity index is proposed to define the range of data to be effectively used, and based on the index, the parameter ϕ can be tuned. The non-linearity index calculates the performance ratio between the nonlinear model (GP regression model) built from the historical data and the linear locally weighted PPCR model, shown in Eq. (9).

$$\rho^{k+1} := \frac{G_0^{GP}(\boldsymbol{u}^k) - G_0^{GP}(\boldsymbol{u}^{k+1^*})}{[G_0^{PPCR} + \boldsymbol{\mu}^k_{\delta G_0}](\boldsymbol{u}^k) - [G_0^{PPCR} + \boldsymbol{\mu}^k_{\delta G_0}](\boldsymbol{u}^{k+1^*})} \tag{9}$$

After calculating the non-linearity index from Eq. (9), similar to the concept of trust-region optimization (del Rio Chanona et al., 2019), three different thresholds are determined to tune ϕ. These three thresholds are $0 < \eta_1 \le \eta_2 < \eta_3 \le 1$. The shrinking and expansion actions to change ϕ are $0 < t_1 < 1 < t_2$ where t_1 and t_2 are shrinking and expansion values, respectively. It has to be noted that these parameters should be tuned before starting the algorithm.

The effective region of the locally weighted PPCR model is updated based on the following steps:

(1) If $G_i^P(\boldsymbol{u}^{k+1^*}) > 0$ for some $i = 1, \cdots, n_g$ or $\rho^{k+1} < \eta_2$ then $\phi := t_1 \times \phi$
(2) Else if $\rho^{k+1} > \eta_3$ then $\phi := \min\{t_2 \times \phi, \phi^{max}\}$
(3) Else $\phi := \phi$

where ϕ^{max} is the maximum allowable value that ϕ can take. Based on the value of ρ, the decision will be made on whether to repeat the optimization, or the obtained optimal point can be used for the next iteration. The decision criterion is as following:

(1) If $G_i^P(\boldsymbol{u}^{k+1^*}) > 0$ for some $i = 1, \cdots, n_g$ or $\rho^{k+1} < \eta_1$ then $\boldsymbol{u}^{k+1} := \boldsymbol{u}^k$
(2) Else $\boldsymbol{u}^{k+1} := \boldsymbol{u}^{k+1^*}$

Based on the aforementioned procedure, the number of effective data points, which will be used for the optimization, will be changed and adjusted based on the performance of the previous iteration. The steps of the proposed algorithm are provided in Algorithm 1.

However, one of the drawbacks of algorithm 1 is that the solution obtained from the optimization can get into the local optimum. To circumvent this problem and pursue the optimization explore, the acquisition function from reinforcement learning and Bayesian optimization is used. del Rio Chanona et al. (2021) proposed using the acquisition functions in the objective function. However, in our proposed method, acquisition functions are used both in objective and constraint functions. Therefore, the LCB

Algorithm 1 Data-driven self-optimization algorithm

Input: Historical data (input and output); initial (query) point, $\boldsymbol{x_q}$; maximum value for ϕ^{max} and an initial value for ϕ; non-linearity threshold parameters $0 < \eta_1 \leq \eta_2 < \eta_3 \leq 1$; expansion and shrinking parameters t_1 and t_2; objective and n_g constraint functions of the optimization problem

Repeat: for $k = 0, 1, \cdots$

1: Build the generalized weighted PPCR model for the given $\boldsymbol{x_q}$ and the historical data
2: Train GP regression modifiers based on the weighted PPCR estimates and the real-time measurements of the plant
3: Solve the modified optimization problem provided in Eq. (8) and obtain $\boldsymbol{u}^{k+1}$
4: Calculate the non-linearity index ρ^{k+1}
5: Update the value of ϕ based on the value of ρ^{k+1}
6: Based on the developed criterion decide to accept the new operating point or to repeat the optimization problem in step 3.
7: $\boldsymbol{x_q} \leftarrow \boldsymbol{u}^{k+1}$ or $\boldsymbol{x_q} \leftarrow \boldsymbol{u}^{k}$ based on the previous step's result

acquisition function is used (Srinivas et al., 2012) and the modified optimization problem is given in Eq. (10):

$$\begin{aligned} &\boldsymbol{u}^{k+1^*} = \arg\min_{u \in \mathcal{U}} \; [G_0^{PPCR} + \boldsymbol{\mu}_{\delta G_0}^k - \beta \boldsymbol{\sigma}_{\delta G_0}^2](\boldsymbol{u}) \\ &\text{s.t. } [G_i^{PPCR} + \boldsymbol{\mu}_{\delta G_i}^k - \beta \boldsymbol{\sigma}_{\delta G_i}^2](\boldsymbol{u}) \leq 0, \quad i = 1, \cdots, n_g \end{aligned} \tag{10}$$

where the variances estimated from the GPR in Eq. (7) are used to move the optimization search to a newer region and may therefore escape the local optimum points by relaxing the constraints. The negative sign before β is consistent with the optimization formulation as the goal is to minimize the objective function. Introducing the LCB acquisition function in the constraints helps to relax these functions while solving the optimization problem. However, if it is needed to tighten the constraints, the UCB acquisition function can be used. With the introduction of acquisition functions, the optimization problem provided in Eq. (10) is solved in the step 3 of Algorithm 1, and the rest of the steps remain the same.

3. CASE STUDY

In this section, the performance of the proposed algorithm is demonstrated by a simulation of a deethanizer column through the Aspen HYSYS V.9 (Belhocine et al., 2020).

3.1 Simulation Example: Deethanizer column

The deethanizer column is a continuous operating distillation column used for extracting ethane as a distillate from a mixed feed that contains light hydrocarbons. Deethanizer column is one of the most important units in refineries and is usually located ahead of other units in the plant.

In Fig. 1, the typical deethanizer column is shown. The objective of the deethanizer column in the refinery plants is to separate C_3+ components from the upstream feed.

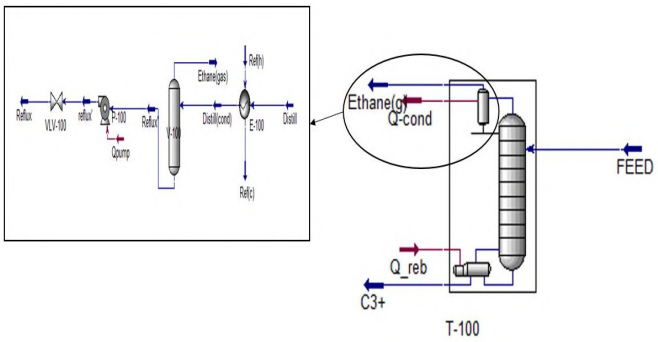


Fig. 1. Schematic of the deethanizer plant (Belhocine et al., 2020)

The main objective of optimization is to minimize the operational cost of the unit, which depends on the energy consumption in the reboilers, the condensers and the pumps. To minimize the energy consumption, the temperature and the feed rate of the input stream need to be regulated. Hence, the objective function is defined as:

$$\begin{aligned} &\min_{T_{feed}, F_{feed}} \; Q_{reb} + Q_{cond} + Q_{pump} \\ &f(Q_{reb}, Q_{cond}, Q_{pump}, F_{bottom}, X_{Ethane,bottom}, T_{feed}, F_{feed}) = 0 \\ &S.T.: \\ &T_{feed} \in [15, 30] \\ &F_{feed} \in [8000, 11000] \\ &F_{bottom} < 2 \times 10^5 [kg/h] \\ &X_{ethane,bottom} < 0.05 \end{aligned} \tag{11}$$

where F_{feed} and F_{bottom} are the flow rate of the feed and bottom product, respectively. T_{feed} is the feed temperature, and $X_{ethane,bottom}$ is the molar fraction of the ethane in the bottom product. Q_{reb}, Q_{cond}, and Q_{pump} are the terms corresponding to the energy consumption of the reboiler, condenser, and pump, respectively. $f(.) = 0$ is the PPCR model that relates input and output variables to each other. The first two constraints i.e., T_{feed}, F_{feed} which are defined in Eq. (11) are the operational constraints, and the F_{bottom} and $X_{ethane,bottom}$ are the planning constraints. In such a setting, 15% of the input and 35% of the output data are missed, and 10% of the data is replaced with outliers to represent the possible errors in data collection through sensors.

By solving the optimization of Eq. (11) with the *optimization module* of Aspen HYSYS, the minimum energy consumption is found to be $1.082 \times 10^8 [W]$ and the decision variables are found to be $T_{feed} = 16.3$ and $F_{feed} = 10485$. The operating region and the actual solution to the optimization of Eq. (11) are presented in Fig. 2

To demonstrate the efficacy of the proposed method in steering the plant to its optimal point, two different initializations (current operating points, COPs) are considered. In Fig. 3(a), the locations of these COPs are shown in Fig. 3(b), where the path and the final solution obtained by the proposed method for each COP are provided.

Based on the results demonstrated in Fig. 3, the proposed algorithm is able to find the optimal path and solution, and steer the plant to the desired point. It is well known that signal to noise ratio (SNR) can affect the performance of the modeling and optimization. To study the effect of measurement noise on the proposed algorithm, eight

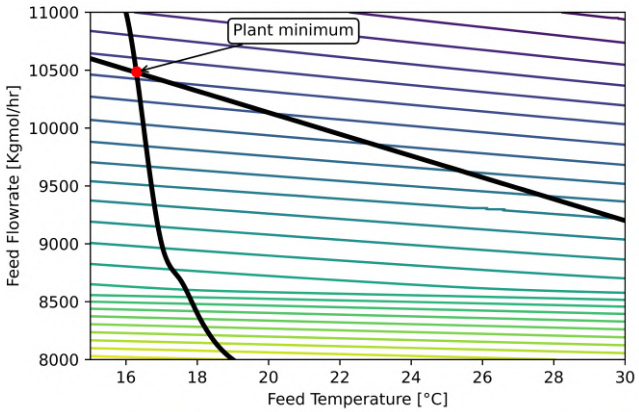
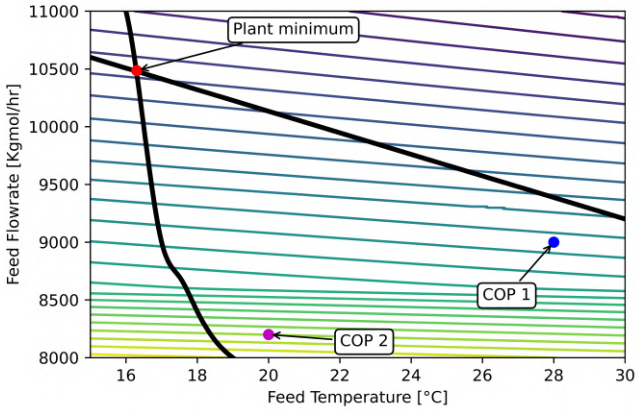
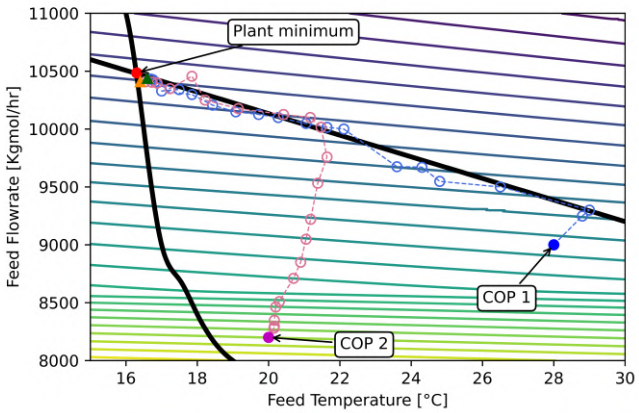


Fig. 2. Operating region and optimal point of the deethanizer problem



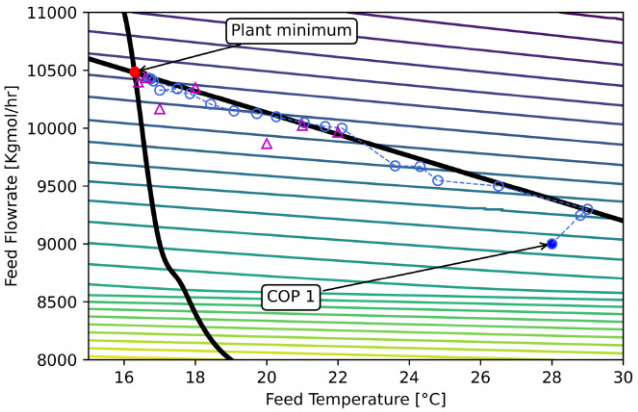
(a) Locations of COP 1 and COP 2



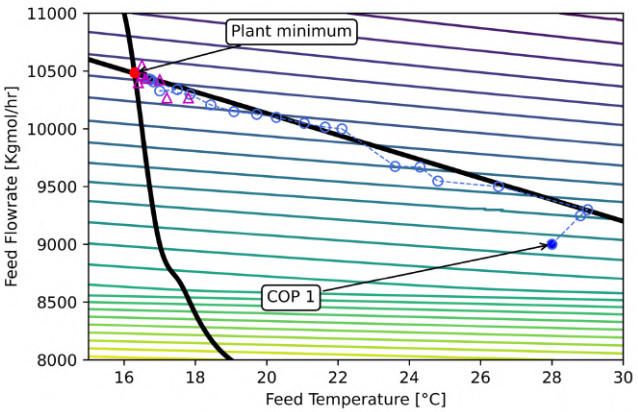
(b) Optimal path and solution from the proposed algorithm

Fig. 3. Initial points and the solutions obtained by the proposed data-driven self-optimization algorithm

different noise levels are considered for this study whose optimal points are shown in Fig. 3. Two different initial operating points are considered. As it can be seen from Figure 4(a), solutions corresponding to noisier data (lower SNR) are getting trapped in local optimum points. To obtain a better solution through the discovery of the



(a) Data-driven self-optimization algorithm without exploration



(b) Data-driven self-optimization algorithm with exploration

Fig. 4. Study the effect of exploration in the optimization problem. Solutions to the optimization problem with 8 different noise levels without exploration (Figure 4(a)) and with exploration (Figure 4(b))

new path by searching a wider optimization region, the exploration as described in Eq. (10) is applied, and the results are shown in Figure 4(b).

From the results of Figure 4, it can be seen that the inclusion of exploration in the optimization as explained in Eq. (10), helps in better convergence to the optimal point, and avoid being trapped in the local optimum. In Figure 4(b), more points with different noise levels are getting closer to the plant minimum compared to the points shown in Figure 4(a).

4. CONCLUSION

In this work, a data-driven self-optimization of the process in the presence of model-plant mismatch is proposed to find the plant optimum along with the path to reach the optimum. The objective of the proposed algorithm is to automate the procedure of finding optimal operating points of a process. It models the plant with a generalized locally weighted PPCR model and the Gaussian process regression model is utilized to identify the model-plant mismatch. A non-linearity index is proposed to adjust the weighted PPCR model to ensure its accuracy at a sufficient

level. Finally, to make a balance between exploitation and exploration, the acquisition function is used in the optimization. The performance of the proposed algorithm is demonstrated on the simulated deethanizer column. Based on the results obtained from the case studies, it can be concluded that the proposed algorithm is able to move the plant towards the plant optimum.

REFERENCES

Belhocine, A., Bendib, R., and Zennir, Y. (2020). Simulation and analysis of a petrochemical process (deethanizer column-mle field) using hysys aspen simulator. *Algerian Journal of Signals and Systems*, 5(2), 86–91.

Chen, S., Dai, L., Liu, J., Gao, Y., Liu, X., Chen, Z., Zhou, J., Cao, C., Han, P., Luo, H., et al. (2013). The visible transmittance and solar modulation ability of vo 2 flexible foils simultaneously improved by ti doping: an optimization and first principle study. *Physical Chemistry Chemical Physics*, 15(40), 17537–17543.

de Avila Ferreira, T., Shukla, H.A., Faulwasser, T., Jones, C.N., and Bonvin, D. (2018). Real-time optimization of uncertain process systems via modifier adaptation and gaussian processes. In *2018 European Control Conference (ECC)*, 465–470. IEEE.

del Rio Chanona, E.A., Graciano, J.A., Bradford, E., and Chachuat, B. (2019). Modifier-adaptation schemes employing gaussian processes and trust regions for real-time optimization. *IFAC-PapersOnLine*, 52(1), 52–57.

del Rio Chanona, E.A., Petsagkourakis, P., Bradford, E., Graciano, J.A., and Chachuat, B. (2021). Real-time optimization meets bayesian optimization and derivative-free optimization: A tale of modifier adaptation. *Computers & Chemical Engineering*, 147, 107249.

Gao, W., Wenzel, S., and Engell, S. (2016). A reliable modifier-adaptation strategy for real-time optimization. *Computers & chemical engineering*, 91, 318–328.

Kiran, B.R., Sobh, I., Talpaert, V., Mannion, P., Al Sallab, A.A., Yogamani, S., and Pérez, P. (2021). Deep reinforcement learning for autonomous driving: A survey. *IEEE Transactions on Intelligent Transportation Systems*.

Liu, T.L. and Chen, H.T. (2004). Real-time tracking using trust-region methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(3), 397–402.

Manhart, A., Vogt, R., Priester, M., Dehoust, G., Auberger, A., Blepp, M., Dolega, P., Kämper, C., Giegrich, J., Schmidt, G., et al. (2019). The environmental criticality of primary raw materials–a new methodology to assess global environmental hazard potentials of minerals and metals from mining. *Mineral Economics*, 32(1), 91–107.

Marchetti, A., Chachuat, B., and Bonvin, D. (2009). Modifier-adaptation methodology for real-time optimization. *Industrial & engineering chemistry research*, 48(13), 6022–6033.

Marchetti, A., Chachuat, B., and Bonvin, D. (2010). A dual modifier-adaptation approach for real-time optimization. *Journal of Process Control*, 20(9), 1027–1037.

Mehta, S., Saxena, T., and Purohit, N. (2020). The new consumer behaviour paradigm amid covid-19: Permanent or transient? *Journal of Health Management*, 22(2), 291–301.

Memarian, A., Varanasi, S.K., and Huang, B. (2021). Mixture robust semi-supervised probabilistic principal component regression with missing input data. *Chemometrics and Intelligent Laboratory Systems*, 214, 104315.

Navia, D., Briceño, L., Gutiérrez, G., and De Prada, C. (2015). Modifier-adaptation methodology for real-time optimization reformulated as a nested optimization problem. *Industrial & engineering chemistry research*, 54(48), 12054–12071.

Oliveira-Silva, E., de Prada, C., and Navia, D. (2021). Economic mpc with modifier adaptation using transient measurements. In *Computer Aided Chemical Engineering*, volume 50, 1253–1258. Elsevier.

Pani, A.K. and Mohanta, H.K. (2011). A survey of data treatment techniques for soft sensor design. *Chemical Product and Process Modeling*, 6(1).

Powell, K.M., Machalek, D., and Quah, T. (2020). Real-time optimization using reinforcement learning. *Computers & Chemical Engineering*, 143, 107077.

Srinivas, N., Krause, A., Kakade, S.M., and Seeger, M.W. (2012). Information-theoretic regret bounds for gaussian process optimization in the bandit setting. *IEEE Transactions on Information Theory*, 58(5), 3250–3265.

Wiebe, J., Cecílio, I., and Misener, R. (2018). Data-driven optimization of processes with degrading equipment. *Industrial & Engineering Chemistry Research*, 57(50), 17177–17191.

Yuan, X., Ge, Z., Song, Z., Wang, Y., Yang, C., and Zhang, H. (2017). Soft sensor modeling of nonlinear industrial processes based on weighted probabilistic projection regression. *IEEE Transactions on Instrumentation and Measurement*, 66(4), 837–845.

Zhang, X., Xie, H., Liu, X., Kong, D., Zhang, S., and Wang, C. (2021). A novel green substrate made by sludge digestate and its biochar: Plant growth and greenhouse emission. *Science of The Total Environment*, 149194.