Smart Optimization of Post-Combustion CO_2 Capture from a Coal-fired Power Plant: A Bayesian Framework with Wavelet Neural Networks

Prince Kumar Jain [*] **, Santhosh Kumar Varanasi** [*]

[*] *Indian Institute of Technology, Jodhpur 342030, Rajasthan, India (e-mail: jain.69@iitj.ac.in, skvaranasi@iitj.ac.in).*

Abstract: Post-Combustion CO_2 capture has been a major focus for decades in efforts to reduce global warming. In this study, CO_2 emissions from a coal power plant are analyzed taking into account an existing process available in Aspen Hysys. In this paper, the main objective is to identify the operating conditions that result in an economical operation of the process. Since the process is complex, instead of relying on a first-principle-based steady-state model, a data-driven approach via a wavelet neural network was considered because of its linearity with respect to the parametric structure. This allows for faster training and provides accurate predictions. Although the model is accurate, due to changes in operating conditions in a process plant, a mismatch between the actual plant output and the predicted model output may exist. To account for this mismatch, Bayesian optimization is employed using Gaussian process regression, which estimates both the mean value and uncertainty of the mismatch. The trust region approach is applied to balance the crucial factors of exploration and exploitation. The efficacy of the proposed method is demonstrated via a Benoit system and a PCC process.

Keywords: Post-combustion CO_2 capture, real-time optimization, model-plant mismatch, Gaussian process regression, Bayesian optimization, wavelet neural network

1. INTRODUCTION

According to the International Energy Agency, coal represents more than a third of global electricity generation, although it is the most carbon-intensive fossil fuel (International Energy Agency, 2024). The emission of CO_2 from coal fired power plants is a significant contributor to the global warming. To address this issue, renewable energy-based power generation has gained significant attention in recent years. Since the complete conversion from fossil fuels to renewable energy takes time, reducing global CO2 emissions, particularly through post-combustion CO2 capture, has been considered a key factor in mitigating global warming.

The primary technologies for CO_2 capture after combustion are classified into two categories: solvent-based and solid sorbent-based methods (Bhattacharyya and Miller, 2017). Among these, post-combustion CO_2 capture (PCC) using amine-based solvents is the most commonly utilized at an industrial scale to mitigate CO_2 emissions. Owing to the high flow rates and low CO_2 concentrations in the flue gas, the solvent-based PCC process demands a significant amount of heat for solvent regeneration, which poses a major challenge to its large-scale implementation. Due to this, it creates a net financial burden for the associated upstream power plants, where profit maximization is a priority (Wu et al., 2020). To minimize economic costs in real-time operations, it's crucial to identify the optimal operating conditions of the process plant via process optimization.

The Plant optimization is primarily achieved through two key approaches: Economic Model Predictive Control (Ellis et al., 2014) and Robust Real-Time Optimization methods (Darby et al., 2011). Economic Model Predictive Control (EMPC) integrates optimization and control objectives into a single framework, wherein a performance criterion which is directly related to the economic operation of the system is considered. The EMPC optimization techniques rely on a dynamic model (unsteady-state model) and are hence typically computationally demanding. On the other hand, RTO utilizes a steady state model of the plant and identifies the optimal process parameters.

The steady-state model utilized in RTO can be derived from two methods: (i) First Principle approach, and (ii) Data Driven Approach (Wiebe et al., 2018; Chen et al., 2013). A comprehensive understanding of plant operations is essential for the first-principles approach (Pani and Mohanta, 2011); however, it is often impractical to possess complete knowledge of all aspects of the plant to develop a first-principles model. Thus, the data-driven RTO approach is employed in this paper for identifying the optimal operating points that minimizes the operational cost of a PCC plant. Wavelet neural networks (WNNs) are preferred due to their ability to efficiently handle nonlinear processes. Owing to its linear-in-parametric form makes WNNs faster to train and implement (Varanasi et al., 2022).

In a typical process plant, owing to the presence of noise or through the drift in operating conditions of the process, a mismatch often occurs between the model and the actual plant. To handle this mismatch, Modifier adaptation (MA) techniques are used. MA subject the model into a non-linear optimization problem which is solved at each RTO iteration. Various variation of MA are used in a typical process plant optimization. The authors in Singhal et al. (2016) used quadratic surrogates as modifier adapters for the predicted cost and constraint functions. The authors in de Avila Ferreira et al. (2018) explored the use of Gaussian processes (GPs) as the cost and constraint modifiers. This idea of Gaussian process regression along with employing the concept of trust-region, captures the uncertainty. Further exploration is guided through acquisition functions while keeping constraints within the acceptable range del Rio Chanona et al. (2021).

The objective of this paper is primarily more focused on finding the optimal point that optimizes the overall operational cost of the plant. For the PCC process, a specific objective is to determine the optimal flow rates of lean MEA solvent and Make-up water that minimizes the operational cost of the plant while ensuring the constraint on streams flow rate. To obtain a steady-state model, a data-based approach through WNNs is incorporated. A bayesian optimization strategy via a Gaussian process regression is employed to account for the model-plant mismatches.

The rest of the paper is organized as follows: In Section 2, a description of post-combustion CO_2 capture (PCC) using amine-based solvents, the economic objective of the PCC process alongside the optimization formulation is given. In Section 3, details of WNNs, Gaussian process regression model and the proposed methodology is given. The efficacy of the proposed approach is demonstrated by considering two case studies: a benoit system and a MEA based CO_2 capture process in Section 4 and the conclusions are drawn in Section 5.

2. MEA-BASED CO_2 CAPTURE PROCESS

2.1 Process description

A typical CO_2 capture plant consists of two components: an absorber and a stripper. In the absorber, flue gas flows upward while a lean MEA solvent flows downward in counter-current fashion. As they interact, the amine absorbs CO_2 from the gas through a chemical reaction, leaving the cleaned gas to be vented from the top. The CO_2-rich solvent is then directed to a heat exchanger, where it recovers energy from the solvent exiting the stripper before moving into the stripper for regeneration.

The stripper, which uses a reboiler for heating, separates the CO_2 from the rich solvent. The CO_2 is collected at the top, while the lean solvent is recycled back to the absorber. A condenser removes any evaporated solution and recycle the water stream into the absorber. The system also includes a buffer tank where fresh MEA solvent and water is added to maintain performance.

The CO_2 capture plant considered in this paper is simulated using Aspen Hysys software and is shown in Fig. 1. Further, the lean solvent flow (Lean Solvent stream in Fig. 1) and Makeup water (Makeup H_2O stream in Fig. 1) are considered as decision variables while % Carbon capture and economic cost are considered as controlled variables in this paper. The Carbon capture % is defined as follows:

$$\text{Carbon Capture (\%)} = \left(\frac{\dot{m}_{\mathrm{CO_2,out}}}{\dot{m}_{\mathrm{CO_2,in}}}\right) \times 100$$

where $\dot{m}_{\mathrm{CO_2,out}}$ is the mass flow rate of CO_2 in the distillate stream of the stripper and $\dot{m}_{\mathrm{CO_2,in}}$ is the mass flow rate of CO_2 present in the dry flue gas.

2.2 Process economics and Constraints

The objective function used to model and evaluate the PCC plant performance was initially introduced by Patrón and Ricardez-Sandoval (2023). It utilizes the steady-state model f_s as referenced in this paper and the economic objective is defined as:

$$\begin{aligned}\phi \;=\; & P_{\mathrm{sales}}\left(\dot{m}^{s}_{\mathrm{CO_2,in}} - \dot{m}^{s}_{\mathrm{CO_2,out}}\right) \\ & + P_{\mathrm{CO_2}}\dot{m}^{s}_{\mathrm{CO_2,out}} + P_{\mathrm{steam}}Q_{\mathrm{reb}}\end{aligned} \tag{1}$$

The objective of process optimization is to minimize the objective i.e.,

$$\min_{\dot{m}^{\mathrm{lean}}_{\mathrm{mea\text{-}solvent}},\,\dot{m}^{\mathrm{makeup}}_{\mathrm{water}}} \phi \tag{2}$$

subject to model and variable constraints as

$$f_s(Q_{\mathrm{reb}}, \dot{m}^{\mathrm{lean}}_{\mathrm{mea\text{-}solvent}}, \dot{m}^{\mathrm{makeup}}_{\mathrm{water}}, \%\ \mathrm{CC}) = 0 \tag{3}$$

$$\dot{m}^{\mathrm{lean}}_{\mathrm{mea\text{-}solvent}} \in [1.6 \times 10^6, 1.603201 \times 10^6]\ \mathrm{kg/hr} \tag{4}$$

$$\dot{m}^{\mathrm{makeup}}_{\mathrm{water}} \in [4.9250 \times 10^4, 5 \times 10^4]\ \mathrm{kg/hr} \tag{5}$$

In this context, P_{sales}, $P_{\mathrm{CO_2}}$, and P_{steam} represent the price of CO_2 sales, the social cost of carbon (SCC), and the price of steam, respectively. These factors are key economic considerations in post-combustion carbon capture (PCC). They are multiplied by the corresponding mass flow rates ($\dot{m}$) or energy duties (Q). The specific values of these prices are provided in Table 1 which was taken from Patrón and Ricardez-Sandoval (2023).

Table 1. Prices used in Objective function

Variable	**Value**
P_{sales}	-50 $ CAD/tn_CO$_2$_sold
$P_{\mathrm{CO\cdot 2}}$	176 $ CAD/tn_CO$_2$_removed
P_{steam}	0.065 $ CAD/kWh

3. WAVELET NEURAL NETWORK ASSISTED SMART OPTIMIZATION OF MEA-BASED CO_2 CAPTURE PROCESS

3.1 Wavelet Neural Network as Surrogate Model

A Wavelet Neural Network (WNN) typically consists of three layers: the input layer, the hidden layer, and the output layer as depicted in Fig. 2.

[1] References and screen images from Aspen HYSYS®are reprinted with permission from Aspen Technology, Inc. AspenTech®, Aspen HYSYS®, Aspen Plus®, Aspen Plus Dynamics®, Aspen Economics Evaluation™, Aspen EDR™, Aspen Energy Analyzer™, and Aspen Properties®, aspenONE®, and the AspenTech leaf logo are

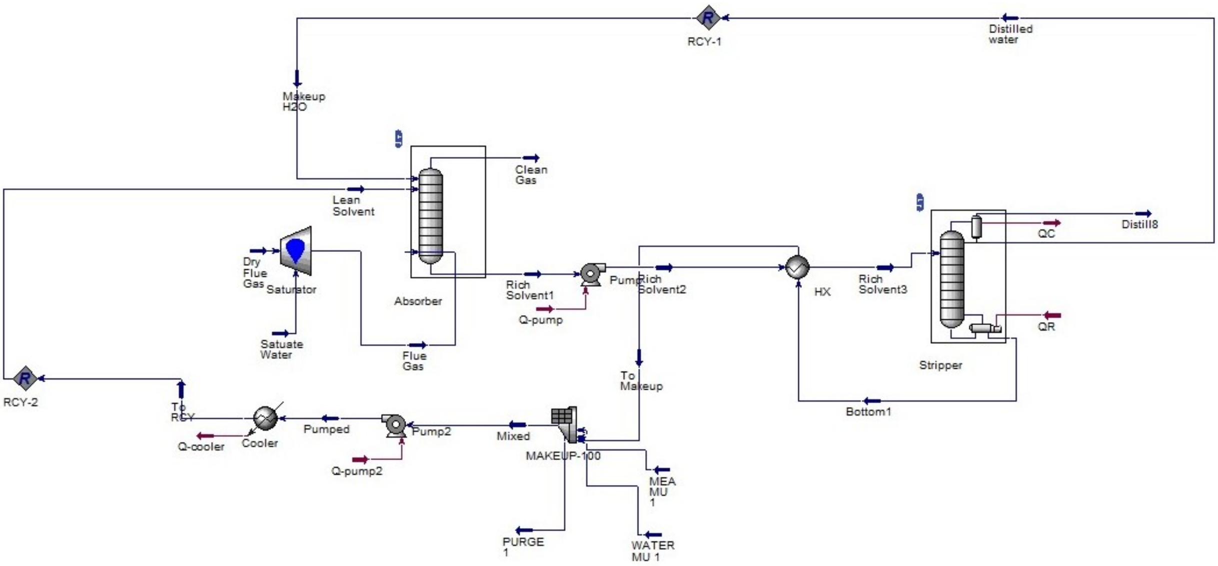

Fig. 1. Flowsheet of PCC Plant using Aspen Hysys (Aspen Technology, Inc., 2023) [1]

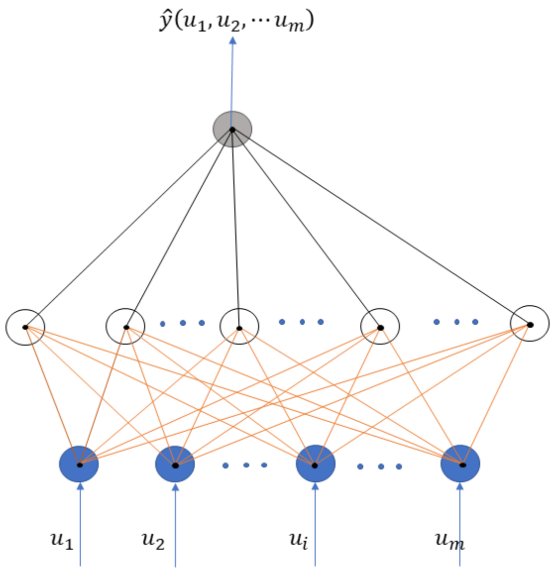

Fig. 2. Wavelet Neural Network

In the input layer, the network receives explanatory variables, also called inputs. The hidden layer contains nodes, known as wavelons, which transform these input variables into a non-linear space. The output layer then uses the transformed variables to approximate the target values. The authors in Alexandridis and Zapranis (2013) proposed a wavelet network that incorporates a multi-dimensional basis and linear connections between the input and output layers, as depicted in Fig. 2, to enhance training performance in highly linear settings.

Since any non-linear function can be approximated using wavelet frame decompositions (Billings and Wei, 2005), similar to the work considered in (Varanasi et al., 2022), the main objective of this work is to represent the parameters in the input-hidden layer using these decompositions. This type of approximation helps us in addressing the issues associated with randomization strategies for training while ensuring the model in linear-in-parametric form thereby making the learning faster, accurate and require less amount of data to train when compared to a traditional single hidden layer feedforward neural network. Further, with the appropriate selection of the dilational and translational parameters associated with the wavelet frame decompositions (Billings and Wei, 2005), the number of neurons to be considered in the hidden layer will be automatically fixed, thereby avoiding the trail and error approach for selection of optimal number of neurons in the hidden layer. The output equation of the wavelet neural network is given as

$$\hat{y}(u_1, u_2, \cdots, u_m) = \sum_{j=j_0}^{j_M} \sum_{k_1 \in K_j} \cdots \sum_{k_r \in K_j} \alpha_{j,k_1,\ldots,k_m} \psi^{[m]}_{j,k_1,\ldots,k_m}(u_1, u_2, \cdots, u_m)$$

where,

$$\psi^{[r]}_{j,k_1,\ldots,k_m}(x(t)) = 2^j \times (m - \|2^j x(t) - v\|^2) \times \exp(-0.5 \times \|2^j x(t) - v\|^2) \tag{6}$$

denote the Mexican hat wavelet function with $j \in \mathbb{Z}$ and $(k_1, k_2, \ldots, k_r) \in \mathbb{Z}^r$ denote the dilation and translation parameters respectively and $v = [k_1, k_2, \cdots, k_m]$. The terms j_0 and and j_f denote be the coarsest resolution and the finest resolution respectively. The term k_i has to be selected in such a way that $-(s_2 - 1) \leq k_i \leq 2^j - s_1 - 1, i = 1, 2, \ldots, r$ where, s_1 and s_2 control the range of the translations. Typical choice of s_1 and s_2 for the Mexican hat wavelet function are -3 and 3 respectively (Billings and Wei, 2005).

3.2 Bayesian Optimization : Gaussian Process Regression

Gaussian Process (GP) regression is an interpolation method used to represent a unknown function $f : \mathbb{R}^{n_x} \to \mathbb{R}$ containing uncertainties (noise) $y = f(x) + \epsilon$, where ϵ

follows a Gaussian distribution $N(0, \sigma^2)$ indicating zero mean and unknown variance noise. It was developed by Krige (1951) and then later used in machine learning field by Williams and Rasmussen (2006).

A Gaussian Process is explained by its mean function and covariance function as

$$f(\mathbf{x}) \sim \mathcal{GP}(m(\cdot), k(\cdot, \cdot))$$

where $m(\cdot)$ denote the vector of mean values of each Gaussian distribution, and the covariance function $k(\cdot, \cdot)$ characterizes the correlations between data points. The most commonly used kernel function is a squared exponential kernel aka Gaussian kernel or RBF kernel (Williams and Rasmussen, 2006):

$$k(\mathbf{x}, \mathbf{x}') = \sigma_f^2 \exp\left(-\frac{1}{2}(\mathbf{x} - \mathbf{x}')^T \Lambda (\mathbf{x} - \mathbf{x}')\right)$$

where σ_f^2 represents the covariance magnitude, and $\Lambda := \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ is a scaling matrix. The key assumption for this kernel is that the inferred function f is smooth and stationary. In addition, we choose mean function as zero as it is complete data driven approach and we do not have prior knowledge of the function. Thus,

$$m(\mathbf{x}) := 0$$

To estimate a GP's hyperparameters $\Psi := \{\sigma_f, \sigma_n, \lambda_1, \ldots, \lambda_n\}$ through maximum likelihood, the noise variance σ_n may either be known or inferred from the data. For the input matrix $\mathbf{X} := [\mathbf{x}_1, \ldots, \mathbf{x}_N]^T \in \mathbb{R}^{N \times n}$ let the noisy output be $\mathbf{y} := [y_1, \ldots, y_N]^T \in \mathbb{R}^N$. The log-likelihood of the given input-output pairs without considering the constant terms is given by:

$$\log p(\mathbf{y}|\mathbf{X}) = -\frac{1}{2}\mathbf{y}^T \left(K(\mathbf{X}, \mathbf{X}) + \sigma_n^2 \mathbf{I}\right)^{-1} \mathbf{y} - \frac{1}{2} \log \left|K(\mathbf{X}, \mathbf{X}) + \sigma_n^2 \mathbf{I}\right|$$

The predicted distribution at test point $\mathbf{x}^*$, given the input-output pair $(\mathbf{X}, \mathbf{y})$ follows a Gaussian distribution as:

$$f(\mathbf{x}^*) \mid \mathbf{X}, \mathbf{y} \sim \mathcal{N}(\mu_f(\mathbf{x}^*), \sigma_f^2(\mathbf{x}^*))$$

where the posterior mean $\mu_f(\mathbf{x}^*)$ and posterior variance $\sigma_f^2(\mathbf{x}^*)$ are given by:

$$\mu_f(\mathbf{x}^*) = \mathbf{k}(\mathbf{x}^*, \mathbf{X}) \left(K(\mathbf{X}, \mathbf{X}) + \sigma_n^2 \mathbf{I}\right)^{-1} \mathbf{y}$$

$$\sigma_f^2(\mathbf{x}^*) = \sigma_n^2 - \mathbf{k}(\mathbf{x}^*, \mathbf{X}) \left(K(\mathbf{X}, \mathbf{X}) + \sigma_n^2 \mathbf{I}\right)^{-1} \mathbf{k}(\mathbf{x}^*, \mathbf{X})^T$$

where $k(\mathbf{x}^*, \mathbf{X})$ denotes the vector of covariances between the test data point and the training data points.

Both the mean function and co-variance function are used in the acquisition functions del Rio Chanona et al. (2021). The acquisition functions are then considered as objective function for the purpose of exploration and exploitation. Two commonly used acquisition functions include lower confidence bound (LCB) and expected improvement (EI). In the current paper, LCB is considered as acquisition function.

Lower Confidence Bound : The LCB function is defined as

$$\mathcal{A}_{\mathrm{LCB}}[\mu_f, \sigma_f](\mathbf{x}^*) := \mu_f(\mathbf{x}^*) - \gamma \sigma_f(\mathbf{x}^*)$$

where γ is the hyperparameter which can be read as exploration weight. The LCB function follows the principle of optimism under uncertainty,

3.3 Proposed Methodology

In this section, we present the modifier adaptation algorithm using Gaussian processes, trust region (TR), and acquisition functions. The detailed steps of the algorithm are outlined below: the use of GPs to describe the plant-model mismatch in an RTO problem was first proposed by de Avila Ferreira et al. (2018). The aim of the GP modifiers to predict the mismatch of the cost and each constraint separately. Below is the initial optimization problem which was used to solve in each RTO iteration without considering any acquisition function and trust-region concept:

$$\mathbf{x}^{k+1} \in \arg\min_{\mathbf{x} \in X} \left[G_0 + \mu_{\delta G_0}^k\right](\mathbf{x}) \tag{7}$$

$$\tag{8}$$

$$\text{s.t.} \ \left[G_i + \mu_{\delta G_i}^k\right](\mathbf{x}) \leq 0, \quad i = 1, \ldots, n_g \tag{9}$$

where $\mu_{\delta G_i}^k$ denotes the mean of the GP trained with the input output data set $\left(\mathbf{X}^k, \delta\mathbf{G}_i^k\right)$; and $\delta\mathbf{G}_i^k$ denotes the mismatch $\delta G_i(\cdot) := G_i^{\mathrm{P}}(\cdot) - G_i(\cdot)$ for inputs in the matrix $\mathbf{X}^k$.

This research includes the addition of the trust region concept with Bayesian optimization, along with the LCB acquisition function. This paper includes the training of Gaussian process and Wavelet neural network in the regime of trust region using K-Nearest Neighbour for better prediction. Now, the modified optimization problem becomes:

$$\mathbf{r}^{k+1} \in \arg\min_{\mathbf{r}} \mathcal{A}\left[G_0 + \mu_{\delta G_0}^k, \sigma_{\delta G_0}^k\right]\left(\mathbf{x}^k + \mathbf{r}\right) \tag{10}$$

$$\text{s.t.} \ \left[G_i + \mu_{\delta G_i}^k\right]\left(\mathbf{x}^k + \mathbf{r}\right) \leq 0, \quad i = 1 \ldots n_g \tag{11}$$

$$\|\mathbf{r}\| \leq \Delta^k, \quad \mathbf{x}^k + \mathbf{r} \in \mathcal{X} \tag{12}$$

The trust-region concept, represented by Δ^k ensures that the current points remain within the permissible operating range, preventing them from exceeding system constraints ($\mathcal{X}$). $\mathcal{A}$ is an LCB acquisition function used for the purpose of exploration and exploitation. A non-linearity index ζ^{k+1} is used to check the performance of WNN model and GP by comparing it with the actual process output, which is given as :

$$\zeta^{k+1} := \frac{G_0^{\mathrm{P}}(\mathbf{x}^k) - G_0^{\mathrm{P}}(\mathbf{x}^k + \mathbf{r}^{k+1})}{[G_0 + \mu_{\delta G_0}^k](\mathbf{x}^k) - [G_0 + \mu_{\delta G_0}^k](\mathbf{x}^k + \mathbf{r}^{k+1})} \tag{13}$$

Based on the value of the non-linearity index, we shrink or expand the trust region, and a new operating point is selected. The shrinking and expansion action is done by $0 < p_1 < 1 < p_2$ where p_1 and p_2 are shrinking and expansion values, respectively. The thresholds for non-linearity index are $0 < \beta_1 \leq \beta_2 < 1$. The algorithm begins by first initializing the above parameters. Below are the conditions on non-linearity index:

$$
\begin{aligned}
&\text{If } \zeta^{k+1} > \beta_2 \wedge \|\mathbf{r}^{k+1}\| = \Delta^k : \\
&\quad \Delta^{k+1} \leftarrow p_2 \times \Delta^k, \\
&\quad \mathbf{x}^{k+1} \leftarrow \mathbf{x}^k + \mathbf{r}^{k+1} \text{ (accept)} \\
&\text{Else If } \zeta^{k+1} < \beta_1 : \\
&\quad \Delta^{k+1} \leftarrow p_1 \times \Delta^k, \\
&\quad \mathbf{x}^{k+1} \leftarrow \mathbf{x}^k \text{ (reject)} \\
&\text{Else:} \\
&\quad \Delta^{k+1} \leftarrow \Delta^k, \\
&\quad \mathbf{x}^{k+1} \leftarrow \mathbf{x}^k + \mathbf{r}^{k+1} \text{ (accept)}
\end{aligned}
\tag{14}
$$

Following is the algorithm used for current research:

Data-driven optimization algorithm including GP, WNN, TR and MA

Input:

Historical input-output dataset, initial operating point $\mathbf{x}^0$ and initial trust-region radius as Δ^0 where $0 < \Delta^0 < \Delta_{\max}$, non-linearity threshold parameters $0 < \beta_1 \leq \beta_2 < 1$, shrinking and expansion parameters p_1 and p_2;

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

(1) Train WNN and GP modifier considering data points
(2) Solve the modified optimization problem provided in Eq. (10) and obtain $\mathbf{r}^{k+1}$.
(3) Check infeasibility: If Eq. (10) is infeasible, then set $\mathbf{x}^{k+1} \leftarrow \mathbf{x}^k$ + some noise for exploration
(4) Calculate the non-linearity index ζ^{k+1}.
(5) Update the value of Δ^{k+1} based on the value of ζ^{k+1}.
(6) Based on the provided conditions in equation (14), decide to accept the new operating point or not.
(7) Update the data set with new operating point.

4. CASE STUDY

4.1 Benoit System

To illustrate the accuracy of the proposed method, a simple two-variable constrained optimization problem termed as Benoit system del Rio Chanona et al. (2021) is considered. The formulation is as follows

$$
\begin{aligned}
\min_{\mathbf{u} \in [-2,2]^2} \quad & y_1(\mathbf{u}) \\
\text{s.t.} \quad & y_2(\mathbf{u}) \leq 0
\end{aligned}
\tag{15}
$$

where

$$
\begin{aligned}
y_1(\mathbf{u}) &:= u_1^2 + u_2^2 + \theta_1 u_1 u_2, \\
y_2(\mathbf{u}) &:= 1 - u_1 + u_2^2 + \theta_2 u_2.
\end{aligned}
$$

The (unknown) plant parameter values are taken as $\theta = [1, 2]$ and the corresponding true plant optimum is u = [0.368, -0.393].

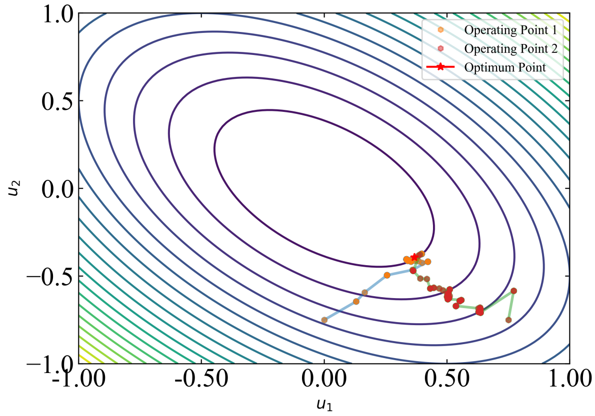

Fig. 3. Smart optimization contour plot of Benoit System when started from two different points. The range of u_1 and u_2 is considered in [-1,1] range in plotting for better visualization.

The objective function was modeled using a Wavelet Neural Network, with the mismatch handled by a Gaussian Process. Two different initial operating points are considered and the contour plots of optimization performance is shown in Fig. 3. From Fig. 3 it can be noted that when the initial operating point 1 is $[0, -0.75]$, the smart optimization yields an optimal point at $[0.38507795, -0.40424324]$, converging close to the true optimal point while adhering to the defined constraints. Similarly, for the initial operating point of $[0.75, -0.75]$, which yields an optimal point at $[0.38372245, -0.38192726]$. These results demonstrate the effectiveness of the model in accurately reaching the optimal points for the Benoit system under different initial conditions.

4.2 MEA Based PCC Process Results

For the PCC Process, the Lean Solvent stream and Makeup water are used as input variables in the wavelet neural network (WNN), while the percentage of carbon capture is the output of the WNN model as the reboiler duty (Q_{reb}) was at constant value. The economic objective function is then calculated using Eq. (1). Using Proposed algorithm, the optimization is performed at four different initial operating points and the following plot is obtained.

The true optimum point is $[1603150.904, 49250.0394]$, with an objective value of 40055.3423.

For the initial operating point of [1602500.2808, 49500.413789 89297] (operating point 1), and the optimization yields an optimum point at [1603201.0, 49252.211728985] with an objective value of 40055.11043. For operating point 2 i.e., [1601001.2808, 49744.4137899],

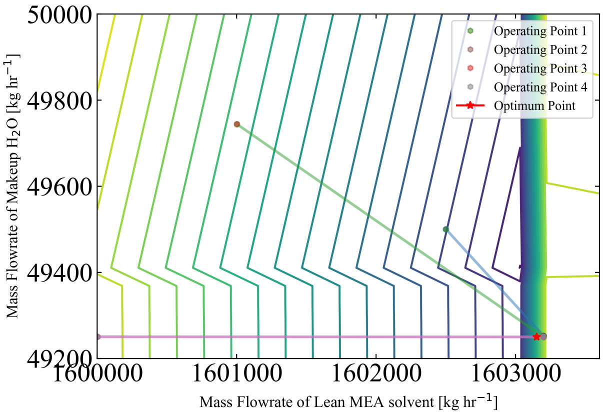

Fig. 4. Smart optimization contour Plots of PCC System when started from four different points

the optimization results in the optimum point [1603201.0, 49250.0394], with an objective value of 40055.1211112. Similarly, the initial operating point 3 is [1600001.2808, 49250.41378989297], and the optimization yields the same optimum point [1603201.0, 49250.0394] with an objective value of 40055.0958202. Finally, for the operating point 4[2] i.e., [1600001.2808, 49250.41378 989297], the optimization results to [1603200.9999999998, 49250.0394] with an objective value of 40055.1190 728. These results highlight the robustness and consistency of the optimization process in reaching the true optimum point across various initial conditions.

5. CONCLUSION

The proposed smart optimization algorithm utilizes Wavelet Neural Network as model to represent the process; Gaussian processes and trust regions methods to account for model-plant mismatch, provides a structured approach to optimize processes in real time. By integrating acquisition functions and just in time learning concept, a balance between exploration and exploitation is achieved, thereby allowing for efficient convergence and improved performance in the optimization tasks. Two case studies were explored to analyze the proposed method. In both cases, the proposed approach successfully reached the true optimal point, demonstrating the efficacy of the methodology.

REFERENCES

Alexandridis, A.K. and Zapranis, A.D. (2013). Wavelet neural networks: A practical guide. *Neural Networks*, 42, 1–27.

Aspen Technology, Inc. (2023). Aspen hysys. `https://www.aspentech.com/products/engineering/aspen-hysys`. Process simulation software.

Bhattacharyya, D. and Miller, D.C. (2017). Post-combustion co2 capture technologies — a review of processes for solvent-based and sorbent-based co2 capture. *Current Opinion in Chemical Engineering*, 17, 78–92. doi:https://doi.org/10.1016/j.coche.2017.06.005. Energy and Environmental Engineering / Reaction engineering and catalysis.

Billings, S.A. and Wei, H.L. (2005). A new class of wavelet networks for nonlinear system identification. *IEEE Transactions on neural networks*, 16(4), 862–874.

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.

Darby, M.L., Nikolaou, M., Jones, J., and Nicholson, D. (2011). Rto: An overview and assessment of current practice. *Journal of Process control*, 21(6), 874–884.

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., 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.

Ellis, M., Durand, H., and Christofides, P.D. (2014). A tutorial review of economic model predictive control methods. *Journal of Process Control*, 24(8), 1156–1178.

International Energy Agency (2024). Electricity - energy system. `https://www.iea.org/energy-system/electricity`. Accessed: 2024-10-07.

Krige, D.G. (1951). *A statistical approach to some mine valuation and allied problems on the Witwatersrand: By DG Krige*. Ph.D. thesis, University of the Witwatersrand.

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).

Patrón, G.D. and Ricardez-Sandoval, L. (2023). Robust real-time optimization and parameter estimation of post-combustion co2 capture under economic uncertainty. *Chemical Engineering Science*, 281, 119124.

Singhal, M., Marchetti, A.G., Faulwasser, T., and Bonvin, D. (2016). Real-time optimization based on adaptation of surrogate models. *IFAC-PapersOnLine*, 49(7), 412–417.

Varanasi, S.K., Daemi, A., Huang, B., Slot, G., and Majoko, P. (2022). Sparsity constrained wavelet neural networks for robust soft sensor design with application to the industrial kivcet unit. *Computers & Chemical Engineering*, 159, 107695.

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

Williams, C.K. and Rasmussen, C.E. (2006). *Gaussian processes for machine learning*, volume 2. MIT press Cambridge, MA.

Wu, X., Wang, M., Liao, P., Shen, J., and Li, Y. (2020). Solvent-based post-combustion co2 capture for power plants: A critical review and perspective on dynamic modelling, system identification, process control and flexible operation. *Applied Energy*, 257, 113941. doi: https://doi.org/10.1016/j.apenergy.2019.113941.

[2] in this case, fewer data points when compared to earlier 3 cases are used in just-in-time learning