A Machine Learning Approach for Chemical Process Optimization

Sergio Bugosena, Carl Lairda, Carleton Coffrinb, and Robert Parkerb\*

a Department of Chemical Engineering, Carnegie Mellon University, Pittsburgh, Pennsylvania, United States

b Los Alamos National Laboratory, Los Alamos, New Mexico, United States

\* Corresponding Author: [rbparker@lanl.gov](mailto:rbparker@lanl.gov).

**Keywords**: Machine learning, Surrogate modeling, Chemical process design, Chemical process optimization, Nonlinear optimization

The United States has set ambitious goals for the future, aiming to reduce greenhouse emissions by 50-52% below 2005 levels by 2030. Additionally, it is looking to reach 100% carbon pollution-free electricity by 2035 [1]. To meet these objectives, chemical processes that manufacture hydrogen from waste or natural gas must be modeled and designed to optimality to satisfy product specifications, as well as safety, economic, and environmental constraints. In today’s context of globalization of trade, heightened competition, and high environmental pressures, it is crucial to prioritize the design of optimal chemical processes [2].

However, classical optimization of a chemical process is challenging due to the number of variables and algebraic equations defining each unit operation, as well as the nonlinear complexity of these first principles models, which makes calculating the derivatives used in the optimization algorithm a difficult and resource-consuming task.

To improve solver convergence reliability and solve times when optimizing a chemical process flowsheet, we demonstrate the implementation of data-driven optimization based on surrogate modeling for the autothermal reforming section of a low-carbon hydrogen manufacturing process. This approach relies on representing complex thermodynamics, transport phenomena and chemical reaction equations with tractable and simple polynomials with reduced dimensionality, which enable faster function evaluations and facilitate the search for optimal solutions in a cost-effective manner.

We perform the comparison between classical (or full space) optimization and surrogate-based optimization by modeling two autothermal reforming flowsheets using Pyomo [3] as the algebraic modeling environment and the IDAES [4] framework. The full space flowsheet is modeled with the first principles Gibbs Reactor and is shown in **Figure 1**. In the second flowsheet, the Gibbs Reactor is replaced by a surrogate block containing polynomials generated by the ALAMO [5] machine learning framework. The optimization problems were solved with IPOPT. We compare convergence reliability, solve time, and solution quality between these two flowsheets. The degrees of freedom for both flowsheets are the steam flow and the bypass fraction. The goal is to maximize H2 concentration in the syngas stream while satisfying the displayed constraints.

A picture containing text, diagram, plan, technical drawing

Description automatically generated**Figure 1.** Full Space Autothermal Reformer Flowsheet.

To demonstrate the potential of surrogate-based optimization, we solve 64 instances parameterized by inlet natural gas pressure and temperature. Our results, shown in **Figure 2**, demonstrate significantly improved convergence reliability and solve time in contrast to the full space formulation, without a significant increase in solution error. The surrogate-based formulation converges all instances in an average of 1.3 s, while the full space formulation converges 49 out of 64 instances in an average of 5.9 s. The mean relative error between objective values is 1.5%.

A screenshot of a graph

Description automatically generated

**Figure 2.** Convergence status for the two formulations.

Regular text uses the **PSE\_Text** style. Please do not add spaces between paragraphs. Paragraphs will indent automatically and will hyphenate automatically. Do not add extra paragraph spaces between sections. Important things to know:

Do not use headers in this abstract, except for references.

**Limit this submission to one page**

The standard column width is 8.5 cm in this template. Insert a single-column figure in its own paragraph, and change the style of that paragraph to **PSE\_FigureAndCaption**. On the next line, type the figure caption, and style it **PSE\_FigureAndCaption**. Place the figure where it best flows with the text.

Diagram, schematic

Description automatically generated

**Figure 1.** Example 1-column figure with a short caption.

Use the Equation Editor in Microsoft Word (Alt + = or Insert | Equation). Use **PSE\_Equation** for the style. Add the equation reference numbers in parenthesis after the equation inline, and use tabs until the equation numbers are aligned properly (equation numbers and whitespace all use **PSE\_Equation**). Do not add additional space above or below equation (handled by the style).

(1)

(2)

You can also use LaTeX format. To add a LaTeX equation, hit Alt + = or Insert | Equation to add a new equation, and make sure the Conversions is set to LaTeX in the Editor ribbon (need this the first time only).

(3)

For tables, use the Insert | Table feature to insert a table with a single column width, inline in the text. Use **PSE\_TableCaption** for the caption and **PSE\_TableContent** for the table itself. Use the **PSE\_TableFormat** template in the Table Design ribbon to apply standard borders. (You will not see **PSE\_TableFormat** in the list of styles, that is found in the Table Design ribbon).

**Table 1:** Caption for the table. Multi-line table captions will be fully justified. The style is **PSE\_TableCaption**.

|  |  |  |
| --- | --- | --- |
| **Stream** | **Description** | **Flow** |
| Feed | A mixture of MeOH, EtOH, and BuOH. | 11.2 kmol/hr |
| Distillate | Mostly methanol | 21.5 kmol/hr |
| Bottoms | Mostly butanol | 110 kmol/hr |

References

1. R. Parker, B. Nicholson, J. Siirola, C. Laird, and L. Biegler, “An implicit function formulation for optimization of discretized index-1 differential algebraic systems,” Computers & Chemical Engineering, vol. 168. Elsevier BV, p. 108042, Dec. 2022.
2. A. Cozad, N. V. Sahinidis, and D. C. Miller, “Learning surrogate models for simulation-based optimization,” AIChE Journal, vol. 60, no. 6. Wiley, pp. 2211–2227, Mar. 13, 2014.
3. A. Wächter and L. T. Biegler, “On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming,” Mathematical Programming, vol. 106, no. 1. Springer Science and Business Media LLC, pp. 25–57, Apr. 28, 2005.
4. A. Lee et al., “The IDEAS process modeling framework and model library—Flexibility for process simulation and optimization,” Journal of Advanced Manufacturing and Processing, vol. 3, no. 3. Wiley, May 13, 2021.