# Machine Learning for Chemical Process Optimization

Sergio Bugosen[1,2], MS; Robert Parker[1], PhD; Carleton Coffrin[1], PhD. | [1]A-1 Information Systems & Modeling; [2]Department of Chemical Engineering, Carnegie Mellon University

## Motivation

### CHEMICAL PROCESS DESIGN AND OPTIMIZATION

A fundamental task in chemical process design is the selection of **optimal operating conditions** that satisfy product specifications, as well as safety and economic constraints. However, classical optimization of a chemical process flowsheet is challenging due to:

(1) **Number of variables and algebraic equations** defining each unit operation.

(2) **Nonlinear complexity** of the first principles models.

(3) The optimization problem heavily depends on converging the **large system of algebraic equations** defining the flowsheet.

### MACHINE LEARNING IN CHEMICAL ENGINEERING

- To reduce the complexity and number of equations, we utilize the Automated Learning of Algebraic Models (ALAMO) tool.
- ALAMO is a machine learning model that provides tractable and simple polynomial functions to represent complex unit operations.
- Optimization of a chemical process flowsheet that uses ALAMO will improve convergence reliability and solve times.

> "Replace complex reaction engineering equations with tractable, simple polynomials."

### OBJECTIVE

- Compare **convergence reliability**, **solve time** and **solution quality** between classical (or full space) and surrogate-based optimization formulations.

### CHEMICAL PROCESS DESCRIPTION

This project will focus on optimizing the autothermal reformer (ATR) section of a low-carbon hydrogen manufacturing process. The objective of the ATR process is to produce syngas, a mixture mainly composed of $H_2$, CO, $CH_4$ and $CO_2$. First, a mixture of natural gas, steam and air is fed into the reactor. The hot syngas is then circulated through a shell and tube heat exchanger, also called the reformer recuperator, to heat the natural gas feed. This natural gas feed is then expanded to generate electrical power and is finally fed into the reactor, closing the loop.

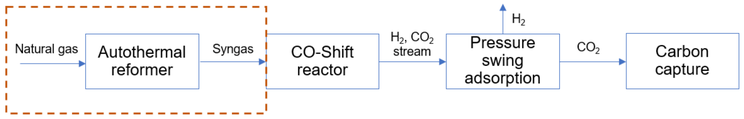


**Chart 1.** Block flow diagram of a low-carbon manufacturing process.

## Methods

### SOFTWARE

Pyomo, an open-source optimization modeling language, and the IDAES framework (Institute for the Design of Advanced Energy Systems) were used to model the ATR flowsheets in Python 3.10.11. NLP solver used was IPOPT 3.13.2, linear solver MA27.

### FULL SPACE ATR FLOWSHEET

The goal is to maximize $H_2$ concentration in the syngas stream while satisfying the constraints shown in Chart 2. The degrees of freedom are the inlet steam flow and the bypass fraction.

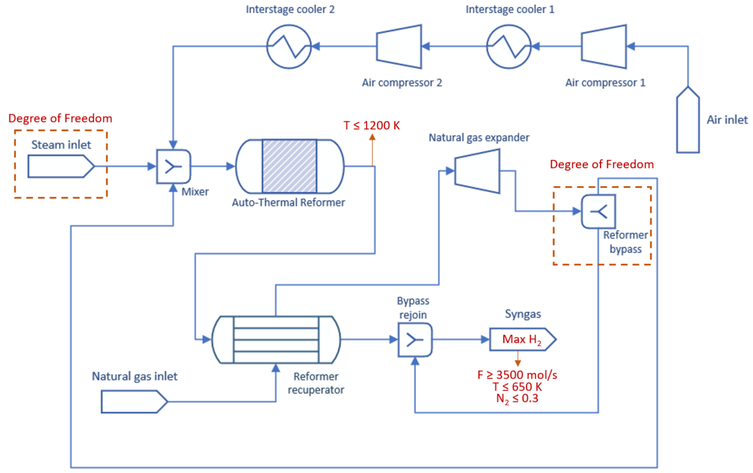


**Chart 2.** Full space ATR flowsheet displaying objective and constraints.

### ALAMO SURROGATE-BASED ATR FLOWSHEET

This flowsheet was modeled exactly as the one in Chart 2. Main difference is that instead of using a first principles reactor, we use a surrogate model block consisting of simple polynomials.

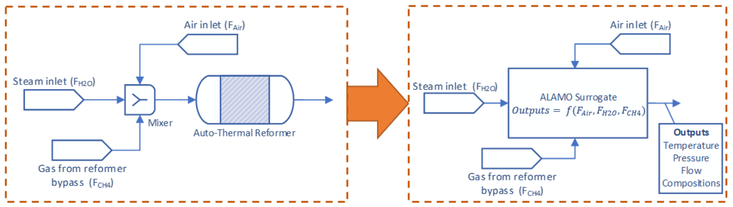


**Chart 3.** Representation of a surrogate reactor block.

### VALIDATION – SOLUTION QUALITY

The procedure shown in Chart 4 was employed to assess the discrepancy between solutions.

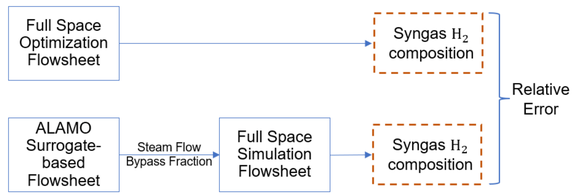


**Chart 4.** Representation of a surrogate reactor block.

## Results

### CONVERGENCE RELIABILITY

The full space formulation was able to converge on 49 out of 64 instances, while the surrogate-based formulation converged in every one of them.

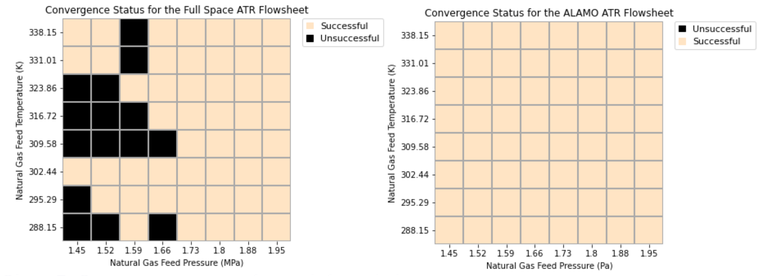


**Chart 5.** Convergence status for each formulation.

### PERFORMANCE

The ALAMO surrogate-based formulation is faster than the full space by a factor of five, and the solution time between instances on this parameter sweep is more uniform, with a standard deviation that is sixteen times lower.

| Formulation | # Variables | # Constraints | Avg. solve time (s) | Standard dev. (s) |
|---|---|---|---|---|
| Full Space | 897 | 895 | 5.97 | 8.02 |
| ALAMO | 451 | 449 | 1.26 | 0.42 |

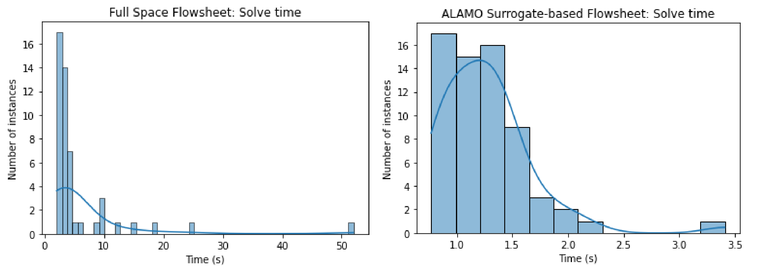


**Chart 6.** Solve time distribution for each formulation.

### SOLUTION QUALITY

The mean relative error between the objective values from the Full Space and ALAMO surrogate-based formulations is 1.47%, with a standard deviation of 0.51 and a maximum error of 3.04%.

## Future Work

1. Include an implicit function formulation in this comparison.
2. Compare these three formulations (Full Space, ALAMO surrogate-based, implicit) for a more complex chemical process, such as the carbon capture section of the plant.
3. Extend the surrogate-based formulation to dynamic optimization problems.

Los Alamos NATIONAL LABORATORY
NNSA National Nuclear Security Administration
Managed by Triad National Security, LLC, for the U.S. Department of Energy's NNSA.
LA-UR# number goes here