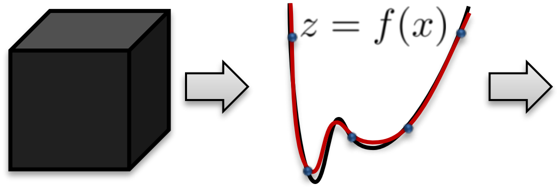


$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & g(x) = 0 \end{aligned}$$

DERIVATIVE-FREE OPTIMIZATION ENHANCED-SURROGATE MODEL DEVELOPMENT FOR OPTIMIZATION

Alison Cozad, Nick Sahinidis, David Miller

CAPD CENTER

Carnegie Mellon University

NETL NATIONAL ENERGY TECHNOLOGY LABORATORY

CCSI™

Carbon Capture Challenge

- **The traditional pathway from discovery to commercialization of energy technologies can be quite long, i.e., ~ 2-3 decades**
- **President's plan requires that barriers to the widespread, safe, and cost-effective deployment of CCS be overcome within 10 years**
- **To help realize the President's objectives, new approaches are needed for taking carbon capture concepts from lab to power plant, quickly, and at low cost and risk**
- **CCSI will accelerate the development of carbon capture technology, from discovery through deployment, with the help of science-based simulations**

Bench Research
~ 1 kWe

↓

Small pilot
< 1 MWe

↓

Medium pilot
1 – 5 MWe

↓

Semi-works pilot
20-35 MWe

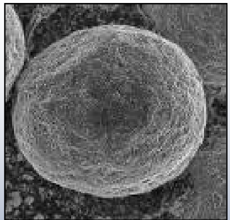
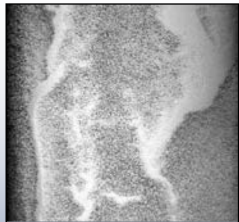
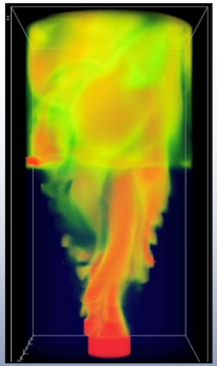
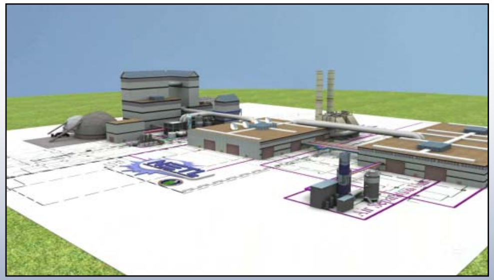
↓

First commercial plant, 100 MWe

↓

Deployment, >500 MWe, >300 plants

Carbon Capture Simulation Initiative
www.acceleratecarboncapture.org

Identify promising concepts → **Reduce the time for design & troubleshooting** → **Quantify the technical risk, to enable reaching larger scales, earlier** → **Stabilize the cost during commercial deployment**

National Labs

- Berkeley Lab
- Pacific Northwest National Laboratory
- NETL
- Los Alamos National Laboratory EST. 1943
- Lawrence Livermore National Laboratory

Academia

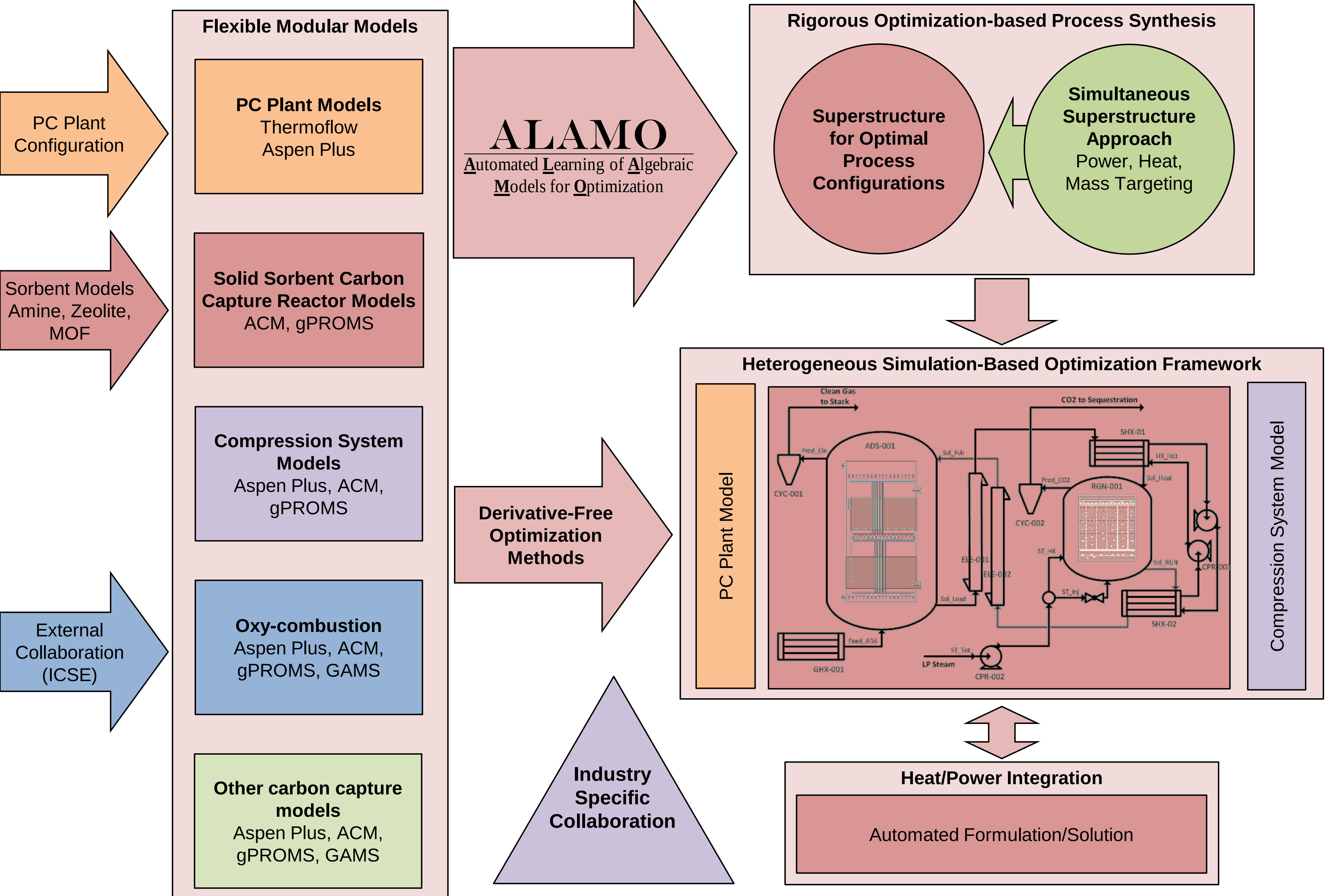
- Carnegie Mellon
- Princeton University
- West Virginia University
- Boston University
- Institute for Clean and Secure Energy, The University of Utah

Industry

- Fluor
- ADA Environmental Solutions
- B&W the babcock & wilcox company
- GE
- EPRI Electric Power Research Institute
- Air Products
- Burns & McDonnell Since 1898
- Southern California Edison An EDISON INTERNATIONAL Company San Onofre Nuclear Generating Station
- Alstom
- Southern Company
- Duke Energy
- URS
- Boeing
- DuPont
- ExxonMobil
- Chevron
- Phillips 66
- WorleyParsons resources & energy
- Eastman

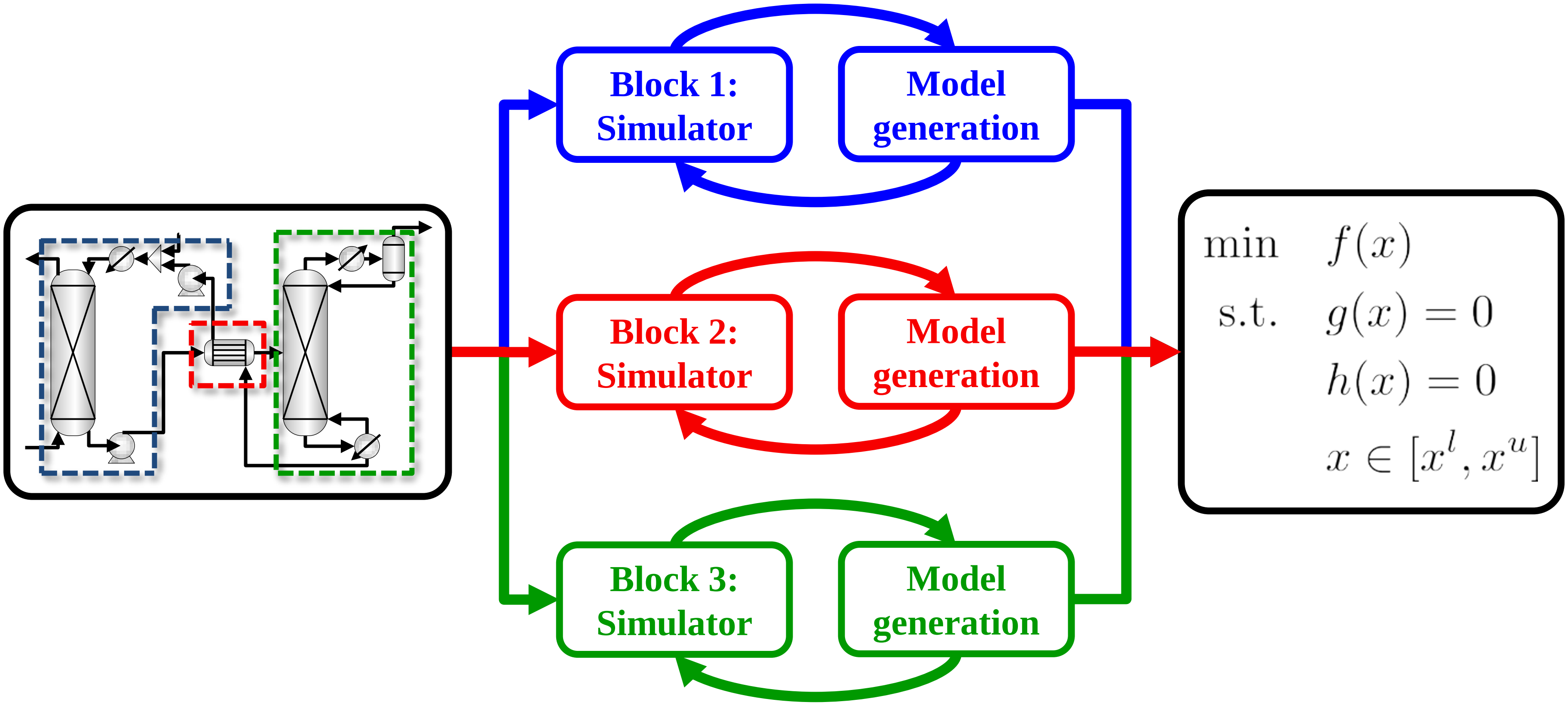
CCSI Carbon Capture Simulation Initiative

NETL · Berkeley Lab · Lawrence Livermore National Laboratory · Los Alamos National Laboratory EST. 1943 · Pacific Northwest National Laboratory · U.S. Department of Energy



CCSI Carbon Capture Simulation Initiative · NETL · Berkeley Lab · Lawrence Livermore National Laboratory · Los Alamos National Laboratory · Pacific Northwest National Laboratory · U.S. Department of Energy

PROCESS DISAGGREGATION



Process Simulation

Disaggregate process into process blocks

Surrogate Models

Build simple and accurate models with a functional form tailored for an optimization framework

Optimization Model

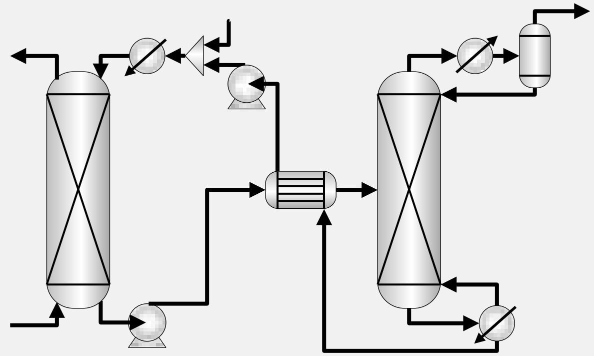
Add algebraic constraints *h(x)*=0: design specs, heat/mass balances, and logic constraints

MODELING PROBLEM STATEMENT

- **Build a model of output variables z as a function of input variables x over a specified interval**

$$x \in \mathbb{R}^D \qquad x^l \le x \le x^u \qquad \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_d \\ \vdots \\ x_D \end{pmatrix}$$

Process simulation

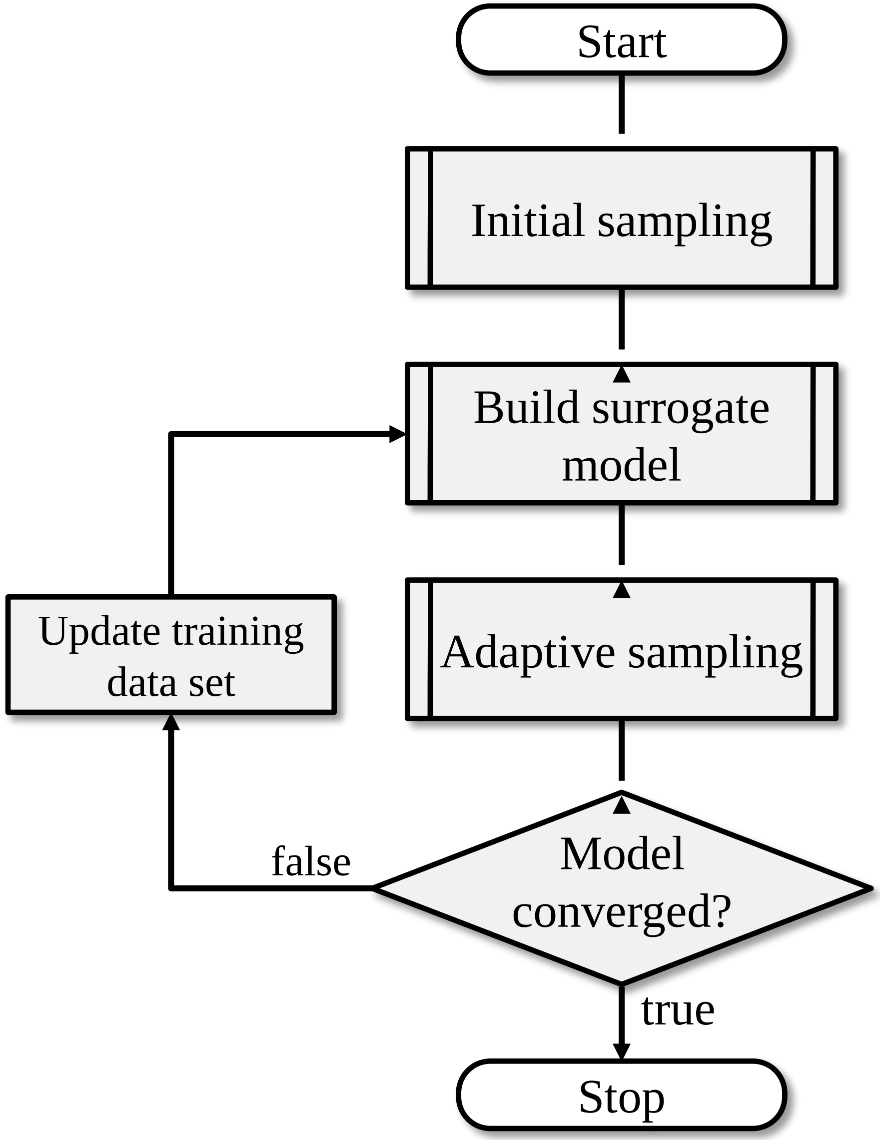


$$\begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_k \\ \vdots \\ z_K \end{pmatrix} \qquad z \in \mathbb{R}^K \qquad z = f(x)$$

Independent variables:
Operating conditions, inlet flow properties, unit geometry

Dependent variables:
Efficiency, outlet flow conditions, conversions, heat flow, etc.

ALGORITHMIC FLOWSHEET

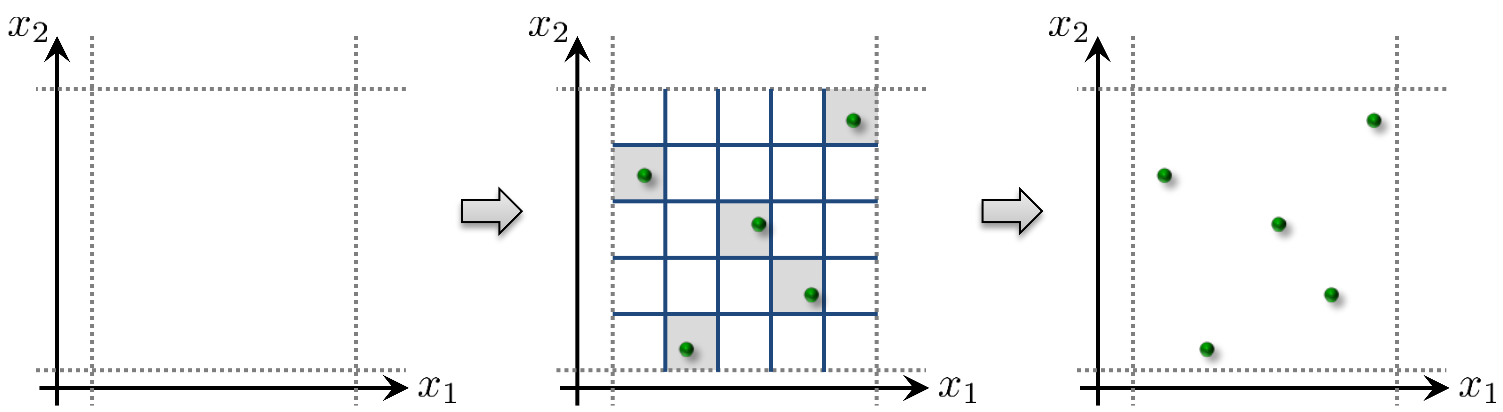


DESIGN OF EXPERIMENTS

- **Goal: To generate an initial set of input variables to evenly sample the problem space**

$$x = \begin{pmatrix} x^1 & x^2 & \cdots & x^i & \cdots & x^N \end{pmatrix} \qquad x^i = \begin{pmatrix} x_1^i \\ x_2^i \\ \vdots \\ x_d^i \\ \vdots \\ x_D^i \end{pmatrix}$$

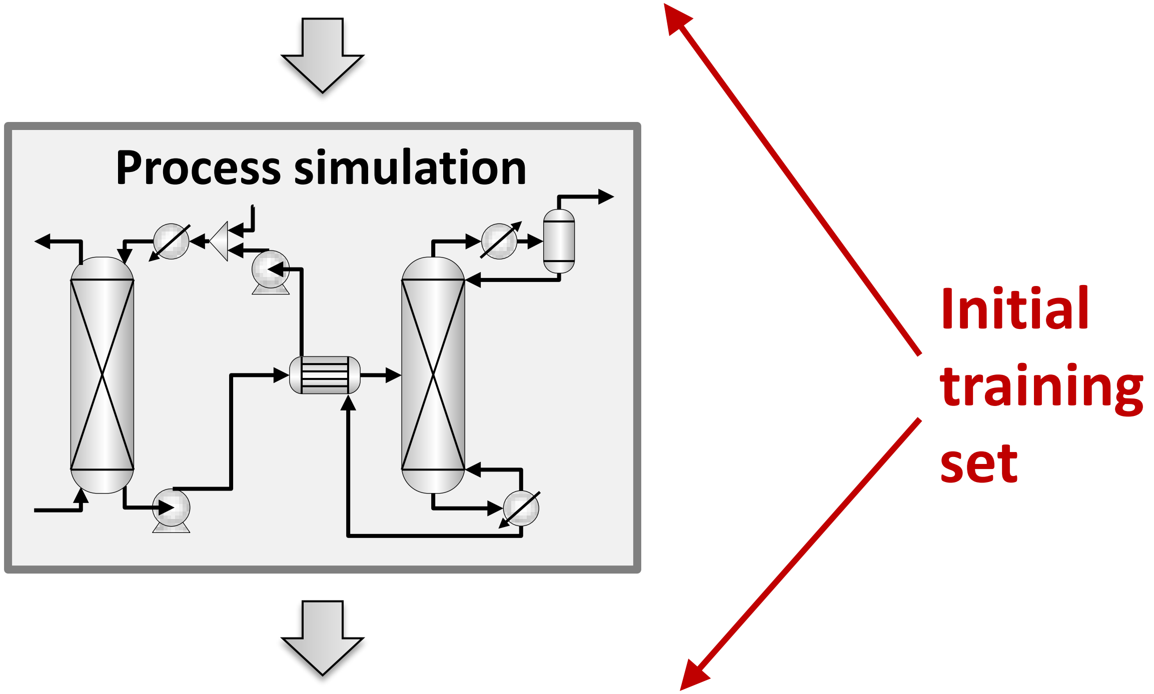
- **Latin hypercube design of experiments** [McKay et al., 79]



INITIAL SAMPLING

- **After running the design of experiments, we will evaluate the black-box function to determine each z^i**

$$x = \begin{pmatrix} x^1 & x^2 & \cdots & x^i & \cdots & x^N \end{pmatrix}$$



$$z = \begin{pmatrix} z^1 & z^2 & \cdots & z^i & \cdots & z^N \end{pmatrix}$$

MODEL IDENTIFICATION

- **Goal: Identify the functional form and complexity of the surrogate models**

$$z = f(x)$$

- **Functional form:**
  - **General functional form is unknown: Our method will identify models with combinations of simple basis functions**

Category		$X_j(x)$
I.	Polynomial	$(x_d)^\alpha$
II.	Multinomial	$\prod_{d \in \mathcal{D}' \subseteq \mathcal{D}} (x_d)^{\alpha_d}$
III.	Exponential and logarithmic forms	$\exp\left(\frac{x_d}{\gamma}\right)^\alpha$, $\log\left(\frac{x_d}{\gamma}\right)^\alpha$
IV.	Expected bases	From experience, simple inspection, physical phenomena, etc.

BEST SUBSET METHOD

- **Surrogate subset model:**

$$\hat{z}(x) = \sum_{j \in \mathcal{S}} \beta_j X_j(x)$$

- **Mixed-integer surrogate subset model:**

$$\hat{z}(x) = \sum_{j \in \mathcal{B}} (y_j \beta_j) X_j(x) \quad \text{such that} \quad \begin{array}{ll} y_j = 1 & j \in \mathcal{S} \\ y_j = 0 & j \notin \mathcal{S} \end{array}$$

- **Generalized best subset problem mixed-integer formulation:**

$$\begin{array}{ll} \min\limits_{\beta, y} & \Phi(\beta, y) \\ \text{s.t.} & y_j = \{0, 1\} \end{array}$$

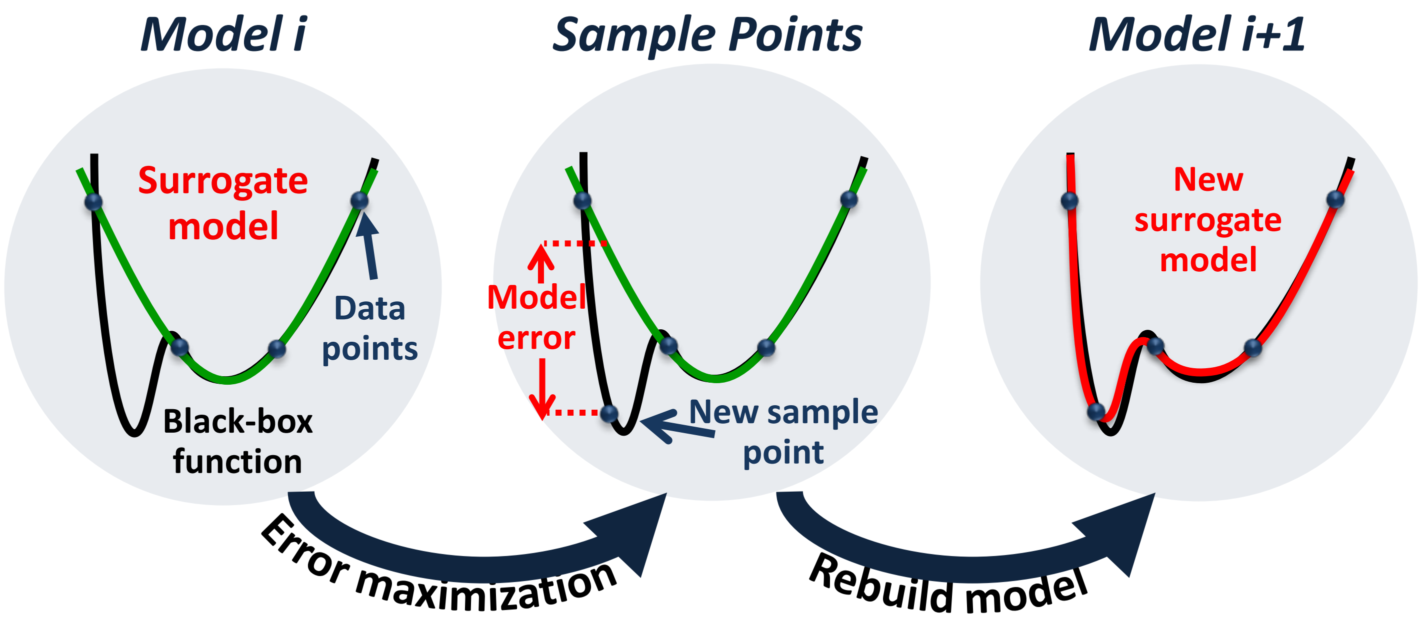
FINAL BEST SUBSET MODEL

$$
\begin{aligned}
\min \quad & SE = \sum_{i=1}^{N} \left| z_i - \sum_{j \in \mathcal{B}} \beta_j X_{ij} \right| \\
\text{s.t.} \quad & \sum_{j \in \mathcal{B}} y_j = T \\
& -U(1 - y_j) \leq \sum_{i=1}^{N} X_{ij} \left(z_i - \sum_{j \in \mathcal{B}} \beta_j X_{ij} \right) \leq U(1 - y_j) && j \in \mathcal{B} \\
& \beta^l y_j \leq \beta_j \leq \beta^u y_j && j \in \mathcal{B} \\
& y_j \in \{0, 1\} && j \in \mathcal{B} \\
& \beta_j \in \left[\beta_j^l, \beta_j^u\right] && j \in \mathcal{B}
\end{aligned}
$$

- **This model is solved for increasing values of *T* until the *AICc* worsens**

ADAPTIVE SAMPLING

- **Goal: Choose new locations to sample that can best be used to improve the model**
- **Solution: Search the problem space for areas of model inconsistency or model mismatch**



ADAPTIVE SAMPLING

- **Goal: Search the problem space for areas of model inconsistency or model mismatch**

- **More succinctly, we are trying to find points that maximizes the model error with respect to the independent variables**

Surrogate model

$$\max_{x} \left(\frac{z(x) - \hat{z}(x)}{z(x)} \right)^2$$

  – **Optimized using a black-box or derivative-free solver (SNOBFIT)** [Huyer and Neumaier, 08]

COMPUTATIONAL TESTING

- **Surrogate generation methods have been implemented into a package:**

ALAMO

(Automated Learning of Algebraic Models for Optimization)

- **Modeling methods compared**
  - **MIP – Proposed methodology**
  - **EBS – Exhaustive best subset method**
    - ***Note: due to high CPU times this was only tested on smaller problems***
  - **LASSO – The lasso regularization**
  - **OLR – Ordinary least-squares regression**

- **Sampling methods compared**
  - **DFO – Proposed error maximization technique**
  - **SLH – Single Latin hypercube (no feedback)**

DESCRIPTION – TEST SET A

- **Two and three input black-box functions randomly chosen basis functions available to the algorithms with varying complexity from 2 to 10 terms**

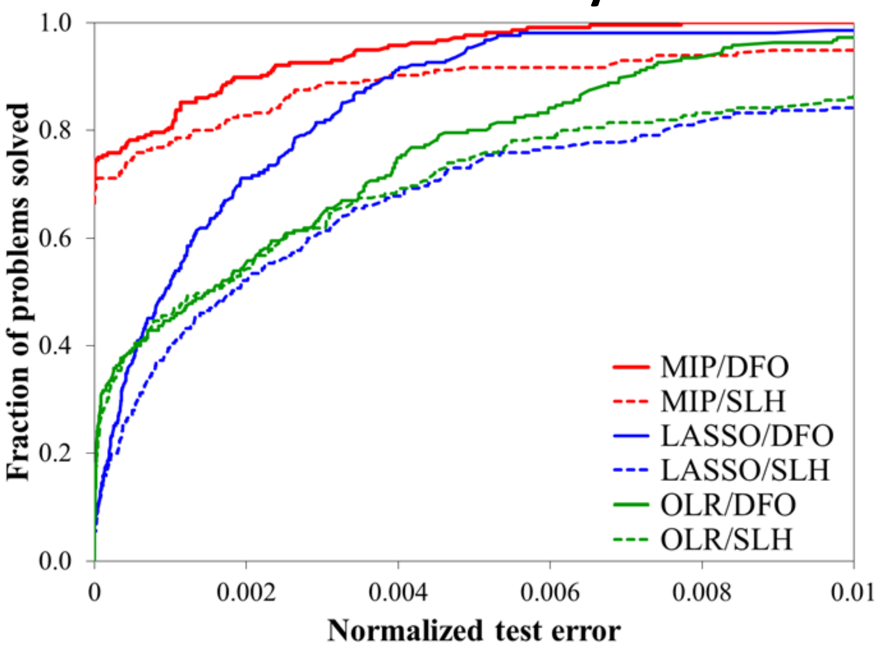
- **Basis functions allowed:**

Category		$X_j(x)$	Parameters used
I.	Polynomial	$(x_d)^\alpha$	$\alpha = \{\pm 3, \pm 2, \pm 1, \pm 0.5\}$
II.	Multinomial	$\prod_{d \in \mathcal{D}' \subseteq \mathcal{D}} (x_d)^{\alpha_d}$	for $\lvert\mathcal{D}'\rvert = 2 \quad \alpha = \{\pm 2, \pm 1, \pm 0.5\}$ for $\lvert\mathcal{D}'\rvert = 3 \quad \alpha = \{\pm 1\}$
III.	Exponential and logarithmic forms	$\exp\left(\frac{x_d}{\gamma}\right)^\alpha, \log\left(\frac{x_d}{\gamma}\right)^\alpha$	$\alpha = 1, \ \gamma = 1$

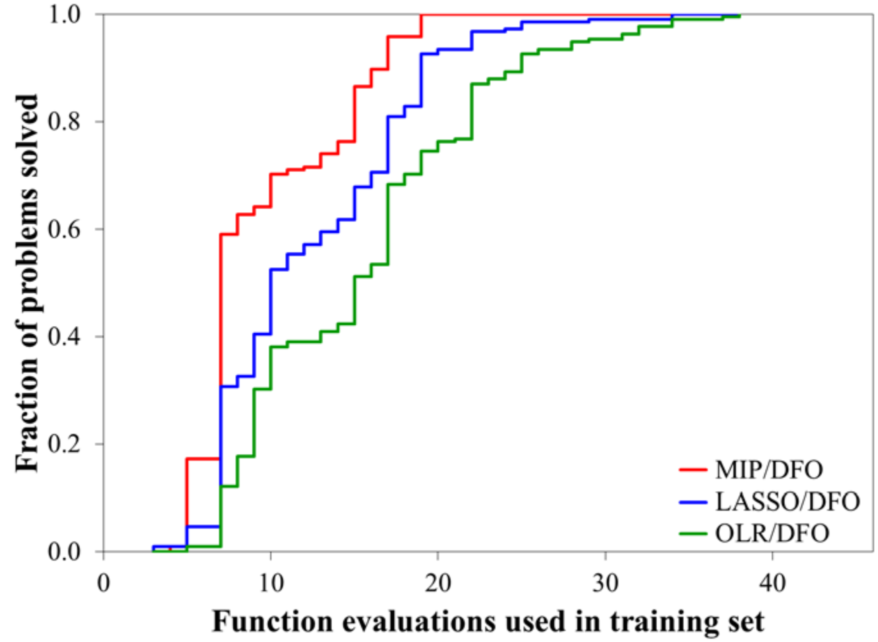
True basis function coefficients were randomly chosen from a uniform distribution where $\beta \in [-1, 1]$.

RESULTS – TEST SET A

Model accuracy



Function evaluations



45 test problems, repeated 5 times, tested against 1000 independent data points

MODEL COMPLEXITY – TEST SET A

No. inputs	No. true terms	MIP/ DFO	MIP/ SLH	EBS/ DFO	EBS/ SLH	LASSO/ DFO	LASSO/ SLH	OLR/ DFO	OLR/ SLH
2	2	2	[2, 2]	2	2	[6, 8]	[6, 11]	[12, 15]	[12, 15]
2	3	3	3	3	3	[5, 12]	[5, 10]	[12, 14]	[12, 14]
2	4	[3, 4]	[3, 4]	[3, 4]	[3, 4]	[8, 11]	[8, 10]	[11, 12]	[11, 12]
2	5	[2, 4]	[2, 4]	[2, 5]	[2, 5]	[3, 12]	[4, 11]	[10, 16]	[10, 16]
2	6	[5, 6]	[6, 6]	[5, 6]	[6, 6]	[7, 10]	[6, 7]	[11, 13]	[11, 13]
2	7	[4, 6]	[4, 6]	[4, 7]	[4, 7]	[7, 11]	[6, 12]	[8, 13]	[8, 13]
2	8	[4, 5]	[5, 6]	[4, 5]	[5, 6]	[6, 8]	[6, 9]	[10, 15]	[10, 15]
2	9	[4, 6]	[4, 6]	NA	NA	[6, 14]	[7, 12]	[10, 17]	[10, 17]
2	10	[4, 8]	[4, 8]	NA	NA	[5, 14]	[7, 14]	[10, 14]	[10, 14]
3	2	[2, 3]	[2, 3]	NA	NA	[6, 12]	[7, 13]	[27, 29]	[27, 29]
3	3	[3, 3]	[3, 3]	NA	NA	[8, 16]	[7, 15]	[19, 22]	[19, 22]
3	4	4	[3, 4]	NA	NA	[10, 13]	[9, 10]	[16, 21]	[16, 21]
3	5	5	5	NA	NA	[11, 17]	[9, 15]	[15, 23]	[15, 23]
3	6	[5, 6]	[6, 6]	NA	NA	[9, 18]	[10, 13]	[15, 26]	[15, 26]
3	7	7	[7, 8]	NA	NA	[10, 22]	[10, 22]	22	22

DESCRIPTION – TEST SET B

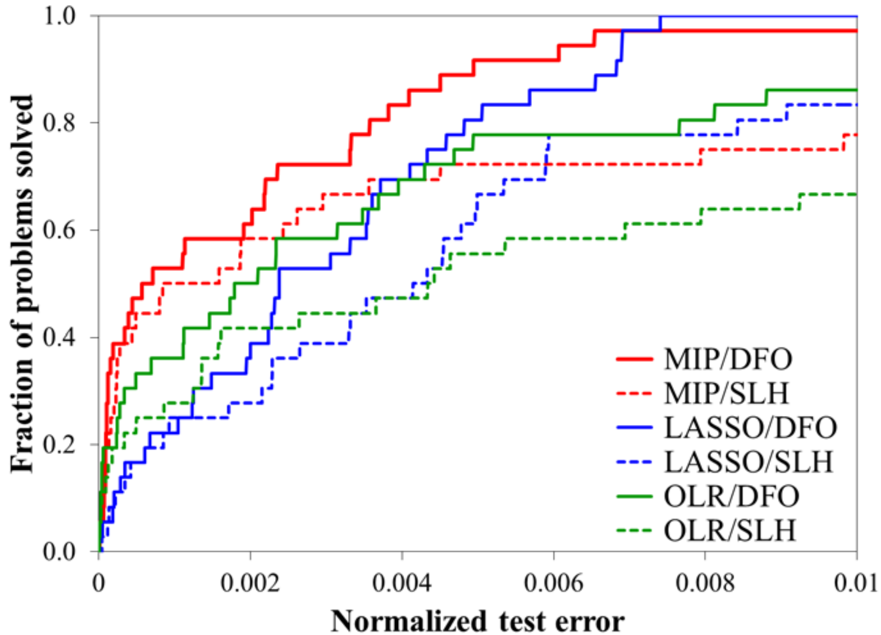
- **Two input black-box functions with basis functions unavailable to the algorithms with**

Function type	Functional form
I	$z(x) = \beta x_i^\alpha \exp(x_j)$
II	$z(x) = \beta x_i^\alpha \log(x_j)$
III	$z(x) = \beta x_1^\alpha x_2^\nu$
IV	$z(x) = \dfrac{\beta}{\gamma + x_i^\alpha}$

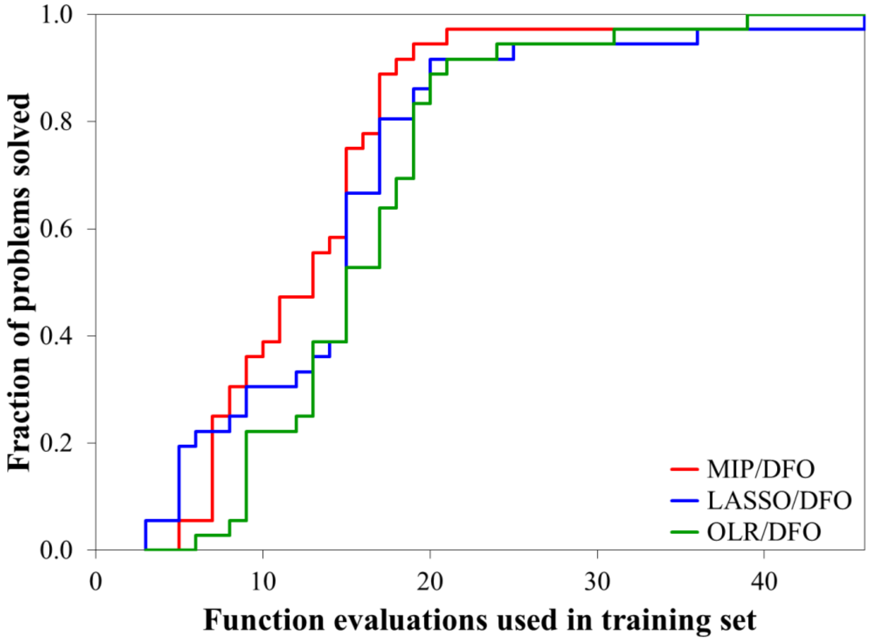
with true parameters chosen from a uniform distribution where $\beta \in [-1, 1]$, $\alpha, \nu \in [-3, 3]$, $\gamma \in [-5, 5]$, and $i, j \in \{1, 2\}$.

RESULTS – TEST SET B

Model accuracy



Function evaluations



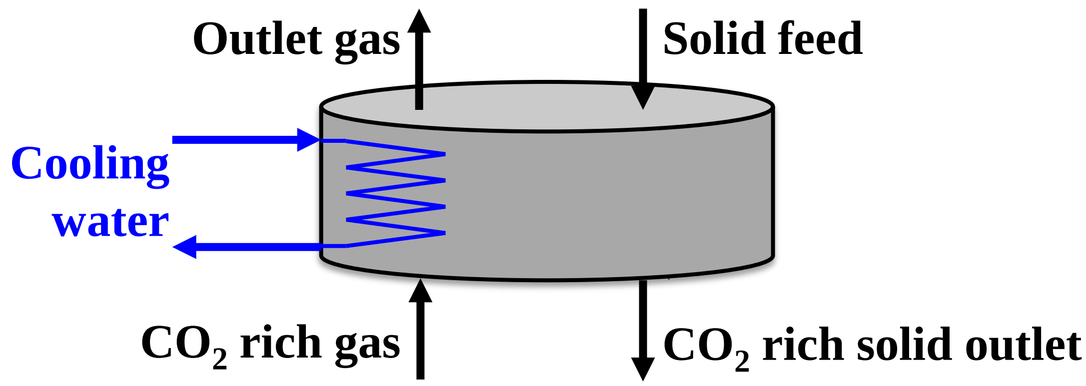
12 test problems, repeated 5 times, tested against 1000 independent data points

MODEL COMPLEXITY – TEST SET B

True function type	Function ID	MIP/ DFO	MIP/ SLH	LASSO/ DFO	LASSO/ SLH	OLR/ DFO	OLR/ SLH
I	a	5	5	[3, 5]	[4, 9]	[6, 17]	[6, 17]
I	b	[4, 10]	[4, 10]	[10, 14]	[5, 8]	[8, 17]	[8, 17]
I	c	[3, 10]	[6, 9]	[8, 9]	[4, 10]	[13, 17]	[13, 17]
II	a	[4, 6]	[4, 10]	[8, 15]	[7, 9]	[15, 19]	[15, 19]
II	b	[1, 7]	[1, 9]	[13, 16]	[11, 17]	[13, 30]	[13, 30]
II	c	[5, 12]	[5, 12]	[9, 13]	[9, 16]	[9, 19]	[9, 19]
III	a	[3, 4]	[1, 4]	[2, 5]	[2, 5]	[9, 20]	[9, 20]
III	b	4	[1, 4]	5	5	[9, 20]	[9, 20]
III	c	[3, 4]	[3, 4]	[5, 8]	[5, 9]	[18, 24]	[18, 24]
IV	a	[7, 8]	[4, 10]	[8, 17]	[11, 18]	[13, 19]	[13, 19]
IV	b	[8, 9]	[9, 10]	[8, 12]	[10, 14]	[9, 17]	[9, 17]
IV	c	[6, 9]	[9, 10]	[5, 13]	[4, 12]	[13, 15]	[13, 15]

BUBBLING FLUIDIZED BED

Bubbling fluidized bed adsorber diagram

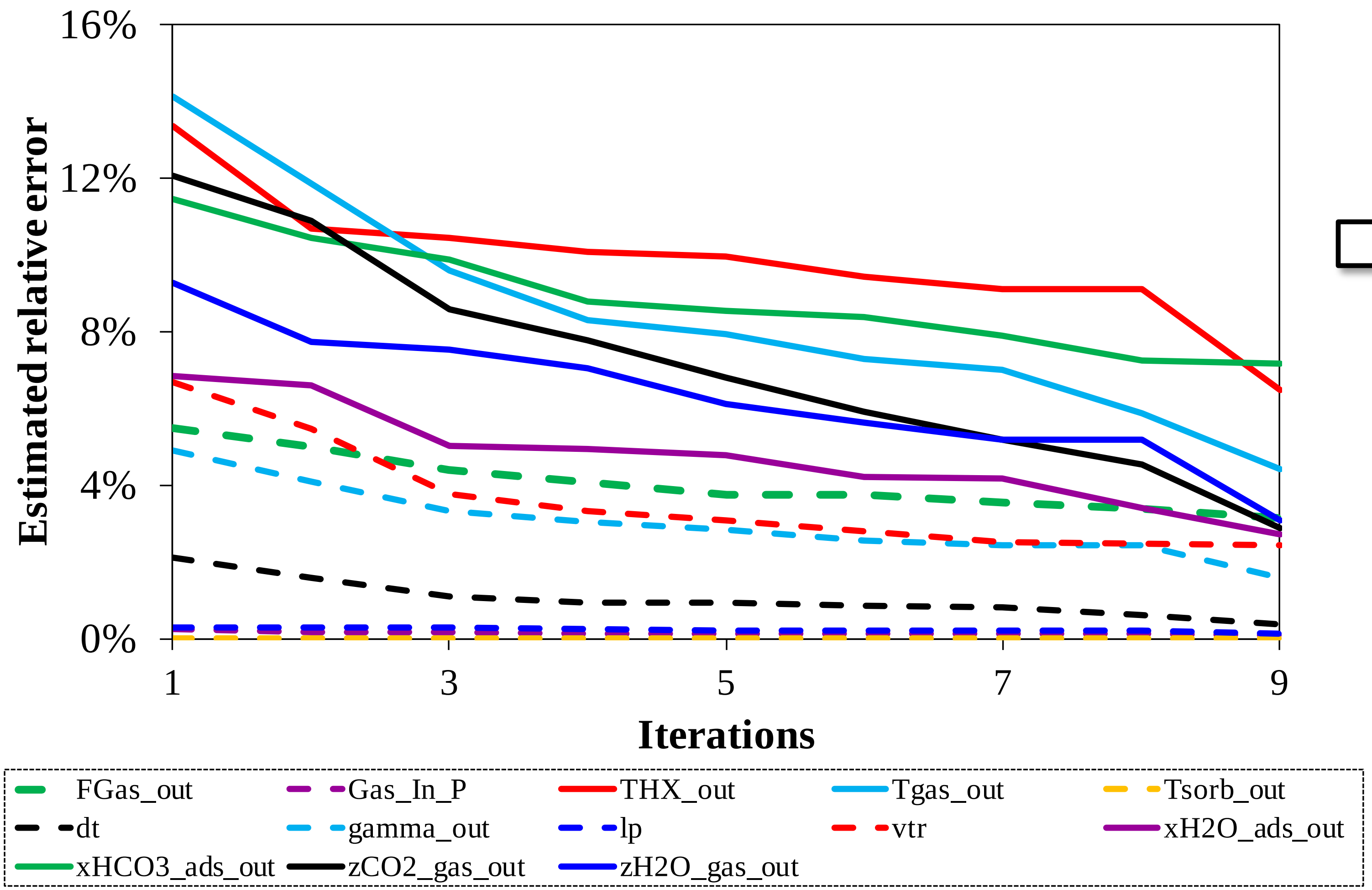


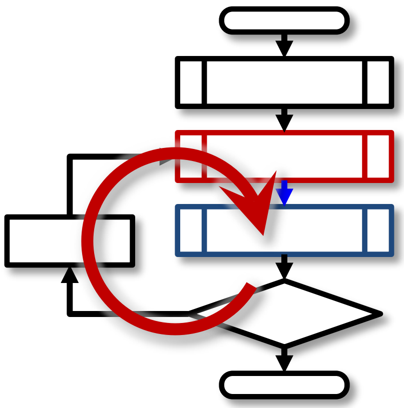
- **Model inputs (14 total)**
  - Geometry (3)
  - Operating conditions (4)
  - Gas mole fractions (2)
  - Solid compositions (2)
  - Flow rates (4)
- **Model outputs (13 total)**
  - Geometry required (2)
  - Operating condition required (1)
  - Gas mole fractions (2)
  - Solid compositions (2)
  - Flow rates (2)
  - Outlet temperatures (3)
  - Design constraint (1)

Model created by Andrew Lee at the National Energy and Technology Laboratory

ADAPTIVE SAMPLING

Progression of mean error through the algorithm

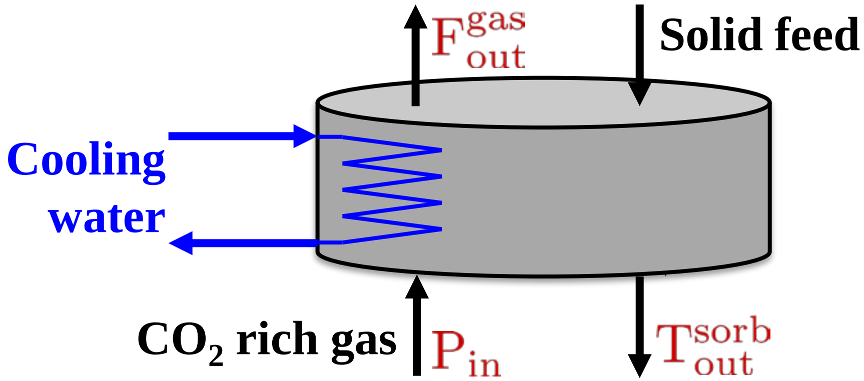




Initial data set: 137 pts

Final data set: 261

EXAMPLE MODELS



$$\mathrm{P_{in}} = 1.0\,\mathrm{P_{out}} + 0.0231\,L_b - 0.0187\,\ln(0.167\,L_b) - 0.00626\,\ln(0.667\,\mathrm{v_{gi}}) - \frac{51.1\,\mathrm{xHCO3^{ads}_{in}}}{\mathrm{F^{gas}_{in}}}$$

$$T^{\mathrm{sorb}}_{\mathrm{out}} = 1.0\,\mathrm{T^{gas}_{in}} - \frac{\left(1.77\cdot 10^{-10}\right)\,\mathrm{NX}^2}{\gamma^2} - \frac{3.46}{\mathrm{NX\,T^{gas}_{in}\,T^{sorb}_{in}}} + \frac{1.17\cdot 10^4}{\mathrm{F^{sorb}\,NX\,xH2O^{ads}_{in}}}$$

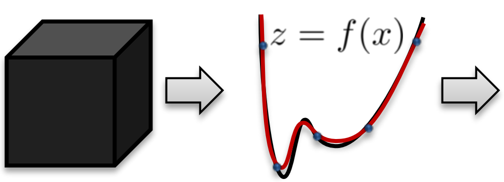
$$F^{\mathrm{gas}}_{\mathrm{out}} = 0.797\,\mathrm{F^{gas}_{in}} - \frac{9.75\,\mathrm{T^{sorb}_{in}}}{\gamma} - 0.77\,\mathrm{F^{gas}_{in}\,xCO2^{gas}_{in}} + 0.00465\,\mathrm{F^{gas}_{in}\,T^{sorb}_{in}} - 0.0181\,\mathrm{F^{gas}_{in}\,T^{sorb}_{in}\,xH2O^{gas}_{in}}$$

CONCLUSIONS

- **The algorithm we developed is able to model black-box functions for use in optimization such that the models are**
  - ✓ **Accurate**
  - ✓ **Tractable in an optimization framework (low-complexity models)**
  - ✓ **Generated from a minimal number of function evaluations**
- **Surrogate models can then be incorporated within a optimization framework with global objective functions and additional constraints**
- **http://archimedes.cheme.cmu.edu/?q=alamo**

ALAMO

Automated **L**earning of **A**lgebraic **M**odels for **O**ptimization



$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & g(x) = 0 \end{aligned}$$

Disclaimer

This presentation was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.