Optimization-based Machine Learning

Marissa Engle[a], Zachary Wilson[a], David Miller[b], Nick Sahinidis[a]
[a]Carnegie Mellon University, [b]National Energy Technology Laboratory

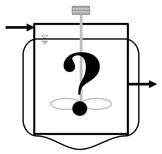
IDAES Institute for the Design of Advanced Energy Systems

ALAMO a black-box modeling tool

RIPE

HELMET

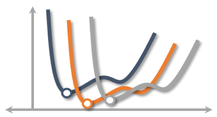
Purpose



Machine learning approaches for automated building of algebraic models

Build models from data ready for optimization or inclusion in flowsheet

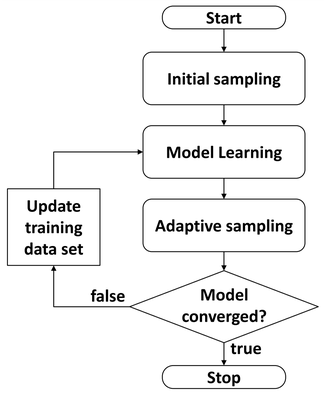
Consideration of alternative reaction mechanism



Incorporation of multiple thermodynamic properties into one equation of state

Estimation of uniquely identifiable parameters and confidence regions

Data Driven Model Learning



Feature selection

$$\hat{z}(x) = \beta_0 + \beta_1 f_1(x) + \beta_2 f_2(x) + \beta_3 f_3(x) + \beta_4 f_4(x) + \beta_5 f_5(x) + \dots$$

$$\hat{z}(x) = 2 + x_2 + 5\, e^{x_1}$$

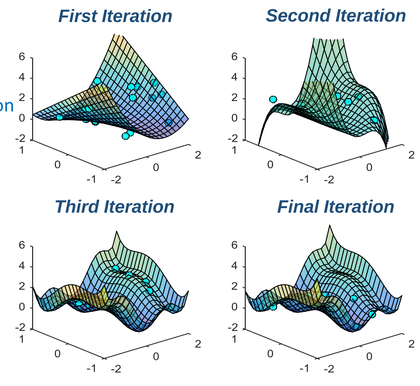
Select subset that balances model fit against model complexity
Obtained via supervised learning algorithm

Error maximization sampling

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

Derivate free design of experiments

- Supervised learning of algebraic models
- ALAMO learns linear models of features tailored for optimization
- RIPE learns reaction network from process data
- HELMET learns a Helmholtz energy equation of state from thermodynamic properties



Iteration	N	R^2_{val}	$\|\beta\|_0$
1	17	0.56	2
2	23	0.61	3
3	31	0.92	11
4	37	0.98	6

Corrupted six hump camel function

$$f(x) = (4 - 2.1x_1^2 + x_1^4/3)x_1^2 + x_1x_2 + x_2^2(4x_2^2 - 4) + \epsilon$$

ALAMO surrogate model

$$f_{Alamo} = 4.56x_1^2 - 3.16x_2^2 - 2.41x_1^4 + 3.07x_2^4 + 0.38x_1^6 + 1.09x_1x_2 - 0.28$$

Known Minimum
$f(0.0898, -0.7126) = -1.0316$

Surrogate Minimum
$f(0.0881, -0.7114) = -1.0291$

Reaction Identification Parameter Estimation

Formulating the RIPE problem

Postulated reactions for target network

$$\hat{r}_{i,s} = \sum_{m=1}^{N_r} \sum_{h=1}^{N_h} v_{sh} k_{mh}(T_i) Act_{ismh} \qquad k_{mh}(T_i) = k^0_{mh} \exp\left(-\frac{E_{mh}}{RT}\right)$$

where m, h index over reaction mechanisms and stoichiometries

Mass balance gives target values

$$\frac{dc_{i,s}}{dt} = \frac{\dot{\boldsymbol{v}}}{V}\left(c^{in}_{i,s} - c^{out}_{i,s}\right) + r^{obs}_{i,s}$$

$$r^{obs}_{i,\boldsymbol{s}} = \frac{\left((c^{out}_{i,s} + \epsilon) - c^{in}_{i,s}\right)}{\tau_i}$$

33 reactions postulated | **True Network identified**

4 alternative mechanisms | **BIC used to size model**

Solution Details

One reaction

$$2A + 2C \rightarrow K + 2E + 3F + H + 2I$$

Two reactions

$$2B + 2C \rightarrow K + 2E + 3F + H + 2I$$
$$A \overset{ads}{\longleftrightarrow} \mathrm{B}$$

• • •

Six reaction network

$$A \overset{ads}{\longleftrightarrow} \mathrm{B} \overset{E}{\rightarrow} F + \frac{1}{2}J \rightarrow 2G + 2I$$

$$B + C \rightarrow K + E + F + I + 2H$$
$$C \leftrightarrow D + E$$
$$B + 2D \overset{E}{\rightarrow} 2F + 2\mathrm{H}$$

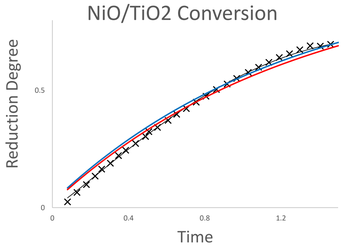
Case Study: Chemical Looping Combustion

Two Catalyst : NiO/Al_2O_3, NiO/TiO_2

$$\frac{dX}{dt} = k(T) f(X) g(y_{gas})$$

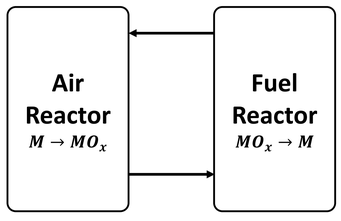
Alamo generates explicit $\frac{dX}{dt}$

All selected models $R^2 = 0.99$



$$\frac{dX_{Alamo}}{dt} = \frac{1}{2}\beta t^{-\frac{1}{2}}$$

$$\frac{dX_{RIPE}}{dt} = \frac{dX_{lit}}{dt} = (1 - X)^{\frac{2}{3}}$$



Air Reactor

Catalyst	NiO/Al_2O_3	NiO/TiO_2
$f(X)$	$(1-X)$	$(1-X)^{\frac{2}{3}}$
k_{lit}	0.77	1.15
k	0.8 ± 0.04	1.2 ± 0.1

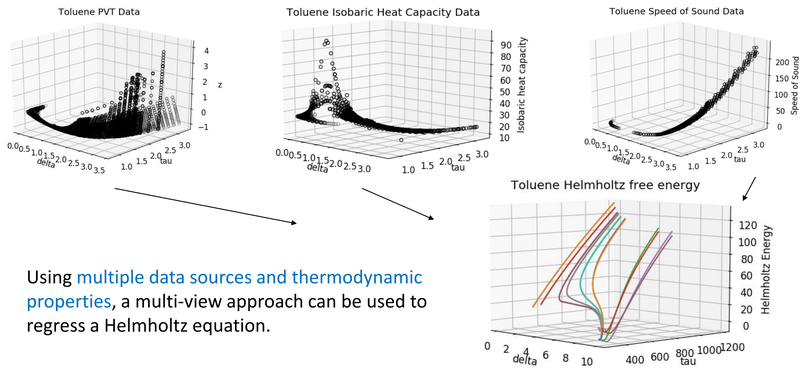
Fuel Reactor

Catalyst	NiO/Al_2O_3	NiO/TiO_2
$f(X)$	$X(1-X)$	$(3/2)(1-X)(-\ln(1-X))^{\frac{5}{6}}$
k_{lit}	0.62	1.66
k	5.6 ± 0.3	1.7 ± 0.03

HELMholtz Energy Thermodynamics

Helmholtz free energy can be related to thermodynamic properties by partial derivatives of density and temperature. This methodology uses a dimensionless Helmholtz function dependent on reduced density, $\delta = \rho/\rho_c$, and inverse reduced temperature, $\tau = T_c/T$.

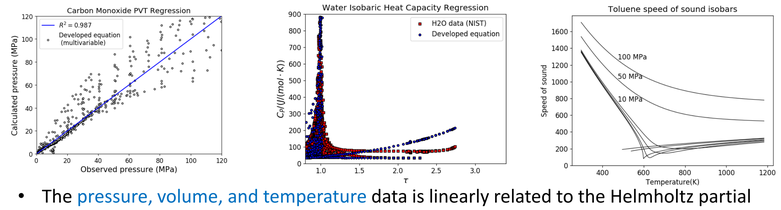
$$\frac{a(\rho,T)}{RT} = \alpha(\delta,\tau) = \overbrace{\alpha^o(\delta,\tau)}^{\text{Ideal}} + \overbrace{\alpha^r(\delta,\tau)}^{\text{Residual}}$$



Using multiple data sources and thermodynamic properties, a multi-view approach can be used to regress a Helmholtz equation.

Case study: water, toluene, & carbon monoxide

Pure chemical compounds, such as water, are **well-studied**, but remain **difficult to characterize**. These systems have lots of different thermodynamic property data that can be simultaneously regressed for us to develop and validate our approach.



- The pressure, volume, and temperature data is linearly related to the Helmholtz partial derivative and with a constrained regression provides a good initial fit
- The fit to isobaric heat capacity has the lowest R^2 due to the rapid change in behavior near the critical point
- The speed of sound is accurately measured and the regression methodology penalizes heavily the residuals resulting in $R^2 \approx 1$

Conclusion

- With ALAMO, RIPE, and HELMET, we use advanced data modeling and leverage machine learning techniques to regress simple, accurate models for use in algebraic optimization
- This regression toolkit extends the power of the IDAES framework to process data, or black-box models lacking an algebraic form

Contact

Marissa Engle mengle@cmu.edu
Zachary Wilson zwilson@cmu.edu
Nick Sahinidis sahinidis@cmu.edu

NATIONAL ENERGY TECHNOLOGY LABORATORY

BERKELEY LAB

Sandia National Laboratories

Carnegie Mellon

WestVirginiaUniversity

U.S. DEPARTMENT OF ENERGY