

# Simultaneous Parameter Estimation in Reactive-Solvent-Based Processes

John C. Eslick<sup>a\*</sup>, Paul T. Akula<sup>b</sup>, Debansu Bhattacharyya<sup>b\*</sup>, David C. Miller<sup>a</sup>

<sup>a</sup>*National Energy Technology Laboratory, 626 Cochrans Mill Rd., Pittsburgh, PA 15236, USA*

<sup>b</sup>*Department of Chemical and Biomedical Engineering, West Virginia University, Morgantown, WV 26506, USA*

*John.Eslick@netl.doe.gov*

## Abstract

Accurate estimation of model parameters is difficult for reactive solvent-based processes due to coupled effects of mass transfer, heat transfer, and reaction kinetics. While commercial process simulators provide capabilities for parameter estimation, it is generally difficult to simultaneously regress parameters of different submodels (e.g. parameters of physical properties and mass transfer models). Morgan et al. (2018) and Chinen et al. (2018) have demonstrated the value of the simultaneous parameter estimation approach in developing highly predictive models. The approach used by Chinen et al. (2018) embedded a simulation of a CO<sub>2</sub> capture process in an external derivative free optimization (DFO) framework. While this approach works well in many cases, it is computationally inefficient, difficult to set up, and limited by the amount of data that can be used. This paper presents an extension to that work where the process model and a generalized parameter estimation tool are developed in the Institute for Design of Advanced Energy Systems (IDAES) process systems engineering (PSE) framework. The IDAES framework allows an entire process model to be implemented in an equation-oriented framework with access to state-of-the-art optimization methods. Improved computational efficiency and ease of implementation are demonstrated.

**Keywords:** Simultaneous parameter estimation, optimization, CO<sub>2</sub> capture, IDAES

## 1. Introduction

Accurate estimation of model parameters is important in the development of predictive process models; however, this is difficult for reactive solvent-based processes due to the coupled effects of mass transfer, heat transfer, and reaction kinetics. Experimental data spanning different scales and operating regimes are required to obtain a process model that is predictive for different scales and widely varying operating conditions. Simultaneous parameter estimation for solvent based CO<sub>2</sub> capture systems has been described and a methodology has been presented using chemical process simulators embedded in a DFO framework (Chinen et al. 2018). The focus of this paper is to extend that work by implementing a generic framework that will support the approach within the IDAES PSE framework, which provides a modular process modeling framework on top of the Pyomo algebraic modeling language (AML) (Hart et al., 2011; Hart et al., 2017).

The novelty of the approach is that it facilitates large-scale parameter estimation for hundreds of parameters at multiple scales embedded in complex process models and submodels in a way that is not generally possible with commercial process simulators.

Depending on the complexity and robustness of the process models, over 100 parameters, thousands of experimental data sets, and several hundred thousand equality and inequality constraints can be handled through decomposition methods. The upper limit of this methodology has yet to be identified.

The IDAES framework provides basic modular flowsheeting features and a model library allowing optimization problems to be implemented much more quickly than in a general AML. The benefit of this approach enables the process model and all submodels to be implemented in a common framework where all variables and parameters are easily accessible. In addition, full derivative information is available enabling the use of state-of-the-art optimization solvers and decomposition methods through Pyomo. These features allow the problem to be implemented in a more straightforward way and solved with significantly reduced computation times and superior optimality conditions. The approach presented here can readily solve problems comprising multiple models and submodels while utilizing data from multiple sources and scales.

## 2. Simultaneous Parameter Estimation

Simultaneous parameter estimation is a technique of regressing parameters of a process model and various interrelated submodels at the same time. This approach allows data from multiple scales and system configurations to be used simultaneously to improve the predictiveness of a model. Features of each coupled sub-model are better identified over a wider domain by this approach. The simultaneous parameter estimation problem formulation used in this work is given by Eq. 1.

$$\min \sum_{i \in R} \sum_{j \in S_i} \frac{\sum_{k \in T_i} (\hat{y}_{i,j,k} - y_{i,j,k})^2}{\sigma_{i,j}^2} \quad (1)$$

$$f_{i,k}(\theta, x_{i,k}, \hat{y}_{i,k}, z_{i,k}) = 0, \quad \forall i \in R, k \in T_i$$

Where:  $R$  = Data sets (e.g., packed column data and wetted wall column data)

$S_i$  = Set of output measurements for data set  $i$

$T_i$  = Set of experiments in data set  $i$

$\hat{y}_{i,j,k}$  = Predicted output for measurement  $j$  for experiment  $k$  in set  $i$

$y_{i,j,k}$  = Measured output for measurement  $j$ , for experiment  $k$  in set  $i$

$f_{i,k}$  = Model equations for experiment  $k$  in set  $i$

$x_{i,k}$  = Measured inputs for experiment  $k$  in set  $i$

$z_{i,k}$  = Other model variables for experiment  $k$  in set  $i$

$\theta$  = Model Parameters (same across all experiments)

$\sigma_{i,j}^2$  = Variance of measurement  $j$  in data set  $i$  (for scaling)

Challenges with the simultaneous parameter estimation approach arise due to the size and complex nonlinear nature of the problem. The number of equations to be solved becomes large when considering large data sets, due to the optimization problem containing equations for a full process model for each experiment. This requires the models to be robust over a wide range of conditions.

To overcome these challenges, the IDAES framework was used for this work. The PySP module in Pyomo provides several decomposition techniques for stochastic programming (Watson et al., 2012). The parameter estimation problem is equivalent to a 2-stage stochastic programming problem with no recourse. In the 2-stage stochastic programming problem, a set of scenarios are generated with different parameter values

based on parameter uncertainty distributions, and the weighted sum of scenario objectives is optimized. First-stage variables must be the same for each scenario, while second-stage variables can be adjusted in each scenario to provide recourse against uncertainty. This problem can be decomposed by solving each scenario separately using one of several algorithms to iteratively converge to a solution. In the case of parameter estimation, each experiment corresponds to a scenario, and the inputs to the experiment correspond to the parameter values for a scenario. The parameters being estimated correspond to first-stage variables, and there are no second-stage (recourse) variables.

### 3. Case study: Amine solvent-based CO<sub>2</sub> absorption

CO<sub>2</sub> capture using aqueous amines is a well-established technology, and a significant amount of bench- and pilot-scale data is available for monoethanolamine (MEA), making it an ideal test case for a parameter estimation framework. This work is applicable to other solvent system as well as other types of process models. Data from two systems are used in this study (Figure 1): (a) a pilot-scale packed column with intercooling (data from the National Carbon Capture Center) and (b) a bench-scale isothermal wetted wall column (WWC) system (Dugas, 2009).

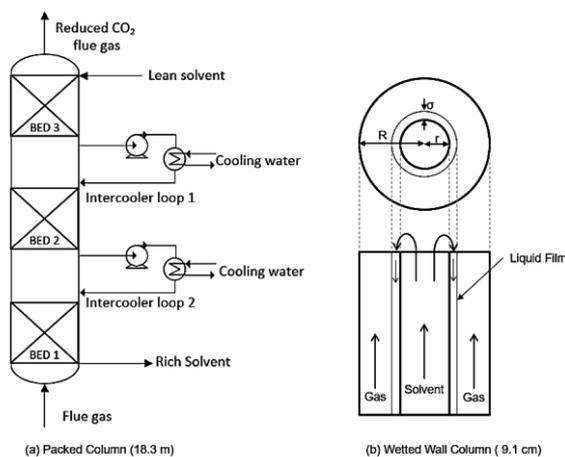


Figure 1: Packed Column and Wetted Wall Column

A 1D rate-based packed absorber column model is implemented. This model is similar to the model of Morgan et al. (2018) with three main differences: (1) an enhancement factor approach is used to characterize the mass transfer resistances through the liquid and gas films as opposed to a rigorous two-film model, (2) the NRTL model is used for calculating the activity coefficients instead of the electrolyte-NRTL (eNRTL) model, and (3) a simplified heat of absorption model is used. A simpler model is used in this work compared to Morgan et al. (2018), and work is ongoing to implement the fully detailed model. The WWC model is developed by considering a column equivalent to the annular space, with appropriate models for the interfacial area and liquid holdup.

Parameters are estimated for the reaction rate constants ( $A$  and  $E$  in Eq. 2), the vapor phase mass transfer coefficient ( $C_v$ ,  $a$ , and  $b$  in Eq. 3), the liquid mass transfer coefficient ( $C_l$  in Eq. 4), and the interfacial area ( $\beta_1$ ,  $\beta_2$ , and  $\beta_3$  in Eq. 5) models. The effective interfacial area correlation is applied to the packed-bed model only.

$$k_i = A_i \exp\left(\frac{E_i}{T}\right), \quad i \in \{\text{H}_2\text{O}, \text{MEA}\} \quad (2)$$

$$k_{v,i} = C_v D_{v,i} \left(\frac{a}{\varepsilon_v d_h}\right)^{1/2} \left(\frac{u_v \rho_v}{a \mu_v}\right)^a \left(\frac{\mu_v}{\rho_v D_{v,i}}\right)^b \quad (3)$$

$$k_{l,\text{CO}_2} = 12^{1/6} C_l \left(\frac{u_l}{\varepsilon_l}\right)^{1/2} \left(\frac{D_{l,\text{CO}_2}}{d_h}\right)^{1/2} \quad (4)$$

$$\left(\frac{a_e}{\beta_1}\right)^{\beta_2} = u_l \beta_3 \left(\frac{\rho_l}{\sigma_l}\right) \quad (5)$$

The objective function is given by Eq. 1. The WWC data set is comprised of 35 experiments where CO<sub>2</sub> flux between the vapor and liquid phase is measured. The packed column data set has 17 points where CO<sub>2</sub> capture fraction is measured. The problem is solved using the IPOPT NLP solver (Biegler and Zavala, 2009).

#### 4. Results

The optimization results are presented in Table 1. The initial parameter values were taken from Aboudheir et al. (2003) (Eq. 2) and Morgan et al. (2015) (Eqns. 3-5). The WWC model R<sup>2</sup> was 81.3 % for the initial and 98.3 % for the optimal parameter values. The packed column R<sup>2</sup> was 75.2 % for the initial and 83.3 % for the optimal parameter values. The root mean squared error is 6.4×10<sup>-5</sup> mol/s/m<sup>2</sup> for the WWC and 1.7 percentage points of CO<sub>2</sub> capture for the packed column models. Some parameters were found to be on their bounds; however, widening the bounds did not yield significant improvement to the fit. Capabilities for parametric uncertainty analysis is currently being developed and will be utilized in future work.

Table 1: Fitted parameters

Model	Kinetics				Mass Transfer				Interfacial Area		
Parameter	$A_{\text{H}_2\text{O}}$	$E_{\text{H}_2\text{O}}$	$A_{\text{MEA}}$	$E_{\text{MEA}}$	$C_v$	$a_v$	$b_v$	$C_l$	$\beta_1$	$\beta_2$	$\beta_3$
Init. value	4.55	3,287	4,610	4,412	0.357	0.750	0.333	0.500	157.0	8.612	1.333
Opt. value*	4.00	4,100	4,800	4,000	0.500	0.495	0.250	0.504	15.0	7.235	-0.50

Figure 2 shows parity plots for the flux in the WWC case and CO<sub>2</sub> capture fraction in the packed column case. The residuals for the packed bed case appear to show a trend suggesting that there are some phenomena not accounted for in the model.

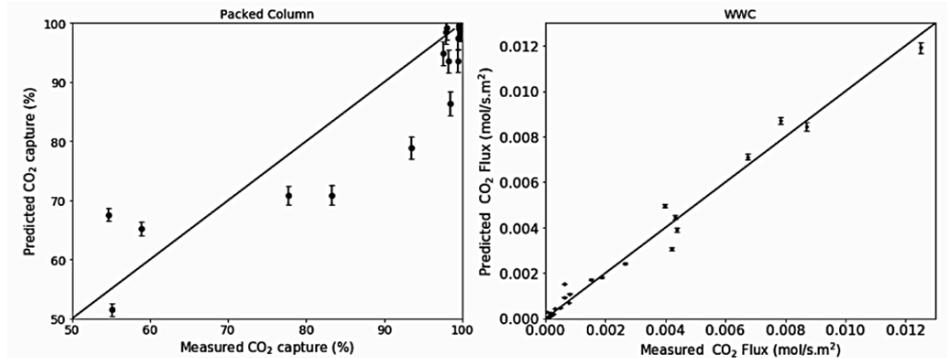


Figure 2: Parity Plots (a) Packed Column, (b) WWC

The packed column temperature profiles in Figure 3 give some insight into this discrepancy among the model used in this work, the Morgan et al. (2018) model, and the plant data. The error bars represent the measurement accuracy based on the type of thermocouples used in the pilot plant. While at low-loading, the temperature profiles agree well, there is a higher discrepancy at high loading. These deviations are within the tolerances given the approximations in the vapor-liquid equilibrium (VLE), mass transfer, and enthalpy models. The e-NRTL model is expected to improve precision by improving the VLE predictions for CO<sub>2</sub> high loading conditions compared with the NRTL model used in this work (Luo et al., 2015; Morgan et al., 2017).

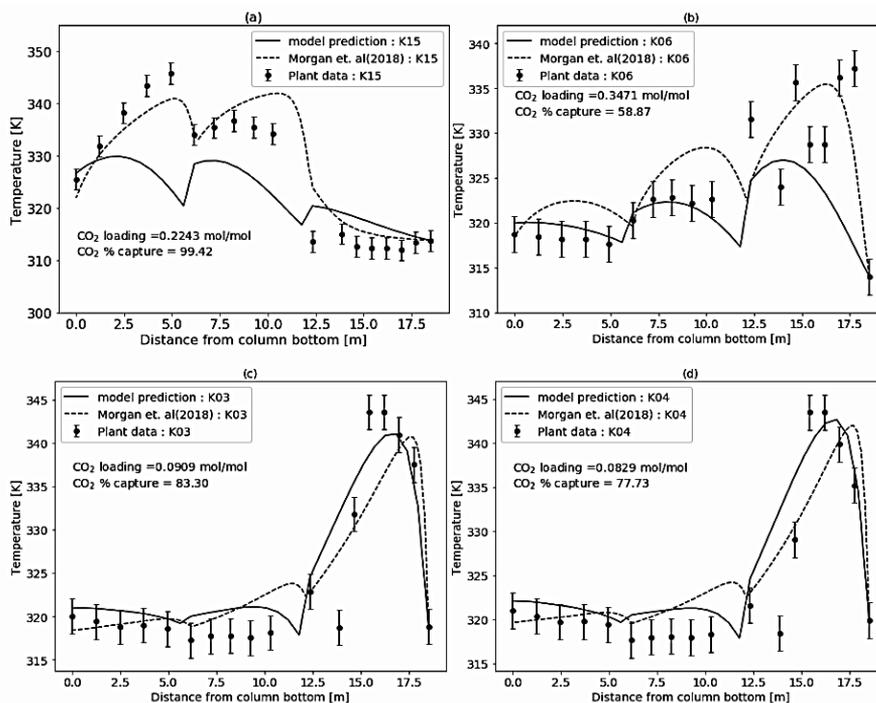


Figure 3: Packed Column Temperature Profiles, this work, National Carbon Capture Center Data, and Morgan et al. (2018)

The parameter estimation problem in this work consisted of 71,768 constraints, many of which are highly non-linear. The computation time required to solve the optimization problem was 206 seconds on an Intel Xeon E3-1505M CPU, which is a substantial improvement compared with the previous parameter estimation framework (Chinen et al. 2018), which took about 36 hours to run on a similar machine.

## 5. Conclusions

A new framework for simultaneous parameter estimation has been demonstrated using a system of industrial significance. The novelty of this system is that it facilitates simultaneous parameter estimation for parameters in multiple submodels embedded in complex process models, which is not generally possible with commercial process simulators. The IDAES framework provides basic modular flowsheeting features and a model library allowing optimization problems to be implemented much more quickly than in a general AML. The result showed a significant improvement in computational

efficiency compared with previous methods presented by Chinen et al. (2018), which used a process simulator embedded in a derivative free optimization framework. It is expected that this will lead to the ability to solve larger, multiscale parameter estimation problems with large amounts of data. Future work will provide enhancements to the parameter estimation framework to enable uncertainty both in parameter estimates and measurements to be considered.

### Acknowledgment and Disclaimer

This research was funded by the Lawrence Berkeley National Laboratory as part of the U.S. DOE's Institute for the Design of Advanced Energy Systems (IDAES). This paper 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.

### References

- A. Aboudheir, P. Tontiwachwuthikul, A. Chakma, R. Idem, 2003, Kinetics of the reactive absorption of carbon dioxide in high CO<sub>2</sub> loaded, concentrated aqueous monoethanolamine solutions, *Chemical Engineering Science*, 58, 5195-5210.
- L. T. Biegler and V. M. Zavala, 2009, Large-scale nonlinear programming using IPOPT: An integrating framework for enterprise-wide dynamic optimization, *Computers & Chemical Engineering*, 33, 3, 575-582.
- A. S. Chinen, J. C. Morgan, B. Omell, D. Bhattacharyya, C. Tong, D. C. Miller, 2018, Development of a gold-standard model for solvent-based CO<sub>2</sub> capture. Part 1: hydraulic and mass transfer models and their uncertainty quantification, Under internal review.
- R. E. Dugas, 2009, Carbon dioxide absorption, desorption, and diffusion in aqueous piperazine and monoethanolamine, Doctoral dissertation, University of Texas at Austin.
- W. E. Hart, J.-P. Watson and D. L. Woodruff, 2011, Pyomo: modeling and solving mathematical programs in Python, *Mathematical Programming Computation*, 3, 3, 219-260.
- W. E. Hart, C.D. Laird, J.-P. Watson, D.L. Woodruff, G.A. Hackebeil, B.L. Nicholson, and J.D. Sirola, 2017, *Pyomo – Optimization Modeling in Python*, Second Edition.
- X. A. Luo, Hartono, S. Hussain, H. F. Svendsen, 2015, Mass transfer and kinetics of carbon dioxide absorption into loaded aqueous monoethanolamine solutions, *Chemical Engineering Science*, 123, 57-69.
- J. C. Morgan, A. S. Chinen, B. Omell, D. Bhattacharyya, C. Tong, D. C. Miller, B. Buschle, M. Lucquiaud, 2018, Development of a gold-standard model for solvent-based CO<sub>2</sub> capture. Part 2: steady-state validation and uncertainty quantification with pilot plant data, Under internal review.
- J. C. Morgan, B. Omell, D. Bhattacharyya, C. Tong, D. C. Miller, 2017, Thermodynamic modeling and uncertainty quantification of CO<sub>2</sub>-loaded aqueous MEA solutions, *Chemical Engineering Science*, 168, 309-324.
- J. C. Morgan, D. Bhattacharyya, C. Tong, D. C. Miller, 2015, Uncertainty quantification of property models: Methodology and its application to CO<sub>2</sub>-loaded aqueous MEA solutions. *AIChE Journal* 61, 6, 1822-1839.
- J.-P. Watson, D. L. Woodruff, W. E. Hart, 2012, PySP: modeling and solving stochastic programs in Python, *Mathematical Programming Computation*, 4, 2, 109-149.