Uncertainty Quantification of Property Models: Methodology and Its Application to CO_2-Loaded Aqueous MEA Solutions

Joshua C. Morgan and Debangsu Bhattacharyya
Dept. of Chemical Engineering, West Virginia University, Morgantown, WV 26505

Charles Tong
Lawrence Livermore National Laboratory, Livermore, CA 94550

David C. Miller
National Energy Technology Laboratory, Pittsburgh, PA 15236

DOI 10.1002/aic.14762
Published online March 9, 2015 in Wiley Online Library (wileyonlinelibrary.com)

Uncertainties in property models can significantly affect the results obtained from process simulations. If these uncertainties are not quantified, optimal plant designs based on such models can be misleading. With this incentive, a systematic, generalized uncertainty quantification (UQ) methodology for property models is developed. Starting with prior beliefs about parametric uncertainties, a Bayesian method is used to derive informed posteriors using the experimental data. To reduce the computational expense, surrogate response surface models are developed. For downselecting the parameter space, a sensitivity matrix-based approach is developed. The methodology is then deployed to the property models for an MEA-CO_2-H_2O system. The UQ analysis is found to provide interesting information about uncertainties in the parameter space. The sensitivity matrix approach is also found to be a valuable tool for reducing computational expense. Finally, the effect of the estimated parametric uncertainty on CO_2 absorption and monoethanolamine (MEA) regeneration is analyzed. © 2015 American Institute of Chemical Engineers *AIChE J*, 61: 1822–1839, 2015
Keywords: uncertainty quantification, property models, monoethanolamine, CO_2 capture

Introduction

The predictive capability of rigorous process models is dependent on the accuracy of the underlying physical property models; however, even the best models have some uncertainty. Thus, it is important that these uncertainties be quantified, so that their effect can be propagated throughout the overall process model to provide reasonable confidence bounds on the model predictions. The sources of uncertainty for property models can be due to the model form and/or the value of the parameters. In addition, uncertainty exists in the measured experimental data on which the model is based. Traditionally, the property models have been developed by proposing some arbitrary or intelligently guessed functional form in terms of one or more dependent or independent variables based on the expected or observed dependencies. The fields of statistical thermodynamics and molecular dynamics are instrumental in developing physical property models from a first principles approach, although some degree of approximation is required.[1,2] For example, the functional forms of quantities such as intermolecular potential are often represented by empirical correlations.[2] As quite often the model parameters, which are generally regressed to some experimental data, have no physical significance, it is difficult to characterize them or provide a bound on them. Therefore, the traditional approach to fitting the model and its parameters can result in uncertainty in the model form and its parameters. This work is focused on development of an approach for determining uncertainty in property models and their parameters using a systematic approach.

The sensitivity of process variables to physical properties has been shown to be very case dependent, as uncertainties in some physical properties have a greater effect on the process performance prediction as compared to others.[3,4] The uncertainty of the physical properties is often more pronounced at lower design temperatures.[4] The high sensitivity of the process model to the VLE model has been reported.[5] Uncertainty analysis is often neglected in many chemical engineering applications, mainly due to lack of systematic approaches for uncertainty quantification (UQ) and capabilities to calibrate the initial estimates of uncertainties by propagating them through the process model and comparing

Disclaimer: This project was funded by the Department of Energy, National Energy Technology Laboratory, an agency of the U.S. Government, through a support contract with URS Energy & Construction, Inc. Neither the U.S. Government nor any agency thereof, nor any of their employees, nor URS Energy & Construction, Inc., nor any of their employees, makes any warranty, expressed 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 U.S. Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the U.S. Government or any agency thereof.

Correspondence concerning this article should be addressed to D. Bhattacharyya at Debangsu.Bhattacharyya@mail.wvu.edu.

© 2015 American Institute of Chemical Engineers

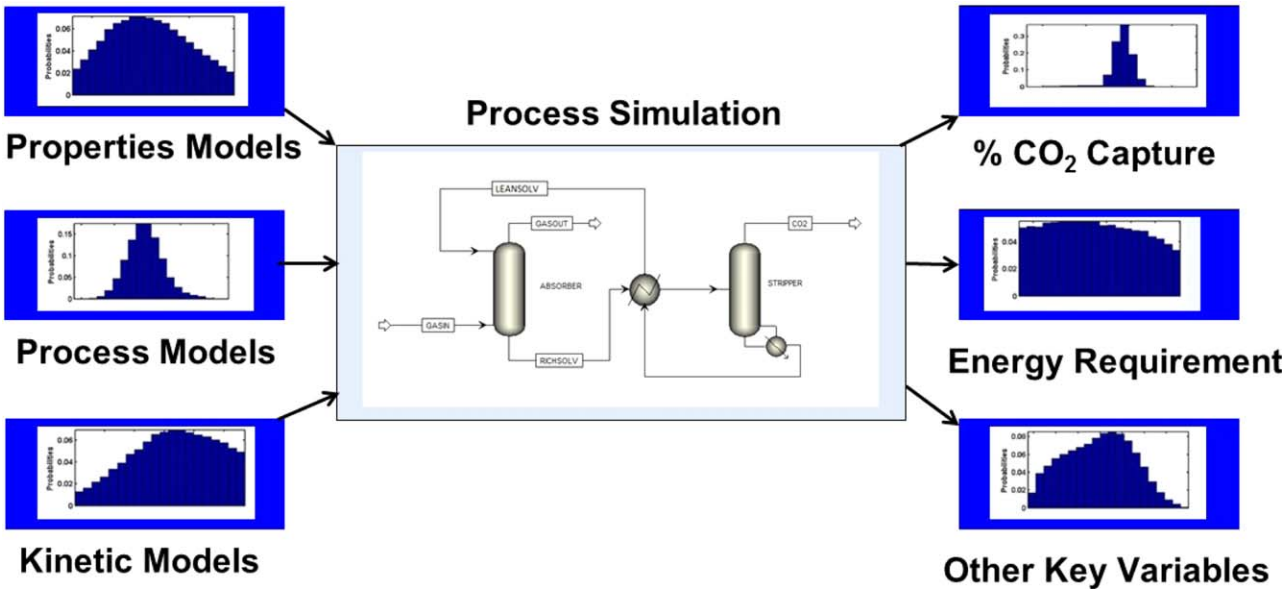

Figure 1. A possible approach to obtain uncertainty bounds in variables of interest from process simulations.

[Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

them with the observed data. Another major challenge to the incorporation of UQ in models is the large parameter space present in complex models, and the frequent lack of data to adequately characterize the model parameters.[6] However, many commercial process simulators provide options for Monte-Carlo simulation for use in uncertainty analysis.[7–9] A direct Monte-Carlo simulation-based approach has been recently applied for uncertainty analysis of coal gasification kinetics.[10] Excellent work on UQ of multiphase CFD models has been recently published.[11,12]

In the existing literature, a systematic approach to UQ of physical properties models is still scarce. In a recent work on phase equilibria, Mathias has developed an intuitive perturbation method to relate uncertainty in equilibrium constants to variability in process design that can easily be applied to process simulation.[8] A systematic approach to UQ of the thermodynamic models for solid-sorbents for CO_2 capture has been presented recently by Mebane et al.[13] In summary, the development of a generalized and systematic methodology for UQ of property models can be highly beneficial as one is not currently available.

Conversely, with growing concern over global CO_2 emissions, it is essential that efficient postcombustion CO_2 capture technologies are developed and deployed. The U.S. Department of Energy's Carbon Capture Simulation Initiative (CCSI) is focused on developing computational tools and models to accelerate the development and commercialization of CO_2 capture technologies.[14] The process modeling team of CCSI is developing high-fidelity process models which industry can use to more rapidly design and scale up such processes while also improving their operability. However, uncertainty in physical and chemical properties can significantly affect the predictability of these models. With this incentive, the UQ methodology is applied to the property models for monoethanolamine (MEA) systems for CO_2 capture. Complex interactions between ion-ion and ion-molecular species can lead to uncertainties in the property models for these electrolyte systems. This, in turn, will help to obtain uncertainty bounds on the variables of interest for such solvent-based postcombustion CO_2 capture systems. A precise knowledge of these uncertainty bounds is important for large-scale deployment of CO_2 capture technologies and for screening potential CO_2 capture technologies.

For application of the UQ methodology, we have focused here on alkanolamines, which are among the most widespread class of solvents for CO_2 capture because of their chemical structure. The hydroxyl group provides water solubility while the amino group provides the alkalinity necessary to capture CO_2.[15] Specifically, aqueous solutions of approximately 30% by weight MEA have been the industrial standard since 1970.[16] Other solvents, however, have been identified as better options for CO_2 capture than MEA in recent research. For example, aqueous piperazine has been found to have a greater rate of CO_2 absorption as well as greater CO_2 capacity.[17] For minimizing the energy requirements in solvent-based CO_2 capture processes, topological and parametric optimization plays a key role. Uncertainty bounds on process simulation results are needed to obtain accurate results from the optimization studies. Uncertainty bounds can be obtained by considering uncertainty in the property models, process models and the kinetic models and then propagating them through the entire process simulation as shown in Figure 1.

In this work, we apply the UQ methodology to the viscosity, density, and surface tension models for a MEA-H_2O-CO_2 ternary system. These properties are essential for designing and evaluating equipment,[15] particularly with respect to column sizing and pressure drop.[18] They are also used for calculating mass-transfer coefficients and interfacial mass-transfer area.[19,20] For these properties, a limited amount of research on deterministic models and experimental data is available for the MEA-H_2O-CO_2 ternary system. Aspen Plus® documentation[21] uses data from Weiland[22] to validate default models for the viscosity, density, and surface tension of CO_2-loaded aqueous MEA solutions at 298.15 K. Weiland et al.[19] presented empirical models for viscosity and density of CO_2-loaded solutions of various alkanolamines, including MEA for solutions at 298 K, 10–40 mass% MEA, and 0–0.5 CO_2 loading (mol CO_2/mol amine). Amundsen et al.[16] expanded on this work by broadening the temperature range from 298.15 to 353.15 K and then used the data to validate the correlations of Weiland

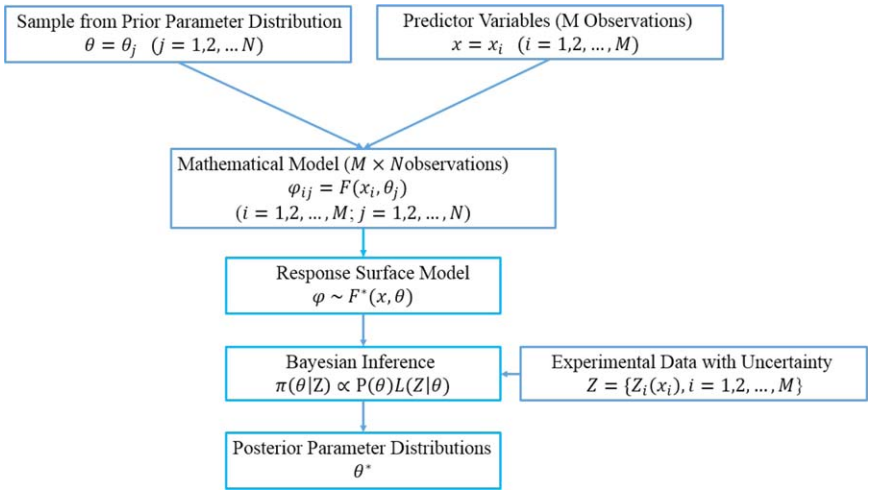

Figure 2. Overview of UQ in properties models.

[Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

et al. However, these models provide poor fits at lower temperatures and higher amine content. Moreover, the density model does not account for electrolyte speciation in the MEA-H_2O-CO_2 mixture, and its effect on the solution molecular weight was not considered. Han et al.[23] presented density data for the ternary system in the temperature range of 298.15–413.15 K for 30–100 mass% MEA. Jayarathna et al.[15] presented data for surface tension and density over a temperature range of 303.15–333.15 K and MEA mass% range of 20–70%, and they fit the surface tension data to a model developed by Connors and Wright.[24] Despite accurately representing the experimental data, the surface tension model presented by Jayarathna et al. is limited by its applicability to solutions of discrete amine compositions.

In this work, some of the available models and experimental data discussed before are used to characterize the viscosity, density, and surface tension of the ternary system of interest both deterministically and stochastically. For density and viscosity, the models developed by Weiland et al. are used as starting points to represent the MEA-H_2O-CO_2 system because these models are specific to CO_2-loaded aqueous amine solutions and satisfactorily fit the experimental data available, but modifications in the model forms are considered to improve the data fit. For the density model, electrolyte speciation in the ternary system is considered as its effect on solution molecular weight is needed to accurately characterize the relationship between the solution density and molar volume. Cross-validation of the density model with data from various sources[15,16,23] is performed. For the surface tension model, the model given by Jayarathna et al.[15] is modified to represent the surface tension as a continuous function of MEA composition while maintaining the accurate fit of the original model.

Overall Approach

For each model, baseline parameter values are first calibrated to fit the model to the experimental data. However, there can be uncertainty in the input variables (input uncertainty), physical property measurements (output uncertainty), functional form of the physical property models (model uncertainty), and the parameters used in the models (parametric uncertainty). Parametric uncertainties are difficult to characterize and quantify; however, the best guess for parametric uncertainty (priors) can be cast in the framework of a Bayesian inference methodology to obtain more informed posterior parameter uncertainty.

The Bayesian inference methodology is implemented as follows in this work. The model output (φ) is a function of a set of predictor variables (x) and a set of model parameters (θ) as shown below

$$\varphi = F(x, \theta) \tag{1}$$

The objective of the methodology is to compute a set of values of θ (denoted as θ^*) so that the output of the function $\varphi = F(x, \theta^*)$ matches with the experimental data (denoted as Z) within a given tolerance. After the predictor variables and model parameters for a physical property model have been identified, the next step is to create a response surface that maps the model inputs (x, θ) to the model output (φ). The reason for this step is that Bayesian inference is computationally expensive as it requires evaluation of the function $F(x, \theta)$ many times, and the use of response surface helps to reduce the overall computational cost by replacing the actual functional form with an emulator, given by

$$\varphi \sim F^*(x, \theta) \tag{2}$$

Response surfaces are built by sampling the process variable space with a sufficient number of sample points and fitting them with some curve fitting tools so that the resulting emulator estimates the true model outputs with sufficient accuracy. In this work, a popular curve fitting method called multivariate adaptive regression splines (MARS) developed by Friedman[25] is used to create the response surfaces, and cross validation is used to assess their quality. The cross validation procedure is necessary because Bayesian inference may lead to misleading results if performed on response surfaces that are not reflective of the true model. In this work, the specific methodology used to create the response surface is as follows. For a given physical property, a total of M experimental observations are available, each of which has a unique set of values of the predictor variables x_i and a value of the output variable $Z_i(x_i)$. For each individual parameter θ, a marginal prior distribution is estimated from the confidence intervals obtained from a least squares regression for the parameter of interest. For the joint prior distribution for the entire set of parameters, all parameters are

Table 1. Comparison of Original and Calibrated Parameter Values for the Solution Viscosity Model

Parameter	Weiland et al. Value	Value Calibrated from Data
a	0	−0.0838
b	0	2.8817
c	21.186	33.651
d	2373	1817
e	0.01015	0.00847
f	0.0093	0.0103
g	−2.2589	−2.3890

assumed to be independent. A sample of size N is obtained from the joint prior distribution. For each sample $j=1:N$, the set of parameter values is denoted as θ_j. The function representing the physical property is then evaluated for each set of x_i and θ_j, giving a total of $M\times N$ observations that are used to develop the response surface model.

After the response surfaces have been created and validated, Bayesian inference is performed, for which the objective is to compute

$$\pi(\theta|Z) \propto P(\theta)L(Z|\theta) \tag{3}$$

where $P(\theta)$ represents the prior distributions of the parameters, or the initial distributions assumed to result in a good match to the data, $\pi(\theta|Z)$ represents the posterior distributions of the parameters, or the revised distributions based on the results of comparing against the data, and $L(Z|\theta)$ is the likelihood function that measures the goodness of the match. A popular likelihood function, which is used in this work, is[26]

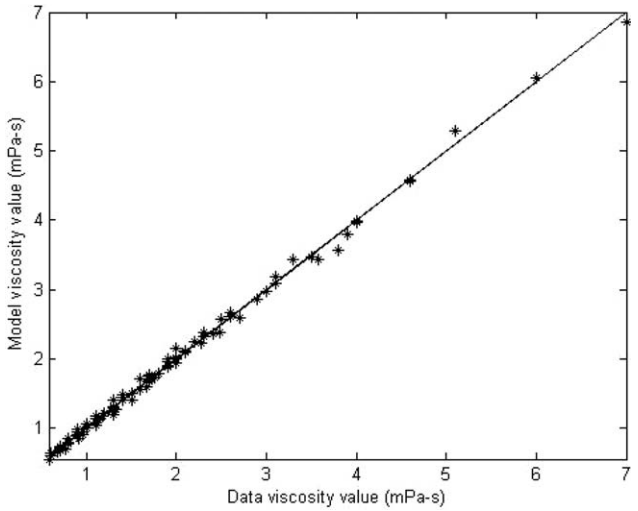

Figure 4. 10-fold cross validation of solution viscosity model.

Straight line represents perfect fit of model value to data value. Stars represent experimental data and model comparison.

$$L(Z|\theta)=\exp\left(-0.5\sum_{i=0}^{M}\frac{[F^*(x_i,\theta)-Z(x_i)]^2}{M\sigma_i^2}\right). \tag{4}$$

where σ_i is the standard deviation of the ith experimental data set. This likelihood function is based on the chi-square statistic, a popular method for data fitting. It can be

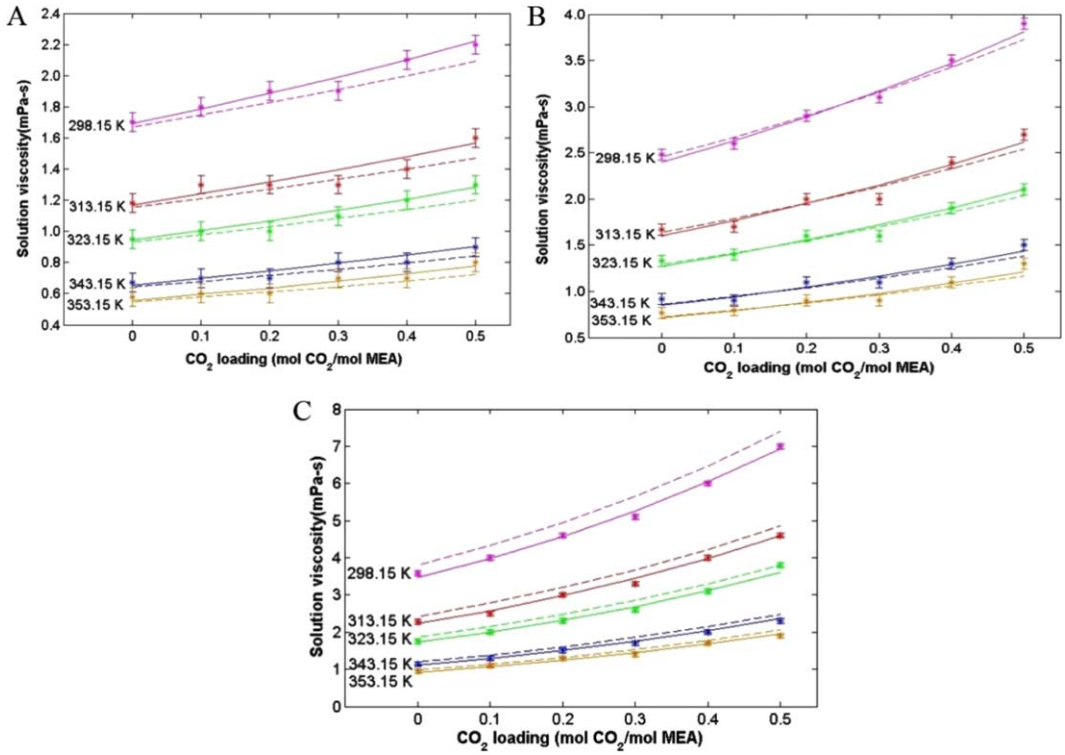

Figure 3. Comparison of models to experimental data for solutions of: (A) $W_{MEA}=20\%$; (B) $W_{MEA}=30\%$; (C) $W_{MEA}=40\%$.

Stars (*) represent experimental data, with error bars representing three standard deviations, from Amundsen et al., dashed lines represent model with parameters given by Weiland et al., and solid lines represent model with recalibrated parameters. Color-coding represents solutions of different temperatures (magenta = 298.15 K, red = 313.15 K, green = 323.15 K, blue = 343.15 K, and orange = 353.15 K). [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

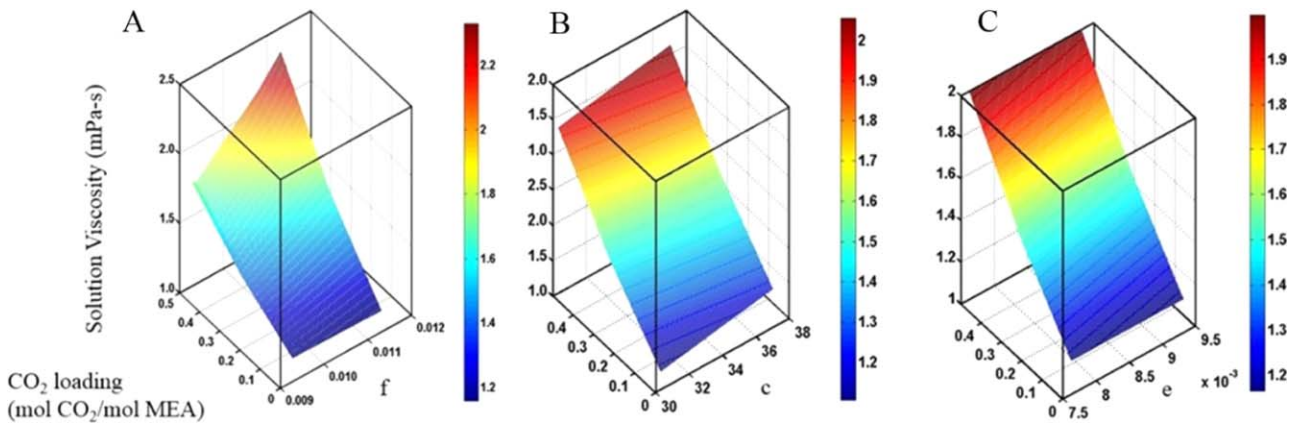

Figure 5. Response surfaces for viscosity model representing: (A) $\mu_{sln} = \mu_{sln}(f, \alpha)$; (B) $\mu_{sln} = \mu_{sln}(c, \alpha)$; (C) $\mu_{sln} = \mu_{sln}(e, \alpha)$.

[Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

considered as a weighted least-squares method that gives more weight to more accurate data. It also assumes that there is no correlation between experiments and that variation of data from their true value follows a normal distribution. The posterior distribution is computed by the Markov Chain Monte Carlo (MCMC) method for which the likelihood function is evaluated many times in traversing the parameter space of θ, and at each evaluation the response surface is also evaluated once.[27] In the MCMC algorithm used in this work, the Gibbs sampling method is used to perform the search and has been determined to be adequate. At the end of the inference, an instantiation of the posterior distributions of the parameters is obtained in the form of a collection of sample points. The methodology used in developing the response surfaces and performing Bayesian inference is illustrated in Figure 2.

Even after using surrogate response surface models, the Bayesian inference approach can still be computationally prohibitive if the parameter space is too large. In addition, there are parameters for which the uncertainty estimate is not improved with additional data, as their posterior distributions are approximately identical to the assumed prior distributions. The response surface corresponding to each parameter for each process variable can be visually examined to determine whether or not it can be excluded from UQ analysis. This approach can be very expensive for systems with a large number of parameters. For downselecting the parameter space, we utilize a sensitivity matrix approach where the candidate parameters for UQ analysis can be automatically selected by constructing a matrix of partial derivatives of the property model with respect to the parameter value. The sensitivity matrix S for a property is of dimension $p \times q$ where a single element is given by

$$S_{ij} = \max_{x_j \in [x_j^L, x_j^U]} \left| \left(\frac{\partial \varphi}{\partial \hat{\theta}_i} \right)_{x_{q \neq j} = x_q^*} \right| \tag{5}$$

where φ denotes the property being considered, p and q correspond to the number of parameters and variables, respectively, in the model and x_j refers to a specific variable and $\hat{\theta}_i$ is a parameter deviation term defined by

$$\theta_i = \bar{\theta}_i \hat{\theta}_i \tag{6}$$

where θ_i refers to a specific model parameter $\bar{\theta}_i$ and refers to the baseline value of the given parameter as obtained from deterministic regression. Also in Eq. 5, the terms x_j^L and x_j^U represent the lower and upper limits of the variable value, determined by the ranges of the data used to develop the model, and x_q^* represents the average of these two values for any given variable. The normalized version of this matrix (N) is defined for a single element as

$$N_{ij} = \frac{S_{ij}}{\max_{i \in [1,p], j \in [1,q]} S_{ij}} \tag{7}$$

For a given parameter θ_k, if $N_{kj} \ll 1 \ \forall j \in [1, q]$, then the parameter θ_k can be excluded from the UQ analysis as the prior belief of this parameter is unlikely to improve due to Bayesian inference. This methodology is validated for all the property models examined here. This procedure may be used to reduce the computational cost when dealing with more complicated models with larger numbers of parameters. This procedure is validated qualitatively by comparing the calculated derivatives with response surfaces generated using MARS.

Viscosity Model

Deterministic model

The model developed by Weiland et al. for the viscosity of the H_2O-MEA-CO_2 system (μ_{sln}) is of the form

$$\mu_{sln} = \mu_{H_2O} \exp\left(\frac{((aW_{MEA} + b)T + cW_{MEA} + d)(\alpha(eW_{MEA} + fT + g) + 1)W_{MEA}}{T^2} \right) \tag{8}$$

where W_{MEA} is the mass percent of MEA in solution on a CO_2-free basis, α is CO_2 loading (mol CO_2/mol MEA), T is temperature (K), and a-g are model parameters. The viscosity of pure water (μ_{H_2O}) in mPa s is given as a function of temperature[28]

$$\mu_{H_2O} = 1.002 * 10^{\frac{1.3272(293.15 - T - 0.001053(T - 293.15)^2)}{T - 168.15}} \tag{9}$$

Experimental data from Amundsen et al. with the variable ranges of $W_{MEA} \in [20\text{–}40]$, $\alpha \in [0\text{–}0.5]$, and $T \in [298.15\text{–}353.15$ K] are used to calibrate the parameters of Eq. 8. The values of the calibrated parameters are compared with the values given by Weiland et al. in Table 1.

Table 2. Calculated 95% Confidence Intervals for Parameters in Viscosity Model

Parameter	Baseline Value	Confidence Interval
a	−0.0838	[−0.2341, 0.0666]
b	2.8817	[−3.1179, 8.8814]
c	33.651	[−13.6178, 80.9212]
d	1817	[−66.8772, 3701.13]
f	0.0103	[0.0071, 0.0135]
g	−2.3890	[−3.4216, −1.3565]

The viscosity model with both sets of parameters is compared to the experimental data in Figure 3.

It is shown in Figure 3 that re-regressing the model parameters resulted in an improved fit at relatively low temperatures and high amine concentrations, particularly for $W_{MEA} = 40$ and $T = 298.15$ K. Although parameters a and b are assumed to be negligible in the works of Weiland et al. and Amundsen et al. for the MEA solvent, their inclusion in the model helps capture the effects of temperature and composition on the solution viscosity model more precisely. For higher temperatures (343.15–353.15 K) and lower MEA content, there is little difference between the fit given by the original parameters and the new parameters. Overall, the absolute average relative deviation (AARD) of the fit of the viscosity model to the data was decreased from 4.87 to 2.69% by adding the two additional parameters to the model and recalibrating. The model given in Eq. 8 is also tested with a 10-fold cross validation, and the results are shown in Figure 4.

As most of the model predictions lie around the 45° line, the model form appears to sufficiently represent the solution viscosity of the H_2O-MEA-CO_2 system. The calculated cross validation AARD is 2.78%.

Parameter screening

Before proceeding with UQ for the viscosity model, the sensitivity matrix method is used to determine the relative importance of the model parameters in an attempt to eliminate some of them from the Bayesian inference procedure. The derivatives given in Eq. 5 are evaluated for the viscosity model over the conditions

$$20 \le W_{MEA} \le 40$$
$$298.15 \le T \le 353.15$$
$$0 \le \alpha \le 0.5$$

where the limits on the model variables are decided based on the ranges of the data for which the model is developed. The normalized sensitivity matrix is calculated as

$$N = \max \begin{bmatrix} \left|\left(\frac{\partial \mu}{\partial \hat{a}}\right)_{W_{MEA},\alpha}\right| & \left|\left(\frac{\partial \mu}{\partial \hat{a}}\right)_{T,\alpha}\right| & \left|\left(\frac{\partial \mu}{\partial \hat{a}}\right)_{T,W_{MEA}}\right| \\ \text{----} & \text{---} & \text{---} \\ \text{----} & \text{---} & \text{---} \\ \left|\left(\frac{\partial \mu}{\partial \hat{g}}\right)_{W_{MEA},\alpha}\right| & \left|\left(\frac{\partial \mu}{\partial \hat{g}}\right)_{T,\alpha}\right| & \left|\left(\frac{\partial \mu}{\partial \hat{g}}\right)_{T,W_{MEA}}\right| \end{bmatrix}$$

$$= \begin{bmatrix} 0.3383 & 0.4601 & 0.2667 \\ 0.3877 & 0.3955 & 0.3057 \\ 0.4556 & 0.5673 & 0.3289 \\ 0.8200 & 0.7658 & 0.5919 \\ 0.0682 & 0.0811 & 0.0757 \\ 0.8244 & 0.8023 & 1.0000 \\ 0.6413 & 0.5714 & 0.7122 \end{bmatrix} \quad (10)$$

Equation 10 shows the normalized sensitivity of the viscosity model to its parameters, evaluated over the range of interest for each variable. The model is clearly most sensitive to parameter f, as the corresponding row of the matrix has the highest value for all three columns, and least sensitive to parameter e.

To confirm the validity of these results, a response surface is generated using MARS with a sample size of 4000 with variable and parameter inputs sampled from a Monte Carlo

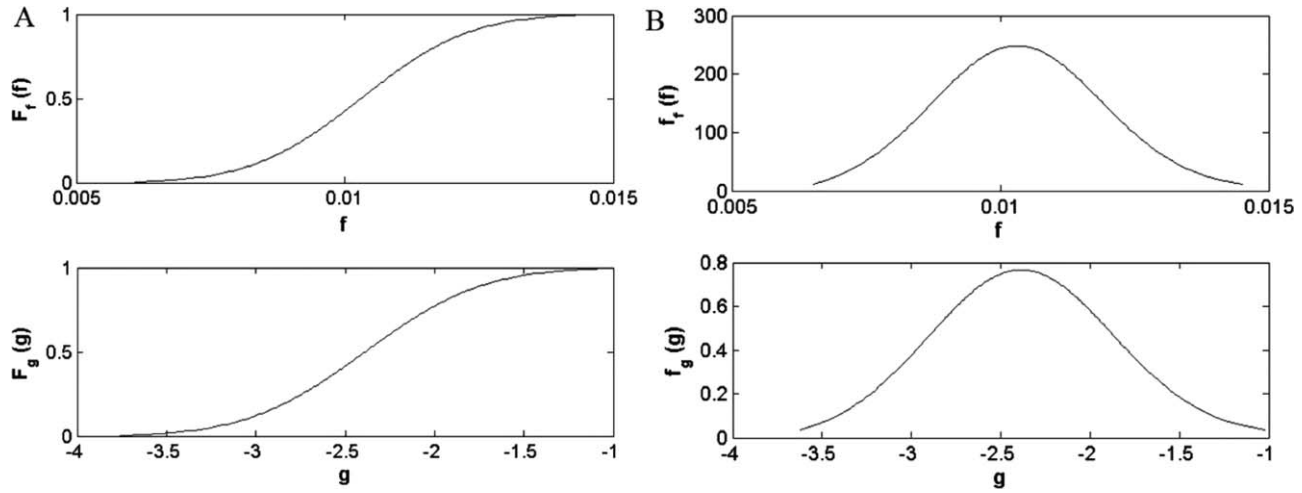

Figure 6. Approximations of (A) cumulative distribution functions and (B) probability distribution functions for viscosity model parameters determined from confidence intervals in the deterministic regression results.

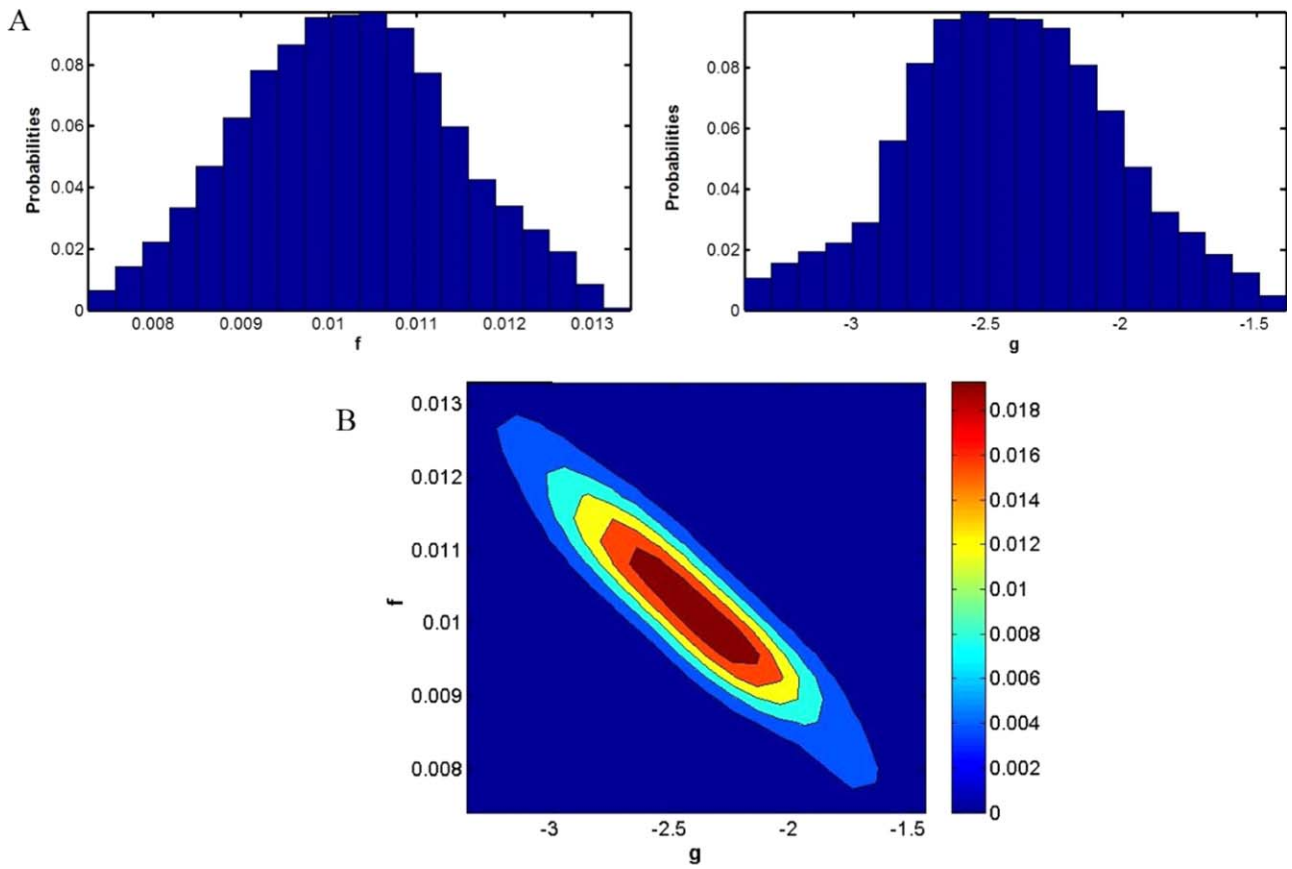

Figure 7. Posterior distributions for viscosity model parameters obtained from Bayesian inference.

(A) Histograms representing marginal distributions of the parameters; (B) contour plot representing joint distribution. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

simulation with uniform distributions (ranges from ±10% of the baseline values for parameters and $W_{MEA} \in [20-40]$, $T \in [298.15-333.15\ \text{K}]$, $\alpha \in [0-0.5]$, and viscosity output calculated from Eq. 8. Figure 5 shows three-dimensional projections from this response surface representing viscosity as a function of CO_2 loading and one of the model parameters. All plots are generated at $T=325.65$ K and $W_{MEA}=30$ (the average values over the variable ranges of interest) and at baseline parameter values given in Table 1, except for the parameter of interest for a given plot.

The relatively high sensitivity of the viscosity model with respect to parameter *f* is confirmed by the shape of the curve in Figure 5A, which depicts large deviation in the viscosity prediction, particularly at high loading values, as a result of a small perturbation in the value of the parameter. There is clearly little sensitivity of the viscosity model to parameter *e* (Figure 5C), as the curve representing viscosity as a function of loading changes negligibly with the parameter value. It can be noted that $N_{ej} \ll 1\ \forall j \in [1,3]$. The parameter *c* is considered as an intermediate case (noting $N_{e\alpha} < N_{c\alpha} < N_{f\alpha}$ as given by Eq. 6, and the response surface of viscosity to this parameter (Figure 5B) is shown to be more subtly peaked than that of parameter *f*. The study shows that the sensitivity matrix approach can be effectively used to avoid unnecessary UQ analysis for parameters without examining the response surfaces visually. For this work, parameters with sensitivity matrix values less than 0.1 for all variables will be omitted from the stochastic model. Based on the specific results obtained for the viscosity model, parameter *e* is not considered further for UQ.

Stochastic model

The prior distributions of the model parameters that are used for Bayesian inference are determined by estimating confidence intervals for the parameters based on the results of the deterministic regression. The 95% confidence intervals for the parameters, noting that *e* has been eliminated as a result of the sensitivity matrix calculation, are given in Table 2.

For this work, parameters which have the value 0 contained in their respective confidence intervals are not considered for UQ. Therefore, the stochastic model for viscosity only considers uncertainty in parameters *f* and *g*. The prior distributions of these parameters are determined by estimating their probability density functions (PDF) from a series of confidence intervals. A confidence interval of significance level γ for the parameter θ may be expressed as

$$\mathrm{P}\left(\theta_\gamma^{\mathrm{L}} \le \theta \le \theta_\gamma^{\mathrm{U}}\right) = 1-\gamma \tag{11}$$

where $\theta_\gamma^{\mathrm{L}}$ and $\theta_\gamma^{\mathrm{U}}$ are the lower and upper limits of the parameter for a confidence interval of $100(1-\gamma)\%$. From the confidence interval, the cumulative distribution function (CDF) of the parameter may be determined by

$$F_\theta\left(\theta_\gamma^{\mathrm{L}}\right) = \frac{\gamma}{2} \tag{12a}$$

$$F_\theta\left(\theta_\gamma^{\mathrm{U}}\right) = 1-\frac{\gamma}{2} \tag{12b}$$

An estimate of the entire CDF of each parameter is obtained by calculating all confidence intervals for levels

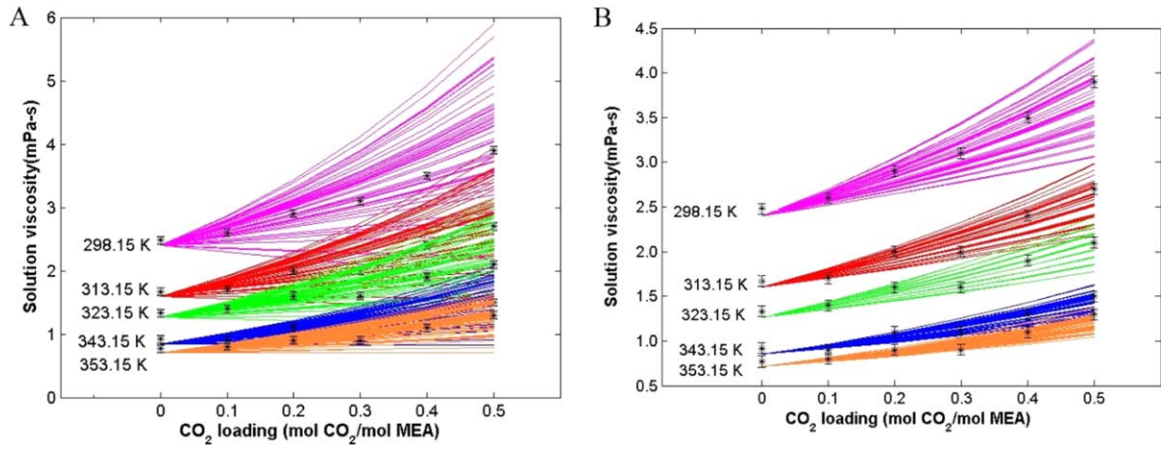

Figure 8. Stochastic viscosity model for $W_{MEA} = 30\%$ considering (A) prior parameter distribution and (B) posterior parameter distribution.

Experimental data represented by (*) and taken from Amundsen et al. Error bars represent three standard deviations. Color-coding represents solutions of different temperatures (magenta = 298.15 K, red = 313.15 K, green = 323.15 K, blue = 343.15 K, and orange = 353.15 K). [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

$\gamma = 0.01 - 0.99$ at increments of 0.01. The PDF of each parameter is determined from a numerical approximation of the derivative

$$f_\theta(\theta) = \frac{dF_\theta(\theta)}{d\theta} \tag{13}$$

The estimated CDFs and PDFs of the two parameters f and g are given in Figure 6.

The marginal prior distributions of the model parameters are taken as the PDFs given in Figure 6, and the joint prior distribution is determined by assuming independence of the parameters. For UQ of the viscosity model, a sample of 100 parameter sets is obtained from the PDFs. These sets are combined with each of 90 data points that correspond to different experimental conditions (combination of T, W_{MEA}, α). Therefore, a total of 9000 data points are used to develop the response surface used for Bayesian inference. This response surface is validated by a 10-fold cross validation procedure, which results in a coefficient of determination $R^2 = 0.9403$. The posterior distributions are obtained from Bayesian inference as a set of sample points. Figure 7 shows the posterior distributions in terms of marginal PDFs in histogram form and a contour plot that represents the joint probability distribution of the two parameters.

By comparing the marginal posterior distributions in Figure 7A with the marginal prior distributions in Figure 6A, the reduction in parameter space as a result of Bayesian inference is not apparent. The contour plot in Figure 7B representing joint distribution $f_{f,g}(f,g)$, however, does clearly demonstrate that the prior distribution assumption that the parameters are independent random variables does not hold in the posterior distribution. Accordingly, the feasible parameter space in the joint distribution gets reduced while performing Bayesian inference.

Figure 8 shows the stochastic model for 30% MEA as a function of CO_2 loading and temperature separately with prior distributions and posterior distributions. All comparisons are made by drawing a sample size of 50 from the parameter distributions of interest. All parameters not included in UQ are treated as the constant values given in Table 1.

In Figure 8, it is shown that the viscosity ranges predicted using the prior distributions are excessively large, and the ranges are reduced considerably as a result of Bayesian inference. Furthermore, it is clear that the two parameter stochastic model is adequate for characterizing the uncertainty of the model.

Density Model

Deterministic model: Existing model

The model for the density of the H_2O-MEA-CO_2 system (ρ_{sln}) given by Weiland et al. is shown below

$$\rho_{sln} = \frac{MW_{sln}}{V_{sln}} \tag{14}$$

$$V_{sln} = X_{MEA}V_{MEA} + X_{H_2O}V_{H_2O} + X_{CO_2}V_{CO_2} + X_{MEA}X_{H_2O}V^* + X_{MEA}X_{CO_2}V^{**} \tag{15}$$

$$V_{MEA} = \frac{MW_{MEA}}{aT^2 + bT + c} \tag{16}$$

$$V^{**} = d + eX_{MEA} \tag{17}$$

where MW_{sln} is the average of the molecular weights of the individual components. The solution molar volume (V_{sln}), as calculated by Eq. 15, consists of partial molar volume terms for the three components as well as terms that account for the interaction between species. The terms X_i and V_i, where i represents one of the three species, are the mole fraction and molar volume, respectively, of the individual species.

Table 3. Parameter Values for Weiland et al. Solution Density Model

Parameter	Value
a	-5.35162×10^{-7}
b	-4.51417×10^{-4}
c	1.19451
d	0
e	0
V_{CO_2} (mL/mol)	0.04747
V* (mL/mol)	−1.8218

Table 4. Summary of Density Data Used for Model Calibration

Data Source	Temperature (K)	CO_2 Loading	r	Number of Observations
Amundsen et al.	298.15–353.15	0–0.5	0.2–0.4	83
Jayarathna et al.	303.15–333.15	0–0.5	0.2–0.4	72
Han et al.	298.15–353.15	0.1–0.56	0.2–0.4	54

V_{MEA} and V_{H_2O} are pure component molar volume values, while V_{CO_2} represents the dissolved molar volume of CO_2, which is a constant value and unrelated to the pure component value. The molar volume of MEA is calculated by Eq. 16 as a function of temperature. The molar volume associated with the interaction between H_2O and MEA is given by V^*, a constant, and the molar volume associated with the interaction between MEA and CO_2 is given by V^{**}, which is dependent on the solution MEA concentration. The values of all constants used in this model are given in Table 3.

Deterministic model: A new model

The original model does not explicitly account for the presence of the ionic species in the H_2O-MEA-CO_2 solution, which can be represented by the reversible reactions[29,30]

$$2MEA + CO_2 \leftrightarrow MEA^+ + MEACOO^- \quad (18)$$

$$MEA + CO_2 + H_2O \leftrightarrow MEA^+ + HCO_3^- \quad (19)$$

This reaction set represents a simplification of the solution chemistry that was developed by Hilliard[30] for typical CO_2 capture process conditions where concentrations of other ions (H_3O^+, OH^-, CO_3^{2-}) are negligibly low. The kinetics for these reactions are given in the work of Plaza.[29]

It is proposed that the model for the density of the solution should take the ionic speciation into account to more accurately represent the solution's molecular weight, as opposed to just calculating the molecular weight based on the apparent mole fractions of the molecular species. For the new model, the molecular weight is calculated as

$$MW_{sln} = \sum_{i=1}^{6} MW_i \bar{X}_i \quad (20)$$

where MW_i represents the molecular weight of a specific component i (from Eqs. 18 and 19) and X_i^J is the true species mole fraction. The true species mole fractions can be determined from experimental data given in terms of solution temperature, CO_2 loading (α), and MEA weight fraction on a CO_2-free basis (r). From this data, the apparent species mole fractions (X_i) can be calculated by

$$X_{MEA} = \left(1 + \alpha + \left(\frac{MW_{MEA}}{MW_{H_2O}}\right)\left(\frac{1-r}{r}\right)\right)^{-1} \quad (21)$$

$$X_{CO_2} = \alpha X_{MEA} \quad (22)$$

$$X_{H_2O} = 1 - X_{MEA} - X_{CO_2} \quad (23)$$

The new model proposed for the solution density is given by

$$\rho_{sln} = \frac{MW_{sln}}{X_{MEA}V_{MEA} + X_{H_2O}V_{H_2O} + aX_{CO_2} + (b + cX_{MEA})X_{MEA}X_{H_2O} + (d + eX_{MEA})X_{MEA}X_{CO_2}} \quad (24)$$

The major differences between this model and the original model (Eqs. 14–17 and 24) are that both interaction terms (V^* and V^{**}) are now assumed to be linear functions of the apparent mole fraction of MEA and that the V^{**} is not assumed to be equal to zero. Only five parameters (a–e) are considered for the deterministic model, because they are related to the nonideality of the electrolytic solution mixture. As the method of calculation of the solution molecular weight was changed, the parameters require recalibration. It is assumed that the molar volumes of H_2O and MEA can be calculated with relatively high precision and certainty. The molar volume of MEA is calculated from Eq. 16, with coefficients given in Table 3. The molar volume of H_2O is calculated similarly, from the equation

$$V_{H_2O} = \frac{MW_{H_2O}}{-(3.2484 \times 10^{-6})T^2 + 0.00165T + 0.793} \quad (25)$$

where the coefficients were regressed from data given in Liley et al.[31]

Three sources of data for the H_2O-MEA-CO_2 solution density are identified. The ranges of essential solution properties and number of data points that are considered for each data source are summarized in Table 4.

The model parameters are calibrated using all data observations from the three sources, as listed in Table 4, with equal weighting. The calibrated parameters are shown in Table 5.

Table 6 shows the values of the AARD for each model with respect to the data from the individual sources and overall. The original and new density models are compared to the experimental data in Figure 9. Model predictions are

Table 5. Calibrated Parameter Values from Density Model Given in Eq. 24

Parameter	Value
a	10.2074
b	−2.2642
c	3.0059
d	207
e	−563.3701

Table 6. Values of AARD for Original and New Density Models

Data Source	Original Model A ARD (%)	New Model AARD (%)
Amundsen et al.	0.24	0.27
Jayarathna et al.	0.37	0.09
Han et al.	0.71	0.28
All Data	0.41	0.20

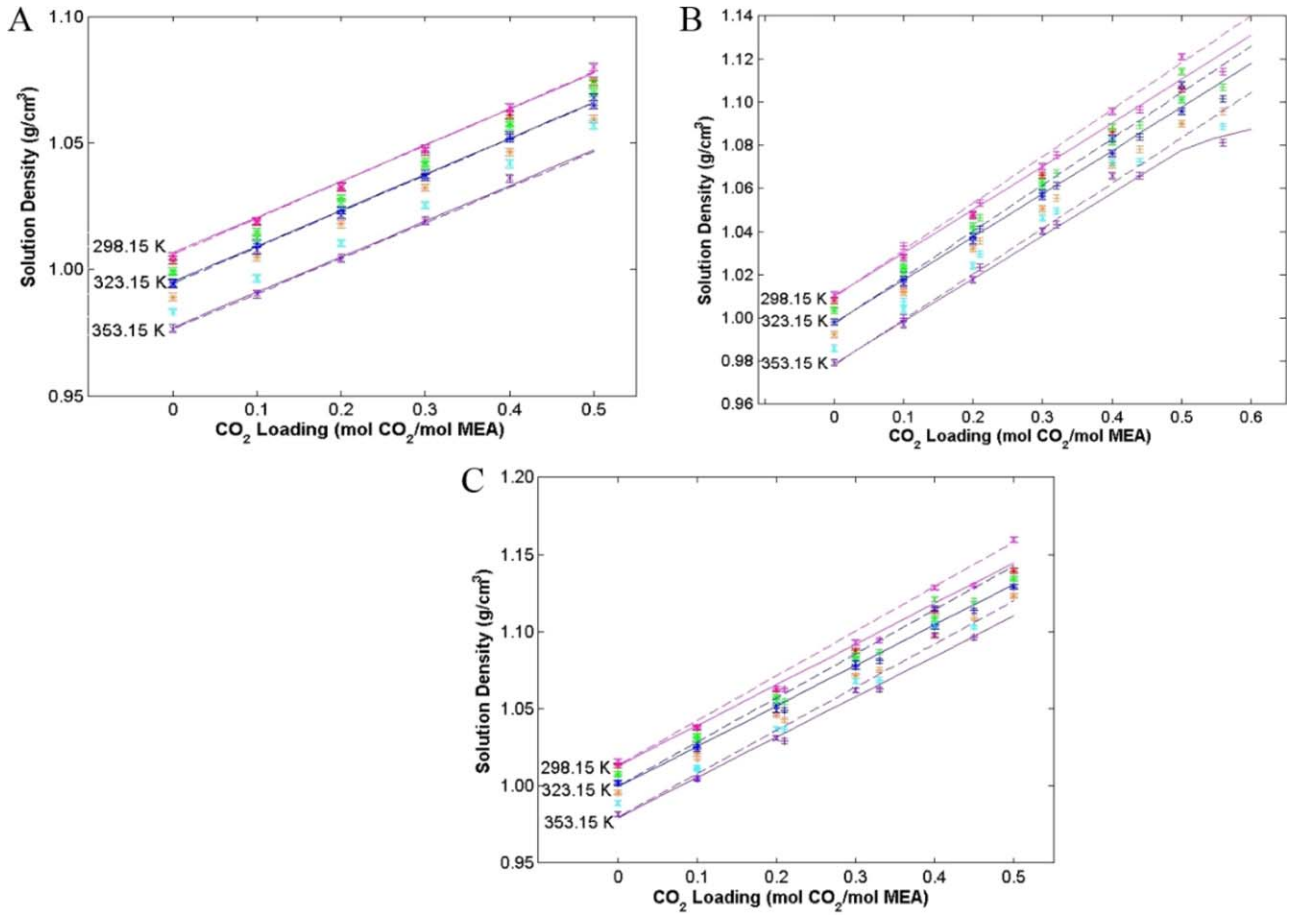

Figure 9. Comparison of models to experimental data for solutions of: (A) $r = 0.2$; (B) $r = 0.3$; (C) $r = 0.4$.

Stars (*) represent experimental data from Jayarathna et al., *x*'s (*x*) represent experimental data from Amundsen et al., pluses (+) represent experimental data from Han et al., and all data are given to three standard deviations. Dashed lines represent density as calculated from the original model and solid lines represent density as calculated from the new model. Color-coding represents solutions of different temperatures (magenta = 298.15 K, red = 303.15 K, green = 313.15 K, blue = 323.15 K, orange = 333.15 K, cyan = 343.15 K, and purple = 353.15 K). [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

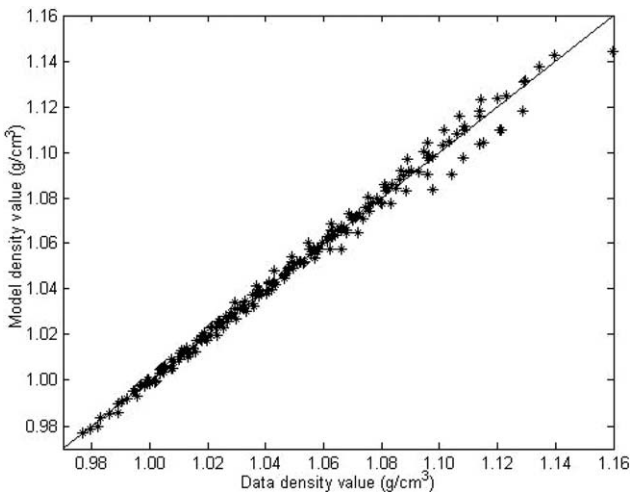

Figure 10. 11-Fold cross validation of solution density model.

Straight line represents perfect fit of model value to data value. Stars represent experimental data and model comparison.

shown only for temperatures of 298.15, 323.15, and 353.15 K to enhance the clarity of the graphs.

For solutions of $r = 0.2$, the density data available for both sources are in agreement with each other, and both the model by Weiland et al. and the new model both provide satisfactory fit to the data; however, for solutions of $r = 0.3$ and $r = 0.4$, there is significant discrepancy between the data given by different sources. As an example, Figure 9B shows that the density of solutions with $r = 0.3$ appears to decrease with increased CO_2 loading at relatively low temperatures (298.15–313.15 K) for $\alpha > 0.5$, despite the otherwise evident trend that solution density increases with loading. This is more likely attributed to measurement inconsistency from different experiments rather than an actual physical phenomenon. In the same graph, a pronounced decrease in the slope

Table 7. Calculated 95% Confidence Intervals for Parameters in Density Model

Parameter	Baseline Value	Confidence Interval
a	10.2074	[4.6508, 15.764]
b	−2.2642	[−3.0579, −1.4704]
d	207	[115.1532, 298.8448]
e	−563.3701	[−941.8307, −184.9095]

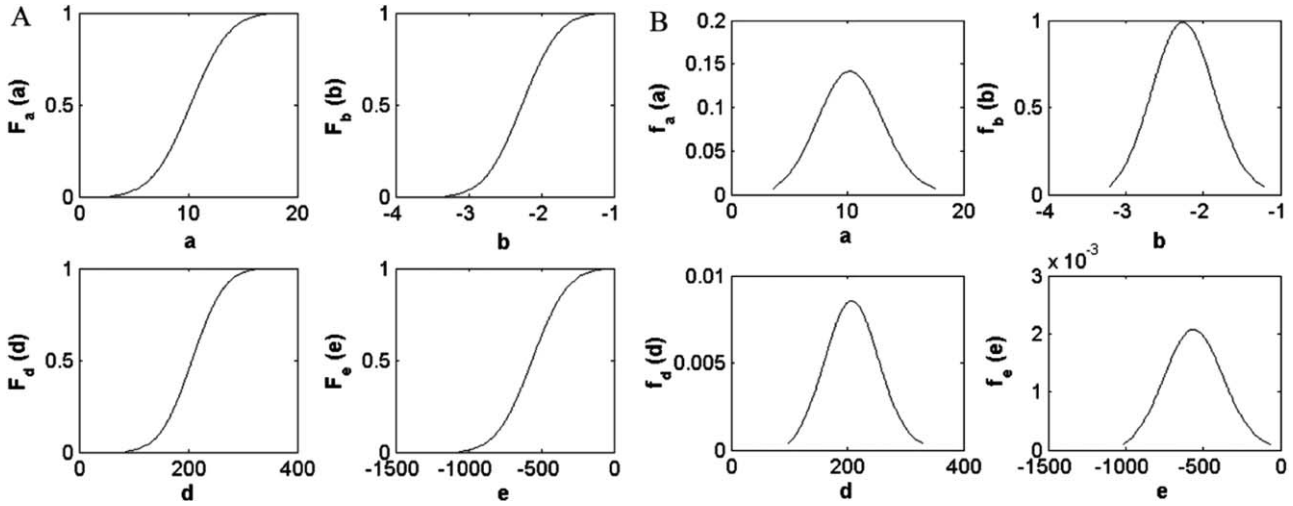

Figure 11. Approximations of (A) cumulative distribution functions and (B) probability distribution functions for density model parameters determined from the confidence intervals in the deterministic regression results.

of the new model for $T = 353.15$ K and $r = 0.3$ occurs at $\alpha = 0.4$ in order to more correctly match the only data point given above this value, at $\alpha = 0.56$. This trend is not observed for curves for lower temperatures ($T \leq 333.15$ K) because data are also given at $\alpha = 0.5$ for which the fit must be optimized. An overall trend is that the data from Amundsen et al. most accurately fits the original model in comparison to data from the other sources probably because these data were originally presented as validation data for the model by Weiland et al. As the overall effect of the modification of this model and the recalibration of the parameters is to minimize the discrepancy between the model and the aggregate data, the AARD of the model compared with the Amundsen et al. data increases, as shown in Table 6. This demonstrates the inherent inability of deterministic models to represent a system for which precise values of a physical property are unknown due to inconsistent data, and implies the need for a stochastic model to effectively incorporate all of these data into the model prediction. The new model is tested using an 11-fold cross validation procedure, and the result is shown in Figure 10.

The cross validation plot shows that the model can accurately predict solution density at relatively low CO_2 loading values; however, the model appears to be more uncertain at higher density. This can be partly attributed to the contradictory data for solutions of higher CO_2 loading from the different data sources. The calculated cross validation AARD for the density model is 0.22%.

Parameter screening

The sensitivity matrix is calculated for the density model with the same procedure used for the viscosity model. The parametric sensitivity, however, is calculated with respect to the molar volume model (the denominator of Eq. 24) because these derivatives can be calculated analytically. As the density model is dependent on the solution molecular weight, which is calculated based on the electrolytic speciation of the system, its derivatives cannot be obtained analytically. As the molecular weight calculation is completely independent of the parameter values, the relative importance of its parameters can be determined by examining the sensitivity of either the density or molar volume models to the parameters. The calculated sensitivity matrix is given by

$$N = \max \begin{bmatrix} \left|\left(\frac{\partial V}{\partial \hat{a}}\right)_{r,\alpha}\right| & \left|\left(\frac{\partial V}{\partial \hat{a}}\right)_{T,\alpha}\right| & \left|\left(\frac{\partial V}{\partial \hat{a}}\right)_{T,r}\right| \\ --- & --- & --- \\ --- & --- & --- \\ \left|\left(\frac{\partial V}{\partial \hat{e}}\right)_{r,\alpha}\right| & \left|\left(\frac{\partial V}{\partial \hat{e}}\right)_{T,\alpha}\right| & \left|\left(\frac{\partial V}{\partial \hat{e}}\right)_{T,r}\right| \end{bmatrix} = \begin{bmatrix} 0.2174 & 0.3139 & 0.4220 \\ 0.1482 & 0.1987 & 0.1577 \\ 0.0214 & 0.0415 & 0.0235 \\ 0.4797 & 1.0000 & 0.9035 \\ 0.142 & 0.4276 & 0.2596 \end{bmatrix} \quad (26)$$

Similar to the viscosity model, these derivatives are evaluated over the variable ranges used to develop the deterministic models. Based on the relative values of the matrix elements, the parameter to which the model is most sensitive is *d*. Parameter *c* is excluded from Bayesian inference because all elements in the respective row are sufficiently close to zero.

Stochastic model

As with the viscosity model, the parameters to be included in the stochastic density model are determined by calculating the 95% confidence intervals, which are shown in Table 7.

Since none of the four parameters contain 0 in their confidence intervals, all are considered for the stochastic model. The marginal prior distributions are determined in the same manner as for the viscosity model, and the CDFs and PDFs are given in Figure 11.

The marginal prior distributions of the model parameters are taken as the PDFs given in Figure 11, and the joint prior distribution is determined by assuming independence of the

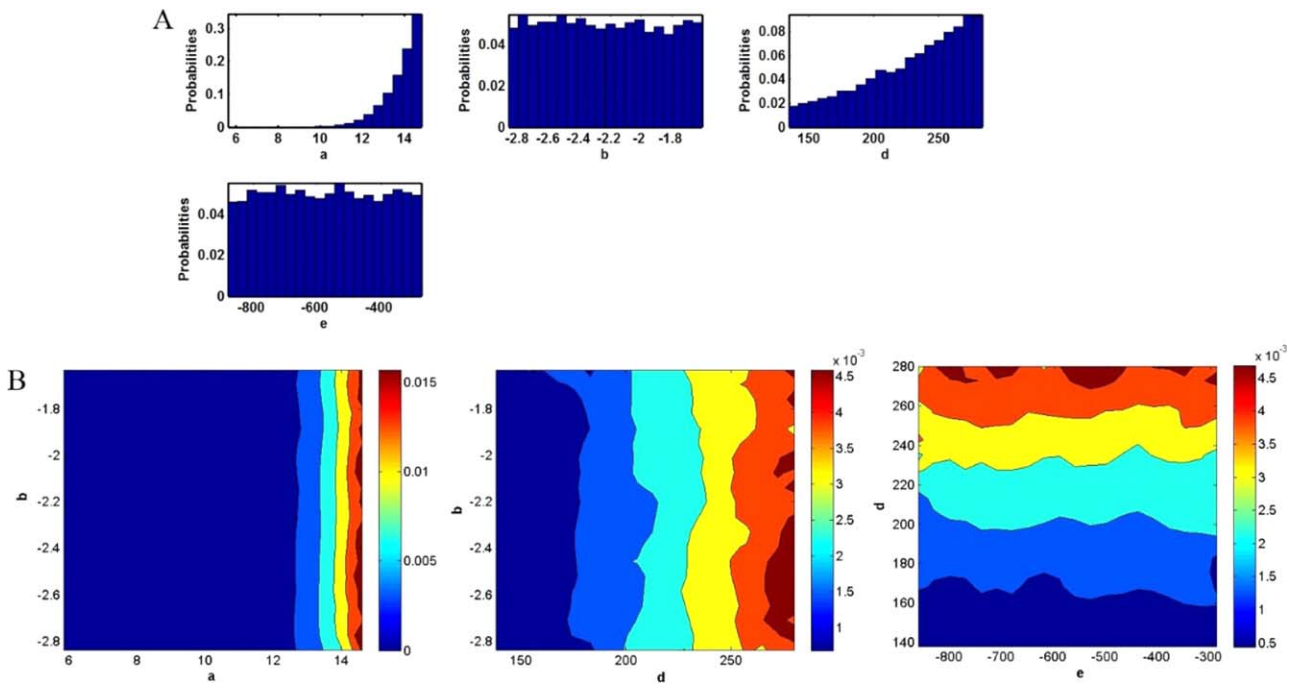

Figure 12. Posterior distributions for density model parameters obtained from Bayesian inference.

(A) Histograms representing single parameter marginal distributions of the parameters; (B) contour plots representing two parameter marginal distributions. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

parameters. A sample of 100 parameter sets is obtained from a numerical approximation of the PDFs. These sets are combined with each of 209 data points that correspond to different experimental conditions, and a total of 20,900 data points are used to develop the response surface model. This model is validated by a 10-fold cross validation procedure with $R^2=0.9999$. Figure 12 shows a representation of the posterior distribution in terms of histograms for all single parameter marginal distributions and contour plots for selected two parameter marginal distributions.

As the parameter space of the density model has four dimensions, it cannot be visualized directly as with the viscosity model. The histograms in Figure 12A represent the marginal distributions of single parameters, and it should be noted that the major change that occurs as a result of Bayesian inference is that the probability mass density of parameters a and d is shifted toward the upper boundary of the 95% confidence interval.

Table 8. MEA Composition Dependent Parameters for Surface Tension Model[15]

	$r=0.2$	$r=0.3$	$r=0.4$
CO_2 ($i=1$)			
a_i	0.3073	0.09409	0.1478
b_i	−0.8574	−0.7392	−0.8982
MEA ($i=3$)			
a_i	1.067	1.114	1.157
b_i	0.1701	0.1757	0.3062

Table 9. MEA Composition Dependent Parameters for σ_{CO_2}[15]

	$r=0.2$	$r=0.3$	$r=0.4$
$S1$	0.08286	0.1605	0.1184
$S2$	4.309×10^{-4}	1.316×10^{-4}	1.954×10^{-4}

Surface Tension Model

Deterministic model: Existing model

The model for surface tension of the H_2O-MEA-CO_2 system (σ_{sln}) developed by Jayarathna et al. is given by

$$\sigma_{sln}=\sigma_2+\sum_{i=1,3}\left(1+\frac{b_iX_i}{(1-a_i)\left(1+\sum_{j=1,3}\frac{a_j}{(1-a_j)}X_j\right)}\right)(X_i(\sigma_i-\sigma_2)) \tag{27}$$

where subscripts 1–3 represent CO_2, H_2O, and MEA, respectively. The surface tension of component i is denoted by σ_i. The apparent mole fraction of species i in the solution is denoted by X_i. The parameters a_i and b_i are dependent on the mass fraction of MEA in the solution. The values of the parameters reported by Jayarathna et al. are shown in Table 8, where r is defined as the mass fraction of MEA in solution on a CO_2-free basis.

Because carbon dioxide does not exist as a liquid in the temperature range considered for this analysis, the surface tension associated with this component is only considered as a fitting parameter in the model.[15] This parameter is given as[15]

Table 10. Parameters for Pure Component Surface Tension Model[32]

	c_1	c_2	c_3	c_4	T_c (K)
H_2O	0.18548	2.717	−3.554	2.047	647.13
MEA	0.09945	1.067	0	0	614.45

Table 11. Parameters for Updated σ_{CO_2} Model

$S1$	−5.987
$S2$	3.7699
$S3$	−0.43164
$S4$	0.018155
$S5$	−0.01207
$S6$	0.002119

Table 12. Regressed Parameter Values for New σ_{sin} Model

Parameter	*Value*	*Parameter*	*Value*
a	2.4558	f	2.3122
b	−1.5311	g	4.5608
c	3.4994	h	−2.3924
d	−5.6398	i	5.3324
e	10.2109	j	−12.0494

$$\sigma_{CO_2} = S1 + S2 * T(\mathrm{K}) \tag{28}$$

where $S1$ and $S2$ are parameters that depend on the MEA content of the mixture, which are shown in Table 9.

The pure component surface tension values for H_2O and MEA were calculated using the model proposed by Asprion[32] with the equation

$$\sigma_i = c_1 \left(1 - \frac{T}{T_c}\right)^{c_2 + c_3(T/T_c) + c_4(T/T_c)^2} \tag{29}$$

where σ_i is the pure component surface tension in N/m, T_c is the pure component critical temperature, and coefficients c_{1-4} are species specific. The values of these model parameters are presented in Table 10.

Deterministic model: A new model

Although this model was shown to accurately represent the available data, it has a major shortcoming in that it is only applicable to solutions with discrete values of MEA composition ($r = 0.2$–0.7 in increments of 0.1) in its original form. Therefore, the model is updated to allow for representation of solution surface tension for a continuous range of compositions. The range $r \in [0.2,0.4]$ is considered for this work to allow for consistency with the models previously developed for viscosity and density. For the updated surface tension model, Eq. 28 is replaced by

$$\sigma_{CO_2} = S1 * r^2 + S2 * r + S3 + T(S4 * r^2 + S5 * r + S6) \tag{30}$$

so that this fitting parameter may be represented in the model as a continuous function of r. The coefficient values are regressed to give the best fit between the values of σ_{CO_2} calculated by Eqs. 28 and 30. In dealing with the MEA

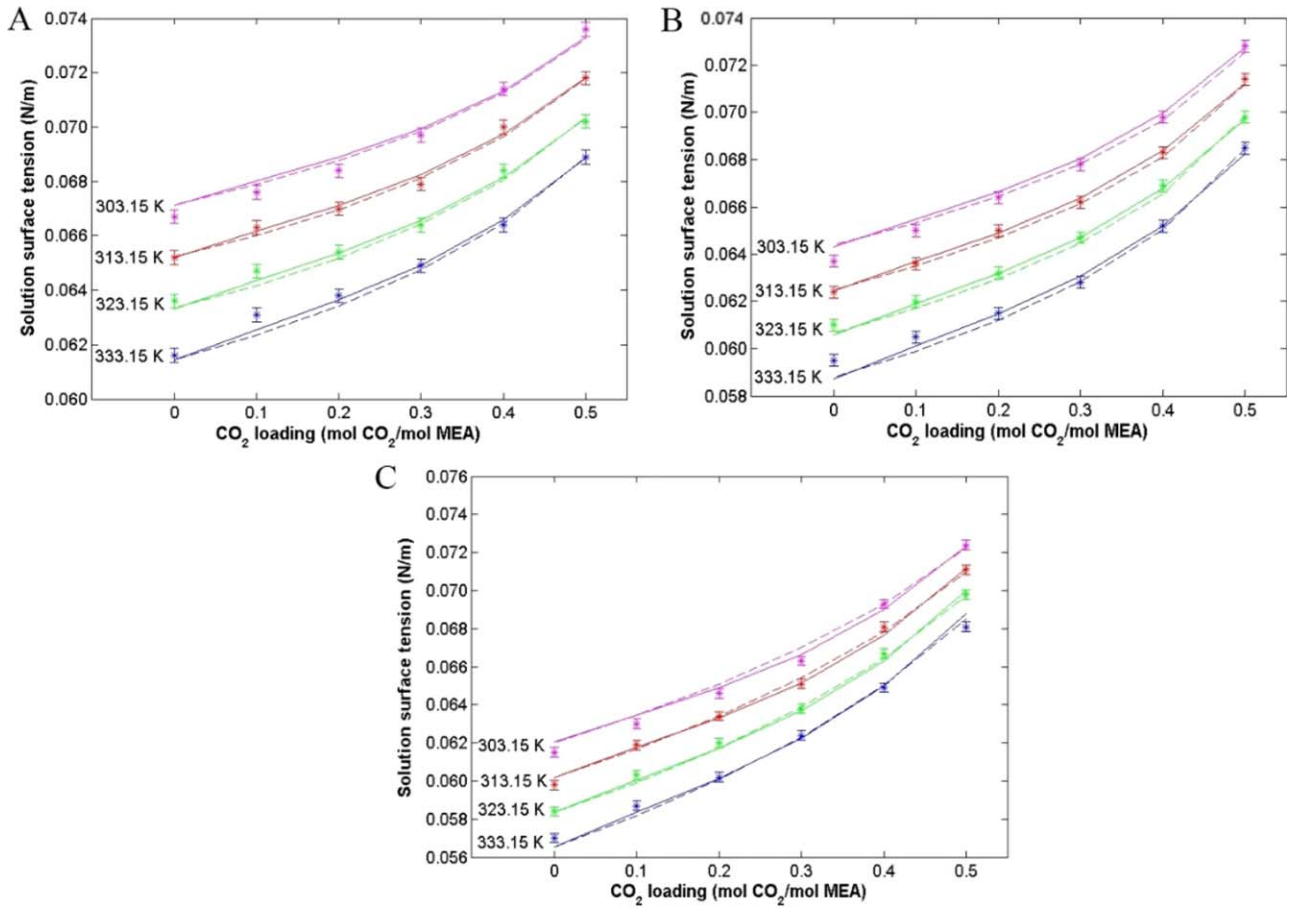

Figure 13. Comparison of models to experimental data for solutions of: (A) $r = 0.2$; (B) $r = 0.3$; (C) $r = 0.4$.

Stars (*) represent experimental data, with error bars representing one standard deviation, from Jayarathna et al., dashed lines represent original model, and solid lines represent new model. Color-coding represents solutions of different temperatures (magenta = 303.15 K, red = 313.15 K, green = 323.15 K, and blue = 333.15 K). [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

concentration-dependent parameters in Eq. 27, a new model form is proposed that is functionally similar to Eq. 27

$$\sigma_{sln}=\sigma_{H_2O}+(\sigma_{CO_2}-\sigma_{H_2O})f(r,\alpha)X_{CO_2}+(\sigma_{MEA}-\sigma_{H_2O})g(r,\alpha)X_{MEA} \tag{31}$$

$$f(r,\alpha)=a+b\alpha+c\alpha^2+dr+er^2 \tag{32}$$

$$g(r,\alpha)=f+g\alpha+h\alpha^2+ir+jr^2 \tag{33}$$

The variables α and r can be related to apparent species mole fractions by Eqs. 21–23. Model parameters (a–j) are regressed with the experimental data from Jayarathna et al. The regressed values of the parameters of Eqs. 30 and 32–33 are given in Tables 11 and 12, respectively.

The fit of the experimental data to the new and original surface tension models is compared in Figure 13.

The fit of the new model to the data is similar to the model by Jayarathna et al., but, as mentioned earlier, the model by Jayarathna et al. is valid only for discrete values of r. The new model can be used to calculate solution surface tension over relevant ranges of solution conditions, specifically temperature and composition making it especially valuable for process simulation. The calculated values of AARD are 0.38 and 0.35% for the original and new models, respectively. The new model is tested with a 9-fold cross validation as shown in Figure 14, and the cross validation AARD is 0.43%.

Parameter screening

The sensitivity matrix for the surface tension model is evaluated using the same criteria as for the viscosity and density models. It is calculated as

$$N=\max\begin{bmatrix}\left|\left(\frac{\partial\sigma}{\partial\hat{a}}\right)_{r,\alpha}\right| & \left|\left(\frac{\partial\sigma}{\partial\hat{a}}\right)_{T,\alpha}\right| & \left|\left(\frac{\partial\sigma}{\partial\hat{a}}\right)_{T,r}\right| \\ --- & --- & --- \\ --- & --- & --- \\ \left|\left(\frac{\partial\sigma}{\partial\hat{j}}\right)_{r,\alpha}\right| & \left|\left(\frac{\partial\sigma}{\partial\hat{j}}\right)_{T,\alpha}\right| & \left|\left(\frac{\partial\sigma}{\partial\hat{j}}\right)_{T,r}\right|\end{bmatrix}$$

$$=\begin{bmatrix}0.5345 & 0.6207 & 1.0000\\ 0.0805 & 0.0977 & 0.3103\\ 0.0475 & 0.0555 & 0.3563\\ 0.3678 & 0.5747 & 0.6897\\ 0.2011 & 0.4138 & 0.3736\\ 0.3448 & 0.4885 & 0.3448\\ 0.1667 & 0.2414 & 0.3218\\ 0.0222 & 0.0257 & 0.0862\\ 0.2356 & 0.4483 & 0.2414\\ 0.1609 & 0.408 & 0.1609\end{bmatrix} \tag{34}$$

Given the results of the sensitivity matrix, the only parameter than can obviously be eliminated before Bayesian inference is h because all elements in the corresponding row are close to zero.

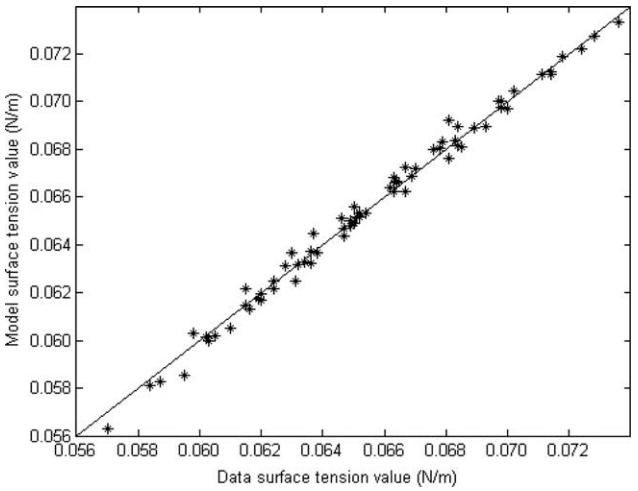

Figure 14. 9-fold cross validation of solution surface tension model.

Straight line represents perfect fit of model value to data value. Stars represent experimental data compared to model values.

Stochastic model

The 95% confidence intervals for the parameters in the surface tension model, excluding parameter h, are given in Table 13.

Parameters b and i are omitted from the stochastic model as the value 0 is included in their 95% confidence intervals. The prior distributions of the remaining seven parameters are estimated using the same methodology employed for the viscosity and density models, and the marginal PDFs and CDFs are given in Figure 15.

The marginal prior distributions are taken as the PDFs given in Figure 15, and the joint prior distribution is determined by assuming independence of the parameters. For UQ of the surface tension model, a sample of 100 parameter sets is obtained from a numerical approximation of the PDFs. A sample size of 72 experimental observations is available so that 7200 model observations are used for to generate the response surface model, which is validated by a 10-fold cross validation procedure with $R^2=0.9710$. Figure 16 shows a representation of the posterior distribution in terms of histograms for all single parameter marginal distributions and contour plots for selected two parameter marginal distributions. As with the density model, the entire joint posterior distribution cannot be visualized directly.

Table 13. Calculated 95% Confidence Intervals for Parameters in Surface Tension Model

Parameter	Baseline Value	Confidence Interval
a	2.4558	[1.5100, 3.4015]
b	−1.5311	[−3.4461, 0.3840]
c	3.4994	[1.9749, 5.0239]
d	−5.6398	[−9.7085, −1.5711]
e	10.2109	[3.5877, 16.8341]
f	2.3122	[1.2025, 3.4219]
g	4.5608	[0.7945, 8.3271]
i	5.3324	[−1.9564, 12.6211]
j	−12.0494	[−23.5634, −0.5355]

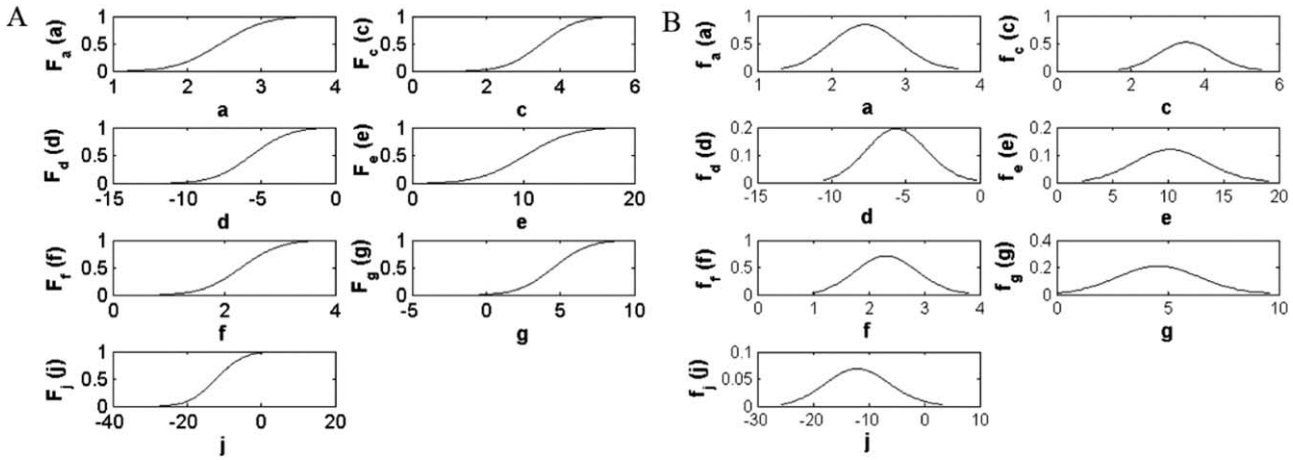

Figure 15. Approximations of (A) cumulative distribution functions and (B) probability distribution functions for surface tension model parameters determined from the confidence intervals in the deterministic regression results.

Effect of Physical Properties Uncertainty on the Process Model Outputs

The objective of this work is to estimate the effect of the parametric uncertainty of the viscosity, density, and surface tension models on the results of Aspen Plus® simulations. The Aspen Plus® model of the MEA-H_2O-CO_2 system that includes the absorber and the regenerator is the Phoenix model that has been developed at the University of Texas at Austin.[29] For the model parameters determined to be important for UQ in this work, a sample size of 200 is drawn from the PDF derived from Bayesian inference. All parameters included in the deterministic models but not in the stochastic models are treated as constants. Absorber and regenerator models are considered separately, and one operating condition is examined for each. For the absorber

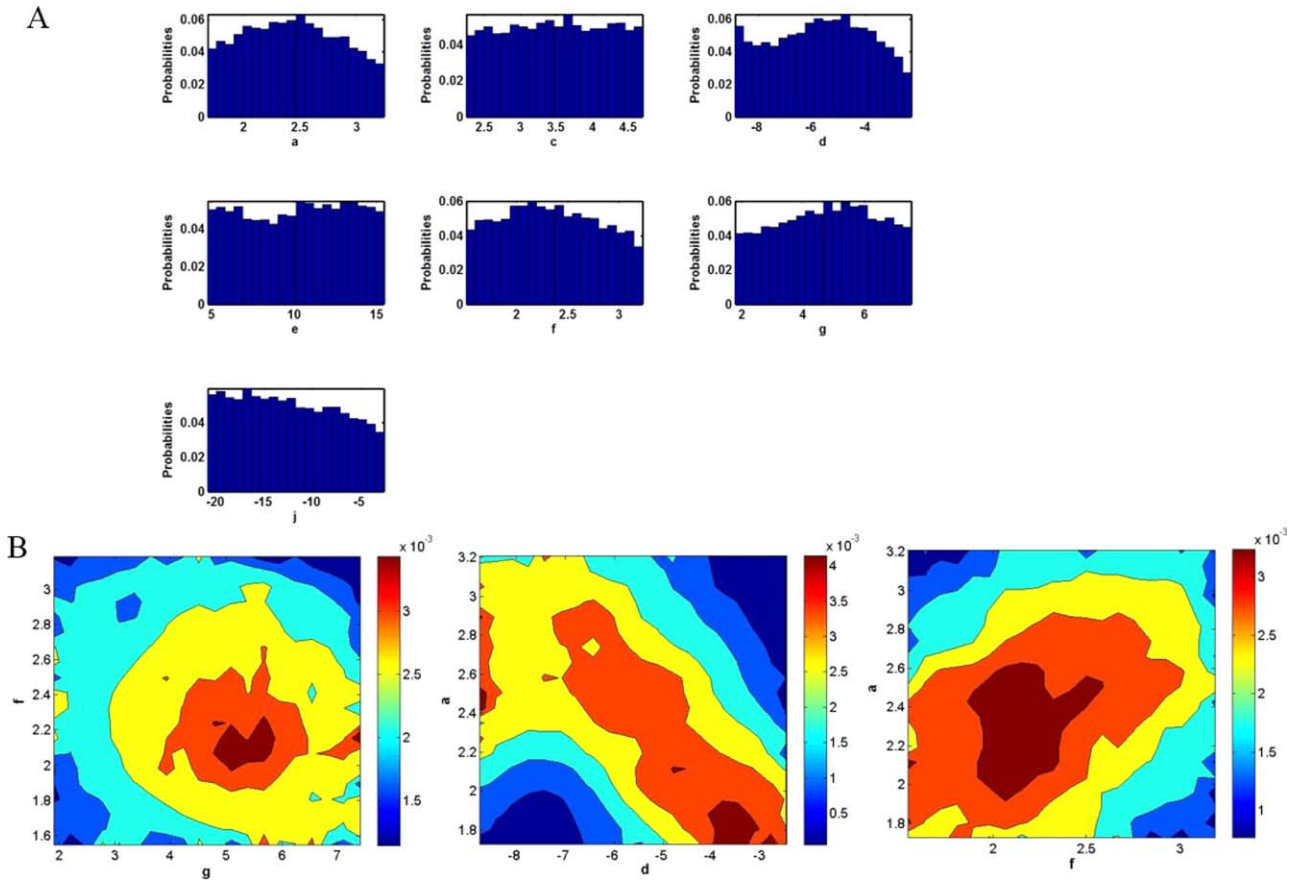

Figure 16. Posterior distributions for surface tension model parameters obtained from Bayesian inference.

(A) Histograms representing single parameter marginal distributions of the parameters; (B) contour plots representing two parameter marginal distributions. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Table 14. Key Input Variables for Absorber Simulation

Inlet lean solvent mass flow rate (kg/h)	3000
L/G (inlet liquid:gas mass ratio)	4.42
Inlet lean solvent temperature (°C)	40
Inlet lean solvent pressure (kPa)	101.3
Inlet lean solvent loading (mol CO_2/mol MEA)	0.35
Inlet lean solvent W_{MEA} (%)	35.4

simulation, the outlet variable of interest is percent CO_2 capture, calculated as

$$\%CO_2, \text{cap} = \left(\frac{\dot{m}_{CO_2,FGin} - \dot{m}_{CO_2,FGout}}{\dot{m}_{CO_2,FGin}} \right) * 100\% \qquad (35)$$

where $\dot{m}_{CO_2,FGin}$ and $\dot{m}_{CO_2,FGout}$ are the mass flow rates of CO_2 in the absorber inlet and outlet gas streams, respectively. Essential input variables for the absorber simulation are shown in Table 14.

The effect of parametric uncertainty in the three models on the uncertainty of the absorber model performance is estimated by simulating the model highlighted in Table 14 a total of 200 times, each representing a different set of the parameters, which were sampled from their respective posterior distributions. The results of this procedure are given in Table 15. The output variables considered are the percentage of CO_2 capture and the values of the physical properties in the inlet lean solvent stream of the absorber, and the ranges, averages, and standard deviations that result from propagating the parametric uncertainty is reported. The uncertainty in the percent of CO_2 capture that results from the propagation of the prior and posterior distributions through the process model is given in the form of histograms in Figure 17.

Table 16. Key Input Variables for Regenerator Simulation

Inlet rich solvent mass flow rate (kg/h)	3100
Inlet rich solvent temperature (°C)	106.5
Inlet rich solvent pressure (kPa)	230.4
Inlet rich solvent loading (mol CO_2/mol MEA)	0.5
Inlet rich solvent X_{MEA} (%)	35.4
Outlet lean solvent loading (mol CO_2/mol MEA)	0.3

In the MEA regenerator simulation, the reboiler duty ($\dot{Q}$) required to reduce the lean solvent loading to a given value ($\alpha = 0.3$) is the major output variable of interest, and the physical properties of the inlet rich solvent are also considered. The simulation input variables are given in Table 16, and the distributions in the output variables are given in Table 17.

As shown in Table 17, the uncertainty in the reboiler duty that results from propagating the parametric uncertainty (even while using the prior) through the regenerator model is negligible. A likely reason that this parametric uncertainty has less effect on the regenerator model in comparison to the absorber model is that the value of viscosity is much less in the regenerator model, and the CO_2 capture process is more sensitive to viscosity at higher values. It should be noted that all three models are extrapolated to the operating temperature of the regenerator, as no data are yet available to validate the models at temperatures at this range.

Although accuracy of viscosity, density, and surface tension models has high impact on the key outputs of an MEA-H_2O-CO_2 system, the results show that the small magnitude

Table 15. Results of Propagating Prior and Posterior Distributions of Physical Property Models Through Absorber Simulation

	Prior Distributions			Posterior Distributions		
	Range	Average	Standard Deviation	Range	Average	Standard Deviation
$\%_{CO_{2,cap}}$	82.97–85.09	84.05	0.46	83.53–84.48	83.97	0.19
μ_{lean} (mPa s)	1.95–5.11	2.95	0.70	2.53–3.31	2.89	0.16
ρ_{lean} (g/cm^3)	1.052–1.113	1.083	0.012	1.048–1.106	1.072	0.012
σ_{lean} (N/m)	0.0553–0.0786	0.0670	0.0051	0.051–0.0776	0.0664	0.0048

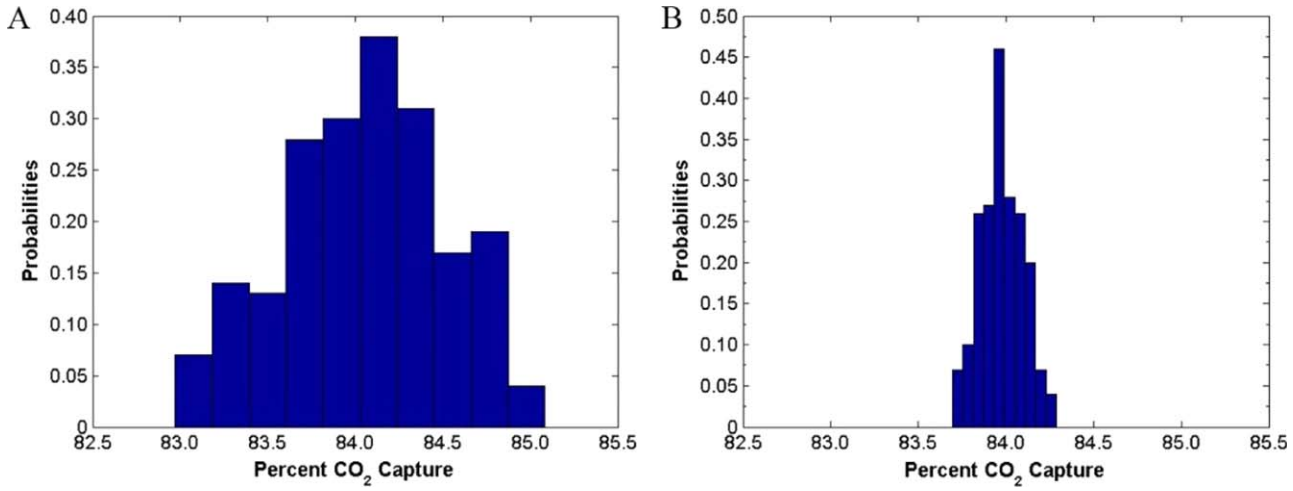

Figure 17. Histograms of percent CO_2 capture resulting from absorber model simulation with (A) prior distributions in model parameters and (B) posterior distributions in model parameters.

[Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Table 17. Results of Propagating Prior and Posterior Distributions of Physical Property Models Through Regenerator Simulation

	Prior Distributions			Posterior Distributions		
	Range	Average	Standard Deviation	Range	Average	Standard Deviation
$\dot{Q}$ (kW)	141.3335–141.3389	141.3366	0.0013	141.3363–141.3378	141.3370	0.0003
μ_{rich} (mPa s)	0.69–1.91	1.12	0.27	0.87–1.18	1.03	0.07
ρ_{rice} (g/cm^3)	0.958–1.034	0.997	0.015	1.000–1.029	1.011	0.007
σ_{rice} (N/m)	0.0451–0.0809	0.0631	0.0076	0.0544–0.0706	0.0629	0.0039

of parametric uncertainty (due to highly accurate deterministic models) does not affect the key outputs appreciably. It should be noted that the effect of the complete physical properties model uncertainty on the integrated plant model and on the dynamic simulation has yet to be studied.

Conclusions

A generalized approach to quantifying the uncertainty of property models is presented here. The Bayesian inference methodology deployed for this purpose seeks to update prior beliefs of parametric uncertainties with due consideration of the experimental data. To reduce the computational cost, response surfaces are used as surrogate models for Bayesian inference using the MCMC method with Gibbs sampling. For downselecting the parameter space, a sensitivity matrix approach based on computing analytical derivatives of physical property models with respect to parameters is developed. The methodology is demonstrated on the viscosity, density, and surface tension models of a MEA-CO_2-H_2O system. To this end, first deterministic models are evaluated over ranges of 0–0.5 mol CO_2/mol MEA, 0.2–0.4 g MEA/g MEA + H_2O, and 298.15–353.15 K for viscosity and density and 303.15–333.15 K for surface tension. The optimal baseline model parameters have been determined for each correlation by minimizing SSE and the model forms have been evaluated using cross validation. The sensitivity matrix approach is found to work effectively for all the models tested and can be used as a tool for downselecting the parameter space. It was observed that the parametric uncertainty in the current models for viscosity, density, and surface tension affects the key outputs of the MEA-based CO_2 capture system insignificantly. In the future, we will employ similar methodologies to complete other physical property models required for MEA systems.

Acknowledgments

As part of the National Energy Technology Laboratory's Regional University Alliance (NETL-RUA), a collaborative initiative of the NETL, this technical effort was performed under the RES contract DE-FE0004000. The authors would like to thank Prof. Gary T. Rochelle from The University of Texas at Austin for sharing the Phoenix model. The authors sincerely acknowledge valuable discussions with Prof. Rochelle and Brent Sherman from The University of Texas at Austin.

Literature Cited

1. Tester JW, Modell M. Statistical mechanical approach for property models. *Thermodynamics and Its Applications*, 3rd ed. Upper Saddle River, NJ: Prentice Hall PTR, 1997:388–455.
2. Angelikopoulos P, Papadimitriou C, Koumoutsakos P. Bayesian uncertainty quantification and propagation in molecular dynamics simulations: a high performance computing framework. *J Chem Phys*. 2012;137:144103.
3. Mah RSH. Effects of thermophysical property estimation on process design. *Comput Chem Eng*. 1977;1:183–189.
4. Streich M, Kistenmacher H. Property inaccuracies influence low temperature designs. *Hydrocarbon Process*. 1979;58:237–241.
5. Chirico RD, de Loos TW, Gmehling J, Goodwin ARH, Gupta S, Haynes WM, Marsh KN, Rives V, Olson JD, Spencer C, Brennecke JF, Trusler JPM. Guidelines for reporting of phase equilibrium measurements (IUPAC Recommendations 2012). *Pure Appl Chem*. 2012; 84:1785–1813.
6. Papadimitriou C. Bayesian uncertainty quantification and propagation in structural dynamics simulations. In: *International Conference on Structural Dynamics*. Porto, Portugal, 2014.
7. Kim SH, Kang J, Kroenlein K, Magee JW, Diky V, Frenkel M. Online resources in chemical engineering education: impact of the uncertainty concept from thermophysical properties. *Chem Eng Ed*. 2013;47:48–57.
8. Mathias PM. Sensitivity of process design to phase equilibrium - a new perturbation method based upon the Margules equation. *J Chem Eng Data*. 2014;59:1006–1015.
9. Turton R, Bailie RC, Whiting WB, Shaeiwitz JA, Bhattacharyya D. Synthesis of a process using a simulator and simulator troubleshooting. *Analysis, Synthesis, and Design of Chemical Processes*, 4th ed. Upper Saddle River, NJ: Prentice Hall, 2012.
10. Gel A, Chaudhari K, Turton R, Nicoletti P. Application of uncertainty quantification methods for coal devolatilization kinetics in gasifier modeling. *Powder Technol*. 2014;265:66–75.
11. Gel A, Li T, Gopalan B, Shahnam M, Syamlal M. Validation and uncertainty quantification of a multiphase computational fluid dynamics model. *Ind Eng Chem Res*. 2013;52:11424–11435.
12. Lane WA, Storlie CB, Montgomery CJ, Ryan EM. Numerical modeling and uncertainty quantification of a bubbling fluidized bed with immersed horizontal tubes. *Powder Technol*. 2014;253: 733–743.
13. Mebane DS, Bhat KS, Kress JD, Fauth DJ, Gray ML, Lee A, Miller DC. Bayesian calibration of thermodynamic models for the uptake of CO_2 in supported amine sorbents using ab initio priors. *Phys Chem Chem Phys*. 2013;15:4355–4366.
14. Miller DC, Syamlal M, Mebane DS, Storlie C, Bhattacharyya D, Sahinidis NV, Agarwal D, Tong C, Zitney SE, Sarkar A, Sun X, Sundaresan S, Ryan E, Engel D, Dale C. Carbon capture simulation initiative: a case study in multiscale modeling and new challenges. *Annu Rev Chem Biomol Eng*. 2014;5:301–323.
15. Jayarathna SA, Weerasooriya A, Dayarathna S, Eimer DA, Melaaen MC. Densities and surface tensions of CO_2 loaded aqeous monoethanolamine solutions with r=(0.2 to 0.7) at T=(303.15 to 333.15)K. *J Chem Eng Data*. 2013;58:986–992.
16. Amundsen TG, Lars EØ, Eimer DA. Density and viscosity of monoethanolamine + water + carbon dioxide from (25 to 80)°. *J Chem Eng Data*. 2009;54:3096–3100.
17. Chen X, Rochelle G. Aqueous piperazine derivatives for CO_2 capture: accurate screening by a wetted wall column. *Chem Eng Res Des*. 2011;89:3096–3100.
18. Eckert JS. Selecting the proper distillation column packing. *Chem Eng Prog*. 1970;66:39–44.
19. Weiland RH, Dingman JC, Cronin B, Browning GJ. Density and viscosity of some partially carbonated aqueous alkanolamine solutions and their blends. *J Chem Eng Data*. 1998;43:378–382.
20. Wang GQ, Yuan XG, Yu KT. Review of mass-transfer correlations for packed columns. *Ind Eng Chem Res*. 2005;44:8715–8729.
21. Rate-based model of the CO_2 capture process by MEA using Aspen Plus. Cambridge, MA: Aspen Technology, Inc., 2008–2012. Distributed with Aspen Plus V8.4.
22. Weiland RH. Physical properties of MEA, DEA, MDEA, and MDEA-based blends loaded with CO_2. *Gas Processors Association*

Research Report No. 152. Tulsa, OK: Gas Processors Association, 1996.
23. Han J, Jin J, Eimer DA, Melaaen MC. Density of water(1) + monoethanolamine (2) + CO_2 (3) from (298.15 to 413.15) K and surface tension of water(1) + monoethanolamine(2) from (303.15 to 333.15)K. *J Chem Eng Data*. 2012;57:1095–1103.
24. Connors KA, Wright JL. Dependence of surface tension on composition of binary aqueous-organic solutions. *Anal Chem*. 1989;61:194–198.
25. Friedman JH. Multivariate adaptive regression splines. *Ann Stat*. 1991;19:1–141.
26. Press WH, Teukolsky SA, Vetterling WT, Flannery BP. Modeling of data. *Numerical Recipes in C: The Art of Scientific Computing*, 2nd ed. New York: Cambridge University Press, 1992:656–706.
27. Beers KJ. Probability theory and stochastic simulation. *Numerical Methods for Chemical Engineering: Applications in MATLAB®*. New York: Cambridge University Press, 2007:372–435.
28. Weast RC. *Handbook of Chemistry and Physics*, 65th ed. Boca Raton, FL: CRC Press, 1984.
29. Plaza JM. Modeling of carbon dioxide absorption using aqueous monoethanolamine, piperazine, and promoted potassium carbonate. PhD Dissertation. Austin, TX: The University of Texas at Austin, 2012.
30. Hilliard MD. A predictive thermodynamic model for an aqueous blend of potassium carbonate, piperazine, and monoethanolamine for carbon dioxide capture from flue gas. PhD Dissertation. Austin, TX: The University of Texas at Austin, 2008.
31. Liley PE, Thomson GH, Friend DG, Daubert TE, Buck E. Physical and chemical data. *Perry's Chemical Engineers' Handbook*, 7th ed. New York: McGraw-Hill, 1997.
32. Asprion N. Surface tension models for aqueous amine blends. *Ind Eng Chem Res*. 2005;44:7270–7278.

Manuscript received July 25, 2014, and revision received Jan. 9, 2015.