Bayesian calibration of thermodynamic models for the uptake of CO\textsubscript{2} in supported amine sorbents using \textit{ab initio} priors

David S. Mebane,*\textsuperscript{ab} K. Sham Bhat,\textsuperscript{c} Joel D. Kress,\textsuperscript{c} Daniel J. Fauth,\textsuperscript{a} McMah\textsuperscript{an} L. Gray,\textsuperscript{a} Andrew Lee\textsuperscript{a} and David C. Miller\textsuperscript{a}

A statistical methodology was applied to the simultaneous calibration and validation of thermodynamic models for the uptake of CO\textsubscript{2} in mesoporous silica-supported amines. The methodology is Bayesian, and follows the procedure introduced by Kennedy and O’Hagan. One key aspect of the application presented is the use of quantum chemical calculations to define prior probability distributions for physical model parameters. Inclusion of this prior information proved to be crucial to the identifiability of model parameters against experimental thermogravimetric data. Through the statistical analysis, a quantitative assessment of the accuracy of various quantum chemical methods is produced. Another important aspect of the current approach is the conditioning of the model form discrepancy – a critical component of the Kennedy and O’Hagan methodology – to the experimental data in such a manner that it becomes an implicit function of the model parameters and thereby connected with the posterior distribution. It is shown that the inclusion of prior information in the analysis leads to a shifting of uncertainty from the posterior distribution for model parameters to this conditioned model form discrepancy. Prospects for more accurate model predictions and propagation of uncertainty in upscaling and extrapolation through a “model-plus-discrepancy” approach are discussed. The synthesis methods and thermogravimetric characterization of hybrid grafted/impregnated mesoporous silica-supported amine sorbents are presented, along with the details of the quantum chemical study, which shows that a carbamic acid–base acceptor complex is the most stable form of adsorbed CO\textsubscript{2} in both alkanol- and ethyleneamines.

1 Introduction

The Carbon Capture Simulation Initiative (CCSI) of the U.S. Department of Energy is a project involving five U.S. National Laboratories, universities and over 20 industrial partners and is focused on developing computational tools and models for accelerating the development of new carbon capture technologies for large-scale power generation.\textsuperscript{2,3} The resulting CCSI toolset will enable companies to more effectively screen, design, scale up and assess potential carbon capture technologies. An essential part of the project is to quantify the uncertainty and technical risk associated with the simulated processes. Although there are many potentially viable technologies under development, CCSI is using solid sorbent-based systems to demonstrate the capabilities of the CCSI toolset.

Solid sorbent carbon capture systems involve phenomena at the quantum scale (through adsorption chemistry), all the way through to the industrial macroscale. One crucial problem faced by the project is how to increase confidence in the predictions of such broad multi-scale simulations. Two important and connected strategies are the use of simulations that adhere as closely as possible to scientific fundamentals, and the development and the application of advanced statistical methodologies for the quantification of uncertainty.

Such an advanced methodology for model validation and uncertainty quantification, one that enables the derivation of maximum engineering utility from a limited set of experimental data and computational resources, is presented and demonstrated...
in this article. The application context is the development of thermodynamic models for the adsorption of CO₂ in amine-based sorbents. A simple, thermodynamically ideal model for the equilibrium adsorption of CO₂ in anhydrous conditions is calibrated to a limited set of thermogravimetrically obtained isotherm data. Ab initio calculations are deployed to create probability distributions for unknown physical model parameters. These distributions are used as prior information in a Bayesian calibration framework. This framework also considers the error inherent in the model itself, through the inclusion of a thermodynamically consistent model form discrepancy. The analysis produces a posterior distribution for the model parameters which arises from the combined influence of experimental observations and ab initio calculations. The analysis also leads to a method for correcting model predictions using the experiment-conditioned discrepancy, in such a manner that the experimental data are reproduced to within the estimated observation error. This “model-plus-discrepancy” prediction capability has important implications for multi-scale modeling.

2 Mesoporous silica-supported amine sorbents

2.1 Background

The great advantage of amines in carbon capture are their nearly perfect chemical selectivity for CO₂ over N₂. Most current industrial carbon dioxide separation processes consist of a thermal swing procedure involving amines in aqueous solution. The difficulty with these processes are the large energy penalties due to the evaporation of water in the regenerator (i.e., the desorption stage of the process), where CO₂ is desorbed at high temperature (≈100–120 °C). Solid sorbents functionalized or impregnated with amines – including the polyamine polyethylenimine (PEI) – were designed to ameliorate this problem by removing amines from aqueous solution and affixing them to solid supports with high surface area (along with adequate pore volume and pore size), and low heat capacity relative to water. Mesoporous silica-supported amines have been extensively studied in the past decade. Typical methods for incorporating amines on the mesoporous silica may include direct co-condensation during synthesis or post-synthesis grafting of primary or secondary aminosilanes onto surface silanol sites or wet impregnation from solution of an amine polymer (e.g., PEI) into the silica mesopores where an ionic bond is formed at the silanol sites. The grafted amine types characteristically have the advantage of faster kinetics, while the impregnated types have a higher sorbency. The sorbent which is the focus of this study is the result of a hybridization of post-synthesis grafting and impregnation, synthesized and characterized at NETL.

2.2 Sorbent synthesis

Analysis of PEI-impregnated mesoporous silica materials found that the sorbent takes up a significant amount of water in simulated flue gas (which is saturated with water vapor). For this reason, an alternate method of synthesis was employed: aminosilanes and PEI were introduced in a one-pot synthesis, in the hopes of improving the deposition of PEI while tailoring the material to be more hydrophobic through the inclusion of moderate concentrations of covalently bound amino-functional groups.

PEI, N1-(3-trimethoxysilylpropyl)diethylenediamine (TMPED) and reagent grade methanol were obtained from Aldrich Chemicals. The silica substrate was obtained from PQ, Inc. with an average particle size of 100 microns, internal surface area of 280 m² g⁻¹ and pore volume of 1.2 ml g⁻¹. The PEI (molecular weight 600 a.u.) and TMPED were dissolved in methanol and combined with PQ 2129 at a 12–28–60 weight ratio. This amine–aminosilane slurry was then placed in a rotary evaporator and agitated for 40–60 minutes under various stages of reduced pressure above 80 °C, which removed the solvent. This immobilization procedure is described in detail in patents from Birbara et al. and Gray et al.

2.3 Sorbent characterization

Substrate characterization was carried out by N₂ adsorption–desorption isotherm analysis at 77 K, using a Micromeritics ASAP 2020™. The measured internal surface area was 285 m² g⁻¹, the pore volume 2.4 cc g⁻¹, and the average pore diameter 33.8 nm. The average particle size was 100 microns. Total organics loading along with CO₂ adsorption isotherms for the synthesized sorbents were experimentally obtained using a Thermo Scientific ThermMax 300™ thermogravimetric analyzer (TGA). The weight fraction of total organics deposited onto the substrate was obtained upon thermally heating the sample under nitrogen from ambient to a final temperature of 600 °C, and was found to be 22.5 wt%. For the isotherms, samples were dried in a vacuum dryer overnight at 70 °C before sorption measurement. In a typical experiment, a 50 mg sample was loaded within a microbalance quartz sample bowl. Prior to adsorption measurements, the samples were heated to 378 K in nitrogen (flow rate: 100 ml min⁻¹) at a heating rate of 5 K min⁻¹ and held isothermally for 90 min to remove pre-adsorbed CO₂ and moisture. The temperature was then increased to the initial adsorption temperature (381 K) before the reactant gas was switched to a dry mixture of 4, 7.5, 10, 18.5 or 100% CO₂, balance N₂ where appropriate, maintaining a flow rate of 100 ml min⁻¹. Subsequent CO₂ adsorption steps were conducted at 373, 365, 338, 330, 322, and 314 K, followed by desorption steps at 338 and 381 K. The hold time was approximately 90 min for each step.

3 Chemical theory and methods

3.1 Quantum chemistry

Both density functional theory (DFT) and wave function-based methods – namely, 2nd- (MP2) and 3rd- (MP3) order Möller–Plesset perturbation theory – electronic structure calculations were performed with Gaussian 09 using a 6-31+G* basis set. For DFT, B3LYP, PBE0 and PBE0/TZVP functional were used. From the analytical Hessian, vibrational frequencies and zero-point energies (ZPEs) were calculated.
Given the number of calculations to be performed in this study and time constraints associated with the broader project, it was not practical to work with macromolecules. Instead, a number of calculations were made with smaller molecules comprised of ethylene or methyl groups with one, two or three amine sites, with both primary and secondary amines represented. The molecules considered were monomethylamine (MMA), dimethylamine (DMA), ethylenediamine (EDA) and diethylentriamine (DETA).

Enthalpies of reaction were calculated as \( \Delta H = \Delta E + P \Delta V \), where \( \Delta E = \) electronic energy + \( \Delta ZPE + \Delta E_{\text{trans}} + \Delta E_{\text{rot}} \). The translational \( \Delta E_{\text{trans}} \) and rotational \( \Delta E_{\text{rot}} \) are calculated for the CO\(_2\) gas phase molecule only. For the work term, ideal behavior was assumed, yielding \( P \Delta V = \Delta nRT = -RT \). This strategy assumes that \( \Delta E_{\text{trans}} \) and \( \Delta E_{\text{rot}} \) to the vibrational \( \Delta E_{\text{vib}} \) are equivalent for reactants and products in the adsorption reactions considered. (This is a good approximation in the limit that each entity is a piece of a long-chain polymer.)

Vibrational modes were calculated using the B3LYP functional, and assignments made based on visualization.

### 3.2 Macroscale thermodynamics

The results of the \textit{ab initio} study (discussed below) led to a model for the uptake of CO\(_2\) in anhydrous conditions according to the following reaction (as written for secondary amines):

\[
R_2NH + R'_2NH + CO_2(g) \rightarrow R_2NCOOH : R'_2NH
\]  

where \( R \) and \( R' \) can be methyl, ethyl, or hydrogen. (See Fig. 1.) Assuming ideal thermodynamics and a monodisperse distribution of sites, the gravimetric data may be interpreted according to the following model:

\[
\kappa = \frac{x^2}{(1 - 2x)\rho} = \exp\left[\frac{\Delta S}{R}\right] \exp\left[-\frac{\Delta H}{RT}\right] / P
\]

\[
w = Mn_vx / \rho
\]

where \( \kappa \) is the equilibrium constant, \( x \) is the site fraction of carbamic acid in the amine (which is equal to the site fraction of associated amines), \( p \) is the partial pressure of CO\(_2\), \( \Delta S \) and \( \Delta H \) are the entropy and enthalpy of the reaction, respectively, \( P \) is the overall pressure, \( w \) is the fractional weight gain, \( M \) is the molecular weight of CO\(_2\), \( n_v \) is the number density of amine sites that are active for the uptake of CO\(_2\), and \( \rho \) is the overall gravimetric density of the sorbent. In the following analysis, \( p \) and \( T \) are referred to as “model inputs,” while \( w \) — which can be thought of as a function of \( p \) and \( T \) — is referred to as the “output.”

The unknown physical parameters that will be estimated through the calibration are \( \Delta S \), \( \Delta H \) and \( n_v \).

### 4 Statistical theory and methods

#### 4.1 Introduction to Bayesian calibration

In statistics, there are two principal paradigms: frequentist and Bayesian. In a typical frequentist application, a probability distribution of a stochastic observed phenomenon is inferred from a set of observations of that phenomenon. But the probability distribution is only known exactly in the limit of an infinite number of observations. Bayesian inference, by contrast, does not depend on an infinite amount of data. Instead, a Bayesian analysis begins with a pre-existing probability distribution, called the \textit{prior} distribution. A finite amount of data can then be taken into consideration, leading to the \textit{posterior} distribution. The mathematical relationship between the prior, the observations, and the posterior is given by Bayes’ theorem,

\[
P(A|B) = \frac{P(B|A)P(A)}{\int_A P(B|A')dA'}
\]

distribution of a stochastic observed phenomenon is inferred from a set of observations of that phenomenon. But the probability distribution is only known exactly in the limit of an infinite number of observations. Bayesian inference, by contrast, does not depend on an infinite amount of data. Instead, a Bayesian analysis begins with a pre-existing probability distribution, called the prior distribution. A finite amount of data can then be taken into consideration, leading to the posterior distribution. The mathematical relationship between the prior, the observations, and the posterior is given by Bayes’ theorem,

\[
P(A|B) = \frac{P(B|A)P(A)}{\int_A P(B|A')dA'}
\]

where \( P(A|B) \) is the posterior, the probability of \( A \) given new information \( B \); \( P(B|A) \), called the likelihood, is the probability of observing \( B \) given \( A \); \( P(A) \) is the prior probability of \( A \); and the denominator is the integrated probability of observing \( B \) over all possible values of \( A \).
Model calibration is the estimation of the unknown parameters of a model through a comparison with data. In a Bayesian context, the model parameters are represented by a joint (or multi-dimensional) probability distribution; this probability distribution reflects our knowledge of what the parameters – which can be thought of as having a single set of “true” values – are. Higher probabilities in the distribution correspond to values more likely to be correct. Given a prior distribution $\mathcal{P}(A)$ for the parameters, a posterior distribution $\mathcal{P}(B|A)$ in light of the data $B$ is sought. The likelihood $\mathcal{P}(B|A)$ implies a stochastic process, depending on $A$, of which the actual observed data $B$ is a realization.

A powerful Bayesian approach to model calibration was published in 2001 by Kennedy and O’Hagan. It is the form of the likelihood that distinguishes the Kennedy–O’Hagan method from its predecessors. It incorporates first a deterministic model of the process that has produced the data. The difference between this model result (given a set of parameter values) and the data is then considered to be the sum of two stochastic terms. One of these terms represents observation error. The other represents the shortcomings of the model with respect to the actual process, called the model discrepancy. A stochastic model for the observations is therefore written

$$Z = Y(\theta) + \delta(\xi) + \varepsilon(\psi)$$

with $Z$ a stochastic vector containing the set of all observations, $Y$ the model output function with parameters $\theta$, $\delta$ the model discrepancy with its parameters $\xi$, and $\varepsilon$ the model for the observation error with its parameters $\psi$. The actual experimental data $Z$ is then considered to be a realization of this stochastic process, to which a probability (or likelihood) can be assigned. The discrepancy $\delta$ and the stochastic function are enforced. A covariance generating function for the GP, and it is through this generating function that the required characteristics of the covariance function are enforced. A covariance generating function of the Matérn family satisfies the necessary requirements in the case of thermodynamic equilibrium. Matérn covariance matrices are stationary, in that the correlation between two input points depends only on the difference between them. Matérn covariances also feature precise control over the differentiability of the sampled functions. Of the Matérn covariance generating functions, the squared-exponential

$$\Sigma_{ij} = \sigma^2 \exp\left[-\frac{(z_i - z_j)^2}{\phi^2}\right]$$

takes a particularly simple form; sampled functions from a GP with a squared-exponential covariance matrix are guaranteed to be infinitely differentiable. Here the range parameter $\phi$ controls the extent of the correlation between two points with respect to their distance from one another, and the sill $\sigma^2$ controls the scale of the variance of the function from its mean. Splitting the state point vector $z$ into its individual components of temperature and partial pressure of CO$_2$, the model discrepancy can be written as a multivariate normally-distributed random vector,

$$\delta \sim \text{MVN}(0, \Sigma)$$

where

$$\Sigma_{ij} = \sigma^2 \exp\left[-\frac{(T_i - T_j)^2}{\phi_T^2}\right] \exp\left[-\frac{(P_i - P_j)^2}{\phi_P^2}\right]$$

$\delta$ in the absence of phase transformations, $\delta$ is thermodynamically only required to be once-differentiable; in this work it is assumed that $\delta$, in conformity with the model (2) and (3), is infinitely differentiable, although this is overly stringent.

---

‡ For general information on Bayesian statistical methods, see ref. 27 and 28.
is the symmetric positive definite covariance matrix, so that $\xi = \{\sigma^2, \phi, \phi_T\}$. The observation error is also described by a GP, with a zero mean vector and a covariance matrix consisting of a single parameter $\psi$ at each entry along the diagonal. This defines an observation uncertainty consisting of an independent, univariate normal distribution of variance $\psi$ centered at each observation point.

With the distributions of all the variables in the statistical model (5) defined, the distribution for the experimental observations $\mathcal{Z}$ can be written:

$$\mathcal{Z} \sim \text{MVN}(Y(\theta), \Sigma[\xi] + \psi I)$$

where $I$ is the identity matrix. This distribution can be used to assign a probability density to the actual data $\mathcal{Z}$, which is the likelihood $\mathcal{L}(\mathcal{Z}|\theta, \xi, \psi)$ appearing in eqn (6).

### 4.3 Discrepancy conditioning and predictions

The distribution (11) associated with the likelihood can be conditioned to the observations $\mathcal{Z}$ such that the construction $Y + \delta$ will predict the observations to within the observation error, with increasing uncertainty for predictions at new locations. This means that the actual observations $\mathcal{Z}$ are taken as known to within observation error, and the resulting distribution is changed to reflect this.

A joint likelihood for any new set of points on the output space $\mathcal{Z}_1$ associated with new inputs $\{\xi\}_1$ and the outputs $\mathcal{Z}_2$ associated with the known inputs $\{\xi\}_2$ is written

$$\begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix} \sim \text{MVN} \left( \begin{bmatrix} Y_1(\theta) \\ Y_2(\theta) \end{bmatrix}, \begin{bmatrix} \Sigma_{11} + \psi I_1 & \Sigma_{12} \\ \Sigma_{12} & \Sigma_{22} + \psi I_2 \end{bmatrix} \right)$$

where each block of the overall covariance matrix $\Sigma(\xi)$ is computed according to (10) using the relevant inputs, and $I_n$ is the identity matrix of the same dimension as $\mathcal{Z}_n$. The conditional distribution for $\mathcal{Z}_i$, given $\mathcal{Z}$, $\theta$, $\xi$, and $\psi$ is therefore $\xi$

$$\mathcal{Z}_i|\mathcal{Z}, \theta, \xi, \psi \sim \text{MVN}(Y_i + \Sigma_{i2}(\Sigma_{22} + \psi I_2)^{-1}(Z - Y_2), \Sigma_{ii} + \psi I_i - \Sigma_{i2}(\Sigma_{22} + \psi I_2)^{-1}\Sigma_{2i})$$

Multiplying the distribution (13) by the posterior distribution $\mathcal{Q}$ and integrating over its domain (i.e., over all model and discrepancy parameters) results in a “model-plus-discrepancy” predictive distribution for any set of inputs:

$$\mathcal{Z}|\mathcal{Z} \sim \int P(\mathcal{Z}_1|\mathcal{Z}, \theta, \xi, \psi)\mathcal{Q}(\theta, \xi, \psi|\mathcal{Z})d\theta d\xi d\psi$$

In the current application, this predictive distribution incorporates both first-principles (through the prior distribution for $\theta$) and experimental ($\mathcal{Z}$) information.

A measure of the size of the conditioned discrepancy as a function of the inputs is of analytical interest. Such a measure may be obtained by subtracting the model predictions from the mean term in the distribution (13), squaring it, and integrating over the full posterior. This will be called the “integrated square discrepancy”.

### 4.4 Implementation

Because of the difficulty in evaluating the integral $Q$ in eqn (6), full posterior distributions can rarely be expressed in closed form, meaning that a numerical simulation must be used in order to obtain samples from the posterior. Markov Chain Monte Carlo (MCMC) was used to obtain such samples, which were then converted into an empirical probability density function.

MCMC was run for 150,000 iterations, which required approximately 13 hours on a a 2.66 GHz 6-Core Intel Xeon on a Mac Pro desktop with 16 GB of RAM; the MCMC code was implemented in R. A “burn-in” of 10,000 samples was first produced and discarded, to ensure that subsequent samples are representative of the posterior distribution. Convergence of our MCMC-based samples to the posterior distribution was verified using a consistent batch means approach. The posterior mean estimates of the model parameters had MCMC standard errors below 0.01 times the mean estimates for both the model and discrepancy parameters.

### 5 Results and discussion

#### 5.1 Thermogravimetry

A sample output from the TGA for sorbent exposed to 18.5% CO$_2$, balance N$_2$, is shown in Fig. 2. From TGA runs such as this at 4, 7.5, 10, 18.5 and 100% CO$_2$, balance N$_2$, equilibrium data was extracted, as shown in Fig. 3.

Like many data sets, this one has some shortcomings that should be acknowledged, but should not otherwise discourage the effort to derive useful information from the data. The gap in the data between 338 and 365 K arises from a desire to focus on temperatures likely to be encountered in the adsorber and regenerator of an actual capture system. It is clear from several studies of PEI-impregnated sorbents of this type that the diffusion of CO$_2$ through the polymer bulk can be a limiting step in the reaction, especially for highly loaded sorbents at dry...
conditions and low temperatures. Although there is some evidence to suggest that there may be a limitation of this type in the sorbent of focus for this study (see the slow uptake of CO₂ at the first desorption step in Fig. 2), since this sorbent has a relatively low loading of active material (22.5 wt% is slightly less than half of the theoretical maximum), it is assumed that these effects are not too severe. It is also clear that there is a significant – though not overwhelming – observation error present in the data.

5.2 Quantum chemistry

5.2.1 Formation of carboxylic acid in alkanol- and ethyleneamines. Most theoretical studies of the interaction between amines and CO₂ have focused on amino alcohols in aqueous solution, particularly monoethanolamine (MEA). 43–51 The alkylammonium carbamate ion pair formed through the adsorption reaction is often assumed to associate through a hydrogen bond between alkylammonium and the nitrogen site of the carbamate anion. 44,46 This is consistent with the often-cited termolecular mechanism for the adsorption of CO₂ in these systems,52,53 wherein the encounter complex between gaseous CO₂ and the amine is immediately (without barrier) deprotonated by a second amine. However, another possibility is the formation of a carboxylic acid, by a transfer of a proton from the nitrogen atom of the encounter complex to one of the oxygens of the CO₂ molecule. This can be accomplished either by a bending of the encounter complex or through a catalytic proton exchange with water or another amine. 54,55 Another, related possibility is the formation of an alkylammonium carbamate associate with an H-bonded relationship at the oxygen of the carbamate group.

The relevant amine chemistry for this study is not amino alcohols, but ethyleneamines. It proved useful nonetheless to perform some calculations on MEA in aqueous solution in order to compare with the results of previous studies. Using the B3LYP DFT exchange–correlation functional and a polarizable continuum model (PCM) to approximate the effects of solvation, 56 the reaction energy for the formation of the alkylammonium carbamate associate was calculated to be −32.6 kJ mol⁻¹, which compares well with the result of −33.4 kJ mol⁻¹ arrived at by Shim et al. 46 Xie et al., another group to report the preferential formation of alkylammonium carbamate in MEA aqueous solution, reported only Gibbs free energies of reaction at 298 K, in which estimates of configurational entropies are likely to figure prominently. Their estimate for alkylammonium carbamate depended on the atomic radii used: UAHF radii yielded −20.5 kJ mol⁻¹ while Pauling radii yielded −15.5 kJ mol⁻¹.

Neither Xie et al. nor Shim et al. considered a carboxylic acid associate involving a hydrogen-bonded relationship between OH of the acid and the nitrogen of another amine molecule, which was found to be the most stable form of adsorbed CO₂ in the present study, both for alkanolamines and ethyleneamines. Such an associate was considered by Arstad et al., and their calculation of −55 kJ mol⁻¹ (using the G3MP2B3 compound method in Gaussian 03) compares reasonably well with our estimate of −72.3 kJ mol⁻¹ for the same associate using B3LYP. All subsequent references to estimates for adsorbed CO₂ in this system will refer to this associate, illustrated in Fig. 1, formed according to the reaction (1).

Table 1 lists the calculated enthalpy of reaction associated with reaction (1) for a number of different chemistries and ab initio methodologies. All of the calculations with results appearing in the table were obtained in the gas phase; calculations in a polar environment equivalent to that of diphenylamine (relative permittivity of 3) made only a negligible difference in the results. The molecular structures for the associate products in the first and fourth reactions in Table 1 appear in Fig. 1.

5.2.2 Accuracy of quantum chemical results. Calculations of binding energies for a set of 22 small molecule pairs (denoted S22) representative of the binding in DNA have compared MP2 with the highly-accurate benchmark coupled cluster (CCSD(T)) approach. 57 For the first seven pairs in S22, which are representative of the kind of binding that occurs in carboxylic acid:amine association complexes, a mean error in binding energy of 1.3 kJ mol⁻¹ was found for MP2 relative to CCSD(T).

<table>
<thead>
<tr>
<th>Reaction</th>
<th>B3LYP</th>
<th>PBE</th>
<th>PBE0</th>
<th>MP2</th>
<th>MP3</th>
</tr>
</thead>
<tbody>
<tr>
<td>CO₂ + 2MMA → MMACO₂:MMA</td>
<td>−52.72</td>
<td>−71.13</td>
<td>−81.59</td>
<td>−52.72</td>
<td>−72.8</td>
</tr>
<tr>
<td>CO₂ + MMA + DMA → MMACO₂:DMA</td>
<td>−46.86</td>
<td>−63.60</td>
<td>−76.99</td>
<td>−53.97</td>
<td>−71.96</td>
</tr>
<tr>
<td>CO₂ + DMA + MMACO₂</td>
<td>−50.21</td>
<td>−66.11</td>
<td>−79.50</td>
<td>−53.14</td>
<td>−72.38</td>
</tr>
<tr>
<td>CO₂ + DMA + MMACO₂</td>
<td>−46.86</td>
<td>−64.02</td>
<td>−77.40</td>
<td>−56.07</td>
<td>−72.80</td>
</tr>
<tr>
<td>CO₂ + DETA + 3EDA → EDACO₂:DETA₃</td>
<td>−37.66</td>
<td>−69.04</td>
<td>−69.04</td>
<td>−55.23</td>
<td>−70.29</td>
</tr>
<tr>
<td>CO₂ + DETA + 3EDA → DETA₃CO₂:EDA</td>
<td>−43.93</td>
<td>−61.50</td>
<td>−75.31</td>
<td>−65.27</td>
<td>−79.50</td>
</tr>
</tbody>
</table>
In a similar DFT study using the same seven pairs, mean errors in binding energy of 7.4, 4.7, and 3.3 kJ mol⁻¹ were found for B3LYP, PBE and PBE0, respectively. The a priori conclusion is that the wave-function based approach (MP2 and MP3) is expected to be the most accurate for the present study, although the hybrid DFT approach of PBE0 is also within “chemical accuracy” of 1 kcal mol⁻¹ (= 4.2 kJ mol⁻¹). Using the MP3 results in Table 1 as an example, we note that the spread in binding energy is less than 9 kJ mol⁻¹. (Calculations involving the 5-site chain triethylene pentamine remained within the spread.) This relatively small spread denotes a weak dependence on both the alkyl chain length (methyl vs. ethyl) and on primary vs. secondary amine for both the carboxylic acid and for the base acceptor. With this weak dependence, we conclude that the results in Table 1 are transferrable to the PEI amine polymer chain.

Further insight as to the accuracy of the methods employed can be derived through a comparison between B3LYP calculations of IR vibrational modes in the carboxylic acid:DMA base acceptor associate (see Fig. 1) in the gas phase with experimental studies on mesoporous silica-supported propyl- and ethyl-eneamines. In the following, peaks appearing in the study by Wang et al. will be listed in descending order of wavenumbers, although all peaks referenced were also observed in the other studies.

A peak found experimentally at 3420 cm⁻¹ and in the calculations at 3469 cm⁻¹ was assigned in both contexts to N–H stretch. A peak at 1650 cm⁻¹, assigned to deformation of NH₂⁺, was not seen in the calculations. Instead, an intermolecular hydrogen bonding mode (the H atom between the carboxylate group and amine in the associate, see Fig. 1) is found in the calculations at 2822 cm⁻¹ as well as an NCO asymmetric stretch/OH wag at 1747 cm⁻¹. A “general” C=O stretch was seen experimentally at 1520 cm⁻¹ and computationally at 1561 cm⁻¹. Finally, a skeletal vibration for NCOO was observed at 1410 cm⁻¹ and is found in the calculations at 1436 cm⁻¹.

The positions of the peaks and the assignment of the modes in the calculations generally agree with the experimental measurements and assignments to the spectra, with the exceptions of the NH₂⁺ deformation, which did not appear in the calculations, and the hydrogen bonding mode and NCO asymmetric stretch/OH wag, which did not appear in the experiment. These results, combined with the energetic calculations discussed above, point toward an alkylammonium carbamate complex, hydrogen bonded between the oxygen in the carbamate group and alkylammonium, in accordance with Scheme 1 in ref. 19. The difference between this complex and the carboxylic acid:amine complex described in the present work is a subtle one, involving the location of the proton involved in the hydrogen-bonded relationship. The calculated energy penalty for moving the proton closer to nitrogen is small. It seems therefore appropriate to use the results appearing in Table 1 and other results of the quantum chemical study in the calibration.

5.3 Calibration

5.3.1 Priors

5.3.1.1 Model parameters. Prior distributions are required for the three unknown parameters found in the model (2)–(3): ΔH, ΔS and nₓ. The prior for ΔH is found by inferring a normal distribution on the data contained in Table 1, as shown in Fig. 4a. The mean of the distribution is −60.84 kJ mol⁻¹ and its standard deviation is 11.8 kJ mol⁻¹. The distribution was bounded between −120 and −30 kJ mol⁻¹.

The entropy of the reaction will be dominated by the loss of translational and rotational degrees of freedom in CO₂ upon adsorption. This can be calculated exactly using closed-form equations implemented in Gaussian 09, with the result equal to −200 J mol⁻¹ K⁻¹. In addition to this entropy change due to the gas-to-adsorbed state transition, there will be additional entropic changes due to changing vibrational and translational degrees of freedom in the polymer. Calculations of vibrational spectra showed that vibrational entropy changes are on the order of less than 10 J mol⁻¹ K⁻¹. On the other hand, changes in the translational degrees of freedom in the polymer could be significant. The formation of carboxylic acid–base associates upon the uptake of CO₂ – along with steric effects – will most likely lead to a negative contribution to the entropy, in accordance with the observed increase in viscosity in amine-based ionic liquids upon the uptake of CO₂. A normal distribution was therefore used, with a mean of −250 J mol⁻¹ K⁻¹ and a standard deviation of 25 J mol⁻¹ K⁻¹, bounded between −500 and −200 J mol⁻¹ K⁻¹, as shown in Fig. 4b.

An estimate for the number of amine sites per unit volume of the sorbent is available through the knowledge about the number of nitrogen sites per unit volume in pure linear PEI, the density of pure silica and the pore volume of the substrate, and

![Fig. 4](image-url) Prior distributions for (a) ΔH, (b) ΔS and (c) nₓ.
is equal to 2351.0 mol m\(^{-3}\). That the actual number of active sites for the uptake of CO\(_2\) will be considerably less than this is clear, given the hyperbranched PEI used (tertiary amines are not active for adsorption of CO\(_2\)) and the known interaction between amine sites and silanol sites on the silica substrate.\(^{19}\) Approximate estimates for these two effects led to a normal distribution with a mean at 1469 mol m\(^{-3}\) and a standard deviation of 295 mol m\(^{-3}\). The distribution is shown in Fig. 4c.

5.3.1.2 Likelihood parameters. As discussed in Section 4.2, there are four parameters in the likelihood that must be estimated along with the model parameters: the sill \(\sigma^2\) and the range parameters \(\phi_P\) and \(\phi_T\) for the model form discrepancy, and the observation error variance \(\psi\). Range parameters were given uniform prior distributions on the interval \([100,40000]\) (Pa) for \(\phi_P\) and \([1,50]\) (K) for \(\phi_T\). The prior distribution for \(\sigma^2\) was derived by producing a large number of model outputs by sampling from the prior distributions for the model parameters, and fitting a GP with zero mean and covariance generating function given by (9) to the stochastic function defined by the difference between the data \(Z\) and the set of model outputs. The result was used in the definition of the inverse gamma (IG) distribution \(\sigma^2 \sim \text{IG}(2,0.2)\) on the interval \([0,0.4]\) (wt%), which was then used as the prior for \(\sigma^2\).\(^{1}\) A prior distribution for the observation error is best inferred from the data itself using repeat measurements, but in the absence of these, a smoothing operation on the available data was used to estimate a prior for \(\psi\), which yielded \(\psi \sim \text{IG}(2,0.01)\) on the interval \([0,0.05]\) (wt%).

5.3.2 Posteriors. Two posterior distributions were generated. The first utilized uniform priors, bounded on the same intervals as the distributions shown in Fig. 4. That distribution is shown in Fig. 5, as a set of three contour plot bivariate distributions.\(^{**}\) The second posterior utilized the informative priors shown in Fig. 4. That distribution is shown in Fig. 6.

The first thing to note is that the uncertainty in the posterior distribution for the case with informative priors is less than that of the case with uniform priors. This is clear from a visual examination of the figures (note the changes in scales between Fig. 5 and 6), but is also reflected in the Shannon entropy of the distributions, which is 4.966 for the case with uniform priors and 4.1 for the informative case. The entropy and enthalpy of the reaction are in fact practically unidentifiable in the uniform prior case, with regions of high probability (the distribution is multimodal) spanning nearly 30 kJ mol\(^{-1}\) of enthalpy and 80 J mol\(^{-1}\) K\(^{-1}\) of entropy. In the informative case, by contrast, the region of highest probability spans considerably less than 10 kJ mol\(^{-1}\) in enthalpy and 15 J mol\(^{-1}\) K\(^{-1}\) of entropy. The distribution in Fig. 6 has also lost much of the multimodality seen in Fig. 5, although the peaks in the former are not quite as sharp.

---

\(^{1}\) While Bayesian statistics is generally disinclined to use the data in any way to determine priors, the primary objective here was to find an estimate of the scale of the sill. Given the lack of alternatives, the authors consider that it is reasonable to use the data for this purpose.

\(^{**}\) Bivariate distributions were produced for ease of visualization. In each case, one physical parameter of the three considered has been integrated out, along with all of the discrepancy parameters (\(\xi\)) and the observation error parameter (\(\psi\)). In other words, each plot is a bivariate marginal distribution.
Comparing the posterior distributions in Fig. 6(a) and (b) with the prior in Fig. 4(a) along with tabulated values in Table 1 shows that the data – along with the effect of other prior distributions operating through the model – tend to indicate a value for the enthalpy of the reaction that is on the high side of the ab initio estimates. The most accurate methods are seen to be PBE0 and MP3, in accordance with the studies on small molecules representative of the binding in DNA base pairs referenced above.

Upon further investigation, it becomes clear that one crucial aspect of the evident constriction in uncertainty upon the introduction of informative prior distributions is the establishment of a maximum for the entropy of the reaction at approximately 200 J mol\(^{-1}\) K\(^{-1}\), as reflected in Fig. 4(b). This cutoff clearly shows that large portions of the distributions – including some areas of highest probability density – shown in Fig. 5(a) and (c) are not physically feasible. There is also constriction at the higher enthalpies when going over from the uniform to informative case, indicating the influence of the distribution in Fig. 4(a).

Expectation values for the posterior distribution with uniform priors are \(\Delta H = -84.49\) kJ mol\(^{-1}\), \(\Delta S = -214.4\) J mol\(^{-1}\) K\(^{-1}\), and \(n_v = 1973.5\) mol m\(^{-3}\). For the posterior distribution with ab initio priors, expectation values are \(\Delta H = -90.63\) kJ mol\(^{-1}\), \(\Delta S = -229.2\) J mol\(^{-1}\) K\(^{-1}\), and \(n_v = 1914.0\) mol m\(^{-3}\).

**5.3.3 Predictions and discrepancy.** Predictions of the model-plus-conditioned discrepancy, along with 95% confidence bounds, for the analysis with the informative priors are found in Fig. 7(a). It is interesting to note the high correspondence between the mean line and the experimental data in light of Fig. 7(b), which shows 95% confidence bounds for the model alone derived from the posterior distribution. Here the influence of the discrepancy is made plain, as the model alone poorly represents the data at 4% and 100% CO\(_2\). The model-plus-discrepancy approach, by contrast, predicts the experimental data to within the estimated observation error, with growing uncertainty in between data points (interpolation uncertainty) and outside of the measured temperature range (extrapolation uncertainty), in accordance with the overall size of the discrepancy.

As mentioned above, whereas the model form discrepancy itself – as reflected in eqn (5) – does not depend on the model parameters, the conditioned discrepancy does depend on the model parameters indirectly, inasmuch as each parameter set drawn from the posterior gives rise to a different “observed” discrepancy with respect to the data. The most informative and useful picture of the discrepancy is one that takes the posterior distribution into account in this way. This is the purpose of the “integrated square discrepancy” described in Section 4.3: it is an absolute gauge, as a function of the inputs (temperature and CO\(_2\) partial pressure in the present case), of how well the model performs in the analysis.

Here it is again useful to look at cases with and without informative priors. Integrated square discrepancies for 4, 18.5 and 100% CO\(_2\) are shown in Fig. 8 and 9 for uniform and informative priors, respectively. For both cases, the discrepancy is higher at the compositional extremes – 4 and 100% CO\(_2\) – than it is for moderate CO\(_2\) compositions. This broadly indicates the existence of phenomena such as multi-site adsorption and adsorbate interactions. The compositional extremes – especially 100% CO\(_2\) – also show the greatest differences between the uniform and informative cases, with larger discrepancies arising in the latter. The expectation value for the integrated square discrepancy across all temperatures and pressures is 0.2391 wt%\(^3\) for the uniform case and 0.2785 wt%\(^3\) for the informative case. When compared with the decrease in entropy of the posterior evident upon introduction of informative priors, this increase in the average model discrepancy indicates that the use of informative priors effects a shift of uncertainty from the model parameters into the model form discrepancy.

**5.4 Discussion**

**5.4.1 Establishment of priors.** The reduction in the model parameter uncertainty upon introduction of the informative priors derived from ab initio calculations is the most striking result of this study. Given such a potentially strong influence of prior distributions on calibration results, the method for establishing prior distributions is deserving of careful consideration. This is an often-ambiguous aspect of Bayesian analysis; inescapably, there will be a certain amount of recourse to qualitative – and sometimes subjective – judgment. The best
practice is to give a clear and open account of how priors were derived, while erring on the side of greater uncertainty.††

In this study, important qualitative decisions have been made in the choice of chemistries to examine theoretically, the \textit{ab initio} methods to employ and the relative weighting of them, and the size of the basis sets used in the calculation. The rationales for these choices are as follows. The chemistries were chosen based on their similarity to the ethyleneimines found in the experimental system; a small number of test calculations on molecules as large as tetraethylenepentamine were found not to deviate significantly from those appearing in Table 1. Equal weighting of different methods as diverse as B3LYP, PBE0 and MP3 can be justified based on the principle of erring on the side of greater uncertainty. The use of small basis sets in the calculation reflects a choice about the allocation of resources: many calculations done with relatively small basis sets was considered to be superior to a small number of calculations with larger basis sets. One might reasonably expect that a larger basis set applied to all calculations would result in a tighter prior distribution, but there is no guarantee of this. In any event, the estimation of error in quantum chemical calculations due to the various approximations employed in those calculations is an emerging area of research in its own right.63–65

5.4.2 The potential usefulness of parameter distributions and model form discrepancy in upscaling and extrapolation.

This work demonstrates how the uncertainty inherent in \textit{ab initio} calculations can be quantified and corrected using experimental data in the context of a chemical equilibrium model. The use of parameter distributions as opposed to computationally-derived point estimates for model parameters has obvious advantages for multi-scale modeling.

But those advantages go beyond parameter uncertainty. Multi-scale modeling with two-way communication means -- frequently -- the use of approximate models of small-scale phenomena within the context of models of large-scale systems. The approximations at the smaller scale can strongly influence the accuracy of the large-scale model. The Bayesian analytical framework herein demonstrated, when combined with a suitable uncertainty propagation scheme, has the power to quantify the effects of such errors. In this light, the shift of uncertainty from model parameters to model discrepancy upon introduction of informative priors emerges as another important outcome of the study. The discrepancy and the posterior distribution respond differently to interpolation and extrapolation of model predictions inherent in model upscaling: the posterior remains unchanged, whereas the uncertainty associated with the discrepancy increases as the system moves further away from regions in the model input space where there is experimental data. Uncertainty propagation using a “model-plus-discrepancy” approach66 therefore stands the best chance of providing a useful metric of uncertainty in a multi-scale model.

6 Conclusions

This work demonstrates how quantum chemistry and experiments can be combined in a Bayesian statistical framework to produce best-available estimates of model parameters, while quantitatively measuring the shortcomings of the associated

†† The latter principle can sometimes be enforced through the use of hyperparameters. Hyperparameters are parameters used in the specification of prior distributions that are assigned bounds instead of fixed values. They are then estimated -- that is, added to the domain of the posterior -- then integrated out.

Fig. 8 Integrated square discrepancy for the case of uniform priors: (a) 4%, (b) 18.5% and (c) 100% CO\textsubscript{2}.

Fig. 9 Integrated square discrepancy for the case of informative priors according to Fig. 4: (a) 4%, (b) 18.5% and (c) 100% CO\textsubscript{2}.
model as a function of model inputs. Posterior estimates of the adsorption enthalpy for the formation of carbamic acid in mesoporous silica-supported ethyleneamines were shown to be on the high side of prior estimates arising from a number of quantum chemical approaches. The inclusion of ab initio data in the calibration helped to differentiate uncertainty in the parameters from uncertainty in the model itself.

Acknowledgements

The authors acknowledge useful discussions with Neil Henson.

Partial support for David Mebane and Andrew Lee were provided through an ORISE postdoctoral fellowship.

Funding for this work was provided by the Department of Energy through the Carbon Capture Simulation Initiative.

This report 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