

Journal of the American Statistical Association

ISSN: 0162-1459 (Print) 1537-274X (Online) Journal homepage: https://www.tandfonline.com/loi/uasa20

Upscaling Uncertainty with Dynamic Discrepancy for a Multi-Scale Carbon Capture System

K. Sham Bhat, David S. Mebane, Priyadarshi Mahapatra & Curtis B. Storlie

To cite this article: K. Sham Bhat, David S. Mebane, Priyadarshi Mahapatra & Curtis B. Storlie (2017) Upscaling Uncertainty with Dynamic Discrepancy for a Multi-Scale Carbon Capture System, Journal of the American Statistical Association, 112:520, 1453-1467, DOI: 10.1080/01621459.2017.1295863

To link to this article: https://doi.org/10.1080/01621459.2017.1295863

	١
-	

Journal of the American Statistical Association

8

View supplementary material 🕝



Published online: 26 Jan 2018.

C C

Submit your article to this journal 🗹

|--|

Article views: 452



View related articles 🗹



View Crossmark data 🗹



Citing articles: 9 View citing articles

Upscaling Uncertainty with Dynamic Discrepancy for a Multi-Scale Carbon Capture System

K. Sham Bhat^a, David S. Mebane^b, Priyadarshi Mahapatra^c, and Curtis B. Storlie^d

^aLos Alamos National Laboratory, Los Alamos, NM; ^bDepartment of Mechanical and Aerospace Engineering, West Virginia University, Morgantown, WV; ^cNational Energy Technology Laboratory, Morgantown, WV; ^dMayo Clinic, Rochester, MN

ABSTRACT

Uncertainties from model parameters and model discrepancy from small-scale models impact the accuracy and reliability of predictions of large-scale systems. Inadequate representation of these uncertainties may result in inaccurate and overconfident predictions during scale-up to larger systems. Hence, multiscale modeling efforts must accurately quantify the effect of the propagation of uncertainties during upscaling. Using a Bayesian approach, we calibrate a small-scale solid sorbent model to thermogravimetric (TGA) data on a functional profile using chemistry-based priors. Crucial to this effort is the representation of model discrepancy, which uses a Bayesian smoothing splines (BSS-ANOVA) framework. Our uncertainty quantification (UQ) approach could be considered intrusive as it includes the discrepancy function within the chemical rate expressions; resulting in a set of stochastic differential equations. Such an approach allows for easily propagating uncertainty by propagating the joint model parameter and discrepancy posterior into the larger-scale system of rate expressions. The broad UQ framework presented here could be applicable to virtually all areas of science where multiscale modeling is used. Supplementary materials for this article are available online.

1. Introduction

The Carbon Capture Simulation Initiative (CCSI) sponsored by the U.S. Department of Energy is focused on accelerating the adoption of new carbon capture technology using modeling and simulation to reduce the amount of physical testing required for development of larger-scale power plants (Miller et al. 2014). This effort relies heavily on computer models for upscaling. These computer models are used to analyze phenomena in carbon capture systems from the quantum scale up through to the industrial macroscale.

The multi-scale uncertainty quantification effort in this article is illustrated on a simple carbon capture process for a "bubbling fluidized bed" adsorber (Lee and Miller 2012), which is built using Aspen Custom Modeler (Aspen 2011). The major driver of the uncertainty in the system is the chemical sorbent model, characterized by one (or more) chemical reactions, dependent on several chemical parameters that describe the equilibrium and kinetic facets of the reactions. In addition, there are certain system conditions (or physical inputs, e.g., temperature and pressure) that affect the behavior of the system. The small-scale model takes in temperature and pressure inputs (describing possible system conditions) and a set of chemistry model parameters and then outputs the sorbent weight gain. Experimental data are collected using thermogravimetric analysis (TGA, e.g., see Mebane et al. 2013, 2014) with the same inputs and outputs. Figure 1 displays the output (sorbent weight gain)

of the small-scale sorbent model as a function one of the inputs (temperature). An important characteristic here is that both the inputs and outputs of the model are functional (in time). Both the small-scale sorbent model, isolated to only a chemical reaction with no fluid dynamics (small scale), and the fully coupled model for a process system (large scale) are governed by the solution of one or more rate-based differential equations.

Computer models such as that in Figure 1 are widely used by scientists and engineers to understand and predict the behavior of complex physical processes (e.g., climate change, nuclear reactor performance, fluid transport, and carbon capture systems) when direct experimentation is difficult, expensive, or impossible. While these computer models are often grounded in scientific theory, they will still have various forms of uncertainty that must be treated appropriately when they are used to predict the behavior of the physical process. These uncertainties may come from many sources, for example, incomplete information about physical constants and/or inadequate quantitative models to describe the physical behavior. Hence, uncertainty quantification (UQ) is widely recognized as essential to analysis of complex computer models (Currin et al. 1991; Kennedy and O'Hagan 2001; Oakley and O'Hagan 2004; Higdon et al. 2008; Storlie et al. 2009). Many of these uncertainties arise in a multi-scale system context, and uncertainties at a small scale may greatly impact the accuracy and reliability of large system

ARTICLE HISTORY

Received January 2015 Revised November 2016

KEYWORDS

Bayesian hierarchical modeling; BSS-ANOVA; Computer model calibration; Extrapolation; Functional data; Propagation of uncertainty



CONTACT K. Sham Bhat Bhat (Subscription) bhat

Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/r/JASA.



Figure 1. Functional input (temperature) in green, functional output (sorbent weight gain) in blue.

predictions. Therefore, it is crucial that this small-scale uncertainty be appropriately represented.

Two major sources of uncertainty in predictions from computer models are parameter uncertainty and model form discrepancy; statistical calibration of the model to experimental (or field) data provides a means to formally quantify these uncertainties. Kennedy and O'Hagan (2001) developed a framework for calibration of computer models, which includes a model discrepancy term describing the deviation from the model to reality. Failure to properly account for model discrepancy can lead to overfitting of the model parameters (Bayarri et al. 2007b), and hence inaccurate predictions. Constructing the model form discrepancy using expert scientific knowledge about how the model does not to conform to reality will improve inference and resulting model predictions (Brynjarsdóttir and O'Hagan 2014) as well as alleviate confounding issues between the parameters and discrepancy (Liu, Bayarri, and Berger 2009). Furthermore, the discrepancy provides a representation of how the model is flawed, which could be useful for model improvement. In this article, we propose an efficient, rigorous approach to propagate forward the uncertainty in model form discrepancy, along with parameter uncertainty, resulting from model calibration at small scale, when making predictions at large scale.

Upscaling uncertainty for a multi-scale system via forward propagation of both the model parameter and discrepancy sources of uncertainty presents a major challenge. Many current approaches for forward uncertainty propagation to a new system (e.g., large scale) do not fully propagate model discrepancy (Miki et al. 2012; McDonnell et al. 2015; Hadjidoukas et al. 2015; Engel et al. 2016). This may be problematic since the calibrated model output, even at the "true" parameters, may not be an accurate representation of the system as it ignores the effect of model form discrepancy. Strong and Oakley (2014) discussed a framework for determining the sensitivity to submodel discrepancies, and Callahan (2016) discussed accounting for model discrepancies between fine and coarse scale models. One approach to propagate model discrepancy is to use a fully nonintrusive "black-box" approach for calibration, that is, obtain a joint distribution of the model parameters and discrepancy, and propagate that distribution to the large-scale system (Bhat et al. 2012). However, this approach is not always feasible in practice. In addition, the estimated model discrepancy at small scale may not be directly applicable to that at large scale. For instance, often the inputs and even the quantity of interest are not comparable from the calibration data to the desired system prediction.

We propose a novel approach to describe the dynamic discrepancy between the small-scale model and reality, which has a clear scientific interpretation about the model shortcomings and is efficient for both calibration and upscaling. Our approach uses an intrusive uncertainty quantification (UQ) approach by including the stochastic discrepancy function within the smallscale sorbent model equations, resulting in a stochastic differential equation(s) (SDE). In addition, the discrepancy terms may be interpreted as a "correction" for a specific deficiency in the small-scale sorbent model (e.g., deviation from the ideal gas assumption). Model form discrepancy here is represented using a Gaussian process with a Bayesian smoothing spline (BSS)-ANOVA covariance. The BSS-ANOVA framework has many advantages within this context; it provides an approximate parametric form that is very convenient for both calibration and upscaling, accounts for the uncertainty due to extrapolation while upscaling, and provides substantial computational gains requiring only $\mathcal{O}(N)$ computational time by eschewing matrix inversions.

Using a Bayesian approach, the small-scale sorbent model is calibrated to data; resulting in a joint sample-based distribution of both model parameters and discrepancy basis function coefficients. The uncertainty in model parameters *and* discrepancy are then easily propagated to scale by passing these posterior samples into the large-scale system of rate expressions and solving them in a Monte Carlo fashion. The output from the large-scale system provides both a quantity of interest (such as the proportion of CO_2 captured) as well as a functional profile of the system conditions (which may also be seen as inputs to the small-scale sorbent model). There currently is not an uncertainty quantification approach that incorporates a stochastic discrepancy in a calibration framework with functional inputs and forward propagates both parameter and model form uncertainty across scales in an efficient manner.

The rest of this article is organized as follows. In Section 2, a broad overview of the proposed framework to upscale uncertainty with dynamic discrepancy is provided, while the details of the proposed calibration, dynamic discrepancy, and upscaling approach are in Section 3. In Section 4, the methodology is then used to provide a UQ analysis on the problem that motivated its design; a bubbling fluidized bed carbon capture system with a chemical kinetics sorbent model. Finally, a discussion, caveats, and avenues for future research are provided in Section 5.

2. UQ for a Multi-Scale System

This section contains a broad overview of our approach to upscaling uncertainty in a multi-scale system. The goal is to make upscaled predictions (with uncertainty) in a large-scale system, with (potentially) many small-scale physical processes embedded in it; where the behavior of both the large-scale system and the small-scale processes are simulated with deterministic computer models. In this article, a single small-scale model is used for the ease of demonstration. A big picture overview of the coupled model for a simple carbon capture system is shown in Figure 2. The uncertainty in this system is largely driven by the uncertainty in the adsorption of CO_2 ; however direct calibration of this large-scale system to learn about this uncertainty is often not feasible due to a dearth of relevant data



Upscaling Uncertainty to Carbon Capture System

Figure 2. Diagram for upscaling uncertainty for a carbon capture system.

and/or computational challenges. However, there is experimental data from the laboratory at a set of system conditions that can be used to constrain the small-scale sorbent model (which describes the adsorption of CO_2 in the same manner in the carbon capture system). Hence, the small-scale sorbent model (with the dynamic discrepancy embedded in the rate equation as described in Section 3) can be calibrated to this data, and the uncertainties (both parametric and model discrepancy) may then be propagated in a Monte Carlo fashion through the carbon capture system.

The details of the small-scale sorbent model and the largescale carbon capture process system are discussed in Sections 2.1 and 2.2, respectively. In Section 2.3, an overview of the proposed Bayesian dynamic discrepancy approach used for UQ analysis is provided, which has several advantages in this particular framework; incorporation of functional inputs, convenience in propagating information across scales, and the inclusion of a dynamic discrepancy function to describe shortcomings of the model.

2.1. Basic Chemical Kinetics CO₂ Adsorption Model

The basic chemistry model describes the adsorption of CO_2 by a solid Amine-based sorbent. It was determined in the course of an ab initio study documented by Mebane et al. (2013) that this adsorption of CO_2 takes place through the formation of carbamic acid according to the reaction:

$$2R_2NH + CO_2(g) \rightleftharpoons R_2NCOOH : R_2NH$$
(1)

A summary of the inputs and outputs may be found in Table 1. There are five model parameters to be estimated: $\theta = [\Delta H, \Delta S, \Delta H^{\ddagger}, \gamma, n_v]$. The equilibrium parameters are ΔH , ΔS , and n_v and the kinetic parameters are ΔH^{\ddagger} and γ . The governing equations of this sorbent model are,

$$\frac{\partial x}{\partial t} = \kappa^{K} [(1 - 2x)^{2} p - x^{2} / \kappa^{E}], \quad x(0) = x_{0},$$

$$w = \tilde{M} n_{v} x / \rho \qquad (2)$$

$$\kappa^{E} = \exp \left(\Delta S / R\right) \exp \left(-\Delta H / RT\right) / P$$

$$\kappa^{K} = \gamma T \exp \left(-\Delta H^{\ddagger} / RT\right). \qquad (3)$$

Table 1. Summary of inputs, outputs, and small-scale sorbent parameters.

The rate equation in (2) is solved on a temporal grid; potential system conditions, temperature (*T*), and partial pressure (*p*) of CO₂, are functional inputs over time (see Figure 1 for an example temperature input), resulting in a functional response w(t) (the weight increase of the sorbent). The response w(t) is a multiple of the chemical state *x*, or the fraction of amine sites occupied by carbamic acid. The initial condition $x(0) = x_0$, where x_0 is often assumed to be 0 (no carbamic acid).

Additional constants within the model are (i) M: the molar weight of CO₂, (ii) ρ : the sorbent density, (iii) R: the ideal gas constant, and (iv) P: the total pressure. The equilibrium constant κ^{E} (which informs on how much of the reaction in (1) has completed) is a function of ΔS and ΔH (equilibrium parameters), while the reaction rate constant κ^{K} (speed of reaction) is a function of γ and ΔH^{\ddagger} (kinetic parameters). As can be seen from Figure 1, as temperature decreases there is an increase in sorbent uptake of CO₂ and a resulting increase in the weight of the sorbent. Similarly, an increase in partial pressure would result in an increase in sorbent uptake of CO₂ and in the weight of the sorbent. The experimental apparatus used to obtain the data here requires a constant partial pressure over time, so the functional pressure input is a constant line. However, several time series observations are collected at different (constant over time) partial pressures.

2.2. Overview of Large-Scale Carbon Capture Process Model

The coupled large-scale system is introduced here with a single quantity of interest x, which in this article is the

chemical state. However, this framework may be easily extended to the multivariate case. The system conditions are $\zeta(t) = [\zeta_1(t), \ldots, \zeta_q(t)]$, where $\zeta_i(t)$ represents the curve for the *i*th condition at time *t*. The function space for these system conditions is very large as it must include any combination of physically feasible curves for all system conditions. A general rate equation describing the small-scale sorbent model is provided below in (4) for convenience, where *x* represents a state variable and $f_s(x, \zeta(t); \theta)$ is the chemical rate function:

$$\frac{\partial x}{\partial t} = f_s(x, \boldsymbol{\zeta}(t); \boldsymbol{\theta}). \tag{4}$$

The large-scale system may be generally expressed as a set of differential equations below in (5)–(6); these equations are solved on a one-dimensional spatial grid (i.e., the vertical location in the carbon capture system, denoted by the variable z). For direct comparison with the inputs and outputs to the smallscale sorbent model, it is convenient to represent these functional curves as functions of time (as opposed to z), which is accomplished numerically using an Eulerian–Langrangian conversion along the flow-velocity field (see supplementary material for details). While the equations for this system will be expressed in terms of z for reasons of scientific rigor, in practice the system is run in the time domain for this article, and upscaled results will be presented as functions of t. Note that the system conditions ζ (e.g., temperature and pressure) are *outputs* for the large scale, but are *inputs* for the small-scale model in (4).

$$\frac{\partial x}{\partial z} = f_{\ell}(x, \boldsymbol{\zeta}(z); \boldsymbol{\theta})$$
(5)

$$g_1(x,\boldsymbol{\zeta}(z);\boldsymbol{\theta}) = 0 \tag{6}$$

$$g_q(x, \boldsymbol{\zeta}(z); \boldsymbol{\theta}) = 0.$$
(7)

The outputs from the large-scale system are functional curves of the quantity of interest x and the system conditions ζ (as a function of z), other possible inputs to this adsorber such as inlet temperature and gas flow rate are fixed. Note here that Equation (5) represents the same CO₂ adsorption process as the sorbent model (as described in Equation (4)) in the carbon capture system.

:

2.3. Framework for Upscaling Uncertainty

A road map for the upscaling methodology is provided here. As shown in the top half of Figure 3, the small-scale model is calibrated to experimental data (see Section 3). The small-scale model inputs and outputs are functional in nature and experimental data observations are usually only available at a handful of different profiles. Thus, the functional input space will typically be only sparsely covered by the experimental design.

A Bayesian calibration approach following the Kennedy– O'Hagan framework (Kennedy and O'Hagan 2001) is employed to calibrate the small-scale model to data (see Section 3). In this approach, prior distributions (using domain expertise and previous results whenever possible) are placed on the uncertain model parameters and a prior distribution is also placed on the discrepancy function (usually a Gaussian Process (GP)).



Figure 3. Overview of the upscaling process.

However, it is infeasible to use the conventional GP approach to model discrepancy for forward propagation, in this case, due to the functional inputs and an entirely different fully coupled system at large scale. Therefore, a novel dynamic discrepancy approach is introduced in Section 3.2 where the discrepancy is embedded within the rate equations as in (4) of the small-scale model.

The end result from calibration is a joint sample-based distribution representing the model parameters and the discrepancy, which is then propagated to the large-scale system, as seen in the bottom half of Figure 3. For each sample, the large-scale system model simultaneously solves for the rate equations, for example, (5)-(6), describing the system response and system conditions at each step, resulting in functional curves for both the system conditions and response. The collection of these solutions (from the sample-based distribution) provides a set of probabilistic predictions of the system conditions and response, while incorporating uncertainty. Aggregated output quantities of interest (such as % CO₂ captured by the system) may then be derived from the response, resulting in probabilistic predictions, characterizing the uncertainty of the output quantities of interest. These predictions are made for the CO₂ capture system described in Sections 4.3 and 4.4. In making upscaled predictions, the system conditions experienced at scale may be far different than those from the small-scale experiments. Thus, it may become important to limit the amount this extrapolation to the extent possible.

3. Calibration with Dynamic Discrepancy

3.1. Computer Model Calibration for Functional Data

Computer model calibration is often used to constrain the computer model to be consistent with experimental data. Hence, the primary goal of calibration is to find a set of model parameter values that best reproduce the reality of experimental (or field) data. In the traditional computer model calibration (i.e., inverse problem) setup (Kennedy and O'Hagan 2001), an output *y* from the physical system is observed (with observational error) at several (*N*) locations of a "controllable" vector of inputs $\boldsymbol{\zeta} = [\zeta_1, \ldots, \zeta_q]$. This physical reality can be approximated by a simulator (i.e., a computer model), $\eta(\boldsymbol{\zeta}, \boldsymbol{\theta})$, where $\boldsymbol{\theta} = [\theta_1, \ldots, \theta_P]$ is a vector of model parameters. If fixed at an appropriate (unknown) value of $\boldsymbol{\theta} = \boldsymbol{\theta}^*$, then $\eta(\boldsymbol{\zeta}, \boldsymbol{\theta})$ will approximate the reality at $\boldsymbol{\zeta}$. The framework also includes a model form discrepancy function δ that admits the possibility of model bias (from reality). Therefore, a general model for the experimental data is

$$Y_n = \rho \eta(\boldsymbol{\zeta}_n, \boldsymbol{\theta}^*) + \delta(\boldsymbol{\zeta}_n) + \epsilon_n, \qquad (8)$$

n = 1, ..., N, where Y_n, ξ_n, ϵ_n are the experimental data, inputs, and observation error for the *n*th observation, respectively, and ρ is a regression parameter, representing multiplicative bias. Due to identifiability issues, $\rho = 1$ will be assumed here. The goal is to estimate θ^* (which is just denoted by θ from now on for ease of notation) and possibly the discrepancy function δ . This is typically done within a Bayesian framework (Higdon et al. 2004), where a prior distribution is placed on θ and δ and then updated by conditioning on the experimental data. A Gaussian process (GP) prior is often selected for the model discrepancy δ (Kennedy and O'Hagan 2001), which will be discussed in more detail in the next section.

For the work in this article, both the inputs ζ_n and the output Y_n are functional in nature. The domain of the functional input/output space consists of a single variable, time, in this case, but the methodology to be discussed is easily generalizable to a higher dimensional domain. Thus, the *n*th functional output curve(s) is expressed as $Y_n(t)$, $t \in [0, T]$, and the inputs are written as $\zeta_n(t) = [\zeta_{n,1}(t), \ldots, \zeta_{n,q}(t)]'$, that is, $\zeta_{n,i}(t)$ represents the input curve at time *t* for the *i*th input of the *n*th observation. In practice, the functional input/output for the *n*th observation is recorded at a dense grid of discretized time points, $\mathbf{t}_n = [t_{n,1}, \ldots, t_{n,M_n}]'$. The simulator output evaluated at model parameters $\boldsymbol{\theta}$, inputs ζ_n , and time *t* can be expressed as $\eta(t; \zeta_n, \boldsymbol{\theta})$. The data model for the calibration framework is then

$$Y_n(t) = \eta(t; \boldsymbol{\zeta}_n, \boldsymbol{\theta}) + \delta(t; \boldsymbol{\zeta}_n) + \epsilon_n(t), \quad t \in [0, T], \quad (9)$$

where the observation error $\epsilon(t)$ is assumed to be a white-noise process with variance σ^2 . In the model to follow, independence is assumed a priori between δ , η , and ϵ . Several others have considered the problem of computer model calibration in the presence of functional output (Bayarri et al. 2007a; Higdon et al. 2008). However, the problem considered here is distinctly different due the functional input. In general, $\eta(t; \zeta_n, \theta)$ and $\delta(t; \zeta_n)$ are dependent on the *entire* input curves ζ_n , not just the inputs at time *t*.

3.2. Dynamic Discrepancy

Model discrepancy is considered here, not merely for better model parameter inference, but also in the context of propagation of uncertainty due to model shortfall in a large-scale system. For this effort, the discrepancy must incorporate scientific understanding of the deficiencies of the model as well as flexibility to be applicable for a wide range of functional responses and account for extrapolation. Due to the functional nature of the inputs, $\delta(t; \zeta)$ in (9) is a function of the entire input curve ζ .

The immediate approach would be to construct an appropriate discrepancy function, calibrate the simulator to experimental data, and upscale the joint posterior distribution $\pi(\theta, \delta)$ to the large-scale system. Bhat et al. (2012) developed such an approach; a small-scale sorbent model was calibrated to data using the framework in (9). The model discrepancy represents deficiencies in the sorbent model for both equilibrium and kinetic behavior. The development of this discrepancy assumed a concurrent functional model (see Ramsay and Silverman 2005, pp. 280–293), that is, $\delta(t; \boldsymbol{\zeta})$ depends only on $\boldsymbol{\zeta}(t)$, which overly simplifies the functional nature of the input/output. The joint posterior $\pi(\theta, \delta)$ was then upscaled by differentiating the posterior realizations of δ w.r.t time, and including them in the rate equations of the large-scale process model in (5)–(6). However, the differentiated posterior discrepancy realizations were very noisy and required heavy, ad hoc smoothing to avoid solver failures.

We now present a novel alternative approach that alleviates the issues mentioned above. The main idea is to include the discrepancy δ within the rate equation of the small-scale model in (4), that is,

$$\frac{\partial x}{\partial t} = f_s(x, \theta, \zeta(t)) + \delta(x, \zeta(t); \beta), \qquad (10)$$

where β denotes the hyperparameters, that is, variance and range parameter(s), of the traditional GP. In (10), δ is referred to as a *dynamic* discrepancy since it allows the dynamic system to change its path depending on the value of δ . The main advantage of this approach is that unlike the output x(t), the rate $\frac{\partial x}{\partial t}$ must be a function of the forcing inputs at only time *t*, that is, $\zeta(t)$, and not dependent on the entire curve ζ . Thus, implementing the discrepancy in the rate equation allows us to justify the use of a concurrent functional model, and sidestep the complications of dependence on the entire functional input.

If a GP prior is placed on δ , then (10) is a stochastic differential equation (SDE). If δ were a traditional GP then, even for a fixed $\boldsymbol{\theta}$ and $\boldsymbol{\beta}$, (10) would result in an SDE. To avoid these complications, a BSS-ANOVA GP (discussed in more detail below in Section 3.3) is used as prior on δ , which among other things, has the advantage of admitting a convenient, approximate parametric form, thus containing its entire stochasticity in its parameters β . In other words, when β is fixed, δ is entirely specified, and the SDE in (10) becomes an ODE. In a Bayesian calibration framework, the SDE can then be easily integrated within the Markov chain Monte Carlo (MCMC) routine as follows: at each MCMC iteration (i) propose a set of model and discrepancy parameters, (ii) obtain a solution of the state variable(s) from the ODE, (iii) evaluate the likelihood, and (iv) accept/reject the sample. This framework provides an avenue to estimate the joint posterior distribution $\pi(\theta, \delta)$ (or equivalently $\pi(\theta, \beta)$) of model parameters and discrepancy, and allows for the forward propagation of uncertainty in the usual sample-based manner. Furthermore, such an approach is generalizable to a broad class of problems. We note that there have been previous efforts to



Figure 4. First eight eigenfunctions for a main effect from the BSS-ANOVA GP.

incorporate model "lack of fit" in a similar manner as in Equation (10) using a B-spline basis for δ (Varziri, McAuley, and McLellan 2008; Hooker 2009), however they were more appropriate for an optimization context instead of upscaling uncertainty.

3.3. BSS-ANOVA Model

The discrepancy function δ is formulated using the BSS-ANOVA GP model (Reich, Storlie, and Bondell 2009; Storlie et al. 2013); using a covariance function that directly uses the functional components from a functional ANOVA decomposition (Gu 2002). The BSS-ANOVA approach has been used previously to model the emulator and discrepancy in the traditional computer model calibration problem (Storlie et al. 2014). This approach has two very useful properties for modeling dynamic discrepancy: (i) it provides a convenient parametric form, which allows the SDE in (10) to be represented as an ODE for each MCMC iteration, providing for easy calibration and uncertainty propagation, and (ii) improves computational efficiency substantially, scaling *linearly* with the number of data points, as opposed to $\mathcal{O}(N^3)$ for a traditional GP. In this section, only the details of BSS-ANOVA model necessary to formulate the discrepancy are provided; further details may be found in the supplementary material and Storlie et al. (2014).

For ease of notation in defining the discrepancy function, let the inputs to the discrepancy function in (10) at a given time *t* be denoted as $\boldsymbol{\omega} = [x(t), \zeta_1(t), \dots, \zeta_q(t)]'$. The discrepancy may then be represented as

$$\delta(\boldsymbol{\omega}) = \beta_0 + \sum_{j=1}^{q+1} \delta_j(\omega_j) + \sum_{j$$

where $\beta_0 \sim N(0, \zeta_0^2)$ and each of the main effect functions δ_j and the two-way interaction functions $\delta_{j,k}$ are mean 0 GPs with the BSS-ANOVA covariance function described by Reich, Storlie, and Bondell (2009). Three-way or higher order interaction functional components can be included as well. Under the BSS-ANOVA construction, the resulting component GPs are such that they will satisfy the functional ANOVA constraints, for example, $\int \delta_r(u) du = 0$ and $\int \delta_{r,r'}(u, v) du = 0$, almost surely. Any realization from this GP also lies in first-order Sobolev space, that is, absolutely continuous with derivative in L_2 .

It was further demonstrated by Storlie et al. (2012) that each functional component in (11) can be further written as an orthogonal basis expansion, for example,

$$\delta_{j}(\omega_{j}) = \sum_{l=1}^{\infty} \beta_{j,l} \phi_{l}(\omega_{j}), \ \beta_{j,l} \stackrel{\text{iid}}{\sim} \mathcal{N}(0, \tau_{j}^{2}).$$
(12)

The ϕ_l terms in the expansion are the eigenfunctions (scaled by the eigenvalues) in the Karhunen–Loéve (KL) expansion (Berlinet and Thomas-Agnan 2004, pp. 65–70). The ϕ_l get increasingly higher frequency and have decreasingly less magnitude as l increases (as depicted in Figure 4), so the expansion in (12) can be truncated at some value L. The choice of L is not critical, as the model will be nearly identical for different L provided it is large enough; our experience suggests that $L \ge 25$ is sufficient for most problems.

The same decomposition in (12) is used for two-way and higher interactions as well. In fact, the ϕ_l for two-way interactions are simply pairwise products of the corresponding main effect basis functions and similarly for three-way and higher interactions. In many problems, it is sufficient to include only main effects and two-way interactions. A few preselected threeway interactions, selected in consultation with domain experts, are included in the model as well for the analysis presented below. Hence, the overall model in (11) can be written in general as

$$\delta(\boldsymbol{\omega}) = \sum_{m=1}^{M} \sum_{l=1}^{L_m} \beta_{m,l} \phi_{m,l}(\boldsymbol{\omega}), \qquad (13)$$

$$\beta_{m,l} \stackrel{\text{ind}}{\sim} \mathcal{N}(0, \tau_m^2),$$
 (14)

where (i) *m* indexes over the *M* functional components included in the discrepancy model, and (ii) *l* indexes over the number of basis functions L_m used for the *m*th functional component of the discrepancy representation. The $\beta_{m,l}$, $\phi_{m,l}$, and τ_m would correspond to a particular term in the expansion of (12) for the *m*th functional component. More specific details of the decomposition of the BSS-ANOVA GP into the linear model in (13) are provided in the supplementary material.

3.4. Dynamic Discrepancy for Sorbent Model

While the additive dynamic discrepancy approach detailed in Section 3.2 can be useful in many situations, there were a large number of solver failures due to either a lack of convergence or solutions outside the physical bounds ($x \in [0, 0.5]$) for the model in this article. Further, the manner in which the chemical model could be deficient was deemed by the modelers (as discussed further below) to be more appropriately modeled with a multiplicative discrepancy for the equilibrium and kinetic processes. This approach, thus, has two advantages; (i) improved convergence of the solver and guarantee physical solutions when the discrepancy is added, and (ii) the multiplicative discrepancy has a clear physical interpretation for sorbent models.

We first derive the model form discrepancy for the equilibrium process. The rate equation and the equilibrium and kinetic constants κ and k for an ideal reaction (which assumes no interaction energy for adsorbates and equivalence among all adsorption sites) are described in (2). Under the assumption of thermodynamic ideality, ΔH and ΔS , the enthalpy and entropy, respectively, are constant with respect to the thermodynamic state space $\zeta = \{p, T\}$. However, when this assumption violated, one frequent way of representing this nonideality is by allowing ΔH and ΔS to depend on the state space ζ , that is, (2) becomes

$$\frac{\partial x}{\partial t} = \kappa_{\text{new}}^{K} \left[(1 - 2x)^2 p - x^2 / \kappa_{\text{new}}^{E} \right], \tag{15}$$

where

$$\kappa_{\text{new}}^{E} = \exp\left(\frac{\Delta S}{R}\right) \exp\left(\frac{-\Delta H}{RT} + \delta^{E}(p, T)\right)$$
$$= \exp\left(\frac{-\Delta S}{R}\right) \exp\left(\frac{-\Delta H}{RT}\right) \exp[\delta^{E}(p, T)]$$
$$= \kappa^{E} \exp[\delta^{E}(p, T)]. \tag{16}$$

Equation (16) shows an equilibrium constant κ_{new}^E that allows for deviations from an ideal reaction, where $\delta^E(p, T)$ is a stochastic function that represents this discrepancy. It is clear from this formulation that for any realization of $\delta^E(p, T)$, the chemical state $x \in [0, 0.5]$. The same idea is applicable for the reaction rate constant, that is,

$$\kappa_{\text{new}}^{K} = \gamma T \exp\left(\frac{-\Delta H^{\ddagger}}{RT} + \delta^{K}(x, p, T)\right)$$
$$= \kappa^{K} \exp[\delta^{K}(x, p, T)]. \tag{17}$$

The kinetic rate discrepancy δ^k in (17) is a function of p, T, and x, since the rate of reaction via κ_{new}^K will necessarily depend on the current value of x as opposed to the equilibrium constant κ_{new}^E , which only depends on p and T.

The two discrepancy functions are expressed using a Gaussian process with a BSS-ANOVA covariance function using (12) as follows,

$$\delta^{E}(p,T) = \sum_{m=1}^{M^{E}} \sum_{l=1}^{L_{m}^{E}} \beta_{m,l}^{E} \phi_{m,l}^{E}(p,T), \quad \delta^{K}(x,p,T)$$
$$= \sum_{m=1}^{M^{K}} \sum_{l=1}^{L_{m}^{K}} \beta_{m,l}^{K} \phi_{m,l}^{K}(x,p,T).$$
(18)

The BSS-ANOVA formulation in (11) includes a constant term β_0 , which is completely confounded with the model parameters ΔS and γ (for δ^E and δ^K); thus β_0 is set to zero. The modified rate equation for the toy sorbent model, with the discrepancies δ^E and δ^K embedded is then,

$$\frac{\partial x}{\partial t} = f_s^*(x, \boldsymbol{\zeta}(t); \boldsymbol{\theta}, \boldsymbol{\beta}) = \kappa^E(\boldsymbol{\theta}) \exp[\delta^K(x, p, T)]((1 - 2x)^2 p - x^2 / [\kappa^K(\boldsymbol{\theta}) \exp[\delta^E(p, T)]]).$$
(19)

0

3.5. Calibration at Small Scale and Upscaling Uncertainty

Here, the Bayesian approach to infer a joint posterior probability distribution of the model parameters and discrepancy parameters of the small-scale model is described. Let $\mathbf{y} = [\mathbf{y}_1^T, \dots, \mathbf{y}_N^T]'$ be the (stacked) vector of experimental observations where \mathbf{y}_n is the output for a functional input $\boldsymbol{\zeta}_n$ (over the time points $\mathbf{t}_n = [t_{n,1}, \dots, t_{n,M_n}]$). Also let $\boldsymbol{\delta} = [\boldsymbol{\delta}^E, \boldsymbol{\delta}^K]$ denote both discrepancy functions. Define the stacked vectors of model output with discrepancy at $\boldsymbol{\theta}$ to be $\boldsymbol{\eta}(\boldsymbol{\theta}, \boldsymbol{\delta}) =$ $[\boldsymbol{\eta}(\boldsymbol{\theta}, \boldsymbol{\delta}, \boldsymbol{\zeta}_1), \dots, \boldsymbol{\eta}(\boldsymbol{\theta}, \boldsymbol{\delta}, \boldsymbol{\zeta}_N)]$, and the observation error variance $\boldsymbol{\epsilon} = [\boldsymbol{\epsilon}_1, \dots, \boldsymbol{\epsilon}_N]' \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_{N^*})$; where $N^* = \sum_{i=1}^N M_i$.

Incorporating the dynamic discrepancy in Section 3.4, the model for **y** is

$$\mathbf{y} = \boldsymbol{\eta}(\boldsymbol{\theta}, \delta) + \boldsymbol{\epsilon}. \tag{20}$$

The goal is to estimate $\{\theta, \delta, \sigma^2\}$ given y. Prior distributions are required for θ , δ , and σ^2 to complete the Bayesian model specification. Priors for some of the model parameters θ are derived from ab initio quantum chemistry calculations and from prior scientific studies or expert judgment. The model for δ in (18) requires a prior specification for τ_j 's from (14). A diffuse inverse Gamma prior is chosen for both the τ_j 's and σ^2 , which results in conjugate updates for these parameters within MCMC procedure. Further discussion regarding the prior specification of θ is provided in Section 4.

Let the variance parameters for the discrepancies be denoted as $\boldsymbol{\tau} = [\boldsymbol{\tau}_1^E, \dots, \boldsymbol{\tau}_{M^E}^E, \boldsymbol{\tau}_1^K, \dots, \boldsymbol{\tau}_{M^K}^K]$ and all of the basis function coefficients $\boldsymbol{\beta} = [\boldsymbol{\beta}^E, \boldsymbol{\beta}^K] = [\boldsymbol{\beta}_1^E, \dots, \boldsymbol{\beta}_{M^E}^E, \boldsymbol{\beta}_1^K, \dots, \boldsymbol{\beta}_{M^K}^K]$, where $\boldsymbol{\beta}_m^E$ and $\boldsymbol{\beta}_m^K$ are the vectors of the regression parameters for the *m*th functional component of the respective discrepancy. The posterior distribution,

$$\pi(\boldsymbol{\theta}, \boldsymbol{\beta}, \boldsymbol{\tau}, \sigma^2 \mid \mathbf{y}) \propto \mathcal{L}(\mathbf{y} \mid \boldsymbol{\theta}, \boldsymbol{\beta}, \sigma^2) p(\boldsymbol{\beta} \mid \boldsymbol{\tau}) p(\boldsymbol{\theta}) p(\boldsymbol{\tau}) p(\sigma^2)$$

is obtained through simulation using Markov chain Monte Carlo (MCMC). The MCMC routine is a hybrid sampling scheme of Gibbs and Metropolis–Hastings (MH) updates, where Gibbs updates are viable for τ and σ^2 with appropriate conjugate priors and MH updates are necessary for θ and β . The updates for β are performed by updating each component separately, for example, all the coefficients for a given main effect (or interaction) component are updated simultaneously. That is, $\beta_m = [\beta_{m,1}, \ldots, \beta_{m,L_m}]$ is updated using a multivariate normal proposal for each $m = 1, \ldots, M^E + M^K$. Block updating with joint proposals rather than updating each parameter individually improves mixing and reduces the number of sorbent model evaluations necessary (each proposal requires a model evaluation), which in turn reduces computational time for the MCMC procedure.

The result of the MCMC is a sample-based distribution of $\pi(\theta, \beta | \mathbf{y})$. A sample of size *V* can then be drawn from the posterior distribution of (θ, β) for the purpose of upscaling to the large-scale system. Let $\theta^{(v)}$, $\beta^{(v)}$ be the *v*th value of this sample and denote by $\delta(\beta^{(v)})$ the discrepancy realization for *v*th sample. Each sample is then propagated forward to the large-scale system model, which involves solving a set of differential equations shown in (21)–(23), where the dynamic discrepancy $\delta(\beta)$ has been embedded into the rate expression $\frac{\partial x}{\partial z}$, as in (19) (subject to a space to time conversion),

$$\frac{\partial x}{\partial z} = f_s^*(x, \boldsymbol{\zeta}(z); \boldsymbol{\theta}, \boldsymbol{\beta})$$
(21)

$$g_1(x,\boldsymbol{\zeta}(z);\boldsymbol{\theta}) = 0 \tag{22}$$

$$\vdots$$

$$g_q(x, \boldsymbol{\zeta}(z); \boldsymbol{\theta}) = 0. \tag{23}$$

The solution to this set of equations for the *v*th sample $(\theta^{(v)}, \beta^{(v)})$ is $\mathbf{x}^{(v)}$ and $\boldsymbol{\zeta}^{(v)}$, which are the system response (i.e., chemical state) and system conditions (i.e., temperature and partial pressure), respectively, and are presented as a function of time *t* (Equations (21)–(23) can be converted to be solved on a time grid, see Section 2.2). The final output of the upscaling is a sample-based distribution, $\{(\boldsymbol{\zeta}^{(1)}, \mathbf{x}^{(1)}), \ldots, (\boldsymbol{\zeta}^{(V)}, \mathbf{x}^{(V)})\}$, that may be used to approximate the distribution of certain quantities of interest, for example, CO₂ capture fraction.

4. Application to a Simple Carbon Capture System

In this section, we apply our methodology on a simple process model, the "bubbling fluidized bed" adsorber (Lee and Miller 2012), driven by a single small-scale chemical sorbent model described in Section 2.1. A post-combustion CO₂ capture system consists of two parts; the adsorber that takes up CO₂ and the regenerator that lifts CO₂ off the sorbent and passes it along for sequestration. This application focuses on just the bubbling bed adsorber. The methodology is first illustrated on a truth known example so that its performance can be assessed. The small-scale model is calibrated to a "reality" function, which is actually a more complicated sorbent adsorption process with two chemical reactions with the same inputs and outputs as the single reaction sorbent model. The calibration results are then upscaled to the process model by propagating the uncertainty forward. The reality function is also upscaled to provide the upscaled "truth" to assess the performance of the proposed approach. Note that the system conditions (temperature and pressure curves over time) are outputs from the process model, but are inputs for the small-scale sorbent model. Finally, the entire analysis is then applied to actual TGA data in Section 4.4. This section only provides a brief overview of the "reality" function and the carbon capture process; more information about this process model is available in the supplementary materials.

4.1. Synthetic Data Generation and Process Model Description

In the synthetic data example, the "reality" function plus iid Gaussian noise is used as a proxy for the experimental data. This exercise is used here to illustrate the proposed methodology without data complications and provide a validation of the upscaled results. However, this is not a "perfect model experiment;" parameters of this reality function and the sorbent model are not directly comparable. This function is based on a two-step adsorption process of CO_2 by the sorbent according to the reactions below. This is a more complicated process than the sorbent model discussed in Section 2.1 in that the single reaction in (1) is in "reality" two separate reactions:

$$CO_2 + R_2NH - \longrightarrow R_2NH^+ - COO^-$$
 (24)

$$R_2NH^+ - COO^- + R_2NH \rightleftharpoons R_2NCOO^- : R_2NH_2^+. (25)$$

There are nine parameters that need to be specified; ΔH_x , ΔS_x , ΔH_z , ΔS_z are enthalpies and entropies for the two reactions, ΔH_x^{\ddagger} , ΔH_z^{\ddagger} , γ_x , and γ_z are the activation energies and preexponential factors for the two reaction, and n_v is the number of active amine sites per unit volume of sorbent; $\boldsymbol{\theta} = [\Delta H_x, \Delta S_x, \Delta H_x^{\ddagger}, \gamma_x, n_v, \Delta H_z, \Delta S_z, \Delta H_z^{\ddagger}, \gamma_z]$. The rate equations of this two-step adsorption model are,

$$\frac{\partial x_z}{\partial t} = \kappa_z^K (sp - x_z/\kappa_z^E) - \kappa_x^K (sx_z - x_c^2/\kappa_x^E)
\frac{\partial x_c}{\partial t} = \kappa_x^K (sx_z - x_c^2/\kappa_x^E), x_z(0) = x_{z0}, x_c(0) = x_{c0}
s = 1 - 2x_c - x_z, \quad w = \tilde{M}n_v(x_c + x_z)/\rho
\kappa_x^E = \exp\left(\frac{\Delta S_x}{R}\right) \exp\left(-\frac{\Delta H_x}{RT}\right),
\kappa_z^E = \exp\left(\frac{\Delta S_z}{R}\right) \exp\left(-\frac{\Delta H_z}{RT}\right)/P
\kappa_x^K = \gamma_x \exp\left(-\frac{\Delta H_x^{\ddagger}}{RT}\right), \kappa_z^K = \gamma_z \exp\left(-\frac{\Delta H_z^{\ddagger}}{RT}\right). \quad (26)$$

The rate equations in (26) are solved on a temporal grid, resulting in a functional response w(t) (the sorbent weight gain) with temperature (*T*) and partial pressure (*p*) of CO₂ as functional inputs over time (see Figure 1 for an example temperature input). The sorbent weight gain w(t) is a multiple of the sum of x_c and x_z , or the fraction of amine sites occupied by carbamic acid and zwitterions, respectively, \tilde{M} is the molar weight of CO₂, and ρ is the sorbent density, *R* is the ideal gas constant, and *P* is the total pressure; the latter four are constant within the model. The equilibrium constants for the two reactions are κ_x^E and κ_z^E , and the reaction rate constants for the two reactions are κ_x^K and κ_z^K . The initial conditions are $x_{c0} = x_{z0} = 0$.

For a given set of functional input profiles for partial pressure and temperature, along with model parameters θ , the reality model solves for output curve w(t). The reality model parameters used here are

$$\boldsymbol{\theta} = [-88671, -67.056, 35148, 141.22, 2000, -32055, -87, 53594, 25657].$$

These values were selected to ensure favorable convergence and system behavior properties when upscaled to the carbon capture process. Upscaling the reality function consists of using this reality function to represent the chemistry in the carbon capture process (i.e., replacing (5) with (26) in Section 2.2).

Since the synthetic data selected should be consistent with the behavior of the experimental apparatus, input profiles with



Figure 5. Temperature input profile for synthetic data.

constant partial pressure and declining temperature over time are selected. In particular, five sets of different inputs, with partial pressures 1, 4, 7.5, 10, and 18.5% CO₂, which are consistent with partial pressures in an adsorber process. All five inputs have a temperature profile shown in Figure 5 where *t* goes from 1 to 60 sec. A small amount of white noise (with standard deviation of 10^{-4}) is added to mimic the effect of observation error to the output of the "reality function." The output from the five input profiles are stacked together to create the synthetic data vector, resulting in a total of N = 305 data points (see Figure 7). The small-scale sorbent model was also run at these same five input profiles in the calibration process.

The upscaled physical process, modeled in Aspen Custom Modeler (Aspen 2011), consists of a single device: a onedimensional, three region "bubbling fluidized-bed" adsorber with internal heat exchangers, through which the sorbent and a mixture of CO₂ and N₂ flow in a co-current configuration; both the sorbent and the gas get injected at the bottom and flow upward. The model predicts the hydrodynamics of the bed and provides axial profiles for all temperature, concentration, and velocities as given by Lee and Miller (2012). Adsorption of CO₂ by the sorbent produces heat in accordance with the heat of reaction, which is equivalent to the adsorption enthalpy ΔH . This heat is removed by a heat exchanger that runs along the length of the adsorber, thereby regulating the temperature within the bed. The full process model is presented in the supplementary material, along with all of the (fixed) process model parameters.

4.2. Implementation for Calibration and Upscaling

The calibration approach discussed in Section 3 is now applied to the synthetic data generated in Section 4.1 using (20). This calibration approach requires embedding the discrepancy functions within the model solver as described in (19), which use (16)–(17). The BSS-ANOVA basis functions are not analytical; they are numerically evaluated on a dense grid and represented as continuous functions using linear interpolation. This choice of interpolant enables the function itself as well as its derivatives to be calculated at any point on its domain. A Crank–Nicolson scheme is used to discretize the ordinary differential equation in (19), meaning that the system can be solved given the input functions *p* and *T* along with the parameter sets θ and β using Newton's method.

As discussed in Section 3.5, MCMC methods are deployed to obtain the posterior distributions of θ , β , and σ^2 . For three model parameters, ΔH , ΔS , and n_v , priors were derived using previous scientific studies; ab initio calculations from quantum chemistry calculations were used to derive the following priors,

 $\Delta H \sim N(-60.84, 125)$ KJ/mol, $\Delta S \sim N(-250, 625)$ J/mol-K, truncated at -200 J/mol-K, $n_v \sim N(1469, 86362)$ mol/m³.

More information about the derivation of the priors may be found in Mebane et al. (2013). An empirical approach using a sensitivity study on model convergence was used to obtain priors for the kinetic parameters $\Delta H^{\ddagger} \sim \text{Unif}(-150, -50)$ kJ/mol, and $\gamma \sim \text{Unif}(0, 5)$. To improve mixing and account for the likely dependency between certain pairs of parameters suggested by the rate equations, we use an adaptive Metropolis proposal (Haario, Saksman, and Tamminen 2001) to update both θ and each β_m .

As discussed in Section 3.5, conjugate prior distributions are selected for τ and σ^2 to ensure Gibbs sampling for these parameters during the MCMC procedure. A sufficiently diffuse inverse gamma parameters was selected for the elements of τ , specifically $\tau_j \sim IG(0.5, 30)$ to allow for adequate flexibility and promote mixing of β . An inverse gamma prior for the observation error parameter is specified as $\psi \sim IG(1, 10^{-8})$.

The MCMC was run for 200,000 iterations, allowing for a burn-in of 50,000 samples; the results were confirmed by running two separate chains which gave similar results. Convergence was also confirmed by small Monte Carlo standard errors using batch means (Flegal, Haran, and Jones 2008). The computer code for the MCMC was implemented in MATLAB using a 2.66 GHz 6-Core Intel Xeon on a Mac Pro desktop with 16GB of RAM. Obtaining the 200,000 samples using MCMC required approximately 190 hr of computer time, the overwhelming majority of which was required to execute the sorbent model. The sorbent model and "reality" function were implemented in C++, with a MATLAB executable file; each run of the sorbent model (with the discrepancy embedded) approximately took 6 sec for each execution. The carbon capture process model is augmented with the discrepancy functions within Aspen Custom Modeler (ACM) following (21)-(23) to facilitate the upscaling of the calibration results. Multiple posterior samples are simulated simultaneously by exploiting the multiple parallel ACM run capability offered by CCSI Turbine Gateway developed as part of CCSI project at Lawrence Berkeley National Laboratory (LBNL) facility.

4.3. Synthetic Data Results

This section presents results from the calibration of the smallscale sorbent model to the reality function described in Section 4.1, and the subsequent upscaling of the calibration results to the simple carbon capture process system. The posterior distribution of the model parameters is displayed via bivariate distributions in Figure 6. The mean and 95% credible regions for the sorbent model parameters (see Table 2) were

Table 2. Posterior mean and 95% credible regions for sorbent model parameters.

Parameter	Mean	95 % Lower bound	95 % Upper bound
ΔH	— 96.271 KJ/mol	- 107.330	- 87.989
ΔS	— 241.59 J/mol	- 263.49	- 219.95
ΔH^{\ddagger}	67.307 KJ/mol	62.966	71.695
γ	2.54	1.89	3.19
n _v	2203.3 mol/m ³	2150.6	2253.6

calculated using the highest posterior density (HPD) method (Chen, Shao, and Ibrahim 2000). There are strong correlations between certain pairs of model parameters in the posterior distribution. The correlation between the equilibrium parameters ΔH and ΔS is 0.82, suggesting a very strong relationship between the equilibrium enthalpy and entropy, which seems



Figure 6. Bivariate marginal posterior distributions for the sorbent model parameters on the off-diagonals, univariate marginal posterior distributions on the diagonals. The parameters are in displayed in the following order: ΔH , ΔS , ΔH^{\ddagger} , γ , n_{p} .

consistent with the equilibrium analysis of this TGA dataset by Mebane et al. (2013). In addition, there is a correlation of 0.91 between the kinetic parameters ΔH^{\ddagger} and γ . There are no "true" parameters here, as the parameters of the reality model are usually not directly comparable with those of the sorbent model.

Posterior predictions and 95% bounds obtained from the full calibration approach with the dynamic discrepancy are shown in Figure 7. These bands were created by obtaining predictions and point-wise bounds at a grid of 61 time locations for each of the five CO₂ composition ratios (1%, 4%, 7.5%, 10%, 18.5%). Thirty randomly chosen posterior prediction curves are also provided in each case. It is clear that when the dynamic discrepancy is included in making predictions, the data are well represented by the predictions. The discrepancy due to equilibrium and kinetic effects, δ^E and δ^K , respectively, are functions of their inputs (p(t), T(t), and x(t)). However, the posterior realizations of δ^E and δ^K given the functional inputs used for calibration (temperature input T(t) in Figure 5 and constant p(t) corresponding to the CO₂ composition ratio) may be expressed as a function of time and suggest a significant nonzero discrepancy that increases over time (see Figures 8 and 9).

The results of upscaling the uncertainty from the model parameters and discrepancy to the carbon capture system, hence obtaining a distributions of the capture fraction and the system conditions are presented below. One thousand posterior samples are propagated from the joint distribution of the model parameters and discrepancy to the carbon capture process, as well as implementing the reality function into the carbon capture process. The 95% credible region for the carbon capture fraction is between 0.79 and 0.90, which clearly covers the "reality" capture rate of 0.85 (see Figure 10). The carbon capture system conditions for both temperature and partial pressure are also displayed in Figure 10; the distribution of the upscaled model results for both system conditions cover the reality system conditions. There appears to be a sizeable uncertainty in the prediction of the upscaled condition for pressure, however, this does not translate into large uncertainty in the primary quantity of interest (i.e., capture fraction).



Figure 7. Posterior fitted plots for sorbent weight gain for all five CO₂ partial pressures. The data (black open circles), reality (black line), and 30 posterior realizations (green lines), calibrated model predictions without discrepancy (blue lines) and 95% credible bands (dashed red lines) are provided for model plus discrepancy predictions.



Figure 8. Posterior equilibrium model discrepancy realizations (posterior mean (blue) along with 20 realizations (green) and 95% credible bands (red)); zero line (dotted black line). Plots correspond to T(t) in Figure 5 and a constant p(t) determined by CO₂ composition.



Figure 9. Posterior kinetic model discrepancy realizations (posterior mean (blue) along with 20 realizations (green) and 95% credible bands (red)); zero line (dotted black line). Plots correspond to T(t) in Figure 5 and a constant p(t) determined by CO₂ composition.



Figure 10. Results from upscaling posterior distributions obtained from calibrating (to the reality function) small-scale sorbent model to the large-scale carbon capture system. Distribution of carbon capture rate (left) with reality capture rate (red dot), along with distributions of system conditions; temperature (middle); and partial pressure (right). Reality function system conditions (black line) and 60 realizations of system conditions from model (red lines) are shown.



Figure 11. Left: entire output from the TGA experiment at 18.5% CO₂ (blue line) and temperature (green line), red lines denote boundaries of the selected snippet. Right: Data from the selected snippet, note that the response shows substantial change over time.

4.4. Analysis and Results for Application to Thermogravimetric (TGA) Data

In this section, the calibration and upscaling with dynamic discrepancy approach in this article is applied to real experimental data obtained from thermogravimetric analysis (TGA) to predict the capture fraction of CO_2 in a large-scale process model. TGA was used to obtain the experimental data to constrain the model; these experiments were conducted at the following ratios, 4%, 7.5%, 10%, 18.5%, and 100% CO_2 (vs. N₂), each has an associated partial pressure that is constant over time. The inputs to the TGA experiment consist of that constant partial pressure and a functional temperature curve. More information about the mechanics about the TGA experiment may be found in Mebane et al. (2013). Since the entire TGA response curve for any particular experiment has a domain of up to 200,000 sec and consists largely of plateau regions (see Figure 11), which yield little information about the kinetics; snippets of the response to large temperature changes are analyzed instead. In particular, 12 snippets have been carefully chosen from the TGA experiments with partial pressure less than 100%.

Posterior predictions and 95% bounds were computed at the 12 previously selected snippets from the TGA experimental data. When the dynamic discrepancy is included in the framework, the data appear to be largely covered by the 95% intervals (see Figure 12 for two snippets). The results of upscaling the uncertainty from the model parameters and discrepancy to



Figure 12. Posterior fitted plots for snippets of TGA data. The data (black open circles), reality (black line), and 30 posterior realizations (green lines), and 95% credible bands (dashed red lines) are provided for model plus discrepancy predictions.



Figure 13. Results from upscaling posterior distributions obtained from calibrating (to the TGA data) small-scale sorbent model to the large-scale carbon capture system. Distribution of carbon capture rate (left) along with distributions of system conditions; temperature (middle); and partial pressure (right). Realizations of system conditions from model (red lines) are shown.



Figure 14. Plot of input combinations of (*p*, *T*) and (*x*, *p*, *T*) that were observed in the experimental data for constraining the small-scale model (blue) and in the final system conditions in the large scale (red) to determine if predictions at scale are an extrapolation. Some of the system conditions appear to be at slightly higher temperatures than the experimental data; so extrapolation may be a concern.

the carbon capture system are presented here. One hundred posterior samples are propagated from the joint distribution of the model parameters and discrepancy to the carbon capture process. It is noted that 18 of these samples failed to converge during upscaling, resulting in 82 samples. It is also possible that the numerical discretization issues may have contributed to the some of these failures (Chkrebtii et al. 2016), although changing the discretization step is difficult in practice here and there are no known subgrid processes that are resolved at a finer resolution.

The mean carbon capture fraction (of the converged samples) is 0.68, and the 95% credible region for the carbon capture fraction is between 0.32 and 0.85 (see Figure 13). The carbon capture system input conditions for both temperature and partial pressure are also described in Figure 13. Certain system design features (such as flow rate) for the carbon capture are treated as fixed here. However, the prediction of performance with uncertainty could easily be explored over the system design space by incorporating these design parameters into the sample of upscaled predictions. A full system design optimization under the presence of uncertainty is a natural next step, but beyond the scope of this article.

5. Conclusions and Future Directions

A novel approach for upscaling uncertainty for a multi-scale system with functional inputs has been presented and demonstrated here. The small-scale model is calibrated to data and then the results of the calibration, a joint distribution of both parameter and dynamic model form discrepancy, are propagated to the large-scale system. The model form uncertainty is represented by a dynamic discrepancy embedded within the rate equation(s) in the small-scale model, and has a clear physical understanding of the deficiency of the model. The dynamic discrepancy is modeled with a BSS-ANOVA framework, which provides a convenient form for calibration and upscaling, accounts for extrapolation uncertainty and has linear complexity with the number of data points. The methodology was demonstrated for a simple carbon capture system driven by a small-scale chemical sorbent model.

One major issue with the proposed approach is the computational time that is required to compute the sorbent model, which is the bottleneck in obtaining the posterior distribution for the model parameters and discrepancy. Each evaluation of the sorbent model requires approximately 6 sec for the synthetic data, and each iteration of MCMC requires 15 model evaluations, requiring \sim 1 week to obtain the full posterior. Both the speed of the sorbent model and the number of model evaluations are a function of the number of inputs and discrepancy coefficients needed. When more complicated models are involved with multiple outputs, the number of differential equations to be solved increases, which in turn increases the number of inputs and discrepancy functions and coefficients. In addition, the computational bottleneck requires us to make difficult choices to reduce the size of the TGA data. The end result is that the computational time required for this approach in the current implementation may be prohibitive. The computational efficiency of this approach may be improved in three ways. First, sorbent models can be implemented using graphics processing units (GPU); GPUs often speed up computer models several-fold. Also, other adaptive proposals to reduce the number of MCMC iterations required to reach convergence may be investigated. Finally, dynamic emulation may be considered as a surrogate for the sorbent model, which would reduce the computational costs of the model evaluations.

Extrapolation over the input function space during upscaling is a difficult challenge; there may be a risk of extrapolation in the upscaling predictions due to different system conditions at scale than the small-scale experimental conditions. However, our dynamic discrepancy approach described in Section 3.4 reduces the large functional input space (e.g., x(t), T(t), p(t)) to the space of scalar (x, p, T) input combinations for the rate equation. Hence, we only need to "train" the discrepancy δ_E and δ_K on the space of scalar (p, T) or (x, p, T) input combinations (obtained from the experiments), respectively; Figure 14 compares these input combinations for the upscaling of the synthetic data results. This does not entirely prevent extrapolation (as the solution of large-scale model may be still dependent on the entire curve), but it may account for a large source of it. Our proposed approach demonstrates the inclusion of the extrapolation uncertainty in the conditions and quantities of interest of the large-scale model. If the uncertainties due to extrapolation are unacceptable, one possible approach to ameliorating this uncertainty due to extrapolation is using a three-step iterative method, which is analogous to the upscaling framework suggested in this article. Data could be gathered at a set of inputs drawn from the sample-based posterior distribution of system conditions after the initial calibration was performed. This could then be followed by a second calibration of small-scale results using data collected at inputs resembling these upscaled system conditions. The second calibration results could then be used to once again upscale to the large-scale system. The entire process could be repeated if necessary, each time producing more relevant small-scale experiments.

Supplementary Materials

The supplementary material consists of three sections that required more detail than could be provided in the manuscript. They are intended to give the reader a more in-depth understanding of these concepts. Section A of the supplementary material describes the derivation of the BSS-ANOVA basis decomposition discussed in Section 3.3. Section B describes the details of the process model, the one-dimensional bubbling fluidized bed (BFB) absorber. The reaction rate equations for the kinetics are provided both for integration of the small-scale sorbent model and the "reality" function into the BFB, as well as the important device parameters and operating conditions. Section C describes the details of the numerically representing the large scale system (rate equation and constraints) using the Eulerian–Langrangian conversion along the flow-velocity field.

Acknowledgments

This work was partially supported by the Carbon Capture Simulation Initiative. The authors also thank Andrew Lee and Dan Fauth for helping to develop the sorbent model and providing us the TGA data, Brian Logdson for the extracting the snippets used in this article from the original TGA data, Joel Kress for ab initio calculations of quantum chemistry data used for priors of model parameters, Brenda Ng and others at LLNL and LBNL for helping us operate the turbine gateway to set up and run the process model, and Dan Fauth and Mac Gray for relevant figures. Funding for this work was provided by the Office of Fossil Energy, US Department of Energy through the Carbon Capture Simulation Initiative. This article 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

- Aspen (2011), Aspen Custom Modeler, Version 21.0. Cambridge, MA: Aspen Tech. Available at http://www.aspentech.com/products/aspen-custommodeler.aspx [1453,1461]
- Bayarri, M., Berger, J., Cafeo, J., Garcia-Donato, G., Liu, F., Palomo, J., Parthasarathy, R., Paulo, R., Sacks, J., and Walsh, D. (2007a), "Computer Model Validation With Functional Output," *The Annals of Statistics*, 35, 1874–1906. [1457]
- Bayarri, M., Berger, J., Higdon, D., Kennedy, M., Kottas, A., Paulo, R., Sacks, J., Cafeo, J., Cavendish, J., Lin, C., and Tu, J. (2007b), "A Framework for Validation of Computer Models," *Technometrics*, 49, 138–154. [1454]

- Berlinet, A., and Thomas-Agnan, C. (2004), Reproducing Kernel Hilbert Spaces in Probability and Statistics, Norwell, MA: Kluwer Academic Publishers. [1458]
- Bhat, K., Mebane, D., Kim, H., Eslick, J., Wendelberger, J., and Miller, D. (2012), "Upscaling Parameter Uncertainty and Model Discrepancy for a Process Model," Technical Report no. LA-UR-12-21855, Los Alamos National Laboratory, New Mexico. [1454,1457]
- Brynjarsdóttir, J., and O'Hagan, A. (2014), "Learning About Physical Parameters: The Importance of Model Discrepancy," *Inverse Problems*, 30, 114007. [1454]
- Callahan, M. D. (2016), Bayesian Parameter Estimation and Inference Across Scales, Ph.D. dissertation, Case Western Reserve University, Cleveland, OH. Available at https://etd.ohiolink.edu/ [1454]
- Chen, M., Shao, Q., and Ibrahim, J. (2000), *Monte Carlo Methods in Bayesian Computation*, New York: Springer. [1462]
- Chkrebtii, O., Campbell, D. A., Calderhead, B., and Girolami, M. A. (2016), "Bayesian Uncertainty Quantification for Differential Equations." *Bayesian Analysis*, 11, 1239–1267. [1465]
- Currin, C., Mitchell, T., Morris, M., and Ylvisaker, D. (1991), "Bayesian Prediction of Deterministic Functions, with Applications to the Design and Analysis of Computer Experiments," *Journal of the American Statistical Association*, 86, 953–963. [1453]
- Engel, D. W., Reichardt, T. A., Kulp, T. J., Graff, D. L., and Thompson, S. E. (2016), "Hierarchical Multi-Scale Approach to Validation and Uncertainty Quantification of Hyper-Spectral Image Modeling," in SPIE Defense+ Security, International Society for Optics and Photonics, pp. 98400N–98400N. [1454]
- Flegal, J., Haran, M., and Jones, G. (2008), "Markov Chain Monte Carlo: Can We Trust the Third Significant Figure?" *Statistical Science*, 23, 250–260. [1461]
- Gu, C. (2002), Smoothing Spline ANOVA Models, New York: Springer-Verlag. [1458]
- Haario, H., Saksman, E., and Tamminen, J. (2001), "An Adaptive Metropolis Algorithm," *Bernoulli*, 7, 223–242. [1461]
- Hadjidoukas, P. E., Angelikopoulos, P., Papadimitriou, C., and Koumoutsakos, P. (2015), "π4u: A High Performance Computing Framework for Bayesian Uncertainty Quantification of Complex Models," *Journal of Computational Physics*, 284, 1–21. [1454]
- Higdon, D., Gattiker, J., Williams, B., and Rightley, M. (2008), "Computer Model Calibration Using High-Dimensional Output," *Journal of the American Statistical Association*, 103, 570–583. [1453,1457]
- Higdon, D., Kennedy, M., Cavendish, J., Cafeo, J., and Ryne, R. (2004), "Combining Field Data and Computer Simulations for Calibration and Prediction," *SIAM Journal on Scientific Computing*, 26, 448–466. [1457]
- Hooker, G. (2009), "Forcing Function Diagnostics for Nonlinear Dynamics," *Biometrics*, 65, 928–936. [1458]
- Kennedy, M., and O'Hagan, A. (2001), "Bayesian Calibration of Computer Models," *Journal of the Royal Statistical Society*, Series B, 63, 425–464. [1453,1454,1456,1457]
- Lee, A., and Miller, D. C. (2012), "A One-Dimensional (1-d) Three-Region Model for a Bubbling Fluidized-Bed Adsorber," *Industrial & Engineering Chemistry Research*, 52, 469–484. [1453,1460,1461]
- Liu, F., Bayarri, M., and Berger, J. (2009), "Modularization in Bayesian Analysis, With Emphasis on Analysis of Computer Models," *Bayesian Anal*ysis, 4, 119–150. [1454]
- McDonnell, J., Schunck, N., Higdon, D., Sarich, J., Wild, S., and Nazarewicz, W. (2015), "Uncertainty Quantification for Nuclear Density Functional Theory and Information Content of New Measurements," *Physical Review Letters*, 114, 122501. [1454]
- Mebane, D., Kress, J., Storlie, C., Fauth, D., and Gray, M. (2014), "Transport, Zwitterions and the Role of Water for CO₂ Adsorption in Mesoporous Silica-Supported Amine Sorbents," *Journal of Physical Chemistry C*, 117, 26617–26627. [1453]
- Mebane, D. S., Bhat, K. S., Kress, J. D., Fauth, D. J., Gray, M. L., Lee, A., and Miller, D. C. (2013), "Bayesian Calibration of Thermodynamic Models for the Uptake of CO₂ in Supported Amine Sorbents Using ab Initio Priors," *Physical Chemistry Chemical Physics*, 15, 4355– 4366. [1453,1455,1461,1464]
- Miki, K., Panesi, M., Prudencio, E. E., and Prudhomme, S. (2012), "Probabilistic Models and Uncertainty Quantification for the Ionization

Reaction Rate of Atomic Nitrogen," *Journal of Computational Physics*, 231, 3871–3886. [1454]

- Miller, D., Syamlal, M., Mebane, D., Storlie, C., Bhattacharyya, D., Sahinidis, N., Agarwal, D., Tong, C., Zitney, S., Sarkar, A., Sun, X., Sundaresan, S., Ryan, E., Engel, D., and Dale, C. (2014), "Carbon Capture Simulation Initiative: A Case Study in Multi-Scale Modeling and New Challenges," *Annual Review of Chemical and Biomolecular Engineering*, 5, 301–323. [1453]
- Oakley, J., and O'Hagan, A. (2004), "Probabilistic Sensitivity Analysis of Complex Models: A Bayesian Approach," *Journal of the Royal Statistical Society*, Series B, 66, 751–769. [1453]
- Ramsay, J. O. and Silverman, B. W. (2005), Functional Data Analysis, New York: Springer. [1457]
- Reich, B., Storlie, C., and Bondell, H. (2009), "Variable Selection in Bayesian Smoothing Spline ANOVA Models: Application to Deterministic Computer Codes," *Technometrics*, 51, 110–120. [1458]
- Storlie, C., Fugate, M., Higdon, D., Huzurbazar, A., Francois, E., and McHugh, D. (2012), "Methods For Characterizing and Comparing Populations of Shock Wave Curves," *Technometrics*, 55, 436–449. [1458]

- Storlie, C., Lane, W., and Ryan, E. (2014), "Calibration of Computational Models With Categorical Parameters and Correlated Outputs via Bayesian Smoothing Spline ANOVA," *Journal of the American Statistical Association*, 110, 68–82. [1458]
- Storlie, C., Swiler, L., Helton, J., and Sallaberry, C. (2009), "Implementation and Evaluation of Nonparametric Regression Procedures for Sensitivity Analysis of Computationally Demanding Models," *Reliability Engineering and System Safety*, 94, 1735–1763. [1453]
- Storlie, C. B., Reich, B., Helton, J., Swiler, L., and Sallaberry, C. (2013), "Analysis of Computationally Demanding Models With Continuous and Categorical Inputs," *Reliability Engineering and System Safety*, 113, 30–41. [1458]
- Strong, M., and Oakley, J. E. (2014), "When is a Model Good Enough? Deriving the Expected Value of Model Improvement via Specifying Internal Model Discrepancies," SIAM/ASA Journal on Uncertainty Quantification, 2, 106–125. [1454]
- Varziri, M. S., McAuley, K. B., and McLellan, P. J. (2008), "Parameter Estimation in Continuous-Time Dynamic Models in the Presence of Unmeasured States and Nonstationary Disturbances," *Industrial & Engineering Chemistry Research*, 47, 380–393. [1458]