SIAM/ASA J. Uncertainty Quantification
Vol. 4, pp. 875–901



# Comparison of Statistical and Deterministic Frameworks of Uncertainty Quantification*

Michael Frenklach†, Andrew Packard†, Gonzalo Garcia-Donato‡, Rui Paulo§, and Jerome Sacks¶

**Abstract.** Two different approaches to the prediction problem are compared employing a realistic example—combustion of natural gas—with 102 uncertain parameters and 76 quantities of interests. One approach, termed bound-to-bound data collaboration (abbreviated to B2B), deploys semidefinite programming algorithms where the initial bounds on unknowns are combined with initial bounds of experimental data to produce new uncertainty bounds for the unknowns that are consistent with the data and, finally, deterministic uncertainty bounds for prediction in new settings. The other approach is statistical and Bayesian, referred to as BCP (for Bayesian calibration and prediction). It places prior distributions on the unknown parameters and on the parameters of the measurement error distributions and produces posterior distributions for model parameters and posterior distributions for model predictions in new settings. The predictions from the two approaches are consistent; a very large degree of overlap exists between B2B bounds and the support of the BCP predictive distribution. Interpretation and comparison of the results is closely connected with assumptions made about the model and experimental data and how they are used in both settings. The principal conclusion is that use of both methods protects against possible violations of assumptions in the BCP approach and conservative specifications and predictions using B2B.

**Key words.** uncertainty quantification, Bayesian analysis, bound-to-bound data collaboration, model prediction, natural-gas combustion, computer models

**AMS subject classifications.** 62F15, 62P35, 68T39

**DOI.** 10.1137/15M1019131

## 1. Introduction.

Numerical modeling of physical phenomena must accommodate sources of uncertainty rooted in the formulation of underlying physical models, their mathematical

*Received by the editors April 29, 2015; accepted for publication (in revised form) May 16, 2016; published electronically August 2, 2016.

http://www.siam.org/journals/juq/4/M101913.html

**Funding:** The first and second authors acknowledge the support of the U.S. Department of Energy, National Nuclear Security Administration, under award DE-NA0002375. The third and fourth authors were funded by Portuguese National Funds through FCT - Fundação para a Ciência e a Tecnologia, project PTDC/MAT/105349/2008. The fourth author was also partially funded by Portuguese national funds through Fundação para a Ciência e a Tecnologia, project CEMAPRE Multi/00491. The third author was also funded by the Spanish Ministry of Economy and Competitiveness under grant MTM2013-42323-P. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

†Department of Mechanical Engineering, University of California, Berkeley, CA 94720-1740 (frenklach@berkeley.edu, apackard@berkeley.edu).

‡Department of Economics and Finance, Instituto de Desarrollo Regional, Universidad de Castilla-La Mancha, Albacete, Spain (gonzalo.garciadonato@uclm.es).

§CEMAPRE and ISEG, Department of Mathematics, University of Lisbon, Lisbon 1200-781, Portugal (rui@iseg.ulisboa.pt).

¶National Institute of Statistical Sciences, Washington, D.C. 20036 (sacks@niss.org).

realization, and numerical implementation, accounting for uncertain/unknown model parameters and experimental calibration data. How to use the experimental data and the numerical realization to enable prediction of new settings and estimate unknowns has garnered much attention in recent years [4], with an important spotlight on coping with model inadequacy. Yet even when a physical model and its numerical realization is, for all purposes, an accurate representation of the phenomenon, the methods and capacity for utilizing experimental data for calibration and prediction has not yet been standardized. Such a setting is found, for instance, in combustion chemistry [12], where the physical model is a complex network of many chemical species related through hundreds and thousands of chemical reactions involving many hundreds, perhaps even thousands of unknown model parameters and mathematically modeled by a large system of ordinary differential equations (ODEs). Moreover, the limited experimental data are heterogeneous (different experimental settings, different laboratories) with sketchy assessments of measurement error.

Two different approaches to the prediction problem, using methane combustion as an example, will be discussed. One approach, developed in a series of studies [14, 15, 8, 27, 25, 26], is termed bound-to-bound data collaboration (B2B). This approach addresses uncertainty quantification (UQ) of a system such as methane combustion by reducing dimension (the number of unknown parameters), approximating ODE solutions through design of computer experiments, and specifying initial bounds on the unknown parameters and on uncertainties (errors) in measurement. Then, by deploying semidefinite programming algorithms [3], the initial bounds on unknowns are combined with the initial bound of experimental data to produce new uncertainty bounds for the unknowns that are consistent with the data and, finally, deterministic uncertainty bounds for prediction in new settings. Details of the B2B process are given in section 4.4.

The other approach is statistical and Bayesian [17, 23], referred hereafter as BCP (for Bayesian calibration and prediction), as initiated in the statistical literature in [20] and implemented in [2]. The BCP approach places prior distributions on unknown parameters, in our case the calibration parameters and the parameters of the measurement error distributions, and produces posterior distributions for the parameters and posterior distributions for model predictions in new settings. Details of BCP are given in section 4.5.

Comparison of the two approaches, B2B and BCP, in the context of the methane combustion example is the content of this paper. We find that predictions from the two approaches are consistent: the B2B predicted bounds and the BCP predictive distribution overlap greatly. Interpretation and comparison of the results is closely connected with assumptions made about the model and experimental data and how they are used in both settings. The principal conclusion is that use of both methods protects against possible violations of assumptions in the BCP approach and conservative specifications and predictions using B2B.

Much of what is done here was done initially on a smaller system of hydrogen combustion [33] but with similar conclusions; the methane problem [29] is more interesting because it is of larger dimension. Section 3 describes the methane combustion dataset; section 4 gives the details of the two approaches, B2B and BCP; results are reported in section 5; and some discussion and comments are given in section 6. We begin, however, with an oversimplified ("toy") example to familiarize the reader with some of the concepts, definitions, and our outlook on the comparison.

**2. A "toy" example.** We consider a set of two consecutive irreversible chemical reactions,

$$\text{A} \xrightarrow{k_1} \text{B} \xrightarrow{k_2} \text{C}, \tag{2.1}$$

where an initial compound A is converted to compound B, and the latter is converted into compound C. The mathematical model of this system takes the form of a set of first-order ODEs,

$$\begin{aligned} \frac{da}{dt} &= -k_1 a, \\ \frac{db}{dt} &= k_1 a - k_2 b, \\ \frac{dc}{dt} &= k_2 b, \end{aligned} \tag{2.2}$$

where $a$, $b$, and $c$ are the concentrations of compounds A, B, and C, respectively, at time $t$; $k_1$ and $k_2$ are the rate constants of the corresponding reactions, and $t$ is the reaction time. The solution of (2.2) (for $k_1 \neq k_2$) with the initial conditions $a = 1$, $b = 0$, $c = 0$ at $t = 0$ is

$$\begin{aligned} a &= e^{-k_1 t}, \\ b &= \frac{k_1}{k_2 - k_1}\left(a - e^{-k_2 t}\right), \\ c &= 1 - a - b. \end{aligned} \tag{2.3}$$

For the example, we assume the nominal values of the rate constants to be $k_1 = 2$ and $k_2 = 1$. Figure 1 shows the computed concentrations for this model at these values.

For the UQ analysis, we further assume that the rate constants are only known to within a factor of two up and down from the nominal values for each $k$; i.e., $k_1$ is assumed to have values within the interval [1, 4] and $k_2$ within [0.5, 2]. In B2B, this information is referred to as *prior knowledge* and denoted by $\mathcal{H}$; it is shown as the gray rectangle in the top left panel of Figure 2. We also assume two measured properties (quantities of interest, QoI):

- The peak value of $b$: $q_1 = b_{\max} = \frac{k_1}{k_2} e^{-k_1 t_{\max}}$; it is illustrated by a red star in Figure 1. For the purpose of this toy example, the uncertainty interval of $q_1$ is taken to be [0.45, 0.55] (10% uncertainty).
- The time of peak of $b$: $q_2 = t_{\max} = \frac{1}{k_2 - k_1} \ln \frac{k_2}{k_1}$. For the purpose of this toy example, the uncertainty interval of $q_2$ is taken to be [0.55, 0.83] (20% uncertainty).

Selecting from the prior-knowledge set of ($k_1$,$k_2$) value pairs only those that reproduce the $q_1$ and $q_2$ values within their respective uncertainty intervals results in a smaller set, which we call the *feasible set* and denote by $\mathcal{F}$; it is shown as the red area in the top left panel of Figure 2. The B2B methodology establishes the feasible set implicitly, as will be discussed in section 4.4.

The bottom right panel of Figure 2 redraws the $\mathcal{H}$ and $\mathcal{F}$ regions, but now with $\mathcal{H}$ occupying the entire box. The $\mathcal{F}$ region is covered with blue dots—these points represent a sample from the posterior distribution of the parameters obtained via a stochastic process, BCP (described in section 4.5), assuming uniform prior distributions on both $k$'s and $q$'s. The

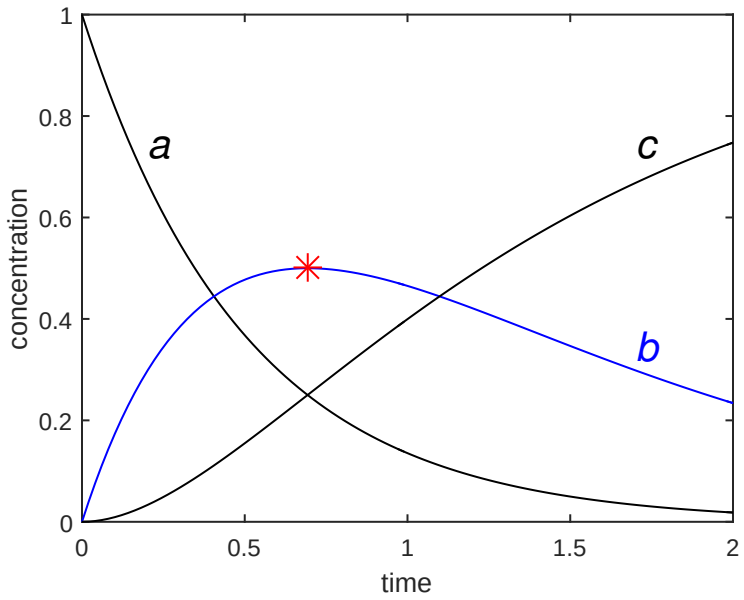


**Figure 1.** *Concentrations of A, B, and C of the toy model computed with* $k_1 = 2$ *and* $k_2 = 1$.

top right and bottom left panels of Figure 2 present the computed posterior BCP distributions along with the predicted B2B intervals for $k_1$ and $k_2$. We will use this latter format of comparison throughout the manuscript.

The use of the uniform priors for both $k$'s and $q$'s in the BCP analysis are made in order to mimic the B2B setup, and hence the agreement between the two approaches seen in Figure 2 is not surprising. However, we draw attention to the fact that B2B and BCP are different formulations. We are not trying to get them to be the same but rather to see what conclusions they reach by "harmonizing" them and embracing their differences. In addition, we repeated the calculations for the toy example, but now assuming that only $k$'s have the uniform distributions, while the $q$'s are Gaussian. The results are reported in Figure 3. We can see that the BCP sampling is now slightly extended beyond the feasible-set region of B2B, as expected. Yet the BCP posteriors, especially the marginals, are not affected much by the choice of priors.

The established feasible set, $\mathcal{F}$, "summarizes" the uncertainty of the model-data system: it is a set of all possible $(k_1, k_2)$ pairs that assures that both model parameters and model predictions of selected-for-analysis experimental observations are each within their respective uncertainty bounds. We can now explore what this information implies for prediction of an unmeasured property. Continuing with the toy example, we select an unmeasured QoI: the ratio of $c/a$ at the time when $b$ reaches the maximum, $t_{\max}$. The results are presented in Figure 4. The left panel reports the computed $(c/a)_{t_{\max}}$ values over the $\mathcal{H}$ domain, displayed as a blue surface. Those value corresponding to $(k_1,k_2)$ pairs of $\mathcal{F}$ are shown as a cyan patch of the surface. Extreme points of this patch determine the B2B predicted interval for $(c/a)_{t_{\max}}$; this interval is shown as the red bar in the right panel of Figure 4, where it is compared to

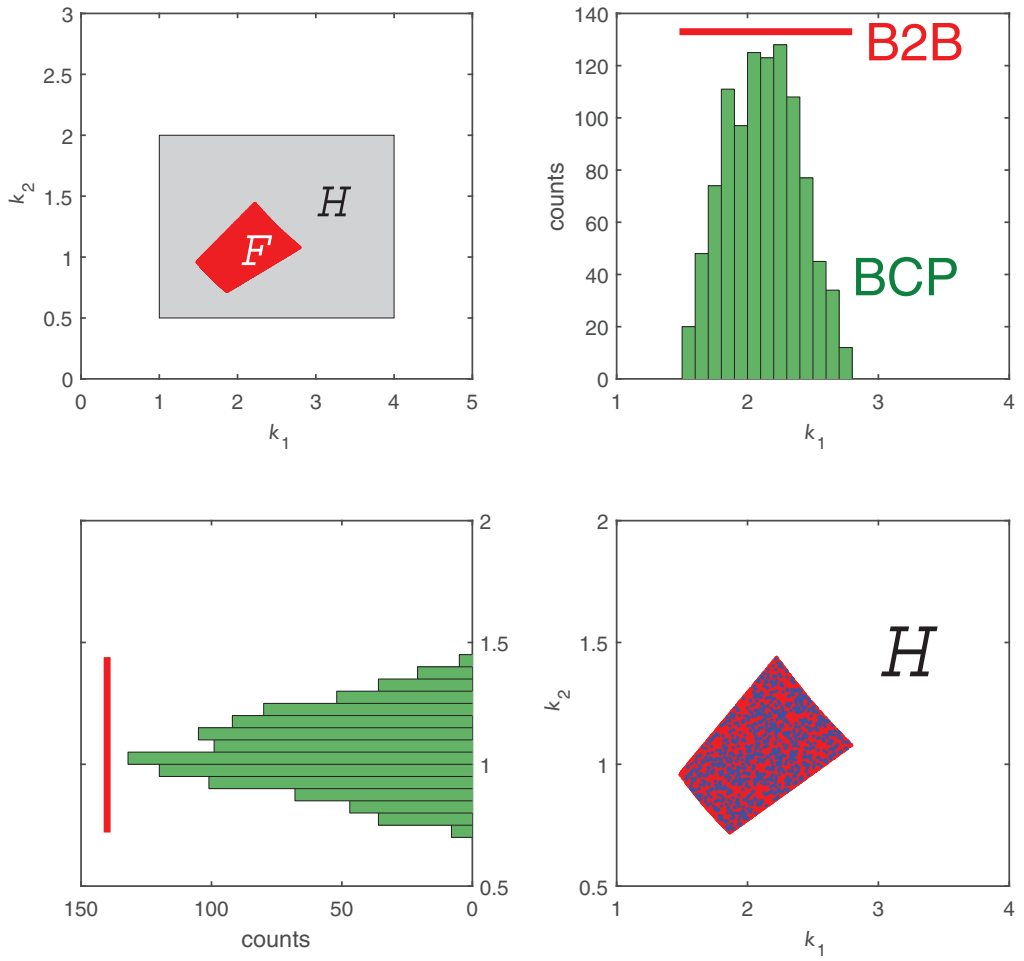


**Figure 2.** *Analysis of the toy system. Top left panel: prior-knowledge, $\mathcal{H}$, and feasible set, $\mathcal{F}$; bottom right panel: same $\mathcal{H}$ and $\mathcal{F}$ with BCP-sampled points marked as blue dots; top right and bottom left panels: posterior BCP distributions (green histograms) and predicted B2B intervals (red bars) for respective $k$'s. The BCP calculations are performed assuming uniform prior distributions on both $k$'s and $q$'s, without bias term included.*

the BCP posterior distribution.

We thus can see that in this simple example, the B2B predicted intervals and the BCP posteriors are consistent with each other (Figures 3 and 4). We remind the reader that this toy example is used merely for introductory purposes: a two-parameter problem is easy to visualize. We now turn to our main analysis platform—a realistic multiparameter, multi-QoI system.

**3. Methane-oxidation dataset.** Our main demonstration platform here is the kinetics of a complex chemical reaction system common in the fields of combustion, atmospheric

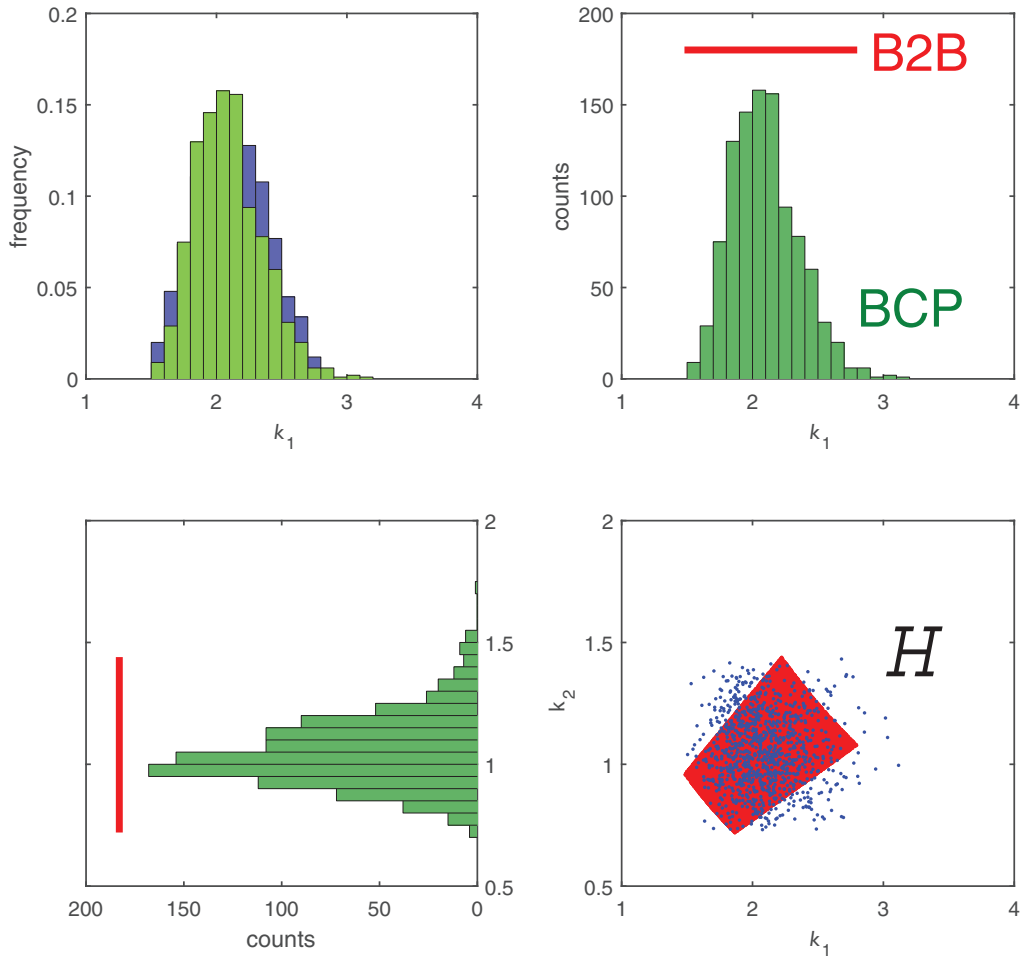


**Figure 3.** *Analysis of the toy system. Top right and bottom panels are similar to those in Figure 2 but with BCP calculations performed assuming uniform prior distributions on k's and Gaussian priors on q's. The top left panel compares the BCP posterior distributions computed for $k_1$ assuming uniform priors on both k's and q's (blue) with those assuming uniform priors on k's and the Gaussian priors on q's (green); again, no bias term is included.*

phenomena, astrophysics, material synthesis, and system biology. The specific dataset we employ here is that of natural-gas combustion [29]. The dataset consists of 76 QoIs, which are selected features from different experiments comprising species concentrations and ignition delays determined in shock tubes, peaks in species profiles and flame velocities determined in flames, shifts in observed species peak positions resulting from changing an initial mixture composition, and the like (see Appendix B). Some of these QoIs are single measurements, while others are averages of a group or series of measurements.

For the purpose of analysis, each QoI value must be accompanied by an assessed uncertainty. Regretfully, in the field of combustion, as in many fields, the reporting of experimental

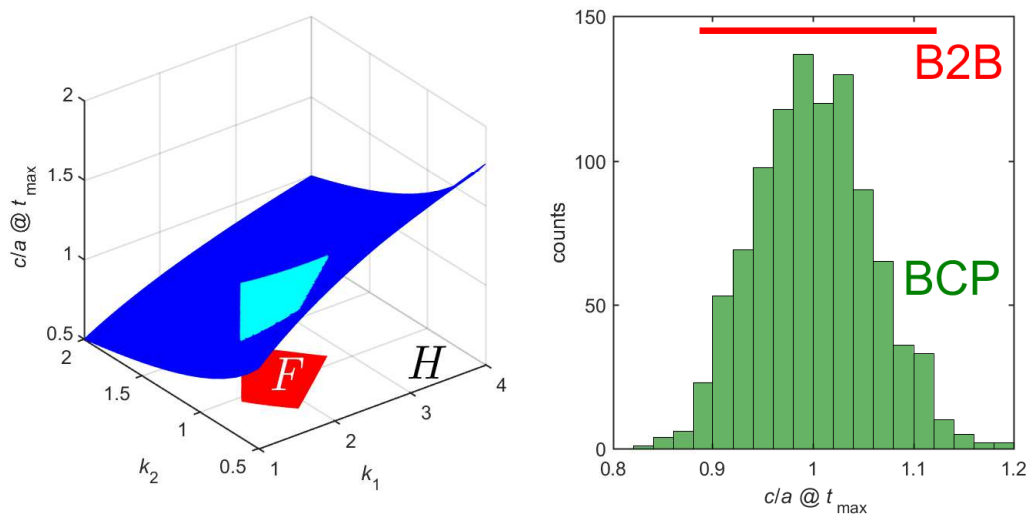


**Figure 4.** *Analysis of the toy system. Left panel: computed values of* $(c/a)_{t_{\max}}$ *over the domain of* $\mathcal{H}$ *(blue surface) and over the domain of* $\mathcal{F}$ *(cyan patch); right panel: posterior BCP distribution (green histogram) and predicted B2B interval (red bar) for* $(c/a)_{t_{\max}}$*. The BCP calculations were performed assuming uniform priors on k's and Gaussian priors on q's, without a bias term.*

(and theoretical) uncertainties is typically insufficient and often nonexistent. Some experimental studies report uncertainties for principal properties such as reaction temperature and then propagate these uncertainties to derived properties of interest, for example, species concentrations. But the details of the "error propagation" are absent. Most reports do not assess experimental errors, except for plotting the observed scatter. Moreover, combustion experiments cannot be *exactly* replicated: a shock-tube or flame experiment occurs at differing initial conditions, some of which cannot be controlled by the experimenter. For example, in a shock-tube experiment the reaction temperature created by the propagating shock front is affected by details of the diaphragm rupture, essentially an uncontrollable phenomenon.

The uncertainties in the measured values of the 76 methane-combustion QoIs were developed over years by a team of experts [29, 34]; these expert-assessed uncertainties must be taken as "tentatively entertained"—it is plausible they can be reassessed with different values.

There is an underlying physical model that is presumed to accurately reproduce each of the 76 experimental QoIs. The model is a complex network of 53 chemical species related to each other through molecular transformations, described by 325 chemical reactions. The mathematical representation of the physical model takes the form of an ODE system, similar to (2.2), but containing species concentration products that makes the system highly nonlinear without closed-form solution. The model depends on uncertain parameters, such as reaction rate constants and species enthalpies of formation. Prior uncertainty assessments of the model parameters are taken from [29, 34]. The underlying physical model is believed to be accurate, so there is no a priori model uncertainty.

**3.1. Dimension reduction: Surrogate modeling.** The methane-combustion dataset [29] was originally built for the purpose of developing an optimized combustion model, namely,

determining a set of parameter values that best fits the experimental QoI values. To speed up model optimization, the numerical solutions for individual QoIs were parameterized by fitting surrogate models in *active variables*, the latter identified by performing screening sensitivity analysis on the ODE model [13]. Each surrogate model was thus developed in its own set of 8 to 12 active variables, covering their respective ranges of uncertainty and holding the rest of the parameters at their nominal values. These parameter sets differ from QoI to QoI. Their union forms a set of 102 active variables.

Quadratic polynomials were found to be sufficiently accurate, all being within a few percent from the respective ODE results. Each quadratic polynomial expresses $\log_{10}$ of the QoI in terms of $\log_{10}$ of its set of active variables (mostly preexponential factors of the Arrhenius reaction rate expressions, but also species enthalpies of formation). Further details can be found in [12, 15, 29, 34]. The QoI surrogate models along with the model-parameter data are available as an HDF5 file in the supplementary materials (gri_mech_30.h5 [local/web 871KB]).

Other, more elaborate surrogate models can be and have been used as approximations [1, 2, 10, 5, 21]. In the present study, we take the quadratics as "truth," which simplifies computations without compromising the implications.

## 4. Methods.

**4.1. General setting.** Both approaches, BCP and B2B, share a common underpinning:

- an underlying physical process and associated numerical model with parametric dependence on unknown/uncertain physical parameters;
- prior assumptions regarding uncertainty in the true parameters;
- a collection of experimental observations with attached uncertainties;
- numerical models of each experimental observation that join the underlying physical process model with accompanying physics models to characterize measured outcomes of the experiment. (The accompanying physical models may introduce additional unknown parameters; they can be accommodated but are fixed for this study.)

The information in the four bullets constitute a *dataset*; the methane example described in section 3 is such an instance.

**4.2. Notation.** The data associated with the 76 experimental QoIs is indexed by $e$, taking values from 1 to $m = 76$. The true outcome for a QoI is $Y_e$, and its measured value is $y_e$. The active variables are denoted as $\mathbf{x} = (x_1, \ldots, x_{102})$. The treatment of measurement and parameter uncertainties is specific to each method and described below in sections 4.4 and 4.5. For QoI $e$, $M_e(\mathbf{x})$ is the output of the surrogate model when $\mathbf{x}$ is input.

**4.3. Preamble to mathematical treatment.** Although the treatments have different mathematical forms, each captures the experimental reality. Each approach combines prior information regarding the unknown parameters with the information provided by linking the experimental observations to their corresponding models, the latter denoted by $M_e(\mathbf{x})$. The result is an enriched understanding of the unknown parameters and prediction of unmeasured QoIs. One clear difference between the two methods is the mathematical form of the prior and posterior descriptions, i.e., intervals to intervals in B2B as compared to prior and posterior probability distributions in BCP. The details of these assumptions are described next, for each approach.

**4.4. Bound-to-bound data collaboration.** *Bound-to-bound data collaboration* (B2B) is an optimization-based framework for combining models and experimental data from multiple sources to explore their collective information content. The methodology tests consistency among data and models [8], explores sources of inconsistency [8], discriminates among differing models [7], makes model interval predictions [14, 15, 26, 6, 31], and analyzes sensitivity of uncertainty propagation [25]. Applications of the approach include combustion science [15, 8, 25, 12, 31] and engineering [24], atmospheric chemistry [28], system biology [7, 9, 32], and quantum chemistry [6].

In B2B, the prior information on $\mathbf{x}$ is expert-assessed uncertainties of the form $x_{i,\min} \leq x_i \leq x_{i,\max}$. This defines a *prior-knowledge* hypercube $\mathcal{H}$, normalized to $[-1, 1]$ for each $x$. The reported experimental data for experiment $e$ (experimentally observed QoI values) consist of two uncertainty bounds: $L_e$ and $U_e$. They specify a range, not merely a single value, of the measured $y_e$.

The computational models (which are parameterized by $\mathbf{x}$) must produce outputs that are consistent with the experimentally observed bounds in the experimental reports. Hence additional constraints that the true parameters must satisfy are

$$L_e \leq M_e(\mathbf{x}) \leq U_e \quad \text{for } e = 1, \ldots, m. \tag{4.1}$$

The subset of $\mathcal{H}$ satisfying (4.1) is called the *feasible set* $\mathcal{F}$ of parameters,

$$\mathcal{F} := \{\mathbf{x} \in \mathcal{H} : L_e \leq M_e(\mathbf{x}) \leq U_e \ \ \forall e\}\,. \tag{4.2}$$

This is simply all parameter values that jointly satisfy all of the prior information and are consistent with all experiment prediction models and actual observed outcomes. A parameter value that is not in $\mathcal{F}$ is at odds with at least one of these constraints.

**4.4.1. B2B as constrained optimization.** The first "bound" in the "bound-to-bound" nomenclature is associated with

- the form of the prior information, namely, that the true parameters must be both contained in the parameter hypercube $\mathcal{H}$ (which is in the form of bounds on the components); and
- the true parameters must result in model predictions of all dataset experiments that are within the measurement bounds declared by the experimenters, namely, $L_e \leq M_e(\mathbf{x}) \leq U_e$ for all $e$.

Together, these are the "bounds" that define $\mathcal{F}$. B2B invokes constrained optimization over the feasible set $\mathcal{F}$,

$$\left[\min_{\mathbf{x}\in\mathcal{F}} f(\mathbf{x}), \quad \max_{\mathbf{x}\in\mathcal{F}} f(\mathbf{x})\right], \tag{4.3}$$

where $f$ is a function of interest, and the computed min and max constitute the "to-bound" aspect of the nomenclature. In short, the bounds that describe the prior information and the bounds on experimental observations are mapped into bounds on prediction. Two common instances are described next.

**4.4.2. Dataset consistency.** The feasible set is a representation of the complete collaborative information contained in a dataset, and questions in the B2B framework are posed as optimization problems over the feasible set. This naturally raises the question of dataset consistency: is $\mathcal{F}$ nonempty? To assess it numerically, a *consistency measure* was introduced [8] that answers the question "What is the largest percentage of uncertainty reduction such that there exists a feasible parameter vector?" Associated with a given dataset D, it is denoted $C_{\mathrm{D}}$ and posed as an optimization problem,

$$
\begin{aligned}
C_{\mathrm{D}} := \max_{\gamma,\mathbf{x}\in\mathcal{H}} \gamma \ \text{ subject to } & \\
(1-\gamma)\frac{L_e-U_e}{2} &\le M_e(\mathbf{x}) - \frac{U_e+L_e}{2}, \\
M_e(\mathbf{x}) - \frac{U_e+L_e}{2} &\le (1-\gamma)\frac{U_e-L_e}{2} \\
\text{for } e &= 1,\dots,m.
\end{aligned} \tag{4.4}
$$

In this definition, the original constraints $L_e \le M_e(\mathbf{x}) \le U_e$ are augmented with a scalar $\gamma$, where positive values of $\gamma$ imply tightening of the constraint, and negative values imply loosening. The consistency measure quantifies how much the constraints can be tightened while still ensuring the existence of a set of parameter values whose associated model predictions match (within their respective bounds) the experimental QoIs. The dataset is *consistent* if the consistency measure is nonnegative, and is *inconsistent* otherwise.

**4.4.3. Model prediction.** Consider a physical configuration (set of conditions) not exercised experimentally but with a property P predicted by model $M_{\mathrm{P}}$. A natural and perhaps the ultimate question of scientific inquiry is, "What is the range of values this model exhibits over the domain of feasible parameter values?" In other words, what is the prediction interval for property P that is consistent with all of the model/observation pairs in the dataset? We refer to this as *model prediction*.

The B2B computation expresses this question into two optimization problems for the lower and upper interval endpoints, $L_{\mathrm{P}}$ and $U_{\mathrm{P}}$,

$$
L_{\mathrm{P}} := \min_{\mathbf{x}\in\mathcal{F}} M_{\mathrm{P}}(\mathbf{x}), \tag{4.5}
$$

$$
U_{\mathrm{P}} := \max_{\mathbf{x}\in\mathcal{F}} M_{\mathrm{P}}(\mathbf{x}). \tag{4.6}
$$

The length $U_{\mathrm{P}} - L_{\mathrm{P}}$ quantifies the amount of uncertainty in $M_{\mathrm{P}}$'s value conditioned on the fact that the true parameter vector is contained in the feasible set $\mathcal{F}$.

As a simple example, consider $M_{\mathrm{P}}(x) := x_k$, the $k$th component of the uncertain parameter vector. The prior information (i.e., $\mathcal{H}$) constrains this as $x_{k,\min} \le x_k \le x_{k,\max}$, whereas the calculations in (4.5) and (4.6) give the posterior range of $x_k$ when restricted to the feasible set $\mathcal{F}$. Geometrically, the predicted interval $[L_{x_k}, U_{x_k}]$ is the projection of high-dimensional set $\mathcal{F}$ onto the $k$th parameter coordinate.

**4.4.4. Optimization on the feasible set.** The optimization problem of B2B, (4.3), involves minimization and maximization of a model over the feasible set $\mathcal{F}$. In Appendix A, we briefly review some key results pertaining to constrained optimization and quantify the

computational importance of polynomial and, more specifically, quadratic parameter dependence in the $M_e$ models. In particular, we note that semidefinite programming enables highly efficient solutions when models $M_e$ are quadratic.

**4.5. Bayesian calibration and prediction formulation.** *The Bayesian calibration and prediction framework* (BCP) is founded on a probabilistic description of the system. We adopt the assumption made previously that the model is an adequate description of reality, $Y_e$, and that its experimental measure is

$$y_e = Y_e + \varepsilon_e \ = M_e(\mathbf{x}) + b_e + \varepsilon_e, \tag{4.7}$$

where $b_e$ is measurement bias and $\varepsilon_e$ is a symmetric component of measurement error. Formally, this is the approach laid out in [20], except that $b_e$ there is model bias (model discrepancy), while here it is measurement bias. We comment later on the issue of model bias. There are $m$ such QoIs and therefore a set $Y_e$, $e = 1, \ldots, m$, of such "realities" and accompanying measurements and measurement errors including biases. The $M_e$'s, as described in sections 4.2 and 4.3, are accepted as the solutions to the ODE system for the QoIs. $M_e$ depends on the unknown model parameters $\mathbf{x}$ (also known as calibration parameters). We denote the vector of bias terms by $\mathbf{b} = (b_1, \ldots, b_m)$.

By specifying distributional assumptions on $\varepsilon_e$, for example, normal with variance $\sigma_e^2$, we get the likelihood

$$L(\mathbf{y} \mid \boldsymbol{\sigma}^{\mathbf{2}}, \mathbf{b}, \mathbf{x}) \propto \prod_{e=1}^{m} \mathrm{N}(y_e \mid M_e(\mathbf{x}) + b_e, \sigma_e^2) \ , \tag{4.8}$$

where $\boldsymbol{\sigma}^{\mathbf{2}} = (\sigma_1^2, \ldots, \sigma_m^2)$ and $\mathbf{y} = (y_1, \ldots, y_m)$. This forms the basis for learning about the unknowns $\mathbf{b}$, $\mathbf{x}$.

For $\boldsymbol{\sigma}^{\mathbf{2}}$, we draw upon the B2B formulation and fix its value by choosing the individual $\sigma_e^2$ so that a normal distribution with mean 0 and variance $\sigma_e^2$ puts probability 0.95 on the interval of width $(U_e - L_e)$ centered at 0, where $L_e$ and $U_e$ are the bounds from (4.1).

Other distributional choices for $\varepsilon_e$ such as having the symmetric component of error uniform, or of using a truncated normal distribution, would be plausible but at some extra computational burden. It is also plausible, as borne out in section 2, that there is little difference to be seen between uniform and matched Gaussian error. Nor is it clear what would be appropriate distributions when the only information is given by bounds that are themselves uncertain. But our main motivation is to see what a typical BCP approach using Gaussian errors does vis-à-vis the typical B2B approach.

To proceed in a Bayesian fashion, one must specify a probability distribution on the unknowns, the so-called prior distribution, which we denote by $\pi(\mathbf{b}, \mathbf{x})$, and then combine this distribution with (4.8) via Bayes' theorem to produce the posterior distribution, $\pi(\mathbf{b}, \mathbf{x} \mid \mathbf{y})$. (The choice of $\pi(\mathbf{b}, \mathbf{x})$ is problem specific; the choice for the methane problem is described in section 5.2.) The posterior distribution summarizes all the available knowledge about the unknowns after observing the experimental data $\mathbf{y}$, combining sampling information and prior information into a single distribution and forming the basis for statistical inference and prediction.

It is clear that there is an unavoidable confounding between $\mathbf{x}$ and $b_e$: a change in one of them adjusted by a change in the other can preserve the value of $Y_e$. It is by including as much expert information as possible in the prior on $b_e$ and $\mathbf{x}$ that we may hope to untangle the dependence between the bias and the calibration parameters—jointly estimating $b_e$ and $\mathbf{x}$ in any case will account for this additional source of uncertainty in prediction. In the context of model bias, predictions and assessment of uncertainty are done in [2]; the same methods are applicable whether $b_e$ is measurement bias or a combination of measurement and model bias, and we follow that approach here. In practice the implementation of this general approach can vary from problem to problem.

**4.5.1. Implementation and interpretation.** The computation of the posterior distribution, $\pi(\mathbf{b}, \mathbf{x} \mid \mathbf{y})$, is generally done by a Markov chain Monte Carlo (MCMC) algorithm [17]. We implemented this algorithm in WinBUGS [22] launched from R [30].

The MCMC applied to (4.7) and (4.8) produces a sample, $\{\mathbf{x}_k, \mathbf{b}_k,\ k = 1, \ldots, K\}$, from the posterior distribution of all unknowns. To predict an outcome of a new QoI, that is, predict $Y_{\mathrm{P}}$ for a new QoI and conditions, we can use the sample from the posterior of $M_{\mathrm{P}}$,

$$M_{\mathrm{P}}(\mathbf{x}_k),\ k = 1, \ldots, K, \tag{4.9}$$

and obtain a so-termed pure-model prediction of $Y_{\mathrm{P}}$ as $\hat{M}_{\mathrm{P}}$, with its mean evaluated as the average of $M_{\mathrm{P}}(\mathbf{x}_k)$. (The pure-model prediction, as described in [2], results from computing the posterior mode or mean, $\hat{\mathbf{x}}$, of the $\mathbf{x}_k$ and then evaluating $M_{\mathrm{P}}$ at this estimated value. This is a minor difference here.)

**4.5.2. What if model bias is present?** Pure-model prediction is called for in the absence of model bias. In many applications, unlike our methane example, model bias is certain to be present and not ignorable [18]. In such cases, if measurement bias is absent, an appropriate predictor [2] is to use the statistical model in (4.7) to construct the so-called bias-corrected prediction. The bias-corrected prediction is obtained from the MCMC (the prior distributions we use are the same no matter how $\mathbf{b}$ is interpreted and the resulting MCMC would then be the same), computing

$$Y_{\mathrm{P}}(\mathbf{x}_k) = M_{\mathrm{P}}(\mathbf{x}_k) + b_{\mathrm{P},k},\ k = 1, \ldots, K, \tag{4.10}$$

and taking the average (or median) of the $Y_{\mathrm{P}}(\mathbf{x}_k)$.

But how do we obtain the samples $\{b_{\mathrm{P},k}\}$? Here we are forced to make assumptions—there are no data to inform about model discrepancy at a new QoI. If, for example, model bias for $Y_{\mathrm{P}}$ is similar to biases of some of the experimented-upon QoIs, then we might choose $b_{\mathrm{P}}$ as the average of those $b_e$, and the MCMC sample from the posterior would then generate a sample for $b_{\mathrm{P}}$.

Matters become more complicated if both measurement and model bias are present. The reason: $Y_{\mathrm{P}}(\mathbf{x})$ has no measurement error, so an average of $b_e$ could result in a misleading $b_{\mathrm{P}}$. The confounding of the two biases leads to ambiguities and affects the interpretation of predictions (either pure-model or bias-corrected). In section 5.5 we point out the differences that may arise if we had assumed that the bias in our methane example was due to model imperfection.

## 5. Results and discussion.

**5.1. Brief recapping.** Before proceeding with the description and comparison of the results, it is pertinent to briefly revisit the commonality and differences in the present analyses:

- Both B2B and BCP use the same quadratic surrogates as models.
- Both B2B and BCP use the same ranges of model parameters. B2B relies on the bounds only, whereas the BCP assumes that the model parameters are uniformly distributed within those bounds.
- Both B2B and BCP rely on the same experimental data. B2B uses the experimenter-reported ranges, whereas BCP uses normal distributions based on the experimenter-reported observation values and uncertainties.
- BCP introduces an additive term and accompanying distribution to each model to account for measurement bias.

The similarities and differences in this list are a template for comparing the results of the analyses described in sections 4.4 and 4.5.

**5.2. Specifying the prior for BCP.** We assume that $\pi(\mathbf{b}, \mathbf{x})$, the prior distribution of the unknowns, is a product of $\pi(\mathbf{b})$ and $\pi(\mathbf{x})$—an independence assumption. We take $\pi(\mathbf{x})$ to be uniform on $\mathcal{H}$.

To construct the prior on $\mathbf{b}$, we grouped experiments from the same laboratory of the same type, except for initial conditions (see Appendix B), and assumed that the biases are the same for experiments in the same group. Additionally, we assumed that the biases were independent across groups. A total of 44 groups were created (Table 1), and the prior on $b_j$, the bias for group $j$, is

$$\pi(b_j) = \int_0^{+\infty} \mathrm{N}(b_j \mid 0, \tau_j^2)\ \pi(\tau_j)\ d\tau_j,$$

where $\pi(\tau_j) \propto 1/(1 + \tau_j^2/\bar{\sigma}_j^2)$ with $\bar{\sigma}_j^2$ the average of the variances over the members of the group. The prior on $\tau_j$ is thus a Cauchy distribution with location parameter zero and scale parameter equal to $\bar{\sigma}_j$. In Bayesian language, the prior on $b_j$ is hierarchical in that it is parameterized by $\tau_j$ (the hyperparameter) and $\tau_j$ itself has prior density (the hyperprior). We are saying that a priori $b_j$ is normal with zero mean, but are uncertain about its standard deviation. The probability distribution describing this uncertainty exhibits heavy tails and a scale similar to the measurement error. This particular choice of prior for $b_j$ is similar to the one used by [1] and has roots on priors for variance parameters in random effects models [16]. This choice of the prior is objective in that it emphasizes the role of the data in obtaining posterior distributions and downplays the impact of the prior.

**5.3. A comment on computations.** Both approaches rely on dimension reduction. Developing surrogate models each in its own set of active variables, which benefits B2B [13], also benefits the MCMC in the BCP analysis. Indeed, only 8 to 12 parameters are involved in any one of the 76 QoIs. Had we needed to use all 102 parameters for each QoI, we would have $102\times102$ matrices rather than $10\times10$ matrices. The gain in computation is a factor of about 100. However, since the group of relevant parameters that are involved in each of the QoIs changes, the computer code was modified to optimally compute the polynomials.

**Table 1**
*Grouping of the bias terms. (QoIs are from [29]; see Appendix B.)*

| QoI | Bias | QoI | Bias | QoI | Bias |
|---|---|---|---|---|---|
| IG.1a, IG.1b | $b_1$ | CH3.T4 | $b_{16}$ | StF8 | $b_{31}$ |
| IG.2 | $b_2$ | CH3.StC6, CH3.StC7 | $b_{17}$ | SNO.C11 | $b_{32}$ |
| IG.6a, IG.6b | $b_3$ | OH.1a, OH.1b | $b_{18}$ | SCH.C11–C13 | $b_{33}$ |
| IG.T1 | $b_4$ | OH.2 | $b_{19}$ | CH.St | $b_{34}$ |
| IG.T2 | $b_5$ | OH.3a–d, OH.ST8 | $b_{20}$ | NFR1 | $b_{35}$ |
| IG.St1a, IG.St1b | $b_6$ | CO.C1a–d, CO.SC8 | $b_{21}$ | NFR2 | $b_{36}$ |
| IG.St3a, IG.St3b | $b_7$ | CO.T1a–d, CO.ST8 | $b_{22}$ | NFR3 | $b_{37}$ |
| IG.St4a, IG.St4b | $b_8$ | BCO.T1–T7 | $b_{23}$ | NF6 | $b_{38}$ |
| CH3.C1a, CH3.C1b | $b_9$ | BCH2O.T1–T3 | $b_{24}$ | NF7 | $b_{39}$ |
| CH3.T1a, CH3.T1b | $b_{10}$ | SR.10c | $b_{25}$ | NF11 | $b_{40}$ |
| CH3.C2 | $b_{11}$ | F1 | $b_{26}$ | NF12/13 | $b_{41}$ |
| CH3.T2 | $b_{12}$ | F2, F3 | $b_{27}$ | NFR4 | $b_{42}$ |
| CH3.C3 | $b_{13}$ | F4 | $b_{28}$ | NFR5 | $b_{43}$ |
| CH3.T3 | $b_{14}$ | F6 | $b_{29}$ | CHNO.St | $b_{44}$ |
| CH3.C4 | $b_{15}$ | SF7 | $b_{30}$ | | |

**5.4. Posterior predictions for model parameters.** Our initial comparison is for the calibration parameter vector, $\mathbf{x}$: BCP posterior distributions versus B2B bounds. The latter are computed by sequential execution of the B2B predictions for $M_i(\mathbf{x}) = x_i$, $i = 1, \ldots, 102$. The results for all 102 variables are shown in Figure 5. The length of the abscissa in all individual plots is the prior uncertainty interval. The B2B predicted intervals are drawn as horizontal lines. Both the outer and inner B2B intervals are given. The difference between them for an individual $x_i$ indicates the uncertainty of the B2B procedure itself; the difference can be narrowed down through additional iterations. The green histograms designate the BCP posterior distributions.

The results presented in Figure 5 demonstrate several features. In 83 out of 102 cases, the B2B predictions for the $x_i$ intervals are of the same length as the prior ones. In other words, the collaboration of all the data produces a feasible set whose projections on most $x_i$ axes span the original ranges. This implies that while the volume of $\mathcal{F}$ is orders of magnitude smaller than the initial prior cube $\mathcal{H}$ [15], the feasible set still spans prior intervals in those $x_i$ directions. Physically this means that the experimental data included in the analysis, while substantially lowering the total system uncertainty, do not contribute to lowering the uncertainty in individual $x_i$'s. The BCP posteriors for 67 of the 102 cases are "uninformative" ($x_3$, for instance), in agreement with the corresponding B2B predictions.

There are also a substantial number of cases where the collaboration of data narrows down the predicted range of an $x_i$, and in many of those cases the two methods are in agreement with each other, for example, $x_{35}$.

There are also differences between the two sets of results. Consider the case of $x_1$. B2B


**Figure 5.** *Comparison of predictions for model parameters. The length of the abscissa is the prior uncertainty interval; green histogram is the BCP posterior; red and blue horizontal bars are the predicted B2B outer and inner bounds, respectively.*

predicts the $x_1$ values to lie on the left part of the interval, while the BCP posterior puts mass on the entire interval. A B2B sensitivity analysis [25] reveals that the upper bound of $x_1$ is essentially controlled by the upper bound of experimental QoI labeled SF7. But the BCP analysis produces a bias for SF7 (not illustrated here) whose posterior distribution lies entirely to the right of 0. In essence BCP trades information about $x_1$ for information about the bias. Which of these conclusions is appropriate is unclear, nor is it clear whether the upper bound for SF7 should be challenged.

There are also cases, like $x_{59}$, where the B2B-predicted bounds do not shrink the initial interval, but the BCP posterior has a clear peak in the middle of the interval. Such results suggest that the feasible set $\mathcal{F}$ has mostly a "Gaussian"-like distribution in the $x_{59}$th dimension but contains "remote" points that pass all the uncertainty constraints from the B2B perspective.

**5.5. Predictions for new QoIs: Leave one out.** Our next comparison involves removing a single QoI from the dataset of 76, and then using the remaining 75 to predict the omitted QoI. This results in 76 analyses and predictions.

The B2B results are obtained by applying the model prediction technique detailed in section 4.4.3; the BCP results follow (4.9). The uncertainties of the predictions do not take into account the measurement error in the omitted QoI because it would generally be unavailable.

We report the predictions in Figure 6; the former exhibits the B2B bounds along with posterior distributions of pure-model predictions, as defined in (4.9). Also reported in the figure, for comparison, are the ranges of the corresponding experimental QoI values, which do include their measurement errors.

Details on how the Bayesian computations were implemented are as follows. For each of the 76 analyses, we carried out the MCMC calculations to obtain a posterior sample $\{\mathbf{x}_k^{(e)}, \mathbf{b}_k^{(e)},\ k = 1, \ldots, K\}$ which does not include the data point corresponding to the $e$th QoI, $y_e$, $e = 1, \ldots, m$. This allows us to obtain a sample from the pure-model prediction of $Y_e$ by computing $\{M_e(\mathbf{x}_k^{(e)}),\ k = 1, \ldots, K\}$. Histograms of these draws are depicted in Figure 6.

Inspection of the results presented in Figure 6 indicates several features. There are cases, like CH3.T2 or OH.3c, where the B2B and BCP are in close agreement with each other as well as with the experimental data. In some cases, e.g., BCO.T4, the B2B predictions are wider than BCP; in some cases, e.g., CH3.C3, the BCP predictions are wider than B2B's. This last observation is due to BCP sampling outside the B2B feasible set, a consequence of the distributional assumption on $e$ in (4.7). In some cases, e.g., NF6, predictions of both methods are wider than the experimental ones, suggesting that the omitted QoI conveys valuable information about $\mathbf{x}$ not obtained in the remaining 75 experiments. In other cases, e.g., NF11, the predicted intervals are narrower than experimental ones, suggesting that such a QoI does not add to the knowledge already present in the rest of the dataset. This comparison underlies a measure of information content of an experiment treated in [15] and [19].

**5.5.1. Predictions when model bias is present.** If we thought that the bias was model discrepancy and not measurement bias, we would be using bias-corrected predictions rather than pure-model predictions. It is interesting to note the effect of so doing and the comparison with the results in section 5.5. The B2B bounds remain the same. The bias-corrected prediction was in section 4.5.2. As discussed there, we need to make an assumption about how

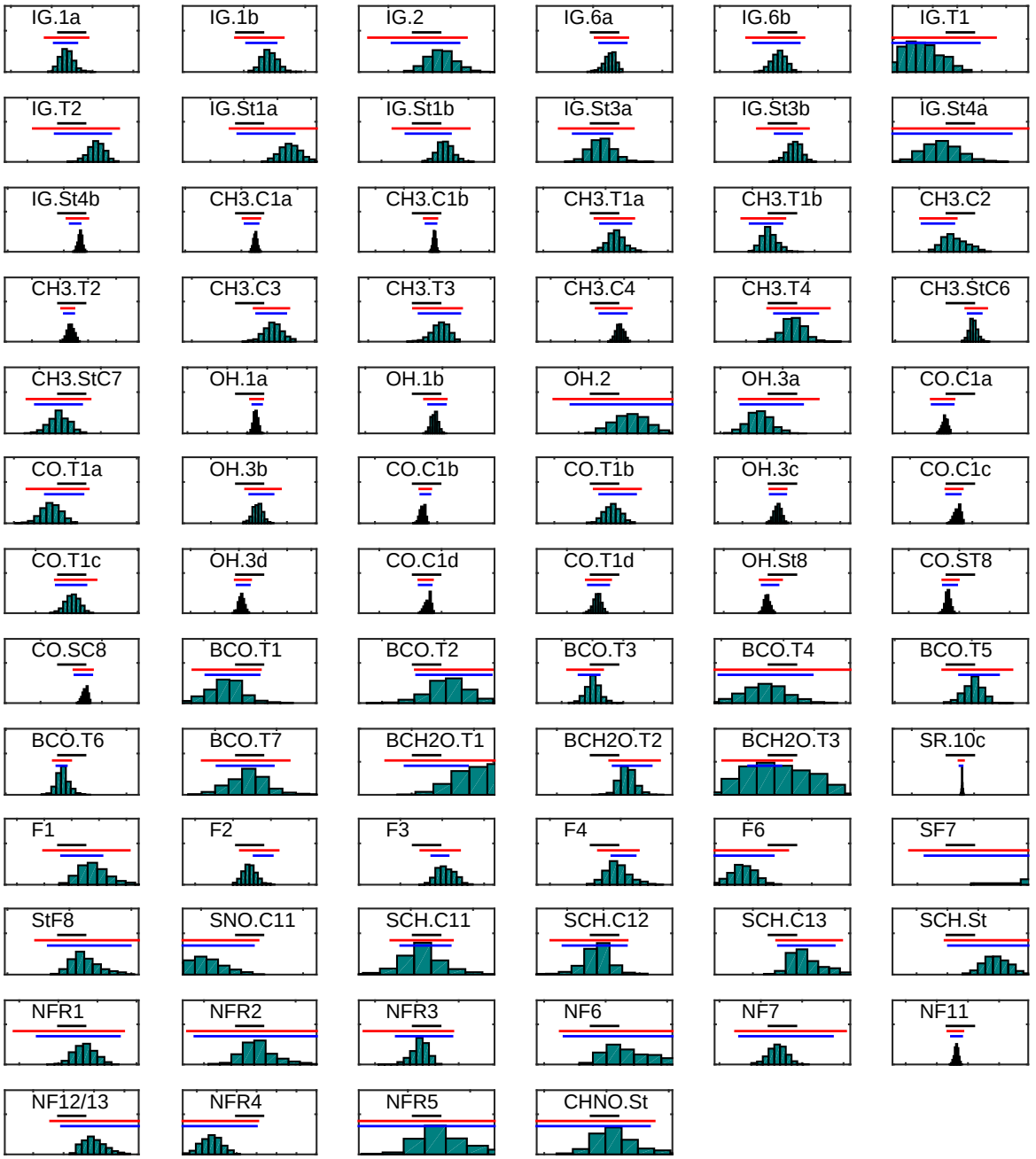


**Figure 6.** *"Leave-one-out" experiment. Comparison of predictions for QoIs. The black horizontal bar is the range of experimental values; red and blue bars are the predicted B2B outer and inner bounds, respectively; green histogram is the BCP posterior pure-model prediction.*

**Table 2**

*Selection of the bias in the leave-one-out experiment. For those QoIs that belong to a group with more than one element, we use the bias of that group, defined in Table 1.*

| QoI | Bias | QoI | Bias | QoI | Bias |
|---|---|---|---|---|---|
| IG.2 | $b_1$ | OH.2 | $b_{18}$ | NFR2 | $b_{35}$ |
| IG.T1 | $b_{20}$ | SR.10c | $b_{42}$ | NFR3 | $b_{35}$ |
| IG.T2 | $b_6$ | F1 | $b_{27}$ | NF6 | $b_{35}$ |
| CH3.C2 | $b_9$ | F4 | $b_{27}$ | NF7 | $b_{35}$ |
| CH3.T2 | $b_{10}$ | SF7 | $b_{27}$ | NF11 | $b_{40}$ |
| CH3.C3 | $b_9$ | StF8 | $b_{27}$ | NF12/13 | $b_{40}$ |
| CH3.T3 | $b_{10}$ | SNO.C11 | $b_{33}$ | NFR4 | $b_{42}$ |
| CH3.C4 | $b_9$ | CH.St | $b_{35}$ | NFR5 | $b_{42}$ |
| CH3.T4 | $b_{10}$ | NFR1 | $b_{35}$ | CHNO.St | $b_9$ |

$b_e$—the bias associated with the left-out QoI—relates to the biases of the experimented-upon QoIs. Our approach is quite natural. Recall that we have grouped the experiments in order to construct the prior on the bias term. That grouping is described in Table 1. If the left-out QoI belongs to a group $j(e)$ that contains more than one experiment, we use the samples from that bias to compute

$$Y_e(\mathbf{x}_k^{(e)}) = M_e(\mathbf{x}_k^{(e)}) + b_{j(e),k}^{(e)} \ .$$

If the left-out QoI is the only experiment in the grouping of Table 1, then we select a "neighbor" experiment from which we borrow the bias. Those are described in Table 2. The resulting draws are depicted in the form of histograms in Figure 7.

It should be noted that the extrapolation used to obtain $b_e$ is done without adequate information on how "neighboring" biases relate to $b_e$, unlike what happens, for instance, in [1]. This is clearly a risky process and one that emphasizes the open question of what constitutes acceptable extrapolation. An interesting and more conservative alternative, raised by a referee, would be using the prior density for the bias term of a QoI that is not part of a group. We have not yet explored such a possibility.

Inspection of Figure 7 indicates that there are cases (OH.3c) when both pure-model and bias-corrected BCP posteriors are essentially the same. There are instances, e.g., F6 or NFR4, when inclusion of the bias term improves the BCP prediction when compared to the experiment, but there are other cases, e.g., SR.10c or NFR2, where it makes it worse.

Which of these BCP predictions is generally appropriate is arguable. If all the biases are measurement bias, i.e., model discrepancy is ignorable, we would have no way of determining the measurement error in a new QoI or its possible bias. In such a case, using a "nearby" bias term (as done in bias-corrected prediction) is inappropriate, and pure-model prediction is called for. On the other hand, if the biases are from model discrepancy, then using bias-corrected prediction is called for; whether the particular way we define and use neighboring QoIs is reasonable is another matter. If both elements are present, it is unclear which, if any, of the two prediction methods is appropriate. Viewing the results in Figures 6 and 7, we note that measurement bias is more pronounced than model discrepancy; of course, the assumptions in our problem point to measurement bias.

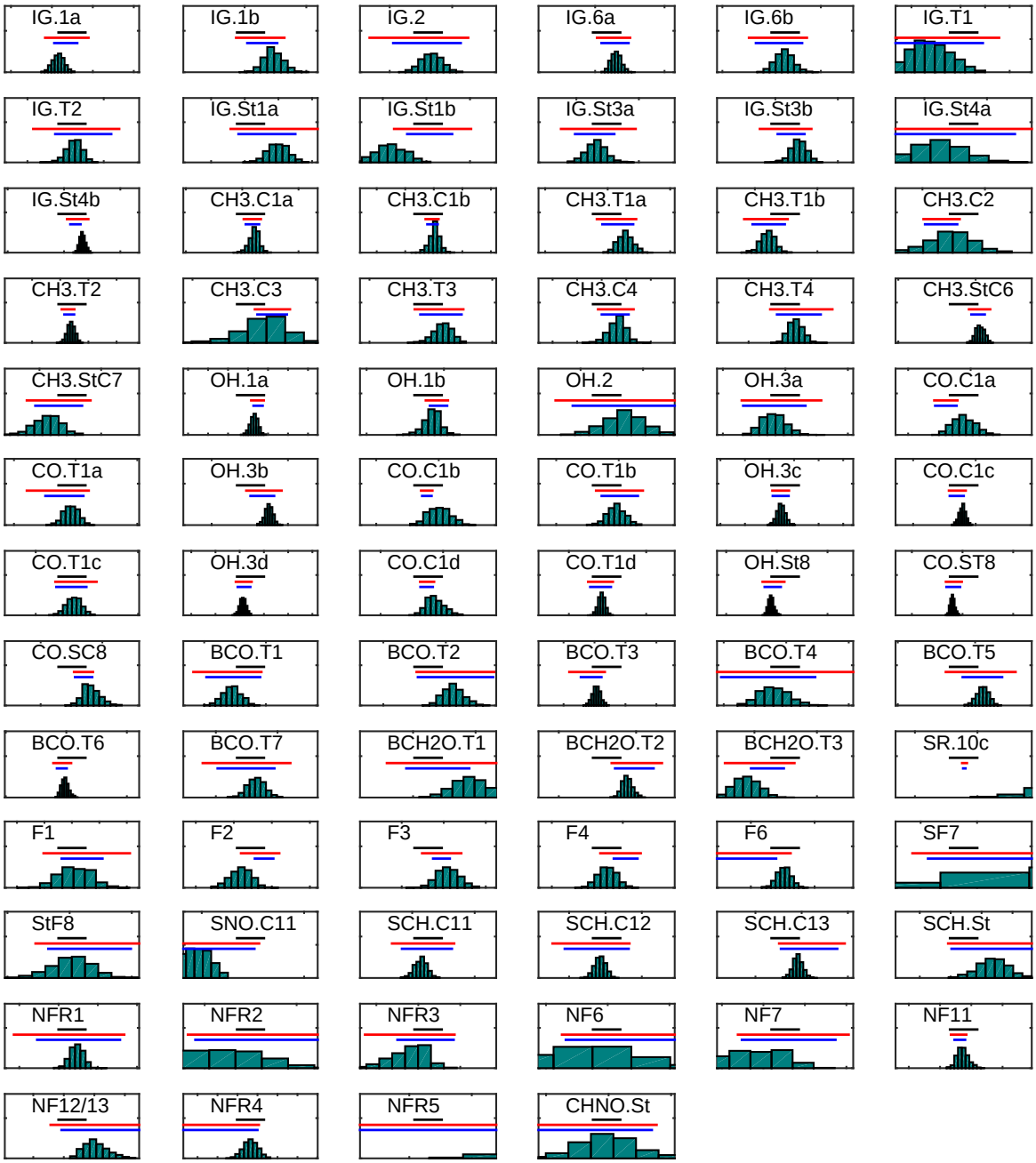


**Figure 7.** *"Leave-one-out" experiment. Comparison of predictions for QoIs. The black horizontal bar is the range of experimental values; red and blue bars are the predicted B2B outer and inner bounds, respectively; green histogram is the BCP bias-corrected posterior.*

**5.6. B2B-BCP connection.** The elicitation of opinion from experts about the nature of measurement uncertainty leading to (4.2) has a Bayesian interpretation but manifests itself in terms of bounds rather than distributions. But there is a theoretical connection between the BCP and B2B approaches: In (4.7), replace the previous assumption about $\varepsilon_e$ with the assumption that they are independent random variables with a uniform distribution on the fixed interval $(-a_e, a_e)$, $a_e = (U_e - L_e)/2 > 0$. The sampling density of the observed data is then

$$f(\mathbf{y} \mid \mathbf{x}, \mathbf{b}, \mathbf{a}) = \prod_{e=1}^{m} \mathrm{U}(y_e \mid M_e(\mathbf{x}) + b_e - a_e, M_e(\mathbf{x}) + b_e + a_e) \ .$$

For $\mathbf{x}$ use, as before, a uniform prior on $\mathcal{H}$. The posterior of $\mathbf{x}$ given the data $\mathbf{y}$ and the parameters $\mathbf{a}$ and $\mathbf{b}$ is

$$\pi(\mathbf{x} \mid \mathbf{a}, \mathbf{b}, \mathbf{y}) \propto 1_{\mathcal{P}}(\mathbf{x}),$$

where

$$\begin{aligned} \mathcal{P} &= \{\mathbf{x} \in \mathcal{H} : \forall e = 1, \ldots, m,\ M_e(\mathbf{x}) + b_e - a_e < y_e < M_e(\mathbf{x}) + b_e + a_e\} \\ &= \{\mathbf{x} \in \mathcal{H} : \forall e = 1, \ldots, m,\ y_e - a_e - b_e < M_e(\mathbf{x}) < y_e + a_e - b_e\} \ . \end{aligned}$$

If we set $b_e = y_e - (U_e + L_e)/2$, then $\mathcal{P} = \mathcal{F}$, the feasibility set.

The distinction with the BCP formulation in section 4.5 is the change in the assumptions on $\varepsilon_e$ and *not* having a prior on $\mathbf{b}$. In the Bayesian formulation, $\mathbf{b}$ has a prior distribution; here a "plug-in" value is used to estimate the "incidental" parameters, $\mathbf{b}$. What is lacking is an adequate assessment of uncertainty arising from the data-dependent choice of the plug-in for $\mathbf{b}$. Nevertheless, the similarity of predictions via B2B or BCP seen in Figures 6 and 7 is thus understood.

**6. Conclusions and comments.** What can we learn from these analyses? The results in Figure 6 suggest that, except for a handful of cases, pure-model predictions are consistent with the B2B bounds. Moreover, both are "equally close" to the QoI experimental bounds. What differences are present may or may not be of practical concern; this is a matter for domain scientists to resolve. Neither the B2B nor the BCP approach dominates: the B2B analysis makes assumptions that are different and plausibly closer to the experimenters knowledge; the BCP approach uses the inputs to the B2B analysis along with additional assumptions and produces a more nuanced assessment of uncertainty.

Despite the assumption that there is no model bias there is always the nagging question of whether such a bias is present. Comparing Figures 6 and 7, it appears bias-corrected predictions are not as good as the pure-model predictions, suggesting that measurement bias is a dominant component of $\mathbf{b}$. In cases where there is wide discrepancy between the B2B bounds and one of the BCP methods (for example, QoIs SR.10c, NFR5, and SF7) it may be useful to opt for the "other" BCP method, creating a hybrid strategy. Shortcomings in the reliability and knowledge of the experimental data can be a more significant factor in interpretation of results than differences between the methods of analysis.

Generalizing these methods to other contexts should take into account that we have made some specific assumptions such as reduction of the ODE system and relatively fast computation of solutions, quadratic polynomial surrogates, and absence of model bias. Alternate

surrogate schemes as noted in section 3.1 are available for other contexts but with added complexity for B2B. MCMC computations may become more burdensome with nonreduced systems and high-order surrogates. The presence of model bias as pointed out in section 4.5.2 can be troublesome for prediction and interpretation.

**Appendix A. Optimization methods.** Without loss of generality, we just consider *minimizations* with inequality constraints, referred to as primal problems,

$$\begin{aligned} p^* = &\min_{x\in\mathbb{R}^n} f(x) \\ &\text{subject to: } g_j(x) \le 0 \quad \text{for } j = 1,\dots,m. \end{aligned}$$

Any direct attempt at this constrained optimization typically only yields an upper bound to $p^*$, since the true minimum may not have been found, due to nonconvexity of the function and/or constraint set, for example. In order to bracket $p^*$, a lower bound is also needed. Lower bounds are often obtained by considering the associated dual (to the primal) problem

$$\begin{aligned} q^* = &\max_{\lambda\in\mathbb{R}^m} \min_{x\in\mathbb{R}^n} f(x) + \lambda^{\mathrm{T}} g(x) \\ &\text{subject to: } \lambda_j \ge 0 \quad \text{for } j = 1,\dots,m, \end{aligned}$$

which *always has* $q^* \le p^*$, giving (if $q^*$ is reliably computed) a lower bound to the primal. As we are considering minimization, lower and upper bounds on $p^*$ are referred to as *outer* and *inner* bounds, respectively. Additionally, the solution of the dual problem informs how $p^*$ is affected by changes in the constraints. Specifically, the primal problem with variable constraints, $v$,

$$\begin{aligned} p^*(v) = &\min_{x\in\mathbb{R}^n} f(x) \\ &\text{subject to: } g_j(x) \le v_j \quad \text{for } j = 1,\dots,m, \end{aligned}$$

is related to the optimal dual variables $\lambda^*$, which act as global (one-sided) sensitivities through the bound $q^* - v^{\mathrm{T}}\lambda^* \le p^*(v)$ for all $v \in \mathbb{R}^m$. Finally, and of critical importance, *if the functions $f$ and all $g_j$ are quadratic (not necessarily sign-definite), then the dual problem is solved efficiently via semidefinite programming* [3, 27]. *Problems with hundreds of variables ($n$) and constraints ($m$) are routine, even in a desktop environment. Taken together, the efficiently computed bound $q^*$ and sensitivities $\lambda^*$ highlight an important consequence for B2B uncertainty quantification with quadratic response-surface models.*

The deterministic perspective of B2B directly leads to the constrained minimization (primal) discussed here. Under the quadratic-dependence restriction on $f$ and all $g_j$, the lower bound $q^*$ provided by the dual problem is equivalent to a stochastic formulation [11]. Specifically, replace the deterministic variable $x$ with a random variable $X$, the only restriction being of finite variance, and modify the cost and constraints of the primal problem to reflect mean values, as

$$\begin{aligned} s^*(v) = &\min_{X} \mathbf{E}\left[f(X)\right] \\ &\text{subject to: } \mathbf{E}\left[g_j(X)\right] \le v_j \quad \text{for } j = 1,\dots,m, \end{aligned}$$

where $\mathbf{E}$ denotes expectation and the minimization is taken over all random variables $X$ with finite variance. Then $s^*(v) = q^*(v)$, where $q^*(v)$ refers to the dual problem for $p^*(v)$, and hence $p^*$ and $s^*$ satisfy global sensitivity relations $s^*(0) - v^T\lambda^* \le s^*(v) \le p^*(v)$ for all $v \in \mathbb{R}^m$. So, while $s^*$ is not a Bayesian estimate as discussed in section 4.5, it offers a probabilistic interpretation of the computed outer bounds in the B2B results. Furthermore, $s^*$ (along with global sensitivities) is computed efficiently, is independent of priors, and satisfies a known relation to the deterministic minimization.

In conclusion, bounds to individually bracket $C_D$, $L_P$, and $U_P$ are solved efficiently using polynomial optimization techniques when all the models are quadratic in the parameters [14, 15, 27]. Further use of branch-and-bound techniques [7] tightens the bracketing interval of each quantity. Solving these optimizations also automatically generates sensitivities of the computed results to the parameter and experiment uncertainties [25]. Extension of the methodology to more general polynomial models is treated in [27] and to nonpolynomial surrogate models, such as Gaussian process, Kriging, or $\epsilon$-SVM, is treated in [9].

**Appendix B. GRI-Mech QoIs.** For the dataset of the present study, 76 experimental QoIs of the GRI-Mech 3.0 release [29] were selected. These QoIs were composed of

- shock-tube (ST) ignition delays;
- species concentration peaks and times of these peaks in shock-heated mixtures;
- species maximum concentrations and times to attain half of the maximum values in shock-heated mixtures;
- flow-reactor observations: temperature rise and species concentrations at the reactor exit;
- laminar flame speeds;
- laminar flame maximum species concentrations;
- laminar flame species ratios.

A brief account of these QoIs is given in Table 3, which also reports the assessed experimental uncertainties [29, 34], lists the experimental series, and identifies laboratories of the experimental studies. Further details can be found in [29, 34]. All these QoI data are also available as an HDF5 file in the supplementary materials (gri_mech_30.h5 [local/web 871KB]).

**Table 3**

*QoIs of the natural-gas combustion dataset (GRI-Mech 3.0 [29, 34]).*

| Key | Description | Measured value | Expt. unc. | Expt. series | Research lab |
|---|---|---|---|---|---|
| IG.1a | ST ign. delay at 1.8 atm, 1500 K | 800 $\mu$s | ±15% | 9.1%$CH_4$-18.2%$O_2$-Ar | Seery, Bowman 1970 |
| IG.1b | ST ign. delay at 2 atm, 1700 K | 105 $\mu$s | ±10% | | |
| IG.2 | ST ign. delay at 3.8 atm, 1600 K | 335 $\mu$s | ±10% | 33.3%$CH_4$-13.3%$O_2$-Ar | |
| IG.St3a | ST ign. delay at 1.6 atm, 1530 K | 500 $\mu$s | ±10% | 4.8%$CH_4$-19.1%$O_2$-Ar | |
| IG.St3b | ST ign. delay at 1.9 atm, 1845 K | 43 $\mu$s | ±10% | | |

**Table 3**
*Continued.*

| | | | | | |
|---|---|---|---|---|---|
| IG.6a | ST ign. delay at 0.25 atm, 1600 K | 436 $\mu$s | $-30\%$, $+10\%$ | 0.5%$C_2H_6$-1.8%$O_2$-Ar | Hidaka, Gardiner 1982 |
| IG.6b | ST ign. delay at 0.25 atm, 1600 K | 222 $\mu$s | $-20\%$, $+10\%$ | 0.5%$C_2H_6$-5.0%$O_2$-Ar | |
| IG.T1 | ST ign. delay at 2.5 atm, 1410 K | 275 $\mu$s | $\pm 25\%$ | 9.5%$CH_4$-1.9%$C_3H_8$-19.0%$O_2$-Ar | Frenklach, Bornside 1984 |
| IG.T2 | ST ign. delay at 7.1 atm, 1640 K | 63.2 $\mu$s | $\pm 20\%$ | 3.4%$CH_4$-0.1%$C_3H_8$-7%$O_2$-Ar | Spadaccini, Colket 1994 |
| IG.St1a | ST ign. delay at 6.1 atm, 1356 K | 1.0 ms | $-10\%$, $+20\%$ | 3.29%$CH_4$-0.21%$C_2H_6$-7%$O_2$-Ar | |
| IG.St1b | ST ign. delay at 7.6 atm, 1688 K | 39 $\mu$s | $-10\%$, $+20\%$ | | |
| IG.St4a | ST ign. delay at 34.6 atm, 1408 K | 534 $\mu$s | $\pm 10\%$ | 2.4%$CH_4$-6.7%$O_2$-Ar | Hanson et al. 1995 |
| IG.St4b | ST ign. delay at 83.9 atm, 1706 K | 160 $\mu$s | $\pm 50\%$ | 0.28%$CH_4$-0.56%$O_2$-Ar | |
| CH3.C1a | ST $CH_3$-peak conc at 1 atm, 2000 K | 162 ppm | $\pm 50\%$ | 0.1%$CH_4$-0.2%$O_2$-Ar | Chang, Davidson, DiRosa, Hanson, Bowman 1994 |
| CH3.T1a | ST $CH_3$-peak time at 1 atm, 2000 K | 397 $\mu$s | $\pm 10\%$ | | |
| CH3.C1b | ST $CH_3$-peak conc at 1 atm, 2400 K | 290 ppm | $\pm 50\%$ | | |
| CH3.T1b | ST $CH_3$-peak time at 1 atm, 2400 K | 50 $\mu$s | $\pm 10\%$ | | |
| OH.1a | ST time to $\frac{1}{2}$ OH max at 1 atm, 2000 K | 700 $\mu$s | $\pm 30\%$ | | |
| OH.1b | ST time to $\frac{1}{2}$ OH max at 1 atm, 2200 K | 255 $\mu$s | $\pm 20\%$ | | |
| CH3.C4 | ST $CH_3$-peak conc at 1 atm, 2264 K | 451 ppm | $\pm 25\%$ | 0.2%$CH_4$-0.1%$O_2$-Ar | |
| CH3.T4 | ST $CH_3$-peak time at 1 atm, 2264 K | 83 $\mu$s | $\pm 10\%$ | | |
| CH3.StC6 | ST $CH_3$-peak time at 1 atm, 2454 K | 27 $\mu$s | $\pm 20\%$ | 0.1%$CH_4$-0.4%$O_2$-Ar | |
| CH3.StC7 | ST $CH_3$-peak time at 1 atm, 1932 K | 510 $\mu$s | $\pm 10\%$ | | |
| CH3.C2 | ST $CH_3$-peak conc at 1.2 atm, 1794 K | 342 ppm | $\pm 10\%$ | 0.03%$C_2H_6$-0.1%$O_2$-Ar | |
| CH3.T2 | ST $CH_3$-peak time at 1 atm, 2264 K | 59 $\mu$s | $\pm 50\%$ | | |
| OH.2 | ST time to $\frac{1}{2}$ OH max at 1.2 atm, 1817 K | 193 $\mu$s | $\pm 10\%$ | | |
| CH3.C3 | ST $CH_3$-peak conc at 1.35 atm, 1684 K | 155 ppm | $\pm 10\%$ | 0.02%$C_2H_6$-Ar | Hanson et al. 1992 |
| SCH.St | ST max CH conc at 1.8 atm, 2800 K | 3.11 ppm | $\pm 20\%$ | 0.008%$CH_4$-0.01%$O_2$-Ar | Hanson et al. 1998 |
| CHNO.St | ST max CH conc at 1.8 atm, 2800 K | 1.57 ppm | $\pm 10\%$ | 0.008%$CH_4$-0.01%$O_2$-0.04%NO-Ar | |

**Table 3**
*Continued.*

| | | | | | |
|---|---|---|---|---|---|
| OH.3a | ST time to $\frac{1}{2}$ OH max at 1.51 atm, 1750 K | 159 $\mu$s | ±10% | 0.4%$CH_4$-20%$O_2$-Ar | Yu, Wang, Frenklach 1995 |
| CO.C1a | ST CO-peak conc at 1.51 atm, 1750 K | 38 nmol/cm$^3$ | ±10% | | |
| CO.T1a | ST time to $\frac{1}{2}$ CO max at 1.51 atm, 1750 K | 155 $\mu$s | ±10% | | |
| OH.3b | ST time to $\frac{1}{2}$ OH max at 1.64 atm, 1900 K | 58 $\mu$s | ±20% | | |
| CO.C1b | ST CO-peak conc at 1.64 atm, 1900 K | 37 nmol/cm$^3$ | ±10% | | |
| CO.T1b | time to $\frac{1}{2}$ CO max at 1.64 atm, 1900 K | 58 $\mu$s | ±10% | | |
| OH.3c | ST time to $\frac{1}{2}$ OH max at 1.51 atm, 1750 K | 551 $\mu$s | ±25% | 1%$CH_4$-3%$O_2$-Ar | |
| CO.C1c | ST CO-peak conc at 1.51 atm, 1750 K | 92 nmol/cm$^3$ | ±10% | | |
| CO.T1c | ST time to $\frac{1}{2}$ CO max at 1.51 atm, 1750 K | 481 $\mu$s | ±15% | | |
| OH.3d | ST time to $\frac{1}{2}$ OH max at 1.64 atm, 1900 K | 201 $\mu$s | ±25% | | |
| CO.C1d | ST CO-peak conc at 1.64 atm, 1900 K | 90 nmol/cm$^3$ | ±10% | | |
| CO.T1d | ST time to $\frac{1}{2}$ CO max at 1.64 atm, 1900 K | 179 $\mu$s | ±20% | | |
| OH.St8 | ST time to $\frac{1}{2}$ OH max at 2.45 atm, 1865 K | 400 $\mu$s | ±30% | 1%$CH_4$-1.5%$O_2$-Ar | |
| CO.ST8 | ST time to $\frac{1}{2}$ CO max at 2.45 atm, 1865 K | 340 $\mu$s | ±30% | | |
| CO.SC8 | ST CO-peak conc at 2.45 atm, 1865 K | 127 $\mu$s | ±10% | | |
| BCO.T1 | ST time to $\frac{1}{2}$ CO max at 1.17 atm, 2124 K | 14.2 $\mu$s | ±10% | 1.46%$CH_2O$-Ar | Eiteneer, Yu, Goldenberg, Frenklach 1998 |
| BCO.T2 | ST time to $\frac{1}{2}$ CO max at 1.51 atm, 1724 K | 88.2 $\mu$s | ±10% | 1.97%$CH_2O$-Ar | |
| BCO.T3 | ST time to $\frac{1}{2}$ CO max at 5.88 atm, 1784 K | 30.7 $\mu$s | ±20% | 1.47%$CH_2O$-0.25%$O_2$-Ar | |
| BCO.T4 | ST time to $\frac{1}{2}$ CO max at 1.89 atm, 1442 K | 158 $\mu$s | ±10% | 0.49%$CH_2O$-1.98%$O_2$-Ar | |
| BCO.T5 | ST time to $\frac{1}{2}$ CO max at 0.91 atm, 1768 K | 27.2 $\mu$s | ±20% | 1%$CH_2O$-5.96%$O_2$-Ar | |
| BCO.T6 | ST time to $\frac{1}{2}$ CO max at 2 atm, 1515 K | 115 $\mu$s | ±25% | 1%$CH_2O$-0.6%$O_2$-Ar | |
| BCO.T7 | ST time to $\frac{1}{2}$ CO max at 1.51 atm, 1720 K | 32.7 $\mu$s | ±10% | 1.5%$CH_2O$-1.5%$O_2$-Ar | |
| BCH2O.T1 | ST time to $\frac{1}{2}$ CO max at 1.55 atm, 1256 K | 282 $\mu$s | ±10% | 4%$CH_2O$-1%$O_2$-Ar | Hidaka et al. 1993 |
| BCH2O.T2 | ST time to $\frac{1}{2}$ CO max at 2.31 atm, 1591 K | 20.8 $\mu$s | ±20% | | |
| BCH2O.T3 | ST time to $\frac{1}{2}$ CO max at 1.81 atm, 1419 K | 126 $\mu$s | ±10% | 2%$CH_2O$-1%$O_2$-Ar | |

**Table 3**
*Continued.*

| | | | | | |
|---|---|---|---|---|---|
| SR.10c | Flow Reactor $T$ @ $CO_2$ = 500 ppm at 1.17 atm, 2124 K | 1439 K | ±10% | 0.147%$CH_4$-0.3%$O_2$-1.9%$H_2O$-$N_2$ | Glarborg et al. 1995 |
| NFR4 | Flow Reactor NO conc at exit at 1.04 atm, 1323 K | 563 ppm | ±30% | 0.29%$CH_4$-0.03%$C_2H_6$-0.51%$O_2$-0.01%NO-2.2%$H_2O$-$N_2$ | Glarborg et al. 1997 |
| NFR5 | Flow Reactor HCN conc at exit at 1.04 atm, 1323 K | 29 ppm | ±10% | | |
| NFR1 | Flow Reactor HCN conc at exit at 1.05 atm, 1165 K | 0.16 ppm | ±0.06 | 0.03%HCN-0.17%CO-2.4%$O_2$-2.8%$H_2O$-$N_2$ | Miller, Glarborg et al. 1994 |
| NFR2 | Flow Reactor NO conc at exit at 1.05 atm, 1165 K | 0.18 ppm | ±0.03 | | |
| NFR3 | Flow Reactor $N_2O$ conc at exit at 1.05 atm, 1165 K | 0.14 ppm | ±0.05 | | |
| F1 | Flame speed at 1 atm, $\phi = 0.98$ | 35.6 cm/s | ±2 | $CH_4$-air | Egolfopoulos et al. 1989-97 |
| F2 | Flame speed at 1 atm, $\phi = 1.43$ | 12.4 cm/s | ±2 | | |
| F3 | Flame speed at 1 atm, $\phi = 0.67$ | 13.5 cm/s | ±2 | | |
| F4 | Flame speed at 3 atm, $\phi = 1$ | 22.7 cm/s | ±2 | | |
| StF8 | Flame speed at 1 atm, $\phi = 1$ | 40.2 cm/s | ±2 | $C_2H_6$-air | |
| F6 | Flame speed at 19.7 atm, $\phi = 1$ | 20.4 cm/s | ±2 | $CH_4$-air | Just 1994 |
| SF7 | Flame speed at 1 atm, $\phi = 1.69$ | 180 cm/s | ±2 | 20.8%CO-20.8%$H_2$-air | McLean et al. 1994 |
| SNO.C11 | Flame max NO conc at 0.033 atm | 17.5 ppm | ±15% | 13.8%$CH_4$-25.9%$O_2$-$N_2$ | SRI 1996-97 |
| SCH.C11 | Flame max CH conc at 0.033 atm | 11.3 ppm | ±10% | | |
| SCH.C12 | Flame max CH conc at 0.033 atm | 4.1 ppm | ±0.7 | 10%$CH_4$-24.7%$O_2$-$N_2$ | |
| SCH.C13 | Flame max CH conc at 0.0395 atm | 21.5 ppm | ±2.9 | 16.3%$CH_4$-25.5%$O_2$-$N_2$ | |
| NF6 | Flame max NO conc at 25 Torr | 1580 ppm | ±160 | 28%$H_2$-9%$O_2$-2%HCN-Ar | Sandia 1984 |
| NF7 | Flame $\frac{[CN]_{max}}{[CN]_{5cm}}$ at 25 Torr | 2.04 | ±15% | | |
| NF11 | Flame $\frac{[CH]_{X=1.2\%NO}}{[CH]_{X=0}}$ at 10 Torr | 0.95 | ±0.1 | 19%$CH_4$-38%$O_2$-X-Ar | Williams, Fleming 1994 |
| NF12/13 | Flame $\frac{[CN]_{X=1.2\%NO}}{[CN]_{X=1.2\%N_2O}}$ at 10 Torr | 2.9 | ±0.6 | | |

## REFERENCES


[1] M. J. Bayarri, J. O. Berger, J. A. Cafeo, G. Garcia-Donato, F. Liu, J. Palomo, R. J. Parthasarathy, R. Paulo, J. Sacks, and D. Walsh, *Computer model validation with functional output*, Ann. Statist., 35 (2007), pp. 1874–1906.

[2] M. J. Bayarri, J. O. Berger, R. Paulo, J. Sacks, J. A. Cafeo, J. Cavendish, C.-H. Lin, and J. Tu, *A framework for validation of computer models*, Technometrics, 49 (2007), pp. 138–154.

[3] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, Cambridge, UK, 2004.

[4] National Research Council, *Assessing the Reliability of Complex Models: Mathematical and Statistical Foundations of Verification, Validation, and Uncertainty Quantification*, The National Academies, Washington, D.C., 2012.

[5] B. J. Debusschere, H. N. Najm, P. P. Pébay, O. M. Knio, R. G. Ghanem, and O. P. Le Maître, *Numerical challenges in the use of polynomial chaos representations for stochastic processes*, SIAM J. Sci. Comput., 26 (2004), pp. 698–719, doi:10.1137/S1064827503427741.

[6] D. E. Edwards, D. Yu. Zubarev, A. Packard, W. A. Lester, and M. Frenklach, *Interval prediction of molecular properties in parameterized quantum chemistry*, Phys. Rev. Lett., 112 (2014), 253003.

[7] R. Feeley, M. Frenklach, M. Onsum, T. Russi, A. Arkin, and A. Packard, *Model discrimination using data collaboration*, J. Phys. Chem. A, 110 (2006), pp. 6803–6813.

[8] R. Feeley, P. Seiler, A. Packard, and M. Frenklach, *Consistency of a reaction dataset*, J. Phys. Chem. A, 108 (2004), pp. 9573–9583.

[9] R. P. Feeley, *Fighting the Curse of Dimensionality: A Method for Model Validation and Uncertainty Propagation for Complex Simulation Models*, Ph.D. thesis, University of California, Berkeley, CA, 2008.

[10] A. Forrester, A. Sobester, and A. Keane, *Engineering Design via Surrogate Modelling: A Practical Guide*, Wiley, Chichester, UK, 2008.

[11] E. Frazzoli, Z. H. Mao, J. H. Oh, and E. Feron, *Resolution of conflicts involving many aircraft via semidefinite programming*, AIAA J. Guidance Control, 24 (2001), pp. 79–86.

[12] M. Frenklach, *Transforming data into knowledge—Process informatics for combustion chemistry*, Proc. Combust. Inst., 31 (2007), pp. 125–140.

[13] M. Frenklach, A. Packard, and R. Feeley, *Optimization of reaction models with solution mapping*, in Modeling of Chemical Reactions, R. W. Carr, ed., Elsevier, Amsterdam, 2007, Ch. 6, pp. 243–291.

[14] M. Frenklach, A. Packard, and P. Seiler, *Prediction uncertainty from models and data*, in Proceedings of the 2002 American Control Conference (Anchorage, AK), IEEE, New York, 2002, pp. 4135–4140.

[15] M. Frenklach, A. Packard, P. Seiler, and R. Feeley, *Collaborative data processing in developing predictive models of complex reaction systems*, Int. J. Chem. Kinet., 36 (2004), pp. 57–66.

[16] A. Gelman, *Prior distributions for variance parameters in hierarchical models*, Bayesian Anal., 1 (2006), pp. 515–533.

[17] A. Gelman, J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin, *Bayesian Data Analysis*, CRC Press, Boca Raton, FL, 2013.

[18] M. Goldstein, *External Bayesian analysis of computer simulators*, in Bayesian Statistics 9, J. M. Bernardo, M. J. Bayarri, J. O. Berger, A. P. Dawid, D. Heckerman, A. F. M. Smith, and M. West, eds., Oxford University Press, London, 2010 (with discussion).

[19] X. Huan and Y. M. Marzouk, *Simulation-based optimal Bayesian experimental design for nonlinear systems*, J. Comput. Phys., 232 (2013), pp. 288–317.

[20] M. C. Kennedy and A. O'Hagan, *Bayesian analysis of computer code outputs*, J. Roy. Stat. Soc. B, 63 (2001), pp. 425–464.

[21] G. Li, C. Rosenthal, and H. Rabitz, *High dimensional model representation*, J. Phys. Chem., 105 (2001), pp. 7765–7777.

[22] D. J. Lunn, A. Thomas, N. Best, and D. Spiegelhalter, *WinBUGS—a Bayesian modelling framework: Concepts, structure and extensibility*, Stat. Comput., 10 (2000), pp. 325–337.

[23] A. O'Hagan and M. West, *The Oxford Handbook of Applied Bayesian Analysis*, Oxford University, Oxford, UK, 2014.

[24] J. Pedel, J. N. Thornock, and P. J. Smith, *Ignition of co-axial turbulent diffusion oxy-coal jet flames: Experiments and simulations collaboration*, Combust. Flame, 160 (2013), pp. 1112–1128.
[25] T. Russi, A. Packard, R. Feeley, and M. Frenklach, *Sensitivity analysis of uncertainty in model prediction*, J. Phys. Chem. A, 112 (2008), pp. 2579–2588.
[26] T. Russi, A. Packard, and M. Frenklach, *Uncertainty quantification: Making predictions of complex reaction systems reliable*, Chem. Phys. Lett., 499 (2010), pp. 1–8.
[27] P. Seiler, M. Frenklach, A. Packard, and R. Feeley, *Numerical approaches for collaborative data processing*, Optim. Eng., 7 (2006), pp. 459–478.
[28] G. P. Smith, M. Frenklach, R. Feeley, A. Packard, and P. Seiler, *A system analysis approach for atmospheric observations and models: The mesospheric $HO_x$ dilemma*, J. Geophys. Res. (Atmospheres), 111 (2006), D23301.
[29] G. P. Smith, D. M. Golden, M. Frenklach, N. W. Moriarty, B. Eiteneer, M. Goldenberg, C. T. Bowman, R. K. Hanson, S. Song, W. C. Gardiner, Jr., V. V. Lissianski, and Z. Qin, *GRI-Mech* 3.0, http://www.me.berkeley.edu/gri_mech/.
[30] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2014.
[31] D. R. Yeates, P. R. Westmoreland, W. Li, W. Speight, T. Russi, A. Packard, and M. Frenklach, *Integrated data-model analysis facilitated by an instrumental model*, Proc. Combust. Inst., 35 (2015), pp. 597–605.
[32] T.-M. Yi, M. Fazel, X. Liu, T. Otitoju, J. Goncalves, A. Papachristodolou, S. Prajna, and J. Doyle, *Application of robust model validation using SOSTOOLS to the study of G-protein signaling in yeast*, in Proceedings of Foundations of System Biology in Engineering, 2005, pp. 133–136.
[33] X. You, A. Packard, and M. Frenklach, *Process informatics tools for predictive modeling: Hydrogen combustion*, Int. J. Chem. Kinet., 44 (2012), pp. 101–116.
[34] X. You, T. Russi, A. Packard, and M. Frenklach, *Optimization of combustion kinetic models on a feasible set*, Proc. Combust. Inst., 33 (2011), pp. 509–516.