Uncertainty Quantification in Chemical Modeling

A. Mirzayeva1*, N.A. Slavinskaya1, M. Abbasi1, J.H. Starcke1, W. Li2, M. Frenklach2

1German Aerospace Center (DLR), Institute of Combustion Technology, 70569, Stuttgart, Germany
2Mechanical Engineering, University of California at Berkeley, Berkeley, CA 94563, USA

Abstract

A module of PrIME automated data-centric infrastructure, Bound-to-Bound Data Collaboration (B2BDC), was used for the analysis of systematic uncertainty and data consistency of the H2/CO reaction model (73/17). In order to achieve this purpose, a dataset of 167 experimental targets (ignition delay time and laminar flame speed) and 55 active model parameters (pre-exponent factors in the Arrhenius form of the reaction rate coefficients) was constructed. Consistency analysis of experimental data from the composed dataset revealed disagreement between models and data. Two consistency measures were applied to identify the quality of experimental targets (Quantities of Interest, QoI): scalar consistency measure, which quantifies the tightening index of the constraints while still ensuring the existence of a set of the model parameter values whose associated modeling output predicts the experimental QoIs within the uncertainty bounds; and a newly-developed method of computing the vector consistency measure (VCM), which determines the minimal bound changes for QoIs initially identified as inconsistent, each bound by its own extent, while still ensuring the existence of a set of the model parameter values whose associated modeling output predicts the experimental QoIs within the uncertainty bounds. The consistency analysis suggested that elimination of 45 experimental targets, 8 of which were self- inconsistent, would lead to a consistent dataset. After that the feasible parameter set was constructed through decrease uncertainty parameters for several reaction rate coefficients. This dataset was subjected for the B2BDC framework model optimization and analysis on. Forth methods of parameter optimization were applied, including those unique in the B2BDC framework. The optimized models showed improved agreement with experimental values, as compared to the initially-assembled model. Moreover, predictions for experiments not included in the initial dataset were investigated. The results demonstrate benefits of applying the B2BDC methodology for development of predictive kinetic models.

1. Introduction

To firmly develop predictive chemical reaction models for complex chemical systems, alliance of excessive amounts of theoretical, computational, and experimental data collected by plentiful number of scientists and researchers is required. The integration involves evaluation of the data consistency, validation of models, and uncertainties quantification for model predictions. This approach to the development of mechanistic reaction models includes assuming the reaction mechanism and comparing the outcome predictions of the generated model to available experimental observations. Mostly, such comparisons lead to mixed outcomes: some demonstrate a fairly close agreement whereas some others do not. In the latter case, the possible inconsistency acquired between the experiment and the model is contended to indicate either that the model is inadequate or that the experiment (or, rather, its interpretation) is incorrect.

Bound-to-Bound Data Collaboration, hereafter abbreviated as B2BDC, is an optimization-based framework for integrating reaction models with experimental data from numerous sources to research and analyze their collective information content. The methodology analyzes consistency among experimental data and reaction models, searches and examines sources of inconsistency, distinguishes differing models, predicts model output interval, and analyzes the sensitivity of uncertainty propagation [1–8].

*Corresponding author. E-mail: Aziza.Mirzayeva@dlr.de © 2018 Al-Farabi Kazakh National University
The common settings are as follows. A physical process is considered that can be represented by a computational model, $M(x)$, with parametric dependence on unknown/uncertain physical parameters, $x$. We assume a prior knowledge (assumptions) on the parameters domain, so constraining each $x$ to an interval $[x_{\text{min}}, x_{\text{max}}]$ and jointly to a hypercube $x \in H$. At the same time, a collection of experimental observations, hereafter referred as Quantities of Interest (QoI), with corresponding uncertainty boundaries, assessed as lower and upper bounds on each QoI values, i.e., $L_e$ and $U_e$ for each $e$-th QoI. The numerical models must give results that are consistent with the QoI bounds in the experimental reports. Thus additional constraints that the true parameters must satisfy are

$$L_e \leq M(x) \leq U_e \quad \text{for all } e. \tag{1}$$

The subset of $H$ appeasing (1) is called the feasible set, $\Phi$, of parameters. $\Phi$ is barely all parameter values that all together satisfy the prior knowledge and are consistent with all prediction models and factual observed results. A parameter value that is not in $\Phi$ is at odds with at least one of these constraints. The mathematical methodology B2BDC appeals to constrained optimization over the feasible set $\Phi$,

$$\min_{x \in F} f(x) \quad \max_{x \in F} f(x)$$

where $f$ is a function of interest, and the calculated $\min$ and $\max$ establish the “to-bound” aspect of the nomenclature. Simply saying, the bounds that describe the prior knowledge and the bounds on QoIs are mapped into bounds on prediction. Some general examples are described below.

### 1.1. Dataset consistency

The feasible set represents the complete collaborative knowledge incorporated in a dataset, and questions in the B2BDC approach are posed as problems of optimization over the feasible set. This subsequently leads to the question of dataset consistency. In order to numerically assess a dataset consistency, a consistency measure was introduced that is able to answer the question on the largest possible percentage of uncertainty lessening to achieve a feasible parameter vector. Correlated with a given dataset $D$, it is noted CD and formulated as an optimization problem,

$$C_D = \max_{y, x \in F} \gamma, \quad \text{subject to (for all } e):$$

$$(1-\gamma)\frac{U_e - L_e}{2} \leq M_e(x) - \frac{U_e + L_e}{2} \leq (1-\gamma)\frac{U_e - L_e}{2}. \tag{2}$$

In this definition, the original constraints $L_e \leq M_e(x) \leq U_e$ are augmented with a scalar coefficient $\gamma$, where positive values of $\gamma$ entail the constraint tightening, and negative $\gamma$ values entail the constraint loosening. The consistency measure quantifies the tightening index of the constraints albeit ensuring the existence of a set of parameter values whose corresponding model predictions describes the experimental QoIs within the bounds. The dataset is considered consistent in case of the non-negative consistency measure, and is inconsistent elseways.

To continue with the analysis, we employed a newly-developed method of computing the vector consistency measure (VCM), similar to Eq. 2, but with original constraints augmented with individual relaxations $\gamma_e$ for each bound. The VCM method determines the minimal bound changes, each bound by its own extent, that result in dataset consistency.

### 1.2. Model prediction

Suppose physical parameters (set of conditions) not exercised experimentally but with a property $P$ predicted by model $M_p$. One of the main questions of scientific analysis is on the range of values this model exhibits over the domain of feasible parameter values. Strictly speaking, the question of what is the prediction interval for property $P$ that is consistent with all of the model/observation pairs in the dataset? This is here assigned as model prediction.

In the B2BDC computation, this question is expressed into two optimization problems for the lower and upper interval endpoints, $L_p$ and $U_p$,

$$L_p := \min_{x \in F} M_p(x) \quad U_p := \max_{x \in F} M_p(x). \tag{3}$$

The length $U_p - L_p$ measures the uncertainty amount inMP’s value with a condition that the true parameter vector is a part of the feasible set $\Phi$.

The suggested analysis propose a sequential procedure with step-by-step outliers identification and sources inspection. A specific direction deter-
mined for improving dataset consistency and an estimate of the extent of possible improvement are provided by the analysis. In general, this computational method offers a tool for estimating experimental observations as well as for model building and improvement.

In the current work, Data Collaboration module of PrIme1 was applied to the H<sub>2</sub>/CO sub-system of the reaction kinetic model [9] to:

1) investigate the computational algorithms, modules and user interface incorporated in PrIME;
2) investigate an algorithm of the consistent dataset construction;
3) test the different chemical kinetic model optimization strategies.

2. PrIME DataSet

2.1. Reaction Model

The H<sub>2</sub>/CO sub-model (6 elements, 17 species, 73 reactions) of C<sub>1</sub>–C<sub>2</sub> reaction mechanism [9] with improvements performed on the data followed from the studies [11‒21] was used to perform systematic uncertainty and consistency analyses with the Data Collaboration module of PrIME to obtain the feasible set sampling for the base H<sub>2</sub>/CO chemistry of DLR reaction data base. The review of the reaction rate coefficients in the analyzed sub-model was performed with additional focus on the pressure depending and multichannel reactions.

The uncertainty factors for rate coefficients were assumed equal to the proposed ones in the sources or evaluated from statistical treatment of the different data:

\[
\begin{align*}
    f_u(T) &= \frac{k_{\text{upper}}(T)}{k_0(T)}, \\
    f_l(T) &= \frac{k_0(T)}{k_{\text{lower}}(T)},
\end{align*}
\]  

(4)

where \( k_0 \) is the nominal rate coefficient, \( k_{\text{lower}} \) and \( k_{\text{upper}} \) are lower and upper bounds.

The active parameters were identified via sensitivity analysis accomplished with uncertainties, represented by the lower and upper bounds. They were assumed equal to those proposed in literature sources or evaluated from a statistical treatment of the literature data.

2.2. Ignition delay time QoI

Before the adjustment of the kinetic parameters can be performed in order to meet the ignition targets it is ultimately required to quantify the uncertainties within the shock tube. In case that it is not possible to describe some active phenomena during the experiments within the shock tube under the assumption of homogeneous conditions (constant \( V, U \) system) after the reflection of the shock, these are classified as “non-idealities” during the experiments in the shock tube [22–31]. Both, setup-dependent effects and phenomena due to energy release can raise the non-idealities and affect the instrument readings, contributing to the uncertainty of experimental results. Considering the syngas mixtures, the two states of ignition should be identified. The first one is the weak ignition: the non-uniform and distributed ignition. The second one is the strong ignition: it is induced by auto ignition at the end wall of the tube and propagates throughout the mixture [27]. Despite the fact that the non-idealities existing in shock tubes have been well discussed [22–31], the quantitative consideration of their impact on the presented data of the ignition delay occurs to be a critical issue. For the evaluation of the uncertainty bounds of the measurements included in the dataset, the empirical algorithm is proposed. Therefore, the most strong non-ideality phenomena [22–31] were identified throughout the investigations. Also the factors related to facility and fuel, which influence these phenomena, have been determined.

It was observed that data collected during experiments with large diameter shock tubes (~ 100 mm), dilute fuel/oxidizer mixtures in monoatomic gases, and short test times (less than about 500 \( \mu \)s) show the lowest uncertainty level. A correlation of the diameter of the shock-tube and weak ignition is found: the bigger diameter leads to an ignition delay similar to that of a homogeneous reactor.

In the best case (diluted mixture, strong ignition, shock tube diameter > 100 mm, \( \tau_{\text{mean}} = 50 \text{ ms} – 500 \text{ ms} \), length of driven section > 8 m) the assumption was made, that the uncertainty is at ~15%. Conditions, deviating from upper mentioned, are evaluated by adding a 5% uncertainty for each criterion not matching the ideal case. For measured ignition delay time longer than one millisecond 5% uncertainty is added for each millisecond. Radical impurities were evaluated as extra 5% uncertainty to the ideal case.

We selected 122 ignition delay targets with obtained uncertainty evaluation from the shock tube experiments [32–39] for the dataset.
2.3. Laminar flame velocity QoI

Flame velocity observations of syngas mixtures at 0.1–0.5 MPa have been analyzed with the help of various techniques known up-to-date [40–41]. There is a scarcity of data on flame speed data at high pressures. The uncertainty of flame speed measurements are assumed by the experimentalists to be in a range between 5–10%. It has to be noted that uncertainty increases with pressure (>0.5 MPa) and fuel-air ratio (φ>2) [40–41].

Uncertainty bounds of experimental data were evaluated from studies [40–41] and analysis of the current syngas atmospheric laminar flame speed data distribution, which can be found in [41]. From the data analysis following from [37–44] the uncertainty of available data can be assumed to be 10% for φ<2, 15% for 2<φ<3, and 20% for φ>3. At higher pressures, the uncertainties for experimental data have been estimated by adding 5%.

The 45 laminar flame speed data included in the dataset are taken from studies [19, 36, 40, 42–50]. They are selected in order to cover a wide range of operating conditions accessible from the literature sources.

A preferred key (or PrIME ID) was determined for each structural element in the chemical reaction model and each experimental target. In this way, each structural element has a “pointer” to the referenced information and/or file. All the experimental and model data were documented in the PrIME Data Warehouse [1]. Selected for analysis experimental QoI are described in the data Attribute files of the PrIME data collection [1]. These QoI together with the corresponding model Me(x) and the experimental and parameter bounds form a dataset. The complete model and experimental data are available in the PrIME Data Warehouse [1].

3. General Results

The ignition delay time and laminar flame speed data was modeled with numerical tools of PrIME [1], numerical packages CHEMKIN II [51] and Chemical Workbench [52]. The ignition delay time was computationally defined by the peak in the OH or OH* concentration, temperature, or pressure. It is pointed in the attribute files of PrIME Warehouse. The thermal diffusion model was applied for calculation of one-dimensional freely propagating laminar premixed flame using CHEMKIN II with over 1000 grid points for each condition.

3.1. Dataset Consistency (Data Quality)

We began the analysis by employing Eq. 2 with the initial dataset, which included all 167 QoI (122 ignition delays and 45 laminar flame speeds) and 55 active parameters. The results indicated a massive inconsistency. Eight QoI, those listed in Table, were found to be self-inconsistent. These were the ignition delay times that were not able to be reproduced within their respective uncertainty bounds by the model employing rate coefficients within their respective uncertainty bounds, H. These eight self-inconsistent QoI were removed from the initially constructed dataset. The latter, however, still remained an inconsistent dataset.

Table
The 8 Self-inconsistent QoI

<table>
<thead>
<tr>
<th>T3 (K)</th>
<th>p3 (atm)</th>
<th>φ</th>
<th>Target PrIME ID</th>
<th>Estimated uncertainty (%)</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1263</td>
<td>1.1</td>
<td>0.5</td>
<td>a00000309</td>
<td>30</td>
<td>[32]</td>
</tr>
<tr>
<td>1695</td>
<td>1.6</td>
<td>0.5</td>
<td>a00000352</td>
<td>30</td>
<td>[36]</td>
</tr>
<tr>
<td>2004</td>
<td>1.6</td>
<td>0.5</td>
<td>a00000355</td>
<td>25</td>
<td>[36]</td>
</tr>
<tr>
<td>1975</td>
<td>1.6</td>
<td>0.5</td>
<td>a00000358</td>
<td>25</td>
<td>[36]</td>
</tr>
<tr>
<td>1436</td>
<td>1.6</td>
<td>0.5</td>
<td>a00000359</td>
<td>25</td>
<td>[36]</td>
</tr>
<tr>
<td>1027</td>
<td>1.6</td>
<td>0.5</td>
<td>a00000360</td>
<td>35</td>
<td>[36]</td>
</tr>
<tr>
<td>1883</td>
<td>1.6</td>
<td>0.5</td>
<td>a00000503</td>
<td>30</td>
<td>[38]</td>
</tr>
<tr>
<td>1008</td>
<td>1.6</td>
<td>0.5</td>
<td>a00000504</td>
<td>50</td>
<td>[38]</td>
</tr>
</tbody>
</table>

To continue with the analysis, we employed a newly-developed method of computing the vector consistency measure (VCM), similar to Eq. 2, but with original constraints augmented with individual relaxations γe for each bound [10]. The VCM method determines the minimal bound changes, each bound by its own extent, that result in dataset consistency. Its application identified such a dataset-consistency point by changing 30 ignition delay times and 7 laminar flame speeds, shown in Figs. 1 and 2, respectively. We emphasize that the VCM-identified feasible parameter set is a single point in H. As this point possesses some optimal attributes, we compare the model predictions obtained with this set of parameters, DLR-SynG 1 dataset, to the optimization results.

To proceed with further features of the B2BDC framework, we created a new dataset by removing the 37 QoI identified by VCM, thus forming the DLR-SynG 2 dataset contacting 122 QoI. This latter dataset is consistent, meaning that all its 122 QoI are consistent with each other and with the 55 active parameters.
3.2. Feasible set construction

While $H$ designates prior information, feasible set $\Phi$ summarizes posterior information: all parameter value combinations that satisfy their own bounds and also the QoI bounds. The size and shape of $\Phi$ compared to those of $H$ represent information gained as a result of the B2BDC analysis.

Projection of $\Phi$ on each of the $x$’s yields the posterior range of the parameter uncertainty [3].
Those that changed are reported in Fig. 3. For the rest of the parameters, the posterior ranges were the same as the prior ones, indicating that the experimental data included in the present analysis did not aid in narrowing down the uncertainty ranges of these parameters individually. However, such an outcome does not necessarily imply no information gain for a given parameter: while the extreme parameter values (bounds) may not change, the feasible set may, and usually does, eliminates some combinations of these parameters with others, which is addressed next.

4. Parameter optimization

While the primary focus of the B2BDC framework is on prediction over the feasible set, it also supports parameter optimization [53]. Four sets of optimized model parameters were investigated and inter-compared in the present study. The first approach is LS-H, a (weighted) least-squared fit constraining parameter values to their initially assessed uncertainty ranges, $H$. This is now a common approach [6, 8, 53]. B2BDC supports two more refined methods of optimization [53], LS-F and 1N-F, where the objective is minimized with $x$’s being constrained to the feasible set $\Phi$. The three problems are easily expressed as mathematical optimizations. The LS methods minimize the familiar sum of weighted least-squared deviations between the surrogate model prediction and the reported measured value, $y_e$. The difference lies in where the search takes place: LS-H considers all of $H$ while LS-F restricts the search to $F$,

$$LS - H : \min_{x \in H} \sum_{e} w_e [ M_e(x) - y_e ]^2$$

$$LS - F : \min_{x \in F} \sum_{e} w_e [ M_e(x) - y_e ]^2$$

By contrast, the 1N-F problem treats the nominal parameter vector, the starting set of parameter values ($x_0 = 0$), as “preferred”. As we have shown in previous sections, this parameter set lies outside the feasible region $\Phi$. The goal of the 1N-F method is to find with least number of changes to $x_0$ a parameter vector that is feasible. Mathematically, the one-norm is a well-known approximation to enforce such sparsity, i.e.,

$$1N - F : \min_{x \in F} \| x - x_0 \|_1$$

The LS-F and 1N-F optimizations were performed with the final dataset, as the two methods are designed to work with an existing feasible set. Inspection of the results highlights several features. All optimization methods result in parameter sets that produce a better agreement with experiment than the original set, composed of literature recommendations. The LS-H optimization, constrained only to the prior uncertainty ranges of parameters, results in the lowest average deviation, as expected, but at the expense of violating uncertainty bounds of 13 experimental QoI.

The average deviation produced by LS-F is larger but not significantly than that of LS-H. The 1N-F method gives a larger average deviation, yet it changes the least number of variables. The LS-F and 1N-F optimization methods, with additional constraints to the QoI uncertainties, do not violate any of the QoI bounds by design unlike to the LS-H. That demonstrates the main difference between two approaches: LS-H optimization can be identified rater as a fitting.

Some of the individual comparisons are shown in Figs. 4–6, with the inclusion of the most recent literature model [54]. Experimental targets of the DLR-SynG dataset in these figures are designated by a star, the 8 self- inconsistent QoI (excluded from the dataset, Table) are colored red and those excluded with VCM (Figs. 1 and 2) are colored green.

The visual observation is that all the optimized models seem to perform with about the same overall quality: some models do better for one set of conditions, while other are closer to other experimental data. The shock-tube ignition delay times show a larger variation between different models. The problem here could lie with the incomplete instrument model used in the simulation of ignition phenomena, as it does not capture the “non-idealities” of shock-tube experiments with sufficient detail, or the development of a mild-ignition regime, which is not entirely driven by chemistry. These factors are especially under suspicion in the inconsistent ignition-delay targets. Generally, the laminar flame speeds are predicted better by all models, with all simulations falling within the uncertainties bounds of experimental observations, reflecting perhaps the higher experimental accuracy of the measurements.

Figures 4–6 demonstrate the benefits of optimization methods LS-F and 1N-F and generally of the B2BDC methodology in comparison to the “conventional” optimization, LS-H, or performed in study [54].
Fig. 4. Ignition delay times: symbols, experimental data [32, 35]; initial model, black line; Varga et al. [54] model, gray line; LS-H, red dotted line; VCM, red dash-dotted line; LS-F, blue dashed line; 1N-F, blue short-dash line. Black
Fig. 5. Ignition delay times: symbols, experimental data [36‒39]; initial model, black line; Varga et al. [54] model, gray line; LS-H, red dotted line; VCM, red dash-dotted line; LS-F, blue dashed line; 1N-F, blue short-dash line. Black stars are targets of DLR-SynG 2 dataset; red stars are self-inconsistent targets; green stars are targets deleted from DLR-SynG 1 dataset.
5. Conclusions

An optimization-based framework B2BDC of an automated data-centric infrastructure, Process Informatics Model (PrIMe) was applied to the syngas reaction mechanism analysis. For this reason, we constructed a dataset based on pertinent experimental observations, chemical kinetics model, and the associated uncertainties. The experimental Quantities of Interest (QoI) were selected through evaluation of ignition delay time and laminar flame speed uncertainties. The composed dataset was subjected to consistency analysis. One outcome of the analysis was identification of a set of experimental QoI that were most difficult or impossible to match with the model; they were removed from the dataset for future investigation. The final consistent dataset with 122 experimental QoI and 55 active variables was used for model optimization on the feasible parameter set. The optimized syngas models produced with B2BDC framework demonstrated an improved agreement with the dataset QoI. The results of the present work, however, demonstrate that the LS-H optimization may miss some critical information of the model-data system. Only parameter optimization performed on the feasible set produces a reaction model which describes the experimental measurements not included in the analysis as well as experimental targets from used dataset. The obtained optimized parameter values indicated parameter inadequacy, and the correlation analysis highlighted the direction of possible parameter modifications and model improvement.

Acknowledgments

The work at UC Berkeley was supported by the U.S. Department of Energy, National Nuclear Security Administration, under Award Number DE-NA0002375.

References


