Proceedings of the World Congress on Momentum, Heat and Mass Transfer (MHMT'16) Prague, Czech Republic – April 4 – 5, 2016 Paper No. CSP 118 DOI: 10.11159/csp16.118

Consistent Chemical Mechanism from Collaborative Data Processing

Nadezda Slavinskaya¹, Jan-Hendrik Starcke¹, Mehdi Abbasi¹, Aisulu Tursynbay¹, Uwe Riedel¹, Wenyu Li², Jim Oreluk², Arun Hedge², Andrew Packard², Michel Frenklach²

¹German Aerospace Center, Institute of Combustion Technology Pfaffenwaldring 38-40, 70569 Stuttgart, Germany Nadja.Slavinskaya@dlr.de ²Mechanical Engineering, University of California at Berkeley, Berkeley, 6105B Etcheverry Hall, Mailstop 1740, CA 94720, USA frenklach@berkeley.edu; wenyuli@berkeley.edu

Abstract - Numerical tool of Process Informatics Model (PrIMe) is mathematically rigorous and numerically efficient approach for analysis and optimization of chemical systems. It handles heterogeneous data and is scalable to a large number of parameters. The Bound-to-Bound Data Collaboration module of the automated data-centric infrastructure of PrIMe was used for the systematic uncertainty and data consistency analyses of the H_2/CO reaction model (73/17) and 94 experimental targets (ignition delay times). The empirical rule for evaluation of the shock tube experimental data is proposed. The initial results demonstrate clear benefits of the PrIMe methods for an evaluation of the kinetic data quality and data consistency and for developing predictive kinetic models.

Keywords: PrIMe, syngas, reaction mechanism, uncertainty, consistency.

1. Introduction

To reliably develop predictive reaction models for complex chemical systems requires integration of large amounts of theoretical, computational, and experimental data collected by numerous researchers. The integration entails assessment of the consistency of the data, validation of models, and quantification of uncertainties for model predictions.

Bound-to-Bound Data Collaboration, abbreviated hereafter as B2B-DC, is an optimization-based framework for combining models and experimental data from multiple sources to explore their collective information content; to test consistency among data and models; to identify sources of inconsistency; to discriminate among differing models [1-8].

The general setting is as follows: numerical model, M(x) of a physical process with parametric dependence on unknown/uncertain physical parameters, x; prior knowledge/assumptions on the domain of parameters, $[x_{min}, x_{max}]$, and all together on a hypercube $x \in \Box$; a collection of experimental observations e, referred hereafter as Quantities of Interest (QoI), with respective uncertainties, assessed as lower and upper bounds on the observed QoI values, i.e., L_e and U_e for each e-th QoI. The computational models 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

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

The subset of *H* satisfying (1) is called the *feasible set*, Φ , of parameters.

Usually, the models M(x) are complex and take the form of differential equations that do not possess a closed-form solution. The integral part of the B2B-DC framework, one that makes the approach practical, is approximation of the M(x) outputs for given QoI by quadratic surrogate models. The above can now be summarized in the definition of the feasible set

$$\Phi := \{ x \in H : L_e \le M_e(x) \le U_e \quad \forall e \} ,$$
⁽²⁾

where $M_e(x)$ designates a surrogate model of *e*-th QoI. Φ defined by (2) 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 Φ is at odds with at least one of these constraints. The true parameters must result in model predictions of all training experiments that are within the measurement bounds declared by the experimenters.

Together, these are the "bounds" that define Φ . The mathematical methodology of B2B-DC invokes constrained optimization over the feasible set Φ , $\left[\min_{x\in\Phi} f(x)\max_{x\in\Phi} f(x)\right]$, where f is a function of interest. *Dataset consistency* analysis identifies the validity and integrity of the data included in the analysis and identifies

Dataset consistency analysis identifies the validity and integrity of the data included in the analysis and identifies experimental QoI or model parameters *x* causing the *dataset D* (i.e., M(x), QoIs and their uncertainties) to be inconsistent. This analysis enhances the quality of the experimental data adopted for model parameter optimization over the feasible region of the parameter space. To assess it numerically, a consistency measure was introduced [2]. Associated with a given dataset *D*, it is notated C_D and posed as an optimization problem,

$$C_{\rm D} = \max_{\gamma, x \in \mathcal{F}} \gamma, \quad \text{subject to (for all } e):$$

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

$$M_e(x) - \frac{U_e + L_e}{2} \le (1 - \gamma) \frac{U_e - L_e}{2}.$$
(3)

In this definition, the original constraints $L_e \leq M_e(x) \leq U_e$ are augmented with a scalar γ , where positive values of γ 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 the bounds) the experimental QoIs. The dataset is consistent if the consistency measure is nonnegative, and is inconsistent otherwise.

Model prediction. The B2B-DC computation expresses the prediction interval for property P by model M_P into two optimization problems for the lower and upper interval endpoints, L_P and U_P ,

$$L_{p} \coloneqq \min_{x \in F} M_{p}(x)$$

$$U_{p} \coloneqq \max_{x \in F} M_{p}(x)$$
(4)

The length U_P-L_P quantifies the amount of uncertainty in M_P 's value conditioned on the fact that the true parameter vector is contained in the feasible set Φ .

The results of the proposed analysis suggest a sequential procedure with step-by-step identification of outliers and inspection of the causes. The analysis identifies a specific direction to follow for improving dataset consistency and provides an estimate of the extent of possible improvement. Altogether, this numerical approach offers a tool for assessing experimental observations and model building and improvement.

In the present paper the B2B-DC module of PrIMe [1] was applied to the H_2 /CO sub-system of the kinetic model [9] to test the numerical algorithms, modules and user interface of the PrIMe and to perform a systematic uncertainty and consistency analysis of the model parameters and related experimental data (ignition delay times), to revise and to optimize the model parameters.

2. PrIMe DataSet

A dataset unit should consist of the measured observation, uncertainty bounds on the measurement and thermo-kinetic data, and a model with the parameter uncertainties. Identification of active parameters via sensitivity analysis and development of a quadratic response surface via computer experiments arranged according to a factorial design. Organized in this manner, the dataset can be subjected to rigorous numerical analysis

2.1. Reaction Model

The CO/H₂ mixture oxidation chemistry is the principal building block in the hierarchy of hydrocarbon oxidation models. In recent years, the role of syngas in sustainable combustion processes triggered further characterization of the CO/H_2 combustion system [10].

The H₂/CO sub-model (6 elements, 17 species, 73 reactions) of C_1 - C_2 reaction mechanism (also the base chemistry of DLR reaction data base) [9] was used to perform systematic uncertainty and consistency analyses with the B2B-DC module of PrIMe. The reaction rate coefficients in the examined sub-model were reviewed with further attention to the pressure depending and multichannel reactions. In comparison to the study [6], the reaction rate coefficients for OH+OH(+M)=H₂O₂(+M) and CO+O(+M)=CO₂(+M) were replaced with values following from [11-13]. The studied sub-model was extended with OH^{*} reaction sub-mechanism from [14] to reproduce more accurately the ignition delay times recorded in shock tubes by the OH^{*} chemiluminescence measurements. Finally, reaction model was presented in the XML format adopted in PrIMe [1]. A preferred key (or PrIMe ID) was prescribed to each structural element in the reaction scheme. Each structural element has a link with the reference information file. Such constructed set of files defines the reaction model M(x) recorded in PrIMe.

2.2. Ignition-delay-time Qol

Quantification of uncertainties in the shock tube is ultimately needed prior to undertaking any tuning of the kinetic parameters to match ignition QoI.

If some active phenomena in the shock tube experiments cannot be described by assuming homogeneous conditions (constant V, U system) behind the reflected shock, they are classified as "non-idealities" in the shock tube experiments [15-24]. Both, facility-dependent effects and energy-release phenomena can increase the non-idealities and influence the instrument readings, thus adding to the uncertainty of experimental data. To evaluate the uncertainty bounds of the measured observations included in the dataset, the empirical algorithm is proposed. For that, the most strong non-ideality phenomena [15-24] were determined across the investigations and the facility-related and fuel-related factors, which affect these phenomena, have been identified.

The dominant non-ideality phenomena were attributed to two gas dynamics effects: i) boundary layer formation after incident shock wave interacts with reflected shock-wave (resulting in inhomogeneities of T and p behind the shock-wave and shock bifurcations); ii) post-shock compression (interaction of the reflected shock-wave with the contact surface). The second most important phenomena influencing the measurements uncertainty in the syngas systems is energy-release: the weak regime (the non-uniform/distributed ignition) and the strong regime (initiated by auto ignition at the end wall of the shock-tube) of ignition [20].

The factors which influence these phenomena are: operating conditions; driven section length; driven section diameter; measured ignition time; mixture dilution and nature of carrier gas (CG). It was found that experimental data obtained by using large diameter shock tubes (~10 cm), dilute fuel/oxidizer mixtures in monoatomic gases, and short test times (less than about 500 μ s) have the lowest uncertainty level. A correspondence with the diameter of the shock-tube and weak ignition is found: a larger diameter leads to an ignition delay close to that of a homogeneous reactor.

It was assumed that in the best case (strong ignition, diluted mixture, $t_{meas} = 50 - 500$ ms, shock tube diameter > 10 cm) the uncertainty can be assumed ~15%. Deviations from these conditions are evaluated by adding a 5% uncertainty for each criterion not satisfied to the ideal case. For measured ignition delay time longer than 1000 µs, 5% uncertainty is added per every 1000 µs. Radical impurities were evaluated as extra 5% uncertainty.

Table 1 displays the selected shock tube experiments [25-29] with the uncertainty intervals evaluated with the proposed empirical rule.

2.3. Active parameters

The surrogate models for numerical algorithms of B2B-DC are generated through modifications of active parameters, i.e., pre-exponential factors of the reaction rate coefficients of the most influential reactions. These reactions are determined through sensitivity analysis performed for each experimental QoI and presented in Table 2.

Ref.	Driven section length, m	Internal diameter, cm	Temperature interval, K	Pressure	φ	Dilution	t _{meas}	Integrated uncertainty %
25	10,7	16,2	< 1000 + 5%	<15 atm	0,5	none + 5%	100-500 µs	30
25	10,7	16,2	< 1000 + 5%	<15 atm	0,5	none + 5%	500-1000 μs + 5%	35
25	10,7	16,2	< 1000 + 5%	<15 atm	0,5	none + 5%	1000-2000 μs + 10%	40
25	10,7	16,2	> 1000	<15 atm	0,5	none + 5%	100-600 μs (+ 5%)	30
26	10,7	16,2	< 1000 + 5%	>15 atm + 5%	0,5	none + 5%	500-1000 μs + 5%	40
26	10,7	16,2	< 1000 + 5%	>15 atm + 5%	0,5	none + 5%	1000-2000 μs + 10%	45
26	10,7	16,2	< 1000 + 5%	>15 atm + 5%	0,5	none + 5%	2000-3000 μs + 15%	50
27	10,7	16,2	< 1000 + 5%	<15 atm	0,5- 1	none + 5%	600-2000 μs + 10%	40
27	10,7	16,2	< 1000 + 5%	<15 atm	0,5- 1	yes	600-1200 μs + 10%	35
28	11,12	9,82	> 1000	>15 atm +5%	0,5	yes	300-500 µs	25
28	11,12	9,82	> 1000	>15 atm + 5%	0,5	yes	500-1000 μs + 5%	30
28	11,12	9,82	> 1000	>15 atm + 5%	0,5	yes	500-1000 μs + 5%	30
29	4,72 + 5%	15,24	> 1000	<15 atm	0,5;1	yes	20 -500µs	25
29	4,72 + 5%	15,24	> 1000	<15 atm	0,5;1	yes	500 - 1000μs + 5%	30
29	4,72 + 5%	15,24	> 1000	>15 atm + 5%	0,5;1	yes	20-500µs	30
29	4,72 + 5%	15,24	> 1000	>15 atm + 5%	0,5;1	yes	500-1000μs + 5%	35
29	4,72 + 5%	15,24	> 1000	>15 atm + 5%	0,5;1	yes	1000- 2000μs + 10%	40

Table 1: Evaluation of uncertainty intervals for the selected shock tube experimental data.

t	Active Variables	#	Active Variables
1	H2+O<=>OH+H	15	H+HO2<=>O2+H2
2	H+O<=>OH	16	H+H2O2<=>HO2+H2
3	H+HO2<=>O+H2O	17	OH+H2O2<=>HO2+H2O
4	H+H2O2<=>OH+H2O	18	HO2+HO2<=>O2+H2O2
5	OH+OH<=>O+H2O	19	O2+CO<=>O+CO2
6	O+H2O2<=>OH+HO2	20	H+HCO<=>H2+CO
7	H+HO2<=>OH+OH	21	OH+CO<=>H+CO2
8	OH+HO2<=>O2+H2O	22	HCO+O2 <=>OH+CO2
9	O+HO2<=>OH+O2	23	HCO+O2 <=>HO2+CO
10	OH+H2<=>H+H2O	24	HO2+CO <=>OH+CO2
11	H+O2<=>HO2	25	O+HCO<=>OH+CO
12	H2+O2<=>2OH	26	O+HCO<=>H+CO2
13	H+O2<=>O+OH	27	HCO<=>H+CO
14	2H<=>H2	28	O+CO<=>CO2

Table 2: Reactions identified for model optimization.

3. General Results

3.1. Dataset Consistency (Data Quality)

The Consistency analysis [2] has been performed for the dataset that included the present reaction model and experimental data from Table 1.

The first results of the dataset consistency analysis indicated the high inconsistency of examined dataset. First, 7 experimental QoI, which did not exhibit a maximum OH* in the calculated OH* profile with the nominal parameter values, were excluded from the dataset, because the ignition delay time could not be fixed in this case. Then, other 5 QoI with relatively large fitting error (> 3% error) have been also deleted from the analysis. All other 83 QoIs of the dataset were fitted with quadratic surrogate models with on-design errors not exciding 1% and off-design error below or about 2%. This investigation revealed that 22 dataset units are self-inconsistent. Self-inconsistency means that no point in the rate constant domain can reproduce the experimental observation within its uncertainty bounds.

To further ensure the consistency of 61 experimental QoI with the studied reaction model, the uncertainty bounds of some experimental QoI have been increased (related to those evaluated with proposed empirical rule, Table 1) through B2B-DC calculations, Fig.1. Finally 43 experimental QoI were included in the further B2B-DC procedure.



Fig. 1: The performed changes of the uncertainty bounds to bring the dataset into consistency.

Fig. 2: Optimal model predictions using optimization methods LS-H, LS-F, and 1N-F.

3.2. Reaction Model Optimization

After a consistent dataset was obtained, the H₂/CO data-mode system was subjected to model-parameter optimization over the feasible region of the parameter space [30]. Optimization on a feasible set using LS-F and 1N-F methods ensured that all model predictions fell within the uncertainties of experimental targets. Method LS-H is a least-squared minimization constrained to the prior-knowledge hypercube, \Box , while method LS-F constrains the minimization to the feasible set, \Box . Method 1N-F is a one-norm minimization constrained to the feasible set, \Box .

Adding experimental uncertainties to parameter uncertainties, the rate coefficients of the 28 most influential reactions, Table 2, have been then optimized, i.e., feasible set for the studied model have been defined. All 43 experimental targets and their corresponding optimal model predictions are shown in Fig. 2.

The ratios of optimal to initial values for 28 rate coefficients obtained with methods LS-H, LS-F and 1N-F are displayed in Fig. 3. Inspection of these results indicates, as expected [30], that the number of the rate constants changed by optimization is the smallest for method 1N-F, whereas methods LS-H and LS-F changed many of them, often pushing their optimized values to the respective upper or lower bounds. Fig.4 demonstrates the ignition delay time [28] calculations performed with original and optimized models.

Fig. 3: Ratios of optimal to initial values for 28 pre-exponential factors of rate coefficients *A* obtained with methods LS-H, LS-F, and 1N-F. Error bars indicate the specified variable ranges.

Fig. 4: Optimal model predictions of data [28] using optimization methods LS-H, LS-F and 1N-F. The bars indicate the specified data and changed ranges.

4. Conclusion

Numerical toll of Process Informatics Model (PrIMe) is mathematically rigorous, numerically efficient approach for analysis of chemical systems handles heterogeneous data and is scalable to a large number of parameters. The B2B-DC module of the automated data-centric infrastructure PrIMe was used to perform systematic uncertainty and consistency analyses of the H₂/CO reaction system (73/17) and 94 experimental targets. For an evaluation of uncertainty bounds of experimental data the empirical rule was developed based on the literature data. The Consistency analysis of the B2B-DC module selected the experimental QoI, which were inconsistent with the dataset units in the given uncertainty bounds. Finally, 43 experimental data of ignition delay times and 28 reaction rate coefficients of active (most influential reactions defined through sensitivity analysis) parameters were included in the dataset and the feasible set of parameter values was found. The B2B-DC framework is applied to the parameter model optimization over the feasible set

It is shown, that B2B-DC module of the PrIMe is effective to examine the influence of parameter and experiment uncertainties on the optimal solution and to identify experimental targets that are most difficult to match as well as model parameter values that are likely to be questionable. The initial results demonstrate clear benefits of applied tool for developing predictive kinetic models.

Acknowledgements

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

References

- [1] http://primekinetics.org
- [2] R. Feeley, P. Seiler, A. Packard, and M. Frenklach, "Consistency of a Reaction Dataset," *J. Phys. Chem.*, vol. 108, pp. 9573-9583, 2004.
- [3] M. Frenklach, A. Packard, P. Seiler, and R. Feeley, "Collaborative Data Processing in Developing Predictive Models of Complex Reaction Systems," *Int. J. Chem. Kinet.*, vol. 36, pp. 57-66, 2004.
- [4] T. Russi, A. Packard, R. Feeley, and M. Frenklach, "Sensitivity Analysis of Uncertainty in Model Prediction," *J. Phys. Chem. A*, vol. 112, no. 12, pp. 2579-2588, 2008.
- [5] T. Russi, A. Packard, and M. Frenklach, "Uncertainty quantification: Making predictions of complex reaction systems reliable," *Chem. Phys. Lett.*, vol. 499, no. 1-3, pp. 1-8, 2010.
- [6] M. Frenklach, A. Packard, and R. Feeley, *Modeling of Chemical Reactions*, R. W. Carr, Ed. Elsevier, pp. 243-291, 2007.
- [7] P. Seiler, M. Frenklach, A. Packard, and R. Feeley, "Numerical approaches for collaborative data processing," *Optim. Eng.*, vol. 7, pp. 459-478, 2006.
- [8] M. Frenklach, "Transforming Data into Knowledge—Process Informatics for Combustion Chemistry," in *Proceedings* of Combustion Symposium, Topical review, 2007, vol. 31, pp. 125-140.
- [9] N. A. Slavinskaya, U. Riedel, S. B. Dworkin, Q. Zhang, and M. J. Thomson, "Detailed Numerical Modeling of PAH Formation and Growth in Non-Premixed Ethylene and Ethane Flames," *Combust. Flame*, vol. 159, pp. 979-995, 2012.
- [10] C. Olm, I. Gy, I. G. Zsély, T. Varga, H. J. Curran, and T. Turányi, "Comparison of the performance of several recent syngas combustion mechanisms," *Combust. Flame*, vol. 162, pp. 1793–1812, 2015.
- [11] R. Atkinson, D. L. Baulch, et al., "Evaluated kinetic and photochemical data for atmospheric chemistry: Volume I gas phase reactions of Ox, HOx, NOx and SOx species," *Atmos. Chem. Phys.*, vol. 4, pp. 1461-1738, 2004.
- [12] H. Wang, X. You, A. V. Joshi, S. G. Davis, A. Laskin, F. Egolfopoulos, and C. K. Law. (May 2007). USC Mech Version II. High-Temperature Combustion Reaction Model of H2/CO/C1-C4 Compounds [Online]. Available: http://ignis.usc.edu/USC_Mech_II.htm.
- [13] J. A. Millerand and C. F. Melius, "Kinetic and thermodynamic issues in the formation of aromatic compounds in flames of aliphatic fuels," *Combust. Flame*, vol. 91, no. 1, pp. 21-39, 1992.
- [14] T. Kathrotia, M. Fikri, M. Bozkurt, M. Hartmann, U. Riedel, and C. Schulz, "Study of the H + O + M reaction forming OH*: Kinetics of OH* chemiluminescence in hydrogen combustion systems," *Combust. Flame*, vol. 157, no. 7, pp. 1261–1273, 2010.
- [15] E. L. Petersen and R. K. Hanson, "Nonideal effects behind reflected shock waves in a high-pressure shock tube," *Shock Waves*, vol. 10, pp. 405–420, 2001.
- [16] D. F. Davidson and R. K. Hanson, "Interpreting Shock Tube Ignition Data," *WSSCI Fall Meeting*, October 20-21, 2003, University of California at Los Angeles, Los Angeles, CA.
- [17] E. L. Petersen, et al., "A facility for gas- and condensed-phase measurements behind shock waves," *Meas. Sci. Technol.*, vol. 16, pp. 1716–1729, 2005.
- [18] E. L. Petersen and R. K. Hanson, "Measurement of reflected-shock bifurcation over a wide range of gas composition and pressure," *Shock Waves*, vol. 15, pp. 333–340, 2006.
- [19] F. L. Dryer and M. Chaos, "Ignition of Syngas/air and Hydrogen/air Mixtures at Low Temperatures and High Pressures: Experimental Data Interpretation and Kinetic Modeling Implications," *Combust. Flame*, vol. 152, pp. 293–299, 2008.
- [20] D. F. Davidson and R. K. Hanson, "Recent advances in shock tube/laser diagnostic methods for improved chemical kinetics measurements," *Shock Waves*, vol. 19, pp. 271–283, 2009.
- [21] M. Ihme, "On the role of turbulence and compositional fluctuations in rapid compression machines: Autoignition of syngas mixtures," *Combust. Flame*, vol. 159, pp. 1592–1604, 2012.
- [22] J. Urzay, et al., "Uncertainty-quantification analysis of the effects of residual impurities on hydrogen–oxygen ignition in shock tubes," *Combust. Flame*, vol. 161, pp. 1–15, 2014.
- [23] A.B. Mansfield and M. S. Wooldridge, "High-pressure Low-temperature Ignition Behavior of Syngas Mixtures," *Combust. Flame*, vol. 161, pp. 2242-2251, 2014.

- [24] K. P. Grogan and M. Ihme, "Weak and strong ignition of hydrogen/oxygen mixtures in shock-tube systems," *Proceedings of Combustion Symposium*, vol. 35, pp. 2181–2189, 2015.
- [25] D. M. Kalitan, J. D. Mertens, M. W. Crofton, and E. L. Petersen, "Ignition and oxidation of lean CO/H2 fuel blends in air," J. Propuls. Power, vol. 23, pp. 1291–1303, 2007.
- [26] E. L. Petersen, D. M. Kalitan, et al., "New syngas/air ignition data at lower temperature and elevated pressure and comparison to current kinetics models," *Combust. Flame*. vol. 149, pp. 244-247, 2007.
- [27] J. D. Mertens, D. M. Kalitan, and E. L. Petersen, "Determination of the Rate of $H + O2 + M \rightarrow HO2 + M$ (M = N2, Ar, H2O) from Ignition of Syngas at Practical Conditions," in *Proceedings of Combustion Symposium*, vol. 32, 2009, pp. 295-303.
- [28] J. Herzler and C. Naumann, "Shock tube study of the ignition of lean. CO/H2 fuel blends at intermediate temperatures and high pressure," *Combust. Sci. and Tech.*, vol. 180, pp. 2015-2028, 2008.
- [29] K. Kéromnès, K. Wayne, A. Metcalfe, N. Heufer, A. K. Das Donohoe, C-J. Sung, J. Herzler, C. Naumann, P. Griebel, O. Mathieu, M. J. Krejci, E. L Petersen, W. J. Pitz, and H. J. Curran, *Combust. Flame*, vol. 160, pp. 995–1011, 2013.
- [30] X. Q. You, T. Russi, A. Packard, and M. Frenklach, "Optimization of combustion kinetics models on a feasible set," in *Proceedings of Combustion Symposium*, vol. 33, 2011, pp. 509-516.