

Towards a high-accuracy kinetic database informed by theoretical and experimental data: $\text{CH}_3 + \text{HO}_2$ as a case study

Carly E. LaGrotta^a, Mark C. Barbet^a, Lei Lei^a, Michael P. Burke^{b,*}

^a Department of Mechanical Engineering, Columbia University, New York, NY 10027, USA

^b Department of Mechanical Engineering, Department of Chemical Engineering, Data Science Institute, Columbia University, New York, NY 10027, USA

Received 7 November 2019; accepted 28 June 2020

Available online 20 September 2020

Abstract

The reliability of predictive simulations for advanced combustion engines depends on the availability of accurate data and models for thermochemistry, chemical kinetics, and transport. In that regard, accurate data are critically important for both their direct use in predictive simulations and for benchmarking improved theoretical methodologies that can similarly produce accurate data for predictive simulations. The use of informatics-based strategies for the determination of accurate thermochemical data with well-defined uncertainties, e.g. the Active Thermochemical Tables (ATcT), has revolutionized the field of thermochemistry—providing thermochemical data of unprecedented accuracy for predictive combustion simulations and has served as a key enabler of *ab initio* electronic structure methodologies of equally impressive accuracy. Clearly, pursuit of informatics-based analogs in chemical kinetics would be similarly worthwhile. Here, we present results and analyses for the kinetics of $\text{CH}_3 + \text{HO}_2$ reactions that demonstrate some key elements of any approach to developing analogs for kinetics, including: the importance of raw data for quantifying the information content of experimental data, the utility of theoretical kinetics calculations for constraining experimental interpretations and providing an independent data source, and the subtleties of target data selection for avoiding unphysical parameter adjustments to match data affected by structural uncertainties. Notably, we find the optimization performed here using the MultiScale Informatics (MSI) approach applied to carefully selected (mostly raw) experimental data yields an opposite temperature dependence for the channel-specific $\text{CH}_3 + \text{HO}_2$ rate constants as compared to a previous rate-parameter optimization. While both optimization studies use the same theoretical calculations to constrain model parameters, only the present optimization, which incorporates theory directly into the model structure, yields results that are consistent with theoretical calculations.

© 2020 The Combustion Institute. Published by Elsevier Inc. All rights reserved.

Keywords: Optimization; Uncertainty quantification; Multiscale informatics; Rate constants

* Corresponding author.

E-mail address: mpburke@columbia.edu (M.P. Burke).

1. Introduction

Predictive simulation offers significant promise as a design tool to accelerate energy technology development. Truly predictive modeling, however, may require unprecedented accuracy in descriptions of the chemical kinetics. For example, engine simulations show that uncertainties in any of several rate constants yield uncertainties in ignition crank angle of $1\text{--}2^\circ$, which is enough to cause misfire [1]. In fact, even knowing all rate constants as accurately as the best characterized rate constant (10–15%) will not yield truly predictive simulation of flame propagation [2] or extinction [3]. Altogether, truly predictive modeling poses significant challenges for the current paradigm in the chemical kinetics field, where uncertainties for even the most studied reactions are still non-negligible and uncertainties for other reactions are much larger (and sometimes go unidentified) even after extensive investigation [4].

A nearly universal theme in experimental rate constant determinations is that their uncertainties generally exceed that dictated by measurement noise alone. Since completely isolating a single reaction (especially at combustion temperatures) is nearly impossible, uncertainties in rate constants for other reactions are often among the largest sources of uncertainty. This interdependence results in a complex network of implicit connections in rate constant determinations – with two significant implications: (1) the systematic uncertainties associated with ignoring these connections lead to inconsistencies among determinations that often exceed experimental precision [4] and (2) there is more information in existing data than has been possible to extract previously using traditional methods.

In this regard, the field of thermochemistry, which has many parallels to chemical kinetics, offers a particularly inspirational precedent. Informatics-based strategies, e.g. the Active Thermochemical Tables (ATcT) [5], have achieved thermochemical data of unprecedented accuracy by quantitatively accounting for the implicit connections in networks of interdependent data.

Achieving analogs for kinetics presents additional complexities, which dictate strict requirements on the elements of any universally applicable approach. For example, any potential approach would likely require not only interpretation of raw experimental data (to unravel the network of implicit connections among rate constant determinations) but also a theoretical kinetics backbone (to enable reliable extrapolation of rate constants whose temperature/pressure/bath gas composition ($T/P/M$) dependence are ill-constrained by data at limited conditions or to handle reactions defying typical rate constant descriptions [6–8]). The vast majority of kinetics informatics studies either infer molecular parameters (within theoretical

calculations) based on rate constant determinations [9–11] or infer rate constant parameters (within semi-empirical rate formulas) based on rate constant determinations and/or multi-reaction observables [3,12–15]. The MultiScale Informatics (MSI) approach [16–18] is unique in that it incorporates both aspects – combining theoretical and experimental data in a manner that both unravels data from multi-reaction systems and, through incorporation of theory, copes with sparse experimental datasets and/or unconventional kinetic behavior. Therefore, we have begun pursuit of a high-accuracy active database for chemical kinetics, leveraging our MSI approach and borrowing general ideas from the ATcT approach (while recognizing that the specifics of the implementation for kinetics will necessarily differ from those for thermochemistry) [5].

This paper outlines some initial steps and demonstrates the importance of various elements of the approach to deriving high-accuracy kinetic parameters. Our approach is to combine theoretical and carefully selected experimental data within MSI. With regard to experimental data selection, we use primarily raw data (to capture the experimental information content most accurately and effectively) from experiments involving small subsystems of at most a few reactions (to avoid introducing too many “weak links” [5]), homogeneous environments (to minimize uncertainties due to transport), and dilute conditions (to minimize structural uncertainties [3,17] in treating kinetics of multi-component, reactive mixtures [6,7,19–25]); we also choose specific subsystems with overlapping reactions in order to grow a network of strongly linked, redundant data (to leverage information from multiple sources and to test the validity of values and their uncertainties – key elements of ATcT [5]). Key elements of the approach are demonstrated here for the $\text{CH}_3 + \text{HO}_2$ reactions, whose temperature dependence is not well established from experimental data alone and whose rate constants contribute substantial modeling uncertainties [26], as a case study – leveraging our previous analysis of pure H_2O_2 [16] for analysis of $\text{H}_2\text{O}_2/\text{CH}_4$ experiments.

2. Approach

The objective of the MSI approach is to identify optimized values and quantified uncertainties for a set of molecular parameters (within theoretical kinetics calculations), rate parameters, and physical model parameters (within simulations of experimental observables) as informed by data from various sources and scales.

The first step of MSI is to construct an active multiscale model and identify a set of targets, y' . This active model consists of a set of active parameters, x , and a structure of physics-based

Table 1

List of active model parameters considered in the optimization.^{a,b}

Reaction		Kinetic Parameters
R1	$\text{H}_2\text{O}_2(+\text{M})=\text{OH}+\text{OH}(+\text{M})$	$A'_{(1)}, n_{(1)}, E_{(1)}$
R2	$\text{H}_2\text{O}_2+\text{OH}=\text{HO}_2+\text{H}_2\text{O}$	$E_{(2)}^\ddagger, v'_{\text{all}(2)}, v'_{\text{tr}(2)}, v'_{\text{ss}(2)}, v'_{\text{imag}(2)}, E_{\text{w}(2)}, \eta'_{\text{H}_2\text{O}_2}, \eta'_{\text{TS}(2)}$
R3	$\text{HO}_2+\text{HO}_2=\text{H}_2\text{O}_2+\text{O}_2$	$E_{(3)}^\ddagger, v'_{\text{all}(3)}, v'_{\text{tr}(3)}, v'_{\text{ss}(3)}, v'_{\text{imag}(3)}, E_{\text{w}(3)}, \eta'_{\text{TS}(3)}$
R4	$\text{HO}_2+\text{OH}=\text{H}_2\text{O}+\text{O}_2$	$E_{(4)}^\ddagger, v'_{\text{all}(4)}, v'_{\text{tr}(4)}, v'_{\text{ss}(4)}, v'_{\text{imag}(4)}, E_{\text{w}(4)}, \eta'_{\text{TS}(4)}$ $E_{(4e)}^\ddagger, v'_{\text{TS}(4e)}, v'_{\text{tr}(4e)}, v'_{\text{ss}(4e)}, v'_{\text{TS}(4e)}, f'_{\text{VRC-TST},c(4)}$
R5	$\text{OH}+\text{OH}=\text{O}+\text{H}_2\text{O}$	$E_{(5)}^\ddagger, v'_{\text{all}(5)}, v'_{\text{tr}(5)}, v'_{\text{ss}(5)}, v'_{\text{imag}(5)}, E_{\text{w}(5)}$ $E_{(5e)}^\ddagger, v'_{\text{TS}(5e)}, v'_{\text{tr}(5e)}, v'_{\text{ss}(5e)}$
R6	$\text{CH}_3+\text{HO}_2=\text{CH}_4+\text{O}_2$	$E_{(6)}^\ddagger, v'_{\text{low}(6)}, (Q_{\text{anh},c,\text{CH}_3})$
R7	$\text{CH}_3+\text{HO}_2=\text{CH}_3\text{O}+\text{OH}$	$f'_{\text{VRC-TST}(7)}, (Q_{\text{anh},c,\text{CH}_3})$
R8	$\text{CH}_4+\text{OH}=\text{CH}_3+\text{H}_2\text{O}$	$A'_{(8)}, n_{(8)}, E_{(8)}$
Macroscopic Observables		Physical Model Parameters
E1-E4	Shock-heated $\text{H}_2\text{O}_2/\text{H}_2\text{O}/\text{O}_2/\text{Ar}$	$T_e', P_e', M'_{\text{H}_2\text{O}_2,ae}, M'_{\text{H}_2\text{O},ae}, M'_{\text{O}_2,ae}, \hat{\sigma}'_{1,\text{H}_2\text{O}_2}, \hat{\sigma}'_{1,\text{HO}_2}$ $e = 1 \dots 4$
E5	Shock-heated $\text{H}_2\text{O}/\text{O}_2/\text{H}/\text{Ar}$	$T_e', P_e', M'_{\text{H},ae}, M'_{\text{H}_2\text{O},ae}, M'_{\text{O}_2,ae}$ $e = 5$
E6-E10	Shock-heated $\text{H}_2\text{O}_2/\text{Ar}$	$T_e', P_e', M'_{\text{H}_2\text{O}_2,ae}, \sigma'_{1,\text{H}_2\text{O}_2}, \sigma'_{2,\text{H}_2\text{O}_2}, \sigma'_{1,\text{HO}_2}, \sigma'_{2,\text{HO}_2}$ $e = 6 \dots 10$
E11	Shock-heated $\text{CH}_4/\text{H}_2\text{O}_2/\text{H}_2\text{O}/\text{O}_2$	$T_e', P_e', M'_{\text{CH}_4,ae}, M'_{\text{H}_2\text{O}_2,ae}, M'_{\text{H}_2\text{O},ae}, M'_{\text{O}_2,ae}, \hat{\sigma}'_{1,\text{H}_2\text{O}_2}, \hat{\sigma}'_{1,\text{HO}_2}$ $e = 11$

^a Note that ' indicates $\ln()$ of that quantity.^b Note that parameters in () are represented by only a single value but influence multiple reactions.

models that make a prediction of the i^{th} target, $f_i(\mathbf{x})$, for given $\mathbf{x}$. The active parameters can include active molecular properties, rate constant parameters, physical model parameters, or any combination of these (e.g. Table 1). The data to be used as targets can include *ab initio* calculations of molecular properties, experimental rate constant determinations, and macroscopic observables from multi-reaction systems (e.g. Table 2). Theoretical kinetics models (e.g. TST-ME [27]) relate active molecular properties to rate constants. Kinetic models, consisting of these rate constants, along with any active rate parameters for other reactions, are then combined with physical models (e.g. adiabatic isobaric reactors) and active physical model parameters for each experiment to predict experimental observables (e.g. species concentrations).

The second step of MSI is to impose the constraints from data onto the active model parameters via “inverse” uncertainty quantification (UQ) [3]. Similar to [16–18], inverse UQ here employs an iterative, uncertainty-weighted, least-squares error minimization.

Each iteration involves construction of a locally linear surrogate model of $f(\mathbf{x})$ in the neighborhood of $\tilde{\mathbf{x}}$

$$f(\mathbf{x}) \approx f(\tilde{\mathbf{x}}) + \mathbf{S}(\mathbf{x} - \tilde{\mathbf{x}}) \quad (1)$$

where $S_{ij} = (\frac{\partial f_i}{\partial x_j})_{\mathbf{x}=\tilde{\mathbf{x}}}$; followed by minimization of the uncertainty-weighted least-squares error

$$E(\mathbf{x}) = \sum_i \left(\frac{y_i^t - f_i(\mathbf{x})}{z_i} \right)^2 \quad (2)$$

to find the optimized values, $\mathbf{x}^*$, where y_i^t is the i^{th} target value, and z_i is the weighting factor for the i^{th} target, where $z_i = \sigma_i/w_i$ is equal to the uncertainty of the i^{th} target, σ_i , divided by the data set weighting factor, w_i (taken to be $w_i = 1/n^{0.5}$ where n is the number of data points in a data set used from a particular study); followed by setting $\tilde{\mathbf{x}}$ to $\mathbf{x}^*$ and repeating until converged.

Once converged, this nonlinear optimization yields an optimized set of active parameters, $\mathbf{x}^*$, and a covariance matrix, Σ , representing the uncertainties in active parameters and correlations among them as informed by all data (i.e. the joint probability distribution function). Prediction uncertainties (in Figs. 2 and 3) were propagated using the covariance matrix. (For reference, many other inverse UQ studies with similar numbers of active parameters, e.g. [3,12,13], use similar, if not identical, mathematical frameworks for quantifying uncertainty – treating prior and posterior distributions as multivariate normal distributions and using a local first-order surrogate model to construct a covariance matrix.)

Data from *ab initio* calculations of molecular properties (Target Class I, cf. Table 2), nominal values for any active rate parameters, and nominal values for active physical model parameters (Target Class IV) impose the prior distributions on model parameters and serve as natural regularization terms (such that the inverse problem is sufficiently constrained). Data for rate constant determinations (Target Class II) and macroscopic observables (Target Class III) serve as additional

Table 2
List of targets used in the optimization.^{a,b}

I. <i>Ab initio</i> calculations		Source
$E_2^\ddagger$ (3 kcal/mol), $E_3^\ddagger$ (2.3 kcal/mol), $E_{4g}^\ddagger$ (2 kcal/mol), $E_{4e}^\ddagger$ (2 kcal/mol), $E_{3g}^\ddagger$ (1.4 kcal/mol), $E_{5e}^\ddagger$ (1.4 kcal/mol), ν'_{alt} (0.03), ν'_{TS} (0.03), ν'_{ir} (0.1), ν'_{ss} (0.18), ν'_{imag} (0.18), E_w (1 kcal/mol), $\eta_{H_2O_2}$ (0.1), $\eta'_{TS(2)}$ (0.26), $\eta'_{TS(3)}$ (0.41), $\eta'_{TS(4g)}$ (0.41), $\eta'_{TS(4e)}$ (0.41), $f^{VRC-TST,c(4)}$ (0.7), $E_6^\ddagger$ (2.3 kcal/mol), $\nu'_{low(6)}$ (0.1), Q_{anh,c,CH_3} (1), $f^{VRC-TST(7)}$ (0.262)		(Harding and Klippenstein Unpublished) [33,34]
II. Rate constant measurements		
k'_2, k'_4, k'_5, k'_8 (see text)		[16,37,38]
III. Macroscopic exp.	IV. Exp. conditions	
$M'_{OH,e}(t)$ (0.05), $M'_{H_2O,e}(t)$ (0.05), $abs'_e(t)$ (0.1) $M'_{OH,e}(t)$ (0.05) $abs'_e(t)$ (0.1)	T'_e (0.01), P'_e (0.02), $M'_{H_2O_2,ae}$ (0.05), $M'_{H_2O,ae}$ (0.05), $M'_{O_2,ae}$ (0.05), $M'_{CH_4,ae}$ (0.05), $\hat{\sigma}'_{1,H_2O_2}$ (0.7), $\hat{\sigma}'_{1,HO_2}$ (0.7) T'_e (0.01), P'_e (0.02), $M'_{H_2O,ae}$ (0.1), $M'_{O_2,ae}$ (0.01), $M'_{H,ae}$ (2.3) T'_e (0.02), P'_e (0.04), $M'_{H_2O_2,ae}$ (0.1), σ'_{1,H_2O_2} (0.7), σ'_{2,H_2O_2} (0.3), σ'_{1,HO_2} (0.7), σ'_{2,HO_2} (0.3)	$e = 1...4, 11$ [39–42] $e = 5$ [43] $e = 6...10$ [44]

^a Note that ' indicates ln() of the quantity.
^b Uncertainties listed in () are intended to reflect two standard deviations.

targets for inverse UQ to impose the posterior distributions.

Uncertainties in this work are intended to reflect two standard deviations, though the lack of specification in most studies make such a designation tentative. Ultimately, once sufficient statistics are obtained through application of this approach to many systems, we expect that the results from our analysis may permit evaluation of the suitability of the uncertainty assignments for various types of data.

3. Implementation

The MSI approach was implemented here for the H₂O₂/CH₄ system, where kinetic parameters for R1-R8 and physical model parameters were assigned to represent uncertainties in the kinetic and physical models (Table 1). All other “secondary” reactions were taken from the Foundational Fuel Chemistry Model (FFCM) [12]. (Identical results were found using other models [26] or including active rate parameters for secondary reactions with uncertainties from [12].) Rate constants and brute-force sensitivity coefficients were calculated via TST-ME [27] in VARIFLEX [28] and PAPR-MESS [29]. Macroscopic observables and sensitivity coefficients were calculated via homogeneous, isochoric or isobaric models in CANTERA [30]. For R1 and R8, whose rate constants over the relevant temperature ranges were sufficiently constrained by macroscopic observables and rate constant targets, Arrhenius parameters (as opposed

to molecular parameters within theoretical kinetics models) were treated as active parameters. For R2-R7, molecular parameters within theoretical kinetics models were treated as active parameters. For R2-R5, the molecular parameters were taken to be the same as [16]. (While not incorporated in the results shown here, alternative analysis using the HO₂ + HO₂ product channels and rate constants from [31] yielded < 10% differences in optimized rate constants for all other reactions here, but we are exploring this reaction further elsewhere.) The active molecular parameters included transition state barrier heights ($E^\ddagger$); scaling factors for all harmonic frequencies of reactants and products (ν_{all}); scaling factors for all (transition-state) harmonic frequencies (ν_{TS}), transitional mode frequencies (ν_{tr}), symmetric stretching mode frequencies (ν_{ss}), and imaginary frequencies (ν_{imag}) in tight transition states; VRC-TST correction factors for loose transition states ($f^{VRC-TST,c}$); the smaller of the two well depths for pre- and post-transition state complexes (E_w); and scaling factors for hindered rotor potentials (η). For R6-R7, a more limited set of parameters was necessary to represent the uncertainties in the theoretical kinetics calculations. For R6, which proceeds through a barrierless entrance transition state to form a van der Waals complex that dissociates to products via an exit transition state with a barrier submerged by only ~ 1.9 kcal/mol relative to the reactants, the rate constant over the temperature range of interest is essentially controlled by only the exit transition state, to which active $E_6^\ddagger$ and $\nu_{low(6)}$ (a scaling factor for the lowest four harmonic

frequencies in the transition state) are ascribed. (Variation of $\eta_{TS(6)}$ and $\nu_{imag(6)}$ within their uncertainty limits yielded negligible differences in the rate constants over the relevant temperature range.) For R7, which proceeds via a barrierless entrance transition state to form a CH_3OOH complex that dissociates to products via another barrierless exit transition state that lies ~ 25 kcal/mol below the reactants, the rate constant is essentially controlled by the barrierless entrance transition state, to which an active $f_{V_{RC-TST,e}}$ is ascribed. (Calculations using roaming estimates from [32] for $J = 0$ show negligible impact of roaming on the present results, so roaming was not included here.)

Previous theoretical calculations for R6 and R7 neglect anharmonicity of the CH_3 umbrella mode [33,34]. To date, estimates using different assumptions indicate anharmonic corrections ranging from nearly zero to 30% [35,36] (Jasper, personal communication) – precluding definitive conclusions for its role. We retain the nominal values from [33,34] but assign another active parameter for R6 and R7 corresponding to a scaling factor for the anharmonic correction from [35], Q_{anh,c,CH_3} – effectively serving as a parameterization of the structural uncertainty related to anharmonicity in the CH_3 umbrella mode.

It is perhaps worth noting that the exact parametric and structural uncertainties to be considered naturally depend on the specifics of the system of interest. For example, when the systems involve larger kinetic mechanisms, where the potential for missing reactions is higher, greater consideration of “missing” reactions may be required (e.g. by using RMG as in [17]). A broader discussion of various sources of parametric and structural uncertainty can be found elsewhere [18], but as discussed elsewhere [18] and in Section 4 below, consideration of all relevant parametric and structural uncertainties is important to obtaining physically realistic parameter adjustments.

Physical model parameters from experiments included as macroscopic observable targets were also among the active model parameters. These parameters included initial temperatures, T_e , initial pressures, P_e , initial species mole fractions, $M_{i,0,e}$, and coefficients for temperature-dependent cross sections of H_2O_2 and HO_2 at 215 nm ($\sigma_{1,\text{H}_2\text{O}_2}$, $\sigma_{2,\text{H}_2\text{O}_2}$, σ_{1,HO_2} , and σ_{2,HO_2}) and 227 nm ($\hat{\sigma}_{1,\text{H}_2\text{O}_2}$ and $\hat{\sigma}_{1,\text{HO}_2}$).

3.1. Target class (I): *ab initio* calculations

Molecular properties from *ab initio* calculations were used as targets for the molecular parameters for R2–R5 from (Harding and Klippenstein, Unpublished) [16], for R6 from [33], and R7 from [34] with uncertainties listed in Table 2 reflecting the level of electronic structure theory and the level of detail in treatment [16–18,27]. In particular, barrier height uncertainty estimates for the various method/basis set combinations were based on

quoted errors from a recent paper by Klippenstein and Cavallotti [27]. It is worth noting that the uncertainty values used here differ from those used in the original MSI study [16], which employed less refined uncertainty estimates of 3, 4, 2, and 3 kcal/mol for R2–R5, respectively. However, alternative MSI analysis conducted with the original barrier height uncertainty estimates [16] yielded very similar results (with less than 20% differences in all optimized rate constants) and identical conclusions.

3.2. Target class (II): Rate constant determinations

Rate constant determinations (from studies other than those treated as macroscopic observables) were included as targets for R2, R4, R5, and R8. Given that uncertainties in rate constant determinations are often higher than typically estimated and prone to systematic errors [4,16,18], conservative uncertainties of a factor of two (or the scatter in the data if larger) were applied.

3.3. Target class (III) and (IV): Macroscopic observables and experimental conditions

Raw data for the macroscopic observables from experiments indicated in Table 2 were included as targets using uncertainties as reported or estimated using typical values [39–43]. (While other studies examined R6 and R7 [33,45,46], further analysis, included in the Supplemental Material, indicated strong influences of several secondary reactions and, therefore, they were not included as targets.) Macroscopic observables used as targets included time-resolved H_2O mole fractions, OH mole fractions, and absorbance at 215 and 227 nm. Reported values for the experimental conditions were used as targets for physical model parameters with uncertainties as reported or using typical values [39–43].

4. Results and discussion

The set of optimized molecular parameters, rate parameters, and physical model parameters constrained by the targets of Table 2 yield values and predictions in reasonable consistency with the targets from *ab initio* calculations, rate constant determinations, macroscopic observables, and reported experimental conditions and other experiments not used as targets. All MSI values, rate constants, and predictions are provided in the Supplemental Material. The discussion below focuses on the macroscopic observables for shock-heated $\text{CH}_4/\text{H}_2\text{O}_2/\text{H}_2\text{O}/\text{O}_2/\text{Ar}$ mixtures, which emphasize R1–R8 in Table 1, and k_6 and k_7 in particular.

Uncertainty-weighted sensitivity analysis of the time profiles for H_2O , OH, and absorbance at 227

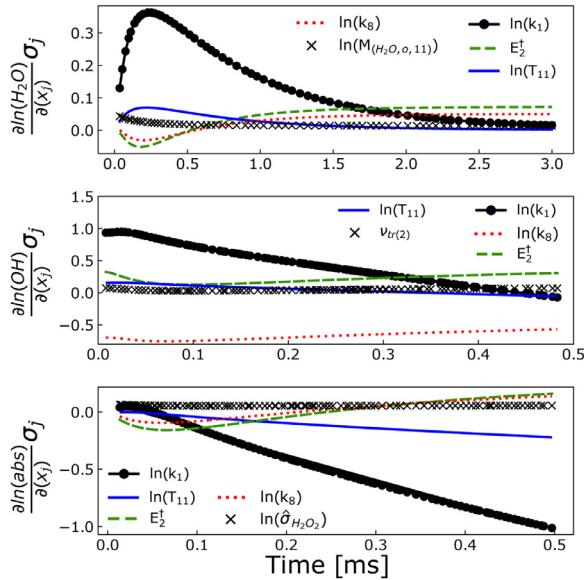


Fig. 1. Uncertainty-weighted sensitivity analysis for the OH, H₂O, and absorbance time profiles of Hong et al. [42] For the figure, k_1 and k_8 are shown with uncertainties of a factor of two and $\delta_{\text{H}_2\text{O}_2}$ and δ_{HO_2} are shown with uncertainties of 10%.

nm are shown in Fig. 1. Uncertainty-weighted sensitivity coefficients indicate the parameters whose uncertainties influence uncertainties in the model predictions and, likewise, which parameters can be constrained by inclusion of data for those observables. Inspection of the five most influential parameters shown in Fig. 1 reveals that predictions of all three observables are predominantly influenced by uncertainties in parameters related to k_1 , k_2 ($E_2^\ddagger$ and $v_{tr(2)}$), k_8 , and physical model parameters (T_{11} , $M_{\text{H}_2\text{O}, o, 11}$, and $\delta_{\text{H}_2\text{O}_2}$). These experiments were used to determine k_6 and k_7 individually [42] and exhibit some sensitivity to their active parameters. However, after including constraints from *ab initio* calculations [33,34], uncertainty-weighted sensitivity coefficients for parameters related to k_6 and k_7 are comparatively minimal. These results demonstrate the implicit interconnections among rate constant determinations for R6–R8 and those for R1–R5 and highlight the advantage of analyzing raw data to capture the information content and unravel the inherent interconnections among rate constant determinations.

Predicted time profiles for H₂O, OH, and absorbance at 227 nm are shown in Fig. 2 for the *a priori* model and MSI model, as compared to experimental data [42]. While the *a priori* model predicts higher OH and lower absorbance than observed experimentally, the MSI model predictions for all three profiles are consistent with experimental data within uncertainties. As one might anticipate based on the predictions being primarily influenced by k_1 and k_2 , which are already con-

strained by H₂O₂ experiments, inclusion of only H₂O₂ experiments as targets (without any CH₄-containing experiments as targets) yielded optimized predictions that are remarkably close to the optimized profiles including all data from Table 2. Inclusion of the CH₄/H₂O₂ experiments as targets did yield modest improvements in the OH peak and absorbance (of roughly 10%). Inclusion of CH₄/H₂O₂ experimental targets yields some minor adjustments in parameters for R6–R8, though predictions with these parameters at their nominal values (and R1–R5 parameters optimized to only H₂O₂ experimental targets) are already reasonably consistent with experimental data. The ability of the model to predict these time profiles so well without including the CH₄/H₂O₂ experiments as targets highlights the advantage of leveraging data from multiple sources. Data for pure H₂O₂ decomposition, theoretical calculations for R6–R7, and k_8 determinations are leveraged here in the interpretations of the raw experimental data to extract more accurate information for R6–R7.

Figure 3 shows comparisons of k_6 and k_7 from the *a priori* model (which are simply the theoretically calculated rate constants from [33] and [34] respectively), MSI model, and model of Olm et al. [13] along with experimental rate constant determinations from [42] (from the experiments shown in Fig. 2) and [33]. Both the values and uncertainties of rate constants in the *a priori* and MSI models are very similar – indicating that the MSI rate constants are predominantly dictated by theory.

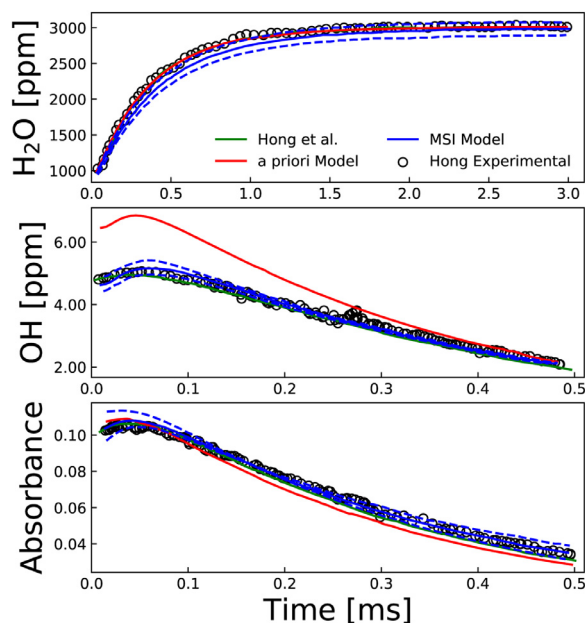


Fig. 2. H_2O , OH , and absorbance time profiles at 1103 K in shock-heated $\text{H}_2\text{O}/\text{H}_2\text{O}_2/\text{CH}_4/\text{Ar}$. Symbols represent experimental data from Hong et al. [42]; lines represent model predictions.

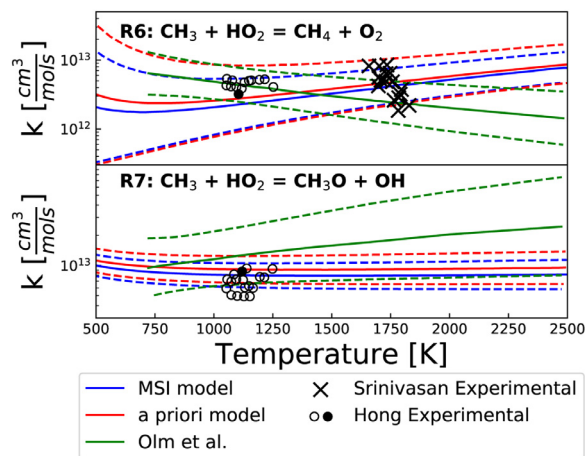


Fig. 3. Rate constants for R6 and R7. Original experimental interpretations [33,42] are shown by symbols (filled in for the conditions of Fig. 2). Dashed lines show prior and posterior uncertainties.

Interestingly, while both the present MSI optimization and the Arrhenius parameter optimization [13] use theoretical calculations [33,34] as targets, the Arrhenius parameter optimization yields a temperature dependence of k_6 and k_7 inconsistent with theoretical calculations [33,34]. These differences likely stem from the (1) selection of experimental data used as targets and (2) way that theoretical calculations are included as targets.

With regard to (1), many targets used in the Arrhenius parameter optimization [13] exhibit

(neglected) structural uncertainties comparable to parametric uncertainties – including chemically termolecular reactions [6], prompt dissociation [8], mixture rules [24], and missing collision efficiencies [47]. Use of targets influenced by structural uncertainties would then impose unphysical constraints on the model parameters. Indeed, inspection of the covariance matrix from [13] reveals that k_6 and k_7 parameters are strongly correlated to parameters for several other reactions whose optimized rate constants are near the edge of the prior

uncertainties ($\text{CH}_2\text{O} + \text{O}_2$, $\text{CH}_3\text{OH} + \text{OH}$), have very different temperature dependence than the prior model ($\text{CH}_3\text{OH} + \text{H}$), or are inconsistent with theoretical calculations ($\text{CH}_3\text{OH} + \text{HO}_2$).

With regard to (2), direct inclusion of theoretical calculations into the MSI model structure ensures that optimized rate constants are inherently consistent with theory. Altogether, the present use of dilute experiments minimizes structural uncertainties in multi-component, reactive mixtures [6,24,47] and direct incorporation of theory into the MSI framework minimizes the potential that kinetic parameters will reach unphysical values.

5. Conclusions

An MSI analysis focused on the $\text{CH}_3 + \text{HO}_2$ reaction is performed – leveraging our previous analysis of pure H_2O_2 [16] for analysis of $\text{H}_2\text{O}_2/\text{CH}_4$ experiments. Results and discussion highlight the key elements of any approach to developing a high-accuracy kinetic database, including: the importance of raw data for quantifying the information content of experimental data, the utility of theoretical kinetics calculations for constraining experimental interpretations and providing an independent data source, and the subtleties of target data selection for avoiding unphysical parameter adjustments to match data affected by structural uncertainties.

Our ongoing studies now focus on more highly studied systems to achieve sufficient data redundancy to test the validity of the optimized values and quantified uncertainties. However, the uncertainty-weighted sensitivity results here give an initial indication of the difficulty in achieving data redundancy in kinetics – such that a large web of interconnected systems may be required.

Declaration of Competing Interest

None.

Acknowledgments

The authors gratefully acknowledge the Department of Energy Gas Phase Chemical Physics program (DE-SC0019487) and National Science Foundation Combustion and Fire Systems program (CBET-1706252) and thank Drs. Ahren Jasper and Stephen Klippenstein for providing input files for the theory calculations [33,34].

Supplementary material

Supplementary material associated with this article can be found, in the online version, at [10.1016/j.proci.2020.06.324](https://doi.org/10.1016/j.proci.2020.06.324)

References

- [1] S. Som, W. Liu, D.D.Y. Zhou, G.M. Magnotti, R. Sivaramakrishnan, D.E. Longman, R.T. Skodje, M.J. Davis, *J. Phys. Chem. Lett.* 4 (12) (2013) 2021–2025.
- [2] M.P. Burke, F.L. Dryer, Y. Ju, *Proc. Combust. Inst.* 33 (1) (2011) 905–912.
- [3] H. Wang, D.A. Sheen, *Prog. Energy Combust. Sci.* 47 (2015) 1–31.
- [4] D.L. Baulch, C.T. Bowman, C.J. Cobos, R.A. Cox, Th. Just, J.A. Kerr, M.J. Pilling, D. Stocker, J. Troe, W. Tsang, et al., *J. Phys. Chem. Ref. Data* 34 (3) (2005) 757–1397.
- [5] B. Ruscic, R.E. Pinzon, M.L. Morton, G. Von Laszewski, S. Bittner, S.G. Nijssure, K.A. Amin, M. Minkoff, A.F. Wagner, *J. Phys. Chem. A* 108 (45) (2004) 9979–9997.
- [6] M.P. Burke, S.J. Klippenstein, *Nat. Chem.* 9 (11) (2017) 1078–1082.
- [7] L. Lei, M.P. Burke, *Proc. Combust. Inst.* (2020), doi:10.1016/j.proci.2020.06.385.
- [8] N.J. Labbe, R. Sivaramakrishnan, C.F. Goldsmith, Y. Georgievskii, J.A. Miller, S.J. Klippenstein, *J. Phys. Chem. Lett.* 7 (1) (2015) 85–89.
- [9] J. Prager, H.N. Najm, J. Zádor, *Proc. Combust. Inst.* 34 (1) (2013) 583–590.
- [10] R.J. Shannon, A.S. Tomlin, S.H. Robertson, M.A. Blitz, M.J. Pilling, P.W. Seakins, *J. Phys. Chem. A* 119 (28) (2015) 7430–7438.
- [11] M.A. Blitz, N.J.B. Green, R.J. Shannon, M.J. Pilling, P.W. Seakins, C.M. Western, S.H. Robertson, *J. Phys. Chem. A* 119 (28) (2015) 7668–7682.
- [12] G. Smith, Y. Tao, H. Wang, Foundational Fuel Chemistry Model Version 1.0 (FFCM-1) (2016).
- [13] C. Olm, T. Varga, É. Valkó, H.J. Curran, T. Turányi, *Combust. Flame* 186 (2017) 45–64.
- [14] D.R. Yeates, W. Li, P.R. Westmoreland, W. Speight, T. Russi, A. Packard, M. Frenklach, *Proc. Combust. Inst.* 35 (1) (2015) 597–605.
- [15] M. Khalil, H.N. Najm, *Combust. Theory Model.* 22 (4) (2018) 635–665.
- [16] M.P. Burke, S.J. Klippenstein, L.B. Harding, *Proc. Combust. Inst.* 34 (1) (2013) 547–555.
- [17] M.P. Burke, C.F. Goldsmith, S.J. Klippenstein, O. Welz, H. Huang, I.O. Antonov, J.D. Savee, D.L. Osborn, J. Zádor, C.A. Taatjes, L. Sheps, *J. Phys. Chem. A* 119 (28) (2015) 7095–7115.
- [18] M.P. Burke, *Int. J. Chem. Kinet.* 48 (4) (2016) 212–235.
- [19] M.P. Burke, R. Song, *Proc. Combust. Inst.* 36 (2017) 245–253.
- [20] M.C. Barbet, K. McCullough, M.P. Burke, *Proc. Combust. Inst.* 37, (2019) 347–354.
- [21] L. Lei, M.P. Burke, *Proc. Combust. Inst.* 37 (1) (2019) 355–362.
- [22] L. Lei, M.P. Burke, *J. Phys. Chem. A* 123 (3) (2019) 631–649.
- [23] L. Lei, M.P. Burke, *Combust. Flame* 213 (2020) 467–474.
- [24] L. Lei, M.P. Burke, *Proc. Combust. Inst.* (2020), doi:10.1016/j.proci.2020.06.187.
- [25] R.E. Cornell, M.C. Barbet, M.P. Burke, *Proc. Combust. Inst.* (2020) doi:10.1016/j.proci.2020.06.354.
- [26] W. Metcalfe, S. Burke, S. Ahmed, *Int. J. Chem. Kinet.* 45 (10) (2013) 638–675.

- [27] S.J. Klippenstein, C. Cavallotti, *Comput. Aided Chem. Eng.* 45 (2019) 115–167.
- [28] S.J. Klippenstein, A.F. Wagner, R.C. Dunbar, D.M. Wardlaw, S.H. Robertson, J.A. Miller, VARIFLEX, 2005.
- [29] Y. Georgievskii, J.A. Miller, M.P. Burke, S.J. Klippenstein, *J. Phys. Chem. A* 117 (46) (2013) 12146–12154.
- [30] D.G. Goodwin, H.K. Moffat, R.L. Speth, Cantera: An Object-Oriented Software Toolkit for Chemical Kinetics, Thermodynamics, and Transport Processes, 2018, Version 2.4.0.
- [31] S.J. Klippenstein, R. Sivaramakrishnan, U. Burke, K.P. Somers, H.J. Curran, L. Cai, H. Pitsch, M. Pelucchi, T. Faravelli, P. Glarborg, *Nat. Meet. Combust. Inst.* (2019) 1A13.
- [32] S.J. Klippenstein, Y. Georgievskii, L.B. Harding, *J. Phys. Chem. A* 115 (50) (2011) 14370–14381.
- [33] N.K. Srinivasan, J.V. Michael, L.B. Harding, S.J. Klippenstein, *Combust. Flame* 149 (1–2) (2007) 104–111.
- [34] A.W. Jasper, S.J. Klippenstein, L.B. Harding, *Proc. Combust. Inst.* 32 (1) (2009) 279–286.
- [35] D.M. Medvedev, L.B. Harding, S.K. Gray, *Mol. Phys.* 104 (1) (2006) 73–81.
- [36] N.J. Labbe, A.W. Jasper, J.A. Miller, S.J. Klippenstein, *10th U.S. Nat. Combust. Meet.* (2017) 1A01.
- [37] T. Gierczak, R.K. Talukdar, S.C. Herndon, G.L. Vaghjiani, A. Ravishankara, *J. Phys. Chem. A* 101 (17) (1997) 3125–3134.
- [38] J.R. Dunlop, F.P. Tully, *J. Phys. Chem.* 97 (43) (1993) 11148–11150.
- [39] Z. Hong, R.D. Cook, D.F. Davidson, R.K. Hanson, *J. Phys. Chem. A* 114 (18) (2010) 5718–5727.
- [40] Z. Hong, A. Farooq, E.A. Barbour, D.F. Davidson, R.K. Hanson, *J. Phys. Chem. A* 113 (46) (2009) 12919–12925.
- [41] Z. Hong, K.-Y. Lam, R. Sur, S. Wang, D.F. Davidson, R.K. Hanson, *Proc. Combust. Inst.* 34 (1) (2013) 565–571.
- [42] Z. Hong, D.F. Davidson, K.-Y. Lam, R.K. Hanson, *Combust. Flame* 159 (10) (2012) 3007–3013.
- [43] Z. Hong, S.S. Vasu, D.F. Davidson, R.K. Hanson, *J. Phys. Chem. A* 114 (17) (2010) 5520–5525.
- [44] Ch. Kappel, K. Luther, J. Troe, *Phys. Chem. Chem. Phys.* 4 (18) (2002) 4392–4398.
- [45] M. Sangwan, L.N. Krasnoperov, *J. Phys. Chem. A* 117 (14) (2013) 2916–2923.
- [46] J.J. Scire Jr., R.A. Yetter, F.L. Dryer, *Int. J. Chem. Kinet.* 33 (2) (2001) 75–100.
- [47] M.C. Barbet, M.P. Burke, *Proc. Combust. Inst.* (2020), doi:10.1016/j.proci.2020.06.178.