

Review

Combustion kinetic model uncertainty quantification, propagation and minimization

Hai Wang ^{a,*}, David A. Sheen ^b^a High Temperature Gas Dynamics Laboratory, Department of Mechanical Engineering, Stanford University, Stanford, CA 94305, USA^b Chemical Sciences Division, National Institute of Standards and Technology, Gaithersburg, MD 20899, USA

ARTICLE INFO

Article history:

Received 23 April 2014

Accepted 13 October 2014

Available online 29 November 2014

Keywords:

Combustion

Chemical kinetics

Kinetic modeling

Uncertainty quantification

UQ

Hydrocarbon fuels

ABSTRACT

The current interest in the combustion chemistry of hydrocarbon fuels, including the various alcohol and biodiesel compounds, motivates this review of the methods and application of kinetic uncertainty quantification (UQ). Our intent is to provide a self-contained review about the mathematical principles and methods of uncertainty quantification and their application in highly complex, multi-parameter combustion chemistry problems. We begin by outlining the reasons why the kinetic uncertainty must be considered and treated as a part of the combustion chemistry development in order to make progress. This is followed by a brief discussion about the sources and classification of kinetic uncertainties and the meanings and definitions of model verification and validation. We discuss the histories of UQ studies with an emphasis on how the combustion community has a long tradition of UQ consideration through standard sensitivity analysis. Such efforts have motivated the advancements of UQ methods specifically tailored to combustion chemistry. They also led to the recent growing interests in applying UQ methods as a part of our recommended long-term solution to the chemical kinetic problem of combustion. We then review and classify the various UQ methods and illustrate their applications for problems involving forward uncertainty quantification and propagation, and as an inverse problem leading to model uncertainty constraining. For the inverse problem, the focus of discussion is in the use of methods originating from Bayes' Theorem. We show that, for combustion chemistry problems, while UQ alone cannot produce precise, individual rate parameters, it can be instrumental in measuring the progress of our understanding of combustion chemistry and in utilizing fundamental combustion property data beyond a simple "agree–disagree" statement. When treated as a Bayesian inference problem, UQ also aids the development of predictive kinetic models in two ways: the use of fundamental combustion property data, global or local, to provide a better constrained kinetic model, and along with forward kinetic uncertainty propagation, to yield estimates for the confidence of a model to make predictions outside of the thermodynamic regimes where the model has been tested. We provide several examples to illustrate the utility of the UQ methods discussed and to demonstrate that, in the field of combustion chemistry, further progress will be better achieved through a combination of fundamental studies of reaction rates through well-defined and designed experiments and *ab initio* theoretical calculations, and of analyses of global experimental measurements. These studies together must be supplemented by UQ analyses, in such a way that a measurable progress can be made over time.

© 2014 Elsevier Ltd. All rights reserved.

Contents

1. Introduction	2
1.1. Historical perspective of kinetic model development: methods and practices	3
1.2. Sensitivity analysis and uncertainty quantification in combustion kinetics	5
1.3. Error versus uncertainty in computer modeling	7
1.4. Source of uncertainties	7

* Corresponding author.

1.5.	Model verification and validation	8
1.5.1.	Verification	8
1.5.2.	Validation	8
2.	Uncertainty quantification	8
2.1.	Estimation of rate constant uncertainties	8
2.2.	Representations of model response and uncertainty	9
2.2.1.	The Taylor series	9
2.2.2.	Polynomial chaos expansion	11
2.2.3.	High-dimensional model representation	12
2.3.	Intrusive methods	13
2.3.1.	Stochastic spectral projection	13
2.3.2.	Sensitivity-analysis-based method	13
2.4.	Non-intrusive methods	14
2.4.1.	Solution mapping	14
2.4.2.	Deterministic equivalent modeling method	15
2.4.3.	Non-intrusive spectral projection	15
2.4.4.	Random-sample high-dimensional model representation	15
2.4.5.	Uncertainty estimation through Monte Carlo integration	16
2.5.	Probability distribution of rate uncertainty versus model uncertainty	16
3.	Uncertainty minimization	16
3.1.	Limitations of forward uncertainty propagation	16
3.2.	The inverse problem	16
3.2.1.	Overview of Bayesian methods	18
3.2.2.	Bayesian analysis	18
3.2.3.	Bayesian analysis using Markov chain Monte Carlo	20
3.2.4.	Uncertainty minimization through polynomial chaos expansion	20
3.2.5.	Data collaboration	21
3.2.6.	Balance between mathematical and physical rigor	22
3.3.	Defining the experimental measurement uncertainty	22
3.3.1.	Experimental uncertainty analysis	22
3.3.2.	Regression error estimate	22
3.3.3.	Bayesian error inference	22
3.4.	Outlier detection and consistency analysis	22
4.	Selected problems of UQ applications in combustion chemistry	23
4.1.	Rate rules and their uncertainties	24
4.2.	Consistency of experimental dataset	24
4.3.	Assessment of systematic experimental uncertainty	24
4.4.	Hierarchical model development	25
4.5.	Model reduction	25
4.6.	Large eddy simulation	26
5.	Perspectives and concluding remarks	26
	Disclaimer	27
	Acknowledgments	27
	References	27

1. Introduction

“Remember that all models are wrong; the practical question is how wrong do they have to be to not be useful.” [1]

G. E. P. Box and N. R. Draper (1987)

In complex problems many aspects of the system are not exactly known and may never be known exactly. Coupled with the high nonlinearity and dimensionality, one must ask how it can be possible to use models of such systems to explore the intricate nature of a phenomenon and make useful predictions. The approach to developing detailed kinetic models (mechanisms) of fuel combustion involves compiling a set of elementary reactions whose rate parameters may be determined from individual rate measurements, reaction-rate theory, or a combination of both. For large hydrocarbon fuels, many of the reaction pathways and rates must be based on extrapolation of knowledge of smaller-species reactions. These methods have uncertainties. Whether the

collective uncertainties can be small enough to meet a certain chemical accuracy and to satisfy a particular combustion simulation continues to be an open question. Of course, model uncertainties may also be the result of incomplete physics and missing reaction pathways. It has been shown and will be reiterated in this review article, that even if the reaction model is complete, the underlying rate-coefficient uncertainty generally precludes the possibility of predicting relevant combustion properties of a fuel *a priori* [2].

Let us consider the rate coefficient of the quintessential reaction of combustion.



with rate constant k_1 . According to the NIST chemical kinetics database [3], k_1 has been examined in at least 77 independent experimental studies in the forward or reverse directions. The rate expressions have been reviewed and evaluated at least 31 times. A historical view of the uncertainty in k_1 , depicted in Fig. 1, is revealing both in terms of the past achievements and future

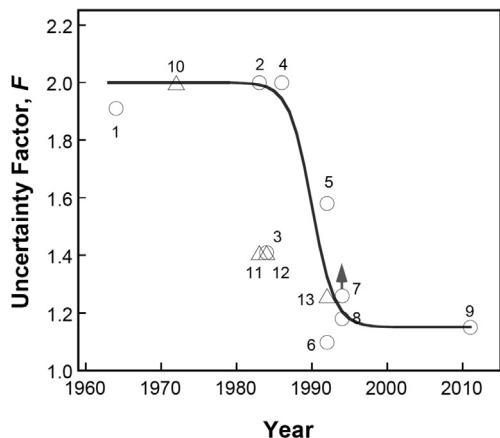


Fig. 1. Selected uncertainty factors quoted in experimental studies of the rate coefficient of $\text{H} + \text{O}_2 \rightarrow \text{O} + \text{OH}$. Circles: forward reaction; triangles: reverse reaction. 1) 300–1600 K [4]; 2) 300–2500 K [5]; 3) 300–2500 K [6]; 4) 300–1250 K [7]; 5) 300–2500 K [8]; 6) 960–530 K [9]; 7) 300–5000 K (the uncertainty factor increases to 3.2 at 5000 K). The plotted value is for 300 K) [10]. 8) 1340–3370 K [11]; 9) 1100–3370 K [12]; 10) 300–2000 K [13]; 11) 200–2500 K [5]; 12) 300–2500 K [6]; 13) 1000–2000 K [8].

challenges. The tremendous advances in laser diagnostics and shock tube techniques in the late 1980s and early 1990s [14] brought significant improvement in the accuracy of k_1 . Today, the best experiments give a two-standard deviation uncertainty in k_1 better than 15% over the temperature range of 1100 K–3370 K [12]. Among elementary reactions of combustion relevance, this precision is by far the highest achieved. It is truly an astonishing achievement considering that k_1 itself spans two orders of magnitude over that temperature range. Yet, such precision may still not be enough for combustion simulations. For example, if a particular combustion response has a logarithmic sensitivity coefficient (see, Section 1.2 for definition) of 0.2 with respect to k_1 —a value typical for the laminar flame speed—the uncertainty of k_1 causes the predicted combustion response to be uncertain by $\pm 3\%$ due to this uncertainty alone. The rate coefficients of other reactions are, in general, less certain than is k_1 , even though they generally do not impact the combustion prediction as much as k_1 .

As a further example, Fig. 2 illustrates a model problem of ethylene oxidation in a perfectly stirred reactor at a pressure of 30 bar and a constant-temperature of 1200 K, and in particular, the prediction uncertainties exhibited with a representative reaction kinetic model [16]. The figure presents a Monte Carlo sampling of the uncertainties of the rate coefficients of the reactions considered in the model and their impact on the concentrations of OH, H₂, H₂O, CO and CO₂ as a function of the mean residence time. Figures in the left panel show the results obtained using uncertainty factors currently known for each reaction, whose values range from 1.15 for k_1 , 1.2 to 2 for 8% of the reactions, 2 and 3 each for one third of the reactions, and 5 to 10 for the remaining 20% of the reactions. The right panel of Fig. 2 shows the corresponding uncertainty assuming all rate coefficients to be hypothetically accurate to within 15%. Two observations can be made from these plots. First, the current uncertainties are still too large for the model to be predictive. For example, in the hysteresis region, the combined uncertainty in the rate parameters produce a prediction uncertainty of more than an order of magnitude in residence time and two orders of magnitude in concentration. Second, even if all rate parameters could be determined to within 15% uncertainty and all reaction pathways were accounted for, the uncertainty of the model prediction is still not negligible. As shown in the figure, the extinction and ignition times can be still

uncertain by as much as a factor of 2. Of course, not every rate coefficient needs to have an uncertainty as small as 15%. The uncertainty in the simulation results usually comes from the uncertainty in the rate parameters of just a handful number of reactions. As will be discussed in further details in Section 3, both observations underscore the key obstacles to combustion chemistry model development and the need for a proper consideration of model uncertainty, both now and the future.

Uncertainty quantification (UQ) is the science for quantification and minimization of uncertainties with computer experiments and simulation [17,18]. It combines physical observations of the system, computational models, and expert opinions to make predictions about the system and, equally importantly, to determine the confidence of these predictions. As a mathematical science, UQ is rooted deeply in statistical mathematics. UQ methods, such as the Bayesian uncertainty approach [19], provide statistical alternatives to deterministic numerical integration.

One typical example in which UQ is essential is weather forecasting. Uncertainty is a fundamental property of weather and seasonal climate predictions. Today, no forecast is complete without a description of its uncertainty [20]. In other words, for such a complex phenomenon, we are forced to make only approximate yet still useful predictions because of the uncertainties intrinsic to the model and its numerical solutions. Although it is not the authors' view that the problem of weather forecasting is identical to combustion in nature, they do share similarities: weather-related prediction uncertainties stem from limitation in our knowledge about the physical processes, while others come from uncertainties in the parameters of the model and the experiments with which we "calibrate" the model, a situation not entirely different from combustion chemistry.

Indeed, one of the key challenges of combustion chemistry lies in the handling and reduction of uncertainties in the reaction pathways and their rate parameters and of an efficient approach to assess the impact of these uncertainties in our ability to predict a combustion phenomenon. This review paper highlights recent advances about UQ applications in combustion chemistry analysis. It also outlines the challenges in future progresses in combustion chemistry in the context of uncertainty quantification.

1.1. Historical perspective of kinetic model development: methods and practices

Combustion chemistry through numerical detailed kinetic modeling started from the work of Dixon–Lewis some fifty years ago [21], though studies that used the principles of elementary reaction mechanisms and kinetics to understand the chemistry of pyrolysis, oxidation and combustion started much earlier (e.g., Refs. [22–25]). Since then, many kinetic models for the combustion of hydrocarbon fuels have been proposed (see, e.g., Refs. [26–90] for work published prior to 2000). The triumph and progress over that period of time have been reviewed from different perspectives (see, e.g., Refs. [91–93]). The success of the earlier work has encouraged a wide range of studies in more recent years into the combustion chemistry of large-molecular weight, liquid hydrocarbon fuels [94–98]. Meanwhile, the community also realized the need to revisit the foundational part of the combustion chemistry from time to time as new, more accurate understanding emerges. Examples include the recent efforts in the H₂/CO reaction mechanism (e.g., Refs. [12,99–111]). Such effort has also manifested in persistent efforts to update rate coefficients through comprehensive evaluations [7,8,10,112–118].

A feature common to past efforts is the lack of quantitative measure of the uncertainties underlying a kinetic model. A feature parallel to this problem is the "many-model" problem, or model

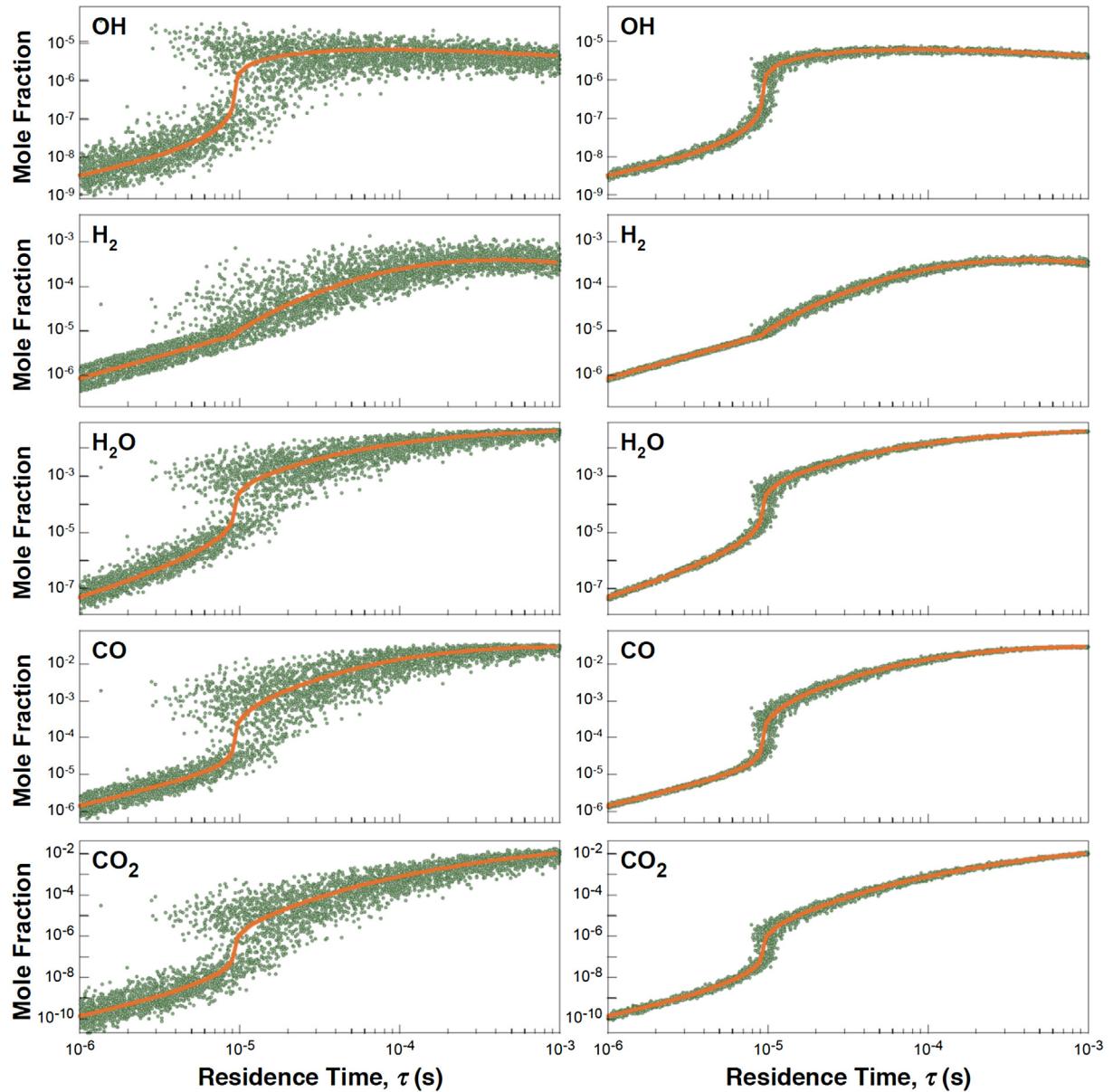


Fig. 2. Monte Carlo sampling (symbols) of the prediction uncertainties for selected major species in ethylene oxidation in air in a constant-temperature perfectly stirred reactor (equivalence ratio $\phi = 0.5$, pressure $p = 30$ bar, and temperature $T = 1200$ K). The lines represent nominal predictions. The simulation used USC Mech II [16] as an example. Left panel: rate parameter uncertainties of an earlier evaluation [15]; right panel: all rate coefficients were assigned a hypothetical, uniform uncertainty factor of 1.15 (or 15% uncertainty).

proliferation [119]. In fact, these “many” models can be viewed as statistical sampling of the underlying parameter uncertainty. Consider some of the reaction models of H_2 oxidation published prior to a comprehensive set of laminar burning rates of hydrogen at high pressures being made available [106]. The reaction models considered here [100,101,103–105,120] share the same set of reactions and were all “calibrated” in one way or another against the laminar flame speed of hydrogen-air mixtures at atmospheric pressure. They differ only in the choices of rate coefficients. When compared to the laminar mass burning rate of hydrogen in an oxygen–helium mixture, however, Burke et al. [106,121] noted the predictions of available reaction models diverge towards elevated pressure. As shown in Fig. 3, the predictions of these models, in fact, lie within the uncertainty band predicted for an arbitrary, nominal model (e.g., the trial model of Davis et al. [101]). Hence, the

differences in the prediction do not imply any missing physics. Rather, they are the result of rate parameter uncertainties whose impact is amplified towards high pressures. While the problem just mentioned is entirely expected, it is also troublesome that a user is often left with the dilemma as to what model to choose, and unfortunately the ultimate selection is often made on an ad hoc basis, because the criterion for choosing a model suitable for a particular application is unavailable.

Development of chemical kinetic models of combustion is not a linear process of reaction pathway and rates postulation followed by a treatment of a mathematical inverse problem. It is very clear that, during the early stage of kinetic mechanism studies, the underlying complexity requires the problem to be tackled piecewise under a degree of chemical isolation. Two important advances have become instrumental to combustion chemistry research: the use of

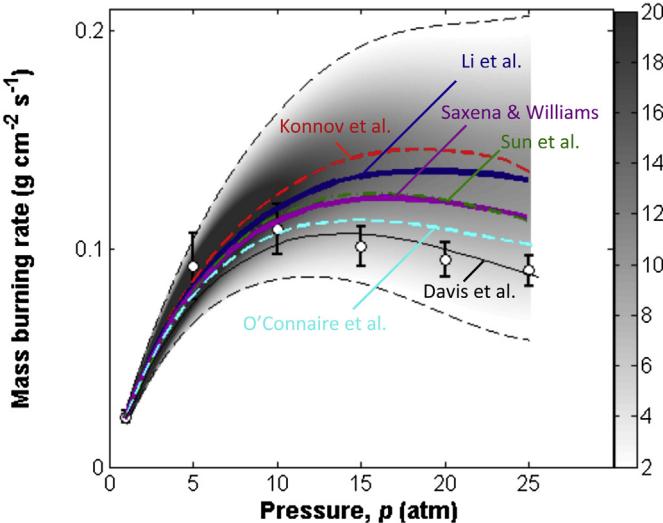


Fig. 3. An illustration of the various reaction models as a sample of rate parameter uncertainty. Symbols: experimental data [106] with the error bars indicate 2σ uncertainty of the data. The shaded-area and the dashed lines bracketing the shaded area are the probability distribution and 2σ uncertainty of the reaction model, respectively. Reaction models used for the simulations are O'Connaire et al. [100], and Davis et al. [101], Saxena & Williams [103], Sun et al. [104], Li et al. [120], and Konnov et al. [105].

electronic structure calculations and reaction rate theories [92,122–124], and advances in laser diagnostics coupled with the shock tube technique [14]. Attempts were also made at addressing the comprehensiveness of a reaction model in terms of the range of thermodynamic state the model covers and the range of combustion phenomena it describes [125].

Given the advances just discussed, one has to ask whether it is possible to rely entirely on first-principles approaches combined with validation against highly accurate experiments to obtain a predictive kinetic model of fuel combustion. Currently, the best demonstrated uncertainty of shock tube experiments is around 15% for rate coefficients at high temperatures [12]. *Ab initio* theories (e.g., the coupled cluster CCSD(T) method) target a chemical accuracy of ± 1 kcal/mol for species thermochemistry and reaction energy barriers (see, e.g., Refs. [126,127]). This accuracy gives an expected uncertainty factor value of 1.65, 1.4 and 1.3 in the rate coefficient at 1000 K, 1500 K and 2000 K, respectively. Combined with the inaccuracies in the vibrational frequencies and the treatment of internal rotations using transition state theory or Rice–Ramsperger–Kassel–Markus theory/master equation modeling (see, e.g., Ref. [122,124,128]), we expect that the total uncertainty in rate predictions to be no better than 50% above 1000 K (see Refs. [129–131] for some recent examples). As was shown in Fig. 2, even if the rate coefficients are all known to within 15%, the uncertainty in predictions of many practical combustion phenomena still can be considerable. In summary, model construction using *ab initio* theories and/or well-designed experiments allowing for submodel isolation can make notable progress, but by themselves will not be enough to achieve truly predictive modeling.

In fact, a complete reaction model, for which each and every rate parameter is sufficiently chemically accurate, will probably never exist. There are two fundamental reasons that support this notion. First, the completeness of a kinetic model cannot be determined a priori. Then, the relevant question to ask is whether the kinetic model can describe known physical observations. Second, even for simple models like that of H₂/CO combustion, the kinetic uncertainties and the dimensionality of the uncertainty space are too large to “pin” the model to a unique point in the uncertainty space.

This situation is exemplified by the discussion made earlier for the mass burning rate of hydrogen. Under this condition, a predictive model can be represented at best by a collection of regimes on an acceptable, hyper uncertainty surface on which every point (or rate parameter combination) gives statistically acceptable predictions for a prescribed set of combustion data. The feasible set, a term from optimization theory recently introduced to the combustion field, defines this rate parameter set [132]. In this context the many competing reaction models in the literature may be viewed as points distributed in the kinetic uncertainty space, some of which lie on the acceptable uncertainty surface, and others that may not be on this surface. A rational goal for developing a predictive reaction model cannot be the identification of a single point (equivalent to a given reaction model), for such a point can be only as arbitrary as any of an infinite number of other points on the acceptable kinetic hyper surface. Rather, the further progress of combustion reaction kinetics must be defined as a progressive reduction of the size of the uncertainty surface using theoretical and experimental tools current available or to be developed in the future.

1.2. Sensitivity analysis and uncertainty quantification in combustion kinetics

A widely used method to understand how the solution of a chemical kinetic model is dependent on the model parameters, and notably the reaction rate coefficients, is sensitivity analysis [17,35,51,133–135]. A first-order local sensitivity analysis calculates the derivative of a model response with respect to model parameters. Higher-order sensitivity determines the joint impact of two or more parameters on a model prediction. As a diagnostic tool, sensitivity analysis is used to uncover those reactions that have the greatest influence on a global combustion property or a local property, e.g., what reactions impact the concentration of a species or temperature. The various sensitivity methods and their relationships to UQ have been discussed in great detail in a review paper by Turányi [133]. UQ is closely related to sensitivity analysis, and in many ways UQ is a natural extension of sensitivity analysis. For example, as first used by Warnatz [136] and further discussed by Turányi et al. [137], the product of the first-order sensitivity coefficient and the uncertainty in the rate coefficient measures of the influence of the uncertainty in an individual rate parameter on the uncertainty in a kinetic model prediction.

One of the commonly used sensitivity measures is the logarithmic sensitivity coefficient. As discussed by Gardiner [138], the sensitivity coefficient of the i^{th} computed quantity or model prediction y_i with respect to the j^{th} rate parameter x_j is

$$S_{i,j} = \frac{\log(y'_i) - \log(y_i)}{\log(x'_j) - \log(x_j)} \quad (1)$$

where x'_j denotes that the parameter j has been altered from some reference value and y'_i denotes the computed quantity i calculated with this modified parameter. If the perturbation is sufficiently small, equation (1) is equivalent to the local slope on a log–log plot of computed quantity versus rate parameter,

$$S_{i,j} = \frac{\partial \log(y_i)}{\partial \log(x_j)} = \frac{x_j}{y_i} \frac{\partial y_i}{\partial x_j} \quad (2)$$

As an invaluable tool, the concept of sensitivity analysis was introduced in the late 1970s and early 1980s. Beyond the brute force method, several numerically efficient methods were introduced [35,139–156]. Shuler and coworkers [140–144] championed the method of Fourier amplitude analysis for multiparameter

model systems. Rabitz and coworkers [145–147] introduced Green's function method of sensitivity analysis in chemical kinetics. The "Direct Method" by Dickinson and Gelinas [148] introduces the use of Jacobian of the model problem in sensitivity calculation. The commonly used method today was originally proposed by Stewart and Sorenson [139], which fits naturally into the solution methods used for problems frequently encountered in combustion chemistry and which works well for both time-independent and time-dependent problems. More importantly, the method was implemented in the CHEMKIN suite of codes [157], which has been the mainstream code for combustion chemistry analysis over the last thirty years. The reader is referred to the review paper of Miller et al. [51] for further details.

Consider a general dynamical system Ξ in vector notation,

$$\Xi(\mathbf{s}, t, \mathbf{y}; \mathbf{x}) = L(\mathbf{s}, t, \mathbf{y}; \mathbf{x}) - f(\mathbf{s}, t; \mathbf{x}) = 0 \quad (3)$$

where t is the time, L represents the conservations of various fluid mechanical and thermodynamic properties and of species mass fractions, $\mathbf{y} = \mathbf{y}(\mathbf{s}, t; \mathbf{x})$ is the solution vector of the system or model response and $f = f(\mathbf{s}, t; \mathbf{x})$ is the source term, both of which depend on the position \mathbf{s} and time t , and are parametrically dependent on the parameters \mathbf{x} , which can be uncertain. Here the source term f is basically a chemical kinetic model applied to a specific problem that is governed by Ξ . In combustion chemistry, Eq. (3) can be a set of conservation equations of a flame, or a set of rate equations describing the time evolution of species and thermodynamic quantities. By the chain rule, we have

$$\frac{\partial \Xi}{\partial y_i} \frac{\partial y_i}{\partial x_j} + \frac{\partial \Xi}{\partial x_j} = 0 \quad (4)$$

In the above equation, $\partial \Xi / \partial y_i$ is the Jacobian J which is readily available as a part of the solution method for both initial value and boundary value problems. Hence, the sensitivity coefficients for a stationary system may be conveniently calculated as

$$J \frac{\partial y_i}{\partial x_j} = - \frac{\partial \Xi}{\partial x_j} \quad (5)$$

Analysis of the first-order sensitivity coefficients, whether they are obtained from brute force calculation or from Eq. (5), quickly

yielded some very useful understanding of combustion chemistry. For example, as Fig. 4 shows, the number of reaction rate coefficients that impact a combustion response is usually very few compared to the total number of reactions considered in a kinetic model; the ranked sensitivity coefficients form the basis for the selection of active rate parameters for model uncertainty minimization, as will be discussed later. Although it is not used frequently, higher-order sensitivity has been discussed [147,158]. In particular, the second-order sensitivity coefficients indicate the importance of parameter coupling, which can be crucial to kinetic model refinement [57].

Among the pioneering work of uncertainty-related analysis in combustion chemistry, the most influential work has been that of Michael Frenklach, who was the first to systematically treat kinetic parameter uncertainties [2,17,35,37,159]. In his studies of systematic optimization of detailed kinetic models [17,159], Frenklach addressed the fundamental issue concerning the role of underlying parameter uncertainty in kinetic modeling and asked whether it is "possible to adjust a large-scale dynamic model in a systematic manner with a reasonable amount of effort." In order to address this question, the method of solution mapping was introduced as a quantitative way to express model predictions as functions of rate parameters; a response surface is generated which contains the first- and second-order sensitivity information explicitly. The role of fundamental combustion experiments was reinterpreted as not only to serve as a measure for comparison with model predictions, but also to provide statistically meaningful information that is used for constraining the joint parameter uncertainty in a kinetic model. It is this reinterpretation that allows for systematic development of the model through multi-parameter optimization against experimental data. Finally, the concept of non-uniqueness of parameter choices, or the feasible region, was introduced. The last point lays the foundation for what is a renewed concept of the feasible parameter set in his later publications [119,132].

The aforementioned studies form the theoretical and practical foundations for the much-celebrated GRI-Mech effort of the 1990s [160–162]—a successful drive of the systematic treatment of kinetic parameter uncertainty through collaborative and coordinated modeling and experimentation, from fundamental rate parameter evaluation, reaction rate theory application, initial trial model testing, to designing experiments for measurements of key rate

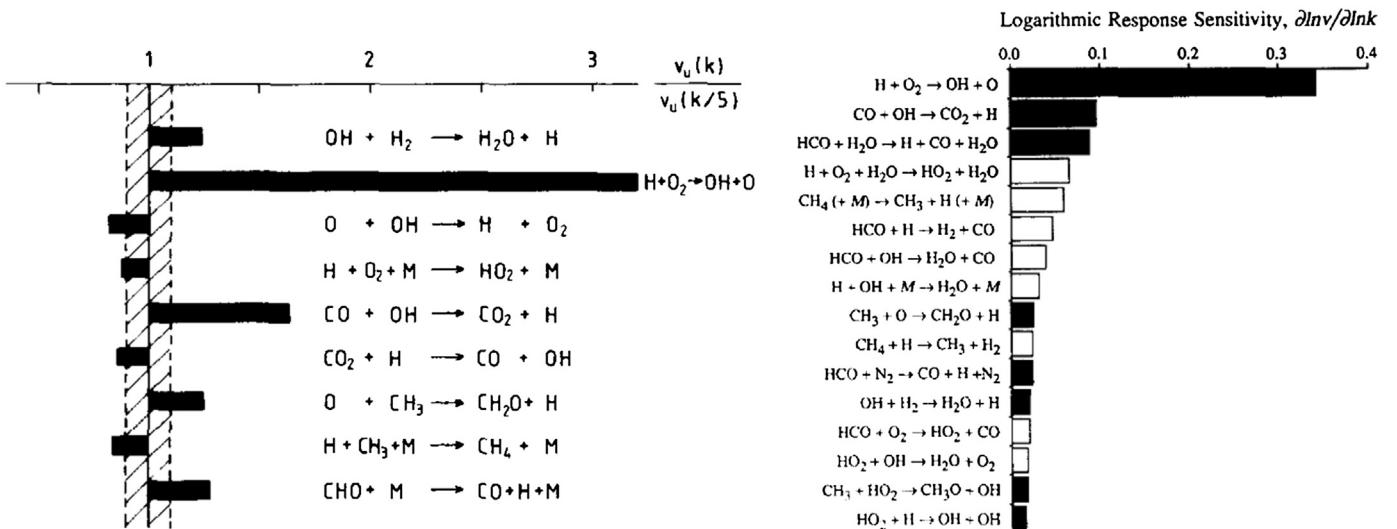


Fig. 4. Examples of sensitivity analyses from two earlier studies on modeling the laminar flame speed of methane-air mixtures using detailed kinetic model. Left panel: brute-force sensitivity spectrum by dividing the rate coefficient by a factor of 5 (adapted from Warnatz [33]); right panel: local sensitivity spectrum (adapted from Frenklach et al. [2]).

parameters and critical combustion properties, and model optimization and interactive dissemination. It also inspired a large range of UQ research that has been conducted over the last decade.

An early effort that bridges the UQ analysis of chemical kinetics to sensitivity analysis is the work of Tamás Turányi. In his 1990 review article, Turányi [133] discussed the various sensitivity methods available at that time, but the real significance of the article is about extending sensitivity analysis to quantitative uncertainty analysis, coupled kinetic parameter estimation, experimental design, and mechanism reduction.

McRae and coworkers [163] championed the method of Deterministic Equivalent Modeling Method (DEMM) to represent the stochastic distribution of kinetic model outputs. The method relies on a representation of parametric uncertainty via polynomial chaos expansions, and utilizes orthogonal collocation to calculate the distributions of the model responses. A series of studies by Najm, Ghenum, and Knio [164–166] further advanced the method of spectral stochastic uncertainty quantification, and showed that polynomial chaos methods and conventional sensitivity analysis provide similar first-order information, but polynomial chaos provide better higher-order information critical to uncertainty quantification [166]. In particular, confidence intervals on sensitivity coefficients, which can be uncertain themselves, may be calculated.

In a recent paper by Frenklach [119], the role of uncertainty analysis and reduction has been discussed in a larger context of model development through Process Informatics Model (PrIMe). A quite comprehensive discussion of the role and methods of sensitivity and uncertainty analyses in combustion chemistry modeling may be found in Tomlin [167]. The intent of the current paper is to provide a self-contained review about the mathematical principles and methods of uncertainty quantification and their application in highly complex, multi-parameter combustion chemistry problems. We intend to provide a comprehensive review and classification for the various UQ methods and illustrate their applications for problems involving forward uncertainty quantification and propagation, and as an inverse problem leading to model uncertainty constraining.

1.3. Error versus uncertainty in computer modeling

The terms “error” and “uncertainty” have often been used interchangeably in many different ways. For instance, as discussed by Raman and co-workers [168], “error” is used in the same sense as the terms “experimental uncertainty” and “model form error” defined below, while “uncertainty” means only “parametric uncertainty.” In this paper, we adopt the following definitions. Uncertainty refers to deficiencies of an experimental measurement or model prediction that is caused by the lack of precise knowledge. Error, on the other hand, is a recognizable, recoverable, deterministic deficiency that is not due to lack of knowledge. Examples of error include a mistyped rate parameter, a mismatch of the thermochemical properties of an isomer due to confusion of species nomenclature, a wrong value in the initial condition of a simulation, rate constants of an elementary reaction not obeying the principle of detailed balancing, and a convergence error that can be resolved by reducing computational error tolerance.

Two sources of uncertainty may exist. They can be categorized into aleatory and epistemic uncertainties. Aleatory uncertainty is inherent to a probabilistic process. This uncertainty is irreducible and must be characterized by a probability distribution. Epistemic uncertainty is due to limited knowledge: e.g., an incomplete understanding of the underlying physics (missing pathways), and imprecise evaluation of a rate constant. Unless otherwise indicated, what we discuss here is the epistemic uncertainty.

1.4. Source of uncertainties

The sources of model uncertainties can be many; and they may be classified into several categories. Oberkampf and Roy [169] divide the sources of uncertainty into three major groups, which are model input uncertainty, model form error, and numerical uncertainty. Numerical uncertainty comes from procedures for the solution of a particular physical problem using a given model. In general, numerical uncertainty can be minimized by careful convergence tests. For example, it has been long known that as an eigenvalue problem, the laminar flame speed or the mass burning rate can be a strong function of the computational domain size and local gradients. In principle, a properly converged solution requires sensitivity tests with respect to the domain size and grid refinement parameters. The other two categories can be subdivided further.

Model input uncertainty refers to uncertainty in the data that is being used to build the model, that is, the notion that a model will always have some inherent “fuzziness.” It can be split into parameter uncertainty and experimental uncertainty. **Parameter uncertainty** stems from chemical kinetic rate parameters, thermodynamic or transport properties whose exact values are unknown, cannot be known beyond a certain accuracy, or are only approximate due to assumptions in the fundamental theory (e.g., rigid-rotor, harmonic oscillator treatment of partition functions). **Experimental uncertainty** refers to the measurement uncertainty of fundamental combustion properties that are used to constrain the model. This uncertainty can be best quantified by careful assessments of similar errors in different apparatus and by repeating measurements across a range of experimental apparatus in one or more laboratories.

Model form error refers to uncertainties in the assumptions inherent in developing the model. It can be split into structural uncertainty and interpolation and extrapolation uncertainty. **Structural uncertainty** is the result of lack of knowledge of the underlying physics. Examples include missing reaction pathways, incomplete reaction description, or ill-formulated conservation equations. **Interpolation and Extrapolation uncertainty** is the result of a lack of available laboratory experiments for a particular set of thermodynamic conditions. The same category of uncertainty is applicable to elements of reaction model development that employs the reaction class rule or analogous reactions. In the context of UQ of chemical kinetics, interpolation and extrapolation uncertainties can be related to parameter uncertainty. For example, extrapolating high-temperature shock-tube measurement of a rate coefficient to low temperatures can lead to tremendously enlarged uncertainty at low temperatures because of the Arrhenius behavior, and in some cases, simple extrapolation can lead to violation of the collision theory and structural uncertainties. It is for this reason that rate coefficient extrapolation are appropriate only when such an extrapolation are guided, at least to some extent, by reaction rate theories.

The focus of the current review is on uncertainty quantification that results from model input uncertainties, although a proper treatment of these uncertainties can lead to useful conclusion about model form errors. For example, failure for a chemical reaction model to predict a particular combustion measurement, within its experimental uncertainty and also within the uncertainty bounds of the model rate parameters (and thermodynamic and transport properties), is a suggestion that the model is deficient beyond parameter uncertainty. The model perhaps has missing reaction pathways, or incomplete reaction descriptions, or the accuracy of the experiment is in question.

1.5. Model verification and validation

Before we start to discuss the various UQ methods for combustion chemistry analysis, we wish to make a brief remark about what is meant by model validation and verification, as these terms have been frequently used in the literature. There has been much discussion about rigorous definitions for these terms. A review is provided in the book, Verification and Validation in Scientific Computing by Oberkampf and Roy [169], which we will summarize here. The first definitions of verification and validation were provided by the Society for Computer Simulation (SCS) [170], in a short article providing concise definitions for a number of terms. In this work, SCS divided modeling and simulation into three key components, which were reality, the conceptual model, and the computerized model. Reality was defined as what you can see and touch. The conceptual model is the mathematical description of reality. The computerized model is the executable code that implements the conceptual model.

1.5.1. Verification

Multiple definitions of verification have been used for verification in computer modeling, differing in detail but similar in substance to the definition of the SCS, which is “substantiation that a computerized model represents a conceptual model within specified limits of accuracy.” As an example, the American Society for Mechanical Engineers’ Guide for Verification and Validation in Computational Solid Mechanics divides verification into two components. These are code verification, which ensures that the computer code is correctly implemented (i.e., there are no mistakes), and solution verification, which ensures that the numerical solution agrees with known solutions to within certain tolerances and also addresses questions of numerical stability. For instance, in certain special cases, the conceptual model will admit analytical solutions, to which solutions of the computerized model must converge for it to be verified.

1.5.2. Validation

As with verification, model validation has many definitions, all similar to the SCS’s, “substantiation that a computerized model within its domain of applicability possesses a satisfactory range of accuracy consistent with the intended application of the model.” This is divided by Oberkampf and Trucano [171] into three aspects, which are comparison with experimental data, extrapolation of the model and associated uncertainty to its intended domain, and assessing whether the model meets certain precision requirements in its intended domain. It is the first aspect that is usually meant by “validation” in the combustion community. As a general practice, unfortunately, the community usually ignores the second and third aspects of validation.

Validation includes extrapolation because experimental data may not be available in the intended domain of the model. This means that the uncertainty of the model must be estimated over that domain. If the model is later to be optimized against the validation data, then the corresponding measurement uncertainty must also be estimated and propagated. Without an uncertainty estimate, it can be said that the model is being compared with or tested against experiment, but it cannot be said to be validated.

An important distinction between verification and validation is that verification is a well-posed problem with a definite answer. In the case of code verification, for instance, the computerized model is being compared with a known (possibly analytical) solution, and the numerical solution either agrees with the analytical solution or it does not. Validation, on the other hand, is poorly posed. An experiment gives an answer, but that answer is subject to interpretations and assumptions. It can never be known whether

those assumptions hold in all cases, and as such the validation data are not “true” in any more concrete sense than are the predictions of the model.

2. Uncertainty quantification

2.1. Estimation of rate constant uncertainties

Rate coefficients are evaluated typically for their most probable values as a function of temperature (and pressure). Such evaluation is not a necessary starting point or even a component of UQ analysis though it has been used in almost all kinetic modeling studies. The uncertainty in the rate coefficient is given typically by the uncertainty factor f . In a symmetric form f may be written as

$$f(T) = \frac{k^0(T)}{k^{\text{lower}}(T)} = \frac{k^{\text{upper}}(T)}{k^0(T)} \quad (6)$$

where k^0 is the nominal rate coefficient with its lower and upper bounds given as k^{lower} and k^{upper} , respectively. The expression states that a rate value outside of the range $[k^0/f, fk^0]$ is highly improbable. This description does not express, however, the probability distribution of the rate coefficient uncertainty. Historically, the uncertainty factor is often interpreted as simple bounds of uncertainty; and the probability distribution of uncertainty is a rectangular function, as in the GRI-Mech effort [160–162]. In other cases, the probability distribution of uncertainty is treated as being lognormal with the uncertainty factor interpreted as the geometric, two standard deviations. As will be discussed later (Section 2.5), the choice for the distribution is not as crucial as one would expect when making forward UQ projections.

The cause for the difficulty in probability distribution assignment is quite simple and can be understood by examining the rate evaluation process and procedure. In Fig. 5, we show an Arrhenius plot for the rate coefficient of reaction $\text{H} + \text{CH}_4 \rightarrow \text{CH}_3 + \text{H}_2$, evaluated by Baulch et al. [118], for which a relatively large number of experimental data are available. It is seen that the quoted $f = 2$ uncertainty limits brackets most of the available experimental data. The evaluation rejects some of the low-temperature values, and this is often the result of a joint consideration of available data and fundamental reaction rate theories. In other cases, as depicted in Fig. 6 for $\text{H} + \text{HCO} \rightarrow \text{CO} + \text{H}_2$, very limited rate data are available and they can scatter over a quite wide range. In this case, certain data or analysis would receive greater weights and the uncertainty band can lie within a range in which a substantial number of data points are rejected. It is not straightforward to give a rigorous description for the probability distribution of the rate uncertainty and the coupling of the Arrhenius parameters when data are sparse, as in this case.

A statistical treatment of the Arrhenius rate parameters becomes possible when a sufficient amount of data becomes available. Turányi and coworkers [172–174] derived an analytical expression that describes the temperature dependence of the uncertainty of k as a function of the elements of the covariance matrix of the Arrhenius parameters. Such a treatment enables the joint uncertainty of the Arrhenius parameters to be defined and leads to a more rigorous representation of the rate coefficient uncertainty. In their treatment, the logarithm of the rate constant $k(T) = AT^n e^{-E/RT}$ is defined as $\kappa(T) = \mathbf{p}^T \boldsymbol{\theta} = [\alpha \ n \ \varepsilon]^T [1 \ \ln T \ -T^{-1}]$, where $\kappa(T) = \ln k(T)$, $\alpha = \ln A$ and $\varepsilon = E/R$. The uncertainty in $\kappa(T)$ is given by $\sigma_\kappa^2(T) = \boldsymbol{\theta}^T \Sigma_p \boldsymbol{\theta}$, where Σ_p is the covariance among the Arrhenius parameters. If the uncertainty in the rate coefficient is specified at enough temperatures, Σ_p can be estimated by least-squares regression. As an example, the uncertainty in the Arrhenius A factor and activation energy for the reaction $\text{O} + \text{N}_2\text{O} \rightarrow 2\text{NO}$ was

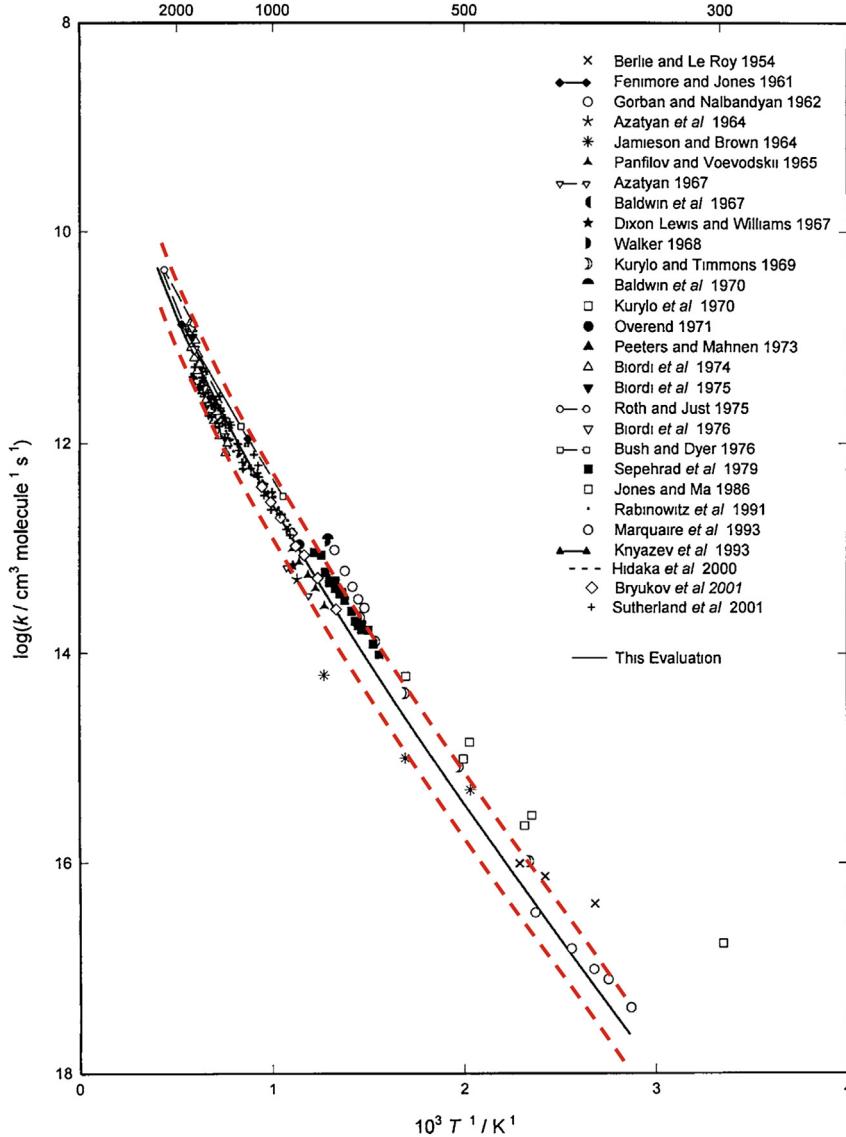


Fig. 5. Sample evaluated rate coefficient: $\text{CH}_4 + \text{H} \rightarrow \text{CH}_3 + \text{H}_2$. The solid line is the rate expression recommended in the evaluation. The dashed lines are added here and indicate the upper and lower bounds according to the evaluated uncertainty factor $f (=2.0)$. The plot is adapted from Baulch et al. [118].

calculated and shown in Fig. 7. In this case, σ_κ is specified to be 0.4 at 1000 K, 0.2 at 2000 K, and 0.3 at 4000 K. As can be seen from the figure, the Arrhenius parameters are tightly correlated. The point of the work of Nagy and Turányi [172] was that reliable estimates of the uncertainties in rate parameters could be made when a sufficient amount of data are available.

In spite of the aforementioned advance, a proper treatment of the rate coefficient uncertainty remains challenging. There is a need for a general protocol to account for statistical representation of measurement uncertainty, theoretical interpretation of the data, and expert opinion.

2.2. Representations of model response and uncertainty

A basic objective of UQ is to analyze the uncertainty in the model response \mathbf{y} as a result of model parameters \mathbf{x} being uncertain. In this section, we review the theories and introduce the various methods that can be used to describe the uncertainty in \mathbf{y} .

2.2.1. The Taylor series

In the simplest case, a model prediction y as a function of \mathbf{x} can be expressed in terms of a Taylor series

$$y = y_0 + \mathbf{g}^T \mathbf{x} + \mathbf{x}^T \mathbf{H} \mathbf{x} + \text{h.o.t.} \quad (7)$$

where $\mathbf{g} = \nabla y$ is the gradient of y , $\mathbf{H} = \frac{1}{2}(\nabla y)$ is the Hessian and J is the Jacobian operator, and h.o.t. represents higher order terms. Note that the model parameters \mathbf{x} are normalized here so that, when $\mathbf{x} = 0$ (a null vector), $y = y_0$. The uncertainty in y , σ_y , is related to the uncertainty in \mathbf{x} . If we consider some set of independent, identically distributed (i.i.d.) random variables ξ , we can express the model parameters as random variables such that $\mathbf{X} = \mathbf{x}^{(0)} + \mathbf{x}^{(1)}\xi$, where $\mathbf{x}^{(0)}$ is the mean of \mathbf{X} , and $\mathbf{x}^{(1)}$ is a transformation matrix such that the covariance matrix is given by $\Sigma = \mathbf{x}^{(1)T} E[\xi\xi^T] \mathbf{x}^{(1)}$, where $E[\cdot]$ is the expectation operator. If the model parameters have some prescribed uncertainty and are independent of each other, $\mathbf{x}^{(1)}$ is diagonal and $x_{ii}^{(1)} = \Theta \sigma_{x,i}$, where $\sigma_{x,i}$ is the uncertainty in the i th model parameter and Θ is a constant that depends on the probability density function (PDF)

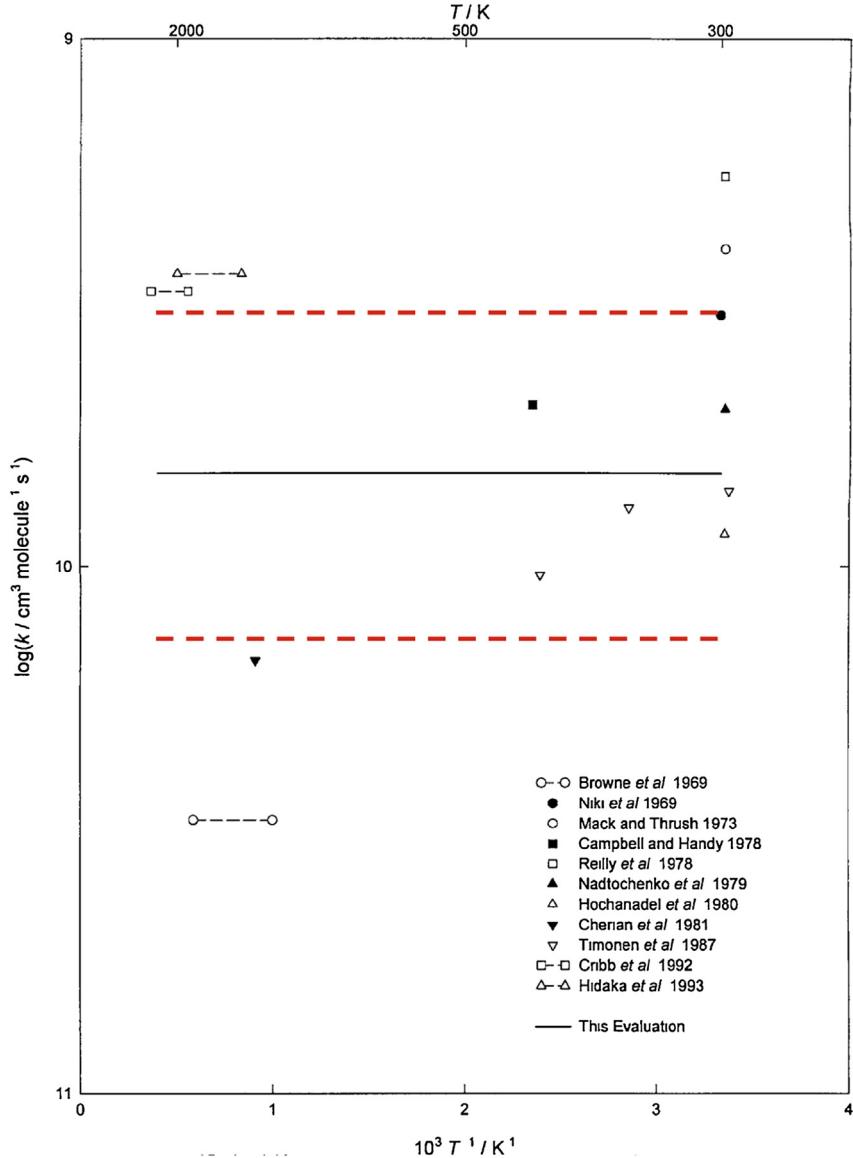


Fig. 6. Sample evaluated rate coefficient: $\text{HCO} + \text{H} \rightarrow \text{CO} + \text{H}_2$. The solid line is the rate expression recommended in the evaluation. The dashed lines are added here and indicate the upper and lower bounds according to the evaluated uncertainty factor $f (=2.0)$. The plot is adapted from Baulch et al. [118].

of ξ . A typical choice is that the members of ξ are normally distributed with mean 0 and variance 1, in which case Θ is 1 or $\frac{1}{2}$, depending on whether $\sigma_{x,i}$ is interpreted as a 1σ or a 2σ uncertainty, respectively.

A critical question is whether the rate parameters are independent, *a priori*. Usually, their relationship depends on the way they have been determined. If the rate evaluation comes from an experimental measurement, usually the rate under consideration can be correlated, perhaps loosely and depending on how well the experiment is designed, with some other secondary parameters. In this case, the correlation is best handled as an inverse uncertainty problem, which is detailed in Section 3. Alternatively, though, the rate evaluations might come from rules for analogous reactions. For instance, in the work of Cai and Pitsch [175] and Prager et al. [176], all reaction rate constants in a given class, such as H abstraction of various hydrocarbons by OH, were assumed to be equal to each other and, therefore, to be perfectly correlated. This assumption resulted in greater model uncertainty but also in a model with

fewer degrees of freedom. The implications of rate rules and uncertainties are discussed further in Section 4.1.

Once the PDF of ξ has been prescribed, σ_y can be computed by replacing \mathbf{x} of Eq. (7) by \mathbf{X} and then taking the variance.

$$\sigma_y^2 = E[yy^T] - E^2[y]. \quad (8)$$

Substituting Eq. (7) into Eq. (8), and exploiting the mutual independence of ξ as well as the symmetry of \mathbf{H} , we obtain

$$\begin{aligned} \sigma_y^2 &= \mathbf{g}^T \Sigma \mathbf{g} + 4\mathbf{x}^{(0)T} \mathbf{H} \Sigma \mathbf{H} \mathbf{x}^{(0)} + 4\mathbf{g}^T \Sigma \mathbf{H} \mathbf{x}^{(0)} \dots \\ &\quad + \sum_{i=1}^n \left(\mathbf{x}^{(1)T} \mathbf{H} \mathbf{x}^{(1)} \right)_{ii}^2 \sigma_{\xi_i}^2 \left[\xi_i^2 \right] \\ &\quad + 4 \sum_{i=1}^n \sum_{j=i+1}^n \left\{ \left(\mathbf{x}^{(1)T} \mathbf{H} \mathbf{x}^{(1)} \right)_{ij}^2 \sigma_{\xi_i}^2 \sigma_{\xi_j}^2 \left[\xi_i \right] \left[\xi_j \right] \right\}, \end{aligned} \quad (9)$$

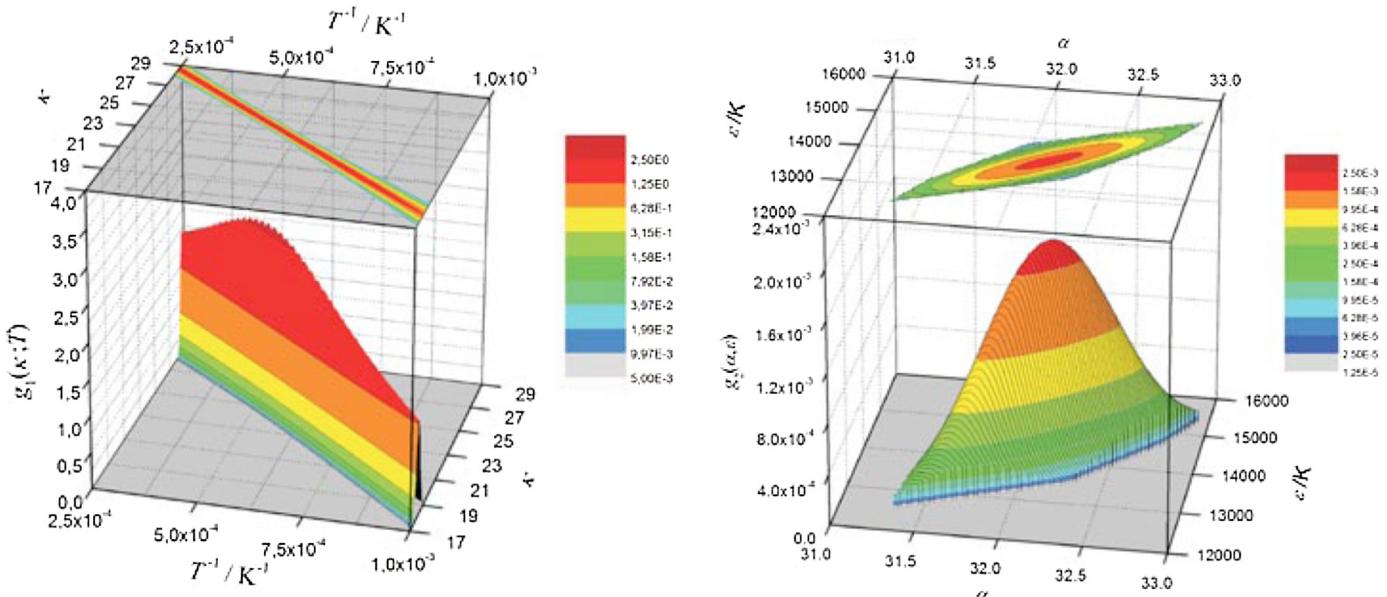


Fig. 7. Three-dimensional representation and two-dimensional projection of the temperature dependent probability density function of $\ln k$ (left) and the joint probability distribution function of the Arrhenius parameters (right), both for reaction $O + N_2O \rightarrow 2NO$. The plots are adapted from Nagy and Turányi [172].

where n is the number of uncertain parameters. In the special case that ξ_i are normally distributed with mean 0 and variance 1, $\sigma^2[\xi_i^2] = 2$ and $\sigma^2[\xi_i] = 1$. Taking $\mathbf{J} = 2\mathbf{H}\mathbf{x}^{(0)} + \mathbf{g}$, Eq. (9) simplifies to

$$\sigma_y^2 = \mathbf{J}^T \Sigma \mathbf{J} + 2\text{tr}(\mathbf{H}\Sigma\mathbf{H}) \quad (10)$$

where $\text{tr}()$ is the trace of a square matrix. This is the most familiar representation of uncertainty, and typically the form found in statistics text. An example of uncertainty calculated with Eq. (10) is presented in the left panel of Fig. 8, adapted from Ref. [178]. In this figure, laminar flame speeds of C_2H_4 /air mixtures at 5 atm and 298 K are presented. Laminar flame speeds were modeled using

USC Mech II [16]. Each A factor was assigned a log-normal distribution about its nominal value. Taylor series terms y_0 , \mathbf{g} , and \mathbf{H} were calculated in increments of 0.2 in equivalence ratio, and the $2\sigma_y$ uncertainty limits were calculated. The probability density between those limits was calculated by fitting a normal distribution with that value of y_0 and σ_y . The uncertainty band was then determined by fitting a cubic spline to the calculated uncertainty and probability density.

2.2.2. Polynomial chaos expansion

The polynomial chaos representation was developed by Wiener [179], originally as a stochastic description of statistical mechanics,

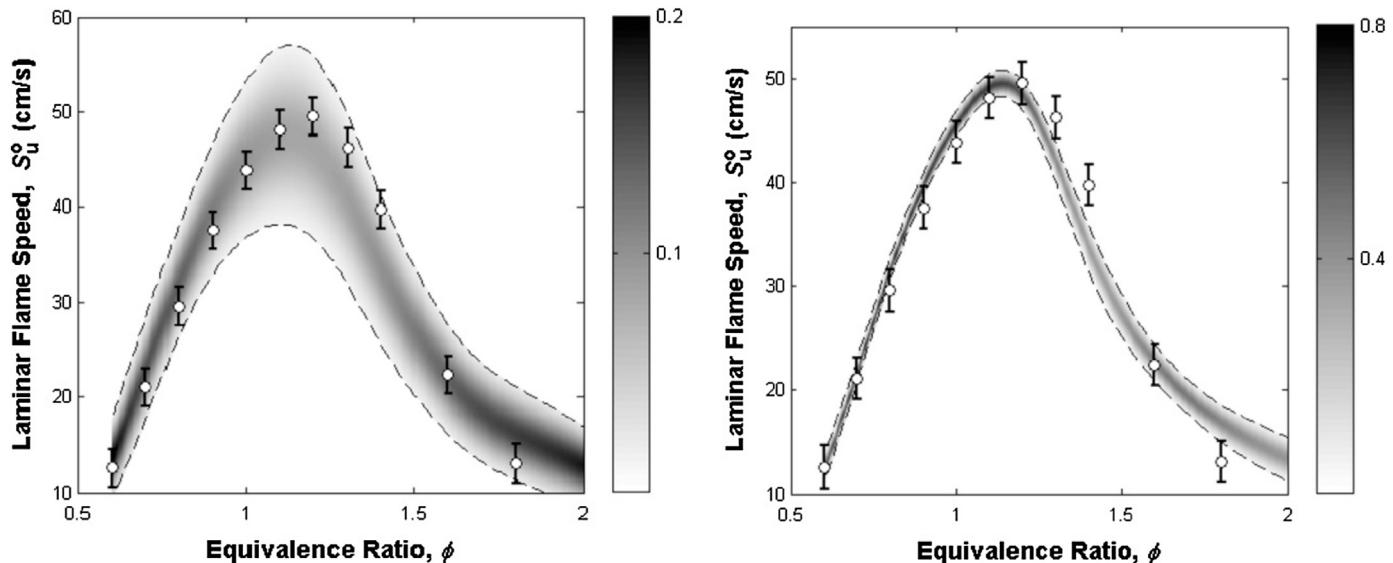


Fig. 8. Laminar flame speed with respect to equivalence ratio for C_2H_4 /air mixtures at 5 atm. Symbols are experimental measurements [177]. The dashed lines indicate the modeled $\pm 2\sigma$ uncertainty limits and the shading indicates the probability density of flame speed as shown by the gray color bar. The left panel depicts the unconstrained model prediction; the right panel is the constrained model prediction. The plot is adapted from Sheen and Wang. [178]. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

and later extended to various other systems. A brief summary, following Najm, Ghanem, and coworkers [180], is presented here.

Consider a set of arbitrary basis random variables ξ . There is then a set of basis polynomials $\Psi(\xi)$ that are orthogonal, i.e., $E[\Psi_i \Psi_j] = 0$ for $i \neq j$. For example, if each ξ_i is normally distributed with mean 0 and variance 1, each Ψ_k is a Hermite polynomial. Together, ξ and Ψ form a polynomial chaos (PC) basis. A complete review of various PC bases was provided by Xiu and Karniadakis [181], showing how different basis random variables are paired with corresponding orthogonal polynomials, as described in what follows.

A PC basis is characterized by its dimension N_d and its order N_o . The size of the expansion P is given by

$$P = \frac{(N_d + N_o)!}{N_d! N_o!} - 1 \quad (11)$$

In principle, N_d and N_o , and therefore P , are infinite, but computational considerations require truncating the expansion at a suitable level. A model prediction y can be decomposed in terms of the basis functions.

$$y = \sum_{k=0}^P y_k \Psi_k \quad (12)$$

The spectral mode strengths y_k can be determined by multiplying the expansion in Eq. (12) by the associated basis function Ψ_k and then taking the expected value.

$$E[y \Psi_k] = \sum_{i=0}^P E[y_i \Psi_i \Psi_k] \quad (13)$$

By the orthogonality relation, y_k is then

$$y_k = \frac{E[y \Psi_k]}{E[\Psi_k^2]} \quad (14)$$

Collectively, the set of spectral mode strengths \mathbf{y} is the polynomial chaos expansion (PCE) for y . By the Cameron–Martin Theorem [182], any random variable can be approximated to an arbitrary precision using a PCE of suitable order and dimension.

The PCE for a random variable y is related to the Taylor series through the polynomial basis functions. In the case of the Wiener–Hermite chaos, ξ are normally distributed with mean 0 and variance 1, and Ψ are the multidimensional Hermite polynomials $H_n(x)$. The first few such functions are $H_0(x) = 1$, $H_1(x) = x$, and $H_2(x) = x^2 - 1$, and so the first few basis functions are

$$\begin{aligned} \Psi_0 &= 1, & \Psi_1 &= \xi_1, & \Psi_2 &= \xi_2, \dots & \Psi_{N_d} &= \xi_{N_d}, \\ \Psi_{N_d+1} &= \xi_1^2 - 1, & \Psi_{N_d+2} &= \xi_1 \xi_2, \dots, & \Psi_{2N_d} &= \xi_1 \xi_{N_d}, & \Psi_{2N_d+1} &= \xi_2^2 - 1, \dots \end{aligned}$$

The basis functions have the property that, for $k > 0$, $E[\Psi_k] = 0$, since $\Psi_k = \Psi_k \Psi_0$. As a result, the expectation of y is the first spectral mode strength y_0 . The uncertainty, which can be calculated from Eq. (8), is then

$$\sigma_y^2 = \sum_{k=1}^P y_k^2 E[\Psi_k^2]. \quad (15)$$

The lower bound of the above summation is $k = 1$ rather than 0 in order to exclude the mean. It should also be noted that every odd-order Hermite polynomial contains a first-order term, and every even-order Hermite polynomial contains a second-order term, and so \mathbf{g} and \mathbf{H} have a contribution from each spectral mode strength.

2.2.3. High-dimensional model representation

The High Dimensional Model Representation (HDMR) was introduced by Sobol' [183] and first used in chemical kinetics by Rabitz [184] as a means to more efficiently represent model outputs. It argues that a polynomial regression such as the Taylor series can be inefficient for arbitrary functions of the model parameters. In HDMR, the model prediction y is expressed as a set of component functions.

$$y(\mathbf{x}) = y_0 + \sum_{i=1}^n y_i(x_i) + \sum_{i=1}^{n-1} \sum_{j=i+1}^n y_{ij}(x_i x_j) + \text{h.o.t.} \quad (16)$$

In this representation, y_i is a function of only the single variable x_i ; likewise the function y_{ij} describes pairwise interactions between variables x_i and x_j . This representation differs from the Taylor series in Eq. (7) in that the form of the functions y_i and y_{ij} are not specified; they can be polynomials of arbitrary order, or for that matter, any function. In this case, h.o.t. represents coupling of three or more variables.

The functions y_i and y_{ij} can be determined through a method following Rabitz [184]. If the model prediction y is defined in some region of the rate parameter space, denoted $K^n \subseteq \mathbb{R}^n$, the components of the HDMR expansion are given by

$$y_0 = \int_{K^n} y(\mathbf{x}) d\mathbf{x} \quad (17)$$

$$y_i(x_i) = \int_{K^{n-1}} y(\mathbf{x}) d\mathbf{x} / dx_i - y_0 \quad (18)$$

$$y_{ij}(x_i, x_j) = \int_{K^{n-2}} y(\mathbf{x}) d\mathbf{x} / dx_i - y_i(x_i) - y_j(x_j) - y_0 \quad (19)$$

where K^{n-1} and $d\mathbf{x}/dx_i$ denote that the integral is taken over all of the variables except for x_i , and likewise for K^{n-2} and $d\mathbf{x}/dx_i dx_j$. These functions are orthogonal to each other, and can therefore be expressed as an expansion of basis orthogonal functions φ , e.g.,

$$y_i(x_i) \approx \sum_{k=1}^l y_{i,k} \varphi_k(x_i) \quad (20)$$

$$y_{ij}(x_i, x_j) \approx \sum_{p=1}^l \sum_{q=1}^l y_{ij,pq} \varphi_p(x_i) \varphi_q(x_j) \quad (21)$$

where the expansion is truncated at a suitable order l . The individual coefficients can be determined by exploiting the orthogonality of the functions φ ,

$$\int_{K^n} \varphi_k(x_i) y(\mathbf{x}) d\mathbf{x} = y_{i,k} \int_{K^n} \varphi_k^2(x_i) d\mathbf{x} \quad (22)$$

which follows from the fact that $\int_{K^n} \varphi_i \varphi_j d\mathbf{x} = 0$ for $i \neq j$. The prediction uncertainty, again calculated from Eq. (8), takes on a similar form in the case of HDMR as for the polynomial chaos expansion, i.e.,

$$\sigma_y^2 = \sum_{i=1}^n \sum_{k=1}^I y_{i,k}^2 + \sum_{i=1}^{n-1} \sum_{j=i+1}^n \sum_{p=1}^I \sum_{q=1}^I y_{ij,pq}^2 + \dots \quad (23)$$

A common choice for φ is a set of polynomials which are orthogonal over K^n [185,186]. In this case, then, the HDMR expression is closely related to the polynomial chaos expansion. The challenge of both HDMR and the polynomial chaos expansion is that they require multidimensional integrals in order to calculate the coefficients of the expansion, namely Eqs. (14) and (22).

2.3. Intrusive methods

When performing numerical simulations, computations are almost always more efficient if an analytic Jacobian can be calculated. Likewise, uncertainty calculations are almost always more efficient if the system is amenable to an analytical treatment of the uncertainty. These methods are termed intrusive because the governing equations must be modified to take advantage of the analytical treatment. Two intrusive methods are presented below: the stochastic spectral projection method and the sensitivity-analysis-based method.

$$\dot{\mathbf{u}}_k + \frac{\sum_{i=0}^P \sum_{j=0}^P (\mathbf{u}_i \cdot \nabla) \mathbf{u}_j E[\psi_i \psi_j \psi_k]}{E[\psi_k^2]} = -\nabla p_k + \frac{\sum_{i=0}^P \sum_{j=0}^P Re_i^{-1} \nabla^2 \mathbf{u}_j E[\psi_i \psi_j \psi_k]}{E[\psi_k^2]} \quad (30)$$

2.3.1. Stochastic spectral projection

The intrusive stochastic spectral projection method is a technique for analytically calculating the terms in a PC representation of a simulation result. Its advantage is that, in principle, it is straightforward to apply to any system of differential equations. This methodology was applied by Ghanem and Spanos [187,188] to problems in civil engineering, and later to fluid mechanics by Najm, Ghanem, and coworkers [180,189] and by Xiu and Karniadakis [190,191] and by Xiu and Karniadakis [190,191]. The application of the method was subsequently extended to chemical kinetics [164–166,192–194].

Recall the generalized dynamical system defined in Eq. (3), $L(\mathbf{s}, t, \mathbf{y}, \mathbf{x}) = f(\mathbf{s}, t, \mathbf{x})$, where L represents the governing physical equations and f is a forcing term, both of which depend on the system state \mathbf{y} , position \mathbf{s} and time t , and are parametrically dependent on the model parameters \mathbf{x} . Because we treat \mathbf{x} as random variables which have PCEs, \mathbf{y} and \mathbf{f} are also random variables and have PCEs, e.g., $\mathbf{y} = \sum_{k=0}^P \mathbf{y}_k \psi_k$. Then the system L itself can also be expressed in a PC form. As an example, following Xiu and Karniadakis [181], consider the first-order system $\dot{y} = -xy$. The constant x and the solution y have PCEs $y = \sum_{k=0}^P y_k \psi_k$ and $x = \sum_{k=0}^P x_k \psi_k$. These PCEs can be substituted into the original system, which yields

$$\sum_{i=0}^P \dot{y}_i \psi_i = \sum_{i=0}^P \sum_{j=0}^P y_i x_j \psi_i \psi_j \quad (24)$$

The spectral mode strengths \dot{y}_k are given by

$$\dot{y}_k = \frac{E[\dot{y} \psi_k]}{E[\psi_k^2]} = \frac{\sum_{i=0}^P \sum_{j=0}^P y_i x_j E[\psi_i \psi_j \psi_k]}{E[\psi_k^2]} \quad (25)$$

As can be seen from Eq. (25), expressing L in terms of the PC basis transforms the single dynamical system $L(\mathbf{y})$ into $P + 1$ coupled dynamical systems $L_k(\mathbf{y})$. These coupled equations are the

PC spectral projection of L . Typically, the uncertainty in the parameters \mathbf{x} would be propagated into the solution \mathbf{y} . For instance, the description of $\mathbf{X} = \mathbf{x}^{(0)} + \mathbf{x}^{(1)}\xi$ of Section 2.2.1 is an example of a PCE for \mathbf{x} , with dimension $N_d = n$ and order $N_o = 1$.

As another example, the PC spectral projection method was applied to the incompressible Navier–Stokes equations [180,191], which give the velocity field \mathbf{u} in terms of the pressure p and Reynolds number Re .

$$\nabla \cdot \mathbf{u} = 0 \quad (26)$$

$$\dot{\mathbf{u}} + (\mathbf{u} \cdot \nabla) \mathbf{u} = -\nabla p + Re^{-1} \nabla^2 \mathbf{u} \quad (27)$$

The PC spectral projection for the Navier–Stokes equations is then given by Ref. [180,191]

$$\nabla \cdot \mathbf{u}_k = 0 \quad (28)$$

$$\dot{\mathbf{u}}_k + \frac{\sum_{i=0}^P \sum_{j=0}^P (\mathbf{u}_i \cdot \nabla) \mathbf{u}_j E[\psi_i \psi_j \psi_k]}{E[\psi_k^2]} = -\nabla p_k + Re^{-1} \nabla^2 \mathbf{u}_k \quad (29)$$

assuming that Re is deterministic, or

if Re is random. For example, Najm, Ghanem, and coworkers [180] considered an uncertain viscosity. Xiu and Karniadakis [190,191] considered uncertain boundary conditions. Najm and coworkers observed difficulties in obtaining convergence with highly-nonlinear systems such as chemically-reacting systems [192], but this problem was solved by using a multi-wavelet expansion method [165,193,194], wherein the basis random variables are piecewise-uniform random variables.

As an example taken from Ref. [194], Fig. 9 shows the time evolution of the PDF of temperature during ignition of a stoichiometric methane/air mixture at 1050 K. At early times, the temperature PDF shows one peak near 1500 K due to the initially slow oxidation process. As time progresses, a peak near 2700 K begins to form due to the system reaching equilibrium, while the low-temperature peak spreads out. There is very little probability of the system being between the two peaks because the ignition, and therefore transition between the low- T and high- T states, is fast. The multi-wavelet formulation captures the bifurcation between the two states.

2.3.2. Sensitivity-analysis-based method

The complete PC representation of the energy conservation equation can prove cumbersome to calculate, as has been noted previously [164,166,192]. An alternative method for calculating the Taylor series is the Sensitivity-Analysis-Based method (SAB) of Wang and coworkers [195]. Many algorithms provide, along with the solution variables \mathbf{y} , the sensitivity coefficients \mathbf{S} , where the sensitivity S_i of a solution variable y to the i^{th} model parameter θ_i is given by

$$S_i = \frac{\partial y}{\partial \theta_i} \frac{\theta_i}{y} \quad (31)$$

The model parameters θ are then related to the normalized parameters \mathbf{x} [2,35] by

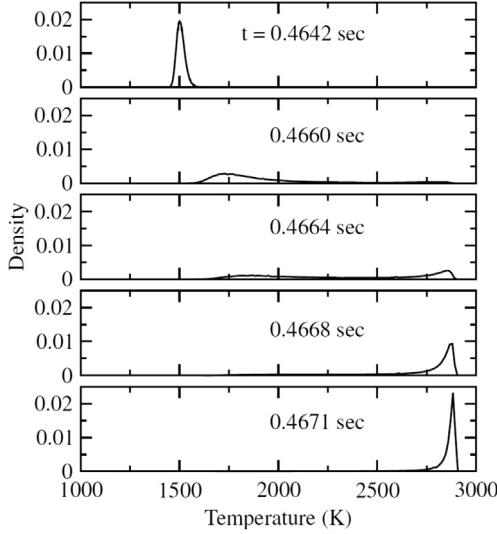


Fig. 9. Time evolution of the PDF of temperature during ignition of a stoichiometric methane/air mixture at 1 atm pressure and an initial temperature of 1050 K. The chemical model is a single-step irreversible model. The plot is adapted from Najm et al. [194].

$$x_i = \frac{\ln(\theta_i/\theta_{i,0})}{\ln f_i} \quad (32)$$

where $\theta_{i,0}$ is the nominal value and f_i is its prescribed uncertainty factor. This normalizes the model parameters so that when $x_i = \pm 1$, θ_i takes on its maximum or minimum values, respectively. It can then be shown that the gradient \mathbf{g} is related to the sensitivity coefficients by

$$g_i = y_0 S_i \ln f_i \quad (33)$$

The terms of the Taylor series are then computed by performing $2n + 1$ sensitivity calculations, one at $x = 0$ and then others at $x_i = \pm \Delta x$, with $x_{j \neq i} = 0$. The perturbation Δx is chosen arbitrarily; ideally, the derivatives should be taken in the limit as $\Delta x \rightarrow 0$, but finite values might be chosen to improve the accuracy of the computed Taylor series over a wider range of \mathbf{x} [195]. The terms of \mathbf{g} and Hessian \mathbf{H} are calculated by finite difference.

$$g_i = \frac{y|_{x_i=\Delta x} - y|_{x_i=-\Delta x}}{2\Delta x} \quad (34)$$

$$H_{ij} = H_{ji} = \frac{1}{2} \frac{\partial^2 y}{\partial x_i \partial x_j} = \frac{1}{2} \frac{\partial g_i}{\partial x_j} = \frac{g_i|_{x_j=\Delta x} - g_i|_{x_j=-\Delta x}}{4\Delta x} \quad (35)$$

The relationship between the first derivative and the sensitivity coefficient can be exploited, so that $g_i|_{x_j=\Delta x} = y_0 S_i|_{x_j=\Delta x} \ln f_i$. Eq. (35) can be rearranged to ensure that \mathbf{H} is symmetric, which yields

$$H_{ij} = H_{ji} = \frac{1}{2} \left(\frac{1}{2} \frac{\partial g_i}{\partial x_j} + \frac{1}{2} \frac{\partial g_j}{\partial x_i} \right) = \frac{\left(S_i|_{x_j=\Delta x} - S_i|_{x_j=-\Delta x} \right) \ln f_i + \left(S_j|_{x_i=\Delta x} - S_j|_{x_i=-\Delta x} \right) \ln f_j}{8\Delta x} y_0 \quad (36)$$

Third-order partial derivatives of the form $\partial^3 y / \partial x_i^2 \partial x_j$ can be calculated by additional finite differencing, and an additional set of $2n + 1$ sensitivity coefficients can be used to obtain the remaining third-order derivatives. It was found that considering derivatives only up to second order was usually sufficient for most combustion

systems [195]. For instance, the uncertainties in Fig. 8 were calculated using this method.

2.4. Non-intrusive methods

The advantage of the intrusive methods is that uncertainties can be directly calculated. However, in many cases, the governing equations are not amenable to an intrusive reconstruction, as was concluded, for example, by Najm and coworkers for the chemical source term [164,166,192]. In principle, the model prediction's Taylor series could be calculated by finite difference, but this is cumbersome. The alternatives to finite-difference calculation of the Taylor series can be grouped into two general classes: quadrature methods and random sampling methods. These methods are essentially variations on statistical regression techniques.

Frenklach and coworkers developed the solution mapping technique, using central-composite factorial designs to calculate the Taylor series [17,159]. McRae and coworkers developed the Deterministic Equivalent Modeling Method (DEMM), which uses a quadrature method to evaluate the PCE coefficients of an uncertain model output [163]. In addition to their intrusive and pseudo-spectral methods, Najm and coworkers developed the Non-Intrusive Spectral Projection (NISP), which uses random sampling to calculate the PCE coefficients in Eq. (14). Rabitz and coworkers developed the Random-Sample High Dimensional Model Representation (RS-HDMR) [186,196], using random sampling to calculate the HDMR coefficients in Eq. (22) [185,186]. RS-HDMR was applied by Tomlin [134,197–199] as well as Chelliah [200,201] and coworkers to a range of uncertainty analysis problems in combustion.

2.4.1. Solution mapping

Solution mapping by Frenklach and coworkers [17,202] was the earliest effort to apply response surface methodologies to combustion model development. As a component of this method, a central-composite (CC) design is generated within the reaction rate uncertainty space and the derivatives are calculated using standard linear regression. A CC design consists of an axial design and a factorial design, which can be fractional. For an n -active parameter problem, the axial design evaluates the model at the $2n + 1$ points in the reaction uncertainty space: $x_i = 0$ ($i = 1, \dots, n$) and then along each major axis $x_i = \pm \Delta x$ with $x_{j \neq i} = 0$. The perturbation $\Delta x = 2^{n/4}$ is chosen to ensure that the design is rotatable, i.e., that the information taken from the model evaluations does not depend on the orientation of the design in model parameter space [203].

The factorial design consists of evaluating the model at all corner points in the parameter space such that $x_i = \pm 1$. There are 2^n such possible combinations, and so a design such as this is referred to as a 2^n design. For a large number of parameters, the number of model evaluations becomes quite large, and so it may be economical to use a fractional factorial design of 2^{n-p} points, with p an arbitrary reduction factor. Such designs are termed 2^{n-p} designs. The resolution of the design is a critical parameter to estimate its performance; if a fractional factorial design has resolution $r = v + w$, then estimates of derivatives of order v contain effects from interactions of order no lower than w . The sparsity of effects principle states that interactions of third order and higher should be negligible, and so to calculate second-order derivatives requires a design of resolution V or higher. There is much literature on developing fractional factorial designs [204], which need not be repeated here, but it can be shown that the size of a resolution V design grows as n^2 , due to the need to calculate $n(n-1)/2$ second-order derivatives. The reduction factor p must be sufficiently small for the design to contain the required number of model evaluations.

2.4.2. Deterministic equivalent modeling method

The Deterministic Equivalent Modeling Method was developed and used by McRae and coworkers [163,205–207]. This method is probabilistic and relies on being able to calculate the expectations in Eq. (14). For the case of a Wiener-Hermite chaos, the integral is evaluated using Gauss-Hermite quadrature. In this method, the dimension $N_d = n$ and order N_0 of the PC expansions are selected.

The number of coefficients in a PC expansion grows approximately as $N_0^{N_d}$, and the orthogonal Wiener-Askey polynomials have N_0 distinct roots, so a grid of points made from possible combinations of these roots has $N_0^{N_d}$ members. The expectation is evaluated according to

$$E[y\Psi_k] = \sum_{i=1}^M w_i y(\xi_i) \Psi_k(\xi_i) \quad (37)$$

where $y(\xi_i)$ is the realization of the uncertain prediction y at the i^{th} quadrature point, of which there are M , and $\Psi_k(\xi_i)$ is the k^{th} multidimensional Hermite polynomial evaluated at the i^{th} quadrature point.

An example of the output from DEMM is shown in Fig. 10, taken from Refs. [163], which plots the H_2 concentration with respect to time in supercritical water oxidation. The initial concentrations are fixed, so there is no uncertainty. Over time, the uncertainty in the species concentrations grows due to uncertainty in the rate constants and in the heats of formation of key species; it eventually shrinks to 0 because the reaction will eventually proceed to completion. Fig. 10 also underscores the utility of having an uncertainty structure such as the PCE or Taylor series, because then the uncertainty in the solution due to a particular model parameter can be easily extracted. For instance, the uncertainty due to a single rate constant ($H_2O_2 \rightarrow 2OH$) and to a key heat of formation (HO_2) is compared to the total uncertainty in the H_2 concentration. These two parameters account for the bulk of the total uncertainty.

2.4.3. Non-intrusive spectral projection

The Non-intrusive Spectral Projection (NISP) method [164,166] was developed to address accuracy concerns in their intrusive methods. This method is similar to the DEMM except that it uses Monte Carlo integration with Latin Hypercube sampling (LHS) instead of Gauss-Hermite quadrature. In this case, the expectations are given by

$$E[y\Psi_k] = \frac{1}{M} \sum_{i=1}^M y(\xi_i) \Psi_k(\xi_i) \quad (38)$$

where, similar to DEMM, $y(\xi_i)$ is the realization of the uncertain y at the i^{th} point in the LHS sequence, which contains M points, and $\Psi_k(\xi_i)$ is the k^{th} Hermite polynomial. An example of the solution

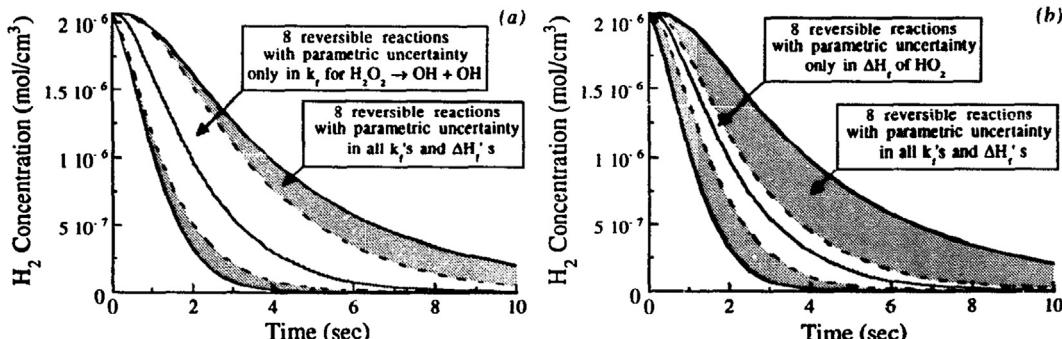


Fig. 10. H_2 concentration and uncertainty (97.5% and 2.5% contour) predicted for supercritical water oxidation with respect to time. The conditions are $2.06 \mu\text{mol}/\text{cm}^3 H_2$, $1.04 \mu\text{mol}/\text{cm}^3 O_2$ and $4.28 \text{ mmol}/\text{cm}^3 H_2O$ at 246 bar and 823 K. The plot is adapted from Phenix et al. [163].

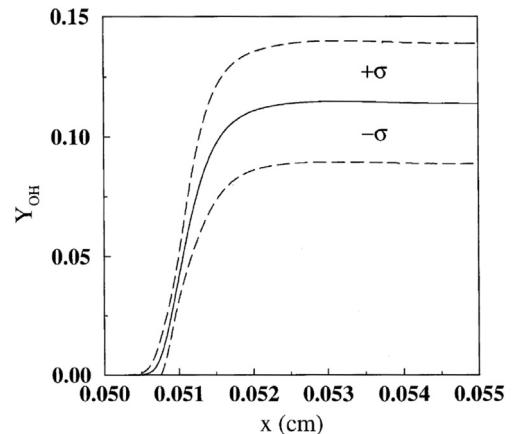


Fig. 11. OH mass fraction and uncertainty in supercritical water oxidation with respect to spatial position. The conditions are the same as Fig. 10. In this case, the system is treated as a one-dimensional premixed flame rather than a homogeneous reactor. The plot is adapted from Reagan et al. [164].

using NISP is shown in Fig. 11, adapted from Ref. [164], which shows the OH mass fraction as a function of position in supercritical water oxidation, under the same conditions as [163]. The uncertainty increases quickly in the reaction zone, then converges to a constant value due to the uncertainty in the heat of formation of a key species.

2.4.4. Random-sample high-dimensional model representation

Random-Sample High-Dimensional Model Representation (RS-HDMR) was developed by Rabitz and coworkers [184–186,196] and implemented in combustion simulations by Tomlin [134,197], Davis [199,208], Chelliah [201] and their respective coworkers. Like Solution Mapping, DEMM, and NISP, it uses black-box model evaluations to calculate the integrals necessary for the uncertainty calculations; the RS-HDMR implementation of Tomlin and coworkers uses Sobol' sampling [209] rather than LHS sampling. In this case, the integral in Eq. (22) is solved by

$$\int_{K^n} \varphi_k(x_i) y(\mathbf{x}) d\mathbf{x} = \frac{1}{M} \sum_{j=1}^M \varphi_k(x_{i,j}) y(\mathbf{x}_j) \quad (39)$$

where, similar to DEMM and NISP, $y(\mathbf{x}_j)$ is the realization of uncertain y at the j^{th} point in the Sobol' sequence, which are consisted of M points, and $\varphi_k(x_{i,j})$ is the k^{th} one-dimensional orthogonal function, evaluated at the j^{th} point in the Sobol' sequence.

An example of an RS-HDMR calculation is shown in Fig. 12 [197], which shows the downstream mole fraction of NO predicted for a

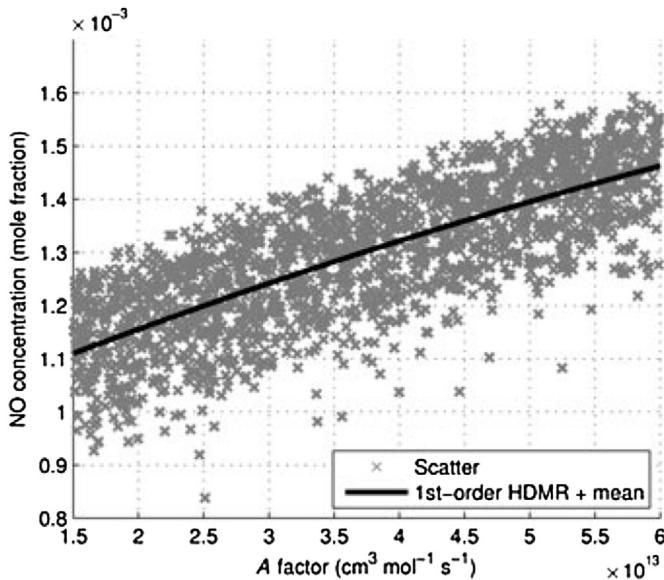


Fig. 12. Simulated NO mole fraction in a methane/air flame at equivalence ratio 1.6, pressure 0.053 atm and unburned temperature 298 K. The flame is doped with 0.5% SO_2 and 1.3% NH_3 . The line shows the first-order contribution to NO mole fraction of the A factor for $\text{SO} + \text{NH} \rightarrow \text{NO} + \text{SH}$. The scatter shows the remaining uncertainty due to other uncertain rate parameters in the model. The plot is adapted from Ziehn and Tomlin [197].

rich CH_4 /air flame doped with SO_2 and NH_3 as a function of the Arrhenius A factor for $\text{SO} + \text{NH} \rightarrow \text{NO} + \text{SH}$. The flame is simulated using the Leeds model [210], updated with the recommendations of Baulch et al. [118] and augmented with SO_x chemistry as described in Ref. [134]. The variation in the NO concentration is determined using Monte Carlo sampling, as shown by the scatter of points in the figure. The Monte Carlo sample was then used as an input to Eq. (39) to calculate the first order contribution of the A factor to the NO concentration, that is, the effect due solely to the reaction $\text{SO} + \text{NH} \rightarrow \text{NO} + \text{SH}$. Note that the first-order effect is nonlinear, because “order” in HDMR refers to the number of parameters involved in the interaction, rather than the exponent of a particular power. The remaining variation in the NO concentration is due to the effects of other rate parameters.

2.4.5. Uncertainty estimation through Monte Carlo integration

In some cases, the gradient \mathbf{g} and Hessian \mathbf{H} might not be straightforward to calculate or might be computationally costly, in which case σ_y can still be determined using Monte Carlo methods [15,102,178,211–215]. Examples of this are shown in Fig. 13, adapted from Ref. [214] and Fig. 14, adapted from Ref. [102]. The left panel of Fig. 13 shows the time histories of C_2H_4 , OH , H_2O and CO_2 during shock-initiated ignition of heptane, which was modeled using the unconstrained JetSurF version 1.0 [224]. The uncertainty is visualized by a random sampling of 1000 points in the model uncertainty space and random reaction time. The reaction time is sampled log-uniformly between 10 and 10 000 μs . Fig. 14 shows the laminar flame speed of H_2 /air mixtures at 1 atm and 298 K, modeled using the Leeds model [210]. The uncertainty is determined by probabilistically weighted Latin Hypercube Sampling within the model uncertainty space.

2.5. Probability distribution of rate uncertainty versus model uncertainty

It is important to understand the difference in results between uniformly and normally distributed rate coefficients ($\log k$). When

it comes to a treatment of the uncertainty distribution of the rate coefficient, they probably represent the two limiting cases. Fig. 15 shows the probability distributions of the flame speed computed for an ethylene–air mixture at the equivalence ratio $\phi = 0.7$, assuming either normally-distributed or uniformly-distributed reaction rate parameters in USC Mech II [16]. The probability distributions of the flame speeds calculated from the two limiting distributions of x are similar. The lack of sensitivity to the assumed probability distribution can be explained by the central limit theorem as discussed in Ref. [178]. This observation is in fact quite essential to the accuracy of forward model uncertainty projection.

3. Uncertainty minimization

3.1. Limitations of forward uncertainty propagation

Much of our discussion, and indeed most of the UQ work, so far has focused on forward uncertainty propagation, where the uncertainty in the rate parameters is prescribed and, in principle, mathematically independent. Forward UQ projection is useful in that the accuracy of a chemical model may be assessed as needed as new and more accurate knowledge from rate measurement or theoretical analysis is made available. Following such an approach, the accuracy of a chemical model can be improved, in principle, as rate parameters are probed with improved accuracy and precision, through UQ-initiated, targeted studies. This is the approach advanced by, for instance, Klippenstein, Tomlin and coworkers [107,130,198,225]. In these studies, screening designs are used to identify reaction steps with the greatest influence on uncertainty, and then those reactions are targeted for investigation in greater detail.

On the other hand, the need to develop combustion models of a wide range of complex fuels (and the rate at which these models are published) has left us with large numbers of rate parameters with large uncertainty. As an example, the uncertainties in the modeled laminar flame speeds shown in the left panel of Fig. 8 are considerably larger than the uncertainties in the combustion experimental measurement. Hence, the measurements contain useful, quantitative information that can provide constraints to the model parameters. Future progress is likely to be achieved more efficiently through efforts from both ends of the spectrum of combustion research: both by “chasing” the accuracy and precision of individual rate coefficients and by pursuing an inverse problem approach whereby the uncertain rate parameters are constrained by combustion experiments to yield a joint probability distribution or reduced sizes of the feasible parameter set. The view expressed above is admittedly controversial and the research community has not reached a consensus about how to engage in an efficient integration of the different approaches and methods. Nonetheless, we shall discuss in what follows the mathematical treatment of the inverse problem that considers the coupled uncertainties in the rate parameters and combustion experiments.

3.2. The inverse problem

Because of the inherent limitations with direct estimates of model rate parameters, there will always be a place for constraining models against measurements, a technique mathematically termed as the inverse problem. In essence, the models are “tuned” against the experimental measurements, adjusting the model rate parameters within their uncertainty bounds until the model can reproduce the observations. In principle, at least, solving the inverse problem results in a model with a set of rate parameters that are closer to reality and have tighter uncertainties. A pitfall, however, is that there will never be enough experimental information

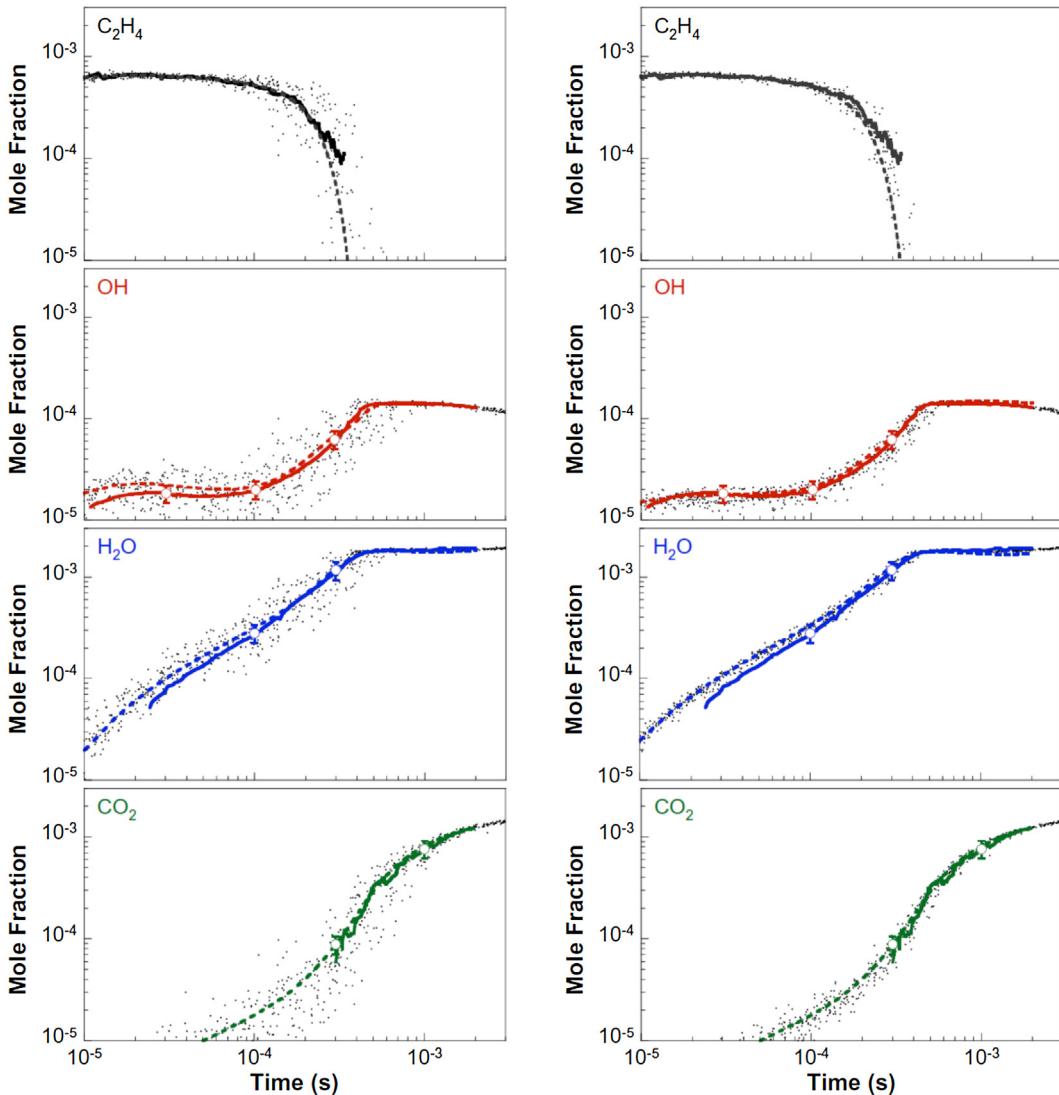


Fig. 13. Mole fraction of selected species during ignition of a n-C₇H₁₆ mixture with respect to time behind a reflected shock wave. The initial post-shock conditions are 0.03% n-C₇H₁₆ and 0.3% O₂ in Ar, with $T = 1494$ K and $p = 2.115$ atm. The solid lines are the experimental measurement [216] presented for comparison, and the open circles with error bars show the uncertainty estimate. Left panel: unconstrained JetSurF 1.0 [224]; right panel: constrained JetSurF 1.0. The dashed lines are the nominal predictions of the models, and the model uncertainty is shown by the small points. The plot is adapted from Sheen and Wang [214].

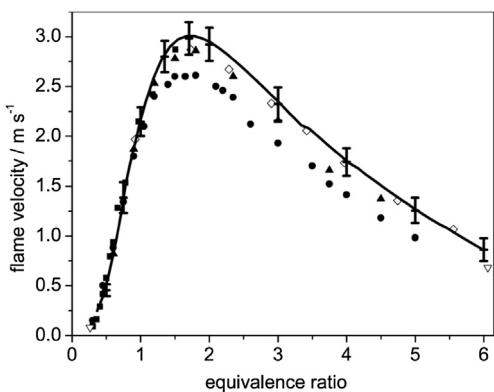


Fig. 14. Laminar flame speed of H₂/air mixtures with respect to equivalence ratio. The solid line is the prediction of the Leeds model [210], and the error bars show the calculated 1σ uncertainty in the model. Symbols are experimental measurements (■ [217,218], ◊ [219], ● [220,221], ▲ [222], △ [223]) presented for comparison. The plot is adapted from Zsély et al. [102].

to fully invert the problem in the realm of combustion chemistry. Historically, the measurements chosen as combustion targets and parameter adjustments against these targets have been treated on an ad-hoc basis. The ad-hoc procedure ensures a set of model parameters that will reproduce the initial set of measurements, but without an uncertainty estimate it is impossible to reliably interpolate or extrapolate outside this set. The proliferation of chemical kinetic models that purport to describe the same chemical systems is a manifestation of the problem, as we discussed earlier.

As an inverse problem, one would wonder whether it is sufficient or even possible to carry out a set of well-controlled measurements in a way that the rate parameters can be “pinned” to a unique set of values. The answer is clearly no. To understand this problem, we show in Fig. 16 the cross sections of the solution surface for a large family of rate parameters that can give the 40 cm/s consensus value of the early 1990s for the atmospheric laminar flame speed of the stoichiometric methane-air mixture. As can be seen, a known experimental quantity reduces the initial dimensionality of the uncertain rate parameter space from a cube to a surface. Hypothetically, one would just need to have at a minimum

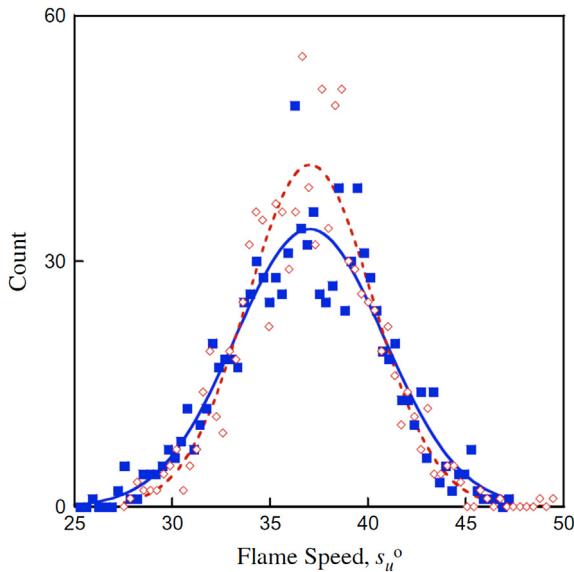


Fig. 15. Comparison of Monte Carlo sampling of the ethylene–air laminar flame speed at equivalence ratio of 0.7 for normally-distributed rate parameters (\diamond , dashed line) and uniformly-distributed rate parameters (\blacksquare , solid line). The computation was carried out using USC Mech II [16]. The plot is adapted from Sheen and Wang [178].

five independent flame speed values, say, over a range of equivalence ratio or pressure, to invert the five-parameter system illustrated. In reality, this is not a feasible approach for at least two reasons. First, most combustion experiments yield observations that are dependent on each other, meaning that the five hypothetical combustion experiments considered do not actually provide a unique solution within the five-dimensional uncertainty space. Second, as the number of thermodynamic and stoichiometric conditions under consideration increases, the number of influential rate parameters increases faster, so that there are more unknown parameters being solved for than known constraints and therefore the solution is non-unique. Therefore, the large parameter space and the difficulties of obtaining independent experimental observation render the problem ill-defined mathematically. Hence, an inverse problem with UQ treatment alone cannot achieve predictive combustion chemistry modeling, but it can produce impactful results, as we demonstrate in Figs. 8 and 13—the fundamental combustion experiments provide constraints on and give a joint probability distribution for the rate parameters within their respective uncertainty bounds.

As an example, we show a plot (Fig. 17) adapted from the work of Miller and Frenklach [35]. In their work, the rate coefficients of two reactions that are most influential to the overall rate of the thermal decomposition of propane at a temperature of 1100 K and molar density of 6×10^{-5} mol/cm³ were treated by an inverse problem against experimental target sets defined by the concentrations measured for five decomposition products at 0.7 ms from the onset of reaction. The two reactions are the C–C fission in propane ($C_3H_8 \rightarrow CH_3 + C_2H_5, k_1$) and H-abstraction of propane by the methyl radical ($C_3H_8 + CH_3 \rightarrow CH_4 + i-C_3H_7, k_2$). In Fig. 17, the various contours represent the 95% confidence region of the (k_1, k_2) combination that can reproduce the different sets of data. Several points can be made by an inspection of the plot. First, all of the viable solutions are characterized by (k_1, k_2) contours of finite size. This is a clear illustration of the difficulties associated with rate coefficient determination under clean, chemically isolated condition. Often, an experiment can only determine coupled rate coefficients or a confidence region in the parameter space.

Under this condition, the choices of the target sets impact the size and shape of the confidence region, as illustrated also in Fig. 17: the different target sets can give very different covariance of the rate parameters. Lastly, a larger target data set does not always lead to a reduced confidence region, because of uncertainties inherent to the experimental data as well as possible structural uncertainties.

Many groups have turned to Bayesian methods to address the problem of uncertainty analysis and rate parameter constraint. Inverse problems in chemical kinetics typically involve models with a given number of uncertain parameters, along with a set of experimental measurements whose simulations each depend on a set of model parameters. Bayesian methods are highly adapted to solving problems of this nature.

3.2.1. Overview of Bayesian methods

A Bayesian study consists of a set of uncertain model parameters \mathbf{x} , defined on some accessible uncertainty space $K^n \subseteq \mathbb{R}^n$ and assigned a PDF $p(\mathbf{x})$, the prior PDF. A training set of measurements is selected, with some measurement data \mathbf{y}^{obs} and associated uncertainty σ^{obs} . A likelihood function $\pi(\mathbf{x}; \mathbf{y}^{obs}, \sigma^{obs})$ can be defined, written as a function of \mathbf{x} and parametrically dependent on \mathbf{y}^{obs} and σ^{obs} . This function describes how likely a particular choice of model parameters is depending on how well it reproduces the experimental observations, and represents the new information coming from those observations. The prior $p(\mathbf{x})$ and likelihood $\pi(\mathbf{x}; \mathbf{y}^{obs}, \sigma^{obs})$ can be combined, using Bayes' Theorem, to determine an updated PDF $p^*(\mathbf{x} | \mathbf{y}^{obs})$, which is the posterior PDF, given by,

$$p^*(\mathbf{x} | \mathbf{y}^{obs}) = \frac{p(\mathbf{x})\pi(\mathbf{x}; \mathbf{y}^{obs})}{\int_{K^n} p(\mathbf{x})\pi(\mathbf{x}; \mathbf{y}^{obs}) d\mathbf{x}} \quad (40)$$

The denominator in Eq. (40) is termed the evidence, which ensures that p^* integrates to 1. Naturally, p^* depends strongly on the form of p and π ; in particular, π might be an implicit function of \mathbf{x} , which makes Eq. (40) difficult to evaluate. Furthermore, computing the response requires performing a high-dimensional integral, especially in the case of chemical models.

Methods for solving Eq. (40) fall into three general classes. The original work on Bayesian model updated by Beck and co-workers [226–232] used Monte Carlo methods to evaluate Eq. (40), work which was followed by Kraft [233–242], Raman [168], Miki [243,244] and their respective coworkers. Sheen and Wang proposed the Method of Uncertainty Minimization using Polynomial Chaos Expansions (MUM-PCE) [15,178,214,215] to calculate credible regions in rate parameter space and the corresponding uncertainty in model predictions, which involves assuming a particular form for the posterior distribution. Turányi and coworkers developed a similar statistical method [245–248], and a similar method was applied by Burke and co-workers [225] and Najm and co-workers [249]. Frenklach and coworkers proposed the Data Collaboration method [132,250–253], which uses robust control to calculate uncertainty bounds on model parameters and predictions based on constraints from experimental observations.

3.2.2. Bayesian analysis

A Bayesian inference study requires making assumptions for the form of p and π . A typical assumption is that the experimental observations and model parameters are jointly normally distributed. This gives the following form for the prior $p(\mathbf{x})$,

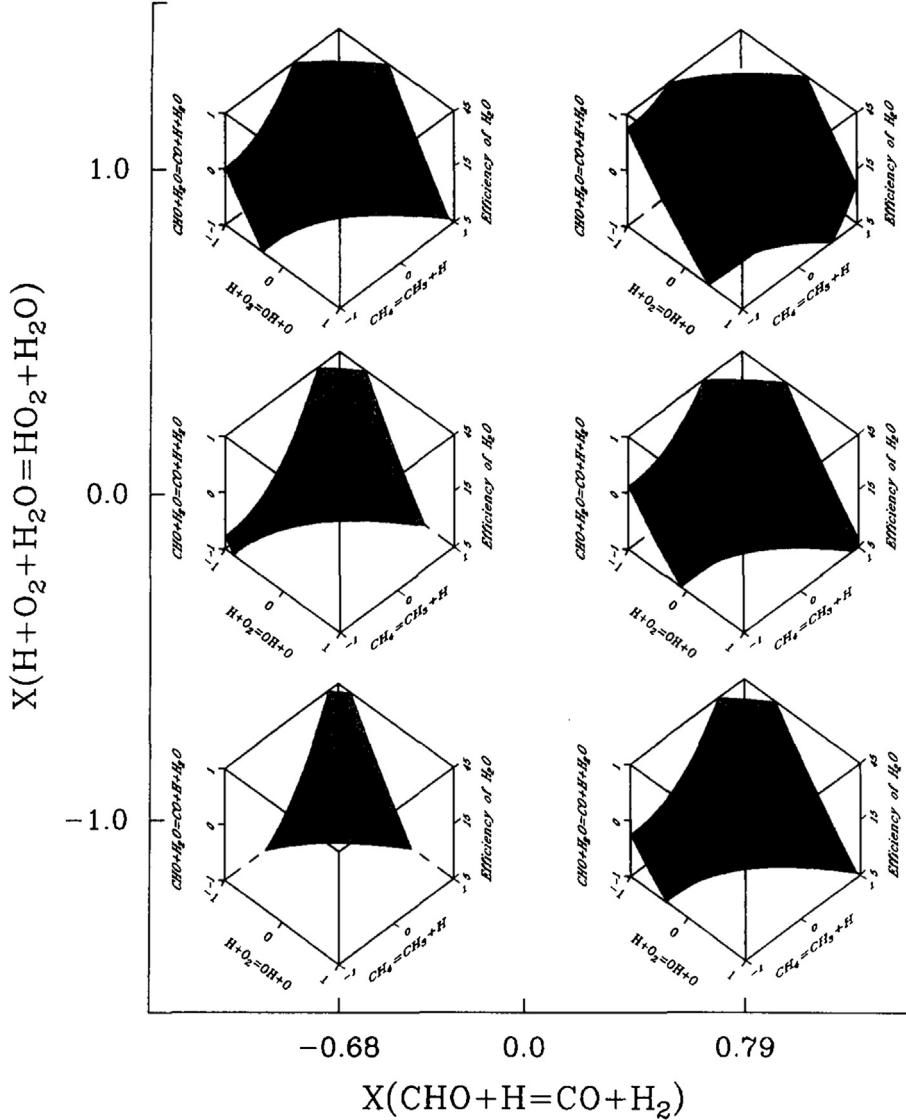


Fig. 16. Cross section of the solution surface. The axes are variables for the normalized rate coefficients x of the reactions indicated. The x values spans from the respective uncertainties of the rate coefficient at or around 1990. The shaded area is the locus of stoichiometric methane-air ($p = 1$ atm) flame velocities equal 40 cm/s. The plots are adapted from Frenklach et al. [2].

$$p(\mathbf{x}) = [(2\pi)^n |\Sigma_{\mathbf{x}}|]^{-1/2} \exp\left(\frac{-\mathbf{x}^T \Sigma_{\mathbf{x}}^{-1} \mathbf{x}}{2}\right) \quad (41)$$

and for the likelihood $\pi(\mathbf{x}; \mathbf{y}^{\text{obs}})$,

$$\pi(\mathbf{x}; \mathbf{y}^{\text{obs}}) = [(2\pi)^{N_e} |\Sigma_{\mathbf{y}}|]^{-1/2} \exp\left\{ -\frac{[\mathbf{y}(\mathbf{x}) - \mathbf{y}^{\text{obs}}]^T \Sigma_{\mathbf{y}}^{-1} [\mathbf{y}(\mathbf{x}) - \mathbf{y}^{\text{obs}}]}{2} \right\} \quad (42)$$

where $\Sigma_{\mathbf{x}}$ is the prior covariance matrix of the model parameters and $\Sigma_{\mathbf{y}}$ is the experimental measurement covariance matrix. As discussed in Section 2.1, it is often assumed that the model parameters and experimental measurements are all independent, so that both $\Sigma_{\mathbf{x}}$ and $\Sigma_{\mathbf{y}}$ are diagonal. Furthermore, if the model

parameters \mathbf{x} are normalized as in Eq. (32), then $\Sigma_{\mathbf{x}}$ is equal to $\Theta^2 \mathbf{I}$, where \mathbf{I} is the identity matrix and Θ is 1 or $1/\sqrt{2}$, depending on whether the uncertainty factors \mathbf{f} represent a 1σ or 2σ uncertainty. Likewise, each experiment is assigned its own measurement uncertainty σ_i^{obs} . The numerator of Eq. (40) can then be rewritten as

$$p(\mathbf{x})\pi(\mathbf{x}; \mathbf{y}^{\text{obs}}) \propto \exp\left\{ -\frac{1}{2} \left[\frac{\mathbf{x}^T \mathbf{x}}{\Theta^2} + \sum_{i=1}^{N_e} \left(\frac{y_i(\mathbf{x}) - y_i^{\text{obs}}}{\sigma_i^{\text{obs}}} \right)^2 \right] \right\} \quad (43)$$

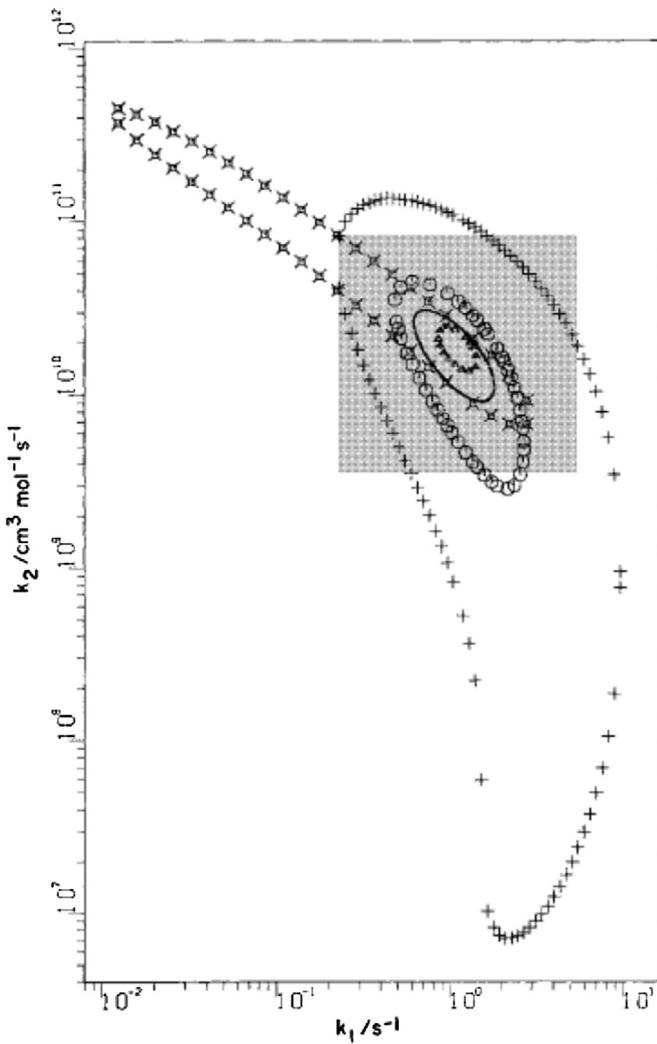


Fig. 17. Variations of the 95% confidence contours of the reaction rate constants of $\text{C}_3\text{H}_8 \rightarrow \text{CH}_3 + \text{C}_2\text{H}_5$ (k_1) and $\text{CH}_3 + \text{C}_3\text{H}_8 \rightarrow \text{CH}_4 + i\text{-C}_2\text{H}_7$ (k_2), determined from an analysis of single-pulse shock tube pyrolysis data of propane (1.6% vol) in argon at 1500 K temperature and 6×10^{-5} mol/cm³ molar density using a 9-step reaction model. The six contours were obtained from considering different combinations of the product concentration values measured at 0.7 ms after the arrival of reflected shock wave, including (1) CH_4 , (2) C_2H_4 , (3) C_2H_6 , (4) C_3H_6 and (5) H_2 . Δ : 1,3,4; +: 2,3,4; x: 1,2,4,5; o: 2,3,4,5; solid line: all five data values. The shaded area represents the initial guesses of the range of the two rate constants at 1500 K. The plot is adapted from Miller and Frenklach [35].

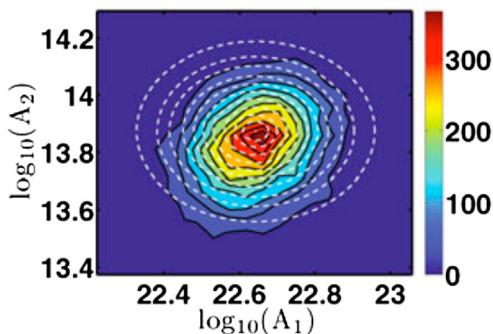


Fig. 18. Joint PDF for the Arrhenius A factors of the reactions $\text{H} + \text{OH} (+\text{M}) \rightarrow \text{HO}_2 (+\text{M})$ (A_1) and $\text{HO}_2 + \text{H} \rightarrow 2\text{OH}$ (A_2). The posterior probability density is indicated by the color bar. The dashed lines are level curves of the prior PDF. The plot is adapted from Braman et al. [168]. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3.2.3. Bayesian analysis using Markov chain Monte Carlo

A straightforward method of solving Eq. (40) is by using Monte Carlo integration. This is the approach proposed by Beck and Katafygiotis [226,227] and employed by Beck [228–232], Kraft [233–241], Raman [168], Miki [243,244] and their coworkers. Many of these studies used Markov-Chain Monte Carlo (MCMC) sampling, specifically the Metropolis–Hastings algorithm [254] to evaluate the posterior. In this method, a random point \mathbf{x}_0 is chosen within K^n , and the function $f(\mathbf{x}_0) = p(\mathbf{x}_0)\pi(\mathbf{x}_0; \mathbf{y}^{\text{obs}})$ is evaluated. Then, a new point \mathbf{x}' is proposed, and $f(\mathbf{x}')$ evaluated. An acceptance probability is defined as $p^{\text{acc}} = \min[1, f(\mathbf{x}')/f(\mathbf{x}_0)]$. The point \mathbf{x}' is accepted with this probability, in which case $\mathbf{x}_1 = \mathbf{x}'$, otherwise it is rejected and $\mathbf{x}_1 = \mathbf{x}_0$. A new \mathbf{x}' is proposed and compared to \mathbf{x}_1 , and an \mathbf{x}_2 is selected. This process continues until N points have been selected. For sufficiently large N , the set of samples \mathbf{x} is equivalent to sampling from the posterior probability p^* . This sample can be used to calculate the mean and variance of any parameter. Furthermore, if there is some model output y , then the expectation $E[y]$ can be calculated using,

$$E[y] = \frac{1}{N} \sum_{i=1}^N y(\mathbf{x}_i) \quad (44)$$

An example of the output from this method is shown in Fig. 18, which compares the prior and posterior joint PDF for two Arrhenius A factors, adapted from Ref. [168]. Because the posterior PDF is being determined based on Monte Carlo sampling, it must be approximated by a two-dimensional histogram. Therefore, what is being plotted in Fig. 18 is the local sample density; for any point, the local sample density counts about how many points in the final set of samples are in some sense close to this point.

3.2.4. Uncertainty minimization through polynomial chaos expansion

Monte Carlo integration has the advantage that it can be used for virtually any choice of p and π . For some well-behaved choices, it can be more efficient to find an explicit form for the posterior p^* , as was done in Refs. [15,178,214,215,245,246,248]. In general, $p^*(\mathbf{x}|\mathbf{y}^{\text{obs}})$ can equivalently be described by a PC expansion for the random model parameter vector $\mathbf{X}^* = \mathbf{X}|\mathbf{y}^{\text{obs}}$ [178],

$$\mathbf{X}^* = \mathbf{x}^{(0)*} + \mathbf{x}^{(1)*}\xi + \mathbf{x}^{(2)*}(\xi) + \mathbf{x}^{(3)*}(\xi) + \text{h.o.t.} \quad (45)$$

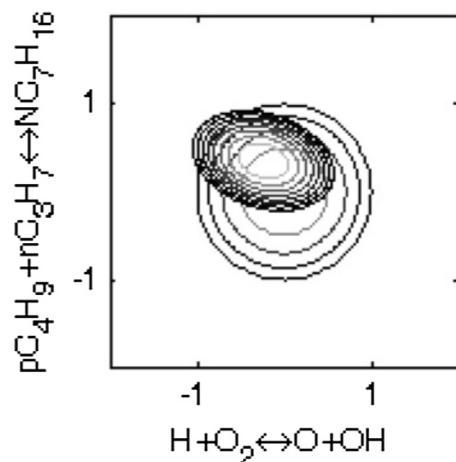


Fig. 19. Joint PDF for the Arrhenius A factors of the reactions $\text{H} + \text{OH} \leftrightarrow \text{O} + \text{OH}$ and $p\text{-C}_4\text{H}_9 + n\text{C}_3\text{H}_7 \leftrightarrow n\text{-C}_7\text{H}_{16}$, adapted from Ref. [214]. The A factors have been normalized using Eq. (32). The plot is adapted from Sheen and Wang [214].

where $\mathbf{x}^{(0)*}$ and $\mathbf{x}^{(1)*}$ were defined in Section 2.2.1 and $\mathbf{x}^{(2)*}$ and $\mathbf{x}^{(3)*}$ represent the 2nd and 3rd order terms of the expansion. If the expansion is truncated at the first order, then $\mathbf{X}^* = \mathbf{x}^{(0)*} + \mathbf{x}^{(1)*}\xi$, which is the form presented in Section 2.2.1. If the basis random variables are normally distributed, then \mathbf{X}^* follows a multivariate normal distribution with mean $\mathbf{x}^{(0)*}$ and covariance matrix $\Sigma^* = \mathbf{x}^{(1)*}\mathbf{x}^{(1)*T}$. $\mathbf{x}^{(0)*}$ can be determined by calculating the maximum probability density in Eq. (43), or equivalently

$$\mathbf{x}^{(0)*} = \operatorname{argmin}_{\mathbf{x}} \left[\frac{\mathbf{x}^T \mathbf{x}}{\Theta^2} + \sum_{i=1}^{N_e} \left(\frac{y_i(\mathbf{x}) - y_i^{\text{obs}}}{\sigma_i^{\text{obs}}} \right)^2 \right] \quad (46)$$

The model predictions can be linearized about $\mathbf{x}^{(0)*}$, $y_i^*(\mathbf{x}) = y_i(\mathbf{x}^{(0)*}) + \mathbf{J}_i^{*T}(\mathbf{x} - \mathbf{x}^{(0)*})$, where $\mathbf{J}_r^* = 2\mathbf{H}_r \mathbf{x}^{(0)*} + \mathbf{g}_r$. This expression for $y_i(\mathbf{x})$ can be substituted into Eq. (43) and rearranged to solve for Σ^* ,

$$\Sigma^* = \left[\frac{\mathbf{I}}{\Theta^2} + \sum_{i=1}^{N_e} \frac{\mathbf{J}_i^* \mathbf{J}_r^{*T}}{\left(\sigma_i^{\text{obs}} \right)^2} \right]^{-1} \quad (47)$$

An example of the output from this method is shown in Fig. 19, which compares the prior and posterior joint PDF for two Arrhenius A factors, adapted from Ref. [214]. Because there is an explicit probability density function, the contours of the PDF are smooth lines.

As was discussed in the introduction, the measure of a model is how useful it is at making predictions. For instance, the prior prediction uncertainty shown in the left panels of Figs. 8 and 13 is large with respect to the experimental measurement uncertainty, and so the prior model makes predictions less precise than the experiment. The posterior parameter uncertainties from Eq. (47), however, can be propagated into Eq. (10), which yields the posterior prediction uncertainties shown in the right panels of Figs. 8 and 13. In these cases, the prediction uncertainties are substantially reduced from the prior, indicating the improved ability of the model to make more precise predictions. Likewise, the rate constant estimate for the reaction $\text{H} + \text{O}_2 + \text{N}_2 \leftrightarrow \text{HO}_2 + \text{N}_2$ is shown in

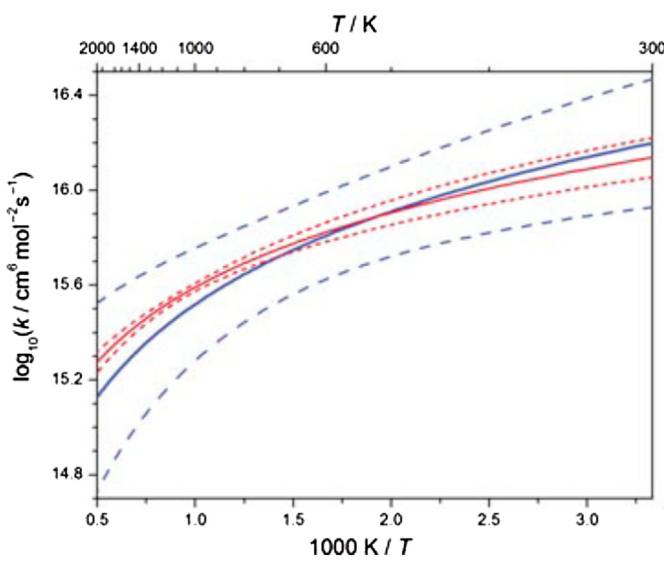


Fig. 20. Rate coefficient of the reaction $\text{H} + \text{O}_2 (+\text{N}_2) \leftrightarrow \text{HO}_2 + (\text{N}_2)$. The prior model estimate, taken from Baulch [118], is shown by the solid line. The uncertainty limits estimated by Turányi et al. [245] are shown by the long-dashed lines. The posterior estimate and uncertainty limits are shown by the solid lines bracketed by two short-dashed lines. The plot is adapted from Turányi et al. [245].

Fig. 20, adapted from Ref. [245]. The prior uncertainty is estimated by a review of literature estimates and is almost a factor of two under some conditions. This value is much larger than the uncertainty estimated by Baulch [118] and certainly very large for this key reaction. The posterior uncertainty is as little as a few percent.

3.2.5. Data collaboration

The Data Collaboration [132,250–253] method differs from other Bayesian methods in that it does not use an explicit probability density function to solve Eq. (40). Instead, it is assumed that the rate parameters and the experimental measurements follow an interval distribution. Each model parameter and measurement is characterized by a mean and an uncertainty. Then the accessible rate parameter space K^n is just the hypercube, $K^n: \{-1 < x_i < 1\}$. Likewise, there is an accessible experimental measurement space $D_e^N: \{y_i^{\text{obs}} - \sigma_i^{\text{obs}} < y_i(\mathbf{x}) < y_i^{\text{obs}} + \sigma_i^{\text{obs}}\}$, where N_e is the number of experimental measurements in the training set, $y_i(\mathbf{x})$ is the model prediction of the i th measurement, y_i^{obs} is the mean value of its measurement and σ_i^{obs} is its measurement uncertainty. The feasible region is then defined as $F^n = K^n \cap D_e^N$, with the superscript n noting that F^n is an n -dimensional subset of the accessible rate parameter space; F^n is then the parameter space consistent with both to the prior parameter estimates and with the training set. The method is still Bayesian, with K^n filling the role of the prior p and D_e^N filling the role of the likelihood π , and the intersection operator filling the role of multiplication. Once the feasible region F^n has been determined, the prediction interval R_i for any prediction y_i is given by $R_i^* = [L_i \ U_i]$, where

$$L_i = \min_{\mathbf{x} \in F^n} y_i(\mathbf{x}) \quad (48)$$

$$U_i = \max_{\mathbf{x} \in F^n} y_i(\mathbf{x}) \quad (49)$$

An example is shown in Fig. 21, which shows F^n for three reaction rate parameters constrained by two experimental measurements. The three reactions are the principal formation and destruction pathways for the CH radical, and the experimental measurements are of the maximum CH concentrations in methane/

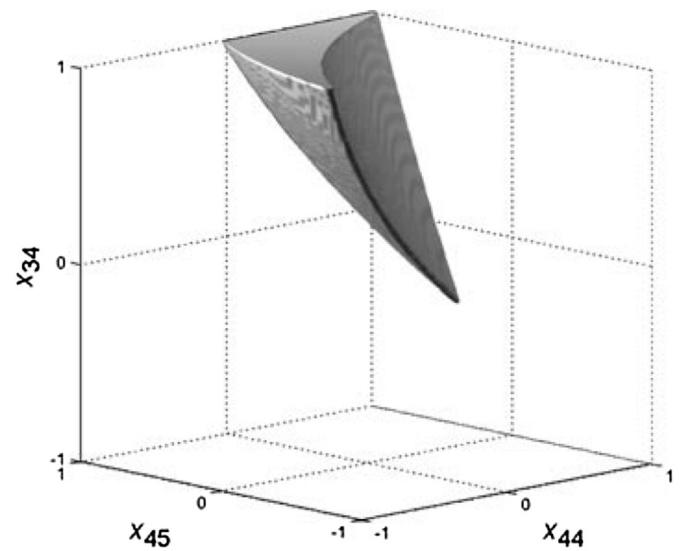


Fig. 21. Feasible region considering two experiments and three rate parameters. x_{44} is the Arrhenius pre-factor for the reaction $\text{CH} + \text{H}_2 \rightarrow \text{H} + \text{CH}_2$, x_{45} that for $\text{CH} + \text{H}_2\text{O} \rightarrow \text{H} + \text{CH}_2\text{O}$, and x_{34} that for $\text{OH} + \text{CH}_3 \rightarrow \text{CH}_2 + \text{H}_2\text{O}$. The experimental measurements are the maximum CH concentrations in methane/air laminar premixed flames. The plot is adapted from Frenklach et al. [132].

air laminar premixed flames. The point of the exercise in Ref. [132] was that, if any pair of rate coefficients is held fixed, it is often not possible to choose a value for the third, variable rate coefficient. The only reliable way to determine the best model is by a simultaneous, joint consideration of the rate parameter uncertainties.

3.2.6. Balance between mathematical and physical rigor

For this section, we briefly introduce two considerations in the solution of inverse problems, those of mathematical rigor and physical rigor. In an inverse problem physical rigor refers to exercises that consider, for example, all chemical pathways relevant to the experimental observations under consideration or the representation of the probability distribution of the rate parameters. If an important pathway is not included, then there will be too few degrees of freedom in the model and solutions of the inverse-problem is not expected to capture all measurements. Mathematical rigor refers to capturing all of the high-order dependence of the simulation on the model parameters. If too many assumptions are made, e.g., excessive truncation of high-order terms, the uncertainty in the rate parameters will be misestimated. In both cases, model parameters may be driven to unreasonable values.

In any inverse-problem study, there is an inherent trade-off between mathematical and physical rigor. For instance, the MCMC methods described in Section 3.2.3 are highly mathematically rigorous because they make no assumptions about how the model predictions depend on the rate parameters, but because of the computational cost they must sacrifice physical rigor and have been limited to a small number of active rate parameters. Conversely, MUM-PCE and similar methods [15,178,214,215, 245,246,248] assume that the model predictions are close to linear in the rate parameters, thus sacrificing mathematical rigor, but the relative smaller computational costs have allowed the study of systems with dozens of parameters [214].

A specific case of the tradeoff of physical rigor for mathematical rigor has been the treatment of temperature-dependent rate uncertainty. As discussed in Section 2.1, the three rate parameters each have their own uncertainty, which allows for a potentially elaborate description of the rate constant uncertainty with respect to temperature [172–174]. In principle, in an inverse-problem study, each of these three parameters would be subject to constraint, thereby constraining the temperature-dependent uncertainty. Because of the computational cost, essentially tripling the number of parameters that must be considered, previous studies considered uncertainty in the Arrhenius A term only. Only recent studies have begun to consider this kind of temperature-dependent uncertainties [215,245,246,248]. Similarly, the inverse-problem studies conducted so far have been restricted to high-temperature chemistry, due to the complexity of low-temperature oxidation chemistry that is more fuel specific than high-temperature chemistry.

3. Defining the experimental measurement uncertainty

3.3.1. Experimental uncertainty analysis

The solution to a Bayesian inference problem, such as that presented in Section 3.2, is highly sensitive to the choice of experimental uncertainty in Eq. (43). In the development of MUM-PCE [15,178], an uncertainty value is estimated for each measurement. In some cases, this value was provided by the experimentalist, and in others it might need to be estimated by statistical analysis. For instance, in the work of Sheen and Wang [178], the uncertainties in ethylene/air laminar flame speeds were set to be 2 cm/s based on consensus, whereas the uncertainties in ethylene/oxygen/argon ignition delay times were determined by a cross-conditional regression.

3.3.2. Regression error estimate

If the measurement uncertainty is specified, it is possible that some will not be consistent with the posterior model, that is, $y_i(\mathbf{x}^{(0)*}) - y_i^{\text{obs}} > 2\sigma_i^{\text{obs}}$. It is precisely these outliers that are interesting and the question is what should be done with them. MUM-PCE uses a consistency analysis scheme to remove these outliers from consideration. In the work of Turányi and coworkers [245,246,248], an additional error term is included in the model based on the discrepancy between the model predictions and the experimental measurements. A total covariance matrix Σ^{tot} is defined as,

$$\Sigma^{\text{tot}} = \Sigma^* + \Sigma^* \sum_{i=1}^{N_e} \left[\left(\frac{y_i(\mathbf{x}^{(0)*}) - y_i^{\text{obs}}}{\sigma_i^{\text{obs}}} \right)^2 \frac{\mathbf{J}_i^* \mathbf{J}_i^{*T}}{\left(\sigma_i^{\text{obs}} \right)^2} \right] \Sigma^* \quad (50)$$

This form for Σ^{tot} has the effect of increasing the model uncertainty in the direction of the outliers, such that the model prediction is more likely to agree with them.

3.3.3. Bayesian error inference

Another means of treating the uncertainty was presented in the work of Raman and coworkers [168]. In this work, the uncertainties are set equal to each other, i.e. $\sigma_i^{\text{obs}} = \sigma$. The likelihood function π is then expressed as

$$\pi(\mathbf{x}; \mathbf{y}^{\text{obs}}) \propto (2\pi\sigma)^{-N_e/2} \exp \left\{ -\frac{1}{2} \left[\sum_{i=1}^{N_e} \left(\frac{y_i(\mathbf{x}) - y_i^{\text{obs}}}{\sigma} \right)^2 \right] \right\} \quad (51)$$

The single uncertainty σ is then subject to optimization during the Bayesian inference process. Because σ is a free parameter, the model uncertainty will depend on $y_i(\mathbf{x}) - y_i^{\text{obs}}$; essentially, σ is an indication of goodness-of-fit. The effects of this change are shown in Fig. 22, taken from Refs. [168], which shows a comparison of a solution using the likelihood in Eq. (51) with one using Eq. (42). The figure shows the measured laminar flame speeds of H₂/CO/He mixtures from Burke et al. [106] along with posterior uncertainty estimates by Raman and coworkers [168]. As can be seen in the figure, in the case of a fixed experimental uncertainty, the model does not agree with some of the experimental measurements; the model cannot be reconciled with every measurement. If the uncertainty is an optimized parameter, the model can be brought into agreement with these measurements, at the cost of a significantly increased uncertainty in the rate coefficients. As was argued [168], this additional uncertainty represents the goodness-of-fit of the model. Expressing the experimental uncertainty in this way, however, has the disadvantage of enforcing homoscedasticity on the model, i.e., the uncertainty is the same for all experiments.

The Data Collaboration method uses a consistency analysis method [250] that is similar in concept to the Bayesian error methods (see, e.g., Ref. [168]). In Data Collaboration, the feasible region F^n will be empty if the model cannot be reconciled with all measurements. A threshold uncertainty C_\emptyset is defined such that F^n will be empty if all of the experimental uncertainties are reduced by C_\emptyset , i.e., the accessible data space is redefined as $D^{N_e} : \{y_i^{\text{obs}} - \sigma_i^{\text{obs}} + C_\emptyset < y_i(\mathbf{x}) < y_i^{\text{obs}} + \sigma_i^{\text{obs}} - C_\emptyset\}$. If C_\emptyset is negative, the experimental uncertainty must be increased until all measurements are consistent with the model.

3.4. Outlier detection and consistency analysis

The error terms used by Turányi et al. [245] and Raman and coworkers [168] could represent a more honest assessment of the

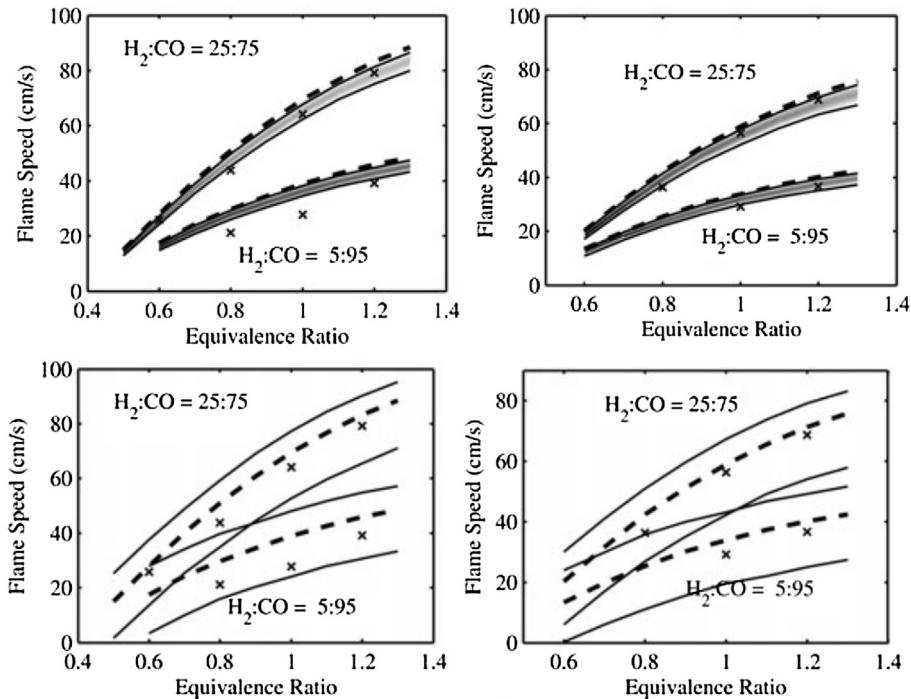


Fig. 22. Prior model predictions and posterior uncertainty estimates for H₂/CO/He flame speeds at 10 atm (left) and 20 atm (right). Symbols are the experimental measurements from Ref. [106]. Dashed lines are the prior model predictions of the Davis H₂/CO oxidation model [101]. Solid lines are the posterior uncertainty estimates by Raman and coworkers [168]. The top panel shows the uncertainty estimates using a fixed experimental uncertainty estimate, as in Eq. (41), and the bottom panel shows the uncertainty estimates using a variable experimental uncertainty estimate, as in Eq. (51). The plot is adapted from Braman et al. [168].

model uncertainty, as they are representative of goodness-of-fit. Taken too far, however, they can be interpreted as putting too much of the responsibility for the error on the model and not enough on the experiment. For instance, if one measurement value is particularly far from the posterior model prediction, that measurement will destroy any precision that might have been obtained from other measurements. In extreme cases, the posterior uncertainty will be greater than that of the prior model.

MUM-PCE does not consider these additional uncertainties; instead, it uses a consistency analysis and outlier detection algorithm. In this method, for every measurement that is inconsistent with the posterior model, a consistency criterion W is calculated as

$$W_i = S_i \left| \frac{y_i(\mathbf{x}^{(0)*}) - y_i^{\text{obs}}}{2\sigma_i^{\text{obs}}} \right|, \quad (52)$$

where S is a strength function indicating the relative importance of the measurement. Measurements are removed from consideration until all measurements are consistent.

Different choices of the strength function are possible, such as

$$S_i = \frac{\mathbf{J}_i^{*T} \mathbf{x}^{(0)*}}{\|\mathbf{J}_i^*\| \|\mathbf{x}^{(0)*}\|}, \quad (53)$$

The above equation derives from the principle that highly inconsistent measurements should drag the posterior model vector into close alignment with their gradient. An alternative strength function can also be defined as,

$$S_i = \mathbf{J}_i^{*T} \frac{\partial \mathbf{x}^{(0)*}}{\partial y_i^{\text{obs}}}, \quad (54)$$

derived on the principle that inconsistent measurements should have a large effect on the posterior parameter values. It can be

shown that this strength function is equal to $(\sigma_i^*)^2 / (\sigma_i^{\text{obs}})^2$, with σ_i^* calculated by substituting Σ^* into Eq. (10). Measurements that have a σ_i^* much less than σ_i^{obs} will be given a lower priority by the consistency analysis algorithm.

Both approaches to outliers are fraught with peril. As argued by Raman and coworkers [168], trusting the experimental uncertainty estimate and allowing the outliers to lie out results in a very precise determination of the rate parameters. In point of fact, because of the need to adjust the uncertainty to account for the outliers, much of the information contained in the experimental measurements is lost; the posterior uncertainty of the model is not much better than the prior, as shown in Fig. 22 [168]. If the experimental measurements are actually in error, then it makes sense not to consider them further [15,178,214]. However, if the disagreement between the measurement and the model is actually due to a structural problem with the model, such as a missing reaction pathway or an underestimated rate coefficient uncertainty, information about the missing structure will be present in the outlier experiments; removing the experiments from consideration destroys this information. There is no easy way to determine which class any particular outlier falls.

In later work, Turányi and coworkers [248] tried to resolve these issues using a hybrid approach, in which certain measurements with potentially large systematic errors were not considered in the optimization to begin with. Furthermore, after the first uncertainty minimization step, measurements that could not be reproduced to within $4\sigma^{\text{obs}}$ were removed from consideration, and a second minimization step was conducted.

4. Selected problems of UQ applications in combustion chemistry

This section highlights some of the recent applications of uncertainty quantification in combustion chemistry. It is shown that

the UQ analysis can yield from useful to critical results that are not obtainable without the use of the UQ methods.

4.1. Rate rules and their uncertainties

All reaction kinetic models of large hydrocarbon fuels use rate rules for their construction. Such a practice is rooted in the chemical similarity rule in which elementary reactions similar in nature usually have similar rate coefficients. The rate-rule assumption is the foundation of all automatic reaction mechanism generation and chemical lumping (see, e.g., Refs. [255–261]). The rate-rule principle removes the need to know the rate coefficient of every reaction before a reaction model can be constructed, as this is impossible in combustion chemistry. The practice also introduces two coupled uncertainties: the measured or computed rate coefficient used for a class of reaction is inherently uncertain, and the application of the rate rule for a similar but different reaction can lead to additional uncertainty. In principle, no two elementary reactions have the same rate coefficients. The chemical similarity rule is accurate only to first order and higher order effects are then the uncertainty in the individual rate parameters among a class of similar reactions. In this case, the rate parameters within the same reaction class are all coupled.

Rate-rule induced uncertainty has received some recent attention, though a general consensus about how to treat this uncertainty and especially the coupling of the rate parameters in a reaction class is still emerging. In a recent paper, Najm, Pitz and coworkers [176] carried out a study of uncertainty quantification of reaction models with a specific aim to account for correlations introduced by rate rules and fitted Arrhenius parameters. Of particular interest is that the authors highlighted the improved fidelity of the model when correlations are accounted for and the computational savings associated with the rate rule due to the reduced dimensionality of the UQ problem. The authors used a recent foundational fuel chemistry model of Curran and coworkers [262,263] and examined rate parameter correlation and rate rule uncertainties on the prediction of ignition delay times of a stoichiometric propane/ethane/air mixture at constant pressure of 10 atm and initial temperature of 1000 K. They found that accounting for dependence among rate parameters of rate rule in origin can lead to different results in the uncertainty predicted for the ignition delay time. Overall, the study suggests that proper accounting for rate rules and rate parameter fitting correlations is critical to robust and accurate estimation of uncertainty in predictions using chemical kinetic models.

In practical reaction model development using an inverse problem, the rate parameters within a reaction class may be assumed to be initial guesses of the rate value and treated as being decoupled from each other. This was the underlying assumption in Sheen and Wang [214] in their modeling study of multispecies time histories of n-heptane oxidation in a shock tube. In a more recent study, Cai and Pitsch [175] took a quite different view. They argued that modification of individual rate parameters within a reaction class violates the consistency of rate parameters of similar reactions and in this context, it is more useful to constrain the rate parameters for the entire class as such an approach would lead to a more consistent model optimization. Using MUM-PCE, they demonstrated that it is possible to make improved prediction and reduce the prediction uncertainty by optimizing a reaction model against a range of shock tube ignition delay data for *n*-pentane.

4.2. Consistency of experimental dataset

All combustion experiments have their associated uncertainties, some of which are random while others are systematic. In applying

any inverse problem approaches to chemical kinetic modeling, a careful examination of the consistency of the experimental target dataset is critical to the successful treatment of the problem. Let us examine Eq. (46) as an example. Each experimental outlier would skew the optimal set of the rate parameters and their feasible ranges in the direction of the outlier. It is conceivable by considering two inconsistent data points, each giving a minimum in the objective function, that the parameter constraining process would yield an optimal parameter set that lies in the region separating the two minima with the uncertainty spanning both sides of the local minima. Detecting data outliers serves two purposes: one that allows us to question the systematic or random error in the experimental measurement in a systematic manner, and the other allowing for elimination of erroneous measurements. The conclusion of such an analysis is admittedly the result of a “majority-rule” principle. The data points deemed inconsistent or in the minority can indicate missing physics for the problem, e.g., including missing reaction pathways or excessively bad rate estimates. Hence, outlier detection can yield very useful results, but it should be used with caution.

As an example, Frenklach and coworkers [250] extended the approach of data collaboration to address the mutual consistency of experimental observations. In their analysis, two major outliers of the GRI-Mech II dataset were detected. Both are of the reaction times to reach half of the maximal OH concentration in shock tube oxidation of methane. It turns out that these were not actual measurements but two “representative” points, at 2000 and 2200 K, extrapolated from a fit to a series of measurements performed over a range of temperatures [264]. The pertinent researchers were contacted about the results that had been flagged by the analysis, though no indication was given about which direction the measurements might be in error. A re-examination of the original experimental data led the researchers to modify the extrapolation and yield new improved estimation for the two data points, both consistent with the direction indicated by the outlier analysis. This example highlights the utility of outlier detection and more importantly, the need for collaborative experimental, theoretical and UQ efforts to achieve measurable progress.

4.3. Assessment of systematic experimental uncertainty

The use of UQ to unravel systematic experimental uncertainty is at its infancy. In a recent paper, Iaccarino, Hanson and coworkers [265] examined the influences of initial radical impurities on the determination of ignition delay times in H₂/O₂ mixtures. Two separate methods were employed to quantify the H atom impurities. The first method involves the measurement of the OH concentration upon shock heating of an oxygen–diluent mixture in absence of the hydrocarbon reactant. The second method infers the H atom impurity concentration by comparing the experimentally measured ignition delay against calculations. The analysis allows them to propose a stochastic Arrhenius model that describes the amount of initial H-radical impurities behind reflected shock waves. Furthermore, theoretical quantifications of the uncertainties induced by the initial H-atom impurities shows that the effect becomes more pronounced in the thermodynamic condition space where the rate of H + O₂ radical chain branching reaction is greater than the rate of H + O₂ chain termination reaction, as expected, and such an effect is magnified with an increased dilution and with decreased temperature or pressure. The effect becomes negligible when a dominant source of radical quenching exists relative to the “natural” radical initiation rate controlled by the H₂ + O₂ → H + HO₂.

There are several points that may be made following Urzay et al. [265]. By treating the H atom impurity as an uncertainty in the

shock tube experiments, the authors were able to isolate the influence of initial radical impurities on the experiment. Monte Carlo simulations showed that the H atom impurity can impact the ignition delay by as much as 30%. Based on the analysis, the authors uncovered the conditions over which the impurity effect can become magnified, and as such, they demonstrate that a careful UQ analysis of the experimental data coupled with kinetic modeling can yield useful information for future experimental design.

4.4. Hierarchical model development

It has been understood for a long while that the mechanism of high-temperature hydrocarbon oxidation has a hierarchical structure [38]. The construction of reaction models of higher hydrocarbon combustion begins with a model for H₂/CO oxidation. This base model is then rigorously checked against a set of data to ensure its accuracy. The model then forms the basis for a CH₄ oxidation model, which is checked against its own set of data. This process is continued to encompass progressively higher-order hydrocarbons. While this is a logical way to proceed, the quantitative accuracy of the result following this approach and the uncertainty propagated from the lower level models of small hydrocarbon combustion to those of the high-order hydrocarbons has never been determined. There is always the concern whether a higher-order hydrocarbon, say propane, may activate otherwise dormant paths in say, the CH₄ oxidation model. Additionally, questions have been raised about what should be an efficient experimental approach to arrive at a predictive combustion chemistry model. For example, if one is interested in having a model to predict accurate the heat release rate in laminar premixed combustion of *n*-octane, what would be the experimental focus, keeping in mind that a direct measurement on the combustion properties of *n*-octane may not be the most efficient or direct path to an accurate *n*-octane model.

Quantitative relations among the hierarchical components of a comprehensive hydrocarbon combustion reaction model were recently examined by Sheen and Wang [266]. Using *n*-octane

combustion as an example and MUM-PCE as the underlying method, they explore the accuracy of the hierarchical model strategy by observing how the model constraining at each step affects the later stages. Fig. 23 shows the prediction uncertainties for the laminar flame speed of *n*-C₈H₁₈-air mixtures at 1 atm and 353 K unburned gas temperature. The prior model (JetSurf 1.0 [224]) predicts the flame speed to have 3–4 cm/s of 1 σ uncertainty. The model was first constrained against an H₂/CO and C₂H₄ experimental dataset, leading to Model III, which gives a significantly reduced uncertainty in the predicted *n*-octane flame speeds. The 1 σ uncertainty reduces from 3 to 4 cm/s to about 1.5 cm/s. With further constraining using a C₃ experimental dataset, the prediction uncertainty decreases to 1 cm/s. Additional constraining, however, did not lead to any noticeable improvements in terms prediction uncertainty. The point of the study is that a high-precision model for a higher-hydrocarbon combustion is usually better achieved by constraining the hierarchical submodels at the H₂/CO/C₂H₄/C₃H₆ level because it is the pyrolysis and oxidation chemistry of these smaller species that govern the overall heat release rates.

4.5. Model reduction

As an enabling tool for the application of combustion chemistry model in computational fluid dynamics, a wide range of model reduction methods have been proposed and put in use (see, e.g., Refs. [125,267]). There are two aspects of UQ analyses in model reduction. The first issue concerns the criteria in deriving or the assessment of the various truncation and approximation errors in a reduced model. Traditionally the accuracy of a reduced or skeletal model is assessed by comparing its performance against a detailed model. Sheen et al. [15] pointed out that without considering the inherent uncertainty in the detailed model, this comparison can over-constrain the reduction criteria. In fact, a reduced model can be said as accurate so long as it reproduces the detailed model within its uncertainty bound. Second, although error propagation algorithms have been proposed [268] within the framework of the Directed Relation Graph (DRG) method [269,270], a basic treatment

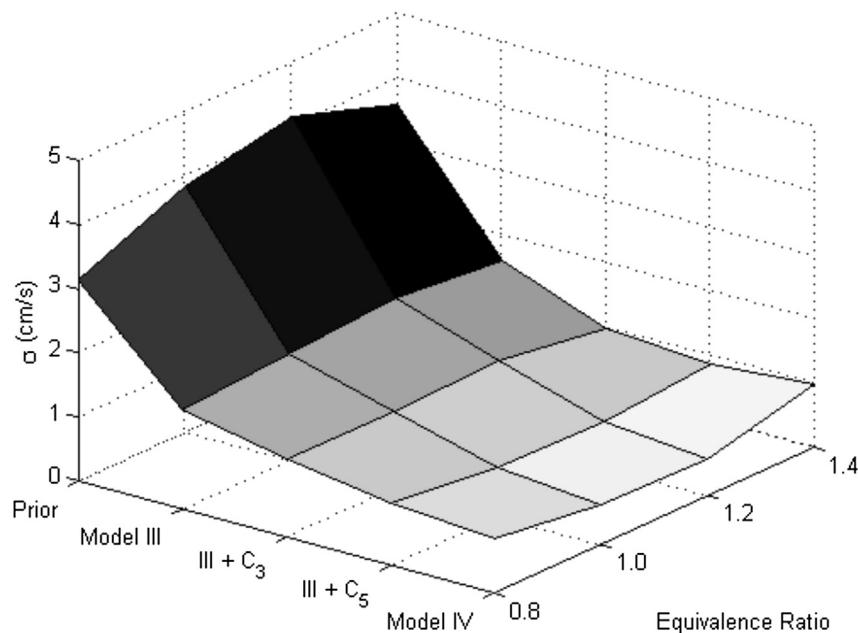


Fig. 23. Prediction uncertainties for the laminar flame speed of *n*-C₈H₁₈-air mixtures at 1 atm and 353 K unburned gas temperature for several model cases. Shown are the prior model (unconstrained JetSurf1.0 [224]) and posterior models constrained against an H₂/CO and C₂H₄ experimental dataset (Model III), with additional constraining from i) C₃ experimental dataset, ii) a C₃+C₅ dataset, and iii) all experiments up to *n*-heptane (Model IV). The plot is adapted from Sheen and Wang [266].

of the underlying uncertainty magnification and propagation in model reduction remains largely unavailable.

In a recent study, Xin et al. [271] used the combustion of *n*-butane and *i*-butane as examples and derived reduced models from a detailed chemistry model through directed relation graph (DRG) and DRG-aided sensitivity analysis (DRGASA) methods. The accuracy of the reduced models was improved by constraining the reduced model against experimental data. The dependence of model uncertainty on the reduced model size was investigated systematically by exploring skeletal models containing different number of species. It was found that model uncertainty is subject to the completeness of the model with respect to the underlying physics. The uncertainty of a complete model is convergent and asymptotic with respect to the model size while an incomplete model is characterized by uncertainty behavior that is difficult to predict. In particular, they showed that for a model having reactions removed that exhibit second-order coupling with the remaining reactions of the model, the accuracy of the model starts to fall apart.

4.6. Large eddy simulation

The ultimate aim of combustion modeling is to be able to understand turbulent combustion in engines or other energy conversion devices. The recent work of Mueller et al. [272] extended the forward uncertainty quantification work into large eddy simulations (LES) of turbulent flames. In this work, they performed an LES solution for Sandia Flame D [273]—a turbulent jet flame consisting of a methane/air core surrounded by a fuel-lean acetylene/hydrogen/air pilot flame. The core is richer than the rich flammability limit, and unburned oxygen from the pilot mixing into the core causes the flame to ignite. The LES was performed using the steady flamelet model of Peters [274]. In this algorithm, a set of laminar diffusion flames at varying strain rate is solved. The solutions are used to generate a lookup table for temperature, specific volume, and composition as a function of the fuel/air mixture fraction and scalar dissipation rate. These are termed the derived properties.

The uncertainty in the LES solutions due to the chemical model uncertainty was calculated using Gauss-Hermite quadrature as in Eq. (38). No screening is used, so the LES uncertainty must be calculated using the uncertainty of the full model. To reduce the computational complexity, brute-force uncertainty analysis was used only on the flamelet lookup table, which allowed the LES to be solved over a much lower-dimensional uncertainty space. The Arrhenius *A* factor of each rate coefficient was assumed to have a lognormal distribution, so that the uncertain parameters could be normalized using Eq. (32). The flamelet lookup table was determined by Latin hypercube sampling within the reaction uncertainty space. It was found that the derived properties in the lookup were approximately linear with respect to the reduced variables \mathbf{x} . This means, for example, that the specific volume v can be expressed as $v = \mathbf{a}\mathbf{x}$, where \mathbf{a} is the vector of sensitivities of specific volume to the rate coefficients. Furthermore, all of the derived properties were closely correlated, so in effect the problem could be treated as being one-dimensional and the quadrature needed relatively few LES evaluations. Results for the LES-predicted temperature and CO mass fraction profiles and their comparison with experimental measurements are shown in Fig. 24.

5. Perspectives and concluding remarks

- (1) As a pillar of predictive science, uncertainty quantification should be an integral part of any modeling effort. UQ is necessary to combustion chemistry development. Making

predictions matters, but a characterization about how good the predictions are matters perhaps even more.

- (2) Significant progress in combustion chemistry has been made over the last few decades, but in comparison to a large portfolio of hydrocarbon fuels one may have to consider, the pace of progress has been slow. One of the major reasons has been the lack of consensus as to how to move forward in obtaining a comprehensive, unified and predictive model of fuel combustion. We argue that the solution is decidedly in the UQ analysis. The various models, at least for the combustion of small hydrocarbon fuels, in fact, can be and should be unified since the different choices of the rate coefficients can be interpreted as the uncertainties of the respective rate parameters. The comprehensiveness of the model can be addressed, in principle, by designing experiments with well-characterized uncertainties targeted to reduce the joint model parameter uncertainty space. The predictiveness of the model is then assessed and ensured if the uncertainty of the model can be defined outside of the thermodynamic regimes tested and constrained.

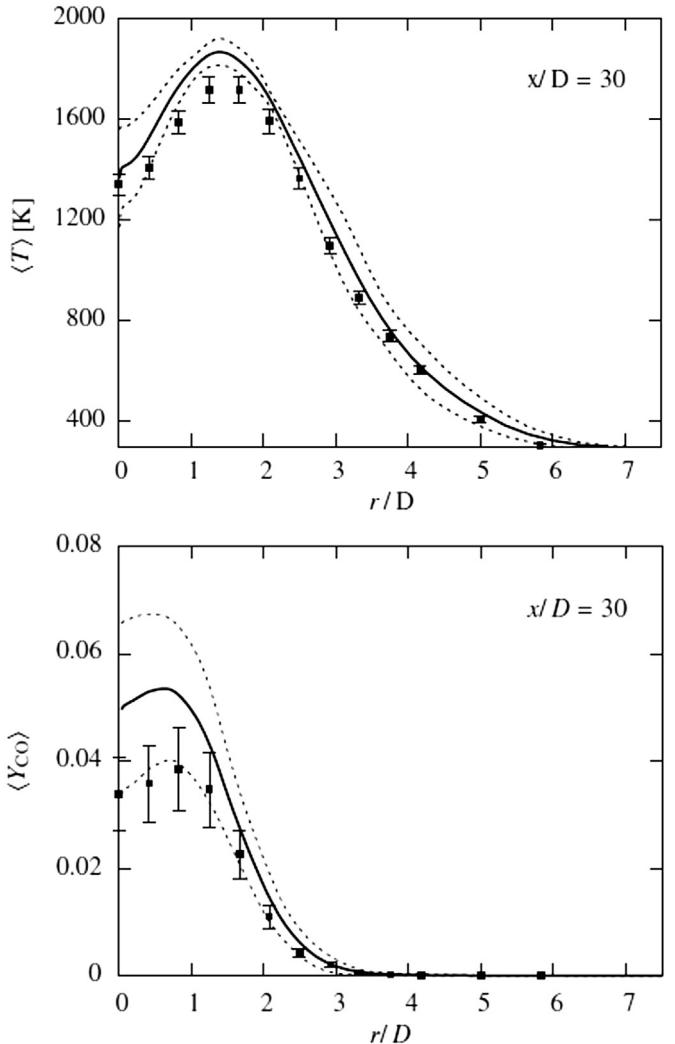


Fig. 24. Mean radial profiles of temperature and CO mass fraction in Sandia Flame D [273]. Symbols are the experimental measurements and uncertainties. The solid line is the mean LES prediction and the dashed lines indicate the 95% confidence interval of the LES. The plot is adapted from Mueller et al. [272].

- (3) The application of UQ is inherently data and measurement intensive. The work is best conducted as a collaborative effort, and such an effort has proven possible in the past (e.g., the GRI-Mech effort). Currently computer codes that allow for various UQ analyses are limited to in-house codes. There is a need to make these codes available to a broader combustion chemistry community for benchmarking, testing and utilization. Prior efforts, including the PrIMe initiative [119], might be important steps in that direction.
- (4) The current paper provides a quite comprehensive review about the mathematical principles and methods of uncertainty quantification. Discussion is given about the sources and classification of kinetic uncertainties and the meanings and definitions of model verification and validation. The various UQ methods are reviewed and classified, and their applications are illustrated for problems involving forward uncertainty quantification and propagation, and as an inverse problem leading to model uncertainty constraining. It is shown that for combustion chemistry problems, while UQ alone cannot produce accurate, individual rate parameters, it can be instrumental in measuring the progress of our understanding of combustion chemistry and in utilizing fundamental combustion property data beyond a simple “agree–disagree” statement. When treated as a Bayesian inference problem, UQ is shown to aid the development of predictive kinetic models in two separate aspects: the use of fundamental combustion property data, global or local, to provide the joint probability density function of the rate parameters leading to a better constrained kinetic model. Along with the forward kinetic uncertainty propagation, UQ can yield estimates of model confidence to make predictions outside of the thermodynamic regimes where the model has been tested.
- (5) This paper has demonstrated that UQ methods are mature enough to become integral parts of combustion chemistry research. These methods extend beyond simple sensitivity analysis and transform detailed kinetic modeling from semi-quantitative tests to the realm of quantitative modeling and comparison. To this end, quantifying the limitations or the uncertainty of a model is critical to identifying key scientific problems. UQ methods should also provide a solution to the “many-model” problem currently facing the combustion chemistry community.
- (6) The primary challenge that lies ahead is the advancement of UQ tools that are widely accessible and easy to use by the community. Currently, UQ methods such as those presented in this paper have been slow to gain traction in the community. In order to gain wider acceptance, tools will need better documentation and easy-to-use interface. The PrIMe workflow, which implements the Data Collaboration method, have been steps in the right direction.

Disclaimer

Certain trade names are identified in this paper in order to specify the procedure adequately. Such identification is not intended to imply recommendation or endorsement by the National Institute of Standards and Technology, nor is it intended to imply that the trade names identified are necessarily the best available for the purpose.

Acknowledgments

Work at Stanford was supported by the Combustion Energy Frontier Research Center (CEFRC), an Energy Frontier Research

Center funded by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences under Award Number DE-SC0001198. The authors are grateful to all anonymous reviewers for offering their perspectives, insights and suggestions.

References

- [1] Box GE, Draper NR. *Empirical model-building and response surfaces*. John Wiley & Sons; 1987.
- [2] Frenklach M, Wang H, Rabinowitz MJ. Optimization and analysis of large chemical kinetic mechanisms using the solution mapping method—combustion of methane. *Progr Energy Combust Sci* 1992;18:47–73.
- [3] Manion J, Huie R, Levin R, Burgess Jr D, Orkin V, Tsang W, et al. NIST chemical kinetics database, NIST standard reference database 17, version 7.0 (Web version), release 1.6. 7 data version 2013.03. Gaithersburg, MD: National Institute of Standards and Technology; 2013.
- [4] Baldwin R, Melvin A. The reaction of hydrogen atoms with oxygen and with ethane. *J Chem Soc* 1964;1785–91.
- [5] Cohen N, Westberg K. Chemical kinetic data sheets for high-temperature chemical reactions. *J Phys Chem Ref Data* 1983;12:531–90.
- [6] Warnatz J. Rate coefficients in the C/H/O system. *Combustion chemistry*: Springer; 1984. p. 197–360.
- [7] Tsang W, Hampson R. Chemical kinetic data base for combustion chemistry. Part I. Methane and related compounds. *J Phys Chem Ref Data* 1986;15: 1087–279.
- [8] Baulch D, Cobos C, Cox R, Esser C, Frank P, Just T, et al. Evaluated kinetic data for combustion modelling. *J Phys Chem Ref Data* 1992;21:411–734.
- [9] Du H, Hessler JP. Rate coefficient for the reaction $H+O_2 \rightarrow OH+O$: results at high temperatures, 2000 to 5300 K. *J Chem Phys* 1992;96:1077–92.
- [10] Baulch D, Cobos C, Cox R, Frank P, Hayman G, Just T, et al. Evaluated kinetic data for combustion modeling. *Supplement I. J Phys Chem Ref Data* 1994;23: 847–1033.
- [11] Yu C-L, Frenklach M, Masten D, Hanson R, Bowman C. Reexamination of shock-tube measurements of the rate coefficient of $H+O_2 \rightarrow OH+O$. *J Phys Chem* 1994;98:4770–1.
- [12] Hong Z, Davidson D, Barbour E, Hanson R. A new shock tube study of the $H+O_2 \rightarrow OH+O$ reaction rate using tunable diode laser absorption of H_2O near 2.5 μm . *Proc Combust Inst* 2011;33:309–16.
- [13] Buxton GV, Greenstock CL, Helman WP, Ross AB. Critical review of rate constants for reactions of hydrated electrons, hydrogen atoms and hydroxyl radicals. *Phys Chem Ref Data* 1988;17:513–886.
- [14] Hanson RK. Applications of quantitative laser sensors to kinetics, propulsion and practical energy systems. *Proc Combust Inst* 2011;33:1–40.
- [15] Sheen DA, You X, Wang H, Lovas T. Spectral uncertainty quantification, propagation and optimization of a detailed kinetic model for ethylene combustion. *Proc Combust Inst* 2009;32:535–42.
- [16] Wang H, You X, Joshi AV, Davis SG, Laskin A, Egolfopoulos FN, et al. USC mech version II. High-temperature combustion reaction model of $H_2/CO/C_1-C_4$ compounds. 2007.
- [17] Frenklach M. Systematic optimization of a detailed kinetic model using a methane ignition example. *Combust Flame* 1984;58:69–72.
- [18] Sacks J, Welch WJ, Mitchell TJ, Wynn HP. Design and analysis of computer experiments. *Stat Sci* 1989;4:409–23.
- [19] Berger JO. *Statistical decision theory and Bayesian analysis*. Springer; 1985.
- [20] Council NR. *Completing the forecast: characterizing and communicating uncertainty for better decisions using weather and climate forecasts*. National Academies Press; 2006.
- [21] Dixon-Lewis G. Flame structure and flame reaction kinetics. I. Solution of conservation equations and application to rich hydrogen-oxygen flames. *Proc R Soc Lond A* 1967;298:495–513.
- [22] Hinshelwood CN. *The kinetics of chemical change in gaseous systems*. 3rd ed. Oxford: Oxford Press; 1933.
- [23] Semenov NN. *Some problems of chemical kinetics and reactivity*. Elsevier; 1958.
- [24] von Elbe G. *Chemical kinetics of hydrocarbon combustion*. Symp (Int) Combust 1955;5:79–85.
- [25] Baldwin R, Mayor L. The mechanism of the hydrogen+ oxygen reaction in aged boric-acid-coated vessels. *Trans Faraday Soc* 1960;56:103–14.
- [26] Seery DJ, Bowman CT. An experimental and analytical study of methane oxidation behind shock waves. *Combust Flame* 1970;14:37–47.
- [27] Bowman CT. A shock-tube investigation of the high-temperature oxidation of methanol. *Combust Flame* 1975;25:343–54.
- [28] Dixon-Lewis G, Goldsworthy F, Greenberg J. Flame structure and flame reaction kinetics IX. Calculation of properties of multi-radical premixed flames. *Proc R Soc Lond A* 1975;346:261–78.
- [29] Dixon-Lewis G. Kinetic mechanism, structure and properties of premixed flames in hydrogen-oxygen-nitrogen mixtures. *Philos Trans R Soc Lond A* 1979;45–99.
- [30] Westbrook CK, Dryer FL. Prediction of laminar flame properties of methanol-air mixtures. *Combust Flame* 1980;37:171–92.
- [31] Cathonnet M, Boettner J, James H. Experimental study and numerical modeling of high temperature oxidation of propane and n-butane. *Symp (Int) Combust* 1981;18:903–13.

- [32] Cherian M, Rhodes P, Simpson R, Dixon-Lewis G. Structure, chemical mechanism and properties of premixed flames in mixtures of carbon monoxide, nitrogen and oxygen with hydrogen and water vapour. *Philos Trans R Soc Lond A* 1981;303:181–212.
- [33] Warnatz J. The structure of laminar alkane-, alkene-, and acetylene flames. *Symp (Int) Combust* 1981;18:369–84.
- [34] Miller JA, Mitchell RE, Smooke MD, Kee RJ. Toward a comprehensive chemical kinetic mechanism for the oxidation of acetylene: comparison of model predictions with results from flame and shock tube experiments. *Symp (Int) Combust* 1982;19:181–96.
- [35] Miller D, Frenklach M. Sensitivity analysis and parameter estimation in dynamic modeling of chemical kinetics. *Int J Chem Kinet* 1983;15:677–96.
- [36] Warnatz J. The mechanism of high temperature combustion of propane and butane. *Combust Sci Technol* 1983;34:177–200.
- [37] Frenklach M, Bornside DE. Shock-initiated ignition in methane-propane mixtures. *Combust Flame* 1984;56:1–27.
- [38] Westbrook CK, Dryer FL. Chemical kinetic modeling of hydrocarbon combustion. *Progr Energy Combust Sci* 1984;10:1–57.
- [39] Westbrook CK, Pitz WJ. A comprehensive chemical kinetic reaction mechanism for oxidation and pyrolysis of propane and propene. *Combust Sci Technol* 1984;37:117–52.
- [40] Brezinsky K. The high-temperature oxidation of aromatic hydrocarbons. *Progr Energy Combust Sci* 1986;12:1–24.
- [41] Dagaut P, Cathonnet M, Rouan J, Foulquier R, Quilgars A, Boettner J, et al. A jet-stirred reactor for kinetic studies of homogeneous gas-phase reactions at pressures up to ten atmospheres (≈ 1 MPa). *J Phys E: Sci Instrum* 1986;19:207.
- [42] Hwang S, Gardiner Jr W, Frenklach M, Hidaka Y. Induction zone exothermicity of acetylene ignition. *Combust Flame* 1987;67:65–75.
- [43] Axelsson E, Brezinsky K, Dryer F, Pitz W, Westbrook C. Chemical kinetic modeling of the oxidation of large alkane fuels: n-octane and iso-octane. *Symp (Int) Combust* 1988;21:783–93.
- [44] Garo A, Westmoreland P, Howard J, Longwell J. Analysis of fuel-lean combustion using chemical mechanisms. *Combust Flame* 1988;72:271–86.
- [45] Westmoreland PR, Howard JB, Longwell JP. Tests of published mechanisms by comparison with measured laminar flame structure in fuel-rich acetylene combustion. *Symp (Int) Combust* 1988;21:773–82.
- [46] Chakir A, Cathonnet M, Boettner J, Gaillard F. Kinetic study of n-butane oxidation. *Combust Sci Technol* 1989;65:207–30.
- [47] Chakir A, Cathonnet M, Boettner J, Gaillard F. Kinetic study of 1-butene oxidation in a jet-stirred flow reactor. *Symp (Int) Combust* 1989;22:873–81.
- [48] Miller JA, Bowman CT. Mechanism and modeling of nitrogen chemistry in combustion. *Progr Energy Combust Sci* 1989;15:287–338.
- [49] Westbrook CK, Warnatz J, Pitz WJ. A detailed chemical kinetic reaction mechanism for the oxidation of iso-octane and n-heptane over an extended temperature range and its application to analysis of engine knock. *Symp (Int) Combust* 1989;22:893–901.
- [50] Wilk R, Cernansky N, Pitz W, Westbrook C. Propene oxidation at low and intermediate temperatures: a detailed chemical kinetic study. *Combust Flame* 1989;77:145–70.
- [51] Miller JA, Kee RJ, Westbrook CK. Chemical kinetics and combustion modeling. *Annu Rev Phys Chem* 1990;41:345–87.
- [52] Chakir A, Belumam M, Boettner J, Cathonnet M. Kinetic study of n-pentane oxidation. *Combust Sci Technol* 1991;77:239–60.
- [53] Emdee J, Brezinsky K, Glassman I. Oxidation of o-xylene. *Symp (Int) Combust* 1991;23:77–84.
- [54] Emdee J, Brezinsky K, Glassman I. High-temperature oxidation mechanisms of m-and p-xylene. *J Phys Chem* 1991;95:1626–35.
- [55] Westbrook CK, Pitz WJ, Leppard WR. The autoignition chemistry of paraffinic fuels and pro-knock and anti-knock additives: a detailed chemical kinetic study. SAE Technical Paper. 1991.
- [56] Yetter R, Dryer F, Rabitz H. Flow reactor studies of carbon monoxide/hydrogen/oxygen kinetics. *Combust Sci Technol* 1991;79:129–40.
- [57] Yetter R, Dryer F, Rabitz H. A comprehensive reaction mechanism for carbon monoxide/hydrogen/oxygen kinetics. *Combust Sci Technol* 1991;79:97–128.
- [58] Chakir A, Bellimam M, Boettner J, Cathonnet M. Kinetic study of n-heptane oxidation. *Int J Chem Kinet* 1992;24:385–410.
- [59] Emdee J, Brezinsky K, Glassman I. A kinetic model for the oxidation of toluene near 1200 K. *J Phys Chem* 1992;96:2151–61.
- [60] Dagaut P, Reuillon M, Cathonnet M. High pressure oxidation of liquid fuels from low to high temperature. 1. n-Heptane and iso-Octane. *Combust Sci Technol* 1993;95:233–60.
- [61] Hidaka Y, Taniguchi T, Tanaka H, Kamesawa T, Inami K, Kawano H. Shock-tube study of CH_2O pyrolysis and oxidation. *Combust Flame* 1993;92:365–76.
- [62] Hunter T, Wang H, Litzinger T, Frenklach M. The oxidation of methane at elevated pressures: experiments and modeling. *Combust Flame* 1994;97:201–24.
- [63] Kim TJ, Yetter RA, Dryer FL. New results on moist CO oxidation: high pressure, high temperature experiments and comprehensive kinetic modeling. *Symp (Int) Combust* 1994;25:759–66.
- [64] Lindstedt R, Skevis G. Detailed kinetic modeling of premixed benzene flames. *Combust Flame* 1994;99:551–61.
- [65] Tan Y, Dagaut P, Cathonnet M, Boettner J-C. Acetylene oxidation in a JSR from 1 to 10 atm and comprehensive kinetic modeling. *Combust Sci Technol* 1994;102:21–55.
- [66] Leung K, Lindstedt R. Detailed kinetic modeling of $\text{C}_1\text{--C}_3$ alkane diffusion flames. *Combust Flame* 1995;102:129–60.
- [67] Lindstedt R, Maurice L. Detailed kinetic modelling of n-heptane combustion. *Combust Sci Technol* 1995;107:317–53.
- [68] Ranzi E, Gaffuri P, Faravelli T, Dagaut P. A wide-range modeling study of n-heptane oxidation. *Combust Flame* 1995;103:91–106.
- [69] Zhang H-Y, McKinnon JT. Elementary reaction modeling of high-temperature benzene combustion. *Combust Sci Technol* 1995;107:261–300.
- [70] Curran H, Simmie JM, Dagaut P, Voisin D, Cathonnet M. The ignition and oxidation of allene and propyne: experiments and kinetic modeling. *Symp (Int) Combust* 1996;26:613–20.
- [71] Dagaut P, Boettner J-C, Cathonnet M. Chemical kinetic study of dimethyl-ether oxidation in a jet stirred reactor from 1 to 10 ATM: experiments and kinetic modeling. *Symp (Int) Combust* 1996;26:627–32.
- [72] Davis S, Wang H, Brezinsky K, Law C. Laminar flame speeds and oxidation kinetics of benene-air and toluene-air flames. *Symp (Int) Combust* 1996;26:1025–33.
- [73] Hidaka Y, Hattori K, Okuno T, Inami K, Abe T, Koike T. Shock-tube and modeling study of acetylene pyrolysis and oxidation. *Combust Flame* 1996;107:401–17.
- [74] Hunter T, Litzinger T, Wang H, Frenklach M. Ethane oxidation at elevated pressures in the intermediate temperature regime: experiments and modeling. *Combust Flame* 1996;104:505–23.
- [75] Lindstedt R, Maurice L. Detailed kinetic modelling of toluene combustion. *Combust Sci Technol* 1996;120:119–67.
- [76] Held T, Marchese A, Dryer F. A semi-empirical reaction mechanism for n-heptane oxidation and pyrolysis. *Combust Sci Technol* 1997;123:107–46.
- [77] Wang H, Frenklach M. A detailed kinetic modeling study of aromatics formation in laminar premixed acetylene and ethylene flames. *Combust Flame* 1997;110:173–221.
- [78] Curran H, Gaffuri P, Pitz WJ, Westbrook CK. A comprehensive modeling study of n-heptane oxidation. *Combust Flame* 1998;114:149–77.
- [79] Curran H, Pitz W, Westbrook C, Callahan G, Dryer F. Oxidation of automotive primary reference fuels at elevated pressures. *Symp (Int) Combust* 1998;27:379–87.
- [80] Curran H, Pitz W, Westbrook C, Dagaut P, Boettner JC, Cathonnet M. A wide range modeling study of dimethyl ether oxidation. *Int J Chem Kinet* 1998;30:229–41.
- [81] Davis S, Law C, Wang H. An experimental and kinetic modeling study of propyne oxidation. *Symp (Int) Combust* 1998;27:305–12.
- [82] Held TJ, Dryer FL. A comprehensive mechanism for methanol oxidation. *Int J Chem Kinet* 1998;30:805–30.
- [83] Davis S, Law C, Wang H. Propene pyrolysis and oxidation kinetics in a flow reactor and laminar flames. *Combust Flame* 1999;119:375–99.
- [84] Davis SG, Law CK, Wang H. Propyne pyrolysis in a flow reactor: an experimental, RRKM, and detailed kinetic modeling study. *J Phys Chem A* 1999;103:5889–99.
- [85] Laskin A, Wang H. On initiation reactions of acetylene oxidation in shock tubes: a quantum mechanical and kinetic modeling study. *Chem Phys Lett* 1999;303:43–9.
- [86] Mueller M, Kim T, Yetter R, Dryer F. Flow reactor studies and kinetic modeling of the H_2/O_2 reaction. *Int J Chem Kinet* 1999;31:113–25.
- [87] Mueller M, Yetter R, Dryer F. Flow reactor studies and kinetic modeling of the $\text{H}_2/\text{O}_2/\text{NO}_x$ and $\text{CO}/\text{H}_2\text{O}/\text{O}_2/\text{NO}_x$ reactions. *Int J Chem Kinet* 1999;31:705–24.
- [88] El Bakali A, Braun-Unkhoff M, Dagaut P, Frank P, Cathonnet M. Detailed kinetic reaction mechanism for cyclohexane oxidation at pressure up to ten atmospheres. *Proc Combust Inst* 2000;28:1631–8.
- [89] Laskin A, Wang H, Law CK. Detailed kinetic modeling of 1, 3-butadiene oxidation at high temperatures. *Int J Chem Kinet* 2000;32:589–614.
- [90] Qin Z, Lissianski VV, Yang H, Gardiner WC, Davis SG, Wang H. Combustion chemistry of propane: a case study of detailed reaction mechanism optimization. *Proc Combust Inst* 2000;28:1663–9.
- [91] Simmie JM. Detailed chemical kinetic models for the combustion of hydrocarbon fuels. *Progr Energy Combust Sci* 2003;29:599–634.
- [92] Miller JA. Theory and modeling in combustion chemistry. *Symp (Int) Combust* 1996;26:461–80.
- [93] Westbrook CK. Chemical kinetics of hydrocarbon ignition in practical combustion systems. *Proc Combust Inst* 2000;28:1563–77.
- [94] Edwards T, Maurice LQ. Surrogate mixtures to represent complex aviation and rocket fuels. *J Propul Power* 2001;17:461–6.
- [95] Violai A, Yan S, Eddings E, Sarofim A, Granata S, Faravelli T, et al. Experimental formulation and kinetic model for JP-8 surrogate mixtures. *Combust Sci Technol* 2002;174:399–417.
- [96] Colket M, Edwards T, Williams S, Cernansky NP, Miller DL, Egolfopoulos F, et al. Development of an experimental database and kinetic models for surrogate jet fuels. 2007.
- [97] Pitz W, Cernansky N, Dryer F, Egolfopoulos F, Farrell J, Friend D, et al. Development of an experimental database and chemical kinetic models for surrogate gasoline fuels. SAE Paper. 2007:0175.
- [98] Pitz WJ, Mueller CJ. Recent progress in the development of diesel surrogate fuels. *Progr Energy Combust Sci* 2011;37:330–50.
- [99] Alamo GD, Williams FA, Sanchez AL. Hydrogen–oxygen induction times above crossover temperatures. *Combust Sci Technol* 2004;176:1599–626.
- [100] O’Conaire M, Curran HJ, Simmie JM, Pitz WJ, Westbrook CK. A comprehensive modeling study of hydrogen oxidation. *Int J Chem Kinet* 2004;36:603–22.

- [101] Davis SG, Joshi AV, Wang H, Egolfopoulos FN. An optimized kinetic model of H₂/CO combustion. *Proc Combust Inst* 2005;30:1283–92.
- [102] Zsély IG, Zádor J, Turányi T. Uncertainty analysis of updated hydrogen and carbon monoxide oxidation mechanisms. *Proc Combust Inst* 2005;30:1273–81.
- [103] Saxena P, Williams FA. Testing a small detailed chemical-kinetic mechanism for the combustion of hydrogen and carbon monoxide. *Combust Flame* 2006;145:316–23.
- [104] Sun H, Yang S, Jomaas G, Law C. High-pressure laminar flame speeds and kinetic modeling of carbon monoxide/hydrogen combustion. *Proc Combust Inst* 2007;31:439–46.
- [105] Konnov AA. Remaining uncertainties in the kinetic mechanism of hydrogen combustion. *Combust Flame* 2008;152:507–28.
- [106] Burke MP, Chaos M, Dryer FL, Ju Y. Negative pressure dependence of mass burning rates of H₂/CO/O₂/diluent flames at low flame temperatures. *Combust Flame* 2010;157:618–31.
- [107] Burke MP, Chaos M, Ju Y, Dryer FL, Klippenstein SJ. Comprehensive H₂/O₂ kinetic model for high-pressure combustion. *Int J Chem Kinet* 2012;44:444–74.
- [108] Li J, Zhao Z, Kazakov A, Dryer FL. An updated comprehensive kinetic model of hydrogen combustion. *Int J Chem Kinet* 2004;36:566–75.
- [109] Olm C, Zsély IG, Pálvölgyi R, Varga T, Nagy T, Curran HJ, et al. Comparison of the performance of several recent hydrogen combustion mechanisms. *Combust Flame* 2014;161:2219–34.
- [110] Kéromnès A, Metcalfe WK, Heufer KA, Donohoe N, Das AK, Sung C-J, et al. An experimental and detailed chemical kinetic modeling study of hydrogen and syngas mixture oxidation at elevated pressures. *Combust Flame* 2013;160:995–1011.
- [111] Hong Z, Davidson DF, Hanson RK. An improved H₂/O₂ mechanism based on recent shock tube/laser absorption measurements. *Combust Flame* 2011;158:633–44.
- [112] Baulch DL, Drysdale DD, Horne DG, Lloyd AC. Evaluated kinetic data for high temperature reactions. Cleveland: CRC Press; 1973.
- [113] Tsang W. Chemical kinetic data base for combustion chemistry. Part 2. Methanol. *J Phys Chem Ref Data* 1987;16:471–508.
- [114] Herron JT. Evaluated chemical kinetic data for the reactions of atomic oxygen O (³P) with saturated organic compounds in the gas phase. *J Phys Chem Ref Data* 1988;17:967–1026.
- [115] Tsang W. Chemical kinetic data base for combustion chemistry. Part 3: propane. *J Phys Chem Ref Data* 1988;17:887–951.
- [116] Tsang W. Chemical kinetic data base for combustion chemistry part 4. Isobutane. *J Phys Chem Ref Data* 1990;19:1–68.
- [117] Tsang W. Chemical kinetic data base for combustion chemistry part V. Propene. *J Phys Chem Ref Data* 1991;20:221–73.
- [118] Baulch D, Bowman C, Cobos C, Cox R, Just T, Kerr J, et al. Evaluated kinetic data for combustion modeling: supplement II. *J Phys Chem Ref Data* 2005;34:757–1397.
- [119] Frenklach M. Transforming data into knowledge—process informatics for combustion chemistry. *Proc Combust Inst* 2007;31:125–40.
- [120] Li J, Zhao Z, Kazakov A, Chaos M, Dryer FL, Scire JJ. A comprehensive kinetic mechanism for CO, CH₂O, and CH₃OH combustion. *Int J Chem Kinet* 2007;39:109–36.
- [121] Burke MP, Dryer FL, Ju Y. Assessment of kinetic modeling for lean H₂/CH₄/O₂/diluent flames at high pressures. *Proc Combust Inst* 2011;33:905–12.
- [122] Miller JA, Pilling MJ, Troe J. Unravelling combustion mechanisms through a quantitative understanding of elementary reactions. *Proc Combust Inst* 2005;30:43–88.
- [123] Miller JA, Melius CF. Kinetic and thermodynamic issues in the formation of aromatic compounds in flames of aliphatic fuels. *Combust Flame* 1992;91:21–39.
- [124] Fernández-Ramos A, Miller JA, Klippenstein SJ, Truhlar DG. Modeling the kinetics of bimolecular reactions. *Chem Rev* 2006;106:4518–84.
- [125] Law CK, Sung CJ, Wang H, Lu T. Development of comprehensive detailed and reduced reaction mechanisms for combustion modeling. *AIAA J* 2003;41:1629–46.
- [126] Feller D, Peterson KA. Probing the limits of accuracy in electronic structure calculations: Is theory capable of results uniformly better than “chemical accuracy”? *J Chem Phys* 2007;126:114105.
- [127] Helgaker T, Ruden TA, Jørgensen P, Olsen J, Klopper W. A priori calculation of molecular properties to chemical accuracy. *J Phys Org Chem* 2004;17:913–33.
- [128] Truhlar DG, Garrett BC, Klippenstein SJ. Current status of transition-state theory. *J Phys Chem* 1996;100:12771–800.
- [129] You X, Wang H, Goos E, Sung C-J, Klippenstein SJ. Reaction kinetics of CO + HO₂ → products: ab initio transition state theory study with master equation modeling. *J Phys Chem A* 2007;111:4031–42.
- [130] Klippenstein SJ, Harding LB, Davis MJ, Tomlin AS, Skodje RT. Uncertainty driven theoretical kinetics studies for CH₃OH ignition: HO₂+CH₃OH and O₂+CH₃OH. *Proc Combust Inst* 2011;33:351–7.
- [131] Goldsmith CF, Tomlin AS, Klippenstein SJ. Uncertainty propagation in the derivation of phenomenological rate coefficients from theory: a case study of n-propyl radical oxidation. *Proc Combust Inst* 2013;34:177–85.
- [132] Frenklach M, Packard A, Seiler P, Feeley R. Collaborative data processing in developing predictive models of complex reaction systems. *Int J Chem Kinet* 2004;36:57–66.
- [133] Turányi T. Sensitivity analysis of complex kinetic systems. Tools and applications. *J Math Chem* 1990;5:203–48.
- [134] Tomlin AS. The use of global uncertainty methods for the evaluation of combustion mechanisms. *Reliab Eng Syst Safe* 2006;91:1219–31.
- [135] Tomlin A, Turányi T. Investigation and improvement of reaction mechanisms using sensitivity analysis and optimization. In: Battin-Leclerc F, Simmie JM, Blurock E, editors. *Cleaner combustion*. London: Springer; 2013. p. 411–45.
- [136] Warnatz J. Resolution of gas phase and surface combustion chemistry into elementary reactions. *Symp (Int) Combust* 1992;24:553–79.
- [137] Turányi T, Zalotai L, Dobe S, Berces T. Effect of the uncertainty of kinetic and thermodynamic data on methane flame simulation results. *Phys Chem Chem Phys* 2002;4:2568–78.
- [138] Gardiner Jr W. The pC, pR, pP, pM, and pS method for formulating the results of computer modeling studies of chemical reactions. *J Phys Chem* 1977;81:2367–71.
- [139] Stewart WE, Sørensen J. Sensitivity and regression of multicomponent reactor models. In: Proc 4th Int Symp Chem React Eng, DECHEMA. Frankfurt; 1976. p. 12–20.
- [140] Cukier R, Levine H, Shuler K. Nonlinear sensitivity analysis of multiparameter model systems. *J Phys Chem* 1977;81:2365–6.
- [141] Cukier R, Levine H, Shuler K. Nonlinear sensitivity analysis of multiparameter model systems. *J Comput Phys* 1978;26:1–42.
- [142] Cukier R, Fortuin C, Shuler KE, Petschek A, Schaibly J. Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients. I. Theory. *J Chem Phys* 1973;59:3873–8.
- [143] Schaibly JH, Shuler KE. Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients. II. Applications. *J Chem Phys* 1973;59:3879–88.
- [144] Cukier R, Schaibly J, Shuler KE. Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients. III. Analysis of the approximations. *J Chem Phys* 1975;63:1140–9.
- [145] Hwang JT, Dougherty EP, Rabitz S, Rabitz H. The Green's function method of sensitivity analysis in chemical kinetics. *J Chem Phys* 1978;69:5180–91.
- [146] Dougherty EP, Hwang JT, Rabitz H. Further developments and applications of the Green's function method of sensitivity analysis in chemical kinetics. *J Chem Phys* 1979;71:1794–808.
- [147] Rabitz H, Kramer M, Daol D. Sensitivity analysis in chemical kinetics. *Annu Rev Phys Chem* 1983;34:419–61.
- [148] Dickinson RP, Gelinas RJ. Sensitivity analysis of ordinary differential equation systems—a direct method. *J Comput Phys* 1976;21:123–43.
- [149] Coffee TP, Heimerl JM. Sensitivity analysis for premixed, laminar, steady state flames. *Combust Flame* 1983;50:323–40.
- [150] Kramer M, Calo J, Rabitz H. An improved computational method for sensitivity analysis: Green's function method with 'AIM'. *Appl Math Model* 1981;5:432–41.
- [151] Kramer MA, Rabitz H, Calo JM, Kee RJ. Sensitivity analysis in chemical kinetics: recent developments and computational comparisons. *Int J Chem Kinet* 1984;16:559–78.
- [152] Dunker AM. The decoupled direct method for calculating sensitivity coefficients in chemical kinetics. *J Chem Phys* 1984;81:2385–93.
- [153] Caracotsios M, Stewart WE. Sensitivity analysis of initial value problems with mixed ODEs and algebraic equations. *Comput Chem Eng* 1985;9:359–65.
- [154] Leis JR, Kramer MA. Sensitivity analysis of systems of differential and algebraic equations. *Comput Chem Eng* 1985;9:93–6.
- [155] Reuven Y, Smooke MD, Rabitz H. Sensitivity analysis of boundary value problems: application to nonlinear reaction-diffusion systems. *J Comput Phys* 1986;64:27–55.
- [156] Smooke M, Rabitz H, Reuven Y, Dryer F. Application of sensitivity analysis to premixed hydrogen-air flames. *Combust Sci Technol* 1988;59:295–319.
- [157] Kee R, Rupley F, Miller J. ChemKin-II: a fortran chemical kinetics package for the analysis of gas-phase chemical kinetics. Livermore, CA: Sandia National Laboratories; 1989.
- [158] Yetter R, Dryer F, Rabitz H. Some interpretive aspects of elementary sensitivity gradients in combustion kinetics modeling. *Combust Flame* 1985;59:107–33.
- [159] Frenklach M. Modeling. In: Gardiner Jr W, editor. *Combustion chemistry*. Springer; 1984. p. 423–53.
- [160] Frenklach M, Wang H, Yu C, Goldenberg M, Bowman C, Hanson R, et al. GRI-mech 1.2—an optimized detailed chemical reaction mechanism for methane combustion. GRI; 1995.
- [161] Bowman C, Hanson R, Davidson D, Gardiner Jr W, Lissianski V, Smith G, et al. GRI-Mech 2.11. 1995.
- [162] Smith GP, Golden DM, Frenklach M, Moriarty NW, Etienne B, Goldenberg M, et al. GRI-Mech 3.0. 1999.
- [163] Phenix BD, Dinaro JL, Tatang MA, Tester JW, Howard JB, McRae GJ. Incorporation of parametric uncertainty into complex kinetic mechanisms: application to hydrogen oxidation in supercritical water. *Combust Flame* 1998;112:132–46.
- [164] Reagan MT, Najm HN, Ghanem RG, Knio OM. Uncertainty quantification in reacting-flow simulations through non-intrusive spectral projection. *Combust Flame* 2003;132:545–55.
- [165] Reagan M, Najm H, Debusschere B, Le Maître O, Knio O, Ghanem R. Spectral stochastic uncertainty quantification in chemical systems. *Combust Theory Model* 2004;8:607–32.

- [166] Reagan MT, Najm HN, Pebay PP, Knio OM, Ghanem RG. Quantifying uncertainty in chemical systems modeling. *Int J Chem Kinet* 2005;37:368–82.
- [167] Tomlin AS. The role of sensitivity and uncertainty analysis in combustion modelling. *Proc Combust Inst* 2013;34:159.
- [168] Braman K, Oliver TA, Raman V. Bayesian analysis of syngas chemistry models. *Combust Theory Model* 2013;17:858–87.
- [169] Oberkampf WL, Roy CJ. Verification and validation in scientific computing. Cambridge University Press; 2010.
- [170] Schlesinger S, Crosbie RE, Gagné RE, Innis GS, Lalwani C, Loch J, et al. Terminology for model credibility. *Simulation* 1979;32:103–4.
- [171] Oberkampf WL, Trucano TG. Verification and validation benchmarks. *Nucl Eng Des* 2008;238:716–43.
- [172] Nagy T, Turányi T. Uncertainty of Arrhenius parameters. *Int J Chem Kinet* 2011;43:359–78.
- [173] Varga L, Szabó B, Zsély I, Zempléni A, Turányi T. Numerical investigation of the uncertainty of Arrhenius parameters. *J Math Chem* 2011;49:1798–809.
- [174] Nagy T, Turányi T. Determination of the uncertainty domain of the Arrhenius parameters needed for the investigation of combustion kinetic models. *Reliab Eng Syst Safe* 2012;107:29–34.
- [175] Cai L, Pitsch H. Mechanism optimization based on reaction rate rules. *Combust Flame* 2014;161:405–15.
- [176] Prager J, Najm HN, Sargsyan K, Safta C, Pitz WJ. Uncertainty quantification of reaction mechanisms accounting for correlations introduced by rate rules and fitted Arrhenius parameters. *Combust Flame* 2013;160:1583–93.
- [177] Jomaas G, Zheng XL, Zhu DL, Law CK. Experimental determination of counterflow ignition temperatures and laminar flame speeds of C_2 – C_3 hydrocarbons at atmospheric and elevated pressures. *Proc Combust Inst* 2005;30:193–200.
- [178] Sheen DA, Wang H. The method of uncertainty quantification and minimization using polynomial chaos expansions. *Combust Flame* 2011;158:2358–74.
- [179] Wiener N. The homogeneous chaos. *Am J Math* 1938;60:897–936.
- [180] Le Maître OP, Knio OM, Najm HN, Ghanem RG. A stochastic projection method for fluid flow I. Basic formulation. *J Comput Phys* 2001;173:481–511.
- [181] Xiu DB, Karniadakis GE. The Wiener-Askey polynomial chaos for stochastic differential equations. *SIAM J Sci Comput*. 2002;24:619–44.
- [182] Cameron RH, Martin WT. The orthogonal development of non-linear functionals in series of Fourier-Hermite functionals. *Ann Math* 1947;48:385–92.
- [183] Sobol' IM. On sensitivity estimation for nonlinear mathematical models. *Matematicheskoe Modelirovaniye* 1990;2:112–8.
- [184] Rabitz H, Aliş Ö. General foundations of high-dimensional model representations. *J Math Chem* 1999;25:197–233.
- [185] Li G, Wang S-W, Rabitz H. Practical approaches to construct RS-HDMR component functions. *J Phys Chem A* 2002;106:8721–33.
- [186] Li G, Rosenthal C, Rabitz H. High dimensional model representations. *J Phys Chem A* 2001;105:7765–77.
- [187] Ghanem RG, Spanos PD. Stochastic finite elements: a spectral approach. New York, NY: Springer-Verlag; 1991.
- [188] Ghanem RG, Spanos PD. Spectral stochastic finite-element formulation for reliability analysis. *J Eng Mech* 1991;117:2351–72.
- [189] Le Maître OP, Reagan MT, Najm HN, Ghanem RG, Knio OM. Stochastic projection method for fluid flow – II. Random process. *J Comput Phys* 2002;181:9–44.
- [190] Xiu DB, Lucor D, Su C-H, Karniadakis GE. Stochastic modeling of flow-structure interactions using generalized polynomial chaos. *J Fluids Eng* 2002;124:51–60.
- [191] Xiu DB, Karniadakis GE. Modeling uncertainty in flow simulations via generalized polynomial chaos. *J Comput Phys* 2003;187:137–67.
- [192] Debusschere BJ, Najm HN, Pebay PP, Knio OM, Ghanem RG, Le Maître OP. Numerical challenges in the use of polynomial chaos representations for stochastic processes. *SIAM J Sci Comput*. 2004;26:698–719.
- [193] Le Maître OP, Najm HN, Pebay PP, Ghanem RG, Knio OM. Multi-resolution-analysis scheme for uncertainty quantification in chemical systems. *SIAM J Sci Comput*. 2007;29:864–89.
- [194] Najm HN, Debusschere BJ, Marzouk YM, Widmer S, Maître OPL. Uncertainty quantification in chemical systems. *Int J Numer Meth Eng* 2009;80:786–814.
- [195] Davis SG, Mhadeshwar AB, Vlachos DG, Wang H. A new approach to response surface development for detailed gas-phase and surface reaction kinetic model optimization. *Int J Chem Kinet* 2004;36:94–106.
- [196] Rabitz H, Aliş OF, Shorter J, Shim K. Efficient input-output model representations. *Comput Phys Commun* 1999;117:11–20.
- [197] Ziehn T, Tomlin AS. A global sensitivity study of sulfur chemistry in a premixed methane flame model using HDMR. *Int J Chem Kinet* 2008;40:742–53.
- [198] Skodje RT, Tomlin AS, Klippenstein SJ, Harding LB, Davis MJ. Theoretical validation of chemical kinetic mechanisms: combustion of methanol. *J Phys Chem A* 2010;114:8286–301.
- [199] Davis MJ, Skodje RT, Tomlin AS. Global sensitivity analysis of chemical-kinetic reaction mechanisms: construction and deconstruction of the probability density function. *J Phys Chem A* 2011;115:1556–78.
- [200] Esposito G, Chelliah H. Effect of binary diffusion and chemical kinetic parameter uncertainties in simulations of premixed and non-premixed laminar hydrogen flames. *Combust Flame* 2012;159:3522–9.
- [201] Esposito G, Sarnacki B, Chelliah H. Uncertainty propagation of chemical kinetics parameters and binary diffusion coefficients in predicting extinction limits of hydrogen/oxygen/nitrogen non-premixed flames. *Combust Theory Model* 2012;16:1029–52.
- [202] Yu CL, Wang C, Frenklach M. Chemical kinetics of methyl oxidation by molecular oxygen. *J Phys Chem* 1995;99:14377–87.
- [203] Box GEP, Hunter WG, Hunter JS. Statistics for experimenters: introduction to design, data analysis, and model building. New York, NY: Wiley; 1978.
- [204] Franklin MF, Bailey RA. Selection of defining contrasts and confounded effects in two-level experiments. *Appl Stat J Roy St C* 1977;26:321–6.
- [205] Tatang MA. Direct incorporation of uncertainty in chemical and environmental engineering systems. Cambridge, MA: Massachusetts Institute of Technology; 1994.
- [206] Hoffmann VH, McRae GJ, Hungerbuhler K. Methodology for early-stage technology assessment and decision making under uncertainty: application to the selection of chemical processes. *Ind Eng Chem Res* 2004;43:4337–49.
- [207] Obrigkeit DD, Resch TJ, McRae GJ. Integrated framework for the numerical solution of multicomponent population balances. 2. The split composition distribution method. *Ind Eng Chem Res* 2004;43:4394–404.
- [208] Zhou DDY, Davis MJ, Skodje RT. Multitarget global sensitivity analysis of n-butanol combustion. *J Phys Chem A* 2013;117:3569–84.
- [209] Sobol' IM. Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Math Comput Simulat* 2001;55:271–80.
- [210] Hughes KJ, Turányi T, Clague AR, Pilling MJ. Development and testing of a comprehensive chemical mechanism for the oxidation of methane. *Int J Chem Kinet* 2001;33:513–38.
- [211] Zádor J, Zsély IG, Turányi T, Ratto M, Tarantola S, Saltelli A. Local and global uncertainty analyses of a methane flame model. *J Phys Chem A* 2005;109:9795–807.
- [212] Zádor J, Zsély IG, Turányi T. Local and global uncertainty analysis of complex chemical kinetic systems. *Reliab Eng Syst Safe* 2006;91:1232–40.
- [213] Zsély IG, Zádor J, Turányi T. Uncertainty analysis of NO production during methane combustion. *Int J Chem Kinet* 2008;40:754–68.
- [214] Sheen DA, Wang H. Combustion kinetic modeling using multispecies time histories in shock-tube oxidation of heptane. *Combust Flame* 2011;158:645–56.
- [215] Sheen DA, Rosado-Reyes CM, Tsang W. Kinetics of H atom attack on unsaturated hydrocarbons using spectral uncertainty propagation and minimization techniques. *Proc Combust Inst* 2013;34:527–36.
- [216] Davidson DF, Hong Z, Pilla GL, Farooq A, Cook RD, Hanson RK. Multi-species time-history measurements during n-heptane oxidation behind reflected shock waves. *Combust Flame* 2010;157:1899–905.
- [217] Egolfopoulos FN, Law CK. An experimental and computational study of the burning rates of ultra-lean to moderately-rich $H_2/O_2/N_2$ laminar flames with pressure variations. *Symp (Int) Combust* 1991;23:333–40.
- [218] Vagelopoulos CM, Egolfopoulos FN, Law CK. Further considerations on the determination of laminar flame speeds with the counterflow twin-flame technique. *Symp (Int) Combust* 1994;25:1341–7.
- [219] Dowdy DR, Smith DB, Taylor SC, Williams A. The use of expanding spherical flames to determine burning velocities and stretch effects in hydrogen/air mixtures. *Symp (Int) Combust* 1991;23:325–32.
- [220] Aung KT, Hassan MI, Faeth GM. Flame stretch interactions of laminar premixed hydrogen/air flames at normal temperature and pressure. *Combust Flame* 1997;109:1–24.
- [221] Aung KT, Hassan MI, Faeth GM. Effects of pressure and nitrogen dilution on flame/stretch interactions of laminar premixed $H_2/O_2/N_2$ flames. *Combust Flame* 1998;112:1–15.
- [222] Kwon OC, Faeth GM. Flame/stretch interactions of premixed hydrogen-fueled flames: measurements and predictions. *Combust Flame* 2001;124:590–610.
- [223] Karpov VP, Lipatnikov AN, Wolanski P. Finding the Markstein number using the measurements of expanding spherical laminar flames. *Combust Flame* 1997;109:436–48.
- [224] Sirjean B, Dames E, Sheen DA, You X, Sung CJ, Holley AT, et al. A high-temperature chemical kinetic model of *n*-alkane oxidation, JetSurF version 1.0. 2009.
- [225] Burke MP, Klippenstein SJ, Harding LB. A quantitative explanation for the apparent anomalous temperature dependence of $OH + HO_2 \rightleftharpoons H_2O + O_2$ through multi-scale modeling. *Proc Combust Inst* 2013;34:547–55.
- [226] Beck JL, Katafygiotis LS. Updating models and their uncertainties. I: Bayesian statistical framework. *J Eng Mech* 1998;124:455–61.
- [227] Katafygiotis LS, Beck JL. Updating models and their uncertainties. II: model identifiability. *J Eng Mech* 1998;124:463–7.
- [228] Vanik MW, Beck JL, Au SK. Bayesian probabilistic approach to structural health monitoring. *J Eng Mech* 2000;126:738–45.
- [229] Beck JL, Au S-K. Bayesian updating of structural models and reliability using Markov chain Monte Carlo simulation. *J Eng Mech* 2002;128:380–91.
- [230] Beck JL, Yuen K-V. Model selection using response measurements: Bayesian probabilistic approach. *J Eng Mech* 2004;130:192–203.
- [231] Cheung SH, Beck JL. Bayesian model updating using hybrid Monte Carlo simulation with application to structural dynamic models with many uncertain parameters. *J Eng Mech* 2009;135:243–55.
- [232] Cheung SH, Beck JL. Calculation of posterior probabilities for Bayesian model class assessment and averaging from posterior samples based on dynamic system data. *Comput-Aided Civ Inf* 2010;25:304–21.

- [233] Vikhansky A, Kraft M, Simon M, Schmidt S, Bart HJ. Droplets population balance in a rotating disc contactor: an inverse problem approach. *AIChE J* 2006;52:1441–50.
- [234] Braumann A, Kraft M. Incorporating experimental uncertainties into multi-variate granulation modelling. *Chem Eng Sci* 2010;65:1088–100.
- [235] Braumann A, Kraft M, Mort PR. Parameter estimation in a multidimensional granulation model. *Powder Technol* 2010;197:196–210.
- [236] Braumann A, Man PLW, Kraft M. Statistical approximation of the inverse problem in multivariate population balance modeling. *Ind Eng Chem Res* 2010;49:428–38.
- [237] Man PLW, Braumann A, Kraft M. Resolving conflicting parameter estimates in multivariate population balance models. *Chem Eng Sci* 2010;65:4038–45.
- [238] Smallbone AJ, Bhave A, Kraft M, Dris A, McDavid R. Moving toward establishing more robust and systematic model development for IC engines using process informatics. *SAE Technical Paper Series*. 2010. 2010-01-0572.
- [239] Braumann A, Man PLW, Kraft M. The inverse problem in granulation modeling—two different statistical approaches. *AIChE J* 2011;57:3105–21.
- [240] Sander M, Patterson RIA, Braumann A, Raj A, Kraft M. Developing the PAH-PP soot particle model using process informatics and uncertainty propagation. *Proc Combust Inst* 2011;33:675–83.
- [241] Kastner CA, Braumann A, Man PLW, Mosbach S, Brownbridge GPE, Akroyd J, et al. Bayesian parameter estimation for a jet-milling model using Metropolis–Hastings and Wang–Landau sampling. *Chem Eng Sci* 2013;89:244–57.
- [242] Mosbach S, Hong JH, Brownbridge GPE, Kraft M, Gudiyella S, Brezinsky K. Bayesian error propagation for a kinetic model of n-propylbenzene oxidation in a shock tube. *Int J Chem Kinet* 2014;46:389–404.
- [243] Panesi M, Miki K, Prudhomme S, Brandis A. On the assessment of a Bayesian validation methodology for data reduction models relevant to shock tube experiments. *Comput Method Appl M* 2012;213–216:383–98.
- [244] Miki K, Prudencio EE, Cheung SH, Terejanu G. Using Bayesian analysis to quantify uncertainties in the $\text{H}+\text{O}_2 \rightarrow \text{OH}+\text{O}$ reaction. *Combust Flame* 2013;160:861–9.
- [245] Turányi T, Nagy T, Zsély IG, Cserháti M, Varga T, Szabó BT, et al. Determination of rate parameters based on both direct and indirect measurements. *Int J Chem Kinet* 2012;44:284–302.
- [246] Zsély IG, Varga T, Nagy T, Cserháti M, Turányi T, Peukert S, et al. Determination of rate parameters of cyclohexane and 1-hexene decomposition reactions. *Energy* 2012;43:85–93.
- [247] Varga T, Zsély IG, Turányi T, Bentz T, Olzmann M. Kinetic analysis of ethyl iodide pyrolysis based on shock tube measurements. *Int J Chem Kinet* 2014;46:295–304.
- [248] Varga T, Nagy T, Olm C, Zsély IG, Pálvölgyi R, Valkó É, et al. Optimization of a hydrogen combustion mechanism using both direct and indirect measurements. *Proc Combust Inst* 2014. <http://dx.doi.org/10.1016/j.proci.2014.06.071>.
- [249] Prager J, Najm HN, Zádor J. Uncertainty quantification in the ab initio rate-coefficient calculation for the $\text{CH}_3\text{CH(OH)}\text{CH}_3 + \text{OH} \rightarrow \text{CH}_3\text{C(OH)}\text{CH}_3 + \text{H}_2\text{O}$ reaction. *Proc Combust Inst* 2013;34:583–90.
- [250] Feeley R, Seiler P, Packard A, Frenklach M. Consistency of a reaction dataset. *J Phys Chem A* 2004;108:9573–83.
- [251] Feeley R, Frenklach M, Onsum M, Russi T, Arkin A, Packard A. Model discrimination using data collaboration. *J Phys Chem A* 2006;110:6803–13.
- [252] Seiler P, Frenklach M, Packard A, Feeley R. Numerical approaches for collaborative data processing. *Optim Eng* 2006;7:459–78.
- [253] Russi T, Packard A, Feeley R, Frenklach M. Sensitivity analysis of uncertainty in model prediction. *J Phys Chem A* 2008;112:2579–88.
- [254] Hastings WK. Monte-Carlo sampling methods using Markov chains and their applications. *Biometrika* 1970;57:97.
- [255] Ranzi E, Faravelli T, Gaffuri P, Sogaro A. Low-temperature combustion: automatic generation of primary oxidation reactions and lumping procedures. *Combust Flame* 1995;102:179–92.
- [256] Ranzi E, Faravelli T, Gaffuri P, Garavaglia E, Goldaniga A. Primary pyrolysis and oxidation reactions of linear and branched alkanes. *Ind Eng Chem Res* 1997;36:3336–44.
- [257] Warth V, Battin-Leclerc F, Fournet R, Glaude P-A, Corme G-M, Scacchi G. Computer based generation of reaction mechanisms for gas-phase oxidation. *Comput Chem* 2000;24:541–60.
- [258] Pierucci S, Ranzi E. A review of features in current automatic generation software for hydrocarbon oxidation mechanisms. *Comput Chem Eng* 2008;32:805–26.
- [259] Allen J, Ashcraft R, Beran G, Goldsmith C, Harper M, Jalan A, et al. RMG (reaction mechanism generator) version 3.2. 2010. p. 2010.
- [260] Van Geem KM, Reyniers MF, Marin GB, Song J, Green WH, Matheu DM. Automatic reaction network generation using RMG for steam cracking of n-hexane. *AIChE J* 2006;52:718–30.
- [261] Harper MR, Van Geem KM, Pyl SP, Marin GB, Green WH. Comprehensive reaction mechanism for n-butanol pyrolysis and combustion. *Combust Flame* 2011;158:16–41.
- [262] Metcalfe WK, Burke SM, Ahmed SS, Curran HJ. A hierarchical and comparative kinetic modeling study of C_1 – C_2 hydrocarbon and oxygenated fuels. *Int J Chem Kinet* 2013;45:638–75.
- [263] Metcalfe WK, Burke SM, Curran HJ. Chemical kinetic mechanism for C_1 to C_3 hydrocarbons, version c3_54.1. 2013.
- [264] Chang A, Davidson D, DiRosa M, Hanson R, Bowman C. Shock tube experiments for development and validation of kinetic models of hydrocarbon oxidation. In: 25th Symposium (International) on Combustion. Irvine, California: Poster 3–23; 1994.
- [265] Urzay J, Kseib N, Davidson DF, Iaccarino G, Hanson RK. Uncertainty-quantification analysis of the effects of residual impurities on hydrogen–oxygen ignition in shock tubes. *Combust Flame* 2014;161:1–15.
- [266] Sheen DA, Wang H. On combustion reaction kinetic model hierarchy. *J Phys Chem* 2014. manuscript in preparation.
- [267] Lu T, Law CK. Toward accommodating realistic fuel chemistry in large-scale computations. *Progr Energy Combust Sci* 2009;35:192–215.
- [268] Pepiot-Desjardins P, Pitsch H. An efficient error-propagation-based reduction method for large chemical kinetic mechanisms. *Combust Flame* 2008;154:67–81.
- [269] Lu T, Law CK. A directed relation graph method for mechanism reduction. *Proc Combust Inst* 2005;30:1333–41.
- [270] Lu T, Law CK. On the applicability of directed relation graphs to the reduction of reaction mechanisms. *Combust Flame* 2006;146:472–83.
- [271] Xin YX, Sheen DA, Wang H, Law CK. Skeletal reaction model generation, uncertainty quantification and minimization: combustion of butane. *Combust Flame* 2014;161:3031–9.
- [272] Mueller ME, Iaccarino G, Pitsch H. Chemical kinetic uncertainty quantification for Large Eddy Simulation of turbulent nonpremixed combustion. *Proc Combust Inst* 2013;34:1299–306.
- [273] Barlow RS, Frank JH. Effects of turbulence on species mass fractions in methane/air jet flames. *Symp (Int) Combust* 1998;27:1087–95.
- [274] Peters N. Laminar diffusion flamelet models in non-premixed turbulent combustion. *Progr Energy Combust Sci* 1984;10:319–39.