Combustion and Flame 190 (2018) 284–292

Contents lists available at ScienceDirect

Combustion and Flame

journal homepage: www.elsevier.com/locate/combustflame

Assessing impacts of discrepancies in model parameters on autoignition model performance: A case study using butanol

Sai Krishna Sirumalla[a], Morgan A. Mayer[b], Kyle E. Niemeyer[b], Richard H. West[a,*]

[a] Department of Chemical Engineering, Northeastern University, Boston, MA 02115, USA
[b] School of Mechanical, Industrial, and Manufacturing Engineering Oregon State University, Corvallis, OR 97331, USA

ARTICLE INFO

Article history:
Received 6 September 2017
Revised 11 October 2017
Accepted 20 November 2017

Keywords:
Butanol
Chemical kinetic models
Model comparison
Uncertainty

ABSTRACT

Side-by-side comparison of detailed kinetic models using a new tool to aid recognition of species structures reveals significant discrepancies in the published rates of many reactions and thermochemistry of many species. We present a first automated assessment of the impact of these varying parameters on observable quantities of interest—in this case, autoignition delay—using literature experimental data. A recent kinetic model for the isomers of butanol was imported into a common database. Individual reaction rate and thermodynamic parameters of species were varied using values encountered in combustion models from recent literature. The effects of over 1600 alternative parameters were considered. Separately, experimental data were collected from recent publications and converted into the standard YAML-based ChemKED format. The Cantera-based model validation tool, PyTeCK, was used to automatically simulate autoignition using the generated models and experimental data, to judge the performance of the models. Taken individually, most of the parameter substitutions have little effect on the overall model performance, although a handful have quite large effects, and are investigated more thoroughly. Additionally, models varying multiple parameters simultaneously were evolved using a genetic algorithm to give fastest and slowest autoignition delay times, showing that changes exceeding a factor of 10 in ignition delay time are possible by cherry-picking from only accepted, published parameters. All data and software used in this study are available openly.

1. Introduction

Detailed kinetic models over a range of temperatures and pressures are essential for predicting the behavior of new fuels. Kinetic combustion models of complicated fuels contain thousands of species and elementary reactions which are described by thermodynamic and rate parameters. Many of these parameters are calculated with semi-empirical methods, estimated, sometimes just guessed, and quite often changed or "tweaked" to alter some global observable. This leads to discrepancies in rates and thermodynamic parameters for the same reaction or species in different models. The work presented aims to determine how these discrepancies affect the performance of a model.

Side-by-side comparison of detailed kinetic models reveals significant discrepancies in the published rates of many reactions and thermochemistry of many species. For example, in the supplementary data of the 2016 Combustion Symposium proceedings, of 2600 reactions we have identified in two or more models, 15% disagree by over an order of magnitude at 1000 K, and some by 31 orders of magnitude; of the species we found in two or more models, 4% of standard enthalpy of formation values span more than 50 kJ/mols. Chen et al. [1] recently used an automated tool to show that many published models have rate coefficients exceeding the collision limit by several orders magnitude. However, the impact of these variations on observable quantities of interest—such as autoignition delay—has not yet been assessed. Each published model has usually been "validated" with and often trained, optimized, or tweaked to match a given set of experimental data. Many reaction rates have been chosen only as part of a whole model and only to match a limited set of experimental data, although they are then frequently used in other models.

Pioneering work by Frenklach et al. [2] advanced the systematic treatment of kinetic parameter uncertainty in combustion modeling. Other notable contributions include those by Wang and Sheen [3], Turányi et al. [4], and Tomlin and Turányi [5], whose reviews, books, and chapters provide a thorough and clear overview of local and global uncertainty analysis in this field.

Recent advances include treatment of correlations between uncertain parameters derived from a common rate rule [6] and the use of multi-scale informatics [7] to propagate uncertainties

* Corresponding author.
E-mail address: r.west@northeastern.edu (R.H. West).

https://doi.org/10.1016/j.combustflame.2017.11.018
0010-2180/

from physically meaningful molecular properties rather than reaction pre-exponential factors. Many approaches involve Monte Carlo sampling within a range of uncertainties, attributed to every parameter by hand or according to some heuristics. However, the systematic assessment of how much uncertainty could be due to discrepancies between parameters in published models has not been attempted, not because the mathematics are complicated but because the data are scattered and hard to reconcile into a common platform. Because species are given different names in different models, it can be hard to find the discrepancies.

In this work we use butanol as a case study. Bio-butanol is a potential renewable biofuel, offering several advantages over bio-ethanol: its higher heating value allows a higher blending rate in gasoline; its lower latent heat of vaporization reduces issues associated with combustion cold start [8]; it is less corrosive, has a higher cetane number, and lower vapor pressure; and it has a similar viscosity to diesel. Butanol research is still of interest to the combustion kinetics community, although not so novel as to be without data for comparison. As well as a popular validated model from Lawrence Livermore National Laboratories by Sarathy et al. [9], upon which we base our investigation, there are plenty of experimental data [10–13]. Agbro and Tomlin [14] recently investigated the Sarathy et al. model used in this work by conducting both local and global uncertainty and sensitivity methods for predicting autoignition delay times and species profiles.

2. Methods

The overall workflow is to take an original model (the LLNL butanol model [9]) in CHEMKIN format, and for the rate of every reaction rate and the thermochemistry of every species, search to see if an alternative has been used in any other recently published kinetic models. This gives a large set of alternative parameters, each of which has been independently "validated," "approved," or at least shared with the community. In one analysis, we consider each variation independently, and measure its impact on the model performance as judged against a broad set of experimental data (475 datapoints across 67 datasets from four papers [10–13]). This allows us to rank the parameters about which there exist disagreement, in order of importance for ignition delay predictions. In a second analysis, we allow many parameters to be varied simultaneously using a genetic algorithm to explore the extrema—fastest and slowest ignition delays—that can be achieved by selecting from among the published alternative parameters.

2.1. Kinetic model curation

The major barrier to comparing published kinetic models is the lack of canonical or even conventional methods of naming the chemical species, combined with the persistent use of a CHEMKIN-compatible data structure designed not to preserve chemical metadata but rather to fit on an 80-column punch card. This has led to many alternative names being used for each species. For example, prop-1-en-2-yl has been referred to in published models as `C*C.C`, `tC3H5`, `CH2CCH3`, `propen2yl`, `ch3cch2`, `T-C3H5`, `CH3CCH2`, `TC3H5`, `C3H5-T`, and `c3h5-t`, making it hard to find and compare all the rates of its reactions. We have developed a tool [15] that helps with this task of identifying the species in a detailed kinetic model [16]. The tool was built using methods from the open-source Reaction Mechanism Generator (RMG) [17] which predicts how identified molecules are expected to react. Comparing this with how the CHEMKIN file says species reacts allows one to deduce which molecule corresponds to which species name.

We have used this tool to partially or fully import 74 detailed kinetic models gathered from the literature [18]. The 74 models include 20 from *Combustion and Flame* (2012–2015), 33 from *Proceedings of the Combustion Institute* (2013–2017), and 21 models from other miscellaneous articles, provided by their authors, or downloaded from repository websites such as AramcoMech, USC-Mech, LLNL.gov, etc. The full list is provided in the supplementary materials, and the models can be downloaded openly [18].

2.2. Alternate model generation

The starting point for model generation was the LLNL comprehensive combustion model for the four butanol isomers by Sarathy et al. [9], which contains 418 species and ~2343 reactions. (Counting the number of reactions in a model is not without complications. The CHEMKIN-to-Cantera conversion script skips reactions with $k \equiv 0$, and converts explicit reverse reaction rates into two irreversible reactions in opposite directions. Furthermore, RMG treats duplicates as a single reaction whose rate is given as the sum of multiple Arrhenius expressions, rather than as independent reactions.)

The large database of kinetic models [18] was first filtered for only reactions containing at least one of the 418 species in the original model, resulting in 55,775 instances of 13,618 rate expressions for 6303 reactions occurring in 74 models. To reduce the risk of introducing errors, kinetics that are represented as the sum of two Arrhenius expressions (i.e., duplicate reactions) were excluded from the analysis. This left 55,058 instances of 13,245 unique rates for 6253 reactions. Many of these reactions are the reverses of each other, but for the current analysis they were not merged or reversed, i.e., rates were only compared across models if the reactions were written in the same direction. However, pressure-dependent and -independent reactions (e.g., A + B ⇌ C and A + B (+M) ⇌ C (+M)) were treated as alternative rate expressions for the same reaction.

For each reaction in the original model, the most common three rates occurring in the database were considered, but only if they appeared in at least two models. For example, the reaction $C_4H_3 - n + OH \rightleftharpoons C_4H_2 + H_2O$ has a rate coefficient of $k = 2 \times 10^{12}$ cm^3/mole/s in the original model and 22 others, but is 2.5×10^{12} cm^3/mole/s in three models, and 1.5×10^{13} cm^3/mole/s in two models. We also found the rate coefficient 3.0×10^{13} cm^3/mole/s in use, but this is the fourth-most common rate and only present in one model, so we excluded it from the current analysis. All of these rates occur in models published in *Proceedings of the Combustion Institute* or *Combustion and Flame* since 2013.

We implemented the requirement that a rate expression has been seen "in the wild" at least twice to reduce the risk of possible errors made when importing the models, and to result in a reasonably conservative estimate of the impact of genuine discrepancies between "accepted" parameters, without being overly influenced by lone outliers. It should be noted, however, that a parameter appearing in many models does not indicate that it is correct. This is not a popularity contest, and often the most accurately determined parameter is not the most commonly used. Furthermore, a complete lack of discrepancy in the literature does not indicate a lack of uncertainty; often an uncertain estimate is adopted universally.

In total there were 300 reactions with one alternative rate considered (besides the original), 471 with two alternatives, and 13 with three alternatives, totaling 1281 kinetic variants on the original model. A similar process was undertaken for the thermochemical parameters. These are provided in the CHEMKIN files in NASA polynomial form describing ΔH, S, and $C_P(T)$. When an alternative is considered, the full set of polynomials for that species are substituted. Out of the 418 total species in the model, 65 species had one set of alternative thermodynamic parameters found in the database, 127 species had two alternatives, and 2 species had three