Comparison of the performance of several recent syngas combustion mechanisms

C. Olm[1,2,3], I. Gy. Zsély[1], T. Varga[1,2], H. J. Curran[4], T. Turányi[1, *]

[1] Institute of Chemistry, Eötvös University (ELTE), Budapest, Hungary

[2] MTA-ELTE Research Group on Complex Chemical Systems, Budapest, Hungary

[3] Chair of Numerical Thermo-Fluid Dynamics, Technical University Bergakademie Freiberg, Germany

[4] Combustion Chemistry Centre, National University of Ireland, Galway (NUIG), Ireland

Keywords: syngas combustion; detailed mechanisms; mechanism testing; mechanism development

Abstract

A large set of experimental data was accumulated for syngas combustion: ignition studies in shock tubes (732 data points in 62 datasets) and in rapid compression machines (492/47), flame velocity determinations (2116/217) and species concentration measurements from flow reactors (1104/58), shock tubes (436/21) and jet-stirred reactors (90/3). In total, 4970 data points in 408 datasets from 52 publications were collected covering wide ranges of temperature T, pressure p, equivalence ratio φ, CO/H_2 ratio and diluent concentration X_{dil}. 16 recent syngas combustion mechanisms were tested against these experimental data, and the dependence of their predictions on the types of experiment and the experimental conditions was investigated. Several clear trends were found. Ignition delay times measured in rapid compression machines (RCM) and in shock tubes (ST) at temperatures below 1000 K could not be well-predicted. Particularly for shock tubes, facility effects at temperatures below 1000 K could not be excluded. The accuracy of the reproduction of ignition delay times did not change significantly with pressure. The agreement of measured and simulated laminar flame velocities is better at low initial (i.e. cold side) temperatures, at fuel-lean conditions, for CO-rich and highly diluted mixtures. The reproduction of the experimental flame velocities is better when these were measured using the heat flux method or the counterflow twin-flame technique, compared to the flame cone method and the outwardly propagating spherical flame approach. With respect to all data used in this comparison, five mechanisms were identified that reproduce the experimental data similarly well. These are the NUIG-NGM-2010, Kéromnès-2013, Davis-2005, Li-2007 and USC-II-2007 mechanisms, in decreasing order of their overall performance. The influence of poorly reproduced experiments and weighting on the performance of the mechanisms was investigated. Furthermore, an analysis of local sensitivity coefficients was carried out to determine the

influence of selected reactions at the given experimental conditions and to identify those reactions that require more attention in future development of syngas combustion models.

1. Introduction

In recent years, there has been an increased interest in studying the combustion of fuel mixtures consisting of carbon monoxide and hydrogen, referred to as syngas or "wet CO". These fuels can be produced from coal and biomass via gasification, and are considered to be a promising option towards cleaner combustion technologies for power generation [1]. The chemistry of syngas combustion forms the basis of the combustion of hydrocarbons and oxygenates, and has been the subject of many experimental and modeling studies for decades. Several new syngas combustion mechanisms have been published in the last ten years. In these publications, the agreement between the measurements and the simulations is typically characterized by plots, in which the experimental data and the simulation results are depicted together. However, quantitative agreement of a large number of simulation results with the corresponding experimental data has not been investigated. A quantitative evaluation enables the modeler to distinguish experiments that are well reproduced by simulations from those that are insufficiently described, which may have implications for model developers in the choice of development targets as well as for experimentalists in the design of new experiments. Furthermore, such a procedure allows for the detection of strengths and weaknesses of the mechanisms in certain ranges of operating conditions. Knowledge about the specific behavior of a mechanism is a necessary first step in attempts to reduce modeling uncertainties during mechanism development and optimization.

A similar comparison of reaction mechanisms for hydrogen combustion based on the quantitative characterization of the reproduction of experimental data was recently elaborated [2]. The present work has several novelties compared to this previous paper and other publications on the investigation of syngas combustion. The comparison performed here, on a very comprehensive set of experiments, is much wider ranging than those used in the previous investigations; various measurement types (ignition delay time, species concentration profile and flame velocity) and experimental techniques (e.g. shock tube and RCM experiments) are included in the analysis. The performance of 16 syngas combustion mechanisms is compared in detail, and the conclusions drawn are supported by reproducible numbers. Furthermore, the influence of data weighting to reduce the effect of multiple repeated measurements is described.

2. Methodology

In this work the agreement of experimental and simulation results is investigated using the following objective function:

$$E_{\mathrm{i}} = \frac{1}{N_{\mathrm{i}}} \sum_{\mathrm{j=1}}^{N_{\mathrm{i}}} \left(\frac{Y_{\mathrm{ij}}^{\mathrm{sim}} - Y_{\mathrm{ij}}^{\mathrm{exp}}}{\sigma(Y_{\mathrm{ij}}^{\mathrm{exp}})} \right)^2$$

$$E = \frac{1}{N} \sum_{\mathrm{i=1}}^{N} E_{\mathrm{i}}$$

where

$$Y_{\mathrm{ij}} = \begin{cases} y_{ij} & \text{if } \sigma(y_{ij}^{\mathrm{exp}}) \approx \text{constant} \\ \ln y_{ij} & \text{if } \sigma(\ln y_{\mathrm{ij}}^{\mathrm{exp}}) \approx \text{constant} \end{cases}$$

Here N is the number of datasets and N_{i} is the number of data points in the i-th dataset. Values $y_{\mathrm{ij}}^{\mathrm{exp}}$ and $\sigma\left(y_{\mathrm{ij}}^{\mathrm{exp}}\right)$ are the j-th data point and its standard deviation, respectively, in the i-th dataset. The corresponding simulated (modeled) value is $Y_{\mathrm{ij}}^{\mathrm{sim}}$ obtained from a simulation using an appropriate detailed mechanism and simulation method. If a measured value is characterized by absolute errors (the scatter is independent of the magnitude of y_{ij}), then $Y_{\mathrm{ij}} = y_{\mathrm{ij}}$. We used this option for laminar flame velocities and measured concentrations. If the experimental results are described by relative errors (the scatter is proportional to the value of y_{ij}), then we used the option $Y_{\mathrm{ij}} = \ln(y_{\mathrm{ij}})$, which is characteristic for ignition time measurements. Error function values E_{i} and E are expected to be near unity if the chemical kinetic model is accurate, and deviations of the measured and simulated results are caused by the scatter of the experimental data only. Note that due to the squaring in the definition of E, a twice as high deviation of the simulated and experimental values of one mechanism in comparison to another leads to a four times higher value of E. This objective function has been used in our previous studies on the comparison of reaction mechanisms [2] and the estimation of rate parameters from experimental data [3-6].

In addition to the average error function E, the average absolute deviation D was used to characterize the behavior of the mechanisms:

$$D = \frac{1}{N} \sum_{\mathrm{i=1}}^{N} \frac{1}{N_{\mathrm{i}}} \sum_{\mathrm{j=1}}^{N_{\mathrm{i}}} \frac{D_{\mathrm{ij}}}{\sigma(Y_{\mathrm{ij}}^{\mathrm{exp}})} = \frac{1}{N} \sum_{\mathrm{i=1}}^{N} \frac{1}{N_{\mathrm{i}}} \sum_{\mathrm{j=1}}^{N_{\mathrm{i}}} \frac{(Y_{\mathrm{ij}}^{\mathrm{sim}} - Y_{\mathrm{ij}}^{\mathrm{exp}})}{\sigma(Y_{\mathrm{ij}}^{\mathrm{exp}})}$$

with the absolute deviation belonging to an individual data point D_{ij} and using the same transformation $y_{\mathrm{ij}} \rightarrow Y_{\mathrm{ij}}$ as given above. In contrast to E, the sign of the difference $Y_{\mathrm{ij}}^{\mathrm{sim}} - Y_{\mathrm{ij}}^{\mathrm{exp}}$ is

maintained in the definition of D. In contrast to our previous work [2], we modified the above equation by dividing D_{ij} by the estimated standard deviation. This allows for a better comparison of data obtained from different types of experiments with respect to their D values. The drawback of the D value is that positive and negative deviations in different data points can cancel each other out and the resulting averaged value would suggest an un-realistically good overall agreement. However, trends such as systematic under- or over-prediction are thereby captured in the D_{ij} values. The D values are displayed in Fig. 11 and in Figs. S2–S5 and S7–S12 of the Supplementary Material. These D value plots may deliver a better understanding of the trends associated with changes of certain operating conditions and should be interpreted alongside their corresponding E value plots.

It is possible to characterize the similarity of simulation results using different mechanisms by calculating Pearson correlation coefficients based on the values of D_{ij}. Similar to the definitions of E and D, correlation coefficients C are calculated for each dataset and then averaged over all of the N datasets. Details of the calculation of C values as well as a brief discussion in the context of comparisons of the performance of mechanisms can be found in [2].

3. The investigated mechanisms

Our aim was to test all major syngas combustion mechanisms that were published in the last decade. Furthermore, GRI-Mech 3.0 [7] was added to the comparison, which was published in 1999 and primarily developed for methane combustion, but is nevertheless widely used in the syngas-related literature. In the forthcoming discussions, an identifier of each mechanism is used, which combines the name of the publishing author(s) or research group and the year of publication.

Earlier mechanisms from the same research group were tested only if they were conceptually different from the latest one. For instance, two mechanisms published by the Galway group, NUIG-NGM-2010 and Kéromnès-2013, were used because the latter features a sub-model for the reactions of the excited OH radical (OH*) largely based on the work of Tamura et al. [8] and updated by Kathrotia et al. [9]. It can be shown that these reactions are of high importance for a more accurate reproduction of shock tube ignition studies at high temperatures. In a similar way, both the SaxenaWilliams-2006 and the SanDiego-2014 mechanisms were included in the comparison. In our previous paper [2] we used the 2011 version of the San Diego mechanism. A recently published update of several reaction rates, fall-off parameters and third body collision

coefficients in the hydrogen chemistry [10], which forms the basis of the new 2014 version, led to a substantially better description of hydrogen ignition delay times, at the cost of less accuracy at the conditions in flames. The same trends apply to syngas combustion, for which hydrogen chemistry is known to be very important.

Several of these mechanisms were originally developed for syngas combustion [11-19], while other mechanisms were elaborated for the combustion of hydrocarbons or oxygenates [7, 20-25], but also used to interpret syngas data. MECHMOD [26] was used to remove unnecessary species and reactions from the mechanisms (e.g. nitrogen chemistry, C_2 and above). Table 1 contains the list of these mechanisms and provides further information about size and included diluents. The numbering of the mechanisms in Table 1 is according to their overall performance from the best (1) to the worst (15). The mechanism of Dagaut-2003, which could not be tested for the complete set of experimental data, was given the number 16.

All mechanisms can handle N_2 bath gas, while only some mechanisms include He as a species. Unlike all other mechanisms, that of Dagaut-2003 [20] does not contain Ar either. This affects the mechanism comparison using many data subsets. As a general rule, we used all experiments except for those in He bath gas as the baseline for all comparisons, and indicated the Ar-free results of Dagaut-2003 wherever possible. In addition to this, flames measured in He-containing mixtures were added to the respective comparison figures. An overall comparison including these additional He data can be found in Fig. S13 of the Supplementary Material.

Thermochemical and transport data were used as published online and/or provided by the authors. The Rasmussen-2008 mechanism [15] was not published with transport data as it was only used for shock tube and flow reactor simulations. However, the transport coefficients of all species involved in syngas combustion were identical in all mechanisms. For this reason, the same transport data as in all other mechanisms were used for Rasmussen-2008 for testing purposes.

4. Collection of experimental data

A large set of experimental data was collected in which the combustion of syngas mixtures was investigated. These types of measurements, called indirect measurements or bulk measurements, are generally used to test detailed reaction mechanisms. We utilized all measurements that were employed for testing the recent mechanisms of Davis et al. [11], Sun et al. [14], Li et al. [22] and Kéromnès et al. [18]. References for measurements were collected from these recent review

articles, and the experimental data were digitized from the original publications. Furthermore, a comprehensive search was carried out to find all other measurements that can be used to test syngas combustion mechanisms. In many practical applications, "syngas" can contain CH_4 or other hydrocarbons, and some authors used syngas to name these more complex fuel mixtures (e.g. in [27]) in their experiments. In the present work, only fuel mixtures of H_2 and CO are considered, as well as pure CO and pure H_2 diluted with CO_2.

An overview of the regions of operating conditions covered and diluents investigated is shown in Table 2. Although each condition among the described ranges is not equally well represented in our experimental database, the numbers provided may serve as an outline for the range of validity of the presented comparison. The detailed list of the data, arranged in datasets, with references is provided as Supplementary Material, Tables A–G. A "dataset" contains those data points that were measured on the same apparatus at the same time at similar conditions except for one that was systematically changed. There are various reasons why the number of utilized and total number of data points in a dataset can differ, e.g. due to temperature constraints on shock tube and JSR data or if single data points could not be modeled or were largely inconsistent with other measurements. These issues are discussed in Chapter 5 and in the Supplementary Material.

The complete set of data include ignition measurements in shock tubes (732 total and 731 utilized data points in 62 datasets from 7 original publications) and RCMs (492/387/40/3), flame velocity measurements (2116/1963/202/31), concentration–time profiles measured in flow reactors (979/508/35/4), outlet concentration profiles from flow reactors (125/125/9/2), shock tubes (436/436/7/1) and JSRs (90/90/9/1).

All relevant experimental conditions and results were encoded into an extended (and fully backwards compatible) version of the PrIMe file format [28], an XML scheme used for the systematic storage of various kinds of combustion experiments. These stored XML data provide all information required for the simulation of the experiments and the calculation of properties observed or derived from experiments (e.g. the ignition delay as defined in the corresponding experimental publication). A MATLAB code called Optima was written [2] which allows for automatic CHEMKIN-II [29] simulations and error evaluations for a specific mechanism on the full data set. The MATLAB code starts the corresponding simulation code of the CHEMKIN-II package (SENKIN [30], PREMIX [31] or PSR [32]), collects and evaluates the simulation results. In principle, the complete investigation of a mechanism against several thousand experimental data can be carried out in a single run.

5. Simulation of experiments

The relationship between the types of experiments, modeling approaches and computational codes to solve the respective problem is shown in Fig. S1 of the Supplementary Material.

5.1 Shock tube simulations

Shock tube experiments were simulated using the SENKIN program [30] of the CHEMKIN-II package. Ignition delay times were extracted from the simulated pressure or concentration profiles, and interpreted as described in the respective publications, e.g. based on the maximum slope of the pressure profile. If an ignition criterion could not be modeled, another similar criterion was chosen. For example, the maximum concentration of the excited radical OH* as measured in [33, 34] cannot be modeled with mechanisms that do not contain this species (i.e. all mechanisms except Kéromnès-2013). The criterion "OH maximum" had to be used instead, which has shown to give similarly good predictions except at very high temperatures behind the reflected shock wave.

Low-temperature shock tube experiments, where the ignition delays are in the order of milliseconds, should be handled with special care by considering the possibility of a pressure change during the induction period [35]. At such reaction times the pressure behind the reflected shock wave increases with time [36]. In both articles it is demonstrated that by taking into account this facility effect, the overall description of the experiments by the models improves. It has also been discussed by Dryer and Chaos [37, 38] that at these conditions the measured ignition delays are extremely sensitive to impurities in the mixture. Impurities in the reactant mixture can cause systematic over-predictions of measured ignition delay times by the simulations (i.e. earlier ignition events). Furthermore, they have shown that by assuming a catalytic conversion between H_2O_2 and OH, the description of the experiments can also be improved [38].

In all cases except for [39], pressure–time histories were not reported, which is why constant volume and adiabatic conditions were assumed. Hence, the above described facility effects could not be taken into account. Alternatively, data points where long ignition delay times (above 1 – 2 ms) were measured and/or data points at low temperatures can be excluded. The latter option was chosen in the present study, with a threshold value of $T \leq 1000$ K. For the simulation of the shock tube data of Thi et al. [39], volume–time histories (VTIM) were provided by the authors and used here.

5.2 RCM simulations

RCM experiments were simulated using the VTIM option of SENKIN [30], to account for the effects of compression and heat loss. Volume–time histories were calculated from the experimental pressure profiles provided by the authors [18, 40, 41] assuming an adiabatic reaction core [42]. Different simulation methods were discussed by Mansfield and Wooldridge [41]. In their work, the authors concluded that the assignment of a single average thermodynamic state to each experiment and a constant volume assumption is similarly accurate than the use of volume–time histories. However, for better comparison with other RCM measurements, volume–time histories calculated from exponential fits to the previously smoothened pressure–time profiles were used here. The measurements of Mittal et al. [40] were published together with volume–time histories expressed in polynomial form. The simulations were performed using these volume–time histories. Some of their experiments lacked published heat loss data and these were excluded from the comparison.

5.3 Laminar flame velocity simulations

Laminar flames were simulated using the PREMIX code [31] of the CHEMKIN-II package. The thermo-diffusion (Soret effect) was taken into account, and molecular diffusion was described with the multicomponent diffusion approach. The number of grid points was always at least 600 to minimize the effect of the grid size on the simulated laminar flame velocity. This lower threshold was determined in a preliminary grid dependence study. Values of 0.1 were chosen for the adaptive mesh parameters GRAD and CURV.

5.4 Simulations of concentration profiles from flow reactors and shock tubes

SENKIN [30] was used to perform the simulations of flow reactor and shock tube experiments in which concentration profiles were measured, assuming constant volume and adiabatic conditions. If the concentrations were reported versus time in turbulent reactor measurements, the half-depletion of the measured concentration of the fuel was matched to the simulated concentration profile to take into account the time shift due to mixing effects. In some experimental datasets, half depletion was not reached or measurements were only taken after half depletion. In these cases, the time-shifting was performed with respect to the average of the minimal and maximal measured fuel concentration values. Only those flow reactor concentration–time measurement points were used where the concentration of the consumed CO was between 10% and 90% of the initial fuel concentration.

5.5 Jet-stirred reactor simulations

The simulations were performed using the PSR code [32] of the CHEMKIN-II package. Wall effects are expected to play a dominant role at low temperatures [20]. The corresponding data points were excluded from the comparisons in the present study, as detailed in Section 6.3.

6 Results and Discussion

6.1 Ignition delay times

In Fig. 1a, the average error function values for the reproduction of shock tube measurements without known volume–time histories ("ST", 1st group of bars), shock tube data using volume–time histories ("ST-VTIM", 2nd group) and rapid compression machine measurements using volume–time histories ("RCM-VTIM", 3rd group) and the overall results (4th group) are compared. The overall best mechanisms for the description of ignition delay times are Li-2015, NUIG-NGM-2010 and SanDiego-2014. Except for these mechanisms, all RCM-related E values are at least three times higher than for ST. Consequently, the overall performance of the mechanisms for ignition simulations will be heavily influenced by their ability to reproduce the ignition delays measured in RCMs.

As the corresponding average σ values (numeric values above the columns) show, this effect is not originated in differences in the standard deviations estimated from the scatter of the experimental results (for details see the Supplementary Material). The ST-VTIM experimental data are in general the most accurately predicted by all mechanisms. However, the E values may still differ up to one order of magnitude, with SaxenaWilliams-2006 being the best mechanism in this category and Ahmed-2007 the worst. Due to the limited amount of ST-VTIM data and the constrained range of conditions covered (mainly high T, high CO/H_2 ratio and φ), a direct comparison of ST-VTIM to ST or RCM is difficult. Yet, it can be shown that the prediction of ignition delay times (τ_{ign}) can be improved if accurate volume–time histories are utilized, especially at low temperatures. The quality of available RCM volume–time histories may suffer from non-idealities (e.g. inhomogeneity, heat loss) that were not captured. Despite the fact that the majority of mechanisms mal-predict RCM ignition delays considerably, the performance of Zsély-2005 is especially bad. This mechanism contains the rate parameters recommended by Baulch et al. [43] without modification. This shows that although the Baulch et al. evaluated rate parameters are widely used, further tuning or optimization is needed for a good description of syngas ignition delay times. Figure S2a in the Supplementary Material corresponds to Fig. 1a,

but shows average absolute deviations instead of *E* values. Three mechanisms (Zsély-2005, CRECK-2012, Davis-2005) under-predict ignition delay times independently of the facility type, i.e. ignitions occur earlier. While the majority of mechanisms over-predicts τ_{ign} from shock tubes (ST and ST-VTIM), 9 of 16 mechanisms under-predict RCM data. Furthermore, the *D* values of RCM data can be much larger than for ST and ST-VTIM, especially for the worst performing mechanisms.

The large differences in the *E* and *D* values between the three sources of ignition delay data require separate investigations with respect to the experimental conditions, which is why the following figures will be split into three parts. If a type of comparison is not meaningful for a certain data subset, e.g. if only very similar mixtures or pressures were studied, it was omitted from the figures. Although the simulation results of Dagaut-2003 for the 7 RCM data points (all Ar-free mixtures) were available, they are not discussed due to their small number.

It can be seen in Fig. 2a to 2c that the average Pearson-type correlation coefficients for the ST subset are higher than for the RCM and the ST-VTIM data. Anti-correlations up to $C = -0.29$ can be observed for pairs formed by Davis-2005 (#3), USC-II-2007 (#5) and Zsély-2005 (#15) with several other mechanisms with respect to the ST-VTIM data. Furthermore, some pairs of SanDiego-2014 (#6) with other mechanisms showed anti-correlations for the RCM data. Especially this latter case (SanDiego-2014 for RCM ignition delays) is a perfect example that deviation from the typical behavior does not have to be a drawback, but can even lead to the best prediction of a sub-group of data. The three highest correlated pairs for each sub-group are Davis-2005/Li-2007 (#3/#5, $C = 0.95$), NUIG-NGM-2010/Li-2007 (#1/#4, $C = 0.91$) and CRECK-2012/SaxenaWilliams-2006 (#7/#12, $C = 0.91$) for ST, GRI3.0-1999/Ahmed-2007 (#10/#14, $C = 0.98$), Kéromnès-2013/Li-2015 (#2/#8, $C = 0.95$) and SanDiego-2014/Starik-2009 (#6/#9, $C = 0.95$) for ST-VTIM and Li-2007/USC-II-2007 (#4/#5, $C = 0.97$), Kéromnès-2013/USC-II-2007 (#2/#5, $C = 0.95$) and Kéromnès-2013/ Li-2007 (#2/#4, $C = 0.95$) for RCM data. Note that some pairs are highly correlated independently of the facility type and that there is just a weak trend towards more correlation among better-performing mechanisms. This trend was much more distinctive in our hydrogen mechanism comparison study [2].

Figure 3 shows the performance of the mechanisms for ignition delay measurements according to ranges of temperature and pressure. As for all partial comparisons, intervals were defined based on a careful inspection of the available data taking into account the specifics of the respective type of measurement and ensuring a statistically significant number of data points in each interval at the same time. For each interval, two numerical values are given referring to the

number of data points considered. The second value refers to the number of data points used for testing Dagaut-2003, while for all other mechanisms the first value is applicable.

In Fig. 3 (top left panel) the dependence of the error function values on the range of temperature behind the reflected shock wave is shown. Most mechanisms reproduce ST experiments similarly adequately above 1000 K. The majority of the mechanisms (except for Zsély-2005 and SaxenaWilliams-2006) perform very poorly below this threshold. These measurements should be interpreted carefully due to the associated facility effect and potential impurity problems as described in Section 5.1. The large over-prediction of τ_{ign} as shown in Fig. S3 can certainly be attributed to these effects. Since recorded pressure changes were not available for these experiments, all shock tube ignition delay time data measured at $T \leq 1000$ K were excluded from further comparisons. This exclusion was not necessary for the ST-VTIM data (Fig. 3, top middle panel), for which only two otherwise bad-performing mechanisms (Ahmed-2007, GRI3.0-1999) show a moderate deterioration with decreasing temperature.

Due to the large scatter in the E values, it is not trivial to identify clear temperature-dependent trends for the RCM subset (Fig. 3, top right panel). Only SanDiego-2014 and Li-2015 perform reasonably at temperatures of 1000 K and lower, which explains their overall superiority for RCM simulations. The two most recent mechanisms (Li-2015 and SanDiego-2014) were released after the publication of the new RCM data of Kéromnès et al. [18]. These data form the largest group among all available RCM data. In the case of the SanDiego-2014 mechanism, Kéromnès et al.'s hydrogen ignition RCM data were utilized to improve the hydrogen sub-chemistry, which apparently affected the prediction of syngas ignitions positively. This underlines the importance of an accurately described H_2 chemistry in the development of syngas mechanisms.

By optimizing rate parameters that are sensitive in this temperature region, the E values can be reduced dramatically, as it has been shown in the hydrogen mechanism optimization study of Varga et al. [6]. A similar behavior can be expected for the optimization of syngas combustion mechanisms. Eight mechanisms, including all top-performing ones except Davis-2005, show a trend towards better agreement with increasing temperature, similar to the shock tube data. The worst-performing ones (e.g. Ahmed-2007 and Zsély-2005) show deviating trends. Figure S3 (top panels) indicates a clear trend towards stronger over-prediction (or less distinct under-prediction) at lower temperatures, which gives a much better phenomenological description than the E values alone. Note that Zsély-2005 and SanDiego-2014 deviate from this trend for RCM data.

As shown in all three bottom panels of Fig. 3, most mechanisms have a maximum E value at an intermediate pressure range of approximately 10 atm, meanwhile the experiments at pressures

below or above this region are much better predicted. It can be seen in Fig. S3 (bottom panels) that most mechanisms tend to more over-prediction (or less under-prediction) with increasing pressure for ST and RCM data, with exceptions such as Davis-2005, USC-II-2007 (both ST), Kéromnès-2013 and SanDiego-2014 (both RCM). For the Zsély-2005 mechanism, which typically does not follow the general trends, this is valid for both the ST and RCM data. With respect to the ST-VTIM data, the mechanisms show various trends in the deviations of the corresponding simulations from the measurements. There is a potential that in shock tubes at low temperatures and long ignition times, mechanisms which tend towards under-prediction of ignition delays balance out the facility effect-related systematic over-predictions at these conditions. As Fig. S3 (top left) shows, such behavior might be expected for Zsély-2005. A closer look at the D values depending on τ_{ign} (threshold 1 ms) confirmed this trend and also revealed that Davis-2005 should be used with caution for the same reason.

In Fig. 4, the influence of the CO/H_2 ratio, φ and X_{dil} on the agreement of ignition delay time simulations and experiments was studied. For ST-VTIM and especially ST data, a trend towards less accuracy at higher CO contents in the fuel (i.e. large CO/H_2 ratio) can be identified, which is not accompanied by a pronounced trend in the D values (Fig. S4, top leftmost panels). For the RCM subset, various largely different, mechanism-specific trends can be found which prohibits drawing general conclusions. This applies to both the E and D values (Figs. 4 and S4, top right panel, respectively), although all mechanisms seem to have a local minimum of D in the intermediate interval of CO/H_2 = 1.4 – 4, suggesting a superposition of other effects, e.g. the influence of temperature or pressure. However, the rise of E and D values towards high CO/H_2 ratios may indicate room for improvement in the carbon-related chemistry of the affected mechanisms. For shock tubes it can be seen (Fig. 4, bottom left panel) that the limited number of measured over-stoichiometric conditions are more accurately predicted (and D are also closer to zero, see Fig. S4, bottom left panel) than under-stoichiometric conditions (D values are typically the highest), while at φ = 1 most mechanisms have higher E values. For the RCM data, the E values drop towards increasing φ. Unfortunately, no fuel-rich data are available here. On the other hand this means that the region with the most remarkable shortcomings for RCM of practically all mechanisms is not only characterized by low temperature, but also a low fuel-to-air ratio. Similar trends were found for hydrogen [2].

Since the diluent system in the RCM experiments was almost always Ar/N_2, the influence of the type of diluent was only investigated for ST and ST-VTIM data (see Fig. S5a for E values and S5b for D values). For the ST subset, higher E values can be found for most mechanisms (except Rasmussen-2008, Kéromnès-2013, Sun-2007 and Li-2007) on data with Ar dilution

compared to diluents that are N_2-based. ST-VTIM measurements were carried out using Ar-based diluents only. While the mechanisms generally show a similar ST-VTIM performance in these diluent systems, measurements in pure Ar are slightly more accurately described by most mechanisms than those in Ar/CO_2 and Ar/N_2. All mechanisms except Zsély-2005 over-predict N_2-based ST data, meanwhile simulations using Zsély-2005, Davis-2005, USC-II-2007, CRECK-2012 and Li-2015 consistently under-predict Ar measurements (ST and ST-VTIM), indicating that the chosen third body collision efficiencies for argon may not be ideal and/or qualitatively different from the others.

6.2 Flame velocity measurements

In Fig. 1b, the performance of the mechanisms with respect to the reproduction of the flame velocity measurements using different experimental techniques and the overall results are presented. Overall, the Davis-2005 mechanism is the best for flames, however, all mechanisms have a similar performance, except Starik-2009, SanDiego-2014, Rasmussen-2008 (which was not developed for flames, see Chapter 3) and Ahmed-2007. The correlation coefficients for flames (Fig. 2d) are higher on average than for ignition delay times, but lower than e.g. for JSR. The highest correlation coefficients can be found for the mechanism pairs Kéromnès-2013/Li-2007 (#2/#4, C=0.97), USC-II-2007/CRECK-2012 (#5/#7, C=0.93) and Li-2015/SaxenaWilliams-2006, (#8/#12, C=0.93). While the first pair is among the higher correlated ones for ignition delays as well, the last two are not very strongly correlated for ignition delay times, which indicates significant differences in the respective choice of rate parameters for reaction pathways governing the ignition process, but do not affect flame propagation strongly. Unsurprisingly, the two worst mechanisms for flames (#14, Ahmed-2007 and #11, Rasmussen-2008) are the least correlated with all other mechanisms.

Syngas flame velocities have been measured by using four different types of methods: the flame cone method (FCM, also referred to as the Bunsen burner method, see e.g. [44] for a discussion of this method), the outwardly propagating/ spherical flame method (OPF) [45], the counterflow twin-flame method (CTF) [46], and the heat flux burner method (HFM) [47]. As Fig. 1b shows, the traditional FCM approach and especially OPF seem to be less accurate in comparison to more advanced techniques (CTF, HFM) based on the agreement of experiments and simulations. The σ values for the OPF are typically much higher than for the other methods. If the σ values were more similar to each other, the differences in the E values would be even more pronounced. Note that the ranges of operating conditions and utilized diluents covered differs largely between the different techniques (see Table 2, particularly regarding p and T_{init}),

which is likely to influence the comparison of the experimental methods. Jayachandran et al. [48] recently reported results of numerical simulations of spherically expanding flames with radiative heat loss which indicate that the standard imaging/measuring approach in the OPF experiments, the shadowgraph/Schlieren technique, could result in a systematic under-estimation of the true laminar flame velocity due to an inward flow induced by the density change in the burned gas. Recently, Varea et al. [49] quantified the uncertainties of the laminar burning velocities from OPF experiments (up to 30% for hydrogen/air) and identified the extrapolation technique as the major source of errors. In addition to this, Yu et al. [50] emphasized radiation effects in spherically propagating flames and suggested an empirical formula for radiation correction. Figure S2b in the Supplementary Material shows that with all mechanisms except Rasmussen-2008, Li-2015, NUIG-NGM-2010 and USC-II-2007 (the mechanisms which generally tend towards the strongest under-prediction of measured flame velocities) medium-to-strong over-predictions can be observed particularly for the OPF subset, i.e. the simulation results were typically above the flame velocities determined in experiments, which supports the observation made in [48]. Correlations of the mechanisms for all four types of experimental facilities are shown in Fig. S6. Note the varying correlations of Ahmed-2007 (#14) for CTF compared to FCM/OPF/HFM. Particularly for the FCM, OPF and CTF data, higher correlations can be found among the overall best performing mechanisms (low identifying numbers) except for NUIG-NGM-2010 (#1) which is just ranked 9th in terms of overall performance for flames.

In several measurements of the flame velocity (238 of 1963 data points, i.e. about one eighth of all data), He was used as the bath gas or a component of the diluent mixture. Only mechanisms having He as the bath gas were used for the simulation of these experiments. The general comparison of the mechanisms with respect to flame velocity data was carried out without these experiments ("no He" in Figs. 5b and 6). A separate investigation of the measurements containing He as a diluent was performed for the nine mechanisms in which this species was defined ("He"). The separated treatment allows for a better evaluation of diluent-specific trends. Note that the results shown in Fig. 1b and 2d refer to the "no He" case only.

Fig. 5a shows the reproduction of the experimental flame velocity according to the type of bath gas/diluent system for the five most frequent choices (see Fig. S8a for the other diluents CO_2, H_2O, Ar, He/H_2O and He/CO_2). Most mechanisms perform better in CO_2/H_2O and N_2/H_2O diluents than in pure He, N_2 and N_2/CO_2. Generally, only Sun-2007, SanDiego-2014 and Li-2015 perform worse for He-diluted mixtures, while all others perform better in He or similarly regardless of the bath gas. Bad reproduction of measurements in specific bath gases (e.g. with increased concentrations of CO_2 or H_2O) can possibly be explained by inappropriate choices of

third body collision efficiencies in the mechanism. A good example for this is the exceptionally bad performance of Starik-2009 for measurements in pure CO_2 (see Fig. S8a), which affects the overall averaged performance of this mechanism for flames. Over- and under-predictions can be observed for all types of diluent systems (see Fig. S7a and S8b), however, the performance of some mechanisms is almost independent of the diluent (see e.g. the generally strong over-prediction of Ahmed-2007, SanDiego-2014 and Zsély-2005, and under-prediction of NUIG-NGM-2010 or Rasmussen-2008). Another dilution effect can be seen in Fig. 5b, which shows the dependence on the sum of the concentrations of all diluents. Except for SanDiego-2014, Li-2015 and Sun-2007 for the He-containing flames, all mechanisms tend to predict measured flame velocities more accurately in more diluted mixtures, which is also reflected in decreasing absolute deviations for most mechanisms (see Fig. S7b).

The performance of the mechanisms for flame velocity data in different ranges of initial temperature, equivalence ratio, CO/H_2 ratio and pressure is shown in Fig. 6. As illustrated in Fig. 6 (top left panel), the errors of the simulated flame velocities increase with the initial temperature for experiments in He-free mixtures. While below 400 K most mechanisms predict similar flame velocities, between 400 K and 600 K the agreement is slightly worse, and above 600 K all mechanisms predict very different flame velocities. The single exception is the CRECK-2012 mechanism that has a similarly good performance for all intervals, while the errors of Starik-2009, Kéromnès-2013 and SanDiego-2014 become very large above 600 K. Note that 29 data points (from Fig. 20a in [51]) were not included in the 310 – 360 K interval, as some of these are badly reproduced by all mechanisms and would impede the identification of temperature-related trends. It is not recommended to use these measurements for mechanism development, particularly not the ones shaded gray in Table C of the Supplementary Material. For He-containing mixtures a similar trend (increase of E with the temperature), although much less pronounced, can be observed with a sharp increase at 600 K for Li-2015, Sun-2007 and especially for SaxenaWilliams-2006 (see Fig. 6, top left panel, right side). NUIG-NGM-2010 and Kéromnès-2013 performed much better for high temperature flames in He. All measurements at 600 K and higher were carried by one group of researchers only (Natarajan et al. [52, 53]) using Bunsen burner flames, which tend to be over-predicted by almost all mechanisms (see Fig. S9 top left panel). Hence, additional studies at similar conditions would be highly desirable to decrease the uncertainty in the description of preheated flames.

Fig. 6 (top right panel) shows that the mechanisms tend to be more accurate at lean conditions for both He containing and He-free mixtures (except SanDiego-2014 and Sun-2007 in He mixtures) and the performances of all mechanisms become worse moving towards fuel-rich

conditions. At high equivalence ratios in He-containing mixtures, the mechanisms produce lower error function values than in He-free systems. However, the deterioration of the accuracy is not dramatic in both cases, except for Ahmed-2007 and Rasmussen-2008 (both "no He"). It is interesting that the corresponding D values (Fig. S9, top right panel) tend to increase with φ in the "no He" case and decrease for He data.

As it can be seen in Fig. 6 (bottom right panel), most mechanisms predict CO-rich or "pure" fuels (H_2 diluted in CO_2 or CO diluted in N_2/H_2O), more accurately than those data points with a low CO/H_2 ratio (i.e. more H_2) and the absolute D values decrease with increasing CO/H_2 ratio (Fig. S9, bottom right panel). Due to the lower content of H_2, slower flame propagation is expected.

No distinctive pressure dependence can be observed for the performance of the models in He-containing flames (Figs. 6 and S9, bottom left panels). However, in He-free mixtures, several mechanisms are more accurate at higher pressures. In contrast to this, the D values increase with p for many mechanisms, including Sun-2007, which was primarily developed to reproduce high-pressure flame velocities. The three other "high-pressure mechanisms", Kéromnès-2013, Li-2015 and Rasmussen-2008, are indeed superior at high pressures in terms of E and D values, with the exception of Li-2015 in He-containing flames. It can be shown that the decrease of E with increasing pressure is associated with the distribution of σ values: at low p, the average σ is 3.4 cm/s, while for intermediate pressure σ = 7.4 cm/s and at high pressure σ becomes 13 cm/s. Hence, those more scattered high-pressure data should be interpreted carefully. We assigned high σ values to the Burke et al. [54] data, which is one of the reasons that low E values were obtained related to these data. The Burke et al., measurements were recently reviewed and criticized by Li et al. [19], who measured new high pressure syngas flame velocity data at a different temperature and using a similar mixture composition. It was found that the level of reproduction of both datasets by the mechanisms was comparable, even if assuming identical σ values. Nevertheless, the Li et al. data are valuable, because these are very discriminatory; mechanisms such as Starik-2009, SanDiego-2014 and Zsély-2005 fail to predict the measured flame velocities dramatically.

Note that the dataset-related property σ was assigned to individual data points here, and the σ value of a subset was calculated based on averaging over all participating datasets. This limits the usefulness of comparing σ values in constellations, where the data points of a dataset are distributed over two or more subsets.

6.3 Concentration profiles from flow reactors, shock tubes and JSRs

Two types of data fall in this category: flow reactor concentration measurements vs. time and outlet concentration measurements from three different types of facilities (flow reactors, shock tubes and JSRs). As Fig. 1c shows, the *E* values corresponding to shock tube and JSR data (for the latter subset, data points at $T < 1000$ K were excluded due to facility effects) are lower than for those coming from flow reactors, with the exception of the reproduction of outlet concentrations by NUIG-NGM-2010, Kéromnès-2013 and SanDiego-2014. In total, the overall five best mechanisms for syngas ignition delay times are also the best for concentration simulations, most notably NUIG-NGM-2010 and USC-II-2007. The latter performs better on shock tube concentration data, while the aforementioned mechanism is superior for flow reactor data. The performance of all mechanisms is very similar for JSR data, which cover a comparatively narrow range of operating conditions, while larger deviations can be observed for the other facility types. Especially Ahmed-2007 and SaxenaWilliams-2006 are inadequate for the description of concentration profiles with a sufficient accuracy. Figure S2c shows that flow reactor concentration–time and outlet profiles as well as JSR outlet concentrations are – in average – under-predicted by all mechanisms (except for Kéromnès-2013, NUIG-NGM-2010 and SanDiego-2014 for flow reactor outlet concentrations). In contrast to this, the shock tube concentration data are over-predicted by a number of mechanisms, most notably Starik-2009. On average, outlet concentrations can be matched better with simulations than temporal profiles. However, as it will be shown in this section, this statement is not valid for flow reactor outlet profiles at certain conditions (e.g. high temperature).

Measured CO and CO_2 concentrations are less accurately matched than the O_2 profiles by all mechanisms in flow reactors and JSRs (see Fig. 7a). The *E* values CO and CO_2 profiles are strongly affected by a datasets that cannot be reproduced by any of the mechanisms. CO and O_2 profiles are similarly well predicted by the mechanisms if they are originated from shock tube measurements, while the worst agreement is obtained for CO_2 profiles (Fig. 7b). This failure of certain mechanisms can be related to the lack of agreement for CO and partially CO_2 profiles, which form the largest groups of data among the measured profiles. Figure S10a shows that the consumed species (CO, O_2, H_2) are consistently under-predicted by the mechanisms and that the opposite trend applies to the produced species (CO_2, OH), which means that the predicted conversion is usually lower than the measured one. This type of behavior was observed for each measurement type except shock tube outlet profile data. Hence, the negative average *D* values for flow reactor concentration–time profiles are in fact the result of merging even lower negative *D* values of consumed CO/O_2 with positive *D* values for CO_2 and OH. In the JSR experiments,

only consumed species were measured, which may explain the relatively large average under-prediction in Fig. S2c. Figure S10b shows the deviating trends for the shock tube data: except SanDiego-2014, all mechanisms over-predict CO and O_2, but under-predict CO_2.

The facility effect associated with JSR data is well illustrated in Fig. 8 (top rightmost panel): low T data cannot be described accurately and are largely under-predicted (see Fig. S11). Note that this effect is the least pronounced for the GRI3.0-1999 mechanism. In all other types of experiments (Fig. 8, three panels from the left), an increase of E with temperature can be observed, the strongest for flow reactor outlet profiles, for which a very strong under-prediction of concentrations can be observed (Fig. S11, top middle left panel). The large E values at high temperatures for flow reactor outlet concentration profiles are caused by a single dataset (*x30001021.xml* in Table E of the Supplemental Material). The overall mediocre performance of mechanisms such as Li-2015 and Starik-2009 can be originated in their complete failure in reproducing some of the high-T flow reactor experiments published by Alzueta et al. [55]. The dependence on the equivalence ratio is not as obvious as the previous relation. Various trends can be found in Fig. 8 (bottom panels). Again, largely varying σ values disguise the effect of the variation of a parameter, particularly for flow reactor outlet concentrations. In general, the overall dependence of the agreement of simulations and concentration measurements on the equivalence ratio is rather small.

While the investigation of the influence of the CO/H_2 ratio was not meaningful due to the fact that just very few different fuel mixtures were investigated in each facility type, the effect of pressure variation could be well visualized (see Fig. 9). For medium-to-high pressures (measured in shock tubes, right panel), the E values increase with increasing pressure, while at low-to-medium pressures (see flow reactor data, left panel), various mechanism-dependent trends can be observed, including a decrease of E with increasing pressure for some of the overall best performing mechanisms. This is an indication that other effects might be more important in this pressure region. Nevertheless, under-prediction tends to become more important in high-pressure regions (see Fig. S12).

As Figs. 2e–h show, the correlation coefficients can be very different for the various types of concentration experiments: JSR correlation coefficients are very high ($C \geq 0.99$), while many pairs of mechanisms are slightly negatively correlated for flow reactor data, i.e. not at all similar in their behavior. Examples in Fig. 2e are the mechanism pairs that the Starik-2009 (#9), NUIG-NGM-2009 (#1) and GRI3.0-1999 (#10) mechanisms form with many other mechanisms and in Fig. 2f those of SanDiego-2014 (#6). There is somewhat more similarity among the performance of the mechanisms for shock tube speciation data, especially for the overall best performing

mechanisms. Also, it is worth noting that for most mechanism pairs, higher correlations were obtained for the species CO_2 and OH than for the consumed species CO, O_2 and H_2 (not shown in Fig. 2).

6.4 Data filtering and weighting considerations

It is unavoidable that some regions of operating conditions are over-represented in our database due to the fact that some measurements were carried out multiple times at almost identical conditions by various groups of researchers, e.g. to test their experimental apparatuses. This introduces a sampling bias that could be mitigated by weighing down or averaging those experimental results. As an example, species concentration data cover only a limited range of operating conditions, particularly just a small number of different mixtures. Hence, many of these data points have a comparatively low content of "new" chemical information, and it can be shown that several data points overlap even within one dataset. One way to minimize the sampling bias due to repeated measurements is to identify those points and decrease their weight in the calculation of E. We considered two measurements to be quasi-identical, if they differed in all species mole fractions by 0.05% or less, and by 3 K or less (for ignition delay times: 1 K or less). For ignition delay times, a constant pressure criterion of 0.1 atm or less was assumed, while for flame and concentration data a relative criterion was used: $(1.01325–1) \times p$[atm], i.e. we considered measurements identical if the same pressure value was reported in either atm or bar units. Naturally, the residence time is an additional criterion for distinguishing data points, when temporal profiles of species concentrations were recorded in the same apparatus. These criteria were the base for the calculations of the normalized data point weights and corresponding dataset weights. The modifications of the error function E with respect to weighting are detailed in Part 2 of the Supplementary Material.

Further problems arise if the measurements carried out at certain conditions are accompanied by large experimental uncertainties or the measured values themselves are problematic, e.g. due to assumptions in the interpretation of a measured signal. Although some measurements were excluded before carrying out the general mechanism comparison (e.g. due to large inconsistencies), others that were retained in the comparison may still contain large systematic errors. A possible approach to identify systematic errors is to optimize the rate parameters, thermochemical data and transport data of an appropriate mechanism within their range of uncertainty on a single dataset only and filtering out those datasets that cannot be reproduced this way, i.e. whose E_i values could not be lowered to values close to unity. This procedure is very time consuming and based on the assumption that all important reaction steps are present in the

mechanism, and the uncertainty domain of all parameters are well assessed. Another strategy is to identify and filter out those measurements that none of the mechanisms could reproduce within a pre-defined uncertainty, e.g. 3σ, which is equivalent to an error function value of $E_i \leq 9$. This procedure does not imply that these measurements necessarily have large systematic errors as it is possible that none of the mechanisms are capable of modeling the chemistry at a specific condition in an accurate manner. However, these measurements should be treated with care in future model development and should potentially receive further attention. In total, 33 of the 358 datasets (corresponding to 533 data points of the previously utilized 4127 data points) in the overall comparison were filtered out based on the $E_i \leq 9$ criterion. In Tables A – G of the Supplementary Material the experimental datasets that do not fulfill the $E_i \leq 9$ criterion are shaded gray.

Figure 10 compares the un-weighted overall results to the weighted results, both for the complete data subset and a reduced subset based on the above filtering criterion. It can be seen that the performances of almost all mechanisms improve if weighting is introduced, meanwhile a negative impact of weighting is observed for Li-2015, SanDiego-2014 and Dagaut-2003 (the latter is not shown in Fig. 10). Hence, the mechanisms react different to the introduction of weighting factors, as the deviations between weighted and un-weighted results show. The Starik-2009 mechanism appears to be insensitive towards weighting, while the E values decrease or slightly increase for the other mechanisms. If improvement is observed, its extent tends to scale roughly linearly with the (old) un-weighted errors, with the exceptions of Starik-2009 (due to the increased relative importance of the problematic CO_2 flames). Due to the quasi-linear scaling of the improvement, the order of performance of the mechanisms changes only slightly, and is mainly among the eight worst mechanisms. A closer look at the shifts in the performances of the mechanism due to the introduction of weighting reveals that the reduction of the overall E values is a consequence of a large drop of E for ignition delays up to –42%, except for SanDiego-2014, Dagaut-2003 (both $\Delta E \approx 0\%$) and Li-2015 ($\Delta E \approx -7\%$). For flame velocity data, $-2.2 \leq \Delta E \leq 2.4$ (except the outlying mechanism Rasmussen-2008 with $\Delta E \approx -9\%$) and for concentration data, ΔE values are always positive and in the range of 0.0% – 1.5%. Obviously some ignition delay data points were disfavored by weighting for which Li-2015, SanDiego-2014 and Dagaut-2003 mechanisms performed well. In other words, these three mechanisms tend to reproduce well-studied operating conditions of shock tubes and RCMs better than those conditions that were investigated less frequently, which is a potential drawback of these mechanisms. The arrows on the bars illustrating the average absolute deviations (Fig. 11) support the above statements with

respect to E values and show that ignition delays are much more sensitive towards weighting than flame velocities and concentration profiles.

By filtering out 33 datasets based on the $E_i \leq 9$ criterion, the performances of almost all mechanisms improved significantly, as it is indicated by green open triangle symbols in Fig. 10. Unlike in the previous hydrogen comparison study [2], an average agreement with the experimental data of less than 3σ could not be achieved this way with any of the mechanisms. If filtering based on $E_i \leq 9$ and weighting are combined, even lower average E values are obtained (green full triangle symbols in Fig. 10). As the relative distances between the pairs of symbols show, both effects are additive. However, after combined data weighting and filtering, the average error for each mechanism is still higher than the average 3σ limit.

6.5 Analysis of the sensitivity coefficients and reaction rate coefficients

In the previous sections we saw that mechanisms can differ significantly in their behavior. In order to improve a chemical model it might be useful to have information about which parts of a mechanism are responsible for a good or bad performance at the various conditions. Furthermore, it can be interesting to have an estimate on how the reproduction of some measurements changes if a reaction rate coefficient is tuned to describe another experiment better. To investigate these relationships, sensitivity analyses were performed at the conditions of all measurement data. Six mechanisms (NUIG-NGM-2010, Kéromnès-2013, SanDiego-2014, GRI3.0-1999, SaxenaWilliams-2006 and Sun-2007) were chosen which represent different levels of predictivity. All but GRI3.0-1999 contain the species He. For the calculation of the sensitivity coefficients, a brute force method was applied by varying the pre-exponential Arrhenius parameter A for all reactions by 5%, one-by-one. The sensitivity values were normalized and then scaled, i.e. divided by the maximum absolute normalized sensitivity coefficient so that this value becomes either –1 or +1. This range was divided into ten equidistant intervals and the relative frequency of the scaled sensitivity values S for each of these intervals was then determined. An additional interval was added containing the frequency of zero values of S.

Seven data subsets were defined for which the values of S were evaluated: (1) shock tube ignition delay times, (2) RCM ignition delay times, (3) flame velocities in He-free systems, (4) flame velocities in He-containing systems, (5) all CO concentration profiles, (6) all O_2 profiles, and (7) all CO_2 profiles, giving a total of 41 test cases (6 mechanisms times 7 data subsets minus one case which could not be investigated due to the absence of He in GRI3.0-1999). Note that further species profiles (H_2 and OH) are not discussed here due to the small size of the subsets.

In Table I of the Supplemental Material, the ten most sensitive (i.e. important) reactions in all mechanisms for each data subset are summarized. Note the superscripts used in Table I and in Figs. 13–16 (which will be discussed later in this section), which describe some of the distinctive features in the way how certain reactions were implemented in respective mechanisms. Furthermore, the averages of the absolute S values are listed for each reaction. The numerical values should not be used to compare the importance of a particular reaction across different mechanisms, as the normalization factors usually differ among the mechanisms. However, they are useful to understand the importance of a reaction relative to other reactions in the same mechanism. As a result of these summaries, four reactions were identified that are the most important reaction in at least one of the 41 test cases. Unsurprisingly, chain-branching $H + O_2 = O + OH$ makes the most appearances here (27 times the most important reaction), followed by $CO + OH = CO_2 + H$ (12 times), which is the major source of heat release in syngas combustion. Chain-branching reaction $H + HO_2 = OH + OH$ has an impact especially in flames where it is always among the top 4 reactions and once the most important, while $CO + HO_2 = CO_2 + OH$ is sensitive mainly at RCM conditions (high pressure, low temperature), where HO_2/H_2O_2 chemistry becomes more important. Similar evidence for the importance of $CO + HO_2 = CO_2 + OH$ was found e.g. by Mittal et al. [56]. All four reactions are shaded dark gray in Table H.

A second group (shaded light gray) includes further important reactions and consists of $H_2 + O = H + OH$, $CO + O_2 = CO_2 + O$, $H_2O_2 + H = H_2 + HO_2$ and the low-pressure limit of $H_2O_2 (+M) = OH + OH (+M)$; the two latter are again most important at RCM conditions. The fact that certain reactions do not appear in these two groups does not mean they can be neglected in mechanism development. However, for the purpose of this work we will focus the above mentioned eight reactions.

Figure 12 shows the Arrhenius plots for reactions of the first group using the parametrizations of rate coefficients from all 16 mechanisms. For a better understanding of these plots, some distinctively deviating rate coefficients are highlighted with their respective mechanism identifiers. For $H + O_2 = O = OH$ (top left panel), most mechanism developers chose very similar rate coefficients, with the overall worst performer Zsély-2005 deviating the strongest at low temperatures. At higher pressures (see magnified view), all well-performing mechanisms (thick lines) except CRECK-2012 include a slightly lower rate coefficient, with $\ln (k / \mathrm{cm^3\ mol^{-1}\ s^{-1}}) \approx 28.5$ at $T = 2000$ K. For the other three reactions in this figure, rate coefficients at temperatures relevant in combustion processes differ much more. In case of $CO + OH = CO_2 + H$ (top right), one of the bad performers Ahmed-2007 has the highest k values up until ~2000 K, while in NUIG-NGM-2010 one of the lowest reaction rate coefficients is defined. The significantly

different slope of some curves (e.g. for Rasmussen-2008) is noteworthy. While in Sun-2007 a somewhat lower rate coefficient for H + HO_2 = OH + OH (bottom left panel) is defined than in all other mechanisms, Zsély-2005, Dagaut-2003, Ahmed-2007, Starik-2009 and CRECK-2012 show the most deviating rate coefficients and all of them are considerably larger at high temperatures. The two latter mechanisms also have a much steeper slope in their ln *k* curves due to an up to 3 times higher activation energy which was chosen. For CO + HO_2 = CO_2 +OH (bottom left), deviations of more than one order of magnitude of *k* on a linear scale can be found, with almost all good performers at the lower end of the range (adopting the lower rate coefficient expression suggested by Mittal et al. [56] in order to achieve a better prediction of RCM ignition delay times), and almost all bad performers at the upper end. As stated above, it is mainly important at low-temperature (RCM) conditions. This statement is supported by an increase in the average absolute *S* values at temperatures below 1000 K in shock tubes (not shown). Due the fact that a large number of data points in subset (2) reflect experiments at low temperature and as a consequence of the results for RCM data found in Section 6.1, a link between the chosen rate coefficient and the performance of the mechanism at these conditions can be established. Although the most apparent for this particular reaction, similar links can be identified for other reactions whose parametrizations were compared in this work. Typically, the best performers possess relatively similar reaction rate coefficients.

Similar to Fig. 12, rate coefficients for the reactions of the second group are visualized in Fig. S22 in the Supplemental Material, which can be interpreted in a similar manner. Most mechanisms include similar rate coefficients for H_2 + O = H + OH, with the least differences in a temperature interval from about 1200 K to 1700 K. Below this, some of the bad performers (Dagaut-2003, Rasmussen-2008, Sun-2007 and Zsély-2005) have lower *k* values, while only CRECK-2012 lies above the majority of the mechanisms. The best agreement of the mechanisms with respect to CO + O_2 = CO_2 + O can be found at high temperatures (except Zsély-2005), while the lower the temperature, the more the graphs scatter, with Rasmussen-2008 and Starik-2009 being the negative and positive extremes. For H_2O_2 + H = H_2 + HO_2, many mechanism developers adopted the rate coefficients of previous studies, resulting in groups of coinciding graphs. A number of mechanisms including almost all bad performers (Starik-2009, Rasmussen-2008, Sun-2007, Ahmed-2007, Zsély-2005, Dagaut-2003) show the most deviating trends, both in their low *k* values at high temperatures as well as in their comparatively low activation energies and therefore shallower curves. Again, a link between low-temperature performance and the chosen rate coefficient seems justified. While most low-pressure limit rate coefficients of H_2O_2 (+M) = OH + OH (+M) overlap (with some spread at very high temperatures >2000 K),

Zsély-2005 has much higher *k* values than all other mechanisms. For Dagaut-2003, no separate low-pressure limit was defined, instead the pressure-independent reaction $H_2O_2 = OH + OH$ was plotted in the bottom right panel, which is unsurprisingly far off the other the graphs.

In the next step, the results of local sensitivity analyses of the six selected mechanisms shall be discussed. In Figs. 13–16, the distributions of the sensitivity coefficients for the reactions in the first group with respect to all seven data subsets are shown. It is worth emphasizing that the sensitivity coefficients highly depend on the experimental conditions, which can differ largely among the types of measurements and facilities, e.g. between shock tubes and RCMs. In other words, different sub-chemistries can be important at various conditions across all types of measurements. Therefore, the ranges of experimental conditions provided in Figs. 13–16 for each data subset should be read together with the graphs.

As observed in Fig. 13, the chain-branching step $H + O_2 = O + OH$ has a promoting effect on the overall reactivity at all combustion conditions. An increase of the rate coefficient of this reaction leads to shorter ignition delays (negative *S* values for ST/ST-VTIM and RCM), faster flame propagation (positive *S* values for flames) and lower concentrations of reactant species CO and O_2 (negative *S* values, i.e. faster consumption) as well as a faster build-up of the product species CO_2. However, the relative importance of this reaction varies depending on the data subset and their respective conditions and also among the mechanisms. While $H + O_2 = O + OH$ is almost always the most influential one at RCM and especially ST conditions, at typical flame conditions it competes with other reactions in flames, leading to a broader and shifted distribution of the *S* values. The most notable exception from these observations is the Sun-2007 mechanism, for which $H + O_2 = O + OH$ is strongly promoting on He-containing flames. An explanation for the below average performance of this mechanism for He-containing flames as it was described in Section 6.2 may be found here. Similar evaluations of the results of sensitivity analyses were performed for the comparison of hydrogen combustion mechanisms [2]. Although a plot corresponding to Fig. 13 in this study for $H + O_2 = O + OH$ was unfortunately not presented there, Fig. 13 can be compared to the numerical values summarized in Tables G1–G5 (columns “R1”) found in the Supplemental Material of [2]. It can be seen that the relative importance $H + O_2 = O + OH$ is lower in H_2 combustion compared to syngas at RCM and flow reactor conditions, while it influences flame propagation stronger for pure H_2. As a conclusion from this, a quantitatively different effect of modifying the rate coefficient $H + O_2 = O + OH$ can be expected for hydrogen and syngas combustion simulations, which gives rise to the usefulness of mechanism development approaches that do not only feature the improvement of rate

coefficients for one combustion system of interest, but also consider hierarchically lower systems such as pure hydrogen.

In contrast, reaction CO + OH = CO_2 +H is mainly important in flames, as the sharp peaks towards large positive *S* values in Fig. 14 show. While this reaction step is very important at least for some of the shock tube data points and ranked 3rd to 5th overall for all mechanisms, it does not even appear in the top 10 of the most sensitive reactions for four mechanisms at RCM conditions and for the other mechanisms it is not very highly ranked either. This limits the informative value of the RCM-related graphs in Fig. 14, where typically two peaks around $S = 0$ can be observed, i.e. the *S* values are typically small, but also non-zero, hence, the reaction is not completely insensitive. For all concentration profiles, this reaction is somewhat important as well. Note that for the two mechanisms that used multiple duplicates for this reaction, only the most sensitive branch was chosen for plotting in Fig. 14.

As Fig. 15 shows, the sensitivity of reaction H + HO_2 = OH + OH is also very dependent on the chosen data subset (i.e. the conditions in the experiments). While in flames chain-branching via the reaction of the radicals H and HO_2 is rather important, it only plays a secondary role for the accurate prediction of species profiles and ignition delay times. The histograms for both flame data subsets are relatively scattered when comparing them with each other, i.e. the relative importance of this reaction differs. This is a rather surprising fact, since Fig. 12 showed that for the creation of four of the six mechanisms (the Galway mechanisms NUIG-NGM-2010 and Kéromnès-2013 as well as SanDiego-2014 and SaxenaWilliams-2006, both published by the San Diego group), identical rate coefficients were chosen. This means that other rate coefficients were modified during the development of the newer mechanisms, which indirectly increased or decreased the concentrations of the radicals reacting in reaction H + HO_2 = OH + OH.

Reaction CO + HO_2 = CO_2 + OH (Fig. 16) strongly affects the prediction of RCM ignition delays, while it is of little relevance at flame conditions. The relative importance of this reaction for shock tube ignition delays and concentration profiles depends on the mechanism and is stronger for the San Diego mechanisms and GRI3.0-1999, the latter being the mechanism that responds the strongest to changes of the rate of CO + HO_2 = CO_2 + OH.

It is apparent that it becomes more difficult to extract meaningful information from analyses of sensitivity coefficients the lower their actual values are, which is why we shall not discuss reactions of the 2nd group in great depth (the corresponding figures can be found in the Supplementary Material, Figs. S23–S26). Reaction H_2 + O = H + OH is as important as some of the reactions in the first group at certain conditions (Fig. S23). For all subsets related to species profiles, Mexican hat-shaped curves can be observed for CO + O_2 = CO_2 + O, which means that

the reaction is either not at all or very important, depending on the specific condition in the respective experiments (see Fig. S24). The two remaining reactions, $H_2O_2 + H = H_2 + HO_2$ and H_2O_2 (+M) = OH + OH (+M), both including hydrogen peroxide, are mainly sensitive at RCM conditions, and partially for the shock tube subset. The significantly deviating histograms for the NUIG-NGM-2010 mechanism with respect to $H_2O_2 + H = H_2 + HO_2$ are due to its unusually low sensitivity towards this reaction, which is even more interesting when bearing in mind its outstanding overall performance.

As some of the above examples showed (e.g. the varying relative importance of $H + HO_2 = OH + OH$ despite the use of identical rate coefficients), a comprehensive comparison of sensitivity coefficients can be a useful tool to determine indirect effects of parameter modifications in addition to the quantification of the mechanism performance discussed in the previous sections. With such analyses, model developers are enabled to identify influential reactions in their mechanisms and by comparing both, rate coefficients and relative sensitivities, the effect of changing a parameter on various observables can be pre-estimated, with the ultimate goal of achieving a better agreement with the measured data. Figure 12 illustrated that the rate coefficients implemented in the investigated mechanisms can vary significantly. Additional theoretical calculations and experimental determinations of the reaction rate coefficients can contribute to the establishment of more reliable and robust kinetic parameters.

7 Conclusions

An accurate description of the combustion of syngas is important from both scientific and industrial viewpoints [1]. An excellent review [18] was elaborated recently in which new developments in this field are discussed. However, a comprehensive investigation and comparison of all recent syngas combustion mechanisms has not been published.

Figure 11 summarizes the absolute deviation values for the various types of measurements. In average, a trend towards under-prediction of concentration profiles can be observed. As stated in Section 6.3 and shown in Figs. S2c and S10, the type of under- or over-predicting behavior depends on the measured profile and partially on the experimental facility. Hence, this average under-prediction is a result of the dominance of measured species such as CO, O_2 and H_2 and despite that opposite trends were obtained for shock tube data. For both ignition delay times and flame velocities, a weak majority tends to over-prediction (i.e. slower ignition and faster flame propagation), however, other mechanisms (e.g. CRECK-2012) show the complete opposite

trend. From a model development standpoint, it is nevertheless important to know about these general trends in order to choose appropriate reaction rate coefficients for counter-balancing.

As shown in Sections 6.1 to 6.3, the best mechanism for the reproduction of ignition delay times is the Li-2015 mechanism, for flames it is the Davis-2005 mechanism, and for concentration profiles the NUIG-NGM-2010 mechanism. Several mechanisms do not perform properly at non-stoichiometric conditions, at high initial temperatures in flames, at low temperatures in shock tubes, and at high pressures. The majority of the published mechanisms are not suitable for an accurate reproduction of RCM-VTIM data. Investigations of the performance at various experimental conditions, such as those presented in the previous chapters, may help the selection of a mechanism for simulations to be carried out at given industrial conditions.

Figure 10 shows the overall performance of each mechanism, tested against all collected experimental data (black open squares). There is a noticeable improvement of the accuracy over calendar time, though some of the overall best mechanisms were published between 2005 and 2007. Currently, the most accurate mechanism for modeling syngas combustion is the NUIG-NGM-2010 mechanism, but Kéromnès-2013, Davis-2005, Li-2007 and USC-II-2007 have a similarly good performance, considering all types of experimental data. It is pertinent to note that three of these best syngas mechanisms include the top three for hydrogen combustion (Kéromnès-2013, NUIG-NGM-2010, Li-2007), while Davis-2005 and USC-II-2007 just take average positions there [2]. Other mechanisms may only excel in certain categories and are not able to provide general reliability across the various types of measurements. Figure S13 in the Supplementary Material shows the overall comparison results in the same manner, except that all data points (i.e. also the measurements in He-containing diluents) are covered. The trends and absolute values of E are similar to those in Fig. 10.

Testing 16 recent syngas mechanisms against virtually all experimental data that were published for syngas combustion we pursued the following major goals:

(i) Identification of generally well performing mechanisms and those that are only good at certain conditions or for certain types of experiments;

(ii) Identification of potentially wrong data due to large experimental uncertainties and/or systematic errors. These data cannot be reproduced by any of the investigated mechanisms while other measurements at similar conditions can be matched better;

(iii) Identification of problematic regions of operating conditions that cannot be well reproduced with even the latest and most advanced syngas combustion mechanisms. These conditions require further attention in mechanism development;

(iv) Investigation of the impact of data filtering and weighting on the performance of syngas mechanisms, overall and relative to each other;

(v) Identification of possible explanations for the different performances of the mechanisms in their detailed chemistries via sensitivity analyses.

Acknowledgements

The authors acknowledge the helpful discussions with Prof. Christian Hasse, Ms. Sandra Hartl (both TU Bergakademie Freiberg) and Prof. Dimosthenis Trimis (Karlsruhe Institute of Technology), as well as the financial support of OTKA grants K84054 and NN100523. The authors are also grateful for the supportive comments of the partners in COST collaboration CM0901 Detailed Chemical Models for Cleaner Combustion.

References

1. T. Lieuwen; V. Yang; R. Yetter (Eds.): Synthesis Gas Combustion. Fundamentals and Applications, CRC Press, Boca Raton, 2010.
2. C. Olm; I. G. Zsély; R. Pálvölgyi; T. Varga; T. Nagy; H. J. Curran; T. Turányi, Combust. Flame 161 (2014) 2219–2234.
3. T. Turányi; T. Nagy; I. G. Zsély; M. Cserháti; T. Varga; B. T. Szabó; I. Sedyó; P. T. Kiss; A. Zempléni; H. J. Curran, Int. J. Chem. Kinet. 44 (2012) 284–302.
4. I. G. Zsély; T. Varga; T. Nagy; M. Cserháti; T. Turányi; S. Peukert; M. Braun-Unkhoff; C. Naumann; U. Riedel, Energy 43 (2012) 85–93.
5. T. Varga; I. G. Zsély; T. Turányi; T. Bentz; M. Olzmann, Int. J. Chem. Kinet. 46 (2014) 295–304.
6. T. Varga; T. Nagy; C. Olm; I. G. Zsély; R. Pálvölgyi; É. Valkó; G. Vincze; M. Cserháti; H. J. Curran; T. Turányi, Proc. Combust. Inst. 35 (2015) in press, http://dx.doi.org/10.1016/j.proci.2014.06.071.
7. G. P. Smith; D. M. Golden; M. Frenklach; N. W. Moriary; B. Eiteneer; M. Goldenberg; C. T. Bowman; R. K. Hanson; S. Song; W. C. Gardiner; V. V. Lissianski; Z. Qin GRI-Mech 3.0. http://www.me.berkeley.edu/gri_mech/.
8. M. Tamura; P. A. Berg; J. E. Harrington; J. Luque; J. B. Jeffries; G. P. Smith; D. R. Crosley, Combust. Flame 114 (1998) 502–514.
9. T. Kathrotia; M. Fikri; M. Bozkurt; M. Hartmann; U. Riedel; C. Schulz, Combust. Flame 157 (2010) 1261–1273.
10. A. L. Sánchez; F. A. Williams, Prog. Energy Combust. Sci. 41 (2014) 1–55.
11. S. G. Davis; A. V. Joshi; H. Wang; F. Egolfopoulos, Proc. Combust. Inst. 30 (2005) 1283–1292.
12. I. G. Zsély; J. Zádor; T. Turányi, Proc. Combust. Inst. 30 (2005) 1273–1281.
13. P. Saxena; F. A. Williams, Combust. Flame 145 (2006) 316–323.
14. H. Sun; S. I. Yang; G. Jomaas; C. K. Law, Proc. Combust. Inst. 31 (2007) 439–446.
15. C. L. Rasmussen; J. Hansen; P. Marshall; P. Glarborg, Int. J. Chem. Kinet. 40 (2008) 454–480.
16. A. M. Starik; N. S. Titova; A. S. Sharipov; V. E. Kozlov, Combust. Explos. Shock Waves 46 (2010) 491–506.
17. CRECK modeling Group Hydrogen/CO mechanism version 1212. http://creckmodeling.chem.polimi.it/kinetic.html/.
18. A. Kéromnès; W. K. Metcalfe; K. A. Heufer; N. Donohoe; A. K. Das; C.-J. Sung; J. Herzler; C. Naumann; P. Griebel; O. Mathieu; M. C. Krejci; E. L. Petersen; W. J. Pitz; H. J. Curran, Combust. Flame 160 (2013) 995–1011.
19. X. Li; X. You; F. Wu; C. K. Law, Proc. Combust. Inst. 35 (2015) in press, http://dx.doi.org/10.1016/j.proci.2014.07.047.
20. P. Dagaut; F. Lecomte; J. Mieritz; P. Glarborg, Int. J. Chem. Kinet. 35 (2003) 564–575.
21. S. S. Ahmed; F. Mauß; G. Moréac; T. Zeuch, Phys. Chem. Chem. Phys. 9 (2007) 1107–1126.
22. J. Li; Z. Zhao; A. Kazakov; M. Chaos; F. L. Dryer; J. J. J. Scire, Int. J. Chem. Kinet. 39 (2007) 109–136.
23. H. Wang; X. You; A. V. Joshi; S. G. Davis; A. Laskin; F. Egolfopoulos; C. K. Law USC Mech Version II. High-Temperature Combustion Reaction Model of H_2/CO/C_1-C_4 Compounds. http://ignis.usc.edu/USC_Mech_II.htm/.
24. D. Healy; D. M. Kalitan; C. J. Aul; E. L. Petersen; G. Bourque; H. J. Curran, Energ. Fuel 24 (2010) 1521–1528.
25. Mechanical and Aerospace Engineering (Combustion Research), University of California at San Diego: Chemical-Kinetic Mechanisms for Combustion Applications, San Diego Mechanism, version 2014-02-17 (http://combustion.ucsd.edu).
26. T. Turányi MECHMOD v. 1.42: Program for the transformation of kinetic mechanisms. http://garfield.chem.elte.hu/Combustion/mechmod.htm/.
27. T. M. Vu; J. Park; O. B. Kwon; J. S. Kim, Int. J. Hydrogen Energy 34 (2009) 6961–6969.
28. M. Frenklach PrIMe Webpage. http://www.primekinetics.org/.
29. R. J. Kee; F. M. Rupley; J. A. Miller, CHEMKIN-II: A FORTRAN Chemical Kinetics Package for the Analysis of Gas-Phase Chemical Kinetics. Sandia National Laboratories Report SAND89-8009B (1989).
30. A. E. Lutz; R. J. Kee; J. A. Miller, SENKIN: A Fortran Program for Predicting Homogeneous Gas Phase Chemical Kinetics with Sensitivity Analysis. Sandia National Laboratories Report SAND87-8248 (1988).
31. R. J. Kee; J. F. Grcar; M. D. Smooke; J. A. Miller, PREMIX: A Fortran Program for Modeling Steady Laminar One-Dimensional Premixed Flames. Sandia National Laboratories Report SAND85-8240 (1985).
32. P. Glarborg; R. J. Kee; J. F. Grcar; J. A. Miller, PSR: A Fortran Program for Modeling Well-Stirred Reactors. Sandia National Laboratories Report SAND86-8209 (1986).

33. C. Naumann; J. Herzler; P. Griebel; H. J. Curran; A. Kéromnès; I. Mantzaras, H_2-IGCC project: Results of ignition delay times for hydrogen-rich and syngas fuel mixtures measured. (2011).
34. J. Herzler; C. Naumann, Combust. Sci. Technol. 180 (2008) 2015–2028.
35. G. A. Pang; D. F. Davidson; R. K. Hanson, Proc. Combust. Inst. 32 (2009) 181–188.
36. M. Chaos; F. L. Dryer, Int. J. Chem. Kinet. 42 (2010) 143–150.
37. M. Chaos; F. L. Dryer, Combust. Sci. Technol. 180 (2008) 1053–1096.
38. F. L. Dryer; M. Chaos, Combust. Flame 152 (2008) 293–299.
39. L. D. Thi; Y. Zhang; J. Fu; Z. Huang; Y. Zhang, Can. J. Chem. Eng. 92 (2014) 861–870.
40. G. Mittal; C. J. Sung; R. A. Yetter, Int. J. Chem. Kinet. 38 (2006) 516–529.
41. A. B. Mansfield; M. S. Wooldridge, Combust. Flame 161 (2014) 2242–2251.
42. A. K. Das; C.-J. Sung; Y. Zhang; G. Mittal, Int. J. Hydrogen Energy 37 (2012) 6901–6911.
43. D. L. Baulch; C. T. Bowman; C. J. Cobos; R. A. Cox; T. Just; J. A. Kerr; M. J. Pilling; D. Stocker; J. Troe; W. Tsang; R. W. Walker; J. Warnatz, J. Phys. Chem. Ref. Data 34 (2005) 757–1397.
44. G. E. Andrews; D. Bradley, Combust. Flame 18 (1972) 133–153.
45. D. Bradley; P. H. Gaskell; X. J. Gu, Combust. Flame 104 (1996) 176–198.
46. C. K. Wu; C. K. Law, Proc. Combust. Inst. 20 (1984) 1941–1949.
47. L. P. H. de Goey; A. van Maaren; R. M. Quax, Combust. Sci. Technol. 92 (1993) 201–207.
48. J. Jayachandran; R. Zhao; F. N. Egolfopoulos, Combust. Flame 161 (2014) 2305–2316.
49. E. Varea; J. Beeckmann; H. Pitsch; Z. Chen; B. Renou, Proc. Combust. Inst. 35 (2015) in press, http://dx.doi.org/10.1016/j.proci.2014.05.137.
50. H. Yu; W. Han; J. Santner; X. Gou; C. H. Sohn; Y. Ju; Z. Chen, Combust. Flame 161 (2014) 2815–2824.
51. X. Zhang; Z. Huang; Z. Zhang; J. Zheng; W. Yu; D. Jiang, Int. J. Hydrogen Energy 34 (2009) 4862–4875.
52. J. Natarajan; T. Lieuwen; J. Seitzman, Combust. Flame 151 (2007) 104–119.
53. J. Natarajan; Y. Kochar; T. Lieuwen; J. Seitzman, Proc. Combust. Inst. 32 (2009) 1261–1268.
54. M. P. Burke; M. Chaos; F. L. Dryer; Y. Ju, Combust. Flame 157 (2010) 618–631.
55. M. U. Alzueta; R. Bilbao; P. Glarborg, Combust. Flame 127 (2001) 2234–2251.
56. G. Mittal; C. J. Sung; M. Fairweather; A. S. Tomlin; J. F. Griffiths; K. J. Hughes, Proc. Combust. Inst. 31 (2007) 419–427.

Table 1. The syngas combustion mechanisms investigated, and the number of species and reactions in these reaction mechanisms. The numbers in parenthesis indicate the corresponding figures in the original mechanisms. All mechanisms can handle N_2 bath gas and the table indicates if the mechanism can also cope with Ar and He bath gases. The values of the average error function for all mechanisms are given for six cases: A – Ignition delay times, all diluents (1041 data points/102 datasets; 122/14 for Dagaut-2003), B – Flame velocities, all diluents except He (1725/175; 1719/174), C – Species concentrations, all profiles (1123/54; 687/47), D – Overall results, all diluents except He (3389/331; 2964;244), E – Overall results, all diluents including He (4127/358). Overall results are indicated by bold letters, results for Dagaut-2003 for a reduced data subset in italics. The top 3 mechanisms in each category are highlighted green, based on the complete subset of data.

No.	Mechanism ID	Ref.	Species number (orig.)	Ar/He	Reactions number (orig.)	Average error function values				
						A	B	C	D	E
1	NUIG-NGM-2010	[24]	15 (293)	x/x	41 (1593)	30.9	9.9	19.4	**17.9**	**17.3**
2	Kéromnès-2013	[18]	15 (17)	x/x	49	44.6	9.4	24.5	**22.7**	**21.4**
3	Davis-2005	[11]	14	x/x	38	58.4	6.8	25.3	**25.8**	**24.2**
4	Li-2007	[22]	15 (21)	x/x	45 (93)	56.3	8.3	39.9	**28.2**	**26.7**
5	USC-II-2007	[23]	14 (111)	x/–	48 (784)	70.9	7.7	20.8	**29.3**	–
6	SanDiego-2014	[25]	15 (50)	x/x	37 (244)	31.5	19.0	77.0	**32.3**	**31.5**
7	CRECK-2012	[17]	14	x/x	34	42.6	11.2	88.6	**33.5**	**31.6**
8	Li-2015	[19]	14	x/x	37	21.3	8.1	147.9	**35.0**	**33.2**
9	Starik-2009	[16]	16	x/x	44	40.0	18.1	106.1	**39.2**	**36.9**
10	GRI3.0-1999	[7]	15 (53)	x/–	48 (325)	85.2	8.5	81.1	**44.0**	–
11	Rasmussen-2008	[15]	15 (24)	x/–	59 (105)	98.7	19.8	105.1	**58.0**	–
12	SaxenaWilliams-2006	[13]	14	x/x	30	88.5	7.7	170.1	**59.1**	**55.2**
13	Sun-2007	[14]	15	x/x	48	150.2	8.4	110.4	**68.7**	**64.7**
14	Ahmed-2007	[21]	14 (246)	x/–	37 (1284)	116.8	23.3	186.7	**78.8**	–
15	Zsély-2005	[12]	13	x/–	44	276.9	10.4	110.3	**108.8**	–
16	Dagaut-2003	[20]	13 (132)	–/–	34 (922)	*26.2*	*8.0*	*46.9*	***16.9***	–

Table 2. Our database of syngas experimental data used in the mechanism comparison by type of measurement and experimental facility. The numbers of included datasets and data points, regions of pressure p, temperature T, equivalence ratio φ and CO/H_2 ratio covered are summarized and the diluents are listed and ordered by their frequency. For flame velocity measurements, temperature T refers to the initial temperature of the mixture and values in brackets refer to measurements in helium. Excluded shock tube ($T \leq 1000$ K) and JSR ($T < 1000$ K) data points are not shown here.

Type of measurement Experimental facility	**No. of datasets**	**No. of data points**	***p* [atm]**	***T* [K]**	***φ***	**CO/H_2**	**Diluents**
Ignition delay time measurements	**102**	**1041**	**0.75–49.4**	**896–2870**	**0.1–6.1**	**0.18–243.4**	**Ar; Ar/N_2; Ar/CO_2; N_2/CO_2; N_2/Ar/CO_2**
Shock tube	50	532	0.75–32.8	1001–2870	0.5–6.1	0.25–243.4	Ar; N_2; N_2/CO_2
Shock tube using VTIM	12	122	1.5–20.5	896–1308	1	0.43–2.82	Ar; Ar/N_2; Ar/CO_2
Rapid compression machine using VTIM	40	387	2.6–49.4	933–1145	0.1–1	0.18–19	Ar/N_2; N_2/CO_2; N_2/Ar/CO_2
Flame velocity measurements	**175 (27)**	**1725 (238)**	**0.5–25 (1–40)**	**293–700 (293–600)**	**0.3–6.8 (0.5–5)**	**0.05–184, pure CO, pure H_2 (0.1–355)**	**N_2; N_2/CO_2; N_2/H_2O; CO_2/H_2O; CO_2; H_2O; Ar (He; He/H_2O; He/CO_2)**
Flame cone method/ Bunsen burner	68 (4)	801 (33)	0.95–1 (10–15)	293–700 (300–600)	0.3–6.8 (0.6–1.2)	0.05–184, pure CO (0.1–4)	N_2; N_2/H_2O; N_2/CO_2; CO_2/H_2O; H_2O (He; He/CO_2)
Outwardly/ spherically propagating flame	60 (20)	563 (179)	0.5–25 (1–40)	293–500 (293–424)	0.4–5.8 (0.55–5)	0.33–335, pure H_2 (1–355)	N_2; N_2/H_2O; N_2/CO_2; CO_2; Ar (He; He/H_2O; He/CO_2)
Counterflow twin-flame	18 (–)	112 (–)	1 (–)	298–323 (–)	0.4–1.4 (–)	0.11–99 (–)	N_2; N_2/H_2O (–)
Heat flux method	29 (3)	249 (26)	0.94–4 (1–9)	298–302 (298)	0.4–5.6 (0.5–0.7)	0.18–19 (0.18–1)	N_2; N_2/CO_2 (He)
Concentration measurements	**54**	**1123**	**1–499.5**	**785–1495**	**0.0005–3.03**	**1–3, pure CO**	**N_2/H_2O; Ar; N_2**
Flow reactor concentration–time profiles	35	508	1–9.6	943–1138	0.001–1.7	pure CO	N_2/H_2O
Flow reactor outlet concentrations	9	125	1.04–1.05	785–1479	0.0005–3.03	pure CO	N_2/H_2O
Shock tube outlet concentrations	7	436	20.9–449.5	995–1495	0.47–1.0	2.4–3	Ar
Jet-stirred reactor outlet concentrations	3	54	1	1000–1400	0.1–2	1	N_2

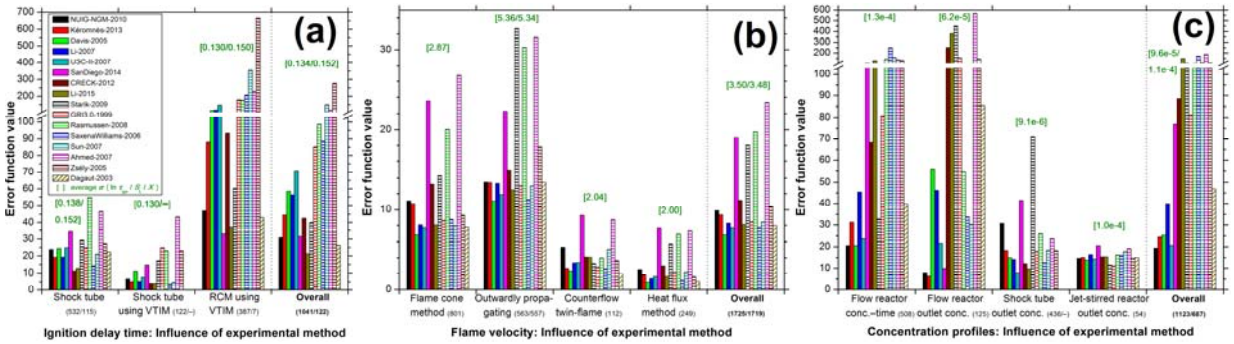

Fig. 1. Errors of the reproduction of (a) ignition delay times, (b) flame velocities and (c) concentration profiles according to the respective type of experiment. The numbers in parentheses indicate the number of data points included. The second values refer to the number of data points for Dagaut-2003. Green numbers above the columns refer to average σ values for each subset. All diluents except He.

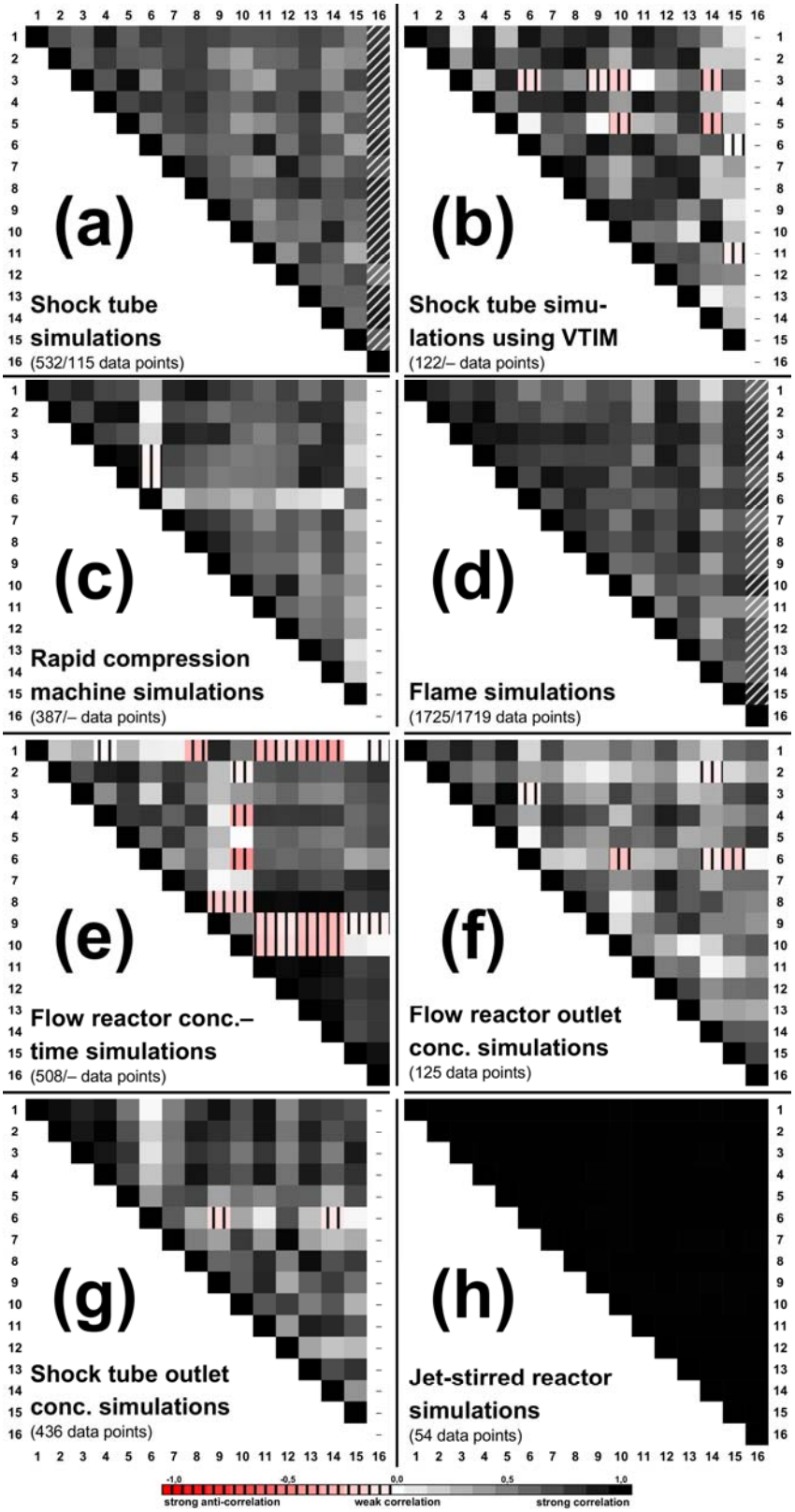

Fig. 2. Matrix of correlations of absolute deviation values for all types of simulations (panels (a) to (h)). Diagonal stripes were used when the comparison was based on a different dataset (for Dagaut-2003: Ar and He excluded, otherwise just measurements in He were excluded). The *x*-

and *y*-axes in each panel refer to the identifying mechanism number (see Table 1).

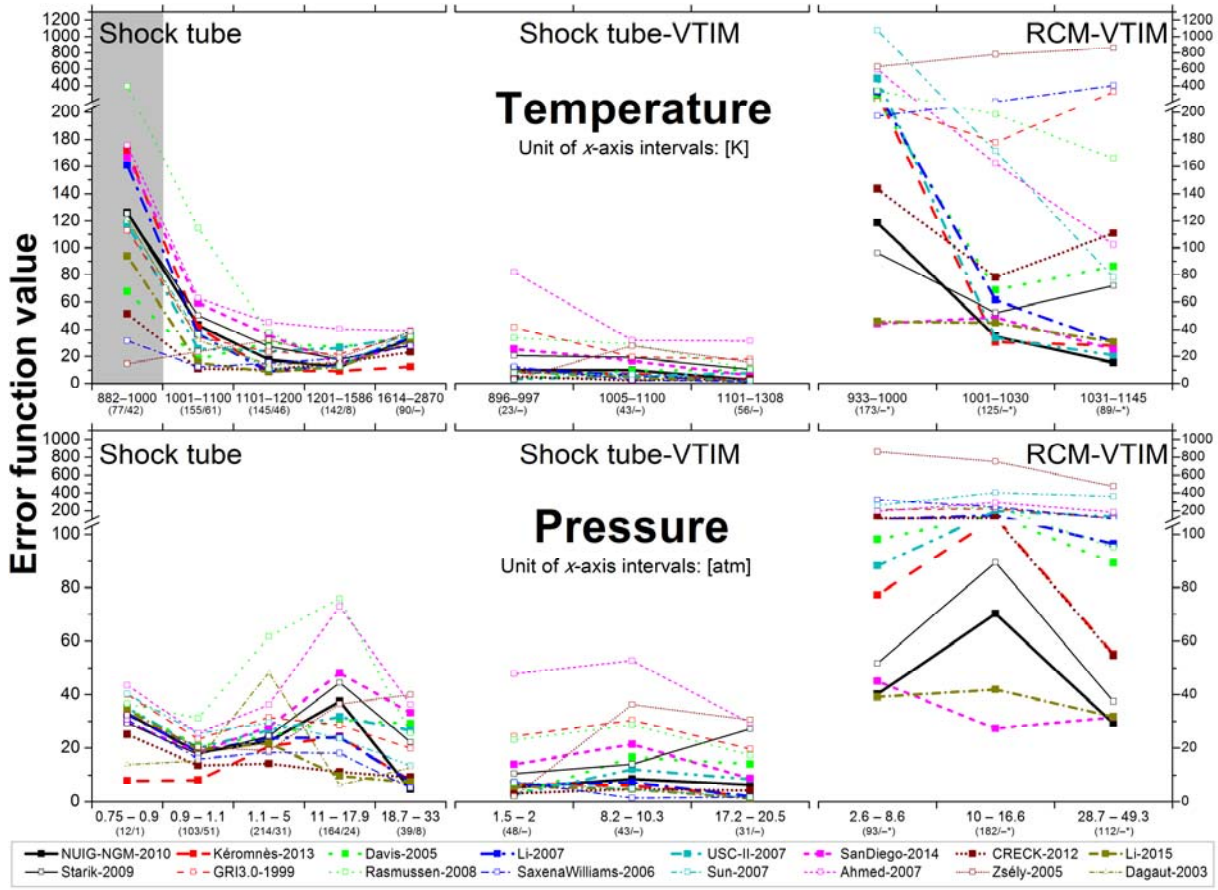

Fig. 3. Performance of the mechanisms for various ranges of temperature and pressure with respect to ignition delay time. Each plot shows the results for shock tubes (left part), shock tubes modeled using VTIM (middle part) and rapid compression machines (right part). Gray shaded: shock tube experiments at $T \leq 1000$ K that were excluded in the general comparison. * – RCM results for Dagaut-2003 not shown (see text).

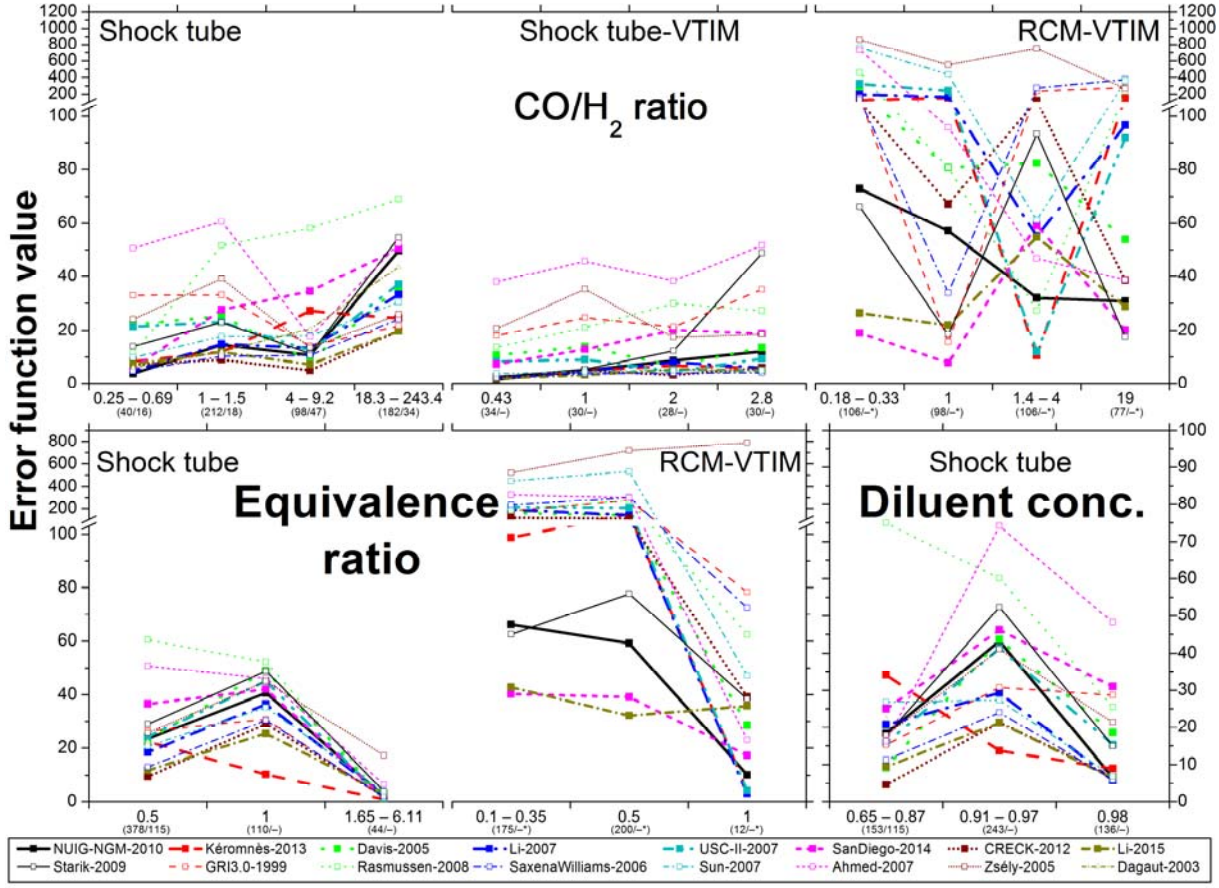

Fig. 4. Performance of the mechanisms for various ranges of CO/H_2 ratio, equivalence ratio and diluent concentration with respect to ignition delay time. Each plot shows the results for shock tubes (left part), shock tubes modeled using VTIM (middle part) and rapid compression machines (right part), if available. * – RCM results for Dagaut-2003 not shown (see text).

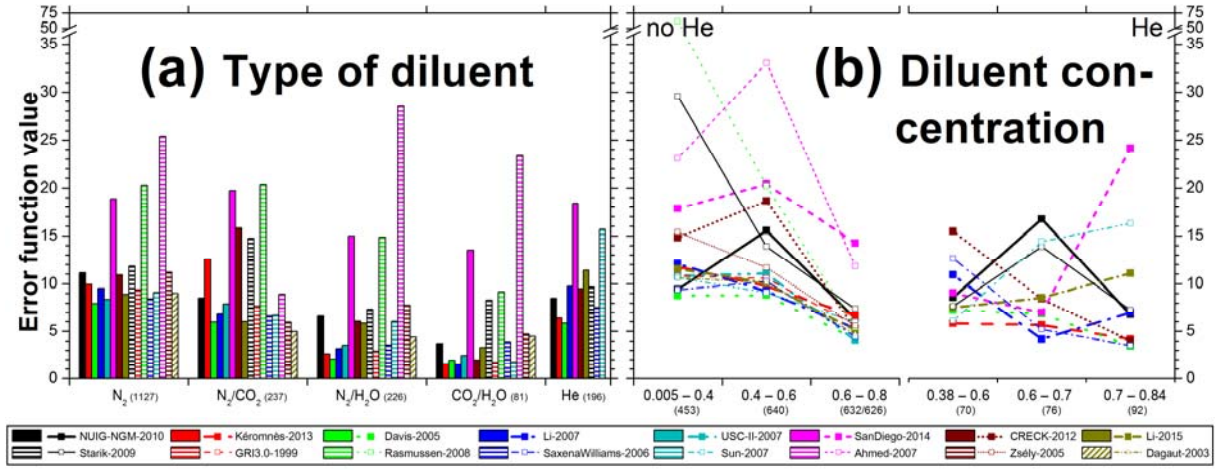

Fig. 5. (a) Error of the reproduction of the flame velocity according to the type of diluent system for the five most important diluent systems, (b) Performance of the mechanisms for various ranges of diluent concentration with respect to flame velocity for He-free (left part) and He-containing mixtures (right part).

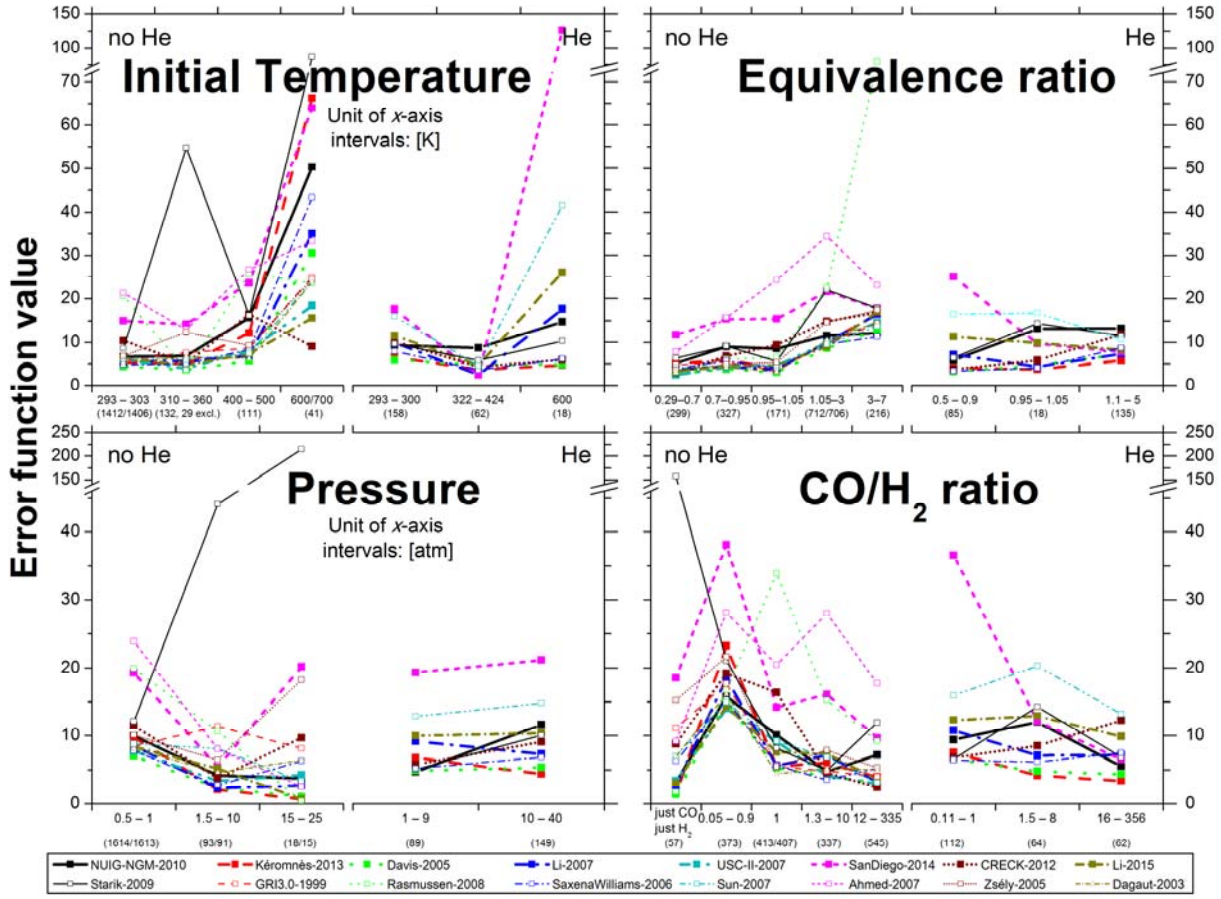

Fig. 6. Performance of the mechanisms for various ranges of temperature, pressure and equivalence ratio and diluent concentration with respect to flame velocities. Each plot shows the results for He-free mixtures (left part) and He-containing mixtures (right part). Note that those mechanisms without He as a defined species do not appear in the right part of the plots.

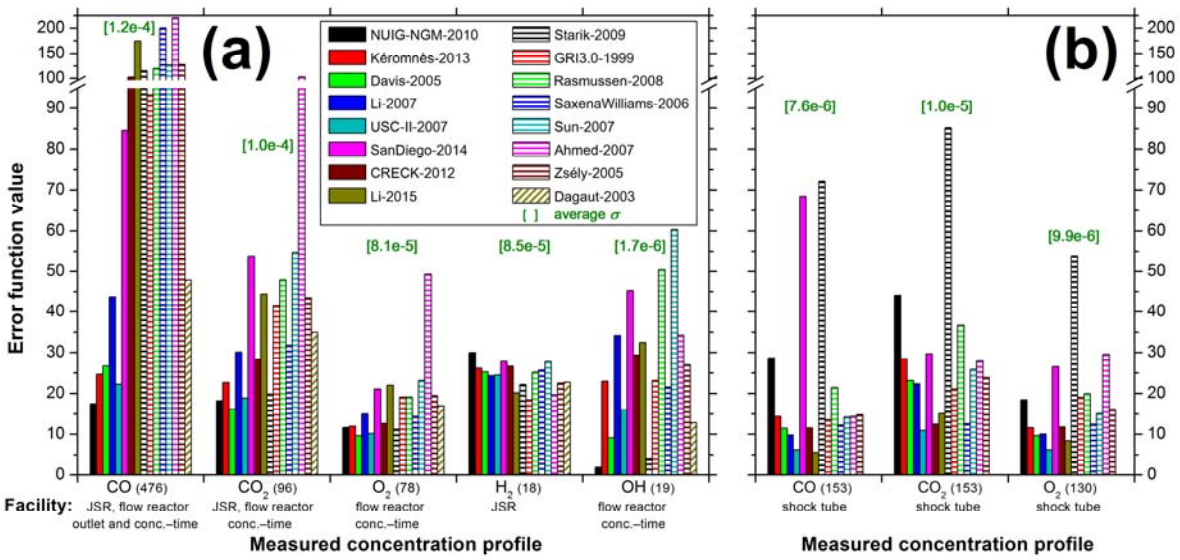

Fig. 7. Errors of the reproduction of concentration profiles depending on the measured profile for (a) JSR and flow reactor data and (b) shock tube data. The type of experimental methods in which the corresponding measurements were facilitated are given for each profile. Green numbers above the columns refer to average σ values for each subset.

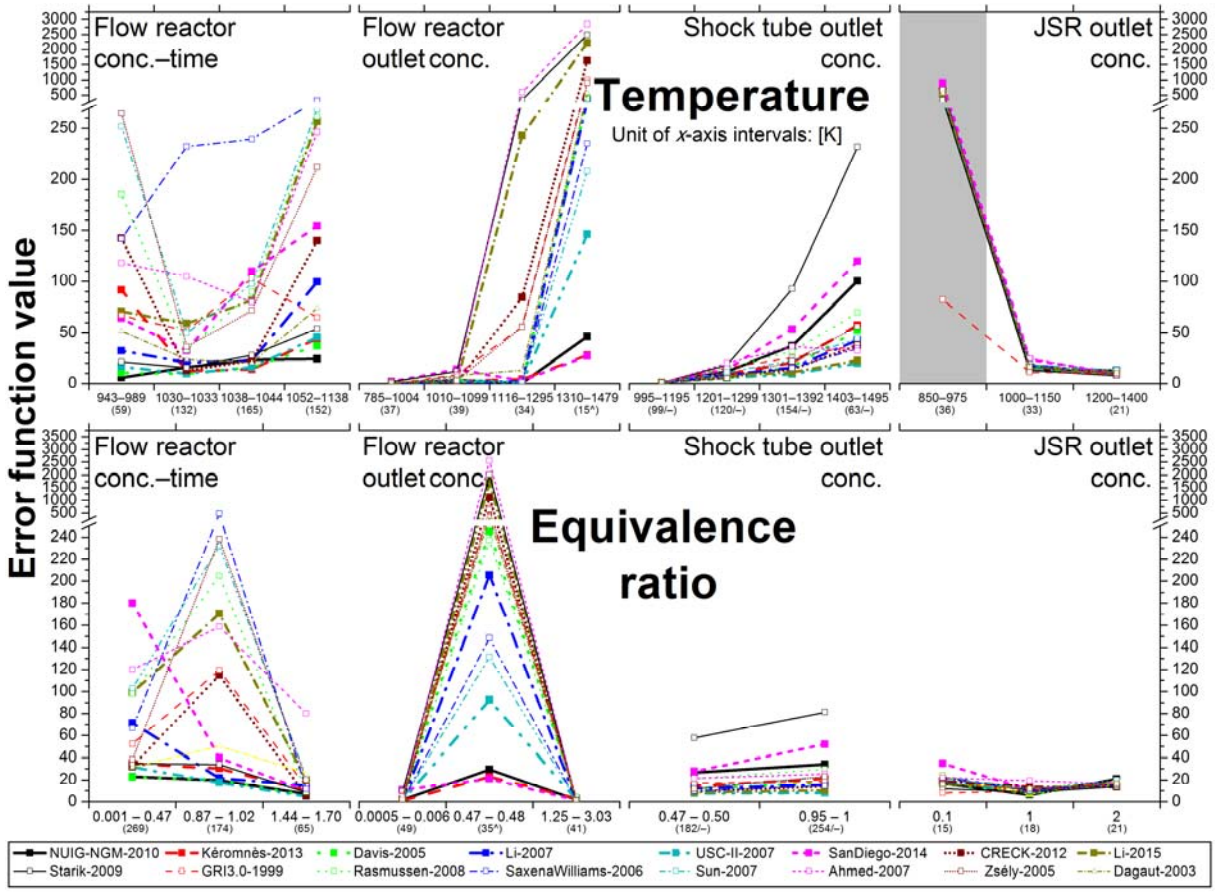

Fig. 8. Performance of the mechanisms for various ranges of temperature (top) and equivalence ratio (bottom) with respect to measured concentrations. Each row shows the results corresponding to flow reactor concentration–time data (leftmost panel), and outlet concentration profiles from flow reactors (middle left), shock tubes (middle right) and jet-stirred reactors (rightmost).

^ Interval in which the badly reproduced high-*T* data of dataset *x30001021.xml* fall (see text).

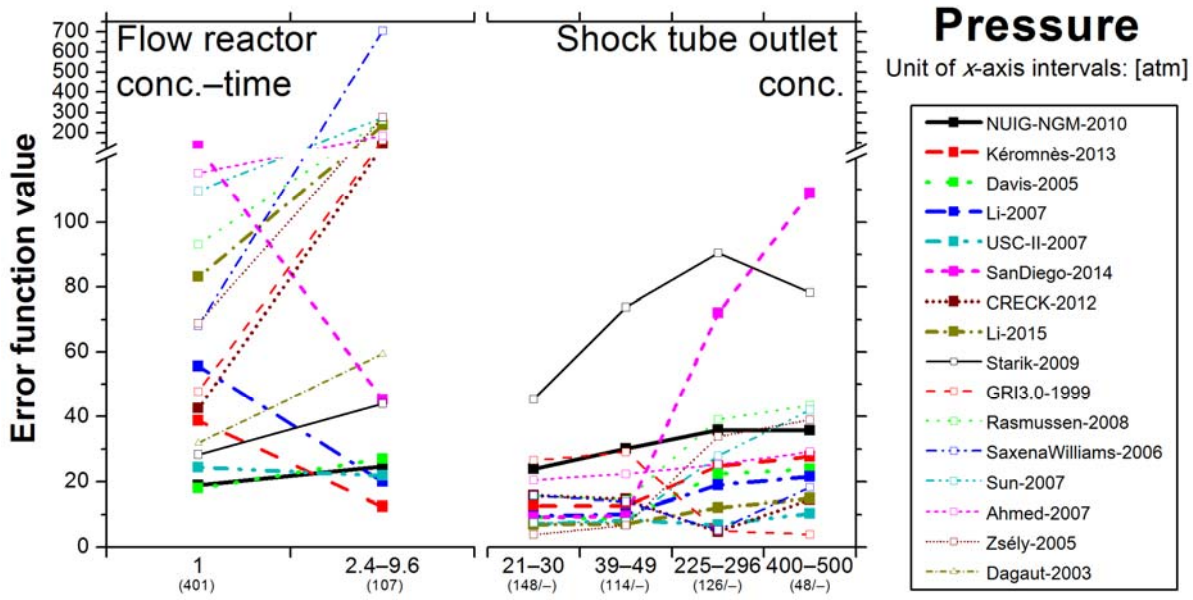

Fig. 9. Performance of the mechanisms for various ranges of pressure with respect to flow reactor concentration–time data (left) and outlet concentration profiles from shock tubes (right).

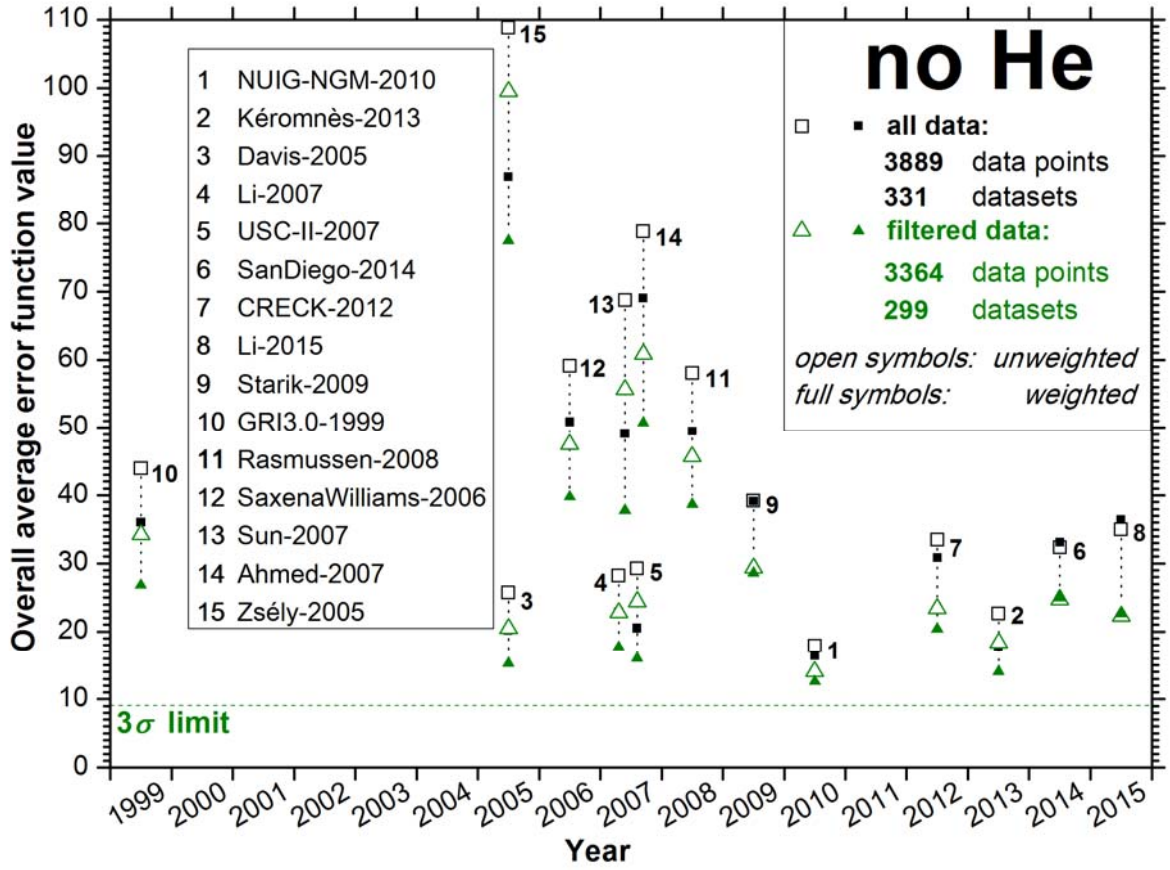

Fig. 10. Un-weighted (open symbols) and weighted (full symbols) overall performance of the mechanisms taking into account all experimental data ("all data", black squares) and only those that were reproduced by at least one mechanism within $E_i \leq 9$ ("filtered data", green triangles) vs. year of publication. All diluents except He. The filtering and weighting approaches as well as the relevance of the 3σ limit are described in Section 6.4.

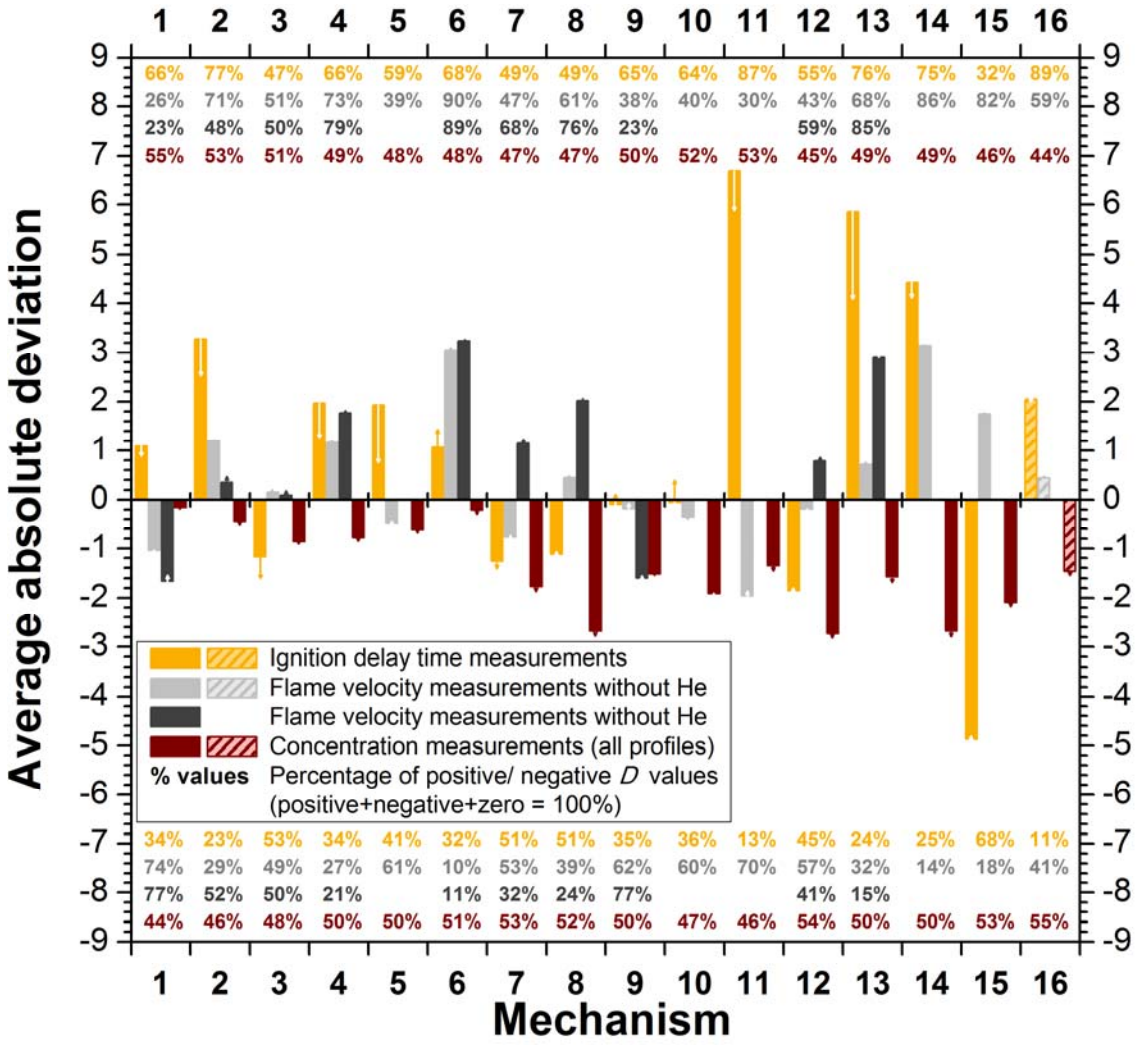

Fig. 11. Summary of average absolute deviation values of all mechanisms for ignition delay times, flame velocities and concentrations. The columns are striped if the number of data points differs from the other mechanisms (for Dagaut-2003, #16: all diluents except Ar and He). Identifying numbers were used for the mechanisms as in Table 1. Arrows indicate the changes of *D* values due the introduction of weighting as described in Section 6.4.

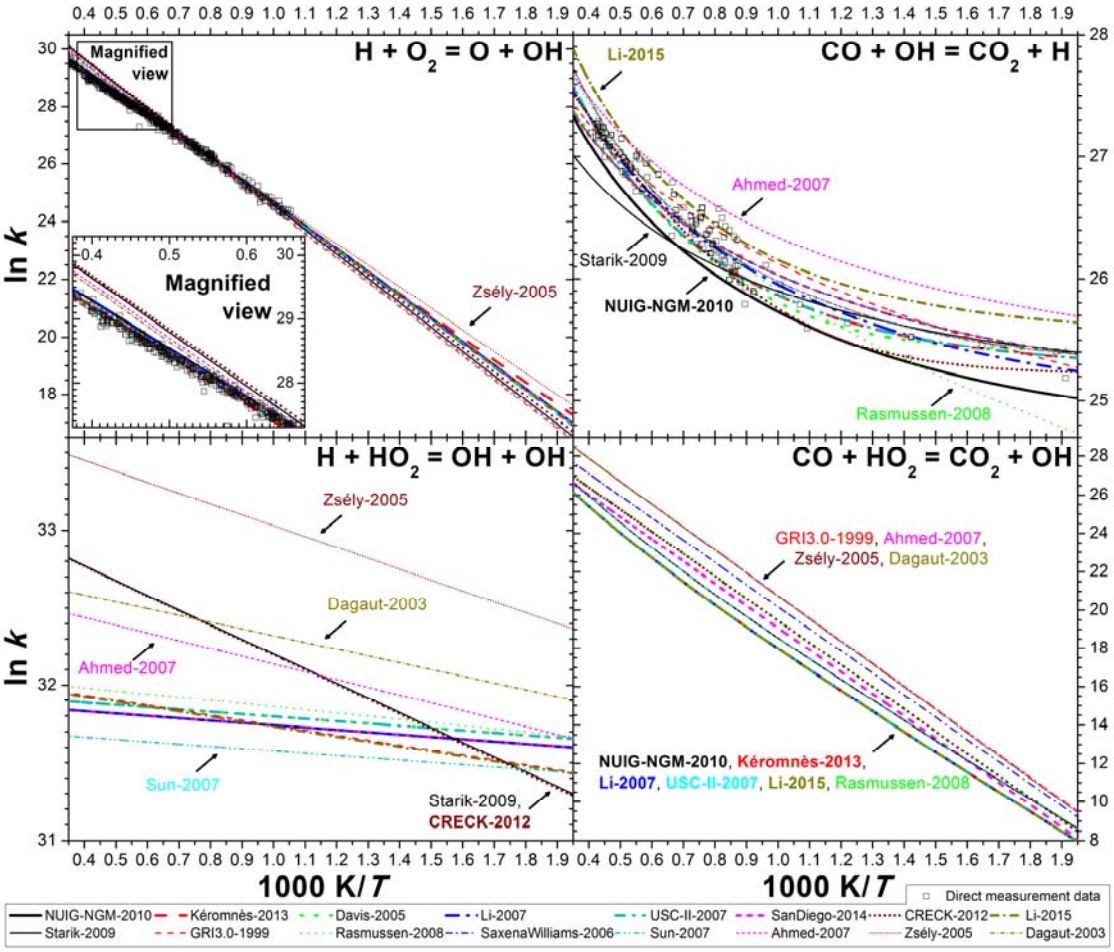

Fig. 12. Comparison of published reaction rate coefficients for the most sensitive reactions: H + O_2 = O + OH (top left panel), CO + OH = CO_2 + H (top right), H + HO_2 = OH + OH (bottom left) and CO + HO_2 = CO_2 + OH (bottom right). References to the direct measurements of reaction rate coefficients (black open squares) can be found in Table J of the Supplemental Material. The units of *k* are expressed in terms of cm, mol and s.

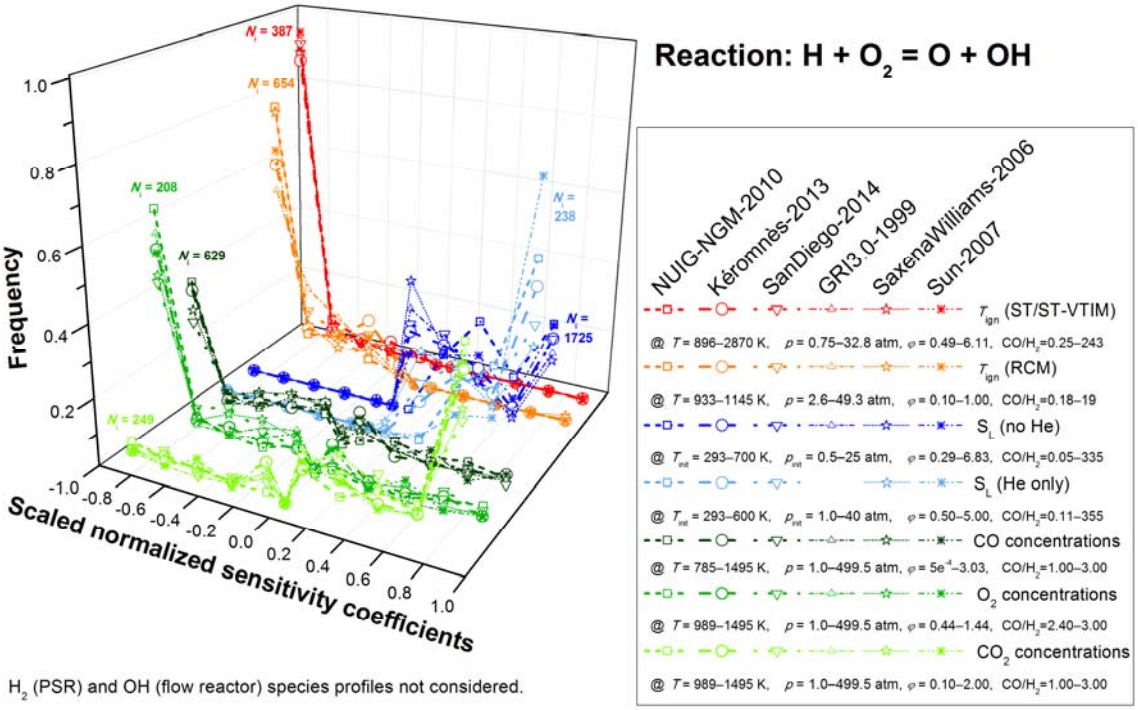

Fig. 13: Frequencies of sensitivity coefficients of reaction H + O_2 = O + OH for various types of measurements. The sensitivity analysis for this reaction was performed with the mechanisms NUIG-NGM-2010, Kéromnès-2013, SanDiego-2014, GRI3.0-1999, SaxenaWilliams-2006 and Sun-2007. The experimental conditions are given for each type of measurement.

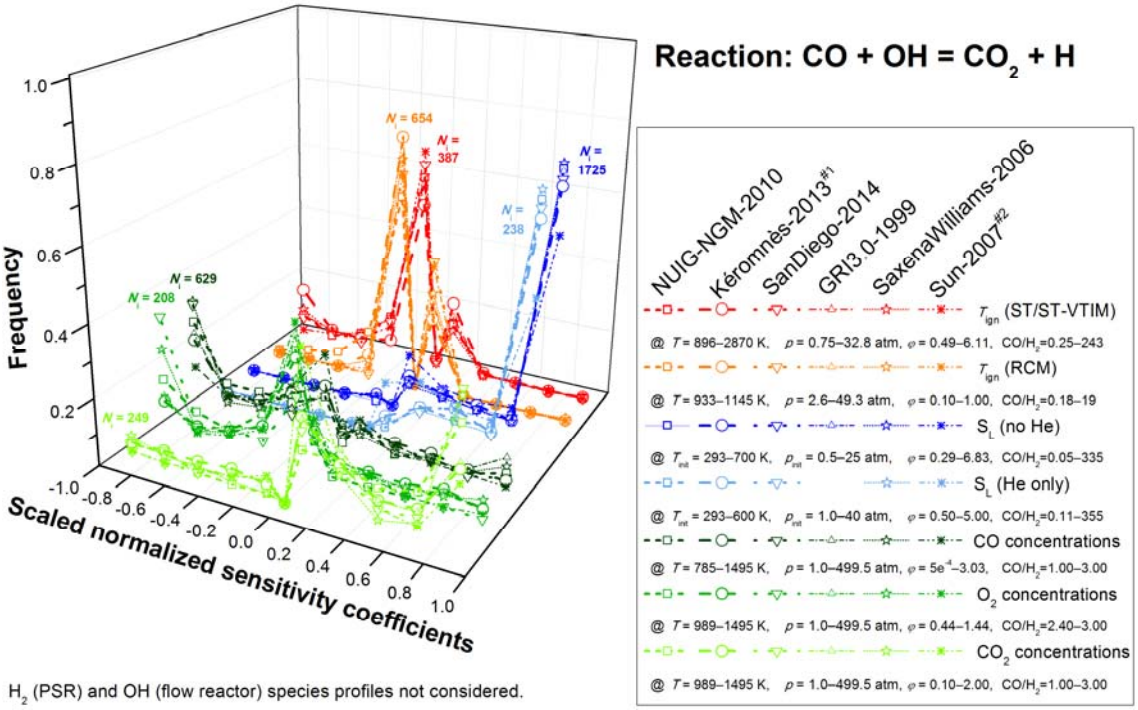

Fig. 14: Frequencies of sensitivity coefficients of reaction CO + OH = CO_2 + H for various types of measurements. The sensitivity analysis for this reaction was performed with the mechanisms NUIG-NGM-2010, Kéromnès-2013, SanDiego-2014, GRI3.0-1999, SaxenaWilliams-2006 and Sun-2007. The experimental conditions are given for each type of measurement.

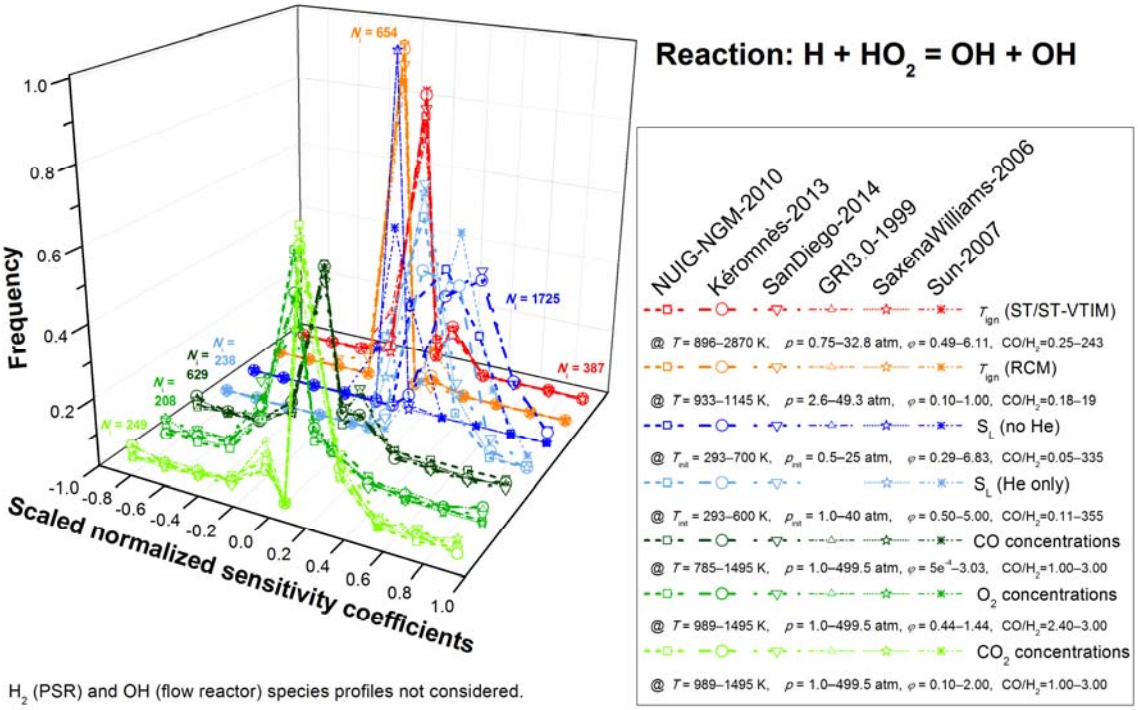

Fig. 15: Frequencies of sensitivity coefficients of reaction H + HO_2 = OH + OH for various types of measurements. The sensitivity analysis for this reaction was performed with the mechanisms NUIG-NGM-2010, Kéromnès-2013, SanDiego-2014, GRI3.0-1999, SaxenaWilliams-2006 and Sun-2007. The experimental conditions are given for each type of measurement.

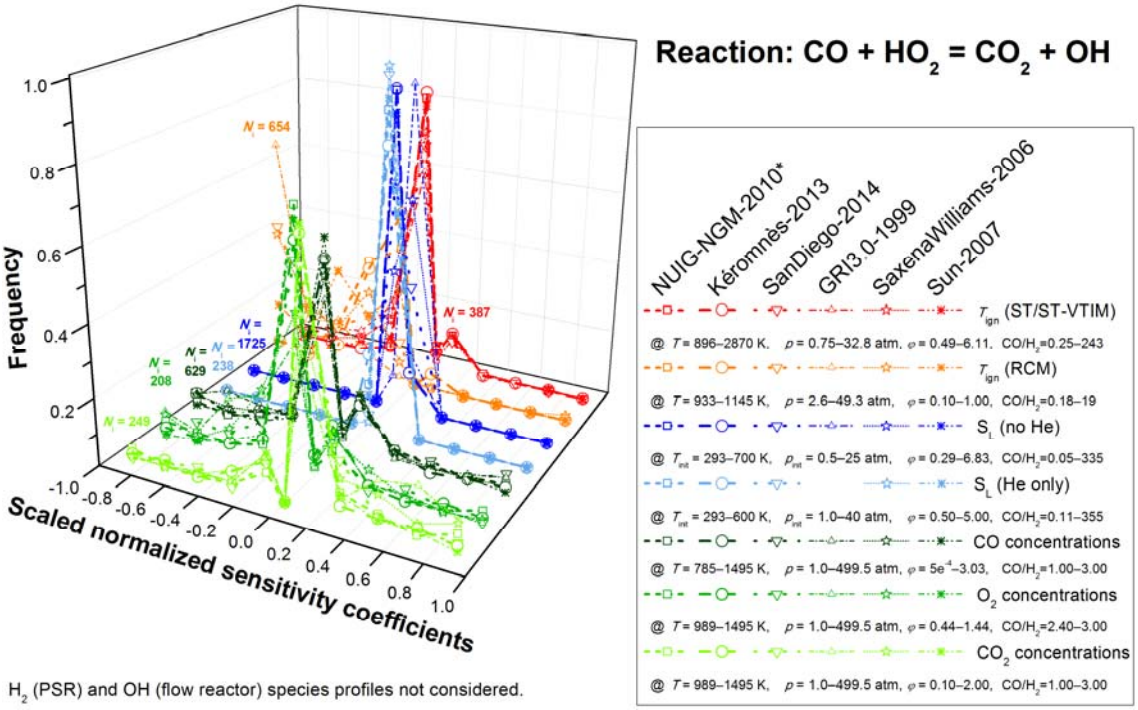

Fig. 16: Frequencies of sensitivity coefficients of reaction CO + HO_2 = CO_2 + OH for various types of measurements. The sensitivity analysis for this reaction was performed with the mechanisms NUIG-NGM-2010, Kéromnès-2013, SanDiego-2014, GRI3.0-1999, SaxenaWilliams-2006 and Sun-2007. The experimental conditions are given for each type of measurement.