

# Investigating novel strategies for parameter optimization on a methanol/NO<sub>x</sub> combustion mechanism

M. Kovács<sup>1</sup>, M. Papp<sup>1</sup>, T. Nagy<sup>\*,2</sup>, T. Turányi<sup>1</sup>

<sup>1</sup>Institute of Chemistry, ELTE Eötvös Loránd University, Budapest, Hungary

<sup>2</sup>Institute of Materials and Environmental Chemistry, ELKH Research Centre for Natural Sciences, Budapest, Hungary

## Abstract

Local sensitivity coefficients and parameter uncertainties are the usual measures used for selecting parameters for the optimization of combustion kinetic models. We identified two further factors and construct four novel measures for parameter ranking and benchmarked them against two known measures in a 10-parameter hierarchical optimization of a methanol/NO<sub>x</sub> mechanism on a large experimental data set. It was found that measures that did not incorporate uncertainty information could reduce the error function initially the most steeply, but in the long run the other methods performed better, though similarly. Regarding posterior uncertainties, one of the novel strategies based on error function derivative and uncertainty information performed the most reliably and can be recommended for future use.

## Introduction

Detailed chemical kinetic mechanisms are widely used to simulate combustion experiments for both industrial and scientific purposes. Once assembled their performance can be improved by updating their parameters using not yet considered or newly published experimental data and theoretical results. First kinetic parameter optimization studies on combustion kinetic mechanisms were done by Frenklach et al. [1,2] and Sheen and Wang [3,4]. A similar method was applied by Pitsch and his coworkers [5]. A method that allowed efficient global optimization of all three Arrhenius parameters in their joint uncertainty domain [6,7] and considered both direct experimental and theoretical determinations and indirect measurements was developed by Turányi et al. [8]. This methodology has been used for the development of several detailed reaction mechanisms (see e.g. refs. [9–12]).

Detailed combustion mechanisms usually contain large number of uncertain parameters and for the most important fuels, large amount of experimental data are available, which renders their optimization a computationally challenging problem. To overcome this, a systematic hierarchical optimization strategy was introduced and applied successfully by Turányi and his coworkers (see e.g. refs. [9–12]). In their method, groups of influential parameters and groups of corresponding sensitive experimental datasets are introduced in the optimization procedure in a stepwise manner, allowing larger parameter changes to be accomplished with lower number of simulations. The method made ranking of parameters and experimental datasets based exclusively on the local sensitivity coefficients. It was recognized by Frenklach et al. that parameters can be tuned only within their prior uncertainty range in a physically meaningful way, thus the absolute value of the product of the sensitivity coefficient and parameter uncertainty is a better measure for parameter selection [13].

The ultimate aim of parameter optimization is to reduce the value of an appropriately defined objective function. Thus, one can also consider the error of experimental data and the deviation between the experimental data and its simulation result to define novel measures for parameter ranking. In this study, beside sensitivity and sensitivity×parameter uncertainty, four new impact measures are proposed and tested in an optimization study of a detailed methanol/NO<sub>x</sub> combustion mechanism.

## The investigated combustion system

The methanol/NO<sub>x</sub> combustion mechanism was chosen to test the novel parameter ranking strategies. Methanol is a promising alternative to fossil transportation fuels, and its interactions with nitrogen oxides (NO<sub>x</sub>) in combustion systems are also important due to environmental regulations. In a recent publication [14], we collected all available experimental data corresponding to methanol/NO<sub>x</sub> combustion, and tested 17 mechanisms against them.

**Table 1** The number and the condition ranges of concentration data collected from methanol/NO<sub>x</sub> combustion experiments (see text for abbreviations)

Experiment	Data sets	Data points	$T / K$	$p / atm$	$\phi$
JSR	72	765	640–1870	0.92–10	0.3–1.34
TFR	160	1648	298–1420	0.99–1.4	0.01–13.5
ST	11	139	1142–1502	0.46–0.66	0.46–2
All	243	2552	298–1870	0.46–10	0.01–13.5
Used	225	2373	298–1502	0.46–10	0.01–13.5

Table 1 summarizes the numbers of collected datasets and points and the ranges of experimental conditions. All the related experiments were concentration measurements carried out in jet stirred reactors (JSR), tubular flow reactors (TFR) and shock tubes (ST). We showed that even the best performing models have

\* Corresponding author: [nagy.tibor@ttk.hu](mailto:nagy.tibor@ttk.hu)

significant inaccuracies and high uncertainties [14]. Some datasets were excluded from the study as none of the mechanisms could reproduce them within  $10\sigma$  experimental uncertainty. The overall numbers for all data and only for the used data are also shown in the table. For more details on the data collection see ref. [14]. In this study, only the kept 2373 data points in 225 datasets were used, which are sufficiently many for a meaningful optimization study. Based on our comparison of recent methanol/NOx mechanisms [14], we selected the ELTE+Glarborg mechanism for optimization due to its relatively good performance and low uncertainty. This mechanism was obtained from mechanism Glarborg-2018 [15] by updating it with optimized rate parameters from our previous works [9–12].

### Optimization method

The collected 243 datasets were stored in 74 ReSpecTh Kinetics Data (RKD) v2.3 data files [16], which is the data format of the ReSpecTh database [17] and which had been developed from the PrIme Kinetics Data Format [18]. For the parameter optimizations, we applied code Optima++ [19], which implements the optimization method developed by Turányi et al. [8], and calls code OpenSMOKE++ [20,21] for the simulations. The method minimizes the following objective function:

$$E(\mathbf{P}) = \frac{1}{N} \sum_{i=1}^N \frac{1}{N_i} \sum_{j=1}^{N_i} \left( \frac{Y_{ij}^{\text{sim}}(\mathbf{P}) - Y_{ij}^{\text{exp}}}{\sigma_{ij}^{\text{exp}}} \right)^2 \quad (1)$$

Here  $N$  is the number of datasets,  $N_i$  is the number of data points in the  $i$ -th dataset, vector  $\mathbf{P}$  contains the used forms of rate parameters (e.g.  $\ln A$ ,  $n$ ,  $E/R$ ), values  $Y_{ij}^{\text{exp}}$  and  $\sigma_{ij}^{\text{exp}}$  are the optionally transformed  $j$ -th data point in the  $i$ -th dataset and its standard deviation, respectively and  $Y_{ij}^{\text{sim}}$  is its simulated value. Thus,  $E(\mathbf{P})$  measures the average of the squared deviation of the simulation results from the experimental data relative to the standard deviation of the experimental data.  $E(\mathbf{P})$  is around one for a perfect model.

In this study, only the pre-exponential factors of selected reactions were optimized in such a way that the Arrhenius curves ( $k(T)$ ) always stayed within their prior uncertainty band (i.e.  $[k_{\min}(T); k_{\max}(T)]$ ). The uncertainty parameter is defined as the radius of a symmetric uncertainty range around the nominal  $k^0(T)$  value on  $\log_{10}$  scale:

$$f_{\text{prior}}(T) = \log_{10} \frac{k_{\max}(T)}{k^0(T)} = \log_{10} \frac{k^0(T)}{k_{\min}(T)} \quad (2)$$

It was determined for each reaction based on directly measured experimental data and theoretical determinations taken from the NIST Chemical Kinetics Database [22]. More details can be found in ref. [14].

### Standard and novel measures of parameter impact

The identification of parameters that can be tuned to improve the model performance the most efficiently is a crucial step prior to the actual optimization task. A common practice is to apply local sensitivity analysis to

quantify the impact of parameters on the simulation results. Local sensitivity coefficients are defined as:

$$S_{ij,l} = \frac{\partial Y_{ij}^{\text{sim}}}{\partial P_l} \quad (3)$$

$P_l$  is the  $l$ -th parameter to be optimized. Assuming that all experimental data are of equal importance, for each ( $ij$ )-th data point the  $S_{ij,l}$  values are normalized by the largest absolute  $S_{ij,l}$  value.

$$\tilde{S}_{ij,l} = \frac{S_{ij,l}}{\max_m |S_{ij,m}|} \quad (4)$$

This value can be used to define an overall sensitivity ( $I_l^S$ ) of all considered simulation results to the  $l$ -th parameter:

$$I_l^S = \sqrt{\frac{1}{N} \sum_{i=1}^N \frac{1}{N_i} \sum_{j=1}^{N_i} \tilde{S}_{ij,l}^2} \quad (5)$$

However, parameters are allowed to be tuned only within their prior uncertainty range, thus Frenklach et al. introduced the parameter impact as the product of the sensitivity coefficient and parameter uncertainty [13]. Based on this measure an overall impact ( $I_l^{\text{SU}}$ , U for uncertainty) can be defined:

$$I_l^{\text{SU}} = \sigma_l I_l^S \quad (6)$$

Here, we introduced  $\sigma_l$  as the standard deviation of parameter  $P_l = \ln A_l$ , and assumed it to be one third of the uncertainty radius of the  $l$ -th rate coefficient:

$$\sigma_l = \frac{\ln 10}{3} \cdot f_{\text{prior},l} \quad (7)$$

As the ultimate aim of model optimization is to minimize the objective function (Eq. (1)), the impact on  $Y_{ij}^{\text{sim}}$  needs to be scaled down by the corresponding  $\sigma_{ij}^{\text{exp}}$  experimental error ("E"), which leads to definition of the following new overall measure for parameter impact:

$$I_l^{\text{SUE}} = \sigma_l \sqrt{\frac{1}{N} \sum_{i=1}^N \frac{1}{N_i} \sum_{j=1}^{N_i} \left( \frac{S_{ij,l}}{\sigma_{ij}^{\text{exp}}} \right)^2} \quad (8)$$

One can also consider the largest effect that can be induced on the summed terms in objective function by  $\pm\sigma_l$  parameter variation using Taylor approximation:

$$I_l^{\text{SUED}} = \sqrt{\frac{1}{N} \sum_{i=1}^N \frac{1}{N_i} \sum_{j=1}^{N_i} \frac{S_{ij,l}^2 \sigma_l^2 + 2|S_{ij,l} D_{ij}| \sigma_l}{(\sigma_{ij}^{\text{exp}})^2}} \quad (9)$$

Here we introduced  $D_{ij}$  as a short notation for  $Y_{ij}^{\text{mod}}(\mathbf{P}_0) - Y_{ij}^{\text{exp}}$  and assumed model  $\mathbf{Y}(\mathbf{P})$  to be linear.

The most obvious local measure for parameter ranking seems to be the absolute value of the partial derivative of the objective function (G for gradient):

$$I_l^G = \left| \frac{1}{N} \sum_{i=1}^N \frac{1}{N_i} \sum_{j=1}^{N_i} \frac{2S_{ij,l} D_{ij}}{(\sigma_{ij}^{\text{exp}})^2} \right| \quad (10)$$

By averaging the signed terms, this measure can take into account the correlated change in various model outputs. However, again, this should be multiplied with the parameter uncertainty to obtain a more reliable measure:

$$I_l^{\text{GU}} = \sigma_l I_l^G \quad (11)$$

## Results of reaction ranking

Brute force local sensitivity analysis on the simulation results of all considered 2373 data points was carried out by symmetrically perturbing all the 562 rate coefficients (25 of them are low-pressure limit (LP) rate coefficients) of the ELTE+Glarborg mechanism with 5% of their nominal values. All the six parameter impacts were calculated and the 10 highest ranked reactions were identified in each case (see Table 2). Note that the optimal order of inclusion of datasets into the optimization could also be determined based on these measures, however, we wanted to focus on the parameter ranking performance of the impact measures thus the whole data collection was used in every optimization steps.

In the case of the 10 most important reactions identified by measure S (local sensitivity analysis), the perturbation of *A* parameters affects a large number of the simulation results significantly. These reactions without exception are also important in the combustion of simpler fuels, such as hydrogen (R1, R13, R9 LP), syngas (R24, R27, R26), methanol (R82, R83, R40) and H<sub>2</sub>/O<sub>2</sub>/NO<sub>x</sub> (R197), whereas none of the reactions specific to the methanol/NO<sub>x</sub> system showed up. These simpler systems have been investigated by our group; thus, the already optimized values of the *A* parameters and their posterior uncertainties were used as initial parameter and uncertainty values, respectively, in this study. Typically, these uncertainties are significantly lower than the prior uncertainties of the other, not yet optimized parameters

taken from the literature. Thus, even though these reactions are influential, their low uncertainties limit their potential to reduce the objective function. Furthermore, we do not intend to optimize these already well-known parameters as their uncertainties probably cannot be reduced further. The G strategy also selected almost exclusively non-system specific reactions (except for the 9<sup>th</sup> and 10<sup>th</sup>) which is due to the fact that this measure does not incorporate parameter uncertainty information.

On the contrary, the other four measures (SU, SUE, SUED and GU), that took parameter uncertainty into account, properly ranked high several reactions playing role in the interaction between carbon containing species and nitrogen oxides: R389, R362 and R363 are specific to the methanol/NO<sub>x</sub> system, whereas R368 is also important in the syngas/NO<sub>x</sub> system. The optimization of these reactions is expected to not only decrease the error function value efficiently, but the uncertainty of the parameters and model predictions as well. These four strategies also identified some of the previously discussed reactions of low uncertainty, but only at lower ranks. The GU strategy which also incorporates all the four factors similarly to SUED, gave the same group of reactions in the first 8 ranks as SUED.

## Results of parameter optimization

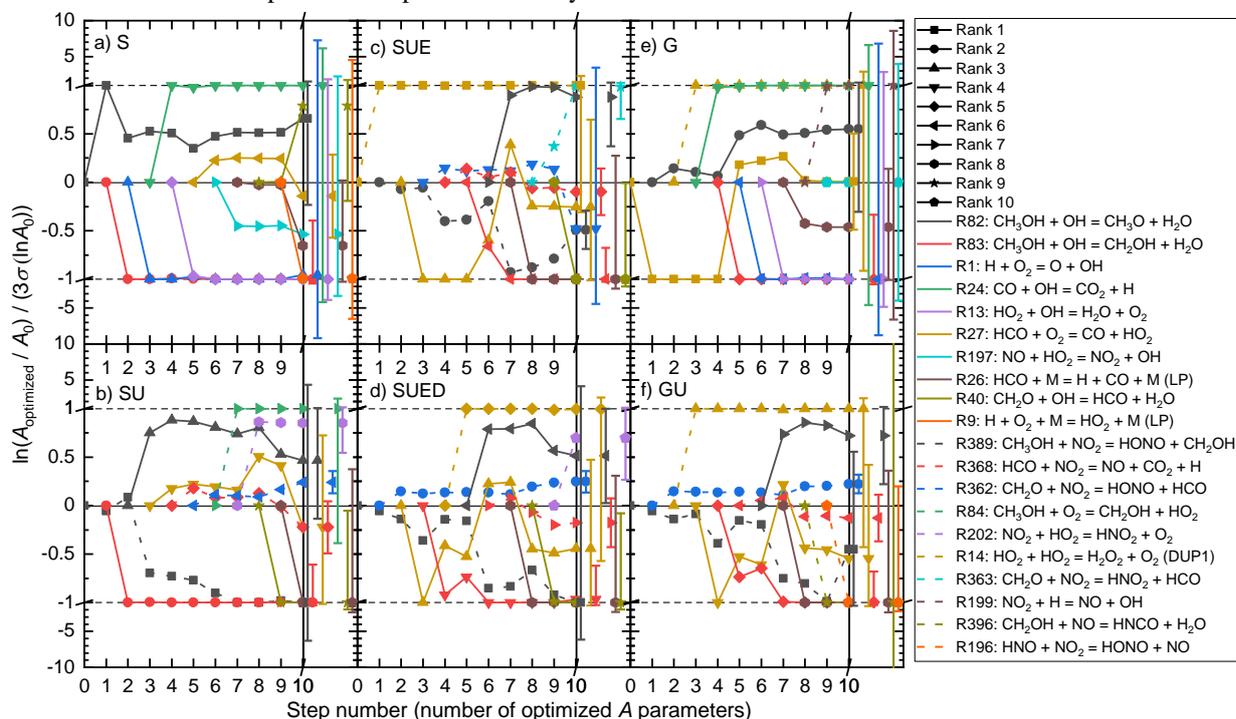
For each determined ranking, a 10-step sequential optimization of the ELTE+Glarborg mechanism was carried out by including the parameters one by one into the procedure and optimizing more and more parameters together.

**Table 2** The 10 most important reactions of the ELTE+Glarborg mechanism identified and ranked by the six parameter impact measures. LP: Low pressure limit, DUPl: 1<sup>st</sup> parameter set of a duplicate reaction

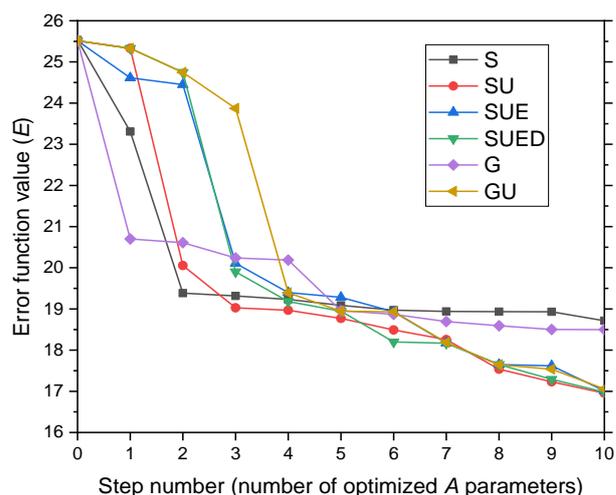
Rank	R#	Reaction	Impact	Rank	R#	Reaction	Impact		
S	1	R82	CH <sub>3</sub> OH + OH = CH <sub>3</sub> O + H <sub>2</sub> O	0.55	1	R389	CH <sub>3</sub> OH + NO <sub>2</sub> = HONO + CH <sub>2</sub> OH	0.67	
	2	R83	CH <sub>3</sub> OH + OH = CH <sub>2</sub> OH + H <sub>2</sub> O	0.55	2	R83	CH <sub>3</sub> OH + OH = CH <sub>2</sub> OH + H <sub>2</sub> O	0.54	
	3	R1	H + O <sub>2</sub> = O + OH	0.46	3	R82	CH <sub>3</sub> OH + OH = CH <sub>3</sub> O + H <sub>2</sub> O	0.40	
	4	R24	CO + OH = CO <sub>2</sub> + H	0.45	4	R27	HCO + O <sub>2</sub> = CO + HO <sub>2</sub>	0.36	
	5	R13	HO <sub>2</sub> + OH = H <sub>2</sub> O + O <sub>2</sub>	0.43	S	5	R368	HCO + NO <sub>2</sub> = NO + CO <sub>2</sub> + H	0.32
	6	R27	HCO + O <sub>2</sub> = CO + HO <sub>2</sub>	0.41	U	6	R362	CH <sub>2</sub> O + NO <sub>2</sub> = HONO + HCO	0.32
	7	R197	NO + HO <sub>2</sub> = NO <sub>2</sub> + OH	0.40	7	R84	CH <sub>3</sub> OH + O <sub>2</sub> = CH <sub>2</sub> OH + HO <sub>2</sub>	0.26	
	8	R26	HCO + M = H + CO + M (LP)	0.39	8	R202	NO <sub>2</sub> + HO <sub>2</sub> = HNO <sub>2</sub> + O <sub>2</sub>	0.25	
	9	R40	CH <sub>2</sub> O + OH = HCO + H <sub>2</sub> O	0.34	9	R40	CH <sub>2</sub> O + OH = HCO + H <sub>2</sub> O	0.24	
	10	R9	H + O <sub>2</sub> + M = HO <sub>2</sub> + M (LP)	0.33	10	R26	HCO + M = H + CO + M (LP)	0.24	
S U E	1	R14	HO <sub>2</sub> + HO <sub>2</sub> = H <sub>2</sub> O <sub>2</sub> + O <sub>2</sub> (DUPl)	1.65	1	R389	CH <sub>3</sub> OH + NO <sub>2</sub> = HONO + CH <sub>2</sub> OH	2.82	
	2	R389	CH <sub>3</sub> OH + NO <sub>2</sub> = HONO + CH <sub>2</sub> OH	1.25	2	R362	CH <sub>2</sub> O + NO <sub>2</sub> = HONO + HCO	2.25	
	3	R27	HCO + O <sub>2</sub> = CO + HO <sub>2</sub>	1.23	3	R27	HCO + O <sub>2</sub> = CO + HO <sub>2</sub>	2.14	
	4	R362	CH <sub>2</sub> O + NO <sub>2</sub> = HONO + HCO	1.22	S	4	R83	CH <sub>3</sub> OH + OH = CH <sub>2</sub> OH + H <sub>2</sub> O	2.06
	5	R368	HCO + NO <sub>2</sub> = NO + CO <sub>2</sub> + H	0.81	U	5	R14	HO <sub>2</sub> + HO <sub>2</sub> = H <sub>2</sub> O <sub>2</sub> + O <sub>2</sub> (DUPl)	1.95
	6	R83	CH <sub>3</sub> OH + OH = CH <sub>2</sub> OH + H <sub>2</sub> O	0.66	E	6	R82	CH <sub>3</sub> OH + OH = CH <sub>3</sub> O + H <sub>2</sub> O	1.85
	7	R82	CH <sub>3</sub> OH + OH = CH <sub>3</sub> O + H <sub>2</sub> O	0.59	D	7	R368	HCO + NO <sub>2</sub> = NO + CO <sub>2</sub> + H	1.77
	8	R26	HCO + M = H + CO + M (LP)	0.43	8	R26	HCO + M = H + CO + M (LP)	1.44	
	9	R363	CH <sub>2</sub> O + NO <sub>2</sub> = HNO <sub>2</sub> + HCO	0.26	9	R40	CH <sub>2</sub> O + OH = HCO + H <sub>2</sub> O	1.01	
	10	R40	CH <sub>2</sub> O + OH = HCO + H <sub>2</sub> O	0.21	10	R202	NO <sub>2</sub> + HO <sub>2</sub> = HNO <sub>2</sub> + O <sub>2</sub>	0.90	
G	1	R27	HCO + O <sub>2</sub> = CO + HO <sub>2</sub>	37.02	1	R389	CH <sub>3</sub> OH + NO <sub>2</sub> = HONO + CH <sub>2</sub> OH	17.25	
	2	R82	CH <sub>3</sub> OH + OH = CH <sub>3</sub> O + H <sub>2</sub> O	31.94	2	R362	CH <sub>2</sub> O + NO <sub>2</sub> = HONO + HCO	11.91	
	3	R14	HO <sub>2</sub> + HO <sub>2</sub> = H <sub>2</sub> O <sub>2</sub> + O <sub>2</sub> (DUPl)	28.71	3	R14	HO <sub>2</sub> + HO <sub>2</sub> = H <sub>2</sub> O <sub>2</sub> + O <sub>2</sub> (DUPl)	10.70	
	4	R24	CO + OH = CO <sub>2</sub> + H	27.07	4	R27	HCO + O <sub>2</sub> = CO + HO <sub>2</sub>	8.53	
	5	R83	CH <sub>3</sub> OH + OH = CH <sub>2</sub> OH + H <sub>2</sub> O	24.27	G	5	R83	CH <sub>3</sub> OH + OH = CH <sub>2</sub> OH + H <sub>2</sub> O	6.54
	6	R1	H + O <sub>2</sub> = O + OH	21.92	U	6	R368	HCO + NO <sub>2</sub> = NO + CO <sub>2</sub> + H	5.90
	7	R13	HO <sub>2</sub> + OH = H <sub>2</sub> O + O <sub>2</sub>	20.12	7	R82	CH <sub>3</sub> OH + OH = CH <sub>3</sub> O + H <sub>2</sub> O	5.57	
	8	R26	HCO + M = H + CO + M (LP)	19.29	8	R26	HCO + M = H + CO + M (LP)	3.29	
	9	R199	NO <sub>2</sub> + H = NO + OH	17.02	9	R396	CH <sub>2</sub> OH + NO = HNCO + H <sub>2</sub> O	3.01	
	10	R197	NO + HO <sub>2</sub> = NO <sub>2</sub> + OH	16.41	10	R196	HNO + NO <sub>2</sub> = HONO + NO	2.72	

In Figure 1, the performance of the strategies is compared by the change of the error function values with the optimization steps. In each strategy, there were one or two major steps in which the fitting error  $E$  of the model decreased the most significantly and only minor improvements were attained in the other steps. This large improvement corresponded to the optimization of the rate parameter of reactions R82 for strategy S (1<sup>st</sup> step), reaction R83 for strategies S (2<sup>nd</sup>) and SU (2<sup>nd</sup>), reaction R27 for strategies G (1<sup>st</sup>), SUE (3<sup>rd</sup>), SUED (3<sup>rd</sup>) and GU (4<sup>th</sup>). Reaction R27 was selected at ranks 6 and 4 for strategies S and SU, respectively, but had a little effect on the error. The best first choice was made by measure G, however by step 4 it became the worst performing strategy. Similarly, strategy S selected two very influential reactions in the beginning, but from then it could hardly improve the error and its performance became the second worst from step 5. The other four strategies, SU, SUE, SUED and GU allowed a steady decrease of the error function value and behaved very similarly in steps 7-10. These strategies produced basically the same final error function values with negligible differences. The common characteristic in these measures (SU, SUE, SUED and GU) is that they all incorporate the uncertainty of the parameters, whereas strategies S and G do not, which clearly demonstrate that uncertainty information needs to be taken into account during parameter selection.

Figure 2 shows the evolution of the optimized parameter values with respect to their prior uncertainty



**Figure 2** Evolution of the optimized  $A$  parameter values with respect to their prior uncertainty range during the hierarchical optimization defined by the six different parameter impact measures. The points are connected with straight lines just to ease the eye to follow their tendencies. In the right little subpanels, the final values of the parameters with their posterior uncertainty ranges are shown, which may look distorted as there is a scale change at the uncertainty limits (at  $\pm 1$ ). The line colors and types correspond to the same reaction across all panels, whereas symbol types denote the ranking of the parameters.



**Figure 1** Evolution of the error function values during the optimization processes according to the different strategies.

range (i.e.  $\ln(A_{\text{optimized}}/A_0)/(3\sigma)$ ). A common feature of all graphs is that the optimized values of most parameters were keep changing whenever additional parameters were included into the optimization, which implies that their values are correlated. For example, one can observe that there is a large (anti)correlation between the rate coefficients of competing methanol consuming reactions R82, R83 and R389 (see Figs. 2b-f).

**Table 3** Initial and optimized  $\ln A$  and  $3\sigma(\ln A)$  values of the investigated reactions in the different strategies. The reduced posterior uncertainties are indicated with bold, the hugely increased ones with italic characters.

Reaction	Initial	S	SU	SUE	SUED	G	GU
R82 $\text{CH}_3\text{OH} + \text{OH} = \text{CH}_3\text{O} + \text{H}_2\text{O}$	-4.57±0.52	-4.22± <b>0.46</b>	-4.32± <b>0.31</b>	-4.11± <b>0.27</b>	-4.30± <b>0.26</b>	-4.28± <b>0.45</b>	-4.19± <b>0.26</b>
R83 $\text{CH}_3\text{OH} + \text{OH} = \text{CH}_2\text{OH} + \text{H}_2\text{O}$	22.71±0.81	21.91± <b>0.49</b>	21.91± <b>0.32</b>	21.91± <b>0.26</b>	21.93± <b>0.28</b>	21.91± <b>0.54</b>	21.91± <b>0.26</b>
R1 $\text{H} + \text{O}_2 = \text{O} + \text{OH}$	36.16±0.07	36.09± <i>0.61</i>	–	–	–	36.09± <i>0.58</i>	–
R24 $\text{CO} + \text{OH} = \text{CO}_2 + \text{H}$	9.72±0.07	9.78± <i>0.34</i>	–	–	–	9.78± <i>0.37</i>	–
R13 $\text{HO}_2 + \text{OH} = \text{H}_2\text{O} + \text{O}_2$	27.59±0.21	27.38± <i>0.60</i>	–	–	–	27.38± <i>0.81</i>	–
R27 $\text{HCO} + \text{O}_2 = \text{CO} + \text{HO}_2$	29.66±0.69	29.56± <b>0.30</b>	29.50±0.65	29.48±0.62	29.35±0.63	29.66± <b>0.34</b>	29.28±0.67
R197 $\text{NO} + \text{HO}_2 = \text{NO}_2 + \text{OH}$	31.45±0.19	31.35± <i>0.53</i>	–	–	–	31.45± <i>0.75</i>	–
R26 $\text{HCO} + \text{M} = \text{H} + \text{CO} + \text{M}$ (LP)	24.62±0.51	24.29± <b>0.34</b>	24.11±0.70	24.11±0.65	24.11±0.67	24.39± <b>0.31</b>	24.11±0.70
R40 $\text{CH}_2\text{O} + \text{OH} = \text{HCO} + \text{H}_2\text{O}$	24.52±0.60	24.99±0.59	23.91±0.57	23.91±0.60	23.91±0.55	–	–
R9 $\text{H} + \text{O}_2 + \text{M} = \text{HO}_2 + \text{M}$ (LP)	45.41±0.12	45.29± <i>0.66</i>	–	–	–	–	–
R389 $\text{CH}_3\text{OH} + \text{NO}_2 = \text{HONO} + \text{CH}_2\text{OH}$	5.01±4.61	–	0.42±24.52	2.75± <b>0.91</b>	0.41±23.62	–	2.93±4.65
R368 $\text{HCO} + \text{NO}_2 = \text{NO} + \text{CO}_2 + \text{H}$	30.77±2.30	–	30.25± <b>0.62</b>	30.54± <b>0.55</b>	30.36± <b>0.58</b>	–	30.47± <b>0.55</b>
R362 $\text{CH}_2\text{O} + \text{NO}_2 = \text{HONO} + \text{HCO}$	-15.78±4.61	–	-14.67± <b>0.52</b>	-17.98± <i>18.14</i>	-14.65± <b>0.51</b>	–	-14.76± <b>0.44</b>
R84 $\text{CH}_3\text{OH} + \text{O}_2 = \text{CH}_2\text{OH} + \text{HO}_2$	32.12±2.16	–	34.28± <i>3.00</i>	–	–	–	–
R202 $\text{NO}_2 + \text{HO}_2 = \text{HNO}_2 + \text{O}_2$	2.94±2.30	–	4.91± <b>0.71</b>	–	4.55± <b>0.99</b>	–	–
R14 $\text{HO}_2 + \text{HO}_2 = \text{H}_2\text{O}_2 + \text{O}_2$ (DUP1)	23.22±1.12	–	–	24.33±1.46	24.33±1.75	24.33± <i>2.14</i>	24.33±1.60
R363 $\text{CH}_2\text{O} + \text{NO}_2 = \text{HNO}_2 + \text{HCO}$	-2.21±2.30	–	–	0.05± <b>0.75</b>	–	–	–
R199 $\text{NO}_2 + \text{H} = \text{NO} + \text{OH}$	32.69±0.09	–	–	–	–	32.77± <i>0.66</i>	–
R396 $\text{CH}_2\text{OH} + \text{NO} = \text{HNCO} + \text{H}_2\text{O}$	25.33±2.30	–	–	–	–	–	23.03± <i>44.71</i>
R196 $\text{HNO} + \text{NO}_2 = \text{HONO} + \text{NO}$	10.69±1.61	–	–	–	–	–	9.08±1.93

For some non-methanol/NO<sub>x</sub> system specific reactions that have been optimized in previous studies and thereby have small prior uncertainties, the  $A$  parameters hit their uncertainty limits and stayed there for subsequent steps. This behavior could be observed for reaction R83 in the case of all strategies, and for reaction R14 in the case of strategies SUE, SUED, G and GU, which all incorporate experimental error information. Strategy S prioritized rate coefficients that were highly constrained by previous studies, thus most of the parameters immediately hit their uncertainty boundary and stayed there for most subsequent steps. A similar picture was starting to develop for strategy G, however, it could not be fully observed due to the missing data points. We can conclude that the error stagnation for strategies S and G is due to their preferred selection of highly influential parameters that usually have small uncertainties, thus, their early hitting of boundaries renders them practically untunable.

Figure 2 also shows the final optimized parameters and their posterior uncertainty with respect to their prior uncertainty range, whereas the actual values are tabulated in Table 3. Table 3 indicates reduced posterior uncertainties with bold, and hugely increased ones with italic characters. At first glance, it is obvious that strategies S and G gave not only wider posterior uncertainties relative to the prior ones compared to other strategies, but they actually predicted hugely larger posterior uncertainties than the prior ones for half of their selected reactions (e.g. 8 times larger for R1). This is also the consequence of their preferred choice of reactions with low prior uncertainties. For the other four strategies, such huge uncertainty increases could be observed only in some single cases: R389 for SU and SUED, R362 for SUE and R396 for GU. The common in these reactions is that they have a high prior uncertainty ( $f_{\text{prior}} = 1$  for R396, 2 for R389 and R362), and accordingly R389 and R362 were selected very early in all four strategies that incorporate uncertainty information. The optimized  $\ln A$  value of R389 changed significantly after the inclusion of each new parameter, and only strategy SUE could reduce

its uncertainty, whereas strategies SU and SUED gave huge posterior uncertainties for it. Reaction R396 was selected only by strategy GU and at a low rank of 9, which is probably due to its large uncertainty and moderate sensitivity. Parameter  $A$  of reaction R362 was consistently optimized by strategies SU, SUED and GU, which suggests that its mistuned value with high uncertainty obtained by strategy SUE should be ignored. The origin of this failure can be clearly identified as the inclusion of the competitive reaction R363, which was selected only by strategy SUE, in the 9<sup>th</sup> step. In Figure 2, the dashed blue line corresponding to reaction R362 evolves for strategy SUE in the same manner as for strategies SU, SUED and GU (see panels c vs. b, d and e) as long as reaction R363 (dashed cyan line) is not included into the optimization, which mistunes the  $\ln A$  value of reaction R362. In summary, one should be careful optimizing parameters with very high uncertainties, as they can have low sensitivity in their whole prior uncertainty range or they can be easily tuned into regions where the simulation results may become locally insensitive to them or when they are tuned together with a competing, correlated reaction, as these eventually can result in huge posterior uncertainties.

All methods selected and gave final optimized  $\ln A$  values and reduced uncertainties in good agreement for reactions R82 and R83, and the match is almost perfect for the latter reaction, however, this is probably because its value hit the lower prior uncertainty limit. Multiple strategies gave final values with significantly reduced uncertainties in good agreement for reactions R368, R362 and R202, thus their results can also serve as new recommendations after suitable averaging.

## Conclusions

Based on the results of a 10-parameter optimization procedure, one can recommend those measures for parameter selection that incorporate parameter uncertainties (SU, SUED, SUE and GU) as they had a significantly better performance in the long run than

those that did not (S and G). However, as the largest initial improvement was achieved by these latter two strategies, the efficiency of them might be significantly improved if parameter impacts are recalculated either regularly or after all major improvements for not all, but for an initially identified larger group of parameters. The SU, SUE, SUEd and GU strategies performed very similarly regarding the minimization of the objective function, thus further studies are needed to differentiate their performance. Regarding posterior parameter and uncertainty estimates, strategy GU can be picked as the most reliable one, as it failed only for one reaction that it ranked only at the 9<sup>th</sup> position, it kept the uncertainties of already well-constrained reactions at low values, and finally it provided the lowest uncertainties for all important rate parameters that were also identified by other strategies incorporating uncertainty information.

### Acknowledgements

This work was supported by the Hungarian National Research, Development and Innovation Office [NKFIH grants K132109 and FK134332] and the New National Excellence Program of the Ministry for Innovation and Technology [ÚNKP-20-3] from the source of the National Research, Development and Innovation Fund.

### References

- [1] D. Miller, M. Frenklach, Sensitivity analysis and parameter estimation in dynamic modeling of chemical kinetics, *Int. J. Chem. Kinet.* 15 (1983) 677–696.
- [2] M. Frenklach, Transforming data into knowledge—Process Informatics for combustion chemistry, *Proc. Combust. Inst.* 31 (2007) 125–140.
- [3] H. Wang, D.A. Sheen, Combustion kinetic model uncertainty quantification, propagation and minimization, *Prog. Energy Combust. Sci.* 47 (2015) 1–31.
- [4] D.A. Sheen, H. Wang, The method of uncertainty quantification and minimization using polynomial chaos expansions, *Combust. Flame.* 158 (2011) 2358–2374.
- [5] L. Cai, H. Pitsch, Mechanism optimization based on reaction rate rules, *Combust. Flame.* 161 (2014) 405–415.
- [6] T. Nagy, T. Turányi, Uncertainty of Arrhenius Parameters, *Int. J. Chem. Kinet.* 43 (2011) 359–378.
- [7] T. Nagy, É. Valkó, I. Sedyó, I.G. Zsély, M.J. Pilling, T. Turányi, Uncertainty of the rate parameters of several important elementary reactions of the H<sub>2</sub> and syngas combustion systems, *Combust. Flame.* 162 (2015) 2059–2076.
- [8] T. Turányi, T. Nagy, I.G. Zsély, M. Cserhádi, T. Varga, B.T. Szabó, I. Sedyó, P.T. Kiss, A. Zempléni, H.J. Curran, Determination of rate parameters based on both direct and indirect measurements, *Int. J. Chem. Kinet.* 44 (2012) 284–302.
- [9] T. Varga, T. Nagy, C. Olm, I.G. Zsély, R. Pálvölgyi, Valkó, G. Vincze, M. Cserhádi, H.J. Curran, T. Turányi, Optimization of a hydrogen combustion mechanism using both direct and indirect measurements, *Proc. Combust. Inst.* 35 (2015) 589–596.
- [10] T. Varga, C. Olm, T. Nagy, É. Valkó, R. Pálvölgyi, H.J. Curran, T. Turányi, I.G. Zsély, É. Valkó, R. Pálvölgyi, H.J. Curran, T. Turányi, Development of a joint hydrogen and syngas combustion mechanism based on an optimization approach, *Int. J. Chem. Kinet.* 48 (2016) 407–422.
- [11] C. Olm, T. Varga, É. Valkó, H.J. Curran, T. Turányi, Uncertainty quantification of a newly optimized methanol and formaldehyde combustion mechanism, *Combust. Flame.* 186 (2017) 45–64.
- [12] M. Kovács, M. Papp, I.G. Zsély, T. Turányi, Determination of rate parameters of key N/H/O elementary reactions based on H<sub>2</sub>/O<sub>2</sub>/NOx combustion experiments, *Fuel.* 264 (2020) 116720.
- [13] M. Frenklach, H. Wang, M.J. Rabinowitz, Optimization and analysis of large chemical kinetic mechanisms using the solution mapping method - combustion of methane, *Prog. Energy Combust. Sci.* 18 (1992) 47–73.
- [14] M. Kovács, M. Papp, I.G. Zsély, T. Turányi, Main sources of uncertainty in recent methanol/NOx combustion models, *Int. J. Chem. Kinet.* in press (2021) <https://doi.org/10.1002/kin.21490>.
- [15] P. Glarborg, J.A. Miller, B. Ruscic, S.J. Klippenstein, Modeling nitrogen chemistry in combustion, *Prog. Energy Combust. Sci.* 67 (2018) 31–68.
- [16] T. Varga, C. Olm, M. Papp, Á. Busai, I.G. Zsély, ReSpecTh Kinetics Data Format Specification v2.3, (2020).
- [17] ELTE Chemical Kinetics Laboratory and MTA-ELTE Complex Chemical Systems Research Group, ELTE Institute of Chemistry, ReSpecTh webpage. <http://www.respecth.hu>.
- [18] M. Frenklach, PrIME Webpage. <http://www.primekinetics.org/>.
- [19] T. Varga, M. Papp, Á. Busai, I.G. Zsély, Optima++ package v2.0.0: A general C++ framework for performing combustion simulations and mechanism optimization, (2021). <http://respecth.hu>.
- [20] A. Cuoci, A. Frassoldati, T. Faravelli, E. Ranzi, OpenSMOKE++: An object-oriented framework for the numerical modeling of reactive systems with detailed kinetic mechanisms, *Comput. Phys. Commun.* 192 (2015) 237–264.
- [21] A. Cuoci, A. Frassoldati, T. Faravelli, E. Ranzi, OpenSMOKE++ 0.12.0. <https://www.opensmokepp.polimi.it/>.
- [22] NIST Chemical Kinetics Database, Version 7.0, Release 1.6.8, Data version 2015.09. <http://kinetics.nist.gov/>.