

A Novel Active Parameter Selection Strategy for the Efficient Optimization of Combustion Mechanisms

M. Kovács¹, M. Papp¹, T. Turányi^{*1}, T. Nagy^{*2}

¹Institute of Chemistry, ELTE Eötvös Loránd University, Budapest, Hungary

²Institute of Materials and Environmental Chemistry, Research Centre for Natural Sciences, Budapest, Hungary

Abstract

In previous mechanism optimization studies, the active parameters were selected based either on local sensitivity coefficients or on products of the local sensitivity coefficient and the uncertainty of the parameters. In this work, we propose a very efficient novel method, called PCALIN, which uses not only the local sensitivities, but also considers the uncertainty and the correlation of parameters, and furthermore the uncertainty and the weights of the experimental data. The method also identifies the relevant subset of the experimental data collection, thereby allows significant savings in computation time at mechanism optimization.

Introduction

Detailed chemical kinetic mechanisms are widely used to simulate combustion systems for both industrial and scientific purposes. In the traditional mechanism development approach, the thermodynamic and kinetic information based on direct experimental data and theoretical results are amalgamated into a detailed chemical kinetic mechanism [1]. However, the performance of such a mechanism can be improved greatly by tuning its parameters within their uncertainty limits using the results of indirect measurements.

The first kinetic parameter optimization studies on combustion kinetic mechanisms were carried out by Frenklach et al. [2,3] and Sheen and Wang. [4,5]. Pitsch et al. [6,7] extended the methodology to the optimization of reaction rate rules and thermochemical parameters. A method that allowed efficient global optimization of all three Arrhenius parameters in their joint uncertainty domain [8,9] and considered both direct experimental and theoretical determinations of rate coefficients, and indirect measurements was introduced by Turányi et al. [10] This methodology has been used to optimize detailed reaction mechanisms of the combustion of several fuels (e.g. syngas [11], methanol and formaldehyde [12], hydrogen doped with nitrogen oxides [13]). In these works, a hierarchical optimization strategy was used, which gradually increased the number of active parameters and experimental data considered.

Detailed combustion mechanisms usually contain large number of uncertain parameters. Frenklach et al. suggested that only those parameters ('the active parameters') need to be fitted that have a high influence on the simulated value of the experimental data. Frenklach et al. noted [14] that at the selection of the active parameters not the local sensitivity information, but instead the product of sensitivity and the range of parameter uncertainty should be considered. Warnatz [15] has used the 'sensitivity-uncertainty index' for the characterization of the importance of a reaction step,

which was defined as the absolute value of the product of the semi-normalized local sensitivity coefficient for model result Y with respect to parameter A_i and uncertainty parameter f_i . Turányi et al. [16] demonstrated that the sensitivity-uncertainty index is proportional to the square root of the contribution of the uncertainty of the rate coefficient of reaction i to the variance of model output Y . Sheen and Wang called [17] the product of uncertainty parameter f_i and normalized sensitivity coefficient as 'uncertainty-weighted sensitivity coefficient'. Cai and Pitsch [6] introduced the term 'optimization potential' for this quantity, and they used it for the selection of the active parameters.

In a recent work [18], we demonstrated that further significant improvements in optimization efficiency can be achieved by considering the correlation of the parameters to select groups of active parameters and relevant subsets of experimental data based on the principal component analysis (PCA) of scaled sensitivity matrices. These methods can be considered as generalizations of the PCA of the sensitivity matrix [19]. Three novel PCA-based active parameter selection strategies of increasing complexity were derived and compared with local sensitivity coefficients and optimization potentials. Here, we present the same case study, but the equations are given in a more general form, which allows weighting of data series in the error function. This is necessary to develop models of balanced performance for various experimental types in the case of an unbalanced data collection, for example, which contains many more data series from jet-stirred reactors than from laminar flame speed measurements.

The investigated combustion system

Optimization of a methanol/NO_x combustion mechanism was chosen to test the novel parameter ranking strategies. Methanol is a promising alternative to fossil transportation fuels, and its interactions with

* Corresponding authors: turanyi@chem.elte.hu
nagy.tibor@ttk.hu

nitrogen oxides (NO_x) in combustion systems are also important due to environmental regulations. In a recent publication [20], we collected all available concentration data measured in jet stirred reactors (JSR), tubular flow reactors (TFR) and shock tubes (ST) corresponding to methanol/NO_x combustion and tested 17 mechanisms against them. After evaluation and filtering, 2360 data points in 225 data series were selected to be used for optimization in the present study. More details about the data collection can be found in the original publication [18]. The collected 225 data series were stored in 73 ReSpecTh Kinetics Data v2.3 (RKD) XML format data files [21]. RKD is the reaction kinetics format specification of the ReSpecTh database [22], which had been developed from the PrIME Kinetics Data Format [3].

Based on the benchmark results [20], we selected the ELTE+Glarborg mechanism for optimization due to its relatively good performance and low uncertainty. This mechanism consists of our previously optimized syngas [11], methanol [12] and hydrogen/NO_x [13] mechanisms extended with the C₁/NO_x reactions obtained from the Glarborg-2018 mechanism [23].

Optimization method

For the parameter optimizations, we applied code Optima++ [24], which implements the optimization method developed by Turányi et al. [10] and uses the numerical optimization algorithm FOCTOPUS developed by T. Nagy [10,25]. In this work we used OpenSMOKE++ [26,27] simulation code, but Optima++ is linked with multiple other solvers. The method minimizes the following error function:

$$E(\mathbf{P}) = \frac{1}{N} \sum_{f=1}^{N_f} \sum_{s=1}^{N_{fs}} w_{f,s,d} \sum_{d=1}^{N_{f,s,d}} \left(\frac{Y_{f,s,d}^{\text{sim}}(\mathbf{P}) - Y_{f,s,d}^{\text{exp}}}{\sigma_{f,s,d}^{\text{exp}}} \right)^2. \quad (1)$$

where N is the total number of data series, N_f is the number of data files, N_{fs} is the number of data series in the f -th file and $N_{f,s,d}$ is the number of data in its s -th data series. In this data series, $Y_{f,s,d}^{\text{exp}}$ is the optionally transformed d -th data and $Y_{f,s,d}^{\text{sim}}$ is its simulated value, which depends on parameter vector \mathbf{P} that contains transformed Arrhenius parameters ($\ln A$, n , E/R). R is the universal gas constant. The $w_{f,s,d}$ factors are the weights, which are all unity in the unweighted case. The experimental uncertainty of a data point is also taken into account via the $\sigma_{f,s,d}^{\text{exp}}$ standard deviation of its determination. The detailed methodology of determining $\sigma_{f,s,d}^{\text{exp}}$ and its values for all the collected data series can be found in the original publication [18]. Thus, square root of $E(\mathbf{P})$ measures the root mean squared deviation of the simulation results from the experimental data relative to the standard deviation of the experimental data. $E(\mathbf{P})$ is around 1 for a perfect model, and below 9 for a model which is accurate on average within three experimental standard deviation.

The selected Arrhenius parameters were optimized in such a way that the rate coefficients calculated with

the optimized rate parameters always remained within their prior uncertainty interval $[k_{\min}(T); k_{\max}(T)]$ in the temperature interval of 800 K to 2500 K. 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 [28] and other sources. All data used are referenced in the *k-evaluation* web site [29] and in the original publication [18] with the description of the determination of prior uncertainty of the rate coefficients. More details can be found also in our previous publication [20]. The standard deviation of $\ln k$ can be calculated by assuming that the extreme $\ln k$ values correspond to 3σ limits [30]:

$$\sigma = \frac{\ln 10}{3} f_{\text{prior}}. \quad (3)$$

When $\ln A$ is the only fitted Arrhenius parameter, this σ is assumed to be characteristic [8,9] for $\ln A$, too.

Simple global measures for parameter selection

In most previous optimization studies, identification of active parameters was based on the matrix of local sensitivity coefficients:

$$\mathbf{S} = \{S_{f,s,d,l}\} = \left\{ \frac{\partial Y_{f,s,d}^{\text{sim}}(\mathbf{P})}{\partial P_l} \right\}. \quad (4)$$

P_l is the l -th parameter out of N_p . For each (f,s,d) -th data, the $S_{f,s,d,l}$ values are normalized and the \mathbf{I}^S importance matrix (S stands for sensitivity) is defined:

$$\mathbf{I}^S = \{I_{f,s,d,l}^S\} = \left\{ \frac{S_{f,s,d,l}}{\max_m |S_{f,s,d,m}|} \right\}. \quad (5)$$

We propose a global importance measure I_l^S using RMS averaging in accordance with Eq. (1):

$$I_l^S = \sqrt{\frac{1}{N} \sum_{f=1}^{N_f} \sum_{s=1}^{N_{fs}} w_{f,s,d} \sum_{d=1}^{N_{f,s,d}} I_{f,s,d,l}^S{}^2}. \quad (6)$$

The optimization potential, which is the product of sensitivity coefficient and parameter uncertainty, is a better measure for the selection of the active parameters, since parameters can be tuned only within their uncertainty range. After the normalization of measure $S_{f,s,d,l}\sigma_l$, the \mathbf{I}^{SU} impact matrix (U stands for parameter uncertainty) can be defined:

$$\mathbf{I}^{\text{SU}} = \{I_{f,s,d,l}^{\text{SU}}\} = \left\{ \frac{S_{f,s,d,l}\sigma_l}{\max_m |S_{f,s,d,m}\sigma_m|} \right\}. \quad (7)$$

Furthermore, we can also define a global impact parameter I_l^{SU} similarly to I_l^S in Eq. (6):

$$I_l^{\text{SU}} = \sqrt{\frac{1}{N} \sum_{f=1}^{N_f} \sum_{s=1}^{N_{fs}} \frac{w_{f,s,d}}{N_{f,s,d}} \sum_{d=1}^{N_{f,s,d}} I_{f,s,d,l}^{\text{SU}}{}^2}. \quad (8)$$

Global importance measures, I_l^{S} and I_l^{SU} , which range between 0 and 1, can be used for ranking and selecting active parameters for optimization. The highest ranked parameters are included one-by-one into the optimization and more and more parameters optimized together in subsequent steps. The procedure is continued until consideration of additional parameters gives only negligible reduction in the error function. In our previous studies, this stepwise procedure could reduce the value of the error function more steadily and reliably than tuning many parameters at once. However, rate coefficients can be highly correlated, thus such a group of correlated parameters needs to be considered together during optimization.

Novel active parameter and experimental data selection strategies based on principal component analysis

To determine parameter groups which have correlated effect on the objective function we use principal component analysis. To derive a compact formalism, we introduce the following data series size-normalized and experimental uncertainty-normalized error vector:

$$\tilde{\mathbf{E}} = \{\tilde{E}_{f,s,d}\} = \left\{ \sqrt{\frac{w_{f,s,d}}{N N_{f,s,d}}} \frac{Y_{f,s,d}^{\text{sim}}(\mathbf{P}) - Y_{f,s,d}^{\text{exp}}}{\sigma_{f,s,d}^{\text{exp}}} \right\}. \quad (9)$$

Thus, the objective function gets a compact form:

$$E(\mathbf{P}) = \tilde{\mathbf{E}}^T \tilde{\mathbf{E}}. \quad (10)$$

Furthermore, the parameter vector can also be normalized by its uncertainty:

$$\tilde{\mathbf{P}} = \{\tilde{P}_i\} = \{P_i/\sigma_i\}. \quad (11)$$

The components of this normalized vector change between -1 and +1 if we vary the corresponding parameter within a $[-\sigma_i, \sigma_i]$ range.

A novel parameter importance matrix \mathbf{I}^{SUE} can be introduced, which also incorporates experimental uncertainty (denoted by superscript E):

$$\mathbf{I}^{\text{SUE}} = \{I_{f,s,d,l}^{\text{SUE}}\} = \frac{\partial \tilde{\mathbf{E}}}{\partial \tilde{\mathbf{P}}} = \left\{ \sqrt{\frac{w_{f,s,d}}{N N_{f,s,d}}} \frac{S_{f,s,d,l} \sigma_l}{\sigma_{f,s,d}^{\text{exp}}} \right\}. \quad (12)$$

The change of the objective function can be expanded up to second order in parameter perturbation $\Delta \tilde{\mathbf{P}}$:

$$\Delta E \approx 2 \tilde{\mathbf{E}}^T \mathbf{I}^{\text{SUE}} \Delta \tilde{\mathbf{P}} + \Delta \tilde{\mathbf{P}}^T \mathbf{I}^{\text{SUE,T}} \mathbf{I}^{\text{SUE}} \Delta \tilde{\mathbf{P}}. \quad (13)$$

It is important to point out that this is not a full second-order expansion as it misses terms containing second derivatives of $\tilde{\mathbf{E}}$ that contains second-order sensitivities, which would be very expensive to calculate.

If we assume that the initial kinetic model is already very accurate (i.e. components of $\tilde{\mathbf{E}}$ are small) then the variation of the objective function simplifies to the

second term of Eq. (12), which is a quadratic form of the positive semidefinite $\mathbf{Q}^{\text{SUE}} = \mathbf{I}^{\text{SUE,T}} \mathbf{I}^{\text{SUE}}$ matrix ($N_p \times N_p$). Principal component analysis (PCA) of \mathbf{Q}^{SUE} matrix is based on its eigenproblem:

$$\mathbf{Q}^{\text{SUE}} \mathbf{U} = \mathbf{U} \cdot \text{diag}(\lambda_1, \dots, \lambda_{N_p}). \quad (14)$$

Matrix \mathbf{U} contains the orthonormal eigenvectors ($\mathbf{u}_i \mathbf{u}_j = \delta_{ij}$) in its columns and λ_i -s are non-negative eigenvalues. Considering unit perturbations along all eigenvectors gives the sum of eigenvalues as the total change of the error function.

$$\sum_{i=1}^{N_p} \mathbf{u}_i^T \mathbf{Q}^{\text{SUE}} \mathbf{u}_i = \sum_{i=1}^{N_p} \lambda_i = \sum_{i=1}^{N_p} \lambda_i \underbrace{\sum_{p=1}^{N_p} u_{ip}^2}_1. \quad (15)$$

This can be approximated usually with relatively low number of principal components belonging to largest λ_i eigenvalues and with the largest absolute u_{ip} eigenvector components that designate the most important groups of correlated parameters which need to be included at once into the optimization procedure.

This formulation also allows the selection of relevant experiments if we recast the error variation caused by perturbation \mathbf{u}_i into the following form:

$$\lambda_i = \mathbf{u}_i^T \mathbf{Q}^{\text{SUE}} \mathbf{u}_i = (\mathbf{I}^{\text{SUE}} \mathbf{u}_i)^2. \quad (16)$$

The components of the $\mathbf{I}^{\text{SUE}} \mathbf{u}_i$ vector has a compound index of (f,s,d) , thus the vector norm can be decomposed into contributions of data files (i.e. experiments) as follows:

$$\lambda_i = \underbrace{\sum_{f=1}^{N_f} \sum_{s=1}^{N_{fs}} \sum_{d=1}^{N_{f,s,d}} (\mathbf{I}^{\text{SUE}} \mathbf{u}_i)_{f,s,d}^2}_{\text{data file } f}. \quad (17)$$

Using this measure, a subset of most important data files can be selected whose simulation results contribute greatly to the variation of the error function induced by a perturbation along a principal component, whereas the contribution of the ignored data files to the error function change will be negligible. Thereby, this PCA-SUE strategy is expected to significantly reduce the computational effort of parameter optimization.

To test the effect of the consideration of experimental uncertainties on the efficiency of the optimization strategy, such PCA analysis can also be done for a similarly defined $\mathbf{Q}^{\text{SU}} = \mathbf{I}^{\text{SU,T}} \mathbf{I}^{\text{SU}}$ matrix (PCA-SU strategy), even though it cannot be directly related to the variation of the error function.

If the initial kinetic model is not very accurate then the first-order variation in Eq. (13) cannot be neglected. Thus, linear contributions due to parameter variations along the principal components should also be considered during ranking of principal components. Let consider the sum of linear and quadratic changes of E induced by ± 1 (parameter) perturbation along eigenvector \mathbf{u}_i :

$$\Delta E \approx \pm 2 \tilde{\mathbf{E}}^T \mathbf{I}^{\text{SUE}} \mathbf{u}_i + \lambda_i \leq 2 |\tilde{\mathbf{E}}^T \mathbf{I}^{\text{SUE}} \mathbf{u}_i| + \lambda_i. \quad (18)$$

For ranking the principal components, the largest induced change should be used, which is obtained if the sign of the linear term is also non-negative. In the following, we refer to this procedure as PCALIN, because it also considers linear changes in the error function. The linear term can also be decomposed into contributions from individual eigenvector components, thus now they are ranked based on the following quantity instead of u_{ip}^2 :

$$\frac{2}{\lambda_i} \left| (\tilde{\mathbf{E}}^T \mathbf{I}^{\text{SUE}})_p u_{ip} \right| + u_{ip}^2. \quad (19)$$

For the selection of data files, both linear and quadratic terms in \mathbf{u}_i need to be considered:

$$2 \left| \tilde{\mathbf{E}}^T \mathbf{I}^{\text{SUE}} \mathbf{u}_i \right| + (\mathbf{I}^{\text{SUE}} \mathbf{u}_i)^2. \quad (20)$$

In the linear term, the scalar product between vectors $\tilde{\mathbf{E}}^T$ and $\mathbf{I}^{\text{SUE}} \mathbf{u}_i$ corresponds to a summing over the (f sd) compound index. Therefore, the maximum variation in the contribution of the f -th data file to the error function due to ± 1 perturbation along principal component \mathbf{u}_i can be given as:

$$\sum_{s=1}^{N_{fs}} \sum_{d=1}^{N_{fsd}} 2 \left| \tilde{\mathbf{E}}_{fsd}^T (\mathbf{I}^{\text{SUE}} \mathbf{u}_i)_{fsd} \right| + (\mathbf{I}^{\text{SUE}} \mathbf{u}_i)_{fsd}^2. \quad (21)$$

This measure allows selection of most relevant data files to be used during optimization thus it can reduce the computational effort.

Note, that these derivations can be generalized for any types of model parameters, for example the thermodynamic and transport parameters of kinetic mechanisms.

Results - ranking of parameters

Local sensitivity analysis on the simulation results of all considered 2360 data points was carried out by perturbing the pre-exponential factors for all the 562 rate coefficients (25 of them are low-pressure limit (LP) rate coefficients) in the ELTE+Glarborg mechanism with +5%. We assumed unit weights for all data points (ie. $w_{fsd} = 1$) in the error function. In all, five optimization strategies were set up and tested. These included (i) strategy S: the selection of the active parameters based on global importance measure I_i^S using the local sensitivity information only; (ii) strategy SU based on global importance measure I_i^{SU} ; (iii) strategy PCA-SU based on the PCA of matrix \mathbf{Q}^{SU} ; (iv) strategy PCA-SUE based on the PCA of matrix \mathbf{Q}^{SUE} ; (v) strategy PCALIN based on the PCA of matrix \mathbf{Q}^{SUE} with linear contributions. In both PCA and PCALIN strategies, principal components, eigenvector

components and experiments were selected to guarantee at least 95%, 90% and 90% reproduction of the objective function and the selected principal components, respectively.

Table 1 shows the comparison of the numbers of the optimization target reactions and the data files used in the optimization steps, using the different active parameter selection strategies. Note that the PCA based strategies use only a selected subset of experiments in the optimization steps, whereas final error values were obtained on the full data collection.

Strategies S and SU were determined based on the direct reaction rankings according to importance measures I_i^S and I_i^{SU} , which were calculated from the local sensitivity coefficients (S) and optimization potentials (SU), respectively.

The ranked list of reactions based on these strategies can be seen in the original publication [18]. In accordance with our hierarchic strategy, one additional parameter was included to the optimization in each optimization step, up the 12th step. In the 13th–14th steps, 3 and 5 new parameters were included together, respectively. The whole data collection was used in each optimization step.

The one-by-one inclusion of parameters based on simple parameter rankings was found to be highly CPU time-ineffective, as it carried out a separate optimization step for each parameter, moreover the subsequent inclusion of highly correlated parameters caused unnecessary readjustments. The proposed principal component analysis-based methods can overcome this issue, as they select groups of correlated reactions and relevant subsets of data files, resulting an optimization process with fewer steps and significantly lower computation times. The PCA-SU strategy identified three principal components corresponding to a 3-step optimization. However, the selection of the data files was still relatively inefficient, as 58 data files out of the 73 was already selected in the first step.

The PCA-SUE and PCALIN methods, which also take into account the experimental errors (E) and the simulation errors of experimental data (in the linear change) and thereby allow a more efficient selection of active parameters and the corresponding relevant data files, identified in 5 and 4 principal components, corresponding to 5-step and 4-step optimization procedures, respectively. The parameter groups selected by the three PCA-based strategies are given in the original publication [18], respectively.

Table 1. The number of active parameters (Par) and experimental data files (Exp) used in the optimization steps (Step) of the various strategies

S			SU			PCA-SU			PCA-SUE			PCALIN		
Step	Par	Exp	Step	Par	Exp	Step	Par	Exp	Step	Par	Exp	Step	Par	Step
1...12	1...12	73	1...2	1...12	73	1	10	58	1	7	40	1	10	36
13	15	73	13	15	73	2	15	61	2	9	45	2	11	51
14	20	73	14	20	73	3	23	69	3	14	55	3	14	53
									4	18	57	4	22	54
									5	23	60			

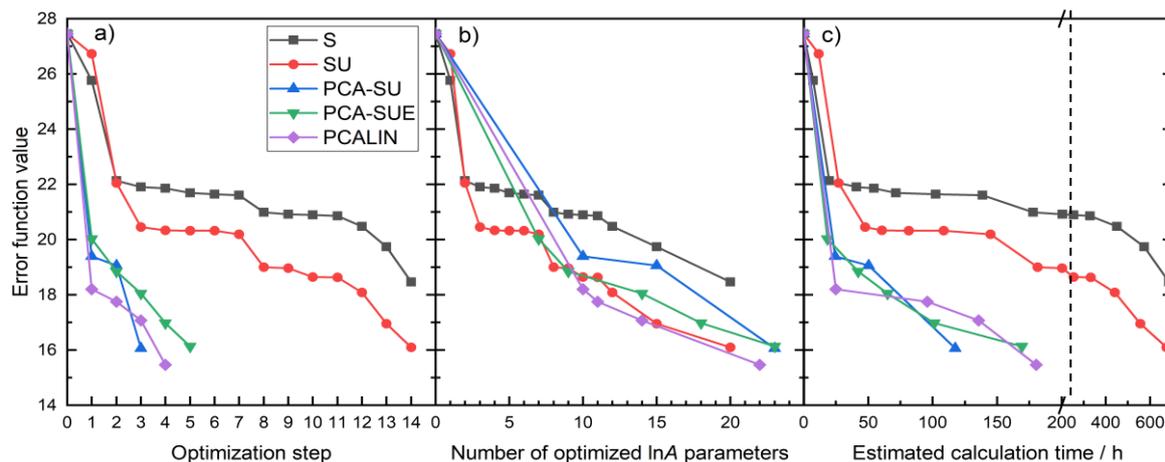


Fig. 1. Evolution of the error function values during the optimization processes according to the different strategies. Note, there is a scale change at 200 hours in panel c).

Results – optimization efficiency of strategies

For the comparison of strategies, we optimized only the $\ln A$ parameters of the selected reactions as local sensitivities coefficients were determined for this parameter only. In Fig. 1, the performance of the strategies is compared by the change of the error function values as a function of the optimization steps, the number of the optimized $\ln A$ parameters and the estimated runtimes required for the computations.

These runtimes were estimated separately for each optimization step based on the product of the mean iteration times and iteration numbers in 48-threaded calculations on AMD EPYC 7451 24-core CPUs.

Fig. 1. a) shows that the optimization of the highest ranked $\ln A$ parameters of S and SU strategies could not improve the accuracy of the mechanism steadily in every steps. The initial $E = 27.45$ value was reduced in the 2nd steps to 22.13 and 22.05 in the case of S and SU, respectively. However, after the 2nd and 3rd steps, respectively, these strategies could not reduce the E value effectively, except only in the 8th step and the 13th-14th lumped steps. The final E values were 18.47 and 16.10 in the cases of S and SU, respectively, showing the superiority of strategy SU over S. A much steadier and steeper decrease in the error function with the number of optimization steps could be achieved using the PCA-SU, PCA-SUE and PCALIN methods. The first steps of these strategies provided huge improvements and the further steps resulted in significant error reductions, too. Final E values of 16.06, 16.13 and 15.46 were reached in the cases of PCA-SU, PCA-SUE and PCALIN, respectively.

Fig. 1. b) also shows that PCALIN had the best performance, as it always had the lowest E values with respect to the numbers of the optimized $\ln A$ parameters. The most significant differences between the performance of the strategies can be seen in Fig. 1 c). The sequential S and SU strategies had an overall computation time around 700 hours, in contrast of the 100-200 hours of the other methods. Due to the lowest final E value and the low runtime, the PCALIN can be

considered as the best investigated parameter selection strategy.

The error function values, the selected reactions, the initial and optimized $\ln A$ parameter values and their prior and posterior uncertainties can be found in the original publication [18]. The optimized values obtained by the different strategies are in good agreement, and the posterior uncertainties are usually much less than the prior ones. Significant differences between the optimal $\ln A$ values according to the different strategies were found only in the cases of R362 and R396. It might be caused by an existing correlated reaction (e.g. R363 for R362), which were not optimized resulting in different optimal values. Note that these reactions had high prior uncertainty ($f \geq 1$).

Conclusions

Based on a novel principal component analysis of the error function, three procedures, denoted as PCA-SU, PCA-SUE and PCALIN, were introduced and tested. It was shown that the PCA-based methods could achieve steady decrease of the model error and by allowing the selection of subsets of relevant experiments they also allowed significant savings in computational time. Especially the PCALIN method proved to be very effective at the selection of the active parameters, as it also takes into account the simulation error of the starting model for the individual experiments. The corresponding codes are available on the respect webpage [22].

Acknowledgements

The authors acknowledge the support of NKFIH OTKA grants FK134332 and K132109 of the Hungarian National Research, Development and Innovation Office.

References

- [1] H.J. Curran, Developing detailed chemical kinetic mechanisms for fuel combustion, Proc. Combust. Inst. 37 (2019) 57–81.
- [2] D. Miller, M. Frenklach, Sensitivity analysis

- and parameter estimation in dynamic modeling of chemical kinetics, *Int. J. Chem. Kinet.* 15 (1983) 677–696.
- [3] M. Frenklach, Transforming data into knowledge - Process Informatics for combustion chemistry, *Proc. Combust. Inst.* 31 (2007) 125–140.
- [4] H. Wang, D.A. Sheen, Combustion kinetic model uncertainty quantification, propagation and minimization, *Prog. Energy Combust. Sci.* 47 (2015) 1–31.
- [5] D.A. Sheen, H. Wang, The method of uncertainty quantification and minimization using polynomial chaos expansions, *Combust. Flame.* 158 (2011) 2358–2374.
- [6] L. Cai, H. Pitsch, Mechanism optimization based on reaction rate rules, *Combust. Flame.* 161 (2014) 405–415.
- [7] F. vom Lehn, L. Cai, H. Pitsch, Sensitivity analysis, uncertainty quantification, and optimization for thermochemical properties in chemical kinetic combustion models, *Proc. Combust. Inst.* 37 (2019) 771–779.
- [8] T. Nagy, T. Turányi, Uncertainty of Arrhenius parameters, *Int. J. Chem. Kinet.* 43 (2011) 359–378.
- [9] 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 H₂ and syngas combustion systems, *Combust. Flame.* 162 (2015) 2059–2076.
- [10] 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.
- [11] T. Varga, C. Olm, T. Nagy, 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.
- [12] 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.
- [13] M. Kovács, M. Papp, I.G. Zsély, T. Turányi, Determination of rate parameters of key N/H/O elementary reactions based on H₂/O₂/NO_x combustion experiments, *Fuel.* 264 (2020) 116720.
- [14] 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.
- [15] J. Warnatz, Resolution of gas phase and surface combustion chemistry into elementary reactions, *Symp. Combust.* 24 (1992) 553–579.
- [16] T. Turányi, L. Zalotai, S. Dóbbé, T. Bérces, Effect of the uncertainty of kinetic and thermodynamic data on methane flame simulation results, *Phys. Chem. Chem. Phys.* 4 (2002) 2568–2578.
- [17] D.A. Sheen, H. Wang, Combustion kinetic modeling using multispecies time histories in shock-tube oxidation of heptane, *Combust. Flame.* 158 (2011) 645–656.
- [18] M. Kovács, M. Papp, T. Turányi, T. Nagy, A novel active parameter selection strategy for the efficient optimization of combustion mechanisms, *Proc. Combust. Inst.* (2022).
- [19] T. Turányi, P. Valkó, S. Vajda, Principal component analysis of kinetic models, *Int. J. Chem. Kinet.* 17 (1985) 55–81.
- [20] M. Kovács, M. Papp, I.G. Zsély, T. Turányi, Main sources of uncertainty in recent methanol/NO_x combustion models, *Int. J. Chem. Kinet.* 53 (2021) 884–900.
- [21] T. Varga, C. Olm, M. Papp, Á. Busai, I.G. Zsély, ReSpecTh Kinetics Data Format Specification v2.3, (2020).
- [22] ELTE Institute of Chemistry, ReSpecTh webpage, (2023). <https://respecth.hu/>
- [23] P. Glarborg, J.A. Miller, B. Ruscic, S.J. Klippenstein, Modeling nitrogen chemistry in combustion, *Prog. Energy Combust. Sci.* 67 (2018) 31–68.
- [24] M. Papp, T. Varga, Á. Busai, I.G. Zsély, T. Nagy, T. Turányi, Optima++ v2.1.0: A general C++ framework for performing combustion simulations and mechanism optimization (2021.05.03), (2021).
- [25] S.K. Goitom, M. Papp, M. Kovács, T. Nagy, I.G. Zsély, T. Turányi, L. Pál, Efficient numerical methods for the optimisation of large kinetic reaction mechanisms, *Combust. Theory Model.* 26 (2022) 1071–1097.
- [26] 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.
- [27] A. Cuoci, A. Frassoldati, T. Faravelli, E. Ranzi, OpenSMOKE++ 0.12.0, (2020).
- [28] NIST Chemical Kinetics Database, Version 7.0, Release 1.6.8, Data version 2015.09, (2013).
- [29] M. Papp, É. Valkó, T. Nagy, T. Turányi, Evaluation of high-temperature gas phase rate coefficients, (2020). <https://k-evaluation.elte.hu/>
- [30] T. Turányi, Sensitivity analysis of complex kinetic systems. Tools and applications, *J. Math. Chem.* 5 (1990) 203–248.