Uncertainty of Arrhenius Parameters

TIBOR NAGY, TAMÁS TURÁNYI

Institute of Chemistry, Eötvös University (ELTE), P.O. Box 32, H-1518, Budapest, Hungary

ABSTRACT: Chemical kinetics databases for many elementary gas-phase reactions provide the recommended values of the Arrhenius parameters, the temperature range of their validity, and the temperature dependence of the uncertainty of the rate coefficient k. An analytical expression is derived that describes the temperature dependence of the uncertainty of k as a function of the elements of the covariance matrix of the Arrhenius parameters. Based on this analytical expression, the various descriptions of the temperature dependence of the uncertainty of k used in the combustion, and in the IUPAC and JPL atmospheric chemical databases are analyzed in detail. Recommendations are given for an improved representation of the uncertainty information in future chemical kinetics databases using the covariance matrix of the Arrhenius parameters. Utilization of the joint uncertainty of the Arrhenius parameters is needed for a correct uncertainty analysis in varying temperature chemical kinetic systems. A method is suggested for the determination of the covariance matrix and the joint probability density function of the Arrhenius parameters from the present uncertainty information given in the kinetics databases. The method is demonstrated on seven gas kinetic reactions exhibiting different types of uncertainty representation. © 2011 Wiley Periodicals, Inc. Int J Chem Kinet 43: 359–378, 2011

INTRODUCTION

The rate parameters of chemical reactions are always determined with some uncertainty in either experiments or theoretical calculations. The published uncertainty of the measured rate parameters usually reflects the scattering of the measurement data only, whereas evaluated rate parameters are based on several (sometimes dozens or even hundreds of) measurements or calculations, and thus can take into account the systematic errors of all determinations. The uncertainty values recommended in the kinetic data evaluations are the results of the critical evaluation by the groups of experts and they do not represent the results of the straight statistical analyses of the entire data sets available from

Correspondence to: Tamás Turányi; e-mail: turanyi@chem .elte.hu.

the original publications. However, the uncertainty information, provided in collections of evaluated data, is a good indicator of the depth of knowledge of a given reaction and the data evaluations attribute a statistical meaning to the uncertainty limits.

In liquid-phase kinetics and atmospheric chemistry, the temperature dependence of a rate coefficient k is described using the Arrhenius expression $k = A \exp(-E/RT)$. In high-temperature gas kinetic systems, such as combustion and pyrolytic systems, the temperature dependence of rate coefficient k is traditionally given in the form of the modified Arrhenius expression $k = AT^n \exp(-E/RT)$. The modified Arrhenius expression is sometimes used in form $k = BT^n \exp(-C/RT)$, to emphasize that parameters B and C are different from preexponential factor A and activation energy E, respectively. In this paper, the form $k = AT^n \exp(-E/RT)$ is used because it allows a common discussion of all Arrhenius-type

expressions. For some gas-phase reactions, the expression $k = AT^n$ describes the temperature dependence of the rate coefficient. Temperature independent rate coefficients are defined by the expression $k = A$. The latter two functions and the original Arrhenius expression can be considered as simplified special forms of the modified Arrhenius expression.

Determination of the main values and the variances of the parameters of Arrhenius-type expressions from experimental k data has been discussed in the classic articles of Cvetanović et al. [1–3]. The uncertainties of Arrhenius parameters have also been investigated by Héberger et al. [4]. They concluded that the Arrhenius coefficients determined in most experiments are highly correlated and that this correlation should be characterized by a covariance matrix. In their work, the uncertainties of the Arrhenius parameters were calculated from simulated rate coefficient data. Estimation of Arrhenius parameters A and E using the logarithmic form of the Arrhenius equation was investigated by Klička and Kubáček [5], and by Sundberg [6], based on different assumptions for the errors of the measured rate coefficients. Rodríguez-Aragón and López-Fidalgo [7] discussed the determination of the Arrhenius parameters from the point of view of experimental design theory. Schwaab et al. [8,9] studied the question of the selection of the optimum reference temperature to decrease the correlation between the parameters of the reparameterized Arrhenius equation. This investigation was extended also to power functions [10]. Najm et al. [11] carried out a numerical simulation of the ignition of a stoichiometric methane-air mixture using a detailed methane combustion mechanism. The calculated CH_4 concentrations were considered to be measured data and random Gaussian noise was added to these values. Then, the joint probability density function (pdf) of transformed Arrhenius parameters $\ln A$ and $\ln E$, belonging to a single-step global reaction, was determined using the Markov chain Monte Carlo (MCMC) method. Strong correlation was found between these transformed Arrhenius parameters and their calculated pdf was almost of a two-dimensional (2D) normal distribution.

The approach presented in this paper is different from those described in the articles above. In all of the above articles, the uncertainty of the Arrhenius parameters was calculated from either real or simulated experimental data. This paper investigates the relationship between the temperature dependence of the uncertainty of rate coefficient k and the temperature independent uncertainty of the Arrhenius parameters A, n, and E.

Many fields of practical importance, including atmospheric chemistry, combustion, and chemical indus-try, use the results of gas kinetic simulations. Uncertain rate parameters cause uncertainty in the simulation results. Not only are the results obtained using the nominal parameter set important but also the predicted uncertainty of the simulation results. It will be shown that a good characterization of the uncertainty of the Arrhenius parameters is a starting point of realistic uncertainty analysis of models using detailed reaction mechanisms.

USUAL DEFINITIONS OF THE UNCERTAINTY OF THE RATE COEFFICIENT

Critical compilations of gas kinetic reactions provide not only the recommended Arrhenius parameters but also report the reliability of the rate coefficients by assigning an uncertainty parameter to them. A temperature range is given that refers to the range of validity of the rate coefficient and its uncertainty. A series of comprehensive evaluations of combustion reactions was published by Baulch et al. [12–14]. In the field of atmospheric chemistry, a series of kinetic evaluations was carried out by the International Union of Pure and Applied Chemistry (IUPAC) Subcommittee for Gas Kinetic Data Evaluation (see its web page [15]). These results were also published in a series of articles by Atkinson et al. [16–19]. The Jet Propulsion Laboratory (JPL) group regularly publishes an updated database of chemical kinetics and photochemical data for use in atmospheric studies. Evaluation Number 15 was the latest database [20] and Evaluation Number 16 will be released soon.

The combustion reviews, the IUPAC and the JPL evaluations use different but related methods for the characterization of the uncertainties of the rate coefficients. In the combustion reviews of Baulch *et al.* [12–14], the uncertainty of the rate coefficient is defined by value f in the following way:

$$f = \log_{10}\left(k^0/k^{\min}\right) = \log_{10}\left(k^{\max}/k^0\right) \qquad (1)$$

where k^0 is the recommended value of the rate coefficient of the reaction, and $k^{\min}$ and $k^{\max}$ are the possible extreme values; rate coefficients outside the $[k^{\min}, k^{\max}]$ interval are considered very improbable by the evaluators. Temperature independent parameter f was defined for some reactions, whereas for other reactions different f values were set at different temperatures or temperature intervals (see Table III for examples). Similar definition of uncertainty information is used in other combustion chemistry data collections (see, e.g., refs. [21–22]).

Assuming that the minimum and maximum values of the rate coefficients correspond to 3σ deviations [23–29] or 2σ deviations [30] from the recommended value on a logarithmic scale, the uncertainty parameter f can be converted [25] at a given temperature T to the standard deviation of the logarithm of the rate coefficient using the equation

$$\sigma(\log_{10} k) = \frac{\sigma(\ln k)}{\ln 10} = \frac{1}{m} f(T) \qquad (2)$$

where $m = 3$ or 2, respectively.

Rigorously speaking, using the notation $\ln k$ is not correct because the logarithm of a quantity with a physical unit cannot be evaluated. The correct notation is $\ln \{k\}$, where the operator $\{\}$ keeps the numerical value of k but removes the physical unit [31]. All similar logarithms will be considered this way, but the curly brackets will be avoided to simplify notation.

In the IUPAC atmospheric chemical kinetic data evaluations [15–19], the uncertainties are characterized by three parameters (T_0, d_0, g) in the following way:

$$d(T) = \Delta \log_{10} k(T) = d_0 + \frac{g}{\ln 10} \cdot \left(T^{-1} - T_0^{-1}\right) \qquad (3)$$

where $T_0 = 298$ K and $d_0 = \Delta \log_{10} k(T_0)$. Uncertainty measures d_0 and g are provided with $\pm$ signs, corresponding to $\pm 2\sigma$ confidence limits in $\log_{10} k$ and E/R, respectively:

$$\sigma(\log_{10} k) = \frac{\sigma(\ln k)}{\ln 10} = \frac{|d(T)|}{2} \qquad (4)$$

$$\sigma(E/R) = \frac{|g|}{2} \qquad (5)$$

Substitution of uncertainty measures d_0 and g with positive signs into Eq. (3) gives a monotonically decreasing function $d(T)$. If $d_0 < g \cdot T_0^{-1}/\ln 10$ then $d(T)$ may become negative within the reference temperature range. For example, for reaction R266 (HO + $CH_3I \rightarrow H_2O + CH_2I$) the corresponding data [19] are $d_0 = \pm 0.2$, $T = 270 - 430$ K, $g = \pm 500$ K, giving $d(430 \text{ K}) = -0.024$. Of course, a negative uncertainty measure is not interpretable. This means that application of Eq. (3) without the appropriate selection of the signs of d_0 and g may give an incorrect result.

The IUPAC uncertainty expression can be transformed to an unambiguous form by taking the absolute values of the error terms, which results in a constructive superposition of the positive errors at any temperature T:

$$d(T) = |d_0| + \left| \frac{g}{\ln 10} \cdot \left(T^{-1} - T_0^{-1}\right) \right| \qquad (6)$$

Using this equation, uncertainty $d(T)$ is always positive and has a minimum value of $|d_0|$ at T_0. This is equivalent to an appropriate temperature-dependent selection of the signs of d_0 and g in Eq. (3) to achieve constructive superposition of terms.

The JPL data evaluations [20] overcome this problem by using absolute value in the function that corresponds to Eq. (3) and by defining the parameters, similar to d_0 and g, to be strictly positive. In the JPL database, no unit is supplied for g, although it should be given as K. In this approach, at any given temperature the uncertainty factor $f_{\text{JPL}}(T)$ may be obtained from the following three-parameter (T_0, f_0, g) expression:

$$f_{\text{JPL}}(T) = f_0 \exp \left(g \,|\, T^{-1} - T_0^{-1}| \right) \qquad (7)$$

where $T_0 = 298$ K; $f_0 = f_{\text{JPL}}(T_0)$ and g are positive parameters. After linearization we get

$$\ln f_{\text{JPL}}(T) = \ln f_0 + g \,|\, T^{-1} - T_0^{-1}| \qquad (8)$$

Note that here g has a meaning different from that used in the IUPAC definition and also the exponent contains an absolute value. The JPL evaluation [20] claims that "an upper or lower bound (corresponding approximately to one standard deviation) of the rate constant at any temperature T can be obtained by multiplying or dividing the recommended value of the rate constant at that temperature by the factor $f(T)$". This means that the standard deviation of $\ln k$ is as follows:

$$\sigma(\ln k) = \ln f_{\text{JPL}}(T) \qquad (9)$$

The JPL database considers one standard deviation as upper and lower bounds, but does not recommend truncation at these bounds. However, definition of a truncation threshold is still needed when the JPL data are used for uncertainty analysis to avoid the consideration of very improbable values.

The IUPAC and the JPL approaches are based on the facts that the rate coefficients of atmospheric chemical reactions are typically known with minimum uncertainty at standard temperature 298 K, and that the overall uncertainty normally increases toward lower and higher temperatures, because there are usually fewer measured data at those temperatures. The temperature dependence of the uncertainty of the rate coefficient is defined by different functions, but these are essentially equivalent.

In all studies where the uncertainties of rate coefficients given in databases of evaluated data have been utilized (see, e.g., refs. [24–30,32,33]), the uncertainty of k was considered to be equal to the uncertainty of preexponential factor A, silently assuming

that the uncertainties of Arrhenius parameters E and n are zero, which is a physically unrealistic assumption. Joint characterization of the uncertainties of the Arrhenius parameters is needed to calculate properly the uncertainties of the simulation results.

TEMPERATURE DEPENDENCE OF THE UNCERTAINTY OF THE RATE COEFFICIENT

Three-Parameter (A, n, E) Arrhenius Equation

The main aim of this section is to show that assuming an Arrhenius-type temperature dependence of the rate coefficient k implies a certain type of function for the temperature dependence of the uncertainty of $\ln k$. The most flexible description of the temperature dependence of k is the modified Arrhenius equation $k(T) = AT^n \exp(-E/RT)$. Taking the logarithm of this equation yields $\ln k(T) = \ln A + n \cdot \ln T - E/R \cdot T^{-1}$, which can be written in a more compact form using transformed parameters $\kappa(T) := \ln k(T)$, $\alpha := \ln A$, $\varepsilon := E/R$ and column vector notations $\mathbf{p} := (\alpha, n, \varepsilon)^{\mathrm{T}}$ and $\boldsymbol{\theta} := (1, \ln T, -T^{-1})^{\mathrm{T}}$, where superscript T denotes transposition.

$$\kappa(T) = \alpha + n \cdot \ln T - \varepsilon \cdot T^{-1} = \boldsymbol{\theta}^{\mathrm{T}}\mathbf{p} = \mathbf{p}^{\mathrm{T}}\boldsymbol{\theta} \quad (10)$$

In this linearized form of the equation, κ is a homogeneous linear function of parameters (α, n, ε). Here κ is a temperature-dependent random variable, and therefore parameters (α, n, ε), which are determined from κ in a temperature range, are also random variables. It was shown in the preceding section that according to all critical data evaluations uncertainty is proportional to the standard deviation of the logarithm of the rate coefficient, thus the uncertainty also has a probabilistic meaning.

The theories of reaction rates do not yield an Arrhenius-type equation, but the temperature dependence of the experimentally or theoretically obtained

Arrhenius expression in a wide temperature range. This is the reason why all kinetics databases use either the modified or the original Arrhenius expression for the characterization of the temperature dependence of the rate coefficient. Within the scope of the Arrhenius expression, parameters α, n, and ε are physically well-defined quantities, therefore their joint pdf, which reflects our current state of knowledge on the rate coefficient, is temperature independent and consequently their expected values $\bar{\mathbf{p}} = (\bar{\alpha}, \bar{n}, \bar{\varepsilon})^{\mathrm{T}}$, standard deviations $\sigma_{\mathbf{p}} = (\sigma_\alpha, \sigma_n, \sigma_\varepsilon)$, and correlations $r_{\alpha n}$, $r_{\alpha \varepsilon}$, $r_{n \varepsilon}$ are also temperature independent. By definition, these have the following properties:

$$0 \leq \sigma_\alpha, \sigma_n, \sigma_\varepsilon, \quad -1 \leq r_{\alpha n}, r_{\alpha \varepsilon}, r_{n \varepsilon} \leq +1 \quad (11)$$

The covariance matrix $\boldsymbol{\Sigma}_{\mathbf{p}}$ of parameters (α, n, ε) is defined and can be expanded as [34]:

$$\boldsymbol{\Sigma}_{\mathbf{p}} = \overline{(\mathbf{p} - \bar{\mathbf{p}})(\mathbf{p} - \bar{\mathbf{p}})^{\mathrm{T}}}$$
$$= \begin{bmatrix} \sigma_\alpha^2 & r_{\alpha n}\sigma_\alpha\sigma_n & r_{\alpha\varepsilon}\sigma_\alpha\sigma_\varepsilon \\ r_{\alpha n}\sigma_\alpha\sigma_n & \sigma_n^2 & r_{n\varepsilon}\sigma_n\sigma_\varepsilon \\ r_{\alpha\varepsilon}\sigma_\alpha\sigma_\varepsilon & r_{n\varepsilon}\sigma_n\sigma_\varepsilon & \sigma_\varepsilon^2 \end{bmatrix} \quad (12)$$

By definition, the covariance matrix is symmetric and positive semidefinit, implying that its determinant is nonnegative, which requires that the following be true:

$$0 \leq \sigma_\alpha^2\sigma_n^2\sigma_\varepsilon^2 \left(1 - r_{\alpha n}^2 - r_{\alpha\varepsilon}^2 - r_{n\varepsilon}^2 + 2r_{\alpha n}r_{\alpha\varepsilon}r_{n\varepsilon}\right) \quad (13)$$

Let us assume that the temperature-dependent pdf of $\kappa(T)$ is known in the temperature range of $[T_1, T_2]$. At a given temperature $T \in [T_1, T_2]$, the expected value and the variance of κ is denoted by $\bar{\kappa}(T)$ and $\sigma_\kappa^2(T)$, respectively. A consequence of Eq. (10) is that there exists a relationship between the expected values of $\kappa(T)$ and (α, n, ε).

$$\bar{\kappa}(T) = \boldsymbol{\theta}^{\mathrm{T}}\bar{\mathbf{p}} = \bar{\mathbf{p}}^{\mathrm{T}}\boldsymbol{\theta} = \bar{\alpha} + \bar{n} \cdot \ln T - \bar{\varepsilon} \cdot T^{-1} \quad (14)$$

Another consequence of Eq. (10) is the following relationship between the variance of $\kappa(T)$ and the elements of the covariance matrix:

$$\sigma_\kappa(T) = \sqrt{\overline{(\kappa(T) - \bar{\kappa}(T))^2}} = \sqrt{\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{\Sigma}_{\mathbf{p}}\boldsymbol{\theta}} \quad (15)$$

$$\sigma_\kappa(T) = \sqrt{\sigma_\alpha^2 + \sigma_n^2 \ln^2 T + \sigma_\varepsilon^2 T^{-2} + 2r_{\alpha n}\sigma_\alpha\sigma_n \ln T - 2r_{\alpha\varepsilon}\sigma_\alpha\sigma_\varepsilon T^{-1} - 2r_{n\varepsilon}\sigma_n\sigma_\varepsilon T^{-1} \ln T} \quad (16)$$

rate coefficients can be accurately described by a two-parameter Arrhenius expression in a less than 100 K range of temperature and by a three-parameter

According to Eqs. (2), (4), and (9), the variance of $\kappa(T)$ is temperature independent if and only if its uncertainty parameter is temperature independent. Equation (16) implies that it is possible only if $\sigma_\kappa^2(T) = \sigma_\alpha^2$,

and that all other standard deviations and correlations are zero. This trivial solution has been applied previously in all uncertainty studies, when the temperature independent variance of κ was assumed to be identical to the variance of α, and the variances of the other Arrhenius parameters and their correlations were not considered. However, this is a physically unrealistic assumption, if the temperature dependence of the rate coefficient is also considered. Equation (16) shows that if the uncertainties of Arrhenius parameters n and/or E are also considered, then the variance of κ cannot be temperature independent. Equation (16) suggests that a temperature independent uncertainty is physically unrealistic, and also shows the required functional form of the temperature dependence of the uncertainty.

The special cases derived from the modified Arrhenius expression by assuming missing parameters are discussed in the next sections.

Two-Parameter (A, E) and (A, n) and One-Parameter (A) Arrhenius Equations

The linearized equations for the two-parameter (A, E) simple, (A, n) power-type, and the one-parameter (A) constant Arrhenius expressions are

$$\kappa(T) = \alpha - \varepsilon \cdot T^{-1} \tag{17}$$

$$\kappa(T) = \alpha + n \cdot \ln T \tag{18}$$

$$\kappa(T) = \alpha \tag{19}$$

The corresponding relationships exist between the variance of $\kappa(T)$ and the elements of the covariance matrix:

$$\sigma_\kappa(T) = \sqrt{\sigma_\alpha^2 + \sigma_\varepsilon^2 T^{-2} - 2r_{\alpha\varepsilon}\sigma_\alpha\sigma_\varepsilon T^{-1}} \tag{20}$$

$$\sigma_\kappa(T) = \sqrt{\sigma_\alpha^2 + \sigma_n^2 \ln^2 T + 2r_{\alpha n}\sigma_\alpha\sigma_n \ln T} \tag{21}$$

$$\sigma_\kappa(T) = \sigma_\alpha \tag{22}$$

Equations (20) and (21) also suggest that a temperature independent uncertainty is physically unrealistic, and they also show the required functional form of the temperature dependence of the uncertainty.

A consequence of Eq. (22) is that if the rate coefficient is temperature independent, then the uncertainty of $\kappa(T)$ should also be temperature independent. Another possibility is to approximate the variance of a temperature independent (constant) rate coefficient using the modified Arrhenius Eq. (10) with random parameters n and E having expected values of zero but nonzero variances and correlations. This way, temperature-dependent uncertainty can be obtained

even for a temperature independent rate coefficient. In a similar way, in the cases of the two-parameter expressions the temperature dependence of the uncertainties can be made more flexible by introducing nonzero variances and correlations for the missing third parameter.

IUPAC Uncertainty Definition

Here, we show that the IUPAC error handling is not in accordance with Eq. (20). If we assume correlation of unit absolute value ($r_{\alpha\varepsilon} = \pm 1$), then Eq. (20) is reduced to the following:

$$
\begin{aligned}
\sigma_\kappa(T) &= \sqrt{\sigma_\alpha^2 + \sigma_\varepsilon^2 T^{-2} - 2r_{\alpha\varepsilon}\sigma_\alpha\sigma_\varepsilon T^{-1}} \\
&= \left|\sigma_\alpha - r_{\alpha\varepsilon}\sigma_\varepsilon T^{-1}\right| = \\
&= \begin{cases} +\sigma_\alpha + \sigma_\varepsilon T^{-1} & \text{if } r_{\alpha\varepsilon} = -1 \text{ and for any } T \\ -\sigma_\alpha + \sigma_\varepsilon T^{-1} & \text{if } r_{\alpha\varepsilon} = +1 \text{ and } T \le \sigma_\varepsilon/\sigma_\alpha \\ +\sigma_\alpha - \sigma_\varepsilon T^{-1} & \text{if } r_{\alpha\varepsilon} = +1 \text{ and } T > \sigma_\varepsilon/\sigma_\alpha \end{cases}
\end{aligned}
\tag{23}
$$

Here, we took advantage of the fact that in this case there is a complete square under the square root sign. The IUPAC uncertainty expression (Eq. (6)) can be transformed in a similar way:

$$
\sigma_{\kappa,\text{IUPAC}}(T) = \frac{\ln 10}{2}d(T) = \frac{\ln 10}{2}|d_0|
$$

$$
+ \left|\frac{g}{2} \cdot \left(T^{-1} - T_0^{-1}\right)\right|
$$

$$
= \begin{cases} \overbrace{\frac{\ln 10}{2}|d_0|}^{+\sigma_\alpha} - \frac{|g|}{2}T_0^{-1} + \overbrace{\frac{|g|}{2}T^{-1}}^{+\sigma_\varepsilon T^{-1}} \\ \qquad\text{if } T \le T_0 > |g| / |d_0| / \ln 10 \\[6pt] \underbrace{\frac{\ln 10}{2}|d_0|}_{-\sigma_\alpha} - \frac{|g|}{2}T_0^{-1} + \overbrace{\frac{|g|}{2}T^{-1}}^{+\sigma_\varepsilon T^{-1}} \\ \qquad\text{if } T \le T_0 < |g| / |d_0| / \ln 10 \\ \qquad (\Rightarrow T \le T_0 < \sigma_\varepsilon/\sigma_\alpha) \\[6pt] \overbrace{\frac{\ln 10}{2}|d_0|}^{+\sigma_\alpha} + \frac{|g|}{2}T_0^{-1} - \underbrace{\frac{|g|}{2}T^{-1}}_{-\sigma_\varepsilon T^{-1}} \\ \qquad\text{if } T > T_0 \\ \qquad (\Rightarrow T > T_0 > \sigma_\varepsilon/\sigma_\alpha) \end{cases}
\tag{24}
$$

All sign combinations of the terms and the corresponding conditions in Eq. (24) can be matched and fulfilled with one of the cases in Eq. (23). In summary, two parameter sets for the joint pdf of the two-parameter Arrhenius equation and thus also two standard deviation functions $\sigma_\kappa(T)$ can be obtained from the IUPAC

uncertainty measures d_0 and g in the following way:

$$\sigma_\alpha = \begin{cases} \left| \dfrac{\ln 10}{2} |d_0| - \dfrac{|g|}{2} T_0^{-1} \right| & \text{if } T \leq T_0 \\[2mm] \dfrac{\ln 10}{2} |d_0| + \dfrac{|g|}{2} T_0^{-1} & \text{if } T > T_0 \end{cases} \qquad (25)$$

$$\sigma_\varepsilon = \frac{|g|}{2} \qquad (26)$$

$$r_{\alpha\varepsilon} = \begin{cases} \text{sign}\left(|g| T_0^{-1} - \ln 10 \cdot |d_0| \right) & \text{if } T \leq T_0 \\[2mm] +1 & \text{if } T > T_0 \end{cases} \qquad (27)$$

$$\sigma_\kappa(T) = \begin{cases} \left| \dfrac{\ln 10}{2} |d_0| + \dfrac{|g|}{2} \cdot \left(T^{-1} - T_0^{-1} \right) \right| & \text{if } T \leq T_0 \\[2mm] \left| \dfrac{\ln 10}{2} |d_0| - \dfrac{|g|}{2} \cdot \left(T^{-1} - T_0^{-1} \right) \right| & \text{if } T > T_0 \end{cases} \qquad (28)$$

This means that there is no unique temperature independent value for σ_α, since different σ_α values can be determined below and above T_0. Also, different correlation coefficient $r_{\alpha\varepsilon}$ is obtained below and above T_0, unless $|g| T_0^{-1} > \ln 10 \cdot |d_0|$. Thus, the IUPAC-type uncertainty definition is not in accordance with the existence of a unique temperature independent joint pdf of the Arrhenius parameters, and therefore it cannot be used for the uncertainty analysis of varying temperature systems.

JPL Uncertainty Definition

As the JPL uncertainty definition is essentially the same as the IUPAC one, it is also not in accordance with Eq. (20). Transforming the JPL uncertainty expression (Eq. (8)) into a form similar to Eq. (23), we get the following:

$$\sigma_{\kappa,\text{JPL}}(T) = \ln f_{\text{JPL}}(T) = \ln f_0 \\ + g \left| T^{-1} - T_0^{-1} \right| \\ = \begin{cases} \overbrace{\ln f_0 - g T_0^{-1}}^{+\sigma_\alpha} \overbrace{+ g T^{-1}}^{+\sigma_\varepsilon T^{-1}} & \text{if } T \leq T_0 > g/\ln f_0 \\[3mm] \underbrace{\overbrace{\ln f_0 - g T_0^{-1}}^{-\sigma_\alpha} \overbrace{+ g T^{-1}}^{+\sigma_\varepsilon T^{-1}}}_{(\Rightarrow T \leq T_0 < \sigma_\varepsilon/\sigma_\alpha)} & \text{if } T \leq T_0 < g/\ln f_0 \\[3mm] \underbrace{\overbrace{\ln f_0 + g T_0^{-1}}^{+\sigma_\alpha} \overbrace{- g T^{-1}}^{-\sigma_\varepsilon T^{-1}}}_{(\Rightarrow T > T_0 > \sigma_\varepsilon/\sigma_\alpha)} & \text{if } T > T_0 \end{cases} \qquad (29)$$

The uncertainty factor f_0 is always greater than one; therefore $\ln f_0 > 0$, and $\ln f_0 + g T_0^{-1} > 0$. All sign combinations of the terms and the corresponding con-

ditions in Eq. (29) can be matched and fulfilled with one of the cases in Eq. (23). In summary, two different covariance matrices of the Arrhenius parameters and thus two standard deviation functions $\sigma_\kappa(T)$ can be determined from the JPL uncertainty measures $\ln f_0$ and g as follows:

$$\sigma_\alpha = \begin{cases} \left| \ln f_0 - g T_0^{-1} \right| & \text{if } T \leq T_0 \\[2mm] \ln f_0 + g T_0^{-1} & \text{if } T > T_0 \end{cases} \qquad (30)$$

$$\sigma_\varepsilon = g \qquad (31)$$

$$r_{\alpha\varepsilon} = \begin{cases} \text{sign}\left(g T_0^{-1} - \ln f_0 \right) & \text{if } T \leq T_0 \\[2mm] +1 & \text{if } T > T_0 \end{cases} \qquad (32)$$

$$\sigma_\kappa(T) = \begin{cases} \left| \ln f_0 + g \left(T^{-1} - T_0^{-1} \right) \right| & \text{if } T \leq T_0 \\[2mm] \left| \ln f_0 - g \left(T^{-1} - T_0^{-1} \right) \right| & \text{if } T > T_0 \end{cases} \qquad (33)$$

This means that there is no unique temperature independent value for σ_α because different σ_α values can be determined below and above T_0. Also, different correlation coefficient $r_{\alpha\varepsilon}$ is obtained below and above T_0, unless $g T_0^{-1} > \ln f_0$. The conclusion is the same as for the IUPAC uncertainty definition. The JPL uncertainty definition is not in accordance with the existence of a unique temperature independent joint pdf of the Arrhenius parameters and cannot be used for the uncertainty analysis of varying temperature systems.

DETERMINATION OF THE COVARIANCE MATRIX OF THE ARRHENIUS PARAMETERS

Equations deduced for the temperature dependence of $\sigma_\kappa(T)$ (see Eqs. (16), (20), and (21)) together with the uncertainty information given in the chemical kinetics databases (see Eqs. (2), (4), and (9)) can be used for the determination of the covariance matrix of the Arrhenius parameters.

To simplify the equations below, a common notation will be used for the definition of uncertainty parameter $F(T)$ and for the proportionality factor M of $\sigma_\kappa(T)$. The meaning of $F(T)$ and M is given in Table I for

Table I Unified Notations for the Uncertainty Definitions of Various Kinetics Databases

Kinetics database	$F(T)$	M
Combustion truncation at 3σ	$f(T)$	$3/\ln 10$
Combustion truncation at 2σ	$f(T)$	$2/\ln 10$
IUPAC	$d(T)$	$2/\ln 10$
JPL	$\ln f_{\text{JPL}}(T)$	1

the different types of kinetics databases. Equation (34) defines the relationship between $F(T)$, M, and $\sigma_\kappa(T)$.

$$F(T) = M \, \sigma_\kappa(T) \qquad (34)$$

In the case of a three-parameter modified Arrhenius expression, based on Eq. (16), the elements of the covariance matrix can be determined by a least squares fit to the uncertainty data using the following function:

the existence of further Arrhenius parameters with expected values of zero but nonzero uncertainties and correlations. Thus, uncertainty equations with two and three Arrhenius parameters (Eqs (35), (36), and (37)) can be used instead to provide a better fit. Constant Rate Coefficient with Uncertain Arrhenius Parameters (α, n) and (α, ε) section provides an illustration for such an approach.

$$F(T) = M\sqrt{\sigma_\alpha^2 + \sigma_n^2 \ln^2 T + \sigma_\varepsilon^2 T^{-2} + 2r_{\alpha n}\sigma_\alpha\sigma_n \ln T - 2r_{\alpha\varepsilon}\sigma_\alpha\sigma_\varepsilon T^{-1} - 2r_{n\varepsilon}\sigma_n\sigma_\varepsilon T^{-1}\ln T} \qquad (35)$$

When determining the parameters, the linear (Eq. (11)) and multivariate nonlinear (Eq. (13)) inequality constraints for the standard deviations and correlations also have to be taken into account. Only a few data fitting software are capable of handling nonlinear constraints and one of them is EASY-FIT Express [35]. A computer code is also available for the determination of the covariance matrix of the Arrhenius parameters from the uncertainty parameter of the rate coefficient in web page http://garfield.chem.elte.hu/ Combustion/Combustion.html.

In the case of the two-parameter (α, ε) and (α, n) expression, the corresponding functions to be used in fitting are

$$F(T) = M\sqrt{\sigma_\alpha^2 + \sigma_\varepsilon^2 T^{-2} - 2r_{\alpha\varepsilon}\sigma_\alpha\sigma_\varepsilon T^{-1}} \qquad (36)$$

$$F(T) = M\sqrt{\sigma_\alpha^2 + \sigma_n^2 \ln^2 T + 2r_{\alpha n}\sigma_\alpha\sigma_n \ln T} \qquad (37)$$

In these cases, only constraints defined in Eq. (11) have to be taken into account and the uncertainty has to be known at least at three temperatures.

If the rate coefficient is described by the constant parameter α, then the corresponding equation is

$$F(T) = M\sigma_\alpha \qquad (38)$$

Even if the number of uncertainty points is equal to that minimally required, the above equations with optimized parameters may not reproduce these points exactly, since Eqs. (35), (36), and (37) are not fully flexible. If the uncertainties are provided at more temperatures than minimally required, then we have an over determined set of parameters. In both cases, the optimal values can be determined by a least squares fit, taking into account the constraints defined in Eqs. (11) and (13).

If the rate coefficient is constant or the temperature dependence is described by a two-parameter Arrhenius expression and the corresponding uncertainty equation (Eqs. (36), (37), and (38)) do not reproduce properly the given uncertainty data, then it is possible to assume

If the uncertainty values were known at fewer points than required, then fixing some of the correlations at certain values (e.g., ± 1) might look like a good idea. In our sample calculations (see the Examples for the Construction of the Joint pdf of the Arrhenius Parameters from the Uncertainty Information of the Kinetics Databases section and Table IV), nearly unit correlations were found in several cases. However, as will be shown in the above-mentioned section, perfect correlation among all Arrhenius parameters results in zero uncertainty of the rate coefficient at some temperature(s), which is physically unrealistic.

Another possibility for fitting the expression $F(T)$ to fewer than required uncertainty data points is to set the variance of some of the Arrhenius parameters to zero. This implies that all correlation coefficients of these parameters are also zero. As parameter A is always present in a rate expression, σ_α should never be zero. The possible cases are given in Table II.

If uncertainty values are available at fewer points than sufficient, another possible approach is to increase the number of points via linear interpolation between two uncertainty values. Note that if a linear interpolation of the uncertainties is applied in a T^{-1} or $\ln T$ scale, then a strong correlation will be generated between Arrhenius parameters (α, ε) or (α, n), respectively, because the corresponding terms within the square root in Eqs. (35), (36), and (37) will be closer to a complete square, which corresponds to a perfect correlation.

THE JOINT NORMAL PDF OF THE ARRHENIUS PARAMETERS

To this point, no particular type of pdf was assumed for either the rate coefficient or the Arrhenius parameters. Some methods of uncertainty analysis [36] calculate the variance of the simulation results from the variances (or the covariance matrix) of the parameters. For these types of methods, the covariance

Table II Simplified Equations for Temperature in Special Cases Regarding the Parameters of the Covariance Matrix of the Arrhenius Parameters

Case ID	σ_n	σ_ε	$r_{\alpha n}$	$r_{\alpha \varepsilon}$	$r_{n\varepsilon}$	$\sigma_\kappa(T)$
0	0	0	0	0	0	σ_α
2r	+	0	$r_{\alpha n}$	0	0	$\sqrt{\sigma_\alpha^2 + \sigma_n^2 \ln^2 T + 2r_{\alpha n}\sigma_\alpha \sigma_n \ln T}$
2+			+1			$\sigma_\alpha + \sigma_n \ln T$
2−			−1			$|\sigma_\alpha - \sigma_n \ln T|$
3r	0	+	0	$r_{\alpha \varepsilon}$	0	$\sqrt{\sigma_\alpha^2 + \sigma_\varepsilon^2 T^{-2} - 2r_{\alpha \varepsilon}\sigma_\alpha \sigma_\varepsilon T^{-1}}$
3+				+1		$|\sigma_\alpha - \sigma_\varepsilon T^{-1}|$
3−				−1		$\sigma_\alpha + \sigma_\varepsilon T^{-1}$
2±3±	+	+	±1	±1	+1	$|\sigma_\alpha \pm \sigma_n \ln T \mp \sigma_\varepsilon T^{-1}|$
2±3∓	+	+	±1	∓1	−1	$|\sigma_\alpha \pm \sigma_n \ln T \pm \sigma_\varepsilon T^{-1}|$

matrix of the Arrhenius parameters provides enough information. Other methods of uncertainty analysis are able to calculate the pdf of the simulation results using the joint pdf of the parameters. Therefore, it is advantageous to define an approximate joint pdf of the Arrhenius parameters.

In several works dealing with the uncertainty of combustion systems [26–28], a normal distribution was assumed for parameters $\ln k$, truncated at $\pm 3\sigma$. A normal distribution for $\ln k$, but with $\pm 2\sigma$ truncation, was used in another combustion uncertainty analysis study [30] and in the uncertainty analysis of atmospheric chemical systems [32,33]. In the data evaluations, the assumed distribution of the rate coefficient is usually not stated explicitly, but implicitly a normal distribution of $\ln k$ is assumed by assigning [15] 95% probability percentages to the uncertainty limits. As a matter of fact, considering a truncated normal distribution for $\ln k$ is also just an assumption. Not enough measurements are available for any rate coefficient to justify this assumption. However, it can be assumed that the recommended value $\ln k^0$ is the most probable value and that the pdf smoothly decreases to zero toward the extreme values $\ln k^{\min}$ and $\ln k^{\max}$.

In this section, the main conclusions made about the relationship between the temperature-dependent pdf of the rate coefficients and the temperature independent pdf of the Arrhenius parameters are listed and the detailed mathematical proofs are given in the Appendices.

It can be shown (see Appendix A) that the joint pdf of the Arrhenius parameters uniquely determines the pdf of the rate coefficient at every temperature.

It can be proven (see Appendix B) that if the temperature dependence of the rate coefficient is described by a three-parameter (two-parameter) Arrhenius equation, then the $\ln k$ values are uncorrelated only up to three (two) temperatures. The random $\ln k$ values at all other temperatures are correlated. If the $\ln k$ values are sampled independently from each other at several

(more than three (or two)) temperatures, then the set of these random $\ln k$ values is physically unrealistic. Using such data leads to misleading uncertainty analysis results.

If $\ln k$ follows a normal distribution at each temperature, it does not necessarily mean that the transformed Arrhenius parameters (α, n, ε) follow a multivariate normal distribution (see Appendix C). If $\ln k$ follows a normal distribution at many temperatures, then the most natural assumption is that (α, n, ε) follow a multivariate normal distribution. A joint multivariate normal distribution of Arrhenius parameters (α, n, ε) implies a normal distribution for $\ln k$ at any temperature.

An algorithm is provided for the generation of random Arrhenius parameter values from their pdf (Appendix D). The pdf should be truncated, that is, the extreme values should be discarded *a posteriori* to ensure that the $\ln k$ values always remain between the $\ln k_{\min}$ and $\ln k_{\max}$ values in the whole temperature interval.

EXAMPLES FOR THE CONSTRUCTION OF THE JOINT PDF OF THE ARRHENIUS PARAMETERS FROM THE UNCERTAINTY INFORMATION OF THE KINETICS DATABASES

The covariance matrices of the Arrhenius parameters of seven reactions were determined using the equations derived in the preceding sections. The chemical reactions were selected to present a wide spectrum of possible uncertainty cases. The examples were chosen from the latest evaluation of Baulch et al. [14], and the latest IUPAC [19] and JPL [20] evaluations. The Arrhenius parameters, the temperature range of validity, and the uncertainty information for sample reactions R1-R7 are given in Table III. The determined corresponding covariance matrices are given in Table IV.

Table III   Data of the Reactions Used in the Examples. Parameter α Is Calculated with Parameter A Given in Units mole, cm, and s

| #	Reaction	α	n	ε/K	Temperature Interval	Uncertainty	M	Ref.
R1	$O + N_2O \rightarrow NO + NO$	32.134	–	13930	1000–4000 K	2000 K: 0.2; 4000 K: 0.3; 1000 K: 0.4	$\frac{3}{\ln 10}$	[14]
R2	$N + OH \rightarrow NO + H$	32.317	−0.2	–	100–2500 K	300 K: 0.1; 100 K: 0.3; 2500 K: 0.4		
R3	$HO_2 + C_3H_5 \rightarrow C_3H_6 + O_2$	28.605	–	–	600–1000 K	600–800 K: 0.3; 1000 K: 0.5		
R4	$O + C_2H_4 \rightarrow$ products	16.422	1.88	92	220–2000 K	300–1000 K: 0.1; 220 K: 0.3; 2000 K: 0.3		
R5	$H + CH_3 \rightarrow H_2 + {}^1CH_2$	37.076	−0.56	1350	300–2500 K	300–1000 K: 0.15; 1000–1700 K : 0.3; 1700–2500 K: 0.2		
R6	$HO + CH_3I \rightarrow H_2O + CH_2I$	23.474	–	1120	270–430 K	IUPAC definition: $T_0 = 298$ K, $d_0 = \pm 0.2$, $g = \pm 500$ K	$\frac{2}{\ln 10}$	[19]
R7	$O(^1D) + H_2O \rightarrow OH + OH$	32.218	–	−60	200–400 K	JPL definition: $T_0 = 298$ K, $f_0 = 1.15, g = 45$ K	1	[20]

Uncertain Arrhenius Parameters (α, ε)

Temperature dependence of the rate coefficient of reaction R1 is defined by a two-parameter (α, ε) original Arrhenius equation. The uncertainty value f is defined [14] at three temperatures and, according to Eq. (36), these values determine the covariance matrix. The optimized parameters of the covariance matrix are given in Table IV (row **R1**) and the $f(T)$ function obtained with these parameter values is plotted in Fig. 1.

Figure 2 presents the temperature-dependent pdf of transformed rate coefficient κ. As Fig. 1 shows, the variance of κ is smaller and therefore pdf is narrower at intermediate temperatures. As the area under the pdf is equal to one at each temperature, the peak of the pdf is higher at intermediate temperatures. Figure 3 shows a three-dimensional (3D) plot of the truncated joint pdf of the Arrhenius parameters. It is interesting that the shape of the border of the allowed values of

Table IV   Determined Elements of the Covariance Matrix of the Arrhenius Parameters the Reactions Defined in Table III. For the Explanation of the Notations see the Text

|	Type	σ_α	σ_n	σ_ε/K	$r_{\alpha n}$	$r_{\alpha\varepsilon}$	$r_{n\varepsilon}$
COMBUSTION	R1	0.355	–	588	–	0.9045	–
	R2	0.970	0.162	–	−0.9979	–	–
	R2x	1.020	0.169	–	fixed −1	–	–
	R3a	5.951	0.909	–	−0.99949	–	–
	R3b	1.110	–	752	–	0.9884	–
	R4	3.427	0.463	254	−0.99972	0.999996	−0.99965
	R4x	4.198	0.558	317	fixed −1	fixed +1	fixed −1
	R5	1.177	0.121	172	−0.999992	0.9814	−0.9806
	R5m	1.008	0.138	66.1	−0.9936	0.9928	−0.99997
IUPAC	R6a	0.609	–	250	–	+1	–
	R6b	1.069	–	250	–	+1	–
	R6	1.564	–	479	–	0.9898	–
JPL	R7a	0.011	–	45	–	+1	–
	R7b	0.291	–	45	–	+1	–
	R7	0.411	–	110	–	0.9415	–
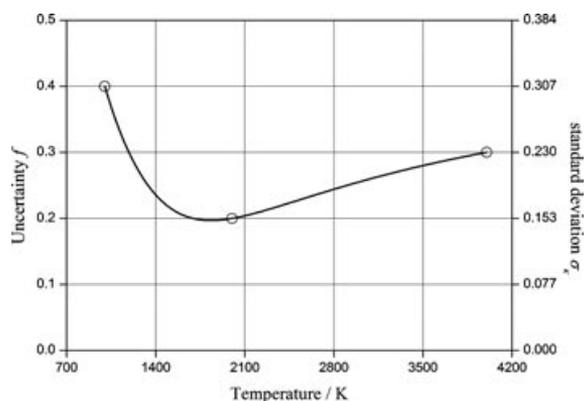
Figure 1 Uncertainties of reaction R1 are defined [14] at temperatures 1000, 2000, and 4000 K (circles). The right axis shows the corresponding standard deviation σ_κ. The uncertainty–temperature function (solid line) is calculated from the covariance matrix of the Arrhenius parameters (see Table IV, row **R1**).

the transformed Arrhenius parameters is significantly different from the elliptical shape of the isocontours. The calculated correlation coefficient $r_{\alpha\varepsilon} = +0.9045$ indicates strong correlation between parameters α and ε and this correlation is also visible in Fig. 3.

Uncertain Arrhenius Parameters (α, n)

Temperature dependence of the rate coefficient of reaction R2 is defined by a two-parameter (α, n) power-type equation. Uncertainty f is defined [14] at three temperatures and these points determine (see Eq. (37)) the three parameters of the covariance matrix of the two Arrhenius parameters. These optimized parame-
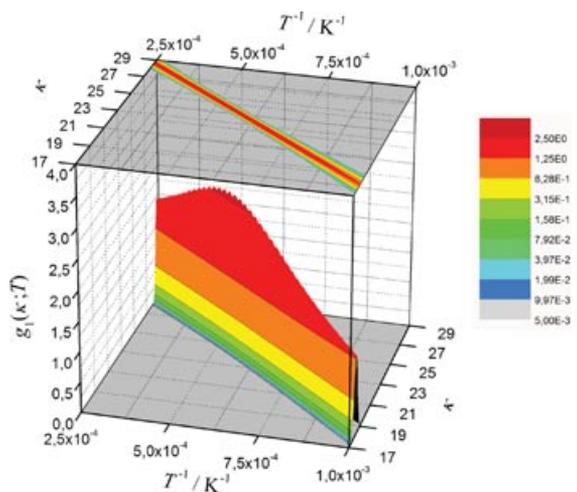
Figure 2 3D graph and its 2D projection of the temperature-dependent pdf of variable $\kappa = \ln k$ for reaction R1. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]
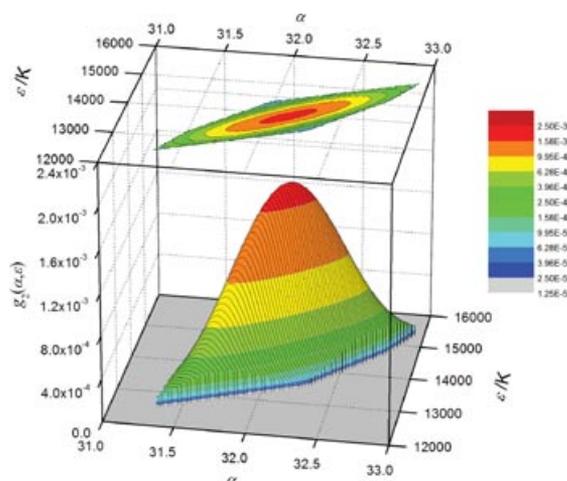
Figure 3 3D graph and its 2D projection of the joint pdf of the Arrhenius parameters for reaction R1. The pdf is truncated to give $k^{\min} < k < k^{\max}$ (truncation at $\pm 3\sigma$) in the entire evaluated temperature range. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

ters are given in Table IV (row **R2**) and the $f(T)$ function obtained with them is plotted in Fig. 4. There is a strong anticorrelation between α and n ($r_{\alpha n} = -0.9979$). However, it is important that correlation $r_{\alpha n}$ is not equal to -1, which would be case **2** in Table II. The dotted line in Fig. 4 shows that the minor change of $r_{\alpha n}$ from -0.9979 to -1 dramatically changes the uncertainty–temperature curve. At temperature $T = \exp(\sigma_\alpha/\sigma_n)$ it goes to zero in a physically unrealistic way. The conclusion is that although parameters α and n are strongly anticorrelated, assuming a
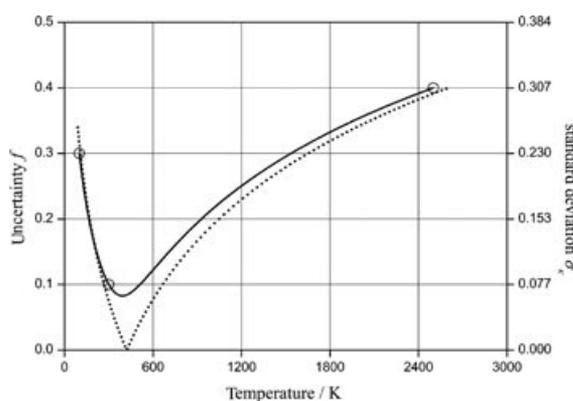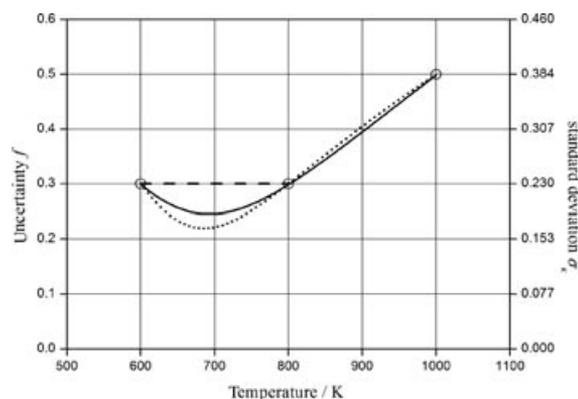

Figure 4 Uncertainties of reaction R2 are defined [14] at temperatures 100, 300, and 2500 K (circles). The uncertainty–temperature function (solid line) is calculated from the covariance matrix of the Arrhenius parameters (see Table IV, row **R2**). When correlation $r_{\alpha n} = -1$ is assumed, then the fitted curve is very different (dotted line; row **R2x** in Table IV).

correlation of $r_{\alpha n} = -1$ leads to qualitatively incorrect results. Nevertheless, the fitted parameters σ_α and σ_n changed only slightly (see row **R2x** in Table IV).

Constant Rate Coefficient with Uncertain Arrhenius Parameters (α, n) and (α, ε)

The rate coefficient of reaction R3 is considered temperature independent in the interval of 600–1000 K that should imply that the uncertainty is constant in the whole temperature range. This is not the case, because its recommended uncertainty is constant [14] in the interval of 600–800 K, whereas a different value was defined [14] at 1000 K. If these uncertainty values are valid, then Arrhenius parameter either n or ε should be regarded as a random parameter having an expected value of zero, but nonzero corresponding elements in the covariance matrix (see the Two-Parameter (α, ε) and (α, n) and One-Parameter (α) Arrhenius Equations section).

Both pairs of Arrhenius parameters (α, n) and (α, ε) are considered. The uncertainty values defined at temperatures 600, 800, and 1000 K are taken into account and the optimized parameters of the covariance matrix are presented in rows **R3a** and **R3b**, respectively, of Table IV. The two fitting functions are different, but the corresponding uncertainty–temperature functions (see Fig. 5) are very close to each other.
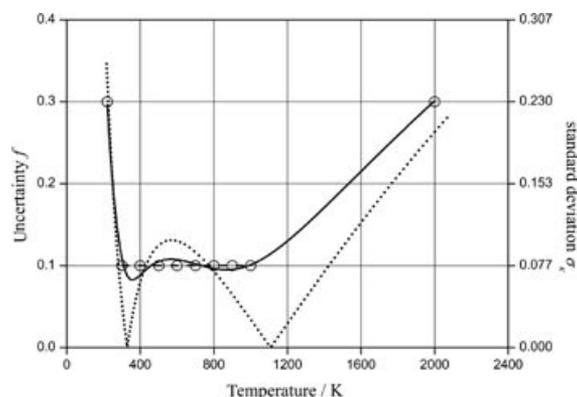

Figure 5 Uncertainties of reaction R3 is defined [14] at 1000 K (circle) and is considered constant [14] in the temperature interval of 600–800 K (dashed line). The uncertainty–temperature function was fitted to the points indicated by circles. To interpret the temperature dependence of the uncertainty parameter, in case (a) uncertain n was assumed with expected value $\bar{n} = 0$, whereas in case (b) uncertain ε was assumed with expected value $\bar{\varepsilon} = 0$. The fitted functions for cases (a) and (b) are denoted by solid and dotted lines, respectively. The corresponding elements of the covariance matrix are given in rows **R3a** and **R3b**, respectively, of Table IV.

There is a strong correlation between parameters α and n ($r_{\alpha n} = -0.99949$), and parameters α and ε ($r_{\alpha\varepsilon} = +0.9884$); however, it is important again that these correlations are not equal to -1 and $+1$, respectively.

Uncertain Arrhenius Parameters (α, n, ε)

The rate coefficient of reaction R4 changes with temperature according to modified Arrhenius Eq. (10). Its uncertainty f was considered [14] to be constant in the temperature interval of 300–1000 K and different f values were defined at temperatures of 220 and 2000 K.

The definition of the covariance matrix of the Arrhenius parameters, according to Eq. (35) requires the knowledge of the uncertainty value at least at six different temperatures. In our calculations, uncertainties were considered at 10 points (denoted by circles in Fig. 6). These included the uncertainties at 220 and 2000 K, and also in eight equally spaced points in the temperature interval of 300–1000 K. The six parameters of Eq. (35) were determined by fitting them to these points using the method of least squares, also taking into account the constraints defined in Eqs. (11) and (13). The value of the $f(T)$ function calculated from the parameters of the correlation matrix is plotted in Fig. 6. According to the optimized parameters (see row **R4** in Table IV), one of the correlations is plus one and the other two correlations are almost of unit absolute values ($r_{\alpha n} = -0.99972$,
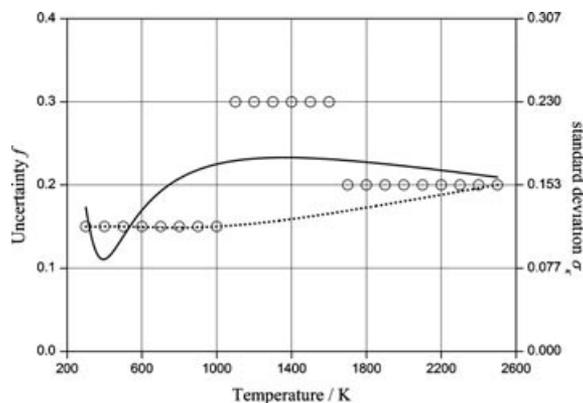

Figure 6 Uncertainty of reaction R4 is defined at temperatures 220 and 2000 K (circles) and is considered constant in the temperature interval of 300–1000 K (dashed line) in ref. [14]. The uncertainty–temperature function (solid line) was fitted to all points indicated by circles. After fixing correlation coefficients at values $r_{\alpha n} = -1$, $r_{\alpha\varepsilon} = +1$, and $r_{n\varepsilon} = -1$, the fitted uncertainty–temperature function (dotted line) reaches twice the x axis. The corresponding elements of the covariance matrix are given in rows **R4** and **R4x**, respectively, of Table IV.

$r_{\alpha\varepsilon} = 0.999996$, $r_{n\varepsilon} = -0.99965$). Again, it is important that the absolute values of the correlations are not equal to one. Fixing values $r_{\alpha n} = -1$, $r_{\alpha\varepsilon} = +1$, and $r_{n\varepsilon} = -1$ (corresponding to case $2 - 3+$ in Table II) during the fitting results in a very different uncertainty–temperature curve (see the dotted line in Fig. 6).

Uncertain Arrhenius Parameters (α, n, ε) with Uncertainty Maximum of the Rate Coefficient at Intermediate Temperatures

Rate coefficients of elementary gas-phase reactions can be determined with relatively low error in the temperature range of about 300–1000 K by direct methods like laser photolysis / laser induced fluorescence (LP/LIF) or discharge flow-LIF. Most of the rate coefficients obtained by the shock tube method are available from about 1700 K, although this method is applicable down to 1000 K. Therefore, the evaluators are tempted to assign low uncertainty values to the rate coefficient in the low and the high temperature range, and a higher uncertainty in the intermediate temperature range. For example, for reaction R5 the recommended [14] uncertainty is $f = 0.15$ in the temperature range of 300–1000 K, $f = 0.3$ in the temperature range of 1000–1700 K, and $f = 0.2$ in the temperature range of 1700–2500 K. These uncertainty data were fitted using Eq. (35). Figure 7 shows that the fitted uncertainty–temperature curve goes far from the recommended uncertainty values. The reason of this failure is that if $k(T)$ can be

Figure 7 Uncertainties of reaction R5 are defined [14] in three temperature intervals and the uncertainty–temperature function was fitted to the values represented by the circles. The best least squares fit (solid line; see row **R5** in Table IV) does not approach these points well, because the recommended uncertainties were not consistent. The alternatively proposed $f(T)$ curve (dotted line) remains below all uncertainty points and has the highest average value of $\sigma_\kappa^2(T)$ in the inverse temperature interval of $[1/T_2, 1/T_1]$ (see text and row **R5m** in Table IV).

described by a single Arrhenius expression, then the value of the uncertainty parameter cannot be significantly higher at intermediate temperatures than at the boundaries of the temperature interval. It can be qualitatively explained in such a way that consideration of the low and the high temperature measurements together limits the uncertainty of the rate coefficient at intermediate temperatures via the Arrhenius expression. Uncertainty–temperature function with a significant maximum at intermediate temperatures cannot be related to any joint pdf of the Arrhenius parameters.

An alternative approach is to consider all uncertainty data as upper estimates. In this case, the fitted curve will not depend on the high uncertainty values at intermediate temperatures, thus, these values are considered as redundant. This can be achieved by fitting the elements of the covariance matrix of the modified Arrhenius parameters in such a way that the $f(T)$ curve remains below all uncertainty points (see dotted line in Fig. 7), while the integral V is also maximized:

$$V = \int\limits_{1/T_2}^{1/T_1} dT^{-1} \sigma_\kappa^2(T) \qquad (39)$$

This integral is proportional to the average variance of κ in the inverse temperature interval of $[1/T_2, 1/T_1]$. When large number of points are approximated, the determined covariance matrix is independent of the number of points sampled from temperature ranges with constant uncertainty. Equation (39) is also applicable if less data points are available than the number of independent elements of the covariance matrix.

## Uncertain Arrhenius Parameters (α, ε)–IUPAC Uncertainty Definition

The rate coefficient of reaction R6 has a temperature dependence defined by the original Arrhenius equation. Its uncertainty is defined in the temperature range of 270–430 K by the IUPAC evaluation [19] using Eq. (6). In this case, the uncertainty was defined by two lines (see the solid gray lines in Fig. 8). Two different pdfs of the Arrhenius parameters can be determined, one that is valid below T_0 and another one that is valid above T_0. The parameters of the corresponding covariance matrices, calculated by Eqs. (25)–(27), are given in rows **R6a** and **R6b**, respectively, of Table IV. Parameter σ_α is different, whereas parameters σ_ε and $r_{\alpha\varepsilon}$ are identical in both cases. The corresponding uncertainty–inverse temperature functions are plotted by dashed and dotted lines, respectively, in Fig. 8. However, the assumption of different pdfs of the Arrhenius parameters below and above T_0 is physically unrealistic.
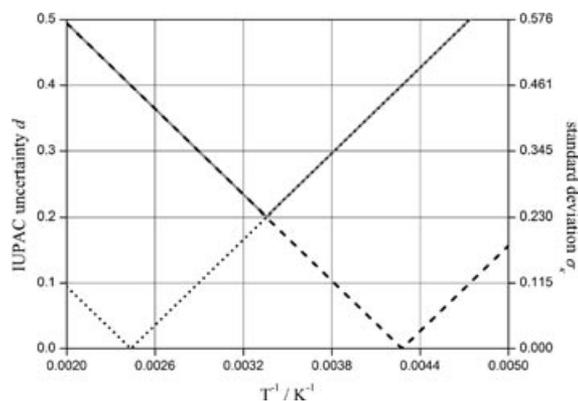
Figure 8 The IUPAC-type uncertainty definition provides minimal uncertainty at 298 K and linearly increasing uncertainties as a function of inverse temperature toward both higher and lower temperatures (solid gray lines). This can be interpreted as a composite of two different uncertainty definitions, one valid below 298 K (dotted line; row R6a in Table IV) and the other valid above 298 K (dashed line; row R6b in Table IV). The plotted functions belong to reaction R6.

To determine the joint pdf of two Arrhenius parameters, the uncertainty of rate coefficient k has to be known at least at three temperatures. Uncertainty values defined by the intersection of the two gray lines (at T_0) and at the extremes of the temperature interval (270 and 430 K) were selected. The parameters of the covariance matrix of the Arrhenius parameters were determined by fitting Eq. (36) to these points. Using this equation with the optimized parameters, the uncertainty of the rate coefficient was calculated in the whole temperature interval (see Fig. 9).

It is impossible to reproduce the lines defined by the IUPAC evaluation this way, but the obtained curve passes through the selected points and predicts a signif-
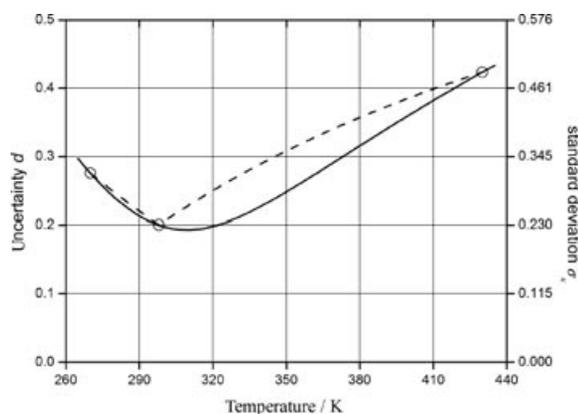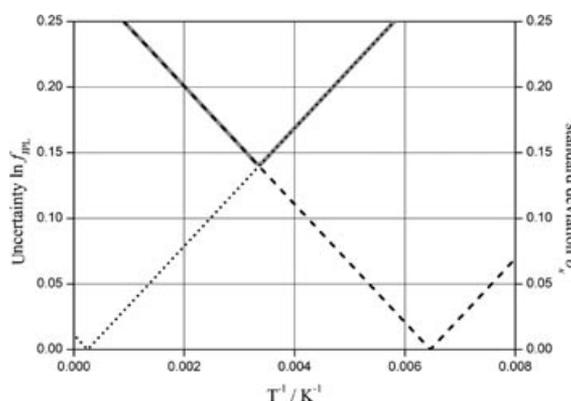

Figure 9 Uncertainties of reaction R6 are defined in the IUPAC evaluation [19] according to the dashed lines. The uncertainty–temperature function (solid line) was fitted to the points indicated by circles.

icantly lower uncertainty at other temperatures. Both the current IUPAC-type uncertainty definition and the uncertainty definition based on the covariance matrix require the same number of parameters; these parameter sets are (T_0, d_0, g) and $(\sigma_\alpha, \sigma_\varepsilon, r_{\alpha\varepsilon})$, respectively. There is a significant correlation ($r_{\alpha\varepsilon} = +0.9898$) between Arrhenius parameters α and ε (see row **R6** in Table IV).

Uncertain Arrhenius Parameters (α, ε)–JPL Uncertainty Definition

The discussion of the JPL-type uncertainty definition is very similar to the IUPAC one. The example is the rate coefficient of reaction R7 that is defined by the two-parameter original Arrhenius equation. Its uncertainty is given in the temperature range of 200–400 K by the JPL collection of evaluated data [20] using Eq. (8). The JPL-type uncertainty was defined by two lines (see the solid gray lines in Fig. 10). Two different pdfs of the Arrhenius parameters correspond to these lines, which are valid below and above T_0. The parameters of the covariance matrices, calculated by Eqs. (30)–(32), are given in rows **R7a** and **R7b** of Table IV; the uncertainty–inverse temperature functions are plotted by dotted and dashed lines, respectively, in Fig. 10.

Again, a proper physical interpretation of the JPL-type uncertainty is to determine the parameters of the covariance matrix by fitting Eq. (36) to the uncertainty values, defined by the intersection of the two gray lines (T_0) and at the extremes of the temperature interval

Figure 10 The JPL-type uncertainty definition for reaction R7 provides minimal uncertainty at 298 K and linearly increasing uncertainties as a function of inverse temperature toward both higher and lower temperatures (solid gray lines). This can be interpreted as a composite of two fully correlated ($r_{\alpha\varepsilon} = +1$) different uncertainty definitions, one valid below 298 K (dotted line; row R7a in Table IV) and the other valid above 298 K (dashed line; row R7b in Table IV). The plotted functions belong to reaction R7.
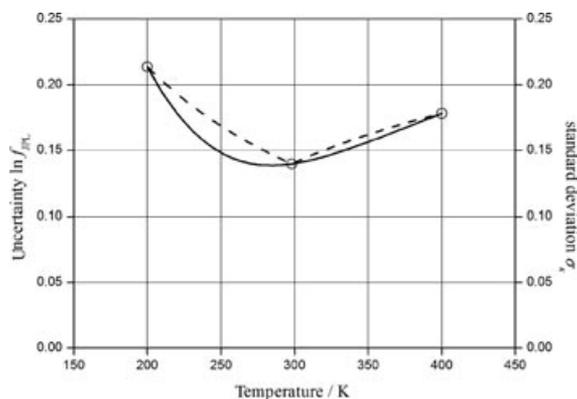
Figure 11 Uncertainties of reaction R7 are defined in the JPL evaluation [20] according to the dashed lines. The uncertainty–temperature function (solid line) was fitted to the points indicated by circles.

(200 and 400 K). The obtained curve (see Fig. 11) passes through the selected points and predicts a lower uncertainty at other temperatures (see row **R7** in Table IV). Given the limited information content behind the defined uncertainties in all kinetic data evaluations, this fitted uncertainty–temperature function defines essentially the same level of uncertainty than the original JPL data, but this function is not arbitrary and has a physical background. Both the current JPL-type uncertainty definition and the uncertainty definition based on the covariance matrix require the same number of parameters; being (T_0, f_0, g) and $(\sigma_\alpha, \sigma_\varepsilon, r_{\alpha\varepsilon})$, respectively.

CONCLUSIONS

Databases of evaluated chemical kinetic data for several thousand gas-phase reactions contain the recommended values of Arrhenius parameters, the temperature interval of applicability, and information on the uncertainty of rate coefficient k. These uncertainty definitions do not take into account that the Arrhenius equation imposes limitations on the temperature dependence of the uncertainty of the rate coefficient.

In the data evaluations commonly used in combustion, the uncertainty is either considered to be constant in some temperature intervals, or it is defined at a few temperatures only. In several cases, the uncertainty information provided in the combustion chemistry databases is not consistent with the Arrhenius form of $k(T)$. In the IUPAC and JPL atmospheric chemical data evaluations, the uncertainties are defined as linear functions of inverse temperature. However, these uncertainty definitions determine different pdfs of the Arrhenius parameters below and above $T_0 = 298$ K, which is physically unrealistic because a unique

Arrhenius expression is considered valid in the whole temperature range.

It is suggested here that the chemical kinetics databases in the future should contain the elements of the covariance matrix of the transformed Arrhenius parameters ($\alpha = \ln A$, n, $\varepsilon = E/R$). This means six figures ($\sigma_\alpha, \sigma_n, \sigma_\varepsilon, r_{\alpha n}, r_{\alpha\varepsilon}, r_{n\varepsilon}$) for a three-parameter modified Arrhenius equation and three figures (σ_α, $\sigma_\varepsilon, r_{\alpha\varepsilon}$ or $\sigma_\alpha, \sigma_n, r_{\alpha n}$) for two-parameter $k(T)$ equations. The elements of the covariance matrix unambiguously determine the temperature dependence of the uncertainty of the rate coefficient in a physically realistic way.

The functional form of the pdf of $\ln k$ is debated. It is assumed in several publications that this pdf can be approximated by a normal distribution at any temperature, truncated at the $\ln k_{min}$ and $\ln k_{max}$ values. The covariance matrix, the bounds for the rate coefficient (defined e.g., by thresholds $\pm 2\sigma$ or $\pm 3\sigma$), and the temperature interval of validity *together* define a truncated multivariate normal distribution for the joint pdf of the transformed Arrhenius parameters. For any temperature, the corresponding pdf of the rate coefficients will be a truncated normal distribution.

In this paper, a method is described for the conversion of the uncertainty information present in the combustion and atmospheric chemical databases to the covariance matrix of the Arrhenius parameters. In some cases, it can be done in an automatic way, but in most cases the uncertainty information present in the databases is either incomplete or inconsistent; therefore a reassessment of the temperature dependence of uncertainties of these rate coefficients is needed. The method is demonstrated on the examples of seven reactions that were selected to show many different types of uncertainty definitions and their related problems.

Using uncertainty analysis, the uncertainty of simulation results can be calculated from the uncertainty of the parameters. Uncertainty analysis of models based on detailed reaction mechanisms is a very important extension to simulation studies. It is needed for the validation of the models, since the uncertainty range of simulation results should overlap with the uncertainty range of measured data. When such models are used in industry, uncertainty analysis indicates the level of reliability and the scope of applicability of the model.

So far, uncertainty information for the rate coefficients has been available in the combustion and atmospheric chemistry databases, but no information was available on the uncertainty of the Arrhenius parameters. Therefore, the kinetic uncertainty studies were based on the uncertainty of k. This approach is correct for constant temperature systems, but as it has

been proven in this paper, in the uncertainty analysis of varying temperature chemical kinetic models, the results obtained from the uncertainty of the rate coefficients are incorrect. Proper uncertainty analysis of nonisothermal chemical kinetic models should be based on the joint pdf of the Arrhenius parameters.

A computer code is available for the determination of the covariance matrix of the Arrhenius parameters from the uncertainty parameter of the rate coefficient in web page http://garfield.chem.elte.hu/Combustion/Combustion.html.

APPENDIX A: CALCULATION OF THE TEMPERATURE-DEPENDENT PDF OF THE RATE COEFFICIENT FROM THE JOINT PDF OF THE TRANSFORMED ARRHENIUS PARAMETERS (α, n, ε)

In this section, the relationship between the pdfs of $\kappa(T)$ and (α, n, ε) is discussed. These are denoted by $\rho_1(\kappa; T)$ and $\rho_3(\alpha, n, \varepsilon)$, respectively.

The pdf of the logarithm of the rate coefficient (κ) at a given temperature can be obtained by integrating the joint pdf of Arrhenius parameters over a subspace corresponding to the constant $\kappa = \theta^{\mathrm{T}}\mathbf{p}$ (see Eq. (10)). This constrained integration can be expressed [37] with the help of the Dirac-delta distribution, as shown below:

$$\rho_1(\kappa; T) = \int_{\Re^3} \mathrm{d}\alpha\, \mathrm{d}n\, \mathrm{d}\varepsilon\; \rho_3(\mathbf{p}) \cdot \delta\left(\theta^{\mathrm{T}}\mathbf{p} - \kappa\right) \quad \text{(A1)}$$

First, let us transform the integration variables from $\mathbf{p}^{\mathrm{T}} = (\alpha, n, \varepsilon)$ to (κ, n, ε) at temperature T, then carry out integration according to variable κ.

$$\rho_1(\kappa; T) = \iiint_{\Re^3} \mathrm{d}\kappa\, \mathrm{d}n\, \mathrm{d}\varepsilon\; \left|\frac{\partial(\alpha, n, \varepsilon)}{\partial(\kappa, n, \varepsilon)}\right|$$
$$\rho_3(\alpha(\kappa, n, \varepsilon; T), n, \varepsilon) \cdot \delta(\theta^{\mathrm{T}}\mathbf{p} - \kappa) \quad \text{(A2)}$$
$$= \iint_{\Re^2} \mathrm{d}n\, \mathrm{d}\varepsilon\; \rho_3\left(\kappa - n\ln T + \varepsilon T^{-1}, n, \varepsilon\right)$$

Here, we made use of the fact that the Jacobian determinant always simplifies to $(\partial\alpha/\partial\kappa)_{n,\varepsilon} = 1$, which shows that this is a one-to-one transformation. Derivations can be carried out in a similar way in the case of the two-parameter Arrhenius equations. No integration is required in the one-parameter case, and $\rho_1(\kappa; T) = \rho_1(\alpha)$ is obtained immediately, which suggests that the probability distribution of the rate coefficient, and therefore its uncertainty should be temperature independent for the constant case. The pdf of rate

coefficient ($\tilde{\rho}_1(k; T)$) can also be calculated from the pdf of $\kappa = \ln k$ ($\rho_1(\kappa; T)$) for any temperature:

$$\int_{\Re} \mathrm{d}\kappa\; \rho_1(\kappa; T) = \int_{\Re^+} \mathrm{d}k\; k^{-1}\rho_1(\ln k; T)$$
$$\Rightarrow \tilde{\rho}_1(k; T) = \rho_1(\ln k; T)\, k^{-1} \quad \text{(A3)}$$

where $\Re$ and $\Re^+$ denote intervals $(-\infty, +\infty)$ and $(0, +\infty)$, respectively. The joint pdf of the transformed Arrhenius parameters $\rho_3(\alpha, n, \varepsilon)$ and the joint pdf of the original Arrhenius parameters $\tilde{\rho}_3(A, n, E)$ uniquely determine each other:

$$\iiint_{\Re^3} \mathrm{d}\alpha\, \mathrm{d}n\, \mathrm{d}\varepsilon\; \rho_3(\alpha, n, \varepsilon)$$
$$= \iiint_{\Re^+\times\Re^2} \mathrm{d}A\, \mathrm{d}n\, \mathrm{d}E\; A^{-1}R^{-1} \cdot \rho_3(\ln A, n, E\,R^{-1})$$
$$\Rightarrow \tilde{\rho}_3(A, n, E) = A^{-1}R^{-1} \cdot \rho_3(\ln A, n, E\,R^{-1})$$
$$\Rightarrow \rho_3(\alpha, n, \varepsilon) = e^{\alpha}R \cdot \tilde{\rho}_3(e^{\alpha}, n, \varepsilon R) \quad \text{(A4)}$$

In this section, we proved that the joint pdf of Arrhenius parameters uniquely determines the pdf of the rate coefficient at all temperatures.

APPENDIX B: INTERRELATION OF RANDOM RATE COEFFICIENTS BELONGING TO DIFFERENT TEMPERATURES

Although so far the joint pdf of the Arrhenius parameters has not been available, during uncertainty analysis of varying temperature chemical kinetic systems (e.g., homogeneous ignition, laminar flame), the uncertainty of κ was always considered to be temperature independent. In Monte Carlo calculations, the parameter A was selected randomly according to a temperature independent uncertainty f, keeping n and E fixed at their nominal values; therefore, at all temperatures k had the same relative deviation from the nominal value. As discussed in the Temperature Dependence of the Uncertainty of the Rate Coefficient section, this approach is physically unrealistic. Moreover, in this section, we show that the proper selection of a random value for k at each temperature cannot be based on the temperature dependence of pdf $\rho_1(\kappa; T)$, but the joint pdf $\rho_3(\alpha, n, \varepsilon)$ of the Arrhenius parameters has to be used instead.

Let κ_{i} denote the random variable of the temperature-dependent random variable $\kappa(T)$

belonging to temperature T_i. Let θ_i denote vector θ (see Eq. (10)) belonging to temperature T_i. The following linear equation can be set up between the Arrhenius parameters and κ_i for i = 1, ..., N, by applying Eq. (10) for $N > 1$ different temperatures and introducing matrix $\Theta = (\theta_1, \ldots, \theta_N)$ and vector $\kappa = (\kappa_1, \ldots, \kappa_N)^T$.

$$\begin{bmatrix} \kappa_1 \\ \vdots \\ \kappa_N \end{bmatrix} = \begin{bmatrix} 1 & \ln T_1 & -T_1^{-1} \\ \vdots & \vdots & \vdots \\ 1 & \ln T_N & -T_N^{-1} \end{bmatrix} \begin{bmatrix} \alpha \\ n \\ \varepsilon \end{bmatrix} \Rightarrow \kappa = \Theta^T \mathbf{p}$$

$$(B1)$$

Monte Carlo analysis of a chemical kinetic system may require the random selection of the rate coefficients at several temperatures simultaneously. However, even if the pdf of a rate coefficient were known at these temperatures, independent random sampling of them would be allowed only if their distributions were independent. If they are not independent (e.g., correlated), then the realistic random sampling of them should be based on the joint pdf of the Arrhenius parameters. The correlation of random variables can be characterized by their covariance matrix.

The covariance matrix Σ_κ of the rate coefficients κ_i belonging to different temperatures relates to the covariance matrix Σ_p of the Arrhenius parameters in the following way:

$$\Sigma_\kappa = \overline{(\kappa - \bar{\kappa})(\kappa - \bar{\kappa})^T} = \Theta^T \overline{(\mathbf{p} - \bar{\mathbf{p}})(\mathbf{p} - \bar{\mathbf{p}})^T} \Theta$$
$$= \Theta^T \Sigma_p \Theta \qquad (B2)$$

First, we exclude the possibility of zero standard deviation of κ for any temperature as it would mean zero uncertainty, which is physically unrealistic. This means that all diagonal elements of matrix Σ_κ are positive.

Rate coefficients κ_i and κ_j belonging to two different temperatures are not correlated if the corresponding (i,j) off-diagonal element of covariance matrix Σ_κ is zero.

If none of the κ_i variables are correlated, then Σ_κ is a diagonal matrix, with positive eigenvalues, in its main diagonal. In this case, it has N linearly independent row vectors (and column vectors), that is its rank would be N, which can be denoted as rank $\Sigma_\kappa = N$. However, the rank of Σ_κ cannot be larger than the minimum of the ranks of matrices in the product $\Theta^T \Sigma_p \Theta$, that is,

$$\text{rank} \Sigma_\kappa \leq \min(\text{rank} \Sigma_p, \text{rank} \Theta) \qquad (B3)$$

Similarly, the rank of matrix Θ cannot be larger than the smaller dimension of the matrix, which means rank $\Theta \leq \min(N, 3)$. However, each column vector θ_i

belongs to different temperatures, thus any two or three of them are linearly independent due to the definition of its elements (see Eq. (10)). This implies equality instead of inequality and thus the rank of matrix Σ_κ can be constrained more precisely:

$$\text{rank} \, \Theta = \min(N, 3)$$
$$\Rightarrow \text{rank} \Sigma_\kappa \leq \min(\text{rank} \Sigma_p, N, 3) \qquad (B4)$$

This means that the rank of matrix Σ_κ cannot be larger than three.

Unit correlation (± 1) between two parameters means complete linear dependence between them, that is, they are a linear function of each other. A composite function of two linear functions is also a linear function (transitivity property), thus if any two of the correlation coefficients among parameters (α, n, ε) have unit absolute value then so does the third one, therefore they fulfil relationship $r_{\alpha n} r_{\alpha\varepsilon} r_{n\varepsilon} = 1$. In this case, the rank of matrix Σ_p is 1 because all of its three column vectors (or row vectors) are parallel with each other. Therefore, the rank of Σ_κ is also one.

If one of the correlation coefficients among Arrhenius parameters (α, n, ε) (denoted by r_{12}), has unit absolute value, then the other two correlation coefficients (denoted by r_{13} and r_{23}) must have equal absolute value and fulfil relationship $r_{13} = r_{12} r_{23}$. In this case, the corresponding two column vectors (or row vectors) of matrix Σ_p will parallel or antiparallel; thus the rank of matrix Σ_p is two. This implies that the rank of Σ_κ is also two. Note that if these two correlation coefficients have an absolute value of unity, then we get back the previous case.

Even if none of the (anti)correlations are perfect, that is none of the column vectors (row vectors) of matrix Σ_p are parallel, they can still be linearly dependent (e.g., $r_{\alpha n} = 2^{-1}$ and $r_{\alpha\varepsilon} = r_{n\varepsilon} = 2^{-1}3^{1/2}$), which also reduces the rank of matrices Σ_p and Σ_κ to 2.

If the column vectors (and also the row vectors) of matrix Σ_p are linearly independent, then the rank of matrix Σ_p is equal to its dimension, which is 3.

Linear dependence of the column vectors means that there exist three coefficients (not all zero) with which the linear combination of the column vectors give a nullvector. This precisely means that the vector formed by these three coefficients is an eigenvector of matrix Σ_p corresponding to a zero eigenvalue. Each independent linear interdependence corresponds to a zero eigenvalue. This means that the rank of a (symmetric) covariance matrix is equal to the number of its nonzero eigenvalues.

A nonzero eigenvalue of a covariance matrix implies that the linear combinations of random variables,

defined by the elements of the corresponding eigenvector, are not uniquely determined. On the contrary, each zero eigenvalue of the covariance matrix uniquely determines the corresponding linear combination of the random variables and thus it reduces the number of dimensions by one, along which the variables can be sampled randomly. The number of independently samplable dimensions is equal to or less than the number of randomly samplable dimensions.

If we assume that $N > 3$ then matrix $\mathbf{\Sigma_\kappa}$, which has a dimension of $N \times N$, cannot be diagonal otherwise its rank would be higher than 3. Therefore, the κ variables belonging to three or more different temperatures are inherently correlated; thus they cannot be sampled independently. This means that the κ variables belonging to temperatures $T_4, \ldots, T_N$ must be correlated with at least one of the κ variables belonging to temperatures T_1, T_2, T_3, and similarly, there must be a correlation between the corresponding κ variables for any selection of these three and $N-3$ temperatures.

As to whether there is a correlation between the κ variables belonging to any three temperatures, assume that $N = 3$, which simplifies equation (B2) to a diagonalization problem of matrix $\mathbf{\Sigma_p}$. As matrix $\mathbf{\Sigma_p}$ is symmetric, an orthogonal basis can be selected from its eigenvectors and the $\mathbf{\Theta}$ matrix formed by them as column vectors can make $\mathbf{\Sigma_p}$ diagonal via transformation $\mathbf{\Theta}^T\mathbf{\Sigma_p}\mathbf{\Theta}$. Note that this diagonal matrix is different from the diagonal matrix of eigenvalues, because vectors $\mathbf{\theta}_i$ are not normalized, due to their fixed form (see Eq. (10)). A sufficient requirement for such a matrix is that its $\mathbf{\theta}_i$ vectors are orthogonal, which requires that $\ln T_i \cdot \ln T_j < 0$ for all i, j pairs, because the first (1) and third $(-T_i^{-1})$ components of these vectors have fixed, positive, and negative signs, respectively. The condition that $\ln T_i \cdot \ln T_j < 0$ cannot be fulfilled for any three temperatures simultaneously.

A similar train of thought applies to the one-parameter and to both of the two-parameter Arrhenius equations with the appropriate substitutions, with an exception only in the power-type expression. In the power-type Arrhenius equation where only two temperatures are considered, the condition that $\ln T_1 \cdot \ln T_2 < 0$ and the orthogonality condition can be fulfilled if one of the temperatures is lower than 1 K, which is chemically irrelevant.

Unfortunately, the orthogonality requirement of vectors $\mathbf{\theta}_i$ is not a necessary one to obtain a diagonal matrix by $\mathbf{\Theta}^T\mathbf{\Sigma_p}\mathbf{\Theta}$ transformation. The matrix $\mathbf{\Theta}$ is invertable, thus assuming a diagonal matrix $\mathbf{\Sigma_\kappa}$ for the κ variables belonging to temperatures defined in $\mathbf{\Theta}$, the corresponding $\mathbf{\Sigma_p}$ matrix can be calculated by equation (B5). This provides a proper covariance matrix, despite the column vectors of matrix $\mathbf{\Theta}$ being nonorthogonal.

$$\mathbf{\Sigma_p} = (\mathbf{\Theta}^T)^{-1}\mathbf{\Sigma_\kappa}\mathbf{\Theta}^{-1} = (\mathbf{\Theta}^{-1})^T\mathbf{\Sigma_\kappa}\mathbf{\Theta}^{-1} \quad \text{(B5)}$$

The requirements for matrix $\mathbf{\Sigma_p}$ and solution for the temperatures can be obtained by expressing the off-diagonal elements of $\mathbf{\Sigma_\kappa}$ from equation (B2) and setting them to zero. This means one and three equations for two and three temperatures in the two- and three-parameter cases, respectively. This confines the number of the solutions to one freely selectable temperature in the two-parameter cases (the dimension of the solution in space $\mathfrak{R}^2$ is one), and zero freely selectable temperatures in the three-parameter cases (the dimension of the solution in space $\mathfrak{R}^3$ is zero). Requirements for matrix $\mathbf{\Sigma_p}$ can also be derived, which show that selection of such temperatures is impossible for all $\mathbf{\Sigma_p}$ matrices. For the $\mathbf{\Sigma_p}$ matrices that fulfil these conditions, some selections of two or three temperature points exist for the two- and three-parameter cases, respectively, where sampling of the rate coefficients can be done without correlation. Thus they might be sampled independently if uncorrelatedness implies independency, which is valid for a normal distribution. However, the selection of temperatures cannot be given in advance and the joint pdf of the Arrhenius parameters should be determined first.

In summary, the random rate coefficients belonging to different temperatures are intrinsically correlated due to the existence of the joint pdf of the Arrhenius parameters. The number of temperatures where $\kappa = \ln k$ variables are uncorrelated is not more than the number of parameters in the Arrhenius equation. Independent sampling of the rate coefficients at different temperatures in a Monte Carlo study leads to systematic errors. In varying temperature chemical kinetic systems, application of the joint pdf of the Arrhenius parameters is crucial for a realistic uncertainty analysis because it allows a correct sampling of κ in the temperature range relevant for the simulation.

APPENDIX C: ASSUMPTION OF A LOGNORMAL DISTRIBUTION FOR THE RATE COEFFICIENT

The joint pdf of a uni- or multivariate normal distribution of parameter vector $\mathbf{x}$ of dimension L is defined in the equation (C1). It is completely determined by the vector $\bar{\mathbf{x}}$ of expectation values and the covariance matrix $\mathbf{\Sigma_x}$.

$$g_L(\mathbf{x}; \bar{\mathbf{x}}, \mathbf{\Sigma_x}) = \frac{\exp\left[-\frac{1}{2}(\mathbf{x} - \bar{\mathbf{x}})^T \mathbf{\Sigma_x}^{-1}(\mathbf{x} - \bar{\mathbf{x}})\right]}{(2\pi)^{L/2}\sqrt{\det\mathbf{\Sigma_x}}}$$

$$\text{(C1)}$$

In the case of multivariate normal distribution, zero correlation between two parameters means total independence.

If the transformed Arrhenius parameters (α, n, ε) have a multivariate normal distribution, then the natural logarithms of rate coefficients κ_i for any $N \geq 1$ number of temperatures follows a univariate $(L = 1)$ or a multivariate normal distribution $(L > 1)$, as they are linear functions of the transformed Arrhenius parameters (see equation (B1)) [35]. The expectation value of vector κ is $\bar{\kappa} = \boldsymbol{\Theta}^T \bar{\mathbf{p}}$ and its covariance matrix can be obtained from equation (B2).

It has been proven in Appendix B that if matrix $\boldsymbol{\Theta}$ is constructed for three (or two) different temperatures, then the column vectors of matrix $\boldsymbol{\Theta}$ are linearly independent in the three- (or two-)parameter cases, respectively, thus the square matrix $\boldsymbol{\Theta}$ is invertable. Therefore, the transformed Arrhenius parameters are also unique linear functions of the κ variables belonging to the corresponding three (or two) temperatures. Thus, if we assume that these κ variables follow a multivariate normal distribution then the transformed Arrhenius parameters (α, n, ε) also follow a multivariate normal distribution. The expectation value of $\mathbf{p}$ is $\bar{\mathbf{p}} = \left(\boldsymbol{\Theta}^T\right)^{-1} \bar{\kappa}$ and their covariance matrix can be obtained from equation (B5). This also implies that the κ variables for any N number of temperatures follow a multivariate normal distribution, and its covariance matrix has a maximum rank of $\min(\text{rank}\boldsymbol{\Sigma}_{\mathbf{p}}, N, 3)$ (or $\min(\text{rank}\boldsymbol{\Sigma}_{\mathbf{p}}, N, 2)$).

Let us assume that at all temperatures the logarithm of the rate coefficients follows a univariate normal distribution, but there is no information about their interdependence. At each temperature, the corresponding κ_i variable represents a linear combination of parameters (α, n, ε) with coefficients $\theta_i^T = (1, \ln T_i, -T_i^{-1})$, denoting a direction in space, along which the distribution is normal. If all linear combinations of parameters (α, n, ε) follow a univariate normal distribution then they also follow a multivariate normal distribution [35]. By changing the temperature, these θ_i vectors designate only a one-dimensional (1D) subspace of directions in the 2D direction space defined by vectors in the 3D space of (α, n, ε). Thus, this condition is not sufficient to state that any possible linear combination of (α, n, ε) parameters are normal, which would mean that they form a multivariate normal distribution [35]. It is also not sufficient in the two-parameter (α, ε) case, where coefficients $(1, -T_i^{-1})$ can't cover directions parallel with vectors $(1, \text{non-negative})$ or $(0, 1)$ in the 1D direction space defined by vectors in the 2D space of (α, ε). However, it is a sufficient assumption in the two-parameter (α, n) case, where coefficients $(1, \ln T_i)$ can cover all

directions apart from the one along vector $(0, 1)$. However, along direction $(1, \ln T_i)$ the distribution can also be considered as normal due to the basic requirement of continuity of the joint pdf of parameters (α, n). Thus, apart from the two-parameter (α, n) case, we have to extend analytically our normal distributions to the rest of the directions in the parameter space to get the multivariate distribution for the transformed Arrhenius parameters, if we don't assume anything on the interrelation between the κ parameters belonging to the different temperatures.

In this section, we showed that $\ln k$ following a normal distribution at each temperature does not necessarily imply that the transformed Arrhenius parameters (α, n, ε) also follow a multivariate normal distribution. If $\ln k$ follows a normal distribution at a given temperature, then a cross-section of the joint pdf of the Arrhenius parameters also follow a normal distribution. Therefore, if $\ln k$ follows a normal distribution at many temperatures, then the most natural assumption is that (α, n, ε) also follows a multivariate normal distribution. A joint multivariate normal distribution of Arrhenius parameters (α, n, ε) implies a normal distribution for $\ln k$ at all temperatures.

APPENDIX D: SAMPLING FROM A MULTIVARIATE NORMAL DISTRIBUTION

One of the applications of the joint pdf of the Arrhenius parameters is the Monte Carlo uncertainty analysis of models based on detailed chemical kinetic mechanism. It is not trivial how a random sampling of the Arrhenius parameters can be carried out that corresponds to their pdf. This topic is discussed here briefly.

Each eigenvalue λ_i and the corresponding eigenvector $\mathbf{o}_i$ of the covariance matrix define an independent normal distribution along a line in the space of variables α, n and ε.

$$\mathbf{O}^T \boldsymbol{\Sigma}_{\mathbf{p}} \mathbf{O} = \boldsymbol{\Lambda} = \text{diag}(\lambda_1, \lambda_2, \lambda_3) \qquad \text{(D1)}$$

where $\mathbf{O} = \begin{bmatrix} \mathbf{o}_1 & \mathbf{o}_2 & \mathbf{o}_3 \end{bmatrix}$ and $\mathbf{O}^T = \mathbf{O}^{-1}$. The random variables t_i are defined in such a way that the variables corresponding to nonzero eigenvalues follow a standard normal distribution and those corresponding to zero eigenvalues have a definite zero value as they follow a Dirac-delta distribution:

$$t_i = \begin{cases} 0 & \text{if } \lambda_i = 0 \\ \mathbf{o}_i^T (\mathbf{p} - \bar{\mathbf{p}}) / \sqrt{\lambda_i} & \text{if } \lambda_i \neq 0 \end{cases} \qquad \text{(D2)}$$

The corresponding random values of Arrhenius parameters can be obtained using the

following equation:

$$\mathbf{p} = \bar{\mathbf{p}} + \mathbf{O}\mathbf{\Lambda}^{1/2}\mathbf{t} = \bar{\mathbf{p}} + \sum_{i:\lambda_i>0} \sqrt{\lambda_i}t_i\mathbf{o}_i \qquad (D3)$$

As discussed above, if the Arrhenius parameters are sampled according to a multivariate normal distribution, the rate coefficients calculated at any temperature will follow a univariate normal distribution.

$$g_1\left(\kappa;T\right) = \frac{1}{\sqrt{2\pi}\,\sigma_\kappa(T)}\exp\left[-\frac{\left(\kappa(T) - \overline{\kappa(T)}\right)^2}{2\sigma_\kappa^2(T)}\right]$$
$$(D4)$$

To avoid using physically unrealistic high and low rate coefficients, the calculated distribution should be truncated at $m \cdot \sigma_\kappa$ ($m = 2$ or 3). The joint pdfs have to be truncated *a posteriori* to ensure that values $\kappa(T)$ always remain within the allowed uncertainty range at any temperature T within interval $[T_1,\ T_2]$:

$$|\kappa - \bar{\kappa}(T)| \le m \cdot \sigma_\kappa(T), \quad T \in [T_1, T_2] \qquad (D5)$$

Assuming a normal distribution, truncation at $\pm 2\sigma$ and $\pm 3\sigma$ excludes 5% and 0.3% of all cases, respectively. Therefore, this truncation only slightly affects the elements of covariance matrix (α, n, ε) and the normalization of their joint pdf. This way, the relationship between the variance of $\kappa(T)$ and the covariance matrix of the Arrhenius parameters (α, n, ε), as discussed in the Temperature Dependence of the Uncertainty of the Rate Coefficient section, remains valid after truncation within a good approximation.

Note that several sources recommend truncation at 2σ or 3σ without making an assumption about the distribution. However, the Chebyshev's inequality theorem shows [34] that the range defined by 1σ, 2σ, and 3σ radii may contain 0%–100%, 75%−100%, and 88.9%–100% of the random values, respectively, depending on the distribution. This implies that the recommendation for a truncation at 2σ may exclude one quarter of the values in the worst case, which is not negligible.

BIBLIOGRAPHY

1. Cvetanović, R. J.; Overend, R.; Paraskevopoulos, G. Int J Chem Kinet 1975, S1, 249–271.
2. Cvetanović, R. J.; Singleton, D. L. Int J Chem Kinet 1977, 9, 481–488.
3. Cvetanović, R. J.; Singleton, D. L.; Paraskevapoulos, G. J Phys Chem 1979, 83, 50–60.
4. Héberger, K.; Kemény, S.; Vidóczy, T. Int J Chem Kinet 1987, 19, 171–181.
5. Klička, R.; Kubáček, L. Chemomet Intell Lab Syst 1997, 39, 69–75.
6. Sundberg, R. Chemomet Intell Lab Syst 1998, 41, 249–252.
7. Rodríguez-Aragón L. J.; López-Fidalgo, J Chemomet Intell Lab Syst 2005, 77, 131–138.
8. Schwaab, M.; Pinto, J. C. Chem Eng Sci 2007, 62, 2750–2764.
9. Schwaab, M.; Lemos, L. P.; Pinto, J. C. Chem Eng Sci 2008, 63, 2895–2906.
10. Schwaab, M.; Pinto, J. C. Chem Eng Sci 2008, 63, 4631–4635.
11. Najm, H.; Debusschere, B. J.; Marzouk, Y. M.; Widmer, S.; Le Maître, O. P. Int J Numer Meth Engng 2009, 80, 789–814.
12. Baulch, D. L.; Cobos, C. J.; Cox, R. A.; Esser, C.; Frank, P.; Just, T.; Kerr, J. A.; Pilling, M. J.; Troe, J.; Walker, R. W.; Warnatz, J. J Phys Chem Ref Data 1992, 21, 411.
13. Baulch, D. L.; Cobos, C. J.; Cox, R. A.; Frank, J. H.; Hayman, G.; Just, T. H.; Kerr, J. A.; Murrels, T.; Pilling, M. J.; Troe, J.; Walker, B. F.; Warnatz, J. Combust Flame 1994, 98, 59–79.
14. Baulch, D. L.; Bowman, C. T.; Cobos, C. J.; Cox, R. A.; Just, T.; Kerr, J. A.; Pilling, M. J.; Stocker, D.; Troe, J.; Tsang, W.; Walker, R. W.; Warnatz, J. J Phys Chem Ref Data 2005, 34, 757–1397.
15. IUPAC, Subcommittee for Gas Kinetic Data Evaluation, in http://www.iupac-kinetic.ch.cam.ac.uk/, 2010.
16. Atkinson, R.; Baulch, D. L.; Cox, R. A.; Crowley, J. N.; Hampson, R. F.; Hynes, R. G.; Jenkin, M. E.; Rossi, M. J.; Troe, J. Atmos Chem Phys 2004, 4, 1461–1738.
17. Atkinson, R.; Baulch, D. L.; Cox, R. A.; Crowley, J. N.; Hampson, R. F.; Hynes, R. G.; Jenkin, M. E.; Rossi, M. J.; Troe, J.; IUPAC_Subcommittee. Atmos Chem Phys 2006, 6, 3625–4055.
18. Atkinson, R.; Baulch, D. L.; Cox, R. A.; Crowley, J. N.; Hampson, R. F.; Hynes, R. G.; Jenkin, M. E.; Rossi, M. J.; Troe, J. Atmos Chem Phys 2007, 7, 981–1191.
19. Atkinson, R.; Baulch, D. L.; Cox, R. A.; Crowley, J. N.; Hampson, R. F.; Hynes, R. G.; Jenkin, M. E.; Rossi, M. J.; Troe, J.; Wallington, T. J. Atmos Chem Phys 2008, 8, 4141–4496.
20. Sander, S. P.; Friedl, R. R.; Ravishankara, A. R.; Golden, D. M.; Kolb, C. E.; Kurylo, M. J.; Molina, M. J.; Moortgat, G. K.; Keller-Rudek, H.; Finlayson-Pitts, B. J.; Wine, P. H.; Huie, R. E.; Orkin, V. L. "Chemical Kinetics and Photochemical Data for Use in Atmospheric Studies. Evaluation Number 15", JPL Publication 06-2, Jet Propulsion Laboratory, Pasadena, 2006.

21. Warnatz, J. In Combustion Chemistry, Gardiner, W. C. (Ed.); Springer: New York, 1984; pp. 197–361.
22. Konnov, A. A. Combust Flame 2008, 152, 507–528.
23. Baulch, D. L. Personal Communication 1995.
24. Brown, M. J.; Smith, D. B.; Taylor, S. C. Combust Flame 1999, 117, 652–656.
25. Turányi, T.; Zalotai, L.; Dóbé, S.; Bérces, T. Phys Chem Chem Phys 2002, 4, 2568–2578.
26. Zsély, I. G.; Zádor, J.; Turányi, T. Int J Chem Kinet 2008, 40, 754–768.
27. Zsély, I. G.; Zádor, J.; Turányi, T. Proc Combust Inst 2005, 30, 1273–1281.
28. Zádor, J.; Zsély, I. G.; Turányi, T.; Ratto, M.; Tarantola, S.; Saltelli, A. J Phys Chem A 2005, 109, 9795–9807.
29. Zádor, J.; Zsély, I. G.; Turányi, T. Reliability Eng System Safety 2006, 91, 1232–1240.
30. Sheen, D. A.; You, X.; Wang, H.; Løvås, T. Proc Combust Inst 2009, 32, 535–542.
31. International vocabulary of metrology – Basic and general concepts and associated terms (VIM), in http://www.bipm.org/, 2008, JCGM 200.
32. Zádor, J.; Wagner, V.; Wirtz, K.; Pilling, M. J. Atm Environ 2005, 39, 2805–2817.
33. Zádor, J.; Turányi, T.; Wirtz, K.; Pilling, M. J. J Atmos Chem 2006, 55, 147–166.
34. Rao, C. R. Linear Statistical Inference and Its Applications, 2nd Ed.; Wiley: New York, 1973.
35. Schittkowski, K. EASY-FIT Express, Version 1.0, in http://num.math.uni-bayreuth.de/~kschittkowski/easy_fitx.htm, 2009.
36. Saltelli, A.; Tarantola, S.; Campolongo, F.; Ratto, M. Sensitivity Analysis in Practice. A Guide to Assessing Scientific Models, Wiley: Chichester, 2004.
37. Bronshtein, I. I. N.; Semendyayev, K.; Musiol, G.; Muehlig, H.; Mühlig, H. Handbook of Mathematics, Springer: Berlin, 2004.