Manual of code *u-Limits2*

Manual version 1.0 / Code version 1.0

Éva Valkó

Laboratory for Chemical Kinetics, Institute of Chemistry,
Eötvös University (ELTE), Budapest, Hungary
MTA-ELTE Research Group on Complex Chemical Systems, Budapest, Hungary
e-mail: vviiccaa@gmail.com

1. Scope

Code *u-Limits* facilitates the assignation of the uncertainty band of the rate coefficient−temperature function of an elementary reaction based on literature data. It also identifies the uncertainty domain of the Arrhenius parameters and stores this information in the form of a covariance matrix. The code provides a visualization of the process and assists the selection from several possible choices. The theory behind this code has been described in the article of Nagy *et al.* [1] and [2].

A separate text input file is needed for each investigated elementary reaction. The first lines of this text input file follow the format of the summary page of the NIST database [3]. This means that each line contains a literature identifier (which is the NIST squib if it exists), the temperature range ([T_{min},T_{max}] in K units), Arrhenius parameters (ln A, n, E/R; units: cm, mole, s, K). These lines can be copied from the NIST summary Web pages. Information obtained from other sources has to be encoded in a similar way. Arrhenius plots referring to different bath gases (*e.g.* data belonging to reactions $H+O_2+N_2=HO_2+N_2$ and $H+O_2+Ar=HO_2+Ar$) can be joined and processed together by assuming a temperature-independent 3rd body collision efficiency of the molecules of the bath gas relative to nitrogen. The input contains the Arrhenius parameters of the selected mean rate coefficient expression and the range of temperature of the analysis.

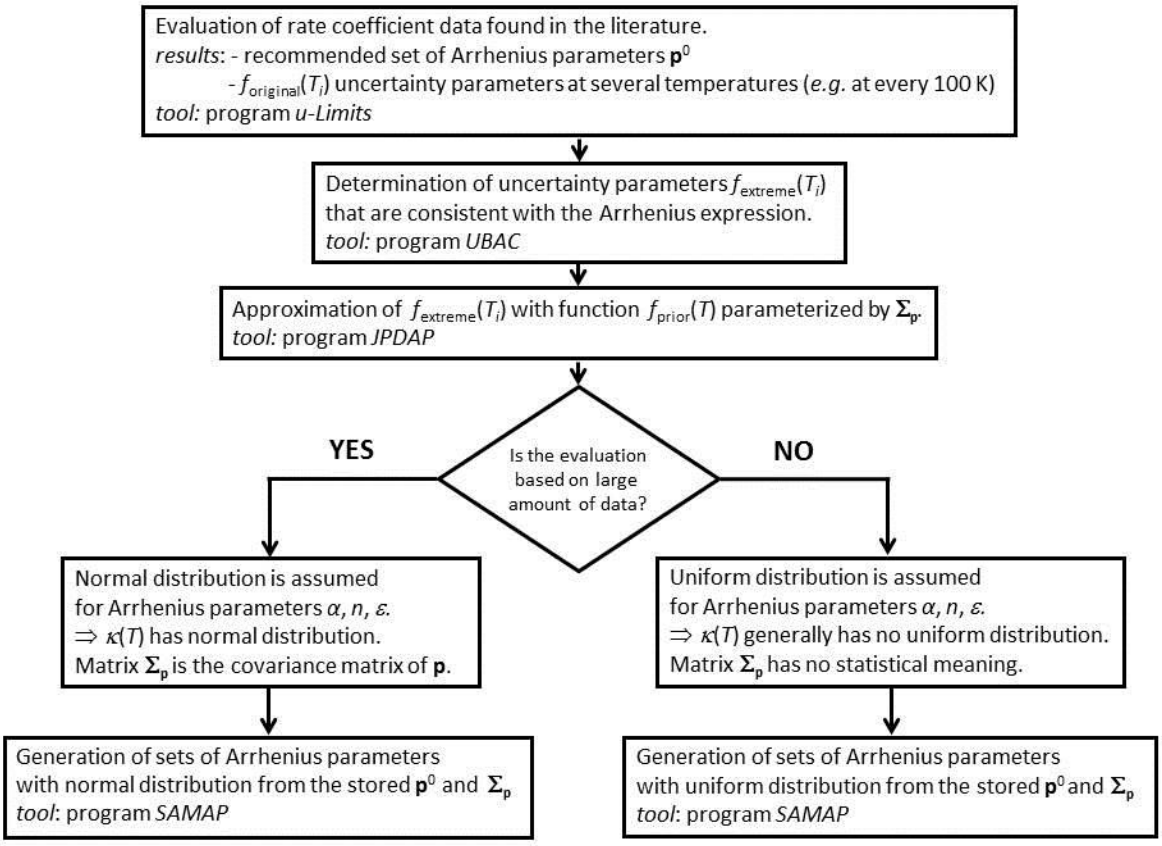
At equidistant temperature points T_i, the program determines empirical uncertainty $f(T_i)$ as the larger of two distances on a decimal logarithmic scale between the mean rate coefficient $\log_{10} k^0(T_i)$ and the upper and lower extreme rate coefficient values $\log_{10} k^{max}(T_i)$ and $\log_{10} k^{min}(T_i)$, respectively. The recommended temperature interval is 100 K, but any other value

can be defined by the user. The automatically calculated $f(T_i)$ values can be manually revised. Such corrections are needed when the automatically determined f values are unrealistically small in a temperature range, which may happen, if in this temperature range all available (typically few) data points are close to the mean curve. Another way of manual intervention is omitting those rate coefficients that unrealistically widen the band of uncertainty. For many elementary reactions, the oldest measurements provided rate coefficient values that are very far from the recently accepted values. Usually, the initially applied experimental method was later superseded by newer techniques, which known to have smaller systematic error. In such cases the values obtained by obsolete methods are not considered at the determination of the uncertainty ranges. These data are not deleted from the input text file, but are flagged as not used ones. The automatically generated f values together with these manual corrections provide the $f_{\text{original}}(T_i)$ uncertainty parameter values.

Matlab code *u-Limits* prepares the input text files for executable (Fortran) codes UBAC and JPDAP, runs these codes, and visualizes their results. One of the generated plots shows the $f_{\text{original}}(T_i)$ points, and the $f_{\text{extreme}}(T)$ and $f_{\text{prior}}(T)$ functions determined by codes UBAC and JPDAP. Another generated figure is an Arrhenius plot that shows all considered $\log_{10} k$ *vs.* $1/T$ functions, together with the mean line, and the upper and lower uncertainty limits calculated from $f_{\text{prior}}(T)$. This allows the user to check if the determined uncertainty range of Arrhenius parameters is consistent with all data considered.

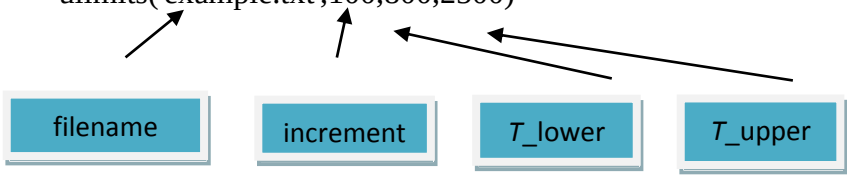
Codes UBAC and JPDAP are separate Fortran codes, which have much additional features beyond the ones used by *u-Limits*. For the further features of these codes the user is directed to the corresponding program manuals [4] [5].

The flowchart of *u-Limits*:



Calling program *u-Limits* in Matlab:

```
>> ulimits('example.txt',100,800,2500)
```



filename	name of the text input file
increment:	resolution of the temperature range (unit: K) (*e.g.* 1 K, 10 K, 50 K, 100 K)
*T*_lower	lower limit of the temperature range (unit: K)
*T*_upper	upper limit of the temperature range (unit: K)

2. The input text file

Columns of the text file have to be separated by spaces. The following lines contain the definition of the columns.

column 1: comments to the data

Possible entries: any text, for example review/experiment/theory or forward/reverse.

If the column contains 0, then the type is identical to those of the previous line.

This information will be part of the label that belongs to the plotted data.

column 2: ID of the paper

Possible entries: any unique identifier (recommended: the NIST squib if it exists)

column 3: flag of the data

Possible entries:	0	these data are not used
	1	these data are used and contain a single Arrhenius expression
	2	these data are used and contain a double Arrhenius expression

column 4: lower limit of temperature range within the data source (T_1) (unit: K)

column 5: upper limit of temperature range within the data source (T_2) (unit: K)

examples:	300	2000	(temperature range of 300 K – 2000 K)
	298	298	(data for T=298 K only)
	298	0	(data for T=298 K only)

column 6: value of the Arrhenius parameter A (unit: cm^3, mole, s)

column 7: value of the Arrhenius parameter n (unit: −)

column 8: value of the Arrhenius parameter E/R (unit: K)

column 9: value of the rate coefficient k at 298 K (unit: cm^3, mole, s)

This column is kept for preserving the NIST summary page format. The entered value is not used and numeric value can be given.

column 10: kinetic order of the reaction

Possible entries are 1, 2, 3. The given value is not used by the code, but it provides info for the user.

column 11: uncertainty given in the original paper is used in the present evaluation

The general recommendation [1] is that the uncertainty of the rate coefficient as defined in the original paper shall not be used for the calculation of the uncertainty band. However, the utilization of it might be necessary when an important measurement is at the edge of the uncertainty band. Therefore, the code allows the consideration of this type of information.

Possible entries:

0	uncertainty parameter f as stated in the original paper is not used.
1	temperature independent uncertainty parameter f is provided in the original paper and used here
2	uncertainty parameter values are stated at two temperature (T_1 and T_2) in the original paper and used here
3	uncertainty parameter f is stated at three temperatures (T_1, T_m and T_2) in the original paper and used here

Using flags 1 to 3, the upper and lower uncertainty limits (as recommended in the original publications) of the rate coefficients are calculated, and the $\log_{10} k^{min}(T)$ and $\log_{10} k^{max}(T)$ lines obtained are handled like the mean $\log_{10} k(T)$ lines obtained from other measurements or calculations. Using flag 1, parameter f is considered to be temperature independent. Using flag 2, the $f(T)$ value is obtained by linear interpolation on the $1/T$ scale from the $f(T_1)$ and $f(T_2)$ values. Using flag 3, the $f(T)$ value is determined by linear interpolation on the $1/T$ scale from either the $f(T_1)$ and $f(T_m)$ or $f(T_m)$ and $f(T_2)$ values. T_m is defined in column 15 (if applicable).

column 12: value of the parameter f at T_1

column 13: value of the parameter f at T_2

column 14: value of the parameter f at T_m

column 15: T_m (unit: K)

The contents of columns 12−15 are used only if a flag other than zero was specified in column 11. Otherwise any numeric values can be placed here, these are not interpreted.

Column 16: third body collision efficiency factor *m*

This column is relevant when the investigated elementary reaction is pressure dependent (using "+M" in Chemkin reaction coding)and the rate parameters refer to the low-pressure limit. The given third body collision efficiency factor is considered to be temperature-independent. It belongs to the bath gas used in the experiment, and it is relative to nitrogen. Examples: for nitrogen bath gas m=1, for argon bath gas m=0.5 (since the ratio of collision efficiencies is $m(\mathrm{Ar})/m(\mathrm{N_2})$= 0.5), etc.

column 17: value of Arrhenius parameter *A* in the second set. (unit: $\mathrm{cm^3}$, mole, s)

column 18: value of Arrhenius parameter *n* in the second set. (unit: −)

column 19: value of Arrhenius parameter *E/R* in the second set. (unit: K)

The values of the second set of Arrhenius parameters are not used, if a single Arrhenius expression is declared according to column 3.

column 20: if extended Arrhenius expression $k = A \cdot T^n \cdot \exp\left(\frac{-E/R\,(T+T_0)}{(T+T_0)^2}\right)$ is used the value of T_0 can be defined in this column. If single Arrhenius expression is assumed please set the value of T_0=0. This extension with parameter T_0 is not available in the case of duplicate Arrhenius expression! (unit: K)

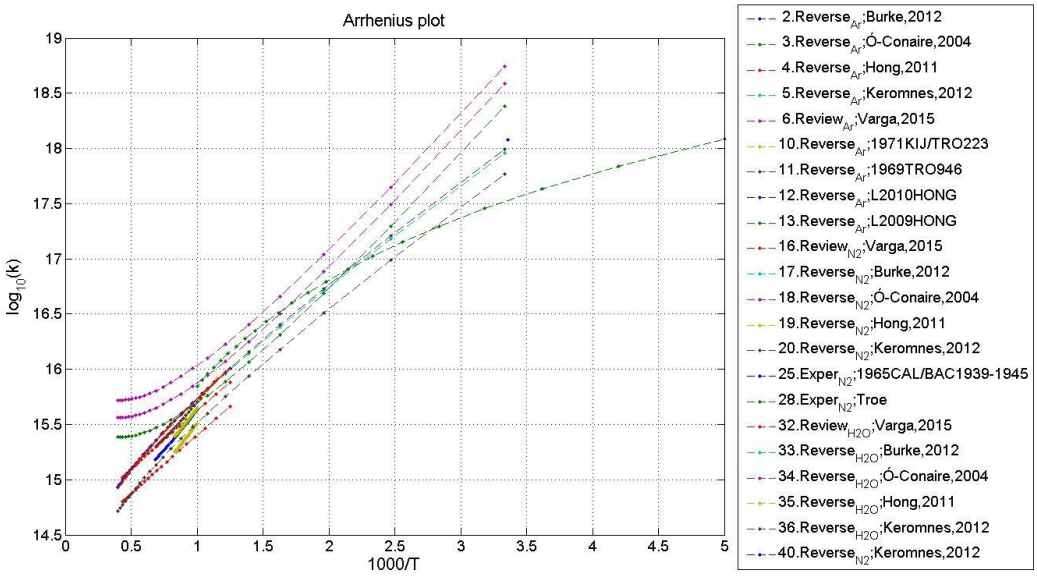
Typical lines in the text file

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Review	Baulch,2005	2	550	1250	4.22E+14	0	6030	0	2	0	0	0	0	0	1	1.32E+11	0	-820	200
Experiment	2006THI/FIT 383-389	1	298	0	1.14E+12	0	0	0	2	0	0	0	0	0	1	0	0	0	0
0	1997SEH/MOG 673-682	0	295	0	2.11E+12	0	0	0	2	0	0	0	0	0	1	0	0	0	0

The last entry defines the mean Arrhenius parameters. This line contains the same number of records as the previous lines, but only the columns containing Arrhenius parameters are interpreted. This means columns 1-3 and 6-8 for a single Arrhenius expression and columns 1-3, 6-8 and 17-19 for a double Arrhenius expression. As discussed in the article of Nagy *et al.* [1], the mean Arrhenius parameter set is not necessarily the best or the recommended one, but it will represent the middle line of the uncertainty band of the rate coefficient.

3. Running the program

Program *u-Limits* creates an Arrhenius plot of all data given in the input text file. Data with zero flag in column 3 are not plotted. The Arrhenius plot spans temperature range *T*_lower – *T*_upper. The Arrhenius expressions (single or double) are plotted as lines in temperature range T_1 – T_2 (see columns 4 and 5), whereas rate coefficient data belonging to a single temperature T_1 (see column 4) are indicated by dots. The legend of the lines is given on the right hand side of the plot. The Arrhenius plot also shows a red line that belongs to mean Arrhenius expression in temperature range *T*_lower – *T*_upper. The *figure* tool of Matlab allows the manipulation of the figure, like zooming in to interesting parts or reading the numerical values belonging to a given point. This tool gives the label of the data by clicking on the corresponding line or dot. If the original experimental or theoretical data are defined by a double Arrhenius expression, then *u-Limits* plots the corresponding curves. These functions are handled in an identical way to those that are defined by a single Arrhenius.



At this point the program asks if the plotted data are acceptable without a manual correction.
>> Do you want to use all plotted data for the determination of the uncertainty band? (y/n)

If the user answers yes (y), then the code continues with the calculation of the *f* values (see section 4). If the user answers no (n), the following message is printed:
>> Press ENTER if you finished the modification of the input data file!

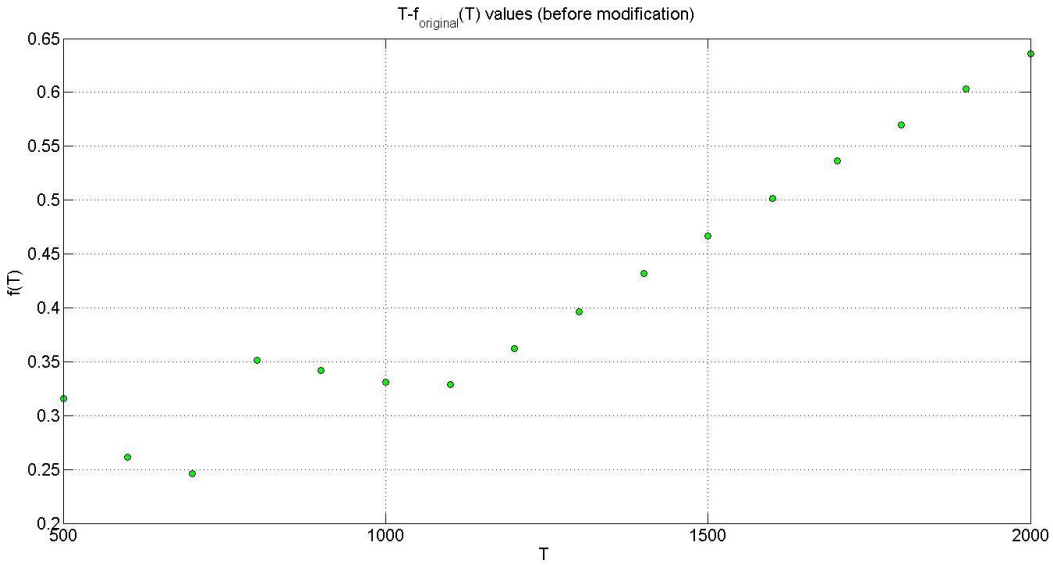
At this point the user is allowed to change the original data file using an external text editor. The user may select some of the data not to be considered by setting the flag to 0 in column 3 of the corresponding line. Having finished the changes, the code may be restarted by pressing <ENTER> and the figure above is redrawn. The procedure for the modification of the input file may be repeated until the user becomes satisfied.

4. Calculation of the $f_{original}$ points

After the selection of the applicable literature rate expressions, the program plots the $f_{original}$ points in the temperature range of *T*_lower to *T*_upper, at every temperature which is a factor of

the increment, that is at every "$T_{start}+n\times increment$" ($n$ = 0, 1, …) temperatures. The T_{start} temperature is the smallest temperature that is higher than T_lower and is a factor of the *increment*. The T_{end} temperature is the highest temperature that is lower than T_upper and is a factor of the *increment*.

For example, if T_lower = 298 K, T_upper = 2510 K and *increment* = 50 K, then $f_{original}$ is plotted at temperatures 300 K, 350 K, 400 K, …, 2500 K. The $f_{original}$ points are saved in data file "T_f_points.txt".



At this point the program asks if the plotted data are acceptable without a manual correction.

>> Would you like to modify the T-f(T) values? (y/n)

If the user answers no (n), then the code continues with the manual selection of the assumed distribution type and the automatic determination of the covariance matrix (see section 6). If the user answers yes (y), then the code lets the user adjust the automatically determined uncertainty parameter values manually (see section 5).

5. Correction of the uncertainty parameters *f*

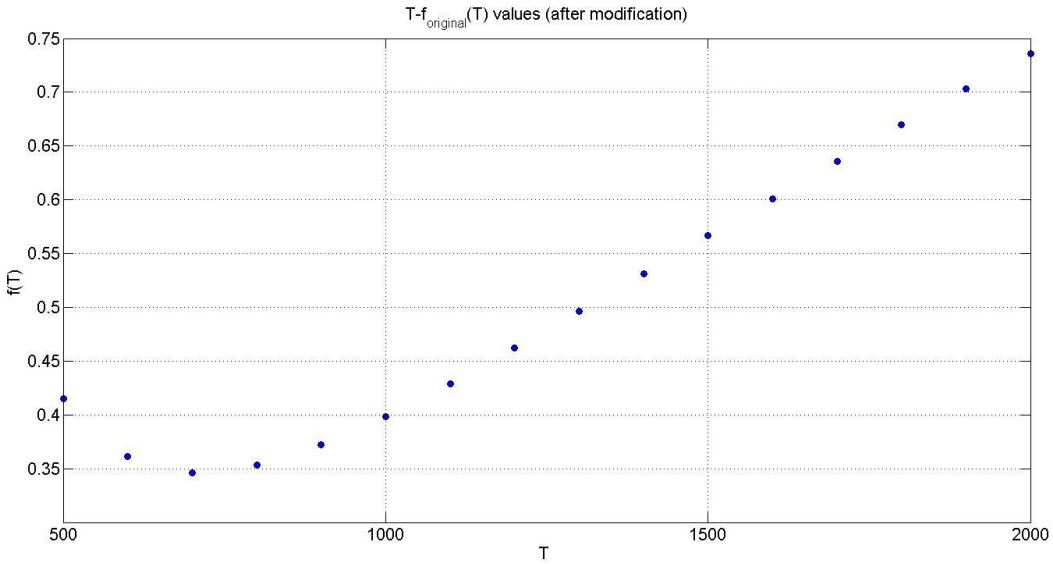
The user may override the automatically determined *f* values by editing the "T_f_points.txt" file with an external text editor. The program is waiting for finishing the correction of the text file.
>> Press ENTER if you have finished the modification of T_f_points.txt!

An example for the adjustments:

Automatically generated file (part):
300 0.52
400 0.08
500 0.44
…

Manually corrected file (part):
300 0.52
400 0.48
500 0.44
…

Pressing enter, the program loads the corrected $T - f$ points from file "T_f_points.txt" and plots them:



The program then asks again:

>> Would you like to modify the T-f(T) values? (y/n)

If the user answers no (n), then the code continues with the manual selection of the assumed distribution type and the automatic determination of the covariance matrix (see section 6). If the user answers yes (y), then the code offers further manual corrections of the uncertainty parameters (see the beginning of section 5).

6. Determination of the assumed distribution and the covariance matrix

From this point the code interprets the actual content of data file "T_f_points.txt".
The next question refers to the assumed distribution:

>> Do you assume Uniform or Normal distribution? (U,N)

Choosing U:
This option should be selected [1] when very little information is available for the uncertainty of the rate coefficient and usually the uncertainty parameter is considered to be temperature independent due to the lack of information. In this case, uniform distribution is assumed for the domain of the Arrhenius parameters based on a temperature independent uncertainty parameter *f*.

>> Please enter the assumed constant *f* value:

The required answer is a decimal number between 0.1 and 1.5 (like 0.7).

The program then calculates the standard deviation of $\log_{10} k$, like

>> Standard deviation of $\log_{10} k = 0.233333$

The information about the uncertainty of the rate coefficient is saved in file "CovMatrix.txt". Although the saved matrix is not the covariance matrix of the transformed Arrhenius parameters, but the uncertainty domain and the probability distribution of the Arrhenius parameters can be reconstructed from it [1]. The program continues with the creation of the final plot (see Section 7.).

Choosing N:
Truncated multivariate normal distribution is assumed for the Arrhenius parameters and truncated normal distribution is assumed for the rate coefficient at each temperature (see [1]).

The user may define if all Arrhenius parameters (*A*, *n*, *E*) are assumed to be random variables or only two of them (*A*, *E*) or (*A*, *n*):

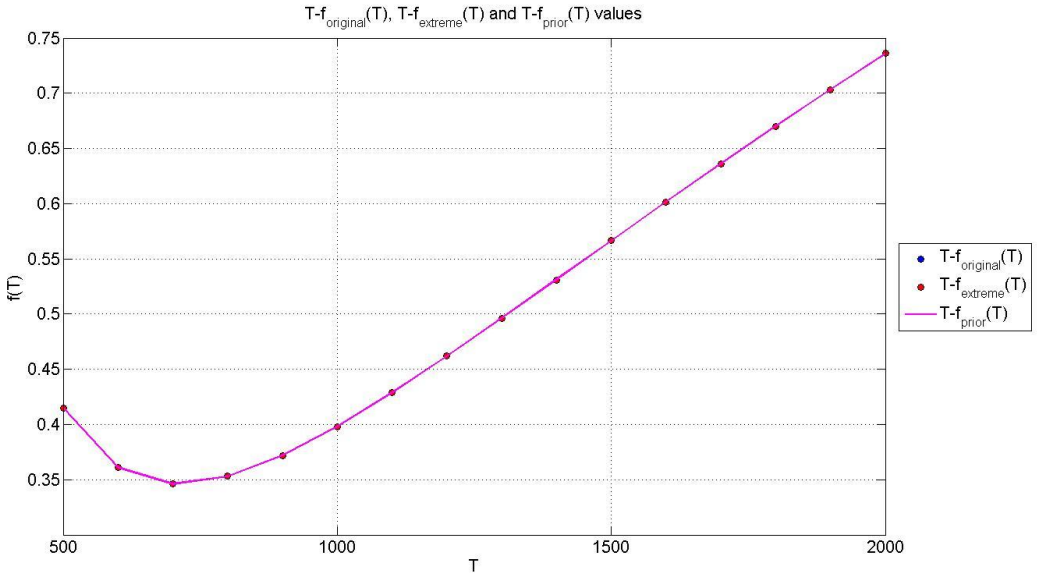
>> Uncertainty of Arrhenius parameters to be determined (AnE, AE, An):

Possible answers: 'AnE', 'AE' or 'An'
If the answer does not match any of the possible answers, then the question is repeated.

The program then calls code UBAC for the determination of the uncertainty band of Arrhenius curves, that is finding the T–f points that are in accordance with the type of the Arrhenius expression [1]. These corrected f values are called f_{extreme} and the program visualizes the T-f_{extreme} points.

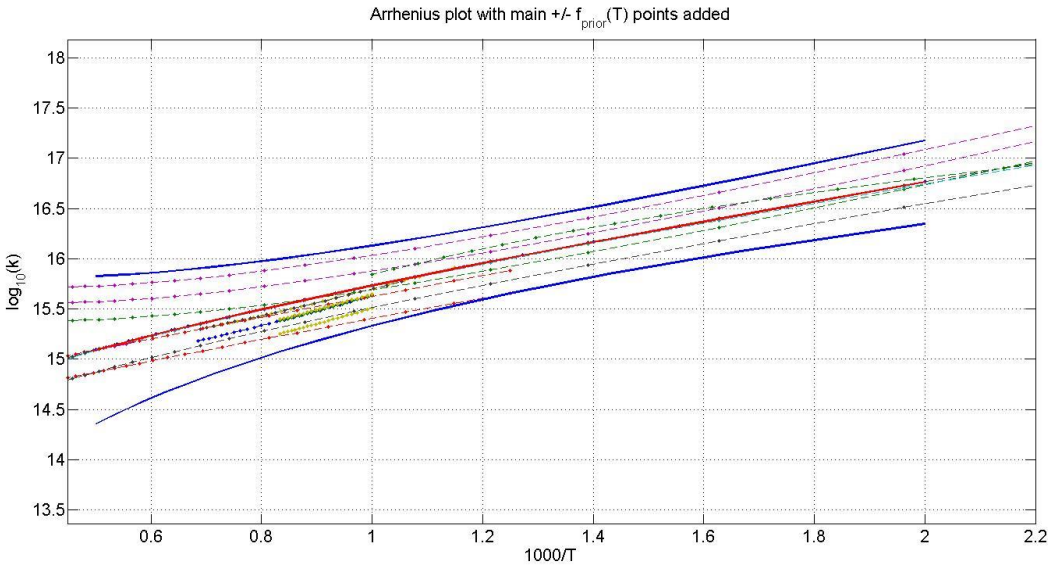
Next, the program calls code JPDAP (acronym for joint probability density of Arrhenius parameters) for processing the T-f_{extreme} points. The results of JPDAP are the parameters of the covariance matrix, which are the variances of the transformed Arrhenius parameters ($\sigma_\alpha^2, \sigma_n^2, \sigma_\varepsilon^2$) and their correlations ($r_{\alpha n}$, $r_{\alpha\varepsilon}$, $r_{\varepsilon n}$). The transformed Arrhenius parameters are α := ln{A}, n and $\varepsilon := E/R$. If only two Arrhenius parameters (*A*, *E*) or (*A*, *n*) are considered as random variables, the corresponding central moments (*e.g.* $\sigma_\alpha^2, \sigma_\varepsilon^2$, and $r_{\alpha\varepsilon}$) are obtained. The parameters of the covariance matrix are saved in the "CovMatrix.txt" file.

Finally, the program visualizes the T-f_{original} and T-f_{extreme} points, and T-f_{prior} function that has been calculated from the prior covariance matrix.



7. Final results

Finally, on an Arrhenius-type plot the program visualizes all used review, experimental, or theoretical Arrhenius expressions (thin coloured lines), the mean Arrhenius curve (thick red line), and the ($\log_{10}(k_{mean}(T)) \pm f_{prior}(T)$) uncertainty limits (thick blue lines). Again, the *figure* tool of Matlab allows zooming in to the interesting parts and reading the label of the data by clicking on the corresponding line. Assuming normal distribution, function $f_{prior}(T)$ is calculated from the prior covariance matrix of the transformed Arrhenius parameters, while assuming uniform distribution $f_{prior} = 3\ \sigma\ (\log_{10} k)$.



1. T. Nagy; É. Valkó; I. Sedyó; I. G. Zsély; M. J. Pilling; T. Turányi, Combustion and Flame (in press) (2015) http://dx.doi.org/10.1016/j.combustflame.2015.01.005.

2. T. Nagy; T. Turányi, Int. J. Chem. Kinet. 43 (7) (2011) 359-378 10.1002/kin.20551.

3. J. A. Manion; R. E. Huie; R. D. Levin; D. R. Burgess Jr.; V. L. Orkin; W. Tsang; W. S. McGivern; J. W. Hudgens; V. D. Knyazev; D. B. Atkinson; E. Chai; A. M. Tereza; C.-Y. Lin; T. C. Allison; W. G. Mallard; F. Westley; J. T. Herron; R. F. Hampson; D. H. Frizzell NIST Chemical Kinetics Database, NIST Standard Reference Database 17, Version 7.0 (Web Version), Release 1.6.6, Data Version 2013.03, National Institute of Standards and Technology, Gaithersburg, Maryland, 20899-8320. . http://kinetics.nist.gov/

4. T. Nagy. Manual of program JPDAP.

5. T. Nagy. Manual of program UBAC.