Uncertainty of reaction kinetics model parameters and the computer code SimLab

András György Szanthoffer
(Slides of Éva Valkó and Tamás Turányi)

2025

1

OUTLINE

1. Introduction
2. Uncertainty of reaction kinetic model parameters
3. Sample generation
4. Correlation of variables
5. Sample generation and statistical analysis with SIMLAB (homework)

2

1. Introduction

- Reaction kinetic models are mathematical models containing many parameters
(kinetic, thermochemical, sometimes transport)
- In every model, the parameters have uncertainty
  - Calculation of the nominal values is not sufficient; the uncertainty belonging to these nominal values has to be defined
  - Correlations between model parameters are important to characterize
  - The effect of this uncertainty on the uncertainty of the model results has to be investigated and quantitatively characterized (uncertainty analysis)

3

3

2. Uncertainty of chemical kinetic model parameters

4

4

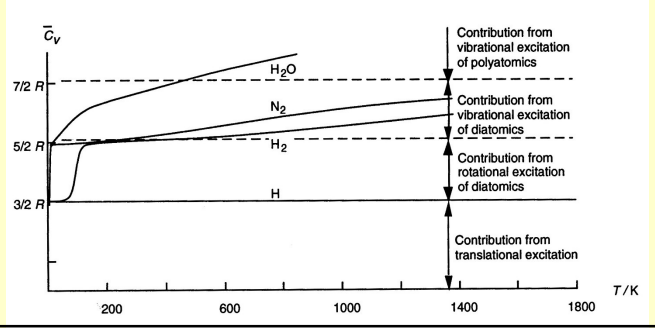
Temperature dependence of thermodynamic data

NASA polynomials

$$\frac{H^\theta}{RT} = a_1 + \frac{a_2}{2}T + \frac{a_3}{3}T^2 + \frac{a_4}{4}T^3 + \frac{a_5}{5}T^4 + \frac{a_6}{T}$$

$$\frac{c_p}{R} = a_1 + a_2T + a_3T^2 + a_4T^3 + a_5T^4$$

$$\frac{S^\theta}{R} = a_1 \ln T + a_2T + \frac{a_3}{2}T^2 + \frac{a_4}{3}T^3 + \frac{a_5}{4}T^4 + a_7$$



5

5

Using thermodynamic data in combustion simulations

ΔH_f $\Rightarrow$ calculation of heat production in a reacting mixture
$\rightarrow$ calculation of temperature changes
$\Rightarrow$ calculation of $\Delta_r G^0$

c_p $\Rightarrow$ calculation of temperature changes

ΔS $\Rightarrow$ calculation of $\Delta G = \Delta H - T\Delta S$
$\rightarrow$ calculation of the equilibrium constant
$\rightarrow$ calculation of the rate coefficient of reverse reactions

6

6

Uncertainty of thermodynamic data

thermodynamic data influence the reaction kinetic calculations in two ways:
- Calculated temperature
- Calculation of the rate coefficients of backward reaction steps

Thermodynamic data used:
- heat capacity (can be calculated using statistical thermodynamics)
- entropy (can be calculated using statistical thermodynamics)
- standard enthalpy of formation (measurement or high level calculation)

7

7

Uncertainty of thermodynamic data

c_p and ΔS can be calculated from the IR spectrum
using methods of statistical thermodynamics

ΔH_f - can be computed
(for small molecules only; not easy)
- can be determined experimentally by
- measuring the equilibrium constant of a reaction
$\rightarrow$ reaction enthalpy $\rightarrow$ enthalpy of formation
- measuring ionization energy by mass spectrometry

8

8

Uncertainty of thermodynamic data

Typical uncertainty of ΔH_f (1σ):

molecules and small radicals: 0.1-0.5 kJ/mole
e.g. CO= 0.17 kJ/mole, CH_4= 0.4 kJ/mole, CH_3=0.4 kJ/mole

large radicals: 1.0 – 5.0 kJ/mole
e.g. HO_2= 3.35 kJ/mole, CH_2OH= 4.2 kJ/mole

less known radicals: 8-10 kJ/mole
e.g. HCCO= 8.8 kJ/mole, CH_2HCO= 9.2 kJ/mole

9

9

Rate coefficient uncertainties

Uncertainty factor *f* as defined in data evaluations
(Tsang, Warnatz, Baulch, Konnov):

uncertainty factor u_j uncertainty parameter f_j

$$u = \frac{k^0}{k^{\min}} = \frac{k^{\max}}{k^0} \qquad f = \log_{10}(u)$$

k^0 recommended value of the rate coefficient of reaction *j*

$k^{\min}$ possible minimal value of *k*

$k^{\max}$ possible maximal value of *k*

$\Rightarrow$ $[k^{\min}, k^{\max}]$ is the physically realistic range for the rate coefficients

assume that ln $k^{\min}$ and ln $k^{\max}$ deviate 3σ from ln k^0

$$\Rightarrow \quad \sigma^2(\ln k) = ((f \ln 10)/3)^2$$

10

10

Uncertainty of *k* at a given temperature

Uncertainty of rate coefficient measurements:

very high quality data	uncertainty factor u= 1.26 ⇔ f=0.1 ⇔ ± 8 % (1σ)
typical good data	uncertainty factor u= 2.00 ⇔ f=0.3 ⇔ ±26 % (1σ)
typical data	uncertainty factor u= 3.16 ⇔ f=0.5 ⇔ ±47 % (1σ)

Uncertainty of (high level) theoretical determinations:
TST/master equation calculations

best systems	uncertainty factor u= 2.00 ⇔ f=0.3 ⇔ ±26 % (1σ)
multi well, main channels	uncertainty factor u= 3.16 ⇔ f=0.5 ⇔ ±47 % (1σ)
multi well, minor channels	uncertainty factor u=10 ⇔ f=1.0

C. F. Goldsmith, A. S. Tomlin, S. J. Klippenstein: Uncertainty propagation in the derivation of phenomenological rate coefficients from theory: A case study of *n*-propyl radical oxidation *Proc. Combust. Inst.*, **34**, 177-185 (2013)

J. Prager, H. N. Najm, J. Zádor: Uncertainty quantification in the *ab initio* rate-coefficient calculation for the $CH_3CH(OH)CH_3$+OH → $CH_3C.(OH)CH_3$+H_2O reaction, *Proc. Combust. Inst.*, **34**, 583-590 (2013)

11

11

3. Sample generation

Why is sample generation important?

In many sensitivity and uncertainty analysis methods (discussed in the next lecture), we need to sample the values of the model parameters (sometimes called variables or factors) within the domain of their uncertainty

→ To get reliable results, we need to sample the space of the parameters as uniformly as possible

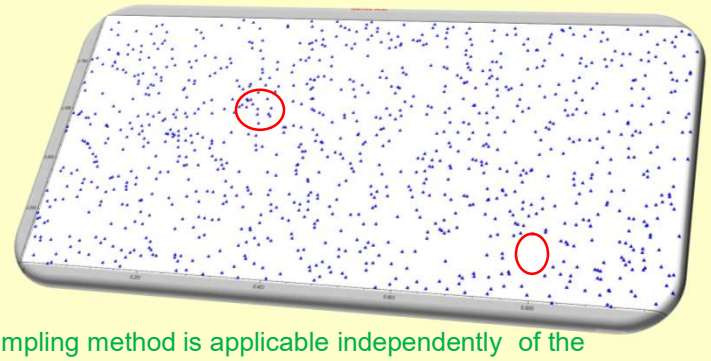
12

12

3. Sample generation

I. Random sampling
(pseudo-random)
unbiased estimation to the expected value and the standard deviation

Problem:
„white places and clustering”



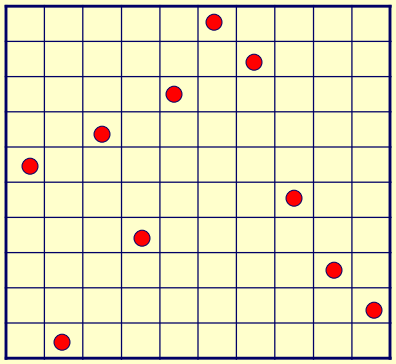
The random sampling method is applicable independently of the distributions of the parameters

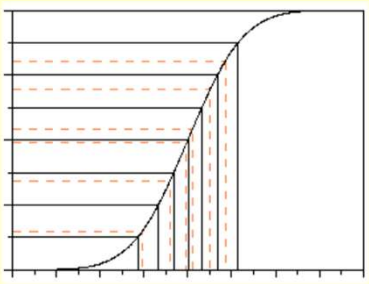
13

13

3. Sample generation

II. Latin Hypercube sampling





In the Latin Hypercube the range of each input factor, X_j, j=1,2…k, is divided into N intervals of equal marginal probability, 1/N, and one observation of each input factor is made in each interval using random sampling within that interval

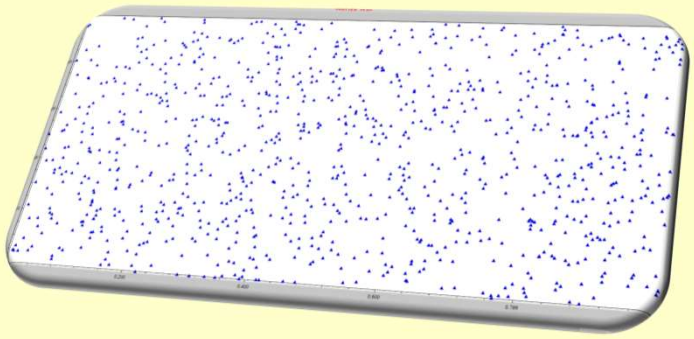
14

14

3. Sample generation

II. Latin Hypercube sampling

More uniform distribution in the sample space



The Latin Hypercube sampling is applicable independently of the distribution of the parameters

15

15

4. Correlation

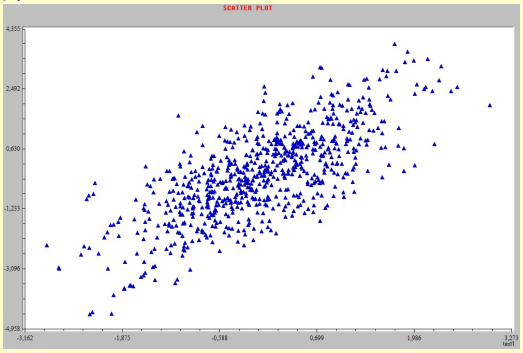
I. Visualization: Scatter plots

One of the easiest methods

Denote x_i (i=1..n) the parameters of the model and y denotes the result of the model

Create a figure belongs to (x_i,y)

Shows the relation between the parameters and the model result (linear, nonlinear, monotonic, non-monotonic), and the strength of relation
Helps to understand the behavior of the model



Disadvantages: Lots of figures have to be created and investigated (in typical chemical models, we may have several tens of thousands of model parameters and predictions)

16

16

4. Correlation

II. Pearson product moment correlation coefficient (PEAR)

Simple method

Denote x_i (i=1..n) the parameters of the model and y denotes the result of the model

Calculate the correlation between x_i and y

$$PEAR(x_i, y) = \frac{\mathrm{cov}(x_i, y)}{\sigma_x \sigma_y} = \frac{E[(x_i - \mu_{x_i})(y - \mu_y)]}{\sigma_x \sigma_y}$$

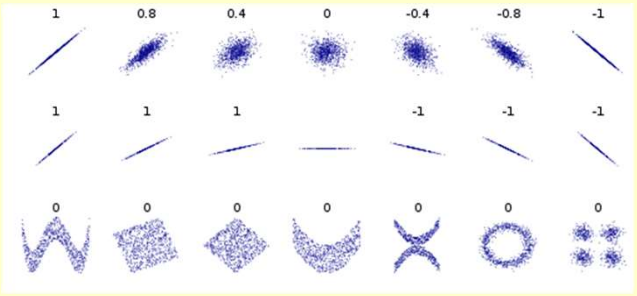
The correlation coefficient is a measure of the linear dependence between two variables x_i and y, giving a value between +1 and −1 inclusive, where 1 is total positive linear correlation, 0 is no linear correlation, and −1 is total negative linear correlation

17

17

4. Correlation

II. Pearson product moment correlation coefficient (PEAR)



The correlation coefficient is a measure of the linear dependence between two variables x_i and y, giving a value between +1 and −1 inclusive, where 1 is total positive linear correlation, 0 is no linear correlation, and −1 is total negative linear correlation

18

18

Homework

- Sample generation and statistical analysis using the program SIMLAB
  (can be downloaded from http://garfield.chem.elte.hu/Turanyi/oktatas/Modelling_in_Chemistry_MSc.html)

- Deadline: 00:00 on November 19, 2025
- Upload the report to Moodle

19

19

Homework 1

Description at:
http://garfield.chem.elte.hu/Turanyi/oktatas/Modelling_in_Chemistry_MSc.html

Name	Neptun ID	Variables (factors)			Correlation
		X	Y	Z	
Balázs Krisztina	JJS8IN	N(0,2)	U(−3,2)	N(2,4)	Corr(X,Z) = 0.5
Burlacu Péter Dániel	UV0XM9	U(0,4)	N(1,1)	N(−1,2)	Corr(Y,Z) = −0.2
Fábián Lizett	JMHILE	N(0,3)	N(0,3)	U(−2,2)	Corr(X,Y) = 0.6
Galanics Kitti Csenge	CE6COC	N(1,4)	N(−1,1)	U(0,2)	Corr(X,Y) = −0.6
Gargya Noémi Éva	JJDXN7	U(0,3)	N(−1,2)	N(0,3)	Corr(Y,Z) = 0.7
Nagy Eszter	RU4P0X	N(−2,2)	U(0,4)	N(0,3)	Corr(X,Z) = −0.7
Palik Dezső István	BGG35F	U(0,3)	N(0,3)	N(−1,2)	Corr(Y,Z) = −0.3
Rácsai Balázs	WM82LC	U(0,4)	N(2,4)	N(2,3)	Corr(Y,Z) = 0.8
Radócz Roland	G73MSE	N(−2,1)	N(1,2)	U(−3,0)	Corr(X,Y) = 0.9
Riznychenko Tetiana	YOJWT5	N(0,3)	N(−1,2)	U(0,3)	Corr(X,Y) = −0.5
Sajósi Benedek	PRVEB8	N(0,2)	U(−1,4)	N(1,4)	Corr(X,Z) = −0.8
Terbák Enikő Krisztina	D4GW6X	U(−1,3)	N(2,4)	N(−3,1)	Corr(Y,Z) = 0.95
Voigt Inga Maria	IML2PL	N(−1,1)	U(−5,−2)	N(3,3)	Corr(X,Y) = −0.95

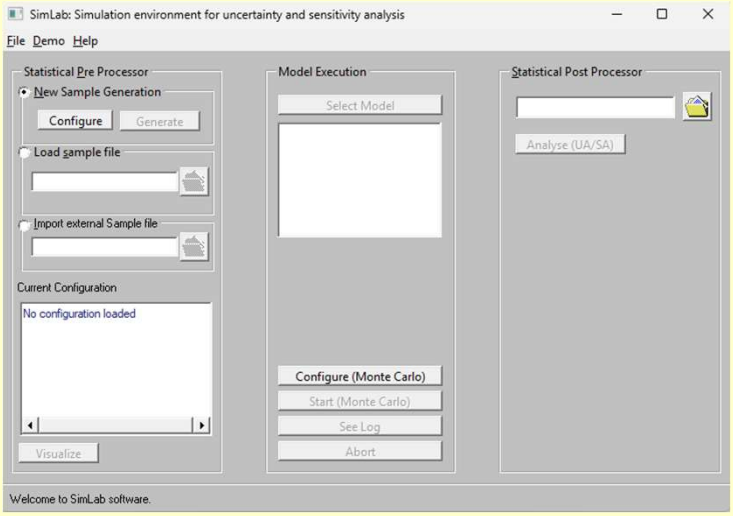
N: normal distribution, $N(\mu, \sigma)$

U: uniform distribution, $U(a, b)$

20

20

Homework 1



21

21

Homework 2

Description at:
http://garfield.chem.elte.hu/Turanyi/oktatas/Modelling_in_Chemistry_MSc.html

22

22

Homework grading

Submit a single PDF document containing your reports for both tasks.

- The two parts will be graded separately on a scale of 0–100%.
- The overall score will be the average of the two scores.
- A 10% deduction is applied from the overall score for every day of delay in the submission (deadline: 00:00 on November 19, 2025).

23

23

Thank you for your attention!

24

24