Mechanism optimization with Optima++

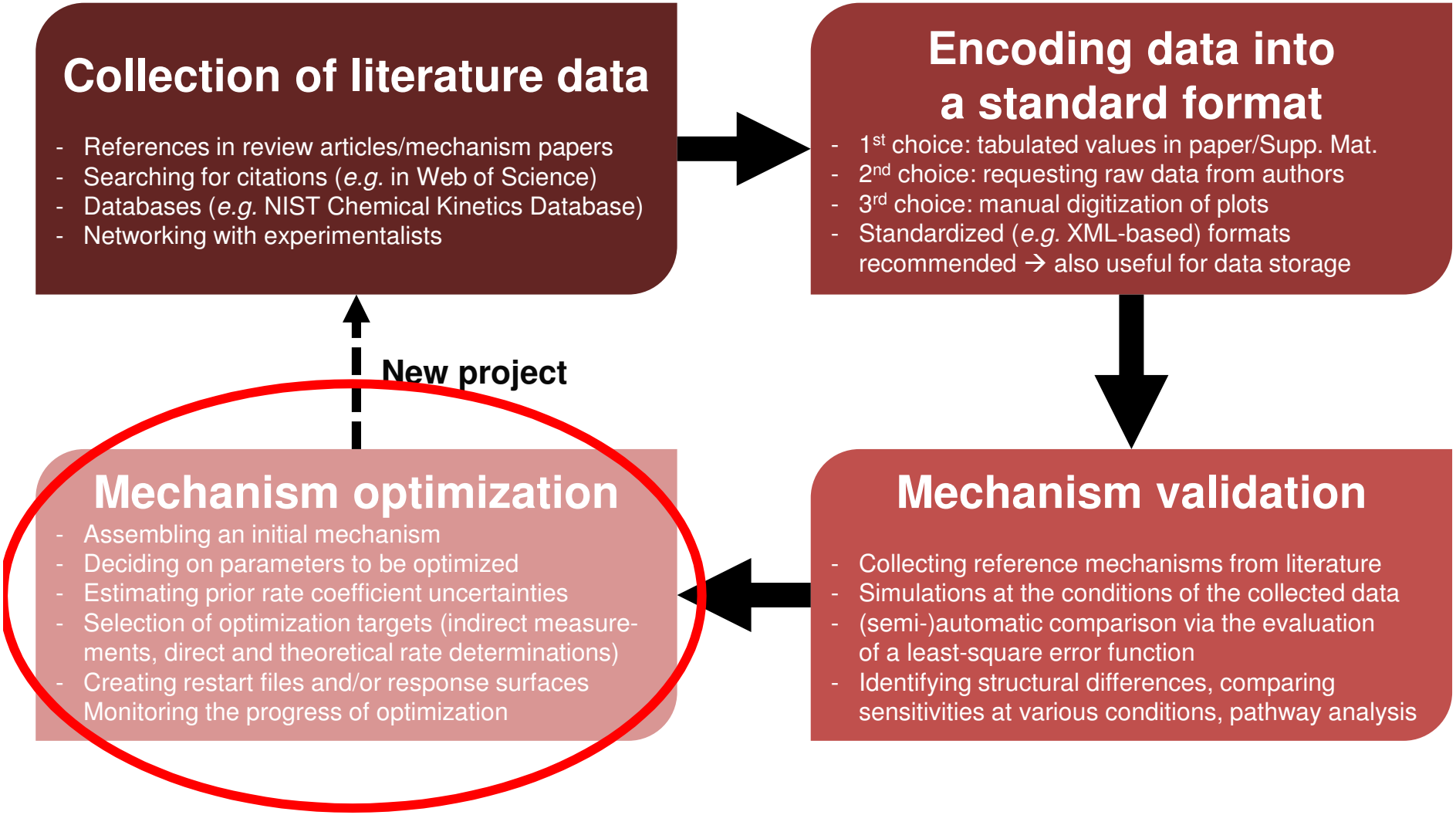
Varga Tamás

Institute of Chemistry, Eötvös University (ELTE), Budapest, Hungary
MTA-ELTE Research Group on Complex Chemical Systems, Budapest, Hungary

Chemical Kinetics Laboratory|ELTE

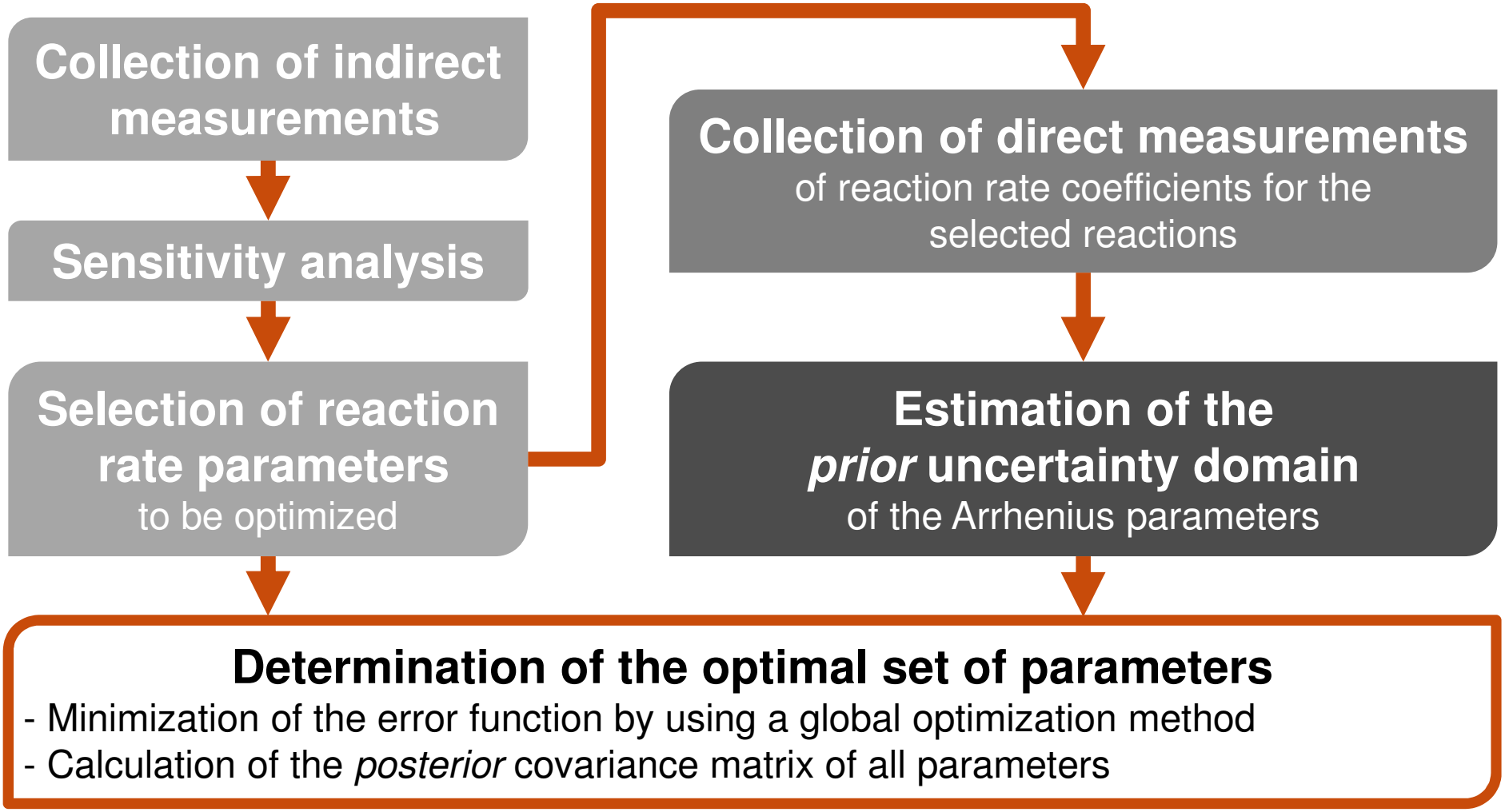
COST Training School on the Analysis of Combustion Mechanisms
July 4 – 7, 2016 | Budapest, Hungary | DAY 4, Practical Session 4-2

General workflow



Optimization methodology of Turányi et al.

T. Turányi et al., *Int. J. Chem. Kinet.* **44** (2012) 284–302



Better rate parameters + better estimation of their accuracy!

Sensitivity analysis

- Useful tool for identifying reactions for which rate parameters can be determined from indirect data
- Local sensitivity analysis

Simulated value

$$S_n = \frac{\partial \ln Y}{\partial \ln p} \approx \frac{p}{Y} \frac{Y' - Y}{p' - p}$$

Model parameter – Typically A Arrhenius parameters

- In Optima++ normalized sensitivity coefficients are calculated using a brute force method

Sensitivity analysis

Input syntax

- Settings
  - MECHANISM – mechanism from a MECHMOD block
  - SOLVER – default “FM”
  - SETTINGS_TAG – default “default”
  - THREAD_LIMIT – default 1
  - SPLIT – default 1
- Specifying target conditions
  - NAME myDataFile1.xml POINTS 1 2 3 4 5
  - NAME myDataFile2.xml POINTS all

Sensitivity analysis

output

- Header containing basic information about the run

- Result lines containing
  - Name of datafile
  - Point number of the experiment
  - Name of measured species (if applicable)
  - Measured results (see manual for units)
  - Simulated results (see manual for units)
  - Normalized sensitivity coefficients

Optimization

- By optimization we mean a fitting of (usually kinetic) parameters of a model, to minimize an error function
- Optimization of chemical kinetic models can serve two, closely related purposes
  - Improvement of reproduction of experiments (at certain conditions)
  - Determination of parameters from complex experiments
- Our methods can be used for both purposes, but the primary advantage is the ability to get realistic rate coefficients with realistic uncertainties from fitting to a large amount of data

Optimization

Error function

$$E = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{N_i}\sum_{j=1}^{N_i}\left(\frac{Y_{ij}^{mod} - Y_{ij}^{exp}}{\sigma(Y_{ij}^{exp})}\right)^2$$

Difference of modeled and experimental value
(characterizes the prediction of one measured value)

Number of data series

Number of points
(division makes data series differing in size equally weighted)

Estimated standard deviation / scatter
(makes different types of experiments comparable, accounts for different reliability of data)

$$Y_{ij} = \begin{cases} y_{ij} & \text{if } \sigma(y_{ij}^{exp}) \approx \text{constant} \\ \ln y_{ij} & \text{if } \sigma(\ln y_{ij}^{exp}) \approx \text{constant} \end{cases}$$

Transformation
(comparison of experiments with absolute and relative errors)

The **overall agreement between simulations and measurements** can be well characterized quantitatively by this error function

Optimization

Transformation

Transformation
(comparison of experiments
with absolute and relative errors)

$$Y_{ij} = \begin{cases} y_{ij} & \text{if } \sigma(y_{ij}^{exp}) \approx \text{constant} \\ \ln y_{ij} & \text{if } \sigma(\ln y_{ij}^{exp}) \approx \text{constant} \end{cases}$$

- Selection of the error scale is always up to the user
- What we recommend is
  - Flame speeds – abs
  - Concentration measuerments – abs
  - Ignition delays – abs ln
  - Rate coefficients – abs ln

Optimization

Estimation of uncertainties

- Calculation of the covariance matrix of the estimated parameters based on linear error propagation:

Covariance matrix of experiments

Discrepancy between experiments and model

$$\boldsymbol{\Sigma}_{\mathbf{p}} = \left[\left(\mathbf{J}_{\mathrm{o}}^{\mathrm{T}}\mathbf{W}\boldsymbol{\Sigma}_{\mathbf{Y}}^{-1}\mathbf{J}_{\mathrm{o}}\right)^{-1}\mathbf{J}_{\mathrm{o}}^{\mathrm{T}}\mathbf{W}\boldsymbol{\Sigma}_{\mathbf{Y}}^{-1}\right]\left(\boldsymbol{\Sigma}_{\mathbf{Y}} + \boldsymbol{\Sigma}_{\Delta}\right)\left[\left(\mathbf{J}_{\mathrm{o}}^{\mathrm{T}}\mathbf{W}\boldsymbol{\Sigma}_{\mathbf{Y}}^{-1}\mathbf{J}_{\mathrm{o}}\right)^{-1}\mathbf{J}_{\mathrm{o}}^{\mathrm{T}}\mathbf{W}\boldsymbol{\Sigma}_{\mathbf{Y}}^{-1}\right]^{\mathrm{T}}$$

- Standard deviations of rate coefficients can be calculated from the covariance matrix of its Arrhenius parameters

$$\underbrace{\ln\{k(T)\}}_{\kappa(\theta)} = \underbrace{\ln\{A\}}_{\alpha} + \underbrace{n}_{n}\cdot\ln\underbrace{\{T\}}_{\theta} - \underbrace{\{E/R\}}_{\varepsilon}\cdot\underbrace{\{T\}}_{\theta}^{-1}$$

$$\boldsymbol{\theta}^{\mathrm{T}} = \begin{bmatrix} 1 & \ln\theta & -\theta^{-1} \end{bmatrix}$$

$$\sigma_{\kappa}(\theta) = \sqrt{\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{\Sigma}_{\mathbf{p}}\boldsymbol{\theta}}$$

Optimization

Algorithm

- Generate random sample of rate parameters
- Evaluate error function *E* at each parameter set
- Select parameter set that gave the lowest *E* values as the present optimal set
- Calculate new covariance matrix of the parameters (optional)
- IF *E* could be not decreased then reduce the scaling of the sampling range
- IF *E* could be decreases then increase the scaling of the sampling range

Optimization

Parameter sampling

- Parameters can be sampled uniformly or normally
- Uncertainty range defined with a mean, covariance matrix, and cutoff in sigma (typically 3)
- Sampling range is defined with (a potentially different) mean and covariance matrix
- The sampling covariance matrix can be scaled according to a "focus level" to help focus on parameter regions where the minimum is expected

$$f_\sigma = N_{\text{sample}}^{-2 \cdot N_{\text{focus}} / N_{\text{parameters}}} \qquad \log f_\sigma = -\frac{2 \cdot N_{\text{focus}}}{N_{\text{parameters}}} \log N_{\text{sample}}$$

- Arrhenius parameters of the same reaction are sampled together
- The cutoff is applied in rate coefficients in a finite temperature range

Optimization

Input syntax

- Same settings as SENSITIVTY, but need to specify parameters for optimization and uncertainties for target experiments
- Parameters
  - Select reaction
  - Select parameters – A, n, E and third body collision effs.
  - Define initial sampling range
  - Define uncertainty bounds for optimization (hard limits)
- Data uncertainties
  - SIGM – standard deviations
  - SIGMSCALE – scale of uncertainties (abs/absln)
  - SIGMUNIT – unit of uncertainties for abs scale

Optimization

Reaction blocks

Start block with REACTION, and close with END

- REACNUM – number of the reaction
- REACSTRING – reaction string, cross-checked with reacnum
- UNC_MEAN and UNC_COVMAT – defines uncertainty range
- INIT and INIT_COVMAT – defines initial sampling range
- TEMPRANGE – defines T limits for uncertainty range
- SIGMARANGE - uncertainty range with in $\sigma(\ln k)$-s

Optimization

Results

- Error function over iterations
- Error functions for each dataset
- Parameter sets
- Optimal parameter set and covariance matrix at each iteration