MCR for process analysis.
Hybrid hard- and soft-modeling. Applications.

Chemometrics workshop (Zanjan, 2011)
Anna de Juan, anna.dejuan@ub.edu
Hybrid MCR. Possible data scenarios.

By component

- **All compounds** or **some of them** obey the fitted model.
  - Off-process contributions to the signal can be present and be left unconstrained (**,+**).

Hybrid MCR. Possible data scenarios.

- **Single global model**
- **Global models for subsets of experiments** (*)
- **Individual models** (+)
- **Model-based and model-free experiments** (x)

Link among models is not necessary or can be unknown

(+ ) A. de Juan, M. Maeder, M. Martínez and R. Tauler. ACA, 442 (2001) 337.
Applications

- **Modeling the signal (S matrix)**
  - ✓ NMR signals.
  - ✓ Voltammetric signals.

- **Modeling concentration profiles**
  - ✓ Kinetic laws.
  - ✓ Equilibrium laws
  - ✓ Enzymatic models.
  - ✓ Convoluted kinetic models in time-resolved spectroscopy.
  - ✓ Adsorption models.
Hybrid hard- and soft-modeling. Possibilities.

- **Process with off-process contributions.**
  - ✔ Chemical contributions (interferences).
  - ✔ Non-chemical contributions (artifacts)
- **Process with non-absorbing contributions.**
  - ✔ Non-absorbing contributions.
  - ✔ Difference spectroscopy.
- **Multiset structures lacking a global model.**
  - ✔ With different unlinked models per experiment
  - ✔ With model-based and model-free experiments.
Hybrid hard- and soft-modeling. Examples

Processes with off-process contributions
- Quantitative analysis of pH-modulated samples
- Solvent effect on the photodegradation of decabromodiphenyl ether.
- Photochemical reaction in a protein reaction center.

Multisets with model-based and model-free experiments
- pH- and time-dependent denaturation of myoglobin
- Ketoprofen photodegradation
Quantitative analysis of pH-modulated samples

- Presence of interfering compounds.
- Different protonation models in aqueous solution and in sample.
- Model parameters and quantitative information.

Quantitative analysis of pH-modulated samples

Measurements

- FTIR derivative spectra coming from samples subject to pH modulation (titration). Analyte + interference.
  - Analyte: malic acid, tartaric acid (diprotic acids).
- FTIR derivative spectra coming from standards subject to pH modulation. Analyte solutions.

Goal

- Quantification of analyte in the presence of interferences.
Quantitative analysis of pH-modulated samples

Data structure

Protonation model

\[
[H_2A] = \frac{C_a[H_3O^+]^2}{[H_3O^+]^2 + K_{a1}[H_3O^+] + K_{a1}K_{a2}}
\]

\[
[HA^-] = \frac{C_aK_{a1}[H_3O^+]}{[H_3O^+]^2 + K_{a1}[H_3O^+] + K_{a1}K_{a2}}
\]

\[
[A^{2-}] = \frac{C_aK_{a1}K_{a2}}{[H_3O^+]^2 + K_{a1}[H_3O^+] + K_{a1}K_{a2}}
\]

Standard model

Fitted parameters: \(K_{a1}, K_{a2}\)

\(C_a\) known (standard concentration)
Quantitative analysis of pH-modulated samples

Data structure

Mixture model

Fitted parameters: $K_{a1}, K_{a2}, C_a$

$C_a$ (sought analyte concentration)

Protonation model

\[
[H_2A] = \frac{C_a [H_3O^+]^2}{[H_3O^+]^2 + K_{a1} [H_3O^+] + K_{a1} K_{a2}}
\]

\[
[HA^-] = \frac{C_a K_{a1} [H_3O^+]}{[H_3O^+]^2 + K_{a1} [H_3O^+] + K_{a1} K_{a2}}
\]

\[
[A^{2-}] = \frac{C_a K_{a1} K_{a2}}{[H_3O^+]^2 + K_{a1} [H_3O^+] + K_{a1} K_{a2}}
\]
Quantitative analysis of pH-modulated samples

Data structure

Interference: soft-modelled
Quantitative analysis of pH-modulated samples

Results

Analyte: malic acid
Interference: diprotic acid.

Results

pK (mixture sample)

\[ pK_{a1} = 3.45 \quad pK_{a2} = 4.83 \]

\[ pK_{1\text{theor}} = 3.40 \quad pK_{2\text{theor}} = 5.11 \]

\[ C_{a,\text{mixture}} = 2.92 \text{ g/L} \]

\[ C_{a,\text{true}} = 3.0 \text{ g/L} \]

% lack of fit = 1.48
Photodegradation of decabromodiphenyl ether
Effect of solvent polarity in photodegradation

- Different kinetic models as a function of solvent composition.
- Interfering contribution (accounting for the solvent effect).

Photodegradation of decabromodiphenyl ether

UV spectroscopic monitoring

BDE-209 (flame retardant)

Solvent: THF/ water mixtures:
- 10% water
- 20% water
- 30% water
- 40% water

Several replicates per solvent composition
Photodegradation of BDE-209

Data structure

\[
\begin{align*}
D_{\text{aug}} & = C_{\text{aug}} \\
 & = \begin{bmatrix} D_{11} & \cdots & D_{31} & \cdots & D_{3n} \\
D_{21} & \cdots & D_{21} & \cdots & D_{2n} \\
\vdots & \ddots & \vdots & \ddots & \vdots \\
D_{31} & \cdots & D_{31} & \cdots & D_{3n} \\
\vdots & \ddots & \vdots & \ddots & \vdots \\
D_{3n} & \cdots & D_{3n} & \cdots & D_{3n} 
\end{bmatrix} \\
C_{\text{aug}} & = \begin{bmatrix} C_{11} & \cdots & C_{21} & \cdots & C_{31} \\
C_{1n} & \cdots & C_{2n} & \cdots & C_{3n} \\
\vdots & \ddots & \vdots & \ddots & \vdots \\
C_{11} & \cdots & C_{21} & \cdots & C_{31} \\
C_{2n} & \cdots & C_{2n} & \cdots & C_{3n} \\
\vdots & \ddots & \vdots & \ddots & \vdots \\
C_{3n} & \cdots & C_{3n} & \cdots & C_{3n} 
\end{bmatrix}
\end{align*}
\]

Global model 1 (10% water)

\[
A \xrightarrow{k_1} B \xrightarrow{k_2} C \xrightarrow{k_3} D
\]

Global model 2 (20% water)

Off-process contribution (spectral solvent effects)
Photodegradation of BDE-209

CONCENTRATION PROFILES

10% water  20% water  30% water  40% water

SPECTRA

% lack of fit = 4.92

A $\xrightarrow{k_1} B \quad B \xrightarrow{k_2} C \xrightarrow{k_3} D$

BDE-209 (1) (2) (3) (4)
## Photodegradation of BDE-209

### Rate constants

<table>
<thead>
<tr>
<th></th>
<th>10 % water</th>
<th>20 % water</th>
<th>30 % water</th>
<th>40 % water</th>
</tr>
</thead>
<tbody>
<tr>
<td>$-\log k_1$</td>
<td>3.5 (1)*</td>
<td>3.5 (2)</td>
<td>3.6 (1)</td>
<td>3.7 (1)</td>
</tr>
<tr>
<td>$-\log k_2$</td>
<td>3.5 (1)</td>
<td>3.7 (1)</td>
<td>4.0 (2)</td>
<td>3.9 (1)</td>
</tr>
<tr>
<td>$-\log k_3$</td>
<td>3.7 (1)</td>
<td>3.8 (2)</td>
<td>4.1 (2)</td>
<td>4.1 (1)</td>
</tr>
</tbody>
</table>

↑ water proportion  →  ↓ $k$
Photoinduced biochemical process in a protein reaction center

- Non-absorbing species in model (difference spectroscopy).
- Interfering contribution (process at a different time scale).
- Different models for different light intensities in experiments.

Protein photochemical reaction

Photochemical kinetic process

\[ \text{Ubiquinone}(Q_1) \xrightarrow{h\nu} \text{ubiquinol}(Q_2) \]

Protein conformational change

\[ \text{Initial conform.}(P_1) \xrightarrow{h\nu} \text{final conform.}(P_2) \]

Photosynthetic reaction center
Rhodobacter Spheroides

Measurement: IR rapid-scan spectroscopy (difference spectra) (1200-1800 cm\(^{-1}\))
Protein photochemical reaction

Kinetics of ubiquinol are modelled in the presence of an interference (protein absorption).

Hard-modeling (ubiquinol formation and decay contribution)
Soft-modeling constraints
HS-MCR and difference spectroscopy

\[ d_o = c_{Q1,o} s_{Q1} + c_{Q2,o} s_{Q2} \]
\[ d_i = c_{Q1,i} s_{Q1} + c_{Q2,i} s_{Q2} \]
\[ c_T = c_{Q1} + c_{Q2} \]

Spectrum \( t = 0 \)
Spectrum \( t = i \)
Mass balance

Light on

\[ Q_1 \xrightarrow{k_i} Q_2 \]
\[ P_1 \rightarrow P_2 \]

Light off

\[ Q_1 \xleftarrow{k_i} Q_2 \]
\[ P_1 \leftarrow P_2 \]

difference spectrum

\[ \Delta d_i = d_i - d_o \]
\[ \Delta d_i = (c_{Q1,i} s_{Q1} + c_{Q2,i} s_{Q2}) - (c_{Q1,o} s_{Q1} + c_{Q2,o} s_{Q2}) \]
\[ \Delta d_i = (c_T - c_{Q2,i}) s_{Q1} + c_{Q2,i} s_{Q2} - (c_T - c_{Q2,o}) s_{Q1} - c_{Q2,o} s_{Q2} \]
\[ \Delta d_i = (c_{Q2,i} - c_{Q2,o}) (s_{Q2} - s_{Q1}) \]

HS-MCR in difference spectroscopy.
Model set.

\[ d_0 = c_{Q1,o} s_{Q1} + c_{Q2,o} s_{Q2} \]
\[ d_i = c_{Q1,i} s_{Q1} + c_{Q2,i} s_{Q2} \]
\[ c_T = c_{Q1} + c_{Q2} \]

Spectrum \( t = 0 \)
Spectrum \( t = i \)
Mass balance

\[ \Delta d_i = d_i - d_0 \]

Difference spectrum

In matrix form,

\[ \Delta D = [c_{Q2} - c_{Q2,o}] [s_{Q2} - s_{Q1}] \]
HS-MCR and difference spectroscopy

Difference spectrum

\[ \Delta d_i = d_i - d_o \]

In matrix form,

\[ \Delta D = [c_{Q2} - c_{Q2,0}] [s_{Q2} - s_{Q1}] \]

Resolution of spectra are difference spectra

Conc. profiles of initial species do not appear

Nr. of absorbing contributions = Total nr. of species – nr. of species at time 0
HS-MCR and difference spectroscopy.

Model set

\[ \Delta D = [c_{Q2} - c_{Q2,0}] [s_{Q2} - s_{Q1}] \]

Initial species are set as non-absorbing species.
Protein photochemical reaction

- Kinetics of ubiquinol formation and decay are modelled (hard-modeling constraint).
  \[ k_1 = 7 \times 10^{-4} \text{ s}^{-1} \]
  \[ k_{-1} = 10^{-4} \text{ s}^{-1} \]

- Photoinduced protein conformational change (model-free) is modelled.
Extension to multiset. Different light intensities.

Protein (soft-modelled)

Two kinetic processes

High and medium light intensity: 1 and 2
Low light intensity: 1.
Extension to multiset. Different light intensities.

**TABLE 3:** Rate Constants Calculated on the Five Data Sets by the HS-MCR Resolution

<table>
<thead>
<tr>
<th>data sets</th>
<th>excitation</th>
<th>relaxation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$k_1 (s^{-1})$</td>
<td>$k_2 (s^{-1})$</td>
</tr>
<tr>
<td>$D_{hi}$</td>
<td>0.09 ± 0.02</td>
<td>1.06 ± 0.08</td>
</tr>
<tr>
<td>$D_{me}$</td>
<td>0.73 ± 0.05</td>
<td>0.25 ± 0.03</td>
</tr>
<tr>
<td>$D_{low1}$</td>
<td>0.39 ± 0.01</td>
<td>−</td>
</tr>
<tr>
<td>$D_{low2}$</td>
<td>0.31 ± 0.01</td>
<td>−</td>
</tr>
<tr>
<td>$D_{low3}$</td>
<td>0.22 ± 0.01</td>
<td>−</td>
</tr>
</tbody>
</table>
Myoglobin denaturation

- Model-free and model-based experiments
- Diverse experiments in data set

Myoglobin denaturation

Steady-state process
UV spectra, pH range 7.0-2.0
N → I_s ? → D

Unknown model

Kinetic process
UV spectra, pH-jump stopped-flow
N →^{k_1} I_t ?^{k_2} → D
First-order consecutive reactions
Myoglobin denaturation

Hard-modelling (kinetic unfolding, 1st order reactions)
Soft-modelling constraints

Model-free and model-based experiments can be analyzed together.
Myoglobin denaturation

Steady-state process
Native (N) → Denatured (D)

Kinetic transient ($I_t$)

Kinetic process

• Formation of a kinetic transient was detected and hard-modelled.
  $k_1 = 4.05 \text{ s}^{-1}$  $k_2 = 0.62 \text{ s}^{-1}$

• Steady-state unfolding was modelled with soft constraints.
Ketoprofen degradation

- Rank-deficiency problem in process.
- Model-based and model-free experiments.
- Diverse experiments in multiset.

Ketoprofen photodegradation

**Spectroscopic monitoring**
- Series of UV spectra collected along the photodegradation process every 3 seconds.

**Chromatographic monitoring**
- Series of LC-DAD chromatograms from aliquots collected at different process times.

Ketoprofen photodegradation

**Spectroscopic monitoring.**

- Variance explained > 99%
- $A \rightarrow B \rightarrow C$
- $k_1 = 0.0444 \text{ s}^{-1}$
- $k_2 = 0.00286 \text{ s}^{-1}$

The process can be more complex if some compounds evolve with identical (or correlated) kinetics.

---

Chromatographic monitoring
Data fusion
Data fusion approaches (multiset analysis)

**LC-DAD-MS.**
- Structural information. Identification of photoproducts.
- Postulation of kinetic mechanism.

**Spectroscopic/chromatographic monitoring (UV detection).**
- Process modelling (elucidation of reaction pathway and estimation of rate constants).
Ketoprofen photodegradation

LC-DAD-MS (I). Multiset structure and information

<table>
<thead>
<tr>
<th>Elution time (min)</th>
<th>Daug\textsubscript{DAD/MS}</th>
<th>Caug</th>
<th>Saug\textsuperscript{T}</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Daug\textsubscript{DAD}</td>
<td>C\textsubscript{1}</td>
<td>S\textsuperscript{T}\textsubscript{DAD}</td>
</tr>
<tr>
<td>( t = 0 ) s</td>
<td>( D_{1}\textsubscript{DAD} ) (Aliquot 1)</td>
<td></td>
<td>( \lambda ) (nm)</td>
</tr>
<tr>
<td>( t = 10 ) s</td>
<td>( D_{2}\textsubscript{DAD} ) (Aliquot 2)</td>
<td>C\textsubscript{2}</td>
<td>S\textsuperscript{T}\textsubscript{MS}</td>
</tr>
<tr>
<td></td>
<td>( D_{7}\textsubscript{DAD} ) (Aliquot 7)</td>
<td>C\textsubscript{7}</td>
<td>m/z</td>
</tr>
<tr>
<td></td>
<td>( D_{1}\textsubscript{MS} ) (Aliquot 1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>( D_{2}\textsubscript{MS} ) (Aliquot 2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>( D_{7}\textsubscript{MS} ) (Aliquot 7)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

NC = Number of components

Kinetic information (peak area vs. process time)

Structural information (photoproduct identification)
Ketoprofen photodegradation

LC-DAD-MS (II). Structural information ($S^T$ matrix)

Identification of photoproducts

Var. Expl. = 98%
Ketoprofen photodegradation

LC-DAD-MS (III). Kinetic information (C matrix).

Photoproduc kinase

Var. Expl. = 98%
Ketoprofen photodegradation

Postulation of kinetic model

\[ A \xrightarrow{k_1} B + C \]
\[ B + C \xrightarrow{k_2} C + D \]
\[ C + D \xrightarrow{k_3} E \]

Low definition of process time axis

Chromatographic/spectroscopic data fusion
Ketoprofen photodegradation

Spectroscopic/chromatographic multiset

All process compounds can be modelled (chromatographic data).

The process time axis is well defined (spectroscopic data).

Hard-modelling constraint is applied to the spectroscopic experiment.
Ketoprofen photodegradation

Spectroscopic/chromatographic multiset

Var. Expl. > 98%

$k_1 = 0.058 \pm 0.005 \text{ s}^{-1}$

$k_2 = 0.0104 \pm 0.0006 \text{ s}^{-1}$

$k_3 = 0.030 \pm 0.002 \text{ s}^{-1}$