

# Analyse af blandinger med MCR

Har man mange spektre til rådighed med varierende sammensætning af nogenlunde de samme analytter, kan man under visse forhold estimere de rene spektre og de dertil hørende relative koncentrationer. Det forbedrer fortolkningsmulighederne ift. PCA, hvor score- og loading-vektorer er ortogonale. Metoden hedder Multivariate Curve Resolution (MCR).

Af José Manuel Amigo, Søren Balling Engelsen, Rasmus Bro, Institut for Fødevarevidenskab, Det Biovidenskabelige Fakultet, Københavns Universitet og Lars Nørgaard, Foss Analytical

Spektroskopiske signaler fra biologiske prøver er i mange tilfælde multikomponent-systemer, hvor det er yderst komplekst at finde selektive signaler fra specifikke analytter. Opsplitning af overlappende signaler er derfor en af hjørnestenene for den analytiske kemi. Det er specielt relevant inden for visse analytiske discipliner som f.eks. kromatografi, hvor det er normalt at finde overlappende toppe, som må adskilles vha. data-analytiske metoder [1].

## Baggrund

Traditionelt baserede teknikker til at opsplitte signalet i de underliggende signaler antager, at spektrene følger en given statistisk fordelingsfunktion. F.eks. kan man forsøge at tilpasse Gauss- eller Lorentz-fordelinger til målte spektre eller elueringsprofiler [1]. Et af de største problemer ved at tilpasse signalet til en funktionel form er valget af en passende funktion (kromatografiske toppe følger sjældent nogen eksakt teoretisk model), det store antal parametre, der skal optimeres for hver enkelt funktion, og ikke mindst valget af antallet af underliggende signaler. Det er f.eks. meget problematisk at opløse signaler, der er gemt under et andet signal. Figur 1 (t.h.) viser et

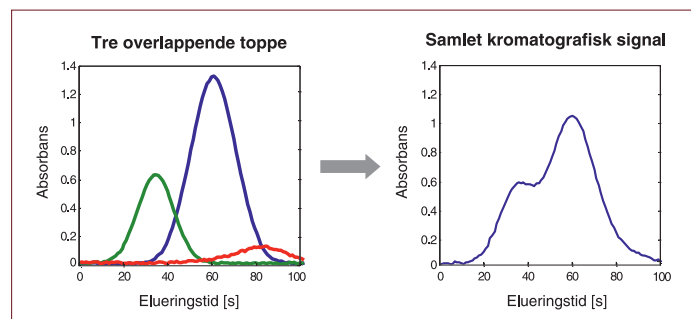
eksempel på sidstnævnte med en simulering af en kromatografisk profil, der indeholder tre gaussiske toppe (t.v.).

Fra et praktisk synspunkt, og uden *a priori* kendskab til antallet af komponenter, virker det fornuftigt at modellere signalet i figur 1 med to forskellige underliggende signaler. Modellering af signalet med to Gauss-funktioner kan imidlertid føre til en forkert konklusion, som vist i figur 2 (t.v.). De to rene Gauss-signaler (røde profiler) giver tilsammen en glimrende tilpasning (grønne prikker) til det komplekse kromatogram. Denne løsning er imidlertid langt fra virkeligheden, da den ikke indeholder oplysninger om den tredje komponent. Når kromatogrammet modelleres med tre gaussiske toppe, opnås der heller ikke en korrekt løsning (figur 2 t.h.).

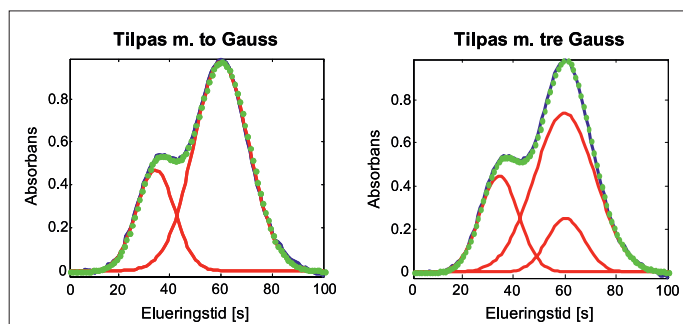
Har man i stedet en multikanal-detektor på sit kromatografiske udstyr, er det fornuftigt at drage fordel af de spektrale informationer for hver elueringsetid. Når man benytter en multikanal detektor, kan den kromatografiske profil visualiseres som et kromatografisk landskab, hvor der til hver elueringsetid er målt ét spektrum [2]. Der er vist et eksempel i figur 3, hvor de simulerede rene spektre for de tre toppe i figur 1 er afbildet i den øverste del af figuren. I landskabet er der lige så mange kromatografiske profiler, som der er bølgelængdekanaler (brune kurver i figur 3).

## Den kemometriske tilgang

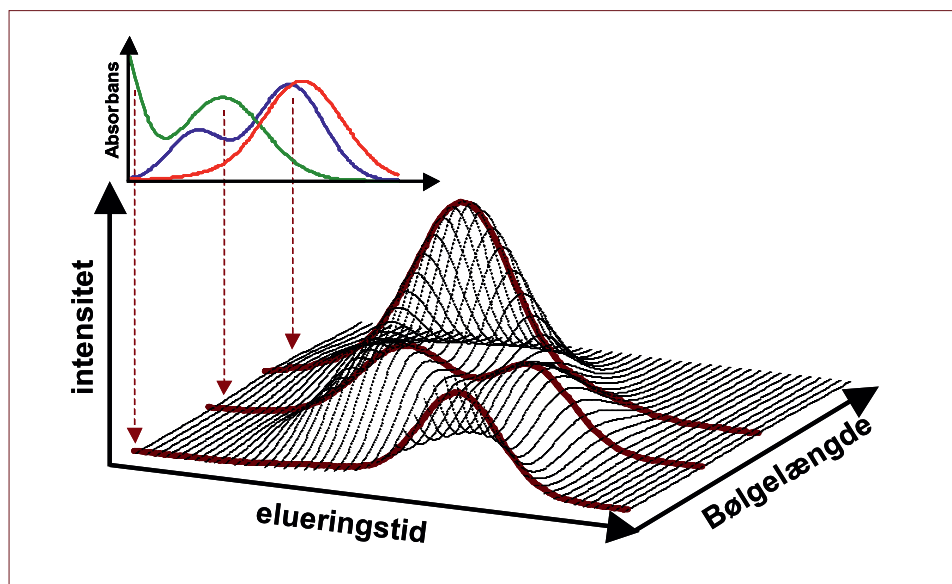
Inden for kromatografi udspringer behovet for udvikling af



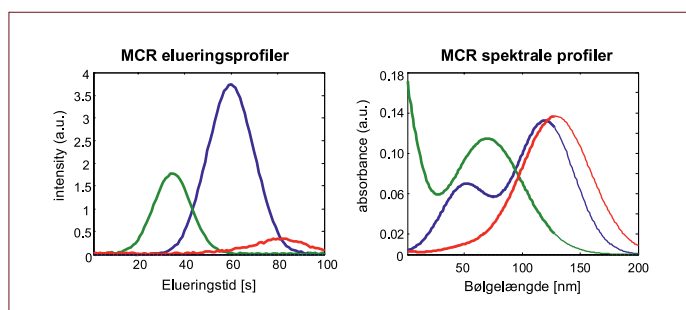
Figur 1. Simulering af et kromatografisk signal sammensat af tre Gauss-toppe og tilfældig støj.



Figur 2. Dekomponering af det kromatografiske signal i hhv. to (venstre) og tre (højre) Gauss-funktioner. Blå: teoretisk kromatografisk signal (under prikket grøn). Prikket grøn: beregnet kromatografisk signal. Rød: beregnede Gauss-toppe.



Figur 3. Rumlig repræsentation af den kromatografiske simulering vist i figur 1, hvor der er målt hele spektrale profiler for hver elueringstid. Den rene spektrale profil for hver top er vist i toppen af figuren. De brune tykke linjer angiver kromatografiske profiler opnået ved tre forskellige bølgelængder.



Figur 4: Eluerings- (venstre) og spektrale- (højre) profiler fundet ved anvendelse af Multivariate Curve Resolution (MCR).

kurvetilpasningsteknikker fra problemet med med-eluering (coelution). Den største fordel ved de kemometriske kurvetilpasningsteknikker er, at de kan rekonstruere den rene elueringsprofil og de rene spektrale profiler (dvs. de underliggende kemiske komponenter) ved at antage, at hver analyt har et unikt spektralt mønster. Et vigtigt aspekt er, at teknikkerne ikke bruger *a priori* empiriske funktioner som Gauss- eller Lorentz-kurver (hverken i elueringsprofilen eller i den spektrale profil). Det kromatografiske landskab i figur 3 kan visualiseres som en matrix  $X$  ( $I \times J$ ) med  $I$  elueringstider og  $J$  bølgelængder. Denne type data kan analyseres med Multivariate Curve Resolution (MCR) [3]. Den generelle MCR-model kan udtrykkes som:

$$X = CS^T + E \quad (1)$$

hvor  $C$  ( $I \times 3$ ) og  $S$  ( $J \times 3$ ) svarer til hhv. elueringsprofilerne og de spektrale profiler. Matricen  $E$  ( $I \times J$ ) indeholder residualerne, som ikke kunne forklares ved MCR-modellen (f.eks. støj og irrelevant struktur). Der er tydelig analogi mellem MCR og PCA (Dansk Kemi 2, 2008). Det er da også kun løsningsbetingelserne, der er forskellige for PCA og MCR. PCA forudsætter bl.a. ortogonale loadings og scores, hvilket ikke er tilfældet med MCR. Helt overordnet vil MCR modellere  $X$  ved et sæt af en elueringsprofil og en spektral profil for hver komponent:

$$X = c_1s_1^T + c_2s_2^T + c_3s_3^T + E \quad (2)$$

hvor eksempelvis  $c_1$  og  $s_1^T$  svarer til hhv. elueringsprofilen og den spektrale profil for den første komponent (den blå top). Ligning 2 afspejler, at metoden er tæt relateret til Lambert-Beers lov, når det gælder additivitet og linearitet. Den globale MCR-løsningsmodel for det kromatografiske landskab er vist i figur 4. Heraf ses det, at der er en god overensstemmelse mellem de beregnede MCR-profiler og de teoretiske (figur 1 og 3). Det er endvidere bemærkelsesværdigt, at MCR er i stand til at finde den røde top, selvom den er næsten fuldstændigt skjult under den blå top.

## Outro

MCR er et nyttigt redskab til at modellere spektrale signaler til deres underliggende rene toppe, og det er muligt at opnå gode estimater. Ulempen er, at man ikke nødvendigvis opnår unikke løsninger. Selvom de beregnede spektre og profiler fra MCR kan se meget "kemiske" ud, så er det ikke altid de rigtige kemiske komponenter, der afspejles. Resultaterne skal derfor altid valideres ift. viden om det kemiske system, der analyseres. Vi har fokuseret på kromatografiske data, men MCR-teknikker er generelt anvendelige i tilfælde, hvor det er muligt at opnå et komplet spektralt mønster som funktion af tid, prøver, sted eller andet. Kort sagt, hvor man kan danne en spektral datamatrix, og hvor Lambert-Beers lov er nogenlunde gældende.

## E-mail-adresser

José Manuel Amigo: jmar@life.ku.dk  
Søren Balling Engelsen: se@life.ku.dk  
Rasmus Bro: rb@life.ku.dk  
Lars Nørsgaard: lno@foss.dk

## Referencer

1. Di Marco VB, Bombi GG, Mathematical functions for the representation of Chromatographic peaks. *Journal of Chromatography A* 2001; 931: 1-30
2. Amigo JM, Skov TH, Bro R, ChroMATHography: Solving Chromatographic issues with mathematical models and intuitive graphics. *Chemical Reviews* 2010; 110:4582-4605
3. de Juan A, Tauler R, Multivariate Curve Resolution (MCR) from 2000: Progress in Concepts and Applications. *Critical Reviews in Analytical Chemistry* 2006; 36: 163-176