

INSTITUT D'ESTUDIS CATALANS
ARXIUS DE LES SECCIONS DE CIÈNCIES, CXV
SECCIÓ DE CIÈNCIES I TECNOLOGIA

ROMÀ TAULER I FERRÉ

ANÀLISI DE MESCLES
MITJANÇANT RESOLUCIÓ
MULTIVARIANT DE CORBES

PREMI ANTONI MARTÍ I FRANQUÈS (MARTÍ D'ARDENYA) 1994

BARCELONA
1997

ANÀLISI DE MESCLES MITJANÇANT
RESOLUCIÓ MULTIVARIANT DE CORBES

INSTITUT D'ESTUDIS CATALANS
ARXIU DE LES SECCIONS DE CIÈNCIES, CXV
SECCIÓ DE CIÈNCIES I TECNOLOGIA

ROMÀ TAULER I FERRÉ

ANÀLISI DE MESCLES
MITJANÇANT RESOLUCIÓ
MULTIVARIANT DE CORBES

PREMI ANTONI MARTÍ I FRANQUÈS (MARTÍ D'ARDENYA) 1994

BARCELONA
1997

Biblioteca de Catalunya. Dades CIP:

Tauler i Ferré, Romà

Anàlisi de mescles mitjançant resolució multivariant de corbes

Premi Antoni de Martí i Franqués (Martí d'Ardenya) 1994. – Bibliografia

ISBN 84-7283-353-4

I. Institut d'Estudis Catalans II. Títol III. Col·lecció 1. Quimiometria

2. Anàlisi multivariable – Informàtica

543.06:519.233

519.233:543.06

© Romà Tauler i Ferré

© 1997, Institut d'Estudis Catalans, per a aquesta edició

Editat per l'Institut d'Estudis Catalans

C. del Carme, 47. 08001 Barcelona

Primera edició: juliol de 1997

Tiratge: 600 exemplars

Compost per Amador Viñolas

Mestre Nicolau, 7, 5è 2a. 08021 Barcelona

Imprès a Limpergraf, SL

Carrer del Riu, 17. 08291 Ripollet

ISBN: 84-7283-353-4

Dipòsit Legal: B. 28530-1997

A proposta d'una ponència formada pels senyors Salvador Alegret i Sanromà, Enric Casassas i Simó i Josep Castells i Guardiola, membres numeraris de l'Institut d'Estudis Catalans, en sessió tinguda el dia 15 de març de 1994, acordà per unanimitat de concedir el Premi Martí i Franquès (Martí d'Ardenya) al senyor Romà Tauler i Ferré pel seu treball *Anàlisi de mescles mitjançant resolució multivariant de corbes*.

En la sessió tinguda el dia 10 d'octubre de 1994, la Secció de Ciències i Tecnologia prengué l'acord de publicar la dita obra, la qual, amb el títol *Anàlisi de mescles mitjançant resolució multivariant de corbes*, és editada a cura del senyor Enric Casassas i Simó.

This One



73PZ-8KQ-CLSR

TAULA

RESUM	11
1. INTRODUCCIÓ	13
1.1. Presentació	13
1.2. Alguns conceptes generals	17
1.2.1. Tipus de mesures en química (figures 1 i 2)	17
1.2.2. Altres conceptes i definicions de treball	21
2. EL MÈTODE DE RESOLUCIÓ MULTIVARIANT DE CORBES	25
2.1. Descripció del mètode	25
2.1.1. Descripció del problema	25
2.1.2. Anàlisi dels components principals	26
2.1.3. Ambigüitats d'intensitat i de rotació	28
2.1.4. Selectivitat	29
2.1.5. Detecció i ús de la informació selectiva	30
2.1.6. Complexitat del problema a resoldre mitjançant el mètode de resolució multivariant de corbes	33
2.1.7. Mètode de mínims quadrats alternats (ALS) i restringits d'una única matriu de dades	34
2.1.8. Mètode de mínims quadrats alternats (ALS) i restringits aplicats simultàniament a múltiples matrius de dades	36
2.2. Comprovació del mètode	42
2.2.1. Descripció de les dades	42
2.2.2. Detecció de la selectivitat	48
2.2.3. Anàlisi d'una matriu de dades pel mètode de mínims quadrats alternats (ALS)	54
2.2.4. Anàlisi simultània de diverses matrius de dades mitjançant resolució multivariant de corbes	58

2.2.5. Conclusions	66
3. L'APLICACIÓ DEL MÈTODE DE RESOLUCIÓ MULTIVARIANT DE CORBES (MVCR) A PROBLEMES REALS	67
3.1. Aplicació del mètode MVCR per a l'estudi de reaccions químiques en solució mitjançant valoracions espectromètriques	67
3.2. Aplicació del mètode MVCR per a l'estudi de la complexació en sistemes macromoleculars mitjançant valoracions espectromètriques	69
3.3. Aplicació del mètode MVCR a alguns sistemes de determinació analítica	79
3.4. Aplicació del mètode MVCR a l'anàlisi de processos industrials de producció química	83
3.5. Aplicació del mètode MVCR al desenvolupament de sensors òptics multivariants basats en reaccions químiques	89
3.6. Conclusions	99
4. DESENVOLUPAMENTS RECENTS DEL MÈTODE DE RESOLUCIÓ MULTIVARIANT DE CORBES	101
4.1. Millora del procediment d'optimització de mínims quadrats alternats restringits (optimització ALS)	101
4.1.1. Millora de la restricció de selectivitat	101
4.1.2. Millora de la restricció de no-negativitat	105
4.1.3. Millora de la restricció d'unimodalitat	106
4.1.4. Millora de la restricció de sistema tancat	107
4.1.5. Millora de la restricció de trilinealitat o d'igualtat de forma dels perfils	108
4.2. Aplicacions recents del mètode de resolució multivariant de corbes	110
4.2.1. Sistemes macromoleculars	110
4.2.2. Sistemes cromatogràfics	111
4.2.3. Sistemes de valoració espectrofluorimètrica	112
4.2.4. Sistemes de valoracions de flux continu i d'anàlisi per injecció de flux (FIA)	113
4.2.5. Sistemes mescla amb deficiència de rang	115
4.2.6. Sistemes voltamperomètrics	116
4.2.7. Sistemes mescla de contribucions de fonts medioambientals	117
5. APÈNDIXS	119
5.1. Programa MATLAB per a l'anàlisi de l'evolució dels factors mitjançant una finestra mòbil de mida fixa (<i>evolving factor analysis with a fixed size moving window FSMWEFA</i>)	121

5.2.	Programa MATLAB per a l'anàlisi de l'evolució dels factors (<i>evolving factor analysis, EFA</i>)	123
5.3.	Programa MATLAB de càlcul de les variables més pures segons el mètode Simplisma	126
5.4.	Programa MATLAB de resolució multivariant de corbes pel mètode d'optimització de mínims quadrats alternats i restringits (ALS)	129
5.5.	Programa MATLAB d'aplicació de la restricció d'unimodalitat	145
5.6.	Programa MATLAB per forçar els perfils de concentració a tenir la mateixa forma (dades amb estructura de segon ordre trilineal)	147
5.7.	Programa MATLAB d'aplicació de la restricció d'unimodalitat	150
5.8.	Programa MATLAB d'aplicació de la restricció de sistema tancat (<i>closure</i>)	154
5.9.	Programa MATLAB de descomposició PCA	156
5.10.	Programa MATLAB d'aplicació de la restricció de no-negativitat (nnls)	157
6.	REFERÈNCIES BIBLIOGRÀFIQUES	159
7.	ÍNDEX DE FIGURES	167

RESUM

En el present treball es descriu un nou procediment de resolució i d'anàlisi de mescles de compostos químics sense separació prèvia mitjançant l'aplicació del mètode de resolució multivariant de corbes. Es discuteixen detalladament quines són les condicions i les restriccions sota les quals es poden resoldre les ambigüitats inherents a les solucions matemàtiques de l'anàlisi de factors (*factor analysis*, FA). Un dels aspectes més rellevants del mètode proposat consisteix en l'aplicació dels mètodes basats en l'anàlisi de l'evolució dels factors (*evolving factor analysis*, EFA) i en la detecció de la selectivitat de les dades analitzades. L'aplicació del mètode de resolució multivariant de corbes descrit permet estimar el nombre d'espècies químiques presents, les seves concentracions i les seves respostes unitàries (per exemple els seus espectres). Quan el mètode és aplicat a l'anàlisi d'una única matriu de dades, l'obtenció de les solucions correctes depèn de la selectivitat del sistema estudiat i de l'aplicació adequada d'una sèrie de restriccions en l'optimització minimoquadràtica alternada (*alternating least squares*, ALS) de les solucions possibles. En canvi, quan l'anàlisi es realitza sobre dades amb estructures tensorials superiors, com per exemple, quan es realitza l'anàlisi simultània de diverses matrius de dades del mateix sistema químic, l'obtenció de les solucions correctes és més fàcil, eventualment sense ambigüitats. S'estudien dades amb estructures tensorials de complexitat diferent. Així mateix, s'estudia detalladament el cas força comú de l'anàlisi simultània de diverses matrius de dades que tenen un dels dos ordres (files o columnes) en comú. El mètode és comprovat primer amb dades simulades i després és aplicat a dades experimentals que provenen de diferents problemes reals que cobreixen un ampli ventall de situacions: *a*) valoracions espectromètriques de sistemes multiequilibri i de sistemes macromoleculars; *b*) sistemes de determinació analítica; *c*) anàlisi de processos industrials; i *d*) desenvolupament de nous sensors químics.

1. INTRODUCCIÓ

1.1. PRESENTACIÓ

L'anàlisi de mostres reals que contenen mescles complexes de compostos químics constitueix un dels reptes fonamentals de la química analítica actual. Aquest repte és encara més important quan aquests compostos o elements químics (analits) es troben a concentracions molt baixes i en presència d'altres components que tenen estructures i propietats similars i que, per tant, interfereixen en la seva determinació. L'anàlisi quantitativa de mescles complexes requereix disposar de mostres patró que tinguin les mateixes característiques i efectes matriu que les mostres reals a analitzar. La detecció de compostos purs (un cop totalment aïllats de la resta) pot fer-se amb diverses tècniques instrumentals de mesura d'alguna propietat de l'analit, mentre que el problema de les interferències, dels efectes matriu i del soroll de fons és un problema més complex que precisa d'una major atenció.

Actualment, el procediment més establert per a resoldre el problema de la manca de selectivitat en mostres complexes consisteix en la separació prèvia dels components químics purs abans d'analitzar-los. Si la separació ha estat total (p. e. mitjançant cromatografia) només cal una mesura de tipus univariant per tenir la informació quantitativa desitjada. Aquesta és l'estratègia que ha tingut més èxit fins a l'actualitat en l'anàlisi de mostres reals complexes. En general, no obstant això, aquesta metodologia analítica és costosa, tant perquè necessita disposar d'una instrumentació analítica sofisticada com pel temps que es necessita per a l'establiment de les condicions òptimes de separació. A més, aquesta estratègia no permet fer les anàlisis multicomponent de mescles de compostos químics *in situ* de forma ràpida i fiable. És en aquest context que té gran interès el desenvolupament de nous

mètodes quimiomètrics de tractament de dades poc selectives que permetin extreure la informació requerida sense la necessitat de la separació prèvia dels diversos components de la mostra.

D'altra banda, el desenvolupament de mètodes per a l'estudi, el seguiment i un millor control de reaccions i processos químics és essencial en molts camps de la química i la tecnologia actuals, tant en la síntesi de nous productes com en la millora del rendiment, la gestió i el disseny dels processos industrials de producció. El coneixement de l'especificitat i del rendiment de les reaccions i els processos químics i l'establiment de les condicions on són absents possibles reaccions laterals, contribueixen notablement a aconseguir un augment del valor afegit dels esmentats productes. És molt important el desenvolupament de nous mètodes que permetin l'anàlisi *in situ* de mostres reals, i que permetin, també, de fer-ho de forma ràpida, fiable i econòmica. En aquest context és especialment interessant el desenvolupament i l'aplicació de les eines matemàtiques que permetin d'extreure la màxima informació química a partir de l'anàlisi de dades experimentals obtingudes mitjançant tècniques espectrals no invasives.

La quimiometria és una disciplina relativament nova en l'àrea de la química, nascuda aproximadament al final dels anys 70,¹ que els darrers anys ha penetrat amb força en diversos camps, especialment els de la química analítica,² la química de síntesi,³ i l'enginyeria química (*Statistical Process Control*).⁴ Han sorgit també disciplines similars a la quimiometria en altres àrees de les ciències socials, com la psicometria⁵ i l'econometria.⁶ El desenvolupament ràpid de la quimiometria es deu fonamentalment a la penetració massiva dels ordinadors en la pràctica generalitzada de la química, tant de síntesi com d'anàlisi. Això és conseqüència de la revolució informàtica, però també de les noves característiques de la instrumentació química. S'ha passat de mesuraments puntuals univariants, d'un sol canal o longitud d'ona, a mesuraments espectrals multivariants multicanal que tenen un gran contingut informatiu. Els mètodes tradicionals d'anàlisi i procés de dades resulten totalment insuficients i l'embut es troba actualment en l'etapa de l'anàlisi de dades instrumentals. Així mateix, la informació requerida dels processos i les reaccions químiques és cada vegada més complexa i difícil d'interpretar. Per exemple, en l'anàlisi de processos de producció industrial (*Process Analysis*)⁷ és corrent el seguiment d'un determinat procés químic de síntesi mitjançant tècniques d'espectroscòpia no invasiva (UV, VIS, NIR, IR, RMN...) amb la finalitat de conèixer l'estat del sistema en cada moment de la producció i determinar-ne, així, l'evolució. D'altra banda, té un gran interès econòmic el fet de poder disposar de mètodes robustos d'anàlisi multivariant de dades, que es puguin aplicar *in situ* durant un procés químic, per controlar i corregir

les incidències en l'elaboració d'un determinat producte. El mateix ocorre en el laboratori o en la planta pilot quan es fa el seguiment de tot tipus de reaccions i processos químics. L'evolució d'una determinada reacció química o d'un procés de separació, com també la caracterització i la identificació dels productes intermedis i dels productes finals d'una reacció química, és una informació essencial en la síntesi de nous productes i en el disseny de nous processos. La quimiometria obre les portes a una nova era en el disseny, el seguiment i la millora de les condicions òptimes de les reaccions químiques.

Un dels camps de treball fonamental de la quimiometria és el del desenvolupament i l'aplicació de nous mètodes d'anàlisi estadística multivariant que permetin extreure la màxima informació química de l'anàlisi de mostres, reaccions i processos químics. El present treball s'emmarca en aquest context i presenta el desenvolupament d'un nou mètode de resolució multivariant de corbes que permet de fer l'anàlisi qualitativa i quantitativa de mostres que contenen mescles complexes, com també l'estudi i el seguiment de reaccions i processos químics. El mètode s'ha aplicat a diferents problemes químics de diversa complexitat, des de l'estudi de sistemes multicomponent en equilibri, fins al desenvolupament de nous sensors òptics basats en reaccions químiques, passant per l'anàlisi de processos industrials de producció i per la determinació analítica quantitativa de compostos coel·luïts en cromatografia HPLC. En tots aquests casos, el problema essencial a resoldre és l'extracció de la informació sobre la composició i les propietats del sistema a partir de senyals o respostes instrumentals amb poca, molt poca o cap selectivitat.

Els darrers anys, en els estudis desenvolupats al Departament de Química Analítica de la Universitat de Barcelona sobre sistemes multiequilibri formats per ions metàl·lics i lligands polifuncionals⁸⁻³⁰ es va prendre consciència de la necessitat de desenvolupar nous mètodes quimiomètrics que permetessin d'obtenir la informació química desitjada sense requerir la postulació prèvia d'un determinat model quimicofísic. L'estudi de reaccions i processos químics mitjançant mètodes tradicionals d'ajustament de dades per mínims quadrats³¹⁻³³ es basa en la postulació inicial d'un determinat model quimicofísic. Es tracta de trobar el model i els paràmetres quimicofísics que millor expliquen les dades experimentals mesurades. El model proposat, bé es coneix amb antelació o bé es postula i es comprova durant l'ajustament minimoquadràtic. Moltes vegades resulta difícil la descripció acurada de la reacció o el procés i es recorre a models locals o a models només vàlids sota condicions restringides. La quantitat de paràmetres a estimar i la qualitat intrínseca de les dades determinen els resultats assolits i poden malmetre'n considerablement la validesa. Aquesta aproximació només és especialment apropiada per a situacions d'estudi al laboratori sota condicions molt controlades.

L'objectiu del present treball és presentar el desenvolupament d'un nou procediment d'anàlisi de mesclres de compostos químics mitjançant tècniques quimiomètriques. La finalitat de la seva aplicació serà l'obtenció de la informació analítica directament. Es tracta de conèixer quantes espècies químiques són presents, quina és la seva naturalesa, quina és la seva concentració i com varia aquesta concentració al llarg del procés o reacció química estudiats. Un cop conegudes aquestes variables, és possible la deducció d'informació química de diferent índole, de tipus estructural (a partir, per exemple, dels espectres isolats) o quimicofísica (lleis d'acció de masses, lleis cinètiques, etc.). Els darrers anys, els mètodes de resolució multivariant de corbes^{34,35} han evolucionat com un grup de tècniques que són especialment adequades per a l'estudi de mesclres no resoltes de components químics sobre els quals no es disposa de cap informació prèvia, no se'n coneixen ni la naturalesa ni la composició química. Encara que els mètodes de resolució de corbes estan relacionats amb els mètodes de calibració multivariant,³⁶ els camps d'aplicació d'aquests dos grups de mètodes han estat considerats tradicionalment com a diferents. En els mètodes de resolució multivariant de corbes, l'interès se centra en la determinació d'informació qualitativa i en la resolució i la recuperació de les respostes unitàries dels components químics a les mostres. En canvi, en els mètodes de calibració multivariant la finalitat primera de l'estudi és l'estimació de la relació quantitativa existent entre mostres i patrons, i no es presta gaire atenció a la recuperació de les respostes pures vertaderes dels analits. Obviament, el mètode ideal serà aquell que permeti la solució d'ambdós problemes, la determinació de la informació quantitativa i la recuperació de les respostes unitàries dels diferents components químics o analits.

Els mètodes de resolució de corbes poden dividir-se en iteratius i no iteratius. Són procediments iteratius, per exemple, els mètodes *iterative factor analysis* (ITTFA),³⁷ *Simple to use Interactive Self-modeling Mixture Analysis* (SIMPLISMA),³⁸ o l'*Evolving Factor Analysis* (EFA).³⁹ En tots aquests casos és necessària una optimització restringida per mínims quadrats alternats (ALS) que millori les estimacions inicials. En canvi, en les aproximacions no iteratives, s'han emprat els conceptes de disminució del rang (*rank annihilation*) i de selectivitat (vegeu més avall). Exemples de mètodes no iteratius són el *rank annihilation evolving factor analysis* (RAEFA),⁴⁰ *window factor analysis* (WFA),⁴¹ i *Heuristic Evolving Latent Projections* (HELP).⁴²

Lawton i Sylvestre⁴³ foren dels primers d'adonar-se que quan s'imposaven les restriccions de linealitat i no-negativitat sobre l'optimització, s'obtenia una banda molt més estreta de possibles solucions. Això es demostrà matemàticament per al cas de mesclres no resoltes de dos components. Més tard, Borgen i

Kowalski⁴⁴ ampliaren aquest tractament en mesclades de tres components. L'aplicació, però, a sistemes de més de tres components és difícil.⁴⁵

S'obté una nova perspectiva del problema quan s'analitzen dades experimentals ordenades en estructures matemàtiques superiors. S'ha demostrat que el problema de la resolució i la quantificació en presència d'interferències es resol quan hom disposa de dades organitzades en estructures superiors de tipus tensor.⁴⁶ És el cas, per exemple, de les dades experimentals obtingudes per l'acoblament de la cromatografia amb l'espectroscòpia multicanal. Mètodes com *Generalized Rank Annihilation* (GRAM),⁴⁷ *Trilinear Decomposition* (TLD),⁴⁸ o els mètodes basats en models *Parallel Factor Analysis* (PARAFAC)⁴⁹ han estat proposats per a aquesta finalitat. El límit d'aquestes aproximacions es troba en els seus mateixos requeriments perquè necessiten d'una estructura de dades de segon ordre trilineal.⁴⁷⁻⁴⁸ Actualment s'està treballant en el desenvolupament de nous mètodes que no pateixin d'unes restriccions massa abusives per a ésser útils en l'anàlisi de dades reals. El present treball s'inscriu en aquest context.

1.2. ALGUNS CONCEPTES GENERALS

1.2.1. TIPUS DE MESURES EN QUÍMICA (FIGURES 1 I 2)

1.2.1.1. *Mesures univariants. Sensors d'ordre zero*

El tipus més senzill de mesures en química és de naturalesa univariant, és a dir, mesures en les quals només s'obté un valor numèric o escalar. L'instrument de mesura emprat pot anomenar-se *sensor d'ordre zero* per analogia amb la denominació matemàtica tensorial on els valors numèrics escalars s'anomenen *tensors d'ordre zero*. Hi ha molts casos en què aquesta situació es dona, per exemple en les mesures potenciomètriques de pH o d'un altre ió determinat mitjançant un elèctrode d'ió selectiu o també el de les mesures espectroscòpiques realitzades a una sola longitud d'ona o freqüència.

Els mètodes matemàtics emprats en l'extracció de la informació química en el cas dels sensors d'ordre zero són l'estadística univariant i la regressió lineal univariant; n'és un exemple la calibració univariant en química analítica.⁵⁰ L'aplicació d'aquests mètodes suposa que la resposta instrumental és conseqüència de la presència d'un sol component químic (selectivitat total), és a dir, que no hi ha cap tipus d'interferència ni soroll de fons sobre el senyal instrumental mesurat. La dependència amb la concentració d'aquest

únic analít generalment és de tipus lineal. Aquest tipus de calibració univariant no es pot aplicar en presència d'interferències ni de contribucions desconegudes del soroll de fons. És obligatòria la separació prèvia dels components interferents i la «neteja» del senyal de fons abans de procedir a l'anàlisi. A més, no és possible l'anàlisi multicomponent o l'anàlisi simultània de diferents components de la mostra.

Tipus de sensors en química

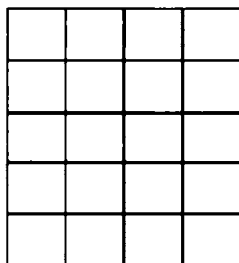
ORDRE 0
ISE, pH,



ORDRE 1
espectres



ORDRE 2
LC-DAD



ORDRE 3
temps-excitació-emissió

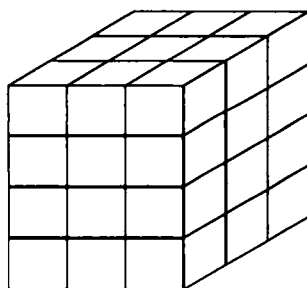


FIGURA 1. Tipus de sensors.

1.2.1.2. *Mesures multivariants*

Mesures vectorials. Sensors de primer ordre

El primer cas de mesures multivariants es dóna quan de cada mostra o objecte analitzat s'obté un vector de dades (p. e., un espectre). La mesura proporciona un conjunt de valors numèrics que es poden ordenar en forma de vector, i l'instrument de mesura s'anomena *sensor de primer ordre* per analogia amb el nom que reben els vectors com a tensors de primer ordre. En són exemples les mesures espectromètriques UV, VIS, IR, NIR, NMR, EPR, etc.

Però també ho són tots aquells mètodes instrumentals que proporcionen una informació numèrica vectorial en l'anàlisi de cada mostra, per exemple a partir de la suma d'un conjunt de sensors univariants o a partir del seguiment de la resposta amb el temps d'un sensor univariant (p. e. cromatografia, potenciometria)

Els mètodes matemàtics emprats són els mètodes derivats de l'anàlisi multivariant de dades i de l'estadística multivariant. En particular, darre-rament s'han desenvolupat amb força els mètodes anomenats *de calibració multivariant*^{36, 51}. En aquest cas no és necessari modelar explícitament les inter-ferències ni el senyal de fons abans d'extreure la informació química requerida. Encara es necessita però, que els patrons emprats per a l'anàlisi tinguin les mateixes interferències i contribucions al senyal instrumental que a les mostres problema. Quan el problema conté interferències no presents en els patrons (cas comú en la determinació de mostres reals), els mètodes de calibració multivariant tampoc donaran resultats satisfactoris, però sí que permetran la detecció d'aquestes interferències. Amb els sensors de primer ordre no es pot resoldre encara de forma completa el problema de les inter-ferències i del senyal de fons.

Mesures d'una matriu de dades. Sensors de segon ordre

El segon cas de mesures multivariants es dona quan per a cada mostra analitzada s'obté una matriu de dades. Aleshores la mesura d'una determina-da mostra o objecte proporciona tot un conjunt de dades que poden ordenar-se en forma de taula o de matriu. En aquesta taula o matriu hi ha dos ordres o direccions, les files i les columnes, que corresponen generalment a dos tipus diferents de mesures. Per exemple, en el cas de l'acoblament cromatografia-espectrometria, cada fila de la matriu (fig. 2) és un espectre a un temps de retenció determinat i cada columna de la matriu és un cromatograma a una longitud d'ona determinada. Hi ha, doncs, dos ordres de mesura, el cromato-gràfic o temporal (temps d'elució) i l'espectral (longitud d'ona). En la notació tensorial, una matriu és un tensor de segon ordre; per això aquest tipus de mesures reben el nom de mesures (realitzades per sensors) de segon ordre. Hi ha altres casos en què també es pot aplicar aquesta denominació, per exemple quan es combinen les mesures d'emissió fluorescent a diferents longituds d'ona d'excitació; quan es fa el seguiment espectral de processos químics en evolu-ció (*process analysis*); i quan es fa el seguiment de respostes de tipus multivariant amb el temps, pH o concentració.

Per a tractar el tipus de dades proporcionades per sensors de segon ordre han estat desenvolupats diversos mètodes matemàtics. Destaquen, entre ells,

els mètodes derivats de l'anàlisi de factors (*Factor Analysis*).⁵² En aquest treball es presenta el desenvolupament del mètode de resolució multivariant de corbes per a tractar dades de tipus matricial.

Mesures d'un cub de dades. Sensors d'ordre superior

L'extensió dels conceptes anteriors ens porta cap als sensors d'ordre superior. En aquest cas per a cada mostra o objecte estudiat obtenim una ordenació de dades en forma de cub o en tres ordres. Un exemple d'aquest tipus de dades és quan es comparen i s'analitzen simultàniament diferents matrius de dades obtingudes amb un sensor de segon ordre com els indicats en l'apartat anterior. Aquest és el cas de la comparació de matrius de dades obtingudes per acoblament cromatografia-espectrometria en l'anàlisi simultània de patrons i problemes. És el cas, també, del seguiment amb el temps d'un procés o d'una reacció química mitjançant emissió fluorescent a diverses longituds d'ona d'excitació. O, també, el cas de comparació de matrius de dades obtingudes en el seguiment espectral de reaccions cinètiques o en el seguiment de valoracions espectrosmètriques.

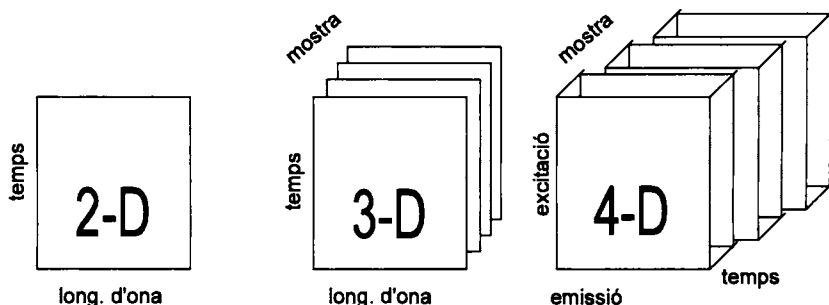


FIGURA 2. Exemples de tensors de dades obtinguts amb sensors d'ordre superior: tensors de segon ordre (matrius) 2-D (en HPLC-DAD, el temps d'elució i la longitud d'ona), tensors de tercer ordre 3-D (en HPLC-DAD, les mostres, el temps d'elució i la longitud d'ona), i tensors de quart ordre 4-D (anàlisi de mostres que evolucionen amb el temps per fluorescència: les mostres, el temps de mesura, la longitud d'ona d'excitació i la longitud d'ona d'emissió).

Els mètodes matemàtics a emprar en aquest cas estan dins l'àrea de l'anàlisi matemàtica de tensors o del que en la literatura anglosaxona s'anomena *three-way data analysis*.⁵³ Aquest és un camp d'investigació actual en què les dades de tipus quàntic tenen unes característiques pròpies, l'explotació de les quals ha permès el desenvolupament de nous mètodes quimiomètrics. El mètode descrit en el present treball permet també fer l'anàlisi de dades amb aquestes estructures d'ordre superior.

1.2.2. ALTRES CONCEPTES I DEFINICIIONS DE TREBALL

Hi ha una sèrie de conceptes i definicions de treball que són útils en la descripció del mètode proposat i que es resumeixen a continuació.

1.2.2.1. *Model lineal*

Un dels primers conceptes que s'utilitza és el del comportament lineal de les dades adquirides en un experiment determinat. Per a dades multivariants, el model lineal s'estén fàcilment. En forma matricial es pot escriure:

$$D = C S + E \quad [1]$$

on D és una matriu de dades experimentals i C i S són dues matrius de dimensions més petites que tenen un determinat significat quàntic (vegeu més avall), i E és la matriu d'error associada a la mesura experimental. Quan és possible una descomposició de les dades experimentals com la proposada, la matriu D s'anomena *bilineal*.

1.2.2.2. *Ordres de mesura d'una matriu o tensor de dades*

La matriu D experimental està ordenada, segons les seves files i les seves columnes, en dos ordres: direccions, maneres [*ways*] o modes [*modes*] de mesura. Els dos ordres de la matriu experimental D estan relacionats amb les dimensions de les matrius C i S en què es descompon. Així, la matriu C té el mateix nombre de files que la matriu D i proporciona informació sobre la variació de les diferents espècies químiques segons l'ordre de les files. Tanmateix, la matriu S té el mateix nombre de columnes que la matriu D , està

associada amb l'ordre de les columnes de D i indica com varien les espècies químiques presents segons l'ordre de les columnes de D . Normalment les columnes de la matriu descriuen l'ordre de mesura espectral (longituds d'ona) i les files descriuen l'ordre de mesura de canvi de concentració.

En l'anàlisi simultània de diferents matrius de dades es poden distingir diferents tipus de situacions:

1. Cap dels ordres es comú entre les diferents matrius. Òbviament en aquest cas no té cap avantatge l'anàlisi conjunta de les diverses matrius de dades. L'estructura de les dades no és d'ordre superior i no hi ha cap avantatge de segon ordre.⁴⁶

2. Un dels ordres és comú entre matrius: cada matriu de dades té un dels dos ordres en comú amb les altres matrius de dades. És el cas, per exemple, de l'estudi simultani de diverses valoracions espectroscòpiques del mateix sistema multiequilibri a condicions inicials diferents⁵⁴⁻⁵⁶ o d'un mateix procés químic seguit en diferents situacions químiques.⁵⁷ Aquest tipus de sistemes són estudiats en detall més endavant. L'estudi simultani de les diferents matrius de dades comporta uns beneficis clars en la resolució de les espècies i en la deducció dels fenòmens químics que ocorren.

3. Dos dels ordres són comuns entre matrius: cada matriu de dades té els dos ordres (files i columnes) comuns entre les diferents matrius de dades que s'analitzen simultàniament. Si les dades mesurades tenen aquesta estructura, es diu que tenen una estructura de segon ordre trilineal.⁴⁶ Aquest és un cas de gran interès, sobretot en química analítica, ja que permet de resoldre totalment el problema de les interferències i del soroll de fons; és el que s'anomena *avantatge de segon ordre (2nd order advantage)*.⁴⁶ L'anàlisi simultània de les diverses matrius de dades permet la resolució directa dels components i dels fenòmens químics existents. És, per exemple, el cas de les tècniques cromatogràfiques acoblades als mètodes espectrals, quan es dona una situació de reproductibilitat i sincronització en l'ordre cromatogràfic.

1.2.2.3. Rang, pseudorang i rang químic d'una matriu de dades

A causa de l'error experimental present en qualsevol mesura, el rang matemàtic de la matriu D serà el nombre més petit entre el nombre de files o columnes de la matriu. No obstant això, el nombre de causes físiques (o químiques) de variació de les dades no associades amb l'error experimental serà un nombre molt més petit, anomenat *pseudorang* de la matriu D . Estem interessats

especialment en aquelles causes de variació associades amb els components químics presents en D o rang químic. Aquest últim es definirà com el nombre d'espècies químiques independents causants de la variació observada en D . Trobar el pseudorang, i especialment, el rang químic no serà sempre fàcil i és un dels primers problemes a resoldre en qualsevol mètode d'anàlisi multivariant de dades com el que aquí es proposa.

1.2.2.4. *Espècie química, component i analit*

En el llenguatge químic sovint s'utilitzen conceptes com espècie, component o analit de forma sinònima. En alguns casos aquestes denominacions són confuses perquè es poden referir a entitats químiques diferents. Així, per exemple, un analit o subjecte de l'anàlisi pot estar constituït per diverses espècies químiques. És el cas, per exemple, de la determinació del contingut total d'un determinat ió metàl·lic quan aquest es troba distribuït en diferents espècies químiques en equilibri o reacció. *Component químic* és un terme genèric i es refereix a una de les entitats que són presents en una determinada mescla. Per tant, tant pot ésser una espècie com un analit. El significat concret d'aquests termes es dedueix a partir del context on es troben. En general quan no s'especifica clarament són termes sinònims.

1.2.2.5. *Selectivitat*

Un altre dels conceptes que s'utilitza més freqüentment en la discussió que segueix és el de la selectivitat. Una mesura és selectiva quan el seu valor respon només a la presència d'un sol component o espècie química. Les zones o regions de mesura selectives d'una matriu seran zones o regions de la matriu (submatrius) que tenen un rang químic igual a 1. Les zones de mesura on no existeixi de forma apreciable cap espècie química seran zones de rang químic igual a 0. Com es demostra més endavant, és molt útil conèixer amb detall la distribució local del rang químic de la matriu de dades D .

1.2.2.6. *Perfils espectrals i perfils de variació de concentració*

En els sensors de segon ordre o d'ordre superior cada espècie química es caracteritza com a mínim per dos vectors resposta que s'anomenen *perfils*, ja

que moltes vegades presenten formes contínues determinades. Un dels dos perfils descriu la resposta unitària de l'espècie química segons l'ordre de mesura espectral, és l'espectre unitari de l'espècie considerada i és útil en la seva identificació. De manera anàloga, si en lloc de mesures de tipus espectroscòpic es realitzen mesures instrumentals multicanal d'un altre tipus, com per exemple mesures electroquímiques, els corresponents perfils unitaris també descriuran la resposta de l'espècie química segons aquesta tècnica i també seran útils en la seva caracterització.

L'altre vector o perfil que caracteritza moltes vegades una determinada espècie química és el seu perfil de concentració, o variació de la concentració de l'espècie química segons una variable del sistema estudiat (pH, temps, etc.). És el cas dels perfils d'elució en els sistemes cromatogràfics, dels diagrames de distribució d'espècies en sistemes multiequilibri o dels perfils temporals en les reaccions i els processos cinètics. Aquests perfils es poden descompondre en dos termes: el perfil unitari pròpiament dit, que expressa la variació per unitat de concentració, i un altre terme que és un valor numèric o escalar, que expressa la contribució o la concentració total de l'espècie.

En alguns sensors o mesures d'ordre superior, cada espècie química pot estar caracteritzada per més d'un vector o perfil espectral. És el cas, per exemple, de les mesures per fluorescència on es pot tenir un perfil espectral d'excitació i un perfil espectral d'emissió per a cada espècie química.

2. EL MÈTODE DE RESOLUCIÓ MULTIVARIANT DE CORBES

2.1. DESCRIPCIÓ DEL MÈTODE

2.1.1. DESCRIPCIÓ DEL PROBLEMA

El mètode de resolució multivariant de corbes assumeix que les dades experimentals segueixen un model de tipus lineal similar a la generalització matricial de la llei de Lambert-Beer en espectroscòpia d'absorció. En forma matricial, aquest model s'escriu de la següent manera:

$$D = C S + E \quad [1]$$

on D és la matriu de dades dels espectres adquirits a valors diferents d'una certa variable (temps, pH, concentració...) durant una reacció o un procés químic; C i S són, respectivament, les matrius dels perfils de concentracions i de respostes unitàries de les espècies químiques actives espectroscòpicament involucrades en el procés o la reacció química, i E és la matriu dels residuals no explicats per les espècies químiques definides en C i S . Aquesta matriu de residuals no explicats per la descomposició en C i S ha d'ésser d'una magnitud similar a l'error experimental associat a les condicions particulars de mesura en l'experiment. Les dimensions d'aquestes quatre matrius són $D(I \times J)$, $C(I \times N)$, $S(N \times J)$ i $E(I \times J)$, on I és el nombre d'espectres analitzats, J és el nombre de canals espectroscòpics i N és el nombre d'espècies químiques presents en D .

* Sense pèrdua de generalitat, en el present treball s'utilitzarà la terminologia usualment emprada per l'espectroscòpia d'absorció, que pot fàcilment estendre's a qualsevol altra situació on es compleixi el model lineal.

Cada dada individual d_{ij} de D és l'absorció mesurada experimentalment a la longitud d'ona j en la mostra-espectre i ; quan el model lineal es compleix, d_{ij} és la suma de l'absorció de les diferents espècies n presents en aquesta mostra-espectre a la longitud d'ona j

$$d_{ij} = \sum_n c_{in} s_{nj} + e_{ij} \quad [2]$$

on c_{in} representa la concentració de l'espècie n en la mostra-espectre i , s_{nj} representa l'absorció unitària de l'espècie n a la longitud d'ona j , i e_{ij} és l'error residual.

A part d'obtenir el nombre d'espècies químiques en C i S (rang químic de D), la finalitat principal dels mètodes de resolució multivariant de corbes és la determinació de les vertaderes matrius C i S a partir només de l'anàlisi de D . Moltes vegades és lícit d'assumir que els diferents espectres de D han estat mesurats de manera contínua i ordenada com a conseqüència de l'evolució d'una reacció química o d'un procés de separació. En aquesta situació les concentracions en C tenen una forma contínua i s'anomenen *perfils de concentració*.

Hi ha dos ordres de mesura per a cada matriu de dades; d'una banda, l'ordre espectral, i de l'altra, l'ordre de canvi de concentració (per exemple el temps d'elució en cromatografia, el pH en les valoracions espectromètriques o el temps en les reaccions cinètiques). L'ordre espectral ve descrit per les columnes de la matriu original de dades (índex j en l'equació [2]) i l'ordre de canvi de concentració, per les files de la matriu original de dades (índex i en l'equació [2]). La informació en cada ordre es correspon amb la informació associada a les dues matrius C i S en què es descompon la matriu original de dades D .

2.1.2. ANÀLISI DELS COMPONENTS PRINCIPALS

La matriu de dades D es pot descompondre mitjançant l'anàlisi de components principals:

$$D = U S V^T + E = D' + E \quad [3]$$

on $U S$ i V^T són respectivament, les matrius *score*, valors singulars (*singular value*) i *loading* de D per un nombre preseleccionat de components; E és la matriu residual d'error que conté la variança en D no explicada pel producte

$U S V^T$; i D^* és la matriu de dades reproduïda. Sota l'assumpció de linealitat, quan s'escull el nombre correcte de components, la matriu residual d'error E és pròxima al soroll o a l'error experimental. La transformació de les matrius $U^* S$ i V^T en C i S , respectivament, no es pot assolir directament si no hi ha una informació addicional. Com que C i S són inicialment desconegudes, l'equació:

$$D^* = U S V^T = U S T T^{-1} V^T = C S \quad [4]$$

té un nombre infinit de possibles solucions per a qualsevol matriu de rotació-transformació T . En les solucions per components principals per trobar C i S hi ha una ambigüitat intrínseca si no es proporciona informació addicional sobre llur naturalesa. La tasca general dels mètodes de resolució multivariant de corbes és restringir el nombre de solucions possibles de manera que eventualment s'aconsegueixi de trobar aquelles solucions físicament i químicament acceptables de les matrius C i S .

Una primera indicació del nombre d'espècies químiques presents en D es troba a partir de l'anàlisi del rang químic o nombre de valors singulars més significatius associats amb la matriu de dades D . Es pot trobar una estimació senzilla del nombre d'espècies quan se sap que l'error experimental és de tipus aleatori i se'n coneix la magnitud (la seva variança). En aquest cas, el nombre d'espècies químiques vindrà donat pel nombre de components principals necessaris en l'equació [3] per a reproduir la matriu original D , és a dir, quan la diferència entre D i D^* és del mateix ordre que l'error experimental. Això no obstant, a la pràctica aquesta selecció de vegades no és senzilla. Si s'accepta que la variança observada en la matriu de dades ve causada fonamentalment pels canvis en la concentració de les diferents espècies químiques presents en les barreges, el nombre d'espècies pot calcular-se a partir del nombre de valors singulars més significatius. A la bibliografia s'han descrit diferents mètodes per a dur a terme aquesta tasca, com són les tècniques de validació creuada de Wold (*cross validation*)⁵⁸ o la teoria de l'error en l'anàlisi de factors de Malinowski (*error in factor analysis*).⁵⁹ En general, aquests mètodes depenen de les característiques i del nivell de soroll i no són útils en la discriminació de les contribucions de les espècies químiques respecte a les contribucions associades amb el soroll de fons, amb els moviments de les línies base (*baseline shift*) i amb les derives dels detectors instrumentals (*detector drifts*). Moltes vegades, part de la variança observada en la matriu de dades no és causada pels canvis de concentració de les espècies químiques, sinó per altres causes com les abans esmentades. Els valors singulars associats amb aquestes causes són més grans que els valors singulars associats amb l'error pur experimental

de tipus aleatori. Per això, poden estimar-se a partir de l'anàlisi de les regions espectrals on no hi ha absorció de les espècies químiques. Mitjançant la comparació d'aquests valors singulars amb els valors singulars trobats en les regions espectrals, on les espècies químiques absorbeixen, es pot realitzar una primera estimació del nombre d'espècies present.

El nombre d'espècies estimat d'aquesta manera pot comprovar-se posteriorment utilitzant altres tècniques com l'anàlisi de l'evolució dels factors (*evolving factor analysis*).³⁹ Durant l'optimització per mínims quadrats alternats (vegeu més avall) buscant sempre aquelles solucions que millor expliquin les dades, que millor les ajustin en un sentit mínim quadràtic i que millor sentit físic tinguin: els perfils de concentració en C i els espectres unitaris en S han de tenir formes reals.

2.1.3. AMBIGÜITATS D'INTENSITAT I DE ROTACIÓ

Hi ha dues classes d'ambigüitats associades als mètodes de resolució multivariant de corbes: les ambigüitats d'intensitat i les ambigüitats rotacionals.⁶⁰ L'ambigüitat d'intensitat és sempre present en les descomposicions per anàlisi de factors o per resolució multivariant de corbes ja que per a qualsevol escalar m ,

$$d_{ij} = \sum_n c_{in} s_{nj} + e_{ij} \quad [5]$$

$$c_{in} s_{nj} = (c_{in} m) (1/m s_{nj}) = c'_{in} s'_{nj} \quad [6]$$

Això vol dir que les concentracions i els espectres estimats es trobaran escalats per un factor desconegut m , la qual cosa no és un problema important en anàlisi qualitativa o en la identificació, d'espectres (*fingerprinting*), però, òbviament és un problema seriós en l'anàlisi quantitativa.

Més important encara és l'ambigüitat rotacional també inherent a les solucions trobades mitjançant l'anàlisi de factors o mitjançant la resolució multivariant de corbes, que ocorre generalment quan hi ha dos o més components linealment independents que es superposen. L'espectre estimat per a qualsevol d'aquests components serà una combinació lineal desconeguda dels veritables espectres

$$s' = \sum_k t_k s_k \quad [7]$$

on t_k són constants de rotació desconegudes; s_k són els vertaders espectres unitaris dels components o espècies químiques presents; i s' són els espectres unitaris estimats. El mateix es pot dir per als perfils de concentració en la matriu C .

Contràriament, per a aquells espectres mesurats on només és present una espècie o en el cas que només una espècie química absorbeixi una determinada longitud d'ona, no hi ha ambigüitat rotacional. Això vol dir que les solucions obtingudes mitjançant anàlisi de components principals dona ja les formes correctes d'espectres unitaris o perfils de concentració. L'ambigüitat rotacional pot resoldre's en les regions de selectivitat total en qualsevol dels dos ordres de mesura experimental. Una vegada s'aconsegueix eliminar l'ambigüitat rotacional només roman l'ambigüitat d'intensitat. Per a qualsevol escalar m ,

$$d_{ij} = u_{i1} v_{1j} = (c_{i1} m) (1/m s_{1j}) \quad [8]$$

$$c_{i1} = u_{i1}/m \quad [9]$$

$$s_{1j} = v_{1j} m \quad [10]$$

Quan només s'analitza una matriu de dades, no es pot resoldre aquesta ambigüitat d'intensitat si no es proporciona informació addicional (per exemple, que el sistema és «tancat»; vegeu mes avall).

2.1.4. SELECTIVITAT

Com ja s'ha fet esment, la selectivitat és la pedra de toc en la resolució de les ambigüitats rotacionals en l'anàlisi d'una única matriu de dades. Les regions selectives són aquelles regions en qualsevol dels dos ordres de la matriu de dades, on només és present una espècie química, és a dir, les regions de la matriu de dades on el rang químic és la unitat. Per exemple, aquelles longituds d'ona on només una espècie química absorbeix, o aquells temps d'elució o pH on només hi ha una espècie. En aquestes regions no hi ha ambigüitat rotacional, i tant la solució per anàlisi de components principals com la solució pel mètode de mínims quadrats alternats (vegeu més avall) convergeixen cap a la solució correcta. Això no és cert per a les regions no selectives, ja que en aquestes l'ambigüitat rotacional persisteix. Quan existeix selectivitat per a una espècie química determinada en els dos ordres de la matriu de dades, aquesta espècie es troba immediatament caracteritzada sense cap ambigüitat rotacional, i tant el seu espectre unitari com el seu perfil de variació de la

concentració es determinen fàcilment. Quan això ocorre per a totes les espècies presents en D , el problema es resol de forma completa.

Si hi ha selectivitat parcial per a una determinada espècie només en un dels dos ordres de mesura, l'ambigüitat es resol per al perfil d'aquesta espècie en l'ordre oposat. En canvi, l'ambigüitat persisteix per a aquells canals del perfil en l'ordre considerat on no hi ha selectivitat. Per exemple, si a una determinada longitud d'ona només un component absorbeix, el perfil de concentració queda totalment determinat a partir de la variació de la resposta del detector a aquesta longitud d'ona, però l'espectre d'aquesta espècie en les longituds d'ona no selectives no es pot conèixer de forma certa. En el cas oposat, quan a un cert temps o pH només hi ha una espècie, l'espectre d'aquesta espècie es determina sense ambigüitats, però el perfil de variació de concentració d'aquesta mateixa espècie a altres valors de pH o de temps no es pot conèixer sense ambigüitats. La raó d'aquesta ambigüitat romanent es deu al fet que en les regions on el rang químic és superior a 1, encara hi són possibles combinacions lineals dels perfils vertaders de les espècies presents.

Un cas particular i interessant és quan pel perfil d'una certa espècie existeix una zona de rang unitari i una zona on aquesta espècie concreta no existeix. Per a una espècie que compleixi aquestes característiques, es poden determinar sense ambigüitats les constants de rotació t_k en l'equació [7] i el perfil complet es recupera. Mètodes que han fet servir aquesta idea han estat HELP de Kvalheim,⁴² RAEFA de Maeder⁴⁰ i WFA de Malinowski.⁴¹

En cas de tenir selectivitat només en un dels dos ordres, la utilització de restriccions addicionals (vegeu més avall) pot ajudar a acabar de resoldre de forma completa les ambigüitats romanents. Algunes de les restriccions, com la unimodalitat, poden tenir un paper fonamental. Quan no hi ha selectivitat en cap dels dos ordres per a cap de les espècies presents en una barreja no resolta, els mètodes de resolució multivariant de corbes encara donaran solucions ambigües. La utilització de restriccions addicionals pot reduir la quantitat d'ambigüitat romanent, però no hi ha garantia que hagi estat totalment resolta.

2.1.5. DETECCIÓ I ÚS DE LA INFORMACIÓ SELECTIVA

Hi ha diverses maneres de detectar les regions o «finestres» on hi ha cada espècie. El primer que s'ha de fer és disposar d'una estimació del nivell de soroll de fons o límit de detecció per sobre del qual una espècie química es pot considerar que és present. Per a tal fi, s'investiguen les regions espectrals on se sap que no hi ha absorció de les espècies químiques o s'investiguen els

temps o pH en què no hi ha formació de cap espècie. A partir de l'anàlisi d'aquestes regions es pot aconseguir una estimació del soroll de fons.

La determinació de les regions selectives o regions en què el rang químic és 1 (només és present una espècie química) es pot aconseguir a partir de diferents mètodes, la majoria dels quals es basen en alguna variant del que hom anomena *anàlisi del rang local* (*local rank analysis*) o dels mètodes de *detecció de variables pures*. Entre aquests mètodes destaca el mètode de l'anàlisi de l'evolució dels factors, especialment en la seva variant de finestra mòbil de mida fixa (*evolving factor analysis with a moving fixed size window*, FSMW EFA).⁴² En aquest mètode, el rang local d'una matriu mòbil de mida fixa es calcula al llarg dels dos ordres de la matriu (vegeu apèndix 5.1). Aquelles regions, on només hi ha un component o espècie química, mostren un valor singular significativament més gran que els valors singulars associats a l'error experimental, que són molt més petits. Les regions on hi ha superposició d'espècies, amb rang superior a 1, mostren més d'un valor singular significatiu per sobre dels valors singulars associats a l'error experimental. Aquesta determinació del rang local pot ésser realitzada indistintament per a qualsevol dels dos ordres o direccions de la matriu experimental.

Aquestes estimacions del rang local poden ésser errònies quan la mida de la finestra mòbil és molt petita i els perfils de les espècies superposades són molt similars (la seva diferència és pròxima a l'error experimental). En aquests casos és millor augmentar la mida de la finestra, però, aleshores, es perd l'habilitat del mètode en la detecció local del rang. Cal trobar un compromís en cada cas. Les capacitats de detecció del mètode disminueixen quan el nivell del soroll associat al senyal augmenta i quan la contribució pròpia de l'espècie al senyal global disminueix (vegeu el concepte de *net analyte signal*).⁶²

Els mètodes de detecció de variables pures³⁸ determinen les variables més selectives per a un conjunt de dades determinat. Aquesta detecció pot realitzar-se també en les dues direccions o ordres de la matriu experimental de dades. A diferència dels mètodes de detecció del rang local, els mètodes de detecció de les variables més pures no estan restringits a un nombre mínim de variables. De tota manera aquests mètodes no permeten assegurar si la variable més pura detectada és completament selectiva o no. Ambdós grups de mètodes (estimacions del rang local i detecció de les variables més pures) poden emprar-se de forma complementària. Una vegada les variables més pures han estat detectades (en qualsevol dels dos ordres), el càlcul del rang local al voltant d'aquestes pot confirmar, o no, la seva selectivitat total.

Tant els perfils recuperats a partir de les variables més pures, com els recuperats a partir de les regions on el rang local és petit (pròxim a 1), són bones estimacions inicials en la resolució multivariant de corbes (vegeu més

avall). Els perfils trobats a partir de variables totalment selectives proporcionen solucions sense ambigüïtat rotacional. A l'apèndix 5.3 es dona el programa per a calcular les variables més pures a partir del procediment proposat per Windig.³⁸

El següent pas a fer és l'estimació de la regió o finestra on cada espècie química existeix i l'estimació de la regió o finestra on cada espècie química no existeix (finestra de concentració pràcticament nul·la). És molt útil en aquest context d'utilitzar el mètode d'anàlisi de l'evolució dels factors en la seva forma estàndard (*evolving factor analysis*, EFA).³⁹ Aquest mètode permet d'obtenir una estimació inicial de les finestres en què cada component o espècie química existeix. EFA es basa en el càlcul dels valors propis (o dels valors singulars que són les seves arrels quadrades) de totes les submatrius construïdes a partir de la matriu original de dades seguint un dels dos ordres d'aquesta. Generalment és més convenient de realitzar EFA en la direcció o l'ordre de la matriu que se sospita que té més selectivitat o resolució. L'evolució dels valors propis de totes les submatrius endavant i endarrere (vegeu apèndix 5.2) dona una descripció del que ocorre en l'experiment. En situacions favorables, a partir de les representacions EFA és possible la detecció de: *a*) regions on només és present una espècie química; *b*) regions on una espècie no és present; i *c*) regions on diverses espècies se superposen. L'avantatge d'aquest mètode és que, a més, proporciona una estimació inicial dels perfils de concentració de les diverses espècies presents en el sistema (o de les seves contribucions espectrals si es realitza en la direcció espectral). Aquestes estimacions es troben a partir del creuament adequat dels valors propis més significatius (o dels valors singulars més significatius que són les arrels quadrades d'aquells) en realitzar l'anàlisi endavant i endarrere (vegeu referències 39, 40, 54, 55, 56 per a una descripció més detallada del mètode). En les zones de rang molt alt, les línies que descriuen l'evolució dels factors poden no estar unívocament relacionades amb les diferents espècies químiques, i per tant les finestres o regions d'existència deduïdes poden ésser errònies. En aquests casos cal certa habilitat en la interpretació de les representacions EFA.³⁹ Un problema addicional es presenta quan la formació successiva de les diferents espècies químiques no segueix un ordre escalonat (*stepwise formation*).

A partir de qualsevol de les aproximacions descrites (detecció del rang local, de les variables més pures o de l'anàlisi de l'evolució dels factors), s'aconsegueixen estimacions inicials dels perfils i de les finestres d'existència de les diferents espècies en un dels dos ordres. A més, a partir d'aquest estudi inicial és possible de conèixer la complexitat del sistema que ha estat motiu d'estudi. Les tècniques descrites poden utilitzar-se sense cap coneixement previ del sistema, i la informació que proporcionen és molt útil per a la resolució correcta de les espècies.

Una forma ja coneguda d'augmentar la resolució i la selectivitat d'aquells sistemes amb molt baixa resolució i selectivitat és mitjançant la utilització de les derivades dels espectres originals. Per exemple, en els màxims o mínims dels espectres unitaris originals, l'espectre derivat s'anul·la i permet de tenir més selectivitat per a altres components superposats. Aquest concepte ha estat emprat per a augmentar el poder de resolució de pics totalment amagats i no resolts en cromatografia amb detecció espectroscòpica.⁶³ Més endavant es mostrarà un exemple d'ús de les derivades en l'augment de resolució en l'anàlisi multivariant de les dades provinents d'un procés industrial de producció.⁵⁷

2.1.6. COMPLEXITAT DEL PROBLEMA A RESOLDRE MITJANÇANT EL MÈTODE DE RESOLUCIÓ MULTIVARIANT DE CORBES

Cada espècie química ve descrita per una díada de vectors o de perfils: els perfils espectrals i els de variació de la concentració. Com que la finalitat del mètode de resolució multivariant de corbes és la recuperació d'aquests perfils, la complexitat del problema a resoldre es descriu mirant les similituds o correlacions entre els perfils de les diferents espècies. Quan els perfils espectrals i de concentració de diferents espècies són molt similars, és difícil resoldre'ls i recuperar-los. Com més similars siguin, més difícil en serà la resolució. La similitud entre dos perfils es pot determinar a partir del cosinus de l'angle entre els dos vectors que descriuen els dos perfils. Així la similitud entre dos perfils,

$$\text{similitud } (x, y) = \cos (\alpha) \quad [11]$$

$$\cos (\alpha) = x y^T / \text{sqrt } [(x x^T) (y y^T)] \quad [12]$$

on x i y són els dos vectors corresponents als dos perfils que es comparen, α és l'angle que formen aquests dos vectors en un espai multivariante i sqrt es refereix a l'arrel quadrada. Perfils molt similars donen valors de cosinus pròxims a la unitat, mentre que perfils molt diferents donen valors de cosinus pròxims a zero.

Ja ens hem referit a la selectivitat com a concepte fonamental per a assolir una resolució sense ambigüitats. Quan es té selectivitat en un o tots dos ordres, les possibilitats d'assolir la resolució correcta s'incrementen considerablement. Per a definir la complexitat d'un cert conjunt de dades des del punt de vista de la seva resolució és necessari conèixer prèviament la possible presència o absència de selectivitat utilitzant algun dels mètodes d'anàlisi local del rang descrits anteriorment. És interessant aconseguir de trobar almenys una variable selectiva per a cadascun dels components en la barreja.

Conceptes més coneguts, com el de resolució cromatogràfica o el de senyal net d'analit (*net analyte signal*),⁶² no inclouen explícitament la informació selectiva, i, encara que són útils, no són tan rellevants en el context del mètode de resolució multivariant de corbes.

2.1.7. MÈTODE DE MÍNIMS QUADRATS ALTERNATS (ALS) I RESTRINGITS D'UNA ÚNICA MATRIU DE DADES.

L'equació [1] es resol de manera iterativa utilitzant el mètode dels mínims quadrats alternats. A cada iteració de l'optimització una nova estimació de la matriu d'espectres S i de la de concentracions C s'obté de manera successiva utilitzant les dues equacions:

$$S = C^* D^* \quad [13]$$

$$C = D^* S^* \quad [14]$$

on la matriu C^* és la matriu pseudoinversa⁶⁴ de la matriu C i la matriu S^* és la matriu pseudoinversa de la matriu S . Òbviament la selecció del nombre correcte de components principals en el càlcul de la matriu D^* és essencial. Emprar la matriu reproduïda per components principals (equació [3]) D^* en lloc de la matriu experimental D , proporciona una major estabilitat en els càlculs, ja que la matriu D^* té part del soroll filtrat (*noise-filtered*) pel nombre escollit de components principals. Generalment, l'optimització iterativa de les matrius S i C a partir de les equacions [13] i [14] comença en l'ordre proposat quan es disposa d'una estimació inicial de la matriu C , obtinguda, per exemple, a partir de l'anàlisi de l'evolució dels factors. Alternativament, en lloc de començar amb l'equació [11] és possible començar amb l'equació [12] si es disposa d'una bona estimació inicial de la matriu S (els espectres de les espècies pures). Aquesta via alternativa és recomanable per a aquells casos en què els perfils de concentració es troben molt sobreposats mentre que els espectres dels components purs són molt diferents amb bandes espectrals característiques. En qualsevol dels dos casos, en cada cicle iteratiu s'aplicaran les següents restriccions:

a) Restricció de no-negativitat

Aquesta és la restricció més general en l'optimització per mínims quadrats. Les concentracions de les espècies químiques han d'ésser sempre valors positius

o zero. Els valors espectrals són generalment no negatius llevat d'alguns casos com l'espectroscòpia de dicroisme circular o les espectroscòpies derivades (com la ressonància d'espín electrònic) que tenen valors positius i negatius. La utilització d'aquesta restricció ha estat àmpliament recollida a la bibliografia.^{34, 35, 37-40} Aquesta restricció es pot dur a terme forçant les solucions de les equacions [13] i [14] a ésser no negatives utilitzant algorismes de mínims quadrats no negatius^{65, 66} o simplement forçant els valors negatius a ésser zero en cada iteració. De tota manera, àdhuc pels casos més simples de només dues espècies sobreposades, les solucions obtingudes sota només la restricció de no-negativitat no són úniques i existeix una banda de possibles solucions.⁴³⁻⁴⁵

b) Restricció de selectivitat i de regions de concentració negligible

Anteriorment ens hem referit a com es poden investigar la selectivitat i el rang local en les etapes preliminars de l'anàlisi de les dades. Una vegada les zones selectives i de concentració negligible per a cada espècie són conegudes s'han de fer explícites en l'optimització minimoquadràtica per evitar, així, la convergència cap a mínims locals no correctes. En el mètode proposat, aquesta restricció s'ha fet explícita forçant a zero la concentració d'una espècie en aquelles regions on se sap que no existeix. Les regions selectives queden definides com aquelles en què només una de les columnes en l'estimació inicial de C té valors no zero. Les zones de concentració negligible per a una espècie determinada tenen els valors corresponents a zero. Quan la selectivitat es troba present en els espectres S se segueix el mateix procediment, però, ara, amb els valors corresponents a les files de S . Durant l'optimització minimoquadràtica el recàlcul de la matriu C amb l'equació [12], s'utilitza una variable opcional (*flag*) per restringir a zero els valors de les concentracions en les mateixes regions que eren zero inicialment. Quan la selectivitat s'ha detectat correctament les ambigüitats rotacionals queden resoltes amb l'esmentat procediment.

c) Restricció d'unimodalitat

La restricció d'unimodalitat es pot emprar per a aquells perfils que es pot suposar que només tenen un màxim o pic. Aquesta situació és força corrent per a perfils de concentració i no tan corrent per a perfils espectrals. Aquesta restricció es pot posar de diferents formes, i en el present treball se'n proposa una de relativament senzilla. Primer es troba el màxim del perfil. Tots els valors a un costat i a l'altre d'aquest màxim han de disminuir de forma monò-

tona i no incrementar localment. En certs casos és permesa una desviació del comportament estrictament monotònic, com per exemple: *a)* quan els perfils de concentració recuperats en *C* mostren un cert soroll associat (pujades i baixades abruptes localment); *b)* quan les regions d'existència estimades inicialment per a les diferents espècies no són gaire acurades a causa del nombre gran d'espècies sobreposades (vegeu algorisme emprat en l'apèndix 5.7).

d) Restricció de sistemes tancats (closure)

Es poden emprar restriccions addicionals de normalització, les quals tenen un efecte important sobre l'ambigüitat d'intensitat o escala. Una de les formes més comunes de normalització és la dels sistemes tancats on la quantitat total dels constituents és constant. Per exemple, els sistemes en reacció en recipients tancats com són en les valoracions espectroscòpiques, la quantitat total dels constituents en reacció és constant i es troba distribuïda en diferents espècies químiques que evolucionen durant l'experiment (vegeu l'exemple d'aplicació a les valoracions espectroscòpiques més avall). Mitjançant aquesta restricció, la suma de concentracions de totes les espècies que tenen el mateix constituent químic en cada punt de la reacció o procés és constant.

Convergència

Cal continuar el procediment d'optimització iterativa restringida descrit fins que s'assoleix convergència o fins que s'arriba a un nombre preseleccionat d'iteracions. La convergència es defineix en termes de canvi relatiu de la funció desviació estàndard dels residuals entre dos cicles consecutius. Generalment, el procediment descrit és ràpid i s'assoleix la convergència (canvi relatiu de menys del 0,1 %) en menys de 50 cicles de l'optimització. Errors d'ajustament de les dades provinents d'una sola matriu per sota de l'1 % són desitjats.

2.1.8. MÈTODE DE MÍNIMS QUADRATS ALTERNATS (ALS) I RESTRINGITS APLICATS SIMULTÀNIAMENT A MÚLTIPLES MÀTRIS DE DADES

El mètode descrit anteriorment en el punt 2.1.7 pot estendre's fàcilment a l'anàlisi simultània de diverses matrius de dades. Suposem que tenim *K*

matrius obtingudes en condicions inicials diferents (p. e., a diferents concentracions inicials dels constituents químics). S'obtenen un conjunt K de matrius D_k de dades:

$$D_k = C_k S + E_k, \quad k = 1, 2, \dots, K \quad [15]$$

on ara C_k és la matriu dels perfils de concentració de les espècies químiques espectralment actives que es formen en la matriu D_k , S és la matriu dels espectres purs d'aquestes espècies i E_k la matriu d'error experimental. Com que les columnes (longituds d'ona) de les matrius D_k són comunes, l'anàlisi es pot realitzar simultàniament sobre més d'una de les matrius de dades de la següent manera:

$$D = \begin{bmatrix} D_1 \\ D_2 \\ \dots \\ D_K \end{bmatrix} = \begin{bmatrix} C_1 \\ C_2 \\ \dots \\ C_K \end{bmatrix} S + E \quad [16]$$

o més simplement

$$D = [D_1; D_2; \dots; D_K] = [C_1; C_2; \dots; C_K] S + E = C S + E$$

La nova matriu augmentada de dades D s'obté posant una sobre l'altra cadascuna de les matrius D_k a analitzar, amb les columnes (longituds d'ona) en comú. El nombre de files resultant és igual al nombre total d'espectres experimentals adquirits en les diferents matrius D_k . En el cas de l'anàlisi d'una sola matriu de dades, aquesta equació és la mateixa que l'equació [1] de l'apartat anterior. En el cas de l'anàlisi simultània de diverses matrius de dades, la nova matriu augmentada serà el producte d'una matriu augmentada de concentracions per la matriu d'espectres unitaris. La matriu augmentada de concentracions conté les diverses submatrius C_k corresponents a la concentració de les diverses espècies químiques presents en cadascuna de les matrius D_k analitzades. Aquesta disposició de les dades no fa cap assumptió sobre com són els perfils de variació de concentració de cadascuna de les espècies presents en les diferents matrius analitzades. El mètode proposat no força les dades a tenir una estructura completa de segon ordre trilineal.⁴⁶ En cas que

aquest tipus d'estructura hi sigui present, el mètode permetrà aprofitar l'estructura de segon ordre trilineal mitjançant una restricció addicional que forçarà els perfils de concentració a tenir una forma igual en les diferents matrius C_k . En qualsevol cas, però, els espectres de les espècies comunes en les diferents matrius analitzades són iguals i descrits en la matriu S . Aquesta assumptió serà certa sempre que les condicions físiques externes, tals com la temperatura o la composició del solvent, romanguin constants per a totes les matrius analitzades.

Com ja hem explicat, l'ambigüitat rotacional es pot resoldre per a aquells components o espècies químiques on existeixi selectivitat en almenys un dels dos ordres de la matriu de dades. El mateix és vàlid per a la matriu augmentada D , la qual cosa vol dir que l'ambigüitat rotacional pot resoldre's per a aquelles espècies en què en una o més de les matrius de dades incloses existeix una regió selectiva. De nou, la presència d'aquesta selectivitat ha d'ésser detectada i utilitzada conseqüentment. En el cas de la no existència de regions selectives per a algunes o totes les espècies presents, l'anàlisi simultània de les diferents matrius encara és millor que l'anàlisi individual de les mateixes matrius, ja que forçarà els espectres de les espècies comunes a ésser iguals en les diferents matrius de dades. El nombre de solucions possibles de l'equació [15] en l'anàlisi simultani està fortament restringit respecte a l'anàlisi individual emprant l'equació [14] per a cada matriu per separat. La naturalesa interna de les matrius D_k incloses en la matriu D afectarà els resultats obtinguts. Un bon disseny experimental millorarà els resultats assolits.

Pel que fa a l'anàlisi quantitativa, l'anàlisi simultània de diverses matrius correlacionades de dades resol l'ambigüitat d'intensitat en termes relatius. Les diferents matrius C_k contenen les contribucions relatives (concentracions) de cada espècie en aquelles matrius. Quan les dades no tenen una estructura de segon ordre trilineal (les formes dels perfils de concentració de la mateixa espècie difereixen en les diferents matrius que la contenen), la informació quantitativa és determinada a partir del quocient entre les àrees dels perfils de concentració dels components o espècies comunes en les diferents matrius de concentració C_k . El mètode pren arbitràriament l'àrea dels perfils de concentració obtinguts per a la primera matriu C_1 com a unitat de comparació. Totes les altres àrees queden referides a aquesta àrea unitària de la primera matriu. En canvi, quan les dades tenen una estructura de segon ordre trilineal, s'utilitza una nova restricció que permet el càlcul directe de les contribucions relatives de cada espècie en les diferents matrius de dades (vegeu més avall, restricció ii). En qualsevol dels dos casos, la recuperació de la informació quantitativa és una conseqüència directa de l'anàlisi simultània de diverses matrius de dades del mateix sistema.

Procediment

Primerament es prepara una estimació inicial de la matriu augmentada de concentracions C a partir d'estimacions inicials de les submatrius $C_1, C_2, \dots, C_K$ obtingudes, per exemple, a partir de l'anàlisi de l'evolució dels factors de les submatrius $D_1, D_2, \dots, D_K$ tal com ja es va descriure. L'estimació de la matriu C s'obté simplement per col·locació de les diverses submatrius C_k una sobre l'altra tal com es mostra en l'equació [16]. La informació sobre la presència de regions selectives i de regions de concentració negligible s'introdueix fent zero els valors corresponents a les regions on se sap que una determinada espècie no existeix. A cada iteració de l'optimització minimoquadràtica s'obté una nova estimació de les matrius S i C (equacions [13] i [14]). Com a l'anàlisi individual, també és possible començar l'optimització a partir d'una estimació inicial de la matriu espectral S , per exemple a partir de la detecció de les variables (longituds d'ona) més pures.³⁵ S'utilitza el mateix procediment que abans, però ara s'aplica a la matriu augmentada D' , reproduïda per components principals. Les restriccions emprades són no-negativitat, unimodalitat, selectivitat i normalització per *closure* descrites anteriorment. A més, a l'anàlisi simultània s'apliquen de forma opcional les següents restriccions per a cada iteració:

I) *Les espècies comunes tenen el mateix espectre en totes les matrius de dades*

Els espectres de les espècies comunes són forçats a tenir el mateix espectre en la matriu S . Aquesta restricció té un efecte important en la resolució dels perfils de concentració i dels espectres unitaris, sobretot en termes quantitatius relatius. Les ambigüitats d'escala i d'intensitat poden resoldre's amb aquesta restricció.

II) *Les espècies comunes tenen el mateix perfil de variació de la concentració en totes les matrius de dades*

La forma dels perfils de concentració d'una espècie determinada en les diferents matrius de dades pot forçar-se perquè sigui igual per a totes elles. En aquest cas, els perfils de concentració de les espècies comunes diferiran només en intensitat i no en forma, la qual cosa permet l'extracció de la informació quantitativa a partir de les seves intensitats relatives. El mètode es basa en les següents equacions (l'algorisme es troba descrit a l'apèndix 5.6) descrites gràficament a la figura 3.

C^n és la columna n de la matriu augmentada C que conté l'estimació actual dels perfils de concentració de l'espècie n en les diferents matrius K . Aquesta columna o vector llarg es pot doblar (*folded*) en una nova matriu C' (I, K), on I és el nombre d'espectres adquirits en cada matriu. S'assumeix que totes les matrius D_k tenen el mateix nombre de files i que existeix sincronització entre les diferents matrius de dades. Sota la suposició d'igual forma, aquesta matriu doblada pot descompondre's en el producte de dos vectors:

$$C = u' t^T \quad [17]$$

on $u' (I)$ és el perfil unitari que conté la informació sobre la forma dels perfils en C , i $t^T(K)$ conté les quantitats relatives d'aquest perfil unitari en les diferents matrius de dades. u' i t^T es calculen fàcilment a partir de la descomposició (primer valor singular) de la matriu C'

$$C' = u_1 s_{11} v_1^T + E \quad [18]$$

$$u' = u_1 \quad [19]$$

$$t^T = s_{11} v_1^T \quad [20]$$

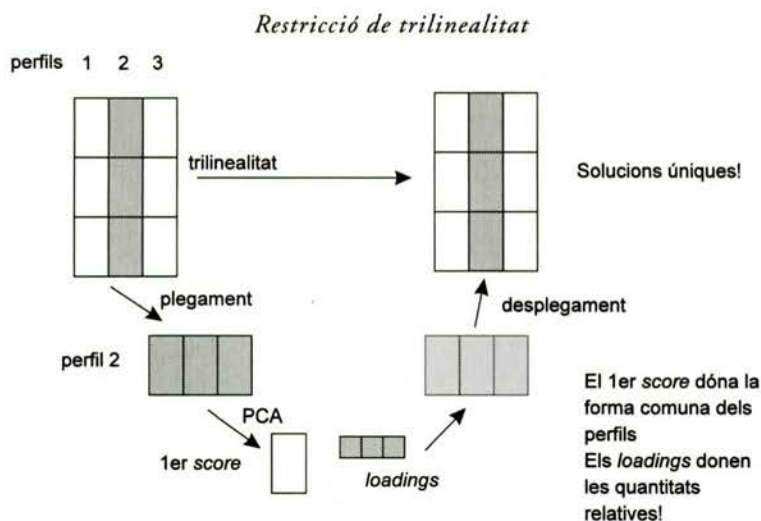


FIGURA 3. Restricció d'igualtat de forma en els perfils de concentració durant l'optimització minimoquadràtica de dades amb estructura de segon ordre trilineal.

El perfil unitari u' és igual al primer *score* u_1 i la concentració relativa de l'espècie es troba a partir del primer *loading* v_1^T escalat pel seu valor singular s_{11} . Aquest mètode assumeix que els perfils de concentració de la mateixa espècie tenen la mateixa forma en les diferents matrius de dades i, consegüentment, que la matriu C' és de rang unitari i E és al nivell de l'error experimental. Una manera senzilla de comprovar aquesta assumpció durant l'anàlisi és comparant la magnitud del segon valor singular següent respecte a la del primer valor singular.

III) *components amb concentració zero*

Quan se sap que una espècie no existeix en una determinada matriu de dades D_i , la concentració d'aquesta espècie en C es força perquè sigui zero. En l'optimització iterativa, les concentracions zero poden forçar-se a romandre zero. La qüestió sobre la presència o l'absència d'una determinada espècie en una determinada matriu de dades es pot resoldre a partir de l'anàlisi individual d'aquesta matriu de dades i de l'anàlisi local del rang.

IV) *correspondència entre espècies*

Amb la restricció anterior hi ha relacionada la restricció de correspondència entre espècies en les diferents matrius. Quan els espectres de dues espècies en diferents matrius de dades són molt similars (*fingerprinting matching*), aleshores, probablement, corresponen a la mateixa espècie, i així es postularà en l'anàlisi simultània. Contràriament, aquelles espècies amb espectres diferents en les diferents matrius de dades corresponen a espècies diferents en l'anàlisi simultània. Aquesta correspondència s'inclou en l'estimació inicial de la matriu C . Quan totes les espècies són comunes a totes les matrius de dades, les columnes de C tenen la concentració de les espècies ordenades de la mateixa manera en què es formen i d'acord amb les files dels espectres unitaris en S per a les mateixes espècies. Això vol dir que la primera columna de C conté la concentració de la primera espècie definida per l'espectre en la primera fila de S , la segona columna de C amb la segona fila de S , i així successivament per a totes les espècies formades. Les espècies comunes queden definides per les columnes comunes de les estimacions inicials de la matriu augmentada C .

L'anàlisi simultània de diverses matrius de dades utilitzant les restriccions descrites (especialment la restricció *ii*) porta al que s'anomena *l'avantatge*

de segon ordre,⁴⁶⁻⁴⁸ i les solucions obtingudes amb el mètode proposat són molt pròximes a les solucions no ambigües obtingudes mitjançant els mètodes de resolució de segon ordre.⁴⁶⁻⁴⁹ Quan la restricció d'igual forma de perfils de concentració no és present, l'avantatge de segon ordre es perd i són possibles solucions ambigües. En aquest cas la presència i la utilització de la selectivitat té encara un paper fonamental en la resolució. A l'apèndix 5.4 es dona l'algorisme complet del procediment descrit.

2.2. COMPROVACIÓ DEL MÈTODE

2.2.1. DESCRIPCIÓ DE LES DADES

A continuació, es mostren alguns exemples d'aplicació del mètode descrit. Com que el més important és comprovar els efectes de la selectivitat i de l'anàlisi simultània de diferents matrius de dades, s'obtenen en primer lloc un conjunt de dades simulades a partir de perfils espectrals i de concentració coneguts. D'aquesta manera, els resultats que s'obtenen per aplicació del mètode es poden comparar i entendre millor. Cada espècie química vindrà definida per una díade de dos vectors, un d'ells és el perfil de concentracions i l'altre és el perfil spectral. El producte extern vectorial (*outer product*) d'aquests dos vectors donarà la matriu resposta de l'espècie química considerada. Les matrius de dades mescla s'obtenen, aleshores, per addició d'aquestes matrius resposta corresponents a cada espècie. S'afegeix un error de tipus aleatori a les matrius mescla. A la taula 1 queda resumida la complexitat matemàtica de les dades simulades.

Exemple 1

El primer exemple escollit serveix fonamentalment per a estudiar l'efecte de la selectivitat en l'anàlisi individual d'una sola matriu de dades. Es tracta d'una reacció química simple entre dos components A i B, per a donar un tercer component AB, $A + B \rightarrow AB$. Només dues de les espècies, B i AB, són actives en la regió spectral sota estudi (A no absorbeix). S'obtenen espectres successius en diferents estats de la reacció entre A i B. El conjunt ordenat d'espectres forma la matriu de dades a analitzar. Els dos ordres de mesura són l'espectral i el de canvi de concentració dels components de la reacció (per exemple, per canvi del pH). A les figures 3, 4 i 5 es mostren els perfils espec-

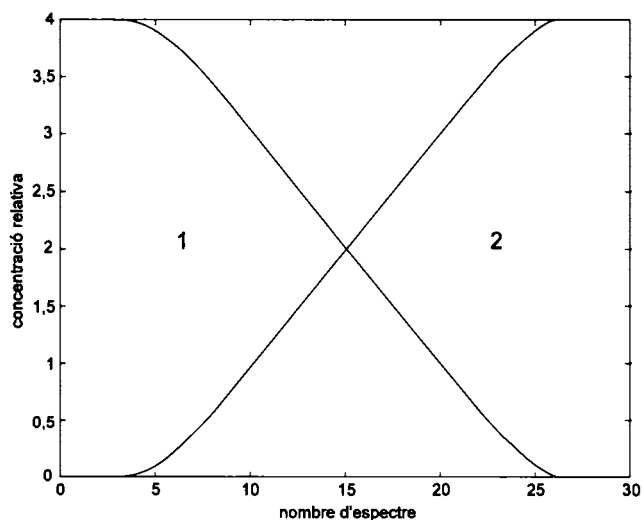


FIGURA 4. Perfils de concentració emprats en la simulació de les matrius de dades 1A i 1B. S'han emprat, també, parts no selectives d'aquests perfils en la simulació de les matrius de dades 1C, 1D, 1E i 1F (vegeu descripció de les dades).

trals i de concentració utilitzats en les simulacions. A la taula 1 es mostren les similituds (correlacions) entre els perfils espectrals i de concentració de les diferents espècies del sistema. Es prenen diferents subgrups de dades a partir del conjunt global de dades, per tal d'estudiar el problema de la selectivitat. En tots els casos es va afegir soroll de mitjana 0,0 i desviació estàndard 0,001 unitats. Es generen 6 matrius diferents de dades (taula 1). A la matriu 1A hi ha selectivitat per a les dues espècies en tots dos ordres (figures 4 i 5). A la matriu 1B hi ha selectivitat en els perfils de concentració de les dues espècies (els mateixos que en 1A) però no n'hi ha en els perfils espectrals (figura 6). A la matriu 1C la selectivitat es troba en els perfils espectrals (figura 5); els perfils de concentració s'obtenen a partir dels perfils de concentració anteriors, però en una zona on no hi ha selectivitat (zona de concentracions en la figura 4 de B/AB des de 3/1 fins a 1/3 o 75 %/25 % fins a 25 %/75 %). La matriu 1D s'obté com la 1B, però amb selectivitat només en el primer component o espècie (rang de concentracions B/AB de 4/0 fins a 1/3 o 100 %/0 % fins a 25 %/75 %). Finalment, en les matrius 1E i 1F no hi ha selectivitat en cap dels dos ordres; els espectres de la figura 6 i els perfils de concentració són els de la matriu 1C.

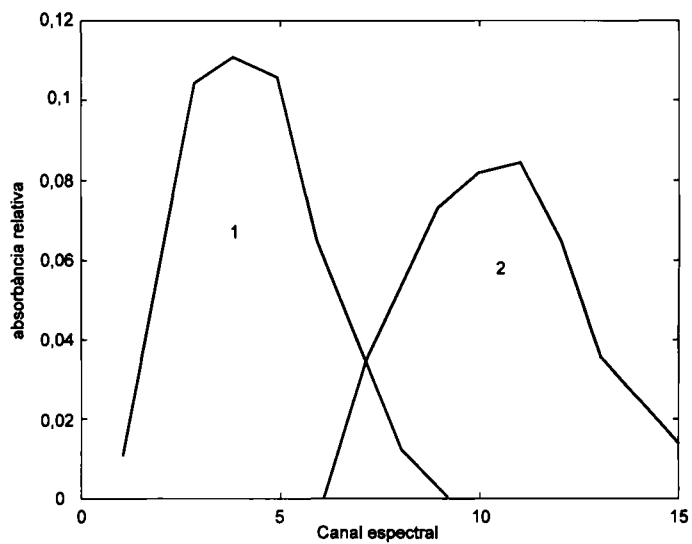


FIGURA 5. Perfils espectrals emprats en la simulació de les matrius 1A i 1C.

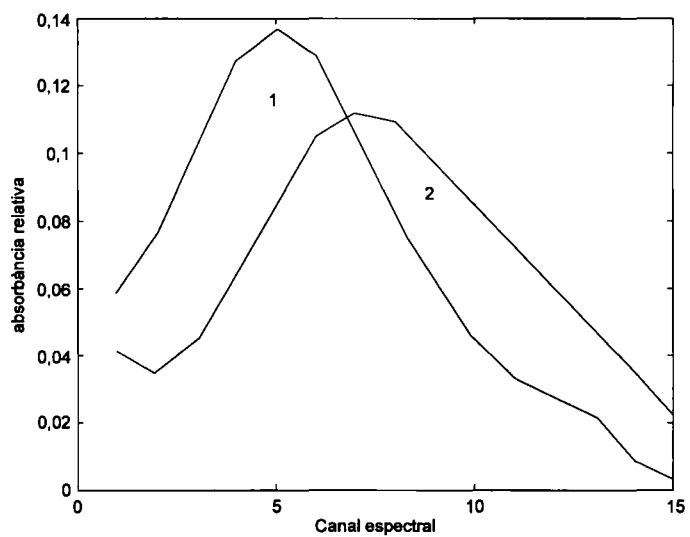


FIGURA 6. Perfils espectrals emprats en la simulació de les matrius de dades 1B, 1D, 1E, 1F, 2C, 2D, 2E i 2F.

TAULA 1
*Descripció de la complexitat de les dades simulades:
 correlació (similitud) entre els perfils de les diferents espècies i presència o no de selectivitat*

Matriu ^a	Cos(c) ^b			Cos(s) ^c espectres			Selectivitat	
	1-2	1-3	2-3	1-2	1-3	2-3	conc ^d	espectra ^e
1A	0,291			0,048			1,2	1,2
1B	0,291			0,888			1,2	n
1C	0,832			0,048			n	1,2
1D	0,627			0,888			1	n
1E,1F	0,832			0,888			n	n
2A	0,435			0,618			1	n
2B	0,867			0,618			n	n
2C, 2D, 2E i 2F	0,867			0,888			n	n
3A	0,322	0,138	0,903	0,995	0,946	0,961	1	n
3B	0,267	0,022	0,683	0,995	0,946	0,961	1	n
3C	0,324	0,007	0,266	0,995	0,946	0,961	1,3	n

a. Matriu de dades experimental. Mateixa notació que en la secció de dades.

b. Correlació o similitud entre les diferents espècies mesurada com el cosinus de l'angle entre els perfils de concentració de les diferents espècies, el nombre de les quals s'indica a la línia següent.

c. Correlació o similitud entre les diferents espècies mesurada com el cosinus de l'angle entre els perfils espectrals de les diferents espècies, el nombre de les quals s'indica a la línia següent.

d. Selectivitat en el perfil de concentració. Els nombres indiquen quines espècies tenen selectivitat. «n», indica que cap perfil de concentració té selectivitat.

e. Selectivitat en el perfil d'espectres. Els nombres indiquen quines espècies tenen selectivitat. «n», indica que cap perfil spectral té selectivitat.

Exemple 2

Aquest exemple serveix per a investigar i comprovar la utilització del mètode de resolució multivariant de corbes en l'anàlisi simultània de diverses matrius de dades amb estructura de segon ordre trilineal. Les dades simulades corresponen a cromatografia líquida amb detecció espectroscòpica. En aquest cas, els dos ordres de mesura són el temps d'elució i la longitud d'ona dels espectres adquirits. S'estudien diferents casos amb diferents nivells de complexitat (taula 1). En tots els casos, s'afegí en cadascuna de les matrius simulades soroll o error aleatori de mitjana 0 i desviació estàndard entre 0,1 i 1 % del màxim d'absorció. El primer exemple estudiat, la matriu 2A, és una barreja de dos pesticides, el carbamat pirimicarb i l'1-naftol. Detalls experimentals

poc importants en la discussió que segueix foren ja publicats en 78 i 79. A les figures 7 i 8 es donen els espectres i els perfils d'elució dels dos components. Pel primer component (pirimicarb) hi ha una certa selectivitat al principi de la seva elució. La matriu 2B s'obté com la 2A, però disminuint la resolució cromatogràfica fins a una superposició completa i absència de selectivitat en l'ordre cromatogràfic (figura 9). Les matrius 2C i 2D s'obtenen amb els mateixos perfils d'elució que la 2B, però amb espectres molt més sobreposats (com els de les matrius 1B, 1D, 1E i 1F de la figura 4). Les matrius 2C i 2D difereixen en les proporcions relatives dels dos components. A més es construeixen les matrius 2C1 i 2C2 que contenen, respectivament, només el primer i el segon component de la matriu 2C, però a una concentració de la meitat. S'utilitzen com a estàndards en les determinacions quantitatives. Finalment, les matrius 2E i 2F són també similars a la matriu 2C, però un dels dos components té una concentració deu vegades més petita. S'utilitzen per a comprovar el funcionament del mètode proposat a baixes concentracions.

Exemple 3

S'escull aquest exemple per a investigar la utilització del mètode de resolució multivariant de corbes en l'anàlisi simultània de matrius de dades que

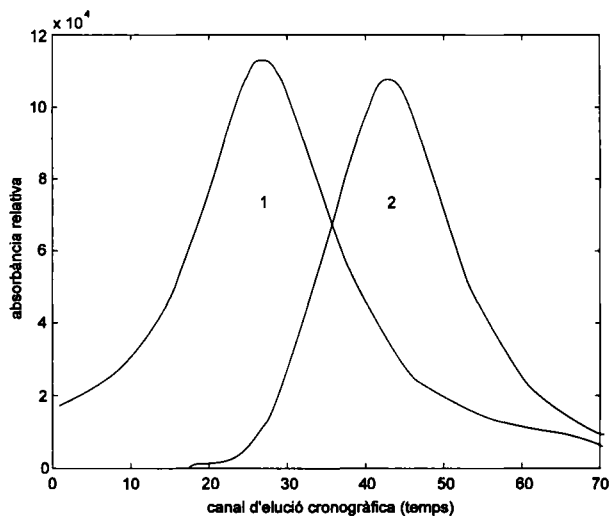


FIGURA 7. Perfils d'elució cromatogràfica emprats en la simulació de la matriu de dades 2A.

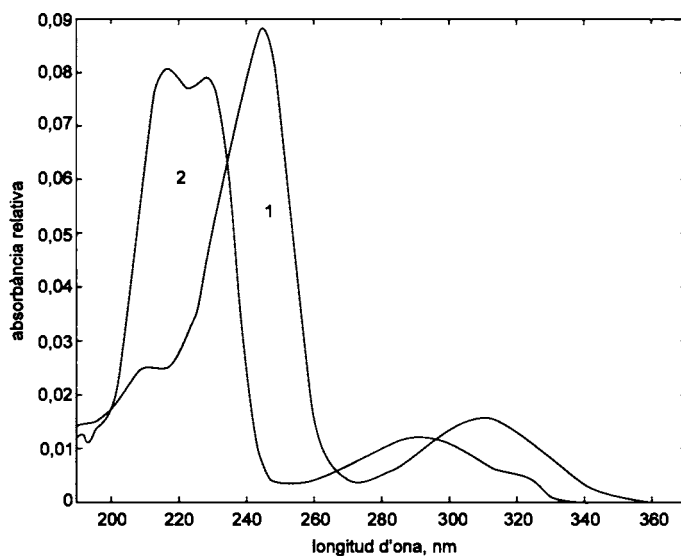


FIGURA 8. Perfils espectrals emprats en la simulació de les matrius de dades 2A i 2B.

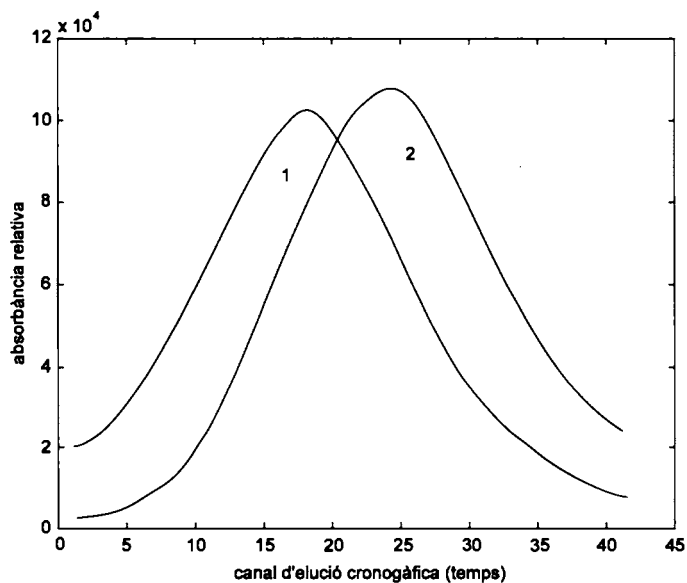


FIGURA 9. Perfils de concentració emprats en la simulació de les matrius de dades 2B, 2C, 2D, 2E i 2F.

no tenen una estructura completa de segon ordre trilineal. S'estudia el sistema de tres valoracions espectroscòpiques d'un sistema multiequilibri de tres espècies que contenen Cu(II) en la seva complexació amb salicilat. Les tres valoracions s'obtenen a diferents relacions de concentracions dels constituents Cu(II) i salicilat, que són 1; 1,5 i 2. En tots els casos, la valoració es du a terme canviant el pH des de 3 fins a 9 en intervals de 0,2 unitats de pH. A cada pH es calcula un espectre que cobreix el rang espectral de 400 nm fins a 900 nm a intervals de 20 nm. Tres espècies diferents presenten absorció: el Cu(II) lliure, i els complexos Cu(sal) i Cu(sal)₂, on *sal* vol dir l'anió salicilat, i on s'han omès les càrregues. Altres espècies presents en el sistema que no presenten absorció de radiació són les espècies protonades del lligand, Hsal i H₂sal respectivament. A les referències 9-12 i 56 es donen més detalls i característiques d'aquest conjunt de dades. Cada valoració espectroscòpica dóna una matriu de dades, 3A, 3B i 3C, respectivament. A la taula 1 es mostren les similituds entre els perfils de concentració (diagrames de distribució) i entre els perfils o espectres unitaris de les tres espècies formades. Aquests tres espectres unitaris són molt semblants entre ells (figura 10) mentre que els perfils de concentració no ho són tant (figura 11), excepte per a la segona i tercera espècie en la primera valoració. En les tres valoracions hi ha selectivitat per al perfil de concentració de la primera espècie i només en la tercera valoració n'hi ha per a la tercera espècie. Per al perfil de concentració de la segona espècie no hi ha, en cap cas, selectivitat; com tampoc n'hi ha per a cap dels tres espectres unitaris. A les figures 10 i 11 es mostren els espectres i diagrames de distribució emprats en la simulació de les tres matrius de dades 3A, 3B i 3C.

2.2.2. DETECCIÓ DE LA SELECTIVITAT

A les figures 12 i 13, es mostren dos exemples d'aplicació del mètode FSWEFA en els dos ordres de la matriu de dades 1B amb una finestra de 3 elements. En l'ordre de les concentracions, s'hi detecten tres regions diferenciades de rang local: la primera i última regió són de rang químic igual a 1 i, entre elles dues, s'hi detecta una regió ampla de rang 2. Quan s'incrementa la mida de la finestra a 5 o més espectres, la detecció de les zones selectives (amb rang unitari) es perd, ja que ara en totes les finestres existeix més d'una espècie. L'anàlisi de les variables pures confirma que les més pures es troben al principi i al final de la reacció. Quan l'anàlisi es du a terme en la direcció espectral, no s'hi detecten zones de rang químic igual a 1. El rang local és sempre més gran d'1 ja que els dos espectres se superposen. Els resultats obtinguts en l'anàlisi del rang local de les altres matrius de l'exemple 1 estan d'acord amb la selectivitat

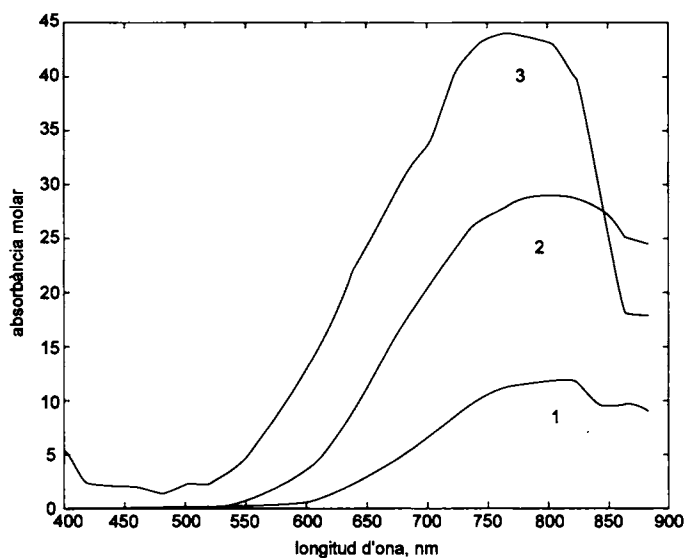


FIGURA 10. Perfils espectrals emprats en la simulació de les matrius de dades 3A, 3B i 3C.

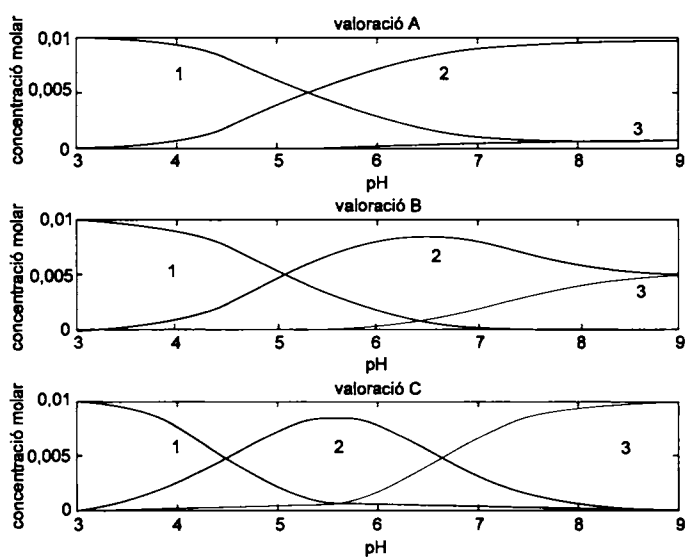


FIGURA 11. Distribució d'espècies (perfils de concentració) emprada en la simulació de les matrius de dades 3A, 3B i 3C.

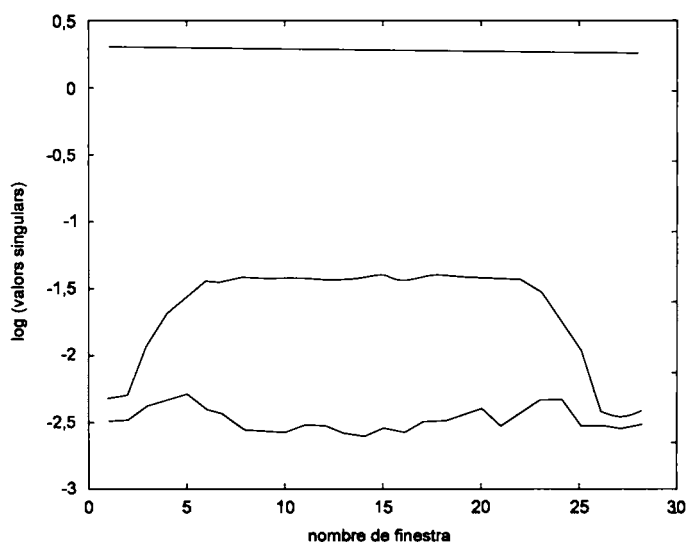


FIGURA 12. Detecció de la selectivitat en l'ordre del canvi de concentracions de la matriu 1B (figura 4) emprant el mètode FSMWEFA.⁶¹

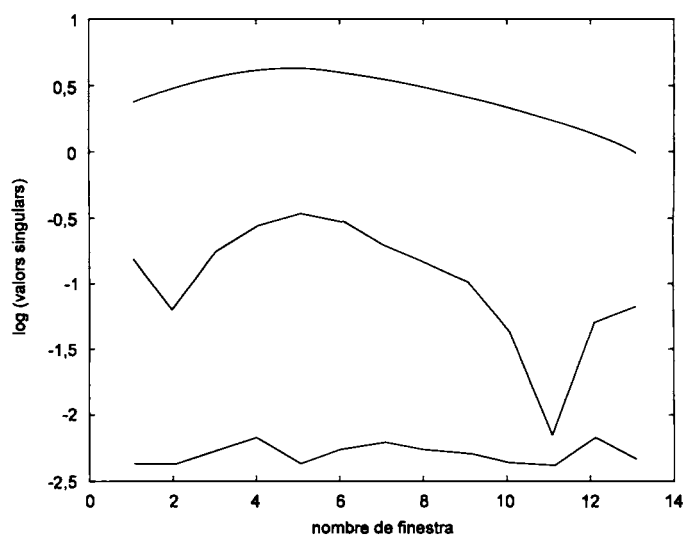


FIGURA 13. Detecció de la selectivitat en l'ordre espectral de la matriu 1B (Figura 6) emprant el mètode FSMWEFA.⁶¹

dels perfils emprats en la simulació. A la figura 14 es dóna un altre exemple de l'aplicació del mètode FSWWEFA per a l'anàlisi de la matriu de dades 2A en la direcció del temps utilitzant una finestra de 4 espectres. S'hi detecten regions de rang 1 i rang 2, que mostren les regions selectives i les no selectives. Per a les matrius 2C, 2D, 2E i 2F no es detecta selectivitat.

A la figures 15 i 16 s'aplica el mateix mètode FSWWEFA per a la matriu 3A en els dos ordres, pH i longitud d'ona. La finestra és de 4 elements. Es va investigar l'efecte del nivell del soroll o error aleatori. Per a un nivell de soroll amb una mitjana de 0 i de desviació estàndard de 0,001 unitats, només es detecta una regió de rang 1 en tot l'ordre de pH, mentre que en l'ordre espectral es detecten regions de rang 1, 2 i 3. Aquesta discordança es deu al fet que els perfils de concentració són més distingibles que els perfils espectrals i, per tant, la detecció del rang local serà millor en l'ordre espectral que en l'ordre de canvi de concentració o pH (vegeu a la taula 1 els valors de correlació entre els diferents perfils). Quan hom disminueix el nivell de soroll (per exemple fins a 0,000001 unitats de desviació estàndard, figura 17) la resolució de les regions de rang 1, 2 i 3 també és possible en l'ordre de pH.

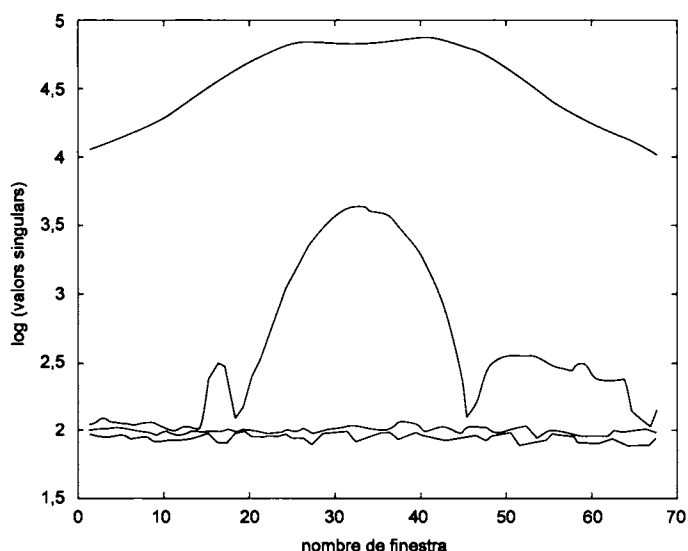


FIGURA 14. Detecció de la selectivitat en l'ordre del canvi de concentracions (elució) de la matriu 2A (figura 7) emprant el mètode FSMWEFA.⁶¹

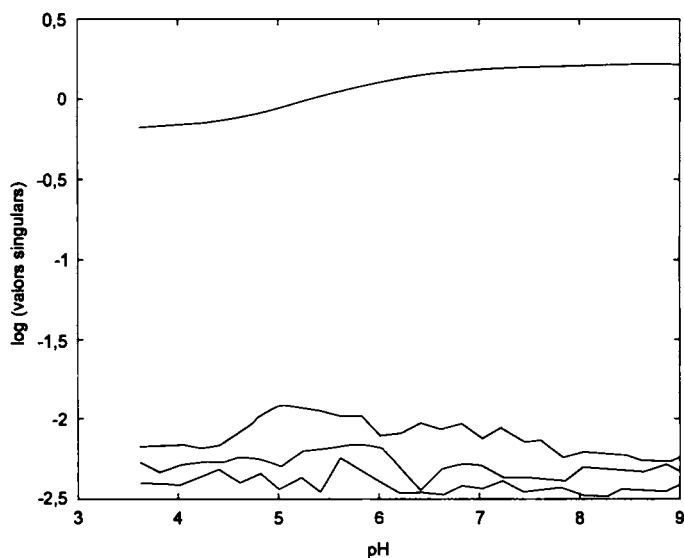


FIGURA 15. Detecció de la selectivitat en l'ordre del canvi de concentracions (pH) de la matriu 3A (figura 11) emprant el mètode FSMWEFA⁶¹ i un nivell de soroll de fons amb una desviació estàndard de 0,001 unitats.

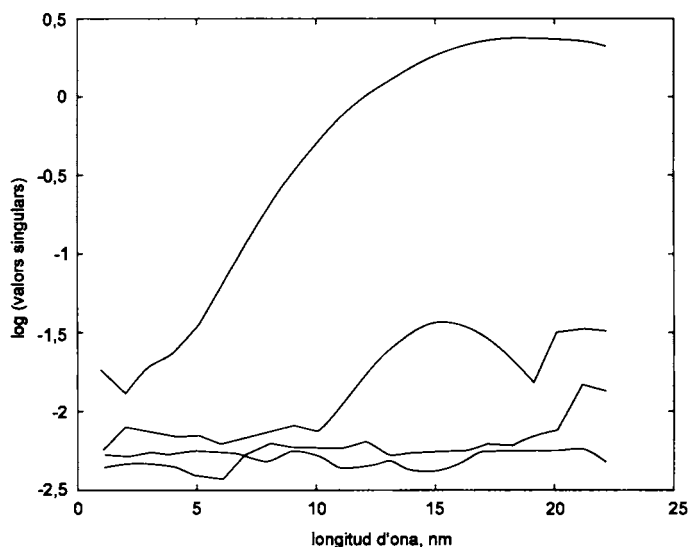


FIGURA 16. Detecció de la selectivitat en l'ordre espectral de la matriu 3A (figura 10) emprant el mètode FSMWEFA.⁶¹

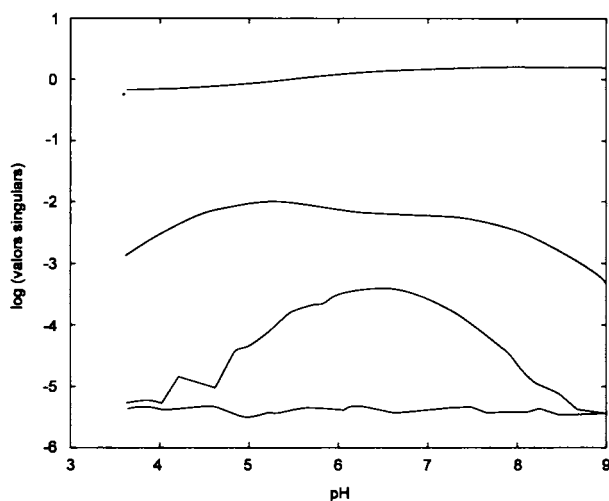


FIGURA 17. Detecció de la selectivitat en l'ordre del canvi de concentracions (pH) de la matriu 3a (figura 11) emprant el mètode FSMWEFA⁶¹ i un nivell de soroll de fons amb una desviació estàndard de 0,0000001 unitats.

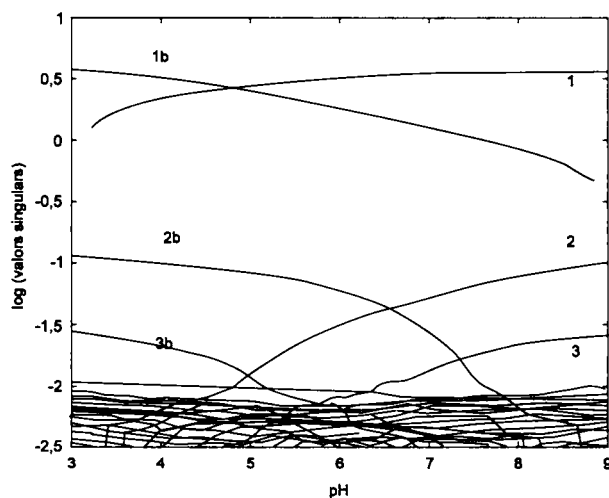


FIGURA 18. Anàlisi de l'evolució dels factors (EFA)³⁹ a la matriu 3A. 1, 2 i 3 són les línies corresponents al primer, el segon i el tercer factor, valors singulars de les submatrius successives obtingudes en augmentar el pH. 1b, 2b i 3b són les línies corresponents al primer, el segon i el tercer factor, valors singulars de les submatrius successives obtingudes en disminuir el pH (vegeu també les referències 54-56 per aplicacions recents de l'esmentat mètode). A l'apèndix 5.2 es dona el programa MATLAB per a fer l'anàlisi EFA.

Els resultats obtinguts per a les matrius 3A i 3B són similars: d'una banda, es confirma la presència de selectivitat per a la primera espècie en començar cadascuna de les tres valoracions i, de l'altra, es detecta una nova regió selectiva per a l'última espècie en els últims espectres de la matriu 3C.

Sigui quin vulgui l'error aleatori, el mètode d'evolució dels factors (EFA) detecta tres espècies per a la matriu 3A (figura 18). La detecció del rang total d'una matriu de dades d'un sistema amb perfils molt superposats és millor per al mètode EFA que per al mètode FSWEFA, ja que la mida de les submatrius en EFA és molt més gran i té molta més capacitat de detecció. A partir d'EFA s'obtenen, a més, les estimacions inicials dels perfils de concentració.

2.2.3. ANÀLISI D'UNA MATRIU DE DADES PEL MÈTODE DE MÍNIMS QUADRATS ALTERNATS (ALS)

Dades de l'exemple 1

L'anàlisi individual de les diferents matrius de dades de l'exemple 1 dona diferents situacions d'acord amb la selectivitat de les dades en cada sistema. A la taula 2 es mostra un resum dels resultats assolits per a cada cas.

MATRIU 1A

Hi ha selectivitat en tots dos ordres per a les dues espècies. Els perfils en tots dos ordres es troben poc correlacionats. No hi ha ambigüitat rotacional. A més, quan s'aplica la restricció de sistema tancat, desapareix l'ambigüitat d'intensitat.

MATRIU 1B

Hi ha selectivitat en els perfils de concentració, però no n'hi ha en els perfils espectrals. L'ambigüitat rotacional es pot trencar per aplicació del mètode d'anàlisi de l'evolució dels factors i de les restriccions de no-negativitat i unimodalitat. No hi ha tampoc ambigüitat d'intensitat.

MATRIU 1C

Hi ha selectivitat en els dos espectres de les dues espècies, però no n'hi ha en els perfils de concentració. Com en el cas anterior, es resol l'ambigüitat rotacional. La selectivitat present en l'ordre espectral s'ha de detectar i emprar

explícitament en les restriccions imposades. Si els espectres són unimodals, també s'aplica aquesta restricció, la qual té un efecte similar a la utilització de la selectivitat.

MATRIU 1D

Només hi ha selectivitat en el perfil de concentracions d'una de les dues espècies i no hi ha selectivitat espectral. Per a l'espècie amb selectivitat en el seu perfil de concentració es recupera correctament el seu perfil espectral. En canvi, el perfil de concentracions d'aquesta mateixa espècie no es recupera correctament en les regions no selectives, ja que en aquestes encara és possible una combinació lineal amb els altres perfils.

Quan per a una de les dues espècies hi ha selectivitat en els dos ordres de mesura, l'ambigüïtat de rotació es resol també per a l'altra espècie. Quan hi ha més de dues espècies, la subtracció de la contribució de l'espècie selectiva de la matriu original de dades dona una nova matriu de rang inferior en una unitat. L'anàlisi es pot començar de nou.

MATRIUS 1E i 1F

Quan no hi ha selectivitat en cap dels ordres i per a cap de les espècies del sistema, l'ambigüïtat rotacional persisteix; àdhuc després d'aplicar-hi les restriccions de no-negativitat, unimodalitat i de sistema tancat.

En cas d'haver-hi més espècies superposades (2, 3 o més), s'obtenen les mateixes conclusions donades fins ara, ja que, per tal de resoldre completament l'ambigüïtat rotacional ha d'existir selectivitat. L'estudi de sistemes amb més espècies es redueix a la detecció de les zones locals on és present un nombre més petit d'espècies. Així mateix, l'estudi d'aquestes regions de forma separada ajuda molt a la resolució de la mescla completa. Els sistemes multicomponent molt superposats poden resoldre's separatament per passos per descomposició de la matriu global de dades en submatrius més petites on poden ésser aplicades les conclusions anteriors. En l'exemple 3 es dona un sistema de tres components on s'arriba a conclusions similars.

Dades de l'exemple 2

L'anàlisi individual de les tres mescles cromatogràfiques dona resultats comparables amb els obtinguts en la secció anterior. Les restriccions emprades per a totes les matrius són de no-negativitat en tots dos ordres i d'unimodalitat en

TAULA 2
Correlació entre els valors teòrics i els valors recuperats pel mètode ALS dels perfils de concentració i dels perfils espectrals en l'anàlisi individual de les matrius de dades

Matriu ^a	Restriccions ^b	Cos(c) ^c			Cos(s) ^d		
		1	2	3	1	2	3
1A	c c y	1,000	1,000	—	1,000	1,000	—
1B	c c y	1,000	1,000	—	1,000	1,000	—
1C	s s y	1,000	1,000	—	1,000	1,000	—
1D	c c y	0,983	1,000	—	1,000	0,999	—
1E	c c y	0,972	0,969	—	0,994	0,992	—
2A	c c n	0,963	1,000	—	1,000	0,981	—
2B	c c n	0,958	0,997	—	0,995	0,973	—
2C	c c n	0,959	0,997	—	0,999	0,999	—
2D	c c n	0,960	0,998	—	0,999	0,987	—
2E	c c n	0,942	0,995	—	1,000	0,920	—
3A	c c y	0,980	0,606	0,993	0,999	0,999	0,968
3B	c c y	0,998	0,963	1,000	1,000	1,000	0,996
3C	c c y	0,998	0,968	1,000	1,000	1,000	0,999

a. Matriu de dades experimental. Mateixa notació que en la secció de dades.

b. Restriccions emprades durant l'optimització en el següent ordre: unimodalitat (c/s), estimacions inicials (c/s), sistema tancat (y/n). (c = concentracions, s = espectres, y = sí, n = no)

c. Correlació entre els perfils de concentració recuperats pel mètode ALS (calculats) i els perfils de concentració vertaders (emprats en la simulació): cosinus de l'angle entre tots dos.

d. Correlació entre els perfils espectrals recuperats pel mètode ALS i els perfils espectrals vertaders: cosinus de l'angle entre tots dos.

els perfils d'elució. Les estimacions inicials dels perfils d'elució s'obtenen amb EFA, i en aquest cas no s'aplica la restricció de sistema tancat. Com que només s'analitza una matriu de dades, les solucions recuperades no contenen informació quantitativa i són ambigües en intensitat.

MATRIU 2A

A causa de la selectivitat present en les dades (taula 1), el perfil d'elució del segon component i l'espectre pur del primer component es recuperen totalment (taula 2). Hi ha encara certa ambigüitat residual en els altres dos perfils, encara que són molt pròxims als vertaders.

MATRIU 2B

La selectivitat només és present per l'espectre pur del segon component. D'acord amb això, el perfil més ben recuperat és el perfil d'elució del segon component. Encara que la resolució cromatogràfica és molt petita (per sota de 0,1) la recuperació dels perfils és satisfactòria.

MATRIUS 2C i 2D

No hi ha selectivitat en cap dels dos ordres, la resolució és baixa i la similitud entre els perfils és alta. Àdhuc en aquesta situació tan desfavorable, els perfils recuperats són bastant bons. Encara hi roman una petita ambigüitat rotacional.

MATRIUS 2E i 2F

Els perfils d'aquestes dues matrius tenen les mateixes formes que els perfils de les matrius 2C i 2D. En aquest cas, la recuperació és molt millor per al perfil espectral del component majoritari que per al component minoritari, a causa, en part, de la selectivitat més gran del perfil d'elució (concentració) corresponent al component majoritari. Quan la concentració d'un dels components baixa de forma considerable, la selectivitat per al perfil d'elució del component majoritari pot créixer considerablement i, per tant, pot fer més fàcil la recuperació del seu espectre pur. En canvi, el component minoritari es troba totalment amagat i no resolt, per la qual cosa la recuperació del seu espectre pur i del seu perfil d'elució és sempre pitjor.

Dades de l'exemple 3

MATRIUS 3A, 3B i 3C

En aquest exemple els espectres de les tres espècies es troben totalment superposats (taula 1). Hi ha més selectivitat per als perfils de concentració, excepte en la primera valoració (figura 11).

A la taula 2 es donen els resultats de l'anàlisi individual de les matrius de dades 3A, 3B i 3C. Les restriccions utilitzades per a les tres matrius són no-negativitat en els dos ordres, unimodalitat en els perfils de concentració, sistema tancat, i selectivitat per a la primera espècie de cada valoració i per a l'última de la tercera valoració. En totes tres valoracions s'obté una bona correlació entre

els espectres calculats i els teòrics. En canvi, s'obtenen resultats més pobres en la recuperació dels perfils de concentració o diagrames de distribució d'espècies, especialment, per a la segona espècie de la primera valoració. En aquest cas, el procediment de mínims quadrats alternats no convergeix cap a la solució vertadera. La raó es troba en la manca de selectivitat d'aquesta espècie i la gran correlació que existeix en els dos ordres, pH i longitud d'ona (taula 1). L'ambigüitat rotacional no es pot resoldre en l'anàlisi individual de la valoració 3A. Una situació semblant, però no tan dramàtica, s'observa per a la valoració 3B. A la valoració 3C hi ha selectivitat en els perfils de concentració de les espècies 1 i 3. Encara que no hi ha selectivitat per a l'espècie 2, quan s'utilitza la restricció de selectivitat adequadament (per a les espècies 1 i 3), l'optimització convergeix cap a la solució vertadera, sigui quin vulgui el soroll de fons.

2.2.4. ANÀLISI SIMULTÀNIA DE DIVERSES MATRIUS DE DADES MITJANÇANT RESOLUCIÓ MULTIVARIANT DE CORBES

Cas I. Cap dels dos ordres és comú entre les matrius de dades

Quan cap dels dos ordres és comú entre les diferents matrius de dades, no hi ha cap avantatge en l'anàlisi simultània ja que no hi ha cap guany d'informació. Les ambigüitats de rotació i d'intensitat persistiran en la mateixa extensió que en l'estudi individual de cada matriu per separat.

Cas II. Dos dels ordres són comuns entre les matrius (dades amb estructura de segon ordre trilineal)

Dades de l'exemple 2

En aquest cas, sigui quina sigui la selectivitat del sistema, es recuperen els perfils de concentració i els espectres correctes. Els resultats obtinguts quan s'aplica la restricció d'igual forma dels perfils de concentració (equacions [17] i [20]) són molt similars als perfils obtinguts per utilització dels mètodes de resolució tensorial com GRAM,⁴⁷ descomposició trilineal⁴⁸ o altres.⁴⁹ Els resultats assolits es resumeixen a la taula 3.

TAULA 3

Correlació entre els valors teòrics i calculats dels perfils de concentració i dels perfils espectrals en l'anàlisi simultània de diverses matrius de dades pel mètode de mínims quadrats alternats (ALS)

Matriu ^a	Rest. ^b	Cos(c) ^c			Cos(s) ^d			Error quant. ^e
		1	2	3	1	2	3	
[1B; 1E]	<i>c c y y</i>	1,000	1,000	—	1,000	1,000	—	estàndard
		1,000	1,000	—	1,000	1,000	—	0,12 %, 0,15 %
[1E; 1F]	<i>c c y y</i>	0,971	0,969	—	0,994	0,992	—	estàndard
		0,971	0,969	—	0,994	0,992	—	0,15 %, 0,20 %
[2E; 2C1]	<i>c c n y</i>	1,000	0,999	—	1,000	0,999	—	0,3 %
								estàndard
[2E; 2C2]	<i>c c n y</i>	1,000	1,000	—	1,000	1,000	—	0,4 %
								estàndard
[2D; 2C2]	<i>c c n y</i>	1,000	1,000	—	1,000	1,000	—	0,2 %
								estàndard
[2D; 2C2]	<i>c c n n</i>	1,000	0,996	—	0,998	1,000	—	19,0 %
								estàndard
[2C; 2D]	<i>c c n y</i>	0,994	0,998	—	1,000	0,997	—	estàndard
		0,994	0,999	—				5,0 %, 2,0 %
[2C; 2D]	<i>c c n n</i>	0,956	0,998	—	0,999	0,991	—	estàndard
		0,980	0,996	—				15,0 %, 4,2 %
[2C; 2D; 2E; 2F]								
	<i>c c n y</i>	1,000	1,000	—	1,000	1,000	—	estàndard
		1,000	1,000	—				0,2 %, 0,04 %
		1,000	1,000	—				12,0 %, 0,01 %
		1,000	1,000	—				0,8 %, 2,0 %
[2C; 2D; 2E; 2F]								
	<i>c c n n</i>	1,000	1,000	—	0,9998	0,9997	—	estàndard
		1,000	1,000	—				1,3 %, 1,4 %
		0,891	1,000	—				52,0 %, 3,8 %
		1,000	0,952	—				5,0 %, 42,0 %
[3A; 3B; 3C]								
	<i>c c y n</i>	1,000	1,000	1,000	1,000	1,000	1,000	
[3A; 3B]	<i>c c y n</i>	0,999	0,999	1,000	1,000	1,000	1,000	
		0,999	0,913					

- a. Matrius augmentades segons les columnes construïdes a partir de les matrius experimentals individuals que s'indiquen (es conserva l'ordre definit per les columnes) per a l'anàlisi simultània de diversos experiments alhora (equació [16]). Mateixa notació que en la secció de dades. Per exemple, la matriu augmentada [2C; 2D; 2E; 2F] es construeix a partir de les matrius individuals 2C, 2D, 2E i 2F situant-les una al damunt de l'altra tot preservant l'ordre de les columnes (també vegeu l'esquema de la Figura 45, pàg. 90)
- b. Restriccions emprades en l'optimització en el següent ordre: unimodalitat (*c/s*), estimacions inicials (*c/s*), sistema tancat (*y/n*), igual forma en els perfils de concentració (o estructura de segon ordre trilineal) (*y/n*). (*c*=concentració, *s*=espectres, *y*=sí, *n*=no)

TAULA 3. (Continuació)

- c. Correlació entre els valors recuperats i vertaders dels perfils de concentració (cosinus de l'angle entre ambdós) per cada espècie resolta i en cada matriu.
- d. Correlació entre els valors recuperats i vertaders dels perfils espectrals (cosinus de l'angle entre ambdós) per cada espècie resolta i en cada matriu.
- e. Error relatiu de les estimacions quantitatives en %. Per a dades amb estructura de segon ordre trilineal, la determinació quantitativa es fa emprant les equacions [17] i [20]. Per a dades sense estructura de segon ordre trilineal, la determinació quantitativa es fa a partir de la relació d'àrees entre els perfils de concentració de la mateixa espècie en les diferents matrius de dades. En tots els casos, una de les matrius de dades és considerada el patró.

Determinació quantitativa utilitzant matrius de l'analit pur com a patró

Aquesta és una situació comuna en l'anàlisi quantitativa. El problema a resoldre és la determinació d'un cert component (analit) en una mescla complexa no resolta amb interferències, utilitzant com a patró una matriu de respostes de l'analit en una mostra pura d'ell sense interferències. Es donen dos exemples:

MATRIU AUGMENTADA [2D;2C2]

La matriu 2D té dos components coeluïts a concentracions equivalents i la matriu 2C2 només conté informació sobre el segon component (estàndard pur). S'analitza la matriu augmentada [2D; 2C2] (equació [16]) amb el procediment descrit abans. Quan s'aplica la restricció d'igual forma en els perfils d'elució de l'espècie comuna, d'una banda es recuperen totalment els perfils de concentració i els perfils espectrals dels dos components en 2D i, de l'altra, es determina correctament la concentració (0,2 % error) del component comú. Quan no s'aplica aquesta restricció, la recuperació dels perfils és una mica pitjor i l'estimació quantitativa és més pobre (error en la predicció del 19 %). Això demostra que el fet de forçar a ser igual la forma dels perfils de concentració quan les dades tenen estructura de segon ordre trilineal, millora no solament la recuperació dels perfils, sinó també, i especialment, la determinació quantitativa. En anàlisi quantitativa seria desitjable aconseguir dades amb estructura de segon ordre trilineal.

MATRIUS AUGMENTADES [2F;2C1] o [2F;2C2]

S'investiguen dos casos diferents. En el primer cas, la matriu estàndard 2C1 té un component en comú amb la matriu 2F, que és el component majoritari.

L'objectiu és la determinació d'aquest component majoritari en presència d'una interferència. En el segon cas, l'estàndard 2C2 conté el component minoritari en 2F. La finalitat és la determinació del component minoritari en presència d'un interferent majoritari. Quan s'aplica la restricció d'igual forma en els perfils de concentració, els resultats són excel·lents en tots dos casos. Això demostra novament que els mètodes de resolució multivariant de corbes poden emprar-se no solament com a mètodes qualitius, sinó també com a mètodes quantitius. En cas de dades amb una estructura de segon ordre trilineal i anàlisi simultània de mostres problema i patrons purs, la determinació quantitativa és possible àdhuc en presència d'interferències desconegudes (avantatge de segon ordre).

Determinació quantitativa utilitzant matrius mescla com a patrons

MATRIU AUGMENTADA [2C; 2D]

Les dues matrius 2C i 2D contenen mescles dels mateixos components a diferents concentracions relatives. La determinació quantitativa es realitza en termes de concentracions relatives dels dos components en les dues matrius 2C i 2D. La matriu 2C es considera el patró. No hi ha selectivitat en cap dels dos ordres, les similituds entre els perfils són altes i la resolució cromatogràfica és baixa (taula 1). Els resultats obtinguts són també força bons (taula 3) amb un error per sota del 5 % (en el pitjor dels casos). Quan es considera que les dades no tenen estructura de segon ordre trilineal (per exemple en casos de manca de sincronització del temps d'elució del pic cromatogràfic del mateix component en les dues matrius) la recuperació dels perfils és pitjor i l'estimació quantitativa també empitjora (fins a un 15 % d'error). El mètode pot estendre's fàcilment a l'anàlisi simultània de diverses matrius de dades, com es mostra en el cas següent.

MATRIU AUGMENTADA [2C; 2D; 2E; 2F]

En aquest cas, tampoc no hi ha matrius d'estàndards o patrons purs. Els perfils se superposen fortament (taula 1). Els resultats es mostren a la taula 3, considerant i no considerant la restricció d'igual forma dels perfils d'elució del mateix component en les diferents matrius. En el primer cas els perfils es recuperen correctament i les estimacions quantitatives, prenent 2C com a estàndard, són bones. Només en la matriu 2E, l'estimació del component minoritari (a una concentració 10 vegades menor que en el patró) és més

pobra (12 % d'error). Quan no se suposa una estructura de segon ordre trilineal (diferent forma dels perfils de concentració), els resultats són encara bastant bons, excepte per als components a concentracions baixes en les matrius 2E i 2F que donen errors grans (del 52 % i del 42 %).

L'anàlisi simultània de les quatre matrius 2C, 2D, 2E i 2F és millor que l'anàlisi simultània de les dues matrius 2C i 2D. En el primer cas les prediccions en 2D es fan amb errors del 0,2 % i del 0,4 % (amb restricció d'igual forma dels perfils d'elució) o del 1,3 % i del 1,4 % (no forçant els perfils d'elució a tenir la mateixa forma), mentre que en el segon cas, els errors augmenten fins al 5,0 % i al 2 % (forçant) o fins al 15 % i al 4,1 % (sense forçar). Així doncs, el mètode proposat es pot estendre fàcilment a la calibració tensorial utilitzant un conjunt de matrius de dades com a patrons que cobreixen el rang de concentracions on s'espera que l'analit sigui present. L'augment del nombre de matrius de dades millorarà la precisió de les estimacions.

Cas III. Només un dels dos ordres és comú entre les matrius

Dades de l'exemple 1

MATRIUS [1B; 1E]

A les matrius 1B i 1E hi ha les mateixes dues espècies i la suma de llurs concentracions és constant (sistema tancat) però en diferents proporcions. Els perfils de concentració tenen forma diferent i no és possible aplicar la restricció de segon ordre trilineal. L'anàlisi simultània de les dues matrius recupera correctament els dos perfils de concentració i els perfils espectrals per a totes dues espècies. Recordem que en l'anàlisi individualitzada de la matriu 1E, l'ambigüitat rotacional persistia ja que no hi havia selectivitat per a cap de les dues espècies. Ara, en canvi, aquesta ambigüitat es resol, fonamentalment a causa de la selectivitat en 1B. Com que aquest és un sistema tancat, l'ambigüitat d'intensitat també es resol. Aquest exemple demostra que les ambigüitats rotacionals presents en l'anàlisi individual d'algunes matrius de dades poden resoldre's quan aquestes s'analitzen conjuntament amb altres matrius que tenen selectivitat. El mètode proposat permet la resolució correcta d'una matriu qualsevol, amb selectivitat o sense, si és analitzada conjuntament amb una altra matriu de dades del mateix sistema en què hi ha selectivitat. Aquesta possibilitat és extremadament important per a aquells sistemes que no tenen una estructura de segon ordre trilineal.

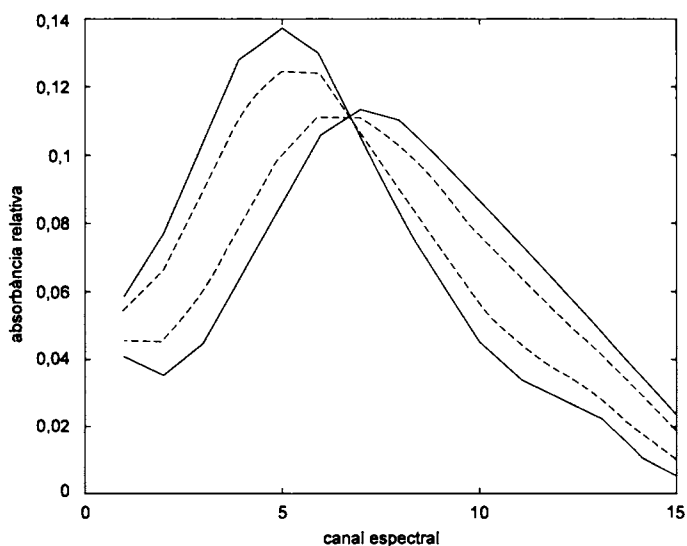


FIGURA 19. Comparació dels perfils espectrals obtinguts en l'anàlisi de la matriu augmentada [1E; 1F]. Les línies discontinües són els perfils recuperats pel mètode ALS i les línies contínues són els perfils teòrics.

MATRIUS [1E; 1F]

Cap de les dues matrius 1E i 1F tenen selectivitat en els seus perfils. La matriu 1F s'ha fet amb els perfils de concentració de la matriu 1E multiplicats per 2; per tant, els dos components no canvien de forma independent en les dues matrius per causa de la restricció de sistemes tancats. En l'anàlisi individual, per a cap de les dues matrius no es recuperaven els perfils correctes. Els resultats assolits en l'anàlisi de la matriu augmentada 1E+1F es donen a la taula 3. En aquest cas, tampoc no es recuperen els perfils correctes (figura 19). La informació quantitativa, en canvi, sí que és recuperada correctament. Amb la qual cosa es comprova que els espectres i els perfils de concentració recuperats descriuen els mateixos subespais vectorials que els perfils correctes, però en són una combinació lineal d'ells. Això sembla indicar que, en aquest cas, la solució assolida és degenerada i que l'ambigüitat rotacional no es resol per causa de la manca de selectivitat i del fet que les concentracions dels dos components no canvien de forma independent en les dues matrius analitzades.

Dades de l'exemple 3

MATRIU AUGMENTADA [3A; 3B; 3C]

Aquest cas no pot ésser estudiat mitjançant els mètodes que assumeixin una estructura de dades de segon ordre trilineal, com són GRAM,⁴⁷ la descomposició trilineal TLD⁴⁸ o els mètodes basats en models tipus PARAFAC.⁴⁹ Com en els casos anteriors, la matriu augmentada es construeix a partir de les matrius 3A, 3B i 3C, mantenint l'ordre de les columnes (longituds d'ona) en comú entre les diferents matrius (equació [16]). S'apliquen les restriccions de no-negativitat, unimodalitat, sistema tancat i igualtat d'espectre per a les espècies comunes. S'aplica, també, la restricció de selectivitat al principi de les tres valoracions per a la primera espècie i al final de la tercera valoració per a la tercera espècie. No s'aplica, en canvi, la restricció d'igual forma en els perfils de concentració, ja que no tenim, en aquest cas, una estructura de segon ordre trilineal.

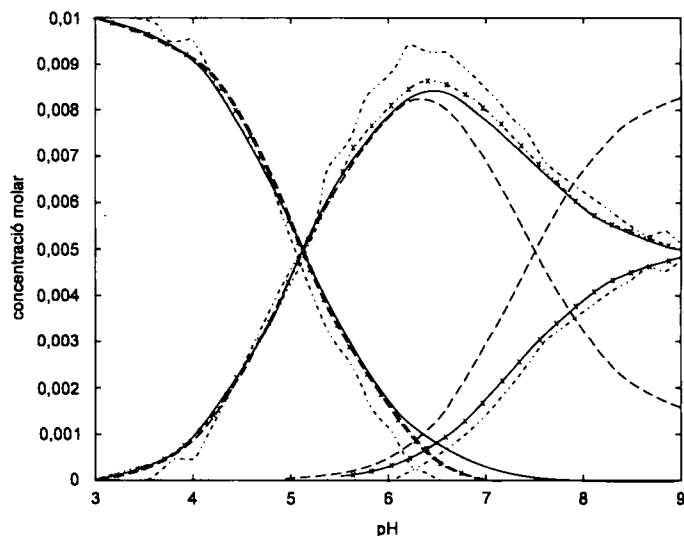


FIGURA 20. Comparació de la distribució d'espècies recuperada a partir de l'anàlisi de la matriu 3B i de la distribució d'espècies teòrica. Les línies contínues són els perfils teòrics; les línies discontinues són els perfils recuperats en l'anàlisi individualitzada de la matriu 3B; les línies —•— són els perfils recuperats en l'anàlisi de la matriu augmentada [3A; 3B; 3C] (nivell de soroll amb una desviació estàndard de 0,000001 unitats).

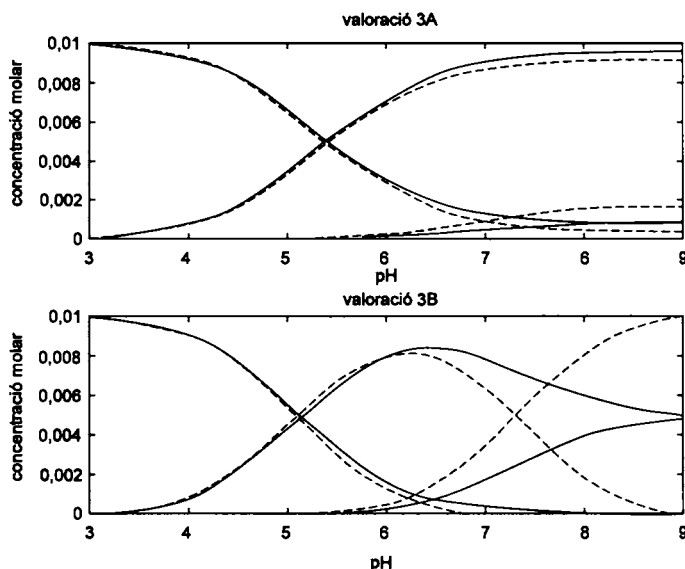


FIGURA 21. Comparació dels perfils espectrals obtinguts en l'anàlisi de la matriu augmentada [3A; 3B]. Les línies discontinües són els perfils recuperats i les contínues són els perfils teòrics.

La qualitat dels resultats depèn del nivell de soroll. A la figura 20 es mostra una comparació dels resultats assolits per a la valoració 3B utilitzant diferents tractaments. En l'anàlisi individual, la recuperació dels perfils de concentració de la segona i la tercera espècie era incorrecta, mentre que en l'anàlisi simultània, la recuperació dels mateixos perfils millora substancialment. Per un error amb una desviació estàndard de 0,001 unitats, els perfils de concentració recuperats tenen formes no totalment resoltes que milloren quan s'abaixa el nivell de soroll. La recuperació dels espectres és, en canvi, sempre bona en tots dos casos (taula 3). L'ajustament de les dades és del mateix nivell que l'error experimental. Així doncs, l'anàlisi simultània de les tres valoracions espectroscòpiques permet de resoldre les ambigüitats rotacionals que existien en l'anàlisi individual, sobretot de les matrius 3A i 3B. En aquest cas, però, els factors determinants en la ruptura de l'ambigüitat rotacional han estat la presència, la detecció i l'ús de la informació selectiva present, sobretot, en la valoració 3C. Per acabar d'esbrinar l'efecte de la selectivitat s'estudià el següent cas.

MATRIU AUGMENTADA [3A;3B]

Per veure quin és l'efecte de la manca de selectivitat, les matrius de dades 3A i 3B s'analitzen simultàniament. En aquestes dues valoracions no existeix selectivitat per a la segona i la tercera espècie. A la figura 21 i a la taula 3 es mostren els resultats assolits. El perfil de concentracions de la segona espècie s'estima de forma incorrecta, especialment en la segona valoració. Això demostra (com en l'estudi de la matriu augmentada [1E; 1F]) que, en cas de manca de selectivitat i absència d'estructura de segon ordre trilineal, l'anàlisi simultània de diferents matrius de dades correlacionades no resol les ambigüitats rotacionals, àdhuc quan s'apliquen les restriccions de no-negativitat, unimodalitat, sistema tancat i igualtat d'espectre per espècies comunes. La manca de selectivitat i d'estructura de segon ordre trilineal no permet, en aquest cas, assolir la resolució correcta.

2.2.5. CONCLUSIONS

L'èxit de l'aplicació del mètode de resolució multivariant de corbes en l'anàlisi d'una matriu única de dades depèn de la presència, la detecció i l'ús de la selectivitat del sistema, com també de l'ús d'un conjunt de restriccions. El mètode de resolució multivariant de corbes pot adaptar-se de forma senzilla a l'anàlisi simultània de diverses matrius de dades correlacionades amb l'estructura de segon ordre trilineal, o no. Es recuperen solucions sense ambigüitat per al cas de dades amb estructura de segon ordre trilineal. En l'anàlisi simultània de matrius de dades sense una estructura de segon ordre trilineal, la recuperació dels perfils correctes dependrà de la selectivitat i de les restriccions aplicades. Així doncs, el cas més difícil serà aquell en què no hi ha selectivitat i l'estructura de les dades no és trilineal. En aquests casos particulars, la recuperació dels perfils correctes i l'extracció d'informació quantitativa és encara difícil, encara que es pot millorar amb l'ús de restriccions com la de no-negativitat, la d'unimodalitat o la de sistema tancat.

3. L'APLICACIÓ DEL MÈTODE DE RESOLUCIÓ MULTIVARIANT DE CORBES (MVCR) A PROBLEMES REALS

A continuació es descriuen una sèrie d'exemples reals en què ha estat aplicat el mètode descrit, els quals cobreixen un ampli espectre de sistemes d'interès en diverses àrees. Els primers exemples es refereixen a sistemes multiequilibri analitzats mitjançant valoracions espectromètriques. Tenen una especial relevància en aquest grup, els sistemes macromoleculars, on precisament l'obtenció d'informació analítica de les diferents espècies existents i la utilització de la llei d'acció de masses és molt difícil o gairebé impossible. En segon lloc, es mostren una sèrie d'aplicacions dins l'àmbit de les determinacions analítiques dels components coeluïts mitjançant mètodes cromatogràfics. En tercer lloc, es mostra l'aplicació del mètode en l'anàlisi de processos industrials. En aquest cas, es tracta de predir el que succeeix en un determinat procés a partir del seu seguiment *in situ* i *on line*. Finalment, es mostren els resultats de l'aplicació del mètode descrit al desenvolupament de nous sensors òptics, basats en la utilització de processos de separació i de reacció per a la detecció i la quantificació *in situ* dels constituents d'una mostra complexa.

3.1. APLICACIÓ DEL MÈTODE MVCR PER A L'ESTUDI DE REACCIONS QUÍMIQUES EN SOLUCIÓ MITJANÇANT VALORACIONS ESPECTROMÈTRIQUES

El primer exemple d'aplicació del mètode que es proposa és l'estudi de valoracions espectromètriques de sistemes multiequilibri. A la figura 22 es mostra un esquema del procediment experimental emprat. Els mètodes de resolució de corbes poden utilitzar-se en l'estudi de valoracions espectromètriques de sistemes multiequilibri com a procediments en els quals no es proposa

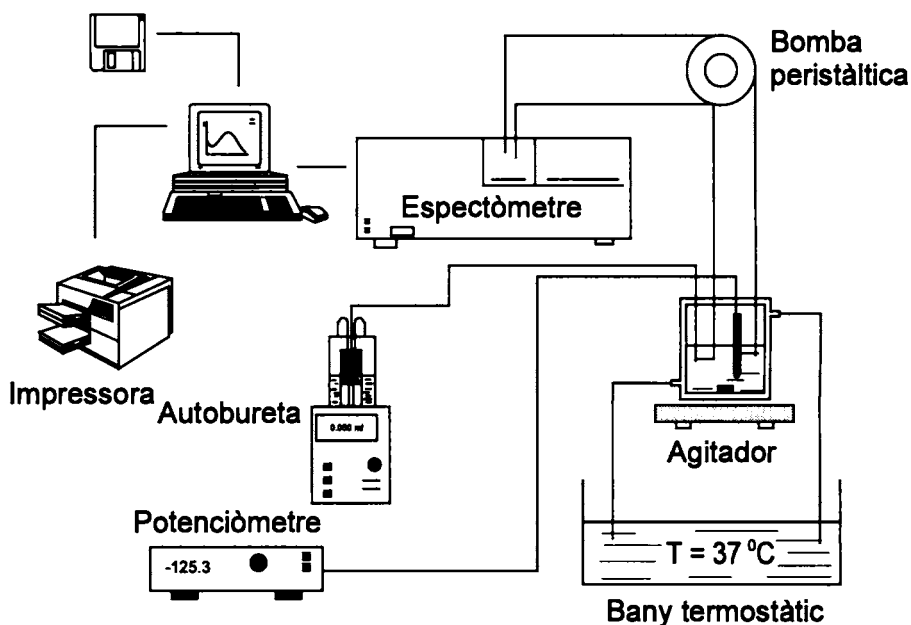


FIGURA 22. Valoracions espectromètriques. Esquema d'adquisició de dades.

un model químic inicial, entès aquest com a definició de les espècies presents, mitjançant llurs coeficients estequiomètrics i constants termodinàmiques de formació. Quan s'utilitza el mètode de resolució multivariant de corbes, no s'apliquen les restriccions químiques derivades del coneixement del model, com és la llei d'acció de masses. Els procediments de resolució multivariant de corbes poden emprar-se de manera complementària als procediments tradicionals d'estudi de sistemes multiequilibri basats en l'ajust minimoquadràtic de models químics determinats.

En la part anterior ja s'ha mostrat l'aplicació al cas del seguiment amb el pH de la reacció de complexació entre l'ió Cu(II) i el salicilat, que serveix com a exemple de les possibilitats d'aplicació del mètode en sistemes similars. Utilitzant aquesta aproximació s'han estudiat sistemes Cu(II)-X i Cu(II)-sal-X , on $\text{X}=\text{mea}$ (monoetanolamina), dea (dietanolamina), tea (triethanolamina); sal és salicilat per espectrometria UV-VIS i ESR.^{9-13, 15, 16, 22, 23, 27}

3.2. APLICACIÓ DEL MÈTODE MVCR PER A L'ESTUDI DE LA COMPLEXACIÓ EN SISTEMES MACROMOLECULARS MITJANÇANT VALORACIONS ESPECTROMÈTRIQUES

L'estudi de complexos metàl·lics macromoleculars en solució és un camp de gran interès, tant per als estudis ambientals com per als biotecnològics.⁶⁷ Aquest estudi, moltes vegades ve dificultat pel fet que les lleis de l'equilibri químic (per exemple, la llei d'acció de masses) només són estrictament vàlides per a cada posició reactiva de la macromolècula, i cal considerar diversos efectes secundaris,⁶⁸ com són l'efecte polielectrolític, l'efecte polifuncional i l'efecte conformacional. Tots aquests efectes contribueixen a l'estabilitat de les espècies formades, i la seva importància relativa és difícil de definir, ja que varia amb el grau d'ocupació de les posicions reactives (per complexació o protonació). La interpretació de les dades experimentals mitjançant mètodes tradicionals d'ajust mínim quadràtic és, en aquests casos, força difícil,⁶⁹ i és necessari el desenvolupament de nous mètodes d'anàlisi de les dades, lliures de les restriccions imposades per la llei d'acció de masses, i lliures de l'assumpció inicial d'un model químic predeterminat. El mètode de resolució multivariant de corbes, descrit en els apartats anteriors, ha permès avançar, considerablement, en l'estudi de les reaccions i els equilibris de les macromolècules en solució. A continuació, es mostren un parell d'exemples referents a la complexació de la polietilenimina (poliE) i de l'àcid poliadenílic (poliA). Cal destacar que en el nostre laboratori ja s'està utilitzant constantment aquest procediment, en l'estudi de: *a*) propietats àcid-base, propietats complexants de metalls i canvis conformacionals de diversos polinucleòtids (àcid poliuridílic,⁷⁰ àcid poliinosínic,⁷¹ àcid policitosínic);⁷² i *b*) propietats àcid-base i propietats complexants de metalls, dels àcids fúlvics aïllats de la matèria orgànica dels sòls.⁷³

a) Sistema Cu(II)-polietilenimina

S'estudia la formació dels complexos de Cu(II) amb la polietilenimina mitjançant valoracions espectromètriques en el VIS. A la referència 74 es donen detalls sobre el procediment experimental i altres característiques del sistema estudiat. Es va detectar la presència de tres complexos metàl·lics macromoleculars de Cu(II), a part de l'ió Cu(II) lliure. A la figura 23 es mostren els resultats de l'aplicació del mètode EFA a dues valoracions espectromètriques analitzades. En aquest cas, fins a pH 4,7, es detecta clarament la presència de tres espècies químiques diferenciades. Quan s'augmenta el pH fins a valors més bàsics es detecta la formació d'una altra espècie.

A partir de l'aplicació del mètode de resolució multivariant de corbes proposat es determinen la distribució (figura 24) i els espectres unitaris (figura 25) de les espècies formades, en variar el pH. Especialment interessant, en aquest cas, és l'estudi de les propietats complexants de la polietilenimina, a partir dels perfils de concentració de les diferents espècies detectades en la distribució d'espècies. A partir de l'evolució del grau de complexació del metall, en variar el pH, es poden avaluar les constants d'estabilitat aparent. El grau de complexació del metall es calcula a partir del quocient de metall unit a la macromolècula respecte del metall total, i/o a partir del quocient de posicions reactives ocupades de la macromolècula, respecte de posicions reactives totals. A la taula 4 es mostra un exemple de com es pot fer una estimació de les propietats de complexació de la macromolècula, a partir dels resultats assolits amb el tractament proposat. A partir de resultats anteriors sobre la protonació de la polietilenimina,⁷⁵ hom pot estimar el nombre de posicions reactives que es troben protonades a cada pH.

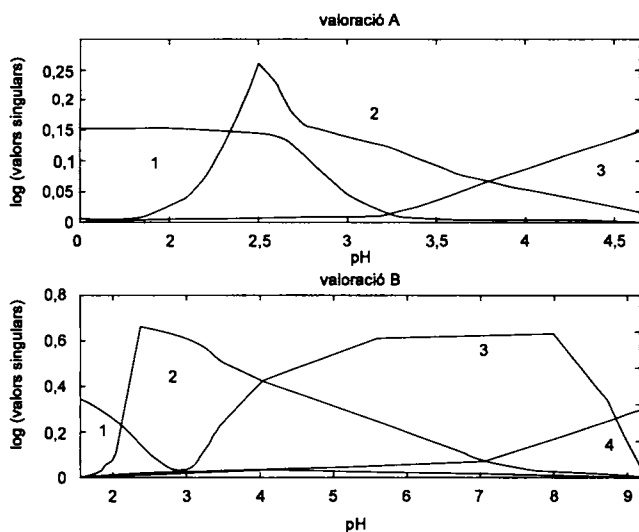


FIGURA 23. Diagrames de distribució abstractes obtinguts mitjançant l'anàlisi de l'evolució dels factors (EFA) de dues matrius de dades, corresponents a dues valoracions espectromètriques (A i B) del sistema Cu(II)-polietilenimina. Les línies més gruixudes són les estimacions inicials dels perfils de concentració de les diferents espècies. Els nombres de l'esquerra i de la dreta donen els factors creixents i decreixents en l'anàlisi EFA, endavant i endarrere [39]. Matriu A pH 1,52-4,67; Matriu B pH 1,43-11,3.

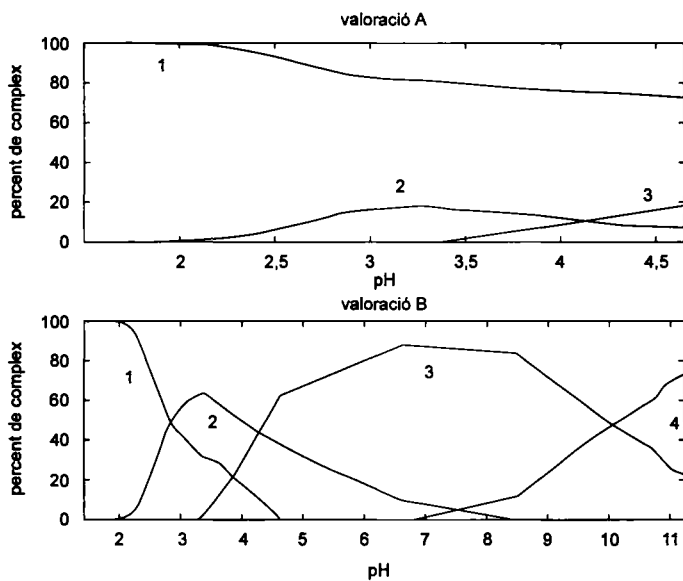


FIGURA 24. Distribució d'espècies obtinguda en el tractament conjunt de les matrius A i B. (1) espècie Cu(II) lliure; (2-4) espècies de complexació successiva del Cu(II) amb la poliE.

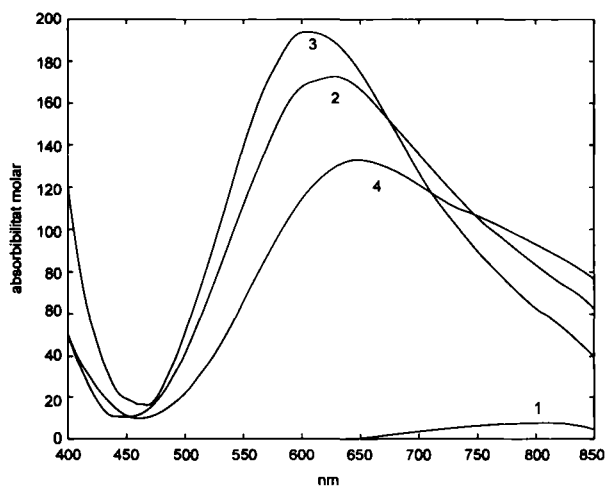


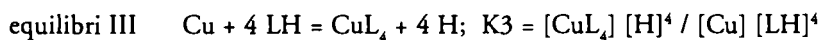
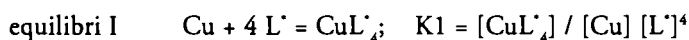
FIGURA 25. Espectres de les 4 espècies detectades i resoltes en les dues valoracions espectromètriques. (1) espècie Cu(II) lliure; (2-4) espècies de complexació successiva del Cu(II) amb la poliE.

TAULA 4
Propietats complexants de la polietilenimina

NS	pH	X0	X1	Y0	Y1	Y2	Y3	LK1	LK2	LK3
1	1,430	1,000	0,000	0,230	1,000	0,770	0,000	—	—	—
2	1,500	1,000	0,000	0,240	1,000	0,760	0,000	—	—	—
3	1,600	1,000	0,000	0,240	1,000	0,760	0,000	—	—	—
4	1,700	1,000	0,000	0,250	1,000	0,750	0,000	—	—	—
5	1,790	0,989	0,011	0,248	0,990	0,752	0,010	3,596	6,004	-3,064
6	1,940	0,983	0,017	0,256	0,985	0,744	0,015	3,796	6,136	-3,441
7	2,050	0,972	0,028	0,254	0,975	0,746	0,025	4,044	6,384	-3,633
8	2,130	0,952	0,048	0,249	0,958	0,751	0,042	4,318	6,658	-3,679
9	2,260	0,897	0,103	0,246	0,910	0,754	0,090	4,762	7,037	-3,731
10	2,350	0,839	0,161	0,232	0,859	0,768	0,141	5,086	7,361	-3,767
11	2,470	0,738	0,262	0,216	0,770	0,784	0,230	5,544	7,755	-3,766
12	2,550	0,655	0,345	0,195	0,697	0,805	0,303	5,885	8,097	-3,744
13	2,680	0,530	0,470	0,164	0,587	0,836	0,413	6,410	8,621	-3,740
14	2,760	0,457	0,543	0,152	0,523	0,848	0,477	6,737	8,887	-3,708
15	2,880	0,376	0,624	0,131	0,452	0,869	0,548	7,138	9,289	-3,787
16	2,980	0,347	0,653	0,124	0,426	0,876	0,574	7,296	9,446	-4,029
17	3,160	0,288	0,712	0,112	0,374	0,888	0,626	7,640	9,731	-4,380
18	3,360	0,228	0,772	0,097	0,322	0,903	0,678	8,035	10,127	-4,785
19	3,610	0,203	0,797	0,093	0,299	0,907	0,701	8,228	10,263	-5,567
20	3,850	0,150	0,850	0,081	0,253	0,919	0,747	8,681	10,661	-6,049
21	4,630	0,000	1,000	0,043	0,121	0,957	0,879	—	—	—
22	6,630	0,000	1,000	0,064	0,121	0,936	0,879	—	—	—
23	8,500	0,000	1,000	0,109	0,121	0,891	0,879	—	—	—
24	9,730	0,000	1,000	0,121	0,121	0,879	0,879	—	—	—
25	10,660	0,000	1,000	0,121	0,121	0,879	0,879	—	—	—
26	11,030	0,000	1,000	0,121	0,121	0,879	0,879	—	—	—
27	11,300	0,000	1,000	0,121	0,121	0,879	0,879	—	—	—

NS és l'espectre considerat; el pH d'aquest; X0 és la fracció del metall lliure (no complexat); X1 és la fracció de metall complexat; Y0 és la fracció de posicions de la macromolècula no complexades pel metall; Y1 és la fracció de posicions lliures de la macromolècula (no complexades, no protonades); Y2 és la fracció de posicions de la macromolècula, protonades o complexades; Y3 és la fracció de les posicions de la macromolècula complexades (però no protonades); LK1 i LK2 són els logaritmes de les constants aparents d'estabilitat, calculades a partir del logaritme dels quocients de concentracions d'ió metàl·lic, complexat i dividit pel producte de la concentració d'ió metàl·lic lliure, i la concentració de posicions no complexades de la macromolècula (LK1), o de la concentració de posicions de la macromolècula no complexades ni protonades (LK2), elevat a l'estequiometria de quatre; LK3 és el logaritme del quocient de concentracions de les espècies de l'equilibri de substitució: $4 \text{ LH} + \text{Cu} = \text{CuL}_4 + 4 \text{ H}^+$, on LH es refereix a la posició protonada de macromolècula i CuL_4 és la concentració de metall, unit a quatre posicions macromoleculars.

Quan s'assumeix que el nombre de posicions reactives de la macromolècula, coordinades a cada metall és quatre, les constants aparents d'estabilitat es poden deduir a partir dels quocients d'equilibri I a III.



on L, L*, L' són, respectivament, les posicions reactives lliures de la macromolècula, les posicions reactives de la macromolècula que no es troben complexades i les posicions reactives de la macromolècula que no es troben protonades. L'equilibri III descriu la reacció competitiva del Cu(II) i els ions d'hidrogen per a les mateixes posicions reactives de la macromolècula. L'evolució d'aquestes constants aparents amb el pH, mostra un marcat efecte polifuncional i polielectrolític. Les dades fins a un pH 3 mostren la capacitat complexant inicial de la macromolècula, abans que en el sistema coexisteixin diversos complexos metàl·lics diferents.

b) Sistema Cu(II)-àcid poliadenilic⁷⁶

El següent exemple es refereix a l'estudi del comportament de l'àcid poliadenilic (poliA) en els seus equilibris àcid-base, de complexació amb l'ió Cu(II) i en canvis conformacionals. Aquest estudi és important, com a pas previ per entendre millor les estructures en solució i les propietats dels àcids nucleics naturals. Els ions metàl·lics tenen efectes profunds en l'estructura i funció dels àcids nucleics, dels mononucleòtids i dels polinucleòtids.⁷⁷

L'estudi d'aquest sistema es va realitzar mitjançant el procediment ja descrit, de valoracions espectromètriques (figura 21). La detecció espectral es va fer mitjançant espectrofotometria UV i espectrofotometria de diroisme circular (CD). Aquesta última és, especialment útil quan el que es pretén és la detecció de canvis conformacionals macromoleculars. Es descriuen detalls experimentals sobre el procediment emprat a la referència 76. Un aspecte particular del procediment emprat en aquest sistema és la possibilitat d'analitzar, conjuntament, la informació obtinguda mitjançant espectrofotometria UV i espectrofotometria CD. A la pàgina següent es mostra un esquema (figura 26) de les possibles matrius augmentades, que es poden emprar en l'estudi d'aquest sistema. L'anàlisi conjunta de la informació proporcionada per les

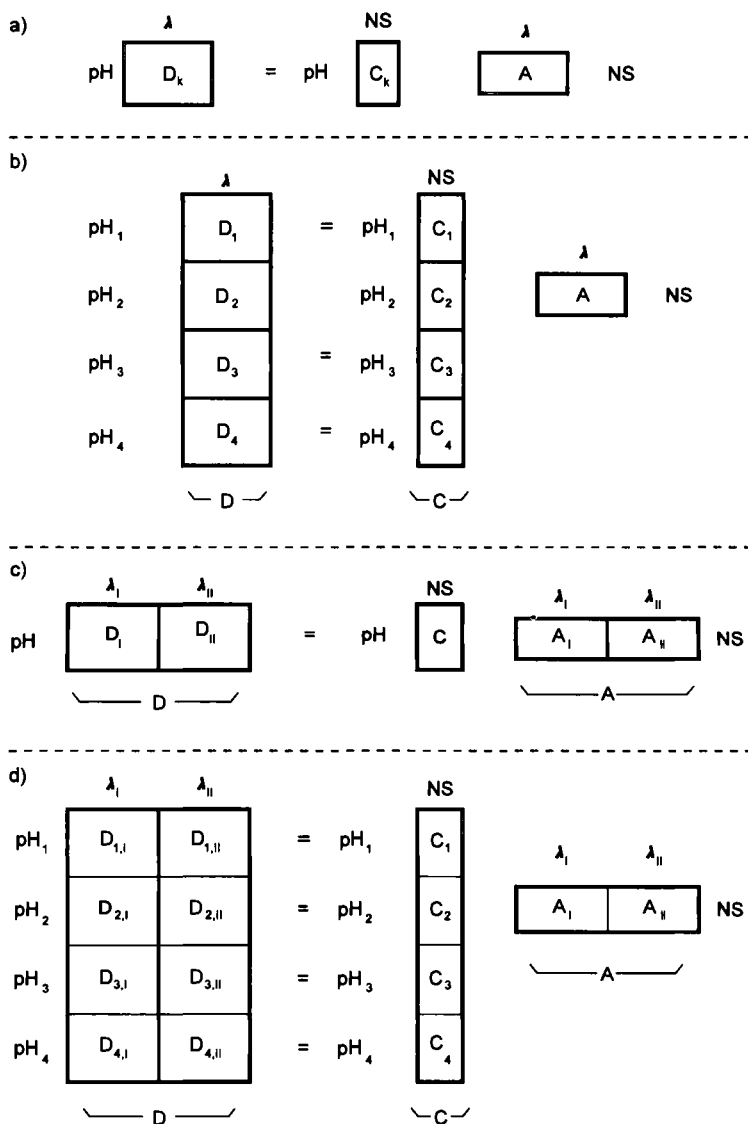


FIGURA 26. Esquema de diferents matrius de dades i disposicions de les matrius: a) tractament individual de cada matriu per separat; b) tractament simultani de les matrius de dades (àcid-base o complexació) obtingudes amb cada tècnica (UV o CD); c) tractament simultani de les matrius de dades (UV i CD) obtingudes amb cada tipus de valoració (àcid-base i complexació); d) tractament simultani de totes les matrius de dades (UV, CD, àcid-base i complexació).

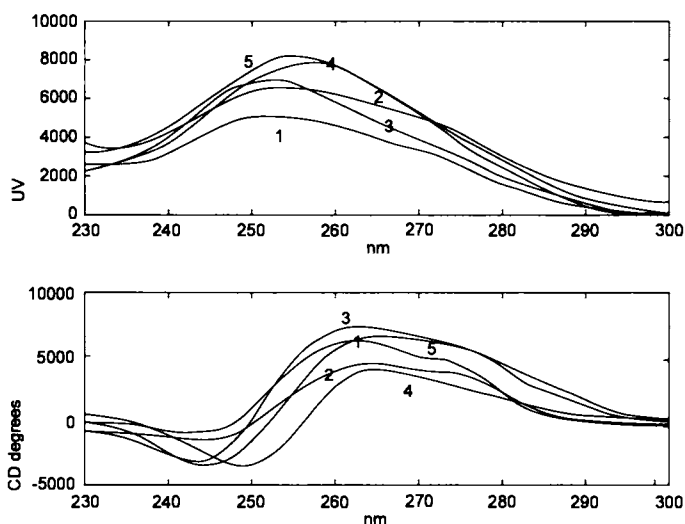
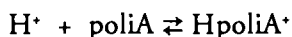


FIGURA 27. Espectres purs resolts en l'anàlisi de la matriu, augmentada amb totes les dades experimentals (valoracions espectromètriques UV, CD, àcid-base i complexació). A l'esquerra, es mostra l'espectre d'absorció i, a la dreta, l'espectre de dicroisme circular. Els nombres 1, 2, 3 corresponen a les diferents conformacions de l'àcid poliadenílic protonat; el 4 correspon a l'espècie totalment desprotonada i el 5, al possible complex de Cu(II) i àcid poliadenílic desprotonat.

dues tècniques espectrals augmenta considerablement el poder de detecció i de resolució del mètode.

En les valoracions àcid-base de l'àcid poliadenílic es detecten 4 espècies espectroscòpicament diferents, en el rang de pH 1,5 a 6,8. A la figura 27 es donen els espectres mixtos UV-CD resolts en el tractament simultani de diverses valoracions espectromètriques, emprant les dues tècniques espectrals. A la figura 28 es donen els diagrames de distribució, obtinguts en l'anàlisi de cadascuna de les valoracions espectromètriques analitzades simultàniament. Tres de les espècies corresponen, probablement, a tres conformacions diferents de l'espècie protonada de l'àcid poliadenílic, i la quarta és l'espècie totalment desprotonada. El punt mitjà de la transició entre les formes protonada i desprotonada és pH 5,8. Aquests resultats estan d'acord amb estudis realitzats per altres autors sobre les diferents conformacions del poliA. Es detecta una espècie nova, en el cas de la presència de Cu(II) que, òbviament, correspon a un complex metàl·lic de la macromolècula. La reacció postulada per a l'equilibri àcid-base és:



on poliA és la forma desprotonada de l'àcid poliadenílic, HpoliA⁺ és la seva forma protonada, i H⁺ és l'ió hidrogen. La constant aparent de l'equilibri corresponent és:

$$(\text{Kapp})_{\text{prot}} = [\text{HpoliA}^+]/([\text{H}^+][\text{poliA}])$$

Les concentracions de les formes protonada i desprotonada es troben directament dels resultats anteriors (espècies 3 i 4 a la figura 26) i la concentració d'ió hidrogen és determinada experimentalment. El grau de protonació

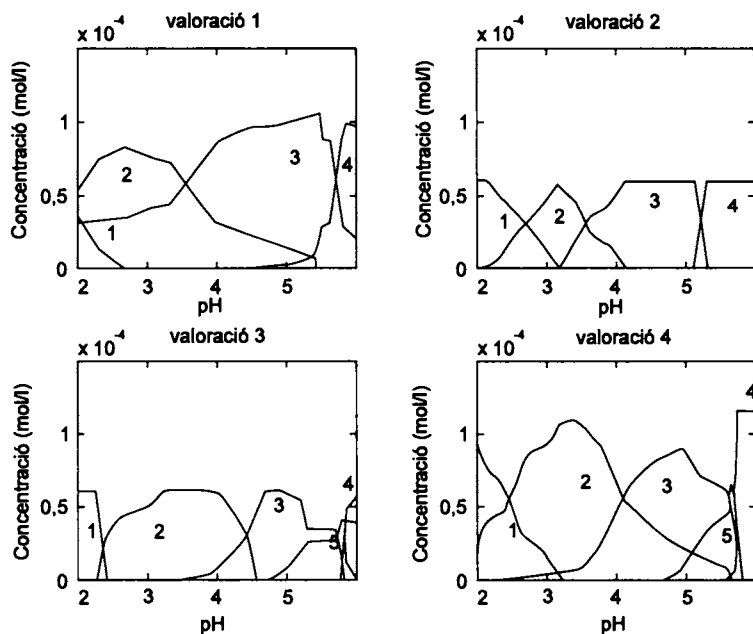
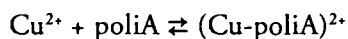


FIGURA 28. Diagrama de distribució d'espècies obtingut en l'anàlisi de la matriu augmentada amb totes les dades experimentals (valoracions espectromètriques UV, CD, àcid-base i complexació). Els nombres de les corbes representen el mateix que a la figura 27. Cada valoració és separada per una línia discontinua. Valoració 1: 17 espectres, des de pH 1,96 fins a pH 5,92 (sense ió Cu(II)). Valoració 2: 21 espectres, des de pH 1,45 fins a pH 6,88 (sense ió Cu(II)). Valoració 3: 21 espectres, des de pH 1,43 fins a pH 6,05 (amb ió Cu(II)). Valoració 4: 22 espectres, des de pH 1,36 fins a pH 6,18 (amb ió Cu(II)).

(α_{prot}) s'obté a partir de la relació entre la concentració d'espècie protonada i la concentració total de poliA:

$$\alpha_{\text{prot}} = [\text{HpoliA}^+]/C_T = [\text{HpoliA}^+]/([\text{poliA}] + [\text{HpoliA}^+])$$

De la mateixa manera, es poden considerar els equilibris de complexació entre l'ió Cu(II) i l'àcid poliadenílic:



amb una constant aparent de complexació $(K_{\text{app}})_{\text{comp}}$

$$(K_{\text{app}})_{\text{comp}} = \frac{[(\text{Cu-poliA})^{2+}]}{[\text{Cu}^{2+}] [\text{poliA}]} = \frac{[(\text{Cu-poliA})^{2+}]}{([\text{Cu}_T^{2+}] - [(\text{Cu-poliA})^{2+}]) [\text{poliA}]} \quad [13]$$

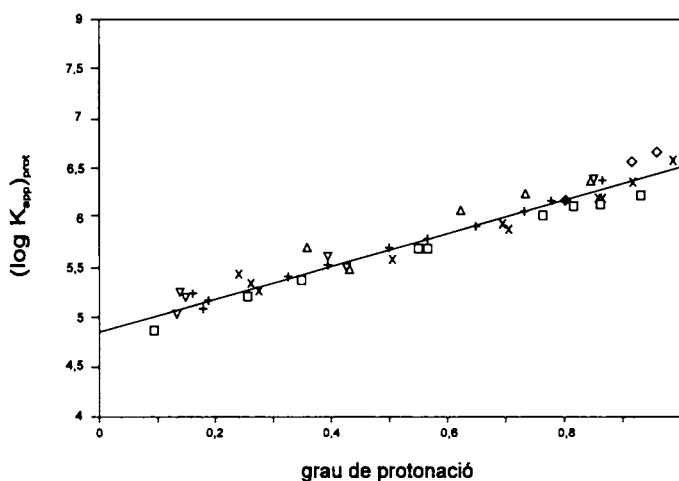


FIGURA 29. Variació del logaritme de la constant de protonació de l'àcid poliadenílic, respecte al grau de protonació en 6 valoracions diferents (UV i CD). La constant és obtinguda a partir del valor del pH, mesurat experimentalment, i de les concentracions d'equilibri de les espècies 3 i 4 de la figura 28.

on ara les concentracions de l'espècie complexada i lliure de polyA són, respectivament, les espècies número 5 i 4 de la figura 28. El grau de complexació α_{comp} s'obté a partir de l'equació:

$$\alpha_{\text{comp}} = [(\text{Cu-poliA})^{2+}] / [\text{Cu}_T]$$

on $[(\text{Cu-poliA})^{2+}]$ és la concentració de l'espècie complexada (espècie 5) i $[\text{Cu}_T]$ és la concentració total de l'ió Cu(II).

La representació del logaritme de les constants aparents de protonació ($\log K_{\text{app, prot}}$, respecte al grau de protonació α_{prot} , es mostra a la figura 29 i el del logaritme de la constant aparent de complexació $\log (K_{\text{app, comp}})$, respecte al grau de complexació, es mostra a la figura 30. A partir d'aquestes figures s'obté que el valor del logaritme de la constant de protonació extrapolada a $\alpha_{\text{prot}}=0$ (quan la macromolècula es troba totalment desprotonada) és 4,9 i el valor del logaritme de la constant de complexació extrapolada a $\alpha_{\text{comp}}=0$ (complexació i protonació nul·la) és 2,9. Aquests valors són una mica superiors als determinats per la mateixa protonació i complexació de l'adenina (4,02 i 2,25 respectivament, 14).

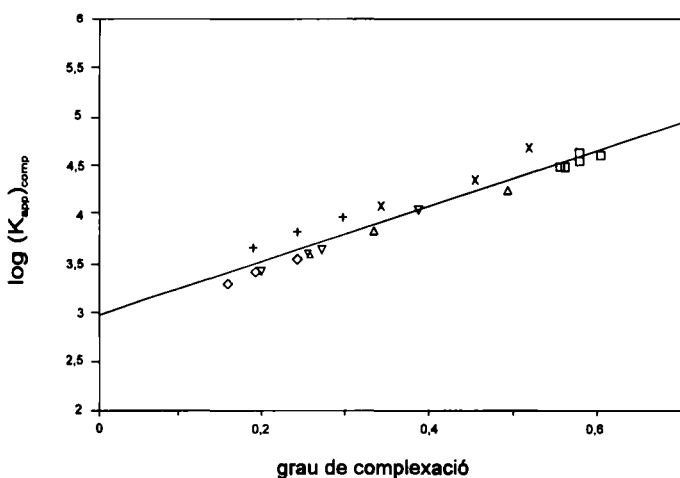


FIGURA 30. Variació del logaritme de la constant de complexació de l'àcid poliadenílic amb l'ió Cu(II) respecte al grau de complexació en 6 valoracions diferents (UV i CD). La constant és obtinguda a partir de la concentració de Cu(II) lliure, mesurada experimentalment (elèctrode d'ió selectiu de Cu(II) i de les concentracions d'equilibri de les espècies 4 i 5 de la figura 28.

La dependència linear obtinguda dels valors d'aquestes dues constants, amb el grau d'ocupació de les posicions reactives de la macromolècula és contrària al que hom pot suposar a partir de l'efecte polielectrolític. El que s'observa és que, a mesura que es protona la macromolècula i per tant, augmenta la càrrega sobre la superfície d'aquella, la constant de protonació augmenta en lloc de disminuir, com seria d'esperar en cas d'efecte polielectrolític pur. L'explicació que es dona en aquest fenomen, deduït directament a partir del tractament amb el mètode proposat de les dades experimentals, és que l'estabilització observada respon a l'aparellament de les bases d'adenina, mitjançant enllaços per pont d'hidrogen. Aquest fet també explicaria els canvis de conformació observats en les zones de pH àcides. L'àcid poliadenílic neutre té una conformació bastant desordenada, mentre que les formes protonades presenten estructures molt més ordenades, en les quals és fonamental la formació d'enllaços per pont d'hidrogen. Es pot fer una interpretació similar dels resultats assolits per la complexació amb Cu(II).

3.3. APLICACIÓ DEL MÈTODE MVCR A ALGUNS SISTEMES DE DETERMINACIÓ ANALÍTICA

A l'apartat 2, s'ha demostrat que el mètode descrit, permet l'extracció d'informació quantitativa, a partir de l'anàlisi simultània de diverses matrius de dades obtingudes en l'anàlisi cromatogràfica, amb detecció espectroscòpica de mescles de components coeluits. A continuació, es mostren els resultats obtinguts en l'estudi del sistema reals cromatogràfics.

A la figura 31, es mostra la representació en tres dimensions de la matriu de dades experimental, obtinguda en l'anàlisi cromatogràfica de la mescla no resolta, dels pesticides naftol i pirimicarb. A les referències 78 i 79, es donen els detalls experimentals de les determinacions realitzades. A qualsevol longitud d'ona existeix superposició dels dos pics cromatogràfics, amb resolucions baixes (per sota de 0,5). A la figura 32, es mostren els resultats de l'aplicació del mètode FSWEFA en l'ordre cromatogràfic. Es detecta clarament la regió de coelució dels dos components, així com la zona on es troben els dos components purs. L'aplicació del mètode de resolució multivariant de corbes permet la resolució unívoca dels dos components, l'obtenció dels seus dos perfils d'elució i dels seus dos espectres unitaris (figura 33). A més, quan es realitza l'anàlisi simultània de la matriu que conté la mescla de components, amb una o les dues matrius dels dos patrons de pirimicarb i naftol, es pot realitzar la seva determinació quantitativa. A la taula 5, es mostra un resum dels resultats assolits; també s'hi han inclòs els resultats assolits pel mètode GRAM (*Generalized Rank Annihilation method*),⁴⁷ el qual es pot considerar

com el procediment estàndard en l'anàlisi de dades amb estructura de segon ordre trilineal. Els resultats assolits per ambdós mètodes són concordants. Cal destacar, però, que GRAM només es pot utilitzar quan existeix sincronització cromatogràfica i quan l'estructura de les dades és de segon ordre trilineal, mentre que el mètode de resolució multivariant de corbes no és tan restrictiu i pot adaptar-se també a l'estudi de dades sense sincronització i sense estructura de segon ordre trilineal (vegeu apartat 2).

TAULA 5

Resultats de la determinació quantitativa dels components del pic cromatogràfic LC-DAD corresponent a la mescla no resolta de pirimicarb (pir) i naftol (naf)^a

Mètode ^b	Pir	Naf	N
1	0,267 (4 %)	0,039 (17 %)	1
2a	0,270 (6 %)	0,045 (3 %)	3
2b	—	0,043 (8 %)	2
2c	0,189 (26 %)	—	2
3a	0,263 (3 %)	0,047 (1 %)	3
3b	—	0,049 (5 %)	2
3c	0,252 (1 %)	—	2
4a	—	0,050 (6 %)	2
4b	0,256 (<1 %)	—	2

- a. Determinació quantitativa de la mescla no resolta pirimicarb-naftol de la figura 30. Les concentracions vertaderes són, respectivament, (en ppm): pirimicarb 0,256 i naftol 0,047
- b. Mètode emprat en la determinació quantitativa de la mescla no resolta: 1) Regressió lineal a partir dels espectres unitaris dels components purs; 2) Anàlisi simultània de diverses matrius de dades utilitzant el mètode ALS de resolució multivariant de corbes i la restricció d'igual espectre unitari per a les espècies comunes en les diferents matrius de dades: 2a) anàlisi simultània de la matriu mescla i de les dues matrius

d'analits purs; 2b) anàlisi simultània de la matriu mescla i de la matriu de naftol pur; 2c) anàlisi simultània de la matriu mescla i de la matriu de pirimicarb pur. 3) Com en 2), però amb la restricció d'espectres unitaris iguals i perfils d'elució amb igual forma per a les espècies comunes; 4) Mètode *generalized rank annihilation* (GRAM): 4a) anàlisi simultània de la matriu mescla i de la matriu de naftol pur; 4b) anàlisi simultània de les matrius mescla i de la matriu de pirimicarb pur.

c. Nombre de matrius de dades emprades en l'anàlisi.

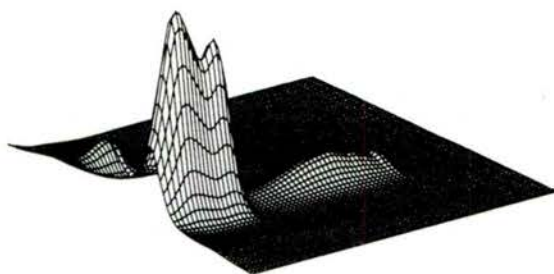


FIGURA 31. Diagrama 3D del pic cromatogràfic LC-DAD, corresponent a la mescla no resolta dels pesticides naftol i pirimicarb.

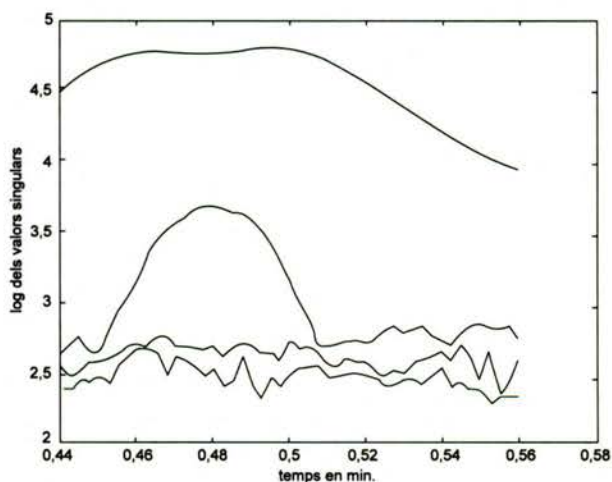


FIGURA 32. Detecció de la selectivitat en l'ordre de les concentracions (elució) de la matriu de dades, corresponent a la figura 31, emprant el mètode FSMWEFA.⁶¹ Finestra de 5 espectres. Es detecten 2 components que coelueixen entre 0,450 i 0,510 min. Existeix selectivitat per als dos components.

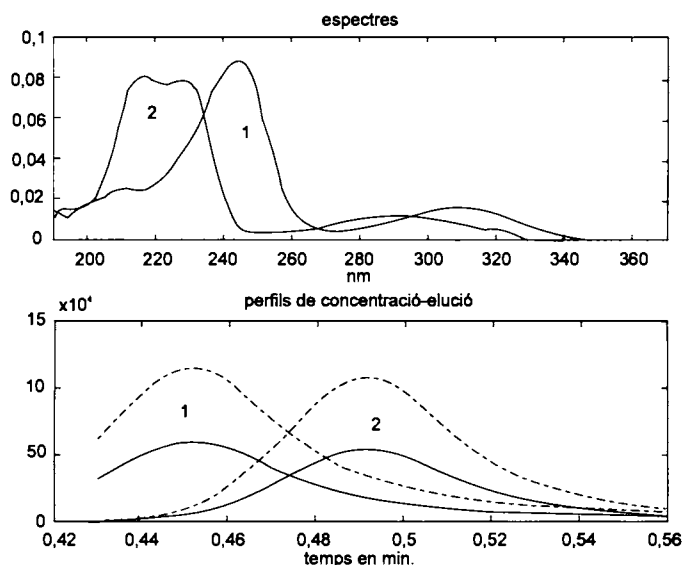


FIGURA 33. Espectres resolts dels dos components coeluïts, pirimicarb i naftol (a la part superior; naftol en línia discontinua i pirimicarb en línia contínua). Perfils d'elució dels mateixos components coeluïts (a la part inferior; mostres dels components purs en línia discontinua i mostra de la mescla en línia contínua; en tots dos casos el perfil més alt és el del pirimicarb).

Actualment (vegeu apartat 4) el mètode proposat s'està aplicant en altres problemes cromatogràfics. Per exemple, s'està realitzant l'estudi dels límits del mètode com a eina quantitativa, quan la resolució dels components coeluïts és molt baixa. L'interès d'aquests estudis està en la possible utilització com a eina de rutina, en aquells casos en què l'increment de la resolució cromatogràfica comporti costos massa alts, tant des d'un punt de vista econòmic, com de temps. Els sistemes que s'estudien estan relacionats amb problemes de tipus ambiental (separació i detecció de pesticides) i farmacològic (separació de mescles d'hormones).

Així mateix, una aproximació semblant a la que s'ha descrit per a la cromatografia s'està desenvolupant (apartat 4) actualment, pels mètodes d'anàlisi en continu i pels mètodes d'anàlisi d'injecció en flux (FIA, *flow injection analysis*). Els resultats preliminars mostren que, en condicions favorables,

l'estructura de les dades és de segon ordre trilineal, la qual cosa permet la determinació quantitativa acurada dels diferents components, en mescles complexes no resoltes, emprant mètodes de detecció multicanal (p. e., tècniques espectromètriques). S'estan desenvolupant mètodes alternatius *in-situ* i *on line* per a l'anàlisi de mescles complexes. Aquesta metodologia aprofita el coneixement acumulat sobre un gran nombre de reaccions químiques amb diferents graus de selectivitat, que es poden incorporar fàcilment a la tecnologia d'anàlisi automatitzada.

Darrerament, al nostre laboratori, hem iniciat l'aplicació del procediment descrit a la resolució dels senyals voltamperomètrics (vegeu apartat 4) obtinguts en l'estudi de les interaccions d'ions metàl·lics amb sistemes macromoleculars. Els resultats preliminars obtinguts fins ara són força esperançadors i poden permetre l'extensió del mètode a situacions més semblants a les naturals, on les concentracions dels ions metàl·lics són extremament baixes, a nivell de traça.

3.4. APLICACIÓ DEL MÈTODE MVCR A L'ANÀLISI DE PROCESSOS INDUSTRIALS DE PRODUCCIÓ QUÍMICA

El tercer exemple és l'aplicació del mètode a l'anàlisi de processos industrials de producció química.⁵⁷ Les dades experimentals s'obtenen en 8 realitzacions successives d'un procés químic industrial. Es recullen espectres FT-IR en diferents intervals de temps iguals, que descriuen l'evolució del procés. Cada realització conté entre 75 i 125 espectres, 795 en total. Les intensitats espectrals obtingudes en 66 canals de l'infraroig es guarden en una matriu de dades per a cada realització del procés, una fila per a cada espectre. La figura 34 mostra els espectres d'una de les 8 matrius processades. Les altres matrius tenien unes característiques similars. Els objectius de l'estudi eren la determinació dels components químics presents durant el procés global, la determinació de les seves respostes espectrals i la determinació dels perfils de variació de la seva concentració amb el temps. L'objectiu addicional era trobar un mètode que permetés la descripció de la situació del procés *on-line*, a qualsevol moment del procés.

Un primer problema que s'havia de resoldre era la presència d'una forta absorció de fons i de línia base, variable segons la matriu de dades considerada. Per disminuir la incidència d'aquests dos efectes, es va recórrer al pretractament de dades descrit a continuació. En primer lloc, per tal de disminuir l'efecte de la línia base sobre el punt inicial de cada realització del procés, es procedeix a la subtracció del primer espectre de cada matriu en els

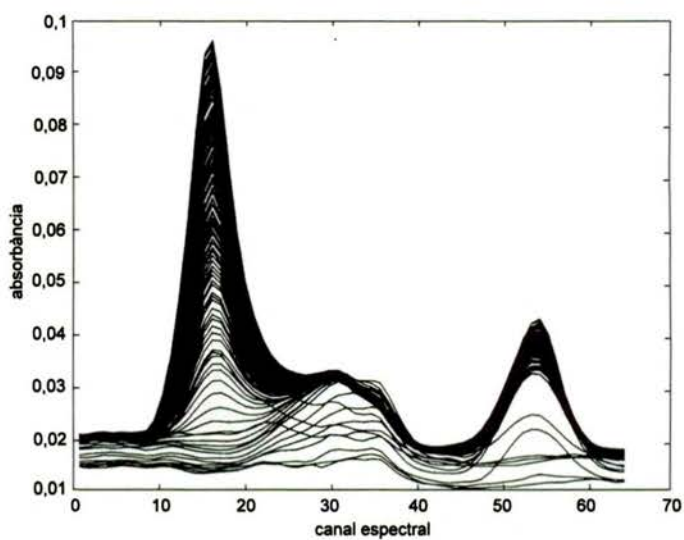


FIGURA 34. Exemple dels espectres d'absorció FT-IR adquirits durant el seguiment d'una realització del procés.

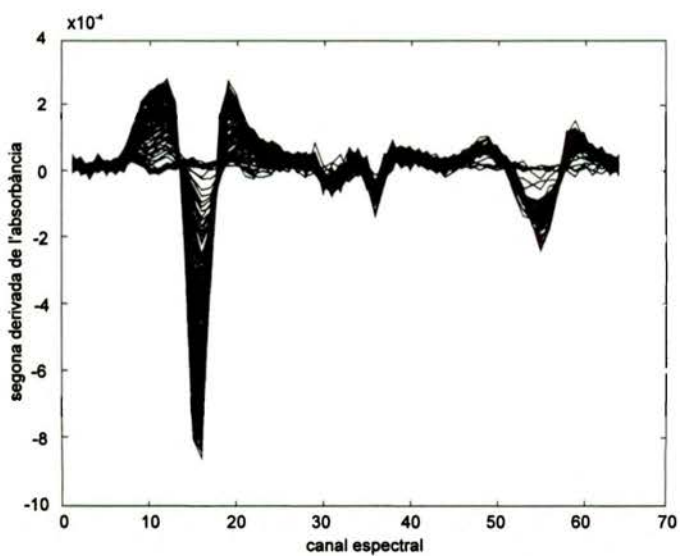


FIGURA 35. Espectres de la 2a derivada dels espectres de la figura 34.

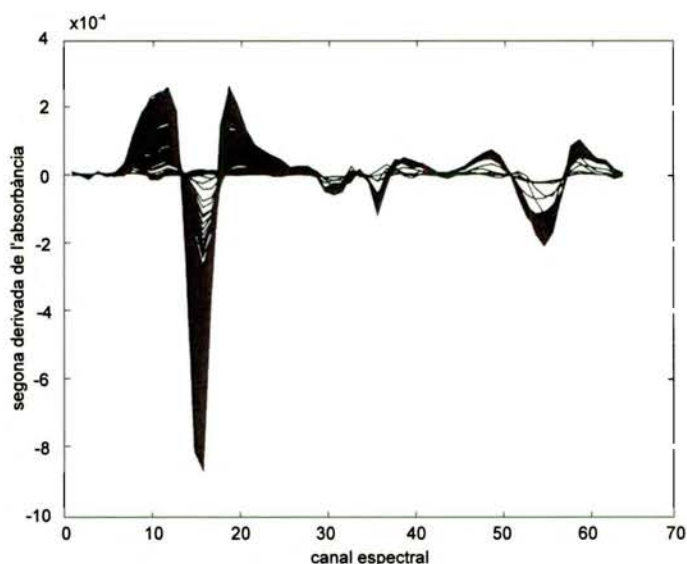


FIGURA 36. Espectres calculats de la 2a derivada (Figura 35) considerant tres components principals.

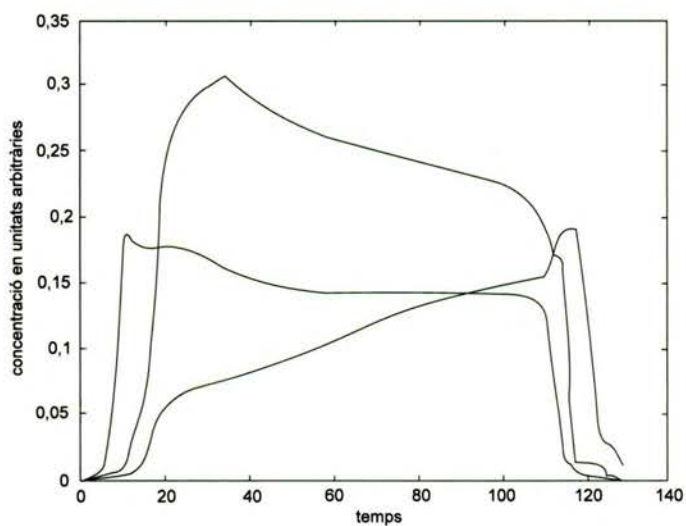


FIGURA 37. Estimació inicial dels canvis de concentració durant una realització del procés a partir de l'anàlisi de l'evolució dels factors (EFA).³⁹

espectres següents de la mateixa matriu. Amb aquest tractament el primer espectre de cada matriu serà 0. En segon lloc, per tal de disminuir l'efecte dels canvis de la línia base i del soroll de fons en una determinada matriu de dades, es calcula la primera i la segona derivada. Aquest pretractament minimitza les contribucions constants al llarg de l'espectre (1a derivada) i també aquelles contribucions que produeixen pendents constants en cada espectre (2a derivada). A la figura 35 es mostren els espectres de 2a derivada corresponents a la figura 34. Com es pot veure, han estat eliminades, la major part de les contribucions de la línia base.

L'estimació del nombre de components espectrals precisava la determinació prèvia de la varianza, associada amb les contribucions del soroll i la línia base. Per aconseguir-ho es calculen els valors propis corresponents a les zones espectrals, on només aquestes contribucions eren presents i no hi havia contribució de les bandes espectrals d'interès. Aquests valors propis es comparen amb els obtinguts a l'anàlisi dels espectres complets. A partir de la seva comparació (tenint en compte les diferents dimensions dels conjunts de dades comparats, emprant els valors propis reduïts, tal com proposa Malinowski,⁸⁰ es va arribar a la conclusió que el nombre màxim de contribucions espectrals independents eren dues o tres. Aquest nombre és només una proposta d'estimació inicial. Posteriorment, el nombre de tres components es va confirmar a partir dels resultats de l'anàlisi de components principals, de l'anàlisi de l'evolució dels factors i dels resultats de l'optimització per mínims quadrats alternats (taula 6). A la figura 36, es mostren els espectres de la figura 35 reproduïts quan es consideren tres components principals. Com es pot veure, la major part de la variació espectral és descrita amb 3 components principals i només s'ha prescindit de la part que pot considerar-se associada a l'error experimental. A la figura 37, es mostra un exemple de l'estimació dels perfils de concentració, obtinguts per aplicació del mètode de l'evolució dels factors³⁹ trobats a l'anàlisi d'una de les vuit matrius de dades. L'anàlisi simultània es realitza, tant per a les vuit matrius de primera i segona derivades com per a les dades originals. Les restriccions aplicades són no-negativitat en les concentracions, no-negativitat en els espectres pel cas de les dades originals, i unimodalitat i igualtat d'espectre per les espècies comunes entre matrius. La restricció d'unimodalitat tenia un efecte important, sobretot en l'anàlisi individual de les diferents matrius. Atesa la forta superposició espectral i l'evolució del procés, es tractava d'evitar formes absurdes dels perfils de concentració, però permetre, alhora, variacions brusques locals degudes, per una banda, a les dificultats de la resolució i, per l'altra, a la mateixa naturalesa del procés. A la figura 38, es mostren les tres contribucions espectrals finalment resoltes per les dades originals, i a la figura 39, l'evolució d'aquestes contribucions al llarg del procés.

La descripció feta del procés va satisfer plenament la indústria que va encarregar aquest treball, ja que donava una descripció força acurada del que ocorria a la planta de producció. A partir de les contribucions espectrals de la figura 38, es varen calcular uns coeficients que, aplicats sobre l'espectre mesurat en un cert moment del procés, permetien d'avaluar la situació del procés en aquell mateix instant. Considerant que les 8 matrius analitzades cobrien el conjunt de situacions que es podien presentar en la producció, la utilització dels coeficients esmentats havia de permetre l'anàlisi de situacions futures. La utilització dels coeficients permetia la correcció quasi instantània de les incidències que es podien presentar durant el procés. Aquest treball ha estat publicat recentment⁵⁷ dins un projecte de col·laboració amb el Center for Process Analytical Chemistry de la Universitat de Washington (Seattle, USA).

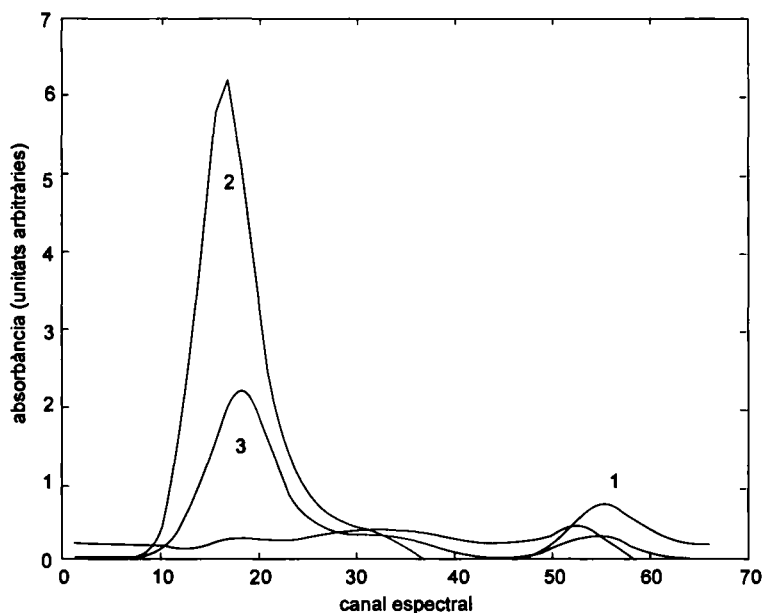


FIGURA 38. Espectres unitaris (purs) dels tres components detectats. Han estat calculats a partir de l'anàlisi simultània de vuit realitzacions del mateix procés (795 espectres).

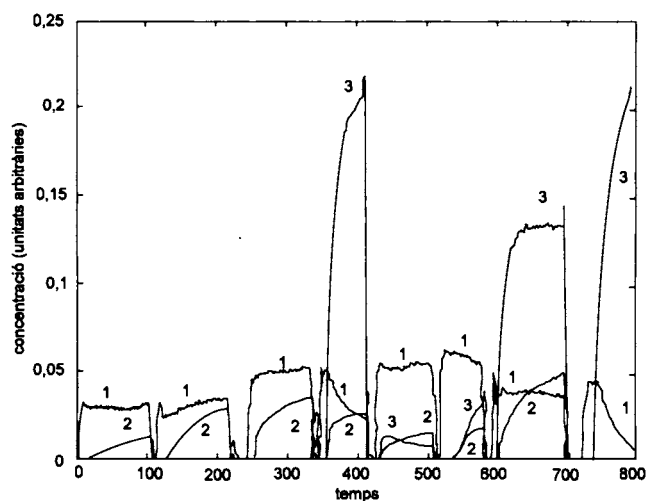


FIGURA 39. Canvis de concentració deduïts a partir de l'anàlisi de les vuit realitzacions del procés. L'eix de les abscisses (temps) dona una indicació del nombre d'espectre considerat dins del nombre total d'espectres considerats (795 espectres en total).

TAULA 6

Desviació estàndard dels residuals obtinguts en l'anàlisi de les dades^a

	2 comp.	3 comp.	4 comp.
exp - PCA ^b	$2,82 \cdot 10^{-3}$	$1,18 \cdot 10^{-3}$	$1,16 \cdot 10^{-3}$
PCA - calc ^c	$5,50 \cdot 10^{-3}$	$3,30 \cdot 10^{-3}$	$6,76 \cdot 10^{-3}$
exp - calc ^d	$6,19 \cdot 10^{-3}$	$3,80 \cdot 10^{-3}$	$6,86 \cdot 10^{-3}$

- Els resultats es donen per dos, tres i quatre components en l'anàlisi simultània de vuit realitzacions (vuit matrius) del procés.
- Desviació estàndard dels residuals entre la matriu de dades experimental i la matriu de dades reproduïda per components principals.
- Desviació estàndard dels residuals entre la matriu de dades reproduïda per components principals i la matriu calculada a partir dels perfils espectrals i de concentració trobats en l'optimització minimoquadràtica restringida.
- Desviació estàndard dels residuals entre la matriu de dades experimental i la matriu calculada, a partir dels perfils espectrals i de concentració trobats en l'optimització minimoquadràtica restringida.

3.5. APLICACIÓ DEL MÈTODE MVCR AL DESENVOLUPAMENT DE SENSORS ÒPTICS MULTIVARIANTS BASATS EN REACCIONS QUÍMIQUES

El desenvolupament de nous sensors químics és un camp d'enorme interès actual.⁸¹ L'aplicació dels mètodes quimiomètrics als sensors químics pot expandir de forma considerable la seva utilització. En aquest exemple es descriu l'aplicació del mètode de resolució multivariant de corbes al desenvolupament d'un nou tipus de sensor químic.

Un sensor multivariant basat en una reacció química acobla una reacció química selectiva amb una tècnica d'anàlisi multicanal (per exemple, de tipus espectroscòpic). A la figura 40 es mostra un exemple de l'esquema d'un sensor òptic, basat en la reacció química de Fujiwara,⁸² per a la determinació de mesclades de compostos organoclorats en aigües residuals subterrànies. Segons aquesta reacció, els hidrocarburs halogenats reaccionen amb la piridina en medi bàsic i donen espècies diferents. Algunes d'aquestes espècies són característiques de cada analit i d'altres són comunes entre els diferents analits. Els tres analits estudiats són el 1,1,1-tricloroetà (TCA), el tricloroetilè (TCE) i el cloroform.

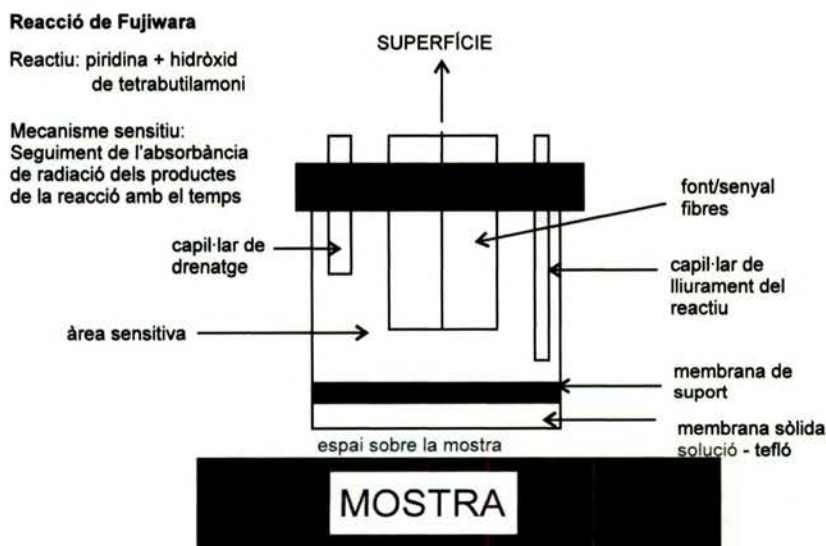


FIGURA 40. Sensor òptic basat en la reacció química de Fujiwara.

Encara que els tres analits per si mateixos no absorbeixen, en el UV, les espècies producte de la reacció de Fujiwara, poden detectar-se amb espectroscòpia UV/VIS. La reacció no és, però, selectiva per compostos organoclorats similars i hi ha un grau molt gran de superposició i correlació en els dos possibles ordres de resposta, que són l'espectral i el temps de reacció. Això fa que els mètodes tradicionals de calibració no es puguin emprar en aquest cas.

El mètode de resolució multivariant de corbes s'ha aplicat, amb èxit, a la resolució dels perfils espectrals i temporals de les diferents espècies produïdes en la reacció de cada analit. Els resultats d'aquesta anàlisi han permès, a més, l'estudi de les reaccions cinètiques associades i el desenvolupament d'un mètode quantitatiu d'anàlisi de mesclures. A continuació es presenten alguns dels resultats assolits.⁸³⁻⁸⁴⁻⁸⁵ És interessant recalcar que els resultats obtinguts fins ara demostren la possible utilització del mètode proposat en sensors químics, basats en qualsevol tipus de reacció química, sia de tipus cinètic o de tipus en equilibri.

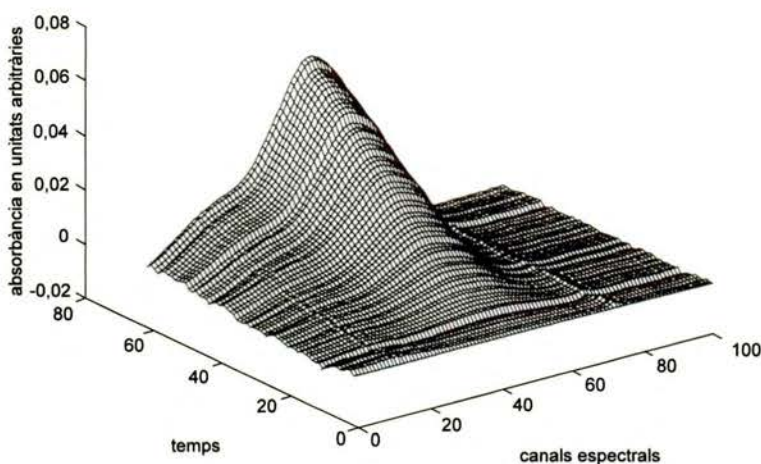


FIGURA 41. Resposta 3D del tricloroetà, TCA, (66,9 ppm) en la reacció de Fujiwara.

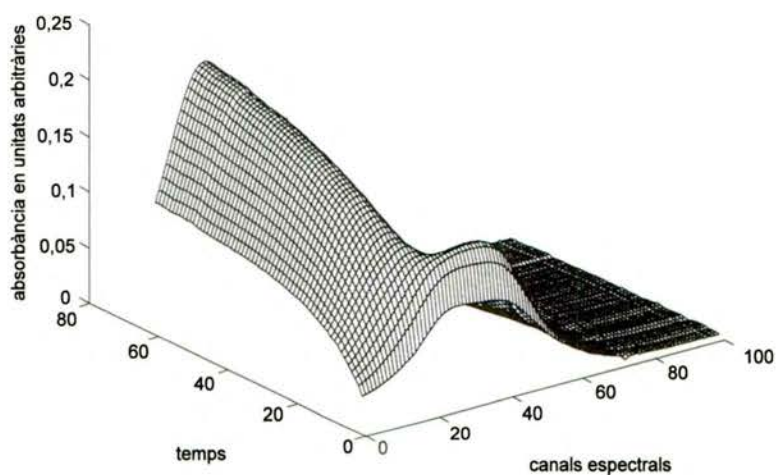


FIGURA 42. Resposta 3D del tricloroetè, TCE (7,31 ppm), en la reacció de Fujiwara.

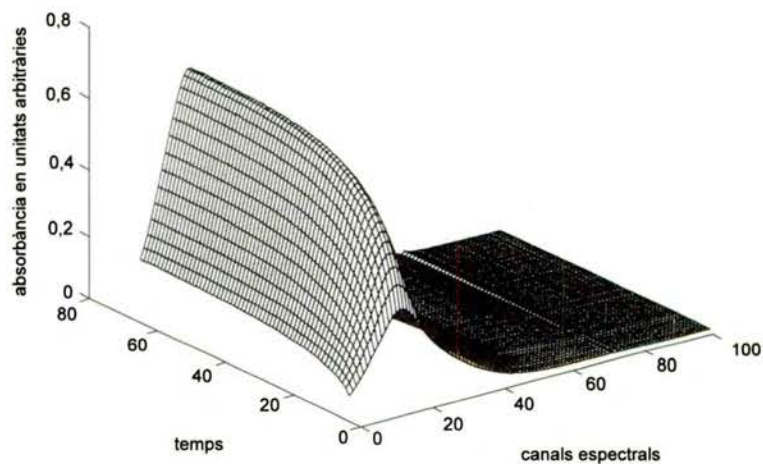


FIGURA 43. Resposta 3D de cloroform, CHCl_3 (4,45 ppm), en la reacció de Fujiwara.

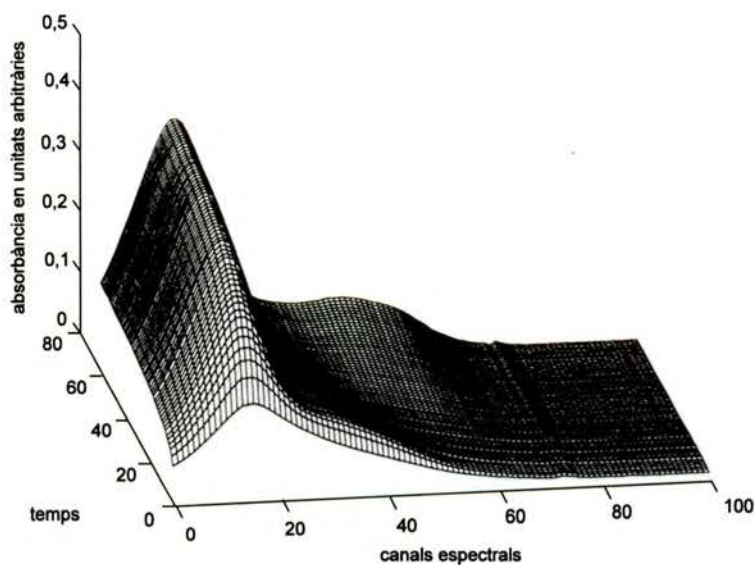


FIGURA 44. Resposta 3D d'una mescla de TCA (66,9 ppm), TCE (2,93 ppm) i CHCl_3 (1,48 ppm) en la reacció de Fujiwara.

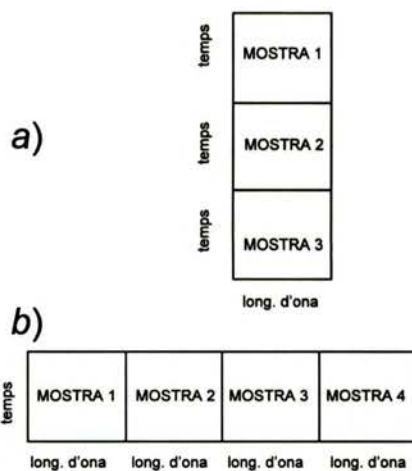


FIGURA 45. a) Matriu augmentada segons les columnes (on es mantenen les columnes o longituds d'ona) formada a partir de tres matrius individuals, corresponents a tres mostres analitzades. b) Matriu augmentada segons les files (on es mantenen les files o temps de reacció) formada a partir de les tres mateixes matrius individuals.

L'experiment consisteix, bàsicament, en l'adquisició ràpida d'espectres en el UV que reflecteixen l'evolució de la reacció de l'analit amb el reactiu. A les figures 41, 42 i 43 es mostren les variacions de les respostes espectrals amb el temps, corresponents a l'anàlisi dels tres analits TCA, TCE i CHCl_3 , per separat. I a la figura 44 es mostra la variació de les respostes espectrals, amb el temps corresponent a l'anàlisi d'una mescla dels tres analits. Cada mostra analitzada proporciona una matriu de dades de dimensions nombre d'espectres per nombre de canals espectrals (long. d'ona). Les finalitats de l'estudi són la determinació del nombre d'espècies químiques formades en les reaccions, la determinació dels seus perfils espectrals i temporals, i la determinació de la quantitat inicial d'analit en les mostres analitzades, abans que es produeixi la reacció. El problema analític és, realment, un repte de primera magnitud.

En primer lloc, es va realitzar l'anàlisi del rang de les matrius de dades i de la complexitat del sistema químic en estudi. Els resultats de l'anàlisi del rang de les diferents matrius de dades permeteren deduir que el TCA només donava una espècie espectroscòpicament activa, mentre que el TCE i el cloroform en donaven dues. Així es va arribar a la conclusió que en una mescla dels tres analits hi hauria, com a màxim, 5 espècies espectroscòpicament actives. El rang de les matrius augmentades, seguint diferents esquemes (figura 45) va confirmar que es formaven les mateixes espècies independentment de quines fossin les concentracions inicials d'analit i que no hi havia interacció entre ells en les mostres mescla. A més, es va confirmar la possibilitat d'estructura de dades de segon ordre trilineal, és a dir, que el perfil temporal de variació de la concentració d'una determinada espècie tenia la mateixa forma en totes les matrius analitzades. Òbviament, també l'espectre d'una determinada espècie era sempre igual.

Una de les assumpcions bàsiques del mètode és que la formació de les diferents espècies en les reaccions depèn, solament, de les concentracions inicials de l'analit. Quan aquesta dependència és lineal, s'està en condicions cinètiques de primer o pseudoprimer ordre. Aquesta situació s'aconseguia, fàcilment, en les condicions de treball del sensor d'excés de tots els reactius, excepte de l'analit.

En l'anàlisi de les matrius dels analits purs es varen trobar els perfils temporals i espectrals de les diferents espècies, formades a la reacció. A les figures 46, 47 i 48 es donen els perfils temporals de les espècies formades pel TCA, el TCE, i el cloroform, quan s'analitzen les diverses matrius de cadascun d'aquests analits. A la figura 49 es donen els perfils espectrals de les mateixes espècies. Com es pot veure, en alguns casos la resolució dels perfils és força difícil, a causa de l'elevat grau de sobreposició i de la manca de selectivitat. Els tres

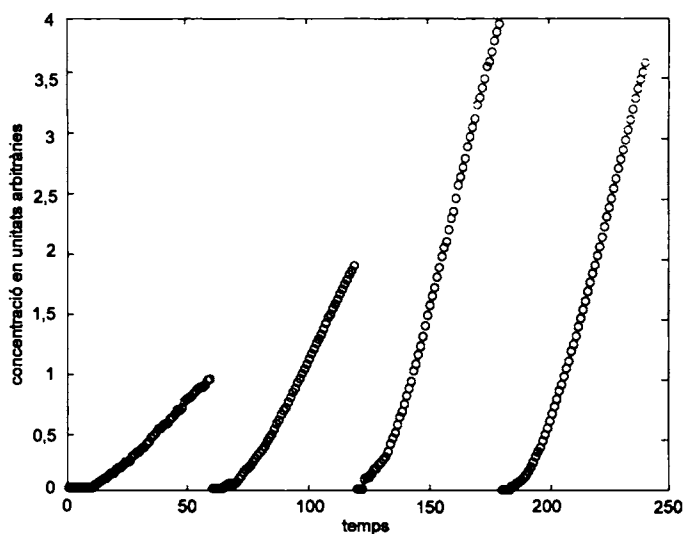


FIGURA 46. Perfils de concentració de l'espècie 1 (figura 49) trobats en l'anàlisi de 4 mostres pures de TCA.

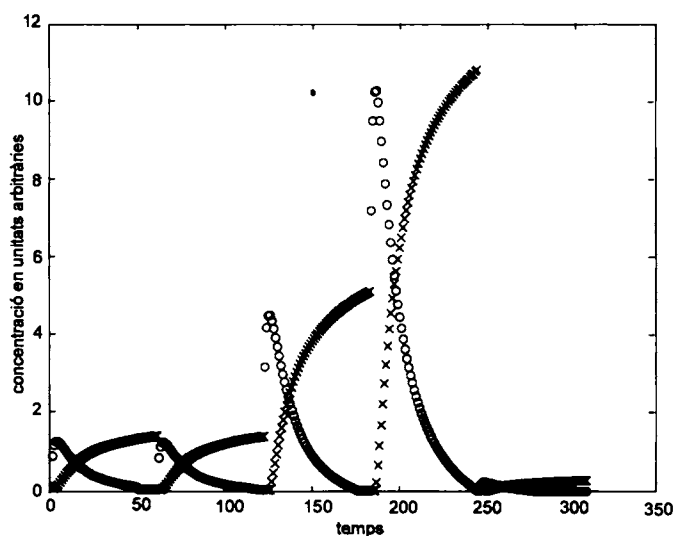


FIGURA 47. Perfils de concentració de les espècies 2 i 3 (figura 49) trobats en l'anàlisi de 5 mostres pures de TCE.

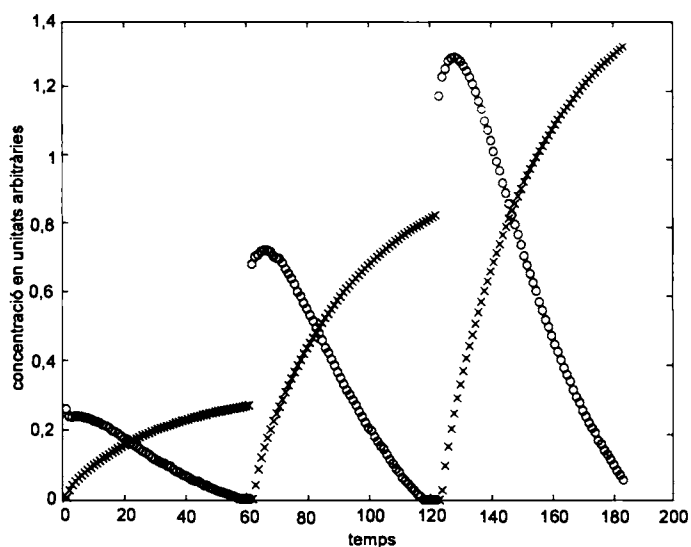


FIGURA 48. Perfils de concentració de les espècies 4 i 5 (figura 49) trobats en l'anàlisi de 3 mostres pures de CHCl_3 .

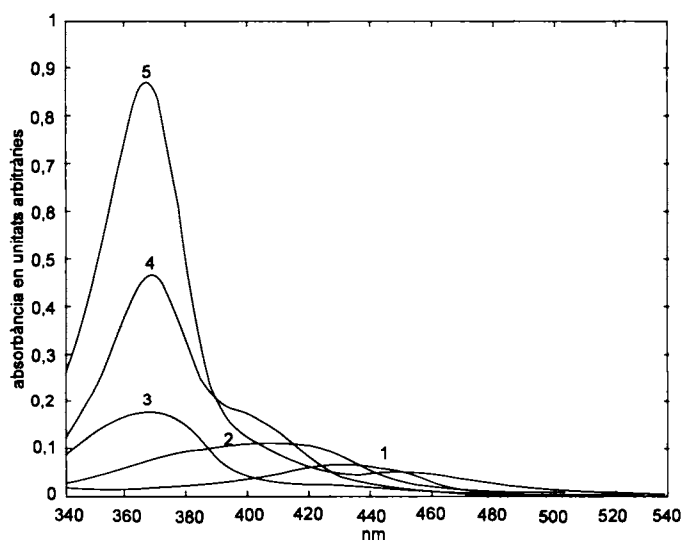


FIGURA 49. Espectres de les 5 espècies detectades en les mostres analitzades. L'espècie 1 és del TCA, la 2 i la 3 són del TCE; la 4 i la 5, del CHCl_3 .

analits tenen comportaments espectrals i temporals força semblants, com queda reflectit a la taula 7, on es comparen les similituds entre els diferents perfils espectrals i de concentració. A la mateixa taula es compara, també, la similitud entre les matrius resposta dels tres analits. La distinció entre la segona espècie del TCE i del cloroform està, fonamentalment, en la diferència de velocitat de formació. Per sort, la primera espècie del TCE té característiques espectrals diferents de la primera espècie del cloroform, la qual cosa va permetre la seva possible distinció. En canvi, el TCA dóna característiques espectrals diferenciades i un perfil de formació de velocitat més lenta. Per això, la determinació quantitativa del TCA és la menys problemàtica.

A la taula 8 es mostra un resum dels resultats de les determinacions quantitatives, realitzades sobre diferents tipus de mostres que contenen els tres analits, prenent-ne algunes com a patrons. Es varen analitzar 15 matrius de dades simultaniament. La quantificació del TCE és, generalment, millor quan s'utilitza la seva primera espècie. La quantificació de les mostres que contenen cloroform i TCE donen resultats pitjors amb errors en les concentracions estimades, que poden arribar fins al 30 %.

Finalment, el mètode s'ha aplicat a la determinació quantitativa d'un cert analit, en presència dels altres analits com a interferències. Les matrius augmentades es construeixen amb dues matrius de dades: la matriu que conté la mostra problema (analit + interferent) i la matriu que conté la mostra pura del patró. A la taula 9 es mostren els resultats obtinguts. Primer, es mostren els resultats de la determinació de TCE en una mescla, que conté, a més, TCA com a interferent. La determinació és, en aquest cas, bona. En el segon cas, es mostra la determinació del TCE en presència de cloroform com a interferent. La determinació quantitativa de TCE és bona, quan s'utilitza la primera espècie d'aquest analit (espècie 2, a la figura 45). Els resultats serien pitjors si s'escollís l'espècie segona de la reacció del TCE (espècie 3, a la figura 45), ja que aquesta última és molt similar a la segona espècie corresponent al cloroform (espècie 5, a la figura 45). El tercer cas és la determinació de cloroform, en presència de TCE com a interferència. Els resultats són ara pitjors, a causa de l'elevada superposició de l'espectre d'ambdues espècies que dóna el cloroform amb el segon espectre que dóna el TCE. Ens trobem, doncs, en els límits de la resolució del mètode i els resultats mostren que per millorar-los s'hauria d'incrementar la selectivitat del sistema. En investigacions posteriors, al Centre for Process Analytical Chemistry (CPAC) de la Universitat de Washington s'està intentant d'aconseguir millor selectivitat per modificació dels perfils temporals d'aquests dos analits. Una de les maneres és incrementar la selectivitat de la membrana, per la qual es difonen els analits (figura 40) a l'interior del sensor on es produeix la reacció. Una altra possibilitat és la modificació de la resposta cinètica, mitjançant canvis de temperatura.

TAULA 7

Similituds entre els tres analits TCA, TCE i CHCl_3 , els seus perfils temporals i els seus perfils espectrals

A) Perfils temporals^a

	TCA (1)	TCE (2)	TCE (3)	CHCl_3 (4)	CHCl_3 (5)
TCA (1)	1	0,15	0,93	0,30	0,94
TCE (2)	0,15	1	0,36	0,95	0,37
TCE (3)	0,93	0,36	1	0,56	0,999
CHCl_3 (4)	0,30	0,95	0,56	1	0,56
CHCl_3 (5)	0,94	0,37	0,999	0,56	1

B) Perfils espectrals^b

	TCA (1)	TCE (2)	TCE (3)	CHCl_3 (4)	CHCl_3 (5)
TCA (1)	1	0,02	0,86	0,09	0,88
TCE (2)	0,02	1	0,13	0,90	0,14
TCE (3)	0,86	0,13	1	0,32	0,997
CHCl_3 (4)	0,09	0,90	0,32	1	0,32
CHCl_3 (5)	0,88	0,14	0,997	0,32	1

C) Analits^c

	TCA	TCE	CHCl_3
TCA	1	0,71	0,68
TCE	0,71	1	0,998
CHCl_3	0,68	0,998	1

- Similituds entre els perfils cinètics temporals de les cinc espècies detectades (mateixa numeració d'espècies que en les figures 46-49). Es mesuren a partir del cosinus de l'angle que formen els vectors dels dos perfils temporals considerats.
- Similituds entre els perfils espectrals de les cinc espècies detectades (mateixa numeració d'espècies que en les figures 46-49). Es mesuren a partir del cosinus de l'angle que formen els vectors dels dos perfils espectrals considerats.
- Similitud entre els tres analits TCA tricloroetà, TCE tricloroetè i CHCl_3 cloroform. Es mesura a partir del cosinus de l'angle que formen els plans definits per les respostes en els ordres (temporal i espectral) de cada un dels dos analits considerats.

TAULA 8
Resultats de les determinacions quantitatives^a

Mostres ^b	Conc. analits ^c	Espècies ^d	Conc. rel. ^e	Conc. absol./ estàndard	Error ^f (%)
TCA	66,9	1	11,96	estàndard	
TCA	133,	1	24,64	137,6	3,4
TCA	267,	1	49,93	279,3	4,6
TCA	267,	1	45,73	255,8	4,2
TCE	7,31	2	32,80	7,54	3,1
		3	72,70	7,47	2,2
TCE	7,31	2	31,82	estàndard	
		3	71,14		
TCE	29,3	2	116,0	26,7	8,6
		3	246,8	25,3	13,0
TCE	58,6	2	270,6	62,2	6,1
		3	614,9	63,2	7,8
TCE	1,46	2	5,765	1,32	9,6
		3	12,47	1,28	13,0
CHCl ₃	1,48	4	20,13	estàndard	
		3	19,14		
CHCl ₃	4,45	4	44,49	3,27	26,0
		3	59,22	4,58	2,9
CHCl ₃	7,42	4	70,14	5,16	30,0
		3	92,89	7,18	3,2
TCA	66,9	1	10,99	61,5	8,1
TCE	2,93	2	14,57	3,35	14,0
		3	24,72	2,55	13,0
CHCl ₃	1,48	4	14,04	1,03	0,2
		5	24,72	1,91	29,0
TCE	2,93	2	17,10	3,93	34,0
		3			
CHCl ₃	4,45	4	41,28	3,03	32,0
		5	57,76	4,47	0,9
TCE	7,31	2	30,89	7,09	2,9
		3	59,10	6,07	17,0
CHCl ₃	1,48	4	15,14	1,11	25,0
		5	23,69	1,83	26,0

a. Resultats de l'anàlisi simultània de quinze mostres, utilitzant el mètode de resolució multivariant de corbes, amb la restricció d'igual forma sobre els perfils de concentració (estructura de segon ordre trilineal).

b. Contingut de les mostres analitzades.

c. Concentració dels analits (en ppm).

d. Espècies utilitzades per a la determinació quantitativa.

e. Concentració relativa de l'analit, estimada a partir del perfil de l'espècie corresponent.

f. Concentració absoluta estimada.

g. Error relatiu en la predicció.

L'exemple descrit mostra les grans possibilitats que ofereix l'aplicació de les tècniques quimiomètriques, al desenvolupament de nous sensors i la possibilitat d'incorporar, en aquest desenvolupament, tots els coneixements bàsics acumulats durant anys, sobre la química en solució i les reaccions químiques en solució.

TAULA 9
Resultats de les determinacions quantitatives en presència d'interferències^a

Mostra ^b	Conc. analit ^c	Espècies ^d	Conc. rel. ^e	Conc. abs. ^f	Error ^g (%)
TCE	7,31	2	22,99	estàndard	17
TCE (CHCl ₃)	7,31 (1,48)	2	19,00	6,04	
TCE	7,31				8,0
TCE	7,31	2	17,98	estàndard	
TCE	7,31	2	16,55	6,72	
(TCA)	(66,9)				
CHCl ₃	4,45	3,5	28,98	estàndard	21,
CHCl ₃	4,45	3,5	22,87	3,51	
(TCE)	(2,93)				

a Resultats de les anàlisis en presència d'interferències, utilitzant el mètode de resolució multivariant de corbes. S'analitzen, simultàniament, dues mostres: la mostra de patró pur de concentració coneguda i la mostra desconeguda, que conté l'analit i/o les interferències (en parèntesis).

b. Contingut de les mostres analitzades.

c. Concentració dels analits (en ppm).

d. Espècie emprada en la determinació quantitativa.

e. Àrea sota el perfil de concentracions de l'espècie corresponent.

f. Concentració estimada de l'analit.

g. Error relatiu en la predicció (en tant per cent).

3.6. CONCLUSIONS

L'aplicació del mètode de resolució multivariant de corbes s'ha demostrat mitjançant diversos exemples reals, d'anàlisi de mesclades de components químics, sense la seva separació prèvia. En concret, es demostra l'aplicació del mètode en els següents casos: *a*) estudis fonamentals sobre reaccions químiques en solució, mitjançant valoracions espectroscòpiques; *b*) estudis fonamentals sobre el comportament de les macromolècules en solució, sense necessitat de postulació de models predeterminats; *c*) determinacions analítiques de components

coeluïts en cromatografia o de components no resolts, en altres metodologies analítiques; *d)* seguiment *on line* i *in situ* de processos industrials de producció química; i *e)* desenvolupament de nous sensors multivariants, basats en reaccions químiques. En tots aquests casos, el mètode proposat ha permès la detecció, la resolució i la determinació quantitativa dels components presents en mescles no resoltes.

4. DESENVOLUPAMENTS RECENTS DEL MÈTODE DE RESOLUCIÓ MULTIVARIANT DE CORBES

El desenvolupament i l'aplicació del mètode descrit en els capítols anteriors ha continuat darrerament. A continuació, es presenta una explicació resumida dels aspectes més rellevants desenvolupats i les seves aplicacions.

4.1. MILLORA DEL PROCEDIMENT D'OPTIMITZACIÓ DE MÍNIMS QUADRATS ALTERNATS RESTRINGITS (OPTIMITZACIÓ ALS)

En primer lloc, cal esmentar la generalització de les restriccions imposades al procediment d'optimització per mínims quadrats alternats, ja descrites en els apartats anteriors i que han permès l'aplicació del mètode a un ventall més ampli de dades químiques. Aquestes nombroses aplicacions del mètode proposat, en constitueixen un dels avantatges més destacats, enfront d'altres mètodes similars descrits a la bibliografia i que, en general, no s'adapten tan bé a tants tipus de dades diferents. S'han generalitzat les restriccions de no-negativitat, de trilinealitat, d'unimodalitat i de sistema tancat, que ara es poden aplicar de forma independent per als perfils (vectors) que defineixen a cada espècie química, en qualsevol dels ordres de mesura i que, també, es poden emprar de forma independent, per a cadascuna de les matrius de dades examinades. Un diagrama de flux del mètode de resolució multivariant de corbes es dona a la figura 50. A la secció d'apèndix es donen els llistats dels programes MATLAB desenvolupats.

4.1.1. MILLORA DE LA RESTRICCIÓ DE SELECTIVITAT

En treballs recents⁸⁶⁻⁸⁷ s'ha confirmat que la resolució sense ambigüetats dels components en mescles no resoltes (*mixture analysis*), a partir de l'anàlisi

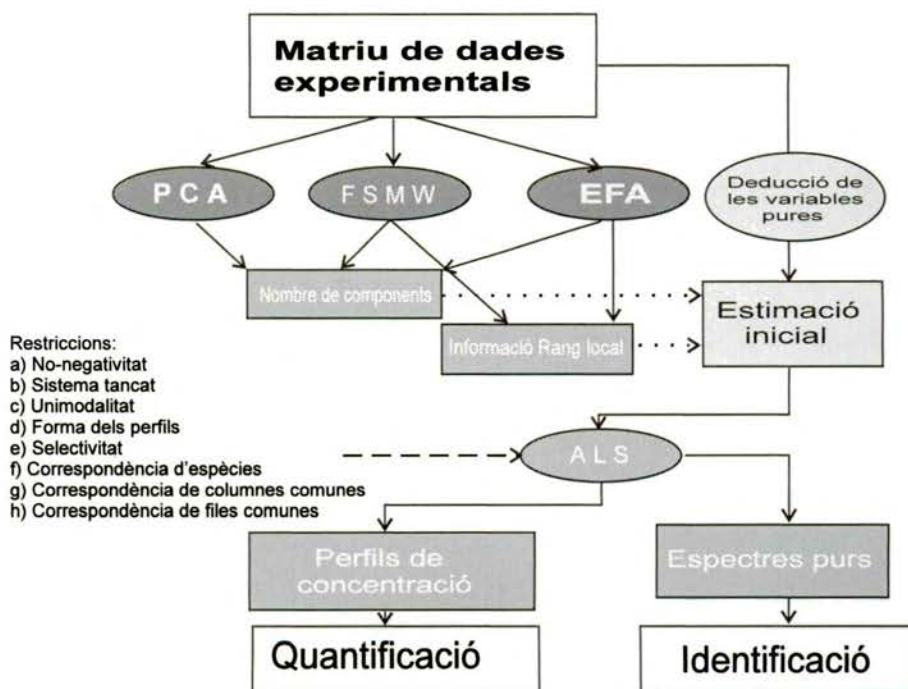


FIGURA 50. Diagrama de flux del mètode de resolució multivariant de corbes.

de matrius de dades, mitjançant mètodes d'anàlisi de factors (*factor analysis*) és possible sota certes condicions relacionades amb l'estructura particular de les dades analitzades i del seu rang químic local.⁸⁸ Aquestes condicions han estat formulades en forma de teoremes de resolució per Manne⁸⁷ i constitueixen una eina de gran importància teòrica. Han estat formulats en el context de les separacions cromatogràfiques, però són fàcilment extrapolables a altres situacions.

Teorema 1

Quan tots els components interferents que apareixen a l'interior de la finestra de concentració de l'analit (regió d'existència) apareixen, també, a fora d'aquesta finestra, aleshores és possible el càlcul del perfil de concentració d'aquest analit.

Teorema 2

Si per a cada interferent, la finestra de concentració de l'analit té una subfinestra, on l'interferent no hi és present, aleshores és possible de calcular l'espectre de l'analit.

Teorema 3

Per a aconseguir la resolució, només a partir de la informació proporcionada pel rang local en la direcció cromatogràfica, les condicions dels teoremes 1 i 2 no són solament suficients, sinó, també necessàries.

Aquests teoremes posen de manifest la importància que té el coneixement del rang local, de la matriu de dades experimental. Per a tal fi, es disposa de diverses tècniques, especialment d'aquelles tècniques que es deriven de l'anàlisi de factors emergents (*evolving factor analysis*),³⁹⁻⁴⁰ com és l'anàlisi de factors emergents amb una finestra mòbil de mida fixa (*evolving factor analysis wit a fixed size moving window*).⁶¹ Un cop conegut el rang local per aplicació de les esmentades tècniques, es pot comprovar si s'assoleixen les condicions estipulades pels teoremes anunciats per Manne⁸⁷ i deduir si és possible la resolució total o parcial del sistema en estudi. Cal tenir en compte, això no obstant, que la formulació dels teoremes és purament teòrica, i que en la pràctica poden sorgir dificultats derivades de la qualitat i la complexitat de les dades estudiades. Un punt feble de l'aplicació mecànica d'aquests teoremes està en la dificultat de definir exactament les finestres d'existència de totes les espècies del sistema, ja que a la pràctica, això es veu fortament afectat pel soroll instrumental i per la poca selectivitat de la resposta, associada a certes espècies. Una altra qüestió important a considerar és com afecta a la resolució l'aplicació d'altres restriccions, com ara la no-negativitat, la unimodalitat o la de sistema tancat. Això vol dir que les conclusions que s'extreuen de l'aplicació dels teoremes esmentats, no poden ésser definitives sobre la possible resolució del sistema, ja que no tenen en compte l'aplicació d'altres possibles restriccions naturals de les dades.

La resolució dels perfils cromatogràfics, a partir de la informació del rang local, no exigeix una selectivitat total en ells, la qual cosa està d'acord amb els resultats prèviament trobats per nosaltres.⁸⁶ A la figura 51 es mostren dues situacions hipotètiques de superposició de perfils d'elució cromatogràfica. En el primer cas, existeixen zones de selectivitat per el primer i el tercer component, però no n'existeix cap per el perfil cromatogràfic del mig. Els teoremes de resolució, això no obstant, ja diuen que serà possible la resolució total del

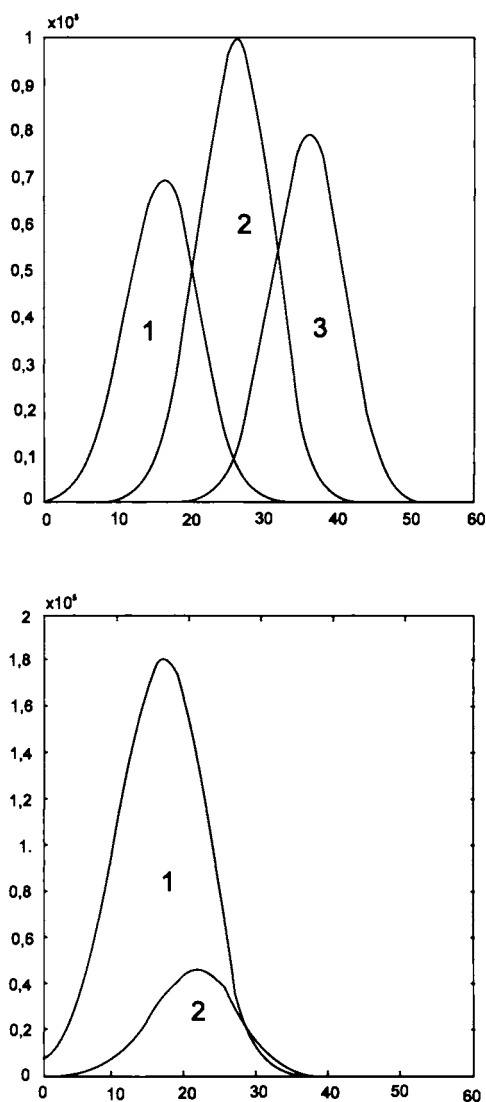


Figura 51. Superposició de perfils d'elució cromatogràfica. En el primer cas, existeixen zones de selectivitat per el primer i el tercer component, però no n'existeix cap per el perfil cromatogràfic del mig. En el segon cas, en canvi, el perfil del component minoritari es troba totalment inclòs (*embedded*) dins el perfil del component majoritari.

sistema, com efectivament s'obté a la pràctica, qualsevol que siguin els espectres dels tres components coeluits. En el segon cas, en canvi, el perfil del component minoritari es troba totalment inclòs (*embedded*) dins el perfil del component majoritari. En aquesta situació els teoremes preveuen que només és possible la resolució de l'espectre del component majoritari i la del perfil d'elució del component minoritari, però que no és possible la resolució del perfil d'elució del component majoritari ni la de l'espectre del component minoritari. Aquests dos últims perfils romanen no resolts, com a combinació lineal dels vertaders, on l'ambigüitat rotacional persistirà i no serà totalment resolta. Aquest resultat teòric està completament d'acord amb els resultats obtinguts per aplicació del mètode proposat, en el cas de l'anàlisi de dades amb perfils totalment inclosos en altres.

Els teoremes de resolució han estat formulats per a l'anàlisi individualitzada de matrius provinents d'un sol experiment. Cal, per tant, fer la seva extensió a l'anàlisi simultània de diverses matrius de dades correlacionades. Aquesta extensió és, especialment senzilla, per a les matrius augmentades segons les columnes, tal com les emprades en l'anàlisi simultània de diverses matrius correlacionades. De fet, les condicions estipulades per a la resolució de les espècies, presents en mescles complexes són molt més fàcils d'assolir quan s'analitzen diverses matrius correlacionades entre si, especialment si es realitza un bon disseny d'experiments. A més, només cal que les condicions de resolució es donin per a una de les matrius, ja que aleshores la resolució també serà possible per a les altres matrius incloses en l'anàlisi simultània, encara que en elles, les esmentades condicions no s'assoleixin. Actualment, s'està aprofundint en aquests aspectes teòrics.⁸⁹

4.1.2. MILLORA DE LA RESTRICCIÓ DE NO-NEGATIVITAT

Pel que fa a la restricció de no-negativitat, s'ha observat que en casos de resolució difícil, el procediment de mínims quadrats no-negatius basat en un algorisme de programació quadràtica (*quadratic programming*),⁹⁰⁻⁹¹ és més segur i proporciona resultats més fiables. Malauradament, l'aplicació d'aquest algorisme és molt més lenta que la de l'algorisme desenvolupat per nosaltres, que consisteix, simplement, a forçar durant el procediment iteratiu els perfils de concentració o espectrals a no ésser negatius, tot canviant-ne els valors negatius a zero. En la majoria de les ocasions, l'aplicació d'aquest darrer algorisme «forçat» proporciona resultats anàlegs al més sofisticat esmentat abans, i amb un guany de temps força considerable. És per això que, en les primeres etapes de l'anàlisi de dades, creiem que és recomanable l'aplicació del mètode

«forçat» i comprovar en les etapes darreres que el mètode més rigorós i sofisticat dóna resultats iguals.

Una segona qüestió important a considerar és l'aplicació de la restricció de no-negativitat, només sobre certs perfils i no en tots ells. D'aquesta manera és possible de modelar de forma natural (no-negativa) una part de la variació de les dades, deixant l'altra part que es modela de forma purament matemàtica (amb valors positius i negatius). Aquesta idea s'assembla a la proposada en el mètode anomenat *de bilinearització residual* (*residual bilinearization*)⁹¹ i pot fer possible l'extensió de l'aplicació del mètode, a aquells casos en què modelar totes les causes de variació de forma natural (no-negativa) resulta molt difícil. Actualment, aquest és, encara, un camp d'investigació.

4.1.3. MILLORA DE LA RESTRICCIÓ D'UNIMODALITAT

També té un interès especial la millora de la restricció d'unimodalitat. A part de poder-se aplicar individualment per a qualsevol dels dos perfils de cada espècie, es proposa la utilització d'una forma alternativa de l'algorisme, que permet de millorar els resultats obtinguts en casos de resolució difícil, com són la resolució de components minoritaris, a partir de mesures experimentals amb soroll alt. A continuació, s'indiquen els passos emprats per a forçar els perfils a ésser unimodals.

1. Es busca el màxim més gran $d_{\max}$ del perfil.
2. Es comprova l'aparició de màxims laterals a la dreta o a l'esquerra del màxim més gran. Es tolera un cert grau de pèrdua d'unimodalitat.
3. S'eliminen els possibles màxims laterals.

L'eliminació dels màxims laterals ha estat duta a terme de dues formes diferents. En la forma que fou proposada inicialment, l'eliminació es féu de manera radical, posant a zero els valors, a partir dels quals es detectava la formació d'un segon màxim en el perfil corresponent. En una segona implementació,⁹² en lloc de forçar a zero els valors del perfil, a partir dels quals es detectava la formació d'un segon màxim, el que es fa és no deixar-los créixer per sobre del valor anterior, al qual se li ha detectat l'inici d'un segon màxim. Aquesta nova formulació de la restricció d'unimodalitat ha resultat avantatjosa per als perfils dels components minoritaris, en presència d'un elevat nivell de soroll superposat.

4.1.4. MILLORA DE LA RESTRICCIÓ DE SISTEMA TANCAT

Pel que fa a les restriccions de sistema tancat, s'ha ampliat el nombre de possibles situacions, on aquest tipus de restricció pot ésser aplicada:

a) Aplicació de la restricció de sistema tancat, en només algunes de les espècies presents;⁹³ aquesta restricció té un interès per a aquells sistemes en els quals, només algunes de les espècies químiques detectades pel sistema de mesura, contenen el mateix reactiu i que, per tant, només per a aquestes espècies és possible de forçar-ne el corresponent balanç de matèria.

b) Aplicació alhora de més d'una restricció de sistema tancat; aquest és el cas de les espècies formades per la interacció de dos o més constituents, cadascun dels quals amb una concentració diferent i coneguda.

La restricció de sistema tancat pot definir-se de dues maneres diferents. La més corrent és la que força la suma de les concentracions de les espècies constituïdes per un determinat reactiu, a ésser igual a la concentració total d'aquest. Aquest és el cas, per exemple, de l'estudi de reaccions mitjançant tècniques de valoració; o del cas de reaccions cinètiques. Aquesta és la restricció que ja ha estat descrita en aquest treball. Un altre cas d'interès és aquell que utilitza les metodologies d'anàlisi en flux continu, on en cada moment la suma acumulada de les concentracions de les diferents espècies pot assolir, com a màxim, un determinat valor prefixat per la concentració inicial d'algun del reactius o bé per l'estequiometria de la reacció, en particular.

Hi ha, també la possibilitat de considerar la condició de sistema tancat (*closure*) sobre els perfils de l'ordre de les columnes (generalment l'ordre de mesura espectral). Hi ha casos particulars en què l'ordenació per files i columnes de la matriu de dades, descrita en els apartats anteriors, no és adequada, aleshores, és millor l'ordenació transposada, és a dir, col·locar la informació sobre la variació de la concentració a l'ordre de les columnes, i la informació sobre la variació de la mesura instrumental (espectral, voltamperomètrica, etc.) a les files. Un exemple d'això es troba en el cas d'examinar dades d'una valoració àcid-base per fluorescència,⁹⁴ on es mesuren els espectres d'emissió a diferents longituds d'excitació. En aquest cas, encara que es demostra que els perfils en els dos ordres tenen la mateixa forma, en les diferents matrius d'emissió, és millor de considerar l'ordre de variació de concentració en les columnes, ja que és el que es manté estrictament igual en intensitat i forma, per a totes les matrius analitzades. Mentre que els espectres d'emissió d'una mateixa espècie, a diferents longituds d'ona d'excitació, encara que tenen la mateixa forma, no tenen, òbviament, la mateixa intensitat i per tant, no són estrictament iguals. Més avall es descriu, amb més detall, aquest cas interessant.

4.1.5. MILLORA DE LA RESTRICCIÓ DE TRILINEALITAT O D'IGUALTAT DE FORMA DELS PERFILS

En el cas de l'anàlisi simultània de diverses matrius de dades correlacionades, l'estructura de les dades pot ésser de diferents tipus. Quan les dades tenen una estructura trilineal, és possible d'incloure la restricció de trilinealitat (igualtat de forma dels perfils de cada espècie en els dos ordres) durant l'optimització ALS, dels perfils de concentració. Aquesta restricció es pot aplicar per a totes o, només per a algunes, de les espècies escollides inicialment, d'entre totes les que componen el sistema. Aquesta possibilitat permet de millorar, per exemple, l'aplicació del mètode a la quantificació de mescles, en sistemes en els quals la igualtat de forma dels perfils de concentració de la mateixa espècie, en les diferents matrius de dades simultàniament estudiades, només es compleix per a algunes de les espècies químiques presents, i no pas per a totes les espècies del sistema. Un exemple d'això es troba en sistemes de reacció en excés de reactiu.⁹³

4.1.5.1. *Restricció de perfils (espectrals) constants*

En alguns casos, pot ésser interessant de mantenir constants (sense optimització) els perfils (espectres) d'aquelles espècies que ja han estat identificades, els quals han estat determinats de forma independent i amb un alt grau de certesa. Aquesta restricció té un interès especial en l'estudi de sistemes complexos, en els quals alguns components són coneguts (amb perfils coneguts). Per exemple, en les etapes inicials de l'estudi de sistemes amb moltes espècies, algunes de les quals tenen espectres coneguts, és convenient, a l'inici de l'anàlisi, de fixar els espectres d'aquestes espècies i avaluar l'ajustament minimoquadràtic que es va assolint en afegir-n'hi de noves. Un cop trobat un joc d'espècies (i dels seus perfils en els dos ordres) que ajusten, de forma satisfactòria les dades experimentals, es realitza l'optimització global, variant tots els perfils de totes les espècies alhora i comprovant la recuperació dels perfils que eren ja coneguts. Aquest procediment ha estat molt útil en l'estudi de les interaccions en sistemes macromoleculars entre diferents constituents (p. e. estudis d'interacció entre el bipiridil i l'àcid heteropolinucleòtid poli(A)-poli(U)).⁹⁵

4.1.5.2. *Restricció de perfils amb formes prefixades per una equació matemàtica*

Com ja s'ha indicat, les possibles restriccions que es poden aplicar depenen, fonamentalment, del tipus de dades estudiades. En alguns casos, se sap

que els perfils de resposta instrumental o de concentració han de tenir una forma determinada. Si aquesta forma determinada es pot descriure, amb una equació paramètrica determinada i aquesta es coneix, és convenient l'aplicació de la restricció corresponent al compliment d'aquesta equació, durant l'optimització ALS. Això, és el que ja s'ha fet en l'estudi de dades d'origen voltamperomètric, en presència de fenòmens d'adsorció.⁹⁶ A causa de la presència d'aquests fenòmens d'adsorció, els perfils de resposta voltamperomètrica que es recuperen (vegeu-ho més avall) tenen formes totalment distorsionades. La forma teòrica que han de tenir els perfils voltamperomètrics es pot descriure bé amb una funció no-lineal asimètrica logística, que depèn de l'amplitud del pic, de la posició del pic, de l'amplada a mitja alçada de pic i del grau d'assimetria del pic. Els valors d'aquests quatre paràmetres, que millor s'ajusten a uns determinats perfils voltamperomètrics poden trobar-se per aplicació del mètode d'ajustament, per mínims quadrats no lineals de Marquardt.⁹⁷ Un cop optimitzats els valors d'aquests quatre paràmetres, es reconstrueix el nou perfil voltamperomètric, que substitueix l'estimació obtinguda en el procediment iteratiu ALS. Mentre que la forma del perfil voltamperomètric, optimitzada segons els quatre paràmetres descrits, es correspon amb la forma que es pot esperar que tingui un perfil voltamperomètric sense adsorció, el perfil obtingut durant l'optimització ALS és un perfil que portarà associada la deformació típica de l'adsorció. En l'aplicació del mètode descrit, la part del senyal corresponent a l'adsorció queda finalment descrita, conjuntament amb el soroll dins dels residuals. Actualment, s'està aprofundint en aquest tractament i generalitzant la seva possible aplicació. Un cas particular d'aplicació important pot ésser, per exemple, l'aplicació d'un procediment similar, per al cas dels perfils d'elució o de concentració cromatogràfics dels components minoritaris.

Una altra possibilitat que estem estudiant és la d'incloure les restriccions derivades del compliment de les lleis de l'equilibri i de la cinètica química, sobre els perfils de concentració d'algun dels components. El compliment d'aquestes lleis, força els perfils de concentració a tenir determinades formes. Per exemple en el cas cinètic, els perfils de concentració han de seguir una equació de tipus exponencial. Els casos en què pot ésser especialment útil la postulació de models són aquells en què la resolució és molt baixa. És per això, que s'està estudiant aquesta possibilitat. El procediment resultant serà un híbrid dels procediments tradicionals d'ajustament minimoquadràtic, basats en models químicofísics (*hard-modelling*)³¹⁻³² i dels procediments lliures de model (*soft-modelling*)⁵² com el descrit en el present treball, de resolució multivariant de corbes. No obstant això, en la majoria dels casos no és necessària l'aplicació de restriccions d'aquest tipus. Per exemple, en l'estudi de reaccions on intervenen substàncies macromoleculares és poc convenient esforçar-se en l'aplicació de

models d'equilibri o cinètics, ja que aquests només tenen una validesa per a cada posició d'enllaç i no poden emprar-se de forma macroscòpica, a causa dels efectes ja esmentats (polielectrolític, polifuncional i conformacional), els quals dificulten, enormement, la postulació d'un model quimicofísic adequat.

4.2. APLICACIONS RECENTS DEL MÈTODE DE RESOLUCIÓ MULTIVARIANT DE CORBES

4.2.1. SISTEMES MACROMOLULARS

A part de l'estudi dels sistemes formats per l'àcid poliadenílic (poli(A))⁷⁶ i polietilenimino (poli(E))⁷⁴ descrits a l'apartat 3 del present treball, els sistemes següents macromoleculars també han estat estudiats recentment pel mètode descrit: *i*) estudi de les propietats àcid-base i complexants d'ions metàl·lics d'àcids fúlvics naturals per espectrometria de fluorescència molecular;⁷³ *ii*) estudi de les propietats àcid-base, de la complexació de metalls i dels canvis conformacionals de diversos àcids homopolinucleòtids: àcid poliuridílic (poli(U)),⁷⁰ àcid poliinosínic (poli(I))⁷¹ i àcid policitosínic (poli(C))⁹⁸ mitjançant la combinació de tècniques d'espectrometria UV, de dicroisme circular i de fluorescència molecular; *iii*) estudi de les propietats àcid-base, de les propietats complexants de metalls i dels canvis conformacionals dels àcids heteropolinucleòtids poliinosínic-policitosínic (poli(I)-poli(C))⁹⁹ i poliadenosínic-poliuridílic (poli(A)-poli(U)).⁹⁵ En tots aquests casos s'han determinat les espècies presents, els seus diagrames de distribució i els seus espectres d'absorció, de dicroisme circular o de fluorescència. Especialment interessant és l'aplicació del mètode descrit en l'anàlisi de dades de dicroisme circular, ja que permet de detectar i resoldre millor, els canvis conformacionals de les diferents espècies macromoleculars. S'han fet estudis comparatius dels resultats obtinguts en l'anàlisi d'aquests sistemes¹⁰⁰ i dels resultats obtinguts en mesclades de solvents amb polaritats diferents.¹⁰¹ S'han comparat els resultats obtinguts per aplicació de mètodes basats en la postulació d'un model químic (*hard-modelling*) amb els obtinguts per aplicació de mètodes no basats en l'assumpció d'un determinat model quimicofísic (*soft-modelling*) en l'estudi de les propietats àcid-base dels polinucleòtids.¹⁰²

Actualment, s'estan estenent aquests estudis a l'àcid ribonucleic (RNA) i a l'àcid desoxiribonucleic (DNA). A més, s'ha iniciat l'estudi de les interaccions dels polinucleòtids amb altres tipus de lligands o molècules, que poden actuar com a intercaladors entre les diferents cadenes dels àcids po-

linucleòtids com ara el bipiridil,⁹⁵ l'etilendiamina o la fenantrolina. La presència o no d'interacció amb aquests intercaladors model permet de treure conclusions sobre el comportament de sistemes bioquímics i farmacològics similars. Una nova aplicació del mètode s'ha desenvolupat per a la interpretació de les corbes de fusió (*melting curves*) dels polinucleòtids¹⁰³ i d'altres sistemes macromoleculars (per exemple, de les proteïnes) en general, a partir de dades multivariants. Aquests estudis són, especialment útils, quan els canvis de temperatura produeixen, simultàniament, diversos canvis conformationals superposats amb els equilibris àcid-base, ja que en aquests casos les aproximacions univariants tradicionals (mesures a una sola longitud d'ona) fallen totalment. Finalment, cal esmentar, en aquest apartat, l'inici de l'estudi de les possibilitats del mètode, per a l'estudi dels canvis conformationals o plegament de les proteïnes.

4.2.2. SISTEMES CROMATOGRÀFICS

S'ha continuat l'aplicació i l'adaptació del mètode descrit en la resolució i la quantificació de components coeluits, en separacions cromatogràfiques. Aquesta aplicació s'ha fet en sistemes d'importància ambiental (anàlisi de pesticides i herbicides) emprant dades obtingudes de l'anàlisi de mostres reals. En primer lloc, s'estudiaren mesclades dels herbicides alaclor i metaloclor,¹⁰⁴ que presenten una coelució molt forta i espectres UV extremament similars. Les determinacions es realitzaren amb procediments de preconcentració en línia (*on-line*) sobre mostres naturals on els nivells de concentració d'aquests constituents són extremament baixos (ppb). La utilització del procediment de resolució multivariant de corbes s'ha ampliat a l'anàlisi de traces de pesticides organofosforats, en mostres d'aigües naturals emprades en exercicis europeus d'intercomparació de mètodes i de laboratoris.¹⁰⁵ El procediment de quantificació a partir de les àrees dels perfils d'elució, un cop resolt, s'ha ampliat emprant el mètode de corba de calibrat. És a dir, la mostra problema s'estudia al mateix temps que una sèrie d'estàndards sotmesos als mateixos procediments de preconcentració i en presència d'interferències. Els patrons analitzats cobreixen un ventall ampli de concentracions. Un cop s'ha assolit la resolució dels components i a partir de l'àrea dels perfils en les mostres de patrons o estàndards, es construeix una corba de calibrat que s'utilitza posteriorment, en la determinació de l'analit a les mostres problema. D'aquesta manera es guanya en precisió respecte al mètode d'anàlisi en què s'empra un sol patró. Els resultats assolits varen estar dins dels intervals de confiança, associats als valors finalment acceptats per l'organització de l'exercici d'intercomparació.

Recentment, s'ha realitzat la validació del mètode de resolució multivariant de corbes. És aplicat en dades cromatogràfiques. S'ha avaluat les possibilitats del mètode d'optimització per mínims quadrats alternats, en la resolució i la quantificació de mesclures d'hidrocortisona i prednisona,¹⁰⁶ que donen pics coeluits (i tenen espectres molt similars) en cromatografia líquida amb detecció multicanal UV de matriu de díodes. Les dades analitzades cobrien un ampli rang de possibilitats, des de resolucions cromatogràfiques properes a 1, fins a resolucions cromatogràfiques de 0,1, i des de nivells d'interferència (prednisona) del 10 % fins a nivells d'interferència del 0,1 %. D'altra banda, s'ha avaluat l'efecte que tenen les estimacions inicials (SIMPLISMA³⁵ o EFA³⁹) sobre la rapidesa i l'exactitud dels resultats. Les conclusions obtingudes en el treball permetien d'assegurar una bona resolució i una bona quantificació del component majoritari, en gairebé tots els casos. Pel que fa a la interferència, la seva resolució i quantificació només és problemàtica a concentracions baixes (1 % o per sota) i a resolucions baixes (0,1).

En un treball més teòric,⁸⁹ s'ha avaluat l'efecte que té l'estructura de les dades cromatogràfiques, sobre la resolució de pics cromatogràfics coeluits. Així, quan la reproducibilitat cromatogràfica entre mostres que contenen els mateixos analits en proporcions diferents és perfecte, tant pel que fa a la forma dels pics, com a la seva posició (temps de retenció) en els diferents cromatogrames (dades trilineals), la resolució i la quantificació es realitzen sense problemes, i els resultats assolits són comparables amb els que s'obtenen quan s'utilitzen els mètodes de resolució tensorial trilineal GRAM⁴⁶⁻⁴⁷ i TLD.⁴⁸ En el cas que les formes dels perfils d'elució d'un determinat analit en els diferents cromatogrames no es conservin (les dades no són trilineals), fet que s'observa freqüentment a la pràctica quan hi ha coelució, el procediment proposat proporciona encara molt bones resolucions i recuperacions de les quantitats d'analit en cada mostra. Finalment, en el cas que les formes dels perfils d'elució d'un determinat analit, en els diferents cromatogrames presentin la mateixa forma, però la seva posició relativa (temps de retenció) sigui diferent, s'ha adaptat el procediment de manera que pugui avaluar el desplaçament dels pics i el corregeixi, abans d'aplicar la restricció corresponent a la igualtat de forma dels perfils d'elució.

4.2.3. SISTEMES DE VALORACIÓ ESPECTROFLUORIMÈTRICA

La detecció fluorimètrica presenta una sèrie d'avantatges, que cal esmentar. De fet, proporciona un dels millors procediments d'obtenció de dades

amb estructura de segon ordre trilineal. Això és així, perquè els espectres d'emissió d'un mateix analit, obtinguts per excitació amb radiació monocromàtica a diferents longituds d'ona d'excitació, tenen la mateixa forma, però intensitats diferents. El canvi d'intensitat produït a les diferents longituds d'ona d'excitació, depèn de l'analit considerat. És per això, que l'anàlisi de mescles d'analits per fluorescència, pot proporcionar fàcilment dades d'estructura de segon ordre trilineal. Com ja s'ha dit, en aquestes condicions és possible la resolució i la quantificació de mescles d'analits, sense ambigüitats. En un treball, que actualment és en curs de publicació, s'han estudiat els equilibris àcid-base de l'àcid saliclic, mitjançant una tècnica de valoració espectrofluorimètrica,⁹⁴ emprant tres longituds d'ona d'excitació. Això vol dir, que en la valoració s'obtenen tres matrius de dades, les quals contenen la mateixa informació química expressada de forma diferent en els espectres d'emissió. L'aplicació del mètode de resolució multivariant de corbes ha permès la resolució del sistema sense ambigüitats i el càlcul acurat dels perfils de concentració de les espècies fluorescents, a partir del qual ha estat possible l'estimació, força acurada, de les dues constants àcid-base de l'àcid saliclic. Les possibilitats del mètode descrit es van comprovar, tant per a dades simulades, com per a dades obtingudes experimentalment.

4.2.4. SISTEMES DE VALORACIONS EN FLUX CONTINU I D'ANÀLISI PER INJECCIÓ DE FLUX (FIA)

La metodologia de valoracions espectrofotomètriques o espectrofluorimètriques, descrita als apartats anteriors ha estat modificada, per tal d'adaptar-la al treball en flux continu¹⁰⁷ i generar dades multivariants a partir de sistemes en reacció. A la figura 52 es mostra un esquema dels diferents procediments experimentals emprats. Amb aquesta metodologia, solucions amb diferents concentracions d'analit poden valorar-se seqüencialment, per simple substitució de la solució d'analit que flueix pel canal A. Aquest procediment millora la reproductibilitat i facilita l'automatització de l'adquisició de dades multivariants; i es va aplicar inicialment per a l'estudi de l'equilibri àcid-base de l'1,2-naftoquinona-4-sulfonat (NQS) en què es varen detectar tres espècies en el rang de pH estudiats (6,9-12,3). Per aplicació del mètode de resolució multivariant de corbes es varen determinar, el diagrama de distribució i els espectres de les tres espècies detectades. Posteriorment, es va estudiar per un procediment similar la reacció existent, entre diferents aminoàcids i el reactiu NQS.⁹³ Per a cada aminoàcid estudiat, es detecta la presència de dues

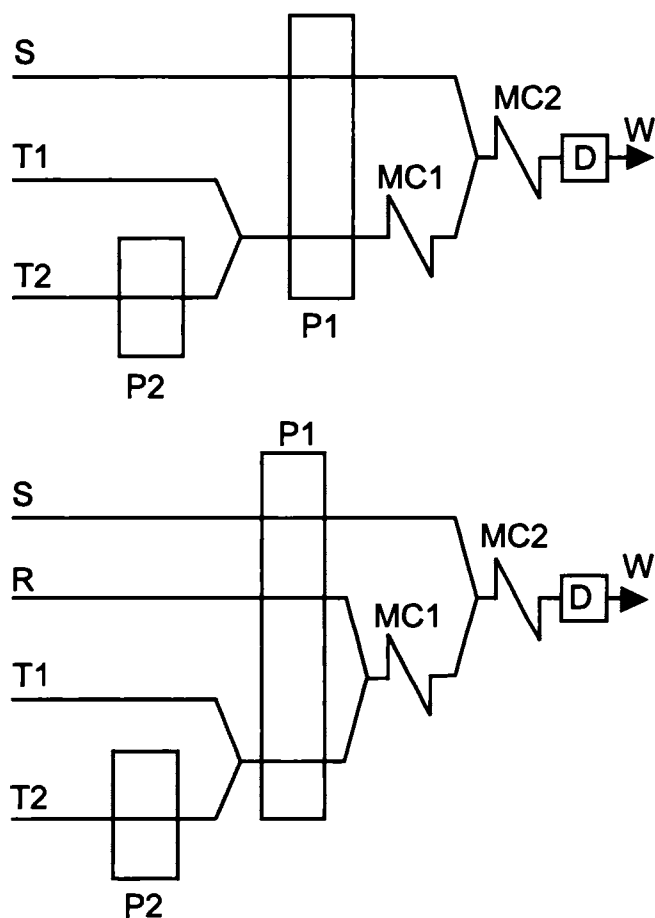


FIGURA 52. Esquema de diferents procediments experimentals d'anàlisi en continu.

espècies addicionals, que provenien de la reacció de derivatització entre el NQS i els aminoàcids, els diagrames de distribució i espectres de les quals també es trobaren per aplicació del procediment de resolució multivariant de corbes. L'aplicació del mètode de resolució es va estendre a l'anàlisi de mesclures d'aminoàcids per FIA (*flow injection analysis* o anàlisi per injecció de flux). El mètode va permetre la determinació quantitativa dels aminoàcids en mesclures d'ells, alguns dels quals considerats interferències.¹⁰⁸ Finalment, ha

estat proposat un mètode de parada de flux (*stopped flow*) per al seguiment de les reaccions cinètiques, que tenen lloc entre el reactiu NQS i els aminoàcids.¹⁰⁹ El procés se segueix espectrofotomètricament emprant un detector de matriu de díodes. Com que la reacció cinètica dels aminoàcids secundaris es més ràpida que la dels aminoàcids primaris, l'anàlisi multivariant de les dades, pel mètode de resolució proposat, ha permès la determinació simultània de mescles d'aquells dos tipus d'aminoàcids. Les respostes en els dos ordres, espectral i temporal, de les espècies, producte de la reacció dels aminoàcids amb el reactiu NQS, presenten certes diferències que permeten llur resolució i quantificació, àdhuc en presència d'interferències.

El desenvolupament de la metodologia adequada per a l'estudi dels sistemes en flux continu ha permès avançar en la posada a punt de les diferents restriccions, que s'apliquen sobre el mètode d'optimització per mínims quadrats alternats (ALS). D'altra banda, ha ajudat a comprendre millor l'estructura de les dades provinents de sistemes de reacció química, tant en equilibri¹⁰⁷⁻¹⁰⁸ com en variació cinètica,¹⁰⁹ així com també de les diferents variants possibles dels sistemes en flux, com són les valoracions en flux continu,⁹³⁻¹⁰⁷ injecció de flux¹⁰⁸ o a flux parat.¹⁰⁹ La discussió sobre les condicions on s'obtenen dades amb estructura de segon ordre trilineals o no trilineals i la discussió sobre les possibles variants de sistema tancat, han estat fonamentals en el desenvolupament darrer del procediment de resolució multivariant de corbes.

4.2.5. SISTEMES MESCLA AMB DEFICIÈNCIA DE RANG

Tenen un especial interès, els sistemes que produeixen matrius de dades amb deficiència de rang químic, és a dir, sistemes on el nombre de components químics que es poden resoldre, a partir de l'anàlisi de les seves matrius de dades experimentals, és inferior al nombre de components químics realment existents en elles. S'ha començat l'estudi sistemàtic d'aquests sistemes, a partir de l'anàlisi de les dades provinents de valoracions espectrofotomètriques àcid-base, de mescles de quatre bases nucleiques:¹¹⁰ adenina, citosina, hipoxantina i uracil. L'anàlisi individual d'una matriu de dades, provinents de l'anàlisi d'una mescla d'aquestes bases dona un rang igual a $n+1$, on n és el nombre de bases nucleiques presents en la mescla. Aquest nombre és inferior a $2n$, que és el nombre d'espècies espectroscòpicament actives, teòricament presents en els sistemes en estudi, ja que se suposa que cada base nucleica donarà dues espècies en la seva valoració àcid-base, una de protonada i una de desprotonada. Aquesta deficiència de rang es pot resoldre mitjançant l'anàlisi simultània de diverses valoracions, amb el mètode proposat de resolució

multivariant de corbes. La recuperació total del rang s'aconsegueix quan la valoració de la mescla de n bases nucleiques, s'analitza conjuntament amb altres $n-1$ valoracions, cadascuna corresponent a la valoració d'una base diferent de les que componen la mescla. Els resultats obtinguts per aplicació d'aquest mètode demostren, que la resolució total del sistema necessita rang complet, és a dir, que el rang químic detectat sigui igual al nombre de components realment presents. En canvi, la resolució i la quantificació de bases individuals, en presència d'interferències és encara possible, per aplicació del mètode de resolució multivariant de corbes (amb errors de predicció inferiors al 2 % en la majoria dels casos) àdhuc en el cas de deficiència de rang.

Actualment, s'estan estudiant mescles reals de bases nucleiques i s'està aplicant la metodologia desenvolupada, per a la seva determinació en fàrmacs. Matrius de dades, provinents de sistemes amb deficiència de rang es presenten freqüentment i el seu estudi és una de les línies de recerca actual.

4.2.6. SISTEMES VOLTAMPEROMÈTRICS

Hi ha molt poques aplicacions de tècniques quimiomètriques per a dades voltamperomètriques. Darrerament, el mètode de resolució multivariant de corbes ha estat aplicat a l'estudi de diferents problemes electroquímics. En primer lloc,¹¹¹ s'estudià la complexació del Zn(II) amb l'àcid polimetacrílic i l'àcid poliacrílic en solució aquosa. Matrius de dades obtingudes mitjançant diferents tècniques electroquímiques (polarografia d'impulsos normal, polarografia d'impulsos inversa, polarografia d'impulsos diferencial i voltamperometria de redissolució anòdica) i per diferents graus de desprotonació del lligand, foren analitzades de forma individual i simultàniament. Per aplicació del mètode proposat es varen determinar el nombre d'espècies presents en els sistemes, els seus diagrames de distribució en variar la concentració del lligand i els seus voltamperogrames individuals, així com també les constants d'estabilitat aparent K i les raons dels coeficients de difusió de les espècies individuals. Els resultats obtinguts estaven d'acord amb els obtinguts pels mètodes basats en l'assumpció d'un determinat model químicofísic.

Una segona aplicació, del mètode de resolució multivariant de corbes a dades voltamperomètriques ha estat l'estudi de la complexació del Cd(II), pel glutatíon (GSH) per polarografia d'impulsos diferencial (DPP).⁹⁶ Aquest sistema fou escollit com a model, en l'estudi de la complexació d'ions metàl·lics, amb pèptids que contenen àtoms de sofre. En aquest treball, es va implementar

una nova restricció sobre la forma dels perfils voltamperomètrics, que va permetre, per una banda, que aquesta forma fos l'esperada per la tècnica DPP, i per altra, banda, que els fenòmens de distorsió del senyal, associats amb processos d'adsorció quedessin en els residuals. Un estudi posterior de la complexació, en variar el pH¹¹² ha permès interpretar millor el mecanisme de complexació i els processos electròdics associats.

Finalment, s'han estudiat els límits del mètode proposat, en l'estudi de sistemes de complexació feble, entre l'ió Pb(II) i el propionat.¹¹³ Aquest sistema és especialment difícil de resoldre, a causa de la variació molt petita del senyal durant la complexació (voltamogrames molt similars) i també, a causa de la feblesa de la complexació, que es tradueix en un diagrama de distribució d'espècies, en què les espècies complexades tenen uns perfils de concentració totalment superposats, sense resolució. Malgrat tots aquests inconvenients, el mètode va permetre l'estimació d'unes constants d'estabilitat, similars a les obtingudes pels mètodes basats en l'assumpció prèvia d'un model fisicoquímic i basats en la llei d'acció de masses.

Tots aquests treballs ens han permès d'aprofundir en les característiques de les dades generades pels sistemes voltamperomètrics. Actualment, hi ha controvèrsia en els àmbits especialitzats sobre el grau de no-linealitat, associat a les mesures voltamperomètriques i sobre la validesa de l'aplicació dels mètodes, anomenats de modelat tou (*soft modelling*) en contraposició als mètodes anomenats de modelat dur (*hard modelling*) basats en l'assumpció prèvia d'un determinat model fisicoquímic. Les conclusions definitives d'aquesta discussió requereixen, encara, un estudi més exhaustiu de les diverses circumstàncies que concorren en l'obtenció de les mesures experimentals voltamperomètriques.

4.2.7. SISTEMES MESCLA DE CONTRIBUCIONS DE FONTS MEDIOAMBIENTALS

El mètode de resolució multivariant de corbes ha estat aplicat, recentment, a l'estudi de la distribució i caracterització de les contribucions de contaminants orgànics, en sediments marins obtinguts de la zona nord-oest mediterrània.¹¹⁴ Aquesta és una aplicació nova, totalment diferent a les descrites fins ara i obre la porta a altres possibles aplicacions futures, del mètode de resolució proposat. S'ha aplicat aquest mètode de resolució multivariant de corbes, a les matrius de dades constituïda per 31 i 51 mostres, sobre les quals es mesuraren 59 compostos orgànics (*n*-alcans, PCB, PAH, pesticides organoclorats) i 96 compostos orgànics (*n*-alcans, PCB, PAH, PAH que contenen sulfur,

pesticides organoclorats, esterols) respectivament. El tractament de les dades es féu en 4 etapes: 1) Pretractament i descripció estadística de les dades obtingudes; 2) Descripció de les dades, mitjançant anàlisi de components principals (PCA); 3) Anàlisi d'agrupaments jerarquitzat (*Hierarchical Cluster Analysis*, HCA); i factorització positiva de la matriu de dades, pel mètode de mínims quadrats alternats (ALS). El mètode ALS va permetre l'assignació de les tres possibles fonts majoritàries dels components orgànics detectats i va permetre, també, el càlcul de la seva contribució sobre cada mostra analitzada.

La consideració que les dades medioambientals, obtingudes per anàlisi química poden contemplar-se des d'una perspectiva d'anàlisi de mescles de contribucions o fonts, obra unes perspectives força interessants d'aplicació del mètode de resolució multivariant de corbes, presentat en aquest treball.

Les matrius de dades obtingudes en aquests estudis, també es poden descompondre en el producte de dues matrius, una d'elles associada als perfils que defineixen la composició de les fonts emissores dels compostos detectats, i l'altra associada als perfils que defineixen la distribució espacial o temporal de les contribucions d'aquestes fonts. L'èxit d'aquesta aplicació del mètode, dependrà dels resultats de les investigacions, que en aquest àmbit hem iniciat amb col·laboració amb el Departament de Química Ambiental del CID (CSIC) de Barcelona.

5. APÈNDIX

PROGRAMES MATLAB* DELS PROCEDIMENTS DESENVOLUPATS EN EL PRESENT TREBALL**

* MATLAB High-Performance Numeric Computation and Visualization Software. The Math Works Inc., 1992, Natick Mass. 01760, EUA

** En el desenvolupament i comprovació dels programes MATLAB que es presenten a continuació hi han col·laborat: Romà Tauler, Anna Izquierdo-Ridersa, Maribel Marqués, Anna de Juan, Xavier Saurina i Raimon Gargallo del grup de quimiometria del Departament de Química Analítica de la Universitat de Barcelona

5.1. PROGRAMA MATLAB PER A L'ANÀlisi DE L'EVOLUCIÓ DELS FACTORS MITJANÇANT UNA FINESTRA MÒBIL DE MIDA FIXA (*EVOLVING FACTOR ANALYSIS WITH A FIXED SIZE MOVING WINDOW FSMWEFA*)

```
function ef=eff(d,ns,nr)
% function evolving factor analysis; fixed size moving window
% d is the data matrix, ns is the fixed size of the moving window
% nr is the number of rows to be considered of the d matrix
```

```
disp '% *****'
disp '% MATLAB program EFF (evolving factor analysis      *)'
disp '% with a fixed size moving window, Keller, 61)      *)'
disp '% written by Roma Tauler, 1993                      *)'
disp '% University of Barcelona                          *)'
disp '% Department of Analytical Chemistry                *)'
disp '% Diagonal 647, Barcelona 08028                    *)'
disp '% e-mail roma@quimio.qui.ub.es                      *)'
disp '% *****'
```

```
clear ef;
hold off;
d=d(1:nr,:);
[nsoln,nwave]=size(d);
minn=min(nsoln,nwave);
```

```
i=1;
n=ns-1;
```

```
while i<=nsoln-n
% built fixed size matrix
x=d(i:i+n,:);
% disp('processing window number: '),disp(i)
l=svd(x);
inum=size(l);
for k=1:inum,if l(k)==0, l(k)=l(k-1); end, end
ef(1:n+1,i)=l(1:n+1,1);
efl(1:n+1,i)=log10(l(1:n+1,1));
i=i+1;
end
ef=ef';
efl=efl';
```

```

% set appropriately the zero log values
minvalue=min(min(efl));
minvalue=round(minvalue)-0.1*abs(round(minvalue));
[ns1,ns2]=size(efl);
for i=1:ns1,
    for j=1:ns2,
        if efl(i,j)==0,
            efl(i,j)=minvalue;
        end
    end
end

%plot the evolving factor analysis
% 1) the singularvalues
xforward=[1:nsoln-n];
disp('press any key to plot the efa results')
pause
plot(xforward,ef)
title('evolving factor analysis with fixed size moving window ')
pause

as=input('do you want to print it (y/n) ? ');
if as=='y', print; end

% 2) plot the log of singularvalues
maxvalue=max(max(efl));
maxvalue=round(maxvalue+1);
minvalue=minvalue;
nwindows=nsoln-n;
v=[1,nwindows,minvalue,maxvalue];
axis(v);
disp('press any key to plot the efa (log) results')
pause
plot(xforward,efl)
title('evolving (log) factor analysis with fixed size window ')
pause

as=input('do you want to print it (y/n) ? ');
if as=='y', print; end

axis
return

```

Notes:

svd és la funció MATLAB per al càlcul de tots els valors singulars d'una matriu de dades.

5.2. PROGRAMA MATLAB PER A L'ANÀLISI DE L'EVOLUCIÓ DELS FACTORS (*EVOLVING FACTOR ANALYSIS, EFA*)

```

function e=efa(d,ns)
% function evolving factor analysis efa
% d is the original matrix where evolving factor analysis is applied
% ns is the number of rows of this matrix d to be included (this allows
% to apply efa to submatrices)
% e is the abstract concentration profiles deduced by efa for the number
% of components are of interest given as input

disp '% *****'
disp '% MATLAB program EFA (evolving factor analysis      *'
disp '% written by Roma Tauler, 1993 (Gampp, 39,40)        *'
disp '% University of Barcelona                            *'
disp '% Department of Analytical Chemistry                 *'
disp '% Diagonal 647, Barcelona 08028                     *'
disp '% e-mail roma@quimio.qui.ub.es                       *'
disp '% *****'

x=d(1:ns,:);
[nsoln,nwave]=size(x);
minn=min(nsoln,nwave);

% forward analysis
% *****

disp('forward analysis')
n=2;
while n<=nsoln
disp('processing row number: '),disp(n)
l=svd(x(1:n,:));

% singular values, not eigenvalues!

nl=size(l);
ef(1:nl,n-1)=l(1:nl,1);
ef(ef==0)=ones(size(find(ef==0)))*1e-10;
efl(1:nl,n-1)=2*log10(ef(1:nl,n-1));
n=n+1;
end
ef=ef';
efl=efl';

```

```

% backward analysis
% *****

disp('backward analysis')
x=x(nsoln:-1:1,:);
n=2;
while n<=nsoln
disp('processing row number (backward): '),disp(n)
l=svd(x(1:n,:));
nl=size(l);

eb(1:nl,n-1)=l(1:nl,1);
eb(eb==0)=ones(size(find(eb==0)))*1e-10;
ebl(1:nl,n-1)=2*log10(eb(1:nl,n-1));
n=n+1;
end
eb=eb';
ebl=ebl';

% evolving factor analysis plots
% *****

% 1) plot the singular values (ef,eb)

xforward=[2:nsoln];
xbackward=[nsoln:-1:2];
hold off;
% disp('press any key to plot the efa results')
% pause
plot(xforward,ef,xbackward,eb)
title('evolving factor (svd) analysis')
as=input('do you want to print it (y/n) ? ');
if as=='y', print; end

% 2) plot the log of eigenvalues (ef ef^2, eb eb^2)

maxvalue=round(max(max(ebl)))+1;
minvalue=round(min(min(ebl)))-1;
[maxvalue,minvalue]
minvalue=input(' min. value of log efa plots ? ');
efl(efl<minvalue)=ones(size(find(efl<minvalue)))*minvalue;
ebl(ebl<minvalue)=ones(size(find(ebl<minvalue)))*minvalue;
plot(xforward,efl,xbackward,ebl)

```

```

title('evolving (log eig) factor analysis')
as=input('do you want to print it (y/n) ? ');

if as=='y',print;end

% 3) the arranged conc. profiles for the num. of factors

nf=input('num. of factors to be considered:? ');
axis;

while nf~=0

for j=1:nf
    jj=nf+1-j;
for i=1:nsoln-1,
    ii=nsoln-i;
    if ef(i,j)<=eb(ii,jj),
        e(i,j)=ef(i,j);
    else,
        e(i,j)=eb(ii,jj);
    end
    if e(i,j)==0.0, e(i,j)=1.0e-30; end
end
end

plot(xforward,e)
title('arranged efa (svd) profiles')
as=input('do you want to print it (y/n) ? ');
if as=='y', print; end
nf=input('other num. of components to be considered:? ');
end

e(2:nsoln,:)=e(1:nsoln-1,:);
end

```

5.3. PROGRAMA MATLAB DE CàLCUL DE LES VARIABLES MÉS PURES SEGONS EL MÈTODE SIMPLISMA

```
function [sp,imp]=pure(d,nr,f)

disp '% *****'
disp '% MATLAB program SIMPLISMA (W.Windig 35,38)   *'
disp '% written by Roma Tauler, 1993                *'
disp '% University of Barcelona                      *'
disp '% Department of Analytical Chemistry           *'
disp '% Diagonal 647, Barcelona 08028                *'
disp '% e-mail roma@quimio.qui.ub.es                 *'
disp '% *****'

% [sp,imp]=pure(d,nr,f)
% sp purest row/column profiles
% imp indexes of purest variables
% d data matrix; nr (rank) number of pure components to search
% if d(nspectra,nwave) imp gives purest nwave => sp are conc. profiles %
(nr,nspectra)
% if d(nwave,nspectra) imp gives purest nspectra => sp are spectra
% profiles (nr,nwave)
% f percent of noise allowed respect maximum of the average spectrum
% it is given in %

[nrow,ncol]=size(d);

% calculation of the purity spectrum

f=f/100;
s=std(d);
m=mean(d);
ll=s.*s+m.*m;
f=max(m)*f;
p=s./(m+f);

[mp,imp(1)]=max(p);

disp('first purest variable: ');disp(imp(1))

% calculation of the correlation matrix
```



```

l=sqrt((s.*s+(m+f).*(m+f)));
for j=1:ncol,
    dl(:,j)=d(:,j)./l(j);
end
c=(dl'*dl)./nrow;

% calculation of the weights
% first weight

w(1,:)=ll./(l.*1);
p(1,:)=w(1,:).*p(1,:);
s(1,:)=w(1,:).*s(1,:);
subplot(3,1,1),plot(m)
subplot(3,1,2),plot(s)
subplot(3,1,3),plot(m)
title('unweighed mean, std and pure first spectra')
pause

% next weights

for i=2:nr,

for j=1:ncol,
    [dm]=wmat(c,imp,i,j);
    w(i,j)=det(dm);
    p(i,j)=p(1,j).*w(i,j);
    s(i,j)=s(1,j).*w(i,j);
end

% next purest and standard deviation spectrum

% plot(p(i,:))
% plot(s(i,:))
% title('sd and purest spectrum')
subplot(2,1,1),plot(p(i,:))
subplot(2,1,2),plot(s(i,:))
title('next weighed std and pure first spectra')
pause
[mp(i),imp(i)]=max(p(i,:));

disp('next purest variable: ');disp(imp(i))
end

```

```
for i=1:nr,  
    impi=imp(i);  
    sp(1:nrow,i)=d(1:nrow,impi);  
end  
  
sp=normv2(sp');  
plot(sp')  
  
end
```

5.4. PROGRAMA MATLAB DE RESOLUCIÓ MULTIVARIANT DE CORBES PEL MÈTODE D'OPTIMITZACIÓ DE MÍNIMS QUADRATS ALTERNATS I RESTRINGITS (ALS)

```

function
[copt,sopt,rtopt,sdopt,areaopt]=alsc(d,x0,ils,nexp,isp,nit,tolsigma);
%
%***** Alternating Least Squares (ALS) *****
%
%   Syntax [copt,sopt,rtopt,sdopt]=als(d,x0,ils,nexp,isp,nit,tolsigma);
%           where
%
%   INPUT VALUES:
%       d : experimental data matrix
%       x0: initial estimation of the concentration profiles (ils=1)
%           or of the species spectra (ils=2)
%       ils: kind of initial estimate provided from efa;
%            1 for initial estimates of concentrations;
%            2 for initial estimates of spectra
%       nexp: number of experiments simultaneously analyzed
%       isp: correspondence between the species in the experiments
%       nit: maximum number of iterations (50 is the default)
%       tolsigma: convergence criteria in the difference in the residuals fitting
%                between iterations (0.1% is the default)
%
%   OUTPUT VALUES:
%       copt: optimized species concentrations
%       sopt: optimized species spectra
%       rtopt: residuals d - copt*sopt at the optimum
%       rtopt: ratio of total concentrations estimated in the quantitation
%       sdopt: standard deviation of fitting residuals at the optimum
%*****
%   different types of closure for every experiment !!!!
%*****
%   other important variables
%   nrow number of rows (spectra) in d
%   ncol number of columns (channels, wavelengths) in d
%   nsign is the total number of significative species
%   nit is the maximum number of allowed iterations in als
%   nspec number of species in each experiment
%   nselect number of selective points
%   ipselect point number with initial selectivity
%   iwselect channel number with initial selectivity

```

```

%      spselect species for which there is selectivity at ipselect
%      ishape = 0,1,2 data structure (see below)
%      nfix number of fixed spectra; ifix [0,1,...0] indices for fixed spectra

disp ('% *****')
disp ('% MATLAB program ALS (alternating least squares)      *')
disp ('% written by Roma Tauler, 1996                          *')
disp ('% University of Barcelona                                *')
disp ('% Department of Analytical Chemistry                      *')
disp ('% Diagonal 647, Barcelona 08028                           *')
disp ('% e-mail roma@quimio.qui.ub.es                            *')
disp ('% *****')

%*****%
PLOT THE RAW DATA

[nrow,ncol]=size(d);
plot(d')

% INIZIALIZATIONS

if nargin<6,nit=50;end
niter=0;
idev=0;
if nargin<7,tolsigma=0.1;end
ans='n';

% *****
while ans=='n'| ans=='N';
% *****

% APPLY SELECTIVITY WHEN ZEROS IN INPUT DATA
iselect=input('apply selectivity where input values are zero initially? yes(1),no(0)? ');

% INPUT INITIAL ESTIMATIONS FROM EVOLVING FACTOR ANALYSIS

% INITIAL ESTIMATION OF THE CONCENTRATION PROFILES
if ils==1,
    conc=x0;
    subplot(2,1,2);plot(conc);title('Initial estimation of the concentration profiles')
    [nrow,nsign]=size(conc);
    abss=conc\d;
    subplot(2,1,1);plot(abss');title('Initial estimation of the unit spectra')

```

```

% SELECTIVITY DETECTION; INPUT VALUES EQUAL TO ZERO
    if iselect==1,
        nselect=0;
        for ip=1:nrow,
            for ispecies=1:nsign,
                if conc(ip,ispecies)==0,
                    nselect=nselect+1;
                    ipselect(nselect)=ip;
                    spselect(nselect)=ispecies;
                end
            end
        end
    end
end

% OR INITIAL ESTIMATION OF THE UNIT SPECTRA
if ils==2,
    abss=x0;
    [ncol,nsign]=size(abss);
    subplot(2,1,1);plot(abss);title('Initial estimation of the unit spectra')
    abss=abss';
    conc=d/abss;
    subplot(2,1,2);plot(conc);title('Initial estimation of the concentration profiles')
    for k=1:nsign,
        for i=1:nrow,
            if conc(i,k)<0.0,
                conc(i,k)=0.0;
            end
        end
    end
end

% SELECTIVITY DETECTION; INPUT VALUES EQUAL TO ZERO
if iselect==1,
    nselect=0;
    for iw=1:ncol,
        for ispecies=1:nsign,
            if abss(ispecies,iw)==0,
                nselect=nselect+1;
                iwselect(nselect)=iw;
                spselect(nselect)=ispecies;
            end
        end
    end
end
end

```

```

close
% *****
% *****
% UNCHANGED SPECTRA
% *****

nfix=input('do you want to keep some input pure spectra fixed (y/n 1/0)
during the optimization? ');
if nfix>0,
    ifix=input('indexes of fixed (1) or changed (0) pure spectra, e.g [0,1,0,1..] ');
end
    clc;disp('***** Press any key to continue *****');

% *****
% initialization of the total concentrations
% *****

totalconc(1:nsign,1:nexp)=ones(nsign,nexp);

if nexp > 1,

    % IN SIMULTANEOUS ANALYSIS OF SEVERAL SAMPLES
    % ENTER NUMBER OF SPECTRA
    ninic(1)=1;
        disp('speciation');isp
        % CONSTRAIN THE CONCENTRATION PROFILES TO HAVE
THE SAME SHAPE
        % ishape = 0 no second order (different shapes of conc profiles)
        % ishape = 1 second order (same shapes of conc profiles and peak matching)
        % ishape = 2 second order (same shapes of conc profiles but no peak match)
        disp('no trilinear, different shape of conc. profiles, ishape=0')
        disp('trilinear, equal shape and synchronization (all species) ishape=1')
        disp('trilinear without synchronization (all species), ishape=2')
        disp('trilinear and synchronization for only some species, ishape=3')

    for i=1:nexp,

        disp(' ')
        disp('*****')
        disp(['Experiment number ', num2str(i)]);
        nsol(i)=input('Enter the number of spectra: ');
        if i>1, ninic(i)=nfin(i-1)+1; end
        nfin(i)=ninic(i)+nsol(i)-1;
    end
end

```

```

% NUMBER OF SPECIES AND CORRESPONDENCE IN isp
icount=0;

for j=1:nsign,
    if isp(i,j)>0,icount=icount+1;end
end

nesp(i)=icount;
% disp(['Number of species: ', num2str(nesp(i))])

ishape=input('ishape for peak matching ? ');

if ishape==3,
    disp('enter the vector of which are trilinear [1,0,1,...] 1=yes , 0=no ')
    spetri(1,1:nsign)=input('');
else if ishape==1 | ishape==2,
    spetri=ones(1,nsign);
end
end
end

else

% WHEN ONLY ONE EXPERIMENT IS PRESENT EVERYTHING IS CLEAR

nsol(1)=nrow;
ninic(1)=1;
nfin(1)=nsol(1);
nesp(1)=nsign;
isp(1,1:nsign)=ones(1,nsign);
ishape=0;

end

% *****
% GENERAL CONSTRAINTS
% *****
% NON-NEGATIVITY CONSTRAINTS
% *****
clc
disp(' ')
ineg=input('Enter the non-negativity constraints (0 = conc / 1 = conc and abs ): ');
disp(' ')

```

```

ialg=input('Enter the kind of algorithm nnls (0 = forced to zero / 1 = algoritm nnls) ');

% *****
% CLOSURE CONSTRAINTS
% *****
clc
disp(' ');
disp('types of closure; give input for every experiment')
disp('      0 = no closure')
disp('      1 = closure for the species concentrations')
disp('      2 = two closures for species concentrations')
disp('      3 = closure in column (spectra) order ')

tclos(1:nexp)=zeros(nexp,1);
tclos2(1:nexp)=zeros(nexp,1);
sclos1(1:nexp,1:nsign)=zeros(nexp,nsign);
sclos2(1:nexp,1:nsign)=zeros(nexp,nsign);
iclos(1:nexp)=zeros(1,nexp);
iclos1(1:nexp)=zeros(1,nexp);
iclos2(1:nexp)=zeros(1,nexp);

for i=1:nexp,
    disp(' ')
    disp(['experiment: ',num2str(i)]),
    iclos(i)=input('closure constraints (0/1/2/3) ');

    if iclos(i)==1 | iclos(i)==2,
        tclos1(i)=input('(first) closure constant is? ');
        iclos1(i)=input('first closure is an equal condition (1) or a lower or
equal than condition (2) ');
        sclos1(i,:)=input('which species are in (first) closure [1,0,1,...] ');

        if iclos(i)==2,
            tclos2(i)=input('(second) closure constant is? ');
            iclos2(i)=input('second closure is an equal condition (1) or
a lower or equal than condition (2)');
            sclos2(i,:)=input('which species are in (second) closure
[1,0,1,...] ');
        end
    end

    if iclos(i)==3,
        tclos(1)=input('closure constant in column (spectra) order ');

```



```

end
end

% *****
% UNIMODALITY CONSTRAINTS
% *****
clc
disp(' ')
disp('Types of unimodality constraints: ')
disp('      0 : none')
disp('      1 : only for the concentration profiles')
disp('      2 : only for the spectrum profiles')
disp('      3 : concentration and spectrum profiles')

imod=input('unimodality none(0), conc (1), spec(2), conc and spec (3) ');

if imod==1|imod==3,
    nmod=input('for how many species unimodality is applied ');

    if nmod==nsign,
        spmod=ones(1,nsign);
    else
        spmod=input('species with unimodality, e.g. [0,1,..] ');
    end

    rmod=input('unimodal constraint tolerance for the conc ');
end

if imod==2|imod==3,
    for i=1:nsign,
        disp('species spectra');disp(i),
        smod(i)=input('unimodal constraint tolerance for the spec ');
    end
end
clc
gr='n';
while gr ~= 'y',
    gr=input('Do you want graphics output during the ALS optimization (y/n)?', 's');
    if gr=='n'| gr=='N',
        gr='n';break;
    elseif gr=='Y';
        gr='y';break;
end

```

```

        end
    end

    disp(' ');
    ans=input('All the input data are correct (y/n)? ','s');
    if ans=='y' | ans=='Y',
        ans='y';
    else
        ans='n';
    end

    % This end comes from the initial while

    % *****
    end
    % *****

    clc

    % *****
    % REPRODUCTION OF THE ORIGINAL DATA MATRIX BY PCA
    % *****
    % dn is the experimental matrix and d is the PCA reproduced matrix
    disp('***** Results obtained after application of PCA to the data matrix *****');
    disp(' ');
    dn=d;
    [u,s,v,d,sd]=pcarep(dn,nsign);

    disp(['Data reproduction with the considered number of species has an error',
    num2str(sd)]);
    disp(' ');disp(' ');disp(' ');
    figure;
    subplot(2,1,1);plot(u*s);title('Scores matrix')
    subplot(2,1,2);plot(v);title('Loadings matrix')

    disp('***** Press any key to start the ALS optimization *****')
    pause

    clc
    sstn=sum(sum(dn.*dn));
    sst=sum(sum(d.*d));
    sigma2=sqrt(sstn);

```

```

% *****
% STARTING ALTERNATING CONSTRAINED LEAST SQUARES OPTIMIZATION
% *****

while niter < nit
niter=niter+1;

% *****
% ESTIMATE CONCENTRATIONS non negative constrained als solutions (nnls)
% *****

if ialg==0,

    conc=d/abss;

    % *****
    % CONSTRAIN APROIATELY THE CONCENTRATIONS
    % *****

    % *****
    % non-negativity
    % *****

    for k=1:nsign,
        for j=1:nrow,
            if conc(j,k)<0.0,
                conc(j,k)=0.0;
            end
        end
    end

end

if ialg==1,
    for i=1:nrow,
        x=nnls(abss',d(i,:)');
        conc(i,:)=x';
    end

end

% *****
% zero concentration species
% *****

for i=1:nexp,

```

```

        for j=1:nsign,
            if isp(i,j)==0,
                for ii=ninic(i):nfin(i),
                    conc(ii,j)=0.0;
                end
            end
        end
    end

% *****
% selectivity
% *****

if iselect==1 & ils==1,
    for ip=1:nselect,
        np=ipselect(ip);
        ns=spselect(ip);
        conc(np,ns)=0;
    end
end

% *****
% equal shapes
% *****

if ishape>=1,
    for j=1:nsign,
        if spetri(j)==1,
            [conc(:,j),t]=shape3(conc(:,j),nexp,ishape);
            totalconc(j,1:nexp)=t;
            rt(j,1:nexp)=totalconc(j,1:nexp)./totalconc(j,1);
        end
    end
end

% *****
% for every experiment unimodality is evaluated separately !!!
% *****

for i=1:nexp,
    % *****
    % unimodality
    % *****

```

```

kinic=ninic(i);
kfin=nfin(i);
conc2=conc(kinic:kfin,:);

if imod==1|imod==3,

    for ii=1:nsign,
        if spmod(ii)==1,
            [conc2(:,ii)]=unimod(conc2(:,ii),rmod);
        end
    end

conc(kinic:kfin,:)=conc2;

end

% *****
% closure
% *****

if iclos(i)==1 | iclos(i)==2,

[conc2]=closure2(conc2,iclos(i),tclos1(i),sclos1(i,:),iclos1(i),tclos2(i),sclos2(i,:),
iclos2(i));
    conc(kinic:kfin,:)=conc2;
end

end

if gr=='y';
    subplot(2,1,1);plot(conc);title('Concentration profiles');
end

% recalculation of total and ratio concentrations if ishape=0 and niter=1

if ishape==0 | niter==1,
    for j=1:nsign,
        for inexp=1:nexp,
            totalconc(j,inexp)=sum(conc(ninic(inexp):nfin(inexp),j));
        end
        rt(j,1:inexp)=totalconc(j,1:nexp)./totalconc(j,1);
    end
end
end

```

```

% areas under concentration profiles
area=totalconc;
% disp(['Relative concentrations respect matrix (sample) 1',num2str(rt)])
% disp(['Areas under concentration profiles ',num2str(area)])

% *****
% ESTIMATE SPECTRA non negative constrained als solutions (nnls)
% *****
%
% *****
% non-negative spectra
% *****

if ineg==1,
    if ialg==0,
        abss=conc\d;
        for k=1:nsign,
            for j=1:ncol,
                if abss(k,j)<0.0,
                    abss(k,j)=0.0;
                end
            end
        end
    end
    if ialg==1,
        for j=1:ncol,
            abss(:,j)=nnls(conc,d(:,j));
        end
    end
else
    abss=conc\d;
end

% *****
% restore (input) pure fixed spectra (not allowed to chang during the optimization)
% *****

```

```

if ils == 2 & nfix~=0,
    for ii=1:nsign,
        if ifix(ii)==1,
            abss(ii,:)=x0(ii,:);
            % disp('spectra kept constant ');ii
        end
    end
end

% *****
% selectivity in spectra
% *****

if iselect==1 & ils==2,
    for iw=1:nselect,
        nw=iwselect(iw);
        ns=spselect(iw);
        abss(ns,nw)=0;
    end
end

% *****
% possibility to constrain the unimodality of spectra
% *****

if imod==2|imod==3,
    for i=1:nsign,
        dummy=unimod(abss(i,:)','smod(i));
        abss(i,:)=dummy';
    end
end

% *****
% possibility of closure in spectra (in case of inverted analysis D'=SC
% *****
% *****
% closure
% *****

if iclos==3,

% tclos closure is fixed for all the experiments and has been given

```

% the value tclos(1); further implementation of this constraint is needed

```

    for k=1:ncol,
        sump(k)=sum(abss(:,k));
        abss(:,k)=(abss(:,k)./sump(k))*tclos(1);
    end

end

if gr=='y';
    subplot(2,1,2);
    plot(abss');title('Unit spectra');
    pause(1);
end

% *****
% CALCULATE RESIDUALS
% *****

res=d-conc*abss;
resn=dn-conc*abss;

% plot(res)
% plot(res')

% *****
% OPTIMIZATION RESULTS
% *****

disp(' ');disp(' ');disp(['ITERATION ',num2str(niter)]);
u=sum(sum(res.*res));
un=sum(sum(resn.*resn));
disp(['Sum of squares respect PCA reprod. = ', num2str(u)]);
sigma=sqrt(u/(nrow*ncol));
sigman=sqrt(un/(nrow*ncol));
disp(['Old sigma = ', num2str(sigma2), ' ——> New sigma = ',
num2str(sigma)]);
disp(['Sigma respect experimental data = ', num2str(sigman)]);
disp(' ');
change=((sigma2-sigma)/sigma);
if change<0.0,
    disp(' ')
    disp('DIVERGENCE !!!')
    idev=idev+1;

```



```

else,
    idev=0;
end
change=change*100;
disp(['Change in sigma (%) = ', num2str(change)]);
sstd(1)=sqrt(u/sst)*100;
sstd(2)=sqrt(un/sstn)*100;
disp(['Fitting error in % (PCA) = ', num2str(sstd(1))]);
disp(['Fitting error in % (exp) = ', num2str(sstd(2))]);

% *****
% if change is positive, the optimization is working correctly
% *****

if change>0 | niter==1,
    sigma2=sigma;
    copt=conc;
    sopt=abss;
    sdopt=sstd;
    ropt=res;
    rtopt=rt';
    itopt=niter;
    areaopt=area;
end

% *****
% test for convergence
% *****
if abs(change) < tolsigma,
    % finish the iterative als because convergence is achieved
    disp(' ')
    disp('CONVERGENCE IS ACHIEVED !!!!')
    disp(' ')
    disp('Relative concentrations respect matrix (sample) 1'),disp(rt')
    % disp(['Areas under concentration profiles ',num2str(area)])
    pause
    return
end

% if divergence occurs 10 times consecutively, stops
if idev > 9,

    disp('Divergence 9 times consecutively, stop!!!')

```

```
        disp(['Results given are at optimum in the iteration ', num2str(itopt)]);
        disp(['Percent of fit at the optimum = ', num2str(sdopt)]);
        disp('Relative concentrations respect matrix (sample) 1'),disp(rt)
        pause
        return
    end

end

% *****
% maximum number of allowed iterations is exceeded
% *****

disp('NUMBER OF ITERATIONS EXCEDED THE ALLOWED!')
disp('Relative concentrations respect matrix (sample) 1'),disp(rt)
pause

return
end
```

5.5. PROGRAMA MATLAB D'APLICACIÓ DE LA RESTRICCIÓ D'UNIMODALITAT

```

function [conc]=unimod(conc,rmod)
[ns,nc]=size(conc);
nmdf=0;

% 1) look for the maximum
for j=1:nc,
[y,imax(j)]=max(conc(:,j));
end

% force unimodality shape
for j=1:nc,
rmax=conc(imax(j),j);

% discard left maxima (tolerance rmod)
for k=imax(j):-1:1,
if conc(k,j)<=rmax,
rmax=conc(k,j);
else,
rmax2=rmax*rmod;
    if conc(k,j)>rmax2,
%         conc(k,j)=1.0E-30;
        conc(k,j)=conc(k-1,j)
        nmdf=nmdf+1;
        rmax=conc(k,j);
    end
end

end

% discard right maxima (tolerance rmod)
rmax=conc(imax(j),j);
for k=imax(j):ns,
if conc(k,j)<=rmax,
rmax=conc(k,j);
else,
rmax2=rmax*rmod;
    if conc(k,j)>rmax2,
%         conc(k,j)=1.0E-30;

```

```
        conc(k,j)=conc(k-1,j)
    nmdf=nmdf+1;
    rmax=conc(k,j);
end
end
end
```

5.6. PROGRAMA MATLAB PER FORÇAR ELS PERFILS DE CONCENTRACIÓ A TENIR LA MATEIXA FORMA (DADES AMB ESTRUCTURA TRILINEAL)

```

function [c,t]=shape(c,ne,ishape)
%function [c,t]=shape(c,ne)
% forces the shapes of the profiles to be equal
% estimates the ratios between the intensity of the profiles
% c input-output unfolded profiles
% ne number of experiments considered

[nst,nsign]=size(c);
ns=nst/ne;
ni=1;
nf=ns;

% folds the unfolded profiles in a matrix
% assumes there is synchronization

for j=1:ne,
cp(1:ns,j)=c(ni:nf,nsign);
ni=nf+1;
nf=ni+ns-1;
end
y=cp;

% *****
% ishape=2 !!!!
% there is peak shifting, matching the peak maxima
% looks for maxima (ymax) and positions (imax)
% *****

if ishape==2,
[ymax,imax]=max(cp)

% looks for the minimum maximum (ymin)
% and takes it as a reference(imin)

imin=1;
ymin=imax(1);

for i=2:ne,
    if imax(i)<ymin & ymax(i)>0,

```

```

        imin=i;
        ymin=imax(i);
    end

    end

    for j=1:ne,

        if j~=imin & ymax(j)>0,

            ishift=imax(j)-ymin;

            if ishift~=0,
                y(1:ns-ishift,j)=cp(1+ishift:ns,j);
                y(ns-ishift+1:ns,j)=zeros(ishift,1);
            else,
                y(:,j)=cp(:,j);
            end

        else,
            y(:,j)=cp(:,j);
        end

    end

end
end
% *****
% this is for ishape = 1 or 2
% svd of the profile matrix
% *****

[u,s,v]=svd(y,0);

% estimates the total concentrations
t=s(1,1).*v(:,1)';

% calculates the new profiles
cp2=u(:,1)*t;
y=cp2

% *****
% this is only for ishape=2
% breaks the matching of peaks to original positions
% *****

```

```

if ishape==2,
for j=1:ne,
    if j~=imin & ymax(j)>0,
        ishift=imax(j)-ymin;
        if ishift~=0,
            y(1:ishift,j)=zeros(ishift,1);
            y(ishift+1:ns,j)=cp2(1:ns-ishift,j);
        else,
            y(:,j)=cp2(:,j);
        end
    else,
        y(:,j)=cp2(:,j);
    end
end
end
% *****
plot(y)
% pause

% unfolds the profiles
ni=1;
nf=ns;
for j=1:ne,
    c(ni:nf,nsign)=y(1:ns,j);
    ni=nf+1;
    nf=ni+ns-1;
end

% calculates the differences
rs=cp-y;
u=sum(sum(rs*rs'));
sigma=sqrt(u/(nst));
disp('std. dev. res. between the estimations');disp(sigma);
disp('singular values ');disp([s(1,1);s(2,2);s(3,3)]);
disp('estimates of the total concentrations are: '),disp(t);
% pause
end

```

5.7. PROGRAMA MATLAB D'APLICACIÓ DE LA RESTRICCIÓ D'UNIMODALITAT

```
% function [conc] = unimod3(conc,fmod,rmod,sunimod,rpeak,z). This is a
% modified unimodality
% constraint. No secondary maxima are allowed, and the peak maxima are
% previously fixed. Small departures of both conditions (unimodality and
% peak maxima) are allowed. conc:concentration profiles to be constrained.
% rmod: Unimodal tolerance. z: Peak maxima. rpeak: Peak position tolerance.
% fmod: Unimodality mode. 1:With fixed peak maxima. 0: Without fixed peak
% max.
function [conc]=unimod3(conc,fmod,rmod,sunimod,rpeak,z);

[ns,nc]=size(conc);
nmdf=0;

% *****
% Looking for the experimental maxima
% *****

% 1) look for the maximum

for j=1:nc,
    [y,imax(j)]=max(conc(:,j));
end

% *****
% Unimodality without defined peak maxima
% *****

if nargin <5

% force unimodal shape

for j=1:nc,
    if sunimod(j)==1
        rmax=conc(imax(j),j);

% discard left maxima (tolerance rmod)

for k=imax(j):-1:1,
    if conc(k,j)<=rmax,
```



```

rmax=conc(k,j);
else,
rmax2=rmax*rmod;
    if conc(k,j)>rmax2,
%     conc(k,j)=1.0E-30;
    conc(k,j) = conc(k+1,j);
%     conc(k,j)=conc(k-1,j)*0.9
    nmdf=nmdf+1;
    rmax=conc(k,j);
    end
end

end

% discard right maxima (tolerance rmod)

rmax=conc(imax(j),j);

for k=imax(j):ns,

if conc(k,j)<=rmax,
rmax=conc(k,j);
else,
rmax2=rmax*rmod;
    if conc(k,j)>rmax2,
%     conc(k,j)=1.0E-30;
    conc(k,j) = conc(k-1,j);
%     conc(k,j)=conc(k-1,j)*0.9
    nmdf=nmdf+1;
    rmax=conc(k,j);
    end
end
end
end
end

% *****
% Unimodality with fixed maxima
% *****

if nargin == 6

% 1) compare the experimental maximum and the fixed value

```

```

if sunimod(1,j)==1
for j = 1:nc
if z(:,j)==0
imax(j)=imax(j);
end
if z(:,j)~=0 & imax(j)>z(:,j)+rpeak|imax(j)<z(:,j)-rpeak
imax(j)=z(:,j);
else
imax(j) = imax(j);
end
end

% force unimodal shape

for j=1:nc,
rmax=conc(imax(j),j);

% discard left maxima (tolerance rmod)

for k=imax(j):-1:1,

if conc(k,j)<=rmax,
rmax=conc(k,j);
else,
rmax2=rmax*rmod;
    if conc(k,j)>rmax2,
%       conc(k,j)=1.0E-30;
        conc(k,j) = conc(k+1,j);
%       conc(k,j)=conc(k-1,j)*0.9
        nmdf=nmdf+1;
        rmax=conc(k,j);
    end
end

end

% discard right maxima (tolerance rmod)

rmax=conc(imax(j),j);

for k=imax(j):ns,

if conc(k,j)<=rmax,

```

```
rmax=conc(k,j);
else,
rmax2=rmax*rmod;
    if conc(k,j)>rmax2,
%    conc(k,j)=1.0E-30;
    conc(k,j) = conc(k-1,j);
%    conc(k,j)=conc(k-1,j)*0.9
        nmdf=nmdf+1;
        rmax=conc(k,j);
    end
end
end
end
end
end
```

5.8. PROGRAMA MATLAB D'APLICACIÓ DE LA RESTRICCIÓ DE SISTEMA TANCAT (*CLOSURE*)

```
function [conc]=closure2(conc,iclos,tclos1,sclos1,iclos1,tclos2,sclos2,iclos2)
% function [conc]=closure(conc,iclos,tclos,sclos,tclos2,iclos3)
```

```
% closure types
```

```
[kfin,nspec]=size(conc);
```

```
% calculation of the total sums
```

```
for k=1:kfin,
```

```
    summ1(k)=0.0;
```

```
    summ2(k)=0.0;
```

```
        for ns=1:nspec,
```

```
            if sclos1(ns) == 1,
```

```
                summ1(k)=summ1(k)+conc(k,ns);
```

```
            end
```

```
        if iclos==2,
```

```
            if sclos2(ns) == 1,
```

```
                summ2(k)=summ2(k)+conc(k,ns);
```

```
            end
```

```
        end
```

```
    end
```

```
end
```

```
maxsum1=max(summ1);
```

```
maxsum2=max(summ2);
```

```
% application of the different closure types
```

```
for k=1:kfin,
```

```
    for ns=1:nspec,
```

```
        if iclos == 1 & iclos1 == 1 & sclos1(ns) == 1
```

```
            conc(k,ns)=(conc(k,ns).*tclos1./summ1(k));
```

```
end

if iclos == 1 & iclos1 == 2 & sclos1(ns) == 1
conc(k,ns)=conc(k,ns).*tclos1./maxsum1;
end

if iclos == 2 & iclos2 == 1 & sclos2(ns) == 1,
conc(k,ns)=(conc(k,ns).*tclos2./summ2(k));
end

if iclos == 2 & iclos2 == 2 & sclos2(ns) == 1,
conc(k,ns)=conc(k,ns).*tclos2./maxsum2;
end

end

end

plot(conc)

end
```

5.9. PROGRAMA MATLAB DE DESCOMPOSICIÓ PCA

```

function [u,s,v,x,sigma]=pcarep(xi,nf)
% function [u,s,v,x,sigma]=pcarep(xi,nf)
% PCA reproduction of the original data matrix x dor nf components
% u,s,v and x as in svd matlab function [u,s,v]=svd(x)
% xi is the original input data matrix (with noise)
% x is the pca reproduced data matrix (filtered no noise)
% nf is the number of components to be included in the reproduction
[u,s,v]=svd(xi,0);
u=u(:,1:nf);
s=s(1:nf,1:nf);
v=v(:,1:nf);
x=u*s*v';
res=xi-x;
sst1=sum(sum(res.*res))
sst2=sum(sum(xi.*xi))
sigma=(sqrt(sst1/sst2))*100
disp('percent of lack of fit with the considered num of PC ');disp(sigma)
pause
end

```

5.10. PROGRAMA MATLAB D'APLICACIÓ DE LA RESTRICCIÓ DE NO-NEGATIVITAT (NNLS)

```
function [x,w] = nnls(E,f,tol)
%NNLS      Non-negative least-squares.
%   X = NNLS(A,b) returns the vector X that solves A*x = b
%   in a least squares sense, subject to x >= 0.
%
%   A default tolerance of TOL = MAX(SIZE(A)) * NORM(A,1) * EPS
%   is used for deciding when elements of X are less than zero.
%   This can be overridden with X = NNLS(A,b,TOL).
%
%   [X,W] = NNLS(A,b) also returns dual vector W where
%   w(i) < 0 where x(i) = 0 and w(i) = 0 where x(i) > 0.
%
%   See also LSCOV, SLASH.

%   L. Shure 5-8-87
%   Revised, 12-15-88,8-31-89 LS.
%   Copyright (c) 1984-94 by The MathWorks, Inc.

% Reference:
% Lawson and Hanson, «Solving Least Squares Problems», Prentice-Hall, 1974.

% initialize variables
if nargin < 3
    tol = 10*eps*norm(E,1)*max(size(E));
end
[m,n] = size(E);
P = zeros(1,n);
Z = 1:n;
x = P';
ZZ=Z;
w = E'*(f-E*x);

% set up iteration criterion
iter = 0;
itmax = 3*n;

% outer loop to put variables into set to hold positive coefficients
while any(Z) & any(w(ZZ) > tol)
    [wt,t] = max(w(ZZ));
    t = ZZ(t);
```

```

P(1,t) = t;
Z(t) = 0;
PP = find(P);
ZZ = find(Z);
nzz = size(ZZ);
EP(1:m,PP) = E(:,PP);
EP(:,ZZ) = zeros(m,nzz(2));
z = pinv(EP)*f;
z(ZZ) = zeros(nzz(2),nzz(1));
% inner loop to remove elements from the positive set which no longer belong
while any((z(PP) <= tol))
    iter = iter + 1;
    if iter > itmax
        error(['Iteration count is exceeded.', ...
              ' Try raising the tolerance.'])
    end
    QQ = find((z <= tol) & P');
    alpha = min(x(QQ)./(x(QQ) - z(QQ)));
    x = x + alpha*(z - x);
    ij = find(abs(x) < tol & P' ~= 0);
    Z(ij)=ij';
    P(ij)=zeros(1,max(size(ij)));
    PP = find(P);
    ZZ = find(Z);
    nzz = size(ZZ);
    EP(1:m,PP) = E(:,PP);
    EP(:,ZZ) = zeros(m,nzz(2));
    z = pinv(EP)*f;
    z(ZZ) = zeros(nzz(2),nzz(1));
end
x = z;
E'*(f-E*x);
end

```


REFERÈNCIES BIBLIOGRÀFIQUES

1. WOLD, S. *Kemisk Tidskrift*, 34 (1972), p. 3. WOLD, S. *Svensk Naturventenskap* (1974), p. 201. KOWALSKI, B. R. *Chem. Inform. & Comp. Sci.*, 15 (1975), p. 201. KOWALSKI, B. R.; BROWN, S.; VANDEGINSTE, B. J. *J. of Chemometrics*, 1 (1987), p. 1.
2. SHARAF, M. A.; ILLMAN, D. L.; KOWALSKI, B. R. *Cheometrics*. Wiley: Nova York, 1986.
3. CARLSON, ROLF. *Design and Optimization in Organic Synthesis*. Amsterdam: Elsevier, 1992.
4. KRESTA, J. K.; MC GREGOR, J. F.; MERLIN, TH. E. «Multivariate Statistical Monitoring of Process Operating Performance». *Can. J. Chem. Eng.*, 69 (1991), p. 35-43; WISE, L. J.; RICKER, L.; KOWALSKI, B. R. *Pro. Contr. Qual.*, 1 (1990), p. 41-51.
5. KROONENBERG, P. M. *Three-mode principal component analysis*. DSWO Press. Leiden: Holand, 1983.
6. CHRIST, C. F. *Econometric Models and Methods*. Wiley: Nova York, 1966.
7. CALLIS, J. B.; ILLMAN, D. L.; KOWALSKI, B. R. «Process Analytical Chemistry». *Anal. Chem.*, 59 (1987), 624A; RIEBE, M. T; EUSTACE, D. J. «Process Analytical Chemistry An Industrial Perspective». *Anal. Chem.*, 62 (1990), 65A.
8. ARIÑO, C.; ESTEBAN, M.; FILELLA, M.; FONRODONA, G.; HERNÁNDEZ, S.; IZQUIERDO-RIDORSA, A.; TAULER, R. «L'estudi dels Equilibris de Formació de Complexos Metàl·lics: Un Aspecte Bàsic de la Química Analítica». *Buill. Soc. Cat. Cien*, 9 (1987), p. 45-50.
9. TAULER, R. «Equilibris de Formació de complexos de coure (II)-salicilat- base nitrogenada». *Tesi doctoral*, publicacions edicions Universitat de Barcelona. Dipòsit legal B: 36. 881, Barcelona, 1984.
10. CASASSAS, E.; TAULER, R. «Mixed Ligand Complex Formation in the Copper (II)-Salicylate-Ion-Ethilenediamine System». *J. Chim. Phys.*, 81 (1984), p. 233-241.
11. CASASSAS, E.; TAULER, R. «Mixed Ligand Complex Formation in the Copper (II)-Salicylate-Ion-Picolinate Ion System». *J. Chim. Phys.*, 81 (1984), p. 558-567.
12. CASASSAS, E.; TAULER, R. «Mixed Ligand Complex Formation in the Copper (II)-Salicylate-Ammonia and in the Copper (II)-Salicylate-Pyridine Systems». *J. Chim. Phys.*, 82 (1985), p. 1067-1077.

13. TAULER, R.; CASASSAS, E.; RAINER, M. J. A.; RODE, B. M. «The Complex Formation of Cu (II) with Triethanolamine in Aqueous Solution». *Inorg. Chim. Acta*, 105 (1985), p. 165-170.
14. TAULER, R.; RAINER, M. J. A.; RODE, B. M. «The Interaction of H^+ , Zn (II) and Cu (II) with Adenine and 9-Methyladenine in Aqueous Solution». *Inorg. Chim. Acta*, 123 (1986), p. 75-82.
15. TAULER, R.; CASASSAS, E.; RODE, B. M. «The Complex Formation of Cu (II) with Mono- and Diethanolamine in Aqueous Solution». *Inorg. Chim. Acta*, 114 (1986), p. 203-209.
16. CASASSAS, E.; TAULER, R. «Potentiometric Study of Mixed Ligand Complex Formation in the Copper (II)-Salicylate-Mono-, Di- and Triethanolamine Systems». *J. Chim. Phys.*, 83 (1986), p. 409-417.
17. CASASSAS, E.; TAULER, R.; FILELLA, M. «A Critical Comparison of Computer Programs for the Potentiometric Determination of Stability Constants». *Anal. Chim. Acta*, 191 (1986), p. 399-411.
18. CASASSAS, E.; FILELLA, M.; TAULER, R. «Assesment of the results Obtained from Different Computer Programs Applied to Potentiometric Complexation Data». *Anal. Chim. Acta*, 191 (1986), p. 413-423.
19. CASASSAS, E.; FONRODONA, G.; TAULER, R. «Potentiometric Study of the Co (II), Ni (II), and Cd (II) Complexes with 3-Hydroxy-2-Naphtalene Carbo-xylic Acid». *Polyhedron*, 6 (1987), p. 1517-1521.
20. TAULER, R.; CASASSAS, E. «A simplex Search for Conditions Giving Maximal and/or Minimal Concentrations of Species in Distribution Plots». *Anal. Chim. Acta*, 206 (1988), p. 189-202.
21. CASASSAS, E.; TAULER, R.; FONRODONA, G. «Mixed Ligand Complex Formation of Copper (II) and Zinc (II) Ions with 3-Hydroxynaphtoate and Picolinate ions in Dioxane-Water Mixtures». *Polyhedron*, 7 (1988), p. 1335-1340.
22. CASASSAS, E.; GUSTEMS, LL.; TAULER, R. «Spectrophotometric Study of Complex Formation in Copper (II)-Mono-, Di- and Triethanolamine Systems». *J. Chem. Soc. Dalton Trans* (1989), p. 569-573.
23. CASASSAS, E.; FONRODONA, G.; IZQUIERDO-RIDORSA, A.; TAULER, R. «Estudio Comparativo de las Propiedades ácido-base y complejantes de algunos ácidos orto-hidroxycarboxílicos aromáticos: ácido salicílico, ácido 3-hidroxi- 2-naftoico y ácido pamoico». *Anales de Química*, 86, serie B (1990), p. 257.
24. TAULER, R.; RODE, B. M. «Reactions of Cu (II) with glycine and glycyglycine in aqueous solution at high concentrations of sodium chloride». *Inorg. Chim. Acta*, 173 (1990), p. 93-98.
25. CASASSAS, E.; IZQUIERDO-RIDORSA, A.; TAULER, J. «Acid-Base and Complex-Forming Properties of Uracil and Uridine in Aqueous Solution». *Inorganic Biochemistry*, 39 (1990), p. 327-336.
26. «Potentiometric and 1H NMR study of the interaction of hypoxanthine and inosine with H^+ , Cu^{2+} and Zn^{2+} ions». *J. of Bioinorganic Chemistry*, 39 (1990), p. 277-285.

27. CASASSAS, E.; IZQUIERDO-RIDORSA, A.; TAULER, R. «ESR and VIS Spectroscopic Studies of Mixed Ligand Complexes of Copper (II) Ion, Salicylate and Ammonia, Mono-, Di-, or Triethanolamine in Aqueous Solution». *J. Chem. Soc. Dalton* (1990), p. 2341-2345.
28. CASASSAS, E.; FONRODONA, G.; DE JUAN, A.; TAULER, R. «Assesment of Solvent Parameters and Their Correlation with Protonation Constants in Dioxane-Water Mixtures Using Factor Analysis». *Chemometrics and intelligent laboratory systems* (1991), p. 29-38.
29. «Study of the acid-base behavior and Cu (II) complexing properties of uracil and hypoxanthin-derived nucleotides in aqueous solution». *J. Inorg. Biochemistry* (1994). [En premsa].
30. IZQUIERDO-RIDORSA, A.; TAULER, R. «Introducció a la modelització en els estudis d'especiació química de ions metàl·lics en els sistemes biològics». *Modelització macroscòpica en ciències experimentals*, cap. 3, Institut d'Estudis Catalans. «Monografies de les Seccions de Ciències». Vol. 10, Institut d'Estudis Catalans, Barcelona.
31. GANS, Peter. *Data Fitting in the Chemical Sciences*. Wiley: Nova York, 1992.
32. LEGGETT, D. J. «Computational Methods for the Determination of Formation Constants». *Modern Inorganic Chemistry*, Plenum Press: Nova York, (1995).
33. MARTELL, A. E.; MOTEKAITIS, R. J. «The determination and use of stability constants». VCH, Weinheim: Germany, (1988).
34. HAMILTON, J. C.; GEMPERLINE, P. J. *J. Chemometrics*, 4 (1990), p. 1-14.
35. WINDING, W. *Chemometrics and Intelligent Laboratory Systems*, 16 (1992), p. 1-16.
36. MARTENS, K.; NAES, T. *Multivariate Calibration*, Wiley: Nova York, 1989.
37. VANDEGINSTE, B. G. M.; DARKS, W.; KATEMAN, G. *Anal. Chim. Acta*, 173 (1985), p. 253-264.
38. WINDING, W.; GUILMET, J. *Anal. Chem.*, 63 (1991), p. 1425-1432.
39. GAMPP, H.; MAEDER, M.; MEYER, Ch.; AD ZUBERBUHLER. *Talanta*, 32 (1985), p. 1133-1139. GAMPP, H.; MAEDER, M.; MEYER, Ch.; AD ZUBERBUHLER. *Chimia*, 39 (1985), p. 315-317. GAMPP, H.; MAEDER, M.; MEYER, Ch.; AD ZUBERBUHLER. *Talanta*, 33 (1986), p. 943-951.
40. GAMPP, H.; MAEDER, M.; MEYER, Ch.; AD ZUBERBUHLER. *Anal. Chim. Acta*, 193 (1987), p. 287-293.
41. MALINOWSKI, E. R. *J. Chemometrics*, 6 (1992), p. 29-40.
42. KVALHEIM, O. M.; LIANG, Y. Z. *Anal. Chem.*, 64 (1992), p. 936-946. LIANG, Y. Z.; KVALHEIM, O. M.; KELLER, H. R.; MASSART, D. L.; KIECHLE, P.; ERNI, F. *Anal. Chem.*, 64 (1992), p. 946-953.
43. LAWTON, W. H.; SYLVESTRE, E. A. *Technometrics*, 13 (1971), p. 617-631.
44. BORGES, O. S.; KOWALSKI, B. R. *Anal. Chim. Acta*, 174 (1985), p. 1-26.
45. BORGES, O. S.; DAVIDSEN, N.; MINGYANG, Z.; OYEN, O. «Mikrocim». *Acta*, 174 (1985), p. 1-26.
46. SANCHEZ, E.; KOWALSKI, B. R. *J. Chemometrics*, 2 (1988), p. 247-265.
47. SANCHEZ, E.; KOWALSKI, B. R. *Anal. Chem.*, 58 (1986), p. 496.

48. SANCHEZ, E.; KOWALSKI, B. R. *J. Chemometrics*, 4 (1990), p. 29-45.
49. SMILDE, A. K.; DOORNBOS, D. A. *J. Chemometrics*, 5 (1991), p. 345.
50. MILLER, J. C.; MILLER, J. N. «Ellis Horwood series in Analytical Chemistry». *Statistic for Analytical Chemistry*, 2a ed. Nova York, (1988).
51. SEKULINK, S.; SEASHOLTZ, M. B.; WANG, Z.; KOWALSKI, B. R. *Anal. Chem.*, 65 (1993), 835A-845A.
52. MALINOWSKI, E. R. *Factor Analysis in Chemistry*, 2a ed. Wiley: Nova York, 1991.
53. COPPI, R.; BOLASBO, S. *Multiway Data Analysis*. North Holland: Amsterdam, 1989.
54. TAULER, R.; CASASSAS, E. «Principal Component Analysis Applied to the Study of Successive Complex Formation Data in the Cu (II) ethanalamine Systems». *J. of Chemometrics*, 3 (1988), p. 151-161. TAULER, R.; CASASSAS, E. «Application of Principal Component Analysis to the Study of Multiequi-libria Systems: Study of Cu (II)-Salicylate-mono-, di-and triethanolamine Systems». *Anal. Chim. Acta*, 223 (1989), p. 257-268.
55. TAULER, R.; CASASSAS, E. «Application of factor analysis to speciation in multiequilibria systems». *Analysis*, 20 (1992), p. 255-268.
56. TAULER, R.; CASASSAS, E.; IZQUIERDO-RIDORSA, A. «Self-Modeling Curve resolution applied to Spectroscopic Titration Data using Factor Analysis». *Anal. Chim. Acta*, 1991, p. 447-458. TAULER, R.; IZQUIERDO-RIDORSA, A.; CASASSAS, E. «Simultaneous analysis of several spectroscopic titrations with self-modeling curve resolution». *Chemometrics and Intelligent Laboratory Systems*, 18 (1993), p. 293-300.
57. TAULER, R.; KOWALSKI, B. R.; FLEMING, S. «Multivariate Curve resolution applied to Process Analysis». *Anal. Chem.*, 65 (1993), p. 2040-2047.
58. WOLD, S. *Technometrics*, 20 (1978), p. 397-405.
59. MALINOWSKI, E. R. *Anal. Chem.*, 49 (1977), p. 612-617.
60. SPJOTVOLL, E.; MARTENS, H.; VOLDEN, R. *Technometrics*, 3 (1982), p. 173-180.
61. KELLER, H. R.; MASSART, D. L. *Anal. Chim. Acta*, 246 (1991), p. 379-390.
62. LORBER, A. *Anal. Chem.*, 58 (1988), p. 1167-1172.
63. LIANG, Y. Z.; KVALHEIM, O. M. *Anal. Chim. Acta*, 276 (1992), p. 425-440.
64. GOLUB, G. H.; VAN LOAN, Ch. F. *Matrix Computations*. Univ. Press, Baltimore: Johns Hopkins, 1989.
65. *MATLAB, Optimization Toolbox. High Performance Numeric Computation and Visualization Software*. Natick, Mass.: The Math Works Inc., 1990.
66. LEGGETT, D. J. *Anal. Chem.*, 49 (1977), p. 276-280.
67. SIGEL, Helmut [ed.]. «Circulation of metals in the environment». *Metal Ions in Biological Systems*. Vol. 18, Nova York: Marcel Dekker, (1984). SIGEL, Helmut [ed.]. «Methods involving metal ions and complexes in clinical chemistry». *Metal Ions in Biological Systems*. Vol. 16, Nova York: Marcel Dekker, (1983).
68. BUFFLE, J.; HORWOOD Sr., Ellis; WILEY, John and Sons. «Complexation Reactions in Aquatic Systems». *An Analytical Approach*, (1988).
69. MARINSKY, J. A. *Coord. Chem. Rev.*, 19 (1976), p. 125. BARBUCCI, R.; CASOLARO, M.; COREZZI, M. S.; FERRUTI, P.; BARONE, V. *Macromolecules*, 19 (1986), p. 37.
70. CASASSAS, E.; GARGALLO, R.; GIMENEZ, I.; IZQUIERDO-RIDORSA, A.; TAULER, R. «Application of an evolving factor analysis-based procedure to speciation in the Copper

- (II)-polyuridilic acid system». *Analytica Chimica Acta*, 283 (1993), p. 538-547.
71. CASASSAS, E.; GARGALLO, R.; IZQUIERDO-RIDORSA, A.; TAULER, R. «Aplication of a new multivariate curve resolution procedure to the simultaneous analysis of several spectroscopic titrations of copper-polyinosinic acid». *Chemometrics and Intelligent Laboratory Systems*, 27 (1995), p. 163-174.
72. CASASSAS, E.; GARGALLO, R.; GIMÉNEZ, I.; IZQUIERDO-RIDORSA, A.; TAULER, R. «Study of the acid-base behavior and Cu (II) complexing properties of uracil and hypoxanthin-derived nucleotides in aqueous solution». *Journal Inorganic Biochemistry*, 56 (1994), p. 187-199.
73. CASASSAS, E.; MARQUÈS, I.; TAULER, R. «Study of acid-base properties of fulvic acids using fluorescence spectrometry and multivariate curve resolution methods». *Analytica Chimica Acta*, 310 (1995), p. 473-484.
74. TAULER, R.; CASASSAS, E. «Spectroscopic resolution of macromolecular complexes using factor analysis: Cu (II)-polyethyleneimine system». *Chemometrics and Intelligent Laboratory Systems*, 14 (1992), p. 305- 317.
75. CASASSAS, E.; BARCELÓ, J.; TAULER, R. [Resultats no publicats].
76. CASASSAS, E.; TAULER, R.; MARQUÈS, M. I. «Interactions of H⁺ and Cu (II) ions with polyadenylic acid: study by factor analysis». *Macromolecules*, 27 (1994), p. 1729-1737.
77. BERGER, N. A.; EICHHORN, G. L. *Biochemistry*, 10 (1971), p. 1857. EICHHORN, G. L.; RIFKIND, Y. A.; SHIN, Y. A. *Adv. in Chem. Ser.*, 162, p. 305.
78. TAULER, R.; DURAND, G.; BARCELÓ, D. «Deconvolution and Quantitation of complex unresolved mixtures of pesticides in liquid chromatography- diode array detection using evolving factor analysis». *Chromatographia*, 33 (1992), p. 244-254.
79. TAULER, R.; BARCELÓ, D. «Multivariate Curve Resolution and calibration applied to liquid chromatography diode array detection». *Trends in Analytical Chemistry*, 12 (1993), p. 319-327.
80. MALINOWSKI, E. R. *J. Chemometrics*, 8 (1987), p. 33-40.
81. MILANIVITCH, F.; GARVIS, D.; ANGEL, S. M.; KLAINER, S. M.; ECCLES, L. *Anal. Inst.*, 15 (1986), p. 137.
82. LUGG, G. A. *Anal. Chem.*, 38 (1966), p. 1532-1536. REITH, J. F.; VAN DITMARSCH, W. C.; RUITER DE, Th. *Analyst*, 99 (1974), p. 652-656.
83. HENSHAW, J. M.; BURGESS, L. W.; KOWALSKI, B. R. «Multicomponent determination of chlorinated hydrocarbons using a reaction-based chemical sensor. Part I Sensor description». *Analytical Chemistry*, 66 (1994), p. 3330-3336.
84. TAULER, R.; SMILDE, A. K.; HENSHAW, J. M.; BURGESS, L. W.; KOWALSKI, B. R. «Multicomponent determination of chlorinated hydrocarbons using a reaction-based chemical sensor. Part II Chemical speciation using multivariate curve resolution». *Analytical Chemistry*, 66 (1994), p. 3337-3344.
85. SMILDE, A. K.; TAULER, R.; HENSHAW, J. M.; BURGESS, L. W.; KOWALSKI, B. R. «Multicomponentndetermination of chlorinated hydrocarbons using a reaction-based chemical sensor. Part III Medium-rank second-order calibration with restricted Tucker models». *Analytical Chemistry*, 66 (1994), p. 3345-3351.

86. TAULER, R.; SMILDE, A.; KOWALSKI, B. «Selectivity, local rank, three- way data analysis and ambiguity in multivariate curve resolution». *Journal of Chemometrics*, 9 (1995), p. 31-58.
87. MANNE, R. « On the resolution problem in hyphenated chromatography». *Chemometrics and Intelligent Laboratory Systems*, 27 (1995), p. 89-94.
88. GELADI, P.; WOLD, S. «Local Principal Component Models, Rank Maps and Contextuality for Curve resolution and Multi-way Calibration Interference». *Chemometrics and Intelligent Laboratory Systems*, 2 (1987), p. 273.
89. TAULER, R. «Multivariate curve resolution applied to second order data». *Chemometrics and Intelligent Laboratory Systems*, 30 (1995), p. 133-146.
90. LAWSON, C. L.; HANSON, R. J. «Solving least squares problems. Prentice hall». *Chapter*, 23 (1974), p. 161.
91. *nns procedure, Optimization Toolbox. MATLAB High Performance Numeric Computation and Visualization Software*. Natick, Mass.: The Math Works Inc., 1990.
92. JUAN DE, A.; HEYDEN, V.; TAULER, R.; MASSART, D. L. «Assesment of new constraints applied to the Alternating Least Squares (ALS) method». *Analytica Chimica acta*. [En premsa].
93. SAURINA, X.; HERNÁNDEZ-CASSOU, S.; TAULER, R. «Multivariate curve resolution applied to continuous flow spectrophotometric titrations. Study of the reaction between amino-acids and 1,2-naphtoquinone-4- sulfonic acid». *Analytical Chemistri*, 67 (1995), p. 3722-3726.
94. TAULER, R.; MARQUES, M. I.; CASASSAS, E. «Second order multivariate curve resolution applied to trilinear data: Study of spectrofluorimetric acid - base titrations of salicylic acid». [Enviat a publicar].
95. TAULER, R.; IZQUIERDO-RIDORSA, A.; GARGALLO, R.; CASASSAS, E.; KUDREV, A. «Study of the protonation and interaction of the polyadenylic -polyuridylic acid with 2,2-bipyridine in aqueous solution». [Enviat a publicar].
96. MENDIETA, J.; DÍAZ-CRUZ, M. S.; TAULER, R.; ESTEBAN, M. «Aplication of Multivariate Curve resolution to Voltammetric Data. Part II: Study of Metal -Binding Properties of the Peptides». *Analytical Biochemistry*, 240 (1996), p. 134-141.
97. MARQUARDT, D. W. *J. Soc. Ind. Appl. Match*, 11 (1963), p. 431-441.
98. CASASSAS, E.; GARGALLO, R.; IZQUIERDO-RIDORSA, A.; TAULER, R. «Application of a multivariate curve resolution procedure to the study of the acid-base and copper (II) complexation equilibria of polycititylic acid». *Reactive Polymers*, 27 (1995), p. 1-14.
99. GARGALLO, R.; TAULER, R.; IZQUIERDO-RIDORSA, A. «Study of the acid- base and copper (II) complexation equilibria of poly (inosinic)-poly (cyttidylic) acid by application of a multivariate curve resolution procedure ». *Biopolymers* (1997). [En premsa].
100. IZQUIERDO-RIDORSA, A.; CASASSAS, E.; GARGALLO, R.; MARQUES, I.; TAULER, R. «A comparative study of polyelectrolytic effects and conformational changes in several purinic and pyrimidinic homopolynucleotides». *Reactive Polymers*, 28 (1996), p. 127-137.

101. DE JUAN, A.; FONRODONA, G.; GARGALLO, R.; IZQUIERDO-RIDORSA, A.; TAULER, R.; CASASSAS, E. «Application of a self-modelling curve resolution approach to the study of solvent effects in the acid-base and copper (II)-complexing of polyuridylic acid». *Journal of Inorganic Biochemistry*, 63 (1996), p. 155-173.
102. GARGALLO, R.; TAULER, R.; IZQUIERDO-RIDORSA, A. «Influence of selectivity and polyelectrolite effects on the performance of self modelling and hard-modelling approaches applied to the study of acid-base equilibria of polyelectrolites». *Analytica Chimica Acta*, 331 (1996), p. 195-205.
103. GARGALLO, R.; TAULER, R.; IZQUIERDO-RIDORSA, A. «Application of a multivariate curve resolution method to the study of the melting behaviour and conformational changes of synthetic and natural polynucleotides». *Anal. Chem.* 1997. [En premsa].
104. LACORTE, S.; BARCELÓ, D.; TAULER, R. «Quantitation and Validation of trace herbicides using on-line solid-phase extraction followed by liquid chromatography diode-array and multivariate curve resolution». *Journal of Chromatography A*, 697 (1995), p. 345-355.
105. TAULER, R.; LACORTE, S.; BARCELÓ, D. «Application of multivariate curve self-modelling curve resolution for the quantitation of trace levels of organophosphorous pesticides in natural waters from interlaboratory studies». *Chromatography A*, 1996, p. 117-183.
106. GARGALLO, R.; CUESTA-SÁNCHEZ, F.; MASSART, D. L.; TAULER, R. «Validation of alternating least squares multivariate curve resolution for the resolution and quantitation of overlapped peaks obtained in liquid chromatography with diode array detection». *Trends in Analytical Chemistry*, 15 (1996), p. 279-286.
107. SAURINA, J.; HERNÁNDEZ-CASSOU, S.; TAULER, R. «Continuous Flow Titration System for the generation of multivariate spectrophotometric data in the study of acid-base equilibria». *Analytica Chimica Acta*, 312 (1995), p. 189-198.
108. SAURINA, J.; HERNÁNDEZ-CASSOU, S.; TAULER, R. «Second order calibration by multivariate curve resolution applied to the flow injection determination of mixture of some aminoacids». *Analytical Chim. Acta*, 335 (1996), p. 41-49.
109. SAURINA, J.; HERNÁNDEZ-CASSOU, S.; TAULER, R. «Kinetic determination of aminoacid mixtures by second order multivariate curve resolution». *Anal. Chem.* 1997 [En premsa].
110. IZQUIERDO-RIDORSA, A.; SAURINA, J.; HERNÁNDEZ-CASSOU, S.; TAULER, R. «Second Order Multivariate Curve Resolution applied to rank deficient data obtained from acid-base spectrophotometric titrations of mixtures of nucleic bases». *Chemometrics Intelligent Laboratory Systems*, (1997). [En premsa].
111. DÍAZ, J. M.; TAULER, R.; GRABARIC, B. S.; ESTEBAN, N.; CASASSAS, E. «Application of Multivariate Curve resolution to the Voltammetric Data. Part I. Study of Zn (II) Complexation with some polyelectrolytes». *Journal of Electroanalytical Chemistry*, 393 (1995), p. 7-16.
112. DÍAZ-CRUZ, M. S.; MENDIETA, J.; TAULER, R.; ESTEBAN, M. «Cadmium binding properties of glutathione: a chemometrical analysis of Voltammetric data». *J. of Inorg. Biochem.* (1997). [En premsa].

113. GRABARIC, B. S.; TAULER, R.; ESTEBAN, M.; CASASSAS, E. «Application of multivariate curve resolution to the voltametric data. Part III: Factor analysis ambiguities in the study of weak consecutive complexation of metal ion with ligand». *Anal. Chim. Acta* (1997). [En premsa].
114. SERRA, J.; TAULER, R.; BAYONA, J.; TOLOSA, I. «Input Characterization of Sedimentary Organic Chemical Markers in the North Western Mediterranean Sea by Exploratory Data Analysis». [Enviat a publicar].

ÍNDIX DE FIGURES

FIGURA 1. Tipus de sensors en química.

FIGURA 2. Exemples de tensors de dades obtinguts amb sensors d'ordre superior: tensors de segon ordre 2-d (HPLC-DAD, temps d'elució i longitud d'ona), tensors de tercer ordre 3-d (HPLC-DAD, mostres, temps d'elució i longitud d'ona), i tensors de quart ordre (anàlisi de mostres diferents per fluorescència: mostres, temps del procés, longitud d'ona d'excitació, longitud d'ona d'emissió).

FIGURA 3. Restricció d'igualtat de forma en els perfils de concentració, durant l'optimització minimoquadràtica de dades, amb estructura de segon ordre trilineal.

FIGURA 4. Perfils de concentració, emprats en la simulació de les matrius de dades 1A i 1B. S'han emprat, també, parts no selectives d'aquests perfils, en la simulació de les matrius de dades 1C, 1D, 1E i 1F (vegeu la descripció de les dades).

FIGURA 5. Perfils espectrals (amb canals selectius) emprats en la simulació de les matrius de dades 1A i 1C.

FIGURA 6. Perfils espectrals (sense canals selectius) emprats en la simulació de les matrius de dades 1B, 1D, 1E, 1F, 2C, 2D, 2E i 2F.

FIGURA 7. Perfils d'elució cromatogràfica, emprats en la simulació de la matriu de dades 2A.

FIGURA 8. Perfils espectrals, emprats en la simulació de les matrius de dades 2A i 2B.

FIGURA 9. Perfils de concentració, emprats en la simulació de les matrius de dades 2B, 2C, 2D, 2E i 2F.

FIGURA 10. Perfils espectrals, emprats en la simulació de les matrius de dades 3A, 3B i 3C.

FIGURA 11. Distribució d'espècies (perfils de concentració) emprada en la simulació de les matrius de dades 3A, 3B i 3C.

FIGURA 12. Detecció de la selectivitat, en l'ordre de les concentracions de la matriu 1B, emprant el mètode FSMWEFA.⁶¹

FIGURA 13. Detecció de la selectivitat, en l'ordre espectral de la matriu 1B emprant el mètode FSMWEFA.⁶¹

FIGURA 14. Detecció de la selectivitat, en l'ordre de les concentracions (elució) de la matriu 2A (fig. 7) emprant el mètode FSMWEFA.⁶¹

FIGURA 15. Detecció de la selectivitat, en l'ordre de les concentracions (pH) de la matriu 3A (fig. 11) emprant el mètode FSMWEFA⁶¹ i un nivell del soroll de fons, amb una desviació estàndard de 0,001 unitats.

FIGURA 16. Detecció de la selectivitat, en l'ordre espectral de la matriu 3A (fig. 10) emprant el mètode FSMWEFA.⁶¹

FIGURA 17. Detecció de la selectivitat, en l'ordre de les concentracions (pH) de la matriu 3A (fig. 11) emprant el mètode FSMWEFA⁶¹ i un nivell del soroll de fons, amb una desviació estàndard de 0,000001 unitats.

FIGURA 18. Anàlisi de l'evolució dels factors (EFA)³⁹ a la matriu 3A (fig. 11). 1, 2 i 3 són les línies corresponents al primer, el segon i el tercer factor, valors singulars de les submatrius successives, obtingudes en augmentar el pH. 1b, 2b i 3b són les línies corresponents al primer, el segon i el tercer factor, valors singulars de les submatrius successives, obtingudes en disminuir el pH (vegeu també les referències⁵⁴⁻⁵⁶ per aplicacions recents de l'esmentat mètode). A l'apèndix 5.2 es dona el programa MATLAB per a fer l'anàlisi EFA.

FIGURA 19. Comparació dels perfils espectrals en l'anàlisi de la matriu augmentada 1E+1F. Les línies discontinues són els perfils recuperats i les contínues, els perfils teòrics.

FIGURA 20. Comparació de la distribució d'espècies, recuperada a partir de l'anàlisi de la matriu 3B (fig. 11) i de la distribució d'espècies vertadera. Les línies contínues són els perfils teòrics; les línies discontinues són els perfils recuperats en l'anàlisi individualitzada de la matriu 3B; les línies - - són els perfils recuperats en l'anàlisi de la matriu augmentada 3A+3B+3C (nivell de soroll amb una desviació estàndard de 0,001 unitats); els símbols + són els perfils recuperats, en l'anàlisi simultània de la matriu augmentada [3A;3B;3C] (nivell de soroll amb una desviació estàndard de 0,000001 unitats).

FIGURA 21. Comparació dels perfils espectrals obtinguts en l'anàlisi de la matriu augmentada 3A+3B. Les línies discontinues són els perfils recuperats i les contínues són els perfils teòrics.

FIGURA 22. Valoracions espectromètriques. Esquema de l'adquisició de dades.

FIGURA 23. Diagrames de distribució abstractes, obtinguts mitjançant l'anàlisi de l'evolució dels factors EFA, de dues matrius de dades corresponents a dues valoracions espectromètriques (A i B) del sistema Cu(II)-polietilenimina. Les línies més gruixudes són les estimacions inicials, dels perfils de concentració de les diferents espècies. Els nombres de l'esquerra i de la dreta donen els factors creixents i decreixents en l'anàlisi EFA, endavant i endarrere.³⁹ Matriu A pH 1,52-4,67; Matriu B pH 1,43-11,3.

FIGURA 24. Distribució d'espècies, obtinguda en el tractament conjunt de les matrius A i B. (1) espècie Cu(II) lliure; (2-4) espècies de complexació successiva del Cu(II) amb la poliE.

FIGURA 25. Espectres de les 4 espècies detectades i resoltes, en les dues valoracions espectromètriques. (1) espècie Cu(II) lliure; (2-4) espècies de complexació successiva del Cu(II) amb la poliE.

FIGURA 26. Esquema de diferents matrius de dades i disposicions de les matrius: *a*) tractament individual de cada matriu per separat; *b*) tractament simultani de les matrius de dades (àcid-base o complexació) obtingudes amb cada tècnica (UV o CD). *c*) tractament simultani de les matrius de dades (UV i CD) obtingudes amb cada tipus de valoració (àcid-base i complexació). *d*) tractament simultani de totes les matrius de dades (UV, CD, àcid-base i complexació).

FIGURA 27. Espectres purs resolts en l'anàlisi de la matriu augmentada, amb totes les dades experimentals (valoracions espectromètriques UV, CD, àcid-base i complexació). A l'esquerra, es mostra l'espectre d'absorció i a la dreta, l'espectre de dicroisme circular. Els nombres 1, 2, 3 corresponen a les diferents conformacions de l'àcid poliadenílic protonat; el 4, correspon a l'espècie totalment desprotonada i el 5, al possible complex de Cu(II) i àcid poliadenílic desprotonat.

FIGURA 28. Diagrama de distribució d'espècies, obtingut en l'anàlisi de la matriu augmentada, amb totes les dades experimentals (valoracions espectromètriques UV, CD, àcid-base i complexació). Els nombres de les corbes representen el mateix que a la figura 27. Cada valoració és separada per una línia discontinua. Valoració 1: 17 espectres, des de pH 1,96 fins a pH 5,92 (sense ió Cu(II)). Valoració 2: 21 espectres, des de pH 1,45 fins a pH 6,88 (sense ió Cu(II)). Valoració 3: 21 espectres, des de pH 1,43 fins a pH 6,05 (amb ió Cu(II)). Valoració 4: 22 espectres, des de pH 1,36 fins a pH 6,18 (amb ió Cu(II)).

FIGURA 29. Variació del logaritme de la constant de protonació de l'àcid poliadenílic, respecte al grau de protonació, en 6 valoracions diferents (UV i CD). La constant és obtinguda a partir del valor del pH, mesurat experimentalment i de les concentracions d'equilibri de les espècies 3 i 4, de la figura 28.

FIGURA 30. Variació del logaritme de la constant de complexació de l'àcid poliadenílic, amb l'ió Cu(II) respecte al grau de complexació, en 6 valoracions diferents (UV i CD).

La constant és obtinguda a partir de la concentració de Cu(II) lliure, mesurada experimentalment (elèctrode d'ió selectiu de Cu(II) i de les concentracions d'equilibri de les espècies 4 i 5, de la figura 28).

FIGURA 31. Diagrama 3D del pic cromatogràfic LC-DAD, corresponent a la mescla no resolta dels pesticides naftol i pirimicarb.

FIGURA 32. Detecció de la selectivitat en l'ordre de les concentracions (elució) de la matriu de dades, corresponent a la figura 31, emprant el mètode FSMWEFA.⁶¹ Finestra de 5 espectres. Es detecten 2 components, que coelueixen entre 0,450 i 0,510 min. Existeix selectivitat per als dos components.

FIGURA 33. Espectres resolts dels dos components (pirimicarb i naftol) coeluïts (a la part superior; naftol en línia discontinua i pirimicarb en línia contínua). Perfils d'elució dels mateixos components coeluïts (a la part inferior, mostres dels components purs en línia discontinua i mostra de la mescla en línia contínua; en tots dos casos, el perfil més alt és el del pirimicarb).

FIGURA 34. Exemple dels espectres d'absorció FT-IR adquirits durant el seguiment d'una realització del procés.

FIGURA 35. Espectres de la 2a derivada dels espectres de la figura 34.

FIGURA 36. Espectres calculats de la 2a derivada (fig. 35) considerant 3 components principals.

FIGURA 37. Estimació inicial dels canvis de concentració durant una realització del procés, a partir de l'anàlisi de l'evolució dels factors (EFA).³⁹

FIGURA 38. Espectres unitaris (purs) dels 3 components detectats. Han estat calculats a partir de l'anàlisi simultània de 8 realitzacions del mateix procés (795 espectres).

FIGURA 39. Canvis de concentració, deduïts a partir de l'anàlisi de les 8 realitzacions del procés. L'eix de les abscisses (temps) dóna una indicació del nombre d'espectre considerat, dins del nombre total d'espectres considerats (795 espectres en total).

FIGURA 40. Sensor òptic basat en la reacció química de Fujiwara.

FIGURA 41. Resposta 3D del tricloroetà, TCA (66,9 ppm) en la reacció de Fujiwara.

FIGURA 42. Resposta 3D del tricloroetà, TCE (7,31 ppm.) en la reacció de Fujiwara.

FIGURA 43. Resposta 3D de cloroform, CHCl₃ (4,45 ppm) en la reacció de Fujiwara.

FIGURA 44. Resposta 3D d'una mescla de TCA (66,9 ppm), TCE (2,93 ppm) i CHCl_3 (1,48 ppm) en la reacció de Fujiwara.

FIGURA 45. a) Matriu augmentada segons les columnes (on es mantenen les columnes o longituds d'ona) formada a partir de tres matrius individuals, corresponents a tres mostres analitzades. b) Matriu augmentada segons les files (on es mantenen les files o temps de reacció) formada a partir de les tres mateixes matrius individuals.

FIGURA 46. Perfils de concentració de l'espècie 1 (fig. 49) trobats en l'anàlisi de 4 mostres pures de TCA.

FIGURA 47. Perfils de concentració de les espècies 2 i 3 (fig. 49) trobats en l'anàlisi de 5 mostres pures de TCE.

FIGURA 48. Perfils de concentració de les espècies 4 i 5 (fig. 49) trobats en l'anàlisi de 4 mostres pures de CHCl_3 .

FIGURA 49. Espectres de les 5 espècies detectades en les mostres analitzades. L'espècie 1 és del TCA, la 2 i la 3 són del TCE; la 4 i la 5, del CHCl_3 .

FIGURA 50. Diagrama de flux del mètode de resolució multivariant de corbes.

FIGURA 51. Superposició de perfils d'elució cromatogràfica. En el primer cas, existeixen zones de selectivitat per el primer i el tercer component, però no n'hi ha cap per el perfil cromatogràfic del mig. En el segon cas, en canvi, el perfil del component minoritari es troba totalment inclòs (*embedded*) dins el perfil del component majoritari.

FIGURA 52. Esquema de diferents procediments experimentals d'anàlisi en continu.

