

UNIVERSIDAD NACIONAL DEL LITORAL
Facultad de Bioquímica y Ciencias Biológicas



**Tesis para la Obtención del Grado Académico de Doctor en
Ciencias Biológicas**

**Aplicación de modelado quimiométrico a datos de
segundo orden como herramienta para la
corrección de problemas de interferencias en la
determinación de analitos en muestras complejas**

**Lic. Agustina Violeta
Schenone**

Dr. Héctor C. Goicoechea
Director de Tesis

Dra. María J. Culzoni
Co-Director de Tesis

Cátedra de Química Analítica I – Laboratorio de Desarrollo Analítico y
Quimiometría - Facultad de Bioquímica y Ciencias Biológicas -
Universidad Nacional del Litoral

2014

A mis padres **Graciela y Pototo**
A mis hermanos **Ignacio y Florencia**
A **Fabio**

*“Nunca consideres el estudio como un deber,
sino como una oportunidad para penetrar en
el maravilloso mundo del saber”*

Albert Einstein

AGRADECIMIENTOS

La realización de esta tesis no habría sido posible sin la valiosa ayuda de mis Directores: Héctor y Julia. A ambos, mi sincero agradecimiento no sólo por la formación académica, la cual sin duda fue muy importante, sino también por la confianza depositada en mí y por su ejemplo de trabajo y dedicación. Especialmente, quiero agradecer a Héctor por brindarme la oportunidad de formar parte de su grupo de investigación y por la motivación y el apoyo recibido a lo largo de estos años. Igualmente, hago extensivo este agradecimiento a Julia por su guía en la programación de tareas a realizar, su valioso aporte en el análisis de resultados y por su experiencia que fue de mucha ayuda en la realización de este trabajo.

A Luciana, por su apoyo y amistad, por estar siempre atenta y dispuesta a colaborar en todo.

También quiero agradecer a todas las personas que han colaborado conmigo y me han acompañado durante estos años de trabajo. A Quela, Carla, Romi, Vane, Mili, Pablito, Yami, Gabi y demás integrantes de la Cátedra de Química Analítica I, del Laboratorio de Control de Calidad de Medicamentos y del Laboratorio de Desarrollo Analítico y Quimiometría. Gracias por poder hacer del trabajo de todos los días un espacio para compartir alegrías, tristezas, problemas, proyectos...

A Nilda, Alberto, Germán y todos los integrantes de Química Analítica II por los momentos compartidos y por el apoyo brindado. Quiero hacer extensiva mi gratitud a Nilda, que siempre confió en mí y me brindó su aliento constantemente.

Agradezco con todo mi cariño y mi amor a las personas que hicieron todo en la vida para que yo pudiera lograr mis sueños, por motivarme y darme la mano siempre que lo necesité, a ustedes papá y mamá. A mis queridos hermanos Ignacio y Florencia.

A Fabio, por compartir tu vida conmigo, por tu paciencia y comprensión, por tu amor incondicional. Gracias por estar siempre a mi lado.

A todas mis amigas y hermanas del corazón, especialmente a Mai, Bay, Cele, Eli, Emi, Nim, Rou, Mari, Regi, Juli M., Cami, Luli, Anto, Juli S. y Delfi, simplemente por ser mis amigas. Por ayudarme y alentarme a ser mejor persona cada día.

A la Facultad de Bioquímica y Ciencias Biológicas (UNL) y al Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET) por el apoyo recibido durante el desarrollo de esta tesis.

A la Comisión Fulbright y la Fundación Bunge y Born por haber financiado mi estadía en el Departamento de Química de la Universidad de Florida Central, Orlando, Estados Unidos.

La presente tesis dio lugar a las siguientes publicaciones y presentaciones a reuniones científicas:

PUBLICACIONES

Fast chromatographic method for the determination of dyes in beverages by using HPLC-DAD data and second order algorithms. María J. Culzoni, Agustina V. Schenone, Natalia E. Llamas, Mariano Garrido, Maria S.Di Nezio, Beatriz S. Fernández Band, Héctor C. Goicoechea. Journal of Chromatography A, 1216, 16 October 2009, pp. 7063-7070.

Determination of tartrazine in beverage samples by stopped-flow analysis and three-way multivariate calibration of non-linear kinetic-spectrophotometric data. Agustina V. Schenone, María J. Culzoni, Nilda R. Marsili, Héctor C. Goicoechea. Food Chemistry, 138, 15 Mayo–1 Junio de 2013, pp. 1928-1935.

Second-order advantage achieved by modeling excitation–emission fluorescence matrices affected by inner filter effects using a strategy which combines standardization and calibration: Reducing experimental and increasing analytical sensitivity. Agustina V. Schenone, María J. Culzoni, María Martínez Galera, Héctor C. Goicoechea. Talanta, 109, 15 de mayo de 2013, pp. 107-115.

Total synchronous fluorescence spectroscopic data modeled with first- and second-order algorithms for the determination of doxorubicin in human plasma. Agustina V. Schenone, María J. Culzoni, Andres D. Campiglia, Héctor C. Goicoechea. Analytical & Bioanalytical Chemistry, 405, octubre 2013, pp. 8515-8523.

REUNIONES CIENTÍFICAS

V Congreso Argentino de Química Analítica. Bahía Blanca, Argentina (2009). *Redes neuronales con ventaja de segundo orden aplicadas a la cuantificación de tartrazina en presencia de interferencias.* Agustina V. Schenone, María J. Culzoni, Nilda R. Marsili, Héctor C. Goicoechea. Tipo de participación: póster.

V Congreso Argentino de Química Analítica. Bahía Blanca, Argentina (2009). *Un método rápido para la determinación de colorantes en bebidas por HPLC-DAD y modelado quimiométrico con ventaja de segundo orden.* María J. Culzoni, Agustina V. Schenone, Natalia E. Llamas, Mariano Garrido, María S. Di Nezio, Beatriz S. Fernández Band, Héctor C. Goicoechea. Tipo de participación: póster.

VI Congreso Argentino de Química Analítica. Santa Fe, Argentina (2011). *Uso de datos de segundo orden y transferencia de calibración para reducir el número de muestras de calibración cuando se tiene presencia de interferencias no modeladas y efecto matriz.* Agustina V. Schenone, María J. Culzoni, Héctor C. Goicoechea. Tipo de participación: póster.

III Workshop de Quimiometría. Arembepe, Brasil (2012). *Uso de datos de segundo orden y transferencia de calibración para reducir el número de muestras de calibración cuando se tiene presencia de interferencias no modeladas y efecto matriz.* Agustina V. Schenone, María J. Culzoni, Héctor C. Goicoechea. Tipo de participación: póster.

XIII Chemometrics in Analytical Chemistry Congress. Budapest, Hungría (2012). "Different strategies for second order data generation and modeling. Applications to solve different analytical problems". L. Vera-Candioti, Y.S. Caro, M.M. De Zan, R. Brasca, M.R. Alcaráz, M. Marchisio, F. Piceh, M.M. Cámara, A.V. Schenone, M.J. Culzoni, H.C. Goicoechea. Tipo de participación: póster.

VII Congreso Argentino de Química Analítica. Mendoza, Argentina (2013).
Desarrollo de un nuevo algoritmo para la obtención de la ventaja de segundo orden en el procesamiento de datos no bilineales. Agustina V. Schenone, María J. Culzoni, Héctor C. Goicoechea. Tipo de participación: póster.



Índice

Índice de Contenidos

	Página
CAPÍTULO 1. Introducción	1
1.1. Calibración multivariada	3
1.1.1. Clasificación de datos instrumentales	3
1.1.2. Algoritmos quimiométricos	7
1.1.3. Diseños de experimentos en calibración multivariada	10
1.1.4. Aplicaciones	11
1.1.4.1. Alimentos	11
1.1.4.2. Contaminantes emergentes	12
1.1.4.3. Matrices biológicas	14
1.2. Herramientas Quimiométricas: algoritmos de calibración multivariados de segundo orden	15
1.2.1. Resolución Multivariada de Curvas	15
1.2.1.1. Resolución Multivariada de Curvas: datos de dos vías	15
1.2.1.2. Resolución Multivariada de Curvas Extendida: datos de tres vías	19
1.2.1.3. Cifras de mérito en MCR-ALS	23
1.2.2. U-PLS/RBL	24
1.2.2.1. Número óptimo de factores (A)	26
1.2.2.2. Predicción basada en el <i>score</i> de la muestra	27
1.2.2.3. Cifras de mérito en U-PL/RBL	30
1.2.3. U-PCA/RBL	31
1.2.4. Análisis Paralelo de Factores (PARAFAC)	33
1.2.4.1. Cifras de mérito en PARAFAC	36
1.3. Otras Herramientas Quimiométricas	37
1.3.1. SIMPLISMA	37
1.3.2. Estandarización Directa Sucesiva (PDS)	39
1.4. Regiones Elípticas de Confianza Conjunta (EJCR)	40
1.5. Objetivos	43
1.5.1. Objetivos generales	43
1.5.2. Objetivos específicos	43

Índice de Contenidos

CAPÍTULO 2: Desarrollo de un método cromatográfico rápido para la determinación de colorantes en bebidas no alcohólicas, utilizando datos obtenidos por cromatografía líquida de alta resolución con detector de arreglo de diodos y algoritmos quimiométricos de segundo orden.	44
2.1. Introducción	45
2.2. Objetivos	47
2.3. Programas	48
2.4. Materiales y Métodos	48
2.4.1. Reactivos	48
2.4.2. Equipamiento	48
2.4.3. Procedimiento	48
2.4.3.1 Estándares de calibración y mezclas de validación	48
2.4.3.2. Muestras de jugos artificiales	49
2.4.3.3. Método de referencia	49
2.4.3.4. Método cromatográfico rápido	49
2.5. Resultados y Discusión	49
2.5.1. Aspectos generales	49
2.5.2. Aplicación de algoritmos de segundo orden	53
2.5.2.1. Mezclas de validación	53
2.5.2.2. Muestras reales	57
2.6. Conclusiones	60
CAPÍTULO 3: Determinación de tartrazina en bebidas mediante análisis por inyección en flujo (FIA) con flujo interrumpido y calibración multivariada de datos cinéticos no lineales.	62
3.1. Introducción	63
3.2. Objetivos	65
3.3. Programas	65
3.4. Materiales y Métodos	66
3.4.1. Reactivos y soluciones	66
3.4.2. Equipamiento y procedimiento instrumental	66
3.4.3. Procedimiento analítico	67
3.4.3.1. Estándares de calibración y mezclas de validación	67
3.4.3.2. Muestras de bebidas comerciales	69
3.4.3.3. Procedimiento cromatográfico	69

Índice de Contenidos

3.5. Resultados y Discusión	70
3.5.1. Características espectrales y cinéticas	70
3.5.2. Cuantificación de tartrazina en las muestras de validación J1 y J2	71
3.5.3. Predicción de tartrazina en bebidas comerciales	78
3.6. Conclusiones	79
CAPÍTULO 4: Obtención de la ventaja de segundo orden a través del modelado de matrices de excitación-emisión afectadas por efecto de filtro interno utilizando transferencia de calibración: reducción del número de experimentos y aumento de la sensibilidad analítica.	81
4.1. Introducción	82
4.2. Objetivos	86
4.3. Programas	86
4.4. Materiales y Métodos	86
4.4.1. Reactivos y soluciones	86
4.4.2. Equipamiento	87
4.4.3. Datos	87
4.4.3.1. Datos simulados	87
4.4.3.2. Datos experimentales	89
4.5. Resultados y Discusión	91
4.5.1. Datos simulados	91
4.5.2. Datos experimentales	94
4.5.3. Muestras de agua	98
4.6. Conclusiones	101
CAPÍTULO 5: Modelado de matrices de fluorescencia sincrónica total con algoritmos de primer y segundo orden. Determinación de doxorrubicina en plasma humano.	103
5.1. Introducción	104
5.2. Objetivos	110
5.3. Programas	110
5.4. Materiales y Métodos	111
5.4.1. Reactivos y soluciones	111
5.4.2. Preparación de muestras de calibración y validación	111
5.4.3. Equipamiento	112
5.5. Resultados y Discusión	112

Índice de Contenidos

5.5.1. Consideraciones generales	112
5.5.2. MCR-ALS	115
5.5.3. U-PLS/RBL	121
5.5.4. Cifras de mérito	122
5.6. Conclusiones	124
CAPÍTULO 6: Ventaja de segundo orden para datos no bilineales. Desarrollo de un nuevo algoritmo adaptado de PLS/RBL.	
6.1. Introducción	126
6.2. Objetivos	127
6.3. Programas	127
6.4. Materiales y Métodos	127
6.4.1. Datos simulados	127
6.4.2. Datos experimentales	130
6.5. Desarrollo teórico del nuevo método U-PLS/RMNB	130
6.6. Resultados y Discusión	133
6.6.1. Datos Simulados	133
6.6.2. Datos Experimentales	137
6.7. Conclusiones	139
CAPÍTULO 7: Conclusiones generales	
CAPÍTULO 8: Bibliografía	145

Índice de Figuras

	Página
Figura 1.1. Datos de orden cero. Se genera un solo dato por muestra, y el conjunto de calibración es un vector que contiene las medidas para las I muestras de calibración. En la curva de calibrado, a cada valor de concentración le corresponde un dato de señal.	4
Figura 1.2. Datos de primer orden. Se genera un vector de datos por cada muestra, por ejemplo un espectro medido a J longitudes de onda. El conjunto de I muestras de calibración genera una matriz de $J \times I$.	5
Figura 1.3. Datos de segundo orden. Se genera una matriz de datos de $J \times K$ por cada muestra I , y el conjunto de calibración resulta ser un arreglo de tres vías.	5
Figure 1.4. Arreglo tridimensional típico en donde se muestra cómo decrece la absorbancia de una muestra de tartrazina en presencia de un oxidante, en función de las longitudes de onda y del tiempo de reacción.	6
Figura 1.5. Representación esquemática de datos de tercer orden adquiridos para una sola muestra. Mediante la técnica de espectroscopia de fluorescencia puede obtenerse una híper matriz de excitación-emisión resuelta en el tiempo.	6
Figura 1.6. Representación del modelo MCR para una serie de I mezclas que contienen hasta N componentes medidos a J longitudes de onda.	17
Figura 1.7. Representación del desdoblamiento de una híper matriz de tres vías en otra de dos vías. Apilamiento de matrices individuales de calibración y una muestra incógnita para ser analizadas simultáneamente. A: aumento en el sentido de las columnas. B: aumento en el sentido de las filas.	20
Figura 1.8. Porcentaje de la variancia explicada en función del número de componentes.	32
Figura 1.9. Desvío estándar de los residuos de la regresión calculado por validación cruzada ($S_{fit_{CV}}$) en función del número de componentes.	32
Figura 1.10. Representación esquemática de un paso de PDS. Se realiza una regresión entre los valores a la longitud de onda j th en la situación A con los valores en la ventana $(j-k..j+k)$ de la situación B. El vector de regresión resultante forma la columna j th de la diagonal de la matriz de transformación F .	40
Figura 2.1. Separación cromatográfica completa (13 min) de los tres colorantes y otros componentes en una muestra de jugo (muestra número 4, Tabla 1.2) al utilizar el método propuesto por Pereira Alves y col. (Pereira Alves y col. 2008).	50
Figura 2.2. (A) Superficie correspondiente a una mezcla ternaria cuando la separación cromatográfica se lleva a cabo en 1.85 min y se registran los espectros entre 440 y 570 nm. (B) Perfiles de elución registrados en los	

Índice de Figuras

máximos correspondientes de cada colorante (T a 450 nm, AO a 490 nm y A a 530 nm) y (C) espectros de los tres colorantes, obtenidos al aplicar MCR-ALS a la matriz de datos de la mezcla (T: línea azul, A: línea verde y AO: línea roja).	51
Figura 2.3. Cromatogramas correspondientes a una mezcla ternaria en línea sólida azul (muestra 1 en Tabla 1.1) y a una muestra real en línea de puntos roja (muestra 5 en Tabla 1.2) registrados a 440 nm.	52
Figura 2.4. Cromatogramas correspondientes a las quince mezclas de validación registrados a 440 nm.	53
Figura 2.5. Cromatogramas (registrados a 500 nm) correspondientes a soluciones estándar de T (8.00 mg L ⁻¹) en línea sólida azul, A (1.53 mg L ⁻¹) en línea de puntos roja y a una mezcla de ambos colorantes en las mismas concentraciones en línea de círculos verde.	53
Figura 2.6. Regiones elípticas para el conjunto de datos globales de predicción aplicando U-PLS/RBL y MCR-ALS a las 15 mezclas de validación de la Tabla 1.1.	57
Figura 2.7. Cromatogramas correspondientes a las siete muestras reales registradas a 440 nm.	58
Figura 2.8. (A) Perfiles de tiempo extraídos por MCR-ALS cuando se analizó la muestra 7 (Tabla 1.2) que contiene los tres colorantes (T: línea de círculos verde, A: línea de cuadrados roja, AO: línea sólida azul) más un interferente (línea de triángulos celeste). (B) Perfiles espectrales.	60
Figura 3.1. Estructura molecular del colorante tartrazina (E-102).	63
Figura 3.2. Espectros de absorción normalizados. T: tartrazina, AB: azul brillante, AO: amarillo ocaso y P: producto de oxidación de AB.	70
Figura 3.3. Curvas cinéticas normalizadas obtenidas para T (línea sólida rosa), AO (línea discontinua con puntos negra), AB (línea discontinua verde) y su producto de degradación (línea de puntos celeste) en las longitudes de onda de sus máximos de absorción.	71
Figura 3.4. Perfiles temporales normalizados de T obtenidos al aplicar MCR-ALS para analizar una muestra de calibración (línea sólida), una muestra del juego J1 (línea discontinua) y una muestra del juego J2 (línea de puntos).	74
Figura 3.5. Perfiles obtenidos al aplicar MCR-ALS para analizar una muestra del juego J2 junto con muestras de calibración. (A) Perfiles cinéticos. (B) Perfiles espectrales. En ambos casos, la línea sólida rosada indica T, la línea discontinua verde indica AB, la línea punteada celeste indica en producto de AB y la línea discontinua con puntos negra indica AO.	75
Figura 3.6. Separación cromatográfica de tartrazina (T) y amarillo ocaso (AO) en una muestra real (muestra BP de la Tabla 2.5) al utilizar el método CLAR ($\lambda = 454$ nm).	79
Figura 4.1. Perfiles de excitación y emisión simulados para el analito puro	

(componente 1, línea sólida azul) e interferente (componente 2, línea discontinua roja), y el analito en una muestra de validación con efecto de filtro interno producido por el componente 2 en ambas dimensiones (línea de puntos negra).	88
Figura 4.2. Representación esquemática de la aplicación de PDS/PARAFAC a una muestra que contiene componentes inesperados y efecto de filtro interno. Ver texto para los detalles de la simbología empleada.	92
Figura 4.3. Gráficos de contorno para una muestra de validación simulada antes (A) y después (B) de PDS/PARAFAC y de una muestra de calibración simulada que sólo contiene el componente 1 (C).	93
Figura 4.4. Gráficos de tres dimensiones de las señales de fluorescencia en función de las longitudes de onda de excitación y emisión para una estándar de 500 ng mL ⁻¹ de FEN en ausencia (A) y presencia (B) de PAR (10.0 µg mL ⁻¹).	95
Figura 4.5. Curvas de calibrado de FEN a $\lambda_{ex} = 215$ nm y $\lambda_{em} = 305$ nm, en ausencia (círculos azules) y presencia (triángulos verdes) de PAR (10.0 µg mL ⁻¹).	96
Figura 4.6. Espectro de emisión de un estándar de FEN (250 ng mL ⁻¹) sin (línea sólida azul) y con (línea de puntos verde) PAR.	96
Figura 4.7. Scores de calibración de FEN (círculos rojos) y de FEN en la muestra de agua de napa analizada (cruz azul).	100
Figura 4.8. Resultados de PARAFAC al predecir FEN (línea sólida azul) en la muestra de agua de napa en presencia de IBU (línea discontinua verde), AAS (línea de puntos roja) y un componente inesperado (línea discontinua con puntos negra): (A) modo excitación, (B) modo emisión	101
Figura 5.1. Superficies de contorno de (A) DOX pura en agua, (B) RF pura en agua y (C) plasma humano desproteinizado.	113
Figura 5.2. (a) Espectros sincrónicos de DOX pura 1.25 µg mL ⁻¹ (línea sólida), plasma con 1.25 µg mL ⁻¹ de DOX (línea discontinua), diferencia entre los espectros de plasma con 1.25 µg mL ⁻¹ de DOX y blanco de plasma (línea discontinua con puntos) y blanco de plasma (línea de puntos), todos a $\Delta\lambda = 80$ nm. (b) Espectros sincrónicos de DOX pura 1.25 µg mL ⁻¹ a varios $\Delta\lambda$ (desde 30 a 160 nm, cada 10 nm). En rojo, el espectro a $\Delta\lambda = 80$ nm.	115
Figura 5.3. Esquema de la matriz C_{sel} utilizada para aplicar la restricción de correlación. NaN: valor no restringido.	117
Figura 5.4. Espectros recuperados en la matriz S^T por MCR-ALS luego de analizar la matriz aumentada que contiene las muestras de validación y calibración. La línea sólida azul corresponde a DOX, la línea discontinua verde a RF y la línea de puntos roja a la señal del plasma. En la imagen insertada, se muestran los espectros de los tres componentes (DOX: línea sólida azul, RF:	

línea discontinua verde, plasma: línea de puntos roja) y el espectro normalizado de DOX pura (línea discontinua con puntos negra), todos a $\Delta\lambda = 80$ nm.	118
Figura 5.5. Scores obtenidos por MCR-ALS correspondientes a DOX (círculos azules), RF (rombos verdes) y plasma (triángulos hacia abajo rojos) en las muestras de calibración. Y DOX (cruces rosadas), RF (cuadrados blancos) y plasma (triángulos hacia arriba grises) en muestras de validación.	119
Figura 5.6. Logaritmo del PRESS en función del número de variables latentes A de U-PLS, obtenidos con datos de calibración (círculos), y logaritmo de los residuos de predicción de U-PLS en función del número de componentes inesperados (N_{int}) (triángulos) al analizar la muestra de validación número dos.	121
Figura 6.1. Espectros de excitación (A) y emisión (B) del analito (línea sólida azul) y del interferente (línea discontinua roja) en el sistema simulado.	128
Figura 6.2. Matriz de excitación-emisión de una mezcla de analito e interferente en el sistema simulado. Las filas de la caja contienen los espectros sincrónicos simulados a diferentes $\Delta\lambda$.	129
Figura 6.3. Espectros sincrónicos del analito (azul) y el interferente (rojo) generados a 25 $\Delta\lambda$ diferentes.	130
Figura 6.4. Representación esquemática de la aplicación de MCR-ALS con la restricción de correlación para la descomposición de una matriz simulada. La matriz resulta del acoplamiento de cinco matrices MFTS de calibración y una de validación desdobladas. El procedimiento se usa para generar ambos perfiles y la concentración del interferente, necesarios para minimizar la Ecuación (5.5).	132
Figura 6.5. Esquema ilustrativo de cómo trabaja el nuevo algoritmo U-PLS/RMNB. Ver el texto para la explicación de los símbolos.	133
Figura 6.6. (A) Primer perfil extraído por RBL a ser aplicado a la muestra 1. (B) Juego de perfiles extraídos que muestran la estructura bilineal de la matriz. (C) Matriz reordenada que se obtuvo al aplicar MCR-ALS para modelar las cinco muestras de calibración junto con la correspondiente a la muestra 1. (D) Juego de perfiles extraídos por MCR-ALS que muestran la estructura no bilineal de la matriz (azul) y los espectros normalizados reales (rojo).	136
Figura 6.7. (A) Contorno correspondiente a la primer MFST estimado por RBL. (B) Perfil extraído por MCR-ALS para ser incorporado en el proceso de minimización en el paso RMNB. (C) matriz MFST real de RF (interferente).	139

Índice de Tablas

	Página
Tabla 2.1. Composición y predicción de las muestras ternarias utilizadas para calibración y validación.	56
Tabla 2.2. Predicción de muestras reales por el método de referencia y el método rápido propuesto acoplado a MCR-ALS y U-PLS/RBL.	59
Tabla 3.1. Composición del juego de validación J1 siguiendo un diseño experimental central compuesto.	68
Tabla 3.2. Composición del juego de validación J2 siguiendo un diseño experimental central compuesto.	68
Tabla 3.3. Resultados de la predicción del juego de validación J1 al aplicar MCR-ALS.	73
Tabla 3.4. Resultados de la predicción del juego de validación J2 al aplicar MCR-ALS.	76
Tabla 3.5. Resultados de la determinación de tartrazina en bebidas comerciales.	78
Tabla 4.1. Composición de las muestras de validación experimentales*.	90
Tabla 4.2. Resultados y cifras de mérito del análisis de datos simulados.	94
Tabla 4.3. Resultados de predicción de FEN para las muestras de validación experimentales.	98
Tabla 4.4. Resultados de predicción de las muestras de agua.	99
Tabla 5.1. Predicción y análisis estadístico de DOX en muestras de validación.	120
Tabla 5.2. Cifras de mérito para MCR-ALS y U-PLS/RBL.	123
Tabla 6.1. Resultados obtenidos al modelar los datos simulados con PARAFAC, U-PLS/RBL y U-PLS/RMNB.	134
Tabla 6.2. Número de componentes inesperados, residuos y concentraciones predichas del analito en la muestra de validación 1 (concentración nominal igual a 1.59) cuando se aplica U-PLS/RBL.	135
Tabla 6.3. Resultados de predicción del juego experimental.	138

Abreviaturas y Símbolos

+	Matriz pseudoinversa (como supraíndice)
^	Valor estimado
$\ \cdot \ $	Norma Euclidea de un vector
\otimes	Producto de Kronecker
γ_n	Sensibilidad analítica
$\ \delta r\ $	Medida del ruido instrumental
A	Número de factores latentes
ALS	Cuadrados Mínimos Alternantes
ANMAT	Administración Nacional de Medicamentos, Alimentos y Tecnología Médica
ANN	Redes Neuronales Artificiales
ANOVA	Análisis de la Variancia
ATLD	Descomposición Trilineal Alternante
APTLD	Descomposición Trilineal Alternante Penalizada
BLLS	Cuadrados Mínimos Bilineales
BLLS/RBL	Cuadrados Mínimos Bilineales seguidos de Bilinealización Residual
CANDECOMP	Descomposición Canónica
CEM	Contaminantes emergentes
CLAR	Cromatografía líquida de alta resolución
CLS	Regresión por Cuadrados Mínimos Clásica
CV	Coefficiente de variación
DAD	Detector de arreglo de diodos
DS	Estandarización Directa
DTLD	Descomposición Trilineal Directa
EC	Electroforesis capilar
EFS	Espectros de fluorescencia sincrónicos
EJCR	Intervalo de Confianza Elíptico Conjunto
FIA	Análisis por inyección en flujo
FL	Detector de fluorescencia
GC	Cromatografía gaseosa
GN	Gauss-Newton
GRAM	Método de Aniquilación de Rango Generalizado

Abreviaturas y Símbolos

GSAM	Método de adición estándar generalizado
IDA	Ingesta diaria aceptable
ILS	Cuadrados Mínimos Inversos
k	Factor de retención
LOD	Límite de detección
LOQ	Límite de cuantificación
LS-SVM	Máquinas de Vectores Soporte con Ajuste por Cuadrados Mínimos
MA	Medioambiente
MCR-ALS	Resolución Multivariada de Curvas mediante Cuadrados Mínimos Alternantes
MEE	Matriz de excitación-emisión
MFST	Matrices de espectroscopia de fluorescencia sincrónicas total
MS	Espectrometría de masa
MLP	Perceptrón Multicapa
NaN	Valores no restringidos
N_c	Número de analitos calibrados
N_{int}	Número interferentes
NLPLS	Cuadrados Mínimos Parciales No Lineales
N-PLS	Cuadrados Mínimos Parciales Multivía
OLS	Cuadrados Mínimos Ordinarios
PARAFAC	Análisis Paralelo de Factores
PARALIND	Factores Paralelos con Perfiles Lineales
PCA	Análisis en Componentes Principales
PCR	Regresión en Componentes Principales
PDS	Estandarización Directa Sucesiva
PLS	Cuadrados Mínimos Parciales
POP	Contaminantes orgánicos persistentes
PRESS	Suma de cuadrados de los errores de predicción
RBF	Funciones de Base Radial
RBL	Bilinealización Residual
REP(%)	Error relativo de predicción porcentual
RMSE	Error cuadrático medio

Abreviaturas y Símbolos

RTL	Trilinealización Residual
SAM	Método de adición estándar
SD	Desviación estándar
SEE	Error estándar de la estimación
SEL	Selectividad
SEN	Sensibilidad
SENASA	Servicio Nacional de Sanidad y Calidad Agroalimentaria
S_{fit}	Desvío estándar de los residuos de la regresión
$S_{fit,CV}$	Desvío estándar de los residuos de la regresión calculado por validación cruzada
SSE	Suma de los errores cuadráticos
SVD	Descomposición en Valores Singulares
SVM	Máquinas de Vectores Soporte
SWATLD	Descomposición Trilineal Alternante Autoponderada
T	Matriz transpuesta (como supraíndice)
TCA	Ácido tricloroacético
TLC	Cromatografía en capa delgada
TLLS/RTL	Cuadrados Mínimos Trilineales seguidos de Trilinealización Residual
U-PCR	Regresión en Componentes Principales Desdoblados
U-PLS	Cuadrados Mínimos Parciales Desdoblados
U-PCA/RBL	Componentes Principales Desdoblados seguidos de Bilinealización Residual
U-PCR/RBL	Regresión en Componentes Principales desdoblados seguidos de Bilinealización Residual
U-PLS/RBL	Cuadrados Mínimos Parciales Desdoblados seguidos de Bilinealización Residual
U-PLS/RMNB	Cuadrados Mínimos Parciales Desdoblados seguidos por Modelado de Residuos de datos No Bilineales
U-PLS/RTL	Cuadrados Mínimos Parciales Desdoblados seguidos de Trilinealización Residual
UV	Ultravioleta
UV-Vis	Ultravioleta-Visible

Abreviaturas y Símbolos

vec

Operación de vectorización

X

Arreglo tridimensional formado por diferentes matrices

Uno de los desafíos de la química analítica moderna es la determinación de distintos tipos de sustancias como aditivos alimentarios, fármacos, metabolitos o agroquímicos que se encuentran presentes en muestras complejas y por lo general en bajas concentraciones. Por otro lado, la búsqueda de nuevos métodos se centra en su confiabilidad y seguridad para determinar estos compuestos con mayor selectividad, sensibilidad y rapidez. En los últimos años la conservación del ambiente ha cobrado mucha importancia, por lo que se busca minimizar el consumo de solventes y reactivos tóxicos.

En la presente tesis se desarrollaron métodos de análisis combinando distintos tipos de señales instrumentales con herramientas quimiométricas adecuadas, que permitieron extraer información relevante de cada muestra y cuantificar los analitos aún en presencia de interferencias no modeladas. Específicamente, en este trabajo se analizaron datos de dos vías con algoritmos que presentan la ventaja de segundo orden para solucionar distintos tipos de problemas analíticos, entre los que podemos mencionar datos no trilineales, efecto de filtro interno y datos con pérdida de bilinealidad.

En los próximos párrafos se resumen los métodos descriptos en los distintos capítulos.

En el Capítulo 1 se desarrolla la Introducción donde se describen los distintos tipos de datos analíticos que se pueden obtener y los métodos y algoritmos quimiométricos utilizados a lo largo de esta tesis.

En el Capítulo 2 se presenta el desarrollo de un método cromatográfico rápido con detección de arreglo de diodos para la determinación de tres colorantes sintéticos: tartrazina, amaranto y amarillo oca, en muestras de bebidas no alcohólicas comerciales. El método consistió en utilizar una fase móvil que permitiera una elución rápida de los componentes de las muestras. La generación de datos de segundo orden presentó corrimiento de picos entre los diferentes cromatogramas, originando datos no trilineales. Este tipo de datos no pudieron ser modelados satisfactoriamente por el algoritmo U-PLS/RBL. Por el contrario, MCR-ALS demostró ser una herramienta adecuada y los resultados obtenidos fueron aceptables. Las predicciones fueron comparadas con las obtenidas por un método cromatográfico clásico. El tiempo de análisis del método propuesto se redujo a un 14.2% con respecto a la separación completa de los colorantes permitiendo

ahorrar no sólo tiempo sino también solventes, lo que está relacionado con la reducción del costo por análisis y el impacto ambiental.

En el Capítulo 3 se evaluó el rendimiento de los algoritmos MCR-ALS y U-PCA/RBL en el análisis de datos no lineales obtenidos a través del seguimiento de una reacción cinética utilizando un sistema FIA de flujo interrumpido. Este sistema se utilizó para determinar tartrazina en presencia de azul brillante y amarillo ocaso en distintos tipos de bebidas. En un primer paso se aplicaron ambos algoritmos para remover la contribución de componentes inesperados no incluidos en la calibración. A continuación, los datos obtenidos fueron modelados con una función polinomial. MCR-ALS fue el único algoritmo que permitió una determinación exacta de tartrazina en las muestras de validación por lo que fue aplicado para la determinación del colorante en muestras reales. Los resultados obtenidos al analizar las bebidas comerciales fueron validados por comparación con los arrojados por un procedimiento cromatográfico clásico. Se pudo establecer que MCR-ALS puede manejar este tipo de datos y generar resultados comparables a los obtenidos con el método separativo. Además, resulta interesante destacar que la tartrazina fue encontrada en las muestras a niveles de concentración cinco veces menores al máximo valor establecido por la legislación Argentina.

En el Capítulo 4 se demostró que se puede disminuir el trabajo experimental para la determinación de un analito en presencia de compuestos inesperados y efecto de filtro interno, sin tener que realizar todo el trabajo experimental necesario para aplicar el método de adición estándar. La metodología propuesta consiste en el modelado de matrices de excitación-emisión con PARAFAC combinado con transferencia de calibración para la cuantificación de felinefrina en muestras de agua en presencia de ibuprofeno, aspirina y paracetamol (que actúa como filtro interno en el rango de trabajo). Además, esta estrategia permitió mantener la sensibilidad de la curva de calibrado en ausencia de efecto de filtro interno evitando la preparación de un gran número de soluciones y manteniendo cifras de mérito aceptables. Los resultados fueron comparados con los obtenidos al aplicar el método de adición estándar convencional combinado con PARAFAC.

Esta metodología puede extenderse a otros sistemas que presenten efecto matriz y, consecuentemente, puede ser una herramienta útil para determinar la

cantidad de fármacos presentes en ambientes acuáticos y para evaluar el efecto de las plantas de tratamiento de efluentes convencionales en la eliminación de compuestos farmacéuticos.

En el Capítulo 5 se presenta el desarrollo de un método para la determinación de doxorubicina en muestras de plasma en presencia de riboflavina utilizando matrices de espectroscopia de fluorescencia sincrónica total. La característica más importante de este trabajo es la obtención de la ventaja de segundo orden por primera vez para datos no bilineales. Esto se logró de dos maneras diferentes: a) aplicando MCR-ALS a datos de primer orden (las matrices se desdoblaron previamente) y b) U-PLS/RBL, que ha demostrado ser una herramienta útil cuando se trabaja con datos que no son trilineales. Con respecto a MCR-ALS se demostró la obtención de la ventaja de segundo orden cuando se modelan datos de primer orden y se aplica de la restricción de correlación. El rendimiento de ambas metodologías, juzgado por la exactitud y precisión de las predicciones de las muestras de validación, demostró ser muy satisfactorio.

Uno de los problemas más desafiantes en cuanto a la estructura de los datos de segundo orden es la pérdida de bilinealidad. En el Capítulo 6 se presenta un nuevo algoritmo que modela este tipo de datos denominado cuadrados mínimos parciales con modelado residual de datos no bilineales (U-PLS/RMNB) donde los perfiles correspondientes a las interferencias son extraídos por modelado de los datos desdoblados con MCR-ALS e incorporados en el paso de minimización que modifica los coeficientes de la regresión PLS. En este capítulo se analizaron datos simulados y experimentales; en el segundo caso, se determinó la concentración de doxorubicina en muestras de plasma, utilizando el mismo sistema que en el Capítulo 4. Al mismo tiempo, se compararon los resultados obtenidos con los arrojados por MCR-ALS, U-PLS/RBL y PARAFAC para mostrar cómo la pérdida de bilinealidad afecta el modelado y la habilidad de predicción de este tipo de algoritmos.

Los métodos desarrollados en este trabajo de tesis demostraron que es posible aplicar técnicas quimiométricas como alternativa a las metodologías oficiales, reduciendo notablemente el tiempo de análisis y el uso de grandes cantidades de solventes, y a su vez, obteniendo resultados comparables.

One of the challenges of modern analytical chemistry is the determination of different types of substances such as food additives, pharmaceuticals, metabolites or agrochemicals in complex samples and at low concentrations. Furthermore, the search for new methods is focused on their reliability and safety to determine these compounds with greater selectivity, sensitivity and shorter analysis time. In the last years the environmental conservation has become increasingly important, so the new developments seek to minimize the consumption of solvents and toxic reagents.

In the present thesis, new analytical techniques were developed with the combination of different kind of signals and suitable chemometrics tools which allowed relevant information from each sample data to be extracted. Moreover, they enabled the quantitation of several analytes even in the presence of uncalibrated interferents. Specifically, two way data was analysed and the second order advantage was achieved in order to solve several analytical issues, among which we can mention: non trilinear data, inner filter effect and the loss of bilinearity.

In the next paragraphs, the methods described in the chapters of this thesis are summarized.

Chapter 1 is an introduction in which different types of analytical data are described, as well as the chemometric methods and algorithms used throughout this thesis.

In Chapter 2, the development of a fast chromatographic method with DAD detection is presented for the determination of three synthetic dyes: tartrazine, amaranth and sunset yellow, in non alcoholic beverages samples purchased in a local supermarket. The procedure consists in using a mobile phase that allows the fast elution of the sample components. The gathered second order data presented time shift between different chromatograms and cannot be conveniently corrected to determine the above-mentioned dyes in beverage samples. This fact originates the lack of trilinearity that cannot be conveniently pre-processed and can hardly be modelled by using U-PLS/RBL algorithm. On the contrary, MCR-ALS has shown to be an excellent tool for modelling this kind of data allowing to reach acceptable results. In contrast with the classic chromatographic procedure, analysis time could be reduced up to 14.2% of the necessary time to perform the complete separation

allowing saving both solvents and time, which are related to a reduction of both the costs per analysis and environmental impact.

In Chapter 3, the performance of MCR-ALS and U-PCA/RBL was evaluated in the modeling of non-linear kinetic-spectrophotometric data acquired by a stopped-flow system for the quantitation of tartrazine in the presence of brilliant blue and sunset yellow FCF as possible interferents. In the present work, MCR-ALS and U-PCA/RBL were firstly applied to remove the contribution of unexpected components not included in the calibration set. Secondly, a polynomial function was used to model the non-linear data obtained by the implementation of the algorithms. MCR-ALS was the only strategy that allowed the determination of tartrazine in test samples accurately. Therefore, it was applied for the analysis of tartrazine in beverage samples with minimum sample preparation and short analysis time. The proposed method was validated by comparison with a chromatographic procedure published in the literature. Furthermore, the levels of tartrazine which were found in the analysed samples were five times lower than its maximum value established by the Argentinian legislation.

In Chapter 4, a methodology based on second-order data (excitation emission matrices) modeling with one of most popular algorithms presenting the second-order advantage, parallel factor analysis (PARAFAC), combined with transference of calibration is proposed to predict the phenylephrine concentration when significant inner filter effects occur, even in the presence of unexpected sample components. The quantitation of phenylephrine in water samples in the presence of ibuprofen, acetyl salicylic acid and paracetamol (which produce inner filter effect across the useful wavelength range) was achieved. The strategy allows reducing the experimental work and increasing the analytical sensitivity, avoiding the preparation of a large number of solutions and maintaining acceptable figures of merit. Results were compared with those obtained after the application of the classical standard addition method combined with PARAFAC. The new methodology can be extended to other systems presenting matrix effect and, consequently, can become in a useful tool to know the amount of pharmaceuticals in the aquatic environment and to evaluate the effect of conventional wastewater treatment plants in the elimination of pharmaceutical compounds.

Abstract

In Chapter 5, we present the development of a method for the determination of doxorubicin in plasma samples in the presence of an unexpected component (riboflavin) by using total synchronous fluorescence spectroscopic data matrices. To the best of our knowledge, this is the first time that the second-order advantage is obtained with this kind of data. Two strategies including unfolding the data and: (a) processing with MCR-ALS as first-order data or (b) processing with U-PLS/RBL procedure were considered. Regarding MCR-ALS application, it was demonstrated that the second order advantage was achieved with first order data and with the used of the correlation constraint. The performance of both methods, judged by the accuracy and precision of the predictions of the validation samples, proved to be very satisfactory.

The most challenging data structure problem to achieve analyte quantitation from second-order data in the presence of uncalibrated components using multivariate calibration methods is the loss in bilinearity. In Chapter 6, a new residual modeling for non-bilinear data is presented, namely Unfolded Partial Least Squares with Residual Modeling of Non Bilinear data (U-PLS/RMNB) in which the profiles corresponding to interfering species are extracted by MCR-ALS modeling of unfolded data and incorporated into a minimization step which modifies the PLS regression coefficient. In this chapter, simulated and experimental data were analysed. In the second case, the concentration of doxorubicin was determined in plasma samples, using the same system from Chapter 4. Both data sets were used to compare the new method to the prediction ability of conventional MCR-ALS, U-PLS/RBL and PARAFAC.

The methods developed in this thesis demonstrated that chemometric techniques can be applied as an alternative tool to official methodologies, reducing significantly the analysis time and the use of large amounts of solvents and obtaining comparable results.



CAPÍTULO 1

Introducción

La Química Analítica es una ciencia metrológica que desarrolla, optimiza y aplica herramientas (materiales, metodológicas y estratégicas) de amplia naturaleza, que se concretan en procesos de medida encaminados a obtener información (bio)química de calidad. Esta información puede ser tanto parcial (presencia/concentración de una especie en una muestra (bio)química) como global sobre materias o sistemas de amplia naturaleza (química, bioquímica, biológica) en el espacio y en el tiempo para resolver problemas científicos, técnicos, económicos y sociales (Valcárcel 1999). Esta disciplina se divide en química analítica cualitativa y química analítica cuantitativa; la primera indica la identidad de una especie química en una muestra, mientras que la segunda determina la cantidad de un dado componente en una sustancia. Por otro lado, los métodos analíticos pueden ser separados en aquellos basados en técnicas clásicas (química húmeda) o instrumentales (Nieman y col. 1998). En el primer caso, se utilizan técnicas de precipitación, extracción y destilación para separar los componentes de una mezcla. El análisis cualitativo se realiza por color, olor o punto de fusión, mientras que el análisis cuantitativo se logra a través de la medida de masa o volumen. En las técnicas instrumentales se usan equipos para medir propiedades físicas de los analitos, como absorción de luz, emisión de fluorescencia o conductividad. En este caso, la separación de los componentes de una muestra se lleva a cabo con técnicas cromatográficas o electroforéticas.

Uno de los objetivos perseguidos por los químicos analíticos es la búsqueda de métodos de análisis más rápidos, selectivos, sensibles y, muchas veces, económicos y amigables con el medio ambiente.

La química analítica tiene aplicaciones en muchos campos de la ciencia, tales como medicina forense, bioanálisis, análisis clínicos, monitorización ambiental y análisis de materiales. A su vez, juega un papel muy importante en la industria farmacéutica, donde, además de asegurar la calidad de los productos, se utiliza para descubrir nuevas drogas. En el caso de las aplicaciones clínicas, es muy importante para entender las interacciones entre una droga y el paciente.

Los métodos de calibración multivariada impactaron significativamente en el desarrollo de métodos analíticos, especialmente en la cuantificación de analitos de interés en muestras complejas, como las mencionadas anteriormente.

1.1. Calibración multivariada

La creación de nuevas herramientas de medición que provean una mejor información química, en especial a través de la implementación de la quimiometría, figura entre las opciones seguidas por la química analítica moderna para la mejora de los métodos.

La quimiometría utiliza métodos de origen matemático, estadístico y otros procedentes del campo de la lógica formal, para resolver problemas en el campo de la química y, en particular, de la química analítica. La quimiometría se puede dividir en tres áreas generales (Massart y col. 1997): 1) optimización de los procedimientos experimentales y mediciones químicas, 2) calibración, validación y significancia de las mediciones analíticas y 3) extracción de la máxima información química de los datos analíticos. El conjunto de métodos quimiométricos, que tienen por objeto la cuantificación de uno o más componentes en mezclas incógnitas de múltiples componentes, se denomina métodos de calibración multivariada (Thomas y Haaland 1988).

La calibración multivariada es una de las mayores aplicaciones de la quimiometría que involucra un tipo de medida para predecir el valor de un parámetro o propiedad. Desde hace décadas, los químicos analíticos han realizado calibraciones para determinar la concentración de analitos en distintos tipos de muestras, como agua, sangre u orina utilizando ensayos químicos. Sin embargo, a partir de los años 70 surgió una nueva escuela de pensamiento basada en métodos multivariados. En lugar de calibrar una única variable (por ejemplo, la altura de un pico) a otra variable única (por ejemplo, concentración), se calibran múltiples variables, como absorbancias a 50 longitudes de onda, a una o más variables (Brereton 2007). La habilidad de utilizar métodos multivariados va de la mano con la creciente disponibilidad de computadoras de laboratorio capaces de adquirir y manejar grandes cantidades de datos de manera eficiente.

1.1.1. Clasificación de datos instrumentales

La calibración multivariada está estrechamente relacionada con los diferentes tipos de instrumentos analíticos, ya que el tipo de calibración que debe realizarse está de acuerdo con el tipo de datos obtenidos. Dentro de este contexto,

podemos clasificar los datos instrumentos del siguiente modo (Booksh y Kowalski 1994, Escandar y col 2007):

Datos de orden cero. Este tipo de datos son generados por instrumentos que producen un único dato por muestra y, por lo tanto, el conjunto de calibración es un vector (Figura 1.1). La terminología “orden cero” se utiliza por analogía con la noción matemática de que un único número es un tensor de orden cero.

Estos datos se analizan mediante calibración univariada. Ejemplos de instrumentos que generan datos de orden cero son los electrodos ión selectivos, los fotómetros que realizan las mediciones de absorbancia a una única longitud de onda, los equipos de absorción atómica, los fotómetros de llama, etc.



Figura 1.1. Datos de orden cero. Se genera un solo dato por muestra, y el conjunto de calibración es un vector que contiene las medidas para las I muestras de calibración. En la curva de calibrado, a cada valor de concentración le corresponde un dato de señal.

Datos de primer orden. Por cada muestra se obtiene un vector de datos (tensor de orden uno), y el conjunto de calibración resulta ser una matriz (Figura 1.2). Cuando los datos de primer orden para un grupo de muestras se juntan en un único arreglo bidimensional, el objeto resultante se conoce como arreglo de dos vías. Con el término vía se indica el número de dimensiones que tiene un arreglo de datos formado por la agrupación de un cierto número de muestras.

Los instrumentos que proveen este tipo de señal son, por ejemplo, los espectrofotómetros, espectrofluorímetros y cromatógrafos, que son capaces de producir múltiples medidas para cada una de las muestras analizadas, tales como espectros de absorción, espectros de emisión de fluorescencia y cromatogramas, respectivamente.

Este tipo de datos se analiza mediante los métodos de calibración multivariada de orden uno.

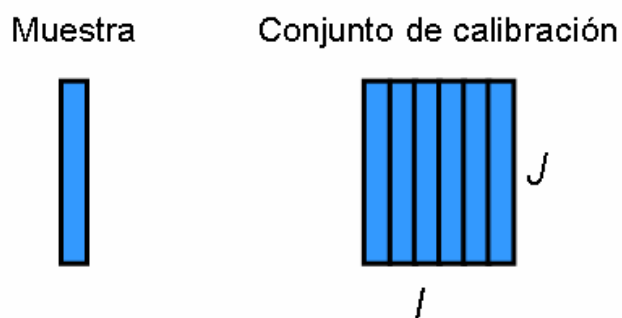


Figura 1.2. Datos de primer orden. Se genera un vector de datos por cada muestra, por ejemplo un espectro medido a J longitudes de onda. El conjunto de I muestras de calibración genera una matriz de $J \times I$.

Datos de segundo orden. En este caso se genera una matriz de datos por cada muestra, y el conjunto de calibración consiste en un arreglo de tres vías (Figura 1.3). Las matrices se pueden producir utilizando un solo instrumento, como, por ejemplo, matrices de excitación-emisión de fluorescencia, espectroscopia UV-vis con detector de arreglo de diodo (DAD), gradiente de pH o monitorización de la cinética de una reacción química (Figura 1.4), cromatografía líquida acoplada a DAD o detector de fluorescencia (FL). Por otro lado, los datos de segundo orden se pueden generar a través de técnicas acopladas o en tándem (*hyphenated techniques*), tales como espectrometría de masas en tándem (MS/MS), cromatografía de gases acoplada a espectrometría de masas (GC/MS), resonancia magnética nuclear bidimensional.

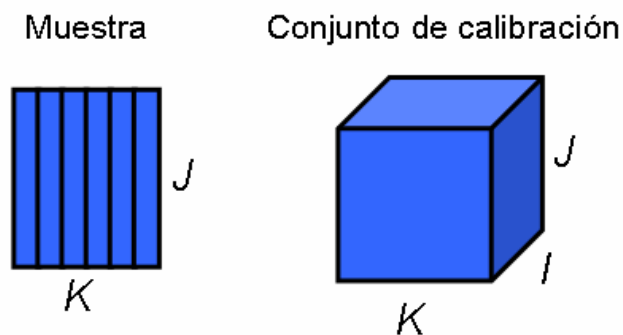


Figura 1.3. Datos de segundo orden. Se genera una matriz de datos de $J \times K$ por cada muestra I , y el conjunto de calibración resulta ser un arreglo de tres vías.

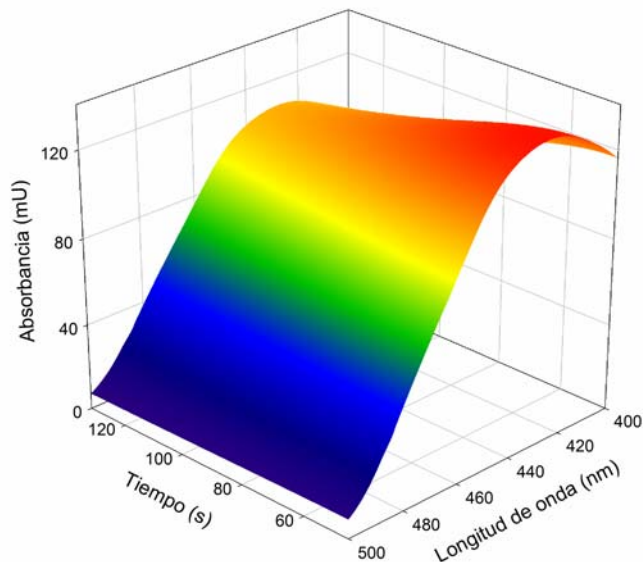


Figure 1.4. Arreglo tridimensional típico en donde se muestra cómo decrece la absorbancia de una muestra de tartrazina en presencia de un oxidante, en función de las longitudes de onda y del tiempo de reacción.

Datos de tercer orden y superiores. Un ejemplo de este tipo de datos es la espectroscopia de fluorescencia de excitación-emisión resuelta en el tiempo al monitorizar una reacción cinética o el decaimiento de la fluorescencia. Este tipo de mediciones provee un tensor de orden tres (Figura 1.5). Para cada muestra se obtiene una h́per matriz de datos, y el juego de calibración es un arreglo de cuatro v́as. Si esta misma muestra adeḿas es registrada a diferentes valores de pH, se obtienen datos de cuatro v́as.

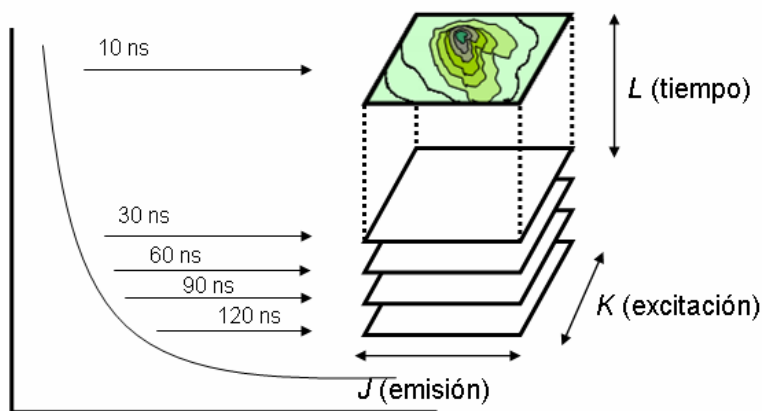


Figura 1.5. Representaci3n esquemática de datos de tercer orden adquiridos para una sola muestra. Mediante la t́cnica de espectroscopia de fluorescencia puede obtenerse una h́per matriz de excitaci3n-emisi3n resuelta en el tiempo.

1.1.2. Algoritmos quimiométricos

La calibración es el proceso matemático y estadístico mediante el que se extra información, por lo general la concentración de un analito, a partir de una señal instrumental. El orden de los métodos de calibración se obtiene del orden de los datos que se recogen por muestra y que se necesitan analizar. De manera general, la obtención de datos de orden n para cada una de las muestras de un juego de calibración genera un tensor de orden $(n + 1)$, que se analiza con un método de calibración o algoritmo de orden n (Booksh y Kowalski 1994). En la evolución desde datos de orden cero a datos de primer y segundo orden, y más, los algoritmos se convierten en herramientas más poderosas ya que aumenta la información que pueden extraer de los datos.

En el caso de datos de orden cero, la relación entre la concentración del analito y la señal se aproxima a modelos teóricos como la ley de Lambert-Beer o a la ecuación de Nernst. Sin embargo, en muchos casos no existen modelos teóricos y, generalmente, se recurre a la regresión lineal. Si en la muestra hay interferentes desconocidos que no están presentes en la etapa de calibración, la estimación de la concentración será errónea, ya que su señal se confundirá o solapará con la señal del analito de interés. Además, en este tipo de calibración es imposible detectar la presencia de un componente inesperado a partir, solamente, de los datos instrumentales, por lo que se requiere de selectividad completa para el analito de interés (Danzer y Currie 1998).

Existe una gran variedad de métodos de calibración multivariada que utilizan respuestas de primer orden (Martens y Naes 1989). Entre los primeros métodos desarrollados se encuentran los de calibración directa: Cuadrados Mínimos Ordinarios (OLS) (Massart y col. 1997), Regresión por Cuadrados Mínimos Clásica (CLS) (Haaland y col. 1999, Galeano y col. 1997, Brown 1991), y el de calibración inversa: Cuadrados Mínimos Inversos (ILS) (Faber 2000, Thomas 1994). El método de regresión por componentes principales (PCR) (Naes y Martens 1998) representa uno de los primeros intentos de reunir las principales ventajas de CLS, OLS e ILS, pero utilizando conceptos de descomposición y factores espectrales. La Regresión por Cuadrados Mínimos Parciales (PLS) mejora los resultados obtenidos por PCR (Martens y Naes 1989). Este último método se convirtió en una herramienta de rutina en el laboratorio.

Por otra parte, cuando las señales instrumentales se comportan de un modo no lineal con respecto a la concentración del analito, se pueden emplear distintos tipos de Redes Neuronales Artificiales (ANNs), como, por ejemplo, el Perceptrón Multicapa (MLP) (Zupan y Gasteiger 1999), las Funciones de Base Radial (RBF) (Derks y col. 1995) o las Máquinas de Vectores Soportes con Ajuste por Cuadrados Mínimos (LS-SVM) (Suykens y col. 2002).

Una limitación de este tipo de algoritmos es que cada especie que presenta señal debe ser incluida en la calibración. Los métodos de calibración de primer orden pueden detectar la presencia de interferentes en las muestras desconocidas. Esta muestra podrá ser identificada como anómala debido a que los residuos espectrales serán anormalmente grandes en comparación con el ruido instrumental pero no podrá ser analizada correctamente o el resultado será erróneo. Esta propiedad se conoce como la ventaja de primer orden (Booksh y Kowalski 1994).

Entre los algoritmos para datos de segundo orden más importantes se encuentran el Análisis Paralelo de Factores (PARAFAC) (Bro 1997), el método de Aniquilación de Rango Generalizado (GRAM) (Sanchez y Kowalski 1986), la Descomposición Trilineal Directa (DTLD) (Sanchez y Kowalski 1990), la Resolución Multivariada de Curvas mediante Cuadrados Mínimos Alternantes (MCR-ALS) (De Juan y col. 2000), los Cuadrados Mínimos Bilineales (BLLS) (Linder y Sundberg 1998, Linder y Sundberg 2002), la Descomposición Trilineal Alternante (ATLD) (Wu y col. 1998) y sus variantes Descomposición Trilineal Alternante Autoponderada (SWATLD) (Chen y col. 2000), la Descomposición Trilineal Alternante Penalizada (APTLD) (Hu y col. 2006, Xia y col. 2005) y el algoritmo de Cuadrados Mínimos Parciales Multivía (N-PLS) (Bro 1996), pero éste último no presenta la ventaja de segundo orden. El modelo PARAFAC2 (Kiers y col. 1999) es una alternativa que se aplica en casos en los que el juego de datos no satisface el modelo PARAFAC; por ejemplo, cuando se presentan corrimientos de picos en datos cromatográficos.

Los métodos de calibración de segundo orden tienen una propiedad especial que permite el análisis de datos aún en presencia de componentes no calibrados, y que se conoce como la ventaja de segundo orden (Booksh y Kowalski 1994). Estos algoritmos pueden predecir la concentración de un analito en

presencia de interferencias. Esta propiedad supera la desventaja de los métodos analíticos de menor orden, es decir, la necesidad de que sean selectivos (en el caso de calibración univariada) y de disponer de muestras de calibración que contengan los potenciales interferentes (en el caso de modelos de primer orden). Además, permite el análisis en simultáneo de varios analitos. Esta última característica también puede lograrse con datos de primer orden, pero se necesita un gran número de muestras y deben tener la misma composición que la muestra incógnita. Por el contrario, los métodos de segundo orden sólo requieren de un número pequeño de estándares de calibración.

Una alternativa para trabajar con datos de segundo orden es reordenar los datos en vectores y aplicar metodologías de primer orden, por ejemplo, Componentes Principales Desdoblados (U-PCR) o Cuadrados Mínimos Parciales Desdoblados (U-PLS) (Wold y col. 1987). Sin embargo, estos métodos desdoblados no proveen la ventaja de segundo orden a menos que se combinen con un procedimiento conocido como Bilinealización Residual (RBL) (Öhman y col. 1990). Linder y Sundberg (Linder y Sundberg 2002) introdujeron RBL como herramienta para obtener esta propiedad. Por otro lado, Olivieri y colaboradores (Olivieri 2005a) presentaron las bases del acoplamiento de N-PLS/RBL para darle a este algoritmo la ventaja de segundo orden.

Cuando las señales registradas presentan una relación no lineal con la concentración de analito se puede obtener la ventaja de segundo orden a partir de información de segundo orden no lineal, mediante la aplicación del algoritmo Componentes Principales Desdoblados seguidos de Bilinealización Residual (U-PCA/RBL). Existen numerosas publicaciones en las que se aplica este método de calibración multivariado en combinación con redes neuronales (García-Reiriz y col. 2007, García-Reiriz y col. 2008, Culzoni y Goicoechea 2007) para resolver sistemas que se desvían de la linealidad.

Una característica importante de los algoritmos de segundo orden es que se pueden extender a datos de órdenes superiores. En este sentido, U-PCR, U-PLS y PARAFAC se pueden usar para analizar datos de N-dimensiones. Para BLS/RBL y U-PLS/RBL se encuentran disponibles extensiones para el tercer orden que dan lugar a TLLS/RTL (Arancibia y col. 2006) y a U-PLS/RTL (Arancibia y col. 2006), respectivamente.

Por otro lado, existe una serie de algoritmos utilizados en transferencia de calibración. Idealmente, las condiciones en que se creó un modelo de calibración deben permanecer estables para situaciones futuras. Sin embargo, en muchos casos existen variaciones químicas y físicas en las muestras y cambios ambientales o instrumentales que pueden afectar estas condiciones. Por lo tanto, la relación entre el modelo viejo y el nuevo no sería óptima. Existen dos soluciones para este problema: realizar nuevamente el modelo de calibración, lo que incrementaría tanto el tiempo de análisis como el costo, o utilizar técnicas quimiométricas capaces de eliminar las diferencias entre ambas situaciones (Feudales y col. 2002). Este tipo de transferencia se denomina estandarización, y entre sus variantes podemos citar: Estandarización Directa (DS) y Estandarización Directa Sucesiva (PDS) (Yang y col. 1991). En el presente trabajo de tesis se aplicó PDS en combinación con PARAFAC para modelar datos de segundo orden en presencia de interferencias y efecto de filtro interno.

1.1.3. Diseños de experimentos en calibración multivariada

Como se mencionó anteriormente en este capítulo, uno de los campos de aplicación de la quimiometría es la optimización de los procedimientos experimentales y las mediciones químicas. Considerando que existe una gran cantidad de factores que influyen en un sistema determinado, es habitual que se requiera de la aplicación de un procedimiento de optimización para lograr un buen desempeño de un dado método. Para lograr este objetivo, se cuenta con varias estrategias posibles, las que pueden dividirse en univariadas y las multivariadas (Massart y col. 1997). En los métodos univariados se estudia la influencia de cada factor por separado variando los niveles de un único factor por vez, mientras que los demás factores permanecen constantes. Por el contrario, en los métodos multivariados se estudian varios factores simultáneamente, variando a la vez los niveles de todos los factores involucrados (Leardi 2009).

La habilidad de realizar predicciones confiables depende, en parte, del diseño y la naturaleza de las muestras de calibración. Generalmente, su conjunto se denomina juego de entrenamiento. Si el juego de calibración no está bien diseñado, el análisis de muestras desconocidas podría ser dificultoso. Si este conjunto de muestras de entrenamiento es pequeño, podrían existir correlaciones

entre las concentraciones de los diferentes compuestos que pueden invalidar modelos de calibración (Brereton 2007). Esto es cierto cuando se preparan mezclas de componentes en la etapa de calibración o, también, cuando se desean preparar muestras de validación que sean representativas.

En la presente tesis se utilizaron distintos diseños de experimentos para elaborar juegos de validación que contienen más de un componente.

1.1.4. Aplicaciones

1.1.4.1. Alimentos

En la actualidad, la oferta de alimentos de todo tipo es creciente. Las técnicas de procesamiento e industrialización aseguran esta mayor diversificación alimentaria brindando al consumidor una amplia gama de posibilidades.

Cuando se trata de la inocuidad de los alimentos, hay que tener en cuenta algunos puntos importantes. En primer lugar, si el producto alimenticio o alguno de sus ingredientes fueron adulterados y, si es así, determinar si el agente adulterante es peligroso o tóxico. En muchas ocasiones se sustituyen materiales por otros de menor costo para obtener mayores beneficios económicos. Desafortunadamente, estas sustituciones no siempre son una opción segura. Por otro lado, las contaminaciones biológicas pueden afectar al producto alimenticio o a un ingrediente en particular, y, generalmente, son el resultado de una contaminación accidental o un almacenamiento inadecuado, que permite que los microorganismos proliferen (Organización Mundial de la Salud, 2007).

A nivel nacional existen numerosas normas como el Código Alimentario Argentino y el Servicio Nacional de Sanidad y Calidad Agroalimentaria (SENASA). El primero regula a todos los alimentos, condimentos, bebidas o sus materias primas y los aditivos alimentarios que se elaboren, fraccionen, conserven, transporten, expendan o expongan, así como a toda persona, firma comercial o establecimiento que lo haga (Código Alimentario Argentino 2010). El SENASA se encarga de ejecutar las políticas nacionales en materia de sanidad y calidad animal y vegetal y de verificar el cumplimiento de la normativa vigente.

A nivel internacional todas estas exigencias están comprendidas en lo establecido por el *Codex Alimentarius (Joint Expert Committee on Food Additives depending on Food and Agriculture Organization (FAO) y World Health*

Organization (WHO)), normas referidas a la producción, elaboración y circulación de alimentos. Su objetivo es asegurar la inocuidad y calidad de los alimentos, proteger la salud del consumidor y promover prácticas equitativas en el comercio internacional. Constituye el patrón de referencia que tienen los países respecto a las exigencias higiénico-sanitarias, bromatológicas y de comercialización de los productos alimentarios.

Es por estos motivos que la supervisión y el análisis de los alimentos se han convertido en tareas indispensables. Para este fin es necesario contar con métodos analíticos confiables, sencillos, rápidos y de bajo costo. En el presente trabajo de tesis se desarrollaron dos métodos para la determinación de colorantes artificiales en bebidas comerciales. Uno de ellos consiste en un método cromatográfico rápido y posterior análisis de los datos con los algoritmos MCR-ALS y U-PLS/RBL. El segundo método consiste en la aplicación de un sistema de inyección en flujo detenido. En este caso, los datos se analizaron con MCR-ALS y U-PCA/RBL.

1.1.4.2. Contaminantes emergentes

Durante los últimos treinta años, el estudio del impacto de la contaminación química estuvo centrado casi exclusivamente en los contaminantes prioritarios convencionales: contaminantes orgánicos persistentes (POPs) y metales pesados. Sin embargo, el uso creciente de fármacos, clasificados dentro de los llamados contaminantes emergentes (CEM), se ha convertido en un nuevo problema ambiental, que despertó gran interés entre los científicos en los últimos años. Estos contaminantes no necesitan ser persistentes para causar efectos negativos, ya que su elevada tasa de transformación y eliminación puede quedar compensada por su continua entrada en el ambiente (Petrovic y col. 2005). Los CEM corresponden, en la mayoría de los casos, a compuestos no regulados como contaminantes, pero que son candidatos para futuras regulaciones debido a sus potenciales efectos sobre la salud y su incidencia en el ecosistema. Como ejemplo de sustancias que surgieron recientemente como contaminantes podemos citar a los antibióticos, analgésicos, anti-inflamatorios, hormonas anticonceptivas, repelentes de insectos, fragancias presentes en los productos de cuidado personal y aditivos de los combustibles (Barceló 2003, Deblonde y col. 2011). Para la mayoría de estas sustancias se desconoce su tasa de incidencia y no existen datos ecotoxicológicos, por lo que

resulta difícil predecir los efectos que pueden ocasionar sobre la salud humana y de los organismos acuáticos (Fatta y col. 2007).

Una de las principales fuentes de entrada al ambiente de estos contaminantes son las aguas residuales urbanas no tratadas y los efluentes de las plantas de tratamiento de aguas residuales. En su mayoría, las plantas de tratamiento no están diseñadas para tratar este tipo de efluentes, y una alta proporción de sustancias contaminantes y sus metabolitos pueden evadir su eliminación y entrar en el ambiente acuático vía efluentes (Petrovic y col. 2005).

Otra cuestión importante a tener en cuenta es la presencia de estos contaminantes en aguas residuales recicladas, ya que a medida que aumenta la presión sobre los recursos hídricos, en muchas partes del mundo se agudizan los problemas de un abastecimiento adecuado de agua. Como consecuencia, se está promoviendo la reutilización de las aguas residuales, principalmente para la irrigación, aunque también para reabastecimiento del agua subterránea (recarga de acuíferos), reciclaje del agua industrial, aumento de los caudales fluviales y otros usos ecológicos, e incluso como agua potable (Hansen 2007).

Debido a su alto consumo, los fármacos y sus metabolitos están incorporándose continuamente a los lodos de las plantas de tratamiento, procedentes de las excreciones, de fármacos caducados o no utilizados y de las descargas industriales (Petrovic y col. 2005). Publicaciones recientes ponen de manifiesto que la tasa de eliminación de algunos fármacos durante el proceso de tratamiento en las plantas depuradoras es baja y, como resultado, estos compuestos están presentes en aguas superficiales, subterráneas y potables (Gros y col. 2006, Deblonde y col. 2011, Clara y col. 2010). Los fármacos pueden generar la misma exposición potencial que los contaminantes persistentes, ya que su elevada tasa de transformación y eliminación puede ser compensada por su continua entrada en el ambiente. Además, las matrices de muestras ambientales como el agua de napa, superficial o de desecho, son muy complejas y, generalmente, contienen otros componentes que pueden actuar como interferentes. Por lo tanto, existe una creciente necesidad de desarrollar métodos analíticos fiables que permitan su determinación en este tipo de muestras (Gros y col. 2006, Peré-Trepat y col. 2007).

La mayoría de los trabajos desarrollados en este campo incluyen técnicas cromatográficas con detección por arreglo de diodos o fluorescencia, o acopladas a MS (Rodil y col. 2009, Santos y col. 2005, Grujic y col. 2009, Raich-Montiu y col. 2007). Estos métodos suelen requerir de técnicas de pretratamiento a la muestra, tener largos tiempos de análisis y utilizar grandes volúmenes de solventes para lograr la separación e identificación de todos los componentes. Sin embargo, cuando se aplican herramientas quimiométricas para analizar múltiples componentes en este tipo de matrices, los tiempos de análisis y el consumo de solventes son menores (Gil García y col. 2009, De Zan y col. 2008, Peré-Trepát y col. 2007, Vosough y col. 2011).

En esta tesis, se desarrolló un método para la determinación de fenilefrina en muestras de agua de canilla y de napa a través del análisis de matrices de excitación-emisión de fluorescencia con una combinación de los algoritmos PARAFAC y PDS.

1.1.4.3. Matrices biológicas

La medición de las concentraciones de drogas terapéuticas en fluidos biológicos, como plasma u orina, y preparados farmacéuticos es un requerimiento en el área de la salud por lo que genera una intensa demanda de nuevas metodologías analíticas. Entre las razones más importantes para analizar fármacos se pueden enumerar las siguientes: a) rango terapéutico estrecho, b) ausencia de parámetros clínicos para seguir sus efectos, c) falta de relación predecible entre dosis y respuesta, d) necesidad de realizar estudios de biodisponibilidad y bioequivalencia en medicamentos genéricos, y e) control de calidad de materias primas y productos comercializados. En la actualidad, la necesidad de realizar estudios de bioequivalencia en medicamentos que han de ser incorporados como genéricos acentuó la necesidad de contar con métodos confiables y aplicables a los laboratorios nacionales. Son varios los fármacos para los que se requieren análisis tanto en las distintas formas farmacéuticas como en materiales biológicos (antibióticos, antiepilépticos, broncodilatadores, anticancerígenos, psicotrópicos, inmunosupresores y cardíacos). En consecuencia, resulta imprescindible generar ensayos confiables, sencillos y de bajo costo para poder analizarlos. Este tema resulta trascendental si se tiene en cuenta el ritmo con el que el mercado

farmacéutico se expande en el mundo, afectando en la misma medida a áreas de la economía y la salud. Esto tiene consecuencias directas sobre nuestro país, en el que la medicina pública juega un rol fundamental en el cuidado de la salud (ANMAT 2002).

Por otro lado, las muestras biológicas suelen mostrar alta viscosidad y/o fuerza iónica, o contener especies interferentes, altas concentraciones de proteínas, etc. Por este motivo, los métodos de calibración multivariada han ganado importancia en la resolución de muestras complejas de múltiples componentes, tales como suero, plasma, orina y tejidos biológicos.

En el presente trabajo de tesis se determinó la droga anticancerígena doxorrubicina en muestras de plasma humano, a través de la obtención de espectros de fluorescencia sincrónicos y posterior análisis con el algoritmo MCR-ALS.

1.2. Herramientas Quimiométricas: algoritmos de calibración multivariados de segundo orden

1.2.1. Resolución Multivariada de Curvas

1.2.1.1. Resolución Multivariada de Curvas: datos de dos vías

La Resolución Multivariada de Curvas (MCR, por sus siglas en inglés) describe a un grupo de técnicas cuyo propósito es obtener los perfiles de respuestas puros (espectros, perfiles de pH, perfiles de tiempo, perfiles de elución, etc.) de constituyentes químicos o especies en una mezcla de componentes no resuelta, obtenidas en procesos químicos, cuando no se dispone de información previa acerca de la naturaleza o composición de estas mezclas (Multivariate Curve Resolution Homepage).

La premisa principal de MCR es la expresión de la ley de Beer para múltiples componentes, tal como se muestra en la siguiente ecuación para la absorbancia de una mezcla de dos especies a una determinada longitud de onda λ :

$$A_{\lambda} = \varepsilon_{\lambda,X}bc_X + \varepsilon_{\lambda,Y}bc_Y \quad (1.1)$$

donde ϵ representa la absorptividad molar de los compuestos X e Y medidos en una cubeta de paso óptico igual a b , y c representa la concentración de cada especie absorbente. Conociendo la absorbancia a una única longitud de onda, A_λ , y el paso óptico no es posible obtener una única solución para los parámetros $\epsilon_{\lambda,X}$, $\epsilon_{\lambda,Y}$, c_X y c_Y . Sin embargo, si medimos una segunda absorbancia a una segunda longitud de onda, y se conocen las absorptividades o las concentraciones de las especies intervinientes, se puede obtener una solución para el resto de las variables resolviendo ambas expresiones simultáneamente. Cabe señalar que cualquier sistema analítico que produce respuestas aditivas linealmente puede ser analizado de esta manera, por ejemplo, espectroscopia de fluorescencia, cromatografía o voltametría.

La serie de ecuaciones a diferentes longitudes de onda se puede representar de manera genérica de la siguiente manera (Rutan y col. 2009):

$$d_j = \sum_{n=1}^N c_n s_{j,n} \quad (1.2)$$

donde d_j representa la absorbancia (o la respuesta instrumental) a la longitud de onda j th, c_n es la concentración de la especie n th en la mezcla y $s_{j,n}$ es la absorptividad molar (u otro factor de sensibilidad instrumental apropiado) a la longitud de onda j th de la especie n th en la mezcla. Si se escribe en formato vectorial tendremos:

$$\mathbf{d} = \mathbf{c} \mathbf{S}^T \quad (1.3)$$

donde \mathbf{d} es un vector de tamaño $1 \times J$ que contiene las mediciones a diferentes J longitudes de onda, \mathbf{c} es un vector de tamaño $1 \times N$ que contiene las concentraciones de los N componentes individuales que contribuyen a la respuesta de la mezcla y \mathbf{S} es una matriz de tamaño $J \times N$ con las absorptividades molares o valores de sensibilidad para cada especie a cada longitud de onda (el superíndice T indica la transpuesta de la matriz).

Para aplicar MCR se requiere un mínimo conjunto de datos de segundo orden, por ejemplo, un juego de espectros de I mezclas medido a J longitudes de onda, donde el modelo lineal para la respuesta está dado por:

$$\mathbf{D} = \mathbf{C} \mathbf{S}^T + \mathbf{E} \quad (1.4)$$

donde la matriz \mathbf{D} ($I \times J$) contiene las absorbancias individuales o las respuestas instrumentales de las I mezclas medidas a J longitudes de onda, la matriz \mathbf{C} ($I \times N$) contiene las concentraciones de cada N componente presente en las I mezclas y \mathbf{E} ($I \times J$) representa la contribución del error de las mediciones. La estructura de este modelo bilineal se representa en la Figura 1.6:

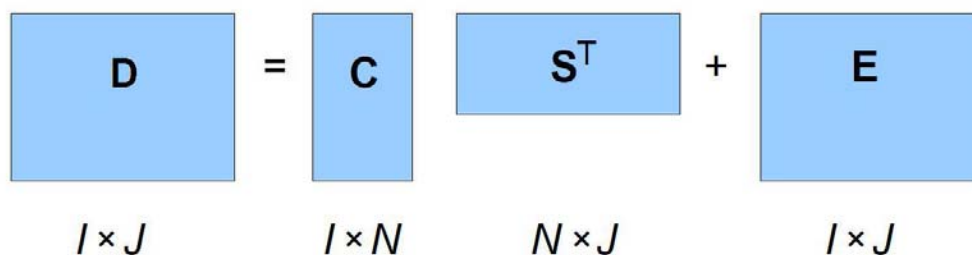


Figura 1.6. Representación del modelo MCR para una serie de I mezclas que contienen hasta N componentes medidos a J longitudes de onda.

Dada una matriz \mathbf{D} con respuestas instrumentales, el método MCR busca obtener tanto las sensibilidades (es decir los espectros) (\mathbf{S}) como las concentraciones (\mathbf{C}). El primer paso de MCR es llevar a cabo una descomposición de la matriz \mathbf{D} para obtener matrices análogas a \mathbf{C} y \mathbf{S} de la Ecuación (1.4), que comúnmente se llaman *scores* y *loadings* de la matriz \mathbf{D} , respectivamente.

En el caso de MCR-Cuadrados Mínimos Alternantes (ALS) (Tauler 1995, Tauler y col. 1995) esta búsqueda se basa en un método iterativo, que representa los datos directamente en términos de estimaciones de las matrices \mathbf{C} y \mathbf{S} , denominados $\hat{\mathbf{C}}$ y $\hat{\mathbf{S}}$, respectivamente. La implementación de esta técnica se basa en el cálculo iterativo de \mathbf{S} y \mathbf{C} como se muestra a continuación:

$$\hat{\mathbf{S}}^T = \mathbf{C}^+ \mathbf{D} = (\hat{\mathbf{C}}^T \hat{\mathbf{C}})^{-1} \hat{\mathbf{C}}^T \mathbf{D} \quad (1.5)$$

$$\hat{\mathbf{C}} = \mathbf{D} \mathbf{S}^{T+} = \mathbf{D} \hat{\mathbf{S}} (\hat{\mathbf{S}}^T \hat{\mathbf{S}})^{-1} \quad (1.6)$$

donde ‘^’ indica una matriz estimada y ‘+’ representa la pseudoinversa de la matriz. Es importante mencionar que MCR-ALS requiere ser inicializado con parámetros del sistema lo más cercanos posible al resultado final. Se pueden proveer, por ejemplo, espectros de las especies obtenidos a partir de sus estándares puros o del análisis de los llamados espectros ‘más puros’ a través de la metodología SIMPLISMA (del inglés, *Simple Interactive Self-modelling Mixture Analysis*), un algoritmo de resolución multivariada de curvas que extrae los espectros puros de los componentes a partir de una mezcla de espectros de composición variada (Windig y Guilment 1991), que se detallará en la Sección 1.3.1. Otra opción es aplicar EFA (del inglés, *Evolving Factor Analysis*) (Maeder 1987, Maeder y Zilian 1988, Maeder y Zuberbuehler 1986) que proporciona estimaciones iniciales de los perfiles de concentración para conjuntos de datos que evolucionan secuencialmente.

Al comenzar las iteraciones, se introducen una serie de restricciones apropiadas, que se describirán más adelante, y que se pueden aplicar tanto a los perfiles en \mathbf{C} como a los espectros contenidos en \mathbf{S} , de manera conjunta o independiente.

El procedimiento general de funcionamiento de MCR-ALS incluye los siguientes pasos (de Juan y col. 2009):

1. Determinación del número de componentes en \mathbf{D} .
2. Generación de las estimaciones iniciales.
3. Cálculo de \mathbf{S}^T bajo restricciones.
4. Cálculo de \mathbf{C} bajo restricciones.
5. Generación de \mathbf{D} a partir del producto de \mathbf{C} y \mathbf{S}^T .
6. Repetición del paso 3 hasta alcanzar la convergencia.

El número de componentes en \mathbf{D} se puede determinar a través del Análisis de Componentes Principales (PCA), por Descomposición de Valores Singulares (SVD) o puede ser conocido de antemano. Las estimaciones iniciales deben tener

un significado químico y se pueden obtener con las metodologías nombradas anteriormente.

El objetivo de MCR-ALS consiste en encontrar dos mínimos de la norma de E , minimizando las dos siguientes ecuaciones bajo restricciones adecuadas:

$$\min_{\hat{C}} \|\mathbf{D} - \hat{C}\hat{S}^T\| \quad (1.7)$$

$$\min_{\hat{S}^T} \|\mathbf{D} - \hat{C}\hat{S}^T\| \quad (1.8)$$

y así, de esta manera, obtener los perfiles que mejor describan el sistema.

El criterio de convergencia en la optimización mediante ALS se basa en la comparación del ajuste obtenido en dos iteraciones consecutivas. Cuando la diferencia relativa es menor a un valor umbral, la optimización finaliza. También se recomienda monitorizar la evolución de la forma de los perfiles para asegurarse de que se obtiene una solución óptima.

1.2.1.2. Resolución Multivariada de Curvas Extendida: datos de tres vías

El algoritmo MCR-ALS se ha convertido en una herramienta quimiométrica ampliamente utilizada para la resolución de múltiples componentes en muestras desconocidas debido a la variedad de tipos de datos que se pueden analizar; esencialmente, cualquier sistema con múltiples componentes que den como resultado tablas o matrices de datos (datos de segundo orden) que puedan describirse por un modelo bilineal (Jaumot y col. 2005).

La deficiencia de rango ocurre cuando el número de componentes significativos detectados es menor al número de contribuciones químicas presentes verdaderamente en el sistema. Un ejemplo típico es cuando existe dependencia lineal entre los perfiles de concentración o espectrales de componentes químicos diferentes en una mezcla, o cuando los perfiles presentan formas idénticas en un caso más trivial. MCR-ALS tiene la capacidad de analizar múltiples matrices de datos simultáneamente, reduciendo ambigüedades intrínsecas del análisis de factores y/o deficiencias de rango. Por otra parte, permite la aplicación flexible y diversa de restricciones que ayudan a la resolución,

mejorando los resultados (Jaumot y col. 2005). En la Figura 1.7 se observan dos situaciones encontradas en la práctica, cuando se analizan varias matrices.

Un juego de datos se considera trilineal cuando los perfiles resueltos del mismo componente en diferentes matrices de datos y en una dirección particular (K o J) tienen la misma forma (Jaumot y col. 2005). MCR-ALS es capaz de analizar juegos de datos que se desvían de la trilinealidad ya que, en vez de formar un arreglo de tres vías, éste es desdoblado a lo largo del modo sospechado de romper la estructura trilineal. Por ejemplo, si la variación ocurre a lo largo de la dirección de las columnas, se crea una matriz aumentada en ese sentido, acomodando la matriz correspondiente a la muestra incógnita y las matrices correspondientes a las muestras de calibración una debajo de la otra (De Juan y col. 2000).

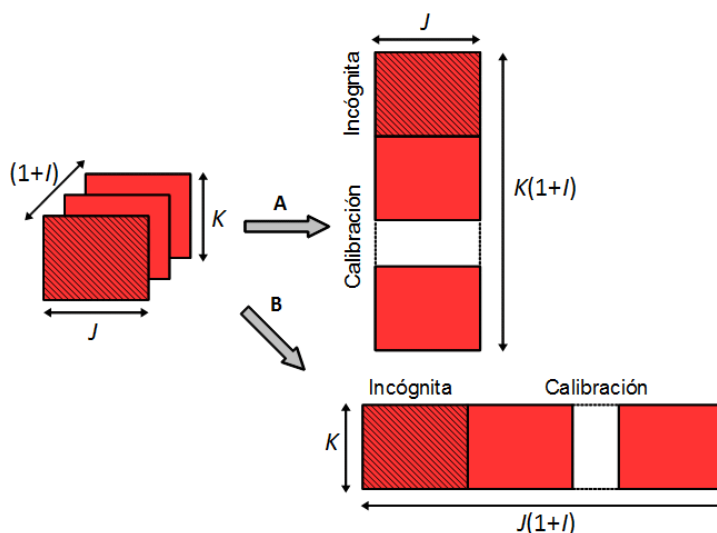


Figura 1.7. Representación del desdoblamiento de una hiper matriz de tres vías en otra de dos vías. Apilamiento de matrices individuales de calibración y una muestra incógnita para ser analizadas simultáneamente. **A:** aumento en el sentido de las columnas. **B:** aumento en el sentido de las filas.

La descomposición bilineal de la matriz aumentada **D** se lleva a cabo según la expresión:

$$\mathbf{D} = \mathbf{CS}^T + \mathbf{E} \quad (1.9)$$

donde las filas de **D** contienen los espectros de absorción medidos en función del tiempo (u otro tipo de señal de segundo orden), las columnas de **C** contienen los

perfiles de tiempo de los componentes involucrados en el proceso, las columnas de \mathbf{S} sus espectros, y \mathbf{E} es la matriz de los residuos no ajustados por el modelo. Las dimensiones apropiadas de \mathbf{D} , \mathbf{C} , \mathbf{S} y \mathbf{E} son $K(1+I) \times J$, $K(1+I) \times N$, $J \times N$ y $J(1+I) \times K$, respectivamente (I es el número de muestras de calibración, J es el número de longitudes de onda, N es el número de componentes que dan respuesta, K el número de tiempos de elución y se asume una sola muestra incógnita). La descomposición de \mathbf{D} se logra a través de la minimización por ALS de la norma de \mathbf{E} , como se describió en la Sección 1.2.1.1 para una sola matriz de datos. La minimización comienza con la provisión de los espectros estimados de los componentes, los que se emplean para estimar $\hat{\mathbf{C}}$ a partir de la Ecuación (1.10):

$$\hat{\mathbf{C}} = \mathbf{D}(\mathbf{S}^T)^+ \quad (1.10)$$

donde '+' indica la inversa generalizada, como ya se mencionó. Con la matriz $\hat{\mathbf{C}}$ de la Ecuación (1.10) y la matriz de datos original \mathbf{D} se vuelve a estimar la matriz \mathbf{S} mediante cuadrados mínimos:

$$\hat{\mathbf{S}}^T = (\hat{\mathbf{C}})^+ \mathbf{D} \quad (1.11)$$

y, finalmente, se calcula \mathbf{E} a partir de la Ecuación (1.9) usando \mathbf{D} y las matrices estimadas $\hat{\mathbf{C}}$ y $\hat{\mathbf{S}}$. Estos pasos se repiten hasta alcanzar la convergencia imponiendo ciertas restricciones adecuadas al sistema en estudio durante el proceso de ALS, como ya se comentó anteriormente. Entre las restricciones más aplicadas se encuentran (Muñoz y De Juan 2007, De Juan y Tauler 2003, Jaumot y col. 2005):

- a) No negatividad: previene la presencia de valores negativos en los perfiles. Se aplica a los perfiles de concentración y a espectros UV-Vis y de fluorescencia, debido a que sus valores son siempre positivos.
- b) Unimodalidad: fuerza la presencia de un único máximo en el perfil de una especie. Se aplica a procesos en donde los perfiles de concentración muestran una forma de aparición-decaimiento (perfiles de reacción, picos cromatográficos y electroforéticos, voltamogramas, etc.).

-
- c) Clausura (cierre o balance de masa): se relaciona con ecuaciones de balance de masa en sistemas cerrados. Se aplica cuando la concentración total de las especies detectables en cada etapa del proceso es constante, pudiéndose fijar a un valor único o variable. Cuando hay especies no absorbentes en el sistema no se puede aplicar, debido a que las especies que contribuyen a la señal medida no forman un sistema cerrado.
- d) Rango local/selectividad: impone la ausencia de algunas especies en un rango específico de perfiles de concentración, por ejemplo, en una cierta ventana de tiempo o pH. Este conocimiento puede tener una base química o ser adquirido mediante la aplicación de métodos de análisis de rango local.
- e) Igualdad: se puede introducir en la optimización información química asociada al conocimiento de los perfiles puros de concentración o espectros. Los perfiles conocidos pueden ser establecidos como invariantes durante el proceso iterativo. No es necesario conocer completamente un perfil; si hay algunos elementos conocidos se pueden fijar.
- f) Modelado rígido: se aplica cuando se conoce la ley química (cinética o termodinámica) asociada a un proceso particular. Los perfiles de concentración asociados a componentes en el proceso se fuerzan a presentar la forma descrita por la ley química.
- g) Trilinealidad: es la restricción de tres vías más esencial, ya que, cuando es aplicable, asegura que los resultados son únicos; en otras palabras, que los perfiles resueltos no tienen ambigüedad. Esta restricción se aplica cuando se analiza una matriz aumentada, o sea, con estructura trilineal.
- h) Correspondencia entre especies: se aplica en el caso del análisis de varias matrices de manera simultánea. Fija la presencia o ausencia de un componente en una matriz particular \mathbf{C} y/o \mathbf{S}^T . La información presencia/ausencia se aplica como código binario, donde 1 indica presencia y 0 ausencia. Como consecuencia, cuando un compuesto en particular no existe en una matriz determinada, los elementos en el perfil correspondiente se fijan en cero.
- i) Correlación: consiste en forzar las concentraciones del analito en las muestras de calibración, c_{cal} , en cada iteración ALS a valores de referencia previamente conocidos del analito, c_{ref} , en estas muestras. Esta restricción se explicará con más detalle en el Capítulo 5.
-

Luego de la descomposición de \mathbf{D} , la información contenida en \mathbf{C} puede ser utilizada para realizar predicciones cuantitativas. Esto se realiza construyendo un gráfico de calibración univariada con las concentraciones relativas de las muestras de calibración contenidas en la matriz optimizada \mathbf{C} . Dado un componente de interés particular, se seleccionan los elementos de \mathbf{C} correspondientes a las I mezclas de calibración desde $C_{1,n}$ hasta $C_{IK,n}$. En nuestro caso, cada serie de valores K corresponde al perfil en la dimensión temporal (tiempos de elución) para una muestra en particular. Si el área bajo el perfil es proporcional a la concentración del componente se construye un modelo según:

$$\left\{ \sum_{k=1}^K C_{kn} \mid \sum_{k=1}^K C_{(k+K)n} \mid \dots \mid \sum_{k=1}^K C_{[k+K(I-1)]n} \right\} = k_{MCR} \mathbf{y} \quad (1.12)$$

donde la sumatoria de cada grupo de K elementos de \mathbf{C} corresponden a un componente en particular en una muestra de calibración dada, e \mathbf{y} es el vector de las concentraciones de calibración.

Mediante la interpolación de los valores provistos por la muestra desconocida puede obtenerse la concentración del analito de interés y_u , es decir:

$$y_u = \frac{\sum_{k=1}^K C_{(k+IK)n}}{k_{MCR}} \quad (1.13)$$

1.2.1.3. Cifras de mérito en MCR-ALS

El proceso más importante para comparar métodos analíticos es la determinación de cifras de mérito. La estimación de estos parámetros en calibración multivariada ha sido objeto de varias publicaciones recientes (Olivieri y col. 2006, Olivieri y Faber 2012, Bauza y col. 2012, Allegrini y Olivieri 2012). La sensibilidad (SEN), definida como el cambio en la respuesta neta para un cambio dado en la concentración de analito, puede ser considerada una de las cifras de mérito más relevantes en el campo de la química analítica, ya que es un factor decisivo en la estimación de otras cifras, tales como límite de detección (LOD),

límite de cuantificación (LOQ), incertidumbre en la predicción de concentraciones (SD), etc.

De acuerdo a Bauza y colaboradores (Bauza y col. 2012), la expresión correcta para estimar la SEN cuando se aplica MCR-ALS es:

$$SEN_{MCR} = m_n [J(\mathbf{S}^T \mathbf{S})_{nn}^{-1}]^{-1/2} \quad (1.14)$$

donde n es el índice para el analito de interés en una mezcla de múltiples componentes, m_n es la pendiente de la gráfica de calibración univariada MCR para este analito, \mathbf{S}^T es la matriz que contiene los perfiles de todos los componentes de la muestra en la dirección de MCR no aumentada, y J es el número de canales en la matriz de la muestra de validación en la dirección de aumento de MCR.

Utilizando el valor de SEN_{MCR} se pueden estimar otras cifras de mérito como LOD, LOQ y la incertidumbre en la concentración estimada, mediante las ecuaciones descritas por Bauza y colaboradores (Bauza y col. 2012).

1.2.2. U-PLS/RBL

El algoritmo Cuadrados Mínimos Parciales Desdoblados contiene una etapa de calibración en donde se emplea información acerca de la concentración del analito sin incluir datos de la muestra desconocida. Por lo tanto, en primer lugar, se utilizan las I muestras de calibración $\mathbf{X}_{c,i}$ (de tamaño $J \times K$, donde J y K son los números de canales en cada dimensión) y las concentraciones de calibración \mathbf{y} ($N_c \times 1$, en donde N_c es el número de muestras de calibración). Las matrices $\mathbf{X}_{c,i}$ se vectorizan y se agrupan dentro de una matriz \mathbf{Z}_X :

$$\mathbf{Z}_X = [\text{vec}(\mathbf{X}_{c,1}) \mid \text{vec}(\mathbf{X}_{c,2}) \mid \dots \mid \text{vec}(\mathbf{X}_{c,I})] \quad (1.15)$$

donde 'vec' denota la operación de vectorización que convierte las matrices de $J \times K$ en vectores de $JK \times 1$. Con esta matriz \mathbf{Z}_X ($JK \times 1$) se calibra un modelo U-PLS convencional que provee dos clases de variables latentes: los *loadings* contenidos en la matriz \mathbf{P} y los pesos de las cargas (*weight loadings*) contenidos en \mathbf{W} , ambos de tamaño $JK \times A$, en donde A es el número de factores latentes; además, se

obtienen los coeficientes de regresión \mathbf{v} ($A \times 1$). Las columnas de \mathbf{W} son ortogonales, mientras que las de \mathbf{P} no necesariamente lo son. Es importante recalcar que las columnas de \mathbf{W} no son autovectores propiamente dichos, sino factores que dependen de las concentraciones de calibración del analito de interés y se llaman variables latentes. La obtención de estas variables se lleva a cabo mediante un algoritmo iterativo cíclico, que puede resumirse en los siguientes pasos:

1. Proyección de la matriz de datos en el vector de concentraciones \mathbf{y}_n para el analito n en las muestras de calibración, obteniéndose el primer *weight loading factor*:

$$\mathbf{w}_1 = \mathbf{Z}_X / \mathbf{y}_n (\mathbf{y}_n^T \mathbf{y}_n) \quad (1.16)$$

En este paso del algoritmo se supone que sólo se conocen las concentraciones de un único componente en las mezclas de calibración, en este caso el analito 1. En otras palabras, \mathbf{w}_1 es una aproximación por cuadrados mínimos al espectro puro del analito 1. En este paso se aprecia la introducción de información concerniente a las concentraciones contenidas en \mathbf{y}_1 en el cálculo del primer factor.

2. Se normaliza el factor \mathbf{w}_1 , dividiéndolo por $\mathbf{w}_1^T \mathbf{w}_1 = \|\mathbf{w}_1\|^2$ para contar finalmente con un juego de factores ortonormales.

3. Obtención del primer *score*:

$$\mathbf{t}_1 = \mathbf{Z}_X^T \mathbf{w}_1 \quad (1.17)$$

Se continúa con la suposición de que únicamente está presente el analito 1, y se calcula qué contribución del primer factor \mathbf{w}_1 está presente en las mezclas de calibración. Estas “concentraciones” forman el vector \mathbf{t}_1 .

4. Obtención del primer coeficiente de regresión v_1 :

$$v_1 = \mathbf{t}_1 \mathbf{y}_n / (\mathbf{y}_n^T \mathbf{y}_n) \quad (1.18)$$

Se calcula el coeficiente de regresión que relaciona el *score* \mathbf{t}_1 calculado en el paso 3 con las concentraciones de calibración.

5. Obtención del primer *loading* \mathbf{p}_1 :

$$\mathbf{p}_1 = \mathbf{Z}_X \mathbf{t}_1 / (\mathbf{t}_1^T \mathbf{t}_1) \quad (1.19)$$

6. Cálculo de los residuos espectrales y de concentración:

$$\begin{aligned} \mathbf{E}_Z^T &= \mathbf{Z}_X^T - \mathbf{t}_1 \mathbf{p}_1^T \\ \mathbf{e}_c &= \mathbf{y}_n - v_1 \mathbf{t}_1 \end{aligned} \quad (1.20)$$

En los pasos 5 y 6 se asegura que los vectores \mathbf{t}_a y \mathbf{w}_a subsiguientes serán ortogonales entre sí. Para ello se calculan los vectores \mathbf{p}_a , llamados *loadings*. Estos vectores no explican la varianza espectral en la matriz \mathbf{Z}_X , sino que representan un intento de explicar dicha varianza, mientras simultáneamente se correlacionan los *scores* \mathbf{t}_a con las concentraciones \mathbf{y}_1 .

7. Se substituyen \mathbf{E}_Z y \mathbf{e}_c por \mathbf{Z}_X y \mathbf{y}_n respectivamente en el paso 1 y se continúa hasta llegar al número de factores deseado (A).

1.2.2.1. Número óptimo de factores (A)

El parámetro A se selecciona usualmente mediante técnicas de validación cruzada (Haaland y Thomas 1988) o utilizando el test propuesto por Faber y colaboradores (Wiklund y col. 2007).

El método de validación cruzada consiste en calibrar con todas las muestras de calibración menos una. Con este nuevo juego de calibración se predice la concentración de la muestra que se dejó de lado; luego se repite dejando de lado otra de las muestras de calibración y así sucesivamente hasta que todas son dejadas de lado. Con las concentraciones predichas y las concentraciones nominales se calcula el error a través de un estadístico llamado *PRESS* (del inglés *prediction error sum of squares*):

$$PRESS = \sum_1^I (c_{\text{nom}} - c_{\text{pred}})^2 \quad (1.21)$$

donde c_{nom} es la concentración nominal del componente en las mezclas utilizadas para la calibración y c_{pred} es la concentración predicha por el modelo.

Normalmente el *PRESS* va disminuyendo hasta alcanzar el valor óptimo, que generalmente coincide con el número de fuentes de variación de las mezclas, y luego comienza a aumentar. Esto ocurre porque los factores comienzan a sumar ruido espectral. El número óptimo de factores correspondería al que tiene menor *PRESS*, aunque no siempre es el más indicado para la predicción. Al seleccionar A con este criterio, pueden ocurrir errores por sobre ajuste. La mejor manera de seleccionar el número de factores es a través del estadístico F , propuesto por Haaland y Thomas (Haaland y Thomas 1988):

$$F_i = \frac{PRESS_i}{PRESS_{\text{menor}}} \quad (1.22)$$

donde $PRESS_i$ es cada uno de los errores calculados y $PRESS_{\text{menor}}$ es el menor error obtenido, para $i = 1, 2, \dots, I/2 + 1$.

El número óptimo de factores corresponderá al F_i tal que $F_i < F_{\alpha, I, I}$, donde $(1-\alpha)$ es el percentil de Senedecor para la distribución F , e I los grados de libertad.

Asumiendo que los errores predichos tienen distribución normal, media igual a cero, que son independientes, y que no existe diferencia significativa entre sus varianzas, la Probabilidad $\{F_i > F_{\alpha, I, I} \mid \sigma_i^2 = \sigma_{\text{menor}}^2\} = 2\alpha$, donde σ_i^2 y σ_{menor}^2 son las varianzas de los errores predichos. En consecuencia, se acepta el primer F_i , cuya probabilidad sea menor que 0.75.

1.2.2.2. Predicción basada en el score de la muestra

Si la calibración resulta conveniente y no hay interferencias no esperadas, \mathbf{v} se puede emplear para estimar las concentraciones de analito en una muestra incógnita, y_u :

$$y_u = \mathbf{t}_u^T \mathbf{v} \quad (1.23)$$

donde \mathbf{t}_u es el *score* de la muestra incógnita, obtenido por proyección de los datos vectorizados para la muestra incógnita \mathbf{X}_u dentro del espacio de los A factores latentes:

$$\mathbf{t}_u = (\mathbf{W}^T \mathbf{P})^{-1} \mathbf{W}^T \text{vec}(\mathbf{X}_u) \quad (1.24)$$

Cuando existen componentes no esperados en \mathbf{X}_u , el *score* de la muestra dado por la Ecuación (1.24) no es adecuado para la predicción del analito utilizando la Ecuación (1.23). En este caso, los residuos del paso de predicción de U-PLS (s_p), representados por la Ecuación (1.25), serán anormalmente grandes en comparación con el ruido instrumental típico:

$$\begin{aligned} s_p &= \|\mathbf{e}_p\| / (JK - A)^{1/2} \\ &= \|\text{vec}(\mathbf{X}_u) - \mathbf{P}(\mathbf{W}^T \mathbf{P})^{-1} \mathbf{W}^T \text{vec}(\mathbf{X}_u)\| / (JK - A)^{1/2} \\ &= \|\text{vec}(\mathbf{X}_u) - \mathbf{P}\mathbf{t}_u\| / (JK - A)^{1/2} \end{aligned} \quad (1.25)$$

donde $\|\cdot\|$ indica la norma Euclidea. En este caso, la muestra incógnita es considerada una muestra anómala u *outlier*. La posibilidad de descartar muestras que presentan interferencias no modeladas y que por ende están sujetas a un nivel de error alto es la denominada “ventaja de primer orden”.

Si la muestra incógnita contiene componentes inesperados, la situación puede ser solucionada aplicando un procedimiento aparte denominado bilinealización residual que se basa en la descomposición de valores singulares (Wold y col. 1987, Olivieri 2005a).

Dada una muestra de validación que contiene componentes no calibrados, su señal puede ser descompuesta en dos partes: una que puede ser modelada usando las variables latentes de la calibración (\mathbf{X}_{mod}) y otra que no puede ser modelada con estas variables ($\mathbf{X}_{\text{nomod}}$):

$$\mathbf{X}_u = \mathbf{X}_{\text{mod}} + \mathbf{X}_{\text{nomod}} \quad (1.26)$$

\mathbf{X}_{mod} se puede expresar en función de las variables latentes de la calibración \mathbf{P} y de los *loadings* de la muestra asociada con \mathbf{t}_u , de la siguiente manera:

$$\text{vec}(\mathbf{X}_u) = \mathbf{P}\mathbf{t}_u + \mathbf{e}_{\text{mod}} + \text{vec}(\mathbf{X}_{\text{nomod}}) \quad (1.27)$$

donde \mathbf{e}_{mod} es el vector de los residuos no ajustados en \mathbf{X}_{mod} mediante el modelo PLS de A factores y tiene elementos del orden del ruido instrumental. Si en $\mathbf{X}_{\text{nomod}}$ existe algo que tenga una estructura bilineal con valores mayores que el ruido experimental, puede ser modelado por SVD, que permite la estimación de los perfiles de las interferencias \mathbf{b}_{int} y \mathbf{c}_{int} mediante la minimización de la norma del vector de los residuos \mathbf{e}_u (calculado mientras se ajustaron los datos de la muestra a la suma de las contribuciones relevantes):

$$\text{vec}(\mathbf{X}_u) = \mathbf{P}\mathbf{t}_{\text{uRBL}} + \text{vec}[\mathbf{g}_{\text{int}} \mathbf{b}_{\text{int}} (\mathbf{c}_{\text{int}})^T] + \mathbf{e}_u \quad (1.28)$$

donde \mathbf{b}_{int} y \mathbf{c}_{int} son los eigenvectores (autovectores) izquierdo y derecho de \mathbf{E}_p , y \mathbf{g}_{int} es un factor de escala:

$$(\mathbf{g}_{\text{int}}, \mathbf{b}_{\text{int}}, \mathbf{c}_{\text{int}}) = \text{SVD}_1(\mathbf{E}_p) \quad (1.29)$$

donde \mathbf{E}_p es la matriz de $J \times K$ obtenida luego de reordenar el vector \mathbf{e}_p de $JK \times 1$ de la Ecuación (1.25), y SVD_1 indica que se tuvo en cuenta el primer componente principal.

Durante el proceso RBL, \mathbf{P} se mantiene constante en los valores de calibración y \mathbf{t}_{uRBL} se varía hasta que $\|\mathbf{e}_u\|$ se minimiza. La minimización se puede llevar a cabo utilizando el procedimiento Gauss-Newton (GN) comenzando desde la Ecuación (1.23). Una vez minimizado $\|\mathbf{e}_u\|$, la concentración del analito se calcula usando la Ecuación (1.24), mediante la introducción del vector final encontrado por el procedimiento RBL, \mathbf{t}_{uRBL} .

El número de interferencias N_{int} se puede evaluar comparando los residuos finales s_u con el nivel de ruido instrumental:

$$s_u = \|\mathbf{e}_u\| / [JK - (N_c + N_{\text{int}})]^{1/2} \quad (1.30)$$

donde \mathbf{e}_u se obtuvo de la Ecuación (1.28) y N_c es el número de analitos calibrados. Generalmente, el gráfico de s_u en función del número de componentes mostrará valores decrecientes, comenzando en s_p cuando el número de componentes es igual a A (número de variables latentes utilizado en la etapa de calibración), hasta que se estabiliza en un valor comparable al ruido instrumental, permitiendo localizar el número correcto de componentes. Debe tenerse en cuenta que para $N_{\text{int}} > 1$, la forma de los perfiles obtenidos para los componentes inesperados está menos correlacionada con el espectro del componente verdadero, debido al hecho de que los componentes principales están restringidos a ser ortonormales (Bortolato y col. 2007).

Por otro lado, la adición de más variables latentes que el número de compuestos químicos al aplicar RBL, podría, de alguna manera, compensar la falta de trilinealidad en los datos (Cañada-Cañada y col. 2009).

1.2.2.3. Cifras de mérito en U-PLS/RBL

Allegrini y Olivieri (Allegrini y Olivieri 2012) desarrollaron una nueva expresión que permite la estimación de la sensibilidad para los métodos de calibración multivariada basados en la regresión de cuadrados mínimos combinada con RBL:

$$\text{SEN}_{JAC} = \{\mathbf{v}^T [\mathbf{P}^T (\mathbf{I} - \mathbf{Z}_{\text{int}} \mathbf{Z}_{\text{int}}^+) \mathbf{P}]^{-1} \mathbf{v}\}^{-1/2} \quad (1.31)$$

donde el subíndice "JAC" representa el método Jacobiano, \mathbf{P} es la matriz de *loadings* de calibración, \mathbf{v} es el vector de coeficientes de regresión de PLS en el espacio de las variables latentes, \mathbf{I} es una matriz unitaria y \mathbf{Z}_{int} contiene información acerca de las especies interferentes.

En la presente tesis se utilizó la expresión recién mencionada para estimar SEN, LOD, LOQ y otras cifras de mérito, según lo descrito por Allegrini y Olivieri (Allegrini y Olivieri 2012).

1.2.3. U-PCA/RBL

El análisis de componentes principales de datos desdoblados no provee información acerca de la concentración de un analito, sino que, en principio, arroja un número de *scores* de entrenamiento y de validación. Con frecuencia los *scores* son utilizados para el entrenamiento de redes neuronales y posterior predicción de muestras incógnitas cuando se analizan datos no lineales. Sin embargo, en la presente tesis demostramos que los *scores* de calibración también pueden ser utilizados para crear un modelo univariado no lineal, útil para realizar predicciones usando los *scores* de las muestras de validación.

PCA es una herramienta muy utilizada en el análisis de datos en varios campos, desde la neurociencia hasta ciencias computacionales, debido a que es un método simple y no paramétrico que extrae información relevante de datos complejos (Shlens 2009). PCA es un procedimiento matemático que utiliza la transformación ortogonal para convertir un conjunto de observaciones de variables posiblemente correlacionadas en un conjunto de valores de variables linealmente no correlacionadas llamadas componentes principales. El número de componentes principales es igual o menor al número de variables originales. El primer componente tendrá la mayor variancia posible, y cada componente siguiente tendrá la mayor variancia posible con la restricción de que debe ser ortogonal (es decir, no correlacionado) respecto de los componentes precedentes. Los componentes principales serán independientes si el conjunto de datos se distribuyen normalmente en forma conjunta.

Como resultado, PCA arroja los *scores* de un componente (los valores de la variable transformada correspondiente a un punto dado en particular), y los *loadings* (el peso por el cual cada variable original estandarizada debe multiplicarse para obtener el *score* del componente) (Shaw 2003).

En el caso de este algoritmo, el número correcto de componentes se obtiene analizando una serie de parámetros: el porcentaje de variancia explicada, el desvío estándar de los residuos de la regresión (S_{fit}), el S_{fi} calculado por validación cruzada ($S_{fit,cv}$) (ver Sección 1.2.2.1) y el radio entre el $PRESS(A)$ (ver Ecuación (1.21)) y el error estándar de la estimación $SEE(A-1)$. En la Figura 1.8 se ilustra a modo de ejemplo una gráfica del % de variancia explicada en función del número

de componentes, y se deduce que el número óptimo es 2, ya que, en ese valor, la ordenada es mayor a 99%.

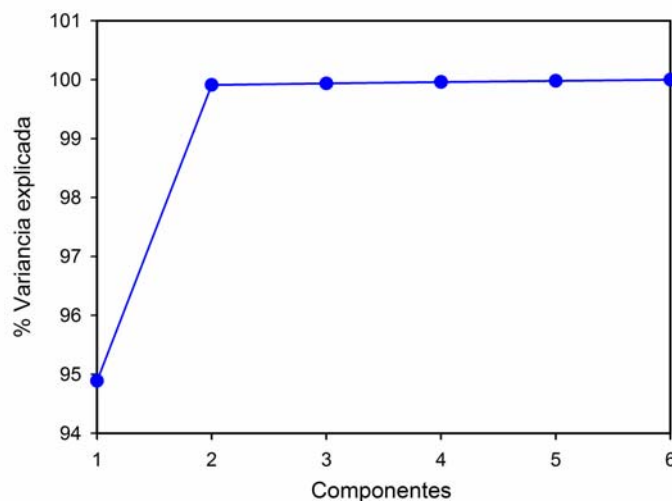


Figura 1.8. Porcentaje de la variancia explicada en función del número de componentes.

En el caso de S_{fit} , se selecciona el número de componentes en donde este parámetro puede ser comparado con el nivel del ruido instrumental. La Figura 1.9, también a modo de ejemplo, muestra cómo el $S_{fit_{CV}}$ se estabiliza en el valor 2 del eje de las abscisas. El valor del radio $PRESS(A) / SEE(A-1)$ debe ser menor a 1.

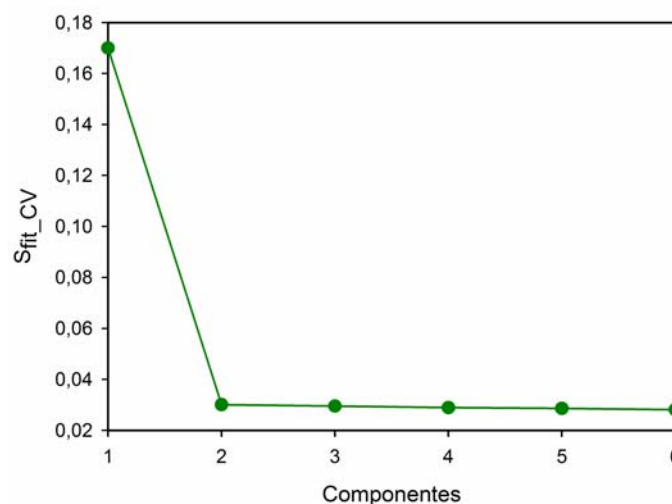


Figura 1.9. Desvío estándar de los residuos de la regresión calculado por validación cruzada ($S_{fit_{CV}}$) en función del número de componentes.

Cuando aparecen componentes no esperados en las muestras de validación, sus *scores* no son adecuados para realizar la predicción. En este caso es necesario

recurrir a una técnica que marque a la nueva muestra como anómala, indicando que se necesitan acciones adicionales antes de realizar la predicción, y que luego aisle la contribución de los componentes no esperados de la de los analitos calibrados, con el objetivo de recalculer *scores* apropiados para las muestras de validación. En U-PCA la muestra será considerada como anómala si los residuos del análisis de los datos de validación desdoblados \mathbf{X}_u [s_p , ver Ecuación (1.25)] son anormalmente grandes en comparación con el ruido instrumental típico. En este caso, \mathbf{P} es la matriz que contiene los primeros A *loadings* obtenidos de la aplicación de PCA a los datos de entrenamiento desdoblados. El resto de los parámetros tienen el mismo significado que los detallados en la Sección 1.2.2. RBL se puede emplear para modelar la presencia de componentes no esperados en las muestras de validación mediante la descomposición de sus señales en dos partes: una modelada usando las variables latentes de calibración (\mathbf{X}_{mod}) y la restante que no puede ser modelada con estas variables ($\mathbf{X}_{\text{nomod}}$). El tratamiento matemático es similar al descrito para U-PLS/RBL en la Sección 1.2.2.2, y comprende las ecuaciones (1.26) a (1.30).

En esta tesis, los *scores* de calibración se emplearon para construir una curva de calibrado univariada no lineal. Luego, los *scores* corregidos por RBL se utilizaron para predecir la concentración del analito en las muestras.

1.2.4. Análisis Paralelo de Factores (PARAFAC)

PARAFAC es uno de los varios métodos de descomposición de datos multivía que descompone los arreglos de datos en juegos de *loadings* y puntuaciones (*scores*) para describirlos de una forma más condensada que la original (Bro 1997). El modelo se propuso independientemente por Harshman (Harshman 1970) y por Carrol y Chang (Carroll y Chang 1970), quienes lo nombraron CANDECOMP (siglas en inglés de *Canonical Decomposition*).

La descomposición de los datos se lleva a cabo de manera similar al análisis en componentes principales, con la diferencia de que cada componente consiste en un vector de *scores* y dos vectores de *loadings*.

Si para una dada muestra se obtienen datos experimentales en forma de matrices de $J \times K$, estamos en presencia de un arreglo de dos vías, en donde J y K denotan el número de datos en la primera y la segunda dimensión,

respectivamente. Si las I matrices de calibración y la matriz de la muestra incógnita se apilan, se genera un arreglo de tres vías $\underline{\mathbf{X}}$, con dimensiones $[(I + 1) \times J \times K]$. De esta manera se agrega al sistema información sobre las posibles interferencias presentes en la muestra, aunque sin necesidad de identificarlas o conocer su concentración. Si $\underline{\mathbf{X}}$ sigue un comportamiento trilineal, el modelo PARAFAC para un arreglo de tres vías está dado por tres matrices *loadings* \mathbf{A} , \mathbf{B} y \mathbf{C} , que contienen elementos a_{in} , b_{jn} y c_{kn} (Bro 1997). El modelo trilineal busca minimizar la suma de los cuadrados de los residuos, e_{ijk} en el modelo:

$$x_{ijk} = \sum_{n=1}^N a_{in} b_{jn} c_{kn} + e_{ijk} \quad i = 1, 2, \dots, I; \quad j = 1, 2, \dots, J; \quad k = 1, 2, \dots, K \quad (1.32)$$

La forma de la matriz del modelo trilineal resulta:

$$\mathbf{X}_{..k} = \mathbf{A} \text{diag}(\mathbf{c}_{(k)}) \mathbf{B}^T + \mathbf{E}_{..k} \quad k = 1, 2, \dots, K \quad (1.33)$$

El modelo también puede describirse según:

$$\underline{\mathbf{X}} = \sum_{n=1}^N \mathbf{a}_n \otimes \mathbf{b}_n \otimes \mathbf{c}_n + \underline{\mathbf{E}} \quad (1.34)$$

en donde \otimes indica el producto de Kronecker, N es el número total de componentes y $\underline{\mathbf{E}}$ es una hiper matriz que contiene los residuos no ajustados por el modelo.

En el campo de la química, las columnas de las matrices *loadings* \mathbf{B} y \mathbf{C} pueden ser asignadas con algún significado físico, como, por ejemplo, espectros de excitación y emisión en espectroscopia de fluorescencia, y perfiles cromatográficos y espectros UV-vis en cromatografía líquida de alto rendimiento (CLAR) con DAD. Las columnas de la matriz de *loadings* \mathbf{A} representan las concentraciones relativas de los componentes en una mezcla (puntuaciones o *scores*).

Una ventaja de PARAFAC es la unicidad de la solución. En los modelos bilineales existe un problema conocido como ambigüedad rotacional. Los *loadings* en un modelo bilineal reflejan los espectros puros de los analitos, pero esto sólo es posible si se incorpora información externa (Bro 1997). Si los datos son trilineales,

se encontrará el espectro verdadero (o lo que constituyan las variables) si se utiliza el número correcto de factores y si la relación señal/ruido es la apropiada (Harshman 1972, Kruskal 1976, Kruskal 1977). El significado matemático de la unicidad es que el modelo estimado por PARAFAC no puede ser rotado sin la pérdida de ajuste, contrario a lo que sucede en el análisis de dos vías, en donde se obtienen distintos *scores* y *loadings* sin cambiar el ajuste del modelo. Por lo tanto, los modelos estimados y verdaderos deben coincidir cuando se selecciona el número correcto de componentes.

En la calibración de segundo orden, en un primer paso PARAFAC trata a las matrices **A**, **B** y **C** como desconocidas. Luego, éstas se estiman mediante el algoritmo de cuadrados mínimos alternantes luego de inicialización al azar. Esto es posible aún si los componentes son químicamente desconocidos, constituyendo la base de la ventaja de segundo orden.

Las concentraciones de los componentes en una muestra incógnita se predicen luego de realizarse la regresión de los *scores* en **A** estimados por PARAFAC, en función de las concentraciones conocidas de las muestras de calibrado, según:

$$(a_{n1} | a_{ni} | \dots | a_{nl}) = k_{\text{PARAFAC}} \mathbf{y} \quad (1.35)$$

donde n corresponde a un determinado componente, \mathbf{y} es un vector que contiene las concentraciones nominales del analito n en las I mezclas de calibración, y k_{PARAFAC} es la constante de proporcionalidad calculada.

Como el modelo usado para la predicción es un modelo de regresión lineal simple, se predice la concentración del analito de interés en la muestra (y_u), solo con el *score* de la muestra ($a_{n(I+1)}$).

$$y_u = a_{n(I+1)} / k_{\text{PARAFAC}} \quad (1.36)$$

Cuando se combina PARAFAC con el método de adición estándar, los datos de segundo orden de la muestra y de las cinco adiciones por triplicado se analizan simultáneamente. La estimación de la concentración del analito en la muestra se puede obtener graficando los *scores* en función de la cantidad de estándar

agregado a la muestra, ajustando los datos y encontrando la intersección con el eje de las abscisas, como se realiza en adición estándar univariada.

1.2.4.1. Cifras de mérito en PARAFAC

Según Olivieri y colaboradores (Olivieri y col. 2012), se puede obtener una expresión general para calcular SEN en cualquier situación de calibración usando señales de segundo orden bilineales según:

$$SEN_n = sn\{[\mathbf{B}_{cal}^T \mathbf{P}_{b,int} \mathbf{B}_{cal}] * (\mathbf{C}_{cal}^T \mathbf{P}_{c,int} \mathbf{C}_{cal})\}^{-1/2}_{nn} \quad (1.37)$$

donde ‘*’ indica el producto de matrices Hadamard, ‘nn’ implica el elemento (n,n) de una matriz dada y sn es la señal total integrada del componente n a una concentración unidad. \mathbf{B}_{cal} y \mathbf{C}_{cal} son los componentes esperados que están presentes en las muestras de calibración, y $\mathbf{P}_{b,int}$ y $\mathbf{P}_{c,int}$ se pueden obtener con la siguiente expresión:

$$\mathbf{P}_{b,int} = \mathbf{I} - \mathbf{B}_{int} \quad (1.38)$$

$$\mathbf{P}_{c,int} = \mathbf{I} - \mathbf{C}_{int} \quad (1.39)$$

donde \mathbf{B}_{int} y \mathbf{C}_{int} contienen los perfiles del componente inesperado (no calibrado) como columnas, e \mathbf{I} es una matriz identidad. Cabe destacar que las matrices \mathbf{B}_{int} y \mathbf{C}_{int} pueden construirse con columnas que representan el verdadero perfil espectral del interferente.

Por otra parte, se puede calcular la sensibilidad analítica (γ_n) según:

$$\gamma_n = SEN_n / \|\delta r\| \quad (1.40)$$

donde $\|\delta r\|$ es la medida del ruido instrumental, que puede ser estimado a través del análisis estadístico de réplicas de una muestra blanco. Este parámetro, como ya se comentó anteriormente, permite la comparación analítica de métodos sin tener en cuenta la técnica específica utilizada, el equipo ni la escala empleada. Además, su inversa (γ_n^{-1}) establece la mínima diferencia de concentración que el método

puede discriminar estadísticamente a lo largo del rango dinámico, considerando como única fuente de error al ruido instrumental.

1.3. Otras Herramientas Quimiométricas

1.3.1. SIMPLISMA

Los métodos iterativos como MCR-ALS requieren de un valor estimado inicial de las matrices \mathbf{C} o \mathbf{S}^T para comenzar el proceso de optimización. Estas aproximaciones de perfiles de concentración o espectrales se irán modificando durante los ciclos iterativos hasta obtenerse una solución final.

Las estimaciones iniciales no aleatorias pueden ser desarrolladas directamente a partir del conjunto de datos originales o pueden ser el resultado de la aplicación auxiliar de métodos quimiométricos. Por lo tanto, la selección de un juego de columnas o filas de los datos en \mathbf{D} , que sea igual en número a los componentes presentes en el sistema, puede dar lugar a estimaciones iniciales de las matrices \mathbf{C} y \mathbf{S}^T , respectivamente. En algunos casos, como por ejemplo datos de cromatografía líquida de alto rendimiento con detección de arreglo de diodos, se pueden utilizar los espectros observados en los picos de los cromatogramas como estimaciones iniciales para \mathbf{S}^T . Por lo tanto, la inclusión de espectros puros conocidos o perfiles de concentración es también posible, cuando están disponibles.

Cuando el conocimiento *a priori* de los datos es escaso o cuando se prefieren procedimientos matemáticos para obtener las estimaciones iniciales se pueden utilizar métodos quimiométricos. SIMPLISMA procede seleccionando de manera secuencial las variables en la dirección de las filas o las columnas que tienen menos información en común que las seleccionadas previamente (Windig y Guilment 1991, Grande y Manne 2000). Si las variables más puras se seleccionan en la dirección de las concentraciones, las estimaciones resultantes son del tipo \mathbf{S}^T , y si las variables seleccionadas son espectros, se generarán estimaciones del tipo \mathbf{C} . Cuando se usan este tipo de métodos, en donde se seleccionan las variables más puras, se obtienen mejores estimaciones cuando la búsqueda se realiza en la dirección en donde exista información con el menor solapamiento; por ejemplo, en

el caso de datos cromatográficos, la mejor opción sería utilizar los espectros estimados relacionados a los tiempos de elución más puros seleccionados.

SIMPLISMA asume que existen variables en los datos (como longitudes de onda o números de onda) que experimentan una contribución de sólo uno de los componentes de la mezcla. Dichas variables se llaman variables puras, y la intensidad espectral en la variable pura representa el perfil de concentración. Si cada mezcla de componentes tiene al menos una variable pura en el juego de datos, el espectro correspondiente puede ser resuelto por medio de una regresión.

Este procedimiento utiliza la función de pureza:

$$p_{i1} = \frac{\sigma_i}{\mu_i + \alpha} \quad (1.41)$$

donde μ_i es la media de la variable i :

$$\mu_i = \frac{1}{c} \sum_{j=1}^c d_{ij} \quad (1.42)$$

y σ_i es la desviación estándar para la misma variable:

$$\sigma_i = \sqrt{\frac{1}{c} \sum_{j=1}^c (d_{ij} - \mu_i)^2} \quad (1.43)$$

donde d_{ij} representa los elementos de la matriz de datos con los espectros en las filas y c es el número de espectros. El segundo subíndice de p en la Ecuación (1.41) indica que es el primer valor de pureza. Windig descubrió que la relación entre la media y la desviación estándar para la misma variable se correlaciona con la pureza de la variable (Windig y Guilment 1991). De acuerdo a la expresión de pureza (Ecuación (1.41)), cuando el valor de la media tiende a cero, la relación entre éste y la desviación estándar tiende a infinito. Por esta razón se adiciona una constante al denominador, α , que se define como un porcentaje de la media, y se suprime la sobreestimación de la pureza.

SIMPLISMA resuelve los componentes uno por uno. La variable pura siguiente es aquella que es más independiente de la primera variable pura. La ecuación general para la pureza de un espectro es:

$$p_{ij} = \left(\frac{\sigma_i}{\mu_i + \alpha} \right) w_{ij} \quad (1.44)$$

donde w_{ij} es un determinante espectral que se utiliza como una función peso.

Los espectros obtenidos son reflejo de la variancia residual en los datos no modelados por los componentes anteriores.

1.3.2. Estandarización Directa Sucesiva (PDS)

PDS relaciona la respuesta \mathbf{r} de una muestra obtenida en una situación A (\mathbf{X}_A) con la respuesta obtenida en una situación B (\mathbf{X}_B). La relación está descrita por la matriz de transformación \mathbf{F} , según:

$$\mathbf{X}_A = \mathbf{X}_B \mathbf{F} \quad (1.45)$$

En el caso estudiado, la situación A corresponde a las señales de excitación-emisión de los estándares preparados en agua MilliQ y la situación B a las señales de excitación-emisión con efecto de filtro interno. PDS construye un modelo multivariado entre la respuesta \mathbf{r} de una muestra medida a una longitud de onda j th en la situación A y diferentes respuestas obtenidas en la situación B. Estas respuestas secundarias están en una ventana que está centrada en la longitud de onda j th. En PDS, se llama ventana a la cantidad de variables secundarias tenidas en cuenta en cada uno de esos cálculos locales, y tolerancia al valor umbral que permite determinar el número óptimo de factores para cada modelo local. La relación se puede describir según:

$$\mathbf{r}_j = \mathbf{R}_j \mathbf{b}_j \quad (1.46)$$

donde \mathbf{R}_j es la matriz de respuesta localizada de las muestras de transferencia y \mathbf{b}_j es el vector que contiene los coeficientes de transformación para el tensor j th. Los

vectores de regresión calculados para cada ventana se ensamblan de tal manera de formar una matriz diagonal \mathbf{F} , según:

$$\mathbf{F} = \text{diag} (\mathbf{b}_1^T, \mathbf{b}_2^T, \dots, \mathbf{b}_j^T, \dots, \mathbf{b}_k^T) \quad (1.47)$$

donde k es el número de sensores (ver Figura 1.10).

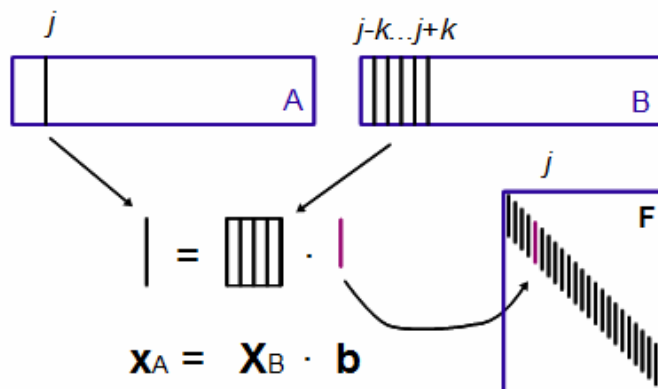


Figura 1.10. Representación esquemática de un paso de PDS. Se realiza una regresión entre los valores a la longitud de onda j th en la situación A con los valores en la ventana ($j-k..j+k$) de la situación B. El vector de regresión resultante forma la columna j th de la diagonal de la matriz de transformación \mathbf{F} .

La respuesta de la muestra incógnita puede ser luego estandarizada según:

$$\mathbf{x}_S^T = \mathbf{x}^T \mathbf{F} \quad (1.48)$$

donde \mathbf{x}_S^T representa la muestra como si hubiese sido medida bajo la situación A (ausencia de efecto de filtro interno).

La metodología descrita es aplicable a señales de primer orden. En consecuencia, las matrices se desdoblaron en vectores y se sometieron a PDS, para luego aplicar PARAFAC a las matrices reconstruidas. Así, los *scores* obtenidos fueron útiles para predecir la concentración del analito en la muestra.

1.4. Regiones Elípticas de Confianza Conjunta (EJCR)

Con el fin de comparar dos metodologías analíticas es posible aplicar un modelo de regresión lineal para evaluar si los resultados obtenidos por ambos métodos son comparables. En general, se recurre a un método de regresión que

tenga en cuenta tanto los errores en el eje de las ordenadas como en de las abscisas, ya que ambos van a contener las concentraciones predichas para un analito determinado por dos métodos analíticos diferentes. El método muy utilizado en estos casos es el de Cuadrados Mínimos Bivariados (BLS por sus siglas en inglés) (Riu y Rius 1996). Esta técnica de regresión busca minimizar la suma de los residuos ponderados (S), expresado en la siguiente ecuación:

$$S = \sum_{i=1}^n \frac{(y_i - \hat{y}_i)^2}{w_i} \quad (1.49)$$

donde \hat{y}_i es el valor estimado de y_i , y el factor w_i tiene en cuenta las variancias de cada par de datos (x_i, y_i) , $s_{x_i}^2$ y $s_{y_i}^2$, según la Ecuación (1.50):

$$w_i = s_{y_i}^2 + \hat{b}^2 s_{x_i}^2 \quad (1.50)$$

donde \hat{b} es la pendiente estimada de la línea de regresión. La estimación del error experimental se define como:

$$\hat{S}^2 = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2 / w_i}{n - 2} \quad (1.51)$$

donde n es el número de niveles de concentración. Por lo tanto, el método le asigna mayor peso (es decir, menos importancia) a aquellos pares de datos con altos valores de $s_{x_i}^2$ y/o $s_{y_i}^2$, es decir, a los pares de datos más imprecisos.

Para comparar, entonces, dos métodos analíticos, se realiza una regresión con los resultados obtenidos para un juego de muestras por ambos métodos. Si ninguno de los coeficientes de la regresión lineal difiere estadísticamente de la unidad (en el caso de la pendiente) o de cero (en el caso de la ordenada al origen), los resultados no se considerarán diferentes a un determinado nivel de confianza α .

La prueba conjunta del intervalo de confianza para la pendiente y la ordenada al origen que considera los errores en ambos ejes se utiliza para analizar si hay diferencias significativas entre los coeficientes de regresión y los valores teóricos uno y cero, respectivamente. Esto consiste en comprobar la presencia de los puntos teóricos dentro de los límites de la región elíptica de confianza conjunta definida por:

$$\sum_{i=1}^n \frac{1}{w_i} (\hat{a} - a)^2 + 2 \sum_{i=1}^n \frac{x_i}{w_i} (\hat{a} - a)(\hat{b} - b) + \sum_{i=1}^n \frac{x_i^2}{w_i} (\hat{b} - b)^2 = 2\hat{s}^2 F_{1-\alpha}(2, n-2) \quad (1.52)$$

donde $F_{1-\alpha}(2, n-2)$ es el valor tabulado de F con 2 y $n-2$ grados de libertad a un nivel de confianza determinado. Sólo cuando los puntos teóricos caen dentro de la región elíptica delimitada por la Ecuación (1.52) se puede concluir que no hay diferencias significativas entre ambos métodos.

El tamaño de la elipse depende directamente de la estimación del error experimental, \hat{s}^2 . Por lo tanto, cuando se dispone de pocos datos experimentales este valor puede sobreestimarse aumentando, entonces, la incertidumbre a causa de la falta de información inherente a un bajo número de pares de datos y, a veces, a la falta de ajuste de los datos experimentales a la línea de regresión BLS. En estos casos, la región elíptica es de gran tamaño y puede incluir al punto (1,0) sólo por azar. Por lo tanto, el tamaño de la elipse da una idea de la precisión del método que se está probando.

El método EJCR se utiliza también para comparar valores predichos de concentración obtenidos por una técnica analítica con los valores nominales de concentración en muestras, por ejemplo, de validación.

1.5. Objetivos

1.5.1. Objetivos generales

Contribuir al desarrollo de la química analítica en general y a la quimiometría en particular, a través de la combinación de datos de orden superior con métodos quimiométricos que permitan la resolución de mezclas complejas, aún con instrumental de mediana complejidad y en un corto tiempo de análisis.

1.5.2. Objetivos específicos

1) Desarrollar aplicaciones prácticas de los algoritmos quimiométricos de calibración multivariada, utilizando en especial señales cromatográficas y espectroscópicas de orden mayor que uno.

2) Desarrollar un nuevo algoritmo quimiométrico que permita modelar datos de segundo orden no bilineales.

3) Analizar y validar la potencial implementación de las metodologías desarrolladas en los siguientes campos:

- a) alimentos, específicamente en el análisis de colorantes en bebidas comerciales,
- b) ambiental, específicamente en el análisis de contaminantes emergentes, y
- c) bioanalítico, específicamente en el análisis de fármacos en muestras biológicas.



CAPÍTULO 2

Desarrollo de un método cromatográfico rápido para la determinación de colorantes en bebidas no alcohólicas, utilizando datos obtenidos por cromatografía líquida de alta resolución con detector de arreglo de diodos y algoritmos quimiométricos de segundo orden

2.1. Introducción

El uso de algoritmos quimiométricos de segundo orden juega un rol crítico en muchos campos analíticos, y puede ser evidenciado mediante una revisión de los artículos publicados en revistas analíticas, quimiométricas y aplicadas de relevancia en los últimos años (Escandar y col. 2007, Gómez y Callao 2008, Escandar y col. 2013).

Una de las formas de obtener datos de segundo orden es mediante el acoplamiento de instrumentos de primer orden, como cromatografía líquida de alto rendimiento, con detector de arreglo de diodos, fluorescencia de barrido rápido o masas; cromatografía de gases con espectrometría de masas (GC-MS), CLAR-CLAR, MS-MS, etc. Específicamente, un número importante de publicaciones se enfocan en la resolución de muestras complejas con el uso de cromatografía líquida, (Martínez Galera y col. 2007, Culzoni y col. 2011, Goicoechea y col. 2011) aprovechando la ventaja de segundo orden (Booksh y Kowalski 1994).

En el desarrollo de nuevas técnicas analíticas, es muy importante considerar la reducción del tiempo de análisis y, en consecuencia, los costos y la cantidad de solventes tóxicos que se utilizan, contribuyendo con la química verde (Tobiszewski y Namieśnik 2012, Gałuszka y col. 2013, de la Guardia 2010). En las últimas décadas, este enfoque ha recibido un gran interés debido a su capacidad para aprovechar la innovación en el área de la química y así cumplir objetivos ambientales y económicos al mismo tiempo (Anastas y Eghbali 2010, Asfaw y col. 2011).

En cromatografía, el factor de retención (k) determina el grado de retención en la columna de un componente en una muestra, y está dado por su partición entre la fase estacionaria y la fase móvil. En la mayoría de los análisis cromatográficos, los analitos eluyen con factores de retención entre 1 y 20, permitiendo su completa separación. Un pico con $k = 0$ se trata de un componente que no interactúa con la fase estacionaria y eluye en el volumen muerto (Dong 2006). Si los analitos se encuentran en su forma ionizada, el valor de k será significativamente menor que el valor que tendría si estuviese en su forma neutra. Un valor de k menor a 1 hace que el componente eluya muy rápido, originando desigualdades tanto en el tiempo de elución, como en la forma de los picos. En consecuencia, se generan datos no trilineales, que hacen necesaria la utilización de

MCR-ALS para su resolución satisfactoria, ya que este algoritmo puede resolver este tipo de problemas a través del aumento de matrices, u otras alternativas como PARAFAC2, una variante de PARAFAC que permite distinguir perfiles de tiempo en cada muestra (Bro 1998). Sin embargo, cuando los datos son pretratados convenientemente para aliviar los problemas mencionados anteriormente, se pueden obtener resultados satisfactorios utilizando PARAFAC o algoritmos basados en RBL (Cañada-Cañada y col. 2009, Gil García y col. 2008, Bortolato y col. 2008, Bortolato y col. 2010, Vásquez y col. 2013).

Generalmente, los colorantes sintéticos se agregan a los alimentos y las bebidas no sólo para mejorar su apariencia, color y textura, sino también para mantener el color natural durante su procesamiento o almacenamiento. Este tipo de colorantes presentan ciertas ventajas en comparación con los colorantes naturales, como, por ejemplo, alta estabilidad frente a la exposición a la luz y el oxígeno, resistencia frente a cambios de pH, uniformidad del color, baja contaminación microbiológica y costos de producción relativamente bajos (Pereira Alves y col. 2008). Sin embargo, muchos de ellos exhiben efectos adversos para la salud humana, tales como alergia, problemas respiratorios, tumores en las glándulas tiroides, daño en los cromosomas, hiperactividad, dolor abdominal, entre otros. (Silva y col. 2007).

Por otro lado, el uso de colorantes en los alimentos es muchas veces indicativo de adulteraciones, como en el caso del agregado de colorantes a jugos de frutas. Por ello, el uso de este tipo de aditivos es estrictamente controlado por leyes y regulaciones, que detallan los valores de ingesta diaria aceptable (IDA) para cada colorante. Amaranto (E123), amarillo ocaso FCF (E110) y tartrazina (E102) están entre los colorantes mayormente utilizados en bebidas no alcohólicas, y los valores de IDA están entre 0.0 y 0.5 mg Kg⁻¹ para amaranto (A), 0.0 y 2.5 mg Kg⁻¹ para amarillo ocaso (AO) y 0.0 y 7.5 mg Kg⁻¹ para tartrazina (T) (*Joint Expert Committee on Food Additives depending on Food and Agriculture Organization (FAO) y World Health Organization (WHO)*).

Algunos inconvenientes encontrados en la determinación de colorantes artificiales se relacionan con la variedad de productos comerciales que presentan mezclas de colorantes, y que, además, contienen otras sustancias que pueden ser consideradas potenciales interferentes. Por lo tanto, las estrategias de análisis se

han enfocado tradicionalmente en métodos que utilizan técnicas separativas. Las técnicas analíticas más frecuentemente utilizadas en la determinación de colorantes incluyen espectroscopia visible (Berzas y col. 1999), cromatografía en capa delgada (TLC) (Oka y col. 1987, Soponar y col. 2008, Milojevic-Opsenica y col. 2003, Dixit y col. 2008), electroforesis capilar (EC) (Jager y col. 2005, Lee y col. 2008, Dossi y col. 2007a, Dossi y col. 2007b), y mayormente CLAR (Ma y col. 2006, Yoshioka y Ichihashi 2008, Dossi y col. 2006, Miniotti y col. 2007, Pereira Alves y col. 2008, Vachirapatama y col. 2008, Yang y col. 2011, Kucharska y Grabka 2010). Sin embargo, los métodos desarrollados con estas técnicas tienen algunas desventajas, como el uso de solventes tóxicos, largos tiempos de análisis y la necesidad de aplicar técnicas de pretratamiento a la muestra. El modelado quimiométrico se utilizó para solucionar la falta de especificidad debida al solapamiento espectral en la determinación espectrofotométrica UV-Vis directa de estos colorantes, representando un método rápido, simple y económico (Ni y col. 2009, Lachenmeier y Kessler 2008). Recientemente, se publicó un método espectrofotométrico con aplicación del algoritmo MCR-ALS para determinar los tres colorantes en bebidas comerciales (Llamas y col. 2009).

En el presente capítulo se presenta un estudio comparativo del desempeño de los algoritmos MCR-ALS y U-PLS/RBL en la determinación de A, AO y T en bebidas comerciales no alcohólicas. Ambos algoritmos de calibración multivariada se aplicaron para modelar datos obtenidos a través de una separación rápida por cromatografía líquida con detección DAD, que presentaron corrimiento de picos entre diferentes corridas cromatográficas.

2.2. Objetivos

- Evaluar los algoritmos de segundo orden MCR-ALS y U-PLS/RBL en el modelado de datos no trilineales.
- Comparar los resultados obtenidos al cuantificar A, AO y T en muestras de bebidas no alcohólicas (jugos de frutas artificiales) con los obtenidos por un método de cromatografía líquida clásico, donde se logra la separación completa de los analitos.

2.3. Programas

MCR-ALS se implementó a través de la interfaz gráfica provista por R. Tauler en su página <http://www.ub.edu/mcr/welcome.html>. U-PLS/RBL se implementó a través rutinas escritas por el Dr. Olivieri incluidas en la interfaz gráfica MVC2 (Olivieri y col. 2009). Ambos algoritmos se ejecutaron utilizando MATLAB 7.1 (MATLAB 7.1, 2005).

2.4. Materiales y Métodos

2.4.1. Reactivos

Todos los experimentos se realizaron con drogas de calidad para análisis y agua MilliQ (sistema purificador Millipore, Bedford, MA, USA). Los colorantes A, AO y T se obtuvieron de Sigma-Aldrich (Bellefonte, PA, USA). Se prepararon soluciones stock de 0.10 mol L⁻¹ disolviendo cantidades apropiadas de cada colorante en agua MilliQ. Las soluciones de estándares y mezclas de los colorantes se prepararon diariamente por dilución de las soluciones stock con agua MilliQ. Se utilizaron metanol y acetato de amonio obtenidos de Sintogran (Buenos Aires, Argentina) y Cicarelli (San Lorenzo, Argentina), respectivamente.

2.4.2. Equipamiento

Se utilizó un cromatógrafo Agilent 1100 Series (Agilent Technologies, Waldbronn, Alemania) compuesto por un degasificador, una bomba cuaternaria, un inyector automático y un DAD. Se empleó una columna ZORBAX Eclipse XDB-C18 (4.6 mm × 150 mm) con tamaño de partícula de 5 µm.

2.4.3. Procedimiento

2.4.3.1 Estándares de calibración y mezclas de validación

Se prepararon cinco muestras de calibración para cada colorante en un rango de concentración de 0.05 y 3.00 mg L⁻¹ para A, de 1.00 y 45.00 mg L⁻¹ para AO y de 0.50 y 8.00 mg L⁻¹ para T. Por otro lado, se prepararon quince mezclas de validación conteniendo los tres colorantes mediante un diseño central compuesto, para evaluar los errores de predicción de los algoritmos quimiométricos (ver Tabla 2.1).

2.4.3.2. Muestras de jugos artificiales

Las muestras reales analizadas fueron siete bebidas comerciales no alcohólicas que fueron homogeneizadas, filtradas usando un filtro de membrana de 0.45 μm de tamaño de poro e inyectadas en el sistema cromatográfico sin dilución previa.

2.4.3.3. Método de referencia

El método de referencia implementado se adaptó del propuesto por Pereira Alves y col. (Pereira Alves y col. 2008). Los cromatogramas se obtuvieron a temperatura ambiente utilizando como fase móvil una mezcla de metanol:acetato de amonio 0.08 mol L⁻¹ (23:77) a una velocidad de flujo de 1 mL min⁻¹ con detección a 454, 484 y 550 nm para T, AO y A, respectivamente. En estas condiciones, el tiempo total de análisis para cada cromatograma fue de 13 min.

2.4.3.4. Método cromatográfico rápido

Con el objetivo de desarrollar un método más rápido que el propuesto en la literatura (Pereira Alves y col. 2008), se invirtió totalmente la composición de la fase móvil, siendo entonces metanol:acetato de amonio 0.08 mol L⁻¹ (77:23). El resto de las condiciones se mantuvieron como en el método de referencia.

2.5. Resultados y Discusión

2.5.1. Aspectos generales

La Figura 2.1 muestra una separación cromatográfica completa de los tres colorantes y otros componentes presentes en una muestra de jugo (muestra número 4 en la Tabla 2.2) al ser analizada por el método propuesto por Pereira Alves y col. (Pereira Alves y col. 2008). Como puede observarse, la separación completa se logra en 13.0 min, y al menos tres interferencias aparecen entre los tiempos de elusión 1.0 y 1.5 min, y entre 3.5 y 5.0 min.

Como se comentó anteriormente, un método cromatográfico más rápido puede ser una mejor opción, ya que, tanto el ahorro de solventes como de tiempo, están relacionados con los costos del análisis y el impacto ambiental. El tiempo de análisis se puede reducir significativamente al cambiar la composición de la fase móvil. Este hecho lleva a que los compuestos eluyan a tiempos parecidos, lo que

produce que los picos se solapen. Los datos así obtenidos deben ser convenientemente procesados por algoritmos de calibración multivariada para lograr selectividad por medios matemáticos.

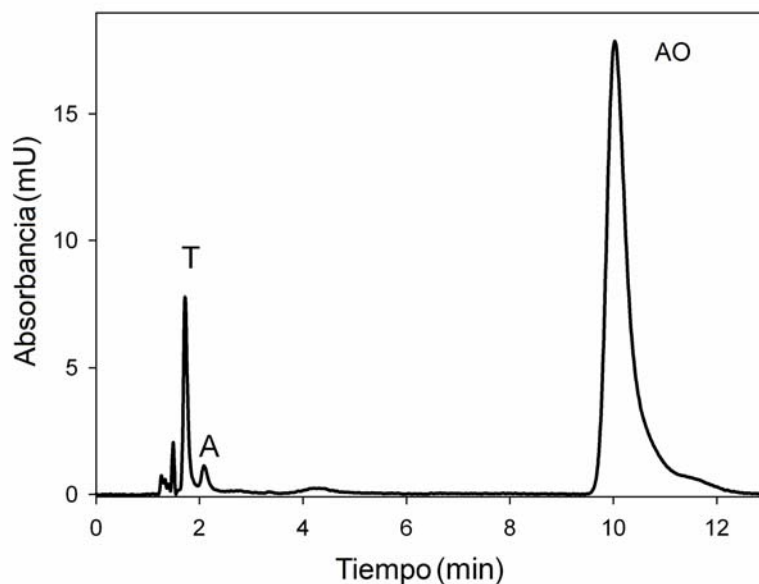


Figura 2.1. Separación cromatográfica completa (13 min) de los tres colorantes y otros componentes en una muestra de jugo (muestra número 4, Tabla 2.2) al utilizar el método propuesto por Pereira Alves y col. (Pereira Alves y col. 2008)

La Figura 2.2 muestra la superficie correspondiente a una muestra de validación que contiene los tres colorantes (muestra 1 de la Tabla 2.1), cuando la separación cromatográfica se llevó a cabo en 1.85 min y se registraron espectros cada 0.41 s en el rango 440 - 570 nm, cada 2 nm. En consecuencia, se obtuvieron matrices de dimensión 100×66 para cada muestra. Esta figura también exhibe los perfiles de elución (registrados en los máximos correspondientes de cada colorante: T a 450 nm, AO a 490 nm y A a 530 nm) y los perfiles espectrales de los tres colorantes luego de analizar la matriz de datos de la mezcla ternaria con el algoritmo MCR-ALS (matrices **S** y **C** de la Ecuación (1.9)). Como se puede apreciar, existe un severo solapamiento entre los tres componentes, por lo que es imposible aplicar calibración univariada para la resolución de estos datos. A su vez, el perfil de elución de AO está extremadamente deformado. Esta situación puede entenderse ya que el factor de retención es cero para todos los analitos y, además, AO se encuentra a muy altas concentraciones en las muestras.

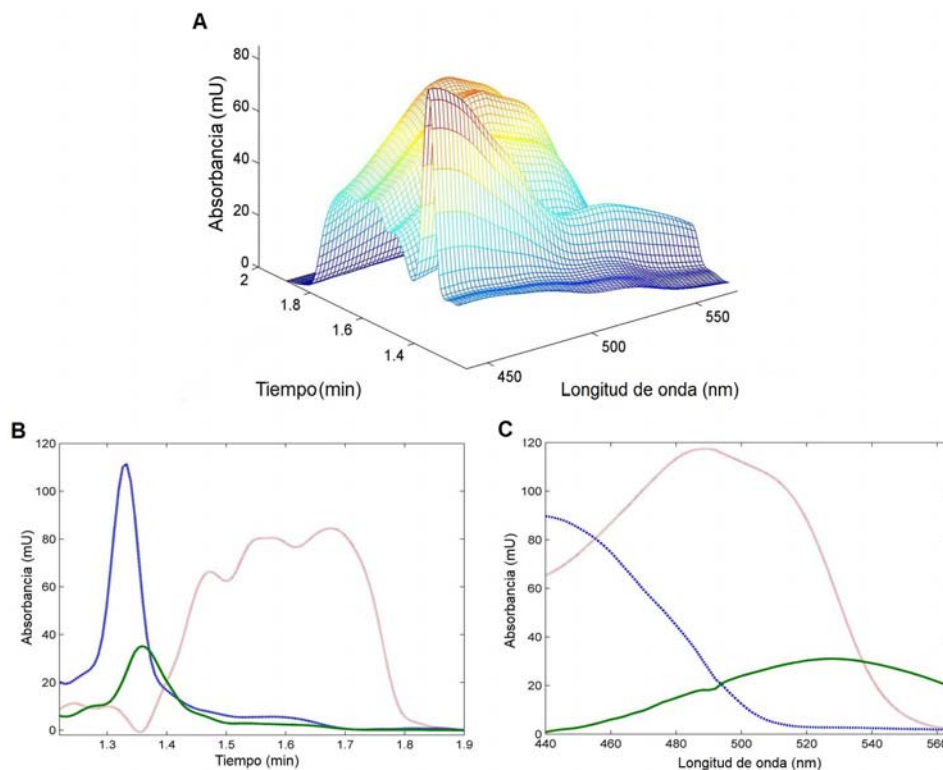


Figura 2.2. (A) Superficie correspondiente a una mezcla ternaria cuando la separación cromatográfica se lleva a cabo en 1.85 min y se registran los espectros entre 440 y 570 nm. (B) Perfiles de elución registrados en los máximos correspondientes de cada colorante (T a 450 nm, AO a 490 nm y A a 530 nm) y (C) espectros de los tres colorantes, obtenidos al aplicar MCR-ALS a la matriz de datos de la mezcla (T: línea azul, A: línea verde y AO: línea roja).

Debemos considerar que las muestras de bebidas comerciales contienen componentes inesperados, como se demostró en la Figura 2.1. Por tal motivo, resulta necesario explotar la ventaja de segundo orden (Booksh y Kowalski 1994). La Figura 2.3 muestra los cromatogramas registrados a 440 nm, correspondientes a una mezcla ternaria (muestra 1 en Tabla 2.1) y a una muestra real (muestra 5 en Tabla 2.2). Como se puede apreciar, aparecen componentes no esperados, especialmente entre 1.4 y 1.9 min.

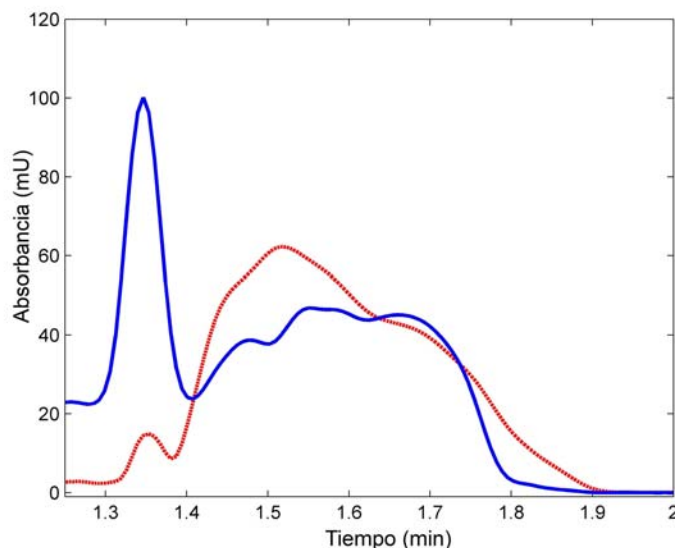


Figura 2.3. Cromatogramas correspondientes a una mezcla ternaria en línea sólida azul (muestra 1 en Tabla 2.1) y a una muestra real en línea de puntos roja (muestra 5 en Tabla 2.2) registrados a 440 nm.

Por otro lado, una inspección visual de la Figura 2.4, donde se muestran los cromatogramas de las quince mezclas de validación registrados a 440 nm, lleva a la conclusión de que se produce un corrimiento en el tiempo entre las diferentes corridas cromatográficas. Este hecho se evidencia especialmente en el rango de 1.5-1.6 min. Por el contrario, las diferencias en la región de 1.3-1.4 min no pueden ser atribuidas solamente al corrimiento en el tiempo, sino también a los cambios en las concentraciones relativas de T y A, cuyos picos están muy solapados. Esto último se puede apreciar mejor en la Figura 2.5, donde se exponen los cromatogramas (registrados a 500 nm) correspondientes a soluciones estándares de T (8.00 mg L^{-1}), A (1.53 mg L^{-1}) y una mezcla de ambos en las mismas concentraciones. Debido a que no es posible alinear los picos sin introducir error, uno de los desafíos de la presente tesis fue la cuantificación de los tres colorantes en muestras de jugo sin la aplicación de ningún pretratamiento a las señales obtenidas.

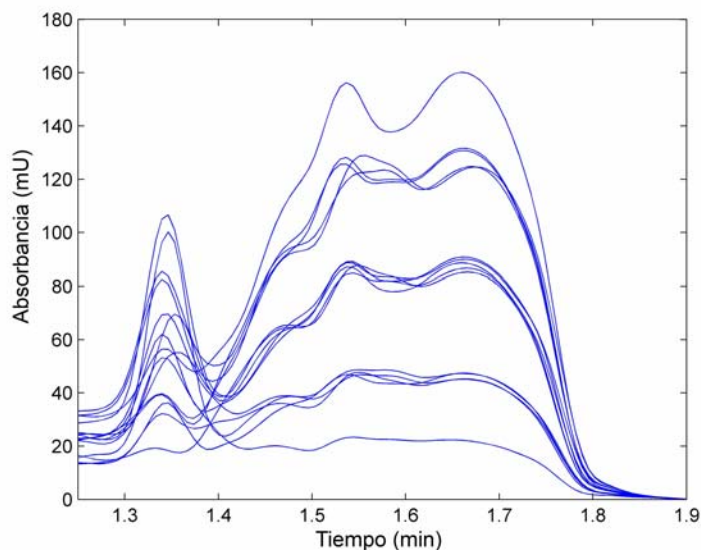


Figura 2.4. Cromatogramas correspondientes a las quince mezclas de validación registrados a 440 nm.

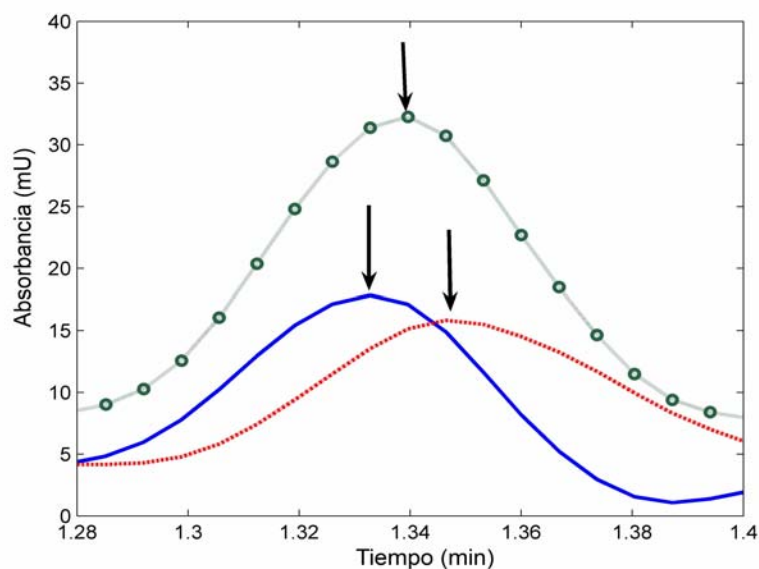


Figura 2.5. Cromatogramas (registrados a 500 nm) correspondientes a soluciones estándares de T (8.00 mg L^{-1}) en línea sólida azul, A (1.53 mg L^{-1}) en línea de puntos roja y a una mezcla de ambos colorantes en las mismas concentraciones en línea de círculos verde.

2.5.2. Aplicación de algoritmos de segundo orden

2.5.2.1. Mezclas de validación

El algoritmo MCR-ALS se aplicó para analizar simultáneamente las 15 mezclas de validación (Tabla 2.1) utilizando los estándares puros descritos en la

Sección 2.4.3.1 de este capítulo. Se generó una matriz aumentada \mathbf{D} apilando en columnas las matrices correspondientes a los espectros registrados durante la corrida cromatográfica de los estándares de calibración junto con las correspondientes a las mezclas de validación.

En el presente estudio, se aplicaron las restricciones de no negatividad en los perfiles espectrales y de concentración para todos los analitos, y unimodalidad en los perfiles de concentración solo para T y A. Resulta importante destacar que para realizar la descomposición de la matriz \mathbf{D} generada como se indicó, los espectros puros de los componentes deben ser los mismos en todos los experimentos, pero los perfiles en cada submatriz \mathbf{C} no necesitan compartir una forma en común, como sucede cuando se pierde trilinealidad (el presente caso).

Antes de comenzar la resolución por MCR-ALS, la determinación del número de componentes se llevó a cabo mediante la descomposición en valores singulares (SVD), mientras que las estimaciones iniciales de \mathbf{S}^T se crearon a través de SIMPLISMA (Windig y Guilment 1991). En algunos casos, los resultados obtenidos de la resolución por MCR no fueron satisfactorios. A pesar de este hecho, los espectros optimizados por MCR se almacenaron y se utilizaron como estimaciones iniciales en análisis subsecuentes, lo que permitió alcanzar resultados con indicadores de calidad analítica satisfactorios. Esta estrategia demostró ser muy efectiva en los casos en donde SIMPLISMA no fue capaz de proveer información inicial adecuada. Los resultados de predicción de las muestras de validación se presentan en la Tabla 2.1 y, como puede apreciarse, son muy satisfactorios.

Por otro lado, cuando se aplicó U-PLS/RBL, el número de variables latentes para cada analito fue igual a uno debido a que se utilizaron las soluciones de estándares puros en la etapa de calibración. La presencia de componentes inesperados (en este caso los otros dos colorantes) tuvo que ser considerada para disminuir los residuos de predicción para las muestras de validación hasta que éstos se estabilizaran en un valor comparable con el ruido instrumental (como se comentó en la Sección 1.2.2.2 del Capítulo 1). Los residuos de predicción en función de valores de prueba N_{int} se monitorizaron para todas las muestras. Cuando N_{int} fue igual a 3, el valor residual alcanzó un valor comparable al ruido instrumental (0.06 en unidades de absorbancia para este sistema). Por lo tanto,

$N_{\text{int}} = 3$ se eligió como la opción correcta de acuerdo al principio de funcionamiento del algoritmo. Evidentemente, el sistema sólo hubiese necesitado dos componentes inesperados, debido a que las mezclas son combinaciones de tres componentes. Sin embargo, un tercer factor fue necesario debido a la pérdida de trilinealidad. No obstante, al utilizar este número de interferentes los resultados obtenidos fueron inaceptables, con errores relativos de predicción de aproximadamente 40 % para los tres colorantes (estos datos no se muestran).

Por lo mencionado anteriormente, se decidió seguir una nueva estrategia para el modelado de los datos con U-PLS/RBL, que consistió en utilizar las mezclas ternarias en el paso de calibración en lugar de los patrones puros. En una publicación reciente, Cañada-Cañada y col. obtuvieron resultados satisfactorios al aplicar un método basado en RBL (N-PLS/RBL) a datos de CLAR con detección fluorescente, utilizando mezclas de dos analitos para calibrar (Cañada-Cañada y col. 2009). En este capítulo, se aplicó el método de validación cruzada, que consiste en realizar 15 modelos de calibración usando cada vez 14 de las muestras descritas en la Tabla 2.1, y prediciendo la restante. Este procedimiento se repitió hasta que se obtuvieron todos los valores de predicción de las muestras, para crecientes valores del número de componentes. En este caso, el número de variables latentes estimado por el método de validación cruzada fue 3, y el número de componentes inesperados fue cero, por lo tanto no fue necesaria la aplicación de la ventaja de segundo orden. Como se puede apreciar en la Tabla 2.1, los resultados presentan cifras de mérito comparables a las obtenidas por MCR-ALS.

Para evaluar la exactitud de los modelos, los valores de concentración predicha obtenidos al aplicar ambos algoritmos se compararon con los valores nominales de cada colorante. Con este propósito, se aplicó una prueba de regiones elípticas de confianza conjunta (EJCR) donde se construyen elipses a partir del ajuste por cuadrados mínimos ordinarios de las concentraciones nominales versus las predichas para cada método, y se compara el valor de la pendiente y la ordenada al origen con sus valores ideales de 1 y 0. Si los métodos son exactos (no presentan errores constantes ni proporcionales), estos valores estarán incluidos en las elipses. Como se sugirió previamente, al realizar un análisis multivariado, es conveniente incluir datos experimentales correspondientes a todos los analitos para mejorar la estimación de la variancia correspondiente a la regresión recién

mencionada. Esto evita el sobredimensionamiento de la región de confianza conjunta debido a errores aleatorios experimentales y, por lo tanto, la probabilidad de incluir al punto (1,0) sólo por azar (Martínez y col. 2000).

Tabla 2.1. Composición y predicción de las muestras ternarias utilizadas para calibración y validación.

Muestra	Tartrazina (mg L ⁻¹)			Amaranto (mg L ⁻¹)			Amarillo ocaso (mg L ⁻¹)		
	Nom.	MCR ^a	RBL ^b	Nom.	MCR ^a	RBL ^b	Nom.	MCR ^a	RBL ^b
1	5.89	6.13	5.85	2.20	2.20	2.05	12.09	11.95	12.31
2	2.61	2.26	2.80	1.05	1.15	1.07	32.91	33.80	34.43
3	5.89	5.81	5.44	1.05	1.11	1.05	12.09	12.50	12.89
4	2.61	2.56	2.48	2.20	2.24	2.25	32.91	33.87	34.31
5	4.25	4.09	4.53	0.65	0.74	0.67	22.50	22.80	22.75
6	2.61	2.72	2.38	2.20	2.20	2.14	12.09	12.00	11.72
7	1.50	1.39	1.58	1.63	1.67	1.56	22.50	23.08	23.94
8	4.25	4.46	4.43	1.63	1.64	1.62	5.00	5.58	5.84
9	4.25	4.47	4.32	2.60	2.70	2.78	22.50	23.37	23.75
10	5.89	5.92	6.26	2.20	2.22	2.17	32.91	34.09	34.45
11	5.89	5.64	5.96	1.05	1.16	1.03	32.91	33.87	34.38
12	2.61	2.60	2.66	1.05	1.12	1.13	12.09	12.27	13.78
13	7.00	7.11	6.96	1.63	1.68	1.66	22.50	23.17	24.43
14	4.25	4.21	4.20	1.63	1.69	1.66	22.50	22.66	24.36
15	4.25	4.00	3.98	1.63	1.69	1.61	40.00	41.09	44.29
Rec. (%) ^c	-	98.8 (5.1)	100.3 (5.3)	-	104.4 (4.1)	100.4 (3.9)	-	102.7 (2.9)	106.3 (5.0)
REP (%) ^d	-	4.1	4.9	-	3.9	4.4	-	3.1	7.5

^a MCR-ALS: tres factores (regiones: ver en el texto).

^b U-PLS/RBL: cinco variables latentes para cada analito (regiones: ver en el texto).

^c Recuperación promedio. Desviación estándar entre paréntesis.

^d Error relativo de predicción, $REP = \frac{100}{\bar{c}} \left[\frac{1}{I} \sum_{i=1}^I (c_{nom} - c_{pred})^2 \right]^{1/2}$, donde I es el número de

muestras, c_{nom} and c_{pred} son las concentraciones nominal y predicha, y \bar{c} es la concentración promedio.

La Figura 2.6 muestra las elipses obtenidas para las comparaciones de las predicciones con los valores nominales del conjunto de datos globales cuando se aplicaron los algoritmos MCR-ALS y U-PLS/RBL. Como se puede observar, ambas elipses contienen los valores teóricos de ordenada al origen (0) para un intervalo de confianza de 95%. Por el contrario, no contienen el valor teórico de la pendiente (1). Este hecho es indicativo de la presencia de error proporcional. Además, el menor tamaño y la mayor cercanía al valor esperado (1) de la elipse correspondiente a MCR-ALS, nos permite concluir acerca de la mayor precisión y menor error proporcional de este algoritmo, al compararlo con U-PLS/RBL (González y col. 1999).

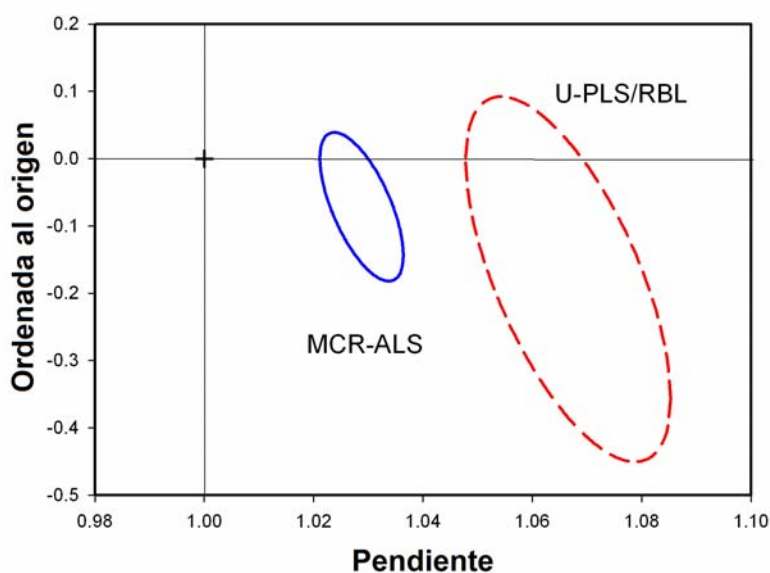


Figura 2.6. Regiones elípticas para el conjunto de datos globales de predicción aplicando U-PLS/RBL y MCR-ALS a las 15 mezclas de validación de la Tabla 2.1.

Debe considerarse que los modelos para MCR-ALS se construyeron sólo con estándares de calibración puros, mientras que para U-PLS/RBL fue necesario construir los modelos con las mezclas de validación. Este hecho demuestra nuevamente que MCR-ALS posee una mayor habilidad para predecir las muestras en los casos en donde los datos presentan importantes desvíos de la trilinealidad.

2.5.2.2. Muestras reales

La complejidad de las muestras reales se puede apreciar en los cromatogramas de la Figura 2.7. Estas muestras fueron analizadas como se indicó

para las 15 mezclas de validación. Los resultados se muestran en la Tabla 2.2, y, como se puede apreciar, las predicciones obtenidas por U-PLS/RBL (se utilizaron las 15 mezclas ternarias en el paso de calibración) son pobres al compararlas con las obtenidas por el método CLAR, el cual garantiza la completa separación de los analitos y las interferencias (ver Figura 2.1). Este hecho sugiere que los datos no pueden ser debidamente modelados debido a la falta de trilinealidad. Además, en la mayoría de las muestras se debieron considerar dos o tres interferentes para alcanzar valores de los residuos comparables con el ruido instrumental.

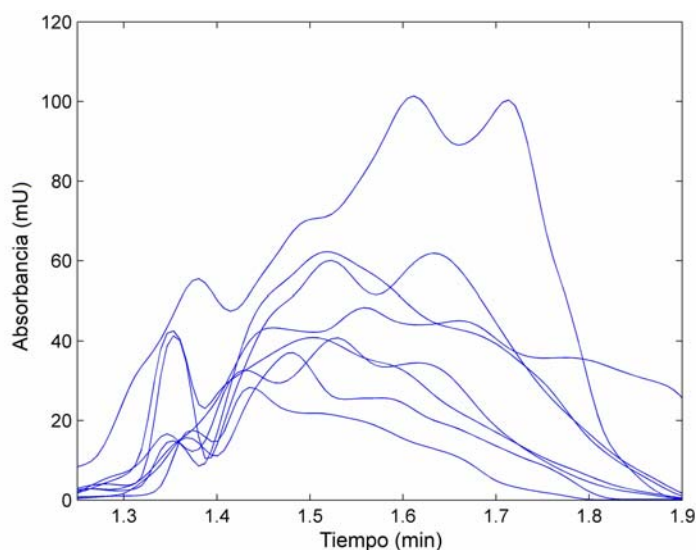


Figura 2.7. Cromatogramas correspondientes a las siete muestras reales registradas a 440 nm.

Por otro lado, la Tabla 2.2 muestra que los resultados que proporciona MCR-ALS se pueden considerar aceptables cuando se los compara con los obtenidos por el método de referencia.

Tabla 2.2. Predicción de muestras reales por el método de referencia y el método rápido propuesto acoplado a MCR-ALS y U-PLS/RBL.

Muestra	Tartrazina (mg L ⁻¹)			Amaranto (mg L ⁻¹)			Amarillo oca (mg L ⁻¹)		
	CLAR ^a	MCR	RBL	CLAR ^a	MCR	RBL	CLAR ^a	MCR	RBL
1	-	-	-	0.22	0.23	0.28	30.65	30.46	30.81
2	-	-	-	0.14	0.16	0.18	15.61	16.05	15.70
3	3.12	2.87	2.60	0.13	0.13	0.14	6.02	5.81	5.12
4	2.29	2.02	2.01	0.77	0.70	0.60	6.33	6.80	4.90
5	0.81	0.75	0.84	-	-	-	14.55	14.75	13.20
6	0.94	0.87	1.17	-	-	-	14.37	14.80	12.50
7	1.40	1.69	1.62	0.11	0.11	0.12	2.34	2.93	2.98
Rec.	-	97.2	103.0	-	101.9	110.1	-	105.1	95.5
(%) ^b		(13.3)	(16.7)		(8.5)	(20.5)		(9.5)	(16.3)
REP (%)	-	10.5	14.8	-	10.1	25.5	-	3.0	8.6

^a Método de referencia CLAR (ver Pereira Alves y col. 2008).

^b Recuperación promedio. Desviación estándar entre paréntesis.

En las Figuras 2.8 A y B se presentan los perfiles de tiempo y espectrales extraídos por MCR-ALS cuando se analizó la muestra 7 (Tabla 2.2) a modo de ejemplificar cómo se implementó este algoritmo. Esta muestra contiene los tres colorantes y, además, una interferencia no esperada. Como se puede ver, este compuesto presenta un espectro similar al de T, mientras que su perfil de tiempo coeluye con AO. Los parámetros estadísticos (porcentaje de falta de ajuste = 3.31 y porcentaje de varianza explicada, $r^2 = 99.8904$) y las cifras de mérito obtenidas son muy satisfactorios y nos indican que MCR-ALS tiene un rendimiento aceptable cuando se procesan este tipo de datos instrumentales tan complejos.

Finalmente, se debe considerar el ahorro de solventes y la disminución de los costos. Por un lado, el tiempo de análisis se pudo reducir a un 14.2 % del tiempo necesario para realizar una separación completa. Por otra parte, la separación completa consume 10.01 mL de metanol y 2.99 mL de acetato de amonio 0.08 mol L⁻¹, mientras que el método cromatográfico propuesto requiere sólo 0.46 mL de metanol y 1.54 mL de acetato de amonio 0.08 mol L⁻¹. Estos datos

apoyan el uso de este tipo de metodologías, más aún si la química verde es seriamente tenida en cuenta (Anastas y Eghbali 2010, Asfaw y col. 2011).

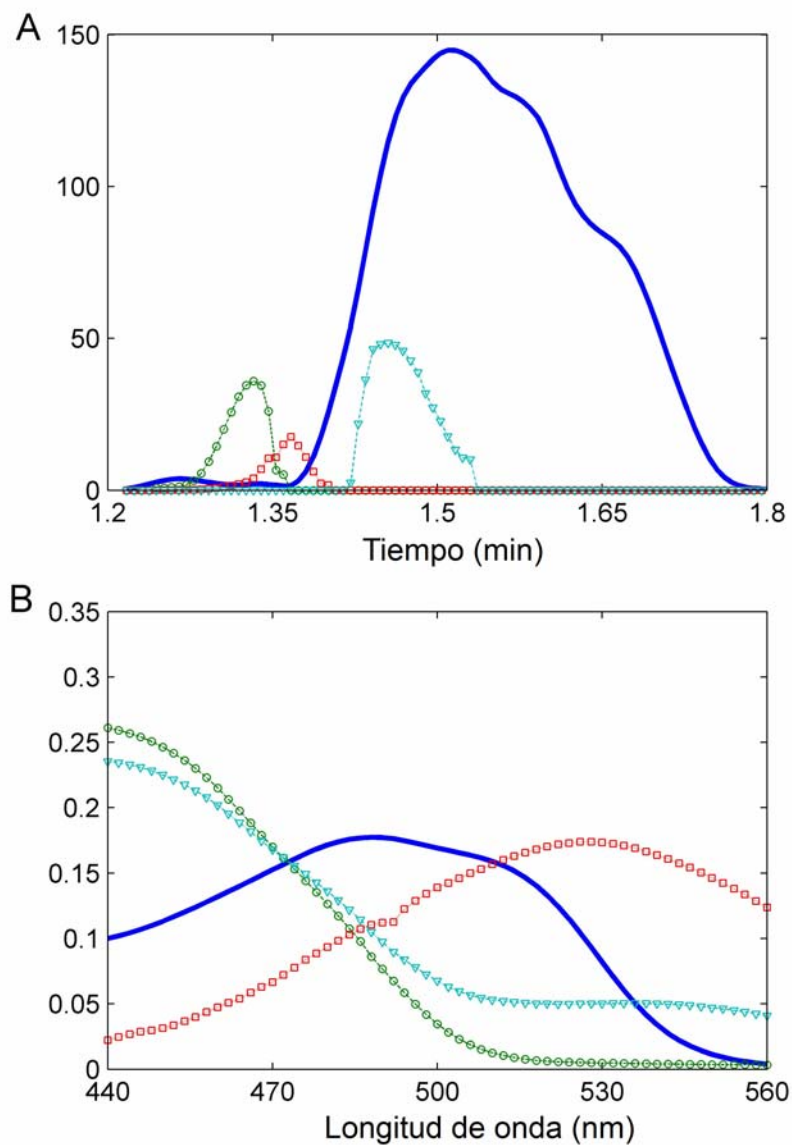


Figura 2.8. (A) Perfiles de tiempo extraídos por MCR-ALS cuando se analizó la muestra 7 (Tabla 2.2) que contiene los tres colorantes (T: línea de círculos verde, A: línea de cuadrados roja, AO: línea sólida azul) más un interferente (línea de triángulos celeste). (B) Perfiles espectrales.

2.6. Conclusiones

La reducción del tiempo de análisis de un método cromatográfico debe tenerse muy en cuenta considerando que se disminuye el uso de solventes, lo que impacta fuertemente en el medioambiente, como así también en los costos del

análisis. Esta reducción de tiempo, además de un menor consumo de reactivos, puede lograrse optimizando estrategias analíticas que involucren la generación de datos cromatográficos de segundo orden y su conveniente modelado con algoritmos adecuados.

Los datos no trilineales obtenidos por CLAR-DAD generados en el presente trabajo no pudieron ser debidamente modelados por U-PLS/RBL. Sin embargo, MCR-ALS resultó una herramienta muy útil para procesar este tipo de datos, permitiendo obtener cifras de mérito aceptables.



CAPÍTULO 3

Determinación de tartrazina en bebidas mediante análisis por inyección en flujo (FIA) con flujo interrumpido y calibración multivariada de datos cinéticos no lineales

3.1. Introducción

La tartrazina (E-102) es un colorante sintético muy utilizado como aditivo alimentario, que se agrega en comidas, bebidas, medicamentos y cosméticos. Su color amarillo se debe a la presencia del grupo funcional azo (N=N) y de grupos aromáticos en su estructura molecular (ver Figura 3.1)

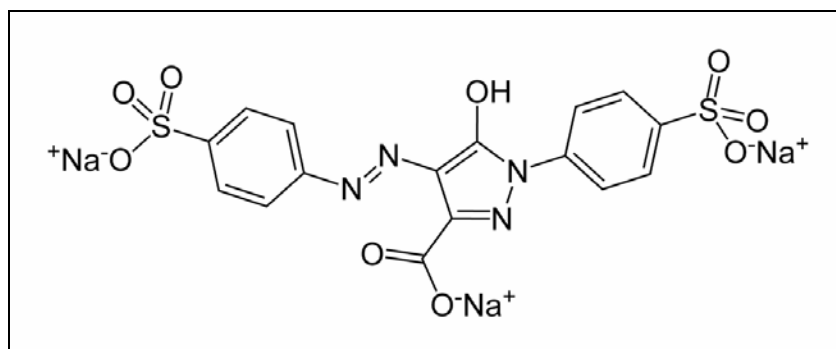


Figura 3.1. Estructura molecular del colorante tartrazina (E-102).

El Código Alimentario Argentino establece que la concentración máxima de tartrazina en bebidas no alcohólicas y alcohólicas es de 100 y 200 mg L⁻¹, respectivamente, y expresa que es obligatorio que en la etiqueta se especifique la presencia de estos productos (Código Alimentario Argentino 2010).

Las técnicas analíticas que se utilizan con mayor frecuencia para la determinación de colorantes utilizan solventes tóxicos, insumen largos tiempos de análisis, y a veces es necesaria la aplicación de técnicas de pretratamiento a la muestra. Por el contrario, la combinación de metodologías simples (como espectroscopia) con modelado quimiométrico para la determinación de colorantes, representa una estrategia rápida, simple y menos costosa (Al Degs 2009, Dinc y col. 2002, Lachenmeier y Kessler 2008, Llamas y col. 2009).

En el presente capítulo se midieron los cambios espectrales en función del tiempo de tartrazina causados por la reacción cinética con bromato de potasio catalizada por Fe(II), con el objetivo de realizar su cuantificación en bebidas. Debido a que el registro de datos en el tiempo se debe controlar cuidadosamente cuando se llevan a cabo este tipo de determinaciones, su automatización es esencial (Araújo y col. 1998). Por lo tanto, un sistema de inyección de flujo detenido es muy adecuado para llevar a cabo esta tarea, ya que se basa en la

medida de tiempos reproducibles. Para este propósito, la muestra y el oxidante se mezclaron automáticamente e inyectaron en el solvente transportador hacia el interior de la cámara de mezcla. Luego, el flujo se detuvo y se registró la evolución temporal de los espectros. En la bibliográfica se encuentran varios trabajos relacionados con el uso de esta metodología (Murillo Pulgarín y García Bermejo 1996, Muñoz de la Peña y col. 2002, Pitonesi y col. 2004, Wang y Lu 2005, Shabania y col. 2011).

Cuando se adquieren espectros en el tiempo durante una reacción cinética, los datos generados pueden ser modelados exitosamente con algoritmos de segundo orden. A medida que se originan nuevos productos en el proceso cinético, las diferencias en los espectros o en las características cinéticas, con respecto a los reactantes, se pueden utilizar para distinguir los componentes de una mezcla (Escandar y col. 2007), en especial, porque se explota la ventaja de segundo orden, que permite conocer la concentración de un analito en presencia de sustancias desconocidas. Recientemente, se desarrolló un método cinético-espectroscópico indirecto para la determinación simultánea de tartrazina y otros cuatro colorantes sintéticos basada en la reducción de Fe(III) por los colorantes seguido de una reacción cromogénica con hexacianoferrato(III). Los datos obtenidos se procesaron mediante diferentes algoritmos quimiométricos (Ni y col. 2009).

Resulta interesante destacar que los datos de dos vías generados por este tipo de métodos cinéticos no requieren de la formulación de modelos cinéticos para realizar la calibración multivariada. Esto representa una gran ventaja con respecto a los métodos cinéticos-espectrofotométricos tradicionales, ya que permite el desarrollo de modelos de calibración empíricos para predecir la concentración del analito en muestras desconocidas (Esteves da Silva y Oliveira 1999). Como se comentó en el Capítulo 1, los datos de segundo orden se procesan tradicionalmente con los algoritmos MCR-ALS (De Juan y Tauler 2001) y PARAFAC (Bro 1997), aunque recientemente, ganaron importancia los algoritmos de segundo orden alternativos U-PLS/RBL y N-PLS/RBL, los que serían más flexibles y proveerían mejores cifras de mérito que sus competidores (Bortolato y col. 2008, Lozano y col. 2008, Bortolato y col. 2010).

Por otro lado, es común que se originen no linealidades en la relación absorbancia-concentración de algunos sistemas cinéticos-espectroscópicos. Este

tipo de comportamientos pueden ser acreditados a distintas causas tales como ruido instrumental y fuentes químicas (por ejemplo, en un equilibrio químico) o físicas (respuesta no lineal del detector, difusión de la luz a altas densidades ópticas y deriva de la línea de base). Este efecto se clasifica como un efecto no lineal real (Gemperline y col. 1991). Los algoritmos quimiométricos mencionados anteriormente están basados en modelos lineales y, por lo tanto, en principio, no resultarían aplicables al sistema en estudio, en el que se comprobó la existencia de una relación no lineal entre la señal y la concentración del analito.

El método de U-PCA/RBL se propuso recientemente como un algoritmo capaz de modelar información de segundo orden no lineal (Olivieri 2005b) que presenta la ventaja de segundo orden (Culzoni y Goicoechea 2007, García-Reiriz y col. 2007, García-Reiriz y col. 2008).

Con el fin de modelar la relación entre los datos instrumentales y la concentración del analito en cada muestra de calibración, se construyó un modelo univariado no lineal con las áreas bajo los perfiles cinéticos (MCR-ALS) o los *scores* (U-PCA/RBL) y sus concentraciones asociadas. Estos modelos se aplicaron para cuantificar tartrazina en muestras de validación que, además, contenían diferentes cantidades de azul brillante (E-133) y amarillo ocaso FCF (E-110) como interferencias, como así también en bebidas comerciales.

3.2. Objetivos

- Estudiar el rendimiento de MCR-ALS al aplicarlo al análisis de datos obtenidos mediante FIA con flujo interrumpido y monitorización espectrofotométrica con el objetivo de determinar tartrazina en bebidas. Comparar los resultados obtenidos con los proporcionados por U-PCA/RBL, ya que este algoritmo aparenta ser el más adecuado para procesar datos de segundo orden no lineales.
- Comparar los resultados obtenidos al cuantificar tartrazina en muestras de bebidas comerciales con los adquiridos con un método clásico de cromatografía líquida, mediante el que los analitos se separan completamente.

3.3. Programas

MCR-ALS se implementó usando la interfaz gráfica disponible en Internet a través de la página <http://www.mcrals.info>. Para la aplicación de U-PCA/RBL se

utilizó la interfaz gráfica MVC2 desarrollada por Olivieri y col. (Olivieri y col. 2009). Los algoritmos multivariados se ejecutaron en MATLAB (Matlab 7.1, 2005).

3.4. Materiales y Métodos

3.4.1. Reactivos y soluciones

Todos los experimentos se realizaron con drogas calidad para análisis y agua MilliQ (sistema purificador Millipore, Bedford, MA, USA).

Los estándares sólidos de tartrazina (T), azul brillante (AB) y amarillo ocaso FCF (AO) se obtuvieron de Ardinet (Buenos Aires, Argentina) con pureza mayor a 95% en todos los casos. Se prepararon soluciones stock de T (0.002 mol L^{-1}), AB (0.001 mol L^{-1}) y AO (0.002 mol L^{-1}) por disolución de cantidades apropiadas de cada droga en agua MilliQ.

Se preparó una solución stock de Fe(II) de 0.02 mol L^{-1} mediante disolución de una cantidad apropiada de $\text{Fe}(\text{NH}_4)(\text{SO}_4)_2 \cdot 6\text{H}_2\text{O}$ (Cicarelli, San Lorenzo, Argentina) en agua, el agregado de 5.0 mL de H_2SO_4 y la dilución hasta la marca en un matraz volumétrico de 500.00 mL con agua MilliQ. Además, se prepararon soluciones de bromato de potasio (Cicarelli, San Lorenzo, Argentina) 0.1 mol L^{-1} y ácido fosfórico (Cicarelli, San Lorenzo, Argentina) 3.0 mol L^{-1} .

El acetato de amonio y el metanol de grado HPLC se obtuvieron de Cicarelli (San Lorenzo, Argentina) y Sintogran (Buenos Aires, Argentina), respectivamente.

3.4.2. Equipamiento y procedimiento instrumental

Se desarrolló un sistema de análisis por inyección en flujo (FIA) con flujo interrumpido utilizando cinco módulos (degasificador, bomba, válvula de inyección, inyector automático y DAD) de un instrumento Agilent 1100 Series (Agilent Technologies, Waldbronn, Alemania). Se implementó un programa de inyección para mezclar de manera automática $95 \mu\text{L}$ de cada muestra con $5 \mu\text{L}$ de bromato de potasio 0.1 mol L^{-1} e inyectar la mezcla dentro de un solvente transportador compuesto por agua MilliQ fluyendo a 1.5 mL min^{-1} , a través de una cámara de mezcla de $800 \text{ mm} \times 0.5 \text{ mm}$ de diámetro interno. Transcurridos 42 s desde la inyección de la muestra, la bomba se paró y se monitorizó la reacción durante 108 s. Una vez que se alcanzó el tiempo establecido, el flujo se reanudó. Para cada pico FI se registraron espectros en el rango de 400 – 650 nm cada 1 nm,

a intervalos regulares de 0.4 s durante un tiempo total de 150 s. Por lo tanto, se generaron matrices de tamaño 375×251 por muestra, aunque luego se seleccionaron regiones en ambas dimensiones antes de emplearlas en calibración multivariada.

El método cromatográfico se adaptó a partir del propuesto por Pereira Alves y col. (Pereira Alves y col. 2008). Los cromatogramas se registraron utilizando el mismo instrumento Agilent 1100 Series, aunque las separaciones se llevaron a cabo usando una columna (4.6×75 mm) ZORBAX Eclipse XDB-C18 de tamaño de partícula $3.5 \mu\text{m}$ (Agilent Technologies).

3.4.3. Procedimiento analítico

3.4.3.1. Estándares de calibración y mezclas de validación

Se preparó un juego de calibración de diez muestras de T en concentraciones igualmente distribuidas entre 2.00 y 20.00 mg L^{-1} . Además, se prepararon dos juegos de validación empleando concentraciones distintas a las utilizadas en la calibración, siguiendo un diseño central compuesto. Uno de ellos contiene nueve muestras (J1) y el otro veinte muestras (J2) con diferentes concentraciones de T, AB y AO en un rango de 2.50 a 19.50 mg L^{-1} (ver Tabla 3.1 y 3.2).

Tanto las muestras de calibración como las de validación se prepararon midiendo alícuotas adecuadas de las soluciones estándares de cada colorante, colocándolas en matraces volumétricos de 10.00 mL para obtener las concentraciones deseadas, agregando 200 μL de ácido fosfórico 3.0 mol L^{-1} y 100 μL de solución stock de Fe(II) 0.02 mol L^{-1} , y completando hasta el aforo con agua MilliQ. Todas las muestras se prepararon por duplicado y se inyectaron en el sistema FIA de manera automática siguiendo el procedimiento descrito en el apartado 3.4.2.

Tabla 3.1. Composición del juego de validación J1 siguiendo un diseño experimental central compuesto.

Muestra	Concentración (mg L ⁻¹)	
	Tartrazina	Azul brillante
1	11.09	2.50
2	16.87	17.24
3	19.28	11.02
4	4.99	17.24
5	2.50	11.02
6	11.09	19.64
7	11.09	11.02
8	16.87	4.99
9	4.99	4.99

Tabla 3.2. Composición del juego de validación J2 siguiendo un diseño experimental central compuesto.

Muestra	Concentración (mg L ⁻¹)		
	Tartrazina	Azul brillante	Amarillo ocaso
1	10.99	11.02	10.94
2	10.99	11.02	10.94
3	10.99	11.02	10.94
4	16.00	16.09	5.95
5	10.99	11.02	2.44
6	16.00	5.94	16.04
7	10.99	11.02	19.55
8	10.99	2.49	10.94
9	10.99	11.02	10.94
10	5.98	16.09	5.95
11	10.99	11.02	10.94
12	10.99	19.54	10.94
13	10.99	11.02	10.94
14	5.98	5.94	16.04
15	19.47	11.02	10.94
16	2.51	11.02	10.94
17	5.98	16.09	16.04
18	5.98	5.94	5.95
19	16.00	5.94	5.95
20	16.00	16.09	16.04

3.4.3.2. Muestras de bebidas comerciales

Las muestras analizadas fueron tres jugos listos para consumo (LC1: Ser Naranja-Durazno, LC2: Gatorade Green Mango, LC3: Style Naranja-Durazno), un jugo en polvo (BP: Boing Manzana) y una bebida alcohólica espumante (BA: Frizze Green Evolution), que contienen tartrazina y otros colorantes, que se compraron en un supermercado local. Todas las muestras contienen ácido cítrico como acidificante. El resto de los ingredientes declarados en sus etiquetas son: endulzantes artificiales (aspartamo, acesulfamo de potasio, sacarina, ciclamato de sodio y sucralosa), saborizantes frutales y conservantes (benzoato de sodio, sorbato de potasio, dióxido de sulfuro y sulfato de magnesio). En todas las muestras se indica contenido de agua, salvo en la muestra en polvo, y, además, dos muestras presentan gas carbónico en su composición.

Las muestras se homogeneizaron y, en caso de ser necesario, se degasificaron mediante sonicado. En el caso de la muestra sólida, se homogeneizó en su propio paquete, se pesó todo su contenido y, finalmente, se disolvió en 1000.0 mL de agua MilliQ a temperatura ambiente.

Las soluciones muestra se colocaron en matraces volumétricos de 10.00 mL con 200 μ L de ácido fosfórico y 100 μ L de solución stock de Fe(II). Posteriormente se centrifugaron a 6000 rpm durante 2 min y se inyectaron en el sistema FIA. A pesar de que una inspección visual de las muestras no revelaba sedimentación o partículas en suspensión, de igual manera, se centrifugaron para asegurar la remoción total de partículas y prevenir obstrucciones en el sistema FIA.

3.4.3.3. Procedimiento cromatográfico

La concentración de T en las muestras comerciales se verificó aplicando el método CLAR adaptado del propuesto por Pereira Alves y col. (Pereira Alves y col. 2008), a temperatura ambiente, utilizando como fase móvil una mezcla de metanol:acetato de amonio 0.08 mol L⁻¹, flujo de 1.0 mL min⁻¹ y detección UV a 454 nm. Se aplicó un programa de gradiente de elución: 15 % de metanol durante los primeros 2.5 min, luego se aumentó a 50 % durante 6.5 min y finalmente se disminuyó a 15 %. En estas condiciones, el tiempo de análisis total de cada cromatograma fue de 10.0 min.

3.5. Resultados y Discusión

3.5.1. Características espectrales y cinéticas

La reacción entre T y bromato de potasio en presencia de Fe(II) se aplicó con anterioridad en trabajos desarrollados en nuestro grupo (Culzoni y col. 2008). Durante esta reacción de oxidación, la absorbancia de T (pico a 430 nm) decrece dando lugar a un componente incoloro. Resulta muy interesante destacar que el colorante AB también reacciona con el bromato, sufriendo la degradación y generación de un producto cuyo espectro es muy parecido al de T, constituyéndose en un interferente muy difícil de eliminar. En el caso de AO, su espectro presenta menor similitud con el de T (aunque presenta un hombro a 420 nm), y su señal también decrece en presencia del oxidante, generando un producto incoloro. Los espectros de los tres colorantes pueden apreciarse en la Figura 3.2, como así también el espectro del producto de reacción de AB. La Figura 3.3 muestra las curvas cinéticas para estos compuestos en sus máximos de absorción (430, 630 y 490 nm para T y el producto de AB, AB y AO, respectivamente).

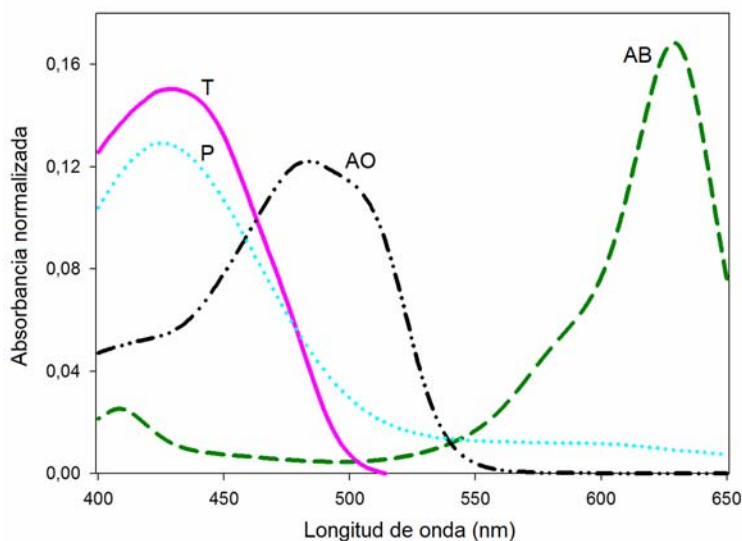


Figura 3.2. Espectros de absorción normalizados. T: tartrazina, AB: azul brillante, AO: amarillo ocaso y P: producto de oxidación de AB.

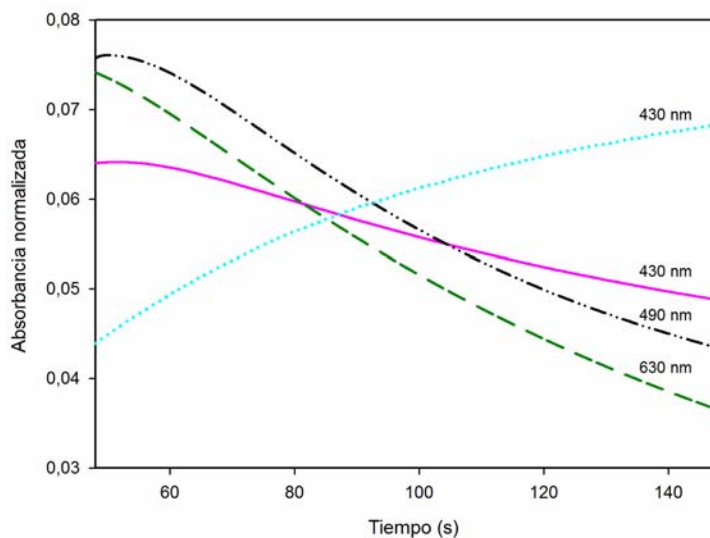


Figura 3.3. Curvas cinéticas normalizadas obtenidas para T (línea sólida rosa), AO (línea discontinua con puntos negra), AB (línea discontinua verde) y su producto de degradación (línea de puntos celeste) en las longitudes de onda de sus máximos de absorción.

3.5.2. Cuantificación de tartrazina en las muestras de validación J1 y J2.

Como se mencionó anteriormente, se seleccionaron regiones en ambas dimensiones (longitud de onda y tiempo) antes de construir los modelos de calibración. Específicamente, el tiempo se restringió entre 46 y 132 s y las longitudes de onda entre 400 y 500 nm, generando matrices de 216×101 por cada muestra analizada. Las matrices correspondientes a las muestras de los juegos de validación se apilaron en columnas junto con las matrices de las muestras de calibración para generar un arreglo bidimensional que sería analizado con MCR-ALS.

Antes de empezar la resolución, se determinó el número de contribuciones presentes en cada matriz de datos aplicando SVD, y se estableció la presencia de tres componentes para el conjunto J1 (correspondientes a las señales de T, AB y su producto de reacción) y cuatro para el J2 (los mismos que J1 más la señal de AO). Las estimaciones iniciales se obtuvieron por selección de las variables más puras basadas en SIMPLISMA (Windig y Guilment 1991).

En este capítulo, se aplicó la restricción de no negatividad en espectros y perfiles de concentración para todos los componentes. La correspondencia entre especies también se empleó para indicar la presencia o ausencia de tres o cuatro

componentes en las muestras de calibración y los juegos de validación J1 y J2, respectivamente.

El algoritmo de calibración multivariada MCR-ALS se describió ampliamente en el Capítulo 1 de la presente Tesis. En esta oportunidad sólo se comentará acerca del significado de cada parámetro perteneciente a la Ecuación 1.9 de la Sección 1.2.1.2 de ese apartado: las filas de la matriz **D** contienen los espectros medidos para las diferentes muestras a varios tiempos, las columnas de **C** contienen los perfiles cinéticos de las especies intervinientes, las columnas de **S** sus espectros y **E** es la matriz de los residuos no ajustados por el modelo.

En un primer paso, MCR-ALS se aplicó para determinar T en J1 aumentando las matrices en ambos modos, temporal y espectral. En el caso de aumento en el tiempo, además de las restricciones mencionadas anteriormente, se utilizó la de unimodalidad en los perfiles de concentración.

Debido a que las áreas bajo los perfiles de concentración obtenidas por MCR-ALS para cada muestra son proporcionales a la concentración de analito, las correspondientes a las muestras de calibración se utilizaron para construir un gráfico univariado. Ya que el comportamiento no lineal también se encontró en las áreas obtenidas, se modelaron con un polinomio de grado dos que luego se utilizó para predecir la concentración de T, x_T :

$$y_T = a + b x_T + c x_T^2 \quad (3.1)$$

donde a , b y c son los parámetros de la regresión e y_T es el área bajo el perfil de concentración del analito. El ajuste por cuadrados mínimos arrojó $a = 4 (4) \times 10^2$, $b = 148 (9) \times 10^1$ y $c = -17 (4)$ (entre paréntesis la desviación estándar en la última cifra significativa) y un coeficiente de correlación $r^2 = 0.9942$ para el caso de aumento en el sentido temporal. En el caso de aumento en el modo espectral, estos valores fueron $a = 13 (3) \times 10^2$, $b = 112 (7) \times 10^1$ y $c = -17 (3)$ y un coeficiente de correlación $r^2 = 0.9924$.

Los resultados de las predicciones se presentan en la Tabla 3.3. Como puede apreciarse, se obtuvieron mejores resultados cuando se aplicó el aumento de matrices en el modo temporal, con un error de predicción del 9.0 %. Por el contrario, se obtuvo un error de predicción del 14.7 % cuando se analizaron las

muestras aumentando las matrices en el modo espectral. Este hecho se podría atribuir a las diferencias en el grado de solapamiento espectral y cinético entre las especies intervinientes.

Tabla 3.3. Resultados de la predicción del juego de validación J1 al aplicar MCR-ALS.

Muestra	Concentración de tartrazina (mg L ⁻¹)		
	Nominal	Modo temporal ^{a,b}	Modo espectral ^{a,b}
1	11.09	11.6 (0.2)	10.6 (1.7)
2	16.87	15.0 (0.1)	14.8 (0.1)
3	19.28	17.2 (0.1)	16.8 (0.1)
4	4.99	5.2 (0.3)	6.2 (0.2)
5	2.50	2.5 (0.1)	3.6 (0.1)
6	11.09	9.9 (0.3)	10.3 (0.2)
7	11.09	11.0 (0.1)	10.90 (0.01)
8	16.87	16.97 (0.01)	14.58 (0.01)
9	4.99	5.3 (0.4)	5.3 (0.3)
Rec. (%) ^b	-	98.1 (7.0)	102.3 (19.1)
REP (%) ^c	-	9.0	14.7

^a Promedio de duplicados.

^b Recuperación promedio. Desviación estándar entre paréntesis.

^c Error relativo de predicción, $REP = \frac{100}{\bar{c}} \left[\frac{1}{I} \sum_1^I (c_{nom} - c_{pred})^2 \right]^{1/2}$, donde I es el número de

muestras, c_{nom} and c_{pred} son las concentraciones nominal y predicha, y \bar{c} es la concentración promedio.

Para calcular el grado de solapamiento se empleó la siguiente expresión:

$$S_{12} = \frac{\|\mathbf{s}_1^T \mathbf{s}_2\|}{\|\mathbf{s}_1\| \|\mathbf{s}_2\|} \quad (3.2)$$

donde \mathbf{s}_1 y \mathbf{s}_2 son los espectros del componente 1 y 2, respectivamente. El valor de S_{12} varía entre cero y uno, correspondiendo a situaciones extremas de no

solapamiento y solapamiento completo, respectivamente. Cuando la Ecuación (3.2) se aplicó para determinar el grado de solapamiento cinético entre los componentes, los espectros se reemplazaron por los correspondientes perfiles de tiempo.

El grado de solapamiento espectral entre T y el producto de oxidación de AB es 0.9961, mientras que el grado de solapamiento cinético es 0.9779. A primera vista, estos valores muestran un mayor solapamiento en la dimensión espectral indicando que un aumento en el sentido de los espectros sería la mejor opción, pero esto es estrictamente cierto cuando no existe falta de sincronización en los perfiles temporales. La Figura 3.4 muestra las curvas cinéticas de T correspondientes a una muestra de calibración y una muestra de cada juego de validación (J1 y J2) obtenidos al aplicar MCR-ALS. Es evidente la diferencia que existe, demostrando que el aumento en el sentido temporal sería la opción correcta.

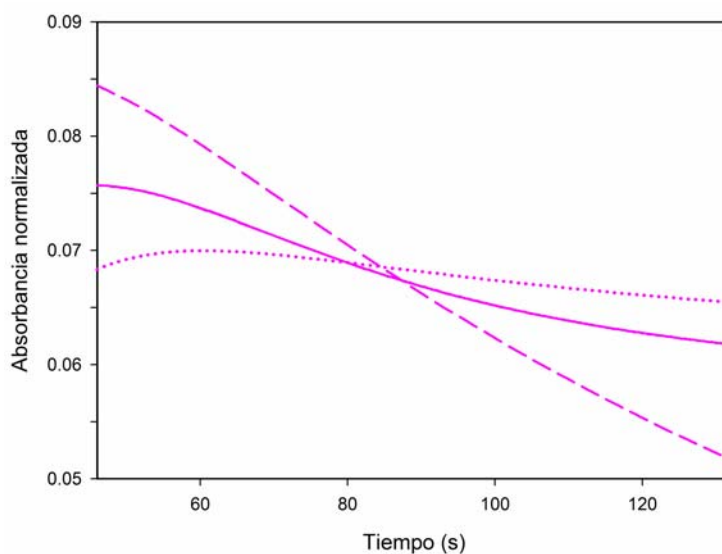


Figura 3.4. Perfiles temporales normalizados de T obtenidos al aplicar MCR-ALS para analizar una muestra de calibración (línea sólida), una muestra del juego J1 (línea discontinua) y una muestra del juego J2 (línea de puntos).

Además, los grados de solapamiento espectral y cinético entre T y AB, calculados en el rango de longitudes de onda analizado, resultaron 0.9017 y 0.9921, respectivamente, mostrando un comportamiento opuesto. Adicionalmente, teniendo en cuenta el colorante AO, este comportamiento se repite y las curvas

cinéticas de T y este colorante se solapan en mayor grado que sus espectros. Además, una inspección la Figura 3.5A muestra que de las cuatro curvas cinéticas obtenidas al aplicar MCR-ALS a una muestra de J2, tres decrecen y sólo una aumenta. Por otro lado, en la Figura 3.5B se observa una mayor diferencia entre los espectros de los cuatro componentes. Por las razones mencionadas, las muestras del juego J2 sólo se analizaron apilando sus matrices en el modo temporal.

Las Figuras 3.5 A y B muestran los perfiles cinéticos y espectrales arrojados por MCR-ALS cuando se procesó una muestra del juego J2 junto con algunas muestras de calibración. En la Figura 3.5 A se observa una señal que crece debido a la aparición del producto de oxidación de AB en la muestra de validación. Además, se pueden apreciar otras dos señales con perfiles que decaen debido a la absorción de AB en la región de longitudes de onda seleccionadas y a la presencia de AO, que produce un componente incoloro durante su degradación. Por otro lado, la señal de decaimiento de T puede verse tanto en la muestra de validación como en las de calibración.

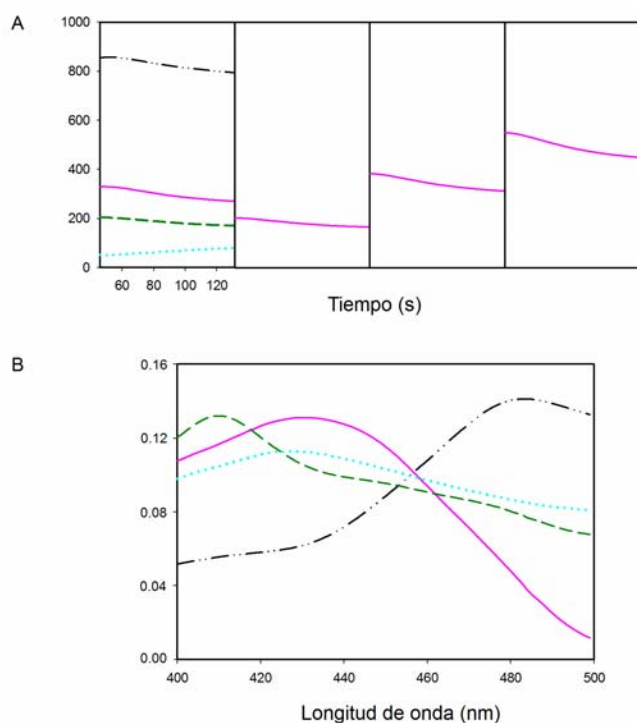


Figura 3.5. Perfiles obtenidos al aplicar MCR-ALS para analizar la muestra 3 del juego J2 junto con muestras de calibración (8, 12 y 16 mg L⁻¹). (A) Perfiles cinéticos. (B) Perfiles espectrales. En ambos casos, la línea sólida rosada indica T, la línea discontinua verde indica AB, la línea punteada celeste indica el producto de oxidación de AB y la línea discontinua con puntos negra indica AO.

La Tabla 3.4 muestra los resultados de predicción de las veinte muestras de validación de J2.

Tabla 3.4. Resultados de la predicción del juego de validación J2 al aplicar MCR-ALS.

Muestra	Concentración de tartrazina / mg L ⁻¹	
	Nominal	MCR-ALS ^{a,b}
1	10.99	9.5 (0.8)
2	10.99	10.34 (0.01)
3	10.99	10.89 (0.01)
4	16.00	14.9 (0.1)
5	10.99	11.3 (0.7)
6	16.00	15.8 (1.2)
7	10.99	10.0 (0.6)
8	10.99	11.86 (0.01)
9	10.99	10.6 (0.1)
10	5.98	4.7 (0.7)
11	10.99	11.5 (0.1)
12	10.99	10.1 (0.4)
13	10.99	11.6 (0.3)
14	5.98	6.9 (0.4)
15	19.47	19.40 (0.01)
16	2.51	2.7 (0.1)
17	5.98	6.3 (0.1)
18	5.98	7.8 (0.3)
19	16.00	17.2 (0.1)
20	16.00	13.85 (0.01)
Rec. (%) ^b	-	99.9 (11.2)
REP (%) ^c	-	6.5

^a Promedio de duplicados.

^b Recuperación promedio. Desviación estándar entre paréntesis.

^c Error relativo de predicción, $REP = \frac{100}{\bar{c}} \left[\frac{1}{I} \sum_1^I (c_{\text{nom}} - c_{\text{pred}})^2 \right]^{1/2}$, donde I es el número de

muestras, c_{nom} and c_{pred} son las concentraciones nominal y predicha, y \bar{c} es la concentración promedio.

Los resultados de predicción de la concentración de T en ambos juegos de validación (Tabla 3.3 y 3.4) pueden considerarse aceptables, logrando una recuperación promedio del 98.1 (7.0) % y 99.9 (11.2) %, y un error relativo de predicción (REP) de 9.0 % y 6.5 % basados en la concentración promedio de calibración para J1 y J2, respectivamente, al realizar un aumento de matrices en el sentido temporal. Estos parámetros indican que el método propuesto es apto para la obtención de la ventaja de segundo orden en casos en donde los componentes de una muestra poseen espectros similares y existe una relación no lineal entre la señal y la concentración.

Debido a que se demostró que U-PCA/RBL es lo suficientemente flexible como para modelar datos de segundo orden no lineales (García-Reiriz y col. 2008), y considerando la no linealidad manifestada anteriormente, trabajamos bajo la hipótesis de que este algoritmo podría ser una herramienta útil para analizar los datos obtenidos con el método cinético.

Con este propósito, las matrices se desdoblaron en vectores columnas de 21715×1 . Los *scores* obtenidos para las muestras de calibración se emplearon para construir una función polinomial (ver Ecuación (3.1)), pero en este caso y_T se refiere a los *scores* de T. Los coeficientes de regresión fueron $a=4(2) \times 10^2$, $b=61(3) \times 10^1$ y $c=-8(2)$, y el coeficiente de correlación $r^2=0.9938$.

Las muestras de validación del juego J1 se sometieron a bilinealización residual para eliminar la contribución de dos componentes no calibrados y a análisis de componentes principales para obtener los *scores*. Cuando se aplicó U-PCA/RBL al juego J2 para adquirir los *scores* libres de interferentes, se encontró sólo un componente no calibrado.

Los errores de predicción fueron mayores a los obtenidos por MCR-ALS, obteniéndose un REP de 14.6 % y 16.3 % cuando se predijeron las muestras de J1 y J2, respectivamente. Esto puede deberse a la falla observada en la extracción de un perfil de interferente apropiado, lo que puede ser responsable de la poca habilidad de U-PCA/RBL para modelar los datos en estudio. Como se mencionó recientemente, RBL puede no ser una herramienta útil para obtener la ventaja de segundo orden cuando existe solapamiento severo (García-Reiriz y col. 2007, García-Reiriz y col. 2008). Consecuentemente, la cantidad de T en presencia de AB

como interferente no puede ser cuantificada con exactitud siguiendo esta estrategia.

3.5.3. Predicción de tartrazina en bebidas comerciales

MCR-ALS se aplicó para analizar muestras de bebidas que contienen T en presencia de otros colorantes (AO y AB). Las muestras también se analizaron por el método CLAR para validar los resultados obtenidos con el método presentado en este capítulo (ver Tabla 3.5).

Tabla 3.5. Resultados de la determinación de tartrazina en bebidas comerciales.

Muestra ^a	Sabor	Colorantes declarados ^b	Tartrazina (mg L ⁻¹)	
			CLAR	MCR-ALS
LC1	Naranja-Durazno	E-102, E-110	4.07 (0.06)	4.2 (0.1)
LC2	Green Mango	E-102, E-133	6.58 (0.01)	6.7 (0.2)
LC3	Naranja-Durazno	E-102, E-110	0.79 (0.01)	0.7 (0.1)
BP	Manzana	E-102, E-110	21.55 (0.01)	21.1 (0.1)
BA	Green Evolution	E-102, E-133	7.73 (0.01)	7.8 (0.9)
$t_{\text{calculado}}^c$			0.56	
Rec. (%) ^d			-	98.8 (4.2)
REP(%) ^e			-	4.2

^a LC1-3: bebidas listas para consumo, BP: bebida en polvo, BA: bebida alcohólica.

^b E-102 (tartrazina), E-110 (amarillo ocaso), E-133 (azul brillante).

^c Valor de t (calculado según Miller y col. (2005)) es menor al valor crítico 2.78 (95% nivel de confianza y 4 grados de libertad).

^d Recuperación promedio. Desviación estándar entre paréntesis.

^e Error relativo de predicción, $REP = \frac{100}{\bar{c}} \left[\frac{1}{I} \sum_{i=1}^I (c_{\text{nom}} - c_{\text{pred}})^2 \right]^{1/2}$, donde I es el número

de muestras, c_{nom} and c_{pred} son las concentraciones nominal y predicha, y \bar{c} es la concentración promedio.

Un cromatograma típico de una de las muestras analizadas que contiene T y AO revela que la separación completa se lleva a cabo en 6.0 minutos (ver Figura

3.6). Además, las muestras que contienen T y AB necesitaron cerca de 10.0 minutos para que todos los componentes sean separados. Por otro lado, sólo 2.5 minutos son necesarios para cuantificar T manera exacta a través de la aplicación del método propuesto.

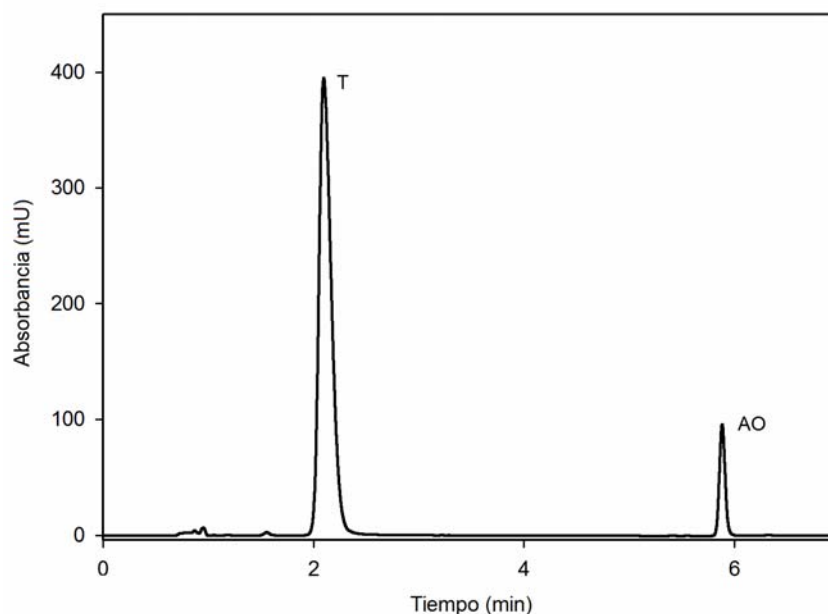


Figura 3.6. Separación cromatográfica de tartrazina (T) y amarillo ocazo (AO) en una muestra real (muestra BP de la Tabla 3.5) al utilizar el método CLAR ($\lambda = 454$ nm).

Para decidir si los resultados de la metodología desarrollada y el método CLAR son comparables, se llevó a cabo un test *t* basado en la prueba estadística de datos pareados (Miller y Miller 2005). Como se puede apreciar en la Tabla 3.5, el valor de *t* calculado (0.56) es menor que el valor de tabla (2.78 con 95% de nivel de confianza y 4 grados de libertad), lo que indica que no hay diferencias estadísticamente significativas entre el método de referencia y la estrategia descrita en este capítulo.

3.6. Conclusiones

La combinación de FIA con la reacción cinética con bromato de potasio, catalizada por Fe(II), que permite generar datos de segundo orden con comportamiento no lineal, pudo ser implementada con éxito para la determinación

de T en presencia de otros colorantes y componentes inesperados de la matriz de las bebidas analizadas.

El método que en principio se pensaba que iba a tener mejor rendimiento (U-PCA/RBL), no pudo modelar la existencia de colinealidades entre los espectros del analito y las interferencias, mostrando una limitación. Por el contrario, se demostró la habilidad de MCR-ALS para modelar datos de segundo orden no lineales, el que arrojó perfiles cuyas áreas pudieron ser ajustadas con una función polinomial, obteniéndose cifras de mérito satisfactorias. Cuando este modelo se aplicó al análisis de muestras reales, los resultados fueron aceptables y estadísticamente comparables con aquellos obtenidos por CLAR. Además, resulta interesante destacar que la T se encontró en las muestras a niveles de concentración cinco veces menores al máximo valor establecido por la legislación Argentina. Si se necesitaran analizar futuras muestras con concentraciones mayores, una simple dilución sería suficiente.

Por todo lo expuesto, se puede concluir que el método propuesto provee una alternativa rápida, exacta y económica a las técnicas separativas.



CAPÍTULO 4

Obtención de la ventaja de segundo orden a través del modelado de matrices de excitación-emisión afectadas por efecto de filtro interno, utilizando transferencia de calibración: reducción del número de experimentos y aumento de la sensibilidad analítica

4.1. Introducción

La determinación de fármacos en muestras complejas requiere de la aplicación de diferentes estrategias que permitan obtener selectividad frente a la presencia de componentes inesperados. El problema es aún mayor cuando, además, se produce un efecto matriz que origina cambios en la respuesta analítica. En estos casos, la curva de calibrado construida con muestras de estándar puro llevará a una determinación incorrecta del analito de interés (Danzer y Currie 1998).

El efecto matriz puede ser originado, entre otras causas, por un problema inherente a muchos procedimientos fluorimétricos: la absorción de radiación de excitación y/o emisión por especies disueltas (fluorescentes o no) o por el fluoróforo en sí mismo (Lakowicz 1983). Este proceso es conocido como efecto de filtro interno, y genera una variación no sólo en la intensidad, sino también en la forma del espectro. Ambos inconvenientes pueden ser corregidos por el método de adición estándar (SAM), que compensa el efecto matriz que surge por la composición específica de la muestra. Sin embargo, el mayor inconveniente ligado a esta metodología es el gran número de experimentos que se necesitan, ya que se deben construir nuevas curvas de calibrado cuando se analizan nuevas muestras. Además, para obtener cifras de mérito satisfactorias se recomienda la adición de al menos cuatro concentraciones diferentes de estándar a la muestra (Danzer y Currie 1998, Lakowicz 1983, Castells y Castillo 2000). Por otro lado, si se quieren determinar varios analitos simultáneamente en una mezcla de multicomponentes, se debe recurrir al método de adición estándar generalizado (GSAM) que permite corregir el efecto matriz y la presencia de interferencias, pero que presenta la desventaja de ser lento y laborioso ya que requiere herramientas matemáticas más elaboradas que SAM (Saxberg y Kowalski 1979, Nepote y Olivieri 2001). En este contexto, se han desarrollado varias aplicaciones interesantes, como por ejemplo, la cuantificación espectrofluorimétrica simultánea de oxatomida y fenilefrina en presencia de paracetamol en un preparado farmacéutico (Nepote y Olivieri 2001), la aplicación de GSAM en un sistema de inyección de flujo para determinar K, Ca y Na por fotometría de llama (Silva y col. 1999) y la determinación de cobre en bebidas alcohólicas de caña de azúcar (Honorato y col. 2002, Silva y col. 1996). Recientemente, se desarrolló un método que combina GSAM con Funciones de

Base Radial-Cuadrados Mínimos Parciales (RBF-PLS) para la determinación de varios analitos en presencia de efecto matriz, solapamiento espectral y no linealidades (Shariati-Rad y col. 2013a, Shariati-Rad y col. 2013b).

Como se mencionó anteriormente, la composición de las muestras reales, por lo general, incluyen componentes no esperados. Estas mezclas de multicomponentes pueden ser resueltas con técnicas de calibración multivariada como, por ejemplo, Cuadrados Mínimos Parciales (PLS-1). Sin embargo, si se desea obtener precisión y exactitud, los algoritmos de primer orden necesitan modelar en el paso de calibración no sólo los componentes que presentan respuesta instrumental, sino también el efecto matriz, para lo cual es necesaria la preparación de numerosas muestras (Navarro-Villoslada y col. 1999).

La cuantificación de un analito o varios en presencia de potenciales interferentes se puede realizar midiendo señales de segundo orden y procesándolas con algoritmos apropiados para obtener la ventaja de segundo orden (Escandar y col. 2007). Para resolver ambos inconvenientes: presencia de compuestos desconocidos y efecto matriz, se propuso la combinación de estos algoritmos y SAM. En la literatura se encuentran publicadas varias aplicaciones del algoritmo PARAFAC cuando la sensibilidad de la respuesta depende de la composición de la matriz (Cañada-Cañada y col. 2009, Rodríguez y col. 2009a, Rodríguez y col. 2009b, Valderrama y Poppi 2011). Otros algoritmos de calibración de segundo orden utilizados en este contexto son MCR-ALS (Peré-Trepat y col. 2005, Trepat y col. 2007, Lozano y col. 2009a), SWATLD (Zhu y col. 2008) y N-PLS/RBL (Lozano y col. 2009b). Sin embargo, en todos los casos, se necesita un gran número de muestras para lograr cifras de mérito aceptables.

Recientemente, se demostró la habilidad de U-PLS/RBL para analizar muestras complejas de especies fluorescentes que presentan efecto de filtro interno (Piccirilli y Escandar 2006, Bohoyo y col. 2006). Es interesante mencionar que se redujo el tiempo de análisis del procedimiento SAM con el modelado del cambio en los espectros a través de la flexibilidad de PLS. Sin embargo, en estos casos, tanto el analito como el componente causante del efecto de filtro interno se modelaron en la etapa de calibración. En consecuencia, fue necesario conocer este componente, lo que no siempre es posible para muestras reales.

El principal objetivo de la estrategia presentada en este capítulo es reducir el número de muestras a preparar cuando se presentan interferentes de manera conjunta con efecto de filtro interno, sin la necesidad de modelar la presencia del agente causante de dicho efecto.

Para resolver el primer inconveniente se aplicó el algoritmo PARAFAC para obtener la señal del analito libre de aquella generada por componentes que no se incluyeron en la calibración. Con el fin de solucionar el efecto de filtro interno realizando el mínimo número de experimentos, se implementó un enfoque novedoso: modelado de segundo orden acoplado a transferencia de calibración. Las técnicas de estandarización se usan frecuentemente para transferir un modelo de calibración creado en un instrumento a una nueva situación (por ejemplo, un instrumento diferente, condiciones alteradas, etc.), y evitar una recalibración completa (Feudale y col. 2002). En la literatura existen varios métodos de transferencia de calibración que tratan de hacer frente a las variaciones no calibradas que se detectan luego de que el modelo ya está en uso (por ejemplo, cambio del instrumento o de una de sus partes) (Siano y Goicoechea 2007, Walczak y col. 1997, Wang y col. 1991). Un desarrollo importante en la transferencia multivariante es la Estandarización Directa Sucesiva (PDS) (Wang y col. 1991), que reconstruye cada punto del espectro medido en el primer instrumento con varias mediciones en una ventana pequeña en el segundo instrumento. Se han publicado algunos trabajos con respecto a la aplicación de PDS en varios contextos, como clasificación (Di Anibal y col. 2012), corrección de variaciones espectrales inducidas por temperatura (Wülfert y col. 2000) y transferencia de calibración para datos de tres vías (Thygesen y van den Berg 2011). Una búsqueda bibliográfica revela que PDS puede ser considerado un método de rutina para transferencia de calibración.

Existen informes de la combinación de PDS con modelado de segundo orden para aliviar el trabajo experimental (Gil García y col. 2008, Gil García y col. 2009, Vosough y col. 2010). Estas publicaciones corresponden a la aplicación de CLAR con DAD para la determinación de contaminantes emergentes en muestras de agua. El juego de transferencia de calibración se obtuvo a través de la estandarización de los datos de calibración preparados en solvente, y se utilizó en el paso de predicción de muestras de agua, en las que era evidente la presencia de

efecto matriz debido al paso de extracción en fase sólida. Como resultado, los perfiles cromatográficos obtenidos por inyección directa pudieron ser transformados en cromatogramas en una situación real mediante la utilización del vector de transformación, con coeficientes de determinación elevados entre las señales antes y después de la transformación ($r^2 \geq 0.99$). En este capítulo se pondrá énfasis en la demostración estadística completa de la aplicabilidad y robustez de la metodología propuesta. Con este objetivo, se crearon juegos de datos de matrices de excitación-emisión (MEE) simulados y experimentales para evaluar el rendimiento de los métodos. Además, el sistema experimental se utilizó para determinar fenilefrina clorhidrato (FEN) en muestras de agua en presencia de ibuprofeno (IBU), ácido acetilsalicílico (AAS) y paracetamol (PAR). FEN se determinó previamente en muestras farmacéuticas por CLAR (Erka y Kartal 1998, Martín y col. 2002), análisis con flujo detenido (FIA) (Knochen y Giglio 2004), espectrofotometría (Mahgoub 1990), electroforesis capilar (Gómez y col. 2002), espectrofluorometría (Arancibia y col. 2000) y calibración multivariada utilizando PLS-1 (Collado y col. 2000). Se demostró que la intensidad de emisión de FEN decrece notablemente en presencia de PAR, ya que éste absorbe en la región de máxima excitación de FEN (Nepote y Olivieri 2001, Arancibia y col. 2000). Por esta razón, se desarrolló la siguiente estrategia quimiométrica: en primer lugar, se aplica PDS a las matrices desdobladas para transferir la señal de la muestra a la situación de la curva de calibrado usando sólo dos muestras en cada situación. Luego, los datos transformados se reordenan en matrices y se someten a PARAFAC para obtener *scores* útiles para la predicción (PDS/PARAFAC). Una ventaja adicional (además de que sólo se usan dos adiciones de estándar a la muestra) es que la sensibilidad de la curva de calibrado se mantiene (la curva con SAM tiene menor pendiente), y la señal de la muestra es transferida para predecir su concentración utilizando la curva de estándares puros. Los resultados se compararon con aquellos obtenidos luego de aplicar el método de adición estándar combinado con PARAFAC (SAM/PARAFAC), preparando cinco adiciones a la muestra por triplicado.

4.2. Objetivos

- Reducir el número de soluciones a preparar cuando la muestra presenta efecto de filtro interno causado por un componente desconocido presente en la matriz, además de otros componentes no calibrados.
- Evaluar el rendimiento de PDS en la transferencia de la señal de segundo orden de una muestra que presenta efecto de filtro interno y otros componentes no calibrados.
- Evaluar la aplicación de PARAFAC en matrices de excitación-emisión para determinar fenilefrina utilizando una curva de calibrado en agua y la señal de la muestra transferida.

4.3. Programas

Para la creación de datos simulados y para el procesamiento de los datos, todas las rutinas se llevaron a cabo utilizando MATLAB 7.1 (Matlab 7.1, 2005). La aplicación de PDS se realizó a través de una rutina incluida en el PLS_Toolbox 3.5 (Wise y col. 2005) y se utilizó la interfaz gráfica MVC2 desarrollada por Olivieri y col. (Olivieri y col. 2009) para la aplicación de PARAFAC.

4.4. Materiales y Métodos

4.4.1. Reactivos y soluciones

Las drogas fenilefrina clorhidrato, ibuprofeno, ácido acetilsalicílico y paracetamol se obtuvieron del Laboratorio de Control de Calidad de Medicamentos (Facultad de Bioquímica y Cs. Biológicas, Universidad Nacional de Litoral). Se prepararon soluciones acuosas de FEN (100 mg L⁻¹), AAS (100 mg L⁻¹) y PAR (200 mg L⁻¹). Además, se preparó una solución de IBU (100 mg L⁻¹) en metanol grado HPLC (Aberkon, Buenos Aires, Argentina). A continuación, una alícuota de 1.00 mL de esta solución se colocó en un matraz de 10.00 mL, el solvente se evaporó con una corriente de nitrógeno y se enrasó el matraz con agua MilliQ.

El agua MilliQ se obtuvo utilizando un sistema purificador Millipore (Bedford, MA, USA).

4.4.2. Equipamiento

Las mediciones fluorimétricas se llevaron a cabo utilizando un espectrómetro Perkin-Elmer LS-55 equipado con una lámpara de descarga de xenón, monocromadores de tipo Monk-Gillieson y un fotomultiplicador conectado a una PC vía una conexión RS232C y usando cubetas de cuarzo de 1.00 cm de paso óptico.

Las matrices de excitación-emisión se registraron variando las longitudes de onda de excitación entre 215 y 240 nm cada 2 nm, y recogiendo los espectros de emisión entre 270 y 360 nm cada 0.5 nm. Por lo tanto, el tamaño de cada matriz fue de 181×13 . Los anchos de banda de las aberturas de ambos monocromadores se fijaron en 10 nm y el voltaje del detector en 650 V.

4.4.3 Datos

4.4.3.1. Datos simulados

Con el objetivo de analizar el desempeño del presente método, se utilizaron datos de excitación-emisión simulados. Se construyó un juego de datos de calibración de segundo orden con un único analito en concentraciones nominales entre 1.0 y 6.0 (unidades arbitrarias) cada 1 unidad, por triplicado. Además, se creó un juego de datos de validación de 51 muestras en las que se incluyó un componente no calibrado junto con el analito, éste último en concentraciones nominales entre 1.0 y 6.0 cada 0.1. El compuesto no calibrado no sólo simula un interferente, sino que, además, actúa como filtro interno, disminuyendo la intensidad de fluorescencia de la señal original en todos los sensores, y modifica el espectro de emisión del analito, que varía entre cada muestra de validación. Asimismo, se realizaron seis adiciones de analito entre 1.0 y 6.0 a cada muestra de validación por triplicado.

La dimensión de cada matriz fue de 31×31 , que intenta imitar las longitudes de onda de emisión y excitación, respectivamente. El nivel de ruido agregado a los datos fue de 1% y 5% en señal y concentración, respectivamente. La Figura 4.1 muestra los espectros de excitación y emisión de los componentes 1 y 2 e ilustra cómo el perfil del analito varía en una muestra de validación en ambas dimensiones.

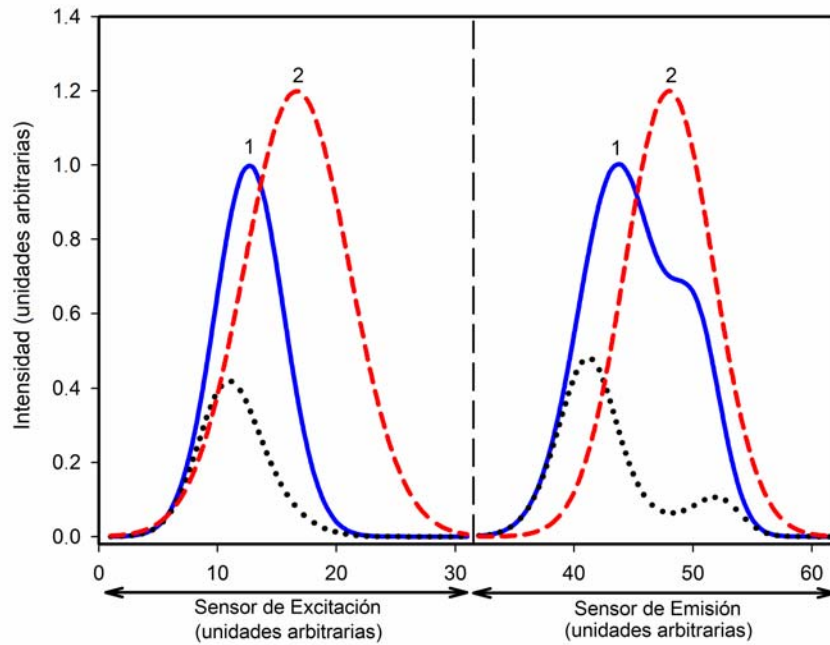


Figura 4.1. Perfiles de excitación y emisión simulados para el analito puro (componente 1, línea sólida azul) e interferente (componente 2, línea discontinua roja), y el analito en una muestra de validación con efecto de filtro interno producido por el componente 2 en ambas dimensiones (línea de puntos negra).

En ausencia de filtro interno, las señales de calibración se pueden computar como:

$$\mathbf{X}_{c,i} = y_{1,c,i} \mathbf{S}_1 + \mathbf{R} s_X \quad (4.1)$$

donde $\mathbf{X}_{c,i}$ es la matriz de $J \times K$ para la i th muestra de calibración, $y_{n,c,i}$ es la concentración nominal de cada analito, y $\mathbf{S}_n = k_n \mathbf{b}_n \mathbf{c}_n^T$ son las matrices de datos correspondientes a concentración unitaria para el analito n (\mathbf{b}_n y \mathbf{c}_n son los perfiles en la primera y segunda dimensión, ambos normalizados, y k_n es un factor de escalamiento), \mathbf{R} es la matriz de números al azar con distribución Gaussiana con desviaciones estándares de tamaño apropiado, y s_X es la desviación estándar del ruido agregado a las señales.

Las señales de las muestras de validación se construyeron según la siguiente expresión:

$$\mathbf{X}_u = y_{1,u} \mathbf{S}_1 + y_{2,u} \mathbf{S}_2 + \mathbf{R} s_X \quad (4.2)$$

donde \mathbf{X}_u es la matriz de $J \times K$ de la muestra incógnita, $y_{n,u}$ es la concentración nominal de cada componente y \mathbf{S}_2 es la matriz de señales del componente no calibrado.

Cuando ocurre el efecto de filtro interno causado por el componente 2 sobre el componente 1, \mathbf{S}_1 es reemplazado en la Ecuación (4.2) por la matriz $\mathbf{S}_{1,fi}$ cuyo elemento genérico (j,k) está dado por la siguiente expresión:

$$\mathbf{S}_{1,fi}(j,k) = \mathbf{S}_1(j,k) \times \exp[-(\varepsilon_{2j} + \varepsilon_{2k})y_{2,u}] \quad (4.3)$$

donde ε_{2j} y ε_{2k} son las absorptividades del componente 2 en los sensores j y k , respectivamente, en cada dimensión. El producto $(\varepsilon_{2j} y_{2,u})$ representa el efecto de filtro interno producido cuando el componente 2 absorbe la intensidad de excitación en el sensor j , mientras que $(\varepsilon_{2k} y_{2,u})$ corresponde a la absorción de la intensidad de emisión en el sensor k . Las absorptividades están dadas por el producto de un factor de escala ajustable y los valores del perfil de excitación \mathbf{c}_2 a cada uno de ambos sensores.

4.4.3.2. Datos experimentales

- *Recolección y preparación de la muestra.* Se recolectaron muestras de agua de canilla y de napa de la ciudad de Santa Fe (Argentina). Antes de ser analizadas se centrifugaron a $5000 \times g$, se filtraron a través de un filtro de $0.45 \mu\text{m}$ y se almacenaron a 4°C en heladera.

- *Preparación de muestras de calibración y validación y muestras fortificadas.* Se preparó una curva de calibrado de cinco estándares de FEN en el rango de 250 a 750 ng mL^{-1} cada 125 ng mL^{-1} , por triplicado. Además, se preparó la misma curva de calibrado pero en presencia de PAR ($10.0 \mu\text{g mL}^{-1}$) para comparar las pendientes de ambas regresiones.

Se preparó un juego de muestras de validación en agua MilliQ siguiendo un diseño central compuesto factorial de cuatro factores y en cinco niveles, con concentraciones diferentes de FEN (entre 248.0 y 400.0 ng mL^{-1}), IBU (entre 251.0

y 400.0 ng mL⁻¹), AAS (entre 50.0 y 79.0 ng mL⁻¹) y PAR (entre 5.0 y 20.0 µg mL⁻¹). La composición de las muestras de validación se presenta en la Tabla 4.1.

Las muestras de agua de canilla y napa, tratadas como se describió anteriormente, se fortificaron con FEN, IBU, AAS y PAR en diferentes niveles de concentración. La composición de estas muestras se explica posteriormente en la Sección “Resultados y Discusión” de este capítulo.

Se realizaron dos adiciones de FEN a cada muestra de validación y muestras de agua (375.0 y 625.0 ng mL⁻¹) para implementar el procedimiento de transferencia. Además, se llevaron a cabo cinco adiciones de FEN (entre 250.0 y 750.0 ng mL⁻¹) por triplicado a tres de las muestras de validación (muestras 2, 12 y 14 de la Tabla 4.1) para poder aplicar el método de adición estándar convencional combinado con PARAFAC.

Tabla 4.1. Composición de las muestras de validación experimentales*.

Muestra	FEN (ng mL ⁻¹) ^a	IBU (ng mL ⁻¹) ^a	AAS (ng mL ⁻¹) ^b	PAR (µg mL ⁻¹) ^c
1	365.0	369.6	54.5	8.04
2	365.0	280.0	74.3	12.5
3	321.2	400.4	64.4	20.0
4	277.4	369.6	54.5	12.5
5	321.2	324.8	79.2	17.0
6	394.2	324.8	64.4	8.04
7	277.4	280.0	54.5	12.5
8	365.0	369.6	74.3	8.04
9	321.2	324.8	64.4	17.0
10	321.2	324.8	64.4	12.5
11	365.0	280.0	54.5	12.5
12	248.2	324.8	64.4	17.0
13	321.2	250.6	64.4	12.5
14	321.2	324.8	49.5	5.0
15	321.2	324.8	64.4	12.5
16	277.4	369.6	74.3	8.04
17	277.4	280.0	74.3	17.0

* Siguiendo un diseño central compuesto factorial de cuatro factores y cinco niveles.

^a Incertidumbre en concentración $s = 7$ ng mL⁻¹ estimada por propagación de errores.

^b Incertidumbre en concentración $s = 6$ ng mL⁻¹ estimada por propagación de errores.

^c Incertidumbre en concentración $s = 0.1$ µg mL⁻¹ estimada por propagación de errores.

Las muestras se prepararon midiendo alícuotas apropiadas de solución estándar de cada droga, colocándolas en matraces volumétricos de 10.00 mL para obtener la concentración deseada y completando hasta el enrase con agua MilliQ (en el caso de las muestras de validación) o con las muestras de agua (en el caso de agua de canilla y napa).

4.5. Resultados y Discusión

4.5.1. Datos simulados

Para ilustrar el desempeño del método propuesto, se construyeron matrices que simulan datos de excitación-emisión con dos componentes, un determinado nivel de ruido y efecto de filtro interno (ver Sección 4.4.3.1).

En este capítulo, se utilizó una combinación de estandarización directa sucesiva (PDS) (Wand y col. 1991) con el algoritmo de segundo orden PARAFAC para transformar los datos de la muestra a la señal correspondiente a la situación de analito puro y para aislar la señal del analito, respectivamente. Esto se llevó a cabo desdoblado las matrices MEE en vectores y sometidas a PDS, como se explicó en la Sección 1.3.2 del Capítulo 1. Una vez que los datos de la muestra fueron transferidos, se rearmó la matriz y se llevó a cabo el método de calibración multivariada PARAFAC (ver Figura 4.2).

Los datos de cada una de las 51 muestras de validación se trataron de la siguiente manera:

1. La matriz de la muestra y de sus dos adiciones (2 y 5 en unidades arbitrarias), y las muestras de calibración se desdoblaron en vectores de 1×961 .
2. El vector de la muestra se restó de los vectores de cada una de las dos adiciones.
3. Se aplicó PDS para obtener **F** comparando las dos adiciones a las que se les restaron la muestra con dos muestras de calibración (con la misma concentración de las adiciones). El tamaño de la ventana fue 3 y la tolerancia fue 0.01.
4. Se utilizó la matriz de transformación **F** para transferir el vector de la muestra.

5. Se rearmó la matriz de la nueva señal de la muestra y se llevó a cabo PARAFAC junto con los estándares de calibración. El número de factores, que se determinó a través del análisis del parámetro conocido como *core consistency* (Bro y Kiers 2003), resultó ser dos.
6. Se utilizaron los *scores* para predecir la concentración del analito en la muestra incógnita.

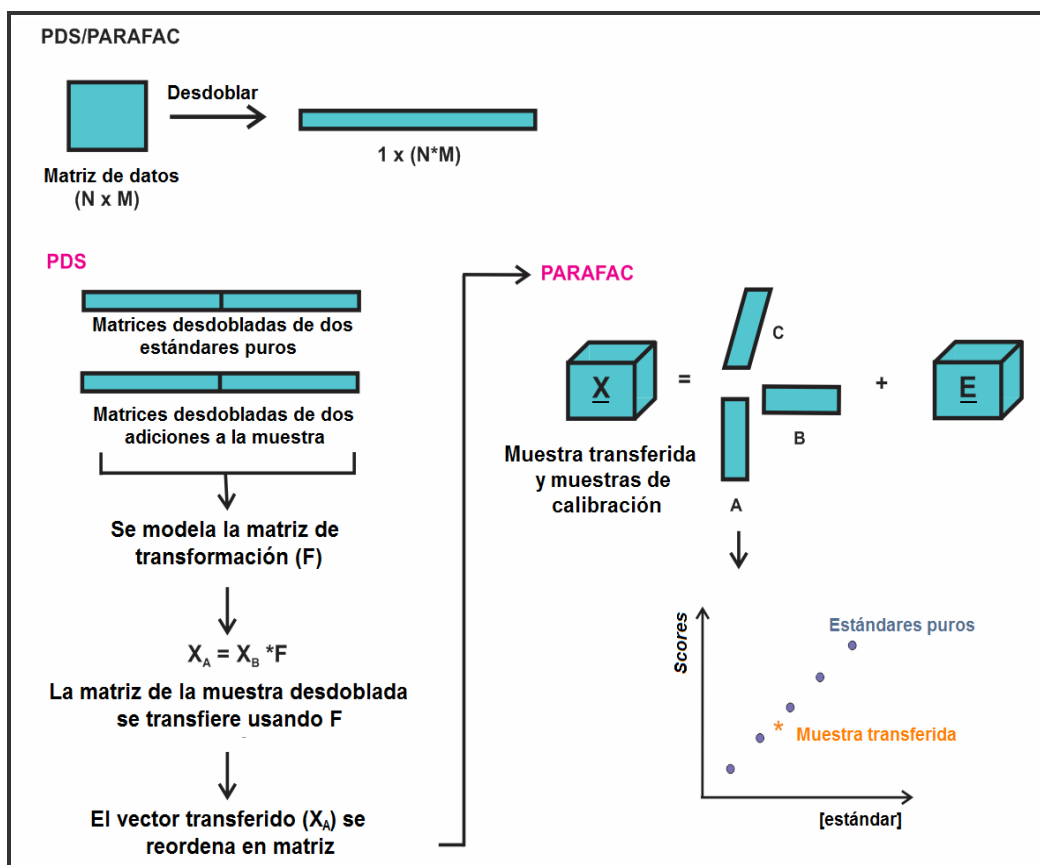


Figura 4.2. Representación esquemática de la aplicación de PDS/PARAFAC a una muestra que contiene componentes inesperados y efecto de filtro interno. Ver texto para los detalles de la simbología empleada.

El impacto de la transformación puede apreciarse en la Figura 4.3, donde se muestran los gráficos de contorno de una muestra de validación simulada antes y después de ser procesada con PDS/PARAFAC, y se comparan con la señal del analito puro. También se aplicó el método SAM/PARAFAC a los datos simulados para comparar los resultados obtenidos.

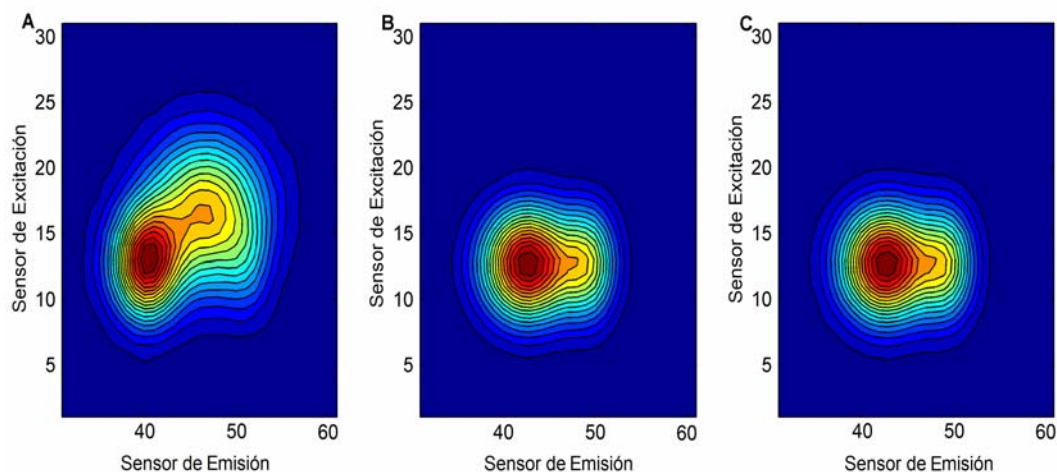


Figura 4.3. Gráficos de contorno para una muestra de validación simulada antes (A) y después (B) de PDS/PARAFAC y de una muestra de calibración simulada que sólo contiene el componente 1 (C).

Los parámetros de predicción de las 51 muestras de validación simuladas están resumidos en la Tabla 4.2. Los resultados fueron satisfactorios, lográndose una recuperación promedio de 100.0% y un error relativo de predicción (REP) de 3.7% cuando se aplicó PDS/PARAFAC. Cuando se aplicó SAM/PARAFAC, los resultados obtenidos fueron de la misma calidad, demostrando que es posible reducir considerablemente el número de experimentos.

La diferencia principal entre ambos métodos puede verse cuando se analiza la sensibilidad analítica (γ) (ver Tabla 4.2). Este valor aumenta más del doble cuando se aplica PDS/PARAFAC. Por lo tanto, su inversa, que significa la menor diferencia de concentración que puede ser medida, cae desde 0.009 a 0.004 (en unidades arbitrarias de concentración), lo que indica que el método desarrollado es más sensible con respecto a SAM/PARAFAC.

Debe realizarse otra observación importante: como se utiliza la matriz de transformación \mathbf{F} (estimada con información del analito con y sin efecto de filtro interno) para transferir la matriz desdoblada de la muestra, uno debería saber si el espectro de la interferencia se modifica durante este paso. Por lo tanto, se calculó el grado de solapamiento espectral (S_{12}) con el fin de comparar los perfiles del interferente obtenido luego de la aplicación de ambas metodologías con su espectro puro, utilizando la Ecuación (3.2) descrita en el Capítulo 3. En este caso \mathbf{s}_1 y \mathbf{s}_2 corresponden al espectro puro del interferente y al perfil obtenido en cada

caso, respectivamente. En la dimensión de excitación, los valores de S_{12} fueron 0.9256 y 0.9999 para PDS/PARAFAC y SAM/PARAFAC, respectivamente, mientras que fueron 0.9038 y 1.000 cuando se analizó la dimensión de emisión por ambos métodos. En el último caso el espectro del interferente obtenido se modifica un poco, pero este hecho no repercute de manera significativa en las cifras de mérito (ver Tabla 4.2).

Tabla 4.2. Resultados y cifras de mérito del análisis de datos simulados.

	PDS/PARAFAC ^a	SAM/PARAFAC ^a
Rec. % ^b	100 (4)	99 (3)
REP % ^c	3.7	2.5
SEN ^d	2.2	1.5
γ^e	221	108
γ^{-1}	0.004	0.009

^a Desviación estándar entre paréntesis.

^b Recuperación promedio.

^c Error relativo de predicción, $REP = \frac{100}{\bar{c}} \left[\frac{1}{I} \sum_1^I (c_{nom} - c_{pred})^2 \right]^{1/2}$, donde I es el número de

muestras, c_{nom} y c_{pred} son las concentraciones nominal y predicha, y \bar{c} es la concentración promedio.

^d Sensibilidad promedio.

^e Sensibilidad analítica promedio.

4.5.2. Datos experimentales

Los datos experimentales se obtuvieron registrando matrices de excitación-emisión para las muestras de calibración y validación comentadas en la Sección 4.4.3.2 de este capítulo. En las Figuras 4.4 A y B se muestran los gráficos en tres dimensiones de las señales de fluorescencia en función de las longitudes de onda de excitación y emisión, en presencia y ausencia de PAR ($10.0 \mu\text{g mL}^{-1}$), respectivamente. Como puede apreciarse en la Figura 4.4B, se produce una disminución significativa en la intensidad de emisión de FEN debido al efecto de filtro interno causado por la presencia de PAR.

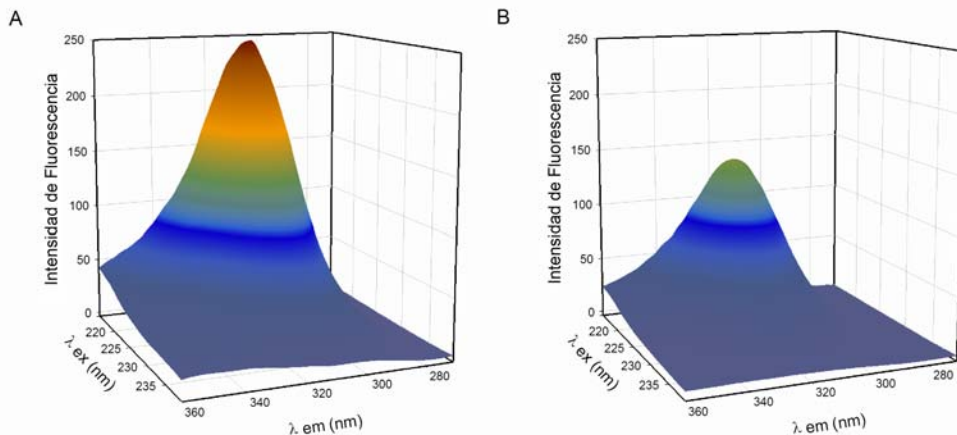


Figura 4.4. Gráficos de tres dimensiones de las señales de fluorescencia en función de las longitudes de onda de excitación y emisión para una estándar de 500 ng mL^{-1} de FEN en ausencia (A) y presencia (B) de PAR ($10.0 \text{ } \mu\text{g mL}^{-1}$).

Este comportamiento también puede observarse en la Figura 4.5, donde se muestran las curvas de calibrado de FEN a $\lambda_{\text{ex}} = 215 \text{ nm}$ y $\lambda_{\text{em}} = 305 \text{ nm}$, en presencia y ausencia de PAR ($10.0 \text{ } \mu\text{g mL}^{-1}$), ya que el cambio en las pendientes es evidente. Por otra parte, en la Figura 4.6 se puede apreciar un pequeño corrimiento en el pico del espectro de emisión de FEN en presencia de PAR. Por lo tanto, la estrategia en estudio se propone con el objetivo de disminuir el trabajo experimental que debe ser llevado a cabo debido al efecto de filtro interno, además de explotar la ventaja de segunda orden, necesaria debido a la presencia de componentes inesperados cuyos espectros están solapados con el correspondiente al analito.

La aplicación de PDS/PARAFAC a los datos experimentales se llevó a cabo siguiendo los pasos descriptos para los datos simulados (Sección 4.5.1 de este capítulo). El número de componentes que se incluyó en el modelo PARAFAC se seleccionó analizando el *core consistency* (Bro y Kiers 2003). En todos los casos, el número de factores fue tres, atribuidos a las señales de fluorescencia de FEN, IBU y AAS. Debido a que PAR no presenta fluorescencia intrínseca en solución acuosa, no es tenido en cuenta como interferente, aunque su presencia origina el efecto de filtro interno. Por otro lado, se aplicó el método SAM/PARAFAC a las muestras 2, 12 y 14 para comparar los resultados. La Tabla 4.3 muestra los resultados de la predicción de las 17 muestras de validación que contienen diferentes mezclas de

FEN, IBU, AAS y PAR (ver Tabla 4.1). Como se puede apreciar, los valores de recuperación promedio son muy satisfactorios y comparables con los obtenidos por SAM/PARAFAC.

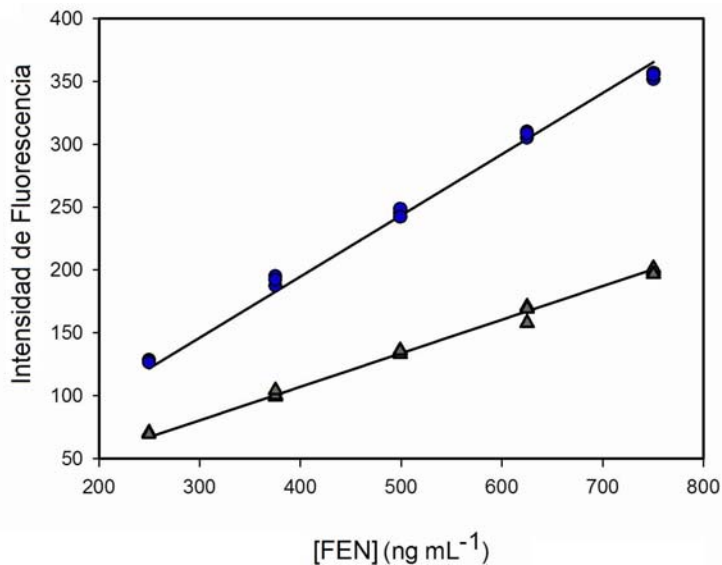


Figura 4.5. Curvas de calibrado de FEN a $\lambda_{ex} = 215$ nm y $\lambda_{em} = 305$ nm, en ausencia (círculos azules) y presencia (triángulos verdes) de PAR ($10.0 \mu\text{g mL}^{-1}$).

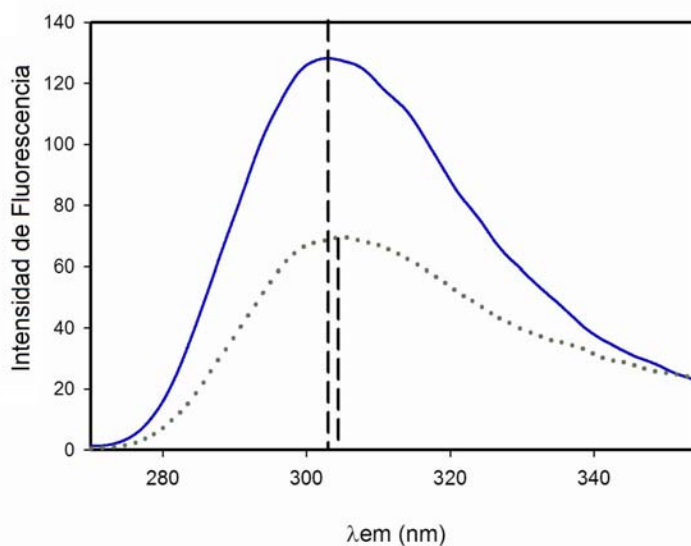


Figura 4.6. Espectro de emisión de un estándar de FEN (250 ng mL^{-1}) sin (línea sólida azul) y con (línea de puntos verde) PAR.

Los errores relativos de predicción (REP) de ambos métodos (REP = 4% para SAM/PARAFAC y REP = 5 % para PDS/PARAFAC) se compararon utilizando la

prueba estadística descrita por H. van der Voet (H. van der Voet 1994). Específicamente, la aplicación de esta prueba-*t* para comparar los REP obtenidos arrojó un nivel de confianza superior a 0.05 (calculado usando 1999 iteraciones). Esta probabilidad indica que el REP para SAM/PARAFAC es estadísticamente igual al REP para PDS/PARAFAC. Por lo tanto, se concluye que la metodología propuesta es comparable al método de adición estándar combinado con PARAFAC.

Resulta importante destacar que se obtienen diferentes valores de pendientes para distintos niveles de PAR. Por ejemplo, el valor de este parámetro es 0.26 mL ng⁻¹ y 0.04 mL ng⁻¹ cuando FEN está en presencia de 10.0 µg mL⁻¹ o 50.0 µg mL⁻¹ de PAR, respectivamente. En consecuencia, se deberían construir diferentes curvas de adición estándar para cada muestra que ingresa al laboratorio, lo que involucra al menos cuatro adiciones de estándar puro en diferentes niveles. Como se mencionó anteriormente, la influencia del compuesto que causa el efecto de filtro interno se logró modelar con U-PLS/RBL por Escandar y colaboradores lográndose resultados excelentes, aunque se debe remarcar que sólo puede lograrse cuando el agente causante del filtro interno es conocido y modelado (Piccirilli y col. 2006, Bohoyo Gil y col. 2006).

En el caso analizado en este capítulo, sólo se deben realizar dos adiciones a la muestra sin réplicas para aplicar PDS/PARAFAC, a diferencia de SAM/PARAFAC, donde se deberían preparar cinco adiciones por triplicado para cumplir las recomendaciones de la IUPAC (Danzer y Currie 1998, Castells y Castillo 2000). Por lo tanto, la estrategia presentada disminuye el trabajo experimental manteniendo cifras de mérito aceptables, ya que se utiliza la misma curva de calibrado en agua MilliQ para predecir la muestra. Además, la Tabla 4.3 muestra que la sensibilidad analítica (γ) aumenta al ser eliminado el efecto de filtro interno sobre la señal del analito en la muestra al aplicar PDS.

Tabla 4.3. Resultados de predicción de FEN para las muestras de validación experimentales.

Muestra	PDS/ PARAFAC ^a (ng mL ⁻¹)	SEN	γ (mL ng ⁻¹)	γ^{-1} (ng mL ⁻¹)	SAM/ PARAFAC ^a (ng mL ⁻¹)	SEN	γ (mL ng ⁻¹)	γ^{-1} (ng mL ⁻¹)
1	380(3)							
2	393(4)				361(3)	0.19	0.32	3.1
3	325(4)							
4	259(3)							
5	315(4)							
6	375(6)							
7	285(3)							
8	389(5)							
9	308(2)	0.33	0.38	2.6				
10	341(1)							
11	378(2)							
12	239(9)				238(3)	0.16	0.30	3.3
13	312(2)							
14	343(3)				337(9)	0.22	0.25	4.0
15	342(2)							
16	289(3)							
17	287(3)							
Rec. (%) ^b	102 (5)				100 (5)			
REP (%) ^c	5				4			

^a Desviación estándar de predicción entre paréntesis.

^b Recuperación promedio. Desviación estándar entre paréntesis.

^c Error relativo de predicción, $REP = \frac{100}{\bar{c}} \left[\frac{1}{I} \sum_{i=1}^I (c_{nom} - c_{pred})^2 \right]^{1/2}$, donde I es el número de

muestras, c_{nom} y c_{pred} son las concentraciones nominal y predicha, y \bar{c} es la concentración promedio.

4.5.3. Muestras de agua

La metodología PDS/PARAFAC se aplicó para la determinación de FEN en dos muestras de agua de diferente procedencia: canilla y napa. Ambas muestras se fortificaron con FEN, IBU y AAS debido a que las matrices de excitación-emisión

correspondientes a las muestras originales mostraron no contener cantidades detectables de estas drogas. Las muestras se prepararon con las concentraciones nominales informadas en la Tabla 4.4.

Al analizar las muestras de agua reales se observaron cambios en la sensibilidad debido a la matriz de cada una (efecto adicional al producido por PAR). Además, también se adicionó PAR para aumentar la complejidad del sistema. Para evaluar la repetibilidad y la precisión intermedia se investigaron tres muestras independientes por triplicado durante tres días consecutivos.

El número de componentes a ser incluidos en el modelo PARAFAC fue cuatro en ambos casos, atribuidos a la señal de fluorescencia de FEN, IBU, AAS y a un componente no esperado de la matriz de las muestras. En la Tabla 4.4 se presentan los resultados de predicción obtenidos al aplicar la estrategia propuesta.

Tabla 4.4. Resultados de predicción de las muestras de agua.

Componente		Muestra			
		Agua de canilla		Agua de napa	
		Agregado	Predicho ^a	Agregado	Predicho ^a
FEN (ng mL ⁻¹)	Día 1	306.6	297 (5)	499.8	484 (5)
	Día 2	306.6	293 (6)	499.8	493 (7)
	Día 3	306.6	304 (5)	499.8	491 (6)
IBU (ng mL ⁻¹)		397.7	-	299.9	-
AAS (ng mL ⁻¹)		70.8	-	50.6	-
PAR (µg mL ⁻¹)		13.0	-	11.0	-
Rec. (%)			97		98

^a Desviación estándar de predicción entre paréntesis ($n = 3$). Se analizaron tres muestras independientes durante tres días consecutivos.

Con el propósito de analizar si las diferencias entre los valores promedios entre los tres días no eran significativas, es decir, si las variaciones eran debidas a un muestreo aleatorio, se aplicó un test ANOVA a cada nivel de concentración. Como resultado de la prueba estadística aplicada, se concluyó que no existe diferencia estadísticamente significativa entre las medias ($p = 0.11$ para agua de canilla y $p = 0.24$ para agua de napa).

En la Figura 4.7 se puede observar la relación entre los *scores* y la concentración nominal de FEN en las muestras de calibración, mientras que la Figura 4.8 muestra los *loadings* de los dos modos arrojados por PARAFAC, correspondientes al análisis de la muestra de napa. Las Figuras 4.8A y B corresponden a los modos de excitación y emisión, respectivamente. La presencia del interferente original presente en la matriz de la muestra se puede atribuir a sustancias húmicas, como los ácidos húmico y fúlvico. Estas sustancias se producen por la biodegradación de la materia orgánica muerta y se encuentran con frecuencia en muestras de agua. Estas moléculas poseen diferentes grupos funcionales y se ha demostrado que presentan fluorescencia (Senesi y col. 1991, Chin y col. 1994, Hessen y Tranvik 1998).

Como puede verse, los resultados presentados en la Tabla 4.4 son comparables con los obtenidos para las muestras de validación (Tabla 4.3), sugiriendo que el método propuesto es apropiado para determinar FEN en muestras de agua.

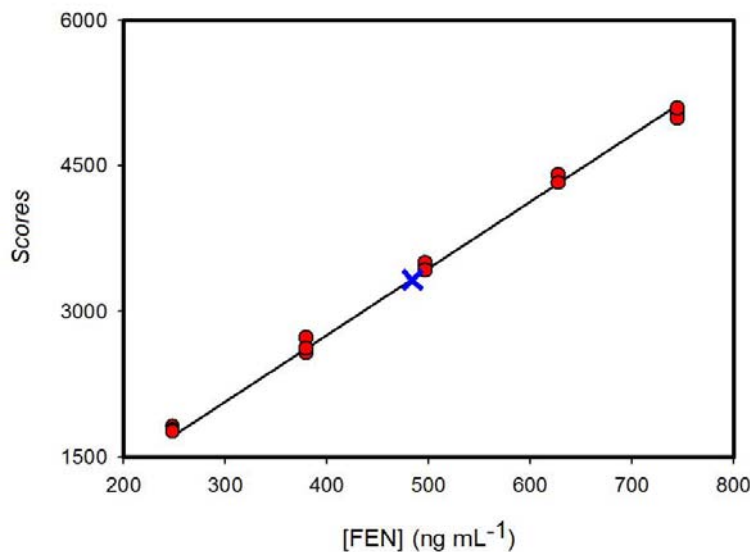


Figura 4.7. Scores de calibración de FEN (círculos rojos) y de FEN en la muestra de agua de napa analizada (cruz azul).

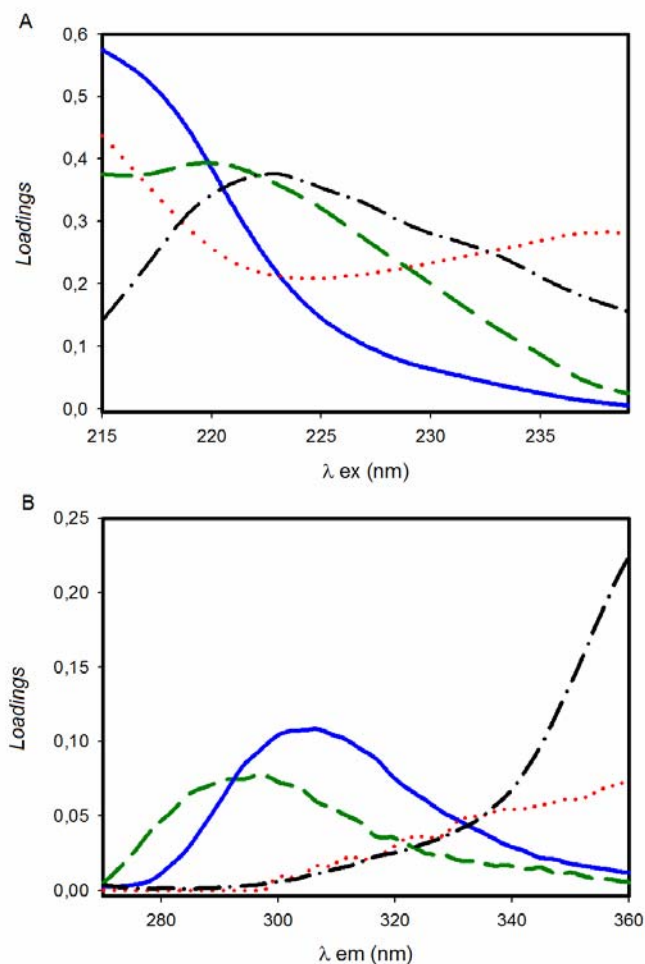


Figura 4.8. Resultados de PARAFAC al predecir FEN (línea sólida azul) en la muestra de agua de napa en presencia de IBU (línea discontinua verde), AAS (línea de puntos roja) y un componente inesperado (línea discontinua con puntos negra): (A) modo excitación, (B) modo emisión.

4.6. Conclusiones

La metodología presentada en este capítulo demuestra que es posible reducir el trabajo experimental para la determinación de un analito en presencia de componentes desconocidos y efecto de filtro interno, evitándose la preparación de un gran número de soluciones y manteniéndose cifras de mérito aceptables. Esta ventaja se logró combinando el algoritmo de segundo orden PARAFAC con la técnica de estandarización PDS. Contrariamente a aplicaciones previas del método de estandarización directa sucesiva combinada con algoritmos de segundo orden (Gil García y col. 2008, Gil García y col. 2009, Vosough y col. 2010), se realizó una demostración estadística relacionada a la aplicabilidad de la metodología propuesta. Esto permite concluir que la estrategia PDS/PARAFAC constituye un

método simple y de bajo costo para la determinación de FEN en muestras de agua con una reducción considerable de tiempo. Esta metodología puede extenderse a otros sistemas que presenten efecto matriz y, consecuentemente, puede ser una herramienta útil para determinar la cantidad de fármacos presentes en ambientes acuáticos y para evaluar el efecto de las plantas de tratamiento de efluentes convencionales en la eliminación de compuestos farmacéuticos.



CAPÍTULO 5

Modelado de matrices de fluorescencia sincrónica total con algoritmos de primer y segundo orden. Determinación de doxorrubicina en plasma humano

5.1. Introducción

Actualmente, existe una amplia aceptación en la comunidad analítica acerca de la creciente necesidad de mejorar los métodos analíticos disponibles para estudiar muestras cada vez más complejas. En el campo de la quimiometría, este hecho está relacionado con la búsqueda intensiva de nuevas formas de obtener datos multidimensionales y, en consecuencia, su modelado con algoritmos adecuados, que se generaron en los últimos años (Olivieri y col. 2011, Escandar y col. 2007, Escandar y col. 2013). Como ya se mencionó en los capítulos anteriores, los métodos de calibración multivariada pueden clasificarse de manera similar a los instrumentos, es decir, de orden cero, de primer orden, segundo orden y de órdenes superiores. A medida que se aumenta el orden de la calibración, los algoritmos se vuelven más poderosos ya que aumenta la información que puede extraer de los datos (Escandar y col. 2007, Escandar y col. 2013). Los algoritmos de primer orden permiten la identificación de nuevas muestras que contienen componentes inesperados; una propiedad conocida como la ventaja de primer orden (Booksh y Kowalski 1994). Estos métodos de calibración pueden compensar posibles interferencias, por lo general, por su inclusión en el conjunto de calibración. También se pueden implementar estrategias de selección de longitudes de onda con el objetivo de eliminar la región espectral afectada por la interferencia (Goicoechea y Olivieri 1999). Por otra parte, los datos instrumentales de múltiples vías llevan el potencial intrínseco de la ventaja de segundo orden, que, en principio, permite la cuantificación de un analito en muestras que contienen componentes inesperados, es decir, compuestos no incluidos en la calibración (Booksh y Kowalsky 1994). Esta propiedad es especialmente interesante cuando se quieren analizar muestras complejas, ya que permite construir modelos de calibración con un número limitado de estándares, manteniendo la calidad en la habilidad predictiva del analito aún en presencia de agentes interferentes, sin la necesidad de aplicar técnicas de separación cromatográficas o electroforéticas o de realizar un pretratamiento a la muestra para eliminar las interferencias.

Una revisión de la literatura en relación a los algoritmos de segundo orden capaces de obtener la ventaja de segundo orden revela un hecho interesante: en su mayoría requieren que los datos cumplan con la propiedad de trilinealidad. De acuerdo a una revisión reciente presentada por Escandar y colaboradores

(Escandar y col. 2011), los siguientes requerimientos son necesarios para cumplir con la propiedad de multi linealidad: (I) perfiles constantes a través de las diferentes muestras para cada componente, (II) relación lineal entre la señal y la concentración del analito, y (III) señal bilineal para una muestra dada. La propiedad de trilinealidad implica que una única matriz de datos para un componente puede ser descompuesta en el producto de dos vectores, conteniendo cada uno el perfil del componente en una de las dos dimensiones de los datos (un ejemplo conocido de este tipo de datos es la matriz de excitación-emisión). Debido a su estructura trilineal interna, PARAFAC parece ser el algoritmo apropiado cuando los datos siguen una estructura trilineal (Bro 1997).

Como se mencionó anteriormente en la Sección 1.2.4 del Capítulo 1, un arreglo de tres vías \mathbf{X} puede ser modelado según la siguiente ecuación:

$$x_{ijk} = \sum_{n=1}^N a_{in} b_{jn} c_{kn} + e_{ijk} \quad (5.1)$$

donde N es el número total de constituyentes químicos que generan la señal medida, a_{in} es la concentración relativa o *score* del componente n en la muestra i th, y b_{jn} y c_{kn} son las intensidades en las dimensiones J y K , respectivamente. Los valores de e_{ij} son los elementos de la matriz \mathbf{E} , que representa el error residual, y tiene las mismas dimensiones que \mathbf{X} . La suma en la Ecuación 5.1 implica que las señales individuales son aditivas. En general, los *loadings* se normalizan a la unidad, y se recogen en las matrices \mathbf{B} y \mathbf{C} , de tamaño $J \times N$ y $K \times N$, respectivamente. Los *scores* se agrupan en la matriz \mathbf{A} (tamaño $I \times N$), que refleja los valores de concentración relativa de varios componentes en todas las muestras. Esta información se puede utilizar para construir un modelo univariado. Este modelo permite la predicción de la concentración del analito en muestras desconocidas al proyectar su *score* en la recta ajustada.

Según Escandar y colaboradores (Escandar y col. 2011), las siguientes propuestas pueden ser utilizadas para tratar datos que no cumplen con la propiedad de multi-linealidad (trilinealidad):

a) En los casos en que no hay reproducibilidad en los tiempos de retención y/o las formas de los cromatogramas o electroferogramas, el analista puede recurrir al

alineamiento/remodelado usando algoritmos adecuados, luego de lo cual se puede aplicar cualquier algoritmo trilineal (Goicoechea y col. 2011). Otra opción es aplicar un modelo no trilineal capaz de hacer frente a las variaciones de los perfiles de los componentes entre las muestras. Esto incluye MCR-ALS (Tauler 1995) o PARAFAC2, una variante de PARAFAC capaz de modelar datos con pérdida en la constancia de los perfiles para cada componente de muestra a muestra (Kiers y col. 1999). Otras posibilidades incluyen U-PLS/RBL y N-PLS/RBL (Olivieri 2005a, Bortolato y col. 2009).

b) Dependencia lineal en los perfiles de los componentes: este es el caso de relaciones de cierre entre concentraciones de especies ácido-base conjugadas en un gradiente de pH, o entre concentraciones de reactivos y productos en una reacción química. En estas situaciones, pueden aparecer más de un mínimo que satisfagan los requerimientos del modelo trilineal. Las soluciones a este problema se pueden alcanzar a través de la aplicación conveniente de parámetros de inicialización y de restricciones adecuadas durante el ajuste de PARAFAC (Sadygov y col. 2006), o modelando con MCR-ALS o PARALIND (una variante de PARAFAC hecho especialmente para este tipo de datos) (Bahram y Bro 2007). Otra posibilidad es utilizar una adaptación del algoritmo Cuadrados Mínimos Bilineales (BLLS), que se aplicó recientemente a sistemas que presentan dependencia lineal (Goicoechea y Olivieri 2005). Una posibilidad más atractiva es, quizás, usar U-PLS/RBL o redes neuronales seguido de RBL (ANN/RBL). Ambos algoritmos han demostrado obtener buenos resultados sin la necesidad de modificaciones (Espinosa-Mansilla y col. 2004, Culzoni y Goicoechea 2007).

Un caso especial de dependencia lineal ocurre cuando el analito y el interferente tienen perfiles idénticos. Esto es particularmente cierto si hay al menos dos componentes con perfiles iguales en uno de los modos, originando una matriz con deficiencia de rango (Smilde y col. 1994). Este tipo de inconvenientes pueden ser evitados solamente utilizando MCR-ALS (Culzoni y col. 2008), PARALIND (Bahram y Bro 2007) o el nuevo algoritmo presentado por Olivieri y colaboradores, que combina la flexibilidad de U-PLS/RBL y la habilidad de MCR-ALS y PARALIND de ignorar la pérdida de selectividad en uno de los modos: el algoritmo U-PLS/RBL-LD (Lozano y col. 2010).

c) Efecto de filtro interno en espectroscopia luminiscente. Este efecto puede causar distorsiones en los perfiles de excitación y/o emisión de las especies de interés. Si el efecto ocurre tanto en el modo de excitación como en el de emisión de fluorescencia, el único algoritmo capaz de modelar convenientemente este tipo de datos es U-PLS/RBL (Bohoyo Gil y col. 2006). Ya que la dimensión de la distorsión depende de la magnitud de la interferencia, su concentración debe ser modelada en el juego de calibración. Otra posibilidad es calibrar la interferencia con la ayuda de un algoritmo trilineal y el método de adición estándar múltiple (Rodríguez y col. 2009a, Valderrama y Poppi 2011).

d) Relación no lineal entre la concentración del analito y la señal. La respuesta no lineal se puede deber a las diferentes fuentes de ruido instrumental, químicas o físicas, que pueden causar curvatura en la función concentración-respuesta. Este efecto se clasifica como efecto no lineal real (Gemperline y col. 1991). Se propuso la aplicación del análisis de componentes principales desdoblados seguido de RBL como algoritmo capaz de lidiar con información de segundo orden no lineal obteniendo la ventaja de segundo orden (Olivieri 2005, Culzoni y Goicoechea 2007). Otra opción es construir un gráfico de calibración univariado no lineal con las áreas bajo los perfiles cinéticos o los *scores* obtenidos por MCR-ALS y sus concentraciones asociadas (Schenone y col. 2013a).

e) Presencia de perfiles diferentes para un componente en una única muestra. Este es uno de los problemas más difíciles de resolver. Sólo se publicaron dos artículos en donde se modelan datos no bilineales para una única muestra (Zampronio y col. 2002, Kumar y Mishra 2011), e incluyen: (1) matrices de espectroscopia de fluorescencia sincrónica total (MFST) y (2) datos de espectrometría de masa en tándem (MS^2). Los datos MFST no son bilineales debido a que cada fluoróforo no tiene un espectro sincrónico único, ya que la forma del mismo cambia con la variación del intervalo de longitudes de onda ($\Delta\lambda$) (Kumar y Mishra 2011, Kumar y Mishra 2012b, Patra y Mishra 2002a, Patra y Mishra 2002b). Los datos MS^2 no son bilineales debido a que cada fragmento de un único componente tiene un patrón MS específico en la segunda dimensión MS, haciendo que sea imposible describir los datos MS^2 en términos de un perfil MS en cada dimensión. Se han obtenido

buenos resultados al analizar este tipo de datos no trilineales utilizando N-PLS combinado con pre-procesamiento de datos. Sin embargo, todos los componentes se proporcionaron tanto en las muestras de entrenamiento como en las de validación (Zampronio y col. 2002).

Los espectros de fluorescencia sincrónicos (EFS) constituyen un tipo de señal de primer orden que se utiliza para resolver muestras complejas con muchos componentes, cuyos espectros de emisión y/o excitación se encuentran severamente solapados. La espectroscopia sincrónica consiste esencialmente en la realización de un barrido en simultáneo de las longitudes de onda de excitación y emisión con una diferencia fija entre ambas ($\Delta\lambda$). Su principal característica es que se produce una reducción del ancho de banda (Lloyd 1971). Por lo tanto, EFS es un método muy simple y eficaz para la obtención de datos para varios componentes en una sola medición. Además, hay una reducción de las dispersiones Rayleigh y Raman y, por lo tanto, no se necesitan correcciones o métodos de pre-procesamiento como muchas veces ocurre en el caso de espectros de excitación y/o emisión (Wehry 1990).

Es importante destacar que aunque los espectros EFS son más angostos en comparación a los espectros de fluorescencia convencionales, no siempre es posible evitar el solapamiento de espectros en algunas mezclas de componentes múltiples. Sin embargo, la combinación de espectros EFS con métodos de calibración multivariada hace que el análisis sea simple y más veloz, sin necesidad de implementar pasos de extracción (Patra y Mishra 2002a).

Es posible obtener matrices de fluorescencia sincrónica total (MFST) registrando espectros sincrónicos para diferentes intervalos de longitud de onda ($\Delta\lambda$). Cada MFST consiste en un arreglo de dos vías de dimensiones longitud de onda de excitación $\times \Delta\lambda$, y puede ser modelada por algoritmos de calibración multivariada de segundo orden, considerando sus propiedades, que como se mostrará a continuación, son diferentes de las matrices de excitación-emisión convencionales.

Cada fluoróforo no posee un único espectro sincrónico debido a que su forma varía cuando se modifica el $\Delta\lambda$ (Kumar y Mishra 2012b). Kumar y Mishra cuantificaron simultáneamente cinco hidrocarburos aromáticos policíclicos (PAHs) usando MFST y métodos de calibración multivariada como N-PLS, U-PLS y MCR-

ALS, pero no se utilizó la ventaja de segundo orden debido a que se incorporaron todos los componentes tanto en las muestras de calibración como en las de validación (Kumar y Mishra 2011). Los mismos autores abordaron la falta de trilinealidad convirtiendo las matrices MFST en matrices de excitación-emisión, para luego aplicar PARAFAC (Kumar y Mishra 2012a).

La doxorubicina (DOX) es un antibiótico del grupo de las antraciclinas producida por la bacteria *Streptomyces peucetius varieta Cassius* y es actualmente utilizada en una variedad de tratamientos antitumorales (Benjamin y col. 1973, Speth y col. 1988). Para la determinación de esta droga y sus metabolitos en preparaciones farmacéuticas y en fluidos biológicos se desarrollaron numerosos métodos de HPLC (Andersen y col. 1993, Gotlieb y col. 2007, Chen y col. 2001, Loadman y Calabrese 2001, Fahmy y col. 2004), y HPLC acoplado a espectrometría de masas (Di Francesco y col. 2007, Lachâtre y col. 2000). Zagotto y colaboradores describieron varias técnicas de separación y cuantificación de DOX y otros agentes anticancerígenos en fluidos biológicos, que incluyen electroforesis capilar, espectroscopia UV-Vis y HPLC (Zagotto y col. 2001). Recientemente, Lu y colaboradores desarrollaron un método rápido para el análisis de DOX y daunorrubicina utilizando electroforesis capilar basada en un microchip con detección fluorescente inducida con láser (Lu y col. 2009).

La determinación de DOX en plasma humano se logró modelando matrices de excitación-emisión de fluorescencia con PARAFAC y N-PLS (Trevisan y Poppi, 2003). Sin embargo, el juego de calibración se construyó en plasma fortificado con DOX y no se llevaron a cabo determinaciones en muestras con componentes desconocidos.

El objetivo de este trabajo es demostrar que los datos MFST modelados con MCR-ALS o U-PLS acoplado a RBL pueden ser utilizados para la cuantificación de DOX en muestras de plasma en presencia de riboflavina (RF) cuyo espectro de fluorescencia se solapa con el del analito, sin pasos de pre-procesamiento y calibrando sólo con DOX. RF es una vitamina soluble en agua que es crítica en el metabolismo y en la producción de energía. Además, presenta emisión de fluorescencia nativa y se encuentra en fluidos biológicos, como el plasma. Debido a que su espectro de emisión se solapa con el de DOX, se seleccionó como interferente no calibrado.

Una de las características más importantes de este trabajo es que se obtuvo por primera vez la ventaja de segundo orden con datos no bilineales. Este objetivo se logró de dos maneras diferentes: a) a partir de respuestas de primer orden (los datos de dos vías se desdoblaron asumiendo que los datos MFST desdoblados tienen estructura bilineal (Kumar y Mishra 2012b)) modeladas con MCR-ALS, y b) a partir de datos de segundo orden modelados con U-PLS/RBL, que ya demostró obtener la ventaja de segundo orden cuando se analizan datos instrumentales que se desvían de la trilinealidad (Culzoni y col. 2006, Goicoechea y Olivieri 2005).

En cuanto a la aplicación de MCR-ALS, se demostró que la obtención de la ventaja de segundo orden, utilizando y procesando datos de primer orden, es posible, especialmente cuando se aplica la restricción de correlación (Culzoni y col. 2006, Goicoechea y Olivieri 2005, Llamas y col. 2009, Mamián-López y Poppi 2013).

5.2. Objetivos

- Evaluar el rendimiento del algoritmo MCR-ALS en el modelado de matrices de fluorescencia sincrónica total, no bilineales y desdobladas, aplicando la restricción de correlación.
- Comparar los resultados de la predicción con MCR-ALS para la determinación de doxorubicina en plasma en presencia de riboflavina, con los obtenidos por el algoritmo U-PLS/RBL.

5.3. Programas

Para el procesamiento de los datos, todas las rutinas se implementaron usando MATLAB 7.1 (MATLAB 7.1, 2005). U-PLS/RBL se implementó a través de rutinas escritas por el Dr. Olivieri e incluidas en la interfaz gráfica MVC2 (Olivieri y col. 2009). Para la aplicación de MCR-ALS con la restricción de correlación se utilizó un código de MATLAB que puede solicitarse al Dr. Tauler (e-mail: Roma.Tauler@idaea.csis.es).

5.4. Materiales y Métodos

5.4.1. Reactivos y soluciones

Se utilizaron reactivos de calidad analítica y agua ultrapura, obtenida a partir de un sistema purificador de agua Millipore (Bedford, MA, USA). Doxorubicina clorhidrato (94.0 %) y riboflavina (96.9 %) se obtuvieron de Laboratorios Richmond (Buenos Aires, Argentina). El ácido tricloroacético (TCA) se obtuvo de Anedra (San Fernando, Argentina) y el metanol grado HPLC de Aberkon (Buenos Aires, Argentina).

Se prepararon soluciones stock de DOX (455.0 mg L⁻¹) y RF (60.0 mg L⁻¹) disolviendo cantidades apropiadas de cada droga en agua MilliQ y metanol, respectivamente. Estas soluciones se almacenaron en oscuridad a 4 °C. Las soluciones estándares de trabajo se prepararon diariamente diluyendo la solución stock con agua MilliQ para obtener las concentraciones deseadas.

Se utilizaron muestras de plasma humano heparinizado de pacientes sanos del Hospital Iturraspe de Santa Fe, Argentina, y se almacenaron a -20 °C hasta que los experimentos se llevaran a cabo.

5.4.2. Preparación de muestras de calibración y validación.

Se preparó un juego de calibración agregando estándar puro de DOX al plasma debido a la presencia de interacciones entre el analito y la matriz de la muestra (ver más abajo). El juego se construyó en dos pasos: primero se prepararon cinco estándares de DOX en agua, por duplicado, con las siguientes concentraciones: 8.0, 24.1, 40.0, 56.9 y 72.8 µg mL⁻¹. A continuación, se colocaron 20.0 µL de estas soluciones en tubos Eppendorf junto con 600.0 µL de plasma. Luego de adicionar 20 µL de TCA 20 % (p/v) para precipitar las proteínas, cada muestra se mezcló 30 s utilizando un vórtex y se centrifugó a 14000 rpm durante 10 min. A continuación, se transfirieron 500.0 µL de sobrenadante a una microcubeta de cuarzo. Los valores de concentración final correspondiente a la droga en plasma fueron: 0.25, 0.75, 1.25, 1.78 y 2.28 µg mL⁻¹. El paso de precipitación de proteínas se realizó para eliminar un posible efecto matriz y para aumentar la sensibilidad del método, ya que las proteínas del plasma proporcionan un apagamiento significativo de la fluorescencia.

Además, se preparó un juego de nueve muestras de validación siguiendo un diseño central compuesto de dos factores y un punto central, con diferentes concentraciones de DOX (entre 16.0 a 64.0 $\mu\text{g mL}^{-1}$) y RF (entre 2.4 y 9.6 $\mu\text{g mL}^{-1}$) en agua. Luego, se siguió el mismo procedimiento descripto para las muestras de calibración. Las concentraciones finales de DOX y RF en las muestras de validación se muestran en la Tabla 5.1 (Sección “Resultados y Discusión”). Estas concentraciones se encuentran dentro de los valores terapéuticos en plasma humano de la droga estudiada.

5.4.3. Equipamiento

Todas las medidas fluorimétricas se llevaron a cabo utilizando un espectrofluorímetro Perkin-Elmer LS-55 equipado con una lámpara de descarga de xenón, monocromadores de tipo Monk-Gillieson y un fotomultiplicador conectado a una PC vía una conexión RS232C, y usando microcubetas de cuarzo de 1.00 cm de paso óptico. Los anchos de banda de las aberturas de los monocromadores de excitación y emisión se fijaron en 10 nm y el voltaje del detector en 650 V. Los espectros EFS se adquirieron en un rango de longitudes de onda de excitación entre 489.5 y 614.5 nm cada 0.5 nm, con $\Delta\lambda$ entre 30 y 160 nm, cada 10 nm entre los monocromadores de excitación y emisión. Por lo tanto, el tamaño de cada matriz fue de 251 \times 14. Además, para analizar las características fluorescentes de DOX, RF y el plasma, se registraron matrices de excitación-emisión entre las longitudes de onda de 400 a 480 nm (cada 5 nm) y de 500 a 600 nm (cada 0.5 nm) para excitación y emisión, respectivamente.

5.5. Resultados y Discusión

5.5.1. Consideraciones generales

Una inspección de la figura de contornos de las MEE obtenidas para las drogas puras DOX y RF, ambas en agua (Figura 5.1) muestra que sus máximos de excitación están en 480 y 445 nm, respectivamente, y sus máximos de emisión en 555 y 520 nm, respectivamente. Además, la figura muestra los espectros del plasma humano desproteinizado, demostrando que existe un solapamiento considerable entre los tres componentes. Por lo tanto, la cuantificación de DOX en plasma en presencia de RF a través del método de calibración univariada es difícil

de realizar. En cambio, pueden utilizarse los métodos de calibración multivariada que involucran el análisis de más de una variable a la vez para estudiar este tipo de muestras sin ninguna técnica de separación.

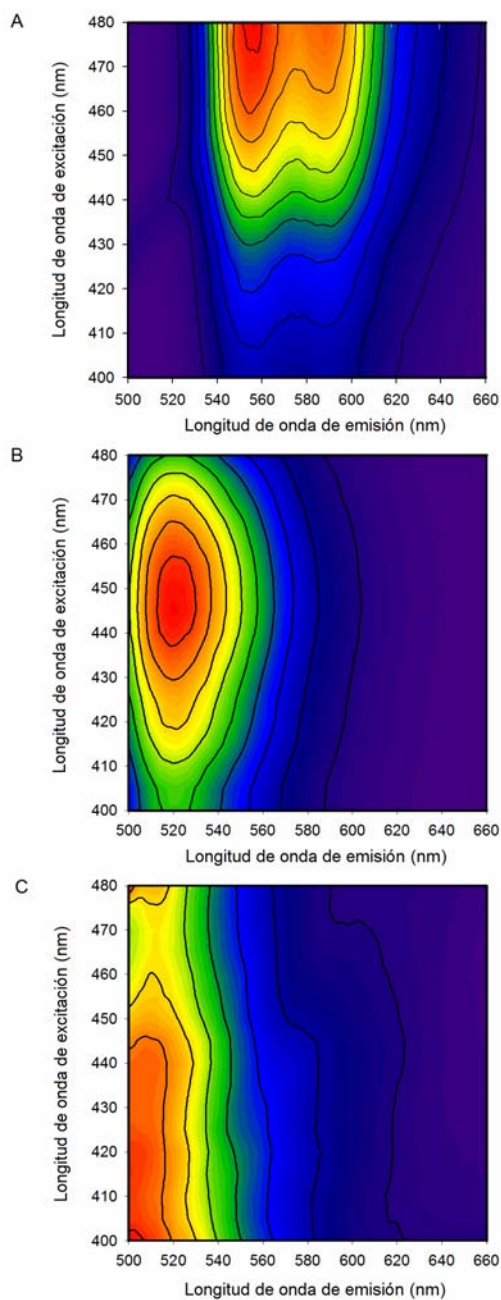


Figura 5.1. Superficies de contorno de (A) DOX pura en agua, (B) RF pura en agua y (C) plasma humano desproteínado.

La Figura 5.2A compara los espectros de fluorescencia sincrónicos a $\Delta\lambda = 80$ nm de una solución acuosa de DOX pura y la diferencia entre los espectros de una muestra de plasma con y sin DOX. El valor de $\Delta\lambda$ elegido corresponde al corrimiento de Stokes, es decir, la diferencia entre los máximos de excitación y emisión de DOX (la diferencia exacta es 75 nm, pero las medidas experimentales se llevaron a cabo entre 30 y 160 nm, cada 10 nm). Se puede apreciar que la intensidad de fluorescencia del analito decrece significativamente debido a su interacción con la matriz del plasma. Por esta razón, se prepararon las muestras de calibración agregando estándar puro de DOX al plasma. Para evaluar si este comportamiento se mantiene en diferentes muestras de plasma, se compararon estadísticamente las pendientes de las curvas de calibrado univariadas (graficando la intensidad de fluorescencia a $\lambda_{\text{ex}}=480$ nm y $\lambda_{\text{em}}=555$ nm vs. la concentración de DOX) construidas con estándar de DOX en tres muestras de plasma distintas. Dado que el valor de p para las pendientes fue 0.912 ($p>0.1$), se concluyó que no existe diferencia estadísticamente significativa con un nivel de confianza del 90 %.

La Figura 5.2B muestra varios espectros sincrónicos de DOX pura a diferentes $\Delta\lambda$. Como puede observarse, la forma de los espectros del fluoróforo cambia con el cambio de $\Delta\lambda$, por lo que un MFST se debe considerar como no bilineal. Por ello, se implementaron dos estrategias para modelar este tipo de datos: a) desdoblamiento de las matrices y modelado con MCR-ALS obteniendo la ventaja de segundo orden con datos de primer orden como se realizó en un número reducido de publicaciones (Azzouz y Tauler 2008, Goicoechea y col. 2010, Llamas y col. 2009, Mamián-López y Poppi, 2013), y b) modelado con U-PLS e intento de explotar la ventaja de segundo orden con la implementación de RBL, aunque se sabe que este paso requiere que los datos sean bilineales (Öhman y col. 1990).

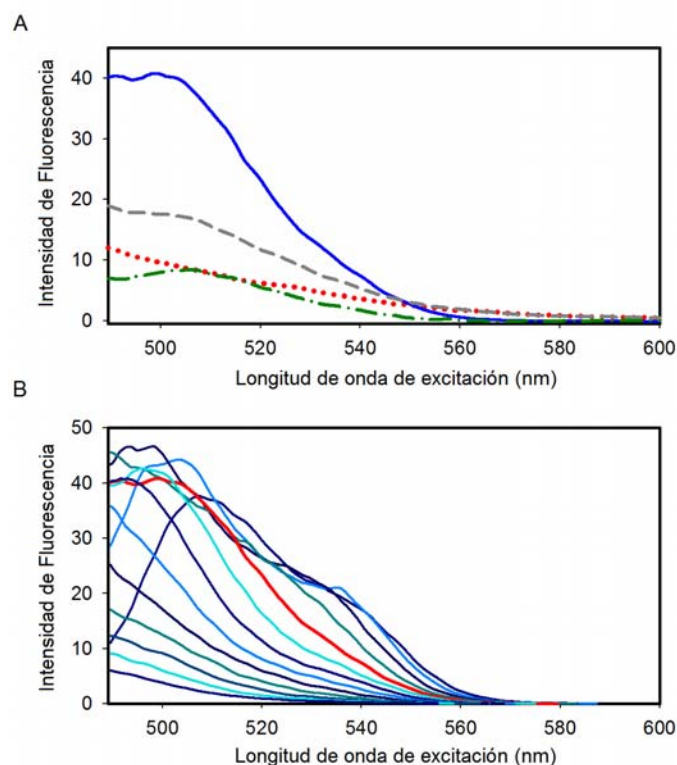


Figura 5.2. (A) Espectros sincrónicos de DOX pura $1.25 \mu\text{g mL}^{-1}$ (línea sólida azul), plasma con $1.25 \mu\text{g mL}^{-1}$ de DOX (línea discontinua gris), diferencia entre los espectros de plasma con $1.25 \mu\text{g mL}^{-1}$ de DOX y blanco de plasma (línea discontinua con puntos verde) y blanco de plasma (línea de puntos roja), todos a $\Delta\lambda = 80 \text{ nm}$. (B) Espectros sincrónicos de DOX pura $1.25 \mu\text{g mL}^{-1}$ a varios $\Delta\lambda$ (desde 30 a 160 nm, cada 10 nm). En rojo, el espectro a $\Delta\lambda = 80 \text{ nm}$.

5.5.2. MCR-ALS

El algoritmo de calibración multivariada MCR-ALS se describió ampliamente en la Sección 1.2.1 del Capítulo 1. En esta oportunidad, se comentará acerca del significado de cada parámetro de la Ecuación (1.9) de ese apartado: las filas de la matriz **D** contienen los espectros sincrónicos desdoblados de las diferentes muestras a varias longitudes de onda de excitación, las columnas de **C** contienen las concentraciones relativas de las especies intervinientes, las columnas de **S** sus espectros desdoblados y **E** es la matriz de los residuos no ajustados por el modelo.

Durante el cálculo iterativo de **C** y **S^T**, se aplicaron una serie de restricciones para dar un significado físico a las soluciones obtenidas. En este capítulo, se usaron la no-negatividad en espectros y concentraciones y se incluyó información acerca de la correspondencia entre especies en cada matriz a través de una matriz binaria,

que indica la presencia o ausencia de un dado componente en las diferentes matrices desdobladas y apiladas (Ruckebush y col. 2006). Además, se aplicó la restricción de correlación, mediante la que se forzaron las concentraciones del analito en las muestras de calibración, \mathbf{c}_{cal} , en cada iteración ALS a valores de referencia previamente conocidos del analito, \mathbf{c}_{ref} , en estas muestras (Antunes y col. 2002, Azzouz y Tauler 2008). Para su implementación se construye un modelo lineal local entre los valores de \mathbf{c}_{cal} y \mathbf{c}_{ref} de modo que:

$$\mathbf{c}_{\text{ref}} = b\mathbf{c}_{\text{cal}} + b_0 + \mathbf{e}_{\text{ref}} \quad (5.1)$$

donde b y b_0 son la pendiente y los valores de corrección que mejor ajustan \mathbf{c}_{cal} a \mathbf{c}_{ref} , obtenidos por regresión lineal de mínimos cuadrados, y \mathbf{e}_{ref} es el error en las concentraciones de referencias (no modeladas). Los valores de concentraciones correspondientes a las muestras de calibración calculadas usando este modelo local son:

$$\hat{\mathbf{c}}_{\text{cal}} = b\mathbf{c}_{\text{cal}} + b_0 \quad (5.2)$$

Para predecir una concentración desconocida del analito en una muestra $\hat{\mathbf{c}}_{\text{des}}$, la ecuación que se utiliza es:

$$\hat{\mathbf{c}}_{\text{des}} = b\mathbf{c}_{\text{des}} + b_0 \quad (5.3)$$

donde b y b_0 son los valores obtenidos previamente en el paso de calibración a partir de \mathbf{c}_{ref} , y \mathbf{c}_{des} son las concentraciones de las muestras predichas por ALS. Entonces, cada iteración ALS se completa luego de actualizar los valores de predicción obtenidos, es decir, por sustitución de \mathbf{c}_{des} por $\hat{\mathbf{c}}_{\text{des}}$.

Previamente al análisis de los datos, cada matriz (de tamaño 251×14) se desdobló en un vector de 1×3514 . Para realizar la cuantificación, las matrices desdobladas correspondientes a las muestras de validación se analizaron simultáneamente con las muestras de calibración. Luego, todos los vectores se apilaron formando una matriz aumentada en columnas de tamaño 19×3514 .

La estimación del número de componentes presentes en el sistema se realizó aplicando SVD a la matriz aumentada, lo que sugirió la existencia de tres factores significativos. Las estimaciones iniciales para MCR-ALS se obtuvieron por selección de las variables más puras basadas en SIMPLISMA (Windig y Guilment 1991).

La implementación de restricciones adecuadas durante la optimización hace que los resultados estén menos sujetos a ambigüedades. Para la implementación de la restricción de correlación se proporcionó una matriz C_{sel} del mismo tamaño que C (19×3) como dato de entrada, la que contiene en las primeras diez filas las concentraciones nominales de DOX en las muestras de calibración, como así también información que indica que el plasma, pero no la interferencia, están presentes. Las siguientes nueve líneas indican la presencia de los tres componentes en las muestras de validación con valores no restringidos (NaN). A continuación, en la Figura 5.3 se muestra una representación de la matriz mencionada:

	DOX	RF	Plasma	
Perfil conocido Información de rango local (ausencia de RF)	0.294	0	NaN	Muestras de calibración
	0.294	0	NaN	
	0.750	0	NaN	
	0.750	0	NaN	
	1.250	0	NaN	
	1.250	0	NaN	
	1.780	0	NaN	
	1.780	0	NaN	
	2.275	0	NaN	
	2.275	0	NaN	
NaN: valores no restringidos	NaN	NaN	NaN	Muestras de validación
	NaN	NaN	NaN	
		⋮		
	NaN	NaN	NaN	
	NaN	NaN	NaN	
	NaN	NaN	NaN	

C_{sel}

Figura 5.3. Esquema de la matriz C_{sel} utilizada para aplicar la restricción de correlación. NaN: valor no restringido.

Luego de treinta y ocho ciclos iterativos se alcanzó el criterio de convergencia seleccionado (cambio relativo en el ajuste para iteraciones sucesivas

= 0.1 %). El valor de la falta de ajuste fue de 2.43 y la variancia explicada 99.94 %. La Figura 5.4 muestra los espectros en S^T recuperados por MCR-ALS. Aunque existe un solapamiento significativo, los espectros correspondientes a DOX y a la interferencia pudieron ser obtenidos adecuadamente. En la imagen insertada en el recuadro de la Figura 5.4 se muestran los espectros sincrónicos obtenidos para los tres componentes y, además, el espectro normalizado de DOX pura, todos a $\Delta\lambda = 80$ nm, que es el valor óptimo para el analito. Como puede observarse, el espectro recuperado para DOX coincide con el puro (con un $r^2 \approx 1$), indicando que el perfil extraído corresponde al del analito de interés.

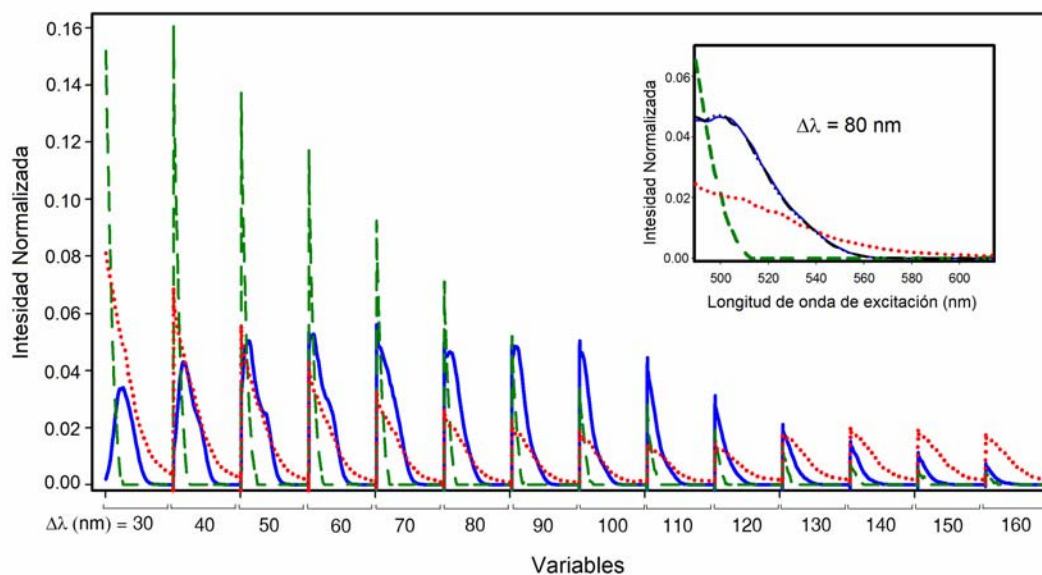


Figura 5.4. Espectros recuperados en la matriz S^T por MCR-ALS luego de analizar la matriz aumentada que contiene las muestras de validación y calibración. La línea sólida azul corresponde a DOX, la línea discontinua verde a RF y la línea de puntos roja a la señal del plasma. En la imagen insertada, se muestran los espectros de los tres componentes (DOX: línea sólida azul, RF: línea discontinua verde, plasma: línea de puntos roja) y el espectro normalizado de DOX pura (línea discontinua con puntos negra), todos a $\Delta\lambda = 80$ nm.

La cuantificación se llevó a cabo a través de la regresión de los *scores* obtenidos en la matriz C para cada muestra de calibración en función de los valores de concentración nominales. La Figura 5.5 muestra la gráfica univariada y los *scores* de las muestras de validación, ilustrando cómo varían los *scores* correspondientes al analito en las muestras de calibración (círculos azules) y de validación (cruces rosadas), mientras que, como se esperaba, los *scores* del plasma

(triángulos grises y rojos) se mantienen casi constantes. En el caso de RF, no hay contribución en los calibradores (rombos verdes; este hecho es esperable ya que fue impuesto como restricción), pero su presencia es notable en las mezclas de validación (cuadrados blancos).

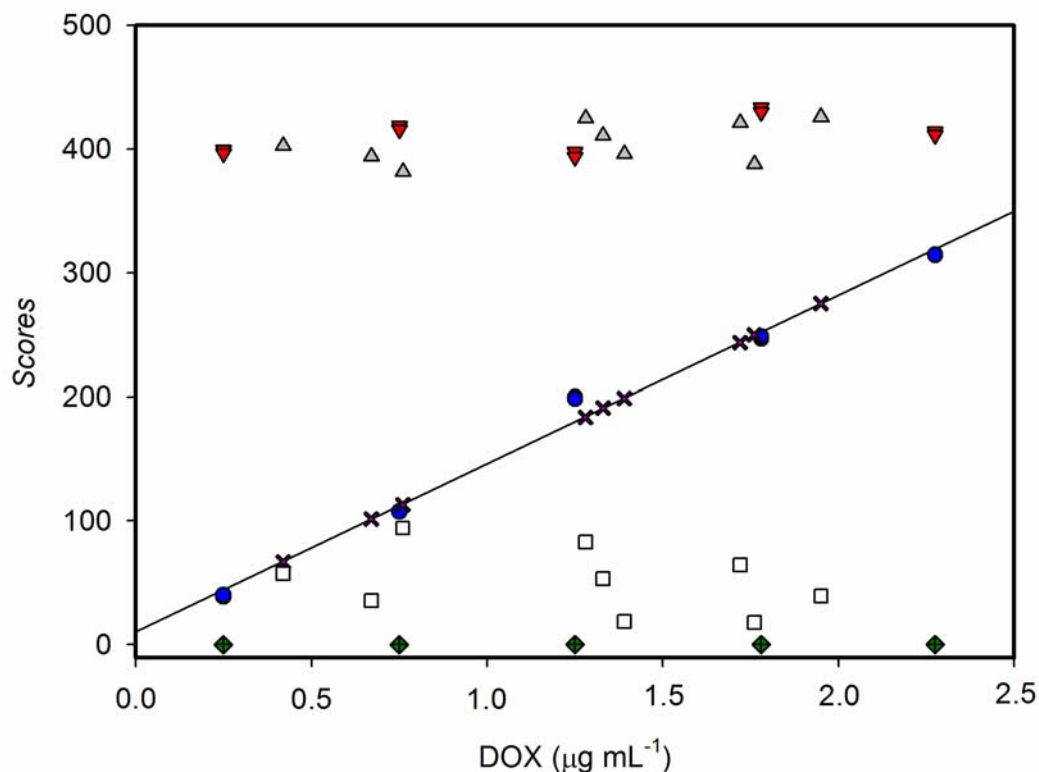


Figura 5.5. Scores obtenidos por MCR-ALS correspondientes a DOX (círculos azules), RF (rombos verdes) y plasma (triángulos hacia abajo rojos) en las muestras de calibración, y DOX (cruces rosadas), RF (cuadrados blancos) y plasma (triángulos hacia arriba grises) en las muestras de validación.

La Tabla 5.1 presenta los resultados obtenidos al analizar las mezclas de validación, mostrando una excelente concordancia entre los valores nominales y predichos: el RMSE es $0.07 \mu\text{g mL}^{-1}$, lo que implica un valor de REP del 5 %. Además, se evaluó la precisión del método a través del análisis de cinco réplicas de la muestra de validación número 5 (ver Tabla 5.1), por el mismo analista, bajo las mismas condiciones y en tres días diferentes. Los coeficientes de variación (CV%) obtenidos fueron todos menores al 15 % (ver Tabla 5.2), tal como lo recomienda la Agencia Europea de Medicamentos en un documento presentado recientemente (Agencia Europea de Medicamentos, 2013). Además, mediante la aplicación de la prueba ANOVA, se compararon las variancias de los tres días y se concluyó que no

hay diferencias significativas ($p = 0.524$). Estos parámetros, junto con la recuperación promedio ($100\% \pm 8\%$), indican que el método propuesto resulta una metodología viable para conseguir la ventaja de segundo orden en el caso de datos MFST no trilineales desdoblados.

Tabla 5.1. Predicción y análisis estadístico de DOX en muestras de validación.

Muestra	Concentración nominal			DOX predicha		
	DOX ($\mu\text{g mL}^{-1}$)	RF ($\mu\text{g mL}^{-1}$)	MCR-ALS ^a ($\mu\text{g mL}^{-1}$)	Rec. (%)	U-PLS/RBL ^{a,b} ($\mu\text{g mL}^{-1}$)	Rec. (%)
	1	0.50	0.19	0.42(1)	84.0	0.43(1)/2
2	1.78	0.27	1.72(1)	97.6	1.73(1)/1	97.2
3	1.25	0.30	1.28(1)	102.4	1.28(1)/2	102.4
4	1.99	0.19	1.95(1)	98.0	1.91(1)/2	96.0
5	1.25	0.19	1.33(1)	106.4	1.32(1)/2	105.6
6	0.72	0.27	0.76(1)	105.6	0.73(1)/2	101.4
7	1.25	0.075	1.38(1)	110.4	1.36(1)/3	108.8
8	0.72	0.11	0.67(1)	93.1	0.69(1)/1	95.8
9	1.78	0.11	1.77(1)	99.4	1.80(1)/1	101.1
Rec. (%) ^{a,c}				100(8)	99(7)	
RMSE($\mu\text{g mL}^{-1}$) ^d				0.07	0.06	
REP (%) ^e				5	5	

^a Desviación estándar entre paréntesis.

^b Número de interferentes que se consideraron para estabilizar el error s_u después de la barra.

^c Recuperación promedio.

^d Raíz del error cuadrático medio, $\text{RMSE} = \left[\frac{1}{I} \sum_{i=1}^I (c_{\text{nom}} - c_{\text{pred}})^2 \right]^{1/2}$, donde $I = 9$, c_{nom} y c_{pred}

son las concentraciones nominal y predicha.

^e Error relativo de predicción, $\text{REP} = 100 \times \text{RMSE} / \bar{c}$, donde \bar{c} es la concentración promedio de calibración.

5.5.3. U-PLS/RBL

La primera fase en el procesamiento de los datos con el algoritmo U-PLS/RBL consiste en la estimación del número de componentes. En este caso, el número de variables latentes A se estima por el método de validación cruzada (Haaland y Thomas 1988). La Figura 5.6 muestra cómo varía la suma de los cuadrados del error de predicción ($PRESS = \sum (c_{nom} - c_{pred})^2$) a medida que aumenta el número de variables latentes cuando se usa el juego de calibración, indicando que $A = 2$ es el valor óptimo, que corresponde a la presencia de DOX y plasma en estas muestras.

El método RBL se utiliza frecuentemente para filtrar la contribución de componentes inesperados (N_{int}) de los datos de validación. Para encontrar este número se analizaron los residuos s_u obtenidos por el procedimiento RBL. La Figura 5.6 indica claramente que los residuos se estabilizan en $N_{int} = 1$ cuando se analiza la muestra de validación número dos, debido a la presencia de RF en las muestras de validación.

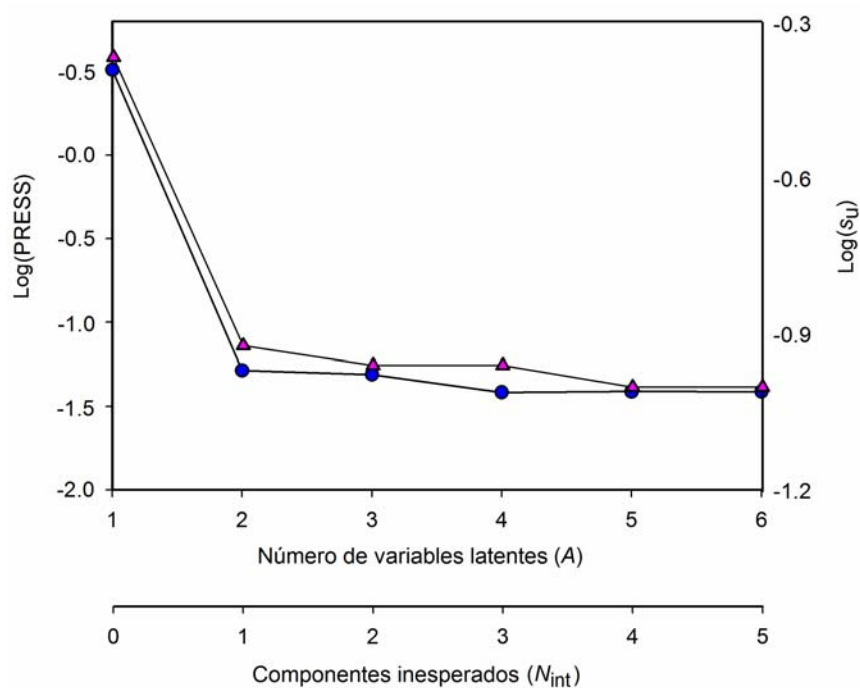


Figura 5.6. Logaritmo del PRESS en función del número de variables latentes A de U-PLS, obtenidos con datos de calibración (círculos), y logaritmo de los residuos de predicción de U-PLS en función del número de componentes inesperados (N_{int}) (triángulos) al analizar la muestra de validación número dos.

U-PLS/RBL se aplicó para predecir las muestras de validación, usando dos componentes para modelar los datos de calibración y un número variable de componentes inesperados para RBL. Los valores de recuperación promedio ($99\% \pm 7\%$), RMSE ($0.06 \mu\text{g mL}^{-1}$) y REP (5%) indican que los resultados alcanzados son satisfactorios (Tabla 5.1).

Con el fin de evaluar la precisión, se analizaron los coeficientes de variación (CV%). Para ello, se llevaron a cabo cinco determinaciones de la muestra número 5 (ver Tabla 5.1) en tres días diferentes. Los valores de CV% obtenidos se muestran en la Tabla 5.2. La prueba ANOVA muestra que las variancias son comparables para los tres días ($p > 0.05$).

A diferencia del modelado con MRC-ALS, en donde se soluciona la falta de bilinealidad de los datos MFST desdoblado las matrices, el modelado con U-PLS/RBL compensa parcialmente esta falta y explota la ventaja de segundo orden.

5.5.4. Cifras de mérito

Las cifras de mérito se utilizan regularmente con el objetivo de comparar varios métodos analíticos. La Tabla 5.2 muestra las cifras obtenidas cuando se aplicó MCR-ALS y U-PLS/RBL, calculadas como se detalló en la Sección 1.2.1.3 y 1.2.2.3 del Capítulo 1, respectivamente. Como puede observarse, la sensibilidad (SEN) para U-PLS/RBL es casi el doble que el valor obtenido para MCR-ALS. Este hecho puede deberse a que se mantiene la estructura de los datos de segundo orden cuando son modelados por el primer algoritmo. Lo mismo se aprecia para la sensibilidad analítica, γ , que puede ser de mayor utilidad para comparar métodos, ya que las unidades de intensidad de fluorescencia son arbitrarias. Ésta permite la comparación de técnicas analíticas independientemente de la técnica específica, el equipamiento o la escala empleada, y su inversa (γ^{-1}) establece la mínima diferencia de concentración estadísticamente discernible por la metodología. En el caso analizado, γ^{-1} es $0.005 \mu\text{g mL}^{-1}$ y $0.003 \mu\text{g mL}^{-1}$ para MCR-ALS y U-PLS/RBL, respectivamente.

Otras cifras de mérito importantes son los límites de detección (LOD) y de cuantificación (LOQ) mostradas en la Tabla 5.2. Se observa que hay una mejora en los LOD y LOQ resultantes para U-PLS/RBL en comparación con los valores

obtenidos para MCR-ALS. Esto es de esperar ya que estos términos están afectados por la sensibilidad.

Tabla 5.2. Cifras de mérito para MCR-ALS y U-PLS/RBL.

		MCR-ALS	U-PLS/RBL
SEN (AFU mL μg^{-1}) ^a		23	44
γ (mL μg^{-1}) ^b		192	370
γ^{-1} ($\mu\text{g mL}^{-1}$)		0.005	0.003
LOD ($\mu\text{g mL}^{-1}$) ^c		0.02	0.01
LOQ ($\mu\text{g mL}^{-1}$) ^d		0.06	0.03
CV (%) ^e	Día 1	8	7
	Día 2	9	7
	Día 3	8	8
ANOVA ^d		$F=0.69$ ($p=0.52$)	$F=0.51$ ($p=0.61$)

^a AFU: unidades arbitrarias de fluorescencia.

^b Sensibilidad analítica.

^c LOD, límite de detección calculado según y Bauza y col. (Bauza y col. 2012) para MCR-ALS y de acuerdo a Allegrini y Olivieri (Allegrini y Olivieri 2012) para U-PLS/RBL, considerando un 95% de probabilidad.

^d LOQ, límite de cuantificación calculado como $\text{LOD} \times (10/3.3)$.

^e CV, coeficiente de variación calculado para la muestra de validación número 5 de la Tabla 5.1 ($n = 5$) durante tres días consecutivos.

^d F crítico_{(3-1);(15-3);0.05} = 3.89.

Es importante remarcar que el procedimiento RBL obtiene una aproximación de las MFST reales correspondientes a la interferencia, ya que se basa en que los datos son bilineales (aplicación de SVD, un algoritmo de descomposición bilineal de las matrices). Sin embargo, esta aproximación es suficiente para alcanzar un error aceptable ($\text{REP} = 5\%$), que es similar al obtenido al modelar los datos desdoblados. Aunque se obtuvieron mejores cifras de mérito cuando se usó U-PLS/RBL, ambos algoritmos constituyen una buena alternativa para resolver muestras de plasma para la determinación de DOX cuando se analizan datos no bilineales.

5.6. Conclusiones

En este capítulo se demostró la posibilidad de modelar datos no bilineales obtenidos por fluorescencia sincrónica, explotando la ventaja de segundo orden, para la cuantificación de DOX en muestras de plasma y en presencia de un componente inesperado (RF). Se consideraron dos estrategias que incluyeron el desdoblamiento de los datos y: a) procesarlos con MCR-ALS como datos de primer orden, o b) procesarlos con PLS aprovechando la ventaja de segundo orden a través del procedimiento RBL. Ambas metodologías alcanzaron valores de REP = 5 %, lo que puede considerarse aceptable para muestras biológicas.



CAPÍTULO 6

**Ventaja de segundo orden para datos no bilineales.
Desarrollo de un nuevo algoritmo adaptado de
PLS/RBL**

6.1. Introducción

En el Capítulo 5 se describieron distintos ejemplos de datos que no cumplen con la propiedad de trilinealidad y diferentes propuestas para modelarlos. Uno de los casos es la presencia de perfiles diferentes para un componente en una única muestra, lo que genera datos no bilineales. Se publicaron sólo tres artículos en donde se modelan datos no bilineales para una única muestra (Zampronio y col. 2002, Kumar y Mishra 2011, Schenone y col. 2013b).

En el análisis de muestras complejas de composición desconocida con múltiples especies fluorescentes, además de resolver los espectros solapados, el algoritmo seleccionado debe ser capaz de modelar señales no bilineales. Aunque se propuso a U-PLS/RBL para cumplir con esta tarea (Escandar y col. 2011), el procedimiento RBL sólo actúa adecuadamente cuando los datos del componente inesperado son bilineales. Esto se debe a que RBL modela la señal del componente inesperado a través de la descomposición en valores singulares (Öhman y col. 1990). Esta propuesta puede alcanzar errores de predicción de hasta 20-30 % en los peores casos.

En el capítulo anterior se presentó una nueva estrategia para modelar matrices de fluorescencia sincrónica total siendo la única aplicación que explota la ventaja de segundo orden con este tipo de datos no bilineales con fines cuantitativos. El método desarrollado se aplicó a la determinación de doxorrubicina en plasma humano en presencia de riboflavina y los resultados obtenidos fueron satisfactorios.

En este capítulo se presenta un nuevo algoritmo para procesar datos no bilineales basado en cuadrados mínimos parciales seguido por el modelado de residuos usando MCR-ALS con la restricción de correlación (Goicoechea y col. 2010). Durante el paso de modelado residual, los perfiles correspondientes a las interferencias son extraídos por modelado de los datos desdoblados con MCR-ALS e incorporados en el paso de minimización que modifica los coeficientes de la regresión PLS. Para comparar la habilidad de predicción de la nueva metodología con la de los algoritmos convencionales U-PLS/RBL, MCR-ALS y PARAFAC, se evaluó un juego de datos simulados y el mismo sistema experimental estudiado en el capítulo anterior.

6.2. Objetivos

- Desarrollar un algoritmo que permita el modelado de datos no bilineales, incorporando el perfil del interferente obtenido mediante la aplicación de MCR-ALS en un paso posterior a U-PLS.
- Comparar los resultados obtenidos con el nuevo algoritmo con los arrojados por otros algoritmos (PARAFAC, MCR-ALS y U-PLS/RBL) para mostrar cómo la pérdida de bilinealidad afecta el modelado y la habilidad de predicción de estos algoritmos, en especial de PARAFAC.

6.3. Programas

Todos los cálculos se realizaron utilizando MATLAB 7.10 (MATLAB 7.10, 2010). PARAFAC y U-PLS/RBL se aplicaron a través de la interfaz gráfica MVC2 escrita en MATLAB por el Dr. Olivieri y colaboradores (Olivieri y col. 2009) y disponible en Internet (www.iquir-conicet.gov.ar/descargas/mvc2.rar). Para la aplicación de MCR-ALS con la restricción de correlación se utilizó un código MATLAB que puede solicitarse al Dr. Tauler (e-mail: Roma.Tauler@idaea.csis.es). Todas las simulaciones realizadas y el nuevo algoritmo U-PLS/RMNB se implementaron utilizando rutinas elaboradas por nuestro grupo de trabajo.

6.4. Materiales y Métodos

6.4.1. Datos simulados

La simulación de matrices MFST se llevó a cabo para ilustrar el comportamiento del nuevo algoritmo de segundo orden que llamamos U-PLS/RMNB (de cuadrados mínimos parciales con datos desdoblados seguido de un paso de modelado de residuos no bilineales) y comparar su rendimiento con otros algoritmos: PARAFAC (que requiere el cumplimiento de bilinealidad) y U-PLS/RBL (que puede compensar la pérdida de bilinealidad agregando un número de factores mayor). El sistema simulado se construyó considerando un único analito presente en las muestras de calibración. Para las muestras de validación se tuvo en cuenta la presencia del analito y un interferente. Por ello, una resolución adecuada del sistema requiere la ventaja de segundo orden. Las matrices MFST se generaron empezando por los perfiles sin ruido en las dimensiones de excitación y emisión que se muestran en la Figura 6.1A y B, respectivamente.

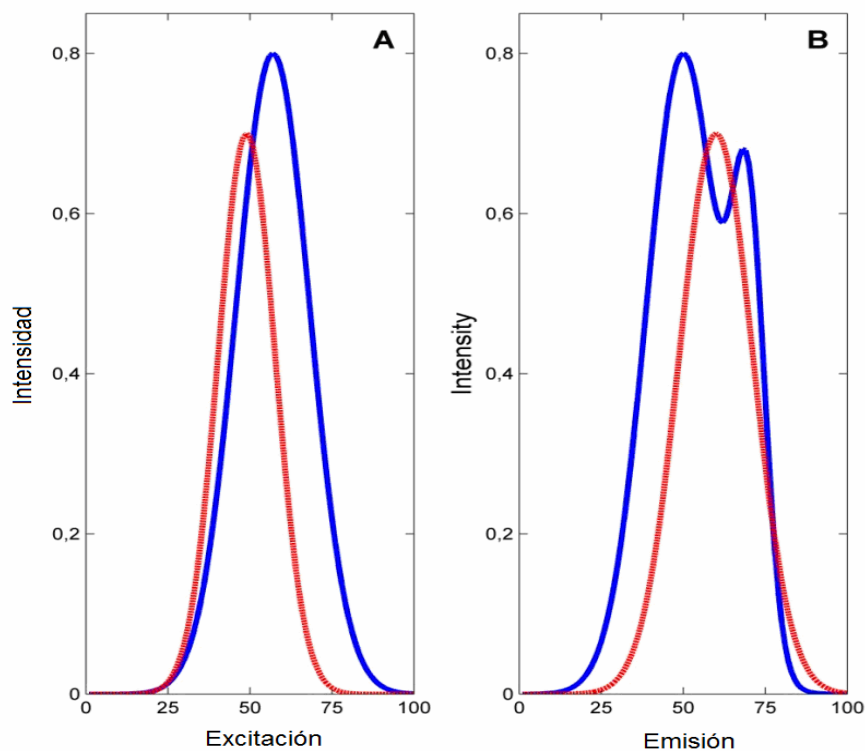


Figura 6.1. Espectros de excitación (A) y emisión (B) del analito (línea sólida azul) y del interferente (línea discontinua roja) en el sistema simulado.

Los mismos perfiles espectrales se utilizaron para generar la matriz de excitación-emisión de cada componente de la muestra. En la Figura 6.2 se puede ver el gráfico de contorno de una MEE correspondiente a una mezcla de ambos componentes. Las matrices construidas de esta manera son de dimensión 100×100 . Cada fila de la caja dibujada dentro del gráfico corresponde a un espectro de fluorescencia sincrónica (elementos diagonales de la matriz) hipotéticamente registrado a un valor de $\Delta\lambda$ dado.

Se construyó un juego de calibración de cinco matrices que contienen sólo el analito en concentraciones nominales de 1.0, 2.0, 3.0, 4.0, 5.0 (en unidades arbitrarias, a.u.). El tamaño de cada matriz fue de 25×50 (50 puntos correspondientes a la dimensión espectral y 25 a la dimensión $\Delta\lambda$). Además, se creó un juego de validación de nueve muestras con concentraciones de analito e interferente según la Tabla 6.1, siguiendo un diseño central compuesto.

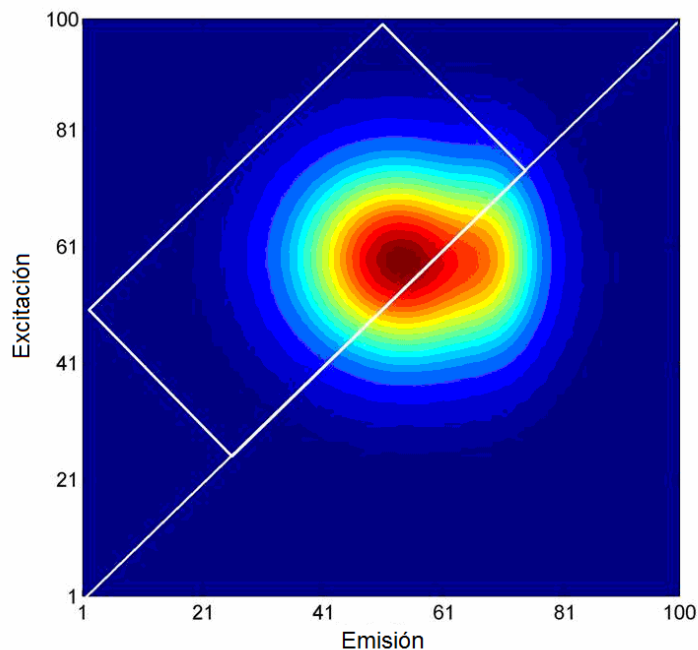


Figura 6.2. Matriz de excitación-emisión de una mezcla de analito e interferente en el sistema simulado. Las filas de la caja contienen los espectros sincrónicos simulados a diferentes $\Delta\lambda$.

La Figura 6.3 muestra los espectros sincrónicos simulados obtenidos para ambos compuestos a cada uno de los $\Delta\lambda$. La tendencia en los espectros es la misma que se observó en los datos experimentales, por lo que los cambios en los perfiles a diferentes $\Delta\lambda$ hacen que las matrices de fluorescencia sincrónica total sean un caso característico de datos no bilineales.

Para una mayor aproximación a la realidad, se adicionó ruido a los espectros simulados siguiendo una distribución Gaussiana con una desviación estándar de 1 % de la señal máxima de calibración.

Los datos de segundo orden no bilineales de cada una de las 9 muestras de validación se unieron con aquellas del juego de calibración, y cada conjunto de 6 matrices se utilizó para construir modelos con los métodos de calibración multivariada U-PLS/RBL, PARAFAC y el nuevo U-PLS/RMNB. Las predicciones se utilizaron para realizar un análisis estadístico y comparaciones futuras.

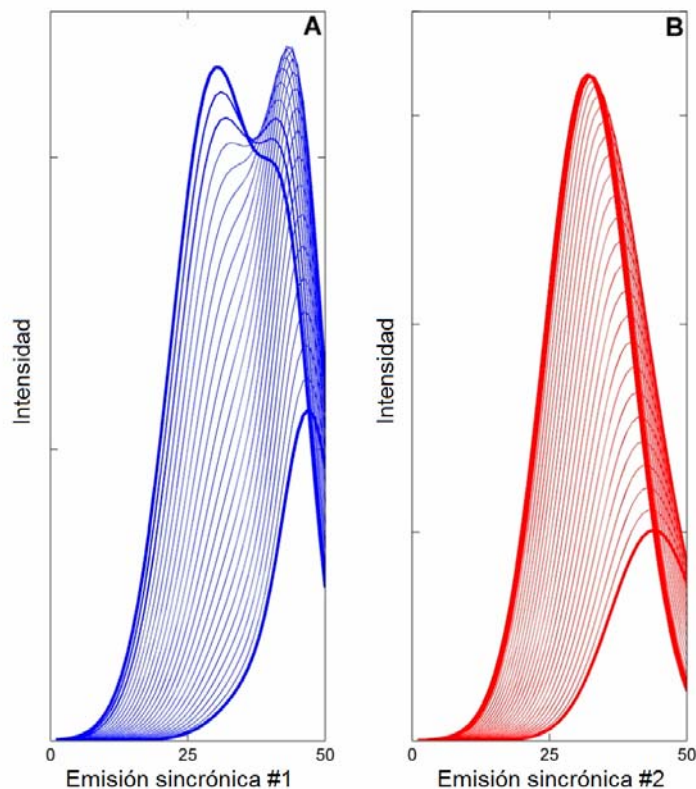


Figura 6.3. Espectros sincrónicos del analito (azul) y el interferente (rojo) generados a 25 $\Delta\lambda$ diferentes.

6.4.2. Datos experimentales

En este capítulo se utilizaron los datos obtenidos en el sistema analizado en el Capítulo 5 para la determinación de doxorrubicina (DOX) en muestras de plasma de composición desconocida.

6.5. Desarrollo teórico del nuevo algoritmo U-PLS/RMNB

El nuevo algoritmo intenta superar la dificultad de RBL para modelar datos no bilineales. La idea básica de RMNB es similar a la del método clásico RBL, que es minimizar la norma del vector de los residuos \mathbf{e}_u obtenido a través del ajuste de las contribuciones relevantes de los datos de calibración y de las interferencias, pero la diferencia principal es el uso de MCR-ALS. Este algoritmo se utiliza para extraer el vector que contiene el perfil de cada interferente de una matriz aumentada \mathbf{D} [de tamaño $(I+1) \times JK$] construida tanto con las I matrices desdobladas del juego de calibración $\mathbf{x}_{c,i}^T$ (de tamaño $1 \times JK$) como con las matrices desdobladas de las muestras incógnitas \mathbf{x}_u^T (de tamaño $1 \times JK$):

$$\mathbf{D} = \left[\mathbf{x}_u^T \mid \mathbf{x}_{c,1}^T \mid \mathbf{x}_{c,2}^T \mid \dots \mid \mathbf{x}_{c,I}^T \right] \quad (6.1)$$

La matriz aumentada \mathbf{D} se descompuesta por MCR, bajo ciertas restricciones:

$$\mathbf{D} = \mathbf{C} \mathbf{S}^T + \mathbf{E}_{\text{MCR}} \quad (6.2)$$

donde \mathbf{C} [de tamaño $(I+1) \times N$, donde N es el número de componentes estimados por PCA que debería ser igual al número de componentes calibrados más el número de agentes interferentes] contiene los *scores*, y \mathbf{S}^T (de tamaño $N \times JK$) contiene los perfiles (espectros sincrónicos obtenidos a diferentes intervalos concatenados uno a continuación del otro) de los componentes calibrados y los interferentes. En el caso de que exista un solo analito y un interferente:

$$\mathbf{S} = \left[\mathbf{s}_{\text{int}} \mid \mathbf{s}_c \right] \quad (6.3)$$

donde \mathbf{s}_{int} ($JK \times 1$) y \mathbf{s}_c ($JK \times 1$) son los vectores que contienen los perfiles del interferente y el analito calibrado, respectivamente (ver Figura 6.4).

La posibilidad de aplicar restricciones para mejorar el modelado de los perfiles en \mathbf{C} y \mathbf{S}^T se realiza durante la optimización mediante ALS. Como ya se detalló en la Sección 1.2.1 del Capítulo 1, se puede aplicar la restricción de no-negatividad en ambos perfiles. Otra posibilidad, que se exploró en el capítulo anterior de esta tesis y será utilizada en este apartado nuevamente, es la de aplicar la restricción de correlación para obtener valores de concentración dentro de la escala de los valores reales de concentración (Goicoechea y col. 2010). En cada ciclo iterativo de la resolución, esta restricción lleva a cabo un modelo de regresión entre las concentraciones calculadas por MCR-ALS (en unidades arbitrarias) y las concentraciones de referencia que son conocidas a partir de las muestras de calibración. El modelo de regresión establecido es luego utilizado para actualizar las concentraciones en \mathbf{C} a la escala de los valores verdaderos. Al final de este paso, se obtiene una correlación directa entre los valores de concentración de la resolución y las concentraciones reales en las muestras.

En el paso final, la contribución del agente interferente, extraído de la matriz \mathbf{S}^T y su valor de concentración correspondiente, es modelada en el procedimiento RMNB según la siguiente expresión:

$$\mathbf{x}_u = \mathbf{P} \mathbf{t}_{u\text{RMNB}} + c_{\text{int}} \mathbf{s}_{\text{int}} + \mathbf{e}_u \quad (6.4)$$

donde c_{int} es el *score* de concentración del interferente en la muestra de validación.

En la Figura 6.4 se ilustra la descomposición de una matriz construida con los datos extraídos de un ejemplo simulado, y compuesta por cinco matrices de calibración desdobladas y una muestra incógnita. Los vectores correspondientes a cada MFTS para el interferente (\mathbf{s}_{int}) y su *score* (c_{int}) son obtenidos por MCR-ALS.

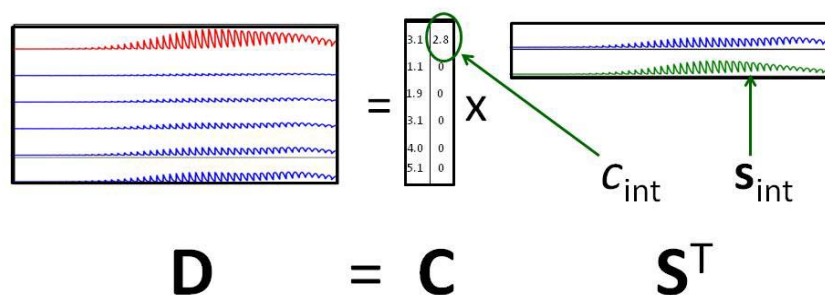


Figura 6.4. Representación esquemática de la aplicación de MCR-ALS con la restricción de correlación para la descomposición de una matriz simulada. La matriz resulta del acoplamiento de cinco matrices MFTS de calibración y una de validación desdobladas. El procedimiento se usa para generar ambos perfiles y la concentración del interferente, necesarios para minimizar la Ecuación (6.4).

De manera similar a la descrita para RBL, la minimización de $\|\mathbf{e}_u\|$ con el procedimiento RMNB se lleva a cabo manteniendo constante \mathbf{P} en los valores de calibración y variando $\mathbf{t}_{u\text{RMNB}}$. Una vez minimizado $\|\mathbf{e}_u\|$ de la Ecuación 6.4, la sustitución del valor final del vector $\mathbf{t}_{u\text{RMNB}}$ en la Ecuación (1.23) (ver Sección 1.2.2.2 del Capítulo 1) que se muestra nuevamente a continuación, provee la concentración del analito en la muestra.

$$y_u = \mathbf{t}_{u\text{RMNB}}^T \mathbf{v} \quad (6.5)$$

El proceso completo se esquematiza en la Figura 6.5.

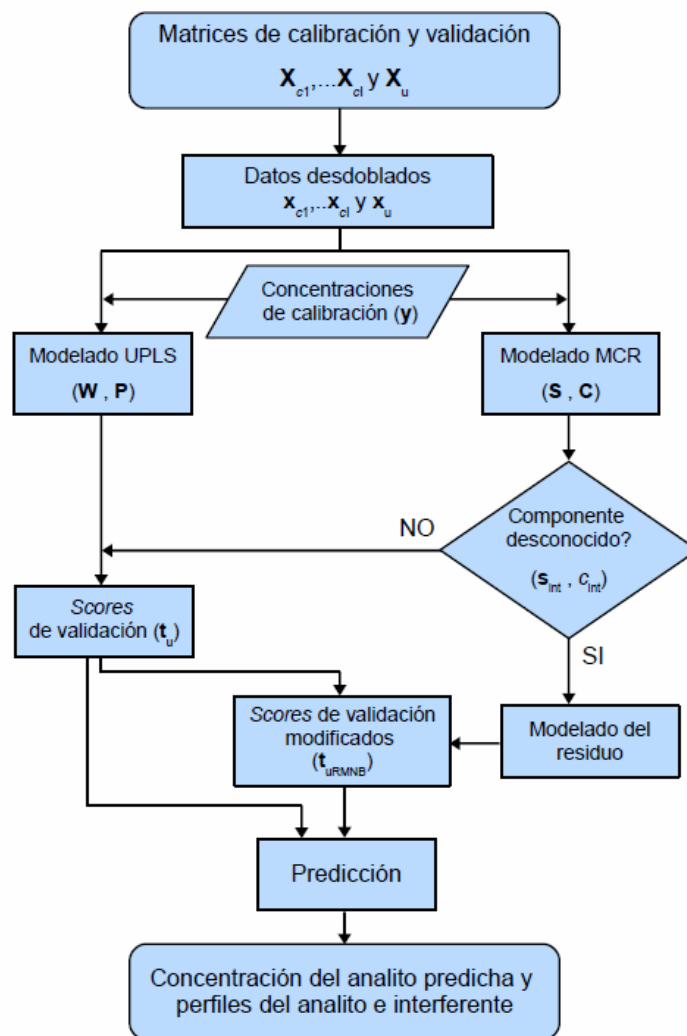


Figura 6.5. Esquema ilustrativo de cómo trabaja el nuevo algoritmo U-PLS/RMNB. Ver el texto para la explicación de los símbolos.

6.6. Resultados y Discusión

6.6.1. Datos Simulados

El solapamiento severo que existe entre los espectros sincrónicos del analito y el interferente en todos los valores de $\Delta\lambda$ (ver Figura 6.3) exige la aplicación de un algoritmo con ventaja de segundo orden. La Tabla 6.1 resume los resultados obtenidos con PARAFAC, U-PLS/RBL y U-PLS/RMNB.

La recuperaciones obtenidas para las nueve muestras de validación al utilizar PARAFAC son muy pobres (REP % > 100) y esto se debe a la incapacidad del algoritmo para modelar estructuras no bilineales. El número de componentes

(2) utilizado para construir el modelo de calibración se obtuvo analizando los residuos del ajuste por cuadrados mínimos del arreglo de tres vías.

Tabla 6.1. Resultados obtenidos al modelar los datos simulados con PARAFAC, U-PLS/RBL y U-PLS/RMNB

Muestra	Analito ^a	Interferente ^a	PARAFAC ^{a,b}	U-PLS/RBL ^{a,c}	U-PLS/RMNB ^{a,d}
1	1.59	1.59	3.31	2.66 (3)	1.50
2	4.41	1.59	6.21	4.72 (3)	4.33
3	3.00	3.00	6.45	4.18 (3)	2.82
4	1.00	3.00	4.38	2.01 (3)	0.82
5	4.41	4.41	9.61	6.55 (3)	4.16
6	5.00	3.00	8.45	7.30 (3)	4.83
7	1.59	4.41	6.76	3.20 (3)	1.32
8	3.00	1.00	4.11	3.98 (3)	2.95
9	3.00	5.00	8.93	5.01 (4)	2.71
REP %	-	-	127.1	51.1	6.4

^a En unidades arbitrarias

^b El número de componentes se eligió por análisis de los residuos del ajuste por cuadrados mínimos del arreglo de tres vías al modelo trilineal ($N = 2$).

^c El valor entre paréntesis corresponde al número de interferentes considerados para estabilizar el error s_u .

^d Un factor extraído por MCR-ALS (ver Ecuación (6.2)).

En la aplicación de U-PLS/RBL, el número de variables latentes A se obtuvo a través del criterio de validación cruzada (Haaland y Thomas, 1988). Debido a que en nuestro caso el analito fue el único componente presente en las muestras de calibración, A fue igual a 1. La contribución de los componentes inesperados (N_{int}) se evaluó aplicando el método RBL a las muestras de validación. El análisis de los residuos s_u obtenidos por este procedimiento se llevó a cabo según la Ecuación 1.30 (ver Sección 1.2.2.2 del Capítulo 1). La Tabla 6.2 muestra la variación de los residuos y las predicciones obtenidas para la muestra 1. Las cifras tabuladas indican que la magnitud de los residuos se estabiliza en $N_{\text{int}} = 3$. Considerando que

las muestras de validación presentan un solo interferente, la necesidad de introducir un mayor número de componentes para mejorar el modelado de su contribución es evidente. De todas maneras, ninguno de los distintos números de componentes inesperados utilizados arrojó predicciones o recuperaciones aceptables al analizar la muestra 1 con U-PLS/RBL.

Tabla 6.2. Número de componentes inesperados, residuos y concentraciones predichas del analito en la muestra de validación 1 (concentración nominal igual a 1.59) cuando se aplica U-PLS/RBL.

N_{int}	Residuo antes de RBL	Residuo luego de RBL	Concentración predicha	Rec (%) ^a
0	0.140	-	3.33	209
1	0.140	0.034	3.25	204
2	0.140	0.018	3.18	200
3	0.140	0.011	2.66	167
4	0.140	0.011	2.89	182
5	0.140	0.011	2.88	181

^a Recuperación (%) = (predicho/nominal) × 100.

Las predicciones de las nueve muestras de validación mediante U-PLS/RBL se detallan en la Tabla 6.1. Aunque los resultados son mejores que los obtenidos por PARAFAC, las concentraciones predichas están lejos de ser satisfactorias. El problema principal con este algoritmo resulta de la dificultad de introducir el número correcto de especies interferentes, que se explica mejor en la Figura 6.6.

La Figura 6.6A muestra la superficie correspondiente al primer perfil extraído por SVD cuando se aplicó el procedimiento RBL a la muestra 1. La estructura bilineal de la matriz se ilustra en la figura 6.6B, en donde todos los espectros extraídos comparten el mismo espectro de emisión. Además, se observan valores negativos incluso sin centrar los datos. La Figura 6.6C y D ilustran los resultados obtenidos a través de MCR-ALS cuando se modelaron las cinco muestras de calibración junto con la muestra 1. La no bilinealidad de los

datos se evidencia por el hecho de que todos los espectros de emisión no comparten el mismo perfil.

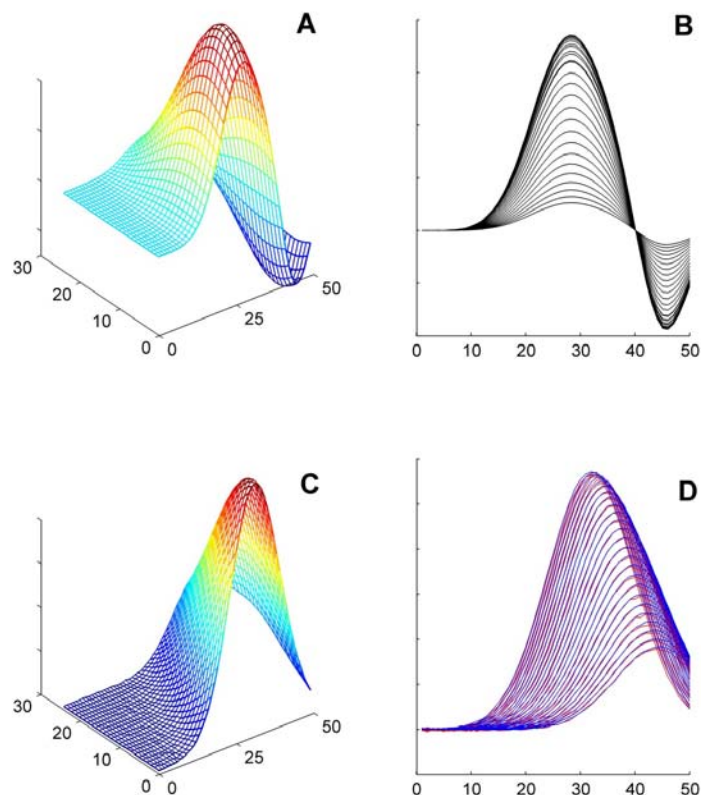


Figura 6.6. (A) Primer perfil extraído por RBL al ser aplicado a la muestra 1. (B) Juego de perfiles extraídos que muestran la estructura bilineal de la matriz. (C) Matriz reordenada que se obtuvo al aplicar MCR-ALS para modelar las cinco muestras de calibración junto con la correspondiente a la muestra 1. (D) Juego de perfiles extraídos por MCR-ALS que muestran la estructura no bilineal de la matriz (azul) y los espectros normalizados reales (rojo).

Resulta interesante señalar la similitud de los perfiles en la Figura 6.6D con los espectros del interferente que se muestran en la Figura 6.3B. Esta semejanza es indicativa de la capacidad del procedimiento MCR-ALS para extraer con exactitud los perfiles de la interferencia. La matriz que se muestra en la Figura 6.6C se desdobló y junto con su correspondiente *score* se introdujeron en la Ecuación (6.4), para obtener el nuevo vector de *scores* PLS (\mathbf{t}_{URMNB}) que se empleó subsecuentemente en la Ecuación (6.5) para calcular una concentración más exacta del analito en la muestra de validación 1.

Los resultados de predicción obtenidos al aplicar el algoritmo U-PLS/RMNB a las nueve muestras de validación simuladas se ubican en la última columna de la Tabla 6.1. Los mejores valores alcanzados en todos los casos representan la capacidad superior del algoritmo para modelar la interferencia no bilineal.

6.6.2. Datos Experimentales

La determinación de DOX en muestras de plasma de composición desconocida a través de espectroscopia FST y modelado mediante U-PLS/RBL se describió en el capítulo anterior de esta tesis. En esta sección, se comparan los resultados obtenidos por este algoritmo, MCR-ALS, PARAFAC y U-PLS/RMNB.

La Tabla 6.3 resume los resultados obtenidos con MCR-ALS, PARAFAC, U-PLS/RBL y U-PLS/RMNB. Aunque PARAFAC provee resultados considerablemente mejores que los reportados en la Tabla 6.1, siguen siendo los menos satisfactorios (REP%=10.8) en comparación con el resto de los algoritmos utilizados. El número de componentes (4) empleado para construir el modelo de calibración se obtuvo como se describió anteriormente para los datos simulados. En el caso de U-PLS/RBL los cálculos también fueron los mismos que se mencionaron previamente. Debido a que las muestras de calibración se prepararon adicionando DOX a muestras de plasma, el número de variables latentes (A) fue 2. Los comportamientos de N_{int} y s_u fueron similares a los observados para el sistema simulado. MCR-ALS se aplicó como se describió en el capítulo anterior. U-PLS/RBL arrojó un mejor REP% que PARAFAC cuando se utilizaron dos componentes para modelar los datos de calibración y un número variable de componentes inesperados (ver Tabla 6.3). Estos valores son aceptables y se deben a la simplicidad del sistema investigado. DOX y RF muestran un menor solapamiento espectral que los espectros simulados analizados anteriormente.

Tabla 6.3. Resultados de predicción del juego experimental.

Muestra	DOX (mg L ⁻¹)	RF (mg L ⁻¹)	DOX Predicha (mg L ⁻¹)			
			MCR-ALS	PARAFAC ^a	U-PLS/RBL ^b	U-PLS/RMNB ^c
1	0.50	0.19	0.42	0.67	0.43 (2)	0.48
2	1.78	0.27	1.72	1.70	1.73 (1)	1.75
3	1.25	0.30	1.28	1.36	1.28 (2)	1.26
4	2.00	0.19	1.95	1.89	1.91 (2)	2.01
5	1.25	0.19	1.33	1.34	1.29 (2)	1.28
6	0.72	0.27	0.76	0.90	0.71 (2)	0.71
7	1.25	0.08	1.38	1.40	1.37 (3)	1.30
8	0.72	0.11	0.67	0.91	0.69 (1)	0.75
9	1.78	0.11	1.77	1.71	1.80 (1)	1.73
REP%	-	-	5.5	10.8	4.9	2.4

^a El número de componentes se fijó a través del análisis de los residuos del ajuste de cuadrados mínimos del arreglo de tres vías al modelo trilineal ($A = 4$).

^b El valor entre paréntesis corresponde al número de interferentes que se consideraron para estabilizar el error s_u .

^c Un factor extraído por MCR-ALS (ver Ecuación (6.2)).

Los mejores resultados se obtuvieron con U-PLS/RMNB, que arrojó un valor de REP% de 2.4. La razón por la que este algoritmo tiene un mejor desempeño se ilustra en la Figura 6.7, donde se comparan los gráficos de contorno FST de la especie interferente (RF) obtenidos vía RBL (A) y MCR-ALS (B) con el contorno real de RF (C). Teniendo en cuenta que el contorno obtenido por MCR-ALS es el que se incorpora en el paso de minimización del nuevo algoritmo, se atribuye el mejor rendimiento de U-PLS/RMNB a su habilidad superior para modelar los perfiles de las interferencias.

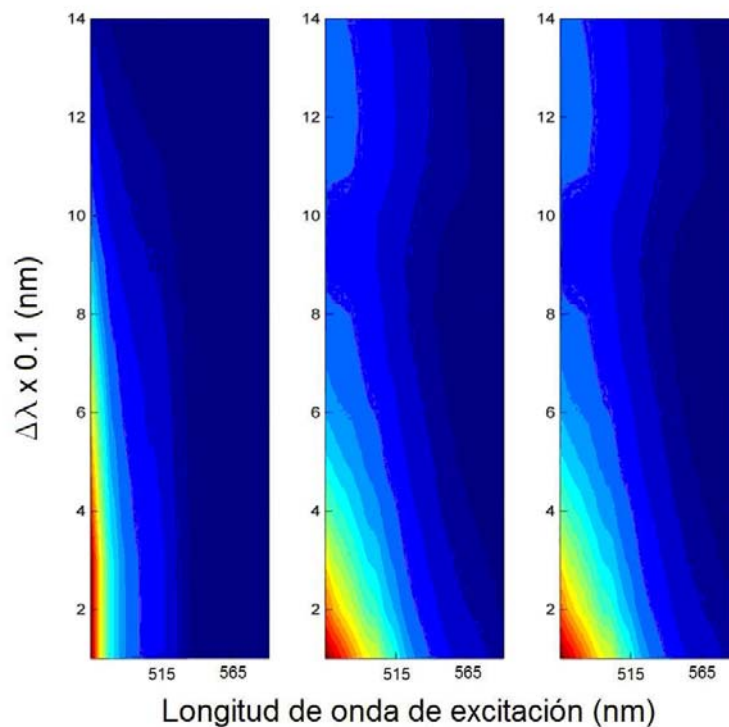


Figura 6.7. (A) Contorno correspondiente a la primer MFST estimado por RBL. (B) Perfil extraído por MCR-ALS para ser incorporado en el proceso de minimización en el paso RMNB. (C) matriz MFST real de RF (interferente).

6.7. Conclusiones

En este capítulo se demostró la obtención de la ventaja de segundo orden al modelar datos con un fuerte solapamiento espectral y estructura no bilineal con el nuevo algoritmo que combina U-PLS con MCR-ALS. La habilidad de predicción del nuevo método - U-PLS/RMNB - se demostró tanto para datos simulados como experimentales. Además, se llevó a cabo una comparación de los resultados con los obtenidos con otros métodos de calibración de segundo orden ampliamente estudiados, MCR-ALS, PARAFAC y U-PLS/RBL. En todos los sistemas analizados, el método presentado en este capítulo logró mejores predicciones que las metodologías establecidas.



CAPÍTULO 7

Conclusiones generales

De los objetivos planteados al inicio de este trabajo de tesis y después de llevado a cabo todo lo descripto en los capítulo anteriores, se desprenden las siguientes conclusiones generales y otras particulares de las diferentes aplicaciones.

El desarrollo y la validación de nuevas metodologías analíticas se llevaron a cabo mediante la combinación de instrumentación sencilla y herramientas quimiométricas que permitió la determinación de distintos compuestos de interés biológico presentes en muestras de composición compleja. La implementación satisfactoria de dichos métodos hizo posible que se abordaran problemas relacionados con la separación incompleta de componentes en una mezcla, la presencia de especies no modeladas y/o efecto matriz. Por otra parte, se demostró que los algoritmos utilizados y las estrategias empleadas son de mucha utilidad cuando los datos no cumplen con la propiedad de trilinealidad, sin la necesidad de aplicar técnicas de pre-tratamiento a las señales obtenidas.

Se demostró que MCR-ALS puede ser aplicado a datos de segundo orden no trilineales, específicamente, cuando se analizan datos cromatográficos que presentan corrimiento de picos. El algoritmo se aplicó para determinar tres colorantes sintéticos presentes en bebidas no alcohólicas con la obtención de datos a través de un método cromatográfico rápido arrojando resultados satisfactorios. El rendimiento de U-PLS/RBL también se evaluó, pero en este caso se construyeron los modelos con mezclas de los tres colorantes. Además, para analizar la mayoría de las muestras reales se debió considerar dos o tres interferentes para alcanzar valores de residuos comparables con el ruido instrumental. Este hecho demuestra que MCR-ALS posee una mayor habilidad para predecir muestras en casos donde los datos presentan fuertes desvíos de la trilinealidad. Por otro lado, se redujo notablemente el tiempo de análisis y el uso de solventes. Este hecho se debe tener seriamente en cuenta ya que el ahorro de solvente tiene un fuerte impacto en el ambiente. En términos generales, este tipo de análisis se puede sugerir para el control de calidad de estos aditivos en muestras de composición compleja.

Dado que U-PCA/RBL aparenta ser el algoritmo más aplicable para procesar datos de segundo orden no lineales, se utilizó para determinar tartrazina a través de un sistema FIA con flujo interrumpido y se compararon los resultados de este

algoritmo con los obtenidos por MCR-ALS. Los datos se obtuvieron al registrar la reacción cinética entre el colorante y el oxidante (bromato de potasio), en presencia de Fe(II). Debido a la relación no lineal entre la concentración de tartrazina y los *scores* obtenidos luego de procesar los datos, se utilizó una función polinomial para la predicción del analito en presencia de otros colorantes (azul brillante y amarillo ocaso). Los parámetros de calidad obtenidos indican que MCR-ALS resulta una metodología apta para la obtención de la ventaja de segundo orden en aquellos casos donde los componentes de una muestra poseen espectros similares y existe una relación no lineal entre la señal del analito y la concentración. Sin embargo, la cantidad de tartrazina en presencia de azul brillante como interferente no pudo ser cuantificada con exactitud al aplicar U-PCA/RBL, mostrando una limitación del método. El análisis de muestras de bebidas reales con MCR-ALS arrojó resultados satisfactorios y estadísticamente comparables con aquellos obtenidos por CLAR, demostrando ser una alternativa rápida, exacta y económica a las técnicas separativas.

Se presentó una estrategia novedosa que combina PARAFAC y transferencia de calibración cuyo objetivo fue reducir el número de muestras a preparar cuando se presentan interferentes de manera conjunta con efecto de filtro interno, sin la necesidad de modelar la presencia del agente causante del efecto mencionado. Para ilustrar el desempeño del método propuesto se evaluaron datos simulados y experimentales y se puso énfasis en la demostración estadística completa de la aplicabilidad y robustez de la metodología propuesta. La estrategia utiliza sólo dos adiciones de estándar a la muestra y posee la ventaja adicional de mantener la sensibilidad de la curva de calibrado de estándares puros. Los resultados se compararon con aquellos obtenidos luego de aplicar el método de adición estándar combinado con PARAFAC. La estrategia presentada disminuye el trabajo experimental manteniendo cifras de mérito aceptables y se demostró que la sensibilidad analítica se aumenta al ser eliminado el efecto de filtro interno sobre la señal del analito en la muestra al aplicar PDS. Esto nos permitió concluir que la estrategia PDS/PARAFAC constituye un método simple y de bajo costo para la determinación de fenilefrina en presencia de otras drogas en muestras de agua con una reducción considerable de tiempo. Esta metodología puede extenderse a otros sistemas que presenten efecto matriz y, consecuentemente, puede ser una

herramienta útil para determinar la cantidad de fármacos presentes en ambientes acuáticos y para evaluar el efecto de las plantas de tratamiento de efluentes convencionales en la eliminación de compuestos farmacéuticos.

Por otro lado, se demostró la posibilidad de modelar datos no bilineales obtenidos por espectroscopia de fluorescencia sincrónica total obteniendo la ventaja de segundo orden para la cuantificación de doxorrubicina en muestra de plasma en presencia de un componente inesperado (riboflavina). Una de las características más importantes de este trabajo es que se obtuvo por primera vez la ventaja de segundo orden con datos no bilineales. Para este fin, se consideraron dos estrategias que incluyeron el desdoblamiento de los datos y posterior procesamiento con MCR-ALS como datos de primer orden, o con U-PLS/RBL. En cuanto a la aplicación de MCR-ALS, se demostró que es posible la obtención de la ventaja de segundo orden utilizando y procesando datos de primer orden, especialmente cuando se aplica la restricción de correlación. Ambas metodologías alcanzaron resultados aceptables para muestras biológicas. Las metodologías desarrolladas podrían ser implementadas en la determinación de doxorrubicina en estudios de farmacocinética y biodisponibilidad de medicamentos.

Finalmente, se desarrolló un nuevo algoritmo que modela datos no bilineales denominado cuadrados mínimos parciales con modelado residual de datos no bilineales (U-PLS/RMNB) donde los perfiles correspondientes a las interferencias son extraídos por modelado de los datos desdoblados con MCR-ALS e incorporados en el paso de minimización que modifica los coeficientes de la regresión PLS. Su habilidad predictiva se demostró analizando tanto datos simulados como experimentales. Los resultados se compararon con los obtenidos al aplicar MCR-ALS, U-PLS/RBL y PARAFAC. U-PLS/RMNB no sólo fue capaz de obtener la ventaja de segundo orden para el caso donde los datos son no bilineales, si no que también arrojó un error de predicción menor a los obtenidos por los algoritmos recién mencionados. Los buenos resultados sugieren que el nuevo algoritmo podría ser extendido al análisis de otro tipo de datos no bilineales, como es el caso de espectroscopía de masa en tándem, para la determinación de uno o varios analitos en muestras complejas en presencia de interferencias.

En el transcurso de esta tesis se evaluaron diferentes temas de interés: las necesidades del análisis ambiental y biológico asociado al desarrollo de nuevos

métodos analíticos, el origen de la señal analítica y el uso de herramientas quimiométricas. La selección de las metodologías analíticas se basó en la mejora de la exactitud, precisión, sensibilidad y/o selectividad usando diferentes estrategias. También se tuvo en cuenta el impacto ambiental, la simplicidad, la factibilidad del análisis en simultáneo, el menor tiempo de análisis y la disminución de los volúmenes de solventes utilizados, que lleva a una disminución de los costos.

A través de la validación de las metodologías propuestas, el cálculo de cifras de mérito y la aplicación de pruebas estadísticas apropiadas se demostró que es posible desarrollar y proponer métodos alternativos a los oficiales y obtener resultados comparables.



CAPÍTULO 8

Bibliografía

Al Degs, Y. S. (2009). *Determination of three dyes in commercial soft drinks using HPLC/GO and liquid chromatography*. Food Chem. 117: 485–490.

Allegrini, F. y Olivieri, A. C. (2012). *Analytical Figures of Merit for Partial Least-Squares Coupled to Residual Multilinearization*. Anal Chem 84: 10823–10830.

Anastas, P. y Eghbali, N. (2010). *Green Chemistry: Principles and Practice*. Chemical Society Reviews 39: 301–312.

Andersen, A.; Warren, D. J. y Slordal, L. (1993). *A sensitive and simple high-performance liquid chromatographic method for determination of doxorubicin and its metabolites in plasma*. Ther. Drug Monit. 15: 455–461.

Anmat, *Boletín para profesionales*, Vol. X, Nº 3-4, 33-64. Agosto de 2002.

Antunes, M. C.; Simao, J. E. J.; Duarte, A. C. y Tauler, R. (2002). *Multivariate curve resolution of overlapping voltammetric peaks: quantitative analysis of binary and quaternary metal mixtures*. Analyst 127: 809–817.

Arancibia, J. A.; Nepote, J. A.; Escandar, G. M. y Olivieri, A. C. (2000). *Spectrofluorimetric determination of phenylephrine in the presence of a large excess of paracetamol*. Anal. Chim. Acta 419: 159–168.

Arancibia, J. A.; Olivieri, A. C.; Bohoyo Gil, D.; Muñoz de la Peña, A.; Durán-Merás, I. y Espinosa Mansilla, A. (2006). *Trilinear least-squares and unfolded-PLS coupled to residual trilinearization: New chemometric tools for the analysis of four-way instrumental data*. Chemom. Intell. Lab. Syst. 80: 77–86.

Araújo, A. N.; Catita, J. A. M. y Lima, J. L. F. C. (1998). *Kinetic determination of uric acid in urine based on single-line flow-system with multi-site detection*. Anal. Sci. 14: 809–813.

Asfaw, N.; Chebude, Y.; Ejigu, A.; Hurisso, B. B.; Licence, P.; Smith, R. L.; Tang, S. L. Y. y Poliakov, M. (2011). *The 13 Principles of Green Chemistry and Engineering for a Greener Africa*. Green Chem. 13: 1059–1060.

Azzouz, T. y Tauler, R. (2008). *Application of multivariate curve resolution alternating least squares (MCR-ALS) to the quantitative analysis of pharmaceutical and agricultural samples*. Talanta 74: 1201–1210.

Bahram, M. y Bro, R. (2007). *A novel strategy for solving matrix effect in three-way data using parallel profiles with linear dependencias*. Anal. Chim. Acta 584: 397–402.

Bauza, M. C.; Ibañez, G. A.; Tauler, R. y Olivieri, A. C. (2012). *Sensitivity Equation for Quantitative Analysis with Multivariate Curve Resolution-Alternating Least-Squares: Theoretical and Experimental Approach*. Anal. Chem. 84: 8697–8706.

Barceló, D. (2003). *Emerging pollutants in water analysis*. Trends Anal. Chem. 22: XIV-XVI.

Benjamin, R. S.; Riggs, C. E. Jr. y Bachur, N. R. (1973). *Pharmacokinetics and metabolism of adriamycin in man*. Clin. Pharmacol. Ther. 14: 592–600.

Berzas, J. J.; Rodríguez Flores, J.; Villaseñor Llerena, M. J. y Rodríguez, Fariñas. N. (1999). *Spectrophotometric resolution of ternary mixtures of tartrazine, patent blue V and indigo carmine in commercial products*. Anal. Chim. Acta 391: 353–364.

Bohoyo Gil, D.; Muñoz de la Peña, A.; Arancibia, J. A.; Escandar, G. M. y Olivieri, A. C. (2006). *Second-order advantage achieved by unfolded-partial least-squares/residual bilinearization modelling of excitation-emission fluorescence data presenting inner filter effects*. Anal. Chem. 78: 8051–8058.

Booksh, K. S. y Kowalski, B. R. (1994). *Theory of analytical chemistry*. Anal. Chem. 66: 782A–791A.

Bortolato, S. A.; Arancibia, J. A. y Escandar, G. M. (2009). *Non-trilinear chromatographic time retention–fluorescence emission data coupled to chemometric algorithms for the simultaneous determination of ten polycyclic aromatic hydrocarbons in the presence of interferences*. Anal. Chem. 81: 8074–8084.

Bortolato, S. A.; Arancibia, J. A.; Escandar, G. M.; Olivieri, A. C. (2007). *Improvement of residual bilinearization by particle swarm optimization for achieving the second-order advantage with unfolded partial least-squares*. J. Chemom. 21: 557–566.

Bortolato, S. A.; Arancibia, J. A.; Escandar, G. M.; Olivieri, A. C. (2010). *Time-alignment of bidimensional chromatograms in the presence of uncalibrated interferences using parallel factor analysis: Application to multi-component determinations using liquid-chromatography with spectrofluorimetric detection*. Chemom. Intell. Lab. Syst. 101: 30–37.

Brereton, R. G. (2007). *Applied Chemometrics for Scientists*. John Wiley & Sons Ltd, Chichester, England.

Bro, R. (1996). *Multway calibration. Multilinear PLS*. J. Chemom. 10: 47–61.

Bro, R. (1997). *PARAFAC: Tutorial & applications*. Chemom. Intell. Lab. Syst. 38: 149–171.

Bro R. (1998). Tesis Doctoral, Universidad de Amsterdam, Países Bajos.

Bro, R. y Kiers, H. A. L. (2003). *A new efficient method for determining the number of components in PARAFAC models*. J. Chemom. 17: 274–286.

Brown, S. D. (1991). *Rapid parameter estimation with incomplete chemical calibration models*. Chemom. Intell. Lab. Syst. 10: 87–105.

Cañada-Cañada, F.; Arancibia, J. A.; Escandar, G. M.; Ibañez, G. A.; Espinosa Mansilla, A.; Muñoz de la Peña, A. y Olivieri, A. C. (2009). *Second-order multivariate calibration procedures applied to high-performance liquid chromatography coupled to fast-scanning fluorescence detection for the determination of fluoroquinolones*. J. Chromatogr. A 1216: 4868–4876.

Cañada-Cañada, F.; Espinosa-Mansilla, A.; Muñoz de la Peña, A.; Jiménez Girón, A. y González-Gómez, D. *Determination of danofloxacin in milk combining second-order calibration and standard addition method using excitation-emission fluorescence data*. (2009). Food Chem. 113: 1260–1265.

Carroll, J. D. y Chang, J. (1970). *Analysis of individual differences in multidimensional scaling via an N-way generalization of "Eckart-Young" decomposition*. Psychometrika 35: 283–319.

Castells, R. C. y Castillo, M. A. (2000). *Systematic errors: detection and correction by means of standard calibration, Youden calibration and standard additions method in conjunction with a method response model*. Anal. Chim. Acta 423: 179–185.

Chen, C. L.; Thoen, K. K. y Uckun, F. M. (2001). *High-performance liquid chromatographic methods for the determination of topoisomerase II inhibitors*. J. Chromatogr. B: Biomed. Appl. 764: 81–119.

Chen, Z. P.; Wu, H. L.; Jiang, J. H.; Li, Y. y Yu, R. Q. (2000). *A novel Trilinear Decomposition Algorithm for Second-Order Linear Calibration*. Chemom. Intell. Lab. Syst. 52: 75–86.

Chin, Y-P., Alken, G. y O'Loughlin, E. (1994). *Molecular Weight, Polydispersity, and Spectroscopic Properties of Aquatic Humic Substances*. Environ. Sci. Technol. 28: 1853–1858.

Clara, M.; Windhofer, G.; Hartl, W.; Braun, K.; Simon, M.; Gans, O.; Scheffknecht, C. Y Chovanec, A. (2010). *Occurrence of phthalates in surface runoff, untreated and treated wastewater and fate during wastewater treatment*. Chemosphere 78: 1078:1084.

Código Alimentario Argentino, Chapter XII (2010.) http://www.anmat.gov.ar/alimentos/codigoa/CAPITULO_II.pdf

Collado, M. S.; Mantovani, V. E.; Goicoechea, H. C. y Olivieri, A. C. (2000). *Simultaneous spectrophotometric-multivariate calibration determination of several components of ophthalmic solutions: phenylephrine, chloramphenicol, antipyrine, methylparaben and thimerosal*. Talanta 52: 909–920.

Culzoni, M. J. y Goicoechea, H. C. (2007). *Determination of loratadine and pseudoephedrine sulfate in pharmaceuticals based on non-linear second order spectrophotometric data generated by a pH-gradient flow injection technique and artificial neural networks*. Anal. Bioanal. Chem. 389: 2217–2225.

Culzoni, M. J.; Goicoechea, H. C.; Ibañez, G. A.; Lozano, V. A.; Marsili, N. R. y Olivieri, A.C. (2008). *Second-order advantage from kinetic-spectroscopic data matrices in the presence of extreme spectral overlapping. A multivariate curve resolution-alternating least-squares approach*. Anal. Chim. Acta 614: 46–57.

Culzoni, M. J.; Goicoechea, H. C.; Pagani, A. P.; Cabezón, M. A. y Olivieri, A. C. (2006). *Evaluation of partial least-squares with second-order advantage for the multiway spectroscopic analysis of complex biological samples in the presence of analyte-background interactions*. Analyst 131: 718–732.

Culzoni, M. J.; Mancha de Llanos, A.; De Zan, M. M.; Espinosa-Mansilla, A.; Cañada-Cañada, F.; Muñoz de la Peña, A. y Goicoechea, H. C. (2011). *Enhanced MCR-ALS modeling of HPLC with fast scan fluorimetric detection second-order data for quantitation of metabolic disorder marker pteridines in urine*. Talanta 85: 2368–2374.

Danzer, K. y Currie, L.A. (1998). *Guidelines for calibration in analytical chemistry. Part 1. Fundamentals and single component calibration*. Pure Appl. Chem. 70: 993–1014.

Deblonde, T.; Cossu-Leguille, C. Y Hartemann, P. (2011). *Emerging pollutants in wastewater: A review of the literature*. Int. J. Hyg. Envir. Heal. 214: 442-448.

de Juan, A.; Casassas, E. y Tauler, R. (2000). *Soft modeling of analytical data*. Meyers, R.A. Encyclopedia of Analytical Chemistry. Wiley, Chichester, UK, p. 9800–9837.

de Juan, A. y Tauler, R. (2001). *Comparison of three-way resolution methods for non-trilinear data sets*. J. Chemom. 15: 749–772.

de Juan, A. y Tauler, R. (2003). *Chemometrics applied to unravel multicomponent processes and mixtures: Revisiting latest trends in multivariate resolution*. Anal. Chim. Acta 500: 195–210.

de Juan, A.; Rutan, S. C. y Taule, R. (2009). *Two-Way Data Analysis: Multivariate Curve Resolution –Iterative Resolution Methods*. En: *Comprehensive Chemometrics Chemical and Biochemical Data Analysis*. Elsevier B.V., Amsterdam, p. 325–344.

de la Guardia, M. (2010). *Green analytical chemistry*. TrAC Trends Anal. Chem. 29: 577.

Derks, E. P. P. A.; Sanchez Pastor, M. S. y Buydens, L. M. C. (1995). *Robustness analysis of radial base function and multi-layered feed-forward neural network models*. Chemom. Intell. Lab. Syst. 28: 49 – 60.

De Zan, M. M.; Gil García, M. D.; Culzoni, M. J.; Siano, R. G.; Goicoechea, H. C. y Martínez Galera, M. (2008). *Solving matrix-effects exploiting the second order advantage in the resolution and determination of eight tetracycline antibiotics in effluent wastewater by modelling liquid chromatography data with multivariate curve resolution-alternating least squares and unfolded-partial least squares followed by residual bilinearization algorithms I. Effect of signal pre-treatment*. J. Chromatogr. A 1179: 106–114.

Di Anibal, C. V.; Ruisánchez, I.; Fernández, M.; Forteza, R.; Cerdà, V. y Callao, M. P. (2012). *Standardization of UV-visible data in a food adulteration classification problem*. Food Chem. 134: 2326–2331.

Di Francesco, R.; Griggs, J. J.; Donnelly, J. y Di Cenzo, R. (2007). *Simultaneous analysis of cyclophosphamide, doxorubicin and doxorubicinol by liquid chromatography coupled to tandem mass spectrometry*. J. Chromatogr. B Analyt Technol Biomed Life Sci. 852: 545–553.

Dinc, E.; Baydan, E.; Kanbur, M. y Onur, F. (2002). *Spectrophotometric multicomponent determination of sunset yellow, tartrazine and allura red in soft drink powder by double divisor-ratio spectra derivative, inverse leastsquares and principal component regression methods*. Talanta 58: 579–594.

Dixit, S.; Khanna, S. K. y Das, M. (2008). *A simple 2-directional high-performance thin-layer chromatographic method for the simultaneous determination of curcumin, metanil yellow, and sudan dyes in turmeric, chili, and curry powders*. J. AOAC Int. 91: 1387–1396.

Dong, M. W. (2006). *Modern HPLC for Practicing Scientists*, Wiley Synomics Pharmaceutical Services, LLC, Wareham, Massachusetts.

Dossi, N.; Toniolo, R.; Susmel, S.; Pizzariello, A. y Bontempelli, G. (2006). *Simultaneous RP-LC determination of additives in soft drinks*. Chromatographia 63: 557–562.

Dossi, N.; Toniolo, R.; Pizzariello, A.; Susmel, S.; Perennes, F. y Bontempelli, G. (2007a). *A capillary electrophoresis microsystem for the rapid in-channel amperometric detection of synthetic dyes in food*. J. Electroanal. Chem. 601: 1–7.

Dossi, N.; Piccin, E.; Bontempelli, G.; Carrilho, E. y Wang, J. (2007b). *Rapid analysis of azo-dyes in food by microchip electrophoresis with electrochemical detection*. *Electrophoresis*, 28: 4240–4246.

Escandar, G. M.; Faber, N. K. M.; Goicoechea, H. C.; Muñoz de la Peña, A.; Olivieri, A. C. y Poppi, R. J. (2007). *Second and third-order multivariate calibration: data, algorithms and applications*. *TrAC, Trends Anal. Chem.* 26: 752–765.

Escandar, G. M.; Goicoechea, H. C.; Muñoz de la Peña, A.; Olivieri, A. C. *Second- and higher-order data generation and calibration: A tutorial*. *Anal. Chim. Acta* In Press, Accepted Manuscript, Available online 11 November 2013.

Espinosa-Mansilla, A.; Muñoz de la Peña, A.; Goicoechea, H. C.; Olivieri, A. C. (2004). *Two multivariate strategies applied to three-way kinetic spectrophotometric data for the determination of mixtures of the pesticides carbaryl and chlorpyrifos*. *Appl. Spectrosc.* 58: 83–90.

Esteves da Silva, J. C. G. y Oliveira, C. J. S. (1999). *Parafac decomposition of threeway kinetic-spectrophotometric spectral matrices corresponding to mixtures of heavy metal ions*. *Talanta* 49: 889–897.

Erk, N. y Kartal, M. (1998). *Simultaneous high performance liquid chromatographic and derivative ratio spectra spectrophotometry determination of chlorpheniramine maleate and phenylephrine hydrochloride*. *Farmaco* 53: 617–622.

European Medicines Agency (2011). *Guideline on bioanalytical method validation*. http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2011/08/WC500109686.pdf

Faber, N. M. (2000). *Exact presentation of multivariate calibration model as univariate calibration graph*. *Chemom. Intell. Lab. Syst.* 50: 107–114.

Fahmy, O. T.; Korany, M. A. y Maher, H. M. (2004). *High performance liquid chromatographic determination of some co-administered anticancer drugs in pharmaceutical preparations and in spiked human plasma*. *J. Pharm. Biomed. Anal.* 34: 1099–1107.

Fatta, D.; Nikolaou, A.; Achilleos, A. y Meric, S. (2007) *Analytical methods for tracing pharmaceutical residues in water and wastewater*. Trends Anal. Chem. 26: 515–533.

Feudale, R. N.; Woody, N. A.; Tan, H.; Myles, A. J.; Brown, S. D.; Ferré, J. (2002) *Transfer of multivariate calibration models: a review*. Chemom. Intell. Lab. Syst. 64: 181–192.

Galeano Diaz, T.; Guiberteau, A.; Ortiz Burguillos, J. M. y Salinas, F. (1997). *Comparison of chemometric methods: derivative ratio spectra and multivariate methods (CLS, PCR and PLS) for the resolution of ternary mixtures of the pesticides carbofuran carbaryl and fenamiphos after their extraction into chloroform*. Analyst 122: 513–517.

Gałaszka, A.; Migaszewski, Z. y Namieśnik, J. (2013). *The 12 principles of green analytical chemistry and the SIGNIFICANCE mnemonic of green analytical practices*. TrAC Trends Anal. Chem. 50: 78–84.

García-Reiriz, A.; Damiani, P. C. y Olivieri, A. C. (2007). *Analysis of amoxicillin in human urine by photo-activated generation of fluorescence excitation–emission matrices and artificial neural networks combined with residual bilinearization*. Anal. Chim. Acta 88: 192–199.

García-Reiriz, A.; Damiani, P. C.; Culzoni, M. J.; Goicoechea, H. C. y Olivieri, A. C. (2008). *A versatile strategy for achieving the second-order advantage when applying different artificial neural networks to non-linear second-order data: unfolded principal component analysis/residual bilinearization*. Chemom. Intell. Lab. Syst. 92: 61–70.

Gemperline, P. J.; Long, J. R. y Gregoriou, V. G. (1991). *Nonlinear multivariate calibration using principal components regression and artificial neural networks*. Anal. Chem. 63: 2313–2323.

Gil García, M. D.; Cañada-Cañada, F.; Culzoni, M. J.; Vera-Candioti, L.; Siano, G. G.; Goicoechea, H. C. y Martínez Galera, M. (2009). *Chemometric tools improving the determination of anti-inflammatory and antiepileptic drugs in river and wastewater by solid-phase microextraction and liquid chromatography diode array detection*. J. Chromatogr. A 1216: 5489–5496.

Gil García, M. D.; Culzoni, M. J.; De Zan, M. M.; Santiago Valverde, R.; Martínez Galera, M. y Goicoechea, H. C. (2008). *Solving matrix effects exploiting the second-order advantage in the resolution and determination of eight tetracycline antibiotics in effluent wastewater by modelling liquid chromatography data with multivariate curve resolution-alternating least squares and unfolded-partial least squares followed by residual bilinearization algorithms: II. Prediction and figures of merit*. J. Chromatogr. A 1179: 115–124.

Goicoechea, H. C.; Culzoni, M. J.; Gil García, M. D. y Martínez Galera, M. (2011). *Chemometric strategies for enhancing the chromatographic methodologies with second-order data analysis of compounds when peaks are overlapped*. Talanta 83: 1198–1107.

Goicoechea, H. C. y Olivieri, A. C. (1999). *Determination of bromhexine in cough-cold syrups by absorption spectrophotometry and multivariate calibration using partial least-squares and hybrid linear analyses. Application of a novel method of wavelength selection*. Talanta 49:793–80.

Goicoechea, H. C. y Olivieri, A. C. (2005). *New Robust Bilinear Least Squares Method for the Analysis of Spectral-pH Matrix Data*. Appl. Spectrosc. 59: 926–933.

Goicoechea, H. C.; Olivieri, A. C. y Tauler, R. (2010). *Application of the correlation constrained multivariate curve resolution alternating least-squares method for analyte quantitation in the presence of unexpected interferences using first-order instrumental data*. Analyst 135: 636–642.

Goicoechea, H. C.; Yu, S; Olivieri, A. C. y Campiglia, A. D. (2005). *Four-Way Data Coupled to Parallel Factor Model Applied to Environmental Analysis: Determination of 2,3,7,8-Tetrachloro-dibenzo-para-dioxin in Highly Contaminated Waters by Solid-Liquid Extraction Laser-Excited Time-Resolved Shpol'skii Spectroscopy*. Anal. Chem. 77: 2608–2616.

Goicoechea, H. C.; Yu, S; Moore A. F. T. y Campiglia, A. D. (2012). *Four-way modeling of 4.2K time-resolved excitation emission fluorescence data for the quantitation of polycyclic aromatic hydrocarbons in soil samples*. Talanta 101: 330–336.

Gomez, M. R.; Olsina, R. A.; Martínez, L. D. y Silva, M. F. (2002). *Simultaneous determination of dextromethorphan, diphenhydramine and phenylephrine in expectorant and decongestant syrups by capillary electrophoresis*. J. Pharm. Biomed. Anal. 30: 791–799.

Gómez, V. y Callao, M. P. (2008). *Analytical applications of second-order calibration methods*. Anal. Chim. Acta 627: 169–183.

González, A. G.; Herrador, M. A. y Asuero, A. G. (1999). *Intra-laboratory testing of method accuracy from recovery assays*. Talanta 48: 729–736.

Gotlieb, W. H.; Bruchim, I.; Ben-Baruch, G.; Davidson, B.; Zeltser, A.; Andersen, A. y Olsen, H. (2007). *Doxorubicin levels in the serum and ascites of patients with ovarian cancer*. EJSO 33: 213–215.

Grande, B. V.; Manne, R. (2000). *Use of Convexity for Finding Pure Variables in Two-Way Data from Mixtures*. Chemom. Intell. Lab. Syst. 50: 19–33.

Grimmer, G.; Brune, H.; Dettbarn, G.; Heinrich, U.; Jacob, J.; Mohtashamipur, E.; Norpoth, K.; Pott, F. y Wenzel-Hartung, R. (1988). *Urinary and faecal excretion of chrysene and chrysene metabolites by rats after oral, intraperitoneal, intratracheal or intrapulmonary application*. Arch. Toxicol. 62: 401–405.

Gros, M.; Petrovic, M. y Barceló, D. (2006). *Development of a multi-residue analytical methodology based on liquid chromatography–tandem mass spectrometry (LC–MS/MS) for screening and trace level determination of pharmaceuticals in surface and wastewaters*. Talanta 70: 678–690.

Grujic, S.; Vasiljevic, T. y Lausevic, M. (2009). *Determination of multiple pharmaceutical classes in surface and ground waters by liquid chromatography–ion trap–tandem mass spectrometry*. J. Chromatogr. A, 1216: 4989–5000.

Haaland, D. M.; Han, L. y Niemczyk, T. M. (1999). *Use of CLS to understand PLS IR calibration for trace detection of organic molecules in water*. Appl. Spectrosc. 53: 390–395.

Haaland, D. M. y Thomas, E. V. (1988). *Partial Least-Squares Methods for Spectral Analyses. Relation to Other Quantitative Calibration Methods and the Extraction of Qualitative Information*. Anal. Chem. 60: 1193–1202.

Hansen, P-D. (2007). *Risk assessment of emerging contaminants in aquatic systems*. Trends Anal. Chem. 26: 1095–1099.

Harshman, R. A. (1970). *Foundations of the PARAFAC procedure: model and conditions for an 'explanatory' multi-mode factor analysis*. UCLA Work. Pap. Phon. 6: 1–84.

Harshman, R. A. (1972). *Determination and proof of minimum uniqueness conditions for PARAFAC*. UCLA Work. Pap. Phon. 22: 111–117.

Hessen, D. O. y Tranvik, L. J. (Editors) (1998). *Aquatic Humic Substances: Ecology and Biogeochemistry*. Springer, Berlin.

Holler, F. J.; Skoog, D. A. y West, D. M. (1996). *Fundamentals of analytical chemistry*. Philadelphia: Saunders College Pub.

Honorato, F. A.; Honorato, R. S.; Pimentel, M. F. y Araujo, M. C. U. (2002). *Analytical curve or standard addition method: how to elect and design-a strategy applied to copper determination in sugarcane spirits using AAS*. Analyst 127: 1520–1525.

Hu, L. Q, Wu, H. L.; Ding, Y. J.; Fang, D. M.; Xia, A. L. y Yu, R. Q. (2006). *Alternating asymmetric trilinear decomposition for three-way data arrays analysis*. Chemom. Intell. Lab. Syst. 82: 145–153.

Jager, A. V.; Tonin, F. G. y Tavares, M. F. M. (2005). *Optimizing the separation of food dyes by capillary electrophoresis*. J. Sep. Sci. 28: 957–965.

Jaumot, J.; Gargallo, R.; De Juan, A. y Tauler, R. (2005). *A graphical userfriendly interface for MCR-ALS: a new tool for multivariate curve resolution in MATLAB*. Chemom. Intell. Lab. Syst. 76: 101–110.

Joint Expert Committee on Food Additives depending on Food and Agriculture Organization (FAO) y World Health Organization (WHO), <http://www.codexalimentarius.net/gsaonline>, 2013.

Jongeneelen, F. J.; Anzion, R. B. M.; Leijdekkers, Ch. M.; Bos, R. P. y Henderson P. Th. (1985). *1-Hydroxypyrene in human urine after exposure to coal tar and a coal tar derived product*. Int. Arch. Occ. Env. Hea. 57: 47–55.

Kiers, H. A. L.; ten Berge, J. M. F. y Bro, R. (1999). *PARAFAC2 - Part I. A direct fitting algorithm for the PARAFAC2 model*. J. Chemom. 13: 275–294.

Knobel, G.; Calimag-Williams, K. y Campiglia, A. D. (2013). *Analysis of polycyclic aromatic hydrocarbon metabolites in cow's milk by liquid-liquid extraction and synchronous room-temperature fluorescence spectroscopy*. Anal. Meth. 5: 1577-1582.

Knochen, M. y Giglio, J. (2004). *Flow-injection determination of phenylephrine hydrochloride in pharmaceutical dosage forms with on-line solid-phase extraction and spectrophotometric detection*. Talanta 64: 1226-1232.

Kruskal, J. B. (1976). *More factors than subjects, tests and treatments: An indeterminacy theorem for canonical decomposition and individual differences scaling*. Psychometrika 41: 281-293.

Kruskal, J. B. (1977). *Three-way arrays: Rank and uniqueness of trilinear decomposition, with application to arithmetic complexity and statistics*. Linear Algebra Appl. 18: 95-138.

Kumar, K. y Mishra, A. K. (2011). *Simultaneous quantification of dilute aqueous solutions of certain polycyclic aromatic hydrocarbons (PAHs) with significant fluorescent spectral overlap using total synchronous fluorescence spectroscopy (TSFS) and N-PLS, unfolded-PLS and MCR-ALS analysis*. Anal. Meth. 3: 2616-2624.

Kumar, K. y Mishra, A. K. (2012a). *Application of parallel factor analysis to total synchronous fluorescence spectrum of dilute multifluorophoric solutions: addressing the issue of lack of trilinearity in total synchronous fluorescence data set*. Anal. Chim. Acta 755: 37-45.

Kumar, K. y Mishra, A. K. (2012b). *Application of 'multivariate curve resolution alternating least square (MCR-ALS)' analysis to extract pure component synchronous fluorescence spectra at various wavelength offsets from total synchronous fluorescence spectroscopy (TSFS) data set of dilute aqueous solutions of fluorophores*. Chemom. Intell. Lab. Syst. 116: 78-86.

Kucharska, M. y Grabka, J. (2010). *A review of chromatographic methods for determination of synthetic food dyes*. Talanta 80: 1045-1051.

Lachâtre, F.; Marquet, P.; Ragot, S.; Gaulier, J. M.; Cardot, P. y Dupuy, J. L. (2000). *Simultaneous determination of four anthracyclines and three metabolites in human serum by*

liquid chromatography-electrospray mass spectrometry. J. Chromatogr. B Biomed. Sci. Appl. 738: 281–291.

Lachenmeier, D. W. y Kessler, W. J. (2008). *Multivariate curve resolution of spectrophotometric data for the determination of artificial food colors*. J. Agric. Food Chem. 56: 5463–5468.

Lakowicz, J. R. (1983) *Principles of Fluorescence Spectroscopy*. Plenum Press, New York, p. 44.

Leardi, R. (2009). *Experimental design in chemistry: A tutorial*. Anal. Chim. Acta 652:161–172.

Lee, K. S.; Shiddiky, M. J. A.; Park, S. H.; Park, D. S. y Shim, Y. B. (2008). *Electrophoretic analysis of food dyes using a miniaturized microfluidic system*. Electrophoresis 29: 1910–1917.

Linder, M.; Sundberg, R. (1998). *Second-order calibration: bilinear least squares regression and a simple alternative*. Chemom. Intell. Lab. Syst. 42: 159–178.

Linder, M. y Sundberg, R. (2002). *Precision of prediction in second-order calibration, with focus on bilinear regression methods*. J. Chemom. 16: 12–27.

Llamas, N. E.; Garrido, M.; Di Nezio, M. S. y Fernández Band, B. S. (2009). *Second order advantage in the determination of amaranth, sunset yellow FCF and tartrazine by UV-vis and multivariate curve resolution-alternating least squares*. Anal. Chim. Acta 655: 38–42.

Lloyd J. B. F. (1971). *Synchronized excitation of fluorescence emission spectra*. Nature Phys. Sci. 231: 64–65.

Loadman, P. M. y Calabrese, C. R. (2001). *Separation methods for anthraquinone related anti-cancer drugs*. J. Chromatogr. B: Biomed. Appl. 764: 193–206.

Lozano, V. A.; Ibañez, G. A. y Olivieri, A. C. (2008). *Three-way partial least-squares/residual bilinearization study of second-order lanthanide-sensitized luminescence excitation-time decay data. Analysis of benzoic acid in beverage samples*. Anal. Chim. Acta 610: 185–195.

Lozano, V. A.; Ibañez, G. A. y Olivieri, A. C. (2009b). *A novel second-order standard addition analytical method based on data processing with multidimensional partial least-squares and residual bilinearization*. Anal. Chim. Acta 651: 165–172.

Lozano, V. A.; Ibañez, G. A. y Olivieri, A. C. (2010). *Second-Order Analyte Quantitation under Identical Profiles in One Data Dimension. A Dependency- Adapted Partial Least-Squares/Residual Bilinearization Method*. Anal. Chem. 82: 4510–4519.

Lozano, V. A.; Tauler, R.; Ibañez, G. A. y Olivieri, A. C. (2009a). *Standard addition analysis of fluoroquinolones in human serum in the presence of the interferent salicylate using lanthanide-sensitized excitation-time decay luminescence data and multivariate curve resolution*. Talanta 77: 1715–1723.

Lu, H.; Yuan, G.; He, Q. y Chen, H. (2009). *Rapid analysis of anthracycline antibiotics doxorubicin and daunorubicin by microchip capillary electrophoresis*. Microchem. J. 92: 170–173.

Ma, M.; Luo, X.; Chen, B.; Su, S. y Yao, S. (2006). *Simultaneous determination of water-soluble and fat-soluble synthetic colorants in foodstuff by high-performance liquid chromatography–diode array detection–electrospray mass spectrometry*. J. Chromatogr. A 1103: 170–176.

Maeder, M. (1987). *Evolving Factor Analysis for the Resolution of Overlapping Chromatographic Peaks*. Anal. Chem. 59: 527–530.

Maeder, M.; Zilian, A. (1988). *Evolving Factor Analysis, a New Multivariate Technique in Chromatography*. Chemom. Intell. Lab. Syst. 3: 205–213.

Maeder, M.; Zuberbuehler, A. D. (1986). *The Resolution of Overlapping Chromatographic Peaks by Evolving Factor Analysis*. Anal. Chim. Acta 181: 287–291

Mahgoub, H. (1990). *Computerized Derivative Spectrophotometric Assay of two Component Mixture Using Least Squares Method*. Drug Dev. Ind. Pharm. 16: 2135–2144.

Mamián-López, M. B. y Poppi, R. J. (2013). *Standard addition method applied to the urinary quantification of nicotine in the presence of cotinine and anabasine using surface enhanced Raman spectroscopy and multivariate curve resolution*. Anal. Chim. Acta 760: 53–59.

Marín, A.; García, E.; García, A. y Barbas, C. (2002). *Validation of a HPLC quantification of acetaminophen, phenylephrine and chlorpheniramine in pharmaceutical formulations: capsules and sachets*. J. Pharm. Biomed. Anal. 29: 701–714.

Martens, H. y Naes, T. (1989). *Multivariate Calibration*. John Wiley & Sons Ltd., Chichester, England.

Martínez, A.; Riu, J.; Busto, O.; Guasch, J. y Rius, F. X. (2000). *Validation of bias in multianalyte determination methods. Application to RP-HPLC derivatizing methodologies*. Anal. Chim. Acta 406: 257–278.

Martínez Galera, M.; Gilgarcía, M. D. y Goicoechea, H. C. (2007). *The application to wastewaters of chemometric approaches to handling problems linked to highly complex matrices*. TrAC Trends Anal. Chem. 26: 1032–1042.

Massart, D. L.; Vandeginste, B. G. M.; Buydens, L. M. C.; De Jong, S.; Lewi, P. J.; Smeyers-Verbeke, J. (1997). *Handbook of Chemometrics and Qualimetrics*. Elsevier. Amsterdam.

MATLAB 7.1. (2005). The MathWorks Inc., Natick, Massachusetts, USA.

MATLAB 7.10. (2010). The MathWorks Inc., Natick, Massachusetts, USA.

Miller, J. N. y Miller, J. C. (2005). *Statistics and Chemometrics for Analytical Chemistry*. New York: Pearson-Prentice Hall (Capítulo 2).

Milojkovic-Opsenica, D. M.; Lazarevic, K.; Ivackovic, V. y Tesic, Z. L. (2003). *Reversed-phase thin-layer chromatography of some foodstuff dyes*. J. Planar. Chromatogr. 16: 276–279.

Minioti, K. S.; Sakellariou, C. F. y Thomaidis, N. S. (2007). *Determination of 13 synthetic food colorants in water-soluble foods by reversed-phase highperformance liquid chromatography coupled with diode-array detector*. Anal. Chim. Acta 583: 103–110.

Muñoz, G. y De Juan A. (2007). *pH- and time-dependent hemoglobin transitions: A case study for process modelling*. Anal. Chim. Acta 595: 198–208.

Muñoz de la Peña, A.; Espinosa-Mansilla, A.; Acedo Valenzuela, M. I.; Goicoechea, H. C. y Olivieri, A. C. (2002). *Comparative study of net analyte signal-based methods and partial least squares for the simultaneous determination of amoxicillin and clavulanic acid by stopped-flow kinetic analysis*. Anal. Chim. Acta 463: 75–88.

Murillo Pulgarín, J. A. y García Bermejo, L. F. (1996). *Flow-injection stopped-flow spectrofluorimetric kinetic determination of paracetamol based on its oxidation reaction by hexacyanoferrate (III)*. Anal. Chim. Acta 333: 59–69.

Multivariate Curve Resolution Homepage. <http://www.mcrals.info/>

Naes, T.; Martens, H. (1998). *Principal component regression in NIR [near-infra-red] analysis: viewpoints, background details and selection of components*. J. Chemom. 2: 155–167.

Navarro-Villoslada, F.; Pérez-Arribas, L. V.; León-González, M. E. y Polo-Díez, L. M. (1999). *Matrix effect modelling in multivariate determination of priority pollutant chlorophenols in urine samples*. Anal. Chim. Acta 381: 93–102.

Nepote, A. J. y Olivieri, A.C. (2001). *Simultaneous spectrofluorometric determination of oxatomide and phenylephrine in the presence of a large excess of paracetamol*. Anal. Chim. Acta 439: 87–94.

Ni, Y., Wang, Y. y Kokot, S. (2009). *Simultaneous kinetic spectrophotometric analysis of five synthetic food colorants with the aid of chemometrics*. Talanta 78: 432–441.

Nieman, T. A.; Skoog, D. A. y Holler, F. J. (1998). *Principles of instrumental analysis*. Pacific Grove, CA: Brooks/Cole.

Öhman, J.; Geladi, P. y Wold, S. (1990). *Residual bilinearization. Part 1: Theory and algorithms*. J. Chemometrics 4: 79–90.

Oka, H. et al. (1987). *Simple method for the analysis of food dyes on reversed-phase thin-layer plates*. J. Chromatogr. A 411: 437–444.

Olivieri, A.C. (2005a). *On a versatile second-order multivariate calibration method based on partial least-squares and residual bilinearization: Second-order advantage and precision properties*. J. Chemom. 19: 253–265.

Olivieri, A. C. (2005b). *A combined artificial neural network/residual bilinearization approach for obtaining the second-order advantage from three-way non-linear data*. J. Chemom. 19: 615–624.

Olivieri, A. C.; Escandar, G. M. y Muñoz de la Peña, A. (2011). *Second-order and higher-order multivariate calibration methods applied to non-multilinear data using different algorithms*. TrAC, Trends Anal Chem 30: 607–617.

Olivieri, A. C. y Faber, N(K). M. (2005). *A closed-form expression for computing the sensitivity in second-order bilinear calibration*. J. Chemom. 19: 583–592.

Olivieri, A. C. y Faber, N(K). M. (2012). *New developments for the sensitivity estimation in four-way calibration with the quadri-linear parallel factor model*. Anal. Chem. 84: 186–193.

Olivieri, A. C.; Faber, N(K). M.; Ferré, J.; Boqué, R.; Kalivas, J. H. y Mark, H. (2006). *Uncertainty estimation and figures of merit for multivariate calibration*. Pure Appl. Chem. 78: 633–661.

Olivieri, A. C.; Wu, H. L. y Yu, R. Q. (2009). *MVC2: a MATLAB graphical interface toolbox for second-order multivariate calibration*. Chemom. Intell. Lab. Syst. 96: 246–251.

Organización Mundial de la Salud (2007). Departamento de Inocuidad de los Alimentos, Zoonosis y Enfermedades de Transmisión Alimentaria. http://www.who.int/topics/food_safety/es/

Patra, D.; Mishra, A. K. (2002a). *Recent developments in multi-component synchronous fluorescence scan analysis*. TrAC, Trends Anal. Chem. 21: 787–798.

-
- Patra, D. y Mishra, A. K. (2002b). *Total synchronous fluorescence scan spectra of petroleum products*. Anal. Bioanal. Chem. 373: 304–309.
- Peré-Trepat, E. ; Lacorte, S. y Tauler, R. (2007). *Alternative calibration approaches for LC-MS quantitative determination of coeluted compounds in complex environmental mixtures using multivariate curve resolution*. Anal. Chim. Acta 595: 228–237.
- Peré-Trepat, E.; Lacorte, S.; Tauler, R. (2005). *Solving liquid chromatography mass spectrometry coelution problems in the analysis of environmental samples by multivariate curve resolution*. J. Chromatogr. A 1096: 111 – 122.
- Pereira Alves, S.; Brum, D. M.; de Andrade, É. C. B. y Netto, A. D. P. (2008). *Determination of synthetic dyes in selected foodstuffs by high performance liquid chromatography with UV-DAD detection*. Food Chem. 107: 489–496.
- Petrovic, M.; Hernando, M. D.; Diaz-Cruz, M. S. y Barceló, D. (2005) *Liquid chromatography-tandem mass spectrometry for the analysis of pharmaceutical residues in environmental samples: a review*. J. Chromatogr. A, 1067: 1–14.
- Piccirilli, G.N. y Escandar, G.M. (2006). *Partial least-squares with residual bilinearization for the spectrofluorimetric determination of pesticides. A solution of the problems of inner-filter effects and matrix interferences*. Analyst 131: 1012–1020.
- Pitonesi, M.; Centurión, M. E.; Ferenández Band, B. S.; Damiani, P. C. y Olivieri, A. C. (2004). *Simultaneous determination of levodopa and benserazide by stoppedflow injection analysis and three-way multivariate calibration of kineticspectrophotometric data*. J. Pharmaceut. Biomed. 36: 541–547.
- Raich-Montiu, J.; Folch, J.; Compañó, R.; Granados, M. y Prat, M. D. (2007). *Analysis of trace levels of sulfonamides in surface water and soil samples by liquid chromatography-fluorescence*. Journal of Chromatogr. A, 1172: 186–193.
- Riu, J. y Rius, F. X. (1996). *Assessing the accuracy of analytical methods using linear regresión with errors in both axes*. Anal. Chem. 68: 1851–1857.
-

Rodil, R.; Quintana, J. B.; López-Mahía, P.; Muniategui-Lorenzo, S. y Prada-Rodríguez, D. (2009). *Multi-residue analytical method for the determination of emerging pollutants in water by solid-phase extraction and liquid chromatography–tandem mass spectrometry*. J. Chromatogr. A 1216: 2958–2969.

Rodríguez, N.; Ortiz, M. C. y Sarabia, L. A. (2009a). *Fluorescence quantification of tetracycline in the presence of quenching matrix effect by means of a four-way model*. Talanta. 77: 1129–1136.

Rodríguez, N.; Real, B. D.; Ortiz, M. C.; Sarabia, L. A. y Herrero, A. (2009b). *Usefulness of parallel factor analysis to handle the matrix effect in the fluorescence determination of tetracycline in whey milk*. Anal. Chim. Acta 632: 42–51.

Ruckebusch, C.; De Juan, A.; Duponchel, L. y Huvenne, J. P. (2006). *Matrix augmentation for breaking rank-deficiency: A case study*. Chemom. Intell. Lab. Syst. 80: 209–214.

Rutan, S. C.; de Juan, A. y Tauler, R. (2009). *Introduction to Multivariate Curve Resolution*. En: *Comprehensive Chemometrics Chemical and Biochemical Data Analysis*. Elsevier B.V., Amsterdam, p. 249–259.

Sadygov, R. G.; Maroto, F. M. y Hümer, A. F. R. (2006). *ChromAlign: A two-step algorithmic procedure for time alignment of three-dimensional LC-MS chromatographic surfaces*. Anal. Chem. 78: 8207–8217.

Sanchez, E. y Kowalski, B. R. (1986). *Generalized rank annihilation factor analysis*. Anal. Chem. 58: 496–499.

Sanchez, E. y Kowalski, B. R. (1990). *Tensorial resolution: A direct trilinear decomposition*. J. Chemom. 4: 29–45.

Santos, J. L.; Aparicio, I.; Alonso, E. y Callejón, M. (2005). *Simultaneous determination of pharmaceutically active compounds in wastewater samples by solid phase extraction and high-performance liquid chromatography with diode array and fluorescence detectors*. Anal. Chim. Acta 550: 116–122.

Saurina, J.; Leal, C.; Compañó, R.; Granados, M.; Prat, M. D. y Tauler, R. (2001). *Estimation of figures of merit using univariate statistics for quantitative second-order multivariate curve resolution*. Anal. Chim. Acta 432: 241–251.

Saxberg, B. E. y Kowalski, B. R. (1979). *Generalized standard addition method*. Anal. Chem. 51: 1031–1038.

Schenone, A. V.; Culzoni, M. J.; Marsili, N. R. y Goicoechea, H. G. (2013a). *Determination of tartrazine in beverage samples by stopped-flow analysis and three-way multivariate calibration of non-linear kinetic-spectrophotometric data*. Food Chem. 138: 1928–1935.

Schenone, A. V.; Culzoni, J. M.; Campiglia, A. D. y Goicoechea, H. C. (2013b). *Total synchronous fluorescence spectroscopic data modeled with first- and second-order algorithms for the determination of doxorubicin in human plasma*. Anal. Bional. Chem. 405: 8515–8523.

Senesi, N.; Miano, T. M.; Provenzano, M. R. y Brunetti, G. (1991). *Characterization, Differentiation and Classification of Humic Substances by Fluorescence Spectroscopy*. Soil Sci. 152: 259–271.

Shabania, A. M. H.; Ellis, P. S. y McKelvie I. D. (2011). *Spectrophotometric determination of iodate in iodised salt by flow injection análisis*. Food Chem. 129: 704–707.

Shariati-Rad, M.; Irandoust, M.; Amin, N. y Shamsipur, M. (2013a). *Solving matrix effect, spectral interferences and nonlinearity by generalized standard addition method coupled with radial basis functions-partial least squares: Application to simultaneous determination of drugs in urine*. Chemom. Intell. Lab. Syst. 120: 77–83.

Shariati-Rad, M.; Irandoust, M.; Amini, T. y Shamsipur, M. (2013b). *Solving matrix effect, spectral overlapping and nonlinearity by generalized standard addition method coupled with radial basis functions-partial least squares: simultaneous determination of atorvastatin and amlodipine in urine*. J. Chemom. 27: 63–69.

Shaw, P. J. A. (2003) *Multivariate Statistics for the Environmental Sciences*. Hodder-Arnold, London.

Siano, G. G. y Goicoechea, H. C. (2007). *Representative subset selection and standardization techniques. A comparative study using NIR and a simulated fermentative process UV data*. Chemom. Intell. Lab. Syst. 88: 204–212.

Silva, E. C.; Araújo, M. C. U.; Honorato, R. S.; Costa Lima, J. L. F.; Guidetti Zagatto, E. A. y Boscolo Brienza, S. M. (1996). *Standard additions in flow injection analysis based on merging zones and gradient exploitation: application to copper determination in spirits*. Anal. Chim. Acta 319: 153–158.

Silva, M. L. S.; García, B. Q.; Lima, J. L. F. C. y Barrado, E. (2007). *Voltammetric determination of food colorants using a polyallylamine modified tubular electrode in a multicommutated flow system*. Talanta 72: 282–288.

Silva, E. C.; Martins, V. L.; Araújo, A.F. y Araújo, M.C.U. (1999). *Implementation of a generalized standard addition method in a flow injection system using merging-zones and gradient exploitation*. Anal. Sci. 15: 1235–1240.

Smilde, A. K.; Wang, Y. y Kowalsky, B. R. (1994). *Theory of medium-rank second-order calibration with restricted-Tucker models*. J. Chemom. 8: 21–36.

Soponar, F.; Mot, A. C. y Sârbu, C. (2008). *Quantitative determination of some food dyes using digital processing of images obtained by thin-layer chromatography*. J. Chromatogr. A 1188: 295–300.

Speth, P. A.; van Hoesel, Q. G. y Haanen, C. (1988). *Clinical pharmacokinetics of doxorubicin*. Clin. Pharmacokinet. 15: 15–31.

Suykens, J. A. K.; van Gestel, T.; de Brabanter, J.; de Moor, B. y Vandewalle, J. (2002). *Least-Squares Support Vector Machines*. World Scientific, Singapore.

Tauler, R. (1995). *Multivariate curve resolution applied to second order data*. Chemom. Intell. Lab. Syst. 30: 133–146.

Tauler, R. ; Smilde, A.K. y Kowalski, B. R. (1995). *Selectivity, local rank, three-way data analysis and ambiguity in multivariate curve resolution*. J. Chemom. 9: 31-58.

Thygesen, J. y van den Berg, F. (2011). *Calibration transfer for excitation-emission fluorescence measurements*. Anal. Chim. Acta 705: 81-87.

Trevisan, M. G. y Poppi, R. J. (2003). *Determination of doxorubicin in human plasma by excitation-emission matrix fluorescent and multi-way analysis*. Anal. Chim. Acta 493: 69-81.

Tobiszewski, M. y Namieśnik, J. (2012). *Direct chromatographic methods in the context of green analytical chemistry*. TrAC Trends Anal. Chem. 35: 67-73.

Thomas, E. V. (1994). *A primer on multivariate calibration*. Anal. Chem. 66: 795A-804 A.

Vachirapatama, N.; Mahajaroensiri, J. y Visessanguan, W. (2008). *Identification and determination of seven synthetic dyes in foodstuffs and soft drinks on monolithic C18 column by high performance liquid chromatography*. J. Food Drug Anal. 16: 77-82.

Valcarcel, M. (1999). *Principios de Química Analítica*. Editorial Springer-Verlag Ibérica, S.A., Barcelona.

Valderrama, P. y Poppi, R. J. (2011). *Second order standard addition method and fluorescence spectroscopy in the quantification of ibuprofen enantiomers in biological fluids*. Chemom. Intell. Lab. Syst. 106: 160-165.

van der Voet, H. (1994). *Comparing the predictive accuracy of models using a simple randomization test*. Chemom. Intell. Lab. Syst. 25: 313-323.

Vásquez, V.; Báez, M. E.; Bravo, M. y Fuentes, E. (2013). *Determination of heavy polycyclic aromatic hydrocarbons of concern in edible oils via excitation-emission fluorescence spectroscopy on nylon membranes coupled to unfolded partial least-squares/residual bilinearization*. Anal. Bioanal. Chem. 405: 7497-7507.

Vosough, M.; Bayat, M. y Salemi, A. (2010). *Matrix-free analysis of aflatoxins in pistachio nuts using parallel factor modeling of liquid chromatography diode-array detection data*. Anal. Chim. Acta 663: 11–18.

Walczak, B. ; Bouveresse, E. y Massart, D. L. (1997). *Standardization of near-infrared spectra in the wavelet domain*. Chemom. Intell. Lab. Syst. 36: 41–51.

Wang, R. Y. y Lu, Y. T. (2005). *Kinetic spectrophotometric method for determination of perphenazine based on monitoring the oxidation intermediate by applying a stopped-flow technique*. Spectrochim. Acta A 61: 791–797.

Wang, Y. D.; Veltkamp, D. J. y Kowalski, B. R. (1991). *Multivariate instrument standardization*. Anal. Chem. 63: 2750–2756.

Wehry E. L. (1990). *Practical Fluorescence*. (Ed.: G.G. Guilbault). Marcel Dekker, New York.

Wiklund, S.; Nilsson, D.; Eriksson, L.; Sjöström, M.; Wold, S. y Faber, K. (2007). *A randomisation test for PLS component selection*. J. Chemom. 21: 427–439.

Windig, W. y Guilment, J. (1991). *Interactive self-modeling mixture analysis*. Anal. Chem. 63: 1425–1432.

Wise, B. M.; Gallagher, N. B.; Bro, R.; Shaver, J. M.; Windig, W. y Koch, R. S. (2005). *PLS Toolbox Version 3.5 for use with MATLAB*, Eigenvector Research Inc., Manson, USA.

Wold, S.; Geladi, P.; Esbensen, K. y Öhman, J. (1987). *Multiway principal components and PLS analysis*. J. Chemom. 1: 41–56.

Wu, H. L.; Shibukawa, M. y Oguma, K. (1998). *An alternating trilinear decomposition algorithm with application to calibration of HPLC-DAD for simultaneous determination of overlapped chlorinated aromatic hydrocarbons*. J. Chemom. 12: 1–26.

Wülfert, F.; Kok, W. Th.; de Noord, O. E. y Smilde, A. K. (2000). *Correction of Temperature-Induced Spectral Variation by Continuous Piecewise Direct Standardization*. Anal. Chem. 72: 1639–1644.

Xia, A. L.; Wu, H. L.; Fang, D. M.; Ding, Y. J.; Hu, L. Q. y Yu, R. Q. (2005). *Alternating penalty trilinear decomposition algorithm for second-order calibration with application to interference-free analysis of excitation-emission matrix fluorescence data*. J. Chemom. 19: 65-76.

Yang, Y. D.; Veltkamp, D. J. y Kowalski, B. R. (1991). *Multivariate Instrument Standardization*. Anal. Chem. 63: 2750-2756.

Yang, Y.; Yin, J. y Shao, B. (2011). *Simultaneous determination of five aluminum lake dyes in chewing gum by HPLC with photodiode array detection*. Food Addit. Contam. A 28: 1159-1167.

Yoshioka, N. y Ichihashi, K. (2008). *Determination of 40 synthetic food colors in drinks and candies by high-performance liquid chromatography using a short column with photodiode array detection*. Talanta 74: 1408-1413.

Zagotto, G.; Gatto, B.; Moro, S.; Sissi, C. y Palumbo, M. (2001). *Anthracyclines: recent developments in their separation and quantitation*. J. Chromatogr. B. 764: 161-171.

Zampronio, C. G.; Gurden, S.P.; Moraes, L. A. B.; Eberlin, M. N.; Smilde, A. K. y Poppi, R. J. (2002). *Direct sampling tandem mass spectrometry (MS/MS) and multiway calibration for isomer quantitation*. Analyst 127: 1054-1060.

Zhu, S. H.; Wu, H. L.; Li, B. R.; Xia, A. L.; Han, Q. J.; Zhang, Y.; Bian, Y. C. y Yu, R. Q. (2008). *Determination of pesticides in honey using excitation-emission matrix fluorescence coupled with second-order calibration and second-order standard addition methods*. Anal. Chim. Acta 619: 165-172.

Zupan, J. y Gasteiger, J. (1999). *Neural Networks in Chemistry and Drug Design*. Wiley, New York, USA.