



**UNIVERSIDAD NACIONAL DEL LITORAL**

**FACULTAD DE BIOQUÍMICA Y CIENCIAS BIOLÓGICAS**

**CÁTEDRAS DE QUÍMICA ANALÍTICA I Y II**

*DESARROLLO DE NUEVOS MÉTODOS ANALÍTICOS  
BASADOS EN TÉCNICAS ESPECTROSCÓPICAS.  
RESOLUCIÓN DE PROBLEMAS RELACIONADOS CON ADITIVOS  
ALIMENTARIOS Y MUESTRAS AMBIENTALES*

Nilda Raquel Marsili

Director de Tesis

Dr. Héctor C. Goicoechea

2005

## AGRADECIMIENTOS

A mi Director Héctor C. Goicoechea por todo lo que aprendí trabajando a su lado, porque encontré en él una persona preocupada no solo por la ciencia, sino por la problemática de la humanidad, y sobre todo por su amistad y comprensión.

Al Dr. Alejandro Cesar Olivieri y su querida esposa Graciela Escandar por su valiosa y sabia contribución, en la presente tesis.

A todas las personas que forman la Cátedra de Química Analítica I, muy especialmente al Profesor Víctor Mantovani por su calidez y buena predisposición, a Juan Carlos Robles, Mercedes De Zan, Marisi Cámara, Gonzalo, Gabriel, Silvia Fernández, Silvia Fabiano, que me brindaron amistad y compañerismo gracias a lo cual pude llevar adelante el trabajo experimental, todo mi agradecimiento y cariño.

Al Departamento de Química de la Universidad Nacional del Sur (Bahía Blanca) por abrirme sus puertas con tanta generosidad, a la Dra. Beatriz Fernández Band, a la Dra. Adriana Lista, María, Miriam y Marcelo.

A mis viejos compañeros Claudio Núñez, José Coronel, Maria Rosa Williner y Juan Alberto Nepote por veinte años de paciencia.

Al Sr. Decano Ángel Amadeo Cellino.

A mi amiga Silvina Sobrero, a mi Doctora Anita, a la escultora Raquel Minetti, y a la pintora Paula Sobrero, cada una sabia en lo suyo, por que siempre están presentes en las buenas y en las malas.

*A mis dos esperanzas, mi sol y mi luna, Antonio y Emilia*

*A mi queridísima hermana Silvia Ester Marsili, y a todos los chicos: Antonio Julián y María Emilia Vergara, Laura y Federico Marsili, Catalina Ana y Clara Lucia Colombo. A la memoria de mi padre Hermes Antonio Marsili, mi querida tía Dainé Ceferina Adria Marsili y Emilia Antonia Villani. A mi madre Catalina Pons. A mi familia Javier Colombo, Víctor Eliseo Marsili e Ida Lilian (Didi) Aliendro. A un viejo amigo Juan Ignacio Vergara, y sus hermanos Araceli, Tomas, Checho y Pablo Vergara. A Martha Beatriz y Cristina Giavi, Raúl y Laurita Núñez.. A Fabio Fernando y Nelda González. A Eugenio y Olga González. A mi hermana Adriana María Maronetti, su esposo Fernando Marcolín y sus hijas María Eugenia, María Fernanda y María Florencia Marcolín. A mi primo Daniel Belofí y su esposa Susana. A Luci Belofí, y las chicas Marsili: Susi, Viví y Chichí. A los Rodríguez. A Paulita Sobrero, Jorge Ulises Mosset Iturraspe, Paula Manuela, Ulises, Camila y Sofía. A Silvia Susana Baizre y su bella Leila. A Maricel Cherry, la princesa Manón, Néstor Calvo. A María Rosa Brollo, Guillermo Fernández, Ignacio y Pablo. María Rosa Pfeiffer, Bárbara, Luisina y Walter. A mi amiga Raquel Minetti, mi ahijado Pedro Sosa, Joaquín Sosa y Emiliano Nahuel Quintana. A los Todopantalla: Lisandro Francucci, Andrés Olivo, Esteban y Emiliano. A las Fans de los Todopantalla: Malena Acosta, Florencia Osuna, Natacha Gomes Rodríguez, Nadia Pérez y Tamara. A la bella Nazaret Giraldi. A Mariel Leiva. A Perla Gomes y sus hijos Nicolás y Lucas. A Ester Cardellino y la memoria del Padre Justo. A mi sobrino Augusto y su madre Isabel Lilian Fassi. A mi querida Doctora Ana Gabbi y su secretario Javier. A María Silvina Sobrero, Jorge Benet, Julia y Genaro Benet. A Juanita Pedro. A las chicas de la escuela Jorgelina, Mirta, Mari, Alicia. A Gustavo Ricardo Galleguillo. A los hermanos Ana y Berardino José Santirocco. A la memoria de Alfredo Jack, Beatriz y Carlos Cerda. A Néstor Valentín, Daniela y Eugenia Vergara. A Fernando Birri, Mirca y Carmen Birri. A Yeye Andino, Susana Ocampo, Julio Cesar Botta, Roberto Favareto Fornner. A Cesar López Claro y su amada María. A la memoria de Miguel Orellano. A Giotto, Cimabue, Albeto Durero, Xul Solar, Fernando Espino, Gomes Cornet, Frida Kfialo por haber existido. A mis vecinos de los barrios: Candiotti, Guadalupe y Sargento Cabral. A mis vecinos los pescadores: al Gringo. A Doña Eufemia, José, Pancho, Antonio. A Raúl Prieto y Carlos Gaggiomo. A los amigos de otros tiempos: Walter, Marcelo, mi negro, Willi, Federico, Mariela, Carina, Paulo, Sebastián, Martín, al Negro Froia, Birdi, Adriana, Cacho. A Nilda Freitas, Pablo y Carlota Steinberg. A María Antonia Gómez, Martín y Pancho Buschiazzo. A mis compañeros de trabajo de QAI: Claudio Omar Núñez, María Rosa Williner, José Coronel, Juan Alberto Nepote y de QAI: Marisi Cámara, Caco, Víctor, Silvia, Mercedes, Gonzalo, Gabriel, Luciana, Julia. A Héctor Casimiro, Diana, Julia y Paula Goicoechea. A Ángel Amadeo Cellino. A Macedonio Fernández, Juana Azurduy, la Gorda, Tritri, Goliat, la Sra. Pearl, Telma, Saylor y la Tegrta. A la memoria de Jack, "el destripador". Al Loco Gatti y su abuelo el Dr. Niel. A Ana Niel, Ruy. Al Kiwi, Richard. A Querelle. A María, Ramona, Catalina y Mari. A mis alumnos de la Mantovani por la alegría de vivir. A Josefina. A Ernesto, Julio, Jorge, Carlitos, Rosita, José, Lola, Alfonsina, por argentinos, Al Guille, Mónica, y toda la banda. A Marito, la Muñe y la tía Betty. A Andrés, Elvira, Peti, Artemio y Coco. A todos los que nombré, a los que olvidé y a los que no puedo nombrar, a todos los que me ayudan a construir mi acotada y desdibujada realidad.*

*"Será que no soy tan estricta con el universo,  
que las cosas se presentan desordenadas frente a mi,  
siempre de la manera menos pensada,  
nunca dejan de asombrarme."  
"Estaría perdida si las cosas fuesen  
solo lo que parecen ser."*

*Kryptonita*

Parte del material de la presente tesis dio lugar a las siguientes publicaciones y presentaciones a reuniones científicas:

Publicaciones:

1. "Spectrophotometric determination of sorbic and benzoic acids in fruit juices by a net analyte signal based method with selection of the wavelength range to avoid non modelled interferences". Nilda R. Marsili, María S. Sobrero y Héctor C. Goicoechea. *Analytical and Bioanalytical Chemistry*, 2003, 376: 126-133.
2. "New method for the determination of benzoic and sorbic acids in commercial orange juice based on second-order spectrophotometric data generated by a pH gradient flow injection technique". Nilda R. Marsili, Adriana Lista, Beatriz S. Fernandez Band, Héctor C. Goicoechea y Alejandro C. Olivieri. *Journal of Agricultural and Food Chemistry*, 2004, 52: 2479-2484.
3. "Evaluation of complex spectral-pH three-way arrays by modified bilinear least-squares: determination of four different dyes in interfering systems". Nilda R. Marsili, Adriana Lista, Beatriz S. Fernandez Band, Héctor C. Goicoechea y Alejandro C. Olivieri. *Analyst*, 2005, 130: 1291-1298.

Presentaciones a reuniones científicas:

1. "Estrategia para el análisis de vitaminas en una preparación compleja para uso veterinario basada en espectroscopía UV y calibración uni y multivariada". Nilda R. Marsili y Héctor C. Goicoechea. II Congreso de Química Analítica. Huerta Grande, Córdoba, octubre de 2003.
2. "Determinación simultánea de tres principios activos en gel óptico de uso veterinario usando espectroscopía ultravioleta y calibración multivariada mejorada con detergentes". Nilda R. Marsili, Juan C. Robles y Héctor C. Goicoechea. XXV Congreso Argentino de Química, "Dr E. Bottani". Olavarría, Bs. As. septiembre de 2004.

3. “Dos estrategias multivariadas para la determinación de sorbato y benzoato en jugos de frutas basadas en respuestas de primer y segundo orden”. Nilda Marsili, Adriana Lista, Beatriz Fernández Band, Silvina Sobrero, Héctor Goicoechea y Alejandro Olivieri. XXVII Reunión anual de CASLAN. Jornadas de nutrición y salud FBCB-UNL, Santa Fe 11-13 de noviembre de 2004.

4. “Determinación de colorantes en jugos de fruta en polvo utilizando datos espectrales obtenidos por gradiente de pH en flujo y un nuevo método de modelado para datos de segundo orden con dependencia lineal”. Nilda R. Marsili, Adriana Lista, Beatriz Fernández Band, Alejandro C. Olivieri y Héctor C. Goicoechea. Presentación oral (30 minutos). 1º Congreso Iberoamericano de Química Analítica, Niteroi, Brasil, 12 al 16 de setiembre de 2005.

# ÍNDICE DE CONTENIDOS

	Pág.
<b>CAPÍTULO 1: INTRODUCCIÓN</b>	1
1.1 CUANTIFICACIÓN DE ADITIVOS QUÍMICOS EN ALIMENTOS	1
1.1.1 Aditivos químicos, su legislación	1
1.1.2 Clasificación de los aditivos químicos	3
1.1.3 Preservativos químicos	5
1.1.4 Colorantes	7
1.2 CALIBRACIÓN MULTIVARIADA	9
1.2.1 Introducción	9
1.2.1.1 Instrumentos de orden cero	9
1.2.1.2 Instrumentos de orden uno	10
1.2.1.3 Instrumentos de orden dos	11
1.2.1.4 Instrumentos de orden tres y superiores	13
1.2.2 Métodos de calibración multivariada	14
1.3 MÉTODOS DE CALIBRACIÓN DE PRIMER ORDEN	17
1.3.1 Regresión por cuadrados mínimos parciales	17
1.3.1.1 Introducción	17
1.3.1.2 Desarrollo teórico matemático	29
1.3.2 Señal neta del analito	29



1.3.2.1	Introducción	29
1.3.2.2	Desarrollo teórico matemático	29
1.3.3	Métodos basados en señal neta	35
1.3.3.1	Introducción	35
1.3.3.2	Desarrollo teórico matemático	36
1.4	ANÁLISIS MULTIVARIADO DE ORDEN SUPERIOR	43
1.5.1	Introducción	43
1.5.2	PARAFAC	44
1.5.2.1	Perfiles proporcionales paralelos	44
1.5.2.2	Producto Khatri-Rao	44
1.5.2.3	Algoritmos	46
1.5.2.4	Fundamentos matemáticos	46
1.5.2	MCR-ALS	52
1.5.2.1	Introducción	52
1.5.2.2	Desarrollo teórico matemático	52
1.5.3	Cuadrados mínimos bilineales (BLLS)	58
1.5.3.1	Introducción	58
1.5.3.2	Desarrollo teórico matemático	58
1.5	OBJETIVOS	63
	<b>CAPÍTULO 2: MATERIALES Y MÉTODOS</b>	64
2.1	MATERIALES	64
2.1.1	Soluciones patrón de conservantes	64
2.1.2	Soluciones patrón de colorantes	64
2.1.3	Muestras comerciales	65
2.1.3.1	Bebidas preparadas a partir de jugos de fruta	65

2.1.3.2	Polvos para preparar bebidas artificiales	65
2.1.4	Jugos para modelar las interferencias de la matriz	66
2.1.5	Muestras artificiales	66
2.1.6	Reactivos reguladores del pH	66
2.1.7	Reactivos método oficial para conservantes	67
2.1.8	Reactivos para optimizar el sistema FIA	67
2.1.9	Equipos	67
2.2	<b>MÉTODOS</b>	68
2.2.1	Método oficial de la AOAC para la cuantificación de SOR y BEN en bebidas analcohólicas	68
2.2.2	Métodos computacionales	69
2.2.2.1	Programa MULTIVAR	69
2.2.2.2	Programas MVC1 Y MVC2	74
2.2.3	Metodología del análisis del en flujo. Sistema FIA	78
2.2.3.1	Sistema FIA para la cuantificación de conservantes	78
2.2.3.2	Sistema FIA para la cuantificación de colorantes	80
	<b>CAPÍTULO 3: RESULTADOS Y DISCUSIÓN</b>	83
3.1	<b>MÉTODO DESARROLLADO PARA LA CUANTIFICACIÓN DE CONSERVANTES (SORBATO DE POTASIO Y BENZOATO DE SODIO) EN MUESTRAS COMERCIALES Y ARTIFICIALES, APLICANDO ESPECTROSCOPIA UV Y CALIBRACIÓN MULTIVARIADA DE PRIMER ORDEN (PLS y HLA/GO).</b>	83
3.1.1	Introducción	83
3.1.2	Calibración	85
3.1.3	Validación	88
3.1.3.1	Juego de validación	89
3.1.3.2	Precisión: Análisis de una muestra comercial con la adición de conservantes	90

3.1.3.3 Exactitud: Análisis de muestras comerciales con adición de conservantes	91
3.1.3.4 Análisis de muestras artificiales con matrices significativamente diferentes a las utilizadas para calibrar	92
3.1.3.5 Análisis de muestras comerciales	95
3.2 MÉTODO DESARROLLADO PARA LA CUANTIFICACIÓN DE SORBATO DE POTASIO Y BENZOATO DE SODIO EN MUESTRAS COMERCIALES, APLICANDO METODOS DE CALIBRACIÓN MULTIVARIADA DE SEGUNDO ORDEN (PARAFAC Y MCR-ALS) A DATOS DE ESPECTROS DE ABSORBANCIA OBTENIDOS EN GRADIENTE DE pH	96
3.2.1 Introducción	96
3.2.2 Calibración	98
3.2.3 Validación	100
3.2.3.1 Juego de validación	100
3.2.3.1 Muestras comerciales	106
3.3 MÉTODO DESARROLLADO PARA LA CUANTIFICACIÓN DE COLORANTES EN JUGOS COMERCIALES, APLICANDO UN NUEVO METODO DE CALIBRACIÓN MULTIVARIADA DE SEGUNDO ORDEN (BLLS) A DATOS ESPECTRALES EN GRADIENTE DE pH	109
3.3.1 Introducción	109
3.3.2 Calibración	113
3.3.3 Validación	113
3.3.4 Muestras reales	117
<b>CAPÍTULO 4: CONCLUSIONES</b>	122
<b>CAPÍTULO 5: RESUMEN</b>	124
<b>BIBLIOGRAFÍA</b>	127

# ÍNDICE DE TABLAS

	Pág.
Tabla 1.1: Clasificación de la calibración multivariada, de acuerdo al orden del instrumento.	14
Tabla 2.1: Algoritmo cuadrados mínimos alternantes.	48
Tabla 3.1: Diseño según el cual se prepararon las mezclas del juego de calibración.	86
Tabla 3.2: Región espectral de trabajo, parámetros de la calibración y cifras de merito obtenidos en la calibración aplicando los métodos quimiométricos PLS1 y HLA/GO.	88
Tabla 3.3: Composición del juego de validación, resultados predichos aplicando el método HLA/GO y porcentajes de recuperación obtenidos.	89
Tabla 3.4: Parámetros estadísticos obtenidos en el estudio de precisión del método.	90
Tabla 3.5: Estadísticos obtenidos al aplicar regresión lineal a los resultados de la recuperación	92
Tabla 3.6: Resultados de las predicciones de las cinco muestras preparadas con jugos no modelados. Se presentan también los rangos de longitudes de onda seleccionados en cada caso, y el EIF correspondiente.	94
Tabla 3.7: Resultados obtenidos aplicando HLA/GO y el método oficial de la AOAC, para SOR y BEN en dos muestras comerciales (JN1 y JN2).	95

Tabla 3.8: Concentraciones para SOR y BEN en las mezclas de calibración, en solución reguladora BRH.	96
Tabla 3.9: Concentraciones para SOR y BEN en las mezclas de validación	100
Tabla 3.10: Pasos que se siguen en la calibración y predicción para los métodos PARAFAC y MCR-ALS utilizando datos de segundo orden.	102
Tabla 3.11: Valores obtenidos al predecir las mezclas de validación mediante los métodos <i>N</i> -PLS, PARAFAC y MCR-ALS.	105
Tabla 3.12: Resultados obtenidos para SOR y BEN al predecir cuatro muestras comerciales mediante tres métodos quimiométricos ( <i>N</i> -PLS, PARAFAC y MCR-ALS), además del método oficial.	107
Tabla 3.13: Resultados obtenidos aplicando el test de comparación de medias entre métodos, los cálculos se realizaron con las concentraciones expuestas en la Tabla 3.12.	108
Tabla 3.14: Predicciones de los cuatro colorantes en las mezclas de validación.	116
Tabla 3.15: Resultados de las predicciones para tres muestras reales y adiciones realizadas en estas tres muestras.	118
Tabla 3.16: Probabilidades obtenidas para el test <i>t</i> por parejas analizando las predicciones para las muestras reales.	120

# ÍNDICE DE FIGURAS

	Pág.
Figura 1.1: Datos de orden cero. Una muestra genera un solo dato, y el conjunto de calibración resulta un vector que contiene las medidas para las $I$ muestras de calibración.	10
Figura 1.2: Curva de calibrado, a cada valor de concentración le corresponde un dato de absorbancia.	10
Figura 1.3: Datos de orden uno. Cada muestra genera un vector de datos, por ejemplo una muestra medida a $J$ longitudes de onda. El conjunto de $I$ muestras de calibración genera una matriz de $J \times I$ .	11
Figura 1.4: Espectro de absorbancia de un analito, a una determinada concentración.	11
Figura 1.5: Datos de segundo orden. Cada muestra genera una matriz de datos, y el conjunto de calibración resulta una matriz cúbica.	12
Figura 1.6: Superficie correspondiente a una muestra de calibración, que contiene sorbato de potasio y benzoato de sodio.	12
Figura 1.7: Representación esquemática de un cubo de datos, adquiridos para una sola muestra. Mediante la técnica de fluorescencia puede obtenerse una matriz de excitación-emisión resuelta en el tiempo.	13
Figura 1.8: Grafico <i>PRESS</i> vs. $A$ . Puede observarse que el menor <i>PRESS</i> corresponde a $A = 3$ , pero para $A = 2$ se cumple que la probabilidad de $F2$ es menor de 0,75, por lo que el número de factores que se elige es 2.	21
Figura 1.9: Espacio definido por todos los analitos menos $k$ ( $\mathbf{R}_k$ ) y la proyección del vector $\mathbf{r}$ en el espacio ortogonal a dicho plano ( $\mathbf{r}_k^*$ ). $\mathbf{s}_k^*$ es la proyección del vector $\mathbf{s}_k$ en el espacio NAS.	31

Figura 1.10: Relación que existe entre la señal neta de la muestra ( $\mathbf{r}_k^*$ ) línea cortada, y la señal neta del analito puro ( $\mathbf{s}_k^*$ ) línea completa, donde la constante de proporcionalidad es $c_{k,un}$ .	32
Figura 1.11: Gráfica de regresión de las NAS (NASRP) para el componente 1 en una muestra hipotética que contiene dos componentes (1 y 2), en la que se observa linealidad.	33
Figura 1.12: Grafico de la función EI en función de la ventana y el primer sensor.	40
Figura 1.13: Gráfico de contorno mostrando un mínimo EI en primer sensor = 31 y ancho de ventana = 33.	41
Figura 1.14: Gráficos de regresión de las NAS para dos regiones: a) total y b) restringida, las interferencias son eliminadas utilizando la función EI.	41
Figura 1.15: Modelo PARAFAC de dos componentes y arreglo de tres vías.	45
Figura 1.16: Representación del desdoblamiento de una hipermatriz $\mathbf{R}$ ( $IJK$ ) de tres vías, en otra $\mathbf{R}_k$ ( $I \times JK$ ) de dos vías.	46
Figura 1.17: Hipermatriz de datos de calibración, a la que se agrega la matriz de la muestra incógnita. Resultando una matriz de dimensiones $(I+1) \times J \times K$ .	49
Figura 1.18: Representación del desdoblamiento de una hipermatriz $\mathbf{R}$ de tres vías $[(I+1)JK]$ , en otra $\mathbf{D}$ de dos vías $[(I+1)K \times J]$ . A la matriz del juego de calibración, se le agregan los datos de la muestra incógnita, de esta manera se incorpora información de las interferencias no modeladas.	52
Figura 1.19: Esquema de la descomposición de una matriz de datos ( $\mathbf{D}$ ), en dos matrices $\mathbf{C}$ y $\mathbf{S}$ que contienen los perfiles de concentración y espectrales de los componentes puros, respectivamente.	53
Figura 1.20: Restricción de no-negatividad en los perfiles de concentración. Como se observa en la figura, deben corregirse los valores negativos.	54

Figura 1.21: Restricción de uni-modalidad de los picos en los perfiles espectrales. Deben suprimirse picos menores.	54
Figura 1.22: Restricción del rango local / selectividad. Se suprimen picos menores de alguno de los componentes, cuando coinciden con los picos de otro de los componentes, de esta manera se mejora la selectividad.	54
Figura 1.23: Restricción balance de masas o restricción de cierre. La sumatoria de las concentraciones de las especies debe ser constante.	55
Figura 1.24: Diagrama de flujo. Pasos seguidos en la descomposición bilineal de la matriz <b>D</b> .	56
Figura 1.25: Esquema del desdoblamiento de las $\mathbf{X}_{c,i}$ matrices de las $I$ mezclas de calibración, para obtener los vectores $\text{vec}(\mathbf{X}_{c,i})$ . En un siguiente paso estos vectores se ordenan para formar la matriz $\mathbf{V}_x$ . La matriz $\mathbf{V}_x$ se descompone en dos matrices mediante el mismo mecanismo que utiliza CLS.	59
Figura 1.26: Reformato de los vectores $\text{vec}(\mathbf{S}_{Nc})$ , descomposición de la nueva matriz $\mathbf{S}_{Nc}$ en valores singulares, y calculo de la matriz para la predicción $\mathbf{S}_{cal}$ .	60
Figura 2.1 Pantalla del programa Multivar.	69
Figura 2.2: Pantalla del programa MVC1.	74
Figura 2.3: Pantalla del programa MVC2.	78
Figura 2.4: Sistema FIA para la cuantificación de conservantes.	79
Figura 2.5: Esquema del perfil del gradiente de pH. Si se registra la señal a 270 nm, se tiene una línea de base mientras circulan por el sistema las muestras o patrones en medio ácido, cuando se inyecta el bolo alcalino la señal se modifica, presentando un pico.	80



Figura 2.6: Sistema FIA para la cuantificación de colorantes.	81
Figura 3.1: Espectros de absorción electrónica de las soluciones en HCl $5 \times 10^{-4}$ mol L <sup>-1</sup> : A) SOR (10,0 mg L <sup>-1</sup> ), B) BEN (10,0 mg L <sup>-1</sup> ) y C) jugos de fruta de diferentes orígenes (diluidas 1/100 con el mismo ácido).	85
Figura 3.2: Espectros de absorbancia de las 15 mezclas de calibración, adquiridos entre 210 y 300 nm.	87
Figura 3.3: Espectros de absorción electrónica para: A) jugo de fruta con el que se modelaron las interferencias, y B) jugo de fruta no modelado: manzana, ambos diluidos 1/100 en HCl $5 \times 10^{-4}$ mol L <sup>-1</sup> . Los otros dos espectros presentes corresponden al SOR y BEN (ambas de 10,0 mg L <sup>-1</sup> ).	93
Figura 3.4: A) NASRP para SOR en una muestra de jugo de manzana, en un rango de 210-300 nm. B) NASRP para la misma muestra pero en el rango seleccionado (220-256 nm) mediante el estudio del EIF. La línea sólida corresponde a la recta de regresión para los puntos calculados.	93
Figura 3.5: Espectros de absorción: A) ácido benzoico, B) benzoato, C) ácido sórbico, D) sorbato, E) jugo de naranja en medio ácido y F) jugo de naranja en medio alcalino. Las concentraciones del ácido benzoico y ácido sórbico son de 10 mg L <sup>-1</sup> y para el jugo 20 %. (concentración equivalente a la que presentan las muestras comerciales). Los valores de pH fueron: pH = 2 para los espectros A y C y E y pH = 6 para los espectros B, D y F.	97
Figura 3.6: Dos gráficos superpuestos que corresponden: uno a la concentración de la especie ácida y el otro a la sal correspondiente, en función del pH. Puede observarse la dependencia lineal que existe entre ambos (la suma de las concentraciones a los diferentes pH es una constante).	98
Figura 3.7: Superficie correspondiente a la señal adquirida para una de las mezclas de calibración.	99

- Figura 3.8: A) Perfiles espectrales normalizados provistos por el método PARAFAC para 5 componentes, cuando se procesó una muestra sintética (T1). B) Perfiles de pH normalizados. En ambos casos, la numeración corresponde al orden asignado por el modelo en términos de la contribución a la varianza espectral. 104
- Figura 3.9: Espectros de absorción de cuatro colorantes. Las líneas sólidas corresponden a los espectros: tartrazina (verde), Amarillo ocaso (naranja), rojo 40 (rojo) e índigo carmín (azul) (todos en concentración de 8,0 ppm), a pH ácido (5,0). Las líneas cortadas corresponden a los espectros a pH básico (12,0). Las líneas grises representan los espectros de una solución que contiene la matriz de la muestra a pH = 5,0 (línea sólida) y a pH = 12,0 (línea cortada). 111
- Figura 10: Gráfico tridimensional de la evolución espectral en función del pH para una muestra artificial (mezcla 1 del juego de validación) que contiene los cuatro analitos. 112
- Figura 3.11: A) Cambios en los residuos de la predicción con BLLS para una muestra de validación en función del número de componentes, considerando cinco diferentes combinaciones de las especies en equilibrio para cada analito. La nomenclatura es  $[N_{s1} N_{s2} N_{s3} N_{s4}]$ , donde los números sucesivos corresponden a las especies T, A, R e I. El plano negro horizontal indica el nivel de ruido instrumental. B) Ídem pero para una muestra real. 115
- Figura 3.12 (A y B): Perfiles espectrales y de pH obtenidos mediante BLLS para una muestra real. Los analitos identificados son tartrazina (T), amarillo ocaso (A), rojo 40 (R) e índigo carmín (I). Las líneas grises corresponden a las interferencias. 119
- Figura 3.13: Intervalo elíptico de confianza para la pendiente ( $b$ ) y la ordenada al origen ( $a$ ) correspondiente a la recta de regresión entre las concentraciones nominales y las predichas para el amarillo ocaso (triángulos grises) y la tartrazina (círculos negros). La cruz negra indica el punto ( $a = 0, b = 1$ ). 121

# SÍMBOLOS

Es necesario aclarar algunas expresiones que se usarán en el desarrollo teórico de los diferentes métodos de calibración multivariada:

$J$ : número de sensores o longitudes de onda a la que se realizan las medidas.

$I$ : número de mezclas de calibración.

$K$ : número de componentes involucrados en la calibración.

$\mathbf{R}$ : matriz respuesta, de dimensiones  $J \times I$ .

$\mathbf{C}$ : matriz de  $I \times K$ , que contiene las concentraciones de los  $K$  componentes en las  $I$  mezclas de calibración.

$\mathbf{c}_k$ : vector de  $I \times 1$ , que contiene las concentraciones del analito  $k$  en las  $I$  mezclas de calibración.

$c_{k,\text{un}}$ : concentración del analito  $k$  en la muestra incógnita.

$\mathbf{c}_{\text{un}}$ : vector de  $K \times 1$ , que contiene la concentración de los  $K$  analitos en una muestra incógnita

$*$ : señal neta de analito (supraíndice).

$\| \cdot \|$ : norma euclidiana de un vector, se calcula como la raíz cuadrada de la suma de los cuadrados de sus elementos.

$+$ : matriz pseudo inversa (como supraíndice).

$\mathbf{0}$ : matriz nula (las dimensiones dependen del caso).

$\mathbf{I}$ : matriz identidad, cuyos elementos son todos 1.

$A$ : dimensionalidad (número de factores).

$\mathbf{b}_k$ : vector de  $J \times 1$ , que contiene los coeficientes de regresión para el analito  $k$ .

**E**: matriz de errores, las dimensiones dependen del caso.

**P**: matriz de  $J \times A$ . *Loadings* para la matriz **R** calculados con  $A$  factores.

**r**: vector de  $J \times 1$ , que contiene las respuestas instrumentales de una muestra incógnita medidas a  $J$  sensores.

$\mathbf{r}_{i,k}^*$ : vector de  $J \times 1$ , que corresponde a la señal neta del analito  $k$  en la muestra de calibración  $i$ .

$\mathbf{R}_{-k}$ : matriz de  $J \times I$ , que contiene las respuestas instrumentales de todos los analitos excepto  $k$  en las  $I$  muestras de calibración.

$\mathbf{r}_{un,k}^*$ : vector de  $J \times 1$ , que corresponde al espectro neto del analito para una muestra incógnita, proyectado en el espacio de la NAS del analito  $k$ .

**S**: matriz de  $J \times K$ , que contiene las respuestas instrumentales de  $K$  analitos en concentración unitaria.

$\mathbf{s}_k$ : vector de  $J \times 1$ , que contiene las respuestas instrumentales del analito  $k$  en concentración unitaria.

$\mathbf{S}_{-k}$ : matriz de  $J \times (K-1)$ , que contiene las respuestas instrumentales de todos los analitos excepto  $k$  en concentración unitaria.

$\mathbf{s}_k^*$ : vector de  $J \times 1$ , que corresponde a la señal neta de analito puro  $k$ .

**T**: matriz de  $I \times A$ . Corresponde a los *Scores* de la matriz **R** calculados con  $A$  factores

**T**: matriz transpuesta (como supraíndice).

$\mathbf{v}_k$ : vector de  $A \times 1$ , contiene los coeficientes que relacionan  $\mathbf{c}_k$  con la matriz **T**.

**W**: ( $J \times A$ ) matriz de *weight loading factors* de la matriz **R** calculados con  $A$  factores en PLS.

$\|\epsilon\|$ : Error instrumental, calculado a partir de 10 blancos.

$\otimes$ : Producto de Kronecker.

# CAPÍTULO 1

## INTRODUCCIÓN

### 1.1 CUANTIFICACIÓN DE ADITIVOS QUÍMICOS EN ALIMENTOS. IMPORTANCIA DEL TEMA

#### 1.1.1 ADITIVOS QUÍMICOS, SU LEGISLACIÓN

El uso de aditivos químicos en la conservación de alimentos tiene una larga historia. La conservación de huevos en vidrio soluble (silicato de sodio) ha sido empleada durante siglos en Europa, así como el agregado de sal, especias, y o vinagre en carnes y vegetales.

Actualmente los aditivos químicos juegan un papel preponderante en la industria alimenticia. En los países subdesarrollados la falta de facilidades para el almacenamiento, y el inadecuado transporte de los alimentos hace necesario el agregado de aditivos para su conservación, al igual que en regiones calurosas y húmedas, este agregado evita la pérdida de los excedentes de producción de temporada. En países altamente desarrollados, donde existe disponibilidad para el transporte y el almacenamiento de alimentos refrigerados, disminuye, si es que no se elimina, la necesidad del uso de preservativos químicos. Sin embargo en estos países, como en el resto del mundo, existe una demanda creciente por alimentos más atractivos, de calidad uniforme, comercializables durante todo el año, aun fuera de temporada y de la zona de producción, esto conlleva un importante aumento en el uso de aditivos.

La regulación en el uso de aditivos químicos ha recibido la atención de los gobiernos de la mayoría de las naciones del mundo. La Organización para Alimentos y Agricultura (FAO) y la Organización Mundial de la Salud (OMS) de las Naciones Unidas han abordado el problema desde el punto de vista internacional. Un comité mancomunado

de expertos sobre aditivos alimenticios, de la FAO y la OMS, en 1956, redactó una serie de normas (Codex Alimentarius Mundial) donde se formulan los principios generales que regulan el uso de aditivos alimentarios.<sup>3</sup> Dicho comité definió “aditivos alimentarios: como sustancias no nutritivas añadidas intencionalmente al alimento, generalmente en pequeñas cantidades, para mejorar su apariencia, sabor, textura o propiedades de almacenamiento”.<sup>1</sup>

En 1959 en el “VII Congreso Latinoamericano de Química” reunido en México, se aprobó el proyecto de Código Latinoamericano de Alimentos, y se creó una Comisión Permanente del Código Latinoamericano de Alimentos, encargada de realizar gestiones para que los países latinoamericanos adopten o tengan en cuenta las normas reglamentarias establecidas en dicho código.<sup>2</sup>

En el “Congreso Argentino de Legislación Alimentaria”, llevado a cabo en la ciudad de Santa Fe, en abril de 1964, se declaró: “oportuna, necesaria y conveniente la sanción de una Ley de alcance nacional que establezca las condiciones higiénico-sanitarias que deben reunir los alimentos”. A consecuencia se redactó un proyecto que fue sancionado con fuerza de ley (Ley 18.284) en el año 1964, relacionado con la implantación de normas comunes para todo el país en lo concerniente a la producción, elaboración y circulación de alimentos de consumo humano en cuanto se refiere a las condiciones higiénico-sanitarias, bromatológicas y de identificación comercial que deben satisfacer. En el artículo 1º de esta Ley se declara vigente en todo el territorio de la Nación el Código Alimentario Argentino.<sup>3</sup> Mientras que en nuestra provincia el 16 de diciembre de 1941 es sancionado con fuerza de ley (Ley 2998) el Código Bromatológico de la Provincia de Santa Fe, y se establecen las funciones que deben cumplir el Instituto Bromatológica y Departamento Químico dependientes del Ministerio de Salud Pública y Bienestar Social.<sup>4</sup>

El tipo de sustancia que se utiliza como aditivo, así como la cantidad en la que se agrega esta perfectamente legislado, y es controlado de manera constante por organismos oficiales.

La legislación justifica el uso de aditivos alimenticios cuando sirve a los siguientes propósitos:

- 1- El mantenimiento de la calidad nutritiva de un alimento.
- 2- El aumento del mantenimiento de la calidad o estabilidad dando como resultado una reducción en la pérdida de alimentos.

3- Hacer atractivos los alimentos al consumidor de tal forma que no lleve a engaño.

4- Proporcionar ayudas esenciales en el procesado de alimentos.

Cualquier decisión para usar un aditivo intencional debe basarse en el juicio considerado por científicos apropiadamente calificados, los que determinan el nivel en el que puede agregarse. La administración del aditivo será sustancialmente más baja que cualquier nivel que pudiera ser perjudicial para los consumidores. Está establecido que la cantidad de aditivo autorizado debe ser la mínima necesaria para producir el efecto deseado. Este mínimo está establecido según: el nivel de consumo estimado del alimento, los niveles mínimos en los que producen desviaciones significativas en el comportamiento fisiológico normal y el margen de seguridad para reducir al mínimo los riesgos en la salud.

El Comité Internacional de la FAO-OMS opina que los aditivos alimenticios deben ser usados para suplementar la efectividad de los métodos tradicionales de conservación y no para su reemplazo. Los riesgos asociados con el incremento de aditivos deben ser sopesados contra los beneficios obtenidos de prevenir las pérdidas de alimentos y en poner a disposición más alimentos en áreas que no son necesarias. La ley estipula que los consumidores deben ser informados de la presencia de aditivos en su alimento y el método más efectivo para lograrlo es la declaración en la etiqueta.

Todo esto determina la importancia de desarrollar nuevos métodos analíticos, más sensibles, rápidos y económicos para la cuantificación de aditivos químicos en alimentos.

### **1.1.2 CLASIFICACIÓN DE LOS ADITIVOS QUÍMICOS**

Los aditivos pueden clasificarse según sus funciones:<sup>3</sup>

1) Preservativos:

- a) Pudrición microbiológica
- b) Deterioración química
- c) Sustancias químicas para el control de insectos y roedores

2) Suplementos nutritivos:

- a) Vitaminas
- b) Aminoácidos
- c) Minerales

- d) Calorías
- 3) Modificadores del color:
  - a) Materias naturales colorantes
  - b) Colorantes artificiales
- 4) Agentes impartidores de sabor:
  - a) Sintéticos
  - b) Naturales
  - c) Aumentadores o extensores del sabor
- 5) Sustancias químicas que afectan las propiedades funcionales de los alimentos:
  - a) Control de las propiedades coloidales
  - b) Agentes afirmadores
  - c) Agentes de maduración
- 6) Sustancias químicas usadas para el procesamiento de los alimentos:
  - a) Para propósitos sanitarios, de salud pública o estéticos
  - b) Para facilitar la eliminación de cubiertas no deseadas (pieles, cueros, plumas, etc.)
  - c) Agentes antiespumantes
  - d) Agentes atrapadores
  - e) Levaduras nutrientes
- 7) Sustancias químicas para controlar la humedad:
  - a) Ceras
  - b) Agentes antitorta
- 8) Sustancias químicas usadas para el control del pH:
  - a) Ácidos
  - b) Alcalis
  - c) Sales
- 9) Misceláneos:
  - a) gases (donadores de presión)

Dentro de esta clasificación nos ocuparemos de los preservativos químicos y de los modificadores del color.



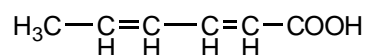
### 1.1.3 PRESERVATIVOS QUÍMICOS

Se designa de esta manera cualquier sustancia química que, cuando es añadida al alimento, tiende a prevenir o retardar su deterioro. Se excluyen los preservativos naturales como la sal común de mesa, vinagres, especias y sus aceites, y aquellas sustancias incorporadas a los alimentos por exposición directa del producto al humo de madera.

Los preservativos agregados a los alimentos deben ser señalados en la etiqueta, estableciendo la sustancia química como un ingrediente, y el nombre debe ser declarado en términos comprensibles por los consumidores. Si dos o más preservativos químicos son añadidos a un alimento cada uno debe ser señalado individualmente.

Entre los preservativos químicos más utilizados se encuentran el ácido sórbico, el ácido benzoico y sus sales.

#### ÁCIDO SÓRBICO

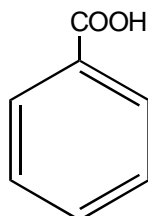


Esquema 1.1: Estructura química del ácido sórbico.

En general los ácidos grasos poli-insaturados son efectivos como agentes antifúngicos. El ácido sórbico, en particular, ha sido encontrado muy útil en el control del crecimiento del moho y de levaduras, aunque su acción no es tan eficaz con las bacterias. Se utiliza en la conservación de todo tipo de alimentos (bebidas refrescantes, caramelos, productos de confitería, conservas vegetales, carnes, etc.).<sup>5</sup> Se ha encontrado que la influencia inhibidora ejercida sobre el crecimiento del moho por el ácido sórbico es debida a la inhibición de la enzima deshidrogenasa. Cuando el nivel de desarrollo del moho es bajo el ácido sórbico tiene efectos fungistáticos y hasta fungicidas sobre los organismos. En presencia de altos niveles de crecimiento de moho, el ácido sórbico es metabolizado y no parece tener ninguna influencia inhibidora.

El ácido sórbico ha sido objeto de estudio intensivo. La demostración de que el metabolismo de este ácido graso poli-insaturado es normal y similar al metabolismo del ácido caproico, es prueba inequívoca de que este preservativo no es nocivo.

## ÁCIDO BENZOICO



Esquema 1.2: Estructura química del ácido benzoico.

El ácido benzoico y sus sales de sodio y de potasio son conservadores aceptados internacionalmente, ya que en todas las pruebas efectuadas se ha comprobado la ausencia de toxicidad.<sup>5</sup>

Se piensa que la molécula no disociada del ácido es la que presenta poder germicida. En un alimento altamente ácido, la acción germicida es del orden de 100 veces más efectiva que en un alimento ligeramente alcalino.

Los benzoatos son usados comúnmente en la preservación de sidra de manzana, pescado, jarabes, encurtidos, jugos, y otros alimentos ácidos, en concentraciones aproximadas de 200 a 1400 partes por millón ( $\text{mg Kg}^{-1}$ ).

El sorbato de potasio y el benzoato de sodio potencian su efectividad cuando son agregados de manera conjunta a los alimentos.

Según el Código Alimentario Argentino, en el Art. 996, estos conservantes (ácido benzoico o su equivalente benzoato de sodio, ácido sórbico o su equivalente sorbato de potasio) pueden ser agregados individualmente con una concentración máxima de  $800 \text{ mg Kg}^{-1}$ , o una mezcla de ambos, con un máximo de  $800 \text{ mg Kg}^{-1}$  calculada como la suma de ambos, en bebidas no gasificadas. En bebidas gasificadas el nivel máximo es de  $600 \text{ mg Kg}^{-1}$ ; y en los Art. 1006 y 1007 donde habla de bebidas alcohólicas elaboradas a base de una determinada especie cítrica, mezcla de cítricos, o con el agregado de jugos de especies frutales no cítricos, u hortícolas, el nivel máximo es de  $1 \text{ g Kg}^{-1}$  para el ácido benzoico o su equivalente benzoato de sodio, o  $1 \text{ g Kg}^{-1}$  de ácido

sórbico o su equivalente sorbato de potasio, o no más de  $1\text{g Kg}^{-1}$  de la mezcla de ambos.

#### 1.1.4 COLORANTES

Los colorantes son sustancias que añadidas a otras les proporcionan, refuerzan, o varían el color. Los colorantes son utilizados por el hombre desde los tiempos más remotos como aditivos de los alimentos. En principio se usaron colorantes extraídos de plantas e incluso minerales. Actualmente se utilizan colorantes artificiales o sintéticos.<sup>5</sup>

Los colorantes pueden clasificarse según su origen:

- Colorantes orgánicos, procedentes de plantas y animales tales como la clorofila, carotenos, riboflavina, etc.
- Colorantes minerales, tales como lacas, sulfato de cobre, cromato de plomo, etc. No son usados actualmente por contener iones metálicos.
- Colorantes artificiales, obtenidos por síntesis química. Se conocen más de 3000, pero se utilizan unos pocos.

Los colorantes artificiales poseen excelentes propiedades: proporcionan un color persistente, ofrecen colores variados, uniformes y de gran intensidad, son de alta pureza y bajo costo.

Los colorantes se utilizan en los alimentos por varias razones:

- 1) Dar un color uniforme. Por ejemplo el jugo de naranja, tiene diferentes colores según la variedad frutal, estado de madurez, procedencia, época del año, etc. lo que hace necesario el agregado de pequeñas cantidades de colorante para uniformar el color.
- 2) Realzar el color natural. Por ejemplo para hacer más atractivos los helados o yogures de frutilla o cereza.
- 3) Ocultar algún defecto. Salvo en casos muy leves no se debe recurrir a los colorantes por esta última razón.

Cada alimento o bebida tiene definido el tipo y número de colorantes que pueden utilizarse en su elaboración. Para algunos alimentos como la leche o el pescado no está permitido el agregado de colorantes.

En las etiquetas o envolturas de los alimentos se deben citar el o los nombres de los colorantes utilizados, o el número correspondiente a dicho colorante, asignado en el código alimentario. Las dosis a utilizar para los colorantes naturales son las que

proporcionan el aspecto deseado, sin las limitaciones impuestas a los colorantes artificiales que no pueden pasar de 300 partes por millón, en general.

Según el Código Alimentario Argentino en el Art. 996 para bebidas analcohólicas están permitidos los colorantes artificiales indicados a continuación y en las condiciones de uso que se señalan:

1) Empleo individual:

- Amaranto. Máximo:  $10 \text{ mg Kg}^{-1}$ .
- Amarillo Ocaso. Máximo:  $25 \text{ mg Kg}^{-1}$ .
- Azul Patente. Máximo:  $8 \text{ mg Kg}^{-1}$ .
- Tartrazina. Máximo:  $25 \text{ mg Kg}^{-1}$ .
- Rojo 40 o Rojo allura. Máximo:  $25 \text{ mg Kg}^{-1}$ .

2) Mezcla de hasta tres colorantes: la suma de los cocientes que resulten de dividir la cantidad a emplear por la máxima permitida no debe ser superior a 1.

## 1.2 CALIBRACIÓN MULTIVARIADA

### 1.2.1 INTRODUCCIÓN

Podemos definir quimiometría, como la aplicación de métodos matemáticos, estadísticos y basados en la lógica para la resolución de problemas químicos. El conjunto de métodos quimiométricos, que tienen por objeto la cuantificación de uno o más componentes en mezclas incógnitas de multicomponentes, son denominados: métodos de calibración multivariada.<sup>6</sup>

El nombre proviene del hecho que, a diferencia de la calibración univariada (en la que se construye una curva de calibración con patrones de un único componente), la calibración multivariada involucra el estudio de mezclas de varios componentes, tanto durante la etapa de calibración como la de predicción.<sup>7</sup>

La calibración multivariada está estrechamente relacionada con los diferentes tipos de instrumentos analíticos. Estos instrumentos, se clasifican de acuerdo a los datos que proporcionan, lo que implica, que de acuerdo al tipo de datos obtenido, es el tipo de calibración que debe realizarse. Dentro de este contexto, podemos clasificar los instrumentos del siguiente modo:<sup>8</sup>

#### 1.2.1.1 INSTRUMENTOS DE ORDEN CERO

Este tipo de instrumentos generan un único dato por muestra, por lo tanto el conjunto de calibración resulta un vector (Figura 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, en la Figura 2 puede observarse un ejemplo de este tipo de calibración (curva de calibrado).

Ejemplos de instrumentos de orden cero son los electrodos ion selectivos, los fotómetros que realizan las mediciones de la absorbancia a una única longitud de onda, los equipos de absorción atómica, los fotómetros de llama, etc.

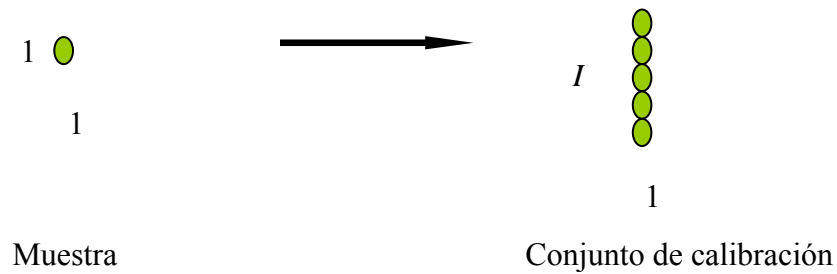


Figura 1.1: Datos de orden cero. Una muestra genera un solo dato, y el conjunto de calibración resulta un vector que contiene las medidas para las  $I$  muestras de calibración.

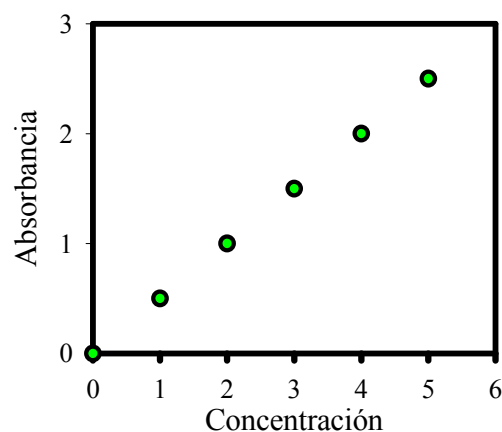


Figura 1.2: Curva de calibrado, a cada valor de concentración le corresponde un dato de absorbancia.

### 1.2.1.2 INSTRUMENTOS DE ORDEN UNO

Proveen un tensor de orden uno, o sea que por cada muestra se obtiene un vector de datos, el conjunto de calibración resulta una matriz (Figura 3).

Los instrumentos, que proveen este tipo de señal son por ejemplo: los espectrofotómetros, espectrofluorómetros, y cromatógrafos, los que son capaces de producir múltiples medidas para cada una de las muestras. En la Figura 4 puede observarse un espectro de absorción a varias longitudes de onda.

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

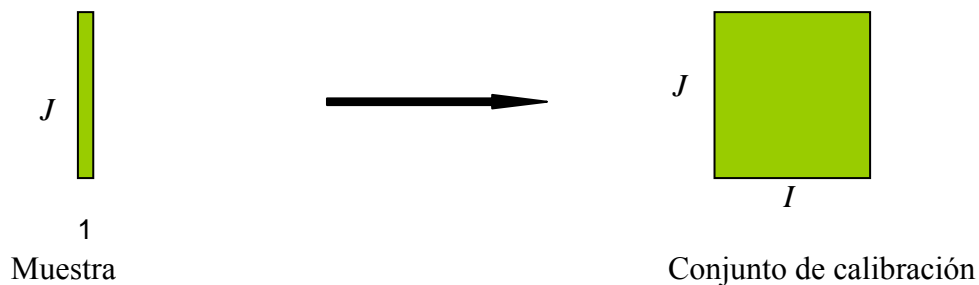


Figura 1.3: Datos de orden uno. Cada muestra genera un vector de datos, por ejemplo una muestra medida a  $J$  longitudes de onda. El conjunto de  $I$  muestras de calibración genera una matriz de  $J \times I$ .

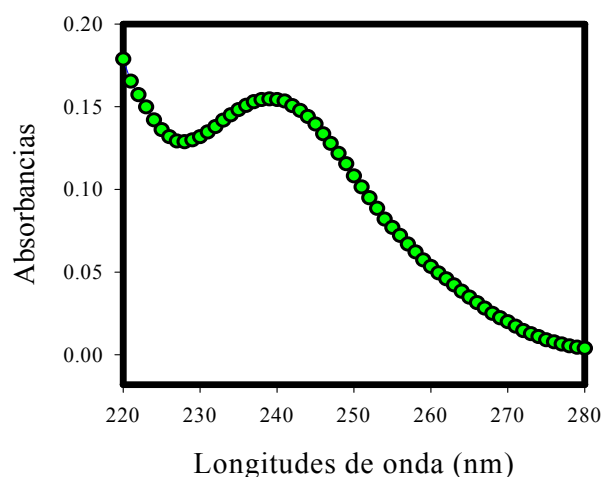


Figura 1.4: Espectro de absorbancia de un analito, a una determinada concentración.

### 1.2.1.3 INSTRUMENTOS DE ORDEN DOS

Generan una matriz de datos (tensor de orden dos) para cada muestra, el conjunto de calibración es un arreglo de tres vías (Figura 5). Este tipo de datos se genera por la aplicación de técnicas acopladas o en tándem (*hyphenated techniques*), por ejemplo, cromatografía de gases acoplada a espectrometría de masas (GC/MS), MS/MS,

resonancia magnética nuclear bidimensional, GC acoplada a espectrofotometría infrarroja, o también matrices de excitación-emisión en fluorescencia, espectroscopia UV con arreglo de diodos en gradiente de pH (Figura 6), cromatografía líquida acoplada a detector de arreglo de diodos, etc. Este tipo de datos se analiza convenientemente mediante técnicas de calibración multivariada con datos de segundo orden, que conceptualmente son diferentes de la calibración multivariada de orden uno.

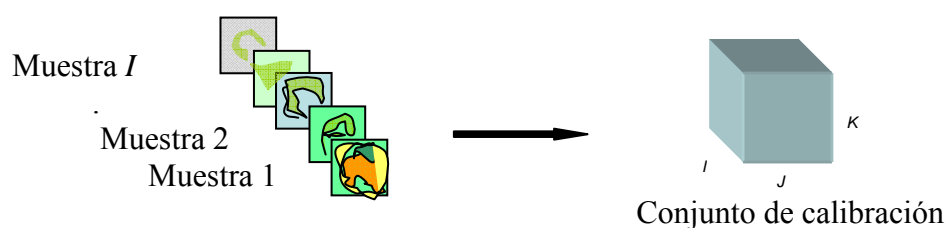


Figura 1.5: Datos de segundo orden. Cada muestra genera una matriz de datos, y el conjunto de calibración resulta un arreglo de tres vías.

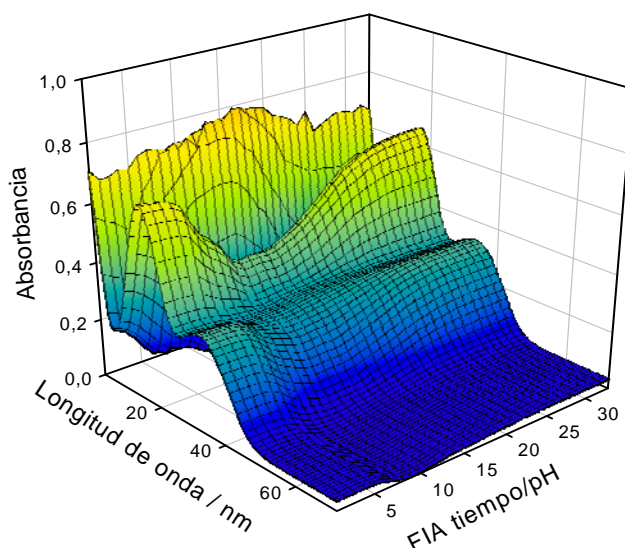


Figura 1.6: Superficie correspondiente a una muestra de calibración, que contiene sorbato de potasio y benzoato de sodio. Los datos fueron adquiridos con un espectrofotómetro con arreglo de diodos, y gradiente de pH generado mediante un sistema en flujo.



#### 1.2.1.4 INSTRUMENTOS DE ORDEN TRES Y SUPERIORES

No existe límite máximo teórico para el orden de un instrumento.

Técnicas como la espectroscopía de fluorescencia de excitación-emisión resuelta en el tiempo proveen un tensor de orden tres (Figura 7). Mediante esta técnica se obtiene una hipermatriz de datos por cada muestra, el juego de calibración resulta un arreglo de orden cuatro. Si esta misma muestra es registrada además a diferentes pH se obtienen datos de orden cuatro.

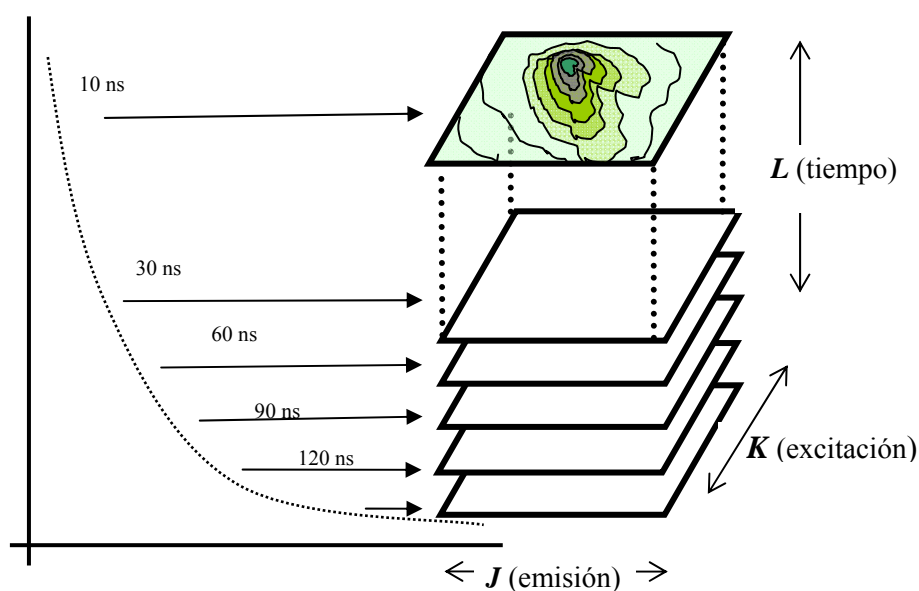


Figura 1.7: Representación esquemática de datos de tercer orden, adquiridos para una sola muestra. Mediante la técnica de fluorescencia puede obtenerse una hipermatriz de excitación-emisión resuelta en el tiempo.

La Tabla 1 resume la clasificación hasta aquí presentada.

<b>Calibración según el tipo de instrumento</b>	<b>Datos que genera una muestra</b>	<b>Conjunto de datos de un juego de calibración</b>	<b>Arreglo</b>
Orden cero	escalar	vector	una vía
Primer orden	vector	matriz	dos vías
Segundo orden	matriz	hipermatriz	tres vías
Tercer orden	hipermatriz	arreglo de 4 <sup>o</sup> orden	cuatro vías
$n$ orden	arreglo de orden $n$	arreglo de orden $n+1$	$n + 1$ vías

Tabla 1.1: Clasificación de la calibración multivariada, de acuerdo al orden del instrumento.

### 1.2.2 MÉTODOS DE CALIBRACIÓN MULTIVARIADA

La quimiometría se ha transformado en una herramienta fundamental para la química analítica. Tal es la importancia de este tema en la investigación, que en la actualidad existen revistas especializadas como el *Journal of Chemometrics* y el *Chemometrics and Intelligent Laboratory Systems*.

En la literatura se presenta una gran variedad de métodos de calibración multivariada, que utilizan respuestas de primer orden.<sup>9,10</sup> Entre los primeros métodos desarrollados se encuentran los de calibración directa: cuadrados mínimos ordinarios (OLS, del inglés *ordinary least squares*),<sup>11</sup> regresión por cuadrados mínimos clásica (CLS, del inglés *classical least squares*),<sup>12-14</sup> y los de calibración inversa: cuadrados mínimos inversos (ILS, del inglés *inversal least squares*).<sup>15,16</sup> Estos métodos en la actualidad han sido dejados de lado por otros más robustos, cuyos rendimientos han sido objeto de numerosas publicaciones.<sup>17</sup> El método de regresión por componentes principales (PCR, del inglés *principal component regresion*)<sup>18-20</sup> 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, del inglés *partial least squares*) mejora los resultados obtenidos por PCR.<sup>21-23</sup> Este último método se ha convertido en una herramienta de rutina en el laboratorio.

En este contexto, el análisis lineal híbrido (HLA, del inglés *hibrid linear analysis*) que se basa en cálculos de la señal neta del analito, resulta una alternativa al método PLS.<sup>24-27</sup>

Por otra parte, otros métodos han sido presentados y utilizados en calibración multivariada, entre ellos pueden ser citados: Regresión de Fourier,<sup>28</sup> Proyecciones Heurísticas,<sup>29</sup> Resolución de Curvas,<sup>30</sup> Curvas de Contenido Aparente,<sup>31</sup> Métodos de Adición Estándar Generalizado,<sup>32</sup> Método de Aniquilación del Rango,<sup>33</sup> Análisis de Factor,<sup>34</sup> Regresión de Hruschka,<sup>35</sup> PCR Cuadrático,<sup>36</sup> Algoritmo Genético<sup>37</sup> y Redes Neuronales.<sup>9,38</sup>

Dentro de los modelos para datos de segundo orden o superior, se han desarrollado una gran variedad de métodos. Entre los más importantes tenemos el modelo CANDECOMP (*canonical decomposition*) modelo trilineal o multilineal, desarrollado por Carroll y Chang en 1970.<sup>39</sup> Al mismo tiempo este modelo fue desarrollado por Harshman, quien lo llamó PARAFAC (*Three-way parallel factor análisis*).<sup>40</sup>

El modelo PARAFAC 2,<sup>41</sup> es una alternativa que se aplica en algunos casos, cuando el juego de datos (idealmente trilineal) no satisface el modelo PARAFAC.

El método desarrollado por Paatero (1994),<sup>42</sup> es un interesante algoritmo para ajustar modelos bilineales, siempre que los datos sean positivos, posteriormente este método fue extendido a tres vías en 1997,<sup>43</sup> asemejándose al modelo PARAFAC. Lo interesante del algoritmo es que no utiliza la alternativa de cuadrados mínimos, sino que ajusta todos los parámetros simultáneamente mediante una aproximación de Gauss-Newton.<sup>44</sup>

Muchos otros algoritmos para datos de segundo orden son descriptos, tales como: GRAFA (*generalizad rank annihilation factor análisis*),<sup>45</sup> GRAM (*generalizad rank annihilation method*),<sup>46-49</sup> TLD (*trilinear decomposition*),<sup>50,51</sup> DTD (*direct trilinear decomposition*),<sup>52-54</sup> NBRA (*nonbilinear rank annihilation*). Algunos de estos métodos son idénticos, pero nombrados de diferentes maneras, tales como GRAFA y GRAM, DTD y TLD. Todos estos métodos se basan en el algoritmo RAFA (*rank annihilation factor analisis*).<sup>55,56</sup>

El método de cuadrados mínimos bilineales, BLLS (del inglés *bilinear least squares*), fue originalmente presentado por Linder y Sundberg en el año 1998<sup>57,58</sup> y posteriormente rescatado por Olivieri et. al.<sup>59</sup> Este método permite calibrar sin tener en

cuenta las interferencias de la matriz de la muestra y obtiene los espectros de los analitos puros con los datos de calibración.<sup>57-60</sup>

MCR-ALS (del inglés *multivariate curve resolution alternating least squares*), es probablemente uno de los mejores métodos de análisis para matrices de datos multidimensionales.<sup>61-64</sup>

En la presente tesis nos dedicaremos específicamente al estudio de métodos de primer orden: PLS y HLA, y métodos de segundo orden: PARAFAC, MCR-ALS y BLLS los que fueron aplicados en el trabajo experimental a diferentes sistemas analíticos.

## 1.3 MÉTODOS DE CALIBRACIÓN DE PRIMER ORDEN

### 1.3.1 REGRESIÓN POR CUADRADOS MÍNIMOS PARCIALES (PLS)

#### 1.3.1.1 INTRODUCCIÓN

El método de cuadrados mínimos parciales, es un método de calibración inversa desarrollado con el objetivo de mejorar el método de regresión por componentes principales, utilizando factores dependientes del componente de interés.<sup>6</sup> Para ello se introducen los valores de las concentraciones de calibración en el cálculo de factores. De esta manera se crean factores dependientes de la concentración.

Este método presenta la ventaja de que permite calibrar, predecir y estudiar estadísticamente los residuos en forma separada para cada componente de una muestra, y es ideal para muestras en las que los componentes tienen intensidades espectrales muy diferentes.<sup>6,65</sup>

Existen dos tipos de métodos PLS: <sup>6,9,66</sup> uno denominado PLS-1, que concentra su atención en un único analito a la vez, y otro llamado PLS-2, que permite calibrar y predecir las concentraciones de varios analitos simultáneamente. Así, PLS-1 debe repetirse para cada analito de interés, insumiendo más tiempo de cálculo. Pero por otro lado, PLS-1 permite optimizar las condiciones de trabajo para cada analito independientemente, lo que representa una gran ventaja. Para la mayoría de las aplicaciones actuales, se prefiere PLS-1, que en adelante llamaremos simplemente PLS.<sup>17</sup>

#### 1.3.1.2 DESARROLLO TEÓRICO MATEMÁTICO

En el método PLS se descompone la matriz respuesta **R** en dos matrices (**W** y **P**), que contienen dos clases de factores espectrales o variables latentes: unos llamados *weight loading factors*, contenidos en la matriz **W**, y otros llamados simplemente *loadings*, contenidos en la matriz **P**.

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.

### **Algoritmo**

La obtención de estos factores se lleva a cabo mediante un algoritmo iterativo cíclico, que puede resumirse en los siguientes pasos:

Paso 1) Proyección de la matriz de datos  $\mathbf{R}$  en el vector de concentraciones  $\mathbf{c}_k$ , obteniéndose el primer *weight loading factor*:

$$\mathbf{w}_1 = \mathbf{R} \mathbf{c}_k / (\mathbf{c}_k^T \mathbf{c}_k) \quad (1.1)$$

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, similar a la que se hubiese realizado en CLS suponiendo la presencia de un único componente en la calibración. En este paso se aprecia la introducción de información concerniente a las concentraciones contenidas en  $\mathbf{c}_1$  en el cálculo del primer factor.

Paso 2). Se normaliza este factor, dividiéndolo por  $\mathbf{w}_1^T \mathbf{w}_1 = \|\mathbf{w}_1\|^2$ .

Este paso se realiza para normalizar  $\mathbf{w}_1$ , para contar finalmente con un juego de factores ortonormales.

Paso 3) Obtención del primer *score*:

$$\mathbf{t}_1 = \mathbf{R}^T \mathbf{w}_1 \quad (1.2)$$

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  $\mathbf{t}_1$ .

Paso 4) Obtención del primer coeficiente de regresión  $\nu_1$ :

$$\nu_1 = \mathbf{t}_1 \mathbf{c}_k / (\mathbf{c}_k^T \mathbf{c}_k) \quad (1.3)$$

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

Paso 5) Obtención del primer *loading*  $\mathbf{p}_1$ :

$$\mathbf{p}_1 = \mathbf{R} \mathbf{t}_1 / (\mathbf{t}_1^T \mathbf{t}_1) \quad (1.4)$$

Paso 6) Cálculo de los residuos espectrales y de concentración:

$$\mathbf{E}_R^T = \mathbf{R}^T - \mathbf{t}_1 \mathbf{p}_1^T \quad (1.5)$$

$$\mathbf{e}_c = \mathbf{c}_k - \nu_1 \mathbf{t}_1 \quad (1.6)$$

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, a diferencia de PCR, no explican la varianza espectral en la matriz  $\mathbf{R}$ , sino que representan un intento de explicar dicha varianza, mientras simultáneamente se correlacionan los *scores*  $\mathbf{t}_a$  con las concentraciones  $\mathbf{c}_1$ .

Paso 7) Se substituyen  $\mathbf{E}_R$  y  $\mathbf{e}_c$  por  $\mathbf{R}$  y  $\mathbf{c}_k$  respectivamente en el paso 1) y se continúa hasta llegar al número de factores deseado ( $A$ ).

### ***Número óptimo de factores (A)***

Se repite el algoritmo cíclico tantas veces como la mitad del número de mezclas más uno ( $I/2 + 1$ ). En cada ciclo se obtienen los parámetros correspondientes ( $\mathbf{w}_i$ ,  $\mathbf{t}_i$ ) y con ellos se realiza una validación interna o cruzada (método de Haaland y Thomas).<sup>6</sup> Este método 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 fue dejada de lado; luego se repite dejando de lado otra de las muestras de calibración y así sucesivamente hasta que todas fueron 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*), ecuación (1.7).

$$PRESS = \sum_1^I (c_{act} - c_{pred})^2 \quad (1.7)$$

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

Esquemáticamente:

para  $A=1$ , se calcula  $PRESS_1$

para  $A=2$ , se calcula  $PRESS_2$

·  
·  
·

para  $A=I/2+1$ , se calcula  $PRESS_{I/2+1}$

Normalmente el  $PRESS$  va disminuyendo hasta alcanzar el valor óptimo, este generalmente coincide con el número de fuentes de variación (analitos) de las mezclas, 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 el número de factores que corresponde al menor  $PRESS$ , 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.

La prueba consiste en calcular  $F$ , [ver ecuación (1.8)], que es el cociente entre cada uno de los errores calculados ( $PRESS_i$ ) y el menor error obtenido ( $PRESS_{menor}$ ), para  $i = 1, 2, \dots, I/2 + 1$ .

$$F_i = \frac{PRESS_i}{PRESS_{menor}} \quad (1.8)$$

El número óptimo de factores corresponderá al  $F_i$ , tal que:

$$F_i < F_{\alpha; I, 1}$$

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  $\{F_i > F_{\alpha; I, 1} \mid \sigma_i^2 = \sigma_{menor}^2\} = 2\alpha$ , donde  $\sigma_i^2$  y  $\sigma_{menor}^2$  son las varianzas de



los errores predichos. Con lo que se acepta el primer  $F_i$ , cuya probabilidad sea menor que 0,75.

Puede graficarse  $PRESS$  versus  $A$ , donde se visualiza cual es el número de factores al que le corresponde el menor  $PRESS$ , (ver Figura 8).

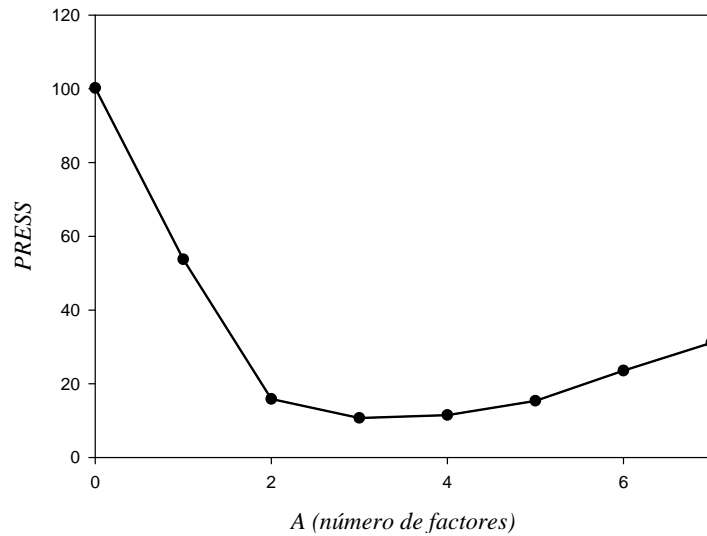


Figura 1.8: Grafico  $PRESS$  vs  $A$ . Puede observarse que el menor  $PRESS$  corresponde a  $A = 3$ , pero para  $A = 2$  se cumple que la probabilidad de  $F_2$  es menor de 0,75, por lo que el número de factores que se elige es 2.

### **Relación lineal interna en PLS**

Las matrices  $\mathbf{R}$  y  $\mathbf{c}_k$  se relacionan con la matriz de *scores*  $\mathbf{T}$  a través de las ecuaciones (1.9) y (1.10):

$$\mathbf{R} = \mathbf{P} \mathbf{T}^T \quad (1.9)$$

$$\mathbf{c}_k = \mathbf{T} \mathbf{v} \quad (1.10)$$

La matriz  $\mathbf{T}$  se obtiene a partir de (1.9) como:

$$\mathbf{T}^T = \mathbf{P}^+ \mathbf{R} \quad (1.11)$$

Siendo  $\mathbf{P}^+$  la matriz pseudoinversa de la matriz  $\mathbf{P}$ , ( $\mathbf{P}^+ = [\mathbf{P}^T \mathbf{P}]^{-1} \mathbf{P}^T$ ).

Sin embargo, las columnas de  $\mathbf{P}$  no son ortogonales, de modo que en PLS no se cumple que  $\mathbf{P}^T = \mathbf{P}^+$ , como en PCR. Para solucionar este problema, se multiplica primero la ecuación (1.9) por  $\mathbf{W}^T$ :

$$\mathbf{W}^T \mathbf{R} = \mathbf{W}^T \mathbf{P} \mathbf{T}^T = (\mathbf{W}^T \mathbf{P}) \mathbf{T}^T \quad (1.12)$$

La matriz cuadrada  $(\mathbf{W}^T \mathbf{P})$  es simple de invertir, de modo que:

$$\mathbf{T}^T = (\mathbf{W}^T \mathbf{P})^{-1} \mathbf{W}^T \mathbf{R} \quad (1.13)$$

Trasponiendo la ecuación anterior,

$$\mathbf{T} = \mathbf{R}^T \mathbf{W} (\mathbf{P}^T \mathbf{W})^{-1} \quad (1.14)$$

De acuerdo a las ecuaciones del modelo (1.9) y (1.10), si multiplicamos (1.9) por  $\mathbf{W}$  obtenemos:

$$\mathbf{R}^T \mathbf{W} = \mathbf{T} \mathbf{P}^T \mathbf{W} \quad (1.15)$$

Esta ecuación nos dice que la proyección de  $\mathbf{R}$  sobre los vectores ortonormales contenidos en  $\mathbf{W}$  da por resultado algo que podría ser definido como un nuevo *score*, *score* de las señales  $\mathbf{T}_r$ :

$$\mathbf{R}^T \mathbf{W} = \mathbf{T} \mathbf{P}^T \mathbf{W} = \mathbf{T}_r \quad (1.16)$$

Es decir que existe una relación entre el *score* de las señales  $\mathbf{T}_r$  y el *score*  $\mathbf{T}$ , al que se podría llamar *score* de la concentración o  $\mathbf{T}_c$

$$\mathbf{T}_r = \mathbf{T}_c (\mathbf{P}^T \mathbf{W}) = \mathbf{T}_c \mathbf{D} \quad (1.17)$$

Esta última ecuación se conoce como relación lineal interna del modelo PLS, donde puede observarse que el método encuentra tanto *scores* en señal, como en concentración; y que ambos se encuentran relacionados linealmente. Esto permite introducir información de concentración en los factores. El modelo podría haberse escrito de la siguiente manera:

$$\mathbf{R}^T = \mathbf{T}_r \mathbf{W}^T \quad (1.18)$$

$$\mathbf{c}_k = \mathbf{T}_c \mathbf{v} \quad (1.19)$$

### ***Coefficientes de la regresión***

Los coeficientes de la regresión pueden obtenerse a partir de la siguiente ecuación:

$$\mathbf{b}_k^T = \mathbf{c}_k^T \mathbf{R}^+ \quad (1.20)$$

$$\mathbf{b}_k^T = \mathbf{c}_k^T (\mathbf{T}^T)^+ (\mathbf{W}^T \mathbf{P})^{-1} \mathbf{W}^T \quad (1.21)$$

Trasponiendo la ecuación anterior:

$$\mathbf{b}_k = \mathbf{W} (\mathbf{P}^T \mathbf{W})^{-1} \mathbf{T}^+ \mathbf{c}_k \quad (1.22)$$

### ***Predicción basada en el score de la muestra***

La concentración del analito en la muestra incógnita se calcula según la ecuación:

$$c_{k,\text{un}} = \mathbf{r}^T \mathbf{b}_k = \mathbf{r}^T \mathbf{W} (\mathbf{P}^T \mathbf{W})^{-1} \mathbf{T}^+ \mathbf{c}_k \quad (1.23)$$

Al producto  $\mathbf{r}^T \mathbf{W} (\mathbf{P}^T \mathbf{W})^{-1}$  se le puede denominar *score* de la muestra. Este *score* se calcula con el número apropiado de factores.

Continuando con la ecuación (1.23):

$$c_{k,\text{un}} = \mathbf{t}_r^T \mathbf{T}^+ \mathbf{c}_k = \mathbf{t}_r^T \mathbf{v} \quad (1.24)$$

### ***Parámetros estadísticos indicadores de la calidad del ajuste***

Para evaluar la calidad de la calibración, una vez elegido  $A$ , pueden calcularse parámetros estadísticos, entre ellos podemos mencionar: <sup>6,9</sup>

a) *PRESS*: corresponde a la suma de cuadrados de los errores en las concentraciones predichas. Este parámetro ya fue descrito anteriormente, ecuación (1,7).

b) *RMSECV* (por *root mean square error of cross validation*), también llamado *SEP*, que como puede observarse, tiene las características de una desviación estándar:

$$RMSECV = \left[ \frac{1}{I} \sum_{i=1}^I (c_{act} - c_{pred})^2 \right]^{1/2} \quad (1.25)$$

donde  $c_{act}$  es la concentración nominal del componente en las  $I$  mezclas de calibración,  $c_{pred}$  es la concentración predicha por el modelo.

c) *REP* (del inglés *relative error of prediction*) error estándar de la predicción

$$REP\% = \frac{100}{\bar{c}} \left[ \frac{1}{I} \sum_{i=1}^I (c_{act} - c_{pred})^2 \right]^{1/2} \quad (1.26)$$

donde  $c_{act}$  es la concentración agregada del componente en las  $I$  mezclas utilizadas para la calibración,  $c_{pred}$  es la concentración predicha por el modelo y  $\bar{c}$  es la concentración promedio de las mezclas de calibración.

d) Coeficiente de correlación al cuadrado, que es un indicador de la calidad del ajuste, ya que indica el grado de correlación entre las concentraciones agregadas y las predichas:

$$r^2 = 1 - \frac{\sum_{i=1}^I (c_{act} - c_{pred})^2}{\sum_{i=1}^I (c_{act} - \bar{c})^2} \quad (1.27)$$

Es común que se definan estadísticos similares relacionando las concentraciones predichas en muestras usadas para validar modelos. Así se pueden calcular:  $RMSE_{Pred}$ ,  $REP_{pred}$  y  $r^2_{pred}$ .

#### **Detección de muestras anómalas (outliers)**

Durante la calibración puede llevarse a cabo la detección de *outliers* o datos fuera de control, mediante el cálculo de un estadístico  $F$ . La relación  $F_{1,I-1}(c_j)$ , es calculada para cada muestra durante la validación cruzada, según:

$$F_{1,I-1}(c_j) = \frac{(I-1)(e_{c_j}^2)}{\sum_{i \neq j} e_{c_i}^2} \quad (1.28)$$

donde  $e_{ci}$  es la diferencia entre la concentración nominal de analito y la concentración predicha para la  $i^{aba}$  muestra de calibración, y  $e_{cj}$  es la diferencia entre el valor de concentración nominal y el valor predicho para la muestra  $j$ .  $I$  es el número de mezclas de calibración.

Si esta relación resulta mayor que uno, la muestra es considerada *ouliers*, siendo conveniente que se deje de lado en la calibración y la predicción.

Una manera de evaluar la presencia de una interferencia en una muestra desconocida, es a través del valor del residuo espectral analizado en la ecuación (1.29). Los

programas comerciales que aplican PLS realizan un cálculo basado en una prueba  $F$ , e informan si una muestra se comporta o no como un *outlier*. La ecuación utilizada para el cálculo de  $F(\mathbf{r})$  (valor de  $F$  en función de la respuesta espectral) es la siguiente:

$$F(\mathbf{r}) = I \sum_{j=1}^J (e_{r,j})^2 / \sum_{i=1}^I \sum_{j=1}^J (e_{i,j})^2 \quad (1.29)$$

donde  $e_{r,j}$  son los residuos espectrales de la muestra desconocida  $\mathbf{r}$ , y  $e_{i,j}$  son los residuos correspondientes de las  $I$  muestras de calibración. Se puede apreciar gráficamente la aparición de interferencias no modeladas examinando los residuos espectrales.

La posibilidad de descartar muestras que presentan interferencias no modeladas y que por ende están sujetas a un alto error, es la denominada “ventaja de primer orden”.

### ***Cifras de merito***

Las cifras de merito son de suma importancia para conocer el rendimiento de un método analítico, así como para compararlo con otros métodos analíticos, incluso desarrollados con técnicas diferentes. Entre las cifras de merito del método podemos citar: limite de determinación, limite de cuantificación, selectividad, y sensibilidad para el analito,<sup>67</sup> cuyas definiciones son las siguientes:

a) Límite de Detección, es la concentración mínima que da una señal instrumental estadísticamente diferente a la señal del blanco o señal de fondo. Se calcula según la ecuación 1.30.

$$\text{LOD}_k = 3,3 S_{bl} \|\mathbf{b}_k\| \quad (1.30)$$

Donde  $S_{bl}$  es la desviación estándar del blanco y puede ser calculada registrando varias veces la señal del blanco, o calculando la norma de la señal del analito para replicados de una muestra y su correspondiente desviación estándar, y  $\mathbf{b}_k$  es el vector de los coeficientes de regresión para el analito  $k$ .

b) Límite de Cuantificación, es la concentración mínima que da una señal instrumental con la precisión suficiente como para hacer una determinación cuantitativa.

$$\text{LOQ}_k = 10 S_{bl} \|\mathbf{b}_k\| \quad (1.31)$$

c) Sensibilidad, se define por analogía con la calibración univariada, (sensibilidad igual a la pendiente de la curva de calibración):

$$SEN_k = 1 / \| \mathbf{b}_k \| \quad (1.32)$$

tiene unidades de señal/concentración.

d) Conviene también definir la sensibilidad analítica, como el cociente entre la sensibilidad y el nivel promedio del ruido

$$\gamma_k = SEN_k / S_{bl} \quad (1.33)$$

Este parámetro tiene unidades de (concentración)<sup>-1</sup>. Su inversa es la menor diferencia de concentración de un dado analito que puede apreciarse a lo largo del rango de calibración, y es útil para comparar métodos en forma independiente de la señal utilizada para cuantificar.

e) Selectividad, se define como la medida de lo bien que se puede discriminar un componente de los otros.

$$SEL_k = 1 / (\| \mathbf{b}_k \| \| \mathbf{s}_{k,LS} \|) \quad (1.34)$$

Donde  $\mathbf{s}_{k,LS}$  se aproxima mediante datos centrados, ecuación (1.35). Este cálculo se explica a continuación.

$$\mathbf{s}_{k,LS} = \mathbf{R} \mathbf{c}_k / (\mathbf{c}_k^T \mathbf{c}_k) \quad (1.35)$$

### ***Preprocesamiento de los datos***

Existe, una serie de procesamientos de los datos primarios, entre los que pueden citarse los siguientes:<sup>6,68</sup>

- 1) Derivación espectral de varios órdenes, lo que permite incrementar la resolución (a costa, sin embargo, de una pérdida de sensibilidad).<sup>69</sup>
- 2) Pesado por la varianza o escalado, que se aplica fundamentalmente en análisis multivariado, cuando se tienen datos con unidades de concentración muy distintas.<sup>70</sup>
- 3) Centrado.

El centrado es el preprocesamiento que más se aplica en calibración multivariada. Consiste en calcular el espectro promedio de la calibración (sumando las respuestas de las  $I$  muestras de calibración, obtenidas a cada sensor, y luego dividiendo por el número de muestras), y a continuación restar este espectro medio de cada uno de los espectros

de calibración. Un procedimiento análogo se aplica a la matriz  $\mathbf{c}_k$ : se resta de cada valor la concentración media de la calibración.

El centrado tiene la ventaja de hacer a los datos menos dependientes de derivas de la línea de base u otros artefactos espectrales, y a veces conduce a modelos más simples que requieren menos factores para ser definidos, lo que vuelve al problema más sencillo de tratar computacionalmente.

Cuando se calcula la concentración de una muestra incógnita, el espectro  $\mathbf{r}$  debe ser sometido a idéntico preprocesamiento que la matriz  $\mathbf{R}$ ; en el caso de centrado de datos, debe restarse de  $\mathbf{r}$  el espectro medio de calibración. Una consecuencia del centrado es que las concentraciones calculadas en las incógnitas están también centradas, y deben ser corregidas para convertirlas a resultados no centrados.

El centrado de datos provee una aproximación al cálculo del espectro puro del analito  $\mathbf{s}_k$ .

La matriz de datos  $\mathbf{R}$  contiene espectros que provienen de una serie de  $K$  analitos, cuyas concentraciones forman la matriz  $\mathbf{C}$ , que en calibración inversa es desconocida. Desconocida no significa que no existe, simplemente que el operador desconoce los valores de las concentraciones de todos los analitos (excepto  $k$ ) en  $\mathbf{C}$ . Si las respuestas siguen un modelo lineal como CLS, entonces la matriz  $\mathbf{R}$  está dada por:

$$\mathbf{R} = \mathbf{S} \mathbf{C}^T \quad (1.36)$$

aunque posiblemente  $\mathbf{S}$  también sea desconocida. Si se multiplican ambos miembros por el vector  $\mathbf{c}_k$  se obtiene:

$$\mathbf{R} \mathbf{c}_k = \mathbf{S} \mathbf{C}^T \mathbf{c}_k \quad (1.37)$$

El producto  $\mathbf{C}^T \mathbf{c}_k$  podría escribirse así:

$$\mathbf{C}^T \mathbf{c}_k = \begin{pmatrix} \mathbf{c}_1^T \\ \mathbf{c}_2^T \\ \mathbf{c}_k^T \\ \dots \\ \mathbf{c}_K^T \end{pmatrix} \mathbf{c}_k = \begin{pmatrix} \frac{\mathbf{c}_1^T \mathbf{c}_k}{\mathbf{c}_2^T \mathbf{c}_k} \\ \frac{\mathbf{c}_2^T \mathbf{c}_k}{\mathbf{c}_k^T \mathbf{c}_k} \\ \dots \\ \frac{\mathbf{c}_K^T \mathbf{c}_k}{\mathbf{c}_K^T \mathbf{c}_k} \end{pmatrix} \quad (1.38)$$

cada fila de la matriz  $\mathbf{C}^T$  corresponde al vector fila  $\mathbf{c}_K$ , mientras que cada elemento del vector  $\mathbf{C}^T \mathbf{c}_k$  es un escalar ( $\mathbf{c}_k^T \mathbf{c}_k'$ ). Si los datos están centrados, y el diseño de calibración es adecuado, estos productos escalares serán todos nulos, excepto cuando  $k = k'$ . Podemos interpretar a ( $\mathbf{c}_k^T \mathbf{c}_k'$ ) como al producto escalar de los vectores  $\mathbf{c}_k$  y  $\mathbf{c}_k'$ . Un

diseño adecuado es aquel que hace que estos productos escalares sean nulos, o que los vectores  $\mathbf{c}_k$  y  $\mathbf{c}_k'$  sean ortogonales entre sí, con lo cual:

$$\overline{\mathbf{C}}^T \overline{\mathbf{c}}_k = \begin{pmatrix} 0 \\ 0 \\ \frac{\overline{\mathbf{c}}_k^T \overline{\mathbf{c}}_k}{\overline{\mathbf{c}}_k^T \overline{\mathbf{c}}_k} \\ \dots \\ 0 \end{pmatrix} \quad (1.39)$$

y, por lo tanto:

$$\overline{\mathbf{R}} \overline{\mathbf{c}}_k = \mathbf{S} \overline{\mathbf{C}}^T \overline{\mathbf{c}}_k = |\mathbf{s}_1| |\mathbf{s}_2| |\mathbf{s}_k| \dots |\mathbf{s}_K| \begin{pmatrix} 0 \\ 0 \\ \frac{\overline{\mathbf{c}}_k^T \overline{\mathbf{c}}_k}{\overline{\mathbf{c}}_k^T \overline{\mathbf{c}}_k} \\ \dots \\ 0 \end{pmatrix} \quad (1.40)$$

donde se ha reemplazado  $\mathbf{s}_k$  por  $\mathbf{s}_{k,LS}$ , haciendo notar que la ecuación anterior es válida dentro de una aproximación de cuadrados mínimos (cuya validez también depende del diseño de calibración). Si dividimos ambos miembros de (1.40) por el escalar  $\overline{\mathbf{c}}_k^T \overline{\mathbf{c}}_k$  obtenemos:

$$\overline{\mathbf{R}} \overline{\mathbf{c}}_k = \mathbf{s}_{k,LS} \overline{\mathbf{c}}_k^T \overline{\mathbf{c}}_k \quad (1.41)$$

La ecuación (1.41) implica que, mediante datos centrados y diseño adecuado, puede obtenerse una buena aproximación a  $\mathbf{s}_k$ . Una pregunta relevante es cómo diseñar adecuadamente las mezclas de calibración para que la ecuación (1.41) proporcione una buena aproximación al espectro puro del analito  $k$ . Los diseños de factorial total son muy adecuados para estos propósitos, pero debe tenerse en cuenta que estos diseños son prohibitivos cuando el número de componentes es alto. Y aún cuando no lo fuesen, en ocasiones se desconocen las concentraciones de otros analitos, por lo que diseñar la calibración para satisfacer la ecuación (1.41) puede volverse un problema complejo. De cualquier modo, la experiencia del operador al diseñar juegos de calibración es un factor importante, que a veces supera de manera satisfactoria los inconvenientes planteados.

### ***Ventajas de PLS***



- 1- Utiliza una calibración inversa, que permite calibrar conociendo las concentraciones de un solo componente de la mezcla.
  - 2- Resuelve colinealidades espectrales, a través del uso de factores ortogonales.
  - 3- Utiliza los espectros completos.
  - 4- Permite el análisis estadístico del ajuste y provee residuos espectrales.
  - 5- Optimiza la información espectral y de concentraciones referidas al componente de interés.
  - 6- Es capaz de detectar muestras *outliers*, ventaja de primer orden.
- Desventaja: Es sensible a interferencias no modeladas.

### 1.3.2 SEÑAL NETA DEL ANALITO (NAS)

#### 1.3.2.1 INTRODUCCIÓN

El concepto de NAS (del inglés *net analyte signal*), fue introducido por Lorber,<sup>71,72</sup> quien define la señal neta para el analito  $k$  ( $\mathbf{r}_k^*$ ) como la parte del espectro que es ortogonal al espacio definido por los espectros de todos los analitos menos  $k$ .

#### 1.3.2.2 DESARROLLO TEÓRICO MATEMÁTICO

##### *Señal neta del analito en calibración univariada*

La respuesta instrumental correspondiente al sensor  $j$  ( $r_j$ ), puede relacionarse con la concentración del analito  $k$  ( $c_k$ ) en la solución que se está midiendo según:

$$r_j = r_{j0} + s_{jk} c_k + e_j \quad (1.42)$$

Donde  $r_{j0}$  es la parte constante en toda calibración y corresponde por ejemplo a la señal del solvente,  $s_{jk}$  o NAS corresponde a la absorptividad o coeficiente de extinción molar del analito  $k$  para el sensor  $j$ , multiplicado por el camino óptico, y  $e_j$  incluye los errores aleatorios y sistemáticos de la medición. De aquí el concepto de NAS como la parte de la señal útil para la predicción.<sup>73,74</sup>

Para la predicción del analito  $k$  se calcula previamente la NAS de  $k$  para  $j$ , de la siguiente manera:

$$s_{jk} = (r_j - r_{j0} - e_j) / c_k \quad (1.43)$$

luego:

$$c_{un} = r_{un} / s_{jk} \quad (1.44)$$

donde  $c_{un}$  es la concentración del analito en una muestra y  $r_{un}$  la señal medida de la muestra según el sensor  $j$ .

El concepto de señal neta del analito puede ser extendido a calibración multivariada tanto en métodos de primer orden, como en métodos de orden superior.<sup>75</sup> En la presente tesis solo analizaremos NAS para métodos de primer orden.

### ***Señal neta del analito en calibración multivariada***

#### ***Señal neta del analito en calibración directa (CLS)***

Si consideramos la ecuación que relaciona la señal con la concentración, en el modelo CLS tenemos:

$$\mathbf{R} = \mathbf{S} \mathbf{C}^T + \mathbf{E} \quad (1.45)$$

donde  $\mathbf{R}$  de  $J \times I$ , es la matriz que contiene las respuestas instrumentales, de las  $I$  muestras de calibración, medidas a  $J$  sensores,  $\mathbf{C}$  de  $I \times K$  es la matriz de concentraciones de los  $K$  analitos presentes en las  $I$  muestras de calibración, y cada columna de la matriz  $\mathbf{S}$  de  $J \times K$  ( $\mathbf{S} = [\mathbf{s}_1, \mathbf{s}_2, \mathbf{s}_k, \dots, \mathbf{s}_K]$ ) contiene las  $J$  respuestas instrumentales de cada analito puro en concentración unitaria. La matriz  $\mathbf{E}$  es la diferencia entre las medidas experimentales ( $\mathbf{R}$ ) y la señal modelada ( $\mathbf{S} \mathbf{C}^T$ ).

En la calibración directa se define la *señal neta del analito  $k$*  separada del resto de los analitos, según:

$$\mathbf{R} = \mathbf{R}_{(\text{otros analitos})} + \mathbf{R}_{(\text{analito } k)} \quad (1.46)$$

Lo que puede explicarse del siguiente modo:

$$\mathbf{R} = \mathbf{R}_{-k} + \mathbf{R}_k = \mathbf{S}_{-k} \mathbf{C}_{-k}^T + \mathbf{s}_k \mathbf{c}_k^T \quad (1.47)$$

donde  $\mathbf{C}_{-k}^T$  es una matriz de  $(K-1) \times I$ , que contiene las concentraciones de todos los analitos menos  $k$ , y  $\mathbf{S}_{-k}$  es una matriz de  $J \times (K-1)$  cuyas columnas son las sensibilidades de todos los analitos excepto  $k$ . La matriz  $\mathbf{R}_{-k}$  representa el espacio definido por todos los analitos menos  $k$ , esto significa que la matriz de proyección  $(\mathbf{I} - \mathbf{R}_{-k} \mathbf{R}_{-k}^+)$  es capaz

de proyectar el espectro  $\mathbf{r}$  de una muestra en un espacio ortogonal a todos los analitos excepto  $k$ . Esta proyección se llama señal neta del analito  $k$  y se representa como  $\mathbf{r}_k^*$ , mientras que la matriz de proyección se simboliza  $\mathbf{P}_{\text{NAS},K}$ , (ver Figura 9).

$$\mathbf{r}_k^* = (\mathbf{I} - \mathbf{R}_{-k} \mathbf{R}_{-k}^+) \mathbf{r} = \mathbf{P}_{\text{NAS},K} \mathbf{r} \quad (1.48)$$

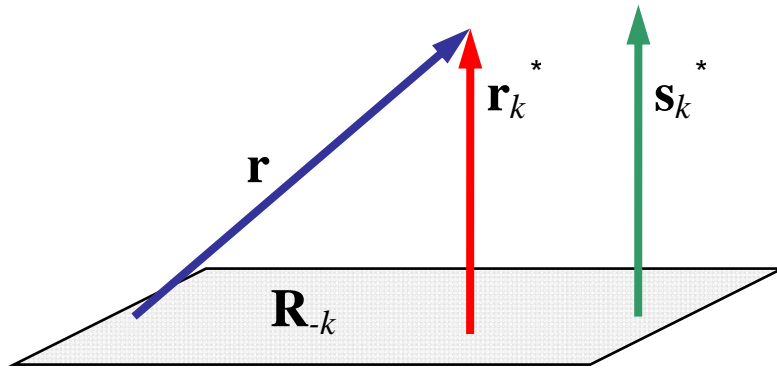


Figura 1.9: Espacio definido por todos los analitos menos  $k$  ( $\mathbf{R}_{-k}$ ) y la proyección del vector  $\mathbf{r}$  en el espacio ortogonal a dicho plano ( $\mathbf{r}_k^*$ ).  $\mathbf{s}_k^*$  es la proyección del vector  $\mathbf{s}_k$  en el espacio NAS.

En CLS, la matriz  $\mathbf{R}_{-k}$  puede obtenerse a partir de la matriz  $\mathbf{C}_{-k}$  que se supone conocida, y de  $\mathbf{S}_{-k}$  que puede construirse con las columnas apropiadas de la matriz  $\mathbf{S}$  ajustada por cuadrados mínimos. Por lo tanto la ecuación (1.48), puede escribirse como:

$$\mathbf{r}_k^* = [\mathbf{I} - \mathbf{S}_{-k} \mathbf{C}_{-k}^T (\mathbf{C}_{-k}^T)^+ \mathbf{S}_{-k}^+] \mathbf{r} = (\mathbf{I} - \mathbf{S}_{-k} \mathbf{S}_{-k}^+) \mathbf{r} \quad (1.49)$$

En CLS la matriz de proyección  $\mathbf{P}_{\text{NAS},K}$  esta dada por:  $(\mathbf{I} - \mathbf{S}_{-k} \mathbf{S}_{-k}^+)$ .

### ***Predicción de la concentración de una muestra***

Considerando el espectro de una muestra ( $\mathbf{r}$ ), tenemos:

$$\mathbf{r} = \mathbf{S} \mathbf{c} + \mathbf{e} \quad (1.50)$$

la parte de  $\mathbf{r}$  que es ortogonal al espacio definido por todos los analitos excepto  $k$  es:

$$\mathbf{r}_k^* = (\mathbf{I} - \mathbf{S}_{-k} \mathbf{S}_{-k}^+) \mathbf{r} = (\mathbf{I} - \mathbf{S}_{-k} \mathbf{S}_{-k}^+) (\mathbf{S} \mathbf{c} + \mathbf{e}) = (\mathbf{I} - \mathbf{S}_{-k} \mathbf{S}_{-k}^+) \mathbf{S} \mathbf{c} + (\mathbf{I} - \mathbf{S}_{-k} \mathbf{S}_{-k}^+) \mathbf{e} \quad (1.51)$$

si consideramos que  $\mathbf{S} = [\mathbf{S}_{-k} | \mathbf{s}_k]$  y  $\mathbf{c} = [\mathbf{c}_{-k} | c_k]$  y utilizando la propiedad de la pseudo inversa:  $\mathbf{S}_{-k} \mathbf{S}_{-k}^+ \mathbf{S}_{-k} = \mathbf{S}_{-k}$

$$\mathbf{r}_k^* = \mathbf{S}_{-k} \mathbf{c}_{-k} - \mathbf{S}_{-k} \mathbf{c}_{-k} + c_k \mathbf{s}_k^* + (\mathbf{I} - \mathbf{S}_{-k} \mathbf{S}_{-k}^+) \mathbf{e} \quad (1.52)$$

$$\mathbf{r}_k^* = c_k \mathbf{s}_k^* + (\mathbf{I} - \mathbf{S}_{-k} \mathbf{S}_{-k}^+) \mathbf{e} \quad (1.53)$$

la ecuación (1.53) nos muestra que cuando el error es igual a cero los vectores  $\mathbf{r}_k^*$  y  $\mathbf{s}_k^*$  son paralelos, están relacionados linealmente, y la constante de proporcionalidad es la concentración del analito  $k$  ( $c_k$ ) en la muestra (ver Figura 10),  $c_k$  puede ser estimado por el método de cuadrados mínimos:

$$c_k = (\mathbf{r}_k^*)^T \mathbf{s}_k^* / \|\mathbf{s}_k^*\|^2 = \mathbf{r}^T \mathbf{s}_k^* / \|\mathbf{s}_k^*\|^2 = \mathbf{r}^T \mathbf{b}_k \quad (1.54)$$

Los coeficientes para el modelo CLS son por lo tanto  $\mathbf{b}_k = \mathbf{s}_k^* / \|\mathbf{s}_k^*\|^2$ .

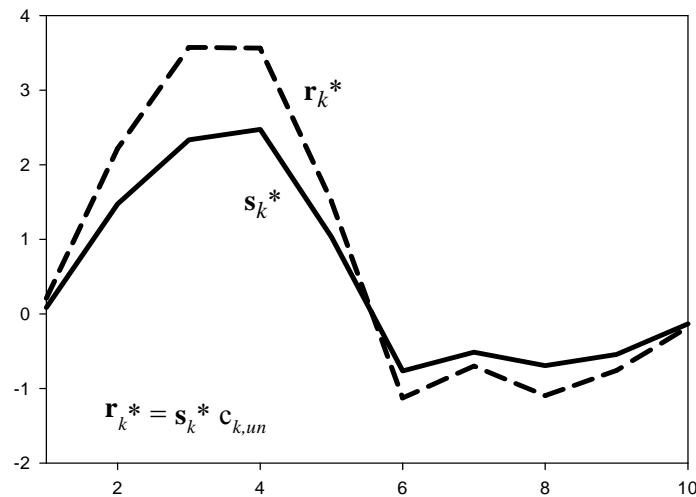


Figura 1.10: Relación que existe entre la señal neta de la muestra ( $\mathbf{r}_k^*$ ) línea cortada, y la señal neta del analito puro ( $\mathbf{s}_k^*$ ) línea completa, donde la constante de proporcionalidad es  $c_{k,un}$ .

### ***Aplicaciones de la NAS en CLS: gráficas de regresión***

En ausencia de interferencias, los espectros netos  $\mathbf{r}_k^*$  y  $\mathbf{s}_k^*$  deben ser paralelos. Puede demostrarse que la señal neta del analito 1 ( $\mathbf{r}_1^*$ ) en una mezcla que contenga los analitos 1 y 2, sin interferencias no modeladas, debe ser proporcional a la señal neta del analito 1 puro ( $\mathbf{s}_1^*$ ). Supongamos que el espectro  $\mathbf{r}$  de la mezcla está dado por:

$$\mathbf{r} = \mathbf{S} \mathbf{c}_{un} = [\mathbf{s}_1 | \mathbf{s}_2] \mathbf{c}_{un} \quad (1.55)$$

$$\mathbf{r}_1^* = \mathbf{P}_{\text{NAS}} \mathbf{r} = (\mathbf{I} - \mathbf{s}_2 \mathbf{s}_2^+) \mathbf{r} = (\mathbf{I} - \mathbf{s}_2 \mathbf{s}_2^+) [\mathbf{s}_1 | \mathbf{s}_2] \mathbf{c}_{\text{un}} \quad (1.56)$$

$$\mathbf{r}_1^* = [(\mathbf{I} - \mathbf{s}_2 \mathbf{s}_2^+) \mathbf{s}_1 | (\mathbf{I} - \mathbf{s}_2 \mathbf{s}_2^+) \mathbf{s}_2] \mathbf{c}_{\text{un}} \quad (1.57)$$

utilizando una de las propiedades de la pseudoinversa,  $\mathbf{s}_2 \mathbf{s}_2^+ \mathbf{s}_2 = \mathbf{s}_2$ , se tendrá:

$$\mathbf{r}_1^* = [(\mathbf{I} - \mathbf{s}_2 \mathbf{s}_2^+) \mathbf{s}_1 | (\mathbf{s}_2 - \mathbf{s}_2)] \mathbf{c}_{\text{un}} = [(\mathbf{I} - \mathbf{s}_2 \mathbf{s}_2^+) \mathbf{s}_1 | \mathbf{0}] \mathbf{c}_{\text{un}} \quad (1.58)$$

$$\mathbf{r}_1^* = [\mathbf{s}_1^* | \mathbf{0}] \mathbf{c}_{\text{un}} = \mathbf{s}_1^* c_{1,\text{un}} \quad (1.59)$$

donde  $\mathbf{0}$  es una columna de  $J \times 1$  ceros.

Este último análisis demuestra lo siguiente:

1. El vector  $\mathbf{r}_1^*$  es paralelo a  $\mathbf{s}_1^*$ .
2. La constante de proporcionalidad es la concentración de 1 en la incógnita, esto es  $c_{1,\text{un}}$ .

De lo anterior, podemos decir que la concentración  $c_{1,\text{un}}$  es la pendiente de una gráfica de regresión cuyas variables son  $\mathbf{r}_1^*$  y  $\mathbf{s}_1^*$ . Dicha gráfica se denomina NASRP (del inglés *net analyte signal regression plot*) (ver Figura 11).<sup>16,76</sup>

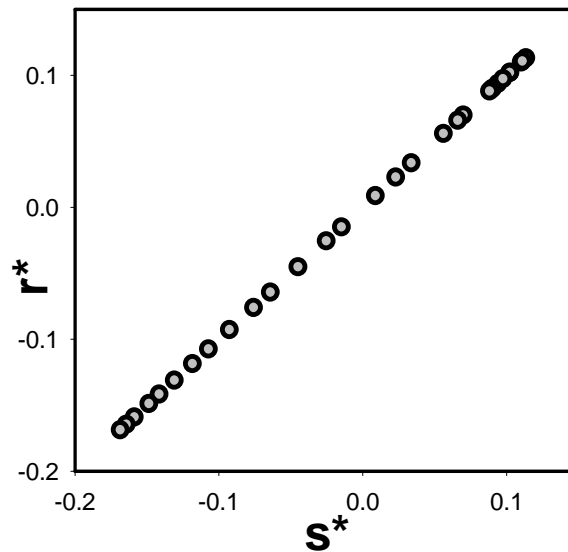


Figura 1.11: Gráfica de regresión de las NAS (NASRP) para el componente 1 en una muestra hipotética que contiene dos componentes (1 y 2), en la que se observa linealidad.

### ***Residuos en el espacio de las NAS***

La existencia de la NAS permite definir un nuevo tipo de residuos: aquellos que corresponden a la pérdida de linealidad en la relación entre  $\mathbf{r}_k^*$  y  $\mathbf{s}_k^*$ . Estos residuos se denominan residuos en el espacio de la NAS, y se calculan como:

$$\mathbf{e}^* = \mathbf{r}_k^* - c_{k,\text{un}} \mathbf{s}_k^* \quad (1.60)$$

Y dado que:  $c_{k,\text{un}} = (\mathbf{s}_k^*)^T \mathbf{r}_k^* / \|\mathbf{s}_k^*\|^2$ :

$$\mathbf{e}^* = \mathbf{r}_k^* - \mathbf{s}_k^* (\mathbf{s}_k^*)^T \mathbf{r}_k^* / \|\mathbf{s}_k^*\|^2 = [\mathbf{I} - \mathbf{s}_k^* (\mathbf{s}_k^*)^T / \|\mathbf{s}_k^*\|^2] \mathbf{r}_k^* \quad (1.61)$$

Esto significa que si definimos  $(\mathbf{s}_k^*)^T / \|\mathbf{s}_k^*\|^2$  como  $(\mathbf{s}_k^*)^+$ :

$$\mathbf{e}^* = [\mathbf{I} - \mathbf{s}_k^* (\mathbf{s}_k^*)^+] \mathbf{r}_k^* \quad (1.62)$$

La ecuación anterior implica que los residuos en el espacio NAS son las proyecciones de la NAS de la muestra en un espacio perpendicular a dicha NAS.

### ***Señal neta del analito en calibración inversa***

En la calibración inversa no es posible utilizar la expresión (1.49), ya que se desconoce, en general, la matriz  $\mathbf{S}_{-k}$ . No obstante, puede recurrirse a la ecuación general de la señal neta del analito, escrita en función de  $\mathbf{R}_{-k}$ :

$$\mathbf{r}_k^* = \mathbf{P}_{\text{NAS},k} \mathbf{r} = (\mathbf{I} - \mathbf{R}_{-k} \mathbf{R}_{-k}^+) \mathbf{r} \quad (1.48)$$

La matriz  $\mathbf{R}_{-k}$  puede obtenerse mediante cualquier procedimiento que permita descontar la contribución del analito  $k$  de los datos. Una posibilidad para el cálculo de  $\mathbf{R}_{-k}$  consiste en el uso de la siguiente ecuación:

$$\mathbf{R}_{-k} = \mathbf{R} - (\mathbf{r}/c_k) \mathbf{c}_k^T \quad (1.63)$$

donde  $\mathbf{r}$  es el espectro de una muestra cualquiera que contiene el analito  $k$  en concentración  $c_k$ , o de modo más general, una combinación lineal de espectros que contienen  $k$ . Este espectro no necesariamente debe corresponder al analito  $k$  en forma pura (lo cual en ocasiones es imposible). Se podría decir que la ecuación (1.63) no solamente cancela las contribuciones de  $k$  en  $\mathbf{R}$ , sino que agrega información respecto de otros analitos. Esto último es cierto, pero el efecto neto es irrelevante. El hecho importante es que (1.63) cancela la contribución de  $k$ , independientemente de si agrega o no información respecto de otros analitos.

Existen varias alternativas para el cálculo de  $\mathbf{R}_{-k}$  [y específicamente para la selección del espectro  $\mathbf{r}$  en la ecuación (1.63)] dependiendo de las ecuaciones de cada modelo.

Esto lleva a la aparición de varios métodos basados en el concepto de NAS en calibración inversa, como se verá a continuación.

### 1.3.3 MÉTODOS BASADOS EN SEÑAL NETA

#### 1.3.3.1 INTRODUCCIÓN

Recientemente ha sido desarrollado un grupo de métodos de calibración multivariada inversa, NMB<sub>s</sub> (del inglés *net method based signal*). Estos métodos toman el concepto introducido por Lorber, ya que calculan un vector a partir de los espectros de calibración (NAS), con el que se puede cuantificar el analito de interés sin tener en cuenta las concentraciones de los demás analitos presentes en la muestra, siempre y cuando estos hayan sido modelados en la calibración.

Este concepto de NAS es tomado del método de calibración univariada,<sup>73,74</sup> como fue descrito anteriormente, y luego extendido a métodos de calibración multivariada, tanto para calibración directa (OLS, CLS), como para métodos inversos (PCR, PLS).

Los NMB<sub>s</sub> presentan varias ventajas que los convierten en métodos robustos y muy útiles para la química analítica. Entre estas ventajas están:

- la posibilidad de calcular cifras de merito como sensibilidad, selectividad, y limite de detección.<sup>77,78</sup>
- La selección del rango de trabajo. A partir de la función error indicador puede seleccionarse el rango de longitudes de onda donde el analito de interés presenta mayor variación.<sup>25,76,79,81</sup> Esto se discutirá más adelante.
- Detección de muestras anómalas. La correlación entre la señal neta y los coeficientes de regresión permite realizar un estudio de muestras anómalas u *outliers*.<sup>82</sup>
- Desarrolla un nuevo método de pre-tratamiento de las señales. Las herramientas de preprocesamiento de NAP son comparables a las utilizadas por PLS para obtener la corrección ortogonal de la señal.<sup>83,84</sup>

- Interpretación de las predicciones en muestras de multicomponentes. Se reducen los errores por solape de espectros utilizando NAS, y mejora los errores en la predicción.<sup>74</sup>
- Se ha descrito la necesidad de calibrar con un menor número de muestras de calibración que para PLS.<sup>85,85</sup>

Todos los métodos derivan del análisis linear híbrido (HLA, del inglés *hybrid linear análisis*),<sup>24,72,87</sup> el que requiere el espectro del analito puro. Las distintas variantes para el calculo de la contribución del analito de interés, a fin de restar su contribución a la matriz  $\mathbf{R}$ , originan los siguientes métodos: a) HLA/XS, desarrollado por Xu y Schechter,<sup>89,90</sup> b) HLA/GO introducido por Goicoechea y Olivieri,<sup>72,91,92</sup> c) NAP/CLS (*net analyte pre-processing/classical least square*) preprocesamiento de la señal combinada con CLS.<sup>83,93</sup>

Resumiendo, los métodos que se denominan HLA, basados en el mismo concepto general, difieren entre si por la manera específica en la que se implementa la ecuación (1.63).<sup>87,91</sup>

$$\mathbf{R}_{-k} = \mathbf{R} - (\mathbf{r}/c_k) \mathbf{c}_k^T \quad (1.63)$$

### 1.3.3.2 DESARROLLO TEÓRICO MATEMÁTICO

A continuación se describirán las características de los métodos antes mencionados:

#### 1) MÉTODO HLA

Como fue mencionado anteriormente, en este método se calcula la matriz con las señales instrumentales de los componentes de la mezcla menos el analito  $k$ , utilizando el espectro del analito puro  $\mathbf{s}_k$ ,<sup>24</sup> por lo tanto se tiene:

$$\mathbf{R}_{-k} = \mathbf{R} - \mathbf{c}_k \mathbf{s}_k^T \quad (1.64)$$

Utilizando como matriz de proyección:

$$\mathbf{P}_{\text{NAS},k} = \mathbf{I} - \mathbf{R}_{-k}^T \mathbf{R}_{-k} \quad (1.65)$$

La concentración del analito  $k$  en una muestra incógnita se obtiene a partir del espectro  $\mathbf{r}$  según la ecuación (1.66), que es similar a la (1.54):



$$c_k = \frac{\mathbf{s}_k^T \mathbf{P}_{NAS,k} \mathbf{r}}{\mathbf{s}_k^T \mathbf{P}_{NAS,k} \mathbf{s}_k} = \frac{\mathbf{s}_k^T \mathbf{P}_{NAS,k} \mathbf{P}_{NAS,k} \mathbf{r}}{\mathbf{s}_k^T \mathbf{P}_{NAS,k} \mathbf{P}_{NAS,k} \mathbf{s}_k} = \frac{(\mathbf{s}_k^*)^T \mathbf{r}_k^*}{\|\mathbf{s}_k^*\|^2} \quad (1.66)$$

Este método es aplicable solo si se cuenta con un espectro del analito puro de alta calidad y cuando no existen interacciones de los demás componentes de la muestra.

## 2) MÉTODO HLA/XS

Este método no requiere conocer el espectro del analito puro, para obtener la matriz  $\mathbf{R}_k$ .<sup>95</sup> Cada espectro de calibración es dividido por la concentración del analito  $k$  en dicha muestra  $c_{i,k}$  (excluyendo aquellas muestras donde la concentración del analito  $k$  es cero) y luego se calcula la media de estos espectros resultantes, según:

$$\bar{\mathbf{s}}_{cal} = \frac{1}{I} \sum_{i=1}^I \frac{\mathbf{r}_{i,cal}}{c_{i,cal}} \quad (1.67)$$

donde  $I$  es el número de muestras de calibración exceptuando las que tienen  $c_{i,k} = 0$ , luego este vector calculado es utilizado para construir la matriz  $\mathbf{R}_k$ , según la siguiente operación:

$$\mathbf{R}_{-k} = \mathbf{R} - \mathbf{c}_k \bar{\mathbf{s}}_{cal}^T \quad (1.68)$$

Posteriormente se utiliza la aproximación de  $\mathbf{s}_k$  por el método de cuadrados mínimos para el cálculo de  $\mathbf{s}_k^*$ :

$$\mathbf{s}_k^* = \mathbf{P}_{NAS,k} \mathbf{s}_{k,LS} = \mathbf{P}_{NAS,k} \mathbf{R}^T \mathbf{c}_k (\mathbf{c}_k^T \mathbf{c}_k)^{-1} \quad (1.69)$$

Finalmente se calcula la concentración del analito a partir de la ecuación (1.66).

Xu y Schechter sugieren el uso de  $I$  factores de  $\mathbf{R}_k$  para la predicción, pero también se puede seleccionar el número óptimo de factores  $A$  mediante validación cruzada.<sup>89,90,95</sup>

Este método es muy similar a HLA, y puede ser usado aun cuando el número de mezclas excede el número de sensores.<sup>91,92</sup>

## 3) MÉTODO HLA/GO

Una tercera variante de los métodos basados en señal neta fue descrita por Goicoechea y Olivieri,<sup>96</sup> y presenta excelentes resultados de predicción.

En el método que estamos describiendo, primeramente se calcula la media no centrada del espectro de calibración:

$$\bar{\mathbf{r}}_{cal} = \frac{1}{I} \sum_{i=1}^I \mathbf{r}_{i,cal} \quad (1.70)$$

donde  $\mathbf{r}_{i,cal}$  es el espectro para la  $i^{aba}$  muestra de calibración.

La contribución del analito  $k$  a la matriz  $\mathbf{R}$  es calculada de la siguiente manera:

$$\mathbf{R}_{-k} = \mathbf{R} - \frac{\mathbf{c}_k \bar{\mathbf{r}}_{cal}^T}{\bar{c}_{k,cal}} \quad (1.71)$$

donde  $\bar{c}_{k,cal}$  es la media no centrada de las concentraciones de calibración del analito  $k$ .

El cálculo de  $s_k^*$  que permitirá predecir la concentración en una muestra incógnita, se obtiene a partir de la siguiente ecuación:

$$\mathbf{s}_k^* = \mathbf{P}_{NAS,k} \begin{bmatrix} \bar{\mathbf{r}}_{cal}^T \\ \bar{c}_{k,cal} \end{bmatrix} \quad (1.72)$$

#### 4) MÉTODO NAP/CLS

Este método utiliza para el cálculo de  $\mathbf{R}_{-k}$  la aproximación por cuadrados mínimos del vector  $\mathbf{s}_k$ ,  $\mathbf{s}_{k,LS}$ . Olivieri demuestra en su libro que el mejor candidato para remplazar  $\mathbf{r}_k$  en la ecuación (1.64) es  $\mathbf{s}_{k,LS}$ .<sup>97</sup>

$$\mathbf{R}_{-k} = \mathbf{R} - \mathbf{c}_k \mathbf{s}_{k,LS} \quad (1.73)$$

Luego se siguen los siguientes pasos: a) pre-procesamiento de los datos de la matriz  $\mathbf{R}$ , proyectándola en el espacio ortogonal a todos los analitos excepto  $k$ , obteniendo la matriz de datos  $\mathbf{R}_k^*$

$$\mathbf{R}_k^* = \mathbf{P}_{NAS,k} \mathbf{R} \quad (1.74)$$

y b) relacionar esta matriz de datos con el vector concentración  $\mathbf{c}_k$  según el método CLS.

Esperamos que una buena matriz de proyección  $\mathbf{P}_{NAS,k}$  realmente cancele las contribuciones de otros analitos contenidos en  $\mathbf{R}$  excepto  $k$ . Es natural esperar que los espectros netos de calibración estén linealmente relacionados con las correspondientes concentraciones, esto es, que se cumpla la ecuación siguiente:

$$\mathbf{R}_k^* = \mathbf{s}_k^* \mathbf{c}_k^T + \mathbf{E} \quad (1.75)$$

Y por lo tanto, una buena aproximación por cuadrados mínimos a  $\mathbf{s}_k^*$  será:

$$\mathbf{s}_k^* = \mathbf{R}_k^* \mathbf{c}_k (\mathbf{c}_k^T \mathbf{c}_k)^{-1} \quad (1.76)$$

A partir de la ecuación (1.60) puede también deducirse que:

$$\mathbf{R}_k^* - \mathbf{s}_k^* \mathbf{c}_k^T = \mathbf{E} \quad (1.77)$$

$$\mathbf{R}_k^* - \mathbf{R}_k^* \mathbf{c}_k \mathbf{c}_k^T (\mathbf{c}_k^T \mathbf{c}_k)^{-1} = \mathbf{E} \quad (1.78)$$

$$\mathbf{P}_{NAS,k} [\mathbf{R} - \mathbf{R} \mathbf{c}_k \mathbf{c}_k^T (\mathbf{c}_k^T \mathbf{c}_k)^{-1}] = \mathbf{E} \quad (1.79)$$

Esta última ecuación nos sugiere una excelente opción para obtener la matriz  $\mathbf{R}_{-k}$ :

$$\mathbf{R}_{-k} = \mathbf{R} - \mathbf{R} \mathbf{c}_k \mathbf{c}_k^T (\mathbf{c}_k^T \mathbf{c}_k)^{-1} \quad (1.80)$$

Si elegimos  $\mathbf{R}_{-k}$  de modo que se cumpla (1.80), la proyección de esta última matriz en el espacio de la NAS será mínima [dentro de la aproximación de cuadrados mínimos, ver (1.79)].

Nótese que (1.80) es idéntica a:

$$\mathbf{R}_{-k} = \mathbf{R} - [\mathbf{R} \mathbf{c}_k / (\mathbf{c}_k^T \mathbf{c}_k)] \mathbf{c}_k^T = \mathbf{R} - \mathbf{s}_{k,LS} \mathbf{c}_k^T \quad (1.81)$$

Esto confirma que  $\mathbf{r}$  en la ecuación (1.63) es la aproximación  $\mathbf{s}_{k,LS}$  obtenida por centrado de datos [nos informa también que la concentración promedio  $c_k$  requerida en (1.63) es unitaria].

Debemos mencionar que la ecuación (1.66) permite calcular la señal neta del analito puro:

$$\mathbf{s}_k^* = \mathbf{R}_k^* \mathbf{c}_k (\mathbf{c}_k^T \mathbf{c}_k)^{-1} = \mathbf{P}_{NAS,k} \mathbf{R} \mathbf{c}_k (\mathbf{c}_k^T \mathbf{c}_k)^{-1} = \mathbf{P}_{NAS,k} \mathbf{s}_{k,LS} \quad (1.82)$$

### ***Cifras de merito para HLA***

Las cifras de merito, LOQ, LOD, SEN,  $\gamma$ , se calculan utilizando las mismas ecuaciones presentadas para el método PLS, pagina 25.

Por otra parte, una definición para selectividad puede ser obtenida a partir del concepto de NAS. Tomando SEL como la parte de la señal del analito que es solapada por la señal de las interferencias, se calcula como la razón entre la señal neta del analito calculada sin solape espectral, y la señal total.<sup>7,98</sup> En consecuencia el rango de variación de SEL estará entre 0 (señal completamente solapada) y 1 (libre de interferencias), y se calcula mediante la siguiente ecuación:

$$SEL = \frac{\|\mathbf{s}_k^*\|}{\|\mathbf{s}_k\|} \quad (1.83)$$

### ***Selección de la región de trabajo basada en la estrategia del ancho de ventana móvil***

Este método permite la selección de una región de longitudes de onda de trabajo donde las muestras desconocidas pueden cuantificarse, evitando zonas de solapamiento espectral, o localizar la zona en que menos interfieran componentes extraños a los

modelados. Esta selección puede realizarse mediante el cálculo de la función error indicador (EIF) utilizando una ventana móvil, y el gráfico de regresión NAS para cada muestra en particular. La expresión para el cálculo de EIF es la siguiente:

$$\text{EIF} = \left[ s_{\text{fit}}^2 \left( 1 + \frac{N^2 s_{\text{fit}}^2}{4 \| \mathbf{r}_k^* \|^2} \right) \right]^{1/2} \| \mathbf{r}_k^* \|^{\text{-}1} \quad (1.84)$$

donde  $s_{\text{fit}}$  es la desviación estándar de la línea de regresión (NASRP) en una región dada,  $N$  es el número de puntos de la recta, y  $\| \mathbf{r}_k^* \|$  es el módulo del vector  $\mathbf{P}_{\text{NAS},k}$ .

Como puede observarse el módulo del vector condiciona EIF, por lo que la recta de regresión no se reduce a un punto.

Según la ecuación (1.84) puede realizarse un gráfico tridimensional como el de la Figura 12, calculado para una ventana espectral cuyo ancho y posición inicial se va modificando.<sup>76,87</sup> Para construir este gráfico, el programa MVC1,<sup>99</sup> que se corre en entorno de MATLAB, optimiza en primer lugar el modelo, luego calcula las NAS para la muestra estudiada ( $\mathbf{r}^*$ ) y para el analito puro ( $\mathbf{s}^*$ ), y construye la NASRP correspondiente a cada una de las subregiones que va seleccionando. Utilizando los datos obtenidos en cada ventana, calcula los distintos valores del error indicador y finalmente selecciona la región que contiene el mínimo EIF. Una mejor manera de visualizar esta región es mediante un gráfico de contorno, que se presenta la Figura 13.

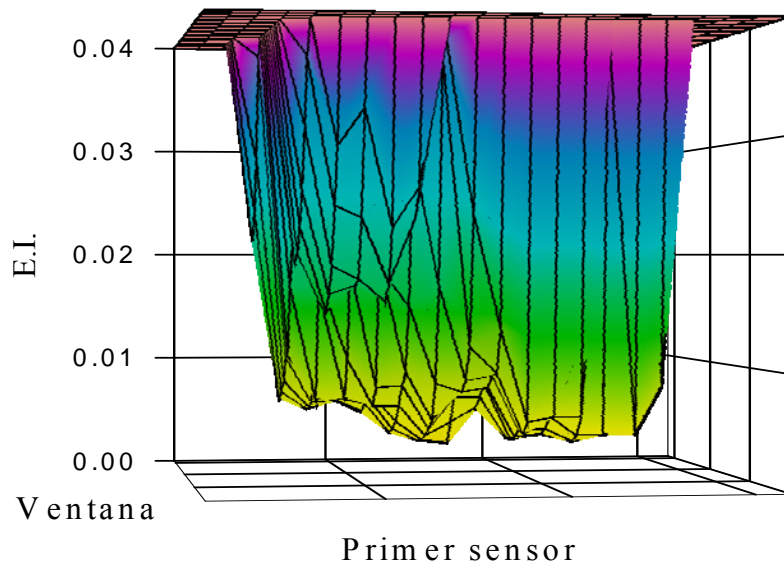


Figura 1.12: Gráfico de la función EI en función de la ventana y el primer sensor.

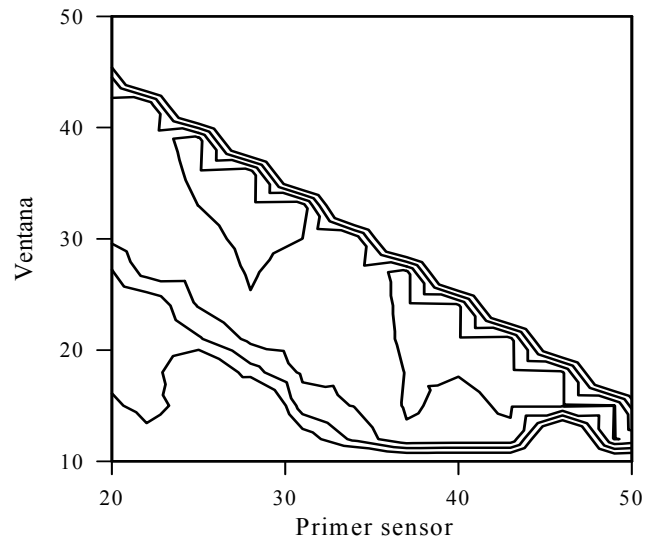


Figura 1.13: Gráfico de contorno mostrando un mínimo EI en primer sensor = 31 y ancho de ventana = 33.

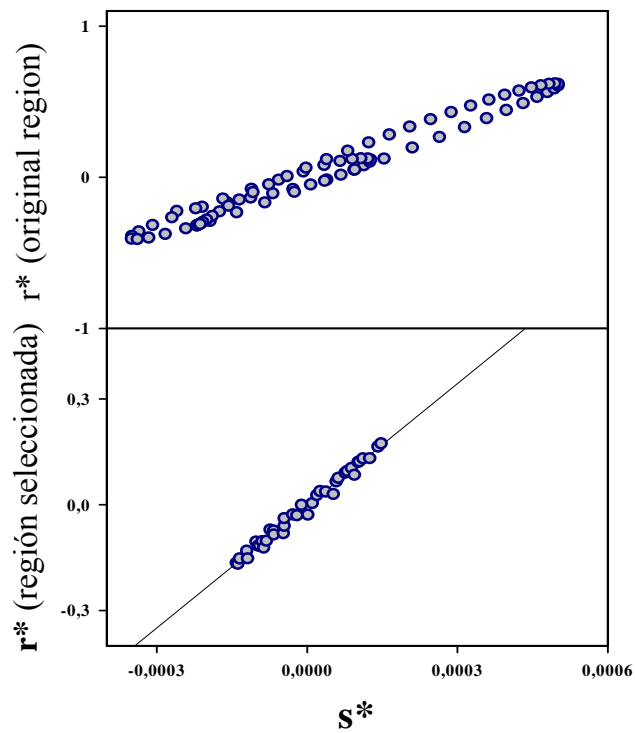


Figura 1.14: Gráficos de regresión de las NAS para dos regiones: a) total y b) restringida, las interferencias son eliminadas utilizando la función EI.

En la Figura 14 pueden observarse dos gráficos  $r^*$  vs.  $s^*$ , en el primero se consideraron todas las longitudes de onda del espectro, y en el segundo se restringe el rango a una determinada región. Claramente se observa que trabajando en un determinado rango de longitudes de onda, aparentemente no hay error en la regresión, indicando que en esta región no existen interferencias no modeladas en la calibración.



## 1.4 ANÁLISIS MULTIVARIADO DE ORDEN SUPERIOR

### 1.4.1 INTRODUCCIÓN

En análisis multivariado de primer orden estudiado hasta aquí, los datos de calibración están ordenados en estructuras de dos dimensiones, en el caso de obtenerse datos ordenados en estructuras de tres dimensiones, o superiores, cada dato puede ser indicado por tres o más índices. Debido a los avances de la tecnología aplicada al desarrollo de instrumentos de laboratorio y la creación de programas de computación más robustos, es cada vez más accesible obtener este tipo de datos.

Los métodos de análisis multi vía son una extensión natural del análisis multivariado de primer orden. Estos métodos son herramientas muy poderosas frente a diferentes problemas analíticos. Una de las más importantes ventajas, es la llamada “ventaja de segundo orden”, gracias a la cual es posible resolver muestras con matrices complejas sin necesidad de modelar las interferencias.

### 1.4.2 PARAFAC

Como ya fue mencionado, este modelo fue desarrollado simultáneamente con los nombres CANDECOMP y PARAFAC. El término PARAFAC es el más utilizado en la literatura, por lo que es el que adoptaremos a continuación.

Este modelo se ha tornado de gran interés para la quimiometría, debido a: 1) su estructura se asemeja a muchos modelos físicos, acordes con los datos instrumentales. 2) su propiedad de unicidad, lo cual implica que una cantidad de datos pueden ser descompuestos mediante este modelo, según las contribuciones individuales de cada uno de los analitos presentes.<sup>100-102</sup> Esto significa que pueden resolverse matemáticamente problemas de solapamiento espectral, o que los perfiles de ambos modos de variación (por ejemplo espectros de excitación y emisión en fluorescencia), pueden separarse según las contribuciones de cada analito en particular.



### 1.4.2.1 PERFILES PROPORCIONALES PARALELOS

El modelo PARAFAC se basa en el principio perfiles proporcionales paralelos (del inglés *parallel proportional profiles*) que fue descrito por Cattell en 1944.<sup>103</sup> Este principio enuncia que una matriz de datos puede descomponerse según un grupo de perfiles o vectores *loading*, que describen la variación en más de un sentido (vía) al arreglo de datos, estos perfiles se encuentran en diferentes proporciones o con diferentes pesos. Este principio es la propiedad fundamental para obtener optimas descomposiciones. Que los perfiles sean paralelos conlleva a que el modelo no tenga libertad rotacional.<sup>104</sup>

Tomemos el ejemplo de una matriz  $\mathbf{X}$ , que está correctamente modelada por  $\mathbf{A}\mathbf{B}^T$ , donde el número de columnas de  $\mathbf{A}$  y  $\mathbf{B}$  es dos, debido a la presencia de dos analitos en la mezcla. Este modelo puede expresarse como:

$$\mathbf{X}_1 = \mathbf{a}_1 \mathbf{b}_1^T c_{11} + \mathbf{a}_2 \mathbf{b}_2^T c_{12} \quad (1.85)$$

donde  $c_{11}$  y  $c_{12}$  son ambas iguales a uno. Supongamos ahora que tenemos otra matriz  $\mathbf{X}_2$ , que puede ser descrita por el mismo juego de vectores *scores* y *loadings* ( $\mathbf{a}_1$ ,  $\mathbf{a}_2$ ,  $\mathbf{b}_1$  y  $\mathbf{b}_2$ ) pero en diferentes proporciones:

$$\mathbf{X}_2 = \mathbf{a}_1 \mathbf{b}_1^T c_{21} + \mathbf{a}_2 \mathbf{b}_2^T c_{22} \quad (1.86)$$

donde  $c_{21}$  y  $c_{22}$  no son necesariamente iguales a  $c_{11}$  y  $c_{12}$ . Los dos modelos constan de los mismos perfiles (paralelos) solo que en diferentes proporciones. Una manera de describir un modelo combinado es:

$$\mathbf{X}_k = \mathbf{A} \mathbf{D}_k \mathbf{B}^T \quad (1.87)$$

donde  $\mathbf{D}_k$  es una matriz diagonal, con los elementos  $c_{kf}$  de  $\mathbf{C}$  en su diagonal. Este es en esencia el principio de perfiles proporcionales paralelos. Cattell fue el primero en demostrar la presencia de estos perfiles, que permiten calcular inequívocas descomposiciones.

### 1.4.2.2 PRODUCTO KHATRI-RAO

El modelo PARAFAC puede ser descrito de diferentes maneras, una muy simple es la que se desarrollara a continuación. Tomando del ejemplo anterior las ecuaciones (1.85) y (1.86)

$\mathbf{X}_1$  y  $\mathbf{X}_2$  pueden ser expresadas como una tercera matriz  $\mathbf{X}$ , tal que  $\mathbf{X} = [\mathbf{X}_1 \ \mathbf{X}_2]$  :

$$\mathbf{X} = [\mathbf{a}_1 \mathbf{a}_2] \begin{bmatrix} \mathbf{b}_1 \mathbf{c}_{11} & \mathbf{b}_2 \mathbf{c}_{12} \\ \mathbf{b}_1 \mathbf{c}_{21} & \mathbf{b}_2 \mathbf{c}_{22} \end{bmatrix}^T \quad (1.88)$$

$$\mathbf{X} = [\mathbf{a}_1 \mathbf{a}_2] \left[ \text{vec}(\mathbf{b}_1 \mathbf{c}_1^T) \text{vec}(\mathbf{b}_2 \mathbf{c}_2^T) \right]^T \quad (1.89)$$

con  $\mathbf{c}_1 = [c_{11} \ c_{21}]^T$  y  $\mathbf{c}_2 = [c_{12} \ c_{22}]^T$ .

donde 'vec' indica la operación de desdoblar una matriz y darle forma de vector.

Según el producto Khatri-Rao, definido por Mc Donald,<sup>105</sup> Rao y Mitra,<sup>106</sup> de dos matrices con el mismo numero de columnas,  $F$ , se tiene:

$$\mathbf{C}|\otimes|\mathbf{B} \equiv \left[ \text{vec}(\mathbf{b}_1 \mathbf{c}_1^T) \text{vec}(\mathbf{b}_2 \mathbf{c}_2^T) \dots \text{vec}(\mathbf{b}_F \mathbf{c}_F^T) \right] = \left[ \mathbf{c}_1 \otimes \mathbf{b}_1 \ \mathbf{c}_2 \otimes \mathbf{b}_2 \ \dots \mathbf{c}_F \otimes \mathbf{b}_F \right] \quad (1.90)$$

Este producto hace posible expresar el modelo PARAFAC como:

$$\mathbf{X} = \mathbf{A}(\mathbf{C}|\otimes|\mathbf{B}^T) \quad (1.91)$$

donde  $\mathbf{X}$  es la matriz de datos,  $\mathbf{A}$  y  $\mathbf{B}$  las matrices *loadings* que contienen los perfiles extraídos por el modelo, y  $\mathbf{C}$  la matriz que contiene los datos de concentración.

Cuando se tienen matrices de orden superior, por ejemplo para cuatro vías, tenemos:

$$\mathbf{X}^{(I \times J \times K \times L)} = \mathbf{A}(\mathbf{D}|\otimes|\mathbf{C}|\otimes|\mathbf{B})^T \quad (1.92)$$

donde  $\mathbf{D}$  es el cuarto modo de la matriz de *loadings*.

Este producto posee una serie de propiedades, una de las más importantes es la propiedad asociativa:

$$(\mathbf{D}|\otimes|\mathbf{C})|\otimes|\mathbf{B} = \mathbf{D}|\otimes|(\mathbf{C}|\otimes|\mathbf{B}) \quad (1.93)$$

la propiedad distributiva:

$$(\mathbf{B} + \mathbf{C})|\otimes|\mathbf{D} = (\mathbf{B}|\otimes|\mathbf{D} + \mathbf{C}|\otimes|\mathbf{D}) \quad (1.94)$$

y por ultimo la propiedad del producto:

$$(\mathbf{A}|\otimes|\mathbf{B})^T (\mathbf{A}|\otimes|\mathbf{B}) = (\mathbf{A}\mathbf{A}^T) \circ (\mathbf{B}\mathbf{B}^T) \quad (1.95)$$

donde el operador,  $\circ$ , es el producto de Hadamard.

Gráficamente, podemos explicar el modelo de tres vías PARAFAC, con tres matrices *loadings*, **A**, **B** y **C**, cuyos elementos son  $a_{if}$ ,  $b_{jf}$  y  $c_{kf}$  respectivamente, (ver Figura 15). Puede redefinirse la matriz de datos **X**, según cada uno de sus elementos, como:

$$\hat{X}_{ijk} = \sum_{f=1}^F a_{if} b_{jf} c_{kf} \quad (1.96)$$

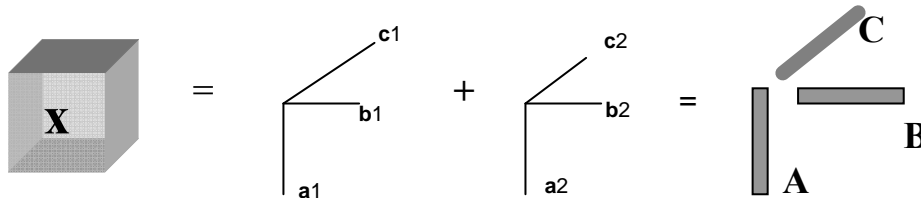


Figura 1.15: Modelo PARAFAC de dos componentes y arreglo de tres vías. Cada elemento de **X** proviene del triple producto  $a_{ij}b_{jl}c_{kl}$ <sup>107</sup>

### 1.4.2.3 ALGORITMOS

En la literatura existen una serie de algoritmos que pueden aplicarse para ajustar el modelo PARAFAC.<sup>108</sup> Entre ellos podemos citar: a) descomposición trilineal directa (DTLD del inglés *direct trilinear decomposition*),<sup>109-113</sup> b) descomposición trilineal alternante (ATLD del inglés *alternating trilinear decomposition*),<sup>114</sup> c) descomposición trilineal alternativa por pesos (SWATLD del inglés *self-weighted alternating trilinear decomposition*),<sup>115,116</sup> d) Pseudo cuadrados mínimos alternantes (PALS del inglés *pseudo alternating least squares algorithm*),<sup>117</sup> e) ACOVER (del inglés *alternating coupled vectors resolution*), f) ACOMAR (*algorithm alternating coupled matrices resolution*),<sup>118</sup> g) ASD (*alternating slice-wise diagonalization*),<sup>119</sup> y h) ALS (del inglés *alternating least squares*).<sup>121,122</sup> Este último es el más frecuentemente aplicado, y es el que describiremos, ya que fue empleado en el tratamiento de los datos experimentales de la presente tesis.

### 1.4.2.4 FUNDAMENTOS MATEMÁTICOS DE ALS

Para aplicar el algoritmo ALS, primeramente es necesario desdoblar la hipermatriz de datos, (**R** de dimensiones *IJK*) en una de sus dimensiones, lo que implica que los datos que inicialmente eran de tres vías, se conviertan en datos de dos vías.

Para explicarlo, tomemos por ejemplo, el modo  $K$  de la hipermatriz  $\mathbf{R}$ , este es fragmentado en “tajadas frontales”, que posteriormente son ordenadas según otro de los modos ( $J$ ), transformando la matriz  $\mathbf{R}$  ( $IJK$ ), en una matriz  $\mathbf{R}_k$  de dimensiones  $I \times JK$ , (ver Figura 16). La matriz  $\mathbf{R}_k$ , queda definida de la siguiente manera:

$$\mathbf{R}_k = [\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_K] \quad (1.97)$$

Donde  $I$  es el número de mezclas de calibración y  $J$  y  $K$  son los modos en que varía el sistema.<sup>123,124</sup>

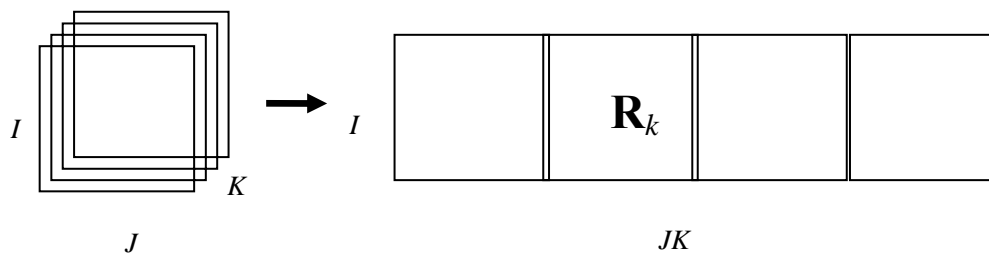


Figura 1.16: Representación del desdoblamiento de una hipermatriz  $\mathbf{R}$  ( $IJK$ ) de tres vías, en otra  $\mathbf{R}_k$  ( $I \times JK$ ) de dos vías. Nótese que desaparece el modo  $K$ .

De esta manera, el modelo PARAFAC puede ser escrito como:

$$\mathbf{R}_k = \mathbf{X} \mathbf{D}_k \mathbf{Y}^T + \mathbf{E}_k \quad (1.98)$$

$$k = 1, \dots, K$$

donde  $N$  es el número total de analitos de interés presentes en las muestras,  $\mathbf{X}$  ( $I \times N$ ) es el primer modo de la matriz de *loadings*, cuyos elementos son  $x_{ni}$ ,  $\mathbf{Y}$  ( $J \times N$ ) es el segundo modo de la matriz de *loadings*, y sus elementos son  $y_{nj}$ . El tercer modo de la matriz de *loadings* es  $\mathbf{Z}$  ( $K \times N$ ), cuyos elementos son  $z_{nk}$ ,  $\mathbf{D}_k$  es una matriz diagonal que contiene los elementos de la  $k^{\text{ava}}$  fila de  $\mathbf{Z}$ , en su diagonal. La matriz  $\mathbf{E}_k$  contiene la variación residual de  $\mathbf{R}_k$ .

Resumiendo, en la primera etapa de cálculo, se desdobla la matriz  $\mathbf{R}$  en sus tres modos, obteniendo:

$$\mathbf{R}_x = \mathbf{X}(\mathbf{Z} \otimes \mathbf{Y})^T + \mathbf{E}_x \quad (1.99)$$

$$\mathbf{R}_y = \mathbf{Y}(\mathbf{Z} \otimes \mathbf{X})^T + \mathbf{E}_y \quad (1.100)$$

$$\mathbf{R}_z = \mathbf{Z}(\mathbf{Y} \otimes \mathbf{X})^T + \mathbf{E}_z \quad (1.101)$$

donde  $\mathbf{R}_x$ ,  $\mathbf{R}_y$ , y  $\mathbf{R}_z$  son matrices de dos vías, de dimensiones  $I \times JK$ ,  $J \times IK$ , y  $K \times IJ$  respectivamente.

### **Algoritmo ALS**

1) El primer paso del algoritmo ALS, es estimar  $\mathbf{X}$ , según la siguiente ecuación:

$$\mathbf{X} = \mathbf{R}_x((\mathbf{Z} \otimes \mathbf{Y})^T)^+ \quad (1.102)$$

donde el superíndice  $^+$  indica la inversa de Moore-Penrose. Se asumen  $\mathbf{Y}$  y  $\mathbf{Z}$  constantes.

2) El segundo paso es estimar  $\mathbf{Y}$ , con la matriz  $\mathbf{X}$  obtenida en el primer paso, y  $\mathbf{Z}$  constante. Aplicando:

$$\mathbf{Y} = \mathbf{R}_y((\mathbf{Z} \otimes \mathbf{X})^T)^+ \quad (1.103)$$

3) Luego, de la manera antes descrita se estima  $\mathbf{Z}$ , con  $\mathbf{X}$  e  $\mathbf{Y}$  estimadas en los pasos anteriores:

$$\mathbf{Z} = \mathbf{R}_z((\mathbf{Y} \otimes \mathbf{X})^T)^+ \quad (1.104)$$

4) Este algoritmo se repite iterativamente hasta obtener la convergencia en un mínimo global para la matriz  $\mathbf{E}$ .<sup>125</sup>

En la Tabla 2, se resume el algoritmo PARAFAC-ALS, para una matriz de datos  $\mathbf{R}$ , y  $N$  componentes de una mezcla.<sup>108</sup>

---

### **Algoritmo PARAFAC-ALS**

Para una matriz de datos  $\mathbf{R}$ , de una mezcla de  $N$  componentes

---

- (1) Inicialmente se fijan  $\mathbf{Y}$  y  $\mathbf{Z}$
  - (2)  $\mathbf{X} = \mathbf{R}_x((\mathbf{Z} \otimes \mathbf{Y})^T)^+$
  - (3)  $\mathbf{Y} = \mathbf{R}_y((\mathbf{Z} \otimes \mathbf{X})^T)^+$
  - (4)  $\mathbf{Z} = \mathbf{R}_z((\mathbf{Y} \otimes \mathbf{X})^T)^+$
  - (5) Se vuelve al paso (2), hasta la convergencia
- 

Tabla 2.1: Algoritmo Cuadrados mínimos alternantes.

Como ya se mencionó, se llega a la convergencia aplicando el método de cuadrados mínimos, es decir minimizando la siguiente suma:

$$S = \sum \| \mathbf{R} - \mathbf{X}(\mathbf{Z} \otimes \mathbf{Y})^T \|^2 = \sum \| \mathbf{E} \|^2 \quad (1.105)$$

De esta manera mediante la aplicación del algoritmo:

- 1) Se encuentra  $\mathbf{X}$ , tal que  $S$  sea mínima, manteniendo  $\mathbf{Y}$  y  $\mathbf{Z}$  constantes.
- 2) Se encuentra  $\mathbf{Y}$ , tal que  $S$  sea mínima, manteniendo  $\mathbf{X}$  y  $\mathbf{Z}$  constantes.
- 3) Se encuentra  $\mathbf{Z}$ , tal que  $S$  sea mínima, manteniendo  $\mathbf{X}$  e  $\mathbf{Y}$  constantes.
- 4) Se vuelve a repetir hasta que el valor de  $S$  no difiera significativamente del anterior.

### ***Ventaja de segundo orden***

El modelo PARAFAC agrega al juego de datos de calibración la matriz con los datos de la muestra incógnita, por lo que al modo  $I$  debe sumarse 1, (ver Figura 17).<sup>126</sup>

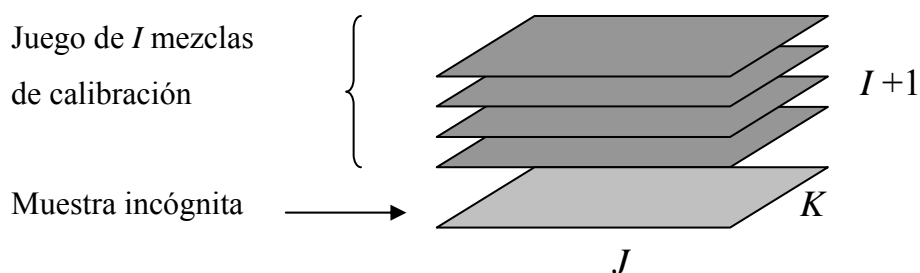


Figura 1.17: Hipermatriz de datos de calibración, a la que se agrega la matriz de la muestra incógnita, resultando una matriz de dimensiones  $(I+1) \times J \times K$ .

De esta manera se agrega información sobre las posibles interferencias de la muestra en la matriz de calibración, aunque no es necesario conocer las interferencias presentes en la matriz, ni su concentración.

### ***Predicción***

Los perfiles obtenidos en la descomposición de  $\mathbf{R}$ , corresponden a los espectros de los  $N$  analitos puros presentes en la mezcla. Estos perfiles son utilizados para construir un modelo de regresión lineal simple para cada analito en particular; lo que transforma finalmente la calibración multivariada en una calibración pseudo-univariada.<sup>15, 127</sup>

Tomando a la ecuación:

$$\mathbf{X}_{ijk} = \sum a_{ni} b_{nj} c_{nk} + \mathbf{E}_{ijk} \quad (1.106)$$

una vez resuelta la descomposición matricial y los perfiles de cada uno de los componentes ha sido calculado, los *scores* para las muestras de calibración asociados con un determinado analito, son linealmente relacionados con las concentraciones nominales del analito, según:

$$(a_{n1} | a_{ni} | \dots | a_{nI}) = k_{\text{PARAFAC}} \mathbf{y} \quad (1.107)$$

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  $k_{\text{PARAFAC}}$  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 ( $a_{n(I+1)}$ ).

$$y_u = a_{n(I+1)} / k_{\text{PARAFAC}} \quad (1.108)$$

El modelo es muy simple en términos de grados de libertad. Calibrando con un número elevado de muestras, no es necesario realizar un test de validación cruzada para evaluar la capacidad de predicción.

### ***Cifras de merito***

Para evaluar la calidad de los resultados obtenidos, pueden calcularse las siguientes cifras de merito:<sup>107,128</sup>

a) Media del cuadrado del error, (MSE del ingles *mean square error*):

$$\text{MSE}_x = \frac{\sum_{i=1}^I \sum_{n=1}^N (x_{in} - \hat{x}_{in})^2}{IN} \quad (1.109)$$

donde  $x_{in}$  corresponde a la concentración nominal y  $\hat{x}_{in}$  a la concentración predicha.

b) Falta de ajuste, (LOF del ingles *lack of fit*), que se expresa como porcentaje:

$$\text{LOF} = 100 \left( \frac{\sum_{i=1}^I \sum_{j=1}^J \sum_{k=1}^K e_{ijk}^2}{\sum_{i=1}^I \sum_{j=1}^J \sum_{k=1}^K r_{ijk}^2} \right) \quad (1.110)$$

c) Sensibilidad de sobre ajuste de factores, lo que involucra el problema de determinar la dimensionalidad correcta, esto puede ser juzgado visualmente o evaluando el modelo a través de los valores de MSE.

d) Habilidad predictiva, (RMSEC del ingles *root means squares error cross-validation*) Para un juego de datos reales se basa en la diferencia entre el valor nominal y el predicho, para un dado componente.

$$\text{RMSEC} = \left[ \frac{\sum_{i=1}^I (y_{ijk} - \hat{y}_{ijk})^2}{I} \right]^{1/2} \quad (1.111)$$

Entre las ventajas del modelo PARAFAC, están:

a) La unicidad de la solución. Los *loadings* obtenidos por descomposición bilineal reflejan los espectros puros de los componentes, pero esto no es posible sin información externa. Harshman (1972)<sup>129</sup> y Leurgans (1993)<sup>107</sup> explican la unicidad de solución cuando los vectores *loadings* son linealmente independientes en dos de sus modos, y en el tercer modo la condición restrictiva es que por lo menos dos de los *loadings* no sean linealmente dependientes.

b) Pueden aplicarse restricciones en caso de ser necesarias, c) Puede ser fácilmente extendido a datos de orden mayor, d) No es necesario calibrar las interferencias, y e) de acuerdo al número de *loadings* obtenidos puede conocerse el número de componentes de la mezcla.

Dentro de las desventajas, una es que el algoritmo es lento, y además tiene el problema, común a los modelos bilineales, de la libertad rotacional.



## 1.4.2 MCR-ALS

### 1.4.2.1 INTRODUCCIÓN

El modelo de resolución de curvas por análisis multivariado, ajustado mediante cuadrados mínimos alternantes, llamado MCR-ALS (del inglés *multivariate curve resolution-alternating least squares*), es también un método adecuado cuando se tienen datos ordenados en arreglos de tres vías. Este método se aplica cuando se pretende conocer los perfiles (espectrales, de pH, de elusión, etc.) de más de un componente en una mezcla, y poder predecir la concentración de los componentes de una muestra desconocida.<sup>130-133</sup>

### 1.4.2.2 DESARROLLO TEÓRICO MATEMÁTICO

Para resolver el modelo, primeramente es necesario desdoblar la matriz de datos (**R**) de tres vías ( $IJK$ ), en otra de dos vías, **D** de  $[(I+1)K \times J]$ , de la misma manera que ya fuese explicado para el modelo PARAFAC (ver Figura 18). En este modelo también se agrega la matriz de la muestra desconocida al juego de matrices de datos de calibración.

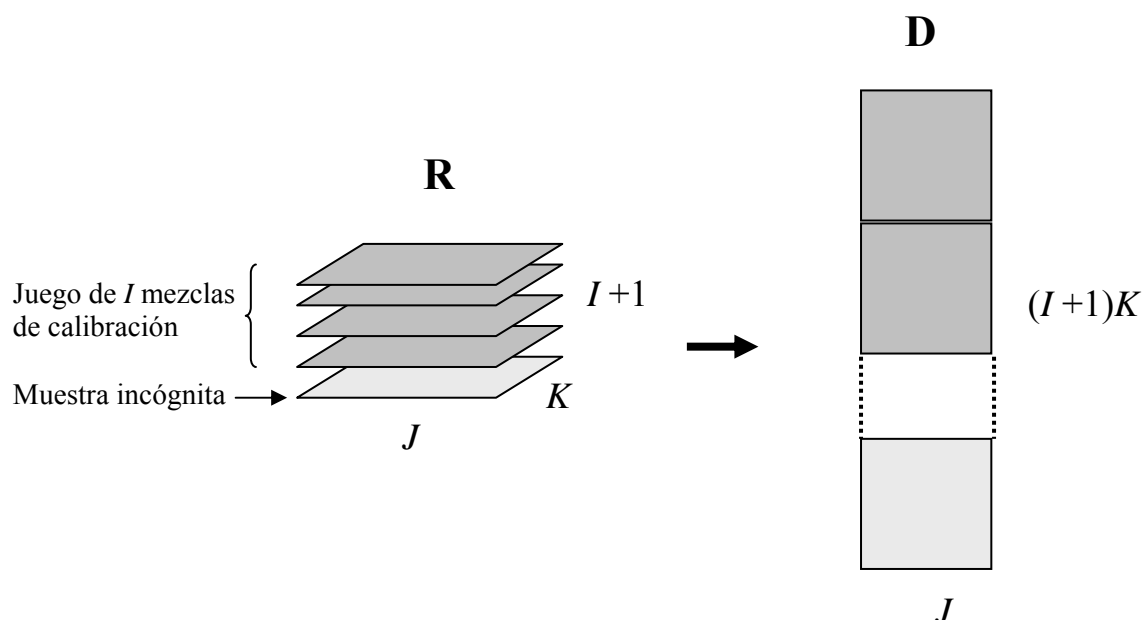


Figura 1.18: Representación del desdoblamiento de una hipermatriz **R** de tres vías  $[(I+1)JK]$ , en otra **D** de dos vías  $[(I+1)K \times J]$ . A la hipermatriz del juego de calibración, se le agregan los datos de la muestra incógnita, de esta manera se incorpora información de las interferencias no modeladas.

Después de desdoblar la hipermatriz  $\mathbf{R}$ , y obtener  $\mathbf{D}$ , se aplica el método de descomposición en valores singulares (SVD) o el método de análisis por componentes principales (PCA) para estimar el número de componentes de las mezclas.

Posteriormente se lleva a cabo la descomposición bilineal de la matriz  $\mathbf{D}$ , mediante un algoritmo iterativo cíclico (ALS). Este algoritmo utiliza el método de cuadrados mínimos, para obtener un mínimo global de  $\|\mathbf{E}\|$ .

La matriz  $\mathbf{D}$  puede expresarse según la siguiente ecuación:

$$\mathbf{D} = \mathbf{C} \mathbf{S}_{\text{MCR}}^T + \mathbf{E} \quad (1.112)$$

donde  $\mathbf{D}$  es la matriz de datos experimentales, las columnas de  $\mathbf{C}$  y de  $\mathbf{S}_{\text{MCR}}^T$  contienen los perfiles extraídos por el algoritmo ALS, y  $\mathbf{E}$  es la matriz de los residuos no ajustados por el modelo. Estas matrices tienen las siguientes dimensiones  $\mathbf{D}$   $[(I+1)K \times J]$ ,  $\mathbf{C}$   $[(I+1)K \times N]$ ,  $\mathbf{S}_{\text{MCR}}$   $(J \times N)$ , y  $\mathbf{E}$   $[(I+1)K \times J]$ , respectivamente. Al igual que en los modelos anteriores,  $I$  es el número de mezclas de calibración,  $N$  el número de componentes, y  $J$  y  $K$  los modos en que varía el sistema.

Esquemáticamente, podremos graficar la descomposición de la matriz de datos ( $\mathbf{D}$ ), en dos matrices que contienen los perfiles: espectrales y de concentración, (ver Figura 19).<sup>128,129</sup>

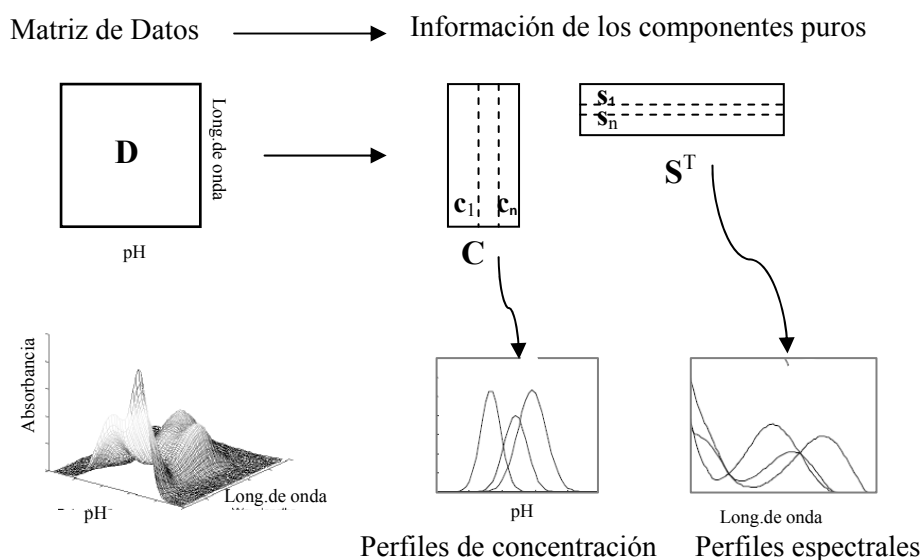


Figura 1.19: Esquema de la descomposición de una hipermatriz de datos ( $\mathbf{D}$ ), en dos matrices  $\mathbf{C}$  y  $\mathbf{S}$  que contienen los perfiles de concentración y espectrales de los componentes puros, respectivamente.

### Restricciones

Son necesarias ciertas restricciones para aplicar el modelo, entre ellas: a) la no-negatividad de los perfiles espectrales y de concentración. Como se observa en la Figura 1.20 deben corregirse los valores negativos, b) de uni-modalidad, como se observa en las Figuras 1.21 y 1.22 se suprimen los picos menores y c) de cierre, se corrige la sumatoria de las concentraciones de las especies para que esta resulte una constante, ver Figura 1.23.

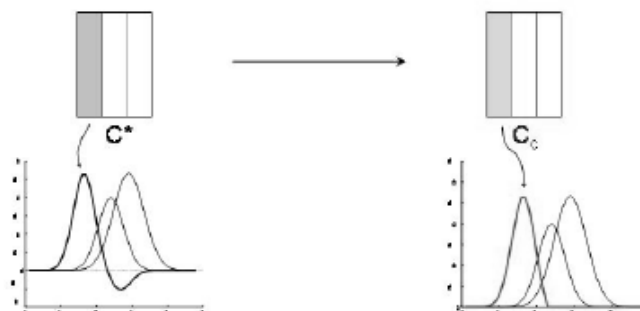


Figura 1.20: Restricción de no-negatividad en los perfiles de concentración.

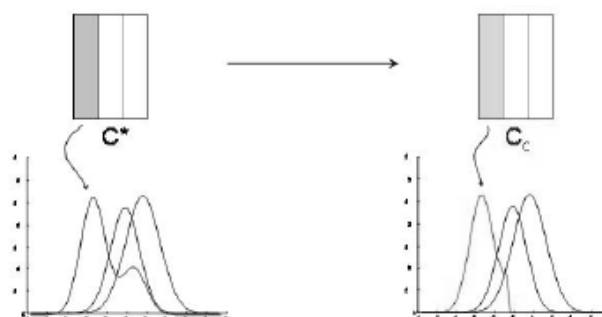


Figura 1.21: Restricción de uni-modalidad de los picos en los perfiles espectrales.

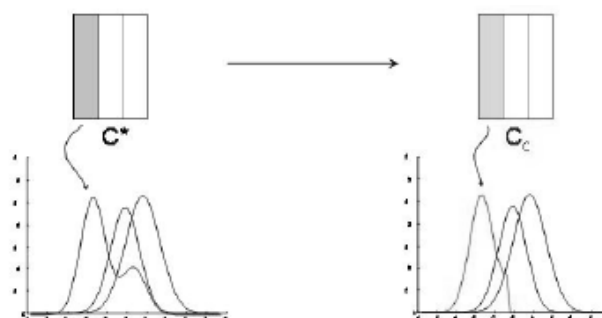


Figura 1.22: Restricción del rango local / selectividad. Se suprimen picos menores de alguno de los componentes, cuando coinciden con los picos de otro de los componentes, de esta manera se mejora la selectividad.

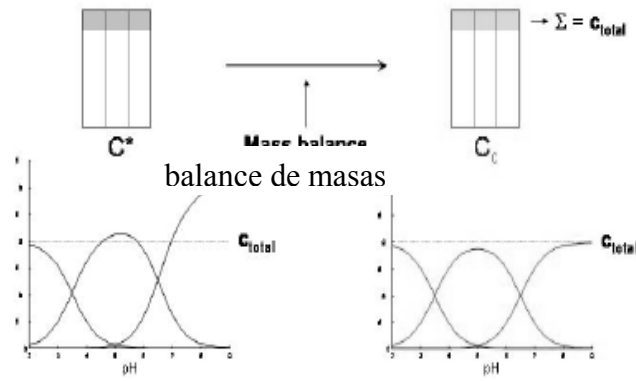


Figura 1.23: Restricción balance de masas o restricción de cierre.

**Algoritmo**

a) El primer paso del algoritmo consiste en estimar la matriz  $C$ , según:

$$\hat{C} = D S_{MCR}^{T+} \tag{1.113}$$

donde  $^+$  indica la pseudo-inversa de la matriz  $S$

b) Con la matriz estimada en la ecuación (1.113)  $\hat{C}$ , y la matriz original de datos  $D$ , se re-estima  $S_{MCR}$  mediante la siguiente ecuación:

$$\hat{S}_{MCR} = (\hat{C}^+ D)^T \tag{1.114}$$

c) Finalmente se calcula  $E$ , según la ecuación (1.112), utilizando la matriz de datos  $D$  y las matrices estimadas,  $\hat{G}$  y  $\hat{S}_{MCR}$ ,

a) Estos pasos se repiten hasta la convergencia. Se buscan dos mínimos de  $\|E\|$ , minimizando las dos ecuaciones (1.113) y (1.114), bajo las restricciones antes mencionadas.

$$\min_{\hat{C}} \|D - \hat{C} \hat{S}^T\| \tag{1.115}$$

$$\min_{\hat{S}} \|D - \hat{C} \hat{S}^T\| \tag{1.116}$$

De esta manera se obtienen los perfiles que mejor describen el sistema.

Toda la secuencia de cálculos anteriores podemos resumirlas en un diagrama de flujo presentad en la Figura 24:<sup>134</sup>

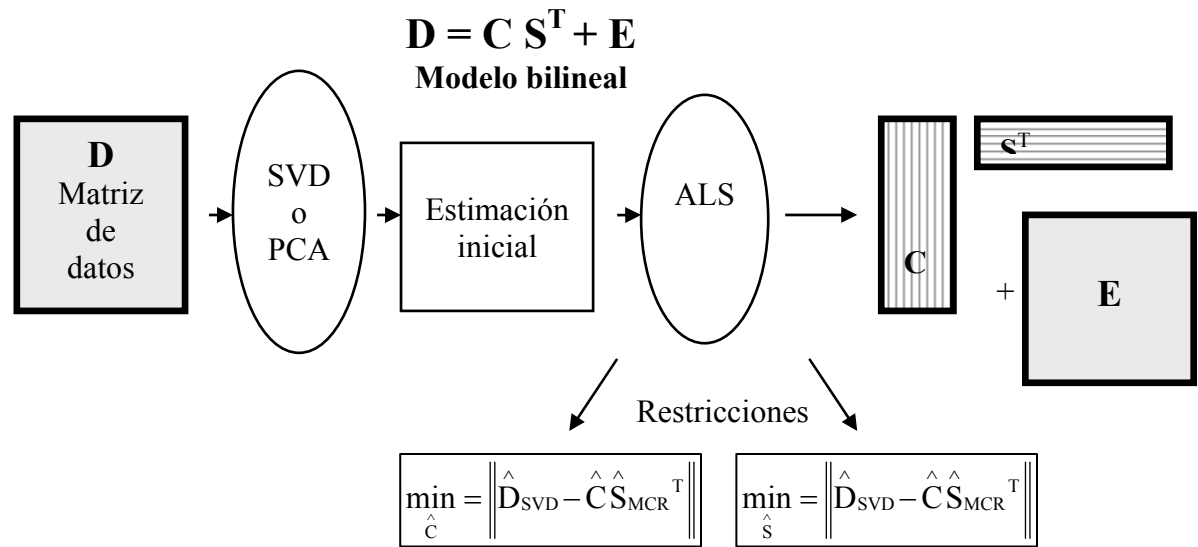


Figura 1.24: Diagrama de flujo que muestra los pasos seguidos en la descomposición bilineal de la matriz **D**.

### *Falta de unicidad de la solución*

Este método, a diferencia de PARAFAC, tiene libertad rotacional, lo que conlleva a que la solución no sea única. Analizando el algoritmo tenemos:

$$D = C S^T + E = D^* + E \quad (1.117)$$

$$S_{nueva}^T = T S^T \quad (1.118)$$

$$C_{nueva} = C T^{-1} \quad (1.119)$$

$$D^* = C S^T = C T^{-1} T S^T = C_{nueva} S_{nueva}^T \quad (1.120)$$

donde **T** es una matriz diagonal. Las dimensiones de **T** pueden variar, inicialmente se toma el número de filas y columnas igual al número de componentes de la mezcla. Además **T** depende de las restricciones que se apliquen al sistema. Esta es la razón por la que la solución no es única.

En la presente tesis, MCR- ALS fue aplicado a un sistema de dos componentes, cuyas especies ácida y básica se encuentran en equilibrio, por lo que dos componentes se transforman en cuatro especies debido a la variación de pH, generando una matriz de

datos con variaciones en pH y absorbancia. En nuestro caso las restricciones para el proceso ALS fueron la no-negatividad de los perfiles de absorción, y pH, y que se mantenga constante la relación de concentración entre los pares ácido-base de los componentes.

Los espectros puros de los componentes son los mismos en todos los experimentos, pero a diferencia de PARAFAC, los perfiles de pH en las diferentes submatrices de  $\mathbf{G}$  no necesariamente tienen la misma forma. Esta es la razón por la cual los experimentos llevados a cabo bajo diferentes condiciones (por ejemplo temperatura, ligeras diferencias de pH, etc.) pueden ser analizados mientras los espectros de los componentes permanezcan invariantes.

Al igual que en PARAFAC, la calibración se lleva a cabo aumentando la matriz de datos con la matriz de la muestra incógnita. Usualmente el número de muestras de calibración es limitado.<sup>62,135</sup>

### ***Predicción***

Después de la descomposición de  $\mathbf{D}$ , la información contenida en  $\mathbf{C}$  puede ser utilizada para realizar los cálculos de predicción, de manera similar a la descrita para PARAFAC. Primero se construye un gráfico de calibración pseudo-univariado con las concentraciones relativas de las muestras de calibración contenidas en la matriz  $\mathbf{C}$  estimada. Cuando se decide cual es el analito a cuantificar, 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 de pH para una muestra en particular. Si el área bajo el perfil es proporcional a la concentración del componente, se construye un modelo según:

$$\left\{ \sum_{k=1}^K C_{kn} \left| \sum_{k=1}^K C_{(k+K)n} \right| \cdots \left| \sum_{k=1}^K C_{[k+K(I-1)]n} \right| \right\} = k_{\text{MCR}} \mathbf{y} \quad (1.121)$$

donde la sumatoria de cada grupo de  $K$  elementos de  $\mathbf{C}$  corresponden a un componente en particular, en una muestra de calibración dada.

Mediante la interpolación de los valores provistos por la muestra desconocida puede obtenerse la concentración del analito de interés.

$$\mathbf{y}_u = \frac{\sum_{k=1}^K C_{(k+IK)n}}{k_{\text{MCR}}} \quad (1.122)$$

### 1.4.3 CUADRADOS MÍNIMOS BILINEALES (BLLS)

#### 1.4.3.1 INTRODUCCIÓN

El método de cuadrados mínimos bilineales, es un método de calibración multivariada utilizado para procesar arreglos de datos de tercer orden como los constituidos por matrices de excitación-emisión, aunque como se verá más adelante puede ser extendido al caso en que existen dependencias lineales como las originadas en los datos obtenidos por barridos espectrales en gradientes de pH. A diferencia de PARAFAC con ventaja de segundo orden, BLLS emplea la información de las concentraciones de los analitos en las mezclas de calibración durante la etapa de calibración, sin utilizar los datos de la muestra desconocida.

#### 1.4.3.2 DESARROLLO TEÓRICO MATEMÁTICO

Primeramente, se estiman las matrices a concentración unitaria ( $\mathbf{S}_n$ ) para los analitos calibrados. Para ello, las  $I$  matrices de calibración  $\mathbf{X}_{c,i}$  (cada una de dimensiones  $J \times K$ ) primero son vectorizadas y agrupadas dentro de una matriz  $\mathbf{V}_X$  (de dimensiones  $JK \times I$ ).<sup>58</sup>

$$\mathbf{V}_X = [\text{vec}(\mathbf{X}_{c,1}) \mid \text{vec}(\mathbf{X}_{c,2}) \mid \dots \mid \text{vec}(\mathbf{X}_{c,I})] \quad (1.123)$$

donde ‘vec’ implica la operación de desplegado de la matriz con el objeto de vectorizarla. Luego se emplea un ajuste por cuadrados mínimos para obtener información de los analitos puros, en un procedimiento análogo a CLS en calibración multivariada de primer orden:<sup>58</sup>

$$\mathbf{V}_S = \mathbf{V}_X \mathbf{Y}^{\text{T}+} \quad (1.124)$$

donde  $\mathbf{Y}$  es una matriz de  $(I \times N_c)$  que colecta las concentraciones nominales de todos los analitos calibrados,  $N_c$  es el número de analitos calibrados y  $\mathbf{V}_S$  ( $JK \times N_c$ ) contiene las matrices  $\mathbf{S}_n$  en forma vectorizada, ver Figura 25:

$$\mathbf{V}_S = [\text{vec}(\mathbf{S}_1) \mid \text{vec}(\mathbf{S}_2) \mid \dots \mid \text{vec}(\mathbf{S}_{N_c})] \quad (1.125)$$

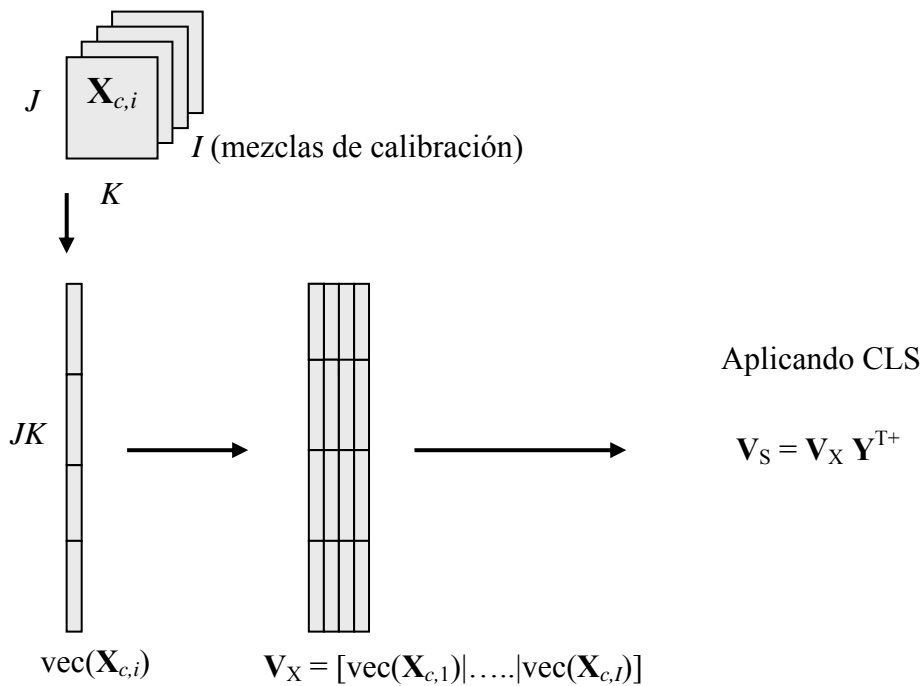


Figura 1.25: Esquema del desdoblamiento de las  $\mathbf{X}_{c,i}$  matrices de las  $I$  mezclas de calibración, para obtener los vectores  $\text{vec}(\mathbf{X}_{c,i})$ . En un siguiente paso estos vectores se ordenan para formar la matriz  $\mathbf{V}_X$ . La matriz  $\mathbf{V}_X$  se descompone en dos matrices mediante el mismo mecanismo que utiliza CLS.

En la literatura se proponen dos procedimientos para obtener los perfiles en ambas dimensiones, presentes en las matrices  $\mathbf{S}_n$  ( $J \times K$ ):<sup>57,60</sup> cuadrados mínimos (LS) y descomposición en valores singulares (SVD). Aparentemente SVD parece ser el más confiable y simple de los métodos. Con este método, los perfiles se obtienen por descomposición en valor singular por componente individualmente ( $\text{SVD}_1$ ) de cada una de las matrices  $\mathbf{S}_n$ , obtenidas después del correspondiente reformato de los vectores  $\text{vec}(\mathbf{S}_n)$ :

$$(g_n, \mathbf{b}_n, \mathbf{c}_n) = \text{SVD}_1(\mathbf{S}_n) \quad (1.126)$$

donde  $g_n$  es el primer valor singular, y  $\mathbf{b}_n$  y  $\mathbf{c}_n$  son los autovectores izquierdo y derecho respectivamente ( $J \times 1$  y  $K \times 1$ ) de  $\mathbf{S}_n$ . De esta manera se completa el proceso de calibración.

Posteriormente, las concentraciones en muestras desconocidas son estimadas a través de ajuste por cuadrados mínimos, suponiendo que no existen interferencias no modeladas:



$$\mathbf{y}_u = \mathbf{S}_{\text{cal}}^+ \text{vec}(\mathbf{X}_u) \quad (1.127)$$

donde  $\mathbf{X}_u$  es la matriz ( $J \times K$ ) de datos para la muestra incógnita,  $\mathbf{y}_u$  es un vector ( $N_c \times 1$ ) que contiene la concentración estimada de los  $N_c$  analitos en la muestra, y  $\mathbf{S}_{\text{cal}}$  es una matriz de calibración de ( $JK \times N_c$ ) dada por:

$$\mathbf{S}_{\text{cal}} = [g_1 (\mathbf{c}_1 \otimes \mathbf{b}_1) \mid g_2 (\mathbf{c}_2 \otimes \mathbf{b}_2) \mid \dots \mid g_{N_c} (\mathbf{c}_{N_c} \otimes \mathbf{b}_{N_c})] \quad (1.128)$$

donde  $\otimes$  implica el conocido producto de Kronecker (ver Figura 26).

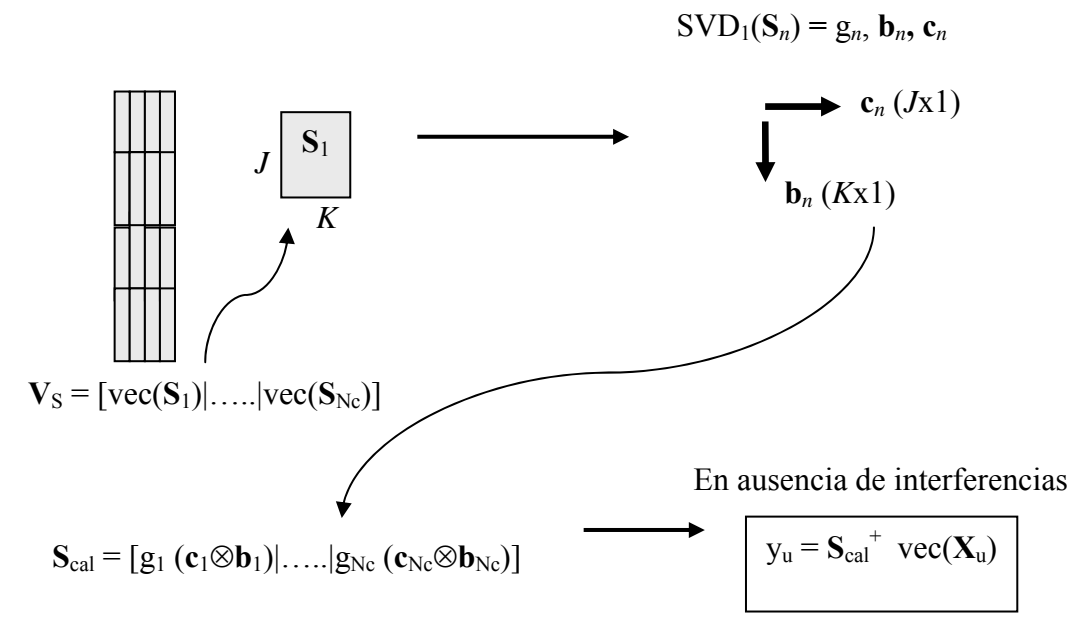


Figura 1.26: Reformateo de los vectores  $\text{vec}(\mathbf{S}_{N_c})$ , descomposición de la nueva matriz  $\mathbf{S}_{N_c}$  en valores singulares, y cálculo de la matriz para la predicción  $\mathbf{S}_{\text{cal}}$ .

Si existe un solo analito,  $\mathbf{S}_{\text{cal}}$  en ecuación (1.128) contiene una columna. Si embargo, si el analito tiene formas dependientes del pH, su correspondiente  $\mathbf{S}_n$  contendrá información de las especies en equilibrio, y por lo tanto será necesario considerar más de un componente para la descomposición singular de la ecuación (1.126). Esto originará, para un analito dado, diferentes valores de  $g_n$ , por ejemplo,  $g_{n1}$ ,  $g_{n2}$ , etc., donde el primer subíndice identifica al analito y el segundo a la especie en equilibrio. De la misma manera, se obtendrán diferentes perfiles para cada dimensión ( $\mathbf{b}_{n1}$ ,  $\mathbf{b}_{n2}$ ,  $\mathbf{c}_{n1}$ ,  $\mathbf{c}_{n2}$ , etc.), que consistirán en combinaciones lineales de los perfiles espectrales verdaderos para las especies en equilibrio del analito.<sup>137</sup>

Para un único analito ( $n = 1$ ), se emplea una ecuación similar a la ecuación (1.127) para estimar la concentración, pero  $\mathbf{S}_{\text{cal}}$  estará dada por:

$$\mathbf{S}_{\text{cal}} = [g_{11} (\mathbf{c}_{11} \otimes \mathbf{b}_{11}) \mid g_{12} (\mathbf{c}_{12} \otimes \mathbf{b}_{12}) \mid \dots \mid g_{1N_s} (\mathbf{c}_{1N_s} \otimes \mathbf{b}_{1N_s})] \quad (1.129)$$

donde  $N_s$  es el número total de especies en equilibrio para cada analito. Aún en el caso de un solo analito,  $\mathbf{y}_u$  es en este caso un vector de ( $N_s \times 1$ ) que contiene la concentración predicha del analito calibrado (hay  $N_s$  valores ya que cada uno se obtiene a partir de cada especie en equilibrio).

Para más analitos, cada uno de los cuales tiene especies en equilibrio, la ecuación general para  $\mathbf{S}_{\text{cal}}$  estará dada por:

$$\mathbf{S}_{\text{cal}} = [g_{11} (\mathbf{c}_{11} \otimes \mathbf{b}_{11}) \mid g_{12} (\mathbf{c}_{12} \otimes \mathbf{b}_{12}) \mid \dots \mid g_{1N_{s1}} (\mathbf{c}_{1N_{s1}} \otimes \mathbf{b}_{1N_{s1}}) \mid \dots \mid g_{21} (\mathbf{c}_{21} \otimes \mathbf{b}_{21}) \mid g_{22} (\mathbf{c}_{22} \otimes \mathbf{b}_{22}) \mid \dots \mid g_{2N_{s2}} (\mathbf{c}_{2N_{s2}} \otimes \mathbf{b}_{2N_{s2}}) \mid \dots] \quad (1.130)$$

donde  $N_{s1}$ ,  $N_{s2}$ , etc. es el número de especies asociadas con cada analito.

La presencia de componentes no calibrados en la muestra incógnita se investiga después de que el paso de calibración se completa. Si hay algún interferente, se realiza una bilinealización residual (RBL, del inglés *residual bilinearization*).<sup>50,58</sup> En este proceso se encuentran los perfiles del o los interferentes, y son incorporados dentro de una versión expandida de  $\mathbf{S}_{\text{cal}}$ :

$$\mathbf{S}_{\text{exp}} = [\mathbf{S}_{\text{cal}} \mid g_{\text{int}} (\mathbf{c}_{\text{int}} \otimes \mathbf{b}_{\text{int}})] \quad (1.131)$$

donde los perfiles  $\mathbf{b}_{\text{int}}$  y  $\mathbf{c}_{\text{int}}$  se obtienen por un procedimiento que involucra la minimización residual de una matriz  $\mathbf{E}_u$ , mientras se ajustan los datos de la muestra con la suma de las contribuciones de varios componentes:

$$\mathbf{X}_u = \sum_{n'=1}^{N_c} g_{n'} \mathbf{b}_{n'} (\mathbf{c}_{n'})^T y_{n',u} + g_{\text{int}} \mathbf{b}_{\text{int}} (\mathbf{c}_{\text{int}})^T + \mathbf{E}_u \quad (1.132)$$

Durante la minimización, los perfiles de la interferencia son estimados por SVD de una matriz residual  $\mathbf{E}_c$ , obtenida sustrayendo las contribuciones del analito a la señal total  $\mathbf{X}_u$ :

$$\mathbf{E}_c = \mathbf{X}_u - \sum_{n'=1}^{N_c} \mathbf{S}_{n'} y_{n',u} \quad (1.133)$$

$$(g_{\text{int}} \mathbf{b}_{\text{int}} \mathbf{c}_{\text{int}}) = \text{SVD}_1(\mathbf{E}_c) \quad (1.134)$$

La ecuación (1.132) es apropiada para analitos que tienen una sola especie. En el caso de que diferentes analitos tengan diferente número de especies, la expresión se modifica para contemplar esta posibilidad:

$$\mathbf{X}_u = \sum_{n'=1}^{N_c} \sum_{s'=1}^{N_s} g_{n's'} \mathbf{b}_{n's'} (\mathbf{c}_{n's'})^T y_{n's',u} + g_{\text{int}} \mathbf{b}_{\text{int}} (\mathbf{c}_{\text{int}})^T + \mathbf{E}_u \quad (1.135)$$

Se ha descrito un método iterativo para aplicar RBL,<sup>60</sup> aunque algunas veces suele presentar divergencia en lugar de converger en la concentración del analito. Otra posibilidad es usar el algoritmo de minimización de Gauss-Newton para minimizar  $\|\mathbf{E}_u\|$  en la ecuación (1.135). Esto se puede implementar bajo MATLAB, donde  $g_{\text{int}}$ ,  $\mathbf{b}_{\text{int}}$ , and  $\mathbf{c}_{\text{int}}$  se encuentran aplicando una descomposición en valores singulares a  $\mathbf{E}_c$ . El procedimiento proporciona el valor de  $y_{ns,u}$  que minimiza  $\mathbf{E}_u$  en las ecuaciones (1.132) y (1.135). Es decir que el procedimiento permite encontrar el valor de la concentración del analito calibrado al mismo tiempo que los perfiles de la/s interferencia/s que hacen que los residuos sean comparables al ruido instrumental.

La idea central de BLLS es la de un proceso de calibración-predicción de dos pasos, donde la predicción es guiada por una minimización por cuadrados mínimos. La ventaja de segundo orden se obtiene en una etapa siguiente en la que los residuos de la matriz problema se analizan para encontrar y estimar los perfiles de la o las interferencias. Estos sirven para expandir las cargas (*loadings*) y poder estimar correctamente las concentraciones de los analitos, aún en presencia de constituyentes no esperados. Debe notarse que no son necesarios restricciones ni inicialización para realizar el procedimiento de minimización.

## 1.5 OBJETIVOS

Al elaborar el plan de trabajo de la presente tesis se plantearon los siguientes objetivos:

- 1) Desarrollar nuevos métodos analíticos para el estudio de aditivos alimentarios y contaminantes en muestras ambientales por aplicación de distintas técnicas espectroscópicas y electroquímicas, acopladas a métodos quimiométricos para aumentar su selectividad
- 2) Estudio de las posibilidades de automatizar los métodos por aplicación de los principios de análisis por inyección en flujo (FIA).
- 3) Transferir los métodos desarrollados a laboratorios bromatológicos y de calidad ambiental.

## CAPÍTULO 2

### MATERIALES Y MÉTODOS

#### 2.1 MATERIALES

##### 2.1.1 SOLUCIONES PATRÓN DE CONSERVANTES

Se prepararon soluciones de sorbato de potasio (SOR) y benzoato de sodio (BEN) ambas de  $1000,0 \text{ mg L}^{-1}$ . Dichas soluciones patrón fueron preparadas pesando  $100,0 \text{ mg}$  de la droga correspondiente [(SOR) marca Milwaukee y (BEN) marca Sigma], disolviendo y enrasando en matraz aforado de  $100,00 \text{ mL}$  con agua bidestilada.

##### 2.1.2 SOLUCIONES PATRÓN DE COLORANTES

Se prepararon cuatro soluciones patrón de  $4,00 \text{ g L}^{-1}$  de los siguientes colorantes:

- 1) Tartrazina: [(trisodio 5-hidroxi-1-(4-sulfofenil)-4-(4-sulfofenilazo) pirazol-3-carboxilato)].
- 2) Amarillo Ocaso: (6-hidroxi-5-[(4-sulfofenil)azo]-2-naphtaleno sulfónico disodio).
- 3) Rojo 40 o Rojo Allura (6-hidroxi-5-[(2-metoxi-5-metil-4-sulfofenil)azo]-2-naftalensulfónico disodio).
- 4) Índigo Carmín (5,5'-(2-(1,3-dihidro-3-oxo-2H-indazol-2-ylideno)-1,2-dihidro-3H-indol-3-ona) disulfonato).

Dichas soluciones fueron preparadas pesando  $400,0 \text{ mg}$  de droga sólida marca Sigma (en todos los casos), disolviendo y llevando a un volumen final de  $100,00 \text{ mL}$  en matraz aforado con agua purificada ( $18 \text{ m}\Omega$ ) mediante el uso de un sistema B-pure.

### 2.1.3 MUESTRAS COMERCIALES

#### 2.1.3.1 BEBIDAS PREPARADAS A PARTIR DE JUGOS DE FRUTA

- 1) Jugo de naranja (concentrado, pasteurizado, con vitamina C), marca “MOCORETÁ” (producción regional). Conservantes declarados: benzoato de sodio y sorbato de potasio. Esta muestra fue designada como JN1.
- 2) Jugo de naranja (concentrado), marca “MIJÚ” (producción regional). Conservantes declarados sorbato de potasio y benzoato de sodio. Esta muestra fue designada como JN2.
- 3) Jugo de naranja (concentrado), marca “COTTI” (producción regional). Conservantes declarados sorbato de potasio y benzoato de sodio. Esta muestra fue designada como JN3.
- 4) Jugo de naranja con el agregado de vitamina C (listo para beber), marca “Hi-C”. Conservantes declarados sorbato de potasio y benzoato de sodio. Esta muestra fue designada como JN4.

#### 2.1.3.2 POLVOS PARA PREPARAR BEBIDAS ARTIFICIALES

- 1) Polvo para preparar bebida analcohólica, artificial, dietética. Fortificada con vitaminas A y C, sabor naranja, marca “Tang” (Kraft Foods Argentina S.A.). Colorantes declarados: tartrazina y amarillo ocaso. Esta muestra fue designada como J1.
- 2) Polvo para preparar bebida analcohólica, artificial, dietética. Fortificada con vitaminas A y C, sabor mandarina, marca “Tang” (Kraft Foods Argentina S.A.). Colorantes declarados: tartrazina, amarillo ocaso y rojo 40. Esta muestra fue designada como J2.
- 3) Polvo para preparar bebida analcohólica artificial, dietética, sabor manzana deliciosa, marca “Clight” (Kraft Foods Argentina S.A.). Colorantes declarados: tartrazina, amarillo ocaso, rojo 40 e índigo carmín. Esta muestra fue designada como J3.

#### 2.1.4 JUGOS PARA MODELAR LAS INTERFERENCIAS DE LA MATRIZ

Se prepararon jugos de naranja, pomelo y limón, a partir de frutas frescas exprimidas al 20 % en agua bidestilada (dilución que generalmente se aplica al jugo de fruta exprimida durante su industrialización).

#### 2.1.5 MUESTRAS ARTIFICIALES

Se prepararon cinco muestras artificiales a partir de jugos de frutas frescas exprimidas (20%) con el agregado de solución patrón de SOR y BEN en cantidad necesaria para obtener una concentración de  $0,50 \text{ g L}^{-1}$  (de cada uno), llevando a volumen final con  $\text{HCl } 5,0 \times 10^{-4} \text{ mol L}^{-1}$ . Las frutas utilizadas para preparar cada una de estas muestras artificiales fueron: manzana, tomate, pera, mandarina y zanahoria, y a dichas muestras se las denominó: M1, M2, M3, M4 y M5 respectivamente.

#### 2.1.6 REACTIVOS REGULADORES DEL pH

- 1) Solución de ácido clorhídrico ( $5,0 \times 10^{-4} \text{ mol L}^{-1}$ ). Esta solución fue preparada por dilución de ácido clorhídrico concentrado ( $13,8 \text{ mol L}^{-1}$ ), marca Mallinckrodt, con agua bidestilada.  $\text{pH} = 3,3$ .
- 2) Solución madre de NaOH  $1,00 \text{ mol L}^{-1}$ . Para la preparación de la solución madre se pesó la masa correspondiente de droga sólida (marca Merk, Darmstadt, Germany), llevando a volumen final con agua bidestilada.
- 3) Solución reguladora Britton-Robinson Ácida (BRH),  $\text{pH} = 2,0$ . Dicha solución contiene  $0,0400 \text{ mol L}^{-1}$  de ácido Fosfórico (marca Mallinckrodt),  $0,040 \text{ mol L}^{-1}$  de ácido Bórico (marca Anedra Buenos Aires, Argentina) y  $0,040 \text{ mol L}^{-1}$  de ácido Acético (marca Mallinckrodt). Todos los reactivos son drogas de grado analítico disueltos en agua purificada ( $18 \text{ m}\Omega$ ) mediante el uso de un sistema B-pure.
- 4) Solución Britton-Robinson Alcalina (BROH),  $\text{pH} = 12,0$ . Solución de NaOH  $0,20 \text{ mol L}^{-1}$ , preparada por dilución de la solución madre de NaOH ( $1,00 \text{ mol L}^{-1}$ ) con agua bidestilada.
- 5) Solución reguladora de pH 5,5. Preparada mezclada mezclando  $96,10 \text{ mL}$  de fosfato di-ácido de potasio (marca Mallinckrodt) ( $9,073 \text{ g L}^{-1}$ ), y  $3,90 \text{ mL}$  de fosfato mono-ácido de sodio (marca Mallinckrodt) ( $9,464 \text{ g L}^{-1}$ ).

- 6) Solución de NaOH 0,10 mol L<sup>-1</sup>, preparada por dilución de la solución madre de NaOH (1,00 mol L<sup>-1</sup>) con agua bidestilada.

### 2.1.7 REACTIVOS METODO OFICIAL PARA CONSERVANTES

- 1) Ácido tartárico, droga sólida Merk Darmstadt, Germany.
- 2) Sulfato de magnesio heptahidratado, marca Merk Darmstadt, Germany.
- 3) Solución de cobre alcalina. Preparada mezclando 1,0 mL de sulfato de cobre heptahidratado (0,1 %) (marca Merk, Darmstadt, Germany), 0,50 g de carbonato de sodio (marca Merk, Darmstadt, Germany), y agua bidestilada c.s.p. 1000 mL.
- 4) Ácido clorhídrico 0,1 mol L<sup>-1</sup>. Preparada por dilución del ácido comercial.

### 2.1.8 REACTIVOS PARA OPTIMIZAR EL SISTEMA FIA

- 1) Verde de bromocresol (BCG), marca Sigma.
- 2) Solución de tartrazina de 50,00 mg L<sup>-1</sup>, preparada por dilución de la solución patrón de 4,00 g L<sup>-1</sup>

### 2.1.9 EQUIPOS

- 1) Espectrofotómetro Perkin-Elmer Lambda 20, con celda de cuarzo de 1 cm de camino óptico. La velocidad de barrido en todos los casos fue 1200 nm min<sup>-1</sup>.
- 2) Espectrofotómetro Hewlett-Packard 8452A con detector de arreglo de diodos y celda de flujo Yema 178-010-QS con un volumen de 18 µL.
- 3) Bomba peristáltica Gilson Minipuls 3. Se empleó con una velocidad de 5 rpm.
- 4) Válvula de inyección Rheodyne 5041.
- 5) Tubos de diámetro interno 0,5 mm de PTFE para el reactor (*coil*), *loop* y líneas de transmisión del sistema FIA.



## 2.2 MÉTODOS

### 2.2.1 MÉTODO OFICIAL DE LA AOAC PARA LA CUANTIFICACIÓN DE SORB Y BEN EN BEBIDAS ANALCOHÓLICAS

Se miden exactamente 20 mL de muestra y se colocan en un valón, se agregan 2,0 g de ácido tartárico y 50,0 g de sulfato de magnesio heptahidratado, posteriormente se realiza una rápida destilación por arrastre con vapor hasta recoger aproximadamente 350-400 mL. La solución obtenida se lleva a 500,00 mL con agua bidestilada en matraz aforado y luego se filtra. A 5,0 mL del filtrado se la agregan 25,0 mL de solución alcalina de sulfato de cobre, se agita enérgicamente con varilla de vidrio 2 o 3 minutos. Finalmente se acidifica agregando 5,0 mL de HCl 0,1 M.

Se miden las absorbancias a 230 y 263 nm, frente a un blanco preparado con 5,0 mL de agua bidestilada y los reactivos empleados.

Los cálculos a realizar se muestran en las ecuaciones (2.1) y (2.2).

$$\text{- Ácido sórbico (g L}^{-1}\text{)} = \frac{(A_2 - 0,087 \times A_1) \times 500}{32 \times V} \quad (2.1)$$

$$\text{- Ácido benzoico (g L}^{-1}\text{)} = \frac{(A_1 - 0,0060 \times C_2) \times 500}{13,4 \times V} \quad (2.2)$$

Donde:

$A_1$  = Abs. A 230 nm –  $A_{f1}$

$A_{f1}$  = Abs. de fondo a 230 nm

$A_2$  = Abs. A 263 nm –  $A_{f2}$

$A_{f2}$  = Abs. de fondo a 263 nm

$C_2$  = concentración de ácido sórbico en  $\text{g L}^{-1}$

$V$  = volumen de muestra en mL

## 2.2.2 MÉTODOS COMPUTACIONALES

Los espectros de absorción electrónica obtenidos con los equipos antes citados fueron convertidos a formato ASCII para su posterior manipulación. Para el procesamiento de los datos en los trabajos realizados se utilizaron los siguientes programas: Multivar, MVC1 y MVC2.

### 2.2.2.1 PROGRAMA MULTIVAR

Para el tratamiento de los datos en la cuantificación de conservantes se utilizó el programa Multivar, escrito en Visual Basic 5.0, e implementado en una microcomputadora tipo Pentium 166.<sup>91,138</sup>

En la Figura 2.1 se observa la pantalla del programa. Esta está dividida en cuatro regiones horizontales, cuyas funciones se describen a continuación:

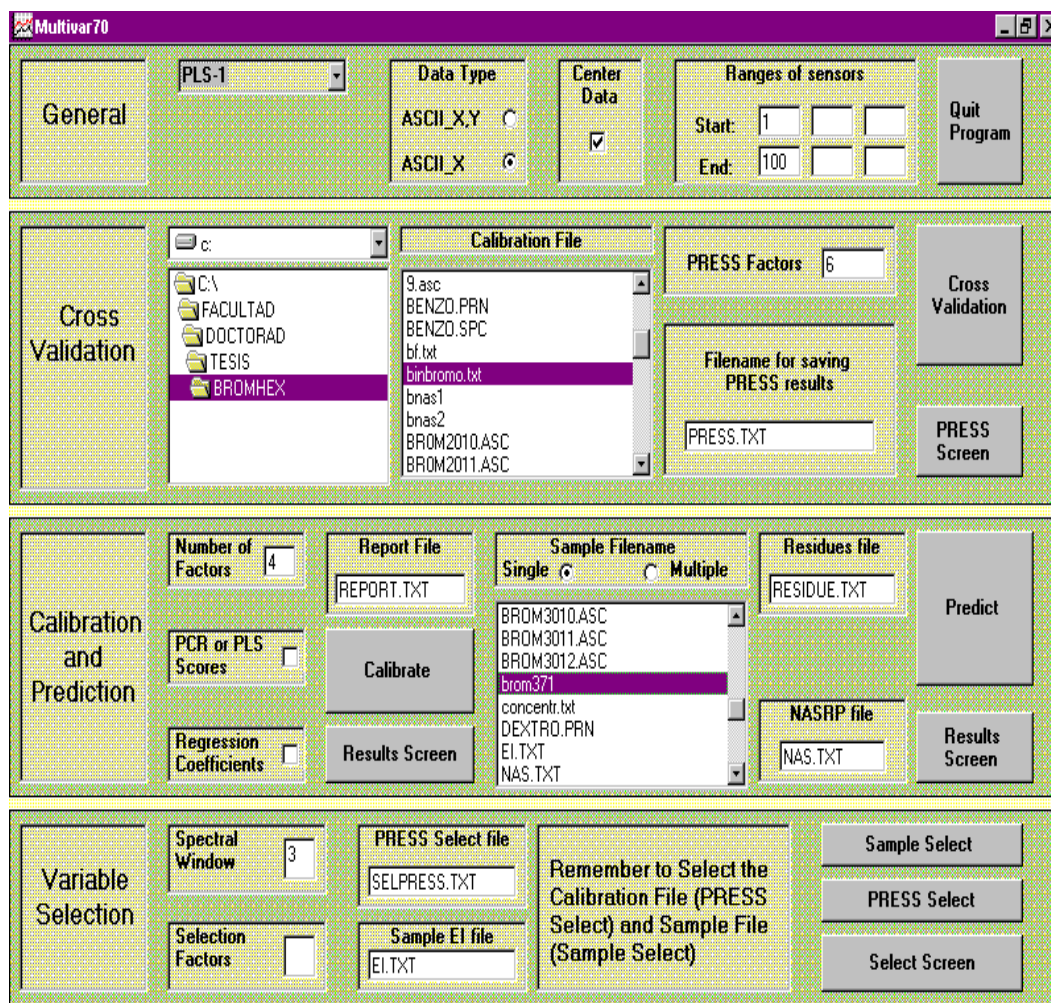


Figura 2.1 Pantalla del programa.

**Funciones:****1. General:** funciones generales y procesamiento de datos.

- 1) Selección del método: permite seleccionar el método multivariado de trabajo (PCR, PLS-1, HLA, etc.).
- 2) Tipo de datos utilizados (*Data Type*): los datos pueden ser tipo ASCII\_X,Y, correspondientes a archivos en formato ASCII conteniendo dos columnas, una con las longitudes de onda y otra con los valores de respuesta (absorbancias), y tipo ASCII\_X, correspondientes a una columna con las respuestas.
- 3) Centrado de los datos (*Center Data*): se activa por defecto y permite realizar el centrado de los datos.
- 4) Rango de los sensores (*Range of sensors*): se pueden definir hasta tres diferentes rangos de los sensores. Estos deben ser especificados en términos del número de entrada en el archivo formato ASCII correspondiente y no del valor específico del sensor. Por ejemplo, cuando el rango es 210-300 nm, corresponde asignar 1-91.
- 5) Salir (*Quit Program*): permite terminar con la aplicación del programa y volver a *Windows*.

**2. Validación cruzada (*Cross Validation*)**

- 1) Archivo de calibración (*Calibration File*): permite seleccionar el disco, directorio y archivo de trabajo. Este último archivo debe ser de formato ASCII y contener la información relativa a las muestras de calibración, concentraciones y nombres de los archivos con los espectros correspondientes. Una estructura típica de un archivo de este tipo se muestra a continuación:

*nombre del componente*

*benz*

*cantidad de mezclas*

*15*

*concentraciones*

*0*

*0*

*1000*

*1000*

*400*

*400*

*400*

*400*

*0*

**1000**  
**0**  
**1000**  
**400**  
**400**  
**400**  
**puntos**  
**91**  
**archivos con los espectros**  
**1.TAB**  
**2.TAB**  
**3.TAB**  
**4.TAB**  
**5.TAB**  
**6.TAB**  
**7.TAB**  
**8.TAB**  
**9.TAB**  
**10.TAB**  
**11.TAB**  
**12.TAB**  
**13.TAB**  
**14.TAB**  
**15.TAB**

- 2) Cálculo del *PRESS* en función del número de factores (*PRESS Factors*): permite asignar el número máximo de factores utilizados para realizar la validación cruzada.
- 3) Nombre del archivo para grabar los resultados del cálculo del *PRESS* (*Filename for saving PRESS results*): se debe colocar el nombre del archivo donde se guardarán los resultados de la validación cruzada. El nombre por defecto es *PRESS.TXT*, pero puede ser cambiado.
- 4) Validación cruzada (*Cross Validation*): botón que permite iniciar los cálculos de validación cruzada. Los resultados aparecen en una nueva ventana, desde la que se puede regresar a la ventana principal con el botón menú principal (*main menu*).
- 5) Ventana *PRESS* (*PRESS Screen*): botón que permite ver la ventana con los resultados de validación cruzada.

### **3. Calibración y predicción (*Calibration and Prediction*)**

- 1) Número de factores (*Number of Factors*): se coloca el número óptimo de factores obtenido por validación cruzada.

- 2) *Scores* de PLS o PCR (*PCR or PLS Scores*): esta opción permite guardar los scores en un archivo (en PCR son llamados T1.TXT, T2.TXT, etc., mientras que en PLS son llamados TA1.TXT, TA2.TXT, etc.). Estos nombres no pueden ser cambiados. En los métodos basados en cálculos de NAS no hay scores.
- 3) Coeficientes de regresión (*Regression Coefficients*): permite guardar los coeficientes de regresión (archivo BF.TXT, cuyo nombre puede ser cambiado).
- 4) Archivo de informe (*Report File*): nombre del archivo para guardar el informe de la predicción (por defecto es REPORT.TXT y puede ser cambiado).
- 5) Calibración (*Calibrate*): permite realizar la calibración. En este caso, se crean archivos llamados R1, R2, etc. La información guardada depende del método de trabajo seleccionado: en PCR contienen factores espectrales de la matriz de datos, en PLS contienen residuos en señales y concentraciones, *weight loading factors* y *loading factors*, y en métodos basados en cálculos de NAS, contienen factores espectrales de la matriz  $\mathbf{R}_k$ . También se crean archivos llamados BNAS1, BNAS2, etc., que contienen factores espectrales de la matriz  $\mathbf{R}_k$  (para PCR y PLS), cuyos nombres no pueden ser cambiados. Los resultados concernientes a sensibilidad y selectividad aparecen en una nueva ventana, desde la que se puede volver accionando *main menu*.
- 6) Pantalla con los resultados (*Results Screen*): permite volver a ver la pantalla con los resultados de calibración.
- 7) Nombre del archivo de la muestra (*Sample Filename*): se tienen dos opciones. Simple (*Single*), que permite seleccionar el archivo conteniendo el espectro de una muestra desconocida y Múltiple (*Multiple*), que permite la selección de un archivo de texto conteniendo una lista de archivos correspondientes a espectros de muestras incógnitas.
- 8) Archivo con los residuos (*Residues File*): nombre del archivo donde se guardan los residuos espectrales. El nombre por defecto es RESIDUES.TXT puede ser cambiado.
- 9) Archivo NASRP (*NASRP file*): nombre del archivo donde se guardan los datos para la construcción de las NASRP (que contiene dos columnas, la primera es la NAS del analito puro y la segunda es la NAS de la muestra). El nombre por defecto es NAS.TXT y puede ser cambiado.
- 10) Predicción (*Predict*): botón para realizar la predicción. Los resultados aparecen en una nueva ventana desde la que se puede regresar accionando *main menu*.

- 11) Pantalla de resultados (*Results Screen*): permite volver a observar la ventana con los resultados de la predicción.

#### 4. Selección de variables (*Variable Selection*)

- 1) Ventana espectral (*Spectral Window*): se especifica el mínimo tamaño de la ventana utilizada para calcular el error indicador (EI).
- 2) Selección de factores (*Selection Factors*): se especifica el número máximo de factores para calcular EI.
- 3) Archivo para selección de *PRESS* (*PRESS Select file*): nombre del archivo para guardar los resultados de la selección del *PRESS*. El nombre por defecto es SELPRESS.TXT y puede ser cambiado.
- 4) Archivo para el error indicador (EI) de la muestra (*Sample EI file*): nombre del archivo para guardar los valores de EI para una muestra dada en función del primer sensor y la ventana. El archivo EI.TXT contiene cinco columnas: primer sensor, ventana espectral, EI, número óptimo de factores y concentración predicha. El nombre puede ser cambiado.
- 5) Selección de la muestra (*Sample Select*): botón para comenzar la selección para una muestra dada. Esta operación puede ser lenta, dependiendo del ancho espectral elegido. Los resultados correspondientes al EI mínimo aparecen en una nueva pantalla, desde la que se puede regresar a la pantalla principal apretando *main menu*.
- 6) Selección del *PRESS* (*PRESS Select*): comienza la selección basada en la búsqueda del *PRESS* mínimo. Esta operación puede ser lenta, dependiendo del ancho espectral elegido. Los resultados correspondientes al *PRESS* mínimo aparecen en una nueva pantalla, desde la que se puede regresar a la pantalla principal apretando *main menu*.
- 7) Pantalla de selección (*Select Screen*): permite volver a ver la pantalla con los resultados de la selección.

### 2.2.2.2 PROGRAMAS MVC1 Y MVC2

#### PROGRAMA MVC1

Para el tratamiento de los datos en la cuantificación de conservantes, también se utilizó el programa MVC1<sup>99</sup> que se corre en entorno de MATLAB.

En la Figura 2.2 se observa la pantalla del programa. Esta está dividida en dos regiones verticales, cuyas funciones se describen a continuación:

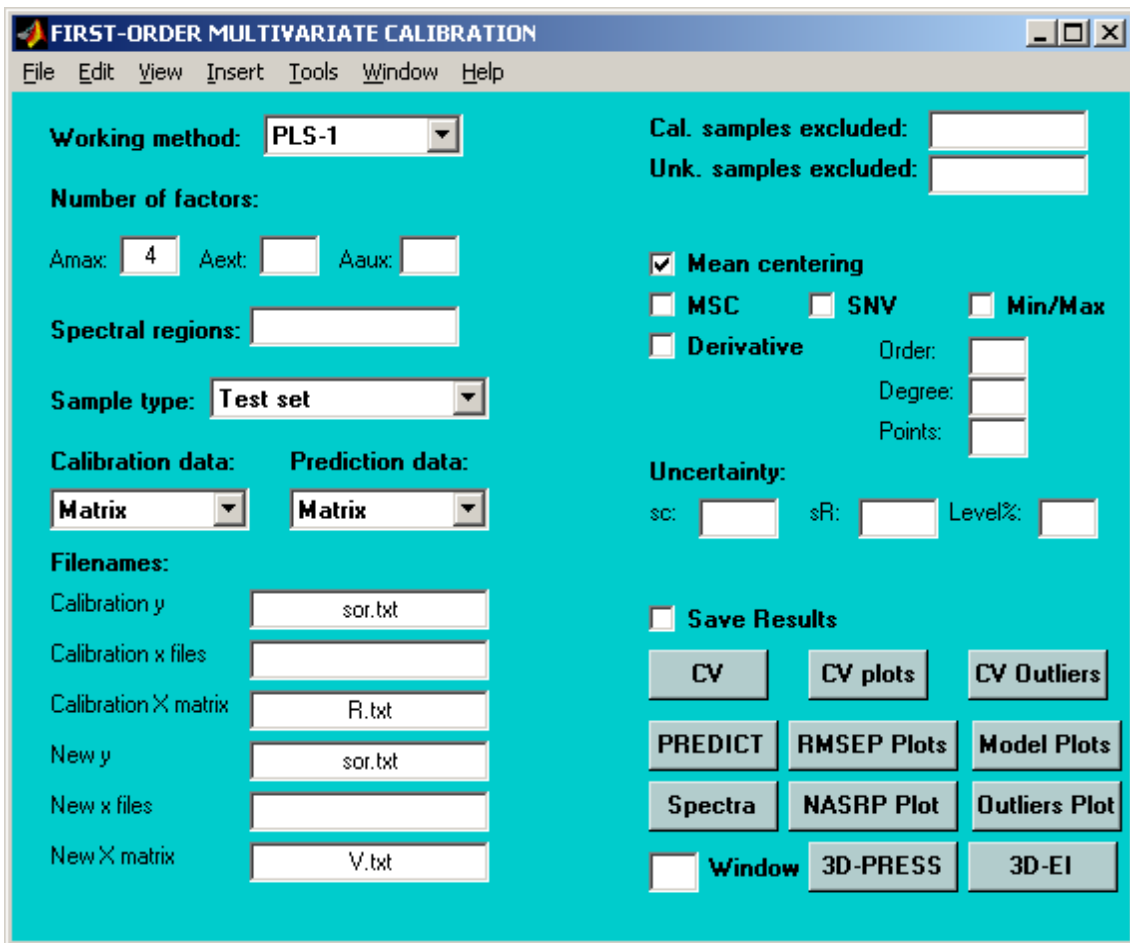


Figura 2.2: Pantalla del programa MVC1.

#### Funciones:

**1. Columna izquierda:** ingreso de datos y selección de las condiciones de trabajo.

- 1) Selección del método (*working method*): permite seleccionar el método multivariado de trabajo (PCR, PLS-1, HLA/GO, HLA/XS, etc.).
- 2) Número de factores (*number of factors*): consta de tres ventanas. En la primera ventana, según el criterio de Haaland y Thomas, se ingresa la mitad del número de

mezclas de calibración más uno, y se realiza la validación cruzada. Una vez obtenido el número óptimo de factores, se ingresa este valor, para realizar la predicción. La segunda y tercer ventana se utilizan en métodos que requieren más de un factor, como los basados en corrección ortogonal de la señal (OSC).

- 3) Región espectral (*spectral regions*): permite seleccionar la región de longitudes de onda de trabajo. En caso de dejar el espacio en blanco el programa realiza los cálculos con todos los datos del espectro.
- 4) Tipo de muestra (*sample type*): presenta dos opciones de trabajo: a) juego de validación (*test set*) o b) (*unknowns*) para muestras desconocidas.
- 5) Datos de calibración (*calibration data*) y datos de predicción (*prediction data*): estas dos ventanas presentan tres opciones: a) *x\_vectors*, b) *x*, *y\_vectors*, o c) *matrix*. Se selecciona una, de acuerdo al tipo de arreglo que tengan los datos.
- 6) Nombre de los archivos de datos (*filenames*): consta de una columna de siete ventanas donde se ingresan los nombres de los archivos de datos.

Calibration y: vector de datos (concentraciones del juego de calibración)

Calibration x vectors: vector de datos (señales de calibración)

Calibration X matrix: matriz de datos (señales del juego de calibración)

Unknown y: vector de datos (concentraciones de la muestra)

Unknown x vectors: vector de datos (señales de la muestra)

Unknown X matrix: matriz de datos (señales de las muestras)

Single unknown x: vector de datos (señal de una muestra)

## 2. Columna derecha: funciones que realiza el programa.

- 1) Exclusión de muestras de calibración (*Cal. samples excluded*): permite excluir muestras anómalas (*outliers*) del juego de calibración, y trabajar con las restantes. En la ventana se coloca el número correspondiente de la muestra a excluir.
- 2) Muestras de validación excluidas (*Unk. samples excluded*): permite excluir muestras anómalas (*outliers*) del juego de validación, y trabajar con las restantes. En la ventana se coloca el número correspondiente de la muestra a excluir.
- 3) Métodos de preprocesamiento de datos consta de cinco ventanas. Proponen varias opciones, en algunos casos permiten eliminar el ruido o en otros darles el mismo peso a todas las variables. Estas ventanas son:
  - a) Centrado de los datos (*Mean centering*): se activa por defecto.
  - b) *Multiplicative scatter correction (MSC)*.



- c) *Different variations of estándar normal variate (SNV)*.
- d) Escalado (*Min/Max*).
- e) Derivada de la señal (*derivate*): para esta opción existen tres ventanas, donde se puede ingresar el orden de la derivada (*order*), el grado del polinomio de Savitzky-Golay (*degree*), y el número de puntos a usar en el suavizado (*points*).
- 4) Errores (*uncertainly*): consta de tres ventanas, *sc*, *sR* y *level %*, correspondientes a la estimación de errores en concentración y en señal. Estos datos son utilizados para el calculo de cifras de merito.
- 5) Guardar resultados (*save results*): consta de una ventana que permite la opción de archivar los resultados. Se crean archivos en el directorio en el que se está trabajando.
- 6) Validación cruzada (Tecla *CV*): al apretar esta tecla el programa realiza los cálculos de la validación cruzada. Los resultados se informan en la pantalla del programa MATLAB.
- 7) Gráficos de la validación cruzada (Tecla *CV plots*): mediante esta tecla se pueden visualizar los gráficos de la validación cruzada en una pantalla que se abre automáticamente: 1) *PRESS vs. A*, 2) *log(PRESS) vs. A*, 3) *SEP vs. A*, 4) *log(SEP) vs. A*.
- 8) Muestras anómalas del juego de calibración (Tecla *CV outliers*): al apretar esta tecla en una nueva pantalla el programa pide que se ingrese el número de factores de calibración (*set number factor*). Una vez ingresado este número se presiona *GO* y el programa grafica (abriendo otra pantalla):  $N^0$  de mezcla vs.  $F_{pract}/F_{crit}$  (*sample vs. Fpract/ Fcrit*).
- 9) Predicción (Tecla *PREDICT*): al presionar esta tecla el programa realiza la predicción de las muestras desconocidas, el resultado se informa en la pantalla del programa MATLAB.
- 10) Estadísticos de la predicción (Tecla *RMSEP Plots*): al presionar esta tecla se abre una pantalla en la que el programa grafica algunos estadísticos de la predicción: 1) *Error vs. Número de factores (RMS vs. A)*, 2) *Resultados predichos vs. Concentraciones nominales (Predicted vs. Actual)*, 3) *Residuos vs. muestra (Residuals: Residuals vs. Simple)*, 4) *Intervalo elíptico conjunto para la pendiente y la ordenada al origen de la regresión del punto 2 [EJCR (Elliptical joint confidence regions): Intecept vs. Slope]*.

- 11) Gráficos del modelo (Tecla *Model plots*): al presionar esta tecla en una nueva pantalla el programa realiza los siguientes gráficos: 1) Sensibilidad vs. Número de factores (*Sensitivity vs. A*), 2) Coeficientes de regresión (*Regresión coeficientes:  $\mathbf{b}_k$  vs. Sensor*), 3) Selectividad vs. Número de factores (*Selectivity vs. A*), y 4) Señal neta (*Net sensitivity:  $\mathbf{s}_k$  vs. Sensor*).
- 12) Gráficos de los espectros (Tecla *Spectral plots*): apretando esta tecla se obtienen, en una nueva pantalla, los siguientes gráficos: 1) grafico de los espectros de calibración vs. Sensores (*Raw calibration*), 2) grafico de los espectros de las muestras desconocidas vs. Sensores (*Raw unknowns*) 3) y 4) espectros de calibración y de predicción preprocesados según opciones ya descritas vs. sensores (*Preprocessed calibration y Preprocessed unknowns*).
- 13) Gráficos de regresión de la NAS (Tecla *NASRP Plots*): al presionar esta tecla el programa abre una pantalla en la que pide que se ingrese el número de la muestra con la que se desea realizar los gráficos. A continuación se aprieta la tecla GO, y el programa abre una pantalla con los siguientes gráficos: 1)  $\mathbf{r}_e^*$  y  $\mathbf{s}_e^*$  vs. Sensores ( $\mathbf{r}_k^*$  and  $\mathbf{s}_k^*$  vs. *Sensor*), 2) NASRP ( $\mathbf{r}_k^*$  vs.  $\mathbf{s}_k^*$ ), 3) y 4) Residuos vs. sensores (*Net residual vs. Sensor y Residual vs sensor*).
- 14) Muestras anómalas de la predicción (Tecla *Outliers Plots*): al presionar esta tecla el programa abre una pantalla con el grafico: muestra vs. Ffrac./Fcrit. (*sample number vs. Fpract/Fcrit*)
- 15) Gráficos del PRESS y del Error indicador en función del primer sensor y de la ventana movil (Teclas *3D-PRESS* y *3D-EI*). Junto a estas teclas aparece una ventana donde debe ingresarse el numero de sensores con el que se realiza el cálculo. El programa informa los resultados numéricos en la pantalla MATLAB, y además abre una ventana en cada caso, con un grafico 3D y un diagrama de contorno.

### *PROGRAMA MVC2*

Es una rutina que se corre bajo el entorno MATLAB, similar a la descripta anteriormente. Se aplica en cálculos con arreglos de tercer orden, (métodos: PARAFAC, BLLS, etc.).

Utiliza las rutinas escritas por Bro para hacer PARAFAC y N-PLS<sup>139</sup> y la rutina escrita por Olivieri para hacer BLLS.<sup>137</sup>

El manejo de MVC2 es semejante a MVC1, ya que tienen la misma estructura.

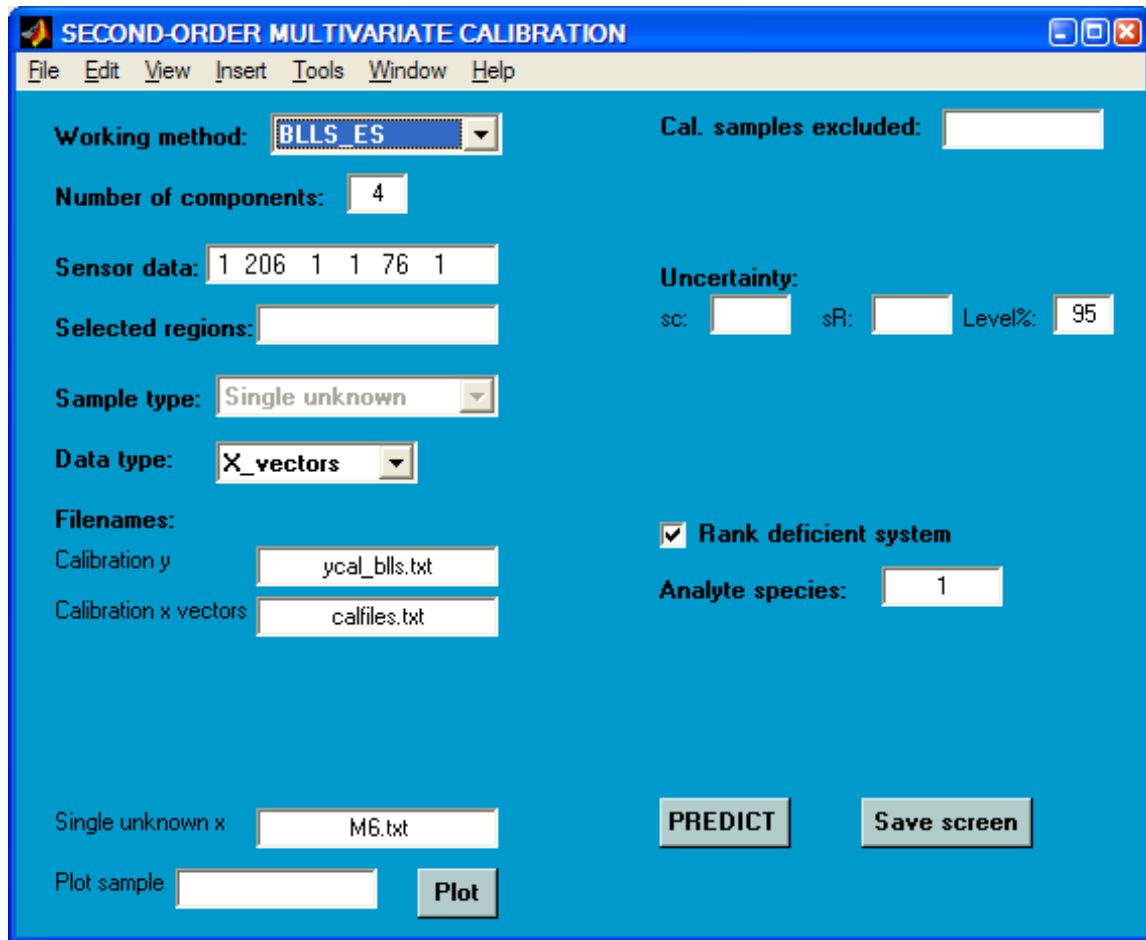


Figura 2.3: Pantalla del programa MVC2.

### 2.2.3 METODOLOGÍA DEL ANÁLISIS EN FLUJO. SISTEMA FIA (*flow injection analysis*)

#### 2.2.3.1 SISTEMA FIA PARA LA CUANTIFICACIÓN DE CONSERVANTES

Para la cuantificación de conservantes (SOR y BEN) en jugos de fruta comerciales, se construyó un sistema FIA simple para generar un doble gradiente de pH en línea como el que se expone en la Figura 2.4.

En primer lugar, se prepararon las soluciones a analizar (mezclas de calibración y muestras comerciales y artificiales), llevándolas a pH = 2,0 con solución Britton-Robinson ácida (*carrier*). Las soluciones fueron introducidas en el sistema FIA

consecutivamente mediante la bomba peristáltica, con una velocidad de 5 rpm y un caudal de  $0,56 \text{ mL min}^{-1}$ , requiriendo un tiempo no mayor de 3 minutos para el análisis completo de cada una de ellas.

Al introducir la solución se registró la señal de la línea de base, con un espectrofotómetro Hewlett-Packard 8452A con detector de arreglo de diodos y una celda de flujo.

Una vez que la mezcla está en línea, mediante la válvula se inyectan  $50 \mu\text{L}$  de solución Britton-Robinson alcalina. La mezcla con el bolo alcalino pasan por el reactor (*coil*), de 1250 mm de largo y 0,5 mm de diámetro interno. Por diferentes fenómenos físicos (flujo laminar, dispersión, etc.) las dos soluciones (ácida y alcalina) se mezclan en la parte anterior y posterior del bolo generando el doble gradiente de pH, necesario para producir el pico FIA en la señal (ver Figura 2.5).

Las señales fueron registradas, en un rango de longitudes de onda de 200 a 300 nm, cada 2 nm y 1,5 segundos. Por cada mezcla de calibración o muestra se obtiene una matriz de  $51 \times 33$ , lo que equivale a 1683 datos de absorbancia.

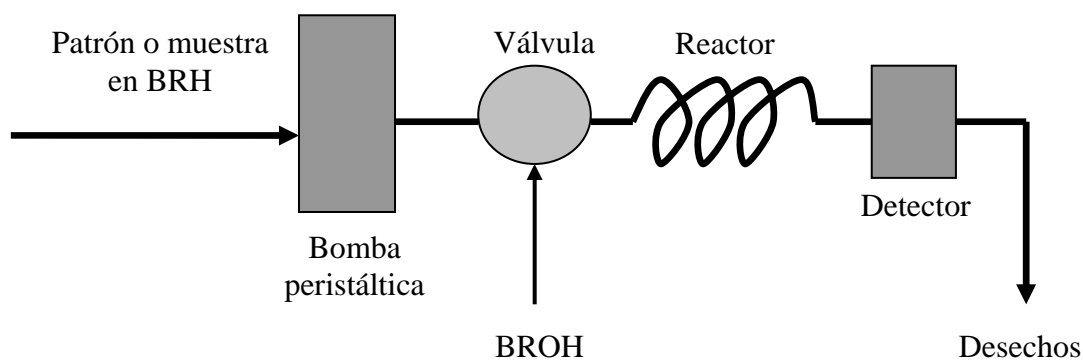


Figura 2.4: Sistema FIA para la cuantificación de conservantes.

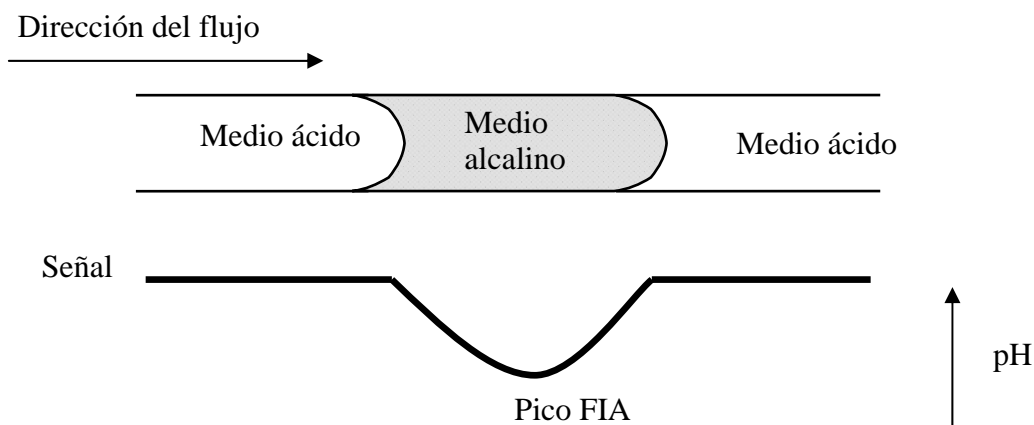


Figura 2.5: Esquema del perfil del gradiente de pH. Si se registra la señal a 270 nm, se tiene una línea de base mientras circulan por el sistema las muestras o patrones en medio ácido. Cuando se inyecta el bolo alcalino la señal se modifica, presentando un pico. Posteriormente la señal vuelve a ser la línea de base.

### ***Optimización del sistema FIA***

Para verificar que efectivamente se generaba del doble gradiente de pH, se procedió de la siguiente manera: 1) se preparó una solución indicadora de verde de bromocresol ( $pK_a = 4,5$ ) en buffer BRH ( $pH = 2,0$ ), 2) se introdujo dicha solución en el sistema FIA mediante la bomba peristáltica con un caudal de  $0,56 \text{ mL min}^{-1}$ , 3) se inyectaron  $50 \mu\text{L}$  de BROH ( $pH = 12,0$ ), 4) se registraron las absorbancias y se comprobó que en el sistema se generaba un rango de pH (2-6), lo suficientemente amplio como para obtener las cuatro especies deseadas: ácido sórbico ( $pK_a = 4,87$ ) y su sal sorbato, ácido benzoico ( $pK_a = 4,19$ ) y su sal benzoato.

Se optimizaron además: a) la velocidad de la bomba peristáltica, variando la velocidad entre 5 y 12 rpm, b) el volumen de BROH inyectado, probando volúmenes entre 30 y  $150 \mu\text{L}$ , c) el largo del reactor, modificando su longitud entre 1000 y 1900 mm. Para dichas optimizaciones se utilizó también el indicador verde de bromocresol, variando las condiciones de trabajo se estudiaron el ancho y el alto del pico FIA, hasta obtener el que brinda la mayor información.

### **2.2.3.2 SISTEMA FIA PARA LA CUANTIFICACIÓN DE COLORANTES**

Para la cuantificación de colorantes se preparó un sistema FIA simple, semejante al anterior (ver Figura 2.6).

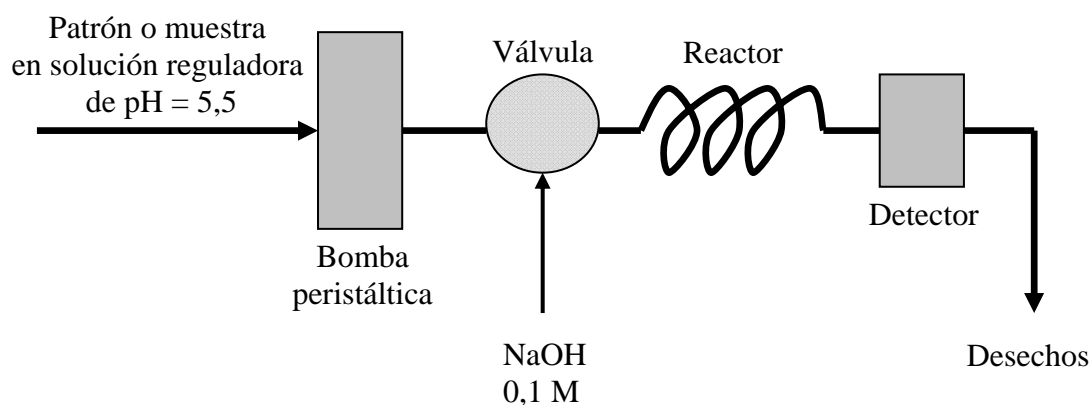


Figura 2.6: Sistema FIA para la cuantificación de colorantes.

Las soluciones de las mezclas de calibración y muestras comerciales se prepararon con solución reguladora de pH = 5,5 ( $\text{KH}_2\text{PO}_4 / \text{Na}_2\text{HPO}_4$ ) (*carrier*). Estas soluciones fueron introducidas al sistema mediante la bomba peristáltica a una velocidad de 5 rpm, con lo que se generó un caudal de  $0,54 \text{ mL min}^{-1}$ .

El gradiente de pH se obtuvo inyectando un bolo de  $50 \mu\text{L}$  de NaOH 0,1 M. El rango de pH obtenido en este sistema está comprendido entre 5,5 (pH del *carrier*) y 12,8 en el centro del pico FIA. El reactor usado fue de 1400 mm de longitud y 0,5 mm de diámetro interno.

Las señales se registraron entre 340 y 750 nm, cada 2 nm, y 2 segundos. Con lo que se obtiene por cada mezcla una matriz de  $206 \times 76$ , o sea 15.656 datos de absorbancia.

### ***Optimización del sistema FIA***

La optimización de este sistema se llevó a cabo utilizando tartrazina y registrando las señales a 430 nm. Se verificó que el rango de pH obtenido estaba entre 5,5 y 12,8 y el mismo fue reproducible para todas las muestras.

Además se optimizaron otras variables, como: a) la velocidad de la bomba peristáltica, modificando el caudal entre  $0,36$  y  $0,72 \text{ mL min}^{-1}$ , b) la longitud del reactor, probando con diferentes reactores cuyas longitudes variaron entre 1300 y 1900 mm, c) el volumen de NaOH inyectado, se estudiaron volúmenes entre 30 y  $100 \mu\text{L}$ , y d) la concentración de NaOH inyectado, la cual se varió entre 0,05 y  $0,15 \text{ mol L}^{-1}$ .

En todos los casos se estudió el ancho y el alto del pico producido por la tartrazina, a 430 nm, modificando las condiciones hasta obtener picos que brindaran la mayor información (cuanto más alto y más ancho sea el pico, mayor información).

## CAPÍTULO 3

### RESULTADOS Y DISCUSIÓN

#### 3.1 MÉTODO DESARROLLADO PARA LA CUANTIFICACIÓN DE CONSERVANTES (SORBATO DE POTASIO Y BENZOATO DE SODIO) EN MUESTRAS COMERCIALES Y ARTIFICIALES, APLICANDO ESPECTROSCOPIA UV Y CALIBRACIÓN MULTIVARIADA DE PRIMER ORDEN (PLS Y HLA/GO).

##### 3.1.1 INTRODUCCIÓN

En la presente tesis se desarrolló experimentalmente un método para la cuantificación de sorbato de potasio (SOR) y benzoato de sodio (BEN) en bebidas comerciales analcohólicas a base de jugos de fruta (naranja, pomelo y limón). Como ya se mencionó en el Capítulo 1, generalmente estos conservantes son utilizados de manera conjunta para potenciar su acción germicida. Se adicionan para preservar una gran variedad de alimentos tales como: jugos, bebidas, mermeladas, confituras, conservas, embutidos, pollos, etc. El límite máximo admitido para bebidas no gasificadas preparadas a base de jugos es de  $1 \text{ g L}^{-1}$  para cada conservante en caso que se agreguen en forma separada, o  $1 \text{ g L}^{-1}$  para la suma de las concentraciones de ambos, en caso de ser utilizados de manera conjunta.

Las muestras comerciales analizadas fueron jugos de naranja concentrados, producidos y comercializados en la zona (JN1 y JN2).

Por otra parte, se prepararon muestras artificiales a partir de jugos de frutas y vegetales de estación (pera, manzana, mandarina, zanahoria y tomate), como fue mencionado en el Capítulo 2 (Materiales). Los jugos que se utilizaron para preparar las muestras artificiales fueron seleccionados teniendo la precaución de que sus matrices presentaran diferencias



significativas con las matrices de los jugos con las que se modelaron las interferencias (naranja, pomelo y limón), con la intención de probar la habilidad predictiva del método HLA/GO, que brinda la posibilidad de seleccionar el rango de longitudes de onda mediante el análisis del error indicador (EIF) para cada muestra en particular.

Además de la variabilidad que presentan las matrices de los jugos de diferentes frutas, existen también diferencias en la matriz para una misma fruta, de acuerdo a la zona de producción, la estación del año, las condiciones climáticas, la alta variabilidad natural (diferentes variedades de una misma especie), estado de maduración, etc. Estos factores, a los que se suma el solapamiento espectral que tiene el espectro de absorción de la matriz con los espectros de los conservantes, hace que el sistema sea muy difícil de resolver analíticamente, sin tener que recurrir a métodos separativos (ver Figura 3.1).

El método fue validado comparando los resultados, con los obtenidos al aplicar el método oficial de la AOAC (*Association of Analytical Communities*). Dentro de la validación se estudió además la precisión, los porcentajes de recuperación, y cifras de merito como: limite de detección, limite de cuantificación y sensibilidad.

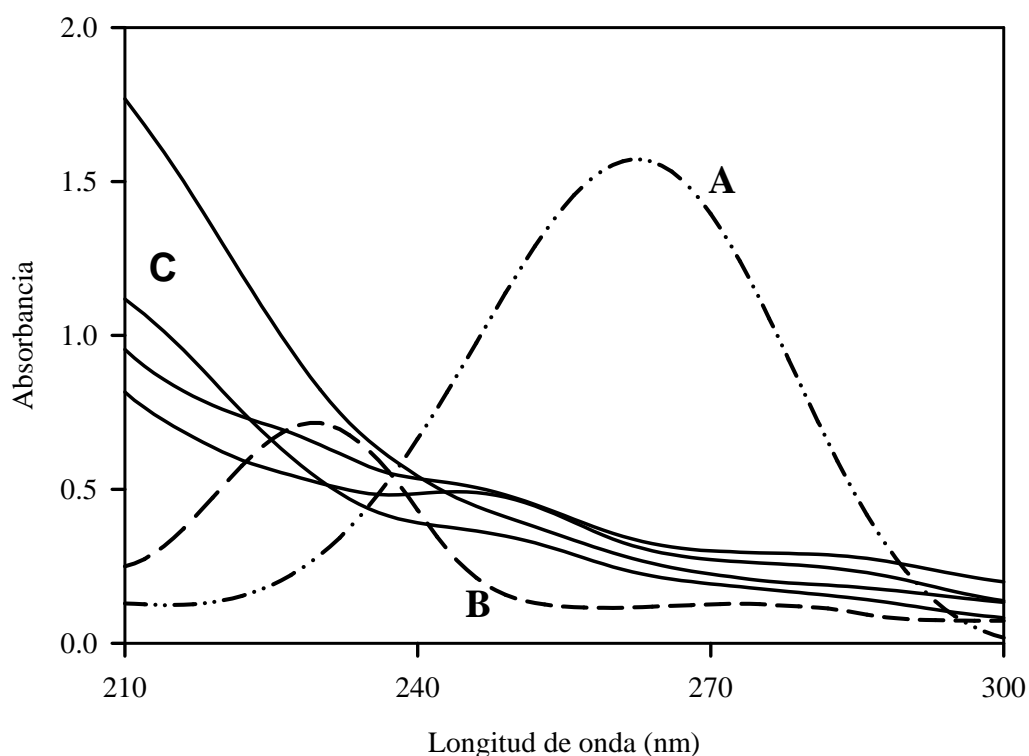


Figura 3.1: Espectros de absorción electrónica de las soluciones en  $\text{HCl } 5 \times 10^{-4} \text{ mol L}^{-1}$ : A) SOR ( $10,0 \text{ mg L}^{-1}$ ), B) BEN ( $10,0 \text{ mg L}^{-1}$ ) y C) jugos de fruta de diferentes orígenes (diluidas 1/100 con el mismo ácido).

### 3.1.2 CALIBRACIÓN

Para la calibración se preparó un juego de 15 mezclas, según un diseño de Box Behnken para tres componentes (SOR, BEN, y jugo), en tres niveles de concentración y con tres muestras centrales.

Debe aclararse que el tercer componente (jugo) corresponde a la matriz de la muestra, la cual se agregó para modelar interferencias. Los tres niveles de este tercer componente corresponden a distintos tipos de frutas: naranja, pomelo y limón, contemplando posibles variaciones en la producción de las bebidas comerciales, (ver Tabla 3.1).

Los niveles de los componentes en las mezclas se encuentran en un rango de 0 a 100 % del límite máximo aceptado para muestras comerciales según recomendaciones del organismo regulador argentino.<sup>3</sup>

Se estudió la linealidad de la respuesta en un amplio rango de concentraciones, y de acuerdo a los resultados obtenidos se decidió el rango de concentraciones para preparar las mezclas de calibración. Las soluciones de ambos analitos cumplen con la ley de Beer en el rango de concentraciones seleccionado.

Las mezclas fueron preparadas por dilución de cantidades convenientes de ambas soluciones patrón, con el agregado del jugo correspondiente (20%), y llevando a volumen final con solución de ácido clorhídrico ( $5 \times 10^{-4} \text{ mol L}^{-1}$ ). El agregado de ácido clorhídrico asegura que las soluciones tengan un pH entre 3,0 y 3,5, en el que ambas especies se encuentren en su forma ácida, evitando el desplazamiento de los picos de absorción.

Calibración mezcla	Nominal ( $\text{g L}^{-1}$ )		Jugo de fruta
	SOR	BEN	
1	0,00	0,00	Naranja
2	1,00	0,00	Naranja
3	0,00	1,00	Naranja
4	1,00	1,00	Naranja
5	0,00	0,40	Limón
6	1,00	0,40	Limón
7	0,00	0,40	Pomelo
8	1,00	0,40	Pomelo
9	0,40	0,00	Limón
10	0,40	1,00	Limón
11	0,40	0,00	Pomelo
12	0,40	1,00	Pomelo
13	0,40	0,40	Naranja
14	0,40	0,40	Naranja
15	0,40	0,40	Naranja

Tabla 3.1: Diseño según el cual se prepararon las mezclas del juego de calibración. Es necesario aclarar que las concentraciones que aparecen en la tabla corresponden a la concentración inicial (simulando las concentraciones de los jugos comerciales). Al efectuar las lecturas de absorbancia estas están diluidas 1/100 con HCl  $5 \times 10^{-4} \text{ mol L}^{-1}$ . El rango de concentraciones para las lecturas es: 0 -  $10 \text{ mg L}^{-1}$ .

Las medidas de absorbancia se registraron con un espectrofotómetro Perkin Elmer entre 210 y 300 nm. En la Figura 3.2 pueden verse los espectros de las mezclas de calibración. Los datos se archivaron en formato ASCII y se procesaron con los programas Multivar y MVC1. Los parámetros y cifras de merito obtenidos se exponen en la Tabla 3.2.

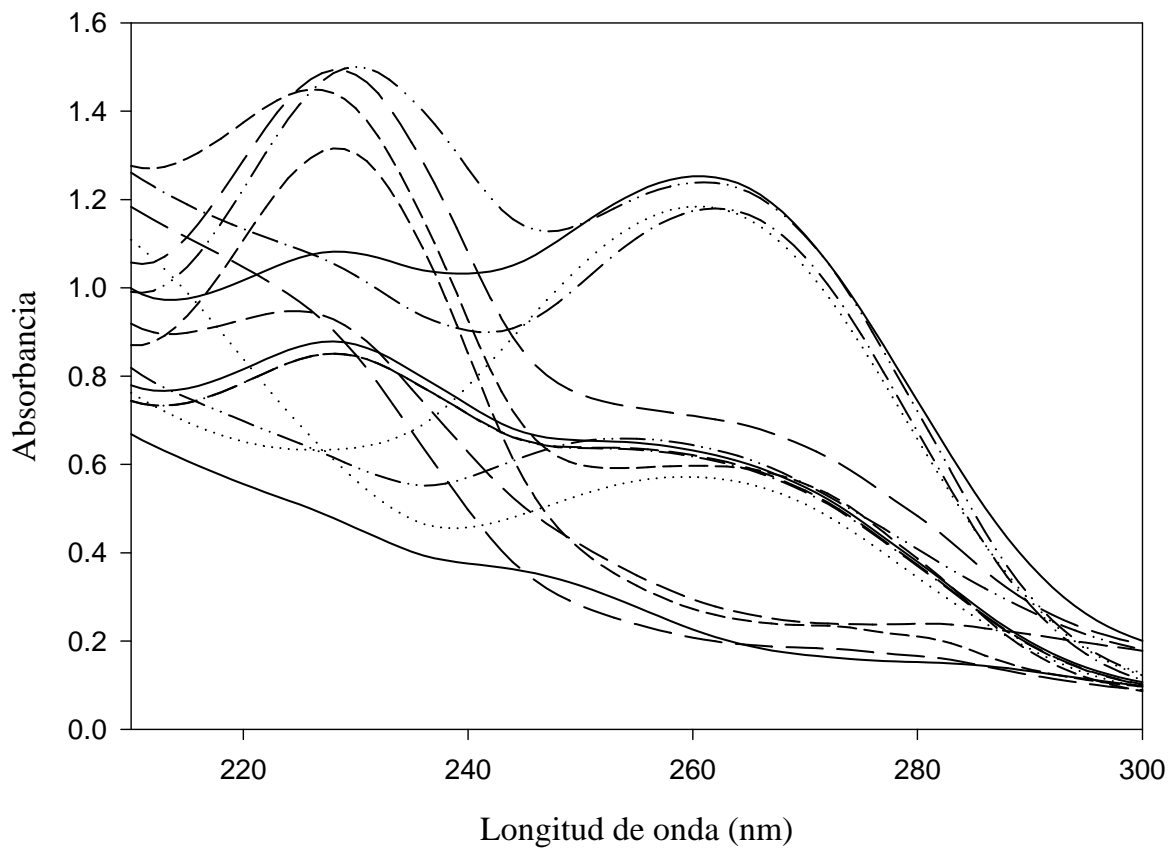


Figura 3.2: Espectros de absorbancia de las 15 mezclas de calibración, adquiridos entre 210 y 300 nm.

	SOR		BEN	
	PLS-1	HLA/GO	PLS-1	HLA/GO
Region (nm)	220 – 300		210 – 300	
N <sup>0</sup> de Factores <sup>a</sup>	5	4	5	4
RMSECV (g L <sup>-1</sup> )	$8,6 \times 10^{-3}$	$7,7 \times 10^{-3}$	$7,0 \times 10^{-3}$	$6,2 \times 10^{-3}$
REP %	1,89	1,70	1,55	1,36
$\ \mathbf{b}_k\ $	0,417	0,429	0,631	0,666
$\ \epsilon\ ^b$	0,051		0,049	
SEN	$2,4 \times 10^{-3}$	$2,3 \times 10^{-3}$	$1,6 \times 10^{-3}$	$1,5 \times 10^{-3}$
SEL	0,46	0,45	0,44	0,41
$\gamma^{-1}$ (mg L <sup>-1</sup> )	0,021	0,022	0,031	0,033
LOD (mg L <sup>-1</sup> )	0,063	0,066	0,093	0,098
LOQ (mg L <sup>-1</sup> )	0,19	0,20	0,28	0,30

<sup>a</sup> El número de factores se seleccionó mediante el método de validación cruzada (Haaland y Thomas)<sup>6</sup> y la región de trabajo mediante la estrategia de ventana móvil.

<sup>b</sup>  $\|\epsilon\|$  corresponde al ruido instrumental y fue calculado a partir de 10 blancos de muestra.

Tabla 3.2: Región espectral de trabajo, parámetros de la calibración y cifras de merito obtenidos en la calibración aplicando los métodos quimiométricos PLS1 y HLA/GO.

Analizando los resultados expuestos en la Tabla 3.2, se observa que los parámetros obtenidos para el método HLA/GO son mejores que los obtenidos para PLS1, tanto para SOR como para BEN. Se observa, por ejemplo, que para los errores (REP y RMSECV), los resultados obtenidos utilizando el método HLA/GO son entre 10 y 13% menores, que para PLS1. Por esta razón se dejó de lado el método PLS1, y los cálculos de predicción y validación se realizaron solo con el método HLA/GO.

### 3.1.3 VALIDACIÓN

La validación se realizó utilizando distintos tipos de muestras artificiales y comerciales, variando las concentraciones de ambos componentes y el tipo de jugo de fruta. Estas muestras para la validación fueron preparadas de la siguiente manera:

**3.1.3.1 JUEGO DE VALIDACIÓN:**

Se preparó un juego de 13 mezclas, según un diseño de Box Behnken para dos componentes en tres niveles y una muestra central. Las concentraciones de ambos analitos en estas mezclas son ligeramente diferentes de las concentraciones con las que se prepararon las mezclas de calibración. Estas mezclas fueron preparadas de manera similar a las mezclas de calibración. Las concentraciones de las mezclas de validación, los valores predichos aplicando el método HLA/GO y los porcentajes de recuperación son expuestos en la Tabla 3.3.

Muestra	SOR			BEN		
	Nominal (g L <sup>-1</sup> )	Predicha (g L <sup>-1</sup> )	Recup. (%)	Nominal (g L <sup>-1</sup> )	Predicha (g L <sup>-1</sup> )	Recup. (%)
1	0,00	0,01	–	0,80	0,76	95,0
2	0,00	0,01	–	0,80	0,77	96,3
3	0,80	0,82	102,5	0,40	0,41	102,5
4	0,80	0,81	101,2	0,40	0,42	105,0
5	0,40	0,41	102,5	0,40	0,42	105,0
6	0,40	0,42	105,0	0,40	0,42	105,0
7	0,40	0,41	102,5	0,00	0,03	–
8	0,40	0,40	100,0	0,80	0,80	100,0
9	0,00	0,01	–	0,00	0,04	–
10	0,80	0,82	102,5	0,80	0,81	101,5
11	0,00	0,01	–	0,40	0,38	95,0
12	0,80	0,82	102,5	0,40	0,38	95,0
13	0,40	0,42	105,0	0,40	0,38	95,0
REP (%)		3,6			5,2	
Media de los porcentajes de recuperación <sup>a</sup>			102,6 (1,6)			99,6 (4,4)

<sup>a</sup> Los valores entre paréntesis corresponden a las desviaciones estándar.

Tabla 3.3: Composición del juego de validación, resultados predichos aplicando el método HLA/GO y porcentajes de recuperación obtenidos.

Analizando los resultados expuestos en la Tabla 3.3, puede observarse que los porcentajes de recuperación son buenos, tanto para SOR como para BEN. Sin embargo el error (REP) para BEN es considerablemente mayor, que para SOR. Esto puede deberse al solapamiento espectral que sufre el BEN por parte de las matrices de los jugos y a la variabilidad intrínseca que estas presentan (ver Figura 3.1). Es importante destacar que los resultados obtenidos son comparables con los obtenidos mediante métodos separativos.<sup>140</sup>

### 3.1.3.2 PRECISIÓN: ANÁLISIS DE UNA MUESTRA COMERCIAL CON LA ADICIÓN DE CONSERVANTES

Se preparó una muestra a partir de una mezcla de jugos comerciales de naranja, pomelo, y limón (20%), a la que se le adicionó una cantidad previamente calculada de solución patrón de SOR y BEN para lograr a una concentración de 0,50 g L<sup>-1</sup> para cada componente, llevando a volumen final con HCl 5×10<sup>-4</sup> mol L<sup>-1</sup>. Esta muestra fue fraccionada en tres submuestras, las que se procesaron durante tres semanas diferentes, realizando diez replicas cada vez. Con los resultados obtenidos se realizó una prueba estadística (ANOVA) para estudiar la precisión intra e inter-ensayos (ver Tabla 3.4).

	SOR	BEN
	Media de las concentraciones predichas (g L <sup>-1</sup> )	
Semana 1 <sup>a</sup>	0,693 (0,007)	0,573 (0,004)
Semana 2 <sup>a</sup>	0,693 (0,006)	0,574 (0,004)
Semana 3 <sup>a</sup>	0,694 (0,005)	0,573 (0,004)
ANOVA	No existe diferencia significativa entre las medias	
	$p = 0,812$	$p = 0,995$
CV % (promedio)	0,83	0,68

<sup>a</sup> Los valores entre paréntesis corresponden a las desviaciones estándar (10 replicas).

Tabla 3.4: Parámetros estadísticos obtenidos en el estudio de precisión del método.

En la Tabla 3.4 son expuestas las medias de los resultados obtenidos al procesar las diez replicas durante tres semanas consecutivas. Con estas medias se realizó la prueba estadística (ANOVA), resultando que no existe diferencia significativa entre las medias de las concentraciones predichas. Se puede concluir que las variaciones que se observan en los resultados son producidas por errores aleatorios. Además, en la tabla aparecen los coeficientes de variación (CV %); como puede observarse los resultados son excelentes.

### **3.1.3.3 EXACTITUD: ANÁLISIS DE MUESTRAS COMERCIALES CON ADICIÓN DE CONSERVANTES**

Se tomaron dos muestras comerciales JN1 y JN2. A la primera se le agregó la cantidad necesaria de solución patrón de SOR como para obtener concentraciones de 0,20; 0,40; 0,60 y 0,80 g L<sup>-1</sup>, en cada caso. De la misma manera se procedió con la segunda muestra pero agregando solución patrón de BEN. Estas muestras fueron empleadas para evaluar la capacidad de predicción del analito en presencia de la matriz (recuperación) aplicando el método estadístico de regresión lineal sobre los resultados obtenidos. Se realizó un ajuste por cuadrados mínimos entre la concentración nominal y la concentración predicha, calculandose: pendiente, ordenada al origen, coeficiente de correlación y demás estadísticos de la regresión.

Para que un método sea exacto debe cumplirse que la ordenada al origen sea cero, y la pendiente uno. Para evaluar la pendiente calculada en ambos casos se empleó la siguiente ecuación:<sup>141</sup>

$$|R-1| \leq t_{\alpha/2, eff} u(R) \quad (3.1)$$

donde:  $R$  es la pendiente calculada,  $t_{\alpha/2, eff}$  es la  $t$  de Student tabulada para  $\alpha/2$ , y  $eff$  grados de libertad, asociada con la incertidumbre  $u(R)$ . La incertidumbre puede estimarse con la desviación estándar del ajuste.

En nuestro caso la  $t$  tabulada es:  $t_{0,025, (15-2)} = 2,13$ . Los valores obtenidos mediante el método de cuadrados mínimos son informados en la Tabla 3.5.



	Pendiente	Desviación estándar ( $S_b$ )	$R - 1$	$t_{\alpha/2, eff} u(R)$	Coefficiente de regresión ( $r^2$ )
SOR	1,02	0,04	0,02	0,085	0,9998
BEN	1,06	0,04	0,06	0,085	0,9998

Tabla 3.5: Estadísticos obtenidos al aplicar regresión lineal a los resultados de la recuperación

De acuerdo a estos valores al aplicar la ecuación (3.3), se obtuvo:  $|R-1| < 0,085$  para ambos analitos. Con lo cual podemos concluir que los resultados obtenidos en la recuperación son aceptables (tanto para SOR, como para BEN).

#### **3.1.3.4 ANÁLISIS DE MUESTRAS ARTIFICIALES CON MATRICES SIGNIFICATIVAMENTE DIFERENTES A LAS UTILIZADAS PARA CALIBRAR**

Se prepararon cinco muestras artificiales con jugos cuyas matrices no fueron modeladas en la calibración: manzana, tomate, pera, mandarina, y zanahoria. En todas ellas se adicionó la cantidad necesaria de solución patrón de ambos conservantes como para obtener una concentración de  $0,50 \text{ g L}^{-1}$  (M1, M2, M3, M4 y M5 respectivamente). Este grupo de muestras fue utilizado para investigar el efecto producido por matrices con interferencias no modeladas.

Como puede observarse en la Figura 3.3 estas matrices presentan espectros de absorción significativamente diferentes a los espectros de las matrices con las que se modeló. Para mejorar las predicciones se seleccionaron las regiones de trabajo estudiando el EIF, como fuera discutido en el Capítulo 1 (punto 1.4.3) y realizando las graficas NASRP. En la Figura 3.4 (A-B), pueden observarse dos gráficos de regresión (NASRP). El superior [Figura 3.4 (A)] realizado con el espectro completo (210-300 nm), donde puede verse la falta de linealidad a causa de las interferencias no modeladas en la calibración, y presentes en esta muestra en particular. Por otra parte, en la Figura 3.4 (B) realizada con la región seleccionada (220-256), se ve claramente como se corrige la linealidad, indicando que en la región encontrada no está presente (o al menos se hace muy pequeño) el efecto de la interferencia no modelada.

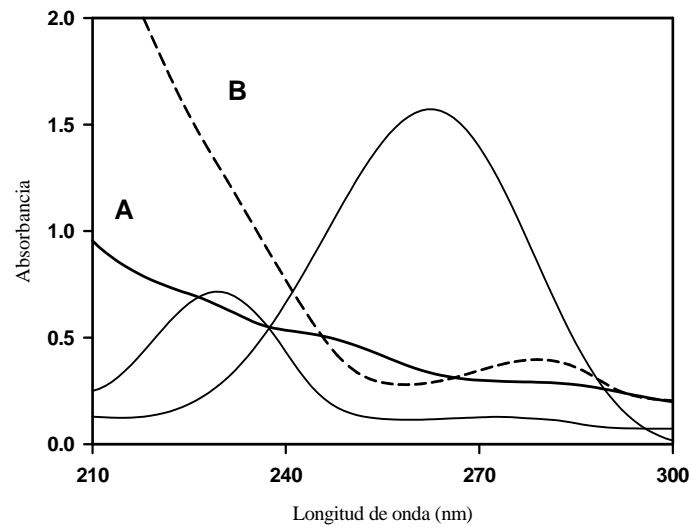


Figura 3.3: Espectros de absorción electrónica para: A) jugo de fruta con el que se modelaron las interferencias, y B) jugo de fruta no modelado: manzana, ambos diluidos 1/100 en  $\text{HCl } 5 \times 10^{-4} \text{ mol L}^{-1}$ . Los otros dos espectros presentes corresponden al SOR y BEN (ambas de  $10,0 \text{ mg L}^{-1}$ ).

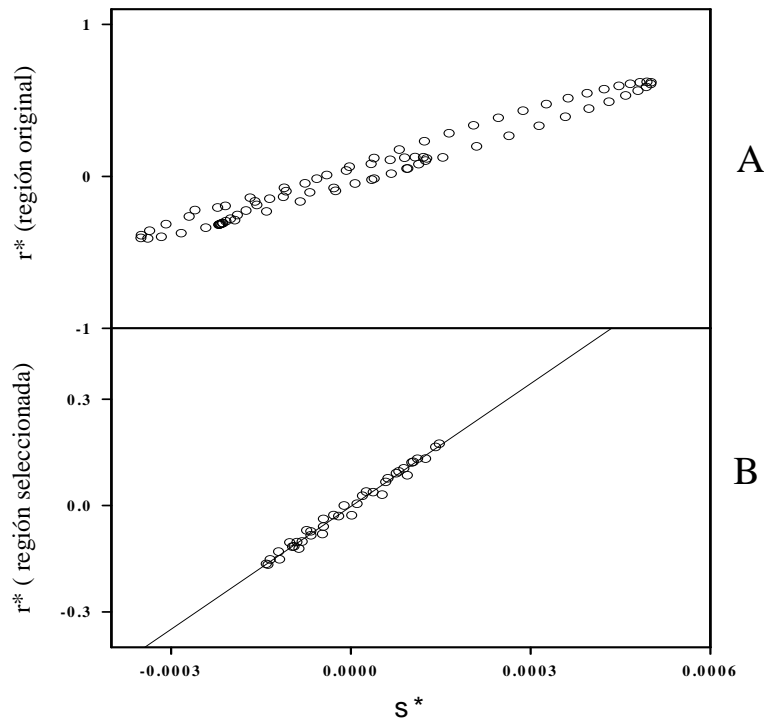


Figura 3.4: A) NASRP para SOR en una muestra de jugo de manzana, en un rango de 210-300 nm. B) NASRP para la misma muestra pero en el rango seleccionado (220-256 nm) mediante el estudio del EIF. La línea sólida corresponde a la recta de regresión para los puntos calculados.

En la Tabla 3.6 se informan las concentraciones predichas con todos los datos del espectro, y las predichas utilizando solo un determinado rango. Además se informa el EIF correspondiente, y los porcentajes de recuperación para cada caso.

SOR				
JUGO	Rango de long. de onda (nm)	EIF	Concentraciones predichas (g L <sup>-1</sup> )	Recuperación (%)
Manzana	220 – 300	0,22	0,58	116
	220 – 256	0,16	0,49	98
Tomate	220 – 300	0,019	0,66	132
	230 – 270	0,017	0,58	116
Pera	220 – 300	0,30	0,60	120
	220 – 270	0,25	0,50	100
Mandarina	220 – 300	0,17	0,50	100
Zanahoria	220 – 300	0,14	0,76	152
	220 – 245	0,09	0,44	88
BEN				
Manzana	210 – 300	0,31	0,71	142
	210 – 250	0,29	0,53	106
Tomate	210 – 300	0,12	0,23	46
	230 – 270	0,10	0,37	74
Pera	210 – 300	0,46	0,47	94
	220 – 280	0,49	0,49	98
Mandarina	210 – 300	0,40	0,41	82
	220 – 250	0,21	0,47	94
Zanahoria	210 – 300	0,30	0,23	46
	210 – 270	0,21	0,49	98

Tabla 3.6: Resultados de las predicciones de las cinco muestras preparadas con jugos no modelados. Se presentan también los rangos de longitudes de onda seleccionados en cada caso, y el EIF correspondiente. Las concentraciones nominales de SOR y BEN fueron 0,5 g L<sup>-1</sup> en todas las muestras.

Analizando los resultados obtenidos, puede observarse claramente como se mejoran las predicciones acotando el rango de longitudes de onda. Esto permite aplicar el método a diferentes tipos de muestras comerciales, aun cuando las frutas utilizadas para su producción presenten marcadas diferencias con las que se utilizaron en la calibración.

### 3.1.3.5 ANÁLISIS DE MUESTRAS COMERCIALES

Dos muestras comerciales (JN1 y JN2), fueron analizadas mediante el método HLA/GO y el método oficial de la AOAC<sup>142</sup> (descrito en el Capítulo 2: “Materiales y Métodos”). Se compararon los resultados, validando de esta manera la exactitud del método.

Las muestras comerciales fueron centrifugadas y diluidas 1/100 con ácido clorhídrico  $5 \times 10^{-4}$  mol L<sup>-1</sup> (pH = 3,3), antes de registrar las medidas. Los resultados obtenidos al procesar tres replicados se informan en la Tabla 3.7.

Muestras comerciales	SOR		BEN	
	Método oficial	HLA/GO	Método oficial	HLA/GO
1 <sup>a</sup>	0,42 (0,09)	0,39 (0,02)	0,00	0,00
2 <sup>a</sup>	0,22 (0,10)	0,20 (0,02)	0,85 (0,10)	0,79 (0,03)

<sup>a</sup> Los valores entre paréntesis corresponden a la desviación estándar (tres replicas)

Tabla 3.7: Resultados obtenidos aplicando HLA/GO y el método oficial de la AOAC, para SOR y BEN en dos muestras comerciales (JN1 y JN2).

Puede observarse que los resultados son razonablemente buenos. Es importante considerar que el método oficial es muy tedioso, siendo necesario largos tiempos de procesamiento (destilación por arrastre con vapor, preparación de diferentes soluciones, etc.). Por otro lado, la precisión obtenida mediante el método de calibración multivariada es considerablemente mejor que la del método oficial.

## **3.2 MÉTODO DESARROLLADO PARA LA CUANTIFICACIÓN DE SORBATO DE POTASIO Y BENZOATO DE SODIO EN MUESTRAS COMERCIALES, APLICANDO MÉTODOS DE CALIBRACIÓN MULTIVARIADA DE SEGUNDO ORDEN (PARAFAC Y MCR-ALS) A DATOS DE ESPECTROS DE ABSORBANCIA OBTENIDOS EN GRADIENTE DE pH.**

### **3.2.1 INTRODUCCIÓN**

Ya fue discutido anteriormente la forma en que se utilizan los conservantes (SOR y BEN), y la importancia de su cuantificación. También se discutió el problema analítico que ocasionan las matrices de los jugos, y la necesidad de recurrir a métodos separativos para su cuantificación. En la bibliografía se encuentran numerosos trabajos publicados sobre diferentes maneras de cuantificar conservantes por métodos separativos como CLAR y cromatografía en capa delgada,<sup>143,144</sup> electroforesis capilar (micelar electrocinética),<sup>145</sup> al <sup>148</sup> extracción con solvente y derivada de la señal espectrofotométrica,<sup>149</sup> polarografía,<sup>150</sup> o determinación enzimática,<sup>151</sup> entre otros.

Todos estos métodos requieren pasos previos para el tratamiento de la muestra, son engorrosos y requieren largo tiempos de análisis, al igual que el método oficial de la AOAC.

En el método descrito en el punto anterior se demostraron los beneficios de la aplicación de métodos multivariados para respuestas de primer orden (HLA/GO), por los buenos resultados obtenidos y la sencillez del procedimiento.

El método que se describe en este punto propone mejorar los resultados ya obtenidos, para lo cual se diseña un sistema donde se generan datos de segundo orden. Estos datos, posteriormente se ordenan en arreglos de tres vías. Esto brinda una mayor información y permite obtener la ventaja de segundo orden, gracias a la cual no es necesario el agregado de jugos en el juego de calibración para modelar las interferencias. De esta manera es posible independizarse del problema de las variaciones propias de la matriz y poder aplicarlo a diferentes bebidas comerciales. El juego de calibración se realizó con mezclas preparadas solamente con los patrones de los conservantes, y fueron analizadas muestras reales.

La matriz de datos se obtuvo mediante un sistema FIA en el cual se genera un doble gradiente de pH, (ver Figura 2.4, Capítulo 2), y realizando las lecturas con un detector de

arreglo de diodos. La muestra, diluida en medio ácido (BRH, pH = 2,0), se introdujo en el sistema mediante una bomba peristáltica, con un flujo de  $0,56 \text{ mL min}^{-1}$ , registrando de esta forma la línea de base. Posteriormente, se inyectó en el flujo un bolo de  $50 \mu\text{L}$  de solución Britton-Robinson alcalina (pH = 12,0). Esto generó el doble gradiente de pH, produciendo el pico FIA en la señal, (ver Figura 2.5).

En el sistema, donde inicialmente a pH ácido, teníamos solo dos especies (ácido sórbico y ácido benzoico), al generar el gradiente de pH, cada una de las especies establece un equilibrio ácido-base con su sal (sorbato y benzoato). Un comportamiento similar tiene la matriz de la muestra (Figura 3.5). Una característica de los datos producidos de esta manera, es lo que se denomina matemáticamente como “deficiencia de rango”, lo que significa que el rango obtenido con los datos del conjunto de los componentes no es igual a la suma de los rangos de las contribuciones de las especies individualmente. Este fenómeno es consecuencia de la correlación que existe entre las concentraciones de la especie ácida y su sal, las que se van modificando a lo largo del gradiente de pH en forma dependiente (ver Figura 3.6).

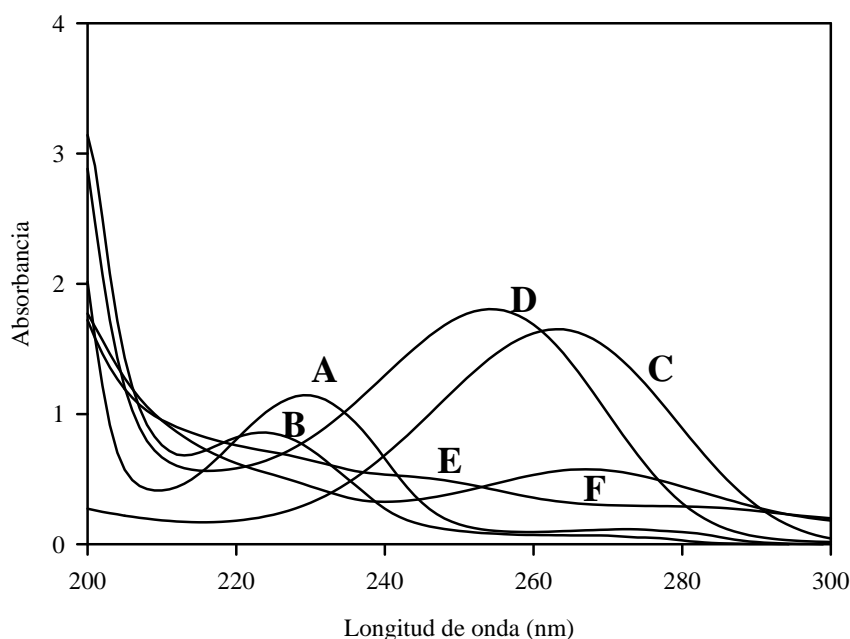


Figura 3.5: Espectros de absorción: A) ácido benzoico, B) benzoato, C) ácido sórbico, D) sorbato, E) jugo de naranja en medio ácido y F) jugo de naranja en medio alcalino. Las concentraciones del ácido benzoico y ácido sórbico son de  $10 \text{ mg L}^{-1}$  y para el jugo 20 %. (concentración equivalente a la que presentan las muestras comerciales). Los valores de pH fueron: pH = 2 para los espectros A, C y E, y pH = 6 para los espectros B, D y F.

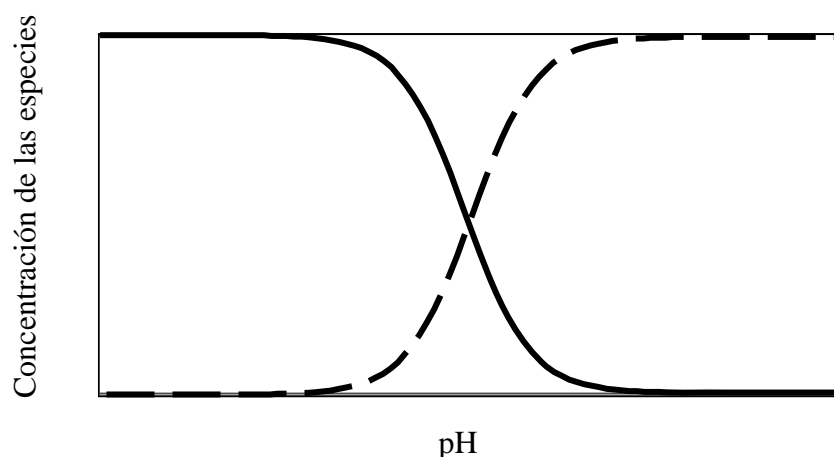


Figura 3.6: Variación de las concentraciones de las especies ácida y básica, en función del pH. Puede observarse la dependencia lineal que existe entre ambos (la suma de las concentraciones a los diferentes pH es una constante).

### 3.2.2 CALIBRACIÓN

Para la calibración se prepararon nueve mezclas, según un diseño central compuesto para dos componentes en cinco niveles. Las concentraciones extremas del diseño fueron 0 y 12 mg L<sup>-1</sup> tanto para SOR como para BEN. Las soluciones se prepararon diluyendo la cantidad necesaria de ambas soluciones patrón (1,0 g L<sup>-1</sup>) como para obtener las concentraciones deseadas, llevando a volumen final con solución reguladora BRH (pH = 2,0). En la Tabla 3.8 se exponen las concentraciones de ambos analitos en las mezclas de calibración.

Muestra	BEN (mg L <sup>-1</sup> )	SOR (mg L <sup>-1</sup> )
C1	1,80	1,80
C2	1,80	10,00
C3	10,00	1,80
C4	10,00	10,00
C5	6,00	0,00
C6	6,00	12,00
C7	0,00	6,00
C8	12,00	6,00
C9	6,00	6,00

Tabla 3.8: Concentraciones para SOR y BEN en las mezclas de calibración, en solución reguladora BRH.

Como fuera comentado anteriormente, la respuesta es lineal con respecto a la concentración, en todo el rango de trabajo.

Se procesaron las muestras con el sistema FIA y se adquirieron las señales en un rango de longitudes de onda de 200 y 300 nm, cada 2 nm y 1,5 segundos. El orden en que se introdujeron las mezclas fue aleatorio. Por cada mezcla de calibración se obtuvo una matriz de 51x 33, lo que equivale a 1683 datos de absorbancia, esto representa en el espacio una superficie como la que se observa en la Figura 3.7. El arreglo de datos de tres vías se obtiene ordenando las nueve matrices de calibración, resultando con las siguientes dimensiones: 51 x 33 x 9. Para aplicar la ventaja de segundo orden, a este arreglo de calibración se le agrega la matriz de la muestra para poder realizar las predicciones. El arreglo resultante tiene las siguientes dimensiones: 51 x 33 x 10.

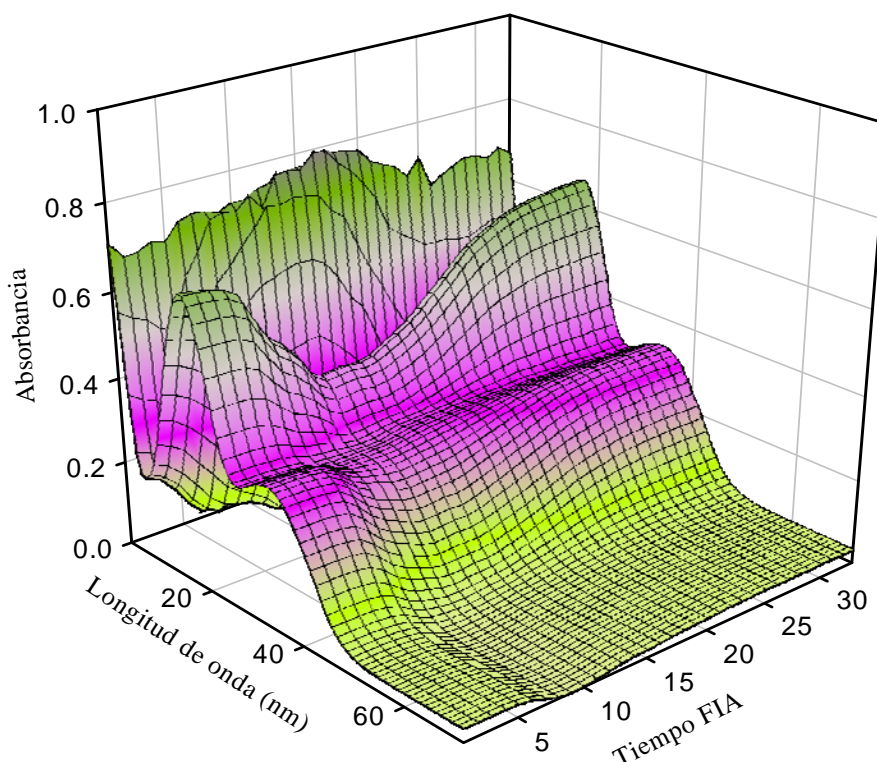


Figura 3.7: Superficie correspondiente a la señal adquirida para la primer mezcla de calibración.



### 3.2.3 VALIDACIÓN

La validación se realizó utilizando distintos tipos de muestras: artificiales y comerciales, variando las concentraciones de ambos componentes y con jugo de diferentes frutas. Estas muestras usadas en la etapa de validación fueron preparadas de la siguiente manera:

#### 3.2.3.1 JUEGO DE VALIDACIÓN:

Se preparó un juego de nueve mezclas de validación según un diseño factorial completo (*full factorial*) para dos componentes en tres niveles (ver Tabla 3.9). Dichas mezclas fueron preparadas por dilución de la cantidad necesaria de solución patrón de  $1,0 \text{ g L}^{-1}$  como para obtener la concentración deseada, en solución reguladora BRH ( $\text{pH} = 2,0$ ). A estas mezclas se la adicionó jugo de naranjas frescas exprimidas, al 20 % (concentración similar a la que presentan las bebidas comerciales), añadiendo jugos de diferentes frutas a cada una de las mezclas. El registro de las señales para estas mezclas se efectuó de manera aleatoria, y en días diferentes en los que se realizó la calibración.

Muestra	BEN ( $\text{mg L}^{-1}$ )	SOR ( $\text{mg L}^{-1}$ )
T1	2,00	2,00
T2	6,70	2,00
T3	10,70	2,00
T4	2,00	6,70
T5	6,70	6,70
T6	10,70	6,70
T7	2,00	10,70
T8	6,70	10,70
T9	10,70	10,70

Tabla 3.9: Concentraciones para SOR y BEN en las mezclas de validación

#### *Procesamiento de los datos:*

Se utilizaron tres métodos quimiométricos para el tratamiento de los datos: *N*-PLS, PARAFAC y MCR-ALS, comparando posteriormente los resultados obtenidos.

***N-PLS:***

Se procesaron los datos obtenidos para el juego de calibración empleando el método *N-PLS*. Primeramente, se realizó una validación cruzada, y aplicando el criterio de Haaland y Thomas para la selección del número óptimo de factores. Se determinó este valor, que para nuestro sistema resultó igual a tres. Este valor es coherente ya que en las mezclas de calibración se tienen cuatro especies, pero el sistema es deficiente de rango en una unidad. Posteriormente se realizaron las predicciones de las concentraciones de las mezclas del juego de validación. Estos resultados se exponen en la Tabla 3.11.

Debe aclararse que el método *N-PLS* es una extensión de *PLS1*, para respuestas de orden superior y que no utiliza la ventaja de segundo orden ya que no agrega información de las muestras en la etapa de calibración.

***PARAFAC y MCR-ALS:***

Posteriormente, se predijo usando los métodos *PARAFAC* y *MCR-ALS*, los que fueron implementados según la manera descrita en la Tabla 3.10.

	PARAFAC	MCR-ALS
1	Construye la hipermatriz de datos $(I + 1) \times J \times K$ ordenando las respuestas (matrices de dos vías: $J \times K$ ) de las $I$ muestras de calibración, desde 1 hasta $I$ , e incorporando la matriz de la muestra desconocida.	Construye la hipermatriz de datos $(I + 1) \times J \times K$ ordenando las respuestas (matrices de dos vías: $J \times K$ ) de las $I$ muestras de calibración, desde 1 hasta $I$ , e incorporando la matriz de la muestra desconocida.
2	Inicialmente es necesario suministrar una estimación de los perfiles <b>A</b> , <b>B</b> y <b>C</b> .	Inicialmente es necesario suministrar una estimación de <b>S</b> y <b>G</b> .
3	Descompone el cubo de datos y obtiene <b>A</b> , <b>B</b> y <b>C</b> , aplicando restricciones de no-negatividad en todos los modos.	Aplicando cuadrados mínimos alternantes obtiene <b>S</b> y <b>G</b> , aplicando restricciones de no-negatividad en todos los modos.
4	Identifica en <b>A</b> y <b>B</b> el par de perfiles correspondientes a ambas especies del analito de interés ( $n_1$ y $n_2$ )	Identifica en <b>S</b> el par de perfiles correspondientes a ambas especies del analito de interés ( $n_1$ y $n_2$ )
5	Por separado realiza una regresión entre el primer elemento de las $I$ columnas $\mathbf{a}_{n1}$ y $\mathbf{a}_{n2}$ y las concentraciones de los estándares $\mathbf{c}_{cal}$ (calibración pseudo-univariada). Obtiene las constantes de proporcionalidad $k_1$ y $k_2$ : $[a_{1,n1} \dots a_{I,n1}] = k_1 \times \mathbf{c}_{cal}^T$ $[a_{1,n2} \dots a_{I,n2}] = k_2 \times \mathbf{c}_{cal}^T$	Por separado realiza una regresión entre el primer elemento de las $I$ columnas $\mathbf{g}_{n1}$ y $\mathbf{g}_{n2}$ y las concentraciones de los estándares $\mathbf{c}_{cal}$ (calibración pseudo-univariada). Obtiene las constantes de proporcionalidad $k_1$ y $k_2$ : $[g_{1,n1} \dots g_{I,n1}] = k_1 \times \mathbf{c}_{cal}^T$ $[g_{1,n2} \dots g_{I,n2}] = k_2 \times \mathbf{c}_{cal}^T$
6	Convierte las concentraciones relativas en absolutas para las muestras desconocidas, utilizando para el cálculo el último elemento de las columnas $\mathbf{a}_{n1}$ y $\mathbf{a}_{n2}$ : $c_{unk,1} = a_{I+1,n1} / k_1$ $c_{unk,2} = a_{I+1,n2} / k_2$	Convierte las concentraciones relativas en absolutas para las muestras desconocidas, utilizando para el cálculo el último elemento de las columnas $\mathbf{g}_{n1}$ y $\mathbf{g}_{n2}$ : $c_{unk,1} = g_{I+1,n1} / k_1$ $c_{unk,2} = g_{I+1,n2} / k_2$
7	Promedia los resultados obtenidos para $n_1$ y $n_2$ : $c_{unk} = (c_{unk,1} + c_{unk,2})/2$	Promedia los resultados obtenidos para $n_1$ y $n_2$ : $c_{unk} = (c_{unk,1} + c_{unk,2})/2$

Tabla 3.10: Pasos que se siguen en la calibración y predicción para los métodos PARAFAC y MCR-ALS utilizando datos de segundo orden.

Debe aclararse que fue necesario aplicar restricciones de no negatividad para los perfiles espectrales, unicidad y no negatividad para los perfiles de concentración, y de cierre para las especies ácido-sal, a los métodos PARAFAC y MCR/ALS. Además, debido a la dependencia lineal que existe entre las concentraciones de cada uno de los ácidos y su correspondiente sal (ver Figura 3.6) fue necesario introducir información adicional (*oldloads*, en la rutina escrita

por Bro para PARAFAC)<sup>139</sup> a los datos de calibración. Se introdujeron para el cálculo los espectros de los dos analitos puros (SOR y BEN) tanto a pH ácido como alcalino, los espectros de la interferencia, también a pH extremos, y la curva de distribución de la concentración de las especies en función del pH.

En el caso de PARAFAC el número de componentes ( $N$ ) puede ser estimado de diferentes maneras: 1) mediante el parámetro interno conocido como “*core consistency*”, el que debe tener un valor menor del 60%,<sup>44</sup> y 2) mediante el estudio del error de la recta de regresión pseudounivariada.<sup>152</sup> En nuestro caso el número de componentes estimado coincide con el número de componentes de las muestras. La construcción de un modelo de cinco componentes fue satisfactorio para todas las muestras sintéticas: estos cinco componentes pueden atribuirse al ácido sórbico y su sal (sorbato), al ácido benzoico y su sal (benzoato), y la contribución de la línea de base (matriz del jugo de frutas). La variación de la línea de base (para el jugo de naranja) con respecto al pH no es lo suficientemente significativa como para requerir la consideración de otro componente en el sistema.

Por otra parte, para las muestras reales, las cuales en algunos casos contienen mezclas de jugos de diferentes tipos de frutas, el número de componentes es mayor.

Aplicando el método PARAFAC para 5 componentes, al procesar la muestra T1, se obtienen los perfiles espectrales que se exponen en la Figura 3.8 (A) y los perfiles de pH expuestos en la Figura 3.8 (B). Al comparar los perfiles espectrales obtenidos con los espectros puros de los estándares, el resultado es muy satisfactorio. Si se observan los espectros puros de la Figura 3.5 con los perfiles de la Figura 3.8 (A), puede comprobarse la semejanza que existe entre el N<sup>0</sup> 1 de la Figura 3.8 (A) con el ácido benzoico, el N<sup>0</sup> 2 con el benzoato, el N<sup>0</sup> 3 con el ácido sórbico, el N<sup>0</sup> 4 con el sorbato y el N<sup>0</sup> 5 con un promedio de la línea de base de las matrices de los diferentes jugos de fruta.

Los perfiles de pH que se muestran en la Figura 3.8 (B) tienen, como era de esperar un comportamiento predecible. Nótese que la intersección entre los perfiles para el ácido benzoico y el benzoato ( $pK_a = 4,19$ ) (componentes N<sup>0</sup> 1 y 2) ocurre a pH menores que para las especies ácido sórbico y sorbato ( $pK_a = 4,87$ ) (componentes N<sup>0</sup> 3 y 4), de acuerdo con el orden que tienen los valores de los  $pK_a$  de estos compuestos.

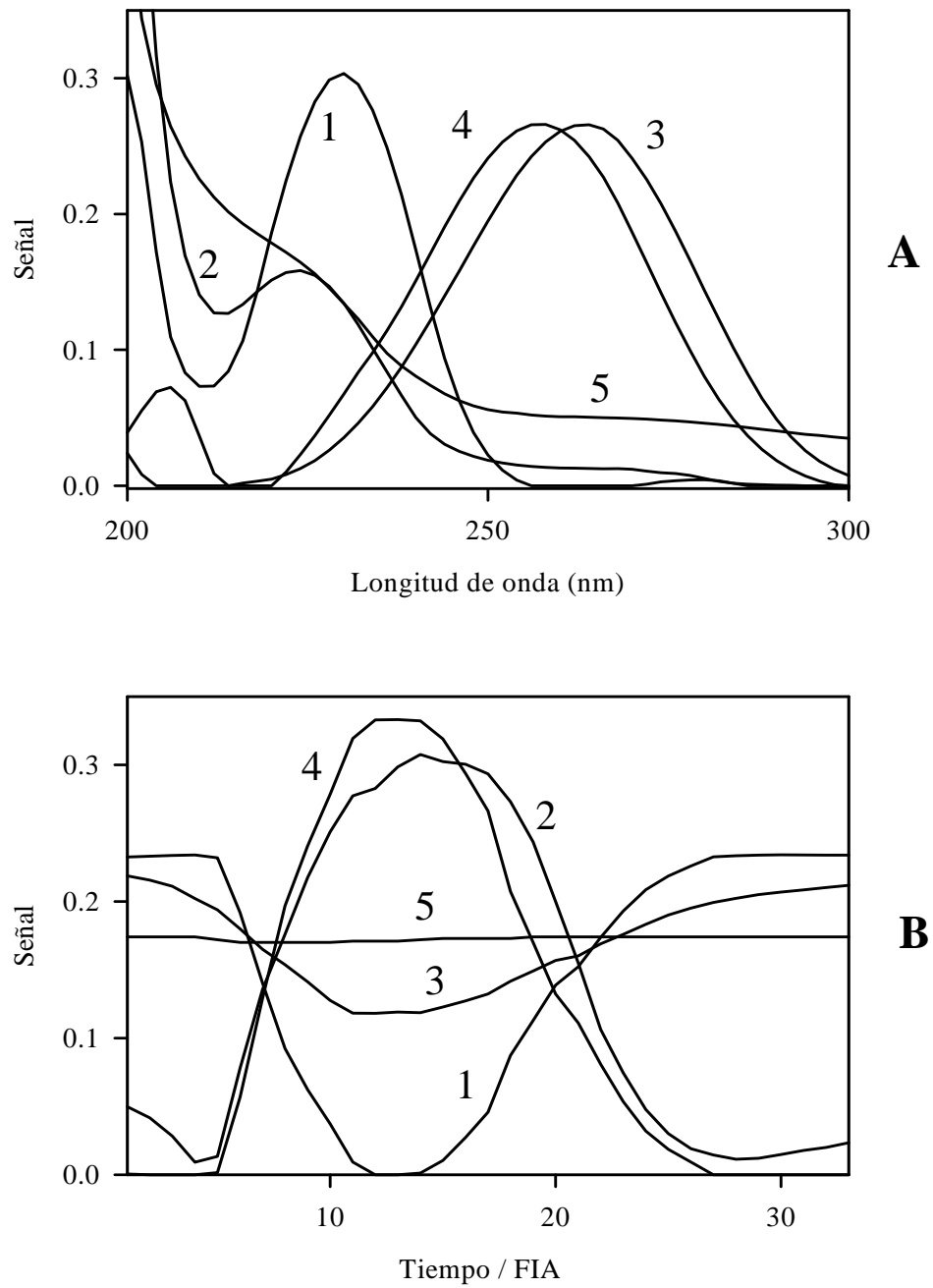


Figura 3.8: A) Perfiles espectrales normalizados provistos por el método PARAFAC para 5 componentes, cuando se procesó una muestra sintética (T1). B) Perfiles de pH normalizados. En ambos casos, la numeración corresponde al orden asignado por el modelo en términos de la contribución a la varianza espectral.

Como ya fuera mencionado, la aplicación de MCR-ALS también requiere de información adicional inicial, sobre los parámetros del sistema. Esta debe tener valores lo más cercanos posible a los valores de los resultados finales. Por ejemplo, se requieren los espectros de los estándares puros, o en el caso de muestras con posibles interferencias es necesario suministrar los espectros de estas interferencias.

Como se menciona en la Tabla 3.10, al igual que con PARAFAC es necesario agregar la matriz de la muestra a la matriz de calibración al construir el arreglo de tercer orden con el que se modela.

Como ventaja, los autores argumentan que MCR-ALS requiere menor cantidad de mezclas de calibración que PARAFAC,<sup>153,154</sup> pero en nuestro caso se utilizó el mismo juego de calibración de nueve mezclas para ambos métodos para poder comparar los resultados obtenidos por ambos métodos.

Los resultados obtenidos en las predicciones de las concentraciones de las mezclas de validación, aplicando los tres métodos son expuestos en la Tabla 3.11:

Muestra	Nominal <sup>a</sup>		N-PLS <sup>a</sup>		PARAFAC <sup>a</sup>		MCR-ALS <sup>a</sup>	
	BEN	SOR	BEN	SOR	BEN	SOR	BEN	SOR
T1	2,00	2,00	4,50	1,91	2,75	2,10	3,70	2,71
T2	6,70	2,00	6,81	1,92	5,95	2,01	4,93	1,83
T3	10,70	2,00	9,03	2,02	9,66	1,95	8,61	1,75
T4	2,00	6,70	4,65	6,04	2,13	6,93	1,15	7,52
T5	6,70	6,70	7,04	6,57	7,02	6,82	6,03	6,73
T6	10,70	6,70	9,13	7,33	11,15	6,74	10,12	6,71
T7	2,00	10,70	4,83	10,75	1,93	11,03	1,84	10,84
T8	6,70	10,70	6,85	10,12	6,74	10,81	6,71	10,46
T9	10,70	10,70	9,01	10,04	10,82	10,61	10,85	10,23
RMSE <sup>a</sup> (mg L <sup>-1</sup> )			1,92	0,45	0,56	0,16	1,22	0,44
REP <sup>a</sup> (%)			29,7	7,0	8,7	2,5	18,9	6,84

<sup>a</sup> Valores expresados en mg L<sup>-1</sup>.

Tabla 3.11: Resultados obtenidos al predecir las mezclas de validación mediante los métodos N-PLS, PARAFAC y MCR-ALS.

### ***Análisis de los resultados:***

Analizando los resultados obtenidos cuando se aplica *N*-PLS (ver Tabla 3.11), puede observarse que el error en la predicción es muy importante. Por ejemplo, para los valores de BEN predichos el % de recuperación en algunos casos es de hasta el 250 %, y el REP: 29,7 %. Este error es debido a las interferencias presentes en la matriz del jugo, que como ya fue discutido anteriormente *N*-PLS es sensible a ellas. Concluimos que es imposible aplicar *N*-PLS para resolver este tipo de problema. *N*-PLS es el más sencillo de los tres métodos, no requiere información adicional, pero su desventaja de no poder implementar la ventaja de segundo orden, lo hace desde el punto de vista analítico, el menos robusto de los tres.

Interpretando los resultados obtenidos al aplicar los dos modelos multivariados: PARAFAC y MCR-ALS, puede observarse que se obtienen parámetros estadísticos notablemente mejores que para *N*-PLS, lo que concuerda con la discusión acerca de la utilización de la ventaja de segundo orden. A su vez, los resultados provistos por PARAFAC son mucho mejores que los obtenidos por MCR-ALS, lo que indica que no existen desviaciones serias de la trilinealidad, condición requerida por la primera de estas técnicas.

#### **3.2.3.1 MUESTRAS COMERCIALES:**

Fueron analizadas cuatro muestras comerciales: JN1) Jugo de naranja concentrado marca “MOCORETÁ”, JN2) Jugo de naranja concentrado marca “MIJÚ”, JN3) Jugo de naranja concentrado marca “COTTI” y JN4) Jugo de naranja listo para beber marca “Hi-C”,

Las muestras fueron preparadas colocando 0,25 mL de jugo comercial en un matraz aforado de 25,00 mL y enrasando con solución reguladora BRH.

Las concentraciones declaradas para cada uno de los conservantes en las muestras comerciales es de  $\cong 0,5 \text{ g L}^{-1}$ , estas concentraciones fueron consideradas para la selección del rango de concentraciones del jugo de calibración, y se realizaron las diluciones correspondientes para que las señales estén en la zona de mínimo error para las lecturas y dentro del rango de linealidad.

Las señales correspondientes a las muestras comerciales fueron registradas en orden aleatorio, y en diferentes días en los que se procesaron las muestras de calibración y validación.

Los resultados obtenidos en la predicción fueron comparados con los obtenidos aplicando el método oficial de la AOAC (este método se encuentra descrito en el Capítulo 2: “Materiales y Métodos”) y se presentan en la Tabla 3.12.

Muestras	Método oficial <sup>a,b</sup>		<i>N</i> -PLS <sup>a,b</sup>		PARAFAC <sup>a</sup>		MCR-ALS <sup>a</sup>	
	BEN	SOR	BEN	SOR	BEN	SOR	BEN	SOR
JN1	0,69	0,24	0,95	0,17	0,71	0,27	0,64	0,22
JN2	0,57	ND	1,10	ND	0,62	0,03	0,62	0,02
JN3	0,71	0,24	1,05	0,41	0,70	0,24	0,76	0,23
JN4	0,28	0,18	0,12	0,24	0,24	0,25	0,24	0,21

<sup>a</sup> Todos los resultados están expresados en g L<sup>-1</sup>.

<sup>b</sup> ND = no detectado.

Tabla 3.12: Resultados obtenidos para SOR y BEN al predecir cuatro muestras comerciales mediante tres métodos quimiométricos (*N*-PLS, PARAFAC y MCR-ALS), además del método oficial.

En la Tabla 3.12 puede observarse que los resultados obtenidos aplicando el método *N*-PLS corroboran lo ya discutido, el método no puede ser utilizado para este sistema.

Por otro lado, los valores obtenidos con PARAFAC y MCR-ALS son comparables con los obtenidos empleando el método oficial. Para probar esto, se realizó un análisis estadístico de comparación de medias entre métodos, concluyendo que no existen diferencias significativas entre los resultados obtenidos mediante los métodos quimiométricos y el oficial (todos los  $t_{\text{calculados}}$  fueron menores que los  $t_{\text{tabulados}}$  para  $\alpha = 0,05$  y 4 grados de libertad). Los resultados obtenidos al realizar el análisis estadístico son expuestos en la Tabla 3.13.



	SOR			BEN		
	Media	SD	$t_{calculado}$	Media	SD	$t_{calculado}$
Método oficial	0,165	0,012		0,562	0,039	
PARAFAC	0,197	0,012	2,26	0,567	0,049	0,25
MCR-ALS	0,170	0,010	0,42	0,565	0,050	0,09

Tabla 3.13: Resultados obtenidos aplicando el test de comparación de medias entre métodos, los cálculos se realizaron con las concentraciones expuestas en la Tabla 3.12. En todos los casos los grados de libertad fueron 4, el límite de confianza 95% y el valor tabulado para  $t = 2,79$ .

Los resultados obtenidos nos permiten proponer estas herramientas quimiométricas (PARAFAC y MCR-ALS) como una alternativa rápida, sencilla y eficiente para el análisis de los conservantes ácido sórbico y ácido benzoico en bebidas comerciales a base de jugos de frutas.

### **3.3 MÉTODO DESARROLLADO PARA LA CUANTIFICACIÓN DE COLORANTES EN JUGOS COMERCIALES, APLICANDO UN NUEVO MÉTODO DE CALIBRACIÓN MULTIVARIADA DE SEGUNDO ORDEN (BLLS) A DATOS ESPECTRALES EN GRADIENTE DE pH**

#### **3.3.1 INTRODUCCIÓN**

En la etapa final de la presente tesis, se planteó el problema de cuantificar colorantes en muestras complejas. Estas muestras fueron polvos artificiales utilizados para preparar jugos de fruta. Dependiendo de la marca y la fruta a imitar, algunas de estas bebidas comerciales pueden llegar a estar formuladas con hasta cuatro colorantes. A esto se suman las interferencias producidas por el resto de los componentes de la bebida deshidratada: ácido ascórbico, edulcorantes, etc. Para resolver el problema se utilizaron datos generados de la misma manera que para el análisis de sorbato y benzoato en jugos.

Como ya fuera comentado, la matriz de datos absorbancia/pH presenta un desafío especial para los algoritmos usados para datos de segundo orden, debido al hecho de que las especies presentan una correlación a lo largo de la dimensión pH, ya que la suma de las concentraciones individuales de la especies permanece constante. Esto produce una dependencia lineal en la dimensión pH, requiriendo la aplicación de restricciones para obtener una calibración y predicción satisfactoria. Para tratar este tipo de datos fueron desarrollados algoritmos específicos que permiten resolver esta dependencia lineal: PARAFAC (PARAFAC 2, PARALIND),<sup>44,155</sup> MCR-ALS, RBL (bilinearización de los residuos),<sup>156</sup> y una reciente modificación del método de cuadrados mínimos bilineales (BLLS).<sup>137</sup>

Como se demuestra en el presente trabajo, estos algoritmos no son aplicables cuando se tienen colinealidades extremas además de la dependencia en la dimensión pH. Por lo tanto, se ha explotado una alternativa: una modificación de BLLS que tiene en cuenta la presencia especies en equilibrio ácido-base. BLLS brinda una elevada eficiencia en el análisis de datos con severas dependencias en las dimensiones espectros y pH.<sup>137</sup>

Con respecto a los colorantes, debe mencionarse que el color y el sabor son características fundamentales para que los alimentos resulten atractivos. En el caso de las bebidas artificiales, los colorantes son los que brindan las propiedades organolépticas que poseen las frutas frescas. Esta es la razón de la utilización de aditivos químicos en la industria alimenticia. Los colorantes figuran entre los aditivos permitidos por los organismos de regulación. Algunos colorantes sintéticos son patogénicos, especialmente cuando se consumen en exceso, por lo que es necesario el control por parte de las industrias y las organismos de salud.<sup>157,158</sup>

La tartrazina (T) [(trisodio 5-hidroxi-1-(4-sulfofenil)-4-(4-sulfofenilazo) pirazol-3-carboxilato], el amarillo ocaso (A) (6-hidroxi-5-[(4-sulfofenil)azo]-2-naftaleno sulfónico disodio), el rojo 40 o rojo allura (R) (6-hidroxi-5-[(2-metoxi-5-metil-4-sulfofenil)azo]-2-naftalensulfónico disodio) y el índigo carmín (I) (5,5'-(2-(1,3-dihidro-3-oxo-2H-indazol-2-ilideno)-1,2-dihidro-3H-indol-3-ona) disulfonato), son los cuatro colorantes sintéticos que brindan los siguientes colores: amarillo (T y A), rojo (R) y azul (I), los que están presentes de manera individual o en diferentes combinaciones en los jugos analizados.

Numerosos métodos fueron desarrollados y publicados para la cuantificación de estos colorantes. En general, se trata de técnicas separativas debido a la complejidad de las muestras. La cromatografía líquida de alto rendimiento (CLAR) es la más recomendada para la cuantificación de mezclas de varios colorantes.<sup>159-162</sup> Otras técnicas utilizadas son: electroforesis capilar,<sup>163,164</sup> espectroscopía con aplicación de herramientas quimiométricas de primer orden,<sup>165,166</sup> polarografía con pulso diferencial,<sup>167</sup> voltimetría con onda cuadrada<sup>168</sup> y espectrometría de masas.<sup>169</sup>

### ***Comportamiento espectral de los analitos:***

La Figura 3.9 muestra los espectros de absorción de tartrazina (T), amarillo ocaso (A), índigo carmín (I) y rojo 40 (R) a pH = 5,0 y a pH = 12,0. A estos dos valores de pH se encuentran las formas ácidas y básicas respectivamente de todos los colorantes. Una de las principales dificultades que presenta el sistema es (como puede verse en la Figura 3.9) que existen extremas colinealidades en la dimensión espectral, a causa de que las formas ácido y base de cada especie no modifican su forma. Estos espectros se solapan entre sí, complicando el análisis por métodos como PARAFAC o MCR-ALS.

El espectro para un blanco de muestra (sin colorantes) también es expuesto en la Figura 3.9 a los pH ya mencionados para los otros analitos. Como puede verse, un intenso solapamiento ocurre entre los espectros de los analitos y la línea de base.

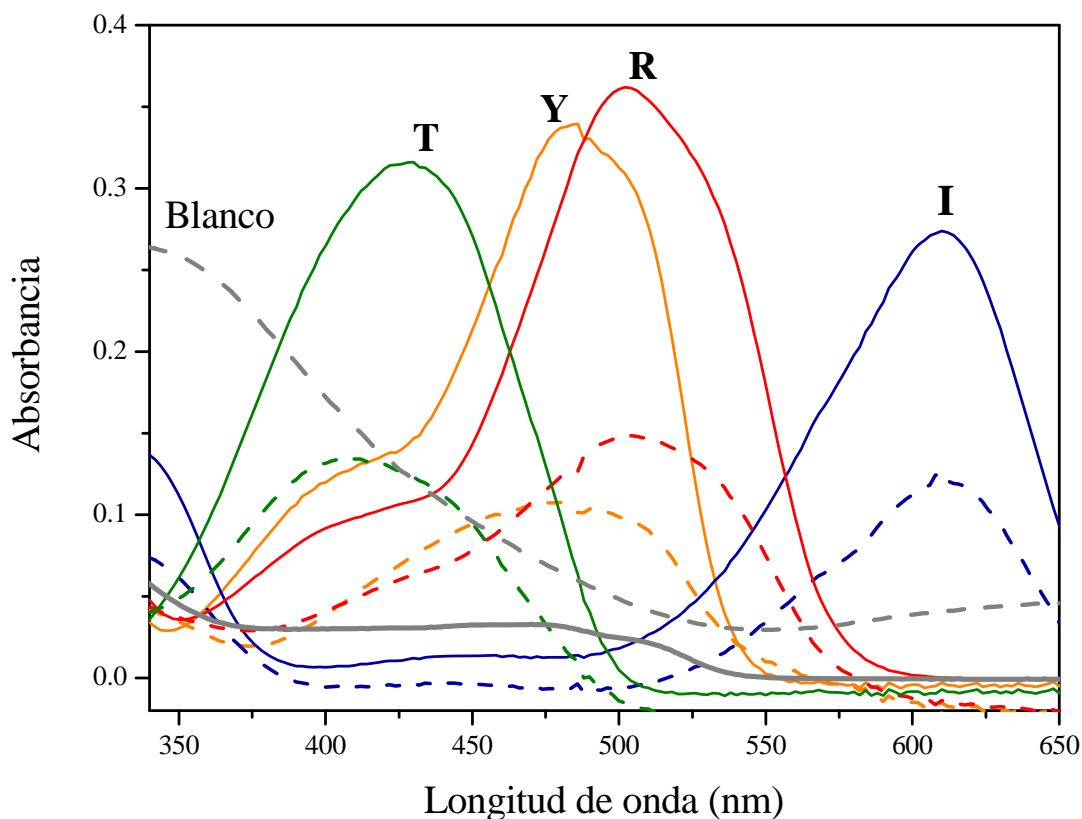


Figura 3.9: Espectros de absorción de los cuatro colorantes estudiados. Las líneas sólidas corresponden a los espectros de tartrazina (verde), Amarillo ocaso (naranja), rojo 40 (rojo) e índigo carmín (azul) (todos en concentración de  $8,0 \text{ mg L}^{-1}$ ), a pH ácido (5,0). Las líneas cortadas corresponden a los espectros a pH básico (12,0). Las líneas grises representan los espectros de una solución que contiene la matriz de la muestra a pH = 5,0 (línea sólida) y a pH = 12,0 (línea cortada).

Para generar los datos se implementó un sistema FIA simple (ver Capítulo 2, punto 2.2.3.2). La muestra en solución reguladora de fosfato (pH = 5,5) se introdujo al sistema, mediante la bomba peristáltica. Una vez en línea, se inyectó un bolo de  $50 \mu\text{L}$  de NaOH 0,1 M. Esto genera el doble gradiente de pH, cuyos valores extremos están comprendidos entre 5,5 y 12,8. Se registraron los espectros de absorción en un rango de 340 a 750 nm, cada 2 nm, y

cada 2 segundos. El tiempo requerido para el análisis de cada muestra es menor a 3 minutos. El número de espectros adquiridos para cada pico FIA es 76, por lo que la matriz de datos por cada muestra tiene dimensiones 206 x 76, es decir 15.656 puntos. El sistema fue optimizado considerando: caudal, pH de la solución reguladora de fosfato y del NaOH, tiempo de inyección y de adquisición de la señal, y longitud del reactor.

La Figura 3.10 muestra un gráfico tridimensional para la matriz de datos de una muestra (muestra 1 del juego de validación, la que contiene los cuatro analitos). La variación de la señal a lo largo del pico FIA puede ser apreciada como un decrecimiento de la absorbancia artificial (mezcla 1 del juego de validación) que contiene los cuatro analitos. en el centro de la dimensión tiempo-FIA.

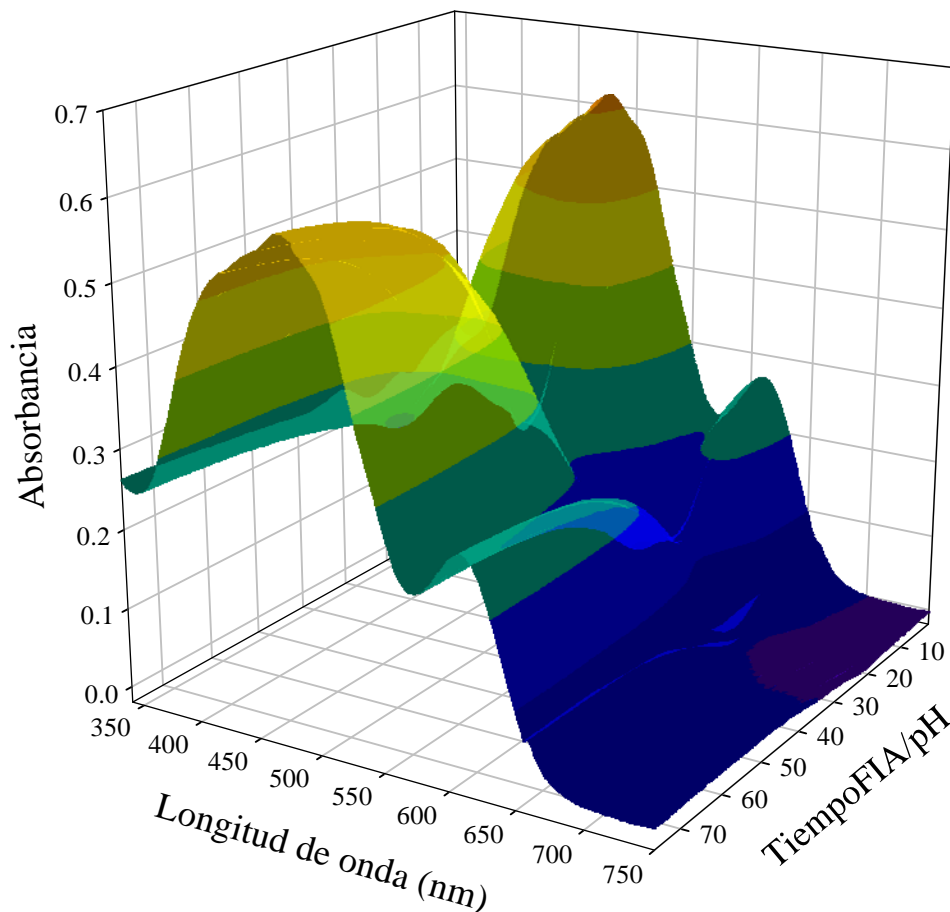


Figura 10: Gráfico tridimensional de la evolución espectral en función del pH para una muestra

### 3.3.2 CALIBRACIÓN

La calibración se llevó a cabo con las siguientes soluciones de patrones puros: se prepararon seis estándares para cada analito en concentraciones: 2,0; 4,0; 8,0; 12,0; 16,0 y 20,0 mg L<sup>-1</sup>, lo que constituye un juego de 24 soluciones.

Para la preparación de estos estándares se midió la cantidad necesaria de solución patrón (del colorante correspondiente para cada caso) de 4,00 g L<sup>-1</sup> y se llevó a volumen final en matraz aforado con solución reguladora de fosfato (pH = 5,5).

Se verificó la linealidad de la respuesta en el rango de concentraciones propuesto.

### 3.3.3 VALIDACIÓN

Se preparó un juego de 10 mezclas para validar el modelo multivariado. Las concentraciones de los analitos para estas mezclas fueron seleccionadas según un diseño de Plackett-Burman (ver Tabla 3.14) con concentraciones extremas de 4,0 y 8,0 mg L<sup>-1</sup> para todos los colorantes. Las mezclas fueron preparadas midiendo cantidades adecuadas de la solución patrón correspondiente y llevando a volumen final en matraz aforado con buffer fosfato (pH = 5,5). Las evoluciones espectrales en función del pH fueron medidas de manera aleatoria, después de la inyección alcalina en el sistema FIA.

#### *Procesamiento de los datos:*

Para la aplicación de BLLS, las matrices del juego de calibración (todas con un tamaño de 206 x 76) fueron primero vectorizadas para poder aplicar las expresiones matemáticas descritas en la teoría (Capítulo 1). La bilinealización residual fue aplicada a todas las muestras de validación a fin de explotar la ventaja de segundo orden.

El número de interferencias fue estimado incrementando el número de componentes hasta estabilizar los residuos de la predicción en un valor compatible con el nivel del ruido instrumental. Este último fue estimado mediante medidas de réplicas del blanco.

#### *Análisis de las mezclas del juego de validación:*

El uso de PARAFAC o MCR-ALS (con las correspondientes restricciones) no proporcionó resultados aceptables para el presente sistema. La causa más probable es el extremo solapamiento espectral que existe entre las formas ácidas y básicas de los analitos, lo que

introduce una fuente adicional de dependencia lineal, además de la que existe para la dimensión pH. Sin embargo, el análisis del juego de validación fue llevado a cabo satisfactoriamente mediante el uso del algoritmo BLLS. Para poder aplicar BLLS al presente problema, deben considerarse dos parámetros: 1) el número de componentes (analitos e interferencias) y 2) el número de especies por cada analito. El número de especies por analito fue optimizado estudiando los valores del residuo de la predicción en función del número de componentes para cada muestra en particular, modificando este número hasta que estos residuos fueron comparables con el nivel de ruido instrumental (calculado con las medidas de 10 de replicas del blanco). Los resultados fueron satisfactorios al considerar cuatro componentes y una especie por cada analito.

El número de componentes obtenido (4), coincide con el número de analitos presentes en las muestras de validación. Además, es comprensible que el número de especies por cada analito sea uno, debido a la elevada colinealidad que existe entre las formas ácidas y básicas.

Como se puede observar en la Figura 3.11 (A), para una de las muestras de validación, el modelo de cuatro componentes y una sola especie por analito muestra que los residuos de la predicción es comparable con el nivel del ruido instrumental. Utilizar un modelo más complejo solo produce un ligero decrecimiento en los residuos lo cual no justifica complejizar el sistema.

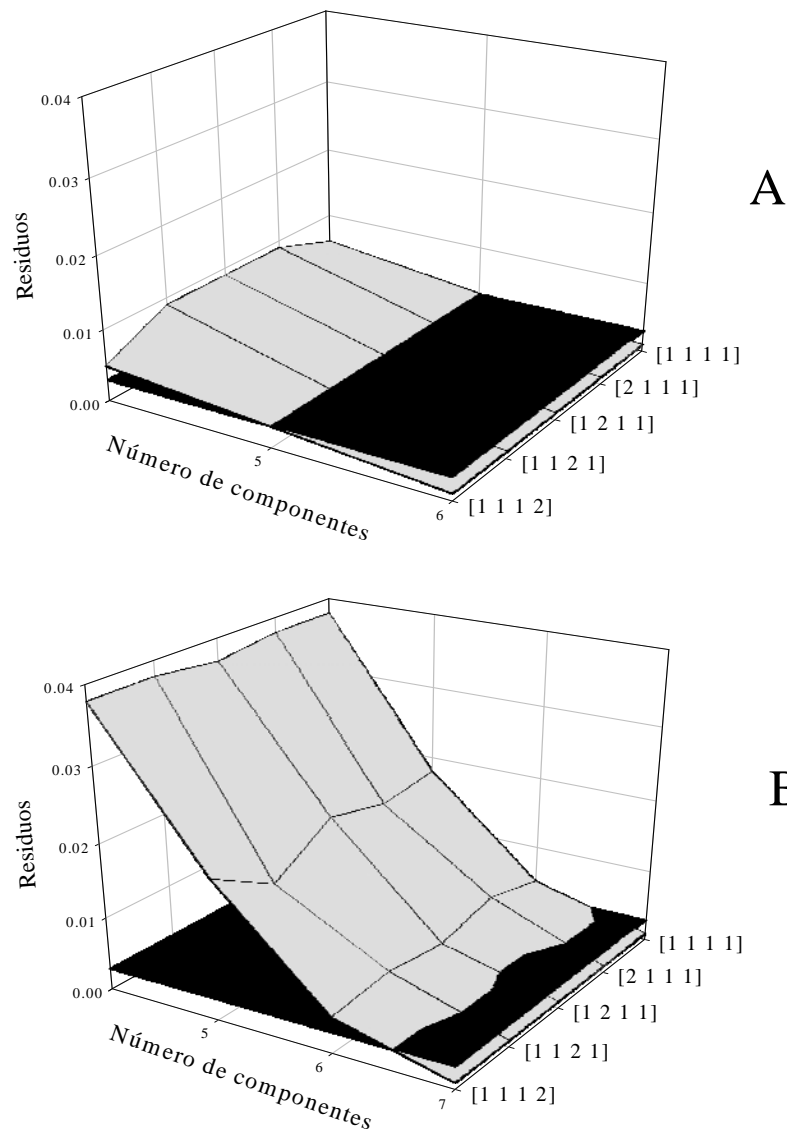


Figura 3.11: A) Cambios en los residuos de la predicción con BLLS para una muestra de validación en función del número de componentes, considerando cinco diferentes combinaciones de las especies en equilibrio para cada analito. La nomenclatura es  $[N_{s1} N_{s2} N_{s3} N_{s4}]$ , donde los números sucesivos corresponden a las especies T, A, R e I. El plano negro horizontal indica el nivel de ruido instrumental. B) Ídem pero para una muestra real.



Mediante la aplicación de BLLS a las muestras de validación se calcularon los perfiles espectrales de los analitos y se estimaron las concentraciones. Todos los perfiles obtenidos mediante el cálculo son coincidentes con los espectros de los estándares.

La predicción de las muestras sintéticas mediante un modelo BLLS de cuatro componentes brinda los resultados expuestos en la Tabla 3.14. Los porcentajes de recuperación entre 94,8 y 109,3 % pueden ser considerados como muy aceptables en vista de la complejidad de las muestras analizadas. Los mejores resultados fueron obtenidos para el índigo carmín, como era de esperar ya que el espectro de este analito muestra una forma significativamente diferente a los espectros de los otros tres.

Muestras	Amarillo ocazo (mg L <sup>-1</sup> )		Índigo carmín (mg L <sup>-1</sup> )		Rojo 40 (mg L <sup>-1</sup> )		Tartrazina (mg L <sup>-1</sup> )	
	Nominal	Pred.	Nominal	Pred.	Nominal	Pred.	Nominal	Pred.
1	4,0	3,8	8,0	8,4	8,0	8,7	8,0	8,4
2	8,0	7,5	8,0	8,5	8,0	8,7	4,0	4,7
3	8,0	7,5	4,0	4,8	8,0	8,7	4,0	4,4
4	8,0	7,5	4,0	4,3	4,0	4,9	8,0	8,2
5	4,0	3,7	8,0	8,1	4,0	4,3	4,0	4,1
6	4,0	3,7	4,0	4,0	8,0	8,3	8,0	8,1
7	8,0	7,5	8,0	8,6	4,0	4,7	8,0	8,5
8	8,0	7,8	4,0	3,9	4,0	4,3	4,0	4,2
9	6,0	5,9	6,0	6,0	6,0	6,3	6,0	6,3
10	6,0	5,8	6,0	6,1	6,0	6,2	6,0	6,2
Media del % de recuperación		94,8		104,7		109,3		105,8

Tabla 3.14: Predicciones de los cuatro colorantes en las mezclas de validación.

### 3.3.4 MUESTRAS REALES

Tres muestras reales fueron estudiadas: J1 (que contiene dos colorantes: T y A), J2 (que contiene tres colorantes: R, T y A) y J3 (conteniendo I, R, T y A).

Las muestras fueron preparadas pesando 24,0 g de polvo comercial y llevando a volumen final de 250,00 mL en matraz aforado, enrasando con solución reguladora de fosfato (pH = 5,5).

Posteriormente, cada muestra se sometió a diversas diluciones para que la concentración final coincidiera con el rango de concentraciones de calibración (esto fue realizado comparando las absorbancias máximas del espectro de la muestra con las magnitudes de las absorbancias de las mezclas de calibración). Las diluciones realizadas fueron las siguientes: J1: 1/10; J2: 2/10 y J3: 3/10.

Por otra parte, se realizaron adiciones estándar de los colorantes a estas muestras para realizar estudios de recuperación. De esta manera en cada muestra por separado se agregaron cada uno de los colorantes declarados, en tres niveles de concentración. Estos niveles fueron 2,0; 4,0 y 8,0 mg L<sup>-1</sup>. Este procedimiento originó seis muestras para J1, nueve para J2 y doce para J3 (ver Tabla 3.15).

Los espectros de todas estas soluciones fueron adquiridos en orden aleatorio, y en diferentes días que la calibración y las muestras de validación.

#### *Análisis de los resultados obtenidos para muestras reales*

Los cálculos para el análisis de las muestras reales fueron realizados de la misma manera que para las muestras sintéticas utilizando BLLS, pero usando seis componentes en lugar de cuatro. En este caso, el número de especies químicas para cada analito fue considerado uno, de acuerdo a los estudios previos realizados con las muestras de validación, pero la presencia de interferencias incrementa el número de componentes en dos. Fue necesario agregar dos componentes para disminuir los residuos de predicción para cada muestra, hasta llevarlos a nivel de ruido instrumental. En la Figura 3.11 (B) se muestra como implementando el modelo con seis componentes y una especie por analito los residuos de la predicción son equivalentes al nivel de ruido instrumental.

Muestra	Amarillo ocaso (mg L <sup>-1</sup> )		Indigo carmín (mg L <sup>-1</sup> )		Rojo 40 (mg L <sup>-1</sup> )		Tartrazina (mg L <sup>-1</sup> )	
	Nominal <sup>a</sup>	Pred.	Nominal <sup>a</sup>	Pred.	Nominal <sup>a</sup>	Pred.	Nominal <sup>a</sup>	Pred.
J1	–	3,8	–	0,2	–	0,3	–	6,6
J1-T2	–	4,2	–	0,2	–	0,1	8,4	8,2
J1-T4	–	3,4	–	0,2	–	0,2	10,4	10,2
J1-T8	–	3,4	–	-0,3	–	-0,3	14,4	15,5
J1-A2	5,7	5,8	–	0,3	–	0,3	–	5,8
J1-A4	7,7	7,3	–	0,4	–	0,4	–	5,9
J1-A8	11,7	12,7	–	0,8	–	1,1	–	7,3
J2	–	2,3	–	3,1	–	0,4	–	1,6
J2-T2	–	2,2	–	3,4	–	0,4	3,3	3,6
J2-T4	–	2,3	–	3,6	–	0,5	5,3	5,6
J2-T8	–	2,4	–	4,1	–	0,7	9,3	9,7
J2-A2	4,3	4,4	–	2,9	–	0,4	–	1,3
J2-A4	6,3	6,4	–	3,2	–	0,5	–	1,2
J2-A8	10,3	10,7	–	3,0	–	0,7	–	0,7
J2-R2	–	2,2	–	3,2	2,4	2,5	–	1,3
J2-R4	–	2,0	–	3,4	4,4	4,7	–	1,3
J2-R8	–	2,0	–	3,0	8,8	9,4	–	1,1
J3	–	3,1	–	0,1	–	0,7	–	3,8
J3-T2	–	3,1	–	0,3	–	0,9	6,0	5,9
J3-T4	–	3,2	–	0,4	–	1,0	8,0	8,6
J3-T8	–	3,1	–	0,4	–	1,0	12,0	11,5
J3-A2	5,1	5,4	–	0,1	–	0,5	–	3,3
J3-A4	7,1	7,6	–	0,4	–	0,8	–	4,8
J3-A8	11,1	11,7	–	0,0	–	0,8	–	3,7
J3-R2	–	3,0	–	0,0	2,8	2,8	–	3,2
J3-R4	–	3,0	–	0,0	4,8	5,1	–	3,4
J3-R8	–	2,9	–	0,0	8,8	9,6	–	3,4
J3-I2	–	3,3	2,2	2,1	–	0,5	–	3,5
J3-I4	–	3,6	4,2	4,4	–	0,6	–	3,9
J3-I8	–	3,5	8,2	9,2	–	0,9	–	3,9

<sup>a</sup> Las concentraciones nominales corresponden a la suma del valor promedio predicho mediante BLS y las cantidades agregadas en el proceso de adición estándar (ver texto).

Tabla 3.15: Resultados de las predicciones para tres muestras reales y adiciones realizadas en estas tres muestras.

La predicción con el modelo BLS proporciona perfiles espectrales y de pH para cada una de las especies, estos perfiles pueden verse en las Figuras 3.12 (A y B), para una muestra real. Comparando las Figuras 3.9 y 3.12 (A), puede observarse que los perfiles obtenidos para cada analito mediante la aplicación de BLS son iguales a los espectros de absorción de

los estándares. En el caso de los perfiles de las interferencias espectrales, estos corresponden a los residuos de predicción, y son una combinación lineal de los espectros reales. En este sentido, ellos no son estrictamente comparables con los que muestran los blancos de muestra expuestos en la Figura 3.9.

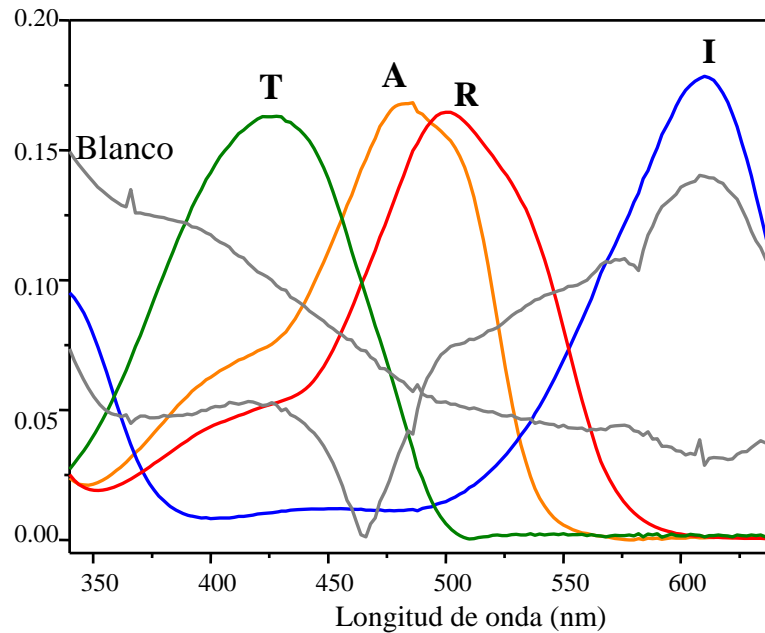
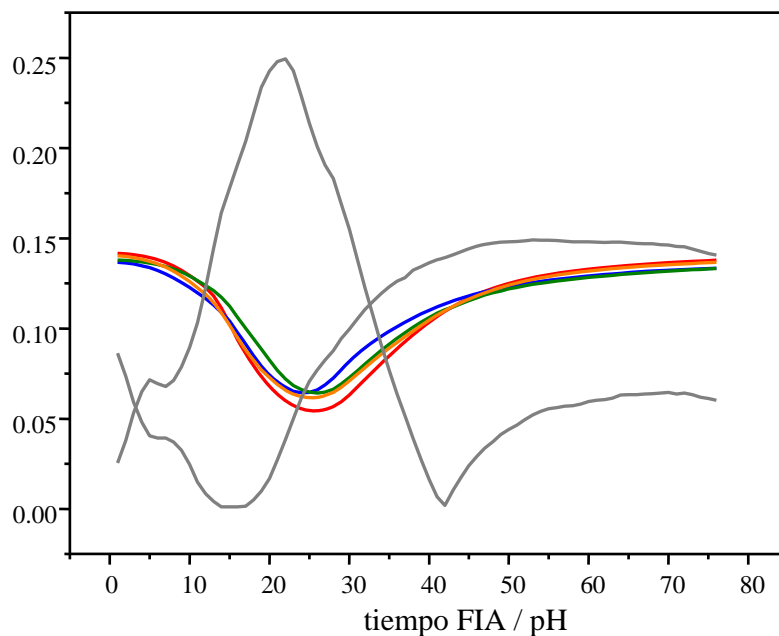
**A****B**

Figura 3.12 (A y B): Perfiles espectrales y de pH obtenidos mediante BLLS para una muestra real. Los analitos identificados son tartrazina (T), amarillo ocaso (A), rojo 40 (R) e índigo carmín (I). Las líneas grises corresponden a las interferencias.

BLLS permite obtener satisfactoriamente los perfiles correctos de los analitos para ser empleados en la predicción de las concentraciones de las muestras reales. Los perfiles de pH obtenidos se muestran en la Figura 3.12 (B). Es evidente la dificultad para el análisis del presente sistema: las formas de los perfiles de los analitos son semejantes a lo largo de la dimensión pH, produciendo una fuerte colinealidad en el modo espectral.

Como puede verse los resultados de la predicción para las muestras reales que se muestran en la Tabla 3.15 son excelentes para la mayoría de las muestras adicionadas.

En la Tabla 3.16 se muestran las probabilidades obtenidas al aplicar una prueba estadística  $t$  por parejas cuando se comparan las cantidades adicionadas y las predichas con el modelo BLLS. Como puede verse solo en dos casos se presentan diferencias significativas (el valor de  $p$  es menor de 0,05). Se realizó también, una regresión lineal entre los valores de concentración agregados y los predichos para el amarillo ocaso y la tartrazina (por ser los analitos para los que se disponen más datos) en las tres muestras reales adicionadas. La estimación de la ordenada al origen y la pendiente ( $\hat{a}$  y  $\hat{b}$  respectivamente) fueron comparados con valores de 0 y 1 utilizando un intervalo de confianza elíptico conjunto (EJCR), utilizando un ajuste por cuadrados mínimos (OLS). La Figura 3.13 muestra el gráfico EJCR para ambas sustancias. Las elipses contienen los puntos 0 y 1 en ambos casos. Esto indica que no existen errores sistemáticos constantes ni proporcionales. Por otro lado, el amarillo ocaso presenta mayor precisión en la estimación de las concentraciones considerando el tamaño de la elipse correspondiente.

Muestra	Amarillo ocaso	Indigo carmín <sup>a</sup>	Rojo 40 <sup>a</sup>	Tartrazina
J1	0,63	–	–	0,64
J2	0,18	–	0,15	0,01
J3	0,03	0,38	0,26	0,90

<sup>a</sup> No fueron realizadas adiciones

Tabla 3.16: Probabilidades obtenidas para el test  $t$  por parejas analizando las predicciones para las muestras reales.

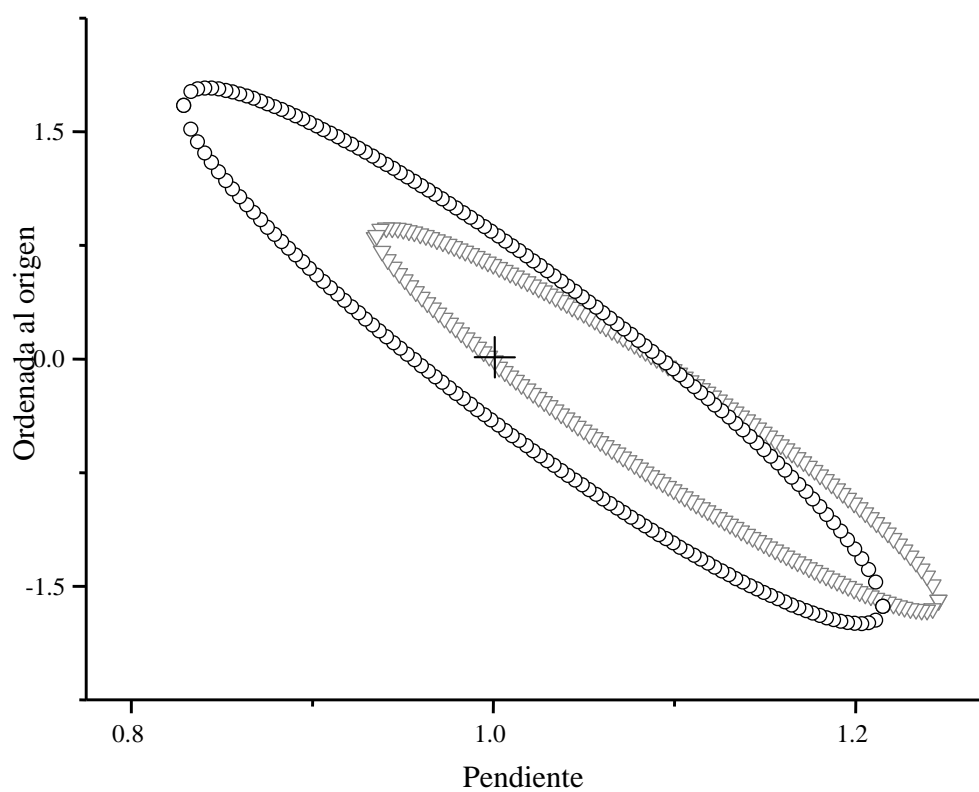


Figura 3.13: Intervalo elíptico de confianza para la pendiente ( $b$ ) y la ordenada al origen ( $a$ ) correspondiente a la recta de regresión entre las concentraciones nominales y las predichas para el amarillo ocaso (triángulos grises) y la tartrazina (círculos negros). La cruz negra indica el punto ( $a = 0, b = 1$ ).

## CAPÍTULO 4

### CONCLUSIONES

La aplicación de espectroscopia UV-Vis acoplada a estrategias de calibración multivariada, tanto de primer como de segundo orden, genera la posibilidad de desarrollar métodos analíticos con interesantes propiedades, entre las que podemos citar las siguientes: rapidez, simplicidad, precisión y exactitud. Estos métodos pueden ser aplicados a numerosos campos, tales como el análisis de aditivos en alimentos, contaminantes en muestras ambientales, fármacos en productos farmacéuticos y fluidos biológicos, entre otros. Concretamente, la explotación de las enormes potencialidades de las técnicas quimiométricas utilizadas en la presente Tesis, permitió la determinación de diferentes analitos (aditivos alimentarios) en mezclas complejas de alimentos.

La cuantificación simultánea de dos conservantes alimentarios: sorbato de potasio y benzoato de sodio, en jugos de fruta comerciales puede ser realizada utilizando respuestas instrumentales de primer y segundo orden. El tratamiento de las respuestas de primer orden, obtenidas por espectroscopía UV, usando un método basado en cálculos de señal neta del analito (HLA/GO), permite implementar la ventaja de primer orden (detección de muestras anómalas u *outliers*) y realizar los estudios de la función error indicador (EIF) en una ventana móvil y del gráfico de regresión NAS para cada muestra en particular. Esto último posibilita cuantificar muestras desconocidas evitando zonas de solapamiento espectral del analito con las interferencias no calibradas, o localizar zonas espectrales donde dichas interferencias se minimizan. La cuantificación de conservantes presentes en jugos concentrados mediante el método HLA/GO brinda excelentes resultados comparables a los obtenidos por métodos basados en CLAR o el método oficial de la AOAC, con la ventaja de ser notablemente más sencillo, rápido, y económico.

Por otra parte, las respuestas de segundo orden obtenidas mediante un sistema FIA con doble gradiente de pH, analizadas mediante los modelos PARAFAC y MCR-ALS, permite aprovechar la denominada ventaja de segundo orden, no siendo necesario modelar las interferencias durante la calibración, como sucede al usar datos de primer orden.

Cuando se trabaja con datos como los antes mencionados, al emplear PARAFAC y MCR-ALS, es necesario aplicar restricciones de no negatividad y de cierre, e introducir información inicial sobre el sistema en estudio (espectros de analitos e interferencias y perfiles de pH). Este hecho suma complejidad al análisis.

Los resultados obtenidos por aplicación de PARAFAC y MCR-ALS son comparables con los obtenidos empleando el método oficial. A su vez, PARAFAC presenta mejor rendimiento que MCR-ALS, lo que indica que no existen desviaciones serias de la trilinealidad en los datos usados, condición requerida por la primera de estas técnicas.

La cuantificación de los colorantes tartrazina, amarillo ocaso, índigo carmín y rojo 40 en polvos artificiales para preparar bebidas, se puede realizar usando datos de segundo orden, pero procesados mediante el método BLS con una modificación en el algoritmo de cálculo que contempla a más de una especie por analito. Este método, además de aprovechar la ventaja de segundo orden, permite resolver: a) problemas de severo solapamiento espectral entre los analitos y las interferencias de la matriz, b) el inconveniente de la dependencia lineal que existe entre las concentraciones de las especies en función del pH, y c) las colinealidades espectrales de las especies ácidas y básicas para todos los analitos. Este modelo resultó ser robusto para cuantificar simultáneamente este tipo de analitos, siendo su aplicación mucho más simple y rápida que la de los modelos MCR-ALS y PARAFAC.

Todos los métodos aquí desarrollados presentan las ventajas de ser sencillos, rápidos, de bajo costo y no requerir pre-procesamiento de las muestras. El uso de datos de segundo orden, aunque requiere de una implementación algo más compleja que los de primer orden, presenta mejores posibilidades de análisis de los analitos de interés en presencia de interferencias no calibradas. Las características que presentan los analitos y las matrices estudiadas, permiten suponer que es posible el desarrollo de métodos analíticos, como los que aquí se presentan, para una gran variedad de sistemas analíticos, entre ellos los constituidos por contaminantes en muestras ambientales.



## CAPÍTULO 5

### RESUMEN

Con el objetivo central de desarrollar nuevos métodos analíticos, resultantes de la aplicación de técnicas espectroscópicas acopladas a métodos quimiométricos, se estudiaron sistemas analíticos constituidos por aditivos alimentarios en diferentes matrices. Se buscó aumentar la selectividad de la espectroscopia UV haciendo uso de señales de primer y segundo orden, tratadas con nuevos modelos quimiométricos.

#### *Problemas analíticos estudiados*

1) Conservantes (sorbato de potasio y benzoato de sodio) presentes en jugos de frutas comerciales. Los jugos estudiados fueron bebidas analcohólicas preparadas a base de jugos de frutas frescas exprimidas (naranja, pomelo y limón), concentrados, de elaboración regional. La cuantificación simultánea se planteó a través del uso de señales instrumentales de primer y segundo orden. Para hacer uso de espectros de primer orden, se usó un modelo basado en el estudio de la señal neta del analito (NAS), concretamente el método de análisis híbrido lineal sugerido por Goicoechea y Olivieri (HLA/GO). En este sistema, el espectro de absorción de la matriz del jugo se solapa con los espectros de los conservantes, y además esta matriz presenta grandes variaciones de acuerdo al tipo de fruta, la variedad, la estación del año, el estado de maduración, etc. Esto hizo necesario modelar las interferencias durante la etapa de calibración. Para esto se agregaron jugos de frutas (naranja, pomelo y limón, al 20 %) a las mezclas de calibración.

Los resultados obtenidos mediante el método HLA/GO fueron comparados con los obtenidos aplicando el método de los cuadrados mínimos parciales (PLS). Se observó que los resultados de HLA/GO fueron notablemente mejores que los de PLS, por lo que este último método fue dejado de lado.

Por otra parte, para estudiar la exactitud se analizaron dos muestras comerciales aplicando los métodos HLA/GO y el oficial sugerido por la AOAC. Al comparar los resultados, pudo concluirse que no existen diferencias significativas entre los dos métodos.

El método desarrollado tiene las ventajas de ser rápido, sencillo, económico, no requiere el agregado de reactivos, solo debe realizarse una dilución de la muestra (1/100) en ácido clorhídrico ( $5,0 \times 10^{-4} \text{ mol L}^{-1}$ ); frente al método oficial que resulta laborioso, requiere largos tiempos de preprocesamiento, ya que consta de una primera etapa de destilación por arrastre con vapor, seguida del agregado de diferentes reactivos.

Una cosa para remarcar es que mediante la aplicación de HLA/GO puede aprovecharse la ventaja de primer orden (detección de muestras anómalas u *outliers*), y además seleccionar la ventana de trabajo estudiando la función error indicador (EIF). Aplicando la selección del rango de longitudes de onda de trabajo, se logró cuantificar satisfactoriamente ambos conservantes en jugos de frutas con matrices significativamente diferentes a las matrices con que se calibró (pera, manzana, zanahoria, mandarina y tomate).

Por otra parte, y continuando el análisis del sistema antes descrito, se desarrolló un nuevo método, pero utilizando respuestas de segundo orden. Para obtener dichas respuestas se diseñó sistema de análisis en flujo en el que se genera un doble gradiente de pH. Mientras la muestra disuelta en solución reguladora ácida (pH = 2,0) circula por el sistema, se inyecta un bolo alcalino (pH = 12,0). De esta manera, se genera un gradiente de concentración de la solución alcalina en los extremos del bolo, lo que a su vez provoca el doble gradiente de pH. Las variaciones espectrales originadas en el seno de este gradiente se registran con un espectrofotómetro con arreglo de diodos entre 200 y 300 nm, cada 1,5 segundos. Así se obtiene una matriz de datos por cada muestra. Las matrices de calibración obtenidas por aplicación de la metodología descrita, se ordenaron en un arreglo de tres vías, al que se le aplicaron dos métodos quimiométricos: análisis paralelo de factores (PARAFAC) y resolución multivariada de curvas seguida de cuadrados mínimos alternados (MCR-ALS).

Estos dos métodos permiten implementar la ventaja de segundo orden, gracias a lo cual no fue necesario modelar las interferencias durante la calibración. Los resultados obtenidos para las predicciones de muestras comerciales y artificiales fueron excelentes, con errores relativos de predicción del 8,7 % para BEN y 2,5 % para SOR.

Los resultados fueron comparados con los obtenidos aplicando el método oficial mediante una prueba estadística de comparación de medias, con la que se llegó a la conclusión de que no existen diferencias significativas entre los métodos.

2) Colorantes (tartrazina, índigo carmín, amarillo ocaso y rojo 40 o rojo allura) presentes en polvos comerciales para preparar jugos artificiales bebibles con sabor a frutas. Para analizar simultáneamente estos cuatro analitos en este sistema, se utilizaron respuestas de segundo orden obtenidas mediante una metodología similar a la empleada anteriormente, aunque con variaciones en los tipos de soluciones reguladoras usadas para conseguir los cambios de pH que más favorezcan las variaciones espectrales.

Este sistema presentó serias dificultades analíticas debidas a: a) problemas de severo solapamiento espectral entre los analitos y las interferencias de la matriz, b) el inconveniente de la dependencia lineal que existe entre las concentraciones de las especies en función del pH, y c) las colinealidades espectrales de las especies ácidas y básicas para todos los analitos. Esto pudo resolverse mediante la aplicación del método de cuadrados mínimos bilineales (BLLS) con una modificación en el algoritmo de cálculo que contempla a más de una especie por analito. Se obtuvieron excelentes resultados, con porcentajes de recuperación del 94,8; 104,7; 109,3 y 105,3 % para amarillo ocaso, índigo carmín, rojo 40 y tartrazina respectivamente. Por otra parte, para investigar la presencia de errores sistemáticos constantes y proporcionales se realizó un estudio que consiste en realizar una gráfica de regresión entre concentraciones nominales y predichas, y contraste de los valores de ordenada al origen y pendiente con los valores 0 y 1 (teóricos), respectivamente. El estudio se hizo a través de la construcción del intervalo elíptico conjunto, y mostró la no existencia de errores sistemáticos en el método desarrollado.

Se concluye que todos los métodos desarrollados presentan las ventajas de ser sencillos, rápidos, de bajo costo y no requerir pre-procesamiento de las muestras. El uso de datos de segundo orden, aunque requiere de una implementación algo más compleja que los de primer orden, presenta mejores posibilidades de análisis de los analitos de interés en presencia de interferencias no calibradas. Las características que presentan los analitos y las matrices estudiadas, permiten suponer que es posible el desarrollo de métodos analíticos, como los que aquí se presentan, para una gran variedad de sistemas analíticos, entre ellos los constituidos por contaminantes en muestras ambientales.

## BIBLIOGRAFÍA

- 1 Norman W. Desrosier. (1984). "Conservación de Alimentos". Ed. C.E.C.S.A. México. Decimotercera edición.
- 2 "VIII Congreso Latinoamericano de Química. Código Latinoamericano de Alimentos". (1962). Quinta edición.
- 3 Código Alimentario Argentino. (1992). Texto de Actualización Acumulado ley N° 18.284 Decreto N° 2126/75. Marzochi Ediciones. Buenos Aires. Tomo II: Capítulo XII; Bebidas hídricas y agua; y agua gasificada, pag.:461-545. Capítulo XVIII; Aditivos alimentarios, pag.: 737-851.
- 4 Código Bromatológico de la Provincia de Santa Fe. Ley 2998. (1941)
- 5 Madrid, A. (1992). "Los aditivos en los alimentos". Ediciones A. Madrid, Mundi Prensa Libros, S.A., Madrid.
- 6 Thomas, E.V.; Haaland, D.M. (1988). "Partial least-squares methods for spectral analyses. 1. Relation to other quantitative calibration methods and the extraction of qualitative information". *Anal. Chem.*, 60: 1193-1202.
- 7 Booksh, K.S.; Kowalski, B.R. (1994). "Theory of analytical chemistry". *Anal. Chem.*, 66: 782A-791A.
- 8 Strang, G. (1980). "Álgebra Lineal y sus Aplicaciones". Addison-Wesley Iberoamericana. Wilmington.

- 9 Martens, H.; Naes, T. (1989). "Multivariate Calibration". John Wiley & Sons. Chichester.
- 10 Lavine, B. K. (1998). "Chemometrics". *Anal. Chem.*, 70: 209R-228R.
- 11 Toasaksiri, S.; Massart, D.L.; Vander-Heyden Y. (2000). "Study of method validation criteria in a capillary electrophoresis method for the separation of non-steroidal anti-inflammatory drugs". *Anal.Chim. Acta.* , 416: 29-42.
- 12 Haaland, D.M.; Han, L.; 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.
- 13 Galeano Diaz, T.; Guiberteau, A.; Ortiz Burguillos, J.M.; 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.
- 14 Brown, S.D. (1991). "Rapid parameter estimation with incomplete chemical calibration models". *Chemom. Intell. Lab. Syst.*, 10: 87-105.
- 15 Faber, N.M. (2000). "Exact presentation of multivariate calibration model as univariate calibration graph". *Chemom. Intell. Lab. Syst.*, 50: 107-114.
- 16 Thomas, E.V. (1994). "A primer on multivariate calibration". *Anal. Chem.*, 66: 795 A-804 A.
- 17 Damiani, P.; Escandar, G.; Olivieri, A.; Goicoechea, H. (2005). "Multivariate calibración: a powerful tool in pharmaceutical analysis". *Curr. Pharm. Anal.*, 1: 145-154.
- 18 Lorber, A.; Kowalski, B.R. (1990). "Alternatives to cross-validatory estimation of the number of factors in multivariate calibration". *Appl. Spectrosc.*, 44: 1464-1470.
- 19 Marbach, R.; Heise, H.M. (1990). "Calibration modelling by partial least-squares and principal component regression and its optimization using an improved leverage correction for prediction testing". *Chemom. Intell. Lab. Syst.*, 9: 45-63.

- 20 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.
- 21 Wang, Z.; Luo, G.O.; Zhou, G.; Yao, C. (1989). "Application of PLS [partial least-squares] method in simultaneous spectrophotometric determination of trace metal ions". *Fenxi. Huaxue*, 17: 317-320.
- 22 Martens, H.; Izquierdo, L.; Thomassen, M.; Martens, M. (1986). "Partial least-squares regression on design variables as an alternative to analysis of variance". *Anal. Chim. Acta.*, 19: 133-148.
- 23 Lindberg, W.; Ohman, J.; Wold, S. (1986). "Multivariate resolution of overlapped peaks in liquid chromatography using diode-array detection". *Anal. Chem.* 58: 299-303.
- 24 Berger, A.J., Koo, T., Itzkan, I. y Feld, M.S. (1998). "An enhanced algorithm for linear multivariate calibration". *Anal. Chem.*, 70: 623-627.
- 25 Goicoechea, H.C.; Olivieri, A.C. (1999). "Wavelength selection by net analyte signals calculated with multivariate factor-based hybrid linear analysis (HLA). A theoretical and experimental comparison with partial least-squares (PLS)". *Analyst*, 124: 725-731.
- 26 Sena, M.M; Chaudhry, Z.F; Collins, C.H; Poppi, R.J. (2004). "Direct determination of diclofenac in pharmaceutical formulations containing B vitamins by using UV spectrophotometry and partial least squares regression". *J. Pharm. Biomed. Anal.*, 36: 743-749.
- 27 Munoz de la Peña, A.; Espinosa Mansilla, A.; Valenzuela, M.I.A; Goicoechea, H.C.; 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.
- 28 Korany, M.A.; Elsayed, M.A.; Mahgoub, H.; Korany, E. (1990). "Computer-assisted spectrophotometry: multicomponent analysis with a discrete Fourier transform". *Talanta*, 37: 1183-1188.
- 29 Liang, Y.; Kvalheim, O.M.; Keller, H.R.; Massart, D.L.; Kiechle, P.; Erni, F. (1992). "Heuristic involving latent projections: resolving two-way multicomponent data. Detection and resolution of minor constituents". *Anal. Chem.*, 64: 946-953.

- 30 Osten, D.W.; Kowalski, B.R. (1984). "Multivariate curve resolution in liquid chromatography". *Anal. Chem.*, 56: 991-995.
- 31 Maurí Aucejo, A.R.; Llobat Estellés, M.; Marín Saez, R.; San Martín Cigés, M.D.; Alvarez Alonso, C. (1993). "Apparent content curves: new analytical applications". *Fresenius J. Anal. Chem.*, 346: 888-895.
- 32 Jochum, C.; Jochum, P.; Kowalski, B.R. (1981). "Error propagation and optimal performance in multicomponent analysis". *Anal. Chem.*, 53: 85-92.
- 33 Sanchez, E.; Kowalski, B.R. (1986). "Generalized Rank Annihilation Factor Analysis". *Anal. Chem.*, 58: 496-499.
- 34 Lorber, A. (1984). "Validation of hypothesis on a data matrix by target factor analysis". *Anal. Chem.*, 56: 1004-1010.
- 35 Hruschka, W.R.; Norris, K. (1982). "Least squares curve fitting of near infrared spectra predicts protein and moisture content of ground wheat". *Appl. Spectrosc.*, 36: 261-265.
- 36 Gemperline, P.J.; Long, J.R.; Gregoriou, V.G. (1991). "Non-linear multivariate calibration using principal components regression and artificial neural networks". *Anal. Chem.*, 63: 2313-2323.
- 37 Shaffer, R.E.; Small, G.W. (1997). "Learning optimisation from nature. Genetic algorithms and simulated annealing". *Anal. Chem.*, 69: 236A-242A.
- 38 Despagne, F.; Massart, D.L. (1998). "Neural networks in multivariate calibration". *Analyst*, 123: 157R-178R.
- 39 Carroll, J.D.; Chang, J. (1970). "Analysis of individual differences in multidimensional scaling via an N-way generalization of "Eckart-Young" decomposition". *Psychometrika*, 35: 283-319.
- 40 Harshman, R.A. (1970). "Foundations of the PARAFAC procedure: model and conditions for an explanatory multi-mode factor analysis". *UCLA Work. Pap. Phon.*, 16: 1-84.

- 41 Wise, B.M.; Gallagher, N.B.; Martin, E.B. (2001). "Application of PARAFAC2 to fault detection and diagnosis in semiconductor etch". *J. Chemom.*, 15: 285-298.
- 42 Paatero, P. (1994). "Positive matrix factorization: a non negative factor model with optimal utilization of error estimates of data values". *Sci. Total. Environ.*, 5: 111-116.
- 43 Paatero, P. (1997). "A weighted non-negative least squares algorithm for three-way "PARAFAC" factor analysis". *Chemom. Intell. Lab. System*, 38: 223-242.
- 44 Bro, R. (1998). "Multi-way analysis in the food industry. Models, algorithms, and applications". PhD thesis, University of Amsterdam (NL).
- 45 Gerritsen, M.J.P.; Tanis, H.; Vandeginste, B.G.M.; Kateman, G. (1992). "Generalized rank annihilation factor analysis, iterative target transformation factor analysis, and residual bilinearization for the quantitative analysis of data from liquid chromatography with photodiode array detection". *Anal. Chem.*, 64: 2042-2056.
- 46 Faber, N.M. (2001). "The price paid for the second-order advantage when using the generalized rank annihilation method". *J. Chemom.*, 15: 743-748.
- 47 Gimeno, R.A.; Comas, E.; Marce, R.M.; Ferre, J.; Rius, F.X.; Borrull, F. (2003). "Second-order bilinear calibration for determining polycyclic aromatic compounds in marine sediments by solvent extraction and liquid chromatography with diode-array detection". *Anal. Chim. Acta.*, 498: 47-53.
- 48 Keinanen, M.M.; Korhonen, L.K.; Martikainen, P.J.; Vartiainen, T.; Miettinen, I.T.; Lehtola, M.J.; Nenonen, K.; Pajunen, H.; Kontro, M.H. (2003). "Gas chromatographic-mass spectrometric detection of 2- and 3-hydroxy fatty acids as methyl esters from soil, sediment and biofilm". *J. Chromatogr. B: Anal. Technol. Biomed. Life Sci.*, 783: 443-451.
- 49 Faber, N.M. (2002). "Towards a rehabilitation of the generalized rank annihilation method (GRAM)". *Anal. Bioanal. Chem.*, 372: 683-687.
- 50 Garrido Frenich, A.; Picon Zamora, D.; Martínez Galera, M.; Martínez Vidal, J.L. (2003). "Application of GRAM and TLD to the resolution and quantitation of real complex multicomponent mixtures by fluorescence spectroscopy". *Anal. Bioanal. Chem.*, 375: 974-980.



- 51 Xie, Y.L.; Baeza Baeza, J.J.; Ramis Ramos, G. (1996). "Second-order tensorial calibration for kinetic spectrophotometric determination". *Chemom. Intell. Lab. Syst.*, 32: 215-232.
- 52 Bezemer, E.; Rutan, S.C. (2002). "Three-way alternating least squares using three-dimensional tensors in MATLAB". *Chemom. Intell. Lab. Syst.*, 60: 239-251.
- 53 De Juan, A.; Tauler, R. (2001). "Comparison of three-way resolution methods for non-trilinear chemical data sets". *J. Chemom.*, 15: 749-772.
- 54 Xie, Y.L.; Hopke, P.K.; Paatero, P. (1998). "Positive matrix factorization applied to a curve resolution problem". *J. Chemom.*, 12: 357-364.
- 55 Booksh, K.S.; Kowalski, B.R. (1994). "Comments on the DATA ANALYSIS (DATAN) algorithm and rank annihilation factor analysis for the analysis of correlated spectral data". *J. Chemom.*, 8: 287-292.
- 56 Hart, S.J.; Hall, G.J.; Kenny, J.E. (2002). "A laser-induced fluorescence dual-fibre optic array detector applied to the rapid HPLC separation of polycyclic aromatic hydrocarbons". *Anal. Bioanal. Chem.*, 372: 205-215.
- 57 Linder, Marie; Sundberg, Rolf. (1998). "Second-order calibration: bilinear least squares regression and a simple alternative". *Chemom. Intell. Lab. Syst.*, 42: 159-178.
- 58 Faber, N. M.; Ferré, J.; Boqué, R.; Kalivas, J.H. (2002). "Second-order bilinear calibration: the effects of vectorising the data matrices of the calibration set". *Chemom. Intell. Lab. Syst.*, 63: 107-116.
- 59 Damiani, P.; Nepote, A.; Bearzotti, M.; Olivieri, A.C. (2004). "A test field for the second-order advantage in bilinear least-squares and parallel factor analyses: fluorescence determination of ciprofloxacin in human urine". *Anal. Chem.*, 76: 2798-2809.
- 60 Linder, M.; Sundberg, R. (2002). "Precision of prediction in second-order calibration, with focus on bilinear regression methods". *J. Chemom.*, 16: 12-27.
- 61 Saurina, J. (2000). "Analytical application of pH-gradients in flow-injection analysis and related techniques". *Rev. Anal. Chem.*, 19: 157-178.

- 62 Martínez, T.; Lartigue, J.; Ávila Pérez, P.; Zarazua, G.; Cabrera, L.; Tejada, S.; Ramírez, A. (2004). "Determination of lead in blood by TXRF and its correlation to environmental lead". *Nucl. Instrum. Method Phys. Res, Sect. B.*, 213: 584-589.
- 63 Tauler, R.; Izquierdo Ridorsa, A.; Casassas, E. (1993). "Simultaneous analysis of several spectroscopic titrations with self-modelling curve resolution". *Chemom. Intell. Lab. Syst.*, 18: 293-300.
- 64 Saurina, J.; Leal, C.; Compano, R.; Granados, M.; Tauler, R.; Prat, M.D. (2000). "Determination of triphenyltin in sea water by excitation-emission matrix fluorescence and multivariate curve resolution". *Anal. Chim. Acta.*, 409: 237-245.
- 65 Thomas, E.V.; Haaland, D.M. (1988). "Partial least-squares methods for spectral analyses. 2. Application to simulated and glass spectral data". *Anal. Chem.*, 60: 1202-1208.
- 66 Thomas, E.V.; Haaland, D.M. (1988). "Partial least-squares methods for spectral analyses. 3. Quantitative infrared analysis of borophosphosilicate films using multivariate statistical methods". *Anal. Chem.*, 60: 1208-1217.
- 67 López de Alba, P.L.; López Martínez, L.; Amador Hernandez, J. (1995). "Simultaneous determination of uranium(VI) and thorium(IV) ions with carminic acid by the partial least-squares (PLS-1 and PLS-2) methods". *J. Radioanal. Nucl. Chem.*, 200: 397-415.
- 68 Bautista, R.D.; Jimenez, A.I.; Jimenez, F.; Arias, J.J. (1996). "Resolution of ternary and quaternary mixtures of drugs in pharmaceutical preparations by use of spectrophotometric data in conjunction with PLS-1 and PLS-2 data processing methods". *Anal. Lett.*, 29: 2645-2665.
- 69 Talski, G. (1994). "Derivate spectrophotometry. Low and higher order". Editorial VCH, Cambridge, UK.
- 70 Kellner, R.; Mermet, J.M.; Otto, M.; Widner, H.M. (1998). "Analytical Chemistry". Ed. Wiley-Vch Verlag GmbH, 69469 Weinheim, páginas 776-777.
- 71 Lorber, A. (1986). "Error propagation and figures of merit for [multi-component] quantification by solving matrix equations". *Anal. Chem.*, 58: 1167-1172.

- 72 Ferré, J.; Faber, N.M. (2003). "Net analyte signal calculation for multivariate calibration". *Chemom. Intell. Lab. Syst.*, 69: 123-136.
- 73 Faber, N.M. (1998). "Mean centering and computation of scalar net analyte signal in multivariate calibration". *J. Chemom.*, 12: 405-409.
- 74 Faber, N.M. (1998). "Efficient computation of net analyte signal vector in inverse multivariate calibration models". *Anal. Chem.*, 70: 5108-5110.
- 75 Faber, N.M.; Lorber, A.; Kowalski, B.R. (1997). "Analytical figures of merit for tensorial calibration". *J. Chemom.*, 11: 419-461.
- 76 Ferré, J.; Rius, F.X. (1998). "Detection and correction of biased results of individual analytes in multicomponent spectroscopic analysis". *Anal. Chem.*, 70: 1999-2007.
- 77 Kalivas, J.K.; Lang, P.M. (1996). "Inter-relationships between sensitivity and selectivity measures for spectroscopic analysis". *Chemom. Intell. Lab. Syst.*, 32: 135-149.
- 78 Faber, N.M. (1999). "Notes on two competing definitions of multivariate sensitivity". *Anal. Chem. Acta*, 381: 103-109.
- 79 Xu, L.; Schechter, I. (1996). "Wavelength selection for simultaneous spectroscopic analysis. Experimental and theoretical study". *Anal. Chem.*, 68(14): 2392-2400.
- 80 Espinosa Mansilla, A.; Acedo Valenzuela, M.I.; Muñoz de la Peña, A.; Salinas, F.; Canada Canada, F. (2001). "Comparative study of partial least-squares and a modification of hybrid linear analysis calibration in the simultaneous spectrophotometric determination of rifampicin, pyrazinamide and isoniazid". *Anal. Chim. Acta*, 427: 129-136.
- 81 Goicoechea, H.C.; 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: 793-800.
- 82 Lorber, A.; Faber, K.; Kowalski, B.R. (1997). "Net analyte signal calculation in multivariate calibration". *Anal. Chem.*, 69: 1620-1626.

83 Goicoechea, H.C.; Olivieri, A. C. (2001). "A comparison of orthogonal signal correction and net analyte preprocessing methods. Theoretical and experimental study". *Chemom. Intell. Lab. Syst.*, 56: 73-81.

84 Olivieri, A.C. (2002). "A simple approach to uncertainty propagation in preprocessed multivariate calibration". *J. Chemom.*, 16: 207-217.

85 Goicoechea, H.C.; Olivieri, A.C. (2002). "Chemometric assisted simultaneous determination of four component nasal solutions with a reduced number of calibration samples", *Anal. Chim. Acta*, 453: 289-300.

86 Goicoechea, H.; Olivieri, A. (2001). "Sustained prediction ability of net analyte preprocessing method using reduced calibration sets. Theoretical and experimental study involving the spectrophotometric determination of multicomponent mixtures". *Analyst*, 126: 1105-1113.

87 Xu, L. y Schechter, I. (1997). "A calibration method free of optimum factor number selection for automated multivariate analysis. Experimental and theoretical study". *Anal. Chem.*, 69: 3722-3730.

88 Green, J.M. (1996). "A practical guide to analytical method validation". *Anal. Chem.*, 68: 305A-309A.

89 Goicoechea, H.C.; Olivieri, A.C. (1999). "Enhanced synchronous spectrofluorimetric determination of tetracycline in blood serum by chemometric analysis. Comparison of partial least-squares and hybrid linear analysis calibrations". *Anal. Chem.*, 71: 4361-4368.

90 Espinosa-Mansilla, A.; Durán-Merás, I.; Galián, R. (2001). "Simultaneous fluorimetric determination of pteridine derivatives: comparison between synchronous, partial least-squares, and hybrid linear analysis methods". *Appl. Spectrosc.*, 55: 701-707.

91 Goicoechea, H.C.; Olivieri, A.C. (2000). "MULTIVAR. A program for multivariate calibration incorporating net analyte signal calculations". *Trends Anal. Chem.*, 19: 599-605.

92 Galera, M.M.; Zamora, D.P.; Vidal, J.L.M.; Frenich, A.G.; Espinosa Mansilla, A.; Muñoz de la Peña, A.; López, F.S. (2003). "Determination of carbendazim, thiabendazole and fuberidazole using a net analyte signal-based method". *Talanta*, 59: 1107-1116.

- 93 Damiani, P.C.; Borraccetti, M.D; Olivieri, A.C. (2002). "Direct and simultaneous spectrofluorimetric determination of naproxen and salicylate in human serum assisted by chemometric analysis". *Anal. Chim. Acta*, 471: 87-96.
- 94 Park, K.S.; Lee, H.S.; Jun, C.H.; Park, K.H.; Jung, J.W.; Kim, S.B. (2000). "Rapid determination of iron oxide content in sinter ores using DRIFT spectra and multivariate calibrations". *Chemom. Intell. Lab. Syst.*, 51: 163-173.
- 95 Xu, L.; Schechter, I. (1997). "A calibration method free of optimum factor number selection for automated multivariate analysis. Experimental and theoretical study". *Anal. Chem.*, 69: 3722-3730.
- 96 Collado, M.S.; Mantovani, V.C.; Goicoechea, H.C.; Olivieri, A. (2000). "Simultaneous spectrophotometric-multivariate calibration determination of several components of ophthalmic solutions; phenylephrine, chloramphenicol, antipyrine, methylparaben and thimerosal". *Talanta*, 52: 909-920.
- 97 Olivieri, A.C. (2001). "Calibración multivariada. Introducción a la programación con MATLAB". Ed. Ediciones Cientificas Argentinas. Buenos Aires.
- 98 Boqué, R.; Rius, F.X. (1996). "Multivariate detection limits estimators". *Chemom. Intell. Lab. Syst.*, 32: 11-23.
- 99 Olivieri, A.C.; Goicoechea, H.C.; Inon, F.A. (2004). "MVC1: an integrated MatLab toolbox for first-order multivariate calibration". *Chemom. Intell. Lab. Syst.*, 73: 189-197.
- 100 Harshman R.A. (1972). "Determination and proof of minimum uniqueness conditions for PARAFAC1". *UCLA Work. Pap. Phon.* 22:111-117.
- 101 Sidiropoulos, N.D.; Bro R. (2000). "On the uniqueness of multilinear decomposition of N-way arrays". *J. Chemom.*, 14: 229-239.
- 102 Ghasemi, J.; Niazi, A. (2005). "Two-and three-wy chemometrics methods applied for spectrophotometric determination of lorazepam in pharmaceutical formulations and biological fluids". *Anal. Chim. Acta*, no hay datos actuales.
- 103 Cattell, R.B.(1944). "Parallel proportional profiles and other principles for determining the choise of factor by rotation". *Psychometrika*, 9: 267-271.

- 104 Faber, K.; Lorber, A.; Kowalski, B.R. (1997). "Generalized rank annihilation method: standard errors in the estimated eigenvalues if the instrumental errors are heteroscedastic and correlated". *J. Chemom.*, 11: 95-109.
- 105 Mc Donald, R.P. (1980). "A simple comprehensive model for the analysis of covariance structures: some remarks on applications". *British J. Mathem. and Statist. Psys.*, 33: 161- 167.
- 106 Rao, C.R.; Mitra, S. (1971). "Generalized inverse of matrices and its applications". Ed. Wiley, New York.
- 107 Leurgans, S.; Ross, R.T.; Abel, R.B. (1993). "A decomposition for three-way arrays". *SIAM J. Matrix Anal. Appl.*, 14: 1064-1071.
- 108 Faber, N.K.M.; Bro, R.; Hopke, P. (2003). "Recent developments in CANDECOMP/PARAFAC algorithms: a critical review". *Chemom. Intell. Lab. Syst.*, 65: 119-137.
- 109 Faber, N.M.; Buydens, L.M.C.; Kateman, G. (1994). "Generalized rank annihilation method: I. Derivation of eigenvalue problems". *J.Chemom.*, 8: 147-154.
- 110 Faber, N.M.; Buydens, L.M.C.; Kateman, G. (1986). "Generalized rank annihilation: III. Practical implementation". *J.Chemom.*, 8: 273-285.
- 111 Sanchez, E.; Kowalski, B.R. (1986). "Generalized rank annihilation factor analysis". *Anal. Chem.*, 58: 496-499.
- 112 Faber, N.M. (1997). "On solving generalized eigenvalue problems using MATLAB". *J. Chemom.*, 11: 87-91.
- 113 Lorber, A. (1985). "Features of quantifying composition from two-dimensional data array by the rank annihilation factor analysis method". *Anal. Chem.*, 57: 2395-2397.
- 114 Wu, H.L.; Shibukawa, M.; 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.

- 115** Chen, Z.P.; Wu, H.L.; Jiang, J.H.; Li, Y.; Yu, R.Q. (2000). "A novel trilinear decomposition algorithm for second-order linear calibration". *Chemom. Intell. Lab. Syst.*, 52: 75-86.
- 116** Chen, Z.P.; Wu, H.L.; Yu, R.Q. (2001). "On the self-weighted alternating trilinear decomposition algorithm the property of being insensitive to excess factors used in calculation". *J. Chemom.*, 15: 439-453.
- 117** Chen, Z.P.; Li, Y.; Yu, R.Q. (2001). "Pseudo alternating least squares algorithm for trilinear decomposition". *J. Chemom.*, 15: 149-167.
- 118** Li, Y.; Jiang, J.H.; Wu, H.L.; Chen, Z.P.; Yu, R.Q. (2000). "Alternating coupled matrices resolution method for three-way arrays analysis". *Chemom. Intell. Lab. Syst.*, 52: 33-43.
- 119** Jiang, J.H.; Wu, H.L.; Li, Y.; Yu, R.Q. (2000). "Three-way data resolution by alternating slice-wise diagonalization (ASD) method". *J. Chemom.*, 14: 15-36.
- 120** Brown, C.D. (2004). "Discordance between net analyte signal theory and practical multivariate calibration". *Anal. Chem.*, 76: 4364-4373
- 121** Hopke, P.K.; Paatero, P.; Jia, H.; Ross, R.T.; Harshman, R.A. (1998). "Three-way (PARAFAC) factor analysis: examination and comparison of alternative computational methods as applied to ill-conditioned data". *Chemom. Intell. Lab. Syst.*, 43: 25-42.
- 122** Andersson, C.A.; Bro, R. (2000). "The N-way toolbox for MATLAB". *Chemom. Intell. Lab. Syst.*, 52: 1-4.
- 123** Kowalski, B.R.; Seasholtz, M.B. (1991). "Recent developments in multivariate calibration". *J. Chemom.*, 5: 129-145.
- 124** Lorber, A. (1984). "Quantifying chemical decomposition from two-dimensional data array". *Anal. Chim. Acta*, 164: 293-297.
- 125** Li, S.; Gemperline, P.J. (1993). "Eliminating complex eigenvectors and eigenvalues in multiway analyses using the direct trilinear decomposition method". *J. Chemom.*, 7: 77-88.

- 126** Wang, Y.; Borgen, O.S.; Kowalski, B.R.; Gu, M.; Turecek, F. (1993). "Advances in second-order calibration". *J. Chemom.*, 7: 117-130.
- 127** Gimenez, D.; Sarabia, L.; Ortiz, M.C. (2005). "The maintenance of a PARAFAC calibration and second-order property: application to the determination of ciprofloxacin in presence of enrofloxacin by excitation-emission fluorescent". *Anal. Chim. Acta*, 544: 327-336.
- 128** Olivieri, A.C. (2004). "Sample-specific standard prediction errors in three-way parallel factor analysis (PARAFAC) exploiting the second order advantage". *J. Chem.*, 18: 1-9.
- 129** Ten Berge, J.M.F. (2004). "Partial uniqueness in CANDECOMP/PARAFAC". *J. Chemom.*, 18: 12-16.
- 130** Casassas, E.; Marques, I.; Tauler, R. (1995). "Study of acid-base properties of fulvic acids using fluorescence spectrometry and multivariate curve resolution methods". *Anal. Chim. Acta.*, 310: 473-484.
- 131** Tauler, R. (1995). "Multivariate curve resolution applied to second order data". *Chem. Intell. Lab. Syst.*, 30: 133-146.
- 132** de Juan, A.; Tauler, R. (2001). "Comparison of three-way resolution methods for non-trilinear chemical data sets". *J. Chemom.*, 15: 749-772.
- 133** Antunes, M.G.; Esteves da Silva, J.G. (2005). "Multivariate curve resolution analysis excitation-emission matrices of fluorescence of humic substances". *Anal. Chim. Acta*, 546: 52-59.
- 134** Tauler, R. (2001). "Calculation of maximum and minimum band boundaries of feasible solutions for species profiles obtained by multivariate curve resolution". *J. Chemom.*, 15: 627-646.
- 135** Andersen, C.M.; Bro, R. (2003). "Practical aspects of PARAFAC modeling of fluorescence excitation-emission data". *J. Chemom.*, 17: 200-215.
- 136** Boque, R.; Ferre, J. (2004). "Using second-order data in chromatographic analysis". *LC. GC. Europe*, 17: 402-407.



**137** Goicoechea, H.C and Olivieri, A.C. (2005). “New robust bilinear least squares method for the analysis of spectral-pH matrix data”. *Appl. Spectrosc.*, 59: 926-933.

**138** Espinosa Mansilla, A.; Muñoz de la Peña, A.; Salinas, F.; Gonzalez Gomez, D. (2004). “Partial least square multicomponent fluorimetric determination of fluoroquinolones in human urine samples”. *Talanta*, 62: 853-860.

**139** Para PARAFAC: <http://www.models.kvl.dk/source/>, para MCR-ALS: <http://www.ub.es/gesq/mcr/mcr.htm>.

**140** Huang, H.Y.; Chuang, C.L.; Chiu, C.W; Yeh, J.M. (2005). “Application of microemulsion electrokinetic chromatography for the detection of preservatives in foods”. *Food Chem.*, 89: 315-322.

**141** Maroto, A.; Boqué, R.; Rius, J.; Rius F.X. (2001). “Estimation of measurement uncertainty by using regression techniques and spiked samples”. *Anal. Chim. Acta*, 446: 133-145.

**142** Métodos oficiales para la cuantificación de aditivos químicos. AOAC, Association of Analytical Communities, <http://www.aoac.org>.

**143** Xie, Y.; Chen, P.; Wei, W. (1999). “Rapid analysis of preservatives in beverages by ion chromatography with series piezoelectric quartz crystal as detector”. *Microchem. J.*, 61: 58-68.

**144** Mihyar, G.; Yousif, A.; Yamani, M. (1999). “Determination of benzoic and sorbic acids in labaneh by high-performance liquid chromatography”. *J. Food Compos. Anal.*, 12: 53-61.

**145** Lin, Y.H.; Chou, S.S.; Sheu, F.; Shyu, Y.T. (2000). “Simultaneous determination of sweeteners and preservatives in preserved fruits by micellar electrokinetic capillary chromatography”. *J. Chromatogr. Sci.*, 38: 345-352.

**146** Frazier, R.A.; Inns, E.L.; Dosis, N.; Ames, J.M.; Nursten, H.E. (2000). “Development of capillary electrophoresis method for simultaneous analysis of artificial sweeteners, preservatives and colours in soft drinks”. *J. Chromatogr. A*, 876: 231-220.

- 147** Boyce, M.C. (1999). "Simultaneous determination of antioxidants, preservatives and sweeteners permitted as additives in food by micellar electrokinetic chromatography". *J. Chromatogr.*, 847: 369-375.
- 148** Pant, I.; Trenerry, V.C. (1995). "The determination of sorbic acid and benzoic acid in a variety of beverage and foods by micellar electrokinetic capillary chromatography". *Food Chem.*, 53: 219-226.
- 149** Garcia Castro, J.C.; Rodríguez Delgado, M.A.; Sánchez, M.J.; García Montelongo, F. (1992). "Simultaneous 2nd order derivative spectrophotometric determination of sorbic and benzoic acids in soft drinks". *Anal. Lett.*, 25: 2367-2376.
- 150** Fung, Y.; Luk, S. (1990). "Polarographic determination of sorbic acid in fruit juices and soft drinks". *Analyst*, 115: 1219-1221.
- 151** Hamano, T.; Mitsuhashi, Y.; Auki, N.; Semma, M.; Ito, Y. (1997). "Enzymic method for the spectrophotometric determination of benzoic acid in soy sauce and pickles". *Analyst*, 122: 259-262.
- 152** Muñoz de la Peña, A.; Espinosa Mansilla, A.; González Gómez, D.; Olivieri, A.; Goicoechea, H. (2003). "Interference-free analysis using three-way fluorescent data and the parallel factor model. Determination of fluoroquinolone antibiotics in human serum". *Anal. Chem.*, 75: 2640-2646.
- 153** Vives, M.; Gargallo, R.; Tauler, R. (1999). "Study of the intercalation equilibrium between the polynucleotide poly(adenylic)-poly(uridylic) acid and the ethidium bromide dye by means of multivariate curve resolution and the multivariate extension of the continuous variation and mole ratio methods". *Anal. Chem.*, 71: 4328-4337.
- 154** Smilde, A.K.; Tauler, R.; Saurina, J.; Bro, R. (1999). "Calibration methods for complex second-order data". *Anal. Chim. Acta*, 398: 237-251.
- 155** Pardo, R.; Helena, B.A.; Cazorro, C.; Guerra, C.; Deban, L.; Guerra, C.M.; Vega, M. (2004). "Application of two- and three-way principal component analysis to the interpretation of chemical fractionation results obtained by the use of the BCR [Community Bureau of Reference] procedure". *Anal. Chim. Acta*, 523: 125-132.
- 156** Reis, M.M.; Gurden, S.P.; Smilde, A.; Ferreira, M.M. (2000). "Calibration and detailed analysis of second-order flow-injection analysis data with rank overlap". *Anal. Chim. Acta*, 422: 21-36.

**157** Prasad, U.V.; Sastry, C.P. (1983). "Spectrophotometric determination of some permitted food colours using Aliquat 336". *J. Food Sci. Technol.*, 20: 263-264.

**158** European Union, Directive 94/36/EC on colours for use in foodstuffs. (1994).

**159** Fuh, M.R.; Chia, K.J. (2002). "Determination of sulfonated azo dyes in food by ion-pair liquid chromatography with photodiode array and electrospray mass spectrometry detection". *Talanta*, 56: 663-671.

**160** Chen, Q.; Mou, S.; Hou, X.; Riviello, J.M.; Z. Ni, Z. (1998). "Determination of eight synthetic food colorants in drinks by high-performance ion chromatography". *J. Chromatogr. A*, 827: 73-81.

**161** Holcapek, M.; Bandera, P.; Zderadicka, P. (2001). "High performance liquid chromatography-mass spectrometric analysis of sulfonated dyes and intermediates". *J. Chromatogr. A*, 926: 175-186.

**162** Scarpi, C.; Ninci, F.; Centini, M.; Anselmo, C. (1998). "High-performance liquid chromatography determination of direct and temporary dyes in natural hair colourings". *J. Chromatogr. A*, 796: 319-325.

**163** Huang, H.; Shih, Y.; Chen, Y. (2002). "Determining eight colorants in milk beverages by capillary electrophoresis". *J. Chromatogr. A*, 959: 317-325.

**164** Perez Urquiza, M.; Beltrán, J.L. (2000). "Determination of dyes in foodstuffs by capillary zone electrophoresis". *J. Chromatogr. A*, 898: 271-275.

**165** Dinc, E.; Baydan, E.; Kanpur, M.; Onur, F. (2002). "Spectrophotometric multicomponent determination of Sunset Yellow, tartrazine and Allura Red in soft drink powder by double divisor-ratio spectra derivative, inverse least-squares and principal component regression methods". *Talanta*, 58: 579-594.

**166** Berzas Nevado, J.; Rodríguez Flores, J.; Villaseñor Llerena, M.; Rodríguez Fariñas, N. (1999). "Simultaneous spectrophotometric determination of tartrazine, patent blue V and indigo carmine in commercial products by partial least-squares and principal-component regression methods". *Talanta*, 48: 895-903.

**167** Combeau, S.; Chatelut, M.; Vittori, O. (2002). "Identification and simultaneous determination of Azorubin, Allura red and Ponceau 4R by differential pulse polarography: application to soft drinks". *Talanta.*, 56: 115-122.

**168** Berzas Nevado, J.J.; Rodríguez Flores J.; Villaseñor Llerena, M.J. (1997). "Square wave adsorptive voltammetric determination of sunset yellow". *Talanta.*, 44: 467-474.

**169** Harada, K.; Masuda, K.; Suzuki, M.; Oka, H. (1991). "Separation and identification of food dyes by thin-layer chromatography - liquid secondary-ion mass spectrometry". *Biol. Mass. Spectrosc.* 20: 522-528.