



CAPÍTULO 1

Introducción

1.1. OBJETIVO DE LA TESIS DOCTORAL

Esta tesis doctoral pretende desarrollar metodologías fácilmente aplicables al análisis de mezclas de componentes en preparados farmacéuticos, que no cuenten con monografía oficial para su control de calidad, o que en el caso de contar con método oficial, éste implique el uso de HPLC o técnicas analíticas complejas. El trabajo está orientado además a la determinación de fármacos en sangre, comúnmente monitoreados por la importancia de sus niveles en la salud de los pacientes.

Estas metodologías están basadas en el acoplamiento de espectroscopía de absorción electrónica y técnicas computacionales.

La ventaja de estas metodologías radica en que se implementan a través de procedimientos sencillos, rápidos, económicos y que presentan un funcionamiento analítico adecuado para ser usadas en laboratorios que realizan control de calidad de medicamentos y para monitoreo de drogas terapéuticas.

Con esta finalidad se han planteado los siguientes objetivos:

- Desarrollo de métodos analíticos basados en la combinación de espectroscopía de absorción electrónica y calibración multivariada, para la determinación de fármacos en preparados farmacéuticos de uso humano y veterinario.
- Desarrollo de métodos analíticos basados en la combinación de espectroscopía de absorción electrónica y calibración multivariada, para el monitoreo de drogas terapéuticas.
- Estudio de diferentes estrategias para optimizar y validar el funcionamiento de los mismos.

1.2. ESTRUCTURA DE LA TESIS DOCTORAL

La memoria de esta tesis se estructura en cuatro capítulos. En el primer capítulo se recogen los fundamentos teóricos que se han utilizado en el desarrollo de la misma. Tras una breve introducción sobre el control de calidad de fármacos y monitoreo de drogas terapéuticas, se revisan los distintos tipos de métodos analíticos, los criterios para su selección y se hace una descripción de la metodología analítica utilizada en la tesis. Posteriormente se introducen los temas de diseño estadístico de experimentos, optimización experimental y validación de metodologías analíticas. Finalmente se presentan los métodos de calibración multivariada, para modelar respuestas lineales y no lineales utilizados en el trabajo de tesis.

En el segundo capítulo se presentan los desarrollos analíticos realizados, basados en la combinación de técnicas espectroscópicas y métodos de calibración multivariada y se estudia la implementación de diferentes estrategias, para mejorar los parámetros estadísticos de la calibración y las cifras de mérito y la capacidad predictiva de los métodos desarrollados.

En el tercer capítulo se presenta un resumen del trabajo realizado y las conclusiones generales extraídas de esta tesis.

En el capítulo cuatro, se presentan las referencias bibliográficas utilizadas.

1.3. NOTACIÓN

LISTA DE SÍMBOLOS Y ABREVIATURAS

Abreviatura o símbolo	Significado
+	Matriz seudoinversa (como supraíndice)
$\ \ $	Norma euclídeana de un vector (raíz cuadrada de la suma de los cuadrados de sus elementos)
®	Marca registrada
a	Ordenada al origen de la recta de calibración
A	Dimensionalidad. Número de factores
ANMAT	Administración Nacional de Medicamentos, Alimentos y Tecnología Médica
ANOVA	Análisis de la varianza (<i>analysis of variance</i>)
AOAC	Asociación Oficial de Químicos Analíticos (<i>Association of Oficial Analytical Chemists</i>)
ASTM	Sociedad Americana de Testeo y Materiales (<i>American Society for Testing and Materials</i>)
b	Pendiente de la recta de calibración
b_k	$(J \times 1)$ Coeficientes de regresión para el analito k
BLS	Método de cuadrados mínimos bivariados (<i>bivariate least-squares</i>)
C	$(I \times K)$ Concentraciones de los K analitos en las I muestras de calibración.
c	Concentración de un analito absorbente
c_k	$(I \times 1)$ Concentración del analito k en I muestras de calibración
$c_{k,un}$	Concentración del analito k en la muestra incógnita.
CLS	Método de cuadrados mínimos clásicos (<i>classical least-squares</i>)
c_{un}	$(K \times 1)$ Concentración de K analitos en una muestra incógnita
CV	Coefficiente de variación
DE	Diseño experimental
D_f	Grados de libertad (<i>degree free</i>)
E	Matriz de errores (las dimensiones dependen del caso)
EJCR	Intervalo de confianza elíptico conjunto (<i>elliptical joint confidence region</i>)

Abreviatura o símbolo	Significado
EPA	Agencia de Protección Ambiental (<i>Environmental Protection Agency</i>)
FB	Farmacopea Británica (<i>British Pharmacopoeia</i>)
FDA	Administración de alimentos y medicamentos (<i>Food and Drugs Administration</i>)
GC	Cromatografía gaseosa (<i>gas chromatography</i>)
GMP	Buenas Prácticas de Manufactura (<i>Good Manufacture Practice</i>)
HPLC	Cromatografía líquida de alto rendimiento
<i>I</i>	Número de muestras de calibración
I	$(J \times K)$ Matriz que relaciona las concentraciones con las respuestas de manera inversa a la ley de Beer
ICH	Conferencia Internacional de armonización (<i>International Conference of Harmonisation</i>)
ILS	Método de cuadrados mínimos inversos (<i>inverse least-squares</i>)
INAME	Instituto Nacional de Medicamentos
ISO	Organización Internacional de Estandarización (<i>International Organization for Standardization</i>)
IUPAC	Unión Internacional de Química Pura y Aplicada (<i>International Union of Pure and Applied Chemistry</i>)
<i>J</i>	Número de sensores
<i>K</i>	Número de analitos
<i>k</i>	Factor o variable controlada en el diseño experimental
LOD	Límite de detección (<i>limit of detection</i>)
LOQ	Límite de cuantificación (<i>limit of quantitation</i>).
MS	Espectrometría de masa (<i>mass spectrometry</i>)
MSe	Variancia del error puro
MSlof	Variancia por falta de ajuste
MSr	Variancia de los residuos
N	Número de experimentos en el diseño experimental
NCCLS	Comité Nacional de Estándares para Laboratorios Clínicos (<i>National Committee for Clinical Laboratory Standards</i>)

Abreviatura o símbolo	Significado
OLS	Método de cuadrados mínimos ordinarios (<i>ordinary least-squares</i>)
P	$(J \times A)$ Loadings para la matriz R calculados con A factores.
PCR	Método de regresión en componentes principales (<i>principal component regression</i>)
P	Potencia del haz de luz luego de atravesar una muestra
P_0	Potencia del haz de luz antes de atravesar una muestra
PMQ	Proceso de Medida Química
PRESS	Error en la predicción (<i>Prediction error sum of squares</i>)
R	Coefficiente de correlación
R	$(J \times I)$ Respuestas instrumentales de I muestras de calibración medidas a J sensores
r	$(J \times 1)$ Respuestas instrumentales de una muestra incógnita medidas a J sensores
R^2	Cuadrado del coeficiente de correlación.
REP(%)	Error estándar de predicción (<i>relative error of prediction</i>)
RMS	Método de Superficie de Respuesta (<i>Responce Surface Method</i>)
RMSECV	Error absoluto en validación cruzada (<i>root mean square error in cross validation</i>)
RSD	Desviación estándar relativa (<i>relative standard desviation</i>)
S	$(J \times K)$ Respuestas instrumentales de K analitos en concentración unitaria.
s	Estimador de la desviación estándar
s_a	Estimador de la desviación estándar de la ordenada al origen de la recta de calibración
s_b	Estimador de la desviación estándar de la pendiente de la recta de calibración
SEL	Selectividad
SEN	Sensibilidad
s_k	$(J \times 1)$ Respuestas instrumentales del analito k en concentración unitaria
s_k^*	$(J \times 1)$ Señal neta del analito puro k
SSe	Suma de cuadrados del error puro
SSlof	Suma de cuadrados por falta de ajuste

Abreviatura o símbolo	Significado
SSr	Suma de cuadrados de los residuos
T	Transmitancia
T	Matriz transpuesta (como supraíndice)
T	($I \times A$) Scores de la matriz R calculados con A factores
USP	Farmacopea de los Estados Unidos de América (<i>United States Pharmacopeia</i>)
UV	Ultravioleta
v_k	($A \times 1$) Vector de coeficientes que relacionan c_k con la matriz T
W	($J \times A$) Matriz de <i>weight loading factors</i> de la matriz R calculados con A factores en PLS
WLS	Método de cuadrados mínimos ponderados (<i>weighted least-squares</i>)
α	Error de tipo I
β	Error de tipo II
α	Parámetro utilizado en el diseño experimental cuyo valor es igual a la raíz cuarta de N
γ	Sensibilidad analítica
ν	Grados de libertad
ϵ'	Coefficiente de extinción molar de la solución absorbente
ϵ_λ	Coefficiente de absorción molar de las especies a la longitud de onda λ
λ	Longitud de onda

1.4. METODOS ANALÍTICOS PARA EL ANÁLISIS DE MUESTRAS COMPLEJAS

La determinación de las concentraciones de drogas terapéuticas en fluidos biológicos y preparados farmacéuticos, resulta un campo de interés en el área salud, y genera una intensa demanda de nuevas metodologías analíticas.

Entre las razones más importantes para analizar fármacos se pueden enumerar las siguientes: a) rango terapéutico estrecho, b) ausencia de parámetros clínicos para seguir sus efectos, c) falta de relación predecible entre dosis y respuesta, d) la necesidad de realizar estudios de biodisponibilidad y bioequivalencia en genéricos, y e) control de calidad de materias primas y productos comercializados.

1.4.1. CONTROL DE CALIDAD DE FÁRMACOS

Las preparaciones farmacéuticas tales como comprimidos, cápsulas, cremas, inyectables, geles, emulsiones, suspensiones orales, gotas oftálmicas y orales, son consideradas muestras complejas, dado que su matriz esta compuesta por lo general por varios excipientes. Por lo tanto, para realizar el control de calidad de las mismas, es necesario contar con métodos de análisis adecuados que permitan resolver los diferentes problemas analíticos que se presentan.

Los fármacos están destinados a atender una necesidad tan esencial como es la salud de las personas. Por tal razón, los mismos deben acreditar su calidad, seguridad y eficacia en todas las etapas que cursan desde su diseño y formulación, pasando por la elaboración y el control de calidad, hasta llegar a su dispensación al público.

La prescripción de medicamentos está bajo la responsabilidad de los médicos, y la dispensación de ellos corresponde sólo a los farmacéuticos.

El sistema de producción y entrega de medicamentos, está sometido a disposiciones legales y reglamentarias. En todos los países del mundo se dictan leyes, normas y reglamentos destinados a garantizar medicamentos que produzcan el mayor y mejor efecto curativo con el menor riesgo posible.

Un fármaco no puede entrar en uso médico sin satisfacer las pruebas o ensayos clínicos, preclínicos o "in vitro" a que deben ser sometidos. Además antes de entrar en el mercado debe contar con la aprobación de las autoridades pertinentes de cada país.

Hay normas de producción, llamadas normas de "Buenas Prácticas de Manufactura" (GMP, en inglés *Good Manufacture Practice*), que establecen requisitos relativos a la calidad de las instalaciones, maquinarias, procedimientos, personal técnico y varios otros aspectos. Estas normas son elaboradas

por organismos relacionados con el tema, como son la *FDA* (Food and Drug Administration) en los Estados Unidos o el *ANMAT* (Administración Nacional de Medicamentos, Alimentos y Tecnología Médica) de nuestro país.

Por otra parte, se exige el control de calidad de materias primas, productos en proceso y productos terminados y para esto se pueden seguir las especificaciones que aparecen en las Farmacopeas.

Las Farmacopeas son textos oficiales que contienen información acerca de los parámetros que deben cumplir los medicamentos con el objeto de asegurar su eficacia e inocuidad.

En las mismas se describen especificaciones para las drogas puras (materias primas a partir de las cuales se fabrican medicamentos) y para productos terminados o preparaciones farmacéuticas; estableciéndose además las metodologías analíticas apropiadas para comprobar que dichas especificaciones se cumplan efectivamente.

Las más importantes son la *USP* (Farmacopea de los Estados Unidos), la *BP* (Farmacopea Británica) y en nuestro país esta vigente la séptima edición de la *Farmacopea Argentina*.

Para llevar a cabo las tareas de control de calidad de ciertas preparaciones farmacéuticas, es necesario desarrollar métodos analíticos en el laboratorio, ya sea porque no se dispone de ellos o porque se necesitan métodos más rápidos y/o económicos.

El desarrollo de métodos de análisis es un área de investigación en Química Analítica de gran actualidad y de suma importancia en el tema del control de calidad de fármacos.

1.4.2. MONITOREO DE DROGAS TERAPÉUTICAS

El monitoreo de drogas terapéuticas constituye una parte esencial en el cuidado del paciente, ya que algunas drogas tienen eficacia y efectividad, si se administran en un rango de concentración terapéutica específico. Es decir que esta disciplina de monitoreo o seguimiento de drogas, permite estimar la dosis exacta para prescribir un medicamento, de modo de maximizar los beneficios para el paciente, minimizando los riesgos de toxicidad (Shaw, 1998; Bowers, 1998).

Para lograr el efecto farmacológico requerido, es preciso alcanzar cierta concentración en la biofase o sitio de interacción entre la molécula de la droga y el receptor. Las dosis o concentraciones que no producen ningún efecto medible o cuantificable se consideran subterapéuticas, mientras que aquellas superiores a la dosis mínima que produce un 100% de efectividad no son garantizables y pueden llegar a ser tóxicas.

El efecto de la mayoría de las drogas se puede describir mediante una curva de logaritmo dosis-respuesta o logaritmo concentración-respuesta, que habitualmente presenta una forma sigmoidea y el intervalo terapéutico es un rango de concentraciones situado en alguna parte del tercio inferior de la curva (Kaplan *et al*, 1986).

La relación entre la dosis tóxica de una droga y la terapéutica, se denomina *índice terapéutico* y puede ser estrecha para algunas drogas y amplia para otras. Por ende, el *intervalo terapéutico* también puede ser estrecho o amplio. El monitoreo de las concentraciones sanguíneas, es particularmente aconsejable, para aquellas drogas que presentan un intervalo terapéutico estrecho y un bajo índice terapéutico, para drogas cuya cinética de eliminación depende de la dosis, o para aquellas que muestran una gran variabilidad individual en su metabolismo (Kaplan, 1986; Goodman Hilman, 1996).

Los fármacos más comúnmente monitoreados son los broncodilatadores, antipiréticos, antiinflamatorios, antibióticos, agentes inmunosupresores, psicoactivos, antipiréticos y cardíacos. Las ventajas alcanzadas mediante la utilización de esta disciplina son (Shaw, 1998; Korpela *et al*, 1998; Bowers, 1998):

- ◆ Reducción del tiempo requerido para el tratamiento.
- ◆ Reducción del riesgo de toxicidad en pacientes vulnerables a efectos adversos.
- ◆ Establecimiento de dosis efectivas cuando se inicia la terapia.
- ◆ Reducción del riesgo por falla terapéutica.
- ◆ Eliminación de interacciones droga-droga.
- ◆ Minimización de los costos del tratamiento.
- ◆ Eliminación de problemas médico-legales.
- ◆ Aumento de la capacidad por parte del médico para el uso efectivo de las drogas.
- ◆ Reducción de los riesgos de desarrollo de resistencia bacteriana a antibióticos por uso de cantidades mayores a las realmente necesarias para el tratamiento de la enfermedad.

Por otra parte, en nuestro país existe una disposición de la Administración Nacional de Medicamentos, Alimentos y Tecnología Médica (ANMAT) de realizar estudios de bioequivalencia en medicamentos que han de ser incorporados como genéricos. Esto ha acentuado la necesidad de contar con métodos confiables y aplicables a los laboratorios nacionales.

1.4.3. SELECCIÓN DE MÉTODOS ANALÍTICOS

Existe una ordenación conceptual y técnica para clarificar el significado de las palabras *Proceso*, *Técnica*, *Método* y *Procedimiento* (Valcárcel, 1999).

La *Técnica* es un principio genérico que se utiliza para extraer información, que implica el uso de un instrumento. El *Proceso Analítico o Proceso de Medida Química (PMQ)* es el conjunto de operaciones que separa a la muestra (sin tomar, sin medir y sin tratar), de los resultados generados y expresados según los requerimientos del *Problema Analítico* planteado. El *Proceso de Medida Química* se puede considerar dividido en tres etapas generales:

- * Etapa de las operaciones previas (etapa pre-analítica), cuyo objetivo es la adecuación de la muestra a la medición propiamente dicha.
- * Etapa de la medida y transducción de la señal analítica (etapa analítica), directamente relacionada con las características/concentración de los analitos. Implica el uso de un instrumento (técnica analítica).
- * Etapa de adquisición de señales y tratamiento de datos (etapa post-analítica), cuyo objetivo es obtener los resultados de acuerdo a lo requerido.

El *Método Analítico* es la aplicación concreta de una técnica analítica para la determinación de un analito o mezcla en una muestra definida. Supone la concreción del Proceso Analítico.

Por último el *Procedimiento* es la descripción pormenorizada del método, es decir las instrucciones escritas para su aplicación.

1.4.3.1. CLASIFICACIÓN DE LOS MÉTODOS ANALÍTICOS

Los métodos analíticos se pueden clasificar de acuerdo a diferentes criterios. Existe una diferenciación basada en la técnica implicada en el proceso analítico, que habla de *métodos clásicos* y *métodos instrumentales*.

Los métodos clásicos precedieron en más de un siglo a los instrumentales y pueden tener una finalidad cualitativa o cuantitativa.

Los métodos clásicos cualitativos sirven para la identificación de iones y moléculas, mediante el empleo de reacciones químicas y los sentidos humanos. Dentro de los métodos clásicos con connotaciones cuantitativas, están las volumetrías y las gravimetrías.

Los métodos instrumentales pueden tener una finalidad cualitativa, cuantitativa o estructural y se basan en la medición de propiedades físicas de los analitos tales como conductividad, absorción o emisión de la luz, razón masa a carga y fluorescencia.

Por otra parte, los procesos analíticos actuales, emplean frecuentemente procedimientos separativos para separar y resolver compuestos estrechamente relacionados, la mayoría de los cuales se basa en cromatografías y electroforesis capilar.

Los métodos analíticos también pueden clasificarse en función de su origen. Se habla de *métodos oficiales*, cuando son métodos que aparecen publicados en textos oficiales o documentos legislativos, como es el caso de los métodos de las Farmacopeas, los publicados por la *FDA* o por la *EPA* (Environmental Protection Agency) en los Estados Unidos. Estos métodos analíticos están rigurosamente validados y se considera que son de alta calidad metrológica.

Los *métodos estándar*, son métodos analíticos desarrollados, validados y publicados por organizaciones internacionales de prestigio, tales como la *AOAC* (Association of Official Analytical Chemists) o la *ASTM* (American Society for Testing and Materials).

Finalmente, están los *métodos publicados en la literatura*, es decir en las revistas científicas y los *desarrollados en el propio laboratorio*. Ambos tipos de métodos deben someterse a procedimientos de validación, antes de su implementación para el análisis de muestras.

Desde el punto de vista de la calibración necesaria y de la trazabilidad, los métodos analíticos se clasifican en *calculables* y *relativos*.

Un método calculable es aquel que origina un resultado mediante cálculos basados en las leyes que rigen los parámetros químicos y físicos, que se materializan en fórmulas matemáticas en las que están implicadas tanto constantes (por ejemplo, pesos atómicos) como las medidas tomadas durante el proceso analítico, como pesos, volúmenes, etc.

Dentro de los métodos calculables se distingue entre los *métodos absolutos* y los *absolutos con estándares analíticos*. Los métodos absolutos no emplean estándares químico-analíticos para generar un resultado, mientras que los métodos absolutos con estándares analíticos son aquellos que necesitan de un estándar químico-analítico que no contiene al analito.

Es decir que los métodos absolutos se relacionan directamente con patrones básicos. Así por ejemplo, un método gravimétrico o un método que utiliza una técnica analítica como la espectrometría de masas, tienen una trazabilidad directa al patrón de masa, una vez calibrado el instrumento.

Los métodos absolutos con estándares analíticos pueden ser estequiométricos, como es el caso de las volumetrías, donde se emplean sustancias patrón que no contienen generalmente al analito que después se determina, y la determinación analítica se basa directamente en la relación estequiométrica de la reacción entre el valorante y el analito de la muestra. Dentro de los métodos absolutos con estándares analíticos, además de las volumetrías, están los que utilizan dilución isotópica-espectrometría de masas, pero en este caso se trata de métodos no estequiométricos.

Los métodos relativos son aquellos que precisan la calibración metodológica del instrumento de medida, respecto de la concentración del componente de interés mediante patrones químico-analíticos que en este caso, sí contienen el mismo analito objeto posterior de determinación. Es decir que el resultado se deduce sin necesidad de cálculos basados en teorías físicas o químicas.

Dentro de los métodos relativos, se distinguen dos grupos. Los *métodos relativos por interpolación*, son aquellos que emplean instrumentos en los que es posible definir una relación clara y contundente entre la señal generada y la concentración del analito, es decir que se puede obtener una recta de calibrado con sustancias puras y con muestras-patrón (materiales tipo matriz con composición semejante a la muestra).

El segundo tipo de métodos relativos se denominan *métodos comparativos* y en este caso, la comparación debe hacerse entre la señal que genera la muestra y la que se origina a partir de un conjunto de muestras-patrón, debido a que la respuesta no solo depende de la concentración del analito sino también de la matriz de la muestra.

También se puede hacer una clasificación de los métodos analíticos de acuerdo a su calidad metrológica, siendo los *métodos definitivos o primarios*, los que tienen las más altas cotas de calidad metrológica, dado que están directamente relacionados con estándares básicos.

Los métodos definitivos tienen asociados procesos de medida con operaciones perfectamente descritas y no necesitan de un estándar de la cantidad composicional (estándar analítico que contiene al analito). Se caracterizan además por tener un error sistemático despreciable y una incertidumbre definida y pequeña.

A los métodos definitivos le siguen los *métodos de referencia*, que son aquellos que han demostrado que sus resultados son correctos, por comparación directa con un método definitivo o un material de referencia primario.

1.4.3.2. CRITERIOS PARA LA SELECCIÓN DE MÉTODOS ANALÍTICOS

El *Proceso Analítico Total* se inicia con el planteamiento del *Problema Analítico* a resolver (Kellner *et al*, 1998), es decir con el establecimiento del objetivo del trabajo, las necesidades, los requerimientos y las disponibilidades.

Al definir el Problema Analítico, el analista debe tener en cuenta los siguientes aspectos:

- Tipo de información analítica requerida: cualitativa, cuantitativa o ambas, elemental, de grupos funcionales, de especies químicas o estructural.
- Niveles de concentración de los analitos de interés: mayoritarios, minoritarios o trazas.
- Tipo de muestra: estado físico, estabilidad, homogeneidad, composición de la matriz, posibles interferencias.
- Cantidad de muestra disponible.
- Error admisible: exactitud y precisión requeridas.
- Tiempo disponible entre el muestreo y la entrega de resultados
- Número de muestras a analizar.

Una vez definido el Problema Analítico, se hace una exploración bibliográfica y se analizan las características de los métodos analíticos encontrados, a los fines de seleccionar el más apropiado.

Un método analítico puede caracterizarse mediante criterios de calidad, que pueden ser de tres tipos: criterios estadísticos, criterios operativos y criterios económicos.

Los criterios estadísticos son: precisión, exactitud, límite de detección, límite de cuantificación, sensibilidad, selectividad, rango dinámico, y robustez.

Dentro de los criterios operativos de la calidad se pueden mencionar: rapidez (duración del análisis), equipamiento requerido (especificaciones técnicas, necesidades energéticas y de espacio), precauciones especiales de seguridad, tipo y cantidad de muestra necesaria, y grado de automatización.

Los criterios económicos de la calidad comprenden: inversión inicial, costo de puesta a punto, costo por determinación, costos relativos al personal.

En este punto pueden darse diferentes situaciones (Compañó Beltrán, 2002):

- Es posible encontrar un método validado, muy apropiado, que se ajusta plenamente a las necesidades y que por lo tanto no requiere modificaciones. En este caso, el laboratorio debe simplemente demostrar que es capaz de reproducir los parámetros de calidad analítica, antes de utilizarlo para analizar muestras.
- Es posible encontrar un método bastante apropiado, pero que solo satisface parcialmente las necesidades y que por lo tanto requiere de modificaciones para ser adaptado. En este caso se deben optimizar las modificaciones realizadas y el método debe ser validado.
- Es posible que no se encuentre ningún método que sea apropiado, por lo tanto se hace necesario desarrollar enteramente un nuevo método. En este caso, durante el desarrollo del método analítico se debe realizar la optimización del mismo y posteriormente se debe validar.

1.4.4. METODOS ANALÍTICOS DE FARMACOPEA

Los métodos analíticos oficiales citados en las Farmacopeas para realizar el análisis de medicamentos se pueden agrupar de acuerdo a la Tabla 1.1.

Los métodos analíticos más utilizados son los cromatográficos y dentro de ellos los que implican el uso de cromatografía líquida de alto rendimiento (HPLC). Esto se debe a que los medicamentos son muestras complejas y estos métodos se caracterizan por su sensibilidad y su idoneidad para la separación de especies no volátiles o termolábiles.

Los métodos cromatográficos han tenido un tremendo impacto en la ciencia, debido a su utilidad para caracterizar mezclas complejas. Sin embargo, cuando se utilizan en el análisis rutinario, presentan el inconveniente de insumir tiempos importantes, requerir equipamiento sofisticado y caro, reactivos de alto costo y utilizar solventes que en algunos casos son contaminantes.

Tabla 1.1: Métodos analíticos oficiales para el análisis de medicamentos

Titulométricos	<ol style="list-style-type: none"> 1. Reacciones ácido-base en solventes acuosos y no acuosos 2. Reacciones de precipitación 3. Reacciones Redox 4. Reacciones de complejación
Gravimétricos	<ol style="list-style-type: none"> 1. Pesada de la droga después de la separación 2. Pesada de un derivado después de la separación 3. Pesada de un residuo después de la ignición
Espectrométricos	<ol style="list-style-type: none"> 1. Absorción visible y ultravioleta 2. Absorción en el infrarrojo 3. Emisión fotométrica de llama 4. Emisión fluorimétrica 5. Absorción atómica
Electroquímicos	<ol style="list-style-type: none"> 1. Voltametría 2. Potenciometría
Microbiológicos	<ol style="list-style-type: none"> 1. Recuento de microorganismos 2. Ensayos de potencia (antibióticos) 3. Ensayos de crecimiento (vitaminas) 4. Ensayos de esterilidad 5. Determinación de pirógenos
Cromatográficos	<ol style="list-style-type: none"> 1. Cromatografía líquida de alto rendimiento (HPLC) 2. Cromatografía de gases 3. Cromatografía en capa delgada
Diversos	<ol style="list-style-type: none"> 1. Medición de rotación óptica 2. Medición de densidad. 3. Medición de actividades enzimáticas 4. Medición de radioactividad

1.4.5. METODOLOGÍA ANALÍTICA UTILIZADA EN LA TESIS

Las muestras biológicas tales como el suero humano, también son consideradas muestras complejas, y en el análisis de las mismas se presenta una situación similar a la mencionada para el caso de los preparados farmacéuticos.

Los métodos espectroscópicos ofrecen una alternativa a los métodos cromatográficos, ya que se caracterizan por ser rápidos, sencillos, de bajo costo y fácilmente adaptables a laboratorios de mediana complejidad. Estos métodos constituyen un amplio grupo de métodos analíticos basados en la interacción de las radiaciones con la materia, siendo los más ampliamente utilizados aquellos que están relacionados con la radiación electromagnética (Skoog, 1994). La radiación electromagnética es un tipo de energía que toma diferentes formas, siendo las más fácilmente reconocibles la luz y el calor radiante y las manifestaciones más difícilmente reconocibles los rayos gamma, los rayos X, las radiaciones ultravioletas, de microondas y de radiofrecuencia. Los diferentes tipos de radiación electromagnética se caracterizan por parámetros ondulatorios tales como: amplitud (A), frecuencia (ν), longitud de onda (λ), potencia (P), intensidad (I).

El espectro electromagnético abarca un intervalo amplio de radiaciones electromagnéticas de diferentes longitudes de onda, frecuencias y energías, y esta dividido en regiones espectrales, de acuerdo a los métodos necesarios para generar y detectar las diversas clases de radiación.

Con frecuencia se denomina métodos ópticos a los métodos que utilizan no sólo radiación visible, sino también a aquellos que utilizan radiación ultravioleta e infrarroja.

Cuando la radiación atraviesa una capa de un sólido, un líquido o un gas, ciertas frecuencias pueden eliminarse selectivamente por *absorción*, un proceso en el que la energía electromagnética se transfiere a los átomos, iones o moléculas que componen la muestra. La absorción provoca que estas partículas pasen de su estado normal a temperatura ambiente (estado fundamental) a uno o más estados excitados de energía superior. De acuerdo con la teoría cuántica, los átomos, moléculas o iones sólo tienen un número limitado de niveles de energía discretos. Para que se produzca la absorción de radiación, la energía de los fotones excitadores debe coincidir exactamente con la diferencia de energía entre el estado fundamental y uno de los estados excitados de las especies absorbentes.

Cada especie tiene diferencias de energía características y muestra espectros de absorción típicos, cuando se hace una representación gráfica de la *absorbancia* (medida de la disminución de la potencia radiante) en función de la longitud de onda.

Los espectros de absorción de partículas monoatómicas, presentan una relativa simplicidad debido al pequeño número de posibles estados de energía. La excitación sólo puede producirse mediante un proceso electrónico en el que uno o más de los electrones del átomo se promocionan a un nivel de energía superior.

Los espectros de absorción de las moléculas poliatómicas son considerablemente más complejos que los espectros atómicos, ya que el número de estados de energía de las moléculas es generalmente enorme, si se compara con el de los átomos aislados.

Para la realización del presente trabajo de tesis, se utilizaron métodos espectroscópicos cuantitativos basados en la absorción molecular (*espectrometría de absorción molecular*), para los cuales se requieren dos medidas de Potencia: una antes de que el haz haya pasado a través del medio que contiene al analito (P_0) y la otra, después (P).

Estas medidas permiten calcular la fracción de radiación incidente transmitida por el medio o *transmitancia* (T) según,

$$T = \frac{P}{P_0} \quad (1.1)$$

y la *absorbancia* (A) de un medio como:

$$A = -\log_{10} T = \log \frac{P_0}{P} \quad (1.2)$$

La espectrometría de absorción molecular se basa en la medida de la *transmitancia* (T) o de la *absorbancia* (A) de disoluciones que se encuentran en cubetas transparentes, que tienen un camino óptico de b cm.

Para una radiación monocromática, la concentración c de un analito absorbente esta relacionada linealmente con la absorbancia, de acuerdo a la siguiente ecuación:

$$A = abc \quad (1.3)$$

Esta ecuación es una representación matemática de la Ley de Beer, donde a es una constante de proporcionalidad llamada *absortividad*.

Si b se da en cm y c en gramos por litro, las unidades de la absortividad son $L g^{-1} cm^{-1}$. Si en cambio la concentración se expresa en moles por litro, las unidades son $L mol^{-1} cm^{-1}$, se denomina *absortividad molar* y se representa por el símbolo ϵ .

La ley de Beer también puede aplicarse a un medio que contenga más de una clase de sustancias absorbentes y siempre que no haya interacción entre las especies, la absorbancia total para un sistema multicomponente esta dada por:

$$\begin{aligned} A &= A_1 + A_2 + \dots + A_n \\ A &= \varepsilon_1 b c_1 + \varepsilon_2 b c_2 + \dots + \varepsilon_n b c_n \end{aligned} \quad (1.4)$$

Donde los subíndices se refieren a los componentes absorbentes $1, 2, \dots, n$.

Los métodos de absorción ultravioleta / visible cuantitativos tienen una gran aplicabilidad y abarcan todos los campos en los que se demanda información química cuantitativa.

Existen numerosas especies orgánicas e inorgánicas absorbentes susceptibles de determinación directa, mientras que para las especies no absorbentes existen reactivos que generan productos que absorben fuertemente en las regiones ultravioleta y visible del espectro.

Estos métodos se caracterizan además por tener buena precisión y buena sensibilidad, siendo posible realizar la adquisición de datos de manera sencilla.

La desventaja de estos métodos es que en numerosas ocasiones, la interacción del analito con el particular entorno químico-físico en el que se encuentra, produce una exaltación o una inhibición de la sensibilidad, lo que se conoce como efecto matriz.

En lo que respecta a su selectividad, puede suceder que una parte de la señal obtenida sea debida a interferencias, es decir a causas distintas de la concentración del analito presente en la muestra. Para mejorar la selectividad de los métodos espectroscópicos utilizados para el análisis de muestras complejas de multicomponentes, se trabaja en forma combinada con diferentes modelos de calibración multivariada.

En este caso, la selectividad se consigue mediante la utilización de una herramienta quimiométrica, como es la calibración multivariada, soslayando de esta manera etapas tediosas de pretratamiento químico de las muestras, para aislar los analitos de interés.

1.5. DISEÑO EXPERIMENTAL

1.5.1. QUIMIOMETRÍA

El desarrollo de la Quimiometría como disciplina científica independiente se remonta al año 1969, cuando aparecen los trabajos de Jurs, Kowalski e Isenhour. Años más tarde se publican las primeras revistas especializadas de investigación en Quimiometría: *Chemometrics and Intelligent Laboratory Systems* (1986) y *Journal of Chemometrics* (1987). (Ramis Ramos, 2002)

La Quimiometría trata, específicamente, de todos aquellos procesos que transforman señales analíticas y datos más o menos complejos en información, utilizando métodos de origen matemático, estadístico y otros procedentes del campo de la lógica formal para conseguir sus fines (Massart, 1997).

La Quimiometría se puede dividir en tres áreas generales (Brereton, 1990)

- * Optimización de los procedimientos experimentales y mediciones químicas.
- * Extracción de la máxima información química de los datos analíticos.
- * Calibración, validación y significancia de las mediciones analíticas.

El análisis de datos reales demuestra que, frecuentemente, estos no contienen la información suficiente. A menudo, las técnicas quimiométricas no pueden extraer información útil, debido a que los datos no son representativos del fenómeno que quiere estudiarse o bien no contienen suficiente variabilidad, siendo la parte aleatoria más relevante que la parte correspondiente a la variación sistemática. Por esta razón, las experiencias que conducen a los datos deben programarse mediante el empleo de un diseño experimental.

1.5.2. DISEÑO ESTADÍSTICO DE EXPERIMENTOS

La experimentación juega un papel fundamental en todos los campos de la investigación y el desarrollo, siendo su objetivo el de obtener información de calidad.

La experimentación se debe planificar cuidadosamente para que proporcione la información buscada. Dicha planificación debe considerar dos aspectos importantes relacionados con toda experimentación:

- * La capacidad de experimentar está limitada por el costo en tiempo y en recursos humanos. Por lo tanto una organización óptima de la experimentación deberá contemplar el menor número de experimentos que permita obtener la información buscada.

* El resultado observado de un experimento tiene incertidumbre, por esto, la *Estadística*, disciplina que proporciona las herramientas para trabajar en ambientes de incertidumbre, juega un papel fundamental en el diseño de los experimentos y en la evaluación de los resultados experimentales.

Para realizar la elección de los experimentos y la evaluación de los resultados se debe utilizar una metodología matemática y estadística que indique cómo planificar, diseñar u organizar la secuencia de experimentos de una forma óptima, de modo de minimizar el costo y la influencia del error experimental sobre la información buscada. Dicha planificación y análisis es el principal objetivo del *Diseño Estadístico de Experimentos o Diseño Experimental (DE)* (en inglés, *DOE Design Of Experiments*).

El Diseño experimental se aplica a sistemas en los cuales se observan una o más variables experimentales dependientes o respuestas, cuyo valor depende de los valores de una o más variables independientes controlables llamadas *factores*. Las respuestas pueden estar influenciadas además por otras variables que no son controladas por el experimentador. Es decir que el diseño de experimentos es una técnica orientada a problemas multivariantes.

En términos generales se puede afirmar que el diseño experimental propone, por un camino eficiente, obtener la optimización de un proceso o producto con un mínimo de experimentos.

En Química Analítica, el Diseño Experimental se aplica en contextos tales como:

- * Determinar la influencia de los factores sobre las respuestas observadas y cómo interaccionan entre ellos.
- * Optimizar respuestas.
- * Determinar la robustez de un sistema.
- * Desarrollar modelos para los experimentos de calibración univariada y multivariada (Martens, 1989; Araujo, 1996).

Normalmente existen diversos factores que intervienen en las experiencias, siendo un *factor*, una variable que puede ser manejada de un modo controlado para estudiar su efecto y que tiene (o podría tener) influencia sobre la respuesta estudiada. Se entiende por niveles de un factor, los diferentes valores que dicho factor toma y el *efecto de un factor* se define como el cambio en la respuesta producida por un cambio en el nivel del factor. Se habla de *efecto principal* cuando nos referimos a los factores de interés primordial del experimento.

Los factores se pueden clasificar en:

Cualitativos: cuando sus posibles valores no se pueden ordenar numéricamente (en un ensayo de distintos solventes: agua, acetona, alcohol etílico, etc).

Cuantitativo: cuando los valores posibles se pueden ordenar numéricamente (pH, temperatura, concentración de reactivos, longitud de onda, caudales de gases o reactivos, volumen de muestra, etc).

El modelo que relaciona la respuesta con el efecto de los factores, se denomina *función respuesta*, se obtiene de los experimentos y puede ser una relación algebraica o gráfica.

En el diseño experimental, cada factor representa una dimensión espacial, de manera que cuando se trabaja con varios factores, se tendrá un espacio multidimensional.

Para muestrear este espacio multidimensional se pueden utilizar diferentes métodos, de acuerdo a las características de los problemas a resolver y las ventajas y desventajas intrínsecas de cada uno.

El método de experimentación que surge de manera más intuitiva para estudiar un sistema, consiste en variar un factor cada vez, es decir que a partir de unas condiciones iniciales, se realizan experimentos en los cuales todos los factores se mantienen constantes excepto el que se está estudiando. Luego el procedimiento se repite para los otros factores. De este modo la variación de la respuesta se puede atribuir a la variación del factor y por lo tanto revela el efecto de ese factor.

Esta estrategia experimental presenta inconvenientes importantes cuando existe interacción entre factores ya que no informa cómo un factor interactúa con los otros factores o como estas interacciones afectan a la respuesta. Además, en el caso de la optimización experimental, puede no proporcionar la posición del óptimo.

La interacción entre dos o más factores, es el fenómeno por el cual el efecto de un factor sobre la respuesta depende del nivel de otro (u otros) factores. Se denomina *interacción doble* a la interacción entre dos factores, *interacción triple* a la interacción entre tres factores y así sucesivamente.

Para solucionar los inconvenientes de este método tradicional, se debe variar más de un factor simultáneamente mediante una experimentación reducida donde estos cambios simultáneos se complementen entre sí y permitan obtener la información buscada.

El *DE* proporciona el marco matemático para cambiar todos los factores simultáneamente y conduce a una planificación con un número reducido de experimentos y máxima eficiencia.

Los principios básicos en el Diseño de Experimentos son el de obtención de réplicas, aleatorización y análisis por bloques (Morgan, 1995).

La réplica es una repetición del experimento básico, que permite tener una estimación del error experimental. Esta estimación permite determinar si las diferencias observadas en los datos son estadísticamente significativas. El uso de réplicas permite además una estimación más precisa del efecto de un factor en el experimento y la apreciación de las interacciones.

Los métodos estadísticos requieren que las observaciones o los errores sean variables aleatorias independientes. Por esta razón el experimento se debe aleatorizar adecuadamente, es decir que la asignación del material experimental y el orden en que se realizan las pruebas individuales se deben determinar aleatoriamente.

El análisis por bloques permite incrementar la precisión del experimento, siendo un bloque una porción del material experimental más homogénea que el total del material. Al realizarse un análisis por bloques, se hacen comparaciones entre las condiciones de interés del experimento dentro de cada bloque.

La aplicación del Diseño de Experimentos requiere considerar las siguientes etapas (Montgomery, 1991):

1- Comprensión y planteamiento del problema.

Al diseñar y analizar un experimento es necesario tener de antemano una idea clara de qué es exactamente lo que se va a estudiar y cuales son los objetivos del experimento. El *DE* es una herramienta para encontrar respuestas a problemas perfectamente identificados y especificados. Cuanto más claramente se plantea el problema y se identifica el propósito o información que se desea conseguir con los experimentos, mayor puede ser la ayuda del *DE*.

Para obtener una comprensión profunda del sistema y del problema es necesario recopilar toda la información disponible sobre el sistema en estudio y que pueda ser relevante para la experimentación que se realizará.

2- Selección de la variable de respuesta.

Se debe seleccionar como respuesta una variable que provea la información necesaria. Según el objetivo perseguido, puede ser necesario observar más de una respuesta y encontrar un compromiso entre ellas.

3- Elección de factores y niveles.

Se deben identificar y listar todos los factores que pueden tener influencia en el proceso y en la respuesta. Se debe tener conciencia de la influencia potencial que podría tener cada factor en la respuesta.

Para cada factor se debe definir el dominio experimental, es decir el intervalo de valores que puede tomar. También se debe considerar la forma en que se controlarán dichos factores para mantenerlos en los valores deseados y cómo se les medirá. Para esto es necesario conocer el proceso estudiado, mediante una combinación de experiencia práctica y comprensión teórica.

4- Elección del diseño experimental.

La experimentación se realiza frecuentemente en etapas secuenciales. La elección del diseño experimental para cada etapa depende de una serie de factores a considerar:

- * Naturaleza del problema, información conocida y tipo de información que se desea obtener según el objetivo planteado.
- * Número de factores e interacciones que se deben estudiar.
- * Complejidad de utilizar cada diseño.
- * Validez estadística y efectividad de cada diseño.
- * Facilidad de comprensión e implementación.
- * Restricciones operativas, de costo y tiempo.

5- Realización de la experimentación.

En esta etapa es necesario asegurar que todo se haga de acuerdo a lo planeado a los fines de evitar errores en el procedimiento que puedan anular la validez experimental.

6- Análisis de datos.

Deben emplearse métodos estadísticos para analizar los datos.

Una vez que se dispone de los resultados experimentales se pueden calcular los efectos de los factores, así como sus interacciones. Los métodos estadísticos permiten comprobar si los efectos calculados son significativos comparándolos con el error experimental. Si se construye un modelo de superficie de respuesta, se pueden calcular los coeficientes por el método de los cuadrados mínimos y se puede evaluar el modelo realizando réplicas de ciertos experimentos y aplicando ANOVA. Este modelo se puede usar para buscar la zona óptima matemáticamente.

7- Conclusiones y recomendaciones.

Es importante notar que la realización de estas etapas es cíclica. La información obtenida al realizar una serie de experimentos se debe integrar para planificar la experimentación posterior. Quizás entonces se comprende mejor el problema y se pueden redefinir o concretar más los objetivos, se pueden descartar factores que no eran importantes o modificar su dominio experimental. Con esto se planea una nueva experimentación y así sucesivamente.

1.5.3. OPTIMIZACIÓN EXPERIMENTAL

La optimización puede definirse como un recurso basado en instrucciones que permiten mejorar los resultados de un procedimiento. Científicamente estas instrucciones se expresan a través de métodos matemáticos que permiten maximizar o minimizar alguna propiedad específica del sistema en estudio. Esta propiedad a ser optimizada se denomina *función objetivo* o respuesta y se simboliza con la letra y .

En muchos métodos analíticos, la respuesta del sistema de medida depende de una variedad de factores experimentales que se simbolizan con la letra x , bajo el control del operador, tales como: temperatura, pH, fuerza iónica, composición química de las soluciones, tiempos de reacción y concentración de los componentes de los reactivos, entre otros.

Es decir que existen dos tipos de variables: las respuestas y los factores; las respuestas son las variables dependientes y los factores las variables independientes. La relación entre ambas variables se puede representar según la siguiente ecuación:

$$(y_1, y_2, \dots, y_p) = f(x_1, x_2, \dots, x_n) \quad (1.5)$$

En la mayoría de los casos, cada respuesta se trata separadamente y se tiene:

$$y = f(x_1, x_2, \dots, x_n) \quad (1.6)$$

El proceso de búsqueda de los niveles óptimos de estos factores se conoce como optimización.

El procedimiento de optimización se lleva a cabo en las siguientes etapas (Kellner, 1998):

1- Definición de la función objetivo o respuesta. A veces el criterio de optimización es simplemente una señal analítica o un tiempo de análisis, otras veces el rendimiento de un método analítico se optimiza respecto a criterios de calidad analítica tales como: precisión, exactitud, sensibilidad, límite de detección o relación señal/ruido. Hay casos más complicados de respuestas múltiples donde la función objetivo puede estar compuesta de varios criterios.

2- Determinación de los factores más importantes, es decir de los factores que presentan influencias significativas sobre la respuesta a optimizar. Para esto se utiliza un diseño experimental simple o diseño de cribaje (*screening*).

3- Optimización propiamente dicha. Determinación de la combinación de los niveles de los factores seleccionados que resulte en la mejor respuesta (maximización o minimización). Para esto se utilizan diseños experimentales secuenciales o simultáneos.

1.5.3.1. DISEÑOS DE CRIBADO (SCREENING-ESCRUTINIO-TAMIZADO)

- FACTORIAL COMPLETO A DOS NIVELES

Por diseño factorial se entiende aquel en el que se investigan todas las posibles combinaciones de los niveles de los factores en cada experiencia. Se utilizan para estudiar el efecto conjunto de varios factores sobre una respuesta y se basan en la variación simultánea de un número limitado de niveles de los factores.

Los planeamientos factoriales del tipo 2^k , donde k es el número de factores y 2 el número de niveles, son los más comunes. El diseño requiere un experimento para cada una de las posibles combinaciones de los 2 niveles y de los k factores considerados. Si se trabaja con dos factores, el número de experimentos asciende a cuatro y la matriz de experimentos para un diseño de este tipo se presenta en la Tabla 1.2. El diseño espacial en dos dimensiones que se obtiene se presenta en la Figura 1.1, donde cada eje representa los niveles de un factor.

Tabla 1.2: Matriz de experimentos para un diseño factorial completo 2^2

	x_1	x_2
1	+	-
2	+	+
3	-	-
4	-	+

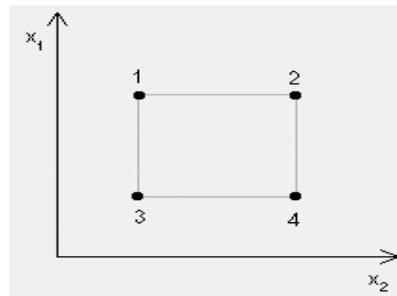


Figura 1.1: Diseño experimental para 2 factores a 2 niveles

Para el diseño general 2^k el número de experimentos a realizar es $n = 2^k$, es decir que cuando los factores aumentan, también aumenta el número de experimentos.

Se acostumbra a representar los niveles de manera codificada usando los signos (+) y (-).

La elección exacta de los niveles se determina principalmente por la experiencia y el conocimiento del experimentador, quien basándose en la experiencia previa, la bibliografía y las necesidades de experimentación, debe escoger qué factores estudiar y qué valores pueden tomar (dominio experimental). Cuando los factores son continuos, su dominio experimental se expresa con los valores máximo y mínimo que pueden tomar.

Si se tiene una variable cualitativa, “alto” y “bajo” se refiere a un par de condiciones diferentes, por ejemplo presencia o ausencia de una sustancia.

Combinando los niveles de los factores se obtiene la *matriz de experimentos*, que luego se concreta en el *plan de experimentación*, cuando se sustituyen los valores de las variables codificadas por los valores de las variables reales.

El modelo estadístico de diseño 2^k incluye k efectos principales, $k/2$ interacciones de dos factores, $k/3$ interacciones de tres factores, ..., y una interacción de k factores, es decir se analizan 2^{k-1} efectos. El problema que aparece en el diseño factorial completo es que el número de experimentos crece rápidamente con el número de factores. Para k factores con dos réplicas para cada combinación de valores, son necesarios 2^{k+1} experimentos.

Para obtener la máxima información deben tenerse algunos cuidados, como son la necesidad de hacer réplicas, es decir repeticiones de algunos ensayos para poder estimar el error experimental. Las réplicas deben ser repeticiones auténticas y todos los ensayos se deben realizar en forma aleatoria. El número de réplicas que un experimentador puede realizar está limitado frecuentemente por los recursos

disponibles. Cuando se hace una repetición individual de un diseño 2^k , se tiene lo que se conoce como diseño factorial no replicado. Con sólo una réplica no hay estimación de error. Una aproximación al análisis de un factorial no replicado consiste en suponer que ciertas interacciones de orden superior son despreciables, y combinar sus cuadrados medios para estimar el error.

El cálculo de los efectos principales y de los efectos de interacción entre factores, puede hacerse manualmente (Strange, 1990) o utilizando un programa de diseño experimental. Los cálculos manuales pueden simplificarse utilizando el algoritmo de Yates.

Para determinar si los efectos calculados son significativos se puede utilizar el ANOVA, calculando las sumas de cuadrados a partir de los efectos estimados. Se comparan luego los cuadrados medios con el cuadrado medio del error (residual). El análisis de la significancia de los efectos también se puede hacer utilizando un programa de diseño experimental.

- FACTORIAL FRACCIONADO

En el diseño factorial, el número de experimentos puede ser muy elevado. Cuando hay más de 3 factores es posible realizar una simplificación suponiendo que las interacciones de orden 3 y superiores son despreciables. Es decir es posible efectuar un diseño factorial parcial, sin que sea necesario la determinación de todos los parámetros de interacción. De esta forma se puede disminuir el número de experimentos y aún determinar los efectos más importantes (principales y de interacciones de segundo orden). La lógica de esta aproximación es que los efectos de orden más grande son normalmente más pequeños que los efectos de los factores principales y los efectos de interacción entre dos factores.

Si las interacciones de orden superior se pueden suponer insignificantes, es suficiente una fracción adecuada de todas las posibles combinaciones de los niveles de los factores.

Este nuevo tipo de diseño debería conservar las siguientes características:

- Ser balanceado en niveles
- Mapear el dominio experimental tan bien como sea posible
- Conservar la ortogonalidad

Un diseño experimental de este tipo, es el diseño factorial fraccionado o incompleto, donde el número de experimentos se reduce por un número p (entero menor que k) de acuerdo a la expresión:

$$n = 2^{k-p} \tag{1.7}$$

Si $p=1$, se obtiene un diseño fraccionado al medio. Por ejemplo, para un diseño factorial completo para tres factores se necesitan $2^3=8$ experiencias y para el diseño fraccionado al medio se necesitan $2^{3-1}=4$ experiencias.

Se pueden hacer diseños aún más chicos (1/4, 1/8), hasta alcanzar el diseño fraccionado “saturado”, con el número mínimo posible de experimentos ($n \geq k$).

Con este tipo de diseños, se produce una pérdida de información debida al menor número de datos y cuanto más fraccionado es el diseño, mayor es la cantidad de factores que se “confunden” entre sí.

- PLACKETT-BURMAN

Estos diseños son factoriales fraccionados de dos niveles utilizados para estudiar hasta un máximo de $k = N-1$ factores en N ensayos, donde N es múltiplo de 4.

Este tipo de diseño se utiliza cuando se quiere estudiar el efecto de muchos factores y se caracteriza porque, si todas las interacciones son despreciables, entonces los efectos principales no están confundidos y tienen las variancias más pequeñas posibles.

Las experiencias surgen de una primera línea generadora del diseño, a partir de la cual se agregan las siguientes por permutación cíclica de esta y al final se agrega una línea completa de niveles “-”.

1.5.3.2. DISEÑOS DE OPTIMIZACIÓN

- SECUENCIALES

La optimización secuencial está basada en un diseño inicial de experimentos con sus respectivas mediciones y resultados, a raíz de los cuales se diseña otra serie de experiencias hasta encontrar en forma ascendente o descendente el punto óptimo.

- MÉTODO UNIVARIADO

Este método de optimización, clasificado como un método secuencial, es también conocido como método del factor único o de estrategia de un factor por vez (*en inglés, one variable at a time, OVAT*). En este método, se fijan a un cierto nivel todos los factores que están siendo investigados, menos uno de ellos. Entonces este último es variado hasta que se encuentre la mejor respuesta. Después de esto este factor se fija y un nuevo factor sufre variación. Este proceso se repite hasta que todos los factores hayan sido adecuados para dar la mejor respuesta.

Dado que un ciclo de variaciones no es suficiente para definir el óptimo con precisión, se deben efectuar tantos ciclos como sean necesarios hasta observar que no ocurre mejora en el resultado de la optimización.

Este tipo de procedimiento es probablemente el método de optimización más comúnmente usado en química, sin embargo no asegura que la región óptima pueda ser localizada (se considera como una mala práctica de optimización).

Este método presenta el inconveniente de que cuanto mayor sea la interacción entre las variables, mayor será la probabilidad de localización de un falso óptimo. Por otra parte, la localización del óptimo real depende de los valores iniciales escogidos para las variables a ser optimizadas.

- **MÉTODO SIMPLEX**

El método Simplex es un método secuencial que se utiliza para maximizar o minimizar una respuesta.

Un simplex es una figura geométrica de n dimensiones, constituido de $n+1$ puntos (Miller, 1993). Cada dimensión corresponde a una variable a ser optimizada. Un simplex en dos dimensiones es un triángulo, en tres dimensiones es un tetraedro y así sucesivamente (Deming, 1973). El método puede ser extendido para mayores dimensiones, pero se complica la representación gráfica del simplex.

El método simplex no se basa en planeamientos factoriales, requiere pocos experimentos para moverse, desplazándose en la dirección del óptimo. El procedimiento de optimización comienza con la elección de los $n+1$ puntos donde será efectuada la evaluación de la respuesta. No hay necesidad de conocer previamente la superficie de respuesta, ya que sólo interesa la respuesta que será obtenida en cada experimento previsto por el simplex. Este resultado será evaluado contra las demás respuestas para que el proceso pueda continuar.

En la Figura 1.2 se observa una representación esquemática de una optimización por el método simplex para dos factores, donde la figura geométrica que en este caso es un triángulo, se representa en un *gráfico de contorno*, formado por curvas de igual valor de respuesta (a semejanza de las curvas de altitud en un mapa geográfico). El óptimo se ubica en el centro de la elipse menor.

El simplex inicial está definido por los puntos 1, 2 y 3. En el primer experimento la respuesta se mide en cada una de las combinaciones de los niveles de los factores dadas por los vértices del simplex.

En este caso, la peor respuesta se encontraría en el punto 3 y es lógico concluir que se obtendría una mejor respuesta en el punto simétrico de 3 respecto a la línea que une 1 y 2, es decir en 4. Si ahora se comparan las respuestas alcanzadas en los puntos 1,2 y 4, se observa que 1 da la peor respuesta. El procedimiento anterior, basado en la reflexión sobre una recta, se repite de nuevo, con lo cual el simplex queda definido por los puntos 2,4 y 5. En la figura 2.1 se observa la continuación de este proceso y se ve que no es posible progresar más allá porque los puntos 6 y 8 dan peor respuesta que el 5 y 7.

Como se ve, el objetivo de este método es forzar al simplex a moverse hacia la región de respuesta óptima, descartando siempre la peor respuesta.

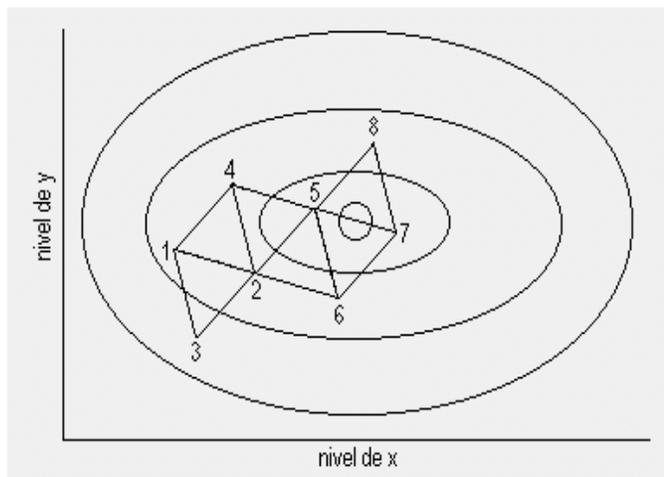


Figura 1.2: Optimización por el método simplex

- SIMULTÁNEOS

En la optimización simultánea, la relación entre la respuesta y los factores se estudia llevando a cabo un diseño experimental, construyendo un modelo matemático e investigando la relación mediante un método de superficie de respuesta (*RMS, del inglés Responce Surface Method*).

La metodología de superficie de respuesta es un conjunto de técnicas matemáticas y estadísticas muy utilizada para modelar y analizar problemas en los cuales una respuesta de interés es influenciada por varias variables, siendo el objetivo optimizar esa respuesta.

La relación entre la respuesta y los efectos de los factores puede estar basada en modelos físicos, fisicoquímicos o en modelos empíricos tales como polinomios o ecuaciones de segundo orden. Estos modelos matemáticos pueden ser capaces de describir una superficie de respuesta lineal o curvilínea.

Si se tienen dos factores x_1 y x_2 , la respuesta y se relaciona con los factores de la siguiente manera:

$$y = f(x_1, x_2) + \varepsilon \quad (1.8)$$

donde ε representa el ruido o error observado en la respuesta y . Si la respuesta esperada se denota por $E(y) = f(x_1, x_2) = \eta$, entonces la superficie de respuesta queda representada por la siguiente ecuación:

$$\eta = f(x_1, x_2) \quad (1.9)$$

Como puede observarse en la Figura 1.3, la superficie de respuesta se representa como una superficie sólida en un espacio tridimensional y para visualizar mejor la forma de una superficie de respuesta se grafican los contornos de dicha superficie.

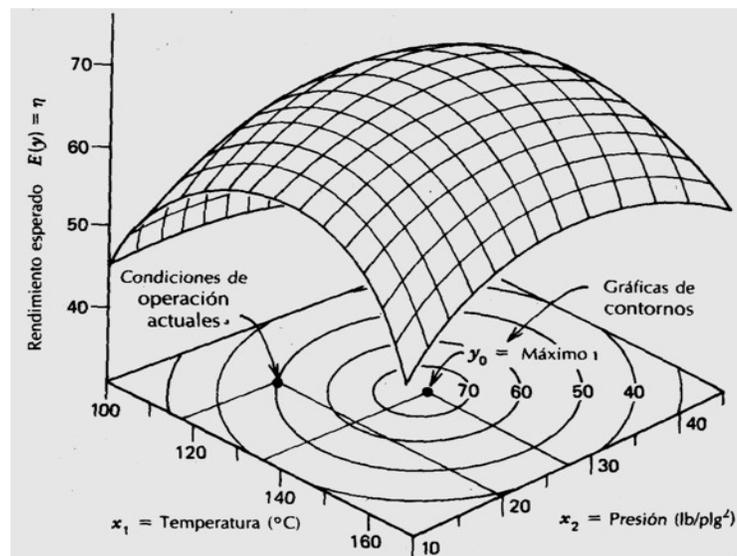


Figura 1.3: Superficie de respuesta tridimensional y su gráfica de contorno

Esta gráfica de contorno se compone de líneas de respuesta constante, que corresponden a una altura específica de la superficie de respuesta.

Generalmente la forma de la relación entre la respuesta y las variables independientes se desconoce, por ello el primer paso en un método de superficie de respuesta es determinar una aproximación a la relación real entre la respuesta y el conjunto de variables independientes. Por lo general se usa un polinomio de orden bajo sobre alguna región de las variables independientes.

Si la respuesta se describe por una función lineal de las variables independientes, la función de aproximación es el modelo de primer orden y queda representada por:

$$\eta = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \varepsilon \quad (1.10)$$

Cuando existe curvatura en el sistema debe usarse un polinomio de mayor grado, por ejemplo un modelo de segundo orden:

$$y = \beta_0 + \sum_{i=1}^k \beta_i x_i + \sum_{i=1}^k \beta_{ii} x_i^2 + \sum_{i < j} \sum_j \beta_{ij} x_i x_j + \varepsilon \quad (1.11)$$

Para el caso de dos factores, el diseño conduce a la estimación de los coeficientes β_i , a través de la siguiente ecuación:

$$\hat{y} = b_0 + b_1 x_1 + b_2 x_2 + b_{12} x_1 x_2 + b_{11} x_1^2 + b_{22} x_2^2 \quad (1.12)$$

La precisión de los coeficientes b_i para estimar los β_i depende del diseño experimental y cuando estos coeficientes se obtienen por regresión múltiple, se pueden usar para predecir la respuesta y como una función de los factores x y construir superficies de respuesta y gráficas de contorno.

Para determinar los niveles de los k factores que optimizan la respuesta, se deberá calcular el conjunto x_1, x_2, \dots, x_k tal que sus derivadas parciales sean cero. Dicho punto se denomina punto estacionario y podría representar un punto de respuesta máxima, un punto de respuesta mínima o un punto silla.

Por supuesto que, es improbable que el modelo polinomial sea una aproximación razonable de la relación funcional real sobre todo el dominio de las variables independientes, pero funciona muy bien en regiones relativamente pequeñas.

El método de cuadrados mínimos sirve para estimar los parámetros del polinomio de aproximación y el análisis de la superficie de respuesta se hace luego en términos de superficie ajustada. El ajuste y el análisis de superficie de respuesta se facilitan con la elección adecuada de un diseño experimental, siendo las características deseables para el mismo las siguientes (Montgomery, 1991):

- 1) Proporcionar una distribución razonable de puntos de datos (y por lo tanto información) en toda la región de interés.
- 2) Permitir investigar la idoneidad del modelo, incluyendo la falta de ajuste.
- 3) Permitir la realización de experimentos en bloques.
- 4) Permitir la construcción secuencial de diseños de orden superior.
- 5) Proporcionar una estimación de error interna.
- 6) No requerir un número grande de corridas.
- 7) No requerir demasiados niveles de las variables independientes.
- 8) Asegurar la simplicidad de los cálculos de los parámetros del modelo.

En química analítica, la relación entre la respuesta y los efectos de los factores se puede plantear mediante modelos empíricos generales de segundo orden y los diseños experimentales que se pueden usar para ajustar un modelo de segundo orden deben tener por lo menos tres niveles de cada factor.

La rotabilidad es una propiedad muy importante de los diseños experimentales que se usan para ajustar superficies de respuesta. Se dice que un diseño experimental es rotable o girable si la variancia de la respuesta predicha y en algún punto x , es función sólo de la distancia al punto desde el centro del diseño y no es una función de la dirección.

Los diseños experimentales que se podrían utilizar para ajustar un modelo de segundo orden se detallan a continuación.

- **FACTORIAL COMPLETO A 3 NIVELES**

El diseño 3^k , es un diseño que consta de k factores con tres niveles cada uno: nivel inferior (0), nivel intermedio (1) y nivel superior (2). Este diseño se utiliza cuando se quiere estimar el efecto cuadrático, es decir que la adición de un tercer nivel permite modelar con una relación cuadrática la relación entre la respuesta y cada factor, pero excepto para pequeños k , el diseño requiere muchos experimentos.

Además, se debe tener en cuenta que el diseño 3^k no es una buena elección como diseño de superficie de respuesta de segundo orden, ya que no es rotable.

- **CENTRAL COMPUESTO**

El diseño central compuesto es probablemente, el diseño experimental más utilizado para ajustar superficies de respuesta de segundo orden. Este diseño consiste en un factorial completo (o factorial fraccionado) 2^k , aumentado por $2k$ puntos axiales y n_0 puntos centrales. Es decir que este diseño en general consiste de tres partes:

- Un diseño factorial completo a dos niveles, que constituye la parte cúbica del diseño y donde el número de puntos o experimentos está dado por $n_c = 2^k$ con niveles -1 y $+1$.
- Un diseño estrella, que permite agregar más niveles y describir curvatura, adicionando puntos que se describen como una estrella y que están situados a una distancia α desde el centro del diseño. El número puntos o experimentos está dado por $n_s = 2k$ y $\alpha = (n_s)^{1/4}$. Los niveles son $-\alpha$ y $+\alpha$.
- Puntos centrales con todos los niveles iguales a cero, del que generalmente se realizan varios replicados. El número de experimentos es n_0 . Esta replicación del punto central permite tener un idea inmediata del error experimental y si se elige un valor adecuado de n_0 , se puede alcanzar la cuasi-ortogonalidad y en este caso se alcanza también la ortogonalidad.

De acuerdo a lo previamente dicho, se puede ver que cada factor se encuentra en cinco niveles posibles: $-\alpha, -1, 0, +1, +\alpha$ y que el número total de experimentos a realizar es:

$$n = n_c + n_s + n_0 \tag{1.13}$$

Como se ve en la Figura 1.4 (derecha), el diseño central compuesto para dos factores implica la realización de nueve experimentos, si el punto central es un único replicado; y para tres factores (izquierda) el número total de experimentos sería igual a quince. Este diseño permite mejorar la economía de experimentos, ya que se trabaja con un número significativamente menor a los utilizados en los diseños factoriales.

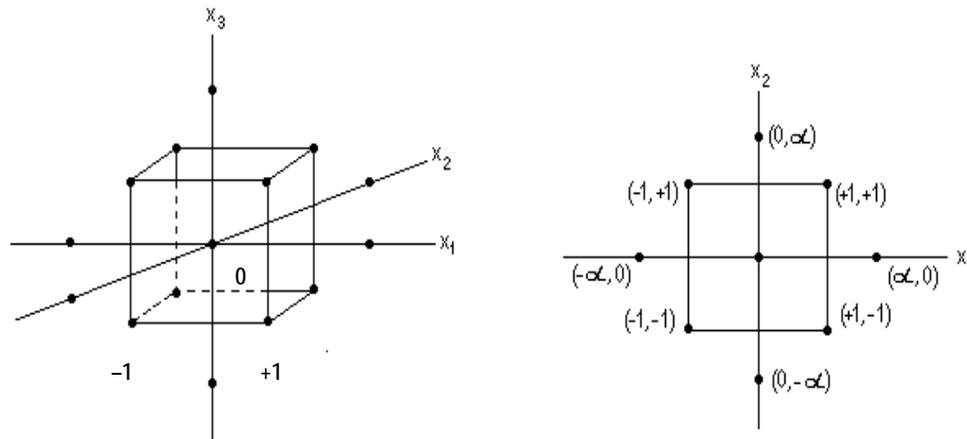


Figura 1.4: Diseños Central Compuestos. Derecha: para dos factores, donde los puntos situados a una distancia α del centro conforman una estrella; izquierda: para tres factores, donde los niveles están indicados sólo para x_1

Una variante del diseño central compuesto es el diseño central compuesto centrado en las caras, en el cual α toma el valor 1. Como puede verse en la Figura 1.5, en este diseño los puntos estrella se localizan en los centros de las caras del cubo. Esta variante del diseño central compuesto, se emplea ocasionalmente, debido a que sólo propone tres niveles de cada factor. Este diseño tiene la desventaja de que no es rotable.

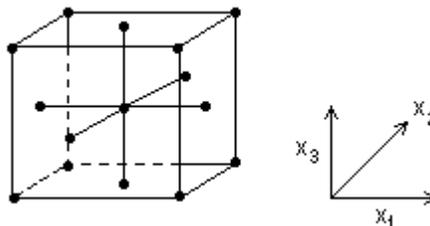


Figura 1.5: Diseño Central Compuesto centrado en las caras para $k=3$

- **BOX-BEHNKEN**

Este diseño fue propuesto por Box y Behnken para ajustar superficies de respuesta y considera que todos los factores tienen exactamente tres niveles (alto, bajo y central). Se interpreta geoméricamente marcando los puntos medios de las aristas de un cubo más un punto central o como el intercalado de tres diseños factoriales de dos niveles más un punto central.

Este diseño suele ser más eficiente en términos del número de corridas (requiere pocas combinaciones de factores), es rotable (o casi rotable) y se forma combinando factoriales 2^k con diseños de bloques incompletos.

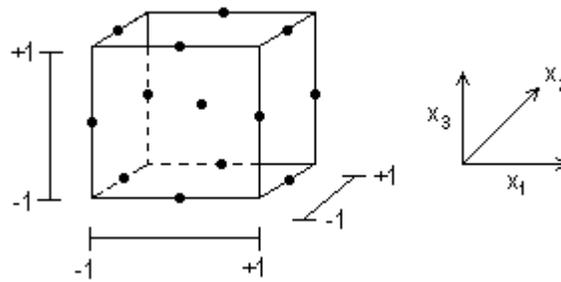


Figura 1.6: Diseño de Box-Behnken para tres factores

Como se puede observar en la Figura 1.6, el Diseño de Box-Behnken es esférico y no contiene puntos en los vértices de la región cúbica (no comprende las esquinas del cubo), sino que se trabaja en la parte central de las aristas del cubo.

- DISEÑO DE MEZCLAS

Los diseños de superficie de respuesta presentados anteriormente, se utilizan en aquellas situaciones donde los niveles de cada factor son independientes de los niveles de los otros factores (Barros, 2001). Cuando los factores son los componentes o ingredientes de una mezcla, sus niveles no son independientes y se deben utilizar diseños de mezcla. En estos diseños, los factores son los componentes de la mezcla y la respuesta es una función de la proporción de cada componente. Si se denota con x_1, x_2, \dots, x_k , las proporciones de k componentes de una mezcla, se tiene que:

$$0 \leq x_i \leq 1 \text{ para } i = 1, 2, \dots, k$$

y

$$x_1 + x_2 + \dots + x_k = 1 \text{ (es decir 100 \%)}$$

El espacio factorial incluye todos los valores de los componentes que se encuentran en el segmento de línea $x_1 + x_2 = 1$, para el caso de dos componentes (cada componente está acotado por cero y uno). Para el caso de tres componentes, el espacio muestral es un triángulo cuyos vértices corresponden a formulaciones que son los componentes puros, es decir mezclas consistentes en 100 % de un solo componente. Cada uno de los lados representa una mezcla que carece en absoluto de uno de los tres componentes (el componente indicado en el vértice opuesto).

Los diseños más usados para estudiar los efectos de los componentes de mezclas en la variable respuesta son los diseños simplex, que fueron introducidos por Scheffé (Scheffé, 1958).

Existen dos alternativas de trabajo diferentes, a saber:

* Diseño de red simplex (k,m) : para k componentes, cada uno de los cuales puede tomar $m + 1$ niveles igualmente espaciados entre cero y uno. Para una ecuación de primer orden, $m = 1$ y los niveles son 0 y 1, mientras que para una ecuación de segundo orden, $m = 2$ y los niveles son $x_i = 0, \frac{1}{2}, 1$.

En la Figura 1.7, se presentan ejemplos de diseños de red simplex para tres y cuatro componentes.

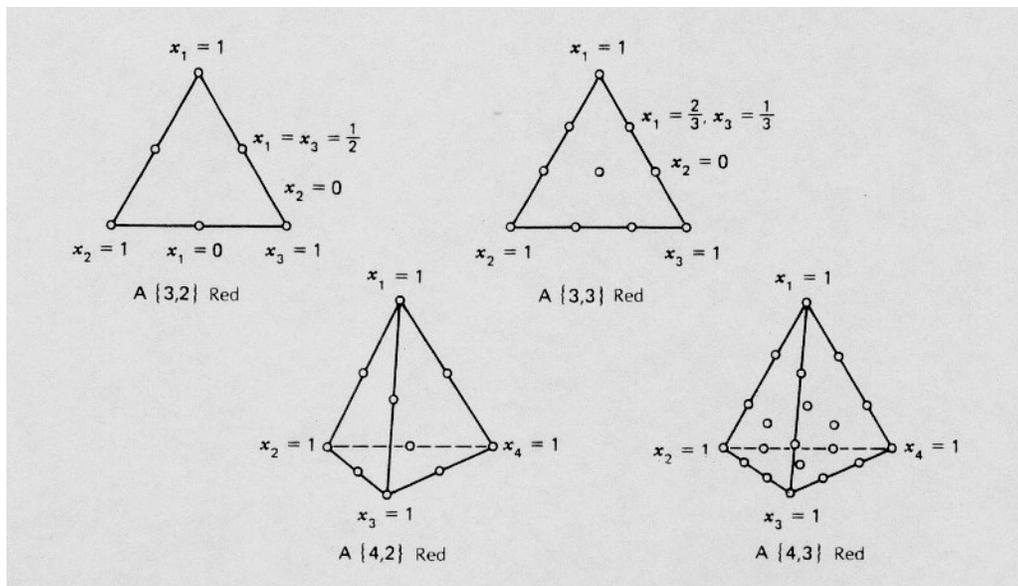


Figura 1.7: Diseños de red simplex para $k=3$ y $k=4$ componentes

* Diseño de centroide simplex: en este diseño hay $2^k - 1$ puntos, que corresponden a las k permutaciones de $(1,0,0,\dots,0)$, las $\binom{k}{2}$ permutaciones de $(\frac{1}{2}, \frac{1}{2}, 0, \dots, 0)$, la $\binom{k}{3}$ permutación de $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, 0, \dots, 0)$, ... y el centroide global $(\frac{1}{k}, \frac{1}{k}, \dots, \frac{1}{k})$.

En la Figura 1.8, se presentan ejemplos de diseños de centroide simplex para tres y cuatro componentes.

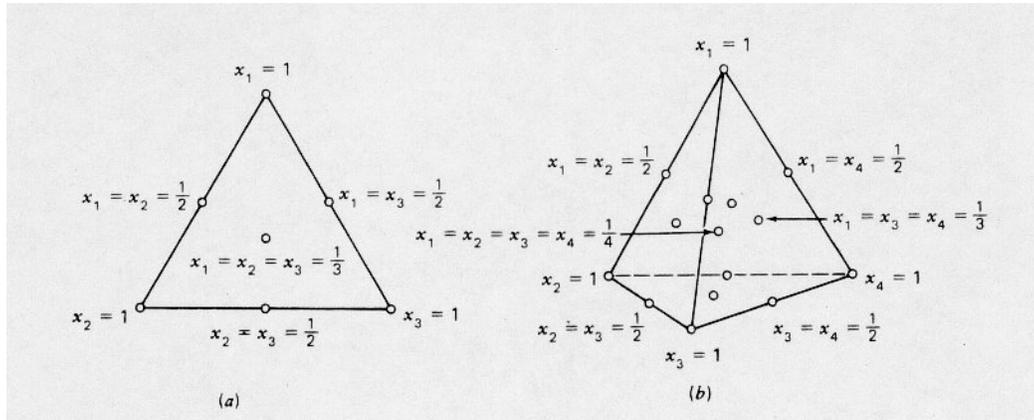


Figura 1.8: Diseños de centroide simplex para $k=3$ (a) y $k=4$ (b) componentes

Los modelos de mezcla difieren de los polinomios usualmente empleados para ajustar superficies de respuesta debido a la restricción $\sum x_i = 1$, siendo los de uso más difundido: el lineal, el cuadrático, el cúbico completo y el cúbico especial, cuyas ecuaciones se presentan a continuación:

* Lineal:

$$y = \sum_{i=1}^k \beta_i x_i \tag{1.14}$$

* Cuadrático:

$$y = \sum_{i=1}^k \beta_i x_i + \sum_{i<j}^k \beta_{ij} x_i x_j \tag{1.15}$$

* Cúbico completo:

$$y = \sum_{i=1}^k \beta_i x_i + \sum_{i<j}^k \beta_{ij} x_i x_j + \sum_{i<j}^k \delta_{ij} x_i x_j (x_i - x_j) + \sum_{i<j<p} \beta_{ijp} x_i x_j x_p \tag{1.16}$$

* Cúbico especial:

$$y = \sum_{i=1}^k \beta_i x_i + \sum_{i<j}^k \beta_{ij} x_i x_j + \sum_{i<j<p} \beta_{ijp} x_i x_j x_p \tag{1.17}$$

1.6. VALIDACIÓN DE MÉTODOS ANALÍTICOS

1.6.1. CONCEPTO DE VALIDACIÓN

Los laboratorios deben demostrar que sus métodos analíticos proporcionan resultados fiables y adecuados para su finalidad o propósito perseguido (ISO/IEC 17025), ya que muchas de las decisiones que se toman están basadas en la información que estos resultados proporcionan.

Validar un método analítico consiste en verificar y documentar su validez, esto es su adecuación a determinados requisitos, previamente establecidos, para poder resolver un problema analítico en particular.

La *validación* puede definirse como un proceso basado en estudios sistemáticos de laboratorio, mediante el cual se pone de manifiesto que un método analítico determinado posee unas características de funcionamiento adecuadas a la aplicación que se le quiere dar (EURACHEM, 1998).

Se pueden distinguir dos tipos diferentes de validación de métodos, la *validación interna* y la *validación interlaboratorios*. La validación interna es la que se circunscribe al ámbito de un único laboratorio, en cambio la validación interlaboratorios, también llamada validación externa, implica la realización de los llamados ejercicios colaborativos (Massart, 1997).

La validación interlaboratorios es la manera más apropiada de evaluar un método analítico, siendo el objetivo de los ejercicios colaborativos, el de comprobar la idoneidad de una nueva metodología o de un método recientemente modificado. En este tipo de ejercicios, participan laboratorios que han debido demostrar su competencia, analizando una misma muestra concreta, siguiendo tan fielmente como sea posible el procedimiento analítico propuesto.

Al validar un método analítico se deben tener en cuenta tres aspectos importantes (Massart, 1997):

- * Se debería validar el proceso analítico en su conjunto, incluyendo las etapas de tratamiento de la muestra (disolución, extracción, preconcentración), previas a la medida analítica.
- * Se debería validar el método en todo el intervalo de concentraciones en que se va a aplicar.
- * Se debería validar el método utilizando un juego representativo de las diferentes matrices que se van a analizar.

Diversos organismos internacionales tales como la *AOAC* (Association of Oficial Analytical Chemists), la *USP* (United States Pharmacopoeia), la *FDA* (Food and Drug Administration) y la *ICH* (Tripartite International Conference on Harmonization), que tienen por objeto la elaboración y difusión de métodos

analíticos de referencia para distintos campos de actividad, ofrecen criterios para la validación de métodos. En la práctica, los protocolos de validación se diseñan teniendo en cuenta el origen, el ámbito de aplicación y el tipo de método.

1.6.2. PARÁMETROS DE CALIDAD

Los métodos analíticos o los procesos de medida, se pueden caracterizar mediante parámetros de calidad o criterios de calidad, que pueden ser de tres tipos (Ramis Ramos, 2002):

- *Estadísticos*
- *Operativos*
- *Económicos*

Los parámetros de calidad que con mayor frecuencia se utilizan para la caracterización de los métodos analíticos y se evalúan en los procedimientos de validación son los que corresponden a los criterios estadísticos:

- ✓ *Precisión*
- ✓ *Exactitud, Veracidad, Sesgo*
- ✓ *Linealidad*
- ✓ *Rango*
- ✓ *Límite de detección*
- ✓ *Límite de cuantificación*
- ✓ *Selectividad*
- ✓ *Sensibilidad*
- ✓ *Robustez*

Estos criterios se denominan estadísticos porque la verificación de los mismos en procedimientos de ensayo, constituye una aplicación importante de la Estadística descriptiva y de la Inferencia estadística y aparecen también en la bibliografía como características de funcionamiento (en inglés, *performance characteristics*) y cifras de mérito (en inglés, *figures of merit*).

1.6.2.1. PRECISIÓN

La *precisión* es una medida del tamaño del error aleatorio (en inglés, *random errors*), y desde el punto de vista estadístico mide la dispersión de los resultados alrededor de la media, sin considerar si esta última es una representación correcta del valor verdadero (Massart, 1997).

Según la *IUPAC* (International Union of Pure and Applied Chemistry) (Thompson, 2002) la precisión se define como el grado de concordancia entre los resultados de ensayos independientes obtenidos en unas condiciones bien definidas. Dado que el número de factores que pueden afectar a la precisión de los resultados es elevada, al evaluar la precisión de un método analítico es conveniente distinguir entre condiciones de mínima variación de los factores de influencia y condiciones de máxima variación de los mismos.

De esta manera, se pueden distinguir dos tipos extremos de precisión, llamadas *repetibilidad* y *reproducibilidad*. La repetibilidad es la precisión obtenida cuando las medidas se llevan a cabo en las mejores condiciones posibles (mismo laboratorio, mismo analista, mismos reactivos, mismos equipos y en un intervalo corto de tiempo) y la reproducibilidad es la precisión obtenida cuando las medidas se llevan a cabo en las condiciones más adversas posibles (diferentes laboratorios, diferentes analistas, diferentes reactivos, diferentes equipos).

Existen situaciones intermedias, en donde las mediciones se realizan dentro del mismo laboratorio, pero variando alguno de los factores antes mencionados (operador, equipo, reactivos, tiempo). En este caso la precisión obtenida se denomina *precisión intermedia*.

Para evaluar la repetibilidad de un método analítico se pueden realizar de seis a ocho determinaciones replicadas, dentro de una misma corrida y por el mismo analista. El resultado se puede expresar como la desviación estándar, σ , estimada analíticamente por s ó como la desviación estándar relativa (*RSD*, del inglés *relative standard deviation*). También se puede utilizar el coeficiente de variación (*CV %*). En algunos trabajos se sugiere repetir la experiencia por lo menos tres veces y luego calcular una desviación estándar promedio.

Por otra parte, es importante tener en cuenta que todos los pasos del procedimiento deben ser replicados y si el blanco es una posible fuente de variación importante, también debe ser replicado. De este modo, el resultado engloba las fuentes de variancia asociadas a todas las etapas del método.

En ciertos casos puede resultar conveniente, medir separadamente la repetibilidad de un paso determinado del procedimiento, por ejemplo la repetibilidad de la inyección, en las determinaciones cromatográficas. Algunos autores hacen una distinción entre *precisión de un método* y *precisión de un*

sistema o precisión instrumental. Para el primer caso es necesario repetir todo el procedimiento, mientras que para el segundo se realizan mediciones replicadas de una preparación estándar, que esta lista para la medición directa y no necesita ningún tratamiento previo.

Otro aspecto que se debe considerar, es que la precisión habitualmente depende de la concentración del analito. Cuando el intervalo de concentraciones es amplio, con frecuencia se observa un fenómeno de heterocedasticidad, donde la desviación estándar aumenta con la concentración y el coeficiente de variación decrece y se estabiliza para altas concentraciones. Por esta razón se recomienda trabajar en tres niveles diferentes de concentración, los que deben incluir los límites superior e inferior del rango lineal.

Por otra parte, la precisión intermedia se obtiene cuando dentro de un laboratorio se varían uno o más factores entre cada uno de los ensayos. De acuerdo a los factores que se varíen, se pueden obtener diferentes tipos de precisión intermedia. Si los ensayos se hacen en diferentes días, se obtiene la precisión intermedia de días diferentes (en inglés, *time-different intermediate precision*). Los ensayos también pueden hacerse variando dos factores como el día y el operador (en inglés, *time+operator-different intermediate precision*) o el día y la calibración del equipo (en inglés, *time+instrument calibration-different intermediate precision*). Si se varían más factores (días, operadores, calibración del equipo, etc), se obtiene la precisión intermedia de días, operadores, calibraciones, etc (en inglés, *time+operator+instrument calibration+...-different intermediate precision*).

Para evaluar la precisión intermedia de días diferentes, se puede determinar la desviación estándar de una serie de mediciones replicadas, realizando una medición por día sobre una muestra homogénea, durante cinco a ocho días. También se pueden realizar experiencias que permiten evaluar repetibilidad y precisión intermedia en forma conjunta. En este caso se realizan replicados en días diferentes, por ejemplo, el NCCLS (*National Committee for Clinical Laboratory Standards*) recomienda duplicados durante 20 días. Luego se aplica ANOVA y se obtiene la repetibilidad (s_r) a partir del cuadrado medio residual (MS_{within}) y la precisión intermedia ($s_{between}$) de acuerdo a:

$$s_{between}^2 = (MS_{between} - MS_{within})/n_j \quad (1.18)$$

Otros organismos, como la SFSTP (*Société Française des Sciences et Techniques Pharmaceutiques*), recomiendan seis replicados durante tres días, u operadores, o instrumentos, de acuerdo al tipo de precisión intermedia que se quiera evaluar (Massart, 1997).

Si se necesita evaluar la precisión intermedia debida a varios factores, se deben utilizar diseños de experimentos anidados. Estos diseños permiten hacer una buena estimación de la precisión intermedia, determinando cuales son los factores que influyen más en la variabilidad de los resultados.

La precisión intermedia también puede obtenerse a partir de procedimientos alternativos, como pueden ser por ejemplo los gráficos de control. En este caso, el material de control se va analizando a lo largo del tiempo y se obtienen resultados que contemplan toda la variabilidad que puede afectar al método analítico, a pesar de que los factores no se hayan variado de forma ordenada, siguiendo un diseño anidado.

Los requerimientos de precisión dependen de la aplicación del método analítico. Para ciertas áreas específicas, tales como análisis clínicos, métodos bioanalíticos y análisis farmacéutico, se han establecido criterios de aceptabilidad.

Para el caso de los métodos bioanalíticos, es decir para los métodos que se usan para cuantificar drogas y sus metabolitos en una matriz biológica como sangre, suero, plasma u orina, se acepta un coeficiente de variación de hasta un 15 % y de hasta un 20 % para el Límite de Cuantificación (Swartz, 2003). Por otra parte, en análisis farmacéutico para la determinación de componentes mayoritarios por métodos cromatográficos, se acepta un coeficiente de variación de hasta un 1% para la precisión instrumental y un coeficiente de variación de hasta un 2 % para la repetibilidad. Si se trata de componentes minoritarios, tales como impurezas, se acepta un coeficiente de variación de hasta un 5 % para la precisión instrumental y de hasta un 10 % para la reproducibilidad (Green, 1996). También existen tablas donde se establecen coeficientes de variación máximos aceptables para un método analítico, en función del por ciento del analito en la muestra, según las teorías de Kolthoff y Hortwitz (Hortwitz,1982)

1.6.2.2. EXACTITUD

Los *errores sistemáticos* (en inglés, *systematic errors*) están caracterizados por términos tales como veracidad y sesgo (en inglés, *bias*) y están relacionados con el término exactitud.

No existe acuerdo en la comunidad científica internacional acerca de la definición de estos términos. Según ISO (ISO, 1993), la *veracidad* (en inglés, *trueness*) es el grado de concordancia entre el valor medio de una serie importante de resultados y un valor aceptado como de referencia y la medición de la veracidad se expresa como sesgo o bias. Es decir que veracidad es el concepto y el sesgo o bias la forma de medirla. Un resultado es veraz, si esta libre de error sistemático (o más rigurosamente, si los errores sistemáticos cometidos son aceptables). En la práctica la veracidad se verifica utilizando referencias, por lo tanto decir que un resultado es veraz, es equivalente a afirmar que es trazable a la referencia utilizada.

Por otra parte, según ISO la *exactitud*, es el grado de concordancia entre el resultado de un ensayo y un valor de referencia aceptado. De acuerdo a este concepto, el término exactitud describe una combinación de componentes de error aleatorio y sistemático. Pudiéndose afirmar entonces que la exactitud es la suma de dos conceptos, como son la veracidad y la precisión (en inglés, *trueness+precision*) y que un resultado es exacto si es veraz (se encuentra libre de errores sistemáticos) y preciso (los errores aleatorios son aceptables).

La estimación del *sesgo o bias*, depende de la extensión del rango de concentraciones del analito en la muestra, de la disponibilidad de un blanco de matriz (matriz a ser analizada sin el analito a determinar), de la posibilidad de adicionar analito a un material de este tipo de manera representativa y de la disponibilidad de materiales de referencia y métodos analíticos adecuados para la comparación de resultados.

Es importante señalar, que los errores sistemáticos pueden ser constantes (absolutos) o proporcionales (relativos) (Massart, 1997). El error constante se expresa en unidades de concentración y es un error sistemático independiente de la concentración del analito. Las principales fuentes de este tipo de error son la selectividad insuficiente (presencia de otros componentes que también producen una respuesta) y la utilización de blancos inadecuados. Por otro lado, el error proporcional depende de la concentración del analito y puede expresarse en unidades relativas, tales como porcentaje. Este tipo de error es causado fundamentalmente por la existencia del efecto matriz, es decir por el cambio de la pendiente de la curva de calibrado (alteración de la sensibilidad), producida por la matriz de la muestra. Además, la presunción incorrecta de linealidad sobre todo el rango de análisis, también puede causar este tipo de error relacionado a la concentración a determinar.

Para estimar el sesgo, se pueden utilizar dos procedimientos diferentes:

Procedimiento A: utilización de valores de referencia (materiales de referencia, muestras fortificadas, matrices fortificadas, adición estándar)

* **Materiales de referencia**

Según el *VIM* (Vocabulary of Basic and General Terms in Metrology), un *material de referencia* (*RM, del inglés Reference Material*) es un material o sustancia de la cual el valor de una o más de sus propiedades es suficientemente homogéneo y bien establecido como para ser usado en la calibración de un aparato, la validación de un método, o para asignar valores a materiales. En el ámbito analítico, los materiales de referencia se incluyen dentro de los patrones analíticos primarios, ya que poseen las características propias de los mismos (Compañó Beltrán, 2002).

Un material de referencia certificado (*CRM, del inglés Certified Reference Material*), se define como un material de referencia acompañado de un certificado, en el cual uno a más valores de sus propiedades están certificados por un procedimiento que establece su trazabilidad a una realización exacta de la unidad en la que se expresan los valores de la propiedad y para lo cual cada valor certificado se acompaña de una incertidumbre con la indicación del nivel de confianza.

Los materiales de referencia certificados pueden ser sustancias puras, sólidas o disueltas en un disolvente adecuado. Este tipo de materiales también son considerados patrones primarios.

Existen además los materiales de referencia certificados tipo matriz, en los que se certifican determinados componentes en matrices concretas, que pueden ser completamente naturales o naturales con los analitos añadidos. En este caso la certificación supone un contraste con patrones primarios, generalmente mediante un ejercicio interlaboratorio.

Un material de referencia debe cumplir una serie de requisitos para poderlo considerar como tal: debe garantizar la trazabilidad del valor de la concentración o el valor del parámetro que se quiera determinar, debe ser homogéneo y debe ser estable durante un tiempo razonable. Cuando se emplean para la validación de un método analítico, su matriz debería ser idéntica o muy similar a la de las muestras reales que se analizarán con el método.

Los principales inconvenientes de los materiales de referencia son su elevado precio y que no existen materiales de referencia certificados para todos los analitos y todos los tipos de matrices de muestras. Sólo se encuentra cubierto entre un 7% y un 10 % de las necesidades actuales de estos materiales.

*** Matrices Fortificadas-Muestras fortificadas-Adición Estándar**

Uno de los procedimientos más comúnmente utilizados para estimar el sesgo de un método analítico, consiste en calcular la recuperación de una cantidad exactamente conocida de analito adicionado.

Cuando se puede obtener el material a ser analizado sin el analito, es posible preparar muestras fortificadas (en inglés, *fortified sample or spiked sample*), adicionando el analito en una cantidad exactamente conocida.

Cuando se puede reproducir la matriz de la muestra, se preparan matrices fortificadas (en inglés, *fortified matrix or spiked matrix*), adicionando el analito a la matriz sintética de la muestra .

Los resultados usualmente se expresan como % de recuperación, es decir se compara la cantidad encontrada con la cantidad adicionada y se expresa como porcentaje.

Imitar la matriz de una muestra solo es posible en el caso de matrices sencillas, con pocos componentes bien conocidos.

Cuando no se dispone de muestra sin el analito y no es posible reproducir artificialmente la matriz de la muestra, se trabaja con el procedimiento de adición estándar, es decir se adiciona el analito a la muestra y se calcula la recuperación.

En este caso, para calcular la cantidad recuperada, se determina el analito inicialmente presente en la muestra y se resta su contribución al valor obtenido con la muestra adicionada. Esto supone el inconveniente de realizar dos determinaciones sobre la muestra (sin el agregado y con el agregado del analito), con el consiguiente incremento en la incertidumbre en el resultado. Por otra parte, es importante señalar que con el experimento de adición estándar se puede detectar un error sistemático proporcional, pero no se puede detectar un error sistemático constante. Esto se debe a que si existiera un error sistemático constante, el mismo afectaría a los dos valores medios, el de la muestra sin analito agregado y el de la muestra con analito agregado y al realizar la resta quedaría eliminado.

Estos procedimientos de preparar muestras o matrices fortificadas o el de adición estándar, constituyen una de las referencias más utilizadas, ya que muchas veces no se dispone de materiales de referencia, pero tienen el inconveniente de que el analito adicionado puede tener un comportamiento diferente al que se encuentra en la muestra.

Es decir que en un procedimiento de este tipo, existe la duda de si el comportamiento en el proceso analítico del analito añadido, es realmente representativo del analito nativo. Por ejemplo, en aquellas metodologías en que se lleva a cabo la extracción de los analitos, es muy probable que las

interacciones entre el analito añadido y la matriz, sean más débiles que las establecidas por el analito nativo y esto evidentemente conduce a una estimación por exceso de la recuperación.

Procedimiento B: utilización de un método analítico independiente para comparar resultados (método definitivo, método de referencia, método oficial, método estándar).

En este caso se trabaja comparando los resultados obtenidos al analizar muestras representativas con el método que se quiere validar y un método independiente, es decir basado en una técnica distinta. Idealmente el método de comparación debería ser un método definitivo, cuyo error sistemático es nulo por definición. Pero estos métodos sólo pueden implementarse en laboratorios de muy alta complejidad, debido a sus requisitos instrumentales y operacionales. Por esta razón es mucho más frecuente utilizar para la comparación, métodos de referencia, métodos oficiales o métodos estándar.

Las principales ventajas de este procedimiento están en que la evaluación se hace con las muestras que serán analizadas en el laboratorio y en que permite estimar error sistemático constante y proporcional. El principal inconveniente es el consumo de tiempo y la necesidad de contar con equipamiento, reactivos y personal capacitado, para la aplicación del método comparativo.

1.6.2.2.1. EVALUACIÓN EN UN INTERVALO REDUCIDO DE CONCENTRACIONES

Cuando el método se aplica en un intervalo reducido de concentraciones, se puede trabajar en un solo nivel, lo que implica asumir que el sesgo es el mismo para todo el intervalo de concentraciones de las muestras de rutina. En este caso se pueden seguir los procedimientos antes mencionados tal como se detallan a continuación.

Procedimiento A: utilización de un material de referencia

Cuando se dispone de un material apropiado, el mismo se analiza repetidamente en condiciones de repetibilidad, utilizando el procedimiento que se usará posteriormente para las muestras desconocidas.

Se obtiene de esta forma un conjunto de n medidas, a partir de las cuales se calcula el valor medio (\bar{x}) y la desviación estándar (s). Este valor medio, en ausencia de errores sistemáticos, debería coincidir con el valor asignado por la organización que proporciona el material de referencia.

Aún en ausencia de errores sistemáticos, existirá una diferencia entre estos valores, debido a la presencia de errores aleatorios y el problema se reduce a conocer si esta discrepancia se debe a la existencia de un sesgo significativo o sólo a errores aleatorios.

El análisis estadístico se puede realizar de diferentes maneras:

* Comparación de medias: en este caso habría que calcular un estadístico t (*t de Student*), pero no se dispone de todos los datos necesarios ya que se desconoce el número de determinaciones replicadas que se realizaron para obtener el valor certificado. Por esta razón no es posible comparar la repetibilidad de las medidas realizadas en el laboratorio con la desviación estándar asociada al valor asignado en el certificado (s_r), mediante una prueba F .

Para solucionar este inconveniente, algunas organizaciones elaboradoras de este tipo de materiales, recomiendan verificar que la repetibilidad de las medidas realizadas en el laboratorio es adecuada, antes de comparar el valor medio obtenido con el valor del certificado. Para esto se verifica que la desviación estándar de la media ($s_{\bar{x}}$) este por debajo de la desviación estándar correspondiente a la certificación del material de referencia (s_r).

Luego, si la repetibilidad es aceptable, recomiendan verificar que la media obtenida cae dentro de los límites de confianza: *valor certificado* $\pm 2 s_r$.

* Comparación de una media con un valor convencionalmente verdadero o valor de referencia aceptado.

En este caso se compara el valor medio obtenido (\bar{x}) con el valor certificado, considerado como valor convencionalmente verdadero (μ). Para esto se calcula un estadístico t (*t de Student*), si el número de repeticiones (n) realizadas sobre el material de referencia es menor que 30 y un estadístico z si es mayor que 30.

$$t_c = |\bar{x} - \mu| \sqrt{n} / s \quad (1.19)$$

$$z_c = |\bar{x} - \mu| \sqrt{n} / \sigma \quad (1.20)$$

El valor absoluto del estadístico calculado, se compara con un valor tabulado de dos colas, para un nivel de significancia α elegido y $n-1$ grados de libertad.

Si el estadístico calculado es menor que el tabulado, se acepta que no existen diferencias estadísticamente significativas entre el valor certificado y el valor medio obtenido al analizar el material de referencia con el método analítico en estudio.

Es importante tener en cuenta las siguientes consideraciones:

- Este tratamiento supone una distribución normal de los datos obtenidos. Si esto no se cumple, se debe comprobar la existencia de valores discrepantes, aberrantes (en inglés, *outliers*), es decir de valores que de alguna manera son inconsistentes con el resto de las observaciones. Si después de eliminar los valores discrepantes, el problema persiste, se debe aplicar alguna técnica de transformación de datos o utilizar ensayos no paramétricos.
- Cuanto mayor sea la desviación estándar de los resultados obtenidos sobre el material de referencia, existen más posibilidades de que el estadístico calculado sea menor que el tabulado y que por lo tanto no se detecten errores sistemáticos. Por esta razón se debe verificar en primer lugar, que el procedimiento empleado este bajo control estadístico en el laboratorio, es decir que tenga una imprecisión aceptable.
- Es conveniente determinar el número de veces (n) que debe analizarse el material de referencia y para esto se debe conocer la precisión del método en estudio y la incertidumbre del valor certificado y se debe fijar el sesgo mínimo que se desea detectar y las probabilidades de errores α y β .

Procedimiento A: preparación de muestras fortificadas, matrices fortificadas o experimento de adición estándar.

Como se dijo anteriormente, en esta experiencia se adiciona una cantidad exactamente conocida del analito ($C_{adicionada}$) a la muestra sin el analito o a la matriz sintética de la muestra. Se realizan replicados con el método en estudio y se obtiene un valor medio (\bar{x}), que es la cantidad de analito recuperada ($C_{recuperada}$).

Para el caso del procedimiento de adición estándar, la muestra que contiene al analito se divide en dos alícuotas iguales. A una de estas alícuotas, se le adiciona el analito (generalmente una solución concentrada en un solvente adecuado) en cantidad exactamente conocida y a la otra alícuota se le adiciona el mismo volumen del solvente utilizado para adicionar el analito (de manera de provocar en ambas alícuotas el mismo efecto sobre la matriz), siendo esta última la muestra de línea de base. Se realizan replicados con el método en estudio sobre ambas alícuotas y se obtiene un valor medio para cada una de ellas. La cantidad recuperada se calcula como la diferencia entre el valor medio obtenido para la muestra adicionada con el analito ($C_{muestra+analito}$) y el valor medio obtenido para la muestra de línea de base (C_{mlb}) según la ecuación (1.21):

$$C_{recuperada} = C_{muestra+analito} - C_{mlb} \quad (1.21)$$

Una vez determinada la cantidad recuperada, el análisis se puede realizar de dos maneras distintas:

1- Calculando el % de Recuperación como:

$$\%R = \frac{C_{recuperada}}{C_{adicionada}} \times 100 \quad (1.22)$$

El valor del % de Recuperación se puede comparar con criterios de aceptabilidad, tal como los establecidos en el análisis farmacéutico del $100 \pm 2 \%$ para componentes principales y $100 \pm 10 \%$ para impurezas.

2- Comparando la cantidad recuperada ($C_{recuperada}$), que sería el valor medio (\bar{x}), con la cantidad conocida agregada ($C_{adicionada}$), que sería el valor convencionalmente verdadero (μ), mediante una prueba t de dos colas y para un nivel de significancia α .

De acuerdo al resultado de esta prueba, se pueden sacar conclusiones acerca del sesgo del método en estudio.

Procedimiento B: utilización de un método comparativo

Si se dispone de un método analítico que reúna las condiciones para ser usado como un método comparativo, se puede verificar que el método en estudio no tiene un sesgo significativo, realizando análisis replicados de una muestra con ambos métodos y obteniendo dos conjuntos de resultados independientes entre si.

Preferentemente, el número de veces (n) que debe analizarse la muestra por ambos métodos, se debe determinar mediante expresiones matemáticas, fijando previamente el sesgo mínimo que se desea detectar y la probabilidad de error β .

Una vez que se ha analizado la muestra con ambos métodos (Método1: comparativo, Método 2: en estudio), se obtienen dos conjuntos de datos con sus valores medios respectivos (\bar{x}_1 y \bar{x}_2) y sus desviaciones estándar (s_1 y s_2).

Antes de realizar la comparación de las medias, se debe aplicar una prueba F (F de Fisher) para determinar si existen diferencias significativas entre las varianzas asociadas a ambos métodos, para lo cual se calcula el estadístico F :

$$F_c = \frac{s_1^2}{s_2^2} \quad (1.23)$$

Este estadístico calculado se compara con un valor tabulado de una cola, para un nivel de significancia α y para (n_1-1) y (n_2-1) grados de libertad.

Cuando el F_c es menor que el F tabulado, se concluye que no existe una diferencia significativa entre las varianzas y se calcula una varianza conjunta (en inglés, *pooled variance*):

$$s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{(n_1 + n_2 - 2)} \quad (1.24)$$

Luego se comparan las medias de ambos métodos, mediante una prueba t , calculando el estadístico según:

$$t_c = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s^2(1/n_1 + 1/n_2)}} \quad (1.25)$$

El valor absoluto del estadístico calculado se compara con un valor tabulado de dos colas, para un nivel de significancia α y $(n_1 + n_2 - 2)$ grados de libertad.

Si el t_c es menor que el tabulado, se concluye que no hay diferencias estadísticamente significativas entre las medias de ambos métodos.

Si la muestra es analizada un número de veces igual o mayor a 30, se trabaja con el estadístico z .

En el caso en que exista una diferencia significativa entre las varianzas, no se puede calcular una varianza conjunta y la comparación entre las medias de los dos conjuntos de resultados se hace mediante la prueba de Cochran, calculando el estadístico t según:

$$t_c = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{(s_1^2/n_1 + s_2^2/n_2)}} \quad (1.26)$$

Este estadístico calculado se compara con un valor tabulado de dos colas, un nivel de significancia α y los grados de libertad resultantes de redondear al número entero más próximo el valor:

$$v = \frac{(s_1^2/n_1 + s_2^2/n_2)^2}{\frac{(s_1^2/n_1)^2}{n_1 - 1} + \frac{(s_2^2/n_2)^2}{n_2 - 1}} \quad (1.27)$$

1.6.2.2.2. EVALUACIÓN EN UN INTERVALO AMPLIO DE CONCENTRACIONES

Cuando no es posible asumir que el error sistemático es el mismo en todo el intervalo de aplicación del método, se debe trabajar en distintos niveles de concentración, como mínimo a una concentración baja, una intermedia y una alta. En este caso se pueden seguir los procedimientos antes mencionados tal como se detallan a continuación.

Procedimiento A: utilización de un valor de referencia

Cuando es posible disponer de materiales de referencia con diferentes concentraciones del analito, se pueden realizar determinaciones réplicas, obteniendo de esta forma un conjunto de n medidas a partir de las cuales se calcula el valor medio (\bar{x}) y la desviación estándar (s), para cada nivel de concentración.

Se debe trabajar por lo menos con tres materiales de referencia con concentraciones diferentes, una alta, una baja y una intermedia, dentro del intervalo de aplicación del método a las muestras problema. En este caso, el análisis estadístico se puede realizar aplicando una prueba t para comparación de una media (\bar{x}) con un valor convencionalmente verdadero o valor de referencia aceptado (valor certificado), para cada material de referencia utilizado.

Si es posible conseguir un número mayor de materiales de referencia, los resultados obtenidos se pueden tratar mediante una técnica de regresión, graficando los valores medios encontrados en el eje y , frente a los valores certificados en el eje x .

Si no existiera error sistemático y no se cometiera ningún error de medición, se tendría la situación ideal donde:

$$y = x$$

que también puede escribirse como:

$$y = 0 + 1x \quad (1.28)$$

es decir una ordenada al origen (a) igual a cero y una pendiente (b) igual a 1.

Debido a la existencia de errores aleatorios, esta situación ideal no se da en la práctica y será necesario estimar los coeficientes a y b mediante regresión:

$$y = a + bx \tag{1.29}$$

Es decir que se trata de comparar parejas de valores que idealmente serían iguales, estudiando el desvío en un contexto estadístico y con un cierto nivel de confianza.

A los fines de seleccionar la técnica de regresión más adecuada, es conveniente verificar si existe homocedasticidad en la variable y , para lo cual se puede calcular un valor F como el cociente entre la varianza máxima y la mínima de las determinaciones réplicas realizadas para cada nivel de concentración:

$$F_c = \frac{\max[s(y_i)^2]}{\min[s(y_i)^2]} \tag{1.30}$$

Este valor se compara con un valor de tablas para $q-1$ y $q-1$ grados de libertad (donde q es el número de niveles de concentración) y un nivel de significancia α (usualmente 0.05).

Si el valor de F_c es menor que el valor tabulado, se considera que hay homocedasticidad y se puede usar regresión lineal ordinaria (OLS, del inglés *ordinary least-squares*).

Si por el contrario, el valor de F_c es mayor que el tabulado, se recomienda estimar los parámetros de la regresión mediante cuadrados mínimos ponderados (WLS, del inglés *weighted least-squares*). En este caso, se calculan la pendiente (b) y la ordenada al origen (a) de la recta ajustada a la ecuación (1.29), minimizando la siguiente suma ponderada de cuadrados (SC):

$$SC = \sum_{i=1}^q w_i (y_i - y_{i\text{calc}})^2 \tag{1.31}$$

Una vez que se dispone de los valores ajustados de a y b con sus respectivas desviaciones estándar, se analiza si las concentraciones encontradas para los materiales de referencia, difieren estadísticamente de los valores certificados. Para esto, se puede construir un intervalo de confianza para la ordenada al origen ($a \pm t_{(n-2),\alpha} \cdot S_a$) y ver si el mismo contiene al valor ideal cero y un intervalo de confianza para la pendiente ($b \pm t_{(n-2),\alpha} \cdot S_b$) y ver si el mismo contiene el valor ideal 1. Si esto es así, se

concluye que las diferencias son debidas únicamente a errores aleatorios, es decir no hay evidencia de error sistemático constante ni proporcional. Por otra parte, si el intervalo de confianza para la pendiente no contiene al uno, se concluye que hay evidencia de error sistemático proporcional y si el intervalo de confianza para la ordenada al origen no contiene al cero, se concluye que hay evidencia de error sistemático constante.

También puede darse el caso de que el intervalo de confianza para la pendiente no incluya al 1 y el intervalo de confianza para la ordenada al origen no incluya al cero, debido a la existencia de error sistemático constante y proporcional.

Actualmente, se considera que este procedimiento es incorrecto ya que no tiene en cuenta que a y b no son variables estadísticamente independientes y que siempre existe un cierto grado de correlación entre ellas. Es decir que debido a la *covarianza* entre la pendiente y la ordenada al origen, la utilización de intervalos individuales, puede llevar a conclusiones incorrectas, cuando se quiere verificar la ausencia de error sistemático constante y proporcional.

El procedimiento correcto es considerar un intervalo de confianza conjunto para la pendiente y la ordenada al origen (*EJCR*, del inglés *elliptical joint confidence region*), que toma en cuenta la correlación entre las estimaciones de ambos parámetros.

Este intervalo es una región en el plano de las dos variables que tiene forma elíptica y la prueba estadística correcta consiste en investigar si el punto $(1,0)$ esta contenido en esta región.

Las regiones elípticas de confianza conjunta están descritas por ecuaciones matemáticas y se dibujan en gráficos bidimensionales, siendo importante señalar que el tamaño de la elipse está controlado entre otros parámetros, por la desviación estándar de la regresión y da una idea de la precisión del método analítico en estudio.

Si existe homocedasticidad y se ha podido aplicar regresión lineal ordinaria, la región elíptica esta descrita por la siguiente ecuación (González, 1999):

$$q(\beta - b)^2 + 2(\alpha - a)(\beta - b) \sum_{i=1}^q x_i + (\alpha - a)^2 \sum_{i=1}^q x_i^2 = 2 S_{y/x}^2 F_{2,q-2} \quad (1.32)$$

donde α y β son las variables que corresponden a las dos dimensiones del plano en que se representa la región elíptica, $s_{y/x}$ es el desvío estándar de la regresión y $F_{2,q-2}$ es el valor del parámetro estadístico F con 2 y $q-2$ grados de libertad para un dado nivel de confianza (usualmente 95%).

Para el caso heterocedástico, en que es necesario aplicar regresión por cuadrados mínimos ponderados, la ecuación que describe la elipse de confianza conjunta es:

$$(\beta - b)^2 \sum_{i=1}^q w_i + 2(\alpha - a)(\beta - b) \sum_{i=1}^q w_i x_i + (\alpha - a)^2 \sum_{i=1}^q w_i x_i^2 = 2 s_{y/x}^2 F_{2,q-2} \quad (1.33)$$

donde w_i es el “peso” o “ponderación” aplicado a cada punto de la regresión, definido como inversamente proporcional a la variancia de la variable en el punto i :

$$w_i = \frac{1}{s(y_i)^2} \quad (1.34)$$

También se pueden preparar muestras fortificadas o matrices fortificadas, agregando cantidades exactamente conocidas del analito ($C_{adicionada}$), de manera de trabajar en varios niveles de concentración. Luego se realizan replicados con el método en estudio y se obtiene un valor medio (\bar{x}), que es la cantidad de analito recuperada ($C_{recuperada}$) y una desviación estándar (s), para cada nivel.

En este caso también es necesario trabajar por lo menos en tres niveles de concentración diferentes, uno cercano al Límite de Cuantificación (LOQ), uno cercano al centro y otro cercano al límite superior del rango lineal.

Para el caso del análisis farmacéutico de materias primas o de componentes mayoritarios de una preparación, se trabaja en tres niveles, que corresponden al 100 % del valor declarado y en los límites extremos que pueden ocurrir en la práctica del 80 % y del 120 %.

El análisis estadístico se puede realizar aplicando una prueba t para comparación de una media (\bar{x}) con un valor convencionalmente verdadero o valor de referencia aceptado ($C_{adicionada}$), para cada nivel de concentración.

Cuando es posible preparar un número mayor de muestras o matrices fortificadas, los resultados obtenidos se pueden tratar mediante una técnica de regresión, graficando los valores medios

encontrados para las cantidades recuperadas en el eje y, frente a los valores de las cantidades adicionadas en el eje x.

El tratamiento estadístico de los datos para detectar la presencia de error sistemático constante y / o proporcional, se hace de la misma manera que para el caso en que los valores de referencia provienen de la utilización de materiales de referencia certificados.

En la Figura 1.9, se muestran las diferentes situaciones que se pueden dar:

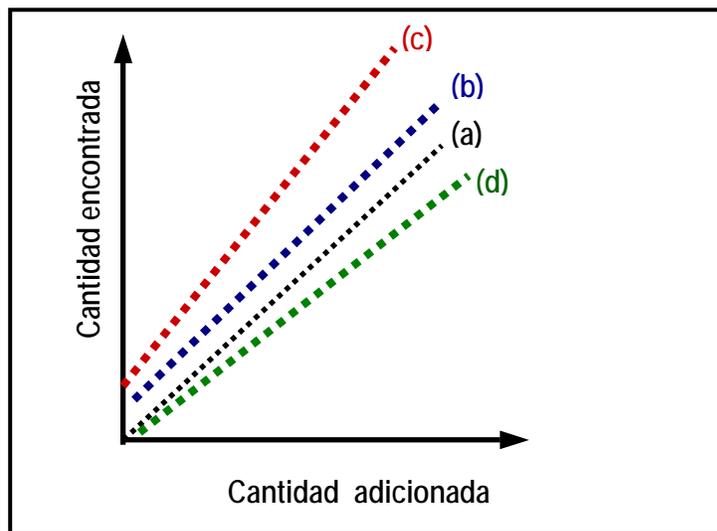


Figura 1.9: Relaciones posibles entre las cantidades encontradas y las adicionadas (a) Ausencia de error sistemático, (b) Error sistemático constante, (c) Error sistemático constante y proporcional, (d) Error sistemático proporcional

Cuando no se dispone de materiales de referencia y no es posible preparar muestras o matrices fortificadas, se puede realizar un experimento de adición estándar, adicionando diferentes cantidades del analito a alícuotas del material a ser analizado y midiendo las señales analíticas generadas para cada nivel. Se aplica una técnica de regresión y se grafican los valores de las cantidades exactamente conocidas de analito adicionadas en el eje x frente a las señales medidas en el eje y.

En este caso no es posible realizar una prueba con la ordenada al origen, ya que la misma corresponde a la señal del analito originalmente presente en la muestra. Por esta razón no es posible detectar error sistemático constante. Posteriormente se analiza la pendiente de la recta obtenida con la adición estándar, respecto de la pendiente de la recta de calibración obtenida con estándares acuosos. Si

ambas rectas tienen la misma pendiente, se concluye que no hay evidencia de error sistemático proporcional.

Tal como se muestra en la Figura 1.10, la recta (a) es la recta obtenida con la adición estándar. Si la recta de calibración con estándares acuosos es (b), se observa que las pendientes son iguales y se concluye que no hay evidencia de error, si la recta de calibración con estándares acuosos es (c), se observa que las pendientes son diferentes y se concluye que existe error sistemático proporcional.

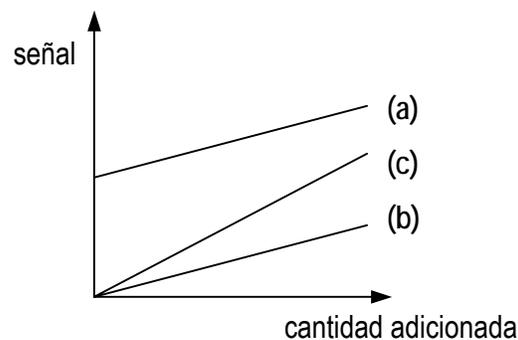


Figura 1.10: Experimento de adición estándar. (a) Curva de Adición Estándar, (b) y (c) curvas de calibrado con estándares acuosos

Procedimiento B: utilización de un método comparativo

Cuando las muestras a ser analizadas con el método en estudio tienen un solo tipo de matriz bien definido, la comparación de los resultados obtenidos con los de un método comparativo puede hacerse trabajando en tres niveles de concentración diferente. Por ejemplo, se podría trabajar en los límites superior e inferior del rango lineal y en un nivel intermedio, o como en el caso del análisis farmacéutico de componentes mayoritarios, en el 80, 100 y 120 % del valor declarado.

Se realizan análisis replicados en cada nivel por ambos métodos y el análisis estadístico se hace mediante una prueba t de comparación de medias.

Cuando las muestras a ser analizadas con el método en estudio presentan tipos diferentes de matrices, se deben analizar muchas muestras con concentraciones uniformemente distribuidas dentro de todo el rango de trabajo.

En este caso, el tratamiento estadístico de los datos se puede realizar de dos maneras diferentes:

a- Comparación de resultados mediante una prueba *t* para datos pareados.

La comparación de los resultados del método en estudio con los de un método comparativo, se puede hacer analizando un conjunto de muestras seleccionadas, por duplicado con ambos métodos.

Se calcula, para cada muestra y para cada método, la media (\bar{x}) y la diferencia absoluta ($|w|$) entre los dos resultados replicados. Los datos se pueden presentar en una tabla del tipo:

Muestra	Método 1		Método 2		Diferencia
	\bar{x}_1	$ w_1 $	\bar{x}_2	$ w_2 $	$d_i = \bar{x}_1 - \bar{x}_2$
1					
2					
i					
n					

Tabla 1.3: Formato para la presentación de los datos provenientes de la comparación de métodos

Se halla la diferencia d_i para cada par de valores y se calcula la media de las diferencias (\bar{d}) y la desviación estándar (s_d) del conjunto de diferencias. Luego se calcula un estadístico t según:

$$t_c = \frac{\bar{d}}{s_d / \sqrt{n}} \quad (1.35)$$

Para determinar si la media de las diferencias difiere significativamente de cero, se compara este estadístico con un valor tabulado de dos colas, con un nivel de significancia α elegido y $n-1$ grados de libertad.

La aplicación de la prueba t para datos pareados, supone que las diferencias (d_i) pertenecen a misma distribución. Si se observa que la magnitud de las diferencias depende de la concentración del analito, no puede asumirse que se cumple con este supuesto y no es correcto aplicar esta prueba. Esto se da, cuando la varianza de los métodos depende fuertemente de la concentración.

Para verificar esta situación, se pueden representar gráficamente las diferencias absolutas entre los replicados ($|w|$) vs. los valores medios (\bar{x}) y observar que no se muestren tendencias, en un gráfico del tipo:

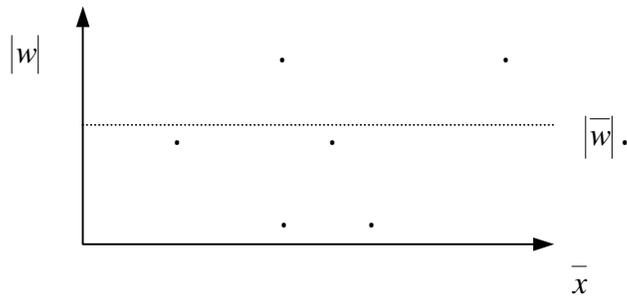


Figura 1.11: Representación visual de las diferencias absolutas entre replicados para uno de los métodos

También se puede construir un tipo de gráfico propuesto por Bland y Altman (Massart, 1997), donde se grafican las diferencias ($d_i = \bar{x}_1 - \bar{x}_2$) vs $\bar{x} = (\bar{x}_1 + \bar{x}_2)/2$, pudiendo observarse las siguientes situaciones:

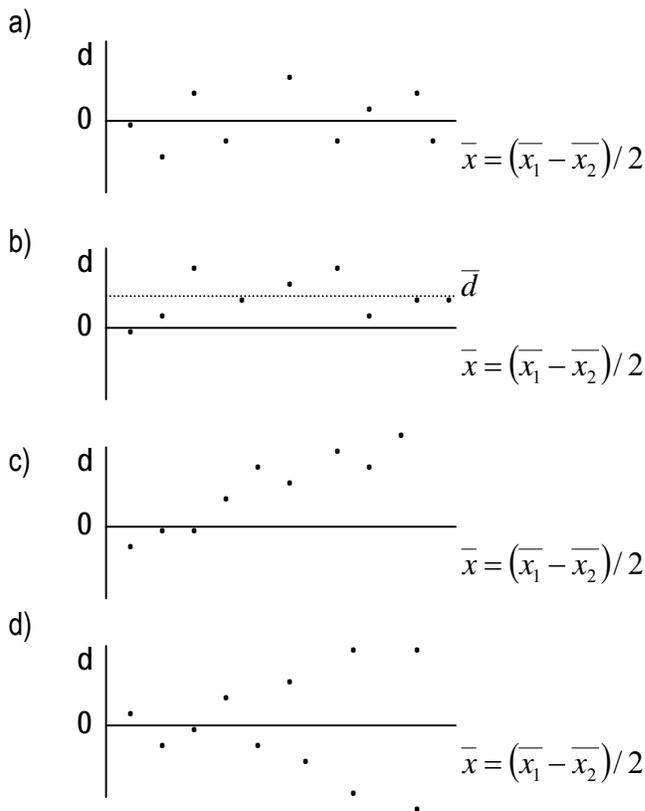


Figura 1.12: Gráficos de Bland y Altman para diferentes situaciones: a) no se observa ninguna tendencia, b) se observa la existencia de un error sistemático constante, c) se observa la existencia de error sistemático proporcional, d) se observa heterocedasticidad

b- Comparación de resultados aplicando regresión lineal

Los resultados obtenidos al procesar las muestras con ambos métodos se pueden tratar estadísticamente aplicando un modelo de regresión lineal para verificar si son comparables.

Los resultados determinados por el método en estudio se consideran la variable y , mientras que los resultados provistos por el método comparativo se consideran la variable x . La regresión lineal de y vs. x se debe hacer con un método que tenga en cuenta los errores en ambos ejes, ya que ambas variables tienen asociada una incertidumbre finita.

Considerando lo anteriormente dicho, no es correcto aplicar, regresión lineal ordinaria (*OLS*) o regresión por cuadrados mínimos ponderados (*WLS*), ya que en ambos ajustes la suposición básica es que no hay error en la variable x .

El método adecuado para tratar este tipo de datos es la regresión por cuadrados mínimos bivariados (*BLS*, del inglés *bivariate least-squares*) (Riu, 1996). En este caso la pendiente y la ordenada al origen de la recta ajustada, se obtienen minimizando una función idéntica a la utilizada para la regresión ponderada, excepto que ahora los pesos son una función de las variancias en ambas variables. Es decir que los pesos de la regresión doblemente ponderada, se eligen como inversamente proporcionales a una combinación de las varianzas en x e y , minimizando la siguiente suma de cuadrados:

$$SC = \sum_{i=1}^q (y_i - y_{i\text{calc}})^2 / (s(y_i)^2 + b^2 s(x_i)^2) \quad (1.36)$$

Cuando la varianza en la variable x es significativamente menor que en la variable y , no es imprescindible aplicar *BLS* y la comparación puede hacerse empleando el método *WLS*, si se tiene la precaución de asignar a la variable x los resultados del método más preciso y a la variable y los resultados del método menos preciso.

Si se puede hacer esta aproximación, se calcula la pendiente y la ordenada al origen mediante *WLS*. Se analiza la región elíptica de confianza conjunta, y si el punto ideal (1,0) esta contenido dentro de la elipse, se concluye que los métodos son comparables estadísticamente en cuanto a la predicción de la concentración del analito en las muestras analizadas.

La interpretación de los resultados obtenidos es similar a la presentada cuando se comparaban resultados del método con valores de referencia. Es decir si hay una desviación de la pendiente respecto de 1, se concluye que hay una discrepancia proporcional entre los dos métodos y si hay una

desviación de la ordenada al origen respecto de cero, se concluye que hay una discrepancia absoluta ó error sistemático constante entre los métodos comparados.

En un experimento de este tipo, se debe tener en cuenta:

- Que se debe trabajar sobre muchas muestras seleccionadas adecuadamente, de manera de que todos los niveles de concentración estén igualmente representados, trabajando dentro del rango lineal de ambos métodos.
- Que se debe considerar el tema de la estabilidad del analito en las muestras y el tiempo requerido para realizar el análisis por ambos métodos.
- Que se debe realizar un análisis visual de los resultados presentados en un gráfico de tipo x - y , donde x se corresponde con los resultados del método comparativo e y con los resultados del método en estudio, donde cada punto representa una muestra analizada por ambos métodos.

1.6.2.3. LINEALIDAD

En las recomendaciones oficiales para el análisis farmacéutico, la *linealidad* de un método analítico se define como la capacidad para generar resultados que son directamente proporcionales a la concentración del analito, dentro de un rango dado. Se especifica además, que en algunos casos puede ser necesario aplicar una transformación matemática para obtener linealidad y que el tratamiento estadístico que se aplica es el cálculo de una recta de regresión por un método de cuadrados mínimos, de los resultados del método vs. la concentración del analito (USP, 2000; ICH, 1996).

En este contexto, se utilizan los métodos ya vistos para determinación de error sistemático.

Muchos procedimientos de medición hacen uso de gráficos de calibración para estimar la concentración del analito en las muestras desconocidas y son diseñados de manera de tener una relación lineal entre la respuesta y la concentración. Por otra parte, la aplicación de algunas herramientas quimiométricas para el tratamiento de los datos (como por ejemplo, ciertos modelos de calibración multivarida), supone la existencia de linealidad y en la práctica es frecuente observar desviaciones.

Por esta razón, una etapa fundamental en el procedimiento de validación de un método analítico, es verificar si los datos presentan un comportamiento lineal, es decir si existe una relación lineal entre la concentración del analito y la respuesta del método.

Para esto, el procedimiento más simple consiste en analizar la linealidad de una recta de calibrado, construida a partir de patrones de concentración conocida del analito.

El primer paso en la construcción de una recta de calibración, es determinar el intervalo de concentraciones dentro del cual interesa verificar el comportamiento lineal. En algunos casos, este intervalo va desde cero hasta los valores más altos que pueden encontrarse en las muestras a ser analizadas. En el análisis farmacéutico de componentes mayoritarios, se trabaja en intervalos que van del 80 al 120 % del valor declarado y para el análisis de impurezas, el intervalo va del 50 al 120 % de las especificaciones.

Una vez que se ha establecido el intervalo de concentraciones en el que se va a trabajar, se preparan los patrones de calibración. Como mínimo se deberían preparar cinco patrones con concentraciones igualmente espaciadas y se debe procesar además un blanco. Posteriormente, se miden las respuestas analíticas por triplicado. También se mide por triplicado la respuesta analítica del blanco (valor cero de concentración).

Luego se hace un análisis de los datos de calibración obtenidos, adoptando una técnica de regresión por cuadrados mínimos ordinarios (*OLS*), que asume que la variable independiente (x , *concentración nominal de los patrones*) no esta sujeta a error y que la variable dependiente (y , *respuesta analítica de los patrones*) es homocedástica. Mediante esta técnica de regresión, se pueden estimar los parámetros del modelo lineal adoptado: ordenada al origen (a) y pendiente (b), según las siguientes ecuaciones:

$$b = \frac{\sum_{i=1}^m (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^m (x_i - \bar{x})^2} \quad (1.37)$$

$$a = \bar{y} - b\bar{x} \quad (1.38)$$

Donde x_i es la concentración de cada uno de los m patrones, \bar{x} es el promedio de esas concentraciones, y_i es la respuesta en cada punto e \bar{y} es el promedio de esas respuestas.

Dado que los datos instrumentales tienen asociado un error que depende del ruido instrumental, es conveniente estimar la incertidumbre de estos parámetros, mediante las siguientes ecuaciones:

$$s_b = \frac{s_{y/x}}{\sqrt{\sum_{i=1}^m (x_i - \bar{x})^2}} \quad (1.39)$$

$$s_a = s_{y/x} \sqrt{\frac{1}{m} + \frac{\bar{x}^2}{\sum_{i=1}^m (x_i - \bar{x})^2}} \quad (1.40)$$

Siendo el parámetro $s_{y/x}$, el desvío estándar de los residuos de la regresión y esta dado por:

$$s_{y/x} = \sqrt{\frac{\sum_{i=1}^m (y_i - y_{icalc})^2}{m - 2}} \quad (1.41)$$

Donde y_{icalc} representa la respuesta estimada en cada punto a partir de la recta de regresión encontrada:

$$y_{icalc} = a + bx_i \quad (1.42)$$

El paso siguiente es la validación de este modelo lineal, es decir la verificación de que es el correcto para ajustar la relación de las dos variables y que se cumplen los supuestos. El análisis de la adecuación de los datos al modelo lineal, se puede hacer de diferentes maneras:

*** Análisis de los residuos**

La distribución de los residuos, es decir, el modo en que los valores de $(y_i - y_{icalc})$ varían con la respuesta, puede proveer información valiosa acerca del cumplimiento de los supuestos y de la falta de ajuste al modelo elegido. Para esto, se construye un gráfico de los residuos en función de x_i , donde pueden observarse diferentes situaciones:

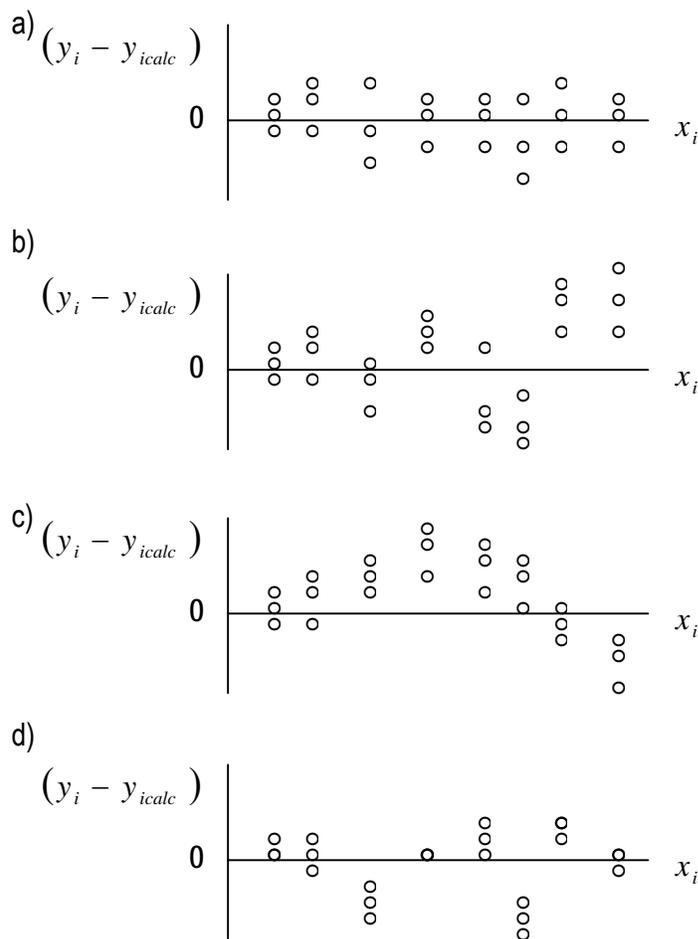


Figura 1.13: Diferentes situaciones para Gráficos de residuos: a) Comportamiento lineal, b) No hay homocedasticidad, c) Comportamiento no lineal, d) Comportamiento lineal con alta incertidumbre en la concentración de los patrones

La Figura 1.13 ilustra casos representativos del comportamiento de los residuos en distintas situaciones. En el caso a) están aleatoriamente distribuidos dentro de una banda horizontal, con un número de residuos positivos aproximadamente igual a los negativos, siendo la variabilidad interna de las réplicas a cada nivel de concentración comparable a la variabilidad global. En este caso se concluye que el comportamiento es lineal. En b) no se cumple la condición de homocedasticidad, la magnitud de los residuos se incrementa con la concentración. En c) se presenta un caso típico de desviación de la linealidad, donde se aprecia visualmente que los residuos poseen un comportamiento parabólico. En este caso se concluye que hay falta de ajuste al modelo lineal. En d) los residuos muestran una variabilidad global significativamente mayor que la que presentan las réplicas a cada nivel. En esta

situación el sistema se comporta linealmente, pero existe mayor incertidumbre en las concentraciones nominales de los patrones que en la señal instrumental.

** Prueba F*

Una prueba estadística que puede utilizarse para determinar si los datos se ajustan a una ley lineal, es la prueba F (Danzer, 1998). En esta prueba se relaciona la varianza de la regresión, medida por $(s_{y/x})^2$ y la del ruido instrumental, medida por $(s_y)^2$. Si la primera es significativamente mayor que la segunda, se supone que hay causas de desvío de la linealidad que son estadísticamente superiores al ruido en la respuesta.

El ruido instrumental se puede estimar mediante la siguiente ecuación:

$$s_y = \sqrt{\frac{\sum_{i=1}^p \sum_{j=1}^r (y_{ij} - \bar{y}_i)^2}{m - p}} \quad (1.43)$$

donde p es el número de niveles de concentración estudiados en la recta, r es el número de réplicas de cada punto, m el número total de datos disponibles, y_{ij} es el valor de la respuesta correspondiente a cada nivel y réplica e \bar{y}_i es el promedio de las respuestas de las réplicas para cada nivel de concentración.

Se calcula un valor experimental de F según:

$$F_c = \frac{(s_{y/x})^2}{(s_y)^2} \quad (1.44)$$

Luego se compara este valor calculado con un valor F tabulado de una cola, para un nivel de significancia α y $m-2$ y $m-p$ grados de libertad. Si el valor calculado es menor que el tabulado, se acepta que los datos se comportan linealmente.

** Análisis de la varianza residual*

Esta prueba F relaciona una varianza de falta de ajuste MS_{lof} (*del inglés, lof: lack of fit*) con la varianza del error puro Mse (Analytical Methods Committee, 1994). Para aplicar esta prueba, se plantean las siguientes sumas de cuadrados (SS) con sus respectivos grados de libertad (Df):

1) Residual:	$SS_r = \sum_{ij} (y_{ij} - y_{i,calc.})^2$	$MS_r = SS_r / Df$	$Df = p \cdot r - 2$
2) Error puro:	$SS_e = \sum_j \sum_i (y_{ij} - \bar{y}_i)^2$	$MSe = SS_e / Df$	$Df = p (r-1)$
3) Falta de ajuste:	$SS_{lof} = SS_r - SS_e$	$MS_{lof} = SS_{lof} / Df$	$Df = p - 2$

donde p es el número de niveles de concentración y r es el número de réplicas de cada concentración. Luego se calcula un $F = MS_{lof} / MSe$ y este debe ser menor a $F_{(p-2), p(r-1)}$ para que exista linealidad con una significancia estadística α . Esta prueba se realiza con valores F críticos de una cola debido a que la hipótesis nula es que las variables no están linealmente relacionadas.

1.6.2.4. LÍMITE DE DETECCIÓN

El *límite de detección (LOD)* es la mínima concentración detectable de manera confiable por el método, es decir es la mínima concentración de analito que proporciona una señal significativamente diferente de la media de las señales del blanco.

El límite de detección se calcula mediante una prueba de hipótesis estadística, donde se fija una concentración llamada nivel crítico (L_c) a partir de donde se toma la decisión de aceptar o rechazar la presencia del analito, con un error suficientemente pequeño. Para concentraciones superiores al L_c , existe una probabilidad α de cometer un error de tipo I o falso positivo, es decir de aceptar erróneamente la hipótesis alternativa, admitiendo que el analito está presente cuando en realidad está ausente. Si se toma α igual a 0.05, una concentración superior al L_c tendrá un 5 % de probabilidad de constituir un falso positivo y la distancia desde L_c y el cero de la escala vale $t_{\alpha, v} s_0$. Donde s_0 es la desviación estándar de la concentración predicha para una muestra blanco (Clayton, 1987).

También existe una probabilidad β de cometer un error de tipo II o falso negativo, es decir de aceptar erróneamente la hipótesis nula, admitiendo que el analito está ausente cuando en realidad está presente. Si se toma β igual a 0.05, la probabilidad de tener un falso negativo será del 5 % y la distancia entre el L_c y la concentración por encima correspondiente a este valor de β , se puede considerar que vale $t_{\beta, \nu} s_0$.

De este modo, se puede calcular el LOD según (Currie, 1999; Currie, 1995):

$$LOD = 2 \times t_{0.05, m-2} \times s_0 \quad (1.45)$$

En la práctica, dado que m es un número relativamente grande, el valor de $(2 \times t_{0.05, m-2})$ tiende a 3.3, por lo que una ecuación aproximada que puede usarse para el cálculo de LOD es:

$$LOD = 3.3 \times s_0 \quad (1.46)$$

Para estimar s_0 se usa la ecuación que permite calcular la desviación estándar de la concentración predicha para una muestra incógnita, a partir de la recta de calibrado:

$$s(x_{inc}) = \frac{s_{y/x}}{b} \sqrt{\frac{1}{n} + \frac{1}{m} + \frac{(x_{inc} - \bar{x})^2}{\sum_{i=1}^m (x_i - \bar{x})^2}} \quad (1.47)$$

Suponiendo que se analiza por triplicado ($n=3$) una muestra en la que el analito no está presente ($x_{inc} = 0$), la ecuación resulta:

$$s_0 = \frac{s_{y/x}}{b} \sqrt{\frac{1}{3} + \frac{1}{m} + \frac{\bar{x}^2}{\sum_{i=1}^m (x_i - \bar{x})^2}} \quad (1.48)$$

1.6.2.5. LÍMITE DE CUANTIFICACIÓN

El *límite de cuantificación (LOQ)* es la mínima concentración cuantificable en forma confiable. En otras palabras, el límite de cuantificación es la concentración que puede determinarse con una desviación estándar relativa (*RSD*) máxima especificada y exactitud adecuada.

El nivel que se toma convencionalmente como la máxima desviación estándar relativa es del 10 %, por lo tanto el límite de cuantificación es la concentración correspondiente a 10 veces la desviación estándar del blanco (en unidades de concentración):

$$LOQ = 10 \times s_0 \quad (1.49)$$

El límite de cuantificación también se puede definir como la concentración para la cual se alcanza una relación señal-ruido (*S/R*) igual a 10, siendo la relación señal ruido el cociente entre la señal neta y su desviación estándar.

1.6.2.6. RANGO

El *rango* de un método analítico se define como el intervalo de concentraciones sobre el cual se obtiene exactitud, linealidad y precisión aceptables. En la práctica la determinación del rango se hace a partir de los resultados de los estudios de estos tres parámetros (Green, 1996).

El *rango dinámico* de un método analítico, es el intervalo de concentraciones en el que es posible calibrar la relación señal-concentración, es decir que es el intervalo entre los niveles más bajo y más alto del analito, en el cual la relación lineal u otra relación de calibración es correcta. En este intervalo, la sensibilidad (*SEN*) existe y tiene un valor determinado. Este es el rango de aplicabilidad del método, ya que en la zona de pérdida de la linealidad, podría aplicarse algún método de regresión de naturaleza no lineal. Se considera que va desde la menor concentración detectable (*LOD*) hasta la pérdida de la relación entre respuesta y concentración.

El *rango lineal* de un método analítico, es el intervalo de concentraciones en el que la sensibilidad (*SEN*), siendo suficientemente distinta de cero y de infinito, se mantiene aproximadamente constante. (Ramis Ramos, 2002). Se considera que va desde un límite inferior correspondiente a la menor

concentración que puede determinarse (*LOQ*), hasta un límite superior correspondiente a la pérdida de la linealidad .

El rango lineal de un método analítico es un parámetro de calidad, porque su extensión determina las muestras que pueden analizarse directamente, sin tener que recurrir a operaciones especiales de preconcentración o de dilución de los analitos (u otro tratamiento que no forma parte del procedimiento usual de medición), las que complican el procedimiento, incrementan la incertidumbre y suponen un riesgo adicional de error sistemático.

1.6.2.7. SENSIBILIDAD

En el análisis cuantitativo, la *sensibilidad* da una indicación de cuanto cambia la señal con la concentración (Massart, 1997). Según la *IUPAC* (Thompson, 2002), en metrología y en química analítica, la sensibilidad se define como la pendiente de la recta de calibrado, porque un método con una pendiente importante es mejor para discriminar entre pequeñas diferencias en la concentración del analito. Esta sensibilidad se denomina *sensibilidad de calibración* ($SEN = b$), sus unidades son: señal \times concentración⁻¹, e indica la variación de la respuesta producida por una unidad de variación de concentración del analito.

No tiene mucho sentido incluir la sensibilidad como parámetro para caracterizar un método analítico, cuando se la define de esta manera. No es suficiente conocer la pendiente de la recta de calibrado, para determinar cuando dos concentraciones podrán discriminarse, sino que es necesario conocer además la desviación estándar de la señal.

Por otra parte, la sensibilidad de calibración no es adecuada para comparar dos métodos analíticos, cuando estos están basados en respuestas de diferente naturaleza (por ejemplo: absorbancia y fluorescencia, absorbancia y medidas electroquímicas, etc). En este caso se prefiere usar la llamada *sensibilidad analítica* (γ), definida por la relación entre la sensibilidad y el ruido instrumental:

$$\gamma = SEN/s_y \tag{1.50}$$

La inversa de la sensibilidad analítica (γ^{-1}), se puede definir como la mínima diferencia de concentración estadísticamente discernible en cualquier punto de la recta de calibrado.

1.6.2.8. SELECTIVIDAD

No existe un acuerdo universal acerca del significado de los términos *selectividad*, *especificidad* e *interferencia*. La *ICH* habla de especificidad y la define como la habilidad para evaluar en forma inequívoca al analito en presencia de otros componentes que podría esperarse que estén presentes (*ICH*). La guía *EURACHEM* dice que la selectividad y la especificidad aseguran la confiabilidad de la medición del analito en presencia de interferencias y que la especificidad se considera generalmente como el 100 % de la selectividad (*EURACHEM*, 1998). Es decir que la selectividad extrema o perfecta, en el sentido de que únicamente el analito es el responsable de la señal medida, se denomina especificidad.

Algunos autores consideran que no es conveniente usar el término especificidad para la caracterización de los procedimientos de análisis cuantitativos, dado que este término se usa en el contexto del análisis cualitativo (*Massart*, 1997).

Por otra parte, una *sustancia interferente* o *interferencia* de un procedimiento analítico, es aquella que, a una concentración dada, causa un error sistemático en el resultado. Las interferencias relacionadas con la selectividad causan un error sistemático constante, en cambio las relacionadas con el llamado efecto matriz, causan un error sistemático proporcional.

En términos generales, también se puede decir que la selectividad es un parámetro que permite asegurar que la señal medida no es influenciada por otras sustancias presentes o, al menos, que la contribución de otras sustancias ha sido eliminada. Se puede definir como el medio de expresar cualitativamente la extensión en que otras sustancias interfieren en la determinación de un analito con un método determinado.

Para mejorar la selectividad de un método analítico, se incluyen operaciones de separación o enmascaramiento y las señales obtenidas se tratan mediante procedimientos electrónicos y matemáticos de filtrado.

Pero por muy sofisticado que sea el procedimiento de medida, una parte de la señal puede ser debida a interferencias, es decir, a causas distintas de la concentración o de la masa del analito presente en la muestra y cuando esto no es tenido en cuenta, los resultados están afectados por un error sistemático.

Las interferencias causan un error sistemático constante, dando lugar a una señal más alta, cuando la señal de las interferencia se añade a la señal del analito o más baja de la esperada cuando dan una señal que se resta de la señal del analito.

Un método es considerado selectivo, cuando es posible demostrar que las otras especies presentes no tienen respuesta por sí mismas. Para esto, habría que hacer una lista de los posibles interferentes y probar experimentalmente que dichas sustancias no tienen efecto sobre los resultados.

En un método cromatográfico, se demuestra que el pico del posible interferente se separa completamente del pico del analito. En el análisis farmacéutico, son posibles interferentes: los excipientes, los productos intermedios de síntesis, los productos de degradación y las impurezas propias del proceso de obtención o elaboración del producto.

Cuando es posible conseguir una matriz sin el analito, se hace un blanco a partir de la misma y si este es suficientemente bajo, se concluye que las sustancias en la muestra no contribuyen a la señal con la que se va a cuantificar el analito. Pero este procedimiento no es seguro, ya que siempre es posible que otra matriz pueda contener otras sustancias presentes, que pueden ser interferentes. Por esta razón, las recomendaciones para los métodos bioanalíticos plantean que se deben analizar por lo menos seis muestras de diferente origen.

Algunas veces, se trabaja haciendo determinaciones sobre un número de muestras, con y sin el posible interferente. Para cada muestra se comparan dos series de mediciones, mediante una *prueba t*.

Si con alguno de estos procedimientos, se detecta la presencia de especies interferentes en el método analítico estudiado, es posible adoptar distintas estrategias de trabajo a seguir. Se podría cambiar de método, recurriendo a otros basados en técnicas cromatográficas, las que son inherentemente más selectivas que otras técnicas analíticas, porque conllevan una separación previa a la medida de la señal analítica. Esto es así, especialmente si la detección presenta una selectividad elevada, como puede ser la espectroscopia de fluorescencia molecular o determinadas espectroscopias atómicas. Por otra parte, la utilización de detectores que realizan el barrido rápido de alguna propiedad, como los detectores de series de diodos o la espectrometría de masas, permite evaluar la pureza de los picos cromatográficos o, lo que es lo mismo, detectar la presencia de especies interferentes que coeluyen con el analito. Si mediante la selección de un método basado en una técnica cromatográfica, tampoco es posible alcanzar el nivel de selectividad buscado, será necesario recurrir a un método que utilice más de una técnica analítica (*métodos acoplados*).

Muchas veces, sucede que este tipo de métodos no son adecuados para resolver el problema analítico planteado y es necesario trabajar con métodos espectroscópicos, mejorando su selectividad.

En algunos casos es posible aplicar una técnica de enmascaramiento, que permite obviar la segregación física de las especies interferentes. El *enmascaramiento* consiste en el bloqueo químico de la interferencia en el propio medio donde se realizan las medidas. Esta aproximación es sencilla y cómoda, pero tiene un campo de aplicación restringido.

Para aislar la señal del analito de la señal de las interferencias, se puede preparar un *blanco*, es decir una solución que tenga exactamente la misma composición que la muestra, pero que no contenga al analito o bien que lo contenga en alguna forma inactiva. Un blanco se puede preparar como blanco interno o como blanco externo. En el blanco externo se imita la matriz, preparando artificialmente una muestra que contiene todos los componentes de la misma a sus respectivas concentraciones, salvo el analito. Imitar la matriz de una muestra, sólo es posible en el caso de matrices muy sencillas, con pocos componentes bien conocidos. Cuando se trabaja con muestras de matrices complejas (por ejemplo: muestras ambientales, o muestras biológicas), resulta imposible preparar un blanco que permita corregir satisfactoriamente las interferencias.

El blanco interno esta formado por la muestra, a la que se le han hecho las modificaciones necesarias para eliminar la señal del analito, sin que se alteren las señales de las interferencias. Estas modificaciones, pueden ser la destrucción o conversión del analito en una forma inactiva por oxidación, hidrólisis, etc, o la alteración del procedimiento de preparación de la muestra, para que no se forme la especie química activa que produce la señal instrumental. Los blancos internos son adecuados, sólo si es posible eliminar o destruir exclusivamente el analito en su totalidad, sin que la muestra experimente ninguna otra alteración que modifique la señal de los compuestos que interfieren.

Para solucionar el problema de la falta de selectividad de algunos métodos analíticos, se utilizan en la actualidad herramientas quimiométricas muy potentes, tales como la calibración multivariada, tema que se desarrolla en el presente trabajo de tesis.

1.6.2.9. ROBUSTEZ

Según la *ICH*, la robustez es la capacidad de un método analítico para que los resultados no se vean afectados por variaciones deliberadas en los parámetros del procedimiento, es decir es la resistencia al cambio de respuesta ante pequeñas variaciones en las condiciones experimentales de trabajo (por ejemplo: pH del medio, fuerza iónica, temperatura, concentración de algunos reactivos, tiempos de incubación, composición de la fase móvil en cromatografía de líquidos, caudal de la fase móvil, etc).

En las pruebas de robustez se introduce deliberadamente el tipo de variación en las variables, que puede esperarse razonablemente que ocurra durante el empleo del método. Se observan los resultados, estudiando estadísticamente la influencia de las variables a los fines de determinar las que tienen un efecto más importante sobre la respuesta (puntos débiles experimentales, factores críticos) y sus intervalos de valores tolerables (rangos de trabajo de las variables).

Un estudio de robustez se debe planificar mediante un diseño experimental, donde la función objetivo o respuesta a seguir es alguno de los parámetros de calidad antes mencionados (Vander Heyden, 2001).

Para cada variable se estudian dos valores (niveles extremos) situados alrededor del valor indicado en el procedimiento (nivel nominal). Estos valores extremos se fijan a una distancia del valor nominal, de manera que la diferencia sea la que razonablemente se puede producir en la práctica y tratando de asegurar que se mida el máximo efecto posible.

Dado que para cada variable se estudian dos niveles y para no tener que realizar muchas experiencias, generalmente se usan diseños experimentales de screening, tales como los factoriales fraccionados o los diseños Plackett Burman. Cuando se sospecha que los efectos son asimétricos, se deben utilizar diseños experimentales más complejos.

1.7. CALIBRACIÓN MULTIVARIADA

1.7.1. INTRODUCCIÓN

El análisis químico instrumental consta usualmente de dos etapas, calibración y predicción. En la etapa de calibración, se investigan las características del método o del instrumento con el objeto de encontrar el modelo de comportamiento ($y = f(x)$), de dos grupos de variables comúnmente llamadas dependiente (y) e independiente (x). La segunda etapa consiste en obtener las variables independientes (x) de una o más muestras, es decir conocer la concentración del analito presente, a partir de la respuesta instrumental medida (y).

Se denominan señales univariantes a aquellas señales analíticas que se miden en un determinado método instrumental en función de una única variable controlada, mientras que las señales multivariantes son aquellas que se miden en función de dos o más variables controladas.

El proceso de calibración con señales univariantes (*calibración univariada*) es relativamente sencillo y consiste en la búsqueda de una relación matemática entre la magnitud medida y el valor de la variable para un conjunto de muestras patrones de concentración conocida (construcción de una curva de calibración con patrones de un único componente).

Los métodos de calibración multivariada, son métodos quimiométricos de análisis multivariante, que tienen por objeto la predicción de la concentración de uno o más componentes en mezclas incógnitas de multicomponentes (Haaland, 1988).

La calibración multivariada involucra el estudio de mezclas de varios componentes, tanto durante la etapa de calibración como la de predicción (Booksh, 1994).

La disponibilidad de instrumentos analíticos que permiten obtener señales multivariantes y programas computacionales para aplicar métodos quimiométricos para el tratamiento de la información, ha hecho posible superar la limitación de las señales univariantes.

Mediante los métodos de calibración multivariada, se puede obtener información cuantitativa selectiva, utilizando respuestas instrumentales no selectivas, como por ejemplo el espectro de una muestra; posibilitando de esta manera la determinación simultánea de la concentración de diversos componentes en muestras complejas.

La calibración es el proceso matemático y estadístico de extracción de información, usualmente la concentración de un analito, a partir de la señal instrumental.

El instrumental analítico utilizado se puede clasificar de acuerdo al tipo de datos que provee, en instrumentos de orden cero, de primer orden, de segundo orden, de n orden (Booksh, 1994).

1.7.1.1. INSTRUMENTOS DE ORDEN CERO

Este tipo de instrumentos generan un único dato por muestra, por lo tanto el conjunto de calibración resulta un vector, tal como se muestra en la Figura 1.14. 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 y en la Figura 1.15 puede observarse un ejemplo de este tipo de calibración (curva de calibrado).

Dentro de los instrumentos de orden cero están los electrodos ión 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.

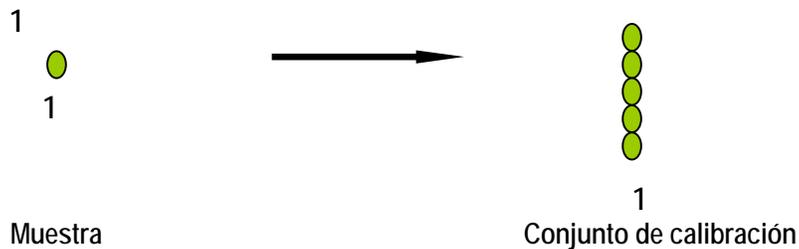


Figura 1.14: Datos de orden cero. Una muestra genera un solo dato, y el conjunto de calibración resulta un vector que contiene las medidas para las n muestras de calibración

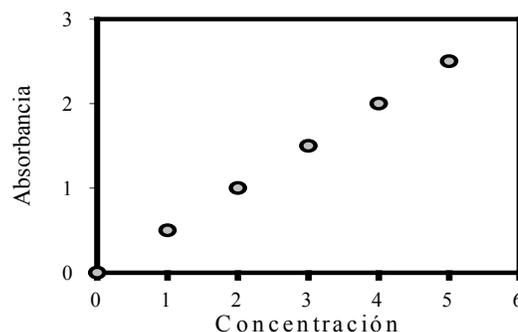


Figura 1.15: Curva de calibrado, a cada valor de concentración le corresponde un dato de absorbancia.

1.7.1.2. INSTRUMENTOS DE PRIMER ORDEN

Son ejemplos de instrumentos de primer orden, todos los tipos de espectrofotómetros y cromatógrafos. Proveen para cada muestra un tensor de primer orden, o sea un vector de datos y el conjunto de calibración resulta una matriz (Figura 1.16).

Estos instrumentos son capaces de generar múltiples mediciones para cada muestra, como por ejemplo un espectro de absorción a muchas longitudes de onda (Figura 1.17).

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

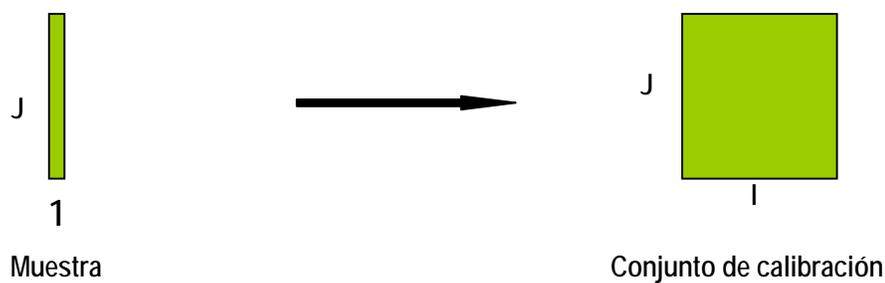


Figura 1.16: Datos de orden uno. Cada muestra genera un vector de datos, por ejemplo una muestra medida a J longitudes de onda. El conjunto de I muestras de calibración genera una matriz de $J \times I$.

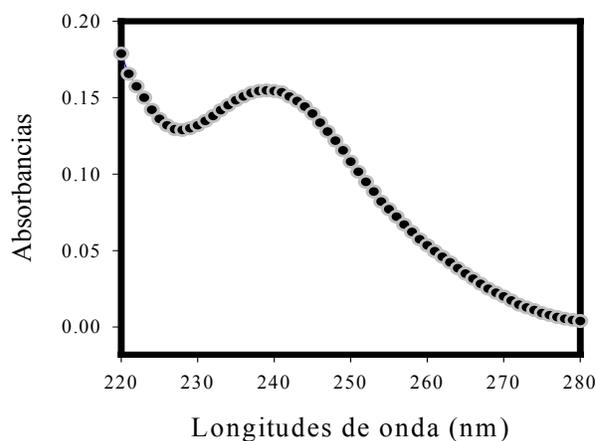


Figura 1.17: Espectro de absorción de un analito, a una determinada concentración.

1.7.1.3. INSTRUMENTOS DE SEGUNDO ORDEN

Estos instrumentos pueden generar una matriz (tensor de segundo orden) de datos por cada muestra y el conjunto de calibración es un arreglo de tres vías (Figura 1.18). Se usan en las llamadas técnicas acopladas (en inglés, *hyphenated or tandem techniques*), que incluyen por ejemplo, cromatografía de gases acoplada a espectrometría de masas (en inglés, GC/MS), espectrometría de masas acoplada a espectrometría de masas (en inglés, MS/MS), cromatografía de gases acoplada a espectrofotometría infrarroja (en inglés, GC/FT-IR), o también en técnicas como la resonancia magnética nuclear bidimensional (en inglés, 2D MNR), la espectrofluorimetría de excitación-emisión, la espectroscopia UV con arreglo de diodos en gradiente de pH (Figura 1.19), la cromatografía líquida acoplada a detector de arreglo de diodos, etc.

Este tipo de datos se analiza convenientemente mediante métodos de calibración multivariada de segundo orden, que conceptualmente son diferentes a los métodos de calibración multivariada de primer orden.

Los instrumentos de segundo orden son analíticamente más potentes que los de orden cero y los de primer orden, ya que permiten el análisis en presencia de componentes en la muestra no incluidos en el modelo de calibración (Booksh, 1994).

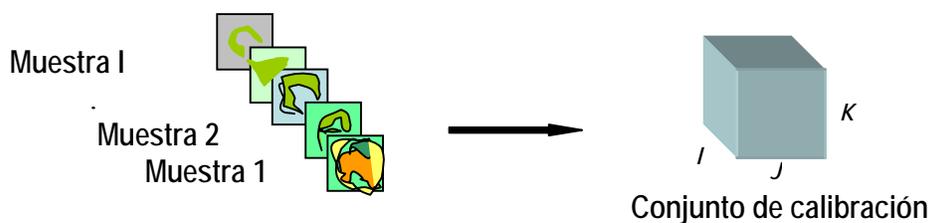


Figura 1.18: Datos de segundo orden donde cada muestra genera una matriz de datos, y el conjunto de calibración resulta un arreglo de tres vías.

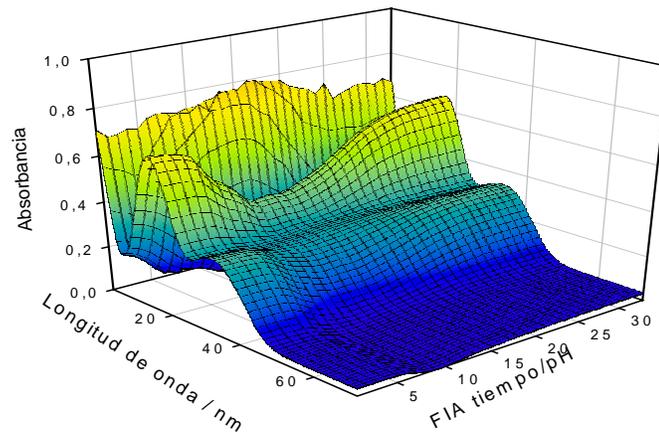


Figura 1.19: Respuesta de segundo orden obtenida con un espectrofotómetro con arreglo de diodos y gradiente de pH generado mediante un sistema en flujo (Marsili, 2004; Marsili, 2005)

1.7.1.4. INSTRUMENTOS DE ORDEN SUPERIOR

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 tercer orden (Figura 1.20). Mediante estas técnicas se obtiene una hipermatriz de datos por cada muestra y el juego de calibración resulta un arreglo de cuarto orden. Si esta misma hipermatriz es registrada a diferentes pH se obtienen datos de cuarto orden que dan origen a arreglos de quinto orden.

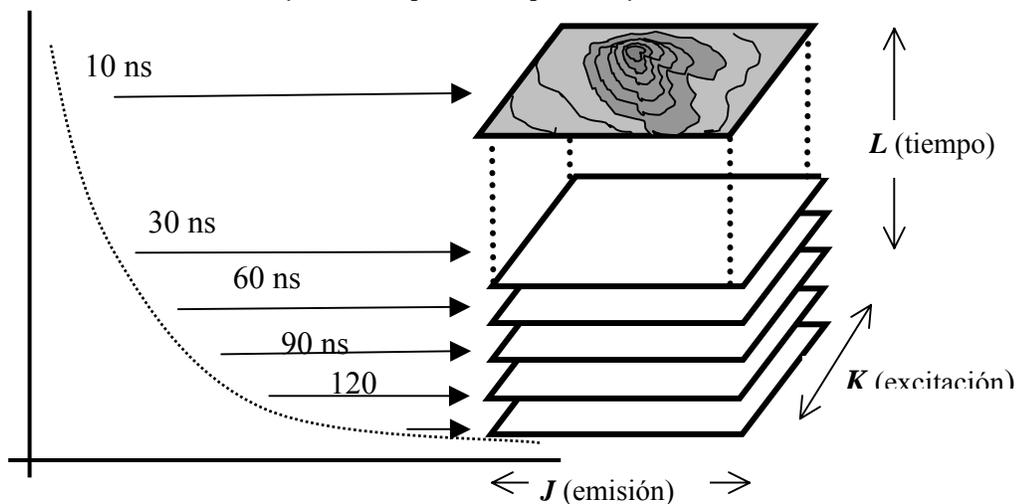


Figura 1.20: Representación esquemática de datos de tercer orden, adquiridos para una sola muestra. Mediante la técnica de fluorescencia puede obtenerse una matriz de excitación-emisión resuelta en el tiempo

Como puede observarse, los métodos de calibración pueden clasificarse de forma similar a los instrumentos analíticos. El orden de un método de calibración esta dado por el orden del tensor de datos que es necesario generar de cada muestra.

La recolección de datos de n -orden de cada una de las muestras de un juego de muestras de calibración, genera un tensor de orden $n+1$, que se puede usar para construir un modelo de calibración de orden n , estimado por un método de calibración de orden n (Booksh, 1994).

En la Tabla 1.4 se resume la clasificación hasta aquí presentada.

Tabla 1.4: Clasificación de los métodos de calibración multivariada, de acuerdo al orden del instrumento.

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

1.7.1.5. MÉTODOS DE CALIBRACIÓN MULTIVARIADA

En la literatura se presenta una gran variedad de métodos de calibración multivariada, para modelar respuestas de primer orden (Martens, 1989; Lavine, 1998). Entre los primeros modelos que fueron desarrollados, se encuentran los métodos de calibración directa: cuadrados mínimos ordinarios (*OLS*, del inglés *ordinary least squares*) (Toasaksiri, 2000), regresión por cuadrados mínimos clásica (*CLS*, del inglés *classical least squares*) (Haaland, 1999; Brown, 1991; Galeano Diaz, 1997); y dentro de los de calibración inversa: cuadrados mínimos inversos (*ILS*, del inglés *inversal least squares*) (Faber, 2000; Thomas, 1994). Estos métodos en la actualidad han sido dejados de lado por otros más robustos, cuyos rendimientos han sido objeto de numerosas publicaciones (Damiani, 2005). El método de regresión por componentes principales (*PCR*, del inglés *principal component regression*) (Lorber, 1990, Naes, 1998; Marbach, 1990) 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 (Wang, 1989; Martens, 1986; Lindberg, 1986; Sena, 2004). Este último método se ha convertido en una herramienta de rutina en el laboratorio, por la calidad de los modelos de calibración obtenidos, por la facilidad de su implementación y por la disponibilidad de software. Permite realizar la determinación de los componentes de una muestra en forma rápida, usualmente, sin necesitar de una separación previa al análisis.

Además, se han propuesto métodos de calibración multivariada basados en el concepto de señal neta del analito (*NAS*, del inglés *net analyte signal*), dentro de los cuales se pueden mencionar: el análisis lineal híbrido (*HLA*, del inglés *hibrid linear analysis*) (Berger, 1998), el HLA/XS desarrollado por Xu y Schechter (Xu, 1997; Goicoechea, 1999, Espinosa Mansilla, 2001) y el HLA/GO y el NAP/CLS desarrollados por Goicoechea y Olivieri (Ferré, 2003; Goicoechea, 2000; Goicoechea, 2001, Damiani, 2002; Muñoz de la Peña, 2002).

Por otra parte, otros métodos han sido presentados y utilizados en calibración multivariada, entre los que se pueden ser citar: Regresión de Fourier (Korany, 1990), Proyecciones Heurísticas (Liang, 1992), Resolución de Curvas (Osten, 1984), Curvas de Contenido Aparente (Mauri Aucejo, 1993), Métodos de Adición Estándar Generalizado (Jochum, 1981), Método de Aniquilación del Rango (Sánchez, 1986), Análisis de Factor (Lorber, 1984), Regresión de Hruschka (Hruschka, 1982), PCR Cuadrático (Gemperline, 1991), Algoritmo Genético (Schaffer, 1997) y Redes Neuronales (Despaigne, 1998; Carroll, 1970).

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

El modelo PARAFAC 2 (Wise, 2001), 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) (Paatero, 1994), es un interesante algoritmo para ajustar modelos bilineales, siempre que los datos sean positivos, posteriormente este método fue extendido a tres vías en 1997 (Paatero, 1997), 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 (Bro, 1998).

Muchos otros algoritmos para datos de segundo orden son descriptos, tales como: GRAFA (del inglés, *generalized rank annihilation factor analysis*) (Gerritsen, 1992), GRAM (del inglés, *generalized rank annihilation method*) (Faber, 2001; Faber, 2002), TLD (del inglés, *trilinear decomposition*) (Garrido French, 2003; Xie, 1996), DTD (del inglés, *direct trilinear decomposition*) (Bezemer, 2002; Xie, 1998), NBRA (del inglés, *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 (del inglés, *rank annihilation factor analysis*) (Booksh, 1994; Hart, 2002).

El método de cuadrados mínimos bilineales, BLLS (del inglés, *bilinear least squares*), fue originalmente presentado por Linder y Sundberg en el año 1998 (Linder, 1998; Faber, 2002) y posteriormente rescatado por Olivieri et. al (Damiani, 2004). 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 (Linder, 1998; Linder, 2002). Este método puede ser entendido como un CLS sobre los datos originales desdoblados (CLS *unfolded*) seguidos de una bilinearización residual que le permite aprovechar la ventaja de segundo orden (Olivieri, 2005). En la actualidad se han presentado publicaciones en las que se aplica la misma metodología pero previa aplicación de PLS sobre los datos desdoblados (PLS *unfolded*) (Culzoni, 2006).

MCR-ALS (del inglés, *multivariate curve resolution alternating least squares*), es probablemente uno de los mejores métodos de análisis para matrices de datos multidimensionales (Saurina, 2000).

Por otra parte, para modelar respuestas no lineales respecto de la concentración del analito, existen algunos métodos basados en la adaptación de los mencionados PCR y PLS, aunque también hay otros modelos que han sido creados especialmente para resolver el problema de la no linealidad. Entre ellos se pueden citar: regresión localmente ponderada (LWR, del Inglés *locally weighted regression*) (Centner, 1998), regresión por búsqueda de la proyección (PPR, del Inglés *projection pursuit regression*) (Sekulic, 1993), expectativa condicional alternante (ACE, del Inglés *alternating conditional expectation*) (Sekulic, 1993), regresión multivariada adaptativa basada en funciones *splines* (MARS, del Inglés *multivariate adaptive regression splines*) (Sekulic, 1993), y las redes neuronales artificiales (ANNs, del inglés *artificial neural networks*). Además debe considerarse que PCR y PLS pueden

contemplar no linealidades por la inclusión de mayor número de variables latentes en el modelo que las requeridas para un sistema lineal (Gemperline, 1991) o, mediante las versiones no lineales NL-PLS (del Inglés, *non linear PLS*) y SPL-PLS (del Inglés, *spline PLS*) (Sekulic, 1993).

En el presente trabajo de tesis se utilizaron métodos de calibración multivariada para modelar respuestas de primer orden, lineales y no lineales respecto de la concentración del analito, los que fueron aplicados en el trabajo experimental a diferentes sistemas analíticos.

En este capítulo se describe el modelo PLS de calibración inversa, utilizado para modelar respuestas lineales y los modelos NL-PLS, SPL-PLS y Redes Neuronales Artificiales utilizados para el análisis de un sistema no lineal constituido por dos fármacos con concentraciones muy desiguales, lo que obligó a trabajar en altas concentraciones del principal analito, las que saturaron el detector del espectrofotómetro UV.

1.7.2. MÉTODOS DE CALIBRACIÓN MULTIVARIADA PARA MODELAR RESPUESTAS LINEALES

1.7.2.1. MÉTODOS DE CALIBRACIÓN INVERSA

Los métodos de calibración inversa reciben este nombre porque se basan, en general, en el uso de la ley de linealidad respuesta-concentración en forma inversa a los métodos clásicos. Estos métodos permiten estudiar mezclas de componentes cuyas concentraciones, espectros y aún su identidad química se desconocen. De este modo se supera una de las grandes desventajas de CLS/OLS: la necesidad del conocimiento de los espectros de todos los componentes presentes en las mezclas incógnitas.

Desde el punto de vista matemático, la calibración directa implica la medida de espectros de calibración contenidos en la matriz R, con concentraciones de analitos contenidas en la matriz C, y obtención de la matriz de sensibilidades a partir de la ley "directa" por ajuste mediante cuadrados mínimos:

$$R = S C^T \quad (1.51)$$

En cambio, en la calibración inversa se utiliza la ley de linealidad "inversa":

$$c_k = R^T b_k \quad (1.52)$$

donde se supone la existencia de una proporcionalidad entre la concentración de un único componente k y la respuesta contenida en R, a través del vector de coeficientes de regresión b_k , que deberá obtenerse por cuadrados mínimos. Una situación similar se tiene en CLS/OLS, pero en los métodos

inversos la ecuación (1.52) se aplica cuando se desconocen los restantes componentes de las muestras.

1.7.2.2. REGRESIÓN POR CUADRADOS MÍNIMOS PARCIALES (PLS)

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 (Haaland, 1988). Para ello se introducen los valores de las concentraciones de calibración contenidos en c_k 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 (Haaland, 1988).

Existen dos tipos de métodos PLS (Haaland, 1988; Martens, 1989): 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 se llamará simplemente PLS (Damián, 2005).

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 W son ortogonales, mientras que las de P no necesariamente lo son. Es importante recalcar que las columnas de W no son autovectores propiamente dichos, sino factores que dependen de las concentraciones de calibración del analito de interés.

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 R en el vector de concentraciones c_k , obteniéndose el primer *weight loading factor*:

$$W_1 = R c_k / (c_k^T c_k) \quad (1.53)$$

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, w_1 es una aproximación por cuadrados mínimos al espectro puro del analito 1. En este paso se aprecia la introducción de información concerniente a las concentraciones contenidas en c_1 en el cálculo del primer factor.

Paso 2)

Se normaliza este factor, dividiéndolo por $(w_1^T w_1)^{1/2} = \|w_1\|$

Este paso se realiza para normalizar w_1 , para contar finalmente con un juego de factores ortonormales.

Paso 3)

Obtención del primer *score*:

$$t_1 = R^T w_1 \quad (1.54)$$

Se continúa con la suposición de que únicamente está presente el analito 1, y se calcula qué contribución del primer factor w_1 está presente en las mezclas de calibración. Estas “concentraciones” forman el vector t_1 .

Paso 4)

Obtención del primer coeficiente de regresión v_1 :

$$v_1 = t_1 c_k / (t_1^T t_1) \quad (1.55)$$

Se calcula el coeficiente de regresión que relaciona el *score* t_1 calculado en el paso 3) con las concentraciones de calibración.

Paso 5)

Obtención del primer *loading* p_1 :

$$p_1 = R^T t_1 / (t_1^T t_1) \quad (1.56)$$

Paso 6)

Cálculo de los residuos espectrales y de concentración:

$$E_{R^T} = R^T - t_1 p_1^T \quad (1.57)$$

$$e_c = c_k - V_1 t_1 \quad (1.58)$$

En los pasos 5 y 6) se asegura que los vectores t_a y w_a subsiguientes serán ortogonales entre sí. Para ello se calculan los vectores p_a , llamados *loadings*. Estos vectores, no explican la varianza espectral en la matriz R , sino que representan un intento de explicar dicha varianza, mientras simultáneamente se correlacionan los *scores* t_a con las concentraciones c_1 .

Paso 7)

Se substituyen E_R y e_c por R y c_k respectivamente en el paso 1) y se continúa hasta llegar al número de factores deseado (A).

Las matrices R y c_k se relacionan con la matriz de *scores* T a través de las siguientes ecuaciones:

$$R = P T^T \quad (1.59)$$

$$c_k = T v \quad (1.60)$$

La matriz T se obtiene a partir de (1.59) como:

$$T^T = P^+ R \quad (1.61)$$

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

Sin embargo, las columnas de P no son ortogonales, de modo que en PLS no se cumple que $P^T = P^+$.

Para solucionar este problema, se multiplica primero la ecuación (1.59) por W^T :

$$W^T R = W^T P T^T = (W^T P) T^T \quad (1.62)$$

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

$$T^T = (W^T P)^{-1} W^T R \quad (1.63)$$

Trasponiendo la ecuación anterior,

$$T = R^T W (P^T W)^{-1} \quad (1.64)$$

De acuerdo a las ecuaciones del modelo (1.59) y (1.60), si multiplicamos (1.59) por W obtenemos:

$$R^T W = T P^T W \quad (1.65)$$

Esta ecuación nos dice que la proyección de R sobre los vectores ortonormales contenidos en W da por resultado algo que podría ser definido como un nuevo *score*, que se puede denominar *score* de las señales T_r :

$$R^T W = T P^T W = T_r \quad (1.66)$$

Es decir que existe una relación entre el *score* de las señales T_r y el *score* T , al que se podría llamar *score* de la concentración o T_c

$$T_r = T_c (P^T W) = T_c D \quad (1.67)$$

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:

$$R^T = T_r W^T \quad (1.68)$$

$$c_k = T_c v \quad (1.69)$$

Coefficientes de la regresión

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

$$b_k^T = c_k^T R^+ \quad (1.69)$$

$$b_k^T = c_k^T (T^T)^+ (W^T P)^{-1} W^T \quad (1.70)$$

Trasponiendo la ecuación anterior:

$$b_k = W (P^T W)^{-1} T^+ c_k \quad (1.71)$$

Predicción basada en el score de la muestra

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

$$c_{k,un} = r^T b_k = r^T W (P^T W)^{-1} T^+ c_k \quad (1.72)$$

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

Continuando con la ecuación (1.72):

$$c_{k,un} = t_r^T T^+ c_k = t_r^T v \quad (1.73)$$

Ventajas de PLS

- 1- Utiliza una calibración inversa, lo que permite calibrar con mezclas que contengan los posibles componentes de las incógnitas, pero en las que solo se requiere conocer las concentraciones de un único componente.
- 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- Permite estudiar componentes cuyo comportamiento se desvíe ligeramente de la linealidad, introduciendo más factores en el análisis.
- 6- Optimiza la información espectral y de concentraciones referidas al componente de interés.

Desventaja de PLS

- Es sensible a interferencias no modeladas.

1.7.2.3. PREPROCESAMIENTO DE DATOS

Existe, una serie de procesamientos de los datos primarios, entre los que pueden citarse los siguientes (Haaland, 1988; Bautista, 1996):

- 1) Derivación espectral de varios órdenes, lo que permite incrementar la resolución (a costa, sin embargo, de una pérdida de sensibilidad) (Talski, 1994).
- 2) Pesado por la varianza o escalado, que se aplica fundamentalmente en análisis multivariado, cuando se tienen datos con unidades de concentración muy distintas (Kellner, 1998).
- 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 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 r debe ser sometido a idéntico preprocesamiento que la matriz R ; en el caso de centrado de datos, debe restarse de 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 s_k . La matriz de datos R contiene espectros que provienen de una serie de K analitos, cuyas concentraciones forman la matriz 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 C . Si las respuestas siguen un modelo lineal, entonces la matriz R está dada por:

$$R = S C^T \tag{1.74}$$

aunque posiblemente S también sea desconocida. Si se multiplican ambos miembros por el vector c_k se obtiene:

$$R c_k = S C^T c_k \tag{1.75}$$

El producto $C^T c_k$ podría escribirse así:

$$C^T c_k = \begin{pmatrix} c_1^T \\ c_2^T \\ c_k^T \\ \dots \\ c_K^T \end{pmatrix} c_k = \begin{pmatrix} \frac{c_1^T c_k}{c_k^T c_k} \\ \frac{c_2^T c_k}{c_k^T c_k} \\ \dots \\ \frac{c_K^T c_k}{c_k^T c_k} \end{pmatrix} \tag{1.76}$$

Cada fila de la matriz C^T corresponde al vector fila c_k , mientras que cada elemento del vector $C^T c_k$ es un escalar ($c_k^T c_k$). 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 ($c_k^T c_k$) como al producto escalar de los vectores c_k y c_k . Un diseño adecuado es aquel que hace que estos productos escalares sean nulos, o que los vectores c_k y c_k sean ortogonales entre sí, con lo cual:

$$\overline{C}^T \overline{c}_k = \begin{pmatrix} 0 \\ 0 \\ \overline{c}_k^T \overline{c}_k \\ \dots \\ 0 \end{pmatrix} \quad (1.77)$$

y, por lo tanto:

$$\overline{R} \overline{c}_k = S \overline{C}^T \overline{c}_k = [s_1 | s_2 | s_k | \dots | s_K] \begin{pmatrix} 0 \\ 0 \\ \overline{c}_k^T \overline{c}_k \\ \dots \\ 0 \end{pmatrix} = s_{k,LS} \overline{c}_k^T \overline{c}_k \quad (1.78)$$

donde se ha reemplazado s_k por $s_{k,LS}$, 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 (1.78) por el escalar $\overline{c}_k^T \overline{c}_k$ obtenemos:

$$\overline{R} \overline{c}_k / (\overline{c}_k^T \overline{c}_k) = s_{k,LS} \quad (1.79)$$

La ecuación (1.79) implica que, mediante datos centrados y diseño adecuado, puede obtenerse una buena aproximación a s_k . Una pregunta relevante es cómo diseñar adecuadamente las mezclas de calibración para que la ecuación (1.79) 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 (1.79) 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.

1.7.2.4. NÚMERO ÓPTIMO DE FACTORES (A)

Se repite el algoritmo cíclico tantas veces como la mitad del número de mezclas más uno ($l:2 + 1$). En cada ciclo se obtienen los parámetros correspondientes (w_i, t_i) y con ellos se realiza una validación interna o cruzada (método de Haaland y Thomas) (Haaland, 1988). 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 (1.80).

$$PRESS = \sum (c_{act} - c_{pred})^2 \quad (1.80)$$

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 $A=l:2+1$, se calcula $PRESS_{l:2+1}$

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 (1.81)], que es el cociente entre cada uno de los errores calculados ($PRESS_i$) y el menor error obtenido ($PRESS_{menor}$), para $i = 1, 2, \dots, l:2 + 1$.

$$F_i = \frac{PRESS_i}{PRESS_{menor}} \quad (1.81)$$

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

$$F_i < F_{\alpha, l, 1}$$

donde $(1-\alpha)$ es el percentil de Senedecor para la distribución F , e l 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 $\{F_i > F_{\alpha, l, 1} \mid \sigma_i^2 = \sigma_{menor}^2\} = 2\alpha$, donde σ_i^2 y σ_{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 1.21).

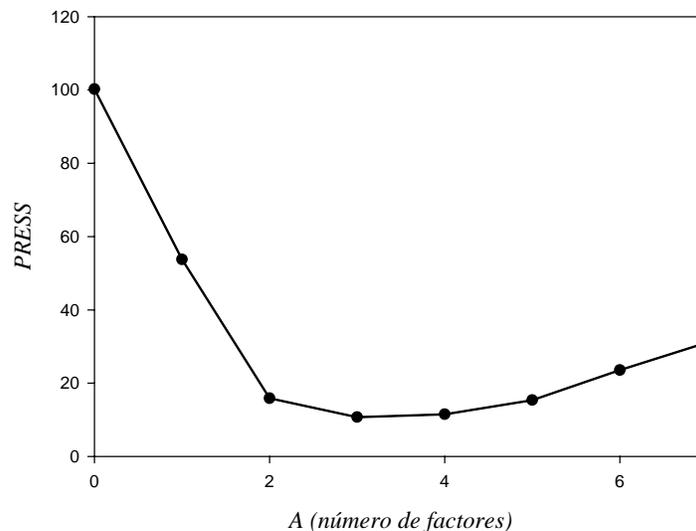


Figura 1.21: Gráfico $PRESS$ vs. A . Puede observarse que el menor $PRESS$ corresponde a $A = 3$, pero para $A = 2$ se cumple que la probabilidad de F_2 es menor de 0,75, por lo que el número de factores que se elige es 2

1.7.2.5. PARÁMETROS ESTADÍSTICOS INDICADORES DE LA CALIDAD DEL AJUSTE

Para evaluar la calidad de la calibración, una vez elegido A , pueden calcularse parámetros estadísticos, entre los que se pueden mencionar (Haaland, 1988; Martens, 1989):

a) *PRESS*: corresponde a la suma de cuadrados de los errores en las concentraciones predichas. Este parámetro ya fue descrito anteriormente, ecuación (1,80).

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

$$RMSECV = \left[\frac{1}{I} \sum_{i=1}^I (c_{act} - c_{pred})^2 \right]^{1/2} \quad (1.82)$$

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

$$REP\% = \frac{100}{\bar{c}} \left[\frac{1}{I} \sum_{i=1}^I (c_{act} - c_{pred})^2 \right]^{1/2} \quad (1.83)$$

donde c_{act} y c_{pred} corresponden a la definición anterior 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:

$$r^2 = 1 - \frac{\sum_{i=1}^I (c_{act} - c_{pred})^2}{\sum_{i=1}^I (c_{act} - \bar{c})^2} \quad (1.84)$$

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_{2pred} .

1.7.2.6.. 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 $F_{I,I-1}(c_j)$, es calculada para cada muestra durante la validación cruzada, según:

$$F_{I,I-1}(c_j) = \frac{(I-1)(e_{c_j}^2)}{\sum_{i \neq j} e_{c_i}^2} \quad (1.85)$$

donde e_{c_i} es la diferencia entre la concentración nominal de analito y la concentración predicha para la i ésima muestra de calibración, y e_{c_j} 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 *outliers*, 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 (1.86). 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 $F(r)$ (valor de F en función de la respuesta espectral) es la siguiente:

$$F(r) = I \sum_{j=1}^J (e_{r,j})^2 / \sum_{i=1}^I \sum_{j=1}^J (e_{i,j})^2 \quad (1.86)$$

donde $e_{r,j}$ son los residuos espectrales de la muestra desconocida 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”.

1.7.2.7. CIFRAS DE MÉRITO EN CALIBRACIÓN MULTIVARIADA

Análogamente al caso univariado, pueden definirse cifras de mérito correspondientes a determinaciones usando múltiples sensores.

Las cifras de mérito son valores que califican al método analítico en su conjunto (Booksh, 1994), siendo de suma importancia para conocer el rendimiento del mismo, así como para compararlo con otros métodos analíticos, incluso basados en técnicas diferentes. También resultan fundamentales para el desarrollo de nuevos métodos multivariados (Olivieri, 2006).

Estas cifras de mérito poseen idénticos significados que para calibración univariada, aunque en este caso se debe considerar para su cálculo la naturaleza multivariada del análisis, utilizando los parámetros correspondientes (Ferré, 1997). Entre las mismas se pueden citar: límite de detección, límite de cuantificación, selectividad, y sensibilidad para un analito, cuyas definiciones son las siguientes:

a) Límite de Detección:

En calibración univariada se define normalmente como la concentración mínima que da una señal instrumental estadísticamente diferente a la señal del blanco o señal de fondo y se calcula mediante la ecuación 1.46. En calibración multivariada podría formularse una expresión similar basada en los coeficientes de regresión para cada componente, según:

$$LOD_k = 3,3 S_{bl} \| b_k \| \quad (1.87)$$

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 b_k es el vector de los coeficientes finales de la regresión para el componente k .

b) Límite de Cuantificación:

En calibración univariada el límite de cuantificación se define como la mínima concentración cuantificable en forma confiable y se calcula mediante la ecuación (1.49). En calibración multivariada podría formularse una expresión similar basada en los coeficientes de regresión para cada componente, según:

$$\text{LOQ}_k = 10 S_{bl} \| \mathbf{b}_k \| \quad (1.88)$$

c) Sensibilidad:

Por analogía con la calibración univariada (sensibilidad igual a la pendiente de la curva de calibración), se define según:

$$\text{SEN}_k = 1 / \| \mathbf{b}_k \| \quad (1.89)$$

Este parámetro tiene unidades de señal/concentración.

d) Sensibilidad analítica:

La sensibilidad analítica se define como el cociente entre la sensibilidad y el nivel promedio del ruido, según:

$$\gamma_k = \text{SEN}_k / S_{bl} \quad (1.90)$$

Este parámetro tiene unidades de (concentración)⁻¹. 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:

La selectividad se define como la medida de lo bien que se puede discriminar un componente de los otros y se calcula de acuerdo a la siguiente ecuación:

$$\text{SEL}_k = 1 / (\| \mathbf{b}_k \| \| \mathbf{s}_{k,LS} \|) \quad (1.91)$$

Donde $\mathbf{s}_{k,LS}$ se aproxima mediante datos centrados, según se especifica a continuación.

$$\mathbf{s}_{k,LS} = \mathbf{R} \mathbf{c}_k / (\mathbf{c}_k^T \mathbf{c}_k) \quad (1.92)$$

1.7.3. MÉTODOS DE CALIBRACIÓN MULTIVARIADA PARA MODELAR RESPUESTAS NO LINEALES

1.7.3.1. NL-PLS Y SPL-PLS

Como se vio con anterioridad, los métodos de calibración multivariada tienen como objetivo encontrar un vector de regresión (b_k) a partir de cierta información (calibración), que permita a su vez encontrar la concentración del analito de interés (c_k) a partir del espectro (r) de una muestra incógnita (predicción), según la siguiente ecuación:

$$c_k = b_k^T r \quad (1.93)$$

En calibración inversa, b_k se obtiene a partir de la pseudoinversa de la matriz R que contiene los espectros de calibración y del vector que contiene las concentraciones del analito de interés (k) en esas mezclas de calibración (c_k). La Ecuación (1.94) muestra esta relación:

$$b_k = R^+ c_k \quad (1.94)$$

De cómo se obtiene la pseudoinversa, surgen los diferentes modelos, entre ellos PCR y PLS. En dicho procedimiento se calculan variables latentes y puntuaciones o *scores* que permiten el cálculo de b_k . Como ya se vio, la principal diferencia entre PLS y PCR es que el primero incorpora información de la concentración en el cálculo de dichas variables latentes. Para el cálculo de la pseudoinversa, las respuestas de calibración son descompuestas según la siguiente ecuación:

$$R = T P^+ \quad (1.95)$$

donde T representa la matriz de *scores* y P la matriz de variables latentes.

En el esquema de la Figura 1.22 puede apreciarse un diagrama de flujo en el que un vector r es afectado a las variables latentes (p) y a los *scores* (t), en este caso se calcularon dos factores, para finalmente encontrar la concentración del analito:

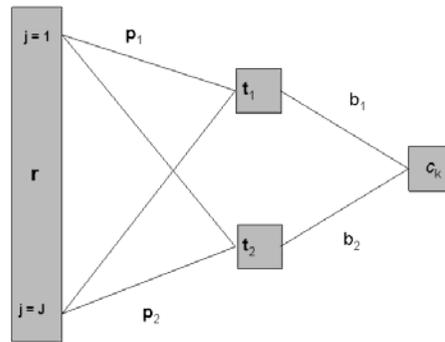


Figura 1.22: Diagrama de flujo que muestra la transferencia de la información durante la etapa de predicción en PCR y PLS

Para implementar las versiones no lineales, se procede a utilizar una matriz de scores aumentada con la contribución de las transformaciones cuadráticas de dichos scores. La Ecuación (1.96) muestra esta matriz aumentada (T_{aum}) para dos componentes, mientras que la Figura 1.23 muestra el correspondiente diagrama de flujo:

$$T_{aum} = [t_1 \ t_2 \ t_1^2 \ t_2^2] \quad (1.96)$$

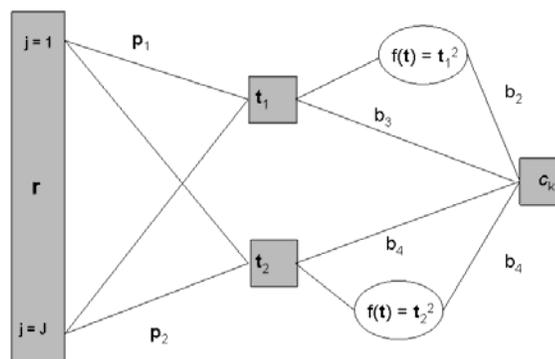


Figura 1.23: Diagrama de flujo de los algoritmos PCR y PLS no lineal durante la etapa de predicción

Se puede incorporar cualquier función en la transformación de los vectores de *scores* para construir el modelo que más convenga (Sekulic, 1993). Por otra parte, se puede evitar la especificación exacta de una transformación aplicando *splines*, originando el método SPL-NPLS (Wold, 1992).

1.7.3.2. REDES NEURONALES ARTIFICIALES

El desarrollo de las redes neuronales artificiales, propone un modelo matemático que pretende imitar el funcionamiento del cerebro. Los elementos de cálculo aritmético equivalen a las neuronas: células que procesan la información en el cerebro y la red en general equivale a un conjunto de neuronas conectadas entre sí. Por lo anterior, a estas redes se las conoce como redes neuronales artificiales (ANNs, del inglés *Artificial Neural Networks*)(Russel, 1996). Las ANNs están constituidas por nodos o unidades que están unidas mediante conexiones. A cada conexión se le asigna un peso numérico. Los pesos constituyen el principal recurso de memoria a largo plazo en las redes y el aprendizaje usualmente se realiza con la actualización de tales pesos (Kröse, 1996; Kasabov, 1998).

Las unidades constan de un conjunto de conexiones de entrada provenientes de otras unidades, un conjunto de vínculos de salida que van hacia otras unidades, un nivel de activación del momento y recursos para calcular cual será el nivel de activación del siguiente paso, con base en sus entradas y pesos respectivos. Lo importante es que en cada una de las unidades se efectúa un cálculo local con base en las entradas que le proporcionan sus vecinas, pero sin que sea necesario un control global en todo el conjunto de las unidades (Russel, 1996; Kröse, 1996; Kasabov, 1998).

1.7.3.2.1. PRINCIPIO DE LAS REDES NEURONALES

En la Figura 1.24 se muestra una unidad típica, donde cada una de las unidades realiza un cálculo sencillo. Esta unidad recibe señales de sus vínculos de entrada y calcula el correspondiente nuevo nivel de activación que envía a través de sus vínculos de salida. El cálculo del nivel de activación está basado en los valores de cada una de las señales de entrada que envía un nodo vecino, así como los pesos de cada uno de los vínculos de entrada. El cálculo está dividido en dos componentes. La primera es una componente *lineal*, denominada función de entrada (ent_i), que calcula la suma ponderada de los valores de entrada de la unidad. La segunda es una componente no lineal conocida como función de activación (g), que transforma la suma ponderada en el valor final que sirve como valor de activación de la unidad (a_i). Por lo general, todas las unidades de la red utilizan la misma función de activación (Russel, 1996; Kröse, 1996; Kasabov, 1998; Zupan, 1999).

La entrada ponderada total es la suma de las activaciones de entrada multiplicadas por sus pesos respectivos:

$$ent_i = \sum_j w_{j,i} a_j = \mathbf{w}_i^T \mathbf{a}_i \quad (1.97)$$

Los pesos correspondientes al nodo i están designados por el vector w_i y al conjunto de valores de entrada se le denomina a_i .

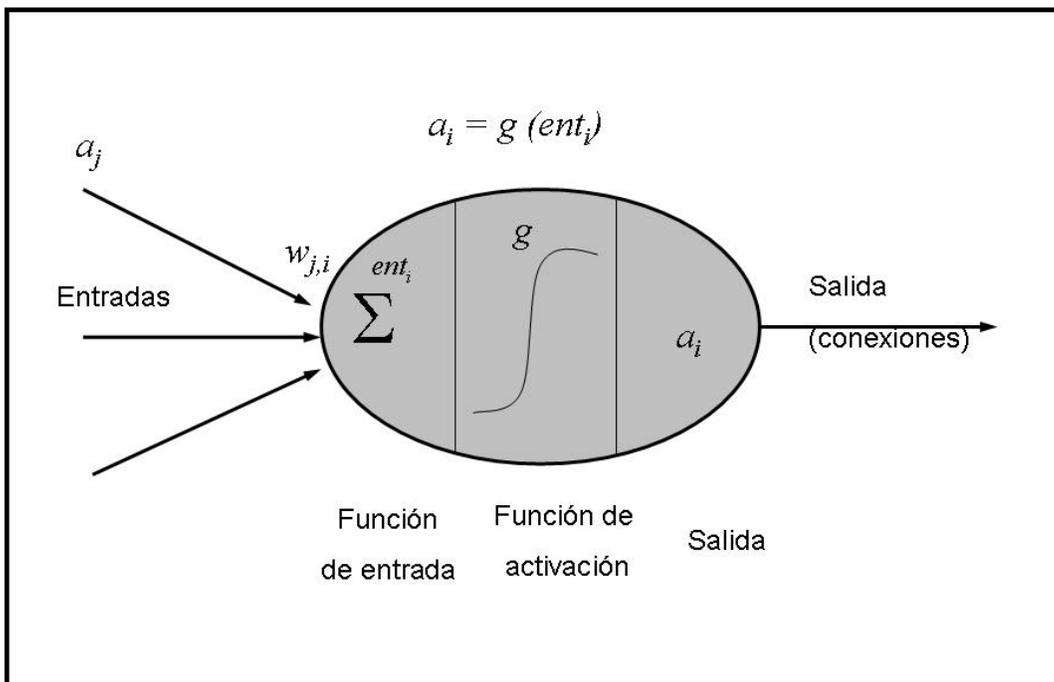


Figura 1.24: Representación esquemática de una unidad de activación o neurona

El nuevo valor de activación de la unidad se calcula mediante un paso de cómputo elemental de cada una de las unidades. Para ello, se aplica la función de activación g al resultado de la función de entrada:

$$a_i \leftarrow g(ent_i) = g\left(\sum w_{j,i} a_j\right) \quad (1.98)$$

Al usar diferentes funciones matemáticas para g se obtienen distintos modelos. Tres de los modelos más comunes son las funciones escalón, signo y sigmoidea, cuyas representaciones esquemáticas se ilustran en la Figura 1.25. La función escalón tiene un límite t , de manera que produce un resultado igual a 1 cuando la entrada es mayor que su límite y, en caso contrario produce 0. El equivalente de la motivación biológica es que 1 representa el envío del impulso al axón, y el 0 representa que éste no es enviado. El límite, o umbral representa la mínima entrada total ponderada necesaria para provocar la activación de la neurona. También pueden definirse versiones límite de las funciones signo y sigmoides.

Matemáticamente:

Función escalón $x = 1, \text{ si } ent \geq t \quad 0, \text{ si } ent < t$

Función signo $x = +1, \text{ si } ent \geq 0 \quad -1, \text{ si } ent < 0$

Función sigmoidea $x = 1 / (1 + e^{-ent})$

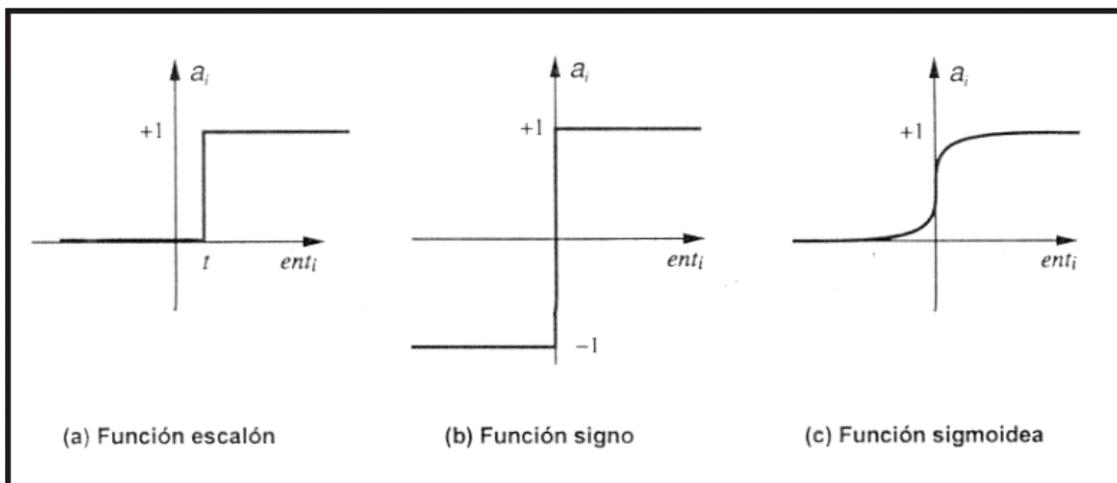


Figura 1.25: Representación gráfica de tres tipos de funciones para activación de unidades

1.7.3.2.2. ESTRUCTURA DE RED ÓPTIMA

La elección de la estructura de la red es un hecho de gran importancia. Si la red elegida es demasiado pequeña, el modelo no podrá representar a la función deseada. Si por otra parte escogemos una red demasiado grande, aunque será capaz de memorizar todos los ejemplos, no será capaz de generalizar bien cuando se presenten entradas que no haya visto anteriormente. Es decir, al igual que todos los modelos estadísticos, las redes neuronales están sujetas a un sobre ajuste cuando los parámetros (esto es, los pesos del modelo) son demasiados. Más adelante se describe como se llega a obtener la

arquitectura óptima en una red usada en calibración multivariada y como se trata de evitar el sobre ajuste cuando se entrena una red.

1.7.3.2.3. REDES NEURONALES EN CALIBRACIÓN MULTIVARIADA

El presente trabajo se restringió al uso de las llamadas redes multicapas con retroalimentación o perceptrón multicapa (*MLP*, del Inglés *multi-layer perceptron*), por lo que de aquí en adelante se describirá el funcionamiento de este tipo de redes. En la Figura 1.26 se muestra un ejemplo de red para cuatro variables descriptoras (x_1, x_2, x_3 y x_4) y una sola salida (y). Las variables son presentadas a las cuatro neuronas de entrada y son pesadas por las conexiones (w_{ij}). Posteriormente, las neuronas escondidas reciben las señales ponderadas y realizan dos tareas: una suma de las entradas y una proyección de esta suma sobre la función de transferencia (f_h), produciendo la *activación*. Luego, estas activaciones son pesadas por las conexiones (w_j'') entre las neuronas escondidas y las de salida, que realizan un tarea similar a las neuronas escondidas proyectando la suma sobre la función de transferencia (f_o). La salida es la respuesta estimada (y_{pred}), que puede ser expresada según la siguiente ecuación:

$$y_{pred} = f_o \left[\theta'' + \sum_{j=1}^{nh} w_j'' f_h \left(\sum_{i=1}^{nd} w_{ij}' x_i + \theta' \right) \right] \quad (1.99)$$

donde nd y nh son los números de neuronas de entrada y escondidas respectivamente. Aunque la red es considerada una herramienta no paramétrica, los modelos están definidos por un conjunto de parámetros ajustables, determinados por un algoritmo. Estos parámetros ajustables son los pesos (w_{ij}' y w_j'') y los umbrales (*biases*) (θ' y θ'') que actúan desplazando horizontalmente las funciones de transferencia. El procedimiento iterativo que permite determinar estos parámetros se llama aprendizaje o entrenamiento (en Inglés, *learning* o *training*). En primer lugar estos parámetros son elegidos al azar y se calcula un valor de la respuesta. Luego, en un paso de retroalimentación, el error entre la respuesta real y la predicha es minimizado por ajuste de los parámetros. Cada par de estos pasos es considerado una época. El procedimiento se repite hasta la convergencia, es decir hasta que se llega a un error pre-especificado o aceptablemente bajo.

Resumiendo, el problema consiste en una optimización en la que se busca un mínimo de error en una superficie multidimensional caracterizada por la presencia de mínimos locales.

Probablemente no se encontrará el mínimo absoluto, pero si uno local que satisfice la solución del problema (Despagne, 1998). El algoritmo de minimización más utilizado es el del gradiente descendente sucesivo, que se basa en la estimación de la primera derivada del error con respecto a cada peso (Fletcher, 1980). En el esquema de la Figura 1.26 puede apreciarse como se realiza el procedimiento.

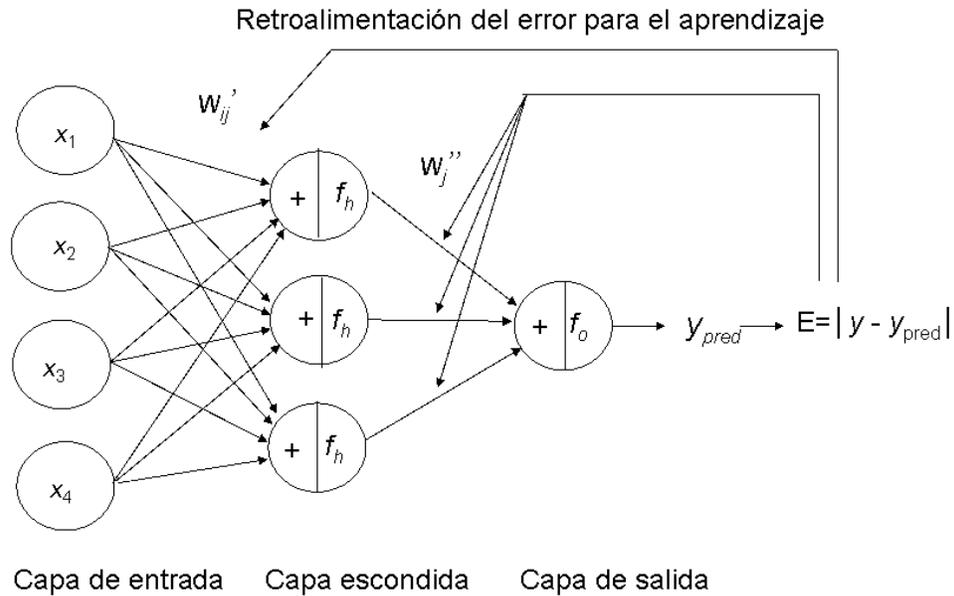


Figura 1.26: Representación gráfica de una red multicapa con retroalimentación.

Las ANNs pueden ser utilizadas para construir modelos de calibración multivariada empíricos de la forma $y = f(X) + e$. Solo se consideraran los modelos de calibración inversa, donde X designa una matriz de medidas o señales analíticas formada por n muestras. Para una muestra dada, las mediciones son descritas por un juego de variables descriptoras, x_i , un ejemplo lo constituyen los valores de absorbancia a una longitud de onda dada (espectro de absorción). Por su parte, y es un vector conteniendo las respuestas de las muestras, por ejemplo, las concentraciones de una analito de interés en un juego de mezclas de calibración (Despagne, 1998).

Es importante remarcar que las ANNs se deben utilizar cuando un juego de datos se sospecha o se sabe que es no lineal (Despagne, 1998). Desde un punto de vista matemático, un verdadero modelo es

no lineal respecto a sus parámetros. En calibración multivariada, también suele considerarse no lineal a un modelo en el que la no linealidad se presenta en la relación entre la respuesta y el descriptor. En espectroscopía, la desviación de la linealidad puede ocurrir si una muestra tiene una elevada absorción o no es homogénea, si el tamaño de la partícula no es constante en todas las muestras (para especies cristalinas) o si algunas señales están solapadas (Gemperline, 1991). Una respuesta no lineal o la presencia de luz parásita, introducen una curvatura en la función de respuesta de la concentración. Estos efectos de no linealidad pueden ser corregidos algunas veces con el uso apropiado de técnicas de preprocesamiento de los datos, aunque estas técnicas presentan limitaciones como la pérdida de sensibilidad (Hadjiiski, 1999; Gemperline, 1991). Se pueden presentar otros efectos no-lineales como equilibrios químicos asimétricos, reacciones intermoleculares, reacciones intermetálicas en electroquímica, uniones puente de hidrógeno y cambios en temperatura o composición de los solventes.

1.7.3.2.4. DESARROLLO DE MODELOS DE CALIBRACIÓN

1. *Detección de no linealidad*

Como regla general, no se debería construir un modelo ANN si no se tiene seguridad de que el sistema es no lineal. Tal como se describió en el punto 1.6.2.3., existen diversas herramientas que permiten establecer si los datos presentan un comportamiento no lineal. Como se vio, el método más simple, y en algunos casos suficiente, consiste en realizar una inspección de los residuos del modelo lineal *versus* la concentración. También ayuda mucho graficar concentración nominal *versus* la predicha por el modelo lineal (Despaigne, 1998). Algunos autores presentaron métodos complejos que han sido aplicados con éxito en algunos casos (Collado, 2002; Centner, 1998).

2. *Muestras de entrenamiento, monitoreo y validación.*

El número de muestras disponibles suele resultar un factor limitante para la aplicación de ANNs. Si se dispone de pocas muestras, el número de parámetros a ser ajustados es tal que conduce a sobreajustes. El número de muestras usadas debería ser superior a 30, pero se pueden obtener buenos resultados solo con 15 (Despaigne, 1998).

Un importante aspecto de las ANNs debe ser considerado, y es la necesidad de contar con un juego de muestras de monitoreo a fin de evitar el sobreajuste. Ocurre que las redes pueden ser entrenadas hasta un error muy bajo. Ese hecho, si bien dice que la calibración realizada es buena, implica la pérdida de capacidad de generalizar, y esto conduce a poca habilidad predictiva del modelo (sobreajuste). La

evaluación del error obtenido en la predicción del juego de monitoreo al mismo tiempo que el de aprendizaje, cuando cada época es desarrollada, permite encontrar un punto de corte. Esto quiere decir que pueden encontrarse los pesos que originan errores de calibración y de monitoreo aceptables al mismo tiempo. La Figura 1.27 muestra como evolucionan los errores y cual sería la época seleccionada de acuerdo al criterio comentado.

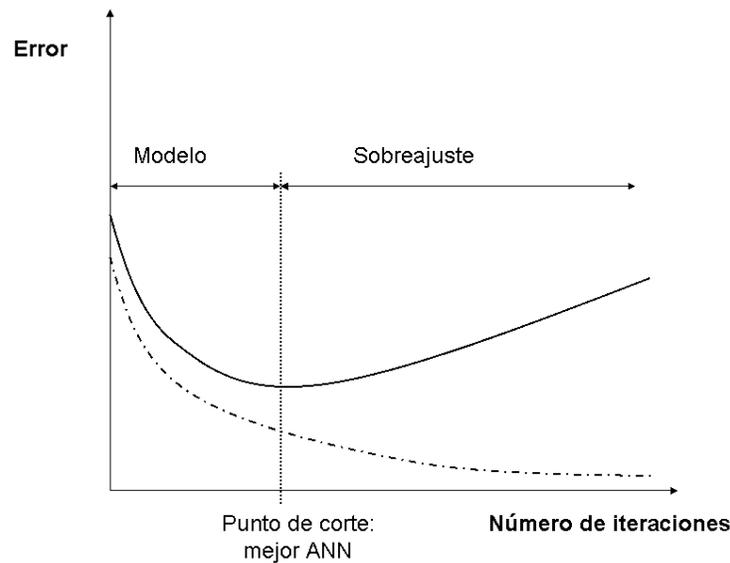


Figura 1.27: Representación gráfica de la evolución típica de los errores de calibración y monitoreo en función del número de iteraciones

Resumiendo, a diferencia de los modelos de calibración como PLS por ejemplo, las ANNs necesita de contar con muestras adicionales agrupadas en el juego de monitoreo. Algunos autores utilizan una parte del juego de muestras de calibración para monitorear. La selección se realiza al azar o aplicando herramientas como el algoritmo de Kennard-Stone que permiten dividir los juegos de una manera inteligente contemplando la variabilidad del sistema bajo estudio (Kennard, 1969).

3. Arquitectura

Generalmente, la arquitectura se representa con números que indican la cantidad de neuronas en cada capa. La red de la Figura 1.26 se representa entonces como 4–3–1. El número de capas escondidas es generalmente 1, no se han encontrado mejoras en los resultados de calibración para más de una capa

escondida (Despaigne, 1998). También es posible usar conexiones directas entre las neuronas de entrada y la de salida, pero estas conexiones generalmente tienen en cuenta las relaciones lineales entre las variables y permiten un aprendizaje más rápido (Despaigne, 1998). Es preferible que el número de neuronas en la capa escondida sea lo más pequeño posible para evitar sobreajuste.

El número de neuronas de entrada puede ser obtenido empezando con un número deliberadamente grande, y removiendo gradualmente algunas (*pruning*) hasta que la calidad del ajuste deja de mejorar. Algo importante a tener en cuenta es que en calibración multivariada se necesitarían tantas neuronas como sensores hay en el espectro (por ejemplo longitudes de onda). La solución se encuentra aplicando análisis de componentes principales (PCA) y utilizando las puntuaciones (*scores*) correspondientes a los autovectores que explique la máxima variancia del sistema.

El número de neuronas de salida debería ser uno ya que se modela una respuesta a la vez. Podría existir la excepción de querer modelar respuestas correlacionadas como las concentraciones de diferentes constituyentes de una mezcla en un sistema cerrado.

5. Entrenamiento de una red

El entrenamiento de una red es un problema de optimización. El algoritmo más utilizado es el del gradiente descendente sucesivo (*steepest-descent*), que consiste en una minimización sucesiva en la superficie del error ajustando los parámetros en un hiperespacio (Rumelhart, 1986). La tendencia a caer en mínimos locales obliga a la utilización de un término de “momento” en la actualización de los pesos. Esto permite suavizar la superficie del error y atenuar las oscilaciones en el fondo de los valles. La velocidad del algoritmo se puede aumentar usando dos parámetros adaptativos: *velocidad de aprendizaje* y *momento*.