

UBU
Universidad de Burgos

UNIVERSIDAD DE BURGOS



7300376493

Deposito

714

UNIVERSIDAD DE BURGOS



DEPARTAMENTO DE QUÍMICA
ÁREA DE QUÍMICA ANALÍTICA



ANÁLISIS MULTIVARIANTE Y TRANSFERENCIA DE CALIBRADO EN VOLTAMPEROMETRÍA

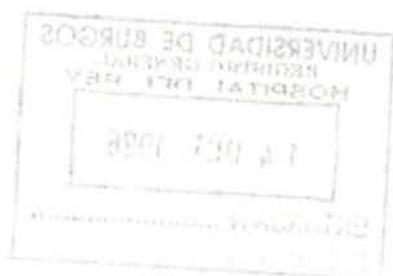
ANA HERRERO GUTIÉRREZ

TESIS DOCTORAL

1996

T
14





X730065919

**ANÁLISIS MULTIVARIANTE Y
TRANSFERENCIA DE CALIBRADO EN
VOLTAMPEROMETRÍA**

Memoria presentada en la Facultad de Ciencia y Tecnología de los Alimentos y Ciencias Químicas de la Universidad de Burgos, por Ana Herrero Gutiérrez, para optar al grado de Doctor en Ciencias Químicas.

Noviembre, 1996



UNIVERSIDAD DE BURGOS
DEPARTAMENTO DE QUIMICA

M. CRUZ ORTIZ FERNÁNDEZ, profesor Titular del Departamento de Química de la Universidad de Burgos

CERTIFICA: Que la presente memoria, titulada '*Análisis Multivariante y Transferencia de Calibrado en Voltamperometría*' y realizada por ANA HERRERO GUTIÉRREZ, ha sido llevada a cabo bajo su dirección en el Área de Química Analítica del Departamento de Química de la Universidad de Burgos.

Y para que así conste, firmo el presente certificado en Burgos, a 14 de octubre de 1996.

Fdo.: M. Cruz Ortiz Fernández





UNIVERSIDAD DE BURGOS
DEPARTAMENTO DE QUIMICA

D. FRANCISCO JAVIER ARNÁIZ GARCÍA, Director del
Departamento de Química de la Universidad de Burgos

CERTIFICA: Que habiéndose recibido en este Departamento la memoria presentada por Ana Herrero Gutiérrez '*Análisis Multivariante y Transferencia de Calibrado en Voltamperometría*', dirigida por la Dra. M. Cruz Ortiz Fernández, con la que opta al grado de Doctor, y no habiéndose presentado ninguna objeción por parte de los miembros de este Consejo de Departamento, extendiendo por la presente el preceptivo informe favorable, a fin de que prosiga su trámite ante la Comisión de Doctorado de la Universidad de Burgos.

Y para que así conste, firmo el presente certificado en Burgos, a 26 de septiembre de 1996.



Fdo.: F. Javier Arnáiz García



A mis padres

AGRADECIMIENTOS

Quiero expresar mi agradecimiento a la profesora M. Cruz Ortiz Fernández por su compromiso y dedicación en la dirección de este trabajo, así como por su entusiasmo e incondicional apoyo, sin los cuales no habría sido posible la realización del mismo.

A los profesores Jesús López Palacios y M. Julia Arcos Martínez, del Área de Química Analítica de la Universidad de Burgos, por su confianza, sus consejos y por la formación necesaria en esta disciplina científica.

A los profesores Luis A. Sarabia Peinador, M. Sagrario Sánchez Pastor y Tomás Pérez Pérez, del Departamento de Matemáticas y Computación, por su inapreciable ayuda y colaboración, así como por haberme facilitado todos los medios a su alcance para finalizar la presente memoria.

Al resto de mis compañeros del Área de Química Analítica y a todas aquellas personas que, de una manera u otra, han colaborado en la finalización de la presente Tesis Doctoral, mi más sincero agradecimiento.

La presente Memoria ha sido subvencionada parcialmente por la Caja de Burgos, entidad a la que deseo expresar por ello mi agradecimiento.

GLOSARIO

<i>ANOVA</i>	Análisis de la varianza	<i>LV</i>	Variable latente
<i>DPASV</i>	Voltamperometría diferencial de impulsos con redisolución anódica	<i>MLR</i>	Regresión lineal múltiple
<i>DPP</i>	Polarografía diferencial de impulsos	<i>MME</i>	Electrodo multimodo
<i>DPSA</i>	Potenciometría diferencial con redisolución anódica	<i>NIR</i>	Espectroscopía de infrarrojo próximo
<i>DS</i>	Método de transferencia de calibrado directa	<i>PCR</i>	Regresión sobre componentes principales
<i>ECE</i>	Reacciones acopladas electroquímica-química-electroquímica	<i>PDS</i>	Método de transferencia de calibrado directa a trozos
<i>GCT</i>	Método de transferencia de calibrado global	<i>PLS</i>	Regresión en mínimos cuadrados parciales
<i>HMDE</i>	Electrodo de gota de mercurio colgante	<i>PRESS</i>	Suma de cuadrados del error residual en predicción
<i>IME</i>	Interface para electrodos de mercurio	<i>Score</i>	Puntuación de un objeto en el nuevo espacio vectorial
<i>Leverage</i>	Peso de las predictoras de un objeto respecto del resto	<i>SE</i>	Error estándar
<i>Loading</i>	Coefficiente de las variables predictoras en el autovector	<i>SEP</i>	Error estándar en predicción
		<i>SET</i>	Error estándar del conjunto de aprendizaje
		<i>SMDE</i>	Electrodo de gota de mercurio estática

ÍNDICE GENERAL

ÍNDICE GENERAL

I.	Introducción	1
II.	Calibración con modelos blandos	9
II.1	Introducción	11
II.2	Reacciones acopladas en la determinación polarográfica de benzaldehído	17
II.2.1	Experimental	19
II.2.2	Cuestiones preliminares	21
II.2.3	Análisis univariante	42
II.2.4	Análisis multivariante	43
III.	Transferencia de calibrado	79
III.1	Introducción	81
III.2	Teoría sobre la transferencia de calibrado	87
III.2.1	Métodos que implican la corrección del modelo de calibrado	88
III.2.1.1	Transferencia de calibrado utilizando un modelo de calibrado clásico	89
III.2.1.2	Transferencia de calibrado utilizando un modelo de calibrado inverso	90
III.2.1.3	Transferencia de calibrado global (GCT)	91
III.2.2	Métodos que transforman las nuevas medidas	92
III.2.2.1	Transferencia de calibrado directa (DS)	93
III.2.2.2	Transferencia de calibrado directa a trozos (PDS)	94

III.3	Determinación simultánea de Cu, Pb, Cd y Zn mediante polarografía	97
III.3.1	Experimental	98
III.3.2	Diseño del experimento	99
III.3.3	Análisis univariante	123
III.3.4	Análisis multivariante	125
III.3.4.1	Construcción y evaluación de modelos PLS	126
III.3.4.2	Transferencia de calibrado mediante el procedimiento PDS	131
III.3.4.3	Transferencia de calibrado mediante el procedimiento GCT	152
III.4	Interferencia de compuestos intermetálicos en determinaciones por voltamperometría de redisolución	163
III.4.1	Experimental. Determinación de Cu, Pb, Cd y Zn	165
III.4.2	Diseño del experimento	166
III.4.3	Análisis univariante	167
III.4.4	Análisis multivariante	191
III.4.4.1	Construcción y evaluación de modelos PLS	191
III.4.4.2	Transferencia de calibrado mediante el procedimiento PDS	198
III.4.4.3	Transferencia de calibrado mediante el procedimiento GCT	213
III.5	Solapamiento de señales en medidas polarográficas	221
III.5.1	Experimental. Determinación de Pb, Sn y Cd	223
III.5.2	Diseño del experimento	224
III.5.3	Análisis multivariante	238
III.5.3.1	Modelos PLS	238
III.5.3.2	Transferencia de calibrado mediante el procedimiento PDS	250
III.5.4.3	Transferencia de calibrado mediante el procedimiento GCT	263

III.6 Modelado de la línea base en la determinación por voltamperometría de redisolución de Tl y Pb	269
III.6.1 Experimental	271
III.6.2 Diseño del experimento	272
III.6.3 Análisis multivariante	287
III.6.3.1 Modelos PLS	287
III.6.3.2 Transferencia de calibrado mediante el procedimiento PDS	294
III.6.3.3 Transferencia de calibrado mediante el procedimiento GCT	304
III.7 Efecto matriz y transferencia de calibrado entre instrumentos	307
III.7.1 Experimental. Determinación de Cu y Fe por DPASV	309
III.7.2 Diseño del experimento	311
III.7.3 Análisis univariante	331
III.7.4 Análisis multivariante	335
III.7.4.1 Modelos PLS	335
III.7.4.2 Transferencia de calibrado mediante el procedimiento PDS	340
III.7.4.3 Transferencia de calibrado mediante el procedimiento GCT	360
IV Selección de variables mediante un algoritmo genético	367
IV.1 Introducción	369
IV.2 Descripción de algoritmo genético utilizado	373
IV.3 Selección de variables en la determinación polarográfica de benzaldehído	377
IV.3.1 Selección de variables	378
IV.3.2 Construcción y evaluación de modelos PLS	391
IV.3.3 Comparación de resultados	404

IV.4 Selección de variables en determinaciones multianalito por voltamperometría	409
IV.4.1 Construcción y evaluación de modelos PLS	409
IV.4.2 Aspectos cualitativos relacionados con la selección de variables	427
V Conclusiones	431
VI Bibliografía	436
VII Apéndices	457

I. INTRODUCCIÓN

Se ha definido la Quimiometría como la disciplina química que utiliza herramientas matemáticas, estadísticas y otras lógico-formales para diseñar procedimientos experimentales óptimos y obtener la máxima información química de los datos experimentales. Es evidente que la estrategia formal más potente y sofisticada tiene interés en Quimiometría siempre que resuelva un problema analítico.

La Quimiometría es una disciplina en fase de expansión, con riesgo de alentar muchas expectativas y provocar grandes decepciones. Sus éxitos están motivados por la formalización de estrategias para resolver problemas, y es esta formalización la que facilita decidir cómo se debe abordar un nuevo problema analítico. Respecto de la realidad del mundo que la Química Analítica estudia, las herramientas tanto instrumentales como quimiométricas pueden conducirnos a cometer tres tipos de errores: i) afirmar que una hipótesis de trabajo es falsa cuando realmente era cierta (error de primera especie); ii) que es cierta cuando era falsa (error de segunda especie) e iii) 'resolver brillantemente un problema mal planteado' (error de tercera especie). Podemos evaluar los errores de primera y segunda especie mediante experimentación diseñada a tal efecto, pero sólo podremos evitar el tercer tipo de error mediante la exigencia explícita de sentido químico al trabajo de investigación.

Carece de sentido generar una tabla de datos experimentales '*para ver qué se obtiene aplicando un tratamiento estadístico multivariante*'. La Quimiometría es una disciplina orientada a objetivos y es por tanto imprescindible la existencia de un problema bien definido para que pueda aportar soluciones con una calidad que nunca será superior a la calidad de los análisis químicos efectuados.

Uno de los logros de la Quimiometría ha sido desarrollar la posibilidad de obtener información química cuantitativa a partir de múltiples sensores aun cuando ninguno de ellos sea muy específico. Esta vía (basada en métodos de regresión multivariantes) se ha explorado, principalmente, para técnicas espectroscópicas y en menor grado para técnicas electroquímicas.

Ahora bien, cuando se utilizan señales obtenidas electroquímicamente, además de problemas analíticos generales tales como falta de selectividad, solapamiento de picos, presencia de otras sustancias que provocan un efecto interferente y/o matriz, se tienen otros. Puede tratarse de la presencia de compuestos intermetálicos cuando se aplican técnicas que incluyen una etapa de preconcentración sobre electrodos de mercurio, o del desplazamiento relativo en el eje de potenciales de los polarogramas y voltamperogramas al modificarse ligeramente el pH o debido a otras causas, etc. En este marco general se encuadra el trabajo de investigación que se presenta en esta Memoria.

Un aspecto no estudiado en la bibliografía es la utilidad de los modelos blandos cuando se producen reacciones electródicas acopladas que, como es sabido, provocan asimetrías y desplazamientos en los picos polarográficos. Cuando se utiliza únicamente el máximo de pico como señal analítica para correlacionarla con la concentración en el proceso de calibrado, es necesario manipular adecuadamente variables instrumentales como el pH, la concentración del electrólito soporte, etc. Sin embargo, si se prescinde de la búsqueda de una señal específica y selectiva sustituyéndola por la idea de que todo el polarograma encierra información, no sólo del analito de interés como tal sino de todo el sistema químico formado por las especies electroactivas, cabe pensar que el calibrado es viable con independencia de las modificaciones de las variables instrumentales o de las condiciones experimentales. Un caso particularmente interesante es el del benzaldehído, cuyas reacciones electródicas acopladas implican que el valor del pH del medio es crítico para poder realizar determinaciones cuantitativas mediante técnicas polarográficas. El estudio de la viabilidad de calibrados multivariantes blandos en este problema y el análisis del significado químico de los modelos obtenidos constituye el Capítulo II de esta Memoria.

Los modelos de regresión sobre variables latentes hacen viable pensar en la posibilidad de utilizar las intensidades registradas a múltiples potenciales para determinar la presencia de varios analitos en una muestra, aun cuando no sean señales completamente específicas de alguno o ninguno de los analitos en cuestión. También es frecuente leer las ventajas de las técnicas electroanalíticas (bajo coste, posibilidad de medir pequeñas concentraciones, en muchas ocasiones no se necesita pretratamiento de la muestra, etc.). Sin embargo, la necesidad de hacer un calibrado completo antes de cada medida es un problema práctico que dificulta su utilización en el caso de la determinación simultánea de varios analitos de modo rutinario por su coste en tiempo y esfuerzo experimental. Formalmente, la situación es que si "a" es el número de analitos a determinar y se utilizan "p" patrones de calibrado para cada

uno de ellos es necesario disponer de p^3 voltamperogramas completos (para determinar tres analitos con cinco patrones de calibrado resulta un total de 125 voltamperogramas). Piénsese lo que significa monitorizar un control de este tipo por ejemplo tres veces al día durante largos períodos de tiempo.

Para reducir el número de voltamperogramas necesario en la etapa del calibrado previa a cada medida de control rutinario, sin que el calibrado sufra una pérdida significativa de la capacidad de predicción, es necesario vincular cada calibrado secundario con uno inicial o calibrado "patrón".

Esta vinculación o transferencia puede realizarse de dos formas: i) seleccionando una parte de las muestras del calibrado patrón y vinculando los voltamperogramas registrados en el calibrado secundario con los correspondientes al patrón (así se dispone de una matriz de transferencia que aplicada a un voltamperograma problema registrado con el instrumento secundario permite transformarlo en un voltamperograma con las características que hubiera tenido en el instrumento patrón y, por tanto, al que puede aplicarse el modelo del calibrado patrón basado en el diseño completo); ii) seleccionando una parte de las muestras del calibrado patrón, determinando con ellas el modelo de calibrado secundario y relacionando matemáticamente ambos modelos. En ambos casos se trata de transferir la información de más calidad obtenida con mayor esfuerzo experimental (calibrado patrón) a otras determinaciones posteriores realizadas con menor esfuerzo experimental de modo que globalmente el resultado mantenga la calidad analítica de la determinación patrón.

Ambas estrategias se han comenzado a diseñar muy recientemente en espectroscopía NIR, técnica que presenta también la necesidad de un extenso número de patrones en la calibración. Han sido el primer paso en una cuestión analítica muy importante como es la estandarización de instrumentos.

Hasta el momento no se ha publicado ningún resultado que haga referencia a la estandarización de instrumentos para voltamperometría. A este respecto se plantean dos cuestiones específicas: i) los desplazamientos que pueden sufrir los voltamperogramas respecto del potencial de un calibrado a otro, incluso cuando se realiza la medida con el mismo instrumento, son más severos que en espectroscopía, lo que se traduce en fuertes fallos de linealidad; ii) el tiempo necesario para la limpieza y puesta a punto del material y el instrumental voltamperométrico aconseja la selección de esquemas de adiciones en los que las concentraciones de cada analito sigan una secuencia creciente. Es decir, sería deseable

introducir restricciones en los diseños experimentales con el fin de reducir el esfuerzo experimental requerido en cada prueba.

Para evaluar la viabilidad de la transferencia de calibrado multivariante en electroanálisis se han aplicado las dos estrategias de estandarización instrumental antes descritas a situaciones experimentales que de por sí encierran dificultades prácticamente insuperables para llevar a cabo un calibrado desde el punto de vista de la señal específica:

- Variaciones no esperadas en la altura de pico. En la determinación simultánea de Cu, Pb, Cd y Zn mediante polarografía se observa que, a pesar de obtenerse picos polarográficos bien definidos, el correspondiente al cobre disminuye cuando se adiciona zinc y viceversa, también es perceptible una disminución de ambos al adicionar cadmio.
- Formación de compuestos intermetálicos en determinaciones por voltamperometría de redisolución sobre electrodo de mercurio.
- Solapamiento de las señales del estaño y el plomo en determinaciones polarográficas de mezclas ternarias Pb(II)-Sn(IV)-Cd(II).
- Efecto causado por la presencia de una sustancia electroactiva en el electrólito soporte que proporciona una señal ajena a los analitos que se adicionan en una voltamperometría de redisolución para determinar Tl(I) y Pb(II) con señales solapadas.
- Efecto matriz debido a la presencia de exceso de Fe(III) en la determinación de Cu(II) mediante voltamperometría de redisolución.

En los cuatro primeros casos la estandarización instrumental se ha efectuado utilizando el mismo instrumento en días distintos, mientras que en el quinto intervienen dos instrumentos distintos, llevándose a cabo el procedimiento de transferencia de calibrado entre ambos también a lo largo del tiempo.

El Capítulo IV se dedica a la selección de variables mediante un algoritmo genético. Es claro que no se espera una mejora significativa de los resultados, pero como en muchos de los sistemas estudiados existen problemas de efecto interferente y/o matriz, cabe esperar que una selección de los potenciales que son realmente necesarios para describir la concentración del analito problema señale aquellas intensidades que son necesarias para explicar dicha concentración, aun cuando no se correspondan con el pico específico del

analito. Es decir, la selección de intensidades señalará al químico analista la presencia de efectos no esperados relacionados con el analito problema así como los potenciales a los que se manifiestan dichos efectos.

El Capítulo V recoge las conclusiones alcanzadas y en el VI se relacionan todas las referencias bibliográficas por orden de aparición en el texto. La Memoria concluye con dos apéndices, el primero de los cuales es un programa, en BASIC, que permite transformar los datos registrados en archivos procedentes de la aplicación de toma de datos Vesuv 10 (Metrohm A.G.) en una tabla de datos con la estructura adecuada para los tratamientos multivariantes realizados tanto con PARVUS como con MATLAB. El segundo apéndice incluye las m-funciones escritas en formato MATLAB ('*transfpds.mat*', '*transfgc.mat*' y '*gct.mat*') que han sido utilizadas para obtener las transferencias de calibrado por los dos métodos indicados.

En cuanto a la estructura formal de esta Memoria se debe señalar que la revisión bibliográfica y la descripción detallada de la metodología experimental y computacional seguida se ha efectuado separadamente en cada Capítulo para facilitar la integración del problema químico planteado, su relación con otros problemas analíticos de interés y la solución propuesta.

II. CALIBRACIÓN CON MODELOS BLANDOS

II.1 INTRODUCCIÓN

El aumento de la complejidad de los procesos químicos y los avances realizados en la ciencia de los materiales y en el diseño y fabricación de sistemas de medida han dado como resultado un enorme crecimiento en el número de datos que pueden obtenerse de un sistema químico de interés. No obstante, con gran frecuencia, muchas de las variables medidas se hallan altamente correlacionadas. Ejemplos concretos [1] se pueden encontrar en las áreas de Química Orgánica y Química Analítica, en el control de procesos, la biotecnología, la ciencia de los alimentos, la investigación farmacológica y medioambiental, etc.

A pesar de que en algunas ocasiones el interés del investigador se centra en el descubrimiento, análisis y descripción de las relaciones estructurales existentes entre las variables medidas, en la mayoría de las aplicaciones el objetivo principal es relacionar un subconjunto de variables, denominadas *variables respuesta*, con el resto de las variables, denominadas *predictoras*. De hecho, la finalidad del análisis suele consistir en modelar la señal correspondiente a un conjunto de sensores en función de la concentración del analito de interés, es decir, lo que en general se denomina calibración [2].

Si el conjunto de datos experimentales consiste en p variables predictoras y $q=1$ variable respuesta medida en cada uno de n objetos, el modelo de regresión multivariante básico que se puede construir se conoce como regresión lineal múltiple (MLR), y se corresponde con la expresión

$$Y = XB + E \quad (1)$$

donde X ($n \times p$) e Y ($n \times q$) son matrices que contienen las variables predictoras y respuesta respectivamente, E ($n \times q$) es la matriz de los errores aleatorios y B ($p \times q$) es la matriz de los coeficientes de regresión. Estos coeficientes se obtienen utilizando el criterio de mínimos cuadrados, por lo que la solución (única cuando $n \geq p$ y $\text{rango}(X) = p$) viene dada por

$$\hat{B} = (X^T X)^{-1} X^T Y \quad (2)$$

La respuesta estimada para un futuro objeto, cuyos valores de las variables predictoras se anotan por \mathbf{x}_0 , vendrá dada por

$$\hat{y}_0 = \mathbf{x}_0 \hat{\mathbf{B}} \quad (3)$$

de modo que el intervalo de confianza al nivel α es

$$\hat{y}_0 \pm t_{\alpha, n-p, \text{bil.}} \text{Var}(\hat{y}_0) \quad (4)$$

donde, siendo $\text{Var}(\mathbf{Y}) = \sigma^2$ y $\hat{\sigma}^2$ es su valor estimado

$$\text{Var}(\hat{y}_0) = \mathbf{x}_0 \text{Var}(\hat{\mathbf{B}}) \mathbf{x}_0^T = \mathbf{x}_0 (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{x}_0^T \hat{\sigma}^2 \quad (5)$$

Nótese que resolver la ecuación (1) haciendo uso de (2) requiere invertir la matriz $\mathbf{X}^T \mathbf{X}$, por lo que el método de mínimos cuadrados resulta adecuado para estimar los coeficientes de regresión cuando las variables predictoras son ortogonales o se encuentran muy cerca de serlo. La existencia de correlaciones entre las variables predictoras lleva a que, en el mejor de los casos, el determinante de $\mathbf{X}^T \mathbf{X}$ sea muy pequeño, con lo que los elementos de la matriz $(\mathbf{X}^T \mathbf{X})^{-1}$ son muy grandes. Y si se consideran las ecuaciones (4) y (5), se observa que esto conduce a intervalos de confianza muy grandes para los coeficientes de regresión y, por tanto, a la construcción de modelos de calibrado muy inestables.

En algunos casos, el problema puede obviarse utilizando un diseño experimental adecuado, que intente minimizar las relaciones internas entre variables predictoras. Sin embargo, muchos instrumentos de análisis proporcionan medidas que inevitablemente están correlacionadas entre sí, como es el caso de un espectrofotómetro de UV o de un polarógrafo [3-5]. En este último, la intensidad registrada a un potencial determinado se encuentra muy correlacionada con la registrada a un potencial próximo.

Asimismo, hay que considerar el hecho de que una de las hipótesis del modelo presupone que la incertidumbre experimental está asociada con la respuesta (concentración), mientras que las variables predictoras se encuentran libres de error aleatorio. Pero en la práctica, si el ruido de las variables predictoras es del mismo orden que las propias variables, el modelo de regresión será poco realista y, en consecuencia, ineficaz.

Estos métodos de regresión, que ajustan los datos a un modelo determinado, se han utilizado en una gran variedad de contextos científicos denominados '*ciencia dura*', basada en modelos duros (hard modelling), aunque quizás deberían haberse llamado '*ciencia endurecida*' [6]. En estos métodos, el modelo lineal es una representación adecuada de la realidad, con un pequeño número de parámetros que deben ser determinados. Habitualmente

se caracterizan porque no suele existir duda en cuanto a lo apropiado del modelo elegido, así como por una relación adecuada entre las variables predictoras y el ruido que las acompaña, y porque el criterio de mínimos cuadrados resulta correcto.

En contraposición a lo anterior, se han desarrollado una serie de aplicaciones científicas, denominadas '*blandas*' (soft modelling), en las que el modelo de calibrado se construye a partir de los propios datos, en lugar de forzar a que los datos se ajusten a un modelo externo [7]. Estas técnicas son útiles cuando existe un gran número de variables predictoras y cuando se dispone de un limitado conocimiento formalizable sobre el problema. En general, lo que se desea principalmente es que el modelo construido tenga una gran capacidad de predicción sin atribuir ningún significado a los coeficientes del modelo. Ejemplos representativos de este tipo de métodos son las regresiones sobre componentes principales, PCR, y por mínimos cuadrados parciales, PLS, en las cuales se utiliza la información aportada por algunas o todas las variables predictoras para construir unas nuevas variables, utilizando como criterio la consecución de una buena calidad en la predicción.

La técnica de regresión por mínimos cuadrados parciales, PLS, se ha presentado a menudo como un algoritmo o familia de algoritmos [8], más que como un modelo multilíneal formal, especialmente en libros relacionados con la Quimiometría y/o la Química Analítica [9-12]. Además, las modificaciones surgidas por variaciones en el algoritmo básico no siempre conducen a los mismos resultados [13], y en algunos casos vienen impuestas por la necesidad de manejar grandes grupos de datos, como ocurren en el análisis de imágenes [14].

Introducida por Wold [15] en 1975, la mayor parte del desarrollo realizado sobre la técnica PLS ha sido realizado en los países escandinavos, basándose principalmente en el procedimiento NIPALS [16, 17]. Éste último es un algoritmo para el cálculo de autovectores y autovalores por el método de las potencias, que calcula uno a uno secuencialmente los autovectores mediante un procedimiento iterativo que esencialmente es un regresión por mínimos cuadrados univariante en cada iteración. PLS actúa hasta cierto punto del mismo modo que la regresión sobre componentes principales [18, 19], PCR, una técnica de regresión ampliamente utilizada en Quimiometría y Química Analítica.

En la regresión sobre componentes principales, en primer lugar se calculan las componentes principales de la matriz X formada por las variables predictoras, y una vez hecho esto, se construye una regresión multilíneal por mínimos cuadrados utilizando las componentes principales como predictoras. Una de las razones fundamentales para utilizar

PCR se basa en el problema de la multicolinealidad entre las variables predictoras y el efecto negativo que ésta ejerce sobre la regresión lineal múltiple, MLR, especialmente en la incertidumbre de los coeficientes y en la correlación entre los mismos, como se ha señalado anteriormente. Como consecuencia de ello, las predicciones hechas con MLR en presencia de colinealidad no son buenas.

El uso de PCR evita por completo esta dificultad porque las componentes principales son variables ortogonales y además permiten efectuar un filtrado del ruido de las variables seleccionando las correspondientes a los mayores autovalores. La gran desventaja de esta técnica de regresión es que la transformación de las variables predictoras (obtención de las componentes principales) es independiente de la respuesta. Una alternativa es la técnica de regresión PLS.

La diferencia fundamental consiste en que PLS construye variables latentes de forma que expliquen simultáneamente la mayor variabilidad posible de las predictoras y de la respuesta [20]. La regresión PLS busca en el espacio de las variables predictoras direcciones de gran variabilidad, pero evitando las no correlacionadas con la respuesta. Esta propiedad es especialmente deseable cuando se analizan datos de tipo electroquímico, ya que al utilizar variables latentes se tienen en cuenta las posibles relaciones internas que intervienen en la intensidad registrada.

Técnicamente el criterio optimizado por PLS es la media geométrica del criterio que conduce a una regresión PCR y del que conduce a la regresión MLR. En este sentido, equilibra la tendencia al sobreajuste (overfitting) de MLR con la tendencia de PCR a desarrollar modelos con variables latentes no significativas.

Por otro lado, recientemente se ha demostrado la enorme importancia de la validación cruzada completa [21] (full crossvalidation) frente a la habitual validación cruzada parcial [22, 23]. Dicho de otro modo, la necesidad de que en el proceso de cómputo del modelo PLS no intervenga de ningún modo el grupo de cancelación, ni siquiera por la vía indirecta de calcular la media y varianza con todos los datos. Esto es especialmente importante cuando se utiliza una regresión blanda, como PLS, ya que la única finalidad de interés que se le puede exigir al modelo construido es que tenga una buena capacidad de predicción, puesto que se ha sacrificado la capacidad de descripción al admitir un sesgo controlado en los coeficientes del modelo multilineal.

Sin embargo, como en Química Analítica los modelos de calibrado se corresponden con sistemas químicos reproducibles, en otro caso carecen de interés analítico, es frecuente que la estructura de las variables latentes muestre características interpretables químicamente. Estas interpretaciones son de gran ayuda al químico analítico especialmente cuando el calibrado presenta anomalías sistemáticas, porque le permiten inducir las posibles razones químicas de las mismas.

Los modelos de regresión PLS han sido aplicados en un amplio rango de áreas científicas diferentes [24]. Esta poderosa herramienta multivariante ha sido aplicada con éxito dentro del ámbito de la Química Analítica en análisis por espectrofotometría ultravioleta [25, 26], en determinaciones cinético espectrofotométricas [27, 28] y fluorimétricas [29], en espectroscopía de absorción electrónica [30, 31], en análisis con inyección en flujo [32], etc.

De hecho, muchas de las técnicas electroanalíticas existentes generan gran número de datos que pueden ser modelados utilizando la regresión PLS. Ésta ha sido aplicada por Henrion y otros [33] para resolver cuantitativamente señales solapadas obtenidas por voltamperometría diferencial de impulsos con redisolución anódica. Otto y Thomas [34] han utilizado esta regresión multivariante en el desarrollo de un modelo de calibrado para una batería de electrodos selectivos utilizados en la determinación simultánea de cuatro cationes. Jagner y otros [35] han aplicado la regresión PLS en la determinación de Fe(III) y Ti(IV) mediante potenciometría de redisolución. Herrero y otros [36] han determinado cobre en vino mediante la regresión PLS, resolviendo el problema planteado por el equilibrio establecido entre los iones Cu(I)/Cu(II) que implica la existencia de relaciones internas entre algunas de las variables predictoras.

Los excelentes resultados ofrecidos por esta técnica de regresión multivariante en diferentes situaciones han conducido a su utilización en la presente memoria. En primer lugar, la regresión PLS se ha aplicado en la determinación polarográfica de un único analito, benzaldehído, bajo diferentes condiciones experimentales que conducen a señales complejas, debido a la existencia de reacciones acopladas a la transferencia electrónica.

En capítulos posteriores, esta técnica de regresión ha sido aplicada también en determinaciones multianalito en las que los análisis se llevan a cabo mediante polarografía y voltamperometría de redisolución. En las mismas, diferentes problemas relacionados con presencia de interferencias, solapamiento de señales, efecto matriz, etc., han sido resueltos mediante la utilización de esta técnica multivariante. Además se han utilizado procedimientos

de transferencia de calibrado para reducir el esfuerzo experimental necesario cuando esas determinaciones multianalito se realizan de forma rutinaria.

II.2 REACCIONES ACOPLADAS EN LA DETERMINACIÓN POLAROGRÁFICA DE BENZALDEHÍDO

Las técnicas polarográficas han alcanzado un notable desarrollo en los últimos años. Ello ha permitido la aplicación rutinaria de esta metodología electroanalítica a la determinación de un considerable número de compuestos orgánicos e inorgánicos mediante la utilización de procedimientos conceptualmente simples y experimentalmente fáciles de realizar. Una de las técnicas polarográficas comúnmente aplicada en análisis de compuestos orgánicos [37-39] es la polarografía diferencial de impulsos (DPP).

Los métodos polarográficos se basan en procesos de transferencia de carga que tienen lugar en el electrodo [40] y su aplicabilidad está sujeta a que el analito en estudio pueda reducirse u oxidarse electroquímicamente dentro del intervalo de potencial definido por el disolvente, el electrólito soporte y el electrodo empleado.

En procesos sencillos la transferencia de carga se produce a partir de un potencial determinado, característico de la especie en estudio, originando un paso de corriente cuya magnitud depende de la intensidad del proceso y, en ausencia de otras complicaciones, es proporcional a la concentración de dicha especie en disolución.

Las reacciones electródicas pueden transcurrir en una o varias etapas. Aquellas reacciones con una única etapa se presentan generalmente cuando la reacción es monoelectrónica. Mientras que, si el número de electrones intercambiados en el electrodo es superior, incluso en reacciones de simple transferencia de carga en las que intervienen varios electrones, éstas suelen desarrollarse en una serie de etapas cuya sucesión es energéticamente más favorable que la reacción formulada en una etapa única. La velocidad de cada una de esas reacciones puede ser diferente, siendo la etapa más lenta la que limita la velocidad global del proceso y representa la etapa determinante de la velocidad de reacción.

En el transcurso del proceso electródico en sistemas orgánicos, frecuentemente, se forman y rompen enlaces, de modo que la transferencia de carga está acompañada de otras

reacciones químicas [41]. A veces, las etapas elementales tienen lugar sucesivamente, es decir, el producto formado en una etapa es el reactivo de la siguiente, y otras veces se desarrollan de modo simultáneo, de manera que una misma especie reacciona por caminos diferentes para dar productos distintos, o se producen a un mismo tiempo varias reacciones en el mismo electrodo. En el primer caso se habla de reacciones consecutivas, y en el segundo, de reacciones paralelas.

El que se produzcan estas reacciones acopladas influye en la forma del polarograma, haciendo que los picos presenten asimetrías, se desplacen, etc. De hecho, algunas etapas de reacción no pueden ser seguidas debido a la existencia de interferencias, que pueden estar causadas por un solapamiento de los picos correspondientes a procesos consecutivos. El resultado de todo ello es una señal de poca calidad analítica para ser tratada de forma clásica.

La utilización de señales específicas y selectivas en estas técnicas es esencial para un correcto establecimiento de la relación entre las señales obtenidas y las correspondientes concentraciones de los analitos, que permita la cuantificación de los mismos. La mejora de la calidad de estas curvas polarográficas se consigue, en general, mediante la modificación adecuada de variables experimentales como pH, concentración de electrólito soporte, etc., e incluso con la utilización de técnicas instrumentales cada vez más sofisticadas. De hecho, la influencia del pH en muchos de los procesos que intervienen en las señales electroquímicas impone severas restricciones en cuanto a los valores del mismo a los cuales se han de realizar las correspondientes determinaciones [42-44].

No obstante, la gran cantidad de datos electroquímicos que proporciona la instrumentación actual posibilita abordar problemas de calibrado con técnicas multivariantes, las cuales permiten utilizar toda la información contenida en el polarograma y no sólo la del máximo de pico como en los métodos univariantes clásicos. La complejidad de las matrices en las que suelen estar contenidos los analitos de interés hace a veces difícil utilizar los planteamientos clásicos de búsqueda de condiciones experimentales que conduzcan a la obtención de señales específicas y selectivas, por lo que la alternativa ofrecida por un tratamiento multivariante parece adecuada para este tipo de problemas.

El uso de la regresión en mínimos cuadrados parciales, PLS, como una técnica de regresión multivariante fue promovido inicialmente dentro del campo de la Quimiometría, y resulta de gran interés siempre que se modela una respuesta como función de múltiples variables predictoras [45, 46]. Ha sido ampliamente utilizada en la resolución de problemas

similares con datos de tipo espectroscópico [9], aunque su uso en el campo del electroanálisis se está viendo incrementado lentamente [5, 49].

El objetivo de este trabajo ha sido la aplicación de las técnicas multivariantes en un problema de calibrado donde se sabe que, además de las reacciones electródicas responsables de los picos polarográficos, coexisten reacciones químicas acopladas. El análisis multivariante se ha aplicado a las señales polarográficas obtenidas en la determinación electroquímica de benzaldehído en medios que difieren en el valor del pH, lo cual provoca en los polarogramas además de desplazamientos de los picos a lo largo del eje de potenciales, la aparición de nuevos fenómenos electródicos y químicos sobre el electrodo.

En análisis de alimentos, principalmente de aquellos elaborados con extracto de almendras, los métodos oficiales de análisis para determinar la cantidad de benzaldehído se basan en la utilización de métodos gravimétricos y métodos de espectroscopía ultravioleta [50]. Así mismo, la determinación de benzaldehído en gas residual [51] se ha llevado a cabo mediante cromatografía de líquidos de alta eficacia, HPLC. Este compuesto orgánico también se puede encontrar como impureza en disoluciones de alcohol bencílico [52], así como un compuesto derivado del mismo, la hidrobencina [53], aunque este último se halla en menor cantidad. El análisis polarográfico [54, 55] supone una alternativa que permite realizar determinaciones con mayor precisión que los métodos gravimétricos, y con una instrumentación menos costosa que la requerida en las determinaciones espectroscópicas.

Especial importancia se ha concedido a la interpretabilidad química de las variables latentes del modelo PLS construido para cada valor de pH.

Por último, se ha realizado un estudio comparativo de los errores obtenidos en el cálculo de las concentraciones de analito en las distintas situaciones experimentales.

II.2.1 Experimental

Reactivos

Todos los reactivos utilizados en este estudio fueron reactivos de calidad para análisis, salvo el correspondiente al benzaldehído que fue de calidad para síntesis (pureza > 99%), sin que se haya realizado una purificación posterior de ninguno de ellos. Todas las disoluciones se

han preparado con agua bidestilada y desionizada obtenida a través del sistema Barnstead NANO Pure II. Como gas inerte para eliminar el oxígeno disuelto en las muestras se ha utilizado nitrógeno del 99.997%.

Todas las disoluciones patrón de benzaldehído fueron preparadas con concentración $9.89 \cdot 10^{-2}$ M en medio alcohólico 20%. Las soluciones tampón utilizadas fueron patrones McIlvaine, obtenidos mediante la combinación adecuada de diferentes cantidades de disoluciones de ácido cítrico 0.1 M y Na_2HPO_4 0.2 M [56] para proporcionar diez valores de pH diferentes, equiespaciados de 2.2 a 7.6.

Instrumental

Las medidas polarográficas fueron realizadas utilizando un procesador Metrohm 646 VA con una unidad 647 VA con un electrodo Metrohm multimodo (MME) usado en el modo de gota de mercurio estática (SMDE). El sistema de tres electrodos se completó con un electrodo de platino que actuó como electrodo auxiliar y un electrodo de referencia Ag/AgCl/KCl (3M). Las adiciones de solución patrón fueron realizadas utilizando una bureta Metrohm 665 Dosimat.

Los análisis de datos fueron realizados utilizando los programas PARVUS [57], STATGRAPHICS [58] y MATLAB [59].

Procedimiento experimental

El procedimiento utilizado para llevar a cabo las medidas polarográficas fue el siguiente: la disolución fue introducida en una celda polarográfica y desoxigenada durante 10 min mediante el burbujeo de nitrógeno. Una vez que se hubo eliminado el oxígeno disuelto en la misma, se realizó un barrido de potencial entre -800 mV y -1502 mV, registrándose el polarograma correspondiente a intervalos de -6 mV.

Todas las medidas realizadas se obtuvieron utilizando el modo diferencial de impulsos, con los siguientes parámetros instrumentales: amplitud del impulso, -50 mV; tiempo de vida de la gota, 0.6 s; área del electrodo, 0.40 mm^2 ; velocidad de barrido -10 mV s^{-1} .

II.2.2 Cuestiones preliminares

Cuando en un proceso electródico se producen reacciones químicas acopladas a la reacción de transferencia de carga se pueden presentar esquemas muy distintos, tanto por el número de etapas en que se desarrollan como por el orden en el que éstas se llevan a cabo [60]. Entre las reacciones químicas acopladas se encuentran los procesos de dimerización, en los que una sustancia producto de la reacción electródica se combina consigo misma para dar lugar a un dímero



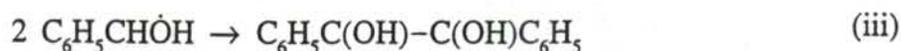
Un ejemplo en el que se presenta este tipo de reacciones acopladas se encuentra en la reducción electroquímica del benzaldehído, compuesto orgánico cuyo complejo mecanismo de reducción sobre electrodo de mercurio ha venido siendo estudiado desde principios de siglo [61-63], junto con otros similares correspondientes a diversos compuestos carbonílicos. Importantes aportaciones a este respecto han sido realizadas principalmente por Zuman y otros [64], Mairanovski [65], y Andrieux, Nadjo y Savéant [66-68]. Diferentes técnicas electroquímicas han sido utilizadas para este mismo fin: polarografía diferencial de impulsos [69-71], polarografía de corriente alterna [72-73], polarografía derivada [74], voltamperometría con diferentes tipos de electrodos [75-77] y sin electrólito soporte [78], voltamperometría de onda cuadrada [79], sonoelectroquímica [80], etc.

En la reducción del benzaldehído sobre un electrodo de mercurio en medio ácido, se observan dos picos monoelectrónicos sucesivos. El primero de ellos se corresponde con la adición reversible de un electrón y un protón, lo cual lleva asociada la formación de un radical libre

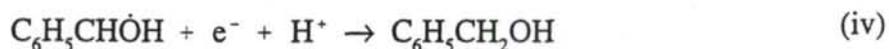


El que la transferencia de protones anteceda o suceda a la transferencia electrónica depende del pH del medio [64]. En aquellos medios más ácidos, se ha evidenciado la existencia de una etapa de preprotonación, mientras que a pH mayores el intercambio de protones sigue a la transferencia electrónica.

Los radicales formados en esta transferencia electrónica rápida dimerizan, originando el correspondiente pinacol, denominado hidrobenczoína, a través de una reacción química irreversible.

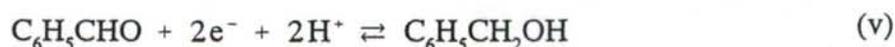


Pero, si el potencial alcanza un valor suficientemente negativo, los radicales se reducen a alcohol bencílico dando lugar a la aparición del segundo pico monoelectrónico, que se ve oscurecido por la descarga de los protones a pH bajos. Este pico aparece a potenciales más negativos, o como mucho iguales, que el primero. Se asume que se trata de una transferencia electrónica lenta [67] que se lleva a cabo del siguiente modo



La influencia de la reacción de dimerización sobre el segundo pico afecta tanto a la forma del mismo como al potencial al que aparece. El hecho de que el segundo pico sea más pequeño que el primero se debe a que se produce una formación parcial del dímero, como han demostrado Andrieux y Savéant [81].

A valores de pH mayores, los dos picos son gradualmente reemplazados por un único pico asociado con una reacción electródica de dos electrones. Este tercer pico se corresponde con la formación del alcohol bencílico, y es el resultado del proceso de tipo ECE siguiente



La interferencia de la reacción de dimerización tiende a separar los potenciales característicos de cada pico, pasando de dos picos monoelectrónicos a un pico bielectrónico a valores de pH suficientemente altos. En el caso del benzaldehído, la variación de la altura de pico depende de la concentración, y esto ha sido interpretado como una interferencia por la simultaneidad de la transferencia electrónica y la dimerización del radical.

Modificando la acidez del medio es posible obtener información sobre las reacciones químicas que acompañan al propio proceso electródico. En los distintos gráficos de las Figuras II.2.1a a II.2.1j se muestra la diferente evolución de los picos polarográficos al variar la concentración de benzaldehído para cada uno de los valores de pH analizados. Las Figuras II.2.2a a II.2.2g permiten observar el efecto de la variación del pH del medio a cada una de las concentraciones de benzaldehído consideradas.

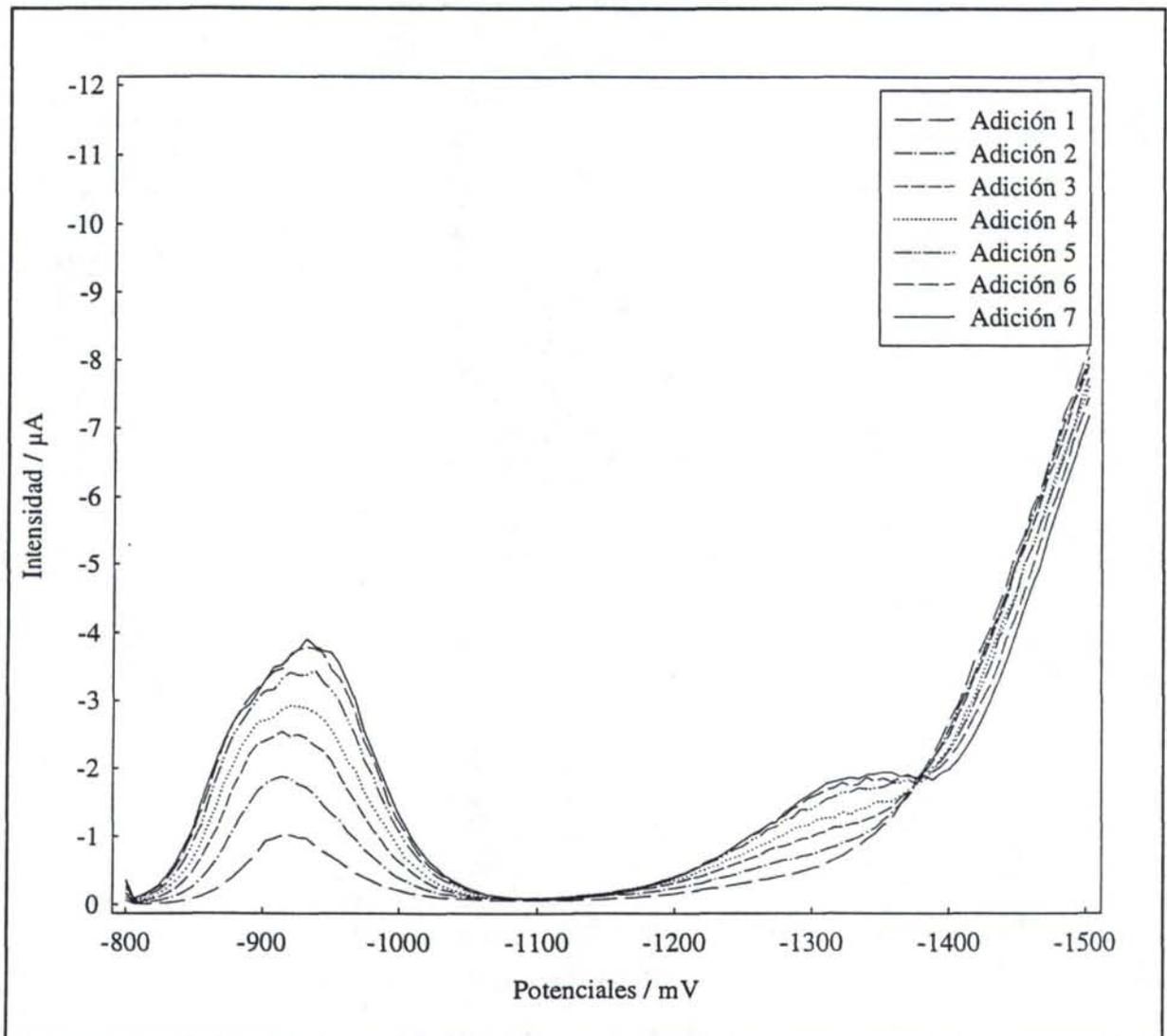


Figura II.2.1a Polarogramas correspondientes a las adiciones de benzaldehído realizadas sobre la disolución tampón de pH 2.2.

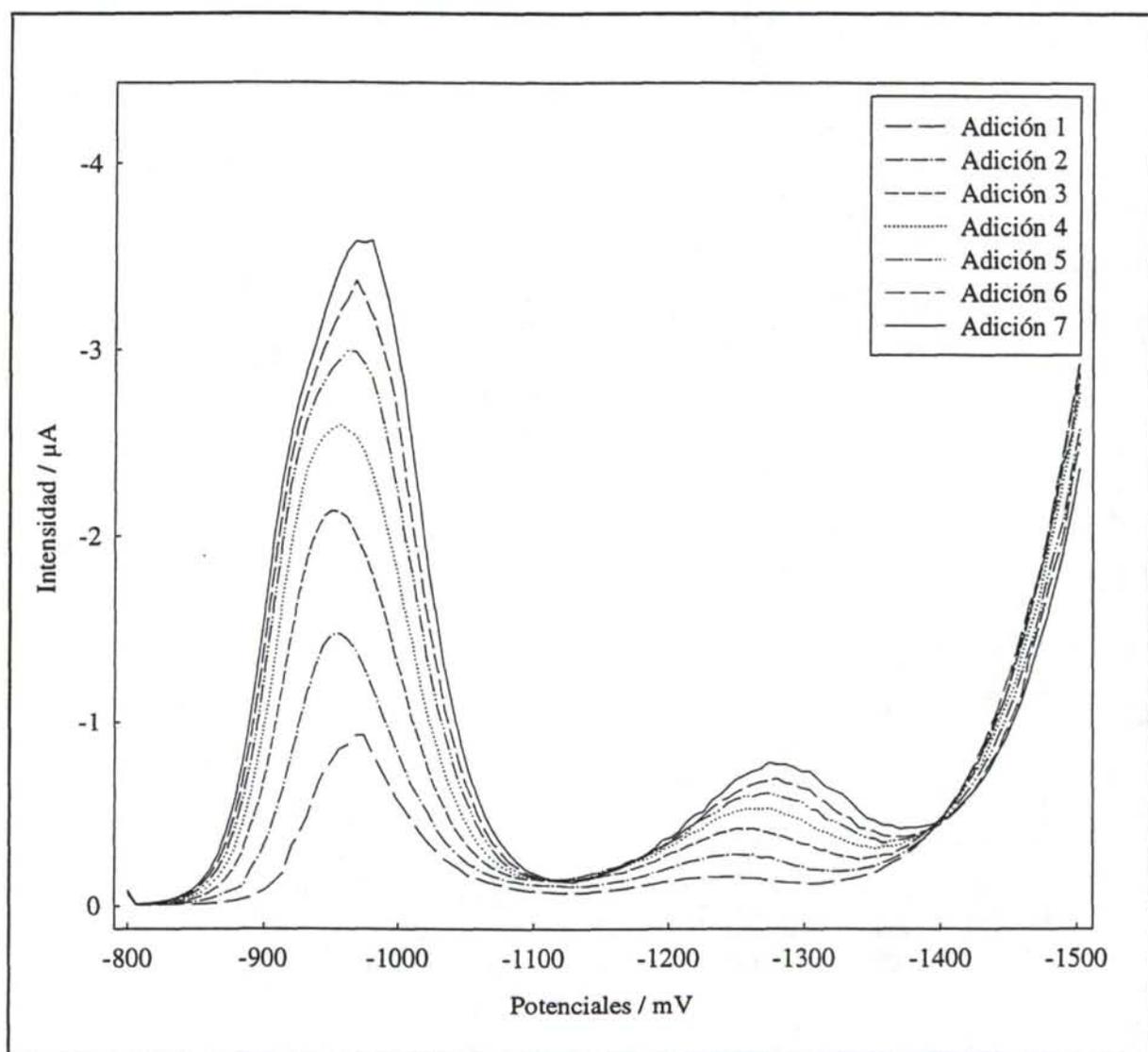


Figura II.2.1b Polarogramas correspondientes a las adiciones de benzaldehído realizadas sobre la disolución tampón de pH 2.8.

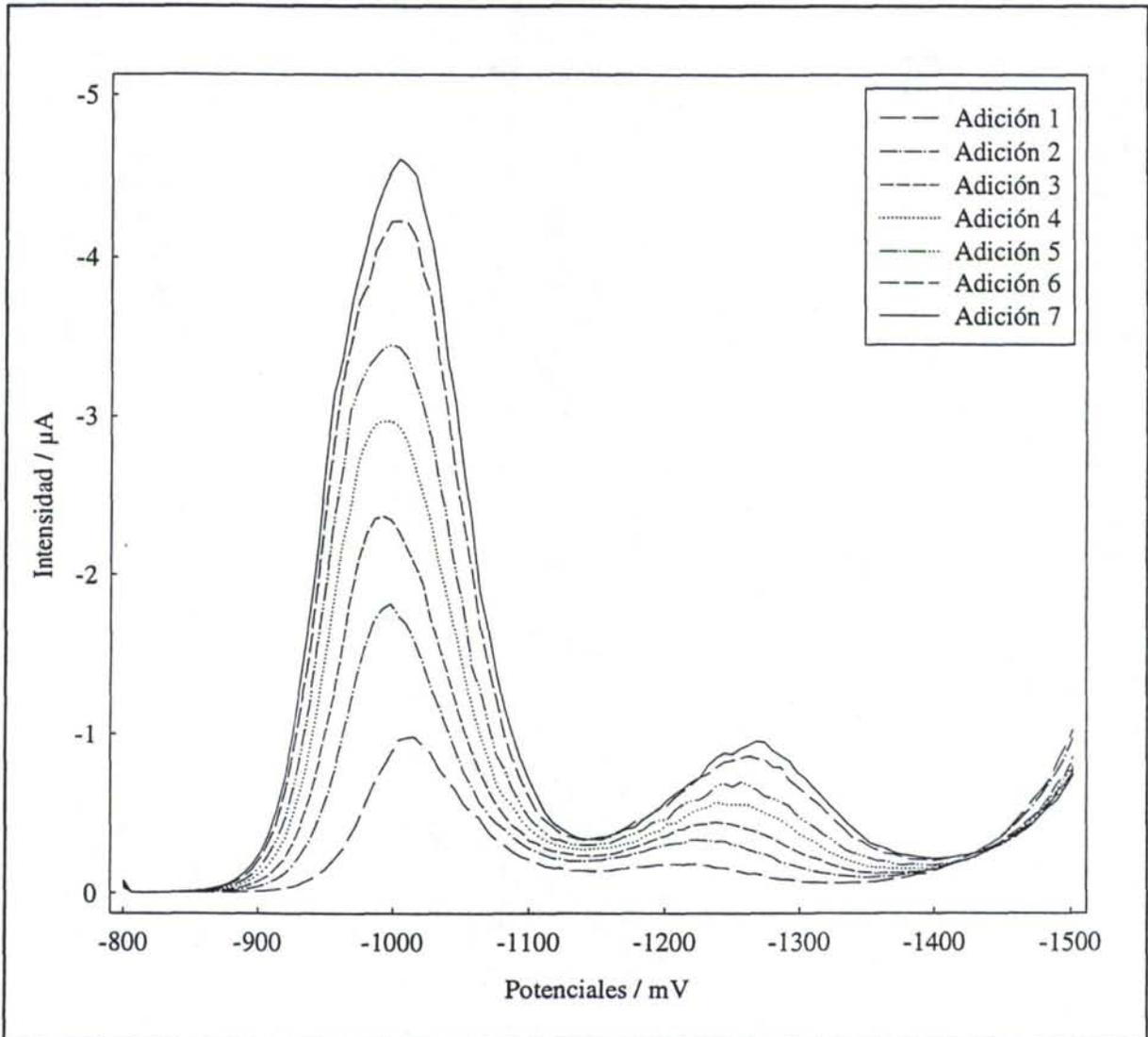


Figura II.2.1c Polarogramas correspondientes a las adiciones de benzaldehído realizadas sobre la disolución tampón de pH 3.4.

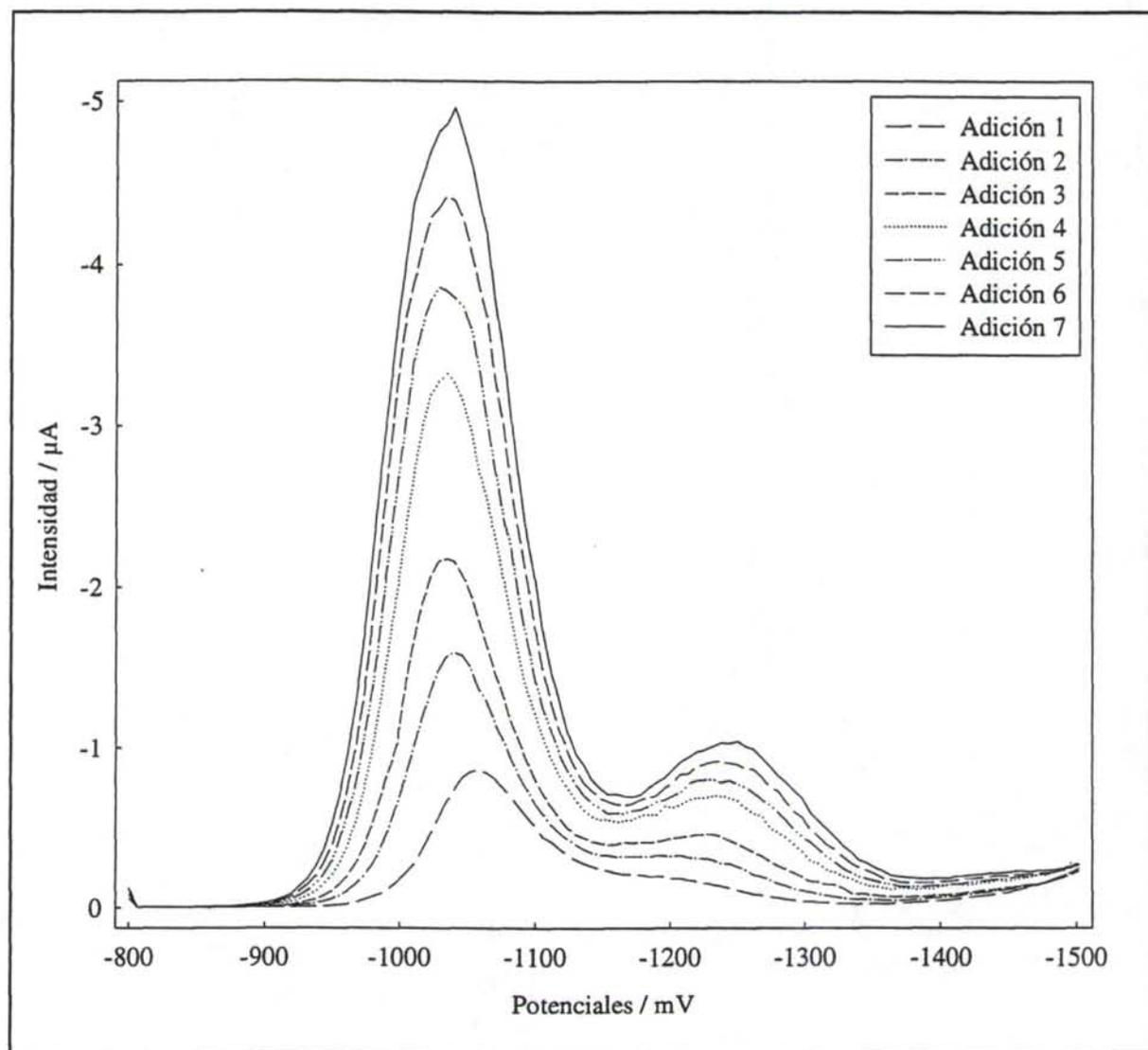


Figura II.2.1d Polarogramas correspondientes a las adiciones de benzaldehído realizadas sobre la disolución tampón de pH 4.0.

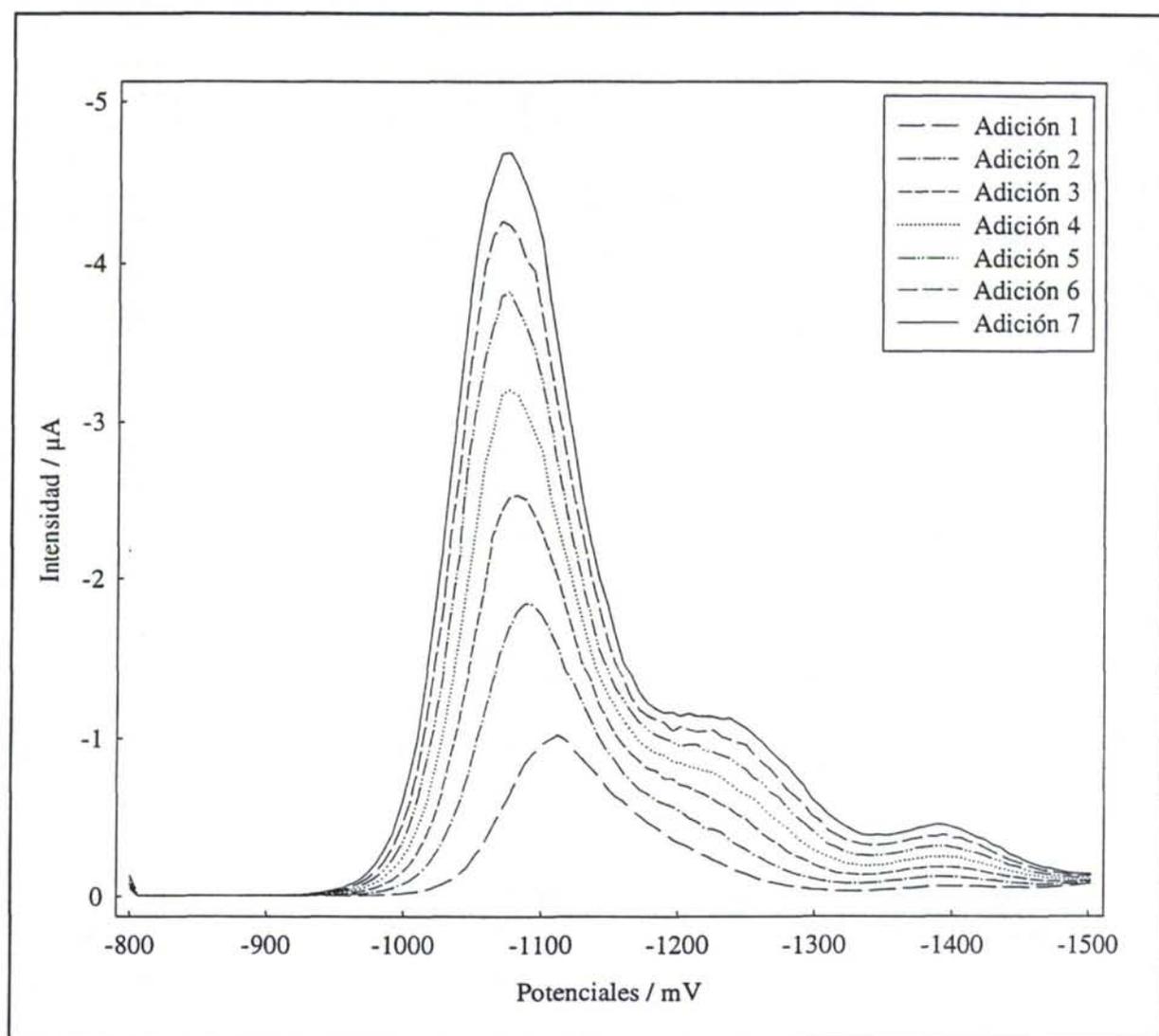


Figura II.2.1e Polarogramas correspondientes a las adiciones de benzaldehído realizadas sobre la disolución tampón de pH 4.6.

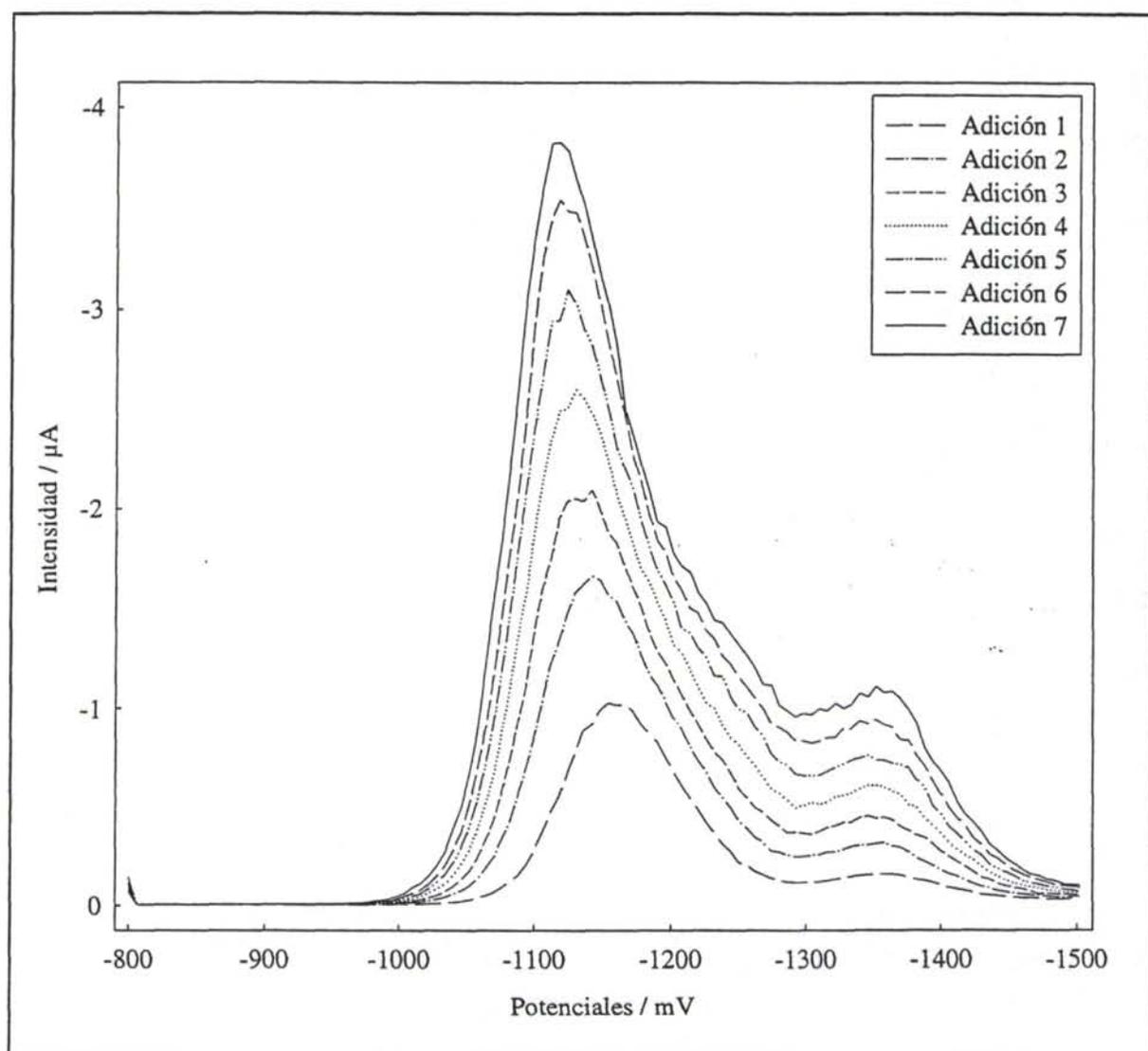


Figura II.2.1f Polarogramas correspondientes a las adiciones de benzaldehído realizadas sobre la disolución tampón de pH 5.2.

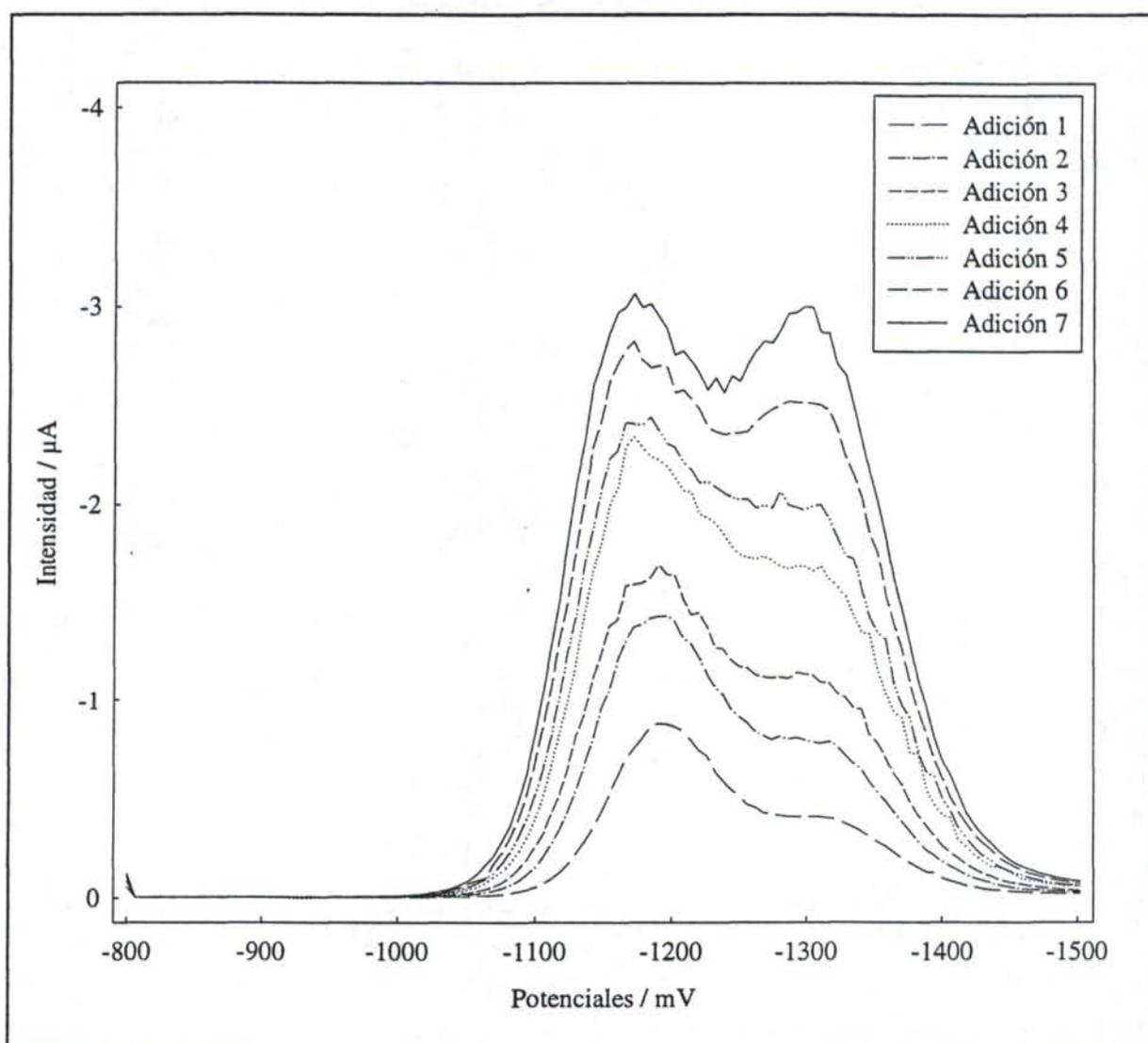


Figura II.2.1g Polarogramas correspondientes a las adiciones de benzaldehído realizadas sobre la disolución tampón de pH 5.8.

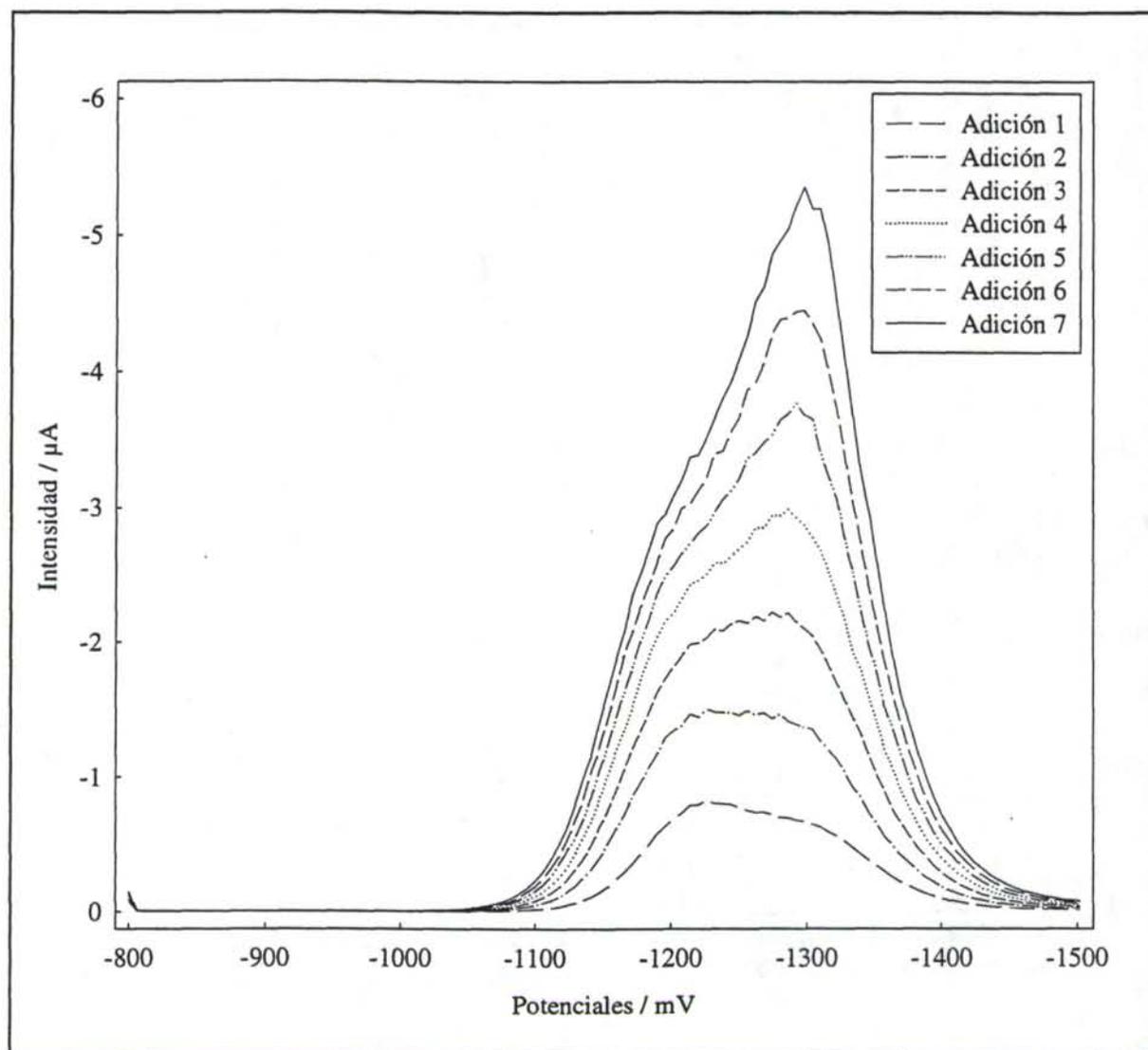


Figura II.2.1h Polarogramas correspondientes a las adiciones de benzaldehído realizadas sobre la disolución tampón de pH 6.4.

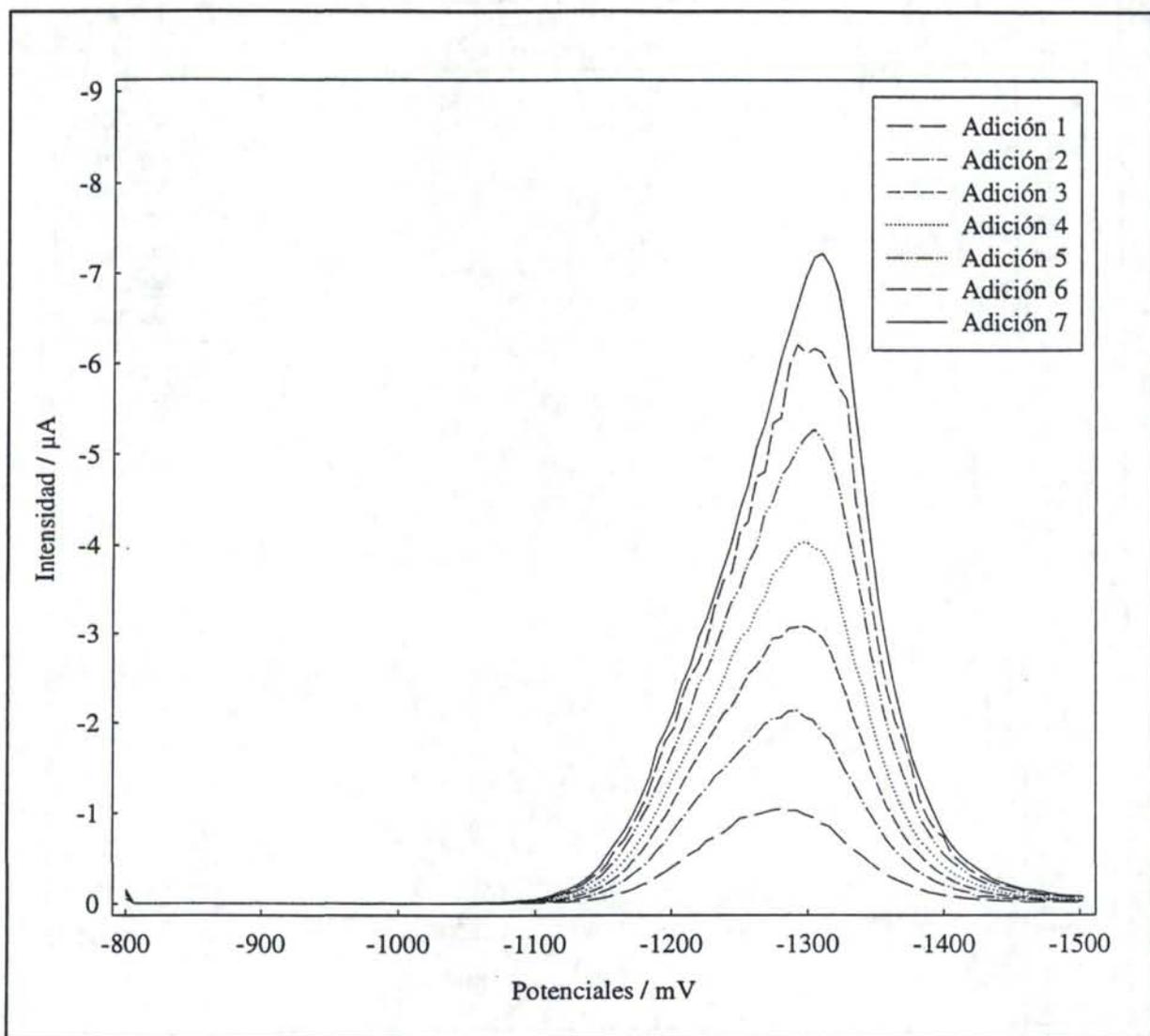


Figura II.2.1i Polarogramas correspondientes a las adiciones de benzaldehído realizadas sobre la disolución tampón de pH 7.0.

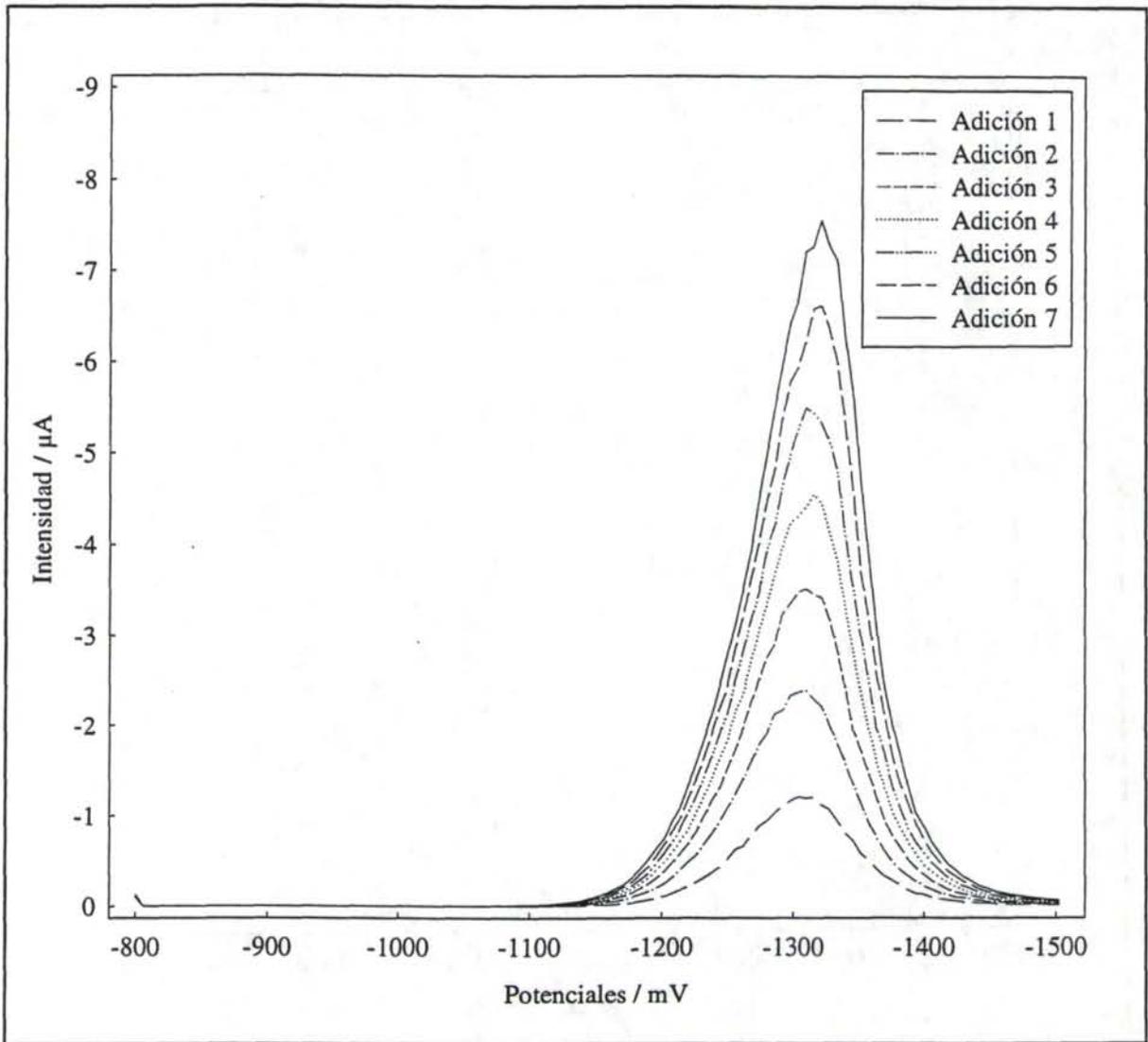


Figura II.2.1j Polarogramas correspondientes a las adiciones de benzaldehído realizadas sobre la disolución tampón de pH 7.6.

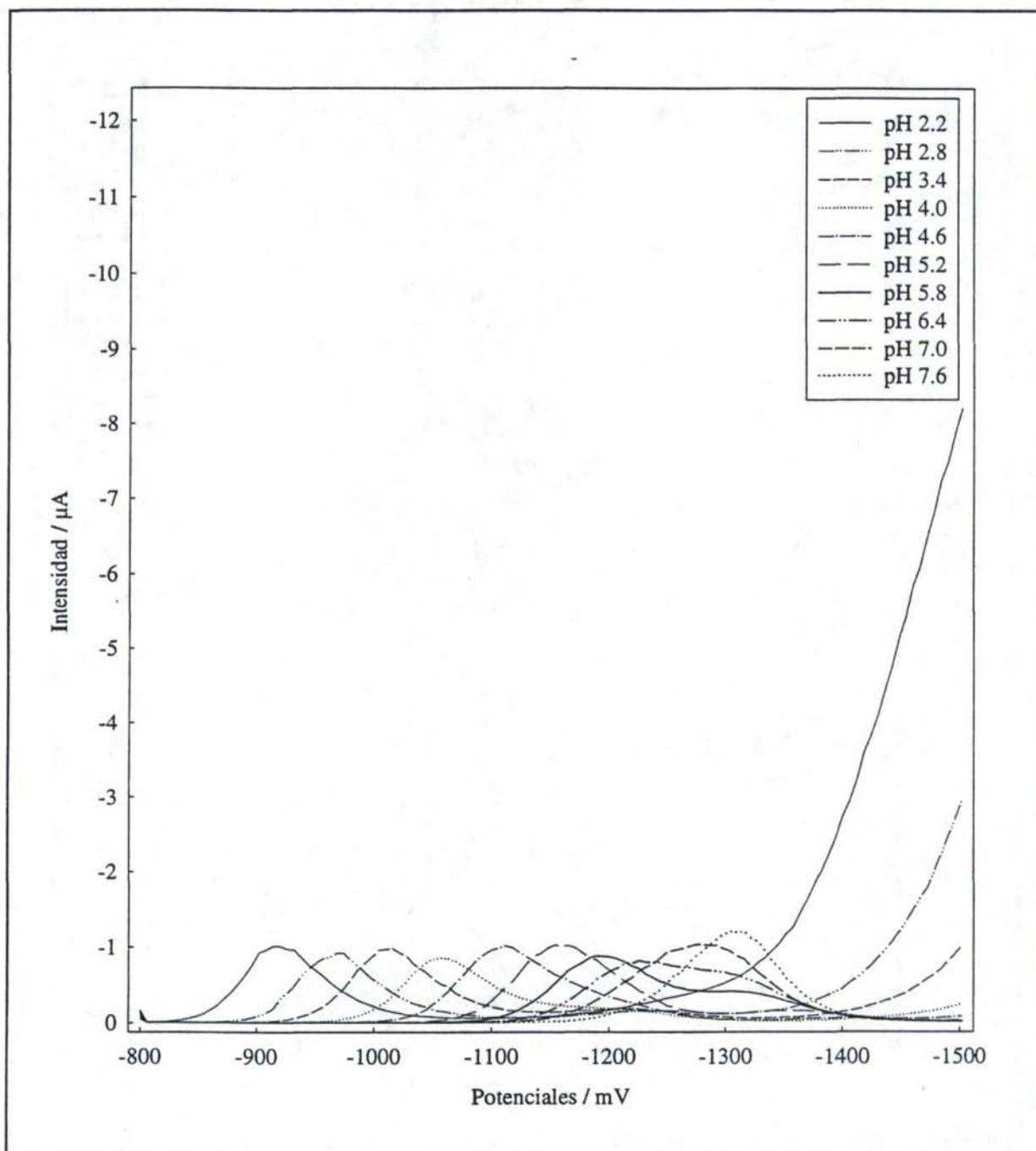


Figura II.2.2a Polarogramas correspondientes a la primera adición de benzaldehído sobre las soluciones tampón con diferentes valores de pH.

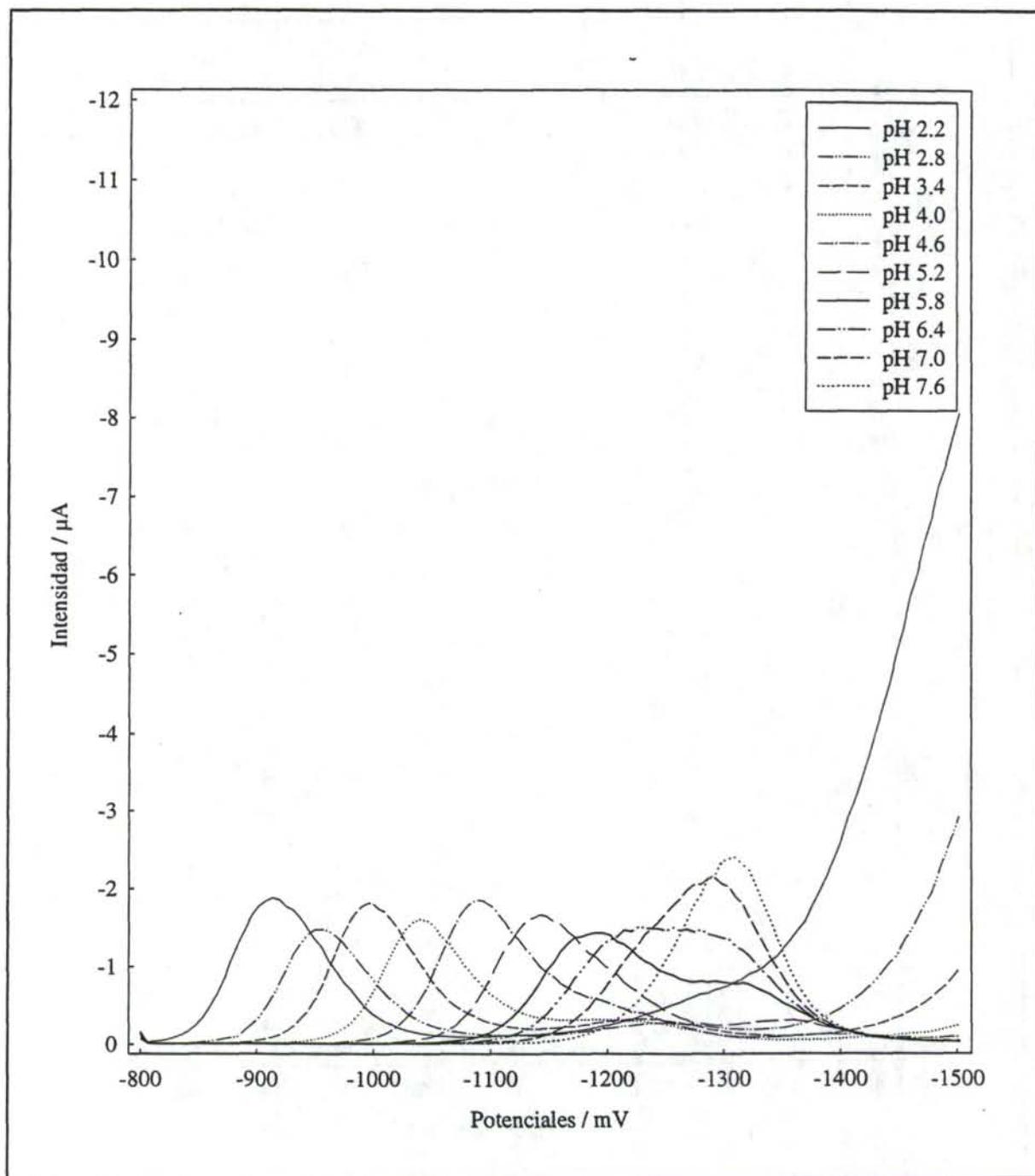


Figura II.2.2b

Polarogramas correspondientes a la segunda adición de benzaldehído sobre las soluciones tampón con diferentes valores de pH.

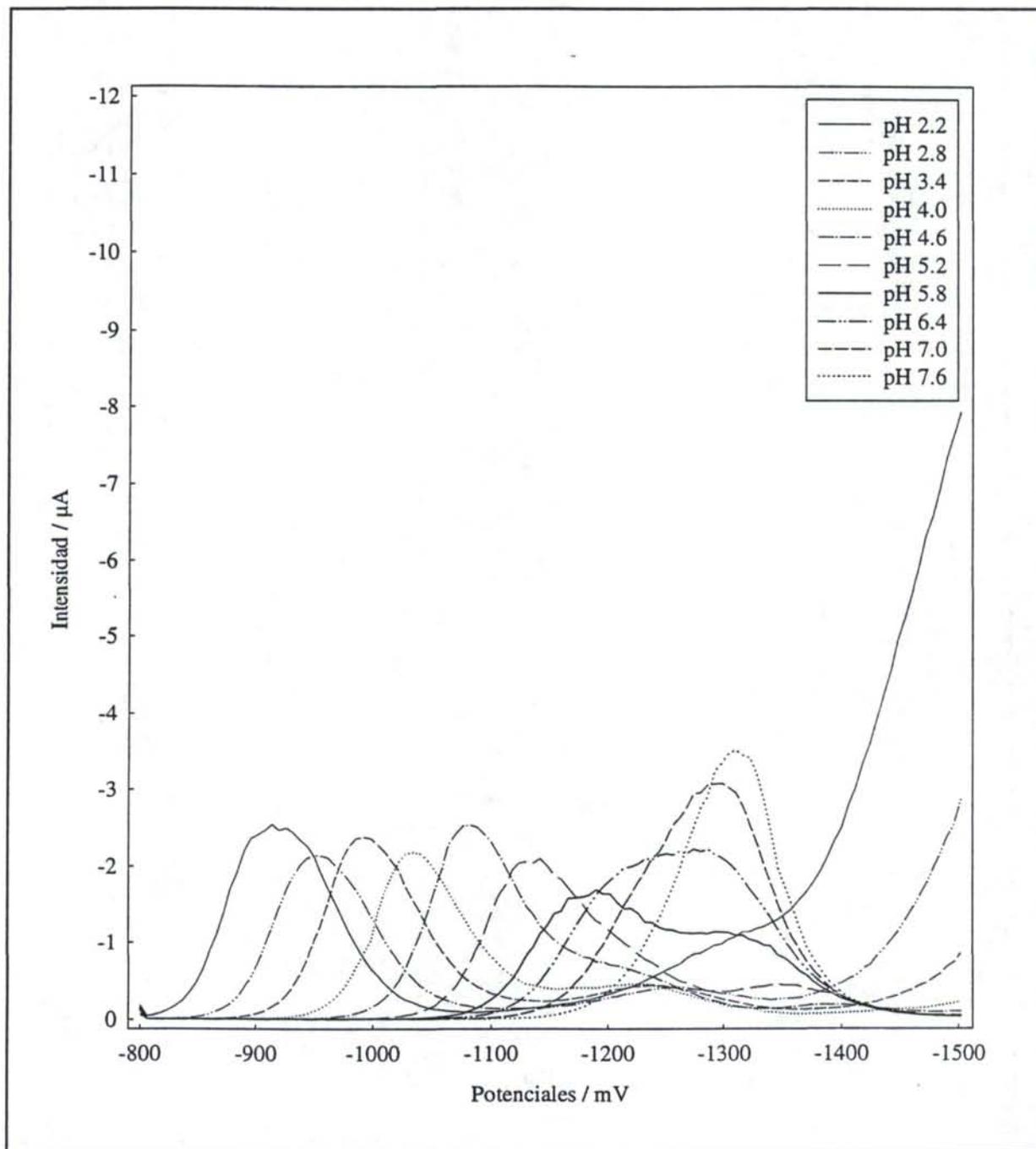


Figura II.2.2c Polarogramas correspondientes a la tercera adición de benzaldehído sobre las soluciones tampón con diferentes valores de pH.

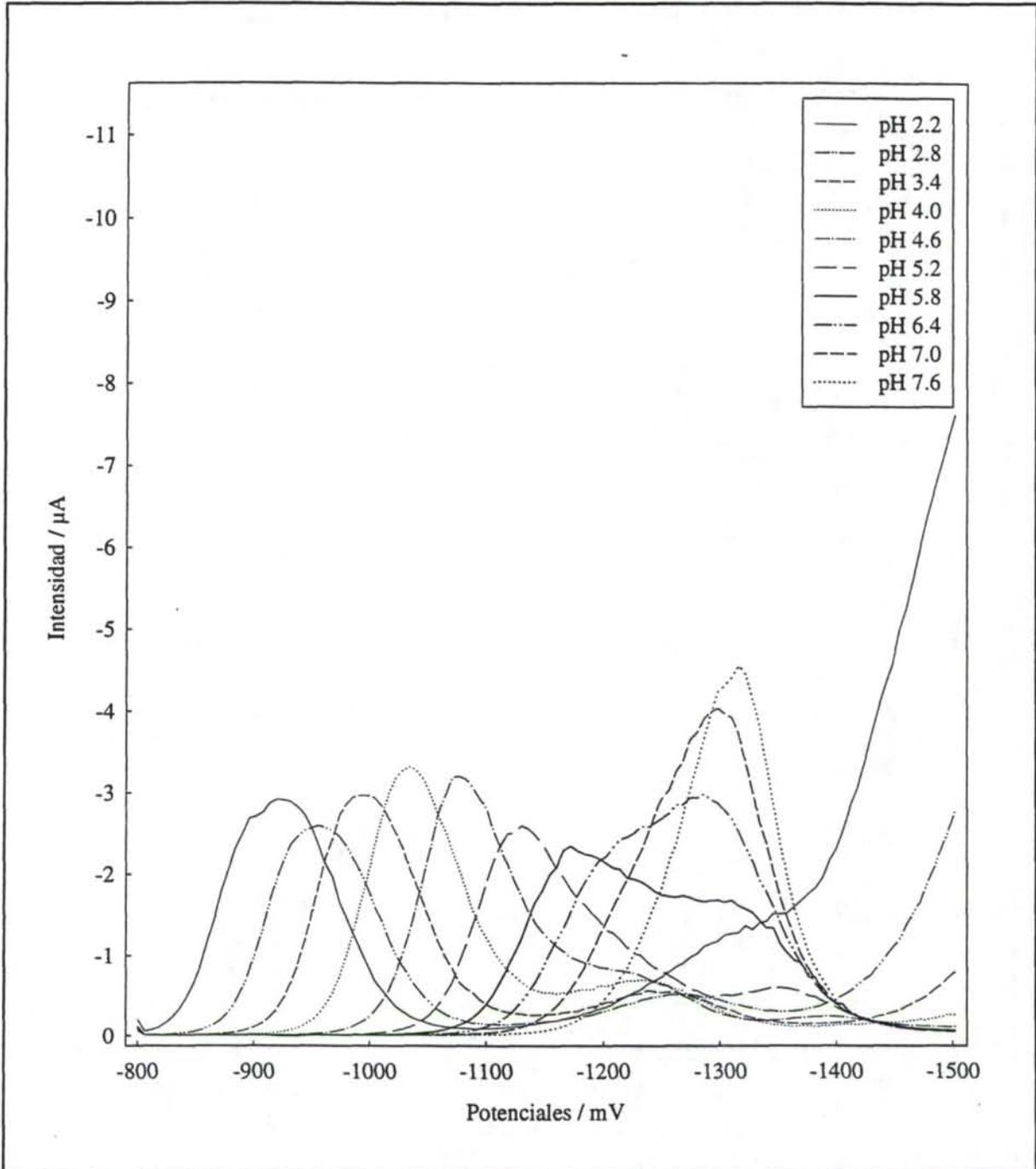


Figura II.2.2d Polarogramas correspondientes a la cuarta adición de benzaldehído sobre las soluciones tampón con diferentes valores de pH.

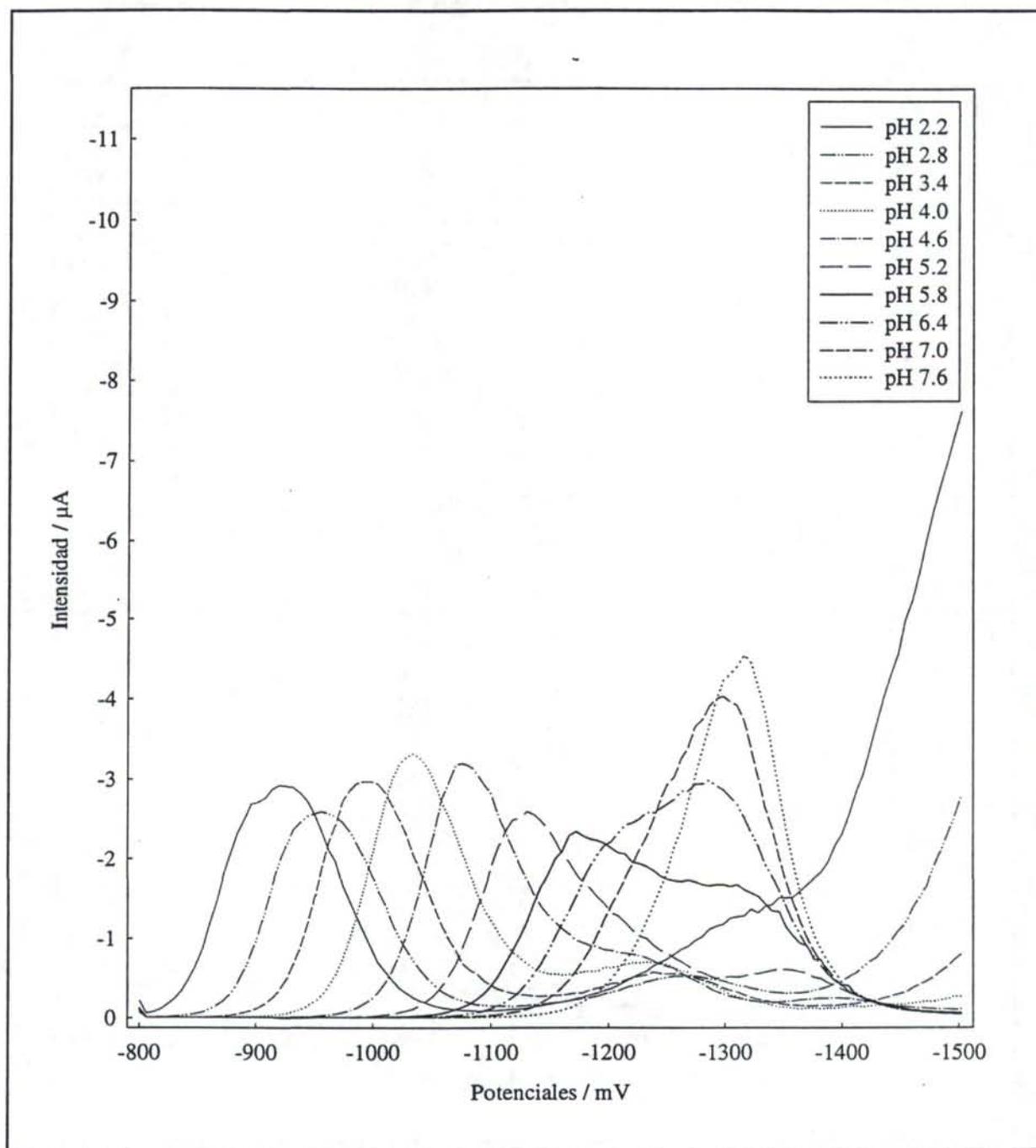


Figura II.2.2e Polarogramas correspondientes a la quinta adición de benzaldehído sobre las soluciones tampón con diferentes valores de pH.

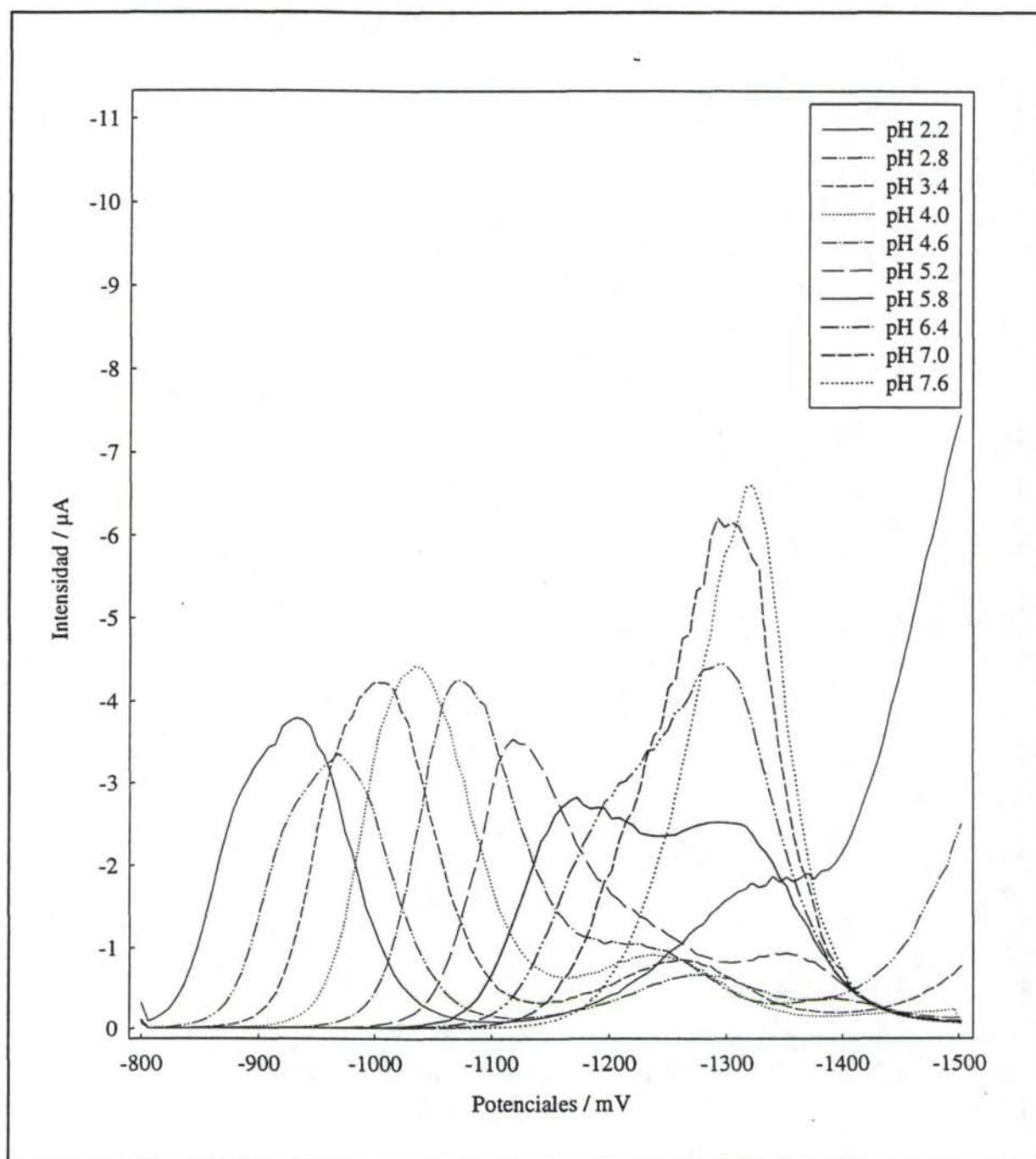


Figura II.2.2f Polarogramas correspondientes a la sexta adición de benzaldehído sobre las soluciones tampón con diferentes valores de pH.

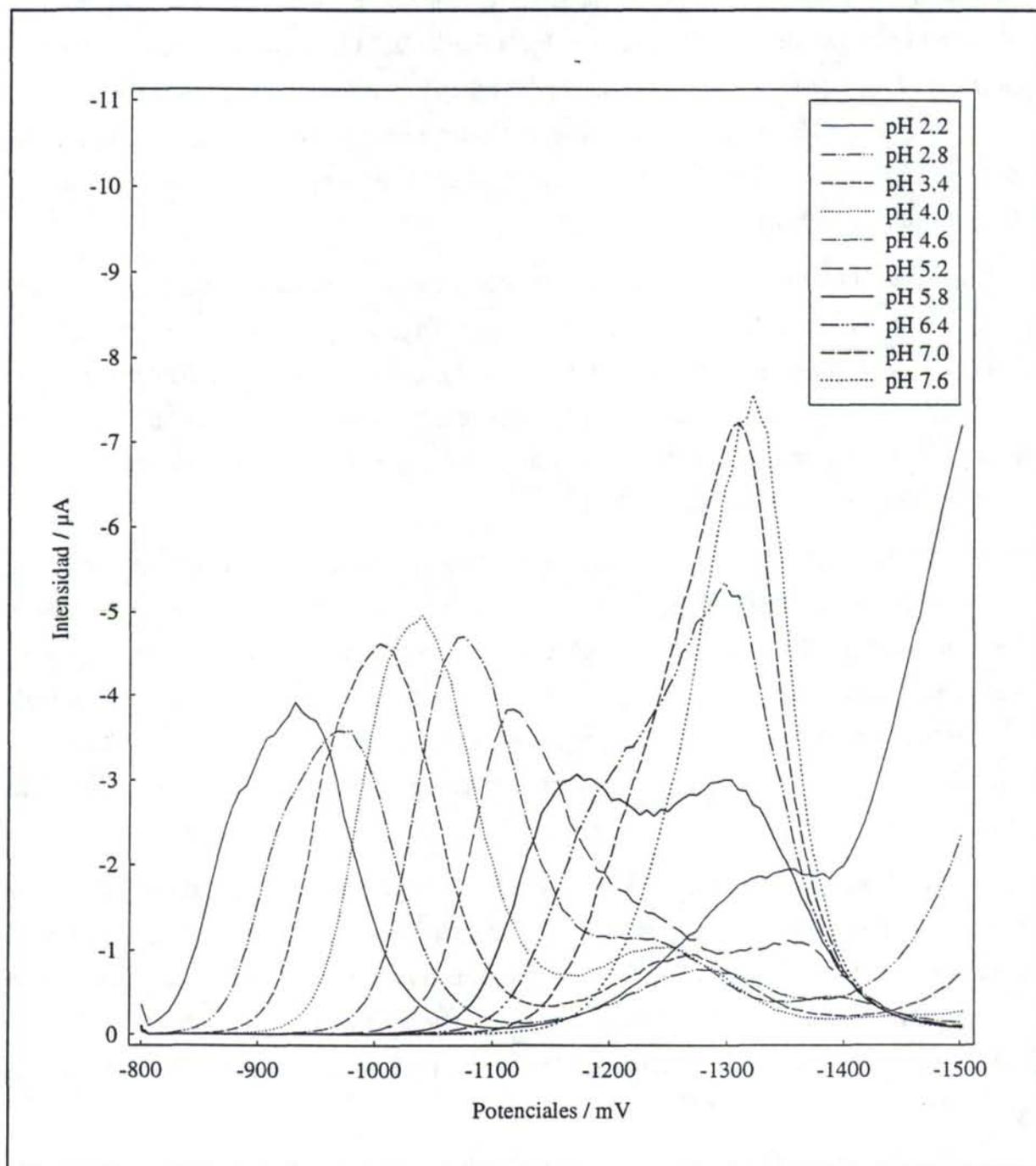


Figura II.2.2g Polarogramas correspondientes a la séptima adición de benzaldehído sobre las soluciones tampón con diferentes valores de pH.

Forma de los polarogramas

Las Figuras II.2.1a a II.2.1j muestran los polarogramas obtenidos con 7 adiciones de benzaldehído realizadas sobre las disoluciones tampón correspondientes a cada uno de los diez valores experimentales de pH analizados. En la Figura II.2.1a, a $\text{pH} = 2.2$, aparecen dos picos polarográficos que se corresponden con los dos procesos monoelectrónicos de las reacciones (ii) y (iv). A este pH se observa que el segundo pico monoelectrónico aparece muy solapado con la barrera de reducción de los protones, sobre todo para las concentraciones de benzaldehído más bajas.

En las Figuras II.2.1a a II.2.1d puede observarse el desplazamiento experimentado por ambos picos, en sentidos opuestos, a lo largo del eje de potenciales, debido tanto a la variación del valor de pH como al incremento de la concentración en las distintas adiciones. También se observa como el pico correspondiente a la reducción del hidrógeno se desplaza hacia potenciales anódicos a medida que aumenta el pH, hasta llegar a desaparecer de la ventana de potenciales estudiados [82].

En la Figura II.2.1e, $\text{pH} = 4.6$, se observa la aparición de un tercer pico que se corresponde con la señal polarográfica de dos electrones debida a la transferencia electrónica de la ecuación (v). También se hace notable el solapamiento de los dos picos correspondientes a los dos procesos monoelectrónicos. Como resultado se obtienen polarogramas en los cuales los tres picos se encuentran muy solapados, tanto más cuanto menor sea la concentración de benzaldehído, confirmando el hecho de que son procesos dependientes de la concentración [67].

La Figura II.2.1f, $\text{pH} = 5.2$, muestra como el pico correspondiente al proceso bieletrónico va adquiriendo importancia en comparación con la magnitud que presenta en los polarogramas de la Figura II.2.1e. Los dos picos monoelectrónicos presentan potenciales de pico cada vez más próximos, originando un solapamiento total de ambos, los cuales por otro lado, también se encuentran muy solapados con el pico correspondiente al proceso de dos electrones.

De la Figura II.2.1g, $\text{pH} = 5.8$, cabe destacar el alto grado de solapamiento de todos los picos polarográficos, que proporciona un polarograma de escasa calidad analítica. Por otro lado, a este pH resulta muy importante la influencia de la variación de la concentración de benzaldehído sobre la señal. A este respecto, se observa que el proceso de dos electrones se ve favorecido a medida que aumenta la concentración de benzaldehído, y pasa de ser poco

importante frente a los dos procesos monoelectrónicos a bajas concentraciones a alcanzar la misma intensidad que estos últimos cuando la concentración es mayor. En la Figura II.2.1h, $\text{pH} = 6.4$, se observa el mismo efecto de la concentración de benzaldehído sobre ese pico polarográfico, pasando el proceso bielectrónico a estar mucho más favorecido que los dos monoelectrónicos a concentraciones altas. En ambas figuras las dos etapas monoelectrónicas aparecen solapadas en un único pico bastante ancho.

Del análisis realizado se deduce que las dos etapas monoelectrónicas se encuentran muy favorecidas a pH menores de 5.8, mientras que la etapa bielectrónica va adquiriendo importancia a pH superiores, a los cuales el aumento de la concentración de benzaldehído favorece mucho esta etapa.

La asimetría del pico en la zona anterior al máximo, en la Figura II.2.1i, implica que el solapamiento sigue existiendo, aunque el proceso que se produce en mayor extensión es el de la etapa de dos electrones. Se observa la misma dependencia de la intensidad de los picos con la concentración de benzaldehído.

La Figura II.2.1j, a $\text{pH} = 7.6$, muestra un único pico que aunque está menos deformado que el anterior también muestra cierta asimetría.

Del estudio de las Figuras II.2.1a a II.2.1j se deduce que el crecimiento de unos picos se realiza a expensas de la disminución de otros, lo cual es indicativo de la existencia de varios equilibrios en los que intervienen compuestos comunes. Los señales polarográficas presentan un comportamiento distinto dependiendo del pH del medio y de la concentración de benzaldehído que esté presente en disolución.

Desplazamiento del potencial de pico

En aquellos procesos electródicos en los que intervienen protones en la reacción electródica el potencial al que aparece el máximo de pico depende de la concentración de éstos en el medio, o lo que es lo mismo, del valor del pH . Las Figuras II.2.2a a II.2.2g muestran la influencia del pH del medio de las distintas soluciones tampón, sobre las que se han realizado adiciones iguales de benzaldehído, en las señales polarográficas registradas. Se observa que la variación del pH del medio provoca un desplazamiento de los potenciales de pico en un sentido u otro dependiendo de la reacción electródica que se considere. Zuman y otros [64] realizan un estudio detallado del desplazamiento del potencial de pico en función del valor del pH para las distintas ondas de reducción de compuestos carbonílicos.

En estas condiciones la determinación del potencial de pico resulta difícil, y en algunos casos imposible, cuando los picos polarográficos se encuentran muy solapados, como por ejemplo ocurre a $\text{pH} = 6.4$.

II.2.3 Análisis univariante

Para cada valor experimental de pH se ha realizado una regresión univariante por mínimos cuadrados, tomando la intensidad máxima de pico como respuesta frente a la concentración de benzaldehído. De los polarogramas correspondientes a cada uno de los valores de pH considerados se ha seleccionado la intensidad de pico del mejor definido, mostrándose en la Tabla II.2.1 los datos experimentales utilizados en el análisis. Los parámetros correspondientes a las regresiones univariantes efectuadas se encuentran en la Tabla II.2.2.

Tabla II.2.1 Intensidades de pico, expresadas en nA, correspondientes a las distintas adiciones de benzaldehído realizadas a diferentes los valores experimentales de pH analizados.

Concentración (mol/L)	pH								
	2.2	2.8	3.4	4.0	4.6	5.2	5.8	7.0	7.6
$4.92 \cdot 10^{-4}$	1017	926	984	862	1024	1029	885	1037	1208
$9.79 \cdot 10^{-4}$	1880	1475	1820	1601	1853	1662	1428	2134	2398
$1.46 \cdot 10^{-3}$	2549	2134	2379	2175	2538	2054	1688	3074	3512
$1.94 \cdot 10^{-3}$	2923	2598	2972	3323	3202	2598	2345	3962	4551
$2.41 \cdot 10^{-3}$	3455	2995	3452	3863	3822	3100	2416	5272	5495
$2.88 \cdot 10^{-3}$	3792	3371	4226	4426	4264	3546	2828	6164	6583
$3.34 \cdot 10^{-3}$	3912	3588	4611	4963	4687	3829	3066	7233	7555

El análisis de los resultados se realiza teniendo en cuenta no sólo el valor de los coeficientes de correlación, sino también la desviación estándar de la regresión, s_{yx} , por estar ésta directamente relacionada con los intervalos de confianza de los parámetros. Se ha analizado también la distribución de los residuos, que en la mayoría de los calibrados no es aleatoria, lo cual indica que no se cumplen las hipótesis sobre la regresión. De acuerdo con

Tabla II.2.2 Resultados del análisis univariante realizado mediante regresiones por mínimos cuadrados a partir de los datos de la Tabla II.2.1.

pH	Pendiente	Término Independiente	Desviación estándar, S_{yx}	Coefficiente de correlación (ρ)	R^2
2.2	$1.0096 \cdot 10^6$	841.23	245.75	0.9463	0.8954
2.8	$9.5067 \cdot 10^5$	606.29	152.17	0.9762	0.9530
3.4	$1.2600 \cdot 10^6$	489.04	104.63	0.9935	0.9870
4.0	$1.4764 \cdot 10^6$	181.04	171.05	0.9874	0.9750
4.6	$1.2855 \cdot 10^6$	574.93	152.15	0.9869	0.9740
5.2	$9.9325 \cdot 10^5$	628.57	86.81	0.9928	0.9857
5.8	$7.5718 \cdot 10^6$	632.49	135.13	0.9705	0.9419
7.0	$2.1668 \cdot 10^6$	-56.59	103.01	0.9979	0.9958
7.6	$2.2087 \cdot 10^6$	209.20	64.82	0.9992	0.9984

estas consideraciones, únicamente las dos últimas regresiones resultarían adecuadas. Incluso en alguna condición experimental ($\text{pH} = 6.4$) se hace imposible la cuantificación del máximo de pico, debido al alto grado de solapamiento de los picos polarográficos, y por tanto el realizar la determinación de la concentración de benzaldehído mediante este procedimiento univariante.

De todo ello se concluye que, cuando se producen reacciones acopladas a la reacción electroquímica, el control de las condiciones experimentales es crítico en la determinación de un analito, ya que normalmente estos sistemas proporcionan señales no específicas para ser analizadas desde un punto de vista univariante ya que conducen a resultados de poca calidad analítica, siendo la alternativa multivariante una posible solución para este problema.

II.2.4 Análisis multivariante

Como se ha expuesto en el epígrafe anterior, una función lineal entre la concentración y la intensidad de pico no muestra suficiente calidad para su uso cuantitativo en la mayoría de las condiciones experimentales consideradas. En consecuencia, no puede calibrarse el sistema de

medida utilizando una relación lineal univariante. Esto obliga al uso de un procedimiento multivariante que tenga en cuenta simultáneamente la presencia de todas las reacciones electródicas y químicas implicadas, así como la competitividad de uno y otro equilibrios.

Por esta razón, se ha decidido emplear una calibración multivariante indirecta en el sentido de hacer adiciones de benzaldehído y utilizar todo el polarograma sin seleccionar sensores específicos. De esta forma en el calibrado se incluyen las relaciones entre todos los equilibrios implicados. Además, cuando se trabaja con datos polarográficos las variables predictoras se encuentran fuertemente correlacionadas entre sí, ya que la intensidad medida a un potencial determinado es parecida a la intensidad medida a otro potencial próximo.

Se ha utilizado una regresión por mínimos cuadrados parciales, PLS, ya que, como se ha señalado anteriormente, realiza la búsqueda de direcciones de máxima variabilidad en las variables predictoras evitando aquellas que no estén correlacionadas con la respuesta. Esta característica inherente a la regresión PLS ofrece la posibilidad de construir modelos de regresión con gran capacidad de predicción.

Construcción de los modelos PLS

La regresión PLS se ha realizado sobre un conjunto de datos formado por siete adiciones de 100 μL de benzaldehído $9.89 \cdot 10^{-2} \text{ M}$ (20% etanol) sobre una muestra de 20 mL de solución tampón McIlvaine. Cada polarograma se ha digitalizado tomando intensidades a 118 potenciales igualmente espaciados, desde -800 mV hasta -1502 mV, y a continuación se le ha restado el polarograma correspondiente al blanco experimental. Este procedimiento se ha aplicado a 10 soluciones de tampón McIlvaine correspondientes a 10 valores de pH diferentes y equidistantes entre sí, de 2.2 a 7.6. Por lo tanto, la matriz de datos original para cada valor de pH consta de 118 variables predictoras (intensidades medidas a 118 potenciales) y de 7 objetos (adiciones de benzaldehído). Los datos utilizados en el análisis multivariante llevado a cabo se muestran, en su conjunto, en las Figura II.2.3a a II.2.3e.

La Tabla II.2.3 es una representación esquemática de las matrices de datos necesarias para llevar a cabo el calibrado multivariante propuesto; en ella, las variables predictoras, X, son las intensidades medidas a cada potencial una vez que ha sido restada la señal correspondiente al blanco experimental, mientras que las distintas concentraciones de benzaldehído constituyen la variable respuesta, Y.

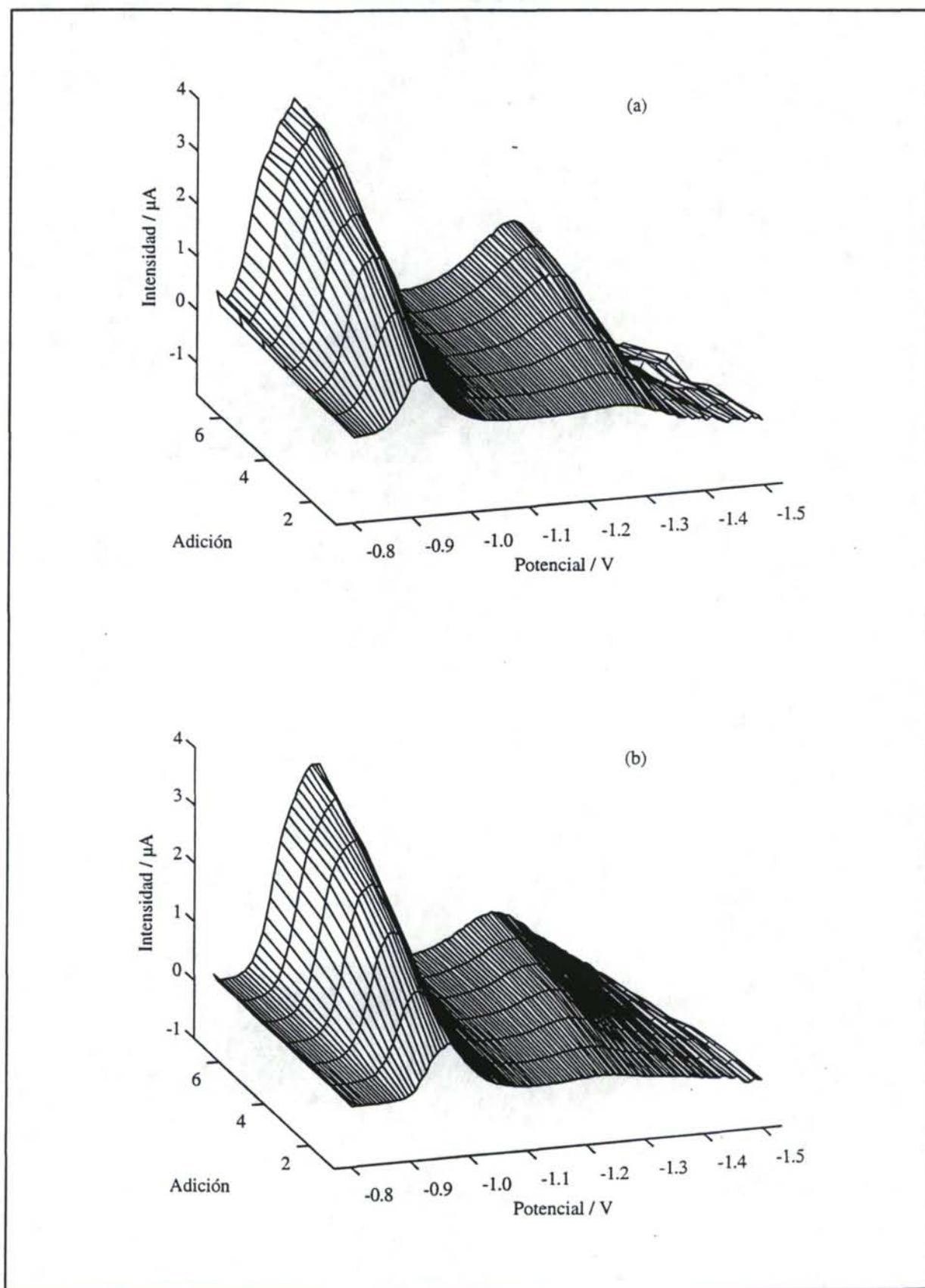


Figura II.2.3a Datos correspondientes a los modelos PLS construidos para los valores de pH 2.2 (a) y 2.8 (b).

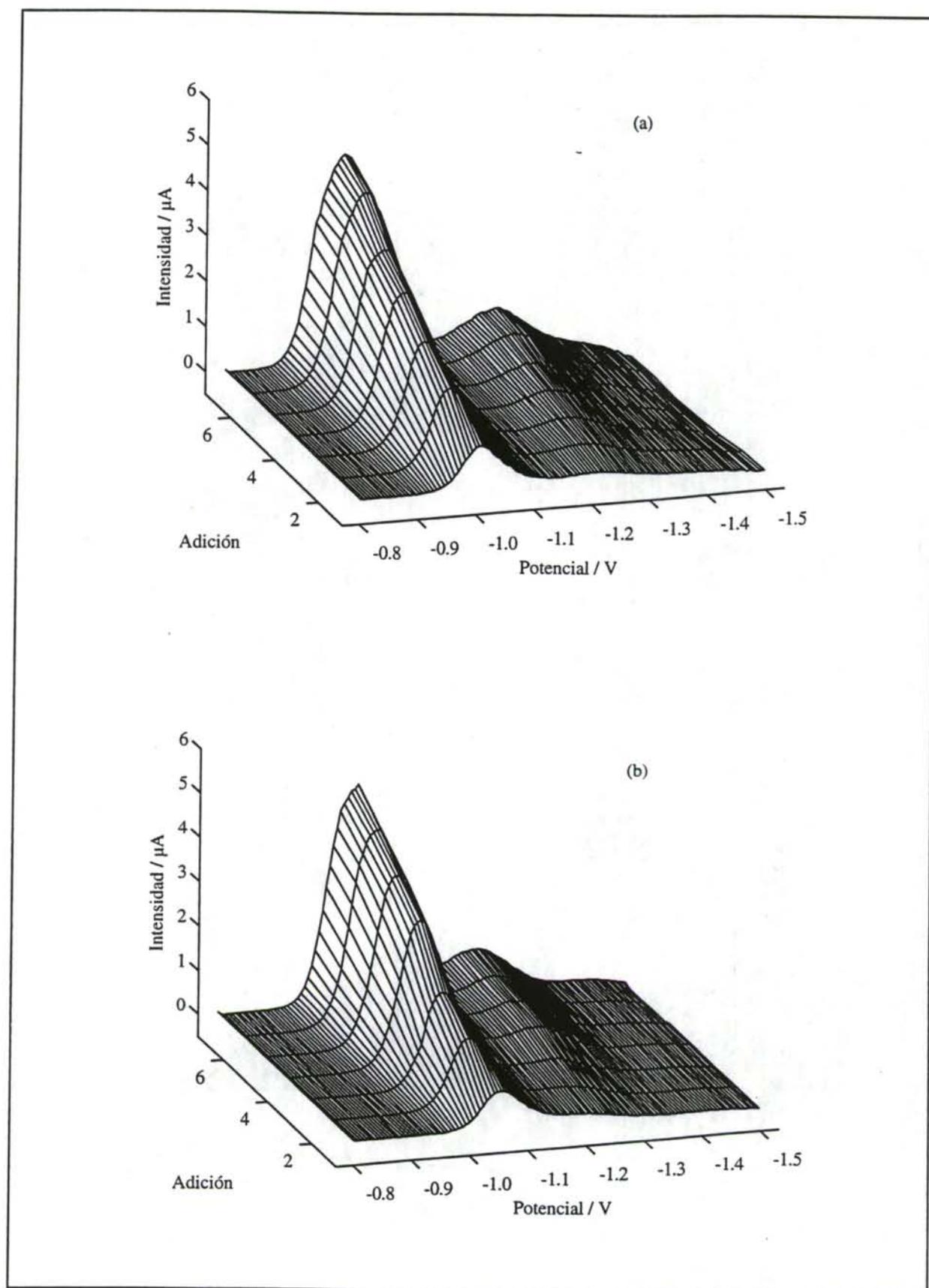


Figura II.2.3b Datos correspondientes a los modelos PLS construidos para los valores de pH 3.4 (a) y 4.0 (b).

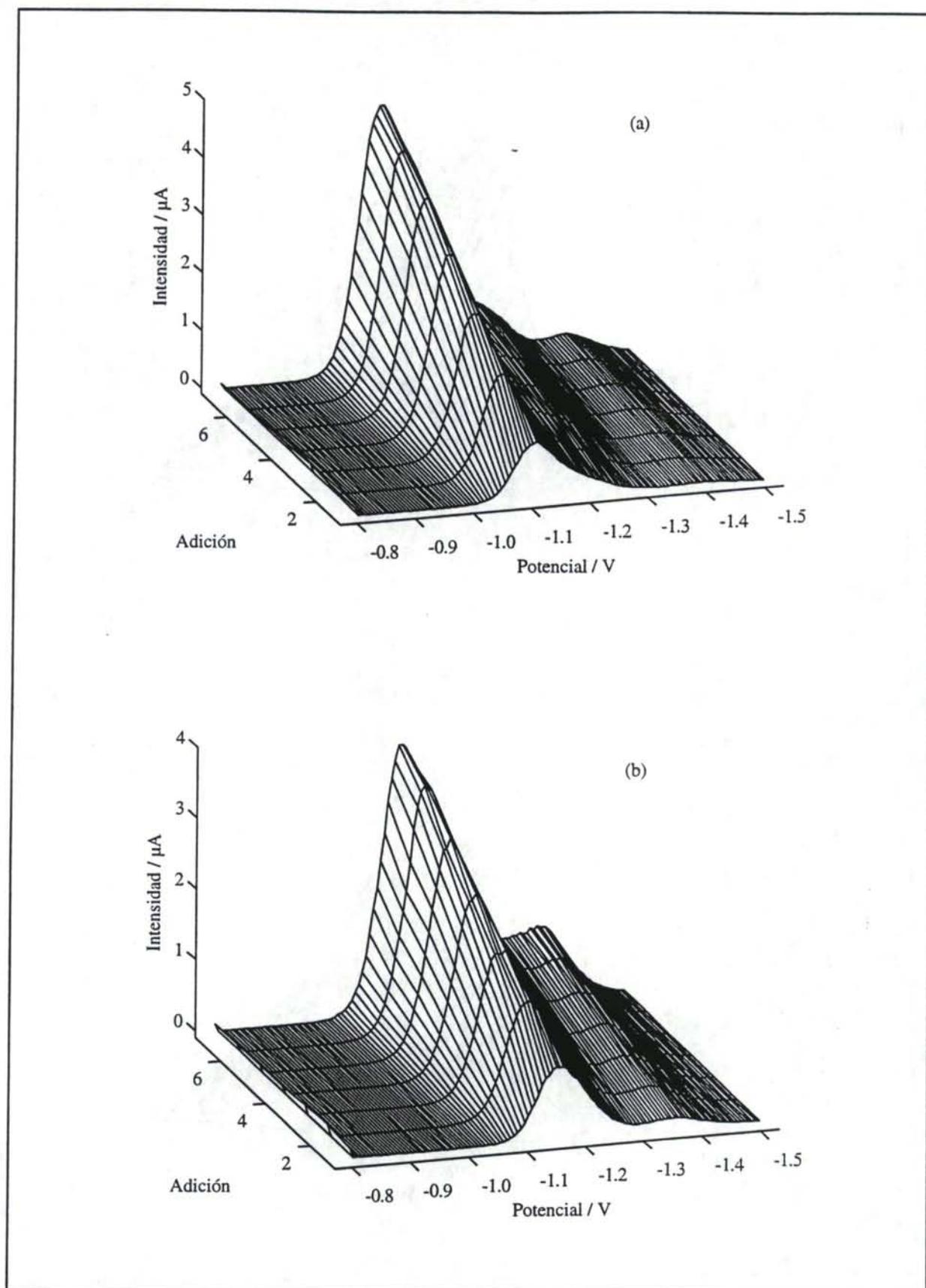


Figura II.2.3c Datos correspondientes a los modelos PLS construidos para los valores de pH 4.6 (a) y 5.2 (b).

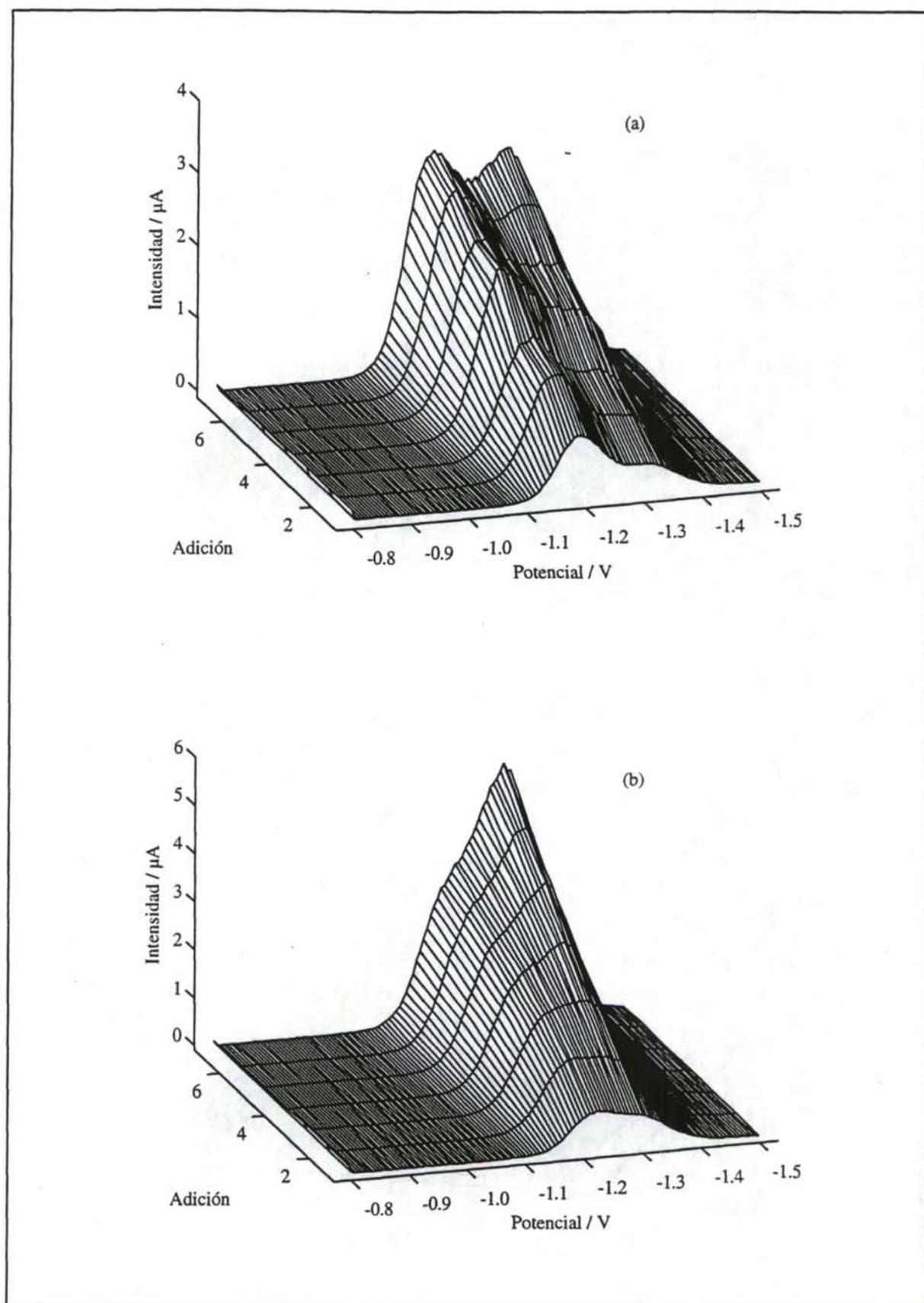


Figura II.2.3d Datos correspondientes a los modelos PLS construidos para los valores de pH 5.8 (a) y 6.4 (b).

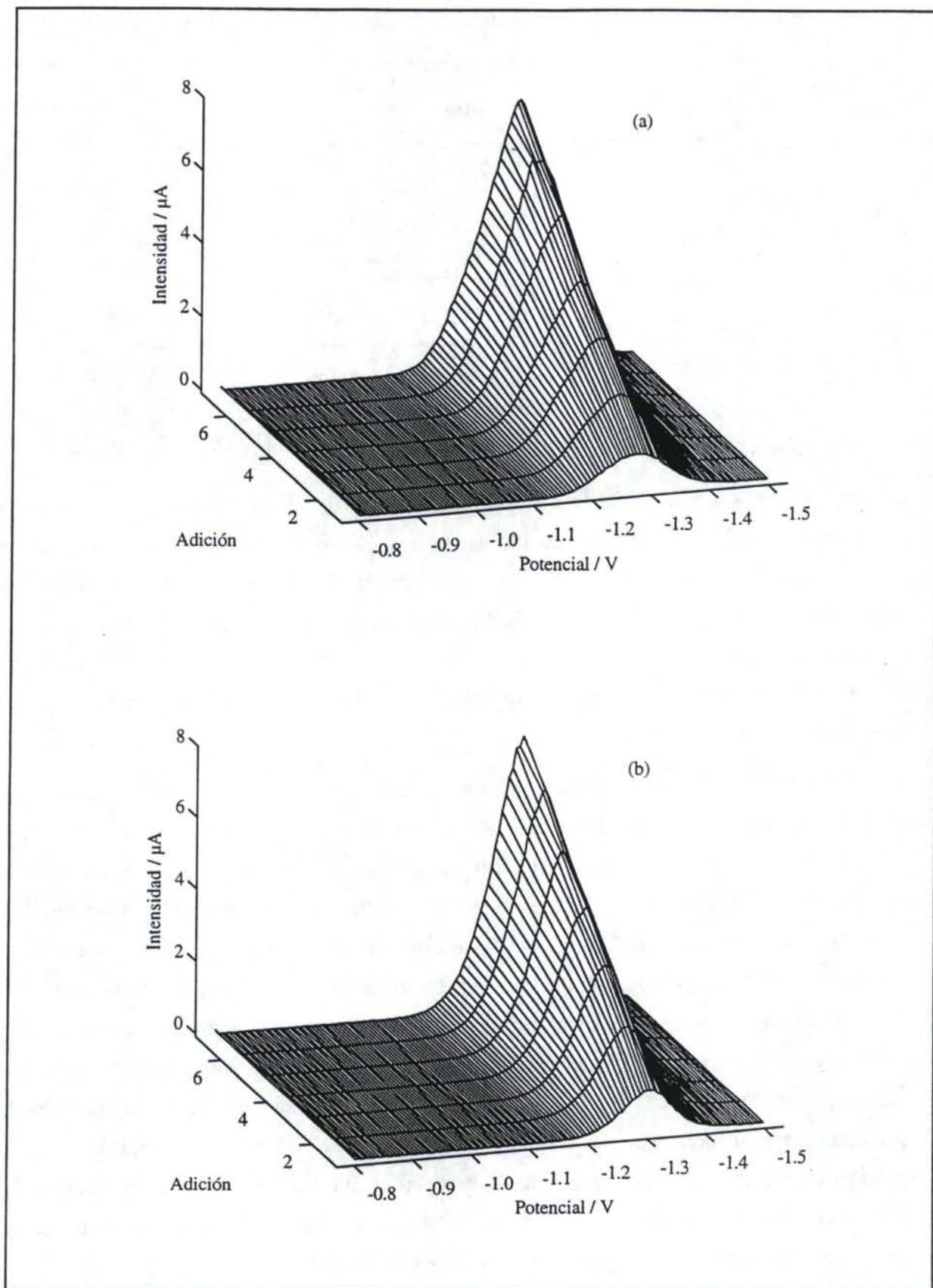


Figura II.2.3e Datos correspondientes a los modelos PLS construidos para los valores de pH 7.0 (a) y 7.6 (b).

Tabla II.2.3 Esquema de las matrices de datos utilizadas en la calibración multivariante.

Bloque X				Bloque Y
$E_{.800}$	$E_{.806}$...	$E_{.1502}$	
$i_{1,1}$	$i_{1,2}$...	$i_{1,118}$	c_1
$i_{2,1}$	$i_{2,2}$...	$i_{2,118}$	c_2
.
.
$i_{7,1}$	$i_{7,2}$...	$i_{7,118}$	c_7

Evaluación de los modelos PLS

El punto crítico para establecer el modelo de calibrado PLS es determinar el número de componentes (variables latentes) que son necesarias. En su determinación se ha de tener en cuenta el porcentaje de varianza explicada tanto en el bloque de las variables predictoras, como en el bloque de la respuesta. De hecho, podría ocurrir que al incrementar el número de componentes se estuviera modelando exclusivamente la variabilidad interna del bloque de las variables predictoras (intensidades), no relacionada con la variabilidad en la respuesta (concentración de benzaldehído).

Además se ha de prestar atención a la capacidad de predicción del modelo construido, es decir, hay que valorar la estabilidad del mismo, lo cual se lleva a cabo a través de la varianza denominada "en predicción". En su cálculo se han utilizado siete grupos de cancelación con validación cruzada completa. Con cada uno de estos conjuntos de aprendizaje se ha construido un modelo PLS con sucesivas componentes y se ha evaluado la varianza no explicada, utilizando para ello las respuestas en los objetos que no han intervenido en la construcción del modelo. Es evidente que se trata de una varianza explicada en predicción, y por lo tanto, cuanto más parecida sea ésta a la varianza explicada por el modelo construido con todos los objetos, más estable será éste. Mientras que por construcción, la varianza explicada crece al incrementar el número de componentes utilizadas en el modelo PLS, la varianza en predicción puede disminuir al aumentar el número de componentes del modelo. Esto sería indicativo de que esas componentes dependen demasiado de muestras u objetos concretos y no han de formar parte del modelo de calibrado.

Por tanto, uno de los criterios utilizados en la construcción de modelos PLS viene dado por la selección de aquellas variables latentes que incrementen tanto los porcentajes de varianza explicada como los correspondientes a la varianza en predicción.

La Tabla II.2.4 resume la evolución del porcentaje de varianza explicada en función del número de componentes del modelo PLS construido para cada uno de los valores experimentales de pH considerados. En negrilla aparecen señalados los valores de varianza explicada correspondientes al número de variables latentes seleccionado para cada modelo aplicando los criterios anteriormente enunciados. En todos los modelos construidos el porcentaje de varianza explicada en el bloque Y es superior al 99.2 %, lo cual indica que estos modelos multivariantes se ajustan muy bien a los datos utilizados en su construcción.

En la mayoría de los casos la inclusión de una nueva componente sólo incrementa apreciablemente la varianza explicada en el bloque de las variables predictoras, indicando que esta última fuente de variabilidad en la intensidades del polarograma no está relacionada con la concentración de benzaldehído añadida.

En la condición de $\text{pH} = 7.0$ no se ha aplicado el criterio estricto de disminución de la varianza en predicción sino que la no aleatoriedad de loadings hace que se tomen dos variables latentes. Esto es coherente con los resultados obtenidos a pH próximos, cuyos loadings presentan estructuras muy similares. Los valores de varianza en predicción son muy parecidos a los correspondientes a la varianza explicada, lo cual pone de manifiesto la estabilidad de los modelos construidos y su buena capacidad de predicción.

Interpretabilidad de las variables latentes

En este tipo de problemas, tiene gran interés un análisis detallado del significado químico de las variables latentes construidas por el modelo de calibrado PLS, puesto que han de estar relacionadas con los equilibrios que intervienen en el problema.

En la interpretación química de las variables latentes se utilizan tanto los valores correspondientes a los loadings como los de los scores. En las Figuras II.2.4a a II.2.4j se muestran los loadings para las cuatro primeras variables latentes de los modelos PLS construidos en cada condición experimental de pH. Mientras que en las Figuras II.2.5a a II.2.5j se representan los scores correspondientes, que vienen dados por las puntuaciones o coordenadas de los objetos en los distintos planos formados por esas cuatro variables latentes.

Tabla II.2.4 Varianza explicada y varianza en predicción (%) en función del número de variables latentes. En negrilla y cursiva se señala el número de variables latentes seleccionadas para cada modelo.

pH	Bloque	Varianza	Variables latentes			
			1	2	3	4
2.2	Y	Explicada	94.72	99.54	99.81	100.00
		En predicción	92.82	98.20	99.31	99.17
	X	Explicada	90.56	94.71	97.72	97.62
2.8	Y	Explicada	97.01	99.58	99.75	99.90
		En predicción	95.90	99.21	98.52	98.98
	X	Explicada	92.34	95.47	96.12	98.72
3.4	Y	Explicada	99.34	99.69	99.66	99.99
		En predicción	99.07	99.25	99.12	99.30
		Explicada	92.93	95.00	98.19	98.51
4.0	Y	Explicada	99.20	99.19	99.10	99.72
		En predicción	99.09	98.59	97.07	97.85
	X	Explicada	92.68	95.49	98.23	98.36
4.6	Y	Explicada	98.36	99.58	99.99	100.00
		En predicción	97.83	99.27	99.95	99.95
	X	Explicada	90.07	96.50	98.55	98.85
5.2	Y	Explicada	96.90	99.90	99.97	99.98
		En predicción	96.14	99.79	99.94	99.86
	X	Explicada	89.14	97.30	98.48	98.55
5.8	Y	Explicada	98.61	99.56	99.89	99.94
		En predicción	98.33	99.48	99.24	99.50
	X	Explicada	92.25	98.03	98.06	98.47
6.4	Y	Explicada	99.78	100.00	100.00	100.00
		En predicción	99.70	100.00	99.99	99.99
	X	Explicada	93.49	98.63	98.73	99.08
7.0	Y	Explicada	99.84	99.85	99.88	100.00
		En predicción	99.75	99.66	99.38	99.23
	X	Explicada	94.11	98.24	98.97	99.02
7.6	Y	Explicada	99.69	99.95	99.99	99.99
		En predicción	99.61	99.94	99.92	99.91
	X	Explicada	94.33	98.58	98.68	99.04

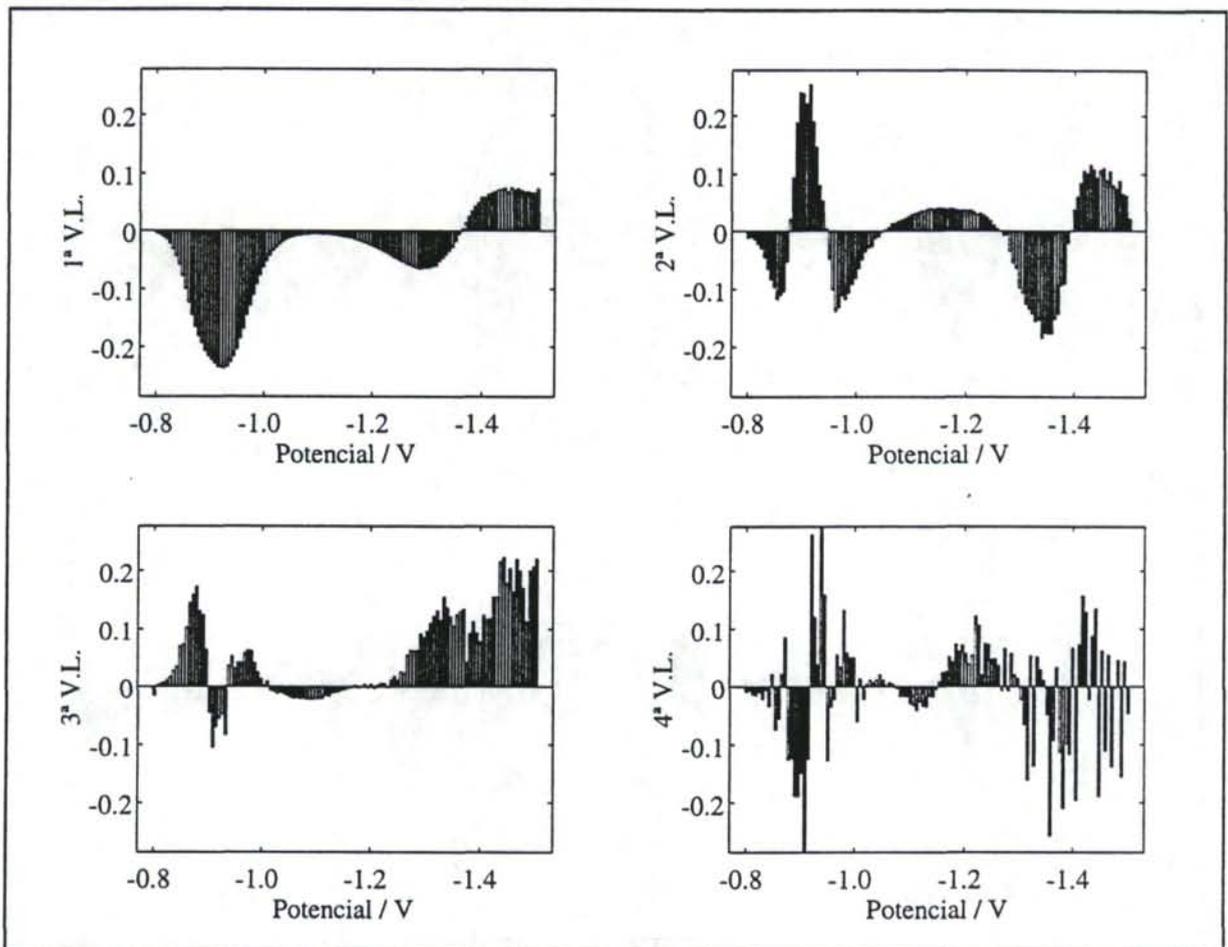


Figura II.2.4a Loadings correspondientes al modelo PLS construido para pH = 2.2.

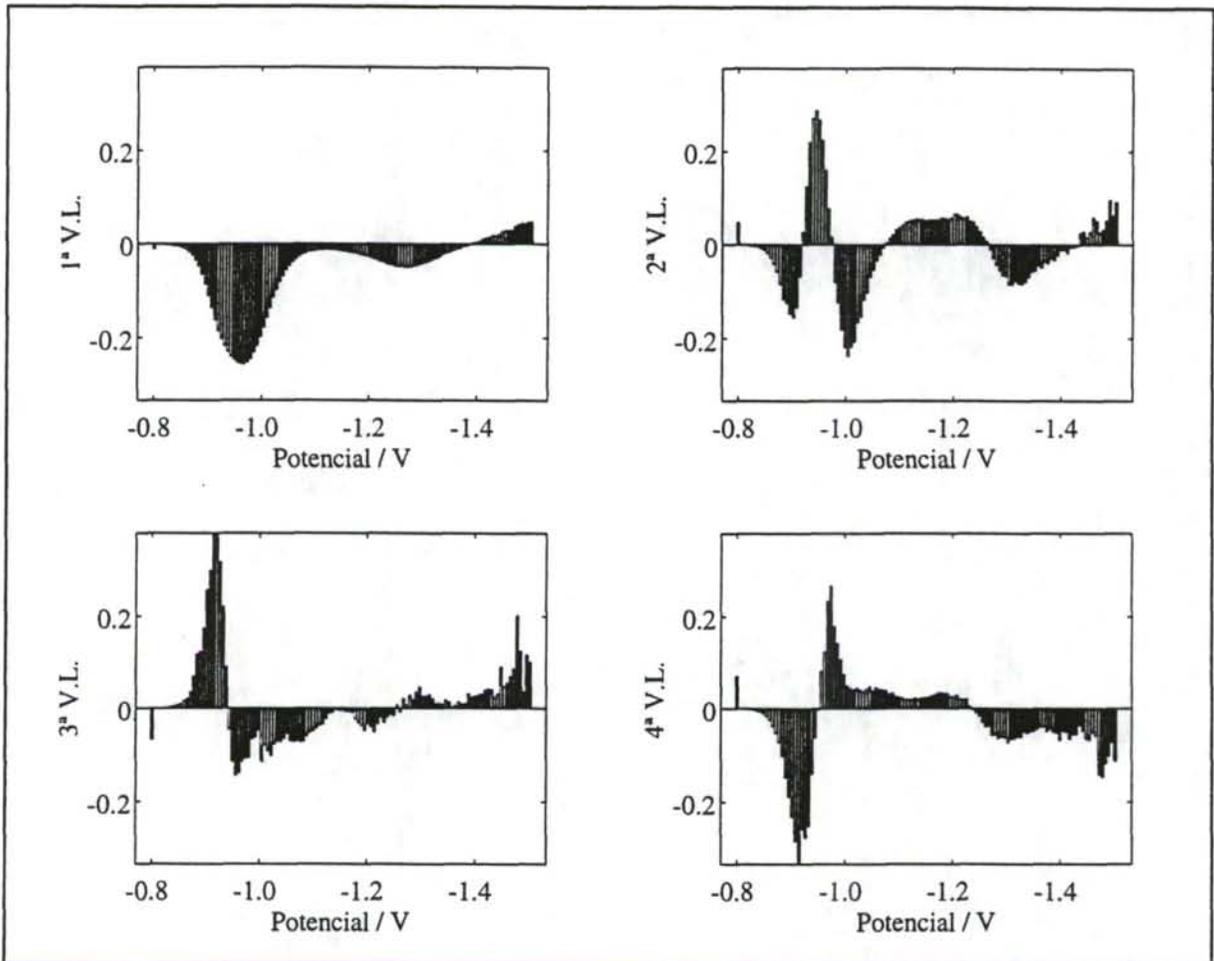


Figura II.2.4b Loadings correspondientes al modelo PLS construido para pH = 2.8.

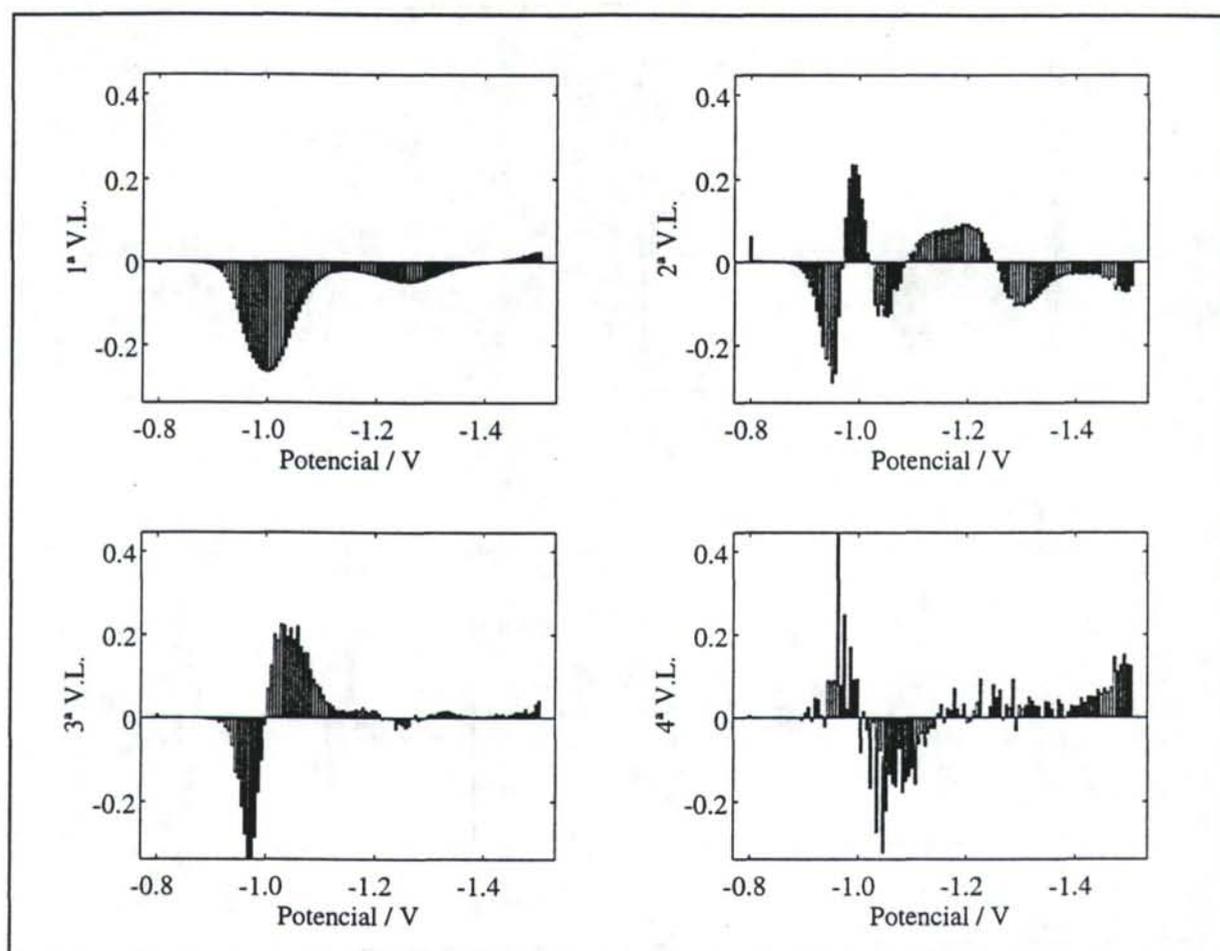


Figura II.2.4c Loadings correspondientes al modelo PLS construido para pH = 3.4.

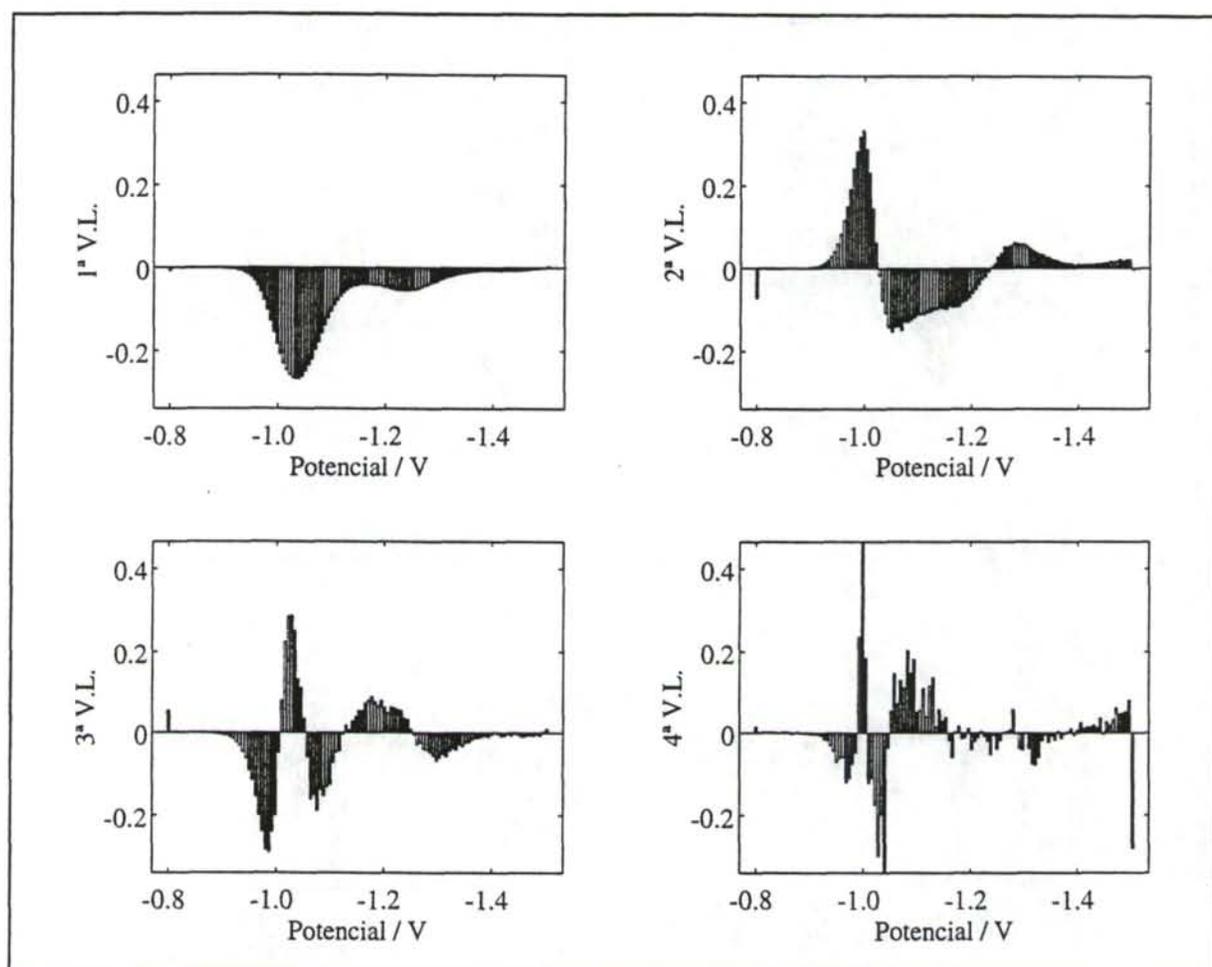


Figura II.2.4d Loadings correspondientes al modelo PLS construido para pH = 4.0.

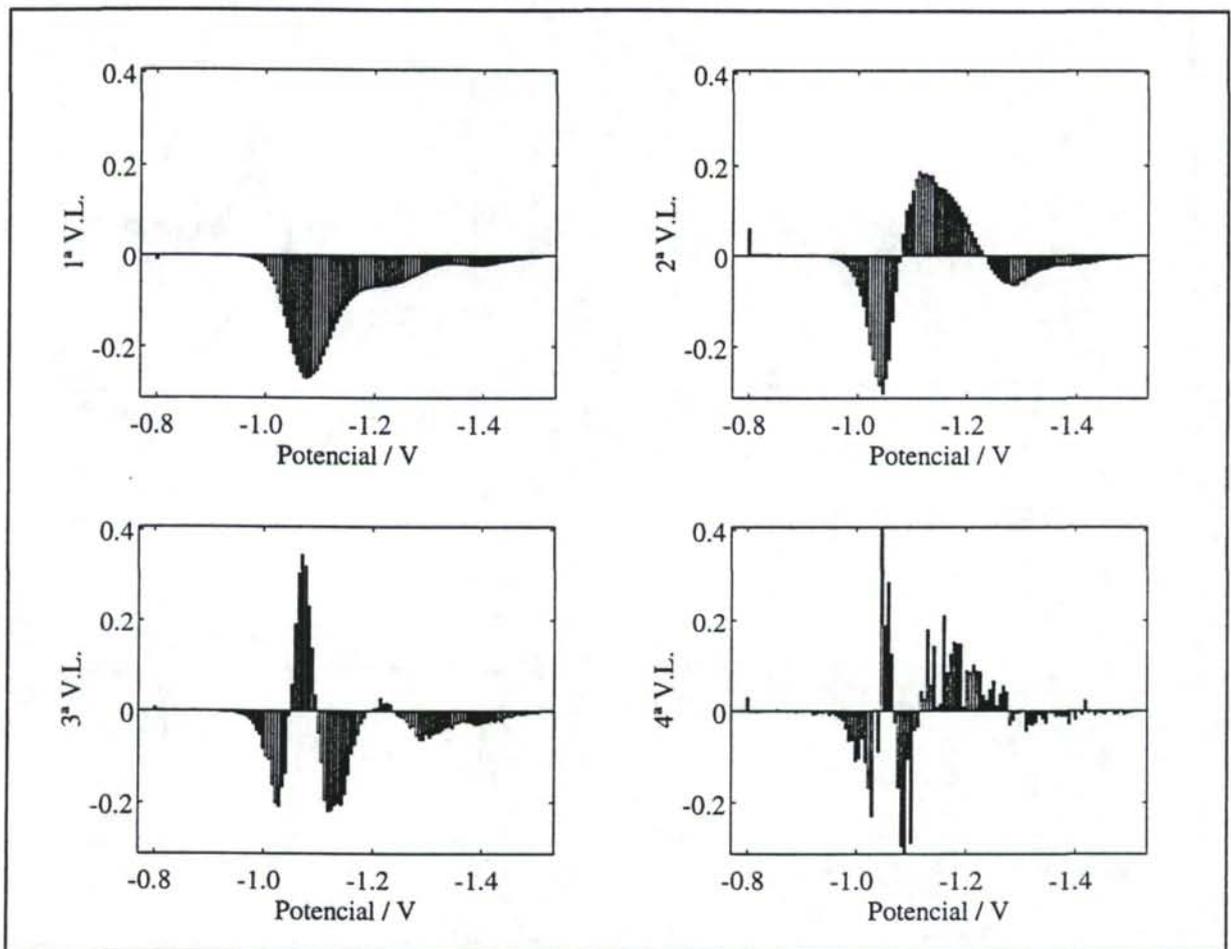


Figura II.2.4e Loadings correspondientes al modelo PLS construido para pH = 4.6.

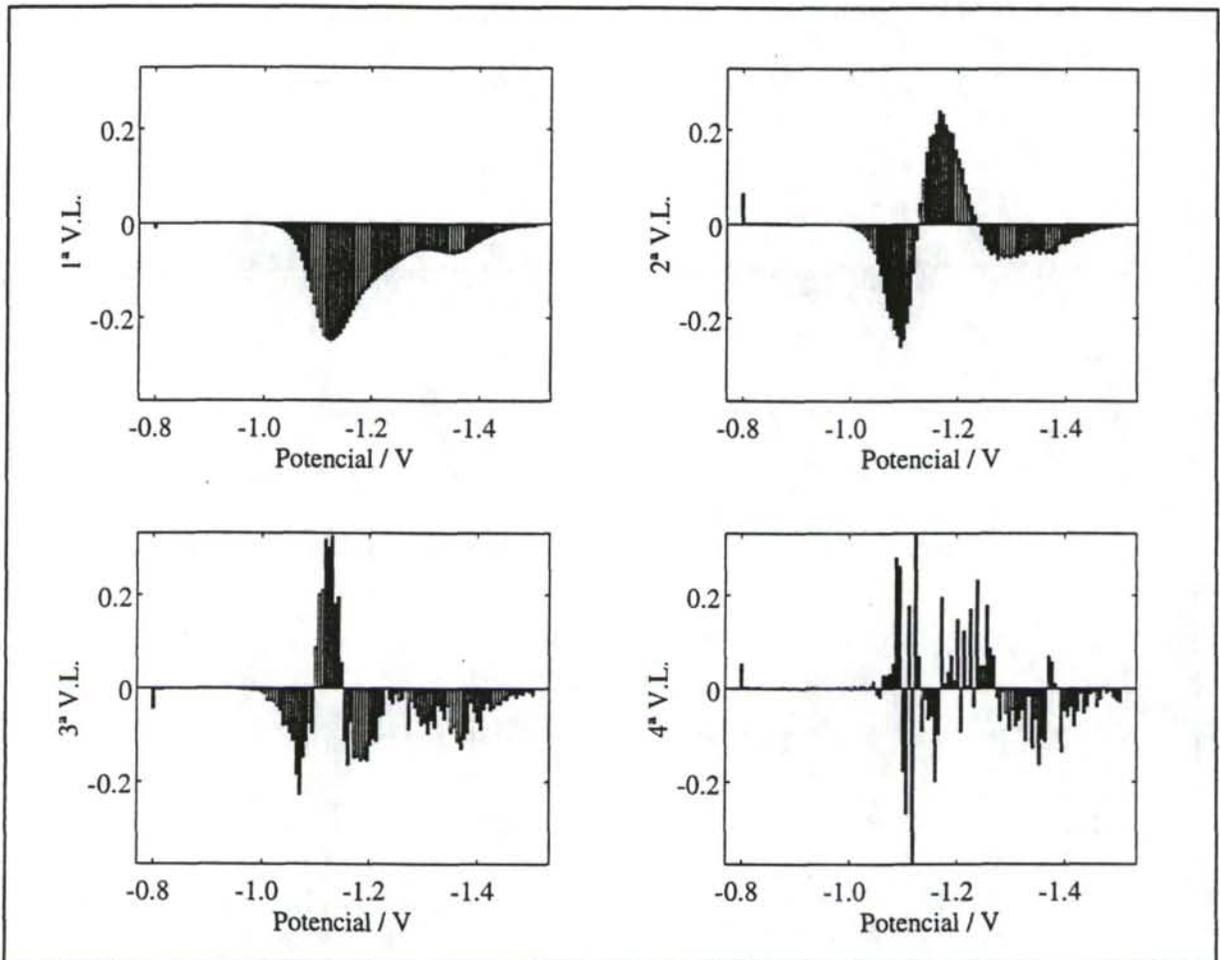


Figura II.2.4f Loadings correspondientes al modelo PLS construido para pH = 5.2.

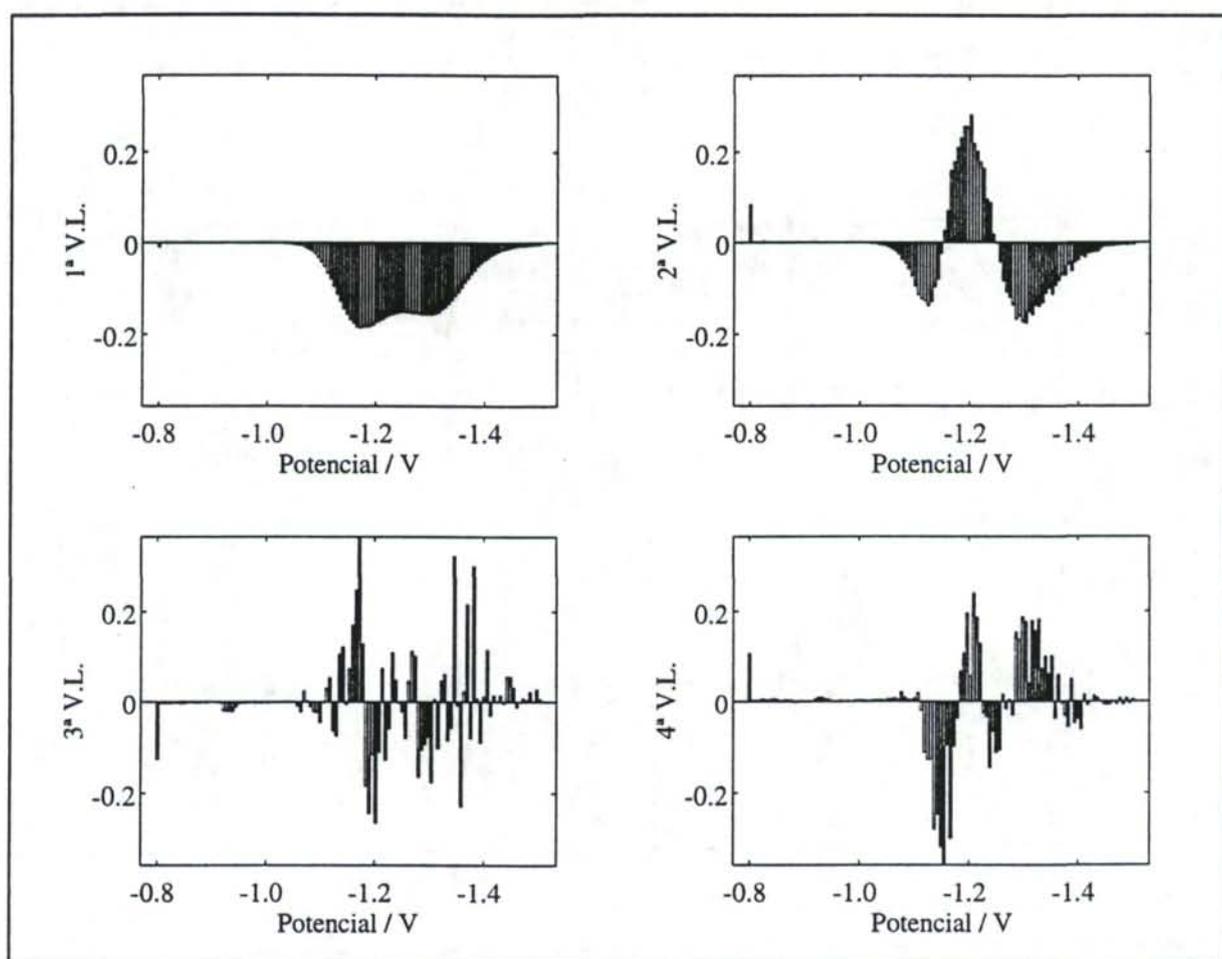


Figura II.2.4g Loadings correspondientes al modelo PLS construido para pH = 5.8.

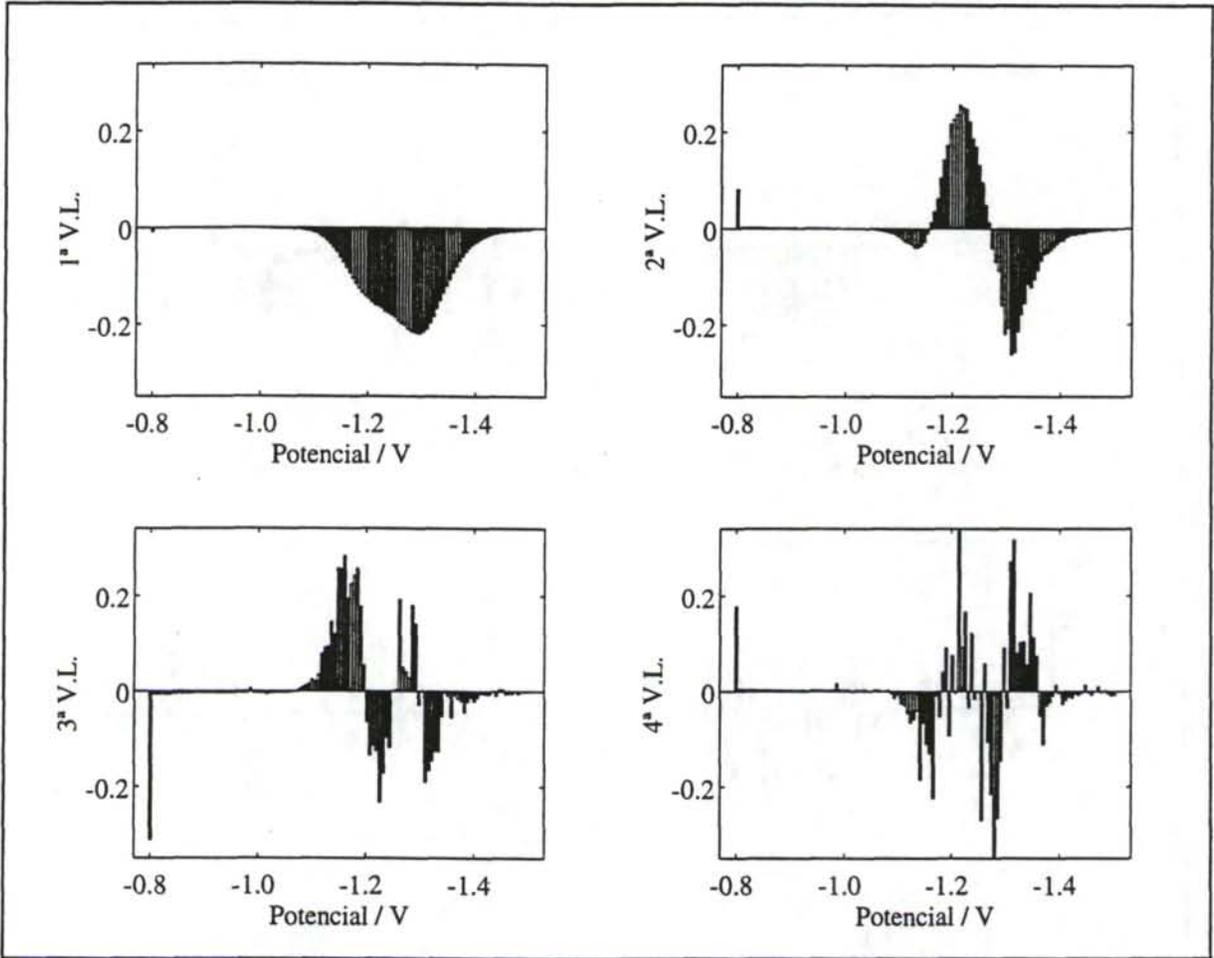


Figura II.2.4h Loadings correspondientes al modelo PLS construido para pH = 6.4.

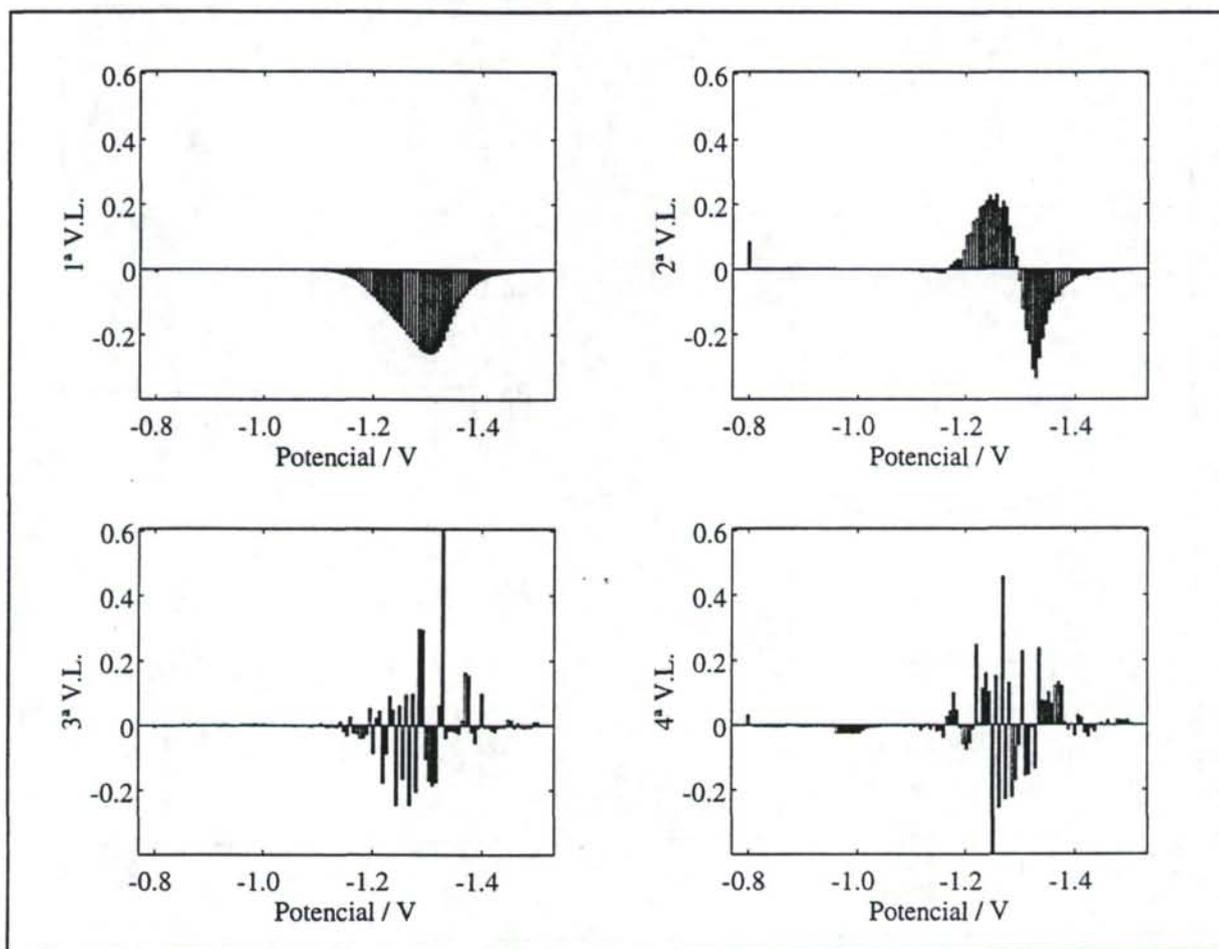


Figura II.2.4i Loadings correspondientes al modelo PLS construido para pH = 7.0.

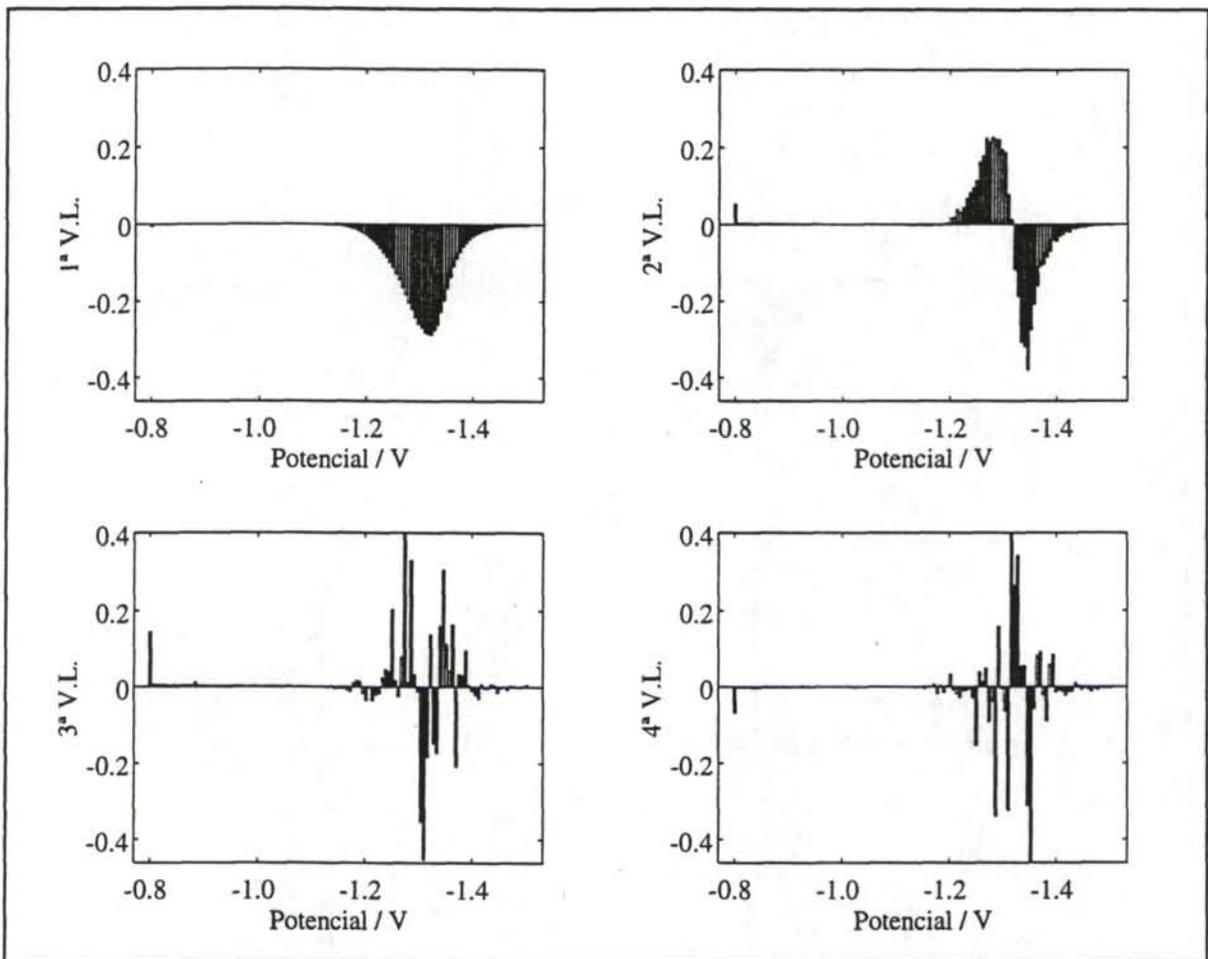


Figura II.2.4j Loadings correspondientes al modelo PLS construido para pH = 7.6.

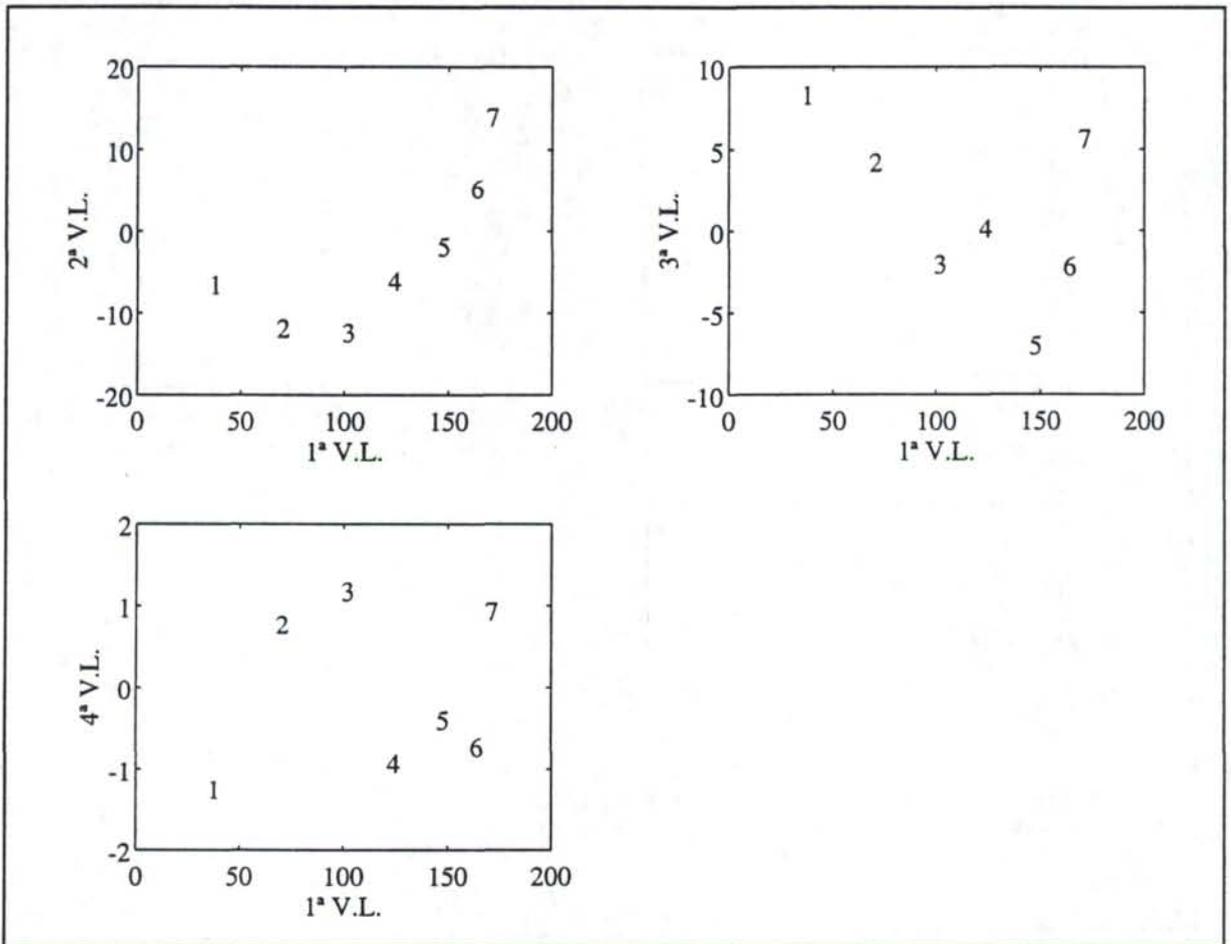


Figura II.2.5a Scores correspondientes al modelo PLS construido para pH = 2.2.

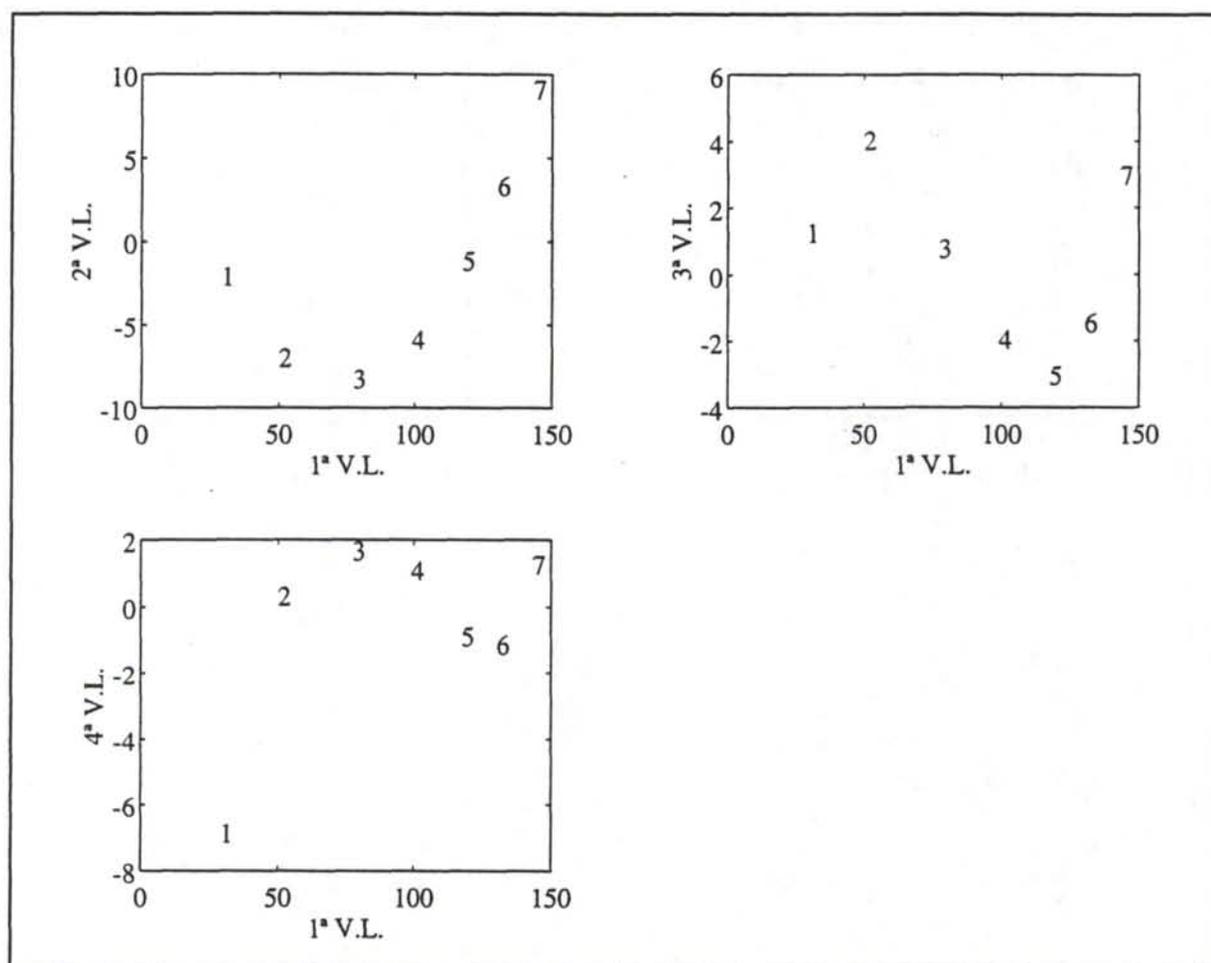


Figura II.2.5b Scores correspondientes al modelo PLS construido para pH = 2.8.

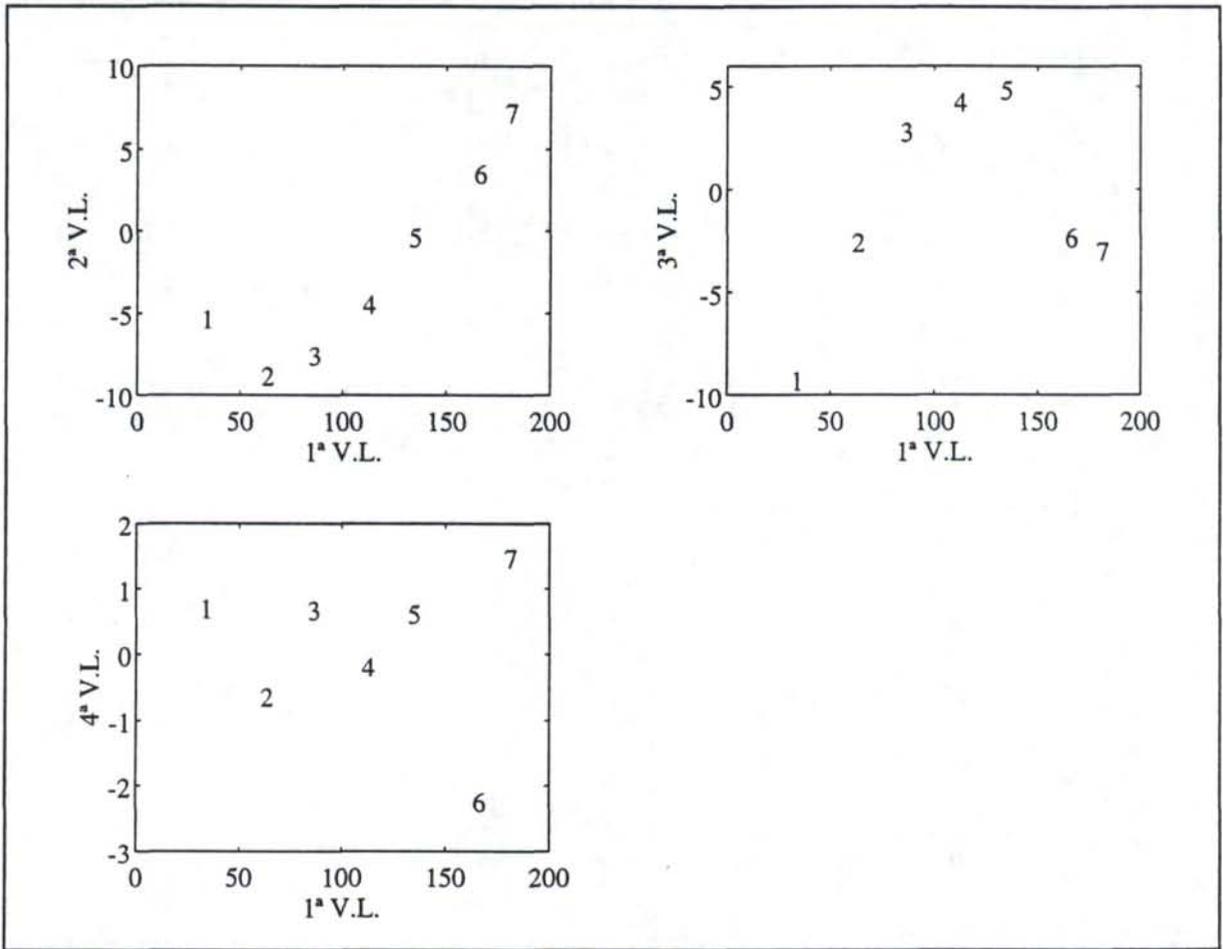


Figura II.2.5c Scores correspondientes al modelo PLS construido para pH = 3.4.

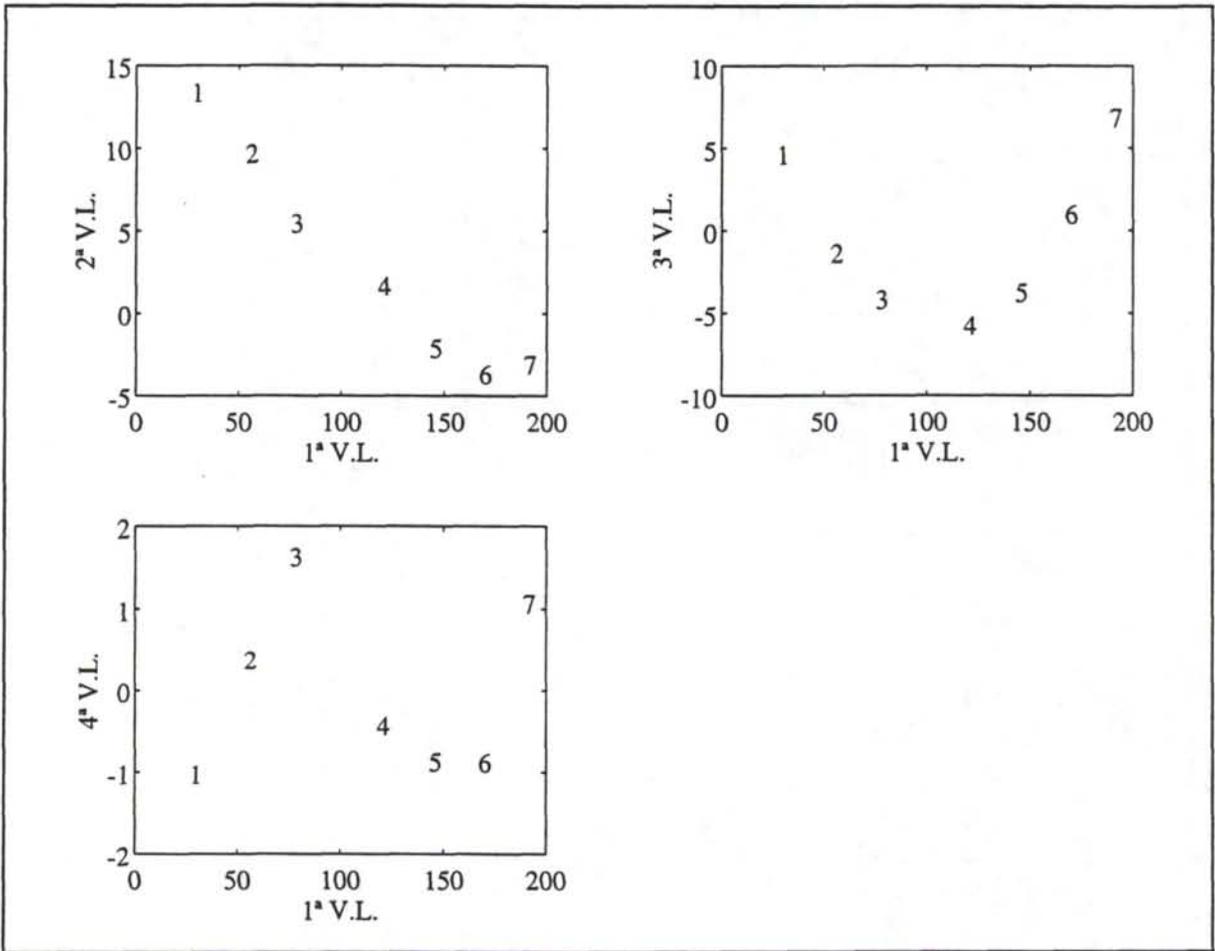


Figura II.2.5d Scores correspondientes al modelo PLS construido para pH = 4.0.

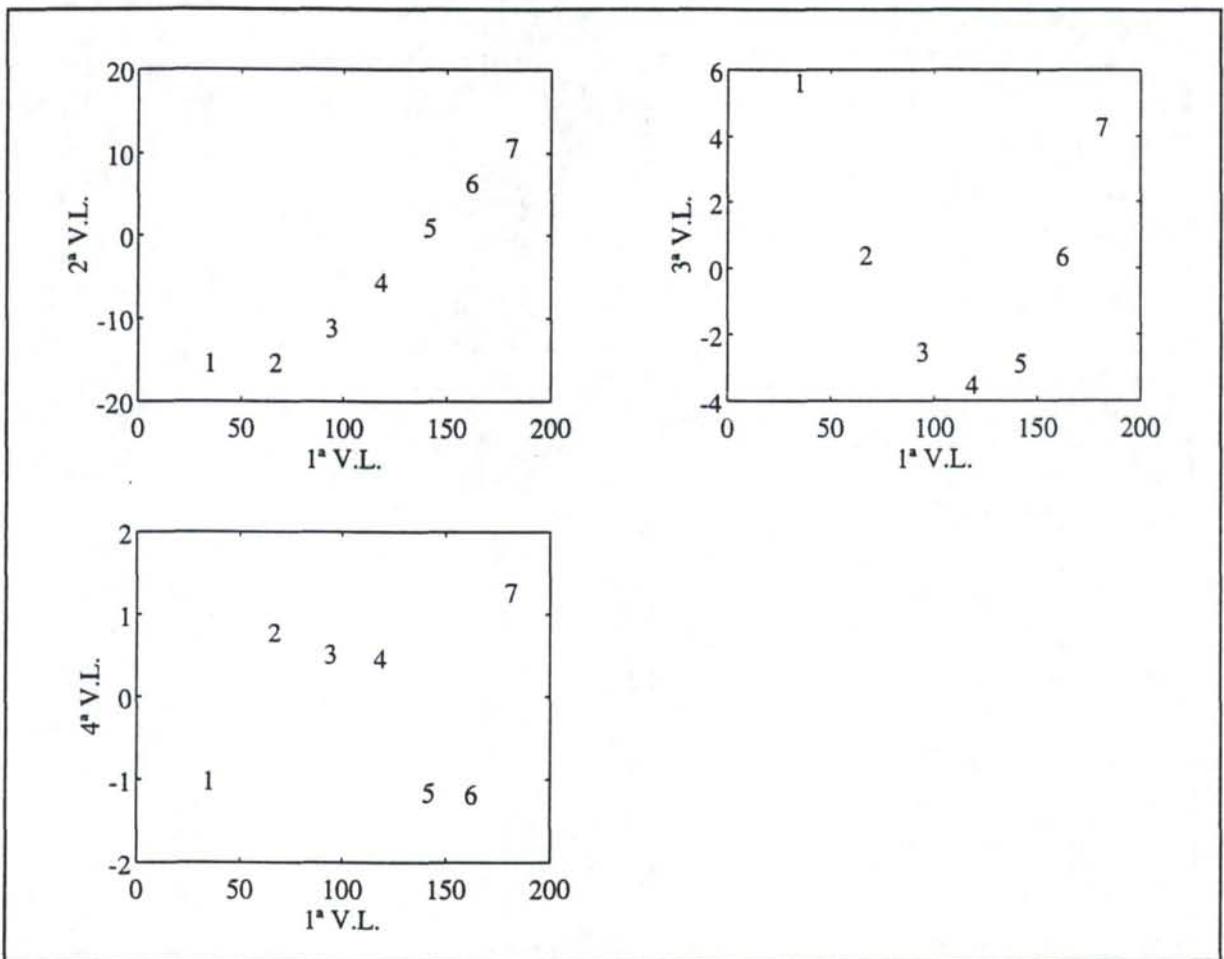


Figura II.2.5e Scores correspondientes al modelo PLS construido para pH = 4.6.

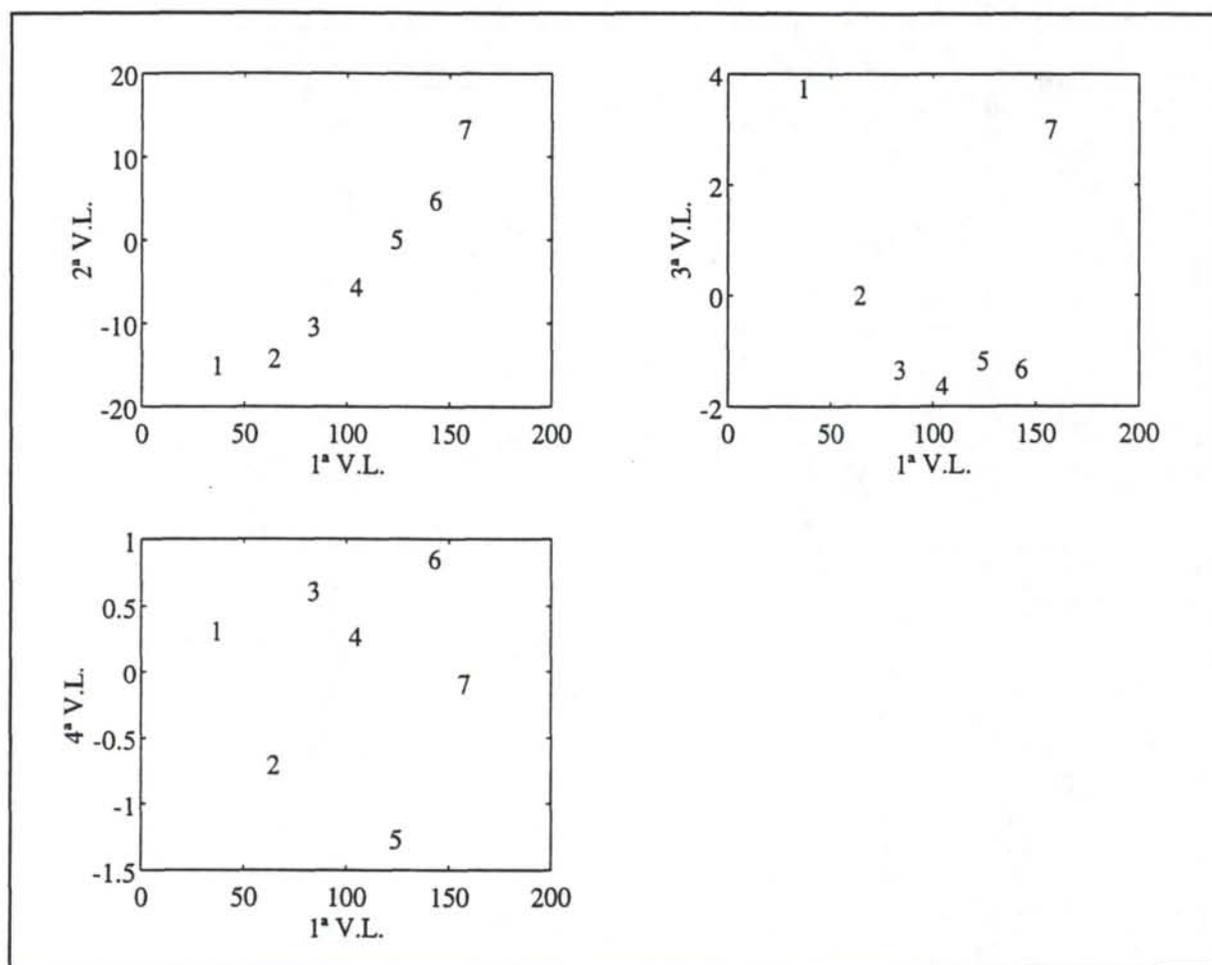


Figura II.2.5f Scores correspondientes al modelo PLS construido para pH = 5.2.

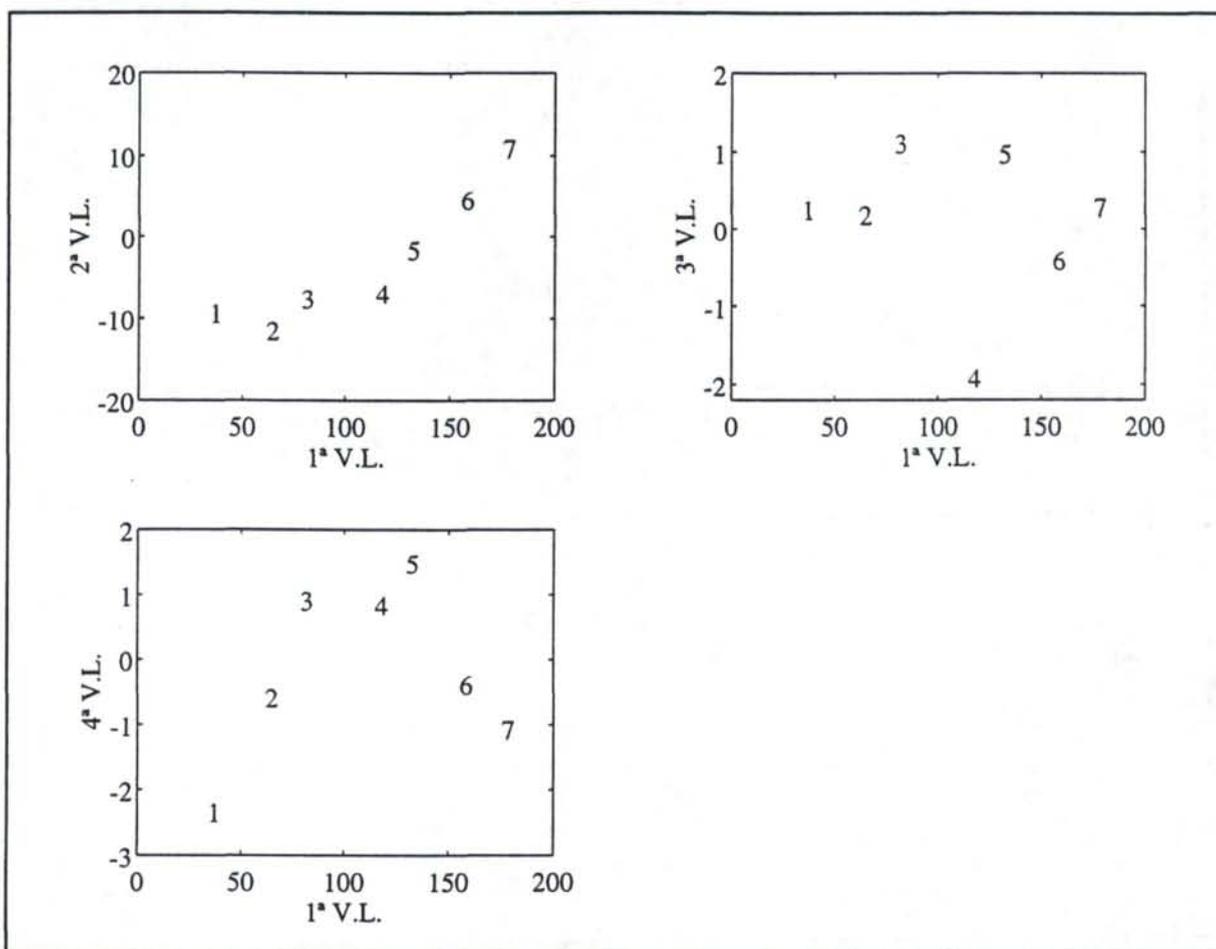


Figura II.2.5g Scores correspondientes al modelo PLS construido para pH = 5.8.

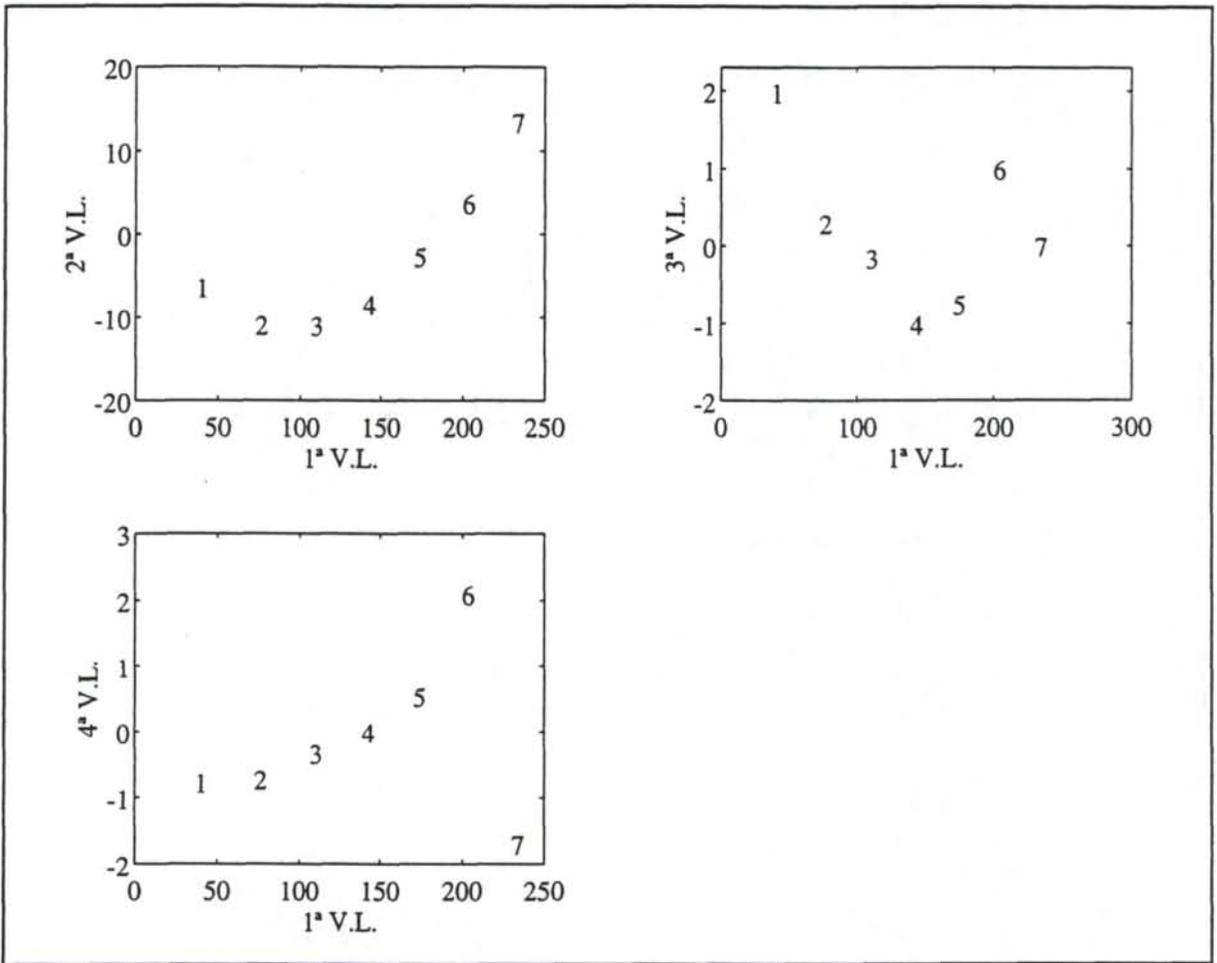


Figura II.2.5h Scores correspondientes al modelo PLS construido para pH = 6.4.

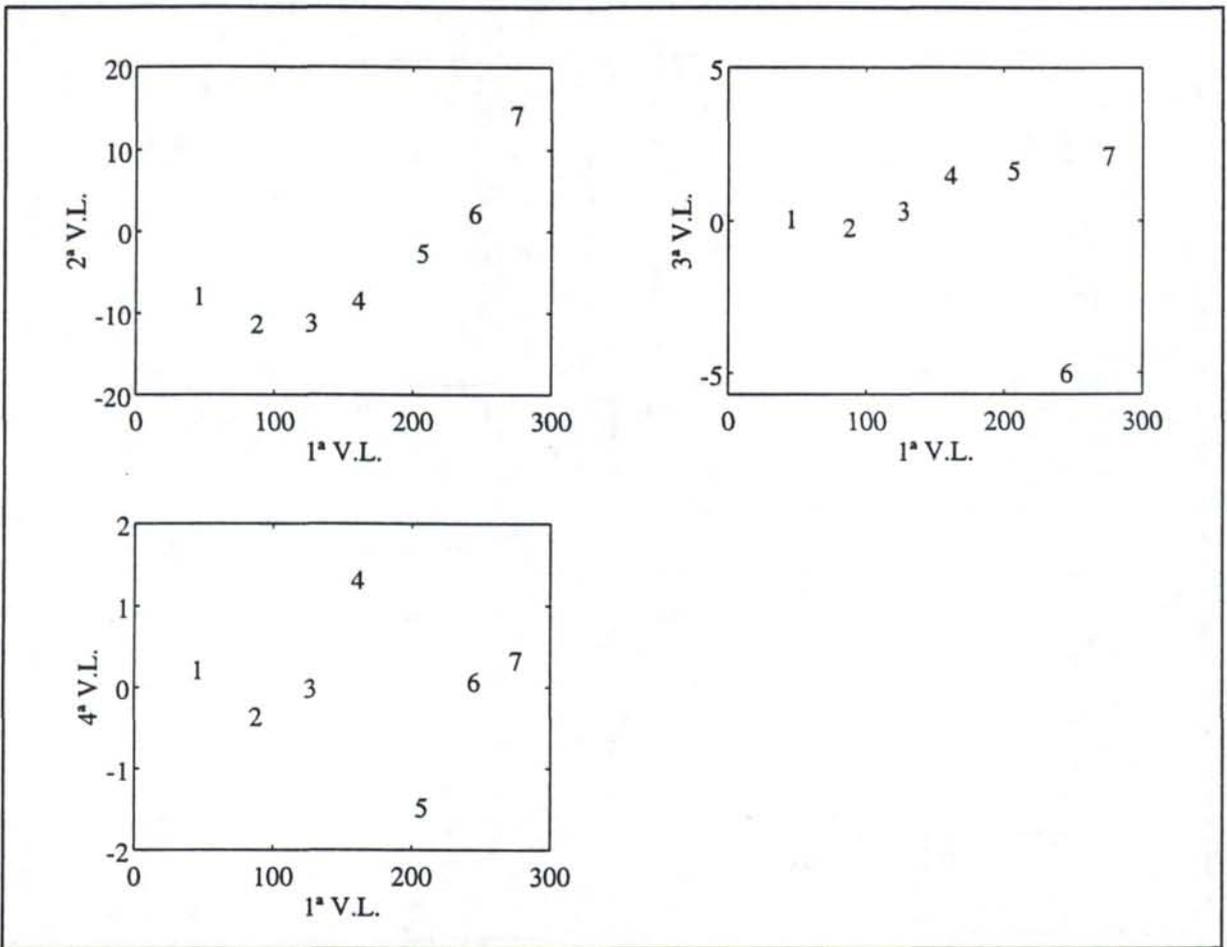


Figura II.2.5i Scores correspondientes al modelo PLS construido para pH = 7.0.

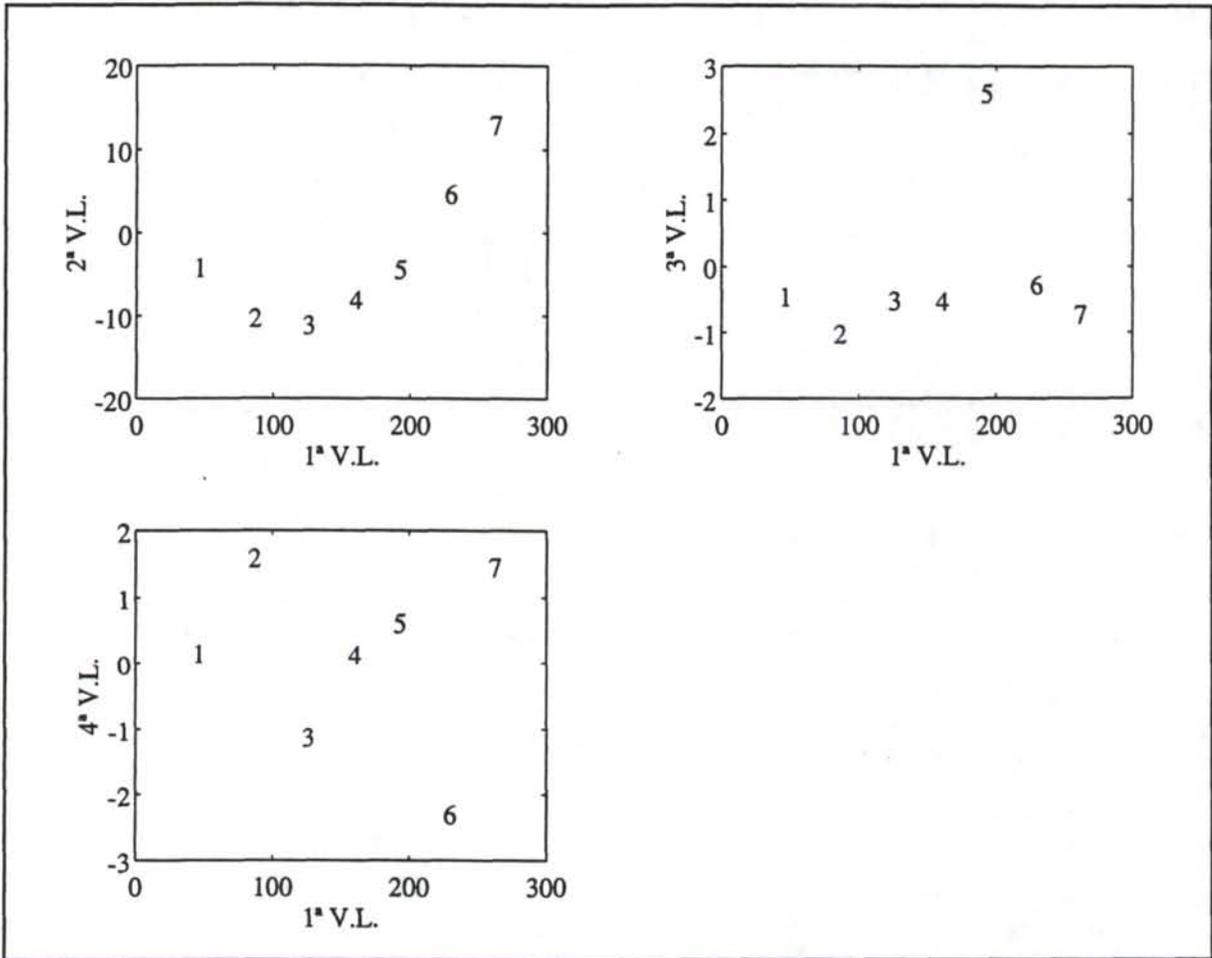


Figura II.2.5j Scores correspondientes al modelo PLS construido para pH = 7.6.

En las Figuras II.2.4a a II.2.4j se observa que, para todos los valores de pH considerados, en la primera variable latente los pesos de las variables en esa componente presentan la forma de los polarogramas que se utilizaron para construir cada modelo PLS. Esto indica que esa primera variable latente es un factor tamaño en todos los casos. Esta primera variable latente modela una fracción muy significativa del benzaldehído añadido (la varianza que explica esta componente está siempre por encima del 94.72 %). Esto se confirma analizando los scores, en las Figuras II.2.5a a II.2.5j, en cuales se observa cómo la puntuación dada por la primera variable latente modela la concentración en todos los casos y, en consecuencia, sobre este eje aparecen de forma correlativa las puntuaciones 1, 2, ..., 7.

En cuanto a la segunda variable latente, hay que señalar que sí existen diferencias entre los distintos modelos construídos para cada valor de pH. En las Figuras II.2.4a a II.2.4c se observa que los loadings correspondientes a los picos polarográficos se encuentran agrupados y presentan signos contrarios. En estos tres casos, la segunda variable latente parece estar relacionada con el desplazamiento del potencial de pico, provocado por la variación de la concentración de benzaldehído. Las Figuras II.2.1a a II.2.1c muestran cómo los potenciales de pico se desplazan en un sentido u otro con las distintas adiciones realizadas.

Esto también aparece reflejado en los scores, Figuras II.2.5a a II.2.5c, donde se observa que los pesos de los objetos en esa segunda variable latente presentan una tendencia distinta en función del desplazamiento que hayan sufrido los potenciales de los picos polarográficos. También se advierte que esta variable latente explica menos varianza cuanto mayor es el pH, de hecho, pasa de explicar un 4.82% en el modelo construído a pH = 2.2 a explicar un 0.35% en el de pH = 3.4, de acuerdo con el proceso químico que hace que las señales polarográficas se desplacen con al variar el pH.

La estructura de los loadings y la distribución de los scores en el primer plano factorial que se observa en los tres primeros pH para la segunda variable latente, pasa a ocupar el lugar de la tercera variable latente en los modelos construídos para valores de pH mayores de 3.4 y menores de 7.0, en los cuales la tercera variable latente explica un máximo de 0.41% de varianza. Esto indica que a estos valores de pH (por encima de 4.0) se produce algún otro fenómeno que tiene mayor relevancia que el expuesto anteriormente.

En este caso, se trata de la aparición de un nuevo pico polarográfico a potenciales más negativos, que se va desplazando hacia potenciales mayores a medida que aumenta el pH, lo cual origina su solapamiento con los picos ya existentes. La aparición de este tercer pico y

su interacción con los otros dos cobra mayor importancia, al variar la concentración de benzaldehído, que el desplazamiento de los potenciales de pico correspondientes, explicando por ello mayor varianza y, por tanto, apareciendo reflejado en la segunda variable latente, ya que éstas se ordenan siguiendo una pauta de mayor a menor varianza explicada.

Esta variable latente está relacionada con la evolución del solapamiento de los picos a medida que el potencial al que aparecen éstos cambia con las adiciones de benzaldehído. De hecho, en los modelos construídos a $\text{pH} = 4.6$ y $\text{pH} = 5.2$, en los cuales se aprecia con claridad el tercer pico, solapado en mayor o menor medida con los otros dos, se hace absolutamente necesaria esa tercera variable latente, llegando a explicar un 0.41% de varianza en el modelo construído a $\text{pH} = 4.6$.

El modelo obtenido para $\text{pH} = 4.0$ marca la transición entre una situación y otra, ya que, a este pH el tercer pico aparece muy tímidamente, Figura II.2.1d. Además, se observa que el desplazamiento de los potenciales de pico no es tan acusado como a pH menores, lo cual hace que sólo sea necesaria una variable latente para construir el modelo PLS correspondiente a ese valor de pH .

La tercera variable latente está relacionada, en los modelos PLS construídos para el primer grupo de valores de pH (entre 2.2 y 3.4), con el fenómeno de reducción de los protones del medio, que por supuesto es más acusado a valores de pH más bajos, en los que la concentración de protones es mayor. De hecho, son necesarias tres variables latentes, que explican un 99.81% de varianza, en el modelo construído para $\text{pH} = 2.2$, ya que para los otros dos valores de pH el porcentaje de varianza explicada por esta variable latente disminuye. Se observa en los loadings de las Figuras II.2.4a a II.2.4c correspondientes a la tercera variable latente como, a los potenciales a los que se produce este fenómeno, la importancia relativa de estos, en cuanto al valor del loading, va disminuyendo a medida que aumenta el pH , siendo mínima a $\text{pH} = 3.4$.

En los modelos construídos para valores de pH mayores de 6.4, la tercera variable latente presenta loadings aleatorios, Figuras II.2.4i y II.2.4j, y la distribución de scores responde a peculiaridades de objetos concretos, Figuras II.2.5i y II.2.5j. Esto indica que esa variable latente únicamente explica la variabilidad debida a objetos que difieren ligeramente del resto y, por tanto, no debe ser tenida en cuenta en la construcción del modelo. La segunda variable latente se corresponde con la estructura indicada anteriormente para la segunda variable latente de los modelos construídos para valores de pH entre 4.0 y 6.4.

Respecto a la cuarta variable latente, en las Figuras II.2.4a a II.2.4j, se observa que la distribución de loadings es aleatoria, y que en cualquier caso, aunque la varianza explicada por esta variable aumenta, la varianza explicada en predicción disminuye, indicando que la inclusión de esta nueva variable introduciría inestabilidad en los modelos. De hecho, la distribución de los scores, en las Figuras II.2.5a a II.2.5j, indica que la cuarta variable latente está muy ligada a objetos concretos y que, por lo tanto, no debe ser tenida en cuenta en la construcción de los modelos PLS.

Análisis de los resultados

En la Tabla II.2.5 se evalúan las diferencias entre las concentraciones de benzaldehído reales y las calculadas por los modelos PLS. Para el cálculo del error relativo (en %) se ha utilizado el valor de concentración de benzaldehído predicha con el número de variables latentes indicado para cada valor de pH en la Tabla II.2.4. Estos errores son bajos en general y se encuentran dentro de lo esperado en las determinaciones realizadas utilizando polarografía diferencial de impulsos, en estos rangos de concentraciones. Evidentemente, los errores mayores se encuentran asociados con las concentraciones de benzaldehído más bajas, así como con los valores de pH más ácidos, ya que en esas condiciones la interferencia debida a la señal de reducción de los protones del medio es más importante.

Los resultados obtenidos avalan la posibilidad de utilizar la polarografía diferencial de impulsos en la determinación de benzaldehído, cualquiera que sea el pH del medio o, lo que es lo mismo, independientemente de la especificidad de la señal experimental.

Por otro lado, se ha evaluado el efecto que tienen la variación del pH del medio y las distintas adiciones de analito sobre el valor de la concentración de benzaldehído calculado. Una forma de llevar a cabo este estudio consiste en realizar un análisis de la varianza (ANOVA), el cual requiere que se verifiquen las hipótesis de normalidad de restos e igualdad de varianzas para que las conclusiones que se extraigan del mismo sean válidas.

Por ello, en primer lugar se ha verificado la normalidad de restos utilizando los siguientes test de hipótesis [58] al nivel de confianza del 95%: el test de la χ^2 ($P_{\text{actual}} = 0.1350$) y el test de Kolmogorov-Smirnov ($P_{\text{actual}} = 0.5804$). En ambos casos, no es posible rechazar la hipótesis nula (H_0 : los datos son normales) ya que la P_{actual} es mayor de 0.05, con lo que hay que considerar que los datos sí son normales.

Tabla II.2.5 Errores relativos (%) obtenidos en el cálculo de la concentración de benzaldehído por los modelos PLS, a los diferentes valores experimentales de pH.

pH	Adición						
	1	2	3	4	5	6	7
2.2	8.64	-3.00	-2.69	1.84	0.25	1.13	-1.07
2.8	14.63	-9.27	-1.62	1.55	2.96	-0.21	-1.32
3.4	1.84	2.28	-2.08	-0.45	-1.55	3.52	-1.64
4.0	-6.54	-5.13	-10.59	5.80	3.00	0.85	-1.99
4.6	0.58	0.11	-0.07	-0.63	0.36	0.41	-0.29
5.2	-1.83	2.68	-1.18	-0.46	0.39	-0.08	0.03
5.8	4.00	1.82	-7.28	4.53	-2.29	0.87	-0.04
6.4	-1.45	-0.13	0.10	0.24	0.20	-0.36	0.11
7.0	-2.83	0.74	0.03	-2.88	1.54	1.50	-1.02
7.6	2.19	0.65	1.48	-0.10	-1.70	0.37	0.23

No obstante, se ha comprobado que no se verifica la hipótesis de igualdad de varianzas, mediante los test de Cochran y Bartlett con ese mismo nivel de confianza, tanto cuando se considera el nivel adición y los errores a los distintos pH como réplicas, como cuando se tiene la situación contraria.

Por lo tanto, los resultados obtenidos a través del ANOVA no serían válidos y, en consecuencia, se hace necesario utilizar otro tipo de análisis, que no requiera la verificación de esta última hipótesis, para estudiar estos efectos. Por ello se ha utilizado un método no paramétrico, el test de rangos de Friedman [83], que no exige las hipótesis previas de normalidad e igualdad de varianzas. La hipótesis nula del test, H_0 , es la siguiente: no existe efecto del factor en estudio. El test concluye que ninguno de los factores es significativo ya que $P_{\text{actual}} = 0.2531 > 0.05$, es decir, que no existe efecto.

Este resultado permite considerar como réplicas los valores de concentración de benzaldehído calculados a los distintos valores experimentales de pH para cada adición de analito. Para cada valor de la concentración de benzaldehído se dispondría de 10 réplicas, asumiendo como tal el valor de concentración calculado a cada uno de los diferentes pH experimentales.

Una regresión por mínimos cuadrados permitiría comparar los valores de la concentración de benzaldehído calculados por los modelos de regresión PLS frente a los valores de concentración reales. Si la pendiente es 1 y el término independiente es 0 implicaría que son iguales. Para cada adición de benzaldehído se consideran 10 réplicas, los valores de concentración calculados a cada uno de los diez valores experimentales de pH. Los resultados del ajuste llevado a cabo aparecen reflejados en la Tabla II.2.6.

Tabla II.2.6 Resultados de la regresión lineal por mínimos cuadrados realizada para comparar los valores de concentración calculados con los modelos PLS y los valores reales.

Variable	Coficiente	S.E.	P _{actual}
Constante	-0.0258	0.1191	0.83
Concentración	1.0006	0.0055	0.00
$\rho = 0.9990$	$S_{yx} = 0.4405$		

En este caso, la hipótesis conjunta [57], que plantea que el término independiente sea igual a 0 y la pendiente igual a 1, se acepta con un nivel de significación del 3.54%. Lo cual indica que, independientemente del valor de pH, o lo que es lo mismo de la calidad de la señal analítica, y para todas las concentraciones utilizadas, los valores de concentración encontrados con los modelos PLS son iguales, desde un punto de vista estadístico, a los valores reales.

III. TRANSFERENCIA DE CALIBRADO

III.1 INTRODUCCIÓN

Las técnicas electroanalíticas aplicadas a la determinación de elementos a nivel de trazas suponen en la actualidad una muy interesante alternativa a los ya tradicionales procedimientos espectroscópicos. El uso de una metodología multivariante, a través de técnicas tales como la regresión sobre componentes principales (PCR) o la regresión en mínimos cuadrados parciales (PLS), permite extraer de forma adecuada la información contenida en las señales electroquímicas, potenciando enormemente las posibilidades de las citadas técnicas electroanalíticas.

No obstante, la aplicabilidad de estas técnicas encuentra una limitación en el hecho de que un modelo de regresión multivariante, construido con un conjunto de datos determinado, generalmente no es válido o aplicable a otros conjuntos de datos diferentes. Esto provoca que en determinadas situaciones puedan plantearse una serie de problemas que dificultan lo que se denomina 'transferencia de calibrado' [84]. En algunos casos, es posible que la respuesta dada por dos instrumentos al medir una misma muestra en iguales condiciones sea diferente, o que exista inestabilidad en la respuesta a lo largo del tiempo (deriva, etc.), que varíen algunas condiciones experimentales tales como humedad o temperatura, o incluso que, por alguna razón, se produzcan cambios en la constitución física de la muestra (cambios de tamaño de partícula o de textura), etc.

Sin embargo, es sabido que en numerosas ocasiones, pequeñas alteraciones de las señales analíticas pueden ser indicio de cambios fundamentales en el sistema en estudio. Por lo tanto, ligeras variaciones entre una misma medida realizada en diferentes situaciones podrían ser suficientemente significativas para impedir la construcción de un modelo de calibrado que pudiera ser transferido, o lo que es lo mismo, empleado posteriormente con muestras que no formaron parte del conjunto de datos utilizado en su construcción.

Una solución obvia a los problemas asociados a una transferencia de calibrado reside en trasladar el conjunto de calibración completo a la nueva situación, lo cual supone repetir todas las medidas electroquímicas con la finalidad de construir otro modelo de calibrado con

las mismas características. Este procedimiento, que es el que tradicionalmente se ha utilizado, supone una importante carga adicional, tanto en términos de esfuerzo experimental como en tiempo de trabajo, sobre todo cuando las muestras son muy numerosas, químicamente inestables, seleccionadas al azar, etc. En cualquier caso, sería deseable poder llevar a cabo la calibración de un instrumento en una nueva situación sin necesidad de repetir la medida de todas las muestras del calibrado por segunda vez.

En este sentido, se han propuesto diferentes procedimientos para reducir el número de muestras de calibrado necesarias para llevar a cabo una regresión multivariante [85]. Se trata de procedimientos de selección de conjuntos de calibrado que, partiendo de un determinado número de datos, eligen aquellos objetos que mejor caracterizan el sistema multicomponente. Procedimientos tales como el análisis de agrupamientos [86], la eliminación iterativa de muestras similares utilizando la distancia de Mahalanobis [87], la selección iterativa de filas (o columnas) o combinaciones de filas (o columnas) más ortogonales entre sí para realizar un análisis factorial [88] o la selección de aquellos objetos que conduzcan a coeficientes de regresión con pequeña variabilidad utilizando el algoritmo de intercambio de Fedorov [89].

A pesar de que este tipo de problemas impone severas restricciones a la aplicación de los modelos de calibrado multivariante no son muchos los ejemplos existentes en la bibliografía al respecto, la mayoría de los cuales han sido publicados en análisis con técnicas de infrarrojo cercano o próximo (NIR). Algunos de estos métodos son aplicaciones a productos agrícolas [90, 91] que suponen la realización del calibrado utilizando una única, o a veces unas pocas, longitudes de onda, por lo que no son generalmente aplicables al calibrado multivariante basado en la utilización de las respuestas del espectro completo.

Existe un método propuesto por Shenk y Westerhaus [92] que trata de corregir la respuesta experimental completa a través de un esquema univariante. Para cada longitud de onda del espectro de un instrumento patrón se detecta la longitud de onda con mayor correlación del espectro de un instrumento satélite. Con ésta y otras dos longitudes de onda del espectro del instrumento satélite próximas a ella se calcula un modelo cuadrático para la longitud de onda del instrumento patrón. Se evalúa la correlación de este modelo con la longitud de onda del instrumento patrón, y la posición que se corresponde con el máximo de correlación es considerada como la longitud de onda del instrumento satélite que mejor reproduce la del instrumento patrón. Una vez corregido el desplazamiento de la longitud de onda, se aplica una corrección para los valores de absorbancia del instrumento satélite, a las

que se quiere transferir el calibrado del instrumento patrón. El algoritmo desarrollado en este método ha sido transformado y mejorado por Bouveresse y otros [93].

Wang y otros [94,95] comparan cinco procedimientos, el de Shenk anteriormente citado y cuatro métodos desarrollados por ellos mismos para llevar a cabo la transferencia de calibrados multivariantes. Estos procedimientos fueron aplicados en la determinación mediante espectroscopía NIR de cuatro analitos en muestras de gasolina utilizando dos espectrofotómetros diferentes, siendo el método de estandarización directa a trozos (piecewise direct standardization, PDS) el método que ofreció mejores resultados [94,96]. Este método se fundamenta en la regresión sobre componentes principales (PCR) o bien en la regresión por mínimos cuadrados parciales (PLS), y también ha sido comparado con una corrección univariante pendiente/sesgo en espectroscopía NIR [97]. Basándose en la medida de un pequeño conjunto de muestras con los dos instrumentos, se determina una matriz de transferencia que se aplica para transformar espectros entre los dos instrumentos, lo cual hace también posible la transferencia de los modelos de calibrado entre ambos.

Bouveresse y otros [98] proponen algunas modificaciones del método PDS aplicado a instrumentación NIR con la finalidad de detectar y reducir las posibles causas de predicciones erróneas. Estudian en primer lugar la selección de muestras de transferencia, llevándola a cabo utilizando el algoritmo propuesto por Kennard y Stone [99] y otro método basado en la maximización de la distancia menor entre puntos, el cual intenta mejorar los resultados obtenidos con la selección basada en la influencia de los puntos (leverage). Una vez resuelta la selección del conjunto de transferencia, presentan un método basado en la predicción de una muestra adicional independiente (el espectro NIR más próximo al espectro medio del conjunto de datos) para detectar y reducir anomalías debidas a malas estimaciones locales.

Sin embargo, como ya se ha señalado, las distintas situaciones en las que resulta interesante realizar una transferencia de calibrado no se limitan a los análisis llevados a cabo con diferentes instrumentos, sino que también pueden tener su origen en diferentes condiciones físicas de las propias muestras, en variaciones en el tiempo en que se realizan las medidas, etc. De hecho, el método PDS ha sido aplicado a la estandarización de señales espectrofotométricas medidas a diferentes temperaturas [100], considerando la variación de temperatura como un tipo especial de variación instrumental, o incluso cuando las medidas se han realizado en días distintos [101].

Otro método, propuesto por Forina y otros [102], sugiere realizar una transferencia de calibrado multivariante a través de la transferencia de la ecuación de regresión. Este procedimiento utiliza la regresión en mínimos cuadrados parciales (PLS) en dos ocasiones: en una primera etapa se calcula la relación entre los espectros de las muestras de transferencia de los dos instrumentos, y a continuación se calcula la ecuación de regresión (relación entre las variables químicas, concentraciones, y las variables espectrales) del instrumento patrón. Estos dos pasos de aplicación de la regresión PLS se combinan para predecir el modelo de regresión del segundo instrumento o instrumento satélite.

Asimismo, se ha desarrollado un método para llevar a cabo la transferencia de calibrado entre instrumentos de segundo orden, aplicándose éste a un ejemplo en cromatografía líquida-espectroscopía ultravioleta (LC-UV) [103]. Este método se basa en el cálculo de dos matrices de transferencia (matrices banda diagonales), utilizando las variables predictoras de una muestra patrón común para corregir simultáneamente el desplazamiento de las variables predictoras así como las variaciones de la respuesta en ambas dimensiones u órdenes. A diferencia de las transferencias de calibrado de primer orden, estas dos matrices de transferencia son estimadas a partir de un conjunto de ecuaciones no lineales simultáneas a través del método de Gauss-Newton.

Aunque la mayor parte de las referencias bibliográficas sobre transferencia de calibrado multivariante se encuentran en el contexto de la espectroscopía NIR, muchas de las estrategias y métodos descritos pueden ser aplicados también a problemas de transferencia de calibrado en otros campos de la Química Analítica [104-107].

En esta memoria se han aplicado dos métodos de transferencia de calibrado diferentes, la transferencia de calibrado a trozos (PDS) y el método propuesto por Forina y otros [102], que se denominará en adelante transferencia de calibrado global (global calibration transfer, GCT), a datos electroquímicos. Éstos se han obtenido en diferentes situaciones; por un lado se han realizado medidas voltamperométricas en días diferentes (transferencia de calibrado a lo largo del tiempo), y en algún caso, se han llevado a cabo determinaciones con instrumentos distintos (transferencia de calibrado entre instrumentos).

En el estudio llevado a cabo se han realizado una serie de calibrados multivariantes utilizando técnicas voltamperométricas para llevar a cabo medidas sobre diferentes sistemas químicos multianalito. Todos ellos constituyen ejemplos representativos de algunos de los problemas más frecuentes con que se enfrenta este tipo de análisis, que hacen que se pierda

la especificidad de las señales. Se han analizado los siguientes casos: existencia de interferencias, formación de compuestos intermetálicos, solapamiento de señales, corrección de la línea base y efecto matriz.

III.2 TEORÍA SOBRE LA TRANSFERENCIA DE CALIBRADO

Los métodos de calibración multivariante pueden ser utilizados para predecir la concentración de uno o más constituyentes de una mezcla compleja, así como parámetros de calidad más abstractos [9]. El modelo de calibrado inverso, $y = f(\mathbf{X}) + e$, donde y es el parámetro que se quiere predecir, \mathbf{X} es la matriz con los datos electroquímicos, y e contiene los residuos, es la forma más versátil de abordar el problema de calibrado [108]. Los parámetros de este modelo se estiman mediante un conjunto de aprendizaje (training set), que contiene n objetos (muestras) de los cuales se dispone de p variables predictoras (intensidades medidas a p valores de potencial) y de la variable a predecir (concentración o parámetro de calidad). En general, se utilizan modelos de regresión lineal multivariante del tipo $y = \mathbf{X}\mathbf{b} + e$, donde los coeficientes de regresión \mathbf{b} se estiman habitualmente a través de una regresión lineal múltiple (MLR) sobre variables seleccionadas, o mediante métodos de calibración que utilizan transformaciones lineales de la señal completa, tales como la regresión sobre componentes principales (PCR) o por mínimos cuadrados parciales (PLS).

El conjunto de aprendizaje debe cubrir todas las fuentes de variabilidad relevantes de tal forma que futuras muestras se encuentren dentro de los intervalos de variación considerados. No obstante, los datos electroquímicos, y por tanto el modelo, también incluyen las fuentes de variabilidad del instrumento y de las condiciones de medida. Esto significa que un modelo de calibrado construido con un conjunto de datos en una situación determinada no sería válido o aplicable a otro conjunto de datos medidos en una situación diferente. En el caso de que se aplique un modelo de calibrado a datos distintos de los que se utilizaron para construirlo, los errores en predicción pueden llegar a ser muy importantes en los casos anteriormente mencionados (cambio de instrumento, deriva, ...), especialmente cuando se utilizan todas las variables predictoras para elaborar el modelo de calibrado (acumulación de errores).

Como se ha indicado previamente el propósito de la transferencia de calibrado es asegurar el intercambio de resultados obtenidos en diferentes situaciones, sin necesidad de

llevar a cabo una recalibración completa en cada una de ellas. En este sentido, el criterio último para evaluar el resultado de una transferencia de calibrado es la calidad de las predicciones que se realizan.

Para llevar a cabo la transferencia de calibrado multivariante existen diferentes aproximaciones que, de forma genérica, pueden clasificarse en dos grupos según realicen la transferencia de calibrado mediante la transformación de los datos de medida o corrigiendo el propio modelo de calibrado. Algunos de los métodos más significativos se muestran con detalle a continuación, utilizando para ello la notación siguiente:

\mathbf{X}_A ($n \times p$) es el conjunto de datos (training set) de la situación inicial o situación A (instrumento patrón, temperatura de referencia, primer día,...) en la que se basa el modelo de calibrado multivariante.

\mathbf{Y} ($n \times q$) es la correspondiente matriz respuesta (concentraciones o parámetro de calidad). Si se trata de un único analito, entonces $q = 1$, y la respuesta es el vector y .

$\bar{\mathbf{X}}_A$ ($m \times p$) contiene las muestras de la situación A que se van a utilizar para llevar a cabo la transferencia de calibrado, es decir, se trata del denominado conjunto de transferencia.

\mathbf{X}_B y $\bar{\mathbf{X}}_B$ son las matrices análogas para la nueva situación o situación B.

III.2.1 Métodos que implican la corrección del modelo de calibrado

El propósito de estos métodos es transferir el modelo de calibrado construido en una situación A de forma que pueda ser utilizado con los datos de una nueva situación (B). El procedimiento puede efectuarse transfiriendo el propio modelo de calibrado, o también transformando el conjunto de aprendizaje de la situación primera al formato que hubieran tenido de haber sido medidos en la nueva situación y construyendo después un nuevo modelo a partir de los datos de la situación inicial transformados. En ambas aproximaciones los modelos de transferencia se construyen en base a las medidas de un pequeño conjunto de muestras patrón de las dos situaciones.

III.2.1.1 Transferencia de calibrado utilizando un modelo de calibrado clásico

Sean \bar{Y} e \bar{Y} las matrices del parámetro que se desea predecir (concentración) del conjunto de calibración completo y del subconjunto de transferencia respectivamente, cuyas dimensiones son el número de muestras por el número de analitos. La relación lineal entre las variables predictoras y la respuesta en la situación A viene dada por

$$\bar{X}_A = \bar{Y}K_A \quad (6)$$

pudiendo plantear un modelo análogo para la situación B de forma que

$$\bar{X}_B = \bar{Y}K_B = \bar{Y}(K_A + \Delta K) \quad (7)$$

donde K_A y K_B son las matrices de sensibilidad para todas las muestras en ambas situaciones y ΔK es la matriz diferencia. La misma formulación debería ser válida para los dos subconjuntos de transferencia,

$$\bar{X}_A = \bar{Y}K_A \quad (8)$$

$$\bar{X}_B = \bar{Y}K_B = \bar{Y}(K_A + \Delta K) \quad (9)$$

a partir de los cuales es posible, resolviendo (8) y (9), estimar ΔK

$$\Delta K = \bar{Y}^+(\bar{X}_B - \bar{X}_A) \quad (10)$$

donde \bar{Y}^+ es la matriz pseudoinversa por la derecha de \bar{Y} . Sustituyendo ΔK en la ecuación (7) y utilizando la ecuación (6), se estima \bar{X}_B como

$$\hat{\bar{X}}_B = \bar{X}_A + \bar{Y}\bar{Y}^+(\bar{X}_B - \bar{X}_A) \quad (11)$$

A partir de $\hat{\bar{X}}_B$ e \bar{Y} se puede construir un nuevo modelo de calibrado que permita hacer estimaciones en la situación B, utilizando para ello cualquier método de calibración multivariante. El resultado de este procedimiento es una estimación de cómo sería el modelo de la situación B, utilizando los datos obtenidos en la situación A, que se transforman previamente de forma adecuada, para construir dicho modelo. El método se muestra de forma esquemática en la Figura III.2.1.

Cabe destacar que en este método se usa explícitamente la matriz de concentraciones del subconjunto de transferencia, de donde se infiere la necesidad de conocer las concentraciones de todos los analitos implicados. Además, el método asume que la misma relación lineal entre las variables predictoras y la variable respuesta se mantiene en ambas situaciones.

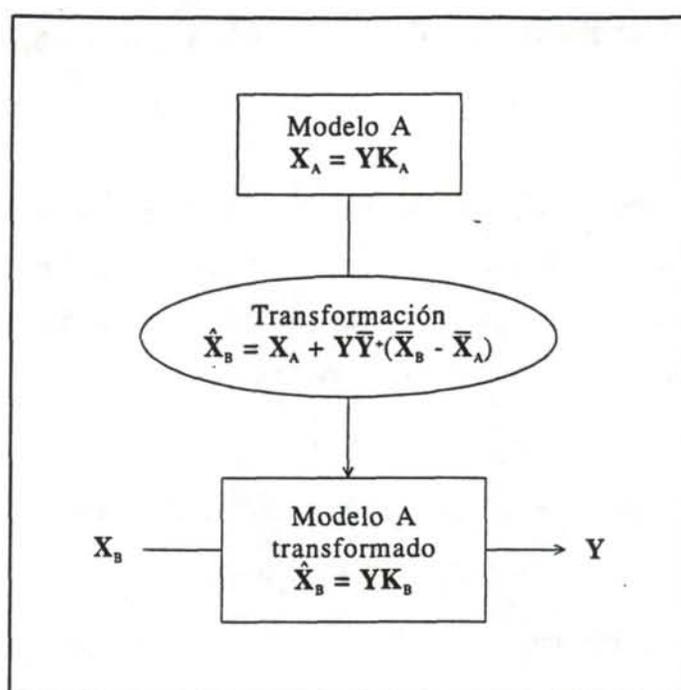


Figura III.2.1 Representación esquemática del método de transferencia de calibrado que utiliza el modelo de calibrado clásico.

III.2.1.2 Transferencia de calibrado utilizando un modelo de calibrado inverso

Este método, en lugar de utilizar el modelo de calibrado clásico, recurre al modelo de calibrado inverso correspondiente a la ecuación

$$y = X_A b_A \quad (12)$$

y de forma similar para la situación B se tiene que

$$y = X_B b_B = X_B (b_A + \Delta b) \quad (13)$$

donde b_A y b_B son los coeficientes de regresión para un analito en ambas situaciones, y Δb es el vector diferencia entre ambos. Ecuaciones análogas pueden plantearse para un subconjunto de transferencia

$$\bar{y} = \bar{X}_A \bar{b}_A \quad (14)$$

$$\bar{y} = \bar{X}_B \bar{b}_B = \bar{X}_B (\bar{b}_A + \Delta b) \quad (15)$$

donde \bar{b}_A y \bar{b}_B son los correspondientes coeficientes de regresión, e \bar{y} representa el vector concentración (u otra propiedad) para ese analito.

Utilizando la ecuación (12) para calcular b_A y combinando las ecuaciones (14) y (15) para estimar Δb , el coeficiente de regresión de transferencia puede ser estimado mediante

$$\begin{aligned}\hat{b}_B &= (b_A + \Delta b) = b_A + (\bar{b}_B - \bar{b}_A) = \\ &= \mathbf{X}_A^+ \bar{y} + (\bar{\mathbf{X}}_B^+ - \bar{\mathbf{X}}_A^+) \bar{y}\end{aligned}\quad (16)$$

Este método ofrece directamente como resultado los coeficientes de regresión \hat{b}_B para la situación B (transferencia de modelo real), los cuales permiten predecir la concentración u otra propiedad cualquiera para las muestras analizadas en esa situación. A diferencia de la transferencia de calibrado que utiliza un modelo clásico, en este caso únicamente es necesario conocer la concentración del analito de interés. El método puede hacerse extensible a problemas con más de un analito simplemente tomando y e \bar{y} en forma de matrices y utilizando una matriz de coeficientes de regresión.

III.2.1.3 Transferencia de calibrado global (GCT)

El método de transferencia de calibrado global (global calibration transfer, GCT), ha sido propuesto por Forina y otros [102] para la transferencia de la ecuación de regresión de un instrumento a otro en espectroscopía NIR. Este método utiliza la regresión PLS en primer lugar para relacionar las variables predictoras de las situaciones A y B con la respuesta

$$y = \mathbf{X}_A b_A \quad (17)$$

$$y = \mathbf{X}_B b_B \quad (18)$$

mientras que para un subconjunto de transferencia dado se tiene

$$\bar{y} = \bar{\mathbf{X}}_A b_A \quad (19)$$

$$\bar{y} = \bar{\mathbf{X}}_B b_B \quad (20)$$

A continuación se establece una relación entre las variables predictoras correspondientes a esas dos situaciones, A y B, de nuevo mediante una regresión PLS y a través de una matriz de transferencia \mathbf{F} , es decir,

$$\bar{\mathbf{X}}_A = \bar{\mathbf{X}}_B \mathbf{F} \quad (21)$$

donde F es una matriz cuadrada de dimensiones $p \times p$, siendo p el número de variables predictoras. Esta última relación permite reescribir la ecuación (19) de modo que

$$\bar{y} = \bar{X}_A b_A = \bar{X}_B F b_A \quad (22)$$

y al comparar las ecuaciones (20) y (22) se tiene una estimación de los coeficientes del modelo de regresión para la situación B dada por

$$\hat{b}_B = F b_A \quad (23)$$

En este procedimiento no se requiere que el número de variables predictoras en ambas situaciones sea el mismo y, por supuesto, es posible utilizar los datos originales o las puntuaciones correspondientes (scores) obtenidas al transformar las variables originales, por ejemplo, mediante componentes principales.

III.2.2 Métodos que transforman las nuevas medidas

En este caso, a los datos analizados en la nueva situación se les aplica una transformación tal que les hace ser equivalentes a haber sido medidos en la situación inicial, y por tanto las predicciones pueden obtenerse utilizando el modelo de calibrado antiguo, tal como se muestra en la Figura III.2.2. Es decir, se transforman los datos de la nueva situación de tal manera que pueda ser aplicado sobre ellos el modelo de calibrado construido con anterioridad.

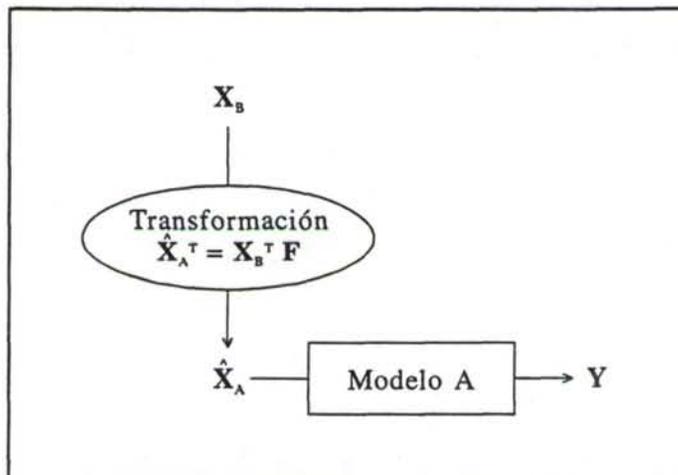


Figura III.2.2 Representación esquemática de aquellos procedimientos de transferencia de calibrado basados en la corrección de los datos medidos en una nueva situación.

El modelo de transferencia que se utiliza en la corrección de los datos de la situación B está basado en una pequeña serie de medidas llevadas a cabo en ambas situaciones, a partir de las cuales se calcula una matriz denominada de transferencia. Estas medidas normalmente son un subconjunto del conjunto de aprendizaje al que se asigna el nombre de conjunto de transferencia.

Estos métodos presentan la ventaja de que no es necesario disponer de los valores de la variable, por lo que también se podrían utilizar muestras patrón para llevar a cabo la transformación de los datos.

Dependiendo de cómo se realice el cálculo de la matriz de transferencia se desarrollan diferentes métodos de transferencia, los cuales se analizan a continuación.

III.2.2.1 Transferencia de calibrado directa (DS)

En la transferencia de calibrado directa, conocida como direct standardization (DS), se establece una relación entre las matrices de las variables predictoras medidas en dos situaciones distintas, A y B, por medio de una matriz de transferencia F , tal y como se indica en la ecuación (21). A partir de dicha ecuación la matriz de transferencia se puede calcular de acuerdo con la siguiente expresión

$$F = \overline{X_B}^+ \overline{X_A} \quad (24)$$

donde $\overline{X_B}^+$ es la pseudoinversa de $\overline{X_B}$, la cual se determina normalmente mediante PCR o PLS.

Utilizando la matriz de transferencia F , las variables predictoras de una nueva muestra medida en la situación B (vector fila x_B^T) pueden ser transformadas al formato esperado para esa muestra en la situación A mediante:

$$\hat{x}_A^T = x_B^T F \quad (25)$$

y entonces las predicciones de los valores de concentración pueden hacerse utilizando el modelo original.

Wang y otros [94] destacan el hecho de que la matriz F se calcula partiendo de un número relativamente pequeño de muestras a través de la ecuación (21) y que esto conduce a que se produzca una mala estimación de F . Sugieren calcular la matriz F utilizando en el procedimiento de transferencia de calibrado las proyecciones de las filas de datos de las

matrices \bar{X}_A , \bar{X}_A y \bar{X}_B sobre un modelo de descomposición en valores singulares de X_A . De esta manera, los scores obtenidos se utilizan en lugar de las matrices de datos originales para llevar a cabo la transferencia de calibrado y la posterior calibración. No obstante, aunque este procedimiento tiene como resultado una reducción de las dimensiones del problema y un filtrado del ruido, también puede conducir a una pérdida de información, sobre todo si las primeras componentes de A y B no describen los mismos fenómenos.

El número de muestras patrón del conjunto de transferencia ha de ser al menos igual al rango de la matriz X_A , para que éstas sean representativas de todas las dimensiones relevantes, aunque no es necesario que el número de datos de las dos situaciones, A y B, sea el mismo.

Una desventaja que presenta el procedimiento de transferencia de calibrado directa (DS) reside en que cada una de las variables transformadas se calcula utilizando todas las variables predictoras, lo que lleva asociado el riesgo de que se produzca un sobreajuste. La estimación de F mediante la ecuación (24) es problemática ya que el número de variables predictoras es muy superior al número de muestras patrón y, en consecuencia, el cálculo de X_B^+ está numéricamente mal condicionado. Estas consideraciones han conducido al desarrollo de un método de transferencia de calibrado directa mediante una aproximación a trozos.

III.2.2.2 Transferencia de calibrado directa a trozos (PDS)

La transferencia de calibrado directa a trozos, piecewise direct standardization (PDS), es un método desarrollado por Wang y otros [94] como una mejora del procedimiento DS descrito anteriormente. Se basa en establecer una relación entre un punto experimental de una situación determinada con una pequeña región de variables próximas a ese mismo punto en otra situación diferente. El método PDS se sirve de una ventana móvil para que cada variable predictora de la situación A sea relacionada con las variables predictoras de la situación B que se encuentren dentro de dicha ventana, próximas a ella.

La ventana Z_i utilizada para la transformación de la variable x_i viene dada por

$$Z_i = [x_{B,i-j}, x_{B,i-j+1}, \dots, x_{B,i+k-1}, x_{B,i+k}] \quad (26)$$

donde la ventana recorre las variables predictoras desde $i-j$ a $i+k$, sin ser necesariamente simétrica alrededor de i . El tamaño óptimo de ventana se puede determinar a partir de los

datos experimentales obtenidos en ambas situaciones, comparando la anchura y la posición de los picos.

Cuando exista un desplazamiento de la señal y una variación lineal de la respuesta experimental de una situación a otra, se tendrá que

$$x_{A,i} \propto x_{B,i} \quad (27)$$

con $i-j \leq l \leq i+k$. Para el caso en que l se encuentre entre dos puntos en los que se produce una variación no lineal en la respuesta experimental, esta relación no se mantendría. Sin embargo, una regresión multivariante local de la forma

$$x_{A,i} = \mathbf{Z}_i \mathbf{b}_i \quad (28)$$

puede llevar a cabo la interpolación y proporcionar una aproximación razonable a la variación no lineal de la respuesta experimental. El que el rango de la matriz \mathbf{Z}_i sea menor que el rango de \mathbf{X}_A da opción a reducir el número de muestras y los problemas de mal condicionamiento.

Cada coeficiente de regresión \mathbf{b}_i puede ser calculado utilizando algún método de calibración multivariante, tal como PCR o PLS. A continuación, todos los vectores \mathbf{b}_i de la regresión se sitúan en la diagonal principal de la matriz de transferencia \mathbf{F} , y se completa con ceros el resto de los elementos de la matriz, obteniéndose como resultado una matriz banda diagonal, tal como se ilustra en la Figura III.2.3.

Comparado con DS, el método PDS también se basa en el cálculo de la matriz de transferencia \mathbf{F} , ecuación (21), pero en este caso se colocan ceros en los elementos de la matriz que se encuentran fuera de una ventana que se desliza a lo largo de la diagonal principal, es decir,

$$\mathbf{F} = \text{diag}(\mathbf{b}_1^T, \mathbf{b}_2^T, \dots, \mathbf{b}_i^T, \dots, \mathbf{b}_p^T) \quad (29)$$

donde p es el número de variables predictoras medidas. Para las nuevas muestras la transformación se lleva a cabo multiplicando la respuesta experimental por la matriz de transferencia tal como se indica en la ecuación (25). Cada variable predictora x_i de la situación A se obtiene al multiplicar las variables predictoras de una ventana \mathbf{Z}_i de la situación B por la matriz banda diagonal \mathbf{F} o matriz de transferencia.

Para algunas de las variables predictoras situadas en los extremos de la señal experimental de la situación A no es factible la construcción de ventanas de determinados tamaños con los datos de la situación B y, aunque estas podrían ser estimadas por extrapolación, es preferible descartarlas.

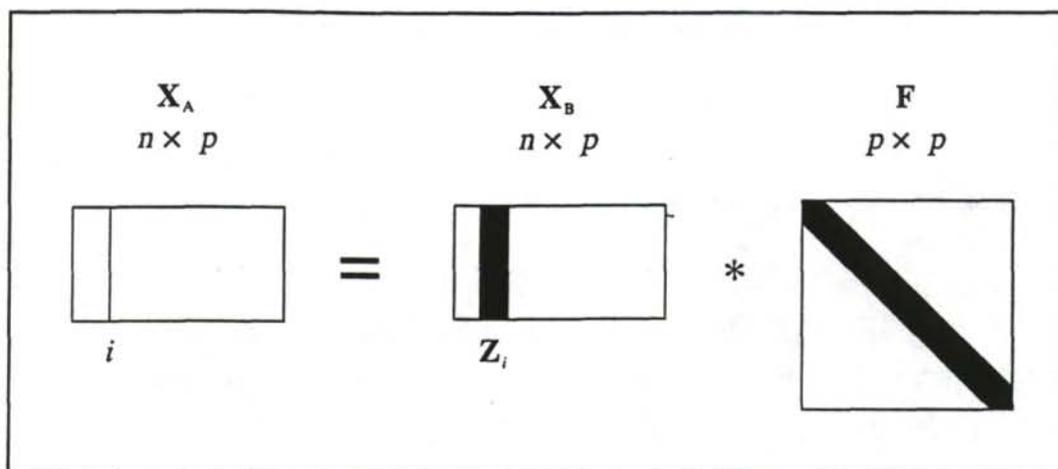


Figura III.2.3 Cada variable x_i de A resulta de multiplicar las var. de una ventana Z_i de B por la matriz banda diagonal F.

En el trabajo presentado por Wang y otros [94], el método PDS fue el que mejores resultados proporcionó en comparación con otros cuatro métodos utilizados para llevar a cabo la transferencia de calibrado en muestras de gasolina analizadas por NIR con dos espectrofotómetros diferentes.

Dado que este tipo de análisis puede ser extrapolable a datos de tipo electroquímico, este procedimiento de transferencia de calibrado, junto con el método GCT anteriormente especificado, han sido utilizados en la presente memoria para llevar a cabo transferencias de calibrado en diferentes determinaciones voltamperométricas multianalito.

III.3 DETERMINACIÓN SIMULTÁNEA DE Cu, Pb, Cd Y Zn MEDIANTE POLAROGRAFÍA

Existe un amplio grupo de elementos, denominados metales pesados, que tienen gran importancia industrial, biológica y medioambiental [109-111]. Cuando estos elementos se encuentran presentes en exceso resultan tóxicos para los organismos vivos, sin embargo, algunos de ellos, tales como Co, Cr, Cu, Mn, Mo y Zn resultan esenciales en concentraciones pequeñas pero críticas para un crecimiento normal de plantas y animales, aun cuando son tóxicos a altas concentraciones.

El término 'metal pesado' se utiliza frecuentemente de forma peyorativa, dando énfasis al aspecto contaminante y tóxico, más que como un nombre general para un grupo de elementos de considerable importancia económica y medioambiental. Los metales pesados que suscitan gran interés por su relación con la salud, la agricultura y el medio ambiente son As, Cd, Hg, Pb, Tl y U. Por otro lado, se encuentran aquellos denominados 'metales pesados esenciales' o 'micronutrientes' que, en cantidades muy pequeñas, forman parte de procesos metabólicos fundamentales en animales y plantas. Dentro de este último grupo se consideran incluidos Fe, Mn, Zn, Cu, Mo, Ni, B, Na y Cl.

Las técnicas instrumentales requeridas para llevar a cabo el análisis de estos elementos pueden clasificarse, de forma global, en métodos que permiten la determinación de un único analito por medida, como es el caso de la espectroscopía de absorción atómica (AAS), o métodos para la determinación simultánea de varios analitos, como la espectroscopía de emisión con plasma acoplado inductivamente (ICP-AES), la espectroscopía de rayos X, etc.

Para realizar un análisis determinado, en la elección de uno u otro método se debe tener en consideración, además del factor indicado, otros como la sensibilidad, la selectividad, la precisión y la exactitud del mismo. Así mismo, sería ventajoso que el método en su totalidad pudiera ser fácilmente automatizado, que fuera susceptible de ser utilizado en medidas de campo (para análisis medioambiental,...), aplicable a un amplio rango de

sustancias, etc. No obstante, en la selección del método pueden prevalecer los costes relativos (principalmente instrumentales) de las técnicas utilizadas.

Los métodos electroquímicos satisfacen completamente los requisitos señalados previamente, destacando principalmente por los bajos costes del instrumental necesario para su puesta a punto, así como por su alta sensibilidad y los bajos límites de detección que permiten alcanzar. Por todo ello, estos métodos son ampliamente utilizados en diferentes campos [112-115], tales como aplicaciones industriales, medioambientales, clínicas, farmacéuticas, etc. Diversas técnicas electroquímicas son aplicadas de forma rutinaria en la determinación de metales pesados [116-119], siendo posible a través de éstas incluso la especiación de los mismos [120].

Cuando las concentraciones de estos metales se encuentran por encima de 1 ppm, éstos pueden ser determinados utilizando alguno de los diversos métodos polarográficos disponibles en la actualidad. Entre las técnicas polarográficas, una de las más utilizadas es la polarografía diferencial de impulsos [121], ya que ofrece una alta discriminación de la corriente de carga y presenta una mejora en la relación señal-ruido.

Esta última técnica electroquímica ha sido la utilizada para llevar a cabo la determinación simultánea de cobre, plomo, cadmio y zinc en muestras acuosas. En el citado análisis se han desarrollado modelos de calibrado multivariante PLS, que posteriormente han sido transferidos a lo largo del tiempo, utilizándose para ello dos procedimientos de transferencia de calibrado distintos, los métodos PDS y GCT anteriormente señalados.

III.3.1 Experimental

Reactivos

Los reactivos empleados en el análisis (calidad p.a.) fueron utilizados sin purificaciones posteriores. Las disoluciones patrón fueron preparadas utilizando agua desionizada obtenida mediante un sistema Barnstead NANO Pure II. Para eliminar el oxígeno disuelto en las disoluciones se utilizó nitrógeno (99.997%) como gas inerte. El electrolito soporte, que presentaba un valor de pH igual a 4.335, estaba constituido por una disolución tampón de ácido acético 2 M e hidróxido amónico 1 M.

Instrumental

Las medidas polarográficas fueron realizadas utilizando un procesador Metrohm 646 VA y una unidad 647 VA, combinados con un electrodo multimodo Metrohm (MME) en el modo electrodo de gota de mercurio estática (static mercury drop electrode, SMDE). El sistema de tres electrodos se completó mediante un electrodo auxiliar de platino y un electrodo de referencia Ag/AgCl/KCl (3M).

Los análisis de datos fueron realizados utilizando los programas PARVUS [57], STATGRAPHICS [58] y MATLAB [59]. Todos los cálculos fueron llevados a cabo en un ordenador con microprocesador 486/66.

Procedimiento experimental

Las medidas polarográficas fueron obtenidas mediante el procedimiento siguiente. La disolución fue introducida en una celda polarográfica y desoxigenada burbujeando nitrógeno durante 10 minutos. Una vez que la disolución había sido desoxigenada, se registró un polarograma en un rango de potenciales desde 0.095 V hasta -1.117 V. Tras cada adición, la solución fue agitada y desoxigenada de nuevo durante 15 s. Todos los resultados presentados fueron obtenidos mediante polarografía diferencial de impulsos, tomando los parámetros experimentales los siguientes valores: amplitud del impulso -50 mV, tiempo de vida de la gota 0.6 s., área del electrodo 0.40 mm² y velocidad de barrido -10 mV s⁻¹.

III.3.2 Diseño del experimento

Las técnicas electroanalíticas son de uso común cuando se trata de determinar varios analitos en una misma muestra, siendo variable el número de experiencias necesarias para llevar a cabo un calibrado multicomponente. Si el número de muestras patrón que constituyen el calibrado es elevado, esto lleva asociado un tiempo considerable en la realización de la experiencia. A este respecto, debe de evaluarse el tiempo necesario para preparar las propias muestras, para la limpieza de la celda electroquímica y del material para cada nueva disolución, el tiempo de desoxigenación de la disolución y el tiempo de barrido, etc. Esto hace necesaria una herramienta que permita optimizar tanto el número de experiencias a

realizar, como el orden en que éstas deben sucederse para que la información que se pueda extraer posteriormente a partir de la señal analítica registrada sea lo más amplia y útil posible, habiendo realizado un esfuerzo experimental razonable.

Por ello, se ha optado por utilizar un experimento diseñado en el que las muestras del calibrado se distribuyen dependiendo de la concentración de cada uno de los analitos según un diseño en el que se intentan aprovechar las buenas cualidades de un diseño central compuesto [122, 123], tal y como se refleja en la Tabla III.3.1. La primera columna de dicha tabla indica el número asignado a cada muestra, mientras que en la segunda aparecen marcadas con un símbolo común aquellas muestras que resultaron de realizar adiciones sobre una única disolución inicial que se tomó como blanco.

La preparación de las muestras se realizó considerando el número mínimo posible de disoluciones de partida, y la obtención del resto de soluciones se llevó a cabo mediante la adición de las disoluciones patrón correspondientes. Este aspecto se tuvo en consideración ya que suponía reducir considerablemente el tiempo de experimentación, teniendo en cuenta que cada disolución de partida necesita 10 minutos de desoxigenación además del tiempo de limpieza y secado del material y de la celda electroquímica.

La distribución en el espacio de los puntos experimentales del diseño central compuesto se muestra en la Figura III.3.1, en la cual aparecen representados únicamente tres de los cuatro analitos. El conjunto de todos los puntos experimentales representa un hipercubo de cuatro dimensiones, en el que cada par de caras paralelas es atravesado por un eje perpendicular a ellas sobre el que se encuentran situados cinco puntos experimentales igualmente espaciados. Este diseño cubre de una forma bastante completa el recinto experimental formado por las concentraciones de los cuatro analitos considerados, de hecho, se recorren cinco niveles de concentración diferentes para cada uno de los analitos.

Un diseño se denomina central compuesto cuando se obtiene como resultado de la yuxtaposición de un diseño factorial (en este caso un diseño 2^4) con otros dos niveles más por eje y un punto central. El punto central del diseño es habitualmente replicado, concretamente se dispone de cuatro réplicas del mismo, para estimar el error experimental puro. En este análisis no se ha utilizado un diseño clásico central compuesto, ni ortogonal rotatable, sino que se han adicionado siempre cantidades proporcionales de los reactivos, según el diseño indicado, con la finalidad de facilitar la automatización del procedimiento experimental (adición de los reactivos desde una bureta automática).

Tabla III.3.1 Diseño experimental utilizado en los calibrados A y B. Los símbolos de la segunda columna señalan la relación de las muestras con las disoluciones iniciales utilizadas.

Muestra	Diseño experimental				Concentraciones (μM)			
	Cobre	Plomo	Cadmio	Zinc	Cobre	Plomo	Cadmio	Zinc
1 ×	2	2	2	2	3.876	3.876	6.077	3.876
2 ×	2	2	2	4	3.846	3.846	6.030	7.692
6 ▲	2	2	4	2	3.846	3.846	12.060	3.846
3 ×	2	2	4	4	3.817	3.817	11.968	7.634
9 ·	2	4	2	2	3.846	7.692	6.030	3.846
10 ·	2	4	2	4	3.817	7.634	5.984	7.634
7 ▲	2	4	4	2	3.817	7.634	11.968	3.817
4 ×	2	4	4	4	3.788	7.576	11.877	7.576
12 ○	4	2	2	2	7.692	3.846	6.030	3.846
13 ○	4	2	2	4	7.634	3.817	5.984	7.634
15 ■	4	2	4	2	7.634	3.817	11.968	3.817
14 ○	4	2	4	4	7.576	3.788	11.877	7.576
16 -	4	4	2	2	7.634	7.634	5.984	3.817
11 ·	4	4	2	4	7.576	7.576	5.939	7.576
8 ▲	4	4	4	2	7.576	7.576	11.877	3.788
5 ×	4	4	4	4	7.519	7.519	11.788	7.519
17 ▼	1	3	3	3	1.923	5.796	9.045	5.796
18 ▼	2	3	3	3	3.831	5.747	9.010	5.747
19 ▼	3	3	3	3	5.725	5.725	8.976	5.725
20 ▼	4	3	3	3	7.605	5.703	8.942	5.703
21 ▼	5	3	3	3	9.470	5.682	8.908	5.682
22 □	3	1	3	3	5.796	1.923	9.045	5.796
23 □	3	2	3	3	5.747	3.831	9.010	5.747
24 □	3	3	3	3	5.725	5.725	8.976	5.725
25 □	3	4	3	3	5.703	7.605	8.942	5.703
26 □	3	5	3	3	5.682	9.470	8.908	5.682
27 ·	3	3	1	3	5.796	5.796	3.015	5.796
28 ·	3	3	2	3	5.747	5.747	6.007	5.747
29 ·	3	3	3	3	5.725	5.725	8.976	5.725
30 ·	3	3	4	3	5.703	5.703	11.922	5.703
31 ·	3	3	5	3	5.682	5.682	14.847	5.682
32 ▽	3	3	3	1	5.796	5.796	9.045	1.923
33 ▽	3	3	3	2	5.747	5.747	9.010	3.831
34 ▽	3	3	3	3	5.725	5.725	8.976	5.725
35 ▽	3	3	3	4	5.703	5.703	8.942	7.605
36 ▽	3	3	3	5	5.682	5.682	8.908	9.470

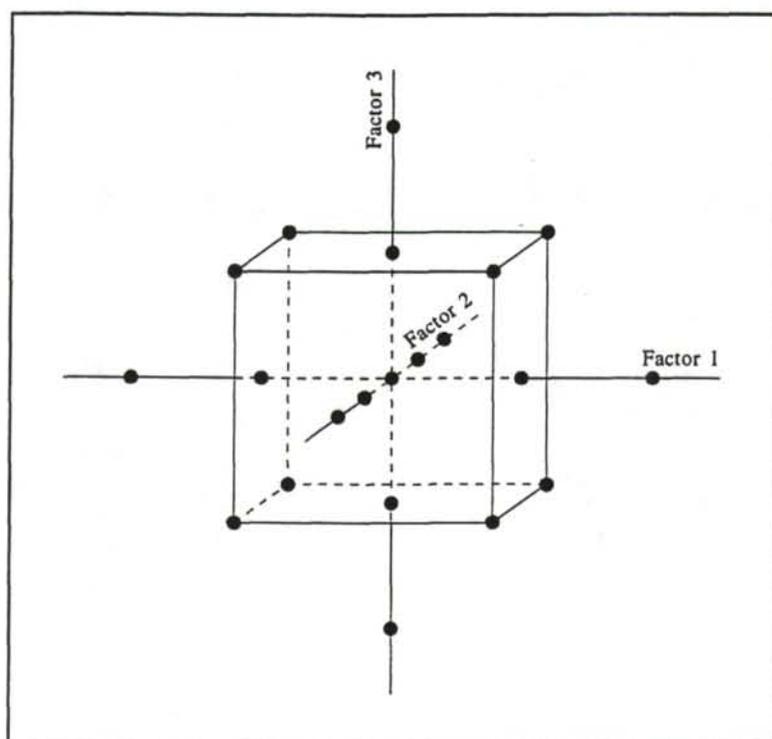


Figura III.3.1 Diseño experimental en un espacio de tres factores. Únicamente aparecen representados tres de los cuatro analitos que constituyen los diseños A y B.

Siguiendo este diseño se han realizado dos calibrados completos en diferentes días, con la finalidad de analizar la variabilidad existente entre ambos debido al efecto tiempo y a continuación llevar a cabo una transferencia de calibrado. Se dispone de dos conjuntos de datos, el calibrado A (realizado en primer lugar) y el calibrado B (realizado días después), cuyos polarogramas se muestran en las Figuras III.3.2a-j y III.3.3a-j respectivamente. Cada una de las figuras consta de diez gráficas que se corresponden con las diez soluciones iniciales (blancos experimentales) utilizadas en cada uno de los calibrados, sobre las que se han realizado distintas adiciones de cada analito según el diseño del experimento.

Comparando las gráficas que componen ambas figuras se observa que existen multitud de similitudes entre los polarogramas que corresponden a cada uno de los calibrados, pero que éstos no son idénticos. Esto confirma la hipótesis de que un mismo instrumento puede proporcionar respuestas experimentales más o menos distintas ante muestras iguales. Además de la variabilidad debida al propio instrumento, también se están recogiendo las eventuales variaciones existentes en las condiciones experimentales, las posibles diferencias entre las

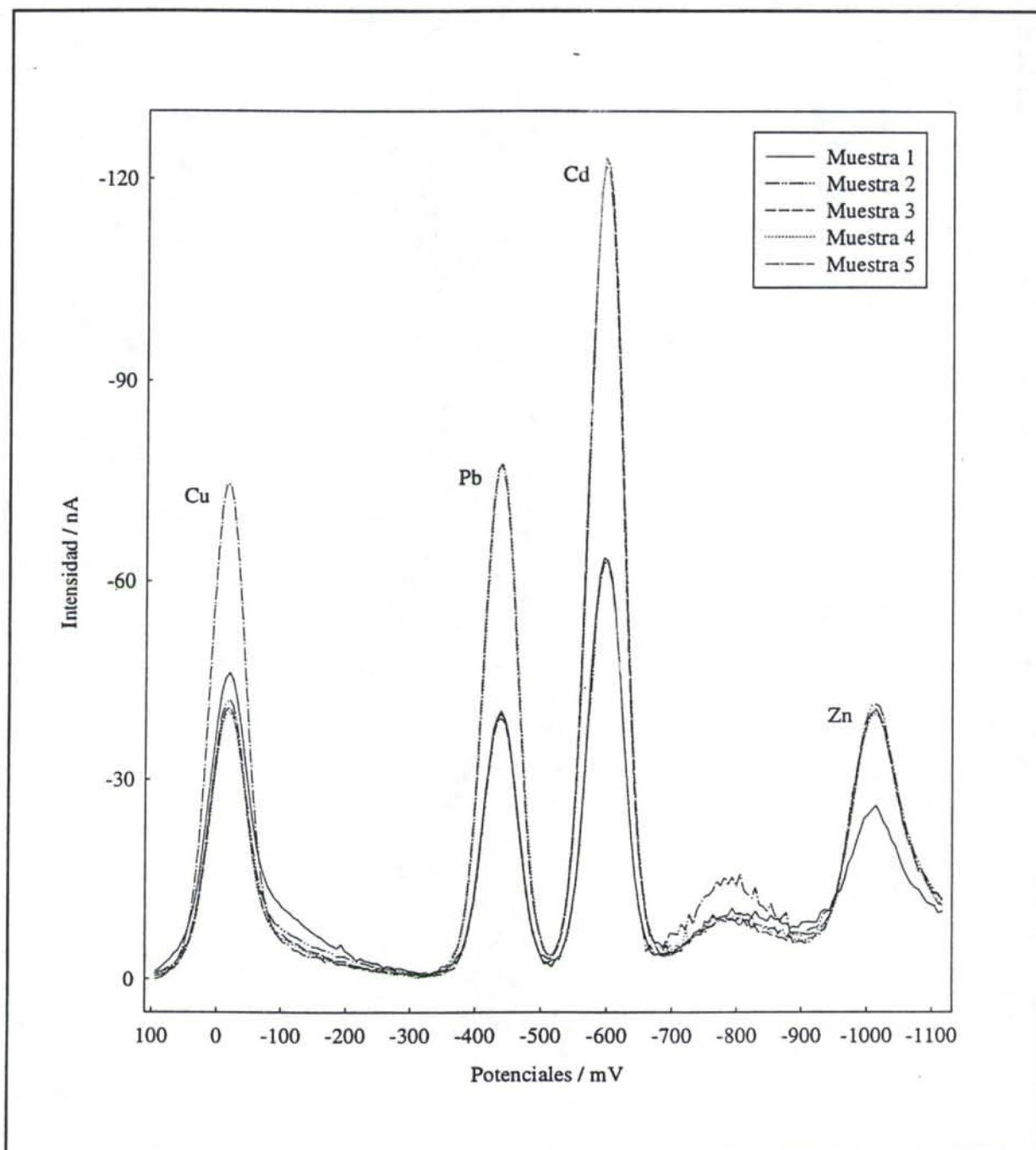


Figura III.3.2a Polarogramas del calibrado A correspondientes a las muestras especificadas en la Tabla III.3.1.

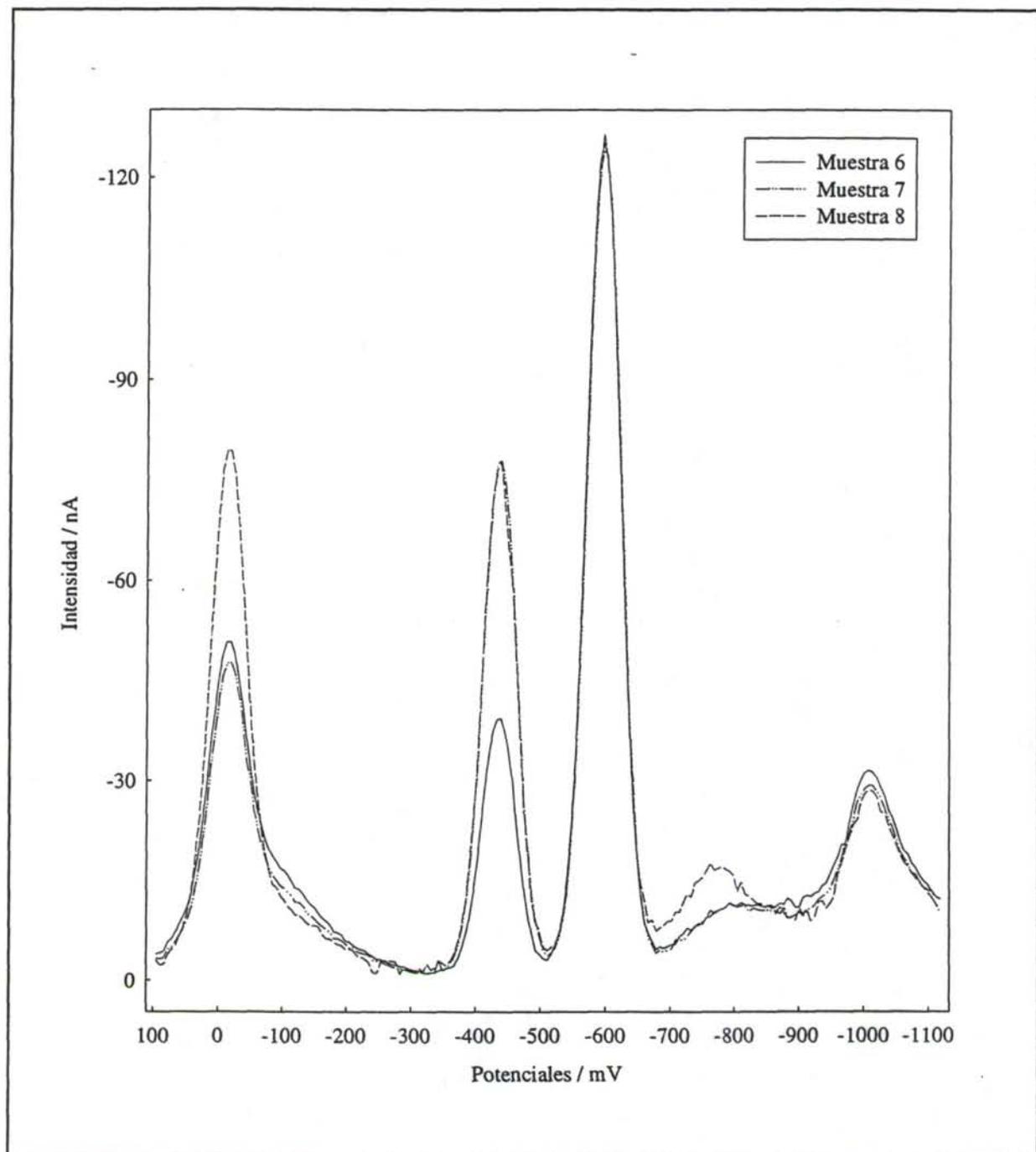


Figura III.3.2b Polarogramas del calibrado A correspondientes a las muestras especificadas en la Tabla III.3.1.

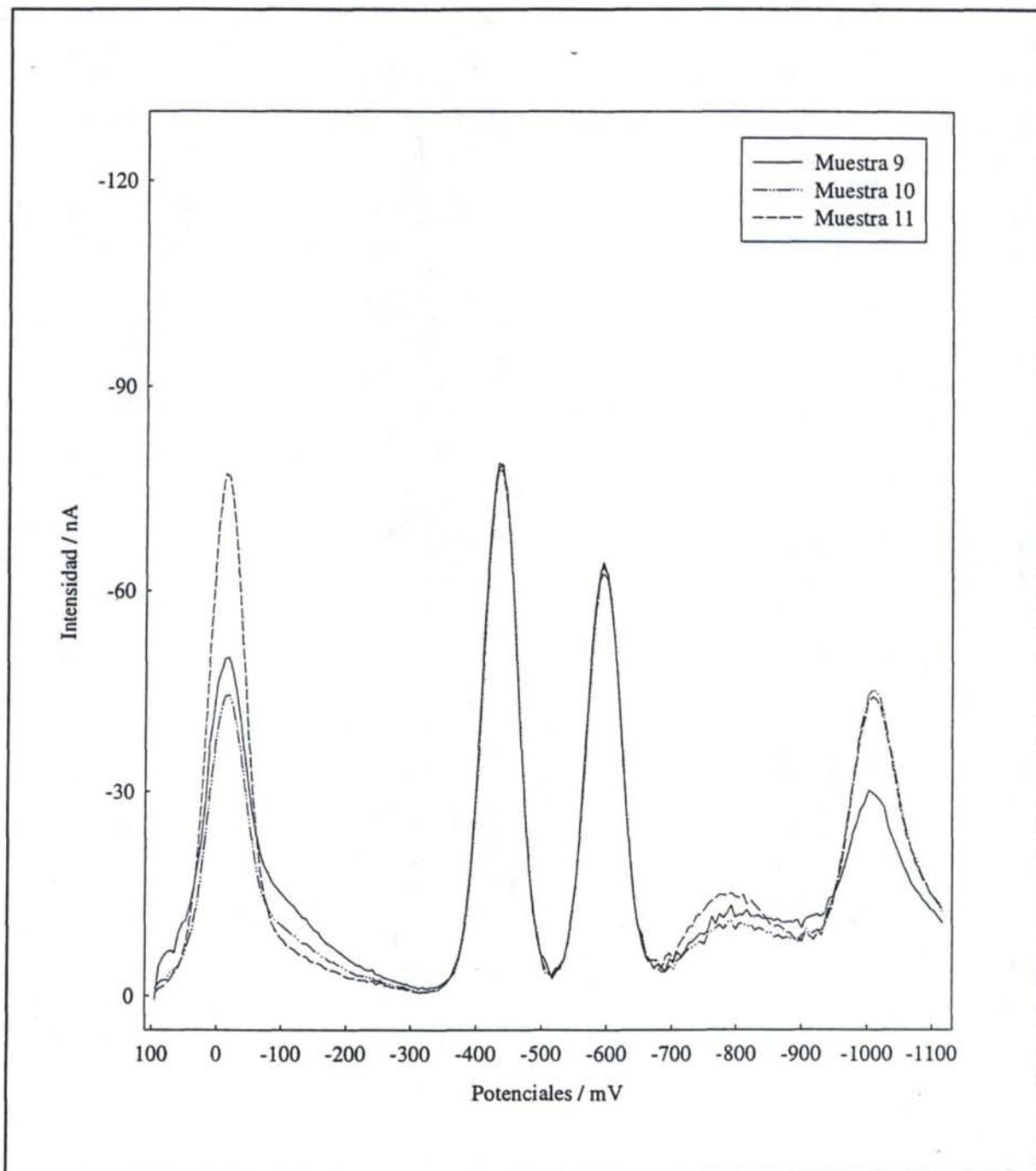


Figura III.3.2c Polarogramas del calibrado A correspondientes a las muestras especificadas en la Tabla III.3.1.

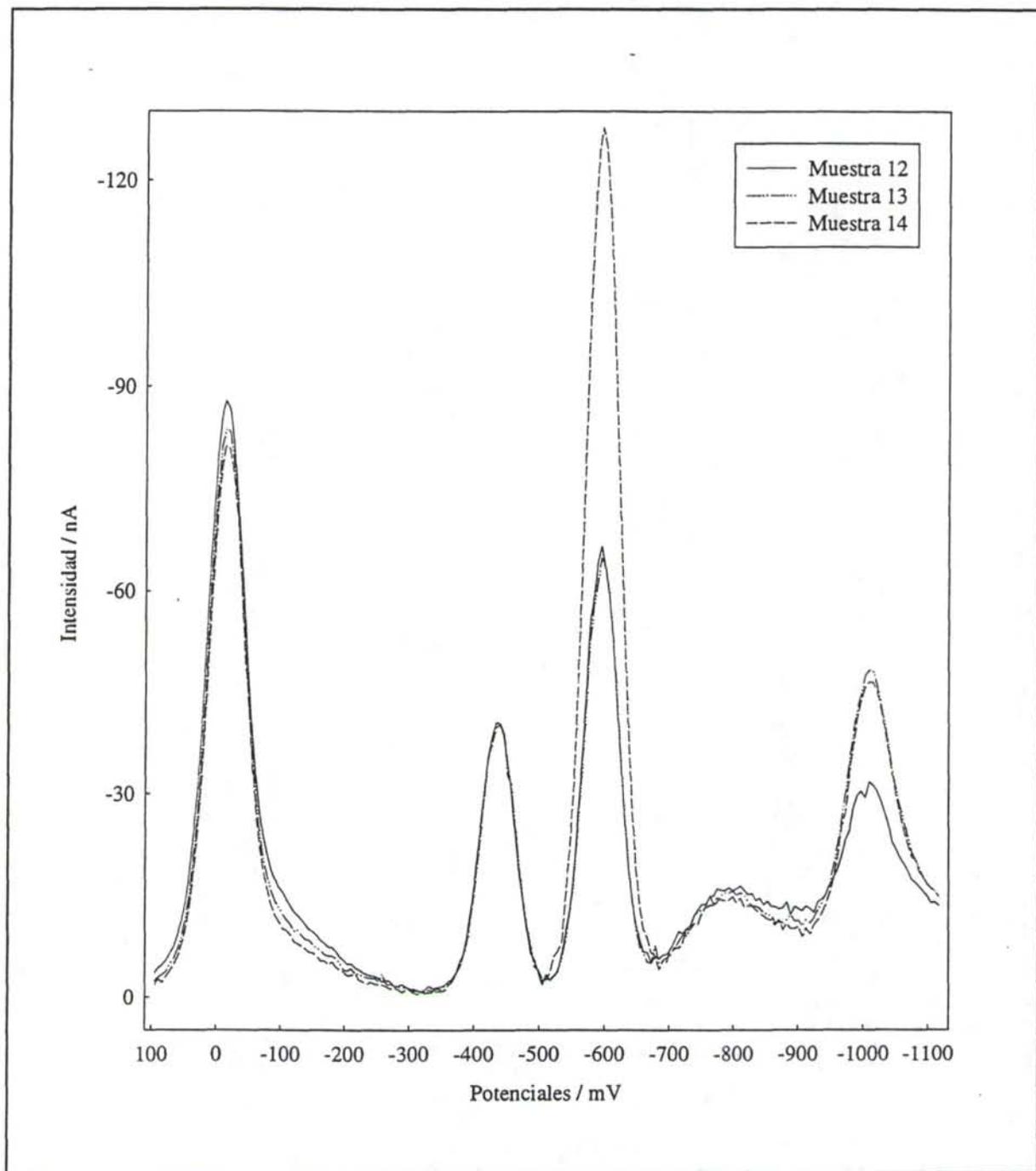


Figura III.3.2d Polarogramas del calibrado A correspondientes a las muestras especificadas en la Tabla III.3.1.

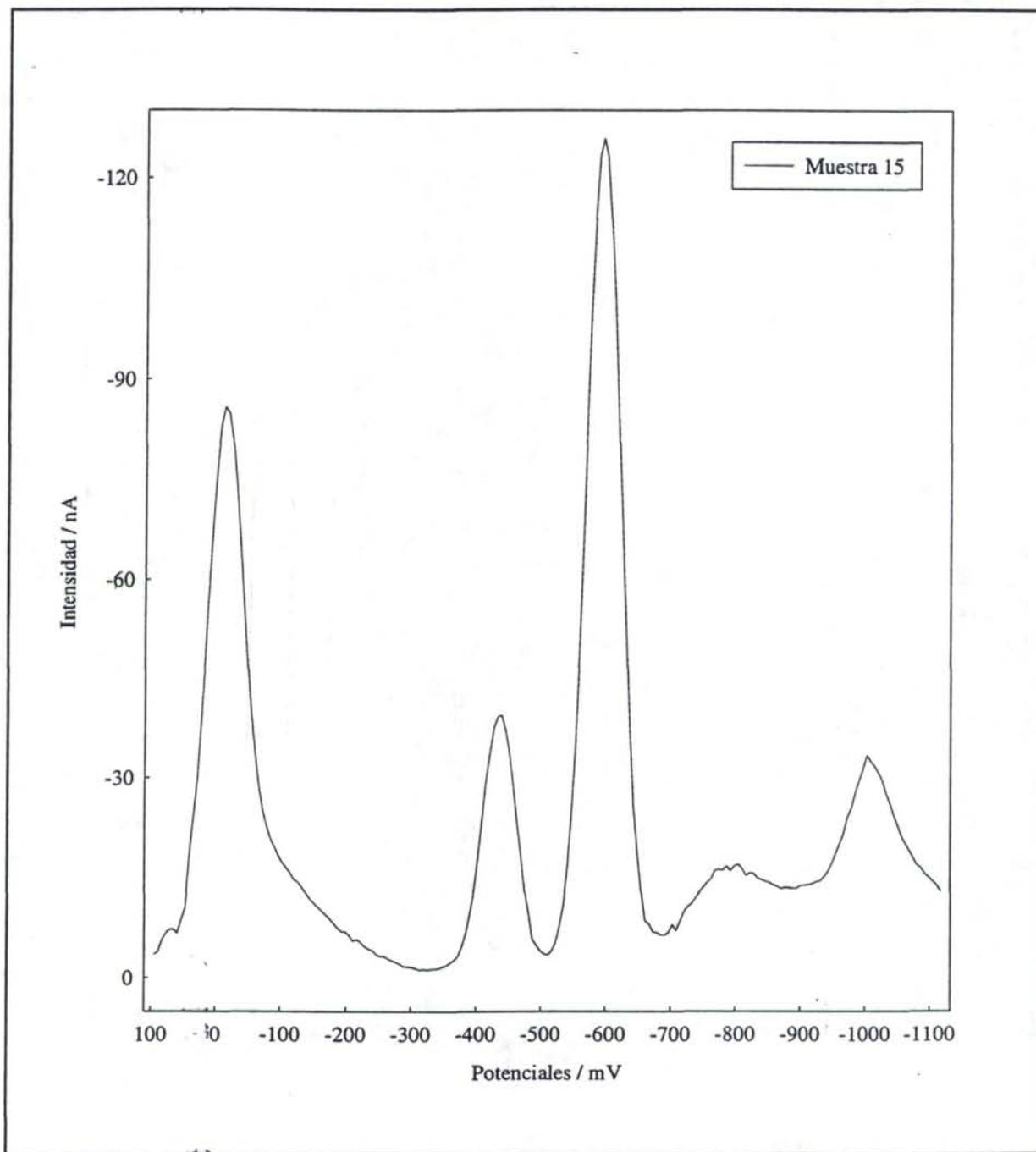


Figura III.3.2e Polarogramas del calibrado A correspondientes a las muestras especificadas en la Tabla III.3.1.

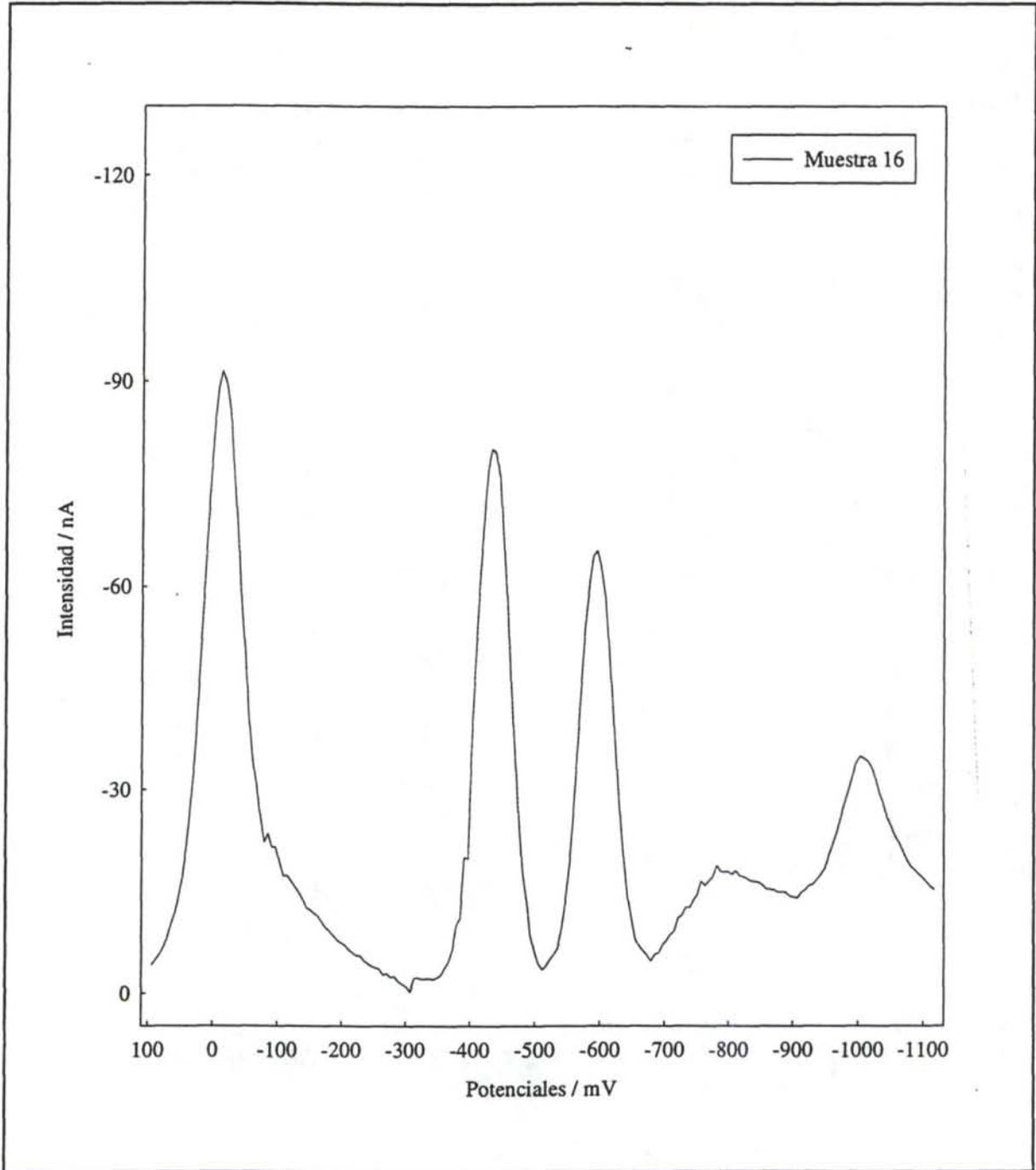


Figura III.3.2f Polarogramas del calibrado A correspondientes a las muestras especificadas en la Tabla III.3.1.

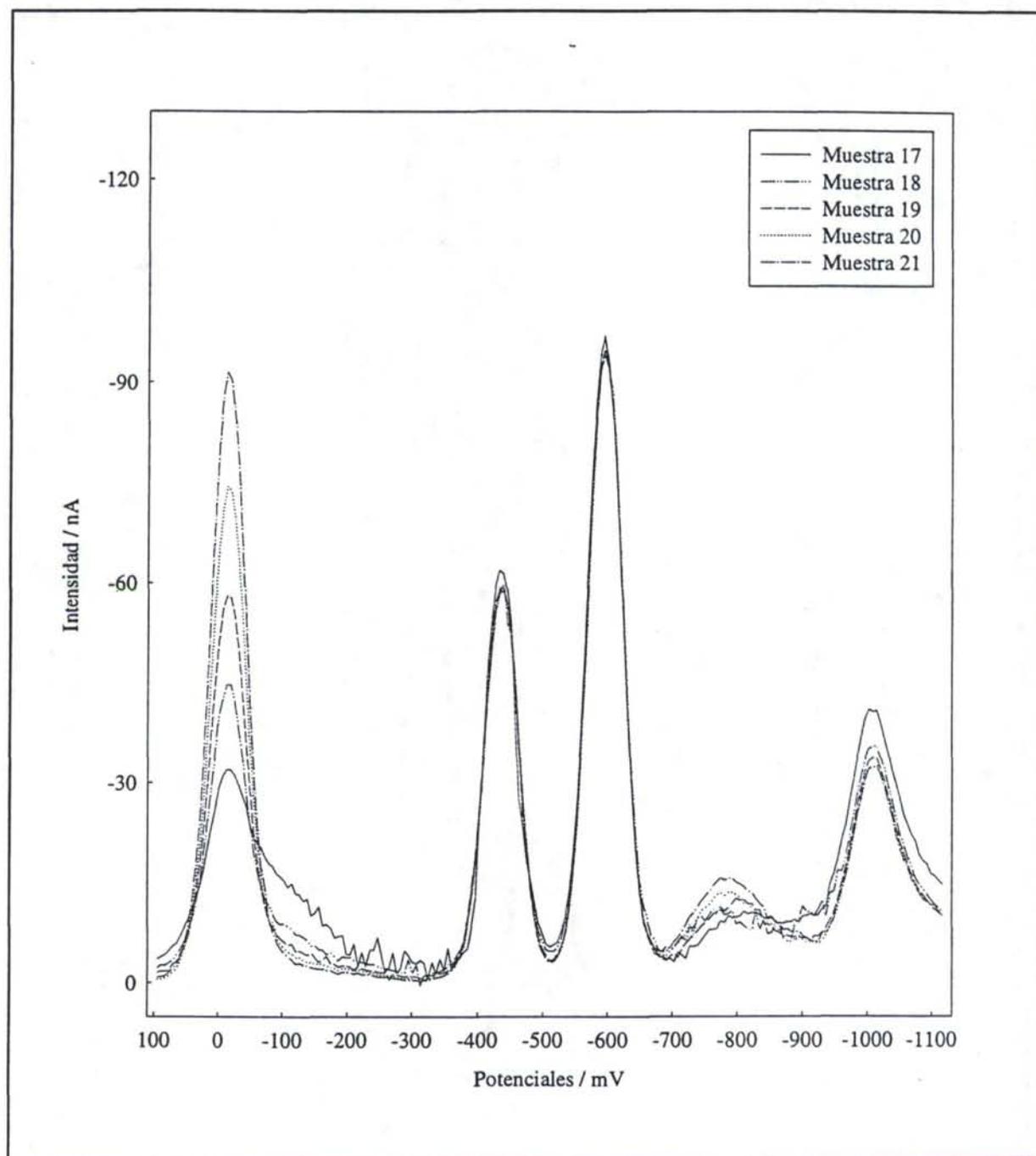


Figura III.3.2g Polarogramas del calibrado A correspondientes a las muestras especificadas en la Tabla III.3.1.

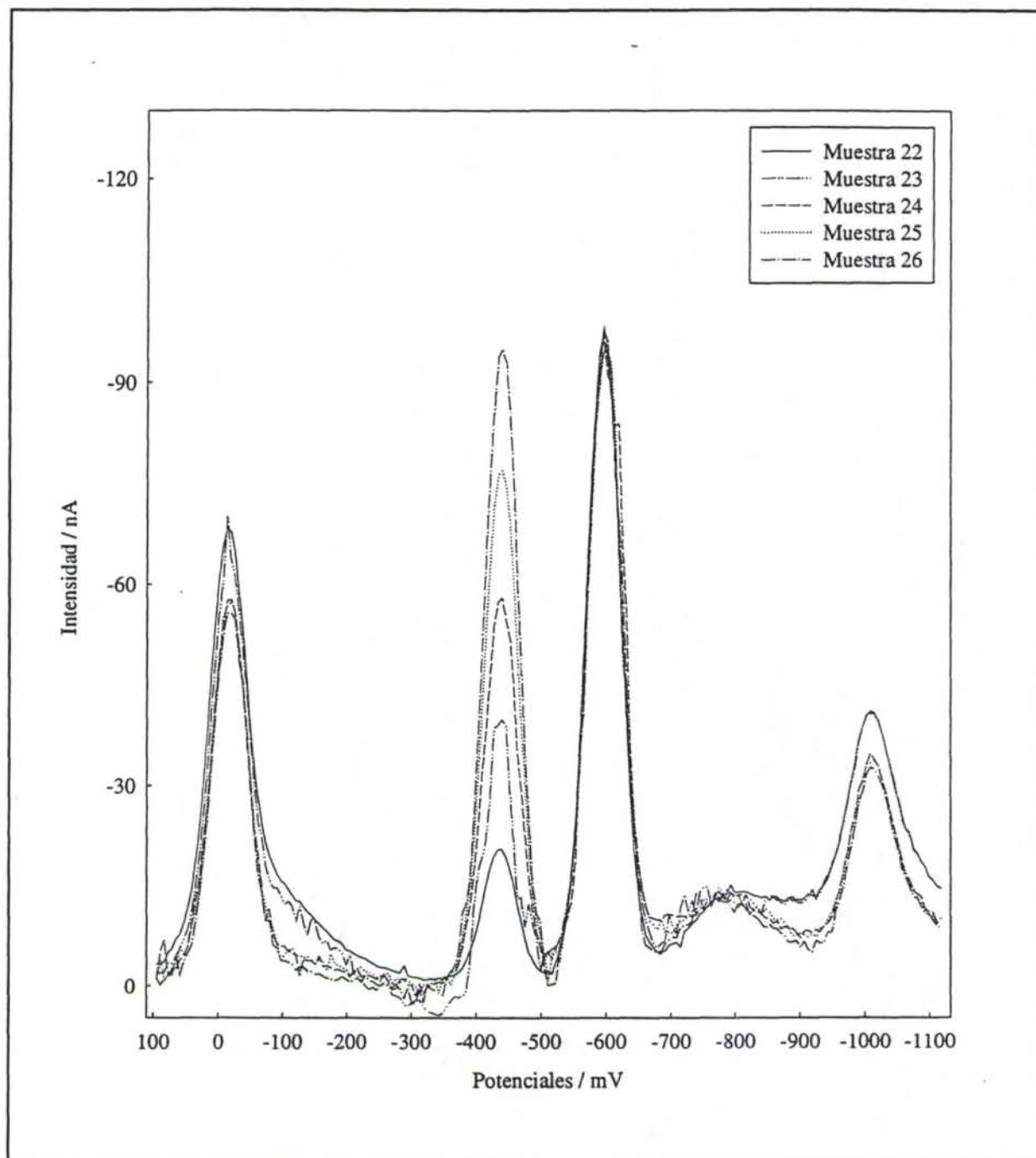


Figura III.3.2h Polarogramas del calibrado A correspondientes a las muestras especificadas en la Tabla III.3.1.

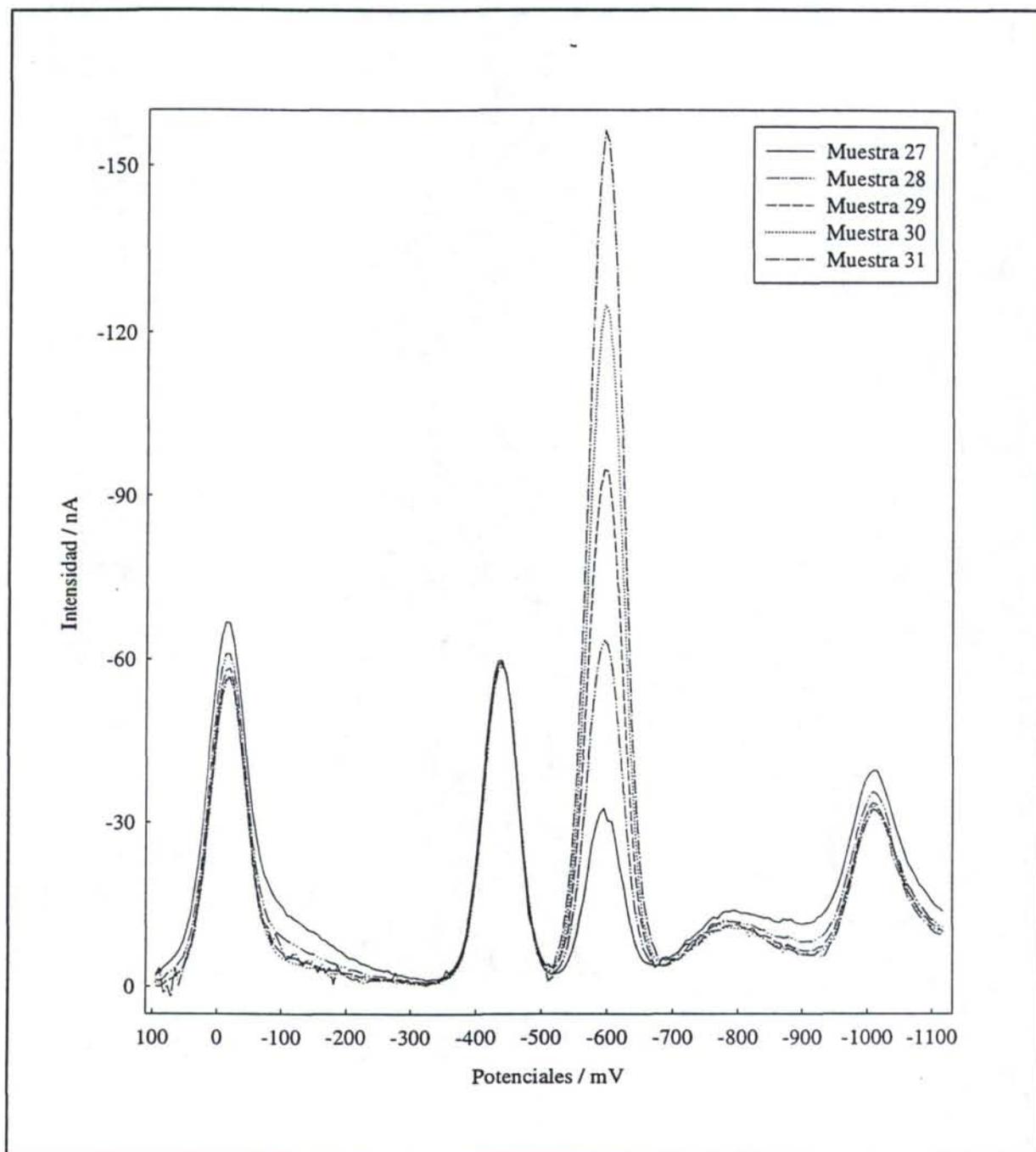


Figura III.3.2i Polarogramas del calibrado A correspondientes a las muestras especificadas en la Tabla III.3.1.

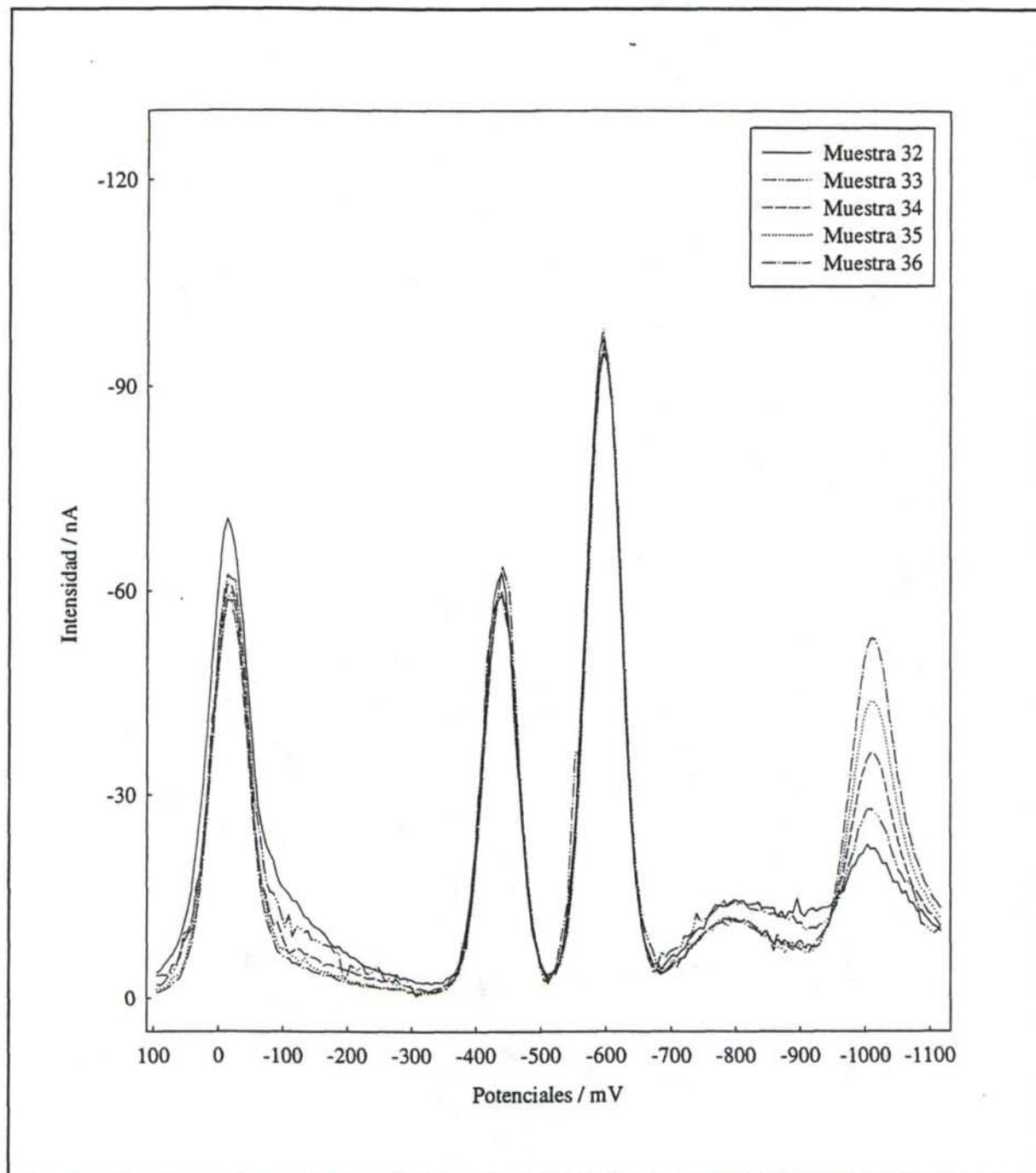


Figura III.3.2j Polarogramas del calibrado A correspondientes a las muestras especificadas en la Tabla III.3.1.

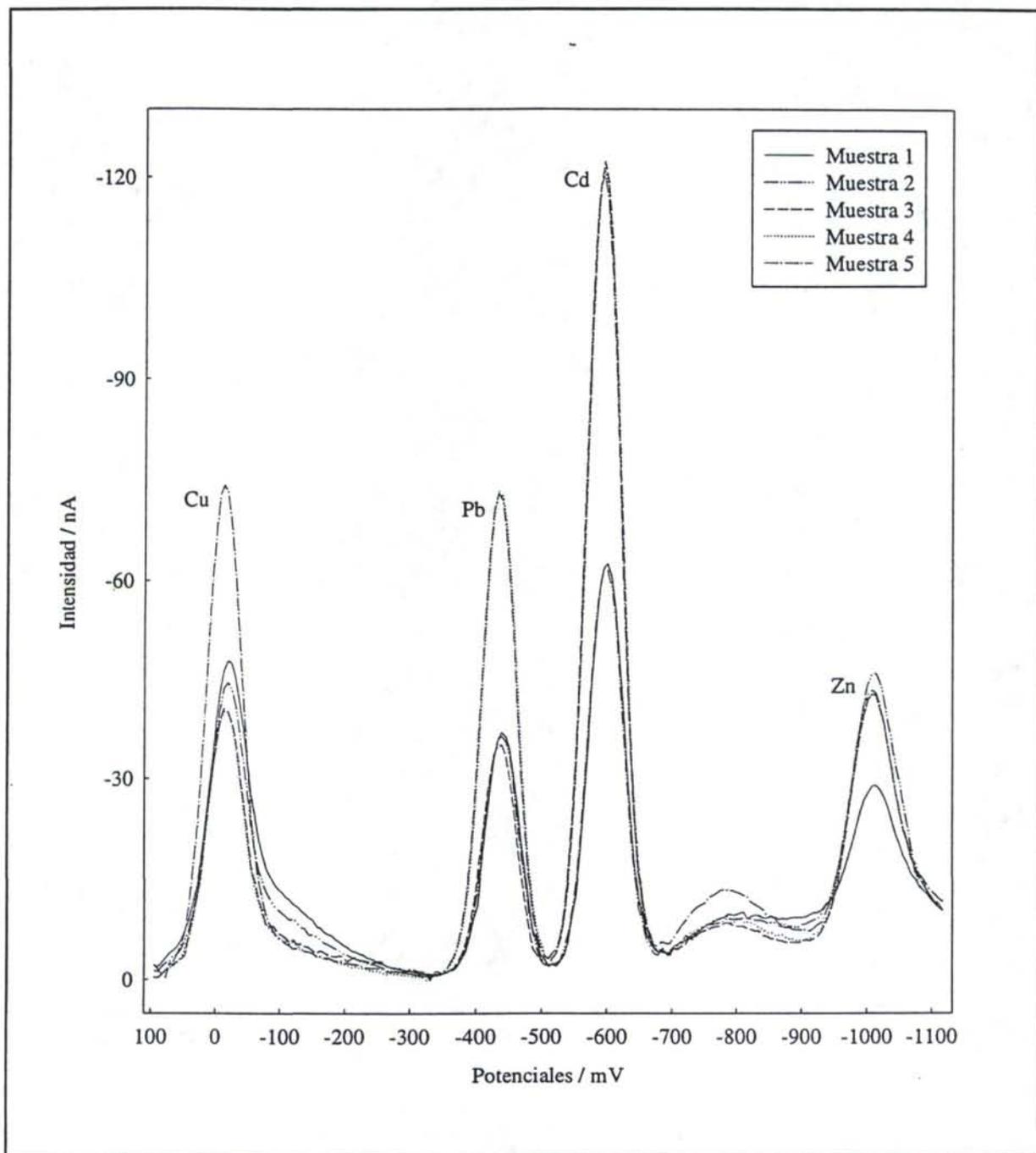


Figura III.3.3a Polarogramas del calibrado B correspondientes a las muestras especificadas en la Tabla III.3.1.

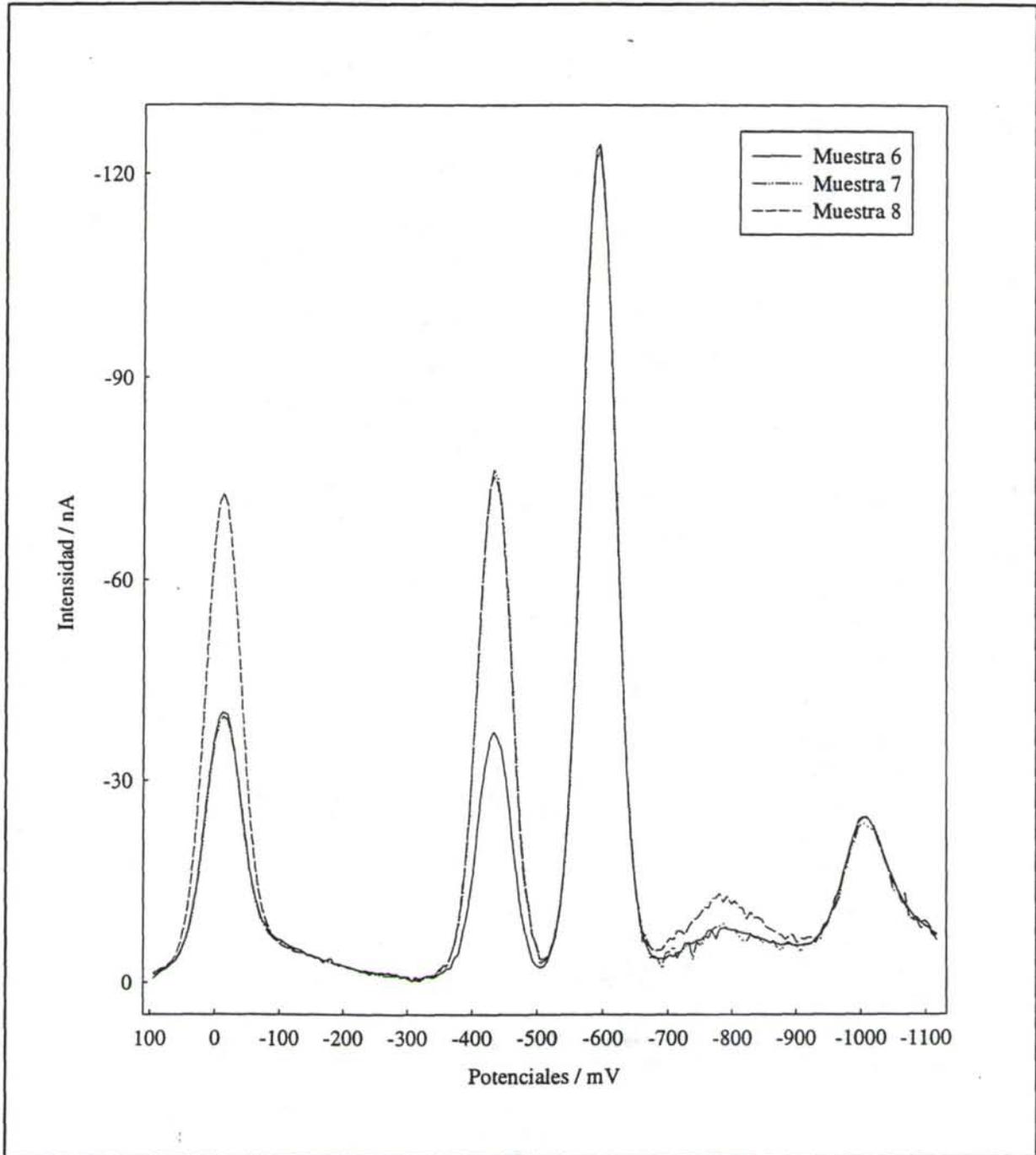


Figura III.3.3b Polarogramas del calibrado B correspondientes a las muestras especificadas en la Tabla III.3.1.

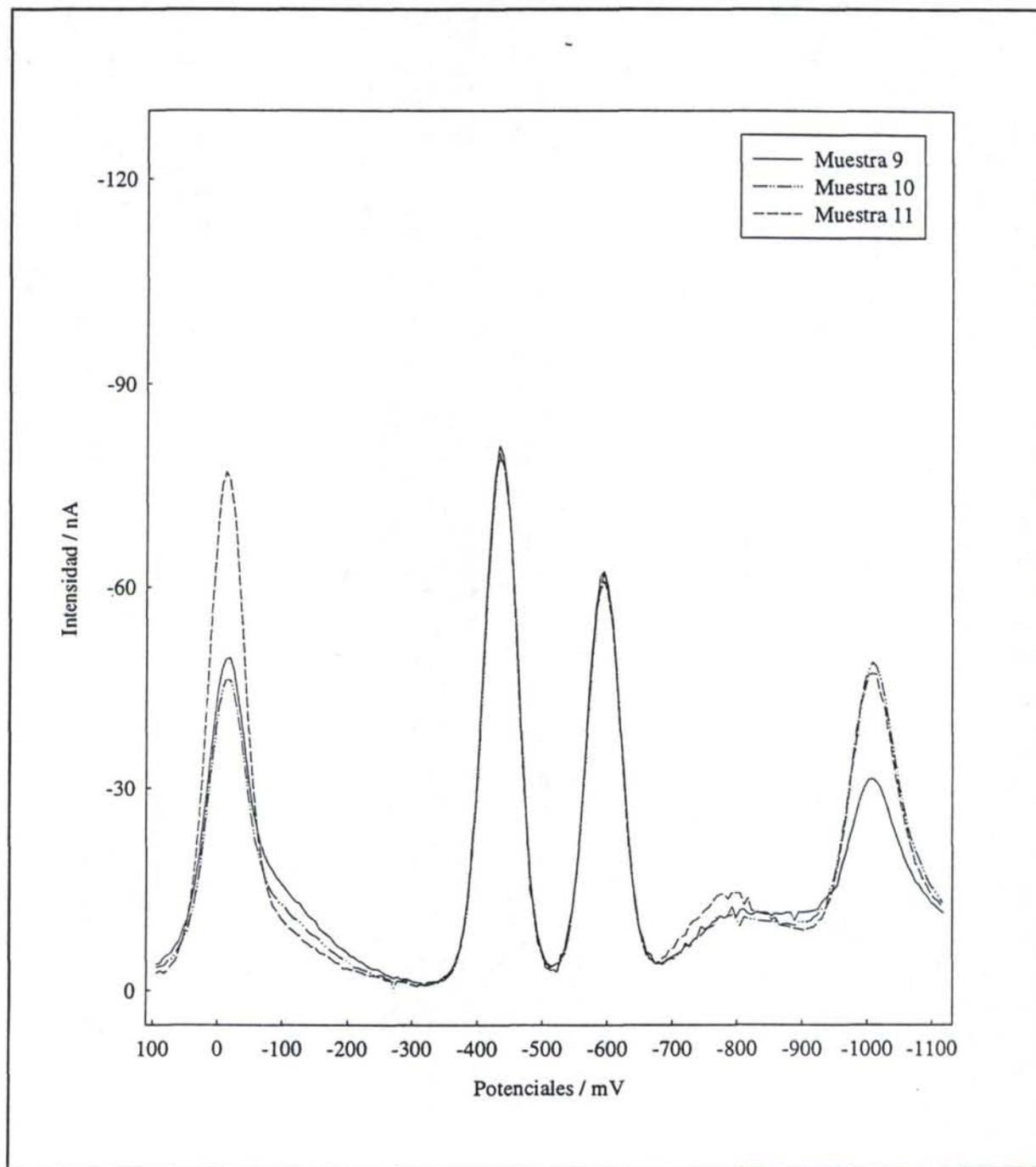


Figura III.3.3c Polarogramas del calibrado B correspondientes a las muestras especificadas en la Tabla III.3.1.

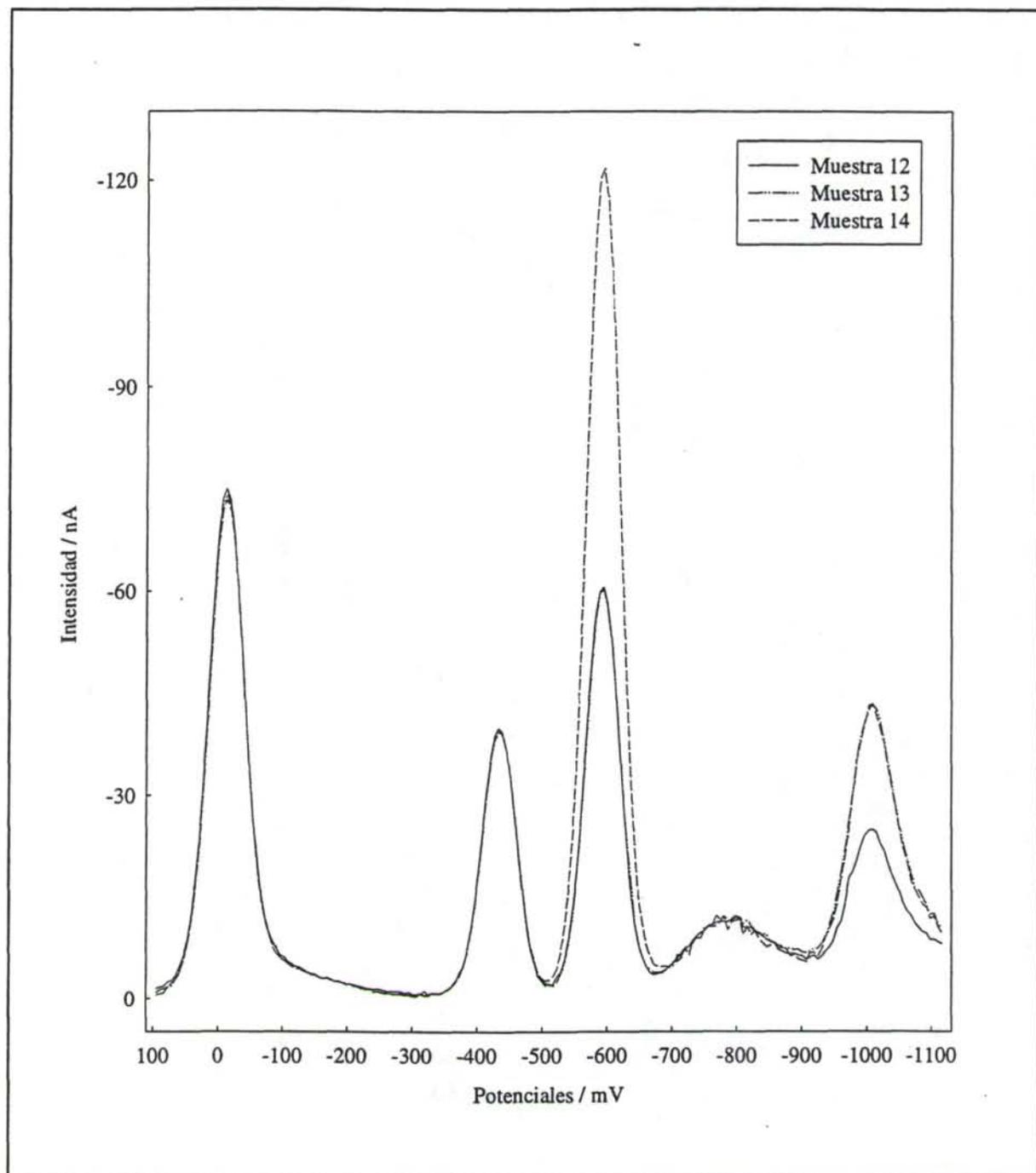


Figura III.3.3d Polarogramas del calibrado B correspondientes a las muestras especificadas en la Tabla III.3.1.

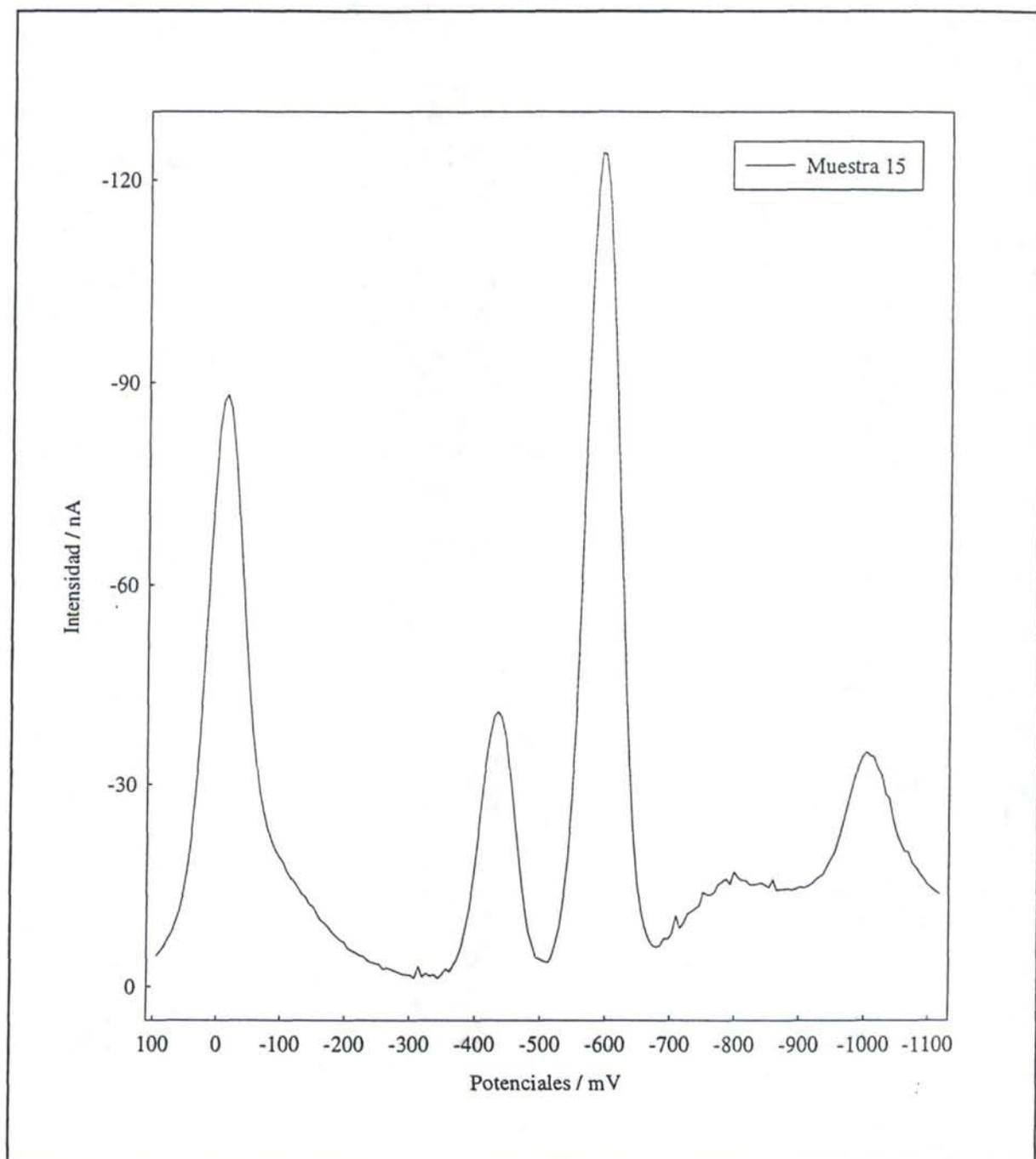


Figura III.3.3e Polarogramas del calibrado B correspondientes a las muestras especificadas en la Tabla III.3.1.

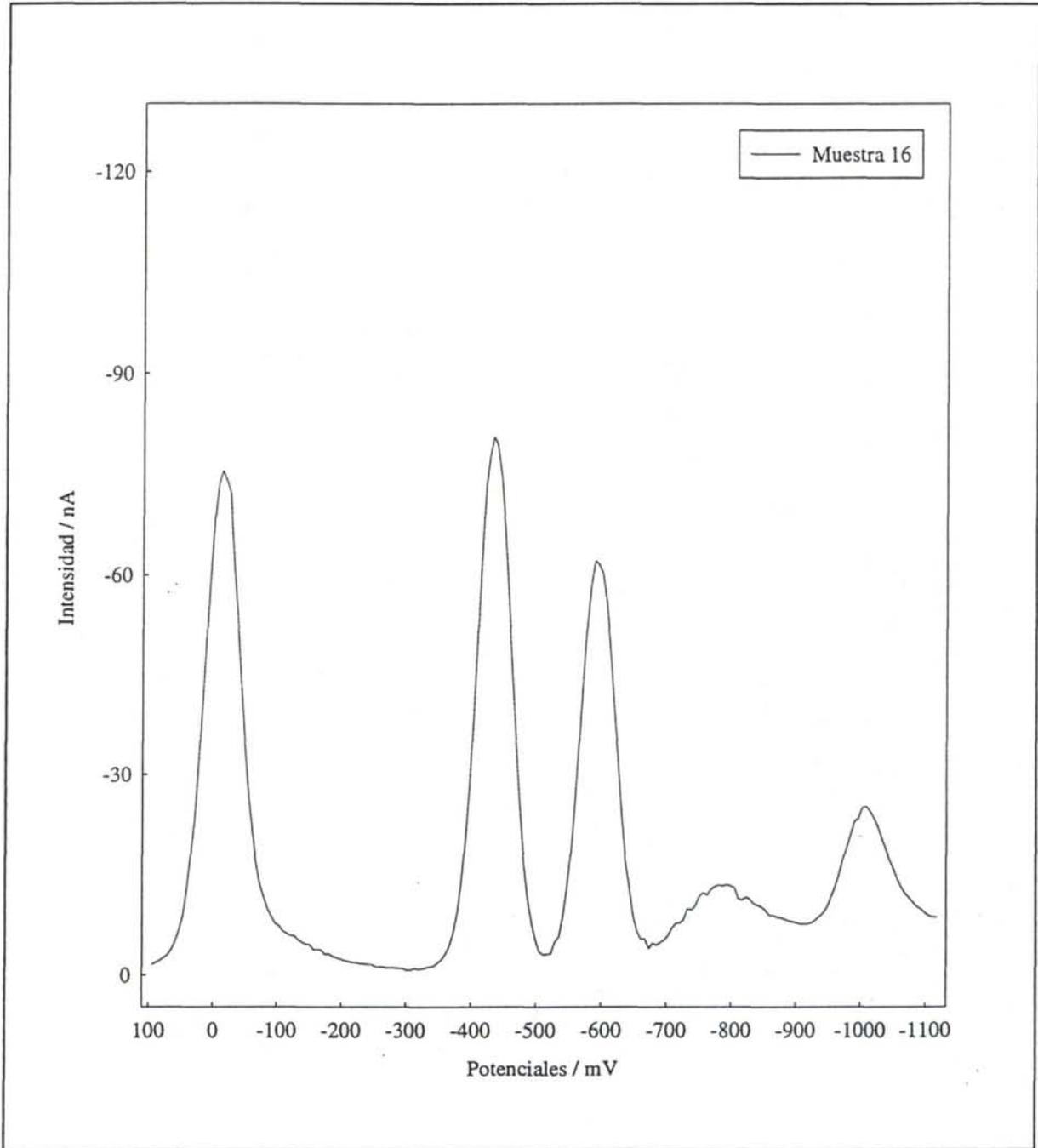


Figura III.3.3f Polarogramas del calibrado B correspondientes a las muestras especificadas en la Tabla III.3.1.

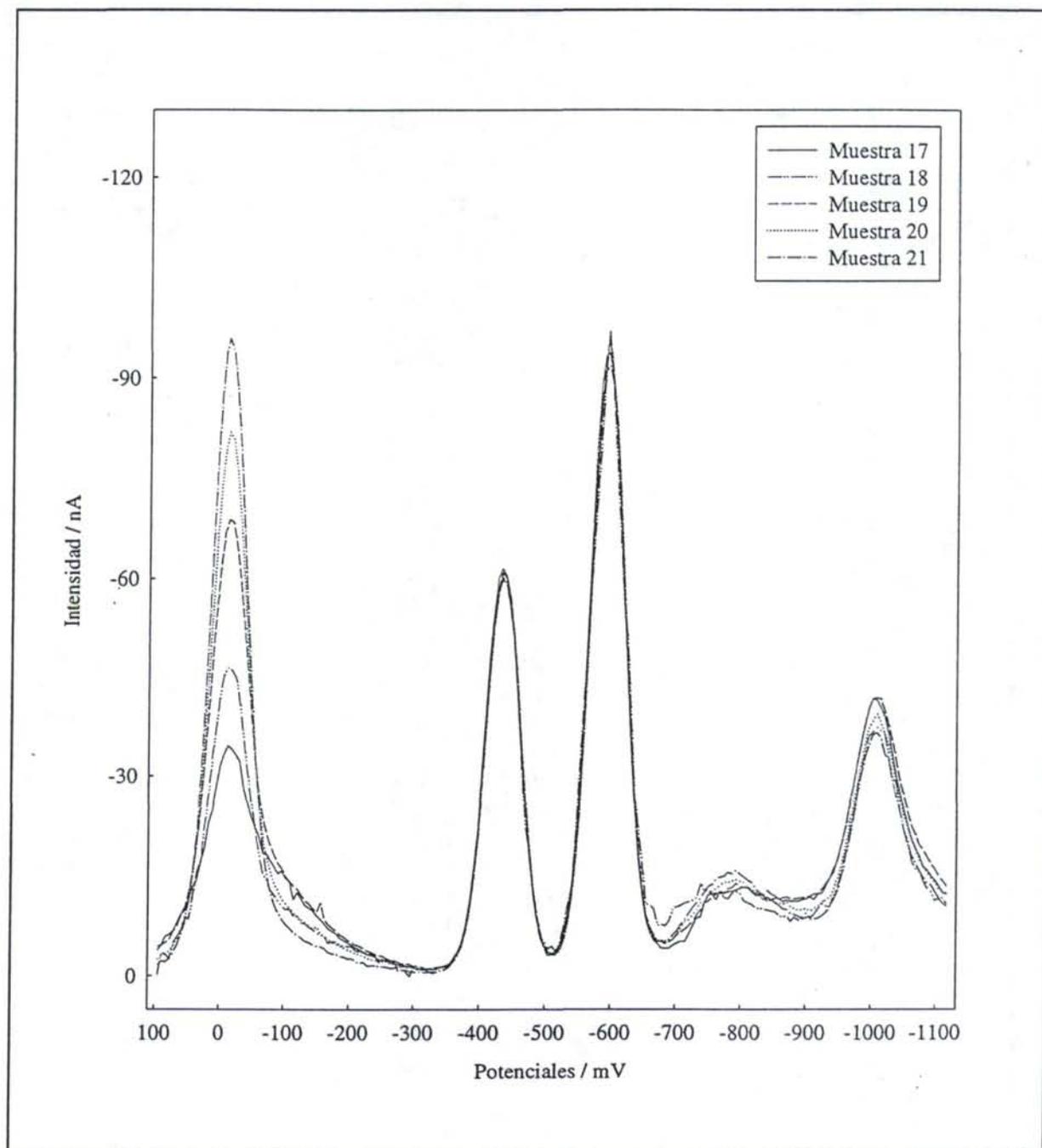


Figura III.3.3g Polarogramas del calibrado B correspondientes a las muestras especificadas en la Tabla III.3.1.

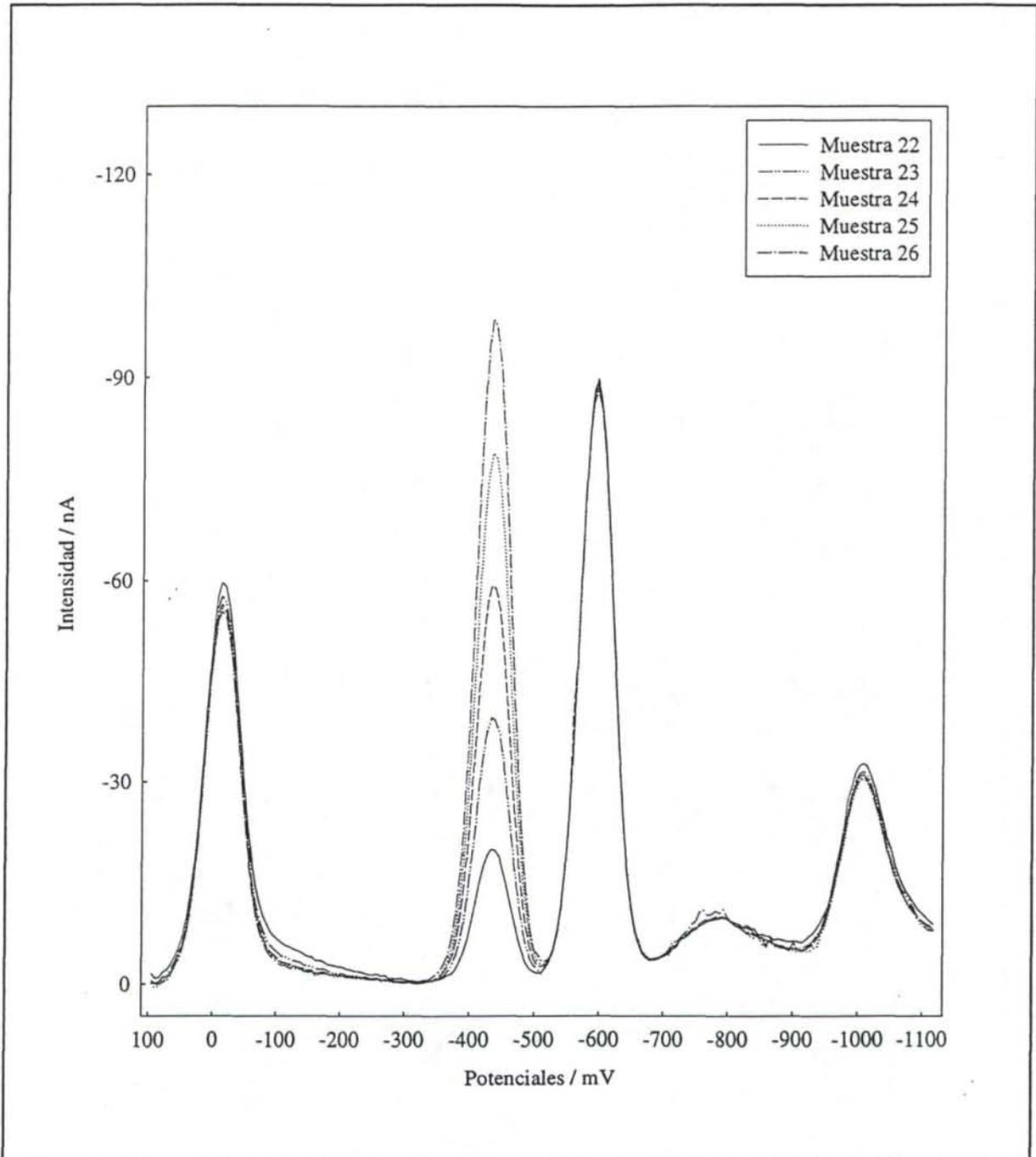


Figura III.3.3h Polarogramas del calibrado B correspondientes a las muestras especificadas en la Tabla III.3.1.

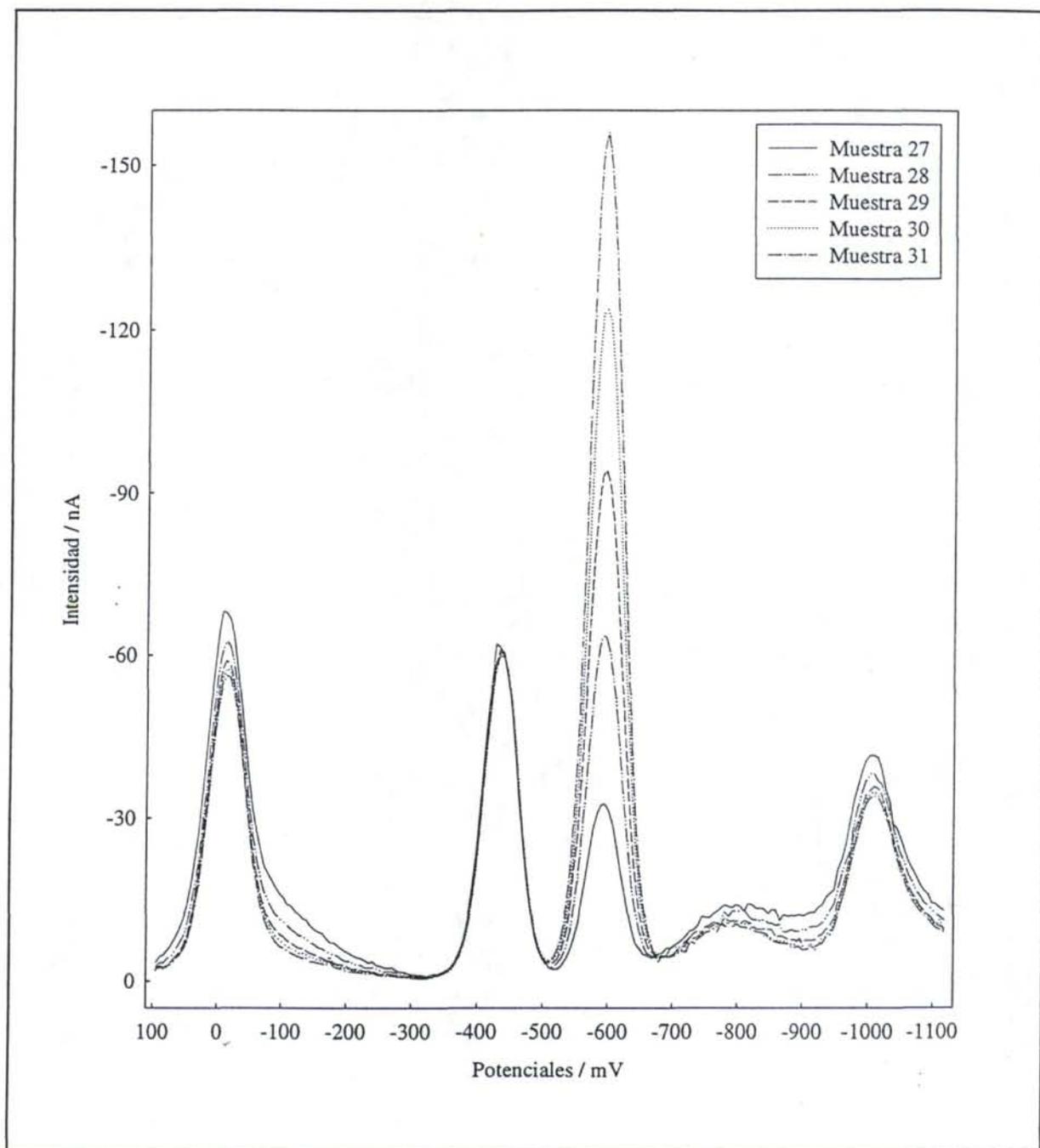


Figura III.3.3i Polarogramas del calibrado B correspondientes a las muestras especificadas en la Tabla III.3.1.

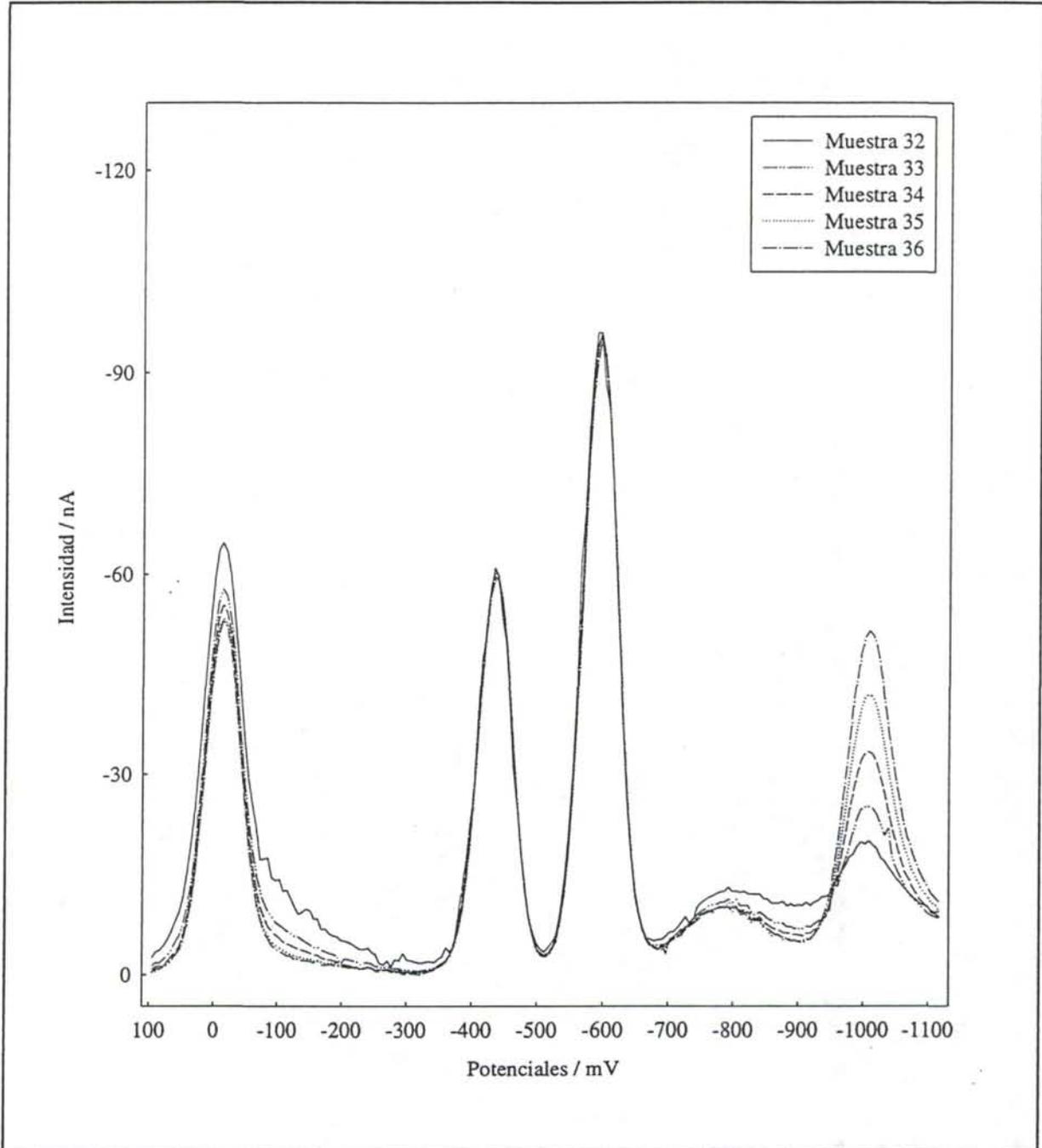


Figura III.3.3j Polarogramas del calibrado B correspondientes a las muestras especificadas en la Tabla III.3.1.

propias muestras, etc., por lo que si se lleva a cabo un procedimiento de transferencia de calibrado todas estas fuentes de variabilidad se habrán tenido en cuenta.

En los polarogramas registrados aparecen cuatro picos bien definidos, que se corresponden con cobre (-0.013 V), plomo (-0.439 V), cadmio (-0.595 V) y zinc (-1.009 V). Observando las adiciones realizadas sobre las distintas muestras se puede concluir que el pico correspondiente al cobre disminuye cuando se adiciona zinc, y el correspondiente al zinc disminuye cuando se adiciona cobre. También se advierte una disminución de los picos de cobre y zinc cuando se adiciona cadmio, aunque menos acusada que la anterior. No obstante, estos fenómenos podrían asociarse con la pequeña dilución que supone cada una de las adiciones realizadas.

III.3.3 Análisis univariante

Para realizar este estudio se ha utilizado una regresión univariante por mínimos cuadrados, en la que se han considerado las intensidades de pico del calibrado A (a las que se ha restado el blanco experimental) frente a la concentración de cada analito. Esta metodología requiere un análisis de regresión independiente para cada uno de los metales, el que cual se anotarán los valores de la intensidad de pico del analito de interés a distintos niveles de concentración del mismo.

Al utilizar el procedimiento univariante, basado en la modificación de la concentración de uno de los analitos mientras se mantiene la de los demás constante, se ha tomado para el calibrado de cobre la intensidad del primer pico (muestras 17, 18, 19, 20 y 21), para el de plomo la del segundo pico (muestras 22, 23, 24, 25, y 26), para el de cadmio la del tercer pico (muestras 27, 28, 29, 30 y 31), y para el de zinc la del último pico (muestras 32, 33, 34, 35 y 36). La numeración de las distintas muestras se corresponde con la que aparece recogida en la Tabla III.3.1. Estos datos son los mismos que se utilizarán posteriormente para llevar a cabo la transferencia de calibrado multivariante.

Los resultados de este análisis de regresión se muestran en la Tabla III.3.2, en la cual puede observarse que tanto para el plomo como para el cadmio los coeficientes de correlación son muy altos, mientras que los correspondientes a cobre y zinc son menores. Aunque para

estos dos últimos analitos los citados coeficientes indican que el ajuste no es malo, se debe concluir lo contrario al tener en cuenta lo amplios que son los intervalos de confianza tanto para la pendiente como para el término independiente en todos los casos. Esta conclusión se reafirma considerando el valor de la desviación estándar de la regresión, S_{yx} , que también es elevado, del orden de seis veces mayor, para cobre y zinc. De hecho, no hay que olvidar que en el calibrado se han utilizado únicamente cinco muestras, y que por tanto cabe esperar que, con tan pocos puntos, los coeficientes de regresión puedan tener valores elevados aun cuando el ajuste no sea muy bueno.

Tabla III.3.2 Resultados del análisis de regresión por mínimos cuadrados llevado a cabo para cada analito.

Regresión	Cobre	Plomo	Cadmio	Zinc
Pendiente	7.89 ± 1.05	9.87 ± 0.17	10.44 ± 0.11	4.11 ± 0.63
Término independiente	15.08 ± 6.61	1.73 ± 1.09	0.93 ± 1.09	14.11 ± 4.96
R^2_{ajustado}	0.9931	0.9999	1.0000	0.9908
S_{yx}	1.97	0.32	0.32	1.18

Para tener un término de comparación de la capacidad de predicción de los modelos univariantes que es posible construir se ha aplicado un método clásico de adición estándar utilizando las intensidades de pico del calibrado B (el mismo que se ha seleccionado en los siguientes apartados para evaluar la capacidad de predicción de los modelos multivariantes de análisis y de los procedimientos de transferencia de calibrado) y se ha determinado la concentración de las muestras 18, 23, 28 y 33 del diseño experimental de la Tabla III.3.1. Lo cual significa que se ha determinado la concentración de cobre en la muestra 18, la de plomo en la muestra 23, la de cadmio en la muestra 28, y la de zinc en la muestra 33. Dichas muestras han sido consideradas muestras problema (en los tratamientos multivariantes siguientes también han sido tomadas como muestras test) sobre las cuales se han realizado adiciones del analito de interés (las muestras marcadas en la Tabla III.3.1 con 3, 4 y 5 en el analito de interés y 3 en los restantes).

Los resultados del procedimiento de adición estándar se recogen en la Tabla III.3.3, así como las concentraciones calculadas para cada analito y los errores relativos correspondientes. Se observa que los errores obtenidos en el cálculo de las concentraciones

de las muestras problema son aceptables en el caso de plomo y cadmio, mientras que, para cobre y zinc son del orden del 40%. El hecho de que de nuevo sean el cobre y el zinc los analitos que proporcionen peores resultados es atribuible a la existencia de algún tipo de interferencia que afecta a la calidad de la señal analítica y a la propia determinación. Como consecuencia de ello, el análisis univariante se muestra insuficiente para resolver este problema y, por tanto, parece adecuado intentar abordarlo utilizando una metodología multivariante que tenga en cuenta las relaciones entre los analitos implicados.

Tabla III.3.3 Resultados del procedimiento de adición estándar univariante realizado con los datos del calibrado B.

Regresión	Cobre	Plomo	Cadmio	Zinc
Pendiente	8.56	10.36	10.35	4.62
Término independiente	48.91	39.57	64.19	24.80
R^2_{ajustado}	0.9761	1.0000	0.9993	0.9981
S_{yx}	4.26	0.12	1.02	0.49
Número de muestra	18	23	28	33
Concentración calculada (μM)	5.7142	4.8195	6.1053	5.3680
Error relativo (%)	49.15	-0.30	1.64	40.12

III.3.4 Análisis multivariante

Debido a la existencia de algunas interferencias entre los analitos no es adecuada la utilización de la regresión lineal múltiple [12, 124]. Es preferible disponer de una técnica de regresión que sea capaz de modelar este tipo de fenómenos y que, al mismo tiempo, obvie el problema que necesariamente van a causar las altas correlaciones y colinealidades existentes en la matriz de las variables predictoras, es decir, la matriz de intensidades. Por ello, se ha optado por usar una técnica de regresión multivariante como PLS ya que, utilizando el polarograma completo es posible incluir en el calibrado la información relativa a las relaciones que puedan existir entre los distintos analitos. Se trata de una calibración

multivariante indirecta, en el sentido de hacer adiciones de cada uno de los analitos, y de utilizar todo el polarograma, sin necesidad de seleccionar sensores específicos, para incluir las relaciones internas antes aludidas.

III.3.4.1 Construcción y evaluación de modelos PLS

Construcción de los modelos PLS

El método de regresión multivariante PLS se ha aplicado a los dos calibrados multicomponentes cuyos polarogramas se muestran en las Figuras III.3.2a-j y III.3.3a-j, el calibrado A y el calibrado B. Como variable respuesta se ha considerado la concentración de uno de los cuatro analitos, mientras que las variables predictoras han sido las intensidades registradas en los diferentes polarogramas, a los que se ha restado previamente el correspondiente blanco experimental. De cada polarograma se han digitalizado las intensidades registradas a 203 potenciales, igualmente espaciados, entre 0.095 V y -1.117 V. En consecuencia, la matriz de datos original, tanto en el calibrado A como en el B, consta de 36 objetos (cada una de las muestras de calibrado que aparecen en la Tabla III.3.1) y de 203 variables predictoras (la intensidad medida a cada potencial).

De las 36 muestras que constituyen cada uno de los calibrados, se han considerado como objetos del conjunto test (no han intervenido en ningún momento en la construcción de los modelos PLS) las anotadas en la Tabla III.3.1 con los números de muestra siguientes: 18, 20, 23, 25, 28, 30, 33 y 35. Estas muestras se corresponden con los puntos centrales de las caras del hipercubo que resulta al representar en el espacio experimental las concentraciones del diseño del experimento (reducido a tres dimensiones en la III.3.1), y se utilizan para evaluar la capacidad de predicción de los modelos de regresión PLS contruídos para cada calibrado. Las muestras restantes, un total de 28, constituyen el conjunto de aprendizaje (training set) con el que se construye cada uno de los modelos PLS.

La regresión PLS se ha llevado a cabo sobre los datos originales (sin autoescalar) para cada analito de forma independiente, utilizando el método de validación cruzada [22, 23] con tres grupos de cancelación. Esto supone, que para cada uno de los dos calibrados, se dispone de cuatro modelos PLS, uno por analito (cobre, plomo, cadmio y zinc), los cuales se van a analizar de forma independiente.

Evaluación de los modelos PLS

Del análisis de los modelos PLS construidos con los 28 datos del conjunto de aprendizaje del calibrado A (del cual se han considerado como objetos test los indicados anteriormente) resultan los porcentajes de varianza explicada que se resumen en la Tabla III.3.4. En la misma se observa que, como cabía esperar, los porcentajes de varianza explicada en el bloque Y (concentración) aumentan a medida que se incluyen nuevas variables latentes en el modelo, mientras que la varianza en predicción disminuye cuando se incorporan al mismo variables latentes que podrían no estar relacionadas con la respuesta, sino estar explicando únicamente varianza en el bloque X (variables predictoras) o incluso estar ligadas a muestras u objetos concretos. En base a este criterio, que considera que el número de variables latentes óptimo es aquel que conduce a un valor de la varianza explicada en predicción mayor, se han seleccionado los modelos construidos con cinco variables latentes para cobre y zinc, y con cuatro variables latentes para plomo y cadmio.

En todos los casos se trata de modelos con buena capacidad de predicción y muy estables, ya que los porcentajes de varianza explicada alcanzados, tanto en evaluación como en predicción, son elevados y se encuentran muy próximos unos de otros.

Cabe destacar el hecho de que se ha mejorado la calidad de las determinaciones (hasta cierto punto serían comparables los porcentajes de varianza explicada por la regresión multivariante con los coeficientes de determinación, R^2_{ajustado} , del análisis univariante de la Tabla III.3.2), sobre todo en lo que se refiere a la estimación en predicción de la concentración, ya que los valores de varianza en predicción son elevados en todos los casos.

Otro modo de evaluar la capacidad real de predicción de estos modelos, además de a través de su varianza en predicción (calculada mediante validación cruzada), consiste en aplicar los modelos a muestras test, que no intervinieron en su construcción, y en evaluar la calidad de la predicción llevada a cabo. En la Tabla III.3.5 se muestran los errores relativos correspondientes a las muestras del conjunto test del calibrado A. En la misma destaca el hecho de que, en 27 de los 32 valores de concentración calculados, el error relativo se encuentran por debajo del 3%, mientras que los restantes quedan muy lejos de los valores de error proporcionados por la regresión univariante (cercaos al 40%).

En un análisis similar llevado a cabo para el calibrado B, asignando los mismos objetos al conjunto de aprendizaje y al conjunto test que en el calibrado anterior, se obtienen resultados perfectamente comparables. Los porcentajes de varianza explicada por los modelos

Tabla III.3.4 Evolución de los porcentajes de varianza explicada en función del número de variables latentes (V.L.) seleccionadas para los modelos construidos para cada analito en el calibrado A.

V. L.	Cobre			Plomo		
	Bloque Y		Bloque X	Bloque Y		Bloque X
	Var. explicada	Var. en predicción	Var. explicada	Var. explicada	Var. en predicción	Var. explicada
1	17.66	19.64	76.29	4.75	–	76.29
2	74.68	74.93	82.86	92.72	88.23	81.09
3	87.84	86.44	88.32	99.35	99.19	89.77
4	98.44	98.60	92.07	99.83	99.77	93.44
5	99.66	99.55	96.84	99.84	99.76	95.99
6	99.72	99.53	97.53	99.89	99.80	97.46
7	99.89	99.52	97.63	99.96	99.83	97.67
8	99.95	99.58	97.68	99.98	99.86	97.85

V. L.	Cadmio			Zinc		
	Bloque Y		Bloque X	Bloque Y		Bloque X
	Var. explicada	Var. en predicción	Var. explicada	Var. explicada	Var. en predicción	Var. explicada
1	59.38	59.28	76.29	–	–	76.29
2	99.70	99.35	83.48	51.10	45.85	79.81
3	99.77	99.72	89.36	90.18	90.62	84.97
4	99.79	99.73	93.34	95.37	97.13	92.40
5	99.84	99.64	94.90	99.14	99.11	96.84
6	99.87	99.68	97.42	99.34	98.94	97.34
7	99.93	99.77	97.78	99.51	98.89	97.75
8	99.96	99.76	97.89	99.69	98.60	97.88

de regresión PLS construídos para este segundo calibrado se encuentran recogidos en la Tabla III.3.6, y son del mismo orden que los correspondientes al calibrado A, Tabla III.3.4, aunque en este caso es necesario mayor número de variables latentes para alcanzar dichos valores. En concreto, sería necesario tomar 8 variables latentes en los modelos construídos para cobre y plomo (los porcentajes de varianza en predicción correspondientes a la novena variable

Tabla III.3.5 Errores relativos (en %) obtenidos para las muestras test por los modelos PLS del calibrado A.

Número de muestra	Analito			
	Cobre	Plomo	Cadmio	Zinc
18	0.52	0.87	2.22	-2.26
20	-1.71	0.18	-0.34	-1.58
23	-1.39	-11.75	2.22	12.87
25	-0.18	-0.26	2.24	0.53
28	-1.22	0.35	-1.33	-0.70
30	-0.53	1.23	0.00	2.81
33	-4.52	7.83	2.99	-5.22
35	0.70	2.28	1.23	1.97

latente son 99.35% y 99.52% respectivamente), mientras que bastarían cinco para el modelo correspondiente al cadmio y seis para el del zinc.

Este hecho muestra la existencia de una variabilidad en la respuesta experimental a lo largo del tiempo, que puede deberse tanto al instrumento en sí mismo como a las propias muestras, y que aparece reflejada en las diferentes estructuras latentes de los modelos construidos para ambos calibrados. La estructura interna del conjunto de calibración cambia de un día a otro, lo cual muestra la imposibilidad de utilizar el modelo de calibrado construido un día con los datos obtenidos utilizando el mismo instrumento una vez que ha transcurrido un cierto tiempo.

Por otro lado, en lo que respecta a los valores de los errores relativos, tampoco existen diferencias apreciables entre los datos correspondientes al calibrado B, Tabla III.3.7, y los obtenidos para el calibrado A, Tabla III.3.5. Los porcentajes de error son del mismo orden que los anteriores aunque, en este caso, para determinarlos se han considerado los valores de concentración calculados mediante los modelos de regresión construidos con el número de variables latentes señalado anteriormente para este segundo calibrado.

Debe recordarse además que los errores relativos obtenidos al calcular las concentraciones en el calibrado B mediante un análisis univariante, Tabla III.3.3, eran 49.15, 0.30, 1.64 y 40.12% para cobre, plomo, cadmio y zinc respectivamente. Estos errores tan

Tabla III.3.6 Evolución de los porcentajes de varianza explicada en función del número de variables latentes (V.L.) seleccionadas para los modelos construidos para cada analito en el calibrado B.

V. L.	Cobre			Plomo		
	Bloque Y		Bloque X	Bloque Y		Bloque X
	Var. explicada	Var. en predicción	Var. explicada	Var. explicada	Var. en predicción	Var. explicada
1	8.88	10.78	75.55	12.38	2.20	75.55
2	69.06	69.35	82.03	86.77	80.03	81.68
3	88.16	85.30	87.86	98.65	98.45	89.01
4	96.33	95.70	92.43	98.82	98.64	93.51
5	99.43	99.16	96.89	98.96	98.71	95.46
6	99.52	99.30	98.33	99.05	98.92	98.32
7	99.78	99.41	98.44	99.78	99.53	98.41
8	99.87	99.45	98.55	99.90	99.54	98.51

V. L.	Cadmio			Zinc		
	Bloque Y		Bloque X	Bloque Y		Bloque X
	Var. explicada	Var. en predicción	Var. explicada	Var. explicada	Var. en predicción	Var. explicada
1	59.43	59.79	75.55	–	–	75.55
2	99.55	99.13	83.35	41.72	40.64	80.75
3	99.56	99.50	85.98	85.74	77.15	85.43
4	99.56	99.46	91.14	97.36	94.23	87.58
5	99.55	99.51	96.35	99.36	98.73	96.89
6	99.57	99.50	98.31	99.42	99.19	98.32
7	99.86	99.49	98.37	99.71	99.12	98.44
8	99.91	99.38	98.42	99.85	99.06	98.53

elevados fueron atribuidos a la existencia de interferencias en el análisis. Sin embargo, en los valores de la Tabla III.3.7 se observa que no existen diferencias entre los porcentajes de error relativo de unos metales y otros, de lo cual se deduce que las regresiones multivariantes PLS llevadas a cabo son capaces de modelar perfectamente ese fenómeno.

Tabla III.3.7 Errores relativos (en %) obtenidos para las muestras test por los modelos PLS del calibrado B.

Número de muestra	Analito			
	Cobre	Plomo	Cadmio	Zinc
18	5.48	12.18	-2.11	-1.74
20	2.10	1.93	0.45	3.68
23	1.22	0.78	-3.55	-4.00
25	2.11	0.39	-3.69	-4.39
28	0.00	-0.17	2.16	0.52
30	0.88	1.05	1.68	1.93
33	-2.96	2.09	2.77	1.83
35	-2.11	1.40	4.36	1.18

III.3.4.2 Transferencia de calibrado mediante el procedimiento PDS

Una vez construídos y analizados los modelos de regresión, tanto del calibrado A como del calibrado B, se ha visto la necesidad de realizar un procedimiento de transferencia de calibrado para evaluar y modelar la variabilidad existente con el tiempo entre los modelos de ambos calibrados. Se ha llevado a cabo la transferencia del calibrado A (considerado éste como situación patrón) al calibrado B (situación satélite), lo cual supone una transferencia de calibrado a lo largo del tiempo. La finalidad de dicha transferencia de calibrado es una disminución del esfuerzo experimental que será necesario en calibrados posteriores, a través de una reducción del número de muestras necesarias para calibrar la obtenidas en la situación satélite. Por esta razón, se han utilizado diferentes combinaciones de algunas de las muestras de ambos calibrados para calcular la matriz de transferencia, evaluándose los resultados obtenidos en cada caso.

Para realizar la transferencia de calibrado entre ambos conjuntos de datos se han utilizado dos de los procedimientos de transferencia especificados en el capítulo anterior. En primer lugar se va a utilizar el procedimiento de transferencia PDS, y a continuación se llevará a cabo una transferencia de calibrado análoga mediante el procedimiento GCT.

Función "pdsgen"

En este caso, para llevar a cabo este procedimiento de transferencia de calibrado se ha utilizado una función (m-function) implementada a tal efecto por Wise [125] en el PLS Toolbox de MATLAB [59]. Se trata de una función denominada "pdsgen", sirve para transferir un modelo de calibrado, obtenido en una situación determinada, a un conjunto de datos correspondientes a otra situación diferente, mediante el procedimiento de transferencia de calibrado directa a trozos (PDS), detallado anteriormente, y que se basa en el procedimiento descrito en las referencias [94] y [96]. Esta función está programada, en principio, para la estandarización de instrumentos de espectroscopía, pero su uso no se ve limitado en absoluto a datos de tipo espectroscópico, sino que puede ser utilizada con datos analíticos de diversa procedencia.

En general, el formato de entrada/salida de la función "pdsgen" es el que se especifica a continuación:

```
[standmat, subset] = pdsgen(spec1, spec2, window, nosamps)
```

donde los parámetros de entrada que aparecen implicados son las dos matrices de datos, *spec1* y *spec2*, y el número de potenciales que determinan la anchura de la ventana seleccionada para el cálculo de la matriz de transferencia, que viene dado por el parámetro *window*. El último es un parámetro opcional, *nosamps*, que puede ser o bien un escalar que indique el número de muestras a utilizar en el subconjunto de transferencia (en cuyo caso las muestras son seleccionadas por el programa de forma aleatoria), o bien un vector fila que especifique los índices de las muestras que van a ser empleadas en el cálculo de la matriz de transferencia. Es necesario que las muestras se encuentren en el mismo orden en las dos matrices que contienen las señales experimentales, al tiempo que, la situación en que se obtuvieron los datos de la primera matriz, *spec1*, es considerada como la situación patrón cuyo modelo de calibrado se transfiere a la situación definida por los datos de la matriz segunda, *spec2*, considerada situación satélite.

El formato de salida de la función incluye el parámetro *standmat*, que especifica el nombre dado a la matriz de transferencia que se calcula, y *subset*, que es un vector que indica los índices de las muestras que forman el conjunto de transferencia, es decir, aquellas muestras utilizadas para realizar la transferencia de calibrado, y que es importante cuando éstas se han elegido aleatoriamente.

Una vez que se ha calculado la matriz de transferencia, multiplicando ésta por una respuesta observada en la situación satélite, como se indica en la ecuación (25), es posible obtener una estimación de cómo hubiera sido esa respuesta en la situación patrón

$$\text{spec1est} = \text{spec2} * \text{standmat}$$

donde la matriz *spec1est* es la señal experimental corregida de la obtenida en la situación satélite.

A esta señal corregida se aplica el modelo de calibrado construido en la situación patrón para predecir los valores de concentración de las muestras problema medidas en la nueva situación. En el caso de que los modelos construidos con los datos de la situación patrón sean modelos de regresión PLS, basta multiplicar *spec1est* por la fórmula cerrada de dichos modelos. Para calcular los valores de concentración se utilizan estos coeficientes, los cuales acompañan a las variables predictoras en la estimación de la respuesta.

Aplicación de la función "pdsgen"

Como ya se ha indicado, se pretende llevar a cabo la transferencia de calibrado de la situación A a la situación B, de forma que sea posible utilizar el modelo PLS construido con los datos del calibrado A para predecir la concentración de las muestras del calibrado B. Además, en este procedimiento cabe la posibilidad de que el número de muestras del subconjunto de transferencia sea menor que el del conjunto de datos completo, lo cual ofrece la opción de no tener que completar todo el diseño del experimento para el nuevo calibrado, si no que sea suficiente con un mínimo número de muestras indispensables para transferir el modelo de calibrado de la situación patrón.

Por ello, se han llevado a cabo transferencias de calibrado utilizando distintos conjuntos de transferencia y diferentes anchuras de ventana para estudiar el efecto de estos dos parámetros sobre los valores de concentración predichos, buscando aquellos que conlleven un esfuerzo experimental razonable e información suficiente para poder realizar la transferencia de calibrado de forma adecuada.

El primer paso de este procedimiento de transferencia de calibrado consiste en la estimación del valor de la matriz de transferencia que relaciona las variables predictoras de ambas situaciones. Para calcular dicha matriz se ha utilizado la función "pdsgen", definida anteriormente, introduciendo los siguientes parámetros:

$$[F, \text{sub}] = \text{pdsgen}(A, B, w, no)$$

cuyo significado se especifica a continuación:

- A:** Matriz que contiene los polarogramas utilizados en la construcción del modelo PLS del calibrado A, es decir, las señales experimentales de la situación A. Consta de 36 filas (un polarograma por cada una de las muestras de la Tabla III.3.1) y de 203 columnas (los potenciales que se han digitalizado en cada polarograma). La matriz A contiene los datos del calibrado que se desea transferir.
- B:** Matriz que contiene los polarogramas utilizados en la construcción del modelo PLS del calibrado B. Es del mismo tamaño que la matriz A, y tanto las filas como las columnas se encuentran en el mismo orden que en esta última matriz. La matriz B contiene los datos a los que se desea transmitir el modelo de regresión del calibrado A.
- w:** Escalar que especifica el número de potenciales incluidos en la ventana móvil que se utiliza en el cálculo de la matriz de transferencia. En las transferencias de calibrado realizadas con distintos conjuntos de transferencia se han tomado para este parámetro los siguientes valores: 3, 5, 7, 9, 11 y 13.
- no:** Vector fila en el que se especifican los índices de las muestras de ambas matrices que se utilizan en la transferencia de calibrado. Estos índices se corresponden con el número de muestra que aparece en la Tabla III.3.1, y señalan las muestras seleccionadas para formar parte de los diferentes conjuntos de transferencia de acuerdo con lo especificado en la Tabla III.3.8. Las dimensiones del vector *no* son $1 \times d$, donde *d* varía dependiendo del número de muestras utilizadas para llevar a cabo la transferencia.

Como resultado de aplicar esta función se tienen las siguientes matrices: la matriz de transferencia (**F**) y un vector fila, *sub*, en el que se especifican los índices de las muestras que han intervenido en la transferencia de calibrado. La matriz de transferencia es una matriz banda diagonal, de dimensiones 203×203 , que es diferente para cada par de valores de los parámetros *no* y *w* introducidos, mientras que el vector *sub* coincide en cada caso con *no*.

Para realizar cada una de las transferencias de calibrado los conjuntos de datos de los calibrados A y B se dividen en tres partes. La primera de ellas está constituida en todos los casos por las ocho muestras test, que también fueron consideradas como tal al evaluar la capacidad de predicción de los modelos PLS construídos para ambos calibrados. De esta forma, es posible comparar los errores en la predicción de las concentraciones de las mismas

muestras a lo largo del análisis realizado. El segundo conjunto de muestras se encuentra constituido por aquellas que se utilizan para calcular el valor de la matriz de transferencia, es decir, es el denominado conjunto de transferencia, que viene dado en cada una de las transferencias de calibrado llevadas a cabo por el vector *no*. Mientras que en el tercero se encuentran recogidas las muestras restantes, es decir, aquellas que no han entrado a formar parte del conjunto de transferencia y que, por no ser las mismas en todos los casos, tampoco se han incluido en el conjunto test.

Es necesario que estos tres conjuntos de muestras sean iguales en los dos calibrados considerados para que sea posible llevar a cabo la transferencia de calibrado de uno a otro. No obstante, una vez que se haya seleccionado el conjunto de transferencia más adecuado, bastaría con disponer de los dos primeros subconjuntos para realizar transferencias de calibrado del calibrado patrón (calibrado A) a otros conjuntos de datos diferentes.

Los distintos valores del parámetro *no* introducidos, que se corresponden con los diferentes conjuntos de transferencia utilizados, se especifican en la Tabla III.3.8. En la misma también se indican, de acuerdo con los símbolos que aparece en la Tabla III.3.1, las disoluciones iniciales necesarias para obtener cada uno de los subconjuntos de transferencia (cada símbolo indica una disolución inicial diferente).

Para cada conjunto de transferencia y cada anchura de ventana se tiene una matriz de transferencia diferente, que se empleará en la corrección de los polarogramas de la situación B. A través de las matrices de transferencia se dispondrá de diferentes valores de concentración, predichos para cada analito según el conjunto de transferencia y la ventana utilizados, lo cual permitirá evaluar el efecto de estos dos parámetros sobre dichas predicciones.

Una vez que se dispone de la matriz de transferencia, es posible transformar las señales del calibrado B al formato que se estima tendrían de haber sido medidas en la situación A. Esta operación se realiza introduciendo la siguiente expresión:

$$\mathbf{Best} = \mathbf{B} * \mathbf{F}$$

donde **Best** es una matriz de dimensiones 36×203 en la que se recogen las señales del calibrado B corregidas.

A continuación, se puede utilizar el modelo PLS construido con los datos del calibrado A para predecir los valores de concentración del calibrado B, siempre a partir de las señales corregidas. Para ello, se ha utilizado la fórmula cerrada de los modelos PLS, calculada en el

Tabla III.3.8 Subconjuntos de muestras del conjunto de aprendizaje utilizadas para realizar la transferencia de calibrado (símbolos como en la Tabla III.3.1).

Conjunto (no)	Muestras del conjunto de aprendizaje	Disolución inicial
1	todas	todas
2	1 a 5	×
3	1 a 8	×▲
4	1 a 11	×▲·
5	1 a 14	×▲·○
6	1 a 15	×▲·○■
7	1 a 16	×▲·○■—
8	6 a 11	▲·
9	6 7 8 12 13 14	▲○
10	9 a 14	·○
11	17 19 21 22 24 26 27 29 31 32 34 36	▼□·▽
12	17 21 22 26 27 31 32 36	▼□·▽

programa PARVUS [57] mediante el procedimiento propuesto por Marengo y Todeschini [126]. Por tanto, para calcular los valores de concentración bastaría con introducir:

$$Y_{cal} = Best * Acer$$

donde Y_{cal} es una matriz, de dimensiones 36×4 , que contiene los valores de concentración calculados para cada analito a partir de las señales del calibrado B corregidas y del modelo PLS construido para el calibrado A. Las filas de esa matriz se corresponden con las 36 muestras de que se dispone en el calibrado B, es decir, el calibrado completo. De ellas, tan sólo algunas han sido utilizadas para realizar la transferencia de calibrado (sus índices estarán especificados en el parámetro *no*), mientras que, entre las demás se encuentran las ocho muestras que fueron asignadas al conjunto test cuando se construyeron los modelos PLS para los dos calibrados. Las columnas de esta matriz incluyen los valores de concentración calculados para cada uno de los cuatro analitos, en el orden siguiente: cobre, plomo, cadmio y zinc, el mismo en el que se encuentran dispuestos los coeficientes de las regresiones PLS en *Acer*.

La matriz **Acer** es una matriz de dimensiones 203×4 , en cuyas filas se recogen los coeficientes de la fórmula cerrada de PLS que acompañan a cada variable predictora en los modelos PLS del calibrado A. Cada columna se refiere a uno de los cuatro analitos y se encuentran dispuestas en el orden indicado anteriormente. Estos coeficientes son los correspondientes a los modelos PLS construidos con los datos del calibrado A, tomando el número de variables que se ha considerado óptimo para cada uno de los analitos, y que ya se especificó cuando se evaluaron los modelos de calibrado correspondientes.

En el modelo PLS construido para el calibrado A, las muestras asignadas al conjunto test (8 muestras en total) no han intervenido en ningún momento en la construcción de dicho modelo, y por tanto, tampoco en el cálculo de la fórmula cerrada del mismo, y además, esas muestras no han sido utilizadas en ningún momento para calcular la matriz de transferencia. Por ello, son muestras completamente ajenas al conjunto de datos utilizado para modelar la variabilidad en el tiempo de la señal electroquímica.

Siguiendo este procedimiento, que se representa en su conjunto de forma esquemática en la Figura III.3.4, se considera finalizada la transferencia de calibrado. Se dispone por tanto, en este caso concreto, de 72 valores de concentración calculados para cada analito en una muestra determinada, es decir, tantos como diferentes *no* por *w* se han utilizado para calcular las matrices de transferencia. Estas concentraciones se han calculado a partir de las señales medidas en una situación, utilizando un modelo de calibrado construido con los datos procedentes de una situación diferente. A continuación, se procede a evaluar los resultados obtenidos, tanto en lo que se refiere a la determinación de las mejores condiciones de transferencia de calibrado, como a la valoración de la calidad de las predicciones realizadas a través de la misma.

Análisis de los resultados

Se dispone de dos parámetros para describir la calidad de los valores de concentración calculados con el modelo PLS del calibrado A a partir de las señales obtenidas en la situación B. Estos parámetros se diferencian entre sí en función de si las muestras utilizadas en su cálculo han formado parte del conjunto test o si, por el contrario, han formado parte del conjunto de aprendizaje de los modelos PLS construidos con anterioridad.

La varianza de la regresión para el conjunto de evaluación (conjunto test) del calibrado B viene dada, para cada analito, por la expresión:

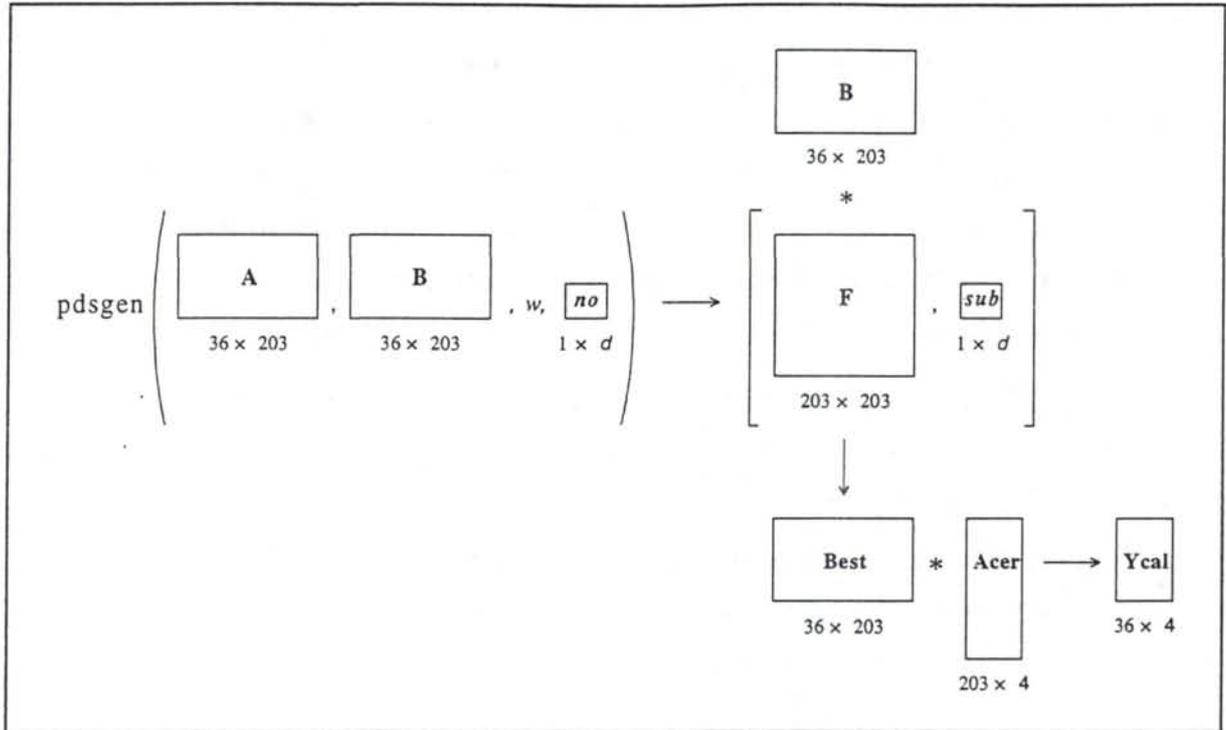


Figura III.3.4 Esquema completo del proceso de transferencia de calibrado y del cálculo de los valores de concentración.

$$s^2 = \frac{\sum_j (y_j - \hat{y}_{B,j})^2}{J} \quad (25)$$

donde y_j e $\hat{y}_{B,j}$ son respectivamente las concentraciones reales y las calculadas por el modelo PLS de la situación A para uno de los analitos, y J indica el número de objetos del conjunto de evaluación. La raíz cuadrada de esta varianza (desviación estándar residual en predicción) es el denominado error estándar en predicción, SEP (standard error of prediction).

De forma análoga, la varianza de la regresión para el conjunto de aprendizaje del calibrado B vendrá dada por la expresión:

$$s^2 = \frac{\sum_i (y_i - \hat{y}_{B,i})^2}{I} \quad (26)$$

donde y_i e $\hat{y}_{B,i}$ representan respectivamente los valores de concentración real y los calculados por el modelo PLS de la situación A para uno de los analitos, e I es el número de objetos del

conjunto de aprendizaje. La raíz cuadrada de esta varianza se ha denominado error estándar del conjunto de aprendizaje, SET (standard error of training set).

El cálculo de ambos errores, así como el procedimiento de transferencia de calibrado que se acaba de llevar a cabo siguiendo el método PDS, se encuentran automatizados en la función '*transfpds.m*'. Dicha función ha sido programada a tal efecto en MATLAB [59] y se encuentra recogida en el *Apéndice 2*.

Para valorar el efecto que tanto la variación del conjunto de transferencia utilizado como el tamaño de ventana tienen sobre las predicciones, se han considerado los valores del error estándar en predicción (SEP), que se representan a través de las correspondientes superficies de respuesta y curvas de nivel de las Figuras III.3.5 a III.3.8 para cada uno de los analitos. Los valores del SET obtenidos en este mismo análisis siguen una pauta similar a la que reflejada en dichas figuras.

En ellas se observa que, en la variación del SEP respecto al tamaño de ventana, las superficies correspondientes al plomo y al cadmio son parecidas (las mayores diferencias surgen en los distintos *no* utilizados), mientras que por otro lado, las correspondientes al cobre y al zinc a su vez presentan similitudes, siendo bastante diferentes de las anteriores. En estos dos últimos casos, la variabilidad del SEP se observa tanto en el eje de ordenadas, como en el de abscisas.

Los diagramas Box plot [58] de las Figuras III.3.9 a III.3.12, en los cuales la línea que divide horizontalmente las cajas representa la mediana, permiten establecer comparaciones visuales más claras entre los diferentes conjuntos de datos. Para construir los citados diagramas se han considerado los valores de SEP obtenidos con los diferentes tamaños de ventana, *w*, independientemente del conjunto de transferencia utilizado, *no*, y viceversa. No obstante, para analizarlos de una forma más objetiva se ha recurrido a un análisis de rangos múltiple a través de los test de agrupamientos para la media de Tukey y de Newman-Keuls [58]. Los resultados de los test de agrupamientos para los valores del SEP correspondientes a cada analito se muestran en las Tablas III.3.9 a III.3.12.

De este análisis se puede concluir que las diferentes anchuras de ventana utilizadas para calcular las matrices de transferencia no conducen a valores de SEP estadísticamente diferentes, según los test de Tukey y de Newman-Keuls, salvo en el caso del cadmio, donde ambos test encuentran agrupamientos entre algunos niveles de este factor. Ello se debe a que la dispersión entre los valores de SEP debida a los diferentes *no* es muy pequeña, lo cual hace

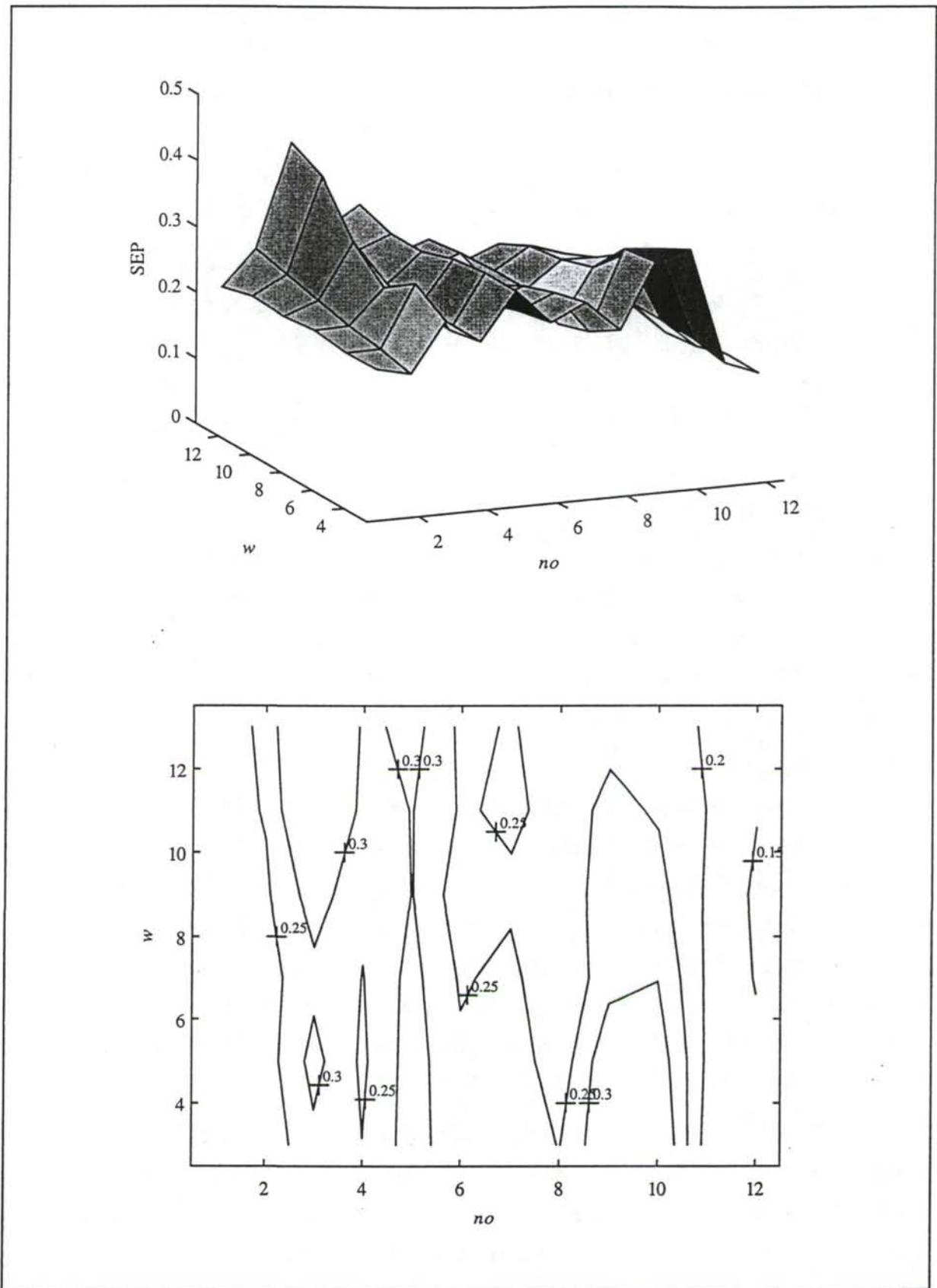


Figura III.3.5 Superficie de respuesta y curvas de nivel del SEP para el cobre.

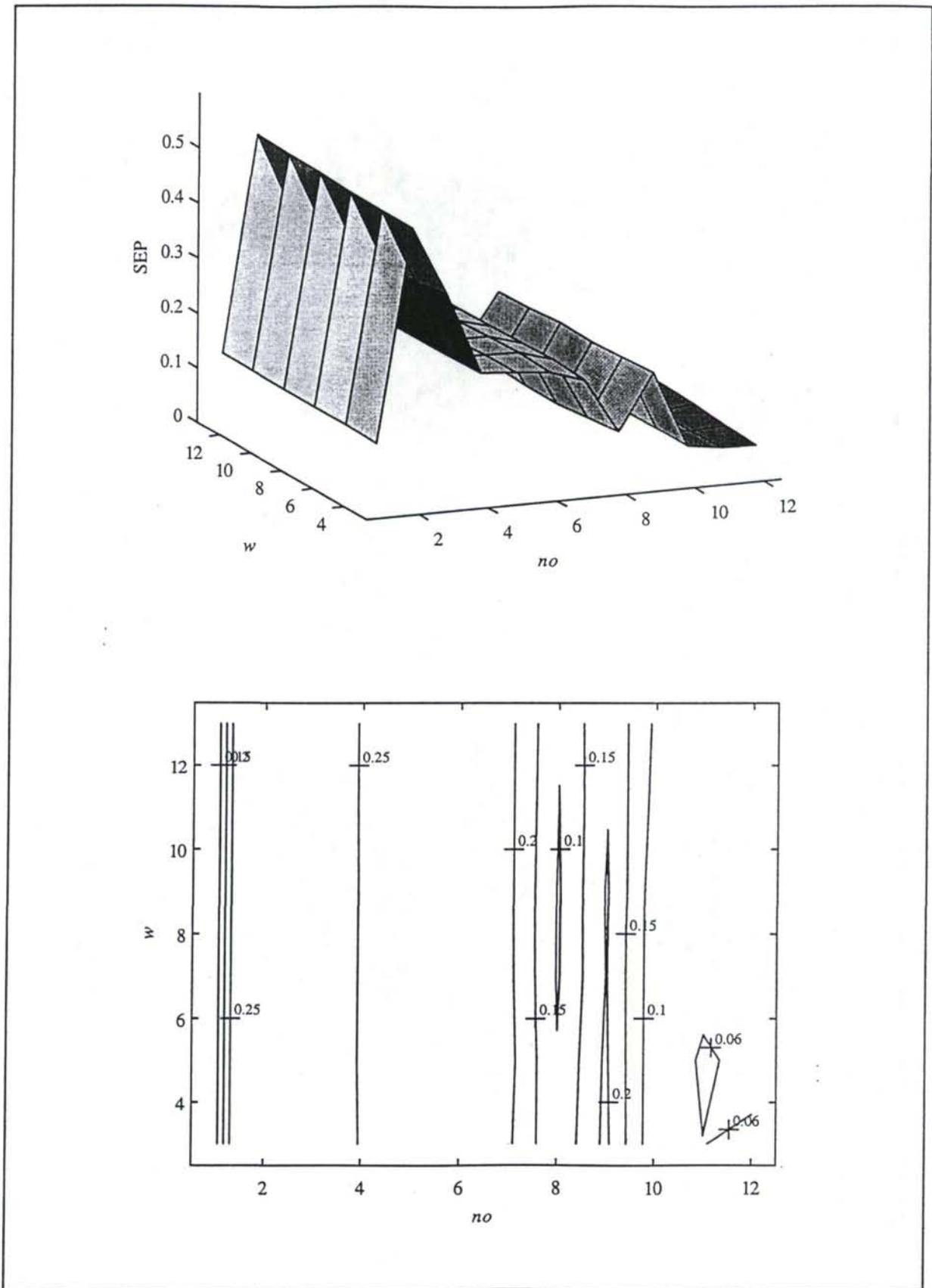


Figura III.3.6 Superficie de respuesta y curvas de nivel del SEP para el plomo.

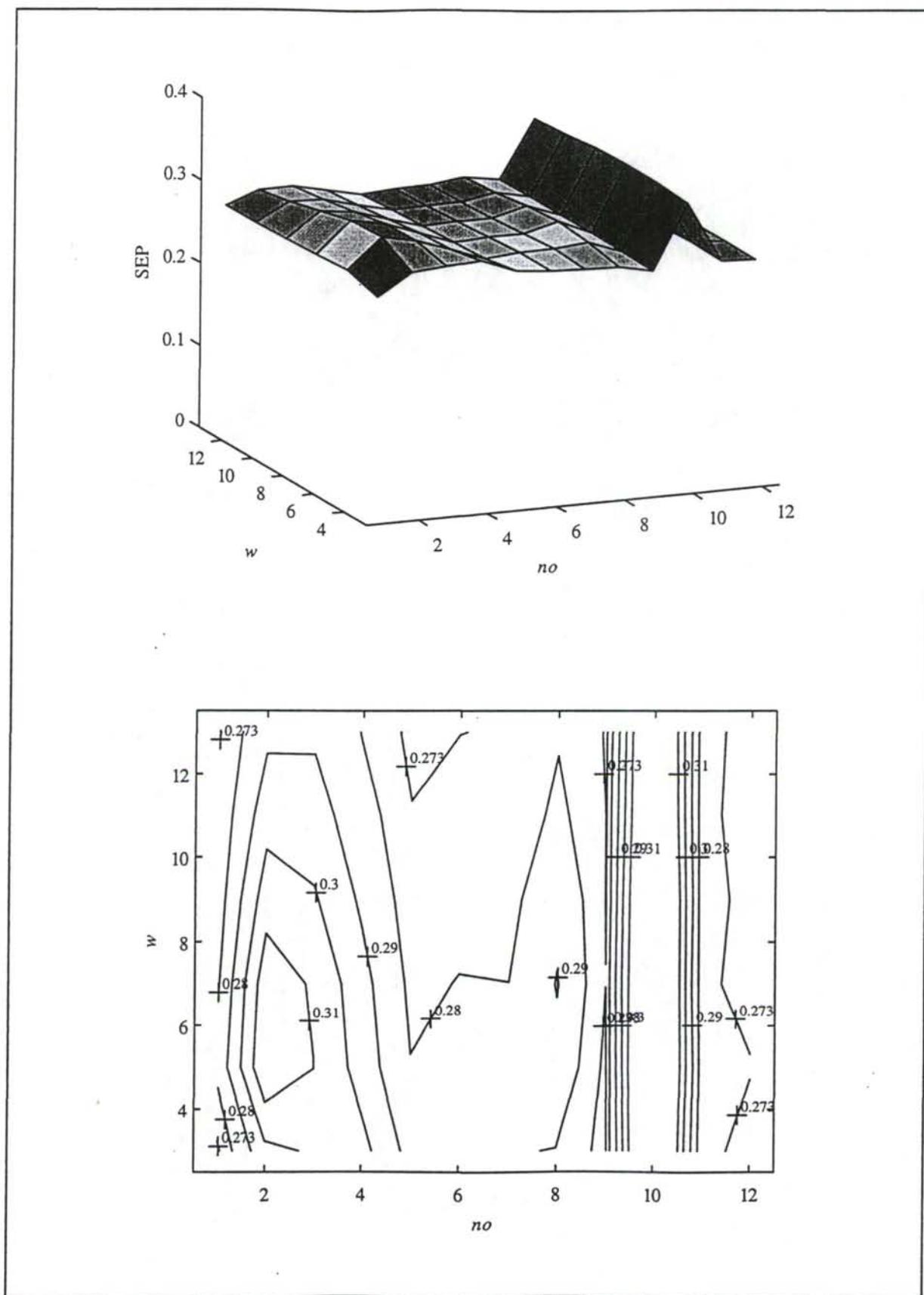


Figura III.3.7 Superficie de respuesta y curvas de nivel del SEP para el cadmio.

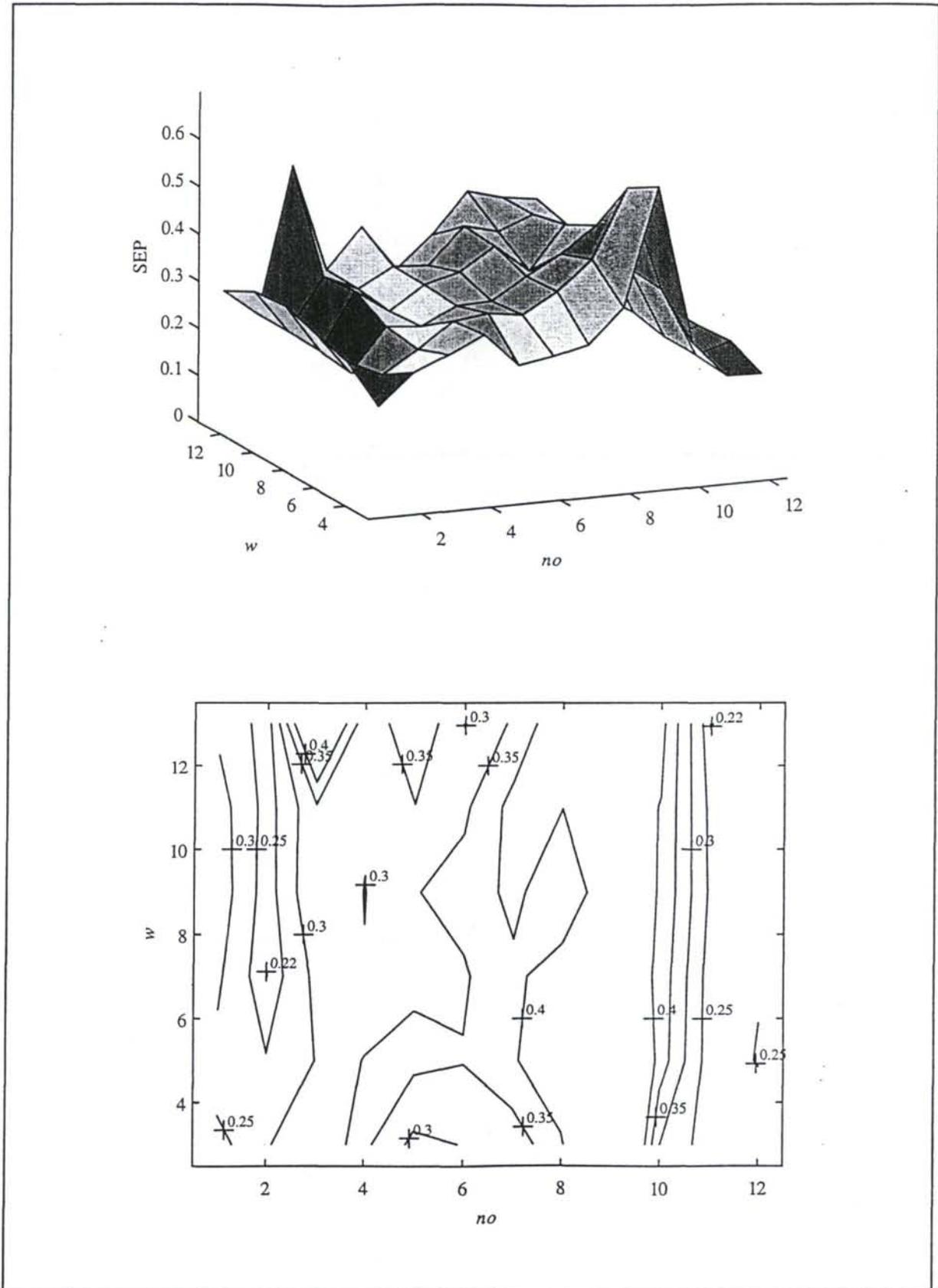


Figura III.3.8 Superficie de respuesta y curvas de nivel del SEP para el zinc.

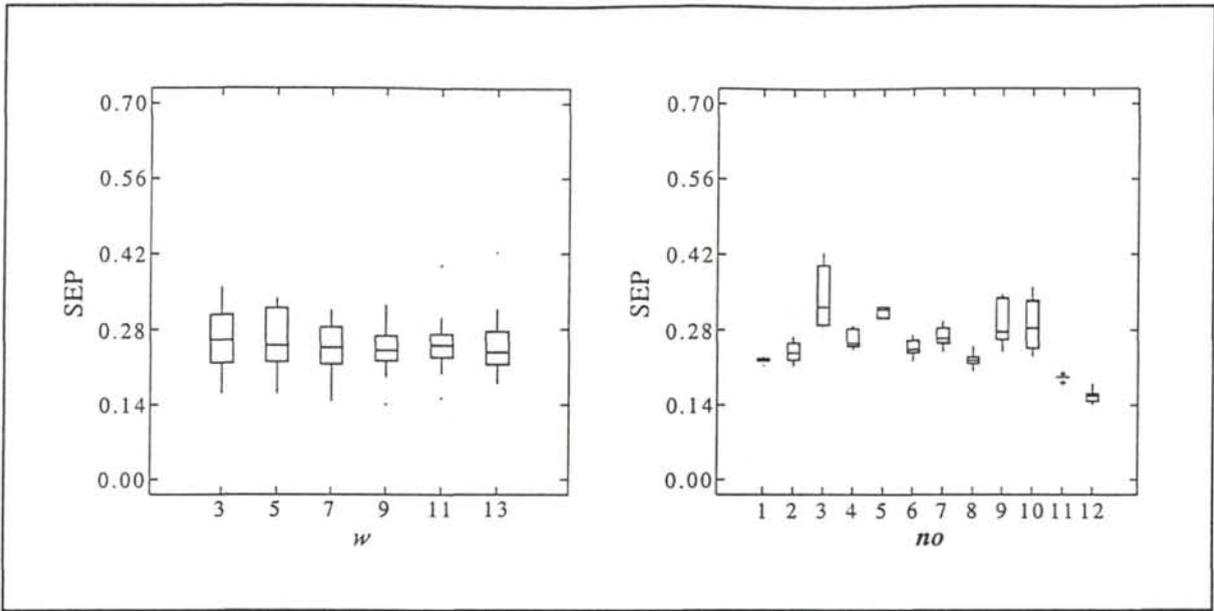


Figura III.3.9 Diagramas Box plot correspondientes a los valores de SEP obtenidos en la transferencia de calibrado del cobre.

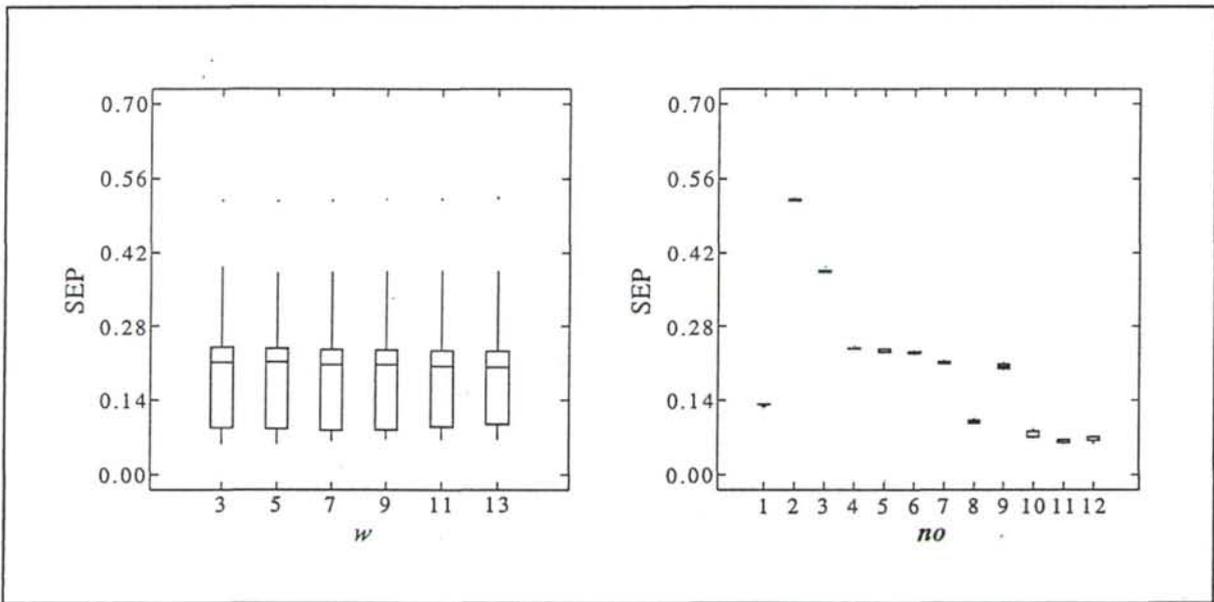


Figura III.3.10 Diagramas Box plot correspondientes a los valores de SEP obtenidos en la transferencia de calibrado del plomo.

que ambos test sean más sensibles con el cadmio que con los otros tres analitos. No obstante, los correspondientes diagramas Box plot muestran cómo los valores de SEP no son muy diferentes. En base a estas consideraciones y al hecho de que utilizar un mismo tamaño de ventana para todos los analitos permite calcular una única matriz de transferencia y realizar

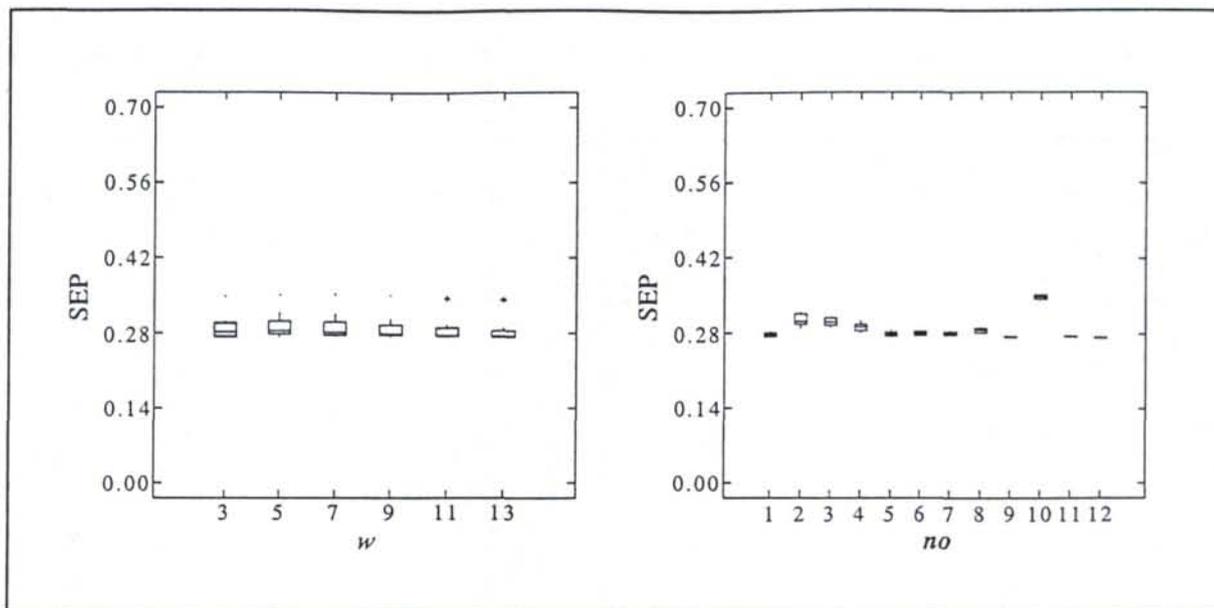


Figura III.3.11 Diagramas Box plot correspondientes a los valores de SEP obtenidos en la transferencia de calibrado del cadmio.

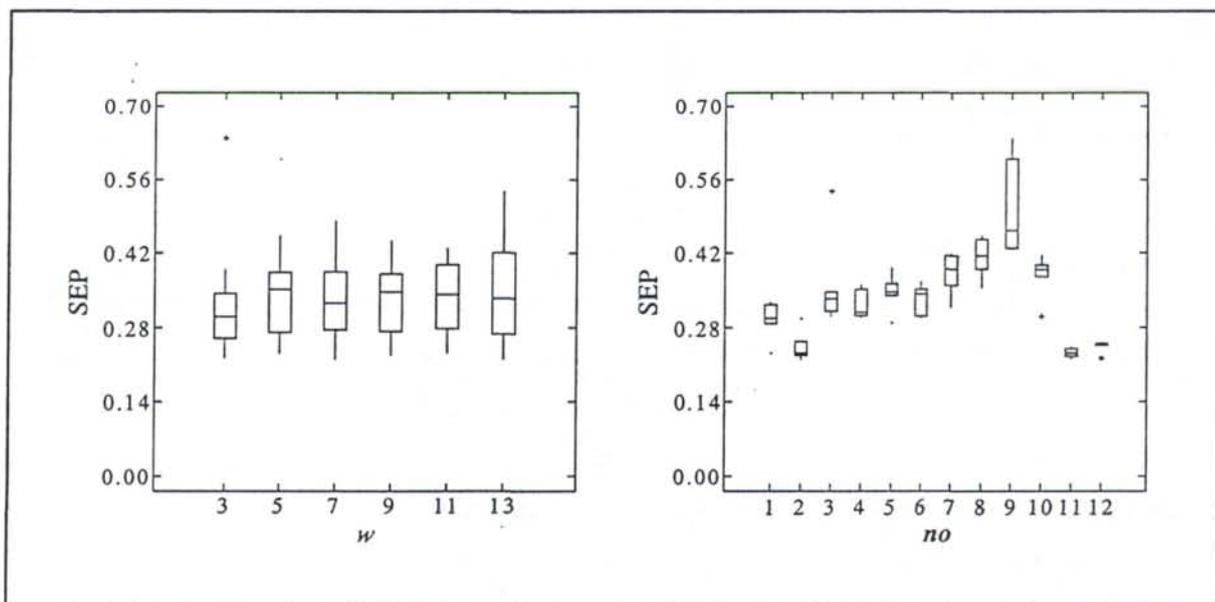


Figura III.3.12 Diagramas Box plot correspondientes a los valores de SEP obtenidos en la transferencia de calibrado del zinc.

la transferencia de calibrado completa para todos ellos a un tiempo, se ha seleccionado para los cuatro analitos la anchura de ventana 5, la cual se encuentra asociada con valores bajos de SEP para varios conjuntos de transferencia, como se puede apreciar en las curvas de nivel correspondientes.

Tabla III.3.9 Test de agrupamientos de Tukey y de Newman-Keuls para los valores del SEP cuando se calcula la concentración de cobre en el calibrado B.

Conjunto de transferencia (no)				Ventana (w)			
Nivel	Media	Grupos homogéneos		Nivel	Media	Grupos homogéneos	
		Tukey	Newman-Keuls			Tukey	Newman-Keuls
12	0.1571	X	X	9	0.2421	X	X
11	0.1910	XX	X	7	0.2451	X	X
1	0.2228	XX	XX	11	0.2543	X	X
8	0.2237	XX	XX	13	0.2543	X	X
2	0.2378	XXX	X	5	0.2593	X	X
6	0.2455	XXX	XX	3	0.2630	X	X
4	0.2607	XXX	XX				
7	0.2663	XXX	XX				
10	0.2890	XXX	XX				
9	0.2893	XXX	XX				
5	0.3136	XX	XX				
3	0.3393	X	X				

Para evaluar el error estándar del conjunto de aprendizaje, SET, se han considerado las concentraciones correspondientes a los polarogramas de las muestras del calibrado B que constituyen dicho conjunto, calculadas a partir de los modelos PLS construidos con los datos del calibrado A para los distintos analitos. El conjunto de aprendizaje utilizado para construir el modelo PLS del calibrado A está formado, en todos los casos, por las 28 muestras de calibrado que no forman parte del conjunto test, y algunas de ellas, dependiendo del valor del parámetro no , cabe la posibilidad de que hayan entrado a formar parte del conjunto de transferencia.

Las muestras incluidas en cada conjunto de transferencia deberían ser suficientemente estables para que fuera posible interpretar la variabilidad entre los datos obtenidos en ambas situaciones como diferencias entre situaciones y no como diferencias entre las propias muestras. Otra característica decisiva de cada conjunto de transferencia reside en lo representativas que sean las muestras que lo componen respecto del conjunto de datos completo.

Tabla III.3.10 Test de agrupamientos de Tukey y de Newman-Keuls para los valores del SEP cuando se calcula la concentración de plomo en el calibrado B.

Conjunto de transferencia (<i>no</i>)				Ventana (<i>w</i>)			
Nivel	Media	Grupos homogéneos		Nivel	Media	Grupos homogéneos	
		Tukey	Newman-Keuls			Tukey	Newman-Keuls
11	0.0634	X	X	7	0.2039	X	X
12	0.0675	XX	X	9	0.2043	X	X
10	0.0746	X	X	5	0.2048	X	X
8	0.1006	X	X	11	0.2048	X	X
1	0.1319	X	X	13	0.2051	X	X
9	0.2033	X	X	3	0.2053	X	X
7	0.2108	X	X				
6	0.2294	X	X				
5	0.2325	X	XX				
4	0.2375	X	X				
3	0.3847	X	X				
2	0.5203	X	X				

A este respecto, las curvas de nivel señalan claramente los conjuntos de transferencia 11 y 12 como aquellos conjuntos que, para todos los analitos simultáneamente, proporcionan valores de SEP mínimos. Los test de agrupamientos encuentran valores de SEP con diferencias estadísticamente significativas entre los distintos niveles analizados, observándose a través de ellos también cuáles son los conjuntos de transferencia que conducen a los mejores resultados.

En la Tabla III.3.13 se muestran los valores de SEP y SET obtenidos con los distintos conjuntos de transferencia (*no*) utilizados en el cálculo de la matriz de transferencia cuando se ha considerado un tamaño de ventana $w = 5$. Además, al final de dicha tabla se incluyen los valores de SEP y SET correspondientes a los modelos PLS construidos con los conjuntos de aprendizaje de los calibrados A y B, anotados como Calibrado A y Calibrado B respectivamente, que permiten a su vez establecer comparaciones con aquellas situaciones en las que no se realizó la transferencia de calibrado.

Tabla III.3.11 Test de agrupamientos de Tukey y de Newman-Keuls para los valores del SEP cuando se calcula la concentración de cadmio en el calibrado B.

Conjunto de transferencia (<i>no</i>)				Ventana (<i>w</i>)			
Nivel	Media	Grupos homogéneos		Nivel	Media	Grupos homogéneos	
		Tukey	Newman-Keuls			Tukey	Newman-Keuls
12	0.2721	X	X	13	0.2821	X	X
9	0.2722	X	X	11	0.2847	XX	XX
11	0.2743	X	X	9	0.2882	XX	XX
1	0.2762	XX	X	3	0.2890	XX	XX
5	0.2776	XX	XX	7	0.2914	X	X
7	0.2784	XX	XX	5	0.2928	X	X
6	0.2786	XX	XX				
8	0.2843	XX	X				
4	0.2906	X	X				
3	0.3006	X	X				
2	0.3037	X	X				
10	0.3478	X	X				

Se observa que entre los valores que aparecen la Tabla III.3.13 no existen grandes diferencias entre SEP y SET en la mayoría de los casos, excepto para algún conjunto de transferencia concreto. Sin embargo, si se consideran los valores de SEP y SET correspondientes a los propios modelos PLS que se construyeron con los calibrados completos, que se encuentran al final de dicha tabla, se observa que los errores asociados con las muestras del conjunto de aprendizaje son menores que aquellos correspondientes a las muestras del conjunto test. Esto es lógico si se tiene en cuenta que un modelo tiende a ajustarse a la estructura que presenten las muestras utilizadas en su construcción, mientras que las muestras correspondientes al conjunto test son, en cualquier caso, ajenas al mismo, luego es habitual que los errores asociados con estas últimas sean mayores.

Teniendo en consideración de los valores de la tabla III.3.13 se deduce que no todos los conjuntos de transferencia conducen a los mismos resultados. De hecho, los conjuntos que presentan unos valores de SEP sensiblemente inferiores al resto son los anotados con los números 11, 12 y 1, como aparece reflejado en los diagramas Box plot correspondientes.

Tabla III.3.12 Test de agrupamientos de Tukey y de Newman-Keuls para los valores del SEP cuando se calcula la concentración de zinc en el calibrado B.

Conjunto de transferencia (<i>no</i>)				Ventana (<i>w</i>)			
Nivel	Media	Grupos homogéneos		Nivel	Media	Grupos homogéneos	
		Tukey	Newman-Keuls			Tukey	Newman-Keuls
11	0.2313	X	X	3	0.3251	X	X
2	0.2428	XX	X	7	0.3322	X	X
12	0.2439	XX	X	9	0.3347	X	X
1	0.2941	XXX	XX	11	0.3356	X	X
4	0.3223	XXXX	XX	13	0.3478	X	X
6	0.3343	XXX	XXX	5	0.3513	X	X
5	0.3469	XX	XXX				
3	0.3609	XX	XXX				
10	0.3784	XX	XX				
7	0.3822	XX	XX				
8	0.4129	XX	X				
9	0.5033	X	X				

El primer conjunto de transferencia se corresponde con todas las muestras del conjunto de aprendizaje, por lo tanto, en estas condiciones la transferencia de calibrado no supondría ninguna mejora o disminución del esfuerzo requerido en la experimentación del procedimiento. Sin embargo, en el conjunto 11 se incluyen, para cada analito, únicamente las tres muestras del conjunto de aprendizaje que se encuentran dispuestas sobre los ejes que atraviesan las caras del hipercubo que configuran las muestras de calibrado, Figura III.3.1, lo que supone un total de 12 muestras. El conjunto 12 es muy similar a este último, pero eliminando del mismo el punto central de cada uno de esos ejes, con lo que dicho conjunto de transferencia queda reducido a 8 muestras.

Una forma de evaluar con más detalle la validez de estos tres conjuntos de transferencia se lleva a cabo a través del cálculo de los porcentajes de error relativo de las muestras del conjunto test, que se muestran en la Tabla III.3.14.

Comparando los errores que aparecen en estas tres tablas con los calculados por los modelos PLS del calibrado B, recogidos en la Tabla III.3.7, que se corresponden exactamente

Tabla III.3.13 Valores de SEP y SET en función del conjunto de transferencia utilizado en el cálculo de la matriz de transferencia (tamaño de ventana: 5). Como Calibrado B y Calibrado A se anotan los valores correspondientes a los modelos PLS construidos con los conjuntos de aprendizaje de los calibrados B y A respectivamente.

Conjunto (no)	SEP				SET			
	Cobre	Plomo	Cadmio	Zinc	Cobre	Plomo	Cadmio	Zinc
1	0.2222	0.1329	0.2827	0.2874	0.1952	0.2004	0.2095	0.3465
2	0.2230	0.5175	0.3190	0.2532	0.1751	0.4466	0.2414	0.2336
3	0.3175	0.3811	0.3101	0.3004	0.4238	0.3238	0.2443	0.4523
4	0.2409	0.2376	0.2959	0.3525	0.2758	0.2222	0.2257	0.3688
5	0.3228	0.2360	0.2803	0.3631	0.3561	0.2235	0.2116	0.4128
6	0.2594	0.2342	0.2833	0.3526	0.2438	0.2229	0.2116	0.4072
7	0.2829	0.2158	0.2837	0.3933	0.2579	0.2153	0.2128	0.4745
8	0.2174	0.1027	0.2871	0.4551	0.2765	0.2147	0.2208	0.4212
9	0.3387	0.2076	0.2716	0.5995	0.6173	0.2260	0.2057	0.9636
10	0.3337	0.0702	0.3511	0.3765	0.3171	0.2266	0.3031	0.4527
11	0.1916	0.0582	0.2756	0.2302	0.1871	0.2407	0.2054	0.3295
12	0.1613	0.0637	0.2732	0.2523	0.1883	0.2590	0.2010	0.2891
Calibrado B	0.1298	0.2570	0.2551	0.1563	0.0536	0.0488	0.1742	0.1211
Calibrado A	0.1111	0.2319	0.1629	0.2879	0.0941	0.0695	0.1240	0.1495

con los mismos objetos test, se observa que son del mismo orden. Es decir, que una vez llevada a cabo la transferencia de calibrado, los resultados obtenidos a través de ella son similares a aquellos correspondientes a una recalibración completa. En este caso, si se toma como referencia el conjunto de transferencia 1, no existen diferencias entre el esfuerzo experimental invertido con transferencia de calibrado o sin ella, ya que en este conjunto se incluyen todas las muestras de calibrado excepto las del conjunto test.

Sin embargo, si se considera cualquiera de los otros dos conjuntos de transferencia, el 11 o el 12, la reducción del esfuerzo experimental por la aplicación de este procedimiento de transferencia de calibrado, comparada con una recalibración completa del instrumento, es importante. Las muestras empleadas para calcular la matriz de transferencia son 12 y 8 respectivamente, frente a las 28 que constituyen el conjunto de aprendizaje. Además, en ambos casos el número de soluciones iniciales (blancos experimentales) requeridas para

Tabla III.3.14 Errores relativos (en %) correspondientes a las muestras del conjunto test en la transferencia de calibrado llevada a cabo mediante el procedimiento PDS.

<i>no</i>	Número de muestra	Analito			
		Cobre	Plomo	Cadmio	Zinc
I	18	11.08	1.32	-3.25	3.60
	20	4.50	3.20	-0.72	2.13
	23	2.91	-0.02	-4.24	-8.40
	25	2.79	1.56	-4.23	-9.80
	28	-0.82	2.51	0.76	1.76
	30	-1.02	3.74	2.39	-2.24
	33	-1.87	2.06	3.01	-4.26
	35	-2.89	1.53	3.36	-0.90
11	18	7.21	-0.38	-1.41	4.03
	20	2.27	1.18	-0.07	5.41
	23	0.31	-2.47	-3.94	-4.87
	25	0.43	-0.63	-4.23	-5.92
	28	-2.38	0.89	1.41	3.63
	30	-2.42	1.34	2.80	2.29
	33	-3.27	0.28	3.36	-1.37
	35	-5.91	-0.80	3.76	1.84
12	18	8.03	-0.63	-1.51	3.77
	20	1.81	0.54	-0.24	6.59
	23	1.04	-3.19	-4.08	-3.13
	25	0.48	-1.12	-4.36	-4.29
	28	-2.03	0.54	1.21	4.36
	30	-0.75	0.83	2.63	4.46
	33	-2.28	-0.06	3.19	3.97
	35	-4.21	1.21	3.62	3.68

calcular la matriz de transferencia es cuatro, frente a las diez que son necesarias para llevar a cabo el calibrado completo.

Esto significa que, una vez completado el calibrado A, únicamente hubieran sido necesarias esas ocho o doce muestras del calibrado B, dependiendo del conjunto de transferencia seleccionado, para calcular la matriz de transferencia, y posteriormente hubiera sido posible predecir todas y cada una de las concentraciones de los cuatro analitos de las muestras del calibrado B restantes.

En el caso del conjunto de transferencia 12 bastaría con realizar únicamente dos adiciones de cada uno de los analito para poder transferir el calibrado. Este valor coincide con el número mínimo de adiciones que serían necesarias si se tratara el problema desde un punto de vista univariante clásico, en un análisis llevado a cabo de una forma muy simple. Sin embargo, al utilizar la transferencia de calibrado multivariante, con idéntico esfuerzo experimental, se llega a resultados muy buenos avalados por toda una metodología multivariante, en la obtención de los cuales se han tenido en consideración las interferencias existentes entre los analitos.

A través de la transferencia de calibrado es posible utilizar el esfuerzo experimental que requiere un calibrado completo, y toda la información que éste aporta, a calibrados realizados posteriormente, con un esfuerzo experimental mucho menor y con una enorme garantía en los resultados obtenidos.

III.3.4.3 Transferencia de calibrado mediante el procedimiento GCT

Se ha utilizado este procedimiento de transferencia de calibrado diferente para llevar a cabo una nueva transferencia de calibrado entre el calibrado A y el calibrado B, por tanto, continua tratándose de una transferencia de calibrado en el tiempo. El objetivo de este procedimiento, como se ha señalado anteriormente, sigue siendo una reducción del esfuerzo experimental que será necesario invertir para realizar calibrados posteriores. Por ello, también en este caso se han utilizado diferentes conjuntos de transferencia para llevar a cabo el análisis, y se han evaluado los resultados obtenidos con cada uno de ellos.

Al utilizar este procedimiento de transferencia de calibrado no tiene sentido considerar el concepto de tamaño de ventana, ya que éste es un método de transferencia de calibrado global, que relaciona los modelos de regresión de las diversas situaciones considerando las

señale completas (aunque no es necesario que el número de variables predictoras medidas en ambas situaciones sea el mismo).

Función "gct"

El procedimiento de transferencia de calibrado global (GCT), propuesto por Forina y otros [102] para realizar transferencias de calibrado en espectroscopía NIR, se ha aplicado a través de una función (m-function) que se ha programado a tal efecto en el programa MATLAB [59]. Se trata de la función que se ha denominado "gct", a través de la cual, como se señaló anteriormente, se desarrolla en primer lugar un modelo PLS que relaciona los dos conjuntos de variables predictoras, y que permite calcular una matriz de transferencia. Esta matriz de transferencia permite estimar, considerando el modelo PLS construido previamente con los datos del calibrado patrón y la concentración correspondiente a uno de los analitos, los coeficientes de ese modelo de regresión para la situación satélite, y a partir de ahí, calcular las concentraciones correspondientes a las nuevas muestras.

Esta función presenta el formato general de entrada/salida que se indica a continuación:

```
[standmat, subset] = gct(spec1, spec2, lv, nosamps, split, cer)
```

donde *spec1* y *spec2* son los bloques de los polarogramas correspondientes al calibrado patrón y al calibrado satélite respectivamente, *lv* es el número máximo de variables latentes a considerar en el cálculo de la matriz de transformación de los coeficientes, *nosamps* es el vector en el que se especifican los índices de las muestras que forman parte del conjunto de transferencia, *split* es el número de conjuntos de cancelación que se desea utilizar, y por último, *cer* es el vector de los coeficientes del modelo PLS del calibrado patrón. El parámetro de salida *standmat* es la matriz de transformación de los coeficientes, o matriz de transferencia, mientras que *subset* contiene los índices de las muestras del conjunto de transferencia, coincidiendo con *nosamps*.

Las filas de *spec1* y *spec2* han de encontrarse dispuestas en el mismo orden pero, a diferencia de lo ocurría en el procedimiento PDS, en este caso no es necesario que tengan el mismo número de variables predictoras. Por lo tanto, en aquellos casos en los que se disponga de diferente número de predictoras en ambas situaciones, la matriz de transferencia resultante del procedimiento de transferencia de calibrado no será una matriz cuadrada.

La matriz de transformación de los coeficientes así calculada se utiliza para corregir los coeficientes de la fórmula cerrada de la situación patrón, estimando cuál sería su valor en la situación satélite, con lo que se tiene como resultado una matriz de coeficientes corregidos. Una vez que se dispone de los coeficientes corregidos para la situación satélite, es posible calcular la concentración de las muestras de esa situación.

Aplicación de la función "gct"

Para llevar a cabo la transferencia de calibrado de la situación A a la situación B, aplicando el procedimiento indicado, que se representa de forma esquemática en la Figura III.3.13, se han utilizado las matrices de datos **A** y **B**, las mismas que en el procedimiento de transferencia de calibrado PDS realizado anteriormente, así como los subconjuntos de transferencia que se especifican en la Tabla III.3.1. De esta forma es posible evaluar la influencia que tienen los distintos conjuntos de transferencia propuestos en el procedimiento de transferencia de calibrado GCT, seleccionando posteriormente aquel que proporcione los mejores resultados e implique un esfuerzo experimental razonable.

En la función "gct" se han introducido los siguientes parámetros:

$$[F, sub] = gct(A, B, 15, no, 3, Acer)$$

En donde los parámetros **A**, **B**, **no** y **sub** son idénticos a los utilizados en el procedimiento de transferencia PDS, es decir, los dos primeros son los polarogramas correspondientes a los calibrados A y B respectivamente, mientras que **no** y **sub** son los vectores que especifican la composición de cada conjunto de transferencia, correspondiéndose con los indicados en la Tabla III.3.8. Por otro lado, el valor 15 señala el número máximo de variables latentes que se han de seleccionar en el modelo PLS construido para calcular la matriz de transferencia, el valor 3 indica el número de grupos de cancelación considerados en la validación cruzada de dicho modelo PLS, y **F** es la matriz de transformación de los coeficientes o matriz de transferencia, que tiene de dimensiones 203×203 .

Cada uno de los conjuntos de transferencia utilizados en este análisis conduce a la obtención de una matriz de transformación de los coeficientes diferente, lo cual supone que se han realizado 12 procedimientos de transferencia de calibrado independientes, tantos como conjuntos de transferencia distintos.

Las matrices de transformación de los coeficientes, **F**, permiten estimar la fórmula cerrada de la situación B a partir de los correspondientes coeficientes del modelo PLS

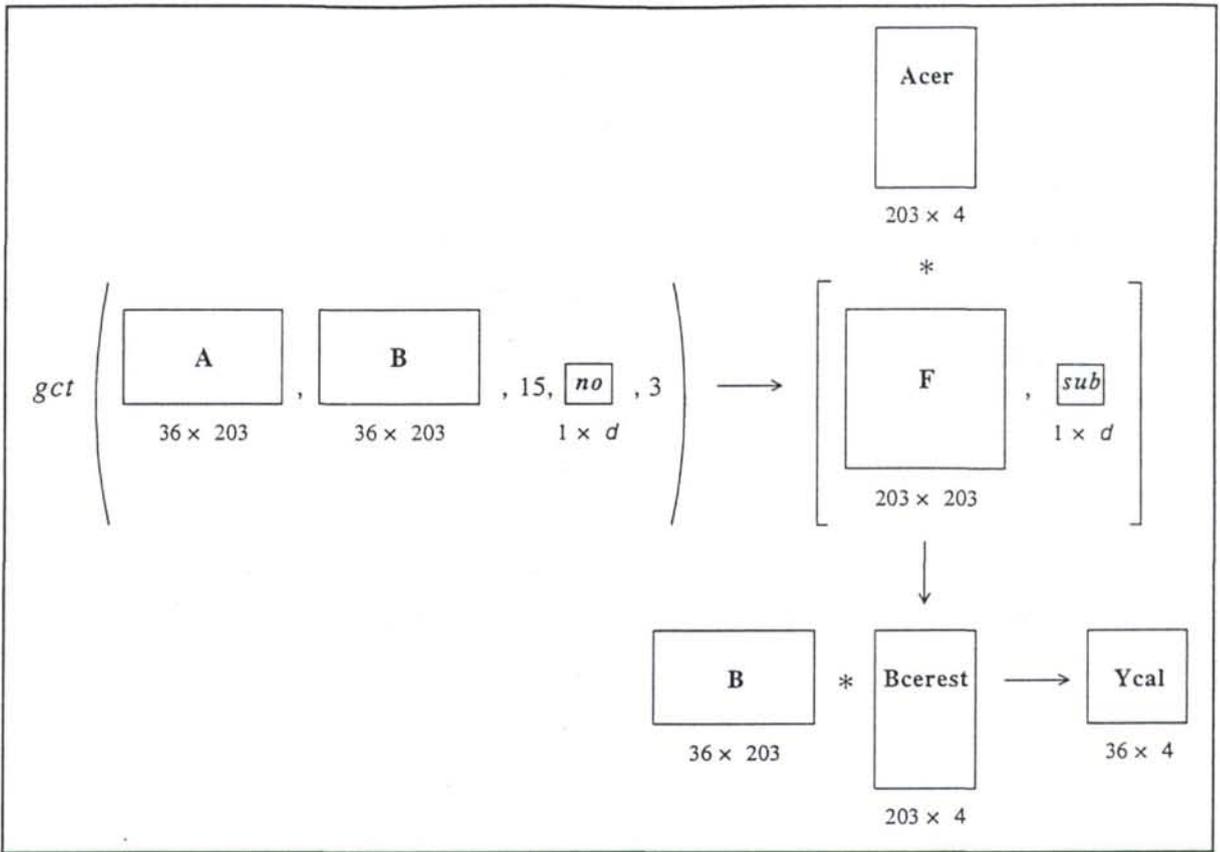


Figura III.3.13 Esquema completo del proceso de transferencia de calibrado GCT y del cálculo de los valores de concentración.

construido con las muestras de calibrado de la situación A, *Acer*, mediante la siguiente relación

$$Bcerest = F * Acer$$

donde *Bcerest* es el vector de los coeficientes estimados para la situación B. Estos coeficientes transformados permiten calcular la concentración de analito de cualquier muestra medida en la situación B a través de

$$Ycal = C * Bcerest$$

donde *Ycal* es la matriz en la que se encuentran los valores de concentración calculados para las muestras del calibrado B.

Este procedimiento de transferencia de calibrado, así como el cálculo de los valores correspondientes a SEP y SET, se realiza de forma automática utilizando la función '*transfpds*', programada en MATLAB, que se muestra en el Apéndice 2. El procedimiento completo ha sido utilizado con cada uno de los conjuntos de transferencia de la Tabla III.3.8,

por lo que se han realizado un total de 12 transferencias de calibrado de forma independiente para cada uno de los analitos.

Análisis de los resultados

Los resultados de los diferentes procedimientos de transferencia de calibrado GCT llevados a cabo se han evaluado también a través de los valores de SEP y SET correspondientes, los cuales se representan en las Figuras III.3.14 a III.3.17, para cada uno de los metales. En dichas figuras se puede observar cómo los conjuntos de transferencia que conducen a los errores más bajos, tanto si se considera el SEP como el SET, son los conjuntos 1 y 7, y también, para plomo y cadmio, el conjunto 11.

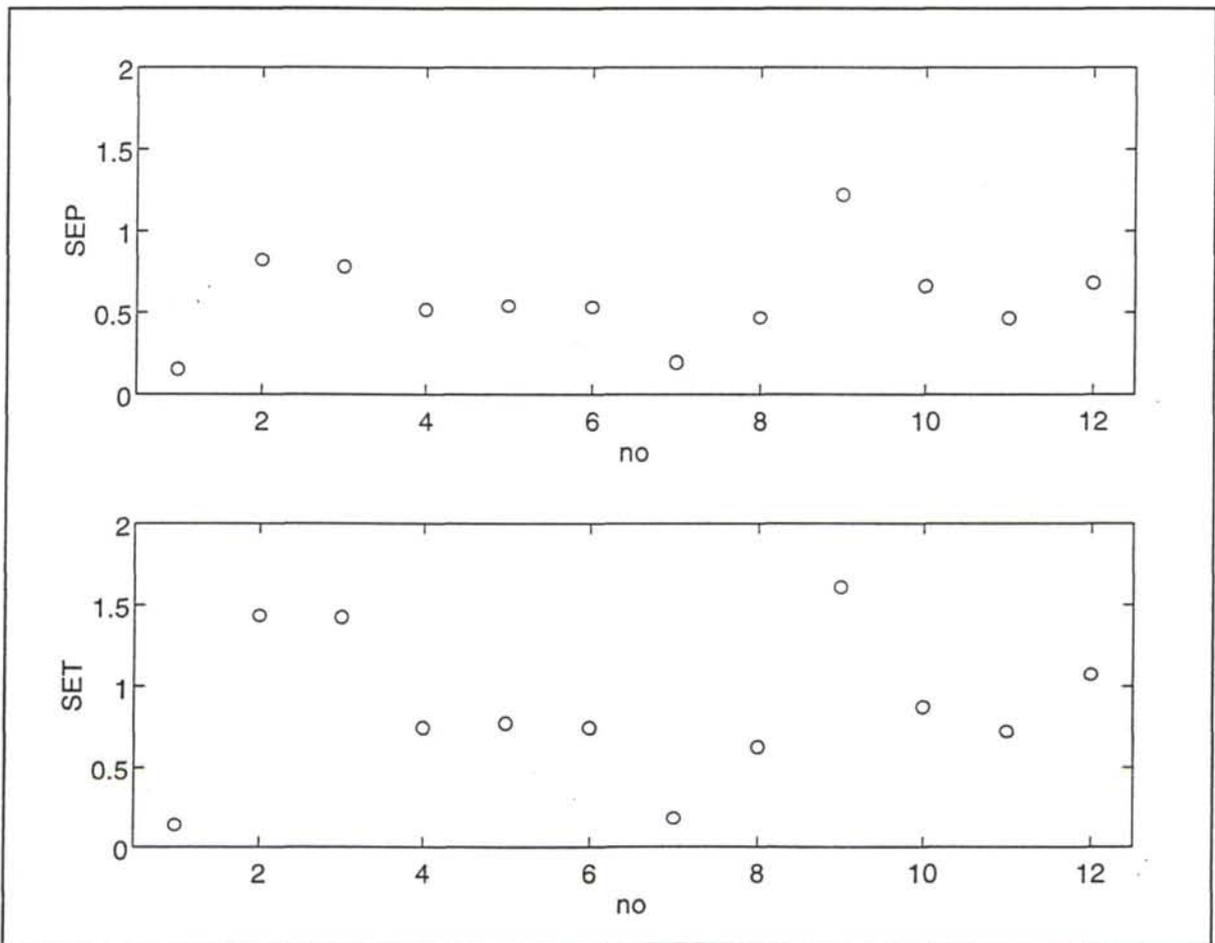


Figura III.3.14 Resultados de la transferencia de calibrado realizada con el procedimiento GCT para el cobre.

Es necesario considerar, por un lado, que el conjunto de transferencia 1 está formado por las 28 muestras del conjunto de aprendizaje del calibrado A, mientras que el conjunto 7

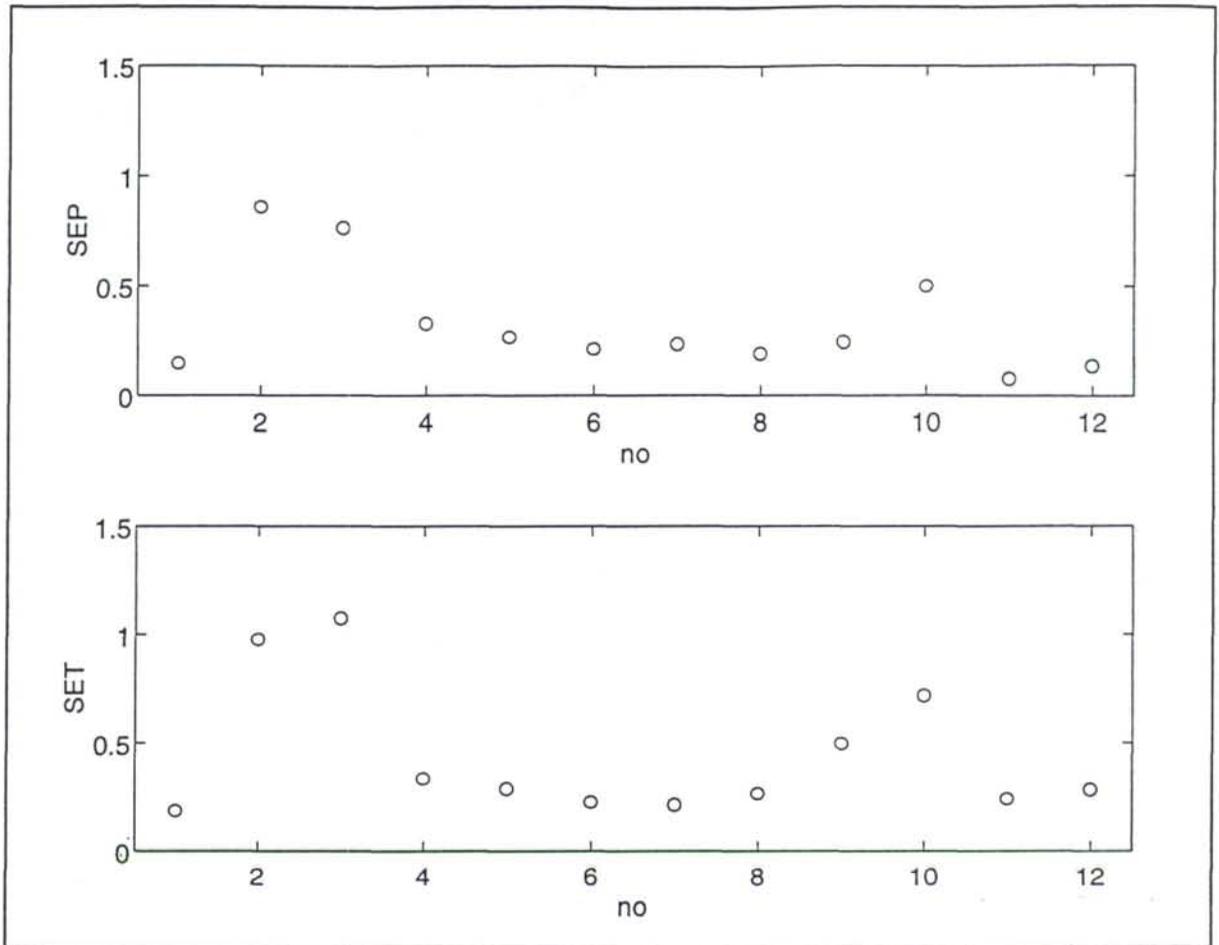


Figura III.3.15 Resultados de la transferencia de calibrado realizada con el procedimiento GCT para el plomo.

está constituido por las 16 primeras muestras de la Tabla III.3.1, y el conjunto 11 únicamente por 12 de las muestras patrón. También hay que tener en cuenta, no sólo el tamaño de los conjuntos de transferencia, si no también la calidad de los resultados obtenidos con ellos a través del procedimiento de transferencia. En la Tabla III.3.15 se muestran los errores de las muestras del conjunto test que se han obtenido como resultado de llevar a cabo el procedimiento de transferencia de calibrado GCT con los tres conjuntos de transferencia indicados.

Los errores correspondientes a $no = 1$ son, en conjunto, los mejores, incluso comparables con los obtenidos por los modelos PLS construidos para el calibrado B, que se muestran en la Tabla III.3.7. No obstante, la transferencia de calibrado llevada a cabo con este conjunto de transferencia equivale a una recalibración completa en cuanto al esfuerzo experimental que requiere y, por tanto, no resulta adecuado dada la finalidad de este análisis.

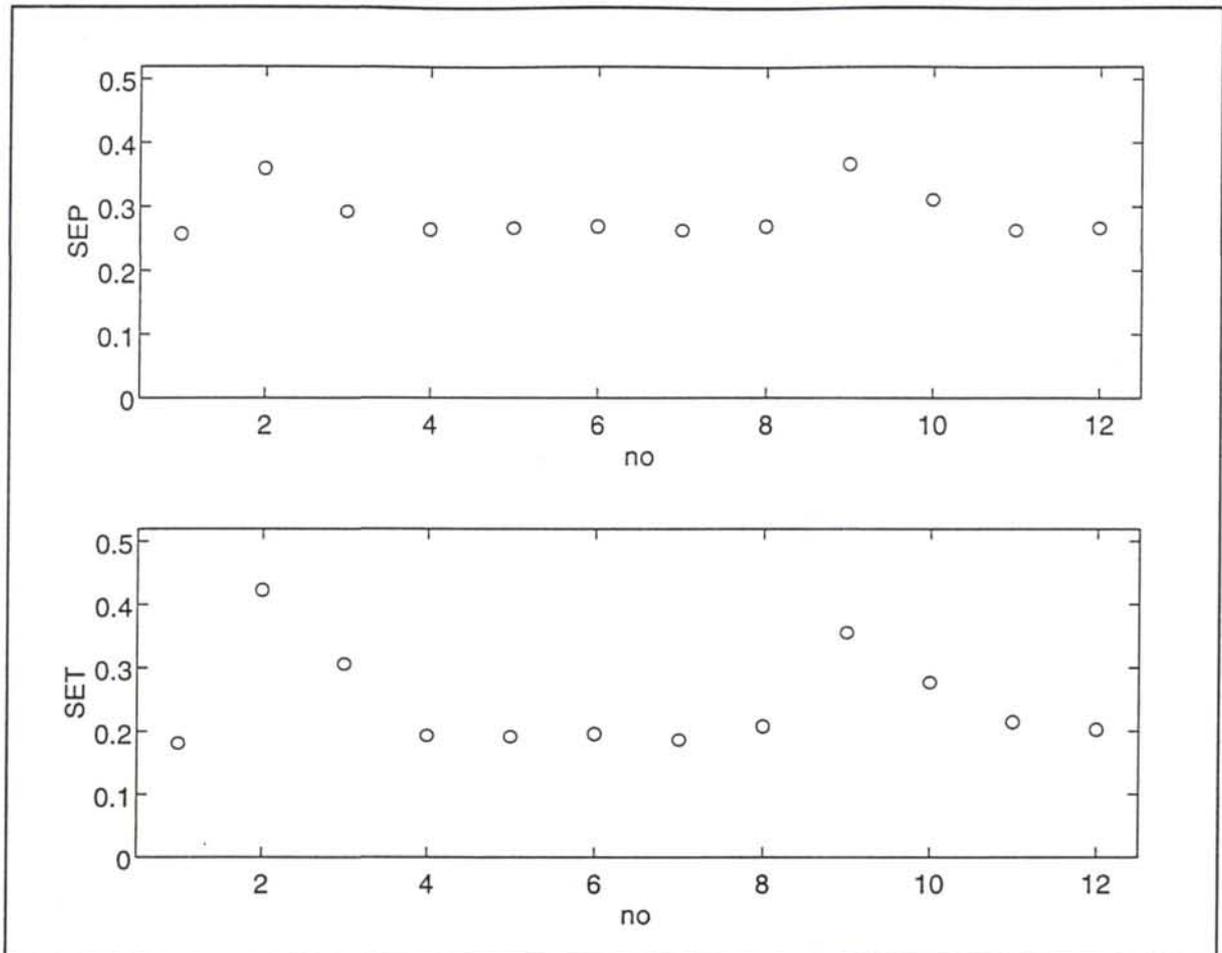


Figura III.3.16 Resultados de la transferencia de calibrado realizada con el procedimiento GCT para el cadmio.

El conjunto de transferencia 7 supone reducir en un 43% el esfuerzo experimental necesario para llevar a cabo el análisis, obteniéndose unos resultados bastante buenos. Sin embargo, el conjunto de transferencia 11, conduce a resultados excelentes para plomo y cadmio, pero no así para cobre y zinc, metales para los que se llegan a obtener errores cercanos al 20%.

En vista de estos resultados, el conjunto de transferencia óptimo para llevar a cabo la transferencia de calibrado mediante el procedimiento GCT, es el conjunto 7. Los errores asociados con este conjunto de transferencia son ligeramente superiores a los obtenidos con el procedimiento de transferencia PDS utilizado en primer lugar (con los parámetros $no = 12$ y $w = 5$), los cuales se encuentran en la Tabla III.3.14. Esto es cierto en todos los casos, excepto en lo que se refiere al zinc, analito para el que el procedimiento GCT ofrece resultados un poco mejores.

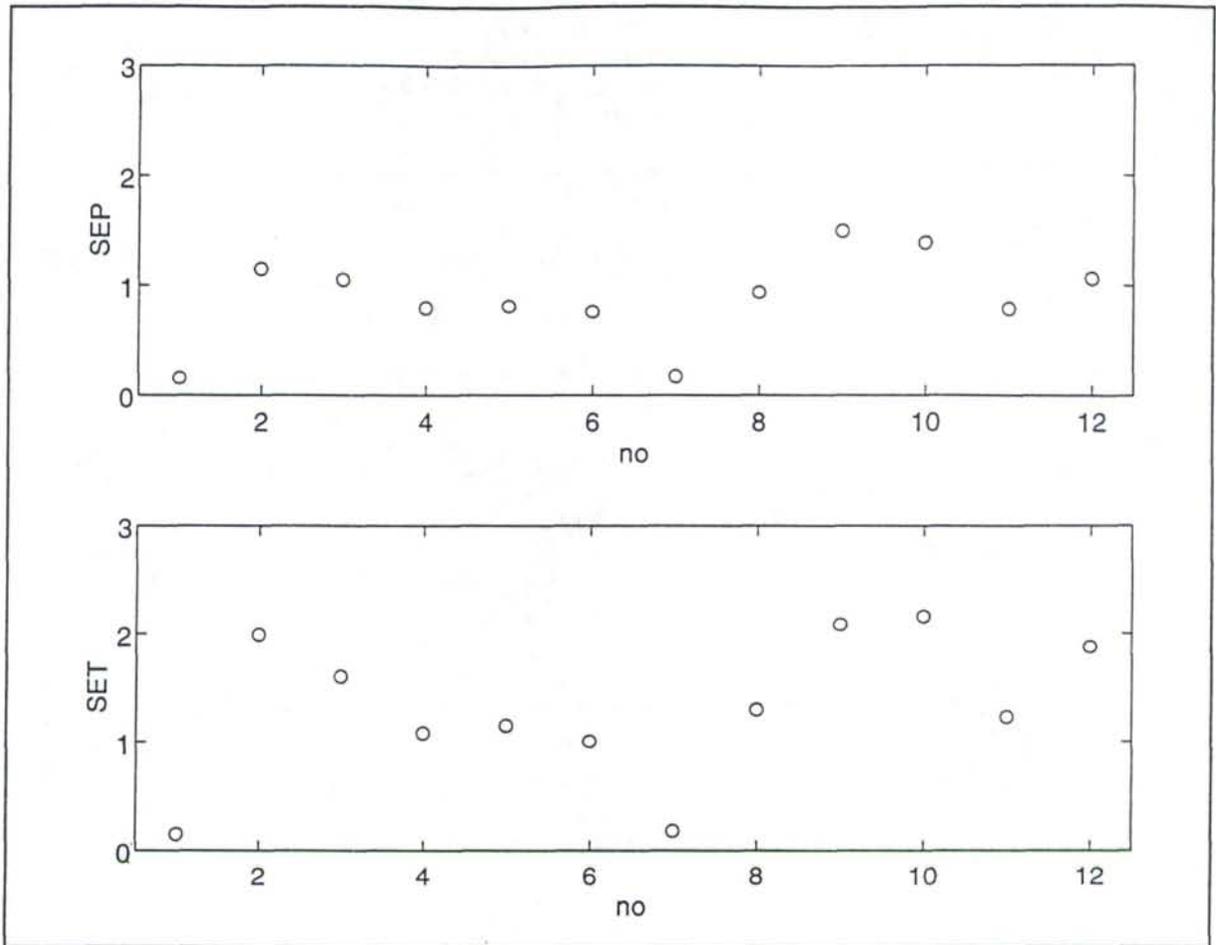


Figura III.3.17 Resultados de la transferencia de calibrado realizada con el procedimiento GCT para el zinc.

Sin embargo, un aspecto que se ha destacado bastante y que resulta un parámetro decisivo para seleccionar un procedimiento de transferencia de calibrado u otro viene dado por el esfuerzo experimental asociado con cada uno de ellos. En este análisis en concreto, las condiciones de transferencia del procedimiento PDS requerían únicamente 8 muestras patrón en la situación satélite, mientras que el correspondiente procedimiento GCT implica el doble de experimentación, 16 muestras en total. Por tanto, dado que los resultados obtenidos a través de ambos métodos son muy similares, habría que decantarse por utilizar el procedimiento de transferencia PDS, ya que es el que implica un número menor de muestras de transferencia.

Del análisis realizado, se podría concluir que el método de transferencia de calibrado PDS, con los parámetros de transferencia óptimos determinados en su momento, es el más indicado para realizar futuras determinaciones de cobre, plomo, cadmio y zinc en muestras

Tabla III.3.15 Errores relativos (en %) de las muestras del conjunto test en la transferencia de calibrado llevada a cabo mediante el procedimiento GCT.

<i>no</i>	Número de muestra	Analito			
		Cobre	Plomo	Cadmio	Zinc
1	18	4.64	1.84	-2.13	-3.51
	20	2.63	2.54	0.39	3.72
	23	2.35	1.74	-3.70	-3.45
	25	3.91	0.66	-3.40	-2.38
	28	-0.71	1.83	2.21	-0.54
	30	-0.36	4.98	1.93	2.71
	33	-3.16	1.73	2.91	2.68
	35	-1.87	2.94	4.29	1.33
7	18	8.36	3.45	-1.97	-4.90
	20	3.81	3.12	0.53	2.87
	23	2.54	3.78	-3.92	-3.93
	25	3.63	2.26	-4.75	-3.48
	28	1.90	2.66	2.63	-2.05
	30	-0.44	7.06	1.48	1.82
	33	-2.78	2.70	2.39	2.09
	35	-1.03	5.65	3.72	0.26
11	18	13.31	-0.77	-1.77	12.70
	20	0.49	2.04	0.07	3.20
	23	-7.66	-0.90	-4.28	-19.04
	25	-4.74	-0.99	-2.11	-16.21
	28	-3.32	0.68	1.85	1.87
	30	0.41	2.12	2.43	4.51
	33	18.72	0.79	3.94	17.515
	35	-18.66	-0.82	3.87	-17.28

acuosas mediante polarografía diferencial de impulsos en las condiciones experimentales utilizadas en el análisis.

III.4 INTERFERENCIA DE COMPUESTOS INTERMETÁLICOS EN DETERMINACIONES POR VOLTAMPEROMETRÍA DE REDISOLUCIÓN

Cuando varios metales se encuentran presentes simultáneamente sobre el electrodo de mercurio, es frecuente la formación de compuestos intermetálicos entre los metales disueltos en el mismo [127]. La formación de estos compuestos se suele producir después de la introducción de un metal más noble en el mercurio (algunos ejemplos: Ag-Zn, Au-Zn, Au-Cd, Cd-Zn, Co-Zn, Cu-Zn, Cu-Sn, Ni-Zn,...) y puede afectar considerablemente a las propiedades de la amalgama, e incluso, si la solubilidad del compuesto intermetálico en mercurio no es muy alta, puede llegar a separarse en una fase sólida originando una amalgama heterogénea.

Aquellos metales con potenciales de pico más negativos (como zinc, cadmio o indio) se ven afectados con más frecuencia por este tipo de interferencias, ello se debe a que como necesitan potenciales de acumulación muy negativos, a esos potenciales de acumulación muchos otros metales pueden permanecer disueltos en el electrodo de mercurio, dando lugar a la formación de compuestos intermetálicos.

La formación de estos compuestos intermetálicos en el electrodo puede reflejarse en la respuesta electroquímica que se registra, existiendo la posibilidad de que se produzca una disminución o un desplazamiento de la misma. Una consecuencia importante de la formación de compuestos intermetálicos es que puede dar lugar a que los resultados de un análisis llevado a cabo mediante el método de adición estándar sean incorrectos [128, 129].

El problema originado por la formación de compuestos intermetálicos es sustancial cuando se utilizan técnicas que requieren una etapa de preconcentración, lo cual es habitual en el análisis de metales [130-134]. Muchos estudios a este respecto están enfocados hacia el compuesto intermetálico formado por cobre y zinc, debido a la presencia de estos metales en gran diversidad de muestras analíticas (particularmente aquellas de interés medioambiental), así como el formado por cobre y cadmio. Las interferencias debidas a

ambos compuestos intermetálicos han sido evaluadas tanto en análisis realizados mediante técnicas potenciométricas con redisolución [135], como en aquellos llevados a cabo utilizando voltamperometría con redisolución anódica [136]. Otros autores [137] han encontrado que un exceso de cobre en solución tiene como resultado la disminución de la respuesta correspondiente al cadmio, hasta en un 58%, cuando éste es preconcentrado sobre un electrodo de película de mercurio y determinado mediante voltamperometría diferencial de impulsos con redisolución anódica. Si la concentración de los analitos en la muestra se encuentra alrededor de 1 ppm, ésta puede ser analizada utilizando polarografía diferencial de impulsos.

Se han sugerido varios métodos para minimizar o eliminar los errores debidos a la formación de compuestos intermetálicos. Frecuentemente este problema se reduce de forma significativa disminuyendo la concentración de los metales en la amalgama. Esto se puede llevar a cabo mediante el uso de electrodos de gota de mercurio en vez de aquellos de película de mercurio, debido a que el volumen de los primeros es mayor, o también mediante la utilización de técnicas electroanalíticas más sensibles, por ejemplo el modo diferencial de impulsos, que permitan acortar el tiempo de acumulación en aquellos casos en los que se utilicen técnicas que incluyen una etapa de preconcentración [138].

También la adición de un tercer elemento que forme un compuesto intermetálico más estable con uno de los componentes del sistema binario es una aproximación a la eliminación de errores debidos a la formación de compuestos intermetálicos [139, 140]. No obstante, la adición de este tercer elemento puede producir efectos negativos sobre la precisión de las determinaciones en aquellos casos en que llegue a interaccionar con los dos componentes del sistema binario [141]. Heineman y otros [142, 143] evitan la formación de compuestos intermetálicos utilizando dos electrodos de trabajo paralelos con acumulación selectiva, de forma que en uno de ellos se mide el cobre, y el zinc o el cadmio en el segundo.

Es posible además utilizar procedimientos experimentales que permitan la separación física de los metales a analizar. Un ejemplo claro de este tipo de análisis lo constituyen los métodos acoplados en los que se lleva a cabo la separación física de los analitos mediante algún tipo de cromatografía, al tiempo que la detección al final de la columna cromatográfica se produce utilizando técnicas electroquímicas [144, 145].

Por otro lado, en una aplicación de la técnica de análisis múltiple con redisolución anódica [146] se propone una corrección de la señal correspondiente al zinc mediante la

adición a la misma de la señal correspondiente al cobre, utilizando éste último como titulante del zinc. También se ha sugerido un método de adición estándar generalizado para caracterizar y corregir las interferencias debidas a la formación de compuestos intermetálicos en análisis con redisolución [147], pero en este caso es necesario conocer todos los interferentes.

Otra forma de abordar el problema se puede encontrar en la aplicación de una metodología multivariante en aquellos casos en que se sospecha que puede tener lugar la formación de compuestos intermetálicos, ya que algunas técnicas de regresión multivariante, como en el caso de PLS, son capaces de evidenciar y modelar las relaciones internas que pueden existir entre las respuestas electroquímicas correspondientes a analitos diferentes.

En el presente capítulo se ha llevado a cabo la determinación simultánea de cobre, plomo, cadmio y zinc en solución acuosa mediante voltamperometría diferencial de impulsos, utilizando una técnica de regresión multivariante, PLS, que ya se ha mostrado útil en la resolución de problemas electroquímicos en los que existían interferencias [36, 148]. La determinación de estos cuatro metales se realiza en los análisis rutinarios de aguas y alimentos [149, 150], siendo una constante en la numerosa bibliografía existente la referencia a la formación de compuestos intermetálicos y a los problemas a que da lugar el fenómeno en el análisis [138, 151].

III.4.1 Experimental. Determinación de Cu, Pb, Cd y Zn por DPASV

Reactivos

Las disoluciones patrón utilizadas en el análisis fueron preparadas a partir de reactivos de calidad para análisis y agua bidestilada y desionizada mediante un sistema Barnstead NANO Pure II. Estas disoluciones fueron preparadas en medio ácido para su preservación. Al tratarse de una determinación en la que la concentración de algunos de los metales se encuentra a nivel de trazas, los reactivos empleados como electrolito soporte fueron de calidad Suprapur, utilizándose disoluciones formadas por ácido acético 2 M e hidróxido amónico 1 M, que presentaban un valor de pH de 4.816. Todas las disoluciones analizadas fueron desoxigenadas antes de la etapa de medida burbujeando nitrógeno (99.997%).

Instrumental

Las medidas voltamperométricas fueron realizadas utilizando el procesador Metrohm 646 VA processor junto con la unidad 647 VA equipada con un electrodo multimodo Metrohm (MME) que actuó en el modo electrodo de gota de mercurio colgante (HMDE), un electrodo de referencia Ag/AgCl/KCl (3M) y un electrodo auxiliar de platino. Se ha utilizado la voltamperometría diferencial de impulsos con redisolución anódica (DPASV) como técnica de medida.

Los análisis de datos fueron realizados utilizando los programas PARVUS [57], STATGRAPHICS [58] y MATLAB [59].

Procedimiento experimental

Cada disolución inicial se introdujo en una celda voltamperométrica y fue desoxigenada durante 10 min mediante el paso de nitrógeno a través de la misma, aplicándose posteriormente un potencial de acumulación de -1.11 V sobre el electrodo de mercurio durante un tiempo de 30 s. Al final de este período de acumulación se desconectó la agitación y, tras un reposo de 20 s, se aplicó un barrido de potencial anódico desde -1.11 V hasta 0.072 V con una velocidad de barrido de 10 mV s⁻¹. Después de cada adición, la disolución se agitó y desoxigenó de nuevo durante 15 s, aplicándose a continuación el mismo procedimiento de medida. Otros parámetros instrumentales fueron los siguientes: amplitud del impulso, -50 mV; duración del impulso, 0.6 s; área nominal 0.40 mm²; velocidad de agitación, 1290 rev min⁻¹.

III.4.2 Diseño del experimento

El diseño experimental utilizado para llevar a cabo la determinación de cobre, plomo, cadmio y zinc por DPASV es similar al empleado en el capítulo anterior, Figura III.3.1, en el que se realizó la determinación simultánea de estos cuatro metales mediante polarografía diferencial de impulsos. Las diferencias existentes entre ambos diseños radican en la magnitud de las concentraciones de los diferentes metales que constituyen los mismos, así como en las proporciones existentes entre ellos. En el presente capítulo, la determinación de cobre, plomo, cadmio y zinc, se ha llevado a cabo en disoluciones en las que estos metales se encontraban

a nivel de trazas, en un rango de concentraciones que va desde 2 ppm de cadmio a 51 ppm de zinc, cantidades que se encuentran habitualmente en aguas superficiales. Las concentraciones de las disoluciones empleadas en el análisis se corresponden con el diseño experimental que se muestra en la Tabla III.4.1.

Según este diseño se han realizado dos calibrados completos, llevados a cabo en días diferentes, que se han denominado respectivamente calibrados C (el realizado en primer lugar) y D (el calibrado posterior). Los voltamperogramas correspondientes a ambos calibrados se muestran en las Figuras III.4.1a-j y III.4.2a-j, encontrándose agrupados en cada una de las gráficas que componen dichas figuras aquellos voltamperogramas obtenidos partiendo de una misma disolución inicial. En ellos, se observan las señales voltamperométricas correspondientes a cada uno de los analitos de interés, resultando evidente que se trata de picos voltamperométricos bien definidos y perfectamente separados unos de otros. De los mismos se desprende también la mayor variabilidad que acompaña a las medidas realizadas por voltamperometría diferencial de impulsos con redisolución anódica, frente a la observada cuando las determinaciones se han llevado a cabo mediante medidas polarográficas directas para estos mismos metales (ver Figuras III.3.2a-j y III.3.3a-j). En ello influye enormemente el que las concentraciones de los metales que están siendo determinados se encuentren en un rango entre 10^{-8} M y 10^{-7} M, razón por la cual se utiliza una técnica electroquímica con etapa de redisolución, mientras que las concentraciones determinadas en el análisis polarográfico anterior se encontraban entre 10^{-6} M y 10^{-5} M.

III.4.3 Análisis univariante

Se ha realizado un estudio univariante análogo al llevado a cabo en el capítulo anterior, es decir, se han efectuado calibrados univariantes considerando como variable dependiente la intensidad del máximo de pico correspondiente a cada analito en aquellas muestras en las que únicamente se incrementaba la concentración de éste. Por tanto, se ha llevado a cabo un calibrado univariante para el cobre considerando la intensidad de pico registrada a -24 mV en las muestras 17 a 21, para el plomo se han tenido en cuenta las señales que aparecen al potencial -444 mV en las muestras 22 a 26, para el cadmio las intensidades a -606 mV en las muestras 27 a 31, y por último, para el calibrado univariante correspondiente al zinc se han

Tabla III.4.1 Diseño experimental seguido en los calibrados C y D. Cada símbolo de la segunda columna indica que se realizaron adiciones sobre una misma disolución inicial. Se han marcado con una *t* aquellas muestras que constituyen el conjunto test.

Muestra	Diseño experimental				Concentraciones (nM)			
	Cobre	Plomo	Cadmio	Zinc	Cobre	Plomo	Cadmio	Zinc
2 ×	2	2	2	2	138.85	51.15	35.39	322.53
2 ×	2	2	2	4	137.11	50.51	34.95	637.00
3 ×	2	2	4	4	135.42	49.89	69.04	629.14
4 ×	2	4	4	4	133.77	98.56	68.20	621.46
5 ×	4	4	4	4	264.31	97.38	67.37	613.98
6 ▲	2	2	4	2	137.11	50.51	69.90	318.50
7 ▲	2	4	4	2	135.42	99.78	69.04	314.57
8 ▲	4	4	4	2	267.54	98.56	68.20	310.73
9 ·	2	4	2	2	137.11	101.03	34.95	318.50
10 ·	2	4	2	4	135.42	99.78	34.52	629.14
11 ·	4	4	2	4	267.54	98.56	34.10	621.46
12 ○	4	2	2	2	274.23	50.51	34.95	318.50
13 ○	4	2	2	4	270.84	49.89	34.52	629.14
14 ○	4	2	4	4	267.54	49.28	68.20	621.46
15 ■	4	2	4	2	270.84	49.89	69.04	314.57
16 -	4	4	2	2	270.84	99.78	34.52	314.57
17 ▼	1	3	3	3	68.56	75.77	52.43	477.75
18 ▼ <i>t</i>	2	3	3	3	136.26	75.30	52.10	474.78
19 ▼	3	3	3	3	203.13	74.84	51.78	471.85
20 ▼ <i>t</i>	4	3	3	3	269.18	74.38	51.46	468.96
21 ▼	5	3	3	3	334.42	73.92	51.15	466.10
22 □	3	1	3	3	205.67	25.26	52.43	477.75
23 □ <i>t</i>	3	2	3	3	204.39	50.20	52.10	474.78
24 □	3	3	3	3	203.13	74.84	51.78	471.85
25 □ <i>t</i>	3	4	3	3	201.88	99.17	51.46	468.96
26 □	3	5	3	3	200.65	123.20	51.15	466.10
27 ·	3	3	1	3	205.67	75.77	17.48	477.75
28 · <i>t</i>	3	3	2	3	204.39	75.30	34.73	474.78
29 ·	3	3	3	3	203.13	74.84	51.78	471.85
30 · <i>t</i>	3	3	4	3	201.88	74.38	68.61	468.96
31 ·	3	3	5	3	200.65	73.92	85.24	466.10
32 ▽	3	3	3	1	205.67	75.77	52.43	159.25
33 ▽ <i>t</i>	3	3	3	2	204.39	75.30	52.10	316.52
34 ▽	3	3	3	3	203.13	74.84	51.78	471.85
35 ▽ <i>t</i>	3	3	3	4	201.88	74.38	51.46	625.28
36 ▽	3	3	3	5	200.65	73.92	51.15	776.83

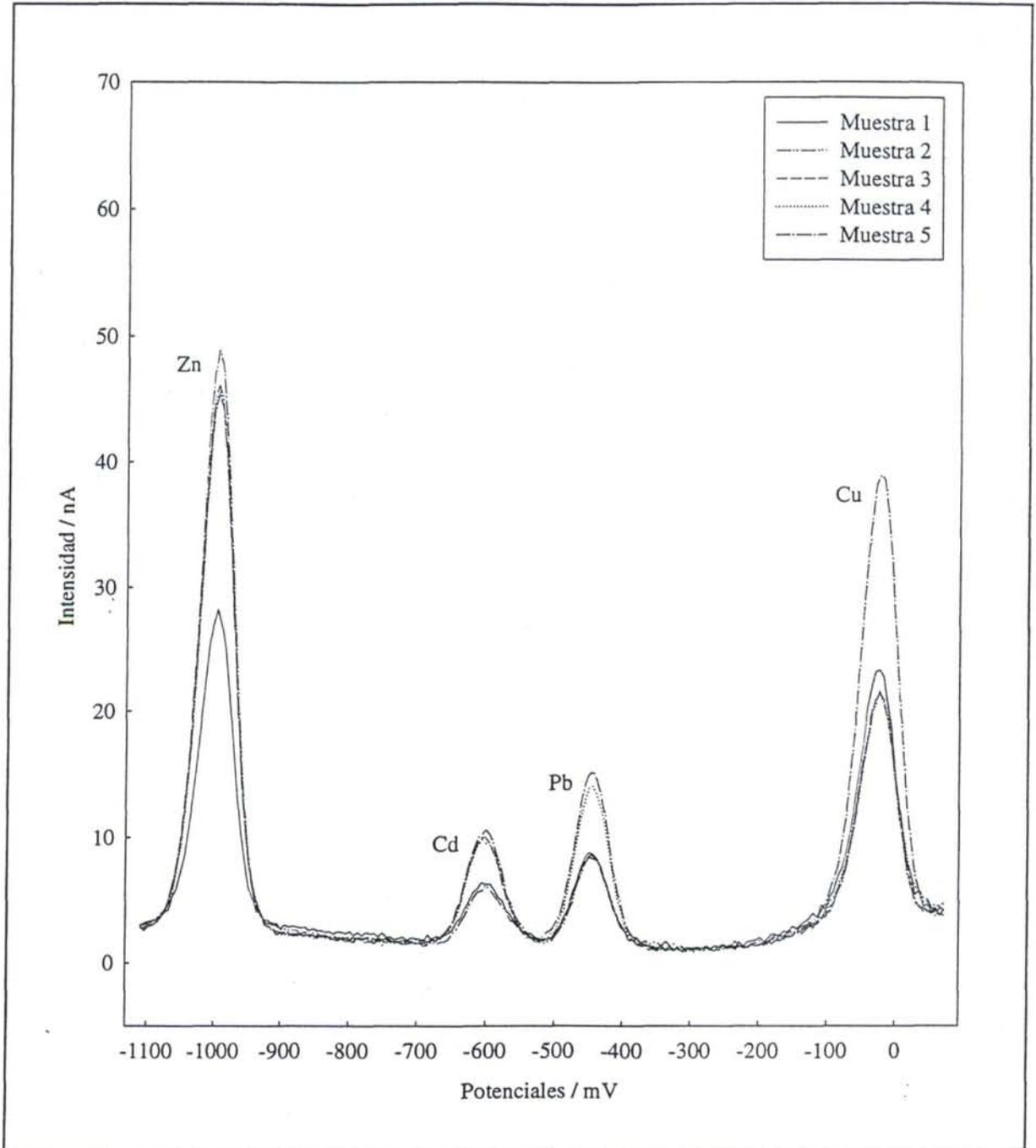


Figura III.4.1a Voltamperogramas correspondientes al calibrado C, obtenidos con una misma disolución inicial.

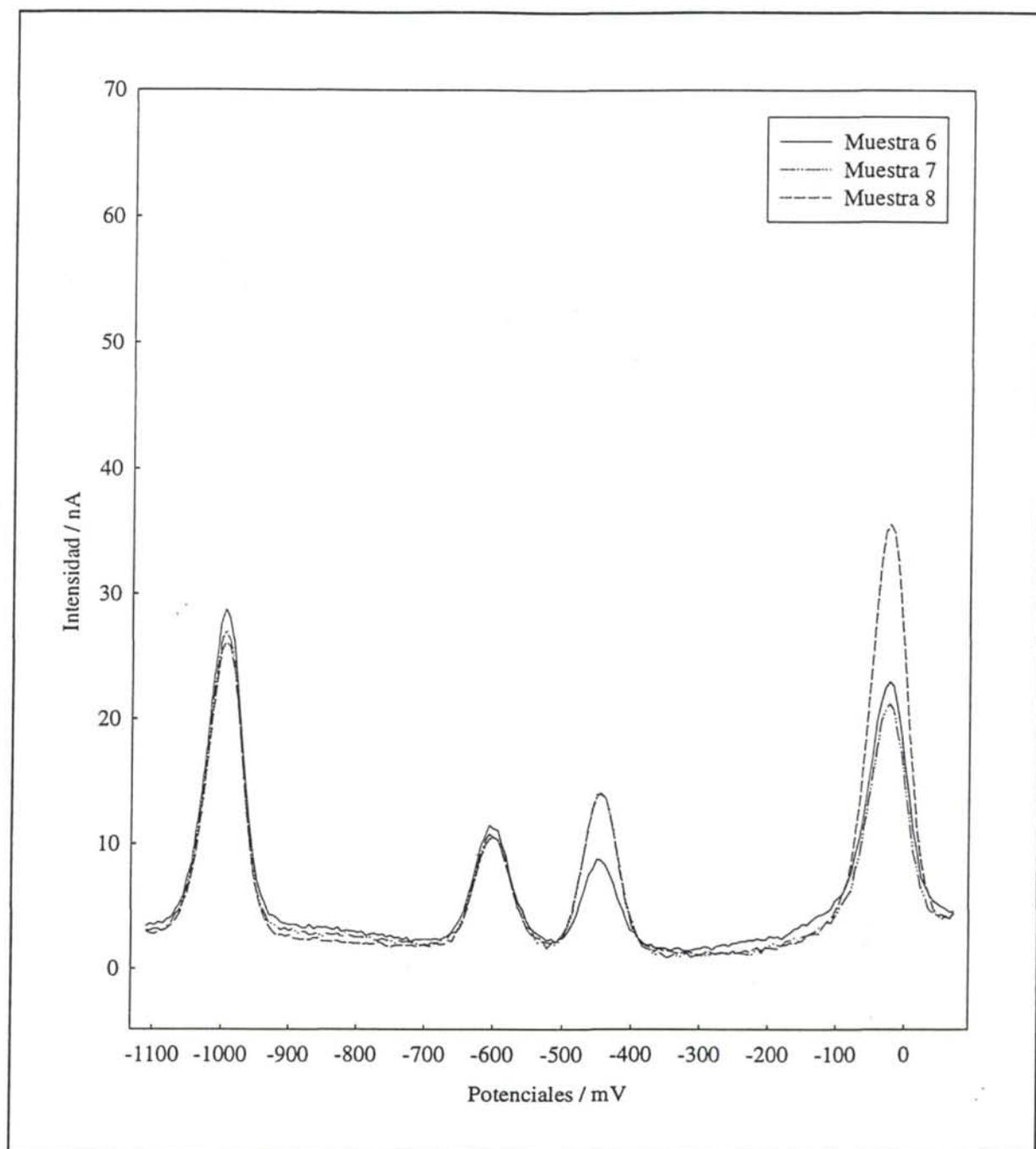


Figura III.4.1b Voltamperogramas correspondientes al calibrado C, obtenidos con una misma disolución inicial.

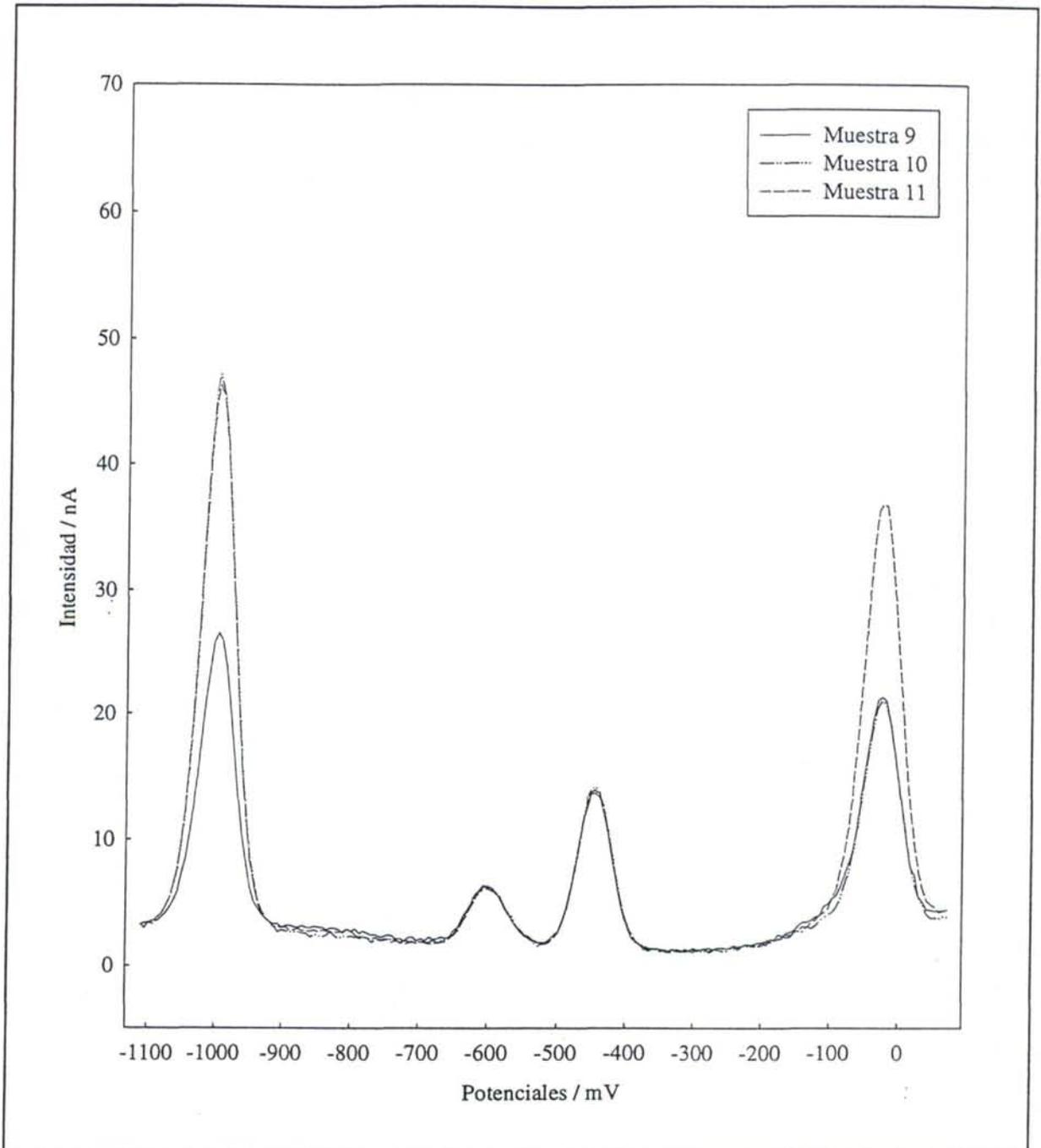


Figura III.4.1c Voltamperogramas correspondientes al calibrado C, obtenidos con una misma disolución inicial.

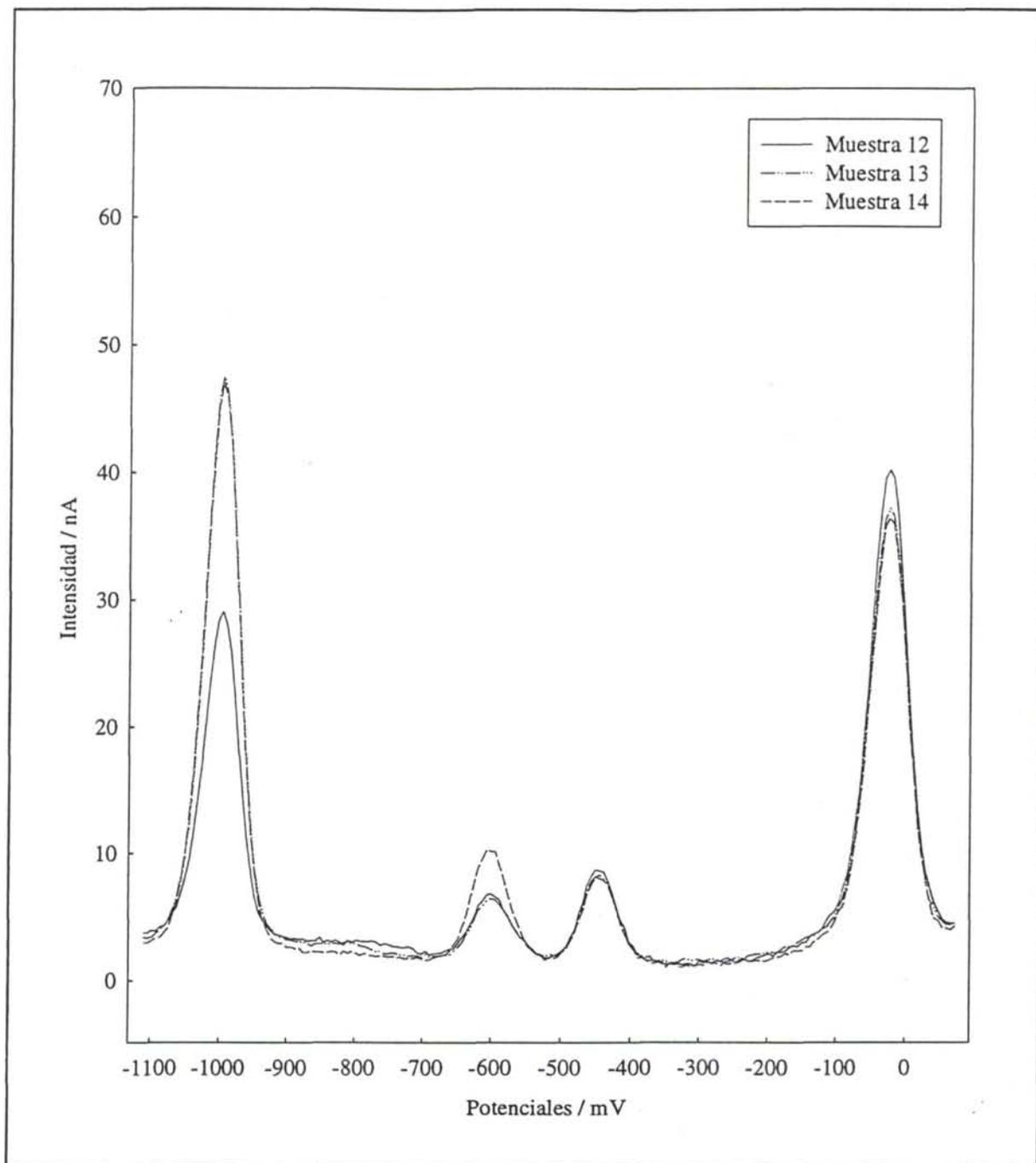


Figura III.4.1d Voltamperogramas correspondientes al calibrado C, obtenidos con una misma disolución inicial.

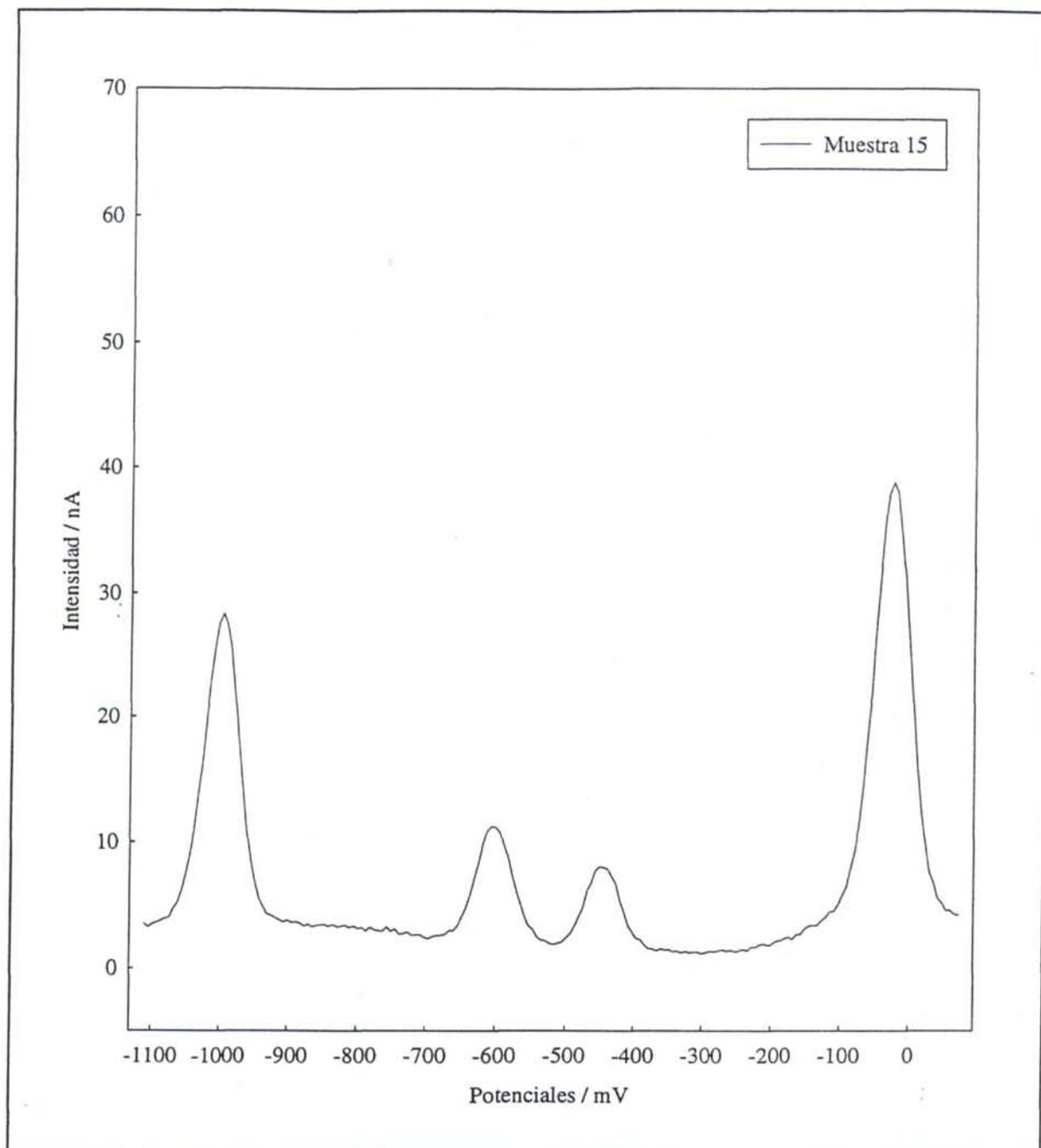


Figura III.4.1e Voltamperogramas correspondientes al calibrado C, obtenidos con una misma disolución inicial.

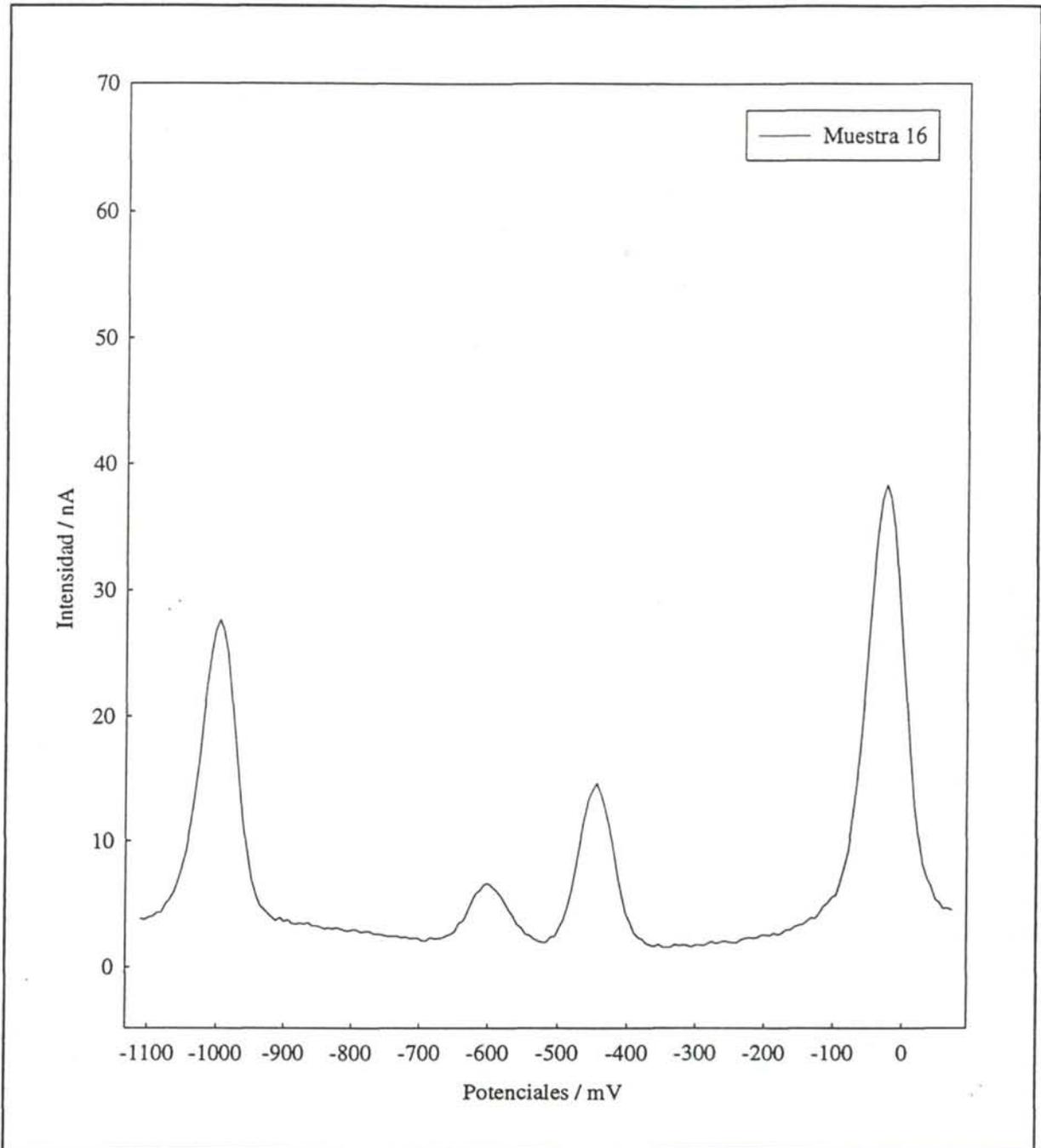


Figura III.4.1f Voltamperogramas correspondientes al calibrado C, obtenidos con una misma disolución inicial.

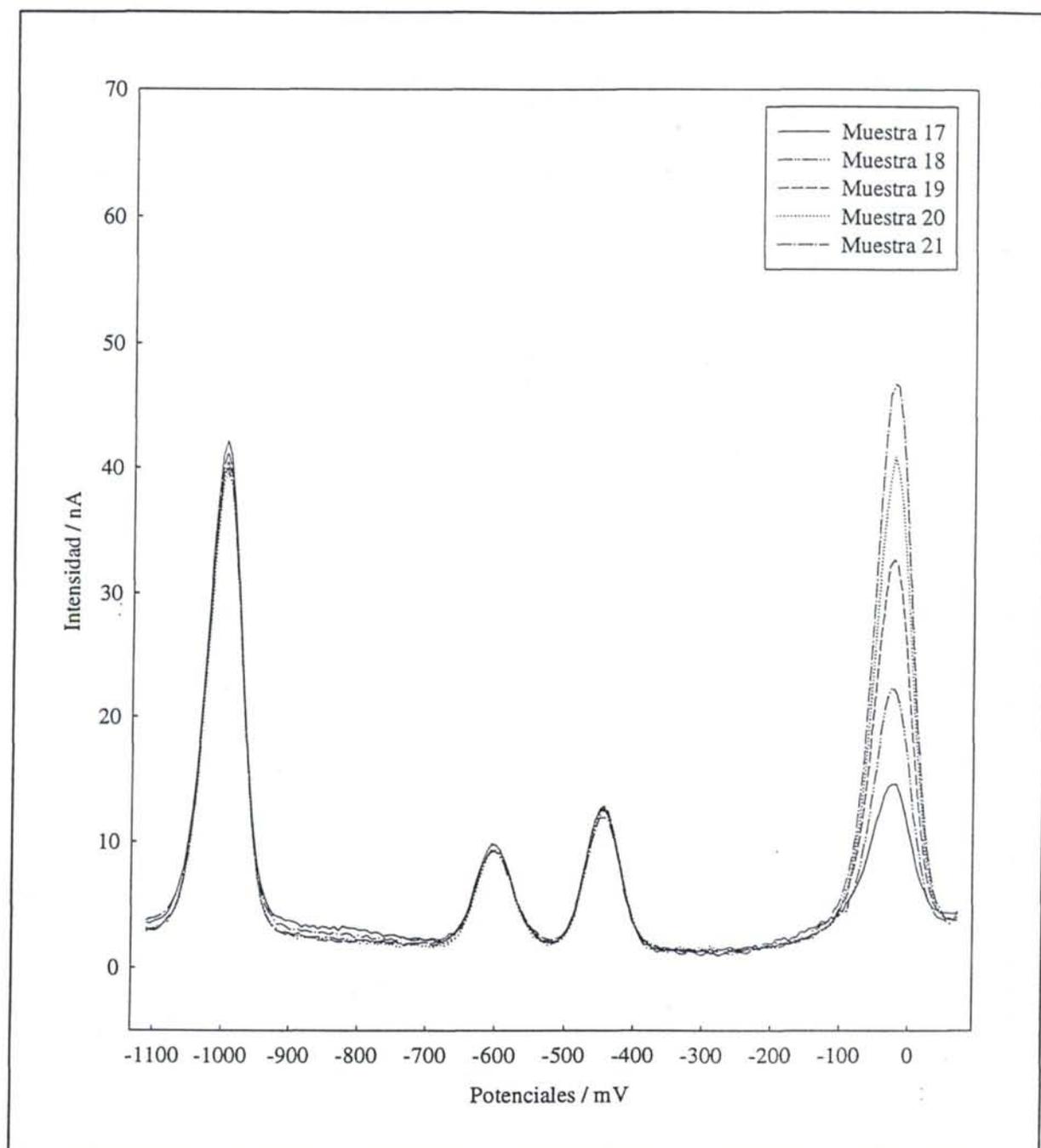


Figura III.4.1g Voltamperogramas correspondientes al calibrado C, obtenidos con una misma disolución inicial.

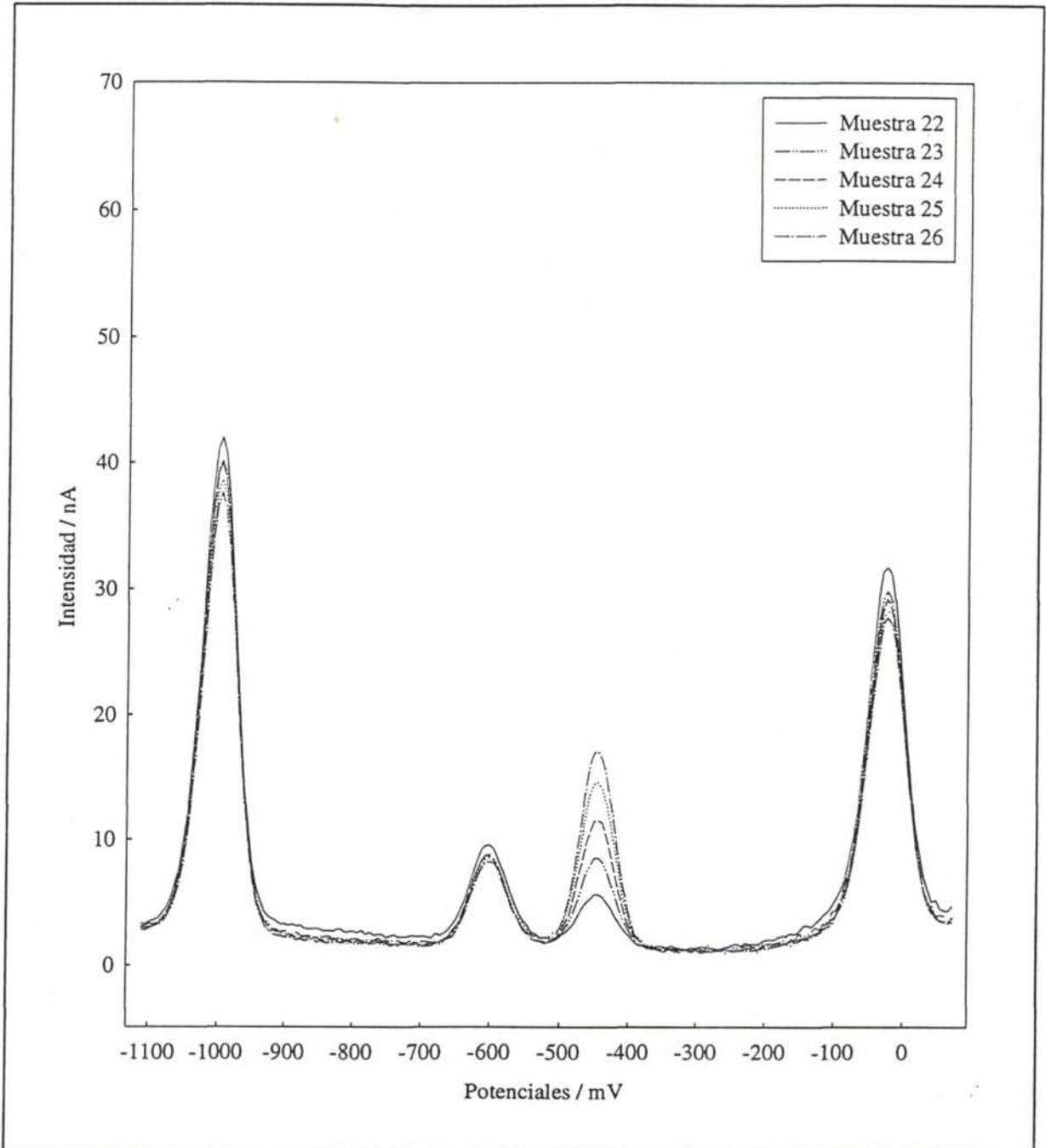


Figura III.4.1h Voltamperogramas correspondientes al calibrado C, obtenidos con una misma disolución inicial.

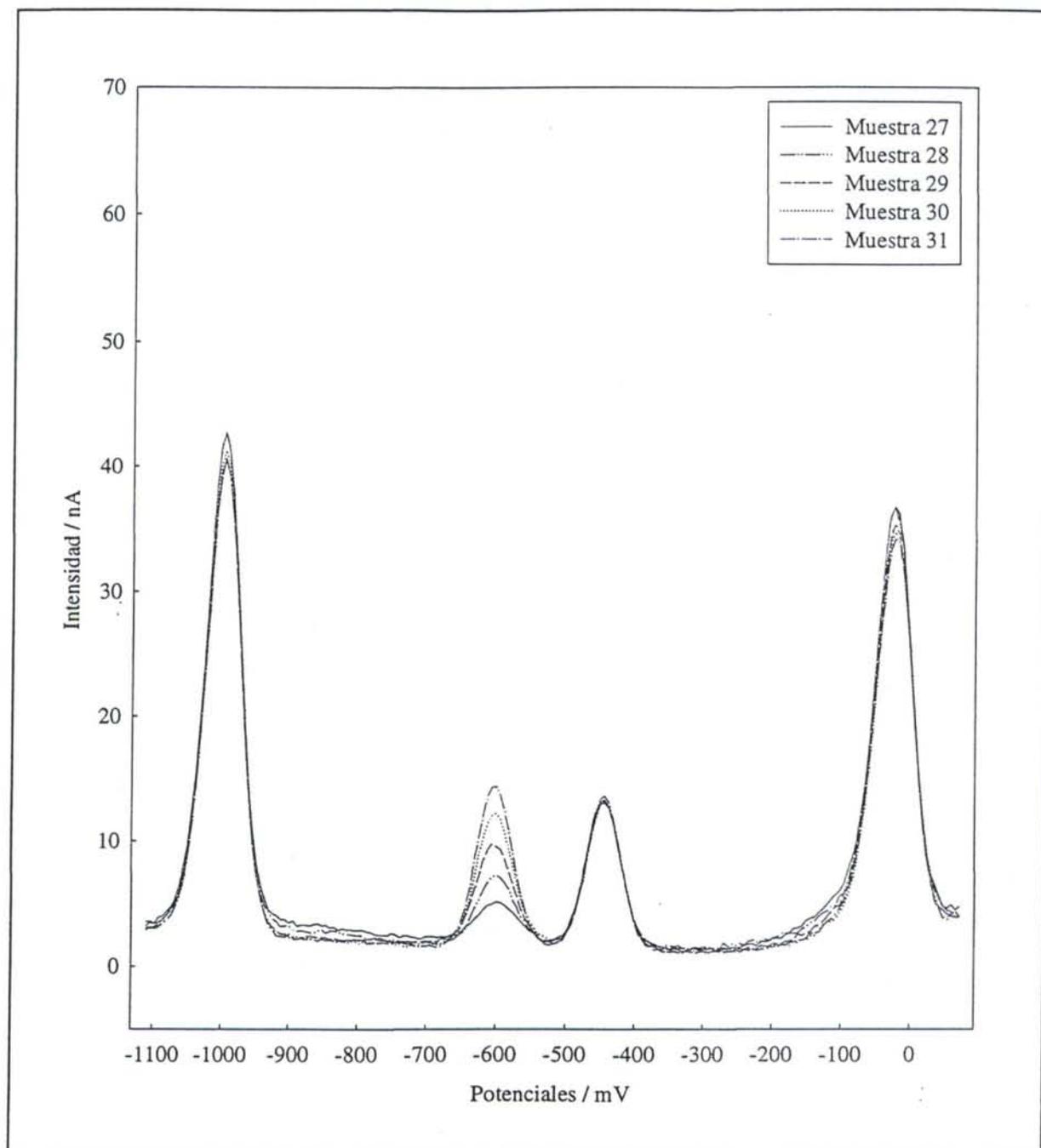


Figura III.4.1i Voltamperogramas correspondientes al calibrado C, obtenidos con una misma disolución inicial.

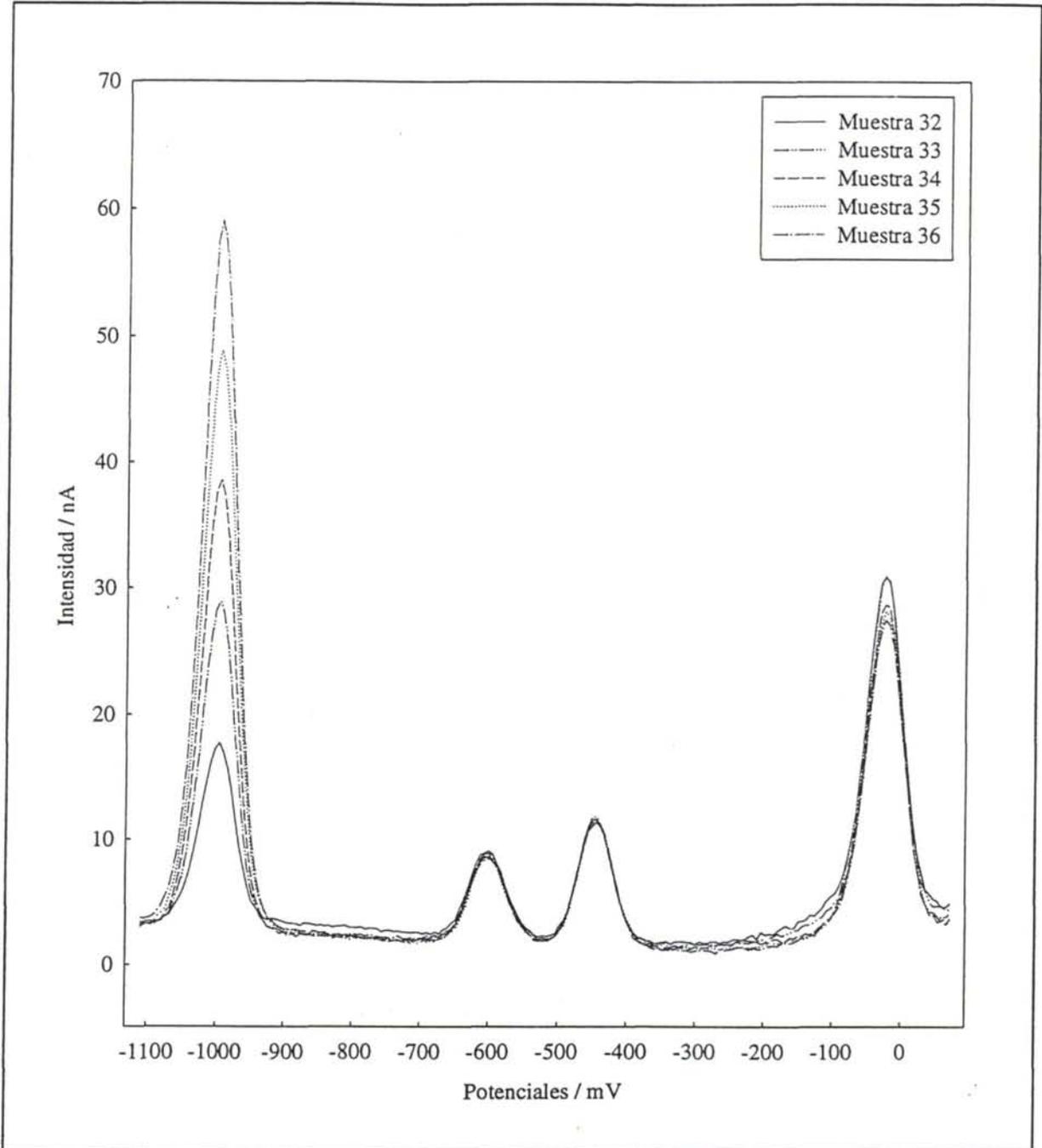


Figura III.4.1j Voltamperogramas correspondientes al calibrado C, obtenidos con una misma disolución inicial.

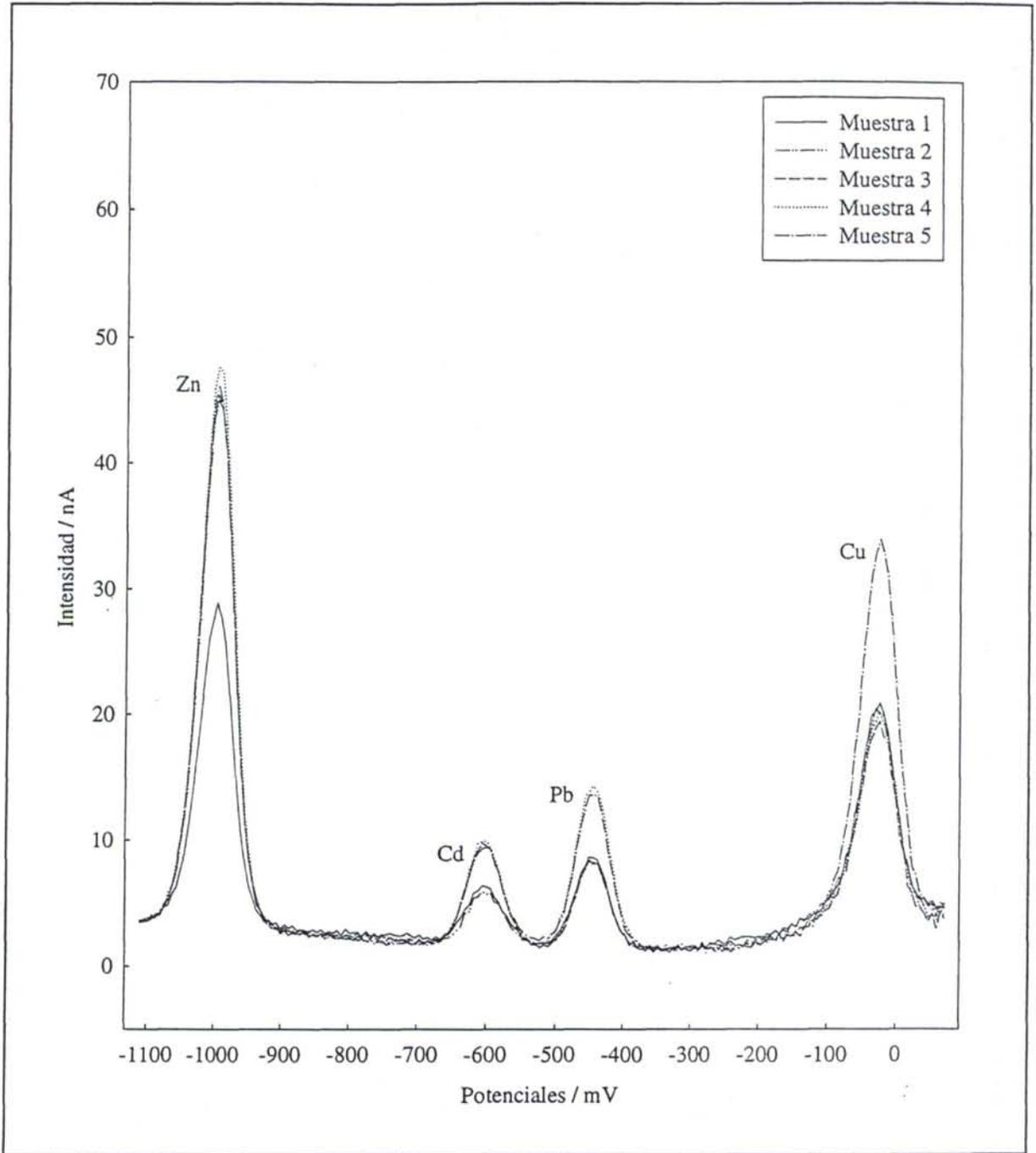


Figura III.4.2a Voltamperogramas correspondientes al calibrado D, obtenidos con una misma disolución inicial.

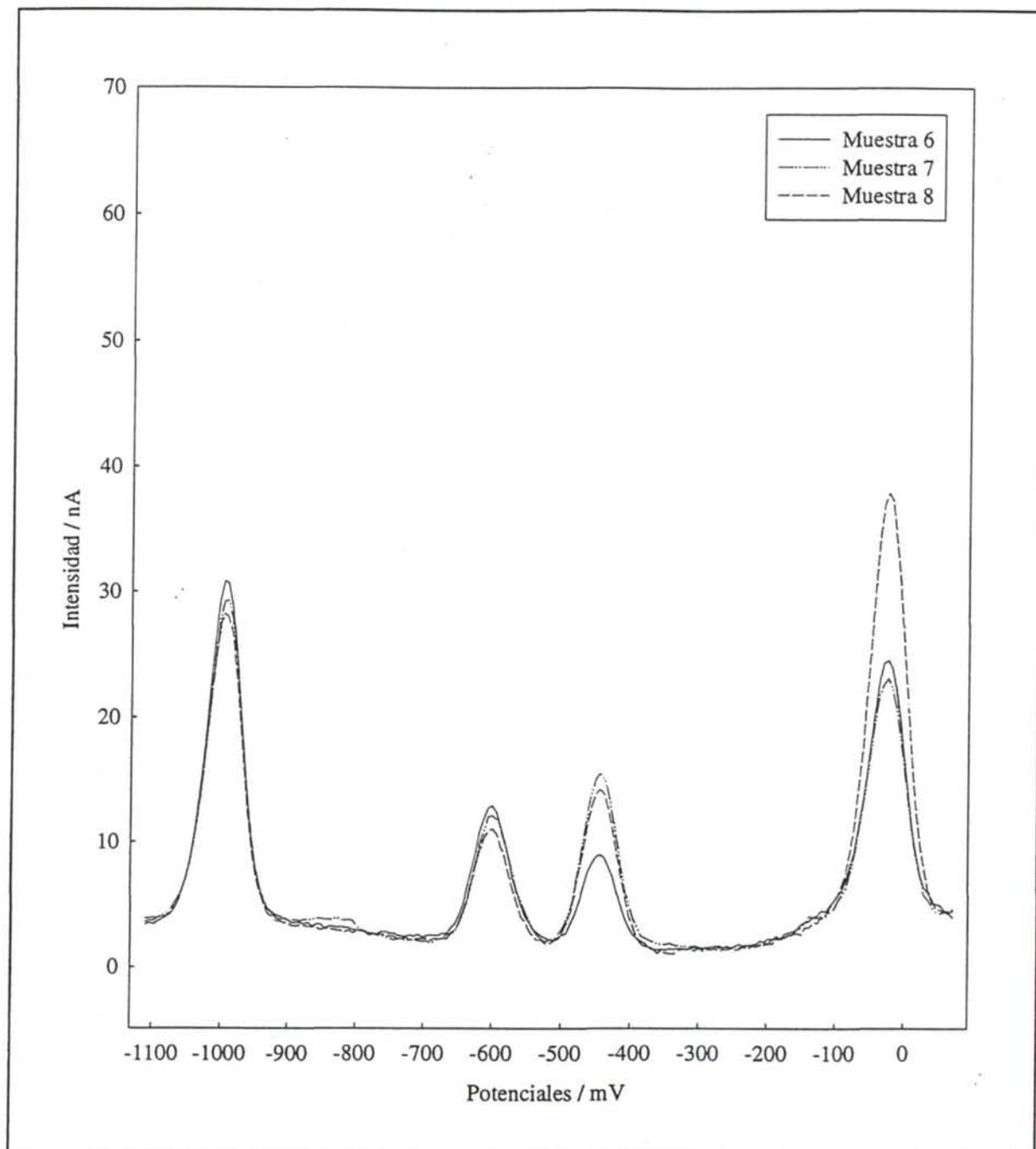


Figura III.4.2b Voltamperograms correspondientes al calibrado D, obtenidos con una misma disolución inicial.

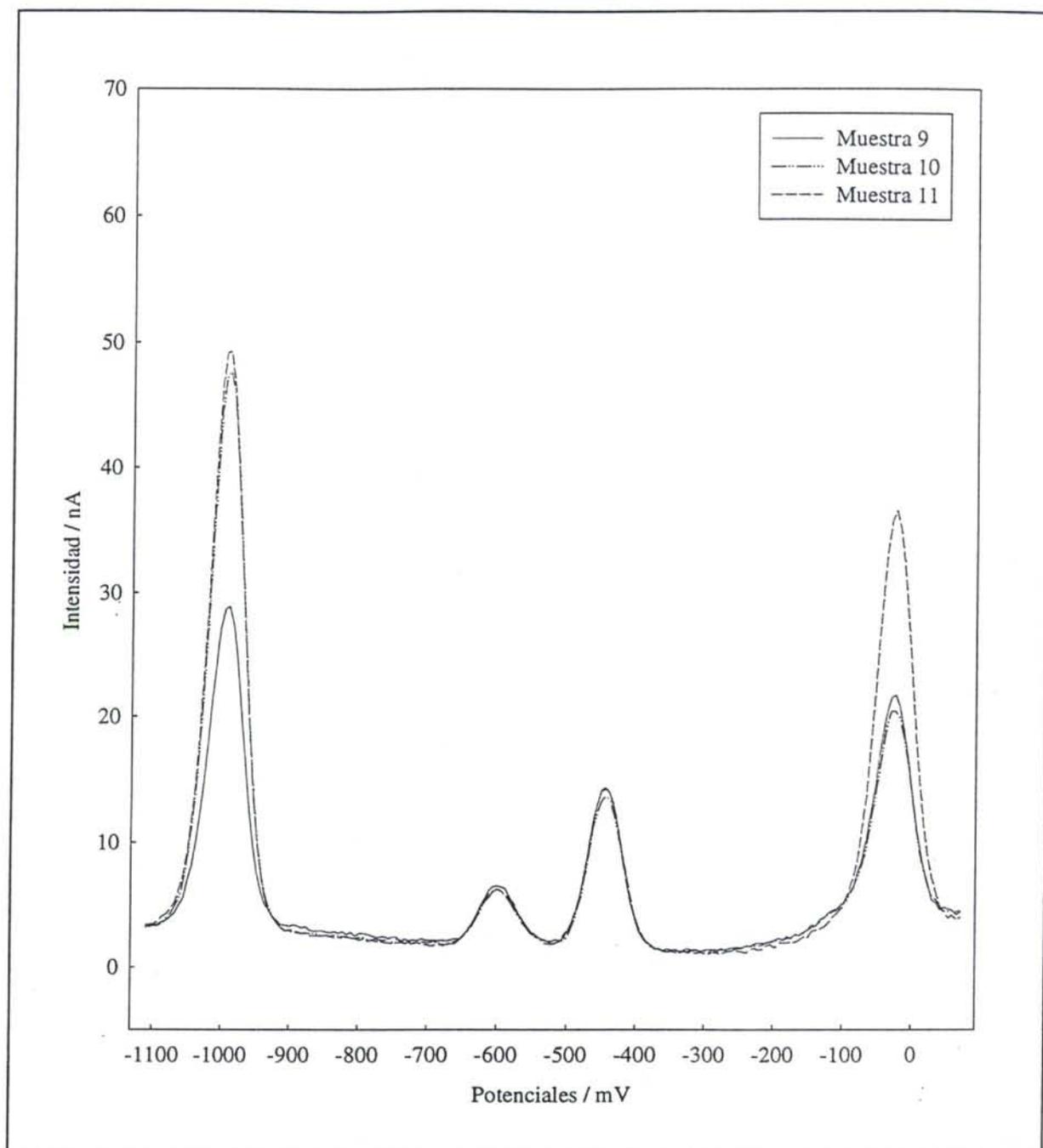


Figura III.4.2c Voltamperogramas correspondientes al calibrado D, obtenidos con una misma disolución inicial.

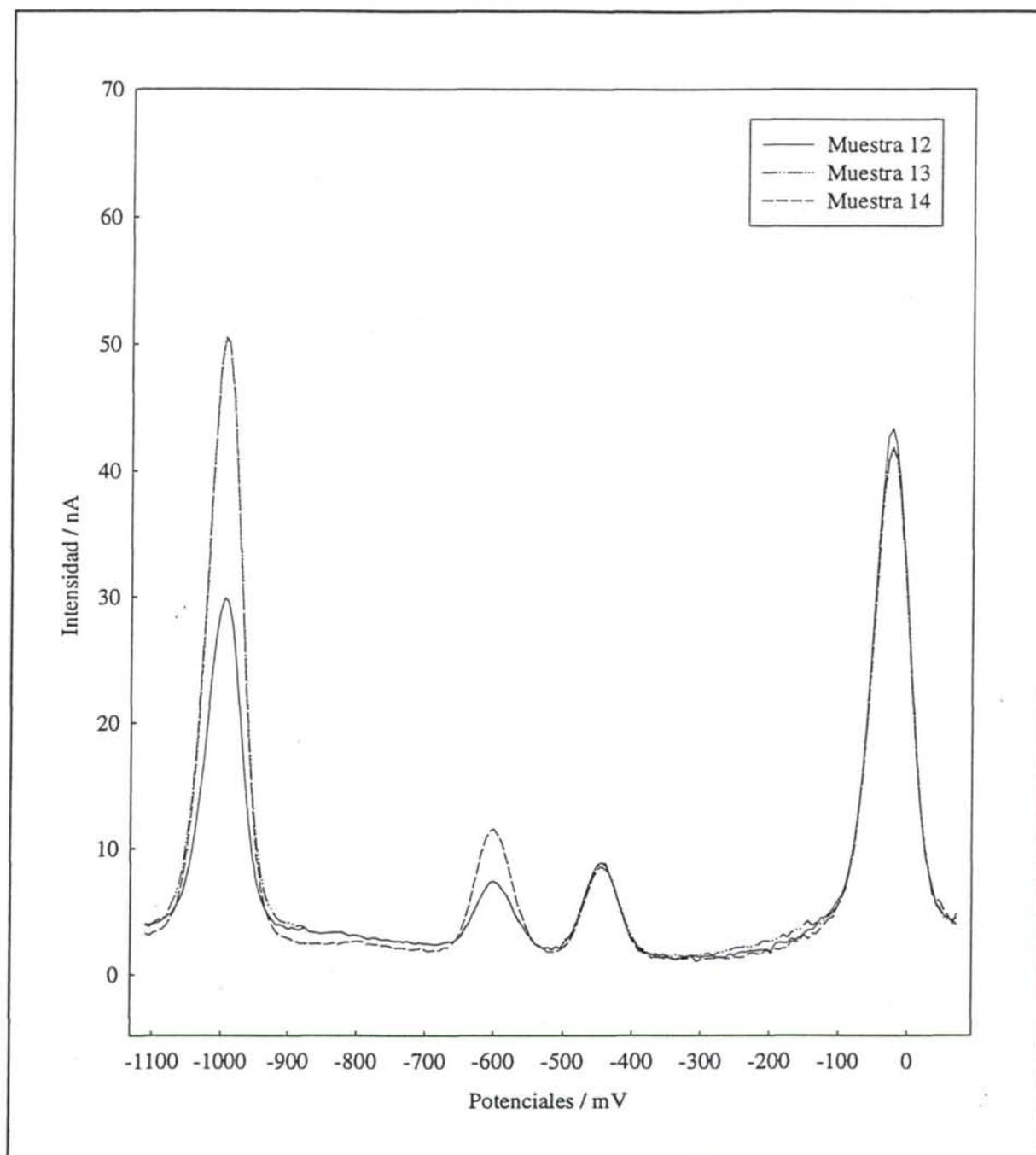


Figura III.4.2d Voltamperogramas correspondientes al calibrado D, obtenidos con una misma disolución inicial.

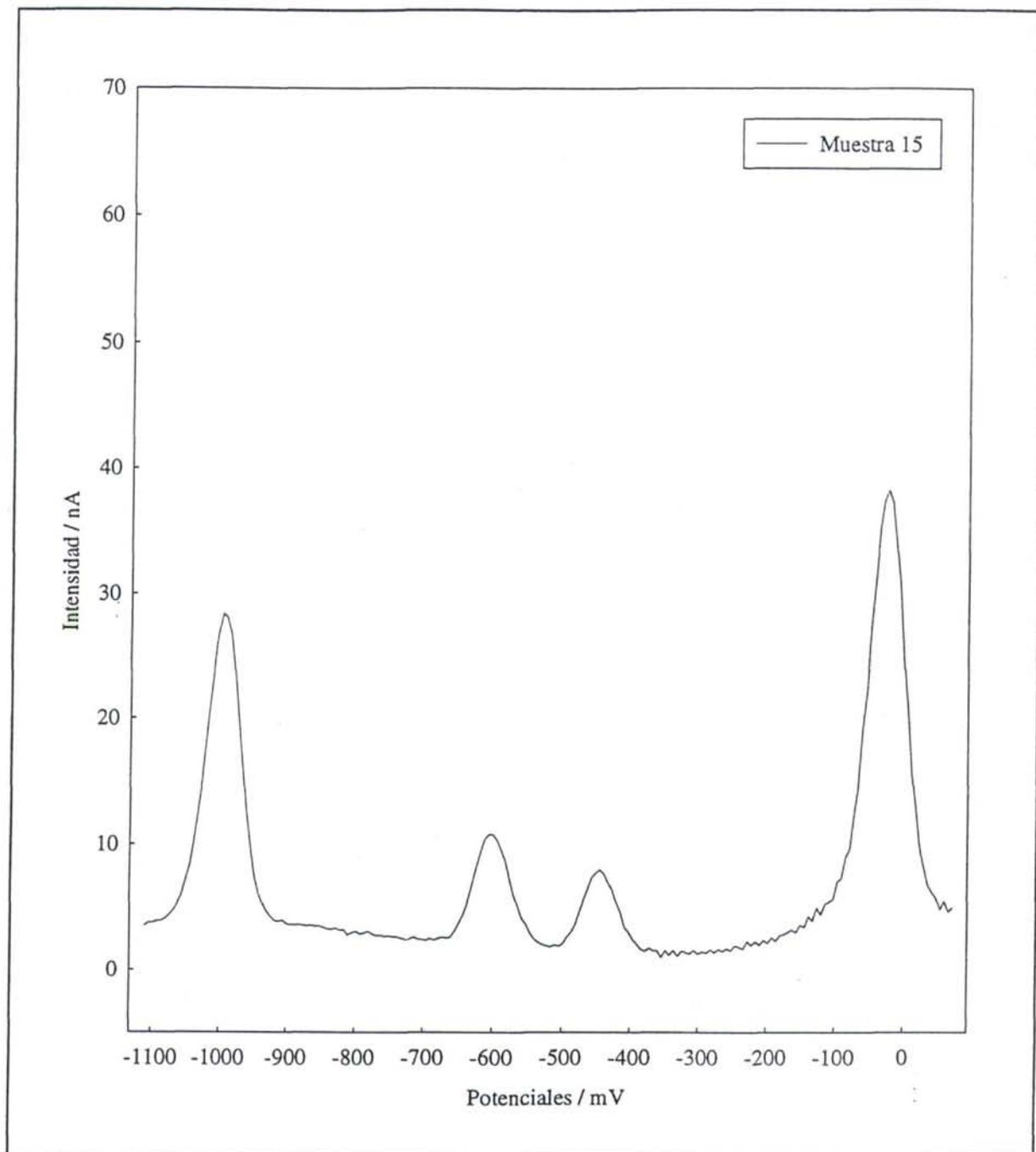


Figura III.4.2e Voltamperogramas correspondientes al calibrado D, obtenidos con una misma disolución inicial.

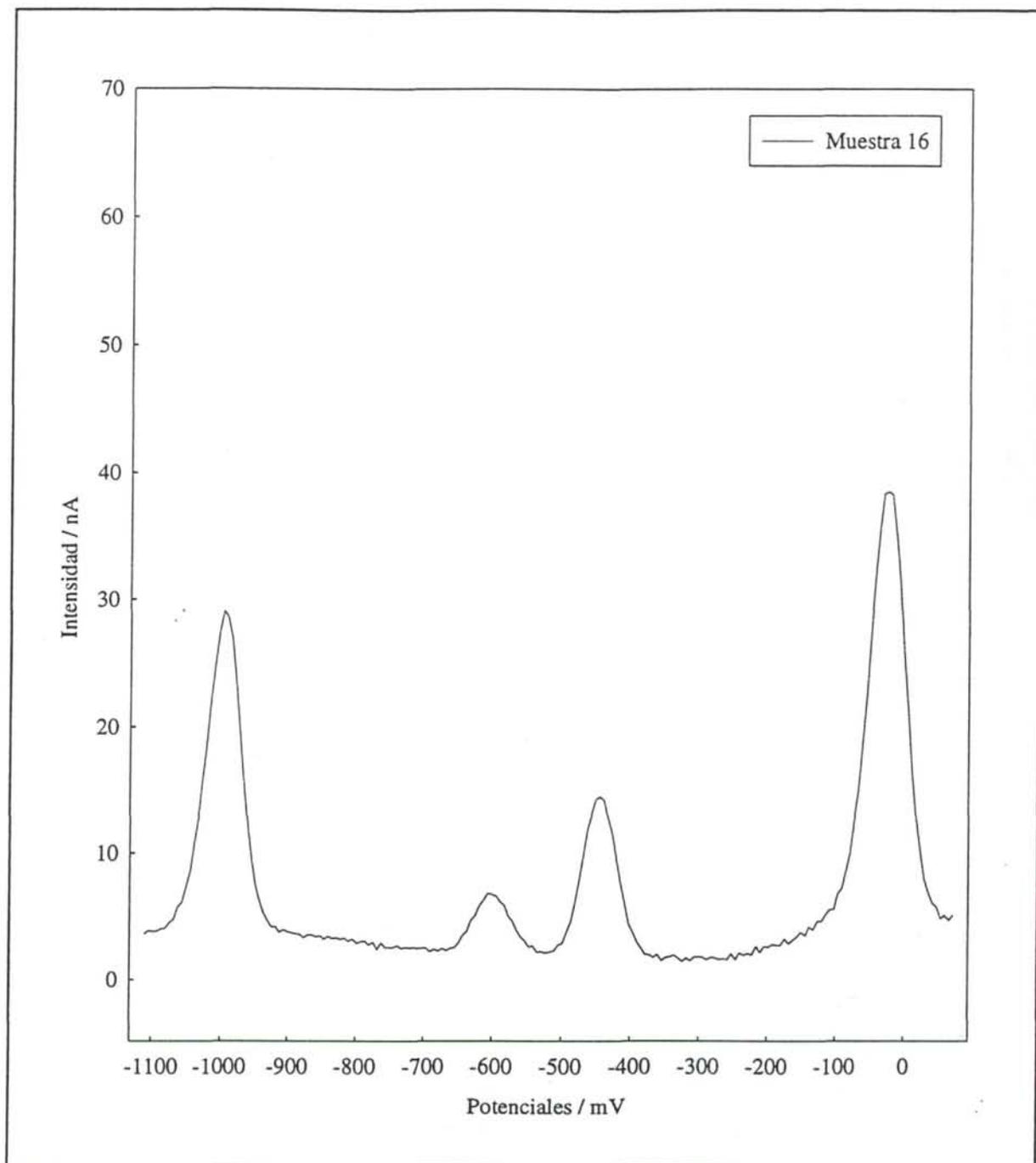


Figura III.4.2f Voltamperogramas correspondientes al calibrado D, obtenidos con una misma disolución inicial.

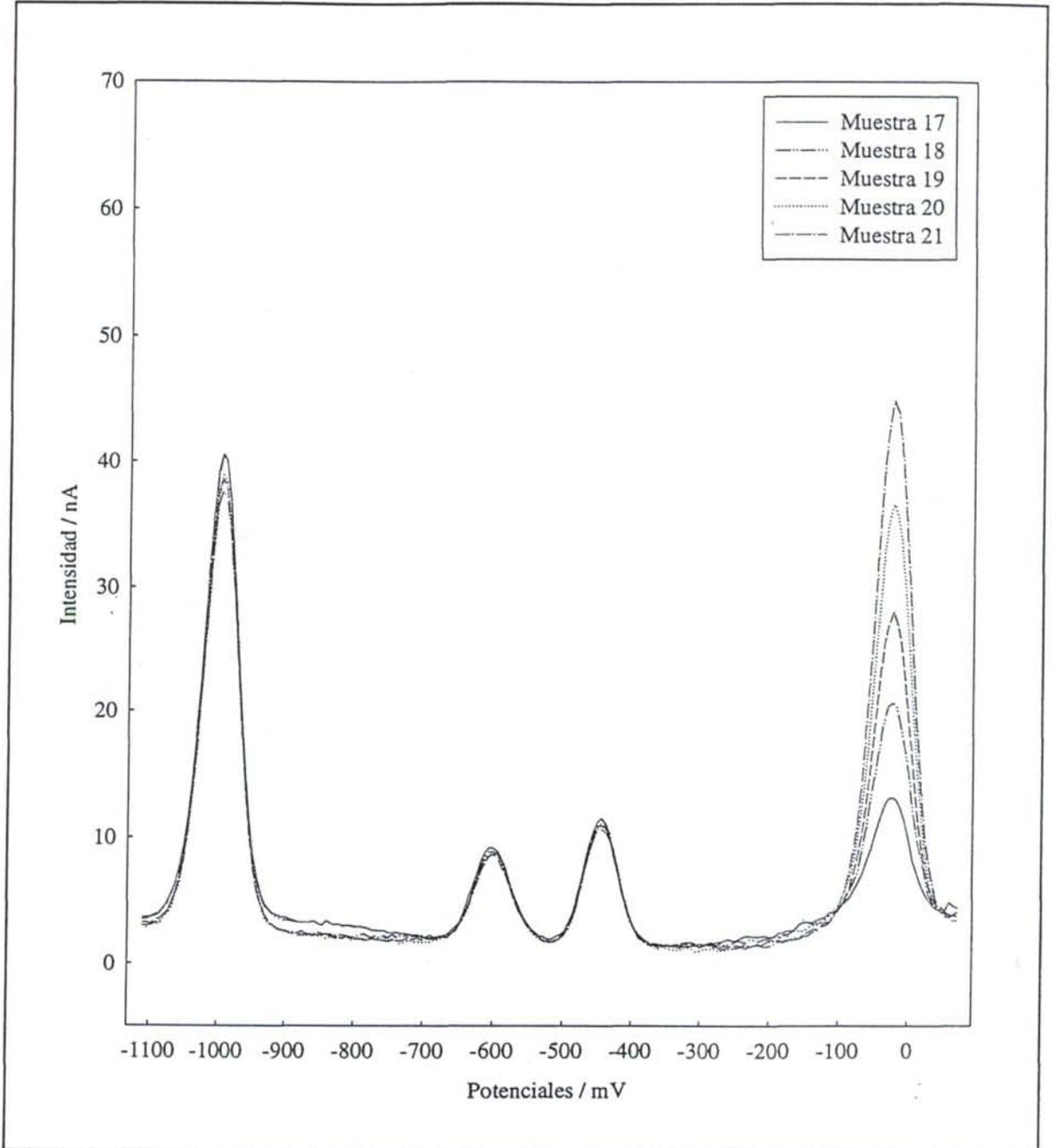


Figura III.4.2g Voltamperogramas correspondientes al calibrado D, obtenidos con una misma disolución inicial.

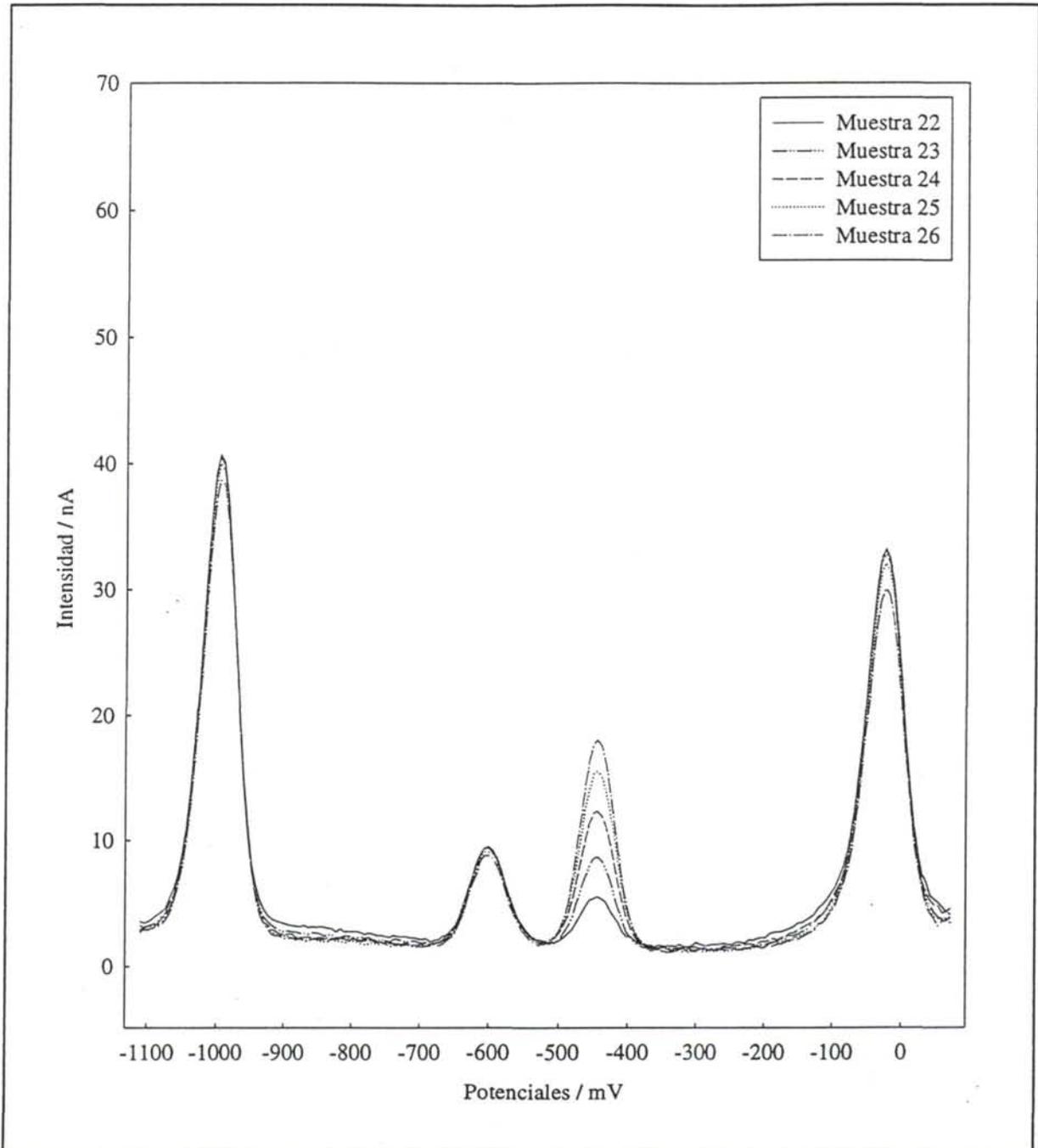


Figura III.4.2h Voltamperogramas correspondientes al calibrado D, obtenidos con una misma disolución inicial.

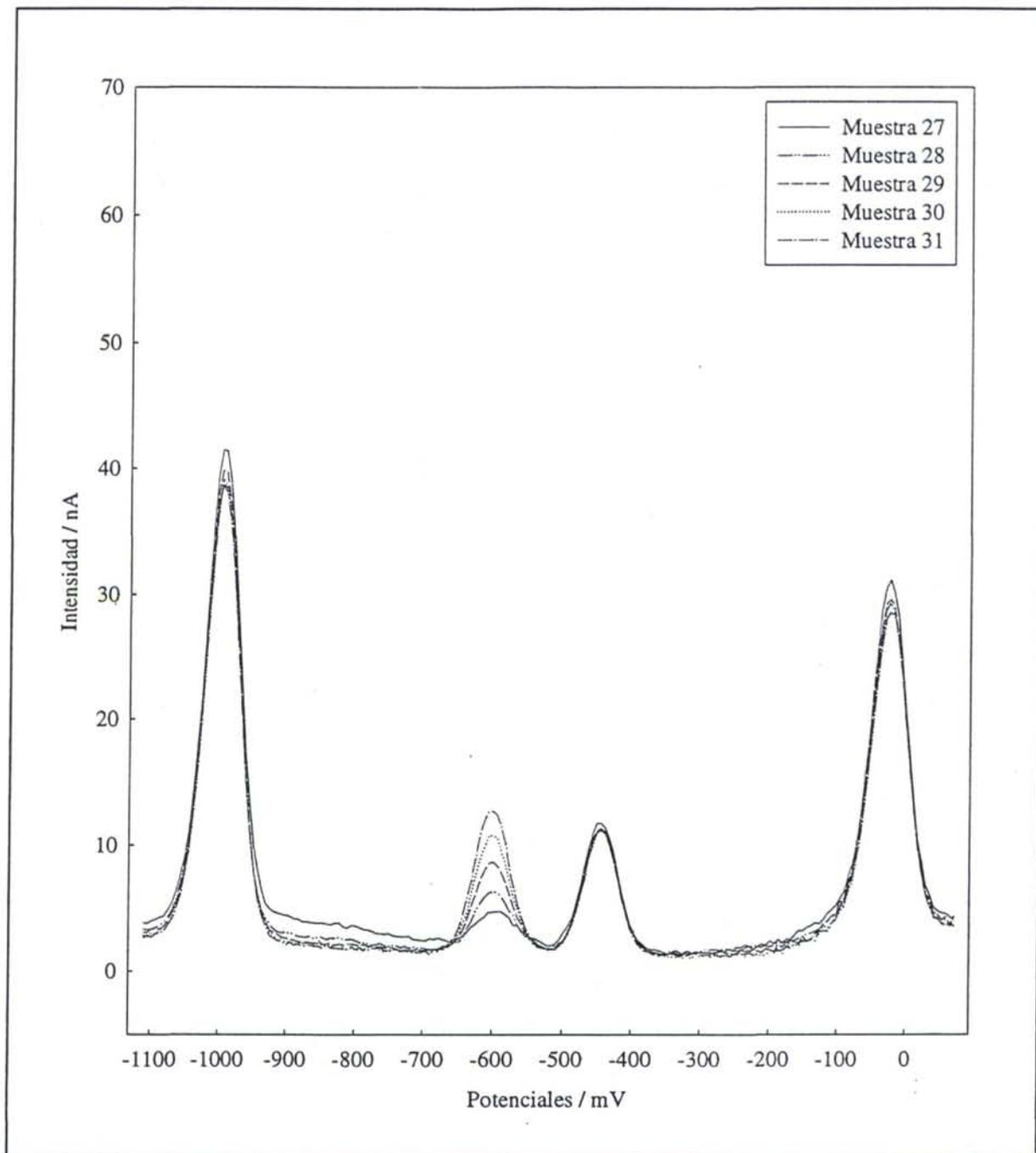


Figura III.4.2i Voltamperogramas correspondientes al calibrado D, obtenidos con una misma disolución inicial.

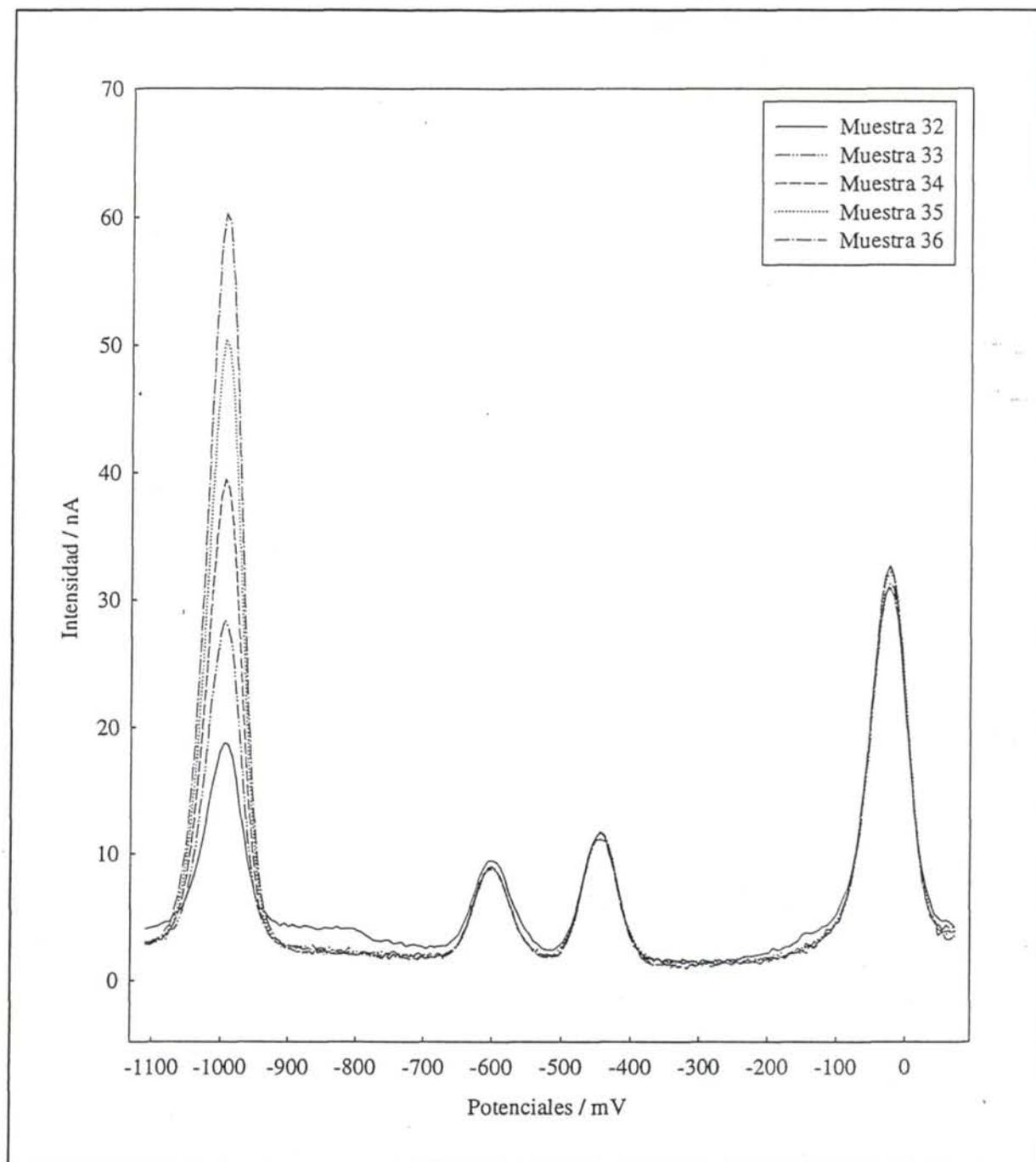


Figura III.4.2j Voltamperogramas correspondientes al calibrado D, obtenidos con una misma disolución inicial.

tomado las intensidades de pico registradas a -996 mV en las muestras 32 a 36. Para realizar el análisis se han utilizado los datos del calibrado C, a los que previamente se ha restado el blanco experimental correspondiente.

Los resultados correspondientes a este análisis se muestran en la Tabla III.4.2, en la que se puede observar que el cobre presenta el coeficiente de correlación más bajo y la S_{yx} mayor, que conduce a intervalos de confianza muy grandes para pendiente y término independiente.

Tabla III.4.2 Resultados del análisis de regresión por mínimos cuadrados realizado con datos del calibrado C.

Regresión	Cobre	Plomo	Cadmio	Zinc
Pendiente	0.12 ± 0.02	0.12 ± 0.01	0.14 ± 0.01	0.07 ± 0.01
Término independiente	2.31 ± 4.25	2.67 ± 0.46	2.64 ± 0.40	1.09 ± 1.14
R^2_{ajustado}	0.9908	0.9991	0.9990	0.9996
S_{yx}	1.26	0.1363	0.12	0.3361

A continuación, se ha utilizado el procedimiento de adición estándar para determinar la concentración de los analitos en muestras concretas. De modo similar a lo llevado a cabo en el capítulo anterior, utilizando los datos del calibrado D y de forma independientemente para cada analito, se ha determinado la concentración de cobre en la muestra 18, la de plomo en la 23, la de cadmio en la 28, y la zinc en la 35. Considerando como adiciones estándar las adiciones realizadas de cada uno de los analitos respectivos cuando no se añadía ningún otro metal, se han construido 4 rectas de regresión univariante cuyos parámetros de ajuste se encuentran en la Tabla III.4.3. En la misma se muestran, además los resultados del método de adición estándar aplicado para cada analito, la concentración calculada a través de dicho procedimiento y el error asociado con ella.

En dicha tabla, se observa que los errores no son tan considerables como los obtenidos en el capítulo anterior, Tabla III.3.3, cuando se utilizó este mismo procedimiento (algunos alrededor de 40%), pero aún así no pueden considerarse despreciables. En este caso, los errores mayores están asociados con los analitos que se hallan en concentraciones menores, plomo y cadmio, mientras que son un poco menores los correspondientes a cobre y zinc. Esto parece indicar que, en el análisis univariante y en las condiciones en las que se ha llevado a

Tabla III.4.3 Resultados del método de adición estándar aplicado a las muestras del calibrado D.

Regresión	Cobre	Plomo	Cadmio	Zinc
Pendiente	0.12	0.13	0.12	0.07
Término independiente	17.66	7.15	4.70	22.48
R^2_{ajustado}	0.9984	0.9929	0.9984	0.9992
S_{yx}	0.42	0.34	0.11	0.39
Número de muestra	18	23	28	33
Concentración calculada (nM)	145.49	53.86	37.90	327.15
Error relativo (%)	6.77	12.46	9.14	3.36

cabo la experiencia, la interferencia debida a la formación de compuestos intermetálicos influye en las determinaciones llevadas a cabo en menor medida que el hecho de que los analitos se encuentren en concentraciones muy bajas.

A pesar de que estos resultados no señalan que el análisis univariante sea inviable, hay que considerar el hecho de que, si se siguiera este procedimiento para determinar la concentración de estos cuatro metales en una misma muestra, el esfuerzo experimental implicado en cada medida sería muy grande. En principio, sería necesario disponer de cuatro alícuotas de la solución problema, sobre cada una de ellas se llevarían a cabo tres adiciones del patrón correspondiente, realizándose tras cada adición una medida voltamperométrica según el procedimiento experimental indicado. Esto equivale a decir que, en conjunto, sería necesario preparar y medir un total de 16 muestras diferentes para llevar a cabo la determinación de los cuatro analitos en una única muestra problema. El procedimiento completo debería repetirse en el caso de que se quisiera determinar la concentración de estos metales en una nueva muestra, lo cual implica la necesidad de realizar un gran esfuerzo experimental con cada una de las muestras problemas de que se disponga.

Con la finalidad de reducir de forma significativa el esfuerzo experimental que se requiere para llevar a cabo la determinación de los cuatro analitos en un conjunto de muestras problema, se propone la utilización de una metodología multivariante que permita construir modelos de calibrado multivariantes capaces de determinar de forma simultánea los cuatro

analitos en cualquier muestra problema, y que para ello tengan en cuenta las posibles relaciones internas que puedan existir entre los distintos analitos.

III.4.4 Análisis multivariante

III.4.4.1 Construcción y evaluación de modelos PLS

Construcción de los modelos PLS

Se han construido modelos PLS para cada uno de los analitos que constituyen los calibrados C y D, por lo que se dispone de un total de 8 modelos de regresión independientes, cuatro para cada calibrado. En el desarrollo de cada modelo PLS se han considerado como variables predictoras los 198 valores de intensidad registrados a intervalos equidistantes de potencial entre -1.11 V y 0.072 V. Las correspondientes matrices de datos, **C** y **D**, están constituidas por 36 filas u objetos (todas las muestras medidas ambos calibrados) y por 198 columnas (las variables predictoras registradas en cada una de las muestras). Los objetos se encuentran distribuidos entre el conjunto de aprendizaje (las 28 muestras de calibrado) y el conjunto test (las 8 muestras que se encuentran marcadas en la Tabla III.4.1 con una *t*). La variable respuesta ha sido, en cada caso, la concentración del analito de interés.

En estas condiciones los modelos PLS se han construido a partir de los datos originales (sin normalizar) y utilizando un método de validación cruzada [22, 23] para evaluar la estabilidad y la capacidad de predicción de los modelos construidos.

Evaluación de los modelos PLS

Los valores de varianza explicada y varianza en predicción correspondientes a los modelos PLS construidos para ambos calibrados se muestran en las Tablas III.4.4 y III.4.5. Analizando en primer lugar los resultados del calibrado C, y según el criterio de hacer máxima la varianza explicada en predicción, se considera que el modelo correspondiente al cobre se debe construir tomando 8 variables latentes, 7 en los modelos correspondientes a plomo y cadmio, y 6 en el modelo desarrollado para el zinc. Se puede observar cómo los valores recogidos en

esta tabla reflejan la gran variabilidad existente entre los voltamperogramas que constituyen el conjunto de aprendizaje del calibrado C.

En cuanto a los errores correspondientes a las muestras del conjunto test del calibrado C, que se encuentran en la Tabla III.4.6, cabe señalar que al menos 22 de ellos son menores del 5%, y tan sólo uno de ellos está por encima del 10%. En este último caso, que se corresponde con la muestra número 30, destaca el hecho de que todos los errores relacionados con ella son elevados, lo cual señala la posibilidad de que esa muestra pudiera ser un dato anómalo dentro del calibrado C.

Resultados análogos se han obtenido con los modelos PLS del calibrado D en lo que se refiere a los valores de varianza explicada y a los errores relativos de las muestras del conjunto test, que se encuentran en las Tablas III.4.5 y III.4.7, respectivamente. En este caso, son necesarias diez variables latentes en el modelo PLS correspondiente al cobre, ya que los valores correspondientes a la varianza explicada y a la varianza explicada en predicción para la decimoprimer variable latente son 99.90 y 95.38% respectivamente. Los modelos restantes estarán constituidos por 4 variables latentes en el caso del plomo, 8 en el del cadmio y 5 en el del zinc. Los errores correspondientes a las muestras del conjunto test de este calibrado son parecidos a los obtenidos para el calibrado C.

Si se comparan los errores de la Tabla III.4.7 con los obtenidos para esas mismas muestras de calibrado mediante el procedimiento de adición estándar univariante llevado a cabo previamente, Tabla III.4.3, se observa que, en todos los casos, el uso de la regresión multivariante PLS ha mejorado los resultados obtenidos. Por otro lado, una vez que se han preparado los 28 patrones de calibrado y construido los modelos PLS, es posible determinar la concentración de cobre, plomo, cadmio y zinc de cualquier muestra problema en la que la concentración de estos metales se encuentre dentro del rango considerado en el calibrado.

En un intento por dar una interpretabilidad química a las variables latentes, se han analizado los loadings correspondientes a cada uno de los modelos PLS construidos para el calibrado C, representados en las Figuras III.4.3 a III.4.6 para los analitos respectivos, siendo éstos muy similares a los obtenidos para el calibrado D. Estos gráficos señalan cómo la primera variable latente está relacionada, en todos los casos, con las direcciones de máxima variabilidad de las variables predictoras, ya que éstas no han sido autoescaladas previamente, y por ello reproducen la forma de los voltamperogramas registrados. De hecho, la varianza explicada por la primera variable latente es considerable en el caso de los modelos

Tabla III.4.4 Varianza explicada en función del número de variables latentes (V.L.) seleccionadas para los modelos PLS construidos en el calibrado C.

V. L.	Cobre			Plomo		
	Bloque Y		Bloque X	Bloque Y		Bloque X
	Var. explicada	Var. en predicción	Var. explicada	Var. explicada	Var. en predicción	Var. explicada
1	34.72	37.04	82.05	-	-	82.02
2	87.96	87.49	90.94	73.11	6.15	83.76
3	92.01	88.12	93.85	85.33	82.87	93.67
4	92.75	88.31	95.55	94.90	92.45	95.87
5	93.92	88.99	97.37	95.40	93.16	97.71
6	98.19	93.74	98.06	98.35	94.67	98.07
7	98.82	95.19	98.24	98.90	95.95	98.26
8	99.21	95.26	98.44	99.39	94.96	98.44
9	99.64	94.96	98.51	99.77	94.86	98.44
10	99.87	94.66	98.60	99.91	94.59	98.66

V. L.	Cadmio			Zinc		
	Bloque Y		Bloque X	Bloque Y		Bloque X
	Var. explicada	Var. en predicción	Var. explicada	Var. explicada	Var. en predicción	Var. explicada
1	-	-	82.02	49.28	48.11	82.02
2	46.03	16.18	84.50	94.31	94.14	90.92
3	68.84	68.34	93.12	96.18	95.25	93.85
4	91.04	84.38	95.69	96.53	94.83	95.69
5	93.34	88.82	97.89	97.55	94.48	96.79
6	96.82	93.29	98.17	98.59	97.49	98.17
7	97.90	94.48	98.35	98.97	97.37	98.35
8	98.87	92.55	98.42	99.59	97.15	98.44
9	99.72	92.73	98.43	99.85	97.15	98.46
10	99.90	92.48	98.44	99.90	97.04	98.59

Tabla III.4.5 Varianza explicada en función del número de variables latentes (V.L.) seleccionadas para los modelos PLS construidos en el calibrado D.

V. L.	Cobre			Plomo		
	Bloque Y		Bloque X	Bloque Y		Bloque X
	Var. explicada	Var. en predicción	Var. explicada	Var. explicada	Var. en predicción	Var. explicada
1	37.51	39.27	81.82	–	–	81.82
2	90.20	89.53	90.36	83.38	56.20	82.91
3	93.53	92.80	93.18	87.48	86.47	91.45
4	94.24	93.22	95.36	97.86	97.60	95.45
5	94.53	93.70	96.91	97.99	97.50	97.27
6	96.06	94.64	97.64	98.73	96.87	97.45
7	97.65	95.16	98.09	99.26	97.28	97.64
8	98.60	95.26	98.36	99.37	97.37	98.36
9	99.50	95.45	98.55	99.65	97.20	98.51
10	99.77	95.54	98.55	99.87	96.94	98.55

V. L.	Cadmio			Zinc		
	Bloque Y		Bloque X	Bloque Y		Bloque X
	Var. explicada	Var. en predicción	Var. explicada	Var. explicada	Var. en predicción	Var. explicada
1	–	–	81.82	48.25	47.62	81.82
2	50.36	–	83.27	94.81	94.99	90.36
3	63.07	63.25	92.18	97.41	97.21	93.55
4	89.12	84.64	95.09	97.63	97.08	95.18
5	91.77	90.84	97.27	97.96	97.29	96.45
6	95.67	91.84	97.55	98.41	96.79	97.73
7	96.81	92.90	98.09	98.85	96.88	98.18
8	97.97	94.08	95.45	99.20	96.75	95.45
9	99.26	93.97	98.51	99.73	96.51	98.49
10	99.62	93.55	98.55	99.90	96.57	98.51

Tabla III.4.6 Errores relativos (en %) obtenidos para las muestras del conjunto test con los modelos PLS del calibrado C.

Número de muestra	Analito			
	Cobre	Plomo	Cadmio	Zinc
18	1.07	6.71	1.96	1.47
20	2.96	2.86	0.63	0.50
23	-4.76	-8.03	-7.78	-0.82
25	1.61	2.00	-2.95	-1.83
28	6.01	3.21	-7.71	8.85
30	11.38	7.90	7.38	5.24
33	-1.55	-0.41	-3.88	0.12
35	3.28	3.94	-0.07	0.11

Tabla III.4.7 Errores relativos (en %) obtenidos para las muestras del conjunto test con los modelos PLS del calibrado D.

Número de muestra	Analito			
	Cobre	Plomo	Cadmio	Zinc
18	-2.09	-9.31	2.84	-3.52
20	-13.08	-5.55	-1.99	1.79
23	0.60	2.11	10.81	6.44
25	4.95	7.81	5.88	5.10
28	-9.00	-1.47	-5.38	-0.30
30	0.40	-2.53	-3.83	2.85
33	5.79	-2.37	0.78	0.05
35	-0.14	2.97	-3.69	3.00

construídos para el cobre y el zinc, mientras que no es significativa en los modelos correspondientes al plomo y al cadmio. Esto se debe a que la magnitud de las variables predictoras relacionadas con estos dos últimos analitos es mucho menor que la de los dos primeros.

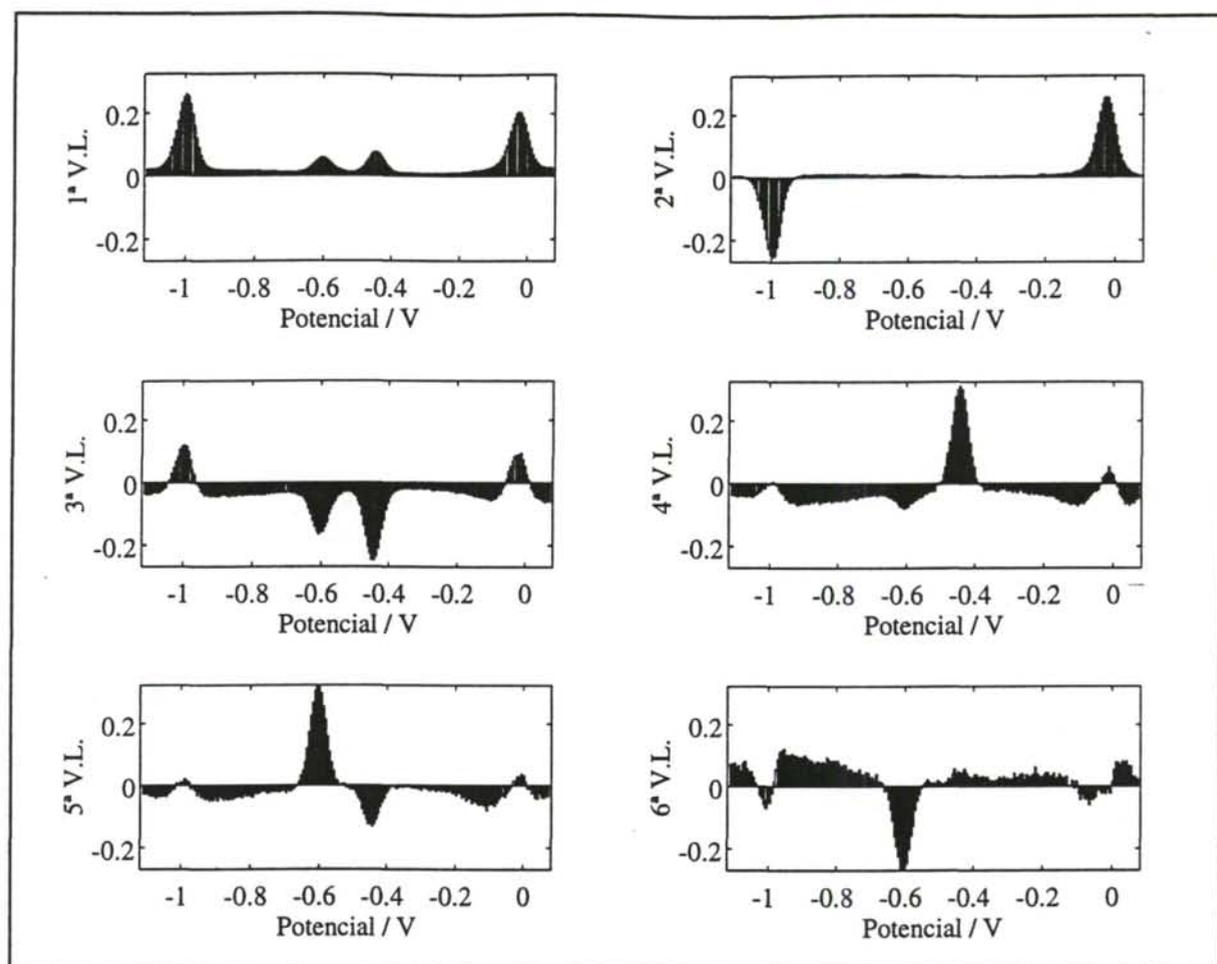


Figura III.4.3 Loadings correspondientes al modelo PLS construido para el cobre en el calibrado C.

Por el contrario, la segunda variable latente parece estar relacionada con la concentración del analito que determina cada uno de los modelos construidos, estando asociada en todos los casos con incrementos importantes en los valores de varianza explicada. En las Figuras III.4.3 y III.4.6, correspondientes a cobre y zinc respectivamente, se puede observar cómo la segunda variable latente está modelando en ambos casos la interacción existente entre ambos analitos, que posiblemente esté relacionada con la formación de compuestos intermetálicos. Los loadings representados en esas dos figuras evidencian la oposición existente entre las concentraciones de ambos analitos y sus señales respectivas, la presencia o ausencia de uno de ellos implica una disminución o un incremento de la señal correspondiente al otro. Las variables latentes restantes poseen estructuras idénticas en ambos modelos, considerándose en éstas la influencia que los demás analitos tienen sobre la determinación del analito de interés y otras posibles fuentes de variabilidad.

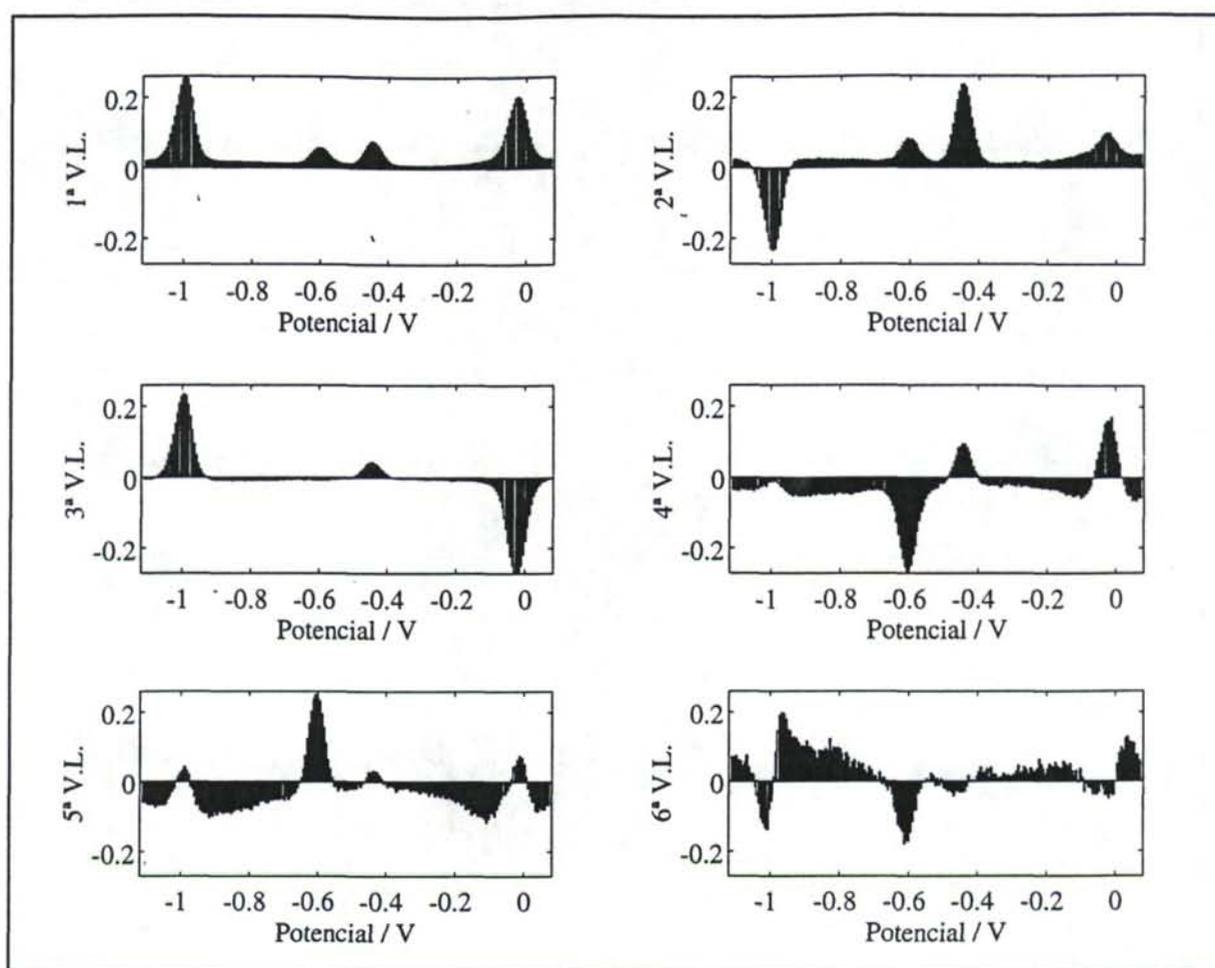


Figura III.4.4 Loadings correspondientes al modelo PLS construido para el plomo en el calibrado C.

En los modelos correspondientes a plomo y cadmio la varianza en predicción para la segunda variable latente no experimenta el mismo incremento que la varianza explicada. Ello parece indicar que, para estos dos analitos, los modelos construidos con dos variables latentes se ajustan excesivamente a los objetos que forman cada uno de los grupos de cancelación utilizados, siendo por esta razón incapaces de modelar fenómenos asociados con objetos que se encuentren fuera de éstos. No obstante, esta componente parece estar relacionada en ambos casos con el modelado de la señal correspondiente al zinc frente al resto del voltamperograma, aunque para completar el modelo PLS se hace necesaria la inclusión de otras variables latentes. De hecho, los loadings correspondientes a la tercera variable latente tienen una estructura muy similar a la observada para la segunda variable latente en los modelos construidos para cobre y zinc, evidenciando el hecho de que también los modelos PLS construidos para plomo y cadmio se ven en la necesidad de modelar esa interferencia.

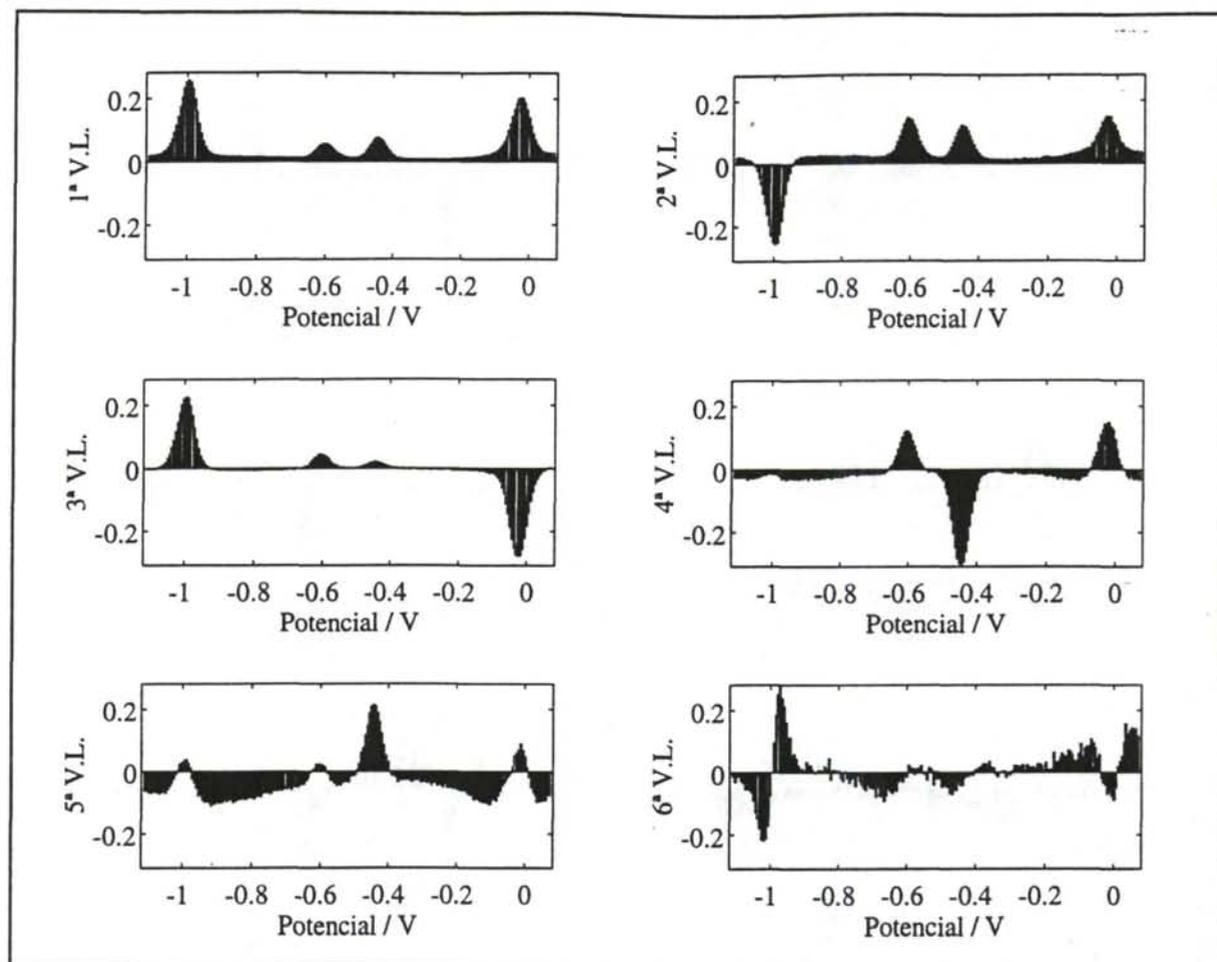


Figura III.4.5 Loadings correspondientes al modelo PLS construido para cadmio en el calibrado C.

La cuarta variable latente tiene un peso importante en los modelos construidos para cadmio y plomo, mientras que las variables latentes restantes suponen pequeños incrementos de varianza explicada y tienen una estructura menos definida, por lo que parecen modelar pequeños matices de variabilidad de las variables predictoras.

III.4.4.2 Transferencia de calibrado mediante el procedimiento PDS

Este método de transferencia de calibrado se va a utilizar para transferir el calibrado C a las muestras obtenidas en la situación D. Por lo tanto, se trata de nuevo de una transferencia de calibrado a lo largo del tiempo, con la que se trata de reducir el número de muestras necesarias para determinar estos cuatro metales en calibrados posteriores.

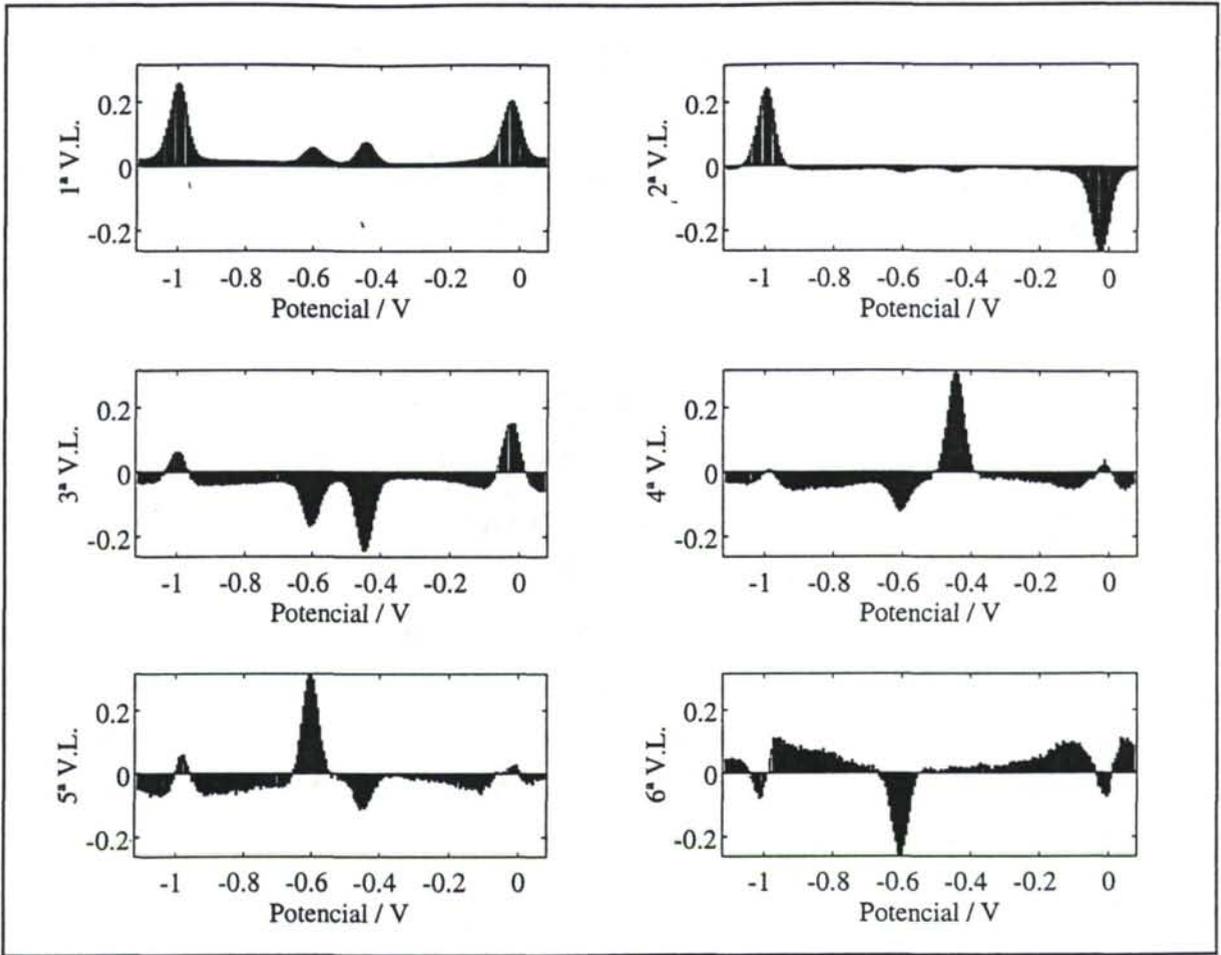


Figura III.4.6 Loadings correspondientes al modelo PLS construido para el zinc en el calibrado C.

Aplicación de la función "pdsgen"

Para llevar a cabo el procedimiento de transferencia de calibrado se ha utilizado nuevamente la función "pdsgen" definida anteriormente. Se ha analizado la influencia sobre el procedimiento de transferencia de calibrado de algunos tamaños de ventana y diferentes conjuntos de transferencia con el fin de encontrar sendos valores de ambos parámetros que resulten adecuados para llevarlo a cabo. En este caso se han utilizado los parámetros de entrada/salida siguientes:

$$[F, \text{sub}] = \text{pdsgen}(C, D, w, n_0)$$

los cuales se corresponden con

C: Matriz con los voltamperogramas del calibrado C, es decir, los registrados en la situación C. Tiene de dimensiones 36 filas (muestras) por 198 columnas (potenciales).

- D:** Matriz con los voltamperogramas del calibrado D, es decir, los obtenidos en la situación D. Tiene idénticas dimensiones que C, y sus elementos se encuentran en el mismo orden.
- w:** Escalar que indica el número de variable predictoras que forman parte de la ventana móvil que se utiliza en el cálculo de la matriz de transferencia. En las transferencias de calibrado llevadas a cabo en el análisis se han utilizado los siguientes tamaños de ventana: 3, 5, 7, 9, 11 y 13.
- no:** Vector fila en el que se indican los índices de las muestras que forman cada uno de los subconjuntos de transferencia, los cuales se hallan especificados en la Tabla III.3.8.

Como resultado de los distintos procedimientos de transferencia llevados a cabo con los distintos *no* y *w* se dispone de un total de 72 matrices de transferencia diferentes e independientes entre sí, todas ellas con dimensiones 198×198 , que se han utilizado posteriormente para transformar los voltamperogramas correspondientes a las muestras del calibrado D. Una vez que se dispone de las señales del calibrado D corregidas, se ha calculado la concentración de los cuatro analitos en las mismas, y a partir de estos valores se han estimado los correspondientes a SEP y SET.

Análisis de los resultados

Para estudiar el efecto que tienen el tamaño de ventana y los subconjuntos de transferencia sobre los procedimientos de transferencia de calibrado llevados a cabo, se han tenido en consideración los valores de SEP y SET obtenidos en dichos procedimientos. Los valores del SEP, cuya variación respecto de *no* y *w* sigue las mismas pautas que los correspondientes al SET, se encuentran representados a través de las superficies de respuesta y las curvas de nivel en las Figuras III.4.7 a III.4.10 para cada uno de los analitos. Parece claro que, en las transferencias de calibrado realizadas con determinados conjuntos de transferencia, la influencia del tamaño de ventana sobre las mismas es nulo o prácticamente despreciable, como ocurre para los subconjuntos de transferencia 1, 4, 5, 6 y 7, que proporcionan buenos resultados en casi todos los casos. Sin embargo, los valores de SEP obtenidos en las transferencias de calibrado llevadas a cabo con otros conjuntos de transferencia se encuentran más ligados al tamaño de ventana seleccionado.

Un análisis más detallado de la dependencia de los valores de SEP con *no* y *w* se ha realizado mediante la representación de los diagramas Box plot correspondientes, Figuras

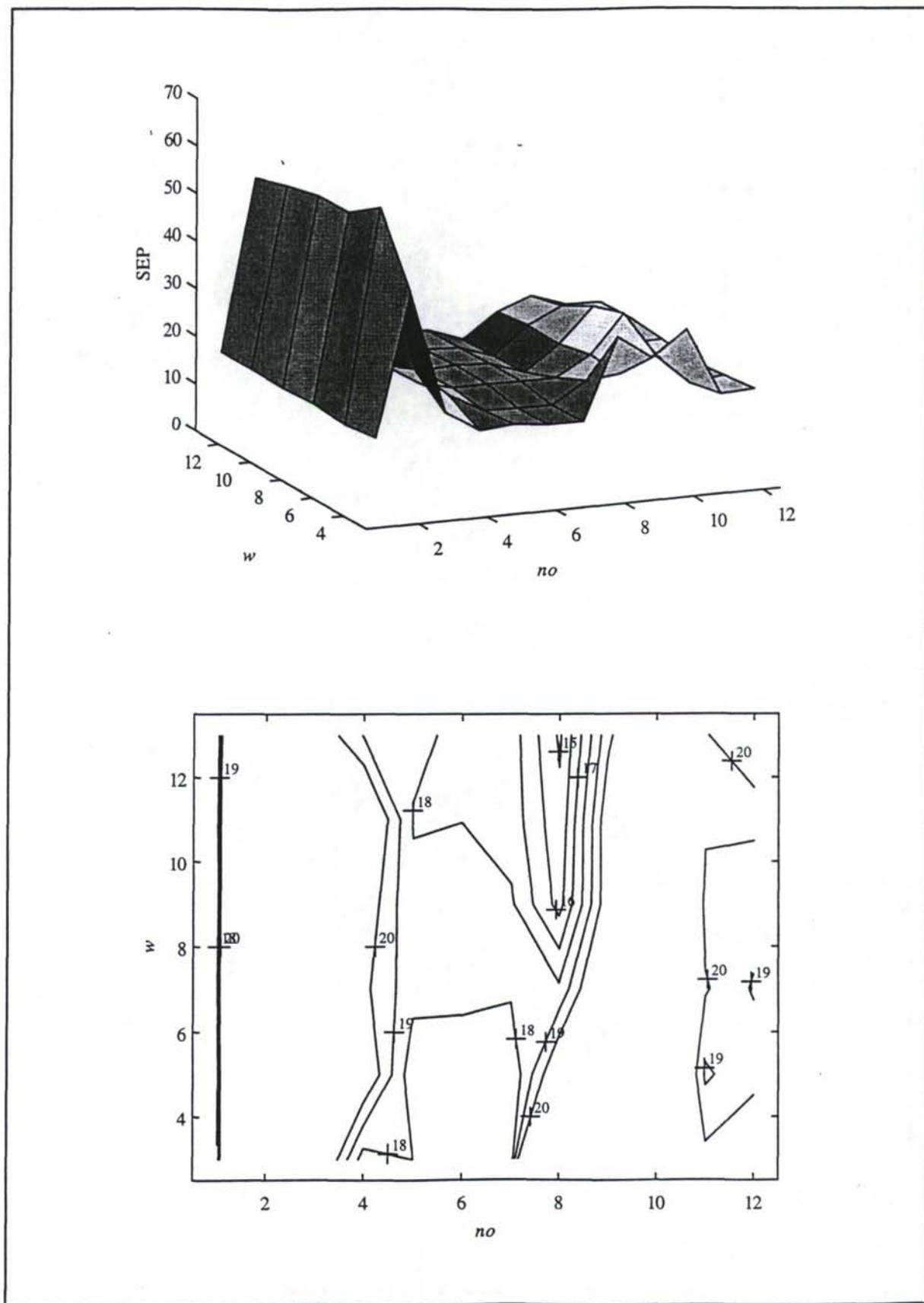


Figura III.4.7 Superficie de respuesta y curvas de nivel del SEP para el cobre.

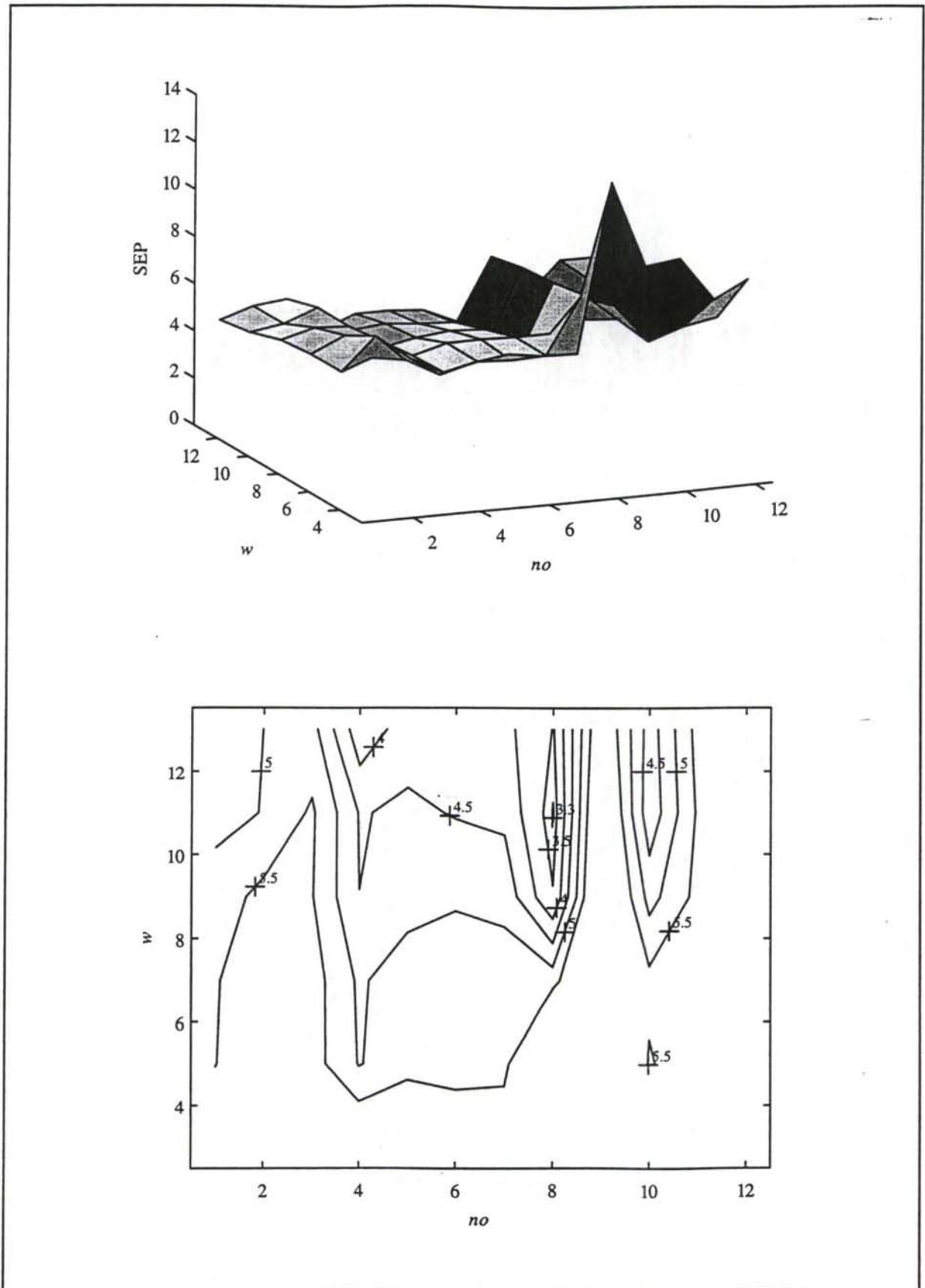


Figura III.4.8 Superficie de respuesta y curvas de nivel del SEP para el plomo.

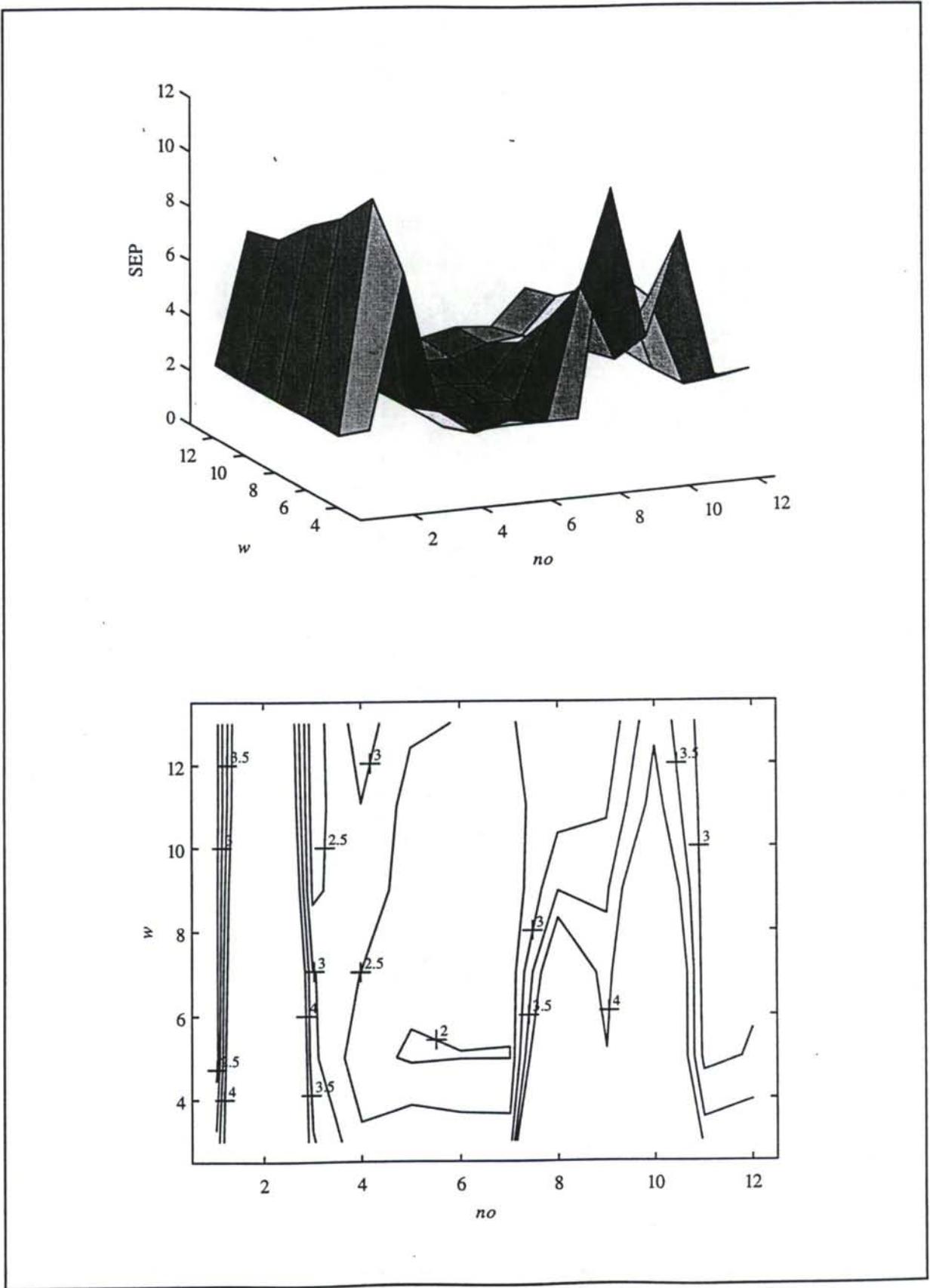


Figura III.4.9 Superficie de respuesta y curvas de nivel del SEP para el cadmio.

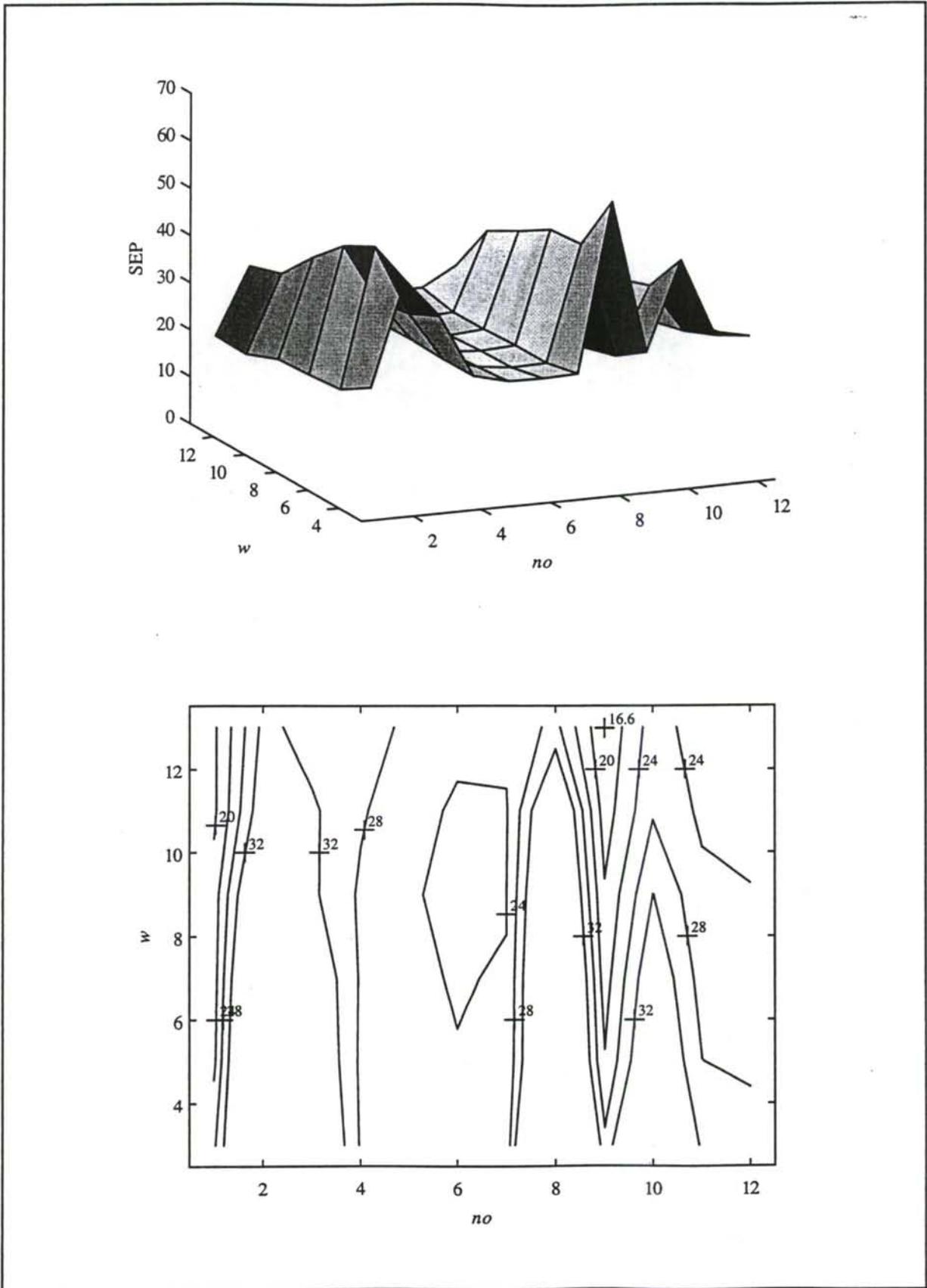


Figura III.4.10 Superficie de respuesta y curvas de nivel del SEP para el zinc.

III.4.11 a III.4.14, y los test de agrupamientos para la media de Tukey y Newman-Keuls, que se encuentran en las Tablas III.4.8 a III.4.11.

En las gráficas correspondientes a los diagramas Box plot, éstos parecen indicar que los tamaños de ventana mayores se asocian con errores más pequeños, salvo en el caso del cadmio, para el que el valor $w = 5$ proporciona el mínimo valor de SEP. Los test de Tukey y Newman-Keuls confirman esta interpretación, señalando el hecho de que no existen diferencias estadísticamente significativas entre los valores de SEP que proporcionan los tamaños de ventana mayores, incluso en lo que se refiere al cadmio. De hecho, las curvas de nivel de las Figuras III.4.7 a III.4.10 indican que, para muchos de los conjuntos de transferencia, los valores mínimos de SEP se pueden relacionar con los tamaños de ventana mayores, y en los casos en que no es así, la influencia de w sobre los valores de SEP obtenidos es mínima. Por tanto, parece lógico analizar con más detalle las transferencias de calibrado que se corresponden con esos tamaños de ventana, por esta razón se ha seleccionado el valor $w = 13$.

Sin embargo, como ya se ha señalado anteriormente, los resultados de los procedimientos de transferencia realizados con los diferentes subconjuntos de transferencia dependen en gran medida del analito de que se trate. En las curvas de nivel correspondientes a cobre y plomo se observa cómo el conjunto de transferencia 8 es el que está ligado en ambos casos con los valores de SEP mínimos, mientras que para el cadmio son los conjuntos 5, 6 y 7 los que proporcionan los SEP más bajos, y por último, es el subconjunto 9 en el caso del zinc. Los diagramas Box plot permiten evaluar más claramente la influencia de estos conjuntos sobre los valores de SEP frente al resto de los conjuntos de transferencia, y conducen a idénticas conclusiones. Incluso los test de agrupamientos llevados a cabo no encuentran diferencias estadísticamente significativas entre ellos, salvo en casos muy concretos.

Con el fin de tener suficientes herramientas de juicio para evaluar la calidad de las determinaciones llevadas a cabo mediante los procedimientos de transferencia de calibrado señalados en último lugar, se han tenido en consideración los errores correspondientes a las muestras del conjunto test. Entre los resultados a que conducen los conjuntos de transferencia 5, 6 y 7, no existen en ningún caso diferencias estadísticamente significativas, según señalan los test de agrupamientos realizados, y en estas condiciones, el conjunto de transferencia más indicado debería ser el que implica menor esfuerzo experimental, es decir, de esos tres conjuntos de transferencia el mejor sería el conjunto 5. Por ello, en la Tabla III.4.12 se

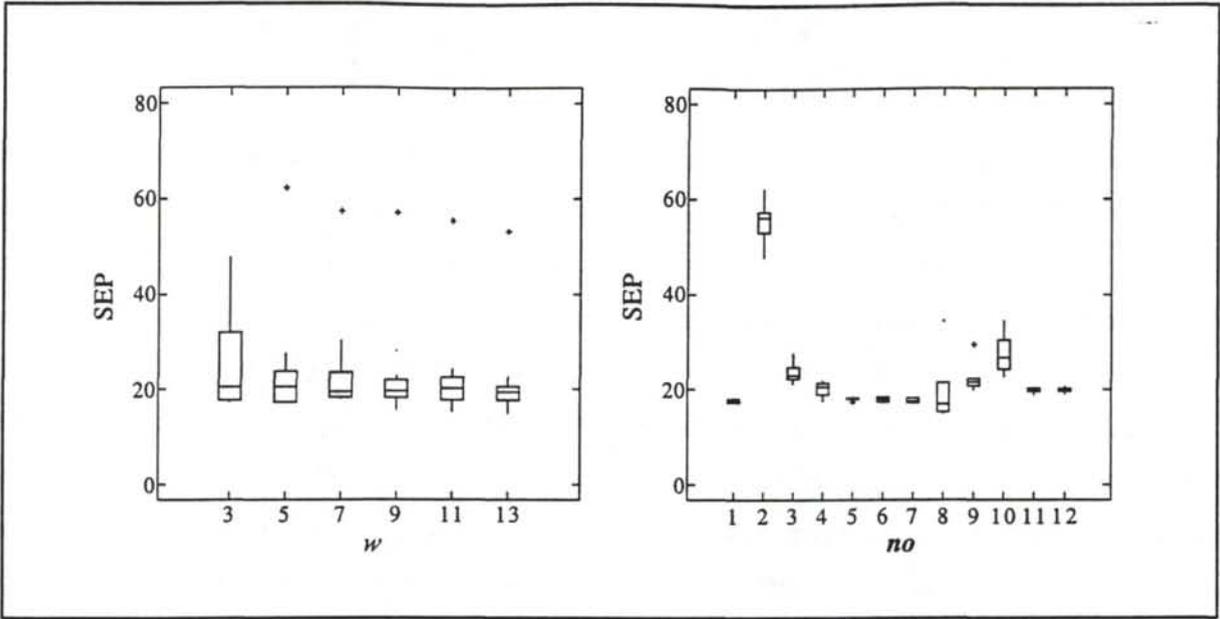


Figura III.4.11 Diagramas Boxplot correspondientes a los valores del SEP obtenidos en la transferencia de calibrado realizada para el cobre.

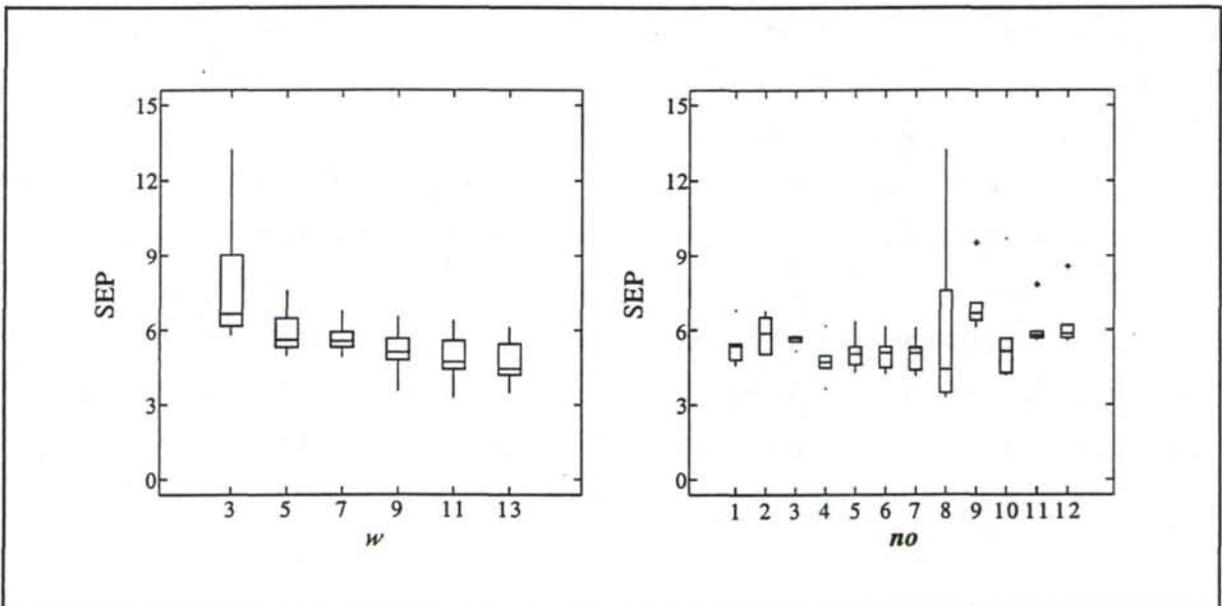


Figura III.4.12 Diagramas Boxplot correspondientes a los valores del SEP obtenidos en la transferencia de calibrado realizada para el plomo.

muestran los errores correspondientes a las muestras del conjunto test de las transferencias de calibrado llevadas a cabo con los subconjuntos de transferencia 5, 8 y 9, y con un tamaño de ventana 13.

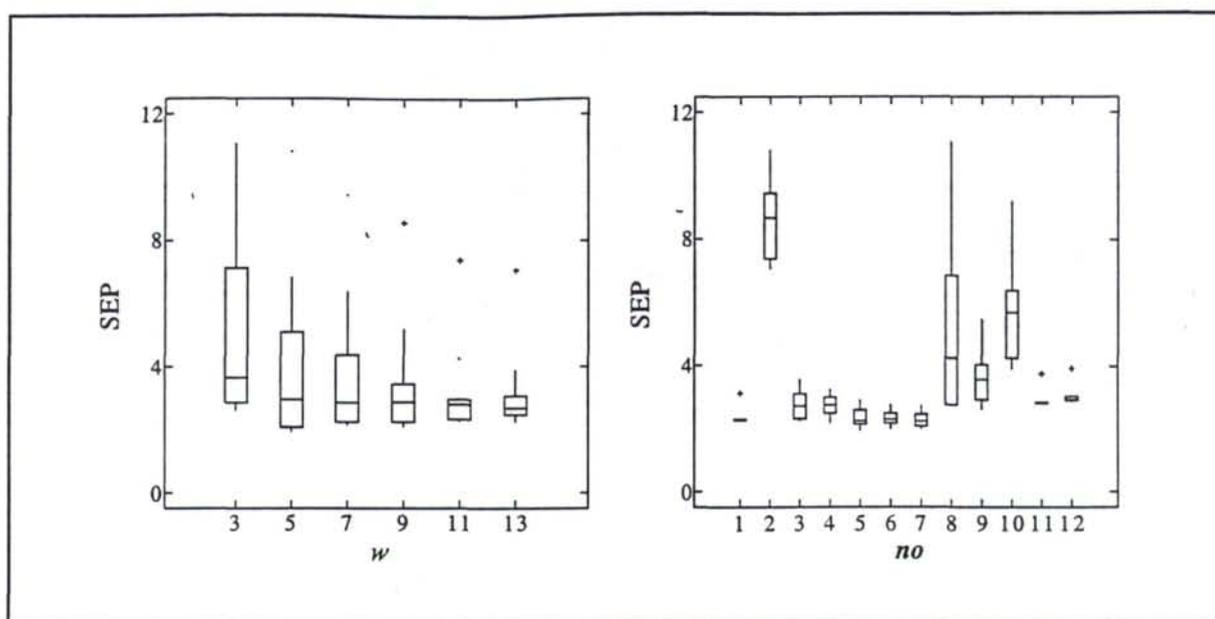


Figura III.4.13 Diagramas Boxplot correspondientes a los valores del SEP obtenidos en la transferencia de calibrado realizada para el cadmio.

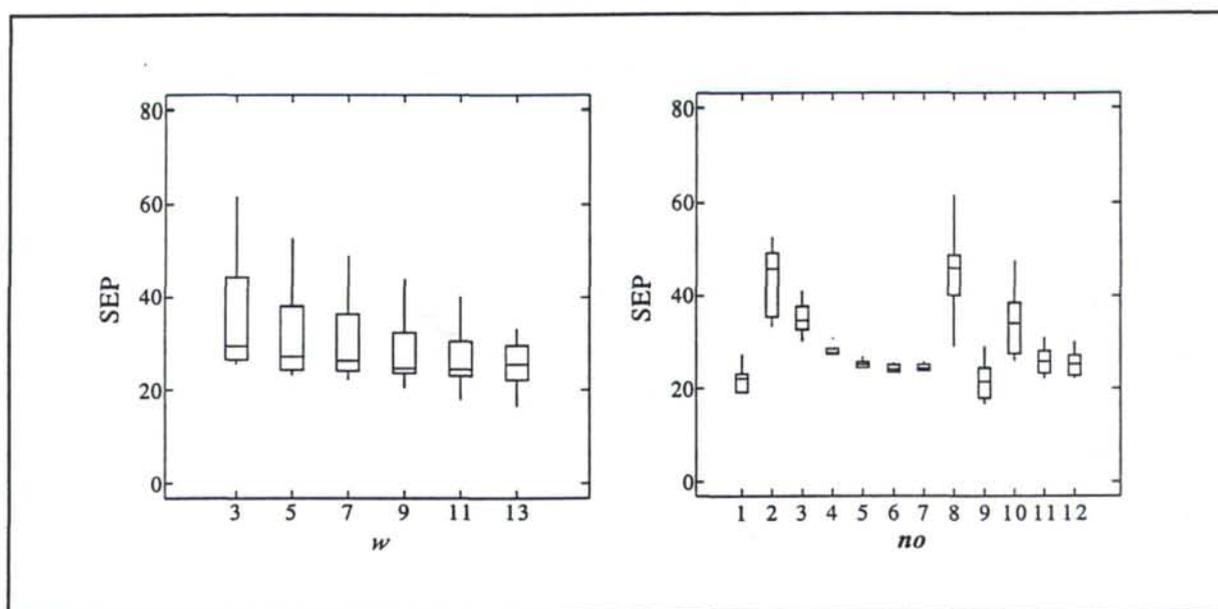


Figura III.4.14 Diagramas Boxplot correspondientes a los valores del SEP obtenidos en la transferencia de calibrado realizada para el zinc.

En la Tabla III.4.13, se muestran los valores de SEP y SET obtenidos con los conjuntos de transferencia señalados anteriormente, en las condiciones de transferencia de calibrado indicadas, además de los calculados a partir de las concentraciones obtenidas por los modelos PLS construídos inicialmente con los calibrados C y D. Para establecer

Tabla III.4.8 Tests de agrupamientos de Tukey y Newman-Keuls para los valores de SEP correspondientes al cobre.

Conjunto de transferencia (<i>no</i>)				Ventana (<i>w</i>)			
Nivel	Media	Grupos homogéneos		Nivel	Media	Grupos homogéneos	
		Tukey	Newman-Keuls			Tukey	Newman-Keuls
1	17.5758	X	X	13	21.7258	X	X
7	17.6106	X	X	11	22.6840	X	X
6	17.8800	X	X	9	23.0345	X	X
5	17.9823	X	X	7	23.7479	X	X
12	19.8091	X	XX	5	23.9433	X	X
11	19.8438	X	XX	3	24.8256	X	X
8	19.9704	X	XX				
4	20.1236	X	XX				
9	22.5322	XX	XX				
3	23.5742	XX	X				
10	27.5196	X	X				
2	55.5011	X	X				

comparaciones con estos últimos valores hay que tener en cuenta el hecho de que los modelos PLS fueron construidos específicamente para cada uno de los conjuntos de datos de los calibrados C y D, y que por tanto, en estos casos los valores del SET deberían corresponderse con el mejor ajuste posible. Por otro lado, los valores de SEP y SET obtenidos como resultado del procedimiento del transferencia son comparables en cualquier caso con los correspondientes al calibrado D, ya que éstos evalúan exactamente las mismas muestras de calibrado.

El análisis conjunto de ambas tablas confirma las consideración realizadas anteriormente. Así, el conjunto de transferencia 8 proporciona los mejores resultados para cobre y plomo, siendo éstos comparables a los obtenidos por sus propios modelos (modelos PLS del calibrado D). A este respecto, destaca lo amplia que es la diferencia existente entre los valores de SEP y SET de los modelos PLS construidos para el cobre en ambos calibrados, respecto del resto de los analitos. Ello se debe al elevado número de variables latentes que componen cada uno de los modelos PLS correspondientes al cobre ya que, cuanto mayor sea

Tabla III.4.9 Tests de agrupamientos de Tukey y Newman-Keuls para los valores de SEP correspondientes al plomo.

Conjunto de transferencia (<i>no</i>)				Ventana (<i>w</i>)			
Nivel	Media	Grupos homogéneos		Nivel	Media	Grupos homogéneos	
		Tukey	Newman-Keuls			Tukey	Newman-Keuls
4	4.7790	X	X	13	4.6882	X	X
7	5.0023	X	X	11	4.8594	X	XX
6	5.0509	X	X	9	5.1483	X	XX
5	5.1029	XX	X	7	5.6345	X	XX
1	5.3874	XX	XX	5	5.9159	X	X
3	5.5882	XX	XX	3	7.7252	X	X
10	5.6713	XX	XX				
2	5.8461	XX	XX				
8	6.0693	XX	XX				
11	6.0880	XX	XX				
12	6.2726	XX	XX				
9	7.0849	X	X				

éste, más se ajustan las regresiones multivariantes a los conjuntos de aprendizaje utilizados en la construcción de las mismas, proporcionando por tanto errores menores para las muestras del conjunto de aprendizaje. Para evitar que se produzca un sobreajuste de la regresión, esta característica de los modelos PLS ha de evaluarse simultáneamente con la capacidad de predicción de los mismos, para evitar la posibilidad de obtener modelos de regresión que evalúen perfectamente el conjunto de aprendizaje utilizado en su construcción, pero con nula capacidad para predecir muestras ajenas al mismo.

En cuanto a las transferencias de calibrado realizadas en la determinación de zinc, indudablemente, el conjunto que habría que seleccionar como óptimo es el conjunto 9, ya que los conjuntos de transferencia 5 y 8 proporcionan valores de SEP mucho mayores, e incluso, el conjunto 8 conlleva un valor de SET muy elevado. En el análisis llevado a cabo para este analito, el conjunto de transferencia 9 mejora los resultados obtenidos incluso por el propio modelo PLS del calibrado D, que se encuentran en la Tabla III.4.7, aunque esto no se puede hacer extensivo al resto de los metales.

Tabla III.4.10 Tests de agrupamientos de Tukey y Newman-Keuls para los valores de SEP correspondientes al cadmio.

Conjunto de transferencia (<i>no</i>)				Ventana (<i>w</i>)			
Nivel	Media	Grupos homogéneos		Nivel	Media	Grupos homogéneos	
		Tukey	Newman-Keuls			Tukey	Newman-Keuls
7	2.2909	X	X	13	3.1091	X	X
6	2.3340	X	X	11	3.1406	X	X
5	2.3438	X	X	9	3.3716	X	X
1	2.4198	X	X	7	3.7100	X	X
4	2.7311	X	X	5	3.9340	XX	X
3	2.7723	X	X	3	4.9949	X	X
11	2.9798	X	X				
12	3.1036	X	X				
9	3.6974	XX	X				
8	5.3228	XX	X				
10	5.8554	X	X				
2	8.6697	X	X				

Sin embargo, los valores de SEP obtenidos para el cadmio con esos tres conjuntos de transferencia no se diferencian apenas entre sí, ni tampoco de los que resultaron con los modelos PLS de ambos calibrados. En este caso, las diferencias más importantes se encuentran en los valores de SET, los cuales señalan como conjunto óptimo el conjunto de transferencia 5, tal y como ya habían concluido los análisis realizados anteriormente para este analito.

En principio, a la vista de los resultados, resulta evidente que no existe un único conjunto de transferencia que proporcione resultados aceptables para los cuatro analitos al mismo tiempo. No hay que olvidar que la finalidad que se persigue al utilizar este procedimiento de transferencia de calibrado es disminuir el esfuerzo experimental requerido cuando se trata de recalibrar un mismo instrumento a lo largo del tiempo. Por tanto, la necesidad de disponer de tres conjuntos de transferencias distintos para llevar a cabo la determinación de los cuatro analitos supondría la necesidad de disponer de un total de 14 muestras de calibrado, ya que los subconjuntos de transferencia 8 y 9 se encuentran incluidos

Tabla III.4.11 Tests de agrupamientos de Tukey y Newman-Keuls para los valores de SEP correspondientes al zinc.

Conjunto de transferencia (<i>no</i>)				Ventana (<i>w</i>)			
Nivel	Media	Grupos homogéneos		Nivel	Media	Grupos homogéneos	
		Tukey	Newman-Keuls			Tukey	Newman-Keuls
9	21.7560	X	X	13	25.4949	X	X
1	22.1254	X	X	11	26.6259	XX	X
6	24.2151	X	X	9	28.4894	XXX	XX
7	24.4741	X	X	7	30.5734	XXX	X
5	25.1898	X	X	5	31.7373	XX	X
12	25.3841	X	X	3	35.1217	X	X
11	25.9384	X	X				
4	28.1853	XX	X				
10	34.6296	X	X				
3	35.1591	X	X				
2	43.7387	X	X				
8	45.2896	X	X				

en el conjunto 5, lo que implicaría la mitad de las muestras correspondientes a una recalibración completa.

Por otro lado, también se podría contemplar la posibilidad de emplear únicamente los subconjuntos de transferencia 8 y 9 en el procedimiento de transferencia de calibrado, utilizando el conjunto 9 para la transferencia de calibrado del cadmio y el zinc, y el 8 en las correspondientes a cobre y plomo. Esto supondría que únicamente sería necesario medir 6 muestras patrón en la nueva situación. En este sentido, se ha aplicado el procedimiento de transferencia de calibrado completo utilizando un conjunto de transferencia formado por los subconjuntos 8 y 9, sin que los resultados hayan mejorado los obtenidos anteriormente al usar los dos subconjuntos de transferencia por separado en los diferentes procedimientos de transferencia de calibrado llevados a cabo.

En este caso, la gran variabilidad observada entre las variables predictoras de uno y otro calibrado, que también quedó reflejada en los modelos PLS construidos en un primer análisis, influye de manera importante en el procedimiento de transferencia de calibrado

Tabla III.4.12 Errores relativos (en %) obtenidos para las muestras del conjunto test en la transferencia de calibrado llevada a cabo con un tamaño de ventana 13.

<i>no</i>	Número de muestra	Analito			
		Cobre	Plomo	Cadmio	Zinc
5	18	10.26	-3.75	10.59	-4.74
	20	-16.64	-9.73	-3.09	2.44
	23	-1.11	-5.52	4.19	7.12
	25	3.00	7.93	3.70	7.74
	28	-5.04	-4.87	-6.13	2.66
	30	-3.89	-1.38	-3.52	5.67
	33	6.21	-0.80	0.47	1.48
	35	3.68	1.24	2.44	6.68
8	18	0.92	-4.79	6.95	-12.07
	20	-8.93	-7.46	0.16	6.03
	23	2.89	-2.85	1.84	-4.21
	25	5.25	4.34	-1.24	-1.96
	28	-3.36	-4.92	-1.48	-6.28
	30	-7.72	-5.82	-9.80	1.12
	33	11.60	-0.11	1.31	-7.99
	35	6.64	0.19	-0.58	-4.14
9	18	-10.34	-13.19	-4.33	-8.48
	20	-14.51	-11.25	-6.93	-2.91
	23	-6.96	-11.71	-2.76	0.45
	25	3.79	4.19	2.08	1.98
	28	-10.54	-9.16	-12.16	-1.19
	30	-10.76	-6.45	-5.44	2.90
	33	4.74	-3.56	-2.02	-2.28
	35	5.38	-0.02	1.46	0.91

Tabla III.4.13 Valores de SEP y SET correspondientes a las transferencias de calibrado llevadas a cabo con el procedimiento PDS.

no	Error	Analito			
		Cobre	Plomo	Cadmio	Zinc
5	SEP	18.12	4.27	2.58	26.77
	SET	17.97	4.78	3.80	22.83
8	SEP	14.83	3.48	2.75	29.14
	SET	16.39	4.45	6.36	40.82
9	SEP	19.79	6.10	2.61	15.58
	SET	23.90	6.13	5.55	17.95
Calibrado D	SEP	15.12	4.19	2.70	17.28
	SET	2.40	3.13	1.92	18.88
Calibrado C	SEP	10.63	3.52	2.66	17.73
	SET	4.71	1.88	2.01	15.33

realizado. Evidentemente, este procedimiento no puede llegar a considerar aquellas fuentes de variabilidad que no pudieron ni siquiera ser modeladas por los calibrados PLS construidos al comienzo del análisis multivariante. Es decir, este procedimiento de transferencia de calibrado se va a encontrar siempre con la limitación impuesta por la calidad del ajuste que se haya realizado en el modelo de regresión multivariante construido con los datos del calibrado patrón. Del mismo modo, posibles fuentes de variabilidad que puedan existir en los datos de la situación satélite, siempre y cuando éstas no hayan sido modeladas de forma simultánea por las regresiones PLS construidas a partir de la situación patrón, ejercerá inevitablemente un efecto desestabilizador en los resultados ofrecidos por este u otros procedimientos de transferencia de calibrado.

III.4.4.3 Transferencia de calibrado mediante el procedimiento GCT

La transferencia de calibrado entre las situaciones C y D se ha realizado también a través del procedimiento GCT. En este caso, para llevar a cabo el análisis se han utilizado nuevamente

las matrices **C** y **D**, así como los subconjuntos de transferencia que se indican en la Tabla III.3.8, es decir, los mismos que en el apartado anterior.

Aplicación de la función "gct"

La función "gct", definida anteriormente, se ha utilizado para transferir el calibrado **C** a la situación **D**. Los parámetros de entrada/salida introducidos han sido:

$$[F, sub] = gct(C, D, 15, no, 3, Acer)$$

En donde los parámetros **C**, **D**, **no** y **sub** coinciden con lo señalado en el procedimiento de transferencia de calibrado utilizado en el apartado precedente y, como ya se ha indicado anteriormente, los valores 15 y 3 señalan, respectivamente, el número máximo de variables latentes y el número de grupos de cancelación del modelo PLS construido para calcular la matriz de transformación de los coeficientes, **F**, que tiene de dimensiones 198×198 .

Este procedimiento de transferencia de calibrado permite calcular las concentraciones de las muestras del calibrado **D**, utilizando para ello el modelo de regresión de la situación **C** corregido. Por tanto, también en este caso se dispone de los valores de SEP y SET, que se han obtenido como resultado de aplicar procedimientos de transferencia independientes, en cada uno de los cuales se utiliza un conjunto de transferencia diferente. Los valores correspondientes a ambos errores se encuentran representados, para cada uno de los analitos a determinar, en las Figuras III.4.15 a III.4.18.

Análisis de los resultados

Dichas figuras señalan cómo el conjunto de transferencia que conduce a los mejores resultados, prácticamente para todos los analitos, es el conjunto 1, en el cual están incluidas las 28 muestras del conjunto de aprendizaje utilizado en la construcción de los modelos PLS. En este caso, este procedimiento de transferencia de calibrado, que como su propio nombre indica, es un procedimiento global, ofrece mejores resultados cuando dispone de toda la información posible.

Ninguno de los conjuntos de transferencia restantes destaca por estar asociado con errores bajos en las transferencias llevadas a cabo para los cuatro analitos. En todo caso, los conjuntos 5 y 9 podrían seleccionarse como ejemplos de las tendencias que se observan en dichas figuras. Los resultados asociados con el conjunto 5 coinciden bastante con los correspondientes a los conjuntos 6 y 7, siendo bajos, en general, los valores de SET para estos

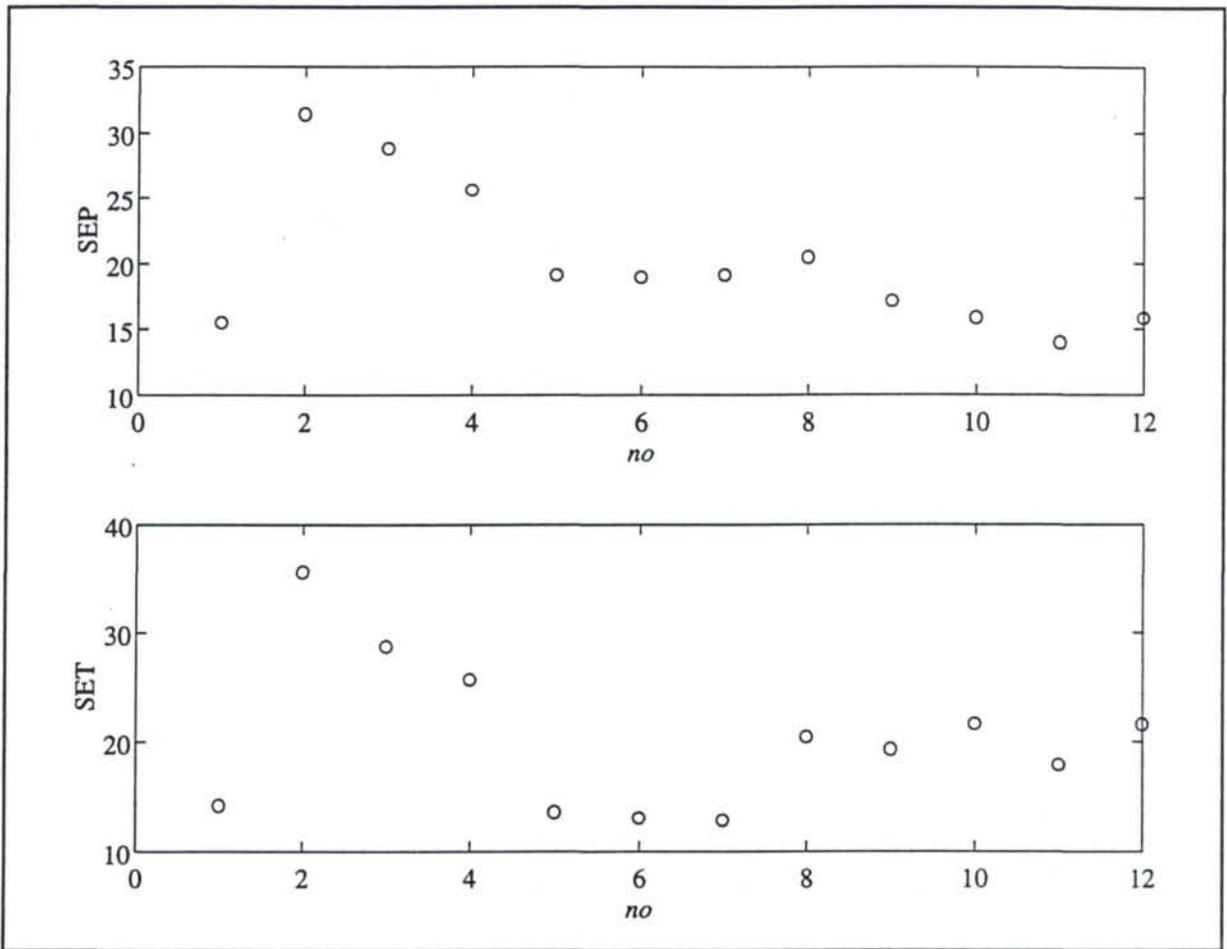


Figura III.4.15 Resultados de la transferencia de calibrado realizada con el procedimiento GCT para el cobre.

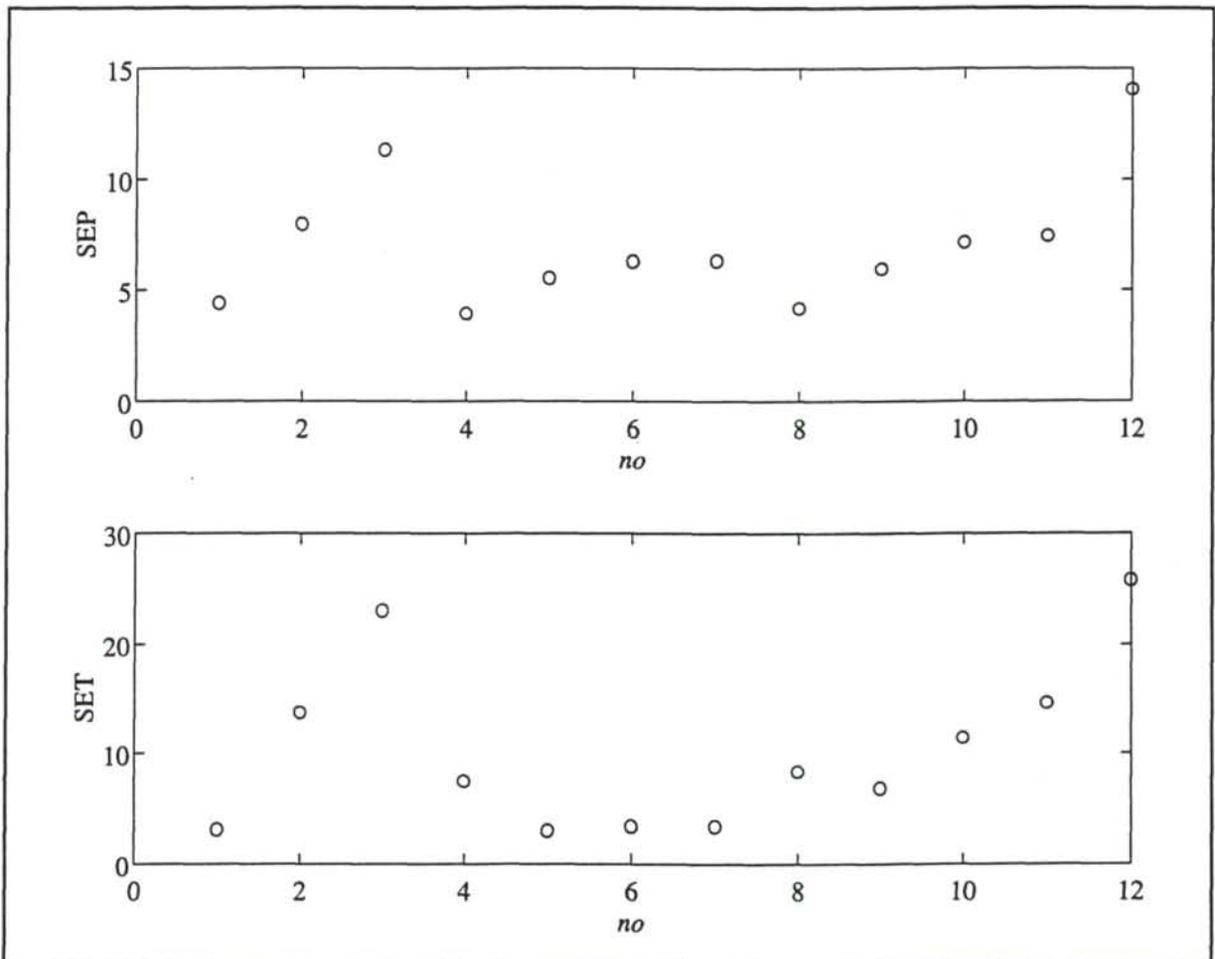


Figura III.4.16 Resultados de la transferencia de calibrado realizada con el procedimiento GCT para el plomo.

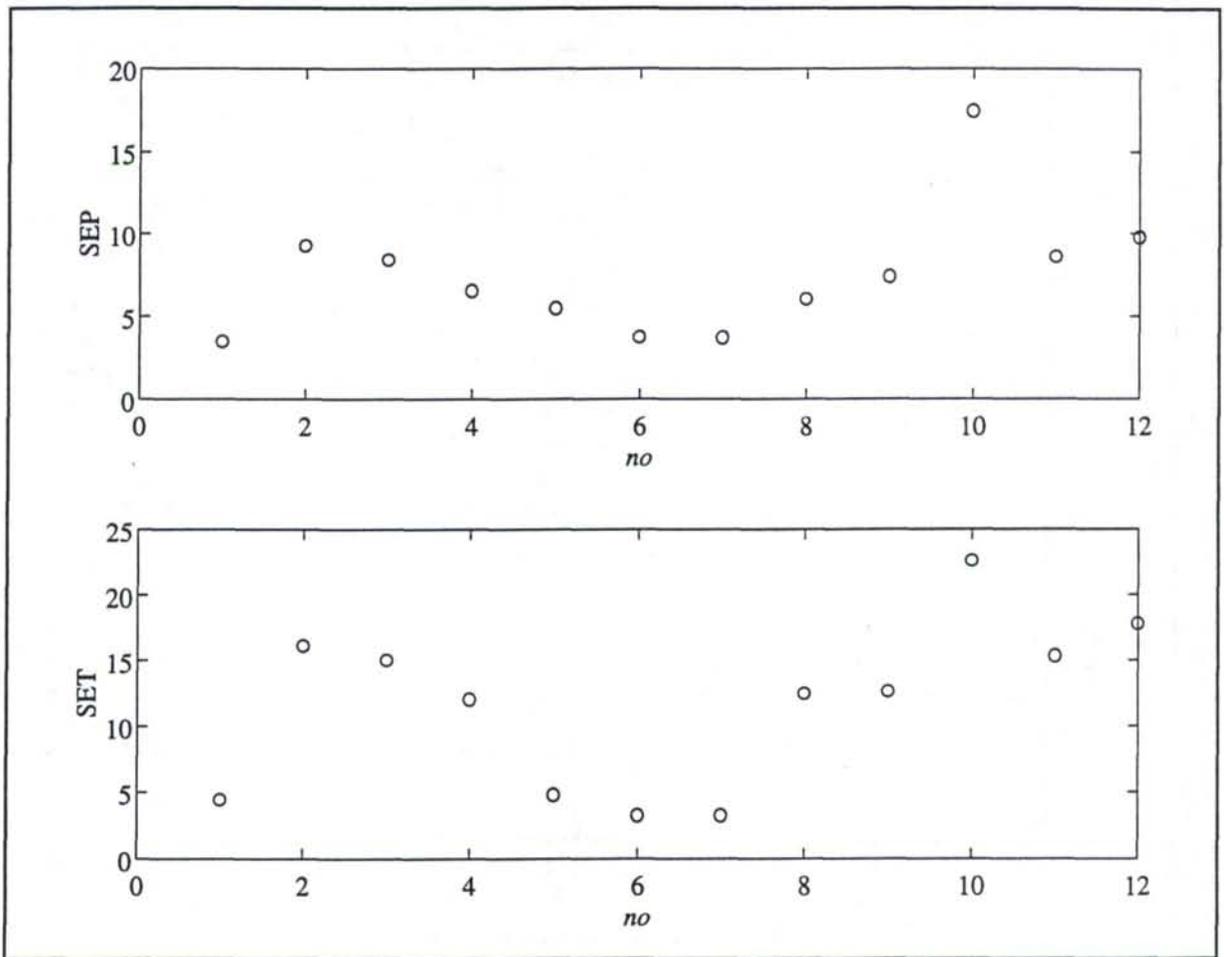


Figura III.4.17 Resultados de la transferencia de calibrado realizada con el procedimiento GCT para el cadmio.

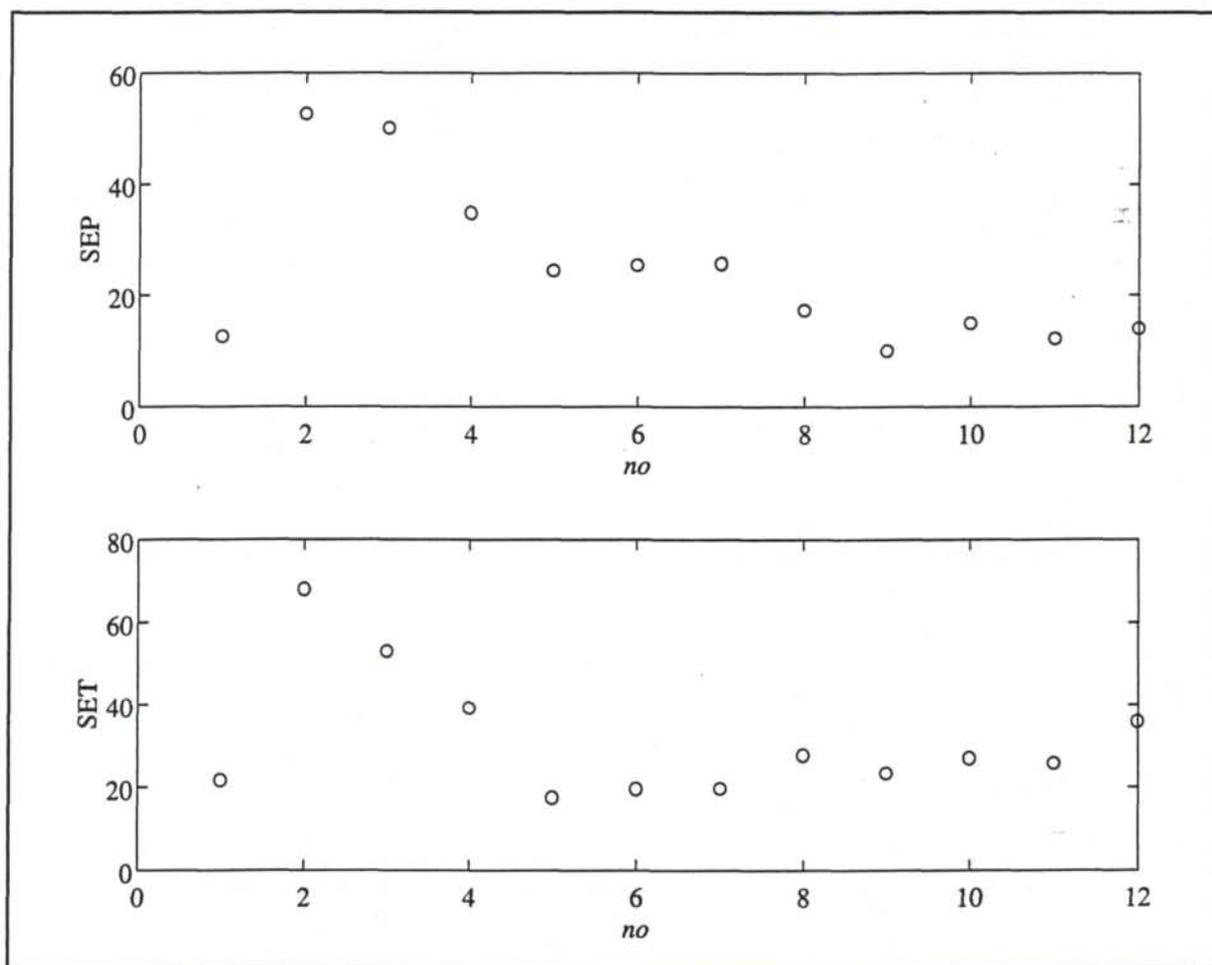


Figura III.4.18 Resultados de la transferencia de calibrado realizada con el procedimiento GCT para el zinc.

conjuntos. Mientras que el conjunto 9 proporciona buenos resultados en las transferencias de calibrado realizadas con el plomo y el zinc.

En la Tabla III.4.14 se muestran los errores correspondientes a las muestras del conjunto test cuando se ha realizado la transferencia de calibrado mediante el procedimiento GCT utilizando los conjuntos de transferencia 1, 5 y 9, dado que éstos parecen ser los subconjuntos de transferencia más idóneos. Al analizar dicha tabla, se confirma el hecho de que $n_0 = 9$ conduce a los errores menores en la transferencia llevada a cabo para el zinc, proporcionando unos resultados excelentes para este analito. Sin embargo, en lo que se refiere a cadmio y plomo, para los cuales $n_0 = 5$ debía proporcionar los mejores resultados después de $n_0 = 1$, se observa que algunos de los errores asociados con la transferencia de calibrado son especialmente elevados. Para el cobre, los resultados proporcionados por ambos conjuntos de transferencia también se encuentran asociados con errores importantes.

En cualquier caso, estos errores son mayores que los obtenidos cuando se aplicó el procedimiento de transferencia PDS para llevar a cabo la transferencia de calibrado entre estos dos conjuntos de datos, Tabla III.4.12, y además implican un número mayor de muestras patrón en la situación satélite. En este caso no es posible reducir el número de muestras necesarias para llevar a cabo la transferencia de calibrado con el conjunto de transferencia 5, que está constituido por un total de 16 muestras.

Por tanto, al igual que en el análisis llevado a cabo para estos cuatro analitos por polarografía, los resultados proporcionados a través del procedimiento de transferencia PDS, así como las condiciones experimentales que se requieren para realizar el mismo, señalan a éste como el procedimiento de transferencia de calibrado más adecuado para llevar a cabo posteriores determinaciones de cobre, plomo, cadmio y zinc por voltamperometría diferencial de impulsos con redisolución anódica.

Tabla III.4.14 Errores relativos (en %) obtenidos para las muestras del conjunto test en la transferencia de calibrado llevada a cabo mediante el procedimiento GCT.

<i>no</i>	Número de muestra	Analito			
		Cobre	Plomo	Cadmio	Zinc
1	18	-4.52	-9.43	7.69	-2.32
	20	-6.13	-4.50	2.77	-0.16
	23	7.16	3.24	9.62	4.85
	25	11.79	8.79	11.26	2.76
	28	-3.01	-2.29	0.88	0.21
	30	-4.33	-1.11	-2.24	1.76
	33	8.88	-1.75	1.05	-2.39
	35	9.70	3.90	7.84	2.84
5	18	-15.40	-14.88	4.52	-4.88
	20	-6.46	-3.43	7.84	4.37
	23	6.65	0.37	9.33	6.95
	25	16.84	9.82	23.08	7.30
	28	-4.67	-3.48	5.94	0.98
	30	-2.16	-0.29	0.37	5.48
	33	10.50	-2.11	4.77	1.08
	35	8.05	3.77	12.57	4.60
9	18	-11.28	-17.11	17.43	-3.49
	20	-11.51	-5.97	-13.83	-0.34
	23	-1.02	-0.90	-5.64	3.45
	25	5.96	4.46	0.69	2.85
	28	-6.92	-8.96	29.48	0.44
	30	-12.90	-6.08	-19.14	0.28
	33	5.37	-1.81	-6.46	-0.66
	35	-0.14	-4.20	-3.32	1.28

III.5 SOLAPAMIENTO DE SEÑALES EN MEDIDAS POLAROGRÁFICAS

La determinación simultánea de varios analitos mediante el uso de técnicas electroanalíticas puede complicarse en aquellos casos en que los potenciales de electrodo a los cuales se produce la oxidación o reducción de las especies electroactivas se encuentran muy próximos, tal y como ocurre en el análisis simultáneo de plomo, estaño y cadmio. En la mayoría de los casos se tiene como resultado un solapamiento de las respuestas experimentales que podría dar lugar a un error analítico cuando se calculan las concentraciones correspondientes a dichos analitos en una disolución.

El solapamiento de las señales polarográficas del estaño y el plomo, así como la marcada tendencia del estaño (IV) a hidrolizarse y polimerizar [152], son serios obstáculos en la determinación simultánea de estos elementos por polarografía diferencial de impulsos. Cuando la determinación analítica se lleva a cabo en medio ácido se evita la hidrólisis del estaño [153, 154]. Se suele utilizar ácido clorhídrico como electrólito soporte, ya que en este medio el estaño se encuentra estabilizado en forma de cloroComplejo. Sin embargo, en estas condiciones experimentales la señal polarográfica del estaño aparece completamente solapada con la correspondiente al plomo y muy próxima a su vez a la señal polarográfica del cadmio.

Para resolver este problema han sido propuestos muchos procedimientos, entre los cuales se incluye la separación completa del estaño mediante extracción o destilación [152, 155], la determinación simultánea de algunos de estos analitos a través de un cambio de medio o de matriz [156-158], la complejación selectiva de cada metal [159-161], la adición de agentes tensoactivos [162] o electrólitos adecuados [163-166], el uso de técnicas adsorptivas con redisolución [167, 168], el uso de electrodos rotatorios [169], etc.

También se han puesto a punto procedimientos que no requieren de un tratamiento o una metodología experimental especial, para resolver problemas en los que aparecían señales solapadas. Inicialmente se desarrollaron los métodos de polarografía substractiva [121] y el propuesto por Bond y Grabaric [170]. La simulación matemática de las señales totales, con

la substracción posterior de las señales individuales simuladas, ha dado peores resultados al ser aplicada a técnicas voltamperométricas [171] que a otro tipo de técnicas, fundamentalmente espectroscópicas. Sin embargo, el desarrollo de las técnicas de polarografía derivada y voltamperometría derivada [121] sí que ha permitido afrontar algunos problemas de este tipo, tales como la resolución de picos solapados correspondientes a mezclas de Cd(II)/In(III) y Tl(I)/Pb(II) respectivamente, mediante polarografía diferencial de impulsos [172].

La voltamperometría de onda cuadrada para casos reversibles [173] permite evaluar mejor la resolución del solapamiento de señales a través de la derivatización de los datos obtenidos con esta técnica electroquímica. No obstante, se ha visto que para aquellos casos en los que el grado de solapamiento es muy elevado, ni siquiera la inspección de la segunda derivada conduce a buenos resultados [174]. Otras técnicas, tales como la voltamperometría de corriente alterna, han sido aplicadas con éxito en la resolución de algunas señales solapadas [175, 176].

También las técnicas semidiferenciales o de deconvolución son de gran interés en problemas con señales solapadas [177, 178]. El uso de la transformada de Fourier en la deconvolución de polarogramas ha resultado una metodología útil. Brown y Brown [179] han demostrado la aplicabilidad de un análisis multicomponente para la resolución de mezclas de Cd(II)/In(III)/Pb(II) mediante polarografía diferencial de impulsos utilizando filtros Kalman. Esta metodología también ha sido aplicada [180] en la determinación de mezclas de Tl(I)/Pb(II), mejorando los resultados obtenidos por Binkley y Dessy [181].

Las técnicas de regresión multivariante sobre variables latentes también se muestran útiles en el desolapamiento de señales, debido a la capacidad que tienen para modelar señales complejas y filtrar el ruido de fondo. Se ha puesto de manifiesto la posibilidad de utilizar una regresión multilineal para desolapar picos procedentes de análisis realizados mediante polarografía diferencial de impulsos [182] y voltamperometría diferencial de impulsos [183]. Ni y otros [184] han comparado los resultados obtenidos con diferentes modelos de regresión multivariante (MLR, PCR y PLS) en la determinación polarográfica simultánea de pirazina y sus derivados metilados, concluyendo que los modelos PLS y PCR conducen a los resultados más satisfactorios. En particular, el método PLS [15] ha sido aplicado ampliamente en problemas de calibración multivariante y/o multicomponente [9]. Entre las aplicaciones de este método destaca su uso en el desolapamiento de señales procedentes tanto de sistemas inorgánicos [33, 148], como de mezclas binarias y ternarias de compuestos orgánicos [47-49].

En este caso, la falta de especificidad en la señal del estaño puede abordarse también desde este punto de vista multivariante, utilizando la información contenida en todo el polarograma, para determinar la concentración de este elemento y, de forma simultánea, la de plomo y cadmio. En este capítulo se ha llevado a cabo la determinación simultánea de estos tres analitos en muestras acuosas por polarografía diferencial de impulsos mediante un calibrado multivariante PLS y, una vez finalizada esta etapa del análisis, se ha realizado una transferencia de calibrado a lo largo del tiempo para intentar reducir considerablemente el esfuerzo experimental requerido en calibrados posteriores.

III.5.1 Experimental. Determinación de Pb, Sn y Cd mediante DPP

Reactivos

Todos los reactivos utilizados han sido de calidad para análisis y no se han realizado posteriores purificaciones. Las disoluciones patrón correspondientes a dichos reactivos fueron preparadas utilizando en todos los casos agua desionizada proporcionada por un sistema Barnstead NANO Pure II.

Como solución tampón se utilizó un medio suficientemente ácido para favorecer la estabilidad del estaño (IV), constituido por una disolución HCl 0.1 M y H₂C₂O₄ 0.1 M, que presentaba un pH = 0.177. El gas inerte utilizado para eliminar el oxígeno disuelto en las disoluciones polarográficas a analizar fue nitrógeno (99.997%).

Instrumental

Las medidas polarográficas han sido llevadas a cabo utilizando un procesador Metrohm 646 VA acoplado a un 647 VA equipado con un electrodo multimodo Metrohm (MME) que operó en el modo de electrodo de gota de mercurio estática (SMDE). El sistema de tres electrodos se completó con un electrodo de referencia Ag-AgCl/KCl 3 M y un electrodo de platino como contraelectrodo. Las adiciones de disoluciones patrón fueron realizadas con una bureta automática Metrohm 665 Dosimat. Las correspondientes medidas de pH se obtuvieron con un Crison 2002 (± 0.001).

Los análisis de datos se han efectuado utilizando los paquetes PARVUS [57], STATGRAPHICS [58] y MATLAB [59].

Procedimiento experimental

Para llevar a cabo los análisis se introdujeron 25 ml. de solución tampón en una celda polarográfica, eliminando el oxígeno disuelto en la misma burbujeando nitrógeno durante 10 minutos. Esta disolución se analizó mediante polarografía diferencial de impulsos (DPP), registrándose la intensidad de corriente frente a 66 valores de potencial igualmente espaciados entre -266 mV y -656 mV. Sobre la misma se realizaron diversas adiciones de las disoluciones patrón, y tras cada una de ellas, la solución polarográfica fue desoxigenada de nuevo durante 15 s, al tiempo que era agitada.

Los parámetros experimentales seleccionados para realizar las medidas fueron los siguientes: amplitud del impulso -50 mV, área del electrodo 0.40 mm², tiempo de vida de la gota 0.6 s y velocidad de barrido -10 mV s⁻¹.

III.5.2 Diseño del experimento

Como se ha indicado anteriormente, la forma en que se lleva a cabo la experimentación en un análisis está directamente relacionada con la información y los resultados que será posible obtener a partir del mismo y con el esfuerzo experimental requerido para ello. Estos aspectos se tienen en consideración cuando se utiliza un experimento diseñado.

El diseño experimental seleccionado para llevar a cabo la determinación simultánea de plomo, estaño y cadmio se encuentra próximo a un diseño central compuesto, como el de la Figura II.3.1, correspondiéndose éste con los valores de concentración de los distintos analitos que se encuentran en la Tabla III.5.1. De este diseño forman parte las muestras 10, 12, 15, 17, 20 y 22, que se han incluido en el mismo como muestras test debido a que su obtención supone un mínimo esfuerzo experimental adicional. Este diseño permite disponer de cinco niveles de concentración para cada uno de los tres analitos que se desea determinar y, como ya se ha señalado, es prácticamente ortogonal.

de cinco niveles de concentración para cada uno de los tres analitos que se desea determinar y, como ya se ha señalado, es prácticamente ortogonal.

Tabla III.5.1 Diseño experimental. Cada símbolo de la segunda columna indica que se realizaron adiciones sobre una misma disolución inicial.

Muestra	Diseño experimental	Concentraciones (μM)					
		Plomo	Estaño	Cadmio			
1	○	2	2	2	3.9063	6.6758	6.1242
2	○	2	2	4	3.8760	6.6240	12.1535
3	○	2	4	4	3.8462	13.1462	12.0600
4	○	4	4	4	7.6336	13.0458	11.9679
5	×	2	4	2	3.8760	13.2481	6.0767
6	×	4	4	2	7.6923	13.1462	6.0300
7	·	4	2	2	7.7519	6.6240	6.0767
8	·	4	2	4	7.6923	6.5731	12.0600
9	▲	1	3	3	1.9455	9.9747	9.1506
10	▲	2	3	3	3.8760	9.9360	9.1151
11	▲	3	3	3	5.7915	9.8977	9.0799
12	▲	4	3	3	7.6923	9.8596	9.0450
13	▲	5	3	3	9.5785	9.8218	9.0103
14	□	3	1	3	5.8366	3.3249	9.1506
15	□	3	2	3	5.8140	6.6240	9.1151
16	□	3	3	3	5.7915	9.8977	9.0799
17	□	3	4	3	5.7692	13.1462	9.0450
18	□	3	5	3	5.7471	16.3697	9.0103
19	·	3	3	1	5.8366	9.9747	3.0502
20	·	3	3	2	5.8140	9.9360	6.0767
21	·	3	3	3	5.7915	9.8977	9.0799
22	·	3	3	4	5.7692	9.8596	12.0600
23	·	3	3	5	5.7471	9.8218	15.0172

Según este diseño se han llevado a cabo dos calibrados multivariantes en diferentes días, obteniéndose los polarogramas que aparecen en las Figuras III.5.1a-f y III.5.2a-f, los cuales se van a corresponder en adelante con los siguientes conjuntos de datos: calibrado E (realizado en primer lugar) y el calibrado G (calibrado obtenido con posterioridad). Cada una de las figuras consta de seis gráficas, que se corresponden con las diferentes disoluciones

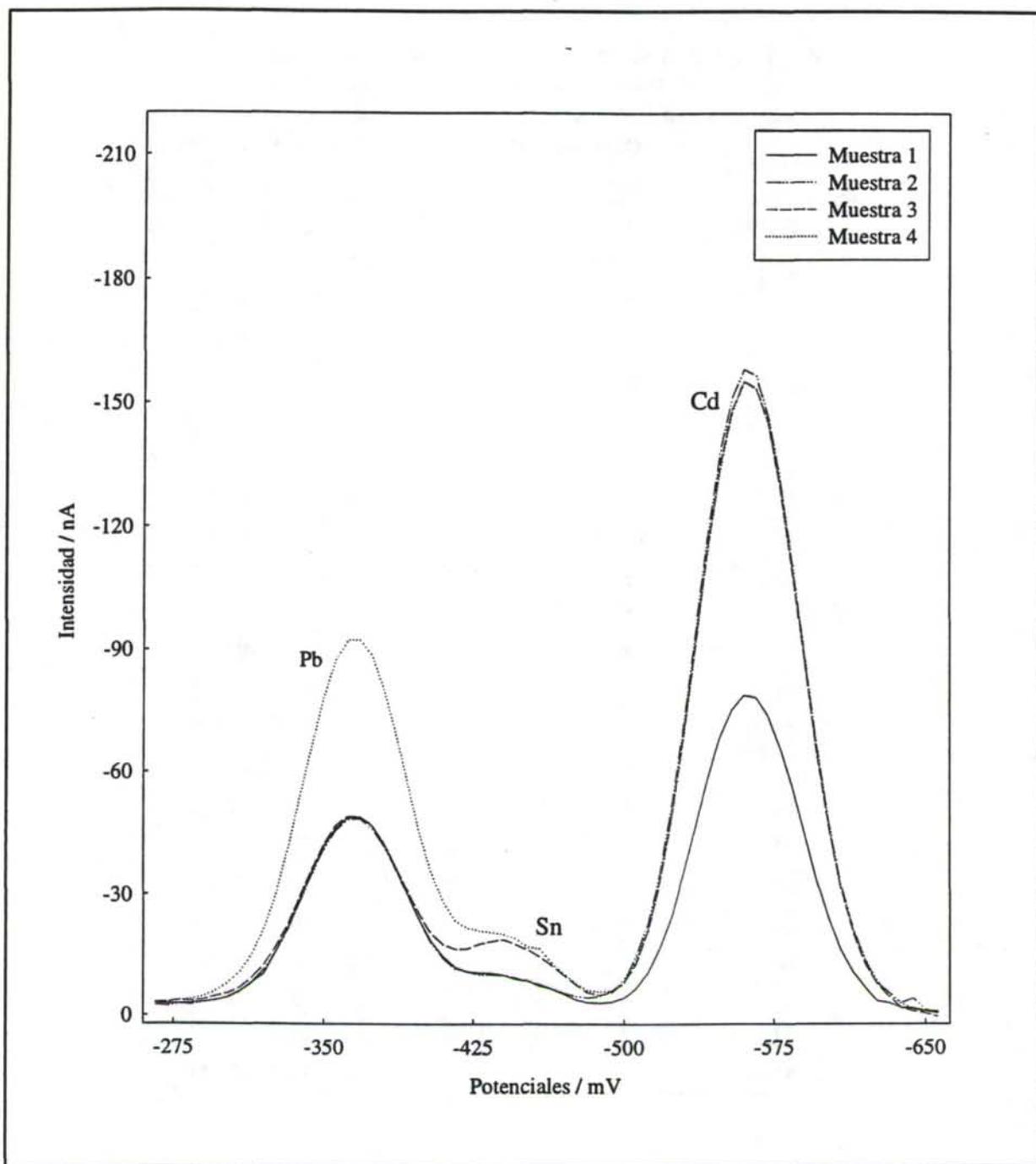


Figura III.5.1a Polarogramas correspondientes al calibrado E, obtenidos con una misma disolución inicial.

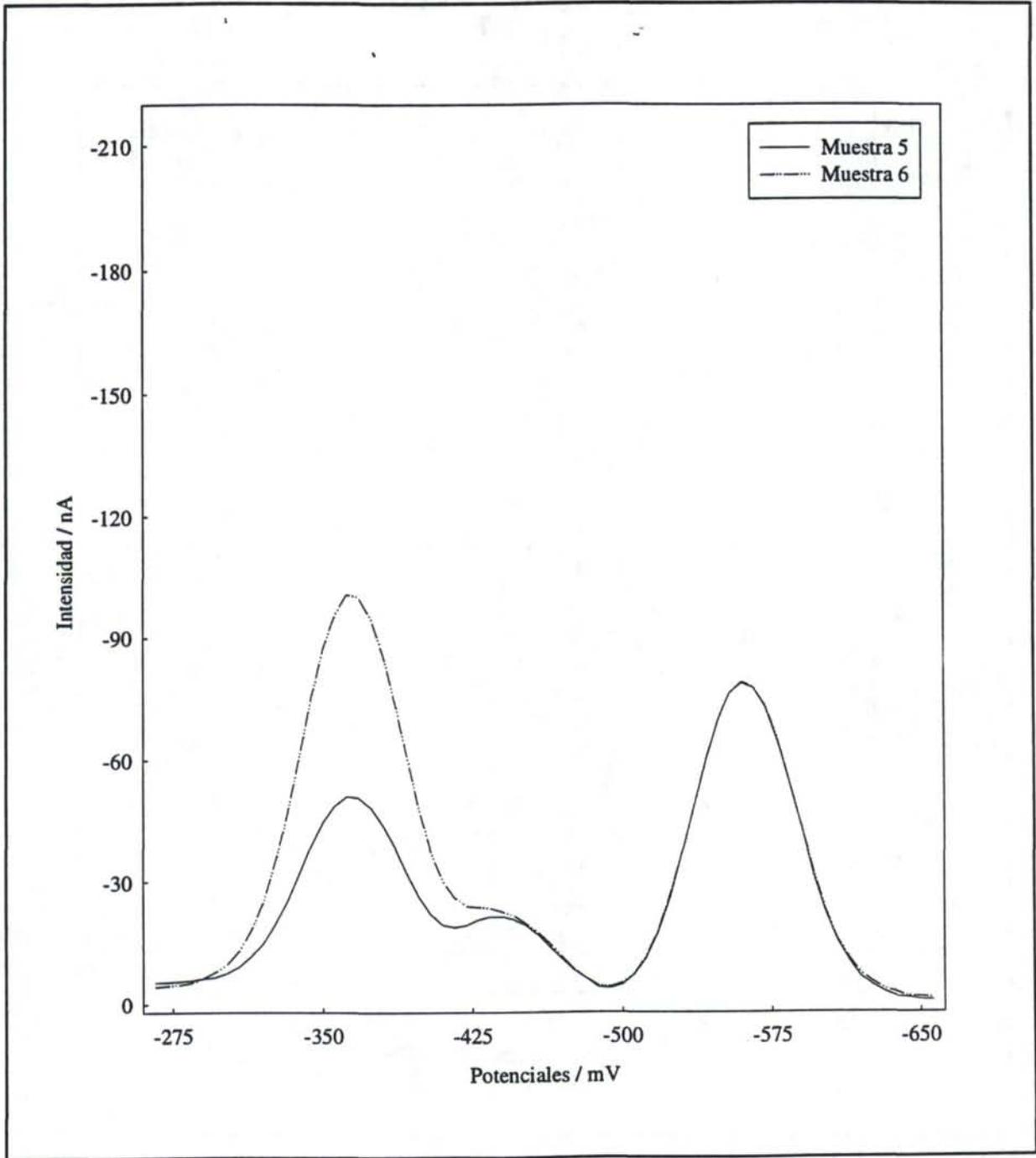


Figura III.5.1b Polarogramas correspondientes al calibrado E, obtenidos con una misma disolución inicial.

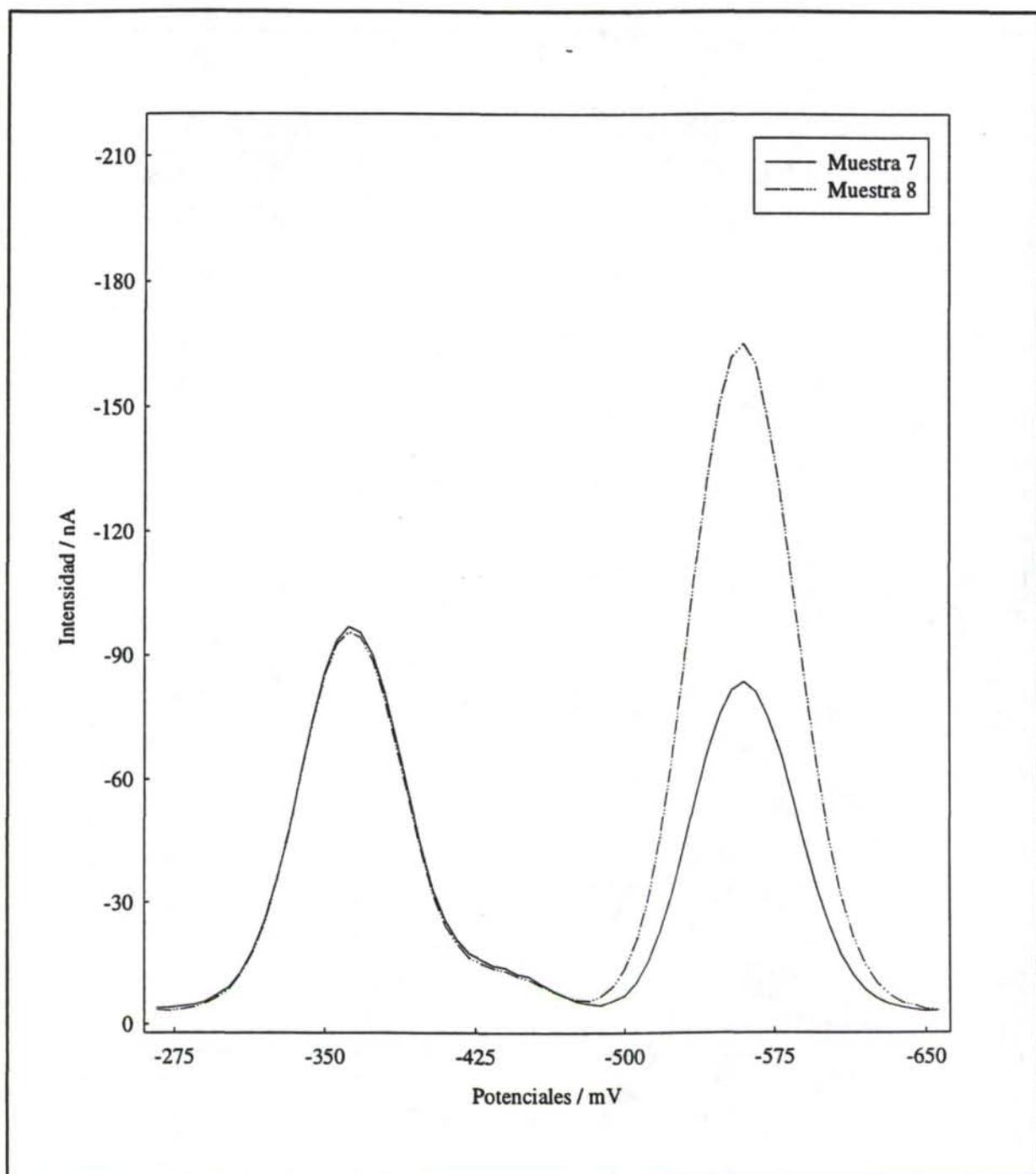


Figura III.5.1c Polarogramas correspondientes al calibrado E, obtenidos con una misma disolución inicial.

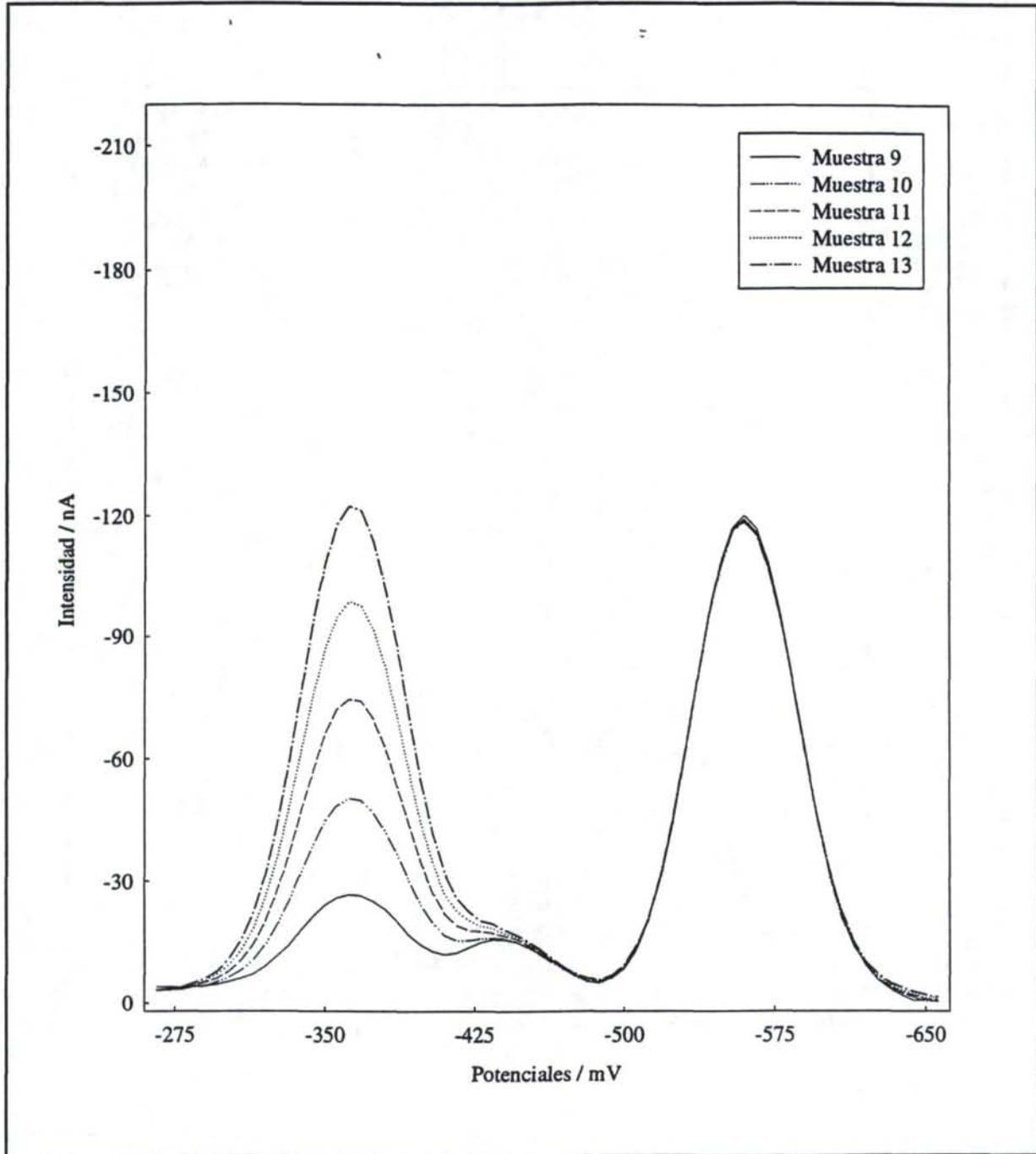


Figura III.5.1d Polarogramas correspondientes al calibrado E, obtenidos con una misma disolución inicial.

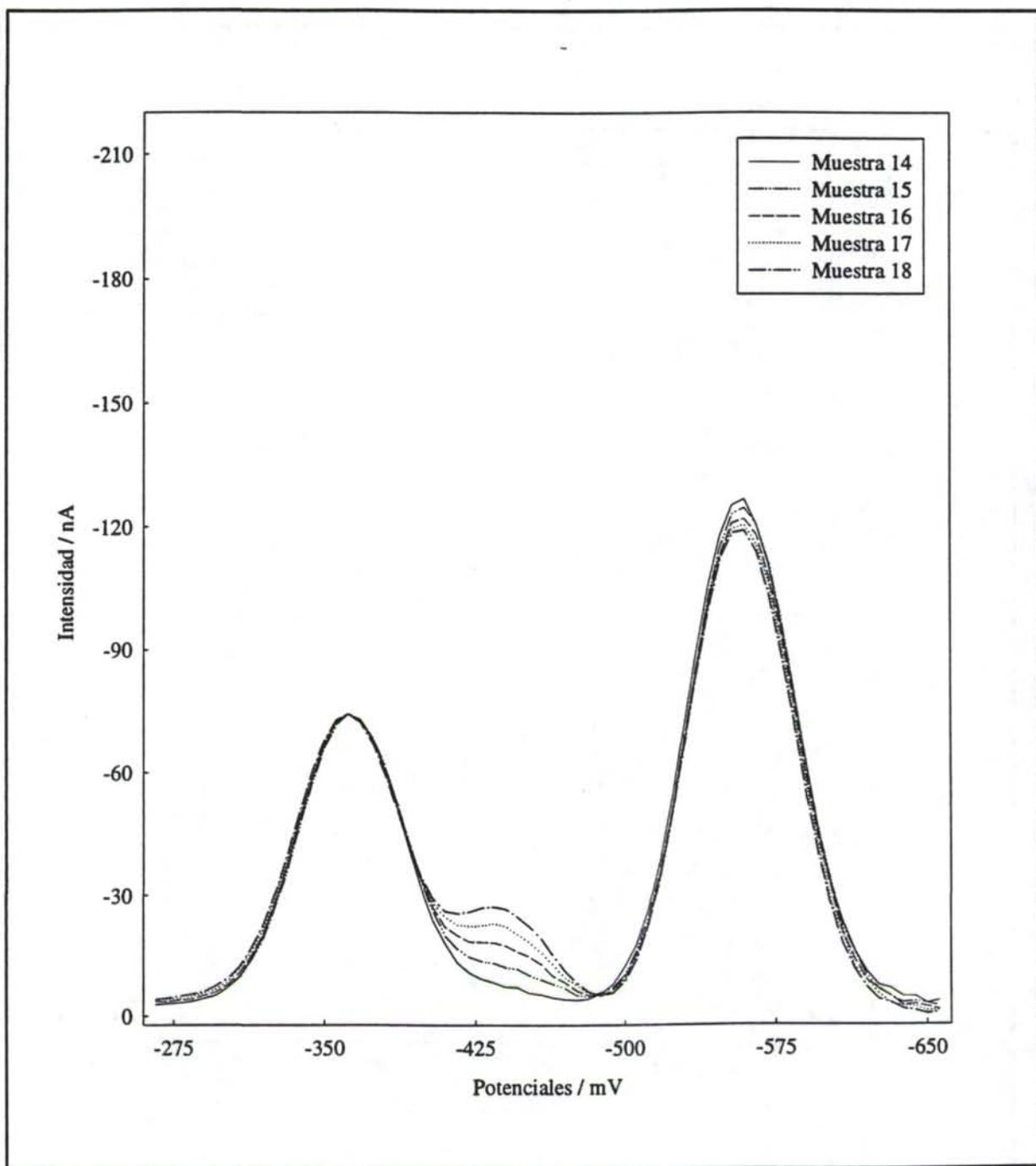


Figura III.5.1e Polarogramas correspondientes al calibrado E, obtenidos con una misma disolución inicial.

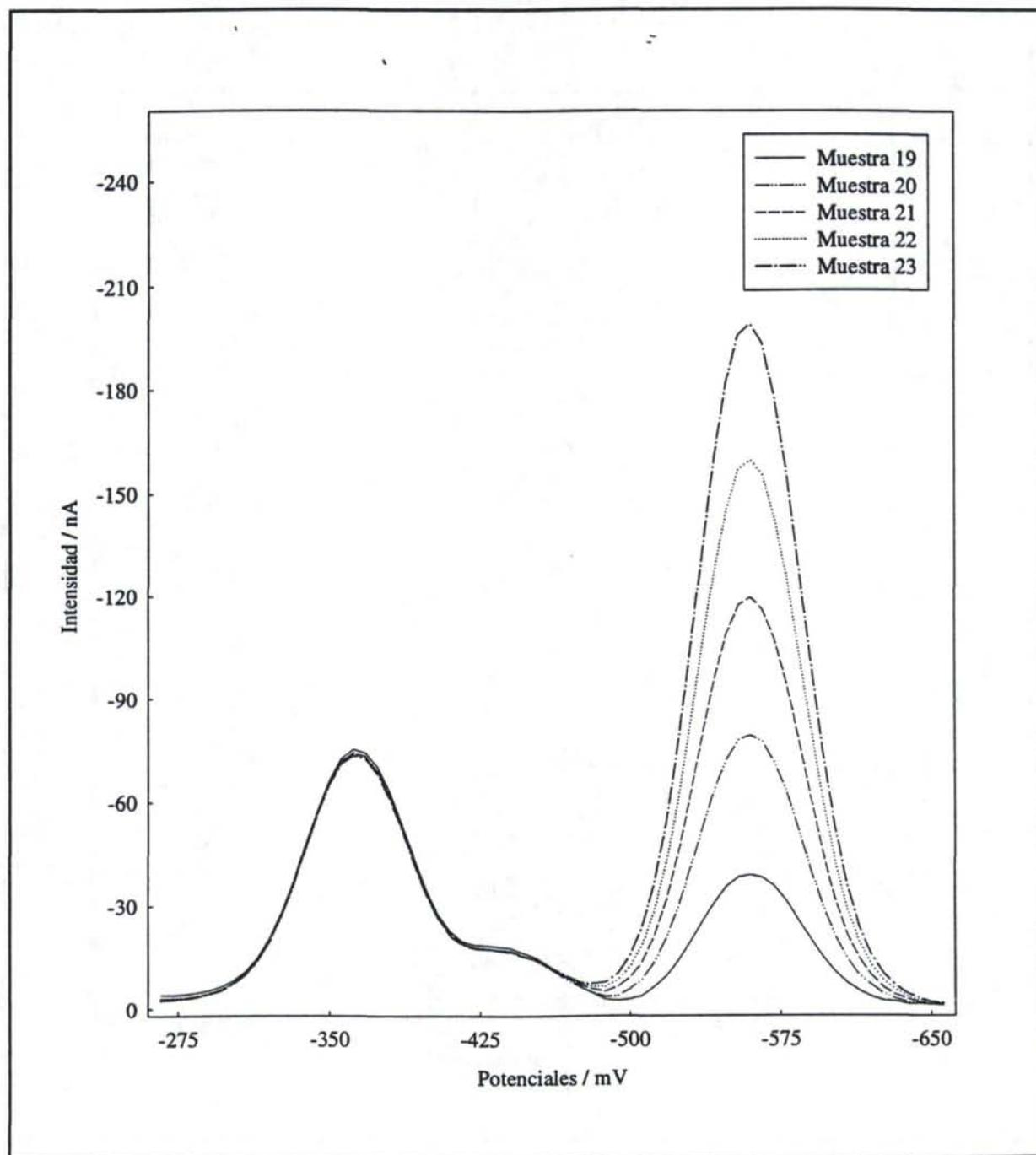


Figura III.5.1f Polarogramas correspondientes al calibrado E, obtenidos con una misma disolución inicial.

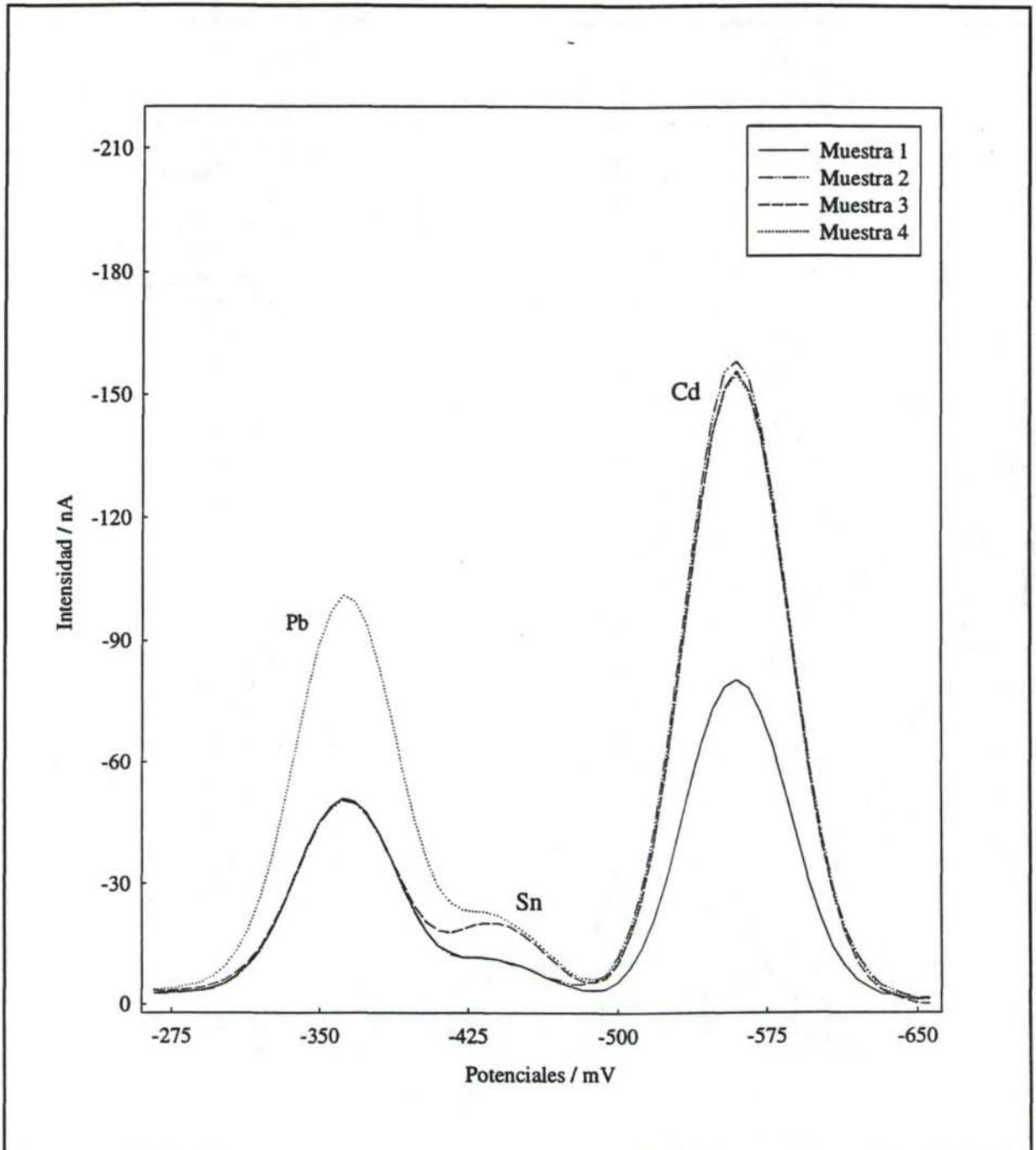


Figura III.5.2a Polarogramas correspondientes al calibrado G, obtenidos con una misma disolución inicial.

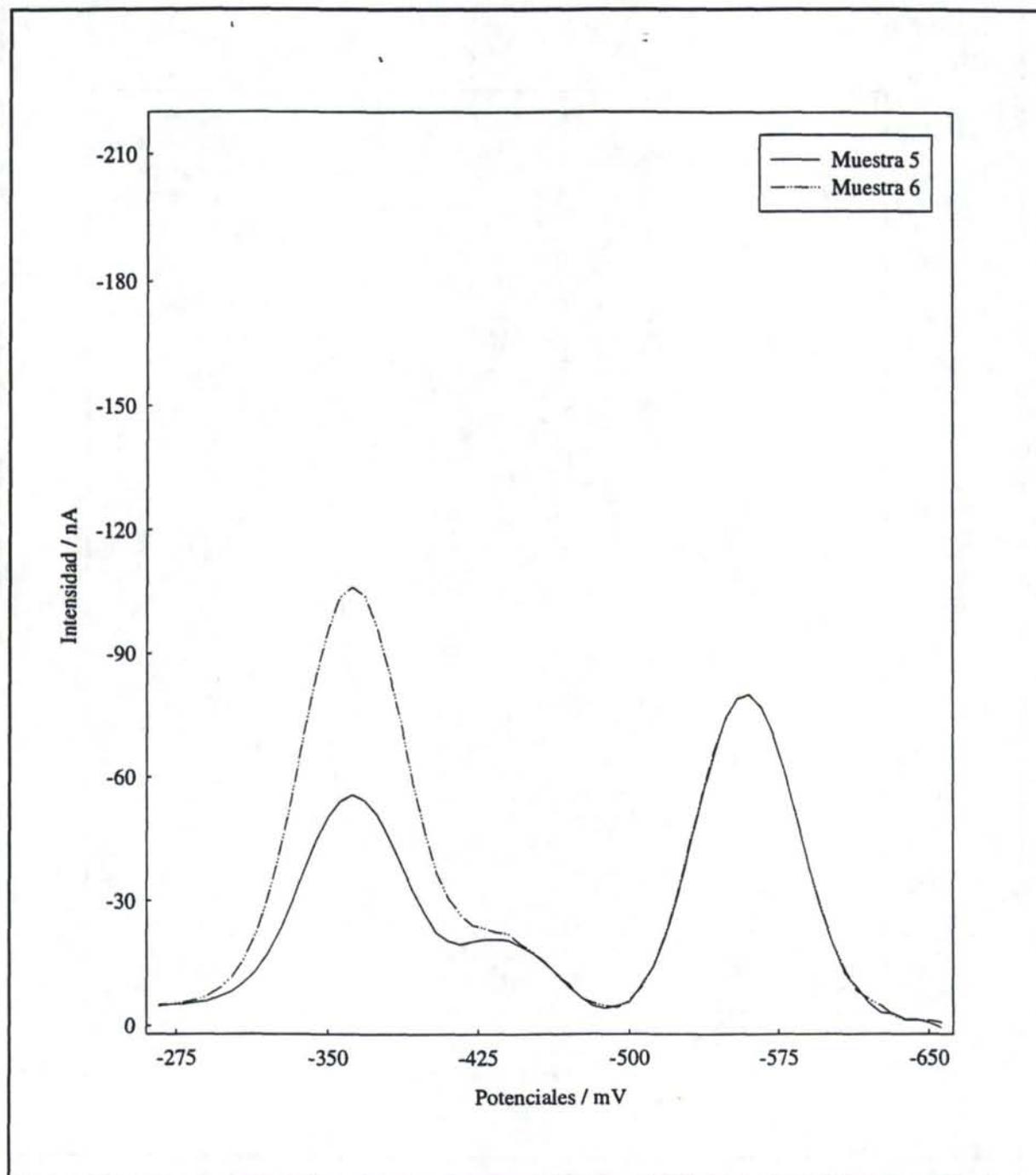


Figura III.5.2b Polarogramas correspondientes al calibrado G, obtenidos con una misma disolución inicial.

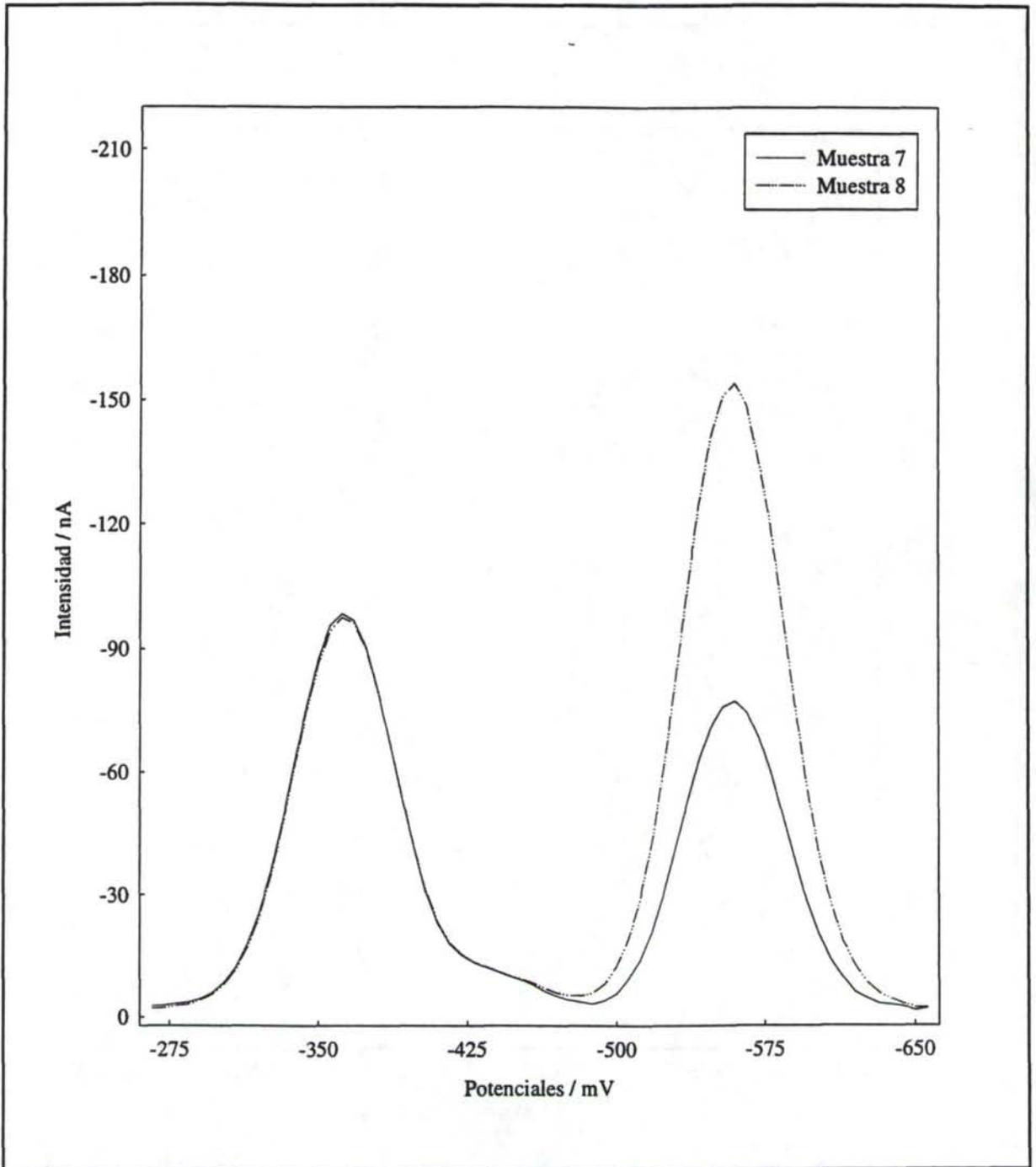


Figura III.5.2c Polarogramas correspondientes al calibrado G, obtenidos con una misma disolución inicial.

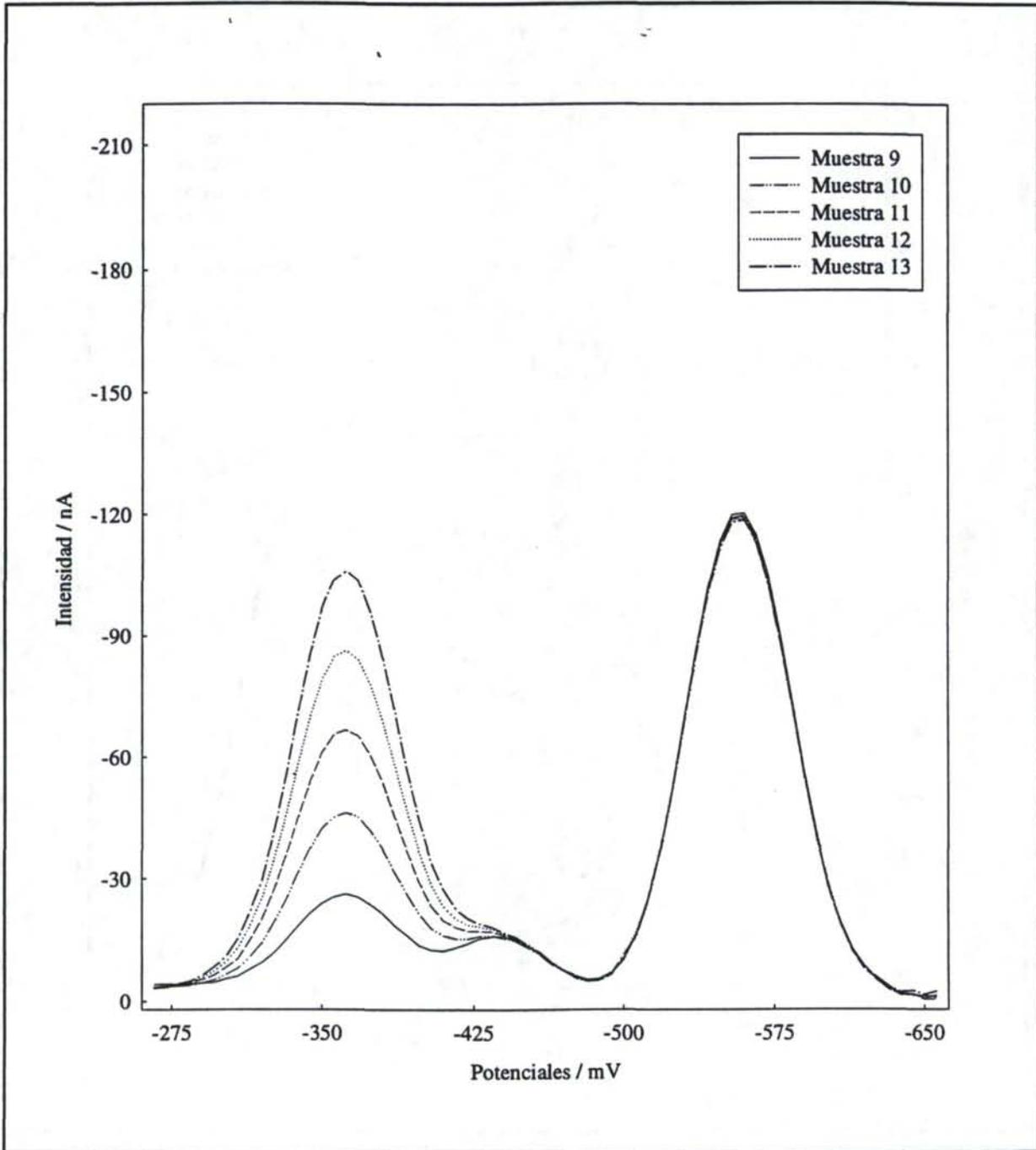


Figura III.5.2d Polarogramas correspondientes al calibrado G, obtenidos con una misma disolución inicial.

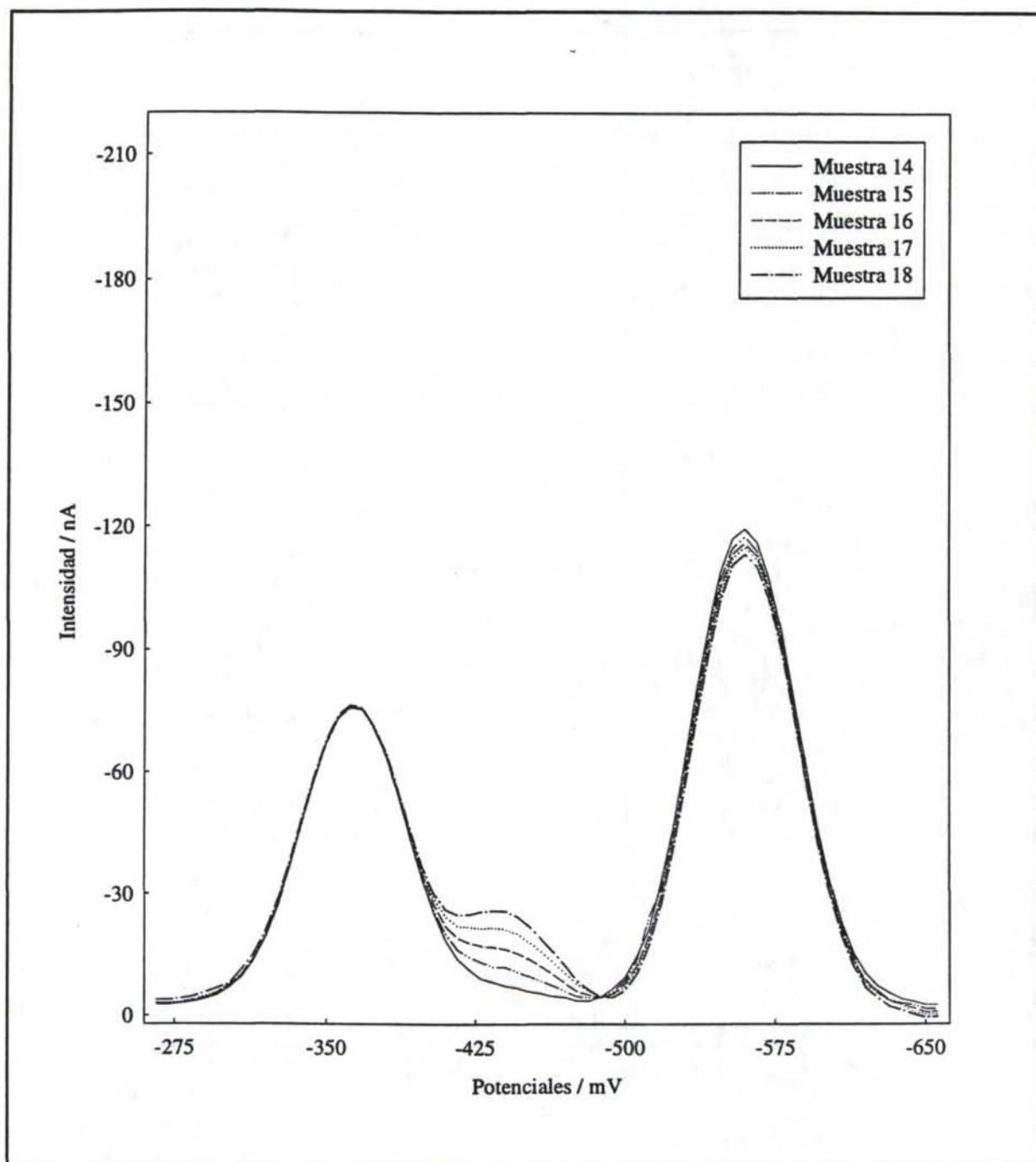


Figura III.5.2e Polarogramas correspondientes al calibrado G, obtenidos con una misma disolución inicial.

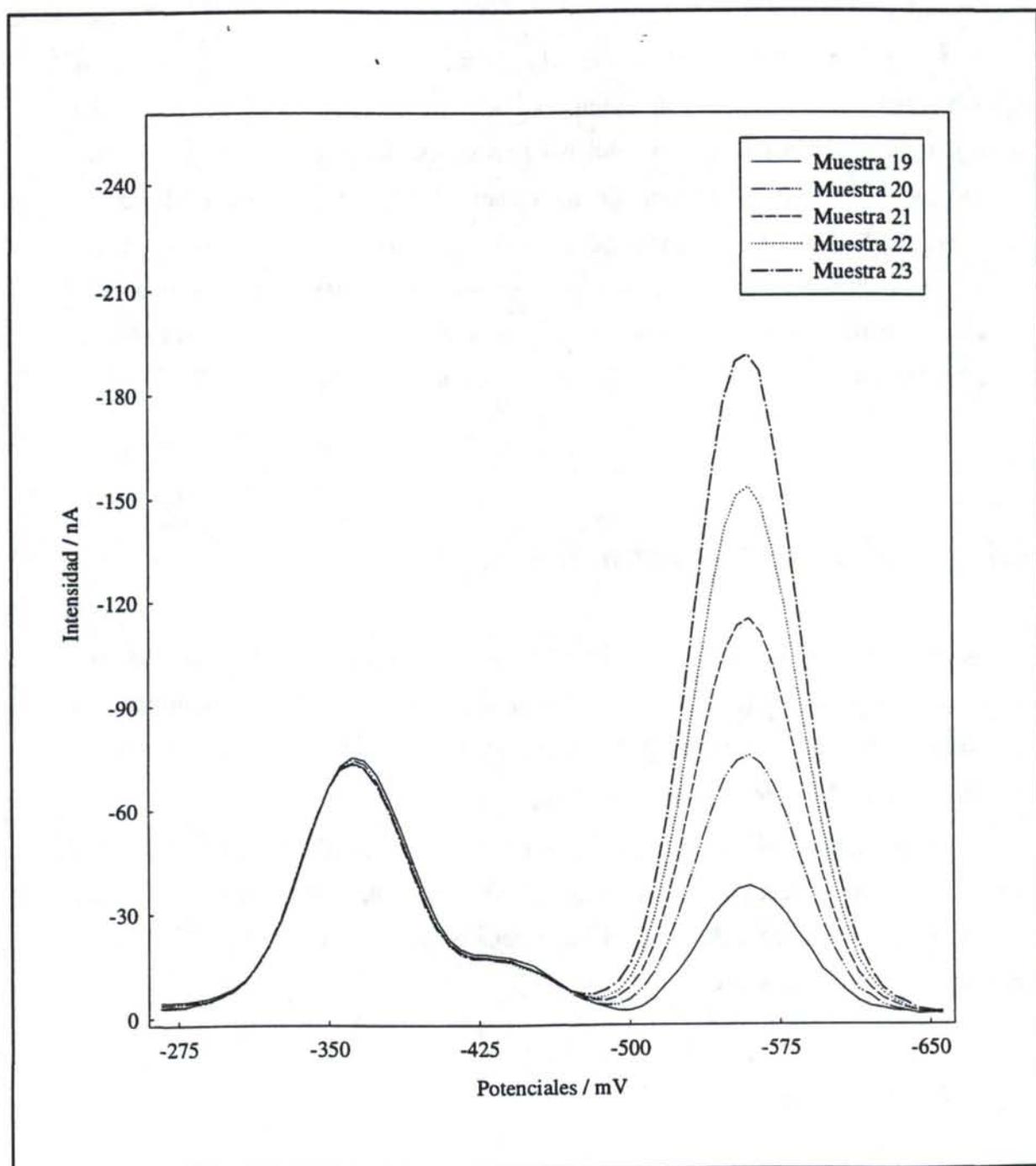


Figura III.5.2f Polarogramas correspondientes al calibrado G, obtenidos con una misma disolución inicial.

En cada uno de los polarogramas registrados aparecen tres picos polarográficos correspondientes respectivamente a plomo, estaño y cadmio. Se observa que los picos polarográficos del plomo y el estaño aparecen solapados, así como parcialmente los del estaño y el cadmio.

En estas circunstancias sería posible estudiar desde un punto de vista univariante las señales correspondientes al plomo y al cadmio, ya que sus señales polarográficas parecen bastante bien definidas. Sin embargo, un análisis univariante del pico del estaño resultaría inviable, ya que no existe un máximo de pico claro, sobre todo si se consideran las concentraciones más bajas. Debido a esta falta de especificidad del estaño en presencia de plomo y cadmio, se puede considerar la posibilidad de llevar a cabo un análisis multivariante que permita, utilizando toda la información contenida en el polarograma, determinar la concentración de estaño y de forma simultánea la de plomo y cadmio.

III.5.3 Análisis multivariante

Dada la capacidad para modelar la información subyacente a una señal compleja mostrada por los modelos de regresión PLS [33], se ha optado por utilizar esta metodología multivariante para resolver el problema originado por el solapamiento de las señales polarográficas correspondientes a plomo, estaño y cadmio.

Una vez construídos los modelos de calibrado PLS correspondientes, se ha llevado a cabo una transferencia de calibrado a lo largo del tiempo, utilizándose para ello los dos procedimientos de transferencia de calibrado aplicados anteriormente, PDS y GCT, cuyos resultados han sido analizados detalladamente.

III.5.3.1 Modelos PLS

Construcción de los modelos PLS

Con la finalidad de llevar a cabo una calibración multivariante que permita la determinación simultánea de plomo, estaño y cadmio, se han construído sendos calibrados con los conjuntos de datos E y G, utilizando en ambos casos una regresión multivariante PLS.

Ambos conjuntos de datos constan de 23 objetos (las muestras del diseño experimental) que se encuentran definidos por 66 variables predictoras (las intensidades registradas a 66 valores de potencial igualmente espaciados). Cada uno de esos objetos se encuentra a su vez asociado con ciertos valores de concentración para cada uno de los analitos, constituyendo éstos la variable respuesta. Por tanto, las dos matrices de datos **E** y **G** tienen dimensiones 23×66 , donde 23 es el número de muestras y 66 los potenciales a los que se ha registrado la intensidad, una vez que se ha restado el blanco experimental correspondiente a cada una de las muestras, según indica la Tabla III.5.1.

En ambos conjuntos de datos han sido seleccionados algunos objetos que se han asignado al conjunto test, son los objetos 10, 12, 15, 17, 20 y 22 de la Tabla III.5.1, mientras que el resto constituyen el conjunto de aprendizaje (training set). Como ya se ha indicado, el modelo de regresión se construye utilizando los objetos del conjunto de aprendizaje, mientras que los objetos del conjunto test no intervienen en esta etapa del análisis. Éstos se utilizan posteriormente para evaluar la capacidad real de predicción del modelo construido.

Los datos del análisis no han sido sometidos a ninguna transformación previa, y se ha utilizado un método de validación cruzada [126] para evaluar la varianza en la predicción de los modelos construidos. Las regresiones multivariantes se han realizado de forma independiente para cada uno de los analitos, es decir, se han construido tres modelos PLS para el calibrado **E** y otros tres modelos diferentes para el calibrado **G**.

Evaluación de los modelos PLS

El análisis de los modelos de regresión PLS construidos para cada analito con el conjunto de datos **E** determina los porcentajes de varianza explicada que se especifican en la Tabla III.5.2. Los valores de varianza explicada alcanzados son muy buenos para un calibrado de este tipo, superan el 99% en todos los casos, lo cual indica que el ajuste llevado a cabo por los modelos construidos es adecuado. Además, la proximidad entre la varianza explicada y la varianza en predicción da idea de la estabilidad de dichos modelos, así como de su capacidad para predecir muestras ajenas a los mismos.

Cabe destacar el hecho de que, a pesar de la falta de especificidad de la señal del estaño, los valores de varianza explicada y en predicción obtenidos con su modelo son del mismo orden que los correspondientes a los modelos construidos para plomo y cadmio, cuyas señales aparecían perfectamente bien definidas. Se pone de manifiesto la capacidad de la

Tabla III.5.2 Porcentajes de varianza explicada obtenidos por los modelos PLS para el calibrado E.

Analito	V.L.	Bloque Y		Bloque X
		Varianza explicada	Varianza en predicción	Varianza explicada
Plomo	1	4.89	3.37	79.34
	2	98.67	98.62	94.55
	3	99.64	99.09	95.99
	4	99.65	99.21	99.03
	5	99.67	99.12	99.30
	6	99.85	98.93	99.34
	7	99.88	98.93	99.49
	8	99.94	98.91	99.50
Estaño	1	-	-	79.37
	2	-	-	94.07
	3	96.56	77.58	96.21
	4	98.60	97.97	99.03
	5	99.43	98.53	99.31
	6	99.61	98.50	99.45
	7	99.91	98.51	99.50
	8	99.98	98.66	99.51
Cadmio	1	83.97	83.62	79.37
	2	99.43	99.46	94.55
	3	99.86	99.87	96.23
	4	99.87	99.90	98.89
	5	99.95	99.94	99.31
	6	99.97	99.95	99.45
	7	99.99	99.95	99.50
	8	100.00	99.94	99.53

regresión PLS para extraer la información útil, excluyendo aquellas fuentes de variabilidad no relacionadas con la concentración del analito de interés.

Las variables latentes que se necesitan en cada uno de los modelos de regresión construídos son aquellas cuyo aporte al modelo conlleva un aumento de la varianza en predicción. Este procedimiento asegura la inclusión de aquellas variables latentes relacionadas con la respuesta. Por lo tanto, el modelo de regresión PLS correspondiente al plomo se construirá con cuatro variables latentes, el del estaño con cinco y, por último, el del cadmio con seis.

Parece claro, en vista de los porcentajes de varianza que explican las primeras componentes, que la primera componente está relacionada principalmente con el cadmio, la segunda con el plomo y la tercera con el estaño, lo cual parece lógico al considerar que los datos no han sido autoescalados y que las direcciones de máxima variabilidad en el bloque de las predictoras van en ese mismo orden. No obstante, la representación de los loadings y los scores correspondientes a los modelos PLS construídos permite una interpretación más detallada de las variables latentes.

En las Figuras III.5.3 a III.5.5 se encuentran representados los loadings correspondientes a cada uno de los modelos PLS del calibrado E, muy similares a aquellos obtenidos para el calibrado G. En dichas figuras se observa que la estructura de los loadings de la primera variable latente es común para todas ellas y reproduce la forma de la señal polarográfica ya que los datos no han sido autoescalados, como se ha indicado anteriormente, es el denominado factor tamaño. Estos loadings presentan valores negativos debido a que las intensidades registradas también eran negativas, el proceso electroquímico seguido ha sido una reducción. Esta variable latente explica, por tanto, un elevado porcentaje de la varianza explicada correspondiente al bloque de las variables predictoras, estando relacionada además con la concentración de cadmio (analito cuya señal polarográfica es la de mayor magnitud).

La estructura de los loadings de la segunda variable latente es común para plomo y estaño, lo cual indica que esta variable latente está considerando simultáneamente ambas señales, de hecho, los loadings asociados con los potenciales correspondientes a las señales de esos dos analitos tienen el mismo signo (están considerando su aportación a la señal). Es la tercera variable latente en ambos modelos la que compensa la participación del "otro analito" a la señal global, teniendo los loadings relacionados con este último signo contrario al correspondiente a la segunda variable latente. Así, los loadings correspondientes al pico del

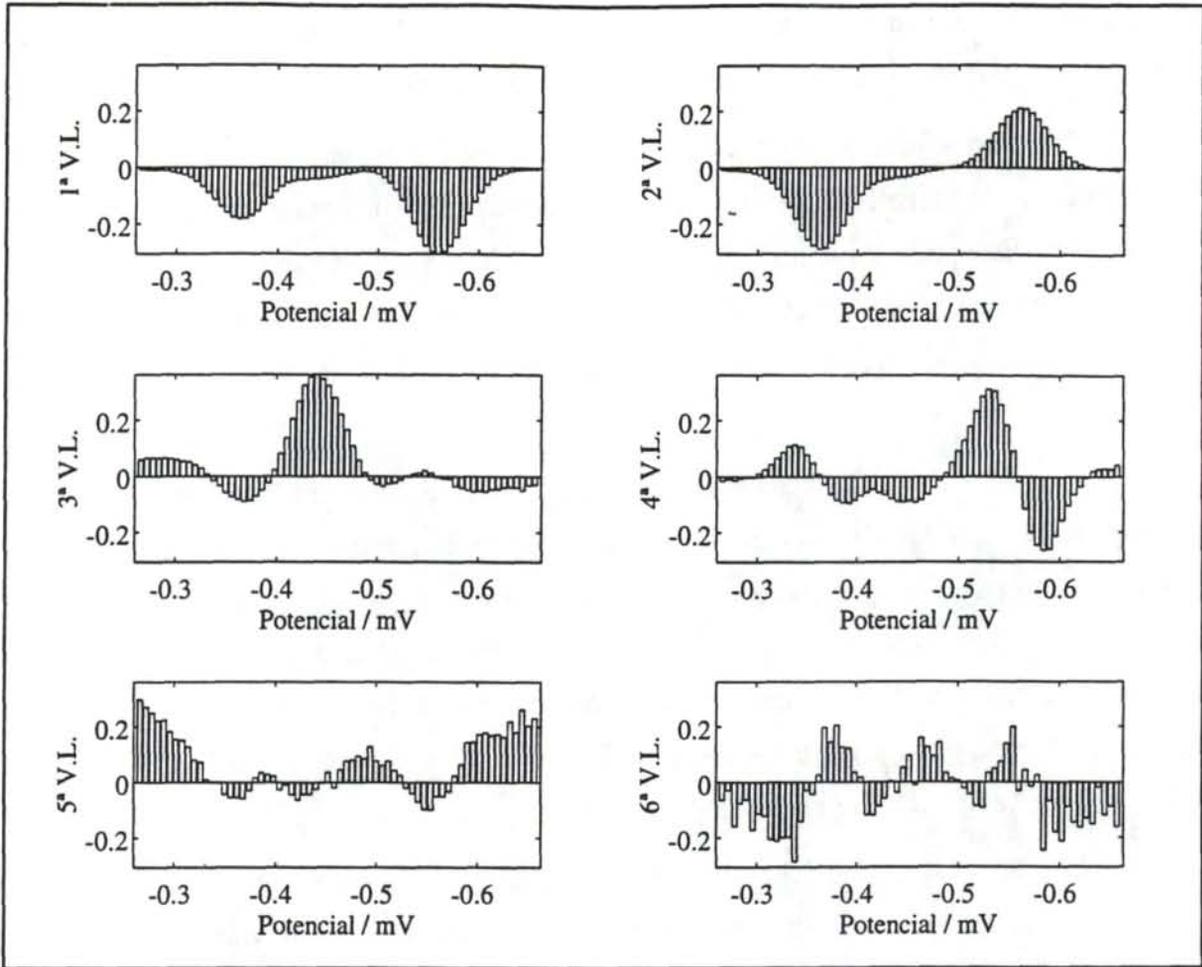


Figura III.5.3 Loadings del modelo PLS construido para la determinación de plomo en el calibrado E.

estaño tienen signo positivo en el modelo construido para el plomo, lo cual indica que están restando la aportación de este analito en la segunda variable latente, donde los loadings relacionados con el plomo tienen signo negativo (se trata de una aportación positiva a la señal). Mientras que, en el modelo PLS correspondiente al estaño, los potenciales a los que aparece el pico del plomo tienen signo positivo, contrario a los correspondientes al analito cuya concentración se modela, el estaño. En ambos modelos, las variables latentes restantes modelan otras fuentes de variabilidad cuyo aporte a capacidad predictiva del modelo es menor.

Sin embargo, los loadings correspondiente a la segunda variable latente del modelo PLS construido para el cadmio presentan una estructura opuesta a la indicada anteriormente para esta misma variable latente en los modelos de regresión de los otros dos analitos. En este caso, los loadings relacionados con el pico del cadmio tienen signo negativo (como la señal)

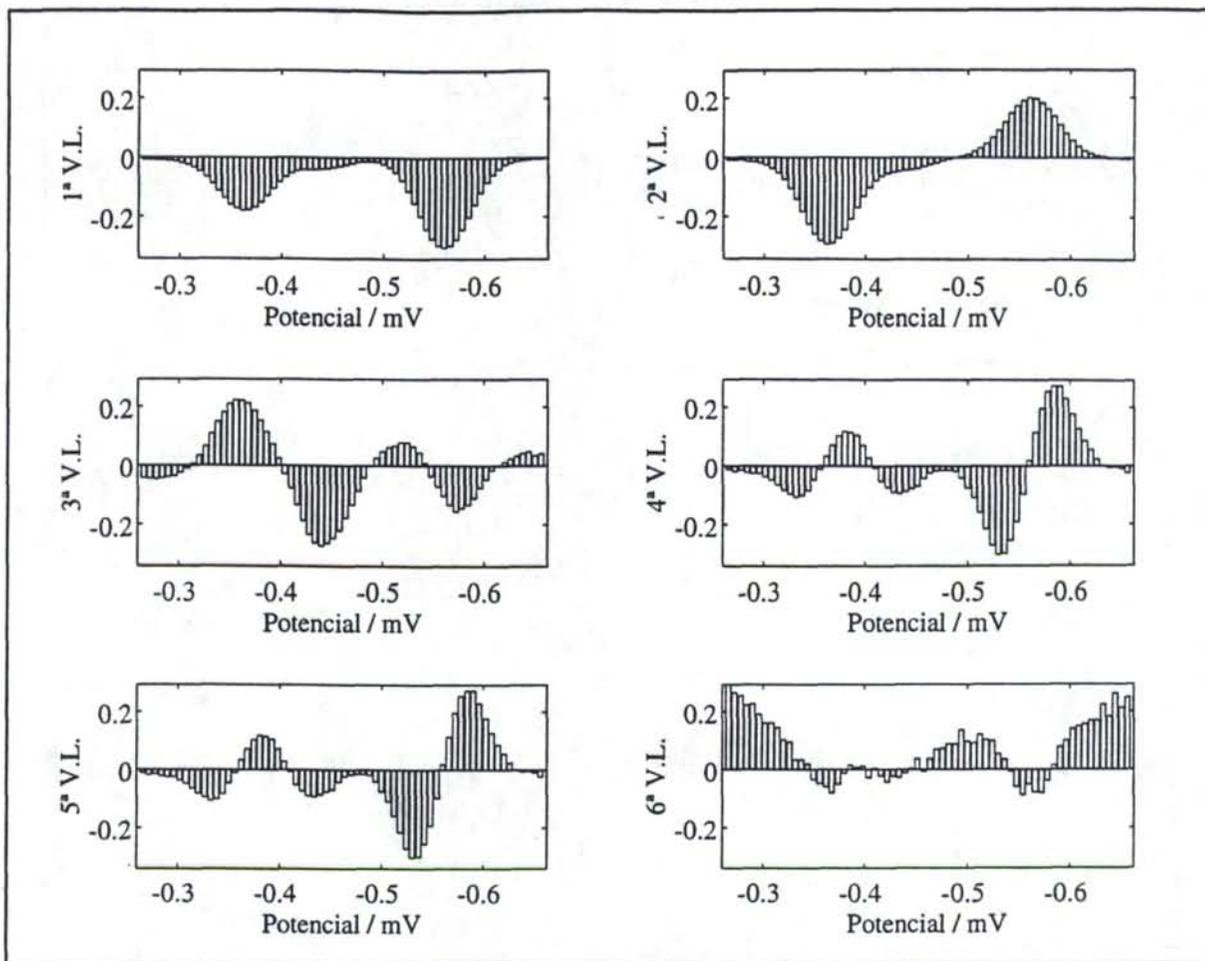


Figura III.5.4 Loadings del modelo PLS construido para la determinación de estaño en el calibrado E.

y los correspondientes al plomo y al estaño, signo positivo (esta variable latente está restando su contribución a la señal). La tercera variable latente supone una corrección del ajuste hecho por la componente anterior, estando principalmente relacionada con los potenciales del pico del estaño, ya que los loadings correspondientes al pico del cadmio presentan signo positivo y negativo, lo cual hace que su contribución a esta variable latente se vea anulada. Las variables restantes, como se ha indicado anteriormente, aportan información al modelo, pero la significación de las mismas es ya menor.

Las Figuras III.5.6 a III.5.8 muestran las proyecciones de los objetos sobre las variables latentes correspondientes a esos mismos modelos de regresión. De nuevo, la estructura que presentan los scores en el primer plano de proyección (formado por la variables latentes primera y segunda) es idéntica para plomo y estaño, resultando ambas estructuras y la correspondiente al cadmio simétricas respecto del eje de abscisas. En dichas gráficas se

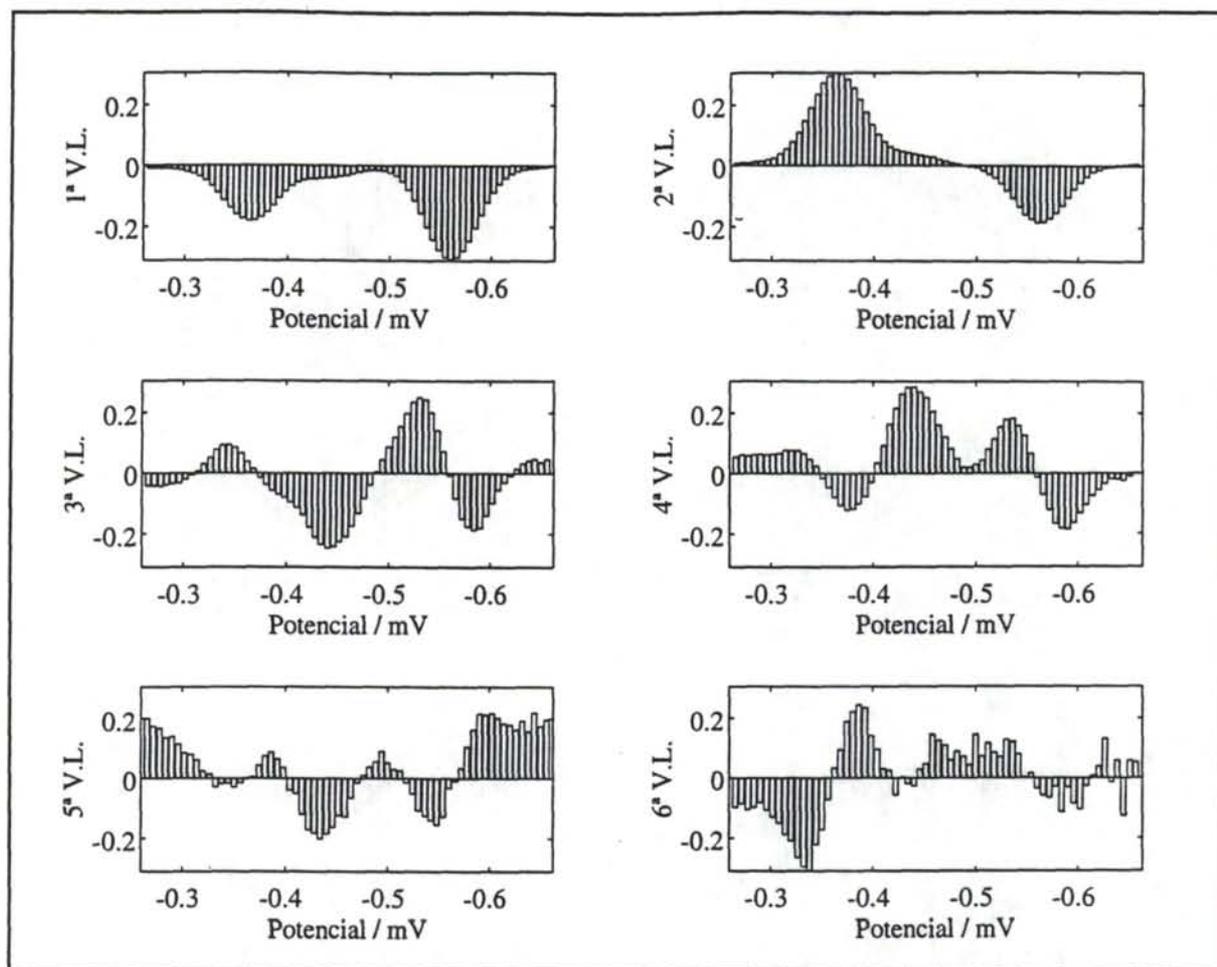


Figura III.5.5 Loadings del modelo PLS construido para la determinación de cadmio en el calibrado E.

podrían trazar sendos ejes, de la muestra 19 a la 23 se encontraría un eje que determina los cinco niveles de concentración del cadmio, mientras que el otro eje, que uniría las muestras 9 y 13, juega ese mismo papel con la concentración de plomo.

Respecto al plano formado por las variables latentes primera y tercera, se observa también la posibilidad de trazar un eje que una las muestras 14 y 18, que determinaría la concentración de estaño, aunque en este caso, la disposición del resto de las muestras respecto de este eje no resulta tan clara como en los dos casos anteriores porque también se nota cierta influencia de la concentración de los otros dos analitos en el mismo.

La capacidad real de predicción de los modelos construidos puede evaluarse a través de las concentraciones calculadas para las muestras del conjunto test. Los errores relativos obtenidos en el cálculo de las mismas, considerando para cada modelo PLS el número de variables latentes indicado anteriormente, se recogen en la Tabla III.5.3. En la misma se

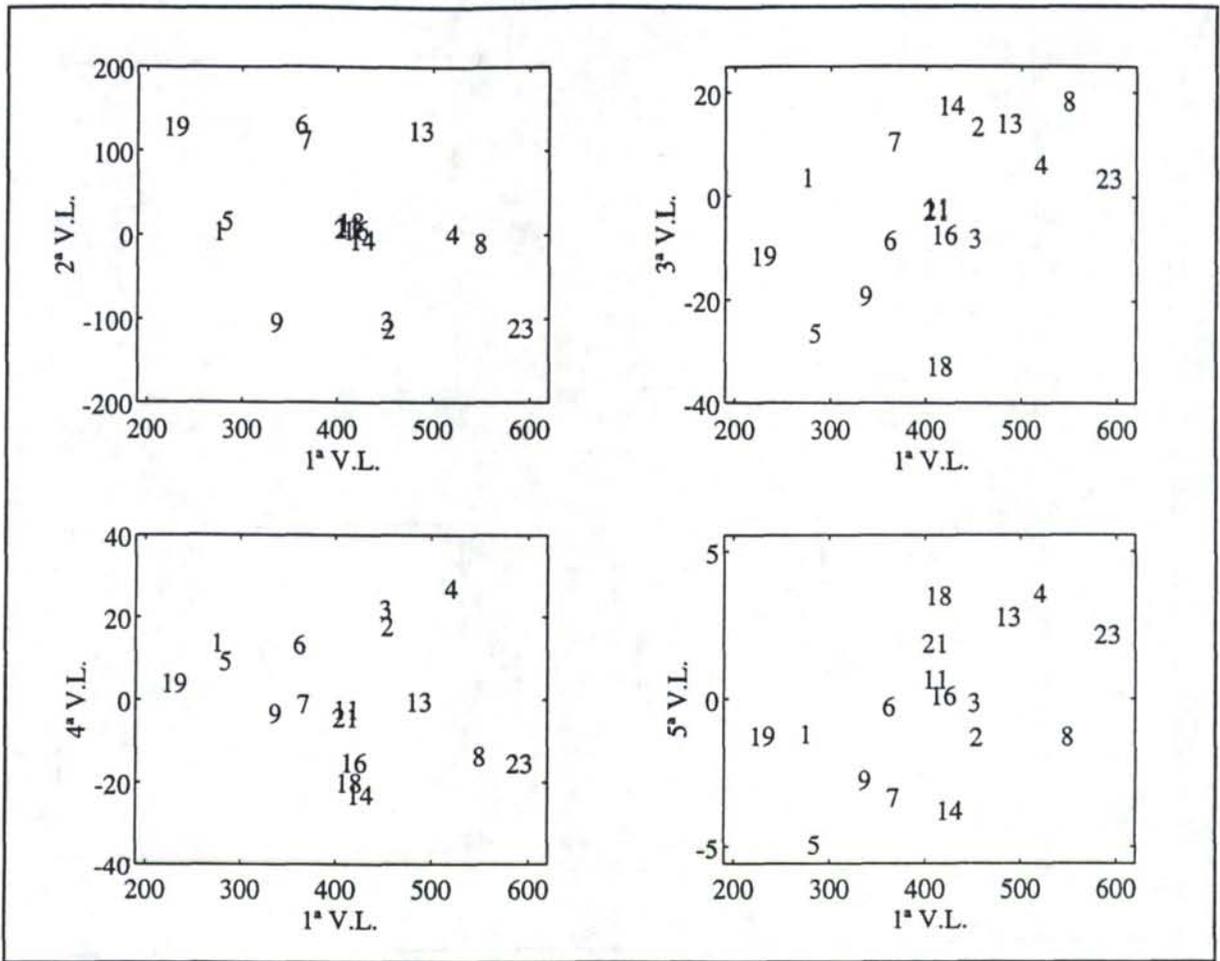


Figura III.5.6 Scores del modelo PLS construido para la determinación de plomo en el calibrado E.

puede observar que la mayor parte de los porcentajes de error se encuentran por debajo o muy próximos al 1%. Se concluye por tanto, teniendo en cuenta los valores de varianza explicada y de varianza en predicción de la Tabla III.5.2 y los errores de la Tabla III.5.3, que los modelos PLS construidos para cada analito en el calibrado E permiten calcular las concentraciones de los distintos analitos de forma muy exacta y poseen una buena capacidad de predicción.

Por otro lado, en un análisis similar llevado a cabo con los datos del calibrado G, los modelos PLS construidos para cada analito en esta nueva situación han conducido a los valores de varianza que se recogen en la Tabla III.5.4. Se observa que éstos son muy similares a los valores de varianza explicada correspondientes al calibrado E, Tabla III.5.2, salvo en el caso del plomo, en el que se llega a explicar el 96.64% de varianza (95.43% de varianza en predicción). Este porcentaje de varianza explicada, sensiblemente inferior al

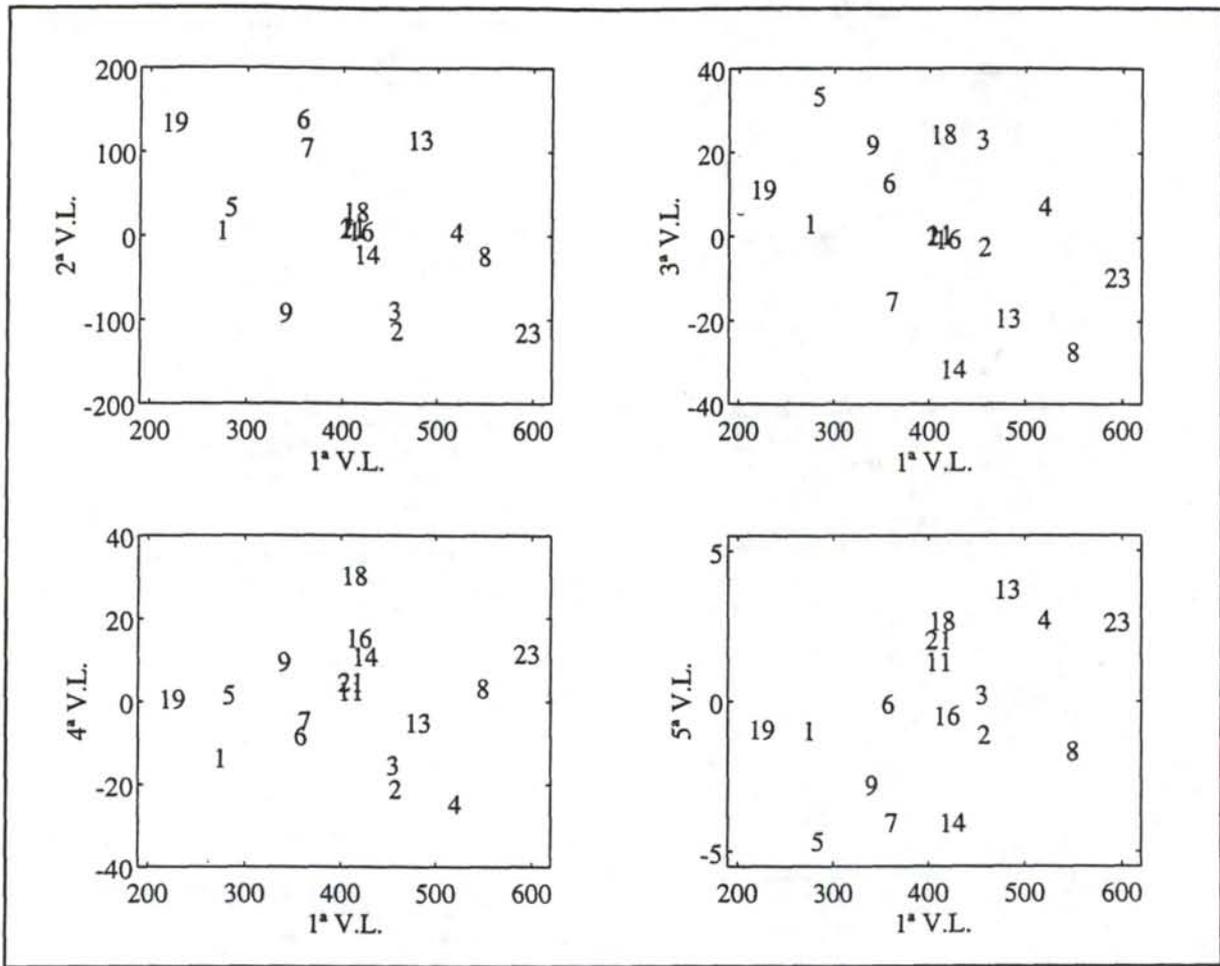


Figura III.5.7 Scores del modelo PLS construido para la determinación de estaño en el calibrado E.

Tabla III.5.3 Errores relativos (%) correspondientes al calibrado E.

Número de muestra	Analito		
	Plomo	Estaño	Cadmio
10	1.15	-0.85	-0.49
12	0.96	-0.01	-0.20
15	0.48	1.39	0.83
17	-1.30	-0.35	0.25
20	-0.56	-3.41	-0.30
22	0.27	1.48	-0.17

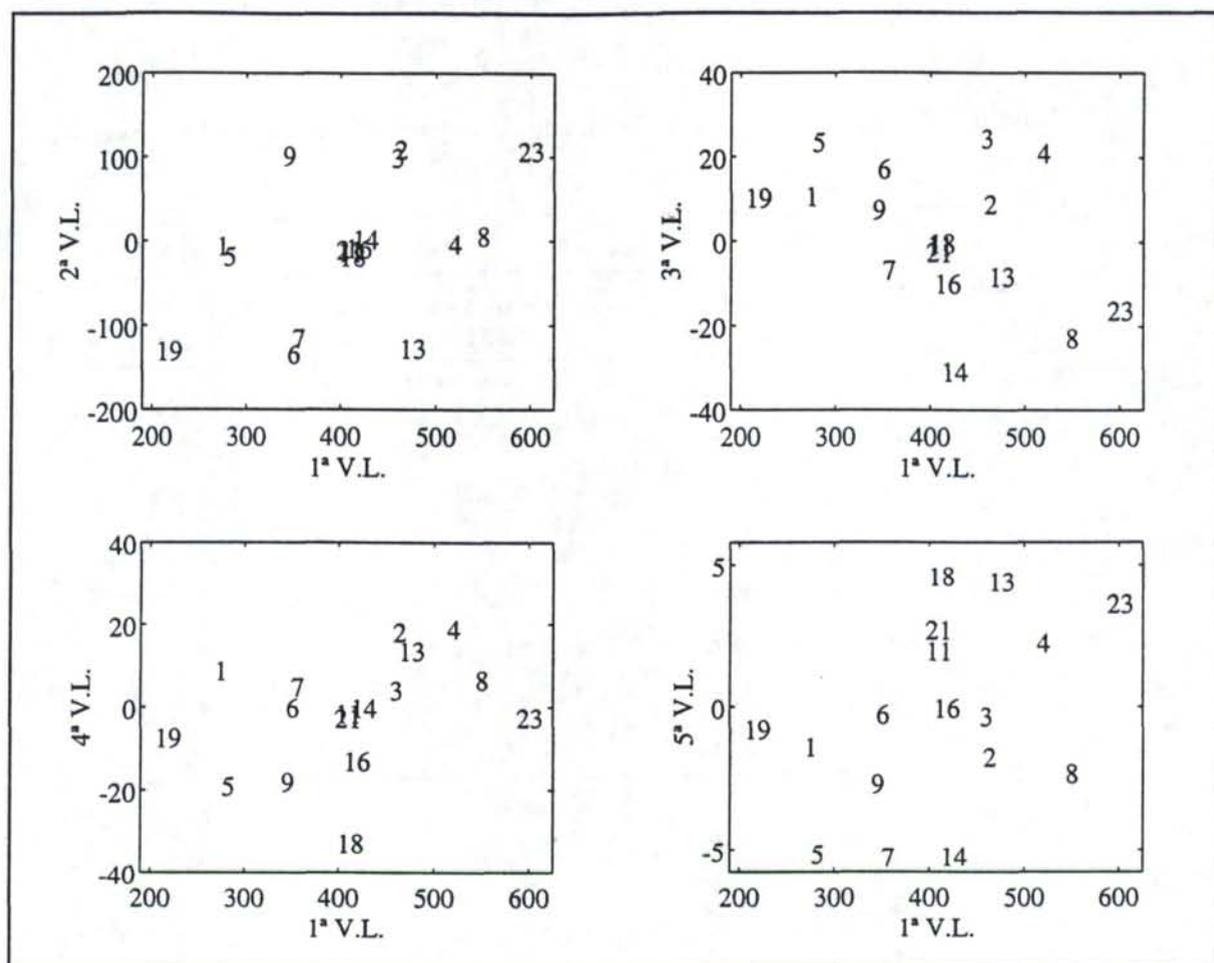


Figura III.5.8 Scores del modelo PLS construido para la determinación de cadmio en el calibrado E.

observado para el calibrado E, ha llevado a un estudio más detallado del conjunto de datos utilizado para construir este modelo, observándose que las intensidades correspondientes al pico del plomo en las muestras 10, 11, 12 y 13, en este segundo calibrado, son anormalmente bajas en comparación con el resto de las adiciones. Esta disminución en el porcentaje de varianza del modelo PLS se debe, por tanto, a un error en la cantidad de plomo que se adicionó a la disolución en dichas muestras, lo cual afecta parcialmente a la señal correspondiente al estaño y parece no interferir en absoluto en el pico polarográfico del cadmio.

Al igual que en el caso anterior, los porcentajes de varianza en predicción se encuentran muy próximos a los valores de varianza explicada correspondientes, quedando patente la estabilidad de los modelos construidos y la buena disposición de estos para predecir muestras ajenas a ellos. En este caso, son necesarias cuatro variables latentes para el modelo

Tabla III.5.4 Porcentajes de varianza explicada obtenidos por los modelos PLS para el calibrado G.

Analito	V.L.	Bloque Y		Bloque X
		Varianza explicada	Varianza en predicción	Varianza explicada
Plomo	1	1.09	0.81	79.51
	2	93.82	92.83	95.43
	3	96.25	94.84	97.38
	4	96.64	95.43	99.07
	5	97.34	94.55	99.30
	6	98.54	94.20	99.40
	7	99.45	94.51	99.49
	8	99.75	94.14	99.53
Estaño	1	-	-	79.51
	2	-	-	95.00
	3	98.54	97.47	97.51
	4	98.72	98.30	98.98
	5	99.49	98.89	99.33
	6	99.73	98.63	99.40
	7	99.88	99.01	99.49
	8	99.95	99.03	99.53
Cadmio	1	85.12	83.99	79.51
	2	99.66	99.65	95.43
	3	99.75	99.74	95.84
	4	99.78	99.71	97.59
	5	99.90	99.88	99.23
	6	99.95	99.89	99.46
	7	99.98	99.93	99.52
	8	100.00	99.93	99.52

correspondiente al plomo, cinco para el del estaño y siete para el cadmio, una más que para el modelo PLS construido para este mismo analito en el calibrado E (el incremento de varianza en predicción de la sexta a la séptima variable latente es únicamente de un 0.05%).

Si se tienen en cuenta los errores relativos correspondientes a las concentraciones de los distintos analitos en las muestras del conjunto test, que se muestran en la Tabla III.5.5, se observa que éstos se encuentran ligeramente por encima de los obtenidos para el calibrado E, Tabla III.5.3, lo cual se debe, como ya se ha indicado anteriormente a un error en la adición de plomo en algunas de las muestras. De hecho, el error mayor (-5.87%) se corresponde precisamente con la muestra número 12, que es la de concentración de plomo más elevada del diseño y una de las asociadas con una adición incorrecta.

Tabla III.5.5 Errores relativos (%) correspondientes al calibrado G.

Número de muestra	Analito		
	Plomo	Estaño	Cadmio
10	-0.18	-0.01	0.85
12	-5.87	1.81	0.57
15	2.04	2.98	0.01
17	-3.90	-1.04	0.47
20	-1.12	-1.30	-3.48
22	1.93	-3.63	-2.15

Con el fin de determinar la validez del calibrado G, se construyeron nuevos modelos de regresión PLS para los tres analitos, eliminando del conjunto de aprendizaje las muestras 11 y 13. El modelo PLS construido para el plomo en estas condiciones consta de ocho variables latentes y explica un 99.99% de varianza (99.18% de varianza en predicción), mientras que para el estaño y el cadmio los modelos obtenidos son similares a los anteriores en cuanto a valores de varianza explicada. En lo que se aprecian diferencias significativas es en los porcentajes de error que resultan al calcular las muestras del conjunto test, éstos aparecen reflejados en la Tabla III.5.6, en los cuales se observa que las muestras 10 y 12 se encuentran asociadas con los errores mayores, sobre todo los correspondientes al plomo.

Tabla III.5.6 Errores relativos (%) del calibrado G al eliminar los objetos 11 y 13 del conjunto de aprendizaje para construir los modelos de calibrado PLS.

Número de muestra	Analito		
	Plomo	Estaño	Cadmio
10	-11.98	-0.67	1.63
12	-14.56	5.11	2.78
15	0.74	1.45	0.28
17	-0.14	-0.03	0.46
20	-2.22	-1.47	-3.80
22	-0.86	-1.51	-1.26

Una vez que ambos calibrados, E y G, han sido estudiados de forma exhaustiva, se ha procedido a llevar a cabo la transferencia de calibrado considerando como calibrado patrón el conjunto de datos E y como calibrado satélite el conjunto de datos G (ya que este último fue obtenido con posterioridad).

III.5.3.2 Transferencia de calibrado mediante el procedimiento PDS

Como ya se ha señalado anteriormente, la finalidad de este procedimiento de transferencia de calibrado es la de utilizar un modelo de calibrado, construido en una situación determinada, con otro conjunto de datos obtenido en una situación diferente. En este caso, se trata de transferir la ecuación de regresión de los modelos construidos con los datos del calibrado E, a un conjunto de datos polarográficos recogidos con el mismo instrumento pero días después, es decir, los datos del calibrado G.

Además, el realizar la transferencia de calibrado puede permitir una reducción del número de muestras experimentales necesarias para realizar determinaciones en diferentes situaciones, ya que únicamente se requiere un número mínimo de muestras para calcular la matriz de transferencia que relacione el calibrado patrón y el calibrado satélite. Una vez que se dispone de la matriz de transferencia, cualquier polarograma obtenido en la situación satélite puede ser corregido, calculándose después las concentraciones de los distintos analitos

correspondientes al mismo utilizando el modelo de regresión obtenido con el calibrado patrón completo.

Aplicación de la función "pdsgen"

Para llevar a cabo la transferencia de calibrado se ha utilizado la función *pdsgen* [125] implementada en el programa MATLAB [59], cuyos parámetros de entrada/salida se han definido anteriormente y que en esta ocasión son los siguientes:

$$[F, sub] = pdsgen(E, G, w, no)$$

cuya asignación se especifica a continuación:

- E:** Matriz con los polarogramas del calibrado E, es decir, las señales experimentales registradas en la situación E. Consta de 23 filas (un polarograma por cada muestra de la Tabla III.5.1) y de 66 columnas (los potenciales digitalizados en cada polarograma). Es decir, los datos del calibrado patrón se encuentran en la matriz E.
- G:** Matriz que recoge los polarogramas utilizados en la construcción del modelo PLS del calibrado G. Tiene las mismas dimensiones que la matriz E, y tanto sus filas como sus columnas se encuentran en el mismo orden que en esta última matriz. La matriz G contiene los datos de la situación satélite, calibrado G.
- w:** Escalar que especifica cuántos potenciales se incluyen en la ventana móvil que se utiliza en el cálculo de la matriz de transferencia. En las transferencias de calibrado llevadas a cabo con diversos conjuntos de transferencia se han utilizado para este parámetro los siguientes tamaños de ventana: 3, 5, 7, 9, 11 y 13.
- no:** Vector fila en el que se indica qué muestras de ambas matrices se utilizan en la transferencia de calibrado. En él se señalan los índices correspondientes a las muestras de la Tabla III.5.1, las cuales han sido seleccionadas para formar parte de los diferentes conjuntos de transferencia de acuerdo con lo especificado en la Tabla III.5.7, de forma que las dimensiones del vector *no* son $1 \times d$, donde *d* varía dependiendo del número de muestras utilizadas para llevar a cabo la transferencia.

Como resultado de aplicar esta función, considerando los dos conjuntos de datos indicados, se obtienen las siguientes matrices: la matriz de transferencia (F) y un vector fila en el que se especifica qué muestras han intervenido en la transferencia de calibrado, *sub*. En este caso, la matriz de transferencia es una matriz banda diagonal, de dimensiones 66×66 ,

que difiere para cada par de valores de los parámetros no y w introducidos, mientras que el vector sub coincide, en cada caso, con los índices señalados en no .

Tabla III.5.7 Subconjuntos de muestras del conjunto de aprendizaje utilizados para calcular la matriz de transferencia teniendo en cuenta los símbolos que aparecen en la Tabla III.5.1.

Conjunto (no)	Muestras del conjunto de aprendizaje	Disolución inicial
1	todas	todas
2	1 a 8	○×·
3	9 11 13 14 16 18 19 21 23	▲□·
4	9 14 16 18 19 21 23	▲□·
5	9 14 18 19 23	▲□·
6	1 2 3 4 7 8	○·
7	9 13 14 18 19 23	▲□·

Una vez que se dispone de la matriz de transferencia, es posible transformar cualquier polarograma obtenido en la situación G e introducirlo en la regresión PLS obtenida con todas las muestras del conjunto de aprendizaje del conjunto E para calcular las concentraciones de plomo, estaño y cadmio correspondientes al mismo. Esta operación se lleva a cabo de forma similar a lo especificado anteriormente para este procedimiento de transferencia de calibrado, utilizándose para ello la fórmula cerrada de la regresión PLS, propuesta por Marengo y Todeschini [126], y considerando el número de variables latentes más adecuado para cada analito en los modelos PLS del calibrado patrón.

El esquema correspondiente a este procedimiento es análogo a los desarrollados previamente para el procedimiento PDS, salvo en lo que se refiere a los nombres y dimensiones de las diferentes matrices, que en este caso se deberán ajustar a los correspondientes conjuntos de datos E y G. Por tanto, como resultado final del procedimiento de transferencia de calibrado se obtiene una matriz de dimensiones 23×3 , Y_{cal} , en la que se encuentran las concentraciones de los tres analitos calculadas por los modelos de regresión PLS del calibrado E para todas las muestras del calibrado G.

En dicha matriz es posible establecer diferencias entre unas muestras y otras. Por ejemplo, aquellas muestras que formaron parte del conjunto test cuando se construyeron los calibrados multivariantes PLS para cada conjunto de datos, se utilizan también en esta etapa del análisis para evaluar la capacidad de predicción de los modelos de calibrado una vez que se ha llevado a cabo la transferencia de los mismos. Por otro lado, aunque las restantes muestras del calibrado E participaron en la construcción del modelo PLS, se puede también distinguir entre aquellas que además han formado parte del conjunto de transferencia y las que quedaron fuera del mismo.

Análisis de los resultados

El error estándar en predicción (SEP) y el error estándar de transferencia (SET) son, como ya se ha indicado anteriormente, dos parámetros que permiten evaluar la calidad de las concentraciones calculadas tras el procedimiento de transferencia de calibrado. En el primero de ellos se consideran únicamente las determinaciones efectuadas en las muestras pertenecientes al conjunto test del calibrado G, mientras que el segundo tiene en cuenta el resto de los objetos del calibrado.

Estos parámetros permiten evaluar, además de las predicciones realizadas a través de la transferencia de calibrado, el efecto que sobre éstas tienen los dos factores que se han variado para llevar a cabo los distintos procedimientos de transferencia de calibrado: el tamaño de ventana y el conjunto de transferencia utilizado. Durante el análisis se ha observado que la dependencia de ambos parámetros con estos dos factores es muy similar, por lo que únicamente se muestran los valores de SEP, obtenidos al aplicar el procedimiento de transferencia de calibrado, a través de las correspondientes superficies de respuesta y curvas de nivel de las Figuras III.5.9 a III.5.11.

En dichas figuras se observa que la influencia de los distintos tamaños de ventana sobre los valores de SEP es importante, sobre todo para algunos conjuntos de transferencia concretos. Los diferentes conjuntos de transferencia de calibrado utilizados también ejercen una fuerte influencia en los valores obtenidos para el SEP, siendo ésta incluso de mayor magnitud que la debida al factor w en la mayoría de los casos.

Se ha realizado un análisis más detallado de los valores de SEP mediante la representación de los correspondientes diagramas Box plot, Figuras III.5.12 a III.5.14. Así mismo, se ha realizado un estudio de estos errores aplicando un análisis de rangos múltiple

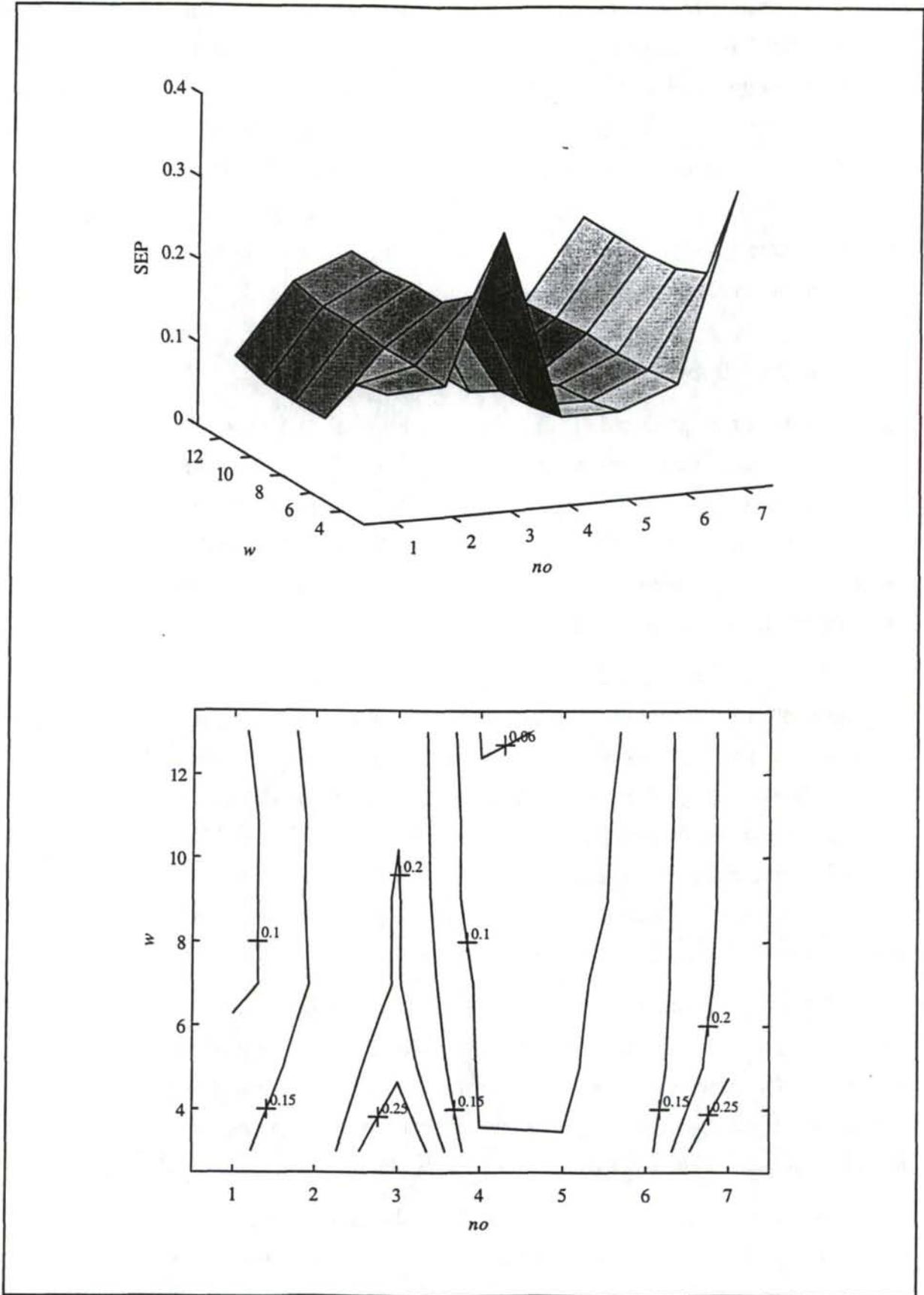


Figura III.5.9 Superficie de respuesta y curvas de nivel del SEP para el plomo.

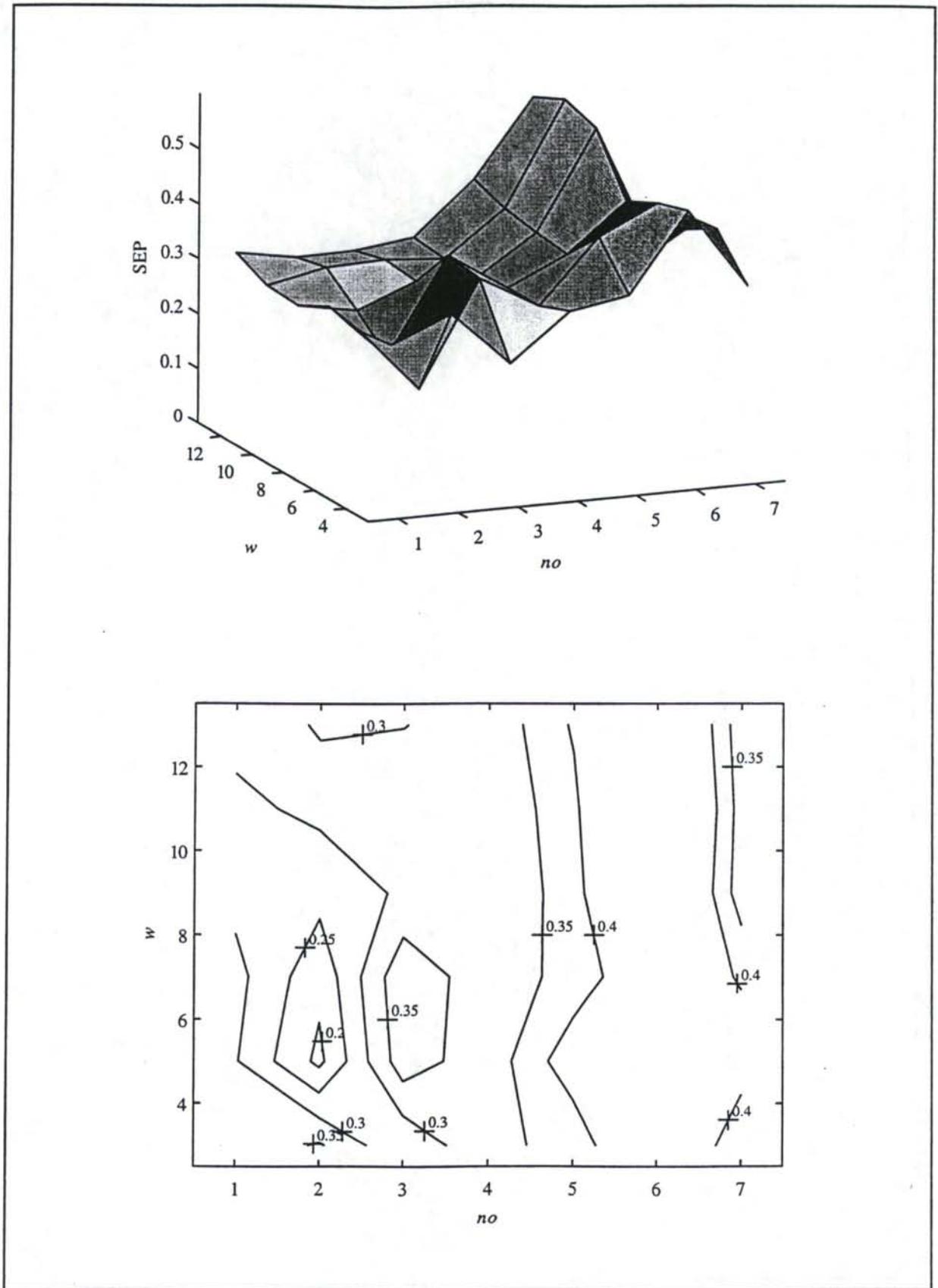


Figura III.5.10 Superficie de respuesta y curvas de nivel del SEP para el estaño.

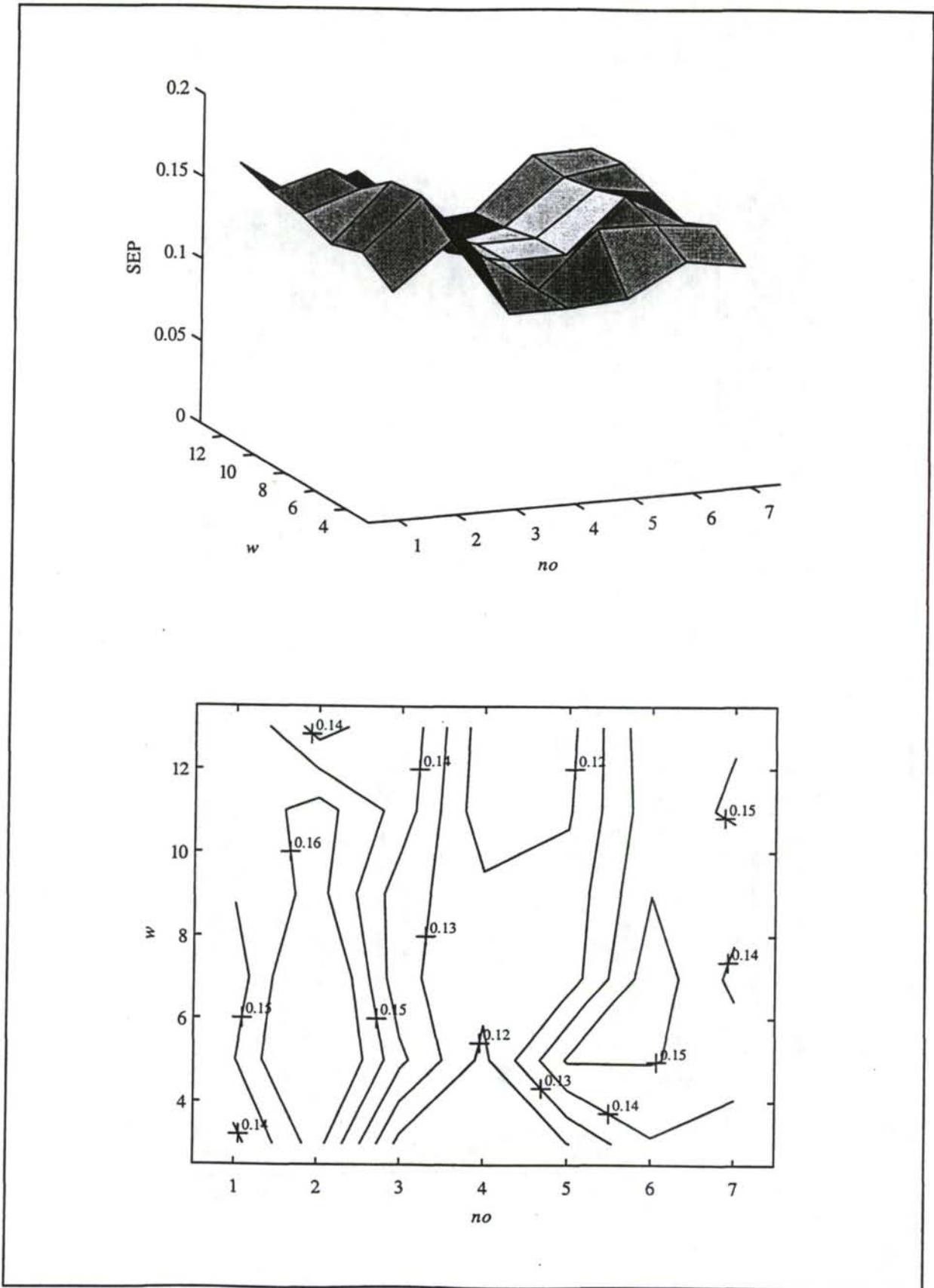


Figura III.5.11 Superficie de respuesta y curvas de nivel del SEP para el cadmio.

mediante los test de agrupamientos para la media de Tukey y de Newman-Keuls, cuyos resultados se muestran en las Tablas III.5.8 a III.5.10.

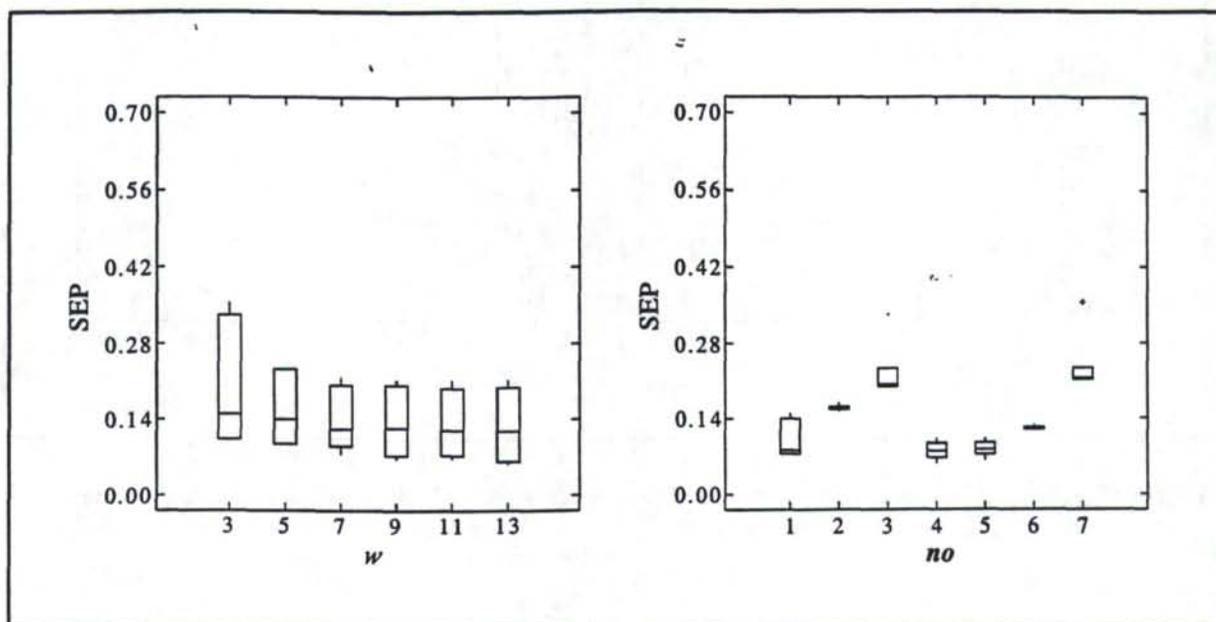


Figura III.5.12 Diagramas Boxplot correspondientes a los valores de SEP obtenidos en la transferencia de calibrado del plomo.

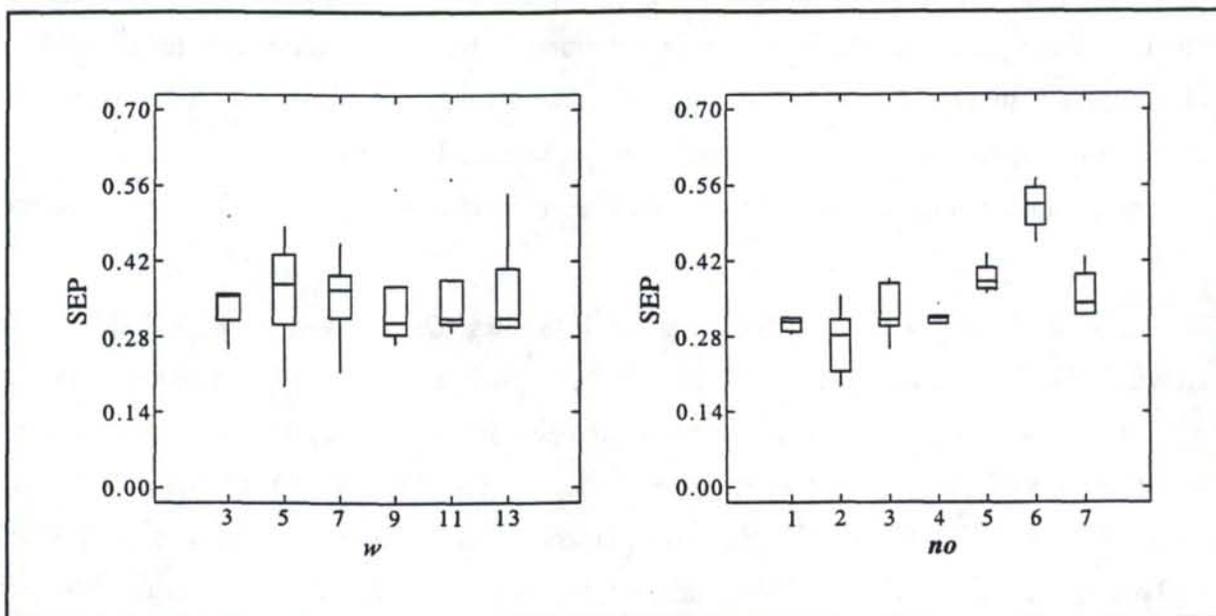


Figura III.5.13 Diagramas Boxplot correspondientes a los valores de SEP obtenidos en la transferencia de calibrado del estaño.

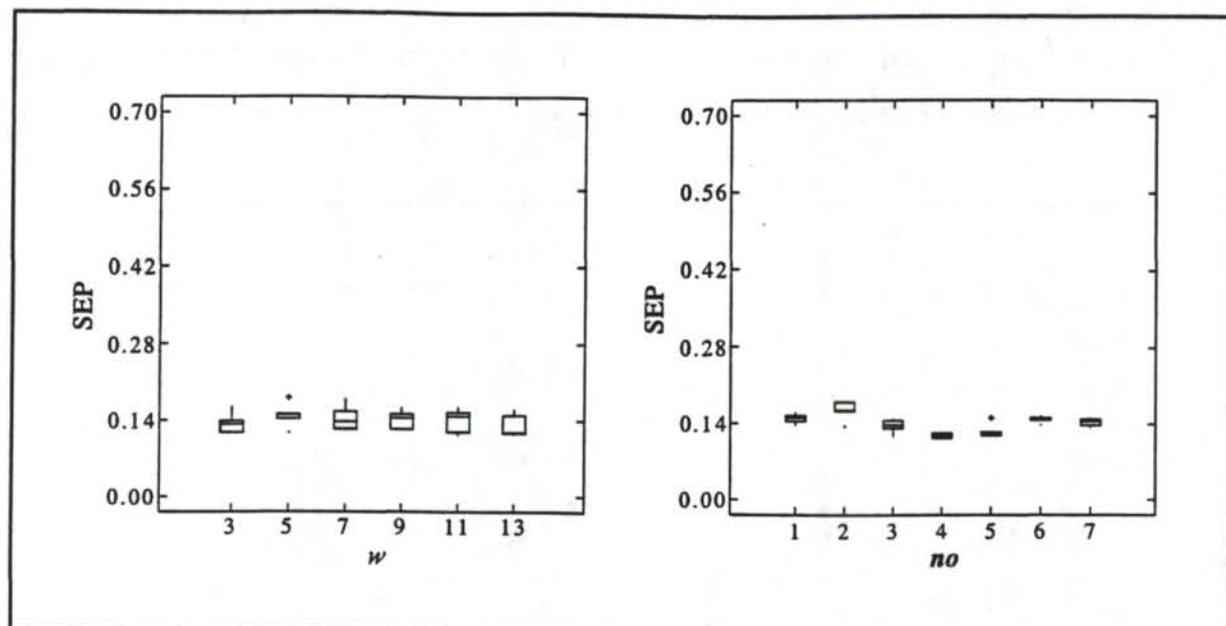


Figura III.5.14 Diagramas Boxplot correspondientes a los valores de SEP obtenidos en la transferencia de calibrado del cadmio.

Los diagramas Box plot señalan cómo la mayor dispersión existente entre los valores de SEP está relacionada con los diferentes conjuntos de transferencia utilizados, más que con los tamaños de ventana. Por otro lado, dado que este último parámetro parece ejercer una influencia limitada sobre los valores de SEP, cabría la posibilidad de seleccionar un mismo tamaño de ventana en las transferencias de calibrado de todos los analitos para así simplificar el análisis computacional. De esta forma, sería posible calcular una única matriz de transferencia que se podría utilizar para corregir las señales polarográficas una sola vez, y posteriormente evaluar la concentración los tres analitos de forma simultánea en un mismo análisis.

Si se tienen en cuenta los resultados de los test de agrupamientos de Tukey y de Newman-Keuls obtenidos para el factor ventana es posible concluir que los únicos tamaños de ventana que en algún caso puede considerarse estadísticamente diferentes del resto son los tamaños 3 y 5, mientras que entre el resto de los tamaños no existen diferencias estadísticamente significativas. Y dado que, como se puede observar en las curvas de nivel de las Figuras III.5.9 y III.5.11, el tamaño de ventana 13 conduce a los valores de SEP más bajos para plomo y cadmio, e incluso, mostrando poca influencia para algunos conjuntos de transferencia en el caso del estaño, se ha seleccionado este tamaño de ventana como el óptimo para llevar a cabo el análisis. Aunque, por supuesto, el valor de w indicado deberá

proporcionar resultados adecuados en combinación con el conjunto de transferencia que se elija finalmente para el análisis.

Tabla III.5.8 Resultados de los test de agrupamientos de los valores de SEP obtenidos para el plomo.

Conjunto de transferencia (<i>no</i>)				Ventana (<i>w</i>)			
Nivel	Media	Grupos homogéneos		Nivel	Media	Grupos homogéneos	
		Tukey	Newman-Keuls			Tukey	Newman-Keuls
4	0.0797	X	X	13	0.1294	X	X
5	0.0824	X	X	11	0.1296	X	X
1	0.1003	X	XX	9	0.1310	X	X
6	0.1226	XX	X	7	0.1367	X	X
2	0.1595	X	X	5	0.1535	XX	X
3	0.2282	X	X	3	0.1889	X	X
7	0.2412	X	X				

Tabla III.5.9 Resultados de los test de agrupamientos de los valores del SEP obtenidos para el estaño.

Conjunto de transferencia (<i>no</i>)				Ventana (<i>w</i>)			
Nivel	Media	Grupos homogéneos		Nivel	Media	Grupos homogéneos	
		Tukey	Newman-Keuls			Tukey	Newman-Keuls
2	0.2722	X	X	9	0.3451	X	X
1	0.3033	XX	XX	7	0.3506	X	X
4	0.3165	XXX	XX	3	0.3552	X	X
3	0.3243	XXX	XX	13	0.3567	X	X
7	0.3585	XX	XX	11	0.3585	X	X
5	0.3887	X	X	5	0.3624	X	X
6	0.5198	X	X				

En lo que se refiere al efecto que sobre los valores del SEP tienen los diferentes conjuntos de transferencia, existen diferencias notables entre unos conjuntos y otros, como queda patente en las figuras correspondientes a los diagramas Box plot. No obstante, no hay

Tabla III.5.10 Resultados de los test de agrupamientos de los valores de SEP obtenidos para el cadmio.

Conjunto de transferencia (<i>no</i>)				Ventana (<i>w</i>)			
Nivel	Media	Grupos homogéneos		Nivel	Media	Grupos homogéneos	
		Tukey	Newman-Keuls			Tukey	Newman-Keuls
4	0.1178	X	X	3	0.1325	X	X
5	0.1257	XX	XX	13	0.1389	XX	XX
3	0.1367	XX	XX	9	0.1412	XX	XX
7	0.1436	X	X	11	0.1418	XX	XX
6	0.1485	XX	X	7	0.1424	XX	XX
1	0.1492	XX	X	5	0.1485	X	X
2	0.1648	X	X				

que olvidar que uno de los objetivos de la transferencia de calibrado que se está llevando a cabo es la reducción del número de muestras del calibrado satélite y, por tanto, éste ha de ser un elemento importante a tener en cuenta en el análisis.

Considerando los conjuntos de transferencia 1, 3 y 7, en los cuales se encuentran incluidas las muestras 11 y 13 del conjunto de aprendizaje del calibrado G (que estaban asociadas con adiciones inexactas de plomo) se observa que están asociados con valores de SEP elevados, sobre todo los conjuntos 3 y 7, en los que el número de objetos es menor y, por tanto, las muestras 11 y 13 tienen mayor peso. De hecho, el conjunto de transferencia que conduce a los mejores resultados, como se desprende de los test de agrupamientos realizados, es el conjunto 4, que tiene en común con el conjunto 3 todas las muestras de transferencia excepto esas dos.

Por otro lado, el conjunto de transferencia 4 está relacionado con los valores de SEP mínimos en las transferencias llevadas a cabo para plomo y cadmio, y con valores medios en el caso del estaño, Figuras III.5.9 a III.5.11. Para este último analito, el conjunto de transferencia con el que obtienen los valores de SEP más bajos es el conjunto 2, que sin embargo, se corresponde con zonas en las que el error es mayor en los resultados obtenidos para el plomo y el cadmio.

El conjunto 4 está formado únicamente por siete muestras, a través de las cuales es posible recoger toda la variabilidad del diseño experimental, así como corregir un polarograma correspondiente a cualquier otra muestra del calibrado G a través de la matriz de transferencia calculada, determinándose la concentración de los tres analitos de forma simultánea. Se observa que este conjunto de transferencia es suficiente para describir el calibrado multianalito llevado a cabo y que no es posible reducir el número de muestras necesarias para calcular la matriz de transferencia ya que, por ejemplo, si se considera el conjunto 5, en el que se han eliminado dos de las muestras del conjunto 4, se observa que para el plomo, Tabla III.5.9, las diferencias entre las medias de SEP obtenidas con ambos conjuntos son estadísticamente significativas.

Tabla III.5.11 Valores de SEP y SET obtenidos con los distintos conjuntos de transferencia utilizados ($w = 13$).

Conjunto (no)	SEP			SET		
	Plomo	Estaño	Cadmio	Plomo	Estaño	Cadmio
1	0.09	0.32	0.16	0.32	0.39	0.17
2	0.17	0.30	0.14	0.49	0.49	0.32
3	0.20	0.30	0.15	0.39	0.39	0.18
4	0.06	0.31	0.11	0.35	0.39	0.17
5	0.06	0.41	0.12	0.36	0.42	0.17
6	0.12	0.55	0.15	0.42	0.85	0.16
7	0.21	0.32	0.15	0.40	0.41	0.17
Calibrado G	0.14	0.16	0.14	0.31	0.20	0.06
Calibrado E	0.05	0.19	0.05	0.10	0.21	0.04

No obstante, los valores de SEP y SET correspondientes a las transferencias de calibrado realizadas con los diferentes conjuntos de transferencia de la Tabla III.5.7 y con una anchura de ventana 13 se muestran en la Tabla III.5.11. Hay que señalar que, en el análisis, para calcular los valores del error estándar en predicción, SEP, se han considerado únicamente las muestras 15, 17, 20 y 22 del conjunto test, las cuales no están asociadas con ningún problema en las adiciones de plomo, como se ha indicado previamente. Esta razón explica el hecho de que los valores de SEP correspondientes al plomo sean en todos los casos mucho

menores que los valores de SET respectivos, ya que, en el cálculo de los valores de SET no se ha eliminado ninguna muestra experimental y, como se comprobó anteriormente, en las muestras 11 y 13 del calibrado G se encontraban asociadas con un error experimental importante.

Al igual que los capítulos anteriores, en la Tabla III.5.11 quedan reflejados también los valores de SEP y SET correspondientes a los modelos PLS construidos con los conjuntos de aprendizaje completos de los calibrados E y G sin mediar ninguna transferencia de calibrado, señalados como Calibrado E y Calibrado G respectivamente. Estos valores son, en general, más bajos que los correspondientes valores obtenidos mediante el procedimiento de transferencia de calibrado.

Estos datos evidencian que las condiciones óptimas para llevar a cabo este procedimiento de transferencia de calibrado vienen dadas por el conjunto de transferencia 4, tanto si se consideran los valores de SEP, como los correspondientes al SET. De hecho, los errores relativos en predicción asociados con estas condiciones de transferencia, que se muestran en la Tabla III.5.12, no son mayores que aquellos proporcionados cuando se utilizaron los modelos PLS construidos con todas las muestras de calibrado, Tablas III.5.5 y III.5.6.

Tabla III.5.12 Errores relativos (%) de las muestras test del calibrado E al considerar $no = 4$ y $w = 13$ para calcular la matriz de transferencia.

Número de muestra	Analito		
	Plomo	Estaño	Cadmio
10	-6.83	0.07	1.27
12	-10.53	-0.09	1.64
15	1.50	1.52	0.40
17	-0.61	-2.12	0.68
20	-1.05	-4.81	-2.06
22	-0.32	-2.76	-1.45

III.5.3.3 Transferencia de calibrado mediante el procedimiento GCT

Aplicación de la función "gct"

Se ha aplicado otro procedimiento de transferencia de calibrado para llevar a cabo la transferencia del calibrado E a la situación G de nuevo. De esta forma se puede comprobar qué resultados se obtienen con este tipo de datos polarográficos cuando la transferencia de calibrado se realiza utilizando el procedimiento GCT. Los parámetros introducidos en este caso en la función "gct" han sido los siguientes:

$$[F, sub] = gct(E, G, 15, no, 3, Ecer)$$

donde **E** y **G** son las matrices de datos indicadas anteriormente, **no** es el vector que señala las muestras de los distintos conjuntos de transferencia utilizados, que se encuentran indicados en la Tabla III.5.1, y **Ecer** es la matriz de los coeficientes de los modelos PLS construídos en la situación E.

Análisis de los resultados

Los resultados de este procedimiento de transferencia de calibrado se muestran a través de los correspondientes valores de SEP y SET, para los distintos conjuntos de transferencia, en las Figuras III.5.15 a III.5.17. En ellas se observa que, en el caso del plomo, los valores del SET son mucho mayores que los respectivos valores de SEP, siendo la causa de esto la misma que se señaló en el procedimiento de transferencia PDS, el hecho de que este error comprenda una evaluación de las estimaciones hechas para las muestras 11 y 13.

Aunque se observan algunas diferencias entre los tres analitos, estas figuras parecen indicar que los mejores resultados se obtienen con el subconjunto de transferencia 4 y, en algunos casos, también con los conjuntos 1 y 7. En la Tabla III.5.13 se muestran los errores de las muestras del conjunto test correspondientes a las transferencias de calibrado llevadas a cabo con esos tres conjuntos de transferencia.

Aunque los errores que aparecen en dicha tabla correspondientes al conjunto 7 se encuentran próximos a los obtenidos con los otros dos conjuntos de transferencia, los errores asociados con la concentración de plomo de algunas de las muestras del conjunto de

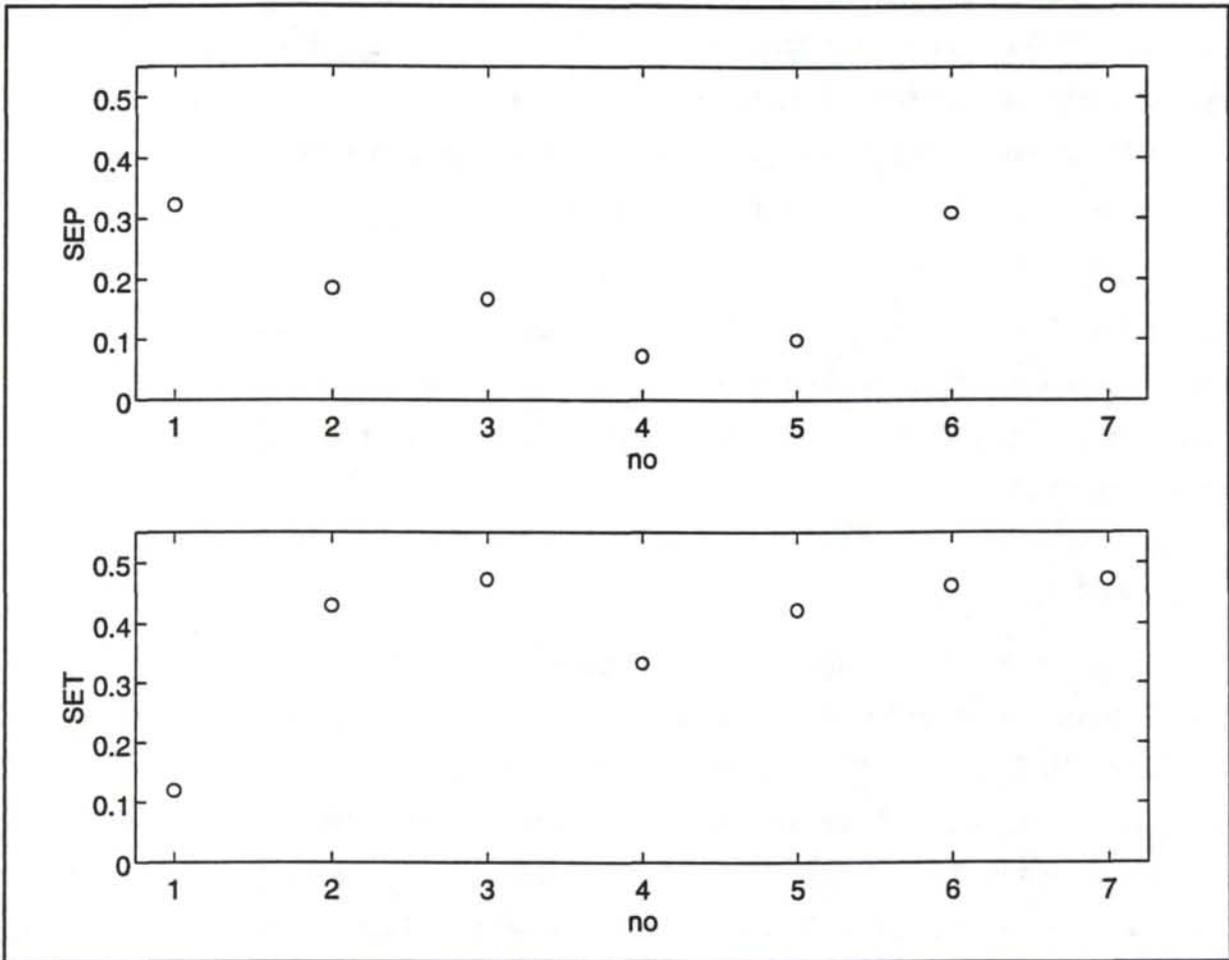


Figura III.5.15 Resultados de la transferencia de calibrado realizada con el procedimiento GCT para el plomo.

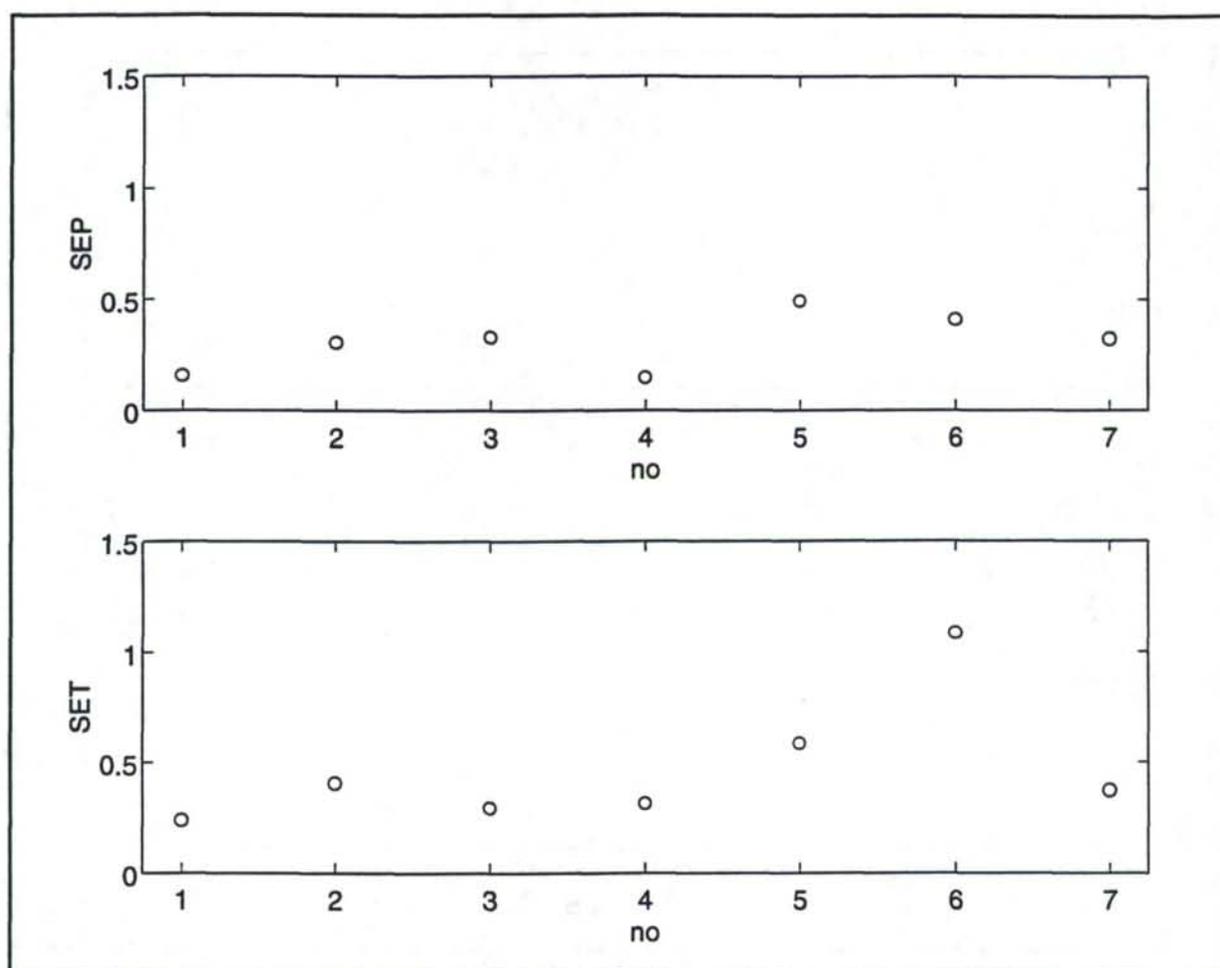


Figura III.5.16 Resultados de la transferencia de calibrado realizada con el procedimiento GCT para el estaño.

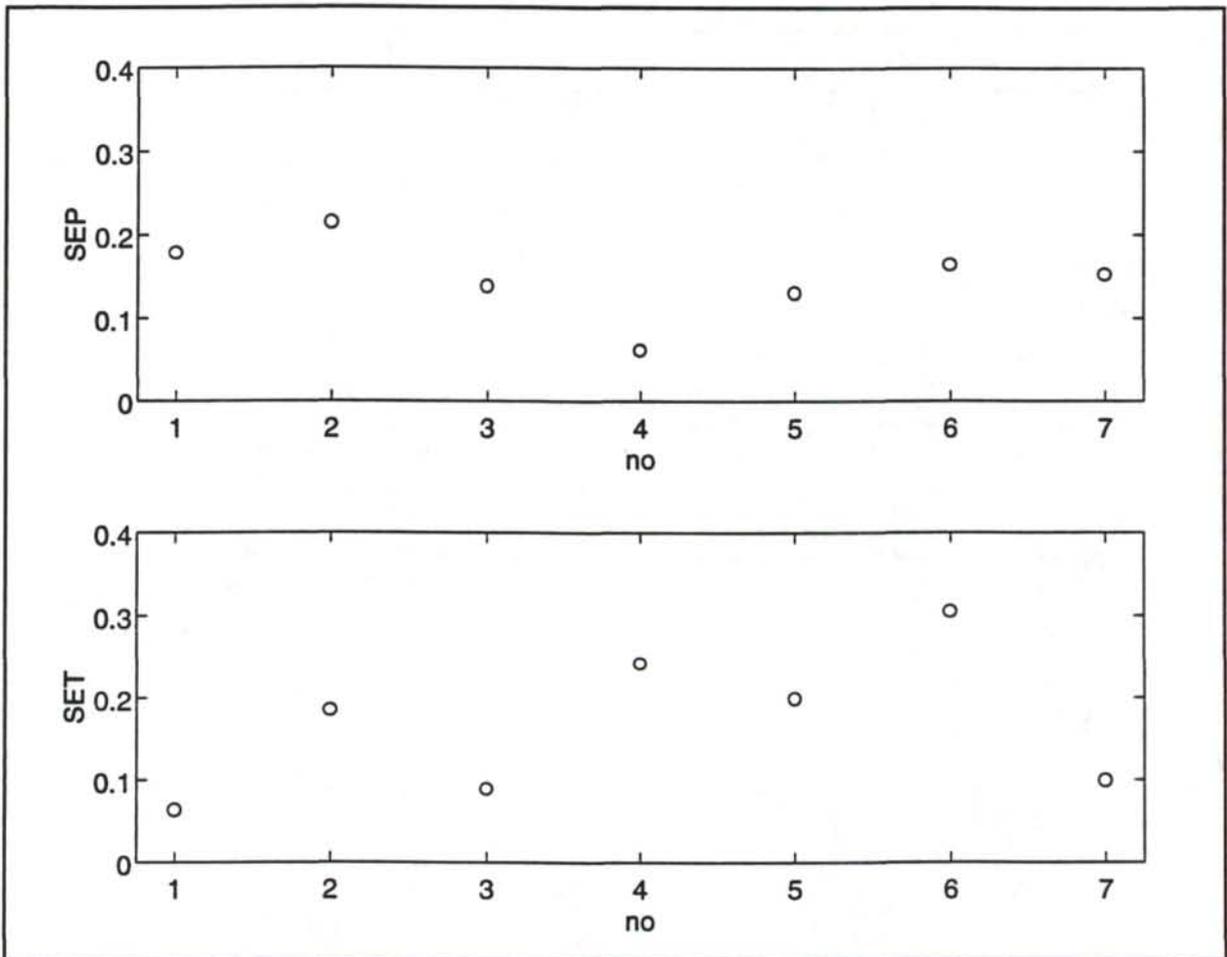


Figura III.5.17 Resultados de la transferencia de calibrado realizada con el procedimiento GCT para el cadmio.

aprendizaje son muy elevados. De hecho, el valor de SET correspondiente a este conjunto de transferencia en el procedimiento aplicado al plomo es el más elevado de todos.

Una situación parecida se tiene con el conjunto de transferencia 4 en el caso del cadmio, sólo que, la señal experimental correspondiente a este analito es bastante mayor, y por tanto, un valor de SET de esta magnitud no se traduce en un error relativo tan elevado como en el caso del plomo.

Tabla III.5.13 Errores relativos (%) de las muestras test del calibrado E, obtenidos al aplicar el procedimiento GCT con los conjuntos de transferencia 1, 4 y 7.

<i>no</i>	Número de muestra	Analito		
		Plomo	Estaño	Cadmio
1	10	-0.60	-0.47	0.65
	12	-2.22	4.17	0.97
	15	7.47	-0.96	0.55
	17	-7.55	0.11	1.53
	20	1.13	-2.34	-4.24
	22	3.14	-1.98	-1.65
4	10	-3.88	1.23	1.88
	12	-7.25	7.24	5.66
	15	1.12	-2.05	0.31
	17	-2.17	-0.83	0.50
	20	0.39	-1.71	-0.64
	22	-0.35	-1.58	-0.85
7	10	3.02	-0.07	0.02
	12	1.75	3.25	0.69
	15	2.10	-4.44	-0.18
	17	-3.63	-1.77	0.94
	20	4.37	-3.78	-3.22
	22	2.32	-3.65	-1.79

En principio, los resultados obtenidos con los conjuntos de transferencia 1 y 4 son bastante similares, excepto en lo que se refiere al plomo. No obstante, la diferencia más importante existente entre esos dos conjuntos de transferencia radica en el número de muestras que los componen. El primero de ellos está formado por 23 muestras, mientras que el segundo, $n_0 = 4$, está constituido únicamente por 7, razón por la cual, éste último parece el más indicado para ser utilizado en transferencias de calibrado que se realicen posteriormente mediante este procedimiento de transferencia.

Por otro lado, al comparar los resultados obtenidos con el procedimiento de transferencia GCT con los correspondientes al procedimiento PDS utilizado en primer lugar, cabe señalar que éstos son muy similares, tanto en lo que se refiere a los errores asociados con las muestras del conjunto test (que también se encuentran evaluados a través del SEP), como a las estimaciones realizadas para las muestras del conjunto de aprendizaje del calibrado G, las cuales vienen dadas por los valores de SET.

Por tanto, en principio no es posible decantarse por uno u otro procedimiento, ya que los resultados alcanzados por ambos son similares.

III.6 MODELADO DE LA LÍNEA BASE EN LA DETERMINACIÓN POR VOLTAMPEROMETRÍA DE REDISOLUCIÓN DE Pb Y Tl

El solapamiento de señales en aquellas determinaciones que se llevan a cabo utilizando técnicas electroquímicas con redisolución supone un obstáculo importante que dificulta el desarrollo de dichos análisis. Este problema se ve incrementado cuando se llevan a cabo determinaciones de impurezas en presencia de gran exceso del otro elemento electroactivo. Un ejemplo clásico de este tipo de falta de selectividad se encuentra en la determinación mediante voltamperometría diferencial de impulsos con redisolución anódica de talio (I) y plomo (II) en diversos medios y matrices [185-186].

Se han propuesto soluciones de tipo químico, matemático e instrumental, al problema originado por el solapamiento de picos en los análisis realizados con técnicas de redisolución. Esta interferencia en las determinaciones puede resolverse, tal y como se ha señalado anteriormente, introduciendo una etapa previa de separación antes de que se lleve a cabo la medida electroquímica. Procedimientos de separación, tales como coprecipitación [152] o extracción [155], han sido aplicados para la separación de estaño y plomo. No obstante, estos procedimientos, además de consumir tiempo, puede dar lugar a la contaminación de la muestra o a que se produzca una recuperación incompleta del metal. Para eliminar los errores relacionados con los procedimientos de extracción, Karbainov y Stromberg [187] proponen la determinación directa de plomo, zinc, cobre, talio y cadmio en la fase orgánica después de su extracción con una disolución de 8-hidroxiquinona en cloroformo.

Una separación por cambio iónico acoplada con la monitorización en línea del efluente de la columna, mediante una técnica de redisolución, ha sido aplicada con éxito en la separación de selenio de otros compuestos metálicos presentes en muestras reales y que actúan como interferentes [188].

En muchos casos, el uso de un electrólito soporte con un agente complejante adecuado puede provocar un desplazamiento del potencial de pico que permita resolver el solapamiento

de las señales. Este procedimiento requiere una cuidadosa elección del ligando en cuestión, así como de la concentración en que ha de encontrarse el mismo, ya que no siempre es posible desplazar el potencial de pico de uno de los metales sin ejercer algún efecto sobre la señal correspondiente al otros. Aunque, también existe la posibilidad de utilizar un agente complejante que actúe como enmascarante selectivo del metal que actúa como interferente. Este es el papel del ácido etilendiaminotetraacético (AEDT) cuando se utiliza en la determinación mediante voltamperometría diferencial de impulsos con redisolución anódica de talio en presencia de plomo y cadmio [189-191].

Otras alternativas para resolver este problema se encuentran en el uso de técnicas electroquímicas con cambio de medio [192, 193] o en la adición de agentes tensoactivos [194]. Incluso, la adición de un tercer elemento que forme, sobre el electrodo de mercurio, un compuesto intermetálico estable con uno de los metales interferentes ha sido aplicada a problemas de solapamiento de picos. Un ejemplo de este último método se encuentra en la adición de cobre para minimizar la interferencia debida a la presencia de indio y cadmio en la determinación de plomo y talio [195].

Garai y otros [196] han llevado a cabo la resolución de picos solapados en análisis por potenciometría diferencial con redisolución anódica (DPSA) considerando las diferencias existentes entre los potenciales de semionda, el número de electrones implicados en la reacción electródica y la relación de concentraciones de los iones que están siendo determinados de forma simultánea.

Collado y otros [197] han desarrollado un procedimiento para eliminar o disminuir el solapamiento de picos en la determinación de cobre, plomo y cadmio por voltamperometría adsorbtiva de redisolución catódica y barridos de potencial rápidos. El método consiste en la substracción de barridos repetitivos en los cuales alguno de los picos se ha eliminado modificando adecuadamente los potenciales y tiempos de acumulación y reoxidación.

La utilización de técnicas de voltamperometría anódica con acumulación ofrece también la posibilidad de obviar el problema desplazando el potencial de acumulación hasta que el ion metálico con potencial de pico más negativo no se acumule sobre el electrodo. Este procedimiento tiene el inconveniente de que, en esas condiciones, ese último analito no puede ser determinado.

Como ya se señaló en el capítulo anterior, existe una gran variedad de referencias bibliográficas relacionadas con metodologías no experimentales utilizadas en la resolución de

problemas originados por la presencia de señales solapadas. Algunas de las fuentes citadas incluso hacían referencia concreta a la determinación electroquímica de plomo y talio [172, 180, 181, 184]. Henrion y otros [33] utilizan la regresión multivariante PLS para resolver el problema de solapamiento de estos dos analitos. Ortiz y otros [198], en cambio, abordan el problema desde un punto de vista más amplio, aplicando la metodología de la regresión continua [6]. Esta técnica de regresión ofrece la posibilidad de seleccionar en un espacio continuo de modelos multivariantes aquellos que conducen a las mejores predicciones, que no siempre han de coincidir con las correspondientes al modelo PLS.

En el presente capítulo se ha optado por utilizar la regresión PLS en la determinación multivariante de plomo y talio por voltamperometría diferencial de impulsos con redisolución anódica, aplicándose posteriormente un procedimiento de transferencia del calibrado multivariante a lo largo del tiempo.

III.6.1 Experimental

Reactivos

Todos los reactivos utilizados han sido de calidad para análisis y no se han realizado purificaciones posteriores sobre ellos. Las disoluciones patrón fueron preparadas usando en todos los casos agua desionizada proporcionada por un sistema Barnstead NANO Pure II. El gas inerte utilizado para eliminar el oxígeno disuelto en las medidas voltamperométricas fue nitrógeno (99.997%). Como electrólito soporte se utilizó una disolución constituida por $\text{H}_2\text{C}_2\text{O}_4$ 0.1 M y HCl 0.1 M, que presentaba un $\text{pH} = 1.037$ y actuaba al mismo tiempo como disolución tampón.

Instrumental

Las medidas voltamperométricas fueron realizadas utilizando un procesador Metrohm 646 VA junto con una unidad 647 VA equipada con un electrodo multimodo Metrohm (MME) que actuó en el modo gota de mercurio colgante (HMDE). El sistema de tres electrodos se completó con un electrodo de referencia Ag/AgCl/KCl (3M) y un electrodo auxiliar de platino. La técnica voltamperométrica empleada en este análisis ha sido la voltamperometría diferencial de impulsos con redisolución anódica (DPASV).

Los análisis de datos se llevaron a cabo utilizando los programas PARVUS [57], STATGRAPHICS [58] y MATLAB [59].

Procedimiento experimental

La disolución a analizar se introdujo en una celda polarográfica, burbujeándose nitrógeno a través de ella durante 10 min con el fin de eliminar el oxígeno disuelto en la misma. A continuación, se aplicó un potencial de deposición de -600 mV durante un período de 30 s, tras lo cual se desconectó la agitación y, transcurrido un tiempo de equilibrio de 20 s, se realizó un barrido de potencial desde -558 mV hasta -282 mV, registrándose la intensidad a intervalos de 6 mV. En aquellos casos en los cuales se llevaron a cabo adiciones, la disolución fue desoxigenada durante 15 s antes de realizar una nueva medida.

Los parámetros instrumentales utilizados en todos los análisis realizados fueron los siguientes: velocidad de barrido, 10 mV s⁻¹; amplitud del impulso, 50 mV; duración del impulso, 0.6 s; velocidad de agitación durante el período de acumulación, 1290 rev s⁻¹; área nominal del electrodo, 0,40 mm².

III.6.2 Diseño del experimento

Se ha señalado anteriormente que la dificultad que entraña normalmente la determinación de talio y plomo mediante técnicas electroquímicas se encuentra en el alto grado de solapamiento que presentan las señales electroquímicas correspondientes a estos dos metales. Esta interferencia es tan importante en el análisis que, cuando en una etapa del mismo se ha utilizado un diseño experimental próximo a un diseño central compuesto, los resultados han indicado que la información aportada por el mismo para la construcción de un modelo de calibrado PLS era insuficiente. Por tanto, se ha optado por usar un diseño experimental que cubra más ampliamente y de forma exhaustiva el dominio experimental, para que de esta forma sea posible recoger toda la información necesaria para construir modelos de calibrado PLS adecuados.

El diseño experimental seleccionado en este caso consiste en una serie de adiciones equimoleculares de plomo y talio sobre una disolución inicial de electrólito soporte, tal y

como se indica en la Figura III.6.1. Siguiendo ese diseño experimental se han llevado a cabo dos calibrados completos en diferentes días, el calibrado H, realizado en primer lugar, y el calibrado I, obtenido días después. Ello permite disponer de dos conjuntos de datos obtenidos con el mismo instrumento, pero en situaciones diferentes, entre los cuales poder establecer una relación mediante un procedimiento de transferencia de calibrado.

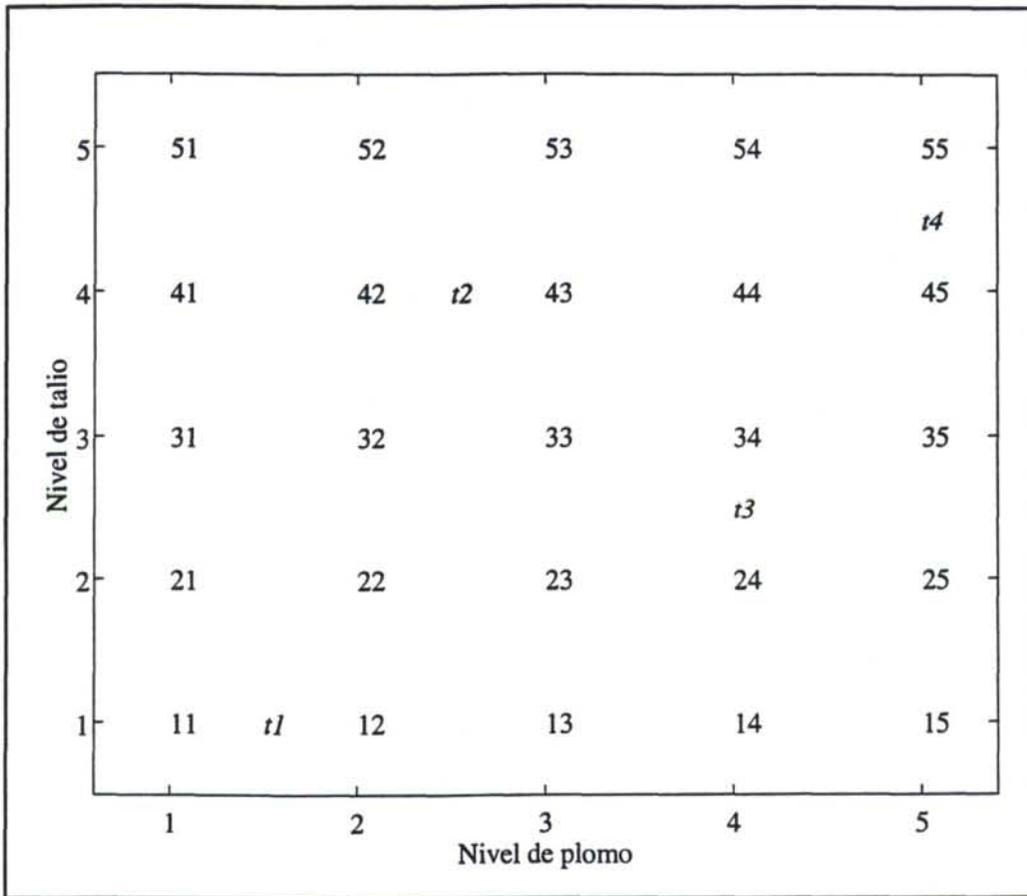


Figura III.6.1 Distribución de muestras de los calibrados H e I. El primer dígito indica el nivel de talio y el segundo el de plomo. Muestras test: t1 a t4.

Además de las 25 muestras de que consta el diseño 5 x 5 indicado, se analizaron en ambos calibrados otras cuatro muestras que van a ser consideradas como muestras test, muestras t1 a t4 en la Figura III.6.1, cuyas concentraciones respectivas se muestran en la Tabla III.6.1 junto con las del resto del diseño.

Los voltamperogramas correspondientes a ambos calibrados se muestran en las Figuras III.6.2a-e y III.6.3a-e, apreciándose en ellos la existencia de dos picos voltamperométricos muy solapados, que ya han sido resueltos con anterioridad mediante técnicas multivariantes

Tabla III.6.1 Diseño experimental de los calibrados H e I. Cada símbolo de la segunda columna indica que se realizaron adiciones sobre una misma disolución inicial.

Muestra		Diseño experimental		Concentraciones ($M \times 10^7$)	
		Talio	Plomo	Talio	Plomo
1	×	1	1	3.2659	2.2909
<i>t1</i>	×	1	1.5	3.2594	3.4295
2	×	1	2	3.2530	4.5636
3	×	1	3	3.2402	6.8185
4	×	1	4	3.2275	9.0557
5	×	1	5	3.2148	11.2754
6	×	2	5	6.4047	11.2315
7	▲	2	1	6.5059	2.2818
8	▲	2	2	6.4803	4.5457
9	▲	2	3	6.4549	6.7918
10	▲	2	4	6.4297	9.0203
<i>t3</i>	▲	2.5	4	8.0214	9.0027
11	○	3	1	9.7205	2.2728
12	○	3	2	9.6824	4.5278
13	○	3	3	9.6445	6.7652
14	○	3	4	9.6070	8.9852
15	○	3	5	9.5698	11.1880
16	■	4	1	12.9098	2.2639
17	■	4	2	12.8594	4.5102
<i>t2</i>	■	4	2.5	12.8343	5.6267
18	■	4	3	12.8093	6.7389
19	■	4	4	12.7597	8.9504
20	■	4	5	12.7104	11.1448
<i>t4</i>	■	4.5	5	14.2717	11.1233
21	•	5	1	16.0742	2.2551
22	•	5	2	16.0117	4.4926
23	•	5	3	15.9496	6.7128
24	•	5	4	15.8880	8.9158
25	•	5	5	15.8269	11.1019

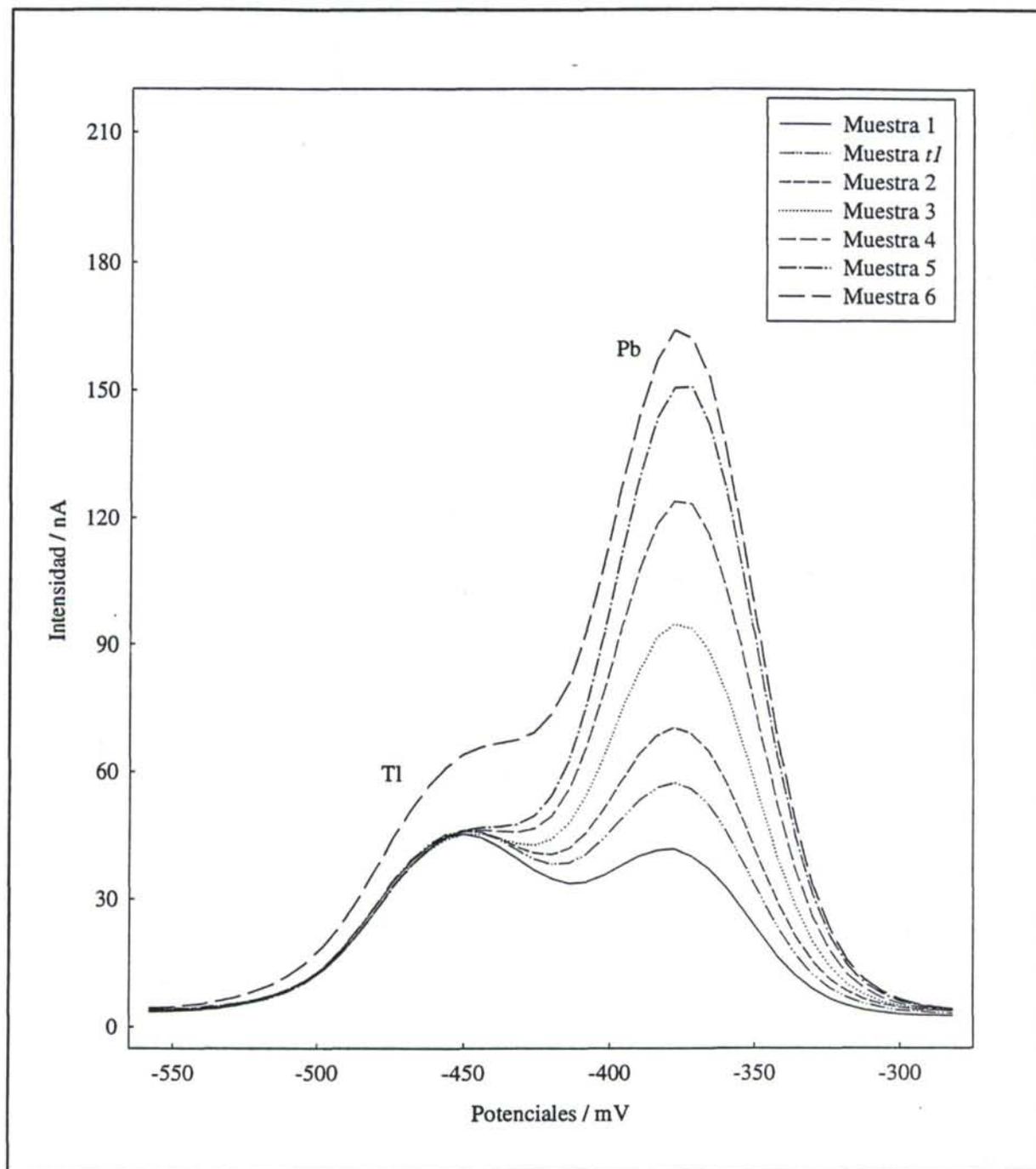


Figura III.6.2a Voltamperogramas correspondientes al calibrado H, obtenidos con una misma disolución inicial.

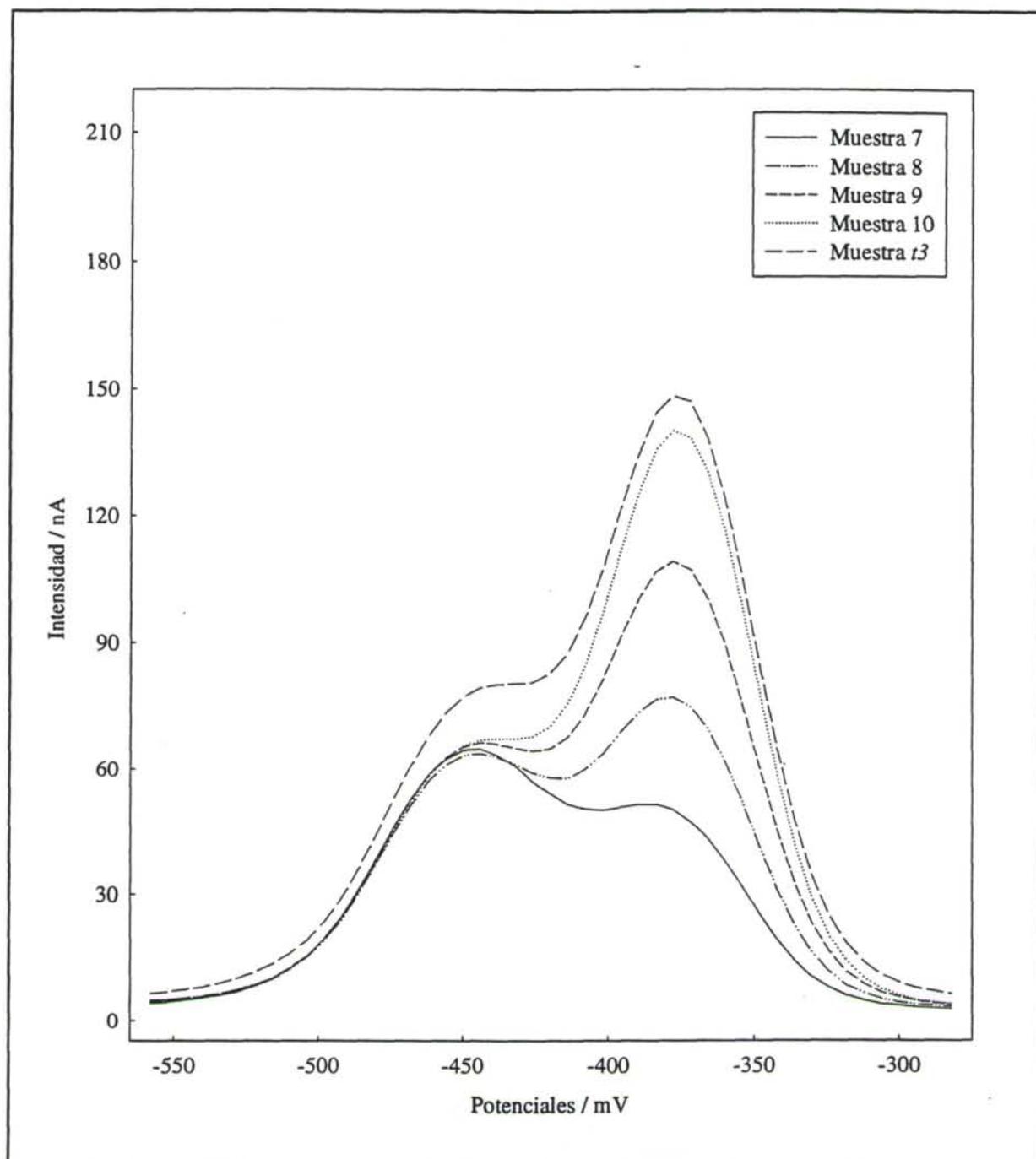


Figura III.6.2b Voltamperogramas correspondientes al calibrado H, obtenidos con una misma disolución inicial.

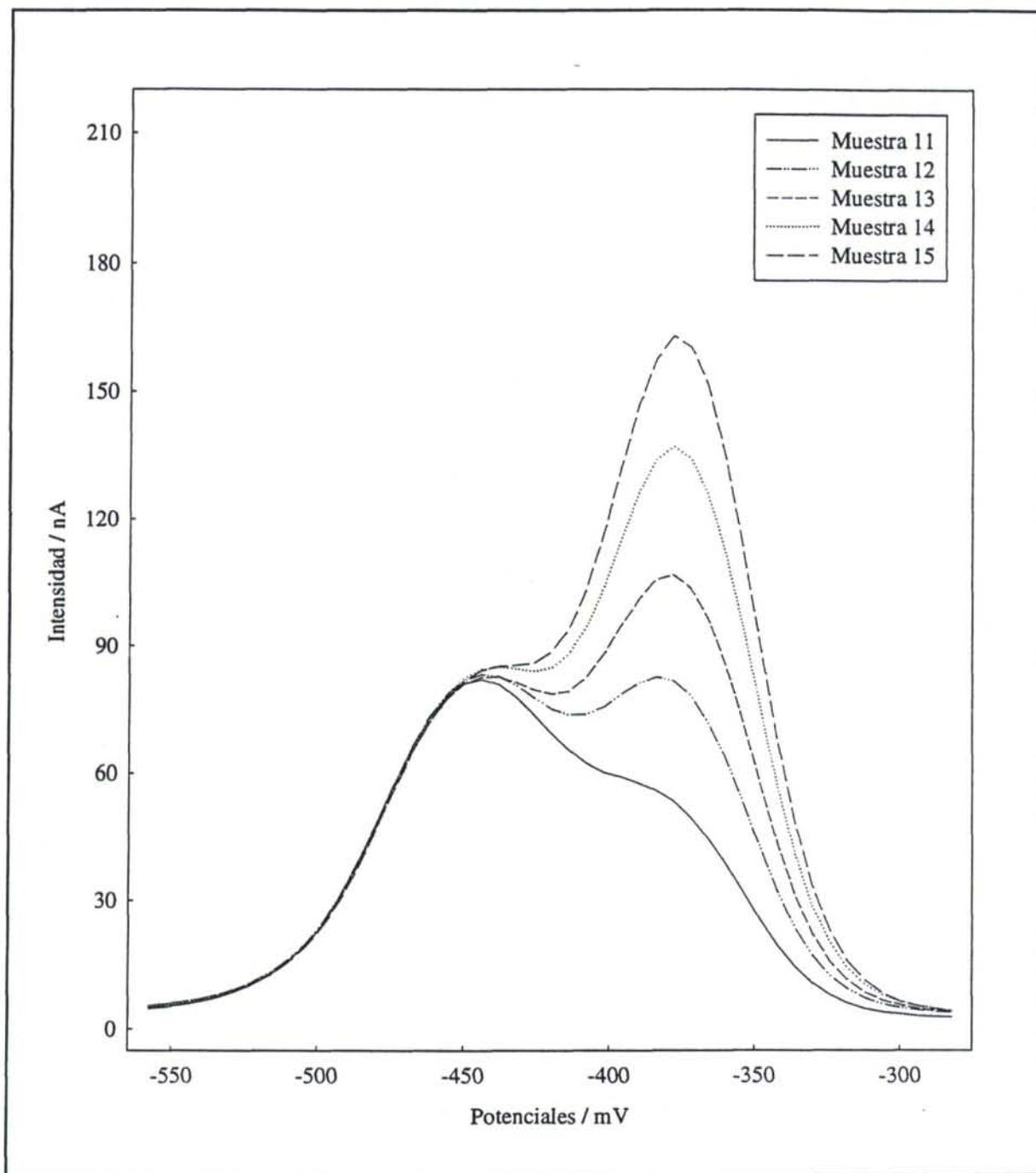


Figura III.6.2c Voltamperogramas correspondientes al calibrado H, obtenidos con una misma disolución inicial.

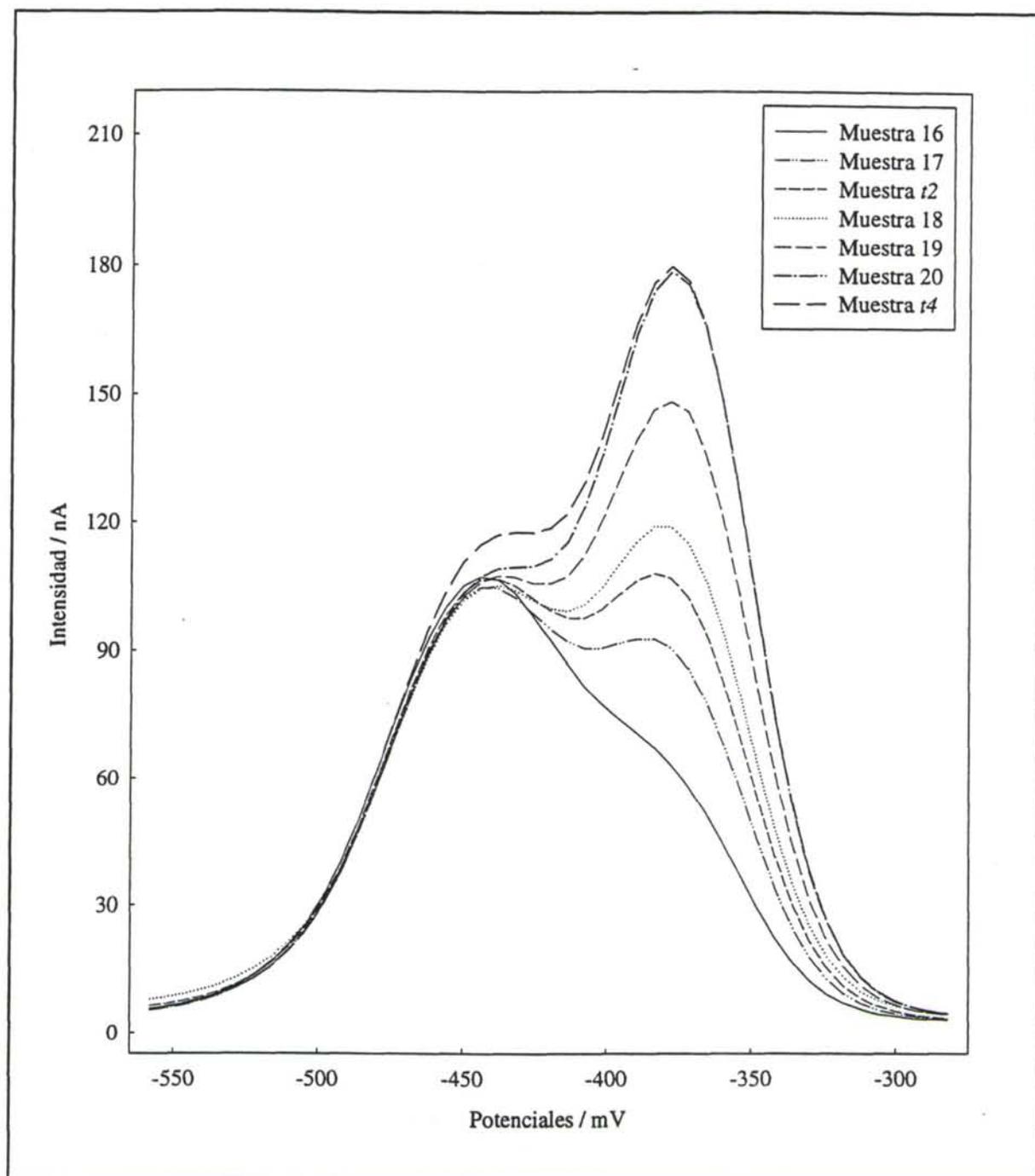


Figura III.6.2d Voltamperogramas correspondientes al calibrado H, obtenidos con una misma disolución inicial.

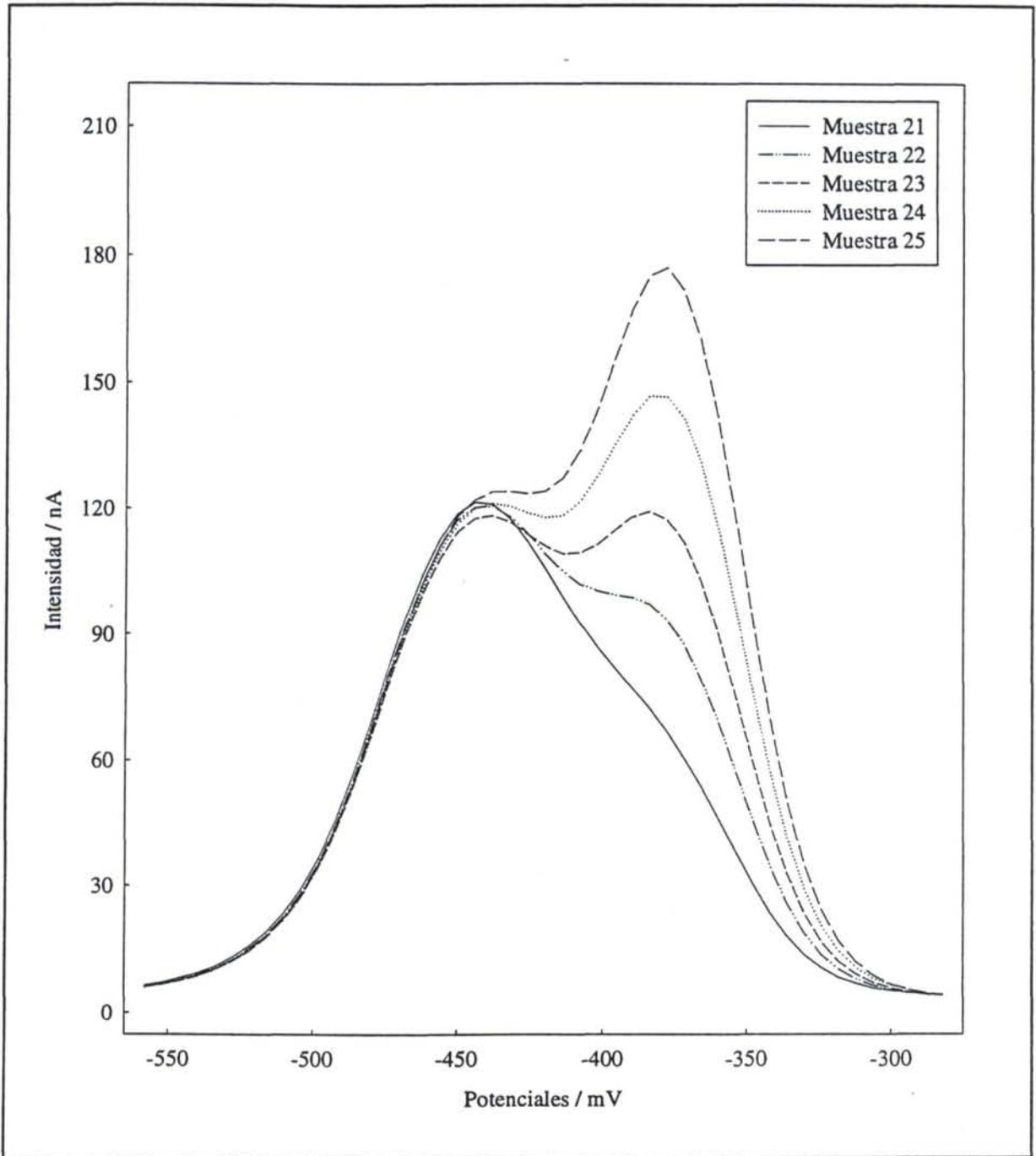


Figura III.6.2e Voltamperogramas correspondientes al calibrado H, obtenidos con una misma disolución inicial.

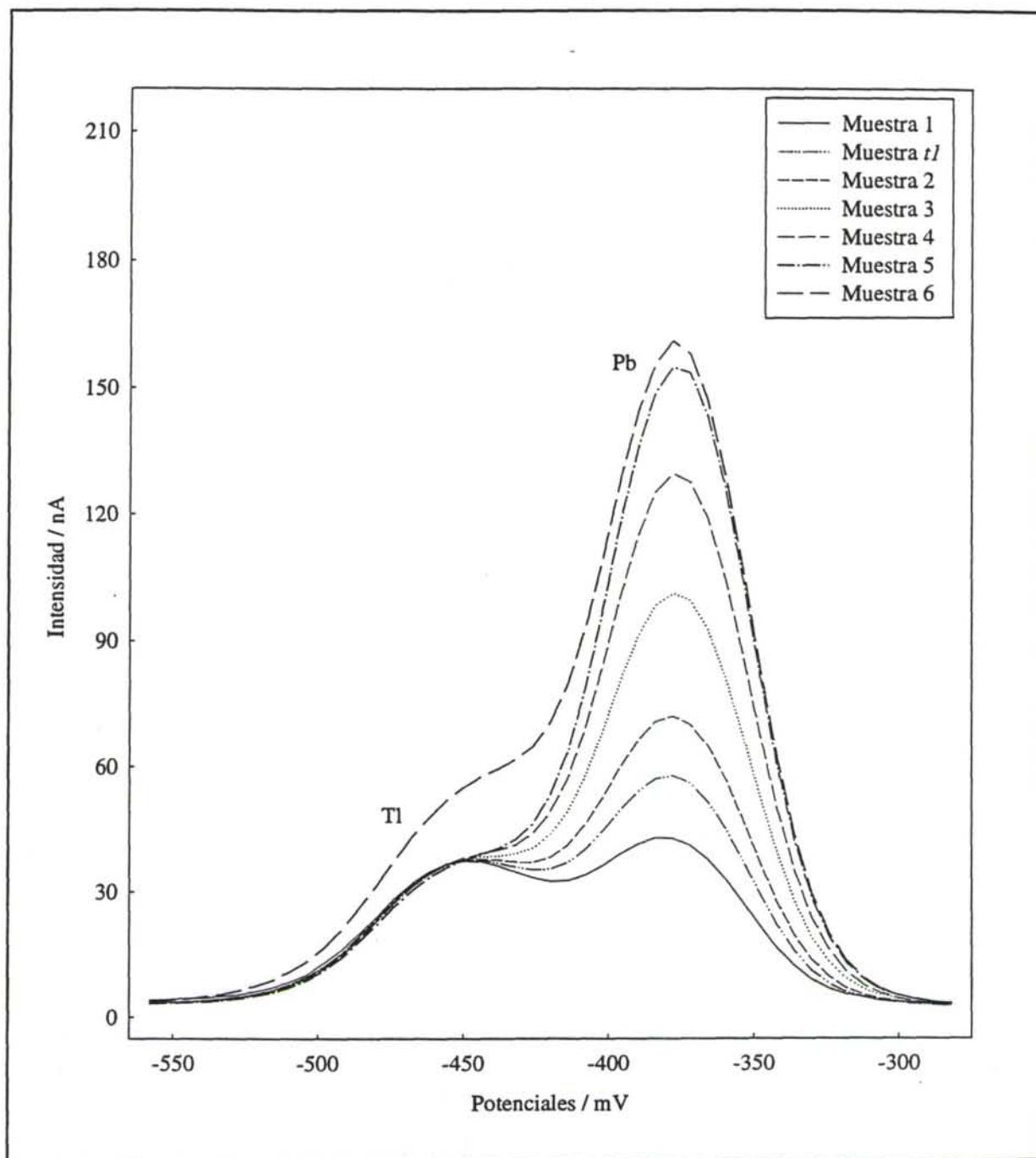


Figura III.6.3a Voltamperogramas correspondientes al calibrado I, obtenidos con una misma disolución inicial.

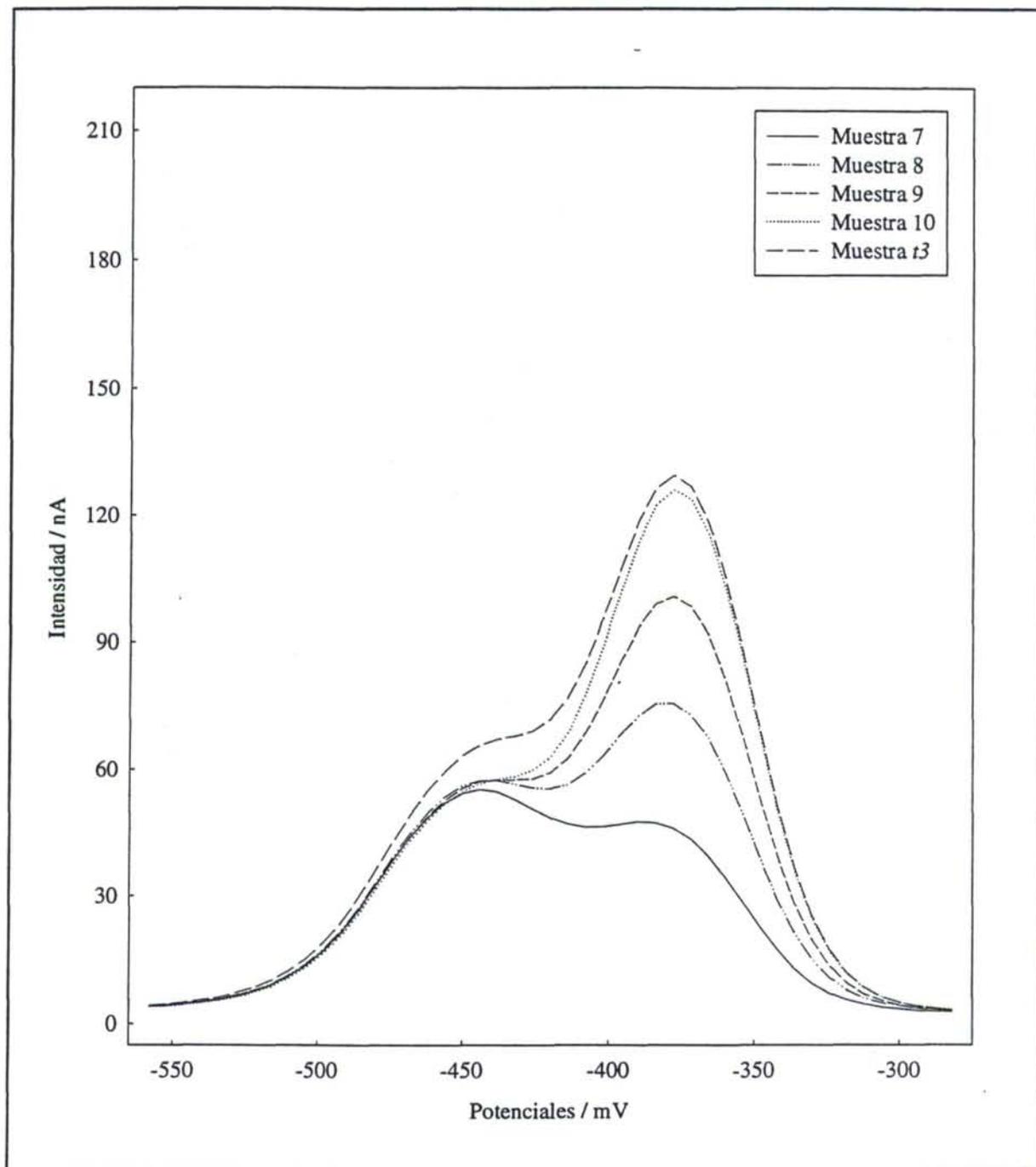


Figura III.6.3b Voltamperogramas correspondientes al calibrado I, obtenidos con una misma disolución inicial.

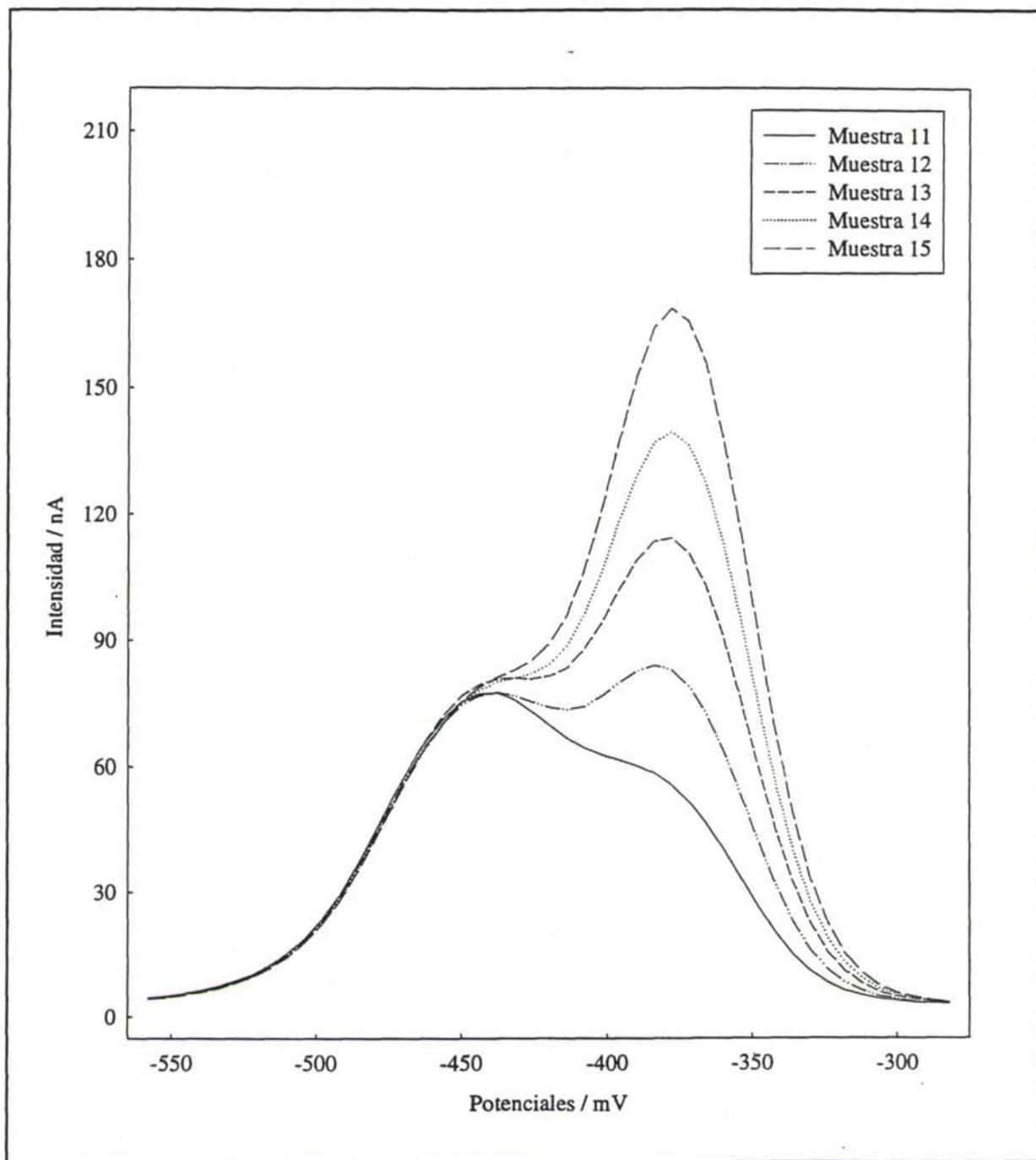


Figura III.6.3c Voltamperogramas correspondientes al calibrado I, obtenidos con una misma disolución inicial.

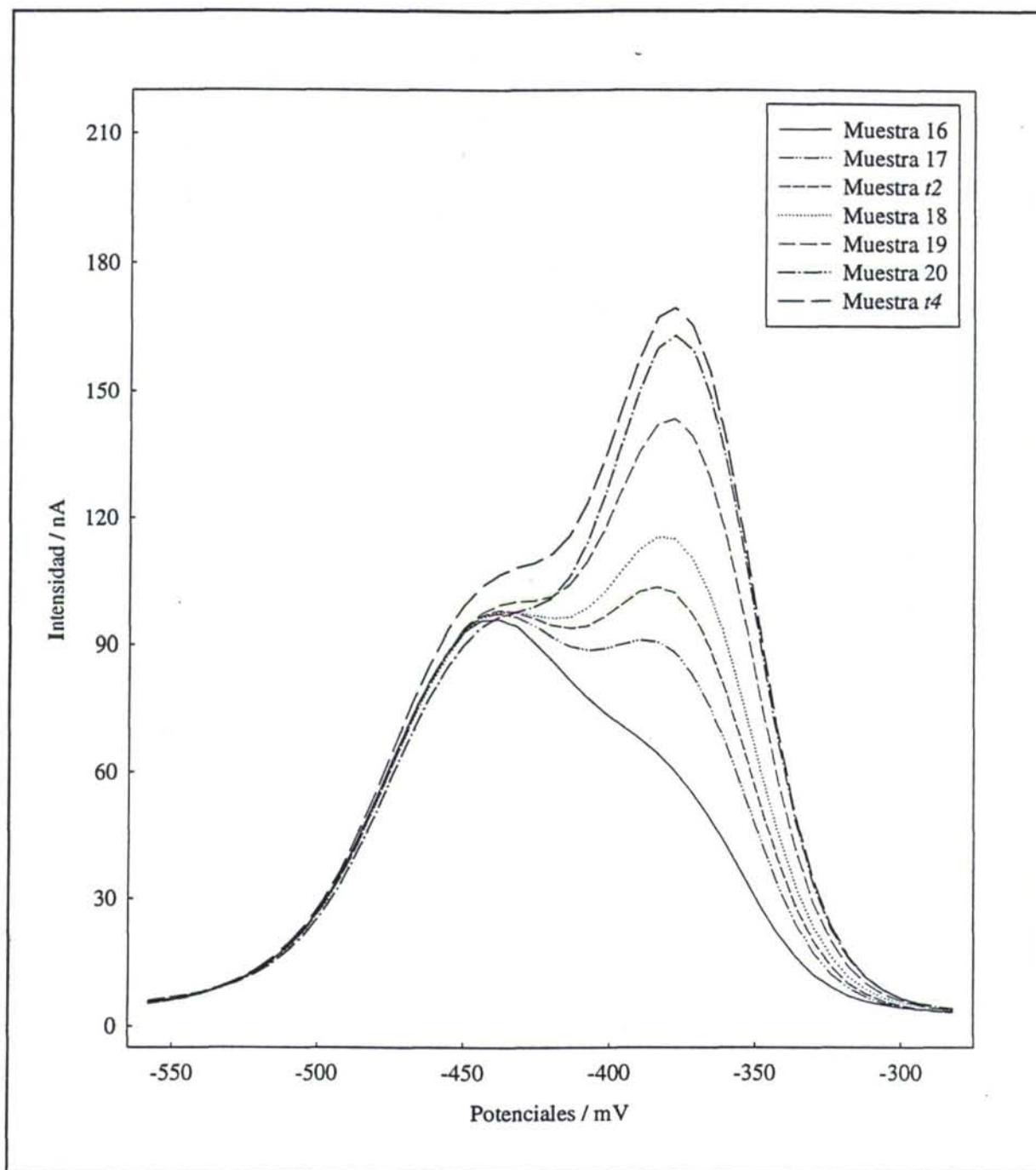


Figura III.6.3d Voltamperogramas correspondientes al calibrado I, obtenidos con una misma disolución inicial.

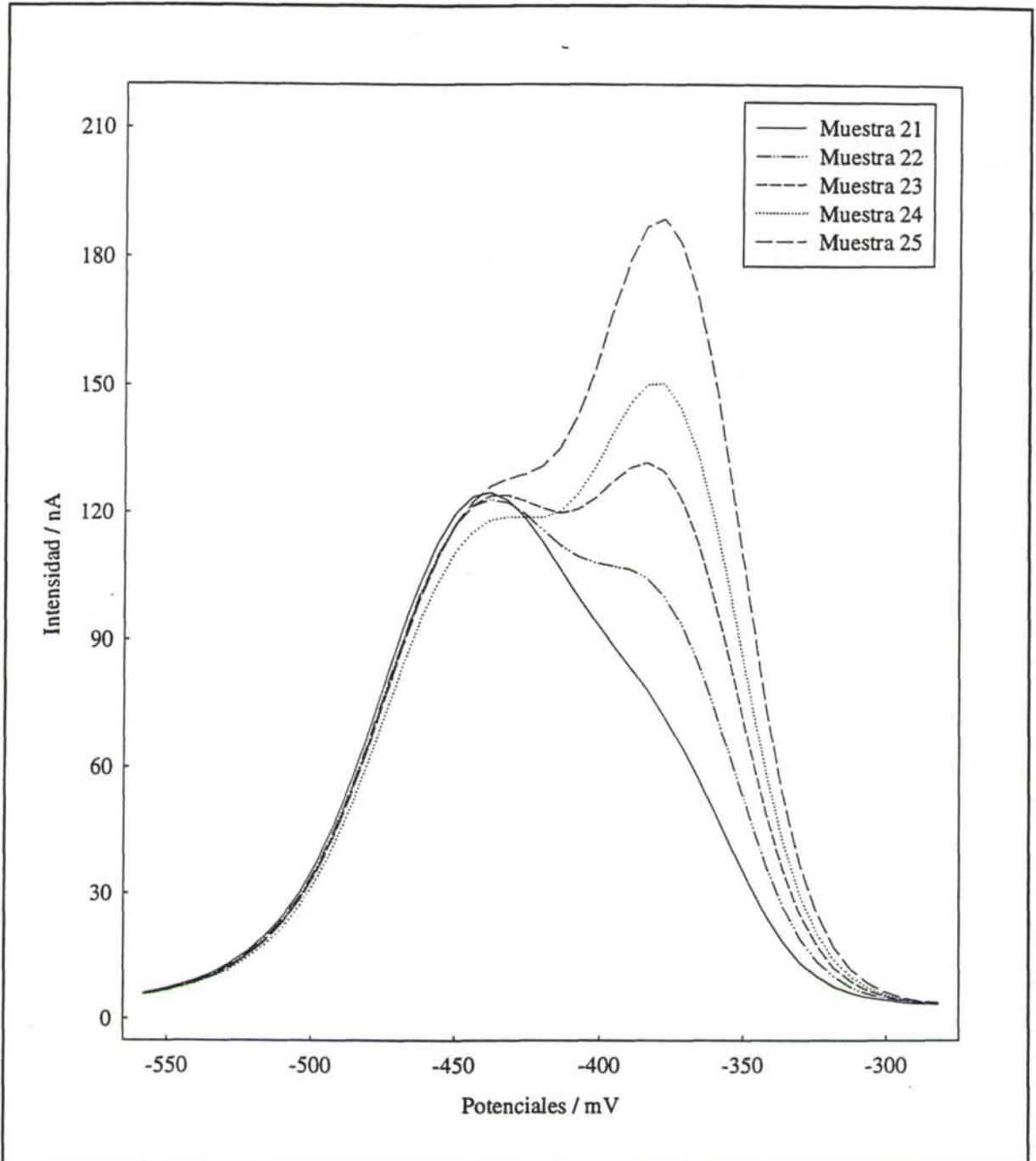


Figura III.6.3e Voltamperogramas correspondientes al calibrado I, obtenidos con una misma disolución inicial.

de regresión [33, 198]. Sin embargo, en el presente análisis se partió de un electrólito soporte que debido posiblemente a la existencia de alguna impureza presentaba, en las condiciones en las que se llevó a cabo el análisis, una señal voltamperométrica en la misma zona en la que aparecían los picos de los analitos de interés, como se puede apreciar en la Figura III.6.4, interfiriendo notablemente en la señal registrada.

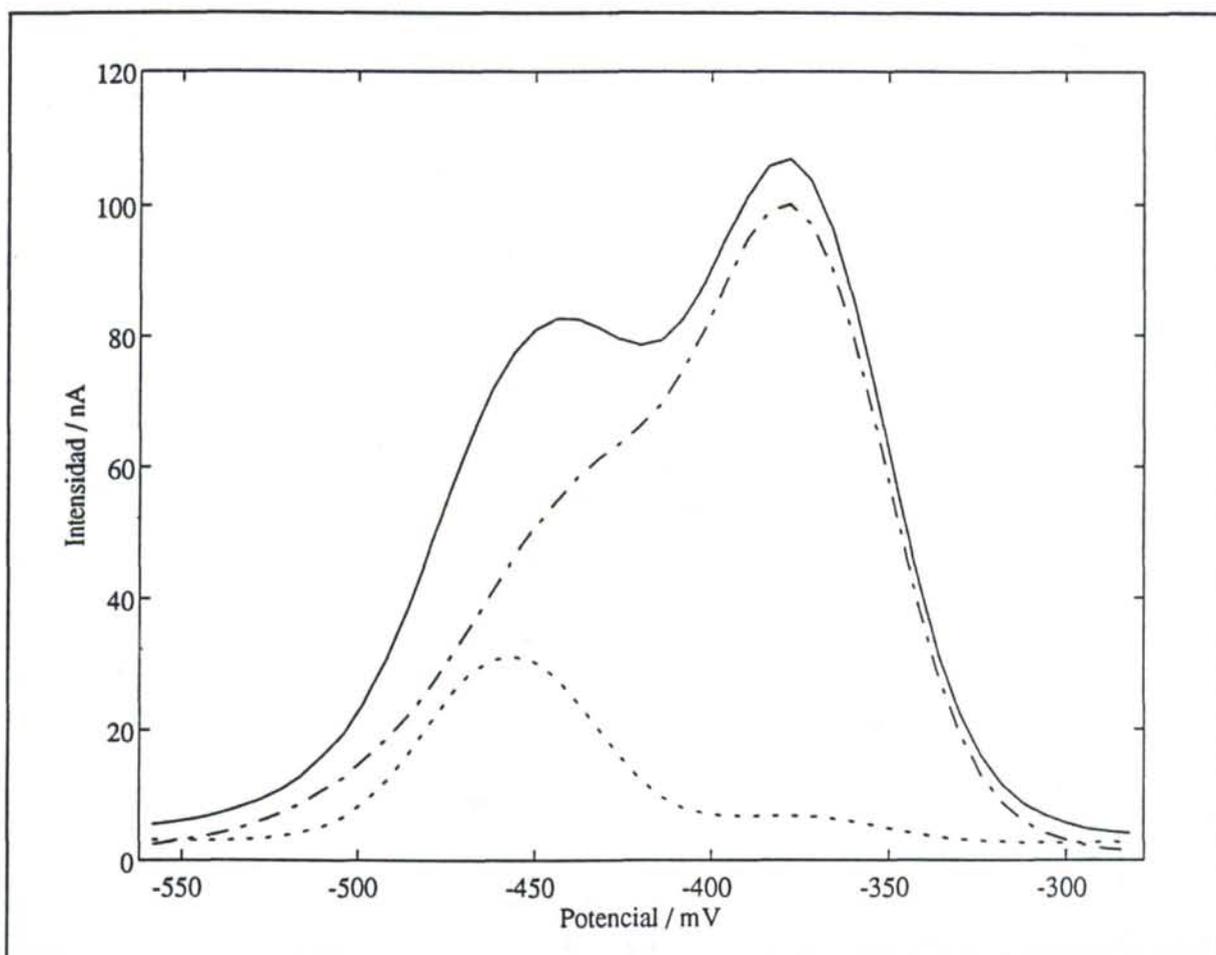


Figura III.6.4 Voltamperogramas correspondientes al blanco (- -) y a la muestra 33 del calibrado H restada de ésta la señal del blanco (- · -) y sin restar (—).

En dicha figura, se muestran los voltamperogramas correspondientes al blanco experimental y a una de las muestras del calibrado H, en trazo punteado y continuo respectivamente, así como el voltamperograma obtenido al restar la señal del blanco de la señal total, línea (- · -). Es posible observar cómo la señal debida al blanco provoca un incremento importante de la intensidad registrada en el voltamperograma original (—) respecto de la señal neta (- · -), obtenida cuando se ha restado la señal del electrólito soporte.

Al mismo tiempo se registra un desplazamiento respecto al eje de potenciales del máximo de pico de la señal correspondiente al talio. El efecto que sobre las señales originales tiene el blanco experimental resulta evidente cuando se comparan las Figuras III.6.2 y III.6.5, ya que en esta última se muestran los voltamperogramas obtenidos al restar a cada una de las muestras del calibrado H el blanco experimental correspondiente.

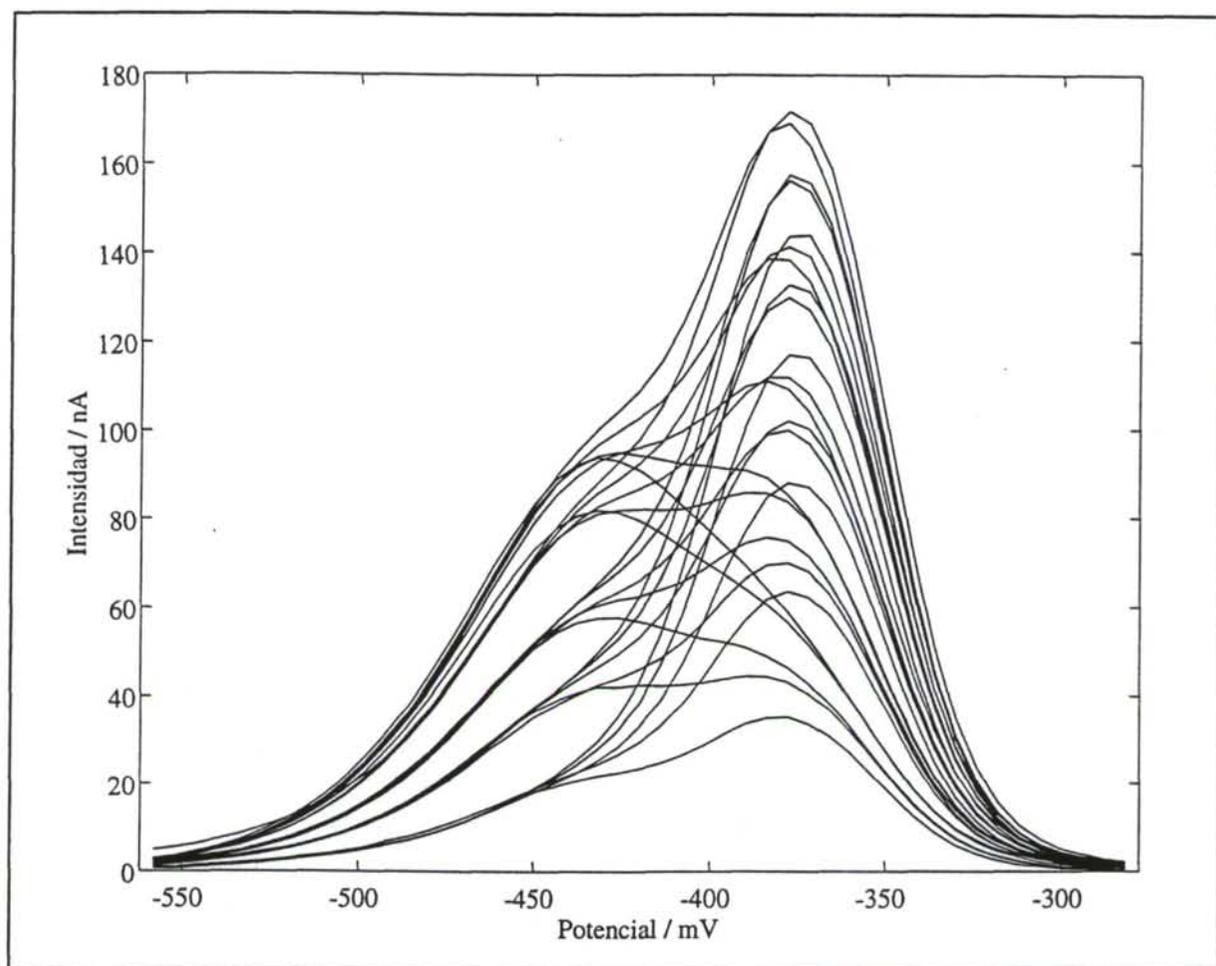


Figura III.6.5 Voltamperogramas obtenidos al restar a las señales originales del calibrado H la señal correspondiente al blanco experimental.

A diferencia de lo que suele ser el procedimiento habitual, en este análisis no se utilizaron los voltamperogramas a los que se había restado la señal del blanco experimental, es decir, la señal correspondiente al electrólito soporte, sino que se optó por utilizar las señales voltamperométricas originales y comprobar si los métodos multivariantes de regresión utilizados eran capaces de modelar esta interferencia.

III.6.3 Análisis multivariante

III.6.3.1 Modelos PLS

Construcción de los modelos PLS

Se han desarrollado modelos de regresión PLS tanto para el talio como para el plomo en cada uno de los dos calibrados que se han llevado a cabo, calibrados H e I. Al igual que en los capítulos precedentes, estos modelos de regresión se han construido de forma independiente para cada uno de estos dos analitos, tomándose como variables predictoras las intensidades registradas a 47 valores de potencial igualmente espaciados entre -558 mV y -282 mV. La variable respuesta viene dada en cada caso por la concentración de uno de los analitos, correspondiéndose la misma con los valores que se encuentran en la Tabla III.6.1, en la cual se puede apreciar que el diseño experimental utilizado recorre 5 niveles de concentración diferentes para cada analito.

Las matrices de datos experimentales, matrices **H** e **I** de dimensiones 29×47 , contienen la información relativa a los voltamperogramas originales, es decir, a las señales originales de las cuales no se ha restado la señal correspondiente al blanco experimental. La matriz de las variables respuesta tiene de dimensiones 29×2 , siendo 2 el número de analitos a determinar en el análisis.

Los modelos PLS se han construido utilizando el método de validación cruzada [22,23] con objeto de evaluar la estabilidad y la capacidad de predicción de los mismos. Con el mismo fin se ha dispuesto de las cuatro muestras test, $t1$ a $t4$, que serán utilizadas como conjunto de evaluación a lo largo de todo el análisis.

Evaluación de los modelos PLS

Una primera aproximación a los modelos PLS construidos para los calibrado H e I la proporcionan los porcentajes de varianza explicada y de varianza en predicción relacionados con dichos modelos. Estos valores se encuentran en las Tablas III.6.2 y III.6.3, para el calibrado H y el calibrado I respectivamente, y se puede observar que, a pesar del alto grado de solapamiento que presentaban las dos señales voltamperométricas, los porcentajes de varianza explicada son bastante altos. Incluso son un indicio claro de que se ha superado la interferencia debida a la señal correspondiente al blanco experimental, que no era en absoluto

despreciable si se tiene en cuenta la relación entre ésta y las señales correspondientes a los analitos de interés, sobre todo en lo que se refiere al talio, con cuyo pico voltamperométrico la señal del blanco se solapa completamente.

Tabla III.6.2 Porcentajes de varianza explicada en función del número de variables latentes (V. L.) seleccionadas en cada uno de los modelos PLS construidos en el calibrado H.

V. L.	Talio			Plomo		
	Bloque Y		Bloque X	Bloque Y		Bloque X
	Var. explicada	Var. en predicción	Var. explicada	Var. explicada	Var. en predicción	Var. explicada
1	47.15	48.93	81.00	54.43	55.71	81.00
2	94.31	94.62	98.10	99.28	99.32	98.10
3	99.68	99.65	99.33	99.36	99.40	99.27
4	99.71	99.67	99.52	99.58	99.57	99.57
5	99.72	99.64	99.72	99.58	99.52	99.75
6	99.72	99.53	99.86	99.59	99.49	99.86

Tabla III.6.3 Porcentajes de varianza explicada en función del número de variables latentes (V. L.) seleccionadas en cada uno de los modelos PLS construidos en el calibrado I.

V. L.	Talio			Plomo		
	Bloque Y		Bloque X	Bloque Y		Bloque X
	Var. explicada	Var. en predicción	Var. explicada	Var. explicada	Var. en predicción	Var. explicada
1	50.99	52.75	80.64	51.98	53.23	80.64
2	97.17	97.37	98.41	99.37	99.35	98.41
3	99.34	99.22	99.24	99.43	99.38	99.16
4	99.58	99.52	99.76	99.46	99.36	99.76
5	99.74	99.51	99.80	99.71	99.23	99.78
6	99.78	99.51	99.86	99.91	99.21	99.78

Los scores correspondientes a los modelos PLS construidos con los datos del calibrado H y representados en las Figuras III.6.6 y III.6.7 para talio y plomo respectivamente, son muy similares a aquellos obtenidos para el calibrado I. En ellos se puede observar cómo las tres

primeras variables latentes están relacionadas en gran medida con la concentración de ambos analitos, como se deduce también del elevado porcentaje de varianza del bloque Y que explican. La primera variable latente, como ocurre habitualmente cuando se analizan datos que no han sido normalizados previamente, está relacionada en gran medida con las direcciones de variabilidad de las propias variables predictoras, es decir, el denominado factor tamaño, hecho que también señalan los valores de varianza explicada por esta primera variable latente en el bloque X.

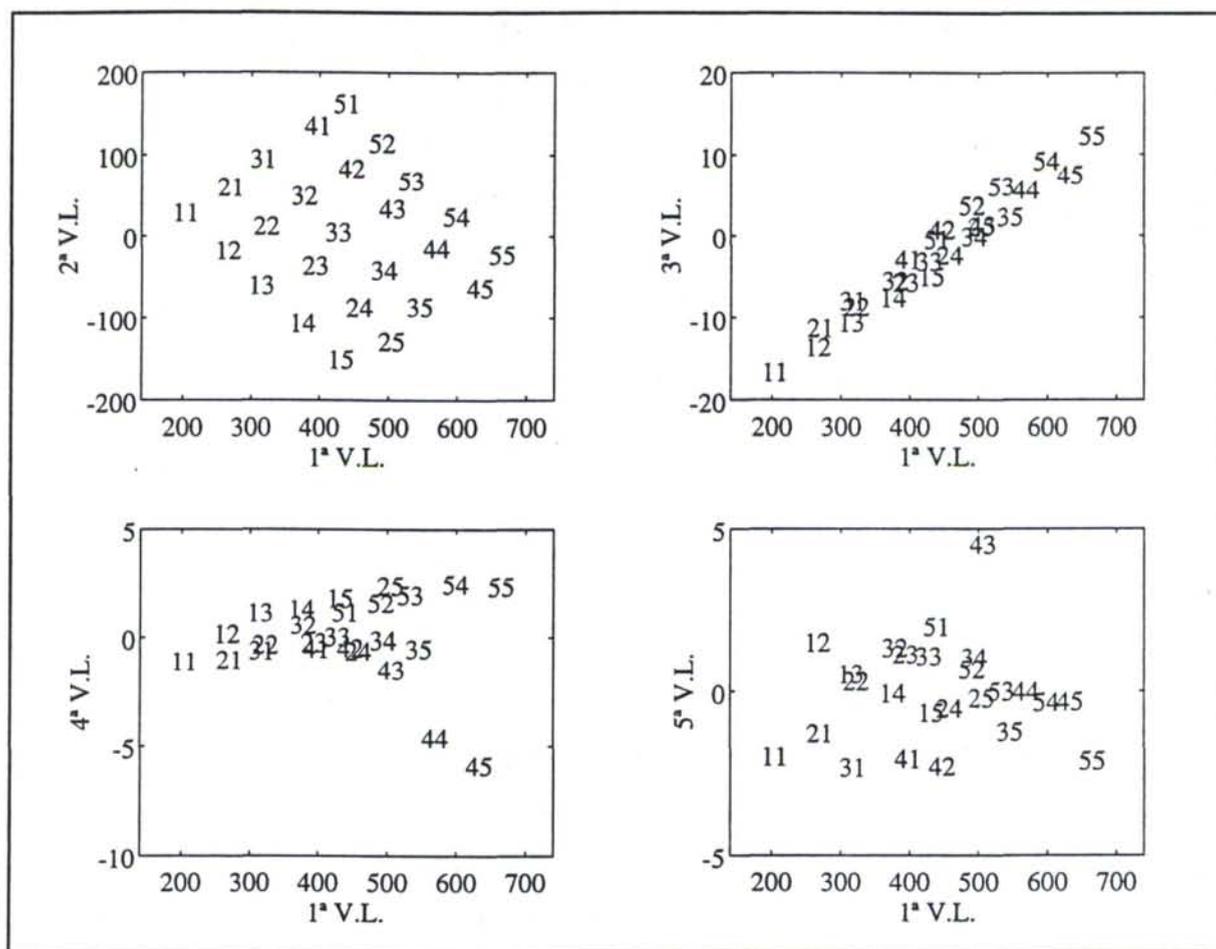


Figura III.6.6 Scores del modelo PLS construido para la determinación de talio en el calibrado H.

Sin embargo, la segunda variable latente recoge información relacionada con la concentración del analito de interés en cada caso y corrige de forma adecuada la interferencia debida al solapamiento de las señales en el modelo construido por la primera variable latente. En el plano formado por las proyecciones de los scores calculados con ambas variables latentes las muestras de calibrado aparecen distribuidas según el diseño experimental utilizado.

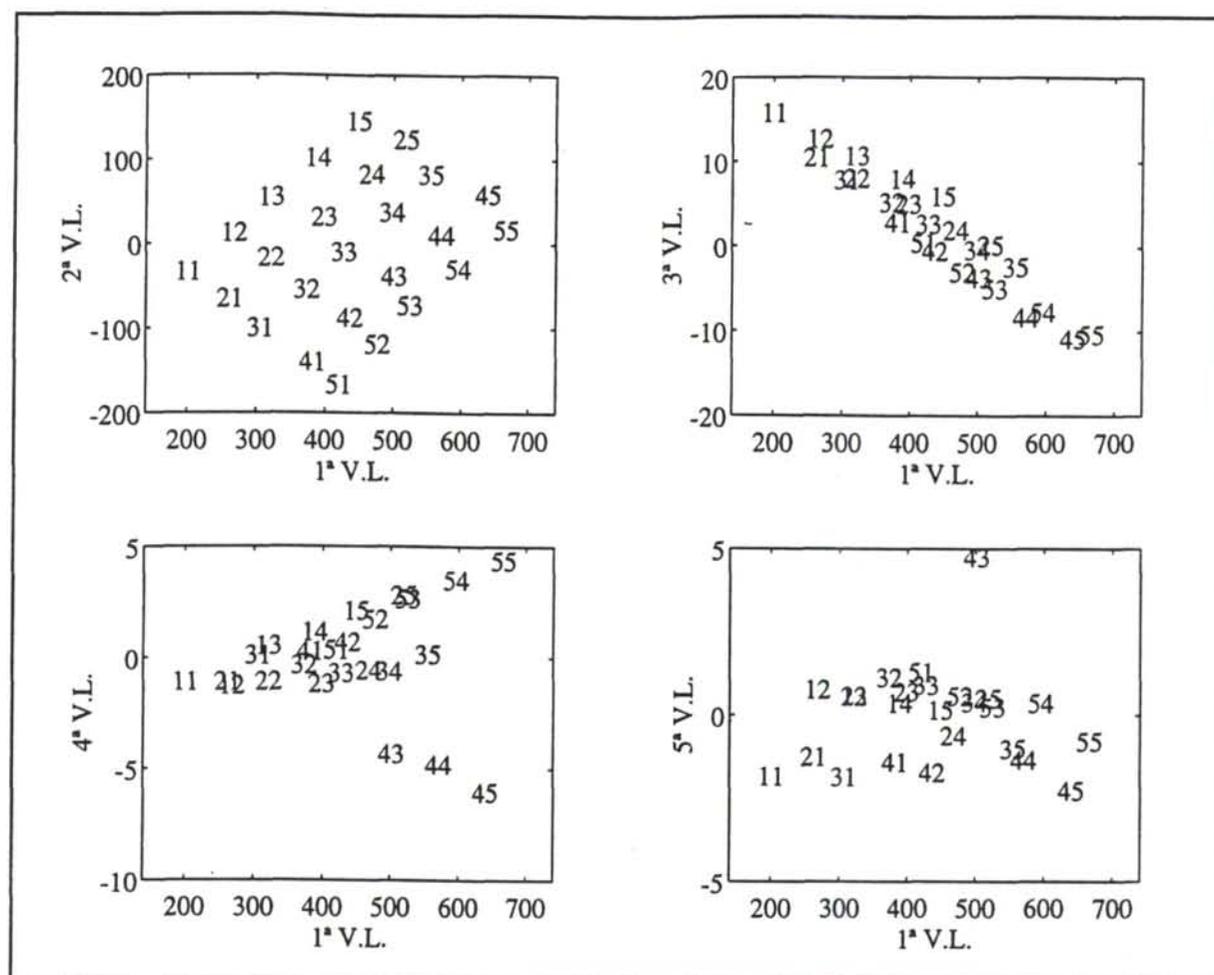


Figura III.6.7 Scores del modelo PLS construido para la determinación de plomo en el calibrado H.

Esta interpretación de las dos primeras variables latentes se ve confirmada cuando se considera la estructura de los loadings correspondientes a cada uno de los modelos PLS construidos para el calibrado H, que se representan en las Figuras III.6.8 y III.6.9. Los loadings de la primera variable latente reproducen la forma de lo que sería un voltamperograma medio, mientras los de la segunda variable latente evidencian la presencia de dos picos muy solapados, presentando éstos valores positivos o negativos dependiendo de que los potenciales correspondientes estén relacionados con uno u otro analito. Tienen valor positivo aquellos loadings que se corresponden con los potenciales en que aparece el pico del analito para el que se construye ese modelo PLS en concreto, mientras que tienen signo opuesto los correspondientes al analito cuya concentración no se está modelando.

Por otro lado, la tercera variable latente también implica una corrección del modelo PLS construido con las dos variables latentes anteriores. La relación lineal existente entre los

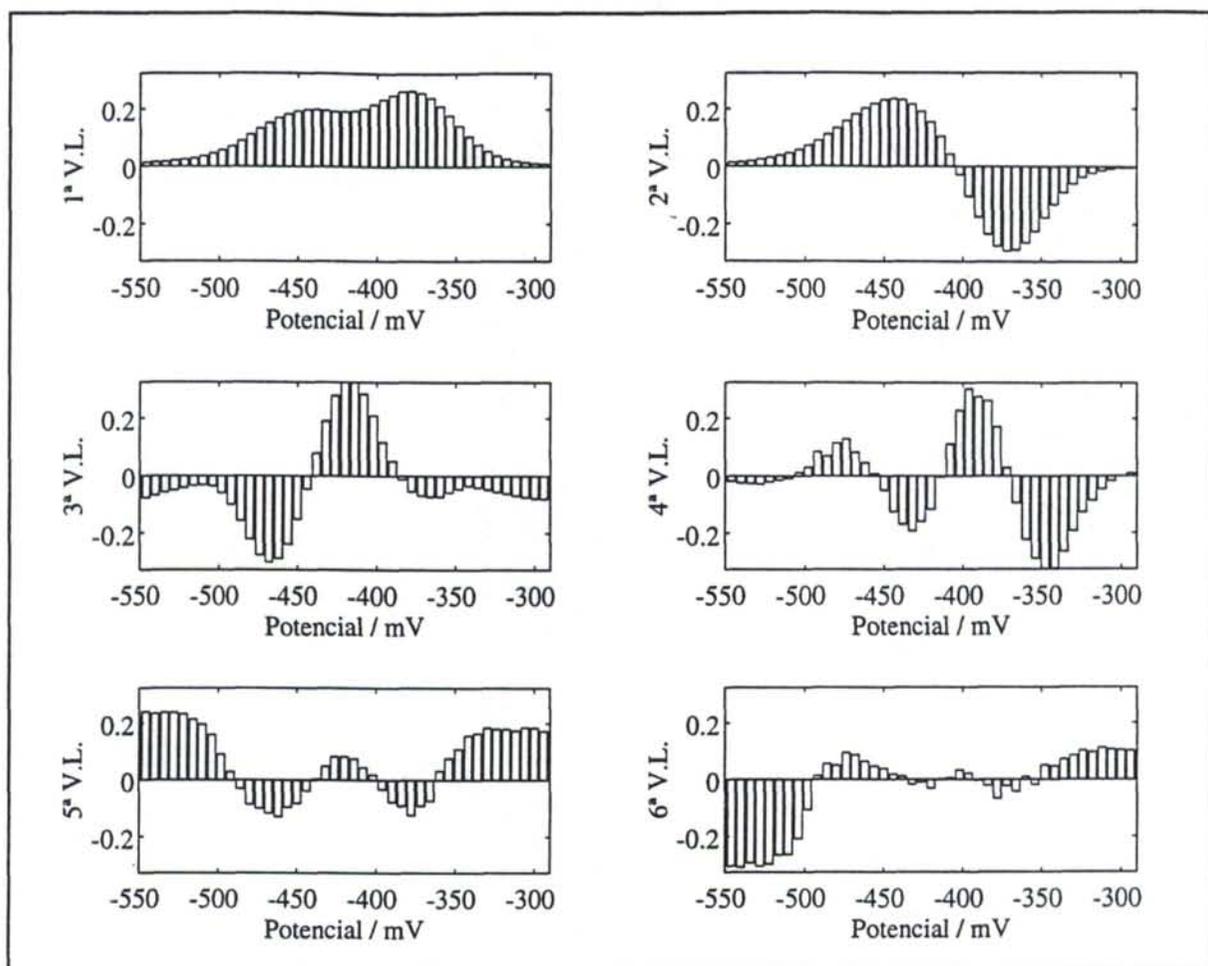


Figura III.6.8 Loadings correspondientes al modelo PLS construido para el talio en el calibrado H.

scores de la tercera variable latente y los de la primera indica que el fenómeno que modela esta tercera variable latente ha de estar íntimamente relacionado con las variables predictoras en sí mismas, es decir, con algo que hace referencia a la señal global y no a la concentración de uno u otro analito en concreto.

De hecho, los scores calculados con la tercera variable latente parecen "ordenar" las muestras de calibrado considerando prácticamente la suma de las concentraciones de ambos analitos, al igual que ocurre con la primera variable latente, aunque en este caso la magnitud de los scores es mucho menor y se puede relacionar con la interferencia debida al blanco experimental. Este extremo se vio confirmado cuando al construir un modelo PLS con las señales a las que previamente se había restado la señal del blanco experimental, es decir, los voltamperogramas de la Figura III.6.5, no se encontró ninguna variable latente con esta estructura.

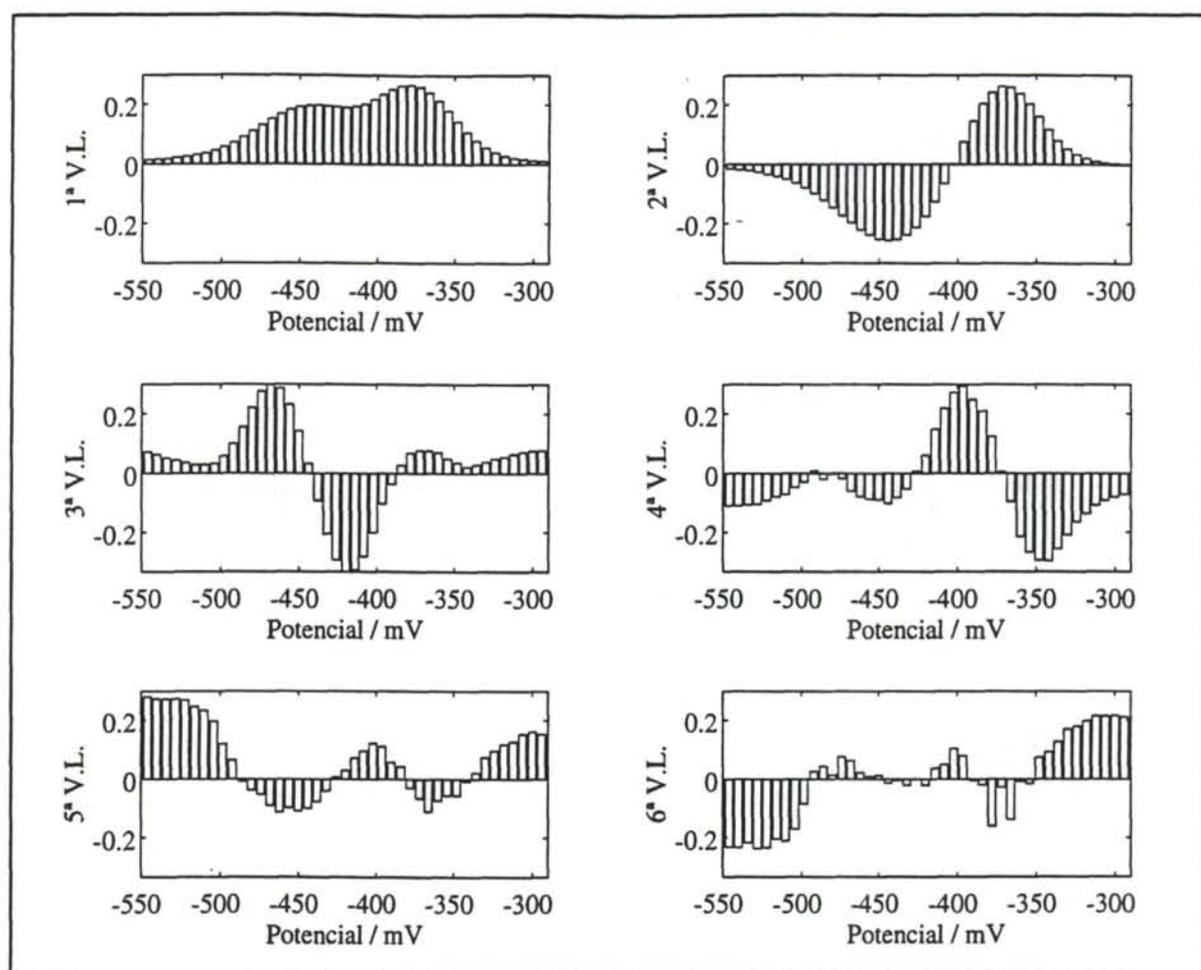


Figura III.6.9 Loadings correspondientes al modelo PLS construido para el plomo en el calibrado H.

Los loadings asociados con esta tercera variable latente parecen establecer una oposición entre los potenciales de la señal del blanco y los correspondientes al pico voltamperométrico del talio. La inclusión de esta tercera variable latente en los modelos PLS construidos supone un incremento mayor de la varianza explicada por los modelos PLS en el caso del talio que en el del plomo, lo cual es lógico si se tiene en cuenta que la señal del blanco afecta principalmente al pico voltamperométrico del talio.

En cuanto a la cuarta variable latente que forma parte de todos los modelos construidos, excepto del correspondiente al plomo en el calibrado I, parece estar más bien relacionada con muestras u objetos concretos, que con algún otro fenómeno más general, como señalan los scores correspondientes a la misma.

La buena capacidad de predicción de estos modelos y la estabilidad de los resultados obtenidos viene dada por los altos valores de varianza en predicción proporcionados por los

mismos y por la proximidad de éstos a los valores de varianza explicada correspondientes. En los modelos PLS construídos para el talio y el plomo con el calibrado H han sido necesarias 4 variables latentes, mientras que en los modelos correspondientes al calibrado I se seleccionaron 4 y 3 variables latentes para los modelos del talio y el plomo.

También es posible llevar a cabo una evaluación de estos resultados considerando verdaderas predicciones a través de los errores asociados con las concentraciones de las muestras del conjunto test, que son completamente ajenas a los modelos PLS construídos. Dichos errores están en las Tablas III.6.4 y III.6.5 para los calibrados H e I respectivamente.

Tabla III.6.4 Errores relativos (en %) asociados con las concentraciones de las muestras del conjunto test del calibrado H.

Muestra	Analito	
	Talio	Plomo
<i>t1</i>	0.06	2.29
<i>t2</i>	4.18	2.77
<i>t3</i>	0.60	5.39
<i>t4</i>	-1.21	-0.73

Tabla III.6.5 Errores relativos (en %) asociados con las concentraciones de las muestras del conjunto test del calibrado I.

Muestra	Analito	
	Talio	Plomo
<i>t1</i>	7.01	7.17
<i>t2</i>	-2.39	-2.79
<i>t3</i>	-2.20	-4.47
<i>t4</i>	-2.85	-3.40

Todos ellos se encuentran prácticamente por debajo del 5%, salvo en el caso de la muestra *t1* del calibrado I en donde se llega a un 7%, por lo que se puede considerar que estos errores se encuentran dentro de lo esperado cuando se utilizan técnicas voltamperométricas que implican una etapa de redisolución.

III.6.3.2 Transferencia de calibrado mediante el procedimiento PDS

Seguidamente se ha procedido a llevar a cabo la transferencia de calibrado mediante los dos métodos de transferencia utilizados previamente, los procedimientos PDS y GCT, tomando el calibrado H como calibrado patrón que se transfiere a la situación I, considerada como situación satélite.

Aplicación de la función "pdsgen"

El procedimiento PDS se ha aplicado utilizando la función "pdsgen" definida anteriormente mediante el formato de entrada/salida que se indica a continuación

$$[F, \text{sub}] = \text{pdsgen}(H, I, w, no)$$

donde se introducen los siguientes parámetros de entrada

- H:** Matriz con los voltamperogramas correspondientes al calibrado H, es decir, los correspondientes a la situación patrón. Tiene de dimensiones 29 filas (un voltamperograma por cada muestra de la Tabla III.6.1) por 47 columnas (potenciales a los que se ha registrado la intensidad).
- I:** Matriz con los voltamperogramas del calibrado I, es decir, los obtenidos en la situación considerada como satélite. Tiene las mismas dimensiones que **H**, y sus elementos se encuentran en idéntico orden.
- w:** Escalar que indica cuántos valores de potencial constituyen la ventana móvil que se utiliza en el cálculo de la matriz de transferencia. En las transferencias de calibrado que se han realizado con diferentes conjuntos de transferencia se han utilizado ventanas que van de 3 a 47 variables predictoras (3, 5, 7, ... , 47), lo que equivale a un total de 23 tamaños de ventana distintos.
- no:** Vector fila en el que se especifican las muestras que se van a utilizar en la transferencia de calibrado; dichas muestras, que constituyen los subconjuntos de transferencia utilizados en la transferencia de calibrado, se encuentran representadas en la Figura III.6.10. Todos los conjuntos de transferencia propuestos, salvo el primero, suponen una reducción importante del número de muestras necesarias para llevar a cabo la determinación de las concentraciones del calibrado satélite.

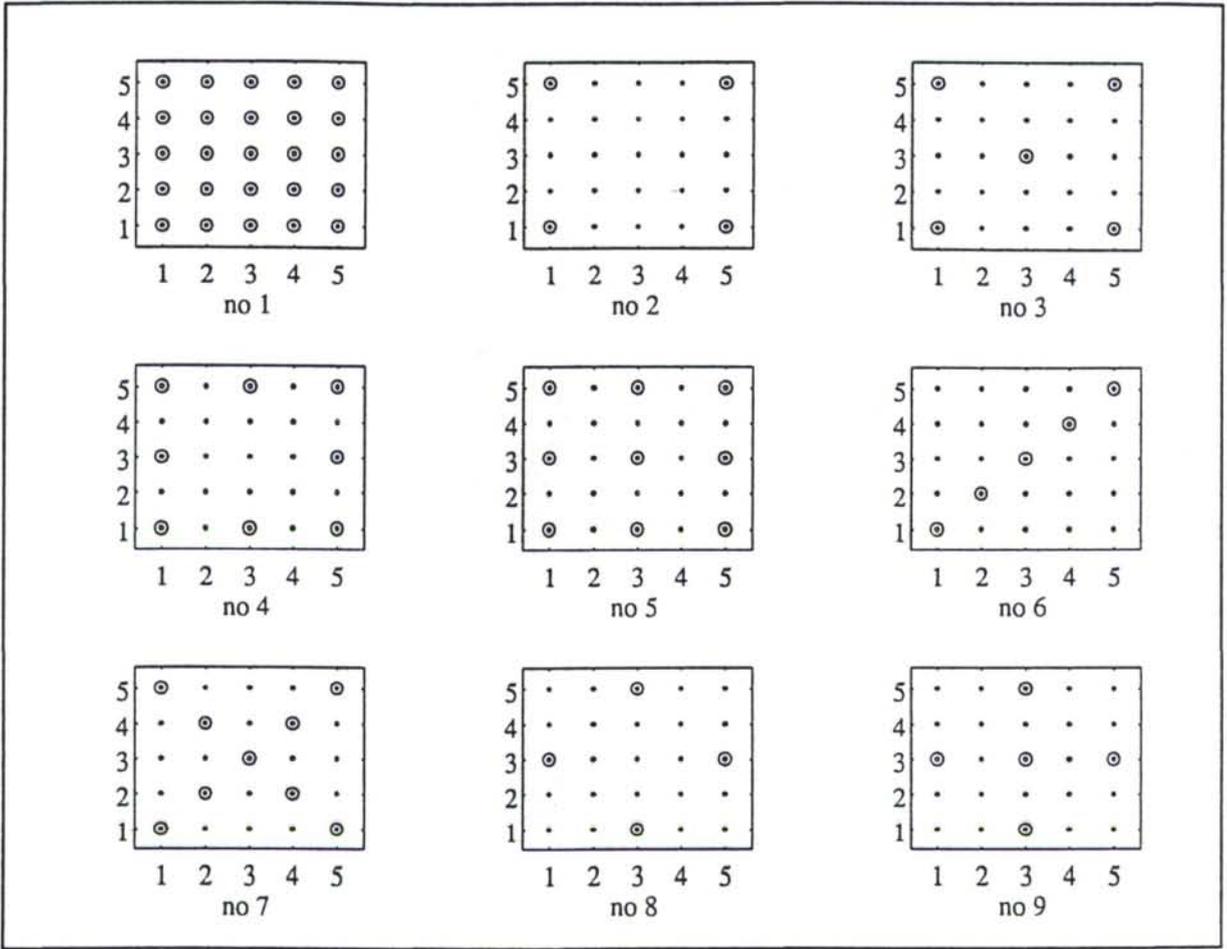


Figura III.6.10 Subconjuntos de transferencia utilizados en la transferencia de calibrado (muestras \odot). En ordenadas el nivel de talio y en abscisas el de plomo.

Los parámetros de salida son, como se ha indicado anteriormente, la matriz de transferencia F que en este caso tiene dimensiones 47×47 , y el vector *sub* que es idéntico al vector *no* introducido y por tanto contiene los índices de las muestras que componen cada uno de los subconjuntos de transferencia.

Se han realizado un total de 207 transferencias de calibrado (23×9) independientes al variar tanto el conjunto de transferencia como el tamaño de ventana utilizados en el procedimiento de transferencia, de forma que es posible estimar cuales son los valores óptimos de estos dos parámetros en la determinación de talio y plomo. Para evaluar la calidad de las transferencias de calibrado llevadas a cabo se han considerado los valores del SEP, que se encuentran representados en las Figuras III.6.11 y III.6.12, para talio y plomo respectivamente, a través de las correspondientes superficies de respuesta y curvas de nivel.

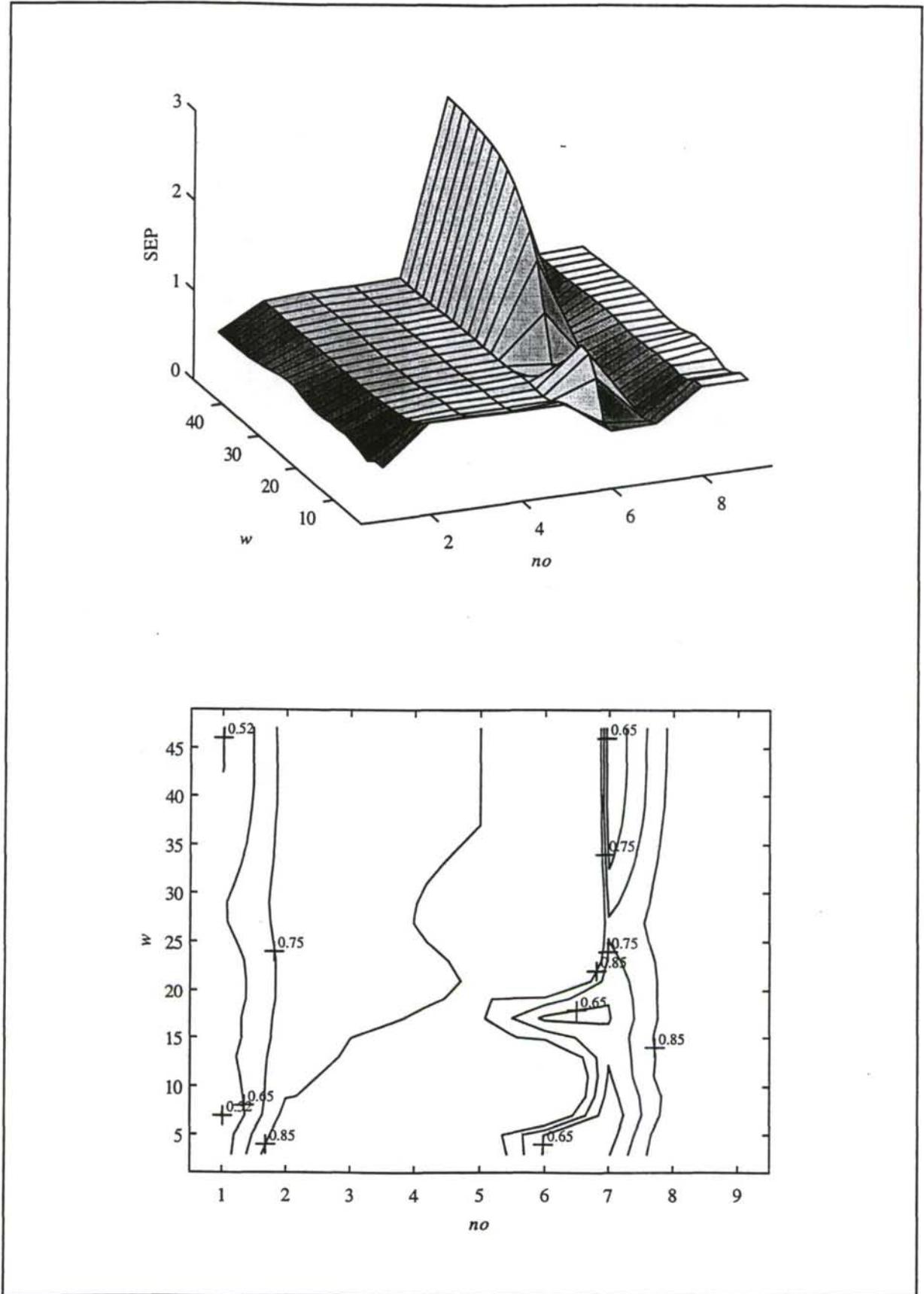


Figura III.6.11 Superficie de respuesta y curvas de nivel del SEP correspondiente al talio.

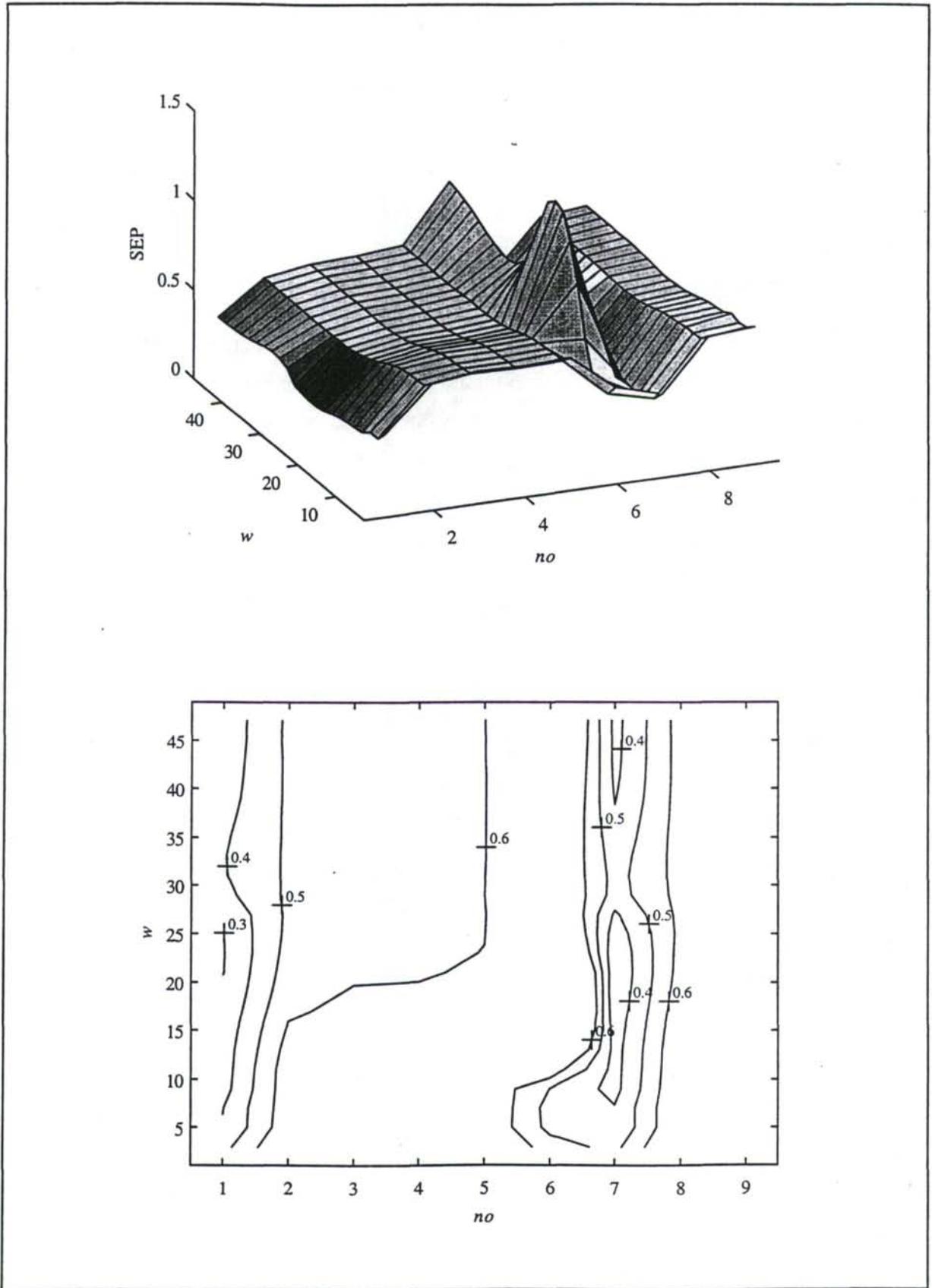


Figura III.6.12 Superficie de respuesta y curvas de nivel del SEP correspondientes al plomo.

Análisis de los resultados

Las Figuras III.6.11 y III.6.12 evidencian claramente que los mejores resultados obtenidos para ambos analitos se corresponden con las transferencias de calibrado llevadas a cabo con los subconjuntos de transferencia 1 y 7, mientras que el subconjunto 6 está asociado con los errores mayores en un amplio rango de tamaños de ventana. Únicamente para tamaños de ventana entre 15 y 20 este último subconjunto de transferencia ofrece resultados aceptables para el talio, pero no así para el plomo, analito para el cual los errores proporcionados por el subconjunto de transferencia 6 son los más altos.

Las peculiaridades asociadas con el conjunto de transferencia 6 se deben a que este conjunto tiene un defecto estructural importante [124], ya que existe una correlación máxima entre las muestras que lo forman. La concentración de ambos analitos crece simultáneamente en todas las muestras que constituyen este conjunto de transferencia, lo cual lleva a una pésima capacidad predictiva del procedimiento de transferencia cuando se le ha utilizado en el cálculo de la matriz de transferencia. En este caso, se podría decir que se ha utilizado un diseño del conjunto de transferencia deficiente ya que, debido a la distribución de las muestras experimentales, no aporta la suficiente información respecto de la situación satélite para un ajustado cálculo de la matriz de transferencia.

Estas conclusiones coinciden con las que se pueden extraer de los diagramas Box plot construídos con los diferentes valores del SEP obtenidos en las transferencias de calibrado llevadas a cabo con distintos valores de n_0 y w , los cuales están representados en las Figuras III.6.13 y III.6.14, así como con los test de Tukey y Newman-Keuls, cuyos resultados se muestran en las Tablas III.6.6 y III.6.7. Entre los resultados producidos por los subconjuntos de transferencia 1 y 7 no existen diferencias estadísticamente significativas y en alguna ocasión, según señala el test de Tukey correspondiente a la transferencia de calibrado realizada para la determinación de talio, éstos son incluso comparables a los correspondientes al subconjunto de transferencia 2.

Dado que el subconjunto de transferencia 1 implica a todas las muestras de calibrado en el proceso de transferencia, el esfuerzo experimental necesario es equivalente a una recalibración completa, mientras que el subconjunto 7 requiere únicamente la medida de 9 muestras patrón para llevar a cabo la transferencia, parece claro que éste último sería el conjunto óptimo para la determinación de talio y plomo a través del procedimiento de transferencia de calibrado que se ha realizado.

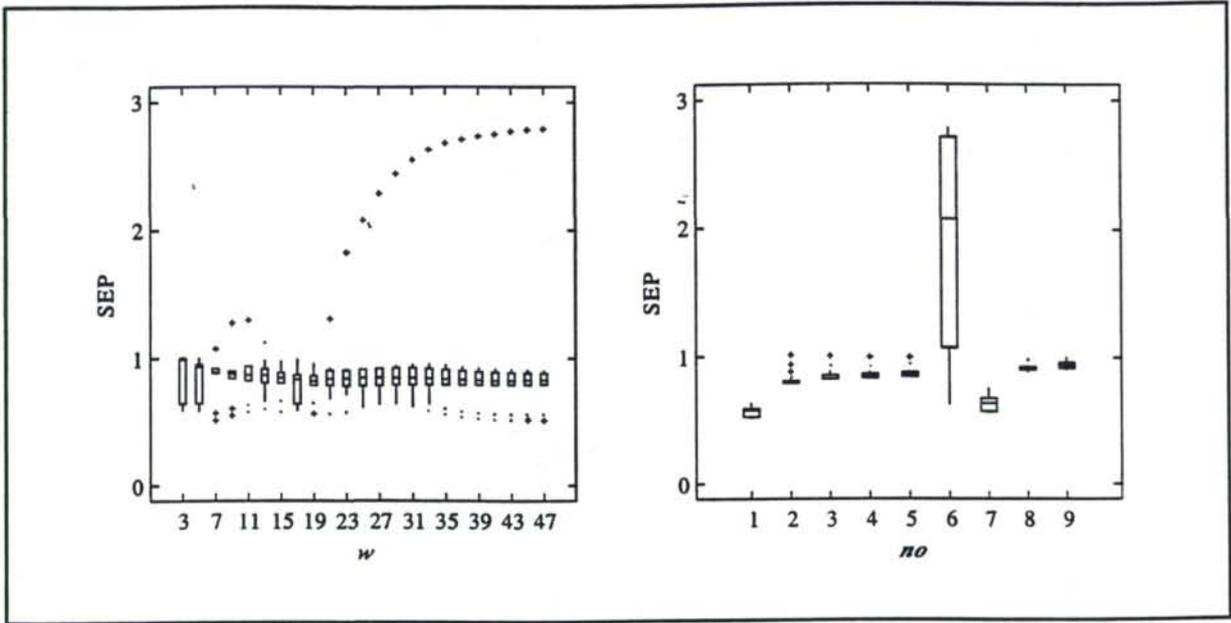


Figura III.6.13 Diagramas Boxplot correspondientes a los valores del SEP obtenidos en la transferencia de calibrado para el talio.

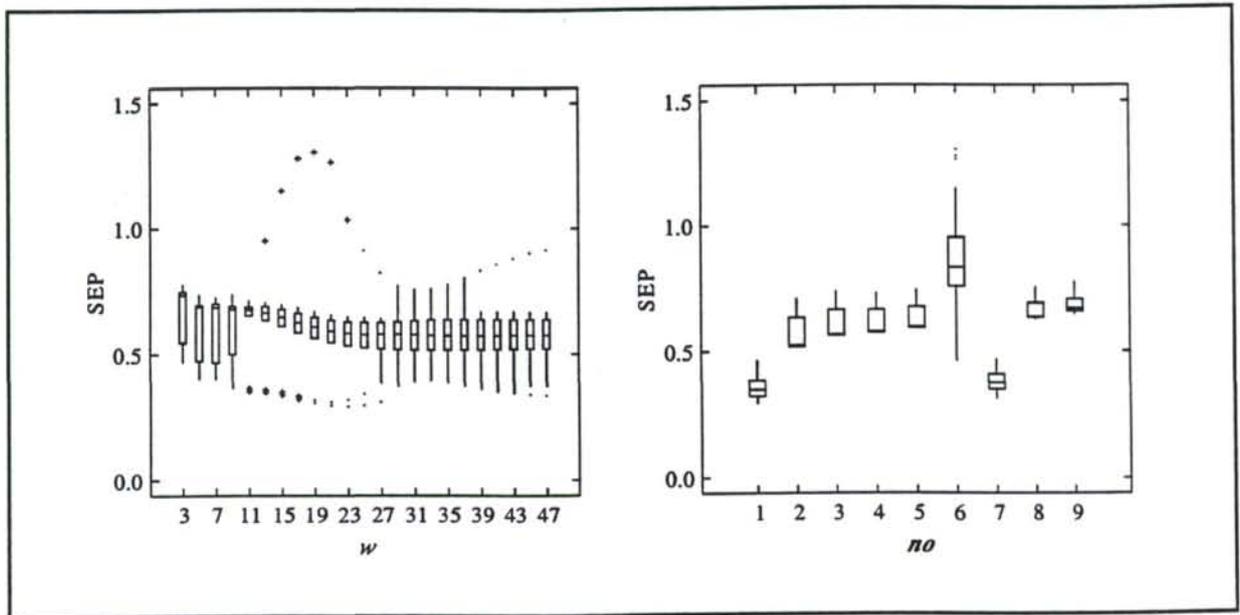


Figura III.6.14 Diagramas Boxplot correspondientes a los valores del SEP obtenidos en la transferencia de calibrado para el plomo.

Esta elección parece muy evidente en lo que se refiere al conjunto de transferencia, sin embargo, en lo que respecta al tamaño de ventana el abanico de posibilidades es menos restringido, tal y como indican los test de rangos llevados a cabo, Tablas III.6.6 y III.6.7, los cuales no aprecian diferencias estadísticamente significativas entre los diversos tamaños de

Tabla III.6.6 Resultados de los tests de agrupamientos de Tukey y Newman-Keuls para los valores de SEP correspondientes a las transferencias de calibrado llevadas a cabo en la determinación de talio.

Conjunto de transferencia (<i>no</i>)				Ventana (<i>w</i>)			
Nivel	Media	Grupos homogéneos		Nivel	Media	Grupos homogéneos	
		Tukey	Newman-Keuls			Tukey	Newman-Keuls
1	0.5705	X	X	17	0.7914	X	X
7	0.6371	XX	XX	19	0.8018	X	X
2	0.8160	XXX	XX	15	0.8217	X	X
3	0.8483	XX	XX	7	0.8402	X	X
4	0.8578	XX	XX	5	0.8410	X	X
5	0.8725	XX	XX	21	0.8540	X	X
8	0.9173	X	X	13	0.8585	X	X
9	0.9443	X	X	9	0.8589	X	X
6	1.8578	X	X	11	0.8727	X	X
				3	0.8747	X	X
				23	0.9165	X	X
				25	0.9540	X	X
				27	0.9852	X	X
				41	0.9935	X	X
				43	0.9942	X	X
				47	0.9955	X	X
				39	0.9956	X	X
				45	0.9956	X	X
				37	0.9993	X	X
				29	1.0016	X	X
				35	1.0045	X	X
				31	1.0078	X	X
				33	1.0083	X	X

ventana utilizados. Ello se debe a que el conjunto de transferencia 6 introduce una enorme dispersión en los datos, como se puede apreciar en los diagramas Box plot correspondientes al tamaño de ventana, Figuras III.6.13 y III.6.14.

Tabla III.6.7 Resultados de los tests de agrupamientos de Tukey y Newman-Keuls para los valores de SEP correspondientes a las transferencias de calibrado llevadas a cabo en la determinación de plomo.

Conjunto de transferencia (<i>no</i>)				Ventana (<i>w</i>)			
Nivel	Media	Grupos homogéneos		Nivel	Media	Grupos homogéneos	
		Tukey	Newman-Keuls			Tukey	Newman-Keuls
1	0.3548	X	X	27	0.5620	X	X
7	0.3852	X	X	25	0.5671	X	X
2	0.5678	X	X	29	0.5720	X	X
3	0.6073	XX	XX	39	0.5720	X	X
4	0.6134	XX	XX	37	0.5722	X	X
5	0.6307	XX	XXX	41	0.5728	X	X
8	0.6603	X	XX	35	0.5732	X	X
9	0.6855	X	X	31	0.5738	X	X
6	0.8551	X	X	33	0.5740	X	X
				43	0.5742	X	X
				45	0.5760	X	X
				47	0.5763	X	X
				23	0.5816	X	X
				9	0.5989	X	X
				7	0.6058	X	X
				5	0.6085	X	X
				11	0.6120	X	X
				21	0.6130	X	X
				19	0.6290	X	X
				13	0.6355	X	X
				17	0.6421	X	X
				15	0.6445	X	X
				3	0.6619	X	X

Llegados a este punto, para seleccionar el tamaño de ventana óptimo se han tenido en consideración los valores de SEP correspondientes al subconjunto de transferencia 7, los cuales se muestran en las Figuras III.6.15 y III.6.16 para talio y plomo respectivamente, acompañados del SET correspondiente. La evolución de los valores del SET en función de

los diferentes conjuntos de transferencia utilizados ha sido en todo momento paralela a la correspondiente al SEP.

En la Figura III.6.15 se observa que en el caso del talio los valores de SEP se encuentran un poco por encima de los correspondientes al SET, pero que los valores mínimos de ambos coinciden en tamaños de ventana grandes, por lo que en la transferencia llevada a cabo para la determinación de este analito se ha seleccionado un tamaño de ventana 47.

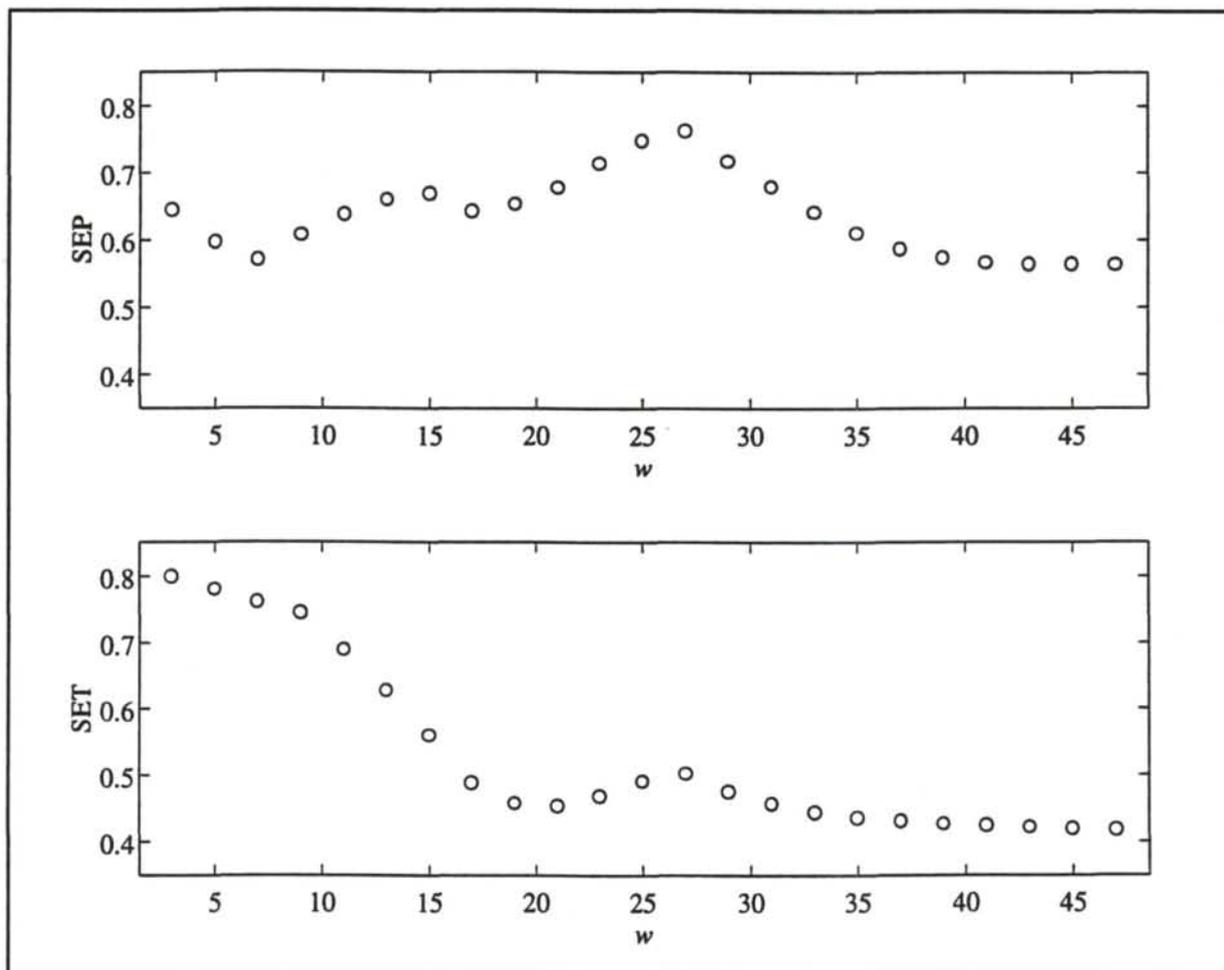


Figura III.6.15 Valores de SEP y SET obtenidos en la determinación de talio mediante el procedimiento de transferencia de calibrado PDS con el conjunto de transferencia 7.

Cuando se considera la situación correspondiente al plomo, Figura III.6.16, se observa que también en este caso los valores del SEP están por encima de los del SET y que siguen exactamente la misma pauta, que señala el tamaño de ventana 21 como óptimo para la determinación de este analito ya que se corresponde con el valor mínimo del SEP.

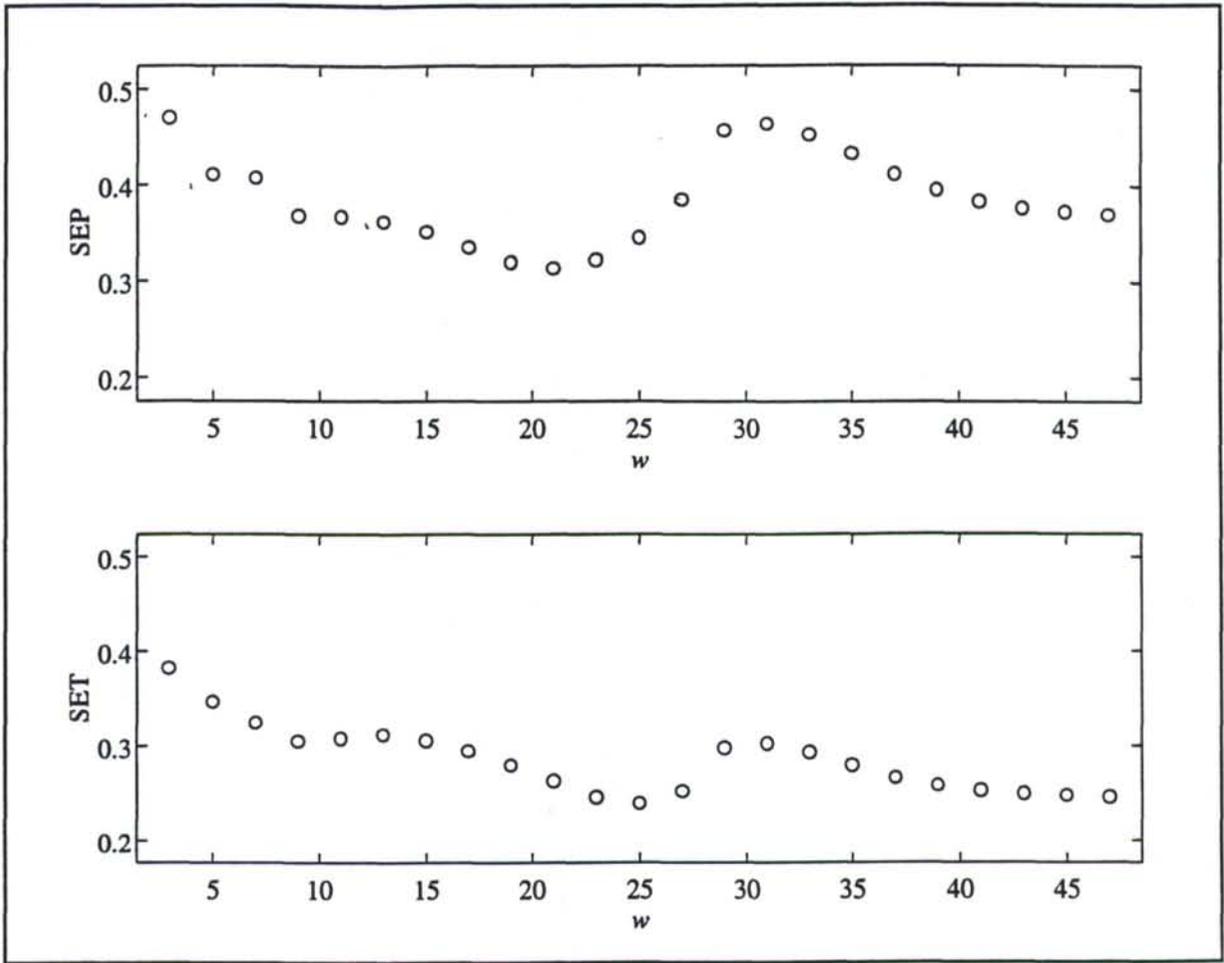


Figura III.6.16 Valores de SEP y SET obtenidos en la determinación de plomo mediante el procedimiento de transferencia de calibrado PDS con el conjunto de transferencia 7.

Los resultados de las transferencias de calibrado llevadas a cabo en estas condiciones quedan reflejados en los errores asociados con las concentraciones calculadas para las muestras del conjunto test del calibrado I, los cuales se muestran en la Tabla III.6.8. Estos errores, obtenidos tras llevar a cabo la transferencia de calibrado, son del mismo orden que los alcanzados para las mismas muestras test por los modelos PLS construidos con todas las muestras del calibrado I, Tabla III.6.5. Sin embargo, cuando las determinaciones se han llevado a cabo mediante el procedimiento de transferencia de calibrado únicamente se han utilizado 9 de las 25 muestras de calibrado. Por lo tanto, el procedimiento de transferencia utilizado conduce a una disminución efectiva del esfuerzo experimental, sin que ello conlleve una disminución significativa de la calidad de las estimaciones realizadas.

Tabla III.6.8 Errores relativos (en %) para las muestras del conjunto test. $no = 7$ y $w = 47$ para el talio y $no = 7$ y $w = 21$ para el plomo.

Muestra	Analito	
	Talio	Plomo
$t1$	5.07	0.71
$t2$	-3.13	-2.95
$t3$	-4.72	-5.24
$t4$	-6.77	-3.38

III.6.3.3 Transferencia de calibrado mediante el procedimiento GCT

Aplicación de la función "gct"

La función "gct" ha sido utilizada para realizar una nueva transferencia de calibrado entre la situación H y la situación I. Para ello se han utilizado las matrices de datos correspondientes a los calibrados H e I indicadas anteriormente, junto con los diferentes conjuntos de transferencia representados en la Figura III.6.10. Los parámetros introducidos en la función "gct" han sido los siguientes:

$$[F, sub] = gct(H, I, 15, no, 3, Hcer)$$

donde **H** e **I** son las matrices de datos señaladas, y **no** es el vector que indica las muestras que forman parte de los conjuntos de transferencia utilizados.

Una vez obtenidas las matrices de transferencia, se transforman los coeficientes de los modelos PLS del calibrado H y con ellos se calculan las concentraciones de las muestras del conjunto test del calibrado I. Los valores de SEP y SET asociados con las transferencias de calibrado llevadas a cabo para talio y plomo mediante el procedimiento de transferencia GCT se muestran en las Figuras III.6.17 y III.6.18.

Análisis de los resultados

Al igual que ocurría en el procedimiento de transferencia PDS utilizado previamente, los subconjuntos de transferencia asociados con los valores de SEP menores son el 1 y el 7, e incluso el conjunto 2 en el caso del talio, como puede observarse en la Figura III.6.17. De

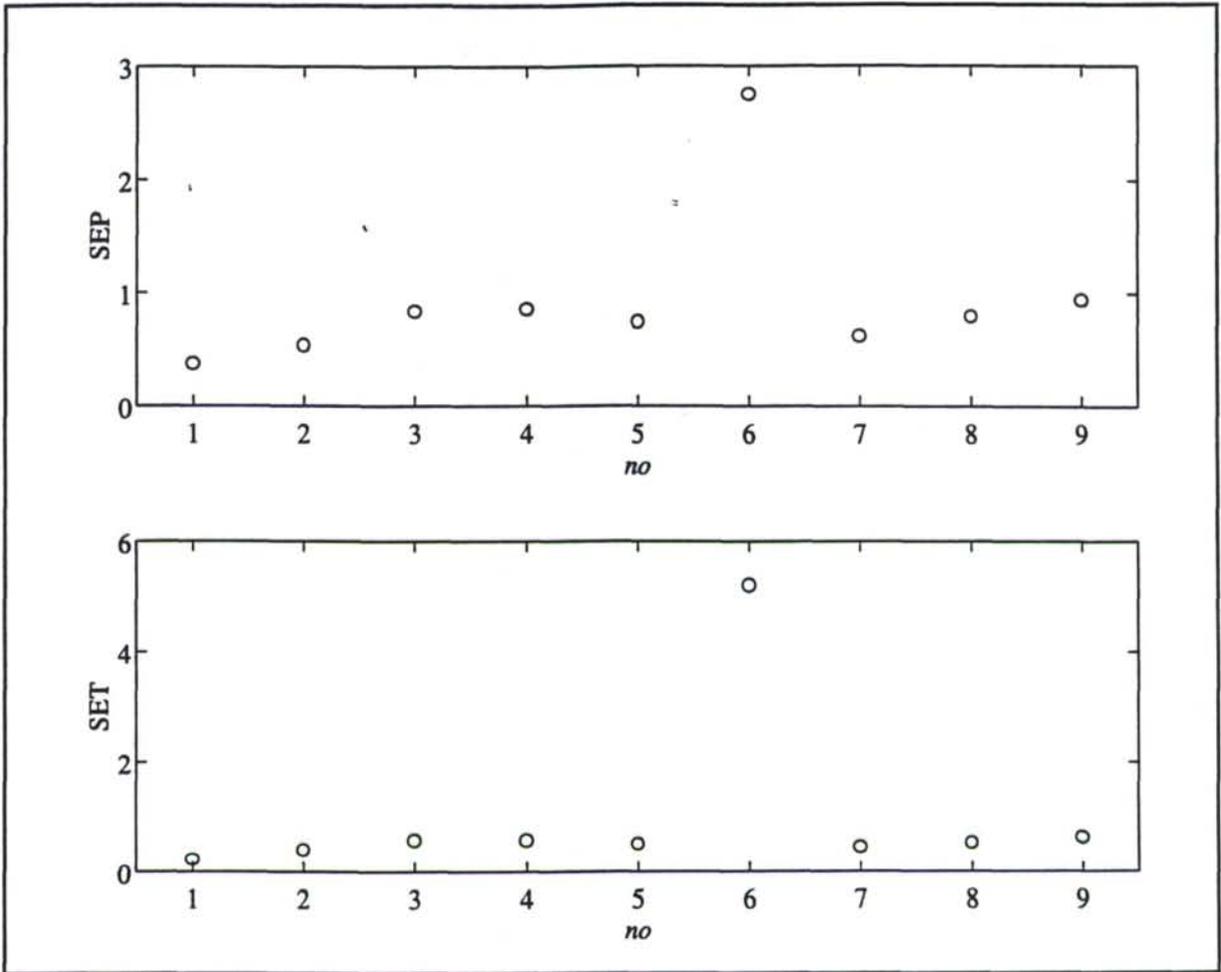


Figura III.6.17 Valores de SEP y SET obtenidos en la determinación de talio mediante el procedimiento de transferencia de calibrado GCT.

hecho, el conjunto 6 continúa proporcionando los errores mayores, sobre todo en el caso del talio, donde incluso el valor del SET dobla el correspondiente valor de SEP.

Los errores asociados con la concentración de talio y plomo de las muestras test del calibrado I cuando ésta se ha calculado mediante este procedimiento de transferencia se encuentran en la Tabla III.6.9. Estos valores están un poco por encima de los errores correspondientes a los propios modelos PLS del calibrado I, Tabla III.6.5, e incluso de los obtenidos cuando se aplicó el procedimiento de transferencia de calibrado PDS, Tabla III.6.8, aunque siguen dentro de lo que cabe esperar cuando se utilizan estas técnicas de análisis.

En principio, el uso de uno u otro método de transferencia de calibrado no parece conducir a resultados significativamente diferentes, incluso en lo que se refiere al esfuerzo experimental requerido para llevar a cabo cada uno de ellos ya que las muestras del conjunto de transferencia son las mismas en uno y otro caso.

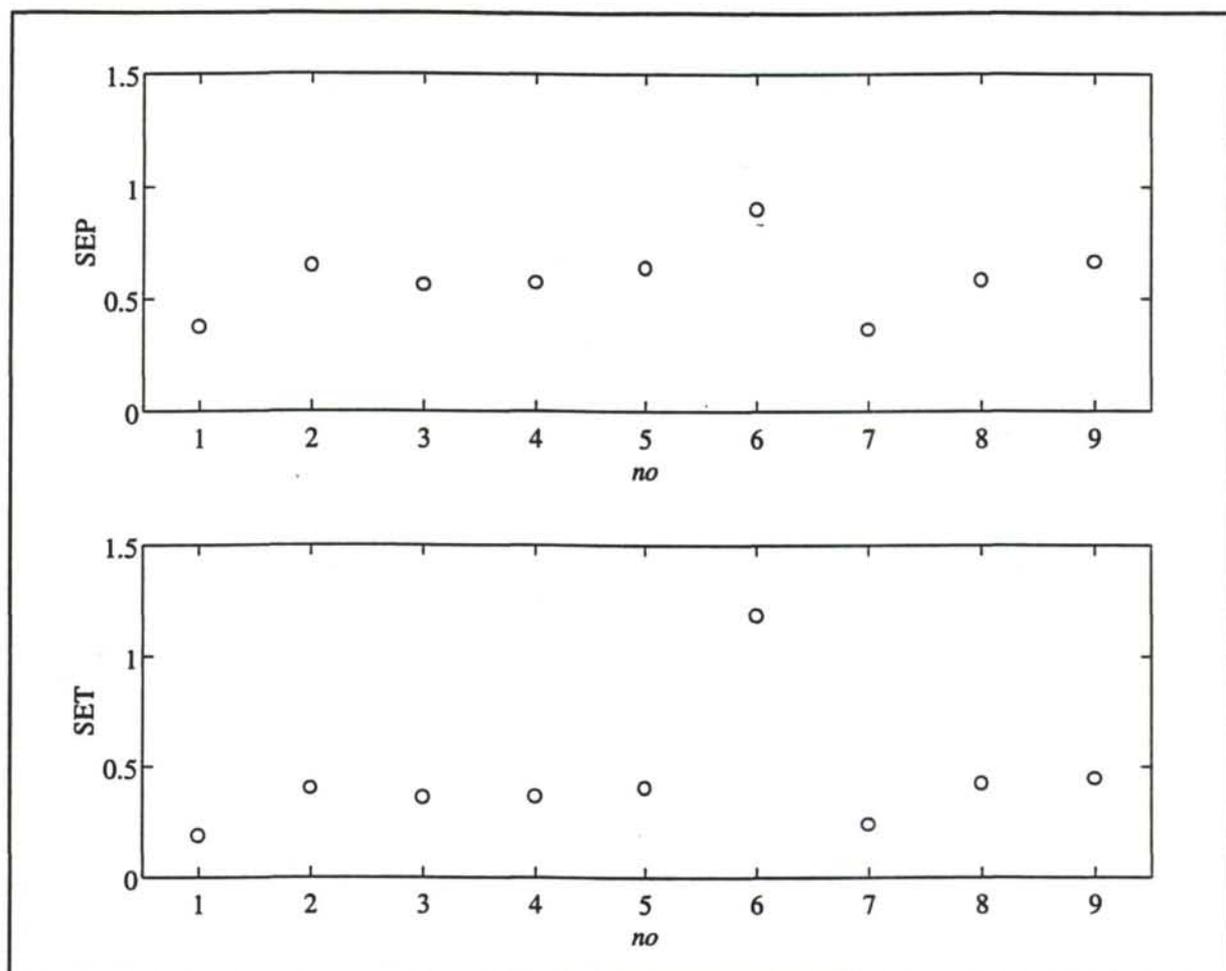


Figura III.6.18 Valores de SEP y SET obtenidos en la determinación de plomo mediante el procedimiento de transferencia de calibrado GCT.

Tabla III.6.9 Errores relativos (%) asociados con las muestras del conjunto test del calibrado I tras la transferencia de calibrado CGT realizada con el subconjunto de transferencia 7.

Muestra	Analito	
	Talio	Plomo
<i>t1</i>	6.39	9.73
<i>t2</i>	-3.01	-3.05
<i>t3</i>	-4.87	-4.57
<i>t4</i>	-7.61	-4.19

III.7 EFECTO MATRIZ Y TRANSFERENCIA DE CALIBRADO ENTRE INSTRUMENTOS

Cuando se utilizan técnicas electroquímicas que incluyen una etapa de redisolución, la aparición de respuestas solapadas puede deberse a la existencia de dos o más especies capaces de formar amalgamas sobre el electrodo de mercurio, o también, a la presencia de especies electroactivas en la propia disolución que también puedan dar lugar a la aparición de señales electroquímicas a los mismos valores de potencial que el o los analitos de interés. En este último caso, el fenómeno se conoce de forma más general como efecto matriz. El ejemplo más común de este tipo de interferencia en los análisis electroquímicos viene dado por la presencia de oxígeno en disolución, ya que la concentración del mismo puede llegar a ser del orden de 10^{-3} M en soluciones acuosas a presión y temperatura ambiente.

El oxígeno disuelto ejerce en los análisis voltamperométricos con redisolución sobre electrodo de mercurio el mismo efecto interferente que en la polarografía clásica ya que, dependiendo del pH, éste es susceptible de reducirse en dos etapas originando sendas señales electroquímicas que pueden solapar las correspondientes al analito de interés. Incluso se pueden producir otras complicaciones debido a la reactividad química de este elemento, que podría oxidar algunos metales de la amalgama formada en la superficie del electrodo o incluso precipitar iones metálicos en las proximidades del mismo. Por estas razones, el oxígeno debe ser eliminado de las disoluciones electroquímicas antes de la etapa de preconcentración.

Otra interferencia que responde a este mismo esquema y que no resulta tan fácil de eliminar como la debida al oxígeno es la que origina la presencia de hierro en la determinación electroquímica de cobre. El cobre es un elemento que se incluye en multitud de análisis rutinarios (aguas naturales y residuales, alimentos, metalurgia, etc.), en los cuales se puede encontrar acompañado de concentraciones elevadas de hierro [199-201]. Este último elemento puede actuar como interferente cuando se lleva a cabo la determinación de cobre [202-203], y en particular cuando para ello se utilizan técnicas electroquímicas [204-207].

Para minimizar las interferencias originadas en el análisis por el efecto matriz se dispone de una serie de métodos cuya finalidad es evitar llevar a cabo separaciones químicas en la muestra. Uno de estos métodos consiste en preparar patrones de calibrado que reproduzcan lo más aproximadamente posible la matriz en la que se va a determinar el analito. Otros por el contrario, proponen procedimientos experimentales alternativos tales como el método del estándar interno o el método de adición estándar. La aplicabilidad de unos y otros se encuentra limitada por la complejidad de la matriz de que se trate y por el tipo de interferencia ejercida por la misma, que puede ser aditiva o multiplicativa [208].

Una solución al problema que suponen las interferencias debidas a la propia disolución, o lo que es lo mismo, al denominado efecto matriz, cuando se trabaja con técnicas voltamperométricas, se encuentra en los procedimientos de intercambio de medio, en los cuales la etapa de preconcentración se lleva a cabo en la propia disolución, mientras que la posterior etapa de redisolución se realiza en un medio más adecuado para el análisis. El cambio a una disolución que contenga electrólito soporte "blanco" eliminaría la interferencia a que se está haciendo referencia. Gottestfeld y Ariel propusieron la aplicación de este procedimiento en la determinación de concentraciones bajas de cobre en aceros de alta aleación [209] depositando el cobre sobre un HMDE en una disolución alcalina de trietanolamina y realizando la determinación, tras un cambio a otra disolución ya desoxigenada de trietanolamina, a partir del voltamperograma de redisolución anódica.

Sin embargo, a pesar de que las ventajas del procedimiento de intercambio de medio son bien conocidas, este método lleva asociados una serie de problemas experimentales que limitan enormemente su uso. Se trata de cuestiones tales como la exposición del electrodo a la atmósfera, la posible cotransferencia de la disolución original, la ruptura del circuito eléctrico, etc.

Otra forma de abordar esta cuestión es la propuesta por Bonelli y otros [210], los cuales utilizan una aproximación substractiva que corrige la contribución del hierro a la intensidad registrada. Estos autores consideran que la intensidad del pico voltamperométrico del hierro es independiente del tiempo de acumulación de la etapa de preconcentración y obtienen un voltamperograma a tiempo de acumulación cero, el cual se resta posteriormente tanto de la señal de la muestra como de las de los patrones. El método completo implica la reducción a cenizas de la muestras por vía húmeda seguida de la reducción del cobre al formar una amalgama sobre el electrodo de mercurio en medio citrato amónico, que actúa simultáneamente como tampón y electrólito soporte.

El uso de electrólitos soporte tales como ácido cítrico-citrato, ácido acético-acetato, etc., que además pueden actuar como sistemas tampón, es algo habitual en las determinaciones electroquímicas de cationes metálicos tales como cobre o hierro. Muchos de estos electrólitos soporte tienen también tendencia a formar complejos con los cationes metálicos, por lo que se han realizado numerosos estudios para conocer el comportamiento de estos metales en presencia de agentes complejantes [211].

También es posible utilizar agentes quelantes para enmascarar selectivamente las señales originadas por la presencia de iones metálicos no deseados, e incluso emplear agentes tensoactivos para enmascarar o desplazar estas señales. Esta posibilidad de inhibición selectiva ha sido utilizada por Lukaszewski y otros [194] para determinar impurezas en presencia de un gran exceso de elemento matriz. De igual forma, Ismail y Etaiw [212], utilizan pirogalol para suprimir la señal correspondiente al Fe(III) en la determinación de algunos metales por polarografía. Pratt y William [213] utilizan ácido fosfórico como soporte para eliminar la interferencia debida la Fe(III) en la determinación de Cr(VI) por amperometría con inyección en flujo.

También se encuentran ejemplos en la bibliografía en los cuales se recurre a procesos de extracción para eliminar el interferente. Meenakumari [205] propone llevar a cabo la extracción de hierro con dietil éter en la determinación voltamperométrica de cobre, cadmio, plomo y zinc en muestras de minerales de hierro. En algunos casos, se recurre a someter a las muestras a tratamientos previos para separar el hierro en forma de hidróxido [214].

En el presente capítulo se ponen de manifiesto algunos problemas que pueden surgir al aplicar el procedimiento substractivo propuesto por Bonelli y otros [210], y se sugiere una estrategia alternativa para resolver el problema del efecto matriz en la determinación de Cu(II) en presencia de exceso de Fe(III) por DPASV.

III.7.1 Experimental. Determinación de Cu y Fe por VDIRA

Reactivos

Todas las disoluciones fueron preparadas partiendo de reactivos de calidad para análisis y de agua bidestilada y posteriormente desionizada mediante un sistema Barnstead NANO Pure II.

Las disoluciones patrón de Cu(II) y Fe(III) fueron preparadas en medio ácido para su preservación. Como electrólito soporte/tampón se utilizó una disolución constituida por ácido cítrico/citrato que presentaba un valor de pH de 4.733. Todas las disoluciones analizadas fueron desoxigenadas burbujeando nitrógeno (99.997%) antes de realizar cualquier medida electroquímica.

Instrumental

Algunas de las medidas voltamperométricas se han llevado a cabo utilizando un sistema μ AUTOLAB de Eco Chemie conectado con una unidad IME 663 (Interface to Mercury Electrodes) y con un 663 VA Metrohm que consta de un electrodo multimodo Metrohm (MME) utilizado en el modo HMDE (0.40 mm^2 de área nominal), un electrodo de referencia Ag/AgCl/KCl (3M) y un electrodo auxiliar de carbono vitrificado (glassy carbon, GC). La adquisición de datos se realizó utilizando el programa General Purpose Electrochemical System (GPES) [215].

El resto de las medidas voltamperométricas fueron realizadas utilizando el procesador Metrohm 646 VA processor junto con el 647 VA equipado con un electrodo multimodo Metrohm (MME) en el modo HMDE (0.40 mm^2 de área nominal), un electrodo de referencia Ag/AgCl/KCl (3M) y un electrodo auxiliar de platino.

En los análisis voltamperométricos llevados a cabo se ha utilizado la voltamperometría diferencial de impulsos con redisolución anódica (DPASV) como técnica de medida.

Los análisis de datos fueron realizados utilizando los programas PARVUS [57], STATGRAPHICS [58] y MATLAB [59].

Procedimiento experimental

Cada disolución fue desoxigenada durante 10 min. burbujeando nitrógeno, aplicándose a continuación un potencial de deposición de -1.3 V sobre el electrodo de trabajo durante 60 s. Al final de este período de acumulación se desconectó la agitación y, tras un tiempo de equilibrio de 30 s, se aplicó un barrido de potencial anódico desde -0.154 V hasta 0.050 V a una velocidad de barrido de 10 mV s^{-1} (duración del impulso, 0.6 s). Después de cada adición la disolución se agitó y desoxigenó de nuevo durante 15 s, y se aplicó el mismo procedimiento de medida.

Cuando se utilizó el sistema μ AUTOLAB la amplitud del impulso fue de -49.95 mV y la velocidad de agitación 1500 rev min^{-1} , mientras que con el equipo Metrohm 646 VA processor la amplitud del impulso que se seleccionó fue -50 mV y la velocidad de agitación 1290 rev min^{-1} .

III.7.2 Diseño del experimento

Se ha utilizado un diseño experimental en el que se recorren 6 niveles de concentración de cobre y hierro diferentes, realizándose adiciones equimoleculares de cada uno de estos analitos de forma que se cubre exhaustivamente el espacio experimental, tal y como se muestra en la Figura III.7.1. La experiencia consta de un total de 36 muestras de calibrado y de cuatro muestras test (muestras $t1$ a $t4$) cuyas concentraciones se indican en la Tabla III.7.1. Como puede observarse en dicha tabla, la concentración de hierro es del orden de cien veces superior a la de cobre.

Siguiendo el diseño experimental de la Tabla III.7.1 se han llevado a cabo tres calibrados independientes. El primero de ellos se ha realizado utilizando el sistema μ AUTOLAB, es decir, el primer instrumento descrito en el apartado *Experimental*, y en adelante será denominado calibrado J. Los dos calibrados siguientes se han realizado utilizando el segundo instrumento, Metrohm 646 VA processor, y en días sucesivos, denominados en adelante calibrados K y L, y correspondiéndose el calibrado K con el realizado en primer lugar y el calibrado L con el siguiente.

Los voltamperogramas correspondientes a los calibrados indicados se muestran en las Figuras III.7.2a-f, III.7.3a-f y III.7.4a-f. La señales obtenidas constan en todos los casos de un único pico voltamperométrico, más o menos ancho, que es el resultado de la suma algebraica de las intensidades correspondientes al cobre y al hierro. Ello se debe a que, como muestran los voltamperogramas correspondientes al nivel 0 de cualquiera de los dos analitos, los potenciales de pico del cobre y el hierro se encuentran muy próximos en estas condiciones de medida, originando por ello un elevado solapamiento de las dos señales.

Por otro lado, cabe destacar que la contribución del hierro a la señal registrada, considerado este analito como elemento matriz, es muy grande frente a la contribución del analito de interés, el cobre, por lo que el efecto matriz es en este caso muy importante.

Tabla III.7.1 Diseño experimental seguido en los calibrados J, K y L. Cada símbolo de la segunda columna indica que se realizaron adiciones sobre una misma disolución inicial.

Muestra	Diseño experimental		Concentraciones (μM)	
	Cobre	Hierro	Cobre	Hierro
1 ×	0	0	0.0000	0.0000
2 ×	1	0	0.2074	0.0000
3 ×	2	0	0.4121	0.0000
4 ×	3	0	0.6141	0.0000
5 ×	4	0	0.8135	0.0000
6 ×	5	0	1.0103	0.0000
7 ▲	0	1	0.0000	20.7417
8 ▲	1	1	0.2061	20.6053
9 ▲	2	1	0.4094	20.4706
10 ▲	3	1	0.6101	20.3377
11 ▲	4	1	0.8083	20.2065
12 ▲	5	1	1.0038	20.0769
13 ▲	5	2	0.9975	39.8981
14 ○	0	2	0.0000	41.2105
15 ○	1	2	0.2047	40.9412
16 ○	2	2	0.4068	40.6753
17 ○	3	2	0.6062	40.4129
18 ○	4	2	0.8031	40.1538
19 □	0	3	0.0000	61.4118
20 □	1	3	0.2034	61.0130
21 □	2	3	0.4041	60.6194
22 □	3	3	0.6023	60.2308
23 □	4	3	0.7980	59.8471
24 □	5	3	0.9911	59.4684
25 •	0	4	0.0000	81.3506
26 •	1	4	0.2021	80.8258
27 •	2	4	0.4015	80.3077
28 •	3	4	0.5985	79.7962
29 •	4	4	0.7929	79.2911
30 •	5	4	0.9849	78.7925
31 ■	0	5	0.0000	101.0323
32 ■	1	5	0.2008	100.3846
33 ■	2	5	0.3990	99.7452
34 ■	3	5	0.5947	99.1139
35 ■	4	5	0.7879	98.4906
36 ■	5	5	0.9788	97.8750
<i>t</i> 1 ▲	1.5	1	0.3081	20.5377
<i>t</i> 2 •	2.5	4	0.5003	80.0511
<i>t</i> 3 ○	4	2.5	0.8005	50.0319
<i>t</i> 4 •	5	4.5	0.9818	88.3636

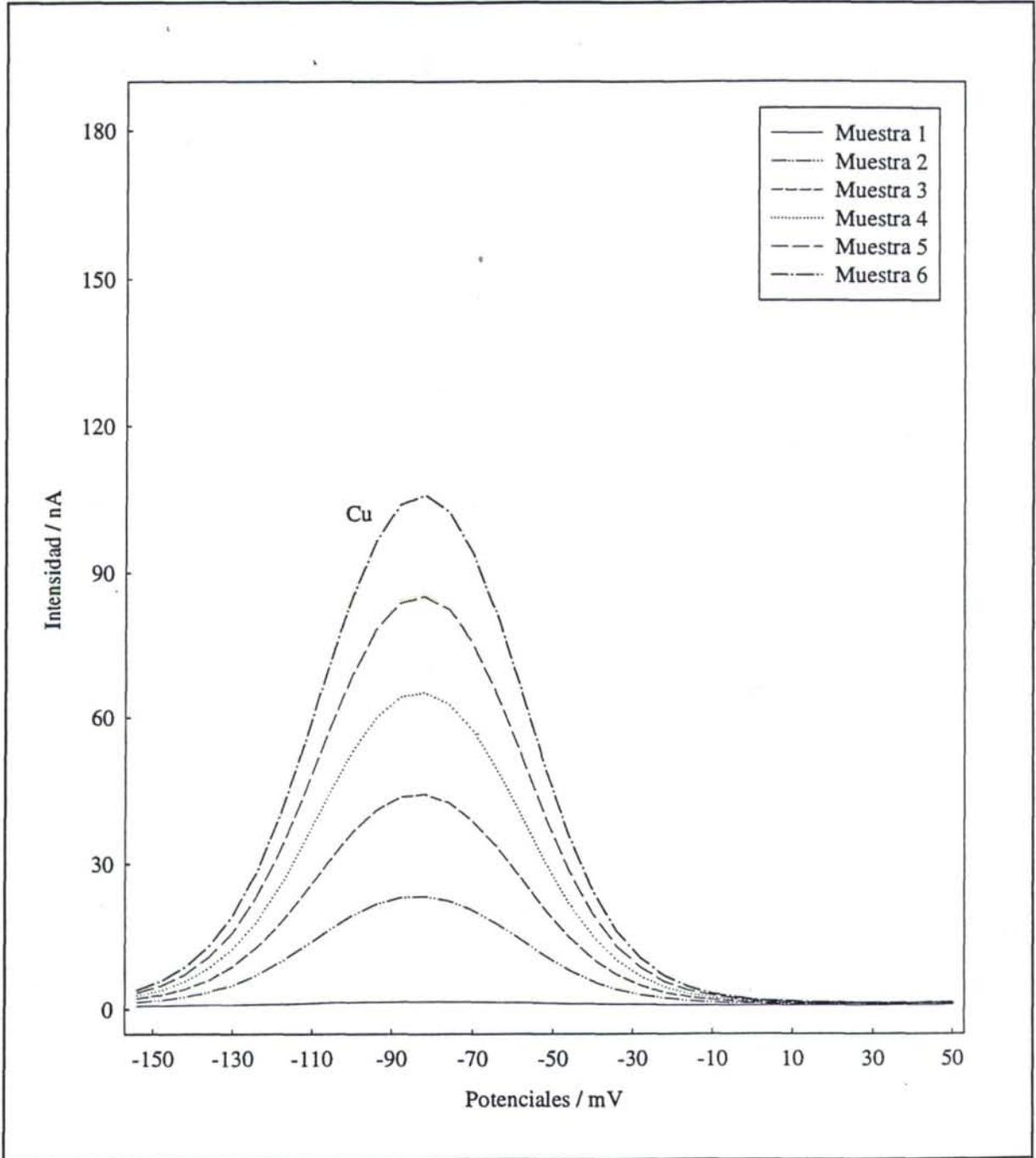


Figura III.7.2a Voltamperogramas correspondientes al calibrado J obtenidos a partir de una misma disolución inicial.

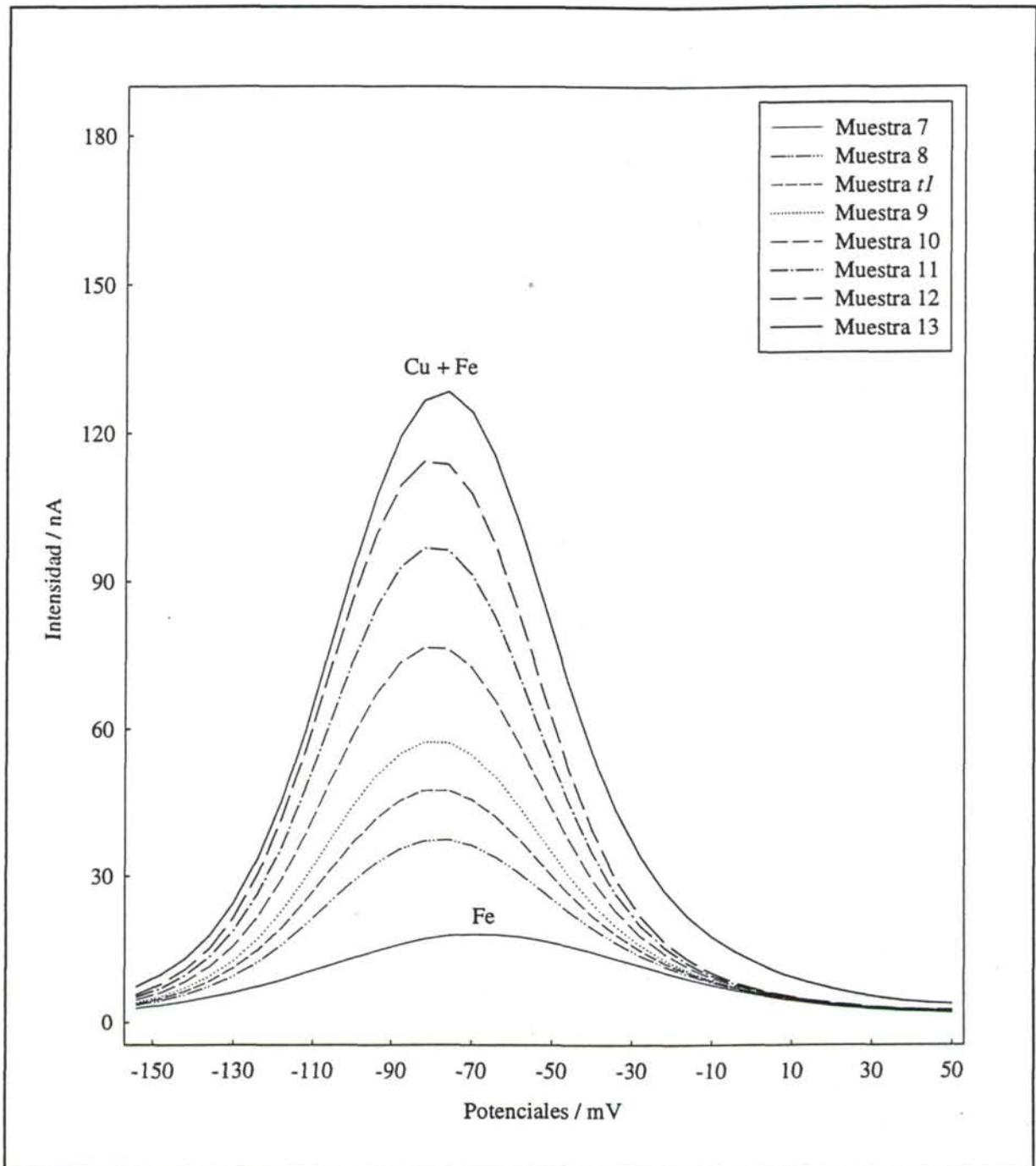


Figura III.7.2b Voltamperogramas correspondientes al calibrado J, obtenidos a partir de una misma disolución inicial.

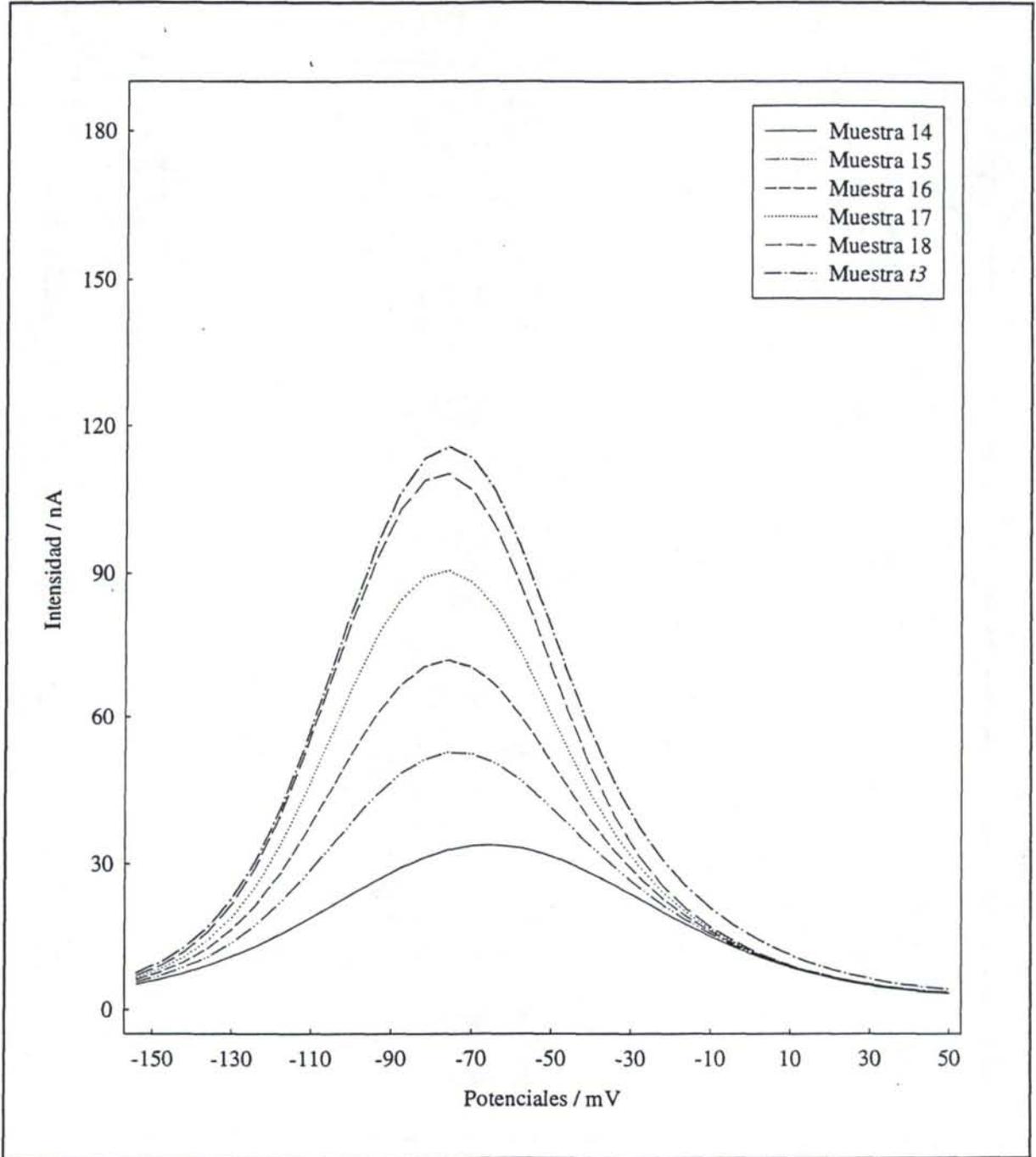


Figura III.7.2c Voltamperogramas correspondientes al calibrado J obtenidos a partir de una misma disolución inicial.

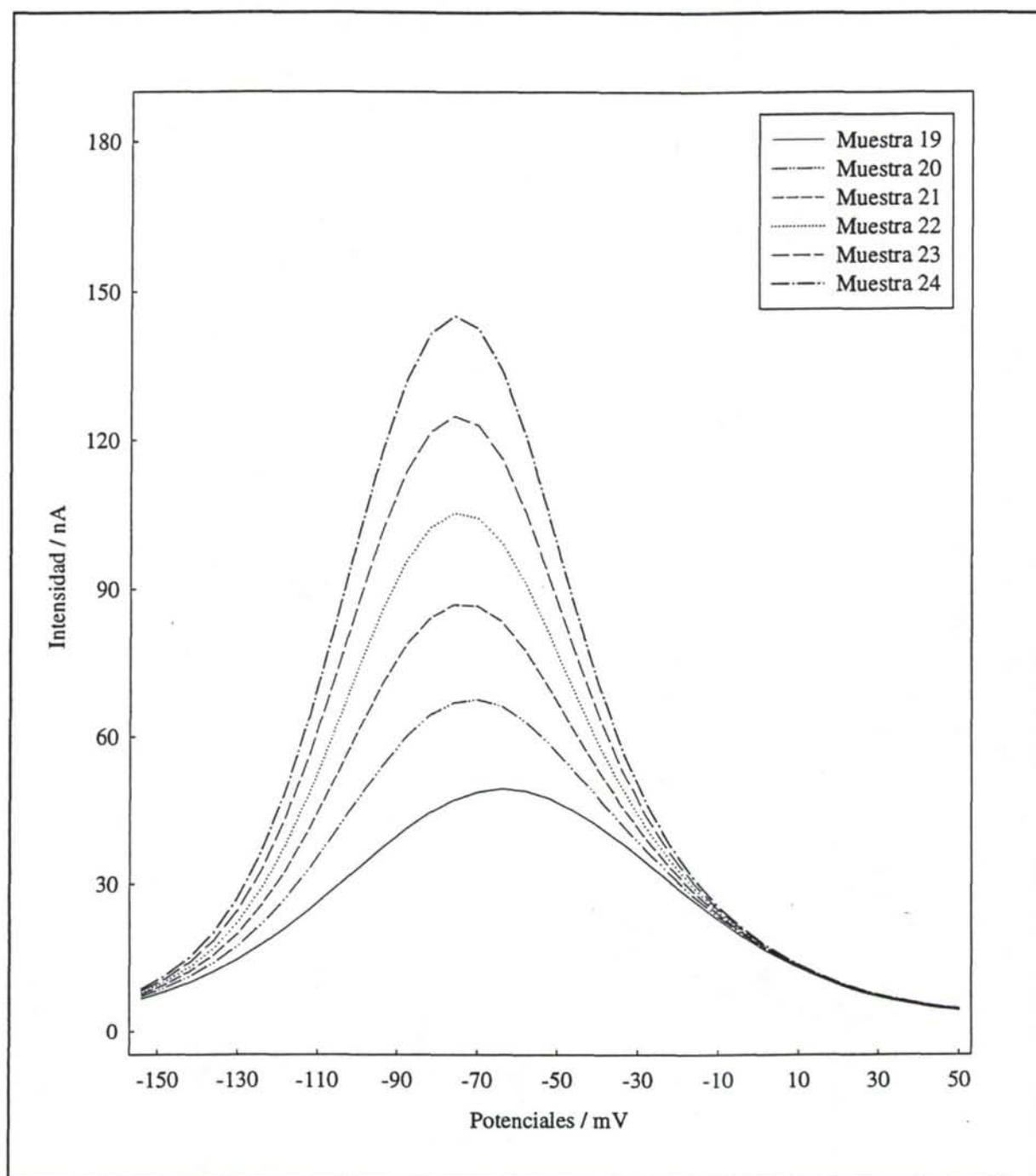


Figura III.7.2d Voltamperogramas correspondientes al calibrado J obtenidos a partir de una misma disolución inicial.

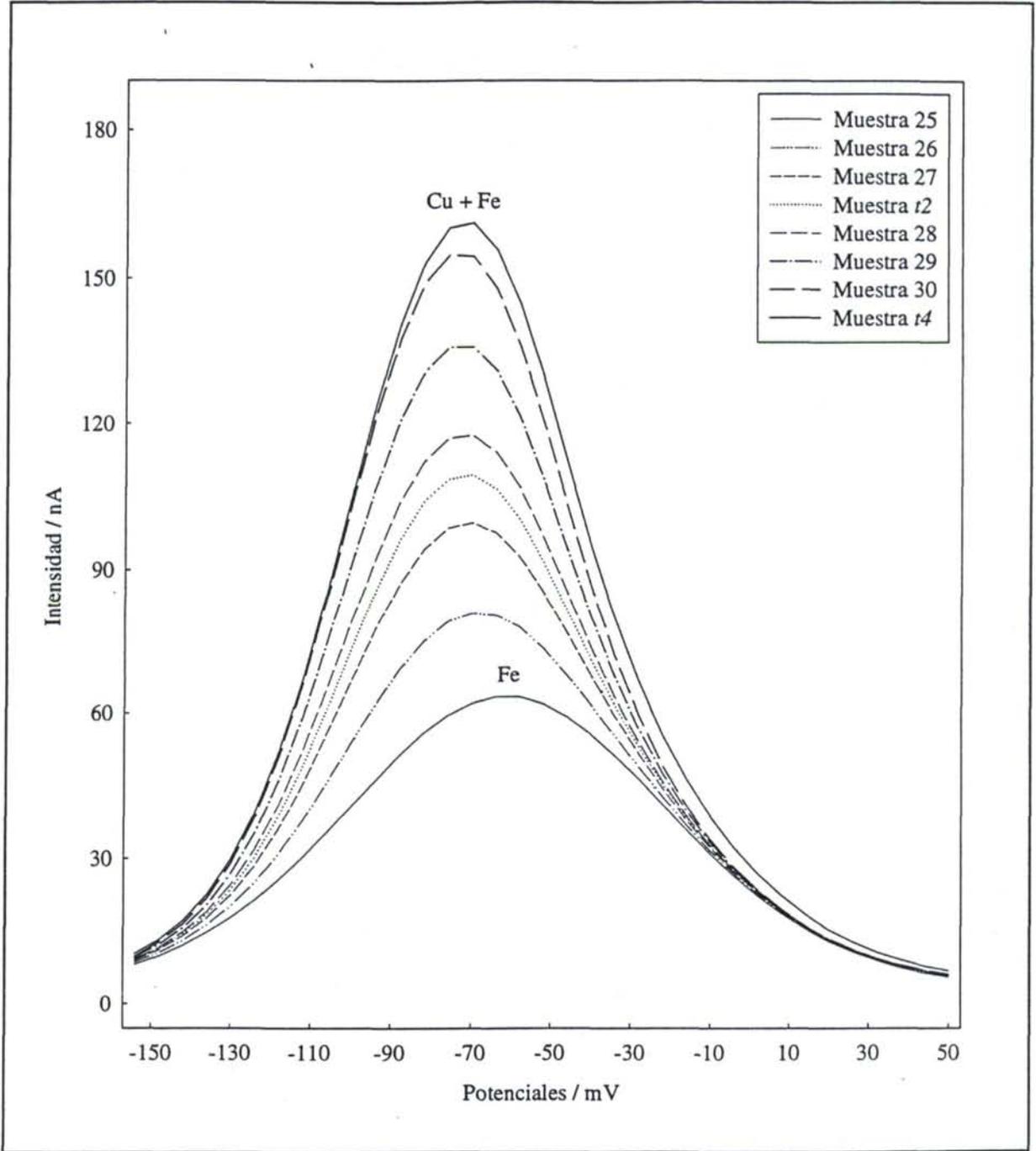


Figura III.7.2e Voltamperogramas correspondientes al calibrado J obtenidos a partir de una misma disolución inicial.

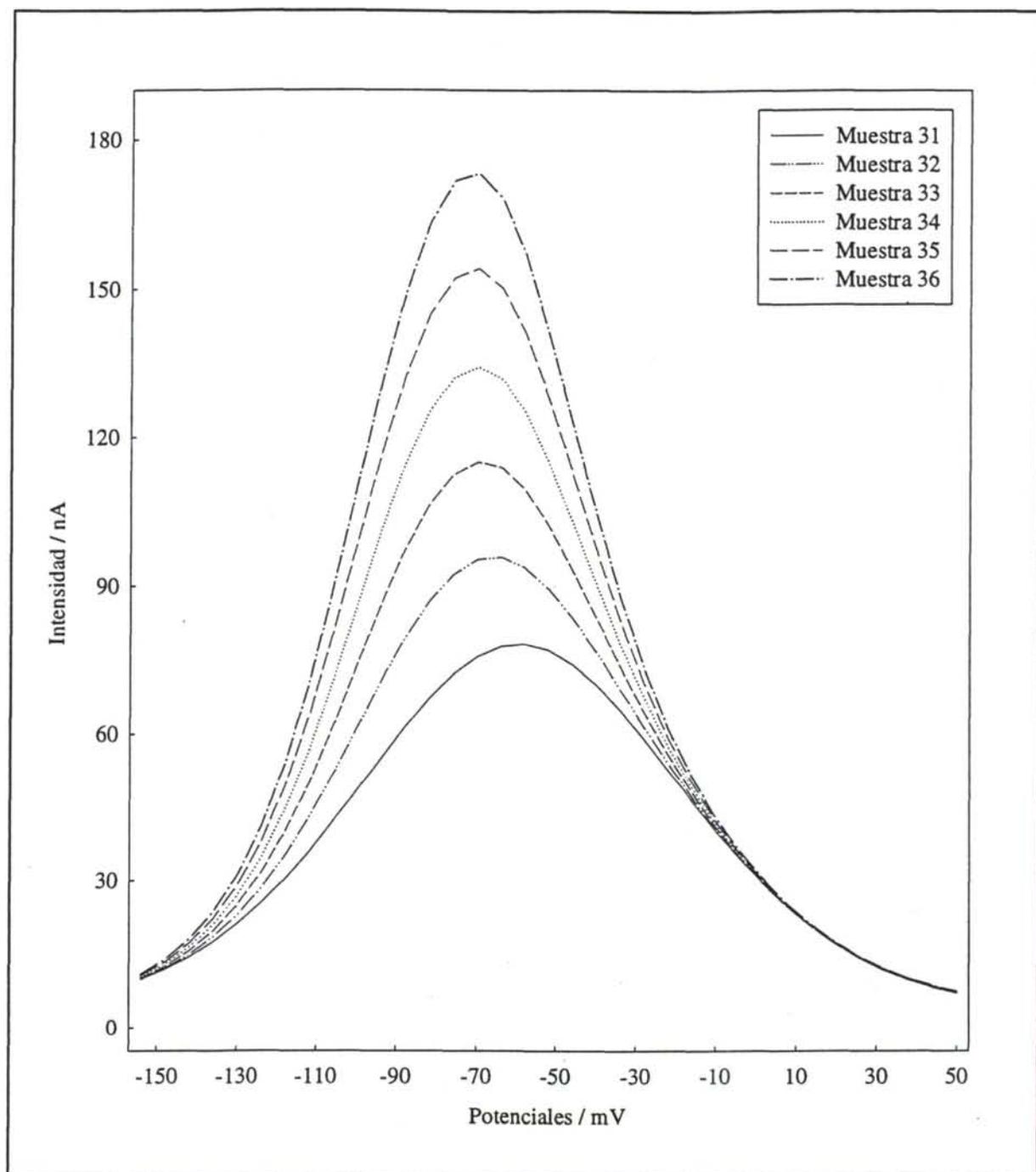


Figura III.7.2f Voltamperogramas correspondientes al calibrado J obtenidos a partir de una misma disolución inicial.

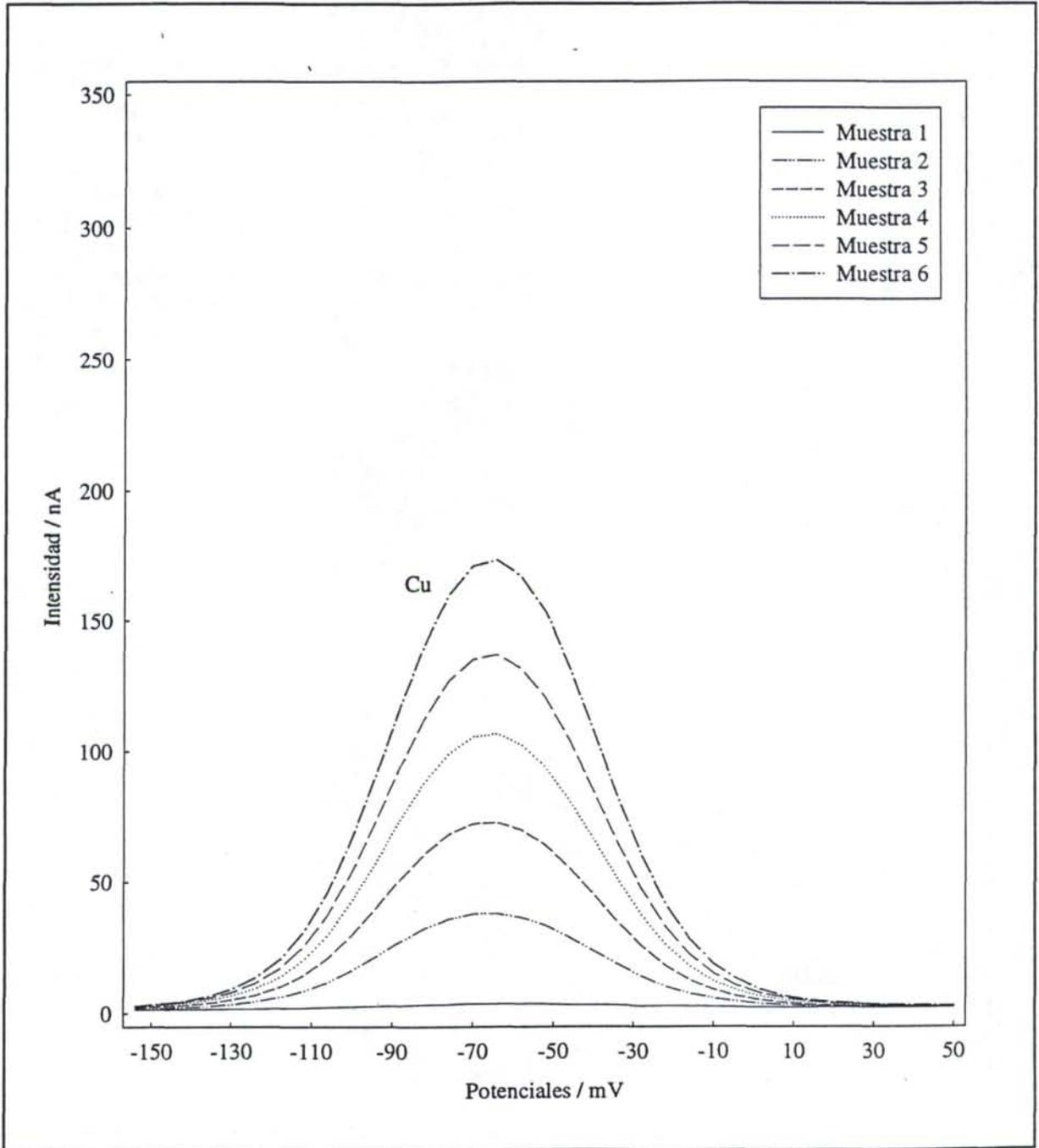


Figura III.7.3a Voltamperogramas correspondientes al calibrado K obtenidos a partir de una misma disolución inicial.

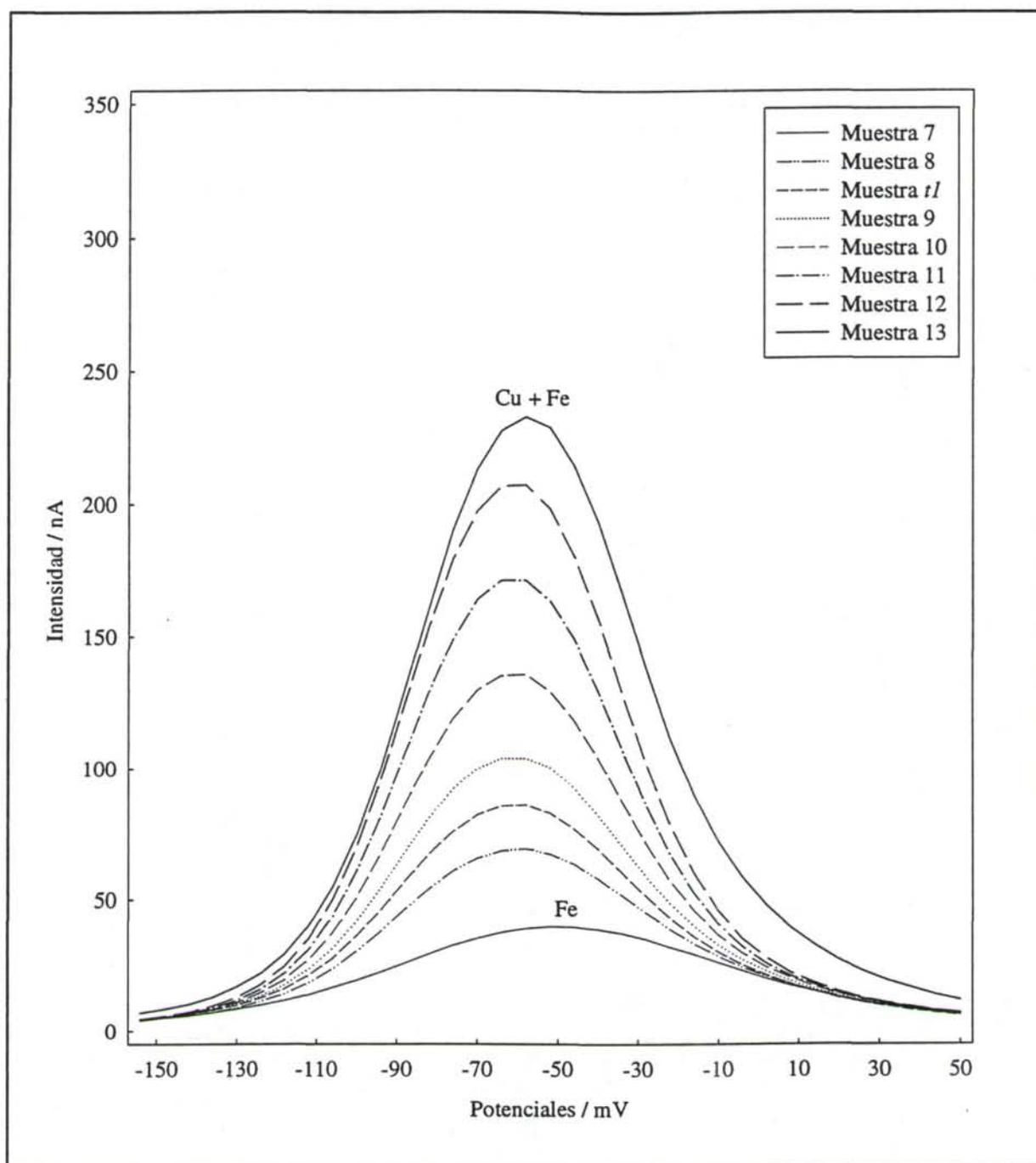


Figura III.7.3b Voltamperogramas correspondientes al calibrado K obtenidos a partir de una misma disolución inicial.

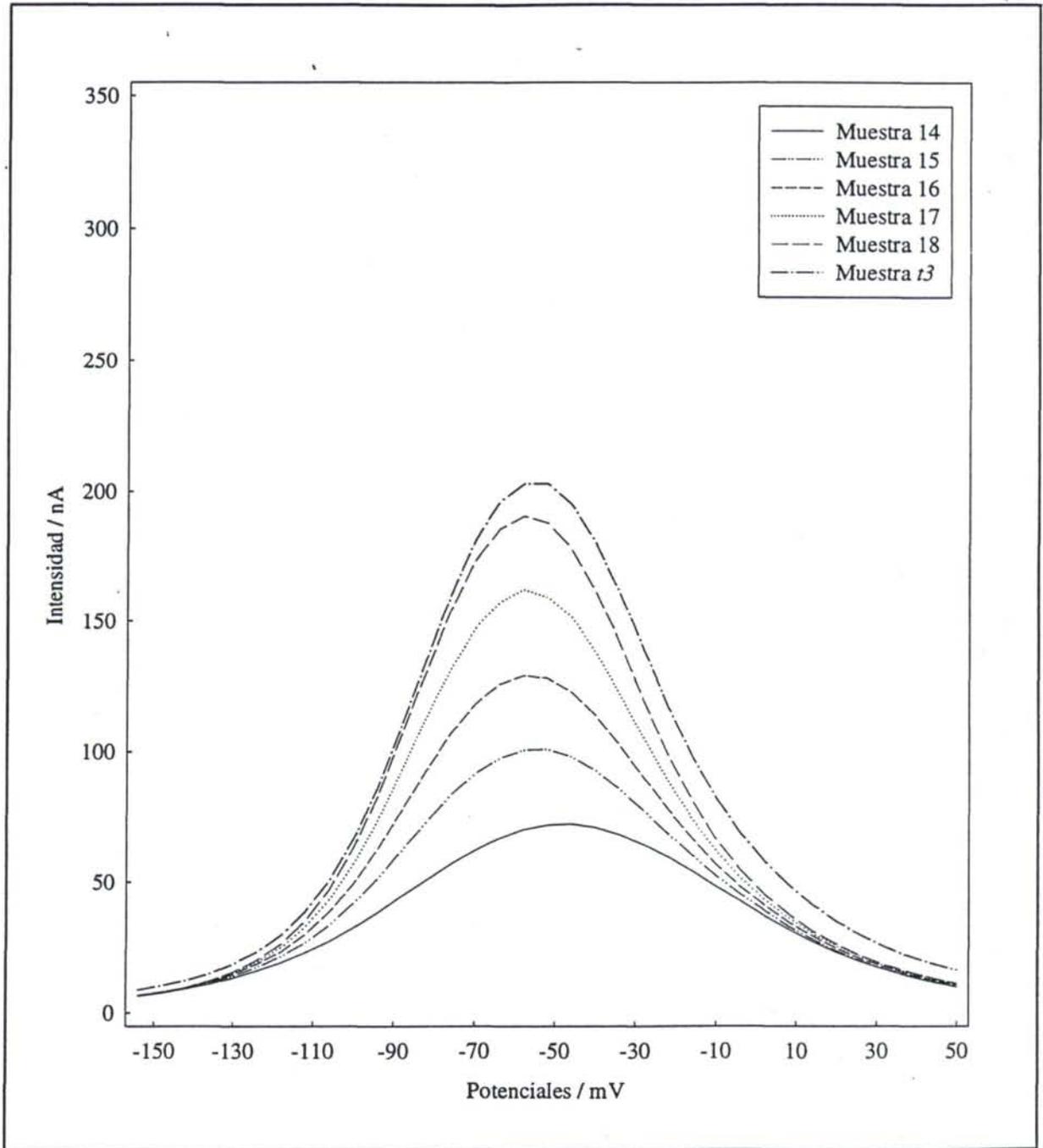


Figura III.7.3c Voltamperogramas correspondientes al calibrado K obtenidos a partir de una misma disolución inicial.

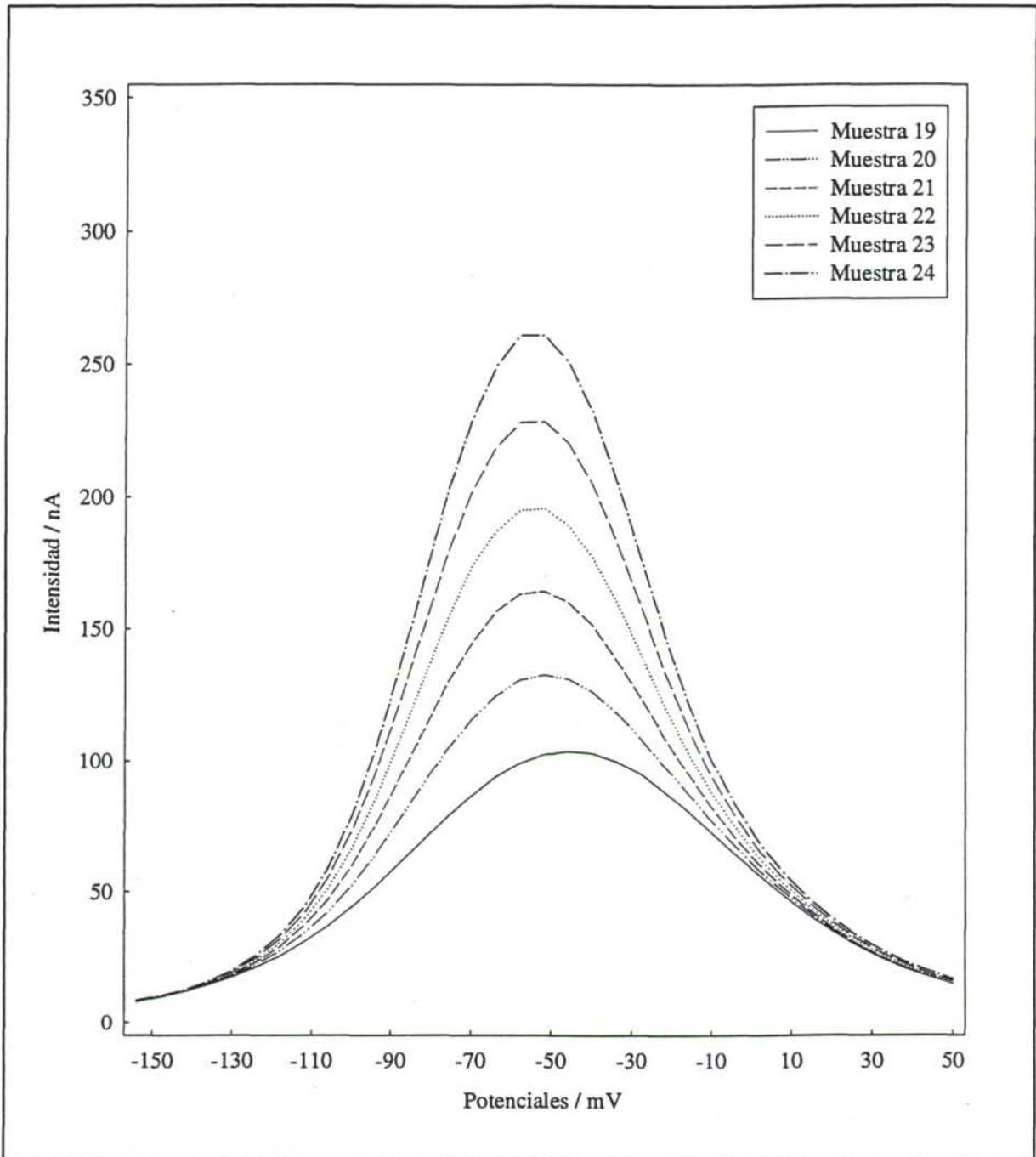


Figura III.7.3d Voltamperogramas correspondientes al calibrado K obtenidos a partir de una misma disolución inicial.

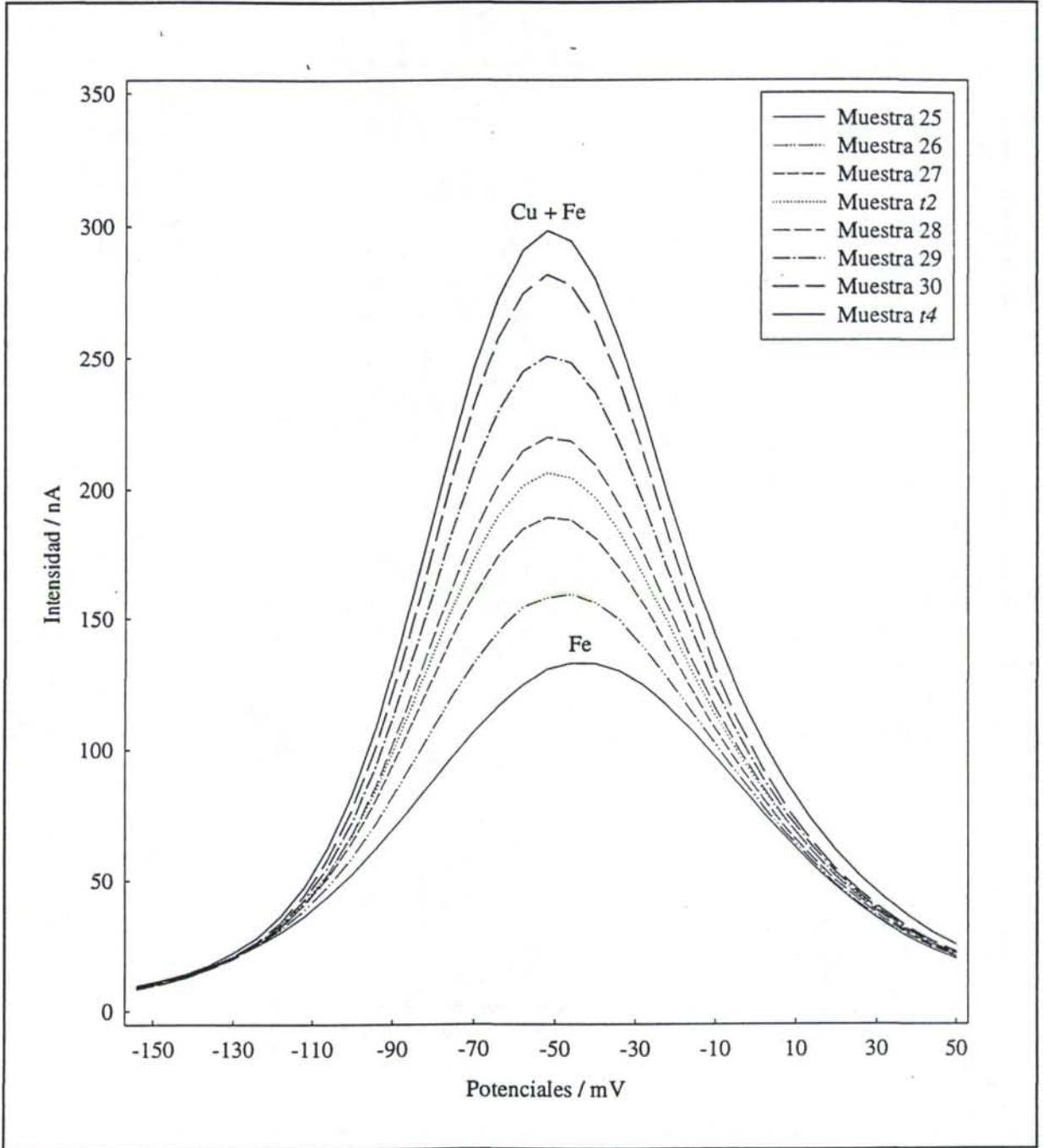


Figura III.7.3e Voltamperogramas correspondientes al calibrado K obtenidos a partir de una misma disolución inicial.

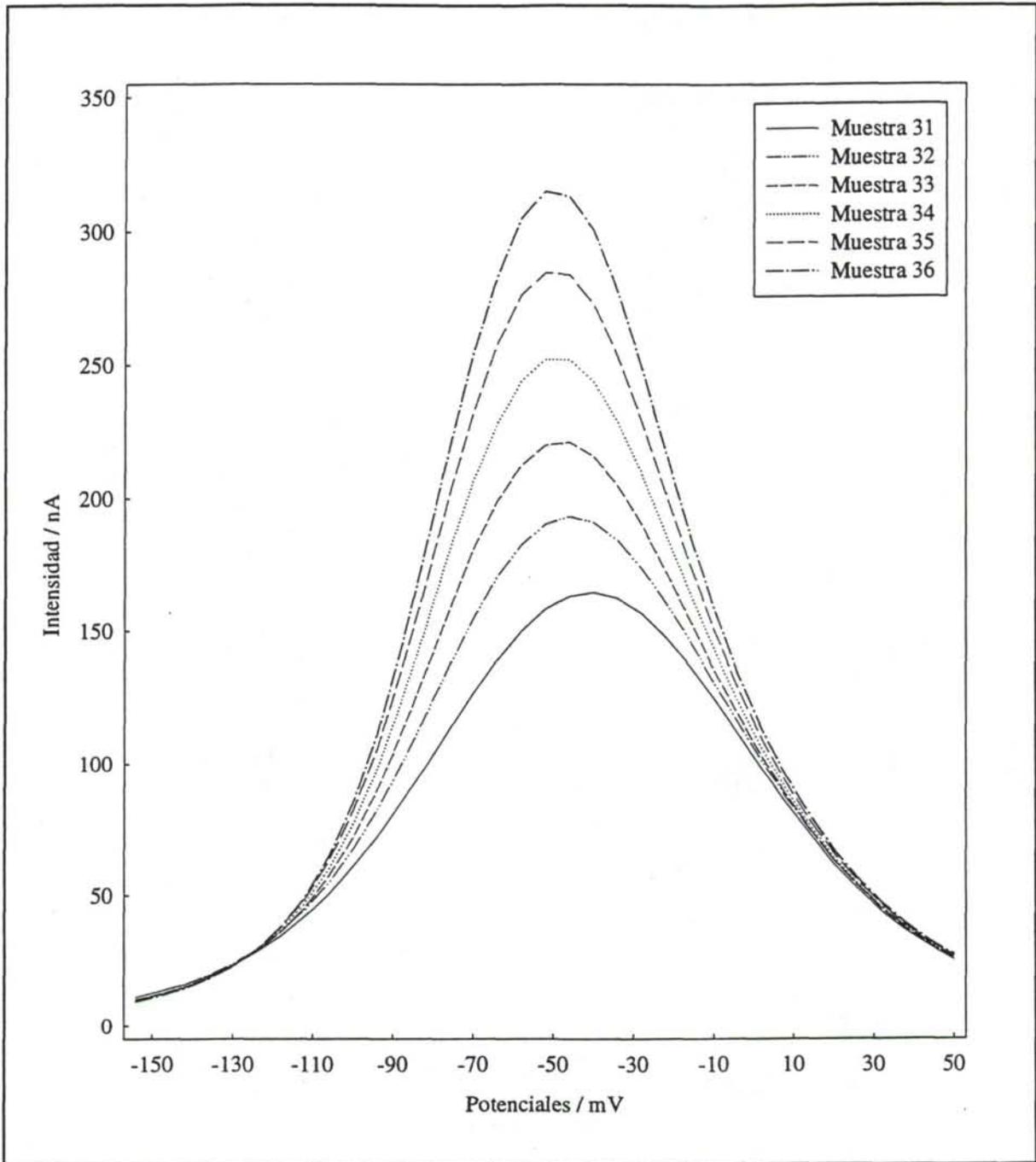


Figura III.7.3f Voltamperogramas correspondientes al calibrado K obtenidos a partir de una misma disolución inicial.

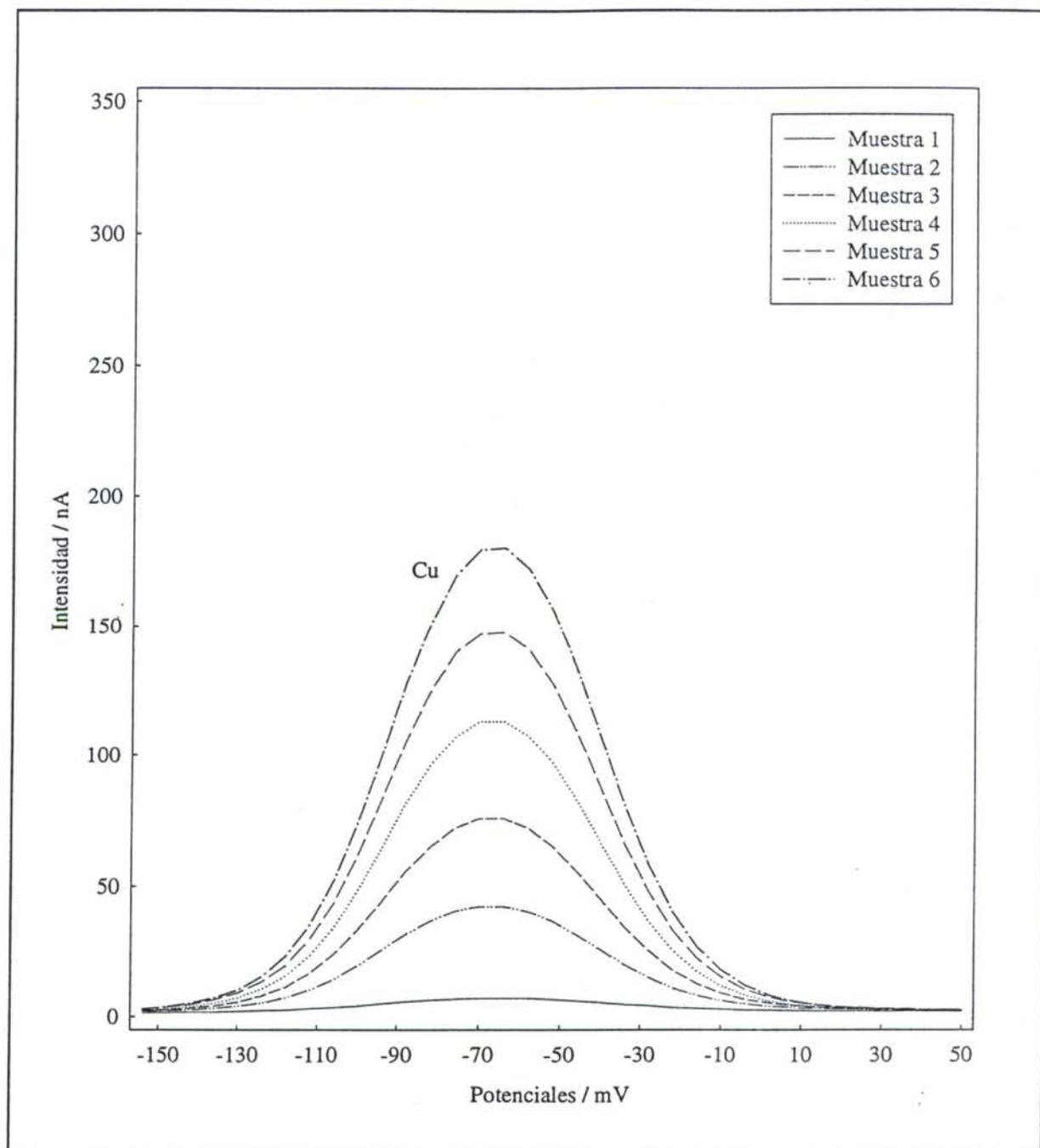


Figura III.7.4a Voltamperogramas correspondientes al calibrado L obtenidos a partir de una misma disolución inicial.

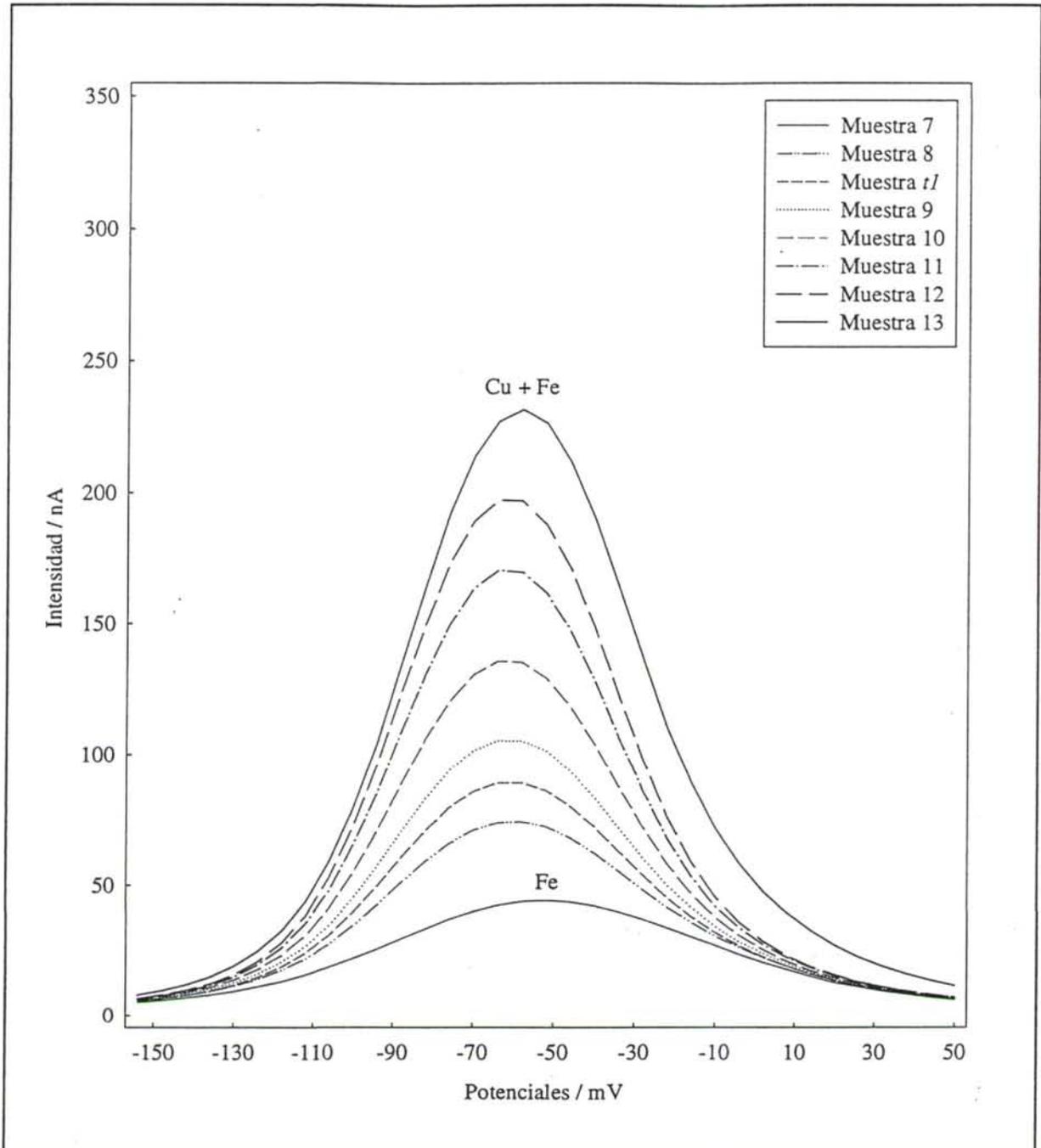


Figura III.7.4b Voltamperogramas correspondientes al calibrado L obtenidos a partir de una misma disolución inicial.

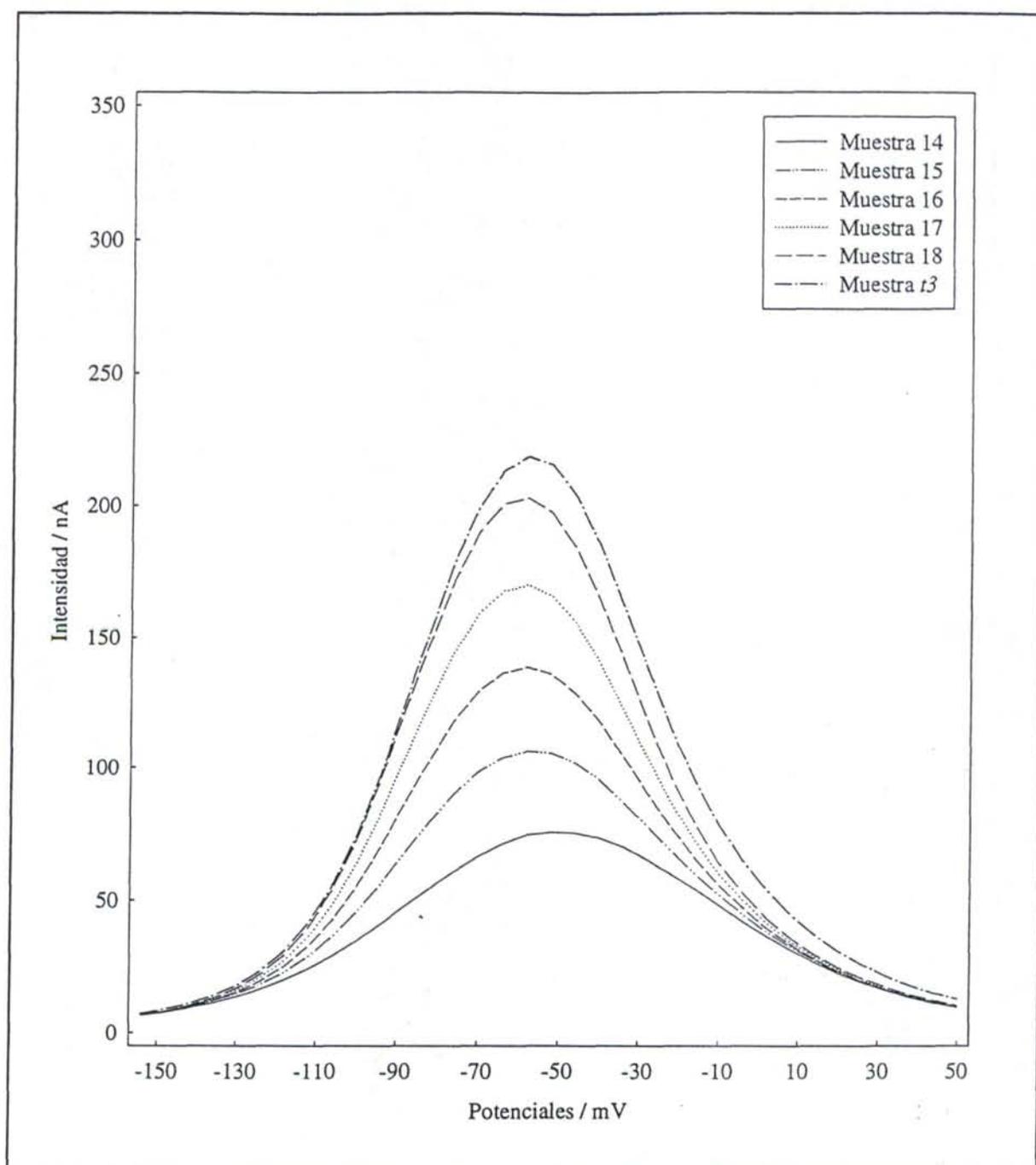


Figura III.7.4c Voltamperogramas correspondientes al calibrado L obtenidos a partir de una misma disolución inicial.

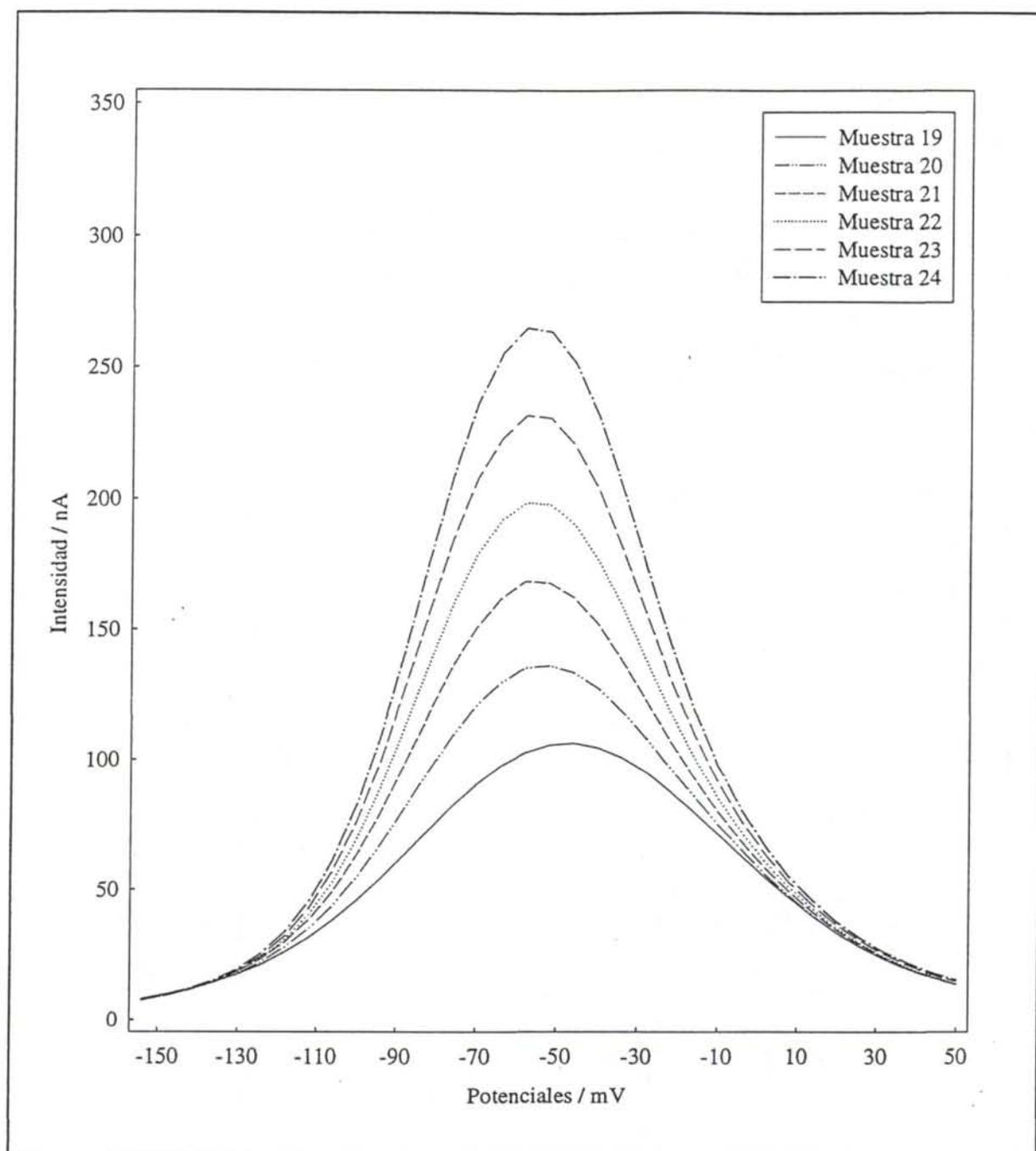


Figura III.7.4d Voltamperogramas correspondientes al calibrado L obtenidos a partir de una misma disolución inicial.

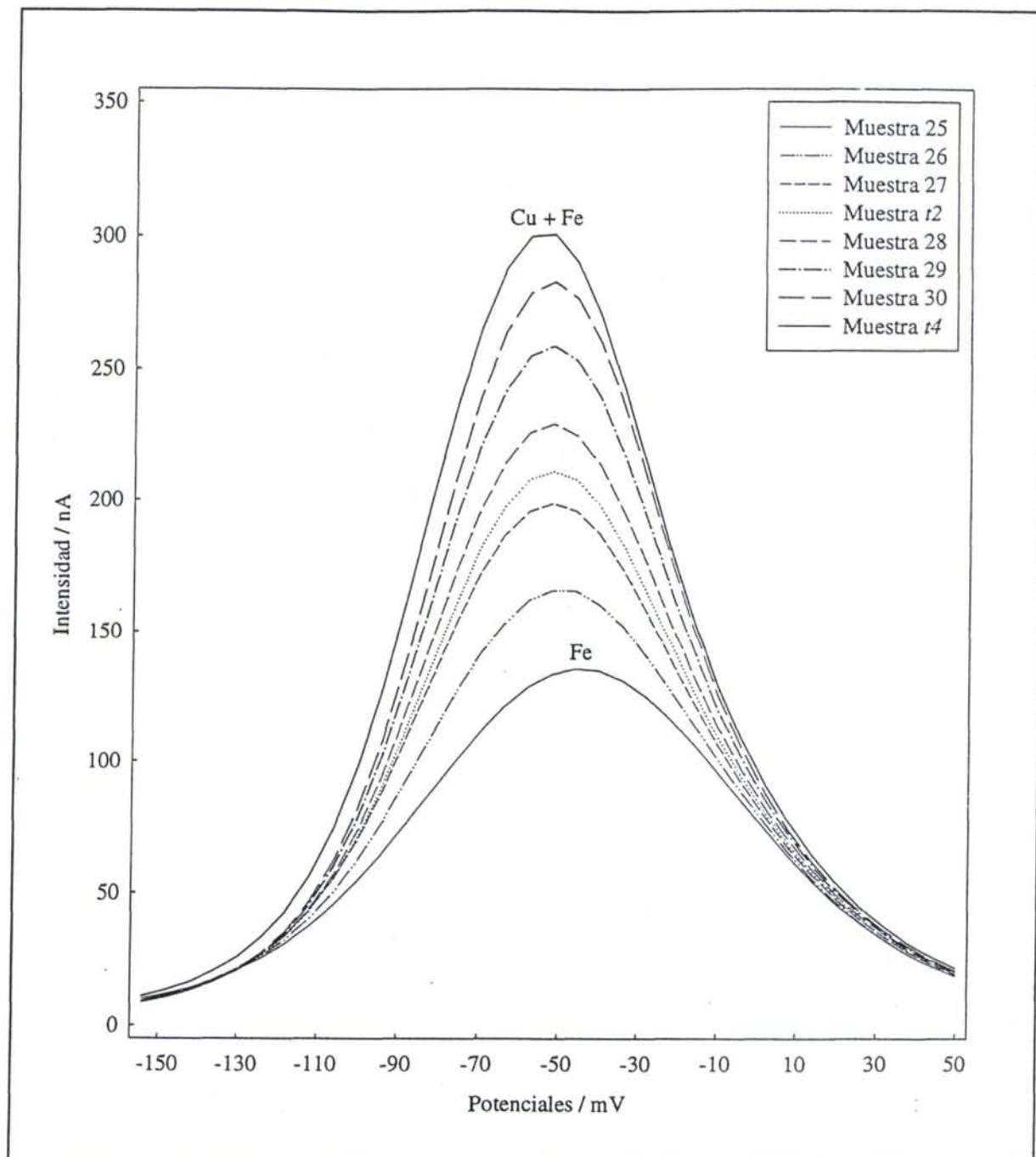


Figura III.7.4e Voltamperogramas correspondientes al calibrado L obtenidos a partir de una misma disolución inicial.

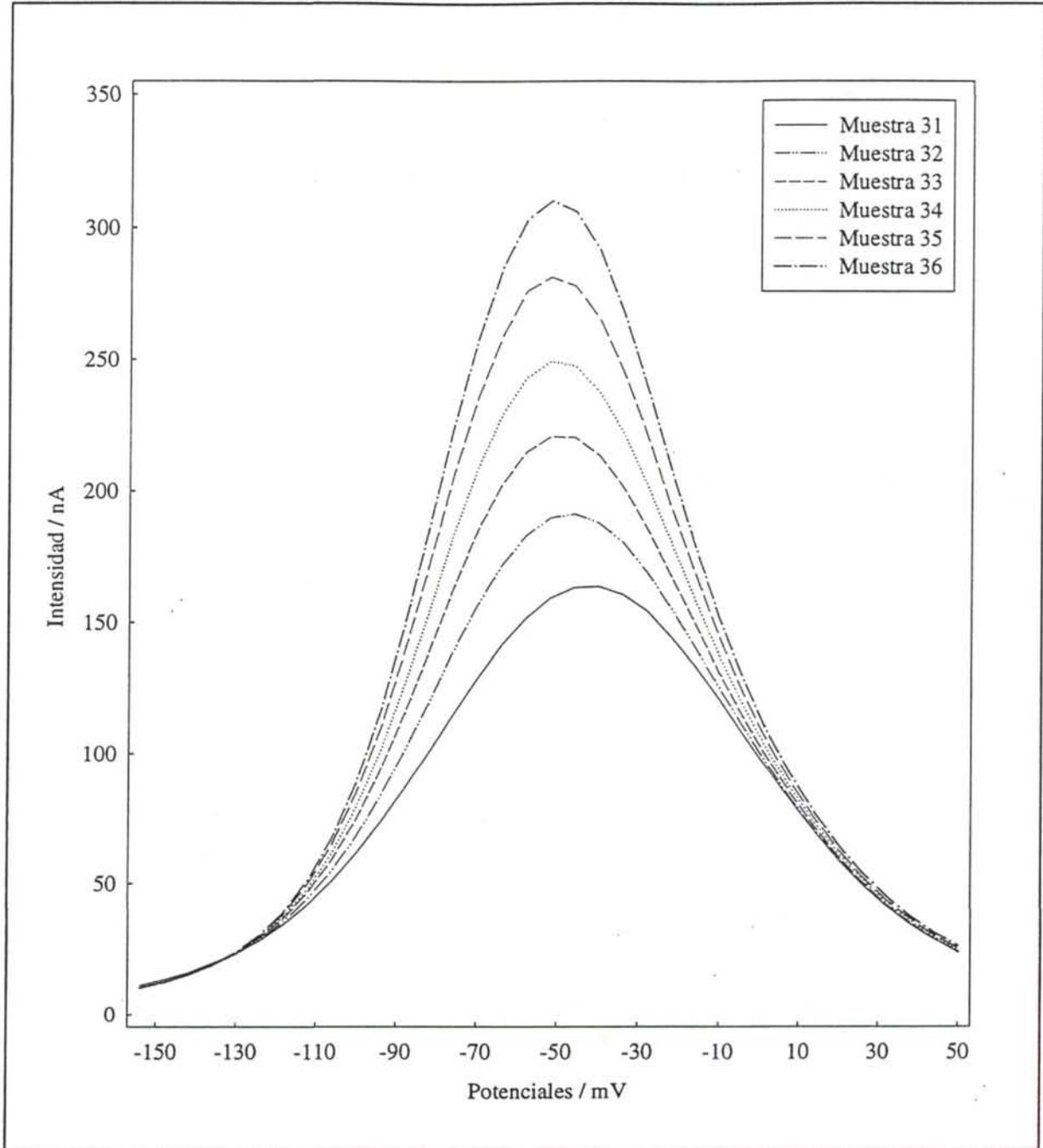


Figura III.7.4f Voltamperogramas correspondientes al calibrado L obtenidos a partir de una misma disolución inicial.

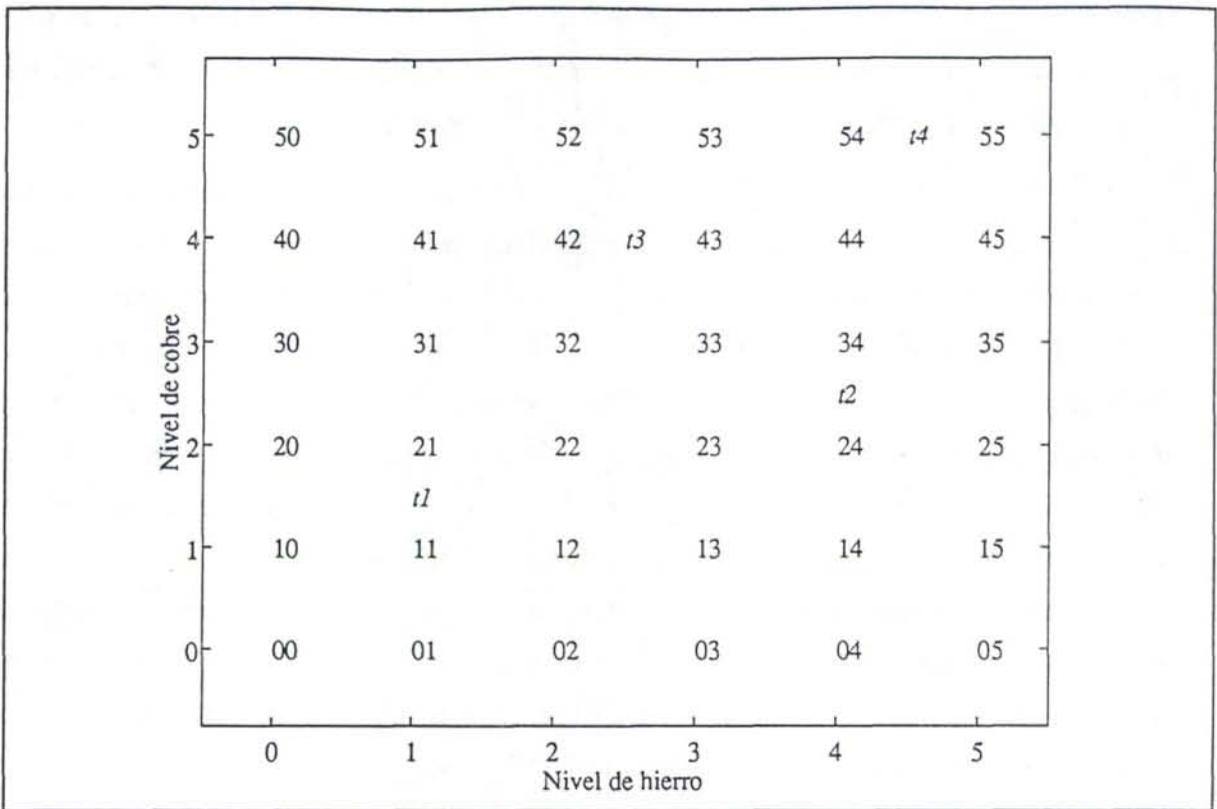


Figura III.7.4 Diseño de las muestras de los calibrados J, K y L. El primer dígito indica el nivel de cobre y el segundo de hierro. Muestras test: $t1$ - $t4$.

Se observa que la diferencia más significativa existente entre los voltamperogramas de los tres calibrados radica en la utilización de un instrumento u otro. Las intensidades registradas para el calibrado J, que se realizó utilizando el sistema μ AUTOLAB, son menores que las correspondientes a los otros dos calibrados, a la par que existe un desplazamiento de las señales respecto al eje de potenciales. Sin embargo, las diferencias entre los dos calibrados llevados a cabo con el mismo instrumento aunque en diferentes días, calibrados K y L, son de menor magnitud.

III.7.3 Análisis univariante

Se ha evaluado la viabilidad de un procedimiento clásico univariante en la resolución del efecto matriz que supone la presencia de hierro en la determinación de cobre según el procedimiento analítico indicado. Este procedimiento consiste, como se ha señalado

anteriormente, en sustraer a la señal global la contribución a la misma que se considera corresponde al interferente y después calcular la concentración del analito de interés en la muestra problema mediante la técnica de adición estándar [210].

La estrategia univariante se basa en considerar que, contando con el error experimental, la composición de la intensidad de pico es la suma algebraica de las intensidades correspondientes al cobre y al hierro si éstos se determinaran separadamente. La intensidad de pico del hierro debe ser independiente del tiempo de acumulación al no formar este elemento una amalgama sobre el electrodo de mercurio, y por tanto, es posible obtener el voltamperograma correspondiente al hierro realizando la medida voltamperométrica con tiempo de deposición cero. A continuación, este voltamperograma se sustrae simultáneamente de los correspondientes a la muestra y a los patrones, obteniéndose una intensidad de pico neta en la que se ha compensado la contribución del hierro y también la acumulación de cobre sobre el electrodo durante la etapa de barrido. Esas son las intensidades de pico netas que se utilizan posteriormente para determinar la concentración de cobre utilizando el método de adición estándar.

Para analizar este método univariante, en la Figura III.7.5 se representan 6 gráficas en las que se muestran los voltamperogramas correspondientes a 15 mL de electrolito soporte, 0.2 mL de Cu(II) 31.32 μ M y x adiciones de 0.1 mL de Fe(III) 3.132 mM (el número de adiciones se indica en cada gráfica entre paréntesis), los cuales se han obtenido utilizando el instrumento Metrohm 646 VA. Los voltamperogramas trazados con línea continua corresponden en cada caso a muestras con la composición indicada y cuya medida se ha realizado con un tiempo de acumulación de 60 s, mientras que aquellos trazados (— -) se corresponden con una medida realizada sobre la misma muestra con tiempo de acumulación cero. Por último, los voltamperogramas restantes (- -) se corresponden con muestras que únicamente contienen hierro al nivel indicado entre paréntesis en la Figura III.7.5, es decir, corresponden a la verdadera señal neta del hierro, la cual, como se puede observar, no coincide en ninguno de los casos con el voltamperograma obtenido con tiempo de acumulación cero para un mismo nivel de hierro en presencia de cobre.

Las discrepancias observadas entre los voltamperogramas correspondientes al hierro solo y los correspondientes a muestras con hierro y cobre, registrados todos ellos con tiempo de acumulación cero, pueden deberse al error experimental a que está sujeta cada una de las medidas realizadas, o también a que una pequeña cantidad de cobre podría amalgamarse sobre el electrodo durante la etapa de medida. Incluso, en cierta medida, es posible que la

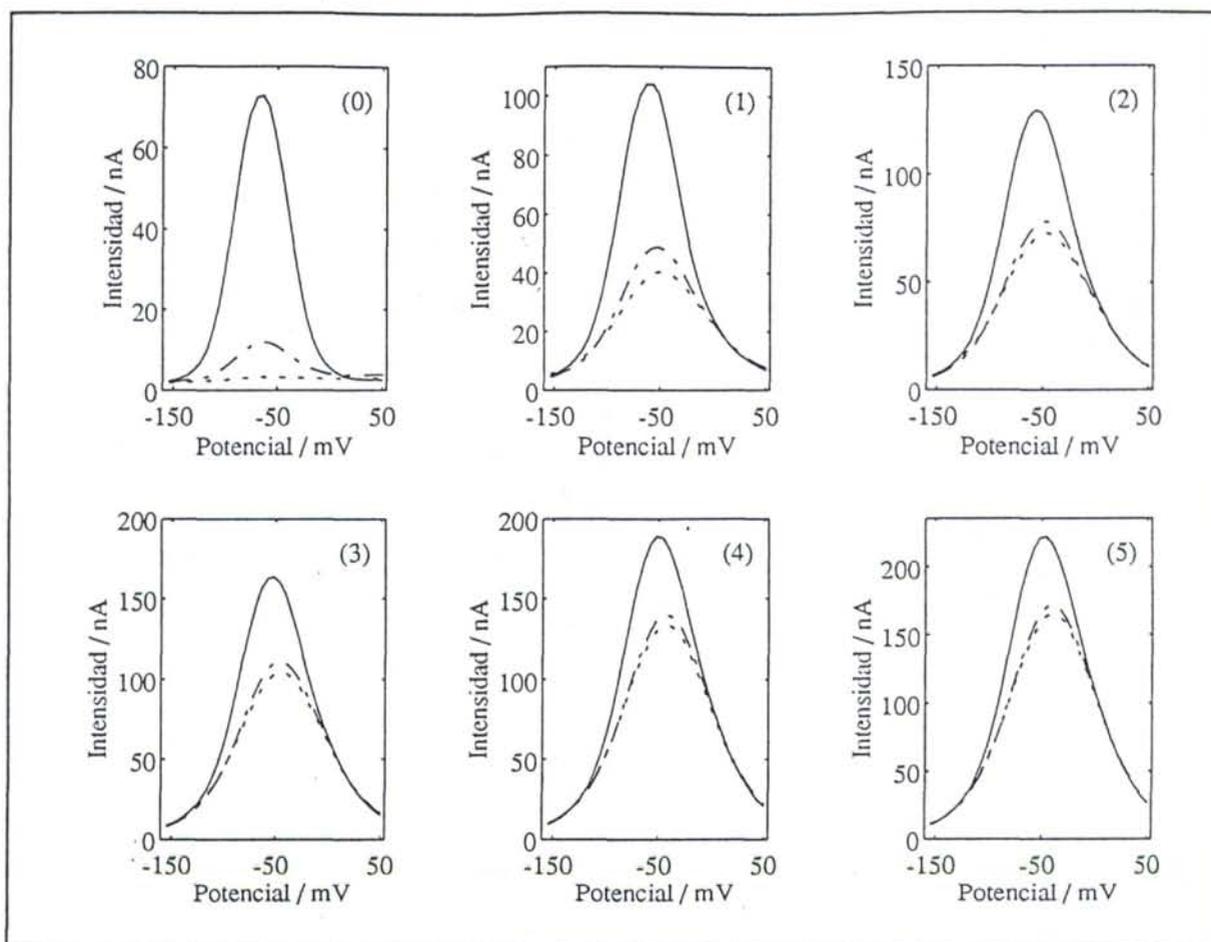


Figura III.7.5 Voltamperogramas de muestras con Cu y Fe (nivel de Fe): $t_{\text{acum}} = 60$ s (—), $t_{\text{acum}} = 0$ s (---) y (- - , sólo hierro).

concentración de cobre sea suficientemente alta para dar lugar a una señal mínima que interfiera en el análisis, aunque ésta pueda considerarse excesivamente pequeña para ser utilizada directamente en la determinación del analito, en cuyo caso, se requiere una técnica de análisis más sensible como son las técnicas de acumulación. No obstante, hay que considerar que en la práctica, cuando se analiza una muestra problema, no se dispone del voltamperograma neto correspondiente a la cantidad de hierro que contiene la misma, y por tanto no es posible establecer estas comparaciones para evaluar el error que puede estar asociado con el procedimiento univariante que se está utilizando.

Por otro lado, como las intensidades de pico del analito de interés y del interferente aparecen a potenciales separados ligeramente, al restar la intensidad de pico obtenida en la medida directa (tiempo de acumulación cero) de las obtenidas con un tiempo de acumulación determinado ocurre que las intensidades de pico que se restan no se corresponden exactamente

con los mismos valores de potencial, lo cual puede resultar erróneo ya que, como es sabido, la intensidad total es la suma algebraica de las contribuciones de los analitos implicados cuando se considera un valor de potencial determinado.

Con el fin de evaluar la calidad de las determinaciones univariantes se ha aplicado el procedimiento especificado previamente para calcular la concentración de cobre en las muestras que aparecen en la Figura III.7.5 a tres niveles de hierro (0, 3 y 5). En las condiciones experimentales en las que se realizaron los calibrados K y L se han llevado a cabo dos medidas en cada una de esas disoluciones, acumulando durante 60 s y sin acumular, y a continuación se han realizado seis adiciones estándar de cobre. El voltamperograma obtenido para la muestra inicial con tiempo de acumulación cero se ha restado de todos los demás para cada nivel de hierro, dando como resultado las señales que posteriormente se han utilizado en el análisis por adición estándar cuyos resultados se muestran en la Tabla III.7.2. Se trata de rectas con buenos coeficientes de correlación e intervalos de confianza pequeños para pendiente y término independiente, pero que sin embargo conducen a determinaciones de la concentración de cobre con errores por encima del 25%.

Tabla III.7.2 Resultados del procedimiento univariante de adición estándar realizado a tres niveles diferentes de concentración de hierro.

Regresión	Hierro		
	Nivel 0	Nivel 3	Nivel 5
Pendiente	176.46 ± 8.33	163.31 ± 2.38	166.91 ± 5.60
Término independiente	57.47 ± 6.00	52.42 ± 1.68	48.25 ± 3.91
R^2_{ajustado}	0.9980	0.9998	0.9990
S_{yx}	3.3975	0.9529	2.2127
Concentración de cobre real (μM)	0.4121	0.4041	0.3990
Concentración de cobre calculada (μM)	0.3257	0.3210	0.2891
Error relativo (%)	26.53	25.89	38.01

En vista de estos resultados, este procedimiento no se puede considerar adecuado para la determinación de cobre en presencia de hierro a las concentraciones indicadas, y por tanto, sería necesario encontrar un procedimiento que de forma general, sin depender tanto de las concentraciones de cada uno de los analito, fuera adecuado para llevar a cabo la determinación de cobre de forma satisfactoria.

En este sentido se puede considerar la posibilidad de llevar a cabo un calibrado univariante, el cual, a diferencia del procedimiento de adición estándar descrito, podría utilizarse para determinar la concentración de cobre en toda una batería de muestras problema. Sin embargo, en este caso la limitación del método reside en que sería necesario conocer la concentración exacta de hierro en esas muestras para así poderla incluir en cada uno de los patrones de calibrado, lo cual reduce de nuevo sensiblemente el número de análisis que se podrían realizar con un mismo calibrado y complica enormemente el análisis.

III.7.4 Análisis multivariante

La determinación de cobre en presencia de hierro por DPASV también puede abordarse considerando la posibilidad de modelar la interferencia originada por el efecto matriz, es decir, introducir el propio interferente en el diseño experimental de forma que en un momento dado el modelo matemático construido sea capaz de discriminar qué contribución a la señal total tiene el interferente. Con esta finalidad se han desarrollado modelos de regresión PLS para cada uno de los tres calibrados de que se dispone, en cada uno de los cuales se recorren 6 niveles de concentración diferentes para cada uno de los analitos.

Por otro lado, esta metodología permite determinar la concentración de cobre y hierro de forma simultánea, sin que ello suponga introducir mayor esfuerzo experimental o de cálculo. La herramienta necesaria para determinar la concentración de cobre en presencia de hierro permite también calcular la concentración de este último elemento.

III.7.4.1 Modelos PLS

Construcción de los modelos PLS

Los modelos de calibrado PLS se han construido considerando como variables predictoras las intensidades registradas a 35 valores de potencial igualmente espaciados desde -0.154 V hasta 0.050 V, y como variables respuesta las concentraciones de los dos analitos (el analito de interés y el interferente) que se muestran en la Tabla III.7.1. Las matrices de datos constan por tanto de un total de 40 objetos, 36 de ellos pertenecientes al conjunto de aprendizaje y

4 al conjunto test (los señalados como $t1$, $t2$, $t3$ y $t4$ respectivamente), de los cuales se dispone de 35 variables predictoras. Es decir, las matrices **J**, **K** y **L** tienen de dimensiones 40×35 .

Dichos modelos han sido estimados partiendo de los datos originales (sin normalizar), utilizando un método de validación cruzada [22, 23] que permite evaluar la estabilidad de los modelos construídos y su capacidad de predicción. Al igual que en apartados anteriores, los modelos PLS se han construído de forma independiente para cada uno de los dos analitos en cada calibrado.

Evaluación de los modelos PLS

La validez de los diferentes modelos PLS construídos para los calibrados J, K y L, tanto para el cobre como para el hierro, queda reflejada en los valores de varianza explicada y varianza en predicción que aparecen en las Tablas III.7.3, III.7.4 y III.7.5 respectivamente. Los scores representados en las Figuras III.7.6 y III.7.7 y los loadings que se muestran en las Figuras III.7.8 y III.7.9, correspondientes a las diferentes variables latentes que constituyen cada uno de los modelos PLS del calibrado J, son similares a los de los modelos PLS construídos con los calibrados K y L.

En el plano formado por las puntuaciones (scores) correspondientes a las dos primeras variables latentes de las Figuras III.7.6 y III.7.7 se refleja la existencia de cierto comportamiento no lineal. Éste se muestra en la curvatura que presentan los scores de ambos analitos, la cual distorsiona una estructura que reproduce el diseño experimental y apunta a una posible interferencia mutua entre los dos analitos. Estas dos variables latentes modelan la mayor parte de la variabilidad, tanto de la respuesta (Bloque Y) como de las variables predictoras (Bloque X).

El valor del score en la segunda variable latente aumenta para todas las muestras con la concentración de cobre en el modelo PLS construído para dicho analito, mientras que la variación de este score tiene signo contrario cuando se incrementa la concentración de hierro, es decir, de alguna forma la segunda variable latente está restando la contribución del hierro en el modelo PLS del cobre, modelando el efecto matriz producido. Esto se confirma a través de los loadings de la segunda componente, los cuales tienen valor positivo en aquellos potenciales que se corresponden con la señal del cobre y negativo en los correspondientes al hierro.

Tabla III.7.3 Porcentajes de varianza explicada en función del número de variables latentes (V.L.) seleccionadas en cada uno de los modelos construidos en el calibrado J.

V. L.	Cobre			Hierro		
	Bloque Y		Bloque X	Bloque Y		Bloque X
	Var. explicada	Var. en predicción	Var. explicada	Var. explicada	Var. en predicción	Var. explicada
1	50.54	48.25	86.31	49.53	47.05	86.31
2	97.92	97.74	99.01	98.39	98.16	97.01
3	99.71	99.73	99.43	99.79	99.80	99.43
4	99.88	99.88	99.80	99.91	99.90	99.80
5	99.89	99.89	99.90	99.98	99.98	99.90
6	99.92	99.89	99.91	99.99	99.99	99.92
7	99.93	99.88	99.93	100.00	99.99	99.93
8	99.96	99.85	99.94	100.00	99.99	99.93

Tabla III.7.4 Porcentajes de varianza explicada en función del número de variables latentes (V.L.) seleccionadas en cada uno de los modelos construidos en el calibrado K.

V. L.	Cobre			Hierro		
	Bloque Y		Bloque X	Bloque Y		Bloque X
	Var. explicada	Var. en predicción	Var. explicada	Var. explicada	Var. en predicción	Var. explicada
1	37.94	34.37	86.02	60.75	58.79	86.03
2	93.18	92.24	95.90	96.88	96.41	95.91
3	98.98	99.04	99.23	99.67	99.68	99.23
4	99.48	99.44	99.69	99.75	99.72	99.69
5	99.74	99.70	99.73	99.93	99.90	99.76
6	99.76	99.75	99.86	99.98	99.97	99.86
7	99.85	99.66	99.86	99.98	99.96	99.87
8	99.87	99.62	99.87	99.99	99.95	99.87

Tabla III.7.5 Porcentajes de varianza explicada en función del número de variables latentes (V.L.) seleccionadas en cada uno de los modelos construidos en el calibrado L.

V. L.	Cobre			Hierro		
	Bloque Y		Bloque X	Bloque Y		Bloque X
	Var. explicada	Var. en predicción	Var. explicada	Var. explicada	Var. en predicción	Var. explicada
1	39.03	35.46	86.31	59.42	57.67	86.32
2	91.58	90.64	96.03	96.95	96.49	96.04
3	98.64	98.72	99.17	99.59	99.59	99.18
4	99.02	99.06	99.71	99.89	99.83	99.20
5	99.39	99.35	99.83	99.90	99.89	99.82
6	99.74	99.35	99.84	99.95	99.95	99.87
7	99.82	99.63	99.88	99.98	99.95	99.88
8	99.92	99.60	99.88	99.99	99.94	99.88

La tercera variable latente parece tener un efecto de compensación sobre las muestras que se encuentran en el primer plano factorial en la zona en la que aparece distorsionado el diseño experimental. Es decir, esta variable latente mejora la estimación de la aportación de cada analito a la señal total al considerar direcciones de varianza ortogonales a las recogidas en las dos variables latentes anteriores. Asimismo, el resto de las variables latentes tienen una estructura clara en este mismo sentido, modelan nuevas fuentes de variabilidad, aunque la magnitud de los scores correspondientes a las mismas va disminuyendo a medida que se considera mayor número de variables latentes, y su aportación a la varianza explicada va siendo menos importante.

Teniendo en cuenta los porcentajes de varianza explicada de las Tablas III.7.3, III.7.4 y III.7.5 se han seleccionado 5 y 7 variables latentes en los modelos correspondientes al cobre y al hierro en el calibrado J, 6 en los modelos construidos para ambos analitos en el calibrado K y 7 en los dos modelos correspondientes al calibrado L. En todos ellos, los porcentajes de varianza explicada son elevados y se encuentran muy próximos a los valores de varianza en predicción, lo cual señala además la estabilidad de los modelos construidos.

Las concentraciones calculadas por los modelos PLS y el error asociado con ellas se encuentran en las Tablas III.7.6, III.7.7 y III.7.8, en las cuales se anotan también los errores

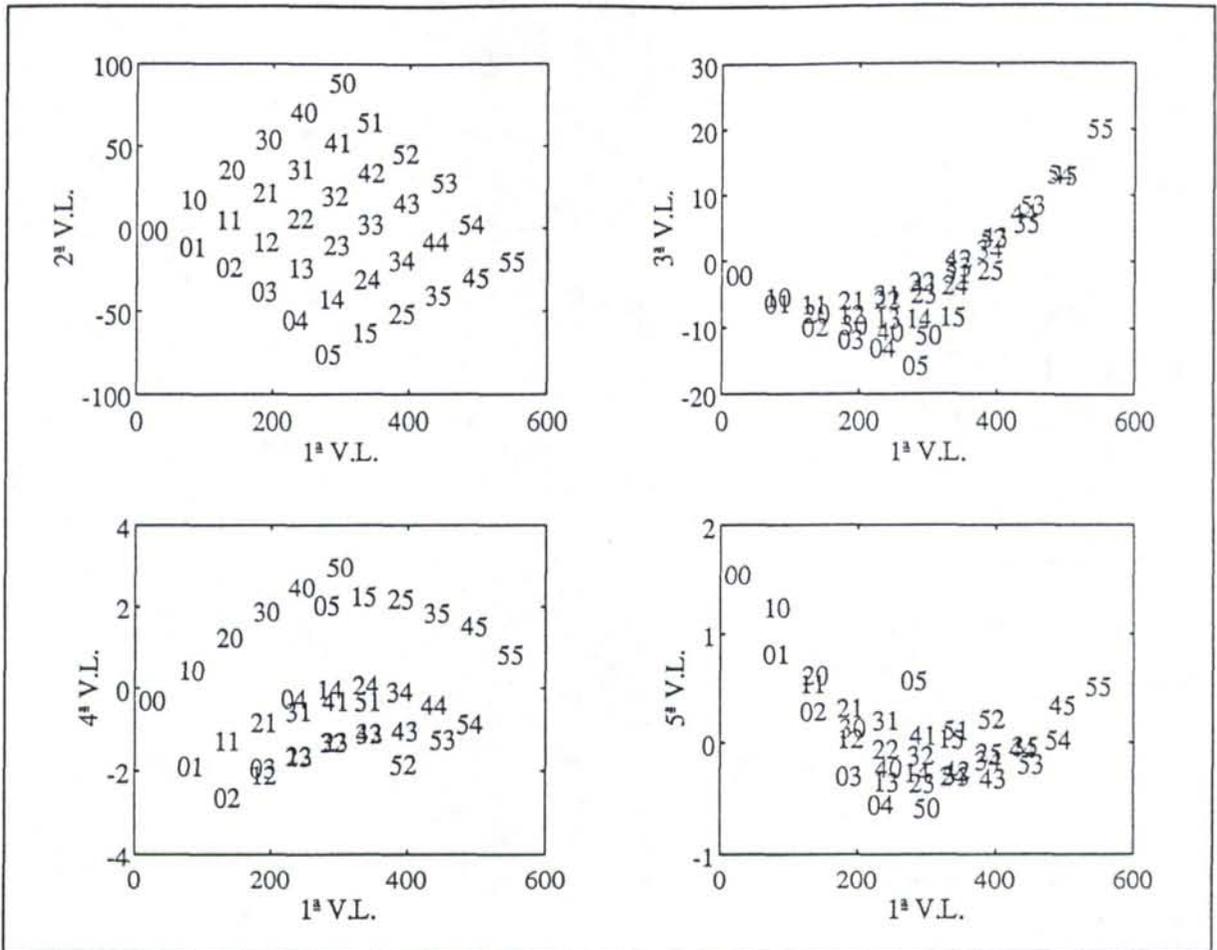


Figura III.7.6 Scores del modelo PLS construido para la determinación de cobre en el calibrado J.

correspondientes a las muestras test. En este caso, se presentan los resultados correspondientes al conjunto test y al resto del calibrado porque en análisis posteriores se tendrán en consideración.

Considerando estos errores se observa que en todos los casos el rango de los mismos es menor para el hierro que para el cobre. Esto se debe a que el hierro es el elemento que se encuentra en mayor concentración y a que éste no tiende a acumularse sobre el electrodo de trabajo, sino que se mide directamente de la disolución, por lo que este analito está desvinculado de la variabilidad asociada con la etapa de preconcentración, que es una de las fuentes de variabilidad más importante de los análisis electroquímicos con redisolución.

Por otro lado, se hace evidente que los errores asociados con el calibrado J son sensiblemente inferiores a los relacionados con los otros dos. Esto podría deberse a que el primer calibrado se llevó a cabo en un instrumento diferente al usado en los otros dos casos.

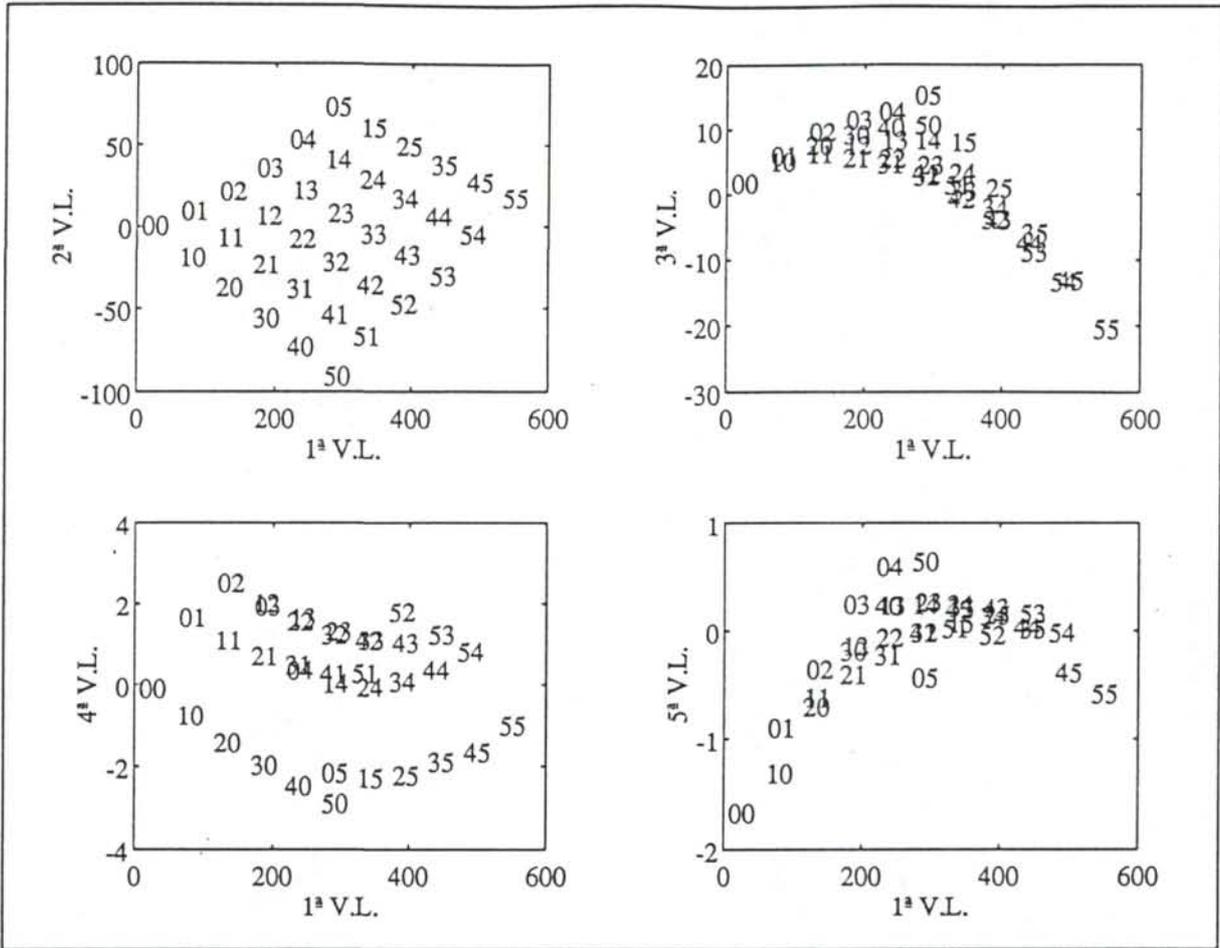


Figura III.7.7 Scores del modelo PLS construido para la determinación de hierro en el calibrado J.

III.7.4.2 Transferencia de calibrado mediante el procedimiento PDS

En vista de los resultados del análisis multivariante realizado con los tres calibrados de que se dispone, se decidió llevar a cabo la transferencia de calibrado entre el calibrado J y los calibrados K y L. Ello implica efectuar una transferencia de calibrado entre instrumentos en la que se va a considerar el sistema μ AUTOLAB como instrumento patrón ya que ha proporcionado los mejores resultados, y el Metrohm 646 VA como instrumento satélite. En primer lugar se ha llevado a cabo la transferencia de calibrado entre los calibrados J y K, y posteriormente el mismo procedimiento de transferencia se ha aplicado a los datos del calibrado L para evaluar la estabilidad de dicha metodología.

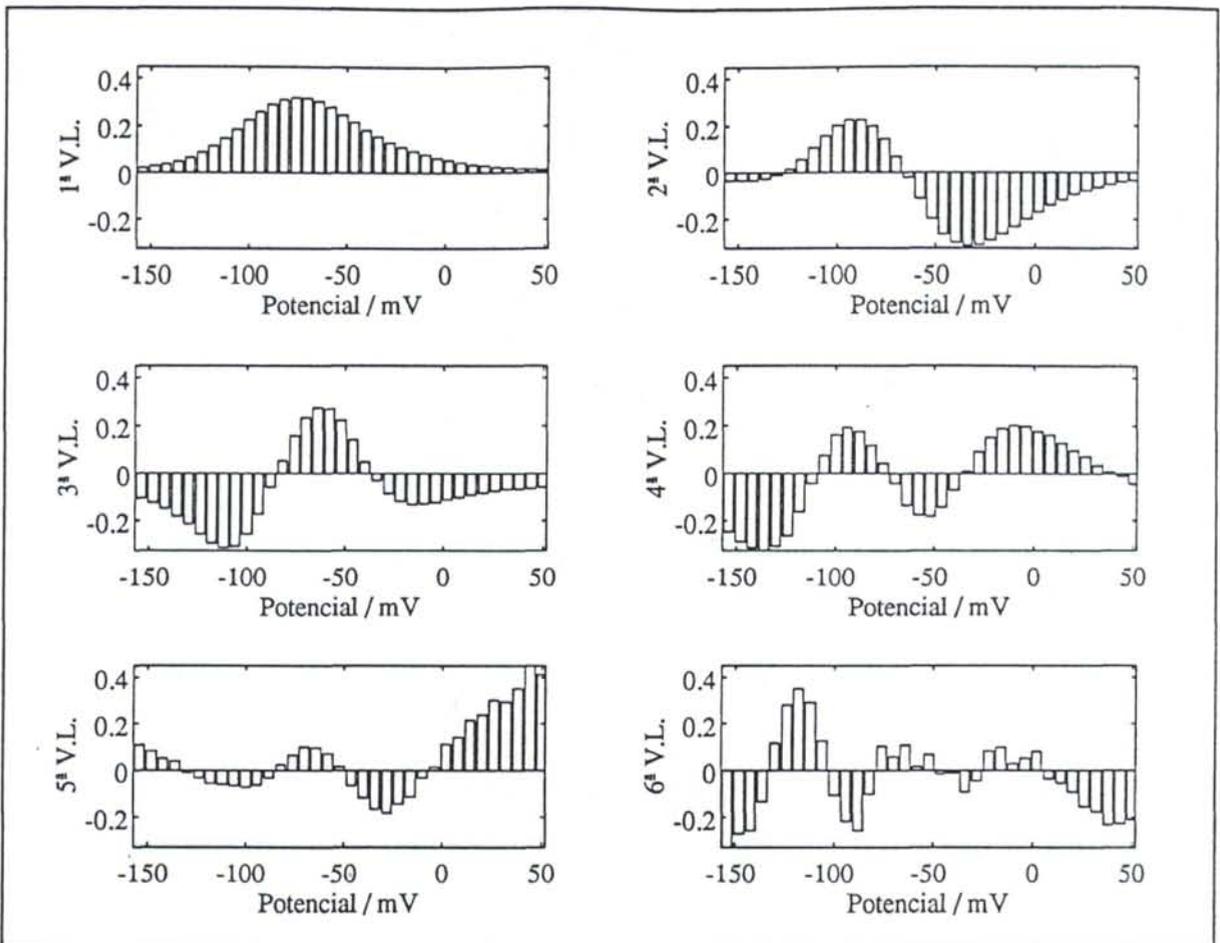


Figura III.7.8 Loadings correspondientes al modelo PLS construido para el cobre en el calibrado J.

Aplicación de la función "pdsgen"

Para efectuar la transferencia de calibrado entre ambos instrumentos, del calibrado J al calibrado K, se ha utilizado la función "pdsgen" definida anteriormente. Como en ejemplos anteriores, en la transferencia de calibrado se han estudiado diferentes tamaños de ventana y subconjuntos de transferencia con la finalidad de mejorar los resultados de este procedimiento. Los parámetros de entrada/salida seleccionados en este caso han sido los siguientes:

» $[F, \text{sub}] = \text{pdsgen}(J, K, w, no)$

correspondiéndose con

J: Matriz con los voltamperogramas del calibrado J, es decir, aquellos registrados con el sistema $\mu\text{AUTOLAB}$ o instrumento patrón. Tiene de dimensiones 40 filas (un polarograma por cada muestra de la Tabla III.7.1) por 35 columnas (potenciales).

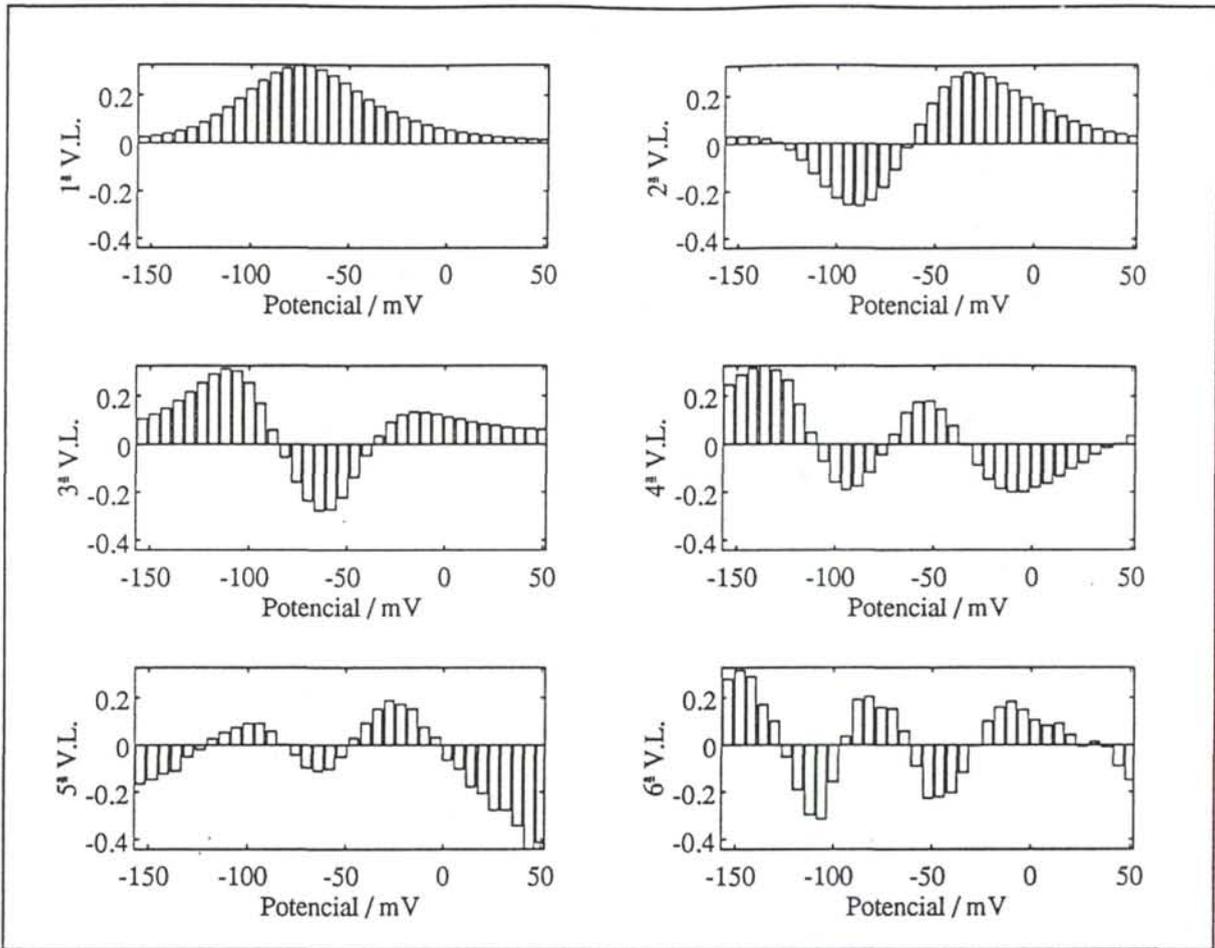


Figura III.7.9 Loadings correspondientes al modelo PLS construido para hierro en el calibrado J.

- K:** Matriz con los voltamperogramas del calibrado K, es decir, los obtenidos utilizando el Metrohm 646 VA o instrumento satélite. Tiene idénticas dimensiones que J, y todos sus elementos se encuentran ordenados de igual forma.
- w:** Escalar que indica cuántos valores de potencial constituyen la ventana móvil que se utiliza en el cálculo de la matriz de transferencia. En las transferencias de calibrado llevadas a cabo con diversos conjuntos de transferencia se han utilizado tamaños de ventana 3, 5, 7, ..., 35, lo cual equivale a un total de 17 ventanas diferentes.
- no:** Vector fila en el que se especifica qué muestras de las matrices de datos se van a utilizar en la transferencia de calibrado. Las muestras que constituyen los subconjuntos de transferencia utilizados en la transferencia de calibrado se hallan en la Figura III.7.10.

Tabla III.7.6 Errores relativos y concentraciones reales y calculadas por los modelos PLS para cobre y hierro en el calibrado J.

Muestra	Cobre			Hierro		
	Conc. real (μM)	Conc. calc. (μM)	Error relativo (%)	Conc. real (μM)	Conc. calc. (μM)	Error relativo (%)
1	0.0000	0.0035	-	0.0000	-0.1516	-
2	0.2074	0.2108	1.6591	0.0000	0.3290	-
3	0.4121	0.4133	0.3009	0.0000	0.0950	-
4	0.6141	0.6177	0.5921	0.0000	-0.0270	-
5	0.8135	0.8154	0.2358	0.0000	0.0130	-
6	1.0103	1.0230	1.2600	0.0000	0.0990	-
7	0.0000	0.0097	-	20.7417	20.6730	-0.3312
8	0.2061	0.2092	1.4930	20.6053	20.5139	-0.4436
9	0.4094	0.4125	0.7548	20.4706	20.3084	-0.7924
10	0.6101	0.6032	-1.1356	20.3377	20.0520	-1.4048
11	0.8083	0.8078	-0.0594	20.2065	20.0440	-0.8042
12	1.0038	0.9837	-2.0025	20.0769	20.0251	-0.2580
13	0.9975	0.9782	-1.9318	39.8981	40.1808	0.7086
14	0.0000	0.0070	-	41.2105	41.3167	0.2577
15	0.2047	0.2114	3.2775	40.9412	41.1173	0.4301
16	0.4068	0.4070	0.0548	40.6753	40.6426	-0.0804
17	0.6062	0.6011	-0.8416	40.4129	40.3423	-0.1747
18	0.8031	0.8037	0.0700	40.1538	40.2030	0.1225
19	0.0000	0.0081	-	61.4118	61.5433	0.2141
20	0.2034	0.2110	3.7419	61.0130	61.1706	0.2583
21	0.4041	0.4177	3.3687	60.6194	60.9070	0.4744
22	0.6023	0.6085	1.0216	60.2308	60.2132	-0.0292
23	0.7980	0.8076	1.1984	59.8471	59.6575	-0.3168
24	0.9911	1.0113	2.0361	59.4684	59.4119	-0.0950
25	0.0000	-0.0016	-	81.3506	81.2510	-0.1224
26	0.2021	0.1979	-2.0584	80.8258	80.2974	-0.6538
27	0.4015	0.3962	-1.3146	80.3077	80.4266	0.1481
28	0.5985	0.5865	-2.0027	79.7962	79.6419	-0.1934
29	0.7929	0.7734	-2.4592	79.2911	79.1879	-0.1302
30	0.9849	0.9606	-2.4635	78.7925	78.9521	0.2026
31	0.0000	-0.0164	-	101.0320	101.0600	0.0277
32	0.2008	0.1960	-2.3934	100.3840	100.3190	-0.0648
33	0.3990	0.4016	0.6519	99.7452	100.0810	0.3367
34	0.5947	0.6005	0.9802	99.1139	99.2250	0.1121
35	0.7879	0.8057	2.2623	98.4906	98.2044	-0.2906
36	0.9788	0.9939	1.5412	97.8750	97.9970	0.1246
<i>t1</i>	0.3081	0.3125	1.4249	20.5377	20.2849	-1.2309
<i>t2</i>	0.5003	0.5001	-0.0406	80.0511	79.9974	-0.0671
<i>t3</i>	0.8005	0.7868	-1.7062	50.0319	49.9233	-0.2171
<i>t4</i>	0.9818	0.9394	-4.3197	88.3636	88.7121	0.3944

Tabla III.7.7 Errores relativos y concentraciones reales y calculadas por los modelos PLS para cobre y hierro en el calibrado K.

Muestra	Cobre			Hierro		
	Conc. real (μM)	Conc. calc. (μM)	Error relativo (%)	Conc. real (μM)	Conc. calc. (μM)	Error relativo (%)
1	0.0000	-0.0053	-	0.0000	0.3944	-
2	0.2074	0.2074	-0.0140	0.0000	0.0800	-
3	0.4121	0.4113	-0.1910	0.0000	0.2740	-
4	0.6141	0.6144	0.0534	0.0000	0.2464	-
5	0.8135	0.7932	-2.4949	0.0000	0.4159	-
6	1.0103	1.0121	0.1752	0.0000	0.2054	-
7	0.0000	-0.0010	-	20.7417	21.7832	5.0213
8	0.2061	0.2143	3.9767	20.6053	19.4181	-5.7616
9	0.4094	0.4140	1.1339	20.4706	20.5666	0.4690
10	0.6101	0.6152	0.8397	20.3377	19.7430	-2.9241
11	0.8083	0.8308	2.7802	20.2065	19.5488	-3.2549
12	1.0038	1.0344	3.0454	20.0769	19.7459	-1.6487
13	0.9975	1.0000	0.2458	39.8981	39.6632	-0.5887
14	0.0000	0.0180	-	41.2105	41.2362	0.0624
15	0.2047	0.2077	1.4514	40.9412	41.0615	0.2938
16	0.4068	0.3809	-6.3719	40.6753	40.9008	0.5544
17	0.6062	0.5734	-5.4152	40.4129	40.9714	1.3820
18	0.8031	0.7569	-5.7477	40.1538	40.1715	0.0441
19	0.0000	0.0180	-	61.4118	60.9993	-0.6717
20	0.2034	0.2112	3.8161	61.0130	61.0640	0.0836
21	0.4041	0.4015	-0.6441	60.6194	61.5558	1.5447
22	0.6023	0.6073	0.8341	60.2308	60.2818	0.0847
23	0.7980	0.8048	0.8525	59.8471	60.2659	0.6998
24	0.9911	1.0021	1.1119	59.4684	59.4369	-0.0530
25	0.0000	0.0110	-	81.3506	80.4958	-1.0508
26	0.2021	0.2012	-0.4325	80.8258	80.3258	-0.6186
27	0.4015	0.3985	-0.7572	80.3077	79.1808	-1.4032
28	0.5985	0.5856	-2.1485	79.7962	80.0035	0.2598
29	0.7929	0.7768	-2.0351	79.2911	79.4421	0.1904
30	0.9849	0.9662	-1.8962	78.7925	78.6448	-0.1875
31	0.0000	-0.0140	-	101.0320	101.7740	0.7344
32	0.2008	0.1952	-2.8123	100.3840	100.4310	0.0468
33	0.3990	0.4004	0.3501	99.7452	99.7318	-0.0134
34	0.5947	0.5995	0.8134	99.1139	99.5144	0.4041
35	0.7879	0.8046	2.1140	98.4906	98.6962	0.2088
36	0.9788	0.9992	2.0851	97.8750	97.6653	-0.2143
<i>i1</i>	0.3081	0.3009	-2.3479	20.5377	20.7670	1.1165
<i>i2</i>	0.5003	0.4915	-1.7521	80.0511	79.7144	-0.4206
<i>i3</i>	0.8005	0.7258	-9.3298	50.0319	49.9566	-0.1505
<i>i4</i>	0.9818	0.9878	0.6139	88.3636	87.6379	-0.8213

Tabla III.7.8 Errores relativos y concentraciones reales y calculadas por los modelos PLS para cobre y hierro en el calibrado L.

Muestra	Cobre			Hierro		
	Conc. real (μM)	Conc. calc (μM)	Error relativo (%)	Conc. real (μM)	Conc. calc. (μM)	Error relativo (%)
1	0.0000	0.0160	-	0.0000	-0.3551	-
2	0.2074	0.2139	3.1514	0.0000	-0.3519	-
3	0.4121	0.3995	-3.0631	0.0000	-0.4542	-
4	0.6141	0.6233	1.4905	0.0000	0.1225	-
5	0.8135	0.8195	0.7384	0.0000	-0.2930	-
6	1.0103	1.0083	-0.1990	0.0000	0.4040	-
7	0.0000	0.0150	-	20.7417	21.1580	2.0071
8	0.2061	0.1980	-3.9369	20.6053	21.4503	4.1009
9	0.4094	0.4004	-2.1915	20.4706	20.9239	2.2144
10	0.6101	0.5960	-2.3157	20.3377	20.7448	2.0017
11	0.8083	0.8032	-0.6272	20.2065	20.3653	0.7859
12	1.0038	0.9929	-1.0812	20.0769	20.2312	0.7685
13	0.9975	0.9817	-1.5857	39.8981	39.8718	-0.0659
14	0.0000	-0.0037	-	41.2105	40.8040	-0.9864
15	0.2047	0.2113	3.2389	40.9412	40.5467	-0.9636
16	0.4068	0.4108	0.9904	40.6753	40.5637	-0.2744
17	0.6062	0.6220	2.6127	40.4129	39.8470	-1.4003
18	0.8031	0.8112	1.0122	40.1538	39.7517	-1.0014
19	0.0000	-0.0100	-	61.4118	61.4826	0.1153
20	0.2034	0.1885	-7.3358	61.0130	61.1619	0.2440
21	0.4041	0.4265	5.5377	60.6194	59.8146	-1.3276
22	0.6023	0.5962	-1.0191	60.2308	60.6282	0.6598
23	0.7980	0.7922	-0.7212	59.8471	60.1211	0.4578
24	0.9911	1.0357	4.5031	59.4684	58.3158	-1.9382
25	0.0000	0.0114	-	81.3506	80.7948	-0.6832
26	0.2021	0.2086	3.1994	80.8258	81.1216	0.3660
27	0.4015	0.3829	-4.6266	80.3077	81.0251	0.8933
28	0.5985	0.5936	-0.8174	79.7962	80.3882	0.7419
29	0.7929	0.8005	0.9552	79.2911	79.7439	0.5711
30	0.9849	0.9760	-0.8992	78.7925	78.5607	-0.2942
31	0.0000	0.0170	-	101.0320	100.4510	-0.5751
32	0.2008	0.1974	-1.6783	100.3840	100.2710	-0.1126
33	0.3990	0.3893	-2.4273	99.7452	100.0890	0.3447
34	0.5947	0.5865	-1.3709	99.1139	99.3056	0.1934
35	0.7879	0.7798	-1.0293	98.4906	98.5718	0.0824
36	0.9788	0.9793	0.0506	97.8750	97.7878	-0.0891
<i>t</i> 1	0.3081	0.2752	-10.6725	20.5377	21.7863	6.0796
<i>t</i> 2	0.5003	0.4755	-4.9526	80.0511	80.9097	1.0726
<i>t</i> 3	0.8005	0.8303	3.7279	50.0319	49.7578	-0.5479
<i>t</i> 4	0.9818	1.0484	6.7865	88.3636	84.2126	-4.6976

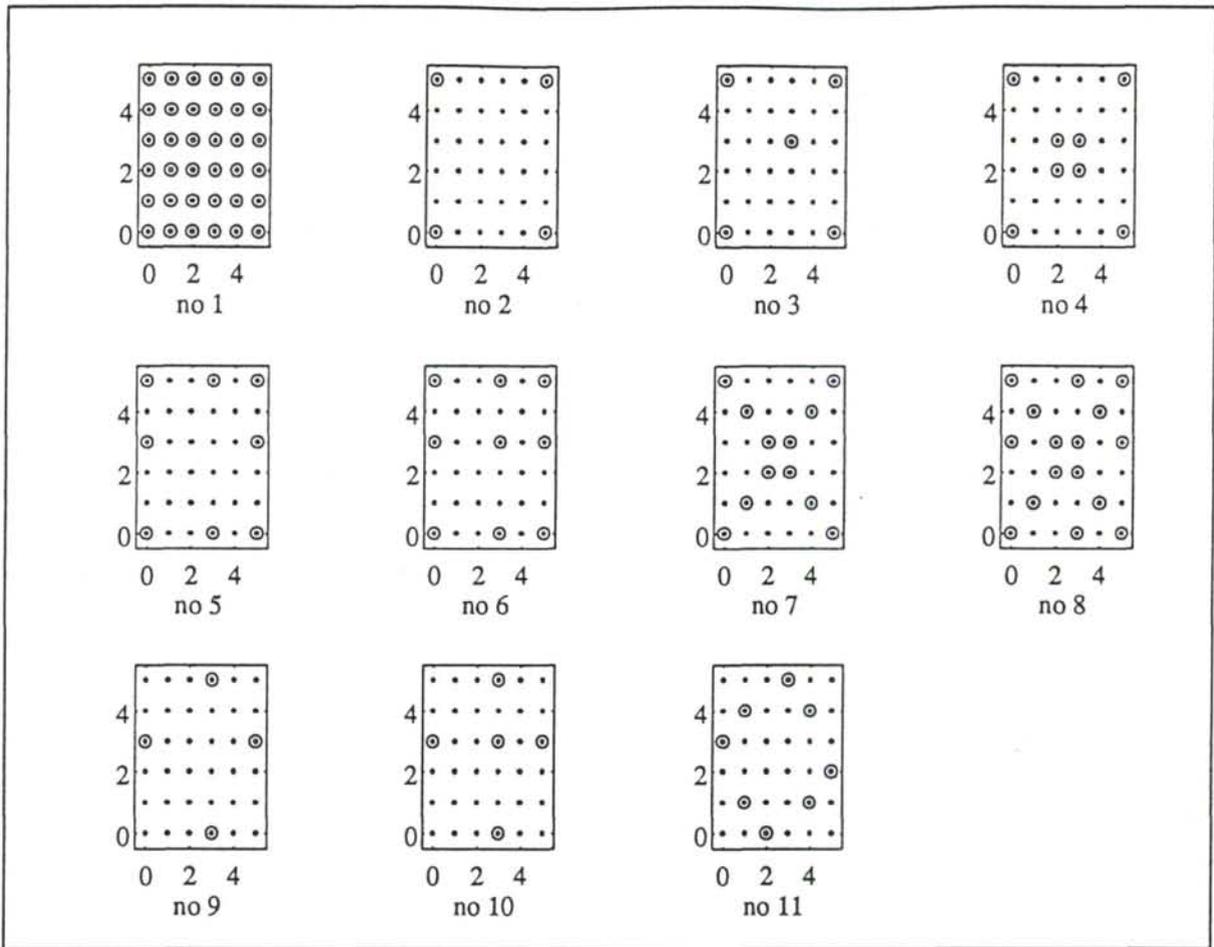


Figura III.7.10 Subconjuntos de transferencia (formados por las muestras \odot). En ordenadas se representa el nivel de cobre y en abscisas el de hierro.

Como resultado del procedimiento de transferencia se tiene una matriz de transferencia de dimensiones 35×35 diferente al variar el tamaño de ventana o el subconjunto de transferencia utilizado. Se han realizado un total de 187 (17×11) procedimientos de transferencia de calibrado distintos, estudiándose posteriormente la validez de los resultados obtenidos a través de los valores de SEP y SET correspondientes, cuya variabilidad en función de los parámetros no y w utilizados en la transferencia se muestra en las Figuras III.7.11 a III.7.14.

En la Tabla III.7.9 se muestran los valores de SEP y SET correspondientes a los propios modelos PLS construidos con cada uno de los calibrados completos. En este caso, los valores del SEP del calibrado K son comparables a los obtenidos para el mismo tras el procedimiento de transferencia de calibrado, ya que este parámetro evalúa muestras completamente ajenas al modelo PLS. Sin embargo, es necesario tener en cuenta que los

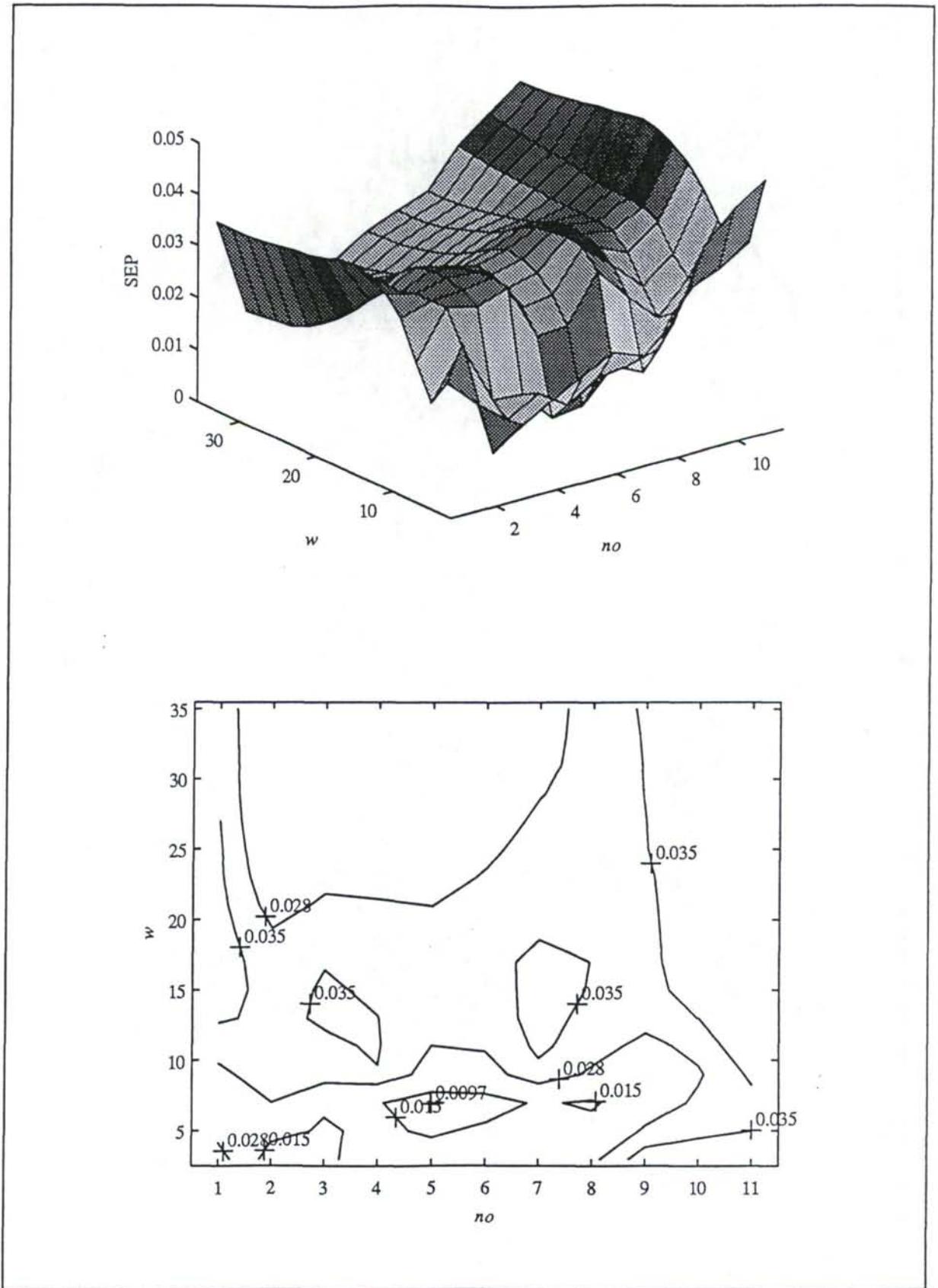


Figura III.7.11 Superficie de respuesta y curvas de nivel del SEP para el cobre.

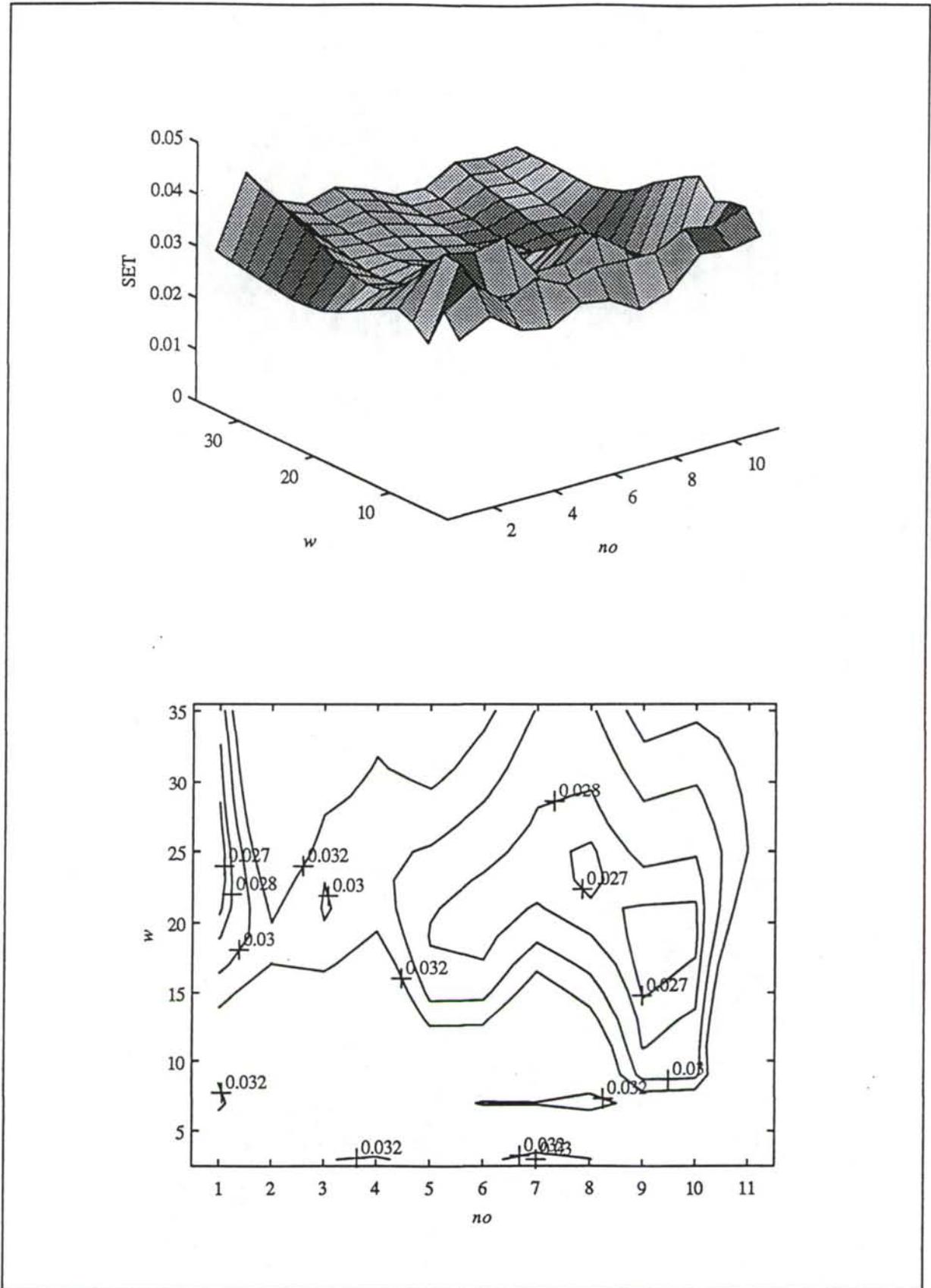


Figura III.7.12 Superficie de respuesta y curvas de nivel del SET para el cobre.

valores del SET se han obtenido considerando las muestras utilizadas en la construcción de los propios modelos PLS.

Tabla III.7.9 Valores de SEP y SET correspondientes a los modelos PLS construidos con los calibrado J y K completos.

Calibrado	Analito	SEP	SET
J	Cobre	0.0224	0.0107
	Hierro	0.2236	0.1815
K	Cobre	0.0379	0.0155
	Hierro	0.4178	0.4919

Una vez que se dispone de todos los valores de SEP y SET, el siguiente paso consiste en decidir qué condiciones son las mejores para llevar a cabo la transferencia de calibrado, es decir, cuáles son los valores de no y w que conducen a los mejores resultados. Para ello, hay que tener en cuenta que el número de experiencias implicadas en el proceso de transferencia sea pequeño, y por tanto suponga una mejora real frente a una recalibración completa, así como que los resultados obtenidos son correctos.

Análisis de los resultados

Dado el elevado número de procedimientos de transferencia llevados a cabo entre los dos instrumentos, el análisis de los resultados de dichos procedimientos se ha realizado a través de los test de agrupamientos de Tukey y Newman-Keuls [58] y de los diagramas Box plot correspondientes. En este análisis se han tenido en consideración tanto los valores del SEP como con los del SET al haberse observado algunas diferencias entre unos y otros con la variación de los parámetros no y w . En las Figuras III.7.15 y III.7.16 se representan los diagramas Box plot realizados con los valores de SEP y SET correspondientes a las transferencias de calibrado llevadas a cabo para la determinación de cobre y hierro, mientras que en las Tablas III.7.10 a III.7.13 se encuentran los resultados de los test de agrupamientos correspondientes.

Al analizar la variabilidad de SEP y SET cuando se considera la transferencia de calibrado llevada a cabo en la determinación de cobre, se observa que la influencia de los diferentes conjuntos de transferencia y de la anchura de ventana es más acusada en los valores

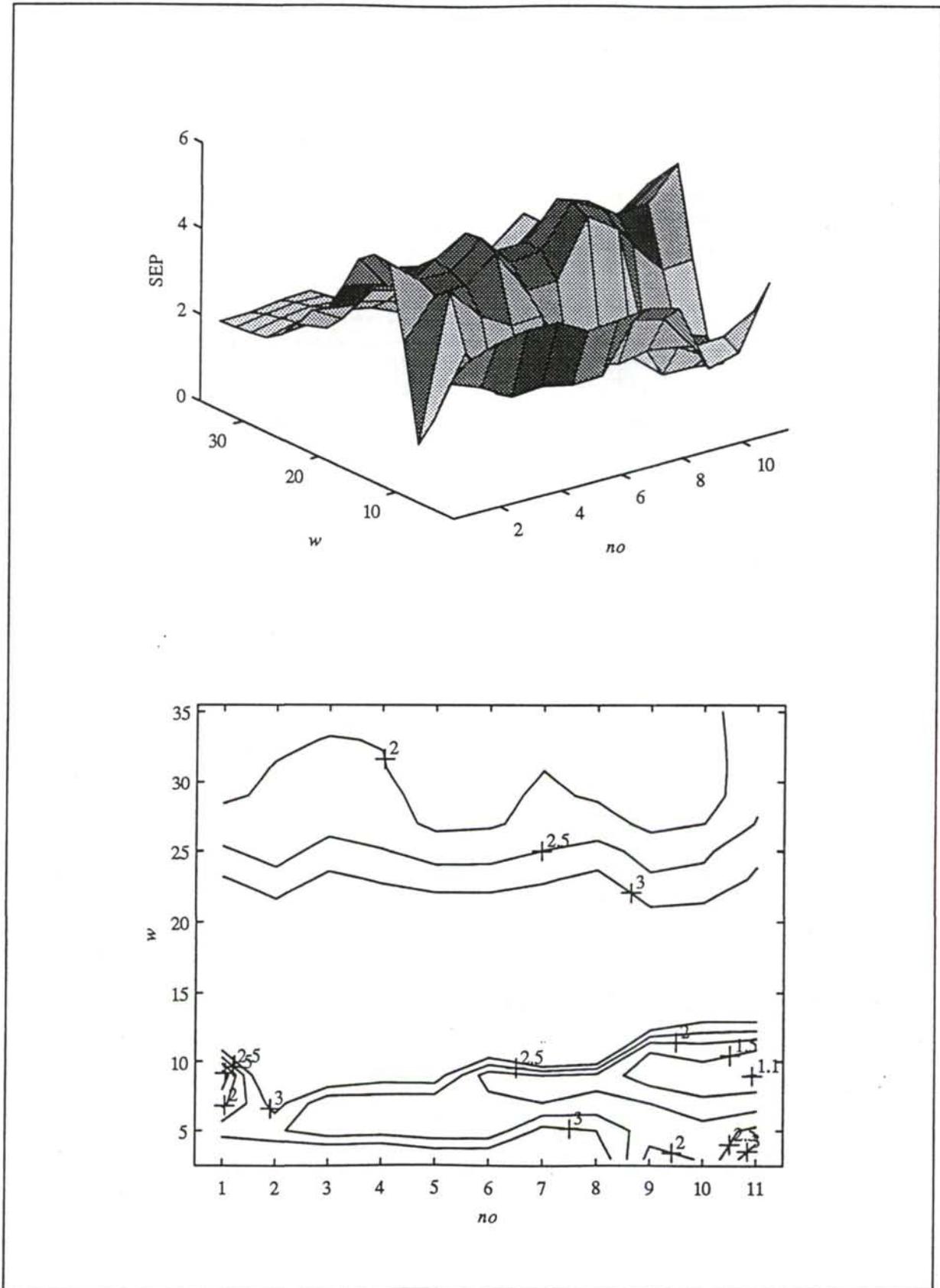


Figura III.7.13 Superficie de respuesta y curvas de nivel del SEP para el hierro.

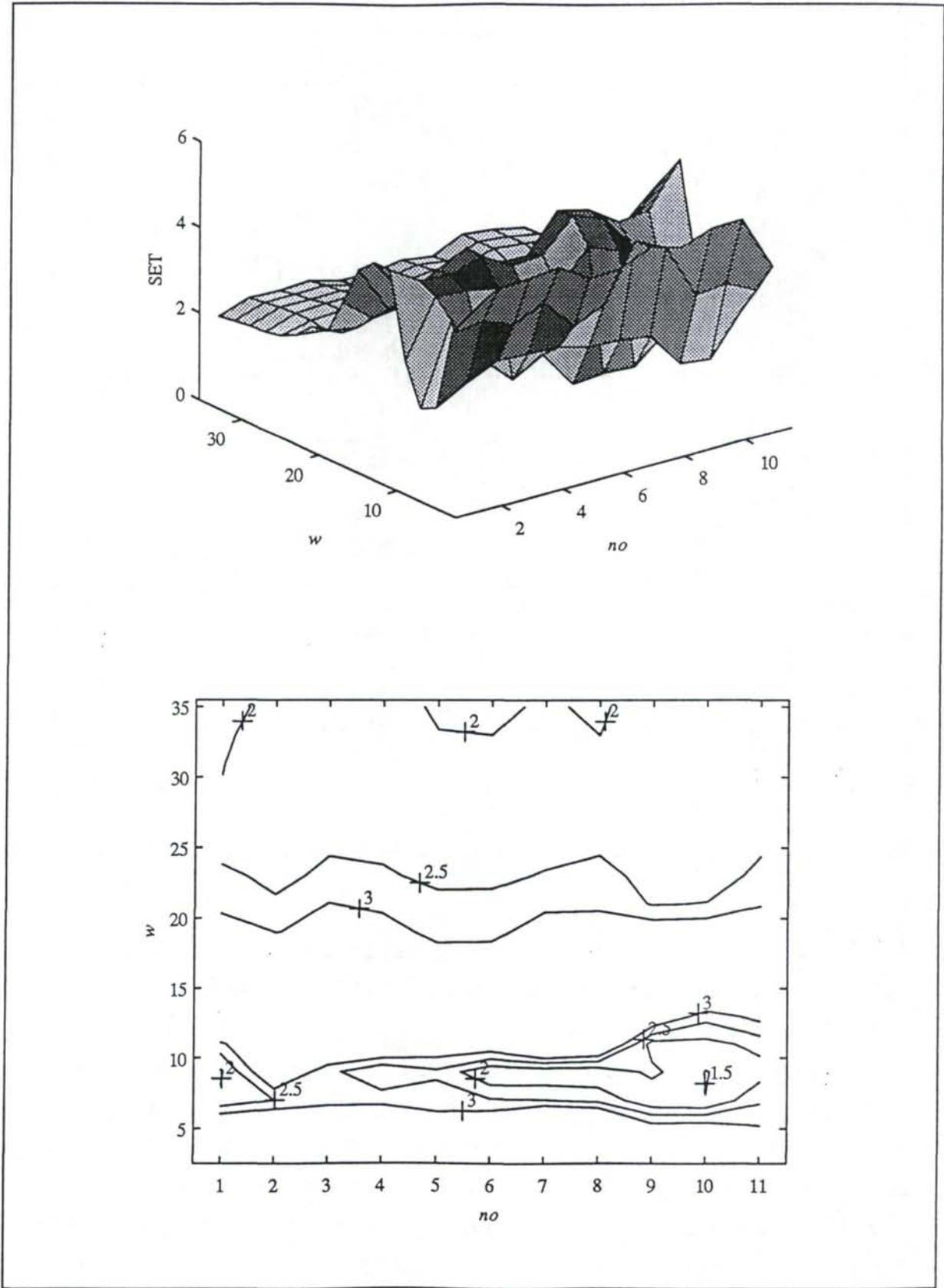


Figura III.7.14 Superficie de respuesta y curvas de nivel del SET para el hierro.

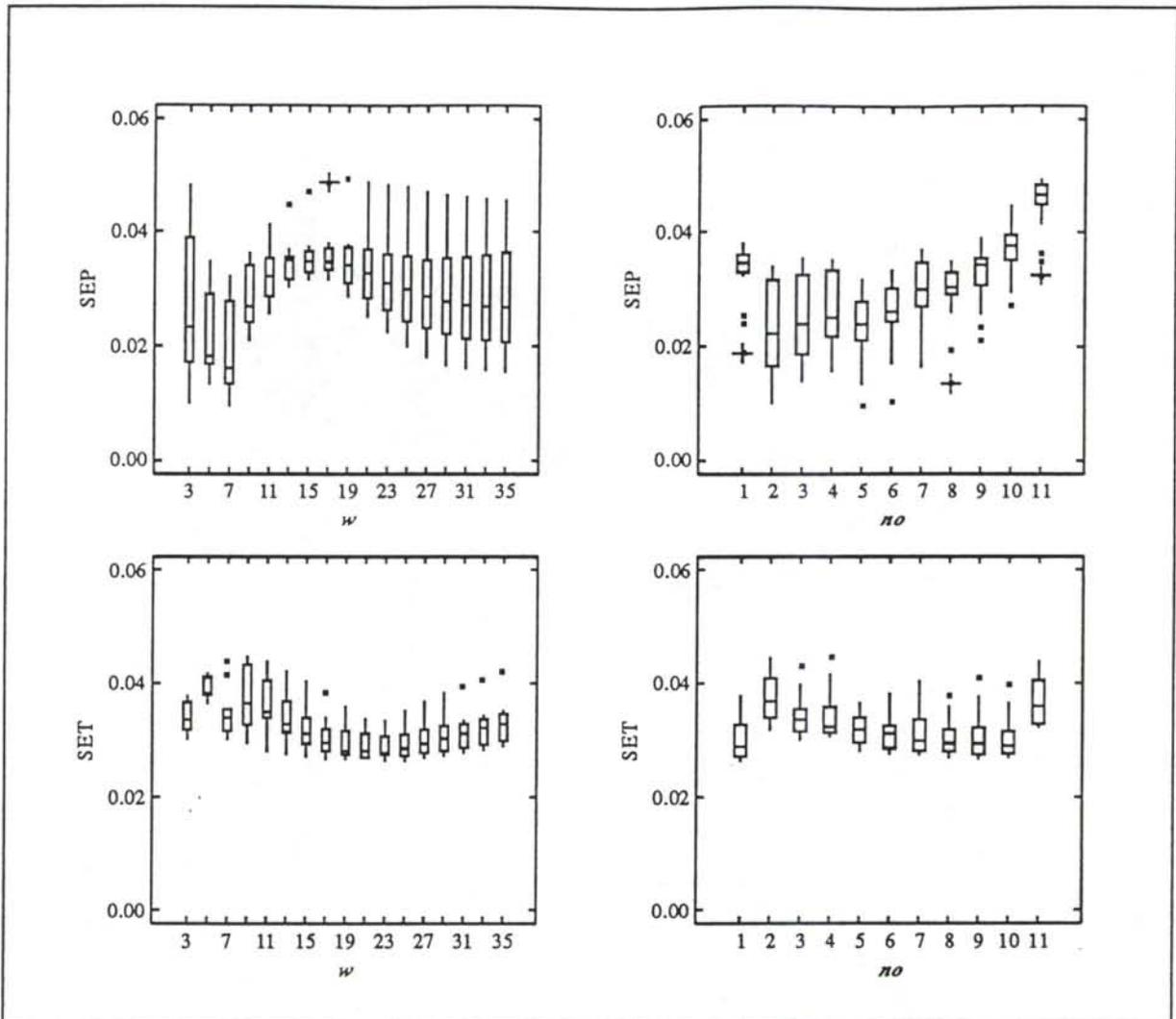


Figura III.7.15 Diagramas Box plot para los valores de SEP y SET de la transferencia de calibrado del cobre.

del SEP que en los equivalentes del SET. En los diagramas Box plot correspondientes, Figura III.7.15, el tamaño de ventana óptimo si se tiene en cuenta el SEP es 7, mientras que la influencia de dicho parámetro en los valores de SET es mucho menor. Los test de agrupamientos de Tukey y de Newman-Keuls, Tabla III.7.11, indican a su vez que los mejores resultados en cuanto al SET están asociados con tamaños de ventana mayores, señalando el tamaño de ventana 21 como óptimo. Con el fin de unificar criterios y analizando las curvas de nivel de las Figuras III.7.11 y III.7.12, se ha tomado como anchura de ventana óptima el valor 7, ya que en dichas figuras se observa que es posible encontrar algunos conjuntos de transferencia para los cuales ese valor de w conduce simultáneamente a valores de SEP y SET aceptables.

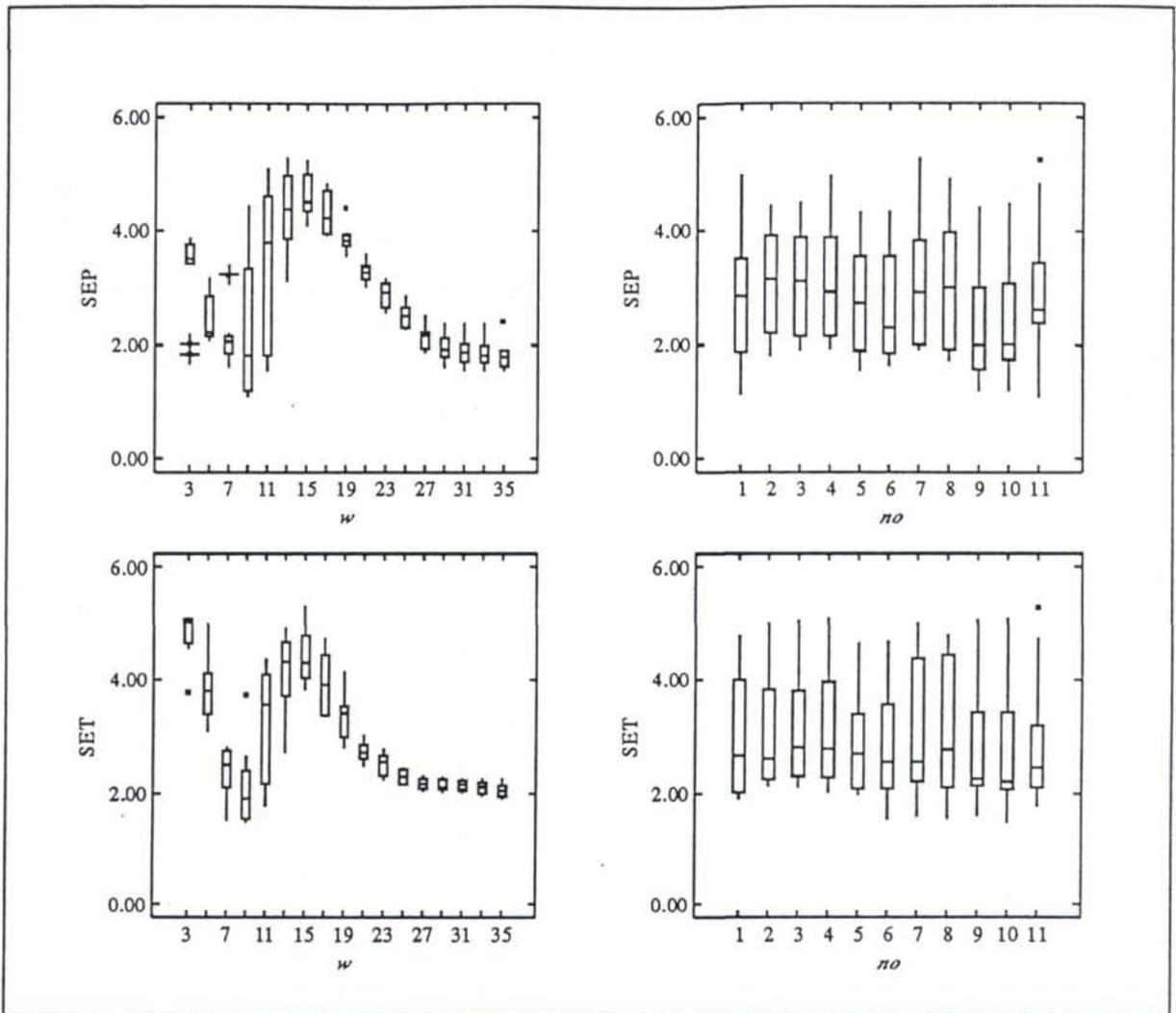


Figura III.7.16 Diagramas Box plot para los valores de SEP y SET de la transferencia de calibrado del hierro.

En cuanto a los diferentes conjuntos de transferencia utilizados, en las curvas de nivel de las Figuras III.7.11 y III.7.12 se observa que para el valor de $w = 7$ los conjuntos 5, 6, 7 y 8 presentan valores aceptables tanto de SEP como de SET. En los test de Tukey y de Newman-Keuls efectuados, Tablas III.7.10 y III.7.11, no se encuentran diferencias significativas entre los conjuntos 5 y 6, los cuales además, se agrupan con los conjuntos que presentan mejores valores medios, tanto en lo que se refiere al SEP como al SET. Como ambos conjuntos conducen a resultados prácticamente iguales, se ha seleccionado el conjunto de transferencia 5 para realizar posteriores transferencias de calibrado, ya que supone un esfuerzo experimental ligeramente menor.

Tabla III.7.10 Tests de agrupamientos de Tukey y Newman-Keuls para los valores de SEP correspondientes al cobre.

Conjunto de transferencia (<i>no</i>)				Ventana (<i>w</i>)			
Nivel	Media	Grupos homogéneos		Nivel	Media	Grupos homogéneos	
		Tukey	Newman-Keuls			Tukey	Newman-Keuls
2	0.0236	X	X	7	0.0194	X	X
5	0.0239	X	X	5	0.0215	XX	X
3	0.0251	XX	X	3	0.0270	XX	X
6	0.0262	XX	XX	35	0.0285	XX	XX
4	0.0262	XX	XX	33	0.0286	XX	XX
8	0.0292	XX	XX	31	0.0287	XX	XX
7	0.0294	XX	XX	9	0.0290	XX	XXX
9	0.0325	XX	X	29	0.0292	XX	XXXX
1	0.0331	XX	X	27	0.0299	XXX	XXXX
10	0.0367	X	X	25	0.0308	XXX	XXXXX
11	0.0448	X	X	23	0.0317	XXX	XXXXX
				11	0.0325	XXX	XXXXX
				21	0.0331	XXX	XXXX
				19	0.0347	XX	XXX
				13	0.0348	XX	XX
				15	0.0358	X	X
				17	0.0358	X	X

Los resultados de aplicar el procedimiento de transferencia para la determinación de cobre, utilizando los parámetros *no* y *w*, indicados se encuentran en la Tabla III.7.14, la cual contiene los valores de concentración calculados y los errores relativos asociados con ellos. Estos errores se encuentran bastante próximos a los obtenidos para el cobre con su propio modelo PLS (calibrado K, Tabla III.7.7), salvo en alguna muestra muy concreta. No obstante, la principal ventaja del procedimiento de transferencia utilizado reside en que únicamente ha sido necesario disponer de las 8 muestras patrón que constituyen el conjunto de transferencia 5 para poder determinar todas las muestras del calibrado K, frente a las 36 muestras de calibrado utilizadas para obtener el modelo PLS propio de dicho calibrado.

Tabla III.7.11 Tests de agrupamientos de Tukey y Newman-Keuls para los valores de SET correspondientes al cobre.

Conjunto de transferencia (<i>no</i>)				Ventana (<i>w</i>)			
Nivel	Media	Grupos homogéneos		Nivel	Media	Grupos homogéneos	
		Tukey	Newman-Keuls			Tukey	Newman-Keuls
1	0.0298	X	X	21	0.0290	X	X
10	0.0302	X	X	23	0.0290	X	X
8	0.0303	X	X	25	0.0294	X	XX
9	0.0304	X	X	19	0.0295	X	XX
6	0.0311	X	X	27	0.0302	XX	XXX
7	0.0316	XX	X	17	0.0306	XXX	XXX
5	0.0319	XXX	X	29	0.0309	XXX	XXX
3	0.0339	XX	X	31	0.0316	XXXX	XXXX
4	0.0344	XX	X	15	0.0321	XXXX	XXXXX
11	0.0367	XX	X	33	0.0323	XXXX	XXXX
2	0.0377	X	X	35	0.0331	XXXX	XXX
				13	0.0340	XXXX	XXX
				3	0.0341	XXXX	XXX
				7	0.0347	XXX	XX
				11	0.0362	XXX	XX
				9	0.0376	XX	XX
				5	0.0390	X	X

En el análisis análogo realizado para el hierro se observa que existen menos diferencias en la influencia de *no* y *w* sobre SEP y SET de las existentes en el caso del cobre, por el contrario, se tienen tendencias muy similares para ambos errores, como muestran las Figuras III.7.13 y III.7.14. Sin embargo, la influencia del tamaño de ventana es mayor que la observada para el cobre y, como evidencian los diagramas Box plot de la Figura III.7.16, SEP y SET siguen idéntica pauta. Las curvas de nivel de las Figuras III.7.13 y III.7.14 muestran que para $w = 9$ los SEP y SET correspondientes a varios conjuntos de transferencia son bajos. Aunque este valor de *w* no es significativamente diferente de otros mayores, según indican los test de Tukey y Newman-Keuls llevados a cabo, Tablas III.7.12 y III.7.13.

Tabla III.7.12 Tests de agrupamientos de Tukey y Newman-Keuls para los valores de SEP correspondientes al hierro.

Conjunto de transferencia (n_0)				Ventana (w)			
Nivel	Media	Grupos homogéneos		Nivel	Media	Grupos homogéneos	
		Tukey	Newman-Keuls			Tukey	Newman-Keuls
9	2.4165	X	X	35	1.8219	X	X
10	2.4543	X	XX	33	1.8650	X	XX
6	2.7465	XX	XXX	31	1.9011	X	XX
5	2.8093	XX	XXX	29	1.9799	X	XX
11	2.9241	XX	XX	7	2.1105	X	XX
1	2.9279	XX	XX	27	2.1404	XX	XX
8	3.0746	X	X	9	2.2590	XX	XX
2	3.0952	X	X	5	2.4804	XX	XX
3	3.1216	X	X	25	2.4966	XX	XX
4	3.1443	X	X	23	2.8764	XX	XX
7	3.1544	X	X	21	3.2968	XX	XX
				3	3.3146	XX	XX
				11	3.5203	XX	XX
				19	3.8518	XX	XX
				17	4.3187	XX	XX
				13	4.3341	XX	XX
				15	4.6842	X	X

En este caso, todos los elementos de análisis señalan que el parámetro más influyente en los valores de SEP y SET obtenidos en las diferentes transferencias de calibrado es la anchura de ventana, ya que la influencia de los diferentes conjuntos de transferencia es mucho menor. No obstante, los conjuntos de transferencia que se asocian con mínimos de SEP y SET, tomando un valor para w de 9, son los conjuntos 10 y 11, proporcionando este último los mejores resultados considerando los errores relativos individuales correspondientes a cada muestra que figuran en la Tabla III.7.14.

Tabla III.7.13 Tests de agrupamientos de Tukey y Newman-Keuls para los valores de SET correspondientes al hierro.

Conjunto de transferencia (<i>no</i>)				Ventana (<i>w</i>)			
Nivel	Media	Grupos homogéneos		Nivel	Media	Grupos homogéneos	
		Tukey	Newman-Keuls			Tukey	Newman-Keuls
10	2.6661	X	X	35	2.0539	X	X
9	2.7539	XX	XX	9	2.0698	X	X
6	2.8091	XX	XX	33	2.1030	X	X
5	2.8475	XX	XX	29	2.1416	XX	X
11	2.8842	XX	XX	31	2.1435	XX	X
1	2.9750	XX	XX	27	2.1571	XX	X
2	3.0672	XX	XX	25	2.2805	XX	XX
3	3.1092	XX	X	7	2.3564	XX	XX
8	3.1242	XX	X	23	2.4814	XX	XX
4	3.1346	X	X	21	2.7360	XX	X
7	3.1906	X	X	11	3.3057	XX	X
				19	3.3165	XX	X
				5	3.8681	XX	X
				17	3.9192	XX	X
				13	4.1426	X	XX
				15	4.4422	XX	X
				3	4.8051	X	X

Aplicación del procedimiento PDS a un tercer calibrado

Con el fin de evaluar la posibilidad de utilizar el procedimiento de transferencia puesto a punto para la determinación simultánea de cobre en presencia de hierro en calibrados posteriores, se ha llevado a cabo la transferencia de calibrado tomando como situación patrón el calibrado J y como situación satélite el calibrado L. En este nuevo procedimiento de transferencia de calibrado se han considerado únicamente las muestras correspondientes a los conjuntos de transferencia 5 y 11 y las del conjunto test, utilizándose respectivamente los valores $no = 5$ con $w = 7$ para determinar la concentración de cobre en las muestras indicadas, y $no = 11$ con $w = 9$ para determinar la concentración de hierro en esas mismas

Tabla III.7.14 Concentraciones reales y calculadas y errores relativos obtenidos como resultado de aplicar el procedimiento de transferencia de calibrado de la situación J a la situación K. Se han tomado los valores $no = 5$ y $w = 7$ para la transferencia del cobre y $no = 11$ y $w = 9$ en la de hierro.

Muestra	Cobre			Hierro		
	Conc. real (μM)	Conc. calc. (μM)	Error relativo (%)	Conc. real (μM)	Conc. calc. (μM)	Error relativo (%)
1	0.0000	0.0124	-	0.0000	-0.8670	-
2	0.2074	0.2013	-2.9547	0.0000	0.2480	-
3	0.4121	0.4025	-2.3392	0.0000	1.0218	-
4	0.6141	0.5969	-2.8002	0.0000	1.5848	-
5	0.8135	0.7788	-4.2600	0.0000	2.1594	-
6	1.0103	1.0094	-0.0934	0.0000	2.7481	-
7	0.0000	0.0028	-	20.7417	21.4008	3.1774
8	0.2061	0.2010	-2.4694	20.6053	19.9343	-3.2564
9	0.4094	0.4211	2.8562	20.4706	20.7909	1.5645
10	0.6101	0.6310	3.4311	20.3377	19.9167	-2.0701
11	0.8083	0.8779	8.6136	20.2065	19.5097	-3.4485
12	1.0038	1.1191	11.4833	20.0769	19.0460	-5.1350
13	0.9975	1.0974	10.0186	39.8981	38.4917	-3.5249
14	0.0000	-0.0016	-	41.2105	41.9887	1.8884
15	0.2047	0.1992	-2.6900	40.9412	41.9203	2.3916
16	0.4068	0.3908	-3.9419	40.6753	41.3792	1.7306
17	0.6062	0.6068	0.0924	40.4129	41.3521	2.3240
18	0.8031	0.8009	-0.2768	40.1538	40.6367	1.2027
19	0.0000	0.0001	-	61.4118	60.7164	-1.1323
20	0.2034	0.2049	0.7487	61.0130	60.9906	-0.0367
21	0.4041	0.4164	3.0493	60.6194	61.3748	1.2461
22	0.6023	0.6342	5.2946	60.2308	60.4245	0.3216
23	0.7980	0.8487	6.3513	59.8471	60.7800	1.5588
24	0.9911	1.0598	6.9277	59.4684	60.3861	1.5432
25	0.0000	0.0004	-	81.3506	78.6844	-3.2774
26	0.2021	0.1942	-3.9278	80.8258	78.7218	-2.6031
27	0.4015	0.3976	-0.9613	80.3077	77.9997	-2.8739
28	0.5985	0.5906	-1.3243	79.7962	79.7430	-0.0667
29	0.7929	0.7766	-2.0519	79.2911	80.8696	1.9907
30	0.9849	0.9575	-2.7856	78.7925	81.6820	3.6673
31	0.0000	-0.0045	-	101.0323	96.3512	-4.6333
32	0.2008	0.2084	3.7763	100.3846	95.8279	-4.5392
33	0.3990	0.4118	3.2196	99.7452	96.7745	-2.9783
34	0.5947	0.6113	2.7882	99.1139	98.0247	-1.0989
35	0.7879	0.8084	2.6034	98.4906	98.6948	0.2074
36	0.9788	0.9744	-0.4512	97.8750	100.2082	2.3838
t_1	0.3081	0.3011	-2.2654	20.5377	21.1727	3.0919
t_2	0.5003	0.5033	0.6093	80.0511	78.4561	-1.9925
t_3	0.8005	0.7866	-1.7348	50.0319	48.7184	-2.6253
t_4	0.9818	0.9924	1.0814	88.3636	88.2227	-0.1595

muestras. Los resultados obtenidos al realizar estas transferencias de calibrado se muestran en la Tabla III.7.15.

Tabla III.7.15 Concentraciones reales y calculadas y errores relativos obtenidos al aplicar el procedimiento de transferencia de calibrado de la situación J a la L, tomando los valores $n_0 = 5$ y $w = 7$ para la transferencia del cobre y $n_0 = 11$ y $w = 9$ en la de hierro.

Muestra	Cobre			Hierro		
	Conc. real (μM)	Conc. calc. (μM)	Error relativo (%)	Conc. real (μM)	Conc. calc. (μM)	Error relativo (%)
1	0.0000	0.0170	-	0.0000	-0.5256	-
3	0.4121	0.3923	-4.8060	0.0000	-2.5928	-
4	0.6141	0.5946	-3.1708	0.0000	-3.0582	-
6	1.0103	0.9886	-2.1514	0.0000	-4.2027	-
8	0.2061	0.1681	-18.4268	20.6053	23.3573	13.3558
11	0.8083	0.7710	-4.6188	20.2065	20.2466	0.1982
13	0.9975	0.9939	-0.3630	39.8981	38.8672	-2.5838
19	0.0000	0.0004	-	61.4118	64.2724	4.6580
24	0.9911	1.0800	8.9706	59.4684	57.1847	-3.8402
26	0.2021	0.2228	10.2630	80.8258	81.1290	0.3751
29	0.7929	0.8435	6.3801	79.2911	78.3087	-1.2390
31	0.0000	-0.0035	-	101.0323	98.7930	-2.2164
34	0.5947	0.6063	1.9570	99.1139	97.3586	-1.7709
36	0.9788	0.9775	-0.1292	97.8750	96.0946	-1.8190
<i>t1</i>	0.3081	0.2613	-15.1852	20.5377	23.2259	13.0892
<i>t2</i>	0.5003	0.5382	7.5784	80.0511	79.1454	-1.1314
<i>t3</i>	0.8005	0.8702	8.7025	50.0319	48.2283	-3.6050
<i>t4</i>	0.9818	1.0424	6.1743	88.3636	83.6652	-5.3171

Del análisis de la misma se deduce que los errores relativos tras la transferencia están un poco por encima de los obtenidos por los modelos PLS construidos con el calibrado L completo, Tabla III.7.8, situación que equivaldría a una recalibración completa. Sin embargo, los errores correspondientes a las muestras 8 y *t1*, que son muy altos, parecen estar más bien relacionados con la naturaleza de las propias muestras que con el procedimiento de transferencia de calibrado en si mismo. De hecho, esas dos muestras proporcionaron errores

relativos del orden del -11 y -4% en la determinación de cobre y del 6 y 4% en la de hierro respectivamente, cuando éstas se llevaron a cabo con los modelos PLS construídos con todas las muestras del calibrado L, Tabla III.7.8.

Por otro lado, esas mismas muestras no presentan errores que discrepen del resto en la transferencia llevada a cabo entre el calibrado J y el calibrado K con idénticas condiciones de transferencia (Tabla III.7.14), por lo que esos errores no parecen deberse a una proporción cobre/hierro determinada, sino a la variabilidad intrínseca de cada muestra de calibrado.

III.7.4.3 Transferencia de calibrado mediante el procedimiento GCT

Aplicación de la función "gct"

El procedimiento de transferencia GCT ha sido utilizado para llevar a cabo la transferencia de calibrado entre el calibrado J y el calibrado K, es decir, una transferencia de calibrado entre instrumentos como la realizada anteriormente. En la transferencia de calibrado se han utilizado las matrices de datos **J** y **K**, y los subconjuntos de transferencia indicados en la Figura III.7.10.

Los resultados correspondientes a este procedimiento de transferencia se muestran en las Figuras III.7.17 y III.7.18, en las cuales están representados los valores de SEP y SET para cobre y hierro respectivamente. Cabe señalar que la pauta seguida por ambos parámetros, SEP y SET, en función del conjunto de transferencia utilizado es muy similar para cobre y hierro, a diferencia de lo que ocurría cuando el procedimiento de transferencia utilizado fue PDS. Ello indica que en este caso, el procedimiento de transferencia se realiza de una forma más global, que no tiene una relación directa con el hecho de que la señal del analito a determinar se encuentre en una zona concreta del voltamperograma.

Por otro lado, algunos de los valores de SEP y SET obtenidos mediante el procedimiento GCT alcanzan valores menores que los mínimos obtenidos en la transferencia de calibrado llevada a cabo a través del procedimiento PDS entre los mismos conjuntos de datos, Figuras III.7.11 a III.7.14. Esto parece indicar que en este caso concreto el procedimiento GCT permite alcanzar resultados más exactos.

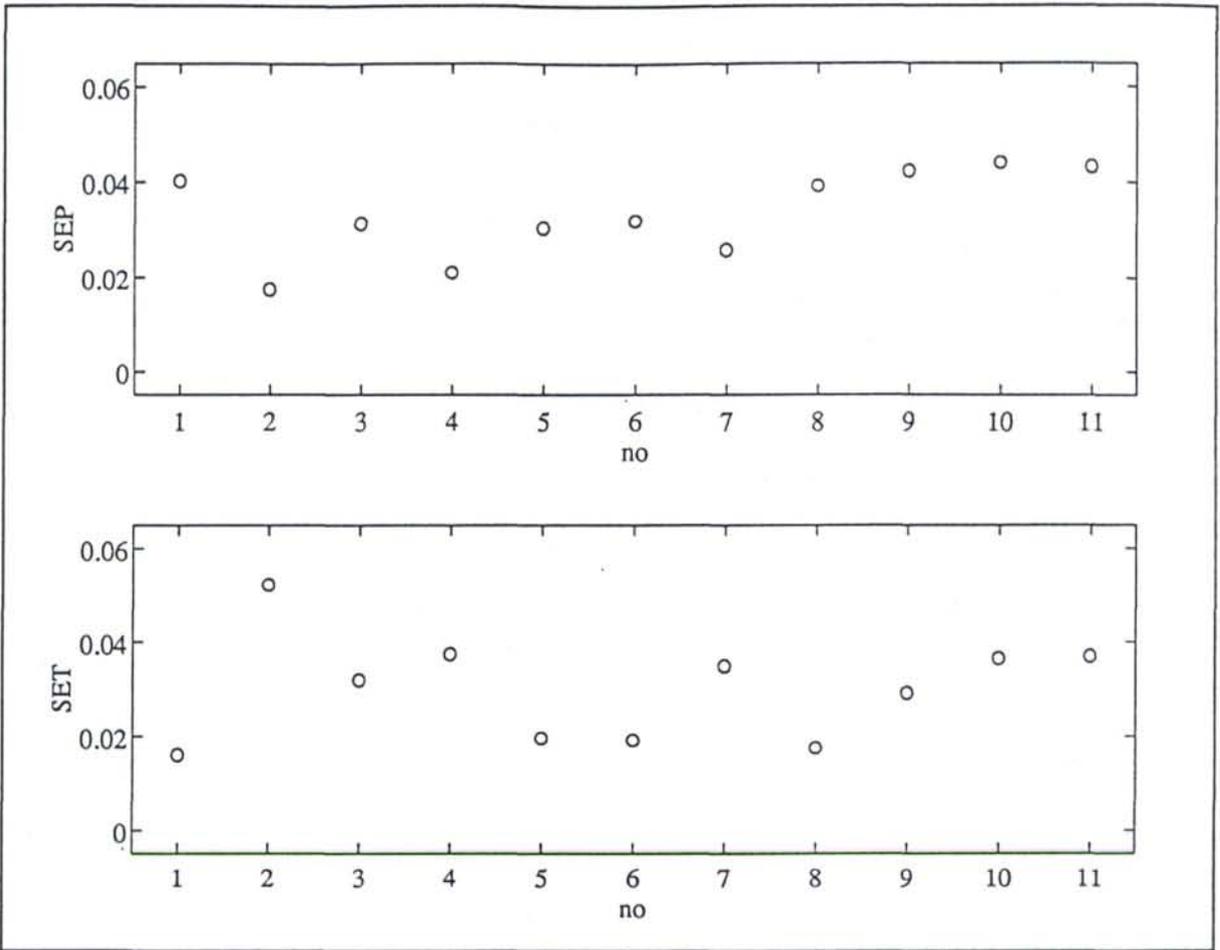


Figura III.7.17 Valores de SEP y SET obtenidos como resultado de la transferencia de calibrado mediante el procedimiento GCT en la determinación de cobre.

Análisis de los resultados

Como es habitual cuando se utiliza este procedimiento de transferencia, los mejores resultados se corresponden con la transferencia de calibrado realizada utilizando el conjunto de transferencia 1, que comprende todas las muestras de calibrado. No obstante, el elevado valor del SEP asociado con este conjunto de transferencia en el caso del cobre, Figura III.7.17, se debe únicamente a que el error relativo correspondiente a la muestra t_3 es del -9.94%. Este error es del mismo orden que el obtenido para esa misma muestra por el modelo PLS construido con todas las muestras del propio calibrado K y, por tanto, no puede responsabilizarse de la magnitud del mismo al procedimiento de transferencia de calibrado realizado ya que puede deberse a un error en la preparación o medida de las muestras.

Teniendo en cuenta la variación de los dos parámetros, SEP y SET, con los diferentes conjuntos de transferencia utilizados y los errores individuales asociados con cada una de las

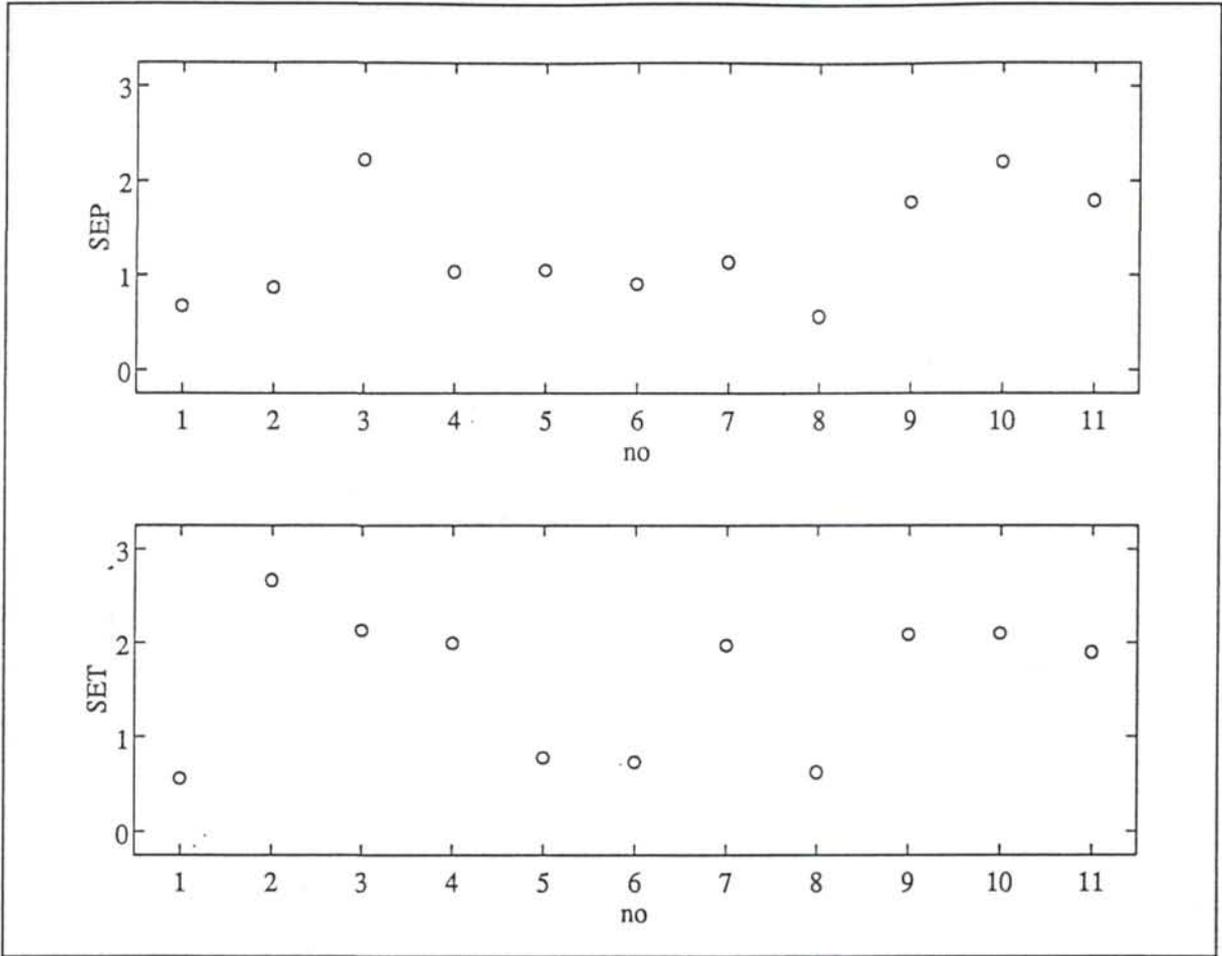


Figura III.7.18 Valores de SEP y SET obtenidos como resultado de la transferencia de calibrado mediante el procedimiento GCT en la determinación de hierro.

muestras, se obtienen resultados similares y bastante aceptables para la determinación de los dos analitos con los conjuntos 5, 6 y 8. De ellos, el conjunto de transferencia 5 es el que representa menor esfuerzo experimental ya que está constituido únicamente por 8 muestras patrón y sería posible utilizarlo de forma simultánea en la determinación de cobre y hierro.

Las concentraciones de las muestras del calibrado K calculadas mediante el procedimiento GCT y los errores asociados con ellas se muestran en la Tabla III.7.16. De forma global, los resultados mejoran los obtenidos con el procedimiento de transferencia PDS, Tabla III.7.14, y se aproximan más a los dados por el propio modelo PLS correspondiente al calibrado K, Tabla III.7.7. No obstante, este procedimiento también se muestra sensible ante algunos objetos que pudieran considerarse anómalos, tal es el caso de las muestras 7, 8, 10, 11, 12 y 13, cuyos errores ya se encontraban por encima del resto cuando las determinaciones se realizaron con los propios modelos PLS.

Tabla III.7.16 Concentraciones reales y calculadas y errores relativos que resultan de aplicar el procedimiento de transferencia de calibrado global GCT de la situación J a la situación K tomando $no = 5$.

Muestra	Cobre			Hierro		
	Conc. real (μM)	Conc. calc. (μM)	Error relativo (%)	Conc. real (μM)	Conc. calc. (μM)	Error relativo (%)
1	0.0000	0.0034	-	0.0000	-0.2484	-
2	0.2074	0.2121	2.2434	0.0000	-0.2105	-
3	0.4121	0.4152	0.7583	0.0000	0.2302	-
4	0.6141	0.6171	0.4871	0.0000	0.2154	-
5	0.8135	0.7995	-1.7180	0.0000	0.1938	-
6	1.0103	1.0232	1.2765	0.0000	-0.0535	-
7	0.0000	-0.0018	-	20.7417	21.8764	5.4705
8	0.2061	0.2130	3.3295	20.6053	19.4717	-5.5016
9	0.4094	0.4209	2.8144	20.4706	20.3408	-0.6339
10	0.6101	0.6268	2.7364	20.3377	19.0851	-6.1592
11	0.8083	0.8526	5.4865	20.2065	18.3520	-9.1777
12	1.0038	1.0684	6.4325	20.0769	18.2976	-8.8623
13	0.9975	1.0282	3.0729	39.8981	38.4643	-3.5935
14	0.0000	0.0094	-	41.2105	41.8360	1.5177
15	0.2047	0.2019	-1.3629	40.9412	41.5134	1.3977
16	0.4068	0.3815	-6.2210	40.6753	40.9959	0.7882
17	0.6062	0.5827	-3.8834	40.4129	40.3314	-0.2017
18	0.8031	0.7683	-4.3344	40.1538	39.5796	-1.4300
19	0.0000	0.0075	-	61.4118	61.4230	0.0182
20	0.2034	0.2060	1.2699	61.0130	61.2288	0.3537
21	0.4041	0.4014	-0.6620	60.6194	61.3677	1.2345
22	0.6023	0.6136	1.8714	60.2308	59.8700	-0.5990
23	0.7980	0.8135	1.9421	59.8471	59.9543	0.1791
24	0.9911	1.0110	2.0104	59.4684	59.3887	-0.1341
25	0.0000	0.0044	-	81.3506	80.4064	-1.1606
26	0.2021	0.1955	-3.2665	80.8258	80.0456	-0.9653
27	0.4015	0.3945	-1.7533	80.3077	78.8455	-1.8207
28	0.5985	0.5827	-2.6411	79.7962	79.6709	-0.1570
29	0.7929	0.7704	-2.8349	79.2911	79.6792	0.4894
30	0.9849	0.9565	-2.8880	78.7925	79.2469	0.5768
31	0.0000	-0.0162	-	101.0323	100.9859	-0.0459
32	0.2008	0.1972	-1.7714	100.3846	99.2577	-1.1226
33	0.3990	0.4015	0.6192	99.7452	98.6641	-1.0839
34	0.5947	0.5988	0.6846	99.1139	99.1213	0.0075
35	0.7879	0.8008	1.6393	98.4906	98.4205	-0.0712
36	0.9788	0.9938	1.5296	97.8750	97.8545	-0.0210
<i>i1</i>	0.3081	0.3030	-1.6656	20.5377	20.7438	1.0035
<i>i2</i>	0.5003	0.4933	-1.3927	80.0511	78.9172	-1.4164
<i>i3</i>	0.8005	0.7415	-7.3668	50.0319	48.8721	-2.3182
<i>i4</i>	0.9818	0.9914	0.9775	88.3636	87.0698	-1.4642

Aplicación del procedimiento GCT a un tercer calibrado

Al igual que en el procedimiento de transferencia PDS, para evaluar la estabilidad del método de transferencia utilizado se ha aplicado el mismo procedimiento al calibrado L, considerando en este caso únicamente el conjunto de transferencia 5 y las muestras del conjunto test. Los errores correspondientes a este nuevo análisis se muestran en Tabla III.7.17, y confirman los buenos resultados obtenidos previamente con este procedimiento al realizar la transferencia de calibrado a la situación K. Incluso, se mantiene la sensibilidad ante objetos anómalos advertida anteriormente, y que se muestra claramente en el error correspondiente a la muestra *t1*, para la que los resultados no mejoran en ningún caso, ni siquiera cuando se utilizan los modelos PLS construídos con su propio calibrado.

Tabla III.7.17 Concentraciones reales y calculadas y errores relativos que resultan al aplicar el procedimiento de transferencia de calibrado GCT de la situación J a la situación L tomando *no* = 5.

Muestra	Cobre			Hierro		
	Conc. real (μM)	Conc. calc. (μM)	Error relativo (%)	Conc. real (μM)	Conc. calc. (μM)	Error relativo (%)
1	0.0000	0.0036	–	0.0000	-0.1521	–
4	0.6141	0.6176	0.5706	0.0000	-0.0242	–
6	1.0103	1.0228	1.2408	0.0000	0.1057	–
19	0.0000	0.0078	–	61.4118	61.4840	0.1175
24	0.9911	1.0108	1.9917	59.4684	59.3504	-0.1985
31	0.0000	-0.0169	–	101.0323	100.9535	-0.0780
34	0.5947	0.6000	0.8851	99.1139	99.1134	-0.0005
36	0.9788	0.9933	1.4794	97.8750	97.8818	0.0069
<i>t1</i>	0.3081	0.2438	-20.8663	20.5377	22.8446	11.2326
<i>t2</i>	0.5003	0.5162	3.1794	80.0511	80.1663	0.1439
<i>t3</i>	0.8005	0.8114	1.3655	50.0319	49.5193	-1.0246
<i>t4</i>	0.9818	0.9724	-0.9563	88.3636	86.0455	-2.6234

Por tanto, en este caso resulta adecuado llevar a cabo la transferencia de calibrado utilizando el procedimiento GCT en lugar del procedimiento PDS, ya que los resultados obtenidos con el primero han sido considerablemente mejores en cuanto a los errores

asociados con las concentraciones calculadas. También en lo que se refiere al esfuerzo experimental requerido, ya que el procedimiento PDS implicaba a diferentes conjuntos de transferencia para cada analito, mientras que el procedimiento GCT no ha establecido diferencias tan marcadas entre los dos analitos y por ello ha sido posible utilizar el mismo conjunto de transferencia en ambas determinaciones, siendo necesarias únicamente 8 muestras patrón para determinar cobre y hierro simultáneamente.

**IV. SELECCIÓN DE VARIABLES MEDIANTE
UN ALGORITMO GENÉTICO**

IV.1 INTRODUCCIÓN

En la construcción de un modelo de regresión la determinación de cuáles son las variables predictoras verdaderamente relevantes es una cuestión de gran importancia [85]. En general, cabría pensar que cuantas más variables entren a formar parte de un modelo mejor será éste [216], ya que incluirá aquellas que estén relacionadas con la respuesta. Este razonamiento no es adecuado [217] ya que la presencia de variables que no aporten información relacionada con la respuesta aumenta el ruido de fondo en el problema y puede obstaculizar el análisis, disminuyendo la capacidad de predicción del modelo.

La finalidad de la selección de variables en un conjunto de datos determinado es lograr una reducción significativa del número de variables. Esto es importante, no desde el punto de vista económico ya que todas las variables predictoras se obtienen en una única experiencia, sino porque de esta manera se obtienen normalmente relaciones mucho más sencillas y estables.

En particular, PLS está diseñado como modelo sesgado con gran estabilidad en las predicciones. La capacidad de predicción de dicho modelo se evalúa por validación cruzada y el número de variables latentes a considerar es el que hace mínima la varianza en predicción, o lo que es lo mismo, la suma de cuadrados del error en predicción (PRESS). Por tanto, cualquier procedimiento de selección de variables ha de ser tal que conduzca a una mejora de la varianza explicada en predicción o cuando menos que la mantenga, sin atender a hacer máxima la varianza explicada, que es la base de los procedimientos habituales en MLR para la selección de variables.

Un aspecto crucial que la selección de variables puede poner de relieve en un calibrado es la necesidad de utilizar señales aparentemente no vinculadas con el analito que se está calibrando. Ello puede ser indicativo de la presencia en la señal registrada de efectos no esperados, como puede ser la existencia de reacciones químicas acopladas, la formación de compuestos intermetálicos, etc. Para mostrar esta posibilidad se ha aplicado un

procedimiento de selección de variables a todos los casos de regresión multivariante analizados en esta memoria.

Una primera aproximación al problema de selección de variables es la que el Profesor Phan-Tan-Luu [218] denomina '*sentimental screening*', que consiste en seleccionar aquellas variables que resultan más atractivas y rechazar aquellas que no cumplen esta condición. Evidentemente, este no es un procedimiento demasiado objetivo, por lo que se han buscado otras aproximaciones al problema.

Las más clásicas y aún frecuentemente utilizadas son los procedimientos de selección de variables paso a paso (*step wise*), bien de adición (se añaden variables si se verifica un determinado criterio), bien de eliminación (eliminando variables que no verifican cierto criterio). En general, los procedimientos paso a paso presentan el inconveniente de no explorar suficientemente las posibles combinaciones de variables. El problema principal asociado con estos métodos reside en los propios algoritmos, ya que cada etapa del análisis condiciona fuertemente las siguientes. De hecho, una vez seleccionadas k variables no es posible descartar ninguna de ellas a pesar de que al introducir una nueva pudieran alcanzarse mejores resultados. La misma consideración se puede hacer cuando el procedimiento paso a paso es de eliminación de variables.

Otra aproximación es la propuesta por Massart y otros [219] como método de selección de variables en determinaciones espectrofotométricas multicomponentes mediante modelos PLS. El procedimiento se basa en la selección de aquellas variables relacionadas con los coeficientes de regresión de mayor magnitud de la fórmula cerrada de PLS, habiendo normalizado y centrado las variables predictoras en una etapa previa. El método ha sido evaluado utilizando diferentes conjuntos de calibración, así como diferentes criterios para seleccionar el número de variables latentes de los modelos PLS. Los resultados sugieren que la selección de variables mejora la capacidad de predicción de los modelos PLS.

Otra alternativa eficaz se basa en el uso de procedimientos de optimización aplicados a la selección de variables, entre ellos se encuentran el enfriamiento progresivo (*simulated annealing*) [220-221], los árboles de clasificación y regresión (CART) [222], o bien el cálculo genético [26, 223-226], que ha sido la aproximación utilizada en este capítulo.

Los algoritmos genéticos [227-230] representan sistemas inteligentes que imitan a los sistemas biológicos que se auto-organizan y adaptan a su entorno. Pueden ser considerados como una técnica de búsqueda alternativa a las tradicionales en inteligencia artificial. El

cálculo genético puede encontrar soluciones casi óptimas que en general son adecuadas para resolver un determinado problema. Entre sus características está la de no exigir ninguna condición a la función a optimizar, ni siquiera al espacio en el que varían las variables. Esta técnica es buena en situaciones complejas, como la presente, en las que el espacio de búsqueda puede llegar a ser prohibitivamente extenso.

Cuando se desea resolver un problema utilizando un algoritmo genético es necesario plantearlo de forma que las posibles soluciones se puedan describir como una cadena de números o caracteres. Cada una de estas cadenas se denomina cromosoma. En la selección de variables, cada cromosoma es un vector fila que contiene tantos elementos como variables, estando cada elemento codificado como 1 si la variable correspondiente ha sido seleccionada, y como 0 si no lo ha sido.

Se debe disponer además de una función de evaluación que produce un valor numérico que representa la habilidad de cada cromosoma para resolver el problema. En el caso de la selección de variables, ésta podría venir dada por el porcentaje de varianza explicada en predicción con el subconjunto de potenciales seleccionados que representa el cromosoma.

Los operadores genéticos son respectivamente: la reproducción, el entrecruzamiento (crossover) y la mutación. La reproducción produce una nueva población de soluciones potenciales y selecciona dos de ellas como 'padres' de otras soluciones posibles. El entrecruzamiento consiste en elegir segmentos de las soluciones padre para intercambiarlos con el propósito de crear nuevas soluciones potenciales. Y por último, la mutación genera cambios arbitrarios ocasionales en candidatos a solución. Es necesaria para garantizar una adecuada 'capacidad de exploración' del espacio de las variables predictoras.

La bondad de un cromosoma se mide evaluando la respuesta que está siendo optimizada por el algoritmo. Esta población de cromosomas iniciada aleatoriamente evoluciona al producirse el entrecruzamiento o combinación de sus elementos y, en menor medida, la mutación, siendo seleccionados posteriormente los miembros más capaces. Este procedimiento se repite un cierto número de veces.

Existen diferentes versiones de algoritmos genéticos, dependiendo de la forma en que el método matemático lleve a cabo la reproducción de los cromosomas, el entrecruzamiento, etc. El algoritmo genético que se ha utilizado en esta memoria, propuesto por Leardi y otros [223, 224], ha sido diseñado específicamente para buscar el grupo de variables predictoras que conlleve la construcción del modelo de regresión PLS con mayor capacidad predictiva. Es el

primer método de selección de variables para modelos de regresión que optimiza la validación cruzada 'casi completa', y por tanto, resulta adecuado para ser aplicado a señales electroquímicas como las analizadas aquí.

IV.2 DESCRIPCIÓN DEL ALGORITMO GENÉTICO UTILIZADO

Este capítulo describe brevemente la variante concreta de algoritmo genético que se ha utilizado en esta memoria para llevar a cabo la selección de variables. Esta variante es conocida por el nombre de '*estado estacionario*' sin duplicado (steady-state without duplicates) [231]. Cada subconjunto de potenciales seleccionados se representará mediante un vector p-dimensional, \mathbf{w} , con coordenadas bivaluadas. Si se selecciona el potencial i-ésimo para formar parte del subconjunto entonces la i-ésima coordenada de \mathbf{w} será un uno, en caso contrario será un cero. Cada vector \mathbf{w} constituye un cromosoma.

Dado un cromosoma \mathbf{w} , se construye un modelo PLS validado completamente por validación cruzada utilizando de cada voltamperograma únicamente las intensidades registradas a los potenciales representados por \mathbf{w} . Mediante el valor $\text{PRESS}(\mathbf{w})$ alcanzado con el modelo se evalúa cada cromosoma \mathbf{w} . El algoritmo genético evoluciona buscando el mínimo de $\text{PRESS}(\mathbf{w})$ en el espacio de todos los posibles cromosomas (es decir, en el espacio de todos los posibles subconjuntos de potenciales) sin establecer a priori la estructura latente del modelo de calibrado, en particular el número de variables latentes, que dependerá de las variables predictoras que intervengan en el problema.

Inicialmente se selecciona una población P_0 formada por n_pop cromosomas elegidos al azar

$$P_0 = \{ \mathbf{w}(j) / j = 1, \dots, n_pop \} \quad (30)$$

y se calcula el valor de $\text{PRESS}(\mathbf{w})$ para cada elemento de P_0 .

La población de cromosomas evoluciona generando poblaciones sucesivas del siguiente modo:

- i) Dada la población P_n , se denomina $\text{PRESS}_{\max} = \max \{ \text{PRESS}(\mathbf{w}) / \mathbf{w} \in P_n \}$.

- ii) Se toman dos cromosomas al azar, w_1 y w_2 , con probabilidad inversamente proporcional al valor de $PRESS(w)$. Esto implica que la probabilidad definida por cada w viene dada por

$$pr(w) = \frac{\frac{1}{PRESS(w)}}{\sum_{w \in P_n} \frac{1}{PRESS(w)}} \quad (31)$$

y de acuerdo con esta probabilidad se realiza un sorteo para elegir los dos cromosomas.

- iii) A los cromosomas w_1 y w_2 se les aplica el operador *entrecruzamiento*, para lo cual se toma la coordenada i -ésima de ambos y se selecciona un número entre cero y uno con probabilidad uniforme, si éste es mayor que p_{cross} , se intercambia la coordenada i -ésima de w_1 con la coordenada i -ésima de w_2 . Al repetir este proceso para todas las coordenadas, desde $i=1$ hasta p , se obtienen dos nuevas cromosomas, w_3 y w_4 .
- iv) A estos nuevos cromosomas se les aplica el operador *mutación*, lo cual implica que cada una de las coordenadas de w_3 y w_4 'muta' a su valor opuesto con probabilidad p_{mut} . De este modo se obtienen los cromosomas w_5 y w_6 .
- v) En este punto se genera la población siguiente P_{n+1} de forma que

Si w_5 no está en la población actual P_n se calcula $PRESS(w_5)$.

Si $PRESS(w_5) < PRESS_{max}$ entonces w_5 sustituye al cromosoma correspondiente a $PRESS_{max}$ y se calcula el nuevo valor de $PRESS_{max}$.

Si w_6 no está en la población actual se calcula $PRESS(w_6)$.

Si $PRESS(w_6) < PRESS_{max}$ entonces w_6 sustituye al cromosoma correspondiente a $PRESS_{max}$.

- vi) Se vuelve al paso i) hasta que se detenga el algoritmo.

La mejor solución, en cada población P_n , es el subconjunto de potenciales w_n tal que $PRESS(w_n) = \min \{PRESS(w) / w \in P_n\}$. En general, si el algoritmo evoluciona correctamente, la última población está formada por subconjuntos de potenciales tales que los correspondientes calibrados tienen valores de $PRESS$ suficientemente pequeños. El algoritmo se detiene cuando se han generado n_{tot} cromosomas descendientes.

Como se indica en la referencia [231] el procedimiento de no reemplazar la población entera por una nueva en cada caso, e introducir cromosomas distintos siempre que su evaluación sea mejor que la del peor de la población (elitismo en grado máximo) da buenos resultados en problemas de optimización en los que la respuesta no tiene ruido, como ocurre en el presente caso, ya que al repetir PLS con el mismo cromosoma se obtiene el mismo valor de PRESS.

Procedimiento global

Este es seguramente el aspecto que más diferencia al algoritmo genético utilizado en esta memoria ya que tiene en cuenta una condición que se considera esencial en los calibrados que se realizan en determinaciones analíticas y que consiste en conseguir un modelo con la máxima capacidad de predicción. Esto implica que el subconjunto de potenciales elegido tiene que ser validado utilizando datos que no hayan intervenido en la construcción del propio modelo. Por esta razón el procedimiento seleccionado incluye una validación externa basada en el SEP definido en la ecuación 25 del *Epígrafe III.3.4.2*.

El espacio definido por todos los cromosomas (todos los posibles subconjuntos de potenciales) es un conjunto discreto, por lo que carece de sentido considerar que los sucesivos cromosomas óptimos en cada población w_n ($n \rightarrow \infty$) convergen al óptimo de PRESS al avanzar en el proceso. Cuando se utilizan algoritmos genéticos, en su lugar se verifican convergencias en probabilidad. Entonces lo que se garantiza es que la probabilidad de que la población n -ésima contenga al óptimo crece hacia uno cuando $n \rightarrow \infty$, pero no existe la certeza de que el cromosoma óptimo $w_{n_{tot}}$ sea el óptimo global, ni se tiene evaluada ninguna medida de lo que dista del óptimo.

Este es el aspecto que más diferencia el modo en que un algoritmo genético evoluciona hacia un óptimo frente a los procedimientos de optimización convencionales. Además, en los problemas que están siendo analizados es imposible buscar una relación funcional entre cada subconjunto de potenciales y el valor de PRESS obtenido cuando se utilizan las intensidades registradas a esos potenciales como variables predictoras de la concentración.

A continuación se describe esquemáticamente el procedimiento global que garantiza la validación 'casi completa' de la selección de potenciales realizada por el algoritmo genético

- i) Un conjunto, D , de m muestras de calibrado (cuyos voltamperogramas constituyen las filas de la matriz de las variables predictoras) se divide en n_sep subconjuntos disjuntos, D_i , tales que

$$D = \bigcup_{i=1}^{n_sep} D_i \quad (32)$$

- ii) Desde que i vale 1 hasta que vale n_sep

Se definen el conjunto de aprendizaje ($D - D_i$) y el de evaluación D_i .

Se optimiza PRESS por validación cruzada completa usando n_can conjuntos de cancelación mediante el algoritmo genético de la sección precedente.

Para cada cromosoma w de la última población se calcula $SEP(w)$ con D_i .

- iii) Se eliminan todos los w tales que el valor de $SEP(w)$ expresado como porcentaje de varianza explicada no alcanza un valor umbral preseleccionado, v_umb , en ninguno de los D_i . De este modo, se garantiza que los n_accept subconjuntos de potenciales seleccionados son capaces de proporcionar un modelo de calibrado de suficiente calidad en predicción, al menos con uno de los conjuntos de evaluación.

- iv) Desde que j vale 1 hasta que vale n_accept

Desde que i vale 1 hasta que vale n_sep

Se calcula el máximo de SEP con los potenciales correspondientes al cromosoma j cuando se recorren los conjuntos $D - D_i$. Llamaremos $SEP(j, i)$ a cada uno de estos valores.

Se toma como índice de calidad del cromosoma aceptado j -ésimo la media geométrica de los SEP que le corresponden, es decir

$$IQ(j) = (SEP(j, 1) \times SEP(j, 2) \times \dots \times SEP(j, n_sep))^{1/n_sep} \quad (33)$$

El cromosoma con mayor valor de IQ se toma como solución por validación cruzada al problema de las selección de variables.

IV.3 SELECCIÓN DE VARIABLES EN LA DETERMINACIÓN POLAROGRÁFICA DE BENZALDEHÍDO

Se ha llevado a cabo la selección de potenciales mediante un algoritmo genético en la determinación polarográfica de benzaldehído a diferentes valores de pH. Los conjuntos de datos con los que ésta se ha realizado coinciden con los señalados en el *Epígrafe II.2.4*, en el que estos datos fueron utilizados para construir modelos de regresión multivariante PLS con todas las variables predictoras.

Como ya se indicó en la *Sección II.2*, la determinación de benzaldehído ha sido llevada a cabo mediante polarografía diferencial de impulsos en las condiciones instrumentales allí especificadas, utilizando disoluciones tampón McIlvaine como electrólito soporte. Sobre un volumen determinado de solución tampón se realizaron una serie de adiciones de benzaldehído y se registraron los polarogramas correspondientes. Este análisis se llevó a cabo considerando diez valores de pH diferentes, obteniéndose para cada uno de ellos un conjunto de datos constituido por 7 objetos (las adiciones de benzaldehído realizadas) y 118 variables predictoras (las intensidades registradas a 118 potenciales que están igualmente espaciados entre -800 mV y -1502 mV).

Analizados los datos correspondientes a las Figuras II.2.3a a II.2.3e se comprobó la inviabilidad de un tratamiento univariante de los mismos, haciéndose necesaria la utilización de una técnica de regresión multivariante como PLS para resolver los problemas de falta de especificidad de los máximos de pico originada por la presencia de reacciones acopladas. Los resultados obtenidos con los modelos PLS construidos fueron muy satisfactorios. Como ya se ha indicado, en este capítulo se va a realizar un procedimiento de selección de variables cuyos resultados podrán ser comparados con los correspondientes al análisis señalado en último lugar. Asimismo, dicho procedimiento ofrece la posibilidad de llevar a cabo evaluaciones cualitativas de los procesos que estén teniendo lugar.

IV.3.1 Selección de variables

Para aplicar el procedimiento de selección de variables descrito en la *Sección IV.1*, y después de realizar algunos ensayos, se han considerado adecuados los siguientes parámetros

Grupos de cancelación externos (n_{sep})	5
Número de cromosomas en la población (n_{pop})	20
Probabilidad de entrecruzamiento (p_{cross})	50%
Probabilidad de mutación (p_{mut})	1%
Grupos de cancelación internos (n_{can})	7
Criterio de parada para cada ciclo externo (n_{tot})	100
Umbral para la varianza en predicción (v_{umb})	80%

Este algoritmo genético ha sido aplicado de forma independiente a los datos correspondientes a cada pH experimental. Los resultados del procedimiento de selección de variables se encuentran en la Tabla IV.2.1, en la que se muestran las variables seleccionadas por el algoritmo genético para cada valor de pH. Las Figuras IV.3.1a-j muestran cómo

Tabla IV.3.1 Selección de variables realizada por el algoritmo genético para los distintos valores de pH.

pH	VARIABLES SELECCIONADAS (potenciales / mV)
2.2	-830, -902, -1004, -1058, -1394, -1412, -1442, -1454, -1496
2.8	-926, -938, -1100, -1136, -1148, -1274, -1298, -1430, -1484
3.4	-848, -926, -1052, -1088, -1232, -1292, -1460
4.0	-830, -1028, -1034, -1244, -1298, -1358, -1406, -1496
4.6	-1040, -1118, -1232, -1274, -1328, -1376, -1472
5.2	-1004, -1124, -1184, -1244, -1262, -1346, -1358
5.8	-986, -1016, -1328, -1490
6.4	-908, -1106, -1130, -1262, -1280, -1304
7.0	-800, -824, -1268, -1298, -1394, -1418, -1454, -1496
7.6	-1064, -1112, -1142, -1184, -1232, -1268, -1298, -1370

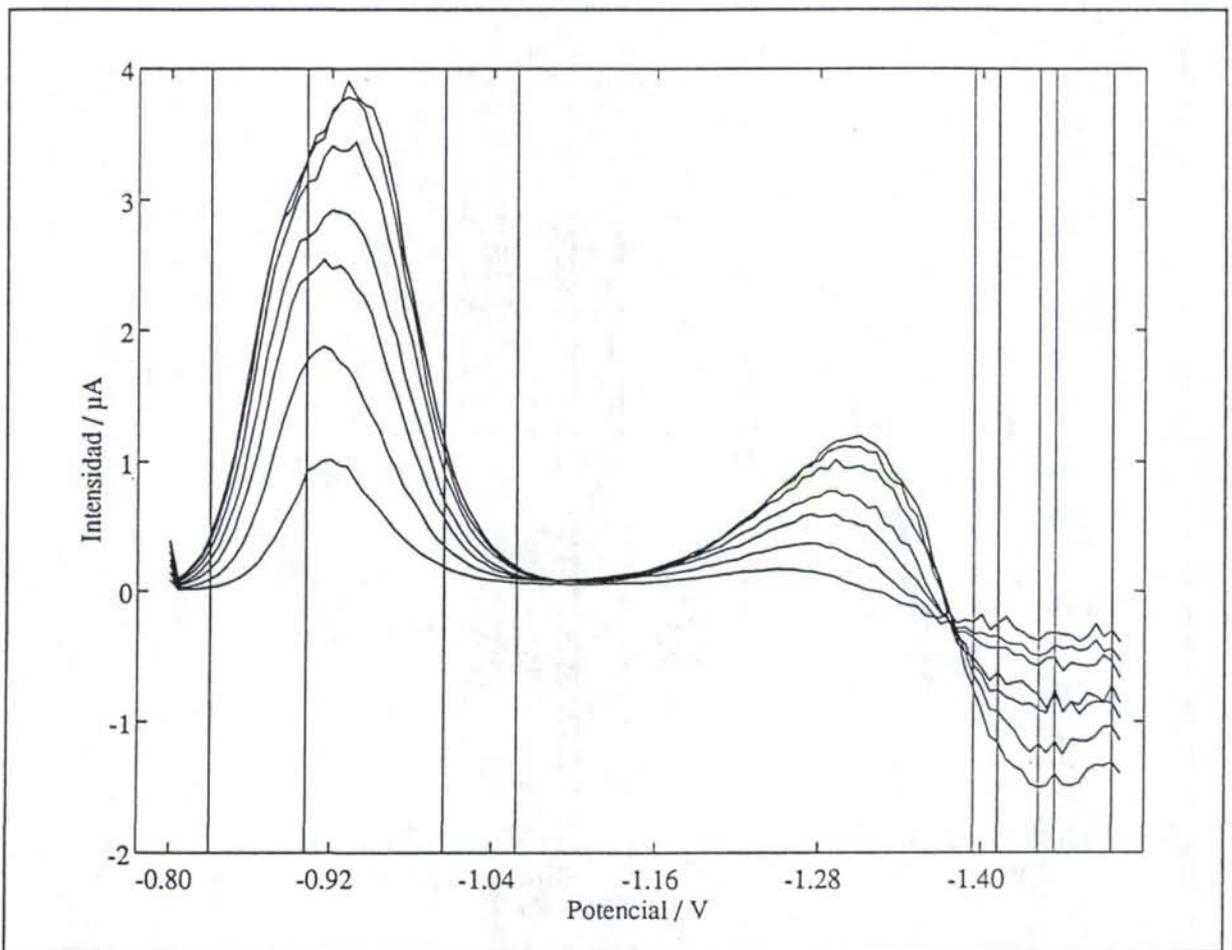


Figura 1a Variables seleccionadas por el algoritmo genético para pH = 2.2.

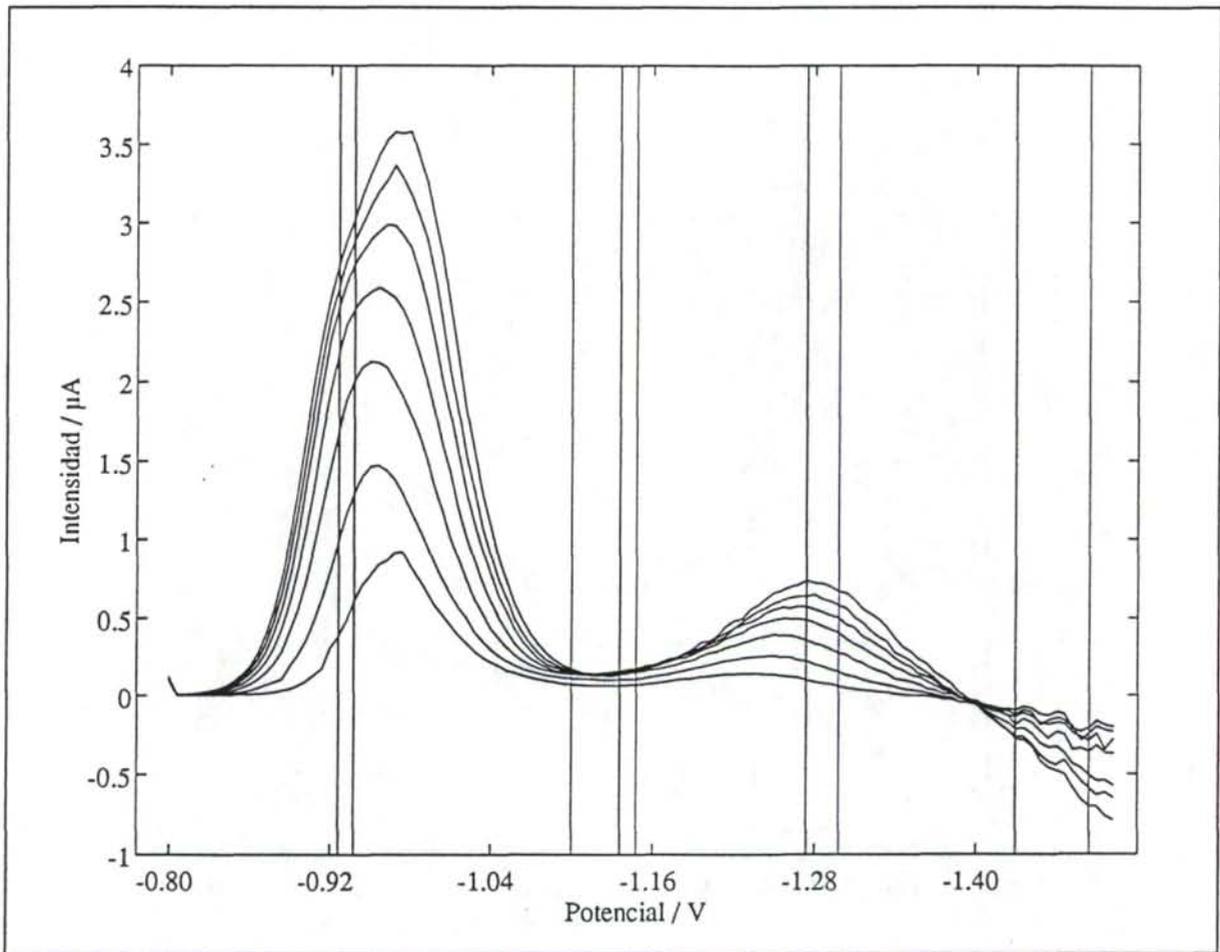


Figura 1b Variables seleccionadas por el algoritmo genético para pH = 2.8.

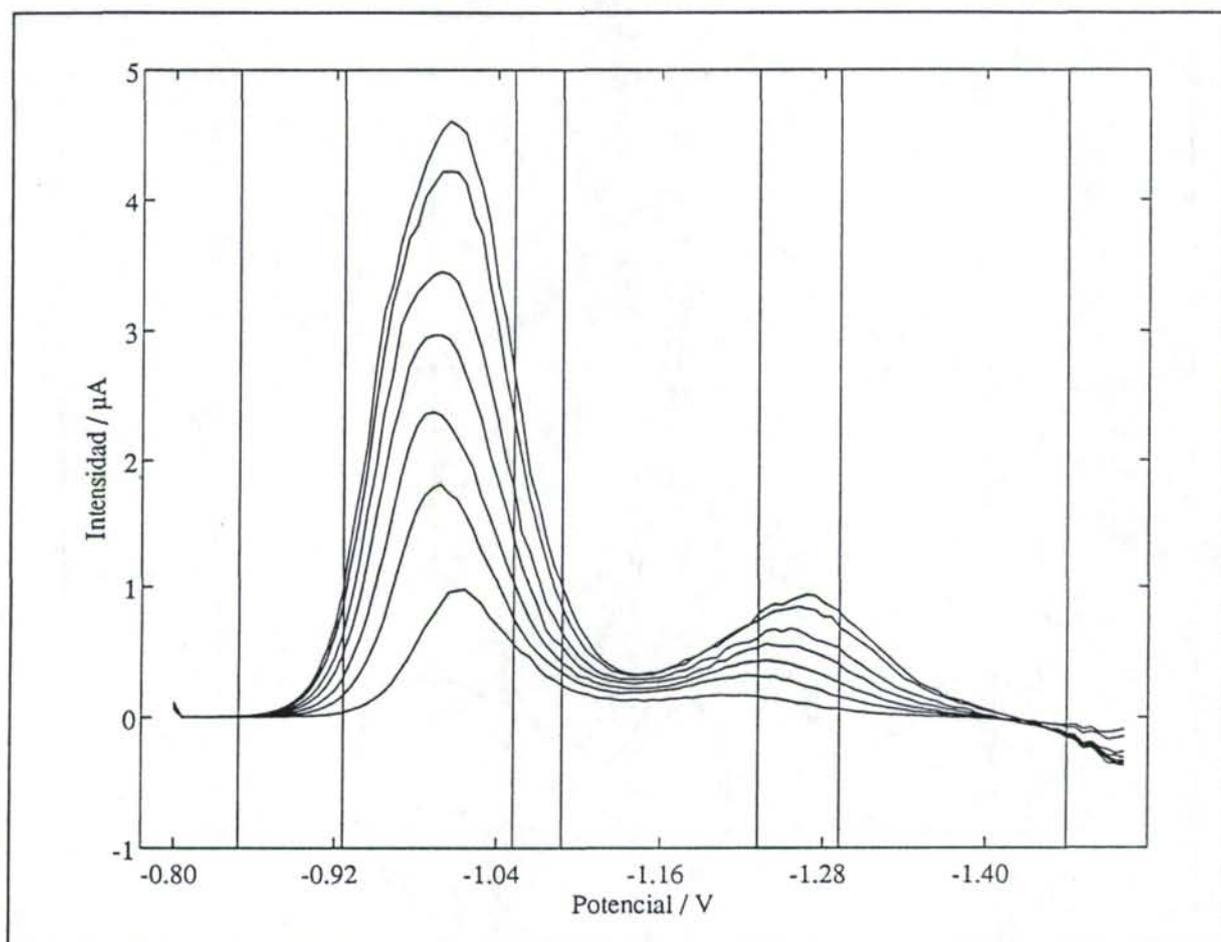


Figura 1c Variables seleccionadas por el algoritmo genético para pH = 3.4.

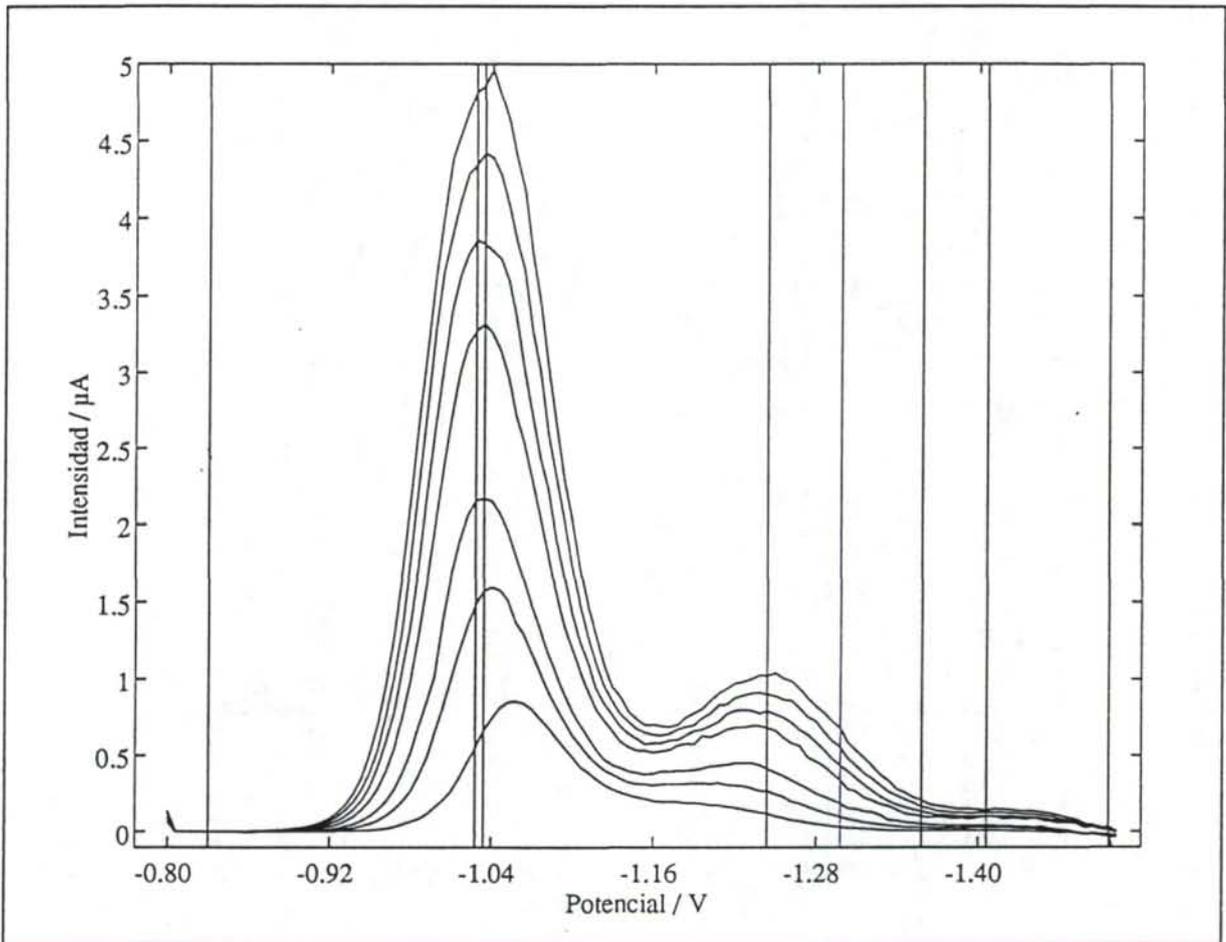


Figura 1d Variables seleccionadas por el algoritmo genético para pH = 4.0.

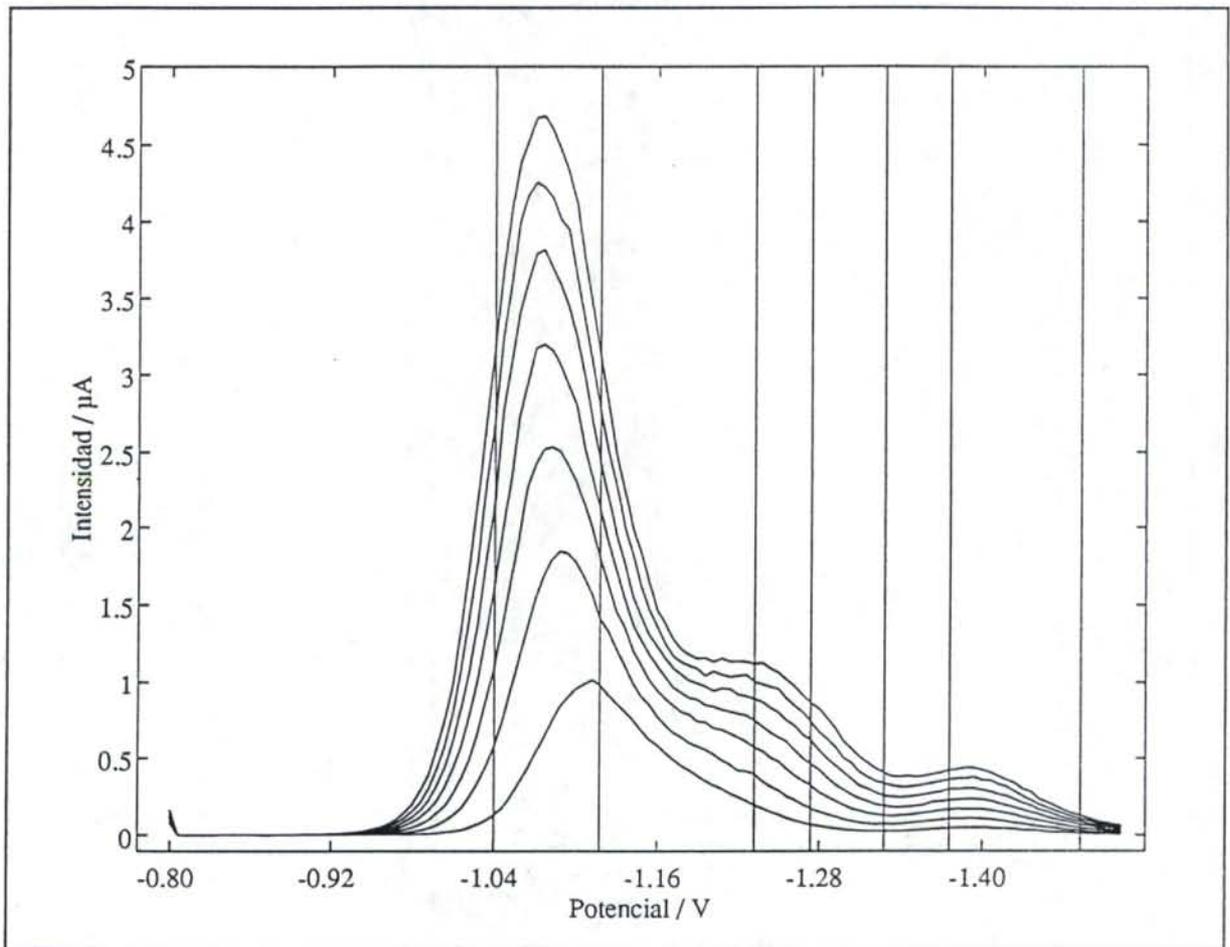


Figura 1e Variables seleccionadas por el algoritmo genético para pH = 4.6.

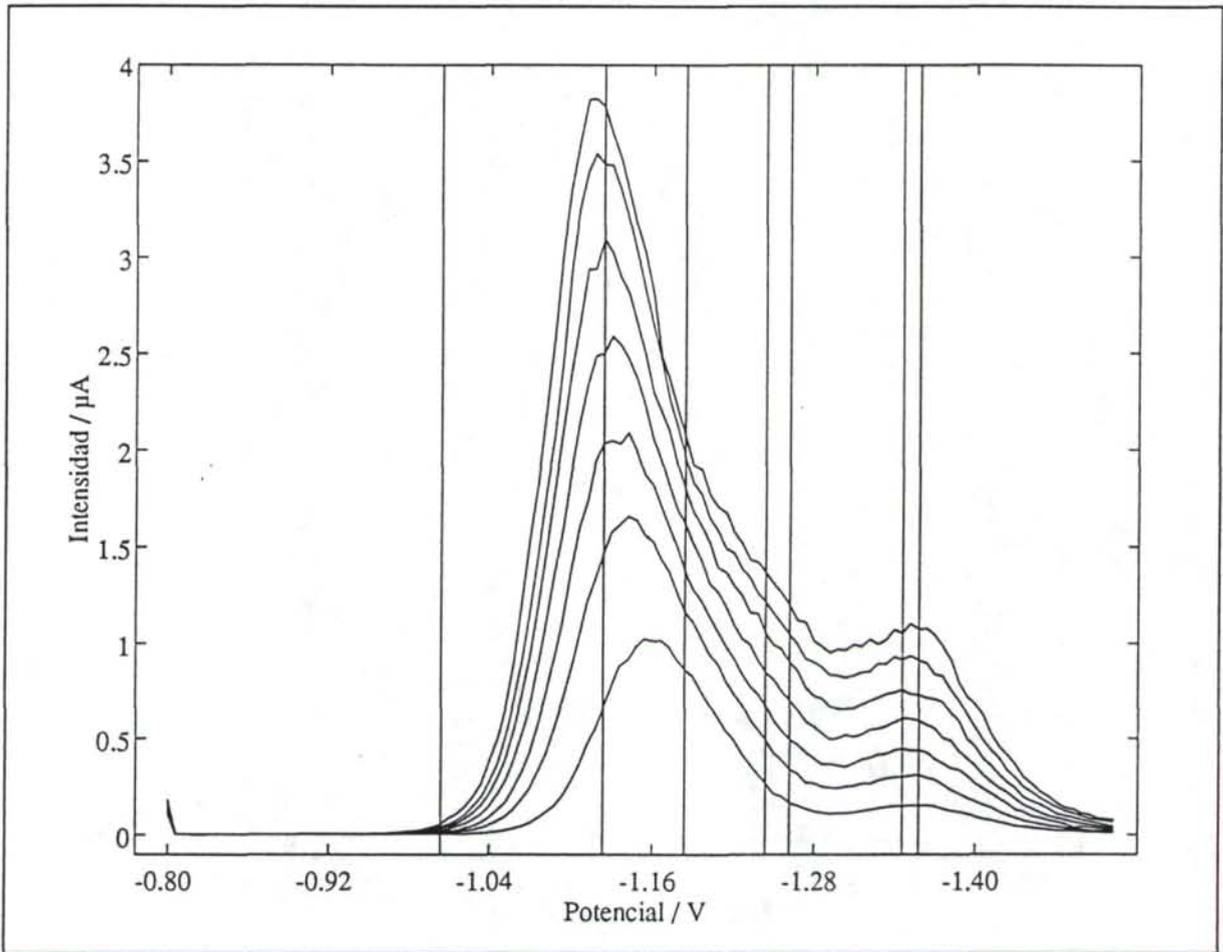


Figura 1f Variables seleccionadas por el algoritmo genético para pH = 5.2.

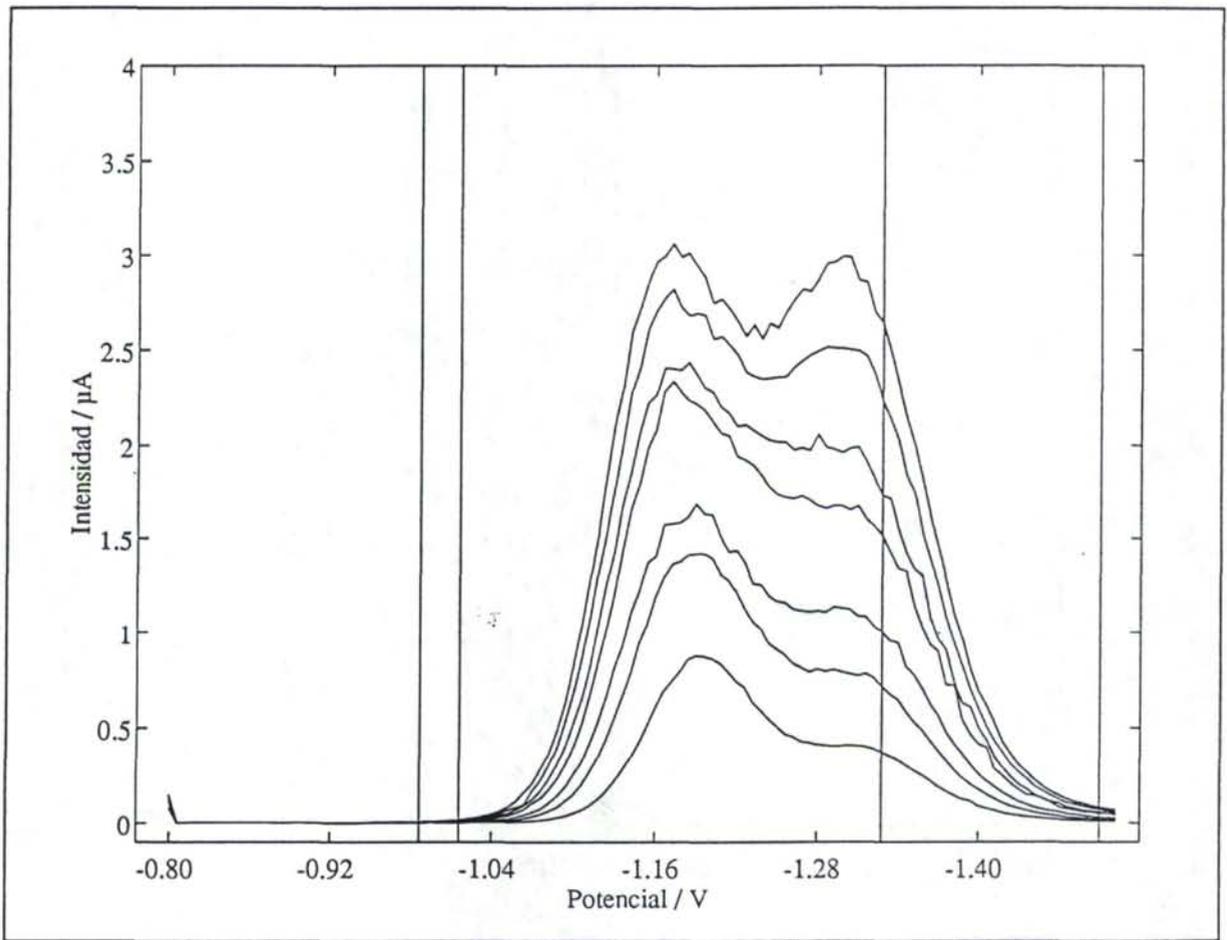


Figura 1g Variables seleccionadas por el algoritmo genético para pH = 5.8.

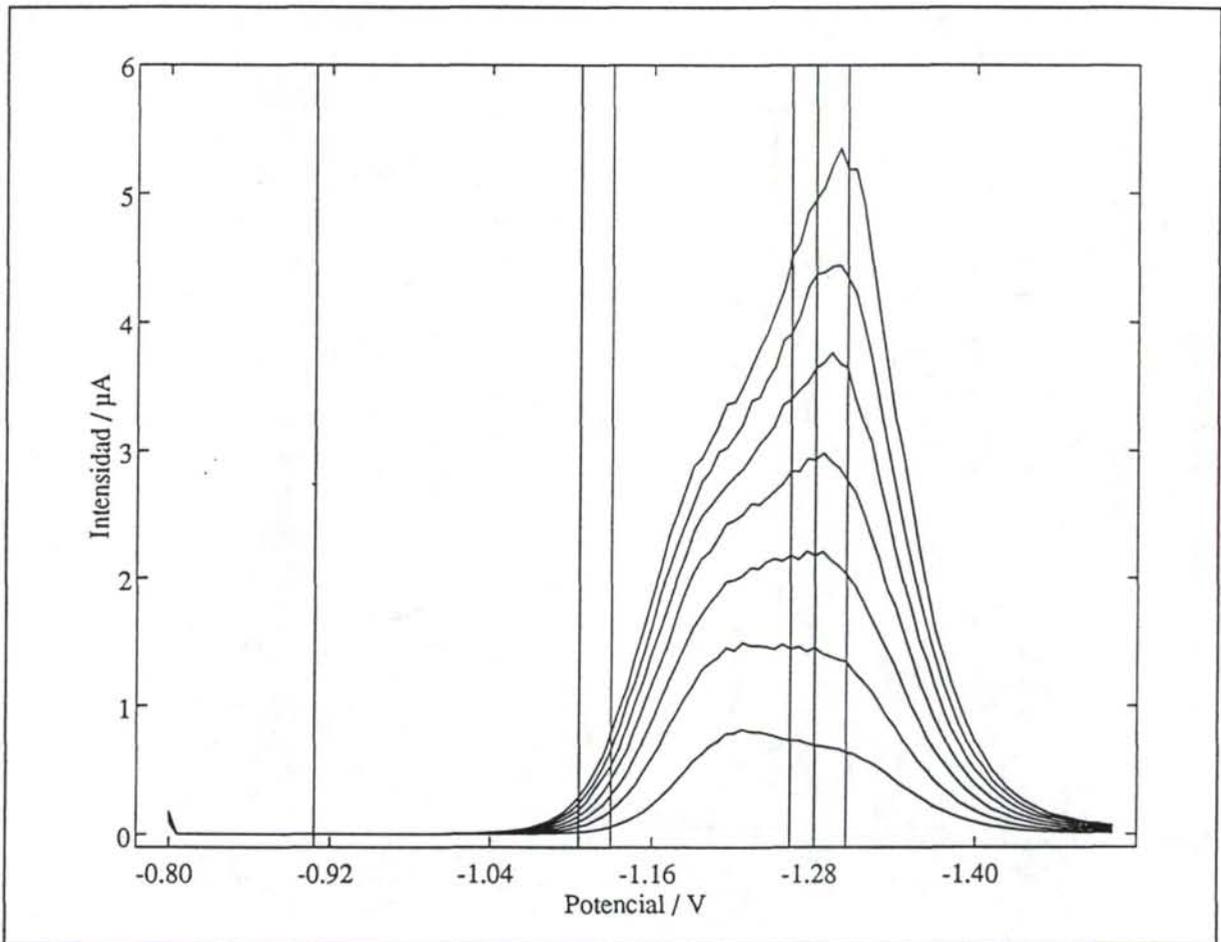


Figura 1h Variables seleccionadas por el algoritmo genético para $\text{pH} = 6.4$.

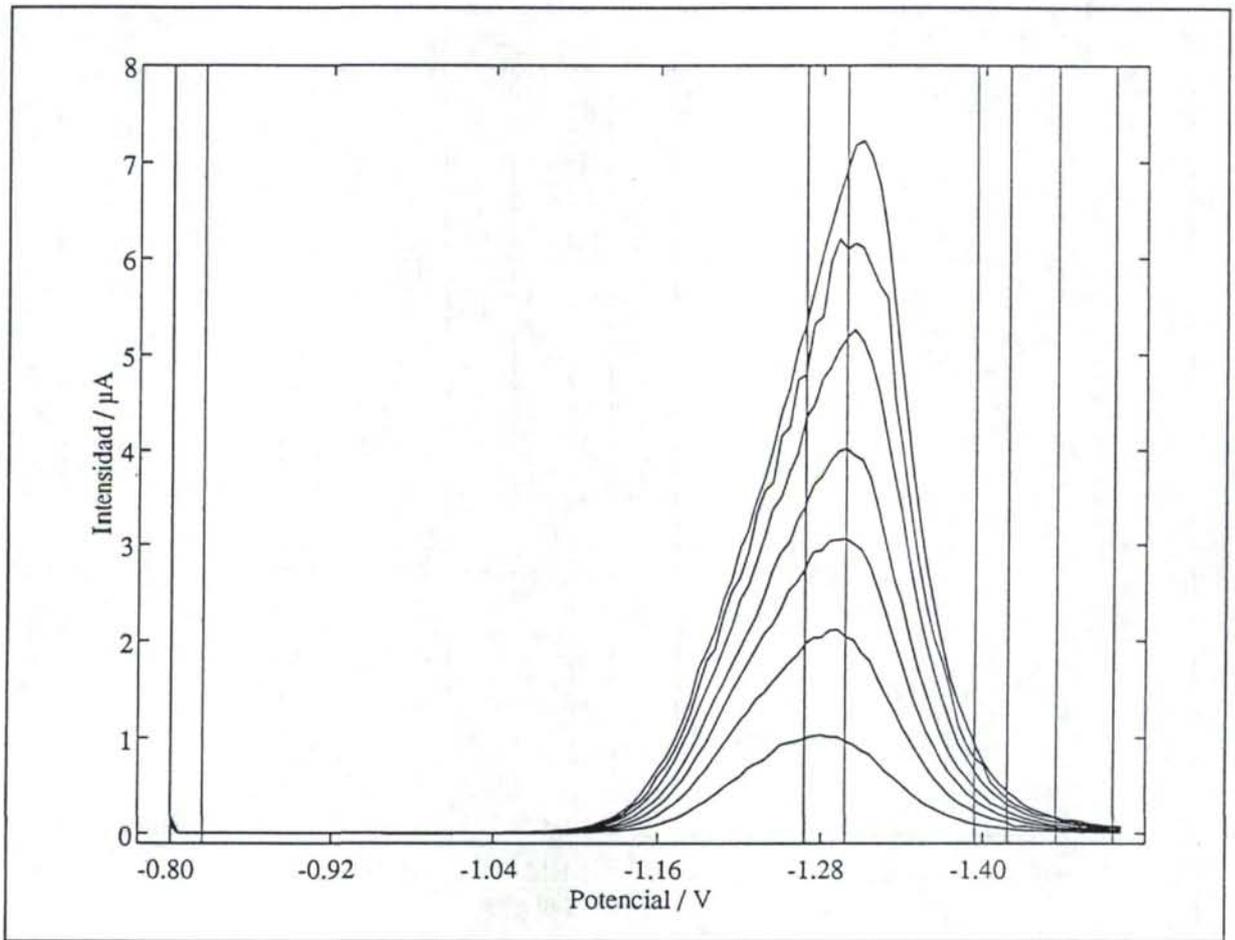


Figura 1i Variables seleccionadas por el algoritmo genético para pH = 7.0.

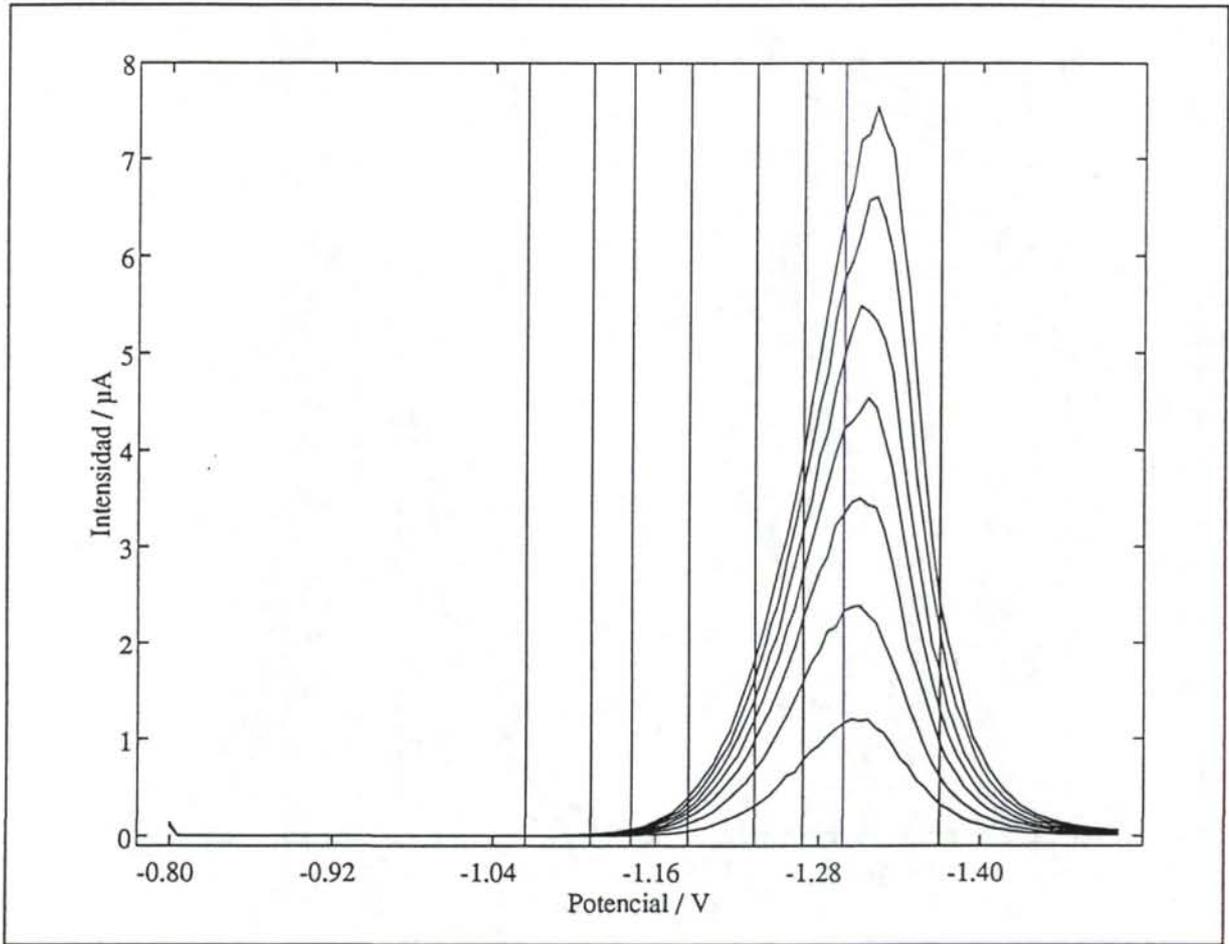


Figura 1j Variables seleccionadas por el algoritmo genético para pH = 7.6.

muchos de los potenciales seleccionados se corresponden con zonas próximas al máximo de pico, mientras que otras recorren las diferentes zonas del polarograma.

Un aspecto sumamente útil del procedimiento de selección de variables es que muestra cuales son las zonas de potencial donde se producen procesos complejos relacionados con el analito de interés. De hecho, se puede observar en las Figuras IV.3.1a-j cómo los valores de potencial seleccionados por el algoritmo genético van variando a medida que cambia el pH de acuerdo con los procesos electródicos ya indicados en el *Capítulo II.2*.

IV.3.2 Construcción y evaluación de modelos PLS

Construcción de los modelos PLS

Con las variables seleccionadas por el algoritmo genético se han construido modelos de regresión PLS independientes para cada valor experimental de pH, considerando como variable respuesta la concentración de benzaldehído correspondiente a las siete adiciones de benzaldehído realizadas. Por tanto, en este caso, el número de variables predictoras que constituyen cada uno de los modelos PLS construidos depende de la selección previa, mientras que el número de objetos considerado en cada calibrado es siempre el mismo.

En la Tabla IV.2.2 se resumen los porcentajes de varianza explicada por los modelos PLS construidos con las variables predictoras seleccionadas por el algoritmo genético para cada valor experimental de pH. En negrilla y cursiva se señalan los valores de varianza correspondientes al número de variables latentes pertenecientes al modelo PLS seleccionado en cada caso.

Estos resultados se pueden comparar con los obtenidos cuando los modelos PLS se construyeron con todas las variables predictoras, Tabla II.2.4. En general, los modelos construidos con las variables seleccionadas por el algoritmo genético están formados por un número de variables latentes mayor, salvo para los modelos correspondientes a los valores de pH 4.6 y 6.4, en los que éste disminuye. Asimismo, en casi todos los casos, los porcentajes de varianza explicada por los modelos son ligeramente más altos cuando se realiza la selección de variables previa. Únicamente tres de los modelos construidos ven disminuido el valor de varianza explicada, respecto del obtenido con todas las variables predictoras, aunque la diferencia mayor es tan sólo de un 0.08%.

Tabla IV.3.2 Porcentajes de varianza explicada por los modelos PLS construidos con las variables seleccionadas por el algoritmo genético, para cada valor experimental de pH, en función del número de variables latentes. En negrilla y cursiva se señala el número de variables latentes seleccionado para cada modelo.

pH	Bloque	Varianza	Variables latentes			
			1	2	3	4
2.2	Y	Explicada	92.29	97.43	99.72	<i>99.86</i>
		En predicción	90.29	97.17	99.44	<i>99.48</i>
	X	Explicada	88.68	95.49	98.00	<i>98.01</i>
2.8	Y	Explicada	91.75	99.52	99.62	<i>99.88</i>
		En predicción	88.89	99.08	99.27	<i>99.66</i>
	X	Explicada	94.83	96.23	99.04	<i>99.30</i>
3.4	Y	Explicada	98.60	99.60	<i>99.75</i>	99.97
		En predicción	98.40	98.79	<i>99.51</i>	99.32
		Explicada	93.25	93.43	<i>98.43</i>	98.25
4.0	Y	Explicada	98.43	98.83	99.49	<i>99.81</i>
		En predicción	97.87	97.34	98.75	<i>99.49</i>
	X	Explicada	97.96	98.62	99.09	<i>99.60</i>
4.6	Y	Explicada	99.58	<i>100.00</i>	100.00	100.00
		En predicción	99.52	<i>99.99</i>	99.98	99.98
	X	Explicada	86.53	<i>98.66</i>	99.21	99.77
5.2	Y	Explicada	96.23	99.41	<i>99.95</i>	99.95
		En predicción	95.27	98.88	<i>99.91</i>	99.86
	X	Explicada	89.57	96.50	<i>98.84</i>	99.38
5.8	Y	Explicada	98.99	99.08	<i>99.48</i>	–
		En predicción	98.68	98.49	<i>99.01</i>	–
	X	Explicada	99.89	99.92	<i>99.95</i>	–
6.4	Y	Explicada	<i>99.98</i>	99.98	99.98	99.97
		En predicción	<i>99.98</i>	99.97	99.83	99.78
	X	Explicada	<i>96.40</i>	98.94	99.40	99.99
7.0	Y	Explicada	99.57	99.86	<i>99.90</i>	99.99
		En predicción	99.37	99.76	<i>99.80</i>	99.61
	X	Explicada	95.99	99.11	<i>99.68</i>	99.63
7.6	Y	Explicada	98.01	99.96	<i>99.97</i>	99.99
		En predicción	97.43	99.89	<i>99.91</i>	99.87
	X	Explicada	96.27	99.45	<i>99.55</i>	99.90

También se observa que, en los modelos construidos tras la selección de variables, las diferencias entre los valores de varianza explicada y varianza en predicción son menores que las correspondientes a los modelos construidos sin selección de variables. Es decir, que los valores de varianza en predicción se encuentran más próximos a los valores de varianza explicada que en el caso anterior, lo cual es indicativo de que los modelos PLS construidos en esta última etapa del análisis tienen mayor capacidad de predicción y estabilidad. No obstante, las discrepancias entre los valores de varianza explicada entre un método y otro son mínimas.

Interpretabilidad de las variables latentes

Evidentemente, las variables latentes de los modelos PLS construidos una vez realizada la selección de variables no son exactamente las mismas que las correspondientes a los modelos desarrollados en el *Epígrafe II.2.4*. De hecho, el número de variables predictoras que han entrado a formar parte de los primeros es significativamente menor, ya que el algoritmo genético utilizado selecciona únicamente aquellas que mejoran la capacidad predictiva de los modelos de regresión. Además, el número de variables latentes que constituyen cada uno de los modelos construidos en este caso difiere del correspondiente a los modelos construidos inicialmente.

En las Figuras IV.3.2a-j se muestran los loadings correspondientes a los modelos PLS obtenidos tras la selección de variables. La estructura de los loadings de la primera variable latente reproduce en los modelos construidos para todos los valores experimentales de pH, como en el caso anterior, la aportación de cada variable predictora al polarograma original. Esta variable latente está fuertemente relacionada con la concentración de benzaldehído, explicando porcentajes de varianza muy elevados en todos los casos.

En cuanto al resto de las variables latentes, su interpretabilidad depende tanto del número de variables predictoras seleccionadas por el algoritmo genético como de la complejidad de los procesos implicados en el análisis. Por ejemplo, la Figura IV.3.2a muestra cómo el algoritmo genético ha seleccionado las variables predictoras relacionadas con la señal principal y con la reducción de los protones del medio en el modelo PLS construido para el valor de pH 2.8. Todas las variables predictoras seleccionadas, un total de nueve, intervienen de forma similar en todas las variables latentes significativas, lo cual dificulta enormemente el establecer una relación entre éstas y alguno de los fenómenos químicos que se están produciendo.

Sin embargo, los loadings correspondientes al modelo PLS construido para el pH 7.0, Figura IV.3.2j, señalan cómo las variables latentes segunda y tercera están relacionadas con los potenciales de las colas del pico polarográfico, modelando la información relativa a la concentración de benzaldehído contenida en esas zonas del polarograma, así como el ruido de fondo.

Por tanto, al efectuar la selección de variables es posible que se pierda parte la interpretabilidad química de las variables latentes. Esto supone que, dependiendo de cuál sea la finalidad de análisis, puede ser o no conveniente la utilización del algoritmo genético para seleccionar variables. Cuando se trate de estudiar un sistema químico y de relacionar la variabilidad de los datos con la realidad química, no es adecuado disminuir tan drásticamente el número de variables predictoras. Mientras que, si lo que se trata es de determinar la concentración de analito, el uso de este algoritmo genético en la selección de variables conduce a la utilización de aquella parte de la información que lleva a una mejor validación cruzada disminuyendo la influencia del ruido de fondo, y por tanto a una predicción mejor.

Análisis de los errores

En la Tabla IV.2.4 se evalúan las diferencias existentes entre las concentraciones de benzaldehído reales y las calculadas por los modelos PLS a través de los correspondientes errores relativos. Para calcular éstos, se han considerado los valores de concentración predichos por los modelos PLS construidos con el número de variables latentes indicado en la Tabla IV.2.2.

Para llevar a cabo el análisis de estos errores, al igual que se hizo en el *Epígrafe II.2.4* en el que los modelos se construyeron con todas las variables predictoras, se ha evaluado el efecto que sobre los mismos tienen la variación del pH del medio y de las adiciones de benzaldehído realizadas. En primer lugar, se ha tratado de hacer un análisis de la varianza (ANOVA) sobre los datos de la Tabla IV.2.3, comprobándose previamente si se cumplían o no las hipótesis de normalidad de restos e igualdad de varianzas.

Para verificar la primera de las hipótesis previas, la de normalidad de restos, se han utilizando los test de hipótesis [58], al nivel de confianza del 95%, siguientes: el test de la χ^2 ($P_{\text{actual}} = 2.01 \cdot 10^{-6}$) y el test de Kolmogorov-Smirnov ($P_{\text{actual}} = 2.6 \cdot 10^{-2}$). Ambos concluyen, a diferencia del caso de los modelos PLS construidos con todas las variables latentes, que los restos no son normales.

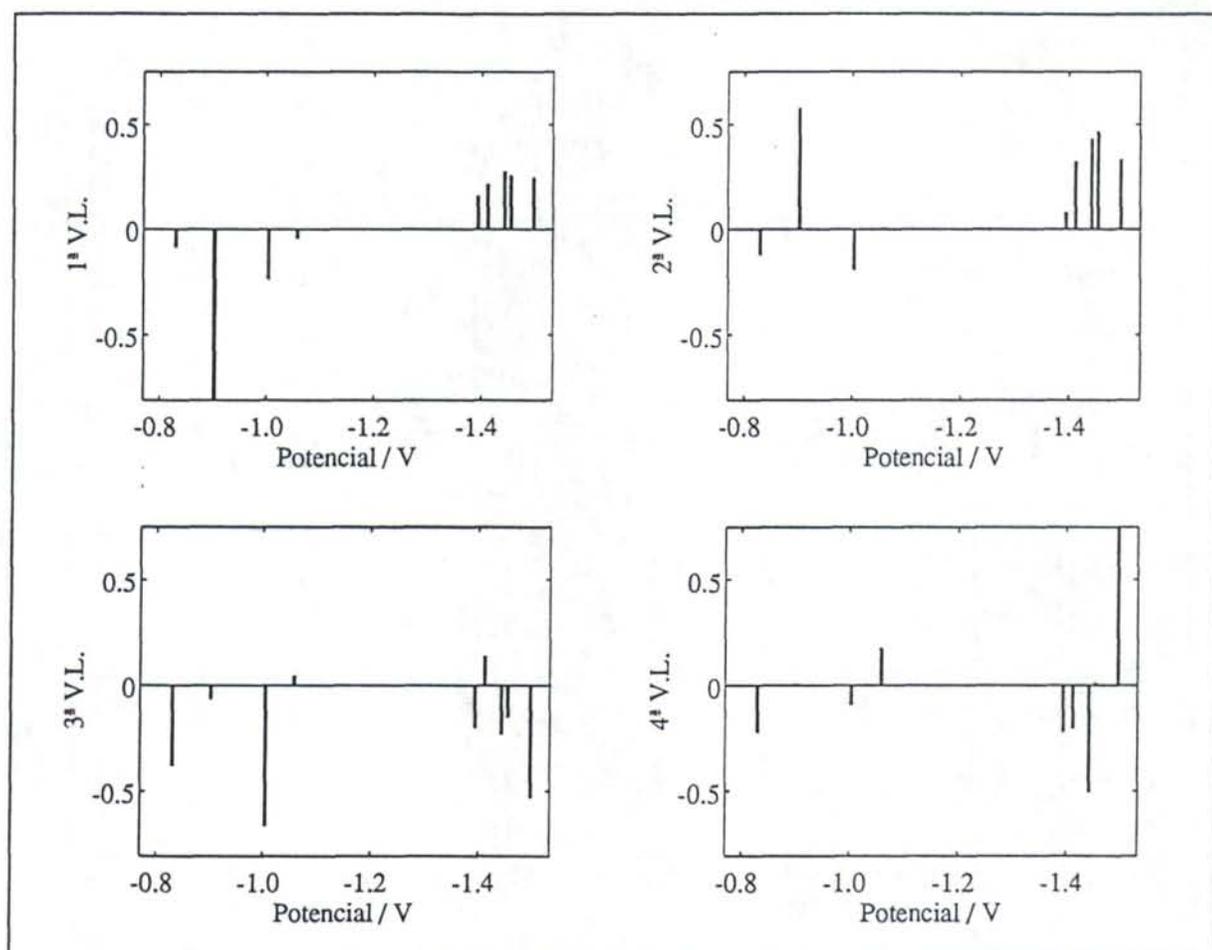


Figura IV.2.2a Loadings correspondientes al modelo PLS construido para $\text{pH} = 2.2$ con las variables seleccionadas por el algoritmo genético.

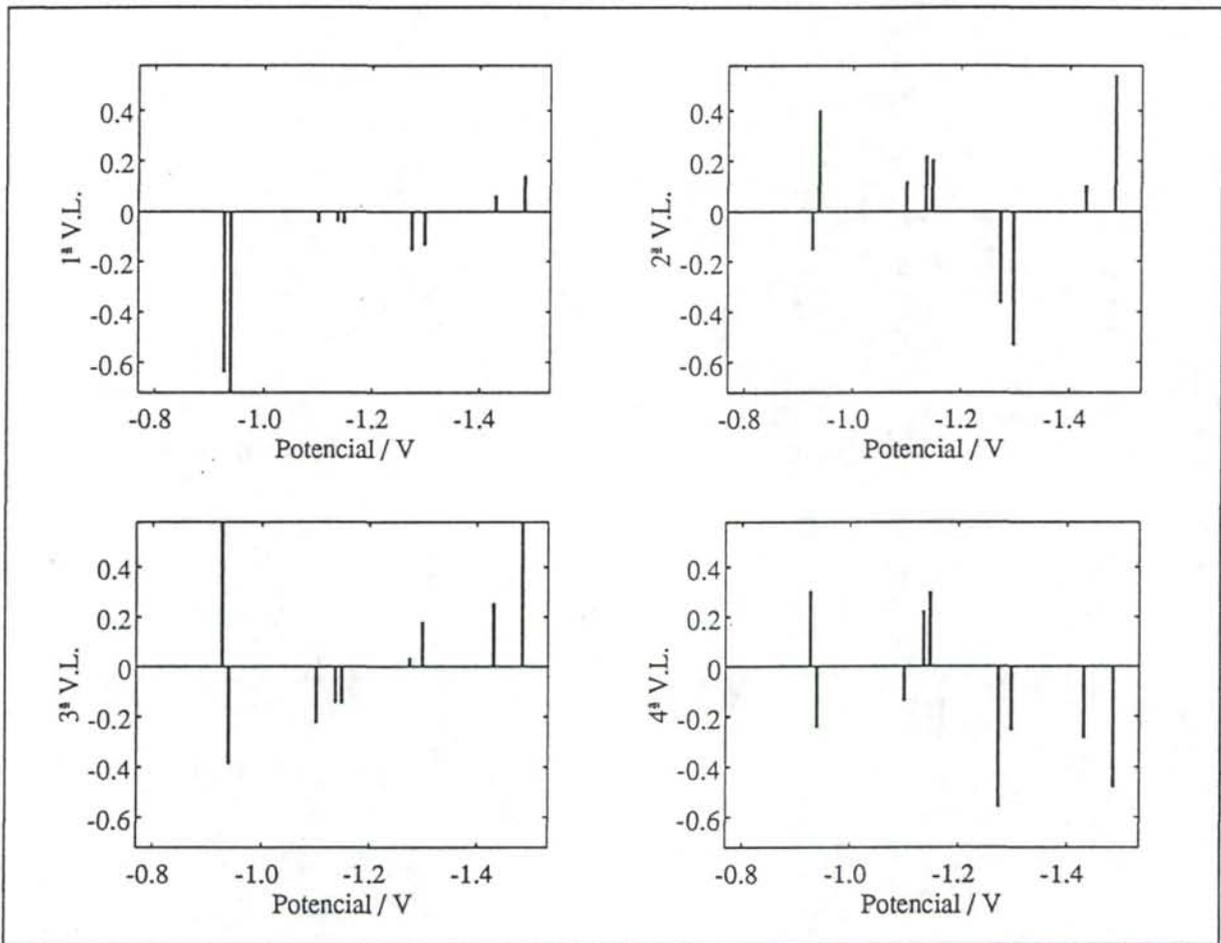


Figura IV.2.2b Loadings correspondientes al modelo PLS construido para pH = 2.8 con las variables seleccionadas por el algoritmo genético.

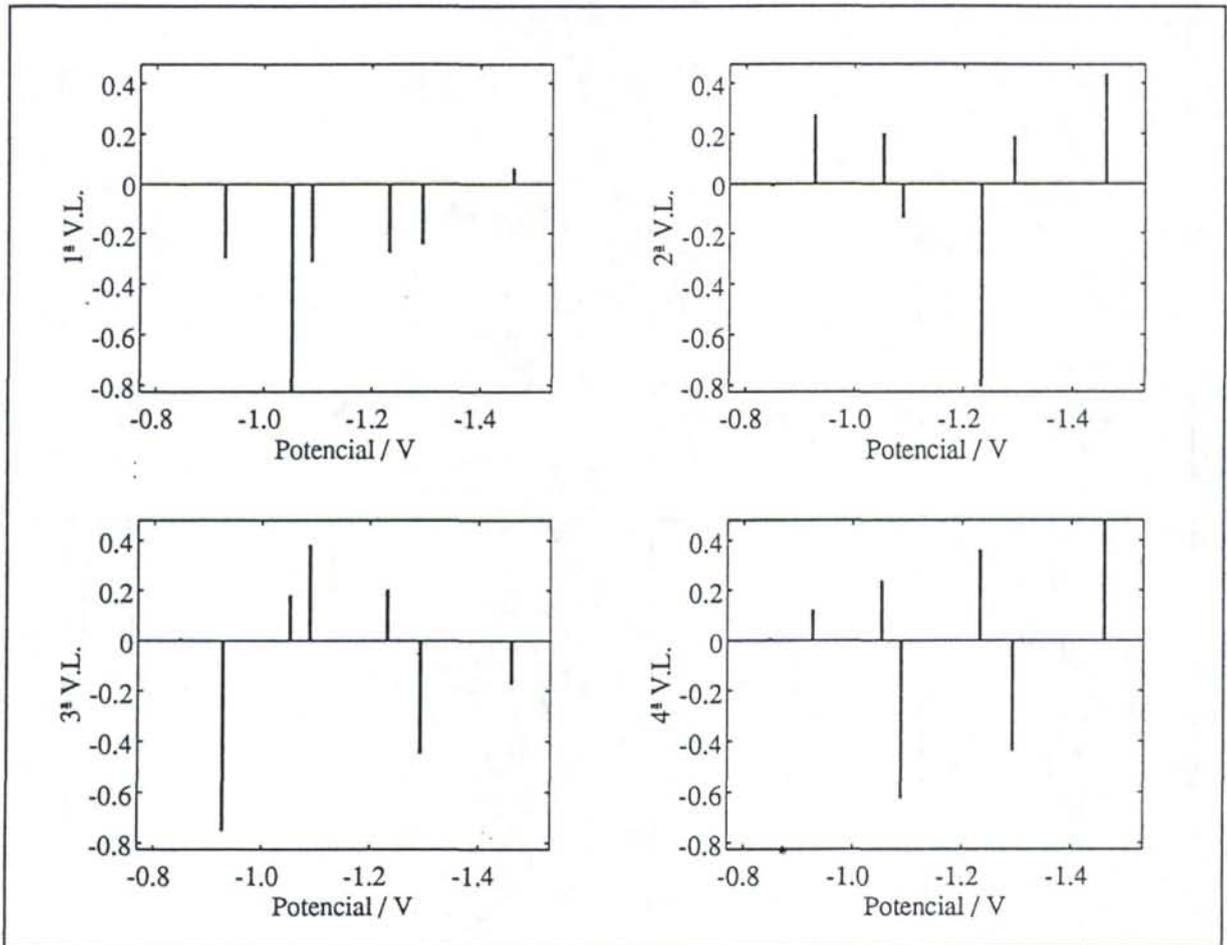


Figura IV.2.2c Loadings correspondientes al modelo PLS construido para $\text{pH} = 3.4$ con las variables seleccionadas por el algoritmo genético.

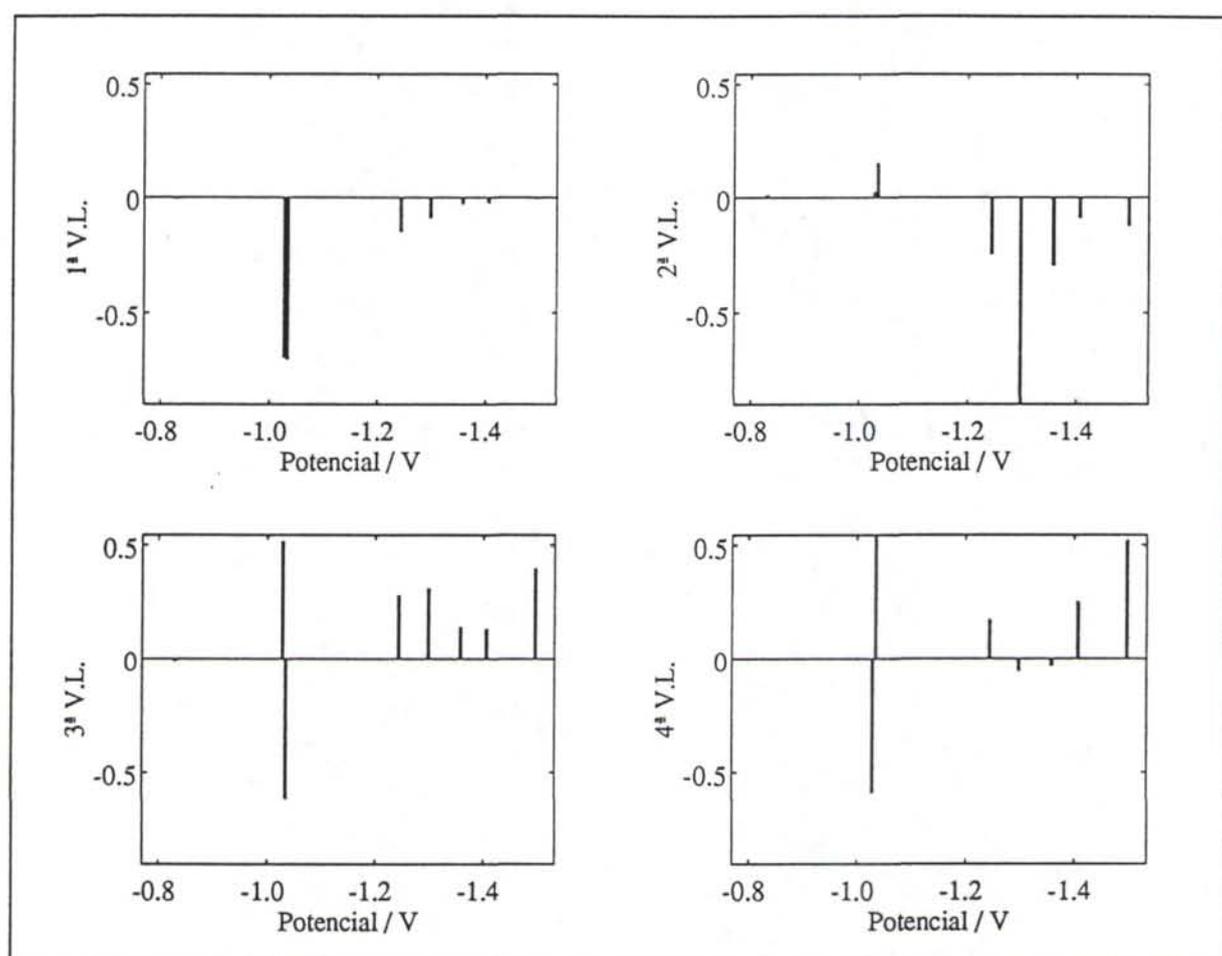


Figura IV.2.2d Loadings correspondientes al modelo PLS construido para $\text{pH} = 4.0$ con las variables seleccionadas por el algoritmo genético.

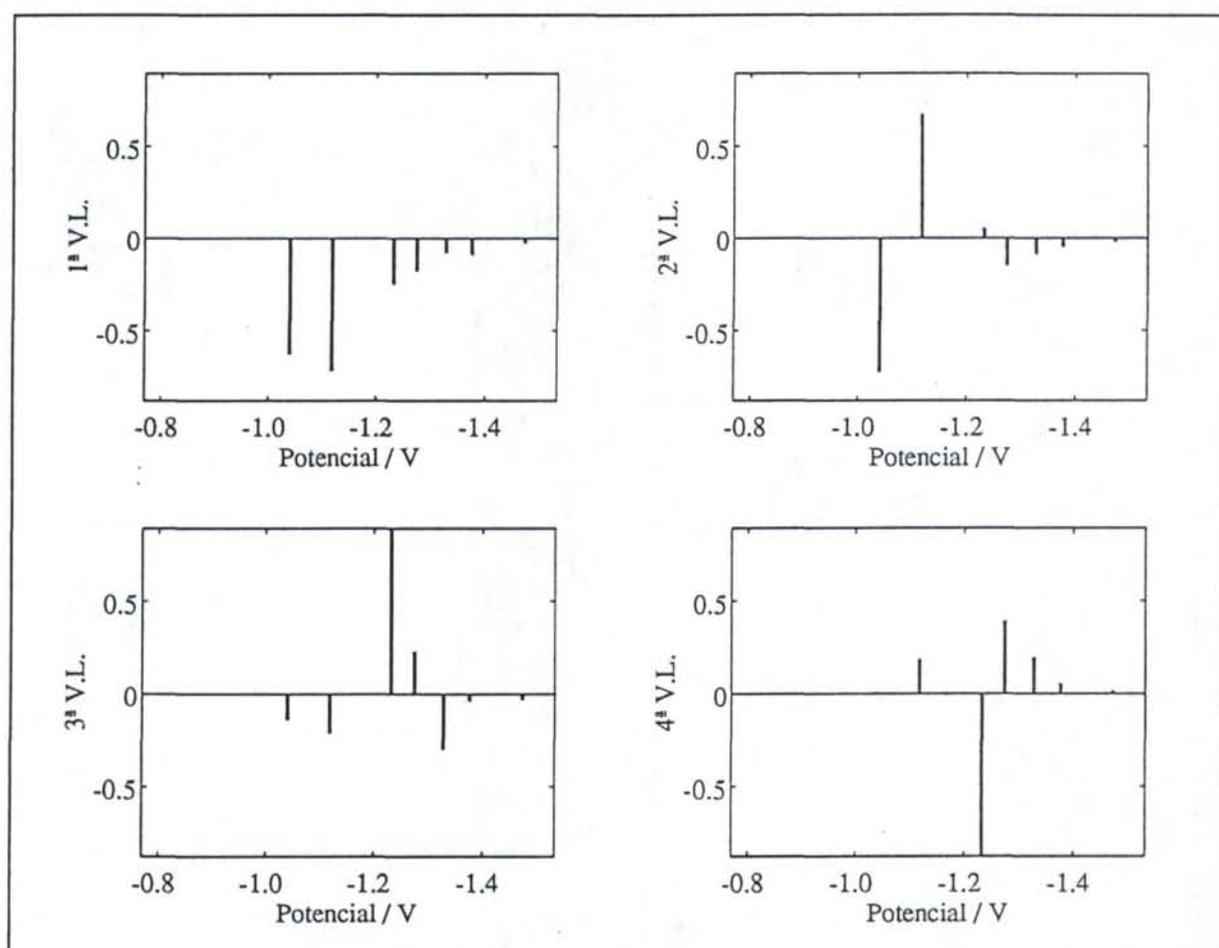


Figura IV.2.2e Loadings correspondientes al modelo PLS construido para pH = 4.6 con las variables seleccionadas por el algoritmo genético.

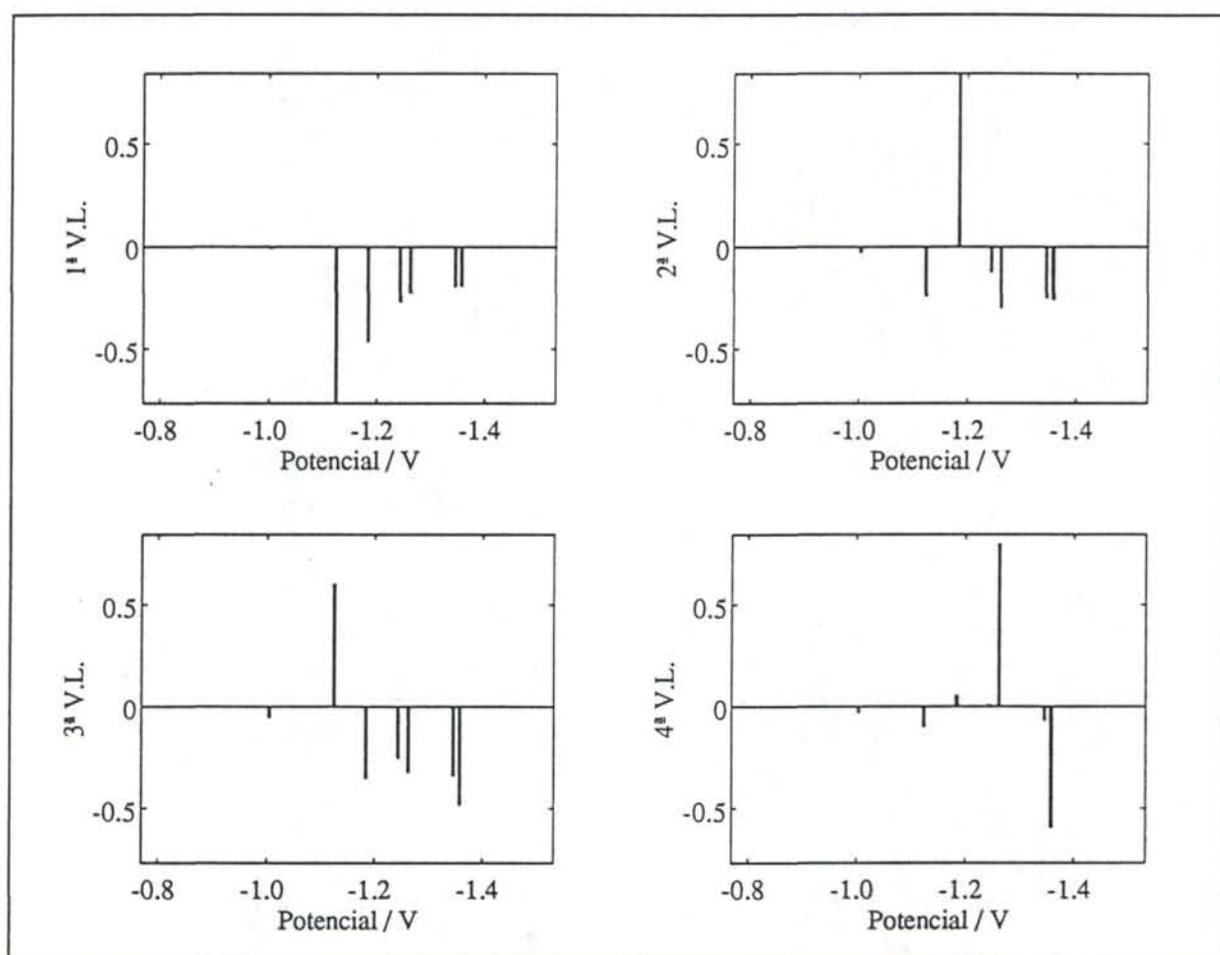


Figura IV.2.2f Loadings correspondientes al modelo PLS construido para pH = 5.2 con las variables seleccionadas por el algoritmo genético.

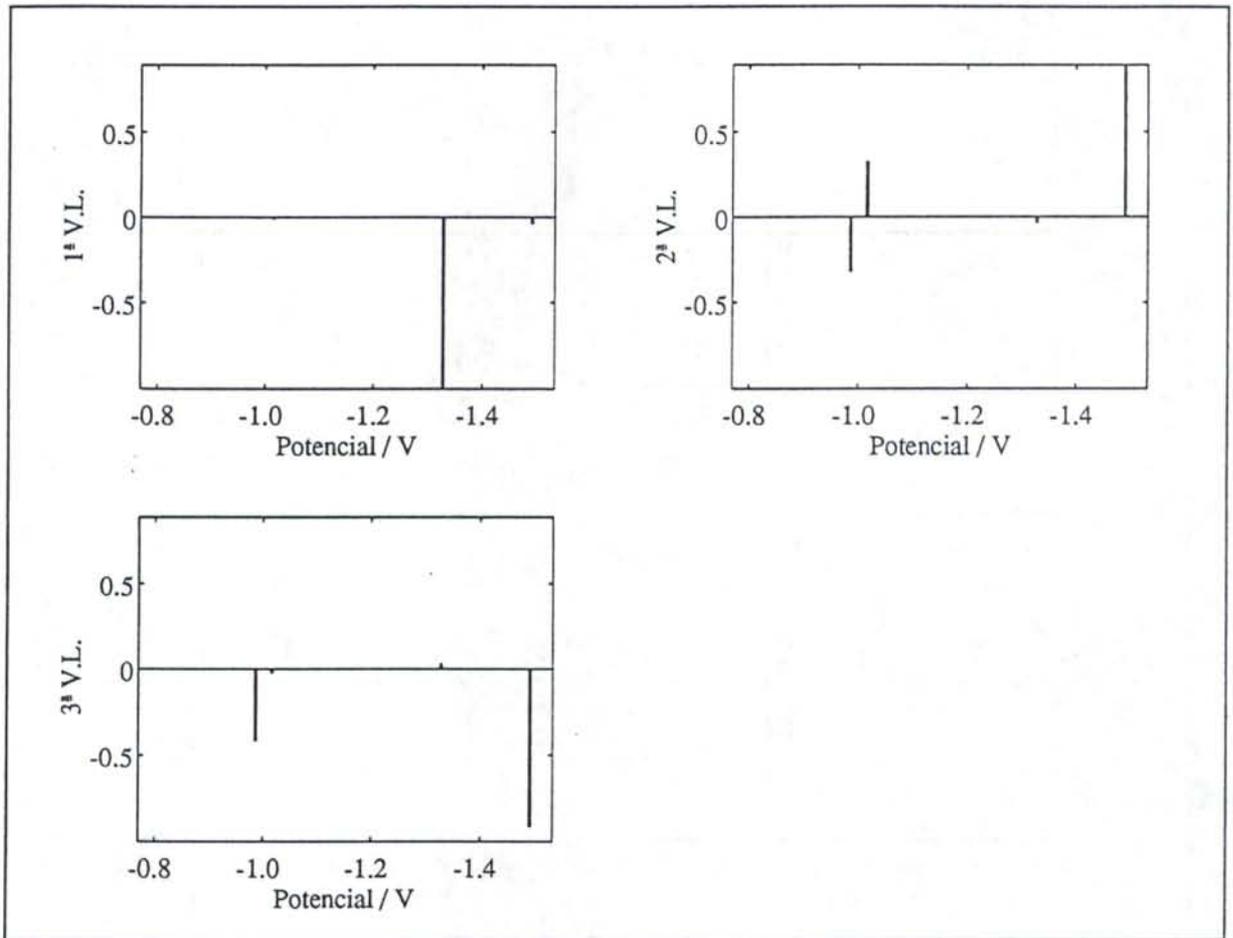


Figura IV.2.2g Loadings correspondientes al modelo PLS construido para pH = 5.8 con las variables seleccionadas por el algoritmo genético.

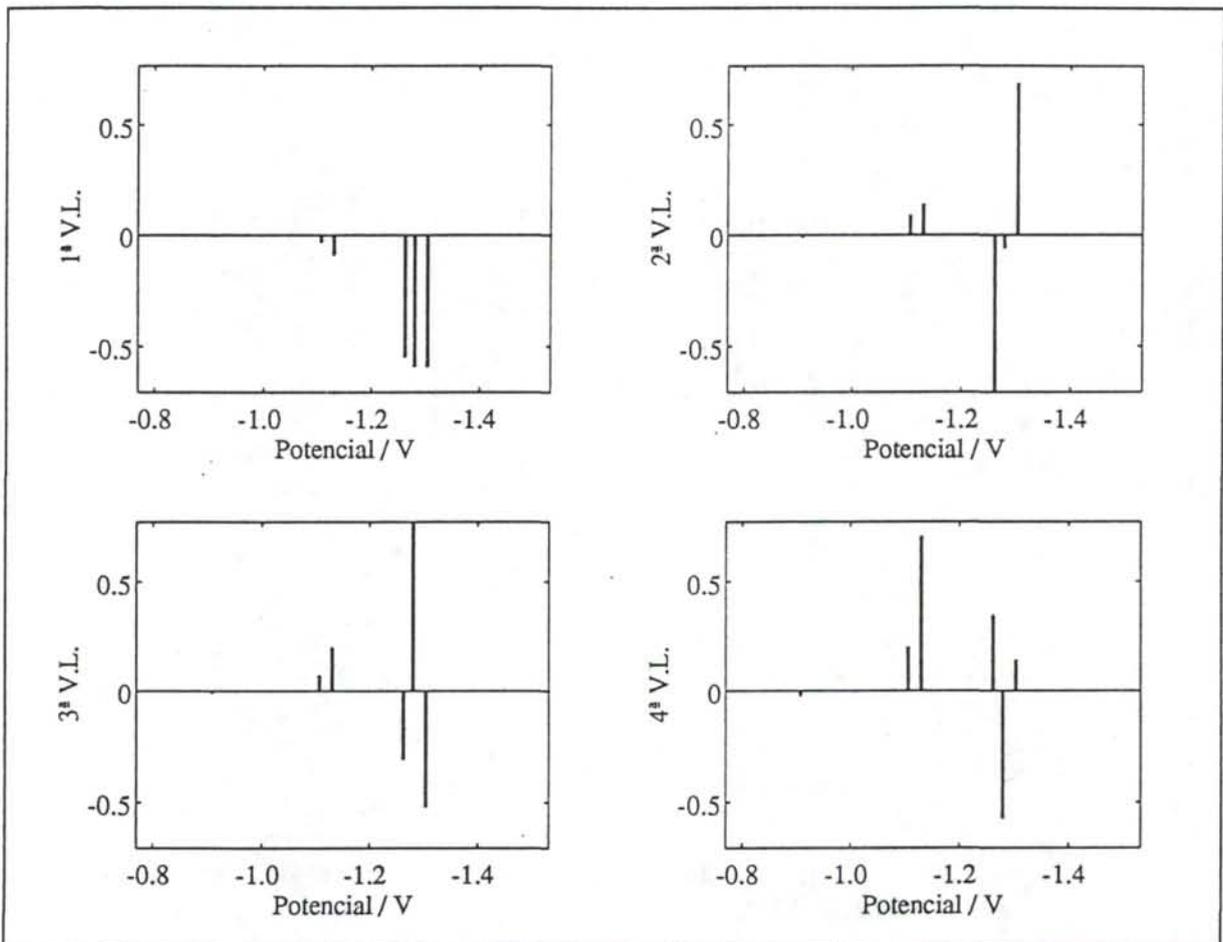


Figura IV.2.2h Loadings correspondientes al modelo PLS construido para pH = 6.4 con las variables seleccionadas por el algoritmo genético.

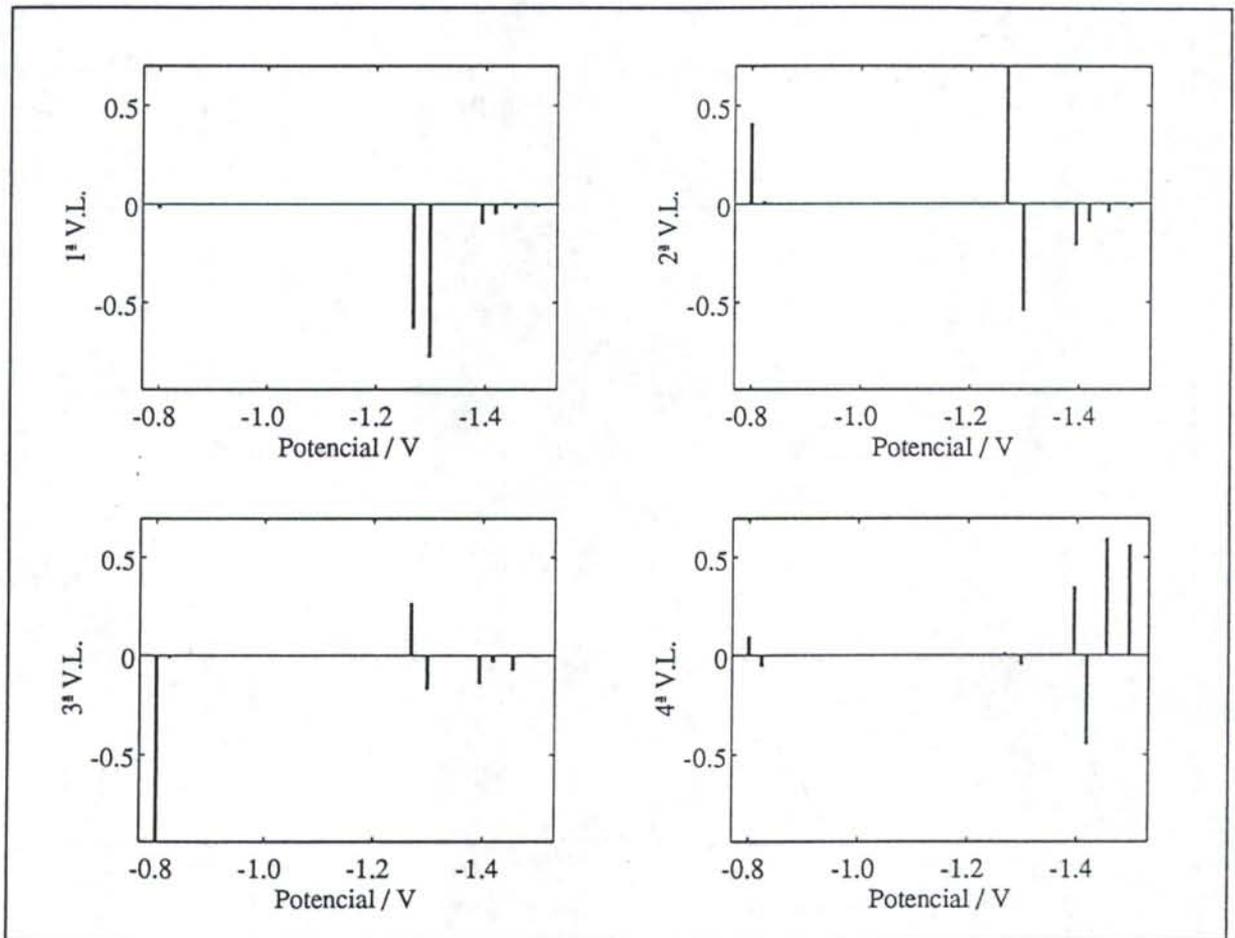


Figura IV.2.2i Loadings correspondientes al modelo PLS construido para pH = 7.0 con las variables seleccionadas por el algoritmo genético.

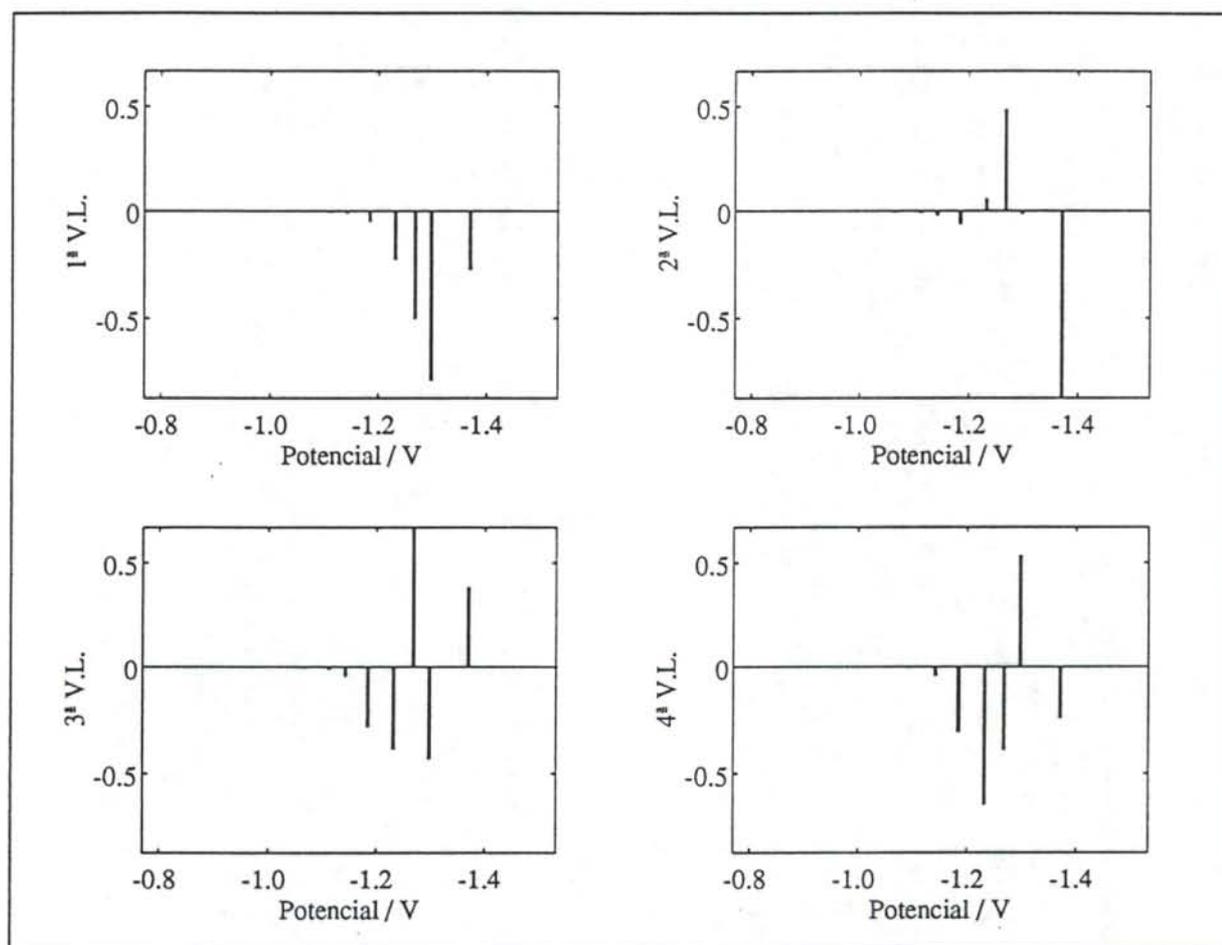


Figura IV.2.2j Loadings correspondientes al modelo PLS construido para pH = 7.6 con las variables seleccionadas por el algoritmo genético.

En cuanto a la hipótesis de igualdad de varianzas, análogamente al caso en el que no se seleccionaron variables, los test de Cochran y Bartlett [58] con ese mismo nivel de confianza indican que no es posible considerar que las varianzas sean iguales, tanto en el caso en el que se toman como nivel las adiciones de benzaldehído y los errores a los distintos pH como réplicas, como cuando se considera la situación contraria.

Tabla IV.3.3 Errores relativos en el cálculo de la concentración de benzaldehído con los modelos PLS contruidos con las variables seleccionadas por el algoritmo genético a diferentes valores experimentales de pH.

pH	Adición						
	1	2	3	4	5	6	7
2.2	6.12	-2.82	0.80	-0.08	-0.82	1.40	-0.66
2.8	1.95	-0.61	-0.97	1.35	0.65	-1.65	0.59
3.4	-0.23	1.61	-0.81	-0.22	-1.35	2.88	-1.43
4.0	-4.14	0.99	-1.30	-1.00	0.43	1.95	-1.13
4.6	1.20	-1.08	0.31	-0.07	0.07	0.22	-0.17
5.2	-0.09	2.27	-1.91	-0.57	0.49	0.61	-0.36
5.8	21.05	1.61	-3.71	2.70	2.92	-0.01	0.54
6.4	-5.02	-1.26	-0.69	0.12	0.80	0.01	-0.12
7.0	0.72	-0.10	-0.84	1.94	1.36	1.13	-0.78
7.6	-1.92	-0.02	-0.76	0.76	-0.21	0.69	-0.48

Por lo tanto, tampoco en este caso es posible realizar el ANOVA con la tabla de datos y se hace necesario un método no paramétrico, como el test de rangos de Friedman [83], para evaluar los efectos. En este caso, el test permite concluir que ninguno de los dos factores es significativo ($P_{\text{actual}} = 0.1366$), es decir, que no existe efecto. De nuevo, es posible considerar los valores de concentración de benzaldehído calculados a los distintos pH, para cada adición de analito, como si fueran réplicas.

Esto permite comparar los valores de concentración calculados por los modelos PLS contruidos con las variables seleccionadas por el algoritmo genético, con los valores de concentración reales mediante una regresión univariante. Este ajuste se ha llevado a cabo por mínimos cuadrados, y los resultados del mismo aparecen reflejados en la Tabla IV.2.4. La

hipótesis conjunta [57], que establece que el término independiente sea igual a 0 y la pendiente igual a 1, se acepta con un nivel de significación del 7.13%.

Tabla IV.3.4 Resultados de la regresión lineal llevada a cabo para comparar los valores de concentración reales con los calculados por los modelos PLS.

Variable	Coefficiente	S.E.	P _{actual}
Constante	0.0286	0.0773	0.71
Concentración	0.9986	0.0036	0.00
$\rho = 0.9996$	$S_{yx} = 0.2860$		

Este resultado, junto con la mayor proximidad de los valores de varianza en predicción a los correspondientes valores de varianza explicada, parece señalar que el procedimiento de selección de variables llevado a cabo en una etapa previa a la construcción de los modelos PLS aumenta la capacidad predictiva de éstos, así como su estabilidad.

IV.3.3 Comparación de resultados

Mediante técnicas de regresión

Habitualmente, cuando se desean comparar dos procedimientos o métodos de análisis, se puede llevar a cabo una regresión por mínimos cuadrados, tomando como variable independiente los resultados obtenidos con uno de los dos métodos, que se considera como método de referencia, y como variable dependiente los correspondientes al otro.

Una de las hipótesis básicas de ese modelo de regresión consiste en considerar que la variable independiente no es aleatoria, lo cual equivale a que, o bien está exenta de error, o su error aleatorio es despreciable frente al de la variable dependiente. Por esta razón, cuando tanto la variable dependiente como la variable independiente están sujetas a error [19, 232], la regresión por mínimos cuadrados no resulta adecuada para llevar a cabo el análisis. En este caso, debería utilizarse una regresión ortogonal [19].

Para comparar los resultados de los dos procedimientos utilizados en la determinación polarográfica de benzaldehído mediante modelos multivariantes PLS, con selección de

variables y sin ella, se ha utilizado una regresión ortogonal robusta. Esta regresión se rige por el criterio de hacer mínima la mediana de cuadrados de las diferencias perpendiculares a la propia recta.

El ajuste se ha llevado a cabo sobre los datos de la Figura IV.3.3, en la cual se representan los valores de las concentraciones de benzaldehído calculadas por los modelos PLS habiendo realizado una selección de variables con un algoritmo genético y sin ella. Se puede observar perfectamente como en ambos ejes, para un mismo valor de concentración real, ambos procedimientos presentan una cierta variabilidad en los resultados.

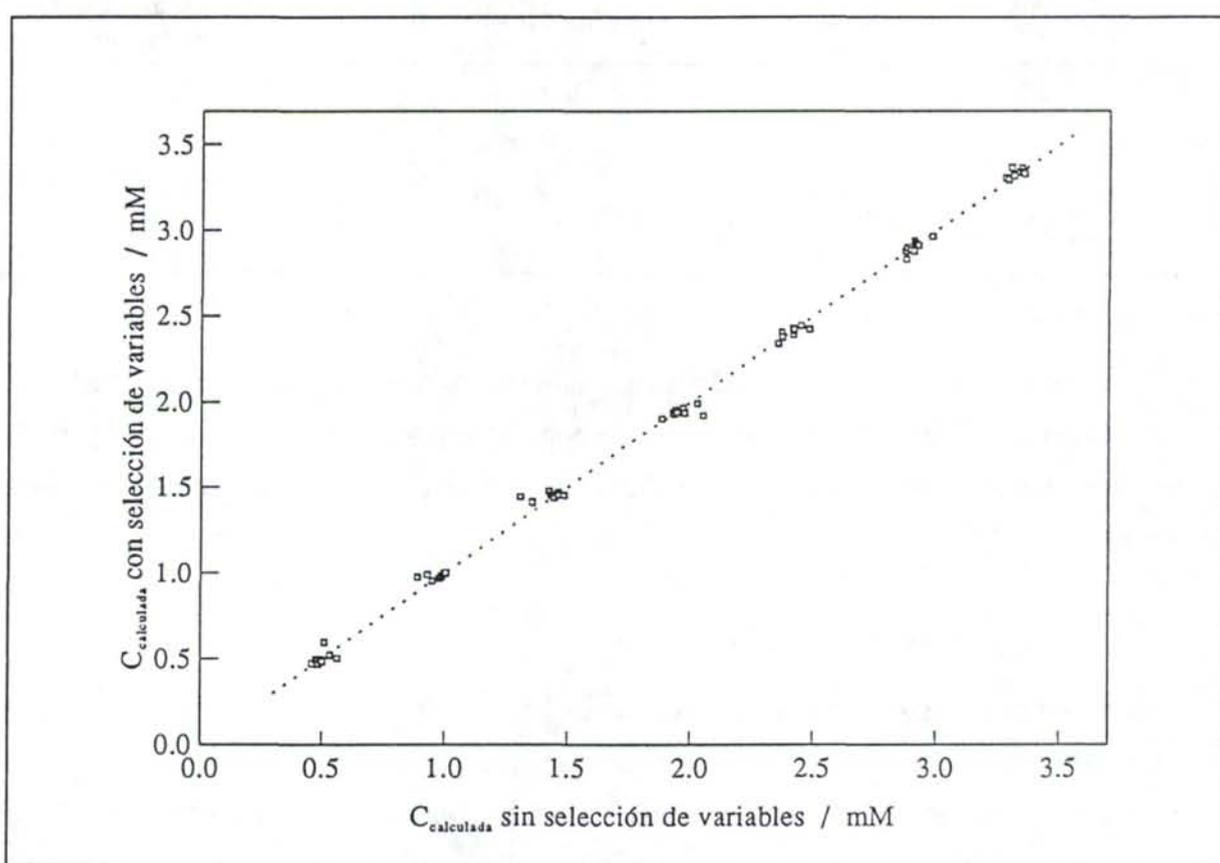


Figura IV.2.3 Concentración calculada por los modelos PLS contruidos con todas las variables o con aquellas seleccionadas por un algoritmo genético.

Se ha utilizado un programa desarrollado por Sarabia y otros [234] que ha permitido realizar el ajuste considerando diversos criterios, según se trate de una regresión por mínimos cuadrados (least squares, LS), de la regresión en mínima mediana de cuadrados (least median squares, LMS) [233], la regresión ortogonal (orthogonal least squares, OLS) [19] o la regresión ortogonal robusta (orthogonal least median squares, OLEM) [234].

El cálculo se ha repetido 14 veces obteniéndose valores muy estables. Los mejores resultados correspondientes a los cuatro métodos de regresión se muestran en la Tabla IV.2.5. En la misma se observa que los valores que más se aproximan a 1 y 0, para pendiente y término independiente respectivamente, son los correspondientes a la regresión ortogonal aunque todos se encuentran muy cercanos.

Tabla IV.3.5 Resultados de los diferentes métodos de regresión univariante utilizados para comparar los resultados con la selección de variables en los modelos PLS y sin ella.

Método	Pendiente	Término independiente	Desviación estándar de la predicción
LS	0.9967	0.0788	0.3580
LMS	0.9990	0.0018	0.3601
OLS	1.0058	-0.1228	0.3692
OLEM	0.9993	0.0003	0.3599

De estos resultados se concluye que, en este caso, no se puede afirmar que existan diferencias significativas entre los valores de concentración calculados por los modelos PLS sin selección de variables o llevando a cabo una selección de las mismas utilizando un algoritmo genético.

Mediante test no paramétricos

Otra metodología que resulta adecuada para analizar el efecto que tiene sobre los resultados del análisis multivariante el realizar o no la selección de variables con el algoritmo genético consiste en aplicar test de hipótesis sobre muestras emparejadas. En este caso, se han calculado las diferencias entre el valor absoluto del error relativo de las concentraciones calculadas cuando se llevó a cabo la selección de variables y las calculadas cuando se construyeron los modelos PLS con todas las variables. Dado que estas diferencias no cumplían la hipótesis de normalidad, se han realizado los test no paramétricos de los signos y de los rangos con signo de Wilcoxon.

Para ambos test, la hipótesis nula es H_0 : la mediana es 0 (no existe efecto por la selección de variables), mientras que la hipótesis alternativa es H_1 : la mediana es distinta de 0 (existe efecto debido a la selección de variables). En ambos casos se rechazó la hipótesis nula a niveles de significación del $4.2 \cdot 10^{-3}$, ya que este valor es menor que 0.05. Además, el

valor de la mediana es -0.255, lo cual indica que la selección de variables reduce el error relativo en -0.255% a ese nivel de significación.

IV.4 SELECCIÓN DE VARIABLES EN DETERMINACIONES MULTIANALITO POR VOLTAMPEROMETRÍA

Un estudio paralelo al anterior se ha llevado a cabo con los distintos ejemplos que aparecen en capítulos anteriores de esta memoria. En concreto, se ha realizado una selección de variables utilizando el algoritmo genético indicado previamente en los calibrados A, C, E, H y J, los cuales fueron utilizados a su vez en las transferencias de calibrado del *Capítulo III* como calibrados patrón.

Las variables seleccionadas por el algoritmo genético en los citados ejemplos para cada uno de los analitos implicados en cada calibrado, tomando los mismos parámetros que en el *Epígrafe IV.3.1*, se señalan en las Figuras IV.4.4 a IV.4.8 mediante líneas verticales.

IV.4.1 Construcción y evaluación de modelos PLS

Construcción de los modelos PLS

Con los potenciales señalados en las Figuras IV.4.4 a IV.4.8 se han construido los correspondientes modelos de regresión PLS para cada uno de los diferentes ejemplos. En dichos modelos tanto la varianza explicada como la varianza en predicción son del mismo orden que los obtenidos para los modelos PLS construidos con todas las variables. Por ello, en las Tablas IV.4.6 a IV.4.10 se muestran únicamente los errores correspondientes a las muestras de los diferentes conjuntos de aprendizaje.

Si se comparan los errores recogidos en dichas tablas con los correspondientes a los modelos PLS construidos con todas las variables predictoras se observa que, en algunos casos, éstos son ligeramente menores para algunos de los analitos determinados mientras que en otros ocurre lo contrario.

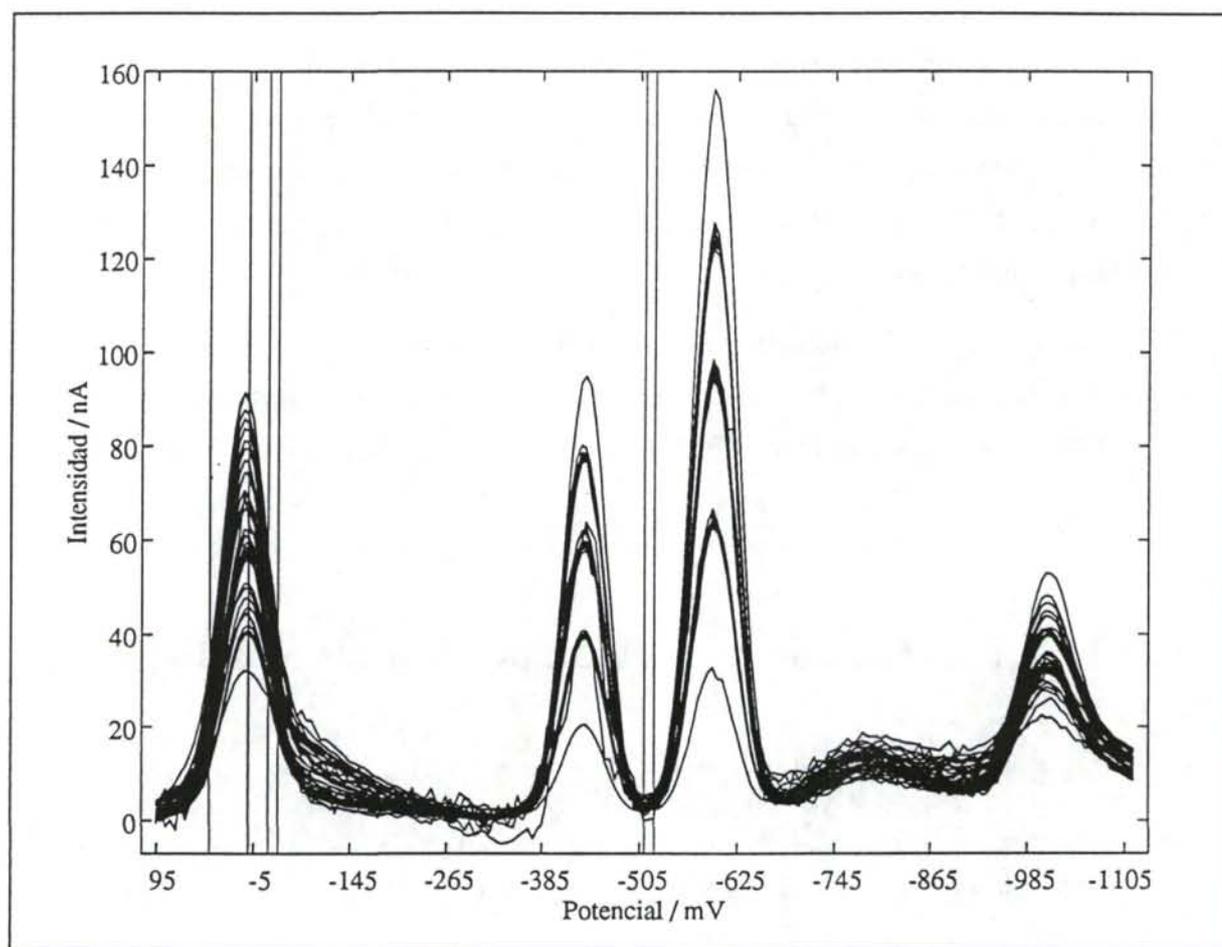


Figura 2a Variables seleccionadas por el algoritmo genético para el Cu (calibrado A). Los picos polarográficos se corresponden de izqda. a dcha. con Cu, Pb, Cd y Zn.

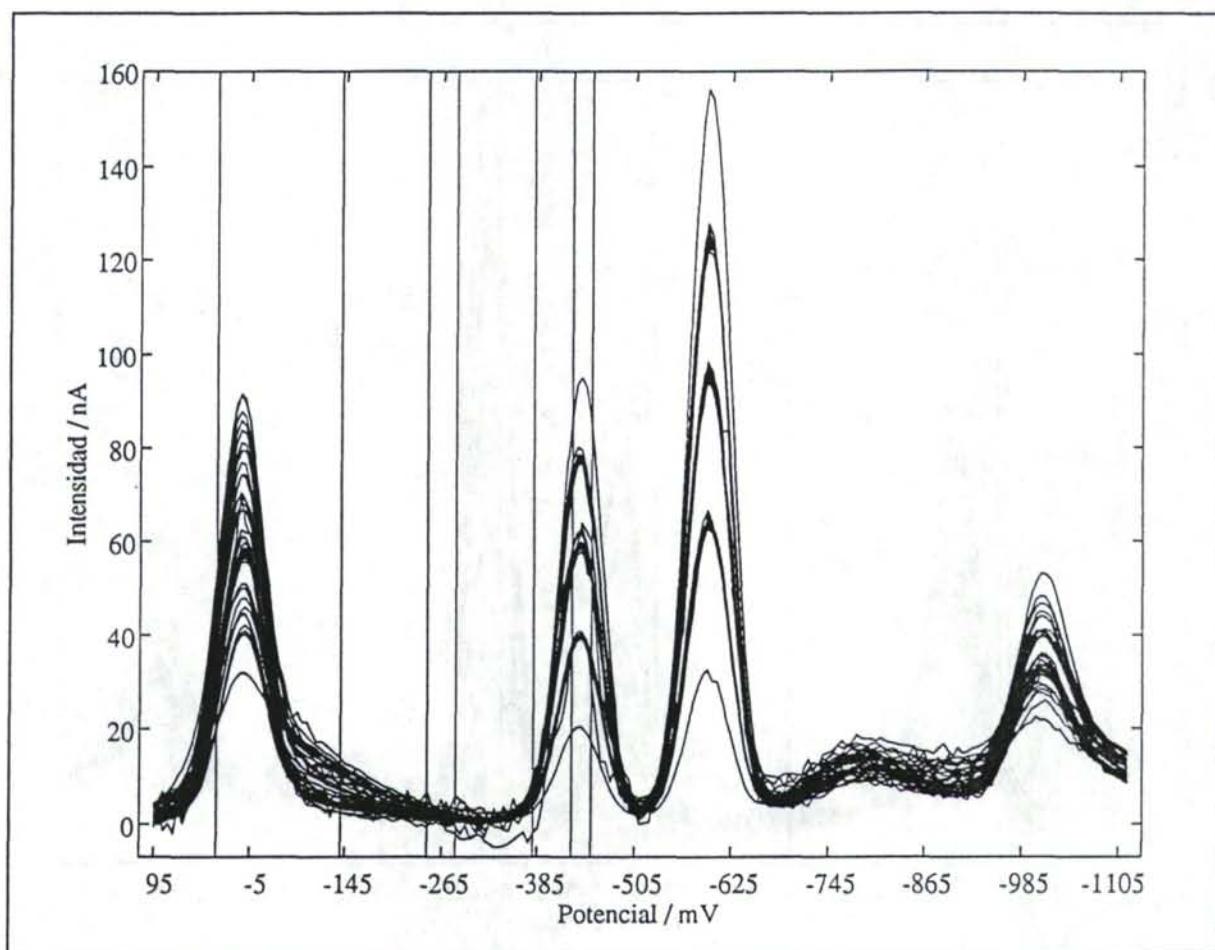


Figura 2b Variables seleccionadas por el algoritmo genético para el Pb (calibrado A). Los picos polarográficos se corresponden de izqda. a dcha. con Cu, Pb, Cd y Zn.

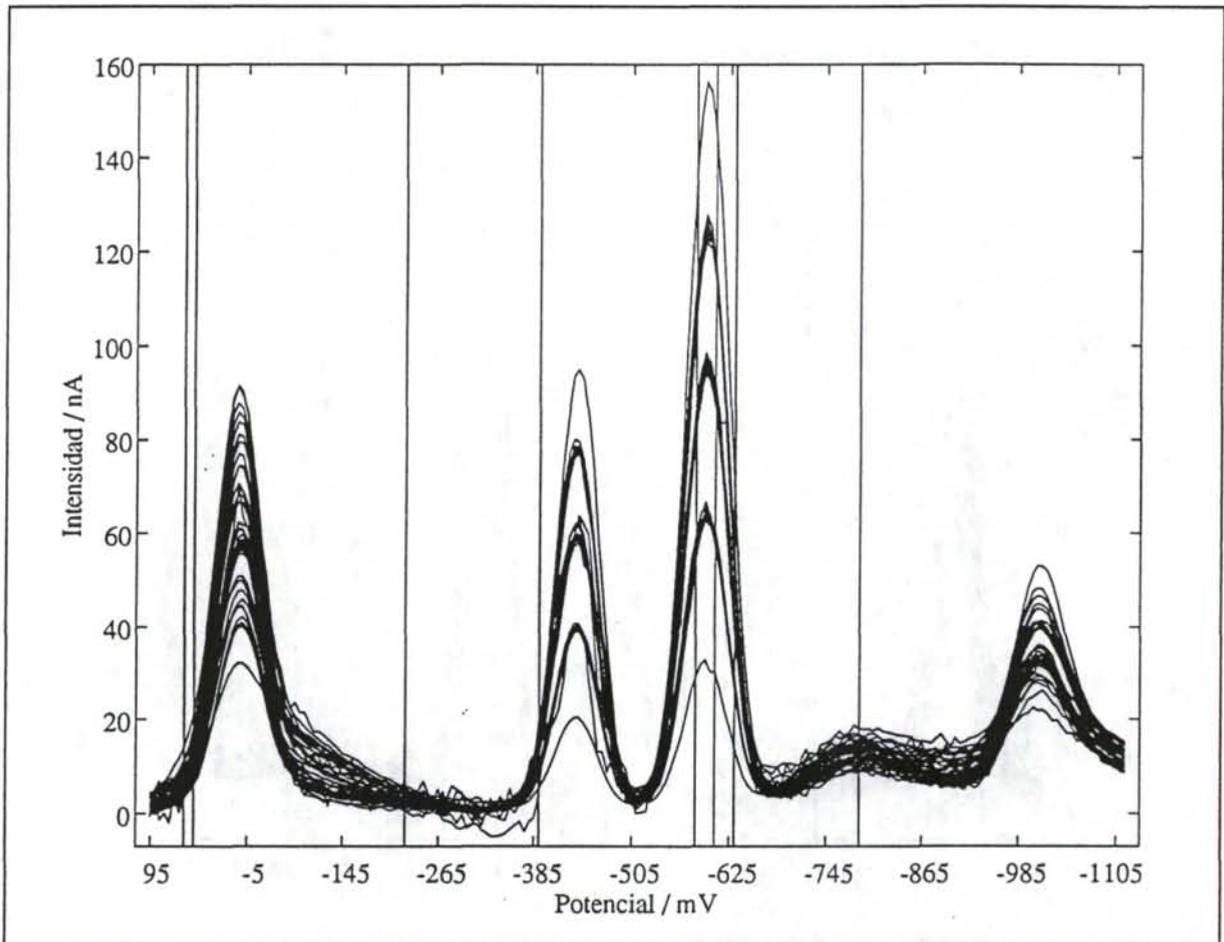


Figura 2c Variables seleccionadas por el algoritmo genético para el Cd (calibrado A). Los picos polarográficos se corresponden de izqda. a dcha. con Cu, Pb, Cd y Zn.

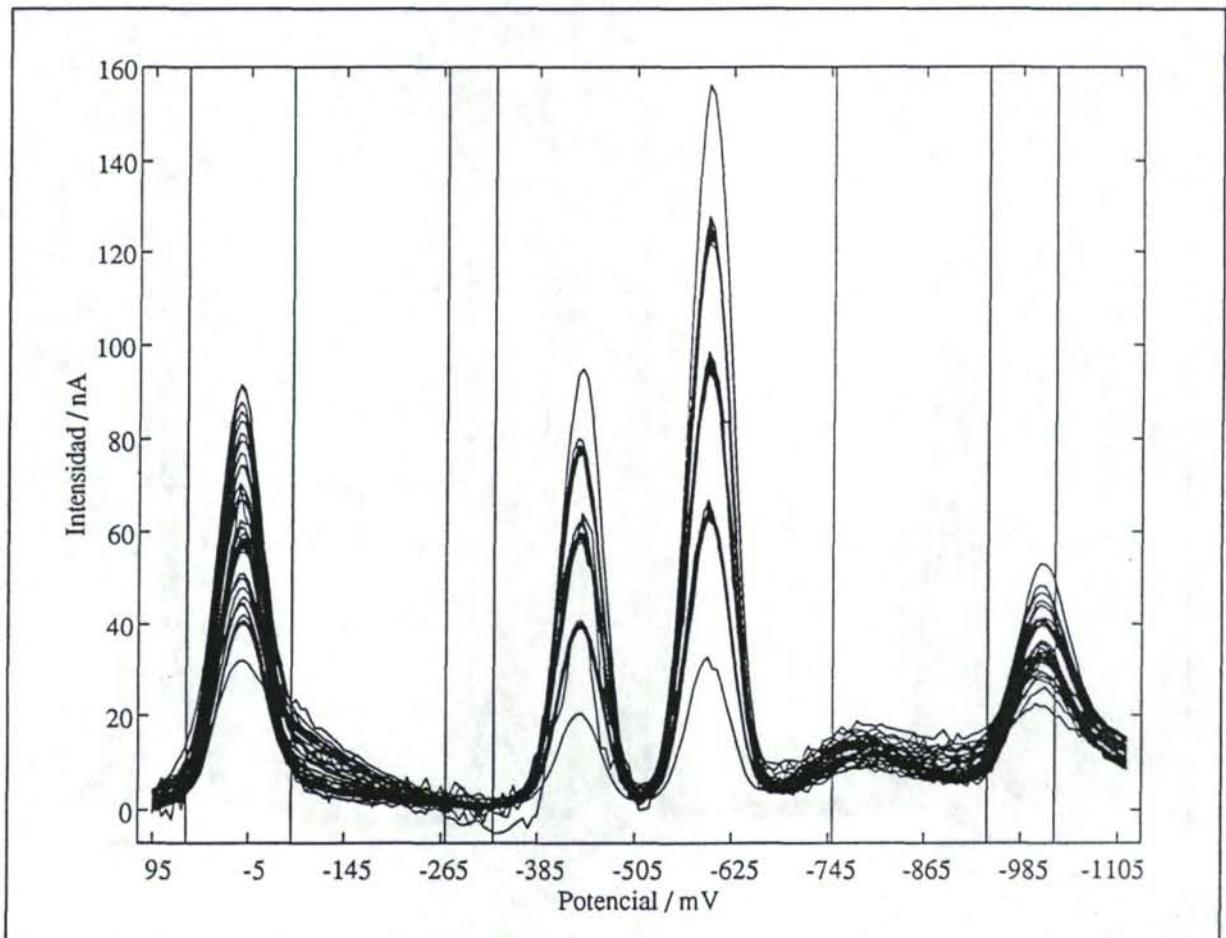


Figura 2d Variables seleccionadas por el algoritmo genético para el Zn (calibrado A). Los picos polarográficos se corresponden de izqda. a dcha. con Cu, Pb, Cd y Zn.

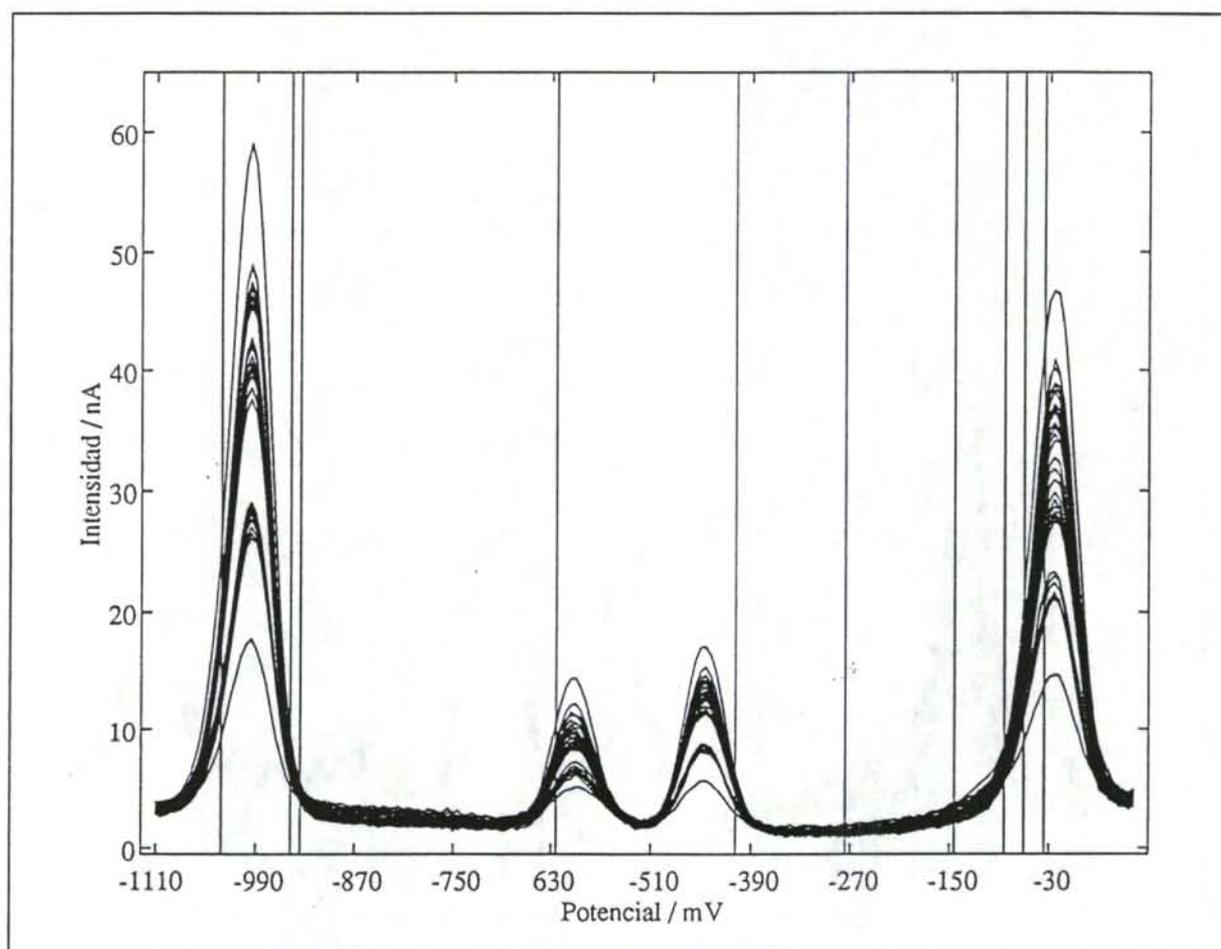


Figura 3a Variables seleccionadas por el algoritmo genético para el Cu (calibrado C). Los picos voltamperométricos se corresponden de izqda. a dcha. con Zn, Cd, Pb y Cu.

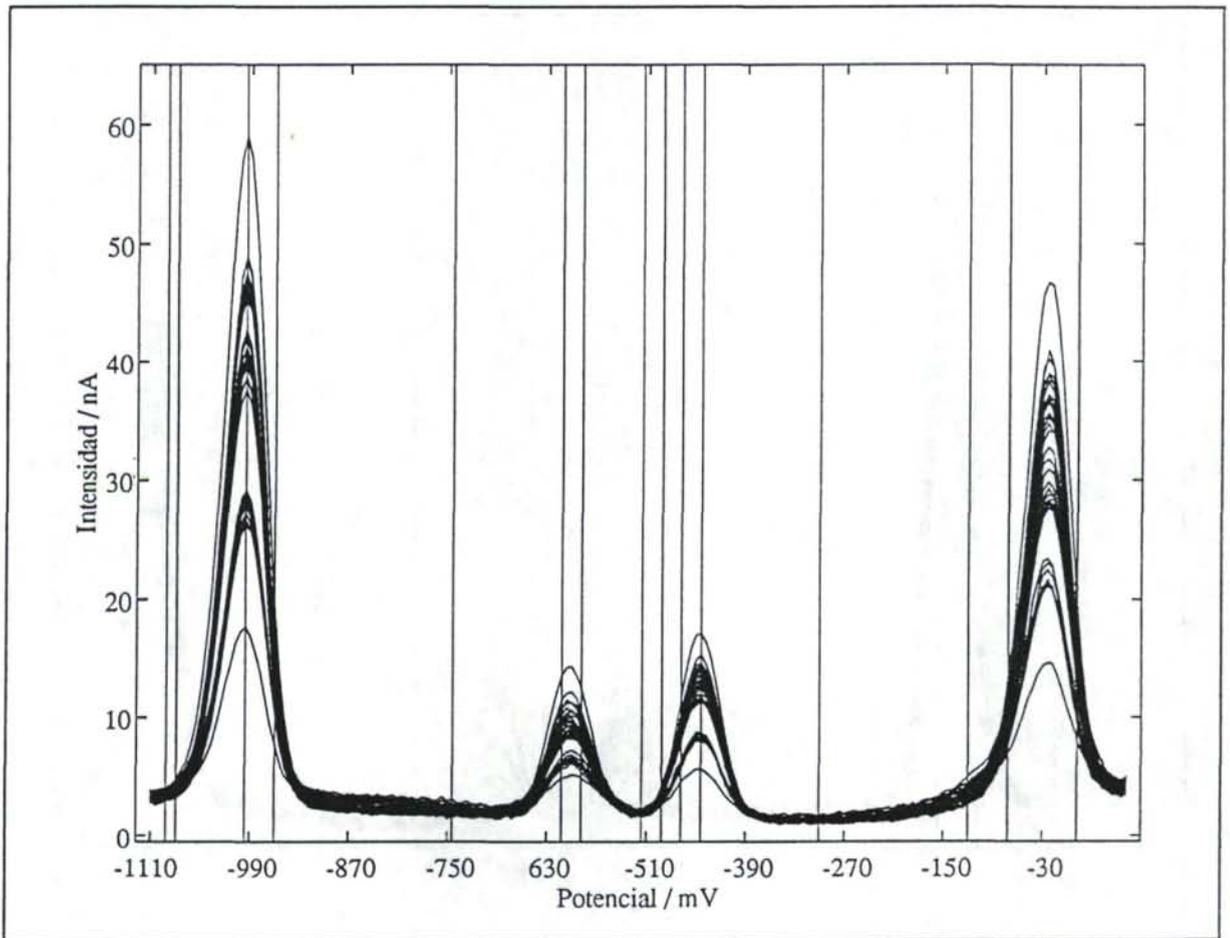


Figura 3b Variables seleccionadas por el algoritmo genético para el Pb (calibrado C). Los picos voltamperométricos se corresponden de izqda. a dcha. con Zn, Cd, Pb y Cu.

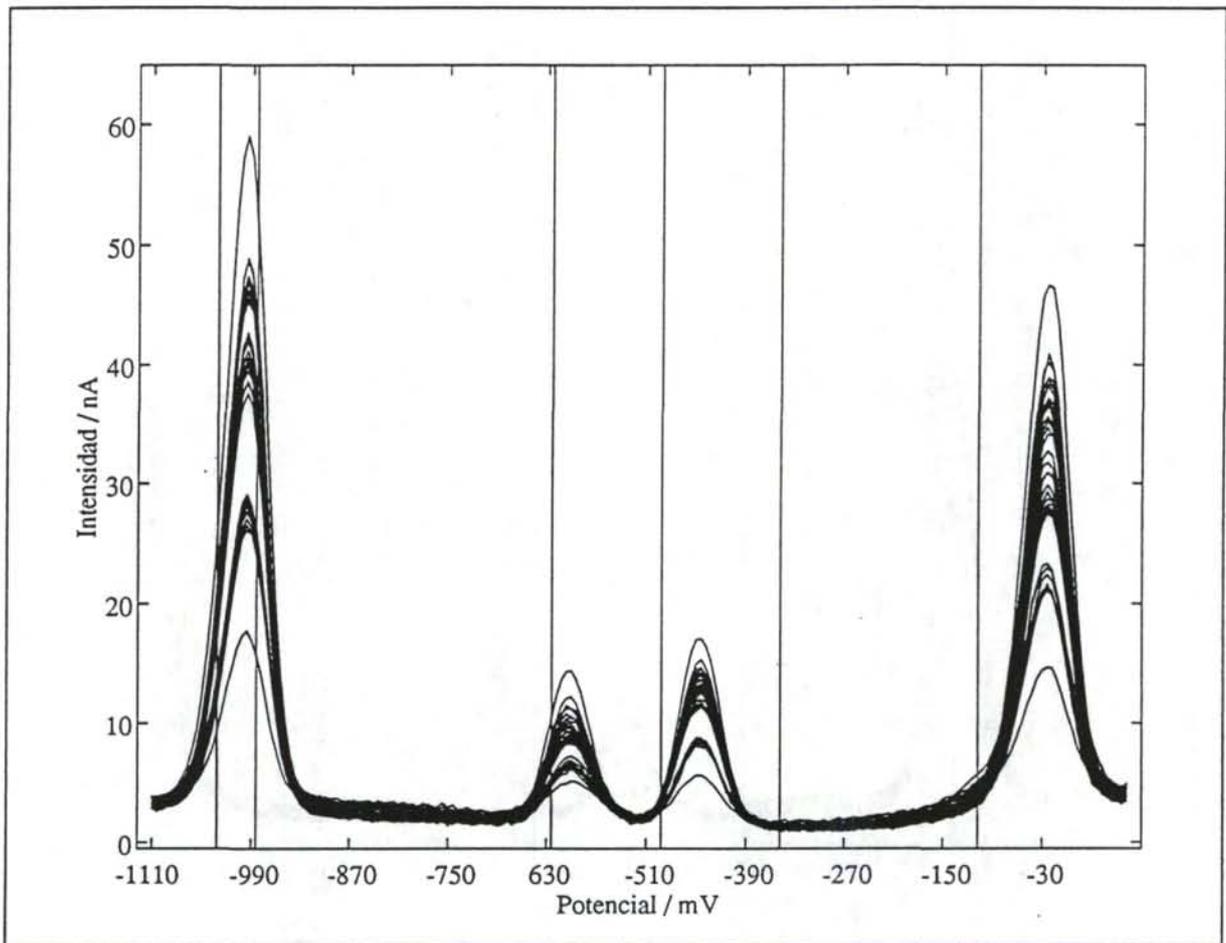
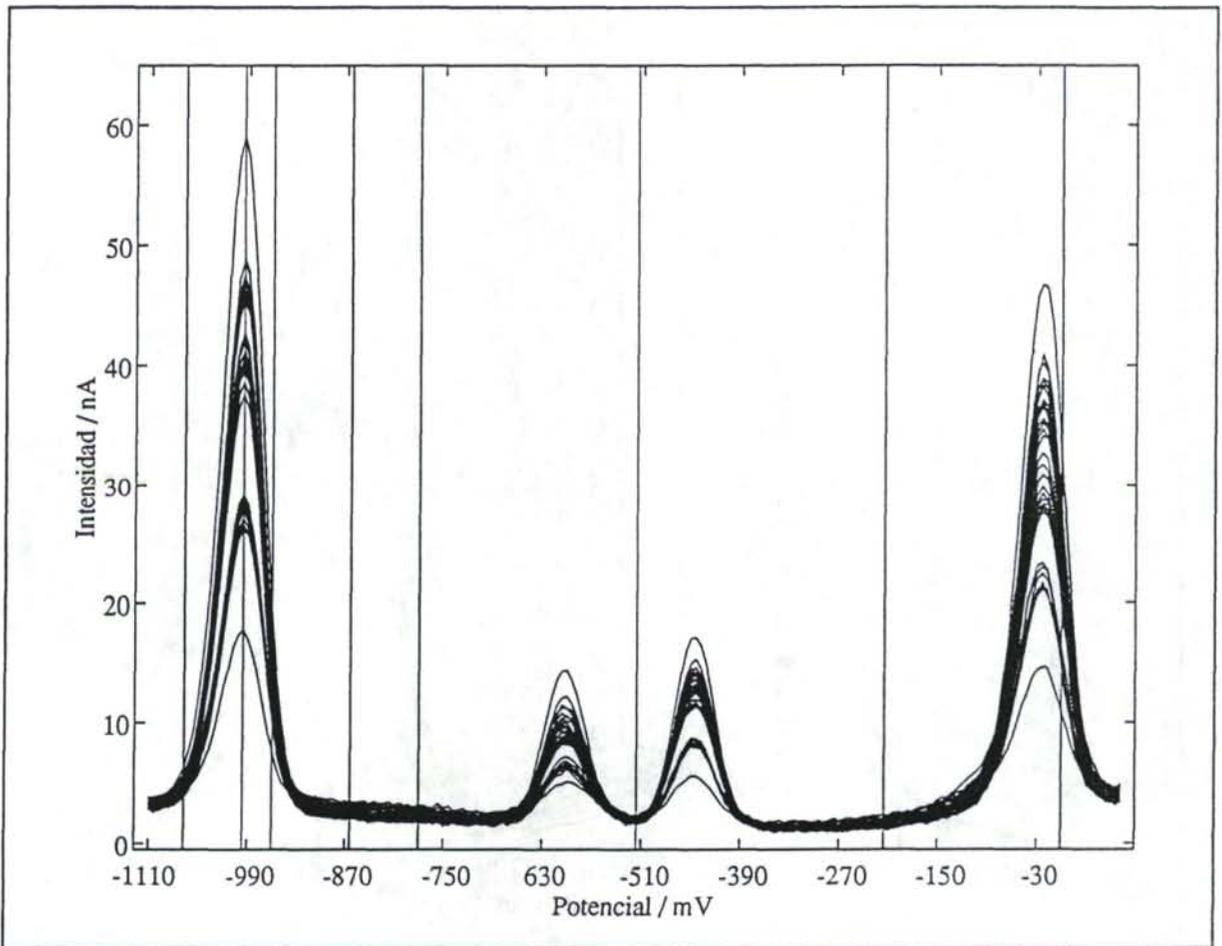
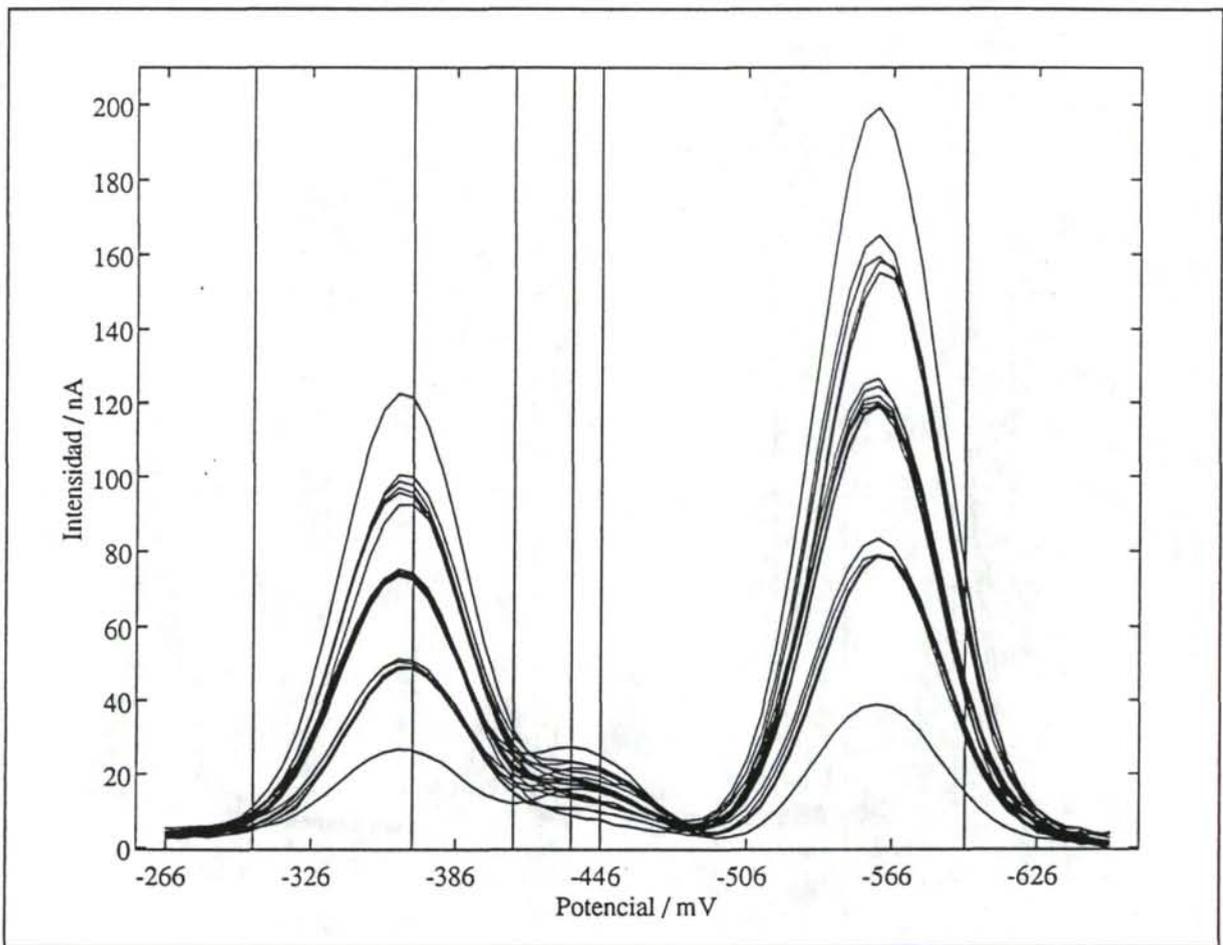


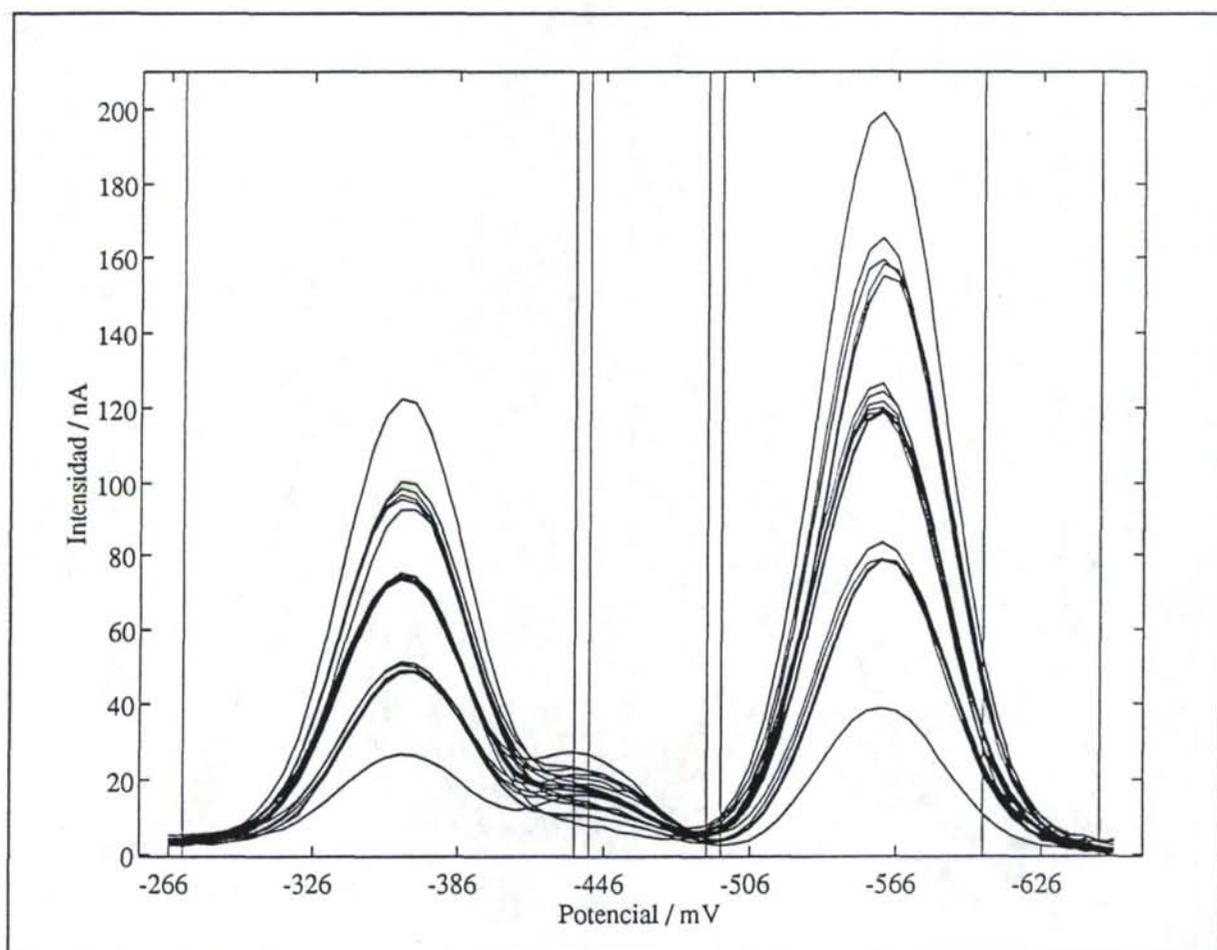
Figura 3c Variables seleccionadas por el algoritmo genético para el Cd (calibrado C). Los picos voltamperométricos se corresponden de izqda. a dcha. con Zn, Cd, Pb y Cu.

**Figura 3d**

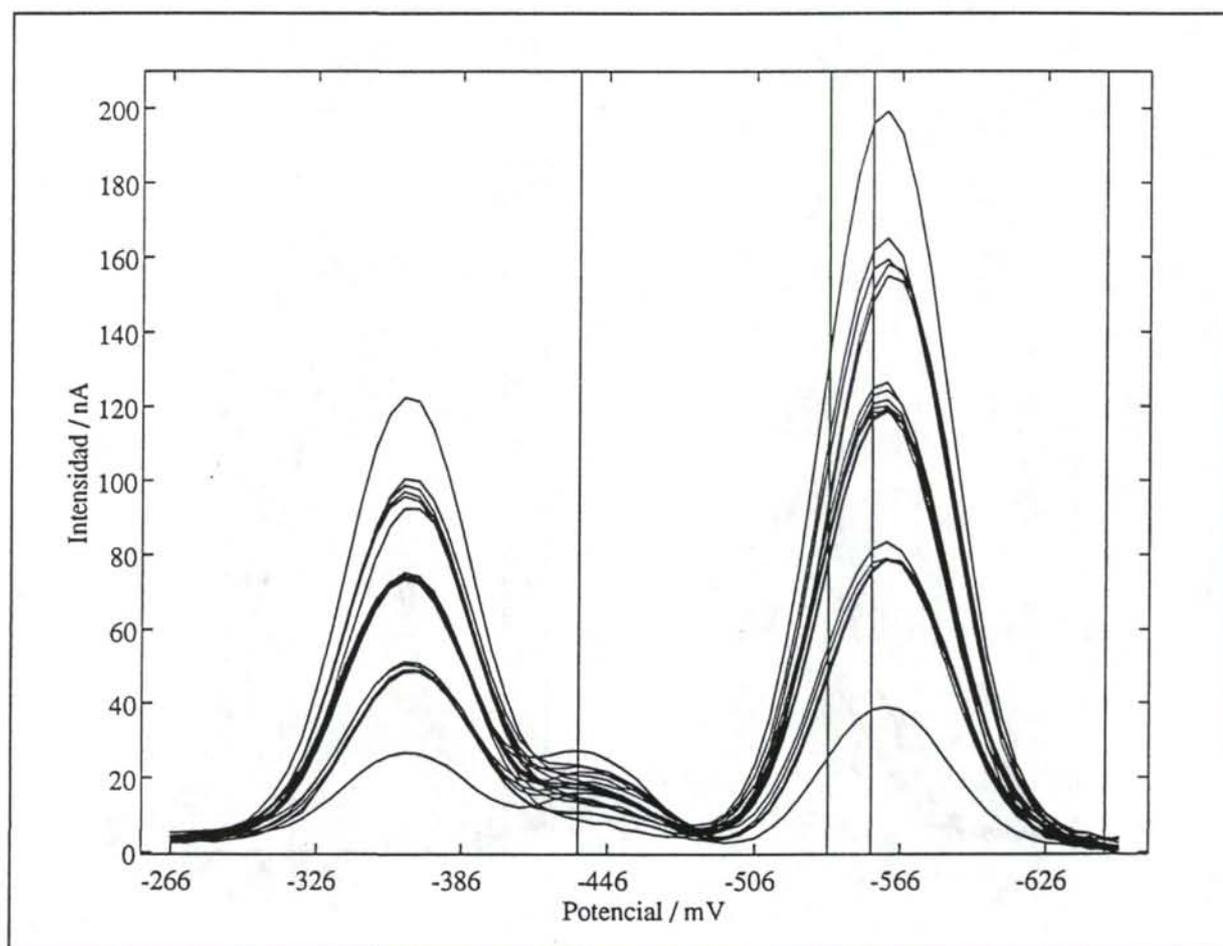
Variables seleccionadas por el algoritmo genético para el Zn (calibrado C). Los picos voltamperométricos se corresponden de izqda. a dcha. con Zn, Cd, Pb y Cu.

**Figura 4a**

Variables seleccionadas por el algoritmo genético para el Pb (calibrado E). Los picos polarográficos se corresponden de izqda. a dcha. con Pb, Sn y Cd.

**Figura 4b**

VARIABLES SELECCIONADAS POR EL ALGORITMO GENÉTICO PARA EL Sn (CALIBRADO E). LOS PICOS POLAROGRAFICOS SE CORRESPONDEN DE IZQDA. A DCHA. CON Pb, Sn Y Cd.

**Figura 4c**

Variables seleccionadas por el algoritmo genético para el Cd (calibrado E). Los picos polarográficos se corresponden de izqda. a dcha. con Pb, Sn y Cd.

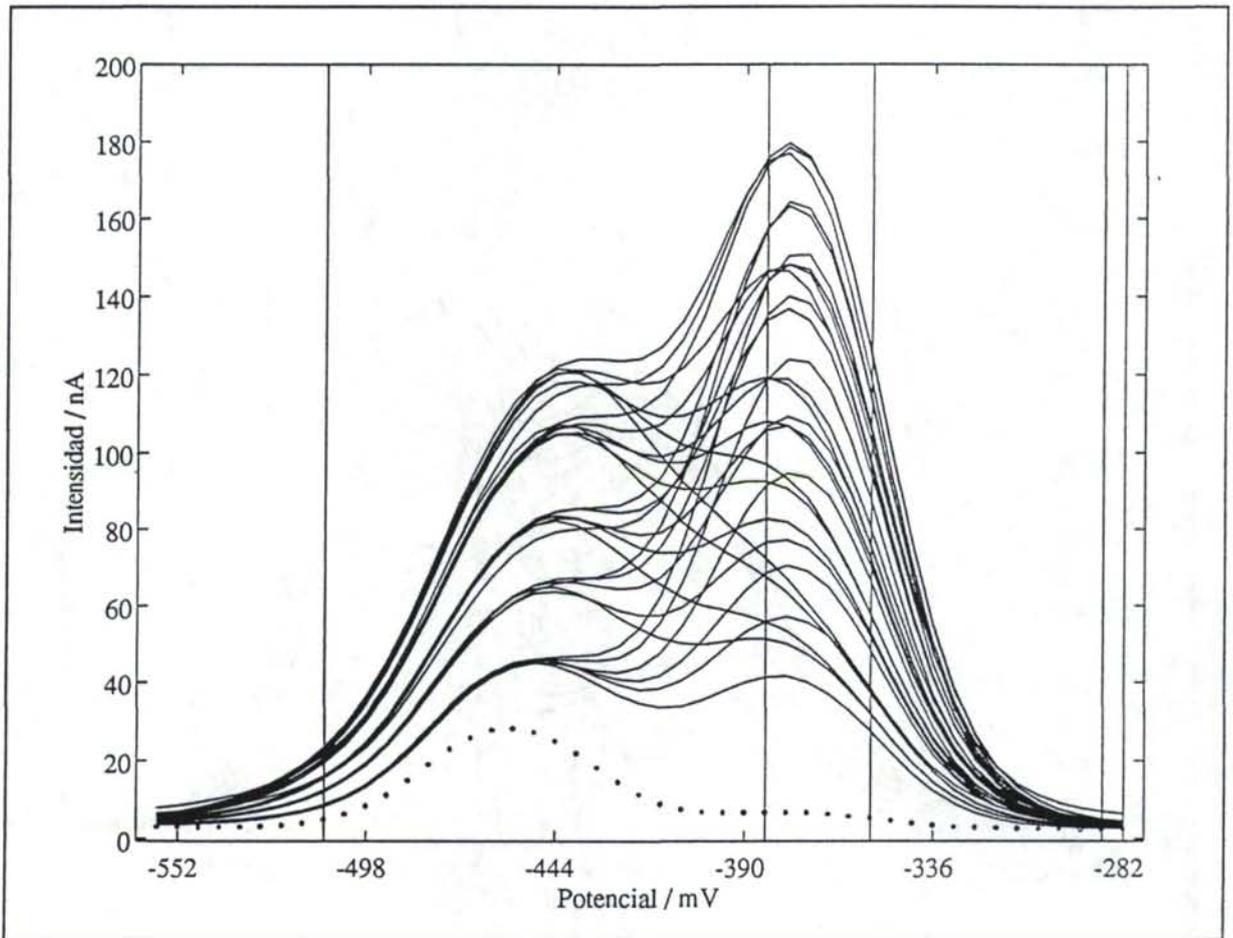


Figura 5a Variables seleccionadas por el algoritmo genético para el Tl (calibrado H). Los picos voltamperométricos se corresponden de izqda. a dcha. con Tl y Pb.

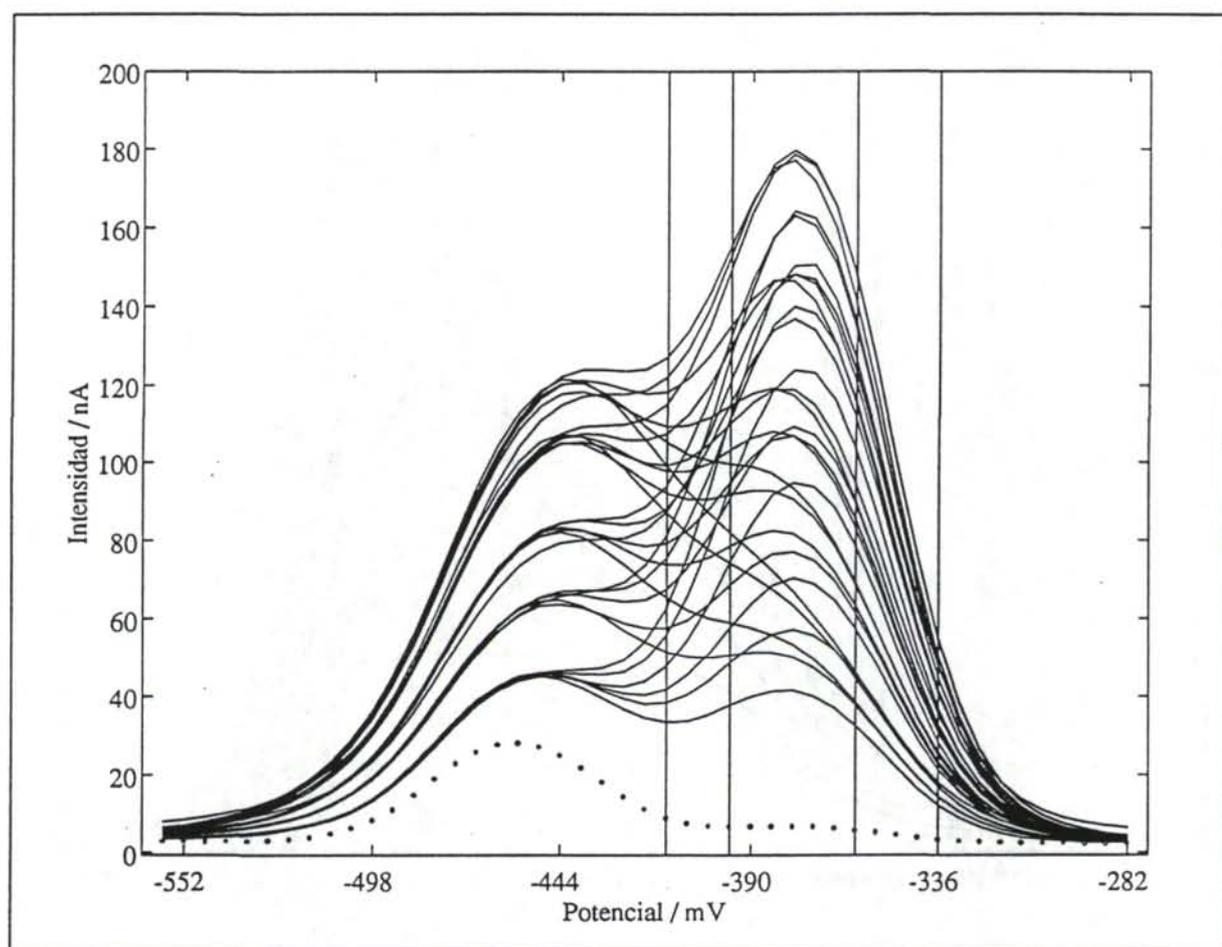


Figura 5b Variables seleccionadas por el algoritmo genético para el Pb (calibrado H). Los picos voltamperométricos se corresponden de izqda. a dcha. con Tl y Pb.

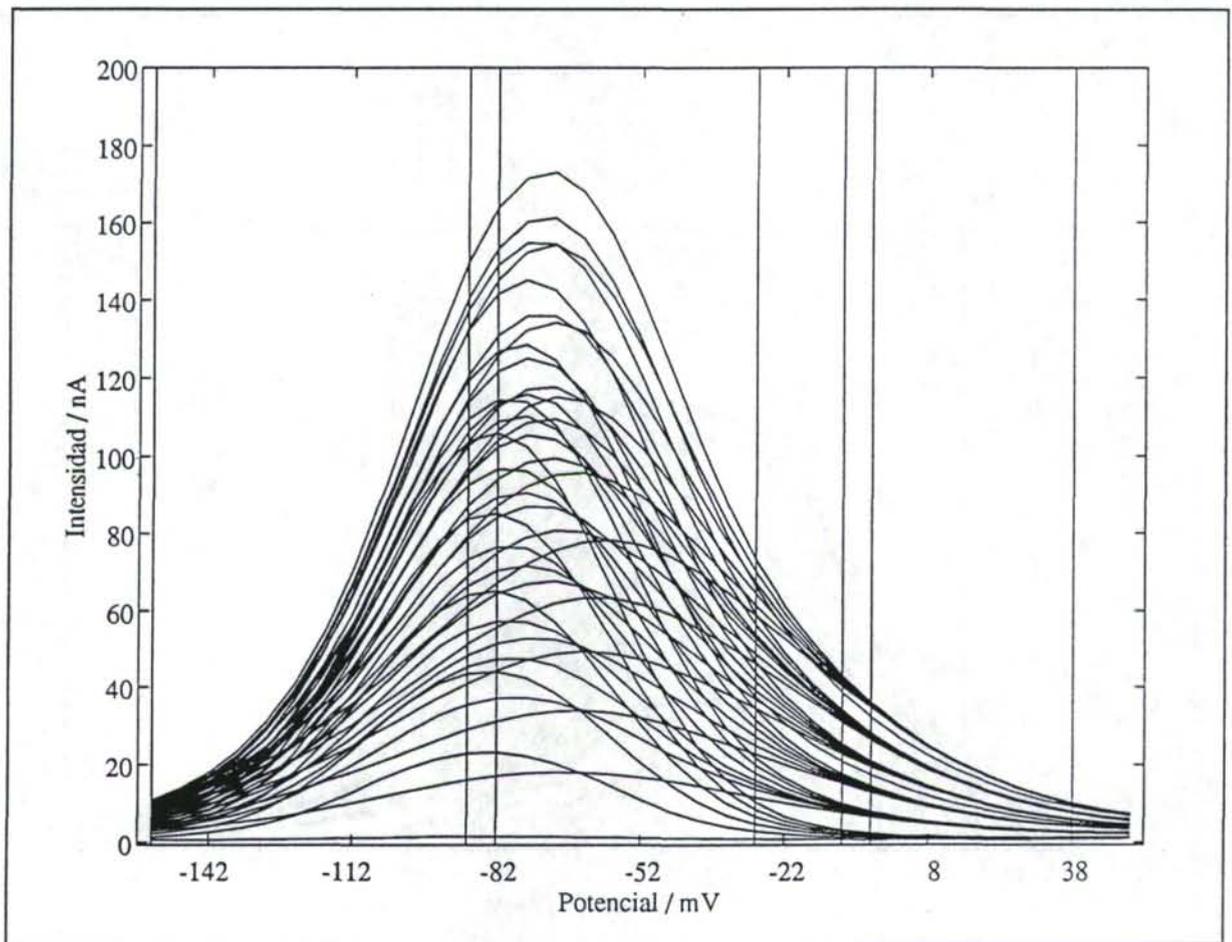


Figura 6a Variables seleccionadas por el algoritmo genético para el Cu (calibrado J). Los picos voltamperométricos se corresponden de izqda. a dcha. con Cu y Fe.

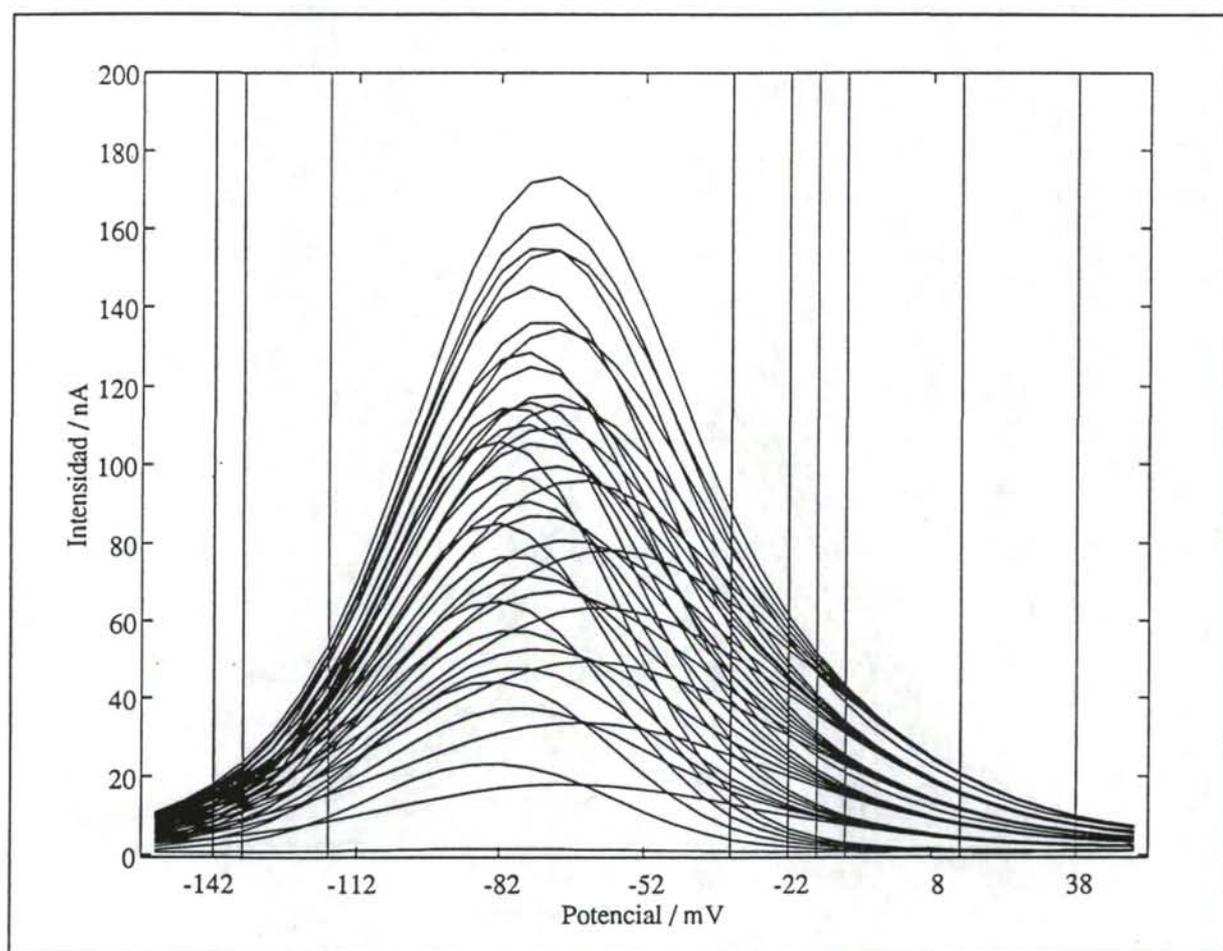


Figura 6b Variables seleccionadas por el algoritmo genético para el Fe (calibrado J). Los picos voltamperométricos se corresponden de izqda. a dcha. con Cu y Fe.

Tabla IV.4.6 Errores relativos (en %) obtenidos para las muestras test del calibrado A tras la selección de variables.

Número de muestra	Analito			
	Cobre	Plomo	Cadmio	Zinc
18	4.49	0.73	2.20	-2.44
20	-1.03	0.15	0.07	-2.33
23	-1.48	-3.24	1.44	17.07
25	-0.01	-0.83	2.99	3.78
28	0.66	0.51	-0.80	1.83
30	-0.06	1.09	0.08	1.12
33	-1.57	4.07	-0.34	-1.10
35	-0.17	1.16	1.24	3.34

Tabla IV.4.7 Errores relativos (en %) obtenidos para las muestras del conjunto test del calibrado C tras la selección de variables.

Número de muestra	Analito			
	Cobre	Plomo	Cadmio	Zinc
18	-3.66	10.30	3.36	-0.26
20	3.02	10.30	1.83	-2.22
23	-5.78	0.66	-9.26	-0.16
25	0.25	2.47	-7.96	-7.80
28	13.43	15.32	1.94	6.87
30	7.44	14.13	9.42	2.36
33	-6.59	1.35	-8.77	-3.05
35	3.15	0.65	-5.65	1.97

Tabla IV.4.8 Errores relativos (%) correspondientes a las muestras test del calibrado E tras la selección de variables.

Número de muestra	Analito		
	Plomo	Estaño	Cadmio
10	1.29	-0.30	-0.58
12	0.97	-2.22	-0.48
15	0.69	0.74	0.91
17	-1.28	1.96	-0.95
20	-0.61	-2.34	-0.91
22	0.99	1.26	0.07

Tabla IV.4.9 Errores relativos (en %) para las muestras del conjunto test del calibrado H tras la selección de variables.

Muestra	Analito	
	Talio	Plomo
t1	10.50	2.89
t2	8.63	3.18
t3	-30.77	4.84
t4	-7.18	-1.91

Tabla IV.4.10 Errores relativos (en %) para las muestras del conjunto test del calibrado J tras la selección de variables.

Muestra	Analito	
	Cobre	Hierro
t1	1.80	-2.06
t2	-0.14	0.03
t3	-1.87	-0.06
t4	-3.39	0.22

Comparación de resultados mediante test no paramétricos

Con el fin de comparar objetivamente los errores obtenidos para una misma determinación llevada a cabo con selección de variables y sin ella, se han utilizado los test no paramétricos indicados anteriormente. Los test de los signos y los rangos con signo de Wilcoxon han sido aplicados para evaluar las diferencias existentes entre el valor absoluto del error relativo de las concentraciones calculadas con esos los dos procedimientos de análisis.

Tabla IV.4.11 Probabilidades actuales de los test no paramétricos de rangos y de signos de Wilcoxon para los calibrados A, C, E, H y J. H_0 : la mediana es 0, y H_1 : la mediana es distinta de cero.

	Calibrados				
	A	C	E	H	J
Mediana	-0.155	1.440	0.085	2.815	0.000
Test de rangos	0.164	0.031*	0.207	0.021*	0.584
Test de los signos	0.112	0.52	0.239	0.077	0.617

* $p \leq 0.05$ implica rechazar H_0

La aplicación de estos test a los resultados obtenidos para todos los ejemplos estudiados permite concluir que no existe evidencia suficiente para rechazar la hipótesis nula en la casi totalidad de los casos, Tabla IV.4.11, a un nivel de significación del 5%. Esto equivale a decir que los resultados obtenidos por ambos métodos son significativamente iguales.

IV.4.2 Aspectos cualitativos relacionados con la selección de variables

Como anteriormente se ha comentado, el procedimiento de selección de variables es capaz de señalar aquellos potenciales a los que se producen procesos complejos relacionados con el analito de interés que, en principio, pueden no ser apreciables a través de la observación

directa de las señales. Por ello, a continuación se ha realizado un análisis sobre los aspectos cualitativos de los procedimientos de selección de variables indicados en el epígrafe anterior.

Las Figuras IV.4.4a-d, correspondientes al calibrado A, muestran cómo el algoritmo genético ha seleccionado no sólo los potenciales correspondientes a la señal de cada uno de los analitos de interés, sino que también ha optado por otras variables que aportan información a los modelos PLS sobre otras fuentes de variabilidad no relacionadas directamente con las concentraciones. Éstas podrían estar relacionadas con el ruido de fondo que se observa en las citadas figuras, el cual debe de tener un peso significativo en todas las intensidades registradas. Esto confirma la necesidad de utilizar varios sensores, incluso en aquellos casos en que éstos no sean específicos del pico correspondiente al analito que se está calibrando.

El siguiente grupo de gráficas, Figuras IV.4.5a-d, se corresponde con la selección de variables realizada en la determinación de zinc, cadmio, plomo y cobre utilizando una técnica de redisolución. Queda patente, a diferencia del caso anterior, que las variables seleccionadas se encuentran localizadas no sólo en la señal correspondiente a cada uno de los analitos de interés, sino también en las de los restantes elementos. Este comportamiento del algoritmo genético está relacionado con la formación de compuestos intermetálicos que hace que existan interferencias importantes entre los distintos analitos que están siendo determinados.

Asimismo, en las variables seleccionadas en las Figuras IV.4.6a-c, correspondientes a la determinación mediante polarografía de Pb(II), Sn(IV) y Cd(II), también se observa la existencia de algún tipo de interacción entre los dos últimos elementos. En los procedimientos de selección de variables aplicados a los calibrados de estos dos analitos, Figuras IV.4.6b y IV.4.6c, se han seleccionado siempre potenciales relacionados con ambos. La figura correspondiente al plomo muestra cómo es necesario considerar potenciales correspondientes al pico del estaño, con el cual está solapado el del plomo, además de un potencial en la zona del polarograma en la que se encuentra la señal del cadmio.

Al analizar el caso del solapamiento de picos en la determinación de plomo y talio por voltamperometría de redisolución, Figuras IV.4.7, la selección de variables realizada por el algoritmo genético muestra cómo este método ha evitado aquellos potenciales relacionados con la señal correspondiente al blanco experimental (línea punteada). Esto resulta especialmente evidente en la selección de potenciales correspondiente al calibrado del talio, para el que únicamente selecciona potenciales situados en las colas del pico

voltamperométrico. En caso de que no se dispusiera de la señal del blanco experimental, éste sería un indicio claro de que en esa zona del voltamperograma se está produciendo algún fenómeno que interferiría en la determinación del talio si únicamente se considerara la intensidad correspondiente al máximo de pico.

Por último, las Figuras IV.4.8a y IV.4.8b correspondientes a la determinación mediante técnicas de redisolución de Cu(II) y Fe(III) cuando este último se encuentra en exceso, señalan cuáles son las zonas conflictivas del voltamperograma. En la figura correspondiente al cobre, las variables seleccionadas se encuentran en las zona del máximo de pico de la señal del cobre y en las colas del mismo, mientras que en la selección de variables correspondiente al hierro únicamente se hace uso de las colas del pico polarográfico registrado. Ésto se debe a que el pico del hierro sufre un desplazamiento hacia potenciales más positivos a medida que aumenta la concentración de este elemento.

V. CONCLUSIONES

1. Las técnicas de calibración 'blanda' se han mostrado capaces de modelar señales obtenidas como resultado de procesos complejos, donde la metodología univariante falla. La regresión multivariante PLS ha sido utilizada para modelar la existencia de reacciones químicas, en concreto reacciones de dimerización, acopladas a los procesos electródicos en la determinación polarográfica de benzaldehído. Este tipo de reacciones acopladas ha dado lugar a la aparición de múltiples señales que se han solapado o que han sufrido desplazamientos a lo largo del eje de potenciales.
2. A diferencia de los procedimientos clásicos que buscan especificidad y selectividad para las señales, la capacidad de modelar señales complejas exhibida por la técnica de regresión PLS ha permitido llevar a cabo la determinación polarográfica de benzaldehído en un amplio rango de valores de pH, por lo que esta condición experimental ha dejado de ser limitante.
3. Ha sido posible dar un significado químico a las variables latentes obtenidas por los modelos PLS construidos. Ello ha indicado que no se trataba de un mero artificio matemático y ha ayudado además a analizar la realidad química del problema y a comprender los procesos que se puedan estar produciendo.
4. Los errores obtenidos en el análisis, que presentan una media para el valor absoluto de los errores relativos igual a 2.10%, han confirmado la viabilidad de la determinación de benzaldehído mediante polarografía diferencial de impulsos utilizando modelos de regresión multivariante PLS.
5. El uso de técnicas multivariantes de regresión permite llevar a cabo determinaciones multianalito simultáneas. Esto implica que a partir de un único conjunto de datos procedentes de una experiencia adecuadamente diseñada es posible determinar la concentración de distintos componentes en una muestra compleja.

6. La regresión PLS se ha mostrado útil en la determinación de Cu(II), Pb(II), Cd(II) y Zn(II) mediante polarografía diferencial de impulsos, superando las dificultades encontradas por las técnicas de regresión univariante utilizadas, que condujeron a errores relativos en las muestras problema de hasta un 40% en algunos casos, frente a un 2.37% de error medio global (media del valor absoluto de los errores relativos). La transferencia de calibrado realizada en este análisis ha permitido reducir el número de patrones de calibrado de 28 a 8, manteniéndose el mismo nivel de error en las determinaciones.
7. La formación de compuestos intermetálicos y otras interferencias que se producen en la determinación de Cu(II), Pb(II), Cd(II) y Zn(II) por voltamperometría diferencial de impulsos con redisolución anódica se pueden modelar utilizando la técnica de regresión multivariante PLS, con errores globales medios del 3.7%. Asimismo se ha mostrado viable la transferencia de estos modelos de calibrado a través del tiempo, lo cual permite realizar la determinación simultánea de esos cuatro analitos mediante la técnica experimental indicada, con independencia de que existan relaciones internas entre las diferentes variables predictoras. Se ha reducido el esfuerzo experimental necesario para llevar a cabo el análisis, pasando en este caso de un total de 28 muestras de calibrado a 16.
8. La regresión PLS ha sido aplicada con éxito (errores medios globales del 0.80%) para llevar a cabo la resolución de señales solapadas en la determinación mediante polarografía diferencial de impulsos de Pb(II), Sn(IV) y Cd(II). Esta regresión multivariante resulta adecuada para extraer la información útil que pueda encontrarse solapada por otras fuentes de variabilidad no relacionadas con el analito de interés. En la transferencia de calibrado en el tiempo correspondiente a esta determinación se ha conseguido recalibrar el sistema utilizando únicamente 7 muestras de calibrado, frente a las 23 que constituían el calibrado patrón.
9. Asimismo, esta técnica de regresión multivariante ha sido capaz de modelar la señal del blanco experimental y de obviar la interferencia que ésta introduce en la determinación por voltamperometría diferencial de impulsos con redisolución anódica de Pb(II) y Tl(I), con errores medios globales del 2.15%. En este caso, el problema se planteaba por el solapamiento de las señales correspondientes a los dos analitos de

interés así como de la correspondiente al propio blanco experimental que, evidentemente, no formaba parte del diseño utilizado. En estas condiciones, la transferencia de calibrado realizada ha permitido reducir el número de patrones de calibrado de 25 a 9.

10. La determinación de Cu(II) en presencia de exceso de Fe(III) por voltamperometría diferencial de impulsos con redisolución anódica ha sido llevada a cabo, permitiendo el procedimiento multivariante aplicado realizar la determinación simultánea de ambos analitos, con errores medios globales del 1.23%. Es decir, la regresión PLS ha permitido obviar el efecto matriz debido a la presencia de un exceso de Fe(III) en las determinaciones de Cu(II) por DPASV. En este caso, se ha estudiado además la viabilidad de llevar a cabo la transferencia de calibrado entre instrumentos.

11. El uso de un algoritmo genético para seleccionar variables en determinaciones voltamperométricas ha permitido construir modelos de calibrado multivariante PLS con la misma o mejor (benzaldehído) capacidad de predicción que los modelos construidos con todas ellas, según concluyen los test no paramétricos utilizados. Desde un punto de vista cualitativo, la información aportada con el análisis de las variables seleccionadas por el algoritmo genético, es útil ya que evidencia la importancia del uso de señales no específicas aun cuando las señales voltamperométricos *'parezcan'* de calidad analítica.

VI. BIBLIOGRAFÍA

1. Geladi P., Notes on the history and nature of partial least squares (PLS) modelling, *Journal of Chemometrics*, 2 (1988) 231-246.
2. Booksh K.S. y Kowalski B.R., Theory of Analytical Chemistry, *Analytical Chemistry*, 66 (1994) 782A-791A.
3. Haaland D.M. y Thomas E.V., Partial least-squares methods for spectral analyses. 1. Relation to other quantitative calibration methods and the extraction of qualitative information, *Analytical Chemistry*, 60 (1988) 1193- 1202.
4. Haaland D.M. y Thomas E.V., Partial least-squares methods for spectral analyses. 2. Application to simulated and glass spectral data, *Analytical Chemistry*, 60 (1988) 1202-1208.
5. Brown S.D. y Bear R.S., Chemometrics techniques in electrochemistry: a critical review, *Critical Reviews in Analytical Chemistry*, 24 (1993) 99-131.
6. Stone M. y Brooks R.J., Continuum regression: cross-validated sequentially constructed prediction embracing ordinary least squares, partial least squares and principal components regression, *Journal of Royal Statistical Society, Series B*, 52 (1990) 237- 269.
7. Burnham A.J., Viveros R. y Macgregor J.F., Frameworks for latent variable multivariate regression, *Journal of Chemometrics*, 10 (1996) 31-45.
8. Martens M. y Martens H., Partial least squares regression, cap. 9 en *Statistical Procedures in Food Chemistry*, editado por J.R. Piggott, Elsevier Applied Science, London, 1986.
9. Martens H. y Næs T., *Multivariate Calibration*, Wiley, 2ª ed., New York, 1992.
10. Brereton R.G., *Chemometrics, applications of mathematics and statistics to laboratory systems*, Ellis Horwood, Chichester, 1990.
11. Meloun M., Militký J. y Forina M., *Chemometrics for analytical chemistry*, Vol. 2, Ellis Horwood, New York, 1994.
12. Brown P.J., Measurement, regression and calibration, in *Oxford statistical science series*, vol. 12, Clarendon press, Oxford, 1993.
13. de Jong S., SIMPLS: an alternative approach to partial least squares regression, *Chemometrics and Intelligent Laboratory Systems*, 18, 251-263 (1993).

14. Lindgren F., Geladi P. y Wold S., The kernel algorithm for PLS, *Journal of Chemometrics*, 7 (1993) 45-60.
15. Wold H., Soft modelling by latent variables: the non-linear iterative partial least squares approach in perspectives in probability and statistics. Paper in honour of M.S. Bartlett, Academic Press, London, 1975.
16. Wold H., Nonlinear estimation by iterative least square procedures, en *Research Papers in Statistics*, editado por F.N. David, Wiley, New York, 1966, pp 411-444.
17. Wold H., Estimation of principal components and related models by iterative least squares in *Multivariate Analysis I*, editado por P.K. Krishnaiah, Academic Press Inc., New York, 1966, pp. 391-420.
18. Massy W.F., Principal components regression in exploratory statistical research, *Journal of American Mathematics and Statistics Association*, 60 (1965) 234-246.
19. Jackson J.E., *A user's guide to principal components*, Wiley, New York, 1991.
20. Garthwaite P.H., An interpretation of partial least squares, *Journal of the American Statistical Association*, 89 (1994) 122-127.
21. Lanteri S., Full validation procedures for feature selection in classification and regression problems, *Chemometrics and Intelligent Laboratory Systems*, 15 (1992) 159-169.
22. Stone M., Cross-validatory choice and assesment of statistical predictions, *Journal of the Royal Statistic Society, serie B*, 36 (1974) 111-147.
23. Wold S., Cross-validatory estimation of the number of components in factor and principal components models, *Technometrics*, 20 (1978) 397-406.
24. Lohmöller J.B., *Latent variable path modeling with partial least squares*, Physica-Verlag, Heidelberg, 1989.
25. Otto M. y Wegscheider W., Spectrophotometric multicomponent analysis applied to trace metal determinations, *Analytical Chemistry*, 57 (1985) 63-69.
26. Arcos M.J., Ortiz M.C. y Villahoz B., Genetic-algorithm-based wavelength selection in multicomponent spectrometry determinations by PLS. Application on indomethacin and acemethacin mixture, *Analytica Chimica Acta*, 34 (1996) 245-262.
27. Xie Y.L., Baeza-Baeza J.J. y Ramis-Ramos G., Kinetic spetrophotometric resolution of binary mixtures using three-way partial least squares, *Chemometrics and Intelligent Laboratory Systems*, 27 (1995) 211-220.
28. De la Guardia M., Khalaf K.D., Hasan B.A., Morales-Rubio A., Arias J.J., García Fraga J.M., Jiménez A.I. y Jiménez F., Simultaneous kinetic spectrophotometric determination of five phenolic compounds by reaction with p-aminophenol, using

- partial least squares data treatment, *Analyst*, 121 (1996) 1321-1326.
29. Blanco M., Coello J., Iturriaga H., Maspocho S. y Redón M., Partial least-squares regression for multicomponent kinetic determinations in linear and non-linear systems, *Analytica Chimica Acta*, 303 (1995) 309-320.
 30. Cirovic D.A., Brereton R.G., Walsh P.T., Ellwood J.A. y Scobbie E., Application of partial least squares calibration to measurements of polycyclic aromatic hydrocarbons in coal tar pitch volatiles, *Analyst*, 121 (1996) 575-580.
 31. Araujo P.W., Cirovic D.A. y Brereton R.G., Multivariate calibration of chlorophyll *a* using partial least squares and electronic absorption spectroscopy, *Analyst*, 121 (1996) 581-584.
 32. MacLaurin P., Worsfold P.J., Norman P. y Crane M., Partial least squares resolution of multianalyte flow injection data, *Analyst*, 118 (1993) 617-622.
 33. Henrion A., Henrion R., Henrion G. y Scholz F., Application of partial least-squares regression for signal resolution in differential pulse anodic stripping voltammetry of thallium and lead, *Electroanalysis*, 2 (1990) 309-312.
 34. Otto M. y Thomas J.R.D., Model studies on multiple channel analysis of free magnesium, calcium, sodium, and potassium at physiological concentration levels with ion-selective electrodes, *Analytical Chemistry*, 57 (1985) 2647-2651.
 35. Jagner D., Renman L. y Stefansdottir S.H., Determination of iron(III) and titanium(IV) as their solochrome violet RS complexes by constant-current stripping potentiometry. 2. Partial least-squares regression calibration procedure for iron (III) and titanium(IV), *Analytica Chimica Acta*, 281 (1993) 315-321.
 36. Herrero A., Ortiz M.C., Arcos J., López-Palacios J. y Sarabia L.A., The Multiple Standard Addition with Latent Variables Method, MSALV. Application to Determination of Cu in Wine by Using Differential-pulse Anodic Stripping Voltammetry, *Analytica Chimica Acta*, 293 (1994) 277-293.
 37. Bersier P.M. y Bersier J., Applied polarography and voltammetry of organic compounds in practical day-to-day analysis. Part. I: applied polarographic and voltammetric techniques, *Critical Reviews in Analytical Chemistry*, 16 (1985) 15-79.
 38. Bersier P.M. y Bersier J., Applied polarography and voltammetry of organic compounds in practical day-to-day analysis. Part. II., *Critical Reviews in Analytical Chemistry*, 16 (1985) 81-128.
 39. Patriarche G.J. y Vire J.C., Applications of polarography and voltammetry in analysis for drugs, *Analytica Chimica Acta*, 196 (1987) 193-204.
 40. Anderson J.E., Bond A.M. y Jones R.D., Differential pulse polarography at the static mercury drop electrode, *Analytical Chemistry*, 53 (1981) 1016-1020.

41. Bard A.J. y Faulkner L.R., *Electrochemical methods, fundamentals and applications*, Wiley, Chichester, 1980.
42. Laviron E., Vallat A. y Meunier-Prest R., The reduction mechanism of aromatic nitro compounds in aqueous medium. Part V. The reduction of nitrosobenzene between pH 0.4 and 13, *Journal of Electroanalytical Chemistry*, 379 (1994) 427-435.
43. Bermejo E., Zapardiel J., Pérez J.A., Huerta A. y Hernández L., Voltammetric studies of a psychotropic drug with nitro groups. Determination of flunitrazepam in urine using HMDE, *Talanta*, Vol.40, No. 11, (1993) 1649-1656.
44. Hernández I., Hernández P. y Nieto O., Determination of thyroxine in urine by cathodic stripping square-wave voltammetry, *Analyst*, 119 (1994) 1579-1583.
45. Wold H., Soft modelling in systems under indirect observation II, editado por K.G. Jöreskog y H. Wold, North-Holland, New York, 1982, pp 1-54.
46. Wold H., Partial Least Squares, *Encyc. Stat. Sci.*, 6, editado por S. Kotz y N. Johnson, Wiley, Inc. New York, 1985, pp. 581-591.
47. Guiberteau A., Galeano T., Espinosa-Mansilla A., López de Alba P.L. y Salinas F., Abilities of differentiation and partial least squares methods in the analysis by differential pulse polarography. Simultaneous determination of furazolidone and furaltadone, *Analytica Chimica Acta*, 302 (1995) 9-19.
48. Guiberteau A., Galeano T., Espinosa-Mansilla A. y Salinas F., Resolution of ternary mixtures of nitrofurantoin, furazolidone and furaltadone by application of partial least squares analysis to the differential pulse polarographic signals, *Talanta*, 41 (1994) 1821-1832.
49. Reviejo A.J., Buyo F.J., Pingarrón J.M. y Peral J.L., Application of partial least-squares regression to the suitability of multicomponent polarographic determination of organochlorine pesticides in emulsified medium, *Electroanalysis*, 5 (1993) 303-309.
50. AOAC Official methods of analysis (1995).
51. Lehotay J., Hromulakova K., Skacani I. y Oktavec D., Determination of benzaldehyde in waste-gas by high-performance liquid-chromatography, *Chemicke Listy*, 90 (1996) 260-263.
52. Hewala I.I., Difference spectrophotometric assay of benzaldehyde in benzyl alcohol, *Talanta*, 40 (1993) 919-923.
53. Chen Y.L. y Chou T.C., Electrochemical reduction of benzaldehyde using Pt-Pb/Naftion as electrode, *Journal of Applied Electrochemistry*, 24 (1994) 434-438.
54. Koshy V.J., Prasad J.V., Kalpana G. y Satish S., Investigations on intermediate acetal formation during polarographic estimation of certain benzaldehyde derivatives, *Analytica Chimica Acta*, 307 (1995) 55-60.

55. Wang L.H. y Chen Z.S., Determination of free formaldehyde in nail preparations by differential pulse voltammetry, *Electroanalysis*, 8 (1996) 842-846.
56. Perrin D.D. y Dempsey B., *Buffers for pH and metal ion control*, Chapman and Hall Ltd., London, 1974.
57. Forina M., Leardi R., Armanino C. y Lanteri S., *PARVUS: An Extendable Package of Programs for Data Exploration, Classification and Correlation*, Ver 1.1, Elsevier Scientific Software, 1990.
58. *STATGRAPHICS*, Ver. 5, STSC, Inc., Rockville, MD, 1991.
59. *MATLAB*, High-Performance Numeric Computation and Visualization Software, Ver. 4.0, The MathWorks Inc., Natick, Mass. 1992.
60. Galus Z., *Fundamentals of electrochemical analysis*, 2^a Ed., Ellis Horwood Series in Analytical Chemistry, Chichester, 1994.
61. Law H.D., Electrolytic reduction. I, Aromatic aldehydes, *Journal of the Chemical Society*, 89 (1906) 1512-1519.
62. Law H.D., Electrolytic reduction. Part III, Electrolysis in alkaline solution, *Journal of the Chemical Society*, 91 (1907) 748-760.
63. Law H.D., Electrolytic reduction. Part IV, Aromatic aldehydes, *Journal of the Chemical Society*, 99 (1911) 1113-1117.
64. Zuman P., Barnes D. y Ryvolová-Kejharová A., Polarographic reductions of some carbonil compounds, *Discussions of the Faraday Society*, 45 (1968) 202-226.
65. Mairanovskii S.G., *Catalytic and kinetic waves in polarography*, Plenum Press, New York, 1968.
66. Andrieux C.P., Nadjo L. y Savéant J.M., Electrodimerization. I. One-electron irreversible dimerization. Diagnostic criteria and determination procedures for voltammetric studies, *Journal of Electroanalytical Chemistry*, 26 (1970) 147-186.
67. Nadjo L. y Savéant J.M., Dimerization, disproportionation and E.C.E. mechanisms in the reduction of aromatic carbonyl compounds in alkaline media, *Journal of Electroanalytical Chemistry*, 33 (1971) 419-451.
68. Nadjo L. y Savéant J.M., Electrodimerization. VIII Role of proton transfer reactions in the mechanism of electrohydrodimerization. Formal kinetics for voltammetric studies (linear sweep, rotating disc, polarography), *Journal of Electroanalytical Chemistry*, 44 (1973) 327-366.
69. De Juan M., Heras A.M., Camacho L., Ávila J.L. y Jiménez C., Differential pulse polarography for a dimerization process, *Journal of Electroanalytical Chemistry*, 191 (1985) 303-310.

70. Rodríguez-Mellado J.M., Domínguez M. y Blazquez M., Derivation and experimental verification of approximate explicit equations in differential pulse polarography. Part II, Second-order processes, *Journal of Electroanalytical Chemistry*, 201 (1986) 237-246.
71. Rodríguez Mellado J.M. y Ruiz Montoya M., CEC mechanisms in the electroreduction of α -dicarbonyl compounds on mercury electrodes, *Journal of Electroanalytical Chemistry*, 365 (1994) 71-78.
72. Hayes J.W., Ružić I., Smith D.E., Booman G.L. y Delmastro J.R., Fundamental harmonic A.C. polarography with irreversible dimerization following the charge transfer step. Theory and experimental results with the benzaldehyde system, *Electroanalytical Chemistry and Interfacial Electrochemistry*, 51 (1974) 269-285.
73. Ružić I., Sobel H.R. y Smith D.E., On the theory for D.C. and fundamental harmonic A.C. polarography with the first-order consecutive ECE mechanism. *Electroanalytical Chemistry and Interfacial Electrochemistry*, 65 (1975) 21-56.
74. Kim M.H., Analysis of the first-order EC process by direct current and derivative polarography, *Analytical Chemistry*, 59 (1987) 2136-2144.
75. Vykhodtseva L.N., Nekrasov L.N. y Shtern M.M., Cathodic reduction of aromatic carbonyl compounds in aprotic solvents. Rotating ring-disk study of the two-electron reduction products of fluorenone, benzophenone and benzaldehyde, *Soviet Electrochemistry*, 28 (1992) 757-764.
76. Cai X. y Kalcher K., Studies on the electrocatalytic reduction of aliphatic aldehydes on palladium-modified carbon paste electrodes, *Electroanalysis*, 6 (1994) 397-404.
77. Liu K.E. y Abruña H.D., Electroanalysis of aromatic aldehydes with modified carbon paste electrodes, *Analytical Chemistry*, 61 (1989) 2599-2602.
78. Chen Y.L. y Chou T.C., Metals and alloys bonded on solid polymer electrolyte for electrochemical reduction of pure benzaldehyde without liquid supporting electrolyte, *Journal of Electroanalytical Chemistry*, 360 (1993) 247-259.
79. Murphy M.M., O'Dea J.J., Arn D. y Osteryoung J.G., Theory for cyclic staircase voltammetry for first-order coupled reactions, *Analytical Chemistry*, 62 (1990) 903-909.
80. Durant A., François H., Reisse J. y Kirschdemesmaeker A., Sonoelectrochemistry. The effects of ultrasound on organic electrochemical reduction, *Electrochimica Acta*, 41 (1996) 277-284.
81. Andrieux C.P. y Savéant J.M., Electro-oxidation of 1,2-ene-diaminas, *Journal of Electroanalytical Chemistry*, 28 (1970) 339-348.

82. Charlot G., Badoz-Lambling J. y Tremillon B., *Las reacciones electroquímicas*, Toray-Masson, Barcelona, 1969, p. 67.
83. Hollander M. y Wolfe D.A., *Nonparametric statistical methods*, Wiley, Chichester, 1973, p 138.
84. de Noord O.E., *Multivariate calibration standardization*, *Chemometrics and Intelligent Laboratory System*, 25 (1994) 85-97.
85. Gemperline P.J., *Mixture analysis using factor analysis I: calibration and quantitation*, *Journal of Chemometrics*, 3 (1989) 549-568.
86. Næs T., *The design of calibrations in near infra-red reflectance analysis by clustering*, *Journal of Chemometrics*, 1 (1987) 121-134.
87. Puchwein G., *Selection of calibration samples for near-infrared spectrometry by factor analysis of spectra*, *Analytical Chemistry*, 60 (1988) 569-573.
88. Schostack K.J. y Malinowski E.R., *Preferred set selection by iterative key set factor analysis*, *Chemometrics and Intelligent Laboratory Systems*, 6 (1989) 21-29.
89. Ferré J. y Rius F.X., *Selection of the best calibration sample subset for multivariate regression*, *Analytical Chemistry*, 68 (1996) 1565-1571.
90. Osborne B.G. y Fearn T., *Collaborative evaluation of universal calibrations for the measurement of protein and moisture in flour by near infrared reflectance*, *Journal of Food Technology*, 18 (1983) 453-460. (pedir)
91. Mark H. y Workman J., *A new approach to generating transferable calibrations for quantitative near-infrared spectroscopy*, *Spectroscopy*, 3 (1988) 28-36. (pedir)
92. Shenk J.S. y Westerhaus M.O., *New standardization and calibration procedures for NIRS analytical systems*, *Crop. Science*, 31 (1991) 1694-1696.
93. Bouveresse E., Massart D.L. y Dardenne P., *Modified algorithm for standardization of near-infrared spectrometric instruments*, *Analytical Chemistry*, 67 (1995) 1381-1389.
94. Wang Y., Veltkamp D.J. y Kowalski B.R., *Multivariate instrument standardization*, *Analytical Chemistry*, 63 (1991) 2750-2756.
95. Wang Z., Dean T. y Kowalski B.R., *Additive background correction in multivariate instrument standardization*, *Analytical Chemistry*, 67 (1995) 2379-2385.
96. Wang Y., Lysaght M.J. y Kowalski B.R., *Improvement of multivariate calibration through instrument standardization*, *Analytical Chemistry*, 64 (1992) 562-564.
97. Bouveresse E., Hartmann C. y Massart D.L., *Standardization of near-infrared spectrometric instruments*, *Analytical Chemistry*, 68 (1996) 982-990.

98. Bouveresse E. y Massart D.L., Improvement of the piecewise direct standardisation procedure for the transfer of NIR spectra for multivariate calibration, *Chemometrics and Intelligent Laboratory Systems*, 32 (1996) 201-213.
99. Kennard R.W. y Stone L.A., *Technometrics*, 11 (1969) 137-149. (buscar título)
100. Wang Y. y Kowalski B.R., Temperature-compensating calibration transfer for near-infrared filter instruments, *Analytical Chemistry*, 65 (1993) 1301-1303.
101. Wang Y. y Kowalski B.R., Calibration transfer and measurement stability of near-infrared spectrometers, *Applied Spectroscopy*, 46 (1992) 764-771.
102. Forina M., Drava G., Armanino C., Boggia R., Lanteri S., Leardi R., Corti P., Conti P., Giangiacomo R., Galliena C., Bigoni R., Quartari I., Serra C., Ferri D., Leoni O. y Lazzeri L., Transfer of calibration function in near-infrared spectroscopy, *Chemometrics and Intelligent Laboratory Systems*, 27 (1995) 189-203.
103. Wang Y. y Kowalski B.R., Standardization of second-order instruments, *Analytical Chemistry*, 65 (1993) 1174-1180.
104. Nørgaard L., Direct standardisation in multi wavelength fluorescence spectroscopy, *Chemometrics and intelligent laboratory systems*, 29 (1995) 283-293.
105. Blanco M., Coello J., Iturriaga H., MasPOCH S. y Rovira E., Wavelength calibration transfer between diode array UV-visible spectrophotometers, *Applied Spectroscopy*, 49 (1995) 593-597.
106. Sales F., Callao M.P. y Rius F.X., Standardization techniques for multivariate models with UV/VIS data, comunicación en el Congreso Internacional: VI Chemometrics in Analytical Chemistry, Tarragona, 1996.
107. Herrero A. y Ortiz M.C., Multivariate calibration transfer for multicomponent polarographic data, comunicación en el Congreso Internacional: VI Chemometrics in Analytical Chemistry, Tarragona, 1996.
108. Kowalski B.R. y Seasholtz M.B., Recent developments in multivariate calibration, *Journal of Chemometrics*, 5 (1991) 129-145. (buscar)
109. Alloway B.J. (Ed.), *Heavy metals in soils*, 2ª ed., Chapman & Hall, London, 1995.
110. Roeva N.N., Rovinskii F.Y. y Kononov E.Y., Special features of the behavior of heavy-metals in various natural environments, *Journal of Analytical Chemistry*, 51 (1996) 352-364.
111. Welch A.J. y Chapman S.K. (Ed.), *The chemistry of the copper and zinc triads*, Royal Society of Chemistry, Cambridge, 1993.
112. Ivaska A., Lewenstam A. y Sara R. (Ed.), *Contemporary electroanalytical Chemistry*, Plenum Press, New York, 1990.

113. Kaldova R., Electrochemical analysis of environmental control, *Electroanalysis*, 2 (1990) 341-346.
114. Ashley K., Electroanalytical applications in occupational and environmental health, *Electroanalysis* (1994) 805-820.
115. Kalcher K., Kauffmann J.M., Wang J., Švancara I., Vytřas K., Neuhold C. y Yang Z., Sensors based on carbon paste in electrochemical analysis: a review with particular emphasis on the period 1990-1993, *Electroanalysis*, 7 (1995) 5-22.
116. Locatelli C., Vasca E., Bigli C., Fagioli F. y Garai T., Determination of metals in a multicomponent system by differential pulse and alternating current anodic stripping voltammetry, *Electroanalysis*, 8 (1996) 165-170.
117. Daniele S., Baldo M.A., Ugo P. y Mazzochin G.A., Determination of heavy metals in real samples by anodic stripping voltammetry with mercury microelectrodes. Application to rain and sea waters, *Analytica Chimica Acta*, 219 (1989) 19-26.
118. Jagner D., Instrumental approach to potentiometric stripping analysis of some heavy metals, *Analytical Chemistry*, 50 (1978) 1924-1929.
119. Adeloju S.B., Bond A.M. y Briggs M.H., Multielement determination in biological materials by differential pulse voltammetry, *Analytical Chemistry*, 57 (1985) 1386-1390.
120. Florence T.M., Trace element speciation by anodic stripping voltammetry, *Analyst*, 117 (1992) 551-557.
121. Bond A.M., *Modern polarographic methods in Analytical Chemistry*, Marcel Dekker, New York, 1980.
122. Anderson V.L. y McLean R.A., Design of experiments: a realistic approach, in *Statistics: textbooks and monographs*, vol. 5, Marcel Dekker, New York, 1974.
123. Box G.E.P., Hunter W.G. y Hunter J.S., *Estadística para investigadores: introducción al diseño de experimentos, análisis de datos y construcción de modelos*, Reverté, Barcelona, 1989.
124. Kalivas J.H. y Lang P.M., Mathematical analysis of spectral orthogonality, in *Practical spectroscopy series*, vol. 17, Marcel Dekker, New York, 1994.
125. Wise B.M., PLS Toolbox for use with MATLAB™, Ver. 1.3. Available from the author.
126. Marengo E. y Todeschini R., A fast method for the calculation of partial least squares coefficients, *Chemometrics and Intelligent Laboratory Systems*, 12 (1991) 117-120.
127. Vydra F., Štulík K. y Juláková E., *Electrochemical Stripping Analysis*, Ellis Horwood, Chichester, 1976.

128. Copeland T.R., Osteryoung R.A. y Skogerboe R.K., Elimination of copper-zinc intermetallic interferences in anodic stripping voltammetry, *Analytical Chemistry*, 46 (1974) 2093-2097.
129. Crosmun S.T., Dean J.A. y Stokely J.R., Pulsed anodic stripping voltammetry of zinc, cadmium and lead with a mercury-coated wax-impregnated graphite electrode, *Analytica Chimica Acta*, 75 (1975) 421-430.
130. Brett C.M.A., Brett A.M.O. y Tugulea L., Anodic stripping voltammetry of trace-metals by batch injection-analysis, *Analytica Chimica Acta*, 322 (1996) 151-157.
131. Golimowski J. y Szczepanska T., Voltammetric method for determination of Zn, Cd, Pb, Cu and Ni in interstitial water, *Fresenius Journal of Analytical Chemistry*, 354 (1996) 735-737.
132. Barbeira P.J.S., Mazo L.H. y Stradiotto N.R., Determination of trace amounts of zinc, lead and copper in sugar cane spirits by anodic stripping voltammetry, *Analyst*, 120 (1995) 1647-1650.
133. Herrero A., Ortiz M.C., Arcos M.J. y López-Palacios J., Optimization of the experimental parameters in the determination of copper(II) by differential-pulse anodic stripping voltammetry and evaluation of the characteristic detection curves, *Analyst*, 119 (1994) 1585-1591.
134. Jagner D. y Westerlund S., Determination of lead, copper, and cadmium in wine and beer by potentiometric stripping analysis, *Analytica Chimica Acta*, 117 (1980) 159-164.
135. Schulze G. y Frenzel W., Störe in durch kupfer bei der potentiometrischen stripping-analyse von zink und cadmium, *Fresenius Z. Analytical Chemistry*, 314 (1983) 459-462.
136. Wise J.A., Roston D.A. y Heineman W.R., The effects of copper-zinc and copper-cadmium intermetallic compounds in different systems used for anodic stripping voltammetry, *Analytica Chimica Acta*, 154 (1983) 95-104.
137. Batley G.E. y Florence T.M., An evaluation and comparison of some techniques of anodic stripping voltammetry, *Journal of Electroanalytical Chemistry*, 55 (1974) 23-43.
138. Wang J., *Stripping Analysis. Principles, instrumentation and applications*. VCH Publishers, Inc., Deerfield Beach, FL, 1985.
139. Adeloju S.B., Sahara E. y Jagner D., Anodic-stripping potentiometric determination of Cu, Pb, Cd and Zn in natural-waters on a novel combined electrode system, *Analytical Letters*, 29 (1996) 283-302.
140. Izquierdo A., Luque de Castro M.D. y Valcárcel M., Simultaneous determination of copper, lead, cadmium and zinc in previously lyophilized biological tissues using

-
- flow-injection analysis/anodic stripping voltammetry, *Electroanalysis*, 6 (1994) 894-902.
141. Neiman E.Y., Petrova L.G., Ignatov V.I. y Dolgopolova G.M., The third element effect in anodic stripping voltammetry, *Analytica Chimica Acta*, 113 (1980) 277-285.
 142. De Angelis T.P., Bond R.E., Brooks E.E. y Heineman W.R., Thin-layer differential pulse voltammetry, *Analytical Chemistry*, 49 (1977) 1792-1997.
 143. Roston D.A., Brooks E.E. y Heineman W.R., Elimination of intermetallic compound interferences in twin-electrode thin-layer anodic stripping voltammetry, *Analytical Chemistry*, 51 (1979) 1728-1732.
 144. Aldstadt J.H. y Dewald H.D., Determination of heavy metals by thin-layer chromatography-square-wave anodic stripping voltammetry, *Analytical Chemistry*, 64 (1992) 3176-3179.
 145. Wen J. y Cassidy R.M., Anodic and cathodic pulse amperometric detection of metal ions separated by capillary electrophoresis, *Analytical Chemistry*, 68 (1996) 1047-1053.
 146. Jagner D. y Kryger L., Computerized electroanalysis. Part III. Multiple scanning anodic stripping and its application to sea water, *Analytica Chimica Acta*, 80 (1975) 255-266.
 147. Gerlach R.W. y Kowalski B.R., The generalized standard addition method: intermetallic interferences in anodic stripping voltammetry, *Analytica Chimica Acta*, 134 (1982) 119-127.
 148. Jagner D., Renman L. y Stefansdottir S.H., Determination of arsenic by stripping potentiometry on gold electrodes using partial least squares (PLS) regression calibration, *Electroanalysis*, 6 (1994) 201-208.
 149. Bruno P., Caselli M., Di Fano A. y Fragale C., Simultaneous determination of copper, lead and zinc in wine by differential pulse polarography, *Analyst*, 103 (1978) 868-971.
 150. Oehme M. y Lund W., Determination of cadmium, lead and copper in wine by differential pulse anodic stripping voltammetry, *Fresenius Z. Analytical Chemistry*, 294 (1979) 391-397.
 151. Abdullah M.I., Berg B.R. y Klimek R., The determination of zinc, cadmium, lead and copper in a single sea-water sample by differential pulse anodic stripping voltammetry, *Analytica Chimica Acta*, 84 (1976) 307-317.
 152. Florence, T.M. y Farrar Y.J., Determination of tin by thin film anodic stripping voltammetry. Application to marine samples, *Journal of Electroanalytical Chemistry*, 51 (1974) 191-200.

153. Kolthoff I.M. y Lingane J.J., *Polarography*, Interscience Publishers, Inc. New York, 2 nd. (1952).
154. Bond A.M., Direct current, alternating current, rapid, and inverse polarographic methods for determination of tin(IV), *Analytical Chemistry*, 42 (1970) 1165-1168.
155. Nghi T.V. y Vydra F., The electrochemical stripping determination of tin in aqueous and nonaqueous media after its extraction, *Journal of Electroanalytical Chemistry*, 71 (1976) 333-340.
156. Mannino S., Potentiometric stripping analysis of lead and tin with a continuous flow system, *Analyst*, 109 (1984) 905-907.
157. Hu A., Dessy R.E. y Granéli A., Potentiometric stripping with matrix exchange techniques in flow injection analysis of heavy metals in groundwaters, *Analytical Chemistry*, 55 (1983) 320-328.
158. Desimoni E., Palmisano F. y Sabbatini L., Simultaneous Determination of tin and lead at the parts-per-billion level by coupling differential pulse anodic stripping voltammetry with matrix exchange method, *Analytical Chemistry*, 52 (1980) 1889-1892.
159. Glodowski S. y Kublik Z., Cyclic and stripping voltammetry of tin in the presence of lead in pyrogallol medium at hanging and film mercury electrodes, *Analytica Chimica Acta*, 104 (1979) 55-65.
160. Maxwell T.J. y Smyth W.F., A study of the stripping voltammetric behavior of selected metal chelates and its application to automated analysis of river waters, *Electroanalysis*, 8 (1996) 795-802.
161. Helms I., Pietsch A. y Scholz F., Electrochemistry and analytical application of the tin(IV) complex of N-benzoyl-N-phenylhydroxamic acid in 2-propanol solution, *Electroanalysis*, 7 (1995) 230-236.
162. Georges J. y Mermet M., Simultaneous determination of tin and lead by anodic stripping voltammetry with addition of surfactant, *Analytica Chimica Acta*, 185 (1986) 363-368.
163. Hofbauerová H., Labuda J., Fišera M. y Vaničková M., Simultaneous voltammetric determination of antimony, copper, lead and tin in steel using a hydrochloric acid-isopropanolic medium, *Electroanalysis*, 8 (1995) 788-790.
164. Mannino S., Simultaneous determination of lead and tin in fruit juices and soft drinks by potentiometric stripping analysis, *Analyst* 108 (1983) 1257-1260.
165. Glodowski S. y Kublik Z., Cyclic and stripping voltammetry of tin at mercury hanging drop and film electrodes in acidic o-diphenol media in the presence of lead and cadmium, *Analytica Chimica Acta*, 115 (1980) 51-60.

166. Shafiqul Alam A.M., Vittori O., Tersigni H. y Courtois G., Dosage simultané du plomb et de l'étain par polarographie à tension-alternative surimposée. Application aux commerciales étain plomb à ame décapante, *Analytica Chimica Acta*, 91 (1977) 325-329.
167. Adeloju S.B.O. y Pablo F., Simultaneous determination of lead and tin in biological and environmental materials by differential pulse cathodic voltammetry on a glassy carbon mercury film electrode, *Electroanalysis*, 8 (1995) 750-755.
168. Wang J. y Zadeii J., Ultrasensitive and selective measurements of tin by adsorptive stripping voltammetry of the tin-tropolone complex, *Talanta*, 34 (1987) 909-914.
169. Kiekens H., Verplaetse L., De Cock L. y Temmerman E., Determination of tin in the presence of lead by stripping voltammetry with collection at a rotating ring-disc electrode, *Analyst*, 106 (1981) 305-310.
170. Bond A.M. y Grabaric B.S., Simple approach to the problem of overlapping waves using a microprocessor controled polarography, *Analytical Chemistry*, 48 (1976) 1624-1628.
171. Bauer K.H. y Neen R., Inversvoltammetrische bestimmung von bismut in biomatrices mit computerunterstutzter meßsignalwertung, *Fresenius' Journal Analytical Chemistry*, 330 (1988) 17-23.
172. Berzas J.J. y Rodríguez J., Use of data treatment software for the resolution of overlapping differential pulse polarographic peaks, *Fresenius' Journal Analytical Chemistry*, 342 (1992) 273-275.
173. Osteryoung J., O'Dea J.J. y Bard A.J. (Editor), *Electroanalytical Chemistry*, volumen 14, Marcel Dekker, New York, 1986, pp. 209-308.
174. Boudreau P.A. y Perone S.P., Quantitative resolution of overlapped peaks in programmed potential-step voltammetry, *Analytical Chemistry*, 51 (1979) 811-817.
175. Locatelli C., Fagioli F. y Garai T., Peak resolution in the determination of cobalt and nickel by differential pulse and alternating current adsorption voltammetry, *Analytical Chemistry*, 63 (1991) 1409-1413.
176. Locatelli C., Fagioli F., Bigli C. y Garai Tibor, Determination of metals by second-harmonic alternating-current voltammetry with a semi-stationary mercury electrode, *Talanta*, 32 (1985) 539-543.
177. Pizeta I., Deconvolution of non-resolved voltammetric signals, *Analytica Chimica Acta*, 285 (1994) 95-102.
178. Raspor B., Pizeta I. y Branica M., Comparative quantitative analysis of overlapping voltammetric signals, *Analytica Chimica Acta*, 285 (1994) 103-111.

179. Brown T.F. y Brown S.D., Resolution of overlapped electrochemical peaks with the use of the Kalman filter, *Analytical Chemistry*, 53 (1981) 1410-1417.
180. Scolari C.A. y Brown S.D., Resolution of strongly overlapped responses in square-wave voltammetry by using the Kalman filter, *Analytica Chimica Acta*, 166 (1984) 253-260.
181. Binkley D.P. y Dessy R.E., Linear parameter estimation of fused peak system in the spatial frequency domain, *Analytical Chemistry*, 52 (1980) 1335-1344.
182. García J.M., Martín E., Jiménez A.I., Jiménez F. y Arias J.J., Application of the multiple linear regression method to differential pulse polarography, *Analytical Letters*, 27 (1994) 1763- (falta el fin de página)
183. Turnes G., Cladera A., Gómez E., Estela J.M. y Cerdá V., Resolution of highly overlapped differential pulse polarographic and anodic stripping voltammetric peaks by multicomponent analysis, *Journal of Electroanalytical Chemistry*, 338 (1992) 49-60.
184. Ni N.Y., Kokot S., Selby M. y Hodgkinson M., Simultaneous polarographic chemometric trace analysis of pyrazine and its methyl derivatives, *Electroanalysis*, 4 (1992) 713-718.
185. Eisner U. y Ariel M., Trace analysis by anodic stripping voltammetry. III. Trace metal determination in biological media. I. The determination of thallium in urine, *Journal of Electroanalytical Chemistry*, 11 (1966) 26-30.
186. Cai Q. y Khoo S.B., Differential pulse stripping voltammetric determination of thallium with an 8-hydroxyquinoline-modified carbon paste electrode, *Electroanalysis*, 7 (1995) 379-385.
187. Karbainov Y.A. y Stromberg A.G., *Izv. Tomsk. Politekhn. Inst.*, 151 (1966) 7. (pedir)
188. Andrews R.W. y Johnson D.C., Determination of selenium(IV) by anodic stripping voltammetry in flow system with ion exchange separation, *Analytical Chemistry*, 48 (1976) 1056-1060.
189. Bonelli J.E., Taylor H.E. y Skogerboe R.K., A direct differential pulse anodic stripping voltammetric method for the determination of thallium in natural waters, *Analytica Chimica Acta*, 118 (1980) 243-256.
190. Allus M.A. y Brereton R.G., Determination of thallium in cement dust and sediment samples by differential-pulse anodic stripping voltammetry: a chemometric approach to linear calibration, *Analyst*, 117 (1992) 1075-1084.
191. Lukaszewski Z., Zembruski W. y Piela A., Direct determination of ultratraces of thallium in water by flow-injection-differential-pulse anodic stripping voltammetry, *Analytica Chimica Acta*, 318 (1996) 159-165.

192. Anderson L, Jagner D. y Josefson M., Potentiometric stripping analysis in flow cells, *Analytical Chemistry*, 54 (1982) 1371-1376.
193. Neeb R. y Kiehnast I.Z., Der einfluß von salzen auf die anodische sptzenhöhe in der inversen voltammetrie, *Fresenius Z. Analytical Chemistry*, 226 (1967) 153-159.
194. Lukaszewski Z., Pawlak M.K. y Cizewski A., Determination of thallium and lead in cadmium salts by anodic stripping voltammetry with addition of surfactants to suppress the cadmium peaks, *Talanta*, 27 (1980) 181-185.
195. Wang J., Farias P.A.M. y Luo D.B., Improved resolution in stripping analysis using the formation of intermetallic compounds, *Analytical Chemistry*, 56 (1984) 2379-2382.
196. Garai T., Nagy Z., Mészáros L., Bartalits L., Locatelli C. y Fagioli F., Resolution of differential potentiometric stripping analysis, *Electroanalysis*, 8 (1996) 381-386.
197. Collado-Sánchez C., Pérez-Peña J., Gelado-Caballero M.D., Herrera-Melian J.A. y Hernández-Brito J.J., Rapid determination of copper, lead and cadmium in unpurged seawater by adsorptive stripping voltammetry, *Analytica Chimica Acta*, 320 (1996) 19-30.
198. Ortiz M.C., Arcos M.J. y Sarabia L.A., Using continuum regression for quantitative analysis with overlapping signals obtained by differential pulse polarography, *Chemometrics and Intelligent Laboratory System*, en prensa.
199. Sturgeon R.E., Liu J., Boyko V.J. y Luong V.T., Determination of copper in environmental matrices following vapor generation, *Analytical Chemistry*, 68 (1996) 1883-1887.
200. Cabrera C., Lorenzo M.L. y López M.C., Electrothermal atomic-absorption spetrometric determination of cadmium, copper, iron, lead and selenium in fruit slurry. Analytical application to nutritional and toxicological quality-control, *Journal of AOAC International*, 78 (1995) 1061-1067.
201. Ryan E. y Meaney M., Determination of trace levels of copper(II), aluminium(III) and iron(III) by reversed-phase high-performance liquid chromatography using a novel on-line sample preconcentration technique, *Analyst*, 117 (1992) 1435-1439.
202. Kabil L.A., Elkourashy A.G. y Elhagrasy M.A., Flame atomic-absorption spectrometric investigation and determination of cobalt and copper using ethanolamine and triethanolamine as chemical modifiers, *Journal of Analytical Atomic Spectrometry*, 11 (1996) 379-387.
203. Martínez L., Baucells M., Pelfort E. y Roura M., Removal of iron interference in the determination of selenium in geological samples by continuous-flow hydride generation ICP-AES, *Fresenius Journal of Analytical Chemistry*, 354 (1996) 126-127.

204. Quentel F., Elleouet C. y Madec Ch., Determination of copper in seawater by adsorptive voltammetry with 1,2-dihydroxyanthraquinone-3-sulfonic acid, *Electroanalysis*, 6 (1994) 683-688.
205. Meenakumari V., Extraction and voltammetric determination of copper, lead, cadmium and zinc in iron ore samples, *Analyst*, 120 (1995) 2849-2851.
206. Lu J., Wang J. y Yarnitzky C., Ultrasensitive adsorptive-catalytic stripping voltammetry of iron in the presence of hydroxamic acid and hydrogen peroxide, *Electroanalysis*, 7 (1995) 79-82.
207. Vázquez Díaz M.E., Jiménez Sánchez J.C., Callejón Mochón M. y Guiraúm Pérez A., Differential-pulse polarographic determination of iron in acids, waters, fruit juices and wines, *Analyst*, 119 (1994) 1571-1574.
208. Vandecasteele C. y Block C.B., *Modern methods for trace element determination*, Wiley, Chichester, 1995.
209. Gottesfeld S. y Ariel M., Trace analysis by anodic stripping voltammetry. IV. The determination of copper in steels, *Journal of Electroanalytical Chemistry*, 9 (1965) 112-116.
210. Bonelli J.E., Skogerboe R.K. y Taylor H.E., Iron interference on the ultratrace determination of copper in natural waters by differential pulse anodic stripping voltammetry, *Analytica Chimica Acta*, 101 (1978) 437-440.
211. Battistini L. y López Palacios J., Multistep mechanism in the Electrochemical Oxidation-Reduction of Fe-citrate complexes, *Analytical Chemistry*, 66 (1994) 2005-2009.
212. Ismail M.I. y Etaiw S.H., Polarographic analytical determination of Hg(II), Cu(II), Cd(II), As(III), Sb(III), Ti(IV) and U(VI) in the presence of Fe(III), *Journal of Electroanalytical Chemistry*, 108 (1980) 229-238.
213. Pratt K.W. y William F.K., Determination of Trace-level chromium(VI) in the presence of chromium(III) and iron(III) by flow injection amperometry, *Analytical Chemistry*, 58 (1986) 124-127.
214. Tuzhi P., Zhe T., Goushum W. y Baoen S., Differential pulse voltammetric determination of lead (II) with benzoin oxime-modified carbon paste electrodes, *Electroanalysis*, 6 (1994) 597-603.
215. GPES: General Purpose Electrochemical System for Windows, ver. 4.3, Echo Chemie B.V., Utrecht, 1996.
216. Lorber A. y Kowalski B.R., The effect of interferences and calibration design on accuracy: implications for sensor and samples selection, *Journal of Chemometrics*, 2 (1988) 67-79.

217. Brown C.W., Lynch P.F., Obremski R.J. y Lavery D.S., Matrix representations and criteria for selecting analytical wavelengths for multicomponent spectroscopic analysis, *Analytical Chemistry*, 54 (1982) 1472-1479.
218. Phan-Tan-Luu R., Designs for screening, Conferencia plenaria en el Congreso Internacional: VI Chemometrics in Analytical Chemistry, Tarragona, 1996.
219. Garrido Frenich A., Jouan-Rimbaud D., Massart D.L., Kuttatharmmakul S., Martínez Galera M. y Martínez Vidal J.L., Wavelength selection method for multicomponent spectrophotometric determinations using partial least squares, *Analyst*, 120 (1995) 2787-2792.
220. Kalivas J.H., Optimization using variations of simulated annealing, *Chemometrics and Intelligent Laboratory Systems*, 15 (1992) 1-12.
221. Kalivas J.H., Roberts N. y Sutter J.M., Global optimization by simulated annealing with wavelength selection for ultraviolet-visible spectrophotometry, *Analytical Chemistry*, 61 (1989) 2024-2030.
222. Ortiz M.C., Herrero A., Sánchez M.S., Sarabia L.A. e Íñiguez M., Modelling the relation between CieLab parameters and sensory scores for quality control of red-wine colour, *Analyst*, 120 (1995) 2793-2798.
223. Leardi R., Boggia R. y Terrile M., Genetic algorithms as a strategy for feature selection, *Journal of Chemometrics*, 6 (1992) 267-281.
224. Leardi R., Application of a genetic algorithm to feature selection under full validation conditions and to outlier detection, *Journal of Chemometrics*, 8 (1994) 65-79.
225. Jouan-Rimbaud D., Massart D.L. y Leardi R., Genetic algorithms as a tool for wavelength selection in multivariate calibration, *Analytical Chemistry*, 67 (1995) 4295-4301.
226. Kubinyi H., Evolutionary variable selection in regression and PLS analyses, *Journal of Chemometrics*, 10 (1996) 119-133.
227. Goldberg D.E., Genetic algorithm in search optimization and machine learning, Addison-Wesley Publishing, MA, 1989.
228. Lucasius C.B. y Kateman G., Understanding and using genetic algorithms. Part 1. Concepts, properties and context, *Chemometrics and Intelligent Laboratory Systems*, 19 (1993) 1-33.
229. Lucasius C.B. y Kateman G., Understanding and using genetic algorithms. Part 2. Representation, configuration and hybridization, *Chemometrics and Intelligent Laboratory Systems*, 25 (1994) 99-145.
230. Hibbert D.B., Genetic algorithms in chemistry, *Chemometrics and Intelligent Laboratory Systems*, 19 (1993) 277-293.

231. Davis L., Handbook of Genetic Algorithms, Van Nostrand Reinhold, New York, 1991.
232. Riu J. y Rius X., Assessing the accuracy of analytical methods using linear regression with errors in both axes, Analytical Chemistry, 68 (1996) 1851-1857.
233. Rousseeuw P.J. y Leroy A.M., Robust regression & outlier detection, Wiley, Chichester, 1987.
234. Sarabia L.A., Ortiz M.C. y Tomás X., Orthogonal least median squares regression, OLEM, Comunicación al Congreso Internacional: VI Chemometrics in Analytical Chemistry, Tarragona, 1996.

VII. APÉNDICES

APÉNDICE 1

Progama 'PASOV.EXE'

REM Programa para transformar archivos de VESUV 10 (*.AMP) al
REM formato ASCii, por filas o columnas

```
CLS
INPUT "Archivo de entrada (con path)      "; ARCH1$
INPUT "Archivo de salida (con path)      "; ARCH2$
INPUT "Por filas (1) o por columnas (2)  "; tipo%

PRINT
PRINT
PRINT "Espere un momento"

i0 = 0
volt% = 0

OPEN ARCH1$ FOR INPUT AS #3

10 i0 = i0 + 1

IF EOF(3) THEN 20

LINE INPUT #3, A$
FOR i = 1 TO LEN(A$)
    B$ = MID$(A$, i, 1)
    IF B$ = "Y" THEN
        volt% = volt% + 1
        GOTO 10
    END IF
NEXT i
GOTO 10
```

```

20 CLOSE
i0 = i0 - 1

DIM A$(i0)
DIM dat$(i0, 12)
DIM dat(i0, 5)

OPEN ARCH1$ FOR INPUT AS #3
FOR i = 1 TO i0
LINE INPUT #3, A$(i)
NEXT i
CLOSE

obj% = 0

FOR i = 1 TO i0
FOR l = 1 TO LEN(A$(i))

22 B$ = MID$(A$(i), l, 1)
IF B$ = "Y" OR B$ = "" THEN ' salta lineas cabecera y en blanco
    i = i + 1
    GOTO 22
END IF
NEXT l

obj% = obj% + 1
k = 0
j = 1

30 A$ = ""

40 IF MID$(A$(i), j, 1) <> " " THEN
    A$ = A$ + MID$(A$(i), j, 1)
    j = j + 1
    dat$(obj%, k) = A$
    GOTO 40
END IF
IF j = LEN(A$(i)) THEN 50
IF MID$(A$(i), j, 1) = " " AND MID$(A$(i), j+1, 1) <> " " THEN
    k = k + 1

END IF
j = j + 1
GOTO 30

```

```
50 NEXT i

FOR i = 1 TO obj%
  k0 = 0
  FOR k = 4 TO 12.1 STEP 2
    k0 = k0 + 1
    IF dat$(i, k) = "pA" THEN dat(i, k0) = VAL(dat$(i, k - 1)) / 1000
    IF dat$(i, k) = "nA" THEN dat(i, k0) = VAL(dat$(i, k - 1))
    IF dat$(i, k) = "uA" THEN dat(i, k0) = VAL(dat$(i, k - 1)) * 1000
    IF dat$(i, k) = "mA" THEN dat(i, k0) = VAL(dat$(i, k - 1)) *
1000000
  NEXT k
NEXT i

ON tipo% GOSUB 1000, 2000
PRINT "He acabado"
END

1000 OPEN ARCH2$ FOR OUTPUT AS #3
coc% = obj% / volt%
A% = 1
B% = coc%

90 FOR i = A% TO B%
  FOR k = 1 TO 5
    PRINT #3, USING "##.#####^ ^ ^"; dat(i, k);
    PRINT #3, " ";
  NEXT k
NEXT i
PRINT #3,
A% = A% + coc%
B% = B% + coc%
IF B% > obj% THEN
  GOTO 100
ELSE
  GOTO 90
END IF

100 CLOSE

RETURN
```

```
2000 OPEN ARCH2$ FOR OUTPUT AS #3
      coc% = obj% / volt%
      FOR k = 1 TO coc%
        FOR l = 1 TO 5
          FOR i = 0 TO volt% - 1
            PRINT #3, USING "##.#####^"^"; dat(k + i * coc%, l);
            PRINT #3, "  ";
          NEXT i
          PRINT #3,
        NEXT l
      NEXT k

      CLOSE

RETURN
```

APÉNDICE 2

Función '*transfpds.mat*'

```
[F, Ccal, SEP, SET] = transfpds (P, S, w, no, Pcer, C, t, p)

% [F, Ccal, SEP, SET] = transfpds (P, S, w, no, Pcer, C, t, p)
%
% Función para realizar transferencias de calibrado con el método PDS
%
% Parámetros de entrada:
%
% P   matriz con las señales del calibrado patrón (objetos por filas)
% S   matriz con las señales del calibrado satélite (objetos por
%     filas)
% w   número de variables predictoras a considerar en la ventana de
%     transferencia
% no  puede ser un único escalar que indique el número de objetos del
%     conjunto de transferencia (las muestras se elegirán por su
%     score) o un vector que señale los índices de esos objetos
% Pcer fórmula cerrada del modelo PLS del calibrado patrón
% C   matriz de concentraciones reales
% t   vector con los índices de las muestras del calibrado patrón que
%     constituyen el conjunto test
% p   vector con los índices de las muestras del calibrado patrón que
%     constituyen el conjunto de aprendizaje
%
% Parámetros de salida:
%
% F   matriz que transforma las señales del calibrado satélite a
%     señales del calibrado patrón
% Ccal matriz de concentraciones calculadas para el calibrado satélite
% SEP error estándar en predicción
% SET error estándar del conjunto de aprendizaje
```

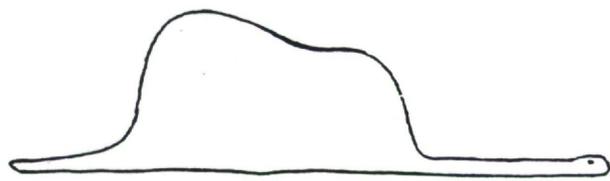
```
[F, subset] = pdsgen (P', S', w, no);
Scorr = S' * F;
Ccal = Scorr * Pcer';
sep=((sum((C(t,:)-Ccal(t,:)).^2))./length(t)).^(1/2);
set=((sum((C(p,:)-Ccal(p,:)).^2))./length(p)).^(1/2);
```

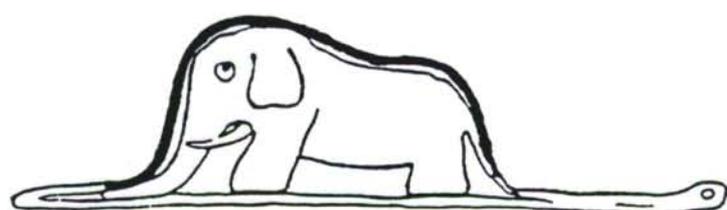
Función 'transfgct.mat'

```
[Scerest, Ccal, SEP, SET] = transfpds (P, S, w, no, Pcer, C, t, p)

% [Scerest, Ccal, SEP, SET] = transfpds (P, S, w, no, Pcer, C, t, p)
%
% Función para realizar transferencias de calibrado con el método GCT
%
% Parámetros de entrada:
%
% P    matriz con las señales del calibrado patrón (objetos por filas)
% S    matriz con las señales del calibrado satélite (objetos por
%      filas)
% no   puede ser un único escalar que indique el número de objetos del
%      conjunto de transferencia (las muestras se elegirán por su
%      score) o un vector que señale los índices de esos objetos
% Pcer fórmula cerrada del modelo PLS del calibrado patrón
% C    matriz de concentraciones reales
% t    vector con los índices de las muestras del calibrado patrón que
%      constituyen el conjunto test
% p    vector con los índices de las muestras del calibrado patrón que
%      constituyen el conjunto de aprendizaje
%
% Parámetros de salida:
%
% Scerest fórmula cerrada estimada para el calibrado satélite
% Ccal matriz de concentraciones calculadas para el calibrado satélite
% SEP error estándar en predicción
% SET error estándar del conjunto de aprendizaje
```

```
X = S(no, :);
Y = P(no, :);
[press, cumpress, minlv, b] = plscv (X, Y, split, lvm, 1);
[O, Q, W, T, U, b ssqdif] = pls (X, Y, minlv);
F = conpred1(b, W, O, Q, minlv);
Scerest = F * Pcer;
Ccal = C * Scerest;
sep=((sum((C(t,:)-Ccal(t,:)).^2))./length(t)).^(1/2);
set=((sum((C(p,:)-Ccal(p,:)).^2))./length(p)).^(1/2);
```





El Principito

Antoine de Saint-Exupéry