



**UNIVERSIDAD AUTÓNOMA DEL ESTADO
DE HIDALGO**



INSTITUTO DE CIENCIAS BÁSICAS E INGENIERÍA

ÁREA ACADÉMICA DE QUÍMICA

**“APLICACIÓN DE MODELOS DE REGRESIÓN MULTIVARIANTE EN
EL DESARROLLO DE MÉTODOS ELECTROQUÍMICOS DE ANÁLISIS
MULTIPARÁMETRO”**

TESIS

**PARA OBTENER EL TÍTULO DE:
MAESTRO EN QUÍMICA**

PRESENTA:

L.Q. EDGAR GONZÁLEZ LÓPEZ

DIRECTOR

DR. CARLOS ANDRÉS GALÁN VIDAL

MINERAL DE LA REFORMA, HGO. JUNIO 2018



UNIVERSIDAD AUTÓNOMA DEL ESTADO DE HIDALGO
Instituto de Ciencias Básicas e Ingeniería
Institute of Basic Sciences and Engineering
Dirección
Dean

Mineral de la Reforma, Hgo., a 18 de junio de 2018.

Número de control: ICBI-D/580/2018.

Asunto: Permiso de impresión de Tesis

M. EN C. JULIO CESÁR LÉINES MEDÉCIGO
DIRECTOR DE ADMINISTRACIÓN ESCOLAR DE LA UAEH

Por este conducto le comunico que el jurado asignado al candidato a Maestro en Química, **Q. Edgar González López**, quien presenta el trabajo "**Aplicación de modelos de regresión multivariante en el desarrollo de métodos electroquímicos de análisis multiparámetro**" después de revisar el trabajo en reunión de sinodales, ha decidido **autorizar la impresión** del mismo, una vez realizadas la correcciones que fueron acordadas.

A continuación se anotan las firmas de conformidad de los siguientes integrantes del jurado:

PRESIDENTE: Dra. María Elena Páez Hernández

SECRETARIO: Dra. María Gabriela Vargas Martínez

VOCAL: Dr. Carlos Andrés Galán Vidal

SUPLENTE: Dr. José Antonio Rodríguez Ávila

Sin otro particular, reitero a Usted la seguridad de mi atenta consideración.

ATENTAMENTE
"Amor, Orden y Progreso"

Dr. Óscar Rodolfo Suárez Castillo
Director del ICBI

ORSC/LHHH

Ciudad del Conocimiento
Carretera Pachuca - Tulancingo km. 4.5
Colonia Carboneras
Mineral de la Reforma, Hidalgo, México, C.P. 42184
Tel. +52 771 7172000 exts 2231, Fax 2109
direccion_icbi@uaeh.edu.mx



www.uaeh.edu.mx

Aplicación de Modelos de Regresión Multivariante en el Desarrollo de Métodos Electroquímicos de Análisis Multiparámetro

Edgar González López



Este proyecto de investigación se realizó en el laboratorio 1 del Área Académica de Química de la Universidad Autónoma del Estado de Hidalgo (UAEH).

La investigación se realizó gracias al financiamiento del Consejo Nacional de Ciencia y Tecnología (CONACYT) como parte del proyecto: **DESARROLLO DE UNA LENGUA ELECTRÓNICA PARA LA CUANTIFICACIÓN SIMULTÁNEA DE GLUCOSA, COLESTEROL Y TRIGLICÉRIDOS**; Clave CB-2014-242677.

L.Q. Edgar González López agradece al CONACYT la beca de proyecto de investigación 596467.

Parte de este trabajo se ha presentado en:

- 51° Congreso Mexicano de Química y 35° Congreso Nacional de Educación Química. “Simulación de respuesta de biosensores multienzimáticos para el análisis multiparamétrico con un solo dispositivo”.
- VI Congreso Iberoamericano de Química Analítica y Encuentro Nacional de Química Ambiental. “Determinación simultánea de paracetamol y naproxeno en fármacos mediante voltamperometría y regresión multivariante”.

Dedicatoria

A quienes hicieron posible este trabajo

PRÓLOGO

La investigación hace posible generar conocimiento, desarrollar tecnología y contribuir al bienestar social mediante la solución de problemas de interés común. El esfuerzo implicado en el ejercicio de esta actividad, generalmente es perceptible hasta conseguir algún logro importante.

Las metas de un científico con relación a un proyecto de investigación no siempre son alcanzadas en el corto plazo debido a la falta de materiales, de conocimientos, de tiempo, o a circunstancias ajenas a su voluntad y control. Como consecuencia, los ajustes y los replanteamientos son necesarios y tienen el poder de cambiar el propósito original de todo el trabajo. Este hecho no es malo en sí mismo porque no significa forzosamente un retroceso en la aportación de conocimientos útiles y de alto impacto para la comunidad científica, más bien, se crean oportunidades de explorar alternativas que generalmente conducen a resultados inesperados e interesantes.

El producto de la investigación registrado en estas páginas es el resultado de una constante documentación bibliográfica, de un proceso continuo de aprendizaje y reflexión que fue concluido en una etapa aún temprana del proyecto. El conocimiento y la experiencia adquirida son recompensas valiosas e intangibles a todo el esfuerzo invertido. En este caso destaca la publicación y exposición de algunos trabajos en congresos nacionales e internacionales que han sido importantes en la estructuración de esta tesis.

Esta rama de la ciencia tiene el potencial de conseguir avances significativos en el desarrollo de dispositivos de interés biomédico en beneficio de la salud humana. Quien desee continuar en esta línea de investigación tiene la oportunidad de contribuir al cumplimiento de las expectativas y objetivos planteados originalmente, con la certidumbre de que el pronóstico para los resultados venideros es favorable.

Edgar González López

ÍNDICE TEMÁTICO

SÍMBOLOS Y ABREVIATURAS.....	1
RESUMEN.....	3
1. SENSORES ELECTROQUÍMICOS Y SUS APLICACIONES	5
1.1. INTRODUCCIÓN	6
1.2. ANÁLISIS VOLTAMPEROMÉTRICOS Y CRONOAMPEROMÉTRICOS	6
1.3. SENSORES Y BIOSENSORES ELECTROQUÍMICOS	7
1.4. BIOSENSORES.....	7
1.4.1. Transductores	8
1.4.2. Elementos de reconocimiento de origen biológico.....	9
1.4.3. Métodos de inmovilización	9
1.4.4. Aplicación de los biosensores enzimáticos	10
1.5. IMPORTANCIA BIOMÉDICA DE LOS BIOSENSORES ENZIMÁTICOS	10
1.5.1. Determinación clínica de glucosa.....	10
1.5.2. Biosensores enzimáticos de glucosa	11
1.5.3. Biosensores enzimáticos de colesterol.....	13
1.5.4. Biosensores enzimáticos de triglicéridos	14
1.6. ANÁLISIS ELECTROQUÍMICOS SIMULTÁNEOS	16
2. REGRESIÓN MULTIVARIANTE.....	17
2.1. REGRESIÓN LINEAL SIMPLE.....	18
2.2. ANÁLISIS MULTIVARIANTE	18
2.3. REGRESIÓN LINEAL MÚLTIPLE (RLM).....	19
2.3.1. Validación del modelo de RLM.....	21
2.3.2. Coeficiente de determinación múltiple	22
2.3.3. Multicolinealidad.....	22
2.3.4. Coeficiente de determinación ajustado	23
2.3.5. Coeficiente de determinación predicho	24
2.4. SELECCIÓN DE LAS VARIABLES DE RESPUESTA	24
2.4.1. Método paso a paso (stepwise)	25
2.4.2. Consideraciones importantes sobre el método stepwise	25
2.5. REGRESIÓN SOBRE COMPONENTES PRINCIPALES (RCP)	26

2.5.1.	Cálculo de los componentes principales	26
2.5.2.	Transformación de las variables predictoras mediante componentes principales.....	29
2.6.	REGRESIÓN SOBRE MÍNIMOS CUADRADOS PARCIALES (MCP)	30
2.6.1.	Algoritmo de la regresión MCP	30
2.7.	PARÁMETROS ANALÍTICOS DE LA REGRESIÓN LINEAL MÚLTIPLE.....	32
2.8.	CONSTRUCCIÓN DE LA MATRIZ DE RESPUESTA.....	33
2.9.	OBJETIVOS	34
2.9.1.	Objetivo General	34
2.9.2.	Objetivos específicos	34
3.	APLICACIÓN DE LA REGRESIÓN MULTIVARIANTE EN SISTEMAS ELECTROQUÍMICOS	35
3.1.	DETERMINACIÓN SIMULTÁNEA DE NAPROXENO Y PARACETAMOL EN MEDICAMENTOS MEDIANTE VOLTAMPEROMETRÍA Y REGRESIÓN MULTIVARIANTE	36
3.1.1.	Introducción.....	36
3.1.2.	Experimentación.....	37
3.1.3.	Resultados y discusión	37
3.1.3.1.	Regresión lineal múltiple (RLM)	40
3.1.3.2.	Multicolinealidad del modelo.....	45
3.1.3.3.	Regresión sobre componentes principales	45
3.1.3.4.	Regresión sobre mínimos cuadrados parciales	49
3.1.3.5.	Resumen de resultados	54
3.1.3.6.	Análisis de la muestra.....	55
3.1.4.	Conclusiones	57
3.2.	DETERMINACIÓN SIMULTÁNEA DE HIERRO (II) Y ÁCIDO ASCÓRBICO EN MEDICAMENTOS POR VOLTAMPEROMETRÍA Y REGRESIÓN MULTIVARIANTE	57
3.2.1.	Introducción.....	57
3.2.2.	Experimentación.....	58
3.2.3.	Resultados y discusión	58
3.2.3.1.	RLM-Stepwise	61
3.2.3.2.	Regresión sobre componentes principales	63
3.2.3.3.	Regresión sobre mínimos cuadrados parciales	65

3.2.3.4.	Resumen de resultados	65
3.2.3.5.	Muestra problema.....	66
3.2.4.	Conclusiones	67
4.	DESARROLLO DE BIOSENSORES MUTIPARÁMETROS	69
4.1.	SIMULACIÓN DE RESPUESTA DE UN BIOSENSOR ENZIMÁTICO PARA LA CUANTIFICACIÓN SIMULTÁNEA DE GLUCOSA, COLESTEROL Y TRIGLICÉRIDOS.....	70
4.1.1.	Introducción.....	70
4.1.2.	Materiales y métodos	70
4.1.3.	Resultados y discusión	71
4.1.3.1.	Regresión lineal múltiple	73
	Regresión sobre componentes principales	75
4.1.3.4.	Regresión sobre mínimos cuadrados parciales	76
4.1.3.5.	Resumen de resultados	76
4.1.4.	Conclusiones	77
4.2.	DESARROLLO DE BIOSENSORES ENZIMÁTICOS PARA LA DETERMINACIÓN SIMULTÁNEA DE GLUCOSA Y ÁCIDO ASCÓRBICO MEDIANTE VOLTAMPEROMETRÍA Y REGRESIÓN MÚLTIVARIANTE EN TABLETAS EFERVESCENTES	77
4.2.1.	Introducción.....	77
4.2.2.	Reactivos y soluciones	79
4.2.2.1.	Construcción de biosensores	79
4.2.2.2.	Instrumentación y software	80
4.2.3.	Resultados y discusión	80
4.2.3.1.	Stepwise RLM	81
4.2.3.2.	Regresión sobre componentes principales	83
4.2.3.3.	Regresión sobre mínimos cuadrados parciales	84
4.2.4.	Resumen de resultados.....	85
4.2.5.	Muestra problema	85
4.2.6.	Conclusiones	86
5.	CONCLUSIONES GENERALES Y PERSPECTIVAS.....	87
5.1.	CONCLUSIONES	88
5.2.	PERSPECTIVAS.....	88

6. REFERENCIAS	89
7. ANEXOS	99
7.1. ÍNDICE DE FIGURAS	100
7.2. ÍNDICE DE TABLAS	101
7.3. ECUACIONES EMPLEADAS PARA EL ANÁLISIS ANOVA	104
7.3.1. Sumas de cuadrados	104
7.3.2. Cuadrados medios	104
7.4. METODOLOGÍA DE SELECCIÓN DE VARIABLES STEPWISE	105

SÍMBOLOS Y ABREVIATURAS

y : Variable de respuesta, Valor real de...

\hat{y} : Valor predicho de "y".

\bar{y} : Valor promedio de "y".

\hat{y}_{-1} : Valor estimado de "y" mediante la estadística SCERP.

x : Variable predictora, Valor real de...

\bar{x} : Valor promedio de "x".

a : Ordenada al origen

b : Pendiente de la recta, coeficiente de regresión lineal

$r_{(x,y)}$: Coeficiente de correlación lineal

S_e^2, σ^2 : Varianza residual

ε : Residuo, diferencia entre el valor real y el pronosticado de "y".

n : Número total de observaciones

m : Número de variables "x".

K : Número de combinaciones diferentes de las variables "x".

p : Número de ecuaciones e incógnitas de la resolución matricial.

X : Matriz de variables independientes.

Y : Vector de observaciones.

β : Vector de coeficientes de regresión.

ϵ : Vector de los residuales.

R^2 : Coeficiente de determinación múltiple.

R_{aj}^2 : Coeficiente de determinación ajustado.

R_{Pred}^2 : Coeficiente de determinación predicho.

RLM (MLR)*: Regresión lineal múltiple.

RCP (PCR)*: Regresión sobre componentes principales.

$MCP (PLS)^*$: Mínimos cuadrados parciales.

$FIV (VIF)^*$: Factor de inflación de varianza.

$SC_{Reg} (SS_{Reg})^*$: Suma de cuadrados de la regresión.

$SC_{ERes} (SS_{Res})^*$: Suma de cuadrados del error residual.

$SC_{FA} (SS_{LOF})^*$: Suma de cuadrados por falta de ajuste.

$SC_{EExp} (SS_{PE})^*$: Suma de cuadrados del error experimental puro.

$SC_T (SS_T)^*$: Suma de cuadrados totales.

$CM_{Reg} (MS_{Reg})^*$: Cuadrados medios de la regresión.

$CM_{ERes} (MS_{Res})^*$: Cuadrados medios del error residual.

$CM_{FA} (MS_{LOF})^*$: Cuadrados medios por falta de ajuste.

$CM_{EExp} (MS_{PE})^*$: Cuadrados medios del error experimental puro.

$CM_T (MS_T)^*$: Cuadrados medios totales.

$SCERP (PRESS)^*$: Suma de cuadrados del error residual pronosticado.

(...)^{*}: Siglas en el idioma Inglés.

RESUMEN

En la búsqueda de una mejor calidad de vida, la preservación de un buen estado de salud es indispensable. El diagnóstico, la prevención y el tratamiento de las enfermedades requieren de métodos de análisis de muestras corporales que puedan ser confiables y del desarrollo de medicamentos eficaces que alivien o controlen el padecimiento y sus síntomas.

En años recientes, la obesidad se ha considerado como un factor desencadenante de diversos padecimientos clasificados dentro del síndrome metabólico, que en su conjunto son una de las principales causas de muerte y discapacidad en el mundo; estos padecimientos son asociados al desequilibrio en las concentraciones de glucosa, colesterol y triglicéridos en el organismo.

Debido al interés biomédico de contar con instrumentos eficientes para el monitoreo constante de estos analitos se han desarrollado nuevas metodologías de cuantificación, entre las que destaca el uso de biosensores enzimáticos. Estos dispositivos presentan muchas ventajas económicas y operacionales respecto a otros métodos de cuantificación; entre ellas su reducido tiempo de análisis y su alta selectividad.

El enfoque clásico de las determinaciones analíticas es el desarrollo de instrumentos altamente selectivos. No obstante, la naturaleza compleja de muestras biológicas y la comercialización de medicamentos que contienen dos o más fármacos, implica comúnmente que la muestra sea analizada por diferentes técnicas, cada una especializada en determinar a un solo analito, comprometiendo tiempo y recursos materiales.

Como una interesante alternativa a la problemática descrita anteriormente, los modelos de regresión multivariante permiten aprovechar a plenitud la señal analítica, determinando simultáneamente dos o más compuestos en la muestra mediante el procesamiento correcto de los datos. Así, estas herramientas estadísticas hacen posible procesar información analítica de dispositivos poco selectivos favoreciendo la simplificación en su construcción y diversificando sus aplicaciones en el análisis multicomponente convirtiendo los interferentes en analitos detectables y determinables.

En química analítica, la regresión multivariante se ha estudiado ampliamente en análisis espectroscópicos con excelentes resultados; sin embargo, en el ámbito de los sensores y de biosensores electroquímicos, la regresión multivariante ha sido poco explorada. En el presente trabajo se describen diversas metodologías electroanalíticas para la cuantificación simultánea de mezclas binarias de fármacos y glucosa, basadas en el uso de electrodos de carbón vítreo y platino; así como de un biosensor voltamperométrico de glucosa.

Esta tesis se divide en 6 capítulos: En el capítulo 1 se describe el funcionamiento de los sensores electroquímicos en general y de los biosensores enzimáticos en particular, sus principales componentes y sus aplicaciones biomédicas actuales. En el capítulo 2 se detallan los fundamentos matemáticos de los modelos de regresión multivariante, su validación estadística y algunas aplicaciones en determinaciones analíticas. En el capítulo 3 se describe la aplicación de los modelos de regresión multivariante discutidos en el capítulo anterior en dos aplicaciones electroquímicas específicas: la primera en la determinación simultánea de naproxeno y paracetamol; la segunda

en la determinación simultánea de sulfato ferroso y ácido ascórbico. Ambos trabajos emplearon electrodos no modificados y sus resultados no superaron el 10 % de error. En el capítulo 4 se estudia la combinación de la detección electroquímica en biosensores con la regresión multivariante para determinar dos o más analitos. Este capítulo se divide en dos subapartados: en el primero se reporta la simulación de la respuesta de un biosensor multienzimático para la determinación simultánea de glucosa, colesterol y triglicéridos; mientras que en el segundo se detalla la aplicación de un biosensor de glucosa con la capacidad de determinar simultáneamente ácido ascórbico y glucosa en muestras farmacéuticas. Los resultados demuestran que es posible determinar simultáneamente al menos dos analitos eficazmente empleando biosensores enzimáticos con detección voltamperométrica y regresión multivariante, especialmente con regresión sobre mínimos cuadrados parciales. El capítulo 5 describe las conclusiones generales del trabajo y las perspectivas en el corto y mediano plazo a partir de esta investigación. Finalmente, el capítulo 6 deja constancia de las referencias consultadas.

1. SENSORES ELECTROQUÍMICOS Y SUS APLICACIONES

En química analítica, los constantes avances tecnológicos han propiciado una evolución en los sistemas de cuantificación que facilitan los análisis de muestras complejas sin sacrificar la calidad de los resultados obtenidos. La necesidad de contar con instrumentos eficientes y económicos para la cuantificación de diversos compuestos en cualquier condición de análisis ha hecho posible el desarrollo de una gran cantidad de sensores y biosensores. Estos dispositivos presentan muchas ventajas económicas y operacionales respecto a otros métodos de cuantificación; como realizar una gran cantidad de mediciones sin alterar la integridad de la muestra, evitar el transporte y el almacenamiento prolongado de la muestra, así como la obtención de resultados en el orden de fracciones de segundo. El biosensor enzimático más desarrollado hasta el momento es el biosensor de glucosa, sin embargo, su esquema de funcionamiento se puede extender a la cuantificación de otros analitos de interés biomédico como colesterol y triglicéridos.

1.1. INTRODUCCIÓN

Las propiedades electroquímicas de las sustancias en disoluciones acuosas pueden ser utilizadas para realizar determinaciones analíticas al estimular una muestra eléctricamente y medir el potencial eléctrico, la intensidad de corriente o la impedancia, asumiendo que el valor resultante es proporcional a la concentración del analito. En la mayoría de las aplicaciones prácticas para la cuantificación se involucra la transferencia de electrones que puede ser reversible o irreversible, dependiendo de si el analito regresa a su estado original al aplicarle el estímulo eléctrico inverso o cambie permanentemente su estructura o composición en el proceso. El cambio químico resultante de la excitación eléctrica generalmente da como resultado la oxidación o la reducción de la sustancia de interés [1].

1.2. ANÁLISIS VOLTAMPEROMÉTRICOS Y CRONOAMPEROMÉTRICOS

Los métodos electroquímicos más comunes utilizados para determinaciones analíticas se basan en procesos que involucran la medición de la intensidad de corriente que circula entre dos electrodos; destacando por su sensibilidad la voltamperometría y la amperometría o cronoamperometría. La voltamperometría consiste en aplicar una variación regular de potencial en un cierto intervalo (o ventana) de potenciales definido. El potencial es graficado respecto a la intensidad de corriente medida en la celda, de tal manera que si en la ventana de potencial se encuentra un analito susceptible ser oxidado o reducido ocurrirá una transferencia de electrones que es detectada mediante un aumento repentino de la intensidad de corriente en forma de un “pico”. Por otra parte la cronoamperometría consiste en aplicar un potencial constante a la disolución. Cuando el sistema se estabiliza la intensidad de corriente permanece constante respecto al tiempo, hasta que ocurra un cambio en la concentración del analito de interés [2]. En la actualidad, ambas técnicas utilizan una celda de tres electrodos (Fig. 1):

- Electrodo de trabajo (ET): Hace contacto con el analito y aplica el potencial deseado de una manera controlada para provocar la transferencia de electrones en el analito.
- Electrodo auxiliar o Contraelectrodo (EA): En este electrodo circula la corriente que equilibra la corriente observada en el electrodo de trabajo.
- Electrodo de referencia (ER): Semicelda con un potencial constante y conocido. Su única función es actuar como referencia en la medición y controlar el potencial entre el EA y el ET. En este electrodo no circula corriente eléctrica.

La determinación de la concentración de un analito es relativamente sencilla. La corriente de oxidación o reducción del analito en la interfase del electrodo al aplicarle cierto potencial es directamente proporcional a la concentración de la muestra, de tal modo que se construye una curva de calibración y la concentración del analito en la muestra es calculada mediante interpolación. En una calibración voltamperométrica se observará el aumento del pico de corriente.

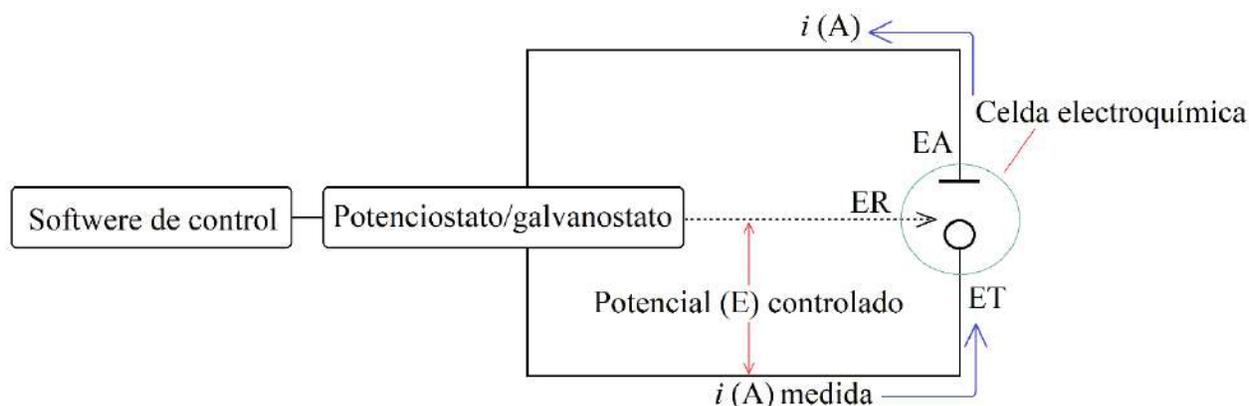


Figura 1. Esquema de detección electroquímica de una celda de tres electrodos.

El empleo de esta estrategia de determinación se encuentra limitada por la complejidad de las muestras y sus interferente electroactivos, por lo tanto, es necesario modificar el electrodo de trabajo para otorgarle selectividad al método. Estos electrodos modificados han evolucionado hasta convertirse en lo que actualmente se define como sensor o biosensor [3].

1.3. SENSORES Y BIOSENSORES ELECTROQUÍMICOS

Un sensor se define como un dispositivo que proporciona una señal cualitativa y/o cuantitativa en respuesta a un estímulo físico o la presencia de una determinada sustancia. En su forma más simple, un sensor se conforma por dos partes fundamentales: 1) el elemento de reconocimiento o detector, que es alterado en primera instancia por el fenómeno o sustancia que reconoce y 2) el transductor, que transforma la respuesta primaria del detector en una respuesta secundaria, normalmente de naturaleza eléctrica. En términos muy generales los sensores pueden clasificarse en tres categorías [4]:

- Sensores físicos: medición de magnitudes físicas como masa, temperatura, carga eléctrica, presión, etcétera.
- Sensores químicos: identificación y cuantificación de sustancias químicas mediante su respuesta a estímulos físicos externos.
- Biosensores: identificación y cuantificación de sustancias químicas empleando un elemento de reconocimiento de origen biológico para la detección altamente selectiva o específica de un analito.

1.4. BIOSENSORES

Los biosensores son sensores químicos que incorporan como elemento de reconocimiento un receptor de origen biológico, su esquema de funcionamiento se observa en la Figura 2. El analito que es detectado por el biosensor puede ser un compuesto inorgánico, orgánico o de origen bio-

lógico. Estos dispositivos son capaces de analizar eficientemente determinadas sustancias químicas de interés medioambiental, biomédico e industrial en muestras complejas sin destruir o alterar la composición de la muestra [5].

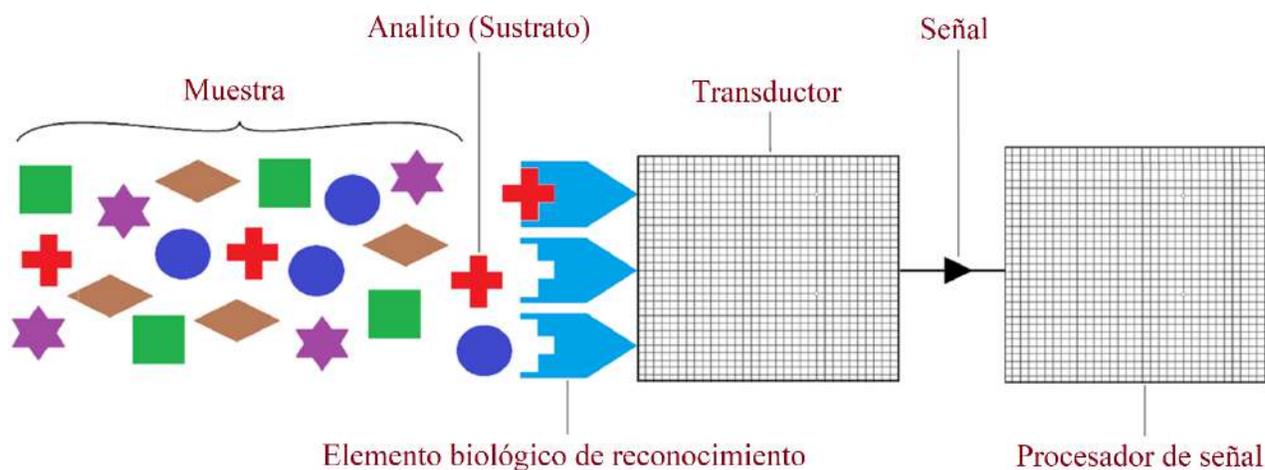


Figura 2. Esquema de funcionamiento de un biosensor.

1.4.1. Transductores

Con base en el tipo de señal primaria que procesan, los transductores pueden dividirse en los siguientes tipos [4]:

- Ópticos: Un transductor óptico convierte un cambio en la radiación electromagnética producida por un evento de reconocimiento, en una señal eléctrica.
- Piezo-Eléctricos: Son sistemas electromecánicos que reaccionan a la compresión. Involucran la generación de un potencial eléctrico debido a la aplicación de una fuerza sobre un cristal como consecuencia de un evento de reconocimiento.
- Térmicos: Todos los procesos químicos y biológicos se encuentran sujetos a la emisión o absorción de energía calorífica, que puede ser medida por medio de dispositivos termosensibles como los termopares.
- Electroquímicos:
 - Potenciométricos. Se emplean en la medición del potencial de una celda en ausencia de corriente eléctrica. El potencial es proporcional al logaritmo de la concentración de la sustancia a determinar.
 - Voltamperométricos. Un barrido de potencial es aplicado a la celda para el análisis de la oxidación y/o reducción de las sustancias químicas, obteniéndose de uno o más potenciales, valores de intensidad de corriente que son directamente proporcionales a la concentración de los analitos.
 - Amperométricos. Si el potencial aplicado es constante, se obtiene el registro de la corriente eléctrica en función del tiempo, conociéndose este modo como un análi-

sis cronoamperométrico. La corriente eléctrica obtenida es directamente proporcional a la concentración del analito de interés.

- Conductimétricos. Un cambio en la composición de la solución conlleva a una variación en la conductividad eléctrica que puede ser medida eléctricamente y relacionada con la concentración de algún analito de interés.

1.4.2. Elementos de reconocimiento de origen biológico

Los elementos de reconocimiento son un aspecto clave en el dispositivo dado que proporcionan la selectividad o especificidad que el sensor requiere para responder a un único analito, con lo que se reduce el efecto de las posibles interferencias de otras sustancias en una muestra. Los elementos de reconocimiento de origen biológico son macromoléculas aisladas de seres vivos con estructuras y funciones definidas; como anticuerpos, ácidos nucleicos y proteínas.

Debido a su relativa economía, los elementos de reconocimiento más comunes en los biosensores son las enzimas; en términos generales son proteínas especializadas en la catálisis de reacciones biológicas [6]. En el caso particular de las oxidasas, su activación requiere de la oxidación de su sitio activo, proceso que puede detectarse electroquímicamente.

1.4.3. Métodos de inmovilización

Para asegurar el funcionamiento óptimo de un biosensor el elemento de reconocimiento debe estar en contacto directo con el transductor; el procedimiento realizado para lograr este cometido se denomina inmovilización. Los métodos más comunes para la inmovilización de las enzimas son los siguientes [7].

- Adsorción en la superficie del transductor
- Microencapsulación: Captura de la enzima entre dos membranas permeables.
- Entrampamiento: La enzima es capturada en una matriz de gel o algún polímero.
- Anclaje covalente: La enzima forma enlaces covalentes con el transductor.
- Entrecruzamiento: Un agente bifuncional es usado como enlace químico entre diversas unidades de la enzima o entre éstas y algún componente adicional con la finalidad de disminuir su solubilidad y favorecer su adsorción sobre el transductor.

Es importante mencionar que la adsorción es el método que menos afecta al elemento de reconocimiento, pero que es el más inestable ya que el receptor puede perderse fácilmente. En contraparte, el anclaje covalente es la metodología que genera dispositivos más estables pero con mayor afectación en la actividad sensorial del elemento biológico de reconocimiento [8].

1.4.4. Aplicación de los biosensores enzimáticos

Las aplicaciones de estos dispositivos se han diversificado de modo que se utilizan en análisis clínicos, medioambientales y en la industria alimenticia. En la actualidad se utilizan para determinar una gran cantidad de analitos, entre los más comunes, compuestos orgánicos simples como alcoholes, aminoácidos, carbohidratos, colesterol y triglicéridos, entre muchos otros [9].

1.5. IMPORTANCIA BIOMÉDICA DE LOS BIOSENSORES ENZIMÁTICOS

En años recientes, la obesidad se ha considerado como un factor desencadenante de diversos padecimientos clasificados dentro del síndrome metabólico; definido como la presencia conjunta de intolerancia a la glucosa, dislipidemia e hipertensión arterial asociada a enfermedades cardiovasculares y diabetes, generalmente asociados al desequilibrio en las concentraciones de glucosa, colesterol y triglicéridos en el organismo [10]. La prevención y el tratamiento de estas enfermedades requiere un monitoreo constante de la concentración de estos tres analitos en la sangre con la finalidad de contribuir al mejoramiento de nuestra dieta y a mantener en niveles normales las concentraciones de estos compuestos [11]. Uno de los métodos más eficaces para su determinación es el uso de biosensores enzimáticos; entre sus características más notables se encuentran su alta selectividad y sensibilidad mediante el uso de instrumentación electroquímica robusta, versátil, portátil y de bajo costo en comparación con los métodos espectroscópicos [12].

1.5.1. Determinación clínica de glucosa

La glucosa, libre o combinada, es el compuesto orgánico más abundante de la naturaleza. Es uno de los principales productos de la fotosíntesis y se considera como la fuente primaria de energía de las células. Este compuesto es el monómero principal de polímeros de importancia estructural como la celulosa y de polímeros de almacenamiento energético como el almidón y el glucógeno. La glucosa se absorbe directamente al torrente sanguíneo durante la digestión y se encuentra presente en la mayoría de los endulzantes comerciales, generalmente en forma de disacárido enlazado a una molécula de fructosa (sacarosa). Cuando el organismo no es capaz de controlar la concentración de glucosa en la sangre provoca una serie de complicaciones médicas que deterioran la salud del individuo [13]. Las determinaciones de glucosa en el laboratorio clínico generalmente se realizan en muestras de sangre venosa. Actualmente, el método más utilizado para determinar glucosa en sangre es el método enzimático de Trinder [14], [15].

La enzima glucosa oxidasa (GOx) oxida a la glucosa en presencia de agua y oxígeno para formar ácido glucónico y peróxido de hidrógeno, además, es específica para la β -D-glucosa, por lo que no existe reacción cruzada con otros azúcares. En solución acuosa, la D-glucosa presenta un equilibrio de mutarrotación entre sus anómeros α y β en una proporción aproximada de 64% y 36% respectivamente. La tasa de interconversión depende del pH, de la temperatura y del tiempo para alcanzar el equilibrio anomérico [16]. El peróxido de hidrógeno en presencia de la enzima peroxidasa (POx) puede condensar fenol y 4-aminoantipirina (4-AAP) para formar la N-antipirina-p-

benzoquinonaimina (QIA). Este compuesto posee un color rojo y su concentración es proporcional a la concentración de glucosa en la muestra (Fig. 3). Para la determinación indirecta de glucosa se emplea un espectrofotómetro y se mide la absorbancia (A) de la QIA en la muestra a una longitud de onda de aproximadamente 500 nm.

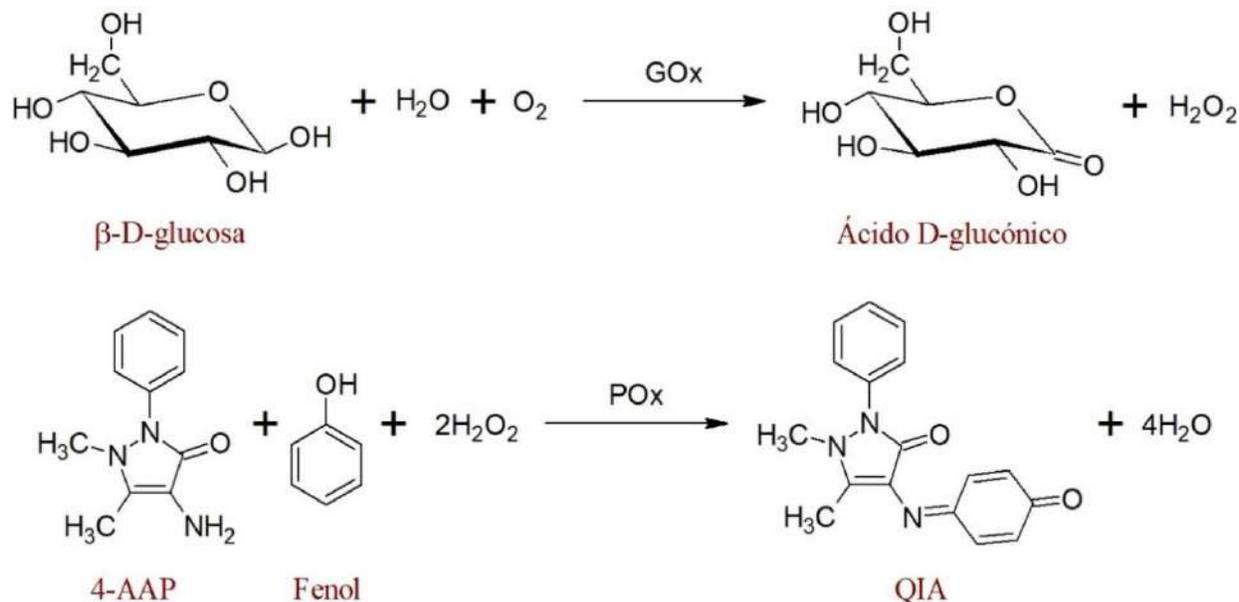
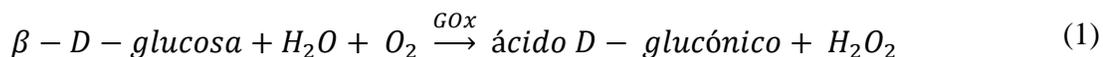


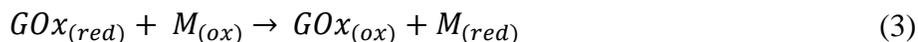
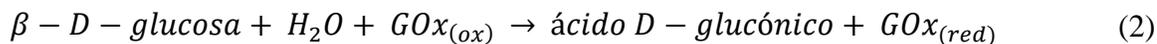
Figura 3. Esquema del método enzimático de Trinder para la cuantificación de glucosa.

1.5.2. Biosensores enzimáticos de glucosa

En 1962, Clark y Lyons propusieron el concepto de electrodos enzimáticos de glucosa, descubriendo algunos principios de la biodetección de la glucosa. Esta metodología resultó ser económica y con un enorme potencial de aplicación en los sustratos de interés mediante el empleo de las enzimas correspondientes [17]. La primera generación de estos biosensores se fundamentó en la detección amperométrica del oxígeno consumido en la oxidación de la glucosa mediante la enzima glucosa oxidasa (GOx) de acuerdo con la ecuación 1.



En años posteriores se describieron biosensores de glucosa que detectaban la formación de peróxido de hidrógeno liberado; sin embargo, su uso está limitado en función de las abundantes interferencias electroquímicas, ya que la medición amperométrica requiere la aplicación de un potencial de trabajo elevado en el que compuestos electroactivos como el ácido ascórbico provocan un aumento de la señal analítica que no corresponde a la oxidación del peróxido de hidrógeno [18]. Para disminuir el efecto de las interferencias, se reemplazó el oxígeno como aceptor de electrones por un mediador sintético (M) que promueve la transferencia de electrones a potenciales más bajos y reducir así los errores provocados por las sustancias interferentes (Ecs. 2-4).



Esta metodología dio lugar a la segunda generación de biosensores de glucosa. El mediador se ancla entre el transductor y la enzima, permitiendo que ésta regenere inmediatamente su sitio activo sin depender de la concentración de oxígeno de la disolución (Fig. 4). La señal analítica corresponde únicamente a la corriente generada por la oxidación de la glucosa [19].

Entre los mediadores más reportados para biosensores enzimáticos electroquímicos de glucosa se encuentran los derivados de ferrocianuro, derivados de quinonas, tetratrafalvaleno (TTF) y 7, 7, 8, 8-tetracianoquinodimetano (TCNQ) [20]. Éste último compuesto es una sal orgánica conductora del cual se han sintetizado cientos de complejos de transferencia de carga así como materiales compósitos que son excelentes receptores de electrones o mediadores debido a que la molécula del TCNQ presenta dobles enlaces conjugados a lo largo de su estructura [21].

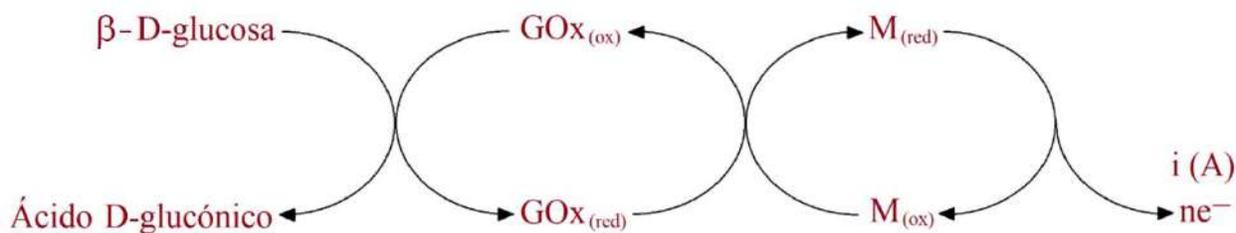


Figura 4. Esquema de transferencia de electrones entre el sustrato y el transductor en un biosensor enzimático de glucosa.

Las reacciones de óxido-reducción reversibles del TCNQ en acetonitrilo se observan en la Figura 5; en disolución acuosa, el radical de la primera reducción del TCNQ es inestable, por lo que se intercambian directamente dos electrones en el proceso [22].



Figura 5. Reacciones óxido-reducción del TCNQ en acetonitrilo.

1.5.3. Biosensores enzimáticos de colesterol

El colesterol es un lípido sintetizado por el hígado y su importancia para el organismo estriba en que es un precursor en la síntesis de hormonas, vitaminas y ácidos biliares; además de ser un elemento estructural presente en la membrana celular de diversos tejidos corporales. No obstante, su acumulación excesiva en el torrente sanguíneo como consecuencia de la ingesta inmoderada de alimentos de origen animal provoca afectaciones importantes en el sistema circulatorio [23].

Su cuantificación puede realizarse del mismo modo que para glucosa empleando biosensores enzimáticos. Recientemente se han desarrollado dispositivos basados en la detección electroquímica de la oxidación del colesterol a colest-4-en-3-ona mediante la enzima colesterol oxidasa (COx) en presencia de oxígeno o de un mediador (Fig. 6) [24]. El biosensor enzimático de colesterol presenta un esquema de reconocimiento análogo a los biosensores de glucosa basados en glucosa oxidasa (Fig. 7).

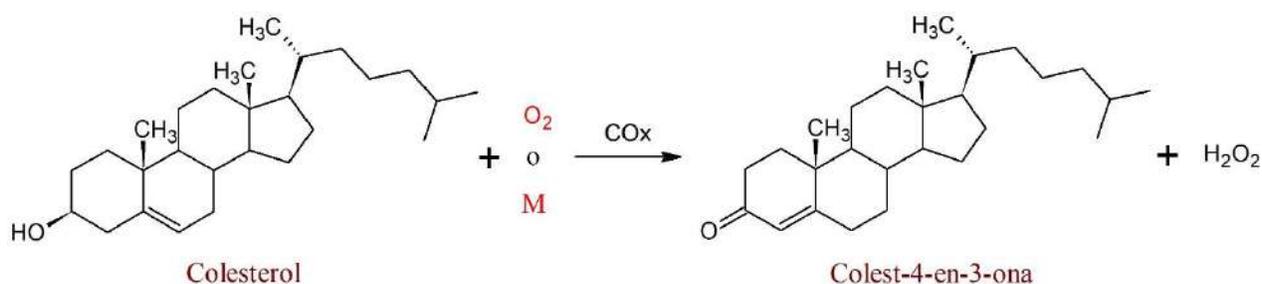


Figura 6. Oxidación enzimática de colesterol en presencia de COx y un receptor de electrones.

Además del oxígeno utilizado para la regeneración de la actividad enzimática de la COx, se utilizan mediadores que favorecen el transporte de electrones a potenciales de trabajo que no produzcan interferencias [25].

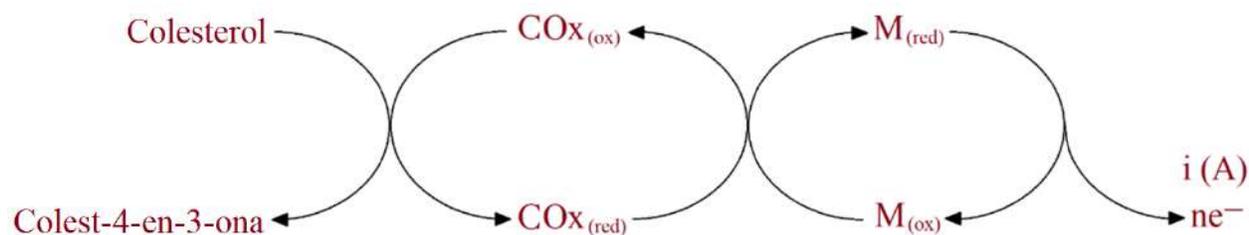


Figura 7. Esquema de transferencia de electrones entre el sustrato y el transductor en un biosensor enzimático de colesterol.

Recientemente, se han desarrollado estrategias para mejorar el desempeño de los biosensores enzimáticos de colesterol; sin embargo, la mínima solubilidad del colesterol en agua y la baja acti-

vidad de las enzimas disponibles comercialmente, limitan su evolución y encarecen su incursión en el mercado de diagnóstico y monitorización [26].

1.5.4. Biosensores enzimáticos de triglicéridos

Los triglicéridos (grasas y aceites) son ésteres de glicerol y de ácidos carboxílicos de cadena larga de origen biológico. Su función primaria es la reserva de energía; células especializadas en su síntesis y almacenamiento forman el tejido adiposo que se concentra principalmente en la cavidad abdominal y en la capa subcutánea. No obstante la importancia de los triglicéridos en el organismo, su acumulación excesiva provoca serios problemas de salud [16].

La cuantificación de los triglicéridos mediante biosensores enzimáticos involucra una preincubación prolongada de la muestra y la transformación química de este analito en una especie detectable electroquímicamente [27]. Los biosensores amperométricos más comunes para triglicéridos se basan principalmente en la actividad conjunta de tres enzimas sobre el analito y sus productos: Lipasa (LPA), glicerol quinasa (GQ) y glicerol-3-fosfato oxidasa (GPOx) [28].

- Lipasa: Esta enzima hidroliza la molécula del triglicérido en una molécula de glicerol y tres moléculas de ácido graso libre (Fig. 8).

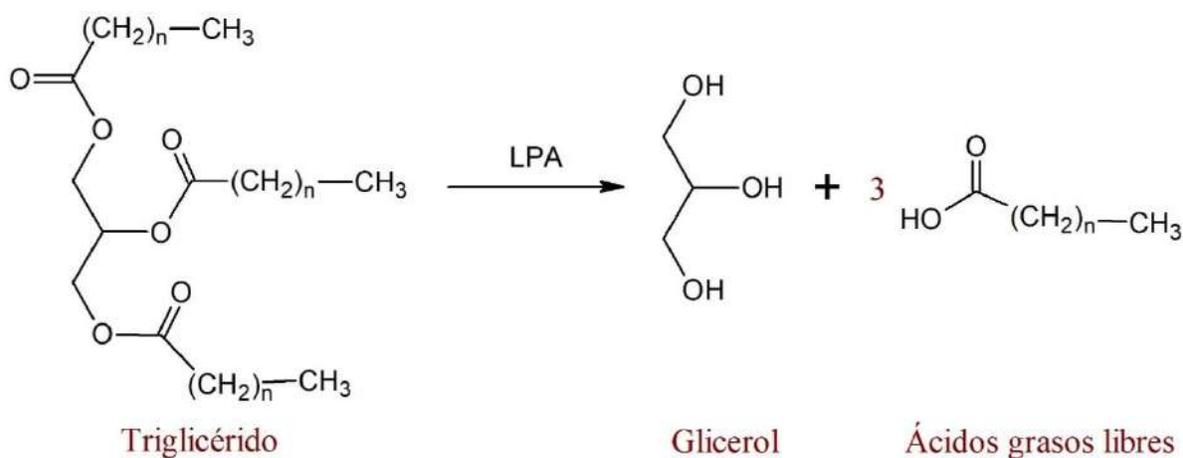


Figura 8. Hidrólisis de un triglicérido a glicerol y ácidos grasos libres.

- Glicerol quinasa: Se encarga de transferir un grupo fosfato de una molécula de trifosfato de adenosina (ATP) hacia una molécula de glicerol, produciendo difosfato de adenosina (ADP) y glicerol-3-fosfato (Fig. 9).

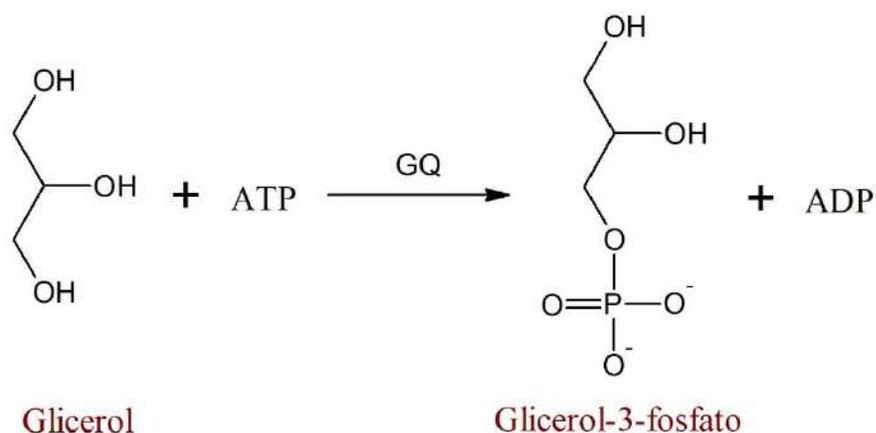


Figura 9. Adición de un grupo fosfato a una molécula de glicerol.

- Glicerol-3-fosfato oxidasa: En presencia de oxígeno oxida al glicerol-3-fosfato produciendo peróxido de hidrógeno y fosfato de dihidroxiacetona (Fig. 9).

De acuerdo a estos mecanismos, se tienen algunas alternativas para determinar la concentración del triglicérido. La primera es la determinación del oxígeno consumido en la solución, la segunda es la detección del peróxido de hidrógeno producido durante la oxidación del glicerol-3-fosfato.

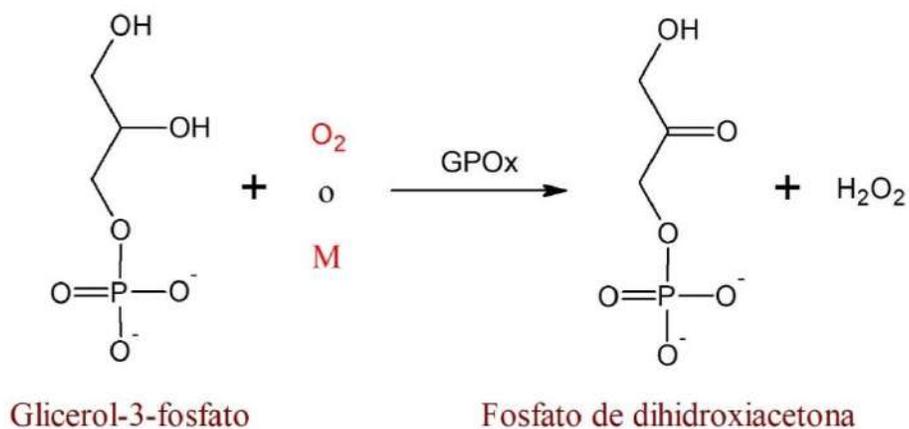


Figura 10. Oxidación de glicerol-3-fosfato a fosfato de dihidroxiacetona.

Finalmente, si se sustituye el oxígeno con un mediador que regenere continuamente el sitio activo de la GPOx se puede determinar indirectamente la concentración de triglicéridos en un esquema de detección análogo a los reportados para biosensores de glucosa y colesterol (Fig. 10) [29].

1.6. ANÁLISIS ELECTROQUÍMICOS SIMULTÁNEOS

La voltamperometría es una técnica particularmente atractiva para el análisis de muestras que contienen dos o más analitos. Siempre que los analitos se comporten de forma independiente, el voltamperograma de una mezcla multicomponente es una suma de los voltamperogramas individuales de cada analito. Si la separación entre los potenciales de oxidación es suficiente, es posible determinar independientemente cada analito en la muestra. No obstante, es frecuente encontrarse con mezclas de analitos que presentan superposición en sus corrientes de oxidación, por lo que se prefiere trabajar con otras técnicas de cuantificación o realizar un pretratamiento de la muestra.

Una estrategia para la determinación electroquímica simultánea de dos o más analitos consiste en el uso de arreglos de electrodos, sensores o biosensores. Estos sistemas analíticos están inspirados en la capacidad sensorial de sabor en mamíferos y son mejor conocidos como lenguas electrónicas, donde algunos receptores pueden responder de una manera distinta a una gran variedad de sustancias. Las señales registradas posteriormente son canalizadas a una etapa compleja de tratamiento de datos que permite cuantificar o clasificar una gran cantidad de analitos. Esta estrategia requiere una mayor infraestructura tecnológica para ser implementada.

Otra estrategia es el empleo de modelos de regresión multivariante para el tratamiento de señales superpuestas para la cuantificaciones de diversos analitos empleando un solo electrodo. Esta estrategia ha sido poco desarrollada y estudiada para técnicas electroquímicas, sin embargo, algunos ejemplos encontrados en la literatura son los siguientes (Tabla 1):

Tabla 1. Determinaciones electroquímicas simultáneas empleando un solo electrodo de trabajo.

Analitos	Técnica	ET	Procesamiento de Datos	Referencia
Cisteína, tirosina, triptófano	Voltamperometría cíclica	Compósito Grafito-Metacrilato	RCP, MCP	[30]
Clorpromazina Prometazina	Voltamperometría de redisolución con diferencial de pulso	Electrodo de carbón vítreo	RCP, MCP, RLM	[31]
Nitrobenceno 2-Nitrofenol 3-Nitrofenol 4-Nitrofenol 2,4-Dinitrofenol	Voltamperometría diferencial de pulsos	Electrodo de gota de mercurio (DME)	RCP, MCP, RLM	[32]
Paratión-Metilo Fenitrotión Paratión	Voltamperometría de redisolución con diferencial de pulso	Electrodo de gota de mercurio (DME)	RCP, MCP, RLM, Redes neuronales	[33]
Ofloxacina Ciprofloxacina Norfloxacina	Voltamperometría lineal	Electrodo de gota de mercurio (DME)	RCP, MCP, RLM	[34]

RLM = Regresión lineal múltiple, RCP = Regresión sobre componentes principales, MCP = Regresión sobre mínimos cuadrados parciales.

2. REGRESIÓN MULTIVARIANTE

La relación fundamental que se observa entre diversos fenómenos en la naturaleza puede reducirse a una simple interacción de causa y efecto. De esta relación puede establecerse un modelo matemático que describa el comportamiento de una propiedad en función de una o más variables relacionadas. La utilidad de estos modelos estriba en la posibilidad de hacer estimaciones y predicciones útiles de un sin número de variables con información previa. El fundamento en la resolución de estos problemas consiste la resolución de ecuaciones mediante álgebra matricial de los datos previamente organizados en matrices.

Muchas propiedades de las disoluciones que son analizadas por métodos instrumentales cumplen con la ley de aditividad, es decir, la señal del instrumento es la suma de las señales individuales de los compuestos que se detectan. En química analítica, esta propiedad ha permitido el desarrollo de metodologías para la determinación simultáneas de dos o más compuestos en una muestra mediante el procesamiento correcto de los datos empleando herramientas quimiométricas como la regresión multivariante.

2.1. REGRESIÓN LINEAL SIMPLE

El modelo de la línea recta establece una relación entre dos variables en la que sus magnitudes son directamente proporcionales. En química analítica es común asociar una propiedad de una disolución a la cantidad de sustancia presente en ella. La propiedad medida entonces es dependiente de la cantidad de analito en la disolución. Actualmente, estas medidas se efectúan en equipos que estimulan una muestra con el objetivo de recuperar una señal analítica medible que se asume, es proporcional a la concentración del analito de interés.

Se comienza con disoluciones que contienen el analito a una concentración conocida a las que se les denomina soluciones patrón o estándar. Estas soluciones son medidas en el instrumento analítico bajo las mismas condiciones que las utilizadas posteriormente para las disoluciones de concentración desconocida [35]. La gráfica de calibrado siempre se construye con la magnitud de la señal en el eje vertical "y", mientras que la concentración del analito se sitúa sobre el eje horizontal "x", bajo la premisa de que los errores se encuentran en los valores de "y". Se asume también que las concentraciones de las disoluciones estándar se encuentran libres de error. La línea de calibrado tiene la forma de una ecuación lineal de primer grado como la representada por la ecuación 5:

$$y = a + bx \quad (5)$$

Donde "b" es la pendiente de la recta y "a" su ordenada al origen. Los puntos individuales sobre la línea se denotan por (x_n, y_n) . La media o promedio de los valores de "x" se designa por " \bar{x} " y de "y" por " \bar{y} ". Para estimar el ajuste de los puntos experimentales a una línea recta se calcula el coeficiente de correlación "r". Una vez establecida la gráfica de calibrado, puede obtenerse la concentración del analito en la solución problema por interpolación a partir de la magnitud de su propiedad o señal.

2.2. ANÁLISIS MULTIVARIANTE

El análisis multivariante comprende los métodos estadísticos que analizan simultáneamente un conjunto de datos cuya magnitud depende de diversas variables medidas. Estos métodos tienen como objetivo principal la discriminación de datos así como su clasificación dentro de grupos de datos que comparten información específica [36]. En química analítica, una de sus principales utilidades es la regresión multivariante donde las variables se pueden dividir en dos grupos; variables de respuesta y variables predictoras. Las variables de respuestas "y", dependen de una serie de variables predictoras $x_1, x_2, x_3 \dots, x_m$, para generar un modelo de regresión donde las respuestas de una muestra problema predicen la concentración del analito de interés en un esquema que se denomina calibración inversa.

Los modelos más empleados para la regresión multivariante son; regresión lineal múltiple (RLM), regresión sobre componentes principales (RCP) y regresión de mínimos cuadrados parciales (MCP) [37].

2.3. REGRESIÓN LINEAL MÚLTIPLE (RLM)

Los modelos multivariados lineales pueden escribirse de forma general como el polinomio de primer grado que se muestra en la ecuación 6:

$$y = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + \dots + b_mx_m = b_0 + \sum_{i=1}^m b_ix_i \quad (6)$$

El procedimiento de mínimos cuadrados puede emplearse para estimar los coeficientes de regresión b_0, b_1, \dots, b_m , en la regresión lineal múltiple. Considerando " n " observaciones $y_1, y_2, y_3, \dots, y_n$, cada una con varianza σ^2 obtenida en " n " combinaciones diferentes de las variables independientes $x_1, x_2, x_3, \dots, x_m$ ($n > m$). Si se establece el modelo multivariado de la ecuación 7 entre la respuesta " y " las " m " variables " x " cada observación puede representarse como se muestra en la ecuación 7:

$$y_i = b_0 + b_1x_{i1} + b_2x_{i2} + \dots + b_mx_{ik} + \varepsilon_i \quad (7)$$

Donde:

- y_i representa la i -ésima respuesta o i -ésima concentración del analito en una determinación analítica ($i = 1, 2, \dots, n$).
- x_{ik} representa el valor de la k -ésima variable independiente para la observación i ($k = 1, 2, \dots, m$) o bien, el valor de la k -ésima señal analítica.
- ε_i es el i -ésimo residuo, es decir, la diferencia entre el valor original de la i -ésima variable de respuesta y_i y su valor calculado o ajustado \hat{y}_i a partir de la ecuación de regresión, es decir $\varepsilon_i = y_i - \hat{y}_i$.

Las estimaciones de mínimos cuadrados de los coeficientes de regresión desconocidos se obtienen minimizando la suma de los residuos cuadrados. Esto requiere la solución de un sistema de " p " ecuaciones normales con " p " incógnitas ($p = m + 1$). Al expresar el problema de regresión en la notación matricial se obtiene una solución que es aplicable a cualquier situación de regresión lineal, incluyendo la línea recta simple [38]. Consideramos el vector de observaciones \mathbf{Y} de orden ($n \times 1$) como:

$$\mathbf{Y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

El vector de los coeficientes de regresión a estimar \mathbf{B} ($p \times 1$) como:

$$\mathbf{B} = \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_m \end{bmatrix}$$

La matriz de variables independientes \mathbf{X} ($n \times p$) como:

$$\mathbf{X} = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1m} \\ 1 & x_{21} & x_{22} & \cdots & x_{2m} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nm} \end{bmatrix}$$

La primera columna a la que corresponden valores de la unidad permite la estimación de los interceptos b_0 , que corresponden al valor de la variable "x" en el primer término de la ecuación 7. El vector de error ϵ ($n \times 1$) se denota como:

$$\epsilon = \begin{bmatrix} \epsilon_0 \\ \epsilon_1 \\ \vdots \\ \epsilon_m \end{bmatrix}$$

El modelo representado por la ecuación 8 se convierte en la expresión matricial dada por la ecuación 8:

$$\mathbf{Y} = \mathbf{XB} + \epsilon \quad (8)$$

Por lo tanto, la estimación de mínimos cuadrados de los coeficientes de correlación se obtiene a partir de la ecuación 9, siempre y cuando la matriz $\mathbf{X}^T \mathbf{X}$ no sea singular.

$$\mathbf{B} = \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_m \end{bmatrix} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y} \quad (9)$$

Esta es la solución de mínimos cuadrados aplicable a todos los modelos que son lineales. Con estos coeficientes de regresión los valores de respuesta estimados $\hat{\mathbf{Y}}$ pueden ser calculados resolviendo la ecuación 10:

$$\hat{\mathbf{Y}} = \begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \vdots \\ \hat{y}_n \end{bmatrix} = \mathbf{X}\mathbf{B} = \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{Y} \quad (10)$$

A partir de los valores estimados de " \hat{y} ", la varianza residual S_ε^2 , que es una estimación del error, se obtiene de la ecuación 11:

$$S_\varepsilon^2 = \frac{\sum \varepsilon_i^2}{n - p} = \frac{\sum (y_i - \hat{y}_i)^2}{n - p} \quad (11)$$

2.3.1. Validación del modelo de RLM

El análisis de varianza (ANOVA) es una herramienta útil para la validación del modelo de regresión lineal múltiple que permite verificar si las variables predictoras pueden explicar una cantidad significativa de la varianza en las variables de respuesta. En la Tabla 2 se generaliza para la regresión múltiple con " p " coeficientes de regresión, " n " número de observaciones y un número " k " de diferentes configuraciones de las variables " x " (cuando no existen observaciones repetidas $k = n$) y se cumple que $m < k \leq n$.

Tabla 2. Análisis de la varianza para RLM.

Fuente de la Variación	Suma de Cuadrados	Grados de Libertad	Cuadrados Medios	F
Variación de la Regresión	SC_{Reg}	$p - 1$	CM_{Reg}	CM_{Reg}/CM_{ERes}
Variación debida al error residual	SC_{ERes}	$n - p$	CM_{ERes}	
Variación debida a falta de ajuste	SC_{FA}	$k - p$	CM_{FA}	CM_{FA}/CM_{EExp}
Variación del error experimental puro	SC_{EExp}	$n - k$	CM_{EExp}	
Variación total	SC_T	$n - 1$	CM_T	

Donde: SC_{Reg} = Suma de cuadrados de la regresión, SC_{ERes} = Suma de cuadrados del error residual, SC_{FA} = Suma de cuadrados por falta de ajuste, SC_{EExp} = Suma de cuadrados del error experimental puro, SC_T = Suma de cuadrados totales, CM_{Reg} = Cuadrados medios de la regresión, CM_{ERes} = Cuadrados medios del error residual, CM_{FA} = Cuadrados medios por falta de ajuste, CM_{EExp} = Cuadrados medios del error experimental puro, CM_T = Cuadrados medios totales.

El contraste "F" obtenido de los cuadrados medios de la regresión y del error residual debería tener un valor elevado en caso de que exista una correlación importante entre las variables. Si se dispone de mediciones repetidas, es posible obtener más información del sistema, como la variación debido a la falta de ajuste y la variación debido al error experimental puro [39]. Para el cálculo de las sumas de cuadrados y de los cuadrados medios para el análisis ANOVA se utilizan las ecuaciones mostradas en el anexo 7.3.

2.3.2. Coeficiente de determinación múltiple

En una línea recta, el coeficiente de correlación lineal entre "x" y "y", así como el cuadrado del coeficiente de correlación lineal ($r_{x,y}^2$), dado por la ecuación 12, representan la proporción de la variación de la variable "y" que es explicada por la variable "x".

En la regresión lineal múltiple, el coeficiente usado para estimar la proporción de la variación de "y" que es explicada por la regresión se denota como "R", y es conocido como coeficiente de correlación múltiple. El cuadrado del coeficiente "R", es decir R^2 , se denomina como coeficiente de determinación múltiple y es calculado por medio de la ecuación 13.

$$r_{x,y}^2 = \frac{SC_{Reg}}{SC_T} = \frac{\sum_i n_i (\hat{y}_i - \bar{y})^2}{\sum_i \sum_j (y_{ij} - \bar{y})^2} \quad (12)$$

$$R^2 = \frac{SC_{Reg}}{SC_T} = \frac{SC_T - SC_{ERes}}{SC_T} = 1 - \frac{SC_{ERes}}{SC_T} \quad (13)$$

Si no existe una relación lineal entre la variable de respuesta y las variables predictoras R^2 será mucho menor a 1; si es una correlación perfecta entonces $R^2 = 1$. El valor de R^2 puede incrementarse adicionando variables predictoras "x" al modelo y puede alcanzar el valor de la unidad si el número de coeficientes en el modelo es igual al número de observaciones ($p = n$).

2.3.3. Multicolinealidad

Cuando dos o más variables predictoras "x" se encuentran correlacionadas linealmente, no son independientes entre sí, es decir, son colineales o multicolineales. Proporcionan información redundante que maximiza el valor de la suma de cuadrados residuales; lo que causa que los coeficientes de la regresión lineal sean imprecisos en la predicción del valor de las variables de respuesta. Ciertos criterios establecidos indican que existen problemas de colinealidad si la correlación ($r_{x,y}$) entre dos variables predictoras es mayor o igual a 0.70 [38]. Este problema puede ser resuelto eliminando las variables más correlacionadas entre sí, o bien, modificando la matriz **X** normalizando las variables o mediante un análisis de componentes principales.

2. Regresión multivariante

Un indicador frecuentemente utilizado para estimar la interdependencia de las variables "x" es la tolerancia; para cada variable x_i es calculado a partir de la ecuación 14, donde $R_{x_i}^2$ es el coeficiente de determinación múltiple para la regresión entre x_i (considerada como variable independiente) y las demás variables independientes.

$$Tolerancia_{x_i} = 1 - R_{x_i}^2 \quad (14)$$

Representa por lo tanto, la variación en x_i que puede ser explicada por las otras variables "x". Un valor de tolerancia pequeño significa que x_i presenta una correlación lineal con las otras variables "x". Un indicador relacionado de la multicolinealidad es el de factor de inflación de varianza (*FIV*) que es el recíproco de la tolerancia y está dado por la ecuación 15. Este indicador se calcula para todas las variables "x" presentes en el modelo de regresión.

$$FIV_{x_i} = \frac{1}{Tolerancia_{x_i}} = \frac{1}{1 - R_{x_i}^2} \quad (15)$$

El valor del *FIV* puede tener valores desde 1 hasta infinito. Un valor de 1 significa que la variable x_i es independiente de cualquier otra variable "x" presente en el modelo de regresión. A medida que el *FIV* aumente su valor, el coeficiente (*b*) de la variable "x" se vuelve inestable debido a que su varianza ha sido "inflada" por estar relacionada con otra variable "x". El valor del *FIV* debe ser menor a 10; es decir $1 \leq FIV < 10$; si el valor de $FIV \geq 10$, la variable en cuestión se encuentra muy correlacionada con otras variables "x" y, por lo tanto, se tiene un problema severo de multicolinealidad [40].

2.3.4. Coeficiente de determinación ajustado

Al igual que el coeficiente de determinación múltiple, el coeficiente de determinación ajustado R_{aj}^2 es usado para estimar la proporción de la variación de "y" que es explicada por la regresión, pero ajustado a los grados de libertad de acuerdo a la ecuación 16. Si se comparan dos regresiones que difieren entre sí en el número de datos y el número de variables, es apropiado compararlas mediante este coeficiente. Su valor también oscila entre 0 y 1 ($0 \leq R_{aj}^2 \leq 1$).

$$R_{aj}^2 = 1 - \frac{\frac{SC_{ERes}}{n - p}}{\frac{SC_T}{n - 1}} = 1 - \frac{CM_{ERes}}{CM_T} \quad (16)$$

2.3.5. Coeficiente de determinación predicho

Indica el grado de predicción del modelo para nuevas observaciones. El coeficiente R_{pred}^2 es calculado mediante la remoción sistemática de una observación del total de las mismas mediante la estadística de suma de cuadrados del error residual predicho (*SCERP*); estimando la ecuación de regresión y determinando que tan bien predice el valor de la observación removida \hat{y}_{-i} conforme a la ecuación 17.

$$SCERP = \sum_i (y_i - \hat{y}_{-i})^2 \quad (17)$$

Este coeficiente de correlación puede tomar valores de entre 0 y 1 ($0 \leq R_{pred}^2 \leq 1$) y es calculado mediante la ecuación 18. Un valor de R_{pred}^2 cercano a la unidad indica que el modelo generado posee una gran capacidad de predicción para nuevas observaciones. Valores superiores a 0.9 obtenidos experimentalmente son considerados como válidos para concluir que el modelo es aceptable [41], [42].

$$R_{pred}^2 = 1 - \frac{SCERP}{SS_T} \quad (18)$$

2.4. SELECCIÓN DE LAS VARIABLES DE RESPUESTA

Para una determinación precisa de las concentraciones de los analitos de interés, es necesario seleccionar adecuadamente las variables predictoras. Una característica típica de los datos obtenidos de las técnicas instrumentales es que la cantidad de variables de respuesta es mucho menor en comparación con el número de variables predictoras; sin embargo, muchas de las variables son irrelevantes o redundantes, por lo que el número de variables a menudo puede ser reducido con poca pérdida de información. La voltamperometría por ejemplo, registra la intensidad de corriente que circula en un electrodo de trabajo a una serie de potenciales definidos. De acuerdo a las condiciones del análisis se tienen decenas o cientos de variables predictoras para una o más variables de respuesta. No resulta práctico probar todas las combinaciones de variables en cuestión para encontrar las variables predictoras idóneas que se correlacionen con las variables de respuesta; se necesitan estrategias de selección de variables para buscar combinaciones apropiadas. Los métodos de selección univariante tratan cada variable de forma independiente; en un modelo lineal, por ejemplo se selecciona el máximo de intensidad de corriente o de absorbancia para garantizar parámetros estadísticos idóneos. En un modelo multivariado comúnmente se realiza la selección de las variables de manera empírica en conformidad con la experiencia y el criterio de analista; sin embargo, esta estrategia no considera la posible colinealidad o multicolinealidad en los datos y es relativamente frecuente que se presente, generando un modelo con poca robustez estadística [43].

2.4.1. Método paso a paso (stepwise)

El modelo de regresión lineal múltiple stepwise; es considerado como un modelo de regresión multivariante altamente aceptado y utilizado para reducir el grado de colinealidad de los datos. En este método, las variables predictoras se seleccionan o eliminan sistemáticamente de acuerdo con criterios estadísticos seleccionados previamente; se basa en introducir o eliminar variables predictoras de la regresión mediante las pruebas "t" para los parámetros de los coeficientes de la regresión. El método finaliza cuando ya no se pueden introducir o quitar justificadamente más predictores.

El estadístico "t" para cada coeficiente da como resultado un valor de "P" que determina la conveniencia de rechazar la hipótesis nula (H_0) en una prueba de hipótesis. En el caso de la regresión, la H_0 establece que el coeficiente es igual a cero y, por lo tanto, no es un coeficiente importante en la ecuación de regresión, mientras que la hipótesis alternativa (H_1) establece que no es igual a cero y, por lo tanto, es un coeficiente que debe incluirse en la ecuación de regresión.

Los valores "P" varían de 0 a 1 y se establece un límite a su valor o nivel de significación como criterio para rechazar o aceptar la H_0 conocidos como α_E y α_R . El valor "P" se calcula a partir de la muestra observada y representa la probabilidad de rechazar incorrectamente la H_0 cuando en realidad es verdadera (error tipo I). En otras palabras, es la probabilidad de obtener una diferencia al menos tan grande como la que existe entre el valor observado y el valor hipotético a través del error aleatorio solo. En el anexo 7.4 se detalla el procedimiento del método stepwise así como un ejemplo utilizando los datos de un sistema real registrados por Brownlee [44].

2.4.2. Consideraciones importantes sobre el método stepwise

No obstante las ventajas del método stepwise, es importante considerar que en algunos casos pudieran presentarse algunos inconvenientes [45]:

- Es posible que modelo final de regresión no sea el óptimo. El procedimiento da como resultado una sola ecuación de regresión, aunque existan otras ecuaciones con una capacidad de predicción similar o inclusive mejor a la encontrada.
- La regresión escalonada no tiene en cuenta el conocimiento previo del investigador sobre los predictores. Puede ser necesario forzar el procedimiento para incluir predictores que se consideren importantes. Es decir, modificar significativamente los valores de α_E y α_R , aunque no sean estadísticamente sólidos.
- El orden en el cuál se seleccionan los predictores carece de significado físico.
- Las pruebas "t" para probar que el valor de "P" sea menor que α_E se llevan a cabo en un procedimiento de regresión escalonada. Por lo tanto, la probabilidad de incluir algunos predictores sin importancia o excluir algunos predictores importantes no es baja.

2.5. REGRESIÓN SOBRE COMPONENTES PRINCIPALES (RCP)

Las variables predictoras que provienen de mediciones sucesivas del mismo sistema en las mismas unidades presentarán cierto grado de multicolinealidad de origen. Una manera de disminuir este problema es realizar un análisis de componentes principales a la matriz de predictores \mathbf{X} y realizar la regresión lineal múltiple de \mathbf{Y} sobre los componentes principales de la matriz \mathbf{X} . Los componentes principales no se encuentran correlacionados, por lo tanto, se elimina el problema de las variables linealmente dependientes. Otra ventaja de los componentes principales es la reducción de variables; éstas son transformadas en otras variables que explican la mayor parte de la variación con una menor cantidad de datos. Esta técnica parte de la idea de que existe un gran número de variables a las que se les denomina dimensiones, entonces de un gran conjunto de variables se buscan aquellas que estén muy relacionadas y las une para formar una nueva dimensión o una nueva variable que, a su vez, tiene nuevos valores llamados puntajes o puntuaciones.

Estas variables nuevas pueden denominarse como características, factores o componentes. La técnica de factores analiza las relaciones existentes entre un gran número de variables con el fin de identificar grupos afines o similares ya que toman las similitudes que hay en la estructura interna de las variables para tratar de explicarlas por medio de estas nuevas dimensiones o factores o componentes. Desde un punto de vista matemático, es una combinación lineal de las variables " x " originales representada por la ecuación 19.

$$C_i = l_1x_1 + l_2x_2 + l_3x_3 + \dots + l_kx_k \quad (19)$$

Donde:

- C_i es el factor o componente i de la observación k .
- l representa el peso que cada variable tiene con respecto al factor encontrado.
- i es el número del factor.
- k es el número de la variable.

2.5.1. Cálculo de los componentes principales

Los componentes principales (CPs) se obtienen de la matriz de covarianza. El término "covarianza" hace referencia a la medida de la varianza conjunta de dos variables. Para el cálculo de la covarianza entre dos variables " x_1 " y " x_2 " se utiliza la ecuación 20:

$$Cov(x_1, x_2) = \frac{\sum_{i=1}^2 \sum_{j=1}^n x_{1i}x_{2j} - n\bar{x}_1\bar{x}_2}{n - 1} \quad (20)$$

2. Regresión multivariante

De tal modo que se construye una matriz de covarianzas \mathbf{A} para las variables involucradas en el cálculo de CPs.

$$\mathbf{A} = Cov[\mathbf{X}] = \begin{bmatrix} Cov(x_1, x_1) & Cov(x_2, x_1) & \cdots & Cov(x_n, x_1) \\ Cov(x_2, x_1) & Cov(x_2, x_2) & \cdots & Cov(x_n, x_2) \\ \vdots & \vdots & \vdots & \vdots \\ Cov(x_1, x_n) & Cov(x_2, x_n) & \cdots & Cov(x_n, x_n) \end{bmatrix}$$

Una vez establecida la matriz de covarianzas se realiza una descomposición espectral $[\mathbf{A} - \lambda \mathbf{I}]$ donde:

- \mathbf{A} es la matriz de covarianzas.
- λ son los valores propios o eigenvalores.
- \mathbf{I} es la matriz identidad.

Se calculan los eigenvalores para la matriz cuadrada de covarianzas de tal modo que satisfagan la expresión 21:

$$[\mathbf{A} - \lambda \mathbf{I}] = 0 \quad (21)$$

Es decir, para determinar los eigenvalores se calcula el determinante de la matriz y se iguala a cero; existirá más de un valor propio dependiendo del rango de la matriz. De las soluciones para cada eigenvalor se tendrá como resultado un vector propio o eigenvector. Para calcular los eigenvalores se desarrolla la ecuación 31 de la siguiente manera:

$$\begin{bmatrix} Cov(x_1, x_1) & Cov(x_2, x_1) & \cdots & Cov(x_n, x_1) \\ Cov(x_1, x_2) & Cov(x_2, x_2) & \cdots & Cov(x_n, x_2) \\ \vdots & \vdots & \vdots & \vdots \\ Cov(x_1, x_n) & Cov(x_2, x_n) & \cdots & Cov(x_n, x_n) \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} = 0$$

En primer lugar, se multiplica el eigenvalor por la matriz identidad:

$$\begin{bmatrix} Cov(x_1, x_1) & Cov(x_2, x_1) & \cdots & Cov(x_n, x_1) \\ Cov(x_2, x_1) & Cov(x_2, x_2) & \cdots & Cov(x_n, x_2) \\ \vdots & \vdots & \vdots & \vdots \\ Cov(x_1, x_n) & Cov(x_2, x_n) & \cdots & Cov(x_n, x_n) \end{bmatrix} - \begin{bmatrix} \lambda & 0 & \cdots & 0 \\ 0 & \lambda & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda \end{bmatrix} = 0$$

Posteriormente, se hace la sustracción de las matrices:

$$\begin{bmatrix} Cov(x_1, x_1) - \lambda & Cov(x_2, x_1) & \cdots & Cov(x_n, x_1) \\ Cov(x_2, x_1) & Cov(x_2, x_2) - \lambda & \cdots & Cov(x_n, x_2) \\ \vdots & \vdots & \vdots & \vdots \\ Cov(x_1, x_n) & Cov(x_2, x_n) & \cdots & Cov(x_n, x_n) - \lambda \end{bmatrix} = 0$$

A continuación se calcula el determinante de la matriz resultante $[\mathbf{A} - \lambda \mathbf{I}]$ y se iguala a 0, dando como resultado un polinomio de grado ("n") con el mismo número de raíces o eigenvalores ($\lambda_1, \lambda_2, \lambda_3, \dots, \lambda_n$) de acuerdo a la ecuación 22.

$$\det[\mathbf{A} - \lambda \mathbf{I}] = |[\mathbf{A} - \lambda \mathbf{I}]| = a\lambda^{n-1} + b\lambda^{n-2} + \dots + c\lambda + d \quad (22)$$

Para encontrar los eigenvectores ($V_1, V_2, V_3, \dots, V_n$) se sustituyen los eigenvalores en la ecuación 23.

$$[\mathbf{A} - \lambda \mathbf{I}]\mathbf{V} = 0 \quad (23)$$

Cabe mencionar que los eigenvectores se obtienen a partir de la resolución de un sistema homogéneo de ecuaciones en las que la que el número de soluciones posibles puede ser infinito. Se debe por lo tanto asignar un valor arbitrariamente a alguno de los componentes del eigenvector y resolver para los restantes. Por lo tanto, si se sustituye en valor del primer eigenvalor λ_1 se tiene el eigenvector V_1 .

$$\begin{bmatrix} Cov(x_1, x_1) - \lambda_1 & Cov(x_2, x_1) & \cdots & Cov(x_n, x_1) \\ Cov(x_2, x_1) & Cov(x_2, x_2) - \lambda_1 & \cdots & Cov(x_n, x_2) \\ \vdots & \vdots & \vdots & \vdots \\ Cov(x_1, x_n) & Cov(x_2, x_n) & \cdots & Cov(x_n, x_n) - \lambda_1 \end{bmatrix} \begin{bmatrix} v_{1,1} \\ v_{1,2} \\ \vdots \\ v_{1,n} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Se aplica el mismo procedimiento para el cálculo de los eigenvectores restantes. Se requiere también que los eigenvectores sean ortogonales y que además estén normalizados (ortonormales) por lo que se divide el vector entre la norma del vector $||\mathbf{V}||$, que se define como la raíz cuadrada del producto del vector transpuesto y el mismo vector (ecuación 24). Para el vector V_1 se tiene que:

$$||\mathbf{V}_1|| = \sqrt{\mathbf{V}_1^T \mathbf{V}_1} = \sqrt{[v_{1,1} \quad v_{1,2} \quad \cdots \quad v_{1,n}] \begin{bmatrix} v_{1,1} \\ v_{1,2} \\ \vdots \\ v_{1,n} \end{bmatrix}} \quad (24)$$

Por lo tanto, el vector ortonormal o componente principal quedaría como:

$$CP_1 = \begin{bmatrix} \frac{v_{1,1}}{||\mathbf{V}_1||} \\ \frac{v_{1,2}}{||\mathbf{V}_1||} \\ \vdots \\ \frac{v_{1,n}}{||\mathbf{V}_1||} \end{bmatrix}$$

Los componentes principales quedan como se indica en la tabla 3:

Tabla 3. Cálculo de componentes principales.

Variable	CP_1	CP_2	...	CP_n
x_1	$\frac{v_{1,1}}{ \mathbf{V}_1 }$	$\frac{v_{2,1}}{ \mathbf{V}_2 }$...	$\frac{v_{n,1}}{ \mathbf{V}_n }$
x_2	$\frac{v_{1,2}}{ \mathbf{V}_1 }$	$\frac{v_{2,2}}{ \mathbf{V}_2 }$...	$\frac{v_{n,2}}{ \mathbf{V}_n }$
\vdots	\vdots	\vdots	...	\vdots
x_n	$\frac{v_{1,n}}{ \mathbf{V}_1 }$	$\frac{v_{2,n}}{ \mathbf{V}_2 }$...	$\frac{v_{n,n}}{ \mathbf{V}_n }$

2.5.2. Transformación de las variables predictoras mediante componentes principales

Ya definidos los componentes principales, las variables originales representadas por la matriz \mathbf{X} son transformadas a las variables no correlacionadas representadas por la matriz \mathbf{Z} calculadas mediante la expresión 25.

$$z_{1,i} = x_{1,i}CP_{1,1} + x_{2,i}CP_{1,2} + \dots + x_{n,i}CP_{1,n} \quad (25)$$

Estas nuevas variables también pueden calcularse mediante la expresión matricial 26, denotando la matriz de componentes principales como \mathbf{C}_p .

$$\mathbf{Z} = \mathbf{X}\mathbf{C}_p \quad (26)$$

El número de variables "z" dependerá de la varianza explicada acumulada de los componentes principales, y generalmente es menor a las variables "x" originales. Por lo tanto, las variables respuesta "y" dependerán de las variables "z" y el modelo de la regresión es el mismo que el de RLM dando como resultado la ecuación 27.

$$y_i = b_0 + b_1z_{i1} + b_2z_{i2} + \dots + b_mz_{im} + \varepsilon_i \quad (27)$$

2.6. REGRESIÓN SOBRE MÍNIMOS CUADRADOS PARCIALES (MCP)

Este método también utiliza la transformación de la matriz de predictores \mathbf{X} , sin embargo, a las variables que muestran una alta correlación se les asigna un peso extra en el polinomio de regresión debido a que serán más efectivas en la predicción. De esta forma se eligen combinaciones lineales de las variables predictoras que estén altamente correlacionadas con las variables de respuesta y también expliquen la variación en las variables predictoras [46].

En general, la regresión sobre mínimos cuadrados parciales (MCP) consta de dos pasos fundamentales. Primero, transforma la matriz de variables predictoras \mathbf{X} de orden $n \times p$, con la ayuda del vector de respuestas \mathbf{Y} de orden $n \times 1$, en una matriz \mathbf{T} de componentes o variables latentes no correlacionadas $T_1, T_2, T_3, \dots, T_p$ de orden $n \times p$, llamados componentes MCP; esto contrasta con el análisis de componentes principales en el cual los componentes son obtenidos usando sólo la matriz de predictores \mathbf{X} . Segundo, calcula el modelo de regresión estimado usando el vector de respuestas original y como predictores, los componentes MCP. La reducción de la dimensionalidad puede ser aplicada directamente sobre los componentes ya que estos son ortogonales. El número de componentes necesarios para el análisis de regresión puede ser menor que el número de predictores.

2.6.1. Algoritmo de la regresión MCP

La regresión MCP utiliza un algoritmo para determinar las componentes MCP denominado “Mínimos Cuadrados Parciales Iterativos No Lineales” (MCPINL) [47]. La entrada de datos corresponde a la matriz \mathbf{X} y el vector de respuesta \mathbf{Y} , previamente estandarizados por columnas. Cabe mencionar que en caso de existir dos o más variables de respuesta, el algoritmo MCPINL solo se debe aplicar independientemente para cada una para calcular sus respectivos coeficientes de regresión. A continuación se describen los pasos principales de este algoritmo:

Paso 1: Se estandarizan los datos de la matriz de predictores \mathbf{X} y el vector de respuesta \mathbf{Y} dando como resultado las matrices, \mathbf{X}^E y \mathbf{Y}^E respectivamente.

Paso 2: Se inicia el cálculo del primer componente T_1 .

Paso 3: Se calcula el primer vector de covarianzas entre las variables predictoras y la variable de respuesta $cov(\mathbf{Y}^E, \mathbf{X}^E)$ de orden $n \times 1$; cada elemento del vector es la covarianza de la variable de respuesta con cada variable predictora. Finalmente, el vector de covarianzas es ortonormalizado para dar lugar al vector \mathbf{w}_1 .

Paso 4: Se calcula el primer componente MCP mediante la ecuación 28.

$$T_1 = \mathbf{X}^E \mathbf{w}_1 \quad (28)$$

2. Regresión multivariante

Paso 5: Se calcula el coeficiente de regresión simple para el primer componente MCP, denotado como " u_1 ", de la variable de respuesta con respecto al vector de componente T_1 (Ecuación 29) calculado en el paso 4.

$$u_1 = \frac{T_1^T \mathbf{Y}^E}{T_1^T T_1} \quad (29)$$

Paso 6: Se calcula el vector \mathbf{b}_1 mediante la ecuación 30. Cada elemento de este vector es el coeficiente de regresión simple de \mathbf{X}^E respecto de la variable latente T_1 .

$$\mathbf{b}_1 = \frac{T_1^T \mathbf{X}^E}{T_1^T T_1} \quad (30)$$

Paso7: Se actualiza la matriz de variables predictoras empleando la expresión 31.

$$\mathbf{X}_1^{MCP} = \mathbf{X}^E - T_1 \mathbf{b}_1 \quad (31)$$

Paso 8: Se actualiza el vector de respuesta, mediante la ecuación 42.

$$\mathbf{Y}_1^{MCP} = \mathbf{Y}^E - T_1 u_1 \quad (32)$$

Paso 9: Se calculan los siguientes componentes MCP a partir del paso 3 con los datos de los pasos 7 y 8.

Los coeficientes de regresión normalizados, representados por la matriz $\mathbf{b}_{i,k}^E$, de la ecuación de regresión para los datos normalizados son calculados a partir de la matriz \mathbf{W} formada por los vectores de covarianzas ortonormalizados $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_n$ y la matriz de coeficientes de regresión simple \mathbf{b} , formada por los vectores de coeficientes de regresión b_1, b_2, \dots, b_n , entre los predictores y las variables latentes con la expresión 33:

$$\mathbf{b}_{i,k}^E = \mathbf{W}(\mathbf{b}\mathbf{W})^{-1}\mathbf{U} \quad (33)$$

Donde \mathbf{U} es la matriz de cargas de \mathbf{Y}^E formada por los valores de coeficientes u_1, u_2, \dots, u_n . Estos coeficientes de correlación $\mathbf{b}_{i,k}^E$ forman una ecuación de regresión que predice nuevos valores a partir de los datos estandarizados de acuerdo con la ecuación 44.

$$y_i^E = b_1^E x_{i,1}^E + b_2^E x_{i,2}^E + \dots + b_m^E x_{i,k}^E + \varepsilon_i \quad (34)$$

Para transformar estos coeficientes para que sean aplicables a los datos sin estandarizar se emplea la expresión 35:

$$\mathbf{b}_{i,k} = \mathbf{b}_{i,k}^E \left(\frac{\sigma_{y_i}}{\sigma_{x_k}} \right) \quad (35)$$

Finalmente, para encontrar el término independiente se tiene que:

$$b_0 = \bar{y}_i - \sum \bar{x}_k \mathbf{b}_{i,k} \quad (36)$$

Por lo tanto, la ecuación de regresión obtenida por mínimos cuadrados parciales queda como:

$$y_i = b_0 + b_1 x_{i1} + b_2 x_{i2} + \dots + b_m x_{ik} + \varepsilon_i \quad (37)$$

2.7. PARÁMETROS ANALÍTICOS DE LA REGRESIÓN LINEAL MÚLTIPLE

En química analítica, regresión lineal simple ha sido ampliamente estudiada y se han estipulado criterios estadísticos que permiten diferenciar entre una señal detectable y una señal que permite la cuantificación del analito. Estos parámetros analíticos son conocidos como límite de detección (LOD) y límite de cuantificación (LOQ) respectivamente. No obstante, en regresión multivariante aún no se establecen criterios estadísticos universales para determinar si una señal es detectable o medible.

En muchos trabajos que emplean regresión multivariante, como los que se encuentran en la tabla 8, es práctica común reportar el porcentaje de error relativo del analito ($\% E_R$) de acuerdo a la ecuación 38. Se considera que los porcentajes de error inferiores al 10 % son aceptables para este tipo de cuantificaciones.

$$\%E_R = 100 \left(\frac{C_t - C_p}{C_t} \right) \quad (38)$$

Donde:

- C_t : Es la concentración de trabajo o estándar del analito.
- C_p : Es la concentración pronosticada o calculada del analito empleando la ecuación de regresión.

Otra forma de dar validez a estas determinaciones es graficando las concentraciones estándar y las concentraciones calculadas de los analitos; de este modo es posible calcular el coeficiente de correlación entre ambas concentraciones. Un coeficiente de correlación cercano a 1 indica que el modelo predice adecuadamente las concentraciones de los analitos, a medida que este valor disminuye, también disminuye la capacidad de predicción de la ecuación de regresión.

2.8. CONSTRUCCIÓN DE LA MATRIZ DE RESPUESTA

En la sección 2.4. se ha discutido la importancia de seleccionar adecuadamente aquellas variables predictoras que mejor se ajusten a las variables de respuesta. En la regresión lineal se realiza la curva de calibración con soluciones de concentración creciente del analito en intervalos regulares; sin embargo, en el caso de la regresión multivariante donde se preparan mezclas de dos o más analitos no se ha establecido un criterio universal para definir las proporciones entre las concentraciones de las mezclas de calibrado, ni el número adecuado de estas mezclas.

Tabla 4. Ejemplos de matrices de concentración empleados en regresión multivariante.

Analitos	Modelo de regresión multivariante	Matriz de Diseño	Número de estándares	Ref.
Aspirina Acetaminofén Cafeína	RMP	Diseño factorial 2^3 (2 niveles, 3 factores)	8	[48]
Rifampicina Isoniazida Pirazinamida	RLM, RMP	Diseño central compuesto (3 factores)	15	[49]
Atrazina Cianacina	RLM, RCP, RMP	Arreglo ortogonal L_9 (3^2)	9	[50]
Acesulfamo-K Aspartamo Esteviósido	RMP	Arreglo ortogonal L_{16} (4^3)	16	[51]
Teofilina Montelukast Loratadina	RCP, RMP	Diseño central compuesto (3 factores)	15	[52]
Ácido ascórbico Dopamina Ácido úrico	RCP, RMP	Diseño de mezclas (Simplex Lattice)	36	[53]

Se pueden preparar estas mezclas de concentraciones al azar, dentro de un intervalo razonable, generalmente establecido por calibraciones univariantes realizadas con anterioridad. Por otra parte, se ha reportado el empleo de matrices de respuesta (o de concentración en determinaciones analíticas) generadas por diseños experimentales y de superficie de respuesta. Estos diseños permiten explicar la mayor cantidad de varianza del sistema con la realización de una mínima cantidad de experimentos. Si las concentraciones de los analitos toman el lugar de los factores se tiene una matriz de respuesta con la suficiente cantidad de mezclas de analitos en diferentes proporcio-

nes para ser analizadas mediante regresión multivariante; esta estrategia ha sido ampliamente utilizada en esquemas de cuantificación espectroscópica [43]. En la tabla 4 se muestran algunos ejemplos, donde se puede observar que dependiendo del tipo de diseño se tienen que preparar entre 8 y 36 mezclas de concentraciones para muestras que contenga tres analitos.

2.9. OBJETIVOS

2.9.1. Objetivo General

Desarrollar esquemas de cuantificación multiparamétrica mediante voltamperometría lineal y herramientas quimiométricas de regresión multivariante aplicadas a la determinación simultánea de analitos de interés biomédico.

2.9.2. Objetivos específicos

1. Determinación simultánea de paracetamol y naproxeno en un medicamento mediante voltamperometría lineal y regresión multivariante.
2. Determinación simultánea de sulfato ferroso y ácido ascórbico en un medicamento mediante voltamperometría lineal y regresión multivariante.
3. Simulación de respuesta de un biosensor multienzimático y multiparamétrico para la cuantificación simultánea de glucosa, colesterol y triglicéridos.
4. Determinación simultánea de glucosa y ácido ascórbico en un medicamento con un biosensor de glucosa mediante voltamperometría lineal y regresión multivariante.

3. APLICACIÓN DE LA REGRESIÓN MULTIVARIANTE EN SISTEMAS ELECTROQUÍMICOS

El enfoque analítico clásico consiste en desarrollar metodologías para la determinación de un único analito dentro de matrices complejas; en este sentido, aspectos como la selectividad son imprescindibles o requieren el establecimiento de procesos de separación complicados, el uso de reactivos costosos y equipos sofisticados. En este contexto estas metodologías resultan poco prácticas y en muchas ocasiones no llegan a ser asequibles.

El análisis simultáneo disminuye la cantidad de disoluciones estándar empleadas para la calibración individual de cada analito y permite reducir el tiempo de análisis de la muestra. En este capítulo se reportan diversos análisis simultáneos de analitos electroactivos mediante métodos electroquímicos y regresión multivariante.

3.1.DETERMINACIÓN SIMULTÁNEA DE NAPROXENO Y PARACETAMOL EN MEDICAMENTOS MEDIANTE VOLTAMPEROMETRÍA Y REGRESIÓN MULTIVARIANTE

3.1.1. Introducción

Gran parte de los medicamentos disponibles en el mercado contiene mezclas de principios activos en sus formulaciones. En este sentido, tanto el naproxeno como el paracetamol suelen encontrarse combinados en diversas proporciones en forma de pastillas, supositorios y suspensiones.

Tanto el naproxeno (protonado o sódico) como el paracetamol pertenecen a la familia de medicamentos antiinflamatorios no esteroideos (AINEs) y se utilizan para aliviar dolores musculares, migraña y otros síntomas característicos de enfermedades infecciosas [54]. Sus propiedades electroquímicas los hacen susceptibles para ser determinados mediante voltamperometría (Fig. 11).

Diversos estudios electroquímicos reportan que el paracetamol tiene un solo producto de oxidación irreversible [55]; sin embargo, el primer producto de oxidación del naproxeno es un radical inestable que se adsorbe sobre la superficie de electrodos de carbono en medios acuosos [56]. Este compuesto, con el tiempo, forma otros productos de ruptura oxidativa que pueden detectarse por voltamperometría [57]. Existen metodologías que permiten cuantificar ambos fármacos de manera independiente así como de manera simultánea empleando técnicas cromatográficas y espectroscópicas [58], [59], [60].

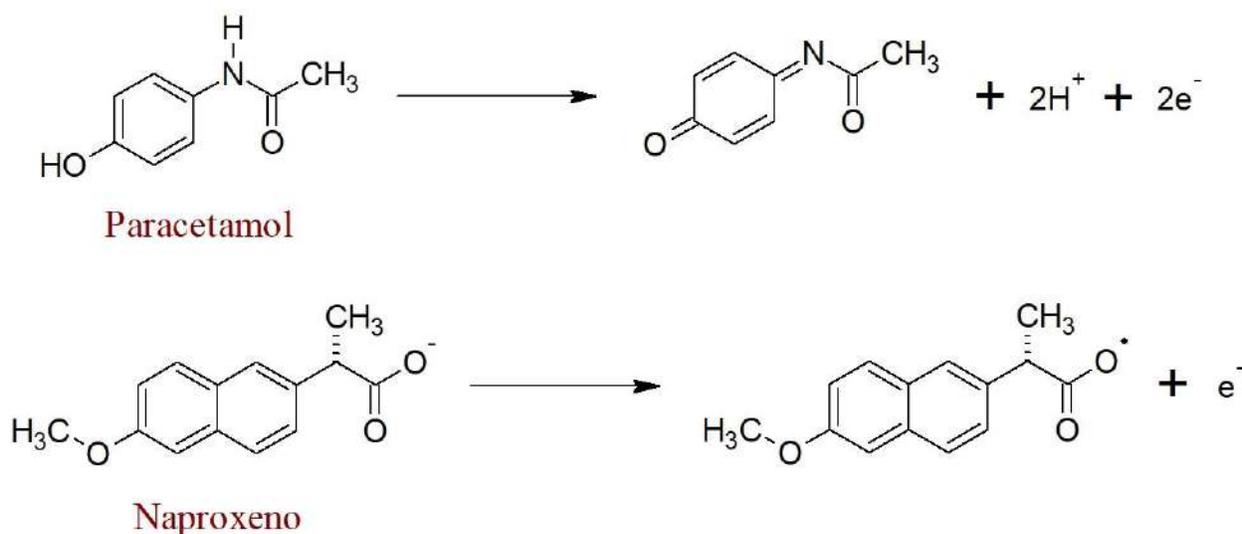


Figura 11. Estructuras de paracetamol, naproxeno y sus productos de oxidación.

Alternativamente en este trabajo se explora la posibilidad de hacer la determinación simultánea de naproxeno sódico y paracetamol en un fármaco mediante voltamperometría lineal empleando los modelos quimiométricos de regresión lineal múltiple (RLM), regresión sobre componentes principales (RCP) y regresión sobre mínimos cuadrados parciales (MCP).

3. Aplicación de la regresión multivariante en sistemas electroquímicos

3.1.2. Experimentación

Todas las soluciones fueron preparadas con reactivos comerciales de grado analítico: naproxeno sódico, paracetamol, perclorato de litio y acetonitrilo (Sigma-Aldrich), metanol (J.T. Baker). El electrolito soporte fue preparado a una concentración de LiClO_4 0.1 M en una disolución de metanol en acetonitrilo al 30 % (V/V). Las soluciones estándar de naproxeno y paracetamol, así como de la muestra fueron preparadas en el electrolito soporte.

Para la preparación de la muestra se pesaron cinco tabletas de una formulación de naproxeno-paracetamol comercial (275 mg y 300 mg por tableta respectivamente) y se pulverizaron. Posteriormente, se pesaron 175.2 mg de la muestra pulverizada y se disolvieron en 20.0 mL. La solución se ultrasonizó durante 20 minutos y se filtró a gravedad con un papel de poro fino (Wathman 42) para retirar el excipiente. Finalmente, la solución filtrada fue aforada a 25.0 mL.

Los voltamperogramas de las disoluciones fueron obtenidos mediante un potenciostato-galvanostato AUTOLAB PGSTAT 302N (Metrohm) controlado mediante el software Nova 2.0. La celda electroquímica se compone de tres electrodos: dos electrodos de platino MF-2013 con un diámetro de 1.6 mm (2.01 mm^2) como electrodo de trabajo y auxiliar respectivamente y un electrodo de plata sólida MF-2

011 (BASi) con un diámetro de 1.6 mm (2.01 mm^2) como electrodo de referencia. Se hicieron los respectivos voltamperogramas lineales de los estándares así como de las muestras, partiendo del potencial de corriente nula hasta 1.60 V a una velocidad de barrido de 10 mVs^{-1} .

Los datos obtenidos son organizados en una matriz de datos para su tratamiento estadístico mediante el procedimiento algebraico matricial de los métodos de regresión multivariante y empleando el software Minitab 17.

En esta sección se detalla paso a paso el procedimiento algebraico matricial de los métodos de regresión multivariante, posteriormente, se realizan los mismos cálculos empleando el software Minitab 17. La matriz se alimenta con los datos de las concentraciones de los analitos en las mezclas estándar y con las intensidades de corriente a los potenciales seleccionados. Posteriormente, se utilizan los modelos de regresión multivariante en modalidad de calibración inversa para generar las ecuaciones que predigan la concentración de una muestra problema. Las concentraciones de trabajo son predichas nuevamente (concentraciones pronosticadas) para determinar el error promedio del método.

3.1.3. Resultados y discusión

En primer lugar se realizaron, a partir de voltamperometría cíclica, las curvas de calibrado para cada analito, tanto de naproxeno como de paracetamol para evaluar su comportamiento electroquímico y la proporcionalidad de la intensidad de pico con respecto a la concentración.

De acuerdo a los voltamperogramas cíclicos hechos a las disoluciones de concentración creciente de naproxeno, este analito se oxida a un potencial de 1.4 V; proceso que corresponde a la reacción mostrada en la Figura 11 (Fig. 12).

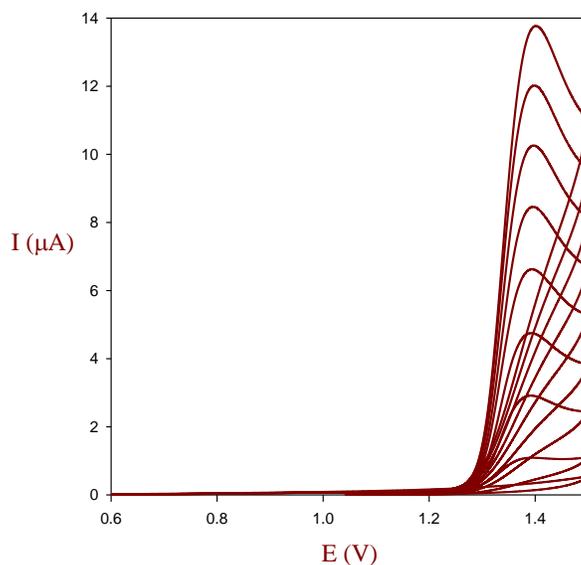


Figura 12. Voltamperogramas de naproxeno sódico en metanol en acetonitrilo/metanol al 30 % (v/v). $\nu = 10 \text{ mVs}^{-1}$.

Por otra parte, en los voltamperogramas cíclicos hechos a las disoluciones de concentración creciente de paracetamol, se observa que este analito se oxida a un potencial aproximado de 1.2 V con un desplazamiento anódico conforme aumenta su concentración. El proceso de oxidación corresponde a la reacción mostrada en la Figura 11 (Fig. 13).

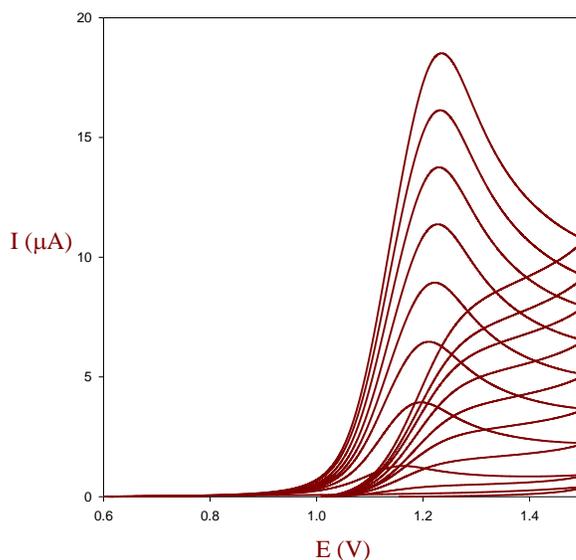


Figura 13. Voltamperogramas de paracetamol en metanol en acetonitrilo/metanol al 30 % (v/v). $\nu = 10 \text{ mVs}^{-1}$.

3. Aplicación de la regresión multivariante en sistemas electroquímicos

A partir de los voltamperogramas obtenidos para ambos analitos se confirma una proporcionalidad directa entre la intensidad de pico y la concentración (Fig. 14, Tab. 5).

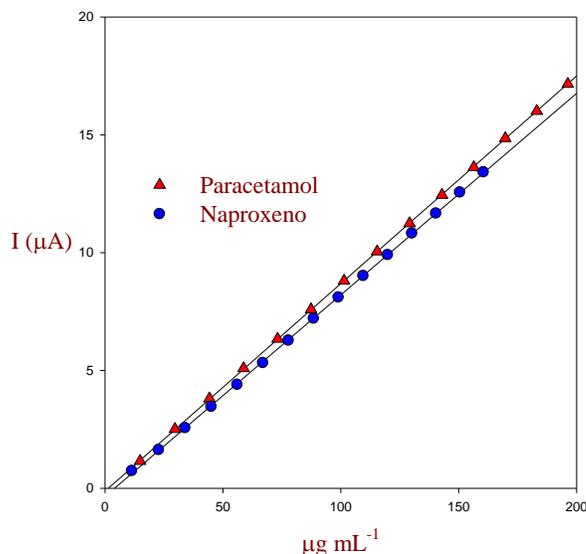


Figura 14. Curvas de calibración lineal para el naproxeno sódico y el paracetamol en acetonitrilo/metanol al 30 % (v/v). $v = 10 \text{ mVs}^{-1}$.

Los coeficientes de correlación de ambos analitos son cercanos a la unidad, lo que confirma la linealidad del sistema en el intervalo de concentraciones seleccionado. Se presenta una sensibilidad similar para ambos analitos a pesar de tener potenciales de oxidación y un número de electrones intercambiados diferente.

Tabla 5. Parámetros analíticos de las calibraciones lineales de naproxeno sódico y paracetamol.

Parámetro	Magnitud	
	Naproxeno	Paracetamol
IRL ($\mu\text{g mL}^{-1}$)	11.4-160.5	14.9-209.3
b ($\mu\text{A } \mu\text{g}^{-1} \text{ mL}$)	0.09 ± 0.00	0.09 ± 0.00
a ($\mu\text{g mL}^{-1}$)	-0.35 ± 0.04	-0.12 ± 0.03
$r_{x,y}^2$	0.999942	0.999988
LOD ($\mu\text{g mL}^{-1}$)	1.6	0.9
LOQ ($\mu\text{g mL}^{-1}$)	5.3	3.2

IRL = Intervalo de respuesta lineal

Con base en los resultados anteriores y para realizar el calibrado multivariante, se prepararon diez mezclas estándar al azar en un intervalo de concentración de cada analito entre 40.0 y 220.0 $\mu\text{g/mL}$ (Tabla 6). Cada mezcla fue analizada mediante un barrido voltamperométrico en modalidad lineal (Fig. 15). Para el análisis multivariante es necesario construir una matriz de datos sobrestimada, por lo tanto, se seleccionan tres potenciales donde se encuentran los máximos de

intensidad de corriente (0.65, 1.00 y 1.40 V) tabulados en la Tabla 6. Esta matriz de datos fue procesada con los modelos de RLM, RCP y MCP.

Tabla 6. Mezclas de concentraciones de los estándares de naproxeno y paracetamol y su respuesta en intensidad de corriente a los potenciales seleccionados.

Estándar	Concentraciones ($\mu\text{g mL}^{-1}$)		I (μA)		
	Naproxeno	Paracetamol	0.65 V	1.00 V	1.40 V
M ₁	201.6	141.7	3.74	10.72	37.99
M ₂	161.3	101.2	2.57	7.80	27.20
M ₃	100.8	121.4	1.73	8.81	24.84
M ₄	40.3	202.4	0.59	17.58	25.45
M ₅	80.6	222.6	1.24	18.48	29.42
M ₆	141.1	60.7	2.51	3.99	28.89
M ₇	121.0	81.0	2.24	5.55	26.00
M ₈	181.4	40.5	2.78	3.09	30.69
M ₉	221.8	161.9	4.37	12.63	38.12
M ₁₀	60.5	182.2	1.11	15.58	25.74

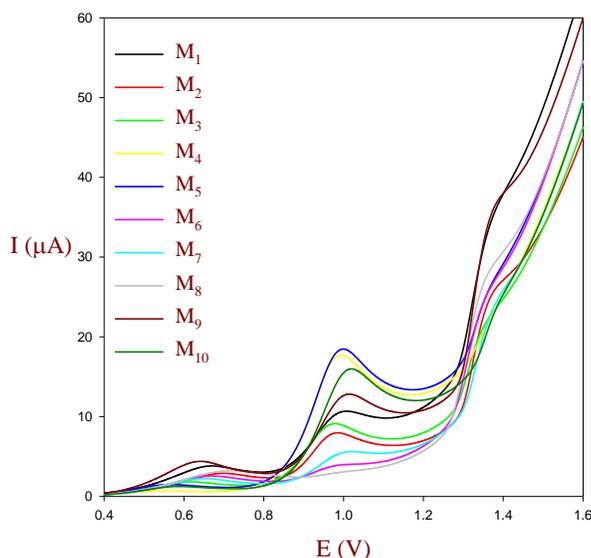


Figura 15. Voltamperogramas lineales de mezclas de paracetamol y naproxeno sódico en acetonitrilo/metanol al 30 % (v/v). $v = 10 \text{ mV s}^{-1}$.

3.1.3.1. Regresión lineal múltiple (RLM)

En el esquema de calibración inversa para la RLM, las variables de respuesta "y" son las concentraciones de naproxeno y paracetamol en las soluciones de concentración estándar, mientras que las variables predictoras "x" son los valores de intensidad de corriente a los potenciales seleccionados (0.65, 1.00 y 1.40 V).

3. Aplicación de la regresión multivariante en sistemas electroquímicos

A la matriz de respuestas \mathbf{X} se le añade una columna con valores de 1 para cada fila para estimación del intercepto b_0 . Las matrices de variables predictoras \mathbf{X} y de respuestas \mathbf{Y} , quedan como:

$$\mathbf{Y} = \begin{bmatrix} 201.6 & 141.7 \\ 161.3 & 101.2 \\ 100.8 & 121.4 \\ 40.3 & 202.4 \\ 80.6 & 22.6 \\ 141.1 & 60.7 \\ 121.0 & 81.0 \\ 181.4 & 40.5 \\ 221.8 & 161.9 \\ 60.5 & 182.2 \end{bmatrix} \quad \mathbf{X} = \begin{bmatrix} 1 & 3.74 & 10.72 & 37.99 \\ 1 & 2.57 & 7.80 & 27.20 \\ 1 & 1.73 & 8.81 & 24.84 \\ 1 & 0.59 & 17.58 & 25.45 \\ 1 & 1.24 & 18.48 & 29.42 \\ 1 & 2.51 & 3.99 & 28.89 \\ 1 & 2.24 & 5.55 & 26.00 \\ 1 & 2.78 & 3.09 & 30.69 \\ 1 & 4.37 & 12.63 & 38.12 \\ 1 & 1.11 & 15.58 & 25.74 \end{bmatrix}$$

Los datos reorganizados para las concentraciones de los analitos generan la matriz de respuestas \mathbf{Y} de modo que:

A partir de las matrices construidas y empleando la ecuación 9, se calculan los coeficientes de la ecuación de regresión:

$$\mathbf{B} = \begin{bmatrix} 11.17 & 2.49 & 0.20 & -0.64 \\ 2.49 & 0.93 & 0.09 & -0.19 \\ 0.20 & 0.09 & 0.01 & -0.02 \\ -0.64 & -0.19 & -0.02 & 0.04 \end{bmatrix} \begin{bmatrix} 1310.40 & 1315.60 \\ 3632.54 & 2751.64 \\ 12044.29 & 16750.12 \\ 40748.73 & 38815.40 \end{bmatrix} = \begin{bmatrix} 3.04 & 22.92 \\ 38.96 & 7.81 \\ -2.12 & 11.74 \\ 2.07 & -1.07 \end{bmatrix}$$

Por lo tanto, empleando el modelo de RLM, las ecuaciones de regresión para el cálculo de las concentraciones de naproxeno (C_{nap}) y de paracetamol (C_{par}) a partir de las intensidades de corriente a los potenciales seleccionados quedan como:

$$C_{nap} = 3.04 + 30.96(I_{0.65V}) - 2.12(I_{1.00V}) + 2.07(I_{1.40V})$$

$$C_{par} = 22.92 + 7.81(I_{0.65V}) + 11.74(I_{1.00V}) - 1.07(I_{1.40V})$$

Las concentraciones recalculadas o predichas partir de los coeficientes de regresión encontrados son denotadas por la matriz $\hat{\mathbf{Y}}$. Estas concentraciones pueden ser calculadas matricialmente a partir de la resolución de la ecuación 10.

Las concentraciones predichas y de las concentraciones estándar pueden ser comparadas mediante el cálculo del porcentaje de error relativo para cada par de datos mediante la ecuación 48. En términos generales se puede calcular el porcentaje de error promedio de todas las mediciones para cada analito a fin de que reconocer la calidad de las mediciones. Los resultados de % de error y el % de error promedio se muestran en la tabla 7.

$$\hat{Y} = XB = \begin{bmatrix} 204.71 & 137.21 \\ 142.97 & 105.37 \\ 103.21 & 113.20 \\ 41.43 & 206.59 \\ 73.07 & 217.98 \\ 152.23 & 58.35 \\ 132.41 & 77.66 \\ 168.39 & 47.96 \\ 225.46 & 164.41 \\ 66.54 & 186.87 \end{bmatrix}$$

Tabla 7. Concentraciones pronosticadas y porcentajes de error promedio en la determinación simultánea de naproxeno sódico y paracetamol mediante RLM.

Mezcla	Concentración estándar ($\mu\text{g mL}^{-1}$)		Concentración pronosticada ($\mu\text{g mL}^{-1}$)		% de Error	
	Naproxeno	Paracetamol	Naproxeno	Paracetamol	Naproxeno	Paracetamol
M ₁	201.6	141.7	204.7	137.2	1.5	3.2
M ₂	161.3	101.2	143.0	105.4	11.3	4.2
M ₃	100.8	121.4	103.2	113.2	2.4	6.8
M ₄	40.3	202.4	41.4	206.6	2.7	2.1
M ₅	80.6	222.6	73.1	218.0	9.3	2.1
M ₆	141.1	60.7	152.2	58.4	7.9	3.8
M ₇	121.0	81.0	132.4	77.7	9.4	4.1
M ₈	181.4	40.5	168.4	48.0	7.2	18.5
M ₉	221.8	161.9	225.5	164.4	1.7	1.5
M ₁₀	60.5	182.2	66.5	186.9	9.9	2.6
% de Error promedio					6.3	4.9

Los errores promedio para ambos analitos son inferiores al 10%, lo que indica que el método es adecuado para una determinación simultánea de ambos analitos. Sin embargo, para explicar detalladamente la idoneidad del modelo se realiza un análisis ANOVA a los parámetros de la regresión. A cada variable de respuesta le corresponde un modelo de regresión y también le corresponde un ANOVA por separado, en el caso de la regresión del naproxeno respecto a las intensidades de corriente a los potenciales seleccionados los resultados se muestran en la Tabla 8.

Tabla 8. Análisis ANOVA para la regresión de naproxeno sódico.

Fuente de la Variación	Suma de Cuadrados	Grados de Libertad	Cuadrados medios	F
Regresión	32655.606	3	10885.202	74.000
Error residual	882.578	6	147.096	
Total	33538.184	9	3726.465	
$I_{0.65V}$	1638.409	1	1638.409	11.138
$I_{1.00V}$	382.340	1	382.340	2.599
$I_{1.40V}$	101.135	1	101.135	0.688

3. Aplicación de la regresión multivariante en sistemas electroquímicos

El análisis ANOVA proporciona las fuentes de variación, sus grados de libertad, las sumas de cuadrados y los cuadrados medios. También incluye el estadístico F y se utiliza para determinar si los predictores o factores están significativamente relacionados con la respuesta. La suma total de cuadrados ayuda a expresar la variación total que puede atribuirse a varios factores.

Al comparar la suma de cuadrados de la regresión con la suma de cuadrados totales, se encuentra la proporción de la variación total que se explica por el modelo de regresión, este valor es mejor conocido como el coeficiente de determinación múltiple $R_{x,y}^2$. De acuerdo a los resultados de la tabla 8 se calcula el coeficiente de determinación múltiple empleando la ecuación 13. En este caso, $R_{x,y}^2$ tiene un valor cercano a 1 e indica que la ecuación de regresión calculada explica adecuadamente el comportamiento de las concentraciones de los analitos.

$$R_{x,y}^2 = \frac{32655.606}{33538.184} = 0.974$$

En la regresión, los cuadrados medios se utilizan para determinar si los términos en el modelo son significativos. Los cuadrados medios se obtienen al dividir la suma de cuadrados por los grados de libertad. Al comparar los cuadrados medios del error residual con los cuadrados medios totales, se obtiene el coeficiente de determinación ajustado R_{aj}^2 ; que representa el porcentaje de variación de la variable de respuesta que se explica por su relación con una o más variables predictoras, ajustado por el número de predictores en el modelo. Este ajuste es importante porque el $R_{x,y}^2$ para cualquier modelo siempre aumentará cuando se agregue un nuevo término. Un modelo con más términos puede parecer que se ajusta mejor simplemente porque tiene más términos. El R_{aj}^2 es una herramienta útil para comparar el poder explicativo de los modelos con diferentes números de predictores y aumentará su valor solo si el nuevo término mejora el modelo. En caso contrario disminuirá cuando un predictor no mejore el modelo de regresión.

$$R_{aj}^2 = 1 - \frac{147.096}{3726.465} = 0.961$$

Para esta regresión R_{aj}^2 tiene un valor cercano a 1 e indica que la ecuación de regresión calculada explica adecuadamente el comportamiento de las concentraciones de los analitos respecto a sus intensidades de corriente.

Ahora bien, para evaluar el grado en el que la ecuación de regresión podrá predecir las concentraciones de muestras que no forman parte de la línea de calibrado se calcula el coeficiente de correlación predicho mediante la estadística *SCER* cuyos resultados se muestran en la tabla 9.

Mientras que $R_{x,y}^2$ y R_{aj}^2 indican qué tan bien el modelo se ajusta a los datos. El coeficiente de regresión predicho, R_{pred}^2 , puede evitar el sobreajuste del modelo y puede ser más útil para comparar modelos de regresión porque se calcula utilizando observaciones no incluidas en la estimación. El sobreajuste se refiere a modelos que parecen explicar la relación entre el predictor y las variables de respuesta para el conjunto de datos utilizado para el cálculo del modelo, pero no proporcionan predicciones válidas para nuevas observaciones.

Tabla 9. Resultados de la estadística SCER para la regresión de naproxeno mediante MLR.

n	$(y_i - \hat{y}_i)^2$									
	i = 1	i = 2	i = 3	i = 4	i = 5	i = 6	i = 7	i = 8	i = 9	i = 10
1	37.768	37.590	7.167	9.696	0.708	27.997	3.137	9.802	54.96	7.246
2	359.974	841.622	305.894	337.902	287.773	310.125	199.816	293.96	232.412	303.06
3	2.520	26.951	10.342	6.711	6.195	15.636	38.162	3.088	6.575	13.686
4	1.333	3.597	2.079	3.568	12.721	1.249	1.207	2.714	0.107	12.182
5	41.537	19.643	57.035	46.771	164.019	60.522	80.878	59.259	52.0175	31.452
6	142.747	96.99	130.676	123.789	128.175	268.016	192.655	1.932	93.305	119.507
7	118.438	13.361	148.281	129.993	158.679	203.739	246.781	64.843	134.969	144.737
8	132.469	136.006	167.174	170.298	171.479	45.761	119.85	651.231	218.988	186.691
9	31.687	9.011	13.672	11.94	11.262	3.506	15.275	49.727	180.730	14.717
10	32.783	7.597	42.856	43.343	9.125	31.871	52.429	65.208	40.039	66.674

Por lo tanto:

$$SCERP = 37.768 + 841.622 + \dots + 66.674 = 2470.752$$

El valor de R_{Pred}^2 es cercano a 1 y sugiere que el modelo no se encuentra sobreajustado y que posee una gran capacidad de predicción para soluciones diferentes a las utilizadas como estándares.

$$R_{Pred}^2 = 1 - \frac{2470.752}{33538.184} = 0.926$$

En el caso el análisis ANOVA para la regresión del paracetamol respecto a las intensidades de corriente a los potenciales seleccionados los resultados se muestran en la Tabla 10.

Tabla 10. Análisis ANOVA para la regresión de paracetamol mediante RLM.

Fuente de la Variación	Suma de Cuadrados	Grados de Libertad	Cuadrados medios	F
Regresión	33544.4	3	11181.5	274.62
Error residual	244.3	6	40.7	
Total	33788.7	9	3754.2	
$I_{0.65V}$	66.0	1	66.0	1.62
$I_{1.00V}$	11676.6	1	11676.6	286.78
$I_{1.40V}$	27.2	1	27.2	0.67

De acuerdo a los datos de la tabla 10, los diversos coeficientes de correlación obtenidos se muestran en la tabla 11.

En el caso de la regresión de paracetamol el modelo arroja valores consistentemente más elevados que en la regresión de naproxeno. El valor de R_{Pred}^2 indica que no hay sobreajuste en el modelo, es decir, las matrices de concentración y la matriz de predictores fue seleccionada correctamente, lo que explica que el porcentaje de error promedio sea menor.

3. Aplicación de la regresión multivariante en sistemas electroquímicos

Tabla 11. Coeficientes de correlación de la regresión del paracetamol.

$R^2_{x,y}$	R^2_{aj}	R^2_{Pred}
0.993	0.989	0.979

3.1.3.2. Multicolinealidad del modelo

Por su naturaleza, es posible que los datos obtenidos a partir de análisis voltamperométricos presenten algún grado importante de multicolinealidad. Por lo tanto, se hace un estudio de multicolinealidad para conocer si existen variables correlacionadas que pudieran eliminarse de la ecuación de regresión sin afectar la calidad de las estimaciones.

Los resultados reportados en la tabla 12. Muestran que el primer potencial posee un FIV mayor, aunque cercano a 10, por lo tanto, este predictor se correlaciona con el resto de predictores, sin embargo, para que la matriz de respuestas sea sobreestimada, no es posible eliminarlo. El resto de predictores no son colineales entre sí.

Tabla 12. Parámetros de multicolinealidad entre las variables predictoras seleccionadas para la determinación simultánea de naproxeno y paracetamol.

Variable	Correlación	R^2_x	Tolerancia	FIV
$I_{0.65V}$	$I_{1.00V}, I_{1.40V}$	0.914	0.086	11.575
$I_{1.00V}$	$I_{0.65V}, I_{1.40V}$	0.693	0.307	3.259
$I_{1.40V}$	$I_{0.65V}, I_{1.00V}$	0.892	0.108	9.257

3.1.3.3. Regresión sobre componentes principales

Para eliminar la multicolinealidad de las variables predictoras y mejorar los parámetros estadísticos de la regresión, se realiza un análisis de componentes principales sobre la matriz de variables predictoras. En primer lugar se construye la matriz de covarianzas \mathbf{A} , calculadas a partir de la ecuación 20 con los valores de la matriz de intensidades de corriente a los diferentes potenciales mostrada en la tabla 6:

$$\mathbf{A} = Cov[\mathbf{X}] = \begin{bmatrix} 1.389 & -2.921 & 4.915 \\ -2.921 & 30.675 & -0.186 \\ 4.915 & -0.186 & 24.215 \end{bmatrix}$$

Al hacer la descomposición espectral empleando la ecuación 21 se tiene que:

$$[\mathbf{A} - \lambda \mathbf{I}] = \begin{bmatrix} 1.389 - \lambda & -2.921 & 4.915 \\ -2.921 & 30.675 - \lambda & -0.186 \\ 4.915 & -0.186 & 24.215 - \lambda \end{bmatrix} = 0$$

Para el cálculo de los valores propios se calcula el determinante de la expresión anterior con la ecuación 22:

$$\det(\mathbf{A}) = |\mathbf{A}| = -\lambda^3 + 56.278 \lambda^2 - 786.279 \lambda + 89.100$$

En la tabla 13 se muestran los valores propios calculados.

Tabla 13. Valores propios calculados para las variables de respuesta de las variables predictoras en la determinación simultánea de naproxeno y paracetamol.

Valor propio o Eigenvalor	Magnitud
λ_1	31.038
λ_2	25.126
λ_3	0.114

Para el cálculo del primer vector propio, se desarrolla la ecuación 23 con el primer valor propio:

$$\begin{bmatrix} 1.389 - \lambda_1 & -2.921 & 4.915 \\ -2.921 & 30.675 - \lambda_1 & -0.186 \\ 4.915 & -0.186 & 24.215 - \lambda_1 \end{bmatrix} \begin{bmatrix} v_{1,1} \\ v_{1,2} \\ v_{1,3} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Entonces:

$$\begin{bmatrix} -29.649 & -2.921 & 4.914 \\ -2.921 & -0.362 & -0.186 \\ 4.914 & -0.186 & -6.822 \end{bmatrix} \begin{bmatrix} v_{1,1} \\ v_{1,2} \\ v_{1,3} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Posteriormente, se llega a un sistema de ecuaciones homogéneo con un número potencialmente infinito de soluciones, por lo tanto, se deben resolver por el método de Gauss-Jordan; quedando la matriz como sigue:

$$\begin{bmatrix} 0 & -4.041 & -36.244 \\ -12.509 & 0 & 13.125 \\ 4.915 & -0.186 & -6.822 \end{bmatrix} \begin{bmatrix} v_{1,1} \\ v_{1,2} \\ v_{1,3} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Finalmente, es necesario asignar un valor a alguno de los valores del vector propio, por ejemplo, si $v_{1,3} = 1$ el resto de valores del vector queda como:

$$v_{1,1} = 1.049 v_{1,3}$$

$$v_{1,2} = -8.968 v_{1,3}$$

De ese modo, el vector propio calculado es:

3. Aplicación de la regresión multivariante en sistemas electroquímicos

$$V_1 = \begin{bmatrix} 1.049 \\ -8.968 \\ 1 \end{bmatrix}$$

Para el cálculo del primer componente principal se debe calcular la norma del vector mediante la ecuación 24:

$$\|V_1\| = \sqrt{[1.049 \quad -8.968 \quad 1] \begin{bmatrix} 1.049 \\ -8.968 \\ 1 \end{bmatrix}} = \sqrt{82.529} = 9.085$$

Por lo tanto, el vector ortonormal o componente principal es el cociente de cada valor del vector propio entre la norma del propio vector:

$$CP_1 = \begin{bmatrix} \left(\frac{1.049}{9.085}\right) \\ \left(\frac{-8.968}{9.085}\right) \\ \left(\frac{1.000}{9.085}\right) \end{bmatrix} = \begin{bmatrix} 0.115 \\ -0.987 \\ 0.110 \end{bmatrix}$$

En la tabla 14, se registran los resultados de los vectores propios ortonormales o componentes principales de la matriz predictora X . El procedimiento matricial para obtener los CP_2 y CP_3 es análogo al anterior y se omite.

Tabla 14. Componentes principales de la matriz X .

Variable	CP_1	CP_2	CP_3
$I_{(0.65V)}$	0.115	-0.186	0.975
$I_{(1.00V)}$	-0.987	-0.130	0.093
$I_{(1.40V)}$	0.110	-0.974	0.200

Una vez calculados los componentes principales se realiza la modificación de las variables predictoras " x " para dar lugar a nuevas variables no correlacionadas " z " mediante la ecuación 26. En algunos casos estas variables no pueden ser transformadas adecuadamente debido a que su colinealidad es alta. Los valores de estas nuevas variables se observan en la tabla 15.

$$\mathbf{Z} = \mathbf{XC}_p = \begin{bmatrix} -5.969 & -39.091 & -2.943 \\ -4.409 & -27.985 & -2.200 \\ -5.763 & -25.662 & -2.455 \\ -14.485 & -27.187 & -2.876 \\ -14.862 & -31.292 & -2.951 \\ -0.469 & -29.123 & -2.949 \\ -2.358 & -26.462 & -2.491 \\ 0.649 & -30.809 & -3.129 \\ -7.767 & -39.584 & -2.177 \\ -12.419 & -27.306 & -2.612 \end{bmatrix}$$

Tabla 15. Variables de respuesta y variables predictoras para la RCP para la determinación simultánea de naproxeno y paracetamol.

Estándar	Concentraciones ($\mu\text{g mL}^{-1}$)		Variables predictoras Z		
	Naproxeno	Paracetamol	z_1	z_2	z_3
M ₁	201.6	141.7	-5.969	-39.091	-2.943
M ₂	161.3	101.2	-4.409	-27.985	-2.200
M ₃	100.8	121.4	-5.763	-25.662	-2.455
M ₄	40.3	202.4	-14.485	-27.187	-2.876
M ₅	80.6	222.6	-14.862	-31.292	-2.951
M ₆	141.1	60.7	-0.469	-29.123	-2.949
M ₇	121	81	-2.358	-26.462	-2.491
M ₈	181.4	40.5	0.649	-30.809	-3.129
M ₉	221.8	161.9	-7.767	-39.584	-2.177
M ₁₀	60.5	182.2	-12.419	-27.306	-2.612

Las ecuaciones encontradas mediante RCP para determinar las concentraciones de naproxeno y paracetamol quedan como:

$$C_{nap} = 3.00 + 6.86(z_1) - 9.02(z_2) + 37.40(z_3)$$

$$C_{par} = 22.90 - 10.80(z_1) - 1.95(z_2) + 8.92(z_3)$$

Las concentraciones de los analitos en las mezclas son calculadas nuevamente; con estos datos es posible estimar el porcentaje de error promedio para regresión. Los resultados se muestran en la Tabla 16 y son similares a los obtenidos mediante RLM, donde los porcentajes de error promedio son inferiores al 10.0 % en ambos compuestos, lo que indica que las ecuaciones encontradas mediante RCP se ajustan al comportamiento de los datos.

Para conocer la idoneidad del modelo de regresión con las variables modificadas se realiza un análisis ANOVA de la regresión y se calculan los coeficientes de correlación ($R_{x,y}^2$, R_{aj}^2 y R_{Pred}^2), correspondientes para cada analito, así como los parámetros de multicolinealidad. Los resultados se muestran en la tabla 17.

3. Aplicación de la regresión multivariante en sistemas electroquímicos

Tabla 16. Concentraciones pronosticadas y porcentajes de error promedio en la determinación simultánea de naproxeno sódico y paracetamol mediante RCP.

Mezcla	Concentración estándar ($\mu\text{g mL}^{-1}$)		Concentración pronosticada ($\mu\text{g mL}^{-1}$)		% de Error	
	Naproxeno	Paracetamol	Naproxeno	Paracetamol	Naproxeno	Paracetamol
M ₁	201.6	141.7	204.7	137.2	1.5	3.2
M ₂	161.3	101.2	142.9	105.3	11.4	4.1
M ₃	100.8	121.4	103.2	113.2	2.4	6.8
M ₄	40.3	202.4	41.4	206.6	2.7	2.1
M ₅	80.6	222.6	73.0	217.9	9.4	2.1
M ₆	141.1	60.7	152.2	58.3	7.9	3.9
M ₇	121	81.0	132.4	77.6	9.4	4.2
M ₈	181.4	40.5	168.4	47.9	7.2	18.4
M ₉	221.8	161.9	225.4	164.4	1.6	1.5
M ₁₀	60.5	182.2	66.5	186.8	9.9	2.5
% de Error promedio					6.3	4.9

Los coeficientes de correlación calculados son cercanos a la unidad en ambos analitos, lo que indica no sólo que los datos de los estándares se ajustan al modelo de regresión, sino también, que el modelo tiene una gran capacidad de predicción para nuevas observaciones (Tabla 20).

Tabla 17. Coeficientes de correlación calculados a partir de análisis ANOVA de la RCP.

Analito	$R_{x,y}^2$	R_{aj}^2	R_{Pred}^2
Naproxeno	0.974	0.961	0.926
Paracetamol	0.993	0.989	0.979

Debido a que las nuevas variables predictoras **Z** son linealmente independientes, el FIV y la tolerancia calculados son prácticamente 1, mientras que los coeficientes de determinación son de cero para todas las variables (Tabla 18). Bajo este modelo de regresión se mejoraron los coeficientes de correlación respecto a la RLM.

Tabla 18. Parámetros de multicolinealidad de las variables predictoras Z.

Variable	Correlación	R_x^2	Tolerancia	FIV
z_1	z_2, z_3	0.000	1.000	1.000
z_2	z_1, z_3	0.000	1000	1.000
z_3	z_1, z_2	0.001	0.999	1.001

3.1.3.4. Regresión sobre mínimos cuadrados parciales

En este caso, el algoritmo MCPINL es aplicado independientemente para cada analito; para ello, se deben calcular tanto el promedio de los datos de las variables como su desviación estándar para su estandarización (Tabla 19).

Tabla 19. Parámetros estadísticos de las variables.

Vector de la variable	Promedio	Desviación estándar
y_1	131.040	61.045
x_1	2.288	1.178
x_2	10.423	5.539
x_3	29.434	4.921

La regresión MCP requiere la previa modificación de las variables de respuesta y de las variables predictoras. La entrada de datos corresponde a la matriz original de datos \mathbf{X} , estandarizada por columnas \mathbf{X}^E y la matriz de respuestas originales \mathbf{Y} , igualmente estandarizada por columnas \mathbf{Y}^E . De esta manera la matriz predictora es \mathbf{X}^E y la matriz de respuesta es \mathbf{Y}^E . El algoritmo MCPINL es aplicado independientemente para cada variable de respuesta.

$$\mathbf{Y}^E = \begin{bmatrix} 1.156 & 0.165 \\ 0.496 & -0.495 \\ -0.495 & -0.166 \\ -1.486 & 1.156 \\ -0.826 & 1.486 \\ 0.165 & -1.156 \\ -0.164 & -0.825 \\ 0.825 & -1.486 \\ 1.487 & 0.495 \\ -1.156 & 0.826 \end{bmatrix} \quad \mathbf{X}^E = \begin{bmatrix} 1.232 & 0.054 & 1.739 \\ 0.239 & -0.474 & -0.454 \\ -0.474 & -0.291 & -0.934 \\ -1.441 & 1.292 & -0.810 \\ -0.889 & 1.455 & -0.003 \\ 0.188 & -1.162 & -0.111 \\ -0.041 & -0.880 & -0.698 \\ 0.418 & -1.324 & 0.255 \\ 1.767 & 0.399 & 1.765 \\ -1.000 & 0.931 & -0.751 \end{bmatrix}$$

En este modelo de regresión no es necesario agregar una columna extra para calcular los interceptos "b" para la matriz \mathbf{X}^E . Así, las variables quedan como se muestra en la tabla 20.

Tabla 20. Variables de respuesta y variables predictoras para la Regresión MCP.

Estándar	Variables de respuesta estandarizadas			Variables predictoras Estandarizadas		
	y_1^E	y_2^E	x_1^E	x_2^E	x_3^E	
M ₁	1.156	0.165	1.232	0.054	1.739	
M ₂	0.496	-0.495	0.239	-0.474	-0.454	
M ₃	-0.495	-0.166	-0.474	-0.291	-0.934	
M ₄	-1.486	1.156	-1.441	1.292	-0.810	
M ₅	-0.826	1.486	-0.889	1.455	-0.003	
M ₆	0.165	-1.156	0.188	-1.162	-0.111	
M ₇	-0.164	-0.825	-0.041	-0.880	-0.698	
M ₈	0.825	-1.486	0.418	-1.324	0.255	
M ₉	1.487	0.495	1.767	0.399	1.765	
M ₁₀	-1.156	0.826	-1.000	0.931	-0.751	

Se calcula el vector de covarianzas $cov(\mathbf{Y}^E, \mathbf{X}^E)$ de orden $m \times 1$; cada elemento es la covarianza de la variable de respuesta con cada variable predictora.

3. Aplicación de la regresión multivariante en sistemas electroquímicos

$$\text{cov}(\mathbf{Y}^E, \mathbf{X}^E) = \begin{bmatrix} 0.980 \\ -0.530 \\ 0.806 \end{bmatrix}$$

Posteriormente, el vector de covarianzas es ortonormalizado para formar el vector \mathbf{w}_1 , aplicando la ecuación 24.

$$|\text{cov} \mathbf{Y}^E, \mathbf{X}^E| = [1.375]$$

Por lo tanto:

$$\mathbf{w}_1 = \begin{bmatrix} \left(\frac{0.980}{1.375}\right) \\ \left(\frac{-0.530}{1.375}\right) \\ \left(\frac{0.806}{1.375}\right) \end{bmatrix} = \begin{bmatrix} 0.713 \\ -0.390 \\ 0.586 \end{bmatrix}$$

A partir del vector \mathbf{w}_1 y la matriz estandarizada de predictores se calcula el primer componente MCP empleando la ecuación 28:

$$T_1 = \mathbf{X}^E \mathbf{w}_1 = \begin{bmatrix} 1.876 \\ 0.087 \\ -0.771 \\ -2.000 \\ -1.197 \\ 0.518 \\ -0.099 \\ 0.958 \\ 2.140 \\ -1.511 \end{bmatrix}$$

El coeficiente de regresión simple " u_1 " de la variable de respuesta con respecto al vector de componente T_1 es obtenido mediante la ecuación 29.

$$u_1 = \frac{T_1^T \mathbf{y}_1^E}{T_1^T T_1} = \frac{12.275}{17.613} = 0.703$$

Posteriormente el vector \mathbf{b} es calculado mediante la expresión 20. Cada elemento de este vector es el coeficiente de regresión simple de X_i sobre T_1 .

$$\mathbf{b} = \frac{T_1^T \mathbf{X}^E}{T_1^T T_1} = \begin{bmatrix} 12.437 & -6.378 & 10.734 \\ 17.613 & 17.613 & 17.613 \end{bmatrix} = [0.706 \quad -0.362 \quad 0.609]$$

Se actualizan las matrices de las variables predictoras y de las variables de respuesta mediante las ecuaciones 21 y 22:

$$X_1^{MCP} = \mathbf{X}^E - T_1 \mathbf{b}_1 = \begin{bmatrix} -0.093 & 0.733 & 0.595 \\ 0.178 & -0.442 & -0.507 \\ 0.072 & -0.571 & -0.463 \\ -0.029 & 0.568 & 0.409 \\ -0.044 & 1.021 & 0.726 \\ -0.177 & -0.974 & -0.426 \\ 0.029 & -0.916 & -0.638 \\ -0.259 & -0.977 & -0.328 \\ 0.256 & 1.173 & 0.461 \\ 0.068 & 0.384 & 0.170 \end{bmatrix}$$

$$Y_1^{MCP} = y_1^E - \mathbf{T}_1 u_1 = \begin{bmatrix} -0.162 \\ 0.434 \\ 0.047 \\ -0.081 \\ 0.015 \\ -0.199 \\ -0.095 \\ 0.152 \\ -0.017 \\ -0.094 \end{bmatrix}$$

Se calcula el próximo componente MCP a partir de una nueva matriz ortonormalizada de covarianzas \mathbf{w}_2 con los datos de las componentes Y_1^{MCP} y X_1^{MCP} .

En este caso, el número de componentes MCP es igual al número de variables predictoras, sin embargo, para matrices que contengan decenas o centenares de estas variables el número de componentes MCP puede ser menor. Por lo tanto, la matriz \mathbf{W} de covarianzas ortonormalizadas de orden (3×3) queda como:

$$\mathbf{W} = \begin{bmatrix} 0.713 & 0.191 & 0.675 \\ -0.386 & -0.698 & 0.604 \\ 0.586 & -0.691 & -0.424 \end{bmatrix}$$

La matriz \mathbf{b} , queda como:

$$\mathbf{b} = \begin{bmatrix} 0.706 & -0.362 & 0.609 \\ -0.041 & -0.904 & -0.546 \\ 0.675 & 0.604 & -0.424 \end{bmatrix}$$

Los coeficientes de regresión $\mathbf{b}_{i,k}^E$ para los datos estandarizados son calculados a partir de los vectores de covarianzas \mathbf{W} y los coeficientes de regresión simple entre los predictores y las variables latentes con la ecuación 33:

3. Aplicación de la regresión multivariante en sistemas electroquímicos

$$\mathbf{b}_{i,k}^E = \mathbf{W}(\mathbf{bW})^{-1}\mathbf{U} = \begin{bmatrix} 0.752 \\ -0.193 \\ 0.167 \end{bmatrix}$$

Donde \mathbf{U} es la matriz de cargas de \mathbf{Y}^E formada por los valores de coeficientes u_1, u_2, \dots, u_n y queda como:

$$\mathbf{U} = \begin{bmatrix} 0.703 \\ 0.053 \\ 0.321 \end{bmatrix}$$

Para transformar estos coeficientes para que sean aplicables a los datos sin estandarizar se emplea la ecuación 35:

$$\mathbf{b}_{i,k} = \mathbf{b}_{i,k}^E \left(\frac{\sigma_{y_i}}{\sigma_{x_k}} \right)$$

Para el primer coeficiente:

$$b_1 = 0.752 \left(\frac{61.045}{1.178} \right) = 38.956$$

Para el segundo coeficiente:

$$b_2 = -0.193 \left(\frac{61.045}{5.539} \right) = -2.124$$

Y para el tercer coeficiente:

$$b_3 = 0.167 \left(\frac{61.045}{4.921} \right) = 2.073$$

Para encontrar el término independiente se utiliza la ecuación 36:

$$b_0 = 131.040 - [(2.288)(0.752) + (10.423)(-0.193) + (29.434)(0.167)] = 3.046$$

Finalmente, la ecuación de regresión para el cálculo de concentraciones del naproxeno a partir de las intensidades a los potenciales seleccionados es:

$$C_{Nap} = 3.046 + 38.956(I_{0.65V}) - 2.124(I_{1.00V}) + 2.073(I_{1.40V})$$

Aplicando el algoritmo para C_{Par} la ecuación de regresión queda cómo:

$$C_{Par} = 22.922 + 7.817(I_{0.65V}) + 11.741(I_{1.00V}) - 1.074(I_{1.40V})$$

Los resultados de concentración pronosticada a partir de las dos ecuaciones anteriores y los porcentajes de error relativo se presentan en la tabla 21.

Tabla 21. Concentraciones pronosticadas y porcentajes de error en la determinación simultánea de naproxeno sódico y paracetamol mediante MCP.

Mezcla	Concentración estándar ($\mu\text{g mL}^{-1}$)		Concentración pronosticada ($\mu\text{g mL}^{-1}$)		% de Error	
	Naproxeno	Paracetamol	Naproxeno	Paracetamol	Naproxeno	Paracetamol
M ₁	201.6	141.7	204.7	137.2	1.5	3.2
M ₂	161.3	101.2	143.0	105.4	11.4	4.1
M ₃	100.8	121.4	103.2	113.2	2.4	6.8
M ₄	40.3	202.4	41.4	206.6	2.8	2.1
M ₅	80.6	222.6	73.1	218.0	9.3	2.1
M ₆	141.1	60.7	152.2	58.4	7.9	3.9
M ₇	121.0	81.0	132.4	77.7	9.4	4.1
M ₈	181.4	40.5	168.4	48.0	7.2	18.4
M ₉	221.8	161.9	225.5	164.4	1.7	1.6
M ₁₀	60.5	182.2	66.5	186.9	10.0	2.6
% de Error promedio					6.4	4.9

Una de las ventajas de la regresión MCP es una mayor versatilidad en el empleo de matrices de variables predictoras. Es posible incluir simultáneamente decenas de estas variables sin necesidad de aumentar el número de experimentos, lo que permite evaluar más información para determinar un modelo de regresión más idóneo a la naturaleza de los datos de entrada y su posterior utilización para la predicción de concentraciones desconocidas. Sin embargo, en este análisis no se evaluaron más potenciales de trabajo; por lo tanto, los coeficientes de la regresión entre RLM y MCP, al igual que los respectivos porcentajes de error, resultaron ser muy similares (Tabla 22).

3.1.3.5. Resumen de resultados

De acuerdo a los resultados de la tabla 22 se logró la cuantificación simultánea de naproxeno y paracetamol obteniendo en todos los casos errores promedio inferiores al 10% en las mezclas de concentración conocida.

Tabla 22. Porcentajes de error promedio de los estándares utilizando los modelos de regresión de trabajo.

Analito	% Error promedio		
	RLM	RCP	MCP
Naproxeno	6.3	6.4	6.4
Paracetamol	4.9	4.9	4.9

3. Aplicación de la regresión multivariante en sistemas electroquímicos

Estos porcentajes de error son similares debido a que la matriz de variables predictoras seleccionadas de manera empírica no era mucho más grande que la de matriz de respuestas; es decir, no hubo un cambio sustancialmente importante en la naturaleza de los datos empleando RCP y MCP.

3.1.3.6. Análisis de la muestra

Para la muestra problema, se utilizó un medicamento comercial que fue diluido hasta conseguir las concentraciones que aparecen en la tabla 23. Se prepararon tres disoluciones a las mismas condiciones y se le hizo el respectivo barrido voltamperométrico (Fig. 16).

Tabla 23. Concentraciones de naproxeno sódico y de paracetamol y sus respectivas intensidades de corriente de soluciones de muestras del medicamento.

Muestra	Concentración estándar($\mu\text{g mL}^{-1}$)		I (μA)		
	Naproxeno	Paracetamol	0.65 V	1.0 V	1.4 V
1	132.4	144.4	2.53	11.41	28.02
2	132.4	144.4	2.59	10.50	26.50
3	132.4	144.4	2.49	8.33	22.56

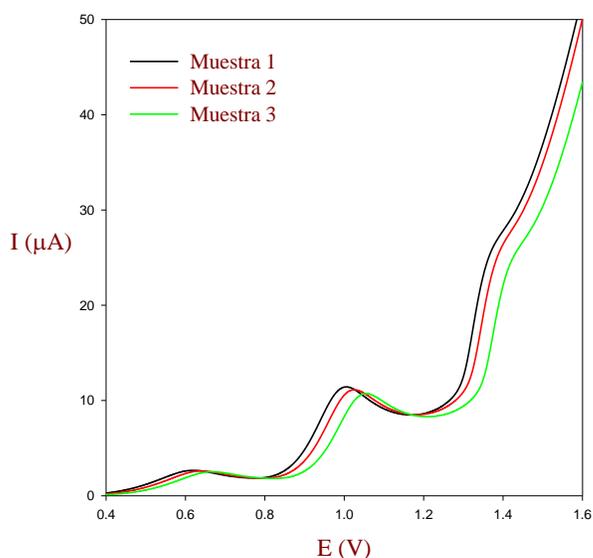


Figura 16. Voltamperogramas de disoluciones de muestras de medicamento en metanol en acetonitrilo/metanol al 30 % (v/v). $v = 10 \text{ mV s}^{-1}$.

Se observa que los voltamperogramas no son equivalentes entre sí; ocurre un desplazamiento del máximo de corriente a potenciales más elevados atribuible a una interferencia importante con el

resto de componentes de la muestra, como excipientes y aditivos que pasivan al electrodo de trabajo.

Este comportamiento provoca que las concentraciones calculadas sean gradualmente menores para el paracetamol, generando errores considerables en la precisión y exactitud del método (Tabla 24).

Tabla 24. Resultados del análisis por triplicado del medicamento mediante RLM.

Análisis	mg de analito por tableta	
	Naproxeno	Paracetamol
1	280.9	304.5
2	283.7	286.6
3	267.8	240.9
Marbete	275.0	300.0
Promedio	277.5	277.3
Desviación estándar	8.49	32.77
$t_{calculado}$	1.48	6.86

$$t_{critico (n-1)} = 4.30 (95 \%)$$

Para contrastar la diferencia entre la cantidad indicada en el marbete del medicamento y la cantidad encontrada en el análisis se utiliza una prueba t de contraste entre la media experimental de los tres análisis y el valor reportado por el fabricante. Para ello se calcula el estadístico t (ecuación 39) y se compara con el valor crítico de tablas [35].

$$t = (\bar{x} - \mu) \sqrt{\frac{n}{s}} \tag{39}$$

El análisis por triplicado de naproxeno da como resultado un valor promedio que no difiere significativamente del valor reportado por el fabricante, sin embargo, para el análisis de paracetamol la prueba de contraste indica que la cantidad de este analito en el marbete y la cantidad encontrada son significativamente diferentes.

Tabla 25. Resultados del análisis del medicamento.

Modelo de regresión	mg encontrados por tableta		% Error promedio	
	Naproxeno	Paracetamol	Naproxeno	Paracetamol
RLM	280.9	304.5	2.1	1.5
RCP	280.9	304.4	2.1	1.5
MCP	280.9	304.5	2.1	1.5
Marbete	275.0	300.0	-	-

La tendencia en el desplazamiento de los máximos de corriente observada en el comportamiento de los voltamperogramas origina un error sistemático que afecta los resultados del análisis de al menos uno de los analitos. En este sentido, se prefiere realizar el análisis sólo con el primer barrido voltamperométrico; de este modo, procesando los datos para calcular las concentraciones de

3. Aplicación de la regresión multivariante en sistemas electroquímicos

los fármacos en el medicamento, los resultados son muy cercanos a los reportados por el fabricante (Tabla 25).

3.1.4. Conclusiones

De acuerdo a los resultados de la cantidad de los analitos en la tableta se logró la cuantificación simultánea de naproxeno sódico y paracetamol mediante los modelos empleados de regresión multivariante, obteniendo en todos los casos errores promedio máximos del 2.1%. Considerando los tres análisis de la tableta (Tabla 24) y realizando un contraste t de Student (95% de significancia) contra el valor del marbete, es posible concluir que la cantidad de ambos fármacos en la tableta es igual a la que señala el fabricante. Los resultados finales son equivalentes independientemente del modelo de regresión empleando.

3.2. DETERMINACIÓN SIMULTÁNEA DE HIERRO (II) Y ÁCIDO ASCÓRBICO EN MEDICAMENTOS POR VOLTAMPEROMETRÍA Y REGRESIÓN MULTIVARIANTE

3.2.1. Introducción

Para el tratamiento de la anemia, se recomienda la ingesta de suplementos alimenticios o multivitamínicos que contengan hierro (II), generalmente en forma de sulfato ferroso heptahidratado ($\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$). En algunos de estos productos el hierro (II) se encuentra mezclado con ácido ascórbico, que actúa como un componente que favorece la absorción del hierro (II) en el organismo y que además, evita su oxidación prematura a hierro (III) [61].

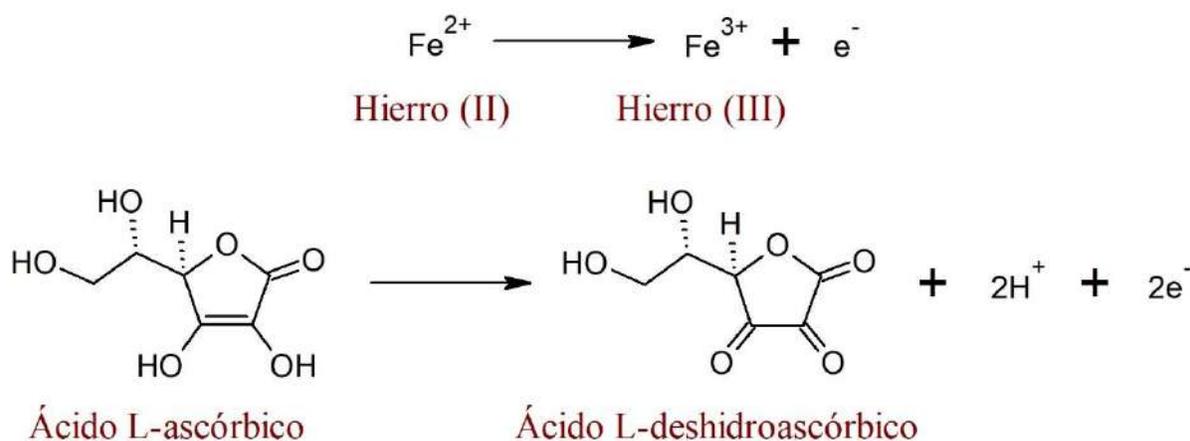


Figura 17. Estructuras del hierro (II), ácido ascórbico y sus productos de oxidación.

Tradicionalmente, los métodos farmacéuticos para la cuantificación de hierro (II) incluyen la detección espectroscópica mediante la formación de un complejo con 1,10-fenantrolina así como su

valoración con sulfato cérico en medio ácido [62]. Para el ácido ascórbico, el método de referencia consiste en su valoración con 2,6-diclorofenolindofenol a pH regulado con ácido metafosfórico y ácido acético [62]. Una variante de este método es la detección espectroscópica del abatimiento de la señal del 2,6-diclorofenolindofenol debida a su reducción en presencia de ácido ascórbico, que forma un producto incoloro [63].

Sus propiedades electroquímicas les permiten ser cuantificados por voltamperometría [64], [65]. El hierro (II) se oxida reversiblemente a hierro (III) en medios ácidos mientras que el ácido ascórbico se oxida irreversiblemente a ácido deshidroascórbico (Fig. 17). Hasta el momento, la determinación simultánea de estos compuestos sólo se ha reportado mediante detección espectroscópica en FIA [61]. En este trabajo se explora la determinación electroquímica simultánea de ambos analitos y su aplicación en un medicamento comercial mediante voltamperometría lineal y regresión multivariante.

3.2.2. Experimentación

Todas las soluciones fueron preparadas con reactivos comerciales de grado analítico: sulfato ferroso heptahidratado y ácido ascórbico, (Sigma-Aldrich), ácido clorhídrico concentrado (J.T. Baker). El electrolito soporte fue preparado a una concentración de HCl 0.1 M con agua desionizada Milli Q (Millipore) con una resistividad $\geq 18 \text{ M}\Omega \text{ cm}$. Las soluciones estándar de hierro (II) y ácido ascórbico fueron preparadas en el electrolito soporte. Para la preparación de la muestra se pesaron y pulverizaron tabletas de una formulación comercial que contiene sulfato ferroso heptahidratado y ácido ascórbico (525 mg y 500 mg por tableta respectivamente). Posteriormente, se pesaron 127.3 mg de la muestra pulverizada y se disolvieron en 20.0 mL. La solución se ultrasonificó durante 20 minutos y se filtró a gravedad con un papel de poro fino (Wathman 42) para retirar el excipiente. Finalmente, la solución filtrada fue aforada a 25.0 mL con el electrolito soporte.

Los voltamperogramas de las disoluciones fueron obtenidos mediante un potenciostato-galvanostato AUTOLAB PGSTAT 302N (Metrohm) controlado mediante el software Nova 2.0. La celda electroquímica se compone de tres electrodos: un electrodo de carbón vítreo BASi MF-2012 con un diámetro de 3.0 mm (7.07 mm^2), un de platino BASi MF-2013 con un diámetro de 1.6 mm (2.01 mm^2) como electrodo auxiliar y un electrodo de referencia de Ag/AgCl de doble camisa (Orion 900002) con solución externa de KNO_3 al 10% (P/V). Se hicieron los respectivos voltamperogramas lineales de los estándares así como de las muestras, partiendo del potencial de corriente nula hasta 1.60 V a una velocidad de barrido de 20 mVs^{-1} . Los datos obtenidos son organizados en una matriz de datos para su tratamiento estadístico empleando el software Minitab 17.

3.2.3. Resultados y discusión

En primer lugar se realizaron, a partir de voltamperometría lineal, las curvas de calibrado para cada analito para evaluar su comportamiento electroquímico y la proporcionalidad de la intensidad de pico con respecto a la concentración. De acuerdo a los voltamperogramas lineales hechos

3. Aplicación de la regresión multivariante en sistemas electroquímicos

a las disoluciones de concentración creciente de hierro (II), este analito se oxida a un potencial de 1.0 V (Figs. 17 y 18). Para el ácido ascórbico, se observa un máximo de corriente a un potencial de 0.75 V (Fig 19).

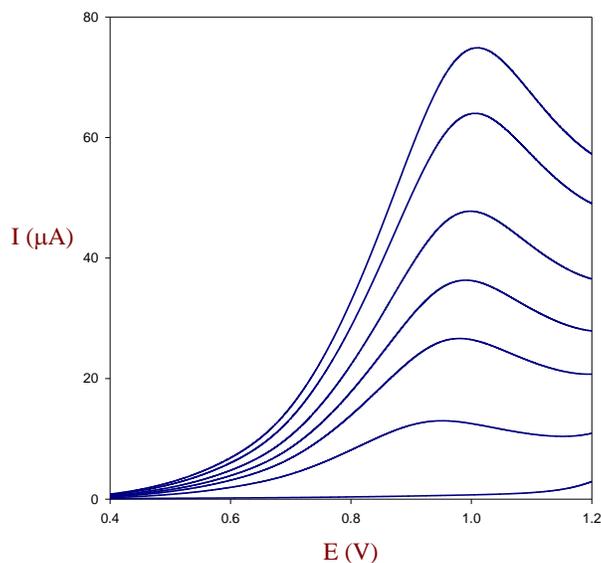


Figura 18. Voltamperogramas de sulfato hierro (II) en HCl 0.1 M. $\nu = 20 \text{ mVs}^{-1}$.

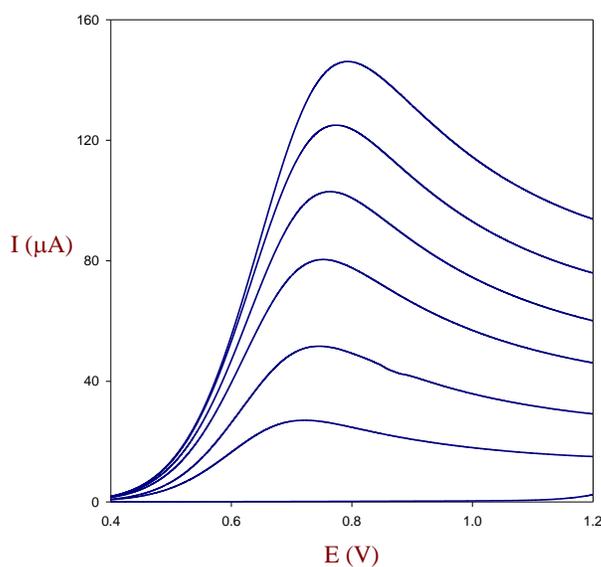


Figura 19. Voltamperogramas de ácido ascórbico en HCl 0.1 M. $\nu = 20 \text{ mV s}^{-1}$.

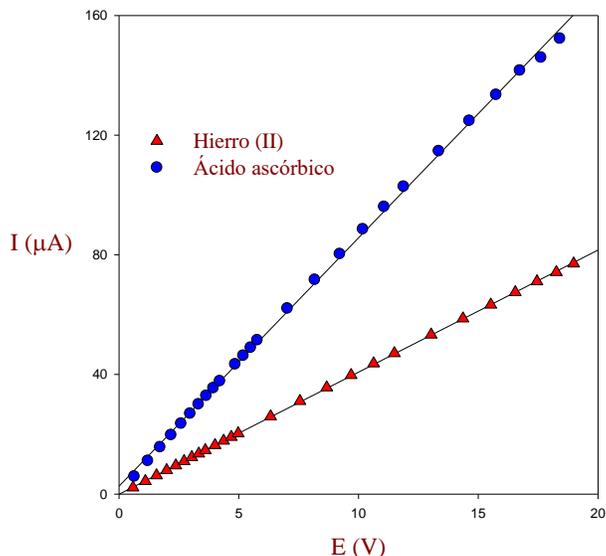


Figura 20. Curvas de calibrado independientes para hierro (II) y ácido ascórbico.

A partir de los voltamperogramas lineales, se construyen las curvas de calibración para cada uno de los analitos confirmándose la linealidad del sistema en el intervalo de concentraciones seleccionado y una sensibilidad mayor para ácido ascórbico; siendo los coeficientes de correlación de ambos analitos son cercanos a la unidad (Fig. 20). Los parámetros analíticos de las curvas de calibración se muestran en la tabla 26.

Tabla 26. Parámetros analíticos de voltamperometría lineal para sulfato ferroso y ácido ascórbico

Parámetro	Magnitud	
	Hierro (II)	Ácido ascórbico
IRL (mM)	0.43-18.98	2.66-18.40
b ($\mu\text{A mM}^{-1}$)	4.08 ± 0.01	8.30 ± 0.08
a (mM)	-0.04 ± 0.10	3.49 ± 0.77
$r_{x,y}^2$	0.999975	0.999603
LOD (mM)	0.13	0.48
LOQ (mM)	0.43	1.60

IRL = Intervalo de respuesta lineal

Una vez establecidos los intervalos de respuesta lineal, para la determinación simultánea de hierro (II) y ácido ascórbico se construye la matriz de respuesta a partir de un diseño central compuesto con la preparación de nueve disoluciones de concentración conocida de ambos analitos; el intervalo de concentraciones de hierro (II) y de ácido ascórbico es de 3.0 a 12 mM y de 5.5 a 18 mM respectivamente (Tabla 27). Cada disolución se analiza a las condiciones descritas anteriormente para cada analito individualmente, los voltamperogramas se muestran en la Figura 21.

3. Aplicación de la regresión multivariante en sistemas electroquímicos

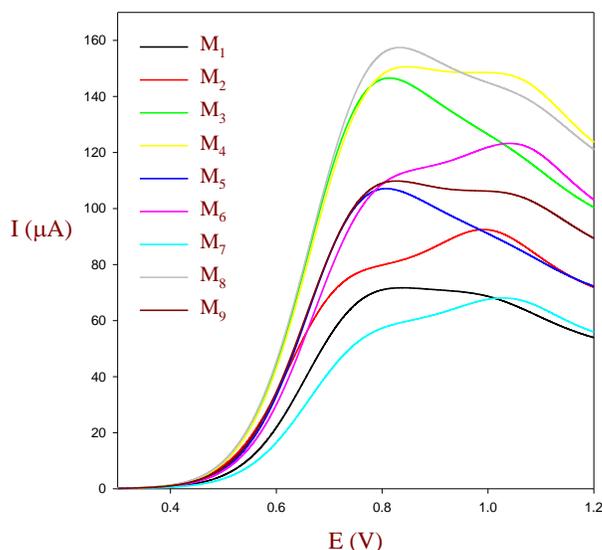


Figura 21. Voltamperogramas de las mezclas de hierro (II) y ácido ascórbico en HCl 0.1 M. $v = 20 \text{ mV s}^{-1}$.

Tabla 27. Matriz de concentración para la determinación simultánea de hierro (II) y ácido ascórbico

Mezcla	Concentración (mM)	
	Hierro (II)	Ácido ascórbico
M ₁	4.32	7.33
M ₂	10.68	7.33
M ₃	4.32	16.17
M ₄	10.68	16.17
M ₅	3.00	11.75
M ₆	12.00	11.75
M ₇	7.50	5.50
M ₈	7.50	18.00
M ₉	7.50	11.75

3.2.3.1. RLM-Stepwise

Para cada analito, se realiza la selección de variables predictoras empleando el método stepwise con la finalidad de seleccionar las variables menos correlacionadas entre sí y que mejor se ajustan a los valores de los estándares. Las ecuaciones de regresión resultantes se muestran a continuación, mientras que los resultados de los valores encontrados para α_E , α_R y de los coeficientes de correlación se presentan en la tabla 28.

$$C_{Fe(II)} = 0.438 + 0.113(I_{0.636V}) - 0.274(I_{0.820V}) + 0.298(I_{1.036V})$$

Aplicando el algoritmo para C_{AA} la ecuación de regresión queda cómo:

$$C_{AA} = 0.373 + 18.160(I_{0.312V}) + 0.135(I_{0.802V})$$

Tabla 28. Resultados de los coeficientes de correlación empleando el modelo de regresión MLR y el método stepwise para la selección de las variables de respuesta.

Parámetro	Hierro (II)	Ácido ascórbico
α_E	0.25	0.15
α_R	0.25	0.15
$R_{x,y}^2$	0.9995	0.9982
R_{aj}^2	0.9991	0.9976
R_{Pred}^2	0.9968	0.9923

La ecuación de regresión para el hierro (II) requirió que el valor de α_E fuese mayor que el utilizado para ácido ascórbico para que se incluyeran variables predictoras al modelo. Aunque los valores de FIV para estas variables son relativamente mayores para dos de ellas (tabla 29), el valor de R_{Pred}^2 cercano a 1 indica que el modelo es capaz de predecir correctamente nuevas mediciones.

Tabla 29. Parámetros de multicolinealidad de las variables predictoras seleccionadas para la regresión de hierro (II).

Variable	Correlación	R_x^2	FIV
$I_{0.636V}$	$I_{0.820V}, I_{1.036V}$	0.9186	12.28
$I_{0.820V}$	$I_{0.636V}, I_{1.036V}$	0.9558	22.60
$I_{1.036V}$	$I_{0.636V}, I_{0.820V}$	0.8748	7.99

En el caso de la ecuación de regresión para el ácido ascórbico, los coeficientes de correlación ($R_{x,y}^2$, R_{aj}^2 , y R_{Pred}^2) son igualmente buenos con sólo dos variables predictoras (Tabla 30). Estas variables seleccionadas no son colineales entre sí y el valor de R_{Pred}^2 calculado es cercano a 1, lo que indica que la ecuación de regresión posee una gran capacidad de predicción para nuevas mediciones.

Tabla 30. Parámetros de multicolinealidad de las variables predictoras seleccionadas para la regresión de ácido ascórbico.

Variable	Correlación	R_x^2	FIV
$I_{0.312V}$	$I_{0.802V}$	0.5551	2.25
$I_{0.802V}$	$I_{0.312V}$	0.5551	2.25

Al aplicar las ecuaciones de regresión para recalculer las concentraciones de los analitos se obtienen excelentes resultados. Los porcentajes de error promedio son inferiores al 1.3 %, por lo tanto, el modelo de regresión se ajusta adecuadamente a los valores de los datos de entrada (Tabla 31).

3. Aplicación de la regresión multivariante en sistemas electroquímicos

Tabla 31. Concentraciones pronosticadas y porcentajes de error en la determinación simultánea de hierro (II) y ácido ascórbico mediante Stepwise-MLR.

Mezcla	Concentración estándar (mM)		Concentración pronosticada (mM)		% de Error	
	Hierro (II)	Ácido ascórbico	Hierro (II)	Ácido ascórbico	Hierro (II)	Ácido ascórbico
M ₁	4.32	7.33	4.38	7.23	1.4	1.4
M ₂	10.68	7.33	10.66	7.26	0.2	1.0
M ₃	4.32	16.17	4.27	16.45	1.2	1.7
M ₄	10.68	16.17	10.77	16.34	0.8	1.1
M ₅	3.00	11.75	2.92	11.6	2.7	1.3
M ₆	12.00	11.75	11.87	11.76	1.1	0.1
M ₇	7.50	5.50	7.52	5.63	0.3	2.4
M ₈	7.50	18.00	7.52	17.66	0.3	1.9
M ₉	7.50	11.75	7.58	11.81	1.1	0.5
% de Error promedio					1.0	1.3

3.2.3.2. Regresión sobre componentes principales

Es posible utilizar el cálculo de componentes principales sobre la matriz de predictores con el fin de crear nuevas variables que se correlacionen de mejor manera con la variable de respuesta. En este caso, los valores de α_F y α_R fueron aumentados a 0.30 para la selección de variables mediante stepwise. Tanto para hierro (II) como para ácido ascórbico, el método selecciona siete potenciales diferentes; mismos que son procesados mediante RCP, encontrando que el número de variables predictoras es reducido a sólo tres variables nuevas para cada variable de respuesta. Para el hierro (II), los parámetros de multicolinealidad de las variables encontradas se muestran en la tabla 32, donde sólo una de las variables predictoras no presenta multicolinealidad con el resto.

Tabla 32. Parámetros de multicolinealidad de las variables predictoras seleccionadas para la regresión de hierro (II) mediante RCP.

Variable	Correlación	R_x^2	FIV
$Z_1^{Fe(II)}$	$Z_2^{Fe(II)}, Z_3^{Fe(II)}$	0.999	1741.55
$Z_2^{Fe(II)}$	$Z_1^{Fe(II)}, Z_3^{Fe(II)}$	0.999	1682.46
$Z_3^{Fe(II)}$	$Z_1^{Fe(II)}, Z_2^{Fe(II)}$	0.660	2.94

Fe (II) = Hierro (II)

Para el ácido ascórbico, los parámetros de multicolinealidad de las variables encontradas se muestran en la tabla 33, donde todas las variables predictoras presentan un alto grado de multicolinealidad entre sí.

Tabla 33. Parámetros de multicolinealidad de las variables predictoras seleccionadas para la regresión de hierro (II) mediante RCP.

Variable	Correlación	R_x^2	FIV
z_1^{AA}	z_2^{AA}, z_3^{AA}	0.999	1237078.68
z_2^{AA}	z_1^{AA}, z_3^{AA}	0.999	1841430.21
z_3^{AA}	z_2^{AA}, z_1^{AA}	0.999	60476.53

AA = Ácido ascórbico

Debido a la alta multicolinealidad de las variables, los coeficientes de correlación $R_{x,y}^2$, R_{aj}^2 y R_{Pred}^2 son cercanos a la unidad, como se muestra en la tabla 34; no existiendo un sobreajuste del modelo.

Tabla 34. Coeficientes de correlación calculados a partir de RCP para hierro (II) y ácido ascórbico.

Analito	$R_{x,y}^2$	R_{aj}^2	R_{Pred}^2
Hierro (II)	0.999	0.999	0.996
Ácido ascórbico	0.998	0.997	0.991

A partir de estas nuevas variables predictoras se calculan nuevamente las concentraciones en las mezclas de los estándares. Los resultados se muestran en la tabla 35.

Tabla 35. Resultados de la determinación simultánea de hierro (II) y ácido ascórbico mediante RCP.

Mezcla	Concentración estándar (mM)		Concentración pronosticada (mM)		% de Error	
	Hierro (II)	Ácido Ascórbico	Hierro (II)	Ácido Ascórbico	Hierro (II)	Ácido ascórbico
M ₁	4.32	7.33	4.36	7.31	1.0	0.3
M ₂	10.68	7.33	10.66	7.36	0.2	0.4
M ₃	4.32	16.17	4.25	16.61	1.7	2.7
M ₄	10.68	16.17	10.79	16.37	1.0	1.2
M ₅	3.00	11.75	2.98	11.67	0.7	0.7
M ₆	12.00	11.75	11.89	11.77	0.9	0.1
M ₇	7.50	5.50	7.51	5.66	0.1	2.9
M ₈	7.50	18.00	7.50	17.82	< 0.1	1.0
M ₉	7.50	11.75	7.60	11.93	1.3	1.5
% de Error promedio					0.8	1.2

El uso de componentes principales no produjo variables no correlacionadas, sin embargo, en ambas regresiones el coeficiente de correlación ajustado (R_{Pred}^2) es cercano a la unidad y es mayor que los calculados utilizando sólo RLM. Debido a ello, los porcentajes de error promedio son menores a los obtenidos empleando MLR.

3. Aplicación de la regresión multivariante en sistemas electroquímicos

3.2.3.3. Regresión sobre mínimos cuadrados parciales

Para esta herramienta quimiométrica, es posible evaluar todas las variables predictoras; cada una aporta un valor a la concentración calculada de los analitos. Sin embargo, se desconoce la capacidad del modelo para nuevas observaciones. Utilizando el software Minitab 17, los resultados de la aplicación de este modelo a los datos de las variables de respuesta se observa en la tabla 36.

Tabla 36. Resultados de la determinación simultánea de hierro (II) y ácido ascórbico mediante MCP.

Mezcla	Concentración estándar (mM)		Concentración pronosticada (mM)		% de Error	
	Hierro (II)	Ácido Ascórbico	Hierro (II)	Ácido Ascórbico	Hierro (II)	Ácido ascórbico
M ₁	4.32	7.33	4.32	7.33	< 0.1	< 0.1
M ₂	10.68	7.33	10.68	7.33	< 0.1	< 0.1
M ₃	4.32	16.17	4.32	16.17	< 0.1	< 0.1
M ₄	10.68	16.17	10.68	16.17	< 0.1	< 0.1
M ₅	3.00	11.75	3.00	11.75	< 0.1	< 0.1
M ₆	12.00	11.75	12.00	11.75	< 0.1	< 0.1
M ₇	7.50	5.50	7.50	5.50	< 0.1	< 0.1
M ₈	7.50	18.00	7.50	18.00	< 0.1	< 0.1
M ₉	7.50	11.75	7.50	11.75	< 0.1	< 0.1
% de Error promedio					< 0.1	< 0.1

Los porcentajes de error promedio son cercanos a 0 en todos los casos, lo que indica que este modelo de regresión se ajusta correctamente al comportamiento de los datos.

3.2.3.4. Resumen de resultados

De acuerdo a los resultados de la tabla 37 se logró la determinación simultánea de hierro (II) y ácido ascórbico en las mezclas de concentración conocida, obteniendo en todos los casos errores promedio inferiores al 10%. Las diferencias en los resultados se deben principalmente a la selección de variables empleadas para los modelos de RLM y RCP, en contraste con MCP donde se utilizaron todos los potenciales de la ventana seleccionada y sus porcentajes de error promedio para el cálculo de las concentraciones de trabajo son inferiores al 0.1 %.

Tabla 37. Porcentajes de error promedio de los estándares de hierro (II) y ácido ascórbico utilizando los modelos de regresión de trabajo.

Analito	% Error promedio		
	RLM	RCP	MCP
Hierro (II)	1.0	0.8	<0.1
Ácido ascórbico	1.3	1.2	<0.1

3.2.3.5. Muestra problema

Se prepararon tres disoluciones de la misma concentración a partir de una muestra pulverizada del multivitamínico a las que se les realizó un barrido voltamperométrico a las mismas condiciones que las mezclas estándar. (Fig. 22).

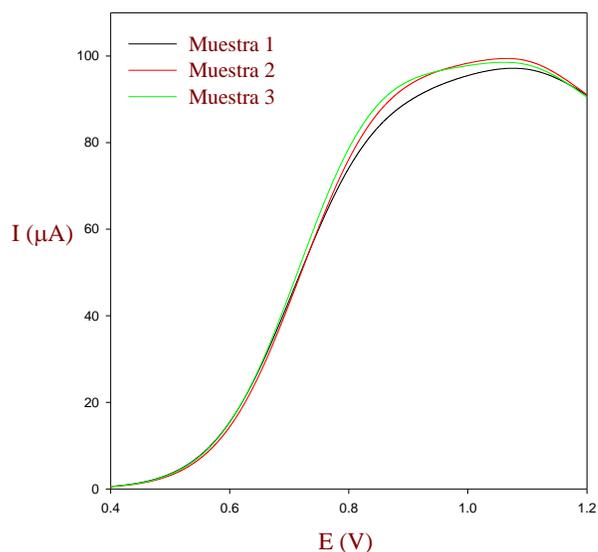


Figura 22. Voltamperogramas lineales de tres muestras de medicamento en HCl 0.1 M. $v = 20 \text{ mV s}^{-1}$.

Los voltamperogramas no presentan el comportamiento esperado a las concentraciones indicadas en el marbete del medicamento. Considerando que el valor encontrado de R_{pred}^2 encontrado mediante RLM y RCP, así como los porcentajes de error obtenidos en MCP, es posible afirmar que el método es robusto y en consecuencia, que la falta de exactitud es debida a los excipientes de la tableta; por ejemplo, la capa entérica que la recubre. Los resultados de las cuantificaciones se muestran en la tabla 38.

Tabla 38. Resultados de la determinación de hierro (II) y ácido ascórbico en una muestra mediante los modelos de regresión de trabajo.

Modelo de Regresión	mg encontrados por tableta		% Error promedio	
	Sulfato ferroso heptahidratado	Ácido ascórbico	Sulfato ferroso heptahidratado	Ácido ascórbico
RLM-Stepwise	686.2	354.9	30.7	27.0
RCP	697.1	365.3	32.8	26.9
MCP	676.3	426.5	28.8	14.7
Marbete	525	500	-	-

3. Aplicación de la regresión multivariante en sistemas electroquímicos

3.2.4. Conclusiones

La cuantificación de hierro (II) y ácido ascórbico en medicamentos comerciales no logró ser efectiva debido a problemas con la composición de la muestra; un pretratamiento a la muestra debe ser implementado para evitar la pasivación electroquímica del electrodo. No obstante este inconveniente, el modelo logra predecir las concentraciones de ambos analitos con porcentajes de error promedio inferiores al 1.3 % en las mezclas de estándares, siendo más exacto el modelo de regresión MCP.

4. DESARROLLO DE BIOSENSORES MUTIPARÁMETROS

El presente trabajo se circunscribe en el marco del proyecto CONACyT CB-2014-242677 en el cual se está trabajando para desarrollar un biosensor multienzimático y multiparamétrico para cuantificar glucosa, colesterol y triglicéridos. En este sentido y con la finalidad de evaluar las metodologías de calibración multivariante a la respuesta esperada de este dispositivo, se realizó la simulación de su respuesta tanto voltamperométrica como cronoamperométrica. Por otra parte se evaluó la respuesta de un biosensor amperométrico de glucosa para cuantificar simultáneamente glucosa y ácido ascórbico.

4.1. SIMULACIÓN DE RESPUESTA DE UN BIOSENSOR ENZIMÁTICO PARA LA CUANTIFICACIÓN SIMULTÁNEA DE GLUCOSA, COLESTEROL Y TRIGLICÉRIDOS

4.1.1. Introducción

Debido a su alta selectividad, los biosensores desarrollados hasta el momento en su gran mayoría son monoparamétricos; sin embargo, de acuerdo al esquema de detección que tienen en común los biosensores de glucosa, colesterol y triglicéridos es posible su determinación simultánea con un solo biosensor que contenga todas las enzimas necesarias para este cometido. Considerando que el desarrollo de un biosensor de esta clase requerirá del uso de reactivos costosos y de un tiempo considerable de trabajo, se puede utilizar un simulador de señales empleando las condiciones de trabajo a las que operarán estos dispositivos para modelar el comportamiento electroquímico esperado de un biosensor triparamétrico que sea capaz de detectar y cuantificar simultáneamente glucosa, colesterol y triglicéridos. De este modo, es posible obtener información valiosa que permita el desarrollo de estos dispositivos en muestras reales en favor de que su empleo en el sector salud contribuya a la prevención y al tratamiento del síndrome metabólico.

4.1.2. Materiales y métodos

Los biosensores enzimáticos electroquímicos utilizan transductores voltamperométricos y cronoamperométricos; por lo tanto, para obtener los voltamperogramas y cronoamperogramas se modeló un biosensor multienzimático construido a partir de un electrodo de carbón vítreo (0.0707 cm^2) sobre el que se depositó una capa de 7,7,8,8-tratracianoquinodimetano (TCNQ) como mediador y sobre ésta una membrana enzimática formulada con glucosa oxidasa, colesterol oxidasa, lipasa, glicerol quinasa y glicerol fosfato oxidasa. Así, al interactuar el biosensor con una mezcla de glucosa, colesterol y triglicéridos, se producirá una intensidad de corriente cuya magnitud es la suma de las intensidades generadas por los tres analitos por separado

Para realizar la simulación se utilizó el software DigiElch 7 considerando en el caso de la voltamperometría lineal una velocidad de barrido de $0.020 \text{ V}\cdot\text{s}^{-1}$ en una ventana de potencial de -0.1 V hasta 1.0 V . Para la obtención de los cronoamperogramas se consideró un potencial impuesto de 0.3 V . Se asume que se cumple la ley de la aditividad donde la señal analítica de una mezcla de analitos es la suma de las señales generadas individualmente para cada analito; por lo tanto, estas señales son sumadas para generar mezclas de concentración conocida. La información generada fue procesada y evaluada mediante el software Minitab 17. Los parámetros alimentados al programa para realizar la simulación se encuentran en la tabla 39.

Tabla 39. Datos y parámetros de entrada para la simulación de un biosensor multienzimático.

Parámetro	Valor	Referencia
Área superficial	0.0707 cm ²	-
Potencial de oxidación del TCNQ (SCE)	0.127 V	[20]
Constante heterogénea de velocidad de transferencia electrónica TCNQ sobre carbón vítreo (k_s)	1.3x10 ⁻⁵ cm·s ⁻¹	[66]
Coefficiente de difusión de glucosa ($D_{Glucosa}$)	6.77x10 ⁻⁵ cm ² ·s ⁻¹	[67]
Coefficiente de difusión de colesterol ($D_{Colesterol}$)	1.8 x10 ⁻¹⁰ cm ² ·s ⁻¹	[68]
Coefficiente de difusión de trioleína ($D_{Trioleína}$)	1.04x10 ⁻¹¹ cm ² ·s ⁻¹	[69]

4.1.3. Resultados y discusión

Para cada concentración se simuló un barrido voltamperométrico desde -0.1 V hasta 1.0 V (Fig. 23) y posteriormente se sumaron las señales de acuerdo a la matriz de concentraciones realizada al azar dentro de los intervalos lineales encontrados en la literatura. A partir de los voltamperogramas generados se obtuvieron las intensidades de corriente para 0.0, 0.1, 0.2, 0.3, 0.4 y 0.5 V, obteniéndose la matriz de resultados de (10x6), (Tabla 40). Esta matriz de variables de respuesta es utilizada en los tres modelos de regresión multivariante.

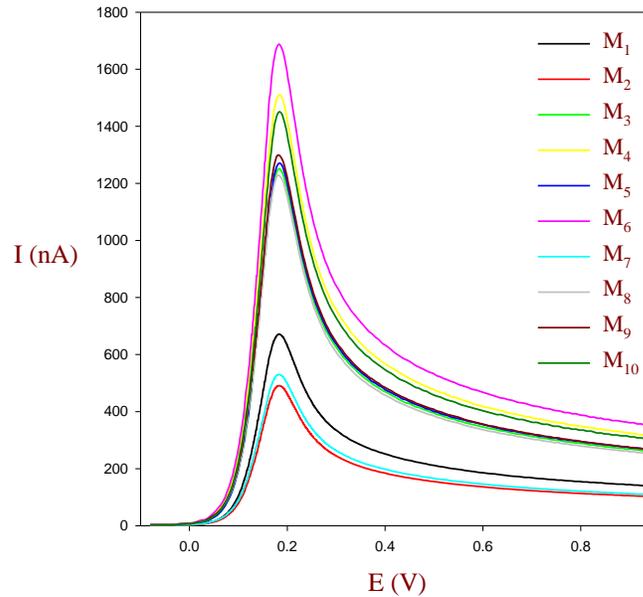


Figura 23. Voltamperogramas lineales de las mezclas de concentraciones de los analitos.

Tabla 40. Matriz de concentraciones y de respuestas de las mezclas de los analitos para un transductor voltamperométrico.

Mezcla	Concentración (mM)			Intensidad de corriente (nA)					
	GLU	COL	TG	I (0.0V)	I (0.1V)	I (0.2V)	I (0.3V)	I (0.4V)	I (0.5V)
M ₁	1.90	5.00	0.70	4.4	155.5	600.8	328.6	251.0	213.1
M ₂	1.20	4.00	0.45	2.9	115.1	439.9	241.9	183.0	154.7
M ₃	5.40	6.00	1.45	7.7	263.8	1146.6	622.5	474.0	401.1
M ₄	6.10	8.00	2.45	9.8	333.2	1389.1	758.8	577.0	487.9
M ₅	3.30	10.00	1.20	8.7	306.0	1141.8	626.1	474.0	403
M ₆	6.80	9.00	1.7	10.3	360.0	1534.3	839.3	637.0	538.3
M ₇	2.60	2.00	0.95	3.2	113.2	497.3	274.9	204.0	170.5
M ₈	4.70	7.00	1.95	7.6	280.3	1127.8	614.2	468.0	395.0
M ₉	0.50	3.00	0.20	2.0	75.4	273.6	151.6	116.0	98.1
M ₁₀	4.00	11.00	2.20	9.9	354.8	1311.0	719.7	552.0	467.2

GLU = glucosa, COL = colesterol, TG = triglicéridos.

De igual manera, se simularon registros cronoamperométrico durante 70 segundos a 0.3 V a las concentraciones de los analitos individualmente, para posteriormente sumar las señales de las mezclas de los analitos. Mediante los cronoamperogramas se obtuvieron las intensidades de corriente a 5, 10, 20, 35, 50 y 65 segundos (Fig. 24), generándose igualmente una matriz de resultados de (10x6), (Tabla 41).

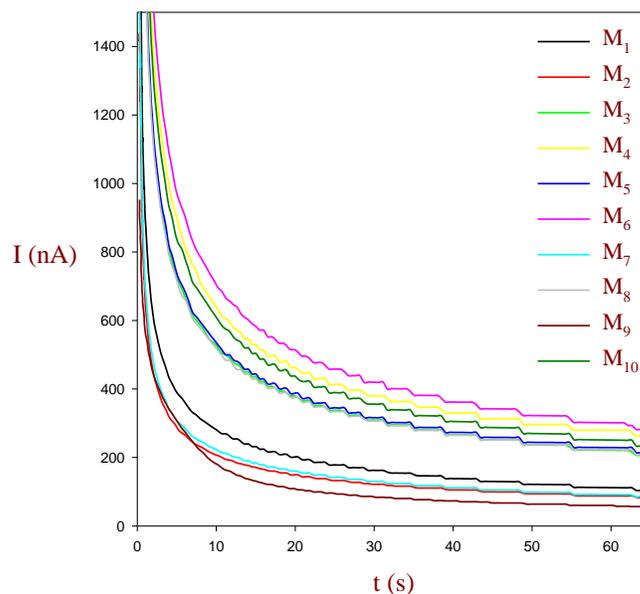


Figura 24. Cronoamperogramas de las mezclas de concentraciones de los analitos.

Tabla 41. Matriz de concentraciones y de respuestas de las mezclas de los analitos para un transductor cronoamperométrico.

Mezcla	Concentración (mM)			Intensidad de corriente (nA)					
	GLU	COL	TG	I (5s)	I (10s)	I (20s)	I (35s)	I (50s)	I (65s)
M ₁	1.90	5.00	0.70	397.5	290.0	205.5	157.4	124.3	106.1
M ₂	1.20	4.00	0.45	301.9	212.5	151.9	118.3	95.7	83.9
M ₃	5.40	6.00	1.45	755.5	544.6	388.5	301.1	242.4	211.4
M ₄	6.10	8.00	2.45	924.9	665.9	478.1	367.5	306.4	271.8
M ₅	3.30	10.00	1.20	769.8	551.4	394.2	307.3	248.7	218.3
M ₆	6.80	9.00	1.70	1012.2	735.3	524.9	399.8	328.6	287.1
M ₇	2.60	2.00	0.95	324.4	233.5	166.1	128.3	102.4	88.5
M ₈	4.70	7.00	1.95	756.1	540.0	386.6	301.2	244.3	214.9
M ₉	0.50	3.00	0.20	184.8	135.3	97.0	76.0	61.9	54.8
M ₁₀	4.00	11.00	2.20	881.6	634.6	451.5	342.2	278.1	240.5

GLU = glucosa, COL = colesterol, TG = triglicéridos.

4.1.3.1. Regresión lineal múltiple

Los resultados de intensidad de corriente para los datos voltamperométricos y cronoamperométricos simulados fueron procesados primeramente por RLM, encontrándose una buena predicción de las concentraciones de trabajo para los datos voltamperométricos para glucosa y colesterol dado que los porcentajes de error promedio son de 1.1 y 1.3 % respectivamente. En el caso de los triglicéridos, sus porcentajes de error son superiores al 10.0 % debido a que sus concentraciones son menores que para glucosa y colesterol; además que su coeficiente de difusión es el menor de los tres analitos (Tabla 42).

Tabla 42. Resultados de porcentajes de error en la determinación simultánea de glucosa, colesterol y triglicéridos mediante cronoamperometría y voltamperometría utilizando RLM.

Mezcla	% error (Cronoamperometría)			% error (Voltamperometría)		
	GLU	COL	TG	GLU	COL	TG
M ₁	5.6	3.5	5.6	3.4	4.0	24.9
M ₂	18.8	8.6	50.4	2.1	2.8	53.5
M ₃	1.1	1.6	3.7	0.7	1.3	3.5
M ₄	1.9	3.0	6.2	0.3	1.3	12.2
M ₅	5.5	3.1	5.5	0.2	0.3	5.2
M ₆	4.7	5.8	11.5	0.4	1.2	8.4
M ₇	35.2	85.5	31.7	0.3	0.3	6.6
M ₈	0.6	1.1	7.0	0.4	0.8	1.4
M ₉	176	52.7	23.9	2.2	0.8	9.2
M ₁₀	13.4	8.5	3.3	0.6	0.6	<0.1
% EP	26.3	17.3	14.9	1.1	1.3	13.9

GLU = glucosa, COL = colesterol, TG = triglicéridos, % EP = porcentaje de error promedio.

En el caso de los datos generados por cronoamperometría, los resultados de porcentajes de error promedio superan el 10 % para los tres analitos. No obstante, en algunas simulaciones logran predecir correctamente la concentración de los analitos, por lo que sería posible optimizar este proceso de transducción cronoamperométrica.

Para conocer la idoneidad del modelo de regresión con las variables modificadas se calculan los coeficientes de correlación ($R_{x,y}^2$, R_{aj}^2 y R_{Pred}^2) correspondientes para cada analito, así como los parámetros de multicolinealidad (Tablas 43 y 44).

Tabla 43. Coeficientes de determinación múltiple, de determinación ajustado y de determinación predicho para detección voltamperométrica y cronoamperométrica de GLU, COL y TG encontrados en la regresión RLM.

Detección	$R_{x,y}^2$			R_{aj}^2			R_{Pred}^2		
	GLU	COL	TG	GLU	COL	TG	GLU	COL	TG
Voltamperométrica	0.999	0.999	0.960	0.999	0.999	0.880	0.996	0.980	0.158
Cronoamperométrica	0.948	0.917	0.950	0.843	0.750	0.860	0.374	0.804	0.353

GLU = glucosa, COL = colesterol, TG = triglicéridos.

En la tabla 43 se observa que comparar los coeficientes de determinación predichos (R_{Pred}^2), éstos son mayores en la detección voltamperométrica en al menos dos analitos. Esto quiere decir que el transductor voltamperométrico tendrá los mejores resultados y un menor porcentaje de error en la predicción de las concentraciones para nuevas mediciones.

Tabla 44. Valores del FIV para la determinación simultánea de GLU, COL y TG de las variables de respuesta seleccionadas para voltamperometría y cronoamperometría.

Detección	Variable	FIV
Voltamperométrica	0.0 V	606.42
	0.1 V	333.35
	0.2 V	33,446.10
	0.3 V	25,957.35
	0.4 V	39,846.39
	0.5 V	49,970.85
Cronoamperométrica	5 s	7,277.21
	10 s	4'458,665.05
	20 s	4'358,795.3
	35 s	3,845.69
	50 s	2'925,5193.47
	65 s	12'192,114.99

Las variables seleccionadas presenta valores de FIV muy por encima del criterio establecido, es decir, las variables son multicolineales, por lo tanto, es posible reducir el número de éstas sin una pérdida importante de información en el proceso.

Regresión sobre componentes principales

Para reducir la multicolinealidad de las variables predictoras con el propósito de obtener mejores modelos de predicción, se realiza un análisis de componentes principales para transformar las variables originales en nuevas variables no correlacionadas. Este procedimiento es aplicado para ambos tipos de detección. Las variables "z_i", encontradas se muestran en la tabla 45 y se observa que éstas no son colineales entre sí.

Tabla 45. Variables predictoras transformadas a partir de componentes principales para datos voltamperométricos y cronoamperométricos.

Detección	Variable	FIV
Voltamperométrica	z_1^V	1.00
	z_2^V	1.00
	z_3^V	1.00
Cronoamperométrica	z_1^C	1.00
	z_2^C	1.00
	z_3^C	1.00

Tabla 46. Coeficientes de determinación múltiple, de determinación ajustado y de determinación predicho para detección voltamperométrica y cronoamperométrica de GLU, COL y TG calculados en la regresión RCP.

Detección	$R_{x,y}^2$			R_{aj}^2			R_{Pred}^2		
	GLU	COL	TG	GLU	COL	TG	GLU	COL	TG
Voltamperométrica	0.992	0.980	0.804	0.989	0.970	0.706	0.959	0.891	0.039
Cronoamperométrica	0.884	0.814	0.816	0.826	0.720	0.724	0.714	0.555	0.319

GLU = glucosa, COL = colesterol, TG = triglicéridos.

Tabla 47. Resultados de porcentajes de error en la determinación simultánea de glucosa, colesterol y triglicéridos mediante cronoamperometría y voltamperometría utilizando RCP.

Mezcla	% error (Cronoamperometría)			% error (Voltamperometría)		
	GLU	COL	TG	GLU	COL	TG
M ₁	10.5	7.6	1.4	3.1	3.8	20.4
M ₂	2.1	2.8	49.0	11.2	8.1	23.0
M ₃	19.7	30.8	11.5	0.3	1.1	10.4
M ₄	1.0	0.1	17.5	1.5	3.5	18.9
M ₅	32.2	20.3	45.0	5.0	4.9	30.8
M ₆	3.4	3.7	20.4	3.3	6.5	29.5
M ₇	39.2	95.4	35.6	11.0	29.3	38.9
M ₈	10.4	13.0	10.8	6.5	8.3	19.4
M ₉	145.0	43.3	26.9	16.0	5.9	80.2
M ₁₀	18.1	10.8	15.5	3.8	3.4	10.1
% EP	28.2	22.7	26.2	6.2	7.5	28.2

GLU = glucosa, COL = colesterol, TG = triglicéridos, % EP = porcentaje de error promedio.

Los coeficientes de correlación ($R_{x,y}^2$, R_{aj}^2 y R_{Pred}^2) correspondientes para cada analito se muestran en la tabla 46, donde se observa que estos parámetros tienen valores inferiores a los encontrados para RLM. Se puede afirmar que los porcentajes de error serán mayores, tanto para recalcular las concentraciones de trabajo como para la estimación de nuevas mediciones.

Al procesar la información cronoamperométrica y voltamperométrica mediante RCP, se encontraron porcentajes de error promedio más elevados en comparación con los calculados mediante RLM (Tabla 47). Estos porcentajes de error indican que las nuevas variables generadas han perdido gran parte de la información útil para obtener un modelo robusto para la regresión.

4.1.3.4. Regresión sobre mínimos cuadrados parciales

En el caso de la regresión MCP, los porcentajes de error relativo alcanzados disminuyeron en el caso voltamperométrico y fueron del mismo orden en el cronoamperométrico (Tabla 48). Debido a que se usaron las mismas variables que las utilizadas en RLM y RCP los resultados son muy similares.

Tabla 48. Resultados de porcentajes de error en la determinación simultánea de glucosa, colesterol y triglicéridos mediante cronoamperometría y voltamperometría utilizando MCP.

Mezcla	% error (Cronoamperometría)			% error (Voltamperometría)		
	GLU	COL	TG	GLU	COL	TG
M ₁	10.2	11.8	12.6	3.4	4.0	16.9
M ₂	22.3	10.8	27.1	2.1	2.9	24.4
M ₃	17.3	26.5	10.7	0.7	1.4	9.4
M ₄	7.3	14.9	17.7	0.3	1.3	19.2
M ₅	38.4	22.7	36.8	0.2	0.3	38.8
M ₆	9.1	10.0	29.3	0.4	1.1	27.9
M ₇	37.6	88.7	34.2	0.3	0.4	34.6
M ₈	4.8	8.7	17.5	0.4	0.8	18.7
M ₉	52.4	12.7	64.3	2.2	0.8	60.2
M ₁₀	30.4	21.8	15.2	0.6	0.6	11.8
% EP	23	22.9	26.5	1.1	1.4	26.2

GLU = glucosa, COL = colesterol, TG = triglicéridos, % EP = porcentaje de error promedio.

4.1.3.5. Resumen de resultados

En la mayoría de las determinaciones se encuentra sistemáticamente que la concentración de triglicéridos reporta mayores errores en su predicción, tanto para datos voltamperométricos como cronoamperométricos ; situación debida a la menor contribución en la intensidad de corriente del biosensor debido a su menor movilidad en disolución acuosa y su menor concentración analítica en los sistemas (Tabla 49).

Tabla 49. Resumen de resultados para la determinación simultánea de glucosa, colesterol y triglicéridos mediante voltamperometría, cronoamperometría y regresión multivariante.

Detección	Regresión	% Error promedio		
		Glucosa	Colesterol	Triglicéridos
Voltamperometría	RLM	1.1	1.3	13.9
	RCP	25.6	24.2	20.6
	MCP	1.1	1.4	26.2
Cronoamperometría	RLM	26.3	17.3	14.9
	RCP	22.9	22.9	26.2
	MCP	23.0	22.9	26.5

4.1.4. Conclusiones

El presente trabajo demuestra que es posible desarrollar un biosensor amperométrico multienzimático para cuantificar simultánea y eficientemente glucosa y colesterol con un solo dispositivo mediante el procesamiento de la información voltamperométrica mediante RLM o MCP. En caso de los datos voltamperométricos, su porcentaje de error promedio para la determinación de los tres analitos es mayor al 10.0 % en todos los casos, lo que en un análisis experimental empleando esta técnica no se esperan resultados tan aceptables como los obtenidos mediante voltamperometría. En el caso de la cuantificación de triglicéridos, su esquema de detección le proporciona porcentajes de error mayores de entre los tres analitos.

Para el desarrollo de un biosensor multienzimático capaz de analizar muestras se sugiere comenzar con un biosensor que determine la concentración de dos analitos simultáneamente empleando un transductor voltamperométrico y realizando el tratamiento de los datos mediante RLM y MCP.

4.2. DESARROLLO DE BIOSENSORES ENZIMÁTICOS PARA LA DETERMINACIÓN SIMULTÁNEA DE GLUCOSA Y ÁCIDO ASCÓRBICO MEDIANTE VOLTAMPEROMETRÍA Y REGRESIÓN MÚLTIVARIANTE EN TABLETAS EFERVESCENTES

4.2.1. Introducción

Hasta el momento, el biosensor de glucosa es el biosensor más desarrollado de su clase. No obstante, uno de los interferentes más problemáticos en su aplicación en muestras es el ácido ascórbico. Éste compuesto y sus sales de sodio, potasio y calcio se utilizan de forma general como antioxidantes. Se encuentra de forma natural o añadida en una gran cantidad de alimentos y bebidas. El cuerpo humano es incapaz de sintetizarlo por sí mismo, sin embargo, sus funciones en el organismo son importantes como por ejemplo en el fortalecimiento del sistema inmunitario y en la conservación de las paredes celulares [70]. En las mediciones empleando biosensores enzimáticos voltamperométricos el ácido ascórbico se oxida sobre la superficie del electrodo provocando un incremento en la intensidad de corriente que no es debida a la acción de la enzima sobre el sustrato, afectando la exactitud del análisis [71]. Para minimizar el efecto de este y otros interfe-

rentes, se han desarrollado métodos sofisticados para la preparación de biosensores altamente selectivos de glucosa.

Tabla 50. Biosensores de glucosa que utilizan TCNQ como mediador y su interferencia con ácido ascórbico.

Transductor:Mediador	Modificación adicional	Sensibilidad (µA/mM)	Relación [AA]/[glucosa] (mM)	% Interferencia con AA	Referencia
Alambre de Platino: TTF-TCNQ	Sin modificación	1.5	1.00/0.25	100	[72]
Electrodo de platino: TTF/TCNQ	Película de polipirrol	0.07	0.10/5.00	3.1	[73]
Electrodo de oro: TTF	Monocapa de ácido mercaptopropiónico	0.219	1.00/1.00	20	[74]
Compósito TiO ₂ /APV/VPD: TTF	Membrana de Nafion 117	0.405	-	0	[75]
Compósito PVC/TTF-TCNQ	Sin modificación	3.5	1.00/1.00	320	[76]
Compósito PVC/TTF-TCNQ/Nanopartículas de oro	Sin modificación	45	1.00/1.00	525	[77]

Una estrategia ampliamente estudiada es emplear una membrana selectiva que minimice el acceso de compuestos oxidables a la superficie del transductor. Para ello se utilizan capas, multicapas o capas mixtas de polímeros con propiedades de transporte basadas en el tamaño, la carga o la polaridad que le otorgan cierta permeabilidad para la oxidación del analito de interés [78].

Otra estrategia es el uso de receptores de electrones que permiten el funcionamiento del biosensor a un menor potencial como en el caso de los biosensores de segunda generación. Sin embargo, mediadores comúnmente utilizados como el TCNQ y el TTF pueden ser reducidos por el ácido ascórbico [79], [80]. Algunos ejemplos se muestran en la tabla 50.

Los biosensores de glucosa que utilizan TCNQ como mediador presentan una interferencia electroquímica importante con el ácido ascórbico debido a que este compuesto reduce espontáneamente parte del mediador en el biosensor. La señal electroquímica que se registra en un medio que contiene glucosa es la misma que se obtiene en un medio que contiene sólo ácido ascórbico porque es debida a la reducción del mediador, que es inducida por dos fenómenos diferentes e independientes entre sí [81].

Los modelos de regresión multivariante permiten aprovechar a plenitud la señal analítica, determinando simultáneamente dos o más compuestos en la muestra mediante el procesamiento correcto de los datos. Así, estas herramientas estadísticas hacen posible procesar información analítica de dispositivos no selectivos favoreciendo la simplificación en su construcción y su aplicación en análisis multicomponente simultáneo.

Este trabajo explora el desarrollo de un biosensor enzimático de glucosa capaz de cuantificar simultáneamente glucosa y ácido ascórbico mediante detección voltamperométrica y regresión multivariante.

4.2.2. Reactivos y soluciones

Todas las soluciones fueron preparadas con reactivos comerciales de grado analítico. KCl y KH_2PO_4 de J.T. Baker. $\text{K}_2\text{HPO}_4 \cdot 3\text{H}_2\text{O}$, glucosa, ácido ascórbico, glutaraldehído, glicina, 7, 7, 8, 8-tetracianoquinodimetano o TCNQ, tetrahidrofurano o THF y glucosa oxidasa (GOx) 163,400 UIg^{-1} , EC 1.1.3.4 de *Aspergillus niger* de Sigma-Aldrich. Amilasa de origen natural. Las disoluciones fueron preparadas con agua desionizada Milli Q (Millipore) con una resistividad $\geq 18 \text{ M}\Omega \text{ cm}$. El electrolito soporte fue preparado con una solución buffer de fosfatos 0.05 M y 0.1 M de KCl a pH 7.0 (PBS).

Se analizaron muestras de tabletas efervescentes comerciales sabor limón y sabor naranja. Las soluciones madre o stock de glucosa se prepararon en el electrolito soporte al menos con 12 horas de anticipación a su uso para alcanzar el equilibrio de mutarrotación [82].

4.2.2.1. Construcción de biosensores

El biosensor de glucosa fue construido sobre la superficie de un electrodo de carbón vítreo (BASi MF-2012) con un diámetro de 3.0 mm (7.07 mm^2). Primeramente se depositaron sobre la superficie del electrodo 10 μL de una solución de TCNQ 0.4 mg mL^{-1} en acetonitrilo para su posterior evaporación a temperatura ambiente, este paso se repite dos veces más. Posteriormente se depositaron 20 μL del coctel enzimático equivalente a 32.8 UI de glucosa oxidasa (GOx), dejando evaporar nuevamente a temperatura ambiente. Una unidad internacional para la enzima GOx está definida como aquella cantidad que oxida 1.0 μmol de β -D-glucosa a D-gluconato y peróxido de hidrógeno por minuto a pH 5.1 y 35 °C [83].

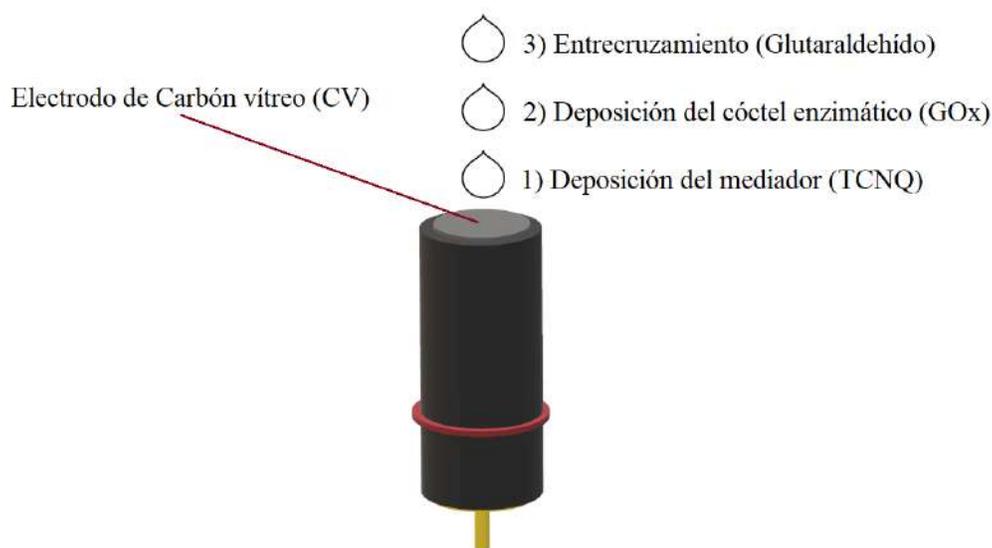


Figura 25. Proceso de fabricación de un biosensor enzimático de glucosa.

Para la inmovilización de la enzima sobre el electrodo se agregan 20 μL de glutaraldehído al 0.4% en PBS dejando evaporar a temperatura ambiente (Fig. 25). Finalmente, son adicionados 30 μL de una disolución de glicina 0.1 M durante cinco minutos, para desactivar el exceso del agente entrecruzante (glutaraldehído). Una vez completado el procedimiento, se enjuaga con agua desionizada.

4.2.2.2. Instrumentación y software

Los voltamperogramas fueron obtenidos mediante un potenciostato/galvanostato Metrohm AUTOLAB Eco Chemie AUT302N.v. con un software de control NOVA 2.0. La celda electroquímica se constituye con el biosensor como electrodo de trabajo (ET), un electrodo auxiliar (EA) de platino (BASi MF-2013) con un diámetro de 1.6 mm (2.01 mm^2), y un electrodo de referencia (ER) de Ag/AgCl (BASi MW-2021).

Para realizar la construcción de la matriz de respuesta y de validación, la selección de las variables predictoras y las regresiones múltiples se utilizó el software Minitab 17.

4.2.3. Resultados y discusión

Para determinar simultáneamente glucosa y ácido ascórbico se diseñó una matriz de concentraciones estándar empleando un arreglo ortogonal L^3 donde $L = 3$ es decir, se prepararon nueve mezclas de tres concentraciones diferentes de cada analito (0.5, 2.0 y 3.5 M). Como se observa en la tabla 51.

Tabla 51. Matriz de concentraciones para la determinación simultánea de ácido ascórbico y glucosa.

Mezcla	Concentración (mM)	
	Ácido ascórbico	Glucosa
M ₁	0.50	0.50
M ₂	0.50	2.00
M ₃	0.50	3.50
M ₄	2.00	0.50
M ₅	2.00	2.00
M ₆	2.00	3.50
M ₇	3.50	0.50
M ₈	3.50	2.00
M ₉	3.50	3.50

A cada mezcla se le hizo un barrido voltamperométrico lineal en sentido anódico a una velocidad de barrido de 40 mVs^{-1} . La ventana de potencial se estableció entre 0.16 y 0.65 V con un tiempo de 5 minutos de preacondicionamiento al potencial de 0.16 V (Fig. 26).

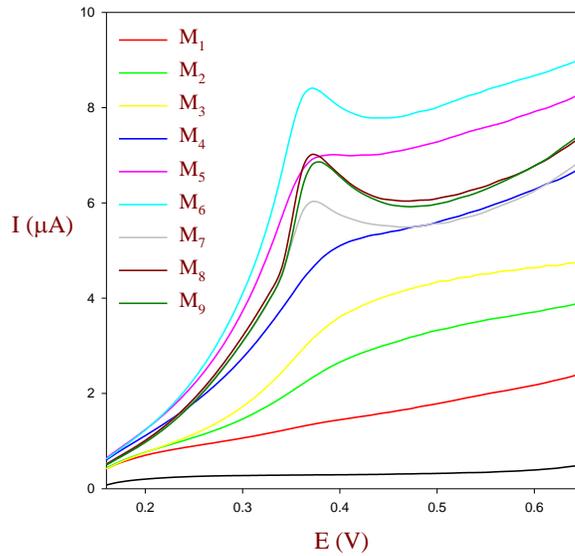


Figura 26. Voltamperogramas lineales de mezclas de concentración conocida de glucosa y ácido ascórbico en PBS pH=7. $v = 40 \text{ mVs}^{-1}$.

4.2.3.1. Stepwise RLM

Los datos fueron tratados en su totalidad con el método stepwise para la selección de las variables predictoras que mejor se correlacionan a la matriz de concentración para generar la ecuación de regresión. Las ecuaciones de regresión resultantes se muestran a continuación, mientras que los resultados de los valores encontrados para α_E , α_R y de los coeficientes de correlación se presentan en la tabla 52.

Tabla 52. Resultados de los coeficientes de correlación empleando el modelo de regresión MLR y el método stepwise para la selección de las variables de respuesta.

Parámetro	Ácido ascórbico	Glucosa
α_E	0.30	0.30
α_R	0.30	0.30
$R_{x,y}^2$	0.998	0.999
R_{aj}^2	0.994	0.999
R_{pred}^2	0.952	0.997

Las ecuaciones de regresión para ambos analitos requirió que el valor de α_E fuese de 0.30, mayor al usual que es de 0.15. El valor de R_{Pred}^2 cercano a 1 en ambos casos indica que el modelo es capaz de predecir correctamente nuevas mediciones. Sin embargo, en el caso de la ecuación de regresión para glucosa se puede observar un sobreajuste debido a que los coeficientes $R_{x,y}^2$ y R_{aj}^2 son muy cercanos a 1 y tienen el mismo valor.

Para verificar la posible multicolinealidad que explique el sobreajuste se calcula el FIV para las regresiones de ácido ascórbico y glucosa (Tablas 53 y 54).

Tabla 53. Parámetros de multicolinealidad de las variables predictoras seleccionadas para la regresión de ácido ascórbico.

Variable	FIV
$I_{0.160 V}$	58.63
$I_{0.261 V}$	323.53
$I_{0.351 V}$	160.68
$I_{0.372 V}$	1729.50
$I_{0.382 V}$	1079.30

Tanto para la regresión del ácido ascórbico como para la regresión de glucosa, todas las variables de respuesta presentan una gran multicolinealidad. Es posible transformar o reducir el número de variables para que el modelo resulte ser robusto para nuevas mediciones.

Tabla 54. Parámetros de multicolinealidad de las variables predictoras seleccionadas para la regresión de glucosa.

Variable	FIV
$I_{0.221 V}$	42.62
$I_{0.331 V}$	98.09
$I_{0.492 V}$	116859.77
$I_{0.502 V}$	452543.11
$I_{0.513 V}$	478436.63
$I_{0.543 V}$	463808.34
$I_{0.553 V}$	91359.57

Tabla 55. Concentraciones pronosticadas y porcentajes de error en la determinación ácido ascórbico y glucosa mediante Stepwise-RLM.

Mezcla	Concentración estándar (mM)		Concentración pronosticada (mM)		% de Error	
	Ácido ascórbico	Glucosa	Ácido ascórbico	Glucosa	Ácido ascórbico	Glucosa
M ₁	0.50	0.50	0.48	0.50	4.6	<0.1
M ₂	0.50	2.00	0.56	2.00	11.9	<0.1
M ₃	0.50	3.50	0.49	3.50	1.3	<0.1
M ₄	2.00	0.50	1.95	0.50	2.6	<0.1
M ₅	2.00	2.00	2.11	2.00	5.5	<0.1
M ₆	2.00	3.50	1.94	3.51	2.8	0.3
M ₇	3.50	0.50	3.48	0.51	0.7	2.0
M ₈	3.50	2.00	3.57	2.00	2.0	<0.1
M ₉	3.50	3.50	3.47	3.50	0.8	<0.1
% de Error promedio					3.6	0.2

Al aplicar las ecuaciones de regresión encontradas para recalculas las concentraciones de los analitos se obtienen excelentes resultados. Los porcentajes de error promedio son inferiores al 3.6 %

y 0.2 % para ácido ascórbico y glucosa respectivamente. Por lo tanto, a pesar de la alta multicolinealidad, el modelo de regresión se ajusta adecuadamente a los valores de los datos de entrada (Tabla 55).

4.2.3.2. Regresión sobre componentes principales

Para reducir la multicolinealidad de las variables predictoras se realiza un análisis de componentes principales para transformar las variables originales en nuevas variables no correlacionadas. Este procedimiento es aplicado para ambos analitos. Las variables " z_i ", encontradas se muestran en la tabla 56 y se observa una importante reducción de los valores de multicolinealidad en comparación con RLM, sin embargo, la mayoría superan el valor de 10.0 considerado como límite.

Tabla 56. Variables predictoras transformadas a partir de componentes principales para ácido ascórbico y glucosa.

Analito	Variable	FIV
Ácido ascórbico	z_1^{AA}	387.15
	z_2^{AA}	341.55
	z_3^{AA}	5.42
Glucosa	z_1^{Glu}	339.12
	z_2^{Glu}	264.94
	z_3^{Glu}	10.32

Los coeficientes de correlación ($R_{x,y}^2$, R_{aj}^2 y R_{Pred}^2) correspondientes para cada analito se muestran en la tabla 57, donde se observa que estos parámetros tienen valores inferiores a los encontrados para RLM. Se puede afirmar que los porcentajes de error serán mayores, tanto para recalcular las concentraciones de trabajo como para la estimación de nuevas mediciones.

Tabla 57. Coeficientes de determinación múltiple, de determinación ajustado y de determinación predicho para la cuantificación de ácido ascórbico y glucosa mediante RCP.

Analito	$R_{x,y}^2$	R_{aj}^2	R_{Pred}^2
Ácido ascórbico	0.813	0.710	0.549
Glucosa	0.864	0.783	0.520

Al procesar los datos mediante RCP, se encontraron porcentajes de error promedio más elevados en comparación con los calculados mediante RLM (Tabla 58). Estos porcentajes de error indican que las nuevas variables generadas han perdido gran parte de la información útil para obtener un modelo robusto para la regresión.

Tabla 58. Concentraciones pronosticadas y porcentajes de error en la determinación ácido ascórbico y glucosa mediante RCP.

Mezcla	Concentración estándar (mM)		Concentración pronosticada (mM)		% de Error	
	Ácido ascórbico	Glucosa	Ácido Ascórbico	Glucosa	Ácido ascórbico	Glucosa
M ₁	0.50	0.50	0.06	0.14	88.5	71.1
M ₂	0.50	2.00	0.73	1.89	45.6	5.5
M ₃	0.50	3.50	1.18	3.74	136.1	7.0
M ₄	2.00	0.50	2.19	0.51	9.3	1.2
M ₅	2.00	2.00	2.27	1.73	13.5	13.7
M ₆	2.00	3.50	2.30	3.15	14.8	10.0
M ₇	3.50	0.50	2.28	1.40	35	179.9
M ₈	3.50	2.00	3.26	2.56	6.8	28.0
M ₉	3.50	3.50	3.75	2.94	7.3	16.0
% de Error promedio					39.6	37

4.2.3.3. Regresión sobre mínimos cuadrados parciales

En la regresión MCP, se modificaron las variables de selección de modo que la sección del voltamperograma seleccionada contenga las variables predictoras que mejor se ajusten a las variables de respuesta. En el caso de ácido ascórbico, se seleccionaron los potenciales desde 0.291 V hasta 0.603 V, mientras que para la glucosa se seleccionaron los potenciales desde 0.210 hasta 0.301 V.

De ese modo los porcentajes de error relativo alcanzados y los porcentajes de error promedio son inferiores a 0.1 % en todos los casos (Tabla 59).

Tabla 59. Concentraciones pronosticadas y porcentajes de error en la determinación ácido ascórbico y glucosa mediante MCP.

Mezcla	Concentración estándar (mM)		Concentración pronosticada (mM)		% de Error	
	Ácido ascórbico	Glucosa	Ácido ascórbico	Glucosa	Ácido ascórbico	Glucosa
M ₁	0.50	0.50	0.50	0.50	<0.1	<0.1
M ₂	0.50	2.00	0.50	2.00	<0.1	<0.1
M ₃	0.50	3.50	0.50	3.50	<0.1	<0.1
M ₄	2.00	0.50	2.00	0.50	<0.1	<0.1
M ₅	2.00	2.00	2.00	2.00	<0.1	<0.1
M ₆	2.00	3.50	2.00	3.50	<0.1	<0.1
M ₇	3.50	0.50	3.50	0.50	<0.1	<0.1
M ₈	3.50	2.00	3.50	2.00	<0.1	<0.1
M ₉	3.50	3.50	3.50	3.50	<0.1	<0.1
% de Error promedio					<0.1	<0.1

4.2.4. Resumen de resultados

De acuerdo a los porcentajes de error promedio mostrados en la tabla 60 para los tres modelos de regresión multivariante, se observa que tanto para el modelo Stepwise-RLM y MCP estos porcentajes son inferiores a al 10.0 %; estos modelos se ajustan al comportamiento de los datos. Sin embargo, para RCP, estos errores son más elevados, debido a que se perdió parte de la información del sistema al reducir y transformar las variables predictoras. En caso de muestras reales se espera por lo tanto, que los modelos que mejor se ajusten al comportamiento de las variables sean Stepwise-RLM y MCP.

Tabla 60. Resumen de resultados de la determinación simultánea de ácido ascórbico y glucosa en las mezclas de concentraciones estándar.

Analito	% Error promedio		
	Stepwise-RLM	RCP	MCP
Ácido ascórbico	3.6	39.6	<0.1
Glucosa	0.2	37.0	<0.1

4.2.5. Muestra problema

Se analizaron dos tipos de muestras: tableta efervescente sabor naranja y tableta efervescente sabor limón. Cada tableta fue pesada y pulverizada. Posteriormente se disolvió parte de la tableta en una disolución de amilasa al 10 % (V/V) en PBS y se le realizaron los respectivos barridos voltamperométricos (Fig. 27).

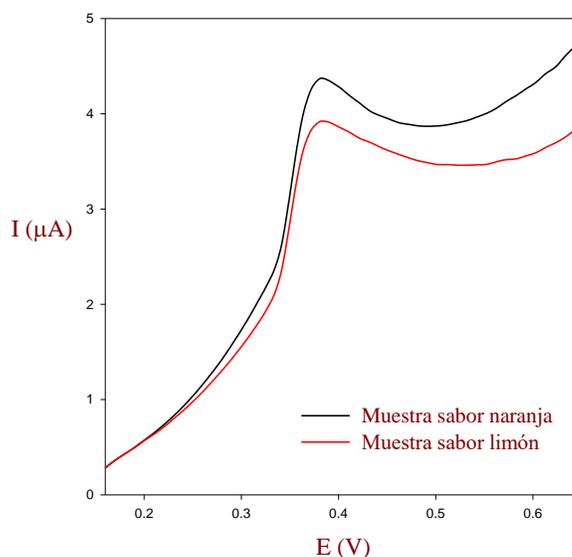


Figura 27. Voltamperogramas lineales de muestras de pastillas efervescentes que contienen glucosa y ácido ascórbico en PBS pH=7. $v = 40 \text{ mVs}^{-1}$.

El marbete indica que el 23.4 % del peso de la tableta es azúcar, no es posible establecer el tipo de azúcar que contiene, sin embargo, se asume la presencia de glucosa en la muestra y por tanto, se espera la aparición de un pico de corriente debito a la reoxidación del TCNQ para regenerar el sitio activo de la enzima. Este comportamiento es observado en la Figura 27 y es posible cuantificar la cantidad de glucosa en la tableta.

Los voltamperogramas de las muestras sabor naranja y sabor limón presentan el mismo comportamiento entre sí y entre los estándares, por lo tanto, es posible determinar la cantidad de glucosa y ácido ascórbico. Para la tableta sabor naranja, se encuentra que la cantidad de ácido ascórbico calculada es cercana al valor reportado no existiendo diferencias considerables entre los métodos. No obstante, la cantidad de glucosa difiere entre cada modelo de regresión (Tabla 61).

Tabla 61. Cantidades de ácido ascórbico y glucosa encontradas por tableta en una muestra de tableta efervescente sabor naranja.

Regresión	g encontrados por tableta		% Error promedio	
	Ácido ascórbico	Glucosa	Ácido ascórbico	Glucosa
RLM-Stepwise	0.833	1.668	16.7	-
RCP	1.027	2.691	2.7	-
MCP	1.059	0.566	5.9	-
Marbete	1	-	-	-

La concentración de ácido ascórbico calculada para la muestra sabor limón es cercana al valor reportado sin diferencias considerables entre los métodos. Sin embargo, al igual que la muestra sabor naranja, la cantidad de glucosa difiere entre cada modelo de regresión (Tabla 62).

Tabla 62. Cantidades de ácido ascórbico y glucosa encontradas por tableta en una muestra de tableta efervescente sabor limón.

Regresión	g encontrados por tableta		% Error promedio	
	Ácido ascórbico	Glucosa	Ácido ascórbico	Glucosa
RLM-Stepwise	0.718	-0.089	28.2	-
RCP	0.889	2.514	11.1	-
MCP	0.98	0.934	2	-
Marbete	1	-	-	-

4.2.6. Conclusiones

Se ha demostrado que el desarrollo de un biosensor enzimático electroquímico capaz de cuantificar simultáneamente glucosa y ácido ascórbico es factible debido a que los porcentajes de error son inferiores al 10.0 % empleando los modelos de regresión multivariante Stepwise-RLM y MCP. Los biosensores monoparamétricos comerciales tienen porcentajes de error en el orden del 20.0 %. Por lo tanto, las estrategias descritas en este trabajo son competitivas y servirán de precedentes para trabajos posteriores.

5. CONCLUSIONES GENERALES Y PERSPECTIVAS

5.1. CONCLUSIONES

En el presente trabajo se demuestra que el análisis multiparamétrico mediante dispositivos electroquímicos, es una alternativa analítica viable para cuantificar compuestos de interés biomédico utilizando procesamiento multivariante.

Se encuentra que el procesamiento de información voltamperométrica brinda mejores resultados que la cronoamperométrica, debido a que esta última exhibe una mayor colinealidad.

Entre los modelos de regresión multivariante aplicados en este estudio: regresión lineal múltiple RLM, regresión sobre componentes principales RCP y regresión sobre mínimos cuadrados parciales MCP, conforme a los resultados es posible recomendar el uso de MCP en virtud de reportar el mejor compromiso en cuanto a calidad de los resultados (menores porcentajes de error relativo) y capacidad de manejar grandes volúmenes de datos de entrada.

Las estrategias de cuantificación planteadas en este estudio, implican un cambio de paradigma en el desarrollo de métodos analíticos dirigidos a obtener la máxima selectividad, ya que hace posible cuantificar los analitos de interés en una muestra al mismo tiempo que los compuestos interferentes; enriqueciendo así la información que se puede extraer del sistema analizado.

5.2.PERSPECTIVAS

Los estudios desarrollados en este trabajo de investigación, requieren ser enriquecidos en cuanto a la validación de los resultados con otras técnicas analíticas como cromatografía y espectroscopia. Asimismo es necesario evaluar la reproducibilidad y ampliar las aplicaciones a muestras reales.

Con base en la experiencia obtenida en este trabajo, se contempla a corto plazo el desarrollo de biosensores bienzimáticos para la determinación simultánea de glucosa y colesterol.

Adicionalmente a las bondades que implica el uso de las técnicas electroquímicas de análisis aplicadas, es posible hacer la transferencia de los sistemas desarrollados a electrodos serigráficos para favorecer el análisis descentralizado a bajo costo y la miniaturización de los dispositivos.

6. REFERENCIAS

- [1] K. A. Rubinson y J. F. Rubinson , *Análisis Instrumental*, Madrid: Pearson Educación, 2001.
- [2] J. Wang, *Analytical Electrochemistry*, New York: Wiley-VCH, 2001.
- [3] R. Alkire, D. Kolb y J. Lipkowski, *Chemical modified electrodes*, Alemania: Wiley-VCH, 2009.
- [4] B. R. Eggins, *Chemical Sensors and Biosensors*, Chichester: John Wiley & Sons, 2002.
- [5] C. A. Galán Vidal, J. Muños, C. Domínguez y S. Alegret, «Chemical sensors, biosensors and thick-film technology,» *TrAC*, vol. 14, nº 5, pp. 225-231, 1995.
- [6] D. L. Nelson y M. M. Cox, *Principies of Biochemistry*, U.S.A.: W. H. Freeman, 2004.
- [7] O. R. Zaborsky, *Immobilized Enzymes*, U.S.A.: CRC Press, 1973.
- [8] W. H. Scouten, J. H. Luong y R. S. Brown, «Enzyme or Protein Immobilization Techniques for Applications in Biosensor Desing,» *Trends Biotechnol.*, vol. 13, nº 5, pp. 178-185, 1995.
- [9] D. R. Thévenot, K. Toth, R. A. Durst y G. S. Wilson, «Electrochemical Biosensors: Recomendmed Definitions and Classification,» *Biosens. Bioelectron.*, vol. 16, pp. 121-131, 2001.
- [10] G. M. Reaven, «Role of Insulin Resistenace in Human Disease,» *Diabetes*, vol. 37, pp. 1595-1607, 1988.
- [11] N. Wacher Rodarte , «Epidemiología del síndrome metabólico,» *Gac. Méd. Méx.* , vol. 145, nº 5, pp. 384-391, 2009.
- [12] G. Rocchitta, A. Spanu, S. Babudieri, G. Latte, G. Madeddu, G. Galleri, S. Nuvoli, P. Bagella, M. I. Demartis, V. Fiore, R. Manetti y P. A. Serra, «Enzime Biosensors for Biomedical Aplications: Strategies for Safeguarding Analytical Performances in Biological Fluids,» *Sensors*, vol. 16, nº 780, pp. 1-21, 2016.
- [13] T. Audesirk , G. Audesirk y B. E. Byers, *Biología. La vida en la tierra con fisiología*, México: Pearson Educación de México, 2017.
- [14] J. M. McMillin, «Chapter 141 Blood Glucose,» de *Clinical Methods: Teh History, Physical, and Laboratory Examinations*, U.S.A., Walker HK, Hall WD, Hurst JW, 1990, pp. 662-665.
- [15] P. Trinder, «Determination of blood glucose using an oxidase-peroxidase system with a non-

- carcinogenic chromogen,» *J. Clin. Path.*, vol. 22, pp. 158-161, 1969.
- [16] T. G. Solomons y C. B. Fryhle, *Organic Chemistry*, U.S.A.: John Wiley & Sons, 2009.
- [17] L. C. Clark y C. Lyons, «Electrode Systems for Continuous Monitoring in Cardiovascular Surgery,» *Ann. NY Acad. Sci.*, vol. 102, pp. 29-45, 1962.
- [18] G. G. Guilbault y G. J. Lubrano, «An enzyme electrode for the amperometric determination of glucose,» *Anal. Chim. Acta*, vol. 64, n° 3, pp. 439-455, 1973.
- [19] J. Wang, «Glucose Biosensors: 40 Years of Advances and Challenges,» *Electroanalysis*, vol. 13, n° 12, pp. 983-988, 2001.
- [20] A. Chaubey y B. D. Malhotra, «Review. Mediated biosensors,» *Biosens. Bioelectron.*, vol. 17, pp. 441-456, 2002.
- [21] C. D. Jaeger y A. J. Bard, «Electrochemical Behavior of Donor-Tetracyanoquinodimethane Electrodes in Aqueous Media,» *J. Am. Chem. Soc.*, vol. 102, n° 17, pp. 5435-5442, 1980.
- [22] T. H. Le, A. Nafady, X. Qu, A. M. Bond y L. L. Martin, «Redox and Acid-Base Chemistry of 7,7,8,8-Tetracyanoquinodimethane, 7,7,8,8-Tetracyanoquinodimethane Radical Anion, 7,7,8,8-Tetracyanoquinodimethane Dianion, and Dihydro-7,7,8,8-Tetracyanoquinodimethane in Acetonitrile,» *Anal. Chem.*, vol. 84, n° 1, pp. 2343-2350, 2012.
- [23] S. M. Grundy, «Cholesterol Metabolism in Man,» *West. J. Med.*, vol. 128, n° 1, pp. 13-25, 1978.
- [24] T. Nakaminami, S. Kuwabata y H. Yoneyama, «Electrochemical Oxidation of Cholesterol Catalyzed by Cholesterol Oxidase with Use of an Artificial Electron Mediator,» *Anal. Chem.*, vol. 69, n° 1, pp. 2367-2372, 1997.
- [25] M. Hernández Cruz, C. A. Galán Vidal, G. A. Álvarez Romero, M. T. Ramírez Silva, M. E. Páez Hernández y J. L. González Vidal, «Behavior of Two and Three Electrode Configurayion and Different Mediators in Working Electrode on Development of Disposable Screen-Printing Biosensors for Determination of Free Cholesterol,» *J. Mex. Chem. Soc.*, vol. 57, n° 1, p. 47.53, 2013.
- [26] S. K. Arya, M. Datta y B. D. Malhotra, «Recent Advances inCholesterol Biosensor,» *Biosens. Bioelectron.*, vol. 23, n° 1, pp. 1083-1100, 2007.
- [27] F. Scheller y F. Schuber, *Techniques and Instrumentation in Analytical Chemistry Volume 11: Biosensors*, Alemania: Elsevier, 1992.

- [28] C. S. Pundir y J. Narang, «Determination of Triglycerides With Special Emphasis on Biosensors: A review,» *Int. J. Biol. Macromolec.*, vol. 61, n° 1, pp. 379-389, 2013.
- [29] C. Y. Joeong, D. Y. Han, J. H. Yoon y H. C. Yoon, «Bioelectrocatalytic Sensor for Triglycerides in Human Skin Sebum Based on Enzymatic Reaction of Lipase, Glycerol Kinase and Glycerophosphate Oxidase,» *J. Biotechnol.*, vol. 175, n° 1, pp. 7-14, 2014.
- [30] J. Saurina, S. Hernández-Cassou, E. Fábregas y S. Alegret, «Cyclic Voltammetric Simultaneous Determination of Oxidizable Amino Acids Using Multivariate Calibration Methods,» *Anal. Chim. Acta*, vol. 405, pp. 153-160, 2000.
- [31] Y. Ni, L. Wang y S. Kokot, «Voltammetric Determination of Chlorpromazine Hydrochloride and Promethazine Hydrochloride with Use of Multivariate Calibration,» *Anal. Chim. Acta*, vol. 439, pp. 159-168, 2001.
- [32] Y. Ni, W. Li y S. Kokot, «Simultaneous Determination of Nitrobenzene and Nitro-substituted Phenols by Differential Pulse Voltammetry and Chemometrics,» *Anal. Chim. Acta*, vol. 431, pp. 101-113, 2001.
- [33] Y. Ni, P. Qiu y S. Kokot, «Simultaneous Determination of Three Organophosphorus Pesticides by Differential Pulse Stripping Voltammetry and Chemometrics,» *Anal. Chim. Acta*, vol. 516, pp. 7-17, 2004.
- [34] Y. Ni, Y. Wang y S. Kokot, «Simultaneous Determination of Three Fluoroquinolones by Linear Sweep Stripping Voltammetry with the Aid of Chemometrics,» *Talanta*, vol. 69, pp. 216-225, 2006.
- [35] J. N. Miller y J. C. Miller, *Statistics and Chemometrics for an Analytical Chemistry*, England: Pearson Education Limited, 2010.
- [36] D. E. Johnson, *Métodos multivariados aplicados al análisis de datos*, Mexico: International Thomson Publishing, 2000.
- [37] R. G. Brereton, «Introduction to multivariate calibration in analytical chemistry,» *Analyst*, vol. 125, n° 1, pp. 2125-2154, 2000.
- [38] J. de la Garza García, B. N. Morales Serrano y B. A. González Cavazos, *Análisis Estadístico Multivariante*, Mexico: McGraw-Hill/Interamericana Editores, 2013.
- [39] D. L. Massart, B. G. M. Vandeginste, L. M. C. Buydens, S. De Jong, P. J. Lewi y J. Smeyers Verbeke, *Handbook of Chemometrics and Qualimetrics: Part A*, Netherlands: Elsevier Science, 1997.

- [40] R. M. O'Brien, «A Caution Regarding Rules of Thumb for Variance Inflation Factors,» *Qual. Quant.*, vol. 41, n° 1, pp. 673-690, 2007.
- [41] D. S. Nascimento, M. Insausti, B. S. F. Band y S. G. Lemos , «Simultaneous Determination of Cu, Pb, Cd, Ni, Co and Zn in Bioethanol Fuel by Adsorptive Stripping Voltammetry and Multivariate Linear Regression,» *Fuel*, vol. 137, n° 1, pp. 172-178, 2014.
- [42] Q. Shen, J.-H. Jiang, G.-L. Shen y R.-Q. Yu, «Variable Selection by an Evolution Algorithm Using Modified Cp Based on MLR and PLS Modeling: QSAR Studies of Carcinogenicity of Aromatic Amines,» *Anal. Bioanal. Chem.* , vol. 375, n° 1, pp. 248-254, 2003.
- [43] T. Rajalahti y O. M. Kvalheim, «Multivariate Data Analysis in Pharmaceuticas: A Tutorial Review,» *Int. J. Pharm.*, vol. 417, n° 1, pp. 280-290, 2011.
- [44] A. Brownlee, *Statistical Theory and Methodology in Science and Engineering*, New York: John Wiley, 1965.
- [45] Y. L. Grossman, S. L. Ustin, S. Jacquemoud, E. W. Sanderson, G. Schmuck y J. Verdebout, «Critique of Stepwise Multiple Linear Regression for the Extraction of Leaf Biochemistry Information from Leaf Reflectance Data,» *Remote Sens. Environ.* , vol. 56, n° 1, pp. 182-193, 1996.
- [46] J. C. Vega-Vilca y J. Guzmán, «Regresión PLS y PCA como solución al problema de multicolinealidad en regresión múltiple,» *Rev. Mate. Teor. Aplic.*, vol. 18, n° 1, pp. 9-20, 2011.
- [47] P. Geladi y B. R. Kowalski, «Partial Least-Squares regression: A tutorial,» *Anal. Chim. Acta*, vol. 185, n° 1, pp. 1-17, 1986.
- [48] R. D. Bautista, F. J. Aberásturi, A. I. Jiménez y F. Jiménez, «Simultaneous Spectrophotometric Determination of Drugs in Pharmaceutical Preparations Using Multiple Linear Regression and Partial Least-Squares Regression, Calibration and Prediction Methods,» *Talanta*, vol. 43, n° 1, pp. 2107-2115, 1996.
- [49] H. C. Goicoechea y A. C. Olivieri, «Simultaneous determination of rifampicin, isoniazid and pyrazinamide in tablet preparations by multivariate spectrophotometric calibration,» *J. Pharm. Biomed. Anal.* , vol. 20, n° 1, pp. 681-686, 1999.
- [50] G. Zhang y J. Pan, «Simultaneous Spectrophotometric Determination of Atrazine and Cyanazine by Chemometric Methods,» *Spectrochim. Acta A*, vol. 78, n° 1, pp. 238-242, 2011.

- [51] Y.-C. He, S. Fang y X.-J. Xu, «Simultaneous Determination of Acesulfame-K, Aspartame and Stevioside in Sweetener Blends by Ultraviolet Spectroscopy with Variable Selection by siPLS Algorithm,» *Maced. J. Chem. Chem. Eng.*, vol. 31, n° 1, pp. 17-28, 2012.
- [52] S. K. Hassaninejad-Darzi, A. Samadi-Maybodi y S. M. Nikou, «UV-Vis Spectrophotometry and Multivariate Calibration Method for Simultaneous Determination of Theophylline, Montelukast and Loratadine in Tablet Preparations and Spiked Human Plasma,» *Iran. J. Pharm. Res.*, vol. 15, n° 3, pp. 379-391, 2016.
- [53] H. Khajehsharifi, E. Pourbasheer, H. Tavallali, S. Sarvi y M. Sadeghi, «The comparison of partial least squares and principal component regression in simultaneous spectrophotometric determination of ascorbic acid, dopamine and uric acid in real samples,» *Arabian J. Chem.*, vol. 10, n° 1, pp. S3451-S3458, 2017.
- [54] P. G. Conaghan, «A turbulent decade for NSAIDs: update on current concepts of classification, epidemiology, comparative efficacy, and toxicity,» *Rheumatol Int.*, vol. 32, n° 1, pp. 1491-1502, 2012.
- [55] B. J. Sanghavi y A. K. Srivastava, «Simultaneous Voltammetric Determination of Acetaminophen, Aspirin, and Caffeine Using An in Situ Surfactant-Modified Multiwalled Carbon Nanotube Paste Electrode,» *Electrochim. Acta*, vol. 55, n° 1, pp. 8638-8648, 2010.
- [56] J. S. Stefano, R. H. O. Montes, E. M. Richter y R. A. A. Muñoz, «Flow-Injection Analysis with Multiple-Pulse Amperometry for Simultaneous Determination of Paracetamol and Naproxen Using a Homemade Flow Cell for Screen-Printed Electrodes,» *J. Braz. Chem. Soc.*, vol. 25, n° 3, pp. 484-491, 2014.
- [57] G. Y. Aguilar Lira, G. A. Álvarez Romero, A. Rojas Hernández, M. E. Páez Hernández, J. A. Rodríguez Ávila y M. A. Romero Romo, «Voltammetric Analysis of Naproxen in Graphite Electrodes and Its Determination in Pharmaceutical Samples,» *Electroanalysis*, vol. 26, n° 1, pp. 1573-1581, 2014.
- [58] THE UNITED STATES PHARMACOPEIAL CONVENTION , Farmacopea de los Estados Unidos de América vol. 2, U.S.A.: USP, 2007.
- [59] THE UNITED STATES PHARMACOPEIAL CONVENTION, Farmacopea de los Estados Unidos de América vol. 3, U.S.A.: USP, 2007.
- [60] J. R. Medina, C. A. García, M. Hurtado y A. M. Domínguez Ramírez, «In Vitro Evaluation of Naproxen Sodium and Acetaminophen from Fixed-Dose Combination Generic Drugs Using the Flow-Through Cell Method,» *Int. J. Pharm. Pharm. Sci.*, vol. 8, n° 2, pp. 239-243, 2016.

6. Referencias

- [61] C. Vakh, E. Freze, A. Pochivalov, E. Evdokimova y M. Kamencev, «Simultaneous Determination of Iron (II) and Ascorbic Acid in Pharmaceuticals Based on Flow Sandwich Technique,» *J. Pharmacol. Toxicol. Methods*, vol. 73, n° 1, pp. 56-62, 2015.
- [62] Secretaría de Salud , Farmacopea de los Estados Unidos Mexicanos 11 Ed. Vol. II, CDMX: Comisión Permanente de la Farmacopea de los Estados Unidos Mexicanos, 2014.
- [63] M. W. Davey, M. Van Montagu, D. Inzé, M. Sanmartin, A. Kanellis, N. Smirnov, I. J. J. Benzie, J. J. Strain, D. Favell y J. Fletcher, «Plant L-Ascorbic Acid: Chemistry, Function, Metabolism, Bioavailability and Effects of Processing,» *J. Sci. Food Agric.*, vol. 80, n° 1, pp. 825-860, 2000.
- [64] D. Merli, A. Profumo y C. Dossi, «An Analytical Method for Fe (II) and Fe (III) Determination in Pharmaceutical Grade Iron Sucrose Complex and Sodium Ferric Gluconate Complex,» *J. Pharm. Anal.*, vol. 2, n° 6, pp. 450-453, 2012.
- [65] A. M. Pisoschi, A. F. Danet y S. Kalinowski, «Ascorbic Acid Determination in Commercial Fruit Juice Samples by Cyclic Voltammetry,» *J. Autom. Methods Manag. Chem.*, vol. 2008, n° ID 937651, pp. 1-8, 2008.
- [66] W. J. Albery, P. N. Bartlett y D. H. Craston, «Amperometric Enzyme Electrodes Part II. Conducting Salts as Electrode Materials for the Oxidation of Glucose Oxidase,» *J. Electroanal. Chem.*, vol. 194, n° 1, pp. 223-235, 1985.
- [67] A. C. F. Ribeiro, O. Ortona, S. M. N. Simoes, C. I. A. Santos, P. M. R. A. Prazeres, A. J. M. Valente, V. M. M. Lobo y H. D. Burrows, «Binary Mutual Diffusion Coefficients of Aqueous Solutions of Sucrose, Lactose, Glucose, and Fructose in the Temperature Range from (298.15 to 328.15) K,» *J. Chem. Eng. Data*, vol. 51, n° 1, pp. 1836-1840, 2006.
- [68] E. L. Cussler, Diffusion; mass transfer in fluid systems, U.S.A. : Cambridge University Press, 2009.
- [69] Z. Q. Ababneh, C. B. Berde, A. M. Ababneh, S. E. Maier y R. V. Mulkern, «In vivo lipid diffusion coefficient measurements in rat bone marrow,» *J. Magn. Reson.*, vol. 27, n° 1, p. 859, 2009.
- [70] N. Smirnov, «Ascorbic Acid: Metabolism and functions of a Multi-Faceted Molecule,» *Curr. Opin. Plant Biol.*, vol. 3, n° 3, pp. 229-235, 2000.
- [71] E.-H. Yoo y S.-Y. Lee, «Glucose Biosensors: An Overview of Use in Clinical Practice,» *Sensors*, vol. 10, n° 1, pp. 4558-4576, 2010.

- [72] Q.-S. Li, B.-C. Ye, B.-X. Liu y J.-J. Zhong, «Improvement of the performance of H₂O₂ oxidation at low working potential by incorporating TTF-TCNQ into a platinum wire electrode for glucose determination,» *Biosens. Bioelectron.*, vol. 14, n° 1, pp. 327-334, 1999.
- [73] F. Palmisano y P. G. Zambonin, «A Disposable, Reagentless, Third-Generation Glucose Biosensor Based on Overoxidized Poly(pyrrole)/Tetrathiafulvalene-Tetracyanoquinodimethane Composite,» *Anal. Chem.*, vol. 74, n° 1, pp. 5913-5918, 2002.
- [74] S. Campuzano, R. Gálvez, M. Pedrero, F. J. M. de Villena y J. M. Pingarrón, «Preparation, Characterization and Application of Alkanethiol Self-Assembled Monolayers Modified with Tetrathiafulvalene and Glucose Oxidase at a Gold Disk Electrode,» *J. Electroanal. Chem.*, vol. 526, n° 1, pp. 91-100, 2002.
- [75] X. Chen y S. Dong, «Sol-Gel-Derived Titanium Oxide/Copolymer Composite Based Glucose Biosensor,» *Biosens. Bioelectron.*, vol. 18, n° 1, pp. 999-1004, 2003.
- [76] M. Cano, J. L. Ávila, M. Mayén, M. L. Mena, J. Pingarrón y R. Rodríguez-Amaro, «A new, third generation, PVC/TTF-TCNQ composite amperometric biosensor for glucose determination,» *J. Electroanal. Chem.*, vol. 615, n° 1, pp. 69-74, 2008.
- [77] G. Sánchez-Obrero, M. Cano, J. L. Ávila, M. Mayén, M. L. Mena, J. M. Pingarrón y R. Rodríguez-Amaro, «A gold nanoparticle-modified PVC/TTF-TCNQ composite amperometric biosensor for glucose determination,» *J. Electroanal. Chem.*, vol. 634, n° 1, pp. 59-63, 2009.
- [78] J. Wang, J. Liu, L. Chen y F. Lu, «Highly Selective Membrane-Free, Mediator-Free Glucose Biosensor,» *Anal. Chem.*, vol. 66, n° 1, pp. 3600-3603, 1994.
- [79] S. N. Murthy A. y A. , «Electrochemical oxidation of L-ascorbic acid on 7,7,8,8-tetracyanoquinodimethane (TCNQ) modified electrode,» *Biosens. Bioelectron.*, vol. 9, n° 1, pp. 439-444, 1994.
- [80] S. Campuzano, Ó. A. Loaiza, M. Pedrero, F. J. M. de Villena y J. M. Pingarrón, «An integrated bienzyme glucose oxidase-fructose dehydrogenase-tetrathiafulvalene-3-mercaptopropionic acid-gold electrode for the simultaneous determination of glucose and fructose,» *Bioelectrochemistry*, vol. 63, n° 1, pp. 199-206, 2004.
- [81] M. Cano, B. Palazuela, J. L. Ávila y R. Rodríguez-Amaro, «Simultaneous Determination of Ascorbic Acid and Uric Acid by Using a PVC/TTF-TCNQ Composite Electrode as Detector in a FIA System,» *Electroanalysis*, vol. 19, n° 7, pp. 973-977, 2007.
- [82] J. M. Nelson y F. M. Beegle, «Mutarotation of Glucose and Fructose,» *J. Am. Chem. Soc.*,

6. Referencias

vol. 41, n° 4, pp. 559-575, 1919.

[83] Sigma-Aldrich, *Product Specification: Glucose oxidase from Aspergillus niger Type-VII*, Saint Louis, 2018.

7. ANEXOS

7.1. ÍNDICE DE FIGURAS

Figura 1. Esquema de detección electroquímica de una celda de tres electrodos.	7
Figura 2. Esquema de funcionamiento de un biosensor.	8
Figura 3. Esquema del método enzimático de Trinder para la cuantificación de glucosa. ...	11
Figura 4. Esquema de transferencia de electrones entre el sustrato y el transductor en un biosensor enzimático de glucosa.....	12
Figura 5. Reacciones óxido-reducción del TCNQ en acetonitrilo.....	12
Figura 6. Oxidación enzimática de colesterol en presencia de COx y un receptor de electrones.....	13
Figura 7. Esquema de transferencia de electrones entre el sustrato y el transductor en un biosensor enzimático de colesterol.	13
Figura 8. Hidrólisis de un triglicérido a glicerol y ácidos grasos libres.	14
Figura 9. Adición de un grupo fosfato a una molécula de glicerol.	15
Figura 10. Oxidación de glicerol-3-fosfato a fosfato de dihidroxiacetona.	15
Figura 11. Estructuras de paracetamol, naproxeno y sus productos de oxidación.....	36
Figura 12. Voltamperogramas de naproxeno sódico en metanol en acetonitrilo/metanol al 30 % (v/v). $v = 10 \text{ mVs}^{-1}$	38
Figura 13. Voltamperogramas de paracetamol en metanol en acetonitrilo/metanol al 30 % (v/v). $v = 10 \text{ mVs}^{-1}$	38
Figura 14. Curvas de calibración lineal para el naproxeno sódico y el paracetamol en acetonitrilo/metanol al 30 % (v/v). $v = 10 \text{ mVs}^{-1}$	39
Figura 15. Voltamperogramas lineales de mezclas de paracetamol y naproxeno sódico en acetonitrilo/metanol al 30 % (v/v). $v = 10 \text{ mV s}^{-1}$	40
Figura 16. Voltamperogramas de disoluciones de muestras de medicamento en metanol en acetonitrilo/metanol al 30 % (v/v). $v = 10 \text{ mV s}^{-1}$	55
Figura 17. Estructuras del hierro (II), ácido ascórbico y sus productos de oxidación.	57
Figura 18. Voltamperogramas de sulfato hierro (II) en HCl 0.1 M. $v = 20 \text{ mVs}^{-1}$	59
Figura 19. Voltamperogramas de ácido ascórbico en HCl 0.1 M. $v = 20 \text{ mV s}^{-1}$	59
Figura 20. Curvas de calibrado independientes para hierro (II) y ácido ascórbico.	60
Figura 21. Voltamperogramas de las mezclas de hierro (II) y ácido ascórbico en HCl 0.1 M. $v = 20 \text{ mV s}^{-1}$	61
Figura 22. Voltamperogramas lineales de tres muestras de medicamento en HCl 0.1 M. $v = 20 \text{ mV s}^{-1}$	66
Figura 23. Voltamperogramas lineales de las mezclas de concentraciones de los analitos. ...	71
Figura 24. Cronoamperogramas de las mezclas de concentraciones de los analitos.	72
Figura 25. Proceso de fabricación de un biosensor enzimático de glucosa.	79
Figura 26. Voltamperogramas lineales de mezclas de concentración conocida de glucosa y ácido ascórbico en PBS pH=7. $v = 40 \text{ mVs}^{-1}$	81

Figura 27. Voltamperogramas lineales de muestras de pastillas efervescentes que contienen glucosa y ácido ascórbico en PBS pH=7. $v = 40 \text{ mVs}^{-1}$85

7.2. ÍNDICE DE TABLAS

Tabla 1. Determinaciones electroquímicas simultáneas empleando un solo electrodo de trabajo.....	16
Tabla 2. Análisis de la varianza para RLM.....	21
Tabla 3. Cálculo de componentes principales.....	29
Tabla 4. Ejemplos de matrices de concentración empleados en regresión multivariante.	33
Tabla 5. Parámetros analíticos de las calibraciones lineales de naproxeno sódico y paracetamol.....	39
Tabla 6. Mezclas de concentraciones de los estándares de naproxeno y paracetamol y su respuesta en intensidad de corriente a los potenciales seleccionados.	40
Tabla 7. Concentraciones pronosticadas y porcentajes de error promedio en la determinación simultánea de naproxeno sódico y paracetamol mediante RLM.	42
Tabla 8. Análisis ANOVA para la regresión de naproxeno sódico.....	42
Tabla 9. Resultados de la estadística SCER para la regresión de naproxeno mediante MLR.	44
Tabla 10. Análisis ANOVA para la regresión de paracetamol mediante RLM.	44
Tabla 11. Coeficientes de correlación de la regresión del paracetamol.	45
Tabla 12. Parámetros de multicolinealidad entre las variables predictoras seleccionadas para la determinación simultánea de naproxeno y paracetamol.	45
Tabla 13. Valores propios calculados para las variables de respuesta de las variables predictoras en la determinación simultánea de naproxeno y paracetamol.	46
Tabla 14. Componentes principales de la matriz X.	47
Tabla 15. Variables de respuesta y variables predictoras para la RCP para la determinación simultánea de naproxeno y paracetamol.	48
Tabla 16. Concentraciones pronosticadas y porcentajes de error promedio en la determinación simultánea de naproxeno sódico y paracetamol mediante RCP.	49
Tabla 17. Coeficientes de correlación calculados a partir de análisis ANOVA de la RCP. ..	49
Tabla 18. Parámetros de multicolinealidad de las variables predictoras Z.....	49
Tabla 19. Parámetros estadísticos de las variables.	50
Tabla 20. Variables de respuesta y variables predictoras para la Regresión MCP.....	50
Tabla 21. Concentraciones pronosticadas y porcentajes de error en la determinación simultánea de naproxeno sódico y paracetamol mediante MCP.	54
Tabla 22. Porcentajes de error promedio de los estándares utilizando los modelos de regresión de trabajo.	54
Tabla 23. Concentraciones de naproxeno sódico y de paracetamol y sus respectivas intensidades de corriente de soluciones de muestras del medicamento.	55
Tabla 24. Resultados del análisis por triplicado del medicamento mediante RLM.....	56

Tabla 25. Resultados del análisis del medicamento.	56
Tabla 26. Parámetros analíticos de voltamperometría lineal para sulfato ferroso y ácido ascórbico.	60
Tabla 27. Matriz de concentración para la determinación simultánea de hierro (II) y ácido ascórbico.	61
Tabla 28. Resultados de los coeficientes de correlación empleando el modelo de regresión MLR y el método stepwise para la selección de las variables de respuesta.	62
Tabla 29. Parámetros de multicolinealidad de las variables predictoras seleccionadas para la regresión de hierro (II).	62
Tabla 30. Parámetros de multicolinealidad de las variables predictoras seleccionadas para la regresión de ácido ascórbico.	62
Tabla 31. Concentraciones pronosticadas y porcentajes de error en la determinación simultánea de hierro (II) y ácido ascórbico mediante Stepwise-MLR.	63
Tabla 32. Parámetros de multicolinealidad de las variables predictoras seleccionadas para la regresión de hierro (II) mediante RCP.	63
Tabla 33. Parámetros de multicolinealidad de las variables predictoras seleccionadas para la regresión de hierro (II) mediante RCP.	64
Tabla 34. Coeficientes de correlación calculados a partir de RCP para hierro (II) y ácido ascórbico.	64
Tabla 35. Resultados de la determinación simultánea de hierro (II) y ácido ascórbico mediante RCP.	64
Tabla 36. Resultados de la determinación simultánea de hierro (II) y ácido ascórbico mediante MCP.	65
Tabla 37. Porcentajes de error promedio de los estándares de hierro (II) y ácido ascórbico utilizando los modelos de regresión de trabajo.	65
Tabla 38. Resultados de la determinación de hierro (II) y ácido ascórbico en una muestra mediante los modelos de regresión de trabajo.	66
Tabla 39. Datos y parámetros de entrada para la simulación de un biosensor multienzimático.	71
Tabla 40. Matriz de concentraciones y de respuestas de las mezclas de los analitos para un transductor voltamperométrico.	72
Tabla 41. Matriz de concentraciones y de respuestas de las mezclas de los analitos para un transductor cronoamperométrico.	73
Tabla 42. Resultados de porcentajes de error en la determinación simultánea de glucosa, colesterol y triglicéridos mediante cronoamperometría y voltamperometría utilizando RLM.	73
Tabla 43. Coeficientes de determinación múltiple, de determinación ajustado y de determinación predicho para detección voltamperométrica y cronoamperométrica de GLU, COL y TG encontrados en la regresión RLM.	74
Tabla 44. Valores del FIV para la determinación simultánea de GLU, COL y TG de las variables de respuesta seleccionadas para voltamperometría y cronoamperometría.	74

Tabla 45. Variables predictoras transformadas a partir de componentes principales para datos voltamperométricos y cronoamperométricos.	75
Tabla 46. Coeficientes de determinación múltiple, de determinación ajustado y de determinación predicho para detección voltamperométrica y cronoamperométrica de GLU, COL y TG calculados en la regresión RCP.	75
Tabla 47. Resultados de porcentajes de error en la determinación simultánea de glucosa, colesterol y triglicéridos mediante cronoamperometría y voltamperometría utilizando RCP.	75
Tabla 48. Resultados de porcentajes de error en la determinación simultánea de glucosa, colesterol y triglicéridos mediante cronoamperometría y voltamperometría utilizando MCP.	76
Tabla 49. Resumen de resultados para la determinación simultánea de glucosa, colesterol y triglicéridos mediante voltamperometría, cronoamperometría y regresión multivariante. .	77
Tabla 50. Biosensores de glucosa que utilizan TCNQ como mediador y su interferencia con ácido ascórbico.	78
Tabla 51. Matriz de concentraciones para la determinación simultánea de ácido ascórbico y glucosa.	80
Tabla 52. Resultados de los coeficientes de correlación empleando el modelo de regresión MLR y el método stepwise para la selección de las variables de respuesta.	81
Tabla 53. Parámetros de multicolinealidad de las variables predictoras seleccionadas para la regresión de ácido ascórbico.	82
Tabla 54. Parámetros de multicolinealidad de las variables predictoras seleccionadas para la regresión de glucosa.	82
Tabla 55. Concentraciones pronosticadas y porcentajes de error en la determinación ácido ascórbico y glucosa mediante Stepwise-RLM.	82
Tabla 56. Variables predictoras transformadas a partir de componentes principales para ácido ascórbico y glucosa.	83
Tabla 57. Coeficientes de determinación múltiple, de determinación ajustado y de determinación predicho para la cuantificación de ácido ascórbico y glucosa mediante RCP.	83
Tabla 58. Concentraciones pronosticadas y porcentajes de error en la determinación ácido ascórbico y glucosa mediante RCP.	84
Tabla 59. Concentraciones pronosticadas y porcentajes de error en la determinación ácido ascórbico y glucosa mediante MCP.	84
Tabla 60. Resumen de resultados de la determinación simultánea de ácido ascórbico y glucosa en las mezclas de concentraciones estándar.	85
Tabla 61. Cantidades de ácido ascórbico y glucosa encontradas por tableta en una muestra de tableta efervescente sabor naranja.	86
Tabla 62. Cantidades de ácido ascórbico y glucosa encontradas por tableta en una muestra de tableta efervescente sabor limón.	86

Tabla 63. Pérdida de apilamiento por día respecto a la velocidad, temperatura y concentración..... 106

Tabla 64. Coeficientes de regresión de pérdida de apilamiento respecto de cada variable predictora (velocidad, temperatura y concentración). 107

Tabla 65. Coeficientes de regresión de pérdida de apilamiento respecto de velocidad, temperatura y velocidad, concentración. 107

Tabla 66. Coeficientes de regresión de pérdida de apilamiento respecto de las variables predictoras (velocidad, temperatura y concentración)..... 108

7.3.ECUACIONES EMPLEADAS PARA EL ANÁLISIS ANOVA

7.3.1. Sumas de cuadrados

$$SC_{Reg} = \sum_i n_i(\hat{y}_i - \bar{y})^2$$

$$SC_{ERes} = \sum_i \sum_j (y_{ij} - \hat{y}_i)^2 = SS_T - SS_{Reg}$$

$$SC_{FA} = \sum_i n_i(\bar{y}_i - \hat{y}_i)^2 = SS_{Res} - SS_{PE}$$

$$SC_{EExp} = \sum_i \sum_j (y_{ij} - \bar{y}_i)^2$$

$$SC_T = \sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{y})^2$$

Donde: SC_{Reg} = Suma de cuadrados de la regresión, SC_{ERes} = Suma de cuadrados del error residual, SC_{FA} = Suma de cuadrados por falta de ajuste, SC_{EExp} = Suma de cuadrados del error experimental puro y SC_T = Suma de cuadrados totales.

7.3.2. Cuadrados medios

$$MC_{Reg} = \frac{SC_{Reg}}{p - 1}$$

$$MC_{ERes} = \frac{SC_{ERes}}{n - p}$$

$$MC_{FA} = \frac{SC_{FA}}{k - p}$$

$$MC_{EExp} = \frac{SC_{EExp}}{n - k}$$

$$MC_T = \frac{SC_T}{n - 1}$$

Donde: CM_{Reg} = Cuadrados medios de la regresión, CM_{ERes} = Cuadrados medios del error residual, CM_{FA} = Cuadrados medios por falta de ajuste, CM_{EExp} = Cuadrados medios del error experimental puro y CM_T = Cuadrados medios totales.

7.4. METODOLOGÍA DE SELECCIÓN DE VARIABLES STEPWISE

A continuación se describe la metodología stepwise:

- **Paso 1:** Establecer un nivel de significación de entrada para decidir cuándo introducir un predictor. Este valor se representa con el símbolo α_E y generalmente es mayor que el nivel establecido para una regresión lineal simple que es de $\alpha_E = 0.05$ de modo que no se eliminen inmediatamente predictores del modelo. También se debe establecer un nivel de significación para decidir cuándo eliminar un predictor, denotado por α_R . Programas estadísticos como Minitab 17. establecen por defecto un $\alpha_E = 0.15$ y un $\alpha_R = 0.15$.
- **Paso 2:** Se hacen las regresiones lineales de la variable de respuesta con cada una de las variables predictoras: "y" respecto de " x_1 ", "y" respecto de " x_2 ", "y" respecto de " x_3 ", etcétera. De los predictores cuya prueba "t" arroja un valor de probabilidad "P" que sea menor que $\alpha_E = 0.15$ será el primer predictor puesto en el modelo stepwise debido a que es el más correlacionado con la variable de respuesta. En caso de que existan más predictores que cumplan con esa condición se seleccionará el que presenta un valor de "P" más bajo. Si ningún predictor tiene un valor de "P" menor que α_E se detiene el procedimiento y se ajustan nuevamente los valores seleccionados de α_E y α_R .
- **Paso 3:** En caso de hallar un predictor " x_i " cuyo valor de "P" de la prueba "t" sea menor a α_E , se realizan regresiones múltiples de "y" respecto de " x_i " y de " x_1 ", de "y" respecto de " x_i " y de " x_2 ", de "y" respecto de " x_i " y de " x_3 ", etcétera. De los predictores cuyo valor de "P" es menor que α_E , el segundo predictor puesto en el modelo stepwise " x_j " es el predictor que tiene el menor valor de "P". Si ningún predictor tiene un valor de "P" de la prueba "t" menor que α_E el proceso se detiene con sólo una variable predictora.
- **Paso 4:** Dado que " x_i " fue el primer predictor en el modelo, se tienen que verificar si al ingresar " x_j " en el modelo stepwise afectó de alguna manera la significación del predictor " x_i ". Es decir, comprobar que el valor de "P" calculado mediante el estadístico "t" sea menor que α_E . Si el valor de "P" de la prueba "t" se ha vuelto mayor que α_E se elimina " x_i " del modelo finalizando el proceso con una sola variable predictora " x_j ".
- **Paso 5:** Si se han seleccionado los predictores " x_i " y " x_j ", se continúa con la realización de regresiones múltiples de "y" respecto de " x_i ", " x_j " y de " x_1 ", de "y" respecto de " x_i ",

" x_j " y de " x_2 ", de " y " respecto de " x_i ", " x_j " y de " x_3 ", etcétera. De los predictores cuyo valor de " P " es menor que α_E , el tercer predictor puesto en el modelo stepwise " x_k " es el predictor que tiene el menor valor de " P ". Si ningún predictor tiene un valor de " P " de la prueba " t " menor que α_E el proceso se detiene con dos variables predictoras.

- **Paso 6:** Dado que " x_i " y " x_j " fueron los primeros predictores en el modelo, se tiene que verificar si al ingresar " x_k " en el modelo stepwise afectó de alguna manera la significación de los predictores " x_i " y " x_j ". Es decir, comprobar que el valor de " P " calculado mediante el estadístico " t " sea menor que α_E . Si el valor de " P " de la prueba " t " se ha vuelto

mayor que α_E para alguna de las variables, ésta se elimina del modelo y el proceso finaliza con dos variables predictoras.

- **Paso 7:** Se deberá continuar con los pasos anteriores hasta que se ya no se encuentre un predictor que produzca un valor de " P " de la prueba " t " por debajo de $\alpha_E = 0.15$.

Tabla 63. Pérdida de apilamiento por día respecto a la velocidad, temperatura y concentración.

Día	Pérdida de apilamiento (y)	Velocidad (x_1)	Temperatura (x_2)	Concentración (x_3)
1	37	80	27	88
2	18	62	23	87
3	18	62	23	87
4	19	62	24	93
5	20	62	24	93
6	15	58	23	87
7	14	58	18	80
8	14	58	17	88
9	13	58	17	88
10	11	58	18	82
11	12	58	19	93
12	8	50	18	86
13	7	50	18	86
14	8	50	18	86
15	8	50	19	80
16	9	50	19	80
17	15	56	20	82

Paso 1: Se establece un nivel de significación de entrada $\alpha_E = 0.15$ y un nivel de significación de eliminación de $\alpha_R = 0.15$.

La tabla 62 enumera una parte del conjunto de datos de pérdida de apilamiento dada por Brownlee [44]. Los datos han sido reorganizados y se obtuvieron de una planta para la oxidación de

amoníaco a ácido nítrico. La variable respuesta "y" es una medida inversa de la eficiencia general de la planta (pérdida de apilamiento). Se ha estudiado durante 17 días en función de tres variables de predicción: " x_1 " es la velocidad de operación de la planta, " x_2 " es la temperatura del agua de refrigeración que circula por las bobinas en la torre de adsorción para el ácido nítrico, y " x_3 " es la concentración de ácido circulante (en unidades arbitrarias). El objetivo de aplicar el método stepwise en esta regresión multivariante es determinar si es posible reducir el número de variables predictoras con un criterio estadístico fundamentado en el valor de la probabilidad de cometer un error tipo 1.

Paso 2: Se realizan las regresiones lineales de la variable de respuesta con cada una de las variables predictoras: "y" respecto de " x_1 ", "y" respecto de " x_2 ", y finalmente "y" respecto de " x_3 ". De acuerdo a los valores registrados en la tabla 63 de " P " respecto a la aplicación del estadístico " t ", las variables " x_1 " y " x_2 " tienen valores de " P " menores que $\alpha_E = 0.15$. Sin embargo, el valor de " P " de " x_1 " es menor al de " x_2 ", por lo tanto, la primera variable seleccionada para formar parte de la ecuación de regresión es " x_1 ".

Tabla 64. Coeficientes de regresión de pérdida de apilamiento respecto de cada variable predictora (velocidad, temperatura y concentración).

Predictor	Coficiente	Estadístico " t "	Valor de " P "
$b_{x_1}^0$	-40.03	-12.32	3.03×10^{-9}
x_1	0.94	16.89	3.58×10^{-11}
$b_{x_2}^0$	-25.37	-3.70	2.14×10^{-3}
x_2	1.96	5.87	3.08×10^{-5}
$b_{x_3}^0$	-39.41	-1.16	0.26
x_3	0.62	1.59	0.13

$b_{x_i}^0$ =término independiente de la regresión.

Paso 3: Una vez encontrado que primer predictor, se realizan regresiones múltiples, primero de "y" respecto de las variables predictoras " x_1 " y " x_2 ". Posteriormente se hace la regresión múltiple para "y" respecto de las variables predictoras " x_1 " y " x_3 ". De acuerdo a los valores registrados en la tabla 64 para la regresión con dos variables, sólo " x_2 " cumple con la condición de que el valor de " P " sea menor al valor de $\alpha_E = 0.15$.

Tabla 65. Coeficientes de regresión de pérdida de apilamiento respecto de velocidad, temperatura y velocidad, concentración.

Predictor	Coficiente	Estadístico " t "	Valor de " P "
b_{x_1, x_2}^0	-41.05	-15.17	4.39×10^{-10}
x_1	0.79	10.95	3.01×10^{-8}
x_2	0.50	2.83	0.01
b_{x_1, x_2}^0	-36.33	-4.31	7.16×10^{-4}
x_1	0.96	15.16	4.41×10^{-10}
x_3	-0.05	-0.48	0.64

b_{x_i, x_j}^0 =término independiente de la regresión.

Paso 4: Con la inclusión de la variable " x_2 ", la variable predictora " x_1 " aún cumple con el criterio de $\alpha_E = 0.15$. Por lo tanto, " x_1 " y " x_2 " son variables significativas para la regresión y son incluidas en la ecuación de regresión.

Paso 5: Finalmente, se hace la regresión de " y " respecto de las variables predictoras " x_1 ", " x_2 " y " x_3 ".

Tabla 66. Coeficientes de regresión de pérdida de apilamiento respecto de las variables predictoras (velocidad, temperatura y concentración).

Predictor	Coficiente	Estadístico " t "	Valor de " P "
b_{x_1, x_2, x_3}^0	-33.77	-5.01	2.35×10^{-4}
x_1	0.80	11.15	5.01×10^{-8}
x_2	0.53	3.05	9.38×10^{-3}
x_3	-0.10	-1.18	0.26

b_{x_1, x_2}^0 = término independiente de la regresión.

Paso 6: De acuerdo a los valores de la tabla 65, " x_3 " no cumple con el criterio de $\alpha_E = 0.15$ y no puede ser considerada para que se integre en la ecuación de regresión. El método stepwise finaliza entonces con dos variables predictoras, " x_1 " y " x_2 ".