



COLEGIO DE POSTGRADUADOS

INSTITUCIÓN DE ENSEÑANZA E INVESTIGACIÓN
EN CIENCIAS AGRÍCOLAS

CAMPUS TABASCO

POSTGRADO EN PRODUCCIÓN AGROALIMENTARIA EN EL TRÓPICO

**“TIPIFICACIÓN Y CARACTERIZACIÓN DE CAFÉ COMERCIAL
MEDIANTE MÉTODOS INSTRUMENTALES Y QUIMIOMETRÍA”**

LORENA VÁZQUEZ HERNÁNDEZ

T E S I S

PRESENTADA COMO REQUISITO PARCIAL
PARA OBTENER EL GRADO DE:

MAESTRA EN CIENCIAS

H. CÁRDENAS, TABASCO

2011

La presente tesis titulada: **TIPIFICACIÓN Y CARACTERIZACIÓN DE CAFÉ COMERCIAL MEDIANTE MÉTODOS INSTRUMENTALES Y QUIMIOMETRÍA**, realizada por la alumna **LORENA VÁZQUEZ HERNÁNDEZ**, bajo la dirección del Consejo Particular indicado, ha sido aprobada por el mismo y aceptada como requisito parcial para obtener el grado de:

**MAESTRA EN CIENCIAS
PRODUCCION AGROALIMENTARIA EN EL TRÓPICO**

CONSEJO PARTICULAR

CONSEJERO:

Dr. Armando Guerrero Peña

ASESOR:

Dr. Adolfo Bucio Galindo

ASESOR:

Dr. Victorino Morales Ramos

H. Cárdenas, Tabasco, México, 29 de agosto de 2011

TIPIFICACIÓN Y CARACTERIZACIÓN DE CAFÉ COMERCIAL MEDIANTE MÉTODOS INSTRUMENTALES Y QUIMIOMETRÍA

Lorena Vázquez Hernández, M.C.

Colegio de Postgraduados, 2011

El control de calidad de los productos agroalimentarios está basado en su caracterización bioquímica, como una de las partes esenciales. Para el caso del café, la composición química está fuertemente relacionada con las condiciones agroecológicas (suelo, clima, altura, etc.) de cada una de las regiones productoras. El café es la segunda materia prima en importancia en el mercado nacional, por lo cual, la calidad es importante para este producto. Para asegurarnos de la calidad del café tanto para los consumidores como productores existe la necesidad de establecer métodos analíticos, rápidos, objetivos y sensitivos.

Para realizar la caracterización del café comercial existen métodos convencionales establecidos en la Norma Mexicana (NMX-F-013-SCFI-2000), para café tostado y molido. Los análisis fisicoquímicos que se realizaron fueron humedad (h%), ceniza (c%) y grasa (g%), en los cuales se realiza una serie de pasos para el tratamiento de la muestras, como lo es el proceso de pesado, dilución, homogenización, extracción y análisis. Lo que conlleva a la destrucción de la muestras, uso de reactivos y alto costo del análisis. Esto implica que la muestra se destruya, sea necesario un uso excesivo de reactivos y alto costo del análisis. Por tal motivo, se propone utilizar a la Espectroscopía de Infrarrojo Cercano (NIRS) para disminuir tiempo, reactivo, costo del análisis así como la mínima manipulación de las muestras; siendo una técnica no destructiva, permite conservar las muestras. La tecnología NIRS, como tecnología moderna se utiliza actualmente para medir compuestos bioquímicos en productos alimenticios y materias primas proporcionando información global de las muestras tanto física como químicamente contenida en sus huellas espectrales para ser tratadas con técnicas quimiométricas.

El objetivo de este trabajo fue evaluar la capacidad de la tecnología NIRS como método instrumental, y comprobar que genera la información útil y suficiente para la caracterización y tipificación de café comercial; además, para realizar un estudio de homogeneidad de cada uno

de los diferentes lotes de café comercial. Para ello se analizaron 120 muestras de café de diferente procedencia, por métodos convencionales de laboratorio (%humedad, %ceniza y %grasa) y NIRS para el desarrollo de huellas espectrales. Los resultados obtenidos en ambos casos fueron analizados quimiométricamente, mediante la aplicación de técnicas no supervisadas (Box-plot y PCA) y supervisada (SIMCA).

Los resultados confirman que la espectroscopia de infrarrojo cercano (NIRS) y la quimiometría pueden aplicarse al estudio de los lotes de café comercial para evaluar la caracterización, tipificación y pruebas de homogeneidad, con resultados similares a los obtenidos con los métodos convencionales.

Palabras claves: *Café, Quimiometría, NIRS, Homogeneidad, PCA, SIMCA.*

CLASSIFICATION AND CHARACTERIZATION OF COMMERCIAL COFFEE BY MEANS OF INSTRUMENTAL METHODS AND CHEMOMETRICS

Lorena Vázquez Hernández, M.C.

Colegio de Postgraduados, 2011

Quality control of food products is based on its biochemical characterization as one of the essential parts. In the case of coffee chemistry is strongly related to agro-ecological conditions (soil, climate, altitude, etc.) each one of the producing regions. Coffee is the second largest commodity in the domestic market; therefore, quality is important for this product. To ensure the quality of coffee for both consumers and producers there is a need for analytical methods, rapid, objective and sensitive.

For the characterization of commercial convention methods are established in the Mexican Norm (NMX-F-013-SCFI-2000), roasted and ground coffee. Physicochemical tests performed were moisture (h %), ash (c %) and fat (g %), in which a series of steps performed for the treatment of the samples, as is the process of weighing, dilution, homogenization, extraction and analysis. This leads to the destruction of the samples, use of reagents and high cost of analysis. This implies that the sample is destroyed, needed excessive use of expensive reagents and analysis. Therefore, we propose using the Near Infrared Spectroscopy (NIRS) to reduce time, reactive, cost analysis and the minimal handling of samples, being a non-destructive technique, preserves the samples. NIRS technology, as modern technology is currently used to measure biochemical compounds in food products and raw materials providing global information of the samples contained both physically and chemically in their spectral signatures to be treated with chemometric techniques.

The aim of this study was to evaluate the ability of NIRS technology as instrumental method, and check that generates the information useful and sufficient for the characterization and classification of commercial coffee, also for a study of consistency of each of the different batches commercial coffee. We analyze 120 samples of coffee from different sources by conventional laboratory methods (%moisture, %ash and %fat) and for the development of

NIRS spectral signatures. The results were analyzed both chemometric by applying unsupervised techniques (Box-plot and PCA) and supervised (SIMCA).

The results confirm that near-infrared spectroscopy (NIRS) and chemometrics can be applied to the study of commercial batches of coffee to assess the characterization, classification and homogeneity tests, with results similar to those obtained with conventional methods.

Keywords: Coffee, Chemometrics, NIRS, Homogeneity, PCA, SIMCA.

DEDICATORIA

Dedico mi tesis a:

A Dios por darme las fuerzas necesarias en los momentos en que más las necesité, por la oportunidad de vivir y sus bendiciones...gracias a ello he logrado cumplir mis metas.

A mi hija Alejandra, por ser quien me impulsa a seguir adelante y la personita que llena mi ser, y porque es ella quien se lleva los créditos de mi esfuerzo...y a quien amo con todas mis fuerzas.

A mis padres, Ramón Vázquez Escudero y Lucía Hernández Hernández por todo lo que me han dado en esta vida, especialmente por sus sabios consejos, por los principios y valores, por estar a mi lado en todo momento, por su apoyo incondicional, por sus muestras de amor y comprensión.

A mis hermanos, hermanas, cuñados, cuñadas, sobrinos y sobrinas por compartir conmigo buenos y malos momentos porque de una u otra manera fueron parte de mi formación profesional.

A todos ustedes les agradezco el tiempo que me han dedicado y sobre todo les tengo presentes porque aun cuando el recorrido fuese difícil estuvieron a mi lado.

“Encomiéndate a Dios de todo corazón, que muchas veces suele llover sus misericordias en el tiempo que están más secas las esperanzas”.

Miguel de Cervantes Saavedra

AGRADECIMIENTOS

Al Consejo Nacional de Ciencia y Tecnología (CONACYT), por el financiamiento otorgado para la realización de mis estudios de maestría.

Al Colegio de Postgraduados Campus Tabasco, por permitirme realizar mis estudios de Postgrado dentro del programa de Producción Agroalimentaria en el Trópico.

A mi director de tesis y sobre todo mi gran amigo el Dr. Armando Guerrero Peña, por compartir sus conocimientos, por el tiempo dedicado al proyecto y por la amistad que me brindo durante la estancia en el Colegio de Postgraduados-Campus Tabasco.

A mi asesor el Dr. Adolfo Bucio Galindo por haberme brindado la oportunidad de trabajar con él, por sus consejos y aportaciones al trabajo de investigación. Por ser una persona con iniciativas a la investigación y por permitir compartir sus experiencias y conocimientos.

A mi asesor el Dr. Victorino Morales Ramos, por la contribución al buen desarrollo de esta investigación, por la oportunidad de contar con su ayuda y experiencias en mi consejo particular.

Al personal de Laboratorio de Suelos, Plantas y Aguas (LASPA) por el apoyo en la realización de los análisis pertinentes de la investigación.

A la línea 12 de “Agregación de Valor” del Colegio de Postgraduados por los apoyos brindados para esta investigación.

A las empresas café de Huimanguillo, café de Córdoba, Veracruz, café de Ixtepec, Oaxaca y café Orgánico Tzisco, Chiapas por la aportación de las muestras de café para la realización de la presente investigación.

A todos mis profesores: Dr. José Jesús Obrador, Dr. Mepivoseth, Dr. Adolfo Bucio, Dr. Víctor Córdova, Dr. Osorio, Dr. Ángel Galmiche, Dr. Ángel Sol, Dr. Armando Guerrero, Dr. Fredy, Dra. Luz del Carmen, Dra. Consuelo, Dra. Ma. del Carmen y la Prof. Juanita Pereyra gracias por compartir sus conocimientos y experiencias, haciéndolos parte de mi formación profesional.

A Héctor Javier Megía Vera, gracias por caminar a mi lado durante este tiempo en la maestría y compartir los buenos y malos momentos juntos, porque de una u otra manera fuiste inspiración para mí.

A mi incondicional compañero Isai Quevedo Damián, que me dio la oportunidad de conocer lo que es una amistad y lo que es saber contar con alguien en los momentos más difíciles.

A mis compañeros de generación PROPAT 2009: Héctor Javier, Isai, Evelyn, Erika, Beatriz A., Carlos A., Mateo, Alberto Naranjo, Germán, Cristóbal, Manuel, Rodrigo, Apolinar, Prisciliano, Ivanna de L., Francis de J., Héctor Sánchez, Josué y Fernando Pacheco por todos los momentos y conocimientos compartidos, por las locuras que hicimos y por las enseñanzas que dejaron cada uno de ellos en mí.

Durante mi estancia en el Colegio de Postgraduados-Campus Tabasco, conocí y compartí momentos con muchas personas que me apoyaron, no tan solo en lo académico y en lo científico, sino también en lo personal. A todas ellas, y sin dejar a nadie en el olvido, quiero agradecerles su tiempo, sus palabras y su apoyo. Gracias.

CONTENIDO

INDICE DE FIGURAS	xiii
INDICE DE CUADROS	xiv
1. INTRODUCCIÓN GENERAL	1
2. PLANTEAMIENTO DEL PROBLEMA	3
3. OBJETIVOS E HIPÓTESIS	4
3.1. Objetivo General	4
3.1.1. Objetivos específicos	4
3.2. Hipótesis General	5
3.2.1. Hipótesis específicas	5
4. REVISIÓN DE LITERATURA	6
4.1. Historia del Café	6
4.2. Cultivo y Producción de Café en México	6
4.3. El Desarrollo de la Cafecultura	8
4.4. Descripción Botánica del Café	10
4.4.1. Tipos de café producidos y calidades mexicanas	11
4.4.2. Composición química del café	13
4.4.3. Componentes aromáticos	14
4.5. Métodos de Tipificación	15
4.5.1. Análisis físico-químico	15
4.6. Métodos Instrumentales	16
4.6.1. Espectroscopia en el infrarrojo cercano (NIRS)	16
4.6.2. Aplicaciones de la tecnología NIRS	16
4.7. Teoría de la Espectroscopia en el Infrarrojo Cercano (NIRS)	17
4.8. Quimiometría	19
4.8.1. Reconocimiento de pautas no supervisadas	21
4.8.2. Reconocimiento de pautas supervisadas	28

5. LITERATURA CITADA	33
CAPITULO I: CARACTERIZACIÓN Y TIPIFICACIÓN DE CAFÉ COMERCIAL MEDIANTE LA TECNOLOGÍA NIRS	41
RESUMEN.....	42
ABSTRACT.....	43
1.1. INTRODUCCIÓN	44
1.2. MATERIALES Y METODOS	45
1.2.1. Localización y selección de las muestras de café comercial	46
1.2.2. Relación de los métodos físico – químicos	47
1.2.3. Software para la interpretación quimiométrica de los métodos físico-químicos y NIRS....	47
1.2.4. Espectrofotómetro de infrarrojo cercano (NIRS)	48
1.3. RESULTADOS Y DISCUSIÓN	49
1.3.1. Espectros NIRS de los cuatro lotes de café.....	49
1.3.2. Análisis de componentes principales (PCA) de los análisis Físicoquímicos	50
1.3.3. Análisis de componentes principales (PCA) en NIRS	52
1.3.4. Modelado independiente de clase (SIMCA) para muestras de café.....	54
1.3.5. Espectros NIRS de las variedades de café de Córdoba.....	55
1.3.6. Análisis de componentes principales (PCA) de los análisis Físicoquímicos	57
1.3.7. Modelado independiente de clase SIMCA para muestras de café de Córdoba, Veracruz. ..	59
1.4. CONCLUSIONES	61
1.5. LITERATURA CITADA	62
CAPITULO II: HOMOGENEIDAD DE MUESTRAS DE CAFÉ TOSTADO Y MOLIDO MEDIANTE TECNOLOGIA NIRS	65
RESUMEN.....	66
ABSTRACT.....	67
2.1. INTRODUCCIÓN	68
2.2. MATERIALES Y METODOS	69

2.2.1. Localización de las muestras de café comercial.....	70
2.2.2. Relación de los métodos físico – químicos	70
2.2.3. Interpretación quimiométrica de los métodos físico-químicos	71
2.2.4. Espectrofotómetro de infrarrojo cercano (NIRS).....	71
2.2.5. Interpretación quimiométrica de los espectros NIRS.....	72
2.3. RESULTADOS Y DISCUSIÓN	72
2.3.1. Descriptores estadísticos de las propiedades fisicoquímicas de café tostado y molido	72
2.3.2. Prueba de homogeneidad mediante análisis físicos y químicos.....	73
2.3.3. Prueba de homogeneidad mediante NIRS.....	76
2.4. CONCLUSIONES	80
2.5. LITERATURA CITADA	81
CAPITULO III: CONCLUSIONES GENERALES DE LA INVESTIGACIÓN.....	85
3.1. CONCLUSIONES GENERALES.....	86

INDICE DE FIGURAS

Figura 1. Distribución por países de la producción mundial del café (Pérez-Grovas, 1998).....	9
Figura 2. Porcentaje de producción de café por estado (Cámara de Diputados, 2001).....	10
Figura 3. Distancia Euclídea y de Manhattan para las muestras k y l.....	22
Figura 4. Unión de un grupo con una muestra aislada de acuerdo con los métodos del centroide (a), de la unión simple (b) y de la unión completa (c).....	24
Figura 5. Dendograma.....	25
Figura 6. Análisis de componentes principales (PCA).....	26
Figura 7. Representación de las puntuaciones de dos clases de objetos en los dos primeros componentes principales.....	28
Figura 8. Modelos de dos clases de objetos A y B en el espacio de sus componentes principales, regiones de confianza y predicción de dos muestras.....	31
Figura 9. Diagrama de Coomans.....	32
Figura 10. Espectroscopio de infrarrojo cercano, NIRs Versatile Agri Analyzer TM 6500.....	48
Figura 11. Espectros generados por cada una de las muestras de café comercial.....	49
Figura 12. Componentes principales para los análisis fisicoquímicos.....	50
Figura 13. Representación en el plano de los componentes principales (<i>biplot</i>) de las variables humedad, ceniza, grasa y las muestras de café de diferente origen geográfico.....	52
Figura 14. Gráfico de los dos primeros componentes principales para NIRS.....	53
Figura 15. Diagrama de Coomans construido con el modelado de clases SIMCA, en muestras de café comercial.....	54
Figura 16. Espectro NIRS observado en el café de Córdoba, Veracruz.....	56
Figura 17. Representación de las huellas espectrales (NIRS) en los dos primeros componentes principales.....	57
Figura 18. Componentes principales para los análisis fisicoquímicos en muestras de café de Córdoba, Veracruz.....	59
Figura 19. Diagrama de Coomans construido con el modelado de clases SIMCA, en muestras de café comercial de Córdoba, Veracruz.....	60
Figura 20. Espectroscopio de infrarrojo cercano, NIRs Versatile Agri Analyzer TM 6500.....	71
Figura 21. Contenido de humedad en muestras de café tostado y molido.....	74
Figura 22. Contenido de ceniza en muestras de café tostado y molido.....	75
Figura 23. Contenido de grasa en muestras de café tostado y molido.....	75

Figura 24. Huellas espectrales de café tostado y molido de: a) Chiapas, b) Tabasco, c) Veracruz y d) Oaxaca.....	77
Figura 25. Diagrama de cajas para CP1 en muestras de café tostado y molido	79

INDICE DE CUADROS

Cuadro 1. Composición química del café verde y tostado (% en base seca).	13
Cuadro 2. Procedencia de muestras de café comercial.	46
Cuadro 3. Métodos de la Norma Mexicana NMX-F-013-SCFI-2000 para café tostado y molido.	47
Cuadro 4. Análisis de componentes principales para los análisis fisicoquímicos.....	50
Cuadro 5. Valores de los <i>Loadings</i> (cargas) en los tres primeros componentes principales.....	51
Cuadro 6. Análisis de componentes principales.....	53
Cuadro 7. Análisis de componentes principales de las huellas espectrales NIR.	57
Cuadro 8. Análisis de componentes principales de los análisis fisicoquímicos de las muestras de café de Córdoba, Veracruz.....	58
Cuadro 9. Procedencia de muestras de café comercial.	70
Cuadro 10. Métodos de la Norma Mexicana NMX-F-013-SCFI-2000 para café tostado y molido.	70
Cuadro 11. Media y desviación estándar de datos obtenidos en el análisis físico-químico de las muestras de café.	72
Cuadro 12. Media y desviación estándar del CP1 obtenido de las huellas espectrales NIR en las diferentes muestras de café.	78
Cuadro 13. Prueba de ANOVA de un factor para Análisis Fisicoquímicos y NIRS.	78

1. INTRODUCCIÓN GENERAL

La evaluación de la calidad de los alimentos constituye una necesidad cada vez más apremiante para las industrias de producción y transformación de los mismos, cuyo objetivo es asegurar la conformación de estándares y garantizar su inocuidad; ya que una de las preocupaciones del consumidor es adquirir productos sanos y que estos tengan los nutrientes necesarios. Para ello, es ineludible mantener un adecuado control de calidad en un producto alimenticio; con base en las normas oficiales; las cuales, generalmente, utilizan métodos convencionales, que incluyen etapas de preparación de la muestra previo al análisis (secado, molienda, tamizado, etc.), lo que hace que estos métodos tengan por característica: a) que sean tediosos, b) demandan tiempo, c) mano de obra, d) con un costo elevado, y e) el error analítico es alto debido a la excesiva manipulación de la muestra. En la actualidad, se tienen otras alternativas metodológicas para evaluar los alimentos con la característica de ser más rápidos, precisos y con menor error analítico que los métodos convencionales (Hernández, 2008).

El desarrollar investigaciones con metodologías que permitan la reducción de tiempo en el tratamiento de las muestras, es uno de los objetivos actuales de la química analítica. Cuya finalidad no es qué contiene o no la muestra, sino determinar su identidad. En ocasiones no es necesario disponer de información individual sobre los compuestos de una muestra. La información contenida en los datos se extrae mediante la aplicación de técnicas quimiométricas. En este caso la manipulación de la muestra se reduce considerablemente, con lo que disminuyen los errores asociados a esta etapa y el tiempo total de análisis es muy corto. Ambas características ponen de manifiesto que esta estrategia de procesado matemático de la información sobre el conjunto de compuestos es muy atractiva (Guerrero, 2006).

Una alternativa que nos proporciona la química analítica son los métodos instrumentales, los cuales presentan un gran interés en procesos industriales y en la realización de investigaciones por su velocidad en los análisis; ejemplos de estas técnicas son el generador de espacio de cabeza acoplado a un cromatógrafo de gases-espectrómetro de masas (HS-GC-MS), espectroscopia en el infrarrojo cercano (NIRS), e infrarrojo medio (MIR), espectroscopia ultravioleta-visible y pirolisis acoplada a espectrometría de masas, entre otros.

El objetivo de esta investigación es evaluar si la información generada con la tecnología NIRS permite caracterizar y tipificar el café comercial con resultados similares a los obtenidos con los métodos convencionales.

2. PLANTEAMIENTO DEL PROBLEMA

Los procedimientos analíticos convencionales utilizados para resolver los problemas analíticos, por ejemplo, la tipificación del café, implican obtener la información química y física de las muestras en estudio. Estos métodos se caracterizan por integrar etapas de preparación de las muestras (secado, molienda, extracción, etc.), previo a la medición analítica. Esta etapa es la parte más costosa del proceso; la más larga y la que aporte mayor probabilidad de error.

El obtener la información individual de la composición química y física de las muestras, no siempre es necesaria. Cuando se requiere determinar la identidad de las muestras y no solo saber lo que contienen, es posible utilizar métodos instrumentales y quimiometría, en lo que se conoce como métodos de *screening* o pruebas de diagnóstico rápido. En esta opción se obtiene señales instrumentales tipo perfil, y la interpretación de los espectros es realizada mediante técnicas quimiométricas. Todo esto constituye en la actualidad la tendencia de la Química Analítica, resolver el problema planteado con el menor esfuerzo posible.

Por tanto, en esta investigación se pretende establecer una metodología instrumental para resolver el problema planteado, tipificación de café comercial tostado y molido, con la ventaja de que sea una propuesta más rápida, con menor error analítico y menor consumo de reactivos, entre otros; cuando se compare con los métodos convencionales.

3. OBJETIVOS E HIPÓTESIS

3.1. Objetivo General

Evaluar la capacidad de los métodos instrumentales para generar información útil y suficiente para la caracterización y tipificación del café comercial.

3.1.1. Objetivos específicos

Realizar la caracterización y tipificación del café comercial mediante métodos físicos y químicos convencionales y la quimiometría.

Realizar la tipificación de café comercial mediante métodos instrumentales y quimiometría.

Realizar un estudio de homogeneidad del café comercial mediante métodos instrumentales y quimiometría.

3.2. Hipótesis General

Los métodos instrumentales producen información suficiente para caracterizar y tipificar el café comercial, con resultados similares a los obtenidos con los métodos convencionales.

3.2.1. Hipótesis específicas

Los métodos químicos convencionales y la quimiometría permiten caracterizar y tipificar el café comercial.

Con la tecnología NIRS, y la quimiometría es posible realizar la tipificación de café comercial.

Con los métodos instrumentales y la quimiometría es posible realizar estudios de homogeneidad en café comercial.

4. REVISIÓN DE LITERATURA

4.1. Historia del Café

De acuerdo con González (1998), el café fue introducido en Arabia procedente de Abisinia antes del siglo X, y se cultivó por primera vez en el siglo XIV. Se conoce con seguridad que en los siglos XV y XVI se producía intensivamente en el distrito Árabe de Yemen, el producto era tan lucrativo que los árabes prohibieron la exportación de sus semillas con el objeto de monopolizarlo; a mediados del siglo XVI el café se consumía en las ciudades de Medina y Meca, desde donde se extendió a Persia, Egipto, Siria, y Turquía.

El café es uno de los cultivos agrícolas de mayor importancia en el mundo, se produce en más de 50 países que se localizan principalmente entre los trópicos de Cáncer y de Capricornio, destacando por su volumen producido Brasil, Colombia, Vietnam, Indonesia y México. El café, como uno de los principales productos agrícolas de México, representa una de las bases económicas agrícolas para el país (ASERCA, 2002).

Fue introducido en México en el año de 1796 a la región de Córdoba, Veracruz, por el señor Juan Antonio Gómez (Regalado, 1996) y en 1823 se importaron algunas semillas al estado de Michoacán. Para el año de 1846 se introdujeron plantas de café al estado de Chiapas con las que se expandió el cultivo hacia otros estados. En 1803, México realizó la primera exportación de café, pero fue hasta 1882, que México pasa a formar parte de los países exportadores de café, situación que permanece hasta nuestros días (Nolasco, 1985).

4.2. Cultivo y Producción de Café en México

El cultivo del café llegó a México por tres rutas comerciales. La primera en 1796 de Cuba a la región de Córdoba, Veracruz; la segunda fue en 1823 desde la región de Moka, Arabia a Morelia, Michoacán y la tercera, en 1847, desde Guatemala a Tapachula, Chiapas. Desde estas regiones, a lo largo del siglo XIX se extiende su cultivo a diversas regiones de México, principalmente de los estados de Veracruz, Chiapas, Puebla, Tabasco y Michoacán (FAO-SAGARPA, 2006).

México produce cafés de excelentes calidades, ya que su topografía, altura, diversidad de climas y suelos le permiten cultivar y producir variedades clasificadas dentro de las mejores en el mundo. Como productor de café, México ocupa el quinto lugar a nivel mundial, después de Brasil, Colombia, Indonesia y Vietnam. Actualmente, México es el primer productor mundial de café orgánico (FAO-SAGARPA, 2006).

Las especies de arbustos de cafeto que se cultivan en el país son dos: la arábica, árabe o arabica y la robusta o canephora. Aproximadamente 98% de los cafetos son variedades arábicas como Bourbon, Caturra, Maragogipe o (Márago), Mundo Novo, Garnica y Typica; siendo esta la que predominaba en México; sin embargo, actualmente se le está reemplazando por variedades de porte bajo y mayor producción como la Catimor y Catuai. Cada variedad posee diferencias de calidad, volumen, rendimiento, aroma, acidez, resistencia a plagas y enfermedades, entre otros (Cámara de Diputados, 2001).

El cultivo y la producción de café se lleva a cabo en doce estados de la República Mexicana, sobre una superficie de 664, 794 ha, que en su conjunto abarcan 398 municipios y están distribuidos en cuatro grandes regiones: Las vertientes del Golfo de México, del Océano Pacífico, la zona Centro-Norte y la del Soconusco en Chiapas, en el sureste mexicano. Cada una de estas regiones muestran características propias y en su mayoría tienen condiciones adecuadas para la obtención de cafés de calidad (Díaz y Castro, 2005).

Inicialmente el producto se utilizaba para masticar, preparar bebidas y licores de los frutos maduros, mientras que la práctica de tostar y preparar bebidas de los granos molidos fue descubierta en Persia, poco después de que el uso del café se había extendido por los países vecinos de Arabia. Rápidamente el café se hizo popular, y pronto aparecieron “casas de café” en muchas ciudades del cercano oriente, donde la gente llegaba a conversar y escuchar relatos de cuentos o música. Durante la primera mitad del siglo XVII, el café se conoció en Europa gracias a los viajeros y comerciantes que regresaban del cercano oriente. Para la segunda mitad de ese siglo, la fama del café como bebida agradable y estimulante se extendió a través de Italia, Francia, e Inglaterra, después a Holanda, Alemania, Austria, Suecia y Dinamarca. En

poco tiempo también se establecieron cientos de “casas de café” en Europa, que se convirtieron en los centros de vida social de la época (<http://www.cafesdemexico.com>).

Desde hace trescientos años, el café, es una de las mercancías más importantes del comercio mundial. Hoy en día existen plantaciones en más de cincuenta países, situándose la mayoría entre el trópico de cáncer y el trópico de capricornio. El consumo se extiende por todo el mundo, siendo Norteamérica y Europa las principales regiones consumidoras (González, 1998).

4.3. El Desarrollo de la Cafeticultura

El impulso de la cafeticultura en México se inicia en la segunda mitad del siglo XX, dentro de un contexto de apoyo a la concentración de tierras en manos privadas. La cafeticultura mexicana sufrió pocas modificaciones hasta la época cardenista en que los grandes latifundios de la zona central del país son repartidos entre los campesinos (Pérez-Grovas, 1998).

En esta perspectiva, México, fue considerado tradicionalmente como un productor de café de menor calidad que otros países; sin embargo, las regiones productoras de México, tiene no solo las cualidades idóneas, sino una vocación natural para la producción de cafés de especialidad sobre todo si se toma en cuenta que la producción es predominante de café de sombra y orgánico, por el resguardo que brindan las selvas, ecosistemas donde se cultiva la mayor cantidad de café en el país. El 92% de los productores de café en México poseen superficies menores a 5 hectáreas. Del total de hectáreas cultivadas de café en México, el 91% son superficies con cafetos e producción, el 4% lo ocupan los cafetos pre-productivos, en tanto que el 5% restante corresponde a caminos e infraestructura agroindustrial (ASERCA, 2002).

Desde el inicio de su cultivo en México, en el siglo XIX, el café juega un papel importante en la generación de empleos para un número considerable de mexicanos, además de convertirse en uno de los principales generadores de ingresos y divisas (Fundación Produce y Tecnológico de Monterrey, 2003). Es un cultivo eminentemente orientado hacia el mercado internacional,

su producción se concentra en cerca de 57 países tropicales; y su principal consumo en los países de clima templado. Más del 75% de la producción mundial se comercializa en el mercado internacional, ubicándose nuestro país como el cuarto productor y exportador, destinándose a Estados Unidos alrededor del 80% del total de nuestras exportaciones (Cih *et al.*, 2001).

Tradicionalmente la producción se ha establecido en los países en desarrollo, que estuvieron bajo regímenes coloniales; y por otro lado la industrialización y la mayor parte del consumo se concentra en los países desarrollados. Cada año se consumen en el mundo 100 millones de sacos de 60 kilogramos de café verde, los cuales son producidos en los países de América Latina, Asia y África. Lo anterior, se puede observar en la Figura 1 (Pérez-Grovas, 1998).

Según la Organización Internacional del Café, la cosecha mundial para el ciclo 2002-2003 fue de 117.3 millones de sacos de 60 kilogramos, cada uno, lo que ocasionó una sobreproducción de café en los siguientes países (millones de sacos): Brasil 47.26, Colombia 11.25, Vietnam 8.90, Indonesia 5.83, India 4.66 y México 4.00; de café árabe y robusta de diferentes calidades. Lo anterior, se complica por el escaso consumo de los países productores, con excepción de Brasil, que ha provocado un desequilibrio en los últimos años entre producción y consumo, y que fue una de las principales causas de la grave crisis del precio del café (Regalado, 2006).

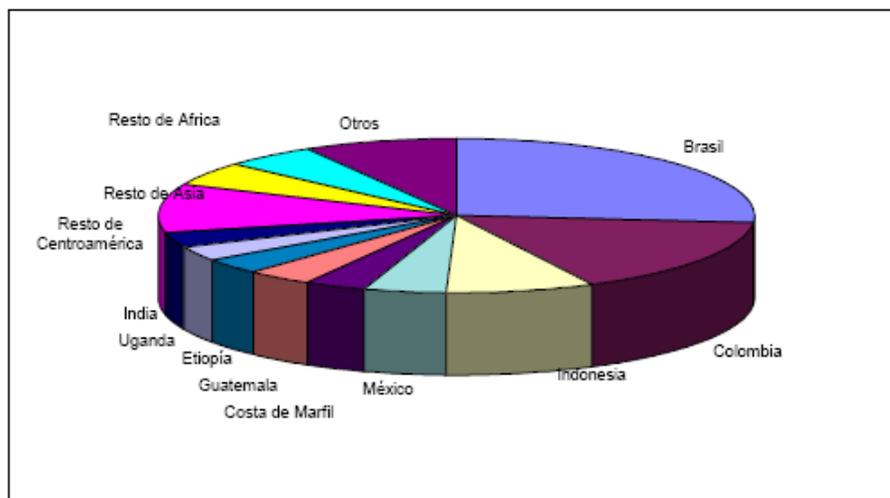


Figura 1. Distribución por países de la producción mundial del café (Pérez-Grovas, 1998).

En la República Mexicana la producción de café se concentra en doce estados. El 92% corresponde a Chiapas, Veracruz, Oaxaca y Puebla, el resto a otras ocho entidades. Lo anterior lo podemos observar en la Figura 2. En México se producen 11 diferentes calidades de café, las mejores son: estrictamente altura, altura, extra prima lavado y prima lavado (Regalado, 2006 y Cámara de Diputados, 2001).

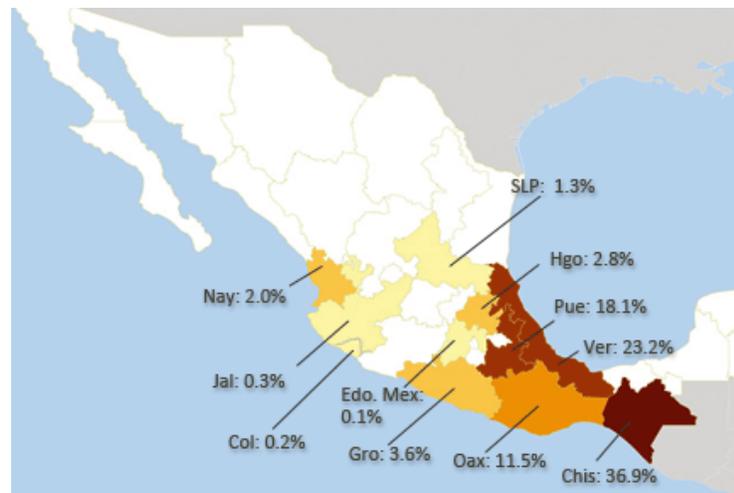


Figura 2. Porcentaje de producción de café por estado (Cámara de Diputados, 2001).

El café que se produce en México es de la especie *Coffea arabica* L., que constituye el 97% de la producción nacional, representada por las variedades Typica (criollo, nacional o arábica), Bourbon, Caturra, Mundo Novo, Garnica, Catuai, Pluma Hidalgo y Maragogype; el 3% de la producción corresponde a la especie *Coffea canephora* Pierre ex Frohener, conocida como robusta, cultivada en zonas bajas de Veracruz (principalmente en los municipios de Tezonapa y Tepatlaxco), Chiapas (sobresale el municipio de Cacahoatán) y Oaxaca (regiones de Tuxtepec y Valle Nacional) (Fundación Produce y Tecnológico de Monterrey, 2003).

4.4. Descripción Botánica del Café

El café pertenece a la familia *Rubiaceae*; la cual, tiene más de 60 especies; el género *Coffea* se puede dividir en dos grupos de acuerdo con el número cromosómico. Estos dos grupos de especies difieren, además del número cromosómico, en aspectos como la forma de polinización cruzada, para la mayoría de las especies diploides, debido a la

autocompatibilidad. Esta característica origina heterogeneidad entre las especies diploides, mientras que las variedades de *C. arabica* son altamente homogéneas, debido a su casi completa autopolinización. La especie *C. arabica* tiene escasa variabilidad genética, posee buen nivel de producción y características o atributos de alta calidad en la bebida. La única variedad producida de *C. canephora* se denomina café robusta, que es de baja calidad comparada con las variedades árabes. A las pertenecientes a *C. arabica* se les conoce como café arábigo. La descripción general de *C. arabica* se llevó a cabo por Carlos Linneo en 1753. La variedad *typica* fue probablemente la que se tomó como base para la descripción de la especie en tallo, ramas, hojas, flores, fruto y semilla (Regalado, 2006).

Es importante resaltar los factores genéticos que influyen en la calidad del café: tamaño, forma, color y composición química son características relacionadas con su genotipo. El tamaño varía según la variedad. Este puede ser grande, mediano o pequeño; a mayor tamaño más calidad de la bebida en la taza. El color de los granos es una variedad del genotipo, la especie *C. arábica* que produce un grano color verde azulado que nos indica café de calidad (estrictamente altura). En cambio, el *C. canephora* produce una fruta marrón o sea de baja calidad. Existen en especies *C. arábicas* colores verde oscuros uniforme de calidad (altura), verde olivo uniforme (prima lavado), los colores claros, oscuros y negros son de muy baja calidad y no se recomiendan.

La composición química de la especie *C. arábica* tiene las características o atributos siguientes: aroma fuerte, bastante acidez, cuerpo liviano, poco amargo y poca cafeína. Difiere de la especie *C. canephora*, que es de baja calidad ya que tiene aroma débil, poca acidez, cuerpo fuerte, muy amargo y mucha cafeína. El sabor está relacionado con el genotipo, la especie *arábica* es más aromática, con bastante acidez y cuerpo ligero, lo que le da un sabor muy agradable de café fino (Costell and Duran, 1981).

4.4.1. Tipos de café producidos y calidades mexicanas

Los tipos de café producidos son: lavados, naturales y semilavados. Las Normas Mexicanas de acuerdo con el Consejo Mexicano del Café A.C. (2002), establecen que los tipos de café

verdes mexicanos se clasifican, para su comercialización, en once calidades, las cuales se mencionan a continuación:

Estrictamente altura. Altura mayor a 900 msnm. *Bouquet* (aroma) rico; acidez completa; cuerpo completo; sabor completo o redondo.

Altura. Altura mayor a 900 msnm. *Bouquet* de completo a rico; acidez de pronunciada a completa; cuerpo de fuerte a completo; sabor de pronunciado a completo o redondo.

Extra prima lavado. Altura de 700 a 900 msnm. *Bouquet* completo; acidez de pronunciada a completa; cuerpo de bueno a fuerte o hasta completo; sabor de medio a pronunciado.

Prima lavado. Altura de 600 a 900 msnm. *Bouquet* de notorio a completo, sin llegar a rico; acidez de media a pronunciada, sin llegar a completa; cuerpo regular, bueno o hasta fuerte, sin llegar a ser completo; sabor medio.

Buen lavado. Altura de 600 a 900 msnm. *Bouquet* de débil o flojo a notorio, sin llegar a completo; acidez de débil a ligera; cuerpo de ligero a regular; sabor de débil a medio.

Lavado. Sin referencia sobre altitud. *Bouquet* de dañado hasta rico; acidez de ausente a completa; cuerpo de ligero a completo; sabor de impuro a completo.

Desmanche. Sin referencia sobre altitud. *Bouquet* de dañado hasta rico; acidez de ausente a completa; cuerpo de ligero a completo; sabor de impuro a completo.

Natural tipo A. Sin referencia sobre altitud. *Bouquet* de dañado hasta rico; sin referencia sobre acidez; cuerpo de ligero a completo; sabor de impuro a completo.

Natural tipo B. Sin referencia sobre altitud. *Bouquet* de dañado hasta rico; sin referencia sobre acidez; cuerpo de ligero a completo; sabor de impuro a completo.

Robusta lavado. Sin referencia sobre altitud. *Bouquet* de dañado hasta rico; acidez de ausente a completa; cuerpo de ligero a completo; sabor de impuro a completo.

Robusta natural. Sin referencia sobre altitud. *Bouquet* de dañado hasta rico; acidez de ausente a completa; cuerpo de ligero a completo; sabor de impuro a completo (Consejo Mexicano del Café A.C, 2002).

Muchas investigaciones han encontrado que la calidad del café está determinada por diversos factores; entre ellos, los ambientales, genéticos, agronómicos y agroindustriales. Cada uno de los factores puede demeritar, modificar o aportar cualidades distintas (Tecnología Integral del

Café S. A., 1991; Santoyo *et al.*, 1996; Salazar *et al.*, 2000; Wintgens, 2004). En México se encuentran regiones que tienen potencial para producir café de calidad suficiente para ingresar a los mercados de especialidad (predominio de cafés árabes, temperaturas “frescas” y cultivo bajo sombra), y así obtener mejores precios en el mercado mundial; sin embargo, se necesita reorganizar la producción bajo este nuevo enfoque y realizar un trabajo de promoción bien estructurado.

4.4.2. Composición química del café

El café tostado es el producto obtenido del café, el cual ha sido sometido a temperatura superior a 150 °C y presenta una pérdida de peso respecto al grano de café verde utilizado del 10 al 24 % en relación masa/masa (m/m); para molerlo se somete posteriormente a una reducción de tamaño de partícula de 6 mm (NMX-F-013-SCFI-2000).

La composición química de los granos de café es extremadamente compleja, se han identificado más de 300 constituyentes; a causa de esta complejidad se desconoce en gran parte la base química del sabor y el aroma del café. En el Cuadro 1 es indicada la composición química promedio del café verde y del café tostado (Fox y Cameron, 1992).

Cuadro 1. Composición química del café verde y tostado (% en base seca).

	Granos verdes	Granos tostados
Proteínas	13	11
Azúcares	10	1
Almidón y dextrinas	10	12
Polisacáridos complejos	40	46
Aceite	13	15
Minerales*	4	5
Ácido clorogénico	7	5
Trigonelina	1	1
Fenoles	0	2
Cafeína (en <i>Coffea arabica</i>)	1	1,3

(Fox y Cameron, 1992).

Los lípidos en los granos de café verde se encuentran localizados en el aceite de café, presente en el endospermo y en pequeñas cantidades en la llamada cera del café, localizada en otras

partes del grano (Folstar, 1985). Durante el tostado, el aceite sale a la superficie de los granos, lo que da el aspecto característico del tostado oscuro, e impide la salida de parte de los componentes volátiles del aroma (Clarke, 1987).

4.4.3. Componentes aromáticos

La mayoría de los componentes volátiles del café derivan de componentes no volátiles presentes en el grano verde, los cuales reaccionan durante el tostado formando mezclas muy complejas. El sabor y aroma característicos del café tostado es el resultado de un conjunto de reacciones, por un lado, entre la que se destaca la pirolisis, y por otro lado, de una serie de interacciones entre componentes muy diversos del grano: azúcares, aminoácidos, ácidos orgánicos y componentes fenólicos. La composición final de volátiles en un determinado café depende de numerosos factores, tales como la especie y variedad del grano, las características del suelo y factores climáticos durante su cultivo, las condiciones de almacenamiento de los granos verdes, el método utilizado para el tostado, así como el tiempo y la temperatura de tostado. Todos estos factores influyen en el aroma final del producto por afectar directamente sobre la cantidad y calidad de precursores no volátiles presentes en el grano verde y su posterior evolución durante el tostado. En café verdes se pueden encontrar varios volátiles, tales como piridinas, quinoleínas, pirazinas, pirroles y poliaminas. Otros componentes (furanos, alcoholes, carbonilos, ésteres, fenoles y tioles), se encuentran en determinadas variedades de café. La presencia de metoxipirazinas, es un compuesto que puede ser el responsable del aroma característico del café verde (Poisson, 1977).

Más de 800 compuestos volátiles se han aislado en el aroma del café tostado, los principales tipos de componentes identificados son: aldehídos, azufres, cetonas, fenoles, furanos, pirazinas, piridinas, pirroles y oxazoles. En el café tostado se han caracterizado varios furanos tiosustituídos. Posiblemente, el compuesto 2-furilmetanotiol tenga mucho peso en café tostado. Su umbral en agua es de 0.01 ppb y a concentraciones de 0.1-5 ppb tiene el aroma del café tostado (Tressl y Silwar, 1981). Los ésteres, aldehídos y cetonas de los grupos tiofenos aportan aromas dulces (Maga, 1975).

Los aldehídos probablemente sean los componentes volátiles más abundantes en el café. El metanal y etanal tienen aromas picantes, pero a medida que el número de átomos de carbono de la cadena aumenta, los aromas tienden a ser frutados y florales. Las cetonas cíclicas derivadas de la caramelización de azúcares, confieren al café aromas dulces, frutados y de azúcar quemada (Hodge, 1967). Los fenoles son característicos de café muy tostados, poseen aromas típicos de ahumado, quemado, picante y amargo, aportando una cierta astringencia a la bebida (Tressl *et al.*, 1978).

Otro compuesto es la trigonelina que está presente en los granos crudos de café a niveles de 1%. Contrariamente a la cafeína, que es termoestable, la trigonelina se descompone fácilmente a las temperaturas de tostado para dar una serie de piridinas y pirroles (Block, 1985). Las pirazinas son abundantes en el aroma del café tostado (Shimoda and Shibamoto, 1990); poseen aromas entre amargo y dulce, pero en sus alquil-derivados se aprecian aromas a nueces, tostado y picante (Maga, 1982). Los pirroles aportan sensaciones de dulce, amargo y aromas tostados (Maga, 1981).

4.5. Métodos de Tipificación

El Diccionario de Real Academia Española (2009), define tipificar como: “Ajustar varias cosas semejantes a un tipo o norma común”. Por tanto, para tipificar el café comercial existen diferentes alternativas metodológicas, para demostrar la semejanza entre grupos, los cuales podemos mencionar:

4.5.1. Análisis físico-químico

Se entiende por análisis de café a la actividad que conduce a la descomposición de las muestras en sus elementos físicos o químicos que la componen, con el propósito de estudiarlos o caracterizarlos en detalle (Guerrero, 2010). El estudio químico de la muestra de café puede ser realizado de dos maneras: 1) para conocer “cuales” componentes se encuentran presentes, se realiza un análisis químico cualitativo; ó 2) para conocer “cuanto” de esos componentes se encuentra presente; esto es un análisis químico cuantitativo.

El análisis físico del café se refiere al estudio de las propiedades de los cuerpos sin modificar su naturaleza íntima. Mientras que el análisis químico busca conocer la composición íntima, es decir, la clase de elementos químicos que las constituyen. La determinación de color, uniformidad de color y tamaño de partícula (una vez molido), son ejemplos de análisis físico mientras que la cuantificación del contenido de mineral y compuestos como humedad, cenizas, grasas, almidones y azúcares, son ejemplos de análisis químico (NOM-169-SCFI-2007, NOM-149-SCFI-2001).

4.6. Métodos Instrumentales

Los métodos instrumentales presentan, actualmente, un gran interés principalmente por su elevada velocidad de análisis. Algunas de las técnicas basadas en este tipo de generación de señales del conjunto de componentes de una muestra son: espectroscopia en el infrarrojo cercano (NIRS) (Wang *et al.*, 1995, Macho y Larrechi., 2002, Zhang *et al.*, 2002, González *et al.*, 2003) e infrarrojo medio (MIR) (Carvalho, 2005), espectroscopia ultravioleta-visible (Thurston *et al.*, 2004) y pirolisis acoplada a espectrometría de masas (Goodacre y Kell., 1996, Radovic *et al.*, 2001, Péres *et al.*, 2002), entre otros. También puede incluirse en este grupo de técnicas la que suele denominarse olfatometría electrónica.

4.6.1. Espectroscopia en el infrarrojo cercano (NIRS)

Dada la necesidad de conocer el origen y la calidad de los alimentos de consumo humano; en la industria alimentaria es necesario realizar la determinación de ciertas propiedades que nos sirvan para caracterizarlos. Para ello, se llevan a cabo los métodos cuantitativos y cualitativos tradicionales de análisis de alimentos, estos son dificultosos, demandan tiempo, mano de obra y son costosos (Murray, 1986; Osborne *et al.*, 1993; Cozzolino, 1998). La espectroscopia (NIR) es una técnica analítica rápida no destructiva, que correlaciona la radiación infrarroja cercana con las propiedades químicas y físicas de los materiales (Chang and Laird, 2002).

4.6.2. Aplicaciones de la tecnología NIRS

La aplicación de la espectroscopia de reflectancia en el infrarrojo cercano (NIRS), se emplea desde la década de los 70' en la industria alimenticia, farmacéutica, petroquímica, como

alternativa a los métodos químicos y químico-biológicos convencionales (Norris *et al.*, 1976; Murray, 1986; Cozzolino *et al.*, 2000). La tecnología NIRS es una técnica rápida, no destructiva, ni contaminante y de gran exactitud, siempre que se sigan los procedimientos adecuados para crear las ecuaciones de calibración (Alomar y Fuchslocher, 1998; Givens y Deaville, 1999).

En la industria se han realizado estudios aplicados en el análisis de alimentos para animales (Cozzolino, 2002); en la predicción de los minerales traza en las legumbres (Cozzolino y Moron, 2004); para determinar la calidad interna de la pera (Liu *et al.*, 2008); en el control de la calidad interna del tomate (Flores *et al.*, 2009); en la determinación de los contenidos de sacarosa en chocolates (Da costa, 2008), el desarrollo y aplicación de la espectrometría de reflectancia en el infrarrojo cercano (NIRS) para la caracterización y modelos de calibración de germoplasmas en *Solanum phureja* (papas cultivadas), para identificar el alto contenido de carotenoides y la evaluación de variedades actuales para el mejoramiento nutricional de la papa (Bonierbale *et al.*, 2009).

4.7. Teoría de la Espectroscopia en el Infrarrojo Cercano (NIRS)

El método utiliza la región del infrarrojo cercano que comprende longitudes de onda entre los 700 y 2500 nanómetros (nm) del espectro electromagnético (Alomar y Fuchslocher, 1998; Murray, 1986; Deaville y Flinn, 2000). En este método, cuando la luz (energía) incide sobre una muestra, una parte de los fotones se transmite a través de la misma, siendo el resto absorbida. La absorción de energía por la muestra produce que los enlaces entre C–H, N–H, O–H, S–H, C=O y C=C componentes principales de la estructura básica de las sustancias orgánicas; vibren en distintas formas (Givens y Deaville, 1999).

La interacción de la energía con la materia obedece a la ley de Lambert-Beer, que establece que la absorbancia a cualquier longitud de onda es proporcional al número o concentración de moléculas absorbentes presentes en el camino recorrido por la radiación (Alomar y Fuschlocher, 1998; Cozzolino, 1998).

Esto determina que para un material de naturaleza química heterogénea, el espectro obtenido en la región del infrarrojo cercano sea una combinación de bandas de absorciones parciales sobrepuestas o muy cercanas, que suelen confundirse en una línea suavizada, en que se encuentran picos, valles y curvaturas en forma de hombro (Alomar and Fuschlocher, 1998; Deaville y Flinn, 2000). Al desarrollar una calibración NIRS, la información espectral (óptica) se relaciona mediante un algoritmo con la información de la composición físico-química (método de referencia) a través de la aplicación de modelos estadísticos como son la regresión múltiple, regresión por componentes principales y los mínimos cuadrados parciales.

La espectroscopia de infrarrojo cercano (NIRS), es usada en muchas áreas, los espectros NIR de una muestra son medidos típicamente por escáneres en cientos de longitudes de onda espaciados igualmente. La información en la curva es usada para predecir la composición química de la muestra por extracción de información relevante de muchos picos sobrepuestos. Muchos autores enfatizan la necesidad de los métodos para eliminar el ruido de los datos y seleccionar las características relevantes. La transformación de pequeñas ondas y el análisis discriminante estándar son aplicadas para remover el ruido. El espectro NIR muestra sobretonos y combinación de las bandas de estos grupos. Desafortunadamente, las bandas de absorción en la región NIR son amplias y se sobreponen, lo cual significa que las técnicas convencionales de calibración univariadas, usando solo una longitud de onda por componente para evaluaciones, no pueden ser aplicadas en casos de sobreposición de bandas. El desarrollo de herramientas estadísticas más complejas, como el método multivariado de regresión de mínimos cuadrados parciales (PLS), da la posibilidad para la amplia aplicación de espectroscopia NIR a muchos laboratorios (Nikolich *et al.*, 2001).

La espectroscopia NIR ofrece un gran número de ventajas para análisis cualitativos y cuantitativos y aplicaciones de control de procesos. Entre las principales ventajas se tiene: no preparar la muestra para el análisis, reducción del error analítico, no hay residuos, costos reducidos, mediciones y análisis más rápidos, componentes ópticos no giroscópicos, mediciones remotas por fibra óptica, mayor exactitud en los análisis y fácil uso.

Con la técnica de reflectancia de infrarrojo cercano, la muestra ha analizar está sujeta a radiación NIR, la superficie de la muestra absorbe parte de la energía, mientras que el resto es dispersada en todas direcciones. La luz dispersada proporciona información sobre la composición de la muestra en la prueba y el espectro resultante es usado en el análisis cualitativo y cuantitativo (Kubelka, 1948). Los mejores intervalos de frecuencia están en la región del espectro en el cual la intensidad de luz reflejada se correlaciona mejor con las concentraciones de las especies químicas analizadas.

4.8. Quimiometría

La Quimiometría puede definirse como una disciplina química que utiliza métodos matemáticos y estadísticos para diseñar procedimientos de medida y para obtener la mayor cantidad de información química posible a partir del análisis de datos (Massart *et al.*, 1998; Ramis, 2001).

Dentro de la Quimiometría pueden citarse los métodos de diseño experimental y exploración de datos, los métodos de reconocimiento de pautas, calibración y regresión simple y multivariante así como los sistemas expertos e inteligencia artificial, entre otros (Massart *et al.*, 1998).

En análisis multivariante cada muestra puede considerarse como un punto en un espacio multidimensional. El objetivo es extraer la información útil de todo el conjunto de las variables medidas en una experiencia. El diseño de experimentos es un complemento de este análisis porque proporciona una estrategia para diseñar la experimentación de un proceso. Esto permite obtener la mayor información útil posible con el menor número de experimentos. Además, la calidad de los parámetros calculados depende exclusivamente del diseño seleccionado y ésta es óptima (menor varianza) cuando la matriz del diseño es ortogonal, es decir, las concentraciones de los diferentes compuestos no están correlacionadas (Massart *et al.*, 1998).

Los métodos de reconocimiento de pautas tienen la finalidad de clasificar o agrupar muestras según las características de las mismas (Brereton, 1990). Pueden distinguirse dos tipos de técnicas de reconocimiento de pautas: supervisadas y no supervisadas.

Las técnicas no supervisadas tienen como objetivo determinar relaciones entre muestras sin utilizar ninguna información previa con respecto a la pertenencia a determinadas clases. Permiten conocer la existencia de estructuras en los datos y tener una idea de la utilidad de la información contenida en los mismos. Entre estos métodos se encuentran el análisis de agrupamientos, una de cuyas modalidades es el agrupamiento jerárquico (HCA) en el que las muestras se asocian en función de unos índices de similitud, y el análisis de componentes principales (PCA), en el que se reduce la dimensión de un conjunto de datos mediante el cálculo de combinaciones lineales de las variables originales.

Las técnicas de reconocimiento de pautas supervisadas tienen como fin desarrollar reglas de clasificación y/o modelado de clases para muestras (objetos) desconocidas a partir de un grupo de muestras conocidas que están caracterizadas por los valores de las variables medidas. El éxito de la predicción depende de que las reglas de clasificación y modelación así como de que los límites marcados por ellas sean adecuados. El objetivo final es predecir si una muestra desconocida pertenece a un grupo (o clase) o no. Algunas de las técnicas más importantes de clasificación y modelado de clases son el análisis discriminante lineal (LDA), modelado independiente de clases (SIMCA), “support vector machines” (SVM), el método de los k vecinos más próximos (k-NN) y las redes neuronales artificiales (ANN), entre otros.

Los métodos de calibración multivariante se basan en concentrar toda la información de las variables medidas en un número menor de ellas sin pérdida de información relevante. Dentro de estas técnicas se engloban la regresión por componentes principales (PCR) y la regresión por mínimos cuadrados parciales (PLS). A continuación se describen brevemente los fundamentos de las diferentes técnicas quimiométricas que se utilizan en este trabajo:

- Como técnicas exploratorias PCA
- Como modelado de clases SIMCA

4.8.1. Reconocimiento de pautas no supervisadas

4.8.1.1. Análisis de agrupamientos jerárquicos

En esta técnica de análisis exploratorio se calcula y compara la similitud entre las muestras. Algunas de las posibilidades para determinar la similitud entre los objetos se describen a continuación (Brereton, 2003):

- Coeficientes de correlación entre las muestras.

Un coeficiente de correlación de 1 implica que las muestras tienen características idénticas. Este coeficiente se obtiene cuando una muestra se compara con ella misma. Los valores pueden oscilar entre -1 y 1 aunque es habitual trabajar con el cuadrado de este número, es decir, el cuadrado del coeficiente de correlación.

- Distancia Euclídea.

La distancia Euclídea entre dos muestras k y l está definida por:

$$d_{kl} = \sqrt{\sum_{j=1}^J (x_{kj} - x_{lj})^2} \quad (1)$$

donde J representa el número total de variables medidas y x_{kj} y x_{lj} son los valores de la variable j para las muestras k y l , respectivamente. Esta medida es, en cierta forma, contraria al coeficiente de correlación y estrictamente es una medida de “no similitud”. No hay restricciones para su valor si bien es siempre positivo. La ecuación (2) es la notación matricial de la ecuación (1):

$$d_{kl} = \sqrt{\left(\begin{matrix} \mathbf{x}_k \\ \mathbf{x}_l \end{matrix} \right) \left(\begin{matrix} \mathbf{x}_k \\ \mathbf{x}_l \end{matrix} \right)^T} \quad (2)$$

- Distancia de Manhattan.

La distancia entre dos muestras k y l se determina según:

$$d_{kl} = \sum_{j=1}^J |x_{kj} - x_{lj}| \quad (3)$$

Las diferencias entre las distancias Euclídea y de Manhattan se ilustran en la Figura 3 en el caso de dos variables. Esta distancia es siempre mayor o al menos igual que la Euclídea.

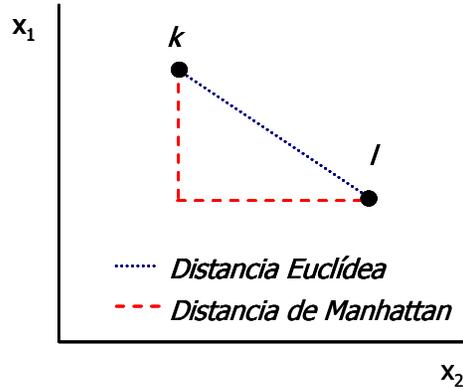


Figura 3. Distancia Euclídea y de Manhattan para las muestras k y l .

- Distancia de Mahalanobis.

Se asemeja a la distancia Euclídea pero tiene en cuenta que algunas variables pueden estar correlacionadas y por lo tanto miden en cierto grado la misma propiedad. La distancia entre dos objetos k y l se puede definir como:

$$d_{kl} = \sqrt{(\mathbf{x}_k - \bar{\mathbf{x}}_l) \mathbf{C}^{-1} (\mathbf{x}_k - \bar{\mathbf{x}}_l)^T} \quad (4)$$

Donde C es la matriz varianza-covarianza de las variables. Es una matriz simétrica cuyos elementos representan las covarianzas entre pares de variables. Los elementos de la diagonal corresponden a la varianza de cada variable. La matriz es de dimensiones $J \times J$.

Este método no es fácilmente aplicable cuando el número de variables excede al de muestras analizadas debido a que la matriz C no tiene inversa. Una posible solución requeriría seleccionar un número más pequeño de variables. Su utilización requiere trabajar con componentes principales en lugar con las variables originales (ver sección, PCA).

El análisis de agrupamientos puede dividirse en dos tipos de técnicas: agrupamientos jerárquicos y no jerárquicos.

En los agrupamientos jerárquicos se generan dendrogramas en los que se visualizan las distancias entre muestras y en los que pueden considerarse diferentes grupos en función de determinados niveles de similitud.

En los agrupamientos no jerárquicos es necesario prefijar el número de grupos en los que se quieren dividir las muestras y éstos se generan mediante un proceso iterativo en el que la pertenencia de cada muestra a un grupo concreto se va modificando hasta llegar a una situación estable.

En el caso del análisis de agrupamientos jerarquizados (HCA), después de calcular las distancias entre todos los pares de muestras, el siguiente paso es generar los agrupamientos. Los pasos que se deben seguir son los siguientes (Brereton, 2003):

- a) Encontrar las dos muestras más similares, esto es, aquellas entre las que existe la menor distancia.
- b) Formar un grupo con estas muestras. Ahora se sustituye el grupo ya formado por un punto que lo represente. Para seleccionar al representante se suelen aplicar los métodos (Ramis, 2001) que se ilustran en la Figura 4. En el método del centroide, los puntos que ya forman

grupo quedan representados por su centroide. Sin embargo en los otros métodos, el grupo queda representado por aquel punto que presente la menor distancia con respecto a la nueva muestra en el caso de la unión simple y por el de mayor distancia si se aplica el método de la unión completa. A continuación se recalculan las nuevas distancias entre el nuevo grupo formado y todas las demás muestras buscando de nuevo el valor más pequeño.

c)

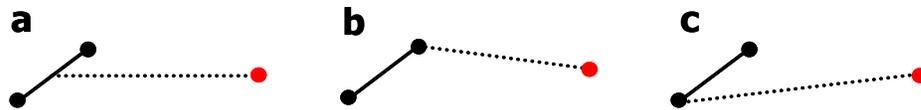


Figura 4. Unión de un grupo con una muestra aislada de acuerdo con los métodos del centroide (a), de la unión simple (b) y de la unión completa (c).

d) Los siguientes pasos consisten en agrupar los datos de manera similar hasta que todos los objetos se unan.

El resultado es una representación gráfica llamada dendograma. Es una estructura en forma de árbol donde la longitud de las ramas es proporcional a la distancia entre los grupos unidos. La forma del dendograma depende del método de unión. Habitualmente, las distancias entre muestras se transforman según la siguiente escala de similitud:

$$S = 1 - \frac{d_{kl}}{d_{\max}} \quad (5)$$

donde d_{\max} es la mayor distancia del conjunto de datos.

En el ejemplo de la Figura 5 los objetos se dividen formando fundamentalmente 2 agrupamientos (A y B). Las muestras 1, 4, 7 y 2 presentan características similares entre sí pero distintas a las del otro grupo, constituido por las muestras 6, 8, 5 y 3.

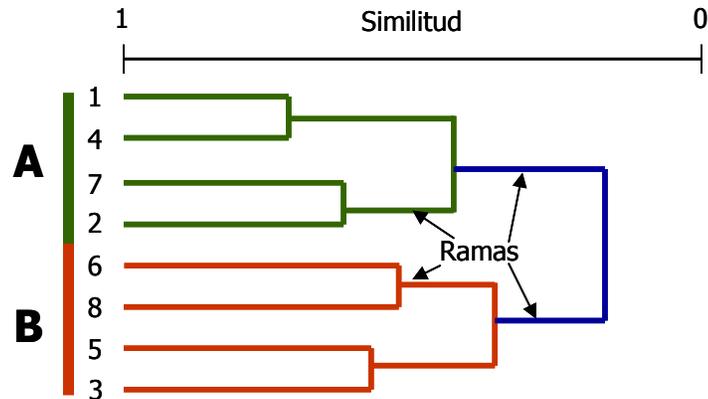


Figura 5. Dendograma.

4.8.1.2. Análisis de componentes principales

Es una técnica de reconocimiento de pautas no supervisadas que permite encontrar la estructura “latente” en una tabla de datos cuya lectura directa lo impide. El método trata de reducir la dimensión del conjunto de datos construyendo unas nuevas variables, ortogonales entre sí, que reciben el nombre de componentes principales (PC) y que son combinaciones lineales de las variables originales.

En este análisis la matriz de datos X formada por I objetos o muestras y J variables (dimensión $I \times J$) se descompone (Massart *et al.*, 1998) en el producto de las matrices T y P que contienen la información relacionada con las muestras y con las variables, respectivamente.

$$X = T \cdot P + E \quad (6)$$

En la expresión (6) la matriz T contiene las puntuaciones, valores (*scores*) de cada muestra en las nuevas variables, de I objetos en A componentes principales. La matriz P contiene los pesos, coeficientes (*loadings*) en las variables originales para generar las nuevas, de J variables sobre A componentes principales. E es una matriz residual. Las matrices de

puntuaciones y pesos consisten en una serie de vectores columna y fila respectivamente según muestra la Figura 6. Ambas matrices están formadas por tantos vectores como PCs puedan calcularse. Estos vectores se representan como t_a y p_a donde a es el componente considerado y reciben el nombre de vectores principales. Si la matriz de datos originales es de dimensiones $I \times J$, el número máximo de PCs que pueden ser calculados es el valor de la dimensión más pequeña.

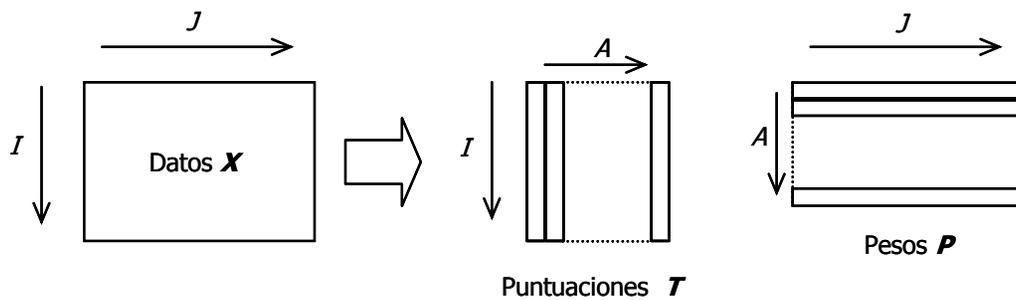


Figura 6. Análisis de componentes principales (PCA).

El objetivo de la técnica es que el número de componentes principales elegido explique la parte determinante de la matriz de datos de modo que los residuales sean lo más pequeños posibles (Blanco and Cerdá, 1988). En el caso de estudiar una mezcla de compuestos (Brereton, 2003), el número de PCs debería ser idéntico al número de compuestos existentes. Sin embargo, esto no es generalmente tan simple y debemos tener en cuenta otros factores como por ejemplo el ruido que pueden presentar los datos. Los primeros componentes son sin duda la mayoría de las veces los más importantes. El primero explica el mayor porcentaje de varianza de los datos, el segundo el mayor de la restante y así sucesivamente. Mediante el proceso de selección del número de PC se retiene la información significativa de las matrices y se separa de aquella relacionada con el ruido.

La importancia de cada componente puede medirse mediante la suma de los cuadrados de las puntuaciones según la expresión:

$$g_a = \sum_{i=1}^I t_{ia}^2 \quad (7)$$

donde g_a es el valor propio (*eigenvalue*) del componente principal a . Además la suma de todos los valores de g_a es la suma de los cuadrados de los valores de la matriz de datos X :

$$\sum_{a=1}^K g_a = \sum_{i=1}^I \sum_{j=1}^J x_{ij}^2 \quad (8)$$

donde K corresponde al valor más pequeño de I ó J . Frecuentemente estos valores se presentan como porcentajes tal y como muestra la expresión (9).

$$V_a = 100 \cdot \frac{g_a}{\sum_{i=1}^I \sum_{j=1}^J x_{ij}^2} \quad (9)$$

donde V_a es el porcentaje de información modelada por el componente a .

El porcentaje acumulado es utilizado a menudo para determinar la proporción de información que ha sido modelada.

Tras aplicar esta técnica, los datos originales quedan matemáticamente modelados por medio de unos pocos componentes principales. Usando un número A de componentes es posible establecer un modelo para cada elemento de X de la forma:

$$x_{ij} = \sum t_{ia} \cdot p_{aj} + e_{ij} = \hat{x}_{ij} + e_{ij} \quad (10)$$

que es la expresión no matricial de la ecuación (6).

Una forma más clara de visualizar esta sección es buscar una interpretación gráfico-geométrica a este análisis (Blanco and Cerdá, 1988). Así, un objeto queda definido por un punto en el hiperespacio constituido por J variables. Esta técnica trata de reducir la dimensión construyendo unas nuevas variables de forma que los objetos estén situados en el nuevo espacio, de forma similar a como lo estaban en el original. Esto produce un giro en el

hiperespacio definido por las J variables de forma que la primera variable nueva (primer eje calculado) contenga la máxima información, la segunda (ortogonal a la primera) el máximo de la restante y así sucesivamente. La matriz de puntuaciones está constituida por las proyecciones de cada uno de los objetos del conjunto de datos en los nuevos ejes o componentes principales. La representación de las puntuaciones de los objetos sobre los componentes principales permite reconocer agrupamientos de las muestras. Esto se ilustra en la Figura 7.

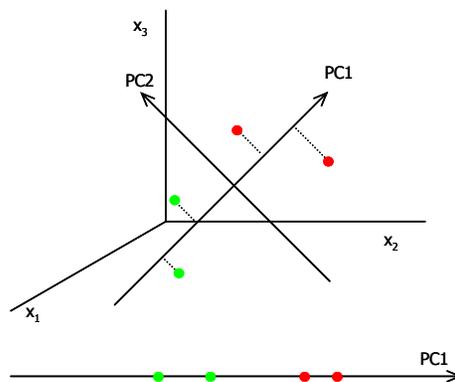


Figura 7. Representación de las puntuaciones de dos clases de objetos en los dos primeros componentes principales.

4.8.2. Reconocimiento de pautas supervisadas

Los métodos de reconocimiento de pautas supervisadas son también conocidos como métodos de clasificación. Su objetivo es clasificar muestras desconocidas en determinados grupos o clases a partir de la información suministrada por un grupo de entrenamiento. Aunque hay numerosos algoritmos es habitual utilizar una estrategia común para clasificar las muestras independientemente de la herramienta matemática que se utilice. La estrategia consta de las siguientes etapas (Brereton, 2003):

- a) Crear un modelo con el grupo de entrenamiento.

El primer paso es generar un modelo matemático que relacione los valores de las variables medidas del conjunto de muestras con el grupo al que pertenecen. Estos objetos reciben el nombre de conjunto de entrenamiento.

b) Evaluación de la capacidad predictiva del modelo.

Después de generar el modelo se determina su capacidad predictiva. Un parámetro representativo puede ser el porcentaje de objetos correctamente clasificados.

Esto puede realizarse siguiendo distintos criterios. Dos de los más habituales se describen a continuación:

- Validación cruzada (*cross validation*).

Es una forma de validación interna del modelo. Una muestra o un conjunto pequeño de ellas se excluyen del conjunto de entrenamiento. Ahora se construye un nuevo modelo sin las muestras excluidas y se comprueba la capacidad de predicción del modelo sobre ellas. El procedimiento se repite hasta que todas las muestras han quedado fuera del modelo al menos una vez.

- Validación externa (*test set*).

Para comprobar la calidad del modelo se realiza la predicción de un grupo independiente de muestras que no ha sido utilizado en ninguna etapa de la construcción del modelo. Se asume que estas muestras son desconocidas y se comparan los valores obtenidos tras aplicar el modelo con los valores reales conocidos.

c) Mejoramiento de los datos.

Cuando el modelo generado no es satisfactorio existen diferentes posibilidades de mejorarlo. Por un lado se puede utilizar otro algoritmo de cálculo o realizar una selección de variables ya

que algunas pueden estar ampliamente influidas por el ruido. En lugar de utilizar todas las variables medidas se seleccionan sólo aquellas que aportan mayor información. El resto podrían estar relacionadas con el ruido.

d) Aplicación del modelo.

Una vez que se dispone de un modelo satisfactorio, éste podrá ser aplicado a muestras desconocidas. Un aspecto importante que se debe tener en cuenta es la estabilidad del modelo con el tiempo, es decir, cuando el conjunto desconocido de muestras se analiza semanas o incluso meses después que las correspondientes del grupo de clasificación. En algunas ocasiones es posible que existan variaciones en la respuesta instrumental registrada debido a problemas de inestabilidad de la señal con el transcurso de los días. Esto puede tener seria influencia sobre la capacidad del modelo para clasificar satisfactoriamente una muestra. En problemas de clasificación es habitual realizar algún tipo de normalización de los datos antes de su análisis.

4.8.2.1. Modelado independiente de clases: “SIMCA”

Uno de los métodos de reconocimiento de pautas supervisadas más utilizados es el modelado de clases SIMCA (Massart *et al.*, 1998; Brereton, 1990).

Mientras que en los métodos de clasificación cada muestra se asigna solamente a una clase y ninguna queda fuera de alguna de ellas, el modelado de clases permite que el resultado de la clasificación para una muestra sea que no pertenece a ninguna de las clases o que pertenece a una o más de una (con la misma o diferente probabilidad). En muchas situaciones reales no es posible asignar una muestra concreta a alguna de las clases o existe la posibilidad de que algunas muestras pertenezcan a dos clases simultáneamente.

El método se basa en construir un modelo independiente para cada una de las clases mediante PCA. El número de componentes principales utilizados para cada clase puede prefijarse si se dispone de información previa o mantener un número de ellos que explique un determinado

porcentaje de la varianza de los datos. En la Figura 8 se presentan dos clases de objetos, A y B. Al aplicar PCA a la clase A (puntos verdes) se obtiene un modelo de dos componentes (un plano), mientras que la clase B (puntos rojos) se ajusta a un modelo de una única variable (una recta). Las regiones de confianza aparecen alrededor de cada clase.

La asignación de los nuevos objetos se hace comprobando si están localizados en la región de confianza de algún modelo. Así el objeto k de la Figura 8 se puede asignar a la clase B mientras que h no pertenece a ninguna clase, es un “outliers”.

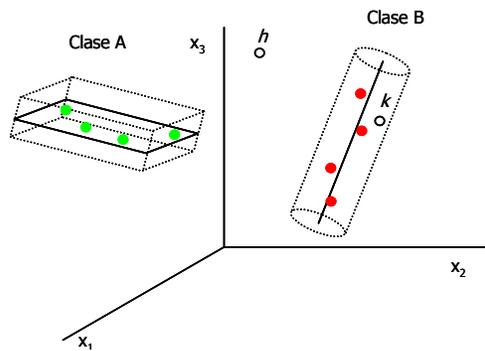


Figura 8. Modelos de dos clases de objetos A y B en el espacio de sus componentes principales, regiones de confianza y predicción de dos muestras.

Para cada clase y su modelo correspondiente se define la sensibilidad como el porcentaje de muestras que perteneciendo a esa clase son reconocidas correctamente por el modelo matemático. Análogamente, se define la especificidad como el porcentaje de muestras que perteneciendo a otra clase son reconocidas como ajenas al modelo.

Los diagramas de Coomans permiten visualizar estos índices ya que son representaciones bidimensionales de las distancias de las muestras a los modelos. La Figura 9 es un esquema de estos diagramas. En el eje de abscisas se representa la distancia a la clase A y en el de ordenadas a la clase B. La recta horizontal discontinua delimita el recinto de la clase B mientras que la línea vertical discontinua delimita el recinto de la clase A. El cuadrado inferior izquierdo (1) es el recinto común a ambas clases y cuantas menos muestras tengan sus coordenadas en él, mayor será la especificidad de los modelos. El cuadrado superior derecho

(2) representa el espacio que queda fuera de ambos modelos. Las muestras con coordenadas en él son muestras extrañas a los modelos generados.

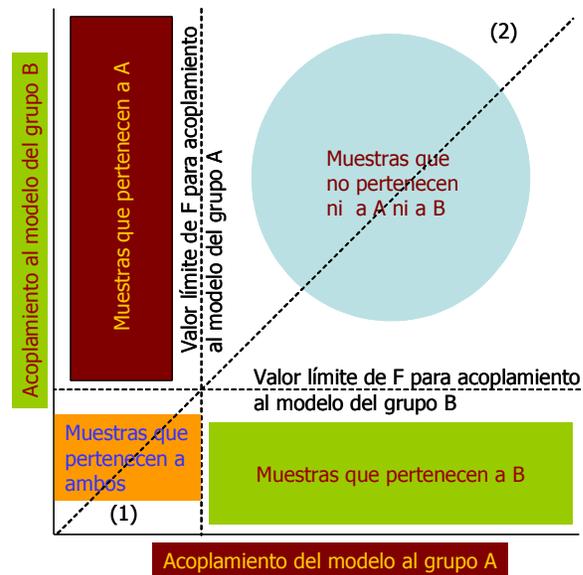


Figura 9. Diagrama de Coomans.

Cuando SIMCA se utiliza únicamente como técnica de clasificación, todas las muestras que se encuentren por encima de la diagonal serán asignadas a la clase A y las muestras por debajo a la clase B.

5. LITERATURA CITADA

- Alomar, D., y Fuchslocher, R. 1998. Fundamentos de la espectroscopia de reflectancia en el infrarrojo cercano (NIRS) como método de análisis de forrajes. *Agrosur*, 26(1):88-104.
- ASERCA. 2002. Claridades agropecuarias en Café; Calidad de exportación. México, D. F. Rev. N°103. 42 p.
- Blanco M. y V. Cerdá. 1988. Quimiometría. Publicaciones de la Universidad Autónoma de Barcelona. Bellaterra. España.
- Block, E. 1985. The chemistry of garlic and onions. *Scientific American*, 252(3): 114-119.
- Bonierbale, M., Grüneberg, W., Amoros, W., Burgos, G., Salas, E., Porras, E., and Felde, T. 2009. Total and individual carotenoid profiles in *Solanum phureja* cultivated potatoes: II. Development and application of near-infrared reflectance spectroscopy (NIRS) calibrations for germoplasm characterization. *Journal of food composition and Analysis* 22:509-516.
- Brereton R. G. 1990. *Chemometrics Applications of Mathematics and Statistics to Laboratory Systems*. Ellis Horwood, Chichester.
- Brereton R. G. 2003. *Chemometrics Data Analysis for the Laboratory and Chemical Plant*. Wiley, Chichester.
- Camacho Reyes Juan. 1985. Cafés de México. (En línea). *Revista Técnica Mensual del Sector Cafetalero. Historia del Café*. Consultado 15 Abril 2010. Disponible en: <http://www.cafesdemexico.com/historia.htm>.
- Cámara de Diputados, San Lázaro. 2001. El mercado del café en México. Centro de estudios de finanzas públicas. México, D.F. Diciembre.

- Carvalho A. R., Nogal Sánchez, M., Wattoon, J., and Brereton, R. G.. 2005. Comparison of PLS and kinetic models for a second-order reaction as monitored using ultraviolet visible and mid-infrared spectroscopy. *Talanta*, 68 (4), pp. 1190-1200.
- Chang, C. W., Laird, D.A., 2002. Near-infrared reflectance spectroscopy analysis of soil C and N. *Soil Science*, 167(2):110-116.
- Cih D., I. R., A. Pineda O., R. Valdivia A. y F. González C. 2001. Efecto del tipo de cambio en las exportaciones de café mexicano en el periodo 1981-1998. *Comunicaciones en Estadística e Informática* 5(1):1-76.
- Clarke, R. 1987. Roasting and grinding. In: Clarke, R., Macrae, R. (eds). *Coffee Technology*, Elsevier Applied Science Publishers, London, 2: 73-108.
- Consejo Mexicano del Café. 2002. www.sagar.gob.mxCmc.café04spl.htm.
- Costell, E.; Durán, L., 1981. El análisis sensorial en el control de calidad de los alimentos. III Planificación, selección de jueces, y diseño estadístico. *Revista Agroquímica y Tecnología de Alimentos*, 21(4): 454-460.
- Cozzolino, D. 1998. Aplicación de la tecnología NIRS para el análisis de calidad de los productos agrícolas. Serie Técnica INIA 97. México.
- Cozzolino D. 2002. Uso de la espectroscopia de reflectancia en el infrarrojo cercano (NIRS) en el análisis de alimentos para animales. *Agrociencia* 6(2) 24-32.
- Cozzolino, D., Fassio, A and Gimenez, A. 2000. The use of near infrared reflectance spectroscopy (NIRS) to predict the composition of whole maize plants. *J. Sci. Food Agri.* 81 (2)140-146.

- Cozzolino D., and Moron A. 2004. Exploring the use of near infrared reflectance spectroscopy (NIRS) to predict trace minerals in legumes. In: *Animal Feed Science and Technology* 111.
- Da Costa Filho P.A. 2008. Rapid determination of sucrose in chocolate mass using near infrared spectroscopy. *Analytica Chimica Acta* 631:206-211.
- Deville E., R. and P. C. Flinn. 2000. Near infrared (NIR) spectroscopy: an alternative approach for the estimation of forage quality and voluntary intake. p. 301-320. In D.I. Givens, E. Owen, R.F.E. Axford and H.M. Omedi (eds.) *Forage evaluation in ruminant nutrition*. CABI Publishing, Wallingford, UK.
- Díaz C., S. y Castro F., A. 2005. *Evaluación Nacional del Programa de Fomento Productivo y Mejoramiento de la Calidad del Café*. México: Universidad Autónoma de Chapingo.
- FAO-SAGARPA, 2006. *Evaluación del fondo de estabilización del café*. México. D.F.
- Flores K, Sánchez Ma-T., Pérez-Martin D. Guerrero J. E y Garrido-Varo A. 2009. Feasibility in NIRS instruments for predicting internal quality in intact tomato. *Journal of food Engineering* 91:311-318.
- Folstar, P. 1985. Lipids. In: *Coffee*. Elsevier Applied Science Publishers, London and New York, I: 203-222.
- Fox, B.; Cameron, A. 1992. *Ciencia de los alimentos, nutrición y salud*. Editorial Limusa, México, 457 pp.
- Fundación Produce Chiapas A. C. y el Instituto Tecnológico y de Estudios Superiores de Monterrey. Campus Chiapas. *Cadena Agroalimentaria del café*. Abril de 2003.

- Givens, D. I., and E. R. Deaville. 1999. The current and future role of near infrared reflectance spectroscopy in animal nutrition: a review. *Aust. J. Agric. Res.* 50:1131-1145.
- González J. A. 1998. Diagnostico de la competitividad de la industria del café en Costa Rica.
- González-Martín I., C. González Pérez, J. Hernández Méndez, and N. Álvarez-García. 2003. Determination of fatty acids in the subcutaneous fase of Iberian breed swine by near infrared spectroscopy (NIRS) with a fibre-optic probe. *Meat Science* 65:713-719.
- Goodacre R. and Kell, D. B.. 1996. Correction of mass spectral drift using artificial neural networks. *Analytical Chemistry* 68:271-280.
- Guerrero, P. A. 2006. Estudio de la contaminación en suelos por petróleo crudo mediante generación de espacio de Cabeza-Espectrometría de masas. Tesis de Doctorado. Universidad de Salamanca. Salamanca, España.
- Guerrero P., A. 2010. Análisis Químico Agrícola. Nota de clase (PAT-626). Maestría en Producción Agroalimentaria en el Trópico. Campus Tabasco-Colegio de Postgraduados. H. Cárdenas, Tabasco.
- Hernández, H. J. M. 2008. Utilización de la tecnología NIRS con sonda de fibra óptica para el control de calidad de alfalfa. Tesis de Doctorado de Química. Universidad de Salamanca, España. 116p.
- Hodge, J. 1967. Origin of flavors in foods - non-enzymatic browning reactions. En: Schultz, H. Day, E. Libbey, L. (eds). *Symposium on Foods: The Chemistry and Physiology of Flavors*. AVI Publishing Co., Westport, Connecticut, 465-491.
- Kubelka, P. 1948. New Contributions to the Optics of Intensely Light-scattering Materials. Part I, *J Opt Soc* 38:448-457.

- Liu Y., Chen X, and Ouyang A. 2008. Nondestructive determination of pear internal quality indices by visible and near-infrared spectrometry. *LWT-Food Science and Technology* 28(10) 2318:2320.
- Macho S. and M. S. Larrechi. 2002. Near-infrared spectroscopy and multivariate calibration for the quantitative determination of certain properties in the petrochemical industry. *Trends in Analytical Chemistry* 21:799-806.
- Maga, J. 1975. The role of sulfur compounds in food flavor. *Food Science and Nutrition*, 6(2): 153-176.
- Maga, J. 1981. Pyrroles in foods. *Journal of Agricultural and Food Chemistry*, 29(4): 691-693.
- Maga, J. 1982. Pyrazines in foods: an update. *Food Science and Nutrition*, 16(1): 1-48.
- Massart, D. L., Vandeginste, G. M., Buydens, L. M. C., De Jong, S., Lewi, P. J., and Smeyers-Verbeke, J. 1998. *Handbook of Chemometrics and Qualimetrics: Part A*. Elsevier, Amsterdam.
- Murray, I. 1986. Near Infrared analysis of forages. In: *Recent advances in animal nutrition*. Edited by W. Haresign and D. J. A. Cole. *Studies in Agricultural and food sciences* Butterworths. Australia.
- Nikolich K., C. Sergides and A. Pittas. 2001. The application of Near Infrared Reflectance Spectroscopy (NIRS) for the quantitative analysis of hydrocortisone in primary materials. *J. Serb. Chem. Soc.* 66(3).
- Nolasco, M. 1985. *Café y sociedad en México*. Centro de Ecodesarrollo. México. D. F. 1ra edición. 454 p.

NMX-F-013-SCFI-2000. Café puro tostado, en grano o molido, sin descafeinar o descafeinado- especificaciones y métodos de prueba. Diario Oficial de la Federación.

NOM-169-SCFI-2007, Café Chiapas-Especificaciones y métodos de prueba. Diario Oficial de la Federación.

Norris, K.H., R.F. Barnes, J.E. Moore, and J.S. Shenk. 1976. Predicting forage quality by infrared reflectance spectroscopy. *J. Anim. Sci.* 43:889-897.

Osborne, B. G.; Fearn, T., and Hindle, P. H. 1993. *Near Infrared Spectroscopy in Food analysis*. Second Edition. Longman scientific and technical.

Pérès, C. C. Viallon, and J. L. Berdagué. 2002. Curie point pyrolysis-mass spectrometry applied to rapid characterization of cheeses. *Journal of Analytical and Applied Pyrolysis* 65:161-171.

Pérez-Grovas, G., Víctor. 1998. Evaluación de la sustentabilidad del sistema de producción de café orgánico en la unión Majomut en la región de los Altos de Chiapas. Universidad Autónoma Chapingo. Chapingo, Estado de México.

Poisson, J. 1977. Aspects chimiques et biologiques de la composition du café vert. En: 8th International Colloquium on the Chemistry of Coffee, ASIC, Paris, 219-234.

Radovic B. S., R. Goodacre, and E. Anklam. 2001. Contributions of pyrolysis-mass spectrometry (Py-MS) to authenticity testing of honey. *Journal of Analytical Applied Pyrolysis* 60:79-87.

Ramis Ramos G., M. C. García Álvarez-Coque. 2001. *Quimiometría*. Síntesis, Madrid. España.

Regalado, A. 1996. Manual para la cafecultura mexicana. 1996. Consejo mexicano del café. Programa café 1995-2000. INCA RURAL. México. 156 p.

Regalado O. A., 2006. ¿Qué es la calidad del café? Universidad Autónoma Chapingo. Chapingo, México.

Salazar, E.; R. Muschler; V. Sánchez y F. Jiménez. 2000. "Calidad de Coffea arabica bajo sombra de Erythrina poeppigiana a diferentes elevaciones en Costa Rica". Agroforestería en las Américas 26:40-42.

Santoyo C., V. H.; S. Díaz C.; E. Escamilla P. y J. D. Robledo M. 1996. Factores agronómicos y calidad de café. Universidad Autónoma Chapingo-Confederación de Productores de Café. Chapingo, México. 21 p.

Shimoda, M.; Shibamoto, T. 1990. Isolation and identification of headspace volatiles from brewed coffee with an on-column GC/MS method. Journal of Agricultural and Food Chemistry, 38: 802-804.

Tecnología Integral de Café. 1991. Curso sobre beneficiado eficiente del café. TECNICAFE S.A. Guatemala, Guatemala. 80 p

Thurston T. J., R. G. Brereton, D. J. Foord, and R. E. A. Escott. 2004. Principal components plots for exploratory investigation of reactions using ultraviolet-visible spectroscopy: application to the formation of benzophenone phenylhydrazone. Talanta 63:757-769.

Tressl, R.; Grünwald, K.; Koppler, H.; Silwar, R. 1978. Phenols in roasted coffees of different varieties. Zeitschrift für Lebensmittel Untersuchung und Forschung, 167(2):108-110.

Tressl, R.; Silwar, R. 1981. Investigation of sulphur-containing components in roasted coffee. Journal of Agricultural and Food Chemistry, 29(5): 1078-1082.

Wang Z., T. Dean, and B. R. Kowalski. 1995. Additive background correction in multivariate instrument standardization. *Analytical Chemistry* 67:2379-2385.

Wintgens, J. N. 2004. "Factors influencing the quality of green coffee". In: *Coffee: growing, processing, sustainable production*. Wintgens, J.N. (ed.). Wiley VCH. Weinheim. pp. 789-809.

Zhang L., C. W. Small, and M. A. Arnold. 2002. Calibration standardization algorithm for partial least-squares regression: application to the determination of physiological levels of glucose by near-infrared spectroscopy. *Analytical Chemistry* 74:4097-4108.

CAPITULO I: CARACTERIZACIÓN Y TIPIFICACIÓN DE CAFÉ COMERCIAL MEDIANTE LA TECNOLOGÍA NIRS

CARACTERIZACIÓN Y TIPIFICACIÓN DE CAFÉ COMERCIAL MEDIANTE LA TECNOLOGÍA NIRS

RESUMEN

El evaluar la calidad, mediante la caracterización y tipificación del café, así como la cuantificación de sus propiedades físicas y químicas, es una de las aplicaciones de la química analítica. En la actualidad existen avances instrumentales como lo es la tecnología NIRS, la cual permite realizar la estimación de los elementos y compuestos que analizan las muestras en solo un minuto de lectura. Para realizar esta aplicación tecnológica, es necesario el uso de la Quimiometría, la cual, es entendida como la aplicación de la matemática, estadística en la interpretación de los espectros NIR. En el presente trabajo fue realizado un estudio de caracterización y tipificación de diferentes muestras de café tostado y molido; mediante los métodos convencionales establecidos en la NMX-F-013-SCFI-2000 para café tostado y molido y se estudio la posibilidad de aplicar la tecnología NIRS en la tipificación de las muestras de café. Los resultados promedios para las diferentes muestras fueron: Chiapas (1.19h%, 4.20 c%, 18.4 g%), Oaxaca (1.55 h%, 3.89 c%, 11.0 g %), Veracruz (1.57 h%, 4.15c%, 13.1g%) y Tabasco (1.91h%, 4.23c%, 7.45g%). Se obtuvieron los diagramas de componentes principales (PCA) para los análisis fisicoquímicos y de las huellas espectrales obtenidas con el NIRS; expresando la mayor variabilidad (99%) de los datos en los dos primeros componentes; pudiendo diferenciar las muestras de café de acuerdo con su origen. Cuando utilizamos los espectros NIRS en el modelado de clase (SIMCA), las diferentes muestras de café se ubican perfectamente en sus aéreas de acoplamiento. Lo que significa que el uso de la tecnología NIRS es una alternativa para caracterizar y tipificar el café tostado y molido con resultados similares a los obtenidos con los métodos convencionales; con la ventaja de ser más rápido, con menor error experimental, etc., comparado con los métodos químicos convencionales.

Palabras claves: *Café, Caracterización, Tipificación, Quimiometría, NIRS*

ABSTRACT

Assessing the quality through characterization and classification of coffee, and the quantification of physical and chemical properties, is one of the properties of analytical chemistry. There are currently instrumental advances such as NIRS technology, which allows the estimation of the elements and compounds that are sampled in only one minute of reading. NIRS technology as it is understood, is a synthesis of spectroscopy, mathematics, statistics and instrumentation, which is a critical element in developing this technology. In the present work was evaluated the physicochemical quality of different samples of roasted and ground, using conventional methods established in the NMX-F-013-SCFI-2000 for roast and ground coffee. The average results for samples were: Chiapas (1.19h%, 4.20 c%, 18.4 g%), Oaxaca (1.55 h%, 3.89 c%, 11.0 g %), Veracruz (1.57 h%, 4.15c%, 13.1g%) and Tabasco (1.91h%, 4.23c%, 7.45g%). Diagrams were obtained principal components (PCA) for physicochemical analysis and NIRS readings, expressing the highest variability (99%) of the existing data in the first two components, to separate class modeling (SIMCA), different coffee samples by NIRS is perfectly located in different areas of coupling. This means that use of NIRS technology is an alternative to characterize and classify the roast and ground coffee to be faster compared to conventional chemical methods.

Keywords: *Coffee, Characterization, Classification, Chemometrics, NIRS*

1.1. INTRODUCCIÓN

El sistema de trazabilidad como una herramienta eficaz puede rastrear la calidad y reducir los problemas de anomalía de información en la selección adversa y riesgo moral en el sistema alimentario (Sikuta, 2005).

La Asociación Mexicana de Estándares para el Comercio Electrónico (AMECE-GS1 México) describe la trazabilidad como el proceso por el cual se obtiene la capacidad de rastrear históricamente el origen y destino, de un producto terminado, desde su creación, pasando por la fabricación de sus productos intermedios, la creación del producto terminado, hasta llegar al punto de venta o consumidor final. Para ello, se hace un registro del origen y destino del producto en cada una de las etapas por las que pasa en la cadena de abasto de los socios comerciales, logrando así un proceso sin ruptura o pérdida de trazabilidad (<http://www.alimentaria-mexico.com>).

Con el desarrollo exponencial en el análisis de los datos electrónicos, han desarrollado sistemas de capacidad de rastreo aplicados en productos alimenticios como pescado (Caswell, 2006; Yang, 2005), alimentos transgénicos (Miraglia *et al.*, 2004), alimentos refrigerados y congelados (Koutsoumanis *et al.*, 2005) así como alimentos de carnes (Bai Lu, and Li, 2005; Li and Yang, 2006; McMeekin *et al.*, 2006). La trazabilidad se refiere a las medidas, pero también a los métodos, instrumentos o laboratorios (Ramis, 2001).

Aunque en nuestro país los sistemas de trazabilidad y aseguramientos de la calidad a nivel de alimentos producidos en su explotación aun están pocos desarrollados, en otros países, como por ejemplo, el Reino Unido, han puesto en marcha sistemas de aseguramientos de la calidad a nivel de explotación. Los procedimientos incluyen tanto a las materias primas, para garantizar que estas sean seguras y cumplan con las especificaciones, como a los procesos del producto (De la Roza y Argamentería, 2009).

Es así como se ha tratado de profundizar en la implementación de la tecnología de infrarrojo cercano (NIRS), para el control de calidad, trazabilidad y seguridad desde el primer eslabón

de la cadena alimentaria, ya que la tecnología NIRS tiene un reconocido valor como método cuantitativo para obtener información analítica simultánea e inmediata de varios parámetros. Existen métodos convencionales y las pruebas analíticas que han sido usados para evaluar productos alimenticios. Sin embargo, el coste y las exigencias del trabajo son excesivos, lo que causa dificultades en el análisis de costes para la investigación u otros objetivos. La espectroscopía de infrarrojo cercano, es una técnica física con el potencial de permitir la determinación rápida y exacta de la composición química y valor nutritivo en diferentes alimentos (Bruno-Soares *et al.*, 1998).

La tecnología NIRS ha tenido aplicaciones en alimentos para análisis de forrajes (Dardenne *et al.*, 1991; Waters and Givens, 1992; Shenk and Westerbaas, 1994) cereales y alimentos (Murray and Garrido, 1990; Givens *et al.*, 1991; Dardenne *et al.*, 1993). Actualmente, se han hecho otros estudios como el uso de la tecnología NIRS para la valoración de la composición química de muestras de carne de buey (Prieto *et al.*, 2006), así como la evaluación del potencial de la espectroscopia de infrarrojo cercano (NIRS) como método rápido para estimar las características de cocción del arroz (Wu and Shi, 2007). Con el uso de la quimiometría se producen además diversos patrones de reconocimiento como el modelo SIMCA que permite clasificar materiales semejantes en grupos definidos y además discrimina muestras desconocidas (Beebe *et al.*, 1998).

El control de calidad en toda la composición del café tostado y molido, su trazabilidad a través de la cadena alimentaria, el rastreo y vigilancia en todo el proceso de producción así como la homogeneidad como mezcla final, es posible realizarlo mediante el registro de las huellas espectrales, como el NIRS, y su interpretación quimiométrica.

1.2. MATERIALES Y METODOS

El desarrollo del experimento se realizó en dos etapas, en esta primer parte se obtuvieron las huellas espectrales mediante la tecnología NIRS de las muestras de café comercial procedentes de los estados de Tabasco, Chiapas, Veracruz y Oaxaca respectivamente.

Esta primer etapa consistió en generar la información básica del café comercial mediante métodos físicos y químicos convencionales establecidos en la Norma Mexicana NMX-F-013-SCFI-2000 y comparar los resultados con los generados con la tecnología NIRS y la quimiometría para la caracterización y tipificación del café.

1.2.1. Localización y selección de las muestras de café comercial

El estudio se llevó a cabo con 120 muestras de café comercial, tal como se muestran en el Cuadro 2, con las cuales se realizó la determinación de humedad, ceniza y grasa con los métodos de referencias establecidos en la NMX-F-013-SCFI-2000. Las muestras fueron identificadas y registradas con un número consecutivo de laboratorio y fueron clasificadas, según el lugar de procedencia; es decir Chiapas, Oaxaca, Veracruz y Tabasco.

Cuadro 2. Procedencia de muestras de café comercial.

Estado	Procedencia	Coordenadas geográficas	No. de muestra
Chiapas	Tziscaco	Latitud: 16°45'57.86''N Longitud: 93°8'56.22''O Altitud: 1500 msnm	30
Oaxaca	Ixtepec	Latitud: 16°33'40.52''N Longitud: 95°5'50.63''O Altitud: 160msnm	30
Veracruz	Córdoba	Latitud: 18°53'39.00''N Longitud: 96°56'5.00''O Altitud: 680 msnm	30
Tabasco	Huimanguillo	Latitud: 17°49'49.00''N Longitud: 93°23'29.00''O Altitud: 30 msnm	30

Las muestras fueron clasificadas de acuerdo a un registro de laboratorio. Para los análisis fisicoquímicos las muestras se utilizaron molidas, las cuales se colocaron en botes de 500 g; para su conservación, fueron etiquetadas de acuerdo al registro de laboratorio.

Los parámetros que se midieron en las muestras de café fueron, humedad, ceniza y grasa las cuales fueron preparadas con base a los métodos de referencias establecidos en la NMX-F-

013-SCFI-2000; siendo cada muestra analizada por triplicado. Con lo cual se obtuvieron 360 datos (matriz de 120 x 3).

1.2.2. Relación de los métodos físico – químicos

Los métodos convencionales que se realizaron son los descritos en las normas mexicanas para análisis físico-químicos de café tostado y molido, mismo que se describen en el Cuadro 3.

Cuadro 3. Métodos de la Norma Mexicana NMX-F-013-SCFI-2000 para café tostado y molido.

Parámetros	Especificaciones	Método de prueba
Humedad (método gravimétrico)	6.0% máximo	NMX-F-013-SCFI-2000. Inciso 11.6
Cenizas (método de calcinación mediante una mufla)	5.0% bs máximo	NMX-F-013-SCFI-2000. Inciso 11.7
Grasa como extracto etéreo (método Soxhlet)	8.0% bs mínimo-18.0% bs máximo	NMX-F-013-SCFI-2000. Inciso 11.8

Para la realización de los análisis fisicoquímicos, las muestras se utilizaron molidas y tamizadas.

1.2.3. Software para la interpretación quimiométrica de los métodos físico-químicos y NIRS

Los datos que se obtengan de los análisis físico-químicos y de las lecturas en el instrumento propuesto (NIRS) serán analizados quimiométricamente, utilizando los programas:

- **ISISCAN**. Para la obtención de los espectros de las muestras.
- **WinISI**. Se utilizó como programa puente entre el ISISCAN y Excel. Los espectros son convertidos a variables de escala.
- **Excel**. Se utilizó como programa puente entre el WinISI y el Pirouette; así como para la realización de una prueba de estadística descriptiva de los análisis físico-químicos (media, desviación estándar, coeficiente de variación).
- **Pirouette Chemometrics**. Software versión 3.11 para Windows de Infometrix Inc. Copyright 2007, para la realización del análisis de componentes principales (PCA), el

modelado independiente de clase (SIMCA) y para la regresión por mínimos cuadrados parciales (PLS).

1.2.4. Espectrofotómetro de infrarrojo cercano (NIRS)

El equipo que se utilizó para generar las huellas espectrales fue un espectrómetro de reflectancia en el infrarrojo cercano modelo Versatile Agri Analyzer™6500 System II Vertical Type 461008, Serial No. 9503 100-240 Vac max 200 W, 50-60 Hz de la empresa FOSS NIRSystems. El instrumento registra el espectro de 400 a 2500 nm en longitudes de onda pares. Los programas utilizados para realizar la interpretación quimiométrica de las huellas espectrales fueron ISISCAN (versión 2081 copyright 2006, WinISI III y Excel, como programas puente; y Pirouette (Chemometrics Software versión 3.11 para Windows de Infometrix Inc. Copyright 2007) (Figura 10).



Figura 10. Espectroscopio de infrarrojo cercano, NIRs Versatile Agri Analyzer TM 6500.

Para realizar la lectura, el equipo cuenta con una lámpara de 4 Volts, ajustada a una placa prealineada, que manda el haz de luz blanca la cual contiene todas las frecuencias al *Grating*, una superficie holográfica de 570 líneas por mm y una distancia de 1.75 μm entre líneas, el cual desvía la luz en ángulos diferentes dependiendo de la longitud de onda; ésta luz reflejada es monocromada y el ángulo de inclinación del *Grating* es controlado por la tarjeta madre al filtro *Order sorter*, que esta sincronizado con el movimiento del *Grating* y es movido por medio de un motor de pasos, el cual permite el paso de la luz por longitud de onda para transmitirla a la muestra, contenida en una copa o celda opaca de cuarzo, y esta a los detectores de Silicio (400 a 1100 nm) y PbS (1100 a 2500 nm), que registran la cantidad de

luz reflejada y/o transmitida, para transformarla en una señal que se registra mediante el programa ISIScan (versión 2.81 Copyright 2006) a una computadora (Acer AMD Sempron™ Processor 3200+ 1.81 GHz, 448 MB de RAM).

1.3. RESULTADOS Y DISCUSIÓN

1.3.1. Espectros NIRS de los cuatro lotes de café

Los espectros del café tostado y molido se obtuvieron en las lecturas de 120 muestras, dado que el equipo genera 1050 variables por muestras en un rango que va de los 400 a 2500 nm en números pares; formando una matriz de 126000 datos o variables de escala. Estas variables representan en forma grafica las huellas espectrales características de cada muestra (Figura 11).

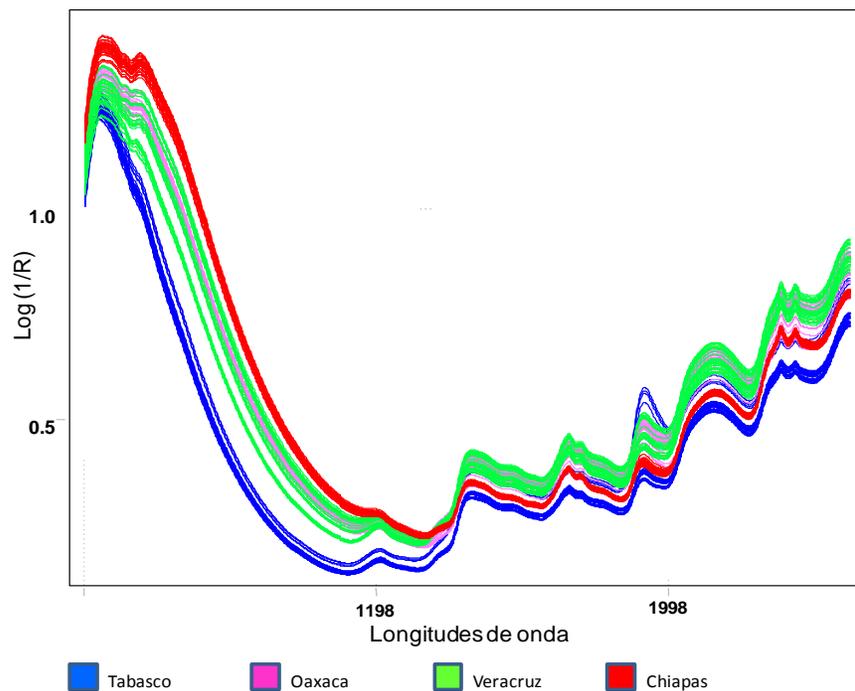


Figura 11. Espectros generados por cada una de las muestras de café comercial.

1.3.2. Análisis de componentes principales (PCA) de los análisis Físicoquímicos

Con los datos de los análisis físicoquímicos (Humedad, Ceniza y Grasa) fue realizado un análisis de componentes principales para las muestras de café comercial (Tabasco, Oaxaca, Veracruz y Chiapas). Los datos obtenidos indican que el componente principal uno (PC1) explica el 99.70% de la varianza total. Por ello, fue utilizado para obtener el coeficiente de variación de cada lote. En la Figura 12., se representa, la distribución de las muestras en los dos primeros componentes principales. Los valores obtenidos en el coeficiente de variación se reflejan en la distribución que tiene cada grupo de muestras de café, siendo el PC1 el componente que explica casi el 100% de la varianza de los datos (Cuadro 4).

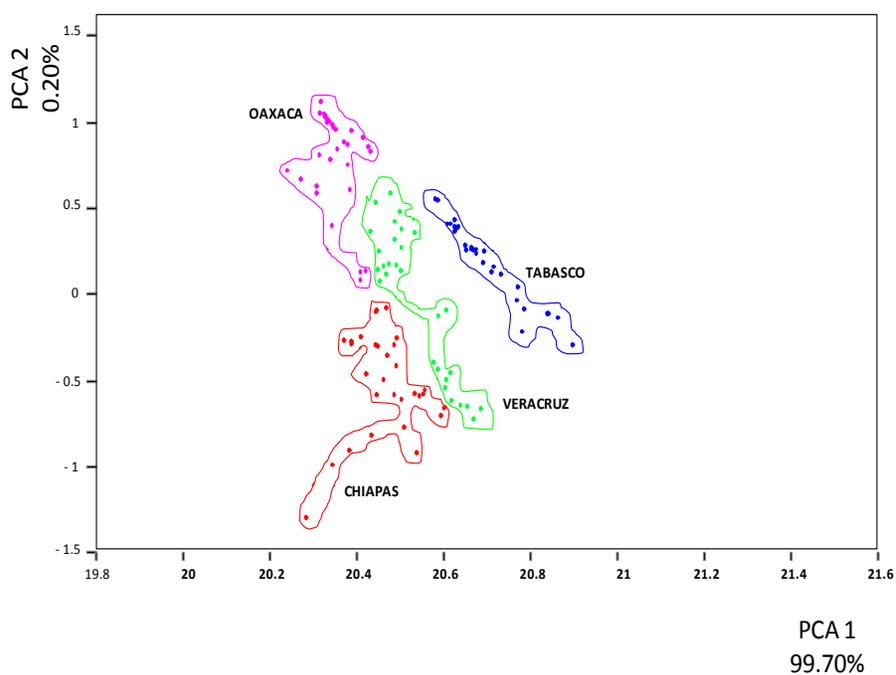


Figura 12. Componentes principales para los análisis físicoquímicos.

Cuadro 4. Análisis de componentes principales para los análisis físicoquímicos.

	Varianza	% Varianza explicada	% Varianza acumulada
PC1	76464.46	99.70	99.70
PC2	152.93	0.20	99.90

La máxima variabilidad existente (PC1), nos indica que los resultados son apropiados para expresar la variabilidad de los datos e indican que todas las muestras son similares entre sí.

De acuerdo con las componentes principales, la variable que modela y que explica la varianza máxima de los datos es la variable grasa (Cuadro 5, PC1); y puede ser un factor determinante para tipificar las muestras de los diferentes tipos de café. Esta cualidad asociada con la información generada por Menchu e Ibarra (1968), quienes indicaron que la grasa está asociada con la altura en la cual se produce el café; valida lo que observamos en las muestras estudiadas, las cuales proceden de zonas con diferentes alturas, y en las cuales, el coeficiente de correlación fue de 0.96 entre la altura y el contenido de grasa.

Cuadro 5. Valores de los *Loadings* (cargas) en los tres primeros componentes principales.

Variable	PC1	PC2	PC3
Humedad	0.101642	0.583163	0.805971
Cenizas	0.282709	0.759833	-0.585432
Grasa	0.953805	-0.287360	0.087634

La representación de los scores y loading en una figura bidimensional (biplot) es presentada en la Figura 13. Los valores de esta representación se obtuvieron del análisis de los componentes principales y nos ayuda a observar cual variable esta mejor asociada con las muestras. En este caso, la distribución de las muestras de café de diferente origen geográfico y las variables medidas en el laboratorio (grasa, humedad y ceniza), nos indican que la grasa es la variable más asociada a las muestras de café.

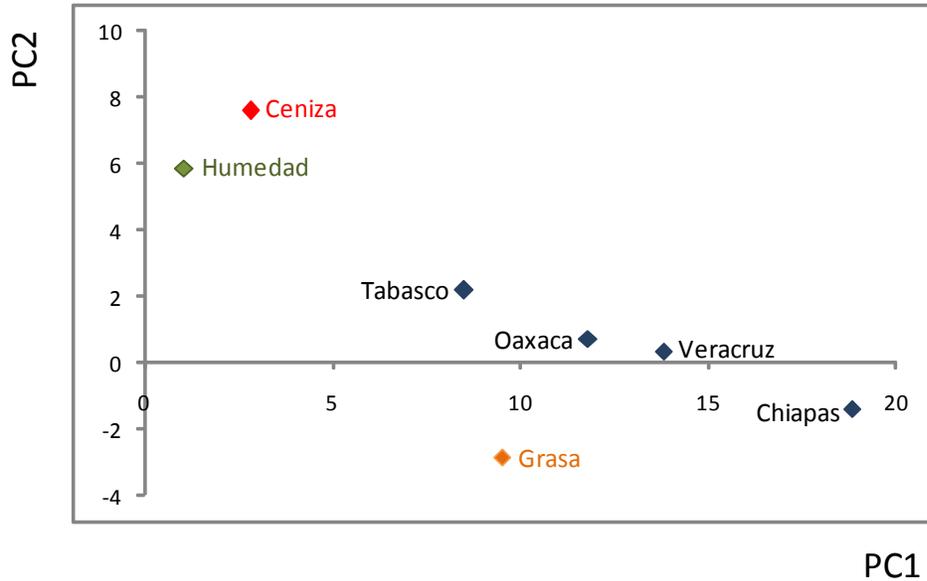


Figura 13. Representación en el plano de los componentes principales (*biplot*) de las variables humedad, ceniza, grasa y las muestras de café de diferente origen geográfico.

1.3.3. Análisis de componentes principales (PCA) en NIRS

El primer análisis quimiométrico de la señal espectral de las 120 muestras analizadas con la tecnología NIRS, fue realizada con técnicas no supervisadas, como es el análisis de componentes principales (PCA); las cuales, se aplicaron con objeto de identificar las estructuras latentes que no es posible observar en una lectura directa de la matriz de datos (120x1050). El objetivo de esta técnica consiste en identificar el menor número de componentes principales que explique la mayor parte de la varianza de las observaciones de modo que los residuales sean los más pequeños posibles.

En el Cuadro 6, se presentan los resultados de las dos primeras componentes principales, los cuales suman el 99.88% de la varianza explicada. Se incluyen la varianza explicada, el porcentaje de varianza explicada y el porcentaje de varianza acumulada, del PC1 y PC2.

Cuadro 6. Análisis de componentes principales.

	Varianza	% Varianza explicada	% Varianza acumulada
PC1	48375.58	99.35	99.35
PC2	254.84	0.53	99.88

La máxima variabilidad existente (PC1), nos indica que los resultados son apropiados para expresar la variabilidad de los datos e indican que todas las muestras son similares entre sí.

En la Figura 14, podemos observar la representación de los dos primeros componentes expresando la máxima variabilidad entre las muestras. Se observa perfectamente la separación de las muestras de café de Tabasco, así mismo, para las muestras de café de Chiapas, siendo las muestras de Oaxaca y Veracruz las que presentan mayor distribución de los datos. Sin embargo y bajo estas condiciones, se puede apreciar muy bien las delimitación de las fronteras entre cada lote de café.

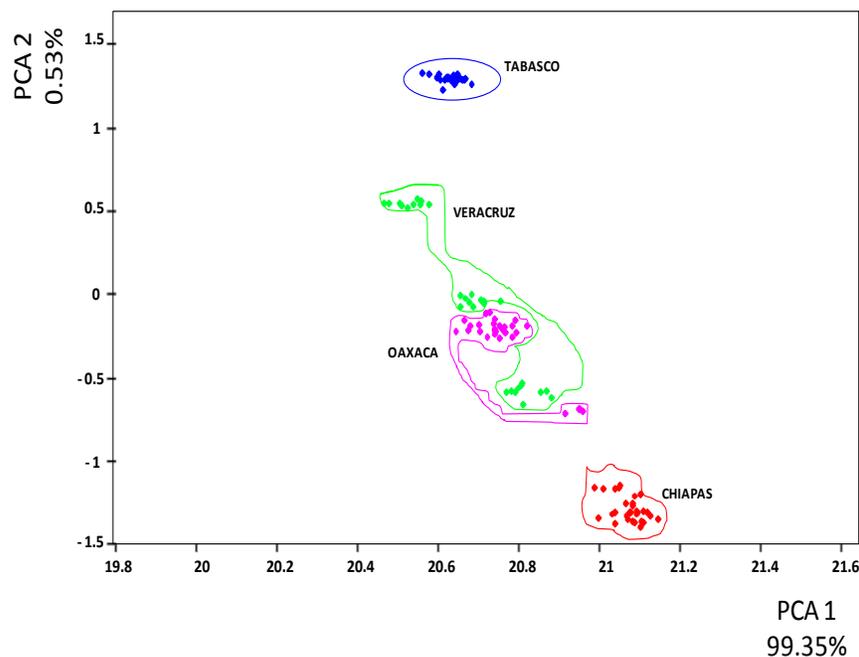


Figura 14. Gráfico de los dos primeros componentes principales para NIRS.

1.3.4. Modelado independiente de clase (SIMCA) para muestras de café

La técnica aplicada en el análisis de los datos generados con el NIRS es la del modelado independiente de clases, la cual mediante el uso del análisis de componentes principales, nos permite construir el diagrama de Coomans construido con el modelado de clases SIMCA, para cada una de las clases que se obtuvieron mediante PCA, tal como se muestra en la Figura 15, en donde podemos observar se generan seis diagramas de Coomans en sus diferentes combinaciones de clases de café comercial analizadas.

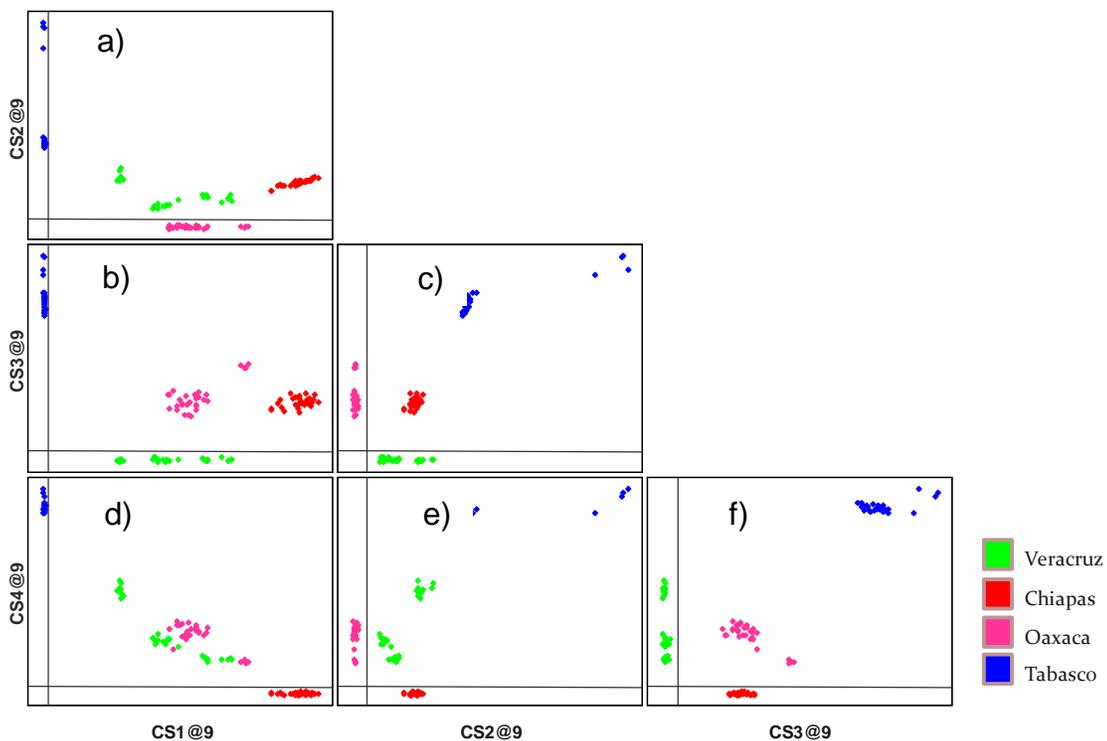


Figura 15. Diagrama de Coomans construido con el modelado de clases SIMCA, en muestras de café comercial.

SIMCA nos permite clasificar materiales semejantes en grupos definidos y además discriminar muestras desconocidas (Beebe *et al.*, 1998). En el diagrama (a) observamos que al comparar la clase 1, correspondiente al café de Tabasco, con la clase 2, correspondiente al café de Oaxaca, cada una de estos lotes de muestras se ubican en sus respectivas área de clase y lo que respecta a las muestras de Veracruz y Chiapas, que no presenta semejanzas ni

similitud con ninguna de las muestras de café mencionadas anteriormente se ubican en el cuadrante que representa a las muestras que no pertenecen ni a Tabasco ni a Oaxaca.

Para el diagrama (b), (c), (d), (e) y (f) se representa el mismo comportamiento que el diagrama (a), los diferentes lotes de café son ubicados en la clase correspondiente, con esto podemos observar que las muestras son ubicadas en sus diferentes aéreas de acoplamiento. Esto nos indica que la tecnología NIRS genera información útil y suficiente para clasificar a las distintas variedades de café, de acuerdo a su origen.

1.3.5. Espectros NIRS de las variedades de café de Córdoba

Con los datos obtenidos de las lecturas de las 30 muestras de café, el equipo generó un total de 1050 datos o variables por muestras; por tanto, se obtuvo una matriz conformada por 31500 datos, que al ser graficados representan la huella espectral característica de cada muestra. En la Figura 16, se puede apreciar la diferenciación espectral entre los lotes de muestras de café (Colombia, Costa Rica-95 y Oro Azteca).

Se tomaron todos los datos obtenidos de los espectros generados por NIRS, creando una base de datos en el programa estadístico Pirouette Chemometrics, para la realización de un análisis de componentes principales (PCA). Técnica que permite reducir la dimensionalidad del conjunto de datos multivariantes. Los datos obtenidos indican que el componente principal uno (PC1) explica el 99.80 % de la varianza total. Por ello, fue utilizado para obtener el coeficiente de variación de cada lote, mismo que explica la homogeneidad de las muestras (Figura 16).

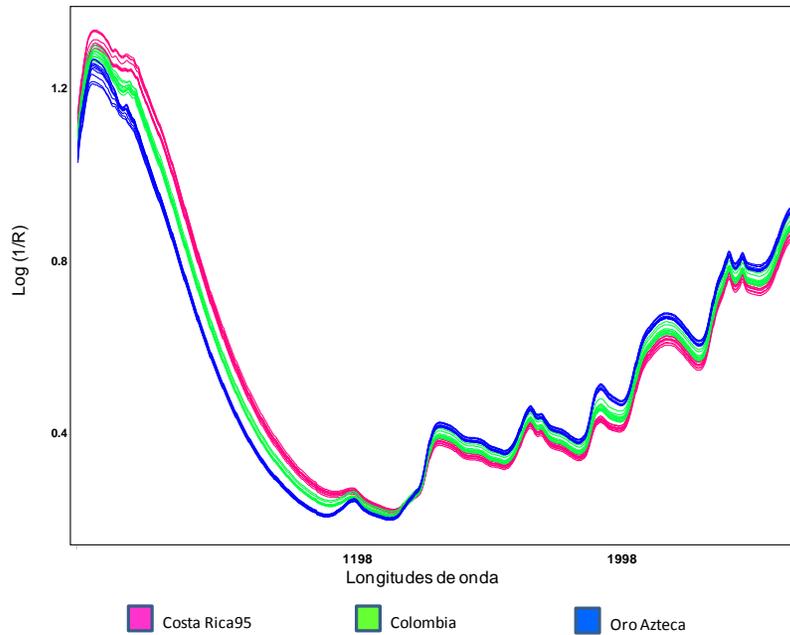


Figura 16. Espectro NIRS observado en el café de Córdoba, Veracruz.

Las huellas espectrales por muestras obtenidas con la tecnología NIRS, fueron convertidas a variables con los programas WINISI y Excel. Con estos datos fue realizado un análisis de componentes principales para visualizar el comportamiento de los datos generados por los espectros NIRS.

Con los datos obtenidos observamos que PC1 explica la máxima varianza, mientras que la posición del PC2, queda fijada por la condición de ortogonalidad con el primer componente. El componente principal uno (PC1) explica el 99.80% de la varianza total. Por ello, fue utilizado para obtener el coeficiente de variación de cada lote. En la figura 17, se representa la distribución de las variedades de café en los dos primeros componentes. Los valores obtenidos en el coeficiente de variación se reflejan en la distribución que tiene cada variedad de café, siendo el PC1 el componente que explica casi el 100% de la varianza de los datos (Cuadro 7).

Cuadro 7. Análisis de componentes principales de las huellas espectrales NIR.

	Varianza	% Varianza explicada	% Varianza acumulada
PC1	12830.943	99.80	99.80
PC2	25.760	0.20	100.00

Para describir la precisión de los resultados analíticos tomamos como base el coeficiente de variación (CV), donde la variedad Oro Azteca representa un lote muy homogéneo (CV=0.78), seguido de Colombia (CV=0.98) y Costa-Rica95 el lote menos homogéneo (CV=1.03). Los valores obtenidos en el coeficiente de variación se reflejan en la distribución que tiene cada grupo de muestras de café (Figura 17).

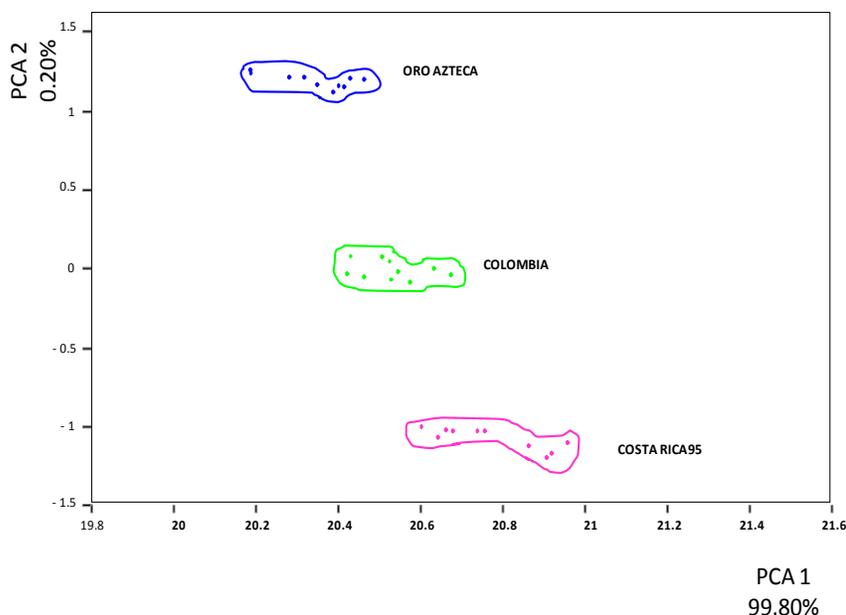


Figura 17. Representación de las huellas espectrales (NIRS) en los dos primeros componentes principales.

1.3.6. Análisis de componentes principales (PCA) de los análisis Físicoquímicos

Con los resultados de los análisis físicoquímicos (humedad, ceniza y grasa) fue realizado un análisis de componentes principales para las variedades de café de Córdoba, Veracruz (Oro Azteca, Colombia y Costa-Rica95). Los datos obtenidos indican que el componente principal uno (PC1) explica el 99.84 % de la varianza total. Por ello, fue utilizado para obtener el

coeficiente de variación de cada lote, mismo que explica la homogeneidad de las muestras. Con el coeficiente de variación (CV) observamos que las muestras presentan diferente comportamiento en cuanto a la homogeneidad entre cada lote, siendo Costa Rica-95 el lote de muestras más homogéneo (CV=2.39), seguido de Colombia (CV=9.67) y Oro Azteca el grupo menos homogéneo (CV=10.35). En la Figura 18, se representa la distribución de las muestras en los dos primeros componentes principales. Los valores obtenidos en el coeficiente de variación se reflejan en la distribución que tiene cada grupo de muestras de café; siendo en el PC1 el componente que explica casi el 100% de la varianza de los datos. Mientras que el CP2, solo se ve reflejado la menor homogeneidad de la variedad de café Oro Azteca (Cuadro 8).

Cuadro 8. Análisis de componentes principales de los análisis fisicoquímicos de las muestras de café de Córdoba, Veracruz.

	Varianza	% Varianza explicada	% Varianza acumulada
PC1	5832.851	99.84	99.84
PC2	9.4012	0.16	100.00

Para describir la precisión de los resultados analíticos tomamos como base el coeficiente de variación (CV), donde la variedad Costa-Rica95 representa un lote muy homogéneo con un CV de 2.39 comparado con las variedades Colombia y Oro Azteca.

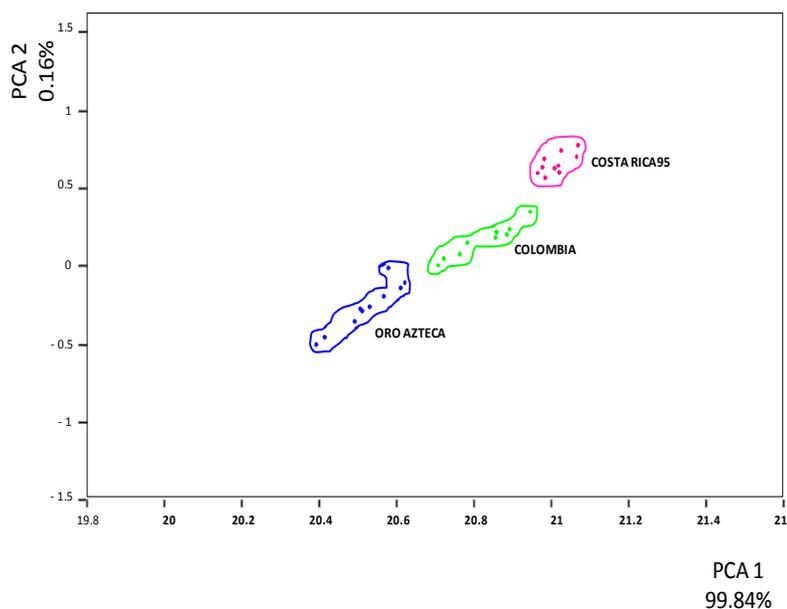


Figura 18. Componentes principales para los análisis fisicoquímicos en muestras de café de Córdoba, Veracruz.

1.3.7. Modelado independiente de clase SIMCA para muestras de café de Córdoba, Veracruz.

La técnica aplicada en el análisis de los datos generados con el NIRS es la del modelado independiente de clases, la cual mediante el uso del análisis de componentes principales, nos permitió construir el diagrama de Coomans construido con el modelado de clases SIMCA, para cada una de las clases que se obtuvieron mediante el PCA, tal como se muestra en la Figura 19, en donde podemos observar se generan tres diagramas de Coomans en sus diferentes combinaciones de clases de café comercial analizados.

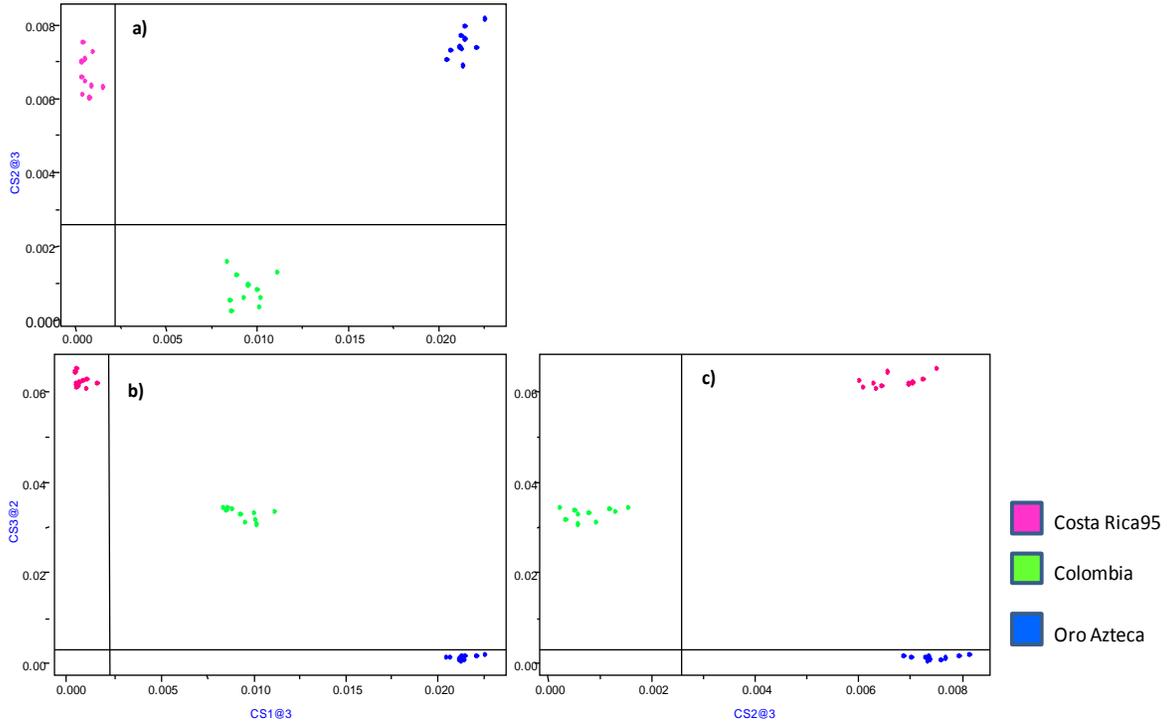


Figura 19. Diagrama de Coomans construido con el modelado de clases SIMCA, en muestras de café comercial de Córdoba, Veracruz.

En el diagrama (a) observamos que al comparar la clase 1, correspondiente a la variedad Costa Rica-95, con la clase 2, correspondiente a la variedad Colombia, cada una de estas variedades se ubica en sus respectiva área de clase y lo que respecta a la variedad Oro azteca, que no presenta semejanzas ni similitud con ninguna de las variedades de café mencionadas anteriormente se ubican en el cuadrante que representa a las muestras que no pertenecen a Costa Rica95 y Colombia.

Para el diagrama (b) y (c), se representa el mismo comportamiento que el diagrama (a), las variedades de café son ubicadas en la clase correspondiente, con esto podemos observar que las muestras son ubicadas perfectamente bien por NIRS en sus diferentes aéreas de acoplamiento. Esto nos indica que la tecnología NIRS permite clasificar perfectamente bien las distintas variedades de acuerdo con su origen.

1.4. CONCLUSIONES

La espectroscopia de infrarrojo cercano (NIRS), es una herramienta analítica no destructiva, que nos permite realizar estudios de homogeneidad de lotes, trazabilidad, y tipificación de café comercial con resultados similares a los obtenidos con los métodos convencionales establecidos en las Normas Mexicanas. Con la ventaja de ser una técnica más rápida, menos costosa y con menor porcentaje de error que los métodos convencionales. Esto lo hace más relevante porque se reduce el uso de reactivos, ahorra tiempo, dinero y es una técnica no destructiva.

Con la interpretación quimiométrica y el análisis de los espectros NIRs fue posible diferenciar a las muestras de café de acuerdo con su origen; así como establecer modelos de clasificación (SIMCA). Adicionalmente, fueron realizadas las mismas pruebas quimiométricas a los espectros NIR de las muestras de café de Córdoba, las cuales estaban integradas por tres variedades; los resultados nos indican que con los espectros es posible diferenciar las variedades Oro Azteca-Colombia-CostaRica 95 así como obtener su modelo de clasificación. Los resultados pueden ser aplicados a nivel industrial para obtener como producto final un café de calidad.

1.5. LITERATURA CITADA

AGRO_Revista Industrial del Campo. 2000. Sistemas de trazabilidad: Eficiencia y competitividad para la industria de alimentos (en línea). Cortesía Alimentaria México 2010. Consultado 27 junio 2011. Disponible en: <http://www.2000agro.com.mx/logistica/sistemas-de-trazabilidad-eficiencia-y-competitividad-para-la-industria-de-alimentos/>.

Arminda M. Bruno-Soares, Ian Murray, Rhona M. Paterson, Jose M. F. Abreu. 1998. Use of near infrared reflectance spectroscopy (NIRS) for the prediction of the chemical composition and nutritional attributes of green crop cereals. *Animal Feed Science and Technology* 75:15-25.

Bai, Y., Lu, C., & Li, C. (2005). The traceability system management of livestock product safety. *Food Science* 8:473–477.

Beebe, K. R., Pell, R. J., and Seasholtz, B. M. 1998. *Chemometrics: A practical guide*. New York, Chichester, Weinheim, Brisbane, Singapore, Toronto.

Caswell, J. (2006). Quality assurance, information tracking, and consumer labeling. *Marine Pollution Bulletin* 53(10-12):650–656.

Dardenne, P., Andrieu, J., Barriere, Y., Biston, R., Demarquilly, C., Femenias, N., Lila, M., Maupetit, P., Riviere, F., Ronsin, T., 1993. Composition and nutritive value of whole maize plants fed fresh to sheep. II. Prediction of the in vivo organic matter digestibility. *Ann. Zootech.* 42(3):251-270.

Dardenne, P., Sinnaeve, G., Biston, R., Lecomte, Ph., 1991. Evaluation of NIRS for predicting fresh forage quality. In: Murray, I., Cowe I.A. (Eds.), *Making Light Work: Advances in NIR Spectroscopy*, VCH. Weinheim, pp. 269-271.

- De la Roza D., B., y Argamentería G., A. 2009. Importancia del control de calidad, trazabilidad y seguridad en las raciones completas (UNIFEED) para ganado vacuno lechero, Aplicación de la tecnología NIRS. Tecnología Agro alimentaria. Boletín Informativo del SERIDA, 2da época. No. 6.
- Givens, D. I., Barker, C.W., Moss, A. R., Adamson, A.H., 1991. A comparison of near infrared reflectance spectroscopy with three in vitro techniques to predict the digestibility in vivo of untreated and ammonia treated cereal straws. *Animal Feed Sci. Technology* 35:83-94.
- Wu, J.,G. and Shi. C.,H. 2007. Calibration model optimization for rice cooking characteristics by near infrared reflectance spectroscopy (NIRS). *Food Chemistry* 103:1054-1061.
- Koutsoumanis, K., Taoukis, P. S., & Nychas, G. J. E. (2005). Development of a safety monitoring and assurance system for chilled food products. *International Journal of Food Microbiology* 100(1–3):253–260.
- Li, X., & Yang, J. (2006). Pig safety traceability system. *Chinese Swine* 2:41–44.
- McMeekin, T. A., Baranyi, J., Bowman, J., Dalgaard, P., Kirk, M., Ross, T. (2006). Information systems in food safety management. *International Journal of Food Microbiology* 112(3):181–194.
- Menchu, J., F., Ibarra, E. (1968). The Chemical composition and the quality of Guatemalan coffee. *Agronomía* 3(7):5-17.
- Miraglia, M., Berdal, K. G., Breram, C., Corbisier, P., Holst-jensen, A., Kok, E. J. (2004). Detection and traceability of genetically modified organisms in the food production chain. *Food and Chemical Toxicology* 42(7):1157–1180.

- Murray, I., Garrido, A., 1990. NIR spectral characteristics of ammonia treated and untreated cereal straws. Proc. 3rd Int. Conference on Near Infrared Spectroscopy Brussels, Belgium, pp. 610-615.
- NMX-F-013-SCFI-2000. Café puro tostado, en grano o molido, sin descafeinar o descafeinado- especificaciones y métodos de prueba. Secretaria de Comercio y Fomento Industrial. Dirección General de Normas.
- Pirouette. 2007. Comprehensive Chemometrics modeligs Software, Infometrix, Versión 3.11, Woodinville, WA.
- Prieto, N., Andres, S., Giraldez, F., J., Mantecón, A., R., Lavín, P. 2006. Potential use of near infrared reflectance spectroscopy (NIRS) for the estimation of chemical composition of oxen meat samples. *Meat Science* 74:487-496.
- Ramis R. G., García A-C. M. C. 2001. Quimiometría. Ed. SINTESIS. Madrid, España. P 238.
- Shenk, J. S., Westerbaas, M. O., 1994. The application of near infrared reflectance spectroscopy (NIRS) to forage analysis. In: Fashey, G. C., Jr. (Ed.), *Forage Quality, Evaluation, and Utilization*, Madison, WI, pp. 406-449.
- Sykuta, M. (2005). Agricultural organization in an era of traceability. *Journal of Agricultural and Applied Economics* 37(2):365–377.
- Waters, C. J., Givens, D. I., 1992. Nitrogen degradability of fresh herbage: Effect of maturity and growth type and prediction from chemical composition and by near infrared reflectance spectroscopy. *Animal Feed Science Technology* 38(4):335-349.
- WinISI III. 2005. Infrasoft International, Copyright[©] by LLC. Version 1.60.
- Yang, L. (2005). The eel quality traceability system based on EAN_UCC. *World Standardization and Quality Management* 3: 3–45.

CAPITULO II: HOMOGENEIDAD DE MUESTRAS DE CAFÉ TOSTADO Y MOLIDO MEDIANTE TECNOLOGIA NIRS

HOMOGENEIDAD DE MUESTRAS DE CAFÉ TOSTADO Y MOLIDO MEDIANTE TECNOLOGIA NIRS

RESUMEN

La composición química del café tostado está fuertemente relacionada con las variedades, las condiciones ecológicas de las regiones productoras y el manejo agronómico; la cual comprende una serie de factores que se unen y producen un tipo o una calidad de café; diferencias de calidad que son reconocidas comercialmente; es por ello que el desarrollo de métodos rápidos en análisis de materias primas, productos semielaborados y terminados han sido uno de los principales objetivos de la industria alimentaria en las últimas décadas. Los métodos analíticos, establecidos en las Normas Oficiales y en los métodos convencionales reportados en la literatura, se caracterizan por ser largos, tediosos, costo elevado y con un alto error analítico debido a la manipulación excesiva de la muestra. Por tanto, la química analítica está orientada a desarrollar estrategias basadas en el análisis instrumental y su interpretación quimiométrica; mismos que son económicos, rápidos y precisos, con menor error analítico capaz de vigilar en tiempo real los procesos industriales ofreciendo a los consumidores productos de mejor calidad y alto valor agregado. En el presente trabajo se utilizó la tecnología NIRS y la Quimiometría para evaluar la homogeneidad del café comercial de diferentes orígenes geográficos; los resultados obtenidos con el método propuesto son similares a los generados con los métodos convencionales de la Norma Oficial Mexicana, con la diferencia de que los análisis con la tecnología NIRS son más rápidos, precisos y no destruyen la muestra.

Palabras claves: *Café, Calidad, Homogeneidad, NIRS, Quimiometría.*

ABSTRACT

The chemical composition of roasted coffee is strongly related to varieties, environmental conditions of the producing regions and crop management, which includes a number of factors that come together and produce a type or quality of coffee quality differences are commercially recognized, that is why the development of rapid methods in the analysis of raw materials, semi finished and finished products have been one of the main objectives of the food industry in recent decades. The analytical methods established in the Official Laws and the conventional methods reported in the literature, are characterized by long, tedious, expensive and highly analytical error due to excessive manipulation of the sample. Therefore, analytical chemistry is focused on developing strategies based on instrumental analysis and interpretation of chemometric; them that are affordable, fast and accurate analytical error less able to monitor real-time industrial processes to offer consumers better products quality and high added value. In this study we used NIRS and chemometrics to assess the homogeneity of commercial coffee from different geographical origins, the results obtained with this method are similar to those generated with conventional methods of NOM, with the difference that the analysis NIRS technology is faster, accurate and not destroy the sample.

Keywords: *Coffee, Quality, Homogeneity, NIRS, Chemometrics.*

2.1. INTRODUCCIÓN

Los métodos utilizados para evaluar la calidad del café comercial se basan en técnicas analíticas convencionales que se caracterizan por altos costos y consumir mucho tiempo y reactivos; además de un elevado porcentaje de error y destruir la muestra (Blanco and Cerdá, 1988). Algunos de los descriptores químicos utilizados para caracterizar variedades de café son: contenido de metales (Martín *et al.*, 1999), los componentes volátiles (Bicchi *et al.*, 1997), los ácidos cloragénicos y la cafeína (Martín *et al.*, 1998), los perfiles de ácidos grasos (Martín *et al.*, 2001), los tocoferoles y los triglicéridos (González *et al.*, 2001). Sin embargo, a pesar de los buenos resultados que proporcionan estos enfoques, es importante tener en cuenta que muchos de los métodos de análisis usados como referencias para determinar los descriptores químicos significativos pueden ser de un tiempo largo de preparación. En la actualidad, la química analítica nos proporciona como alternativa a los métodos convencionales los métodos instrumentales modernos, los cuales, presentan un gran interés en procesos industriales y en la realización de investigaciones por su velocidad en los análisis. Algunas de las técnicas basadas en este tipo de generación de señales del conjunto de componentes de una muestra son: espectroscopia en el infrarrojo cercano (NIRS) (Wang *et al.*, 1995, Macho and Larrechi., 2002, Zhang *et al.*, 2002, González *et al.*, 2003) e infrarrojo medio (MIR) (Carvalho, 2005), espectroscopia ultravioleta-visible (Thurston *et al.*, 2004) y pirolisis acoplada a espectrometría de masas (Goodacre and Kell., 1996, Radovic *et al.*, 2001, Pérés *et al.*, 2002), entre otros.

La espectroscopia de infrarrojo cercano (NIRS), es una técnica analítica no destructiva (Chang *et al.*, 2001). Las razones por la que NIRS está siendo adoptado como el método analítico preferido en muchos laboratorios está relacionada con: mínima preparación de la muestra, análisis rápido, varios constituyentes pueden ser analizados simultáneamente, no destrucción de muestras, bajo costo, no usa productos químicos y los resultados son más precisos (Guerrero *et al.*, 2007), ha sido relevante en áreas de alimentos, como en el desarrollo y aplicación de la espectrometría de reflectancia en el infrarrojo cercano para la caracterización y modelos de calibración de germoplasmas en *Solanum phureja* (papas cultivadas), para identificar el alto contenido de carotenoides y la evaluación de variedades actuales para el

mejoramiento nutricional de la papa (Bonierbale *et al.*, 2009), así como la determinación de arsénico en pimiento rojo (Moros *et al.*, 2008), la calidad fisicoquímica del camote (Guoquan *et al.*, 2006), predicción fisicoquímica de carne (Tejerina *et al.*, 2009).

Es por ello, que es fácil comprender la enorme importancia de la disponibilidad de los métodos analíticos adecuado combinados con herramientas quimiométricas apropiadas, a fin de determinar si existe homogeneidad entre diferentes tipos de café (Esteban-Díez *et al.*, 2004). Una de las formas para verificar si las muestras son uniformes y que sean representativas es mediante un análisis de homogeneidad. Este estudio puede llevarse a cabo de dos maneras: entre botes y dentro de los botes (Van der Veen, Linsinger y Pauwels, 2001). Los resultados de homogeneidad pueden ser representados mediante diagrama de cajas y bigotes. Según Ramis (2001), los diagramas de cajas y bigotes son un ensayo gráfico importante y útil para evaluar la homogeneidad de varias series de resultados. En la actualidad no existen reportes sobre la determinación de homogeneidad del café de diferentes procedencias mediante NIRS, por tanto el objetivo del presente trabajo fue realizar un análisis del café utilizando la tecnología NIRS y la Quimiometría para evaluar la homogeneidad del café comercial y con los resultados obtenidos con el método propuesto verificar si son similares a los generados con los métodos convencionales establecidos en la Norma Oficial Mexicana NMX-F-013-SCFI-2000 para café tostado y molido.

2.2. MATERIALES Y METODOS

En esta segunda etapa se realizó un estudio de homogeneidad del café comercial mediante métodos instrumentales y quimiometría.

El estudio se llevó a cabo con 120 muestras de café tostado y molido, con las cuales se realizó la determinación de humedad, ceniza y grasa con los métodos de referencias establecidos en la NMX-F-013-SCFI-2000. Las muestras fueron identificadas y registradas con un número consecuente de laboratorio y fueron clasificadas, según el lugar de procedencia: Chiapas, Oaxaca, Veracruz y Tabasco. Cada lote fue integrado por 30 muestras de café.

2.2.1. Localización de las muestras de café comercial

Las muestras fueron identificadas y registradas con un número consecutivo de laboratorio y fueron clasificadas, según el lugar de procedencia; es decir Chiapas, Oaxaca, Veracruz y Tabasco.

Cuadro 9. Procedencia de muestras de café comercial.

Estado	Procedencia	Coordenadas geográficas	No. de muestra
Chiapas	Tziscaco	Latitud: 16°45'57.86''N Longitud: 93°8'56.22''O Altitud: 1500 msnm	30
Oaxaca	Ixtepec	Latitud: 16°33'40.52''N Longitud: 95°5'50.63''O Altitud: 160msnm	30
Veracruz	Córdoba	Latitud: 18°53'39.00''N Longitud: 96°56'5.00''O Altitud: 680 msnm	30
Tabasco	Huimanguillo	Latitud: 17°49'49.00''N Longitud: 93°23'29.00''O Altitud: 30 msnm	30

2.2.2. Relación de los métodos físico – químicos

Los métodos convencionales que se realizaron son los descritos en la Norma Mexicana NMX-F-013-SCFI-2000 para análisis físico-químicos de café tostado y molido, los cuales se describen a continuación:

Cuadro 10. Métodos de la Norma Mexicana NMX-F-013-SCFI-2000 para café tostado y molido.

Parámetros	Especificaciones	Método de prueba
Humedad (método gravimétrico)	6.0% máximo	NMX-F-013-SCFI-2000. Inciso 11.6
Cenizas (método de calcinación mediante una mufla)	5.0% bs máximo	NMX-F-013-SCFI-2000. Inciso 11.7
Grasa como extracto etéreo (método Soxhlet)	8.0% bs mínimo-18.0% bs máximo	NMX-F-013-SCFI-2000. Inciso 11.8

Para la realización de los análisis fisicoquímicos, las muestras se utilizaron molidas y tamizadas, las cuales fueron colocadas en bolsas herméticas para su conservación.

2.2.3. Interpretación quimiométrica de los métodos físico-químicos

Los datos que se obtuvieron de los análisis físico-químicos y de las lecturas con el instrumento NIRS fueron analizados quimiométricamente utilizando los programas ISIScan (para la obtención de los espectros de las muestras), WinISI (se utilizó para transformar en variables de escala los espectros obtenidos), Excel (se utilizó como programa puente entre WinISI y el programa Pirouette, así como para la realización de una prueba de estadística descriptiva de los análisis físico-químicos), Statistica (Data Analysis Software System), versión 6 StstSoft, Inc. 2003, (se utilizó para realizar los diagramas de cajas y bigotes) y el Software Pirouette (para el análisis de componentes principales-PCA sin transformaciones y sin preprocesado), se obtuvo el ANOVA de un factor tanto para los resultados de los análisis fisicoquímicos y para PCA.

2.2.4. Espectrofotómetro de infrarrojo cercano (NIRS)

El equipo que se utilizó para generar las huellas espectrales fue un espectrofotómetro de reflectancia en el infrarrojo cercano modelo Versatile Agri Analyzer™6500 System II Vertical Type 461008, Serial No. 9503 100-240 Vac max 200 W, 50-60 Hz de la empresa FOSS NIRSystems. El instrumento registra el espectro de 400 a 2500 nm en longitudes de onda pares (Figura 20).



Figura 20. Espectroscopio de infrarrojo cercano, NIRs Versatile Agri Analyzer™ 6500.

Para realizar la lectura, el equipo cuenta con una lámpara de 4 Volts, ajustada a una placa prealineada, que manda el haz de luz blanca la cual contiene todas las frecuencias al *Grating*, una superficie holográfica de 570 líneas por mm y una distancia de 1.75 μm entre líneas, el cual desvía la luz en ángulos diferentes dependiendo de la longitud de onda; ésta luz reflejada es monocromada y el ángulo de inclinación del *Grating* es controlado por la tarjeta madre al filtro *Order sorter*, que esta sincronizado con el movimiento del *Grating* y es movido por medio de un motor de pasos, el cual permite el paso de la luz por longitud de onda para transmitirla a la muestra, contenida en una copa o celda opaca de quartzo, y esta a los detectores de Silicio (400 a 1100 nm) y PbS (1100 a 2500 nm), que registran la cantidad de luz reflejada y/o transmitida, para transformarla en una señal que se registra mediante el programa ISIScan (versión 2.81 Copyright 2006) a una computadora (Acer AMD Sempron™ Processor 3200+ 1.81 GHz, 448 MB de RAM).

2.2.5. Interpretación quimiométrica de los espectros NIRS

Los programas utilizados para realizar la interpretación quimiométrica de las huellas espectrales fueron ISISCAN (versión 2081 copyright 2006, WinISI III y Excel, como programas puente; y el Software Pirouette (Chemometrics Software versión 3.11 para Windows de Infometrix Inc. Copyright 2007).

2.3. RESULTADOS Y DISCUSIÓN

2.3.1. Descriptores estadísticos de las propiedades fisicoquímicas de café tostado y molido

Los resultados promedio y desviación estándar de los análisis fisicoquímicos del café (humedad, ceniza y grasa) establecidos en la NMX-F-013-SCFI-2000 se expresan en % en peso; y son presentados en el Cuadro 11. Los valores obtenidos indican que las muestras presentan uniformidad entre lotes (n=120).

Cuadro 11. Media y desviación estándar de datos obtenidos en el análisis físico-químico de las muestras de café.

Muestras	Humedad (%)		Ceniza (%)		Grasa (%)	
Chiapas	1.19	0.19 b	4.20	0.07 ab	18.4	2.95 a
Oaxaca	1.55	0.17 ab	3.89	0.05 c	11.0	2.11 c
Veracruz	1.57	0.13 ab	4.15	0.12 b	13.1	1.88 b
Tabasco	1.91	1.57 a	4.23	0.10 a	7.45	0.42 d

SD= desviación estándar

Los resultados se expresan como media \pm desviación estándar de 120 muestras de café tostado y molido de diferente procedencia; realizando un análisis estadístico para evaluar la conformidad de la distribución de las medias.

De acuerdo con los parámetros especificados en la norma NMX-F-013-SCFI-2000 para café tostado y molido (H=humedad 6.0% máximo, C=cenizas 5.0% base seca máximo, G=grasa 8.0% base seca mínimo-18.0% base seca máximo), en las tres variables evaluadas, las muestras de Veracruz y Oaxaca cumplen con lo establecido, pero se observan diferencias significativas (Tukey, 0.05) entre el valor promedio de las medias. Las muestras procedentes de Tabasco no cumplen con el valor mínimo de grasa y las de Chiapas se exceden.

2.3.2. Prueba de homogeneidad mediante análisis físicos y químicos

2.3.2.1. Análisis de humedad

En la variable humedad se presenta en la Figura 21. De acuerdo con la NMX-F-013-SCFI-2000 para café tostado y molido la concentración máxima de humedad en el café debe ser 6.0%, por lo tanto observamos que para las muestras de diferentes procedencia (Tabasco, Chiapas, Oaxaca y Veracruz) cumplen con lo establecido en dicha norma, una humedad excesiva puede perjudicar la conservación del producto.

Los valores para las muestras de Chiapas y Tabasco presentan simetría ubicando la mediana al centro de las cajas, además indican mayor dispersión de los datos. En el café de Tabasco observamos la presencia de cuatro muestras cuyos contenidos de humedad están fuera del

intervalo general de este grupo, lo que comúnmente se conoce como *outliers*; estas muestras corresponden a los siguientes valores: 1525 (5.46% de H), 1523 (5.80% de H), 1516 (5.96% de H) y 1510 (6.10% de H), lo que hace que las caja sea más heterogéneas con respecto a la mediana de los datos. A diferencias de las muestras de Oaxaca y Veracruz, presentan mayor homogeneidad en sus datos con respecto a su mediana.

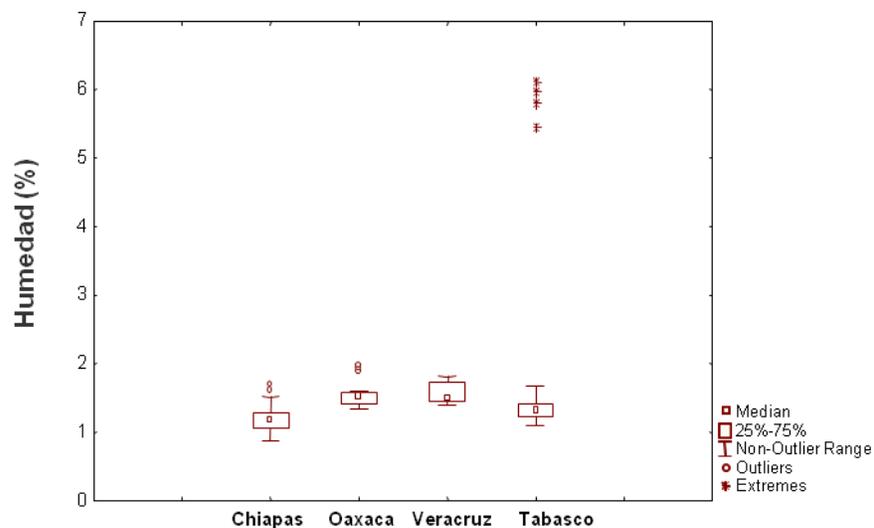


Figura 21. Contenido de humedad en muestras de café tostado y molido.

Con base en el rango intercuartílico, el ordenamiento de las muestras en una secuencia de mayor a menor amplitud es: Chiapas, Tabasco, Veracruz y Oaxaca. Esto nos indica que el grado de homogeneidad de las muestras con base en el valor de humedad es: Oaxaca, Veracruz, Tabasco y Chiapas.

2.3.2.2. Análisis de cenizas

Para realizar el análisis estadístico de los valores de ceniza determinados en las muestras de café, fueron utilizados los diagramas de cajas y bigotes (*Box-plot*). Observamos que las muestras de Tabasco y Veracruz, presentan en el análisis de ceniza, una mayor dispersión de los datos con respecto a su mediana, comparadas con las muestras de Oaxaca y Chiapas. La amplitud de las cajas y bigotes de los datos de ceniza en las muestras Tabasco y Veracruz, indica que son los grupos de muestras más heterogéneas. En el caso de las muestras de Oaxaca y Chiapas la amplitud de la caja y el tamaño de los bigotes es menor comparada con las muestras de Veracruz, esto indica que son datos más homogéneos en ambas (Figura 2.2).

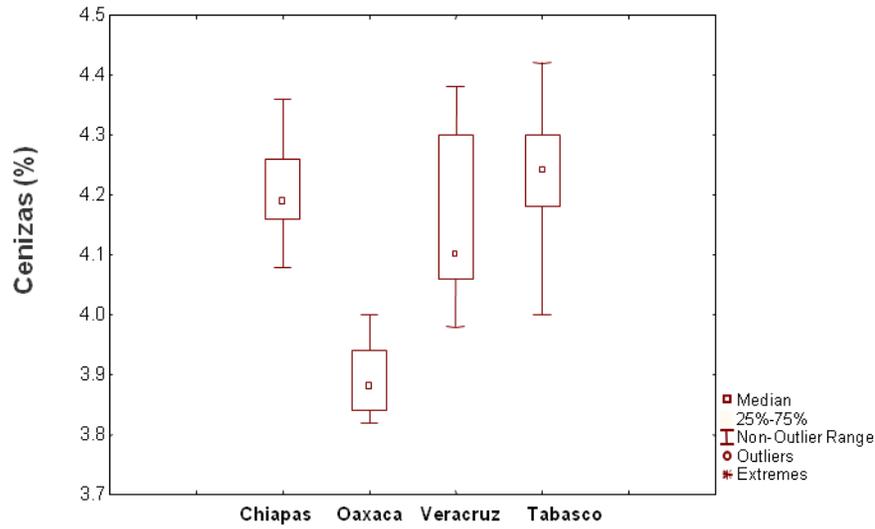


Figura 22. Contenido de ceniza en muestras de café tostado y molido.

2.3.2.3. Análisis de grasa

La descripción estadística para el análisis de grasa, en las muestras de Tabasco se observa menor dispersión y mayor homogeneidad (Figura 23). Los grupos de muestras que presentan mayor variabilidad entre los datos y presentan mayor heterogeneidad, con respecto a su mediana, son Oaxaca y Veracruz, las cuales presentan mayor amplitud en sus cajas y bigotes.

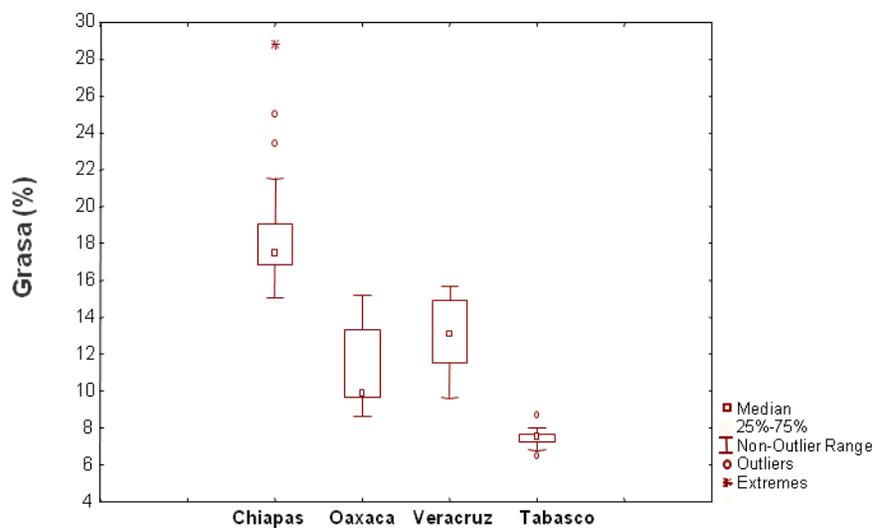


Figura 23. Contenido de grasa en muestras de café tostado y molido.

Por lo anterior, podemos deducir que existe una relación positiva entre la cantidad de grasa en el café con la altitud de las plantaciones. De acuerdo con los datos obtenidos de las alturas de las plantaciones de los cafés de diferente origen geográfico y la media del contenido de grasa (%), existe una correlación positiva ($R^2=0.96$), lo cual indica que entre más altura tengan dichas plantaciones, el contenido de grasa (%) es mayor. Esto comprueba lo establecido por Menchu e Ibarra (1968), quienes indican que el contenido de grasa, cafeína y nitrógeno total aumentan conforme se incrementa la altura a la que se ha producido el café y por ende con la calidad.

2.3.3. Prueba de homogeneidad mediante NIRS

2.3.3.1. Información espectral de las muestras de café tostado y molido

Los espectros de cada procedencia (Chiapas, Oaxaca, Veracruz y Tabasco) se observan en la Figura 24. Este registro de huellas espectrales es comprendido entre los 1100 y 2400 nm.

En los espectros de las diferentes muestras de café se observa que son más homogéneas las muestras procedentes de Chiapas, seguidas de Veracruz y Oaxaca. Se observó que las muestras que presentan mayor heterogeneidad en su comportamiento espectral son las muestras de Tabasco, las cuales están registradas como CH1510, CH1516, CH1523 y CH1525 lo podemos comprobar con la separación en las huellas espectrales, mismas que se presentan como puntos *outliers*, en los diagramas de cajas y bigotes en el parámetro de humedad.

Con la finalidad de asociar las longitudes de onda que se obtuvo en el espectro NIR y la grasa determinada con el método de referencia (Soxhlet); fueron usadas aquellas longitudes que corresponden a grasa en el espectro y que fueron definidas por el programa WINISI. Para ello, se asocio el valor de reflectancia en cada longitud de onda con el de grasa, ambos correspondientes a cada muestra. Los valores de los coeficientes de correlación presentaron un promedio de 0.60, lo cual deduce una baja correlación. Esto puede atribuirse, a que la huella espectral NIR es información global de una muestra, por ello, es poco común utilizar solo una longitud de onda para realizar modelaciones.

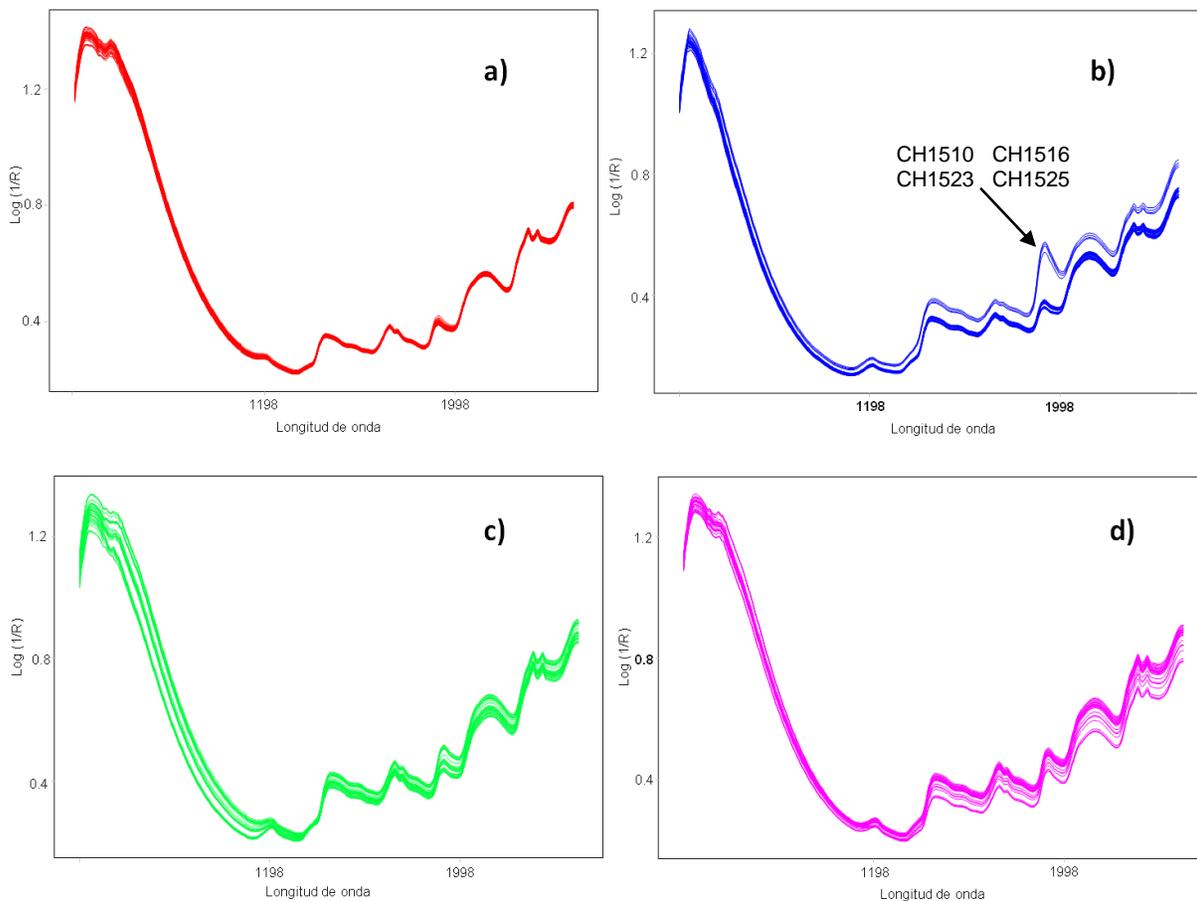


Figura 24. Huellas espectrales de café tostado y molido de: a) Chiapas, b) Tabasco, c) Veracruz y d) Oaxaca.

En los espectros de las diferentes muestras de café tostado (Figura 2.4), se observa que los lotes más homogéneos son los procedentes de Chiapas (a), Veracruz (c) y Oaxaca (d) a diferencia de las muestras del estado de Tabasco (b) que presenta más heterogeneidad en su comportamiento espectral, mostrando una separación de dos grupos de huellas espectrales, estas muestras corresponden a las registradas como CH1510, CH1516, CH1523 y CH1525, y son las muestras que se presentan como *outliers* en los diagramas de cajas y bigotes para el análisis de humedad y en PCA.

En el análisis de componentes principales (PCA) se determinó que la mayor variabilidad es explicada por el primer componente (PC1); obteniendo una varianza acumulada mayor al 99%. De las muestras de café analizadas, las cuales son presentadas en el Cuadro 12, las

muestras procedentes del estado de Tabasco son las que presentan un coeficiente de variación superior a los demás.

Cuadro 12. Media y desviación estándar del CP1 obtenido de las huellas espectrales NIR en las diferentes muestras de café.

Muestras	PCA 1			
	\bar{X}	SD	CV	% varianza acumulada
Chiapas	21.62	± 0.20	0.94	99.99
Oaxaca	20.94	± 0.39	1.86	99.92
Veracruz	20.68	± 0.40	1.91	99.80
Tabasco	16.96	± 0.57	3.35	99.93

En el Cuadro 13, podemos observar que las muestras del estado de Tabasco son las que tienen un valor F_{cal} mayor al resto de grupos 1.98, 1.86, 1.08 y 1.30 en las variables humedad, ceniza, grasa y del PCA de las huellas espectrales NIR.

Cuadro 13. Prueba de ANOVA de un factor para Análisis Fisicoquímicos y NIRS.

Café	Humedad (%)	Ceniza (%)	Grasa (%)	NIRS
	F_{cal}	F_{cal}	F_{cal}	F_{cal}
Chiapas	0.27	0.26	0.48	0.54
Oaxaca	0.37	0.05	0.23	0.30
Veracruz	0.30	0.14	0.54	0.27
Tabasco	1.98	1.86	1.08	1.30

F_{cal} = obtenido de ANOVA

$F_{Tab} (3.35) (P >) (p=0.05)$. $F_{cal} < F_{tab}$ = muestras homogéneas.

En el Cuadro 13, se observa que en la prueba de hipótesis no existe diferencia entre las muestras de café tostado y molido de cada procedencia; ya que la $F_{calculada}$ de todos los tratamientos fue menor que la $F_{tabulada}=3.35$; por lo tanto se acepta la hipótesis nula y se comprueba que existe homogeneidad en cada uno de los lotes de café tostado y molido.

Debido a que el valor de P es mayor al (de 0.05) en todos los casos se acepta la igualdad de varianzas en la muestra y entre muestras de cada lote de café.

2.3.3.2. Cajas y bigotes de PCA mediante NIRS

En la descripción estadística del análisis de PC de los espectros obtenidos mediante NIRS, las muestras de Tabasco se diferencian del conjunto de muestras estudiadas presentando simetría y homogeneidad en sus datos en comparación con la mediana de sus datos, seguidas por las muestras de Chiapas, Veracruz y Oaxaca. En las muestras de Tabasco observamos puntos extremos los cuales se conocen como puntos *outliers*, estas muestras corresponden a: CH1510, CH1516, CH1523 y CH1525 (Figura 25).

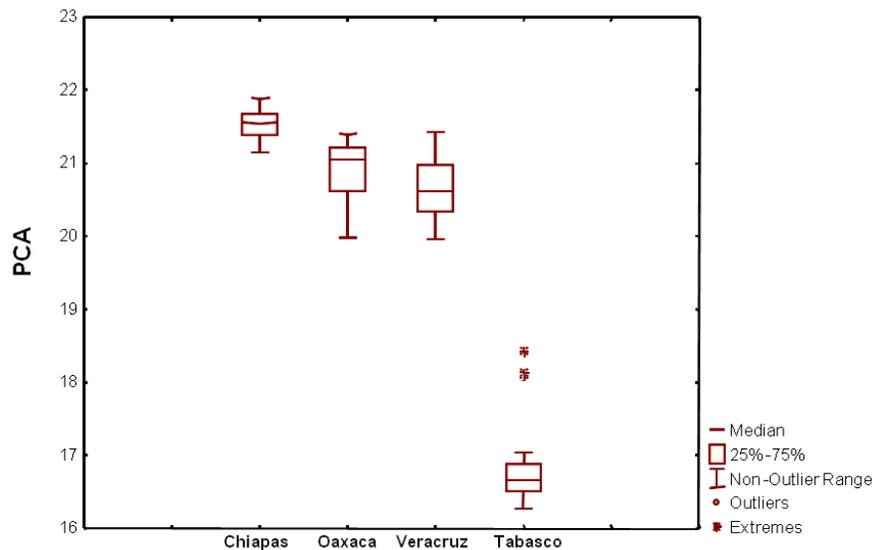


Figura 25. Diagrama de cajas para CP1 en muestras de café tostado y molido

2.4. CONCLUSIONES

Se concluye que con base a las pruebas de homogeneidad todos los lotes son homogéneos, presentado una mayor dispersión en las muestras procedentes del estado de Tabasco en el análisis de ceniza y humedad, son las que presentan mayor variabilidad en el PC1, así como variación en las huellas espectrales NIR; por lo tanto podemos decir que mediante la tecnología NIRS se puede comprobar homogeneidad entre y dentro de los grupos de muestras; así como las diferencias en homogeneidad entre orígenes geográficos.

Mediante la tecnología NIRS y la Quimiometría se estableció una metodología para evaluar la homogeneidad del café comercial; los resultados obtenidos con el método propuesto son similares a los generados con los métodos de referencia de la Norma Mexicana NMX-F-013-SCFI-2000.

2.5. LITERATURA CITADA

- Bicchi C.P., Ombretta M.P., Pellegrino G. and Vanni A.C. 1997. Characterization of roasted coffee and coffee beverages by solid phase microextraction-gas chromatography and principal component analysis. *J. Agric. Food Chem.* 45:4680-4686.
- Bonierbale, M., Grüneberg, W., Amoros, W., Burgos, G., Salas, E., Porras, E, and Felde, T. 2009. Total and individual carotenoid profiles in *Solanum phureja* cultivated potatoes: II. Development and application of near-infrared reflectance spectroscopy (NIRS) calibrations for germplasm characterization. *Journal of food composition and Analysis* 22:509-516.
- Blanco, M. and V. Cerdá. 1988. *Quimiometría*. Publicaciones de la Universidad Autónoma de Barcelona. Bellaterra, España.
- Carvalho A. R., Nogal Sánchez, M., Wattoon, J. and Brereton, R. G. 2005. Comparison of PLS and kinetic models for a second-order reaction as monitored using ultraviolet visible and mid-infrared spectroscopy. *Talanta*, 68 (4):1190-1200.
- Chang, Ch., Laird, D.A., Mausbach, M.J. and Hurburgh, Ch. 2001. Near infrared reflectance spectroscopy principal components regression analyses of soil properties. *Soil Sci. Soc. Am. J.* 65:480-490.
- Estaban-Díez I., Gonzalez-Saiz J.M and Pizarro C. 2004. An evaluation of orthogonal signal correction methods for the characterisation of *arabica* and *robusta* coffee varieties by NIRS. *Analytica Chimica Acta* 514: 57–67.
- Goodacre R. and Kell, D. B. 1996. Correction of mass spectral drift using artificial neural networks. *Analytical Chemistry* 68:271-280.

- González A.G., Pablos F., Martin M.J., Leon-Camacho M and Valdenebro M.S. 2001. HPLC analysis of tocopherols and triglycerides in coffee and their use as authentication parameters. *Food Chemistry*. 73:93–101.
- González-Martín I., C. González Pérez, J. Hernández Méndez, and N. Álvarez-García. 2003. Determination of fatty acids in the subcutaneous fase of Iberian breed swine by near infrared spectroscopy (NIRS) with a fibre-optic probe. *Meat Science* 65:713-719.
- Guerrero, C., Mataix-Solera, J., Arcenteguir, V, Mataix-Beneyto, J and Gómez, I. 2007. Near-infrared spectroscopy to estimate the maximum temperatures reached on Burned soils. *Soil Sci. Soc. Am. J.* 71(3):1029-1037.
- Guoquan L., Huang, H., Zhang, D. 2006. A Prediction of sweetpotato starch physiochemical quality and pasting properties using near-infrared reflectance spectroscopy. *Food Chemistry* 94:632-639.
- Macho S. and M. S. Larrechi. 2002. Near-infrared spectroscopy and multivariate calibration for the quantitative determination of certain properties in the petrochemical industry. *Trends in Analytical Chemistry* 21:799-806.
- Martin M. J., Pablos F and A. G. González A. G. 1999. Characterization of *arabica* and *robusta* roasted coffee varieties and mixture resolution according to their metal content. *Food Chem.* 66: 365–370.
- Martin M. J., Pablos F and A. G. González A. G. 1998. Discrimination between *arabica* and *robusta* green coffee varieties according to their chemical composition. *Talanta* 46:1259–1264.
- Martin M. J., Pablos F., A. G. González A. G., Valdenebro M. S. and Leon-Camacho M. 2001. Fatty acid profiles as discriminant parameters for coffee varieties differentiation. *Talanta* 54:291–297.

- Menchu, J., F., Ibarra, E. (1968). The Chemical composition and the quality of Guatemalan coffee. *Agronomía* 3(7):5-17.
- Moros J., Llorca, I., Cervera, M. L., Pastor, A., Garrigues, S., and M. de la Guardia. 2008. Chemiometric determination of arsenic and lead in untreated powdered red paprika by diffuse reflectance near – infrared spectroscopy. *Analytica Chimica Acta* 613: 196 - 206.
- NMX-F-013-SCFI-2000. Café puro tostado, en grano o molido, sin descafeinar o descafeinado- especificaciones y métodos de prueba. Secretaria de Comercio y Fomento Industrial. Dirección General de Norma.
- Pérès, C. C. Viallon, and J. L. Berdagué. 2002. Curie point pyrolysis-mass spectrometry applied to rapid characterization of cheeses. *Journal of Analytical and Applied Pyrolysis* 65:161-171.
- Pirouette. 2007. Comprehensive Chemometrics modeligs Software, Infometrix, Versión 3.11, Woodinville, WA.
- Radovic B. S., R. Goodacre, and E. Anklam. 2001. Contributions of pyrolysis-mass spectrometry (Py-MS) to authenticity testing of honey. *Journal of Analytical Applied Pyrolysis* 60:79-87.
- Ramis R. G., García A-C. M. C. 2001. *Quimiometría*. Ed. SINTESIS. Madrid, España. P 238.
- Tejerina D., López-Parra, M.M., García-Torres, S. 2009. Potential used of infrared reflectance spectroscopy to predict meat physico-chemical composition of guinea fowl (*Numida meleagris*) reared under different production systems. *Food Chemistry* 113: 1290 – 1296.

Thurston T. J., R. G. Brereton, D. J. Foord, and R. E. A. Escott. 2004. Principal components plots for exploratory investigation of reactions using ultraviolet-visible spectroscopy: application to the formation of benzophenone phenylhydrazone. *Talanta* 63:757-769.

Van der Veen, A. M. H., Linsinger, T., & Pauwels, J. (2001). Uncertainty calculations in the certification of reference materials. 2. Homogeneity study. *Accreditation and Quality Assurance*. 6: 26–30.

Wang Z., T. Dean, and B. R. Kowalski. 1995. Additive background correction in multivariate instrument standardization. *Analytical Chemistry* 67:2379-2385.

WinISI III. 2005. Infrasoft International, Copyright[©] by LLC. Version 1.60.

Zhang L., C. W. Small, and M. A. Arnold. 2002. Calibration standardization algorithm for partial least-squares regression: application to the determination of physiological levels of glucose by near-infrared spectroscopy. *Analytical Chemistry* 74:4097-4108.

CAPITULO III: CONCLUSIONES GENERALES DE LA INVESTIGACIÓN

3.1. CONCLUSIONES GENERALES

De acuerdo con el objetivo de la investigación, evaluar la capacidad de los métodos instrumentales para generar información útil y suficiente para la caracterización y tipificación del café comercial, se concluye que:

En los métodos convencionales para los diferentes tipos de café comercial, se observaron variaciones significativas en los diferentes parámetros evaluados (humedad, ceniza y grasa) al ser comparados con los valores establecidos en la NMX-F-013-SCFI-2000 para café tostado y molido. Las mediciones convencionales realizadas y las interpretaciones quimiométricas fueron suficientes para la tipificación del café comercial de acuerdo con la procedencia.

El uso de técnicas multivariantes como lo es el análisis de componentes principales (PCA), demostró que con las señales espectrales generadas con la tecnología NIRS, permiten visualizar la distribución de las muestras en el espacio de los dos primeros componentes, permitiendo encontrar la estructura relevante en una tabla de datos cuya lectura directa lo impide. Lo cual, facilita la exploración de los datos de los diferentes tipos de café.

La técnica de modelado independiente de analogías de clase (SIMCA), es una herramienta útil en la clasificación de café tostado y molido (café comercial) en grupos definidos, discriminando muestras desconocidas, esta técnica nos permitió encontrar los ejes de los componentes en los espacios de las mediciones multivariantes para cada categoría, indicando a qué grupo pertenece cada muestra de café según su origen. Los espectros NIRS fueron información útil y suficiente para tipificar las muestras de café comercial de acuerdo con la procedencia.

Por lo tanto, se concluye que la espectroscopia de infrarrojo cercano (NIRS) y la quimiometría pueden aplicarse al estudio de café comercial para evaluar su calidad, tipificación y caracterización; con resultados complementarios a los obtenidos con los métodos convencionales, con la ventaja de ser una técnica no destructiva, con menor consumo de reactivos y tiempo, así como un menor porcentaje de error analítico.