DEPARTAMENTO DE POSGRADOS

MAESTRÍA EN GESTIÓN DE LA CALIDAD Y SEGURIDAD ALIMENTARIA

"DETECCIÓN DE ADULTERACIÓN POR AGUADOS EN LECHE CRUDA MEDIANTE ESPECTROSCOPÍA DE INFRARROJO MEDIO PARA EL CANTÓN CUENCA"

TRABAJO DE GRADUACIÓN PREVIO A LA OBTENCIÓN DEL TÍTULO "MAGÍSTER EN GESTIÓN DE LA CALIDAD Y SEGURIDAD ALIMENTARIA"

AUTOR: ING. MARÍA SUSANA CAMPOVERDE MALDONADO.

DIRECTOR: DR. PIERCÓSIMO TRIPALDI

CUENCA, ECUADOR

2013

DEDICATORIA

A mis padres y hermanos quienes siempre me han brindado su apoyo incondicional.

A mi esposo e hijos quienes son las personas más importantes de mi vida.

A mis amigos quienes me han apoyado en todo momento.

AGRADECIMIENTO

Agradezco al Dr. Piercósimo Tripaldi quien me ha guiado y ayudado a culminar con éxito el presente trabajo.

RESUMEN

El presente estudio tiene como objetivo detectar la presencia de adulterantes (agua y sueros fisiológicos) en la leche natural mediante Espectroscopía de infrarrojo FT-MIR-ATR. Para el efecto se utilizaron 24 fuentes de leche natural, las cuales fueron sometidas a adulteraciones conocidas a nivel de laboratorio, generando 98 muestras. Se obtuvieron los espectros FT-MIR-ATR en la región de 750 a 1800 cm^{-1}. Para la discriminación entre la leche natural y la leche adulterada se utilizó el método de clasificación KNN en combinación con el método de algoritmos genéticos. Este modelo fue evaluado mediante una validación externa. Las longitudes de onda seleccionadas y los pre-tratamientos de los datos investigados son reportados en este estudio. Los resultados demuestran que FT-MIR-ATR puede ser utilizado para detectar agua y sueros fisiológicos en muestras de leche natural.

PALABRAS CLAVE:

FT-MIR-ATR, adulterantes de leche, suero fisiológico

ABSTRACT

The goal of the present study is to detect the presence of possible adulterants (water and saline solutions) .-in raw milk through the use of FT-MIR-ATR infrared spectroscopy. For this purpose, we employed 24 sources of raw milk, which underwent laboratory adulterations and we generated 98 samples. We were able to obtain FT-MIR-ATR spectrums in the 750 to 1800 cm-1. In order to differentiate raw milk from adulterated milk we employed the KNN classification method in combination with the genetic algorithm method. This model was evaluated through an external validation. The wavelength and the pre-treatments of the research data are reported in this study. The results show that FT-MIR-ATR can be employed to detect the presence of water and saline solutions in samples of raw milk.

Key Words: FT-MIR-ATR, milk adulterants, saline solution

Translated by,
Diana Lee Rodas

UNIVERSIDAD DE
AZUAY
DPTO. IDIOMAS

INDICE DE CONTENIDO

Página.

Contenido

DEDICATORIA ii

AGRADECIMIENTO iii

RESUMEN Y PALABRAS CLAVE.....iv

ABSTRAC AND KEYWORDS……..……..v

INDICE DE CONTENIDO.....vi

INDICE DE FIGURAS.....viii

INDICE DE TABLAS…….ix

INTRODUCCIÓN……..……1

Espectroscopía de infrarrojo…….....2

Espectroscopía infrarrojo con transformadoras de Fourier (FT- IR)……2

Espectroscopía infrarroja en el rango medio (MIR)……3

Reflectancia total atenuada (ATR)……..3

Método de Crioscopía……..4

OBJETIVO GENERAL……..5

OBJETIVOS ESPECÍFICOS……5

CAPÍTULO I……6

MATERIALES Y MÉTODOS……6

1.1. Localización del Estudio……6

1.2. Origen de las Muestras……..6

1.3 Preparación de Adulterantes……6

1.3.1 Adulteración con Agua y Sueros Fisiológicos……7

1.4 Análisis de las Muestras……7

1.5 Análisis por el Método de Crioscopía……...……...7

1.6 Análisis por Espectroscopía de Infrarrojo FT-MIR……7

1.6.1 Pre-procesamiento de Espectros……..8

1.7 Análisis Estadísticos……10
1.7.1 Método de Clasificación KNN (K- Nearest Neighbors)……10
1.7.2 Método de Algoritmos Genéticos……10
1.7.3 Validación Cruzada……11
1.7.4 Validación Externa……11
CAPÍTULO II……12
RESULTADOS……12
2.1 Método de Crioscopía……12
2.2 Método de Espectroscopia de Infrarrojo (FT-MIR-ATR)……13
2.3 Diagrama de Flujo para aplicación del método de detección de adulteración mediante FT-MIR……21
CAPÍTULO III……22
DISCUSIÓN……22
CONCLUSIONES……25
REFERENCIAS BIBLIOGRÁFICAS……26
REFERENCIAS ELECTRÓNICAS……28
ANEXOS……30

INDICE DE FIGURAS

Página.

Figura 1. Esquema de un instrumento FTIR 2

Figura 2. Esquema representativo de la técnica de muestreo FTIR: ATR 4

Figura 3. Espectros FT-MIR-ATR obtenidos para una muestra de agua destilada, leche natural, leche adulterada con el 10% de agua y leche adulterada con el 15% de solución de sacarosa en la región de 750 a 1800cm^{-1} 14

Figura 4. Calidad del modelo en función del número de variables 15

Figura 5. Espectro de la leche con las regiones de sus componentes mayoritarios 17

Figura 6. Espectros FT-MIR-ATR obtenidos para una muestra leche adulterada con el 20% de solución fisiológica de cloruro de sodio y el espectro de su segunda derivada, en la región de 750 a 1800 cm^{-1} con las longitudes de onda seleccionadas por el algoritmo genético 18

Figura 7. Asignación de clases mediante el método KNN con validación cruzada 19

INDICE DE TABLAS

Página.

Tabla 1. Resultados del Método de Crioscopía para Leche cruda-1 y leche cruda-2 adulteradas con el 5% de diferentes tipos de adulterantes. ..... ……12

Tabla 2. Resultados del Método de Crioscopía para Leche-cruda-1-1 y leche cruda 2, adulteradas con el 10% de diferentes tipos de adulterantes. ..... …12

Tabla 3. Resultados del Método de Crioscopía para Leche-cruda-1-1, adulterada con el 15% de diferentes tipos de adulterantes. 13

Tabla 4. Resultados de Método de Crioscopía para Leche-cruda-1-2, adulterada con el 20% de diferentes tipos de adulterantes. 13

Tabla 5. Condiciones de los métodos KNN y algoritmos genéticos programados en Matlab 7 .. 14

Tabla 6. Bandas seleccionadas por el algoritmo genético..... 15

Tabla 7. Parámetros de Calidad del Modelo para el Training Set 15

Tabla 8. Resultados del training set..... 16

Tabla 9. Resultados del test set..... 17

Tabla 10. Matriz de Confusión de 4 variables para el Training Set 18

Tabla 11. Parámetros de Calidad del modelo de clasificación por clases para el training set.... 19

Tabla 12. Parámetros de calidad del modelo para el conjunto de validación externa..... 20

Tabla 13. Matriz de Confusión para el conjunto de validación externa 20

Campoverde Maldonado María Susana

Trabajo de graduación

Dr. Piercosimo Tripaldi

Abril, 2013

"DETECCIÓN DE ADULTERACIÓN POR AGUADOS EN LECHE CRUDA MEDIANTE ESPECTROSCOPÍA DE INFRARROJO MEDIO PARA EL CANTÓN CUENCA"

INTRODUCCIÓN

La adulteración de la leche natural de vaca es la dilución de ésta con un elemento menos costoso (generalmente agua o suero fisiológico), este problema se agudiza en la industria láctea por la falta de conciencia de los proveedores, falta de políticas de control de la leche natural, el tiempo que implica la determinación de la adulteración y sobre todo la falta de tecnología in situ, todos estos agravantes hacen que este fraude se prolifere sin control.

La calidad y seguridad alimentaria por la adulteración de la leche cruda se ve afectada debido a que se reduce su valor nutricional y se corre el riesgo de introducir peligros significativos por microorganismos provenientes de aguas contaminadas que ponen en riesgo la salud del consumidor final, siendo prioritario para las entidades reguladoras y de investigación determinar nuevas técnicas rápidas y fiables para su control.

La adulteración de la leche natural con sueros fisiológicos es importante debido a que no se puede determinar por el método crioscópico por la razón de que el punto de congelación es función de la concentración de las moléculas disueltas en la leche cruda (lactosa, cloruros y sales minerales), y el suero fisiológico es una disolución que contiene la misma concentración de solutos que la leche natural, manteniéndose el equilibrio en la disolución final, por lo tanto el punto de congelación de la leche adulterada con sueros fisiológicos no varía. (Amiot J. *et al*., 1991).

Por lo anteriormente expuesto el objetivo del presente estudio es evaluar el uso de la espectroscopía de infrarrojo con transformadoras de Fourier en la región media y reflectancia total atenuada (ATR), para la detección de la adulteración de la leche cruda por sueros fisiológicos en el cantón Cuenca.

Espectroscopía de infrarrojo:

La espectroscopía de infrarrojo es una técnica que estudia la interacción de las ondas electromagnéticas con la materia, por lo tanto cuando la radiación infrarroja incide sobre una muestra es capaz de provocar cambios en los movimientos vibracionales de las moléculas induciendo absorción por parte de la muestra que depende del tipo de enlaces y grupos funcionales presentes en la misma, por lo que estas bandas de absorción han sido usadas para análisis cualitativo y cuantitativo de numerosas moléculas debido a la identificación y atribución de estas bandas a grupos químicos específicos que dan información específica del producto investigado. (Stuart y Barbara, 2004).

Los alimentos están compuestos por moléculas complejas resultantes de la polimerización de monómeros como por ejemplo carbohidratos, aminoácidos, etc., estas moléculas contienen grupos químicos específicos como las funciones carboxilo, amino y otros de los cuales se conoce que cada grupo químico absorbe radiación en la región infrarroja permitiendo identificar claramente las características de las bandas de absorción de estos grupos en la región del infrarrojo cercano y medio, se conoce que la mayoría de las bandas de absorción han sido identificadas y atribuidas a grupos químicos específicos para la región media. (Da-Wen Sun, 2009).

Espectroscopía infrarrojo con transformadoras de Fourier (FT- IR):

Instrumentación FTIR:

Los principales componentes de un hardware FTIR son: la fuente (IR source), el haz separador (Beam splitter), el interferómetro, el detector (IR detector) y un laser de referencia (Laser).

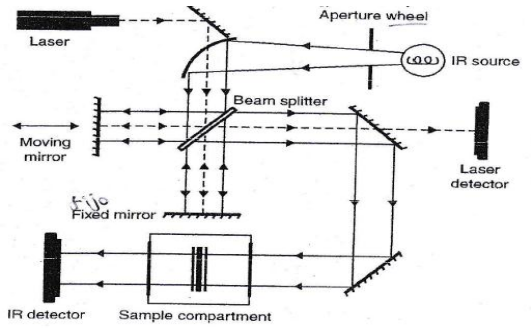


Figura 1. Esquema de un instrumento FTIR

Fuente: Da-Wen Sun, 2009.

La luz de la fuente pasa a través de una abertura (Aperture Wheel) e impacta a un espejo que la dirige hacia el interferómetro el cual consta de un haz separador (Beam splitter) que divide la luz IR en dos. Este desdoblador tiene la propiedad de trasmitir la mitad de la radiación y reflejar la otra mitad, el haz transmitido y el reflejado se dirigen a los 2 espejos (fijo y móvil) orientados perpendicularmente a ambos haces, estos espejos están alineados con la finalidad de reflejar las ondas de luz en una dirección que permiten la recombinación de las 2 ondas (transmitida y reflejada) en el haz separador. La luz recombinada del interferómetro es dirigida por los espejos hacia el compartimiento de la muestra y es reconocida por el detector que la convierte en señal eléctrica y finalmente es procesada por un sistema de computación para obtener los datos espectrales. (Da-Wen Sun, 2009)

El interferómetro nos proporciona como respuesta el interferograma que es una función en el espacio de las frecuencias el cual necesita ser convertido a un espectro mediante una transformada de Fourier (procedimiento matemático) para obtener una función en el espacio de las amplitudes (espectro). Esencialmente las FT descomponen el interferograma provisto por el interferómetro en ondas sinusoidales para cada longitud de onda, estas ondas sinusoidales son arregladas sobre las longitudes de onda produciendo el espectro convencional. (155).

Espectroscopía infrarroja en el rango medio (MIR):

La espectroscopía infrarrojo está dividida en 3 regiones: Infrarrojo cercano (13333 - 4000cm^{-1}), infrarrojo medio (4000 - 400cm^{-1}) e infrarrojo lejano (400 - 10 cm^{-1}).

La espectroscopía MIR involucra el estudio de la absorción de los componentes en el rango MIR (4000 – 400 cm^{-1}) del espectro electromagnético. La espectroscopía FT-MIR monitorea la vibración fundamental y el estrechamiento rotacional de las moléculas las cuales producen perfiles químicos de la muestra. La región MIR es muy robusta y reproducible por lo que pequeñas diferencias en la composición de las muestras pueden ser medidas de manera confiable. Cuando las moléculas absorben la energía MIR muestran movimientos como: estiramientos, doblamientos, torsión, balanceo y movimientos de tijeras en uno o más localizaciones en el espectro que dependen de varios factores como por ejemplo la configuración del enlace, su localización, etc., proporcionando información para el análisis de la composición y la estructura química de las moléculas. (Griffiths P. y Haseth J., cop. 2007).

Reflectancia total atenuada (ATR):

Cuando un haz IR va desde un medio de alto índice refractivo (cristal ZnSe) a un medio de bajo índice refractivo (muestra) una cantidad de luz es reflejada. A cierto ángulo de incidencia casi todas las ondas de luz son reflejadas, a este fenómeno se conoce como reflexión total interna, en esta condición alguna cantidad de energía de luz escapa del cristal y extiende una pequeña distancia (0,1 a 5 µm) más allá de la superficie en forma de ondas. Esta onda invisible es

llamada onda fugaz. La intensidad de la luz reflejada reducida a este punto se conoce como reflectancia total atenuada; es decir cuando la muestra es aplicada sobre el cristal ATR, una cierta cantidad de radiación IR penetra más allá del cristal y es absorbida por la muestra, esta absorbancia es traducida al espectro IR de la muestra. (Da-Wen Sun, 2009).

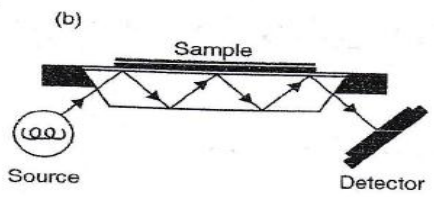


Figura 2. Esquema representativo de la técnica de muestreo FTIR: ATR
Fuente: Da-Wen Sun, 2009.

Método de Crioscopía:

La leche es un sistema coloidal constituido por una solución acuosa de lactosa (5%), sales y minerales en estado de disolución (0,7%), proteínas en estado de suspensión (3,2%) y materia grasa en estado de emulsión, su punto de congelación es una constante física estable y esto se debe a que su presión osmótica se mantiene en equilibrio con la de la sangre del organismo productor de la misma. El descenso del punto de congelación está en relación directa con la concentración de solutos en una solución y puede oscilar entre -0.52°C y -0.56°C, por lo que variaciones superiores a -0.52°C indican aguado y es detectable a partir del 3%. (Amiot J. *et al*., 1991)

La adulteración por suero fisiológico no es detectada por este método debido a que los solutos contenidos se mantienen en equilibrio con la presión osmótica de la leche y su punto de congelación tiende a permanecer constante.

OBJETIVO GENERAL:

Determinar una técnica de espectroscopía de infrarrojo (FT-MIR-ATR), para detectar adulteración por sueros fisiológicos presentes en la leche natural, para el cantón Cuenca.

OBJETIVOS ESPECÍFICOS:

- Demostrar que el método de crioscopía no detecta la adulteración de la leche con sueros fisiológicos.
- Determinar un modelo de clasificación K.N.N. apropiado que permita diferenciar la leche cruda natural de la leche adulterada con agua o sueros fisiológicos.
- Validar el modelo que permita predecir la adulteración con suero fisiológico en muestras desconocidas.
- Crear una metodología de aplicación para poder determinar si una muestra presenta o no adulteración.

CAPÍTULO I.

MATERIALES Y MÉTODOS

1.1. Localización del Estudio:

El análisis de los espectros IR de las muestras de leche natural se realizó en el laboratorio de química perteneciente a la facultad de Ciencia y Tecnología de la universidad del Azuay en la ciudad de Cuenca.

1.2. Origen de las Muestras:

Las muestras de leche natural fueron recolectadas de las haciendas de las principales parroquias rurales productoras de leche del cantón Cuenca, como se indica a continuación:

- Tarqui
- Victoria del Portete
- Cumbe
- Sayausí
- San Joaquín

Se escogió a 5 proveedores de cada parroquia rural.

El volumen de la muestra fue de 1 a 2 litros de leche cruda por cada proveedor, para obtener una muestra confiable, la toma de la muestra fue inspeccionada para garantizar que no se trata de una muestra adulterada. Las muestras recolectadas fueron almacenadas hasta ser llevadas al laboratorio para realizar su análisis.

1.3 Preparación de Adulterantes

Los adulterantes (agua y sueros fisiológicos) [Guidotti M. cop. 2003-2009] utilizados a nivel de laboratorio fueron:

- Agua destilada
- Cloruro de Sodio al 0,9%
- Bicarbonato de Sodio al 1,39%
- Fosfato bisódico de Sodio al 4,62%
- Citrato de Sodio al 3,44%
- Glucosa al 5,51%

- Sacarosa al 9,25%.

Se preparó 1000 cm^3 de cada solución con porcentaje m/m, para ello se utilizó agua destilada y compuestos químicamente puros.

1.3.1 Adulteración con Agua y Sueros Fisiológicos:

En el diseño experimental para la obtención del modelo de clasificación se utilizaron 24 muestras de leche cruda pura y muestras preparadas con los adulterantes indicados (agua y sueros fisiológicos) con 4 niveles de concentración cada uno de ellos. Los niveles de concentración utilizados para cada uno de los adulterantes a nivel de laboratorio fueron: 0%, 5%, 10%, 15% y 20%, estas muestras preparadas fueron homogenizadas previamente antes de su análisis.

1.4 Análisis de las Muestras:

Para la determinación de la adulteración con sueros fisiológicos presentes en la leche natural se utilizaron dos métodos: método crioscópico y el método de espectroscopía de infrarrojo con transformadoras de Fourier, en el rango medio (FT-MIR) con reflectancia total atenuada (ATR).

1.5 Análisis por el Método de Crioscopía:

Para obtener las lecturas de las muestras puras y de las muestras adulteradas en el laboratorio se utilizó un crioscopio marca Funke Gerber modelo Cryo Star 1, de acuerdo a las instrucciones de su manual de usuario.

1.6 Análisis por Espectroscopía de Infrarrojo FT-MIR

Para el análisis de las muestras y obtención de los espectros se utilizó el equipo F-TIR de espectroscopía de infrarrojo en el rango medio (MIR) con reflectancia total atenuada ATR modelo Thermo Nicolet FT-IR 100 con aditamento ATR de seleniuro de zinc (ZnSe), este equipo está conectado a un PC equipado con un software EZ- Omnic. Cada espectro de muestra incluye 16 escáneres en la región de 750 a 1800 cm^{-1}, a un intervalo de 2 cm^{-1}.

Antes de realizar las lecturas de las muestras se realizó un background (blanco) con aire y la cubeta de ATR limpia, se continuó con el análisis de las muestras preparadas que fueron colocadas directamente en la cubeta del ATR en el modo de absorbancia para la obtención de los espectros. La limpieza del cristal ATR después del análisis de cada muestra se realizó con agua destilada y se secó con papel absorbente. Las muestras de leche cruda y las adulteradas a las diferentes concentraciones se homogenizaron previamente y se tomó 3 ml de estas muestras para el análisis en el FT-MIR que fueron colocadas directamente en la celda ATR y los espectros obtenidos fueron guardados en el PC en una carpeta del software Ez-Omnic 32 y el resto se analizó por el método de crioscopía.

Para la recuperación de los datos espectrales se utilizó el operador de datos Omnic 32, el cual permite convertir los espectros obtenidos en valores de una hoja electrónica Excel que contendrá la base de datos para la construcción del modelo.

1.6.1 Pre-procesamiento de Espectros:

El pre tratamiento de los espectros tiene como finalidad mejorar la capacidad de predicción del modelo y eliminar o reducir los efectos o contribuciones no asociadas con la información buscada, las cuales afectan negativamente la precisión y exactitud de las predicciones del modelo, su eficiencia en la reducción del ruido depende de su naturaleza y del método de pre tratamiento usado.

Los métodos de pre procesamiento de datos se dividen en tres categorías:

Métodos para eliminación del ruido y filtrado: estos métodos no tienen supervisión y se aplican antes de la modelización, ayudan a eliminar la influencia de la variación aleatoria sin afectar la señal, por ejemplo: suavización, corrección de la línea base.

Métodos clásicos para la normalización espectral y la diferenciación, por ejemplo: la transformación estándar normal variable, primera y segunda derivada, estos métodos son también sin supervisión y se aplican antes de modelar sin tener en cuenta la respuesta.

Métodos para la selección de variables y reducción de la dimensionalidad eliminando sistemáticamente la variación presente en X que no es relevante para la predicción de Y, estos métodos están basados en modelos e involucran un conjunto de entrenamiento para estimar los parámetros que van a eliminar la variación no deseada. (Trygg J.; Gabrielsson J.; Lundstedt T., cop. 2009).

Los pre-tratamientos aplicados a los espectros FT-MIR-ATR obtenidos en este estudio se detallan a continuación:

Savitzky-Golay, suavización: Se utiliza para reducir la vibración instrumental y consiste en estimar un polinomio de orden reducido utilizando todos los puntos dentro de una ventana en movimiento para estimar el valor del punto central de la misma. El fundamento es que la estimación de la respuesta promedio en el centro del intervalo de análisis tiene menos incertidumbre que el ruido que afecta a los datos. (M. S. Reis, P. M. Saraiva, B. R. Bakshi, cop. 2009)

Baseline Correction: Esta técnica se utiliza para corregir el desplazamiento de la línea base del espectro obtenido que puede ser debido a variaciones instrumentales, para ello se utiliza uno o más puntos de datos en una región espectral exenta de variaciones significativas para calcular la corrección del desplazamiento (un valor constante que es restado de todos los

puntos de la señal), que puede ser la intensidad media de una serie de puntos cercanos. (Zeaiter M. and Rutledge D., cop. 2009)

Differentiation: Los principales objetivos de la derivación son mejorar la resolución pico y eliminar la absorción de fondo. La primera derivada permite que los efectos aditivos constantes de la línea base sean eliminados, la segunda derivada elimina las desviaciones lineales de la línea base y los cambios constantes de la misma. La interpretación geométrica de la derivada de una función es la pendiente de la recta tangente a punto, así la derivada refleja la tasa de cambio f (x) en el punto x. Si un espectro contiene una tendencia lineal, entonces la primera derivada de la tendencia es un valor constante y la segunda derivada es 0. La interpretación geométrica de la derivada segunda de una función está relacionada con la curvatura de la función, un alto valor positivo de la derivada segunda significa una curvatura fuerte hacia arriba. La diferenciación es utilizada para mejorar las variaciones que no son claramente visibles en las señales originales y el trazado de la curvatura que es la segunda derivada puede mejorar en gran medida la resolución espectral. El algoritmo más utilizado es el de Golay S. donde los datos dentro de una ventana móvil son ajustados a un polinomio de grado dado para generar el diferencial de grado elegido. (Taavitsainen V.M., cop. 2009). Para este estudio se aplicó la segunda derivada de los espectros FT-MIR-ATR obtenidos.

Estándar normal Variate (SNV): Esta técnica realiza un auto escalado en cada espectro calculando el promedio y la desviación estándar entre las absorbancias. Se la utiliza para que todas las variables tengan una misma escala y para corregir diferencias en la intensidad global de las señales, resultantes de factores de no interés como: efectos multiplicativos de la dispersión de la luz, tamaño de partícula, variaciones en la penetración de la radiación, creando un cambio en los niveles de absorbancia que pueden hacer difícil la interpretación de los espectros, por lo tanto esta técnica tiene como objetivo dar el mismo peso a cada señal por corrección de las diferencias en la intensidad global. (Zeaiter M. and Rutledge D., cop. 2009). La ecuación básica utilizada es la siguiente:

$$Xsnv = \frac{X - X\ avg}{Xsd}$$

En donde:

X_{snv} : Es el elemento auto escalado del espectro

X: Es el elemento original del espectro

X_{avg}: Es el promedio de los espectros

X_{sd}: Es la desviación estándar de los espectros

1.7 Análisis estadísticos:

En este estudio se utilizó el método de clasificación KNN en combinación con el método de algoritmos genéticos y una validación externa.

Para crear el modelo de clasificación que nos permite diferenciar entre la leche cruda natural de la leche adulterada con agua o suero fisiológico se utilizó el programa "GA_TOOLBOX" en Matlab implementado por el grupo de investigadores de Milano Chemometrics and QSAR Research Group de la universidad de Milan-Bicocca.

1.7.1 Método de Clasificación KNN (K- Nearest Neighbors):

El método de clasificación KNN es del tipo no lineal que está basado en clasificar un objeto de acuerdo a la clase de objetos más cercanos, para ello es necesario calcular y analizar una matriz de distancias, este método calcula la distancia euclidiana entre los objetos y los resultados los ordena de acuerdo a su valor. Para determinar el valor óptimo de K se utilizó un procedimiento de validación cruzada en el test set, dando valores de K de 1 a 10 y se seleccionó el valor óptimo que tenía el error más bajo de clasificación. (Da-Wen Sun, 2009)

1.7.2 Método de Algoritmos Genéticos:

Este método sirve para escoger las longitudes de onda (variables) que proporcionan la mejor clasificación KNN. Esta técnica de búsqueda global se basa en la teoría de la evolución natural y lo que hace es llevar a cabo una simulación en la computadora de lo que ocurre en la naturaleza, para ello las condiciones experimentales de cada variable puede ser codificadas por una secuencia de 0 y 1 (bits), para generar el cromosoma correspondiente, el cual estará siempre asociado a la respuesta que se desea obtener.

Los algoritmos genéticos constan de tres pasos fundamentales:

1. Creación de la población inicial, 2. Reproducción, 3 Mutaciones

Se utiliza una población inicial de individuos (lecturas de los espectros de las muestras) para explorar el espacio de posibles soluciones, las características personales de cada individuo corresponden a los valores que se ajusta en la optimización del problema que es una cadena de elementos (0,1) denominados "genes". La aptitud de un individuo de sobrevivir y generar descendencia es medida mediante una función objetivo que debe reflejar el índice de desempeño que se maximiza en el problema de optimización

En la reproducción, se forma un grupo de apareamiento tomando parejas de los individuos de la población, los individuos con una mejor respuesta se les da una mayor probabilidad de ser seleccionados para el grupo de apareamiento. Los cromosomas de cada pareja se combinan mediante operadores genéticos para generar descendencia, finalmente una nueva población está formada mediante la sustitución de la generación anterior con su descendencia. Este

procedimiento es iterativo hasta que se satisface un criterio de parada y de esta manera se obtienen las mejores respuestas para las condiciones experimentales dadas.

En el caso de la mutación, el computador simula una mutación con una probabilidad extremadamente baja en la que se cambia a un bit de un cromosoma, a diferencia del crossover en donde se cambian genes (variables), y se evalúa la respuesta promedio de esta generación, si esta es mayor a la respuesta promedio de la primera generación, se seleccionan los mejores cromosomas, luego se repite todo el proceso hasta obtener una nueva generación, estas nuevas generaciones se crearán hasta alcanzar un criterio de parada satisfecha que generalmente es el valor de la respuesta objetivo. (Leardi R., 2009).

1.7.3 Validación Cruzada:

Es una técnica de validación que utiliza solo el conjunto de entrenamiento, en donde los datos se dividen en N conjuntos de entrenamiento y pares del conjunto de predicción, donde N es igual al número de muestras en los datos, cualquier muestra particular sólo estará presente en el conjunto de predicción, después de desarrollar un discriminante a partir de un conjunto de entrenamiento y ponerlas a prueba en el conjunto de predicción correspondiente, en otras palabras, la primera muestra es eliminada de los datos y el entrenamiento se realiza de los demás datos. El discriminante desarrollado a partir de los datos cortados se prueba utilizando la primera muestra que fue eliminada, la primera muestra es retornada entonces al conjunto de datos y la segunda muestra se elimina de los datos, este procedimiento se repite hasta que todas las muestras se han suprimido una vez, la tasa de éxito de clasificación para el clasificador se calcula utilizando las muestras eliminadas como muestras de validación. (Lavine B. K., 2009).

1.7.4 Validación Externa:

La bondad del modelo de clasificación se valida mediante un conjunto de muestras externas, es decir muestras que no se utilizaron para el entrenamiento del modelo; para poder comprobar cómo las muestras de clases conocidas son calculadas por el modelo, entonces las diferencias entre los predichos y los pertenecientes a las clases reales de estas muestras nos darán la tasa de error que se podría esperar cuando este clasificador se utilice para analizar muestras reales desconocidas. (Lavine B.K., 2009)

CAPÍTULO II.

RESULTADOS

Los resultados obtenidos de las metodologías antes indicadas son:

2.1 Método de Crioscopía:

Los resultados obtenidos de la leche natural y de la leche adulterada con sueros fisiológicos a nivel de laboratorio se detallan a continuación:

Tabla 1. Resultados del Método de Crioscopía para leche cruda-1 y leche cruda-2 adulteradas con el 5% de diferentes tipos de adulterantes.

Aguados con el 5% de diferentes tipos de soluciones de sueros fisiológicos	**Para leche cruda 1**		**Para leche cruda 2**	
	% de Adulterante	**Punto Crioscópico en °C**	**% de Adulterante**	**Punto Crioscópico en °C**
Agua destilada	1,84%	-0,5026	2,86%	-0,4974
Cloruro de sodio al 0,9%	0,00%	-0,5304	0,00%	-0,5255
Bicarbonato de sodio al 1,39%	0,00%	-0,5210	0,00%	-0,5204
Fosfato bisódico de sodio al 4,62%	0,00%	-0,5593	0,00%	-0,5579
Citrato de Sodio al 3,44%	0,00%	-0,5409	0,00%	-0,5324
Glucosa al 5,51%	0,00%	-0,5364	0,00%	-0,5297
Sacarosa al 9,25%	0,00%	-0,5311	0,00%	-0,5286

Tabla 2. Resultados del Método de Crioscopía para Leche-cruda-1-1 y leche cruda 2, adulteradas con el 10% de diferentes tipos de adulterantes.

Aguados con el 10% de diferentes tipos de soluciones de sueros fisiológicos	**Para leche cruda 1-1**		**Para leche cruda 2**	
	% de Adulterante	**Punto Crioscópico en °C**	**% de Adulterante**	**Punto Crioscópico en °C**
Agua destilada	7,50%	-0,4736	8,68%	-0,4676
Cloruro de sodio al 0,9%	0,00%	-0,5291	0,00%	-0,5253
Bicarbonato de sodio al 1,39%	0,00%	-0,5198	0,00%	-0,5164
Fosfato bisódico de sodio al 4,62%	0,00%	-0,5969	0,00%	-0,5901
Citrato de Sodio al 3,44%	0,00%	-0,5395	0,00%	-0,5392
Glucosa al 5,51%	0,00%	-0,5402	0,00%	-0,5336
Sacarosa al 9,25%	0,00%	-0,5355	0,00%	-0,5301

Tabla 3. Resultados del Método de Crioscopía para Leche-cruda-1-1, adulterada con el 15% de diferentes tipos de adulterantes.

Aguados con el 15% de diferentes tipos de soluciones de sueros fisiológicos	**Para leche cruda 1-1**	
	% de Adulterante	**Punto Crioscópico en °C**
Agua destilada	12,65%	-0,4472
Cloruro de sodio al 0,9%	0,00%	-0,5317
Bicarbonato de sodio al 1,39%	0,00%	-0,5158
Fosfato bisódico de sodio al 4,62%	0,00%	-0,6279
Citrato de Sodio al 3,44%	0,00%	-0,5566
Glucosa al 5,51%	0,00%	-0,5416
Sacarosa al 9,25%	0,00%	-0,5375

Tabla 4. Resultados de Método de Crioscopía para Leche-cruda-1-2, adulterada con el 20% de diferentes tipos de adulterantes.

Aguados con el 20% de diferentes tipos de soluciones de sueros fisiológicos	**Para leche cruda 1-2**	
	% de Adulterante	**Punto Crioscópico en °C**
Agua destilada	19,51%	-0,4121
Cloruro de sodio al 0,9%	0,00%	-0,5263
Bicarbonato de sodio al 1,39%	0,00%	-0,5121
Fosfato bisódico de sodio al 4,62%	0,00%	-0,657
Citrato de Sodio al 3,44%	0,00%	-0,5494
Glucosa al 5,51%	0,00%	-0,5398
Sacarosa al 9,25%	0,00%	-0,5315

2.2 Método de Espectroscopia de Infrarrojo (FT-MIR-ATR):

Para detectar la adulteración de la leche natural con sueros fisiológicos, se obtuvieron los espectros FT-MIR-ATR, los cuales fueron sometidos a un pre tratamiento (suavización, corrección automática de la línea base y se aplicó la segunda derivada) luego fueron digitalizados y auto escalados; los espectros así obtenidos se presentan en el Anexo No 1.

En la figura 3, se puede observar una comparación de los espectros del agua, leche natural, leche adulterada con el 10% de agua y leche adulterada con el 15% de una solución de sacarosa, obtenidos durante los ensayos de laboratorio.

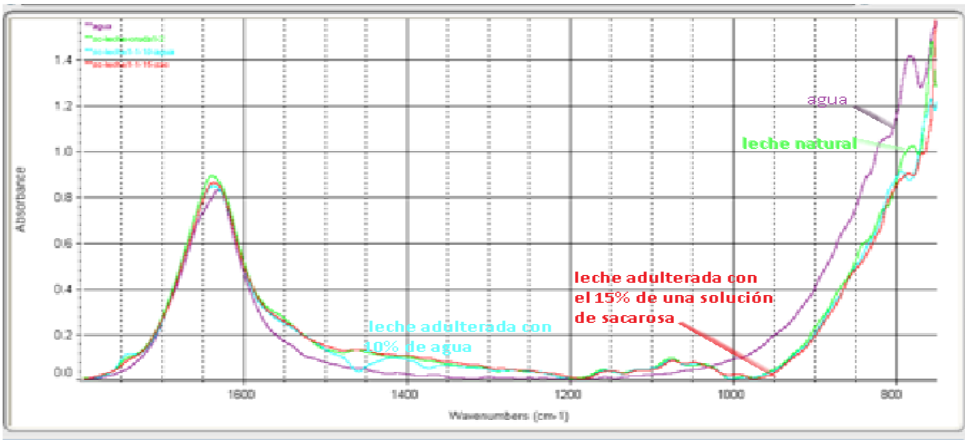


Figura 3. Espectros FT-MIR-ATR obtenidos para una muestra de: agua destilada, leche natural, leche adulterada con el 10% de agua y leche adulterada con el 15% de solución de sacarosa para la región de 750 a 1800cm^{-1}.

Una vez obtenida la base de datos para la construcción del modelo, se utilizó el método de algoritmos genéticos para seleccionar las longitudes de onda que nos permite obtener la mejor clasificación KNN. Para realizar esta selección se utilizó el programa Matlab Vs7, con las condiciones indicadas en la tabla 5.

Tabla 5. Condiciones de los métodos KNN y algoritmos genéticos programados en Matlab 7

method:	'knn'
scal:	'auto'
cv_groups:	10
cv_type:	'vene'
num_chrom:	50
startvar:	5
maxvar:	30
probmut:	0,0100
probcross:	0.5000
runs:	100
num_eval:	150
kernel:	'linear'
class_prob:	2
freq_back:	100
num_windows:	1
dist_type:	'euclidean'
domax:	0

Las mejores longitudes de onda obtenidas por el Algoritmo Genético se presentan en la tabla 6:

Tabla 6. Bandas seleccionadas por el algoritmo genético

N°	Longitud de onda seleccionada. (cm^{-1})
1	1762,6
2	1552,4
3	1429,0
4	1496,5

Los parámetros de calidad del modelo de clasificación obtenido para las longitudes de onda seleccionados se detallan en la tabla 7:

Tabla 7. Parámetros de Calidad del Modelo para el Training Set

N° de objetos en training set	N° variables escogidas	NER CV	N° grupos para cv
79	4	0,911392	10

En donde NER CV: Es la tasa de éxito de clasificación o la tasa de no Error, con validación cruzada para el training set.

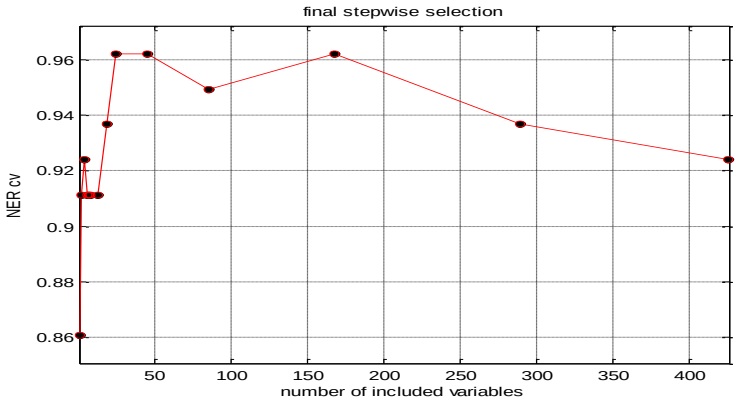


Figura 4. Calidad del modelo en función del número de variables

En la figura 4 se indica la tasa de éxito de clasificación del modelo en relación al número de variables, se puede observar que la tasa de éxito de clasificación disminuye cuando aumenta el número de variables.

En las tablas 8 y 9 se presentan los resultados del algoritmo de clasificación sobre el training set y test set respectivamente.

Tabla 8: Resultados del training set

Nº Progresivo	Muestra	Cm^{-1}				Variables respuesta	
		1762,6	1552,4	1429,0	1496,5		
		1	2	3	4	Clase Verdadera	Clase predecida
1	leche1_1_5_cit	0,9238	-0,8917	-0,5620	-0,8365	2	2
2	leche1_1_5_gluc	0,5801	0,7603	-0,2753	0,4662	2	2
3	leche1_1_10_agua	0,2818	-0,0055	-0,6203	-1,8976	2	2
4	leche1_1_10_bic	-0,4272	0,6290	-1,4389	-0,4862	2	2
5	leche1_1_10_cit	0,2066	0,1110	0,0549	-0,5755	2	2
6	leche1_1_10_fosf	1,2622	0,7521	-0,6793	-2,5696	2	2
7	leche1_1_10_gluc	1,3257	-0,1809	-0,4649	0,4451	2	2
8	leche1_1_10_sac	-2,6054	1,3454	-0,4156	-2,3068	2	2
9	leche1_1_10_sal	0,5759	-0,3717	-0,0864	-0,7261	2	2
10	leche1_1_15_bic	0,8889	-1,4255	0,3804	-0,5300	2	2
11	leche1_1_15_gluc	-0,9832	0,1956	-1,9871	1,8836	2	1
12	leche1_1_15_sac	-0,1957	1,4409	-0,2296	-0,0674	2	2
13	leche1_1_15_sal	-0,0828	0,8773	1,4532	-0,7968	2	2
14	leche1_2_20_agua	0,2672	1,0293	1,0998	-0,7625	2	2
15	leche1_2_20_bic	-0,1377	-0,4294	0,7589	0,2248	2	2
16	leche1_2_20_cit	-0,6201	-1,1035	0,0210	-0,4582	2	2
17	leche1_2_20_gluc	-0,4440	1,4947	2,1863	0,0593	2	2
18	leche1_2_20_sac	0,3791	0,0768	0,7153	-0,7453	2	2
19	leche1_2_20_sal	0,1014	0,4460	0,4434	-0,0677	2	2
20	leche1_5_agua	0,3533	0,1103	1,3774	-0,5011	2	2
21	leche1_5_bic	0,6375	0,4219	0,2598	-0,0102	2	2
22	leche1_5_fosf	0,8132	1,3887	0,2376	0,1774	2	2
23	leche2_5_bic	0,6307	-0,0421	-0,2802	-0,0546	2	2
24	leche2_5_fosf	0,4325	-0,4409	0,9824	-0,5452	2	2
25	leche2_5_gluc	1,2990	1,5002	1,7624	-1,1797	2	2
26	leche2_5_sac	-0,0135	0,9052	0,5087	-0,3279	2	2
27	leche_cruda1	-0,1303	-1,1893	0,3324	-0,3347	1	2
28	leche_cruda1_1	-0,2699	0,0384	-0,5289	0,5318	1	2
29	leche_cruda1_2	0,1000	0,0245	-0,2154	0,5510	1	2
30	leche_cruda_s1	-0,5169	0,7134	0,0752	0,4825	1	2
31	leche_cruda_t1	0,1819	0,3617	1,5172	1,5215	1	2
32	leche_cruda_t2	0,6666	0,2724	0,9596	-0,1073	1	2
33	leche_s1_15_agua	0,5571	0,7581	-0,0020	-0,8167	2	2
34	leche_s1_15_bic	0,4638	0,5445	1,2150	0,5497	2	2
35	leche_s1_15_fosf	-0,1728	0,8159	0,5528	-0,2131	2	2
36	leche_s1_15_gluc	0,6548	0,5037	0,0858	0,1277	2	2
37	leche_s1_15_sac	0,6331	0,6527	-0,0763	-0,2442	2	2
38	leche_s1_15_sal	0,6459	0,4961	0,6267	-0,7356	2	2
39	leche_s1_20_agua	0,8566	1,5372	0,1204	-0,7101	2	2
40	leche_s1_20_cit	0,3241	-0,3567	0,4132	-1,0870	2	2
41	leche_s1_20_fosf	1,3029	0,4608	0,3370	0,3189	2	2
42	leche_s1_20_gluc	0,5296	-0,2822	-0,4353	-0,7053	2	2
43	leche_s1_20_sac	-0,0153	0,7071	0,5316	0,2331	2	2
44	leche_s1_20_sal	0,1025	2,1669	1,0705	-1,1681	2	2
45	leche_t1_5_agua	0,7197	-0,1538	0,7395	0,1278	2	2
46	leche_t1_5_bic	-0,7452	0,7338	-0,1218	0,1623	2	2
47	leche_t1_5_cit	0,1838	-0,5024	0,9342	0,8229	2	2
48	leche_t1_5_fosf	1,1417	0,1853	0,5870	-0,7010	2	2
49	leche_t1_10_bic	0,0844	-0,6374	0,7680	-0,6343	2	2
50	leche_t1_10_cit	0,7606	0,7079	0,7846	-1,2506	2	2
51	leche_t1_10_fosf	0,3799	0,3798	0,8264	-0,8476	2	2
52	leche_t1_10_gluc	0,7328	0,2416	1,0238	-0,8312	2	2
53	leche_t1_10_sac	0,2660	0,3949	1,4203	0,3688	2	2
54	leche_t1_10_sal	0,5305	-0,0469	1,4561	-0,0854	2	2
55	leche_t2_15_agua	0,5008	-0,2933	-0,3216	-0,0006	2	2
56	leche_t2_15_bic	-0,2404	0,0963	-0,2067	-0,0012	2	2
57	leche_t2_15_cit	1,7510	2,0344	-0,6388	-0,7760	2	2
58	leche_t2_15_fosf	-0,1595	0,5635	-0,6389	-0,1422	2	2
59	leche_t2_15_gluc	0,3132	0,4191	0,3508	0,2837	2	2
60	leche_t2_15_sac	0,3246	0,3187	-0,0114	0,1730	2	2
61	leche_t2_15_sal	0,7462	0,3089	-0,9332	0,8228	2	2
62	leche_t2_20_bic	0,4708	0,7472	-0,1154	0,0823	2	2
63	leche_t2_20_cit	0,3226	-0,8362	0,1337	-1,5671	2	2
64	leche_t2_20_fosf	0,5490	0,2973	-0,1912	0,0031	2	2
65	leche_t2_20_gluc	0,4921	0,8856	0,1240	-0,7275	2	2
66	leche_t2_20_sac	-0,4138	0,4122	0,1451	-0,5640	2	2
67	leche_cruda_s2	-2,3636	-1,6002	-1,1977	1,2476	1	1
68	leche_cruda_s3	-1,1117	-0,6132	-1,2250	0,5841	1	1
69	leche_cruda_s4	-1,0526	-1,3881	-0,8264	1,3493	1	1
70	leche_cruda_sj2	-2,7236	-1,9753	-0,4517	2,0935	1	1
71	leche_cruda_sj_mezcla	-2,7335	-2,2475	-1,6663	0,4476	1	1
72	leche_cruda_t3_1	-3,0001	-3,4103	-2,3571	2,3038	1	1
73	leche_cruda_t3_2	-2,9039	-2,4677	-2,1416	1,2735	1	1
74	leche_cruda_t3_3	-2,1987	-1,5162	-0,6723	1,6654	1	1
75	leche_cruda_t3_mezcla	-1,6560	-1,1608	-0,8456	0,9128	1	1
76	leche_cruda_tq3_1	-2,1034	-1,7461	-2,1747	1,9720	1	1
77	leche_cruda_tq4_1	-0,6607	-1,4383	-2,0505	2,8344	1	1
78	leche_cruda_tq4_2	-0,8247	-1,7176	-2,7866	2,4738	1	1
79	leche_cruda_tq_mezcla	-0,9801	-1,1233	-1,4057	1,3097	1	1

Tabla 9. Resultados del test set

N° Progresivo	Muestra	Cm^{-1}				Variables respuesta	
		1762,6	1552,4	1429,0	1496,5		
		1	2	3	4	Clase Verdadera	Clase predecida
1	leche1_1_15_cit	-0,06532439	0,23943948	0,79374339	-1,10174128	2	2
2	leche1_1_15_fosf	-0,54648209	0,04136427	-1,86864444	0,58568478	2	1
3	leche1_5_sac	1,2602804	1,09075108	0,18935014	0,6110998	2	2
4	leche1_5_sal	1,11959016	0,49606963	-0,48222656	-0,54952856	2	2
5	leche2_5_agua	0,15817945	0,10511774	0,74021586	0,5634508	2	2
6	leche2_5_cit	0,75939099	0,38391258	0,49772059	-0,37088618	2	2
7	leche2_5_sal	0,2616252	-0,30818564	-0,0640931	-0,73172244	2	2
8	leche_cruda2	-0,27873969	1,15231001	1,03246321	-1,53426528	1	2
9	leche_s1_15_cit	0,29054318	-0,30727291	0,78031491	-0,66103958	2	2
10	leche_s1_20_bic	0,74329848	0,38273584	-0,47547314	-0,80019846	2	2
11	leche_t1_5_gluc	0,7629525	-0,27483079	0,5557532	0,48072091	2	2
12	leche_t1_5_sac	0,19739536	0,15420355	0,59089763	-1,0993458	2	2
13	leche_t1_5_sal	1,01351137	-0,13274147	1,20015285	-0,37258991	2	2
14	leche_t1_10_agua	0,21348064	0,30840138	1,62255309	-0,99374575	2	2
15	leche_t2_20_agua	0,80043162	0,19453948	-0,04275816	1,3508146	2	2
16	leche_t2_20_sal	0,4256241	0,36329429	0,8269086	0,5681433	2	2
17	leche_cruda_sj1	-0,86576313	-0,99714425	-1,60536998	1,53414601	1	1
18	leche_cruda_sj3	-0,7957379	-1,83594038	-1,45081421	0,84640777	1	1
19	leche_cruda_tq3_2	-1,21542716	-1,72687637	-0,93514062	0,48345859	1	1

Según el estudio realizado por Ortiz Ramírez (2009) en el equipo FTIR, permite apreciar el espectro de la leche de vaca con las regiones de sus componentes mayoritarios.

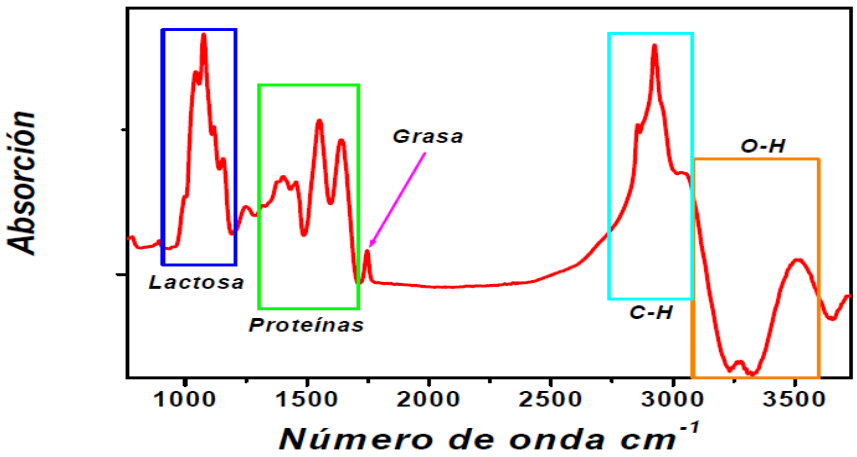


Figura 5. Espectro de la leche con las regiones de sus componentes mayoritarios.

Fuente: Ortiz Ramírez A, 2009.

De acuerdo a la figura 5, la región de la lactosa está desde 800 a 1200cm^{-1}, la región de las proteínas (1200 a 1700 cm^{-1}), el pico de las grasas (1680 a 1800cm^{-1}), los picos de los C-H (2700 a 3100cm^{-1}) y el pico asociado a los O-H (3100 a 3600cm^{-1}).

En la figura 6 se puede observar las longitudes de onda seleccionadas por el algoritmo genético. La longitud de onda de 1762cm^{-1} que corresponde al pico de las grasas, las longitudes de onda de 1552 cm^{-1}, 1429 cm^{-1} y 1496,5 cm^{-1}, corresponden a la región de las proteínas.

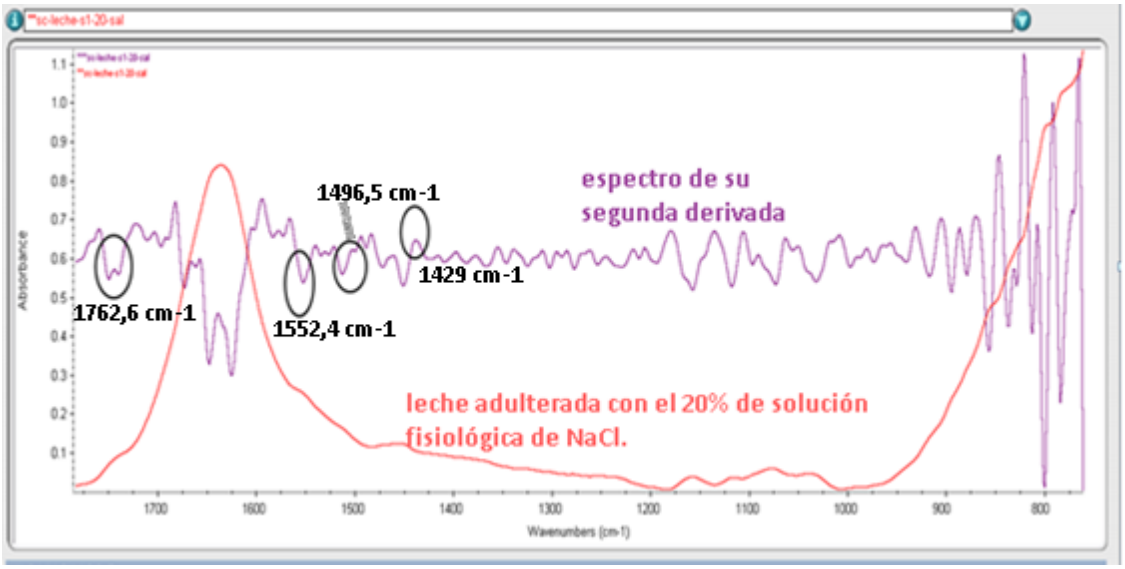


Figura 6. Espectros FT-MIR-ATR obtenidos para una muestra leche adulterada con el 20% de solución fisiológica de cloruro de sodio y el espectro de su segunda derivada, en la región de 750 a 1800 cm^{-1} con las longitudes de onda seleccionadas por el algoritmo genético.

La matriz de confusión nos indica los objetos clasificados correctamente (aquellos objetos pertenecientes a la diagonal principal) y aquellos objetos que han sido asignados a otras clases se encuentran en los demás elementos de la matriz, por lo tanto la tasa de error del clasificador para cada clase se puede evaluar mediante esta matriz de confusión como se indica en la tabla Nº 10

Tabla 10. Matriz de Confusión de 4 variables para el Training Set

	Clases calculadas		
Clases verdaderas	**1**	**2**	**Total**
1	13	6	19
2	1	59	60

En donde la clase 1: pertenece a la leche natural y la clase 2 a la leche adulterada con agua o sueros fisiológicos.

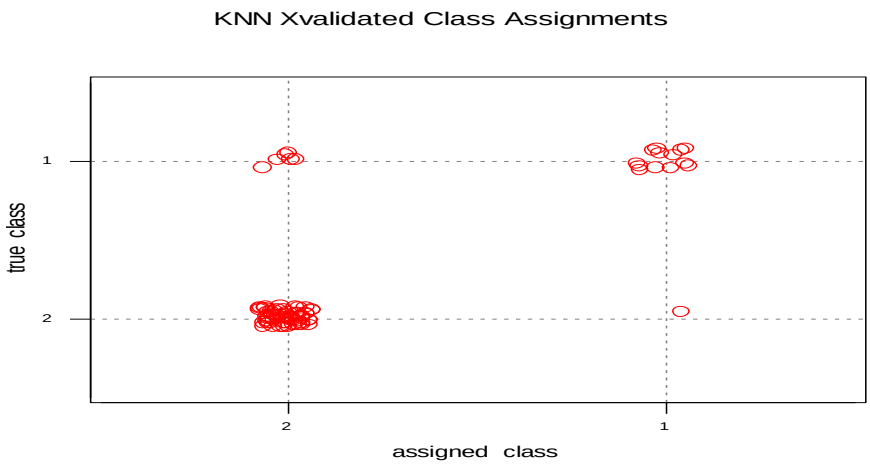


Figura 7. Asignación de clases mediante el método KNN con validación cruzada

De la matriz de confusión del training set se puede obtener los parámetros de calidad del modelo de clasificación para cada una de las clases, estos parámetros son: la precisión, sensibilidad y especificidad para las variables 1 y 2, las cuales se detallan en la tabla N° 11.

Tabla 11. Parámetros de Calidad del modelo de clasificación por clases para el training set

Parámetros de Calidad	**Clases**	
	1	**2**
Prec.	0,928571	0,907692308
Sens.	0,684211	0,983333333
Espec.	0,983333	0,684210526

La precisión es la capacidad del modelo de clasificación de no incluir objetos de otras clases en la clase considerada, por lo tanto viene a ser el porcentaje entre los objetos asignados a esta clase y el número total de objetos asignados para esta clase; la sensibilidad se describe como la capacidad del modelo de reconocer correctamente a los objetos pertenecientes a una clase, y está dada por el porcentaje entre los objetos asignados a esta clase y el número total de objetos pertenecientes a esta clase; mientras que la especificidad es la capacidad de una clase de aislar desde otra clase los objetos de aquella clase, es decir es su grado de pureza.

Resultados para el Test Set:

Como ya se dijo anteriormente la validación externa nos sirve para poder comprobar cómo las muestras de clase conocidas son calculadas por el modelo, entonces las diferencias entre los predichos y los pertenecientes a las clases reales de estas muestras nos darán la tasa de error real que se podría esperar cuando este clasificador se utilice para analizar muestras desconocidas; los parámetros de calidad obtenidos para el test set (conjunto de validación) se detallan en la tabla 12.

Tabla 12. Parámetros de calidad del modelo para el conjunto de validación externa

Ner	0,894736842	
	1	**2**
Sens	0,75	0,93333333
Esp	0,75	0,93333333

En donde la clase 1: pertenece a la leche natural y la clase 2 a la leche adulterada con agua o sueros fisiológicos.

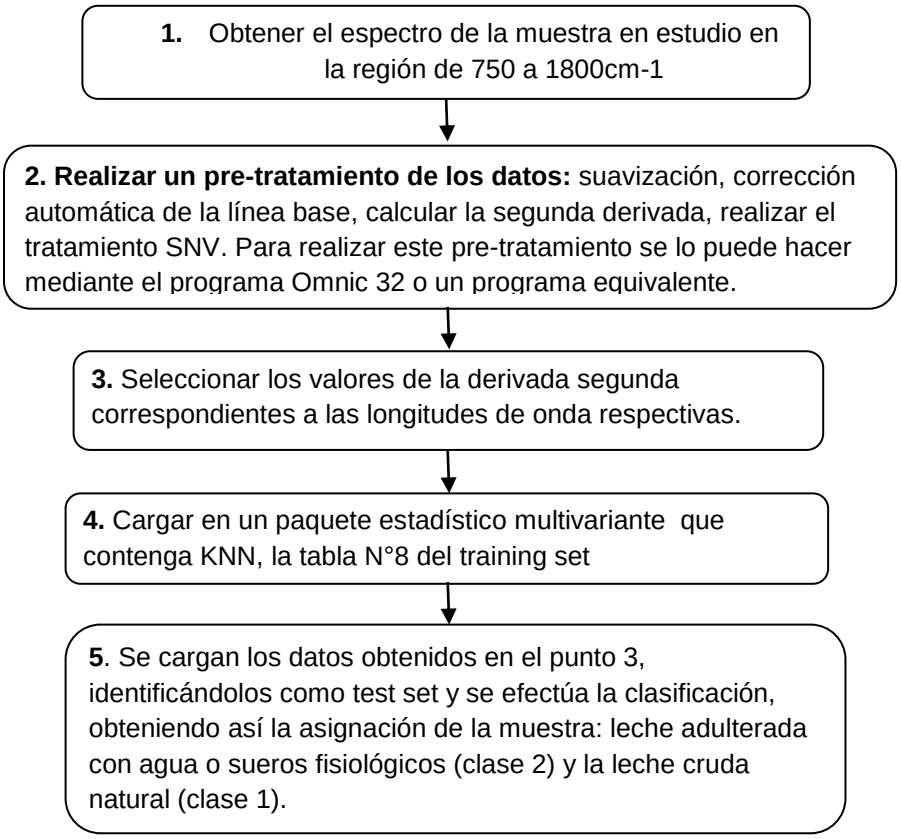
El porcentaje de clasificación correcta para el conjunto de validación externa es un importante factor que expresa el rendimiento de la clasificación para este estudio. En la tabla N°13 se detalla la matriz de confusión obtenida por el modelo, de la cual han sido obtenidos los parámetros de calidad para el modelo de clasificación.

Tabla 13. Matriz de Confusión para el conjunto de validación externa

	Clases Calculadas		**Total**
Clases Verdaderas	**1**	**2**	
1	3	1	4
2	1	14	15
Total	4	15	19

A continuación se presenta un esquema de aplicación del método para detectar si una muestra presenta o no adulteración con sueros fisiológicos

2.3 Diagrama de Flujo para aplicación del método de detección de adulteración mediante FT-MIR.



Este proceso requiere aproximadamente 30 minutos para determinar el posible aguado de la muestra.

CAPÍTULO III.

DISCUSIÓN

El método de Crioscopía (tablas 1, 2, 3 y 4) sólo detecta la adulteración con agua destilada y cuando se trata de una solución de suero fisiológico no lo detecta, debido a que la leche adulterada con suero fisiológico contiene la misma concentración de sustancias disueltas que la leche natural, por lo tanto se demostró que el método de crioscopía no discrimina la leche adulterada con suero fisiológico.

Con el método de espectroscopía de infrarrojo FT-MIR-ATR se logró discriminar la leche natural de la leche adulterada con agua y sueros fisiológicos, mediante el análisis de su espectro, debido a que su grupo químico puede absorberse en la región infrarroja y por lo tanto permite identificar claramente las bandas de absorción características para estos grupos en las regiones de infrarrojo medio (Da-Wen Sun, 2009)

Para poder llevar a cabo este estudio se utilizó el método de clasificación KNN en combinación con el método de algoritmos genéticos para determinar las mejores longitudes de onda (variables) que nos permitieron la mejor clasificación KNN.

En la tabla Nº 6 se puede observar las bandas seleccionadas por el algoritmo genético:

La longitud de onda seleccionada de 1762cm^{-1} que corresponde al grupo carbonilo C=0 presentes en los enlaces anhídrido está muy cerca de la teórica 1760cm^{-1} (Da-Wen Sun, 2009) y corresponde al pico de las grasas, esta selección puede deberse a que la leche natural presenta un mayor porcentaje de grasa que las leches adulteradas con sueros fisiológicos o agua.

La longitud de onda de 1552cm^{-1} que corresponde al grupo N-H, está asociado con los enlaces peptídicos que son principalmente responsables de la absorción ocurrida entre 1700 y 1500 cm^{-1}. Las bandas de las amidas I, II y III se conoce que son sensibles a la conformación adoptada por principales proteínas, por lo tanto la longitud de onda seleccionada por el algoritmo genético está muy cercana a la teórica 1560 (Amida II) y corresponde a la región de las proteínas(Da-Wen Sun, 2009) esta selección puede deberse a una disminución de esta proteína en la leche adulterada, al igual que las longitudes de onda de 1429cm^{-1} y 1496cm^{-1}, las cuales también corresponden a la región de proteínas.

En la práctica el método trabaja utilizando la información de dos componentes de la leche: grasas y proteínas simultáneamente y por esto logra ser eficiente en discriminar las leches adulteradas con agua o sueros fisiológicos de las muestras de leche natural.

Los parámetros de calidad del modelo para el conjunto de entrenamiento están dados por la tasa de no error de clasificación con validación cruzada (NERcv), este parámetro nos dice la bondad del modelo que para el conjunto de entrenamiento es de 0.911, es decir que el porcentaje de error de clasificación del modelo es del 8.9%.

La matriz de confusión para el training set (tabla N°10) nos permite evaluar la tasa de error del clasificador por clases y de esta matriz se obtienen los parámetros de calidad como sensibilidad, especificidad y precisión.

La sensibilidad nos permite determinar el porcentaje de clasificación correcta del clasificador, y se puede observar que para la clase 1 (leche natural), el modelo clasifica a 13 de los 19 elementos dentro de esta categoría, lo que nos da un porcentaje equivalente al 68,42% de clasificación correcta para esta variable, y para la clase 2 (leche adulterada con agua o sueros fisiológicos), el modelo clasifica a 59 elementos de un total de 60 dentro de esta categoría, que es equivalente a un porcentaje 98,33% de clasificación correcta para esta categoría.

Como se puede ver este modelo de clasificación es muy selectivo para discriminar la leche adulterada, mientras que para la leche natural es posible que haya sido adulterada con agua de los recipientes es por ello que el modelo clasifica a 6 elementos de esta categoría como leche adulterada.

La Precisión nos permite determinar el porcentaje de clasificación asignada por el clasificador y se puede observar que para la clase 1 (leche natural), el modelo clasifica a 13 de 14 elementos dentro de esta categoría, lo que nos da un porcentaje equivalente al 92,857% de clasificación asignada para esta variable, y para la clase 2 (leche adulterada con agua o sueros fisiológicos), el modelo clasifica a 59 elementos de un total de 65 dentro de esta categoría, que es equivalente a un porcentaje 90,769% de clasificación asignada para esta categoría.

Para evitar sobreestimar la capacidad predictiva del modelo o tener una mala interpretación del mismo, se evalúa la capacidad de predicción del discriminante, para ello se desafía al discriminante con muestras de clase conocida que no formaron parte del conjunto de entrenamiento y la diferencia entre los predichos frente a la pertenencia a las clases reales nos da la tasa de error real que se podría esperar cuando este clasificador se utilice para analizar muestras reales desconocidas.(B.K. Lavine, 2009).

Los parámetros de calidad obtenidos para el conjunto de validación externa se detallan en la tabla Nº 12

Podemos ver que la tasa de no error para el test set es de 0,894 y por lo tanto el error del modelo es de 10,6% que vendría a ser la tasa de error real para la clasificación de muestras desconocidas.

La sensibilidad del modelo para la clase 1 (leche natural) es de 75% y para la clase 2 (leche adulterada con agua o suero fisiológico) es de 93,3%.

La especificidad para el conjunto de validación coincide con la sensibilidad, es decir que la clasificación asignada por el clasificador es igual a la clasificación real.

Para el conjunto de validación externa el modelo sigue siendo muy sensible para la categoría 2 (leche adulterada) es decir que la probabilidad de que una leche adulterada no sea discriminada es mínima; mientras que para la categoría 1 (leche natural) el modelo puede generar verdaderos falsos, o como ya se dijo anteriormente puede deberse a que la leche contenía ya un pequeño porcentaje de adulteración debido al agua contenida en los recipientes utilizados para la recolección de la muestra.

CONCLUSIONES

- El método crioscópico no detecta la adulteración de leche con suero fisiológico.
- Mediante el método de espectroscopía de infrarrojo FT-MIR-ATR, es posible discriminar la leche natural de la leche adulterada con agua o sueros fisiológicos.
- Los parámetros de calidad del modelo obtenido tanto para el training set y el conjunto de validación son aceptables, por lo tanto esta técnica permite detectar la adulteración en muestras desconocidas.
- El modelo es sensible para el caso de leche adulterada, por lo que es poco probable que una muestra que contenga adulterantes (agua o suero fisiológico) no sea discriminada.
- La espectroscopía infrarroja es una técnica que presenta varias ventajas frente a otras, entre las principales ventajas tenemos la rapidez, simplicidad del análisis ya que es una técnica no destructiva y requiere una mínima cantidad de muestra.

REFERENCIAS BIBLIOGRÁFICAS

Amiot J. y colaboradores. *Ciencia y Tecnología de La Leche.* Acribia S.A. Zaragosa. 1991. 547p.

Da-Wen Sun. *Infrared Spectroscopy for Food quality Analysis and Control*. Dublin (Ireland). Elsevier Inc. cop. 2009. 424p.

Galva˜o R. K. H. and Araujo M-C.U. "Variable Selection". En: *SD Chemometrics.* Brazil. Elsevier B.V. cop. 2009.235-281p, 3vol.

Jonsson H. and Gabrielsson J. "Evaluation of Preprocessing Methods". En: *SD Chemometrics*. Umea, Sweden. Elsevier B.V. cop. 2009. 199-205 p, 2 vol.

Kohler A.; Zimonja M. and Segtnan V.; Martens H. "Standard Normal Variate, Multiplicative Signal Correction and Extended Multiplicative Signal Correction Preprocessing in Biospectroscopy". En: *SD Chemometrics.* Norway. Elsevier B.V. cop. 2009.

Leardi R. "Genetic Algorithms". En: *SD Chemometrics*.. Elsevier B.V. cop. 2009. 631-651p, 1 vol.

Lavine B. K. and Rayens W.S. "Classification: Basic Concepts". En: *SD Chemometrics* Elsevier B.V. cop. 2009. 507-514p, 3 vol.

Lavine B.k. "Validation of Classifiers". En: *SD Chemometrics.* Elsevier B.V. cop 2009.587-598p,3 vol.

Lavine B. K. "Feature Selection: Introduction". En: *SD Chemometrics.* Elsevier B.V. cop. 2009. 601-605p. 3 vol.

Lavine B.K. and Davison C.E. "Multivariate Approaches to Classification using Genetic Algorithms". En: *SD Chemometrics*. Elsevier B.V. cop 2009. 619-644p, 3 vol.

M. S. Reis and P. M. Saraiva, University of Coimbra, Coimbra, Portugal
B. R. Bakshi, Ohio State University, Columbus, OH, USA. "Denoising and Signal-to-Noise Ratio Enhancement: Wavelet Transform and Fourier Transform". En: *SD Chemometrics.*Elsevier B.V. cop. 2009

Olivieri A. C.; Faber N.M. “Validation and Error”. En: *SD Chemometrics.* Elsevier B.V. cop 2009. 92-113 p, 3 vol.

Reis M.S. and Saraiva P.M., “Denoising and Signal-to-Noise Ratio Enhancement: Wavelet Transform and Fourier Transform.” En: *SD Chemometrics.* Elsevier B.V. cop. 2009, 26-54 p. 2 vol.

Taavitsainen V.M. “Denoising and Signal-to-Noise Ratio Enhancement: Derivatives”. En: *SD Chemometrics.* Elsevier. B.V. cop. 2009. 57-65p. 2 vol.

Tomasi G. and Bro R. “Multilinear Models: Iterative Methods”. En: *SD Chemometrics.* Elsevier B.V. cop. 2009. 412-448p, 2 vol.

Trygg J.; Gabrielsson J.; Lundstedt T. “Background Estimation, Denoising, and Preprocessing”. En: *SD Chemometrics.* Elsevier B.V. cop.2009. 1-6 p, 2 vol.

Zeaiter M. and Rutledge. “Preprocessing Methods”. En: *SD Chemometrics*. Elsevier B.V. cop. 2009. 123-179p, 3 vol.

REFERENCIAS ELECTRÓNICAS

"Determinación de la leche con Agua, Cloruros y Sacarosa Guía Práctica". (En línea). Universidad de Zulia. Maracaibo (Venezuela). http://www.revistavirtualpro.com/files/ti23_200512.pdf. (Consulta: 10 de Julio del 2012).

Fana J., Boscan L. (1992). "Detección de adulteración de leche con suero mediante la relación proteína sérica/caseinato". (En línea). Zulia, Venezuela. http://revistas.luz.edu.ve/index.php/rc/article/viewFile/5137/4986 (Consulta: 9 de Mayo del 2012).

Griffiths Peter and Haseth James (cop 2007). "Fourier Transform Infrared Spectroscopy". (En línea). http://books.google.com.ec/books?id=C_c0GVe8MX0C&printsec=frontcover&hl=es&source=gbs_atb#v=onepage&q&f=false (Consulta: 05 de Enero del 2013)

Galindo Amaya L.M., Valbuena Colmenares E. y Rojas Villarroel E. (2006) "Estandarización de la detección del Glicomacropéptido Por Page-SDS como índice de Adulteración de Leche". (En línea). Zulia, Venezuela. http://www.scielo.org.ve/pdf/rc/v16n3/art13.pdf. (Consulta: 05 Abril del 2012).

Guidotti Marcello. Tecnología Farmacéutica: preparazioni per via endovenosa (en línea). http://www.galenotech.org/iniettive.htm (Consulta: 5 de Mayo del 2012).

Irudayaraj, Joseph Reh, Christoph.(2008). Nondestructive Testing of Food Quality. (En línea). http://site.ebrary.com/lib/ucuenca/docDetail.action?docID=10296645&p00=infrared%20spectroscopy (Consulta: 10 Diciembre del 2012)

Kasemsumran S.; Thanapase W., and Kiatsoonthon A. "Feasibility of Near-Infrared Spectroscopy to Detect and to Quantify Adulterants in Cow Milk." (En línea). https://www.jstage.jst.go.jp/article/analsci/23/7/23_7_907/_pdf (Consulta: 10 de Julio del 2012).

Larrahondo, J.E., F.J. Palau y C. Ramírez. 2001. Espectroscopía de infrarrojo cercano - NIR (en línea). Centro de investigación de la caña de azúcar de Colombia. http://www.cenicana.org/publicaciones/serie_pro_indus/serie_pro_indus.php?opcion=2. (Consulta: 02 Julio del 2012).

Ortiz Ramirez A. (2009). “Determinación de los Componentes y Adulteración de la leche, empleando la espectroscopía de F-TIR”. [En línea] Ciba, Tlaxcala. http://www.repositoriodigital.ipn.mx/bitstream/handle/123456789/8648/Alicia%20Ortiz%20Ramirez%20-%20Tesis%20de%20Maestria.pdf?sequence=1 (Consulta: 8 de Enero del 2013).

Panreac Química S.A., Métodos Analíticos en Alimentaria: “Métodos oficiales de Análisis de Leche.” (en línea). http://es.scribd.com/doc/3955081/Metodos-Oficiales-de-Analisis-leche Metodos oficiales 2012-04-16. (Consulta: 16 de Abril del 2012)

Parra Huertas R.A. “Lactosuero: Importancia en la industria de alimentos.” (En línea). Medellín, Colombia. http://www.scielo.org.co/pdf/rfnam/v62n1/a21v62n1.pdf. (Consulta: 31 enero del 2013)

Stuart, Barbara (2004). Infrared Spectroscopy: Fundamentals and Applications. (En línea).http://site.ebrary.com/lib/ucuenca/docDetail.action?docID=10113963&p00=infrared%20spectroscopy . (Consulta: 02 de Diciembre del 2012)

Tabla de Correlaciones en espectroscopía Infrarroja. (En línea). http://es.wikipedia.org/w/index.php?title=Anexo:Tabla_de_correlaciones_en_espectroscopia_infrarroja&oldid=59626931 . (Consulta: 15 de Diciembre del 2012).

T. Theophanides. “Fourier Transform Infrared Spectroscopy. Química Industrial y Aplicaciones Bioquímicas”. (En línea). http://link.springer.com/book/10.1007/978-94-009-6418-1/page/1 . (Consulta: 10 de Diciembre del 2012).

ANEXOS

Anexo 1:

Los espectros FT-MIR-ATR sometidos a un pre tratamiento de datos (suavización, corrección automática de la línea de base y aplicada la segunda derivada) luego digitalizados y auto escalados son presentados en formato digital, que se encuentra disponible en la Biblioteca Hernán Malo de la Universidad del Azuay en la ciudad de Cuenca.

Anexo 2:

Tabla de correlaciones en espectroscopía Infrarroja para la determinación de la estructura de los componentes orgánicos.

Enlace	Tipo de enlace	Tipo específico de enlace	Rango e intensidad de absorción
C-H	alquilo	metilo	1380 cm^{-1} (débil), 1460 cm^{-1} (fuerte) y 2870, 2960 cm^{-1} (ambos, de fuerte a medio)
		metileno	1470 cm^{-1} (fuerte) y 2850, 2925 cm^{-1} (ambos, de fuerte a medio)
		metino	2890 cm^{-1} (débil)
	vinilo	$C=CH_2$	900 cm^{-1} (fuerte) y 2975, 3080 cm^{-1} (medio)
		C=CH	3020 cm^{-1} (medio)
		alqueno monosustituido	900, 990 cm^{-1} (ambos fuerte)
		alqueno cis-disustituido	670-700 cm^{-1} (fuerte)
		alqueno trans-disustituido	965 cm^{-1} (fuerte)
		alqueno trisustituido	800-840 cm^{-1} (fuerte a medio)
	aromático	benceno/benceno sustituido	3070 cm^{-1} (débil)
		benceno monosustituido	700-750 cm^{-1} (fuerte) y 700±10 cm^{-1} (fuerte)
		benceno orto-disustituido	750 cm^{-1} (fuerte)
		benceno meta-disustituido	750-800 cm^{-1} (fuerte) y 860-900 cm^{-1} (fuerte)
		benceno para-disustituido	800-860 cm^{-1} (fuerte)
	alquino		3300 cm^{-1} (medio)
	aldehído		2720, 2820 cm^{-1} (medio)
C=C	C=C acíclico	alqueno monosustituido	1645 cm^{-1} (medio)
		alqueno 1,1-disustituido	1655 cm^{-1} (medio)
		alqueno cis-1,2-disustituido	1660 cm^{-1} (medio)
		alqueno trans-1,2-disustituido	1675 cm^{-1} (medio)
		alqueno tri y tetrasustituido	1670 cm^{-1} (débil)
	C=C conjugado	dienos	1600, 1650 cm^{-1} (fuerte)
		con anillo de benceno	1625 cm^{-1} (fuerte)
		con C=O	1600 cm^{-1} (fuerte)
	C=C aromático		1450, 1500, 1580, 1600 cm^{-1} (fuerte a débil) - siempre los 4
	C≡C	alquino terminal	2100-2140 cm^{-1} (débil)
		alquino disustituido	2190-2260 cm^{-1} (muy débil, a veces no visible)
C=O	cetona/aldehído	alifáticos saturados/ciclos de 6 miembros	1720 cm^{-1}
		α,β-insaturado	1685 cm^{-1} (también va para cetonas aromáticas)
		ciclo de 5 miembros	1750 cm^{-1}
		ciclo de 4 miembros	1775 cm^{-1}

		aldehído	1725 cm^{-1} (influencia de la conjugación como en las cetonas)
	derivados de ácido carboxílico	ácido carboxílico saturado	1710 cm^{-1}
		ácido carboxílico insaturado/aromático	1680-1690 cm^{-1}
		ésteres y lactonas	1735 cm^{-1} (influencia de la conjugación y el tamaño del anillo como en las cetonas)
		anhídridos	1760 y 1820 cm^{-1} (ambos)
		halogenuros	1800 cm^{-1}
		amidas	1650 cm^{-1} (amidas asociadas)
		sales de carboxilatos	1550-1610 cm^{-1} (también va para zwitteriones aminoacídicos)
O-H	alcoholes, fenoles		3610-3670 cm^{-1} (la concentración de la muestra ensancha la absorción y la mueve a 3200-3400 cm^{-1})
	ácidos carboxílicos		3500-3560 cm^{-1} (la concentración de la muestra ensancha la absorción y la mueve a 3000 cm^{-1})
N-H	aminas primarias		doblete entre 3400-3500 cm^{-1} y 1560-1640 cm^{-1} (fuerte)
	aminas secundarias		sobre 3000 cm^{-1} (medio a débil)
	iones amonio		se ensancha con múltiples picos entre 2400-3200 cm^{-1}
C-O	alcoholes	primarios	1050±10 cm^{-1}
		secundarios	alrededor de 1100 cm^{-1}
		terciarios	1150-1200 cm^{-1}
	fenoles		1200 cm^{-1}
	éteres	alifáticos	1120 cm^{-1}
		aromáticos	1220-1260 cm^{-1}
	ácidos carboxílicos		1250-1300 cm^{-1}
	ésteres		1100-1300 cm^{-1}
C-N	aminas alifáticas		1020-1220 cm^{-1} (frecuentemente sobrepuestas)
	C=N		1615-1700 cm^{-1} (efectos de conjugación similares a C=O)
	nitrilos (enlace triple C-N)		2210-2260 cm^{-1} (no conjugado 2250, conjugado 2230 cm^{-1})
	isonitrilos (enlace R-N-C)		2165-2110 cm^{-1} (2140 - 1990 cm^{-1} para R-N=C=S)
C-X (X=F, Cl, Br, I)	fluoroalcanos	ordinarios	1000-1100 cm^{-1}
		trifluorometilo	dos fuertes, se ensancha entre 1100-1200 cm^{-1}
	cloroalcanos		540-760 cm^{-1} (medio to débil)
	bromoalcanos		below 600 cm^{-1}
	iodoalcanos		below 600 cm^{-1}
N-O	compuestos nitro	alifático	1540 cm^{-1} (más fuerte) y 1380 cm^{-1} (más débil) - Siempre ambas.
		aromático	1520, 1350 cm^{-1} (la conjugación usualmente baja el número de onda)