



PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE
ESCUELA DE INGENIERÍA

MEDICIÓN DE LA DENSIDAD BÁSICA DE MADERA USANDO TÉCNICAS ESPECTROSCÓPICAS

CRISTÓBAL GALLEGUILLOS HART

Tesis para optar al grado de
Magíster en Ciencias de la Ingeniería

Profesor Supervisor:
ANDRÉS GUESALAGA

Santiago de Chile, (Marzo, 2010)

© 2010, Cristóbal Galleguillos Hart



PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE
ESCUELA DE INGENIERÍA

MEDICIÓN DE LA DENSIDAD BÁSICA DE MADERA USANDO TÉCNICAS ESPECTROSCÓPICAS

CRISTÓBAL GALLEGUILLOS HART

Tesis presentada a la Comisión integrada por los profesores:

ANDRÉS GUESALAGA

M. PAULINA FERNÁNDEZ

MIGUEL TORRES

DOMINGO MERY

Para completar las exigencias del grado de
Magíster en Ciencias de la Ingeniería

Santiago de Chile, (Marzo, 2010)

© 2010, Cristóbal Galleguillos Hart

A mi familia, gracias por el constante apoyo.

AGRADECIMIENTOS

Quisiera agradecer a mi profesor guía Andrés Guesalaga y a la profesora Paulina Fernández por su tiempo, motivación y colaboración en el desarrollo de este trabajo.

También a Ester San Martín, Rebeca Sanhueza y a toda la gente de Forestal Mininco quienes participaron en la recolección de muestras de *Eucalyptus*, una pieza fundamental para esta investigación.

Por último, a mis compañeros de oficina, con quienes aclaré una infinidad de preguntas y problemas que me permitieron seguir avanzando.

INDICE GENERAL

	Pág.
DEDICATORIA	ii
AGRADECIMIENTOS	iii
INDICE DE TABLAS	vi
INDICE DE FIGURAS	vii
RESUMEN.....	ix
ABSTRACT.....	x
 1 INTRODUCCIÓN.....	 1
1.1 Motivación	1
1.2 Objetivo.....	3
1.3 Experiencias anteriores	4
 2 MARCO TEÓRICO	 6
2.1 Densidad de la madera	6
2.2 Espectroscopía Raman	7
2.2.1 Espectro Raman	9
2.3 Espectroscopía NIR.....	15
2.3.1 Espectro NIR	15
2.4 Regresión multivariable PLS	17
 3 MATERIALES Y MÉTODOS.....	 23
3.1 Muestras	23
3.2 Espectrómetro Raman	24
3.3 Espectrómetro NIR.....	29
3.4 Diseños óptico-mecánicos para la adquisición de espectros.....	30
3.4.1 Diseño óptico-mecánico Raman	31
3.4.2 Diseño óptico-mecánico NIR	32
3.5 Método analítico tradicional	33

3.6	Adquisición de espectros.....	35
3.6.1	Adquisición del espectro Raman	35
3.6.2	Adquisición del espectro NIR.....	36
3.7	Procesamiento de espectros.....	36
3.7.1	Promedio y sustracción del espectro de oscuridad (dark)	37
3.7.2	Selección del rango espectral.....	38
3.7.3	Filtro de Savitsky Golay	39
3.7.4	Media centrada.....	39
3.7.5	Eliminación de fluorescencia.....	40
3.8	Calibración y validación de los modelos de regresión.....	41
4	RESULTADOS Y DISCUSIÓN	48
4.1	Modelos Raman	50
4.2	Modelos NIR.....	59
5	CONCLUSIONES	66
6	TRABAJO FUTURO	68
	NOMENCLATURA	71
	BIBLIOGRAFÍA	72

INDICE DE TABLAS

	Pág.
Tabla 3-1: Características técnicas del espectrómetro Raman BTR111 MiniRam.....	28
Tabla 3-2: Características técnicas del espectrómetro NIR USB2000.....	29
Tabla 4-1: Densidad básica de las muestras [Kg m^{-3}].....	48
Tabla 4-2: Modelos previos Raman y NIR con los datos crudos.....	50
Tabla 4-3: Modelos Raman sin fluorescencia.....	57
Tabla 4-4: Modelos sin fluorescencia para el rango espectral entre $130\text{-}580\text{ cm}^{-1}$	58
Tabla 4-5: Modelos a partir de la fluorescencia del espectro Raman, rango completo ...	59
Tabla 4-6: Parámetros del modelo de calibración PLS para los espectros NIR.....	64

INDICE DE FIGURAS

	Pág.
Figura 2-1: Efecto Rayleigh, Raman y Fluorescencia	9
Figura 2-2: Bandas Anti-Stokes, Rayleigh y Stokes en un espectro Raman	11
Figura 2-3: Aumento de la SNR con un aumento del tiempo de integración	13
Figura 2-4: Espectro NIR de una muestra de <i>Eucalyptus nitens</i>	16
Figura 3-1: Extracción del tarugo y sus dimensiones	23
Figura 3-2: Corte longitudinal realizado a los tarugos.....	24
Figura 3-3: Diagrama esquemático de un espectrómetro Raman	25
Figura 3-4: Configuración de Czerny-Turner	27
Figura 3-5: Espectrómetro Raman BTR111 y sonda BAC100.....	29
Figura 3-6: Espectrómetro NIR USB2000.....	30
Figura 3-7: Esquema del montaje para la adquisición de espectros Raman	31
Figura 3-8: Detalle del contacto entre la sonda Raman y el tarugo	32
Figura 3-9: Esquema del montaje para la adquisición de espectros NIR.....	33
Figura 3-10: Montaje para medición del volumen del tarugo por desplazamiento de agua	34
Figura 3-11: Espectros de oscuridad de cada instrumento; a) Raman b) NIR.....	38
Figura 3-12: Gráfico para la selección del número de variables latentes óptimo del modelo PLS.....	42
Figura 3-13: Gráfico de dispersión valores medidos vs predichos	43
Figura 3-14: Gráfico de Muestras vs Residuos en la predicción para la detección de outliers.....	44
Figura 3-15: Ejemplo de muestras con alto residuo Q y alto T^2	45
Figura 3-16: Gráfico del vector de regresión b para un modelo NIR	45
Figura 3-17: Gráfico de los VIP Scores	46
Figura 4-1: Histograma de la densidad básica de las muestras.....	49
Figura 4-2: Espectro Raman a) datos crudos b) sin dark c) suavizado SG d) detalle del suavizado.....	51
Figura 4-3: Muestras <i>outliers</i> Raman con alto residuo de predicción	52

Figura 4-4: Muestras <i>outliers</i> Raman con valores Q y T ² fuera de los rangos de confianza	52
Figura 4-5: Gráficos de dispersión del modelo preliminar; a) Con todas las muestras b) Eliminando <i>outliers</i>	53
Figura 4-6: <i>VIP Scores</i> para el espectro Raman completo.....	54
Figura 4-7: Gráfico de dispersión modelo preliminar Raman, utilizando el rango 0-1200 cm ⁻¹	55
Figura 4-8: Eliminación de fluorescencia: a) Espectro original y fluorescencia b) Espectro sin fluorescencia.....	56
Figura 4-9: Valores residuales de predicción de las muestras NIR	60
Figura 4-10: Modelo de calibración utilizando todos los datos; a) Con <i>outliers</i> b) Sin <i>outliers</i>	61
Figura 4-11: Espectro errático de la muestra 120 vs el promedio NIR.....	62
Figura 4-12: Efecto del suavizado SG de ancho 15 y orden 2 en un espectro NIR	62
Figura 4-13: <i>VIP Scores</i> del espectro NIR completo.....	63
Figura 4-14: Rango espectral final utilizado en la calibración NIR. 750-1000 nm	64
Figura 4-15: Gráfico de dispersión para densidad básica a partir de espectros NIR	65
Figura 4-16: <i>VIP scores</i> para el modelo final NIR	65
Figura 6-1: Adquisición del espectro NIR mediante transmitancia difusa en el árbol en pie.....	69
Figura 6-2: Diseño preliminar de la sonda NIR	70

RESUMEN

El objetivo de esta investigación es la validación del uso de técnicas espectroscópicas para la medición de la densidad básica de la madera de *Eucalyptus nitens* a través de modelos de regresión multivariable PLS (*Partial Least Squares*).

Las características de las muestras y métodos utilizados buscan emular las condiciones de terreno con el fin de sentar las bases para el futuro desarrollo de un instrumento portátil que permita medir las propiedades de la madera a partir del árbol en pie, lo que implicaría una gran reducción de costos para las empresas forestales ya que permitiría realizar las mediciones directamente en las plantaciones y de forma masiva, reemplazando el lento y costoso método tradicional de medición de la densidad básica en laboratorio.

Se realizaron mediciones sobre muestras húmedas, extraídas de árboles de 8 años de edad aportadas por Forestal Mininco S.A, empresa con la cual se trabajó en conjunto desde el inicio de esta investigación. A partir de cada muestra se obtuvieron dos tipos de espectros, Raman y NIR (*Near Infra Red*); para el primero se utilizó el espectrómetro BTR-111 MiniRam de B&WTEK™ y para el segundo el espectrómetro USB2000 de Ocean Optics™. En la adquisición de los espectros se utilizaron sistemas óptico-mecánicos construidos especialmente para cada caso.

Sobre los espectros se aplicaron diversas técnicas de procesamiento de datos con el fin de mejorar los modelos de predicción de densidad básica obtenidos mediante el método de regresión multivariable PLS. El criterio de selección del mejor modelo de predicción de densidad básica se basó en altos índices de correlación en validación (R_v^2) y un mínimo RMSEP (*Root Mean Square Error of Prediction*).

Los resultados obtenidos demuestran un gran potencial para el caso de la espectroscopía NIR con índices R_v^2 de hasta 0,87 y RMSEP de 8,6 [Kg/m³]. Por otro lado, los resultados para la técnica Raman no fueron satisfactorios ya que la capacidad de predicción de los modelos resultó extremadamente baja.

Palabras claves: Raman, NIR, espectroscopía en madera, densidad básica, regresión PLS, *Eucalyptus nitens*

ABSTRACT

The aim of this investigation is to validate the use of spectroscopic techniques for the measurement of basic wood density in *Eucalyptus nitens*, by means of multivariable regression models i.e. PLS (Partial Least Squares).

The methods and samples used are meant to simulate the conditions found in the field, so the current design can be later extended to a portable instrument that could eventually determine the characteristics of a standing tree. The latter could dramatically reduce sampling costs in the forest industry by replacing the tedious, lengthy and cumbersome laboratory procedures used nowadays.

Model calibration was carried out using wet samples taken from 8 year-old trees, provided by Forestal Mininco S.A. Each sample was scanned to acquire two types of spectra: Raman and Near InfraRed (NIR). In the first case, the B&WTEK BTR-111 MiniRam™ portable spectrometer was used. The NIR spectra were obtained using the Ocean Optics USB2000™ spectrometer. Opto-mechanical arrangements had to be designed for each type of acquisition.

Several processing techniques were applied to the available data in order to obtain the best possible prediction model for basic density when using the PLS multivariable regression technique. The model selection criteria were based on maximizing the correlation indices for validation (R_v^2) and minimizing the Root Mean Square of Prediction (RMSEP).

The results show a tremendous potential of the NIR techniques for this purpose. Values of R_v^2 up to 0.87 and RMSEP down to 8.6 Kg/m³ were achieved. On the other hand, Raman spectroscopy proved to be unfit for the application, yielding extremely low correlation values and high quadratic errors.

Keywords: Raman, NIR, wood spectroscopy, basic density, PLS regression, *Eucalyptus nitens*

1 INTRODUCCIÓN

La industria forestal en Chile exportó el año 2007 USD \$4.952 millones, generando en promedio un 3.3% del PIB anual (CIF, 2008). Debido a las características climáticas y geográficas de nuestro país, se cuenta con un potencial excepcional para el desarrollo de esta actividad. Las principales especies que se manejan en Chile son el Pino Radiata (*Pinus Radiata D. Don*) y diversas especies de Eucalipto (*Eucalyptus nitens (Deane & Maiden)* y *Eucalyptus globulus Labill* principalmente); estas últimas se caracterizan por su alta densidad de fibras cortas, una propiedad ideal para la fabricación de papeles de calidad.

Para las empresas forestales es de vital importancia el continuo monitoreo de las propiedades químicas y físicas de la madera de sus árboles, ya que en base a ellas se toman importantes decisiones en aspectos tales como: producción, manejo silvícola, mejoramiento genético y tipo de productos. Ésta es una actividad permanente debido a la gran extensión de sus predios que alcanzan los 2.000.000 de hectáreas en nuestro país (CIF, 2008).

1.1 Motivación

En la actualidad, la determinación de las propiedades físicas y químicas de la madera se realiza en laboratorios, utilizando complejas técnicas de alto costo que requieren una cantidad considerable de tiempo. Esto impide contar con resultados representativos en el corto plazo, limitando la información disponible para la toma de decisiones en los distintos niveles de las empresas.

Cuando la organización cuenta con sus propios laboratorios y no necesita contratar a terceros para realizar el trabajo de muestreo y medición de las propiedades de sus plantaciones, el proceso tradicional para estimar la densidad básica de la madera consiste en:

- Definir la superficie y material genético a muestrear.
- Encargar la toma de muestras a una cuadrilla de operadores, formadas por un mínimo de dos personas.
- Los operadores, dependiendo de la lejanía, superficie de los predios a muestrear y equipo del que dispongan, extraen un promedio de veinte muestras diarias.
- Al final del día, las muestras son llevadas al laboratorio.
- Cuando se cuenta con un número mínimo de muestras, se ingresan a un impregnador por 12 a 24 horas, con el fin de saturar las muestras con agua destilada (todas por igual).
- Después de impregnadas, se mide el volumen de cada muestra, usando desplazamiento de agua (una a una). Se suelen realizar hasta tres mediciones por muestra con el fin de mejorar la precisión.
- Posteriormente se secan las muestras en un horno a temperatura controlada por al menos 12 horas.
- Finalmente se mide la masa de cada muestra seca con una balanza de precisión.

El resultado final del proceso anteriormente descrito corresponde a la relación entre la masa seca y el volumen verde de la muestra, esto se define como la densidad básica de la madera y es una de las variables de mayor interés para la industria forestal. En base a este y otros parámetros se decide si el árbol en cuestión se destina a la producción de celulosa, madera aserrable, clonación o se deja madurar hasta la siguiente temporada de cosecha (Megraw, 1986).

El proceso anterior cuenta con varios “cuellos de botella” que limitan el número de muestras que se pueden obtener, afectando el control de la empresa sobre toda la línea de producción y sobre el total del patrimonio forestal. Aumentar la cantidad de muestras o la frecuencia del muestreo mediante el método tradicional implica un importante

aumento en los costos del proceso, tanto en lo referido a recursos humanos y materiales, como en el tiempo designado a esta tarea. Lo anterior impulsa a desarrollar nuevas técnicas de medición que permitan disminuir los costos y aumenten el volumen de muestreo.

La espectroscopía aparece como una técnica ideal para reemplazar los métodos tradicionales de laboratorio. La capacidad de realizar un gran número de mediciones en cortos intervalos de tiempo, la poca preparación y la no destrucción de las muestras son algunas de las potenciales ventajas. Más aún, con los desarrollos recientes de la tecnología asociada, hoy en día es posible encontrar una amplia gama de espectrómetros portátiles de bajo costo, que podrían permitir la adquisición del espectro directamente en terreno.

1.2 Objetivo

Esta investigación tiene como objetivo principal validar una técnica de adquisición y procesamiento de espectros de madera verde de *Eucalyptus nitens*, que sienta las bases para el desarrollo un futuro instrumento portátil capaz de medir la densidad básica de la madera en terreno junto con otras propiedades de esta materia prima. Se espera obtener modelos matemáticos que permitan predecir la densidad básica de la madera a partir de su espectro Raman o NIR.

Es muy importante mencionar que la gran mayoría de los estudios previos relacionados con la predicción de propiedades físicas y químicas de la madera con ambas técnicas espectroscópicas, han sido desarrollados utilizando muestras que requieren complejos procesamiento de laboratorio (Jones, et al., 2008; Noupouponen et al., 2005; Schimleck et al., 2005; Ona et al., 2003; Schimleck et al., 1999). Este procesamiento incluye, por ejemplo, el secado y posterior molienda de las muestras con el fin de prepararlas para la adquisición del espectro con espectrómetros de gran tamaño, no portables.

Las diferencias fundamentales de esta tesis con los estudios previos radica, por una parte, en la naturaleza de las muestras utilizadas, las cuales fueron mínimamente intervenidas con el fin de mantener las propiedades de la muestra recién extraída, simulando de la mejor manera posible las condiciones en la cuales operaría el instrumento portátil a desarrollar, y por otra parte, en las características físicas de los espectrómetros utilizados, los cuales son portátiles, livianos y fáciles de utilizar en terreno, pero con capacidades técnicas en cuanto al rango espectral en la adquisición, la potencia del láser Raman y la sensibilidad del detector más limitadas en comparación con otros instrumentos más complejos.

Si bien este estudio se realizó sobre muestras extraídas de ejemplares vivos, las técnicas utilizadas están pensadas de manera tal, que permitan a futuro la adquisición espectros en terreno directamente sobre el árbol en pie, sin necesidad de extraer biomasa del ejemplar, utilizando, por ejemplo, sondas que se introduzcan directamente en el tronco del árbol, guiando la luz de excitación y recolectando la luz dispersada hacia el espectrómetro.

1.3 Experiencias anteriores

El uso de espectroscopía Raman para la determinación de propiedades de la madera es relativamente nuevo en comparación con la espectroscopía NIR. Sin embargo, el trabajo previo es promisorio, utilizando espectrómetros de laboratorio y con un complejo manejo de la muestra previo a la toma del espectro Raman; se han logrado estimar con altos índices de correlación diferentes propiedades físicas de la madera tales como: proporción de tipos de células, morfología y dimensiones de la célula y mecanismos de deformación en fibras (Ona et al., 1999). También se han realizado estudios para determinar los constituyentes químicos de la madera tales como: α -celulosa, holocelulosa, hemicelulosa, lignina y otros (Ona et al., 2003), obteniendo resultados contundentes en la capacidad de predicción de éstas. Para la estimación de la densidad

básica en *Eucalyptus*, los resultados previos son auspiciosos, obteniendo índices de correlación en validación de hasta 0,902 (Ona et al., 2003).

En cuanto a la espectroscopía NIR, en los últimos años se han realizado diversos trabajos relacionados con la medición de las propiedades de la madera, obteniendo excelentes resultados con métodos de laboratorio (Schimleck et al., 2005). En la estimación de la densidad básica se han alcanzado índices de correlación en validación de hasta 0,84 en *Pinus taeda* (Mora et al., 2008), 0,85 en *Pseudotsuga menziesii* (Pino Oregón) (Acuna et al., 2007) y 0,8 en *Eucalyptus globulus* (Schimleck et al., 1999). Como se mencionó anteriormente, estos estudios utilizan técnicas e instrumentos complejos, propios de un laboratorio bien equipado e implican, en la mayoría de los casos, la destrucción del árbol en la toma de muestras. Por lo tanto, no reemplazan fácilmente las técnicas tradicionales para la medición de las propiedades de la madera ya que mantienen el lento proceso de toma y preparación de muestras en laboratorio.

El objetivo de este estudio es implementar técnicas capaces de estimar con exactitud las características deseadas, con una mínima intervención del árbol y que además puedan ser llevadas a terreno, realizando la medición sobre el árbol en pie, sin necesidad de destruirlo o dañarlo significativamente.

Este trabajo se inicia con una revisión de los aspectos teóricos relacionados con la investigación (capítulo 2). Luego se detallan los materiales y métodos empleados en el desarrollo de cada etapa de este estudio (capítulo 3). Posteriormente se presentan los resultados obtenidos (capítulo 4) y finalmente, las conclusiones y propuestas para mejorar los métodos empleados y los pasos necesarios para traspasar estos resultados al desarrollo de un instrumento portátil (capítulo 5 y 6).

2 MARCO TEÓRICO

2.1 Densidad de la madera

Existe una tendencia mundial a la mecanización de las operaciones de cosecha en la industria forestal. Los nuevos equipos de cosecha vienen equipados con diferentes sensores que permiten estimar las características físicas de los rollizos, generalmente diámetro y longitud. El estudio de nuevas tecnologías para medir la calidad de la madera en terreno está progresando con diferentes niveles de éxito, ejemplos de éstos son: acústica, escáner láser y óptico, rayos X, microondas, ultrasonido, resonancia electromagnética y espectroscopía NIR (Tian 1999, So et al. 2002, Carter et al. 2005). Existen diversas propiedades de la madera que afectan la calidad final de los productos. De éstas, la densidad básica es considerada una de las características físicas más importantes, la cual puede ser utilizada para predecir resistencia, rigidez, dureza y calidad de los productos fabricados a partir de esta materia prima (Acuna et al., 2007).

El contenido de humedad de la madera afecta la relación entre masa y volumen. Por esto se definen distintas densidades, entre ellas la Densidad Básica (DB), que es la relación entre la masa de la Madera Seca (MS) y su Volumen Verde (VV) (saturado de agua).

$$DB = \frac{MS}{VV} [Kg \cdot m^{-3}] \quad (2.1)$$

La densidad básica se emplea como un estimador del material leñoso de una especie y del rendimiento que alcanzará en el pulpaje, así como para determinar la facilidad con la que ésta se puede trabajar (cepillar, cortar, moldular, secar, impregnar). Esta propiedad varía entre especies, entre individuos de una misma especie, con la edad de cada ejemplar y con las condiciones edafoclimáticas y de manejo silvícola presentes en

la plantación (Zobel et al., 1989). Una correcta evaluación de la densidad básica en tiempo real representa un verdadero desafío para los productores de madera, quienes desean segregar la cosecha en diferentes clases basándose en esta propiedad interna.

2.2 Espectroscopía Raman

La espectroscopía Raman es una técnica fotónica de alta resolución que proporciona en pocos segundos información química y estructural de casi cualquier material permitiendo así su análisis. Se basa en la dispersión inelástica de la luz, postulada por Smekel en 1923 y demostrada por Chandrasekhara Venkata Raman en 1928. En el experimento original se enfocó el sol a través de un telescopio en unas muestras de agua y alcohol purificadas. A través de un sistema de filtros se recolectó la luz dispersada por las muestras y se logró comprobar la existencia de fotones con frecuencias distintas a la de la luz original. Por este descubrimiento, Raman recibió el premio Nobel de Física en 1930 (Singh, 2002).

En la interacción de la luz con la materia, pueden ocurrir tres fenómenos distintos: Los fotones incidentes son absorbidos, dispersados o no interactúan y pasan a través de ella. En el primer caso la molécula es llevada de un estado electrónico fundamental a un estado electrónico excitado donde la diferencia energética entre estos dos estados corresponde a la energía de los fotones absorbidos. Este fenómeno es utilizado en la espectroscopía de absorción, donde se mide la pérdida de intensidad del haz de luz que incide sobre una muestra. En este fenómeno, la absorción del fotón se produce cuando la diferencia de energía entre dos estados electrónicos “estables” de la molécula es igual a la energía del fotón incidente. En el segundo caso los fotones incidentes son dispersados por la molécula, en este caso las moléculas de la muestra absorben y emiten fotones. Dependiendo de la energía del fotón emitido con respecto a la energía del fotón absorbido, se definen dos fenómenos distintos: Dispersión Rayleigh, cuando se emiten fotones de igual energía a los absorbidos, y dispersión Raman, cuando la energía de los fotones emitidos es distinta a la de los absorbidos. En la dispersión Raman, los fotones

interactúan con la nube de electrones de la molécula y la llevan a un estado electrónico “virtual” el cual es rápidamente abandonado y la molécula vuelve a uno de sus estados electrónicos “estables”, se trata de un fenómeno inelástico de intercambio de energía entre los fotones y la nube de electrones de la molécula. La energía del fotón emitido depende de la diferencia de energía entre el estado electrónico inicial y final de la molécula excitada. Cuando los fotones emitidos tienen mayor energía (menor longitud de onda) que los incidentes se habla de dispersión Anti-Stokes y en el caso contrario de dispersión Stokes. Ambos efectos son débiles ya que ocurren en uno de cada 10^6 - 10^8 fotones dispersados. La diferencia de energía entre los fotones incidentes y los fotones emitidos se traduce en un cambio en la energía vibracional de las moléculas, cuyos estados dependen de la masa de los átomos que la componen y la fuerza de unión de sus enlaces, por lo tanto, los fotones emitidos guardan una particular relación con la naturaleza de la muestra iluminada, obteniendo a través del espectro Raman una “huella digital” que representa el conjunto de frecuencias característico de la estructura molecular de la muestra. A temperatura ambiente es más intensa la dispersión Stokes que la Anti-Stokes debido a que existen más moléculas en estados vibracionales de baja energía.

El fenómeno de dispersión más común en la naturaleza es la interacción elástica entre los fotones y moléculas. Esto ocurre cuando la energía de los fotones emitidos es igual a la energía de los fotones absorbidos y la molécula excitada vuelve al mismo estado electrónico en el cual se encontraba previo a la interacción. Este fenómeno se conoce como dispersión Rayleigh y su efecto sobre la luz solar en la atmósfera es la principal razón de que el cielo sea azul.

Existe un tercer fenómeno de dispersión estrechamente relacionado con la espectroscopía Raman pero muy distinto en su naturaleza: La fluorescencia. Este efecto ocurre cuando un fotón incide sobre una molécula y ésta pasa a un estado electrónico excitado (no virtual) donde permanece unas decenas de nanosegundos, para luego saltar a otro estado electrónico excitado pero de menor energía, liberando un fotón de frecuencia menor al incidente. En los espectros Raman la fluorescencia suele

representarse como una suave curvatura de la línea base. En la figura 2-1 se muestra un diagrama de los procesos de dispersión descritos.

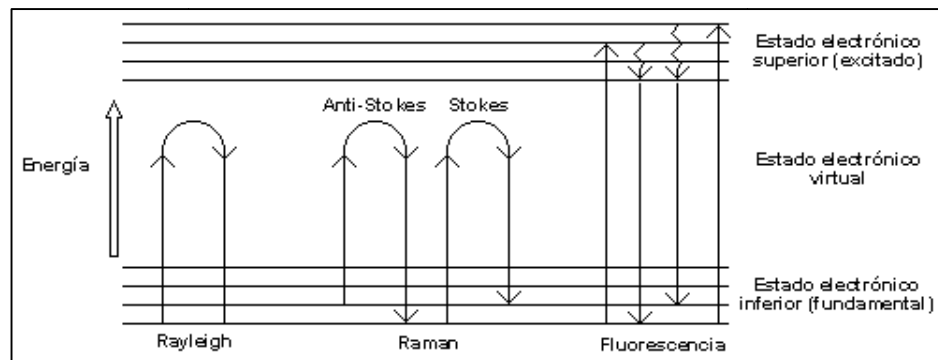


Figura 2-1: Efecto Rayleigh, Raman y Fluorescencia

La intensidad de la luz dispersada, tanto en los efectos Raman, Rayleigh y fluorescencia, es proporcional a la potencia del haz de luz incidente y del número de moléculas iluminadas.

2.2.1 Espectro Raman

El espectro Raman se compone de todo el conjunto de fotones dispersados por la muestra y generalmente se representa como una banda central intensa de la misma frecuencia del láser incidente (que corresponde a los fotones dispersados por el efecto Rayleigh) y bandas laterales de menor intensidad, que corresponden al efecto Raman. Estas bandas se separan en dos grupos: las que corresponden al efecto Raman Anti-Stokes, de energía superior a la del láser incidente y las que corresponden al efecto Raman Stokes, de energía inferior a la del láser incidente. Las bandas Stokes y Anti-Stokes son simétricas con respecto a la banda central Rayleigh y aparecen en pares, aportando cada una la misma información. Sin embargo, la banda Stokes es generalmente más intensa que la Anti-Stokes, lo que la hace más fácil de detectar.

Normalmente no se utiliza la banda Anti-Stokes y el espectro Raman se compone de la banda Rayleigh y las bandas Raman Stokes.

Por convención, las bandas en el espectro Raman no se representan por su longitud de onda λ , ni por su frecuencia ν , sino por su número de onda $\bar{\nu}$. El número de onda es una unidad de frecuencia que indica el número de veces que vibra una onda en una unidad de distancia. Sus unidades en el sistema internacional son los ciclos por metro, pero en la espectroscopía resulta más conveniente utilizar los ciclos por centímetro. El número de onda se obtiene como el inverso de la longitud de onda de los fotones dispersados, expresada en centímetros. La relación entre el número de onda, la longitud de onda, la velocidad de la luz c y la frecuencia de los fotones dispersados se expresa como:

$$\bar{\nu} = \frac{1}{\lambda} = \frac{\nu}{c} \quad [cm^{-1}] \quad (2.2)$$

En el espectro Raman se suelen representar las bandas Stokes y Anti-Stokes con unidades desplazadas con respecto al número de onda del láser de excitación $\bar{\nu}_0$ (banda Rayleigh). Esta representación recibe el nombre de desplazamiento Raman (*Raman Shift*).

$$\text{Raman Shift} \equiv \bar{\nu} - \bar{\nu}_0 \quad [cm^{-1}] \quad (2.3)$$

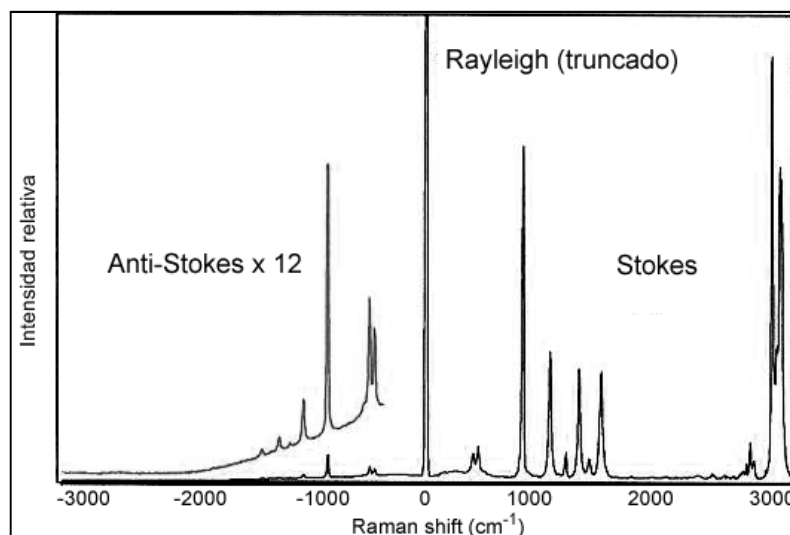


Figura 2-2: Bandas Anti-Stokes, Rayleigh y Stokes en un espectro Raman

La información Raman se puede utilizar para realizar estudios cuantitativos y cualitativos de una muestra. Los estudios cuantitativos se basan en la intensidad de las bandas que componen el espectro completo, mientras que el estudio cualitativo busca determinar los componentes presentes en la muestra al comparar su espectro con otros contenidos en bases de datos de referencia. El conjunto de bandas Raman es una característica unívoca del material analizado y una adquisición precisa de la posición de cada banda es fundamental para la determinación de la estructura molecular que lo compone. Lo anterior no siempre es fácil ya que existen múltiples fuentes de ruido capaces de enmascarar completamente al efecto Raman. La calidad de la información obtenida del espectro de una muestra depende no sólo de las bandas Raman detectadas, sino también de la relación de intensidades entre la señal útil y las distintas fuentes de ruido inherentes al proceso de medición y a la naturaleza de la muestra analizada. En general los fotones dispersados inelásticamente son escasos y pueden ser fácilmente cubiertos por el ruido. En la adquisición del espectro Raman intervienen diferentes fuentes de ruido, las más importantes se nombran a continuación:

- i. **Ruido de señal (*signal noise*):** En ausencia de otras fuentes de ruido, la desviación estándar σ_s del número de fotones dispersados inelásticamente S está determinada por el ruido de señal:

$$\sigma_s = \sqrt{S} \quad (2.4)$$

Este límite es fundamental en la espectroscopía Raman y ocurre en el conteo de eventos aleatorios Poisson. Dicho de otra manera, la desviación estándar del conteo de cualquier evento aleatorio es igual a la raíz cuadrada del número de eventos contados. Como analogía se puede considerar el conteo de autos en una autopista cuando los autos pasan aleatoriamente, pero a una tasa constante. Si se cuentan sólo unos pocos autos la desviación estándar relativa (σ_s / S) es alta. Al aumentar el número de autos contados (incrementando el tiempo de conteo), la desviación estándar σ_s aumenta, pero la desviación estándar relativa σ_s / S disminuye y la razón señal a ruido $SNR = S / \sigma_s$ (*Signal to Noise Ratio*) aumenta. La precisión en la estimación de la tasa autos/minuto claramente mejora al considerar tiempos de conteo mayores.

El conteo de fotones dispersados sigue la misma lógica. Si el ruido de señal es la única fuente de perturbación presente en la adquisición del espectro y consideramos que $S = S' \cdot t$ obtenemos la siguiente expresión para SNR :

$$SNR = S / \sigma_s = \sqrt{S} \quad (2.5)$$

$$SNR = \sqrt{S' \cdot t} \quad (2.6)$$

En el ruido de señal observamos un aumento de SNR con $\sqrt{t}$. Es importante destacar que utilizando un espectrómetro perfecto, que sólo incluye esta fuente de ruido, la SNR sigue estando limitada por la ecuación 2.5. En la figura 2-3 se muestran tres espectros Raman de ascorbato de Calcio (una forma menos ácida de vitamina C) donde se observa como aumenta la SNR al aumentar el tiempo de integración. En este caso la SNR está determinada principalmente por el ruido de señal.

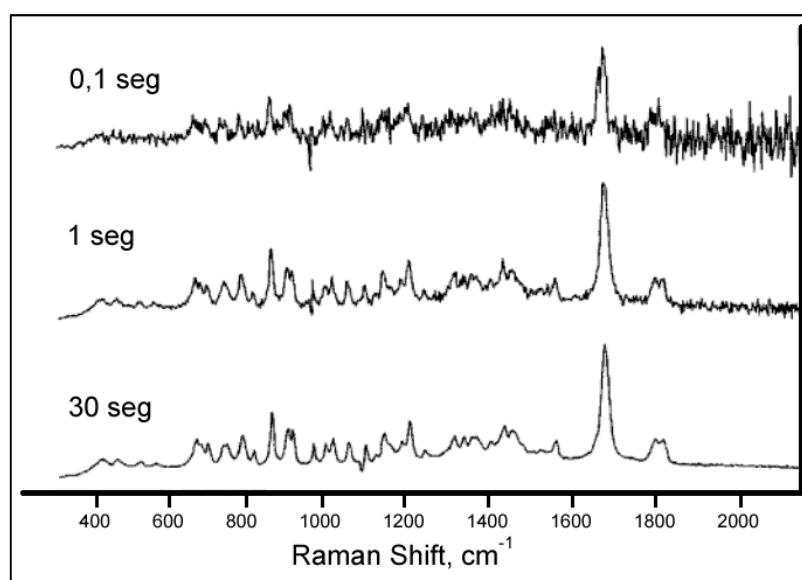


Figura 2-3: Aumento de la SNR con un aumento del tiempo de integración

- ii. **Ruido de fondo (*background noise*):** Este fenómeno se refiere a cualquier fotón detectado proveniente del láser o de la muestra que no tenga relación con el efecto Raman. El ruido de fondo incluye la fluorescencia, dispersión Rayleigh, emisión térmica de la muestra y de los instrumentos, reflejos de la luz en cuerpos extraños como partículas de polvo presentes en la óptica y cualquier otro tipo de dispersión que no haya sido filtrada. Es importante destacar que esta fuente de ruido depende de la cantidad de moléculas

iluminadas y de la intensidad del láser incidente. En la fluorescencia, el fotón incidente es absorbido por la molécula y ésta pasa a un estado electrónico excitado. Después, la molécula salta a otro estado electrónico excitado, pero de menor energía que el anterior, emitiendo en este proceso un fotón de menor energía (mayor longitud de onda) que el absorbido. Este proceso ocurre en millonésimas de segundo por lo que se observa solamente durante la iluminación y es muy común en la espectroscopía Raman. Para reducir la fluorescencia en la adquisición del espectro Raman se utilizan láseres de excitación de elevada longitud de onda cercana a los 1000 nm, dentro del rango NIR.

- iii. **Ruido de oscuridad (*Dark noise*):** Este término se refiere a los fotones o electrones generados espontáneamente en el detector de luz del espectrómetro. Todos los detectores utilizados en la espectroscopía Raman tienen una señal de ruido oscuro asociada principalmente al ruido térmico. También se incluyen en esta clasificación los ruidos generados por píxeles defectuosos, ruido de lectura y la respuesta en frecuencia del detector.

Por los puntos mencionados anteriormente es posible prever que existen muchas fuentes de ruido capaces de cubrir la señal Raman. Para extraer la información útil del espectro, es imprescindible contar con una buena razón señal a ruido. Para lograr esto se utilizan diversas técnicas, tanto en la óptica del espectrómetro, como en el procesamiento computacional de la señal adquirida. En la óptica se utilizan láseres potentes, altamente monocromáticos y de longitud de onda elevada, con el fin de obtener el mayor nivel de señal posible. En el procesamiento computacional del espectro adquirido, se cuenta con numerosos algoritmos que han demostrado ser muy útiles en la reducción del ruido presente en los espectros.

2.3 Espectroscopía NIR

Como alternativa a la espectroscopía Raman, en los últimos años se ha incrementado considerablemente el uso de la espectroscopía NIR para el análisis de productos en la industria alimenticia, bioquímica y farmacéutica. El éxito de esta técnica puede atribuirse en gran parte a su habilidad para realizar análisis rápidos -cuantitativos y cualitativos- de muchos componentes en una muestra. La espectroscopía NIR es una técnica que forma parte del campo de estudio de la espectroscopía molecular, la cual analiza la interacción de la radiación electromagnética con la materia. El rango espectral que estudia la espectroscopía NIR está comprendido entre los 750-2500 nm.

En una molécula se presentan tres diferentes niveles energéticos: vibracionales, rotacionales y electrónicos. Para producir cambios en cada uno de estos niveles es preciso utilizar diferentes cantidades de energía, lo que hace que el efecto de los fotones absorbidos se observe en diferentes regiones del espectro. Cada enlace químico vibra a una frecuencia específica, lo que provoca que absorba energía en una longitud de onda determinada. Los espectros NIR están constituidos por la representación gráfica de la energía absorbida en función de la longitud de onda, esta energía cambia dependiendo del tipo de vibración permitida y de la cantidad y concentración de moléculas en la muestra. El espectro NIR presenta bandas de absorción que corresponden a armónicos o a la combinación de bandas originadas en otros sectores del espectro electromagnético. Por esto, generalmente la información sobre el contenido de un compuesto en la muestra se encuentra distribuida en un amplio sector del espectro y no en longitudes de onda específicas, como es en el caso de la espectroscopía Raman.

2.3.1 Espectro NIR

Existen dos técnicas ópticas diferentes para adquirir el espectro NIR, reflectancia difusa y transmitancia difusa. La primera consiste en captar la radiación que es reflejada

por la muestra mediante un receptor ubicado en el mismo lado que la fuente de iluminación. En la segunda la recolección se lleva a cabo en el lado opuesto al de la fuente de iluminación, es decir, se captura la radiación que atraviesa la muestra. En ambos casos se utiliza una fuente de luz policromática, cuyo espectro debe ser conocido y lo más estable posible para evitar introducir ruido externo en la adquisición del espectro. En la representación del espectro NIR se busca capturar la intensidad relativa detectada (cuentas del detector) versus la longitud de onda correspondiente a esa absorción. En la figura 2-4 se muestra el espectro NIR de una de las muestras de *Eucalyptus nitens* utilizadas en esta investigación.

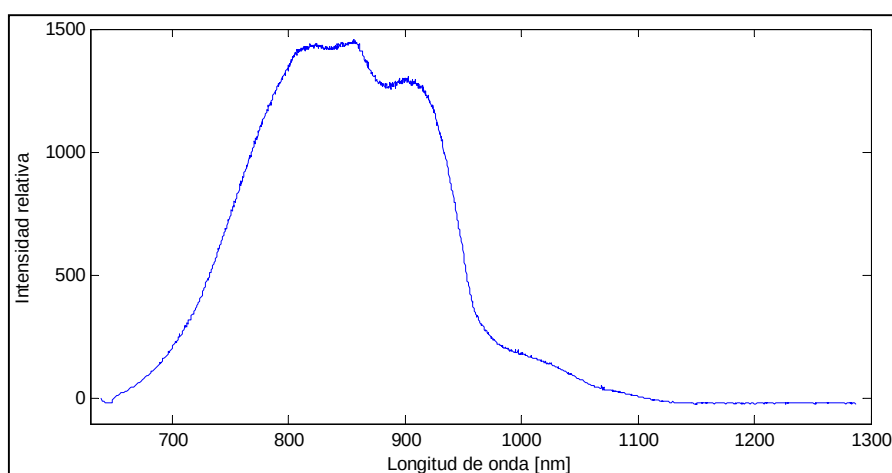


Figura 2-4: Espectro NIR de una muestra de *Eucalyptus nitens*

En la espectroscopía NIR los instrumentos y sondas de medición suelen ser significativamente más simples que en la espectroscopía Raman. Esto se debe a la naturaleza del fenómeno en estudio que consiste en la detección de la intensidad relativa de la luz que atraviesa o es reflejada por la muestra. En cuanto al ruido, las principales fuentes que distorsionan el espectro corresponden al ruido oscuro del detector y a las fuentes externas, como la luz ambiental, que alteran la intensidad del espectro adquirido.

Por este motivo, las mediciones deben realizarse bajo condiciones ambientales reguladas.

2.4 Regresión multivariable PLS

Para entregar información cuantitativa sobre la densidad básica de una muestra a partir de la medición de un espectro de ésta, es necesario establecer una relación entre las variables medidas y la propiedad de interés. Existen casos en que la correlación se hace observando una sola variable (una longitud de onda) y se compara con los cambios en la densidad. Esta es conocida como calibración simple o univariada, donde el modelo es construido a partir de una simple regresión lineal, lo cual conlleva a que el método tenga muchas limitaciones ya que el modelo obtenido para la predicción sólo puede entregar información correcta cuando no existen contribuciones a la variable medida desde otras fuentes y sólo la propiedad objetivo sea la que afecta a la variable.

La calibración multivariada, en cambio, toma en cuenta todas las variables (que en este caso corresponden a las longitudes de onda del espectro adquirido) entregando así una mejor visión del sistema completo, lo cual presenta varias ventajas, por ejemplo, una razón señal a ruido aumentada debido a la información redundante involucrada en el proceso de calibración. Además, este tipo de modelo puede realizar una muy buena cuantificación incluso en la presencia de otros compuestos y variables físicas o químicas que pudieran interferir en la medición. Al tener una visión del espectro completo, es posible cuantificar más de una propiedad.

Por lo tanto, la calibración multivariada puede construir un modelo que compense las especies que interfieren en la medición excepto aquellas que influyen en longitudes de onda relevantes para la estimación. Para construir el modelo es necesario un conjunto de mediciones representativas (conjunto de espectros de calibración) con valores conocidos de la propiedad de interés que se intenta predecir (densidad básica de referencia de cada muestra). Existen varias técnicas de calibración para construir modelos multivariable como por ejemplo, CLS (*Classical Least Squares*), ILS (*Inverse*

Least Squares) y MLR (*Multiple Linear Regression*). Sin embargo, hoy en día se ha hecho muy popular el uso de métodos más complejos como son el PLS (*Partial Least Squares*), PCR (*Principal Component Regression*) y RR (*Ridge Regression*).

La calibración por mínimos cuadrados parciales (PLS) fue introducida por Herman Wold en 1975 para ser aplicada en ciencias económicas y sociales. Sin embargo, gracias a las contribuciones de su hijo Svante Wold (Wold et. al., 2001), ganó popularidad en el área de la química conocida como Quimiometría o *Chemometrics*, en donde se analizan datos que se caracterizan por muchas variables, como es el caso de la espectroscopía. Se ha demostrado que el método PLS es una buena alternativa a otros métodos más clásicos de regresión multivariable, principalmente por ser más robusto, en el sentido de que los parámetros del modelo no cambian demasiado cuando se introducen nuevas muestras de calibración (Geladi y Kowalski, 1986).

La idea motivadora de PLS fue heurística, por este motivo, algunas de sus propiedades son todavía desconocidas, a pesar de los progresos alcanzados por Hellan (1988), Hoskuldson (1988), Stone y Brooks (1990), entre otros (Geladi y Kowalski, 1986). La metodología PLS generaliza y combina características del Análisis de Componentes Principales (*Principal Component Analysis*, PCA) y Análisis de Regresión Múltiple (*Multiple Regression Analysis*, MRA). La demanda de esta metodología y la evidencia de sus resultados, han ido en aumento, convirtiéndose así en uno de los métodos de calibración más utilizados en la actualidad para el análisis de espectros.

Los métodos de calibración multivariable que utilizan modelos de regresión lineal parten desde el supuesto, de que existe una relación lineal entre el espectro y alguna característica de la muestra analizada que pueda ser predicha. Básicamente, los métodos de regresión lineal establecen un modelo como un polinomio de la siguiente forma:

$$y = b_0 + b_1x_1 + \cdots + b_nx_n + e \quad (2.7)$$

Donde y es la variable predicha, b_0 es el *offset*, b_n los coeficientes de regresión, x_n la respuesta medida de la variable (longitud de onda) n y e es el residuo. También podemos representarlo matricialmente como:

$$\mathbf{y} = \mathbf{X} \cdot \mathbf{b} + \mathbf{e} \quad (2.8)$$

Donde $\mathbf{y}$ es un vector columna con las mediciones de calibración (valores de referencia), $\mathbf{X}$ es la matriz de espectros (m mediciones por n longitudes de onda) y $\mathbf{b}$ el vector de regresión. Para obtener este vector, es necesario calcular la inversa $(\mathbf{X}^T \mathbf{X})^{-1}$ por lo que ésta debe existir, es decir, el determinante no puede ser cero. Esto agrega la restricción de que el número de mediciones debe ser igual o mayor que el número de variables involucradas en la calibración, lo cual muchas veces es un impedimento. En estos casos es conveniente utilizar métodos como PLS, PCR o RR donde este requerimiento no se hace necesario.

El PLS realiza una regresión lineal luego de descomponer la matriz de espectros $\mathbf{X}$ en matrices de *scores* y factores PLS o variables latentes. Esto puede ser interpretado como una proyección en un nuevo sistema de coordenadas con menores dimensiones que el espacio original de variables independientes. Las Variables Latentes (VL) son determinadas por la máxima varianza de las variables independientes que son relevantes para predecir las variables dependientes. Existen varios algoritmos para calcular un modelo PLS, siendo los más comunes el SIMPLS y NIPALS (Geladi et. al., 1986). Este último es el más intuitivo y su aproximación matemática para el caso de una variable de predicción, se puede representar como:

$$\mathbf{X} = \mathbf{T} \mathbf{P}^T + \mathbf{e} \quad (2.9)$$

$$\mathbf{y} = \mathbf{U} \mathbf{q}^T + \mathbf{f} \quad (2.10)$$

Donde $\mathbf{e}$ y $\mathbf{f}$ son las matrices residuales, $\mathbf{T}$ y $\mathbf{U}$ las matrices de *scores*, y $\mathbf{P}^T$ y $\mathbf{q}^T$ los *loadings* o matrices de factores. El objetivo del PLS es modelar la relación entre los *scores* de la matriz $\mathbf{X}$ y los *scores* del vector $\mathbf{y}$, de tal manera de encontrar factores capaces de capturar varianza y además lograr correlación. Un conjunto de vectores conocidos como *weights* $\mathbf{W}$, es el que establece las relaciones entre los *scores* de $\mathbf{X}$ e $\mathbf{y}$.

$$\mathbf{U} = \mathbf{T}\mathbf{W} \quad (2.11)$$

Todas estas matrices son calculadas secuencialmente en un proceso iterativo, donde se busca la convergencia para calcular el vector de relaciones para cada variable latente. Cuando el modelo completo es calculado, el vector de regresión resultante se obtiene de la siguiente forma:

$$\hat{\mathbf{b}} = \mathbf{P}(\mathbf{P}^T\mathbf{P})^{-1}\mathbf{W}\mathbf{q}^T \quad (2.12)$$

$$\mathbf{y} = \mathbf{X}\hat{\mathbf{b}} \quad (2.13)$$

SIMPLS entrega exactamente el mismo resultado que NIPALS para regresiones simples y una mínima diferencia en el resultado para regresiones multivariadas. El objetivo explícito de SIMPLS es maximizar la covarianza, además este algoritmo es más rápido y con mejores resultados, ya que está basado en el criterio de covarianza mientras que NIPALS no.

Generalmente, los parámetros estimados para los modelos de calibración son elegidos usando criterios que minimicen el error de calibración y validación. Existen cuatro parámetros clásicos para determinar el nivel de error de predicción de un modelo:

- RMSEC (*Root Mean Square of Error of Calibration*): Mide el nivel de error entre la respuesta del modelo de calibración $\hat{y}_i$ y su correspondiente valor de

referencia y_i para las m muestras utilizadas en la calibración del modelo. Su fórmula es:

$$\text{RMSEC} = \sqrt{\frac{1}{m} \sum_{i=1}^m (y_i - \hat{y}_i)^2} \quad (2.14)$$

- **RMSEV o RMSEP (*Root Mean Square Error of Validation or Prediction*)**: Mide el nivel de error entre los valores predichos $\hat{y}_i$ y sus respectivos valores de referencia y_i para el conjunto de p muestras de validación. Su valor es calculado como:

$$\text{RMSEV} = \sqrt{\frac{1}{p} \sum_{i=1}^p (y_i - \hat{y}_i)^2} \quad (2.15)$$

- **RMSECV (*Root Mean Square Error of Cross-Validation*)**: Al igual que el anterior, mide el error de predicción, pero utilizando el método de validación cruzada para las muestras de calibración. Este método separa el conjunto de muestras en varios subconjuntos y realiza iterativamente una calibración dejando uno de estos conjuntos fuera para usarlo como validación. El método de validación cruzada es uno de los más conocidos para evaluar el número de variables latentes que se usará en el modelo final.
- **R^2 (Índice de Correlación Múltiple)**: Indica el porcentaje de variación de la variable dependiente explicado por el modelo. Existe un índice de correlación para el conjunto de calibración R_c^2 y uno para el conjunto de validación R_v^2 . Estos son calculados de la siguiente manera:

$$R^2 = \left(\frac{\sigma_{\hat{y}y}}{\sigma_{\hat{y}}\sigma_y} \right)^2 \quad (2.16)$$

Donde σ_y corresponde a la desviación estándar de los valores de referencia, $\sigma_{\hat{y}}$ a la desviación estándar de los valores predichos y $\sigma_{\hat{y}y}$ a la covarianza entre los valores de referencia y predichos.

3 MATERIALES Y MÉTODOS

A continuación se detallan los materiales y métodos utilizados para la obtención de los modelos de regresión multivariable que permiten la estimación de la densidad básica en la madera de *Eucalyptus nitens* a partir de su espectro Raman y NIR.

3.1 Muestras

Las muestras utilizadas en el proceso corresponden a tarugos de *Eucalyptus nitens* de 8 años de edad provenientes de distintos predios de Forestal Mininco S.A., ubicados en la VIII región. Estas muestras fueron obtenidas utilizando un taladro de extracción que perfora el árbol de lado a lado a la altura del DAP (Diámetro Altura Pecho, 1,3 m sobre el suelo), obteniendo tarugos de 12 mm de diámetro y 15 cm de largo en promedio. Inmediatamente después de ser extraídos fueron envueltos con Alusaplast para conservar la humedad. En total se obtuvieron 157 tarugos.



Figura 3-1: Extracción del tarugo y sus dimensiones

Debido a las características de la sonda del espectrómetro Raman utilizado, fue necesario realizar un corte longitudinal a cada tarugo, con el fin de presentar una cara

plana donde apoyar la sonda y evitar la entrada de luz externa. Este corte se efectuó con una sierra de banda y se tuvo especial cuidado en realizarlo de manera perpendicular al sentido de las fibras ya que estas actúan como guías de luz. Por otro lado, debido a la potencia por área del láser del espectrómetro Raman, las muestras se quemaban al ser iluminadas, haciendo imposible la adquisición del espectro y dañando permanentemente los tarugos. Para solucionar este problema se saturaron los tarugos con agua destilada a 15°C durante 24 horas. Este no es un problema en el eventual instrumento portátil a desarrollar en el futuro, pues la madera se encuentra en estado saturado en el árbol en pie.

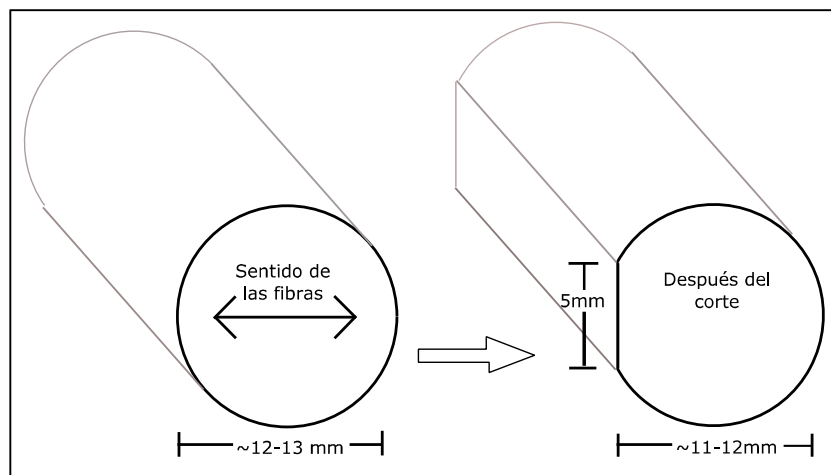


Figura 3-2: Corte longitudinal realizado a los tarugos

3.2 Espectrómetro Raman

Los espectrómetros Raman modernos cuentan con una serie de elementos ópticos que permiten observar el débil efecto Raman a partir de una gran variedad de muestras. Así como actualmente existe una extensa gama de aplicaciones para la espectroscopía Raman, existe una gran variedad de instrumentos capaces de captar este efecto. El funcionamiento básico se diferencia por la complejidad del diseño óptico, y por la

calidad y capacidad de los componentes internos. En la gran mayoría de los casos, la muestra es iluminada usando un rayo láser, la luz dispersada se recoge con un lente y se envía a un detector a través de un elemento dispersor. Las longitudes de ondas cercanas a la del láser, dispersadas debido al efecto Rayleigh, son filtradas, mientras que el resto de la luz se dispersa en el detector. Un esquema generalizado de un espectrómetro Raman se muestra en la figura 3-3.

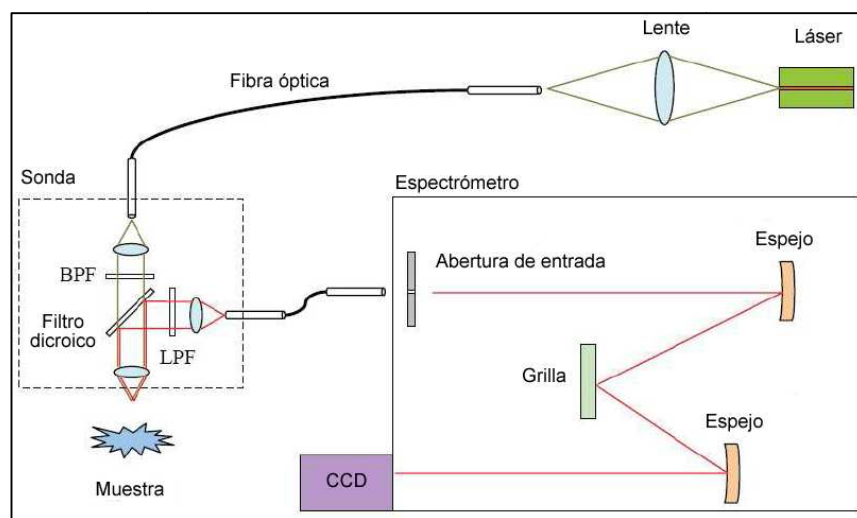


Figura 3-3: Diagrama esquemático de un espectrómetro Raman

Un espectrómetro Raman se compone de diversos elementos que afectan directamente la calidad y características de los espectros adquiridos.

- **Fuente de excitación:** Generalmente se utiliza un láser en la espectroscopía Raman moderna debido a la alta energía que se requiere para producir la dispersión Raman con la intensidad suficiente para poder detectarla. Los parámetros más importantes del láser son su longitud de onda, el ancho espectral del haz y la potencia. La longitud de onda debe ser elegida considerando la intensidad de la fluorescencia de la muestra, con el fin de evitar que enmascare el efecto Raman que se desea observar. Para mejorar la resolución del espectro

adquirido, el láser debe tener un ancho espectral tan estrecho como sea posible y debe contar con una potencia que produzca una señal detectable pero que no dañe la muestra.

- **Sonda:** La función principal de la sonda es dirigir el haz del láser hacia la muestra, recoger la señal y separarla de la intensa dispersión Rayleigh producida por el láser, la cual es una de las principales dificultades ópticas de esta técnica espectroscópica. Teniendo en cuenta que la dispersión Raman es $10^{-6} - 10^{-8}$ veces menos intensa que la dispersión Rayleigh, son necesarias algunas técnicas de filtrado especiales. La más utilizada consiste en filtros rechaza-banda (*Notch*), los cuales dejan pasar todo el espectro, excepto una pequeña banda, la cual se centra en la longitud de onda del láser para así eliminar la dispersión Rayleigh. Otra técnica utilizada cada vez menos, es incorporando dos o tres monocromadores los cuales se configuran de tal forma que separen la dispersión Raman de las otras radiaciones y además aumenten la separación de los peaks Raman entre sí.

Ambas técnicas son eficientes presentando cada una tanto ventajas como desventajas. Las sondas basadas en filtros resultan más simples y menos costosas ya que aprovechan el avance tecnológico y la efectividad de los filtros rechaza banda. Las sondas que usan monocromadores presentan la ventaja de permitir ajustar en un rango la frecuencia del láser utilizado, así como observar el espectro cercano de la línea de la excitación láser.

En la sonda, además, se encuentran otros filtros los cuales cumplen la función de mejorar la calidad de la señal Raman eliminando cualquier otro tipo de interferencia.

- **Elemento de dispersión:** Los espectrómetros Raman convencionales utilizan una configuración basada en la dispersión *Czerny-Turner* (ver figura 3-4). En ésta, la radiación incidente pasa primero por una rendija de entrada para luego

ser colimada en un espejo convexo. La radiación, una vez colimada, es dispersada por una grilla de difracción, y luego es recogida por otro espejo convexo, formándose así múltiples imágenes de la radiación incidente a diferentes longitudes de onda en un detector CCD. Las intensidades relativas de esos componentes de longitud de onda son recibidas por el detector, para formar así el espectro de la muestra examinada. La resolución espectral del instrumento viene determinada por una serie de parámetros, siendo los más importantes el número de pixeles del CCD, la densidad de la grilla, el ancho de la rendija de entrada y la aberración espectral de los espejos.

Otras configuraciones más complejas usadas en algunos espectrómetros pueden ser del tipo *Echelle* o con interferómetro de transformada de Fourier.

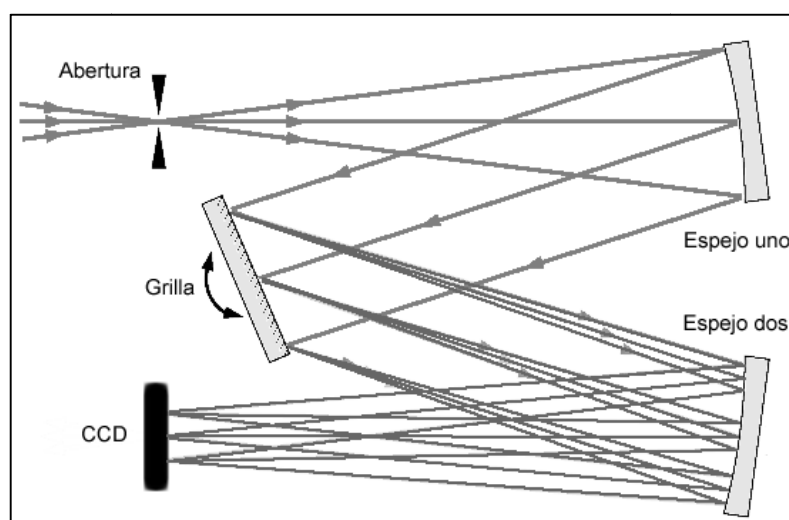


Figura 3-4: Configuración de Czerny-Turner

- **Detector:** Hoy en día casi la totalidad de los espectrómetros Raman utilizan un CCD (*Charge Coupled Device*) como detector. Éste cumple la función de convertir los fotones recibidos en una señal eléctrica, la cual es enviada a través de la conexión del espectrómetro a un computador u otro equipo portátil para ser procesadas y así realizarle el análisis correspondiente. Los principales

parámetros de un detector CCD son el número de píxeles, el cual afectará la resolución del espectro; y la respuesta en frecuencia del detector, la cual define el rango espectral del instrumento.

El espectrómetro Raman utilizado en esta investigación es el BTR111 MiniRam con la sonda BAC100; ambos producidos por B&W Tek Inc. El BTR111 es un espectrómetro Raman miniatura, rápido, ligero y portable, características esenciales para esta investigación. Posee un detector CCD de 2048 píxeles, está equipado con un láser NIR de 785nm y cubre el rango espectral entre los 150 – 3100 cm^{-1} . Se comunica con un computador vía USB que permite configurar los parámetros de la adquisición del espectro, por ejemplo, el tiempo de integración. Se alimenta con una fuente externa que puede ser reemplazada por una batería portable.

La sonda BAC100 posee dos fibras ópticas, una de excitación con un núcleo de 105 μm y otra de colección cuyo núcleo es de 200 μm , ambas son de 150 cm de largo. En el extremo manipulado por el usuario, la sonda cuenta con un obturador mecánico que abre y cierra el paso de la luz.

Tabla 3-1: Características técnicas del espectrómetro Raman BTR111 MiniRam

Rango espectral	175 -3100 cm^{-1} (<i>Raman shift</i>)
Resolución óptica	10 cm^{-1} FWHM
Resolución digital	16 bit, 250kHz
<i>Dark Noise</i>	< 60 RMS a 9ms
Long. Onda láser	785 nm (± 1 nm)
Ancho haz láser	0,2 nm FWHM
Potencia láser	> 300 mW
Tiempo integración	9-65535 mseg
Detector	CCD, 2048 píxeles

Dimensiones	22,6 x 16,2 x 8,2 cm
Peso	2 kg
Interfaz PC	USB



Figura 3-5: Espectrómetro Raman BTR111 y sonda BAC100

3.3 Espectrómetro NIR

El espectrómetro NIR utilizado corresponde al USB2000 fabricado por Ocean Optics Inc. Este espectrómetro recibe la luz transmitida por una fibra óptica y la dispersa mediante una grilla sobre un detector CCD de 2048 píxeles en el rango de 640 –1100 nm. La señal eléctrica se transfiere a un computador vía USB. Este espectrómetro se alimenta directamente del puerto USB, sin necesidad de otra fuente externa.

Tabla 3-2: Características técnicas del espectrómetro NIR USB2000

Rango espectral	640 – 1100 nm
Resolución óptica	0,3-10 nm FWHM
Resolución digital	16 bit

<i>Dark Noise</i>	50 cuentas RMS
Tiempo integración	3 – 65000 mseg
Detector	CCD, 2048 píxeles
Dimensiones	89,1 x 63,3 x 34,4 mm
Peso	0,19 Kg
Interfaz PC	USB



Figura 3-6: Espectrómetro NIR USB2000

3.4 Diseños óptico-mecánicos para la adquisición de espectros

Dependiendo de la técnica espectroscópica utilizada, fue necesario implementar un sistema óptico-mecánico que permitiera estandarizar el proceso de adquisición y minimizar los errores en la toma de espectros. El elemento basal utilizado en cada configuración corresponde a un riel milimétrico que permite un desplazamiento preciso en un eje direccional. En el caso Raman se fijó la sonda al riel con el fin de posicionarla de manera precisa sobre las muestras y en el caso NIR se fijaron los tarugos al riel, los cuales eran desplazados sobre la zona de adquisición del espectro NIR.

3.4.1 Diseño óptico-mecánico Raman

En este caso fue necesario afirmar los tarugos y la sonda de manera tal que se evitara la entrada de luz externa. Con este fin se realizó el montaje presentado en la figura 3-7. La sonda Raman, montada sobre el riel milimétrico, se apoya perpendicular al eje principal del tarugo y se traslada paralela a éste. El tarugo se afirma sobre una base metálica utilizando pinzas de aluminio.

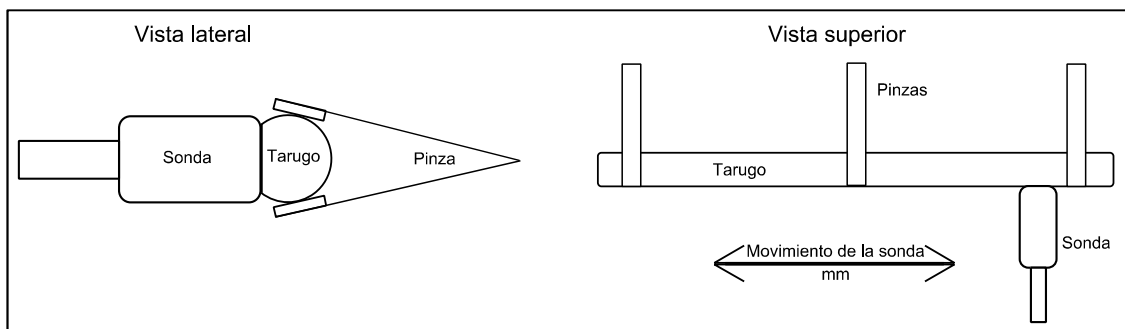


Figura 3-7: Esquema del montaje para la adquisición de espectros Raman

En la figura 3-8 se observa como entra en contacto el tarugo con la sonda Raman. De esta forma se logró evitar la entrada de luz externa en las mediciones realizadas.

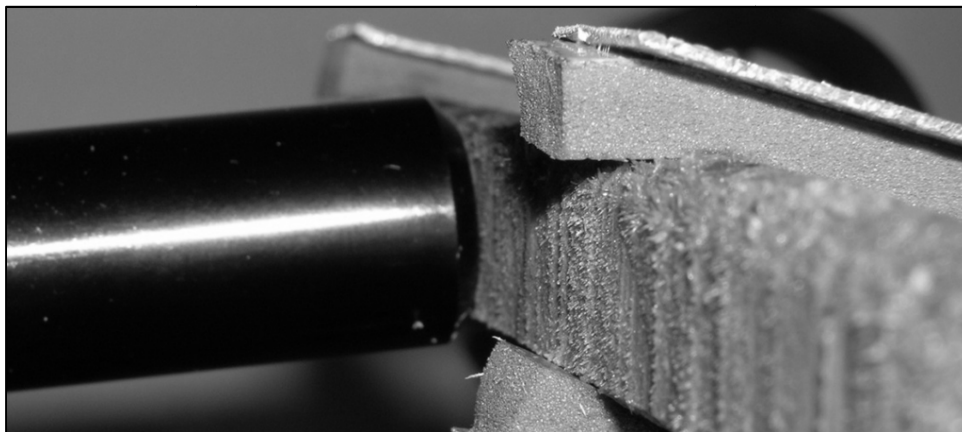


Figura 3-8: Detalle del contacto entre la sonda Raman y el tarugo

3.4.2 Diseño óptico-mecánico NIR

Para la recolección de la luz se utilizó una fibra óptica modelo P400 de Ocean Optics™ con un núcleo de 400 μm de diámetro, cubriendo el rango espectral 300- 1100 nm. En la punta de esta fibra se instaló un colimador transparente en el rango espectral que se está trabajando (640-1300 nm) y como fuente luminosa se utilizó una ampollita halógena de 50 Watts. En este montaje el tarugo adosado al riel milimétrico se desplaza sobre la abertura y se aplica la técnica de adquisición por transmitancia, recolectando la luz que atraviesa la muestra para dirigirla al espectrómetro a través de la fibra óptica.

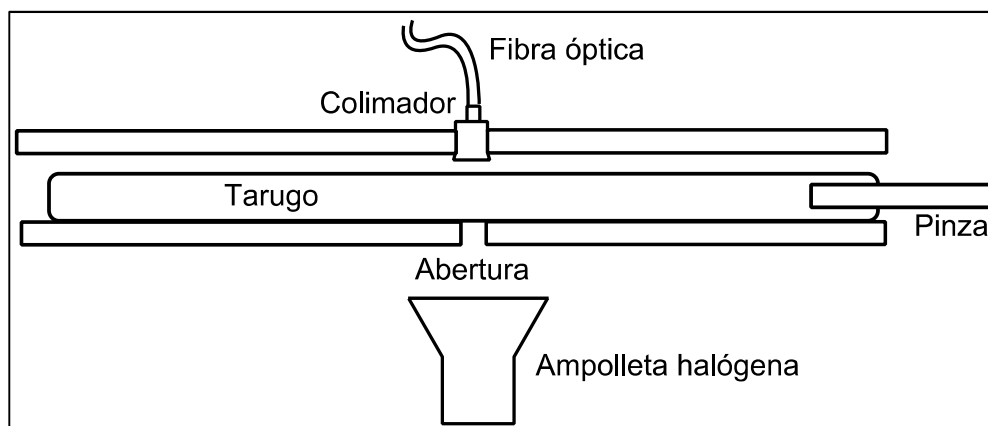


Figura 3-9: Esquema del montaje para la adquisición de espectros NIR

3.5 Método analítico tradicional

El método utilizado para la medición de referencias está basado en la norma Tappi 258 om-92 de la industria de la madera (TAPPI, 2002), para la medición de la densidad básica de la madera, definida como el cuociente entre la masa seca y el volumen verde (ecuación 2.1). El volumen verde se refiere al volumen del tarugo saturado de agua. Para medirlo, se sumergen los tarugos en bidones con agua destilada a 15°C durante 48 horas o hasta que las muestras se hundan en el contenedor.

Posteriormente se mide el volumen del tarugo saturado de manera indirecta a través de la masa de agua que desplaza al ser sumergido en una probeta con agua destilada. Asumiendo que la densidad el agua es 1 g/cc, se obtiene el volumen del tarugo del valor de masa que entrega la balanza.

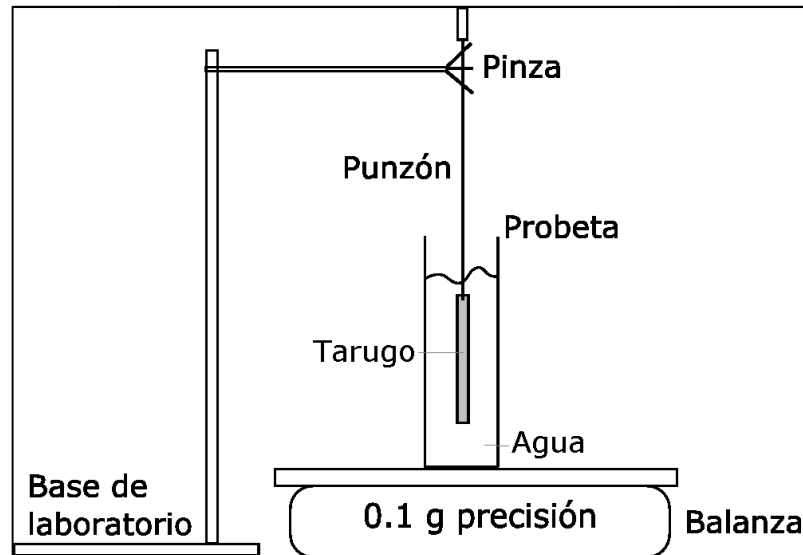


Figura 3-10: Montaje para medición del volumen del tarugo por desplazamiento de agua

Como masa seca se entiende la masa del tarugo después de ser secado a 105 °C por el tiempo necesario hasta que la masa sea estable. Al principio del secado la masa del tarugo disminuirá rápidamente debido a la fácil eliminación de agua que se lleva a cabo. A medida que pasa el tiempo en el secador, la masa tiende a estabilizarse llegando a un punto en que la variación de masa en el tiempo es mínima, este punto se conoce como masa estable. Se determinó que para las muestras de este estudio, el tiempo necesario para alcanzar la masa estable es de 6 horas en el secador a una temperatura controlada de 105 °C. Es importante mencionar que apenas se sacan las muestras secas estas comienzan a absorber humedad del ambiente, es por esto que el proceso de medición de masas debe ser lo más rápido posible y al sacarlas, estas deben ser depositadas en un contenedor impermeable.

Al terminar esta etapa los tarugos sufren modificaciones irreversibles en su estructura física (colapsan las paredes de las células estructurales) y es por esto que la medición de espectros y volumen verde debe ser realizada con anterioridad.

3.6 Adquisición de espectros

Para ambas técnicas espectroscópicas se realizaron mediciones a lo largo del tarugo cada 10 mm, de esta forma se obtuvieron entre 10 a 22 espectros por muestra. Posteriormente se promedian todos los espectros de cada conjunto obteniendo finalmente un espectro Raman y un espectro NIR para cada muestra. Durante la toma de espectros se mantuvieron las condiciones ambientales estables (temperatura, luminosidad, humedad ambiental) y los tarugos se extrajeron directamente del bidón saturador, secándose superficialmente antes de realizar la adquisición del espectro.

3.6.1 Adquisición del espectro Raman

El tiempo de integración del espectrómetro Raman fue definido en 20 segundos. Para establecer esto se hicieron pruebas preliminares con diferentes muestras, aumentando el tiempo de integración pero teniendo cuidado en no saturar el detector, por exceso de fotoelectrones acumulados, ya que si esto ocurre se pierde la información en el rango espectral saturado.

Para la adquisición de espectros Raman de las muestras se realizó el siguiente procedimiento:

- Se extrae el tarugo del bidón saturador y se seca superficialmente.
- Se fija el tarugo con las pinzas del montaje Raman.
- Se acerca la punta de la sonda a uno de los extremos del tarugo hasta que la boquilla entra en contacto con la superficie plana de la muestra, impidiendo la entrada de luz ajena al láser.
- Se adquiere el espectro Raman de la zona iluminada, a través del software del fabricante.
- Se desplaza la sonda 10 mm para una nueva adquisición.

- Se repiten los dos pasos anteriores hasta cubrir el tarugo completo.
- Se devuelve el tarugo al bidón saturador.

Al término del proceso anterior se cuenta con un conjunto de espectros Raman asociados al tarugo utilizado, guardados como archivos de texto en un computador.

3.6.2 Adquisición del espectro NIR

En el caso NIR el tiempo de integración fue menor: 250 ms. Esto porque la intensidad de la luz halógena satura al detector para tiempos mayores. Los tarugos fueron colocados con la cara plana enfrentando la fuente de luz halógena y se desplazaron en pasos de 10 mm, para cada espectro adquirido. Los pasos en la adquisición son similares al caso Raman y el resultado final del proceso es un conjunto de archivos de texto que representan los espectros NIR de la muestra.

3.7 Procesamiento de espectros

Existen algunas técnicas matemáticas para preparar los espectros, las cuales se aplican a los datos que serán utilizados en la calibración del modelo multivariable con el fin de eliminar la información redundante y aumentar las diferencias relevantes entre cada una de las muestras. En muchos casos es fundamental filtrar computacionalmente los espectros obtenidos antes de calibrar los modelos de regresión, ya sea para eliminar el ruido que existe en el proceso de medición o en el caso en que se desee suprimir algún otro fenómeno, como por ejemplo, el efecto de la fluorescencia en los espectros Raman o bandas con poca información relevante en el caso NIR. Ciertas técnicas de procesamiento fueron particularmente útiles en este estudio y mejoraron notoriamente los resultados obtenidos, cada una de ellas se describe a continuación.

3.7.1 Promedio y sustracción del espectro de oscuridad (dark)

Como se dijo anteriormente, se obtuvieron dos espectros por muestra, uno Raman y uno NIR. Cada uno de estos espectros proviene del promedio de todos los espectros obtenidos (uno cada 10 mm) a lo largo del tarugo. Al promediar los espectros se busca eliminar el ruido aleatorio presente en la adquisición.

Otra fuente importante de ruido que debe ser eliminada corresponde al espectro de oscuridad del espectrómetro utilizado, este fue definido en 2.2.1.iii para el caso de la espectroscopía Raman. Para adquirir este espectro, se debe realizar una medición con el instrumento operando sin fuente de excitación (obturador de la sonda cerrado en el caso Raman). Este espectro puede ser un promedio de varios espectros de oscuridad adquiridos en distintos momentos durante la medición. El resultado final es posteriormente restado del espectro promedio de cada tarugo, con el fin de eliminar cualquier información que pudiera interferir desde la sonda o del instrumento. En este estudio se utilizó un promedio de cinco espectros de oscuridad por instrumento. En la figura 3-11 se observa el espectro dark de cada instrumento. Es interesante notar que en términos de unidades de intensidad relativa, el ruido de oscuridad Raman es mucho más intenso que el ruido del espectrómetro NIR. Esto se debe principalmente a que el instrumento Raman está calibrado para detectar un efecto de muy baja intensidad, lo que lo hace más sensible al ruido que se filtra en la adquisición del espectro.

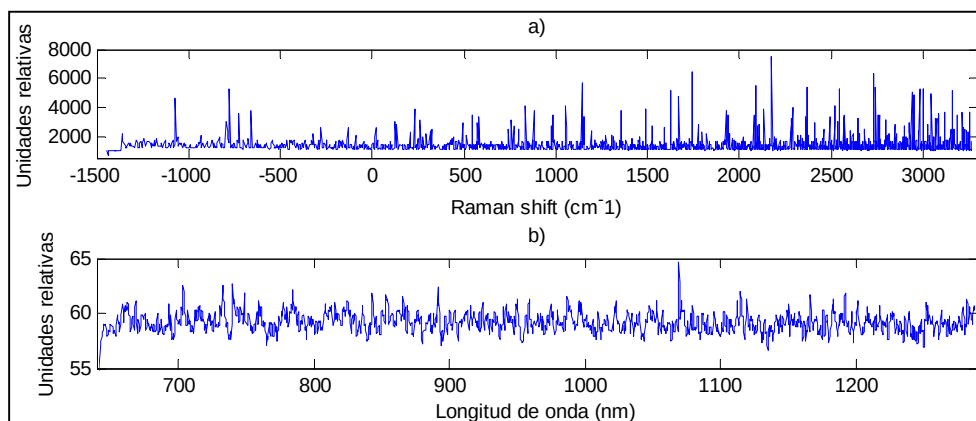


Figura 3-11: Espectros de oscuridad de cada instrumento; a) Raman b) NIR

3.7.2 Selección del rango espectral

Debido a la respuesta espectral de los detectores CCD, normalmente existe una atenuación en los extremos del rango espectral adquirido. Esta distorsión afecta los modelos de estimación ya que introduce variables sin relación con el fenómeno que se intenta predecir. Por otro lado, la información relevante para la predicción de las propiedades de interés no siempre está distribuida en todo el espectro sino que a veces se concentra en ciertos tramos, los cuales deben ser detectados para eliminar los sectores irrelevantes. Debido a lo anterior, se utilizan ventanas espectrales dentro del espectro completo, que buscan reducir los efectos negativos en la predicción seleccionando aquellos tramos del espectro que tengan la mayor relación con las variables de referencia, eliminando los tramos que no contienen información relevante y solamente introducen ruido en la predicción.

3.7.3 Filtro de Savitsky Golay

El filtrado a través del método de Savitsky-Golay es uno de los más utilizados en señales espectroscópicas para la eliminación de ruido aleatorio y suavización de la señal (Člupek et al., 2007). La gran ventaja de este filtro, en comparación con otros (FT, ventana móvil, LPF) es que tiende a conservar las características de la señal como máximos y mínimos relativos y anchos de los peaks. El método se basa en ajustar un polinomio de orden k en un pequeño intervalo de al menos $k+1$ puntos, con la posibilidad además de aplicarle derivadas de primer o mayor orden, lo cual en algunos casos es útil para eliminar ruidos de fondo constantes y para acentuar las diferencias entre los datos espectrales.

3.7.4 Media centrada

El centrado de medias se aplica generalmente cuando se hace un modelo de calibración multivariable. Se calcula un espectro promedio de todo el conjunto de espectros y éste se le resta a cada uno de los espectros individuales. Se hace lo mismo con los valores de referencia, se calcula el valor promedio y se resta este valor a cada una de las muestras. Mediante el centrado de medias se acentúan las diferencias entre los valores de referencia y la respuesta espectral de las muestras. Esto lleva a que se realice una calibración del modelo más precisa y con mejores predicciones. Este método se calcula como:

$$x_{i_{corr}} = x_i - \frac{\sum_i x_i}{n} \quad (3.1)$$

3.7.5 Eliminación de fluorescencia

El algoritmo utilizado para la eliminación de fluorescencia corresponde al del Círculo Rodante (*Rolling Circle*). Éste es un método basado en un algoritmo para la eliminación de la señal de fondo (*background*) de un espectro (Mikhailyuk y Razzhivin, 2003). En general los métodos de eliminación de *background*, como ajuste de curva, o derivaciones e integraciones sucesivas, requieren de un número considerable de parámetros y consideran muchas restricciones para el polinomio que será ajustado como línea de base. Este método está basado en las diferencias geométricas entre el *background* y la señal útil. El algoritmo es el siguiente, por cada punto (x_i, y_i) de la señal original, se construye un círculo de radio R con coordenada x_c del centro igual a x_i . La curva original y cada uno de los círculos, debe tener por lo menos un punto en común, mientras que el resto de los puntos debe tener coordenadas y menores que el respectivo punto en la curva original. Luego se debe calcular la diferencia entre las coordenadas y_i de la señal con respecto al hemisferio superior del círculo para cada x_i . La menor de estas diferencias es almacenada en un vector, el cual corresponderá al espectro corregido sin la señal de *background*. Este algoritmo puede ser interpretado intuitivamente como un círculo que se desplaza bajo la curva original y que va eliminando la región por la que pasa.

La característica principal de este método, es que introduce una mínima distorsión si el radio R es igual al mínimo radio de curvatura de la señal de *background*, o menor a ésta, siempre y cuando sea mayor que el ancho de cualquiera de los peaks de la señal útil. Además, tiene la ventaja de que utiliza un solo parámetro R , el cual puede ser elegido fácilmente analizando la condición de distorsión mínima dada por Mikhailyuk, en una curva del área bajo la señal corregida vs el radio R del círculo.

El algoritmo del *Rolling Circle* fue programado de forma independiente en Matlab, ya que no está incluido en el conjunto de herramientas utilizado.

3.8 Calibración y validación de los modelos de regresión

La elección de los modelos óptimos para cada caso se realizó en función de los índices de correlación R^2 (ecuación 2.16) tanto en calibración como en validación, y del error cuadrático medio de predicción RMSEP (ecuación 2.15). Los índices de correlación son variables adimensionales, mientras que el RMSEP se mide en las mismas unidades que el valor de referencia, en este caso [Kg/m³].

Como se mencionó anteriormente, el método de regresión por mínimo cuadrados parciales (PLS) es el método de calibración más adecuado cuando se quiere relacionar una señal multivariable (espectro) con su respectivo valor de referencia. Este algoritmo se aplicó a los datos utilizando el software Matlab a través del *toolbox* PLS Toolbox 5.2.2 de Eigenvector Research IncTM. Se utilizó principalmente el GUI (*Graphic User Interface*) de esta aplicación, el cual permite realizar de manera intuitiva una calibración completa de los modelos de regresión, incluyendo la separación de los conjuntos de calibración y validación, el filtrado de los datos de entrada, los métodos de validación cruzada, la detección de *outliers* y selección del número óptimo de variables latentes para cada modelo.

En la calibración de los modelos se trabajó con los espectros Raman y NIR por separado, desarrollando un modelo de predicción para cada caso. Los datos de entrada para el modelo son los espectros procesados junto con los correspondientes valores de referencia, mientras que de salida se obtiene el vector de regresión **b**, además de algunos valores estadísticos para la evaluación del modelo.

Los datos de entrada, tanto espectros como referencias deben ser separados en dos grupos para obtener un modelo válido. Uno corresponde a los datos de calibración, mientras que el otro es el de validación. El primero contiene alrededor de tres cuartos de las muestras y el otro el resto. Las muestras de cada uno de los grupos fueron seleccionados para abarcar el rango completo de valores en ambos conjuntos.

Los datos de calibración son los que se usan para construir el modelo en sí, mientras que los de validación sirven para comprobar la efectividad del modelo sobre muestras

que no están en el conjunto de calibración. Luego de ingresar los datos al programa, se ajustan las configuraciones deseadas, se ejecuta el cálculo del modelo y se debe elegir el número de Variables Latentes (VL) que éste usará. Para lo cual se realiza una validación cruzada, en donde aleatoriamente se saca un subconjunto del total de muestras y se utiliza como conjunto de validación, mientras que el resto queda para la calibración. Esto se repite un número predeterminado de veces y se calcula el RMSECV para cada una de ellas. Con ayuda de estos valores a través de un gráfico RMSECV vs VL, como el de la figura 3-12, se elige el número de variables latentes óptimo para el modelo, el cual corresponde al punto en que se quiebra la curva luego de la primera bajada.

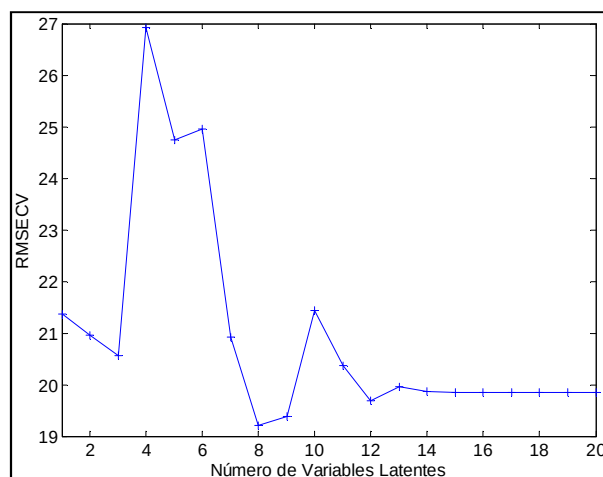


Figura 3-12: Gráfico para la selección del número de variables latentes óptimo del modelo PLS

Con el número de variables latentes seleccionado, se pueden analizar los gráficos estadísticos, en los que se ve el desempeño del modelo ajustado. Estos gráficos corresponden a la revisión de los *scores* y *loadings* del modelo. Dentro de los *scores*, se tiene el gráfico de dispersión en donde cada muestra aparece con su valor de referencia predicho vs el valor medido (ver figura 3-13). Las muestras de calibración se muestran como círculos, mientras que las de validación como triángulos.

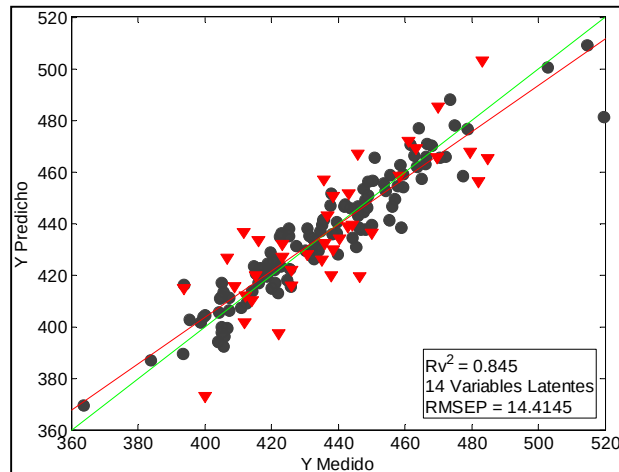


Figura 3-13: Gráfico de dispersión valores medidos vs predichos

Además se muestra la recta de tendencia de las muestras junto a la recta de pendiente 1 que representa la correlación perfecta. Junto al gráfico se observan datos estadísticos como R^2 , RMSEC y RMSEV. La importancia de este gráfico radica en que se puede ver que tan bien ajustados están los datos en el modelo, tanto los de calibración como los de validación.

Otro gráfico importante de analizar dentro de los *scores* es el gráfico de residuos de predicción (*Y residuals*) para cada muestra, como el que se muestra en figura 3-14. En este gráfico se muestra básicamente que tan alejada está cada muestra de la recta de tendencia del modelo PLS, por lo que es fácil identificar alguna muestra que no esté representada por éste. Estas muestras son llamadas *outliers* y su exclusión puede mejorar considerablemente los resultados de predicción del modelo.

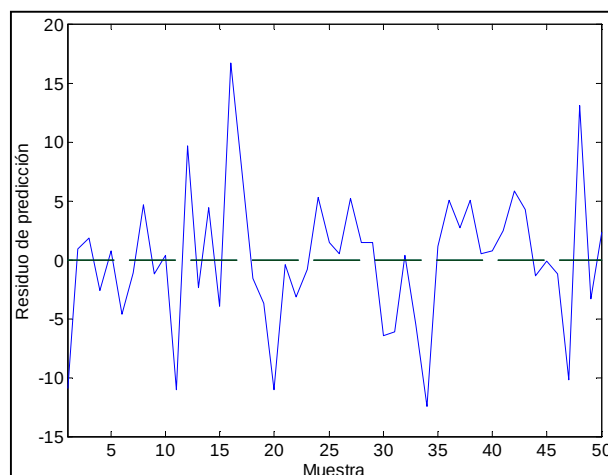


Figura 3-14: Gráfico de Muestras vs Residuos en la predicción para la detección de outliers

Para la detección de *outliers* dentro del todo el conjunto de muestras, se realiza una calibración del modelo PLS con todos los datos y utilizando la técnica de validación cruzada se determina que muestras escapan de la tendencia general, observando: los gráficos de Muestras vs Residuos en predicción (ver figura 3-14), la diferencia entre los valores de referencia predichos y los medidos en un gráfico de dispersión, los valores de T^2 de cada muestra, que representan variaciones inusuales dentro del modelo dado un determinado rango de confianza y los residuos Q que se relacionan con variaciones ajenas al modelo establecido por el número de variables latentes seleccionado.

En la figura 3-15 se observa un ejemplo donde se ha desarrollado un modelo a partir de datos que consideran la medición de tres variables independientes. El modelo obtenido cuenta con dos variables latentes las cuales forman un plano que contiene a la mayoría de las muestras utilizadas, excepto dos. Una de ellas presenta un alto valor T^2 , ya que está contenida en el plano definido por las VL pero lejos de ambas, es decir presenta variaciones inusuales contenidas en el modelo. La otra muestra presenta un alto valor Q ya que no está contenida en el plano definido por las VL, es decir presenta variaciones inusuales fuera del modelo.

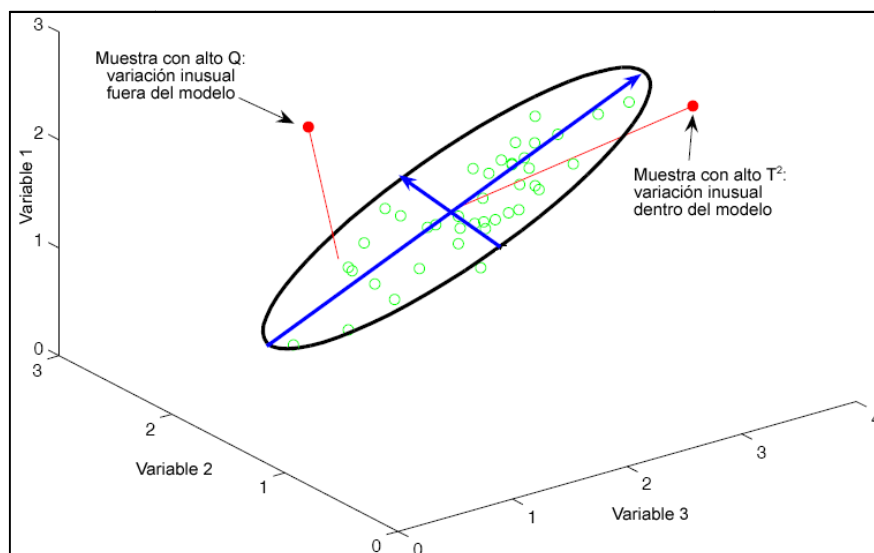


Figura 3-15: Ejemplo de muestras con alto residuo Q y alto T^2

En el caso del análisis de los *loadings*, se tienen dos gráficos importantes para la interpretación del modelo. El primero es el que muestra el vector de regresión $\mathbf{b}$, vector que se multiplica por cada espectro para obtener la respectiva predicción. Este gráfico se puede ver en la figura 3-16.

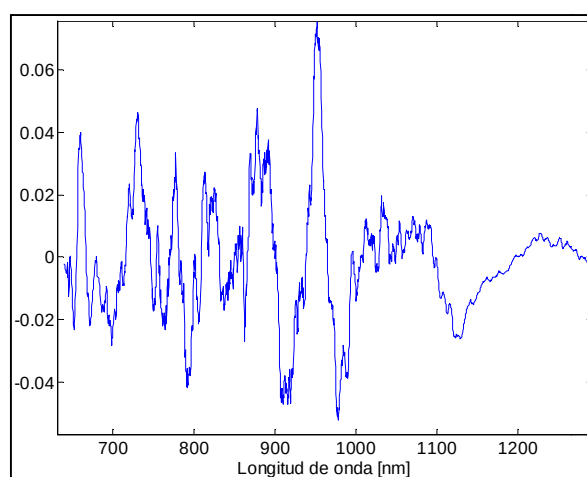


Figura 3-16: Gráfico del vector de regresión $\mathbf{b}$ para un modelo NIR

El segundo (figura 3-17), es el que muestra las VIP (*Variable Importance in Projection*). Este vector entrega una estimación de la importancia relativa de cada variable independiente, en cada una de las variables latentes utilizadas en el modelo PLS. Estos coeficientes entonces representan la importancia de cada variable x en el ajuste de los *scores* de las matrices X e Y , según el algoritmo del PLS (Chong y Jun, 2005).

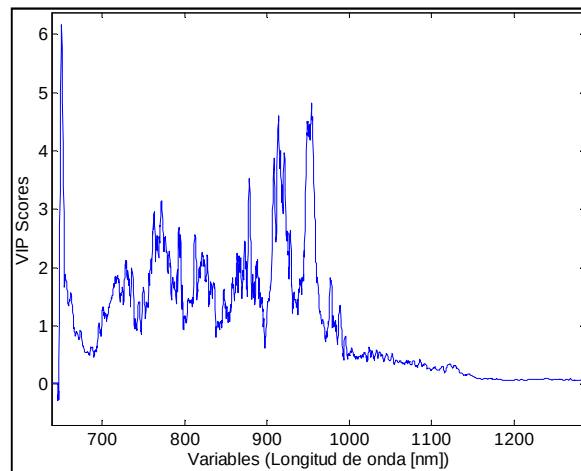


Figura 3-17: Gráfico de los VIP Scores

Las muestras *outliers* normalmente representan errores en el proceso de medición de la densidad básica o algún problema en la adquisición de su espectro y suelen ser eliminados del conjunto total de muestras, previo a la calibración de los modelos finales de predicción. Luego de esta eliminación, se separan los datos originales en dos conjuntos, uno de calibración y otro de validación con un 75% y 25% de las muestras respectivamente. Estos conjuntos se generan aleatoriamente a partir del universo de datos pero debe tenerse cuidado en incluir los valores extremos de las referencias en el conjunto de calibración. Lo anterior se debe a que las estimaciones mejoran

notablemente cuando el modelo de predicción no es sesgado y ha sido calibrado utilizando todo el rango de valores de referencias disponibles. Otro punto importante en la creación de los conjuntos de calibración y validación es mantener la distribución de los datos originales, cuidando de que los grupos sean representativos de la distribución original de los datos de referencia.

4 RESULTADOS Y DISCUSIÓN

Se desarrollaron modelos PLS para ambas técnicas espectroscópicas utilizando diversas combinaciones en los conjuntos de calibración y validación. Se probaron diversas técnicas de procesamiento de los datos de entrada (suavizado Savitsky Golay, centrado de medias, eliminación de fluorescencia y ventanas espectrales) con el fin de mejorar los resultados obtenidos.

Como se demuestra a continuación para el caso Raman, los resultados no fueron satisfactorios. La capacidad de predicción de los modelos desarrollados fue insuficiente, obteniendo bajos índices de correlación tanto en validación como en calibración a pesar de los esfuerzos realizados que serán detallados más adelante.

Por otro lado, en el caso NIR, los resultados obtenidos fueron promisorios y se comprobó la factibilidad de la técnica para predecir la densidad básica de la madera de *Eucalyptus nitens* a partir de muestras húmedas.

El punto de partida del análisis estadístico realizado es con los espectros y valores de referencia de las muestras. En este estudio se contó un total de 157 muestras. Los valores de densidad básica obtenidos con el método tradicional explicado anteriormente se detallan en la siguiente tabla:

Tabla 4-1: Densidad básica de las muestras [Kg m⁻³]

Nº de muestras	Mínimo	Máximo	Promedio	Desviación Estándar
157	363,68	519,50	435,85	25,39

El rango normal de densidades en *Eucalyptus nitens* adultos (10 años) varía entre los 490 y los 520 [Kg m⁻³] (INFOR, 2008). A partir de la tabla 4-1 y la figura 4-1, se observa que en este caso el rango si bien es amplio y bien distribuido, está desplazado

hacia valores inferiores. Esto se explica por la juventud de los ejemplares utilizados (8 años).

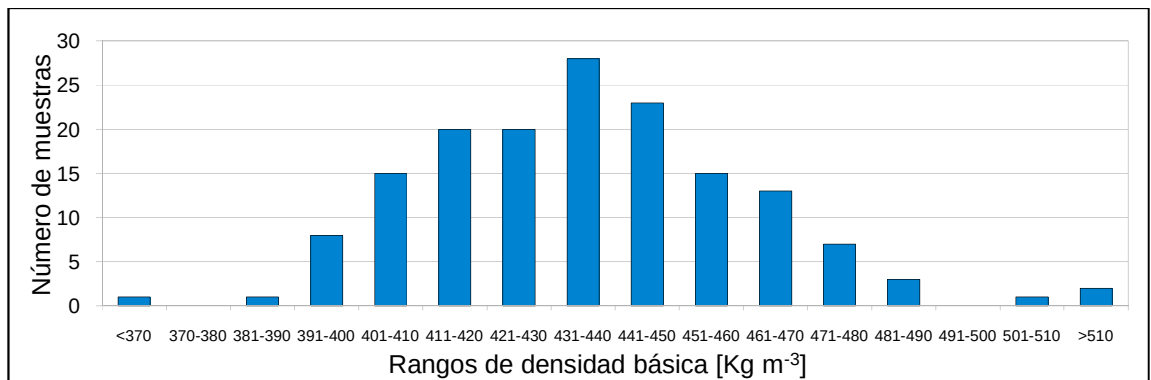


Figura 4-1: Histograma de la densidad básica de las muestras

Se eliminaron las muestras *outliers* y se generaron aleatoriamente 10 pares calibración-validación, tomando en cuenta las precauciones mencionadas en el capítulo anterior. Se calibró un modelo PLS cada conjunto con los datos crudos, sin aplicar procesamientos, obteniendo los índices de correlación en calibración y validación; y de error cuadrático medio de predicción. El número óptimo de VL fue determinado aplicando la técnica de validación cruzada en cada caso tomando en cuenta el gráfico de RMSECV vs VL al realizar una calibración con el total de los datos. La tabla 4-2 resume los datos de los modelos obtenidos para cada conjunto.

Tabla 4-2: Modelos previos Raman y NIR con los datos crudos

N° Set	Raman				NIR			
	Rc ²	Rv ²	RMSEP	VL	Rc ²	Rv ²	RMSEP	VL
1	0,357	0,130	28,6	8	0,487	0,223	33,6	13
2	0,395	0,194	30,5	8	0,461	0,309	28,2	13
3	0,363	0,135	27,4	8	0,467	0,268	45,8	14
4	0,420	0,095	30,1	8	0,500	0,248	53,6	13
5	0,440	0,120	32,3	8	0,262	0,145	46,9	11
6	0,366	0,089	28,5	8	0,534	0,300	27,5	14
7	0,121	0,253	23,6	7	0,650	0,227	31,7	14
8	0,352	0,117	28,5	8	0,660	0,406	27,1	15
9	0,402	0,142	30,7	8	0,698	0,239	27,0	15
10	0,083	0,183	39,1	5	0,569	0,556	23,7	14
Media	0,3299	0,14582	29,93	7,6	0,5288	0,2921	34,51	13,6

Los modelos finales para la predicción de densidad básica a partir de los espectros Raman y NIR mejoraron notablemente gracias a los procesamientos de datos aplicados a los espectros, los cuales serán detallados a continuación.

4.1 Modelos Raman

La sustracción del espectro de oscuridad del instrumento Raman, el suavizado de espectros con un filtro de Savitsky-Golay de ancho 15 puntos y orden 2 y el centrado de medias (CM) de los datos de entrada, fueron procesamientos que mejoraron sistemáticamente los modelos de predicción obtenidos. Estos procesamientos ya han sido aplicados a los espectros crudos para todos los casos que siguen a continuación. El efecto de la sustracción del espectro de oscuridad y el suavizado Savitsky-Golay (SG) se observan en la figura 4-2.

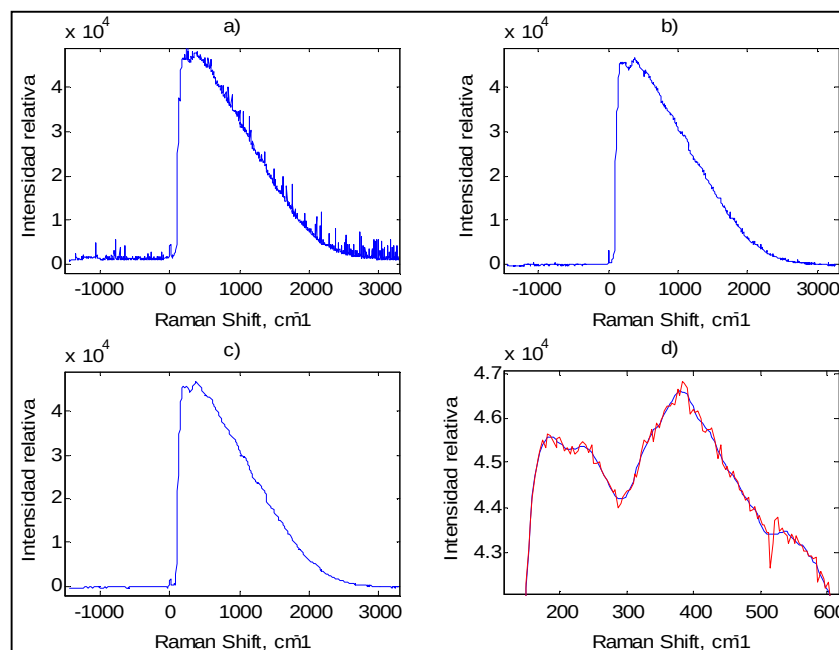


Figura 4-2: Espectro Raman a) datos crudos b) sin dark c) suavizado SG d) detalle del suavizado

Los resultados preliminares al utilizar los datos de entrada sin ningún tipo de procesamiento, entregaban en el mejor de los casos, R_v^2 de 0,25 (ver tabla 4-2). El primer paso para aumentar la capacidad de predicción de los modelos corresponde a la eliminación de las muestras *outliers* reflejadas en los datos estadísticos del modelo al realizar una calibración con el total de los datos originales. Para ello se observó el gráfico de residuos de predicción para cada muestra y se consideró como *outlier* aquella que mostrara los mayores valores (figura 4-3). En segundo lugar se consideraron las muestras que presentaran altos valores tanto en su residuo Q como en su valor T^2 , superiores al límite de confianza del 95% en ambos casos (figura 4-4). De esta forma se encontraron 7 muestras *outliers* y su eliminación significó un leve aumento en el índice de correlación de calibración del modelo preliminar y una disminución del RMSEC (figura 4-5).

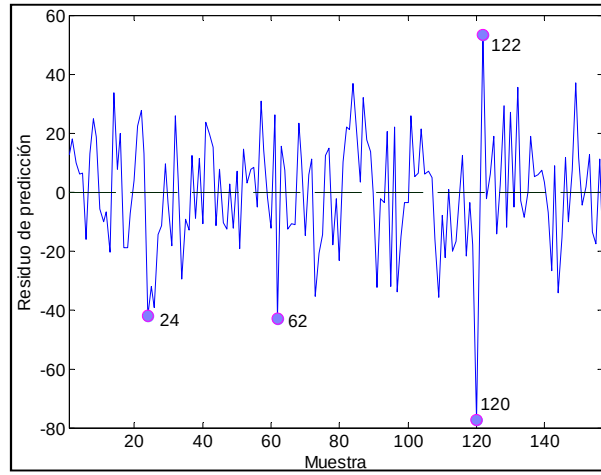


Figura 4-3: Muestras *outliers* Raman con alto residuo de predicción

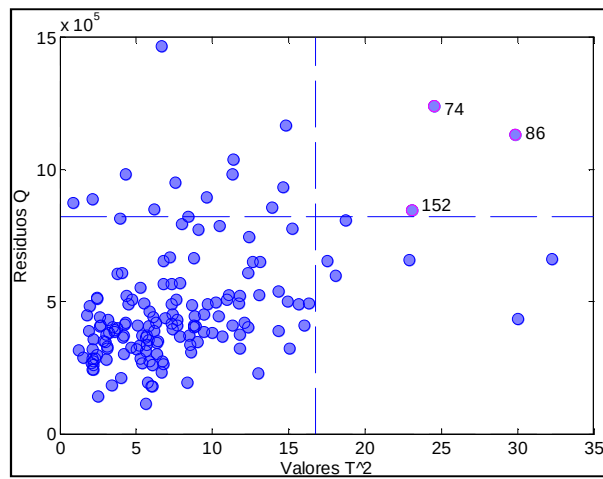


Figura 4-4: Muestras *outliers* Raman con valores Q y T^2 fuera de los rangos de confianza

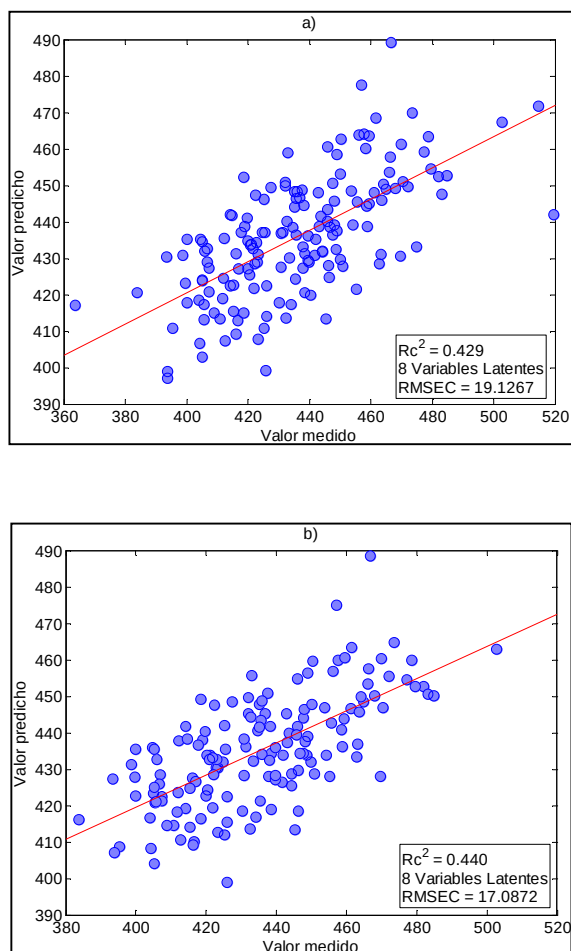


Figura 4-5: Gráficos de dispersión del modelo preliminar; a) Con todas las muestras b) Eliminando outliers

El siguiente paso consistió en la selección del rango espectral que optimizara la capacidad de predicción de los modelos PLS. Como punto de partida se utilizó el gráfico de las *VIP Scores* presentado en la figura 4-6 donde se puede observar claramente un protagonismo del rango comprendido entre los 0-1200 cm^{-1} . Al utilizar solamente este rango en la calibración del modelo preliminar, se logra una mejora en el índice de correlación de calibración, esto se debe a que se eliminan sectores del espectro que únicamente aportan ruido en la calibración. Se realizaron pruebas para obtener un rango

más preciso que optimizara el índice de correlación de calibración, pero el rango inicial comprendido entre los 0-1200 cm^{-1} resultó ser el más adecuado.

La correlación del modelo de calibración mejoraba notoriamente al utilizar este rango espectral (ver figura 4-7), sin embargo, la capacidad de predicción sobre las muestras del conjunto de validación (aquellas que no se utilizaron en el ajuste del modelo) se vio afectada negativamente. Esto se comprobó en varios de los conjuntos iniciales y se optó por continuar utilizando el espectro Raman completo.

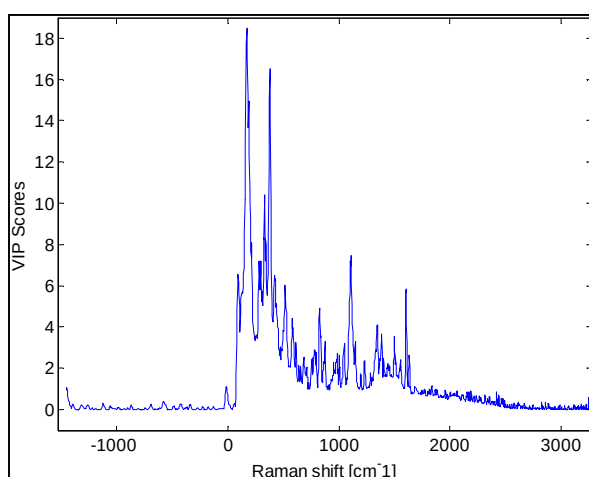


Figura 4-6: *VIP Scores* para el espectro Raman completo

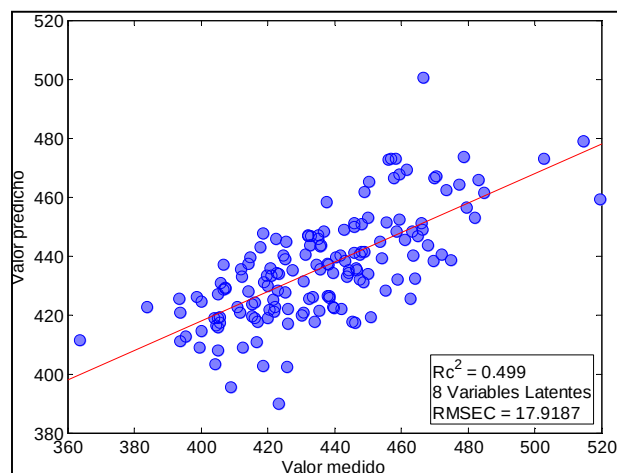


Figura 4-7: Gráfico de dispersión modelo preliminar Raman, utilizando el rango 0-1200 cm^{-1}

A pesar de los esfuerzos hasta este punto por calibrar un modelo de predicción suficiente, el mejor índice de correlación en validación de los modelos obtenidos no superaba el 0,4.

El siguiente paso en la búsqueda de una solución al problema planteado consistió en la eliminación de la fluorescencia de los espectros Raman adquiridos, identificada gracias a la suave curvatura de fondo observada. Para ello se aplicó el método del Circulo Rodante, cuya ventaja radica en la definición radio del círculo como único parámetro. Como regla general para la elección del radio, se debe conocer a priori el ancho máximo de los peaks Raman que contienen información relevante en el espectro y se debe elegir un círculo de diámetro mínimo igual a este ancho y radio máximo igual al radio de curvatura de la señal de fluorescencia. En este caso, estos valores no se conocen y son diferentes en cada espectro, por lo tanto se utilizó el criterio descrito por Milkhailyuk y Razzhivin (2003) para encontrar el radio óptimo del círculo que provoca una mínima distorsión en la señal y máxima eliminación de la señal de fluorescencia. Para esto, se grafica el área bajo la curva de la señal sin fluorescencia en función del radio del círculo, y luego se elige el punto más a la derecha (radio mayor) de la meseta observada. En la figura 4-8 se muestran los espectros originales y el resultado de esta eliminación.

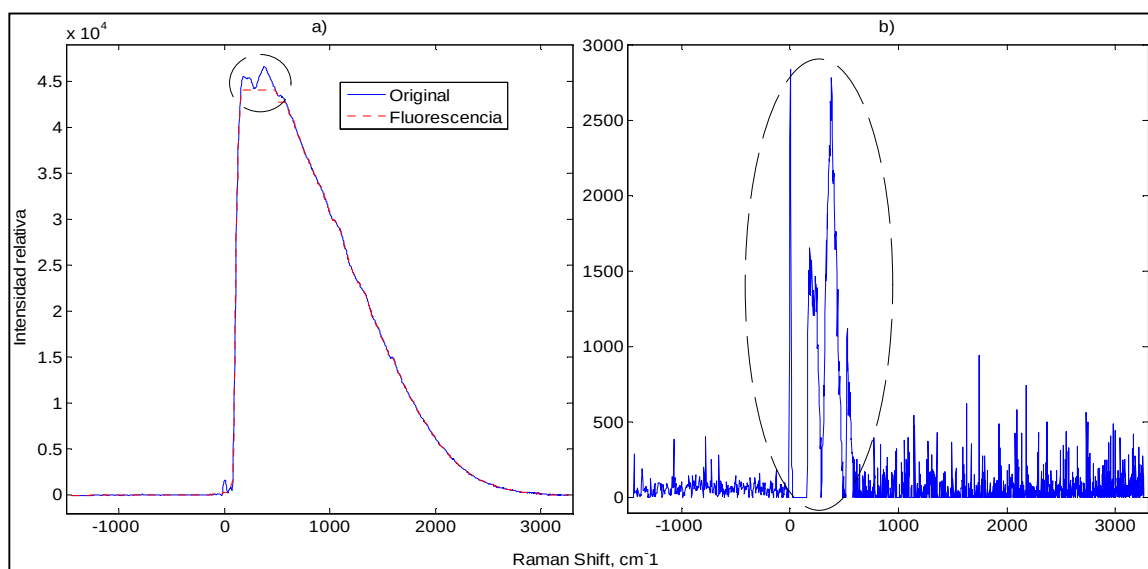


Figura 4-8: Eliminación de fluorescencia: a) Espectro original y fluorescencia b) Espectro sin fluorescencia

Posteriormente se ajustaron modelos utilizando los espectros sin fluorescencia. Se observaron leves mejoras en la capacidad de predicción sobre nuevas muestras para algunos modelos, sin embargo los resultados en general seguían siendo poco satisfactorios, manteniendo índices de correlación bajos y una capacidad de predicción insuficiente. En la siguiente tabla se muestran los principales parámetros de los modelos obtenidos.

Tabla 4-3: Modelos Raman sin fluorescencia

Nº Set	Rc²	Rv²	RMSEP	VL
1	0,44	0,33	22,4	5
2	0,37	0,04	24,3	4
3	0,64	0,10	23,7	6
4	0,17	0,06	21,0	2
5	0,26	0,16	20,5	3
6	0,40	0,13	23,0	4
7	0,27	0,43	17,4	4
8	0,31	0,21	19,8	4
9	0,50	0,15	25,4	5
10	0,30	0,28	23,4	4
Media	0,37	0,19	22,1	4,1

El siguiente paso consistió nuevamente en la búsqueda de un rango espectral que optimizara los resultados, pero esta vez utilizando solamente los espectros sin fluorescencia. Los mejores resultados se obtuvieron para una ventana entre los 130 y 580 cm^{-1} , correspondientes al sector encerrado en el círculo de la figura 4-8. Los resultados se presentan en la tabla 4-4.

Tabla 4-4: Modelos sin fluorescencia para el rango espectral entre 130-580 cm⁻¹

N°Set	Rc²	Rv²	RMSEP	VL
1	0,32	0,37	21,4	6
2	0,44	0,14	22,0	7
3	0,39	0,10	23,3	5
4	0,13	0,06	19,9	2
5	0,38	0,17	20,4	6
6	0,28	0,22	21,5	4
7	0,33	0,38	18,1	6
8	0,40	0,26	20,1	7
9	0,33	0,27	23,2	5
10	0,21	0,26	23,5	3
Media	0,32	0,22	21,3	5,1

Utilizando la ventana espectral mencionada anteriormente, se notó una leve mejora de los modelos sin fluorescencia. Sin embargo las mejoras no son significativas al compararlas con los modelos obtenidos con el rango completo, detallado en la tabla 4-3.

En última instancia se intentó calibrar modelos solamente con la fluorescencia de los espectros Raman, es decir, con el residuo que se obtenía después de aplicar el método de eliminación del Círculo Rodante. Se probó también con la fluorescencia dentro del rango 130-580 cm⁻¹ determinada en el punto anterior, pero los resultados eran equivalentes al considerar el rango completo. Los resultados nuevamente fueron insatisfactorios y se presentan en la tabla 4-5.

Tabla 4-5: Modelos a partir de la fluorescencia del espectro Raman, rango completo

N° Set	Rc²	Rv²	RMSEP	VL
1	0,28	0,14	25,2	5
2	0,33	0,05	23,7	5
3	0,28	0,10	23,4	5
4	0,29	0,08	20,7	5
5	0,28	0,20	19,6	5
6	0,37	0,03	25,2	6
7	0,45	0,33	19,0	8
8	0,29	0,14	21,4	5
9	0,41	0,20	23,9	6
10	0,27	0,23	24,0	5
Media	0,32	0,15	22,6	5,5

4.2 Modelos NIR

Al contrario del caso Raman, los resultados obtenidos para la predicción de la densidad básica a partir de los espectros NIR fueron promisorios. La aplicación de procesamientos a los espectros crudos produjo importantes mejoras en los índices de correlación en calibración y validación disminuyendo considerablemente el *RMSEP*. A continuación se describen aquellos que produjeron los mejores resultados en los modelos finales, aumentando considerablemente los índices de correlación de validación y disminuyendo el error de predicción en los modelos finales.

En la búsqueda de *outliers* se encontraron dos muestras con características extrañas, que presentaban los mayores residuos en la predicción (ver figura 4-9) al realizar una calibración del modelo PLS con el total de los datos utilizando únicamente el centrado de medias como procesamiento previo. Estas muestras fueron eliminadas del conjunto original y se mejoraron notablemente los índices de correlación de calibración y

validación junto con una leve disminución en los errores de predicción RMSEP de los modelos finales (ver figura 4-10).

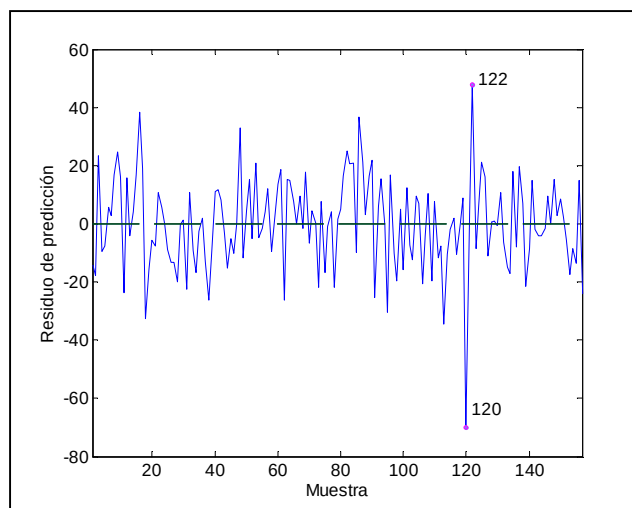


Figura 4-9: Valores residuales de predicción de las muestras NIR

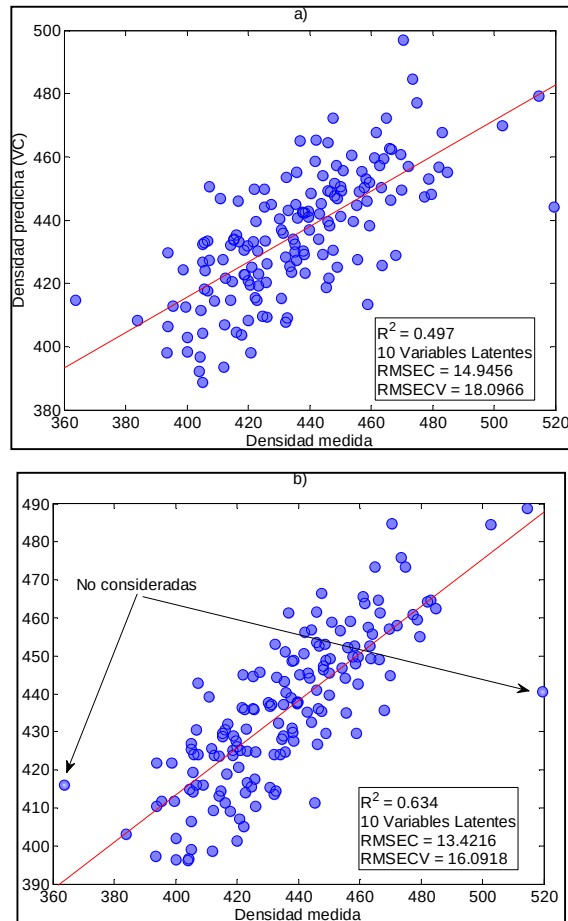


Figura 4-10: Modelo de calibración utilizando todos los datos; a) Con outliers b) Sin outliers

Al analizar los espectros de las muestras *outliers* se observa que su forma es muy distinta a la del promedio de los espectros NIR como muestra la figura 4-11. Esto se debe probablemente a algún problema en la medición y valida la detección de muestras extrañas mediante la simple observación del gráfico de residuos de predicción.

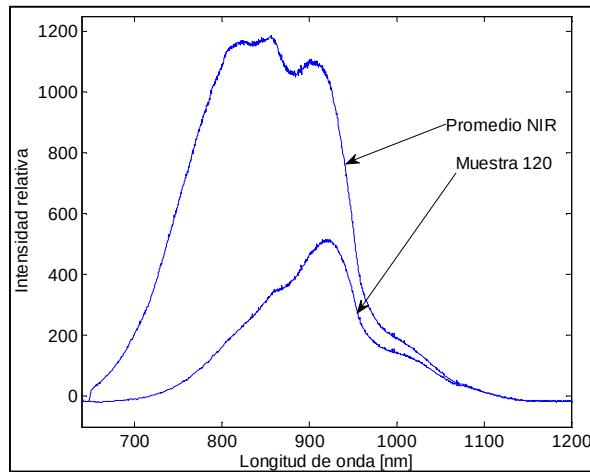


Figura 4-11: Espectro errático de la muestra 120 vs el promedio NIR

Posteriormente se utilizó un filtro de suavizado Savitsky-Golay de ancho 15 y orden 2 el cual mejoraba considerablemente los modelos al disminuir el ruido de los espectros originales, como muestra la figura 4-12. Aumentar el ancho y el orden del filtro afectaba los modelos, pero de manera marginal.

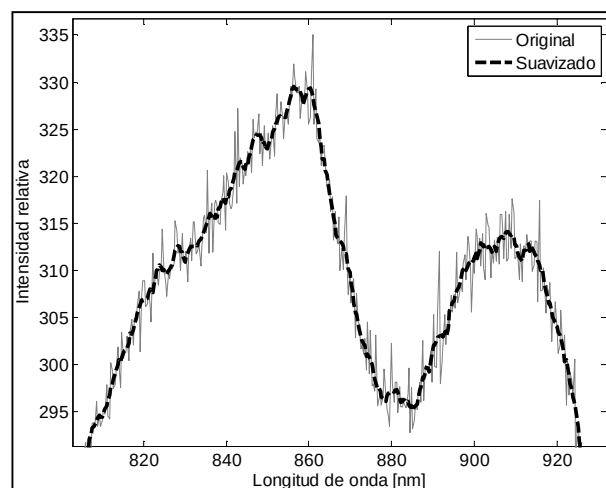


Figura 4-12: Efecto del suavizado SG de ancho 15 y orden 2 en un espectro NIR

La selección de un rango espectral específico dentro del espectro original mejoró de manera significativa la capacidad de predicción de los modelos PLS. Observando el gráfico de *VIP Scores* de todas las variables de entrada (espectro completo) se eligió un rango espectral inicial ubicado entre los 600-1050 nm (ver figura 4-13). Realizando varios ensayos de prueba y error sobre distintos conjuntos de calibración y validación, se logró ajustar este rango inicial optimizando sistemáticamente los resultados en los modelos calibrados. Esta ventana espectral se ubicó entre los 750-1000 nm y representa el sector delimitado en la figura 4-14.

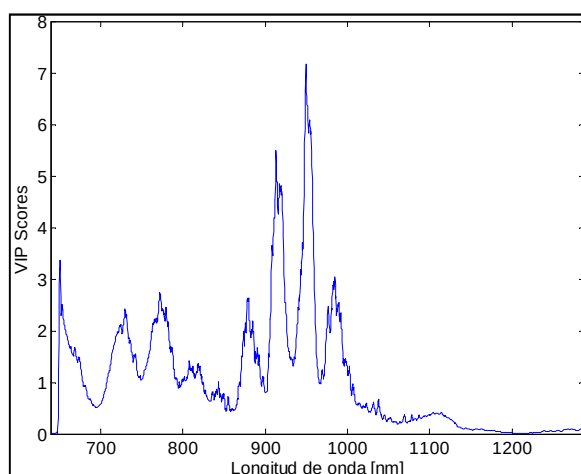


Figura 4-13: VIP Scores del espectro NIR completo

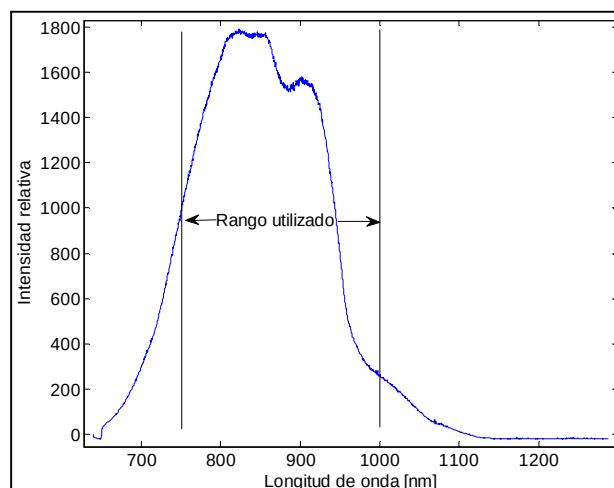


Figura 4-14: Rango espectral final utilizado en la calibración NIR. 750-1000 nm

Gracias a los procesamientos mencionados anteriormente se logró elaborar un modelo de predicción satisfactorio para la densidad básica del *Eucalyptus nitens* a partir del espectro NIR. En la tabla 4-6 se resumen los principales parámetros de este modelo. El gráfico de dispersión para las muestras del modelo se observa en la figura 4-15.

Tabla 4-6: Parámetros del modelo de calibración PLS para los espectros NIR

VL	R_c^2	R_v^2	RMSEC	RMSEP
12	0,82	0,87	10,18 Kg/m ³	8,61 Kg/m ³

Al analizar los *VIP scores* del modelo final NIR (figura 4-16) se observan las longitudes de onda que tienen una mayor importancia en la predicción de la densidad básica. Comparándolo con el de la figura 4-13 se observa que no existen diferencias tan grandes entre los VIPs de las variables consideradas en esta ventana espectral.

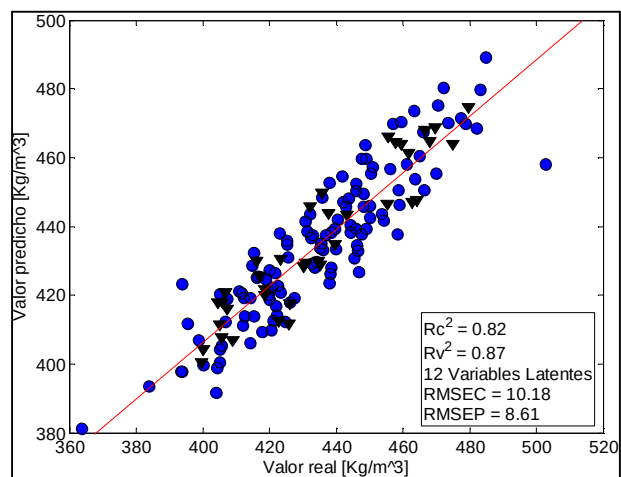


Figura 4-15: Gráfico de dispersión para densidad básica a partir de espectros NIR

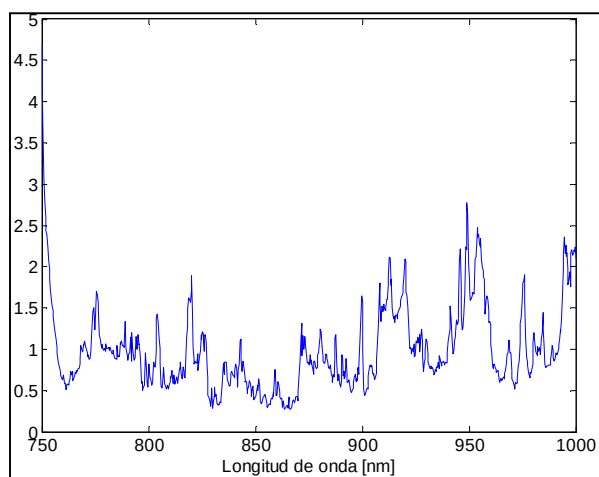


Figura 4-16: *VIP* scores para el modelo final NIR

5 CONCLUSIONES

Los resultados obtenidos en la estimación de densidad básica de *Eucalyptus nitens* mediante espectroscopía NIR, confirman esta técnica como un potencial reemplazante de los métodos tradicionales. Más aún, permiten imaginarla como el núcleo de trabajo de instrumentos portátiles que puedan ser llevados directamente a terreno, simplificando el análisis de laboratorio.

Los modelos NIR descritos en el capítulo anterior tienen índices de correlación y valores de error cuadrático medio que permiten compararlos con otros trabajos previos (Schimleck et al, 1999) basados en técnicas de laboratorio con un complejo y lento procesamiento de la muestra previo a la toma del espectro. Los modelos obtenidos a partir de espectros NIR de tarugos húmedos, con R_v^2 de 0,87 y R_c^2 de 0,82, permiten concluir que este método podría ser utilizado para la estimación de la densidad básica del árbol en pie mediante un sistema de medición en terreno. Esto abre la posibilidad de utilizar un espectrómetro NIR portátil en la industria forestal, aprovechando las facilidades de la técnica, se extiende la posibilidad de ampliar los modelos y utilizarlos para el análisis de otras propiedades de interés.

En cuanto a la técnica Raman, no fue posible desarrollar modelos que permitieran predecir confiablemente la densidad básica de la madera. Los pobres resultados obtenidos mediante la aplicación de esta técnica se explican en base a las características particulares observadas en la adquisición del espectro y a la naturaleza de este fenómeno inelástico. El problema aparente es que se está intentando obtener un espectro representativo de la muestra a partir de un volumen de medición muy pequeño que no refleja necesariamente la composición global del tarugo ya que este es sumamente heterogéneo en su estructura física. De esta manera el espectro final adquirido no guarda relación con las propiedades globales de la muestra, como es la densidad básica. También es probable que la dispersión Raman, debido a la alta sensibilidad y complejidad óptica de los elementos involucrados en la detección de este fenómeno, no sea la base más adecuada en la búsqueda de un método simple, robusto y confiable de

adquisición de espectros sobre muestras sin procesamiento. Por otro lado, la utilización de espectrómetros Raman más complejos y el procesamiento de las muestras previo a la adquisición de espectros (secado, molienda u otros) con el fin de evitar los problemas detectados, se aleja del objetivo final de esta investigación que corresponde a la determinación de una técnica simple que pueda ser utilizada en terreno, rápida y eficazmente.

6 TRABAJO FUTURO

Existe una serie de actividades que podrían ser desarrolladas para extender el trabajo realizado en esta investigación, algunas de ellas se listan a continuación.

- a) Comprobar si la misma técnica NIR permite la predicción de otras variables de interés en la madera tales como el contenido de humedad, el rendimiento pulpable, al ángulo microfibrilar y el contenido de lignina, entre otras.
- b) Verificar si es posible ajustar varios modelos que funcionen en paralelo a partir de una sola adquisición del espectro, lo cual daría pie a un instrumento que midiera un perfil de características de la madera a partir de una sola medición del espectro.
- c) Verificar si se obtienen resultados similares utilizando la técnica NIR por reflectancia. La importancia de esto radica en que la construcción de una sonda basada en reflectancia debería ser más simple que una basada en transmitancia, ya que tanto el emisor como el receptor estarían del mismo lado de la muestra.
- d) Construir un prototipo del instrumento con el modelo PLS embebido, capaz de predecir en terreno los valores de densidad básica utilizando el espectrómetro NIR portable.
- e) Aplicar las mismas técnicas sobre otras especies de interés para la industria forestal, tales como el Pino Radiata y el Pino Oregón, comprobando la capacidad de predicción de los modelos.

- f) Utilizar otro tipo de sonda Raman para el estudio, capaz de variar la potencia del láser sobre la muestra y que permita abarcar una zona mayor de muestreo, de manera de reducir el efecto de la variabilidad espacial inherente de la madera.
- g) Construir una sonda que se introduzca directamente en el árbol en pie y permita la toma de espectros NIR, sin necesidad de la extracción de tarugos. Esta sonda debe utilizar la técnica de adquisición por transmitancia difusa y realizar el menor daño posible en el ejemplar, como se muestra en la figura 6-1. Como una alternativa se presenta un diseño preliminar que permitiría la adquisición de espectros NIR en terreno. El diseño cuenta con fibras ópticas, reflectores y un colimador que cumplen la función de dirigir la luz desde la fuente de excitación, a través de la muestra y de regreso al espectrómetro portátil (figura 6-2).

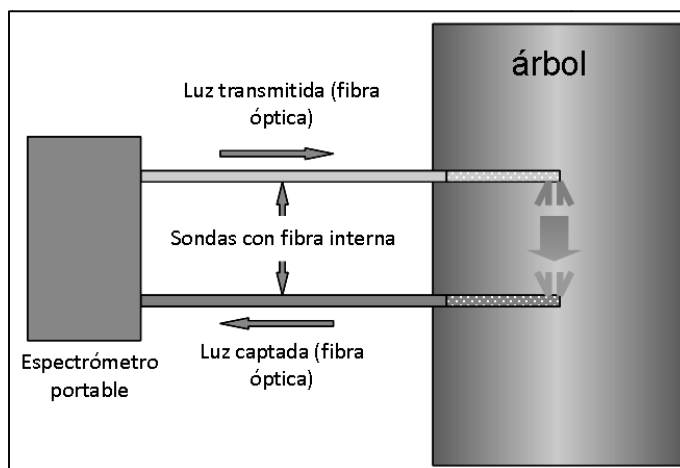


Figura 6-1: Adquisición del espectro NIR mediante transmitancia difusa en el árbol en pie

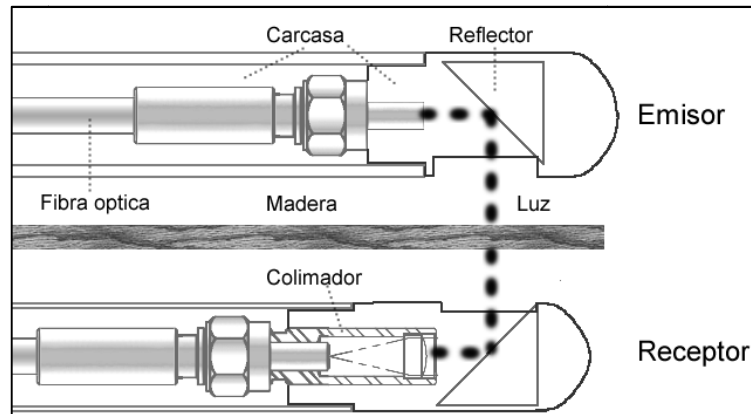


Figura 6-2: Diseño preliminar de la sonda NIR

NOMENCLATURA

R_v^2	Índice de correlación del set de calibración
R_c^2	Índice de correlación del set de validación
RMSEC	Raíz del error medio cuadrático del set de calibración
RMSEP	Raíz del error medio cuadrático del set de predicción
$\sigma_{\hat{y}y}$	Covarianza entre $\hat{y}$ y y
σ_y	Desviación estándar de y
PLS	Mínimos cuadrados parciales (<i>Partial Least Squares</i>)
NIR	Infrarrojo cercano (<i>Near Infra Red</i>)
CCD	Dispositivo de cargas eléctricas interconectadas (<i>Charge Coupled Device</i>)
ν_0	Frecuencia del láser de excitación
VL	Variable Latente (Factores del modelo PLS)
VIP	Variable Importance in Projection
DB	Densidad Básica
MS	Masa Seca
VV	Volumen Verde

BIBLIOGRAFÍA

Acuna M., y Murphy G. (2007). Uso de espectroscopía infrarroja y análisis multivariado para predecir la densidad de la madera de pino oregón. *Bosque*. 28, p. 187-197.

B&W TEC Inc. (2007). User Manual, BTR111 MiniRam miniature NIR Raman spectrometer. *Manual de referencia*, Newark, DE.

Carter P., Briggs D., Ross RJ. and Wang X. (2005). Acoustic testing to enhance western forest values and meet customer Wood quality needs. *In Harrington CA, SH Schoenholtz eds. Productivity of Western forests: a forest products focus. Gen. Tech. Rep. PNW-GTR-642*. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station.

Chong, I., Jun, C. (2005). Performance of some variable selection methods when multicollinearity is present. *Chemometrics and Intelligent Laboratory Systems*, 78, p. 103-112.

CIF (2008). Boletín estadístico 121, Anuario Forestal 2008 (cap III). *Centro de Información Forestal, sede metropolitana*. Recuperado el 11 de diciembre de 2009, de http://www.gaf.cl/descargas/index.php?mekat=General&user_sort_top=&down_id=18

Člupek, Matějka and Volka. (2007). Noise reduction in Raman spectra: Finite impulse response filtration versus Savitzky-Golay smoothing. *Journal of Raman spectroscopy*, 38, p. 1174-1179.

CONAMA (1999). Catastro y evaluación de recursos vegetacionales nativos de Chile. *Proyecto CONAF-CONAMA-BIRF*. Recuperado el 11 de diciembre de 2009, de www.conaf.cl/cd_uso_suelos/nacional.pdf

Eichhorn S.J., Sirichaisit J. and Young R.J. (2001). Deformation mechanisms in cellulose fibres, paper and wood. *Journal of materials science*, 36, p. 3129-3135, UK.

Geladi P. and Kowalski B. Partial Least-Squares: A tutorial (1986). Laboratory for Chemometrics and Center for Process Analytical Chemistry. *Department of Chemistry, University of Washington, Seattle, USA*.

Hsieh E., Uy N. and Wallbacks L. (2006). Development of a portable spectroscopic sensor to measure Wood and fibre properties in standing mountain pine beetle-attacked trees and decked logs. *Pulp and Paper Research Institute of Canada, Mountain Pine Beetle Initiative PO# 8.14*.

INFOR (2008). Datos tecnológicos del *Eucalyptus nitens*. *Documentos técnicos*. Recuperado el 11 de diciembre de 2009, de www.nitens.cl/datos%20tecnologicos.pdf

Jones P.D., Schimleck L., Daniels R.F., Clark III A. and Purnell R.C. (2008). Comparison of *Pinus taeda* L. whole -tree Wood property calibrations using a variety of sampling options. *Wood Sci Technol* vol 42 p. 385-400.

McCreery R.L. (2000). Raman spectroscopy for chemical analysis. *Wiley interscience publication. United States*. p. 49-70

Megraw R.A. (1986). Douglas-fir wood properties. In Oliver C, D Hanley, J Johnson eds. Douglas-fir: Stand management or the future Institute of Forest Research. *College of Forest Resources, University of Washington, Seattle*. p. 81-96.

Microbeam. Espectrómetros y mini-espectrómetros Raman. *Manual de referencia, Barcelona*. Recuperado el 11 de diciembre de 2009, de www.microbeam.es/bwtek/raman.pdf

Mikhailyuk and Razzhivin. (2003). Background Subtraction in Experimental Data Arrays Illustrated by Example of Raman Spectra and Fluorescent Gel Electrophoresis Patterns. *Instrument and Experimental Techniques*, 46, p. 765-769.

Mora R., Schimleck L. and Isik F. (2008). Near Infrared calibration models for the estimation of wood density in *Pinus Taeda* using repeated sample measurements. *Journal of near infrared spectroscopy*, 16, p. 517-528.

Nuopponen M.H., Birch G.M., Skyes R.J., Lee S.J. and Stewart D. (2005). Estimation of wood density and chemical composition by means of diffuse reflectance mid-infrared Fourier transform (DRIFT-MIR) spectroscopy. *Journal of agricultural and food chemistry* 2006, 54, p. 34-40.

Ona T., Sonoda T., Ito K., Shibata M., Ootake Y., Ohshima J., Yokota S. and Yoshizawa N. (1999). In Situ Determination of Proportion of Cell Types in Wood by Fourier Transform Raman Spectroscopy. *Analytical Biochemistry*, 268, p. 43-48.

Ona T., Sonoda T., Ohshima J., Yokota S. and Yoshizawa N. (2003). A Rapid Quantitative Method to Assess *Eucalyptus* Wood Properties for Kraft Pulp Production by FT-Raman Spectroscopy. *Journal of pulp and paper science*, 29, no.1.

Schimleck L., Kube P.D., Raymond C.A, Michell A.J. and French J. (2005). Estimation of whole-tree kraft pulp yield of *Eucalyptus nitens* using near-infrared spectra collected from increment cores. *Canada, Research press*, 35.

Shimleck L., Michell A.J., Raymond C.A. and Muneri A. (1999). Estimation of basic density of *Eucalyptus globulus* using near-infrared spectroscopy. *Canada Journal for Research*, p. 194-201.

Singh R. (2002). C. V. Raman and the discovery of the Raman Effect. *Physics perspective, Birkhäuser Verlag, Basel*, p. 399-420.

Smith, E. y Dent, G. (2005). Modern Raman spectroscopy: a practical approach. *Hoboken, NJ: John Wiley and Sons*.

So C.L., Groom L.H., Rials T.G., Snell R., Kelley S. and Meglen R. (2002). Rapid assessment of the fundamental property variation of wood. In *Outcalt KW eds. Proceedings of the 11th Biennial Southern Silvicultural Research Conference*. USDA Forest Service, Southern Research Station, General Technical Report SRS-48.

Technical Association of the Pulp and Paper Industry (2002). T 258 om-02 Norm. *Norma internacional para la medición de la densidad básica en madera en la industria*.

Tian X. (1999). An application of computer vision technologies to log defect determination. *PhD Thesis*. Lincoln University, New Zealand.

Wise, Gallagher, Bro and Shaver. (2006). PLS_Toolbox 4.0 for use with MATLAB TM. *Manson, Eigenvector Research Inc. Manual de referencia*.

Wold S, Esbensen K. and Geladi P. (1987). Principal Component Analysis. *Chemometrics and Intelligent Laboratory Systems*, 2, p. 37-52.

Wold S., Sjöström M., Eriksson L. (2001). PLS regression: a basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems*, 58, p. 109-130.

Zobel B. and Van Buijtenen J. P. (1989). Wood variation: Its Causes and Control. *Springer-Verlag, Berlin*. 363 p.