



Universidad de Valladolid

FACULTAD DE CIENCIAS

TRABAJO FIN DE GRADO

Grado en Física

Análisis espectroscópicos de minerales análogos marcianos para dar soporte a las actividades científicas de las misiones ESA/ExoMars y NASA/Mars 2020

Autor/a: Sofía Julve González

Tutor/es/as: Guillermo López Reyes, José Antonio Manrique

Año 2022

Agradecimientos.

Me gustaría agradecer el esfuerzo y la cooperación recibida por parte de todos los miembros del grupo de investigación ERICA con el que se ha realizado este proyecto y con el que he tenido la suerte de formar parte. En particular agradecer tanto a los tutores de este proyecto G. López y J.A.Manrique como a M.Veneranda por la atención y ayuda proporcionada para la realización de este estudio.

También dar las gracias por el apoyo recibido por parte de a mis padres, familiares, compañeros, profesores y amigos que me han acompañado durante estos años en mi etapa del grado.

Abstract

This study has been carried out with the aim of understanding the mineralogy of the Martian surface using combined spectroscopic techniques. For this purpose, chemometric analyses based on semiquantitative analysis of binary mixtures of different carbonates by means of univariate and multivariate methods for LIBS and Raman spectroscopic techniques have been used. These have allowed us to obtain different models with which it is possible to evaluate the data collected by SuperCam on the Mars2020 mission.

In addition, a more qualitative study based on PCA analysis has been carried out to classify natural carbonate samples according to their elemental and molecular characteristics.

All these studies have been carried out for LIBS and Raman techniques individually and for the combination of the data obtained in both, thus demonstrating the complementary nature of these two techniques and the importance they acquire in planetary explorations.

Resumen

Este estudio ha sido impulsado con el objetivo de conocer la mineralogía de la superficie marciana haciendo uso de técnicas espectroscópicas combinadas. Para ello se han empleado análisis quimiométricos basados en análisis semicuantitativos de mezclas binarias de diferentes carbonatos por medio de métodos univariantes y multivariantes para las técnicas espectroscópicas LIBS y Raman. Estos nos han permitido obtener diferentes modelos con los cuales nos sea posible evaluar los datos recogidos por SuperCam en la misión Mars2020.

Además, se ha realizado un estudio más cualitativo basado en el análisis de PCA con el cual se han podido clasificar muestras naturales de carbonatos atendiendo a sus características elementales y moleculares.

Todos estos estudios se han hecho para las técnicas LIBS y Raman de manera individual y para la combinación de los datos obtenidos en ambas, demostrando con ello el carácter complementario de estas dos técnicas y la importancia que adquieren en exploraciones planetarias.

Índice general

1. Introducción.	1
1.1. La misión Mars 2020.	1
1.2. El Rover Perseverance.	2
1.3. Lugar de aterrizaje.	4
1.4. Análisis de carbonatos con SuperCam.	4
2. Fundamento teórico.	6
2.1. Fundamentos de la técnica LIBS.	6
2.1.1. Proceso de emisión espectral inducida.	6
2.1.2. Instrumento técnico LIBS	7
2.2. Fundamentos de la técnica Raman.	8
2.2.1. Teoría semiclásica del efecto Raman	9
2.2.2. Espectros Raman	10
2.2.3. Instrumento técnico Raman	11
3. Objetivos.	12
4. Materiales y métodos.	13
4.1. Selección y preparación de muestras.	13
4.2. Instrumentación.	14
4.2.1. Equipo LIBS	14
4.2.2. Equipo Raman	15
4.3. Recopilación de espectros.	15
4.3.1. Espectros LIBS	15
4.3.2. Espectros Raman	16
4.4. Análisis de datos.	17
4.4.1. Método Univariante	17
4.4.2. Métodos Multivariantes	17
5. Elección de picos.	19
5.1. Tratamiento de espectros.	19
5.1.1. Observaciones	19
5.2. Picos LIBS.	20
5.3. Picos Raman.	22
6. Resultados obtenidos.	24
6.1. LIBS.	24
6.1.1. $CaCO_3 - Mg_5(CO_3)_4(OH)_2 \cdot 4H_2O$	24
6.1.2. $CaCO_3 - FeCO_3$	26
6.1.3. $FeCO_2 - Mg_5(CO_3)_4(OH)_2 \cdot 4H_2O$	28
6.1.4. Muestras Naturales: PCA LIBS	29
6.2. Raman.	31

6.2.1.	$CaCO_3 - Mg_5(CO_3)_4(OH)_2 \cdot 4H_2O$	31
6.2.2.	Muestras Naturales: PCA Raman	32
6.3.	Fusión de datos LIBS-Raman.	33
6.3.1.	Modelo Univariante	34
6.3.2.	Modelo Multivariante	35
6.3.3.	Muestras Naturales: Combinación de datos con PCA	36
7.	Conclusiones.	38

Capítulo 1

Introducción.

Nuestro planeta vecino Marte es el segundo planeta más pequeño del sistema solar. Con la mitad de diámetro al de la Tierra, posee una atmósfera poco densa compuesta principalmente de dióxido de carbono al 96 % y añadiendo que está más alejado del sol que la Tierra, hace que las temperaturas sean muy frías, además su geología presenta hierro oxidado que le da ese color rojizo.

Desde la década de los 70, la NASA ha estado realizando diferentes misiones espaciales para inspeccionar la superficie del planeta rojo en busca de vida presente, pasada y posiblemente futura. Sobre estos años, diversas misiones aseguraron que la fina capa de la atmósfera no hacía favorable la habitabilidad de microorganismos vivos, aunque se reveló una rica historia geológica parecida a la de la Tierra, donde se vieron evidencias pasadas de lagos, ríos y océanos líquidos en la superficie marciana. A día de hoy, el agua existe en forma de nubes delgadas, aunque también en algunas laderas se han encontrado pruebas de agua líquida salada. Por otro lado, la composición rocosa del planeta, muestra evidencias de cráteres y volcanes actualmente inactivos, aunque la caracterización de rocas han mostrado que estos procesos fueron agentes geológicos mucho más activos en el pasado lejano.

Todas estas evidencias geológicas en la posible búsqueda de vida, han motivado una serie de misiones in-situ a lo largo de los años, para las cuales diferentes rovers han aterrizado en la superficie de Marte con el objetivo de investigar la estructura, composición, mineralogía y la química que compone este planeta ante la posibilidad de que fuese habitable en su pasado geológico.

1.1. La misión Mars 2020.

Sojourner, Spirit and Opportunity, Curiosity y Perseverance son los Rovers que han participado (por orden cronológico) en las diferentes misiones espaciales en busca de evidencia de vida en Marte a lo largo de los años. Todas ellas han sido parte de un proceso evolutivo para el entendimiento de las características del planeta rojo. El Perseverance ha logrado pisar la superficie marciana gracias a la misión Mars 2020. El enfoque principal de esta misión es buscar señales de biofirmas presentes en las rocas que componen la superficie [1]. Más concretamente, esta misión se centra principalmente en tres objetivos clave:

1. Desarrollar una comprensión y caracterización de la geología del lugar de aterrizaje y de exploración para estudiar la evolución histórica de las rocas marcianas.
2. Comprendiendo esta geología, el rover debe identificar rocas con alta probabilidad de habitabilidad microbiana antigua, es decir, reconocer trazas de vida sobre la mineralogía de la superficie marciana.

3. Realizar una selección y recogida de muestras de interés astrobiológico entendiendo que, estas muestras están más allá de las capacidades analíticas del Rover, identificando aquellos minerales donde los posibles biofirmas tengan una mejor conservación para su posible retorno a la Tierra en la futura misión *MARS Sample Return Mission* [2]. De esta manera se podrán llevar a cabo análisis químicos y físicos más profundos por medio de instrumentos de laboratorio.

Con todos estos objetivos, la misión debe ofrecer el conocimiento necesario para el desarrollo de futuras exploraciones con tecnología más avanzada y un avance de los conocimientos necesarios para ello.

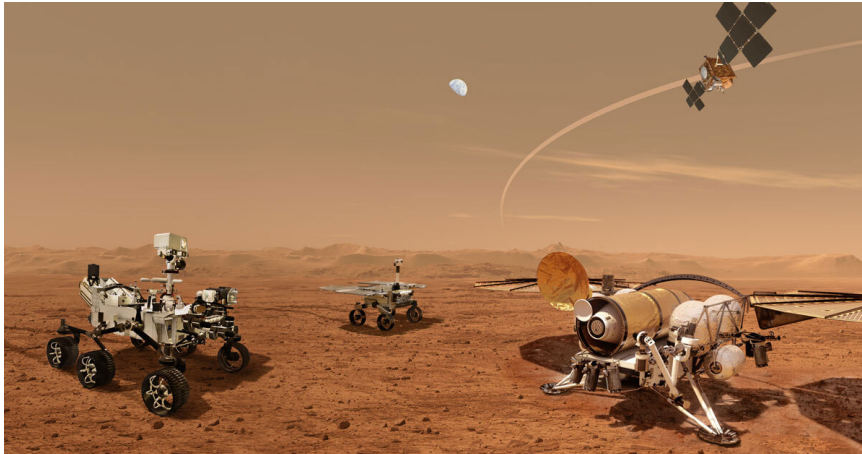


Figura 1.1: *MARS Sample Return: Se representan un conjunto de futuros rover que trabajarán para transportar muestras de la superficie de Marte recogidas por el rover de la NASA Mars Perseverance*

1.2. El Rover Perseverance.

La misión Mars Science Laboratory (MLS) fue lanzada el 26 de noviembre del 2011, en ella el Curiosity alcanzó la superficie de Marte compuesto de equipos técnicos con un alto grado de complejidad [3]. A partir de los exitosos resultados que obtuvieron sus componentes principales, se desarrolla el rover Perseverance, usado en la misión Mars 2020, el cual aterrizó en el planeta rojo el 18 de febrero del 2021.

La principal diferencia entre Curiosity y Perseverance se encuentran en la implementación de un nuevo sistema para tomar y almacenar muestras de rocas para un futuro retorno a la tierra. Además, Perseverance es capaz de realizar mejores análisis de las muestras, con más detalle de estudio gracias a la implementación de técnicas espectroscópicas más sofisticadas y punteras.

El Rover Perseverance consta de siete instrumentos, todos ellos enfocados para un determinado fin. En la figura 1.2 aparece la posición donde se encuentra cada una indicando su función principal.

Sherloc [4], Pixl [5] y SuperCam [6] se basan en diferentes técnicas espectroscópicas para caracterizar, mediante la combinación de estas técnicas, la composición del suelo marciano.

RIMFAX (Radar Imager for Mars' Subsurface Exploration) [7] es un radar de penetración en el suelo con el objetivo de interpretar las rocas sedimentarias estratificadas y las propiedades dieléctricas del subsuelo marciano para la identificación de entornos pasados de vida.

Moix (Mars Oxygen in Situ Resource Utilization (ISRU) Experiment) [8] es capaz de generar O_2 a partir de CO_2 . Además, Meda (Mars Environmental Dynamics Analyzer) [9] es capaz de

INSTRUMENTOS DE PERSEVERANCE

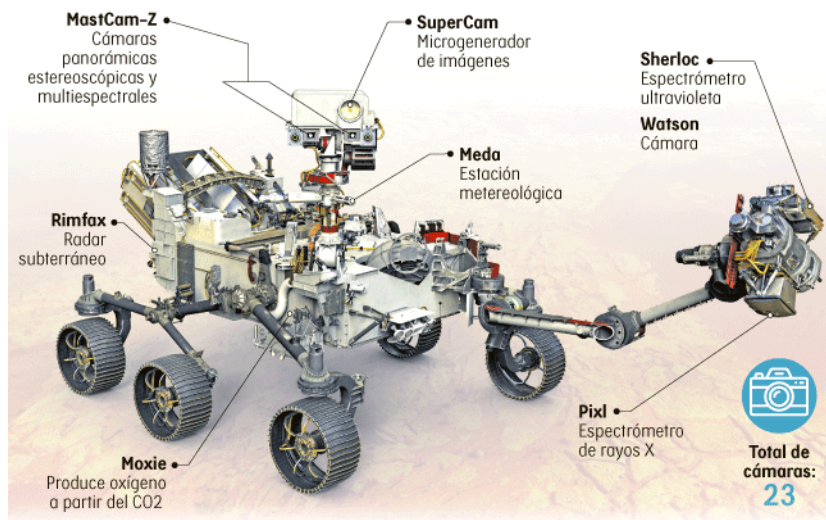


Figura 1.2: *Rover Perseverance*, con sus siete instrumentos principales indicando la situación de estos y su funcionamiento principal

medir las características climáticas. Meda incluye una cámara orientada hacia arriba con un ojo de pez (Skycam), la cual mide los ciclos diurnos y estacionales de la opacidad atmosférica y también de las nubes y sus morfologías.

Adicionalmente, otro de los instrumentos cuyos componentes principales son dos cámaras que funcionan como dispositivo de imagen estereoscópica multispectral es MaxCam-z [10]. Estas cámaras son capaces de resolver características de aproximadamente 1 mm de tamaño en un rango cercano y de 3-4cm a distancias de 100cm.

Desde nuestro estudio nos enfocamos en el instrumento SuperCam, ya que este es el encargado de realizar los análisis que nosotros simularemos para recrear nuestro modelo quimiométrico de optimización.

En detalle, consiste en un conjunto de diferentes elementos usados para la teledetección del color, la textura, la química y la mineralogía de la superficie de Marte. Para ello, SuperCam utiliza un conjunto de cinco técnicas: Espectroscopia de Ruptura Inducida por Láser (LIBS), Raman y Luminiscencia Resueltas en el Tiempo (TRR/L), Espectroscopía Visible-Infrarroja (VISIR) y Micro-Imagen Remota (RMI). Además cuenta con un micrófono de grabación de sonido [11].

Todos los dispositivos encargados para realizar todas estas técnicas están distribuidos en tres unidades del rover. En la unidad colocada sobre el mástil (MU) se encuentra un telescopio, el laser, el espectrómetro IR y un micrófono [12]. En la unidad del cuerpo (BU) están el LIBS, Raman y el hardware de mando y manejo de datos [6]. Por último, en la parte trasera, está el target Calibration (CT) [13].

Este último ha sido diseñado y procesado por el equipo de investigación Erica de la Universidad de Valladolid. El CT contiene muestras geológicas de interés primario en la exploración planetaria, objetivos para la calibración de la composición elemental y objetivos geométricos y de color para la calibración del generador de imágenes.

Las técnicas anteriormente mencionadas se encuentran co-alineadas de manera que todas ellas toman datos espectroscópicos sobre el mismo punto de interés. De esta manera se consigue optimizar las muestras minerales por medio de la complementariedad entre diferentes análisis espectroscópicos.

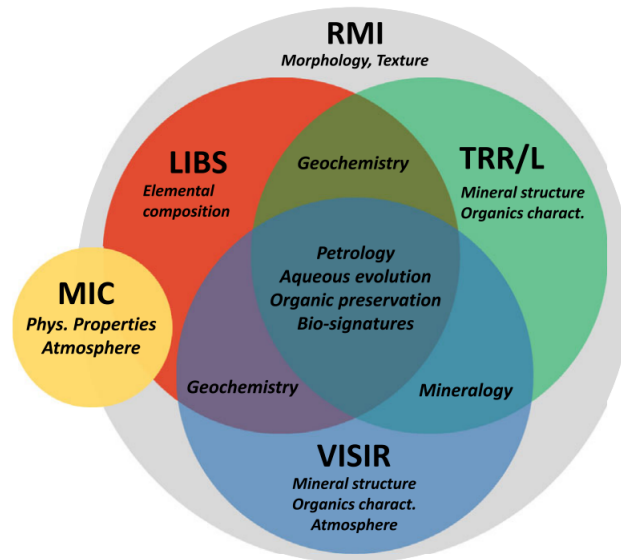


Figura 1.3: *Diferentes técnicas de las que consta SuperCam y la complementariedad analítica entre ellas.*

1.3. Lugar de aterrizaje.

Jezero es uno de los cráter de Marte de la edad *Noachian* de la cuenca de Isidis. Este cráter se formó por la colisión de meteoritos y cuyo tamaño es de 45km de diámetro.

El gran interés de este cráter es por su situación, ya que se encuentra dentro de una de las mayores unidades geológicas de carbonatos identificado en Marte, extendiéndose este por los bordes y algunos puntos internos del cráter Jezero [14] [15]. Además, se sabe que Jezero albergó en su día un paleolago de cuenca abierta, basándose en la presencia de valles de entrada y salida cortados en el borde del cráter. Ciertos estudios han demostrado que el cráter se llenó y desbordó durante un largo periodo del tiempo en el pasado del planeta. Esta presencia pasada de agua, hace un ambiente favorable ante el posible desarrollo de vida.

Por lo tanto, el cráter Jezero es el lugar de estudio donde se centra SuperCam y más en concreto nuestro estudio, debido a la presencia de carbonatos en la geología del planeta, procedentes de ambientes acuosos pasados.

Uno de los deltas es el conocido delta occidental que como su nombre indica se sitúa en la zona más occidental y pertenece al delta de un río en el cual se han visto consistencias espectrales con Fe y Mg en diferentes zonas del suelo del cráter.

Las manifestaciones de carbonatos en Jezero se producen en el margen interior noroeste del cráter, entre el delta occidental y el borde del cráter[16]. Los carbonatos lacustres podrían ser bioseñales de actividad viva en el pasado ya que, en la Tierra, el registro geológico de carbonatos suelen estar mediados biológicamente y pueden preservar eficazmente una variedad de bioseñales y texturas macro y microscópicas, isótopos, sustancias orgánicas y biominerales. Por tanto, una clasificación minuciosa de la detección de carbonatos en Jezero es una señal clave para alcanzar los objetivos de la misión.

1.4. Análisis de carbonatos con SuperCam.

Los antiguos carbonatos de Marte son un objetivo importante para futuras investigaciones y para el retorno de muestras de este planeta, tanto porque su composición isotópica podría servir como registro de la pérdida atmosférica en Marte como porque la precipitación de carbonatos en ambientes acuosos son un excelente mecanismo para la preservación de biofirmas [17].

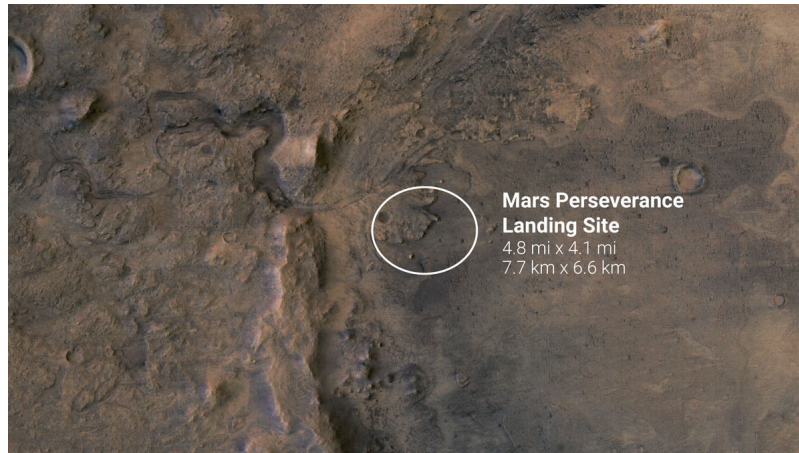


Figura 1.4: Localización del crater Jezero son de se muestra el lugar de aterrizaje de Perseverance y el delta del río.

Estudios recientes, demuestran la aparición de fases minerales de carbonatos en el crater Jezero así como la determinación de cuales de esas fases han sido ya analizadas y estudiadas en el laboratorio. Esta se muestran en la figura 1.5, en ella observamos el recorrido realizado por el rover y los resultados obtenidos de la composición de las muestras elegidas por el rover como posibles carbonatos.

Como complemento de estudios de carbonatos a este trabajo [18] se plantea nuestro estudio, en el que se amplía el análisis de fases minerales de carbonatos de laboratorio con el que se puedan comparar los resultados que nos llegan de Marte y así tener un mayor abanico de muestras.

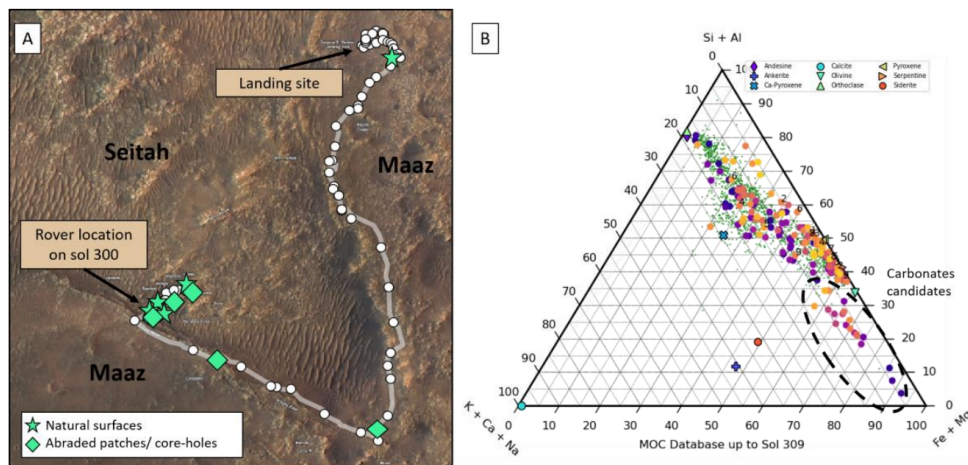


Figura 1.5: Vista preliminar de la detección de carbonatos en el cráter Jezero con SuperCam. A) Vista orbital del lugar de aterrizaje con el recorrido del rover durante los primeros 300 soles de la misión Mars2020, y la ubicación de nuestra lista actual de candidatos a carbonatos. B) Composición de los blancos que contienen los candidatos a carbonatos, basada en las Composiciones de Óxidos; nótese que los blancos de calibración de siderita y ankerita mostrados en el gráfico contienen 10-15 % de SiO_2

Capítulo 2

Fundamento teórico.

2.1. Fundamentos de la técnica LIBS.

LIBS (Laser-Induced Breakdown Spectroscopy) es una técnica espectroscópica de emisión atómica basada en la expansión del plasma inducido producido por la focalización de un pulso láser sobre el material [19]. Esta técnica permite conocer la composición elemental para muestras en cualquier estado de agregación. Los motivos que hacen a LIBS de gran interés para el análisis in-situ, se deben a sus características para la adquisición de espectros, ya que estos se originan de manera instantánea y los ensayos para tomarlos son mínimamente destructivos. Además, muestra un gran interés por la fácil complementariedad que tiene para combinarse con otras técnicas.

2.1.1. Proceso de emisión espectral inducida.

La emisión espectral inducida es el fundamento principal de este método analítico elemental. Para entender mejor como se produce esta emisión nos vamos a centrar de manera rigurosa en como es la creación del plasma sobre la materia y como este se expande produciendo las emisiones atómicas.

El láser es uno de los instrumentos principales que componen cualquier equipo LIBS. Principalmente se usan láseres pulsados de alta energía que sean capaces de desintegrar una pequeña cantidad de material y convertirlo en plasma.

Para entender el proceso de interacción láser-materia-plasma, introducimos el término de ablación. Este se define como el proceso por el cual una pequeña parte de material sufre cambios de fase progresivos hasta llegar al estado de vapor. La ablación se produce cuando el pulso láser focaliza sobre la muestra. La energía del láser es absorbida por el material y se transforma en energía térmica produciéndose, en este punto, las transiciones de fase y por tanto la vaporización del material. Si la energía del láser es suficiente para superar la energía umbral característica del material y así poder producir la vaporización, se genera el famoso plasma inducido. El tiempo que le lleva al material vaporizarse es menor que el tiempo del pulso láser, de forma que la interacción del láser con la nube de vapor ioniza la materia y se crea el plasma. Esta ionización y cambio de temperatura es tan rápida que la presión que ejerce el plasma producido con las condiciones ambientales externas es tan alto que se generan ondas de choque apreciables al oído humano.

Cuando el láser deja de incidir, el plasma sufre una expansión lo que produce una reducción en la presión en su interior y un decaimiento de la temperatura. Esto provoca que los átomos e iones que se habían excitado decaigan a estados de energía menores, emitiendo una radiación característica elemental y por tanto generando la emisión espectral del plasma.

El decaimiento de los átomos a estados de energía inferiores es recolectado por un sistema óptico que envía la información por medio de fibra óptica al espectrómetro, instrumento encargado de

transformar y digitalizar la información [20]. Todo este proceso se visualiza de forma esquemática en la figura 2.1.

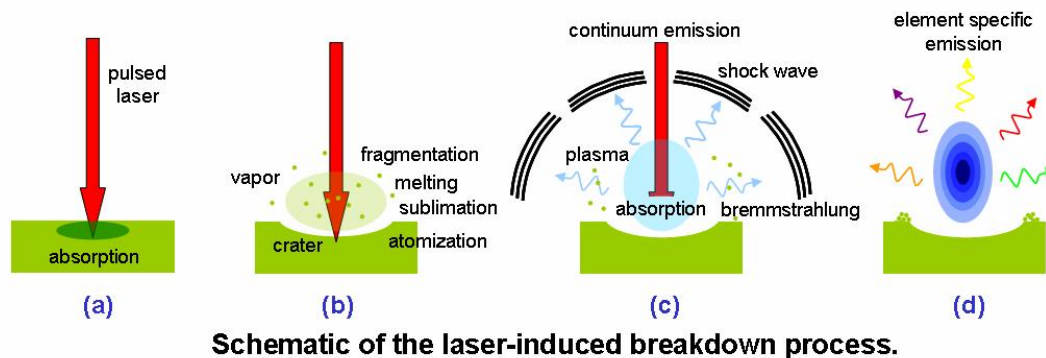


Figura 2.1: *Etapas de formación de plasma inducido: a) El láser incide sobre un punto de la muestra. b) Se produce la ablación, calentando el material y transformándolo hasta su estado de vapor. c) La interacción del láser calienta la nube de vapor generando el plasma y produciendo ondas de choque. d) Al terminar el pulso láser el plasma se expande y enfría produciendo las emisiones atómicas características.*

Aunque la técnica LIBS tiene un carácter analítico de gran utilidad, presenta una serie de problemas con las que tiene que combatir.

En primer lugar, la formación del plasma está influenciado por diversos factores medibles, las características del pulso láser como el spot (que viene dado por la longitud de onda), el ancho y la energía son algunos factores que afectan a la formación.

También, el sistema óptico de focalización influye; en función de la distancia focal que presente o las condiciones ambientales donde nos encontremos causarán un cambio de medidas del plasma. Otro de los problemas que más comúnmente afectan a las mediciones LIBS son los efectos de matriz, pudiendo ser estos tanto problemas físicos de la muestra como químicos del compuesto.

Por otra parte, los efectos de matriz también se ven afectados por la radiación electromagnética de frenado producido por los átomos y elementos del plasma. Ocurre justamente después de la ablación, en la fase de enfriamiento del plasma. Esta radiación de frenado se conoce como radiación de Bremsstrahlung y compite con nuestras emisiones atómicas elementales. El producto de este efecto es un continuo de radiación en los primeros instantes impidiendo la observación de las líneas atómicas.

La radiación de fondo Bremsstrahlung se corrige añadiendo un retardo entre la creación del plasma y la adquisición del espectro.

2.1.2. Instrumento técnico LIBS

Existen dos tipos de instrumentos LIBS. Por un lado está el LIBS resuelto a tiempo, técnica que se utiliza más comúnmente en los laboratorios. Por otro lado el LIBS a distancia, que se caracteriza porque no es necesario que la muestra estudiada se encuentre cerca ya que, gracias a sistemas ópticos se consigue focalizar el láser sobre un punto de interés. La ventaja de LIBS resuelto a tiempo radica en la posibilidad de variar los retardos y el sincronismo del pulso láser con la obtención del espectro.

De manera general, en la figura 2.2 se representan un esquema con los elementos principales de los que dispone un dispositivo experimental LIBS. En ella aparece una fuente láser que se focaliza sobre la muestra por medio de una lente convergente o un sistema de focalización. El

principal tipo de láser que se usa es el láser Nd:YAG [20]. La radiación que genera el plasma inducido en la muestra es recogida por medio de un sistema óptico colector, esencialmente se hace uso de cámaras colectoras conectadas por medio de fibra óptica a un espectrómetro. Este último aparato hace pasar la luz que le llega por una rejilla la cual se ve difractada por diferentes espejos, consiguiendo con ello la detección de las diferentes componentes espectrales que llegan al detector que transforma la información (fotones que le llega) en señales eléctricas para poder ser leídas e interpretadas por un ordenador.

El espectrómetro más común que se usa en esta técnica es el espectrómetro Echelle, además el detector puede estar basado en dos tipos de detección, CCD o ICCD.

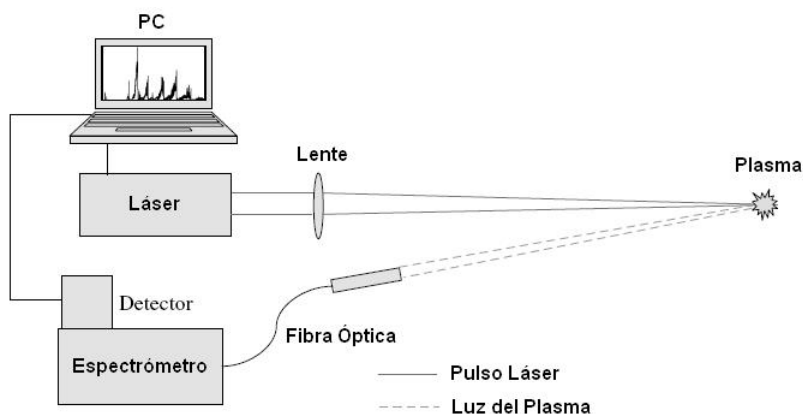


Figura 2.2: Esquema del equipo instrumental típico LIBS.

2.2. Fundamentos de la técnica Raman.

El estudio de la espectroscopia vibracional comprende el análisis de la estructura de la materia, estudiando la interacción entre radiación-materia. La materia está constituida por átomos y moléculas enlazadas entre si las cuales vibran en torno a sus posiciones de equilibrio. Estas vibraciones se realizan con una cierta frecuencia que es característica de la naturales de cada átomo que la compone.

Cuando un campo electromagnético incide sobre la superficie de cierto material, este experimenta excitaciones de diferentes tipos(ver figura 2.3). La excitación del momento dipolar permanece se estudia por medio de espectroscopia de Infrarrojo y excitación de la nube electrónica que compone la materia induciendo la polaridad se estudia por medio de espectroscopia Raman [21]. Es de esta última en la que se centra nuestro estudio.

El principio fundamental del cual se basa la técnica Raman consiste en la dispersión inelástica de la luz monocromática incidente de frecuencia (ν_0), provocado por la excitación de los fonones (coordenadas normales de vibración moleculares) propios de la materia. La radiación dispersada presenta un cambio en su frecuencia debido al intercambio de la energía entre la radiación incidente con el material. Estos cambios en la frecuencia dispersada son característicos de la naturaleza de cada compuesto y son las que nos dan el conocimiento necesario para identificar el tipo de muestra y su estructura molecular [22]. Este cambio de energía se puede calcular empleando $\Delta E = h\Delta\nu$.

Hay que notar, que cuando la luz interacciona con la muestra la, mayor parte de la radiación incidente es absorbida y solo una parte se dispersa. Además, la gran mayoría se dispersa elásticamente (dispersión Rayleigh), y solamente una mínima parte lo hace de manera inelástica produciendo la dispersión Raman. Además de todo esto, una de las mayores características que presenta esta técnica es debido a que no todos los modos de vibraciones moleculares se pue-

den apreciar con Raman. Existen unas reglas llamadas reglas de selección las cuales nos dicen que, para que una vibración molecular sea activa a Raman debe inducirse una variación en su polaridad: $\frac{\partial \chi}{\partial Q} \neq 0$.

2.2.1. Teoría semiclassical del efecto Raman

Cuando la luz monocromática incide sobre el material el fotón incidente produce una excitación de la nube electrónica de los átomos y moléculas. Las moléculas excitadas emiten un fotón con el fin de regresar a un estado de energía menor. Si este estado final es el mismo que el inicial se produce la dispersión elástica, pero si este estado tiene una energía mayor o menor que la que tenía desde un inicio entonces se produce una dispersión elástica y la molécula ya no estará en el mismo estado rotacional y vibracional. La frecuencia del fotón dispersado es la diferencia entre el estado de energía inicial y final. Por lo tanto, en función de si la frecuencia dispersada aumenta o disminuye nos encontramos con dos dispersiones Raman posibles, dispersión Raman Stokes, en la cual la frecuencia del fotón emitido es menor y por tanto se transfiere energía a la molécula, y dispersión Raman Anti-Stokes, en la cual la frecuencia emitida aumenta disminuyendo la energía de la molécula. La diferencia de energía se debe a la creación o destrucción de un fonón y da información de la energía de vibración en el centro de la primera zona de Brillouin.

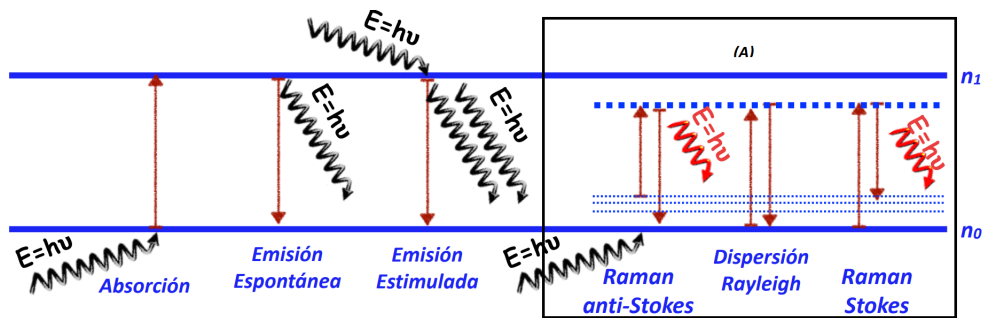


Figura 2.3: Diferentes tipos de intercambio de energía entre la interacción luz-materia. (A) Representación de las líneas de excitación Stokes, anti-Stokes y Rayleigh

Analizando las vibraciones moleculares por medio de sus frecuencias propias, es posible obtener información del tipo de molécula que se trata. Estas frecuencias dependen de la fuerza de enlace entre sus átomos y de la masa de estos.

Por tanto, como la materia está constituida por un conjunto de átomos y enlaces los cuales vibran y rotan con frecuencias características, cuando un trozo de materia entra en contacto con un campo eléctrico en este se induce un dipolo eléctrico generando un momento dipolar inducido.

$$\vec{P} = \chi \vec{E} \quad (2.1)$$

Donde χ representa la polarizabilidad del cristal como respuesta del campo eléctrico incidente y da cuenta de la facilidad de polarización del cristal. Podemos escribir la ecuación de onda de la luz incidente y del movimiento vibracional de los átomos como:

$$\vec{E} = \vec{E}_0 e^{(k_0 \vec{x} - \omega_0 t)} \quad (2.2)$$

$$\vec{Q} = \vec{Q}_0 e^{(q_0 \vec{x} - \omega t)} \quad (2.3)$$

El vector \vec{Q} indica la coordenada normal de las vibraciones reticulares (fonones), que sitúa a los átomos en movimiento a lo largo de su eje molecular. El vector \vec{E}_0 es la amplitud del campo

eléctrico incidente, \vec{k}_0 y \vec{q}_0 los vectores de onda de la radiación incidente y del fotón dispersado, ω_0 y ω las frecuencias incidente y dispersada, x la posición y t el tiempo. Las vibraciones que se presentan en la red cambian las distancias entre átomos y pueden expresarse en desarrollo de serie de Taylor:

$$\chi = \chi_0 + \sum_{i=1}^{i=n} \left(\frac{\partial \chi}{\partial Q_i} \right) + \frac{1}{2} \sum_{i,j=1}^{i,j=n} \left(\frac{\partial^2 \chi}{\partial Q_i \partial Q_j} \right) Q_i Q_j + \dots \quad (2.4)$$

Que sustituyendo en la ecuación 2.1 nos queda:

$$\vec{P} = \underbrace{\chi_0 \vec{E}_0 e^{i(k_0 \vec{x} - \omega_0 t)}}_{\text{Dispersión Rayleigh}} + E_0 \underbrace{\sum_{s=1}^{s=n} \left(\frac{\partial \chi}{\partial Q} \right) Q_0 e^{i[(k_0 \pm \vec{q}) \vec{x} - (\omega_0 \pm \omega) t]}}_{\text{Dispersión Raman}} + \dots \quad (2.5)$$

Para calcular la frecuencia de radiación dispersada Raman aplicamos las leyes de conservación de la energía considerando que, en la ecuación (2.5) el signo - hace referencia a la dispersión Stokes y el signo + a la dispersión Anti-Stokes. Por tanto, la dispersión Raman Stokes viene dada por $\hbar\omega = \hbar(\omega_0 - \omega)$ y la Anti-Stokes por $\hbar\omega = \hbar(\omega_0 + \omega)$. Como a energía esta cuantizada, las bandas Stokes y y Anti-Stokes se sitúan simétricas respecto de la Rayleigh, por este motivo se toma como referencia ω_0 .

2.2.2. Espectros Raman

Un espectro Raman es una representación de la intensidad Raman en función del número de onda ($\nu = 1/\lambda$). La intensidad de las bandas Stokes son proporcionales a la sección eficaz de dispersión (2.6).

$$\frac{d\sigma}{d\Omega} = V \omega_0^4 | \vec{e}_s (TR) \vec{e}_0 |^2 \{n + 1\} \quad (2.6)$$

V es el volumen de interacción, y n la población de fonones. TR es el tensor Raman el cual contiene información sobre la simetría de la red y viene dado por la derivada del tensor polarizabilidad. Además, \vec{e}_s y \vec{e}_0 corresponden a los vectores de polarización del haz dispersado y del incidente respectivamente.

El término $n+1$ indica que la intensidad de líneas de Stokes involucra el elemento correspondiente a la creación de fonones. A diferencia de esto, la intensidad de líneas anti-Stokes interviene en la aniquilación del fonón.

Analizando la intensidad de los picos presentes en la representación espectral Raman, es de clara identificación que el pico correspondiente a la dispersión Rayleigh tendrá una intensidad mucho mayor que la que se obtiene para dispersión Raman ya que hay una dispersión raman por cada 10^5 dispersiones Rayleigh (la cantidad de fotones dispersados elásticamente es mucho mayor que inelásticamente). Además de esto, es interesante mencionar que la intensidad de las bandas Raman Stokes son más significativas que las Anti-Stokes ya que, a temperatura ambiente, la gran parte de las moléculas partían de estados con mínima energía (los electrones se encuentran en su estado fundamental) y de esta manera la probabilidad de encontrar estados finales de mayor energía aumenta. Esta afirmación se expresa en la ecuación (2.7), con N_i el número de moléculas en el estado i y ΔE la diferencia de energía entre el estado inicial y el estado final i [23].

$$\frac{N_i}{N_0} = e^{\frac{-\Delta E}{k_B T}} \quad (2.7)$$

Los espectros Raman pueden presentar una calidad espectral baja debido a diversas interferencias provocadas principalmente por ruido proveniente de fondo, de la absorción de radiación o del fenómeno de fluorescencia[24], siendo este uno de los principales problemas con los que se combate. Este último fenómeno se presenta como una curvatura en la línea base la cual puede llegar a niveles de enmascarar las bandas Raman. Estas interferencias se miden con la relación señal-ruido (SNR) [25].

2.2.3. Instrumento técnico Raman

El esquema más típico de un dispositivo experimental Raman es el que se representa en la figura 2.4. Está formado por diferentes instrumentos, entre ellos destacan una fuente de excitación láser que focaliza sobre la muestra gracias a un cabezal óptico el cual funciona como lente y contiene un filtro polarizador que elimina las frecuencias no deseadas. Tras la interacción radiación-materia, la luz dispersada es recogida por un cabezal óptico cuya información recolectada se envía al espectrómetro por medio de cable coaxial. El haz dispersado llega al espectrómetro haciendo pasar este previamente por una rejilla de tal forma que solamente es posible el paso del primer haz. El espectrómetro es el encargado de separar las distintas frecuencias que le llega por medio de un monocromador de doble track. Las diferentes frecuencias se reflejan en un espejo de manera que se dirijan al detector CCD [26].

El detector se encarga de digitalizar los fotones que le llegan de manera que podamos visualizar una imagen espectral en el software de control propio del espectrómetro usado.

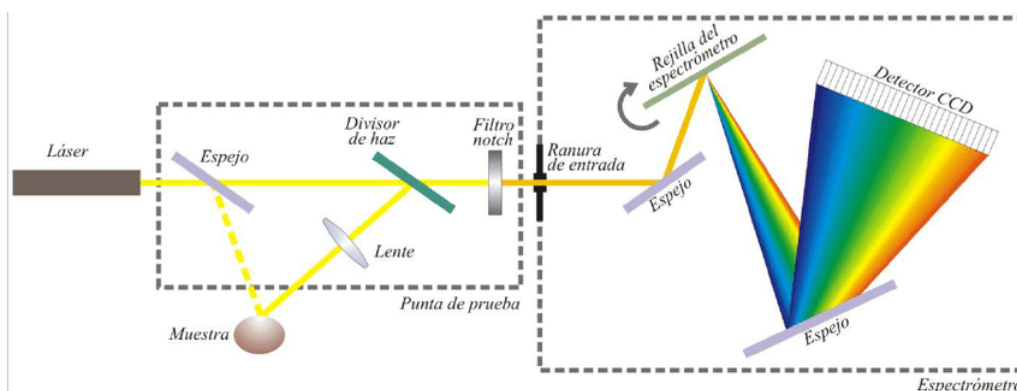


Figura 2.4: Esquema del equipo instrumental Raman

Capítulo 3

Objetivos.

Ante una visión general de los retos científicos planteados hasta el momento y ante la importancia descrita sobre la aparición de carbonatos como indicadores de bioformas de vida, se plantea nuestro estudio de laboratorio basado en análisis multivariante y en la combinación de datos Raman y LIBS para mejorar la discriminación de carbonatos en Marte.

La finalidad de este planteamiento se centra en una correcta identificación elemental y molecular de estas fases minerales, de ahí la elección de desarrollar un modelo basado en muestras binarias de carbonato, gracias a su gran importancia para la habitabilidad de las zonas investigadas. Para conseguir este fin se han realizado diferentes partes en nuestro estudio.

- Por un lado, se han creado mezclas binarias de carbonato con el fin de construir modelos de calibración para las técnicas de Raman y LIBS.
- Con los datos obtenidos se desarrollan diferentes algoritmos basados en el análisis univariante y multivariante para cada técnica.
- Además, se crean modelos de fusión de datos los cuales den cuenta del carácter complementario que presentan ambas técnicas.
- Mostrar el poder de condensación de información en un número reducido de componentes por medio del análisis con PCA, para ello se realiza una clasificación de muestras naturales de carbonatos observando la discriminación de las fases minerales para cada técnica espectroscópica.

Capítulo 4

Materiales y métodos.

4.1. Selección y preparación de muestras.

Los carbonatos son compuestos de poca dureza y colores diversos, se caracterizan por la presencia del anión común CO_3^{2-} . Sus componentes mayoritarios son la calcita, la dolomita y mármoles cristalinos.

Como ya se ha comentado, el estudio de fases minerales de carbonatos juega un gran papel en la astrobiología, de ahí nuestro interés por este tipo de compuestos. Para el desarrollo de un modelo de calibración basado en la identificación de estas fases minerales, se ha hecho uso de dos tipos de muestras diferentes.

Por un lado, se han realizado muestras sintéticas, a partir de reactivos de laboratorio completamente puros, de carbonato de calcio, carbonato de hierro e hidromagnesita ($CaCO_3$, $FeCO_3$ y $Mg_5(CO_3)_4(OH)_2 \cdot 4H_2O$). En particular, se han fabricado mezclas binarias de estos tres compuestos en diferentes proporciones, con el fin de conseguir muestras entre todos los puntos intermedios por cada par de compuestos. Así conseguimos un abanico de muestreo bastante amplio para nuestro estudio.

El formato elegido para la realización de las mezclas ha sido en forma de pastilla. En concreto, se ha creado un gramo de muestras mezclando los compuestos con las cantidades indicadas en la tabla 4.1. Hay que resaltar que cuando hablamos de compuesto A y B son los que corresponden a los carbonatos usados para las crear cada recta, es decir $CaCO_3 - FeCO_3$, $CaCO_3 - Mg_5(CO_3)_4(OH)_2 \cdot 4H_2O$ y $FeCO_3 - Mg_5(CO_3)_4(OH)_2 \cdot 4H_2O$. Además, se tiene mayor cantidad de muestras en los límites, donde la concentración de uno de los compuestos tiene mayor peso sobre el otro.

Compuesto A/mg	1000	990	950	900	750	625	500	375	250	100	50	10	0
Compuesto B/mg	0	10	50	100	250	375	500	625	750	900	950	990	1000

Tabla 4.1: Cantidad de compuesto para las diferentes pastillas

El proceso de fabricación de las pastillas se ha realizado en diferentes fases. El primer paso y de fundamental importancia, ha sido la mezcla de los dos compuestos. Para ello se pesan las cantidades indicadas haciendo uso de una balanza de precisión. Seguidamente se colocan las cantidades en un mortero donde ambas se mezclan y machacan. Con esto se consigue un doble objetivo, mezclar los compuestos para conseguir muestras lo más homogéneas posibles y reducir posibles errores a la hora de su identificación, y la disminución del grano, haciendo que todos los planos cristalográficos y niveles atómicos de las diferentes fases minerales estén presentes en la muestra a la hora de examinarlas con LIBS o Raman. Un par de minutos serán suficientes para conseguir que las muestras sean completamente homogéneas. A continuación, se toma un gramo de la mezcla resultante y se coloca en un molde cilíndrico de 1,5 cm de diámetro y

pocos milímetros de espesor, el cual da forma a nuestras pastillas. Por medio de una prensa hidráulica manual de 10 toneladas, se prensa el polvo de la muestra hasta formar una pastilla completamente sólida. En total se han obtenido 36 pastillas de mezclas binarias sintéticas. Se han presentado dificultades a la hora de prensar algunas de las muestras, especialmente en las que tenían un mayor porcentaje de $CaCO_3$. En algunos casos, estas estallaron al aplicar la presión indicada con lo que tuvimos que repetir algunas de las pastillas.

El otro tipo de muestras usadas han sido minerales completamente naturales de carbonatos puros, como la Calcita, Aragonita, Siderita y Magnesita, y de fases minerales entre ambas, incluyendo es este grupo la Anquerita, Huntita y Dolomita. Todos estos compuestos se han analizado previamente con espectrometría de masas para verificar su pureza. Además, todas estas muestras de carbonatos vienen de la base de datos ADAMM [27].

Ya que lo que se pretende es recrear un ambiente similar al que encuentra SuperCam en Marte, los análisis de estas muestras se han hecho de dos maneras diferentes. La primera ha sido directamente sobre el mineral en cuestión, indicando el punto exacto de medición sobre cada compuesto aunque no se garantiza la homogeneidad. Por ello, se han creado pastillas siguiendo un proceso idéntico al expuesto anteriormente, tomando como material de muestra un trozo del mineral a examen, aunque para este caso analizar un gramo de mineral no necesariamente está relacionado con la fase mayoritaria.

En total se han usado 27 muestras naturales diferentes. En concreto, 8 Sideritas y 4 Magnesitas han sido medidas directamente sobre el mineral, tomando medidas sobre diferentes caras. Además, se han creado 15 pastillas de diferentes muestras naturales: 2 Magnesitas, 3 Sideritas, 3 Calcitas, 2 Aragonitas, 3 Dolomitas y dos mezclas de minerales de calcita y magnesita puras (MITGamooham) y dolomita y cuarzo (MITPongola).

En conclusión, se han creado un total de 51 pastillas, 36 de ellas a partir de reactivos del laboratorio y 15 de minerales naturales, además de usar otros 12 compuestos naturales para su directa medición. Esto nos genera un amplio espacio muestral del cual podemos tomar infinitud de medidas, pudiendo general diferentes técnicas estadísticas para el análisis de los datos y una mayor interpretación de los mismos.

4.2. Instrumentación.

Tanto para la técnica LIBS como en Raman, el esquema general para la adquisición de espectros tiene bastantes rasgos comunes. En ambos se usa un láser de alta energía, un sistema de focalización, un equipo óptico de colección y un espectrómetro detector.

Los equipos usados para la elaboración de este estudio han sido proporcionados por la unidad asociada. Para llevar a cabo un correcto uso de los instrumentos de medida, se han tenido que adquirir conocimientos previos tanto del manejo como de los funcionamientos físicos en los que estos se basan.

4.2.1. Equipo LIBS

Los diferentes componentes con sus respectivas características que forman nuestro equipo LIBS son las siguientes:

- Laser pulsado *Nd:YAD* (2nd Harmonic) de 532nm, con un pulso de 120mJ, operando a 2Hz.
- Lente convergente de 55 mm de diámetro y 75 mm de focal. Es la encargada de focalizar los pulsos láser en un punto concreto de la muestra.

- Cámara colectora *iStar CCD (Andor)*: se trata de una cámara CCD intensificada y rápida, con matriz de 1024x1024, cuyo tiempo de exposición es de 2ns o mejor. Toma la radiación generada por el plasma y la transforma para llevar la información hasta el espectrómetro.
- *Espectrómetro Echelle (Andor)*: con resolución de 0,1nm y un rango de detección entre 200 y 975 nm. Su distancia focal es de 195mm. Este componente se encarga de la lectura de la lectura y análisis de la radiación procedente del plasma.
- *Software Andor Solis*: Obtiene, de manera automática, el espectro calibrado de longitud de onda completa a partir de una imagen echelle compleja. Ofrece un amplio rango de funciones para la adquisición de datos y su procesado.

4.2.2. Equipo Raman

Los componentes del equipo usados en esta técnica son :

- Laser de onda continua *BWF-OEM* de 532nm: diodo láser de alta potencia y compacto con refrigerador eléctrico térmico.
- Cabezal Raman *BWTEK BAC100-532E*: se trata de una sonda de focalización conectada al espectrómetro por medio de fibra óptica
- Espectrómetro de red de doble track para tomar medidas espectrales.

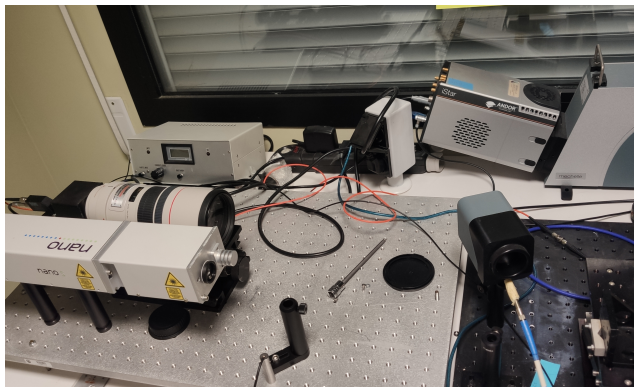


Figura 4.1: *Dispositivo experimental LIBS*

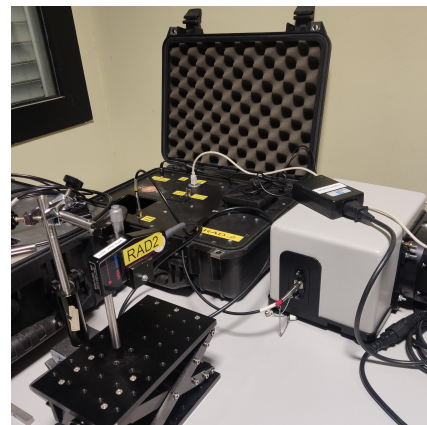


Figura 4.2: *Dispositivo experimental Raman*

4.3. Recopilación de espectros.

La adquisición de los espectros, tanto para LIBS como Raman, se ha realizado de manera metódica y cuidadosa, supervisando los espectros uno a uno según se iban tomando. Los instrumentos usados para cada técnica son los indicados en la sección 4.2 de este documento.

Se han tomado espectros de las 63 muestras disponibles, todas ellas usando los mismos parámetros de adquisición.

4.3.1. Espectros LIBS

El dispositivo LIBS consta de un láser pulsado de alta energía y 532nm el cual incide perpendicularmente sobre el punto de interés de la muestra con ayuda de una microlente. Sobre este mismo punto, se encuentra alineada una cámara que funciona como detector CCD.

Para hacer coincidir estos dos aparatos sobre el mismo punto, se usa un láser rojo que está conectado a la cámara detectora por medio de un cable coaxial. Para conseguir la alineación, se da una baja intensidad del láser pulsado, y con el láser rojo enfocando a la pastilla, movemos la muestra de manera que el láser pulsado genere plasma en la posición del láser rojo, justo en la superficie de la pastilla. Una vez conseguido el alineamiento, se cambia el láser rojo por la fibra óptica conectada al espectrómetro Echelle y asegurándonos de que el punto de máxima irradianza (donde se genera el plasma en la atmósfera) queda dentro de la muestra, a fin de minimizar la huella atmosférica en las medidas.

Se establecen ciertos parámetros que han de ser constantes durante toda la obtención de los diferentes espectros, estos son:

- *Parámetros del Láser y la cámara colectora:* La frecuencia elegida del láser ha sido 2Hz por pulso y los parámetros del gate de la CCD han sido $2\mu\text{s}$ de tiempo de retardo entre el pulso láser y la recolección del espectro y $10\mu\text{s}$ de tiempo de activación del intensificador.

Otro parámetro imprescindible, el cual ha jugado un papel importante, es la intensidad del láser. Lo que se ha ido modificando es el % de la intensidad máxima (120mJ) de potencia del láser por cada pulso, teniendo especial cuidado, ya que, si la muestra está expuesta un tiempo relativamente grande a alta energía, las pastillas terminaban quemándose y perforándose, pudiendo llegar a tomar espectros del porta de vidrio donde se encuentran colocadas en vez de las muestras en cuestión. Pero también, la energía de los pulsos de luz tienen que ser suficientemente altos como para poder formar plasma sobre la muestra. De este modo, se ha tenido que encontrar un equilibrio en la proporción de intensidad del pulso, en concreto este parámetro ha sido del 80 % de la intensidad máxima, es decir, pulso de 96mJ de energía.

- *Parámetros del espectrómetro:* Los parámetros que se usan para la adquisición de los espectros se colocan en el software correspondiente. Se han tomado 20 acumulaciones por punto medido, tomando un total de 40 puntos por pastilla. Para cambiar el punto donde incide el láser se hace de manera manual, por lo que hay que tener cuidado de no modificar la alineación del pulso de luz láser con el detector.

En total se han tomado un total de 2480 espectros LIBS de los cuales 1400 corresponden a las muestras sintéticas y los 1080 restantes a las naturales.

4.3.2. Espectros Raman

Para una buena adquisición de los espectros Raman ha sido imprescindible seguir una serie de pasos:

El primer paso, y el más tedioso, es el correcto enfoque en el punto deseado. El enfoque se realiza manualmente ajustando el tiempo de integración de manera que nos garantice el uso del 70 % del rango dinámico del detector. El tiempo de exposición en cada caso va a cambiar dependiendo del compuesto que tengamos, por este motivo un adecuado enfoque y elección del tiempo de exposición es un paso crucial para las medidas de las diferentes pastillas.

El software nos proporciona un espectro que relaciona el número de cuentas con el número de píxeles. Para hacer la transformación a intensidad y RamanShift se usa un software de calibración que ha sido previamente diseñado por miembros de la unidad asociada.

Es importante que antes de cada medida, se tome el background correspondiente, ya que este será posteriormente restado a nuestros espectros. Además, este dark tiene que tener los mismos parámetros de adquisición que posteriormente usemos para cada medida. Se toman 20 espectros por muestra en diferentes puntos, cada uno con un total de 10 acumulaciones.

En total se han adquirido 560 espectros, de los cuales 260 son de muestras sintéticas y los 300 restantes a muestras naturales.

4.4. Análisis de datos.

El análisis, la representación y la interpretación de los datos tomados se generan en base a diferentes técnicas matemáticas y estadísticas. Por un lado, se plantean métodos semicuantitativos basados en cálculos quimiométricos, con los que seamos capaces de desarrollar un modelo que detecte la composición mineralógica de diferentes fases minerales de carbonatos.

Estos modelos se crean a partir de las muestras sintéticas preparadas en el laboratorio, de las cuales conocemos su pureza y concentración. En particular, se desarrollan métodos univariantes y multivariantes. El modelado del método univariante se desarrollada mediante técnicas de dispersión de datos y regresiones polinómicas. Por otro lado, el método multivariante se basa en técnicas de Inteligencia Artificial y de Aprendizaje Automático.

Con esto, vamos a poder interpretar la bondad en términos diferentes para cada método mencionado, ya que el modelado usado para cada una será diferente.

Además del desarrollo de modelos semicuantitativos, se usan los datos recolectados de las muestras naturales para generar un estudio cualitativo de estas basado en el uso de PCA y representación gráfica por componentes principales.

La principal herramienta usada para el desarrollo de estos métodos ha sido el software de Matlab.

4.4.1. Método Univariante

El análisis univariante es la técnica más sencilla en el tratamiento de datos, ya que solamente se usa una de las variables que describen al sistema. Su objetivo principal es describir los datos y encontrar patrones de ellos.

Por este motivo, la primera idea fue usar la estadística descriptiva con el objetivo de clasificar y analizar los datos obtenidos a partir del análisis de una de las características que describen nuestros espectros. En particular, se ha usado como variable de entrada la intensidad de los picos más significativos relacionándolos con las concentraciones de cada muestra.

Los picos seleccionados se usan para realizar la base estadística del modelo, generando distintos indicadores para cada muestra. El calculo de cada indicador se ha realizado como se indica en la ecuación (4.1), tomando como variables A y B las intensidades normalizadas de los picos que corresponden a los diferentes carbonatos puros usados para la creación de cada mezcla binaria.

$$IND(A) = \frac{Variable(A)}{Variable(A) + Variable(B)} \quad (4.1)$$

Se crean indicadores para todas las posibles combinaciones de los diferentes picos de corresponden a cada compuesto. Además, se realizan gráficas de dispersión las cuales nos representan el valor del indicador en función de la concentración de cada compuesto para todos los indicadores creados. Para elegir la curva de calibración que mejor resultados proporciona, nos fijamos en la forma de la dispersión de los datos, aproximando al polinomio de grado menor que más se ajuste a este y nos proporcionamos un coeficiente de correlación mayor. De esta manera se obtienen tres curvas de calibración correspondientes a cada par de muestras.

4.4.2. Métodos Multivariantes

Los métodos multivariantes se basan en técnicas estadísticas cuyo objetivo es analizar simultáneamente conjuntos de datos, es decir, una distribución de probabilidad de varias variables que funcionan como respuesta de las distribuciones de los datos obtenidos. Esta técnica estadística

es mucho más compleja que la descrita anteriormente, aunque genera un mayor y mejor entendimiento del objeto a estudio.

En concreto, esta técnica se ha llevado a cabo por medio de Redes Neuronales Artificiales (ANN). Las ANN consisten en un código de programación el cual enseña a procesar los datos cambiando la estructura de la red en función de la información que esta adquiera, es decir, programar al ordenador para que trabaje simulando un cerebro humano en el proceso de aprendizaje.

Se eligen, como variables de entrada de la red, las intensidades de los picos más significativos de cada compuesto puro. Con una previa selección de todos ellos, se preparan los datos en forma de matriz, de tal modo que recoja todas las entradas seleccionadas (intensidades de los picos normalizadas) para todos los espectros tomados y haga una división del 70-15-15 % de los datos que se toman para cada muestra. El 70 % se usa para entrenar la red, el primer 15 % se usa para validación y prevenir el sobreentrenamiento de las redes neuronales, esto es que el modelo sea muy bueno separando espectros de un grupo empleando facetas no necesariamente útiles, como el ruido. El segundo 15 % se usa para testear el modelo, aplicando sobre este grupo el modelo previamente entrenado para evaluar su desempeño.

Además de las entradas, se le presenta a la red los targets a usar, es decir, una matriz que recoge las concentraciones de los compuestos que forman cada mezcla binaria. Como resultado, el modelo que se genera nos proporciona como salida (output) las concentraciones calculadas de las muestras reservadas para test una vez pasadas por el modelo entrenados en función de los targets indicados (concentraciones teóricas). Con esto se consiguen medidas directas de las concentraciones de las muestras.

Adicionalmente, se ha realizado un análisis más cualitativo por medio del análisis de datos con PCA (Principal Component Analysis) [28]. Este es un método estadístico multivariante que permite simplificar la información de entrada de múltiples variables en solo unas pocas componentes, generando prácticamente el total de información que se presentaba inicialmente, es decir, explica la varianza de los datos con el menor número de variables. Esta técnica trabaja comparando e identificando las posibles variables a usar, reduciendo automáticamente el número de variables que describen nuestras mezclas y que permite un estudio cualitativo.

Nosotros hemos empleado las tres primeras PCs, las cuales nos dan información de todo el sistema. PCA mira el total de datos, tratando de explicar las diferencias entre todos los espectros con un nuevo conjunto de variables. De este modo nos permite describir el sistema por medio de una representación gráfica de las componentes principales y así tener una idea más intuitiva de como se diferencian las distintas muestras naturales. Se han representado estas tres PCs en el espacio ortogonal de coordenadas de PC para cada técnica, de manera que se forma un patrón en el cual los espectros con características similares están agrupadas.

Capítulo 5

Elección de picos.

La variación de las características de los espectros LIBS y Raman para diferentes muestras hace de gran interés su análisis. En este apartado nos centramos en explicar las correcciones previas que se han llevado a cabo para los diferentes espectros, así como la elección de los picos seleccionados propios de cada compuesto o elemento las cuales serán usadas como variables de entrada para los distintos métodos analíticos descritos anteriormente.

Entendiendo el espectro como la representación gráfica de la distribución de la intensidad en función de su longitud de onda o energía asociada, se pueden realizar diferentes análisis partiendo de diferentes características que posee cada pico, como su intensidad, anchura, área bajo la curva, Gaussiana, etc.

Como ya se ha indicado, nuestro estudio se centra en la intensidad de los picos principales característicos de cada compuesto y propios de cada técnica espectroscópica. Todo el proceso de elección de bandas y tratamiento de espectros se ha llevado a cabo gracias al software de control '*SpectPro*' desarrollado por el grupo de investigación ERICA [27].

5.1. Tratamiento de espectros.

Además de una correcta elección de picos, con *SpectPro* se han realizado diferentes observaciones cualitativas de los espectros. Gracias a las diferentes herramientas que posee, se ha podido determinar los parámetros exactos para una buena corrección de la línea base y una buena normalización.

La normalización se realiza quitando offset y dividiendo por la intensidad máxima del espectro. Una vez realizada, se han ido comparando los diferentes picos de las mezclas binarias entre sus intensidades relativas correspondientes para ver la correlación entre cantidad de compuesto e intensidad de picos.

Otro parámetro que ha sido de gran utilidad es la SNR, que mide la relación de señal-ruido del espectro.

Para el caso de las muestras naturales en Raman, se ha ajustado la calibración para corregir posibles desplazamientos en la posición de las bandas, ya que estos han sido alterados a la hora de tomar las medidas con el láser Raman [29].

5.1.1. Observaciones

Se ha tenido en cuenta las limitaciones espectrales que poseen ambas técnicas ya que a la hora de tomar los espectros se ha lidiado con diversas dificultades.

En cuanto a los espectros tomados con LIBS, una de las mayores objeciones a simple vista, que se observó durante el proceso de adquisición de espectros, ha sido la gran variación en la intensidad de los espectros de una adquisición a otra. Esto puede ser debido a diferentes causas,

entre ella destaca los efectos de matriz que no se han tenido en cuenta, además de las posibles alteraciones en la potencia del láser en cada pulso provocando cambios en el plasma inducido sobre la muestra.

Asimismo, un número alto de adquisiciones mostraron que algunos de los espectros se encontraban saturados. Como consecuencia, se repitieron las medidas, disminuyó el número de adquisiciones por punto a 20, para las muestras que habían presentado esta dificultad.

En el caso de Raman, el mayor problema que se nos ha presentado ha sido la luminiscencia que muestran cierto tipo de compuestos, nos referimos a los compuestos sintéticos con un alto contenido en carbonato de hierro, para los cuales ha sido inviable la identificación de sus picos principales. Los peores resultados obtenidos para las pastillas con alto contenido en hierro se deben a una menor SNR en los picos, que limitan los límites de detección y la capacidad de separación.

5.2. Picos LIBS.

La elección de los picos usados para caracterizar los elementos principales que constituyen las diferentes muestras, se realizan por simple comparación entre todos los espectros tomados para las muestras sintéticas puras con las muestras naturales que corresponden a ese mineral, es decir, se compara cada carbonato sintético puro con sus naturales correspondientes, y así poder hacer una buena determinación de los picos propios de cada elemento. Además de las comparaciones entre espectros tomados en el laboratorio, se ha corroborado la identificación de estos picos gracias al uso de bases de datos [30].

Con esto, lo que se obtiene es una caracterización, por medio de los picos y sus intensidades correspondientes para cada carbonato.

Los picos se eligen de manera que no solapen entre los diferentes compuestos que tenemos, ya que esto nos conduciría a resultados erróneos provocados por el estudio de la superposición entre bandas diferentes. También nos hemos fijado en que ninguno de estos picos estuvieran saturados ni presentasen auto-absorción.

Siguiendo este criterio, las posiciones de los picos seleccionados para cada elemento enlazado al anión CO_3^{2-} son los mostrados en la tabla 5.1.

Ca/nm	442,575	445.526	443,546	458,6	612,4	616 ,4
Fe/cm	404.583	406.372	427.154	438.362	440,493	495,753
Mg/nm	383.272	382.973	518.427	517.345	516.804	

Tabla 5.1: *Picos LIBS seleccionados como variables de entrada de los diferentes métodos de análisis desarrollados para los diferentes compuestos.*

Como se puede comprobar, se han elegido 6 picos propios del calcio, 6 del hierro y 5 del magnesio.

1. **Calcio.** Los picos más intensos aparecen en torno al 612nm y 616nm. En la figura 5.1 vemos la representación solamente de la primera parte del espectro ya que lo que se quiere mostrar es el método seguido para elegir estos. Además, nuestro equipo de trabajo LIBS, tiene una baja percepción a las frecuencias ultravioletas presentes en la parte baja del espectro, aunque comparándolos con las bases de datos vemos como estos picos toman más importancia en la teoría que en los resultados obtenidos.
2. **Hierro.** En la figura 5.2 vemos la representación tanto de los espectros de carbonato de hierro sintético puro como natural (siderita), en el rango de posiciones seleccionadas

para crear nuestro modelo. Como se ve, ambos tipos de muestras coinciden bastante bien preservando la intensidad relativa entre todos estos picos.

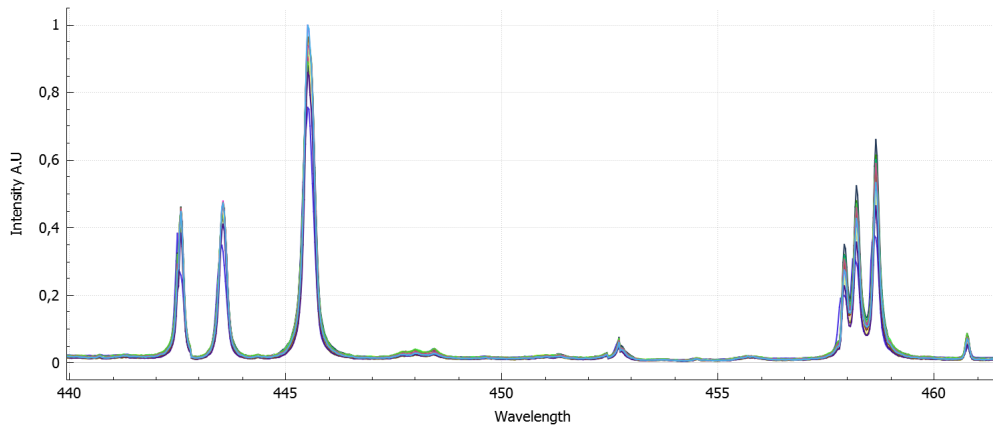


Figura 5.1: Representación de los picos LIBS elegidos para calcio representando el 100 % del carbonato de calcio sintético con la calcila natural.

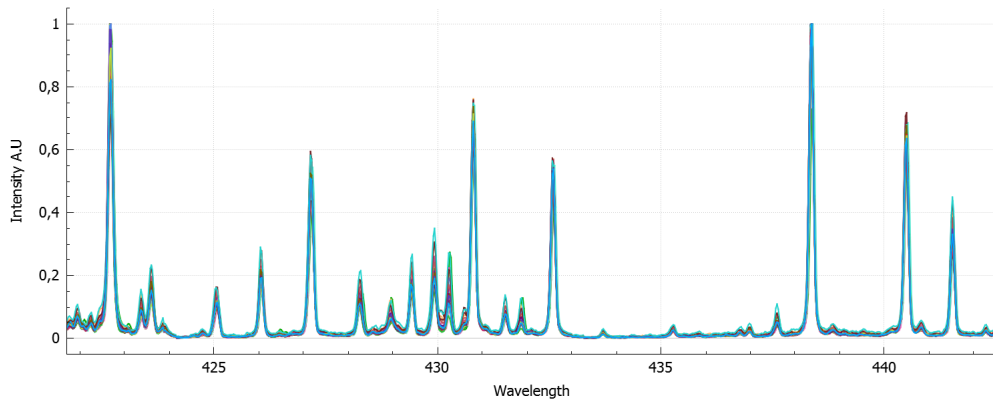


Figura 5.2: Espectros LIBS para la obtención de picos de hierro por medio de las muestras binarias con 100 % de FeCO₃ y siderita.

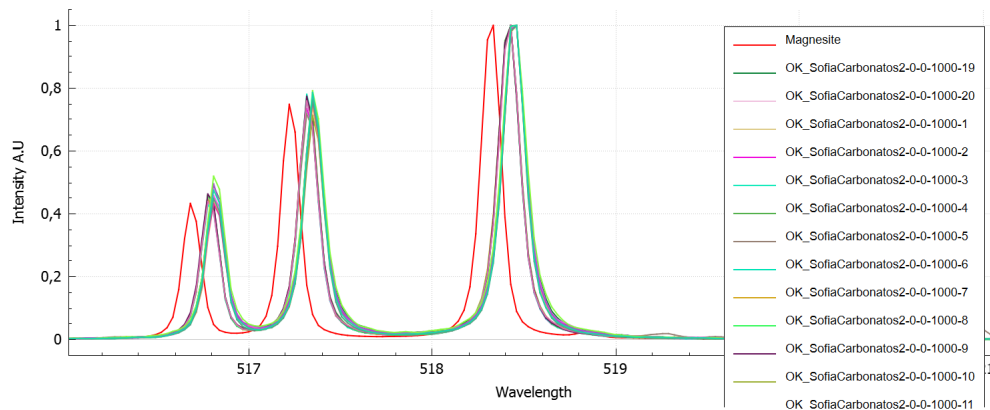


Figura 5.3: Representación de los picos LIBS elegidos para el magnesio comparando el hidrato carbonato de magnesio sintético puro con la magnesita natural.

3. **Magnesio.** Los picos más intensos en esta muestra se ven en las bandas en las posiciones de 516,8 cm, 517,3cm y 518,4 cm. Aunque estas tres estén bastante próximas, la identificación de todas ellas es claramente identificable como se comprueba en la figura 5.3. Cabe resaltar que en esta representación que el espectro de la magnesita natural (Espectro Rojo) se encuentra desplazado a la derecha, esta calibración se revisó y fue necesario adaptar los datos para hacerlos comparables.

5.3. Picos Raman.

La técnica Raman se usa con el objetivo de caracterizar la estructura del anión CO_3^{2-} . Las bandas que caracterizan este grupo se encuentran alrededor de los modos 1440, 1088, 715 y 278 cm^{-1} .

El pico en torno a $1088cm^{-1}$ es debido a las vibraciones de estiramiento simétrico, mientras que el antisimétrico se asocia a la banda que aparece en torno al $1400cm^{-1}$.

Por el contrario, los picos que se muestran en $715cm^{-1}$ son debidas a las flexiones sobre el plano que pueden estar distorsionadas, proporcionando más de un pico en esta zona. Lo importante en este estudio reside en la variación en la posición de estos modos normales de vibración en función del catión con el que este grupo funcional esté enlazado [31]. Por tanto, la caracterización de la calcita, magnesita, dolomita, y siderita van a residir en la variación de la posición de la banda principal y las secundarias.

De la misma forma a la expuesta para los espectros LIBS, podemos caracterizar los espectros de las muestras sintéticas Raman. Ahora la diferencia radica en el tipo de muestra que tenemos, ya que una mezcla binaria sintética de carbonato de calcio y carbonato de magnesio no va a corresponder a la estructura cristalográfica de una dolomita por lo que la banda que caracteriza a la dolomita no aparecerá en la muestra mezcla, ahora lo que tendremos serán bandas de ambos compuestos con intensidades proporcionales a las cantidades del mineral presente en cada muestra.

Este ejemplo se visualiza claramente en la figura 5.4, en la cual hemos tomado un espectro natural de dolomita y un espectro de una de las mezclas de 50% de carbonato de calcio y 50% de hidromagnesita. Como podemos comprobar, en la dolomita solo aparece un pico en la banda $1098,32cm^{-1}$, mientras que en las muestras binarias aparecen dos picos principales de menor intensidad. Uno de ellos se encuentra en torno a la posición del pico principal de la calcita natural pura ($1086,46 cm^{-1}$) y el segundo se encuentra desplazado hacia la derecha y se identifica con el pico principal para la hidromagnesita ($1118,1 cm^{-1}$).

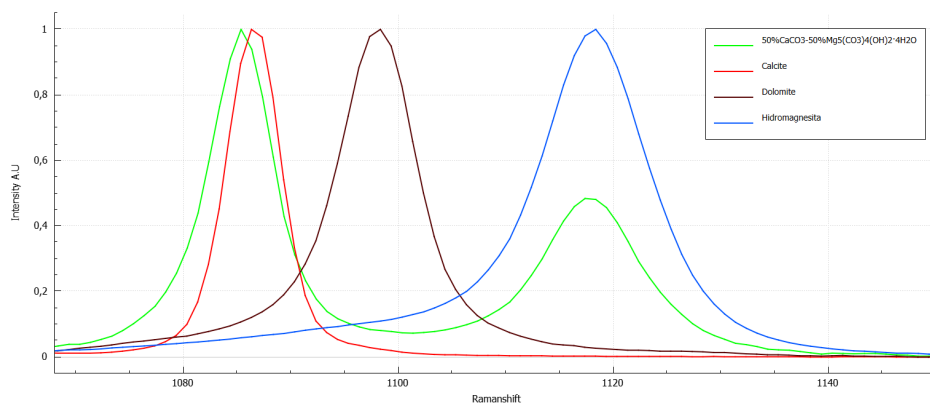


Figura 5.4: Comparación de los picos principales de la dolomita, calcita y magnesita con muestras sintéticas de carbonatos de calcio y magnesio en diferentes proporciones.

Para la identificación de los picos correspondientes con cada mineral, se han seleccionado los 20

espectros tomados para cada muestra sintética y se ha realizado una media de todos ellos para obtener un menor error. Se han representado los espectros media de las muestras sintéticas puras de CaCO_3 y $\text{Mg}_5(\text{CO}_3)_4(\text{OH})_2 \cdot 4\text{H}_2\text{O}$ y la que tiene una proporción media de ambos compuestos (50/50). Esta representación la podemos ver en la figura 5.5, donde se distinguen claramente los picos correspondientes con el carbonato de calcio en la posición 1085,51 cm^{-1} para la banda principal y 279,46 cm^{-1} para la secundaria. De igual forma la posición principal asociada a la hidromagnesita es la correspondiente a la posición 1118,32 cm^{-1} sin ningún pico secundario claramente identificable. Para corroborar toda esta información se usan bases de datos teóricos [32].

Vamos a notar que, aunque la muestra sintética corresponda a un valor del 50% - 50% los ratios de intensidades no siguen esta proporción, aunque al aumentar o disminuir el porcentaje de muestra en el material estas intensidades si varían en la misma proporción.

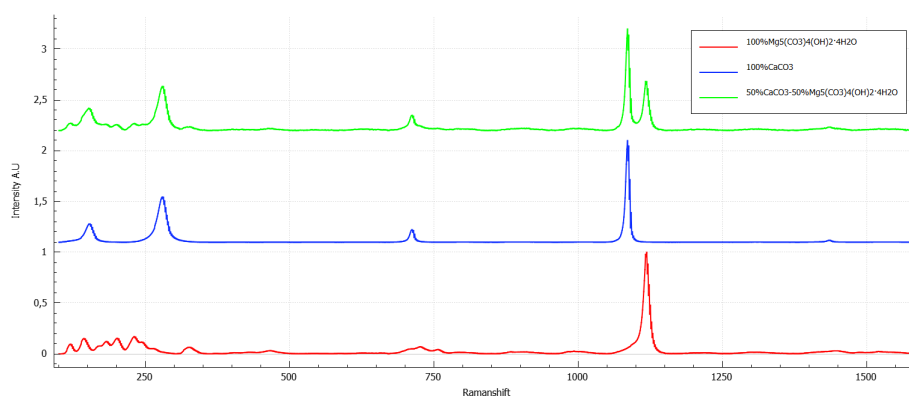


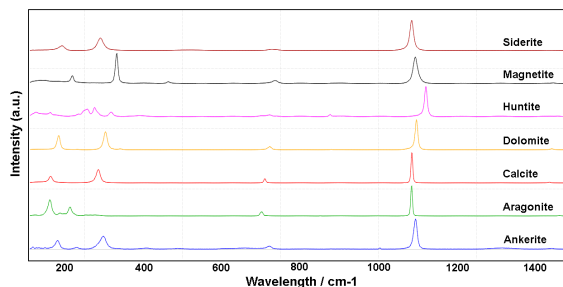
Figura 5.5: Espectros de muestras sintéticas de mezclas binarias de CaCO_3 y $\text{Mg}_5(\text{CO}_3)_4(\text{OH})_2 \cdot 4\text{H}_2\text{O}$.

Por tanto, los picos seleccionados para la identificación de las muestras sintéticas son:

Ca/RamanShift	279,46	1085,51
Mg/RamanShift		1118,32

Tabla 5.2: Picos Raman seleccionados como variables de entrada de los diferentes métodos de análisis desarrollados para los diferentes compuestos.

Para el caso de las muestras naturales, la posición de los picos principales que se usan para caracterizar cada componente son las que se muestran en la tabla de la figura 5.7, además en la figura 5.6 se comparan todos los espectros Raman de las diferentes fases minerales, mostrando como la banda principal del carbonato se ve desplazado en función del catión con el que se enlaza.



Calcita/RamanShift	164	286	710	1085
Aragonito/RamanShift	151	204	705	1084
Magnesita/RamanShift	218	351	736	1093
Siderita/RamanShift	192	290	729	1084
Dolomita/RamanShift	176	300	725	1097
Huntita/RamanShift	162	276	721	1120
Ankerita/RamanShift	180	297	721	1093

Figura 5.6: Espectros Raman de muestras naturales de los diferentes carbonatos

Figura 5.7: Bandas principales y secundarias de Raman para carbonatos naturales

Capítulo 6

Resultados obtenidos.

En este capítulo se presentan los diferentes resultados obtenidos para los tres tipos de mezclas binarias generadas a partir de las muestras sintéticas. Para cada par de muestras se exponen las curvas de calibración del modelo univariante de cada técnica individual y de la combinación de datos entre de ambas. Además de las rectas de las medidas de las concentraciones usando Redes Neuronales. También se muestra la clasificación de las muestras naturales hechas con PCA.

Para exponer de forma organizada estos resultados dividimos por el tipo de técnica usada y dentro de esta comparamos los diferentes modelos en función de los componentes usados para cada mezcla.

6.1. LIBS.

6.1.1. $CaCO_3 - Mg_5(CO_3)4(OH)_24H_2O$

Modelo Univariante

El modo gráfico de representar el modelo univariante resultante obtenido para las mezclas binarias de $CaCO_3 - Mg_5(CO_3)4(OH)_24H_2O$, se han realizado como una función de dispersión en la cual se representa la cantidad de uno de los compuestos en función de uno de los indicadores previamente calculados.

Combinado los seis picos característicos del Ca con los 5 del Mg, en total se han creado 60 indicadores, los cuales se obtienen de manera que para cada pico seleccionado de calcio forme un indicador con el resto de picos del magnesio. Se crea una matriz la cual posea la cantidad de compuesto de cada espectro y saque los diferentes indicadores dependiendo del pico que se use. De este modo tenemos todos los indicadores organizados para cada espectro.

La curva de calibración que mejor resultado han dado en el creado a partir del indicador formado a partir de la intensidad del pico situado en la posición $445,53 \text{ cm}^{-1}$ del Ca y el $382,97 \text{ cm}^{-1}$ para el Mg. Aunque estos picos son los que mejor resultado proporcionan, no son los de mayor intensidad. El modelo trabaja mejor con los picos que mejor relación de ratios entre su intensidad y concentración mantengan.

Los datos se han aproximado a una función cuadrática generándonos nuestra curva de calibrado la cual se observa en la figura 6.1. Además de la dispersión de los datos y la aproximación polinómica de segundo orden, se ha representado las fronteras de predicción (líneas rojas discontinuas) las cuales nos dan una idea de lo bien que funciona el modelo si metiéramos una muestra problema para identificar su concentración. En concreto, representan el intervalo de confianza del 95 %.

El coeficiente de determinación para esta recta es de $R^2=0.983$ valor que está relacionado con el grosor del intervalo de confianza mencionado, cuanto mayor coeficiente de correlación, menor

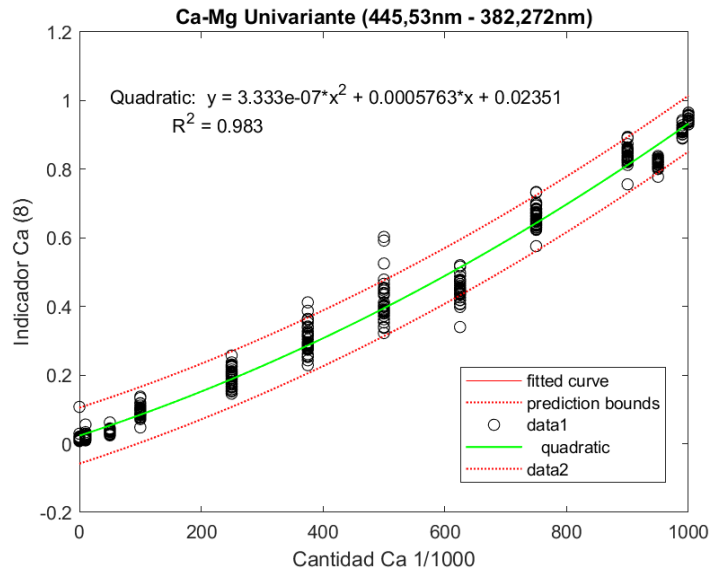


Figura 6.1: Modelo Univariante LIBS: Recta de calibrado $CaCO_3$ vs $Mg_5(CO_3)_4(OH)_2 \cdot 4H_2O$ haciendo uso de indicadores.

distancia entre los márgenes de error que representan las fronteras de predicción.

Fijándonos mejor en esta gráfica, se observa como los valores de concentraciones intermedias tienen una mayor dispersión que en los extremos ya que hay ciertos datos que se salen de los límites de confianza. Esto ocurre de manera más pronunciada en la pastilla con 500mg de $CaCO_3$ y 500 mg de $Mg_5(CO_3)_4(OH)_2 \cdot 4H_2O$ para la cual no nos sería posible una distinción entre muestras con un porcentaje del 50, 35 y 65 % de carbonato de calcio. A diferencia de lo que ocurre en el centro, los datos extremos de la curva se encuentran más agrupados entre ellos, sin datos que sobresalgan de las fronteras de predicción, lo que hace posible la identificación de Ca y Mg en concentraciones menores al 5 %.

Modelo Multivariante

La matriz de entrada de la red es una matriz de 400x11 en la que se representan las intensidades de los 2 picos propios del Ca y 5 del Mg para todos los espectros tomados de las mezclas binarias creadas de los compuestos de $CaCO_3$ y $Mg_5(CO_3)_4(OH)_2 \cdot 4H_2O$. Los target, es este caso, son las concentraciones de $CaCO_3$ para cada mezcla. Como salidas se obtienen las concentraciones directas de este mismo mineral del 15 % de los datos reservados para testear el modelo.

El resultado de este proceso se muestra en la figura 6.2, con un coeficiente de correlación de $R=0.99931$. Fijándonos directamente sobre la gráfica, en ella destaca la gran agrupación de los resultados en función de cada target. Es decir, cada target representa las proporciones teóricas de $CaCO_3$ con los que hemos creado las pastillas, las cuales están representadas en función de la concentración resultante (para las muestras especificadas de test) una vez pasadas por el modelo generado en el entrenamiento y validación.

La mayor dispersión se encuentra en los datos obtenidos para la pastilla con valores intermedios de concentración, pudiendo aparecer un solapamiento entre concentraciones en torno al 40 - 50 %. Los valores de los extremos presentan unos buenos resultados, haciendo posible la discriminación entre un 5 % y un 10 % de cantidad de calcio y magnesio.

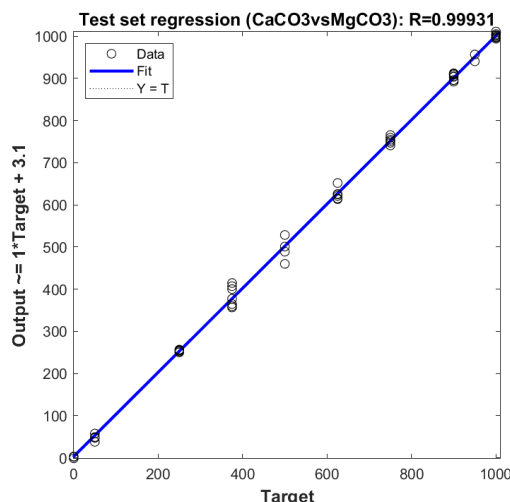


Figura 6.2: Predicción de las concentraciones para muestras test entre $CaCO_3$ vs $Mg_5(CO_3)_4(OH)_2 \cdot 4H_2O$ para LIBS usando ANN.

6.1.2. $CaCO_3$ - $FeCO_3$

Método Univariante

Para el caso de mezclas creadas a partir de los compuestos de $CaCO_3$ y $FeCO_3$ se han usado un total de 12 picos (6 propios del Ca y otros 6 del Fe) con los que se generan un total de 72 indicadores y por tanto 72 gráficas de dispersión, debido a la generación de cada indicador como combinación entre pico de los dos compuestos.

Una vez comparadas todas las gráficas de dispersión, el indicador que mejor resultados proporciona es el creado a partir de los picos que se encuentran en las posiciones 443.546 cm^{-1} para el $CaCO_3$ y el de la posición 495.753 cm^{-1} para el $FeCO_3$. La curva resultante se ve reflejada en la figura 6.4. En este caso la aproximación se ha realizado a una función cubica, la cual genera un coeficiente de $R^2 = 0,9806$.

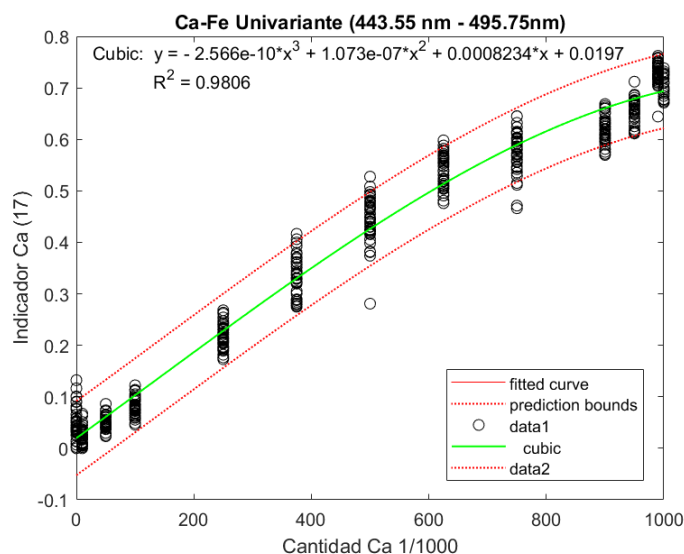


Figura 6.3: Recta de calibrado LIBS $CaCO_3$ vs $FeCO_3$, método univariante

A simple vista, datos generan una mayor dispersión que lo obtenido para la curva de calibración generada por Ca y Mg.

Fijándonos más minuciosamente en esta gráfica vemos como en los valores intermedios sigue existiendo una amplia separación de resultados. También en el límite de concentraciones vemos una dispersión bastante relevante a la hora de calcular concentraciones de otras muestras. Para prácticamente todo conjunto de datos, existen valores los cuales se salen del 95 % de la confianza del modelo, haciendo difícil poder diferenciar entre límites de concentraciones limítrofes. Aunque si es verdad que esta diferencia sigue siendo más significativa en valores intermedios.

Método multivariante

Usando redes neuronales el resultado que se obtiene para las concentraciones de los espectros de las mezclas binarias de carbonato de calcio y carbonato de hierro, se ve representado en la figura 6.4. Se ha usado como variables de entrada las intensidades de 12 picos propios de cada espectro. Para ello, previamente se genera una matriz de tamaño 500x12 con las intensidades de los 500 espectros que tenemos y los 12 picos seleccionados. Además a la red le indicamos las dos concentraciones correspondientes a los dos componentes, $CaCO_3$ y $FeCO_3$ y usa como target de la representación la cantidad de carbonato de calcio en las muestras.

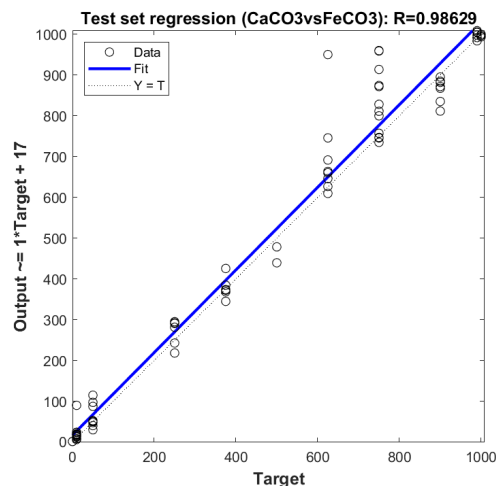


Figura 6.4: Predicción de las concentraciones para muestras test entre $CaCO_3$ vs $FeCO_3$ usando LIBS y empleando un método multivariante.

Interpretando la bondad de este nuevo modelo, se ve como las medidas de las concentraciones tienen valores más dispersos que en el modelo generado con ANN para la identificación de la concentración de Ca-Mg. Para valores altos del target no es posible distinguir entre concentraciones cercanas, teniendo incertidumbre en la concentración elemental de las muestras que presentan proporciones de Ca entre 60-80 %.

Valores de concentración con su target más bajo, presentan mejores resultados, haciendo posible una discriminación elemental más exacta para muestras con Ca bajas proporciones y Fe en mayor cantidad.

La alta dispersión de resultados para valores altos, nos ha generado que la recta resultado se desplace del valor teórico que se espera, este valor es el representado por una línea discontinua en la figura.

6.1.3. $FeCO_3$ - $Mg_5(CO_3)_4(OH)_2 \cdot 4H_2O$

Método Univariante

El último par de muestras sintéticas usadas para las mezclas binarias son carbonato de hierro e hidromagnesita.

El modelo se ha creado por medio de un método univariante, usa los indicadores formados por los 11 picos característicos del elemento de los dos compuestos a estudiar (Fe y Mg). Los picos se combinan entre ellos generando un total de 60 indicadores diferentes.

Después de representar todos estos indicadores en función de la cantidad de compuesto presente en la muestra, se obtiene que el indicador que mejor resultado proporciona es el creado a partir de la intensidad del pico en la posición $440,493 \text{ cm}^{-1}$ para el hierro y $382,973 \text{ cm}^{-1}$ para el magnesio y se puede ver en la figura 6.5.

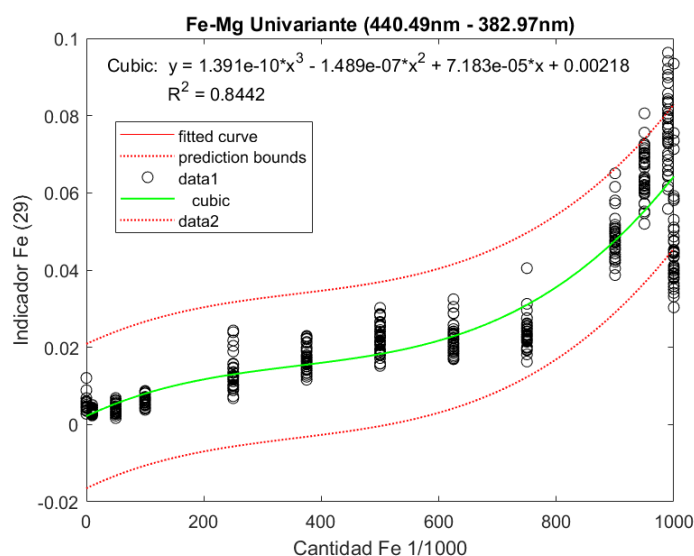


Figura 6.5: Recta de calibrado LIBS $FeCO_3$ vs $Mg_5(CO_3)_4(OH)_2 \cdot 4H_2O$, método univariante

En este caso, los datos se han aproximado a una función polinómica de grado tres (cúbica) con su coeficiente de determinación de $R^2 = 0,8442$.

Vemos como las fronteras de predicción se han aumentado la distancia entre ellas, esto es debido a que para altas cantidades de carbonato de hierro los datos sufren mucha dispersión. En este modelo es prácticamente imposible hacer distinciones de la concentración para muestras problema que tengan un alto contenido en hierro. Por otro lado, estos valores se mejoran para bajas concentraciones de hierro. Por lo general, al tener unas predicciones frontera más alejadas entre si se tiene una menor R^2 , y por tanto nos indica que nuestro modelo va a ser tan preciso para este par de muestras.

Método Multivariante

Para este par de compuestos, se usa como variables de entrada de la red las 11 intensidades características de cada elemento mayoritario presente en los 500 espectros y como target las 500 concentraciones teóricas de $FeCO_3$ para cada espectro. Entrenando la red neuronal, finalmente se obtiene la recta se muestra en la figura 6.6 cuyo coeficiente de correlación que aparece para esta es de $R=0.9924$.

Aquí la agrupación de todos los resultados es muy buena en todos los rangos de concentraciones. En este caso es posible hacer una discriminación de la concentración elemental presente en las muestras menores al 5%.

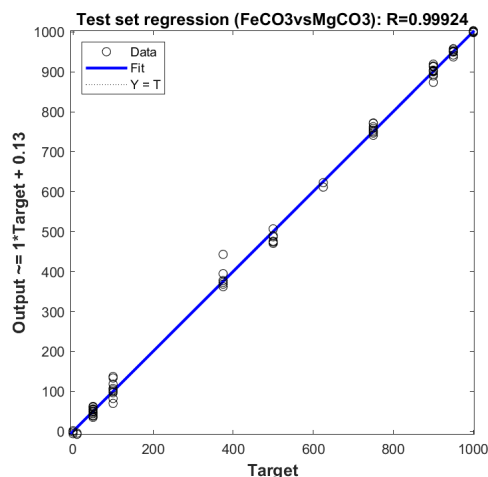


Figura 6.6: Predicción de las concentraciones para muestras test entre $FeCO_3$ vs $Mg_5(CO_3)_4(OH)_2 \cdot 4H_2O$ para LIBS usando ANN.

Por tanto, con un modelo basado en ANN es posible la identificación y clasificación elemental de muestras binarias de interés entre $FeCO_3$ y $MgCO_3$.

6.1.4. Muestras Naturales: PCA LIBS

Para continuar el análisis y clasificación de espectros atendiendo a tipo de compuesto que se trata, se propuso por el equipo de investigación un análisis de las componentes principales (PCA) del conjunto de muestras naturales seleccionadas, con el objetivo de conseguir el agrupamiento de muestras.

Por tanto, en esta sección nos centramos de forma cualitativa en la clasificación de las muestras naturales para espectros LIBS.

El método de PCA se ha llevado a cabo usando un proceso de Matlab utilizando como entrada los espectros LIBS enteros normalizados de todas las muestras y como salida las diferentes PCs. Los PCs se usan como vectores propios del espacio euclídeo, de estos PCs hemos seleccionado los tres primeros para poder realizar una representación ortogonal entre ellos. Esta representación la podemos observar en la figura 6.7, donde se han dividido por formas y colores los diferentes tipos de muestras naturales. Para el caso del LIBS, los tres primeros PCs nos explican el 88 % de la varianza de todos los espectros.

Con este tipo de representación se agrupan los diferentes espectros atendiendo a sus características. Además, se ha ordenado al programa, que separe por colores atendiendo al nombre de cada muestra. De este modo, se realiza una clasificación de los diferentes compuestos en función de su composición elemental.

Como resultado de la separación de fases minerales atendiendo a su composición únicamente elemental, se distinguen claramente tres zonas o grupos.

- Por un lado, representado por círculos amarillos, se encuentran todos los espectros de las diferentes sideritas que tenemos (3 tipos de sideritas en concreto). Estas están aisladas del resto de minerales sin ningún tipo de cruce con ninguna otra.
- Por otro lado tenemos los triángulos naranjas hacia arriba, los cuales están representando a las magnesitas. Al igual que pasaba con la siderita, estos espectros también han sido agru-

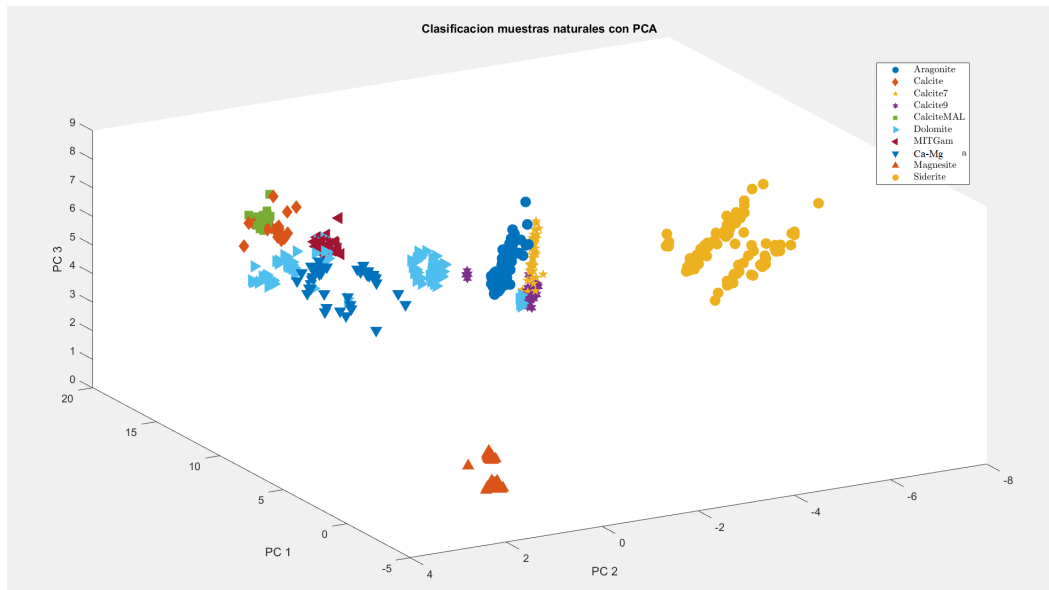


Figura 6.7: Gráfico 3D de la clasificación de las muestras naturales de carbonatos analizadas con LIBS en función de las tres primeras PCs.

pados y separados del resto de minerales, pero a diferencia de las anteriores, las diferentes magnesitas abarcan una zona menor del espacio estando estas más agrupadas.

- La tercera zona tiene un tamaño mayor y más difusa que las anteriores. Esta es la que abarca al resto de formas y colores, que representan las diferentes calcitas, aragonitos y dolomitas. En este caso no hemos agrupado las diferentes muestras de calcitas en un solo grupo (nombre), ya que al hacer esto nos separaban las calcitas en diferentes zonas dentro de esta más grande.

Como se puede comprobar hay una agrupación de muestras en el espacio pero equidistantes entre sí.

En la zona central del gráfico se agrupan todos los Aragonitos, las muestras de Calcita 7, Calcita 9 y una parte de los espectros de la Dolomita.

A la izquierda de estos podemos distinguir otra agrupación de CalcitaMal y Calcita.

Por último, en el espacio intermedio a los dos grupos anteriores se presentan el resto de espectros de Dolomita, separados entre ellos y dos secciones más agrupadas de muestras correspondientes a la muestra MITGamohaam (mezcla de calcita con cuarzo) y una mezcla binaria entre Carbonato de calcio y Magnesio, con la primera en mayor proporción (MITPongola).

Esto es debido a que LIBS no distingue entre grupos funcionales enlazados con el catión CO_3^{2-} , si no que solo nos da una determinación elemental del compuesto mayoritario de la muestra.

Los resultados obtenidos por medio únicamente de espectros LIBS nos generan buenos resultados si lo que se quiere es únicamente una clasificación atendiendo a la composición elemental ya que menos visto tres grandes grupos diferenciando entre la composición principal de cada muestra (Ca-Fe-Mg).

Ahora bien, el uso de PCA con LIBS no es la más idónea para la distinción de polimorfos ya que se ha visto que no es capaz de separar entre aragonito y calcita.

Además de esto, otra dificultad hallada ha sido la poca eficiencia para separar los grupos de compuestos mezclas, como la dolomita o la mezcla binaria de Ca-Mg. El resultado deseado sería una agrupación de todas estas muestras en una zona cercana a las calcitas y aragonitos pero sin mezclarse con ellas, cosa que no ocurre. Por otro lado, la separación en dos zonas diferentes de las diferentes muestras de calcita indican también poca eficiencia de este método en esta zona. Esto se podría ver mejorado con el uso de más componentes principales ya que las tres elegidas solamente nos dan información del 88 % de los espectros, motivo por el cual podrían no distinguirse bien entre dolomitas, calcitas, aragonitos y la separación entre diferentes calcitas.

6.2. Raman.

De manera paralela a lo expuesto para la técnica LIBS se han realizado los mismos métodos analíticos de datos para los espectros Raman.

En cuanto a crear un modelo basado únicamente en datos Raman se han usado solamente las muestras sintéticas de mezclas binarias formadas por $CaCO_3$ y $Mg_5(CO_3)_4(OH)_2 \cdot 4H_2O$.

6.2.1. $CaCO_3$ - $Mg_5(CO_3)_4(OH)_2 \cdot 4H_2O$

Modelo Univariante

El espectro Raman de carbonatos se caracteriza por diferentes bandas principales y secundarias como se indicó en la sección 5.3. Del mismo modo que en LIBS, se han concatenado todos estos picos de las diferentes muestras de manera que se genera una matriz en la cual aparecen los diferentes indicadores calculados para cada muestra y la cantidad de cada componente presente en ella. Como se han elegido dos bandas raman para el $CaCO_3$ y una que caracteriza al $Mg_5(CO_3)_4(OH)_2 \cdot 4H_2O$, se han obtenido un total de 4 indicadores.

Se representan todos estos en función de la cantidad de $CaCO_3$ de cada mezcla y se comparan entre ellos para encontrar el que mejor resultados proporcione.

En la figura 6.8 se ve la gráfica generada por el indicador creado a partir de los picos en la posición $458,6cm^{-1}$ para el $CaCO_3$ y el $382,973 cm^{-1}$ para la hidromagnesita.

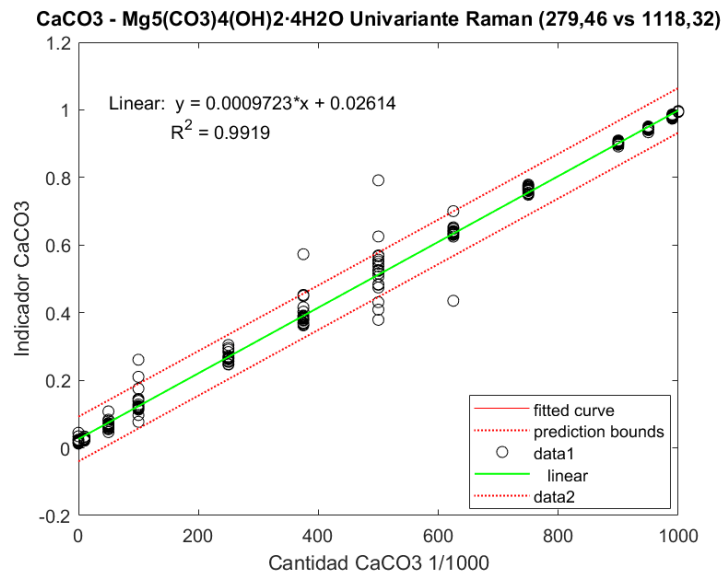


Figura 6.8: Recta de calibrado Raman univariante $CaCO_3$ vs $Mg_5(CO_3)_4(OH)_2 \cdot 4H_2O$.

La dispersión de los datos ahora sigue una tendencia lineal, con un coeficiente de $R^2=0.9919$.

A vista general, se ve como la dispersión entre los datos es más significativa en valores intermedios de la concentración y estos presentan mejores agrupaciones para valores más extremos. Para las pastilla con 50-40-60 % de $CaCO_3$ y 40-50-60 % de $Mg_5(CO_3)_4(OH)_2 \cdot 4H_2O$, aparecen datos fuera de los límites de predicción. En este modelo estas fronteras de confianza están relativamente próximas entre si, dando a entender una mejora en el funcionamiento de este modelo.

El estudio de muestras problemas nos va a proporcionar buenos resultados para valores extremos de concentraciones de $CaCO_3$, pudiendo diferenciar entre concentraciones próximas.

Por tanto, el uso de este modelo con datos Raman, nos permite distinguir entre proporciones de concentraciones pequeñas en el rango de bajas y altas concentraciones de carbonato de calcio pero dificultades para distinguir entre proporciones medias de los dos carbonatos a estudiar.

Modelo Multivariante

El resultado hallado con el uso de las redes neuronales se refleja en la figura 6.9. Para llevar a cabo este proceso se ha usado como entrada de la red una matriz con las intensidades seleccionadas para todos los espectro es decir, una matriz de tamaño 500x3 en la que se ha distinguido los datos que la red tiene que usar como entrenamiento, validación y los datos test que usan para generar el resultado.

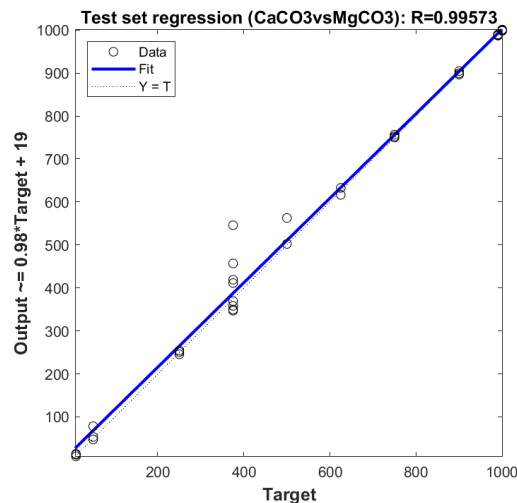


Figura 6.9: Predicción de las concentraciones para muestras test entre $CaCO_3$ vs $MgCo_3$ para Raman usando ANN

La salida que la red proporcionada se representan en función del target, en este caso se ha especificado la cantidad de carboanto de calcio de cada pastilla, de esta forma el coeficiente de correlación de este modelo toma el valor de $R=0.99573$.

De la representación obtenida, destaca la dispersión que aparece para un conjunto de datos perteneciente a un target en particular. Este es la pastilla de 400mg de $CaCO_3$. Las salidas para este conjunto de muestras presentan concentraciones muy dispares, no siendo posible hacer una distinción en la concentración en un rango del 10-20 %. A diferencia de lo que ocurre en este punto, el resto de datos si generan muy buenos resultados, ya que todos ellos están agrupados en concentraciones similares a las esperadas (targets).

6.2.2. Muestras Naturales: PCA Raman

En esta sección pasamos a presentar el análisis cualitativo por medio de componentes principales(PCA) realizado para la clasificación de los espectros Raman de las muestras naturales.

La representación ortogonal de las tres primeras PCs, las cuales dan información del 75 % de los espectros, se puede apreciar en la figura 6.10. Se ha obtenido una clasificación de los espectros separándolos según el tipo de muestra.

El uso de los espectros Raman para la clasificación y agrupamiento del mismo conjunto de muestras naturales usada en LIBS hace posible la discriminación entre grupos funcionales. Como resultado, se ha logrado una separación bastante efectiva de todos los tipos de muestras, pudiendo hacer ahora una distinción entre calcita y aragonito. Además, todas las muestras de calcita se han renombrado como 'calcita' ya que todas se agrupan tendiendo a una preferencia lineal, cosa que en LIBS no sucedía. Siguiendo esta misma tendencia lineal encontramos las muestras de magnesita, aragonito y MITGamohaam (mezcla de calcita con cuarzo). Todas ellas están separadas y son claramente identificables en el plano tridimensional. Además para la dolomita y MITPongola (que corresponde a la muestra de mezcla binaria entre carbonato de calcio y de magnesio) se encuentran aisladas y agrupadas de manera que hace fácil su identificación.

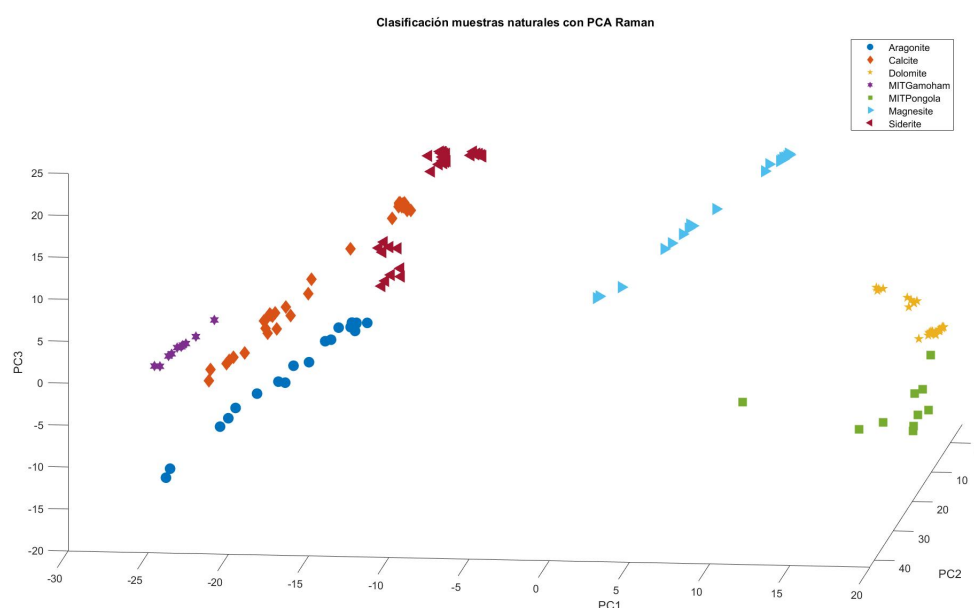


Figura 6.10: Representación euclídea tridimensional de las muestras naturales de carbonatos en función de las tres primeras PCs para la técnica Raman.

En vistas generales, el uso de datos Raman en PCA proporcionan resultados más eficientes, incluso con un porcentaje de información de la varianza menor que para LIBS al usar solo las tres primeras PCs. Por tanto, la clasificación de espectros Raman atendiendo al análisis de sus componentes principales, generan una mejor discriminación entre Ca-Fe-Mg en carbonatos.

6.3. Fusión de datos LIBS-Raman.

Las técnicas LIBS y Raman son herramientas con una gran potencia para la identificación elemental y molecular usados para la investigación de cuerpos planetarios como Marte.

La combinación de ambas técnicas juega un papel fundamental en la caracterización geoquímica, gracias a su carácter complementario. Este hecho ha sido estudiado para otro tipo de compuestos también relevantes en investigación planetaria [33].

El estudio de la caracterización de carbonatos separando técnicas LIBS y Raman no es suficiente para una completa discriminación de fases minerales.

En esta nueva sección nos centramos en exponer los resultados de los diferentes métodos analíticos utilizados haciendo una combinación entre datos LIBS y Raman, tanto de manera semi-cuantitativa para las muestras sintéticas, como de forma más cualitativa usando PCA para las muestras naturales.

6.3.1. Modelo Univariante

El modelo univariante desarrollado en la combinación de técnicas se basa en el mismo principio que en los casos monoanalíticos. Esto es, una representación de la dispersión de los datos generados (indicadores) frente a la cantidad de carbonato presente en cada mezcla.

La diferencia ahora radica en el cálculo de los diferentes indicadores. En concreto, los indicadores combinación se logran concatenando los diferentes indicadores seleccionados para cada técnica individual pertenecientes a las muestras con las mismas concentraciones. Así, para cada indicador LIBS de una concentración, se realiza la suma algebraica y se divide entre dos con todos los indicadores Raman de la pastilla con esa misma concentración. De este modo se consiguen un total de 800 indicadores para una misma mezcla, que sale como el producto de los 40 indicadores de una mezcla LIBS con los 20 indicadores de una mezcla Raman y 400 indicadores fusión en el caso de las pastillas LIBS de las cuales solo tomamos 20 puntos por muestra.

Hay que notar que para llevar a cabo una correcta y más adecuada combinación de indicadores, habría que realizar la suma usando los espectros LIBS y Raman tomados en el mismo punto de la muestra. Por este motivo se amplía este método de combinación, porque aún no apuntando al mismo punto, reduciremos el error y así podremos comparar los resultados de la combinación con los obtenidos individualmente en cada caso.

La representación de los indicadores en función de la cantidad de $CaCO_3$ de cada muestra genera la gráfica de dispersión que se observa en la figura 6.11, en ella podemos ver como la tendencia de los datos es lineal y el coeficiente de determinación $R^2=0.9929$.

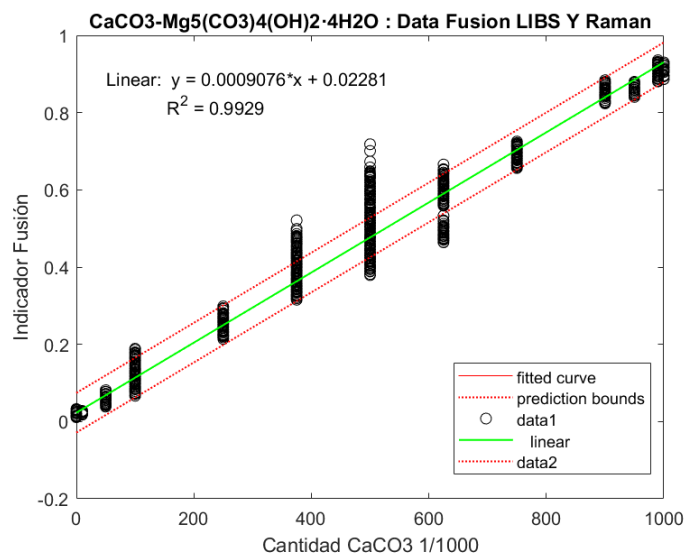


Figura 6.11: Recta de calibrado resultante de la fusión de datos LIBS y Raman

Los resultados combinación generan dispersión de resultados en las zonas intermedias de concentración, como veníamos acostumbrados a ver para cada técnica por separado. Esta dispersión mejora en los valores de los extremos.

Comparando esta gráfica con las resultantes para LIBS y Raman monoanalíticas vemos como esa complementariedad de técnicas si que optimiza la clasificación del modelo aunque no de forma drástica. Se aprecia una gran mejora en las predicciones frontera, ya que estas son más estrechas en el caso de la combinación y por tanto la bondad del modelo será mejor que en los modelos univariantes de LIBS y Raman.

6.3.2. Modelo Multivariante

El nuevo modelo combinación usando el método multivariante con ANN, combina los datos LIBS y Raman generando, previamente, una matriz la cual contenga las intensidades de las bandas principales y secundarias de Raman correspondientes al $CaCO_3$ y al $Mg_5(CO_3)_4(OH)_2 \cdot 4H_2O$ y las que corresponden a los picos LIBS de los elementos Ca y Mg, es decir, concatenando las dos matrices creadas como entradas de las ANN para cada técnica por separado, de forma que para cada punto de cada mezcla, se generan una concatenación entre todos los picos LIBS de ese punto con todos los Raman de todos los puntos de una misma mezcla. Así ampliamos la cantidad de variables de entrada de la red.

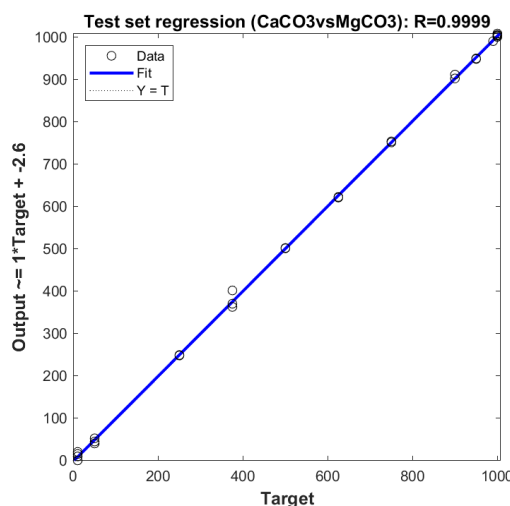


Figura 6.12: Predicción de las concentraciones para muestras test resultante de la fusión de datos de las técnicas LIBS y Raman haciendo uso de ANN

La salida que nos ha generado esta nueva red neuronal se refleja en la figura 6.12. A simple vista se puede apreciar la gran mejora que se obtiene en comparación con los resultados al aplicar las técnicas de manera individual. Se observa como el coeficiente de correlación ha aumentado al $R = 0,9999$ incremento significativo con respecto a las anteriores. Además no se aprecia apenas una dispersión de los resultados en ninguna de las muestras usadas para testear el modelo, la única apreciación que se puede comentar en el target de 400mg de $CaCO_3$, en la cual uno de los valores presenta un resultado de la concentración algo superior al resto de puntos de ese target, incrementando así la dispersión en esa posición. Esto se puede deber al peso que aún tienen el Raman, que como veíamos en el modelo multivariante Raman para este par de compuestos, este mismo punto presentaba variaciones de concentración.

Por lo general, para la combinación de técnicas usando análisis multivariantes, podemos hacer una precisa determinación de las cantidades de los compuestos presentes en una muestra problema para mezclas entre carbonato de calcio e hidromagnesita. Por tanto, ahora si se ve una clara optimización para la clasificación de carbonatos mejorando la discriminación que realizaba

las técnicas de manera individual.

Esta optimización de los resultados demuestra la complementariedad existente entre técnicas espectroscópicas elementales y moleculares a la hora de discriminar fases minerales.

6.3.3. Muestras Naturales: Combinación de datos con PCA

La combinación de datos en PCA se hace mediante la concatenación de las intensidades de todo el espectro de LIBS y Raman normalizadas para cada muestra guardando estas intensidades en una nueva matriz.

En LIBS se tomaron 40 espectros por muestra natural y en el caso de Raman fueron 10. Para que los pesos del conjunto de ambas técnicas fueran las mismas en el nuevo modelo de clasificación, lo que se ha hecho es una media de los espectros LIBS, de modo que se van cogiendo grupos de cuatro en cuatro espectros y calculando el espectro media entre ellos. De esta forma se reducen los datos LIBS a 10 espectros por muestra.

El conjunto de datos combinados con todas las intensidades se filtran haciendo uso de la técnica PCA, de las cuales se obtiene que las 10 primeras PCs generan el 99 % de la varianza.

La combinación de datos genera una nueva clasificación de las muestras naturales las cuales se observan en la figura 6.13. En este caso se ha realizado una representación bidimensional de las cuatro primeras PCs.

Analizando cada una de ellas vemos como para la figura 6.13 (A) la representación de la PC1 en función de la PC2 genera una buena agrupación de las diferentes fases minerales excepto para el aragonito, la calcita y la muestra MITGamohaam. Estas tres se agrupan entre ellas presentado una tendencia más lineal. Este resultado se mejora empleando la PC1 en función de la PC3 (6.13 (B)), en la cual el grupo de aragonito y calcita se ven claramente separados siguiendo del mismo modo con la tendencia lineal.

Este resultado nos indica que la PC3 tiene una mayor información de Raman que la PC2 ya que es capaz de distinguir entre diferentes fases minerales de un mismo compuesto.

Esto mismo ocurre si examinamos la representación de la PC1 en función de la PC4 (6.13 (C)). En ella existe una muy buena discriminación de los carbonatos naturales, agrupándose en diferentes zonas del espacio bidimensional de estas dos PCs y obteniendo una buena clasificación y consiguiendo resultados sutilmente mejores a los de la representación (B).

Terminado de analizar el resto de representaciones generadas, se observa como en todas ellas la separación entre aragonito y calcita tiene resultados óptimos en todas ellas. En la figura 6.13 (D), (E) y (F), la siderita no se encuentra agrupada en una sola zona si no que se encuentra dividida en dos, aunque estas siguen siendo diferenciables del resto. Además, para la representación (E), la distinción de la dolomita y parte de las muestras de calcita no es muy precisa.

El análisis de estas representaciones demuestra que en dos primeras PCs el LIBS tiene un mayor peso ya que no le es posible distinguir entre fases cristalinas de Calcita y Aragonito. Por consiguiente, la PC3 y PC4 tienen un mayor peso de Raman.

La representación de dos PCs cuyo peso mayoritario corresponda a una misma técnica, va a generar unos resultados más parecidos a los encontrados en el estudio de las PCs de las técnicas individuales. La representación de PCs, las cuales tengan mayor peso de técnicas diferentes, generan una discriminación más efectiva, realizando una caracterización mineralógica y geoquímica de muestras de carbonatos mucho más eficiente, incluso haciendo uso únicamente de las cuatro primeras PCs que generan el 84 % de la varianza en los datos.

La complementariedad de técnicas optimiza la caracterización de minerales carbonatos, mejorando con ello la discriminación en las técnicas individuales, demostrando con ello la complementariedad de técnicas espectroscópicas elementales y moleculares.

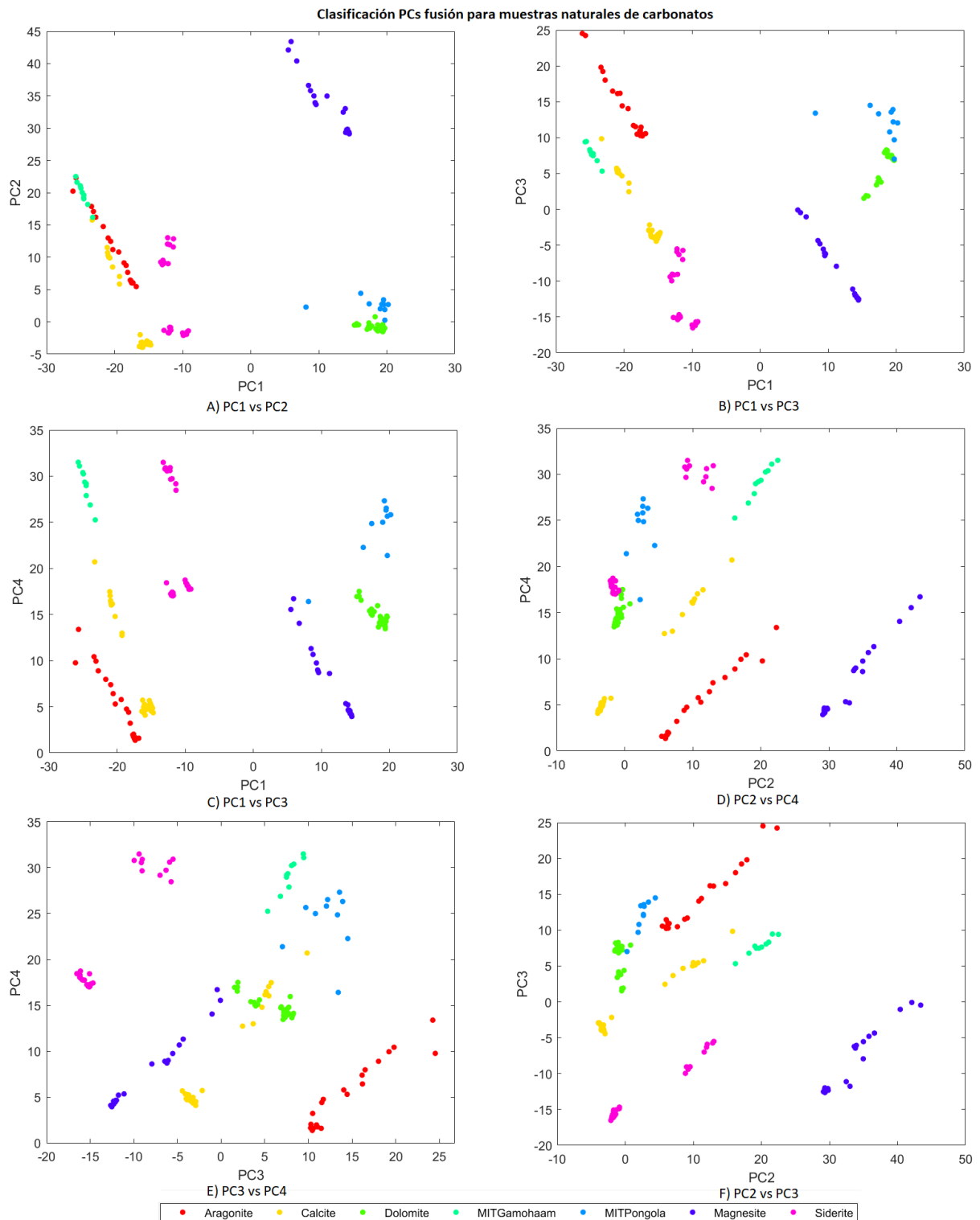


Figura 6.13: Representación bidimensional de las cuatro primeras PCs para clasificar muestras naturales de carbonatos atendiendo a la combinación del conjunto de datos obtenidos por LIBS y Raman

Capítulo 7

Conclusiones.

De manera general se ha visto el gran poder analítico que presentan tanto la técnica LIBS como Raman, entendiendo con ello el gran interés que se tiene en el uso de técnicas espectroscópicas en misiones de exploración planetarias. Nuestro estudio demuestra esta capacidad de análisis del cual podemos sacar ciertas conclusiones.

- El uso de un método univariante como base para la calibración de un modelo el cual nos permita, en un futuro, la determinación de la concentración de muestras problemas, proporcionan resultados adecuados en los rangos extremales para las mezclas sintéticas creadas a partir de los compuestos $CaCO_3$ y $Mg_5(CO_3)_4(OH)_2 \cdot 4H_2O$. Pero para los otros dos conjuntos de muestras estos resultados han sido más dispersos, con los cuales no se podría hacer una buena discriminación de la concentración elemental de los compuestos Fe-Mg. Además para todas las mezclas, los resultados intermedios han presentado muchas dispersiones.

Por tanto, el modelo creado para $CaCO_3 - Mg_5(CO_3)_4(OH)_2 \cdot 4H_2O$ es el que mejor resultado ofrece de los tres conjuntos. Esto lo podemos ver reflejado interpretando la confianza que representan las predicciones frontera, las cuales son menos distantes en el modelo de identificación elemental del Ca-Mg. Este acercamiento entre los límites de las fronteras viene relacionado con la menor R^2 .

- Por lo general, se ha visto como el uso de carbonato de hierro en muestras sintéticas genera en el modelo resultados más dispersos, tanto para modelos univariantes como multivariantes.
- Los resultados de la calibración del modelo univariante creado para Raman de las mezclas con $CaCO_3$ y $Mg_5(CO_3)_4(OH)_2 \cdot 4H_2O$, presentan aún mejores resultados que el creado para LIBS, aunque sigue habiendo dispersiones intermedias lo que daría resultados errores en la concentración en este rango.
- Los modelos basados en Redes Neuronales Artificiales tanto en LIBS como en Raman, han creado modelos de medidas directas de las concentraciones cuyos resultados son muy aproximados a lo que se esperaba teóricamente.
- Gracias a la información complementaria proporcionada por los datos Raman, y LIBS, las investigaciones de laboratorio descritas en este trabajo demuestran como efectivamente la combinación de datos ofrece claras ventajas para la elaboración de un modelo capaz de diferenciar mismas fases minerales con diferente composición estructural. Por tanto, se muestra como la semicuantificación de mezclas se mejora con el uso de fusión de técnicas, tanto en el uso de un método univariante como multivariante.

- Por otro lado, examinando los resultados obtenidos de la caracterización mineralógica y geoquímica de nuestros carbonatos naturales (PCA), podemos concluir que existen ciertas limitaciones usando las técnicas individuales.

La discriminación de fases minerales no se pueden lograr con LIBS, ya que estas solo difieren de su estructura mineral y no elemental, no pudiendo hacer distinción entre materiales como calcita o aragonito, cosa que si se consigue con los datos Raman.

- La combinación de los datos ha ayudado a la discriminación de carbonatos naturales con PCA. Se logra una excelente separación de los grupos minerales haciendo uso solamente de las cuatro primeras PCs. Además de poder identificar el peso que adquiere cada técnica en cada PC.
- El uso de PCA como técnica de clasificación para las muestras naturales nos proporciona un visión general de como esta técnica funciona y, como línea futura, la combinación de PCA con otros clasificadores como ANN o en otras configuraciones como por ejemplo Linear Discriminant Analysis, podría permitir realizar clasificaciones / identificaciones más cuantitativas y precisas.
- Por tanto, la combinación de datos mejora la calidad espectral corriendo los errores asociados a los sistemas analíticos empleados y eliminando las posibles variaciones que no están asociadas a las fases minerales investigadas.

La complementariedad entre diferentes técnicas espectroscópicas fomenta marcos como Super-Cam, con la combinación de diferentes instrumentos con la posibilidad de medir en el mismo punto de interés.

Se ha observado que si bien las técnicas de análisis automatizadas facilitan el procesamiento de múltiples datos, no son inmunes a los efectos de datos erróneos o excepcionales. Por ello, es conveniente, al menos en estadios tempranos del desarrollo de modelos de identificación y clasificación, que el investigador tenga un conocimiento profundo de la naturaleza de los datos. En este sentido, sería útil estudiar en detalle especialmente aquellos espectros que proporcionan resultados o dispersiones fuera de la tónica general.

Impacto del TFG

Los resultados que se han obtenido en este trabajo se han presentado en diversas conferencias, por un lado la conferencia Lunar and Planetary Science Conference 2022, celebrada en Marzo de este año, se presento un abstract bajo el título *Semi-Quantification And Identification Of Carbonate Mixtures Using A Combination Of Libs And Raman Techniques*. También, en la conferencia Georaman 2022 se presento un abstract y un poster con el título *Carbonates discrimination in planetary missions: combination of Raman-LIBS data by univariate models and Artificial Neural Networks*.

Además, parte de lo resultados se emplearon en la elaboración de un artículo próximo a ser enviado y aprobado por el grupo de ciencia de Mars2020. Por otro lado, actualmente se está preparando otro artículo con la autora de este TFG como candidata a primera autora para el *Journal of Raman Spectroscopy*.

Bibliografía

- [1] FARLEY, K.A., WILLIFORD, K.H., STACK, K.M. ET AL. *Mars 2020 Mission Overview*, Space Sci Rev 216, 142 (2020). <https://doi.org/10.1007/s11214-020-00762-y>
- [2] *Concepts for Mars Sample Return*, NASA Science, MARS exploration programe, 2020
- [3] MAURICE, S. ET AL. *The ChemCam instrument suite on the Mars Science Laboratory (MSL) rover: Science objectives and mast unit description.*, Space Sci. Rev. 170, 95–166 (2012).
- [4] FARLEY, K.A., WILLIFORD, K.H., STACK, K.M. ET AL. *Instrument Payload: SHERLOC*, Mars 2020 Mission Overview, Space Sci Rev 216, 142 (2020). <https://doi.org/10.1007/s11214-020-00762-y>
- [5] ALLWOOD, A.C., WADE, L.A., FOOTE, M.C. ET AL. *PIXL: Planetary Instrument for X-Ray Lithochemistry.*,Space Sci Rev 216, 134 (2020). <https://doi.org/10.1007/s11214-020-00767-7>
- [6] WIENS, R.C., MAURICE, S., ROBINSON, S.H. ET AL. *The SuperCam Instrument Suite on the NASA Mars 2020 Rover: Body Unit and Combined System Tests.*, Space Sci Rev 217, 4 (2021). <https://doi.org/10.1007/s11214-020-00777-5>
- [7] HAMRAN, SE., PAIGE, D.A., AMUNDSEN, H.E.F. ET AL. *Radar Imager for Mars' Subsurface Experiment—RIMFAX.* ,Space Sci Rev 216, 128 (2020). <https://doi.org/10.1007/s11214-020-00740-4>
- [8] HECHT, M., HOFFMAN, J., RAPP, D. ET AL. *Mars Oxygen ISRU Experiment (MOXIE).*, Space Sci Rev 217, 9 (2021). <https://doi.org/10.1007/s11214-020-00782-8>
- [9] RODRIGUEZ-MANFREDI, J.A., DE LA TORRE JUÁREZ, M., ALONSO, A. ET AL. *The Mars Environmental Dynamics Analyzer, MEDA. A Suite of Environmental Sensors for the Mars 2020 Mission.*, Space Sci Rev 217, 48 (2021). <https://doi.org/10.1007/s11214-021-00816-9>
- [10] KINCH, K.M., MADSEN, M.B., BELL, J.F. ET AL. *Radiometric Calibration Targets for the Mastcam-Z Camera on the Mars 2020 Rover Mission.*, Space Sci Rev 216, 141 (2020). <https://doi.org/10.1007/s11214-020-00774-8>
- [11] WIENS, R. C., MAURICE, S. PEREZ, F. R. *The SuperCam remote sensing instrument suite for the mars 2020 rover: A preview. Spectrosc.*, (Santa Monica) 32, 50–55 (2017).
- [12] MAURICE, S., WIENS, R.C., BERNARDI, P. ET AL. *The SuperCam Instrument Suite on the Mars 2020 Rover: Science Objectives and Mast-Unit Description*, Space Sci Rev 217, 47 (2021). <https://doi.org/10.1007/s11214-021-00807-w>
- [13] MANRIQUE, J.A., LOPEZ-REYES, G., COUSIN, A. ET AL. *SuperCam Calibration Targets: Design and Development*, Space Sci Rev 216, 138 (2020). <https://doi.org/10.1007/s11214-020-00764-w>

- [14] BRIONY H.N. HORGAN, RYAN B. ANDERSON, GILLES DROMART, ELENA S. AMADOR, MELISSA S. RICE. *The mineral diversity of Jezero crater: Evidence for possible lacustrine carbonates on Mars*, *Icarus*, Volume 339 (2020), <https://doi.org/10.1016/j.icarus.2019.113526>.
- [15] ZASTROW, A. M. *Distinct Carbonate Lithologies in Jezero Crater , Mars*, *Geophys. Res. Lett.* 48, e2020GL09236 (2021).
- [16] TIMOTHY A. GOUDGE, JAMES W. HEAD, JOHN F. MUSTARD, CALEB I. FASSETT. *An analysis of open-basin lake deposits on Mars: Evidence for the nature of associated lacustrine deposits and post-lacustrine modification processes*, *Icarus*, Volume 219, Issue 1, 2012, <https://doi.org/10.1016/j.icarus.2012.02.027>.
- [17] GRICE, K., HOLMAN, A. I., PLET, C. TRIPP, M. *Fossilised biomolecules and biomarkers in carbonate concretions from konservat-lagerstätten.*, *Minerals* 9, 1–15 (2019).
- [18] E. CLAVÉ, K. BENZERARA, P. BECK³, *CARBONATE DETECTION WITH SUPERCAM IN THE JEZERO CRATER, MARS.*, 53rd Lunar and Planetary Science Conference (2022).
- [19] F. ANABITARTE, A. COBO, AND J. M. LOPEZ-HIGUERA. *Laser-Induced Breakdown Spectroscopy: Fundamentals, Applications, and Challenges*, *ISRN Spectroscopy*, 2012 1-12. <https://doi.org/10.5402/2012/285240>.
- [20] DAVID A. CREMERS, LEON J. RADZIEMSKI. *Handbook of Laser-Induced Breakdown Spectroscopy*, 2006
- [21] *Tema 2: Espectroscopía vibracional.* <https://edejesus.web.uah.es/resumenes/DECI/tema2.pdf>
- [22] ROBIN R. JONES, DAVID C. HOOPER, LIWU ZHANG, DANIEL WOLVERSON AND VENTSLAV K. VALEV , *Raman Techniques: Fundamentals and Frontiers*, Jones et al. *Nanoscale Research Letters* (2019), <https://doi.org/10.1186/s11671-019-3039-2>
- [23] VANESSA GUTIERREZ CANO, JAVIER OTERO MÁRQUEZ, *Espectroscopía Raman: Fundamento y aplicaciones.*
- [24] ING. ANALÍA SICARDI SEGADÉ y DR. MARCO ANTONIO MENESES NAVA. *Supresión de fluorescencia en espectroscopía Raman de piezas arqueológicas mediante SSRS y SERDS* , Centro de Investigaciones en Óptica (CIO) A.C.
- [25] RICHARD L. MCCREERY. *SIGNAL-TO-NOISE IN RAMAN SPECTROSCOPY*, *Raman Spectroscopy for Chemical Analysis*, 2000, <https://doi.org/10.1002/0471721646.ch4>
- [26] R. PÉREZ PUEYO. *Capítulo 1: FUNDAMENTOS DE ESPECTROSCOPIA RAMAN*, <https://www.tdx.cat/bitstream/handle/10803/6887/03Rpp03de11.pdf>
- [27] MARCO VENERANDA. *Analytical database of Martian minerals (ADaMM): Project synopsis and Raman data overview*, 2017. <https://doi.org/10.1002/jrs.6215>
- [28] BRO, R. SMILDE, A. K., *Principal component analysis*, *Anal. Methods* 6, 2812–2831 (2014).
- [29] AURELIO SANZ-ARRANZ,^{A,B} JOSE A. MANRIQUE-MARTINEZ,^{A,B*} JESUS MEDINA-GARCIA^{A,B} AND FERNANDO RULL-PEREZA. *Amorphous zinc borate as a simple standard for baseline correction in Raman spectra*, 2017.
- [30] *NIST LIBS database*, <https://physics.nist.gov/PhysRefData/ASD/LIBS/lib-form.html>

- [31] JUNMIN SUN, ZEGUANG WU, HONGFEI CHENG, ZHANJUN ZHANG, RAY L. FROST. *A Raman spectroscopic comparison of calcite and dolomite*, Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, Volume 117, 2014, Pages 158-162, ISSN 1386-1425, <https://doi.org/10.1016/j.saa.2013.08.014>.
- [32] *RRUFF database*, <https://rruff.info/>
- [33] P. SOBRON y UNIDAD ASOCIADA UNIVERSIDAD DE VALLADOLID-CENTRO DE ASTROBIOLOGÍA CSIC-INTA., *RAMAN AND LIBS: A DEFINITIVE COMBINATION FOR THE CHARACTERIZATION OF NATURAL SAMPLES FROM THE RIO TINTO MARS ANALOG.*, 40th Lunar and Planetary Science Conference (2009).