



---

**Universidad de Valladolid**

**FACULTAD DE CIENCIAS**

TRABAJO FIN DE GRADO

Grado en Física

**EVALUACIÓN Y COMBINACIÓN DE DATOS ESPECTROSCÓPICOS DE LOS  
GRANATES COMO ANÁLOGOS MINERALES PARA EL ESTUDIO DE MARTE**

Autora: Pilar Renedo Sánchez-Girón

Tutores: Guillermo López Reyes

Sofía Julve González

2024

# Índice

<b>Abstract</b>	<b>1</b>
<b>1. Introducción</b>	<b>1</b>
1.1. La exploración de Marte . . . . .	1
1.2. Espectroscopia Raman y FTIR. Ventajas y limitaciones . . . . .	2
1.3. Algoritmos de clasificación . . . . .	4
<b>2. Objetivo</b>	<b>5</b>
<b>3. Materiales, instrumentación y análisis de datos</b>	<b>5</b>
3.1. Muestras utilizadas . . . . .	5
3.2. Medidas Raman y FTIR . . . . .	6
3.3. Metodología . . . . .	7
3.3.1. Tratamiento espectral . . . . .	7
3.3.2. Preparación de datos . . . . .	8
3.3.3. Métodos de clasificación usados . . . . .	8
3.3.4. Filosofía de los modelos . . . . .	9
<b>4. Resultados y discusión</b>	<b>10</b>
4.1. Caracterización de muestras . . . . .	10
4.2. Análisis cualitativos Raman: correlación entre parámetros de bandas y cationes . . . . .	15
4.3. Raman . . . . .	17
4.3.1. Parámetros de bandas . . . . .	17
4.3.2. PCA . . . . .	22
4.4. RIR (fusión datos Raman+ FTIR) . . . . .	27
4.4.1. PCA . . . . .	27
4.5. Resumen de resultados . . . . .	31
<b>5. Conclusiones</b>	<b>31</b>
<b>6. Impacto científico</b>	<b>33</b>
<b>Referencias</b>	<b>34</b>
<b>Anexo</b>	<b>36</b>
A1. Script de MATLAB para PCA . . . . .	36
A2. Tabla de datos de parámetros de bandas para entrenamiento . . . . .	39
A3. Tabla de datos de parámetros de bandas para test . . . . .	44

## Abstract

Mars exploration is gaining more relevance as new advances in technology have allowed us to study its composition through rovers capable of landing on its surface. Due to the large amount of mineralogical data they collect, it is necessary to implement machine learning techniques to classify these samples. Thanks to the SuperCam tool of the Perseverance rover, it is possible to perform Raman and FTIR spectrometries in situ on the red planet. In this work, several classification algorithms for garnets trained using band parameters, PCA, and fusion of Raman and FTIR data are proposed with the aim of determining which one presents higher precision.

## 1. Introducción

### 1.1. La exploración de Marte

El estudio de Marte y su superficie comenzó en 2001 con el lanzamiento de la misión 2001 Mars Odyssey. Este tenía como objetivo estudiar la superficie del planeta rojo, de qué elementos químicos estaba conformada y qué minerales se habían formado. (National Aeronautics and Space Agency [NASA], s.f.-a) Desde entonces, se ha seguido con la búsqueda de conocer la corteza de Marte, dando un paso adelante y consiguiendo implementar la tecnología de análisis espectroscópicos en las siguientes misiones.

El rover Mars 2020 Perseverance, lanzado en julio de 2020 hacia el cráter Jezero, se ha equipado para que sea capaz de realizar técnicas analíticas para el estudio de la superficie de Marte con el objetivo de detectar biomarcadores de vida. Para ello estudia el origen y evolución de la geología del planeta. Esto lo consigue a través de siete instrumentos analíticos: Mastcam-Z, MEDA, SHERLOC, RIMFAX, MOXIE, PIXL y SuperCam. Esta última es una cámara, láser y espectrómetro capaz de identificar muestras de minerales a través de: LIBS, la cual determina la composición de la muestra; VISIR, se utiliza para la determinación de minerales; RMI, determina la textura de las muestras, y Raman. Gracias a la fusión de estos métodos se busca estudiar muestras geológicas consiguiendo determinar por completo su apariencia, estructura y composición. Además, las muestras con un alto impacto científico son recolectadas para, en una futura misión (Mars Sample Return), traer a la Tierra. El objetivo de Mars 2020 es buscar rastros de vida microbiana debido a las evidencias de la existencia de agua líquida pasada en el cráter Jezero haciendo un ambiente favorable para el desarrollo y preservación de vida microbiana antigua. De haber existido esta vida, deberían existir signos de la misma en los sedimentos de este antiguo lago. Por ello, esta misión pretende tanto la búsqueda de posibles biomarcadores, como el estudio de la evolución y composición de la geología del planeta. (NASA, s.f.-b)

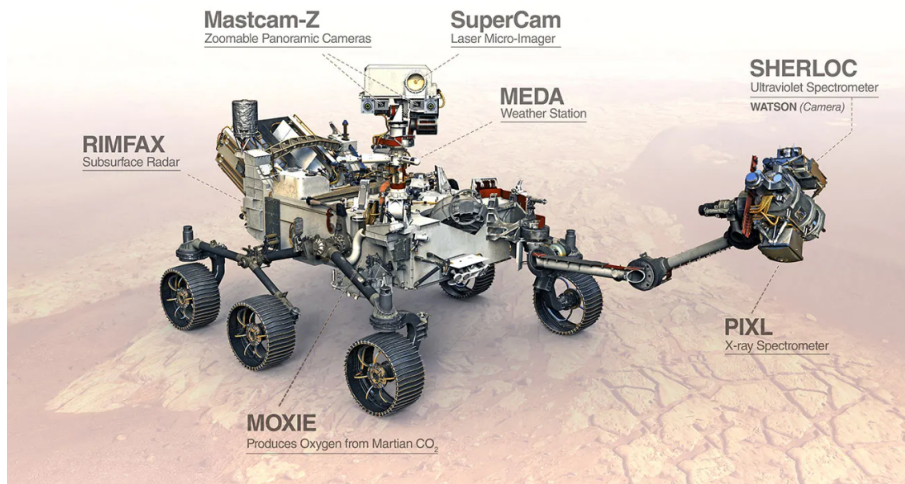


Figura 1: Componentes del rover Perseverance. [2]

La misión ExoMars, liderada por la Agencia Espacial Europea (ESA), también tiene como objetivo el estudio del planeta rojo. En este caso, la misión se compone de dos partes: el Trace Gas Orbiter, lanzado en 2016, y la segunda parte, programada para 2028, que contará con el rover Rosalind Franklin. Los objetivos son: hallar trazas de vida, tanto presente como pasada, y, para ello, se centra en la caracterización del entorno marciano, tanto sus condiciones ambientales como el estudio de la geología que presenta. Además de estudiar la superficie, esta misión incorpora una novedad para el estudio de la geología con un taladro que permite acceder al subsuelo a dos metros de profundidad. Entre los instrumentos que presenta este rover se encuentra el RLS, un espectrómetro Raman en cuyo desarrollo ha colaborado la Universidad de Valladolid. (European Space Agency, [ESA], s.f.)

Como se explica en Venerada, Sanz Arranz et al (2022), en la superficie de Marte podemos encontrar una variedad de rocas volcánicas del tipo de feldespatos, piroxenos y olivinos principalmente, pero también se han formado epidotes o granates. Por esto, una base de datos que permita clasificar el tipo de granate en función de su espectro será de gran utilidad para seguir avanzando en el estudio de este planeta.

Los estudios de espectroscopía Raman enfocada a los granates comenzaron en los años 60. Desde entonces, siguió el interés en el campo de la mineralogía creado en torno a los granates, dirigiéndose varios estudios a entender sus propiedades, su formación y cómo afecta su composición química, entre otros aspectos. (Kolesov et al, 1997)

## 1.2. Espectroscopia Raman y FTIR. Ventajas y limitaciones

La espectroscopía Raman se comenzó a aplicar a la gemología desde hace alrededor de 30 años, aunque ha cobrado más relevancia en esta última década, aumentando el número de laboratorios que cuentan con esta técnica. (Bersani et al, 2009) Esto se debe a las grandes ventajas que presenta este método frente a otro tipo de espectroscopía. Es no invasiva, lo que significa que no se necesita destruir la muestra para realizar su estudio, y no requiere de mayor preparación previa, el tiempo de medida varía en función de cada muestra pero es relativamente rápido. Es por esto que "estas características la convierten en ideal para el análisis in situ en misiones espaciales" (Instituto Nacional de Técnica Aeroespacial [INTA], s.f.)

Como desventaja podemos señalar la dificultad para reconocer la correspondencia con cada pico. Es decir, no disponemos de una base de datos donde poder contrastar qué nos indica el espectro a analizar, si no que deberemos evaluarlo de manera particular y comparativamente con otros minerales ya

identificados.

La espectroscopía Raman se basa en el estudio de la luz monocromática que se dispersa inelásticamente al interactuar con el objeto de estudio. Este cambio en la frecuencia de los fotones dispersados en función de los emitidos se debe a que estos son absorbidos por la muestra y luego se emiten nuevamente tras haber cedido o tomado energía de los mismos. Esta diferencia de energía nos permite obtener información de la composición química y la estructura que conforma el mineral que analizamos. (INTA, s.f.) En el espectro Raman observaremos tres tipos de picos correspondientes al scattering Rayleigh, Raman Stokes y Raman anti-Stokes. El pico correspondiente a Rayleigh será el de dispersión elástica, por lo que no existe un cambio en la frecuencia de los fotones, son emitidos y absorbidos con la misma energía. Cuando el material absorbe energía y los fotones son emitidos con mayor longitud de onda es scattering Raman Stokes, mientras que si son emitidos con mayor energía de onda será Raman anti-Stokes. Cada uno de estos picos Raman nos informa acerca de las transiciones vibracionales y rotacionales de las moléculas que forman la muestra.

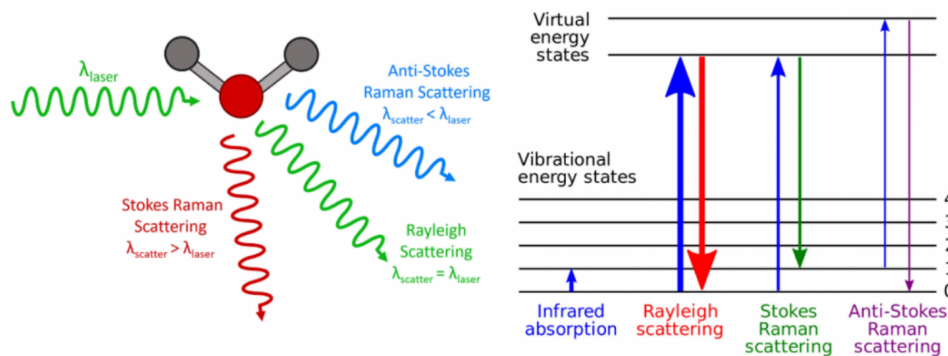


Figura 2: Transiciones de energía para el scattering Raman.

Por otra parte, tenemos la espectroscopía FTIR (Fourier transform infrared spectroscopy), que también determina las vibraciones en las moléculas que conforman el mineral. Esta funciona excitando la muestra con un haz de radiación infrarroja. Una vez interactúa con el mineral, este absorbe frecuencias de este haz características a cada componente. Finalmente, se detecta la señal que la muestra no ha absorbido y, mediante los picos que esta presenta, se identifica cada uno con las frecuencias de vibración de cada enlace molecular. El espectro obtenido se encuentra entre  $4000$  y  $666\text{ cm}^{-1}$  de los cuales podemos separar tres regiones acordes a enlace simple, doble y triple. (Mohamed et al, 2017)

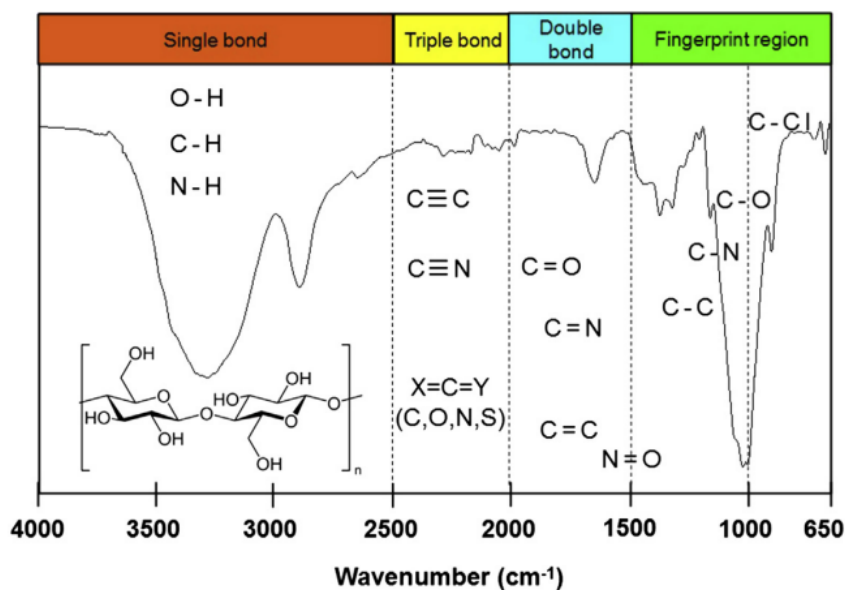


Figura 3: Ejemplo de FTIR de una membrana de celulosa con varios tipos comunes de enlaces. [10]

Para interpretar un espectro FTIR veremos la longitud de onda a la que presenta picos de absorción esa muestra y, con una tabla que nos identifique cada una de las vibraciones, iremos identificando los enlaces que conforman nuestra molécula.

Este método de análisis de muestras también tiene la ventaja de su rápido tiempo de adquisición.

A pesar de que ambos métodos son técnicas vibracionales que nos dan información molecular acerca de la muestra, existen diferencias que hacen que sea verdaderamente útil trabajar con ambas. En la técnica FTIR, la absorción depende del cambio existente en el dipolo permanente de los enlaces químicos de la molécula. La espectroscopía Raman, por el contrario, depende del cambio en el enlace químico de la polarizabilidad. Por tanto, aquellas bandas que tengan una intensidad menor en espectroscopía Raman pueden cobrar más importancia en FTIR y al revés. (Kappler et al, 2016) Estas dos formas de espectroscopía pueden ser útiles por separado pero como verdaderamente podemos aprovecharnos de sus ventajas es si utilizamos ambas de manera complementaria.

### 1.3. Algoritmos de clasificación

Nuestra sociedad siempre ha buscado avanzar técnicamente para buscar facilitar nuestro trabajo. Una vez comenzó la revolución tecnológica, el desarrollo de las computadoras llevaron a que en los años 50 comenzase a hablarse de inteligencias artificiales. Estas se consideraban a aquellas máquinas capaces de realizar un proceso lógico. Esto fue hasta que en la década de los 70 se dejó de considerar esto suficiente y empezó la búsqueda de máquinas capaces de adquirir conocimiento.

En 1980 se comenzaron a desarrollar los primeros algoritmos de los cuales los más relevantes en cuanto a inteligencias artificiales fueron los árboles de decisión y el aprendizaje basado en la lógica (logic-based learning), que simulaban los procesos humanos al minimizar la entropía minimizando la cantidad de información. A mediados de los 90, se popularizó el uso de las redes neuronales como técnica de machine learning. Al mismo tiempo comenzaba también el interés por el aprendizaje estadístico, que rápido ganó importancia con el llamado SVM (Support Machine Vector). (Zhi Hua, 2021)

Tras esto, llegamos a la época del deep learning a principio del milenio donde la capacidad de computación y la gran cantidad de datos necesarias obligó realizar un gran avance en este ámbito.

Zhi-Hua (2021) asegura que en las últimas dos décadas hemos sido testigos de un rápido avance en la recolección, almacenamiento, transmisión y procesamiento de grandes datos. Es por esto que afirma

que no es de extrañar que el machine learning haya ganado tanta atención, pues provee de soluciones que permiten aprovechar el gran poder de los macrodatos.

En concreto, entre estos métodos de machine learning, el trabajo se centrará en los llamados algoritmos de clasificación. Estos mecanismos son capaces de clasificar una serie de datos en categorías previamente señalizadas. Esto se realiza en dos procesos: el primero es un proceso de aprendizaje que será seguido por un test. En esta primera parte, la mayoría de los datos (suele ser en torno al 80% que se divide entre el set de entrenamiento y el set de validación) se dispone para que, de acuerdo con el algoritmo seleccionado, la máquina logre seguir el proceso de aprendizaje para diferenciar las categorías que se han indicado. Tras esto se hará la validación mediante k-folds. A continuación, el resto de datos se utiliza para el test, donde se pondrá a prueba lo aprendido anteriormente intentando clasificar estos datos con determinado nivel de eficacia.

PCA (análisis de componentes principales) es un método estadístico de machine learning cuyo objetivo es el reducir la dimensión del espacio de observación. Esto se logra realizando una transformación que convierte unas variables observables correlacionadas en una serie de variables lineales sin correlación. A estas variables las llamamos componentes principales y nos permiten reducir su dimensión sin perder grandes cantidades de información para conseguir realizar la computación más rápido y de manera más sencilla. (Mahesh, 2019)

## 2. Objetivo

Como ya se ha comentado anteriormente, la espectroscopía Raman es una herramienta muy útil para el estudio de minerales, pero su clasificación resulta algo más compleja. Debido al reciente avance en el estudio de Marte, crear un algoritmo que discrimine el tipo de mineral supondría una mejora realmente significativa en la rapidez del procesamiento de datos de estas misiones.

Por ello el objetivo del presente trabajo es conseguir un modelo de clasificación que discrimine las muestras de minerales granates en base a su espectroscopía Raman y su mejora con la combinación de espectroscopía Raman y FTIR, tanto para muestras tomadas con espectroscopios con misma resolución espectral como para muestras problema tomadas con diferentes equipos.

## 3. Materiales, instrumentación y análisis de datos

### 3.1. Muestras utilizadas

Los granates son un grupo de minerales neosilicatos que cristaliza en un sistema cúbico. Su composición química se compone por el grupo silicato y dos cationes de la siguiente manera:



Los cationes más comunes que ocupan la posición X son  $Ca^{2+}$ ,  $Mg^{2+}$ ,  $Fe^{2+}$  y  $Mn^{2+}$ . Para el catión trivalente Y, los elementos más comunes son  $Al^{3+}$ ,  $Fe^{3+}$  y  $Cr^{3+}$ . De acuerdo con Henderson y Downs (2009) según el convenio de International Mineralogical Societies, podemos dividir los granates en dos grupos. El primero corresponde al catión  $Al^{3+}$  en la posición Y y se le asigna el nombre de pyralspites. Por otro lado tenemos los llamados ugrandites que se corresponden a aquellos con  $Ca^{2+}$  en la posición X.

Se han medido en el laboratorio un conjunto de granates pertenecientes a la colección de minerales ADAMM, (Venerada, 2020), una base de datos del grupo ERICA, compuesto por un total de once minerales diferentes. De cada uno de estos minerales tenemos un número de muestras diferentes que se especifican en la lista más abajo, haciendo un total de veintiséis muestras. Dentro de este grupo de granates, distinguimos dos subgrupos dependiendo de su composición. En subgrupo de pyralspite lo

forman los almandine, pyrope spessartine y rhodolite. Dentro de las ugrandites encontramos andradite, grossular, hessonite, melanite, polyadelphite, shorlomite y uvarovite.

- Almandine 2, almandine 3, almandine 4, almandine 6, almandine 7
- Andradite 2, andradite 3
- Grossular 1, grossular 2, grossular 3, grossular 4
- Hessonite 1, hessonite 2
- Melanite 1
- Polyadelphite 1
- Pyrope 1, pyrope 2
- Rhodolite 1
- Schorlomite 1, schorlomite 2
- Spessartine 2, spessartine 3, spessartine 4
- Uvarovite 1, uvarovite 2, uvarovite 3

Se seleccionaron un conjunto de datos de RRUFF (base de datos espectroscópicas abierta) con la que se ha podido aumentar el número de muestras por cada mineral que tenemos. Estos minerales han sido estudiados con espectroscopios distintos por lo que la resolución espectral de cada uno de ellos será diferente a los espectros obtenidos en el laboratorio. En esta lista se resumen las muestras tomadas de RRUFF, obteniendo siete minerales y veintiocho muestras.

- Almandine 1, almandine 2, almandine 3, almandine 4, almandine 5
- Andradite 1, andradite 2, andradite 3, andradite 4, andradite 5
- Grossular 1, grossular 2, grossular 3, grossular 4, grossular 5
- Pyrope 1, pyrope 2, pyrope 3, pyrope 4, pyrope 5
- Spessartine 1, spessartine 2, spessartine 3, spessartine 4, spessartine 5
- Uvarovite 1, uvarovite2
- Schorlomite 1

### **3.2. Medidas Raman y FTIR**

Los espectros Raman han sido tomados con dos equipos de medida diferentes. Por un lado, se precisaba de los espectros Raman de la base de datos ADdAMM (Venerada et al, 2020). En total, tres espectros Raman por cada muestra disponible en el laboratorio. Estos espectros han sido tomados con el equipo Horiba-JY Induram usando un láser verde de Elforlight G4-PSU, un cabezal Raman Horiba-JY Superhead 532 nm y un microscopio Zeiss AxioTech 30. Debido a la poca cantidad de espectros por cada muestra, se usó el equipo del Simulador del equipo RLS (Lopez-Reyes et al, 2021) , un sistema automático de recolección de espectros. Para la toma de muestras se necesita disponer de cada mineral en polvo fino. Para ello, se tomaron las que se encontraban en polvo y con ayuda de un mortero terminaron de machacar. Después se metieron en una matriz prensados desde donde el láser, de 532nm modelo BWTEK BWN-532-100 comercial con 100-mW de potencia regulable, seleccionará 15 puntos.



Las medidas Raman se tomaron con un BWTEK BTC162 comercial, un cabezal Raman BWTEK INC y un objetivo de enfoque Nikon 50mm F1.4. La adquisición de espectros por espectroscopía infrarroja se realizó con las muestras en polvo pulverizadas mediante el equipo Spectrum 100 FT- IR tomando 3 espectros por muestra con unos parámetros de adquisición de 16 scam.

### 3.3. Metodología

#### 3.3.1. Tratamiento espectral

La visualización, manejo e interpretación de los espectros se ha llevado a cabo a través del programa IDAT/SpectPro. Este software, desarrollado por la Universidad de Valladolid, busca recibir, calibrar y verificar las muestras obtenidas por el RLS en Marte. Ofrece funciones para el tratamiento de espectros tales como cortar, sumar, normalizar y corrección de la línea de base. (Venerada, Saiz et al, 2020) Dado que son muestras naturales, presentan inhomogeneidades. Esto se traduce en variaciones de un punto a otro de la misma muestra, dando lugar a espectros de minerales diferentes. Se ha realizado un análisis preliminar de los espectros, donde se han seleccionados aquellos en los que se obtenía una buena señal con picos definidos de los minerales que se estudian.

Una vez seleccionados los espectros que se tratarán, se realizó la corrección de la línea de base y su normalización. Para determinar los parámetros se ha utilizado IDAT/SpectPro. Tras esto, se ha llevado a cabo la preparación de los datos con el software matemático MATLAB (Matrix Laboratory) haciendo uso de los parámetros dados por SpectPro. El ajuste de bandas permite obtener un espectro más limpio y definido de manera que los parámetros de bandas sean más claros. La corrección de los espectros es esencial en este estudio ya que se han usado datos Raman que han sido tomados con equipos diferentes y parámetros espectrales diferentes, lo que genera que, tanto la resolución espectral como la intensidad o el ruido de los espectros varíen de unos espectros a otros en una misma muestra. También se ha realizado una normalización en base a la intensidad de los picos y una interpolación, para que todos los espectros tengan la misma resolución, lo cual es esencial para poder comparar los espectros entre sí.

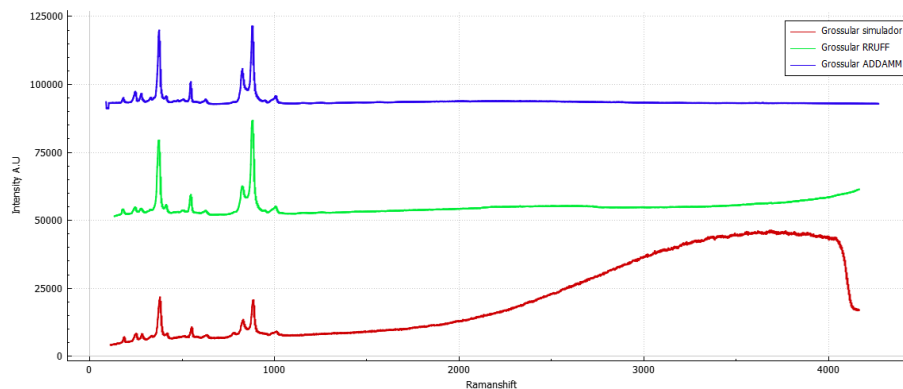


Figura 4: Ejemplo de espectro Raman sin procesar para una muestra de grossular tomada con los tres distintos espectrómetros utilizados.

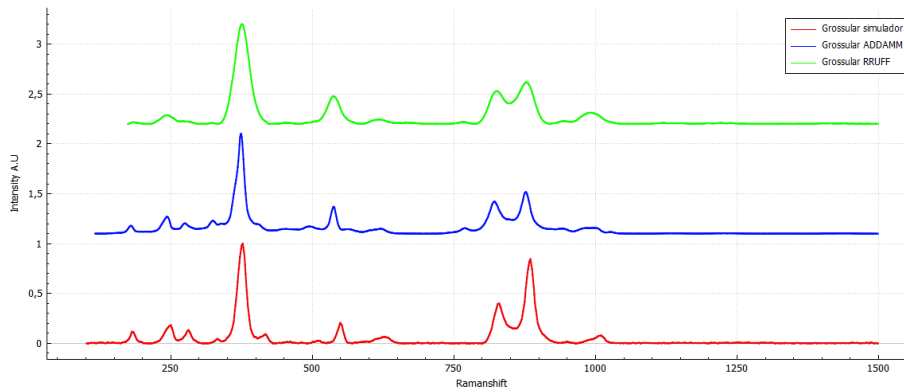


Figura 5: Ejemplo de espectro Raman tras el procesamiento de datos para la misma muestra de grossular tomada con los tres distintos espectrómetros utilizados.

### 3.3.2. Preparación de datos

#### 3.3.2.1 Parámetros de bandas

Una vez obtenidos los espectros ya tratados, se ha realizado un estudio a través de los parámetros de bandas de cada uno de los espectros. Para ellos se han seleccionado los cinco picos que describen el enlace con cada catión y que se repetían en todos los minerales para poder realizar la comparación. Con la herramienta de SpectPro se ha determinado la posición del pico, medida en Raman shift ( $cm^{-1}$ ), sus intensidades respectivas, ya normalizadas anteriormente. A través de los cambios en la posición de estos picos y las intensidades relativas, que dependen de cada catión presente en la su estructura, podemos discernir entre los diferentes minerales. La matriz de datos utilizada, tanto para validación como para test, se adjunta en los anexos A2 y A3 respectivamente.

#### 3.3.2.2 PCA

De manera análoga, se ha decidido realizar un estudio de características principales para comparar con los resultados obtenidos a través de parámetros de bandas. Para realizar este estudio, con la ayuda de MATLAB, se han seleccionado las primeras componentes principales que describen el 95 % de la varianza del espectro. Además se ha realizado una fusión de datos Raman con FTIR. Para eliminar los posibles errores derivados de la diferencia de resolución espectral, se ha realizado también una interpolación en los espectros. Se ha realizado con MATLAB una concatenación de espectros. Cada espectro Raman se ha concatenado con cada espectro FTIR de la misma muestra, de manera que la población de espectros se ve aumentada y sobre los datos concatenados se realiza PCA. Los datos de FTIR son todavía más reducidos, pues para los tomados en el laboratorio necesitamos los minerales en polvo y para las tomadas en RRUFF solo hay datos para un número reducido de muestras. Por ello, al disponer de una población de muestras tan reducida, los resultados de los algoritmos se verán afectados. El script utilizado se presenta en el anexo A1.

### 3.3.3. Métodos de clasificación usados

Para realizar la clasificación se ha empleado la herramienta de MATLAB que permite entrenar algoritmos de clasificación a través de machine learning (Classification Learner en MATLAB). Esta herramienta permite entrenar una gran variedad de algoritmos diferentes de entre los cuales se han seleccionado los siguientes: KNN, ensemble, SVM y redes neuronales.

- KNN (K-nearest neighbor): este modelo se basa en la comparación entre el dato a clasificar con los datos más cercanos. En particular, se analizan los K-vecinos más cercanos y se mide su distancia

al punto. Estudiando la categoría de estos vecinos más cercanos, se establecerá la categoría del nuevo punto. (MATLAB, s.f.-a)

- Algoritmos de ensamble o métodos de conjunto: se basa, como su propio nombre indica, en el ensamble de varios modelos entrenados de manera independiente. Esto permite que, tras juntarlos, se reduzca el error producido por un solo conjunto. (MATLAB, s.f.-b)
- Redes neuronales: busca asemejarse al funcionamiento del sistema de neuronas de nuestro cerebro. De esta manera, se crean redes con nodos conectados que reconocen patrones aprendiendo de datos anteriores para ser capaces de prever o clasificar datos futuros. (MATLAB, s.f.-c)
- SVM (Support vector machine): en este algoritmo se busca encontrar el hiperplano que separe las dos características que buscamos clasificar de manera que delimite el espacio que pertenece a cada clase. Se utiliza para sistemas binarios o múltiples que resulten de la combinación de varios binarios. (MATLAB, s.f.-d)

La precisión del entrenamiento se realiza mediante K-folds con validación cruzada, como ya se ha comentado. Este proceso divide el set de datos de entrenamiento en k particiones. Cada una de ellas servirá para hacer el test del modelo entrenado con el resto de los grupos de muestras. De esta manera, se obtendrán k iteraciones para la validación. (MATLAB, s.f.-e)

#### 3.3.4. Filosofía de los modelos

En la parte de entrenamiento de los modelos se emplearán los datos que se han seleccionado y medido. Con el objetivo de entrenar modelos que nos permitan una clasificación de espectros Raman independientemente del equipo experimental y la resolución espectral, se plantea la idea de entrenar modelos usando los espectros tomados con los equipos de laboratorio, dejando para el conjunto del test todos los datos de RRUFF los cuales tienen resoluciones espectrales diferentes. También se plantea el estudio de algoritmos con espectrómetros de similares características. Para ello, se reservará de la totalidad de datos medidos en el laboratorio una serie de muestras que servirán para el test en este apartado.

Este proceso se realizará siguiendo los dos métodos de reducción de variables de entrada al modelo anteriormente mencionados: parámetros de bandas, utilizando la intensidad y la posición de los picos, y PCA (análisis de componentes principales).

En ambos procesos realizaremos cuatro etapas de clasificación. Primero se pondrán a prueba los datos y los algoritmos seleccionados a través de la diferenciación de subgrupos, de modo que nos separará los minerales en pyralspites y ugrandites. A continuación, con estos subgrupos bien diferenciados, se procederá a realizar la determinación del mineral para cada subgrupo. Por último, también se pondrá a prueba la discriminación de cada mineral sin haberlos separados previamente.

Por tanto, se disponen de las siguientes clasificaciones:

- 1 - Modelo único para clasificar todos los tipos de granates
- 2 - Combinación de modelos en dos fases, primero clasificar tipo, luego el granate dentro del tipo

A - Entrenando con los datos del laboratorio y testeando con RRUFF

B - Entrenando y testeando con los datos del laboratorio (reduce la población de espectros para entrenamiento)

A modo de resumen, se ha probado la siguiente combinatoria:

- Modelos basados en datos Raman (4.3)
  - Usando como método de reducción de parámetros de entrada parámetros de bandas (PB) (4.3.1) modelos R-PB-1A, R-PB-2A, R-PB-1B, R-PB-2B.
  - Usando como método de reducción de parámetros de entrada PCA (4.3.2), modelos R-PCA-1A, R-PCA-2A, R-PCA-1B, R-PCA-2B.
- Modelos basados en fusión de datos Raman + FTIR (4.4)
  - Usando parámetros de bandas (4.4.1) modelos RIR-PCA-1A, RIR-PCA-2A, RIR-PCA-1B, RIR-PCA-2B.

A continuación se adjunta una tabla donde se resume de manera más clara los diferentes apartados:

Modelo	Preclasificación		Tipo de datos		Método		Fuente de datos	
	No	Sí	Raman	Raman+FTIR	PB	PCA	ERICA	ADAMM+RRUFF
R-PB-1A	x		x		x			x
R-PB-2A		x	x		x			x
R-PB-1B	x		x		x		x	
R-PB-2B		x	x		x		x	
R-PCA-1A	x		x			x		x
R-PCA-2A		x	x			x		x
R-PCA-1B	x		x			x	x	
R-PCA-2B		x	x			x	x	
RIR-PCA-1A	x			x		x		x
RIR-PCA-2A		x		x		x		x
RIR-PCA-1B	x			x		x	x	
RIR-PCA-2B		x		x		x	x	

Tabla 1: Resumen de modelos planteados

Los diferentes algoritmos de clasificación serán comparados mediante su precisión y se analizarán los resultados en base a su matriz de confusión para determinar cuál de ellos resulta más útil cuando se trata de clasificar minerales a través de su espectroscopía.

## 4. Resultados y discusión

### 4.1. Caracterización de muestras

A continuación se detalla cada muestra con su composición química pura y la composición obtenida mediante el análisis XRF. Es importante recalcar que se trata de minerales naturales, no puros, por lo que las inhomogeneidades e impurezas son esperables en estas muestras. No se dispone de la composición de uvarovites y pyropes ya que, como se muestra en 6, son muestras pequeñas en un sustrato más grande por lo que no se puede obtener polvo de estos minerales para realizar esta prueba.



Figura 6: Ejemplo de muestra de uvarovita (parte verde) en un sustrato más grande (parte blanca).

$\text{Fe}_3 \text{Al}_2 (\text{SiO}_4)_3$					
	Almandine 1	Almandine 3	Almandine 4	Almandine 6	Almandine 7
Al <sub>2</sub> O <sub>3</sub>	43,46 %	18,93 %	19,27 %	19,05 %	16,14 %
Ca	3,89 %	1,00 %	0,56 %	0,86 %	0,64 %
Cl	2,33 %	0,75 %	0,75 %	0,68 %	0,83 %
Fe	18,50 %	34,22 %	34,79 %	30,76 %	29,96 %
MgO	0,00 %	3,35 %	0,00 %	5,67 %	0,00 %
Mn	3,79 %	1,15 %	1,63 %	1,03 %	14,33 %
SiO <sub>2</sub>	25,75 %	39,59 %	39,82 %	40,62 %	37,56 %
Ti	0,00 %	0,23 %	2,58 %	0,85 %	0,12 %
Total	97,71 %	99,21 %	99,39 %	99,50 %	99,58 %

Tabla 2: Composición almandines

$\text{Ca}_3\text{Fe}_2(\text{SiO}_4)_3$		$\text{Ca}_3\text{Al}_2(\text{SiO}_4)_3$		
	Andradite 2	Andradite 3	Hessonite 1	Hessonite 2
Al <sub>2</sub> O <sub>3</sub>	9,72 %	5,01 %	22,04 %	23,07 %
Ca	24,42 %	28,40 %	25,55 %	27,77 %
Fe	14,67 %	24,14 %	4,33 %	3,44 %
MgO	2,52 %	0,00 %		
SiO <sub>2</sub>	46,50 %	40,73 %	46,30 %	44,47 %
Ti	0,12 %	0,07 %		
Total	97,95 %	98,36 %	98,23 %	98,75 %

Tabla 3: Composición andradites y hessonites

$\text{Ca}_3\text{Al}_2(\text{SiO}_4)_3$					$\text{Ca}_3(\text{Fe}^{3+}, \text{Ti})_2(\text{SiO}_4)_3$
	Grossular 1	Grossular 2	Grossular 3	Grossular 4	Melanite 1
Al <sub>2</sub> O <sub>3</sub>	19,68 %	21,53 %	20,46 %	21,28 %	8,34 %
Ca	28,42 %	27,96 %	25,78 %	27,82 %	26,04 %
Fe	6,00 %	3,09 %	3,16 %	4,34 %	20,70 %
MgO	0,00 %	1,44 %	1,96 %	0,00 %	
SiO <sub>2</sub>	43,82 %	44,65 %	47,12 %	44,67 %	39,82 %
Ti					2,44 %
Total	97,91 %	98,67 %	98,48 %	98,11 %	97,34 %

Tabla 4: Composición grossulares y melanite

$(\text{Ca}, \text{Mn})_3\text{Fe}_2(\text{SiO}_4)_3$		$\text{Ca}_3(\text{Fe}, \text{Ti})_2(\text{Si}, \text{Ti})_3\text{O}_{12}$		
	Polyadelphite 1		Shorlomite 1	Shorlomite 2
Al <sub>2</sub> O <sub>3</sub>	10,75 %	Al <sub>2</sub> O <sub>3</sub>	2,95 %	4,07 %
Ca	10,63 %	Ca	23,64 %	27,05 %
Fe	6,03 %	Fe	22,61 %	19,22 %
MgO	1,51 %	P	0,08 %	1,57 %
Mn	30,91 %	SiO <sub>2</sub>	43,38 %	30,49 %
SiO <sub>2</sub>	33,66 %	Ti	4,80 %	11,91 %
Zn	4,72 %	Zr	0,21 %	3,53 %
Total	98,21 %	Total	97,67 %	97,83 %

Tabla 5: Composición polyadelphites y shorlomites

$\text{Mn}_3\text{Al}_2(\text{SiO}_4)_3$				$(\text{Mg}, \text{Fe})_3\text{Al}_2(\text{SiO}_4)_3$
	Spessartine 2	Spessartine 3	Spessartine 4	Rhodolite 1
Al <sub>2</sub> O <sub>3</sub>	17,82 %	21,26 %	12,12 %	17,30 %
Ca	0,78 %	24,93 %	0,83 %	1,86 %
Fe	7,18 %	5,02 %	15,99 %	23,61 %
MgO	1,33 %	0,00 %	0,00 %	7,11 %
Mn	36,82 %	0,86 %	18,16 %	1,24 %
SiO <sub>2</sub>	34,74 %	46,39 %	51,10 %	44,91 %
K <sub>2</sub> O				2,33 %
Total	98,67 %	98,46 %	98,19 %	98,36 %

Tabla 6: Composición spessartines y rhodolite

El grupo de las almandines no presenta grandes impurezas, salvo en la almandine 7, que se observa cantidades similares de  $\text{Al}_2\text{O}_3$  e impurezas de  $\text{Mn}$ . En el caso de las andradites, presentan unas impurezas de baja concentración de  $\text{Al}_2\text{O}_3$ . Además en la muestra de andradite 3 también se observan leves impurezas de MgO. Las hessonites, grossulares y spessartines añaden a su composición química pequeñas partes de hierro que en la muestra spessartine 4 se hacen más notables. Además, en la spessartine 3 también dispone de impurezas relevantes para su composición de Ca. Tanto la melanite y la polyadelphites como las schorlomites 1 y 2 presentan algo de  $\text{Al}_2\text{O}_3$ , que no se obtiene en las muestras puras de estos minerales. El zinc también está presente en baja proporción (3,53 % y 4,72 % respectivamente) en la schorlomite 2 y polyadelphite 1. La única muestra de rhodolite presenta impurezas que no alcanzan el 6 % en total entre calcio, manganeso y  $\text{K}_2\text{O}$ .

A continuación se presenta la composición química de los datos seleccionados.(RRUFF, s.f.) En la suma total algunas muestras superan el 100 % ya que estos valores son una media de varias medidas con WDS (diffraction-based wavelength-dispersive X-ray spectrometer).

	Almandine 1	Almandine 2	Almandine 3	Almandine 4	Almandine 5
SiO2	40,13	37,84	35,49	40,19	37,12
TiO2	0,06		0,06		
Al2O3	22,69	21,43	20,09	22,81	20,8
FeO	21,99	34,35	21,45	21,41	24,74
MnO	0,5		20,61	0,41	8,34
MgO	11,35	4,51		11,2	4,78
CaO	4,09	2,09	0,72	4,66	2,87
Total	100,8	100,21	98,4	100,68	98,66

Tabla 7: Composición en porcentaje muestras de almandines tomados de RRUFF

	Andradite 1	Andradite 2	Andradite 3	Andradite 4	Andradite 5
SiO2	34,84	36,28	36,58	32,27	25,85
TiO2	0,02	0,06	0,54	5,61	11,79
Al2O3	0,43	6,82	6,06	2,45	0,36
Cr2O3	0,01				0,02
Fe2O3	30,48	21,4	23,55	22,85	28,23
MnO	0,16	10,11	0,95		0,09
MgO	0,03			1,03	0,57
CaO	32,82	25,5	32,39	33	32,63
ZrO2				1,97	
Total	98,78	100,16	100,06	99,17	99,52

Tabla 8: Composición en porcentaje muestras de andradites tomados de RRUFF

	Grossular 1	Grossular 2	Grossular 3	Grossular 4	Grossular 5
SiO2	39,3	36,93	38,31	37,79	39,18
Al2O3	20,85	11,18	19,38	12,91	20,45
TiO2		1,93	0,62	1,03	0,22
CaO	35,92	34,5	35,91	34,79	35,26
FeO	3,86				4,7
Fe2O3		14,96	4,54	13,09	
MgO		0,59	0,11	0,42	0,11
MnO	0,28	0,09	0,2	0,31	0,13
Total	100,21	100,18	99,06	100,33	100,05

Tabla 9: Composición en porcentaje grossulares tomados de RRUFF

	Pyrope 1	Pyrope 2	Pyrope 3	Pyrope 4	Pyrope 5
SiO <sub>2</sub>	40,37	36,29	40,77	41,44	40,09
Al <sub>2</sub> O <sub>3</sub>	22,74	20,23	23,63	22,7	22,7
FeO	19,15	26,74	10,98	12,61	20,28
MnO	0,69	0,27	0,23	0,45	0,83
MgO	13,41	15,84	18,12	16,93	13,24
CaO	3,69	0,42	5,56	4,99	2,57
Total	100,04	99,8	99,28	99,91	99,7

Tabla 10: Composición en porcentaje pyropes tomados de RRUFF

	Spessartine 1	Spessartine 2	Spessartine 3	Spessartine 4	Spessartine 5
SiO <sub>2</sub>	33,02	35,71	37,06	36,54	39,01
Al <sub>2</sub> O <sub>3</sub>	20	20,64	20,77	20,67	22,23
MgO	0,06		0,77	0,39	8,6
CaO	0,54	0,45	9,59	0,32	4,46
MnO	39,06	41,48	24,05	41,07	19,56
FeO	4,35	1,61	6,9	1,09	6,32
Total	99,25	99,9	99,13	100,07	100,37

Tabla 11: Composición en porcentaje spessartines tomados de RRUFF

	Uvarovite 1	Uvarovite 2		schorlomite 1
Na <sub>2</sub> O	0,02	0,02	CaO	31,4
SiO <sub>2</sub>	36,3	37,03	SiO <sub>2</sub>	23,47
MgO	0	0,4	Fe <sub>2</sub> O <sub>3</sub>	18,76
Al <sub>2</sub> O <sub>3</sub>	8,52	7,54	TiO <sub>2</sub>	11,73
CaO	34,65	33,47	ZrO <sub>2</sub>	10,19
TiO <sub>2</sub>	1,31	0,3	Al <sub>2</sub> O <sub>3</sub>	1,84
Fe <sub>2</sub> O <sub>3</sub>	0,23		MgO	1,54
MnO	0	0,11	MnO	0,34
Cr <sub>2</sub> O <sub>3</sub>	17,12	19,13	Na <sub>2</sub> O	0,15
FeO		0,43	K <sub>2</sub> O	0
Total	98,16	98,4	Total	99,27

Tabla 12: Composición en porcentaje uvarovites y schorlomite tomados de RRUFF

La almandine 3 es la muestra de este mineral que mayores impurezas presenta, llegando al 20,61 % en MnO. Almandine 1, 2 y 4 también contienen pequeñas cantidades de MgO y almandine 5, además del MgO, también óxidos de calcio y de manganeso. La andradite 1 es una muestra bastante pura ya que apenas presenta elementos a parte de los componentes químicos ideales.  $Al_2O_3$  está presente en andradite 3 y andradite 2. Además, este último también se compone de MnO en un 10,11 %. En las últimas dos andradites destaca el  $TiO_2$ . Las inhomogeneidades en los grossulares son con el hierro, concretamente con el FeO para el grossular 2, 3 y 4, y con el  $Fe_2O_3$  en grossular 1 y grossular 5. Con una proporción más elevada, el FeO vuelve a presentarse como impureza en el caso de los pyropes. Las spessartines añaden a su composición química pura pequeños porcentajes de CaO y FeO. En las uvarovites y las schorlomites solo se presentan impurezas relevantes de  $Al_2O_3$  y  $ZrO_2$  respectivamente.

En total se disponen de once minerales para la fase de aprendizaje, de los cuales solamente tenemos



siete para la parte de test. A pesar de haber utilizado la base de datos de RRUFF para ampliar la población de muestras, sigue existiendo un desbalance entre los minerales que disponemos para cada fase.

#### 4.2. Análisis cualitativos Raman: correlación entre parámetros de bandas y cationes

De acuerdo con Fu et al (2022), podemos agrupar la variación en los parámetros de bandas en dos grupos, dependiendo de si se conserva el mismo catión X o Y de la siguiente manera.

En el caso de un mismo X y variando Y, existe una dependencia entre el radio del catión variable y la frecuencia de la luz emitida. Es decir, si el mineral se compone de un catión X con un radio mayor los picos se observarán en una posición de menor Raman shift que los correspondientes a un mineral con un catión en esta posición de menor tamaño. Con el grupo de ugrandites (Ca en la posición X), observamos 3 principales cationes en la posición Y: Al, Fe<sup>3+</sup>, Ti, siendo el Ti el radio de mayor tamaño, seguido por el Fe y por último Al. Como se puede observar en la siguiente tabla, si comparamos andradites, hessonites, grossulares y melanites se puede corroborar este desplazamiento en los picos seleccionados.

	Melanite	Andradite	Hessonite	Grossular
Catión Y	Ti <sup>3+</sup> , Fe <sup>3+</sup>	Fe <sup>3+</sup>	Al <sup>3+</sup>	Al <sup>3+</sup>
Peak nº	Raman shift (cm <sup>-1</sup> )	Raman shift (cm <sup>-1</sup> )	Raman shift (cm <sup>-1</sup> )	Raman shift (cm <sup>-1</sup> )
Peak 1	236	242	250	250
Peak 2	359	373	376,5	378
Peak 3	521,5	527	552	550
Peak 4	872	879	885	885
Peak 5	979	1003	1009	1011

Tabla 13: Comparación entre picos de ugrandites

En la siguiente figura se muestra gráficamente con un ejemplo de cada mineral cómo varían los principales picos de acuerdo a lo explicado anteriormente.

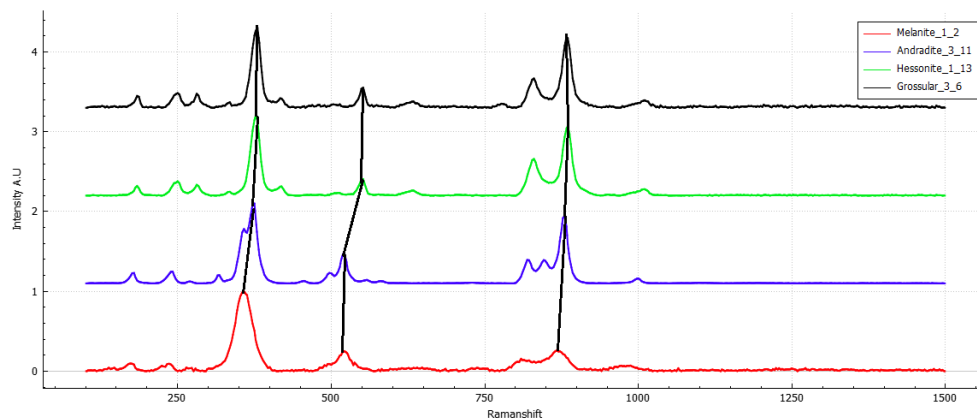


Figura 7: Muestra gráfica de la variación en la posición de los picos en función del catión Y.

Para las pyralspites, es decir fijando el catión Y como aluminio, podemos diferenciar dos casos en el desplazamiento de los picos. Se ha añadido también el catión del calcio, correspondiente al mineral grossular y hessonite del grupo ugrandites, ya que también obedece esta regla al tener aluminio en la misma posición de las pyralspites.

1. A mayor masa atómica del catión X se demuestra un desplazamiento hacia mayor Raman shift en los dos primeros picos que se han analizado. Ordenados de mayor a menor masa atómica

encontramos los siguientes elementos más recurrentes:  $\text{Fe}^{2+}$ , Mn, Ca y Mg. Se adjuntan los datos en la siguiente tabla.

	Almandine	Spessartine	Hessonite	Grossular	Rhodolite
Catión X	$\text{Fe}^{2+}$	$\text{Mn}^{2+}$	$\text{Ca}^{2+}$	$\text{Ca}^{2+}$	$\text{Mg}^{2+}, \text{Fe}^{2+}$
Peak n <sup>o</sup>	R. shift ( $\text{cm}^{-1}$ )	R. shift ( $\text{cm}^{-1}$ )	R. shift ( $\text{cm}^{-1}$ )	R. shift ( $\text{cm}^{-1}$ )	R. shift ( $\text{cm}^{-1}$ )
Peak 1	224	229	236	242	193
Peak 2	355	359	359	373	351

Tabla 14: Comparación entre picos con minerales con aluminio en su composición

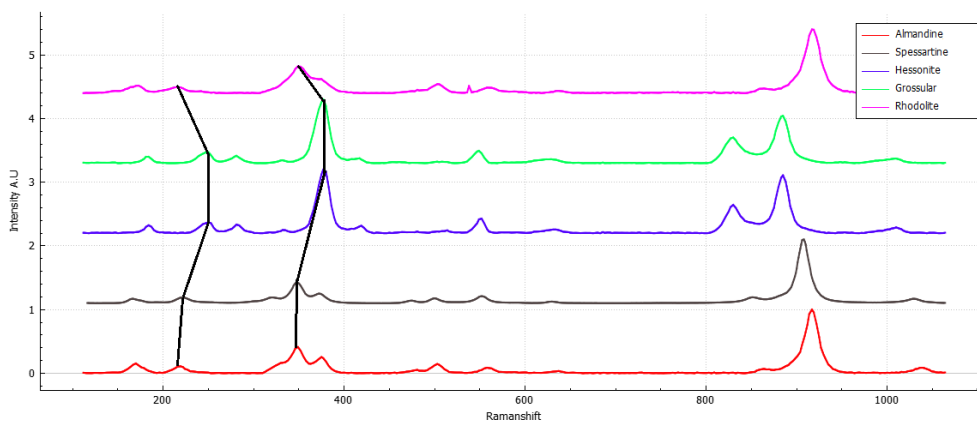


Figura 8: Muestra gráfica de la variación en la posición de los primeros picos en función del catión X.

En este caso, podemos ver una disonancia en el mineral rhodolite. Esto es debido a la gran cantidad de hierro en su composición en comparación a la cantidad de magnesio, lo cual beneficia la aparición del hierro en su espectro.

- En cuanto a los últimos 2 picos seleccionados, estos guardan relación de nuevo con el radio del catión involucrado. En orden creciente serían: Ca, Mn, Fe y Mg. A mayor radio se encuentran a mayor Raman shift, como se observa a continuación.

	Hessonite	Grossular	Spessartine	Almandine	Rhodolite
Catión X	Ca	Ca	Mn	Fe	Mg, Fe
Peak n <sup>o</sup>	R. shift ( $\text{cm}^{-1}$ )	R. shift ( $\text{cm}^{-1}$ )	R. shift ( $\text{cm}^{-1}$ )	R. shift ( $\text{cm}^{-1}$ )	R. shift ( $\text{cm}^{-1}$ )
Peak 4	885	885	904	912	916
Peak 5	1009	1011	1026	1036	1037

Tabla 15: Comparación de la posición de los últimos picos en función del radio del catión en Y

De nuevo, se cumple la importancia del hierro en las rhodolites comparándolos con los almandines ya que presentan una diferencia no demasiado acusada.

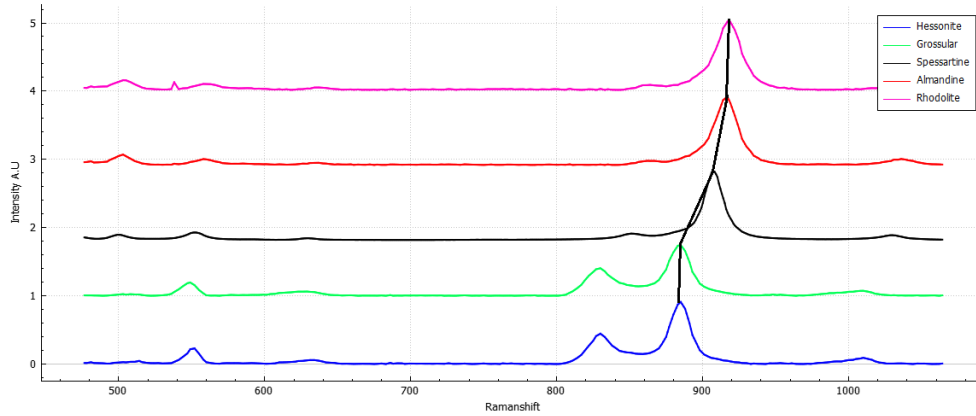


Figura 9: Muestra gráfica de la variación en la posición de los dos últimos picos en función del catión X.

Sabiendo esto, podemos concluir que existe una tendencia en cuanto a la composición de la muestra y el espectro Raman que presenta pudiendo, a través de la comparación con demás minerales, determinar sus cationes y, por tanto, el mineral.

### 4.3. Raman

#### 4.3.1. Parámetros de bandas

##### 4.3.1.1 R-PB-1A

Estos son los resultados de los algoritmos entrenados con datos Raman de laboratorio y RRUFF con parámetros de bandas para la clasificación de todos los minerales.

Los resultados se muestran en la tabla 16 para los modelos testados con base de datos de RRUFF.

RRUFF Test	Validation				Test			
	SVM	KNN	Ensemble	Neural	SVM	KNN	Ensemble	Neural
Mineral	84.56 %	81.88 %	82.21 %	83.22 %	11.11 %	29.63 %	29.63 %	33.33 %
Type	94.97 %	91.61 %	94.63 %	94.30 %	77.78 %	59.26 %	81.48 %	48.15 %
Ugrand	90.57 %	90.57 %	88.05 %	80.50 %	0 %	8.33 %	25.00 %	25.00 %
Pyrrals	93.53 %	94.24 %	93.53 %	84.17 %	40.00 %	40.00 %	53.33 %	33.33 %

Tabla 16: Tabla comparativa modelos entrenados con parámetros de bandas con RRUFF en el test

En la clasificación de minerales, los resultados de la precisión en la fase de test no resultan demasiado satisfactorios.

A nivel general, se puede observar que la precisión disminuye enormemente cuando se realiza el test. No destaca un grupo de minerales que genere mayor confusión, si no que parece existir un error generalizado en los datos de test.

Con el objetivo de mejorar estos resultados, se realiza la preclasificación por tipo de minerales (ugrandites y pyralspites) anterior.

##### 4.3.1.2 R-PB-2A

A continuación se muestran los resultados realizados con los datos del laboratorio y RRUFF con parámetros de bandas para las clasificaciones por tipos.

Los resultados en este caso se espera que sean mejores dado que únicamente se tienen 2 clases que clasificar por lo que el error es menor.

Efectivamente, los resultados en este apartado, que se reflejan en la tabla 16, son mejores que los obtenidos para el resto de apartados, llegando a precisiones que superan el 80 % para el test, aunque todavía no son los esperados.

Una vez determinados los minerales que pertenecen a cada subgrupo, procedemos a realizar la nueva clasificación para cada uno de ellos.

- Ugrandites: Los resultados para las ugrandites son de menor precisión que los de pyralspites. Se observa una precisión del 0,0 % para el algoritmo de SVM. Además, ningún mineral fue correctamente identificado en su totalidad en ninguno de los algoritmos seleccionados. Los mejores resultados han sido del 25 % para el caso de los modelos entrenados con ensamble y redes neuronales.
- Pyralspites: En el caso de las pyralspites, y centrándose en el test, se concluye que el mineral spessartine es el que más problemas genera a la hora de clasificarlo. Los almandines y pyropes, por su parte, dependiendo del modelo se clasifican correctamente o dan lugar a confusión, mayormente entre ellas. Para las redes neuronales y ensamble los almandines son correctamente separados. En el caso de los pyropes, estos son identificados en KNN y SVM. Esto nos lleva a precisiones no demasiado bajas pero que, sin duda, deben mejorarse. En esta categoría el modelo de ensamble es el que mejores resultados ofrece, con una precisión en el modelo superior al 50 %.

Tras las primeras clasificaciones se sacan las siguientes conclusiones:

- Las precisiones en todos los modelos bajan demasiado al realizar el test. Esto implica que los datos tomados para esta fase no son los adecuados. Se espera una mejora en las precisiones de la última fase cuando se utilicen datos tomados en el laboratorio en vez de provenientes de RRUFF para el test.
- Las precisiones más altas son aquellas referidas a la clasificación en subgrupos, dado que únicamente existen dos clases para elegir.
- Se pensaba que la clasificación por minerales una vez separados los subgrupos resultaría con mejores resultados que aquella realizada con todos los minerales juntos. Esto es así para las pyralspites. En el caso de las ugrandites, se observa que la precisión es baja, incluso más que cuando se clasifican todos juntos. De esto podemos deducir que las ugrandites se clasifican con mayor dificultad que las pyralspites.

Se plantean los modelos entrenados con parámetros de bandas para modelos con A la hora de entrenar los modelos, en ciertos minerales no se tienen más que una única muestra incluyendo los datos de RRUFF. Por este motivo los modelos han sido entrenados con cierta información que no se ha podido introducir en el test dando, en muchos casos, resultados erróneos en este apartado con ciertas muestras que no han sido testeadas pero si entrenadas.

#### **4.3.1.3 R-PB-1B**

En este apartado se estudian los modelos entrenados con parámetros de bandas con muestras de laboratorio para la clasificación de todos los granates.

Por el error anteriormente comentado debido a la falta de minerales en la fase de test, se han entrenado estos modelos con minerales de los que se disponen al menos dos muestras. Con lo cual, se han

eliminado melanite y rhodolite.

En el caso de tomar como targets de los modelos todos los minerales sin una previa pre clasificación por subgrupos, se observa como la precisión de la clasificación baja en comparación con la realizada separando en subgrupos.

BF mineral		
Modelo	Train	Test
Neural network	79,65 %	22,22 %
Ensamble	72,09 %	38,27 %
KNN	76,16 %	60,49 %
SVM	78,49 %	40,74 %

Tabla 17: Tabla comparativa de precisiones para modelos que clasifican minerales entrenados con parámetros de bandas

Además, comparando estos valores con los obtenidos usando RRUFF en el test, se ve que nuevamente se obtienen valores con unas precisiones bastante más elevadas aún reduciendo el número de espectros y muestras que se usan para el entrenamiento. A pesar de esto, el porcentaje de acierto en validación es ligeramente inferior en todos los casos, debido nuevamente al número reducido de muestras que se usa para el entrenamiento. En la figura 10 se observa la matriz de confusión para el algoritmo KNN, ya que ha sido el modelo que mejor tasa de acierto ha dado, con una precisión del 60 %. A pesar de haber errores menores en la clasificación de los minerales, la mayoría de las muestras se han entrenado correctamente..

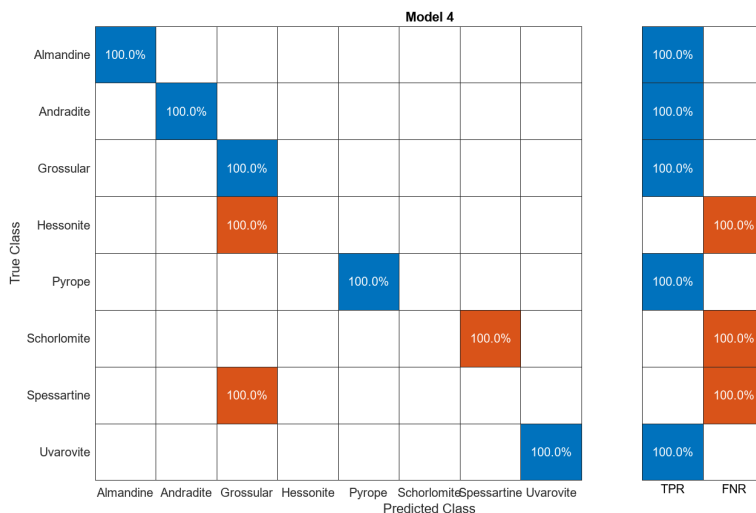


Figura 10: Matriz de confusión de KNN para test en algoritmos de clasificación entrenados con parámetros de bandas para clasificación de minerales con todas las muestras

Como se puede ver, las confusiones se dan al clasificar hessonites como grossulares. Esto es debido a su composición similar, lo que hace que las variaciones entre sus parámetros de bandas sean mínimos. Por otro lado, la muestra de spessartine se ha clasificado como grossular también. Si nos fijamos en la tabla 6 de composición de las spessartines, se observa una muestra (spessartine 3) con una elevada impureza de calcio. Probablemente haya sido esta muestra la destinada a la parte de test y ese catión ha sido el que ha llevado a la confusión ya que es el que toma la posición X en las grossulares y ambas muestras comparten catión Y.

Además, hessonites, schorlomites y spessartines son los minerales de los que menos muestras se tienen.

Al tener solo una muestra para entrenar y otra en test, los modelos pueden estar fallando ya que no son sólidos para la clasificación de estas muestras.

#### 4.3.1.4 R-PB-2B

Modelos entrenados con parámetros de bandas con datos de laboratorio con preclasificación por tipos.

En la tabla 18 vemos como la precisión para los datos en validación es algo más baja en comparación con el porcentaje de aciertos que se obtenían al clasificar por tipo usando los datos RRUFF en el conjunto de test. Esto tiene sentido ya que al testear con datos RRUFF, todos los espectros de ADAMM están disponibles para el entrenamiento ya que no hay que reservar algunos para el entrenamiento

BF type		
Modelo	Train	Test
Neural network	84,88 %	81,48 %
Ensamble	91,28 %	79,01 %
KNN	91,28 %	81,48 %
SVM	89,53 %	79,01 %

Tabla 18: Tabla comparativa de precisiones para modelos de clasificación entrenados con parámetros de bandas

Centrando la atención en los resultados obtenidos para esta segunda fase, se puede ver como el porcentaje de aciertos aumenta en la mayoría de los modelos entrenados, siendo los mejores algoritmos para entrenar los que se obtienen mediante redes neuronales y KNN, con un porcentaje de error del 18,5 % en ambos casos. Además, fijándose en la matriz de confusión para el conjunto de datos de test, se ve como el grupo de las ugrandites las clasifica con un 100 % de aciertos, mientras que es el grupo de las pyralspites son las que generan la tasa de error en este caso, clasificando el 46,9 % de las muestras de pyralspites erróneamente. Por otro lado, se puede representar en un sistema de ejes cartesianos dos de las características espectrales elegidas como inputs para el modelo KNN (en la figura 11), representando con x las muestras que se han clasificado erróneamente. En esta figura se ve como únicamente los espectros de una única muestra perteneciente al tipo de pyralspites son los que se confunden con las ugrandites.

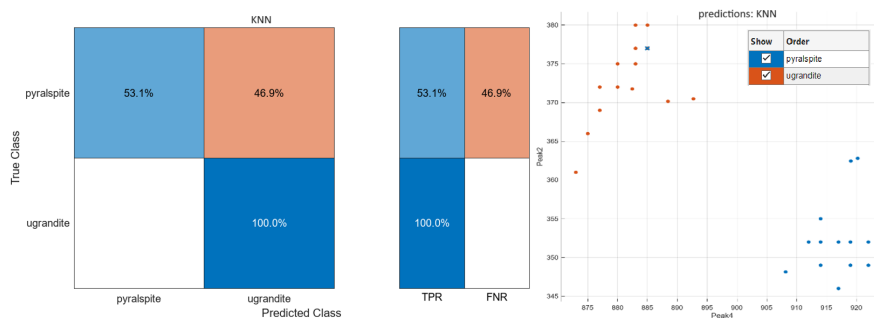


Figura 11: Muestra gráfica de la correlación entre inputs y la relación con la matriz de confusión para KNN en clasificación por tipo.

BF ugrandites		
Modelo	Train	Test
Tree	76,25 %	95,92 %
Neural network	85,00 %	59,18 %
Ensamble	78,75 %	67,35 %
KNN	81,25 %	65,31 %
SVM	82,50 %	69,39 %

Tabla 19: Tabla comparativa de precisiones para modelos de clasificación entrenados con parámetros de bandas

- Ugrandites: En este caso el que mejor modelo resulta es el de arboles de decisión. Este modelo se ha seleccionado debido a su alta precisión en comparación con los algoritmos elegidos. De su matriz de confusión se puede ver como únicamente las uvarovites dan una tasa de error más elevada que la del resto de minerales de este tipo. El árbol de decisiones confunde la clasificación de uvarovites erróneamente con grossulares en un 50 %, lo mismo que ocurre con el modelo de ensemble o SVM, confundiendo este último el 50 % de los datos de test con andradites. En los casos de redes neuronales y KNN, las uvarovites se clasifican correctamente con una tasa de acierto del 100 % en ambos modelos.

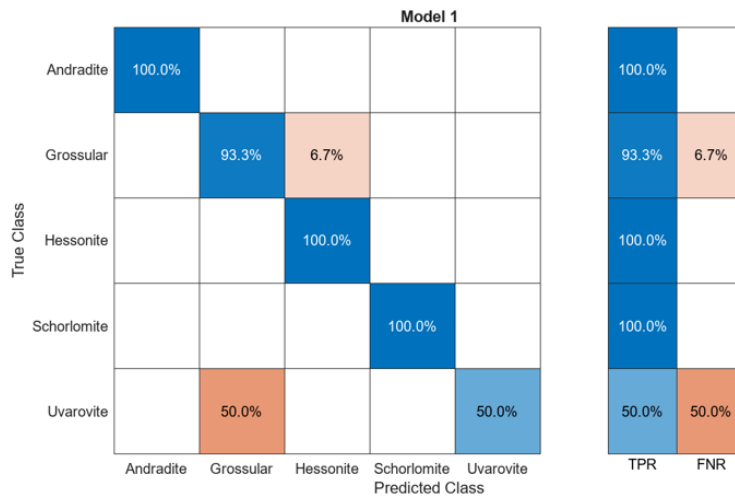


Figura 12: Matriz de confusión del algoritmo de árbol de decisión para clasificación de ugrandites.

- Pyralspites:

BF pyralspites		
Modelo	Train	Test
Neural network	97,83 %	96,74 %
Ensamble	95,65 %	100,00 %
KNN	93,48 %	100,00 %
SVM	98,91 %	100,00 %

Tabla 20: Tabla comparativa de precisiones para modelos que clasifican pyralspites entrenados con parámetros de bandas

Evaluando los resultados de la tabla 20 se puede observar que el resultado de clasificación de las

pyralspites tiene una eficacia del 100 % en tres modelos seleccionados. La matriz de confusión, por tanto es completamente diagonal, como se observa en la figura 13.

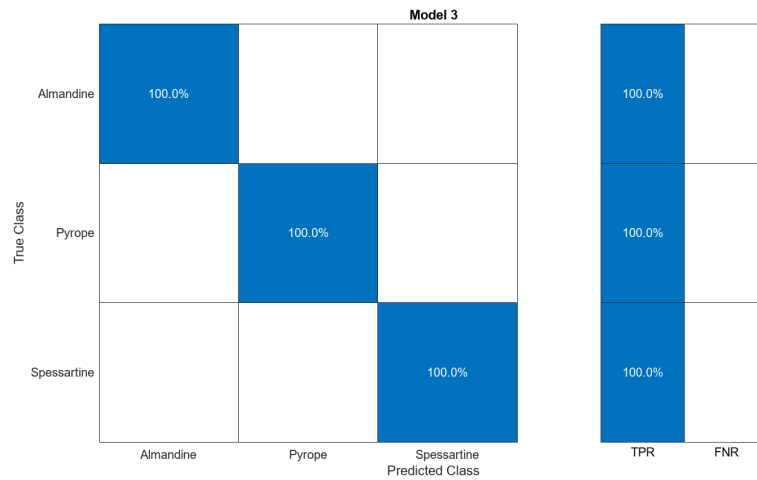


Figura 13: Matriz de confusión de todos los modelos para test en algoritmos de clasificación entrenados con parámetros de bandas para clasificación de pyralspites

#### 4.3.2. PCA

##### 4.3.2.1 R-PCA-1A

Se muestran los resultados entrenados con PCA con datos de RRUFF para el test para clasificación de todos los minerales juntos.

PCA	Validation				Test			
	SVM	KNN	Ensemble	Neural	SVM	KNN	Ensemble	Neural
Mineral	87.25 %	84.23 %	82.55 %	84.90 %	44.44 %	18.52 %	22.22 %	18.52 %
Type	93.96 %	92.95 %	90.94 %	94.30 %	44.44 %	96.30 %	100.00 %	85.19 %
Ugrand	94.34 %	94.97 %	96.23 %	94.97 %	25.00 %	25.00 %	25.00 %	16.67 %
Pyral	98.56 %	94.24 %	94.24 %	100.00 %	33.33 %	13.33 %	0 %	40.00 %

Tabla 21: Tabla comparativa modelos entrenados con PCA

La mayor precisión se obtiene con el algoritmo de SVM con un 44,4 %, aunque esto sigue sin resultar a la altura de lo esperado, existe una mejora respecto a los parámetros de bandas anteriores.

En el caso de KNN y redes neuronales se percibe una tendencia de clasificar todos como un solo mineral, spessartine o grossular, respectivamente. Es por esto que dan un error tan alto. Puede indicar que los parámetros seleccionados no son suficientes para determinar las diferencias entre unos minerales y otros.

##### 4.3.2.2 R-PCA-2A

En estas clasificaciones con PCA para datos de laboratorio y RRUFF con clasificación por tipos obtenemos los siguientes resultados:

Para la clasificación en subgrupos mediante PCA destaca el algoritmo de ensamble, que proporciona un 100 % de precisión. Si se compara con los anteriores modelos determinamos que se obtienen mejores resultados que con parámetros de bandas con datos de RRUFF en la sección de test (como se está



haciendo en este apartado). Pero no todos los modelos resultan con tanta precisión. Para el caso del modelo SVM, todos los minerales se clasifican como ugrandites, por tanto necesitamos mayor descripción en cada mineral para distinguirlo.

- Pyralspites: Como muestra la tabla 21, existen unas precisiones muy diferentes entre modelos. Las redes neuronales clasifican mejor que con parámetros de bandas, llegando al 40 % de precisión. En general se confunden almandines y spessartines y se clasifican erróneamente como pyrope. La precisión se hace nula cuando entrenamos el modelo de ensamble.
- Ugrandites: En este caso se presentan unas precisiones más parecidas para todos los modelos pero más bajas. Como se comentó para parámetros de bandas, la clasificación de ugrandites resulta la más complicada de realizar correctamente. La mayoría de errores vienen por clasificar andradites y grossulares como uvarovites.

Se puede concluir que lo comentado anteriormente a cerca del test con RRUFF sigue cumpliéndose para PCA. La tendencia de las ugrandites siendo la clasificación más complicada también continúa.

#### 4.3.2.3 R-PCA-1B

Para la clasificación con todos los minerales con datos solamente de laboratorio y sin preclasificación existe una disminución bastante relevante en comparación con la precisión obtenida para la clasificación de pyralspites. Sin embargo, se han obtenido resultados algo mejores que los obtenidos para las ugrandites. En este caso el mejor modelo resulta ser el algoritmo de ensamble con un 54,88 %.

PCA mineral		
Modelo	Train	Test
Neural network	90,38 %	20,73 %
Ensamble	91,03 %	24,39 %
KNN	91,67 %	36,59 %
SVM	87,18 %	58,88 %

Tabla 22: Tabla comparativa de precisiones para modelos que clasifican minerales entrenados con parámetros de bandas

La matriz de confusión de la figura 14 muestra que almandines, andradites y grossulares se clasifican correctamente en todos los casos. Sin embargo, el resto de minerales no han sido capaces de identificarse, clasificándolos de manera errónea en grossulares, almandines y spessartines.

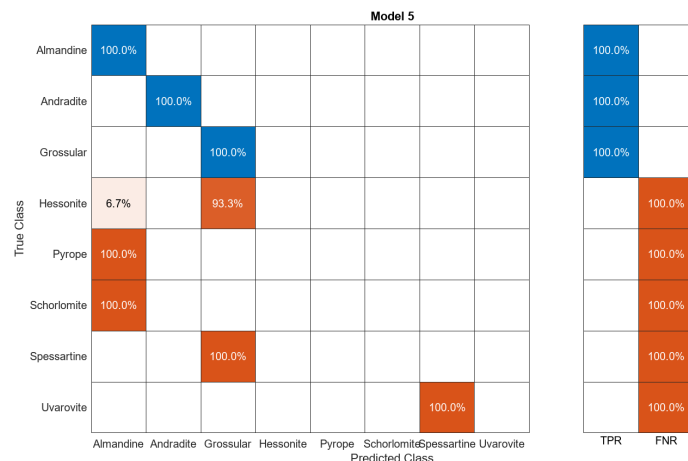


Figura 14: Matriz de confusión para el algoritmo de ensamble para clasificar todos los minerales juntos utilizando PCA.

Hessonites y spessartines continúan clasificándose erróneamente al igual que en el apartado anterior para parámetros de bandas. La confusión con pyropes y almandine, por lo que se explicó en la clasificación de pyralspites, también sigue presente.

#### 4.3.2.4 R-PCA-2B

Se presentan los resultados de las clasificaciones realizadas con PCA para datos de RRUFF con clasificación en ugrandites y pyralspites.

En este caso se observa que los datos para la validación han sido clasificados con total precisión, pero esta disminuye al realizar el test. Se obtienen resultados algo peores en comparación con los parámetros de bandas para este apartado.

PCA type	Train	Test
Modelo	Train	Test
Neural network	100,00 %	63,41 %
Ensamble	100,00 %	63,41 %
KNN	100,00 %	60,98 %
SVM	100,00 %	63,41 %

Tabla 23: Tabla comparativa de precisiones para modelos que clasifican por tipo entrenados con PCA

La precisión de todos los modelos es bastante parecida y han clasificado cometiendo los mismo errores. Como se ve en la matriz de confusión (figura 15) 63,8 % pyralspites han sido confundidos por ugrandites.

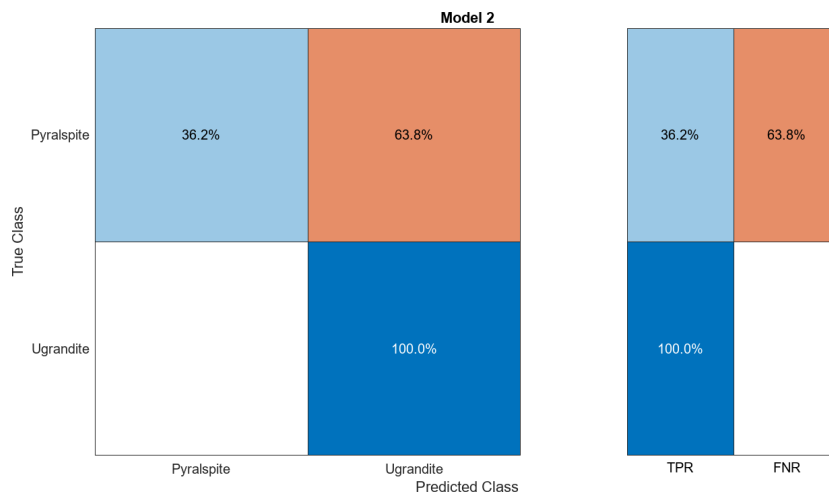


Figura 15: Matriz de confusión para el algoritmo SVM de clasificación de subgrupos para PCA.

- Ugrandites: Los resultados de la precisión en esta sección no han sido los esperados. Hay una precisión del 0 % para el modelo de ensamble y la mayor precisión es únicamente un 30 %. Las ugrandites están siendo las más complicadas de clasificar dado que suelen llevar a más confusiones. Los resultados en la parte de validación tampoco mejoran en gran medida los obtenidos realizando el test con RRUFF.

PCA ugrandites	Train	Test
Modelo		
Neural network	91,14 %	16,00 %
Ensamble	87,34 %	0,00 %
KNN	87,34 %	30,00 %
SVM	87,34 %	30,00 %

Tabla 24: Tabla comparativa de precisiones para modelos que clasifican ugrandites entrenados con PCA

Como se observa en la gráfica de correlación de parámetros representada en la figura 16, los minerales guardan poca dependencia en base a su grupo, por ello los resultados son tan bajos, lo que ha hecho que se clasifiquen en el test todos los minerales como hessonites.

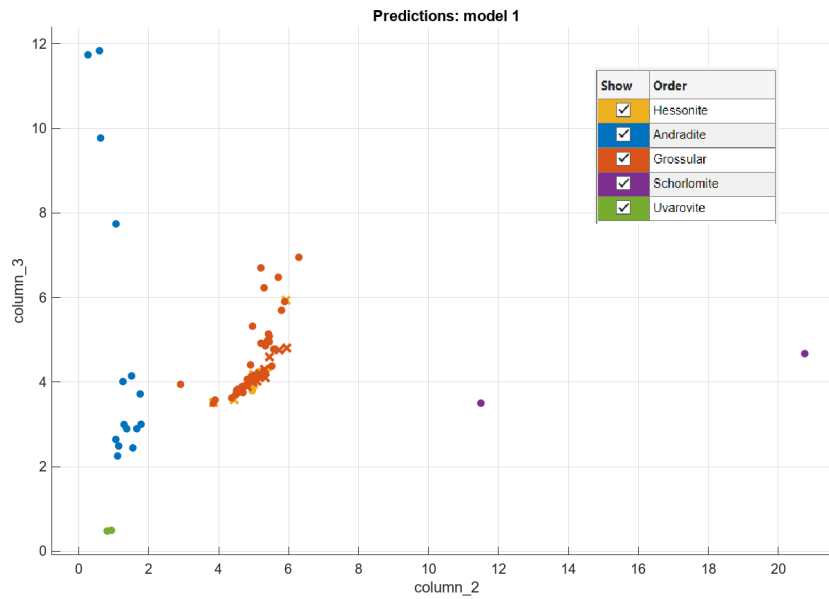


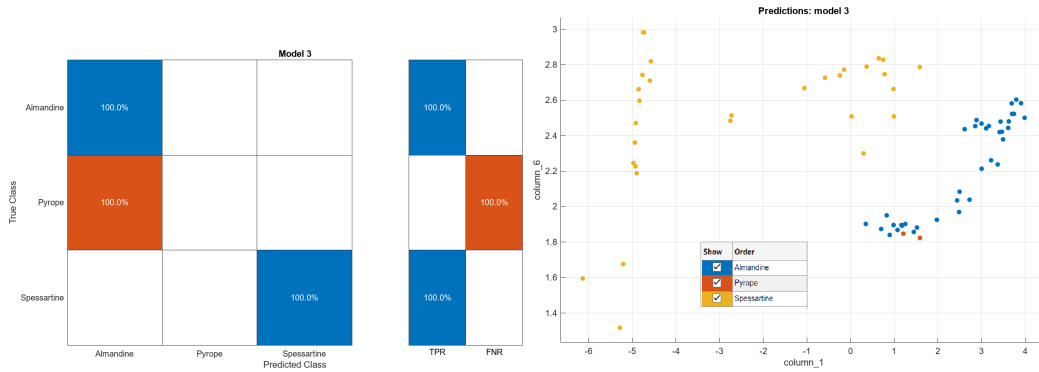
Figura 16: Correlación de los parámetros 3 y 4 para la clasificación de ugrandites en PCA.

- Pyralspites: Continuando con lo observado anteriormente, los resultados para la clasificación de pyralspites resultan mejores que la precisión obtenida con el mismo método en ugrandites. Para los datos de validación la precisión es del 100 %, lo cual se reduce al realizar el test, sobretodo en los algoritmos de redes neuronales y SVM que reducen más de la mitad. En el caso de KNN y ensamble se sigue manteniendo por encima del 90 %.

PCA pyralspites		
Modelo	Train	Test
Neural network	100,00 %	46,88 %
Ensamble	100,00 %	93,75 %
KNN	98,70 %	90,62 %
SVM	100,00 %	46,88 %

Tabla 25: Tabla comparativa de precisiones para modelos que clasifican pyralspites entrenados con PCA

En la figura 17a se observa que tanto almandines como spessartines se han podido clasificar correctamente con redes neuronales. La confusión la provocan los pyropes, que se confunden con almandines. En la gráfica de la figura 17b se puede comprobar cómo estos parámetros han podido generar el error que ha llevado a la incorrecta clasificación de los pyropes.



(a) Matriz de confusión para el modelo de (b) Correlación en los parámetros 1 y 6 paredes neuronales que clasifica pyralspites con ra redes neuronales que clasifican pyralspites con PCA.

Figura 17: Descripción del modelo de ensamble para pyralspites

Se representan los puntos azules como almandines y los rojos como pyropes. La confusión entre almandines y pyropes puede deberse a la poca desviación en los espectros que hay entre el magnesio (pyrope) y el hierro (almandine).

#### 4.4. RIR (fusión datos Raman+ FTIR)

Estos resultados anteriores pueden mejorarse realizando una fusión de datos Raman con FTIR ya que los resultados obtenidos solamente con Raman quedan escasos. Para ello se realiza una nueva matriz de componentes principales que engloben los datos que ofrecen ambas espectroscopías.

##### 4.4.1. PCA

###### 4.4.1.1 RIR-PCA-1A

Se muestran los datos realizados con fusión Raman y FTIR con PCA para modelos testeados con datos de RRUFF clasificando todos los minerales.

Fusión	Validation				Test			
	SVM	KNN	Ensemble	Neural	SVM	KNN	Ensemble	Neural
Mineral	98.95 %	99.21 %	97.11 %	99.08 %	18.18 %	27.27 %	9.09 %	45.45 %
Type	100.00 %	100.00 %	99.61 %	100.00 %	54.55 %	100.00 %	100.00 %	100.00 %
Ugrand	98.80 %	98.07 %	97.35 %	98.31 %	0 %	0 %	16.67 %	16.67 %
Pyralis	100.00 %	99.71 %	96.81 %	98.55 %	0 %	0 %	60.00 %	40.00 %

Tabla 26: Tabla comparativa modelos entrenados con fusión de datos

La gran limitación de los espectros FTIR disponibles en RRUFF hace que la combinación de datos Raman FTIR genere una población de datos reducida, dando problemas a la hora del entrenamiento y testeo de los modelos. Más en concreto, los datos FTIR disponibles en RRUFF son de las muestras de la siguiente lista:

- Almandine 1, almandine 2, almandine 3, almandine 4
- Andradite 1, andradite 2, andradite 3, andradite 4

- Grossular 1, grossular 2, grossular 3
- Spessartine 1

Debido al limitado número de datos no se dispone una cantidad suficiente de espectros como para realizar un entrenamiento que permita evaluar el funcionamiento del modelo con datos específicos para testeo. Por ello, la comparación de estos modelos se realiza con la precisión obtenida en el proceso de validación. Como se ve, únicamente hay dos minerales diferentes por cada tipo de granate, por lo que los datos no van a ser igual de representativos que los de los modelos entrenando Raman individualmente.

Observando la tabla 26, se ve como los modelos en todos los casos funcionan extraordinariamente bien en la parte de validación comparación con los resultados que se obtenían con datos Raman únicamente en la tabla 21. La mayor mejora se percibe cuando se clasifican todos los minerales juntos, llegando al 99 % en dos de los modelos seleccionados. Únicamente se confunden algunas muestras de de hessonites con grossulares. Esto es debido a que poseen la misma composición química ideal. También confunden spessartine con almandine, como se ha visto en el análisis cualitativo, estas muestras están compuestas por manganeso y hierro respectivamente. La confusión puede deberse a que en los cambios en base al catión, hierro y manganeso tienen menor diferencia en los desplazamientos.

Estas mejoras en la parte de validación en comparación a usar únicamente los datos de Raman lleva a pensar que el aumento de información en cuanto a los datos espectrales (obtenidos en este caso con dos técnicas complementarias) que se introducen como inputs en el entrenamiento mejora satisfactoriamente el resultado de los modelos entrenados.

#### 4.4.1.2 RIR-PCA-2A

A continuación se muestran los datos de fusión Raman y FTIR con PCA para datos en test de RRUFF con preclasificación.

Aunque la falta de suficientes datos en la parte del test para realizar la combinación limita para poder testear estos modelos correctamente, se ve como la clasificación por tipo general resultados sin ningún porcentaje de error, mejorando los resultados usando solo datos Raman.

- Ugrandites: Las ugrandites siguen dando bastantes errores, sin ser capaces los modelos de crear una clasificación correcta.
- Pyralpites: En el caso de las pyralpites, la tasa de acierto es algo mejor si se usa el modelo entrenado por algoritmos de ensamble, con una tasa de acierto mucho mejor que cualquier modelo entrenado con datos Raman para las pyralpites. Además, la clasificación del test también mejora atendiendo a la clasificación de todos los minerales contrastando con lo que se obtenía para los datos de Raman.

Por lo tanto, se puede ver como, la clasificación por ugrandites sigue dando peores resultados aun con la combinación de datos y siendo el modelo predominante y el que mejor resultados da la clasificación por tipos de minerales seguidos de una segunda clasificación por cada tipo.

Además, aún con la poca cantidad de muestras que quedan en el set de test, se ve que todos los modelos mejoran en comparación con solo usando datos Raman, tanto en la parte de validación como en la de test, a excepción de la clasificación de ugrandites.

#### 4.4.1.3 RIR-PCA-1B

En este caso se muestran los resultados de la fusión de datos Raman y FTIR con PCA para datos entrenados con muestras de laboratorio y todos los minerales juntos.

Las precisiones en el caso de validación resultan bastante altas, siendo la más baja para el algoritmo de ensamble y SVM con un 98 %. En la gráfica de correlación de parámetros se puede observar que existe una tendencia relativamente clara para cada tipo de minerales, facilitando su clasificación. Aunque la precisión a la hora de testear los modelos es algo inferior en comparación con el uso de PCA para los datos Raman.

fusión minerales		
Modelo	Train	Test
Neural network	99,06 %	30,65 %
Ensamble	98,12 %	31,66 %
KNN	99,06 %	47,24 %
SVM	98,12 %	15,58 %

Tabla 27: Tabla comparativa de precisiones para modelos que clasifican minerales entrenados con fusión Raman FTIR

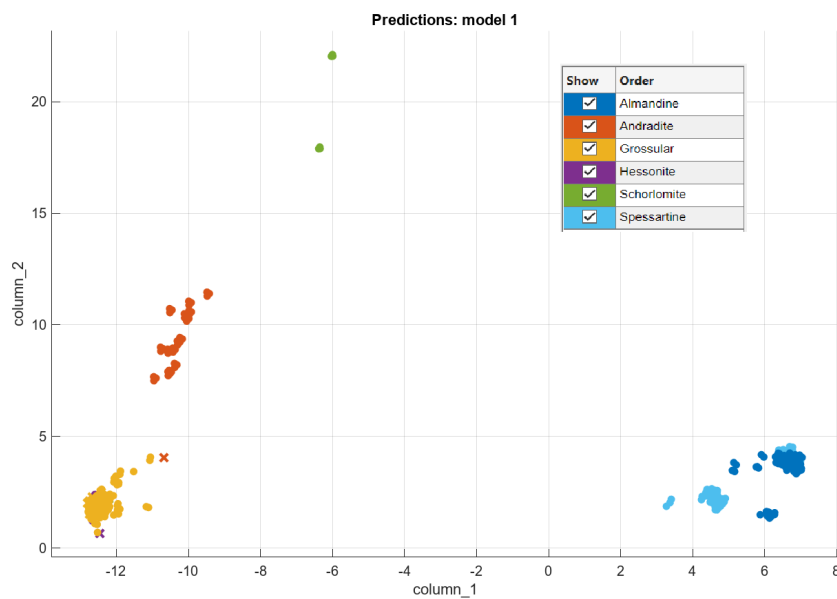


Figura 18: Correlación de parámetros 1 y 2 para KNN que clasifica minerales con fusión Raman FTIR.

Comparando estos datos con la clasificación mediante datos del laboratorio usando PCA sobre sus espectros Raman, se ve que la mejora es notoria en todos los casos. La precisión aumenta pasando de un 60 % a un 77 % de aciertos en la clasificación por tipo, de un 30 % a un 43 % para las ugrandites. Aunque estos resultados no se han podido mejorar en el caso de las pyralspites y la clasificación de todos los minerales.

#### 4.4.1.4 RIR-PCA-2B

En este caso se muestran los resultados de la fusión de datos Raman y FTIR con PCA para datos entrenados con muestras de laboratorio y con clasificación en ugrandites y pyralspites.

Continuando con la línea anterior, las precisiones de validación salen con un 100 % de acierto mientras que cuando evaluamos el test esto disminuye, siendo los mejores modelos redes neuronales y ensamble con un 77,39 %. Tal y como ocurre en el caso de PCA en la clasificación por tipos con SVM, la confusión de estos modelos viene dada por clasificar erróneamente el 50 % perteneciente a las pyralspites en el grupo de ugrandites. Como ya se ha comentado, los valores tan bajos en la precisión de los modelos en test se debe a una descompensación en los datos entre el grupo de validación y de test, además de una población muestral muy reducida.

fusion type		
Modelo	Train	Test
Neural network	100,00 %	77,39 %
Ensamble	100,00 %	77,39 %
KNN	100,00 %	75,38 %
SVM	100,00 %	54,77 %

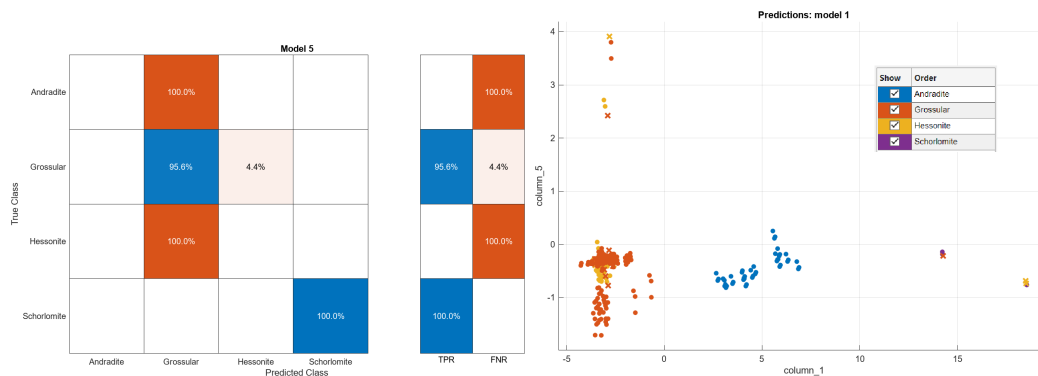
Tabla 28: Tabla comparativa de precisiones para modelos que clasifican por tipos entrenados con fusión Raman FTIR

- Ugrandites: Al igual que en las anteriores clasificaciones de ugrandites, las precisiones no resultan muy buenas. Como se observa en la tabla 29 la correlación entre parámetros vuelve a no resultar del todo clara para realizar la clasificación.

fusion ugrandites		
Modelo	Train	Test
Neural network	96,76 %	41,28 %
Ensamble	93,06 %	6,42 %
KNN	95,83 %	3,67 %
SVM	96,76 %	43,12 %

Tabla 29: Tabla comparativa de precisiones para modelos que clasifican ugrandites entrenados con fusión

Tomando el algoritmo SVM, dado que es el que mejores resultados presenta, se observa que andradites y hessonites se confunden con grossulares. Se puede ver en la tabla 3 la composición de las andradites. En la muestra 2 aumenta el aluminio a la vez que disminuye el hierro respecto a la muestra de andradite 3. Por tanto, esta andradite 2 parece ser la utilizada a en la fase de test ya que su composición se aleja de las andradites mientras se acerca hacia las grossulares.



(a) Matriz de confusión para el modelo SVM (b) Correlación en los parámetros 1 y 5 para que clasifica ugrandites con fusión Raman ra SVM que clasifican ugrandites con fusión Raman FTIR.



- Pyralpsites: En este caso, mediante redes neuronales, se ha sido capaz de obtener un 61,1 % de precisión para la clasificación de pyralpsites.

fusión pyralpsites		
Modelo	Train	Test
Neural network	100,00 %	61,11 %
Ensamble	100,00 %	51,11 %
KNN	100,00 %	16,67 %
SVM	100,00 %	21,11 %

Tabla 30: Tabla comparativa de precisiones para modelos que clasifican pyralpsites entrenados con fusión Raman FTIR

Teniendo en cuenta que únicamente entrenamos dos clases de pyralpsites y los resultados en la validación son del 100 %, vuelven a aparecer las limitaciones presentadas para la clasificación mediante la fusión de datos Raman y FTIR. El descenso tan acusado en la precisión es una causa de los resultados no concluyentes debido a la falta de muestras para este apartado.

Finalmente se ha podido comprobar que los datos obtenidos de la base de datos de RRUFF suponen una limitación a la hora de evaluar los algoritmos. Tras el cambio en cuanto a introducir los datos descargados en el entrenamiento se ha podido comprobar una mejora notoria. Esto podría seguir mejorando con el aumento del número de muestras y datos ya que no se necesitaría datos externos que aporten mayor confusión a los modelos.

#### 4.5. Resumen de resultados

En este trabajo se ha realizado una caracterización de las muestras cuyos espectros se utilizaron posteriormente para entrenar diferentes modelos. Esta caracterización consistió por un lado en el análisis de los resultados analíticos que se habían analizado en el laboratorio (recogidos en el apartado 4.1), así como una caracterización cualitativa de la correlación entre los parámetros de bandas de los espectros, y los cationes presentes en la fórmula química (4.2). Esta caracterización muestra que efectivamente la clasificación basada en datos espectrales es posible.

Posteriormente, se entrenan y evalúan un número notable de modelos con diferentes características (4.3 y 4.4).

Los resultados de todos estos modelos entrenados han mostrado como, aún en las clasificaciones con peores resultados, siempre se ha obtenido una tasa de error menor a la hora de realizar un pre proceso de clasificación por tipos que por directamente la clasificación con todos los granates.

Se ha visto como, usando los parámetros de bandas como inputs a nuestros algoritmos de clasificación, se obtienen una precisión mayor que en el caso de uso de PCA, tanto para los datos Raman, con y sin RRUFF, como para la fusión de datos RAMAN y FTIR, aunque la combinación de estos datos es prometedora.

Además, algunos de los errores cometidos en las clasificaciones se han podido explicar con impurezas presentes en las muestras. Este error se podría disminuir con un mayor número de muestras que cubran un rango de impurezas y muestras puras mayor, de modo que el modelo logre una correcta identificación.

## 5. Conclusiones

Gracias al avance en las nuevas tecnologías, la exploración de Marte es cada vez mayor. Debido al número de datos que se procesan y la cantidad de minerales y muestras que se recogen gracias a los rovers, es necesario automatizar el proceso de identificación de estos minerales. Por ello el objetivo de

este trabajo es facilitar y agilizar el tratamiento de estos datos.

Tras realizar un análisis cualitativo de los datos espectrales crudos se ha podido establecer la correlación entre la posición e intensidad de las bandas espectrales y los cationes que forman parte del mineral, cuantificando el desplazamiento de los picos principales en función del catión correspondiente. Por tanto, se ha concluido que a través de machine learning es posible entrenar un algoritmo de clasificación que pueda diferenciar estos desplazamientos y discriminar así el mineral del que se trata. Además, de esta misma manera se han podido apreciar las diferencias entre los dos subgrupos que conforman los minerales granates: pyralspites y ugrandites.

El objetivo final de un modelo de clasificación es que sea capaz de discernir cuál es el mineral con el que se corresponde un espectro, a priori, desconocido en función de sus características espectrales. Para ello se ha planteado realizarlo de dos maneras. La primera consiste en realizar un único modelo que permita la clasificación de todos los granates. La segunda opción atacada ha sido mediante una filosofía de modelos encadenados de complejidad creciente, de forma que el primer modelo clasificaba entre los dos grandes grupos de granates, que posteriormente son clasificados en los diferentes minerales.

Además, se han utilizado diferentes combinaciones de parámetros para el entrenamiento de los modelos, tal y como se recoge en la tabla 1, utilizando diferentes tipos de datos (Raman, o Raman+FTIR), diferentes técnicas de reducción de datos de entrada (Parámetros de bandas o PCA), o incluso selección de conjuntos de datos de varias bases de datos (ADAMM vs ADAMM+RRUFF)

De los resultados obtenidos se deduce que la clasificación por subgrupos ha resultado muy útil, ya que, al tener únicamente 2 targets, es la que mejor precisión presenta en cualquiera de los modelos entrenados. También se puede observar que la clasificación en minerales hecha una vez separados ugrandites y pyralspites obtiene mejores resultados, sobretodo para pyralspites, que la realizada con todas las muestras juntas. Esto es debido a la existencia de características comunes que comparte cada subgrupo. De modo que al estudiar inputs que en la clasificación conjunta resultaban mezclados, al reducir el espacio que se tiene en cuenta pueden aparecer pequeñas diferencias que antes pasaban desapercibidas.

Se puede concluir, a partir de los resultados obtenidos, que la metodología de reducción de parámetros de entrada al modelo de PCA resulta en una mayor precisión de los algoritmos a nivel general. Concretamente, los mejores resultados en la fase de validación se obtienen en la fusión Raman con FTIR, tanto utilizando datos tomados en el laboratorio como datos descargados de RRUFF para el test. También en la fase de test se obtienen buenos resultados aún con los pocos datos de los que se disponen.

En líneas generales, la clasificación llevada a cabo con muestras tomadas con el mismo espectroscopio resulta en mejores precisiones para todas las clasificaciones, tal y como por otro lado cabría esperar. Para los modelos testeados con datos de RRUFF los resultados no son en general satisfactorios, al haber observado un cierto sobre entrenamiento a los datos de la base de datos ADAMM, que es la que se utiliza para el entrenamiento. Los trabajos futuros deberán centrarse en mejorar el set de datos, mejorando en cuanto a la cantidad de muestras tomadas con diferentes equipos espectrales y la cantidad de muestras que tenemos es también limitada. No obstante, los resultados obtenidos permiten concluir que los modelos desarrollados, al presentar un buen comportamiento frente a datos obtenidos con los mismos instrumentos, serán valiosos para realizar futuros análisis con el instrumento RLS de la misión Exomars, cuyas características son similares al equipo utilizado para adquirir los datos de la base de datos ADAMM.

Por ello, se propone una línea de investigación futura en la que se continúe desarrollando este proyecto. Para ello se necesitan un mayor número de muestras que sea posible obtener en polvo, para que puedan

ser evaluadas con el mayor número de espectroscopías disponible. En esta línea de investigación se propone seguir con la clasificación por fusión de datos Raman y FTIR, ya que se ha visto que es la que mejores resultados genera, de las dos maneras analizadas en el presente trabajo. Este mayor número de muestras podría ser además analizado con diferentes equipos con variadas características (longitudes de onda, resolución espectral, etc), de forma que los futuros modelos desarrollados puedan ser robustos frente a un mayor número de instrumentos y equipos, de forma que puedan ser aplicados en un ámbito investigador más general.

## **6. Impacto científico**

Dada la ayuda prestada por la Universidad de Valladolid a la misión ExoMars de la Agencia Espacial Europea desarrollando el simulador del RLS (Raman Laser Spectrometer), se ha enfocado la investigación a este ámbito. El objetivo de este trabajo apunta a mejorar el estudio de planetas, concretamente de Marte, a través de misiones espaciales que implementen tecnología Raman. Esto mejoraría la rapidez de procesamiento de datos que son capaces de recoger los rovers pudiendo analizar las muestras y nombrarlas con tan solo un algoritmo de clasificación previamente entrenado.

Tras los resultados obtenidos al trabajar con el RLS Sim, tanto para datos de validación como test, y tal y como se ha mencionado, se puede concluir que estos algoritmos presentan buenos resultados cuando se trabaja con una misma resolución espectral. De esta manera se han entrenado unos algoritmos de clasificación con datos conocidos de la Tierra con el RLS Sim y en un futuro se podrían clasificar mediante estos mismos algoritmos minerales marcianos desconocidos. Por tanto, se presentan resultados prometedores y novedosos en cuanto al avance en la investigación de este planeta se refiere. Además, el presente trabajo será presentado de manera oral en el congreso GeoRaman2024 el cual tendrá lugar el próximo mes de septiembre. Para ello se ha presentado un abstract el cual ha sido aceptado.

## Referencias

- [1] National Aeronautics and Space Agency [NASA]. (s.f.-a). *Mars Odyssey*. Recuperado el 2 de mayo de 2024. <https://science.nasa.gov/mission/odyssey/>
- [2] National Aeronautics and Space Agency [NASA]. (s.f.-b). *Mars 2020: Perseverance rover*. Recuperado el 2 de mayo de 2024. <https://science.nasa.gov/mission/mars-2020-perseverance/>
- [3] European Space Agency [ESA]. (s.f.). *exomars. Has life ever existed on Mars?*. Recuperado el 15 de junio de 2024. [https://www.esa.int/Science\\_Exploration/Human\\_and\\_Robotic\\_Exploration/Exploration/ExoMars](https://www.esa.int/Science_Exploration/Human_and_Robotic_Exploration/Exploration/ExoMars)
- [4] M. Veneranda, A. Sanz Arranz, J. A. Manrique, J. Saiz, C. Garcia-Prieto, E. Pascual-Sanchez, J. Medina, M. Konstantinidis, E. Lalla, A. Moral, L. M. Nieto, F. Rull, G. Lopez Reyes, *J Raman Spectroscopy*, 2022, 53(3), 364. <https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/10.1002/jrs.6215>
- [5] R.R. Henderson, R.T. Downs. (2009) *Determining chemical composition of the silicate garnets using Raman spectroscopy*. [Tesis del master in science, Universidad de Arizona]
- [6] Albert Einstein, Un punto de vista heurístico sobre la producción y transformación de la luz, *Annalen der Physik* **17**, 132-148 (1905)
- [7] B.A. Kolesov, C.A. Geiger, S. Quartieri, G. Artioli, G. Antonioli, *Phys Chem Minerals*, 1998, 25:142, 151 [https://www.researchgate.net/publication/235448023\\_A\\_temperature\\_dependent\\_X-ray\\_Absorption\\_Fine\\_Structure\\_study\\_of\\_dynamic\\_X-site\\_disorder\\_in\\_almandine\\_A\\_comparison\\_to\\_diffraction\\_data](https://www.researchgate.net/publication/235448023_A_temperature_dependent_X-ray_Absorption_Fine_Structure_study_of_dynamic_X-site_disorder_in_almandine_A_comparison_to_diffraction_data)
- [8] D. Bersani, S. Andò, P. Vignolac, G. Moltifiori, I.G. Marinoa, P.P. Lottici a, V. Diellac, *Spectrochimica Acta Part A* **73**, 2009, 484-491 <https://www.sciencedirect.com/journal/spectrochimica-acta-part-a-molecular-and-biomolecular-spectroscopy>
- [9] Instituto Nacional de Técnica Aeroespacial, (s.f.). ExoMars Raman <https://www.inta.es/ExoMarsRaman/es/instrumento-rls/espectroscopia-raman/>
- [10] M.A. Mohamed, J. Jaafar, A.F. Ismail, M.H.D. Othman, M.A. Rahman, *Membrane Characterization* 2017, Pages 3-29, <https://doi.org/10.1016/B978-0-444-63776-5.00001-2>
- [11] Veneranda, M., Saiz, J., Lopez-Reyes, G., Manrique, J. A., Sanz Arranz, A., Garcia-Prieto, C., Werner, S. C., Moral, A., Madariaga, J. M., Medina, J., and Rull, F.: *PTAL, ADAMM and Spect-Pro: novel tools to support ExoMars and Mars 2020 science operations*, *Europlanet Science Congress 2020*, 21 September-9 Oct 2020, EPSC2020-1018, <https://doi.org/10.5194/epsc2020-1018, 2020>
- [12] Fu, M.; Dai, J.; Zhao, L. *A Study on the Raman Spectral Characteristics of Garnet from the Jiama Copper Polymetallic Deposit in Tibet*. *Minerals* 2022, **12**, 1578. <https://doi.org/10.3390/min12121578>
- [13] Zhou, Zhi Hua *Machine learning*, 2021, pages 11-17. <https://doi.org/10.1007/978-981-15-1967-3>
- [14] Mahesh, Batta *Machine Learning Algorithms- A review*, 2019, 10.21275/ART20203995, <https://www.researchgate.net/publication/344717762>

- [15] MathWorks. (s.f.-a). *Classification KNN*. Recuperado el 15 de junio de 2024. <https://es.mathworks.com/help/stats/classificationknn.html#d126e201649>
- [16] MathWorks. (s.f.-b). *Ensamble algorithms*. Recuperado el 15 de junio de 2024. <https://es.mathworks.com/help/stats/ensemble-algorithms.html>
- [17] MathWorks. (s.f.-c). *Redes neuronales*. Recuperado el 15 de junio de 2024. <https://es.mathworks.com/discovery/neural-network.html>
- [18] MathWorks. (s.f.-d). *Support Vector Machine*. Recuperado el 15 de junio de 2024. <https://es.mathworks.com/discovery/support-vector-machine.html>
- [19] MathWorks. (s.f.-c). *cvpartition*. Recuperado el 15 de junio de 2024. <https://es.mathworks.com/help/stats/cvpartition.html>
- [20] RRUFF. (s.f.). *Search RRUFF sample data*. Recuperado el 1 de junio de 2024. <https://rruff.info/>
- [21] K appler, A., Fischer, D., Oberbeckmann, S. et al. *Analysis of environmental microplastics by vibrational microspectroscopy: FTIR, Raman or both?*. Anal Bioanal Chem 408, 8377–8391 (2016). <https://doi.org/10.1007/s00216-016-9956-3>
- [22] G. Lopez-Reyes, M. Veneranda, J. A. Manrique,  . Gonz alez Mart n, A. Moral, C. Perez-Canora, J. A. Rodr guez Prieto, A. Sanz Arranz, J. Saiz Cano, E. Lalla, M. Konstantinidis, O. Prieto-Ballesteros, J. Medina, M.  . Gonz alez, E. Charro, J. M. Lopez, F. Rull, *J Raman Spectrosc* 2022, 53(3), 382. <https://doi.org/10.1002/jrs.6281>

## Anexo

### A1. Script de MATLAB para PCA

```
% Raman

tableDataRaman_Train = table;
[nbOfFiles, pathname, ficheros]=selectSpectraToOpen_fn('Espectros Raman ', 'path', '*.');

%Se interpolan los datos para tener el mismo tamaño en los datos tomados y los de RRUFF
intensityMatrix_Train = zeros(nbOfFiles, numel(rShift_i));
MatrizIntensidadesRaman_Train = zeros(nbOfFiles, numel(rShift_i));
SamplesRamanTrain = zeros (numel(nbOfFiles), 1);

for i=1:1:nbOfFiles

    filename=ficheros{i};
    filedata = strsplit(filename, {'_', '.xy'});
    tableDataRaman_Train.index(i)=i;
    tableDataRaman_Train.Sample(i) = filedata(1);
    tableDataRaman_Train.Point(i)=str2double(filedata(2));
    SampleComplete = [filedata{1}, '_', filedata{2}];
    SampleComplete = string(SampleComplete);
    tableDataRaman_Train.Name(i) = SampleComplete;
    if any(strcmp(tableDataRaman_Train.Sample{i}, {'Andradite', 'Grossular', 'Hessonite',
    'Melanite', 'Polyadelphite', 'Schorlomite', 'Uvarovite'}))
        tableDataRaman_Train.Type{i} = 'Ugrandite';
    else
        tableDataRaman_Train.Type{i} = 'Pyralspite';
    end
    [rshift, intensity] = openSpectrum_wholeSpectrum_fn(pathname, filename, '');

    intensity_i = interp1(rshift, intensity, rShift_i);
    intensityMatrix_Train(i,:)=intensity_i;
    MatrizIntensidadesRaman_Train = fillmissing(intensityMatrix_Train, 'constant', 0);
end

%FTIR

tableDataFTIR_train = table;
[nbOfFiles, pathname, ficheros]=selectSpectraToOpen_fn('path', '*.');

intensityMatrixFTIR_Train = zeros(nbOfFiles, numel(lambda_i));
MatrizIntensidadesFTIR_Train = zeros(nbOfFiles, numel(lambda_i));
```

```

for i=1:1:nbOfFiles

    filename=ficheros{i};
    filedata = strsplit(filename, {'-', '_', '.xy'});
    tableDataFTIR_train.index(i)=i;
    tableDataFTIR_train.Sample(i)=filedata(1);
    tableDataFTIR_train.Point(i)=str2double(filedata(2));
    SampleComplete = [filedata{1}, '_', filedata{2}];
    SampleComplete = string(SampleComplete);
    tableDataFTIR_train.Name(i) = SampleComplete;
    if any(strcmp(tableDataFTIR_train.Sample{i}, {'Andradite', 'Grossular', 'Schorlomite',
    'Uvarovite'}))
        tableDataFTIR_train.Type{i} = 'Ugrandite';
    else
        tableDataFTIR_train.Type{i} = 'Pyralspite';
    end
    [wavenumber, intensity_FTIR] = openSpectrum_wholeSpectrum_fn(pathname, filename, '');
    MaxIntensityFTIR = max(intensity_FTIR);
    for j=1:1:length(intensity_FTIR)
        if i==1
            MatrizIntensidadesFTIR_Train=zeros(nbOfFiles, numel(intensity_FTIR));
        end
        MatrizIntensidadesFTIR_Train(i,j)= intensity_FTIR(j)/MaxIntensityFTIR;
    end

end
target_trainFTIR = tableDataFTIR_train.Name;

%Fusion Raman FTIR

NamesRAMAN = vertcat(tableDataRaman_Train.Name(:));
NamesRAMAN = string(NamesRAMAN);
NamesFTIR = vertcat(tableDataFTIR_train.Name(:));
NamesFTIR = string(NamesFTIR);

GarnetList=unique(NamesRAMAN);

MatrizIntensityFusion_Train = zeros();
contadorMatrizFus = 1;
tableFus = table;

for j = 1:numel(GarnetList)
    currentGarnet = GarnetList(j);

    indicesRaman =
    tableDataRaman_Train.index(strcmpi(tableDataRaman_Train.Name, currentGarnet));
    indicesFTIR =

```

```

tableDataFTIR_train.index(strcmpi(tableDataFTIR_train.Name, currentGarnet));

if ~isempty(indicesRaman) && ~isempty(indicesFTIR)

    for k = 1:length(indicesRaman)
        ramanData = MatrizIntensidadesRaman_Train(indicesRaman(k), :);
        for m = 1:1:length(indicesFTIR)
            ftirData = MatrizIntensidadesFTIR_Train(indicesFTIR(m), :);
            fusedData(contadorMatrizFus, :) = cat(2, ramanData, ftirData);

            tableFus.indice(contadorMatrizFus)=contadorMatrizFus;
            tableFus.Name(contadorMatrizFus)=currentGarnet;
            CorrectName = strsplit(currentGarnet, {'_'});
            tableFus.Sample (contadorMatrizFus) = CorrectName (1);
            tableFus.Point (contadorMatrizFus) = str2double(CorrectName (2));
            contadorMatrizFus=contadorMatrizFus+1;
        end
    end
end
end

contadorUgrnadite = 1;
contadorPyralspyte = 1;
for i = 1:1:numel(fusedData(:,1))

    if any(strcmp(tableFus.Sample{i}, {'Andradite', 'Grossular', 'Hessonite',
'Melanite', 'Polyadelphite', 'Schorlomite'}))
        tableFus.Type{i} = 'Ugrandite';
    else
        tableFus.Type{i} = 'Pyralspite';
    end
    tableFus.Tipo(i) = string(tableFus.Type{i});

    if tableFus.Tipo(i) == 'Ugrandite'
        matrizIntesity_UgranditeFused (contadorUgrnadite,:) = fusedData (i,:);
        contadorUgrnadite = contadorUgrnadite + 1;
    else
        matrizIntesity_PyralspiteFused (contadorPyralspyte,:) = fusedData (i,:);
        contadorPyralspyte = contadorPyralspyte + 1;
    end
end
end

targetFused_all = tableFus.Sample(:);
targetTypeFused = tableFus.Tipo(:);
targetUgranditeFused = tableFus.Sample(tableFus.Tipo == 'Ugrandite');
targetPyralspiteFused = tableFus.Sample(tableFus.Tipo == 'Pyralspite');

```



%PCA

```
[loadings,scores_orig,variances, t2, explanation] = pca(fusedData);
scoresF_train = fusedData*loadings;
```

```
PcsFused_Training_All = scoresF_train(:,1:5);
```

## A2. Tabla de datos de parámetros de bandas para entrenamiento

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	mineral	sample	type	Peak 1	Peak 2	Peak 3	Peak 4	Peak 5	Peak 1	Peak 2	Peak 3	Peak 4	Peak 5
2	Almandine	1	pyralspite	250	377	549	885	1008	0,172024	1	0,219486	0,839141	0,0824994
3	Almandine	1	pyralspite	250	377	549	885	1010	0,172989	1	0,193701	0,743375	0,0720737
4	Almandine	1	pyralspite	250	377	549	885	1008	0,178484	1	0,211204	0,816303	0,0715856
5	Almandine	1	pyralspite	250	377	549	885	1008	0,177596	1	0,205392	0,83573	0,076621
6	Almandine	1	pyralspite	250	377	549	885	1010	0,169334	1	0,200639	0,736887	0,0696785
7	Almandine	1	pyralspite	250	377	549	885	1010	0,164295	1	0,229669	0,83072	0,0784779
8	Almandine	1	pyralspite	250	377	549	885	1008	0,166241	1	0,185137	0,735606	0,0696013
9	Almandine	1	pyralspite	250	377	549	885	1008	0,172425	1	0,208923	0,847685	0,0806349
10	Almandine	1	pyralspite	250	377	549	885	1010	0,182566	1	0,210027	0,861025	0,0818679
11	Almandine	1	pyralspite	250	377	549	885	1008	0,168053	1	0,207675	0,857582	0,0790611
12	Almandine	1	pyralspite	250	377	549	885	1010	0,176748	1	0,220762	0,879909	0,0828994
13	Almandine	1	pyralspite	250	377	549	885	1010	0,177263	1	0,202783	0,800012	0,0752123
14	Almandine	1	pyralspite	250	377	549	885	1010	0,182021	1	0,205013	0,844214	0,0798541
15	Almandine	1	pyralspite	250	377	549	885	1013	0,175045	1	0,21689	0,874669	0,0794611
16	Almandine	1	pyralspite	250	377	549	885	1010	0,176904	1	0,211301	0,87453	0,0821285
17	Almandine	3	pyralspite	216	349	503	919	1039	0,084324	0,397673	0,152779	1	0,0879881
18	Almandine	3	pyralspite	216	349	503	919	1039	0,0800402	0,402615	0,151133	1	0,0851535
19	Almandine	3	pyralspite	216	349	503	919	1041	0,0830721	0,393469	0,146567	1	0,0905195
20	Almandine	3	pyralspite	219	349	505	919	1044	0,0731678	0,385828	0,138909	1	0,0850962
21	Almandine	3	pyralspite	219	349	503	919	1039	0,0848933	0,399313	0,155255	1	0,0897708
22	Almandine	3	pyralspite	219	349	503	919	1039	0,0816806	0,382641	0,152948	1	0,0874701
23	Almandine	3	pyralspite	216	349	503	919	1039	0,0873506	0,396951	0,156908	1	0,0872196
24	Almandine	3	pyralspite	219	349	503	919	1044	0,0791829	0,393163	0,150164	1	0,0926391
25	Almandine	3	pyralspite	216	349	503	919	1039	0,0831888	0,390987	0,140191	1	0,0834811
26	Almandine	3	pyralspite	216	349	503	919	1039	0,0896427	0,40118	0,153936	1	0,088365
27	Almandine	3	pyralspite	216	346	503	917	1039	0,0941226	0,409714	0,155604	1	0,0876296
28	Almandine	3	pyralspite	219	349	503	919	1041	0,0865896	0,395424	0,152139	1	0,0877559
29	Almandine	3	pyralspite	219	349	503	919	1044	0,0721576	0,397284	0,132841	1	0,0820896
30	Almandine	3	pyralspite	216	349	503	919	1041	0,0819664	0,402823	0,147681	1	0,080151
31	Almandine	3	pyralspite	219	349	503	919	1044	0,081709	0,389148	0,14916	1	0,0873074
32	Almandine	4	pyralspite	216	352	505	922	1047	0,0626767	0,353345	0,126846	1	0,0842241
33	Almandine	4	pyralspite	216	352	505	922	1044	0,0703692	0,371453	0,140938	1	0,0832268
34	Almandine	4	pyralspite	216	352	505	922	1047	0,065348	0,351552	0,134343	1	0,0866469
35	Almandine	4	pyralspite	216	352	505	922	1047	0,0648213	0,353396	0,139701	1	0,0829683
36	Almandine	4	pyralspite	216	352	505	922	1047	0,0627108	0,356935	0,127764	1	0,0842589
37	Almandine	4	pyralspite	216	349	505	922	1044	0,0638148	0,361348	0,126886	1	0,0819186
38	Almandine	4	pyralspite	219	352	505	922	1047	0,0608328	0,361492	0,115746	1	0,0769103
39	Almandine	4	pyralspite	216	352	505	922	1044	0,0633923	0,355813	0,123	1	0,0769103
40	Almandine	4	pyralspite	216	349	505	922	1047	0,0706287	0,360456	0,133454	1	0,0839917
41	Almandine	4	pyralspite	219	352	505	922	1044	0,0700791	0,372185	0,136912	1	0,0811239
42	Almandine	4	pyralspite	216	352	505	922	1047	0,0640589	0,362131	0,142051	1	0,0814061
43	Almandine	4	pyralspite	216	349	505	919	1044	0,0686324	0,385968	0,127396	1	0,0814117
44	Almandine	4	pyralspite	213	349	505	922	1044	0,0690899	0,370973	0,125573	1	0,0813731
45	Almandine	4	pyralspite	216	352	505	922	1047	0,0556407	0,360072	0,121147	1	0,0806342
46	Almandine	4	pyralspite	219	352	505	922	1047	0,0470278	0,277182	0,116778	0,775002	0,0636937
47	Almandine	6	pyralspite	219	352	505	922	1047	0,064551	0,362055	0,126262	1	0,0838605
48	Almandine	6	pyralspite	216	352	505	922	1047	0,0626767	0,353345	0,126846	1	0,0842241
49	Almandine	6	pyralspite	216	352	505	922	1044	0,0703692	0,371453	0,140938	1	0,0832268
50	Almandine	6	pyralspite	216	352	505	922	1044	0,0557432	0,317149	0,128447	0,908137	0,077885
51	Almandine	6	pyralspite	216	352	505	922	1047	0,065348	0,351552	0,134343	1	0,0866469
52	Almandine	6	pyralspite	216	352	505	922	1047	0,0648213	0,353396	0,139701	1	0,0829683
53	Almandine	6	pyralspite	216	352	505	922	1049	0,0639048	0,364337	0,128095	1	0,0835609
54	Almandine	6	pyralspite	216	352	505	922	1047	0,0627108	0,356935	0,127764	1	0,0842589
55	Almandine	6	pyralspite	216	352	505	922	1047	0,0640589	0,362131	0,142051	1	0,0814061
56	Almandine	6	pyralspite	216	349	505	919	1044	0,0686324	0,385968	0,127396	1	0,0814117
57	Almandine	6	pyralspite	213	352	505	922	1047	0,070306	0,365974	0,144209	0,99908	0,0830385
58	Almandine	6	pyralspite	216	352	505	922	1044	0,067892	0,365458	0,137059	1	0,0783317
59	Almandine	6	pyralspite	216	349	505	922	1047	0,0695912	0,362005	0,12809	1	0,0829208

	A	B	C	D	E	F	G	H	I	J	K	L	M
60	Almandine	6 pyralspite	216	352	505	922	1044	0.0683975	0.367587	0.129023		1	0.0818274
61	Almandine	6 pyralspite	219	352	505	922	1047	0.0470278	0.277182	0.116778	0.775002		0.0636937
62	Almandine	7 pyralspite	219	349	503	917	1036	0.120417	0.408065	0.151468		1	0.0868067
63	Almandine	7 pyralspite	219	349	503	917	1036	0.124247	0.42806	0.153874		1	0.0837772
64	Almandine	7 pyralspite	219	349	503	917	1036	0.115368	0.400308	0.144961		1	0.0823508
65	Almandine	7 pyralspite	219	349	503	917	1039	0.109209	0.411015	0.144753		1	0.0861755
66	Almandine	7 pyralspite	219	349	503	917	1036	0.109847	0.398166	0.1499		1	0.0849652
67	Almandine	7 pyralspite	219	349	503	917	1036	0.110762	0.401898	0.147097		1	0.083051
68	Almandine	7 pyralspite	219	349	503	917	1036	0.115556	0.40495	0.146278		1	0.0882275
69	Almandine	7 pyralspite	219	349	503	917	1039	0.112466	0.399978	0.141852		1	0.0861516
70	Almandine	7 pyralspite	219	349	503	917	1039	0.101828	0.399347	0.150493		1	0.0922632
71	Almandine	7 pyralspite	219	349	503	917	1036	0.116467	0.435084	0.15588		1	0.0829236
72	Almandine	7 pyralspite	219	349	503	917	1036	0.109801	0.39997	0.14217		1	0.0849839
73	Almandine	7 pyralspite	219	349	503	917	1039	0.108577	0.413269	0.14476		1	0.0844556
74	Almandine	7 pyralspite	219	349	503	917	1036	0.110318	0.411138	0.14337		1	0.0876735
75	Almandine	7 pyralspite	219	349	503	917	1036	0.114559	0.414629	0.147663		1	0.0827751
76	Almandine	7 pyralspite	219	349	503	917	1036	0.115472	0.408094	0.151862		1	0.0825719
77	Andradite	2 ugrandite	244	375	536	880	1008	0.146011		0.171036	0.396197		0.53429
78	Andradite	2 ugrandite	242	372	530	877	998	0.171592		0.1793	0.505646		0.0509346
79	Andradite	2 ugrandite	244	375	538	880	998	0.173192		0.219045	0.511582		0.0598213
80	Andradite	2 ugrandite	242	369	538	877	1008	0.177257		0.286418	0.442277		0.0514923
81	Andradite	2 ugrandite	244	375	538	880	998	0.161252		0.489371	0.436346		0.061589
82	Andradite	2 ugrandite	244	372	533	880	998	0.178542		0.201804	0.462655		0.0504666
83	Andradite	2 ugrandite	244	375	536	880	1008	0.151266		0.191259	0.472006		0.413785
84	Andradite	2 ugrandite	244	375	538	880	998	0.164353		0.209564	0.488445		0.0604346
85	Andradite	2 ugrandite	244	375	538	883	998	0.163811		0.208775	0.561044		0.0583789
86	Andradite	2 ugrandite	244	375	538	883	1016	0.160004		0.209159	0.515104		0.0882793
87	Andradite	2 ugrandite	244	375	538	880	1016	0.167099		0.293687	0.483332		0.121038
88	Andradite	2 ugrandite	244	375	536	880	1000	0.169327		0.213768	0.507354		0.0583114
89	Andradite	2 ugrandite	244	375	538	880	1013	0.136178		0.191465	0.37571		0.702819
90	Andradite	2 ugrandite	244	375	538	880	1016	0.156873		0.412157	0.507463		0.11551
91	Andradite	2 ugrandite	244	372	536	880	998	0.166926		0.216343	0.462629		0.0537277
92	Andradite	3 ugrandite	239	372	519	877	998	0.157026		0.355844	0.850794		0.059935
93	Andradite	3 ugrandite	242	375	522	880	1000	0.146826		0.267522	0.788266		0.0651274
94	Andradite	3 ugrandite	242	375	522	880	1000	0.148568		0.322359	0.830325		0.0622194
95	Andradite	3 ugrandite	242	375	522	880	1000	0.154189		0.269867	0.764352		0.0626788
96	Andradite	3 ugrandite	242	375	522	880	1000	0.155322		0.320873	0.799442		0.0598183
97	Andradite	3 ugrandite	242	375	522	880	1000	0.149288		0.25287	0.751341		0.0656959
98	Andradite	3 ugrandite	242	375	522	880	1000	0.148697		0.315018	0.795025		0.0588479
99	Andradite	3 ugrandite	242	375	522	880	1000	0.147823		0.303814	0.792148		0.0615248
100	Andradite	3 ugrandite	242	375	522	880	1000	0.150673		0.345217	0.846983		0.0601567
101	Andradite	3 ugrandite	242	375	522	880	1000	0.146225		0.312454	0.811314		0.0604719
102	Andradite	3 ugrandite	242	375	519	880	1000	0.147629		0.366187	0.845168		0.0612463
103	Andradite	3 ugrandite	242	375	522	880	1000	0.14871		0.31578	0.812139		0.0627798
104	Andradite	3 ugrandite	242	375	522	880	1000	0.146993		0.33021	0.808778		0.0591792
105	Andradite	3 ugrandite	242	375	522	880	1000	0.145085		0.308465	0.792981		0.0596564
106	Andradite	3 ugrandite	242	375	522	880	1000	0.146642		0.33936	0.81438		0.0624899
107	Grossular	1 ugrandite	250	377	549	885	1008	0.172024		0.219486	0.839141		0.0824994
108	Grossular	1 ugrandite	250	377	549	885	1010	0.172989		0.193701	0.743375		0.0720737
109	Grossular	1 ugrandite	250	377	549	885	1008	0.178484		0.211204	0.816303		0.0715856
110	Grossular	1 ugrandite	250	377	549	885	1008	0.177596		0.205392	0.83573		0.076621
111	Grossular	1 ugrandite	250	377	549	885	1010	0.169334		0.200639	0.736887		0.0696785
112	Grossular	1 ugrandite	250	377	549	885	1010	0.164295		0.229669	0.83072		0.0784779
113	Grossular	1 ugrandite	250	377	549	885	1008	0.165241		0.185137	0.735606		0.0696013
114	Grossular	1 ugrandite	250	377	549	885	1008	0.172425		0.208923	0.847685		0.0806349
115	Grossular	1 ugrandite	250	377	549	885	1010	0.182566		0.210027	0.861025		0.0816679
116	Grossular	1 ugrandite	250	377	549	885	1008	0.168053		0.207675	0.857582		0.0790611
117	Grossular	1 ugrandite	250	377	549	885	1010	0.176748		0.220762	0.879909		0.0828994
118	Grossular	1 ugrandite	250	377	549	885	1010	0.177263		0.202783	0.800012		0.0752123

119	Grossular	1 ugrandite	250	377	549	885	1010	0,182021	1	0,205013	0,844214	0,0798541
120	Grossular	1 ugrandite	250	377	549	885	1013	0,175045	1	0,21689	0,874669	0,0794611
121	Grossular	1 ugrandite	250	377	549	885	1010	0,176904	1	0,211301	0,87453	0,0821285
122	Grossular	2 ugrandite	250	380	552	883	1010	0,211425	1	0,236669	0,913684	0,141791
123	Grossular	2 ugrandite	250	377	552	883	1013	0,193201	1	0,245417	0,933414	0,105212
124	Grossular	2 ugrandite	250	380	552	885	1010	0,186094	1	0,250242	0,969282	0,103061
125	Grossular	2 ugrandite	250	380	552	885	1010	0,214556	1	0,272502	0,958885	0,133125
126	Grossular	2 ugrandite	250	380	552	883	1008	0,215464	1	0,244661	0,936869	0,0997979
127	Grossular	2 ugrandite	250	380	549	883	1010	0,206046	1	0,259953	0,90095	0,101125
128	Grossular	2 ugrandite	250	380	552	883	1013	0,198584	1	0,251144	0,940518	0,095607
129	Grossular	2 ugrandite	250	380	549	883	1010	0,195209	1	0,287065	0,941706	0,103305
130	Grossular	2 ugrandite	250	380	552	883	1010	0,195294	1	0,252807	0,940062	0,0968463
131	Grossular	2 ugrandite	250	380	552	883	1010	0,198208	1	0,243788	0,957749	0,129342
132	Grossular	2 ugrandite	250	380	552	885	1005	0,207686	1	0,253334	0,965642	0,112781
133	Grossular	2 ugrandite	250	380	552	883	1008	0,206023	1	0,243834	0,936412	0,115948
134	Grossular	2 ugrandite	250	380	552	883	1010	0,188086	1	0,25417	0,900306	0,10248
135	Grossular	2 ugrandite	250	380	552	885	1010	0,192794	1	0,247815	0,975472	0,104071
136	Grossular	2 ugrandite	250	380	552	883	1008	0,201187	1	0,259811	0,909324	0,081546
137	Grossular	3 ugrandite	250	380	552	883	1016	0,203936	1	0,248692	0,906689	0,199204
138	Grossular	3 ugrandite	250	380	549	885	1016	0,190528	1	0,245279	0,903551	0,116712
139	Grossular	3 ugrandite	250	380	552	885	1005	0,181184	1	0,247216	0,903297	0,0875908
140	Grossular	3 ugrandite	250	377	549	883	1005	0,203856	1	0,226283	0,75768	0,100274
141	Grossular	3 ugrandite	250	380	552	885	1010	0,187913	1	0,247743	0,905399	0,100093
142	Grossular	3 ugrandite	250	380	552	883	1010	0,184047	1	0,2519	0,885586	0,0938067
143	Grossular	3 ugrandite	247	377	544	883	1010	0,174931	1	0,208293	0,589037	0,0660113
144	Grossular	3 ugrandite	250	380	552	885	1016	0,187961	1	0,234415	0,968039	0,205234
145	Grossular	3 ugrandite	250	380	552	885	1021	0,175205	1	0,24829	0,904826	0,416811
146	Grossular	3 ugrandite	250	380	552	885	1010	0,189318	1	0,424512	0,882669	0,104588
147	Grossular	3 ugrandite	250	380	552	883	1010	0,176601	1	0,23438	0,92528	0,097184
148	Grossular	3 ugrandite	250	380	552	885	1013	0,195405	1	0,235668	0,929163	0,112729
149	Grossular	3 ugrandite	250	380	552	885	1016	0,215028	1	0,264388	0,973043	0,213609
150	Grossular	3 ugrandite	250	380	552	885	1010	0,194391	1	0,237336	0,899072	0,108439
151	Grossular	3 ugrandite	250	380	552	885	1010	0,200134	1	0,236581	0,923248	0,134014
152	Grossular	4 ugrandite	250	377	552	885	1010	0,182568	1	0,224368	0,876168	0,0950048
153	Grossular	4 ugrandite	250	377	552	885	1010	0,161101	1	0,209579	0,891925	0,0835868
154	Grossular	4 ugrandite	250	377	552	885	1010	0,191401	1	0,217553	0,918857	0,0891189
155	Grossular	4 ugrandite	250	377	552	885	1010	0,177491	1	0,235163	0,904452	0,108045
156	Grossular	4 ugrandite	250	377	552	885	1010	0,179105	1	0,229673	0,962644	0,0898724
157	Grossular	4 ugrandite	250	377	552	885	1008	0,160228	1	0,222809	0,92008	0,0823638
158	Grossular	4 ugrandite	250	377	552	885	1010	0,173985	1	0,232867	0,960391	0,0904538
159	Grossular	4 ugrandite	250	377	552	885	1010	0,175956	1	0,219663	0,923008	0,0925254
160	Grossular	4 ugrandite	250	377	549	885	1010	0,181495	1	0,213688	0,837599	0,0773223
161	Grossular	4 ugrandite	250	377	552	885	1010	0,1724	1	0,219842	0,903298	0,0895224
162	Grossular	4 ugrandite	250	380	552	885	1013	0,174052	1	0,239265	0,974206	0,0907521
163	Grossular	4 ugrandite	250	377	552	885	1010	0,179009	1	0,22637	0,905168	0,0858426
164	Grossular	4 ugrandite	250	377	552	885	1010	0,178881	1	0,23496	0,951479	0,0955692
165	Grossular	4 ugrandite	250	377	552	885	1010	0,178173	1	0,227915	0,94643	0,0916141
166	Grossular	4 ugrandite	250	377	552	885	1010	0,172258	1	0,215626	0,907338	0,0881534
167	Hessonite	1 ugrandite	250	377	552	885	1010	0,19459	1	0,224242	0,91697	0,103231
168	Hessonite	1 ugrandite	250	377	549	885	1010	0,185802	1	0,345447	0,90658	0,0972346
169	Hessonite	1 ugrandite	250	377	552	885	1008	0,181162	1	0,214838	0,877172	0,0855687
170	Hessonite	1 ugrandite	250	377	552	885	1010	0,18835	1	0,226047	0,904543	0,103151
171	Hessonite	1 ugrandite	250	377	552	885	1010	0,180969	1	0,222309	0,890985	0,0865564
172	Hessonite	1 ugrandite	250	377	552	885	1010	0,212573	1	0,240841	0,938193	0,0970386
173	Hessonite	1 ugrandite	250	377	552	885	1010	0,186911	1	0,231448	0,921073	0,0910117
174	Hessonite	1 ugrandite	250	377	552	885	1010	0,171323	1	0,228153	0,909791	0,0907814
175	Hessonite	1 ugrandite	250	377	549	885	1008	0,17957	1	0,210867	0,880318	0,0747421
176	Hessonite	1 ugrandite	250	377	552	885	1010	0,189377	1	0,222074	0,919994	0,0860474
177	Hessonite	1 ugrandite	250	377	549	885	1010	0,171395	1	0,211148	0,897324	0,076828

178	Hessonite	1 ugrandite	250	377	549	885	1010	0,209173	1	0,217305	0,878684	0,0832024
179	Hessonite	1 ugrandite	250	377	549	885	1010	0,178464	1	0,207687	0,867433	0,0798925
180	Hessonite	1 ugrandite	250	377	549	885	1010	0,177053	1	0,220271	0,911889	0,0861772
181	Hessonite	1 ugrandite	250	377	549	885	1008	0,179299	1	0,215599	0,887894	0,079698
182	Hessonite	2 ugrandite	250	377	552	885	1010	0,167029	0,954777	0,195973	1	0,0921906
183	Hessonite	2 ugrandite	250	377	552	885	1010	0,171426	0,909559	0,188197	1	0,113442
184	Hessonite	2 ugrandite	250	377	552	885	1010	0,156386	0,879897	0,20553	1	0,0901327
185	Hessonite	2 ugrandite	250	375	552	885	1010	0,18673	0,992185	0,205231	1	0,100856
186	Hessonite	2 ugrandite	250	377	552	885	1010	0,167416	0,919161	0,197723	1	0,0956923
187	Hessonite	2 ugrandite	250	377	552	885	1013	0,167464	0,914394	0,201057	1	0,091144
188	Hessonite	2 ugrandite	250	377	555	885	1010	0,143705	0,858037	0,197719	1	0,088837
189	Hessonite	2 ugrandite	250	377	552	885	1010	0,168078	0,893749	0,199479	1	0,0942358
190	Hessonite	2 ugrandite	250	377	552	885	1010	0,161474	0,875253	0,209749	1	0,10498
191	Hessonite	2 ugrandite	250	377	555	885	1010	0,157495	0,883389	0,203883	1	0,0987879
192	Hessonite	2 ugrandite	250	377	555	885	1005	0,155633	0,827298	0,196048	1	0,114006
193	Hessonite	2 ugrandite	250	377	552	885	1010	0,172959	0,921997	0,201864	1	0,0952238
194	Hessonite	2 ugrandite	250	377	552	885	1010	0,169603	0,926074	0,203542	1	0,0930906
195	Hessonite	2 ugrandite	250	377	552	885	1010	0,150582	0,89141	0,193283	1	0,0983411
196	Hessonite	2 ugrandite	250	377	555	885	1010	0,16396	0,875653	0,196976	1	0,0968172
197	Melanite	1 ugrandite	233	358	519	870	972	0,108314	1	0,239637	0,278177	0,0671907
198	Melanite	1 ugrandite	236	358	522	867	987	0,0923472	1	0,248412	0,25958	0,0698151
199	Melanite	1 ugrandite	239	363	527	877	985	0,117072	1	0,261038	0,282538	0,0874607
200	Melanite	1 ugrandite	236	363	525	880	982	0,111979	1	0,29046	0,292913	0,0913764
201	Melanite	1 ugrandite	233	358	519	872	992	0,119957	1	0,253302	0,270711	0,105285
202	Melanite	1 ugrandite	239	363	522	872	982	0,0923999	1	0,273483	0,301162	0,0980538
203	Melanite	1 ugrandite	236	358	522	867	979	0,105089	1	0,239628	0,262104	0,07222771
204	Melanite	1 ugrandite	233	355	516	870	979	0,0949276	1	0,234887	0,257556	0,0705162
205	Melanite	1 ugrandite	233	358	522	867	979	0,096563	1	0,248393	0,260698	0,0873334
206	Melanite	1 ugrandite	236	363	522	870	979	0,0951141	1	0,26967	0,272626	0,0796361
207	Melanite	1 ugrandite	233	355	519	864	985	0,0990253	1	0,219536	0,2391	0,0641806
208	Melanite	1 ugrandite	233	355	519	867	979	0,10876	1	0,224413	0,260042	0,0695247
209	Melanite	1 ugrandite	233	355	516	864	974	0,101344	1	0,224082	0,239333	0,0670388
210	Melanite	1 ugrandite	233	358	522	867	985	0,101358	1	0,245645	0,274624	0,0689523
211	Melanite	1 ugrandite	233	355	519	867	966	0,0979199	1	0,216704	0,235631	0,068875
212	Polyadelph	1 ugrandite	233	363	549	898	1010	0,199907	1	0,373996	0,954338	0,159923
213	Polyadelph	1 ugrandite	233	363	549	898	1005	0,206815	0,974714	0,368389	1	0,174103
214	Polyadelph	1 ugrandite	233	363	549	901	1005	0,207336	0,931817	0,34409	1	0,16767
215	Polyadelph	1 ugrandite	230	363	549	901	1016	0,187259	0,834773	0,33897	1	0,246908
216	Polyadelph	1 ugrandite	233	363	549	901	1003	0,162147	0,842131	0,332558	1	0,28963
217	Polyadelph	1 ugrandite	233	363	549	901	1010	0,179886	0,892004	0,356512	1	0,160279
218	Polyadelph	1 ugrandite	233	361	549	898	1016	0,21093	1	0,343563	0,941207	0,179597
219	Polyadelph	1 ugrandite	230	361	538	896	998	0,182552	1	0,714181	0,82351	0,627182
220	Polyadelph	1 ugrandite	227	361	549	901	1016	0,243167	0,918392	0,366437	1	0,217596
221	Polyadelph	1 ugrandite	233	361	549	901	1005	0,207699	0,989444	0,347646	1	0,171191
222	Polyadelph	1 ugrandite	227	361	549	901	1016	0,215514	1	0,356196	0,957505	0,176778
223	Polyadelph	1 ugrandite	227	358	547	896	1005	0,210371	1	0,294552	0,766353	0,129163
224	Polyadelph	1 ugrandite	233	361	549	898	1016	0,20874	1	0,342922	0,966637	0,152138
225	Polyadelph	1 ugrandite	233	363	549	898	1016	0,200203	1	0,366907	0,940973	0,162811
226	Polyadelph	1 ugrandite	222	358	549	896	998	0,228127	1	0,325634	0,891303	0,35477
227	Pyrope	1 pyralspite	210,02	362,45	558,37	919,06	1052,44	0,04	0,376	0,283	1	0,063
228	Pyrope	1 pyralspite	207,29	362,8	559,27	920,21	1053,96	0,044	0,391	0,281	1	0,064
229	Pyrope	2 pyralspite	209,85	358,96	559,91	920,21	1052,04	0,035	0,331	0,188	1	0,064
230	Pyrope	2 pyralspite	209,21	358,96	559,91	920,21	1052,04	0,044	0,322	0,187	1	0,083
231	Rhodolite	1 pyralspite	199	352	469	919	1044	0,273565	0,37851	1	0,850781	0,0877419
232	Rhodolite	1 pyralspite	216	349	503	917	1039	0,0927789	0,420007	0,141174	1	0,089299
233	Rhodolite	1 pyralspite	216	352	467	919	1039	0,111299	0,409211	0,259125	1	0,0867514
234	Rhodolite	1 pyralspite	199	352	469	919	1047	0,262571	0,398983	0,58552	1	0,0993277
235	Rhodolite	1 pyralspite	216	352	503	917	1039	0,10592	0,42561	0,138528	1	0,0943496
236	Rhodolite	1 pyralspite	187	349	464	914	1036	0,302272	0,523852	0,354826	1	0,0883832

237	Rhodolite	1 pyralspite	216	352	505	919	1039	0.090634	0.402387	0.145658		1	0.0931885
238	Rhodolite	1 pyralspite	210	349	467	912	1026	0.196949	0.524727	0.238952		1	0.103964
239	Rhodolite	1 pyralspite	196	352	467	917	1039	0.403852	0.479243		1	0.960302	0.11247
240	Rhodolite	1 pyralspite	170	349	505	917	1039	0.1047	0.397298	0.136117		1	0.0853455
241	Rhodolite	1 pyralspite	170	349	503	917	1036	0.135563	0.484944	0.218022		1	0.0944045
242	Rhodolite	1 pyralspite	213	349	500	914	1031	0.197498	0.525491	0.161908		1	0.101727
243	Rhodolite	1 pyralspite	193	352	503	917	1039	0.282137	0.462557	0.241508		1	0.108624
244	Rhodolite	1 pyralspite	170	349	503	914	1036	0.145319	0.454531	0.159533		1	0.0895073
245	Rhodolite	1 pyralspite	187	352	464	914	1036	0.340147	0.58251	0.746624		1	0.112621
246	Schorlomite	1 ugrandite	233	366	516	875	972	0.0941111	0.979778	0.426444	0.484111		0.132222
247	Schorlomite	1 ugrandite	232	361	514	873	929	0.167286	0.995	0.489	0.597143		0.299
248	Schorlomite	2 ugrandite	231	335	511	860	1069	0.246333	0.557878	0.312328	0.422444		0.0330847
249	Schorlomite	2 ugrandite	233	335	511	861	1060	0.252412	0.584	0.319235	0.409294		0.0263529
250	Spessartine	2 pyralspite	222	352	505	912	1031	0.149122	0.344281	0.115569		1	0.112503
251	Spessartine	2 pyralspite	224	352	505	912	1034	0.119132	0.319315	0.111128		1	0.102613
252	Spessartine	2 pyralspite	221	348	499	908	1029,14	0.09	0.325	0.076		1	0.069
253	Spessartine	2 pyralspite	224	352	505	912	1031	0.118971	0.304315	0.105694		1	0.100403
254	Spessartine	2 pyralspite	224	352	503	912	1031	0.124938	0.32701	0.118064		1	0.116189
255	Spessartine	2 pyralspite	222	352	505	912	1031	0.13522	0.329298	0.115311		1	0.103299
256	Spessartine	2 pyralspite	222	352	505	912	1036	0.127036	0.316176	0.11317		1	0.106653
257	Spessartine	2 pyralspite	222	352	505	912	1036	0.123804	0.306679	0.10063		1	0.102952
258	Spessartine	2 pyralspite	224	352	505	914	1036	0.123886	0.308839	0.114624		1	0.11242
259	Spessartine	2 pyralspite	224	352	505	912	1036	0.113071	0.292586	0.0986407		1	0.0885898
260	Spessartine	2 pyralspite	224	355	505	914	1031	0.118067	0.296083	0.114004		1	0.0979889
261	Spessartine	2 pyralspite	222	352	505	912	1036	0.126769	0.326854	0.118106		1	0.10824
262	Spessartine	2 pyralspite	224	352	505	912	1036	0.118425	0.300771	0.0923699		1	0.0882724
263	Spessartine	2 pyralspite	222	352	508	914	1031	0.150879	0.306882	0.120725		1	0.116181
264	Spessartine	2 pyralspite	222	352	505	914	1031	0.14271	0.318146	0.134652		1	0.139853
265	Spessartine	3 pyralspite	247	375	549	827	1008	0.177865	1	0.201518	0.380391		0.0823836
266	Spessartine	3 pyralspite	250	375	549	827	1010	0.184344	1	0.19577	0.364131		0.0865214
267	Spessartine	3 pyralspite	247	377	549	830	1010	0.181782	1	0.223565	0.407644		0.0887674
268	Spessartine	3 pyralspite	250	377	549	830	1010	0.181956	1	0.204121	0.385997		0.088431
269	Spessartine	3 pyralspite	250	377	549	830	1010	0.179391	1	0.204914	0.389088		0.0835178
270	Spessartine	3 pyralspite	247	375	549	827	1008	0.174806	1	0.206477	0.396176		0.0757978
271	Spessartine	3 pyralspite	250	377	552	830	1010	0.177918	1	0.223642	0.409695		0.089208
272	Spessartine	3 pyralspite	247	375	549	830	1008	0.178237	1	0.278048	0.380902		0.0878897
273	Spessartine	3 pyralspite	247	375	549	827	1008	0.175569	1	0.200985	0.383371		0.0822639
274	Spessartine	3 pyralspite	250	375	549	827	1005	0.181747	1	0.208458	0.385149		0.0864412
275	Spessartine	3 pyralspite	250	375	549	830	1010	0.187022	1	0.212736	0.385151		0.0883772
276	Spessartine	3 pyralspite	247	375	549	827	1008	0.173396	1	0.20088	0.361046		0.0892524
277	Spessartine	3 pyralspite	250	375	549	830	1010	0.177388	1	0.20281	0.388061		0.0852268
278	Spessartine	3 pyralspite	250	375	549	830	1005	0.181313	1	0.205646	0.378955		0.0855213
279	Spessartine	3 pyralspite	250	375	549	830	1010	0.179941	1	0.209236	0.392091		0.0850977
280	Spessartine	4 pyralspite	213	352	505	919	1041	0.111238	0.425856	0.147635		1	0.102411
281	Spessartine	4 pyralspite	213	349	500	914	1034	0.0978969	0.451002	0.150926		1	0.0797143
282	Spessartine	4 pyralspite	216	352	505	919	1041	0.0889489	0.417713	0.138031		1	0.0927117
283	Spessartine	4 pyralspite	219	352	505	919	1044	0.0706294	0.38929	0.16188		1	0.0832987
284	Spessartine	4 pyralspite	219	352	505	919	1044	0.0834574	0.390384	0.133371		1	0.0803612
285	Spessartine	4 pyralspite	216	352	505	919	1044	0.056384	0.383706	0.120685		1	0.077995
286	Spessartine	4 pyralspite	216	352	503	917	1039	0.0923257	0.419264	0.14394		1	0.0871509
287	Spessartine	4 pyralspite	216	352	505	919	1039	0.0725138	0.395467	0.131845		1	0.0792914
288	Spessartine	4 pyralspite	213	349	503	914	1036	0.10037	0.456512	0.148509		1	0.0842993
289	Spessartine	4 pyralspite	216	352	503	919	1041	0.0722089	0.415073	0.134594		1	0.080854
290	Spessartine	4 pyralspite	213	352	503	917	1039	0.0789532	0.403266	0.135351		1	0.0836599
291	Spessartine	4 pyralspite	216	352	505	919	1044	0.0710332	0.393909	0.129576		1	0.0788661
292	Spessartine	4 pyralspite	216	352	505	917	1039	0.0848013	0.4277	0.137363		1	0.0870681
293	Spessartine	4 pyralspite	216	352	505	919	1044	0.0708384	0.385151	0.135627		1	0.0814892
294	Spessartine	4 pyralspite	213	352	503	917	1036	0.0794566	0.413851	0.140351		1	0.0853652
295	Uvarovite	1 ugrandite	243,13	371,76	533,03	882,45	1003,4	0,089		1	0,352	0,329	0,026
296	Uvarovite	2 ugrandite	242,04	370,16	528,46	888,43	994,88	0,071	0,96	0,33	0,239		0,015
297	Uvarovite	2 ugrandite	243,13	370,48	529,83	892,69	1001,48	0,071	0,96	0,33	0,239		0,015
298	Uvarovite	3 ugrandite	251,01	366,91	517,85	889,11	996,49	0,087	0,965	0,363	0,445		0,037
299	Uvarovite	3 ugrandite	239,29	367,92	519,59	890,13	998,28	0,064	0,947	0,356	0,368		0,027

### A3. Tabla de datos de parámetros de bandas para test

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	mineral	sample	type	Peak 1	Peak 2	Peak 3	Peak 4	Peak 5	Peak 1	Peak 2	Peak 3	Peak 4	Peak 5
2	Almandine	1	pyrralspite	215.733	356.733	559.733	919.733	1048,73	0,0151879	0,228465	0,110078	1	0,0705052
3	Almandine	2	pyrralspite	216.387	351.387	559.387	918.387	1042,39	0,0206966	0,277076	0,0683594	1	0,0648895
4	Almandine	3	pyrralspite	214.423	361.423	563.423	921.423	1049,42	0,0151079	0,248433	0,100132	1	0,055005
5	Almandine	4	pyrralspite	223.994	352.994	558.994	915.994	1037,99	0,0298244	0,244613	0,0732824	1	0,0477485
6	Almandine	5	pyrralspite	276,543	359,17	558,17	913,17	1041,17	0,0116273	0,265504	0,11678	1	0,0674669
7	Andradite	2	ugrandite	233,96	364,96	519,96	871,96	977,96	0,0406907	1	0,472856	0,397147	0,0926187
8	Andradite	3	ugrandite	239,96	372,96	526,96	881,96	1000,96	0,0771876	1	0,31613	0,644839	0,0773943
9	Andradite	4	ugrandite	237,06	372,06	530,06	886,06	1002,06	0,073444	1	0,393004	0,662528	0,0776869
10	Andradite	5	ugrandite	234,186	368,186	514,186	873,186	994,186	0,102963	1	0,436382	0,673744	0,0594177
11	Grossular	1	ugrandite	244,86	372,86	547,86	881,86	1005,86	0,0584482	0,653047	0,152341	1	0,0751874
12	Grossular	2	ugrandite	238,86	370,86	532,86	875,86	989,86	0,111886	1	0,264455	0,542946	0,100283
13	Grossular	3	ugrandite	246,96	375,96	547,96	881,96	1006,96	0,0931021	0,793701	0,187428	1	0,0902107
14	Grossular	4	ugrandite	242,922	375,922	536,922	877,922	989,922	0,0870552	1	0,276707	0,41759	0,111597
15	Grossular	5	ugrandite	243,9	372,9	544,9	879,9	1005,9	0,0632818	0,784053	0,197594	1	0,0728361
16	Pyrope	1	pyrralspite	206,86	353,86	557,86	919,86	1049,86	0,0239338	0,226749	0,102287	1	0,0819955
17	Pyrope	2	pyrralspite	206,558	360,558	559,558	919,558	1051,56	0,0213534	0,228108	0,139024	1	0,0895524
18	Pyrope	3	pyrralspite	210,143	343,143	554,143	913,143	1037,14	0,0252795	0,262621	0,0837197	1	0,0797634
19	Pyrope	4	pyrralspite	209,645	353,645	554,645	915,645	1045,64	0,0160897	0,247261	0,122934	1	0,059442
20	Pyrope	5	pyrralspite	208,86	359,86	557,86	918,86	1050,86	0,0208971	0,24713	0,147442	1	0,0850074
21	Schorlomite	1	ugrandite	165,558	356,558	518,558	872,558	965,558	0,147198	1	0,434855	0,427754	0,177541
22	Spessartine	1	pyrralspite	214,23	359,17	558,17	913,17	1041,17	0,011293	0,265504	0,11678	1	0,0674669
23	Spessartine	2	pyrralspite	222,771	351,771	554,771	909,771	1031,77	0,0434866	0,197196	0,0840014	1	0,0569935
24	Spessartine	3	pyrralspite	228,771	359,771	552,771	903,771	1025,77	0,047854	0,309364	0,100526	1	0,0920502
25	Spessartine	4	pyrralspite	158,771	352,771	544,771	908,771	1024,77	0,0889864	0,096564	0,0821056	1	0,0777243
26	Spessartine	5	pyrralspite	226,138	349,138	550,138	911,138	1030,14	0,0499506	0,139043	0,0968973	1	0,0855242
27	Uvarovite	1	ugrandite	272,898	371,898	532,898	880,898	1005,9	0,0574846	1	0,364214	0,333339	0,0300519
28	Uvarovite	2	ugrandite	244,771	371,771	534,771	884,771	1005,77	0,07484	1	0,334462	0,584987	0,0420563