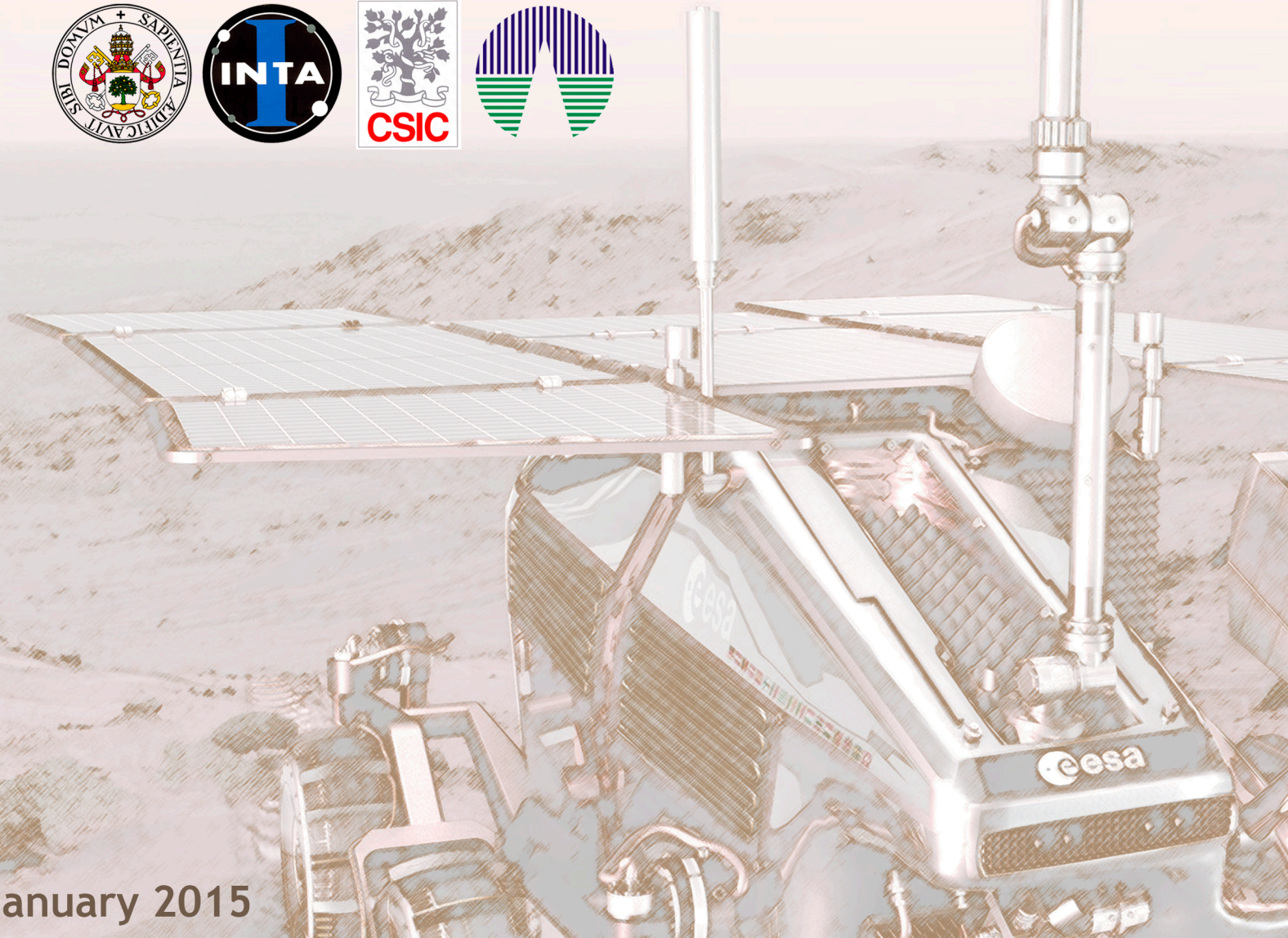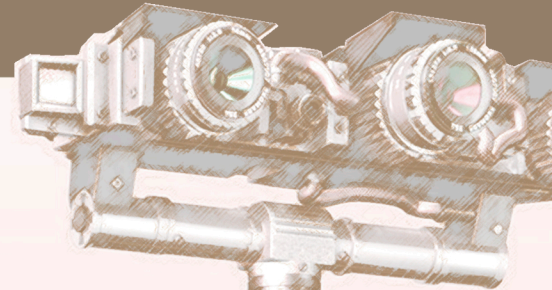PhD Thesis

Development of algorithms and methodological analyses for the definition of the operation mode of the Raman Laser Spectrometer instrument

Guillermo Lopez-Reyes

January 2015

Universidad de Valladolid

FACULTAD DE CIENCIAS

DEPARTAMENTO DE FÍSICA DE LA MATERIA CONDENSADA

TESIS DOCTORAL:

Development of algorithms and methodological analyses for the definition of the operation mode of the Raman Laser Spectrometer instrument

Presentada por Guillermo E. López Reyes para optar al grado de doctor por la Universidad de Valladolid

Dirigida por:
Fernando Rull Pérez

ACKNOWLEDGEMENTS

INDEX

Epilogue

Scope and structure

In situ exploration of Mars, in collaboration with remote analysis of the red planet, is key to understanding its geologic evolution and establishing whether water and life once existed (or still exist) on the Martian surface. The Raman Laser Spectrometer (RLS) instrument, part of the payload of ESA ExoMars mission, will be the first Raman spectrometer ever to be sent on an interplanetary mission.

The potential of Raman spectroscopy as a non-invasive or destructive procedure applied to all kinds of materials defines it as a very important technique for all types of materials, minerals or biological analyses. This spectroscopic technique has remained, however, unprecedented in planetary exploration, so it has been traditionally operated with the intervention of a human operator.

The operation mode of the RLS instrument is understood as the series of operations and algorithms that the instrument needs to perform autonomously for obtaining a Raman spectrum of the highest quality. This must not be confused with the development of the algorithms necessary for the instrument to become a flight instrument (telemetry, communication with the rover, data conversion, etc.). The definition of the operation mode is a purely scientific task, which shifts the actions that the human operator performs with the instrument (laser focusing on the sample, adjustment of the spectrum's acquisition time, elimination of fluorescence…) to a system able to autonomously decide which is the point of focus, to detect whether the sample under study is fluorescent or not and therefore it requires a special treatment, what will the optimal integration time be, etc.

The definition of all these algorithms is based on scientific experiments that help to define objective parameters which are measurable by the instrument, to enable it to autonomously carry out those tasks the human operator manually performs based on his experience.

For the development and optimization of these algorithmic tools, it was necessary to develop an adequate setup. The aim was to provide the science team with a platform in which to perform the experiments in a rover-like scenario, taking into account how the RLS instrument will interact with the rover, setting the basis for the development of the operation mode.

The so-called RLS ExoMars Simulator that we have developed at Unidad Asociada UVa-CSIC-CAB [1, 2] (Figure 0-1) has enabled to perform many scientific experiments under conditions similar to those provided by the Analytic Laboratory Drawer (ALD) of the ExoMars rover, including the possibility of analyzing samples both in laboratory conditions and Martian conditions, with the Martian Simulation Chamber (whose description, design and results are described in detail in a different thesis project). Thanks to it, it has been possible to define how to acquire spectra autonomously,

without the presence of a human operator. So far, some sophisticated commercial Raman equipment were able to work in a pseudo-automatic way [3], by performing a series of measurements at different points in a sample, though treating all points equally. The use of the RLS ExoMars simulator has allowed the definition of an automatic acquisition system that optimizes the parameters in each one of the points, thereby optimizing the use of the available time and resources.
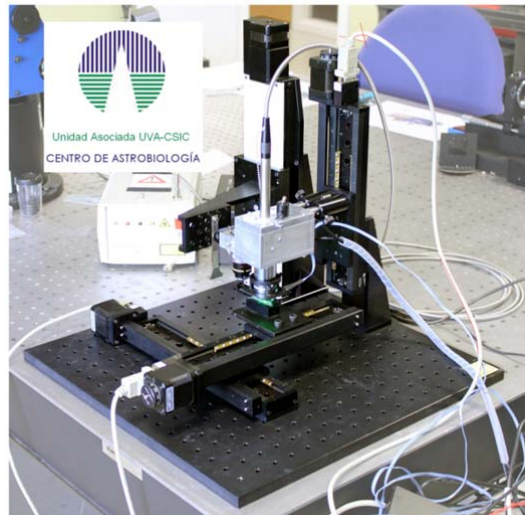
Figure 0-1. RLS ExoMars simulator prototype

On the other hand, the use of this system has allowed performing experiments that have helped understand the possibilities offered within the framework of planetary exploration, in order to maximize the scientific return of the resulting products from the operation of the instrument on Mars. In particular, the development of the RLS ExoMars simulator has paved the way for the systematic study of the operation on powdered samples, clearing the path for the future development of analytical tools.

These tools will form part of the Instrument Data Analysis Tool of RLS (IDAT), a ground segment system that will be used during the operational phase of the mission, not only for the instrument health assessment, but also for short (and long) term science exploitation. Short term use of IDAT will, on a daily basis, provide the scientific team operating the rover with very fast scientific results in order to take tactical decisions for the following rover activities. Having this information analyzed after a very short time (minutes) after the reception of the data from the spacecraft is of paramount importance to take the best tactical course of action. However, this kind of operation also imposes very strict requirements regarding the robustness and reliability of the analytical methods applied for the analyses. Furthermore, it also requires the development and population of databases, and the implementation of routines for the automated analysis of the results. Of course, all these tools will also be usable by the scientific team for the detailed (and slower) analysis of the data, providing support for the long-term science exploitation of the RLS data products.

In this framework, the present thesis intended to evaluate the scientific capabilities of the RLS instrument, with two main objectives.

First, the automation of the instrument and the optimization of the spectra acquisition. To do so, it is necessary to define novel algorithmic procedures which provide the instrument with autonomous decision logic in order to optimize the available resources while increasing the quality and quantity of acquired spectra.

Second, the data exploitation from the RLS instrument. To accomplish this objective, the effects on the acquired data of the engineering constraints imposed by space operation are considered. Then, different analytical techniques for the analysis of the final spectra acquired by the instrument are studied and applied. This aims at identifying the samples, but also at quantifying their abundance in mixtures, and assessing the capabilities of this kind of techniques.

These two objectives are connected and related to the RLS instrument development based on the following logic: first, it is necessary to prepare the RLS instrument for automated operation and spectra acquisition optimization. Then, the operational constraints imposed by the space / rover operation are taken into account and studied to optimize the overall acquisition process. Finally, the data obtained with the RLS instrument (under its operational constraints) are studied with different techniques. This is performed in order to assess and evaluate the possibilities of data exploitation, including statistical analyses, which can be performed due to the particular operational circumstances of RLS: automatic mapping on powdered samples of the instrument

This work has been divided into two parts, related to its two main objectives.

The first part (chapters 2 and 3) describes the development process to create a laboratory prototype that allows performing experiments in a representative scenario from the operating point of view: sample flattening and positioning, spectra acquisition, repeatability and automation, plus imaging capabilities to provide better inputs to the scientific team. This prototype has been developed both from hardware and software points of view, including the automation and optimization of the instrument operational sequences.

The second part (chapters 4 and 5) is centered on the analysis of the data obtained in a scenario constrained by the RLS operation framework. This includes studying the consequences of the requirements imposed by the engineering constraints, in addition to applying analytical techniques for the identification and quantification of the samples that help obtain the maximum science return from the instrument.

Guillermo Lopez-Reyes

According to these objectives, this thesis project has been structured in 6 chapters:

Introduction

- **Chapter 1:** An introduction to Raman spectroscopy and the exploration of Mars, including a description of the ExoMars mission and the RLS instrument onboard the ExoMars Rover, in which this thesis is framed.

Part I

- **Chapter 2:** A description of the RLS ExoMars Simulator system, including its hardware and software capabilities in a user manual fashion.
- **Chapter 3:** Description of the operational framework imposed by engineering constraints of the project, as well as the detailed definition of the algorithmic necessary to operate the instrument in the most efficient way.

Part II

- **Chapter 4:** This chapter describes the consequences of the sample preparation and distribution procedure onboard the ExoMars rover (crushing and flattening issues). Furthermore, we present the spectral differences on Raman spectra and potential science return of the instrument after analyses on powdered samples vs. bulk samples.
- **Chapter 5:** In this chapter, several analytical methods are presented to assess their ability to perform qualitative and quantitative identification of mineral phases in mixtures from multiple Raman spectra acquired with the RLS instrument.

Epilogue

- **Chapter 6:** This chapter summarizes the thesis, and outlines the future research that should follow this work.

Introduction

Chapter 1.

RAMAN SPECTROSCOPY AND THE EXPLORATION OF MARS

Introduction to Raman spectroscopy and the ExoMars mission

1.1 THE AURORA PROGRAMME OF THE EUROPEAN SPACE AGENCY (ESA)

The Aurora programme [4] is part of Europe's strategy for space, endorsed by the European Union Council of Research and the ESA Council in 2001. The primary objective of this programme is to create, and then implement, a European long-term plan for the robotic and human exploration of the solar system, with Mars, the Moon and the asteroids as the most likely targets. A second objective is to search for life beyond the Earth. Future missions under the program will carry sophisticated exobiology payloads to investigate the possibility of life forms existing on other worlds within the solar system.

It is clear from these objectives that the interdependence of exploration and technology is the basis of the Aurora Program. On the one hand the desire to explore provides the stimulus to develop new technology while on the other, it is the introduction of innovative technology that will make exploration possible.

1.2 THE EXOMARS PROGRAMME 2016-2018

Establishing if life ever existed on Mars is one of the outstanding scientific questions of our time. To address this important goal, the European Space Agency (ESA) has established the ExoMars programme in collaboration with the Russian Space Agency (Roscosmos), and contributions from NASA. This programme will investigate the Martian environment and demonstrate new technologies paving the way for a future Mars sample return mission in the 2020's [5]. This program fits into the international roadmap for the exploration of Mars, shown in Figure 1-1. Currently, the red planet has several satellites and orbiting instruments, as well as ground vehicles. The Mars Exploration Rovers (MERs) were sent to Mars to look for evidences of water. Mars Science Laboratory (MSL) looks for past or present evidences of habitability, i.e., sites in which life could exist or have existed. Future missions will look for evidences of past or present life.

Figure 1-1. Past, present and planned missions to the red planet. Credits: ESA.

Two missions are foreseen within the ExoMars program: one consisting of an Orbiter plus an Entry, Descent and Landing Demonstrator Module, to be launched in 2016. The other featuring a rover with a launch date in 2018.

The objectives of this program can be subdivided in technological and scientific aspects [1]. In the technological plane, the ExoMars program will demonstrate a number of essential flight and *in situ* enabling technologies that are necessary for future exploration missions, such as an international Mars Sample Return mission. These include (extracted from [5]):

- Entry, descent and landing (EDL) of a payload on the surface of Mars;
- Surface mobility with a rover;
- Access to the subsurface to acquire samples; and
- Sample acquisition, preparation, distribution and analysis.

At the same time a number of important scientific investigations will be carried out, for example [5]:

- Search for signs of past and present life on Mars;
- Investigate how the water and geochemical environment varies;
- Investigate Martian atmospheric trace gases and their sources.

2016 mission has two different supplementary objectives, from the technological and scientific points of view. It will include an Entry, Descent and Landing Demonstrator Module (EDM), plus a Trace Gas Orbiter (TGO). The Orbiter will carry scientific instruments to detect and study atmospheric trace gases, such as methane. The EDM

will contain sensors to evaluate the lander performance as it descends, and additional sensors to study the environment at the landing site.

For the 2018 mission, ESA is developing a Mars rover carrying a comprehensive suite of analytical instruments dedicated to exobiology and geochemistry research named after Pasteur. The Rover will travel searching for traces of past and present signs of life, collecting and analyzing samples from within rocky outcrops and from the subsurface, down to a depth of 2m with the help of a drill (Figure 1-2).

An effective chemical identification of biomarkers requires access to well-preserved organic molecules. Because the Martian atmosphere is more tenuous than Earth's, three important physical agents reach the surface of Mars with adverse effects for the long-term preservation of biomarkers [6]:

- The ultraviolet (UV) radiation dose is higher than on our planet and will quickly damage potential exposed organisms or biomolecules in the shallow subsurface (depth: ~ 1 mm).
- UV-induced photochemistry is responsible for the production of reactive oxidant species that, when activated, can also destroy biomarkers; the diffusion of oxidants into the subsurface is not well characterized and constitutes an important measurement that the mission must perform (depth: ~ 1m).
- Ionizing radiation penetrates into the uppermost meters of the planet's subsurface (depth: ~ 1.5 m).

Figure 1-2. ExoMars rover on the Mars surface.

All these factors cause a slow degradation process that, over many millions of years, can alter organic molecules beyond the detection sensitivity of analytical instruments.

Thus, the rover's surface mobility and the 2m vertical reach of the drill are both crucial for the scientific success of the mission and for the preparation of the Mars Sample Return (MSR) mission, as it will be able to reach potentially non-radiated samples.

1.3 THE EXOMARS ROVER INSTRUMENT SUITE

Table 1-1 summarizes the instruments forming part of the current configuration of the Pasteur Payload (PPL), name given to the instrument suite that will be included in the ExoMars 2018 mission rover.

Instrument	Description	Responsible country	Characteristics
PanCam	Wide-angle stereo camera (WAC) High-resolution camera (HRC) *Geological context Rover traverse planning Atmospheric studies*	ESA UK	WAC: 35º Field of View (FoV) HRC: 5º FoV
ISEM	IR spectrometer on mast – IR spectrometer for ExoMars (ISEM) *Bulk mineralogy of outcrops Target selection*	Roscosmos Russia	λ = 1.15–3.3µm 1º FoV
WISDOM	Ground penetrating radar – Water Ice and Subsurface Deposit Observation on Mars (WISDOM) *Mapping of subsurface stratigraphy*	ESA France	Penetration: 5m Resolution: 2cm
ADRON-RM	Neutron spectrometer *Mapping of subsurface water and hydrated minerals*	Roscosmos Russia	--
Drill + Ma_MISS	IR borehole spectrometer *In-situ mineralogy information*	ESA Italy	λ = 0.4–2.2µm
CLUPI	Close-up Imager (CLUPI) *Geological environment Microtexture of rocks Morphological biomarkers*	ESA Swizterland	20µm at 50cm Focus: 20cm-∞
MicrOmega	VIS-IR Spectrometer *Mineralogical characterization Pointing for other instruments*	ESA France	λ = 0.5–3.5µm 5x5mm image 265x256 pixels 20µm per pixel
RLS	Raman Laser Spectrometer *Geochemical composition Detection of organic pigments*	ESA Spain	200-3800 cm^{-1} Min resolution: <6 cm^{-1}
MOMA	Laser Desorption Extraction and Mass Spectrometry (LMDS) Pyrolisis extraction, gas chromatography, mass spectrometry (Pyr-Dev GCMS) *Broad range organic molecules in ppb Chirality determination*	ESA Germany France NASA	--

Table 1-1. Current configuration of the Pasteur Payload [7]. The blue background shows the instruments inside the rover analytical laboratory drawer (ALD).

Below, a short description of the instruments is provided [5, 8]:

The **PanCam** - panoramic camera - has been designed to search for textural information on rocks (for example, laminations or pitting) that can be related to the presence of organisms on Mars. PanCam will also capture other information that will help in revealing the geological characteristics of the Martian environment.This instrument will provide stereo and 3D imagery of the terrain around the Rover, for the benefit of the mission as a whole.

ISEM is the Infrared Spectrometer for ExoMars. This instrument, located in the rover mast, is designed to assess the mineralogical composition of surface targets. Working with PanCam, ISEM will contribute to the selection of suitable samples for further analysis by the other instruments.

WISDOM (Water Ice Subsurface Deposit Observation on Mars) is a radar-based instrument that will provide a detailed view of the Red Planet's shallow subsurface structure by sounding the upper layers of its crust. It will study the nature of the subsurface remotely, using radar pulses from a UHF ground penetrating radar, covering the frequency range from 500 MHz to 3 GHz, to map the subterranean layers.

ADRON-RM is a neutron spectrometer that will be used to search for subsurface water and hydrated minerals. It will be used in combination with WISDOM to study the subsurface beneath the rover and to search for suitable areas for drilling and sample collection.

Ma_MISS (Mars Multispectral Imager for Subsurface Studies) is located inside the ExoMars Rover's drill. It will be the instrument in closest contact with the Martian subsurface. Ma_MISS will image the walls of the borehole created by the drill to study Martian mineralogy and rock formation. This will provide valuable information for the study of subsurface soil and rock layers (i.e., stratigraphy), the distribution and state of water-related minerals, and will help to characterize the geophysical Martian environment.

CLUPI (Close-up imager) consists on a camera system to acquire high-resolution color close-up images of rocks, outcrops, drill fines and drill core samples. This instrument will be the only one to access the core samples before they are crushed.

MicrOmega is a micro-imaging system designed to identify, at grain scale, the mineralogical and the molecular composition of the Martian samples collected by the ExoMars drill.

RLS (Raman Laser Spectrometer) is a Raman spectrometer which provides a powerful tool for the definitive identification and characterization of minerals and biomarkers. Raman spectroscopy is sensitive to the composition and structure of any mineral or

organic compound. This capability provides direct information of potential organic compounds that can be related with present or past signatures of life on Mars as well as general mineralogical information for igneous, metamorphous, and sedimentary processes, especially water-related geo-processes.

MOMA is the largest instrument in the ExoMars Rover, and the one directly targeting biomarkers. MOMA will answer questions related to the potential origin, evolution and distribution of life on Mars. This instrument will detect organic molecules, even at very low concentrations, and try to establish their biotic or abiotic origin by molecular identification in terms of chirality. In addition to studying the samples collected by the drill, MOMA will also analyze gases in the Martian atmosphere.

Drill and SPDS: A key part complementing this suite of instruments is the drill and sample preparation and distribution system (SPDS) of the rover. The drill will obtain samples from the Martian subsurface, at sites defined based on data from PanCam, ISEM, WISDOM and ADRON-RM. This sample will be extracted from depths as low as 2 meters under the surface, allowing the analysis of 'fresh' samples that have not been hit by ultraviolet light, minimizing the chances of destruction of potential organic materials. The drill is equipped with the Ma_MISS IR spectrometer to analyze the samples at different depths during the drilling process. Finally, the drill will extract a core sample from the subsurface to be further analyzed by the onboard instruments. In order to provide key textural and context information of the sample, CLUPI will acquire close-up images of the core sample prior to delivery to the SPDS.

The SPDS is a system that will acquire the core sample and crush it to present the instruments in the ALD with a flattened surface for analysis. The first instrument to analyze the sample will be MicrOmega, the infra-red spectrometer. This instrument will acquire multi-spectral images of the sample in a relatively wide area, providing a general view of the materials in the sample. In case something of special interest is found in the sample surface, the SPDS will position the sample under the RLS instrument spot for detailed analysis with Raman spectroscopy. Nevertheless, the RLS instrument will perform a random mapping on the sample surface, as it has demonstrated a good mapping capability by analyzing several points in homogeneous powdered samples. Finally, the combination of RLS and MicrOmega data will provide the necessary inputs to decide if the sample is analyzed with MOMA. The reason is that the number of ovens available for the GCMS is limited, so only a subset of samples will be analyzed with this instrument.

This work only focusses on the RLS instrument. However, the interpretation of the data acquired during the mission will be performed using the whole payload, so it is important to keep in mind the necessary complementarity between RLS with the other instruments of the ALD.

1.4 RAMAN SPECTROSCOPY AND THE EXPLORATION OF MARS

The potential of Raman spectroscopy as a non-invasive or destructive procedure applied to all kinds of materials defines it as a very important technique for all types of materials, minerals or biological analyses. The following paragraphs, mostly extracted from [9], provide an explanation about the feasibility of the application of Raman spectroscopy for space exploration.

Raman spectroscopy is well recognized as a powerful tool for the chemical and structural identification of materials in the solid, liquid or gas state. Its analytical capabilities, without the need to perturb a sample, have made this technique unique for many applications where the materials are scarce or very valuable and rare. As consequence of these capabilities a great expansion in applications has been seen in the last 25 years across many fields ranging from chemistry to mineralogy and geology, art and archaeology, forensic sciences, polymers and biomaterials [9].

This technique can be used to identify the composition and structure of minerals and rocks at the mineral grain scale, by observing rotational, vibrational and other low-frequency modes of a sample. The spectral information received by the spectrometer (number of peaks, position and relative intensities) is determined by the molecular structure and composition of a compound, enabling the identification and characterization of minerals and rocks at molecular level. This method allows analyzing the sample in a non-destructive way, which means that the sample remains intact for further analysis by other instruments. These identification capabilities provide a definitive characterization of a target material.

The scientific community is showing increasing interest on Raman spectroscopy for the identification of minerals, as demonstrated by collaborative projects such as RRUFF [10], which provides an online database for the identification of minerals from Raman spectra, counting with thousands of instances in the database. Indeed, the Raman spectrometer onboard the ExoMars mission will be used to identify organic compounds and search for signatures of life, to identify the mineral products and indicators of biological activities, to characterize mineral phases produced by water-related processes and to characterize igneous minerals and their products resulting from alteration processes (e.g. oxidation) [5].

A Raman spectrometer basically consists of two elements: a monochromatic laser source and a spectrometer. The laser is used to excite the sample under analysis to its rotational and vibrational modes. The sample will scatter the incident light, mostly at the same frequency of the excitation source (which is known as Rayleigh emission). However, this interaction of the light with the crystalline matter will cause low-probability shifts on the emitted light, which are related to the different modes of the

17

molecule s of the sample. This secondary radiation (as referred by Raman in [11]) is known as Raman emission, and occurs at frequencies shifted from the excitation source by an amount dependent on the molecular structure of the samples. This light will be captured by the spectrometer, which will provide a spectrum with the information emitted by the samples.

A typical Raman spectrum and its characteristics can be seen in Figure 1-3. The x axis represents the Raman shift (wavenumber of the Raman vibration with respect to the laser wavenumber). The units for the intensity of the peaks does not directly correspond to the number of photons received by the CCD due to inefficiencies of the devices as well as the mediation of and ADC (Analog to Digital converter), reason why the counts are measured in a. u. (arbitrary units). The laser excitation is always positioned at 0 cm^{-1} (as the Raman shift is referenced with respect to the excitation source position), and filtered by a notch filter (also referred to as band-stop filter) which weakens the excitation light received by the spectrometer in several orders of magnitude. Usually, Raman signals are often swamped or obscured by the background or baseline, due to fluorescence emission, and do not provide any useful information for the Raman spectrum. Thus, the peak intensity of the Raman peaks is defined as the number of counts at the position of the peak, minus the counts of the baseline or background. In order to simplify the interpretations of the spectra, determine the peak positions, estimate the peak width, and measure the peak intensity accurately, the elimination of the background is necessary [12].
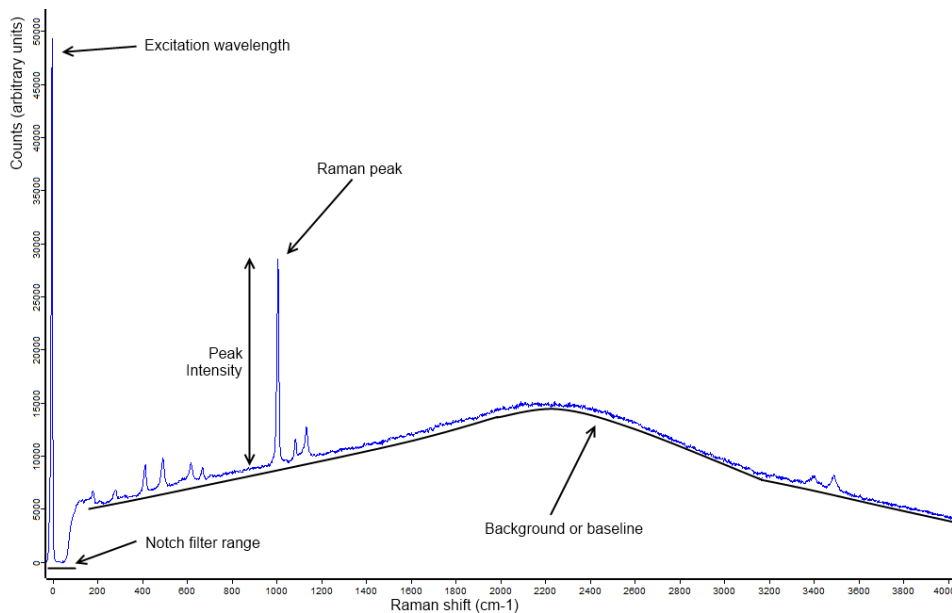
Figure 1-3. Typical Raman spectrum

In the framework of the ExoMars mission, Raman spectroscopy is a key technique, as it will provide tools for the analysis of samples from an exobiology perspective, which is

the final aim of the ExoMars mission. In this direction, lots of research has been carried out in order to study the capabilities of Raman spectroscopy for the analysis of samples from an astrobiological point of view. Especially interesting is the resonant effect of carbonaceous matter, which is thus easily detected with Raman spectroscopy, and has been amply studied [13, 14]. Furthermore, the interest in carbonaceous matter is not only based on the carbon-based composition of organic materials, but also supported by analysis in Martian meteorites [15]. However, there is some controversy regarding the biotic – abiotic origin of carbon based-only on the Raman spectra of the carbonaceous matter. Raman is considered necessary, but not sufficient to assess the biotic origin of this matter [16, 17]. However, by complementing the Raman information with other analyses, such as optical images and structural morphology studies of the carbonaceous matter can univocally assess the biotic origin of carbonaceous matter [18, 19]. In addition to carbonaceous matter *per se*, Raman spectroscopy has also been applied for the detection of biomarkers and organic materials [20-27].

All the research performed with Raman spectroscopy indicates that this technique is a key tool for *in situ* investigation of the Martian surface, for mineralogy as well as for exobiology, though for some cases, it is necessary to complement this technique with others, which highlights the importance of the complementarity of the instruments onboard the ExoMars rover. Thus, the RLS instrument will support the scientific measurements by correlating its spectral information with other spectroscopic and imaging instruments such as the MicrOmega Infrared Spectrometer, or the CLUPI images.

1.5 THE RLS INSTRUMENT AND THE SPDS

The RLS instrument will operate on samples collected from the surface and the subsurface down to 2 meters, by means of the drill included in the ExoMars rover. Then, the samples will be crushed into a fine powder and by means of a dosing station the powder will be presented to RLS and the other instruments in the ALD [28, 29].

There shall be two operational modes [1]: an automatic scanning and a smart scanning.

The mode selection will be performed by spacecraft commands from ground. During automatic scanning, the rover shall place the target following a preconfigured sequence of movements in front of RLS optical head. RLS shall take at least 20 shots per sample at regularly spaced spots on the target. During smart scanning, MicrOmega IR images will be processed by the rover to determine if there is any target of interest which can be reached by the RLS instrument. In affirmative case, the rover would place the target of interest below the RLS optical head. In case no target of interest is found, the preconfigured sequence of movements would be performed instead.

Figure 1-4 shows the sample preparation and distribution system (SPDS) of the rover. This system will crush the sample to a powder with a grain distribution of 200-250 microns and below [30], and deposit it in a sample container in the carrousel. The carrousel will then move to flatten the surface of the powder before presenting it to the ALD instruments for analysis.
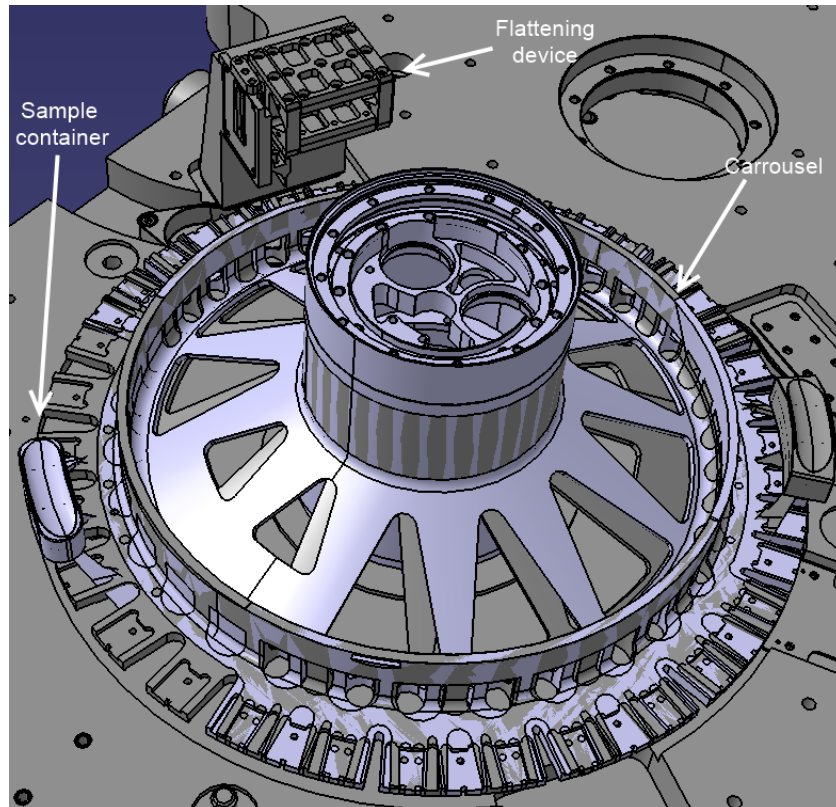
Figure 1-4. Schematics of the rover Sample Preparation and Distribution System

In spite of the technological development that has led to the expansion of Raman spectroscopy, and the capabilities of this technique, Raman spectroscopy has been, and to a great extent still is, considered as a laboratory technique. Nevertheless, in recent years several technological developments in laser sources, optical elements, spectrometers and detectors have led to the possibility of developing robust, compact and miniaturized spectrometers, which have many of the spectral capabilities of laboratory-based systems.

These developments, thus, also paved the way for the potential use of Raman spectroscopy in planetary exploration as part of rover or lander instrumentation, in particular for the exploration of Mars. Furthermore, at present, Raman spectroscopy is considered as a fundamental, next-generation instrument for the characterization of mineralogical and organic material during the exploration of Mars [9]. Previous works

on technical developments of Raman spectrometers for space exploration support this statement [31-36]. However, the RLS instrument will be the first Raman spectrometer to land on Mars.

Due to constraints related to space environment (thermal, mechanical, radiation, planetary protection...) within the rover analytical laboratory, the RLS instrument consists of three independent units connected via electrical and optical harnesses [1], as shown in Figure 1-5 and Figure 1-6.The laser is in the Instrument Control and Excitation Unit (ICEU). Its collimated beam is conveyed to the Internal Optical Head (iOH), where it is focused on the target. This unit must incorporate a mechanism to adjust the optical focus because of irregularities of the sample surface. The Raman signal is collected through the same front optics, the excitation wavelength is filtered out, and the signal is transmitted to the Spectrometer Unit (SPU) for dispersion and capture on a two-dimensional Charged Coupled Device (CCD).

Figure 1-5 shows how the sub-units are mounted in the rover Analytical Laboratory Drawer (ALD). This configuration is necessary because the target sample has to be kept within an ultra-clean zone in the rover to comply with the planetary protection policies (aimed at avoiding biological contamination of extraterrestrial bodies), and thus it is not in physical contact with RLS. These subunits are further described below [1]:

- Spectrometer unit (SPU). The dispersion of Raman light is obtained by transmission using a holographic grating. The Raman spectrum is registered on a 2268x512 pixels CCD, which is kept at -10ºC by means of a thermo-electric cooling device.
- Control and excitation unit (ICEU). It includes de DC/DC power converters and the data processing capability (micro-controller, RAM memory, clock and CANBus). Its role is also to capture the RLS health parameters and to control the thermal management. To support other functionalities of the instrument, it comprises the laser with two redundant excitation outputs, the CCD front-end electronics and the autofocus driver to focus the laser beam on the surface of the sample.
- Optical head (iOH). The range of focus is 2mm in a confocal configuration, with a working distance of 13 mm.

The laser source is continuous at a wavelength of 532 nm. This laser will be able to provide a total of 400 individual sets of measurement cycles (10 minutes maximum each). The laser spot size will be around 50 microns on the surface of the sample. This spot size procures an irradiance level on the surface of the sample fixed to remain below the threshold of powder grain thermal damage, mainly in oxides and hydroxides. The goal of the spectrometer is to cover the spectral shift from around 200 to 3800 cm^{-1}, to capture the range of fundamental vibration of rock forming minerals,

oxyanionic anions, functional groups of organic species, water and OH vibrations in hydrates. The Raman spectral resolution goal is required to be 6 cm^{-1} in the fingerprint spectral region below 2000 cm^{-1} and around 8 cm^{-1} above this limit [28]. Figure 1-6 shows the schematics of the RLS instrument components.
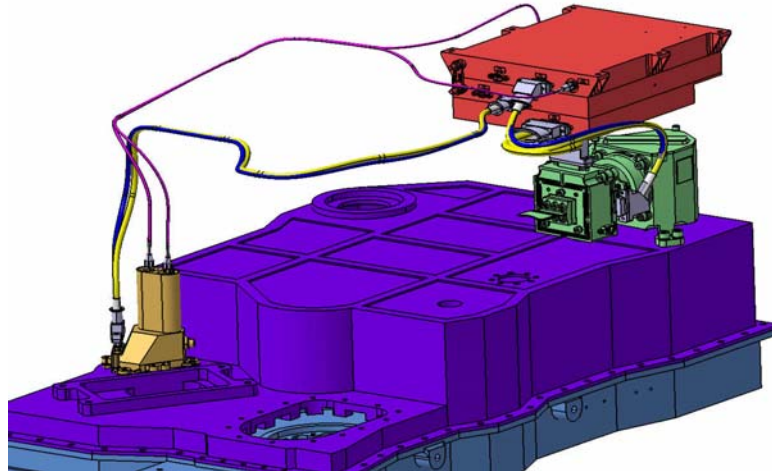
Figure 1-5 Schematics of the implementation of the RLS instrument in the ExoMars rover. Credits: ESA
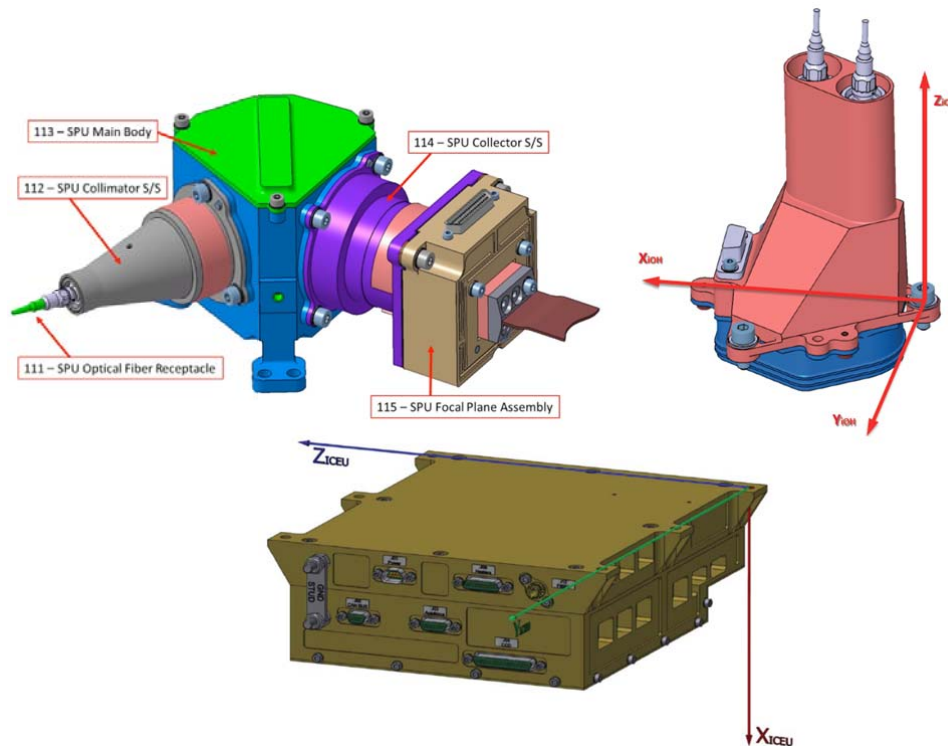
Figure 1-6. Schematics of the RLS spectrometer units: SPU (up, left), iOH (up, right) and ICEU (bottom). Credits: INTA (up), IRAP (bottom).

In this context, the present work was aimed at optimizing all the processes that can influence the scope of the scientific information obtained from the RLS instrument. This is achieved by addressing two different issues. On one hand, it was necessary to implement techniques for the RLS instrument to be operated automatically, without the intervention of a human operator. On the other hand, the analytical process of the data obtained from the instrument also needed to be optimized by applying novel techniques for the analysis of Raman spectra.

The optimization of the acquisition process is achieved by addressing the spectral acquisition on a point by point basis, but also optimizing the best possible mapping performance. In this work I have implemented several algorithms that allow optimizing the spectral acquisition, but also incremented the mapping performance of the instrument by reducing the computational time necessary to acquire spectra. Thus, it will be possible to perform better mappings of the sample surface by increasing the number of points of the sample that are analyzed. In this line, we have studied the influence of the grain size distribution and the problematic associated with the sample preparation and distribution, which will also contribute to the definition of the best mapping parameters. In order to be able to perform the experiments, I built and programmed the RLS ExoMars simulator described in Chapter 2, which allows acquiring spectra in a SPDS-like scenario.

In addition to the optimization of the acquisition process, there is also an important influence from the way the data obtained from the instrument is analyzed. This work presents a study of the influence on the Raman spectrum due to the singularities of the acquisition process, which is of utmost importance in order to characterize the identification capabilities of RLS. Furthermore, these singularities allow applying several techniques for the analysis of the data products from RLS which potentially can provide more complete information from the sample. Thus, this thesis intends to study and optimize, not only all the parameters related to the spectra acquisition, but also optimizing the analytical process of the spectral data obtained under the RLS constraints. Regarding to this second aspect, several techniques are studied and applied to RLS data, providing information regarding the identification accuracy, and the possibility of quantifying the abundances of minerals in mixtures.

Part I

Chapter 2.

Development of the RLS ExoMars Simulator

Description of the RLS ExoMars simulator system performance and capabilities

2.1 INTRODUCTION

In order to achieve the maximum scientific return from the RLS instrument once on Mars, it is necessary to perform experiments under the operation conditions imposed by the rover. To do so, as part of the work of this thesis, we have built and programmed the so-called RLS ExoMars simulator at Unidad Asociada UVa-CSIC-CAB (Figure 0-1 and Figure 2-1). This Simulator has enabled the scientific team to perform many scientific experiments under conditions similar to those provided by the Analytic Laboratory Drawer (ALD) of the ExoMars rover. It also provides the means for the definition and development of the necessary algorithms for the automation of the instrument. The following sections describe how the system is implemented, as well as its capabilities from the hardware and software perspectives.
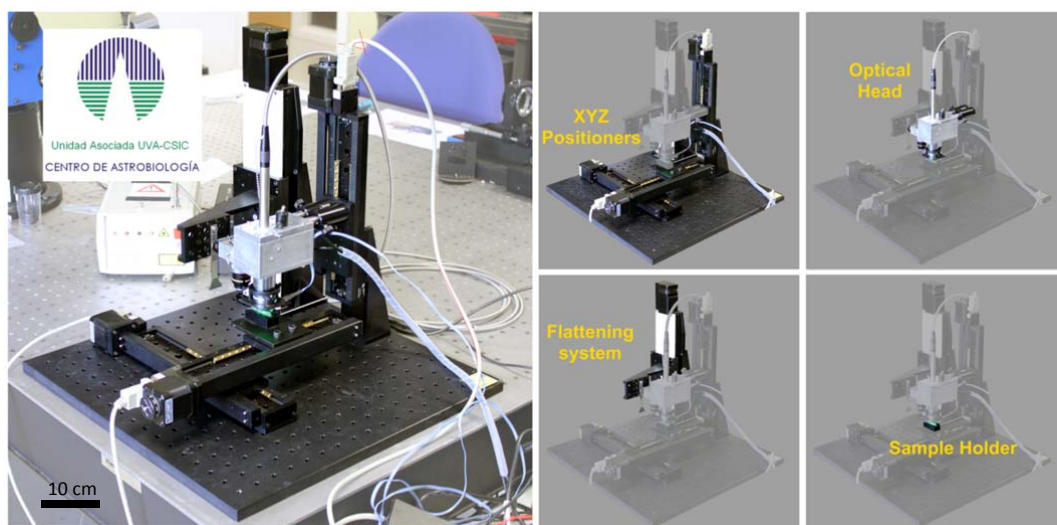
Figure 2-1. RLS ExoMars simulator prototype and schematics
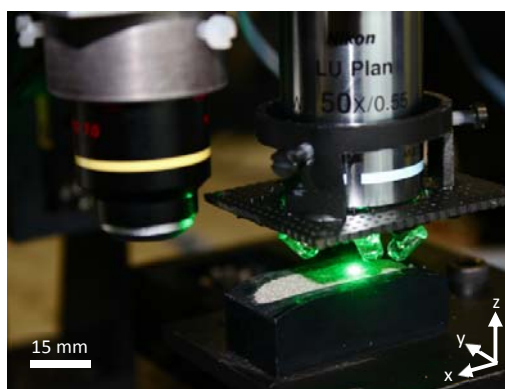
Figure 2-2. Detail of the sample analysis

2.2 POSITIONING

The basic configuration of the RLS ExoMars Simulator consists of an XYZ micrometric positioning system with an optical head attached to the Z axis, and a refillable container attached to the XY positioners. This way, a sample can be positioned under the optical head and be analysed (see Figure 2-1 and Figure 2-2).

The rover Sample Preparation and Distribution System (SPDS) carrousel and the focusing system in RLS are emulated with linear positioners (X and Z axis, respectively), assuming that the carrousel radius is big enough to be approached by a straight line. Furthermore, another axis (Y axis) was added to be able to address issues as collaborations with other instruments, in a collaborative science mode configuration. The positioners are Standa 8MT175-200 (20cm range) for the X and Y axis, and 8MT175-150 (15cm range) for the Z axis, configured with stepper motors. These devices offer a spatial resolution of 2.5 microns. They also give a very high repeatability and accuracy parameters. The microcontrollers used are Standa 8SMC1-USBhF, 1.5A Microstep Driver with USB Interface, powered with 12V. The positioners are equipped with hard stop switches to avoid damage to the stepper motors when reaching the range limits. In addition, the hard stop switches have been configured in order to protect the optical head from damages, providing an effective range of the Z axis of 8 cm, which is more than enough to serve its purpose.

2.2.1 SECURITY STOP

There are three ways to stop the positioner motion. The limit switches of the positioners will automatically stop the motion when activated (either manually or by the stage). There is a switch than the operator can activate to completely stop the motion of all the three positioners. Lastly, removing the current to the micro-controllers will stop the motion of the system.

2.3 IMAGING

The optical head implemented in the RLS ExoMars simulator has been provided with two different objectives which can be positioned over the sample to allow making better studies of the sample. The first objective is a Nikon 10X/0.25 Pol, WD 7.0, which is used only for imaging purposes. This objective provides a field of view of 1600 x 1200 squared microns, and intends to simulate the collaborative science with the MicrOmega instrument on the ExoMars mission. The second microscope objective used in the simulator is a Nikon L Plan SLWD 50X/0.45 EPI, WD 13mm. It provides a field of view of 280 x 210 squared microns. This objective is used both for imaging purposes as well as to focus the Raman laser and collect the Raman signal in a confocal configuration. This way, both the image and Raman excitation share a common optical path, allowing the scientist to see exactly where the laser hits the sample. The hardware configuration is based on a beamsplitter placed with a 45º angle (see Figure

2-4), allowing the Raman probe and the camera to use the same optical path to focus the sample. With this configuration, and with the BWTek laser and a BWTek BAC-100 probe, a spot size of around 58 microns is obtained on the sample surface, when the nominal spot size of the RLS instrument is 50 microns. The experimentally measured spot size is depicted in Figure 2-3, with both the beam and Gaussian adjustment measurements at 13.5% power. Both the cameras used in this layout are web cams Vimicro USB2.0 UVC PC Camera, which were conveniently adapted to be mounted in the setup.
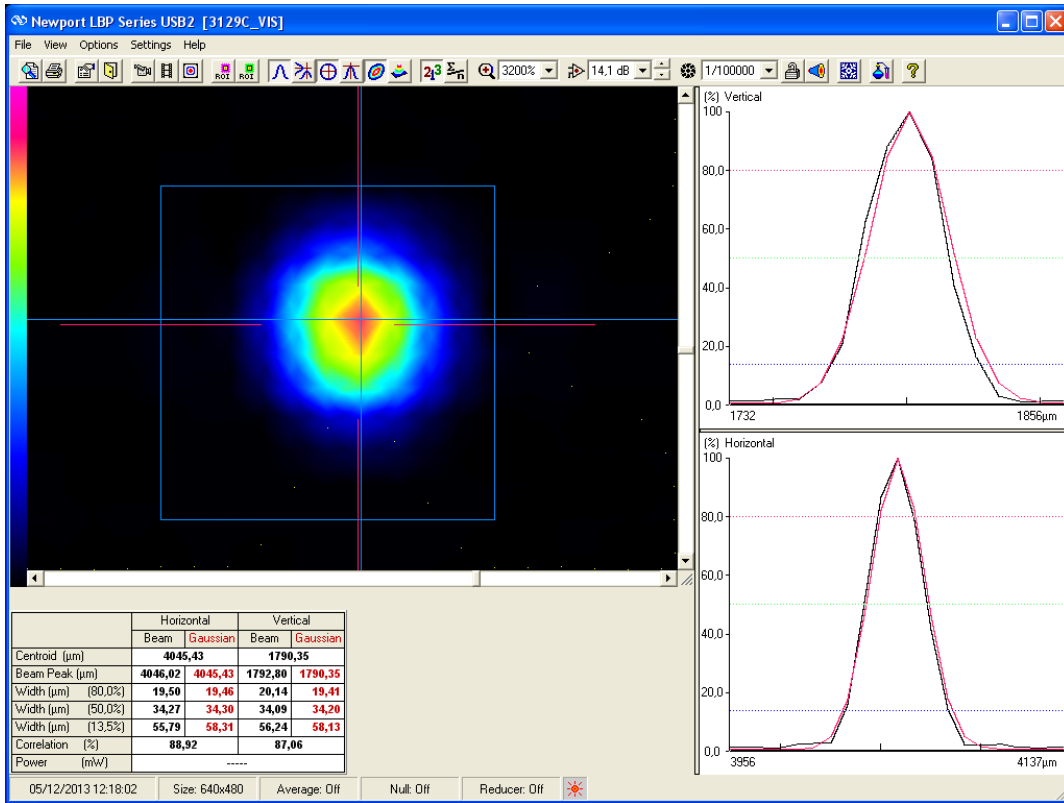
Figure 2-3. Laser beam profile measured with the BWTek laser and the BWTek BAC-100 optical head
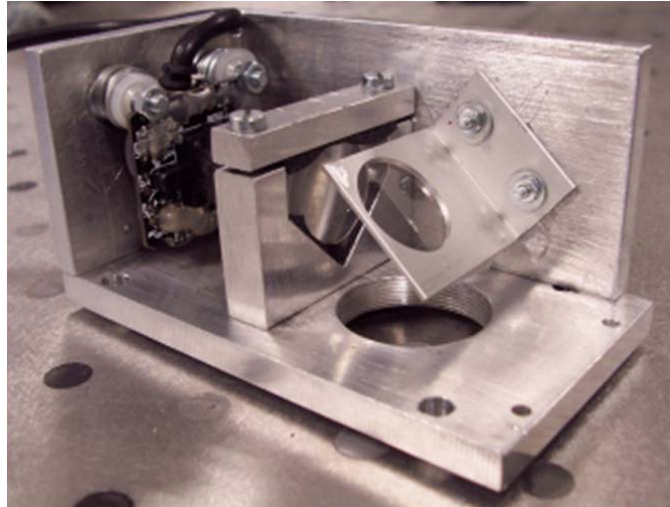
Figure 2-4. Beamsplitter in the Optical Head to share the image and Raman optical path

2.4 FLATTENING

The system is also designed to flatten the sample powder with a rover-like design and position the sample under the instrument optical head (Figure 2-5), which will focus on the sample surface. The flattening blade height can be adjusted by means of a Newmark NSC-1 controller and positioning stage, as well as manually adjusted with the flattening blade screws.
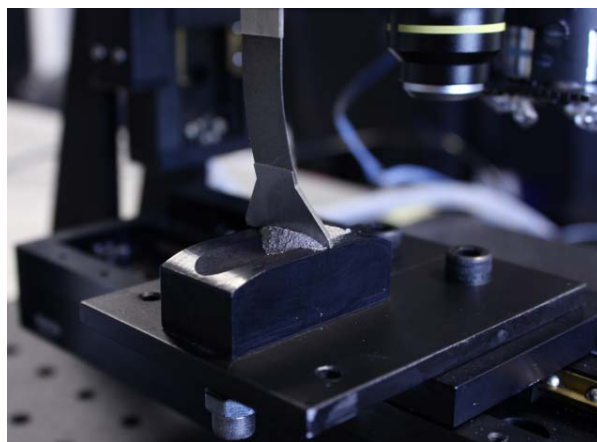
Figure 2-5. Flattening system in operation

2.5 SOFTWARE CAPABILITIES

The control software for the simulator has been developed in Visual C++ with Visual Studio 2008. This software tool is shown in Figure 2-6. However, a second version of the software has been implemented in LabView in order to facilitate the integration with other systems (SPDS, for example). This version of the software is described in section 2.5.3. The description of the software is aimed at describing the system capabilities, as well as to be used for further reference in a "user manual" fashion.

To acquire Raman spectra, a control module for spectrometers and lasers was included in the RLS ExoMars Simulator software. Currently, BWTEK i-Raman 532 nm, BWTEK Compass T and Andor Newton DU940N-BV spectrometers can be controlled from the simulator in manual mode, as well as the Monocrom Prototype II and the BWTEK Laser Laz-USB 532 nm lasers. The designed optical head allows the coupling of many different types of Raman probes to the simulator, including the BWTEK, InPhotonics or KOSI probes available at Unidad Asociada UVa-CSIC-CAB.

NOTE: The automatic mode is optimized to be used with the Compass-T spectrometer and the BWTEK laser device, though it will work with any device configuration, as long as they are correctly selected in the software.
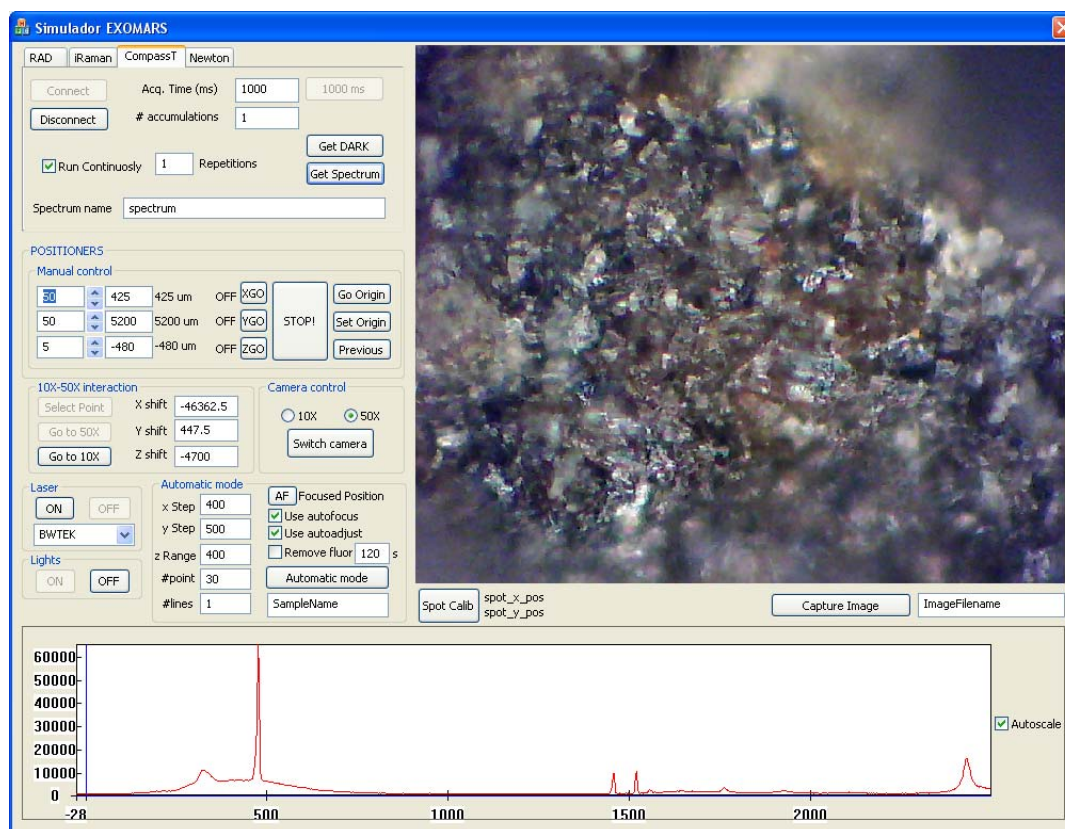
Figure 2-6. Software developed for controlling the Simulator (Visual Studio version)

2.5.1 MANUAL OPERATION

The RLS ExoMars simulator can be operated manually to perform all the basic operations of the system.

2.5.1.1 START-UP CONFIGURATION

There are several parameters that have to be configured each time the software is started and before executing any other activities. These configuration steps consist on the selection of the correct camera and the laser spot calibration.

2.5.1.1.1 SELECTION OF THE CORRECT CAMERA

The camera that is actually being shown should be correctly mapped to the radio button in the Camera Control field. This is due to the fact that both 10X and 50X cameras are equal, and the Operative System cannot distinguish between them by default, as they use the same driver. To be able to distinguish both cameras, it has to be noted that the 50X camera has a dark spot on the upper-left part of the image, while the 10X doesn't.

Figure 2-7. 50X image with the dark spot on the upper-left part

2.5.1.1.2 LASER SPOT CALIBRATION

By default, the laser spot position is set to the center of the 50X image. However, the high number of freedom angles that the positioning of the Raman probe provides, discourages the hard-coding of the laser spot position. Thus, with each new execution of the software, the spot position has to be recalibrated. This is done by clicking the 'Spot Calib' button and selecting the position in which the laser spot is found on the image. The resulting position will be indicated in the 'spot_x_pos' and 'spot_y_pos' text fields.

2.5.1.2 CAMERA SELECTION
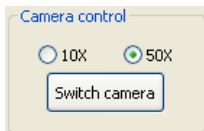
Figure 2-8. Camera control panel

The 'Switch Camera' button (Figure 2-8) can be used to change the viewing camera of the system. There are two possibilities consisting on the 10X and the 50X cameras. These devices are referenced by Windows with the driver name, thus, they both have the same name (Vimicro USB 2.0 UVC PC Camera) on the pop-up menu that appears when pressing the button (see Figure 2-9). However, they can be distinguished based on the black spot that can be appreciated in the upper-left part of the 50X field of view. The radio buttons can be clicked to ensure that the current camera corresponds to the correct radio button.
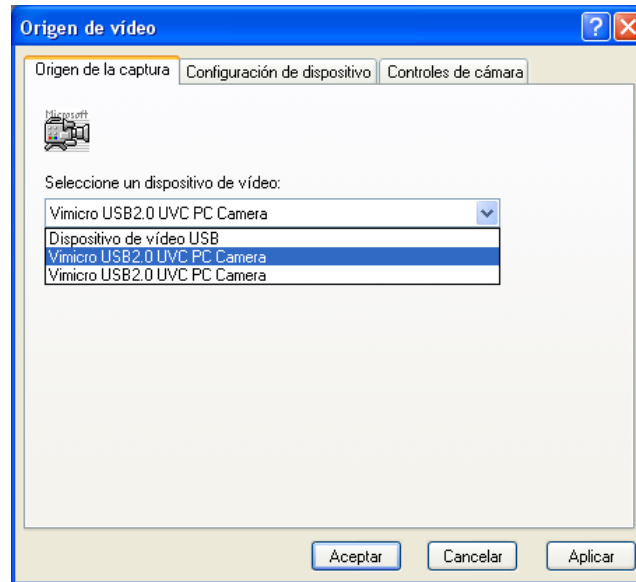
Figure 2-9. Camera selection menu

2.5.1.3　IMAGE CAPTURING

To capture the current view to an image, the Capture Image button can be used. The image is saved in .bmp format with the name indicated in the text field on the right of the button, and in the folder C:\Resultados_SimuladorExoMars\ManualMode\Images.

2.5.1.4　POSITIONING
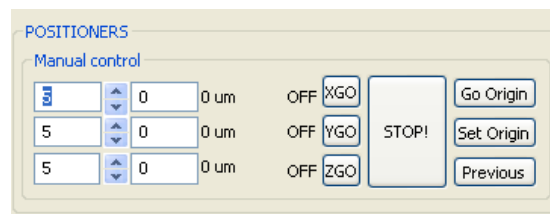
Figure 2-10. Manual positioning controls

The edit boxes on the left indicate the distance increment that the positioner will move when clicking the arrows on their right (5 microns in Figure 2-10). The minimum increment should be set to 2.5, as this is the spatial resolution of the system. Any desired position can also be manually introduced in the boxes found to the right of the arrows, and the positioner will move once the corresponding 'GO' button is pressed. The actual position of the encoder is shown in the text field by the right of the position edit boxes.

The system can be moved to position (0,0,0) by pressing the 'Go Origin' button. The current position of the encoder can be set as the origin reference by pressing the 'SET Origin' button. The previous button will move all three axis to their corresponding previous position. The 'STOP' button will stop the motion of any moving positioners,

based on the motor status (ON/OFF), shown in the text fields to the left of the 'GO' buttons. All units are presented in microns.

2.5.1.5   10X-50X INTERACTION (COLLABORATIVE OPERATION)

To simulate the collaborative operation between different instruments of the ExoMars payload, an interaction tool between the different cameras and objectives was implemented. Figure 2-11 shows the control panel for this function. The underlying principle is that, if something of interest was found on the wider field of view of the 10X camera (simulating the MicrOmega instrument onboard the ExoMars rover), this point should be selected and placed under the 50X camera and Raman optical head (RLS instrument), exactly under the spot position, to be analyzed by the Raman instrument (see Figure 2-12). To do so, the 'Select point' button will be active when the 10X camera is activated. After pressing it, a point on the image can be selected, and the offset automatically calculated taking into account the shift from the center of the 10X image, as well as the shift of the calibrated laser spot from the center of the 50X field of view. This offset is based on the shift values that can be seen in the text fields in Figure 2-11, which correspond to the distance between the centers of the 10X and 50X images. These values can be modified in case a recalibration is needed, but they will be reset to the original values when the program is restarted, unless modified in the source code.
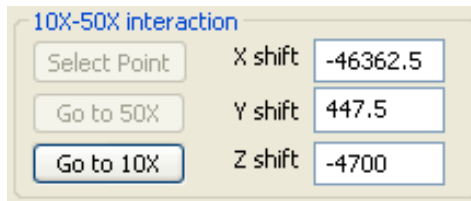
Figure 2-11. Control panel for 10X-50X travelling
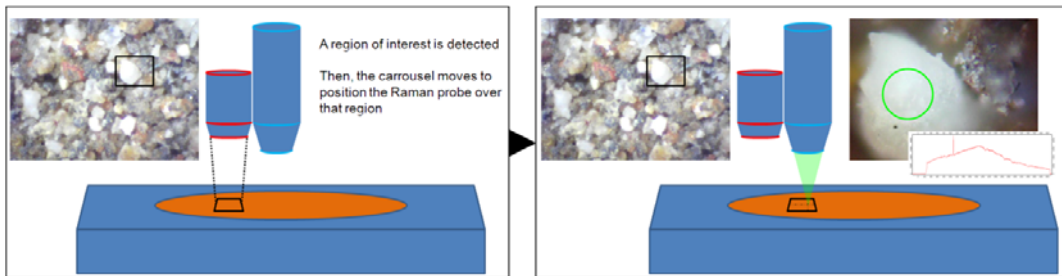
Figure 2-12. Depiction of the collaborative mode operation.

Once an interesting point has been selected, the 'Go to 50X' point button will take the selected point under the RLS spot.

When the 50X camera is activated, there is the option of going back to the 10X position by pressing the 'Go to 10X' button.

2.5.1.6   ILLUMINATION

Figure 2-13. Illumination (LED) and laser control.

The software provides 'ON' and 'OFF' buttons for the LED illumination of the system. When pressed, the LED illumination will be turned ON and OFF. When turning ON, the laser is automatically turned OFF (Figure 2-13).

2.5.1.7   LASER
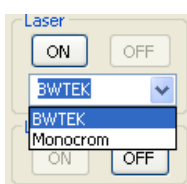
Figure 2-14. Laser control panel

The software also provides 'ON' and 'OFF' buttons for controlling the laser device (Figure 2-14). When the 'ON' button is activated, the LED illumination is automatically turned off to avoid errors during the acquisition. The RLS instrument will not have the capability to control the laser power, thus, in the RLS ExoMars Simulator, it needs to be set by means of the external software LazUSB when using the BWTek laser device (or using the LabVIEW version of the software). The Monocrom device does not provide power control electronics.

2.5.1.8   SPECTRA ACQUISITION

The spectra acquired by the different spectrometers is automatically calibrated. However, the calibration parameters are, in all cases, hard-coded in the software source code, so, in case they have to be changed, these parameters have to be updated in the code.

2.5.1.8.1   RAD1 SPECTROMETER

The software does provide the possibility to include the RAD1 spectrometer (Raman Demonstrator 1, a spectrometer developed with characteristics similar to the flight instrument, with standard components). However, due to problems with the SDK software of this device, and the lack of support from Hamamatsu, the CCD camera manufacturer, it is not possible to correctly acquire spectra from this instrument. Thus, the corresponding RAD tab is empty and cannot be used.

2.5.1.8.2   IRAMAN AND COMPASST SPECTROMETERS
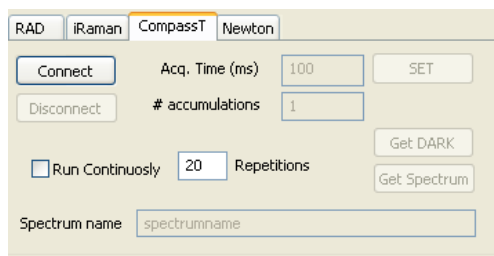
Figure 2-15. Spectra acquisition tab for the CompassT spectrometer (similar to iRaman)

The BWTek spectrometers (iRaman and CompassT) have a similar control panel (Figure 2-15). To manually operate these devices, it is first necessary to connect to it by pressing the 'Connect' button. The acquisition time and number of accumulations are indicated in the corresponding edit boxes. To fix the acquisition time value the set button has to be pressed.

To acquire several spectra in a continuous way, the 'Run continuously' checkbox can be activated, and this number of spectra will be acquired one after the other, and represented in the spectrum drawing box. The aim of this capability is to be able to observe the evolution of the spectrum among different acquisitions (for example to asses if the sample is fluorescent).

Buttons for the acquisition of a dark spectrum and the acquisition of the final spectrum are also enabled once the integration time has been correctly set to the spectrometer. When pressing the 'Get Spectrum' button, a spectrum will be acquired with the indicated acquisition parameters. This spectrum will be saved in the C:\SimuladorExomars\ManualMode\Spectra folder with the name introduced in the 'name' edit field. The spectrum will be saved as .xy (r-shift, intensity) and .txt2 (multiple column) formats. If a dark spectrum has previously been acquired, it will be subtracted from the intensity vector in the .xy file, and the corresponding column of the .txt2 file.

2.5.1.8.3   ANDOR NEWTON SPECTROMETER
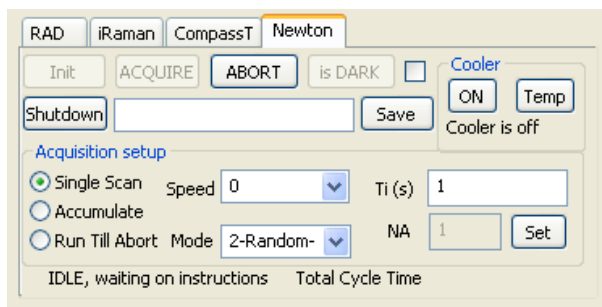
Figure 2-16. Spectra acquisition tab for the Andor Newton CCD spectrometer

2.5.1.8.3.1 ACQUISITION SETUP

The Andor Newton CCD spectrometer is controlled and setup with the control panel shown in Figure 2-16. The basic functionality is similar to the BWTek spectrometers in 2.5.1.8.2, with some differences. This spectrometer, after initialization, has the possibility to set-up several parameters. All the parameters will be set by pressing the 'Set' button, which has to be pressed each time that any parameter wants to be changed.

There are three modes of operation: 'Single Scan', 'Accumulate' and 'Run Till Abort'. In the 'Single Scan' mode, only one spectrum of Ti seconds is acquired. In the 'Accumulate mode', several accumulations are added to the spectrum in order to reduce the spectral noise. Finally, the 'Run Till Abort' mode will acquire spectra of Ti seconds until the abort button is pressed. These spectra will be shown in the software representation chart. Contrary to the BWTek spectrometers, the software is not blocked until all the accumulations are acquired, so this acquisition mode can be used for performing the focus on the sample by looking at the Raman signal.

Figure 2-17 shows the different selectable values for the readout speed and mode of the CCD. The lower the speed value, the lower the readout noise. However, as there is no shutter in the spectrometer, during the readout process light will still be acquired by the CCD, and lower readout speed will imply longer exposition to the light, even if unwanted. 0, 1 and 20 correspond to 50kHz, 1MHz and 2.5 Mhz, respectively.

Regarding the readout mode, there are 5 different possibilities. Full Vertical Binning (FVB) will accumulate all the rows of the CCD in the output register. Multi-Track mode will acquire multiple tracks at evenly spaced distances, covering the whole surface of the CCD. The Random-Track mode will acquire two different tracks at hard-coded positions. These tracks are calibrated to the current configuration of the spectrometer and correspond to the tracks of the Kayser diffraction grating on the CCD. If these tracks need to be changed, the parameters have to be hard-coded in the software source-code. The Single-Track mode will acquire only one track on a determined position, and the Image mode will acquire a 2-D image of the whole CCD. The recommended configuration is at minimum speed readout rate and Random-Track readout noise, and this is the configuration used by default for the automatic acquisition.
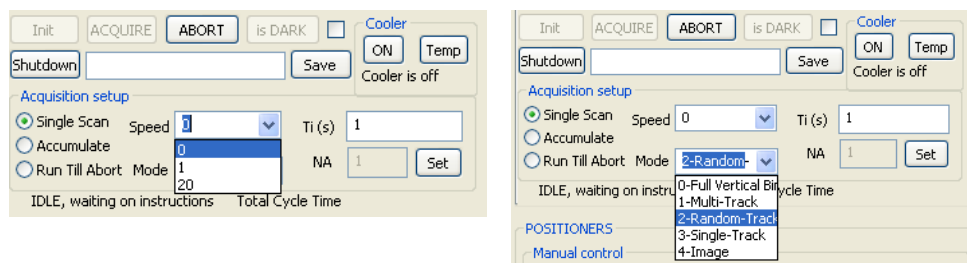
Figure 2-17. Different configuration options for the readout speed (left), and mode (right) of the CCD

2.5.1.8.3.2 TEMPERATURE SETUP

The CCD of the Andor Newton spectrometer can be cooled down by means of a thermo-electric cooler. This temperature can be set by clicking on the button Temp and selecting the desired temperature (see Figure 2-18). The cooler is turned on and off using the ON/OFF button. The cooler device PID can take several minutes to

stabilize the temperature of the CCD, though this does not imply that the spectrometer cannot acquire spectra during that time.
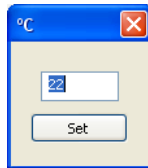
Figure 2-18.

Temp. menu

NOTE: It is important to be aware that the Andor Newton CCD should not be powered-off before the CCD has reached temperature values over at least 0ºC, to avoid damages to the CCD due to condensation. The software is blocked before closing, after shutting down the spectrometer and after automatic acquisition until the safety conditions are reached.

2.5.1.8.3.3 ACQUISITION PROCESS

To acquire spectra, the acquisition parameters have to be fixed, and the temperature (optionally) set to the desired value. Once the spectrometer is correctly configured ('Set' button is pressed), it will be possible to acquire spectra using the 'Acquire' button. Once the spectrum is acquired and the data is available, it will be possible to save the spectra with the 'Save' button, using the name input in the corresponding field.

2.5.1.8.3.4 DARK SPECTRUM ACQUISITION

By default, the dark spectrum that is subtracted to the newly acquired spectra is 0. In order to acquire a dark spectrum to subtract to the subsequent acquired spectra, the 'isDark' button and the checkbox by its side have to be used. The 'isDark' button will save the last acquired spectrum as a dark, and all the subsequent acquired spectra will be saved with this dark subtracted from the intensity signal. This will happen while the checkbox is activated, or while the acquisition conditions do not change. Deactivating the checkbox or changing the acquisition settings (pushing the 'Set' button) will delete the current dark spectrum.

2.5.1.9 Autofocus

2.5.1.9.1 THEORY

The RLS ExoMars simulator software includes an autofocusing algorithm. This algorithm performs image analysis, based on the grey level variance of a region of the image to decide whether the focusing level is better or not. This way, taking several images in different points, it will be possible to decide in which one of them the focusing level is best, based on the following algorithm:

Variance is defined as $\sigma^2 = \frac{1}{N^2}\sum_{x=1}^{N}\sum_{y=1}^{N}(f(x,y) - m)^2$, where m is the mean of the distribution of the grey levels (luminance) of the whole image, *f(x,y)* is the intensity of the grey level of the pixel in position *(x,y)* and the selected squared focusing area has sides of *N* pixels. The bigger the variance, the better focused the image.

For the RLS ExoMars simulator, this algorithm has been implemented using a focusing surface of 32 by 32 pixels, which is equivalent to about a 15 by 15 microns surface, in the center of the laser spot. This is to maximize the focus position where the laser intensity is higher (the center of the spot).

The analysis and tests performed with this system showed that, with a 532 nm laser, the visual focusing position and the Raman focus position (in terms of maximum peak intensity) were quite well related (as shown in Figure 2-19), and thus, the visually detected focusing position is directly used as the optimal focus for the Raman acquisition.
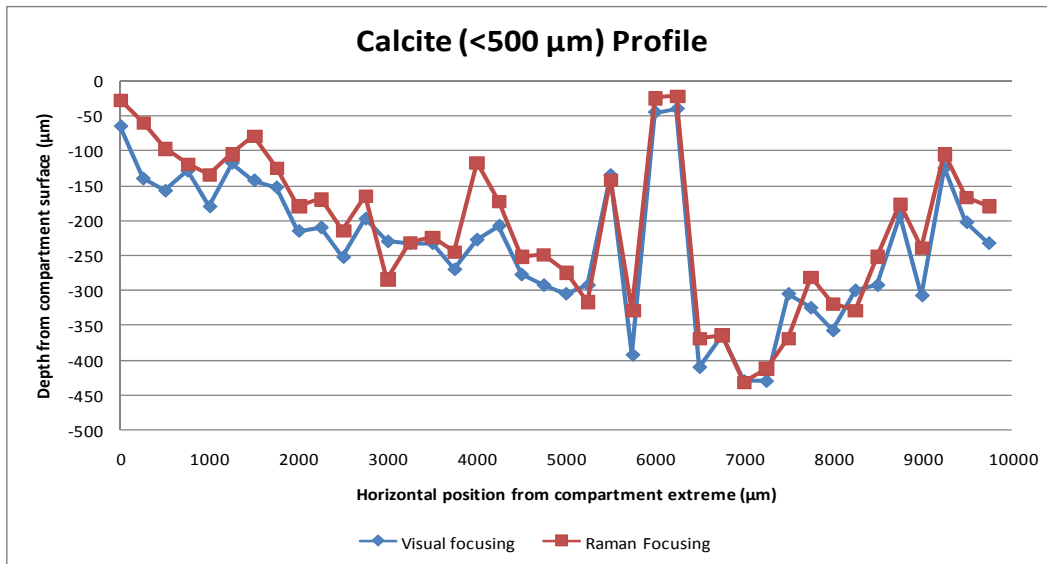
Figure 2-19. Visual vs Raman (maximum peak intensity) focus

2.5.1.9.2   HOW IT WORKS
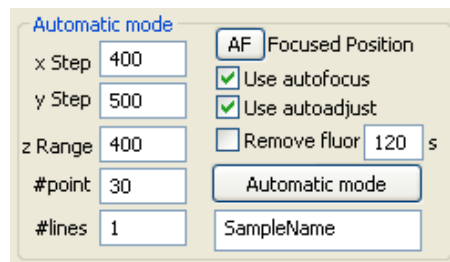
Figure 2-20. Automatic mode menu. Only AF button and z Range edit field are applicable to the manual control mode

The AF button will perform a sweep from the current z position along a range indicated in the z-range edit text box (400 microns in Figure 2-20), moving in the z positive direction (up), contrary to what is performed during the automated operation (see section 2.5.2). A new image will be acquired each 10 microns, calculating on each picture the focusing index with the technique described in section 2.5.1.9.1. Once swept, the optimal calculated focus position will be indicated in a text field and the positioner will be automatically moved to the focused position.

Precautions:

1- The 10 microns step is optimized to be used with the 50X image. To focus the other camera, it is more practical to do it manually, given that the 10X objective depth of field is very long.
2- User has to make sure the laser spot has been indicated in order to focus exactly on the laser spot position, and not on the center of the image.

2.5.2 AUTOMATIC OPERATION
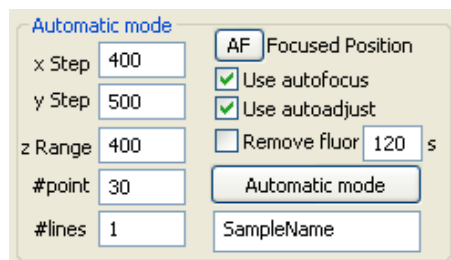
The integration of all the elements under the same software provides the possibility of developing a software tool capable of performing all the activities in automatic mode, provided that there are algorithms that can take live "smart" decisions, based on the results from the current operation. The final aim is to develop, integrate and test all the necessary algorithms that the instrument might need to perform a full-scale operation in an automated way, while optimizing both the spectral quality and scientific results, and the operation time and power consumption. To do so, several algorithms have been proposed (see Chapter 3), and some of them developed and integrated into the RLS ExoMars Simulator.

The nominal operation of the RLS instrument onboard the ExoMars rover consists on acquiring 20 to 39 spectra along a line of a powdered sample, autofocusing at each point, and autoadjusting the acquisition parameters to adapt to the sample characteristics, in order to save time when possible. Each point is limited to an acquisition time (including the dark frame acquisition) of 5 minutes per point.

Figure 2-21 shows the control panel for the automated mode. During the automated operation, the acquisition results, including logs and acquired spectra will be saved in the folder C:/Resultados_SimuladorExomars/AutomaticMode/SampleName, where SampleName is the name indicated by the user in the sample name text field. In case the folder already exists, the data will be overwritten.

To start the execution of the automatic mode, the Automatic mode button has to be pressed, once all the parameters have been correctly set.

NOTE: To ensure the quality of the automatic mode functioning, the desired spectrometer tab has to be selected (without initializing the spectrometer), the laser that is intended to be used, the 50X radio button and the 50X camera have to be activated and the laser spot position correctly calibrated.

Figure 2-21. Automatic mode control panel

2.5.2.1   AUTOMATED POSITIONING: 2-D MAPPING

In its current version, the automatic mode can perform 2-D mappings with Autofocus at each point. The mapping grid is created based on the text fields on the left of the Automatic mode control panel (see Figure 2-21). The x and y step fields indicate the separation between points, the number of points field corresponds to the number of points that will be acquired in the x direction, while the number of lines indicates the number of points that will be acquired in the y direction.

The 2-D mapping is possible using all the spectrometers, though, for the Andor Newton, after each line, the spectrometer will be turned off (with temperature stabilization included). Thus, for this spectrometer, a one-line-only operation is recommended.

2.5.2.2   AUTOFOCUS DURING AUTOMATED OPERATION

The z range field defines the range in which the Autofocus will be performed. Contrary to manual mode autofocus described in section 2.5.1.9.2, during the automatic operation the z Range will be applied in the z negative direction. That is, the z axis will be swept between the original position minus the z Range value, to the original position. This is done to facilitate the sample preparation, as it will only require focusing the image on the sample container surface, providing an easy way to prepare the setup. The rest of the autofocus operation and the theoretical principle is the one described in section 2.5.1.9.

The automated operation will perform autofocus at each point of the 2-D grid only when the 'Use autofocus' check-box is selected. When unchecked, all the spectra will be acquired at a constant z position.

The combination of the 2-D mapping and the autofocus capability allows acquiring surface maps of the samples, both with the profile as well as the Raman information of the sample (with a spatial distribution indicated in the x-y grid parameters), as exemplified in Figure 2-22. However, this kind of operation might require very long operation times.
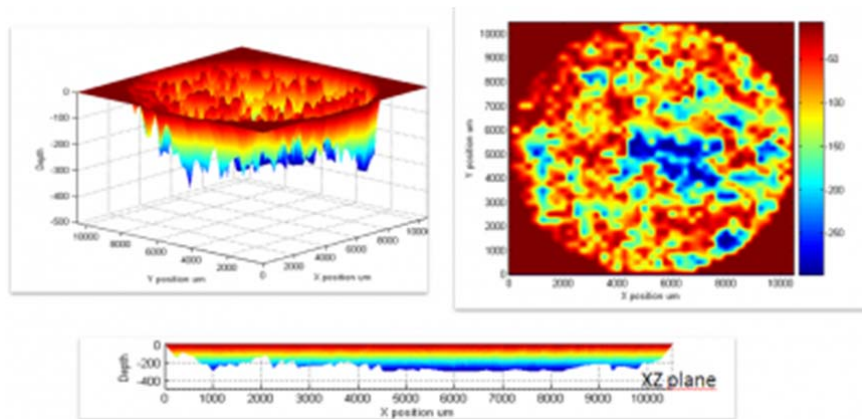
Figure 2-22. Example of 3-D mapping. This picture shows the profile of the sample on a circular sample container.

2.5.2.3 ADJUSTMENT OF THE ACQUISITION PARAMETERS

The status of the 'Use Autoadjust' check-box defines how the system will acquire the spectra. If checked, the software will acquire a reference spectrum, calculate the maximum intensity and SNR of the reference spectrum, and estimate the optimal acquisition parameters (Integration time and number of accumulations) based on these values and on the spectrometer performance, with a limit of 5 minutes per point. However, when unchecked, all the spectra will be acquired with the integration time and number of accumulations indicated on the corresponding boxes in the spectrometer tab (see Figure 2-15 and Figure 2-16). Further description of the algorithm operation can be found in Chapter 3.

2.5.2.4 FLUORESCENCE REDUCTION

The algorithm for the reduction of the fluorescence during the RLS instrument operation will automatically adjust the quenching time based on the background reduction rate at each point (as explained in Chapter 3). However, the algorithm currently implemented in the Visual Studio version of the RLS ExoMars Simulator provides a more simple approach: when the corresponding check-box is activated, the laser will be turned on after the focusing process, and the system will wait as many seconds as indicated in the corresponding text field of the control panel. This time will be applied to all points, no exception. In addition, if the spectrometer that is being used for the spectral acquisition is one of the BWTek systems, a spectrum will be acquired and saved each 4 seconds. This way, it is possible to evaluate the variation of

the sample background with time. When using the Andor Newton CCD, no spectra will be acquired during the fluorescence quenching time.

2.5.3 LABVIEW VERSION OF THE SOFTWARE

This section summarizes the modifications of the software that have been implemented in the LabVIEW based version of the RLS ExoMars Simulator software, with respect to the Visual Studio version.

The use of LabVIEW provides a better graphical interface and, more importantly, much easier integration of all the different hardware units (laser, spectrometer, cameras, positioners…). Furthermore, a LabVIEW interface will make integration with other systems such as the SPDS system much easier, for future integration and joint tests. The software interface is shown in Figure 2-23.

The LabVIEW version of the software is not capable of controlling the Andor Newton, it is only capable of controlling the BWTek spectrometers. However, in this version, it is possible to control the laser power, allowing the automation of experiments with different irradiances levels (which, however, are not in the scope of this work). In addition, a position calibration procedure was implemented to avoid damage to the flattening blade, which the system will execute each time it is started.

Regarding the implemented algorithmic for the automatic mode, in addition to an optimization of the focusing process in terms of focusing time, this version of the software includes the final fluorescence reduction algorithm, which adapts the quenching time as a function of the background reduction rate in fluorescent samples. The detailed process is presented and justified in Chapter 3.
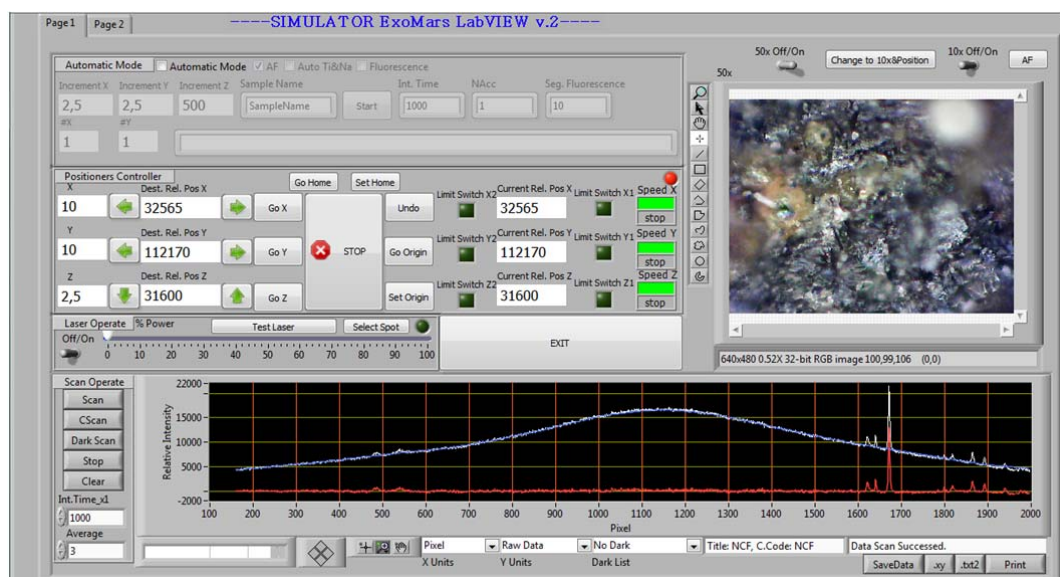
Figure 2-23. Software developed for controlling the Simulator (LabVIEW version)

Generally, the final user of the RLS ExoMars Simulator is scientific personnel, not necessarily familiar with programming techniques. On the other hand, there was a need to adjust our tools to the operational scenario of the instrument. Thus, a tool for defining activity plans for the instrument was included as part of this version of the software. Figure 2-24 shows the activity plan generation tool. The main advantage of this tool is that it allows defining any possible activity without the need to actually program anything, only by selecting among the different available functions. To ensure that the execution flow does not incur in illogical steps, a set of pre- and post-conditions were defined to ensure a safe execution.
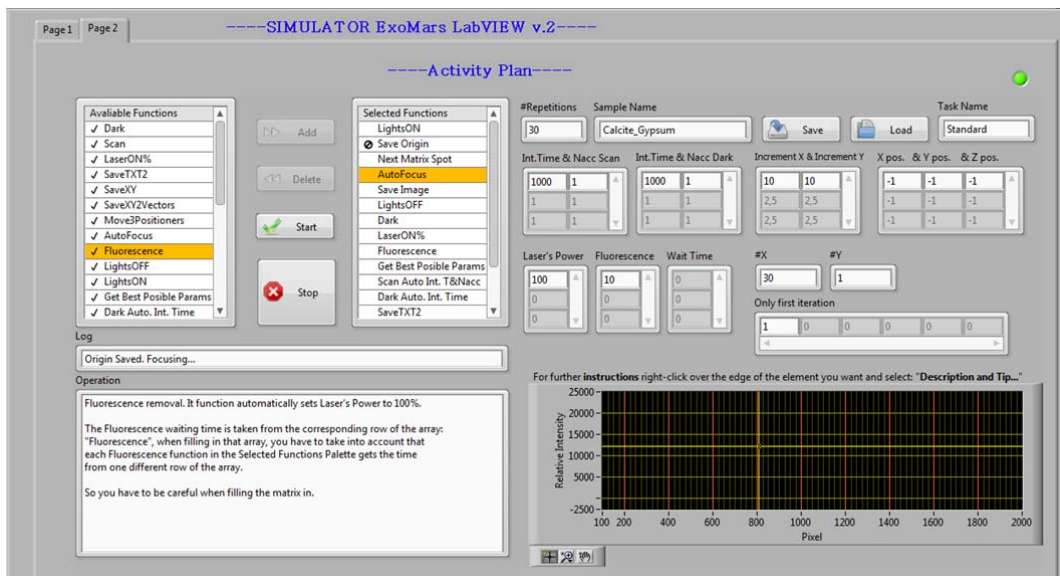
Figure 2-24. Activity plan generator included in the LabVIEW version of the software

2.6 DEVELOPMENT OF A RAMAN DEMONSTRATOR SPECTROMETER

The engineering process for the development of the RLS instrument is a very complex task, both from the technical and logistical points of view. The technical difficulties are in general related to the qualification of the system for space, which imposes very strict constraints and many test campaigns to guarantee the development process along the project life. The problems related to the logistic of a project like RLS inside ExoMars are mostly related with the interaction of several engineering and scientific teams of different institutions and countries, with different funding agencies and national managers. These issues result in a slow development process of the instrument. However, it is very interesting, from the scientific point of view, to be able to work with spectrometers with similar characteristics to those of the flight instrument.

In this framework, it was decided to build a Raman demonstrator (RAD spectrometer), which had the same technical characteristics as the RLS instrument (Figure 2-25), only without all the requirements related to space survival. Thus, we developed two RAD models with the same magnification, diffraction grating, and geometry as the RLS instrument. The main difference among them was the optics, which in the RAD models was based on components off-the-shelf, while the optics for the RLS instrument are custom made to fit the operating temperature requirements, as well as other issues related with mass, volume and weight. These spectrometers allow performing laboratory experiments to assess and validate the performance of the instrument, or at least its configuration. The prototype of the RAD1 spectrometer is presented in Figure 2-26, while the RAD2 spectrometer is depicted in Figure 2-27.

The characteristics of the end to end performance of the RLS with respect to RAD1 and RAD2 spectrometers are summarized in Table 2-1.

Parameter	RLS	RAD1	RAD2
CCD	Custom E2V	Hamamatsu S10141/1109S	Thorlabs LC100
CCD operating temperature (ºC)	-10	-9.5	Laboratory conditions
CCD size	2143 x 512	2068 x 512	2048 x 1
Pixel size (µm^2)	15x15	12x12	14x56
Optical magnification	0.7	0.7	0.7
Diffraction grating	Wasacht photonics:1800 lines per mm Input and Output angle: 32.84º, Efficiency >70%		
Spectral range (nm)	537 - 667	491 - 704	473 - 714
Spectral range (cm^{-1})	200 to 3800	-1532 to 4606	-2320 to 4791
Resolution (cm^{-1})	8 (rs < 2000 cm^{-1}) 7 (rs > 2000 cm^{-1})	12.6 (rs = 483 cm^{-1}) 8.2 (rs = 2574 cm^{-1}) 6.4 (rs = 4567 cm^{-1})	11.6 (rs = 483 cm^{-1}) 7.7 (rs = 2574 cm^{-1}) 6.4 (rs = 4567 cm^{-1})

Table 2-1. Comparative of the RLS specifications with the RAD models measured characteristics (rs stands for 'Raman Shift')

In order to provide better usability of the equipment, RAD1 model was integrated in a case to allow its transportation and to allow operating it with batteries in external sites, as for example in field trips. The RAD1 was integrated with a 532 nm laser source, with batteries and all the necessary electronics to operate it with an external laptop. The integration result can be seen in Figure 2-28. With this configuration, it has been possible to perform several field trips and travels, while acquiring spectra with a spectrometer with similar characteristics to those of the RLS instrument.
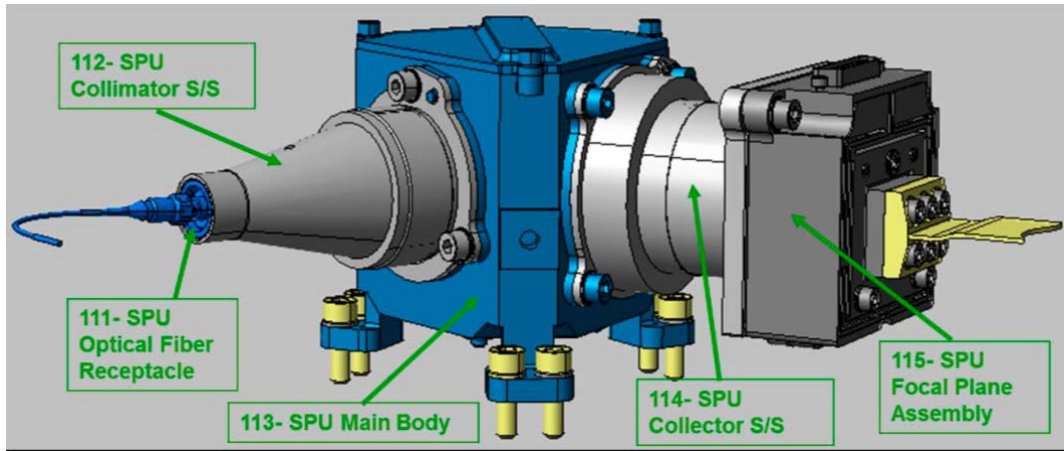
Figure 2-25. RLS instrument 3D model. Credits: INTA.
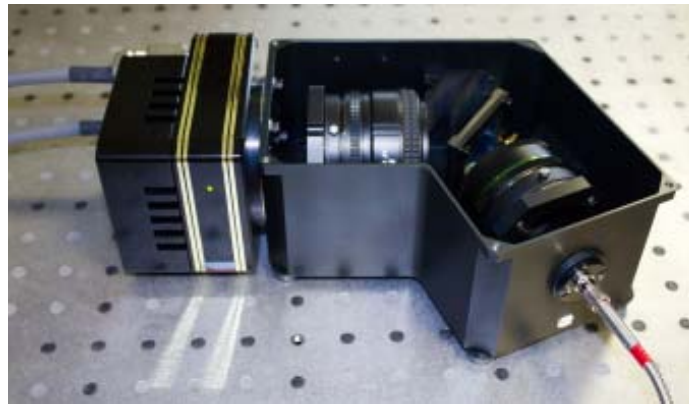
Figure 2-26. RAD1 spectrometer prototype
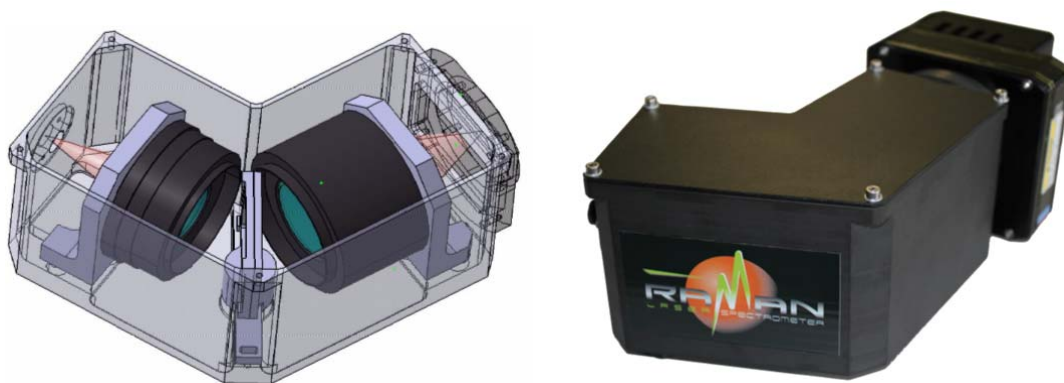
Figure 2-27. RAD2 spectrometer 3D model (left) and prototype (right)

Figure 2-28. Integration the RAD1 spectrometer in a field setup

Chapter 3.

Definition of the RLS instrument operation mode

Description of the operational framework imposed by engineering constraints, and definition of the algorithms necessary to autonomously operate the instrument

3.1 INTRODUCTION

The potential of Raman spectroscopy as a non-invasive or destructive procedure applied to all kinds of materials defines it as a very important technique for all types of materials, minerals or biological analyses. This spectroscopic technique has remained, however, unprecedented in planetary exploration, so it has been traditionally operated with the intervention of a human operator.

The operation mode of the RLS instrument is understood as the series of operations and algorithms that the instrument needs to perform autonomously for obtaining a Raman spectrum of the highest quality. This must not be confused with the development of the algorithms necessary for the instrument to become a flight instrument (telemetry, communication with the rover, data conversion, etc.). The definition of the operation mode is a purely scientific task, which shifts the actions that the human operator performs with the instrument (laser focusing on the sample, adjustment of the spectrum's acquisition time, elimination of fluorescence…) to a system able to autonomously decide which is the point of focus, to detect whether the sample under study is fluorescent or not and therefore it requires a special treatment, what will the optimal integration time be, etc.

The definition of all these algorithms is based on scientific experiments that help to define objective parameters which are measurable by the instrument, to enable it to autonomously carry out those tasks the human operator manually performs based on his experience. This experience is in general necessary to deal with the intrinsic differences found among different samples, or even different points of the same sample. These usually present strong variations in the spectrum, not only due to the existence of mixtures, but also to the different crystal orientations that can be found in powdered samples.

The onboard operation of the instrument foresees the adaptation of the acquisition parameters to the sample characteristics. This will allow saving operational time when analysing good scattering samples, which will potentially be used to analyse other points, while maintaining a total maximum operational time. The optimization of the operation will greatly increase the science return from the instrument, while not requiring more resources.

Thus, in order to provide the instrument with decision tools, the chosen operational paradigm consists on acquiring reference spectra that can be used to assess the spectral quality of the sample, and to adjust the acquisition parameters to reach a determined quality standard, saving operation time when those standards are met.

In Raman spectroscopy, there are many instrument-level parameters, especially related with the CCD device, that can be configured to optimize the acquisition and the

resulting acquired spectra, such as gain, readout speed, or even temperature. Most of these parameters can be optimized for each type of sample, to improve the performance of the instrument. However, it is totally impossible to optimize all the parameters for each sample, given that the sample will not be known in advance: thus, a thorough characterization of the instrument response is necessary to provide the overall best performance.

However, there are two operational-level parameters, integration time (Ti) and Number of Accumulations (NA), that can be easily modified without affecting the instrument performance, but the total operation time. However, several undesired effects such as fluorescence and cosmic rays can severely modify the reference spectra, and thus impair the Ti and NA calculation performance. So, these effects need to be detected and avoided or minimized, when possible.

Taking into account these considerations, we have identified several issues that need to be addressed in order to efficiently acquire Raman spectra. This chapter covers all the issues below, proposing methods for the instrument to deal with them:

- Saturation avoidance: if spectra saturate, it can lead to errors in the operation flow. Thus, it is of paramount importance that the spectra used for reference do not saturate. This issue is covered together with the saturation issue (following bullet) in section 3.2.

- Fluorescence detection and removal: fluorescence effects are commonly found when analyzing Raman spectra (though probably not as common in Mars, due to the lack of organics). Fluorescence can greatly reduce the SNR of the analyzed spectra. Given that fluorescence background can decrease when the sample keeps illuminated with the laser, in section 3.2 we propose a method for automatically calculating the fluorescence quenching time based on the background decreasing rate, saving a considerable amount of time.

- Cosmic ray detection and removal: Section 3.3 refers a proposal for detecting and removing cosmic rays and spikes that can appear in the reference Raman spectra, which could cause failures in the algorithms if considered as peaks.

- SNR calculation for onboard operations: Section 3.4 covers the development of an algorithm for the calculation of the spectral quality of a spectrum. This data will be used as a parameter of the acquisition parameters adjustment algorithms.

- Acquisition parameters adjustment: the adjustment of the integration time and the number of accumulations is covered in section 3.5, providing a method to optimize the overall operation time with a spectral-quality-based method. This will potentially allow acquiring more points of the sample.

All the algorithms proposed for onboard operation need to be computationally simple, while ensuring the best possible performance, and they have been design taking this constraint into account. The general operational flow is depicted in Figure 3-1.

```
┌─────────────────────────┐
│   Saturation and        │
│ Fluorescence detection  │
│     and removal         │
└─────────────────────────┘
            │
            ▼
┌─────────────────────────┐
│   Cosmic ray detection  │
│      and removal        │
└─────────────────────────┘
            │
            ▼
┌─────────────────────────┐
│                         │
│    Calculate final Ti   │
│                         │
└─────────────────────────┘
            │
            ▼
┌─────────────────────────┐
│                         │
│      Calculate NA       │
│                         │
└─────────────────────────┘
```
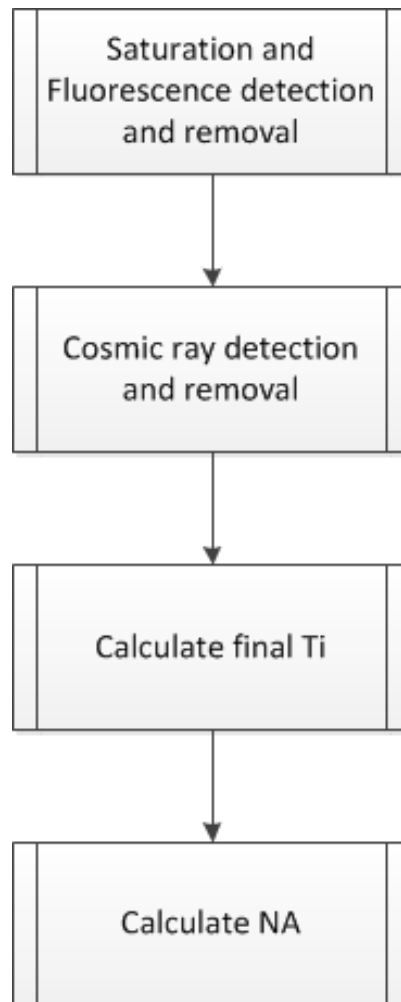
Figure 3-1. General operation flowchart

3.2 SATURATION AND FLUORESCENCE DETECTION AND REMOVAL

3.2.1 SATURATION AVOIDANCE

As already state, the operational flow of the RLS instrument will be based on acquiring reference spectra that allow the instrument to take decisions based on the spectral characteristics of those spectra. However, it is impossible to know in advance whether the samples will be good or bad Raman scatterers, or if the sample will be fluorescent. Though we have defined a relatively short integration time for the reference spectra (Ti_{ref} = 1 second), it is still possible that a spectrum saturates in such a time.

In order to avoid it, RLS will check if the reference spectra are saturated, and, in that case, will acquire new reference spectra with a shorter integration time. This is important to ensure that the rest of the operation will work correctly.

Guillermo Lopez-Reyes

3.2.2 THE FLUORESCENCE EFFECT IN RAMAN SPECTRA

When light interacts with an optically transparent material, some of the light will pass through the material without interference (transmitted fraction), but part of it will be scattered in all directions. The scattered light basically consists of two types: One, called Rayleigh scattering, has the same frequency as the incident beam (w_0), and occurs with high probability (recombination at the excitation wavelength is very probable). The other type of emission, called Raman scattering, is very weak (probability in the order from 10^{-5} and 10^{-13} [37]) and has frequencies $w_0 \pm w_{fi}$, where w_{fi} is a vibrational frequency of a molecule. This is called the Raman effect. The difference in the incident vs. emitted frequencies is due to the fact that in Raman spectroscopy, the excited state is far below the first electronic excited state, in a "virtual state", different from the real excited state (see Janblonski diagram in Figure 3-2a). These virtual states correspond to molecular or crystalline network vibrations, which are specific of the excited material, so these emitted photons will conform the Raman response of the material.
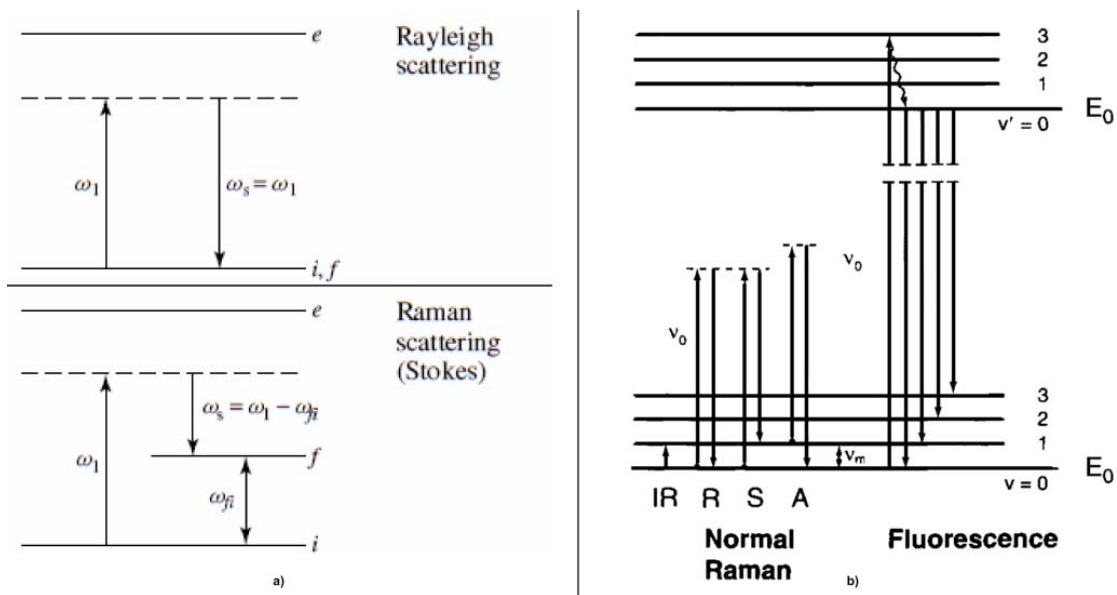
Figure 3-2. Transitions for the Raman and Rayleigh scattering and fluorescence effects
(a - extracted from [38]; b - modified after [39])

However, apart from the inelastic scattering that generates the Raman bands, there are several other effects also induced by the excitation of the sample. One of particular relevance for its influence on the acquired spectrum is fluorescence. Fluorescence is a form of luminescence, which is the emission of light from any substance, occurring from electronically excited states. As explained by [40], "in excited singlet states, the electron in the excited orbital is paired (by opposite spin) to the second electron in the ground-state orbital. Consequently, return to the ground state is spin allowed and occurs rapidly by emission of a photon". In other words, fluorescence occurs when the

molecule is excited from the ground state to a discrete level of the electronic excited state. Then the excited state molecule decays to the lowest energy level via radiationless transitions and then emits radiation, as shown in Figure 3-2b and Figure 3-3. As molecules may drop down into any of several energy levels caused by deep traps on the material, the emitted photons will have different energies, and thus frequencies. This causes the appearance of radiation in different wavelengths, causing the baseline of the spectrum to grow. The lifetime of the excited state in Raman is very short ($\sim 10^{-12}$ to 10^{-13} s), while those in fluorescence are much longer ($\sim 10^{-7}$ to 10^{-9} s) [39]. Thus, by using pulsed excitation sources it would be possible to reduce the fluorescence levels by resolving in time. However, the RLS instrument is based on a continuous wave laser excitation source, and this kind of fluorescence rejection is not possible. One of the main reasons for fluorescence effects in earth analysis is due to biological remains on the samples, which will probably be unlikely in Mars, potentially reducing the incidence of fluorescence in the samples. However, it is still considered necessary to be prepared for potentially fluorescent samples.

It has been observed that, in fluorescence samples, if the sample keeps illuminated by the excitation source, the lower energy states tend to saturate, making the fluorescence transitions happen with lower probabilities. This is known as photobleaching, an undesired effect when performing fluorescence measurements [40]. As a result, the fluorescence emission decays and eventually stabilizes after some time. This means that the background level of consecutively acquired spectra decreases with time and then stabilizes (see Figure 3-4). The decay and stabilization time is dependent on the sample, with times ranging from several minutes to negligible times.
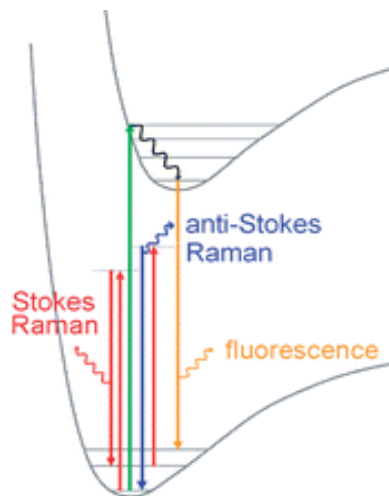
Figure 3-3. Janblonski diagram representing the energy transitions for the Raman and fluorescence effects

In general, the reduction of the background level of the samples implies an increase on the SNR of the spectrum, as shown in Figure 3-5, where it can be seen that there is an inverse correlation between the SNR value and the decreasing spectrum background level. Thus, an algorithm for the detection and quenching of fluorescence prior to the acquisition of spectra on the RLS instrument is necessary to ensure the highest possible quality of the spectra. Such an algorithm needs to deal with the problem that different materials behave differently, so it is not possible to set a general fluorescence reduction time for all samples, and the algorithm will need to dynamically adapt to the sample under analysis. This will be useful to save operation time and power resources.

This algorithm will be implemented and executed onboard the instrument during operation, so it should be as simple as possible. Given the correlation existing between the SNR values and the background level of the spectrum, we have developed a procedure which is focused on comparing the variation of the background of consecutively acquired spectra until it is small enough or until an escape variable (total acquisition time or maximum number of acquired spectra) is reached. Basing the decision on the background level instead of the SNR value of the spectra highly reduces the operational complexity and resources, while increasing the robustness of the algorithm.

In the operation cycle of the RLS instrument, this algorithm will be used at the beginning of the adjustment of the acquisition parameters and it is essential in order to assure that the final integration time is optimally calculated.
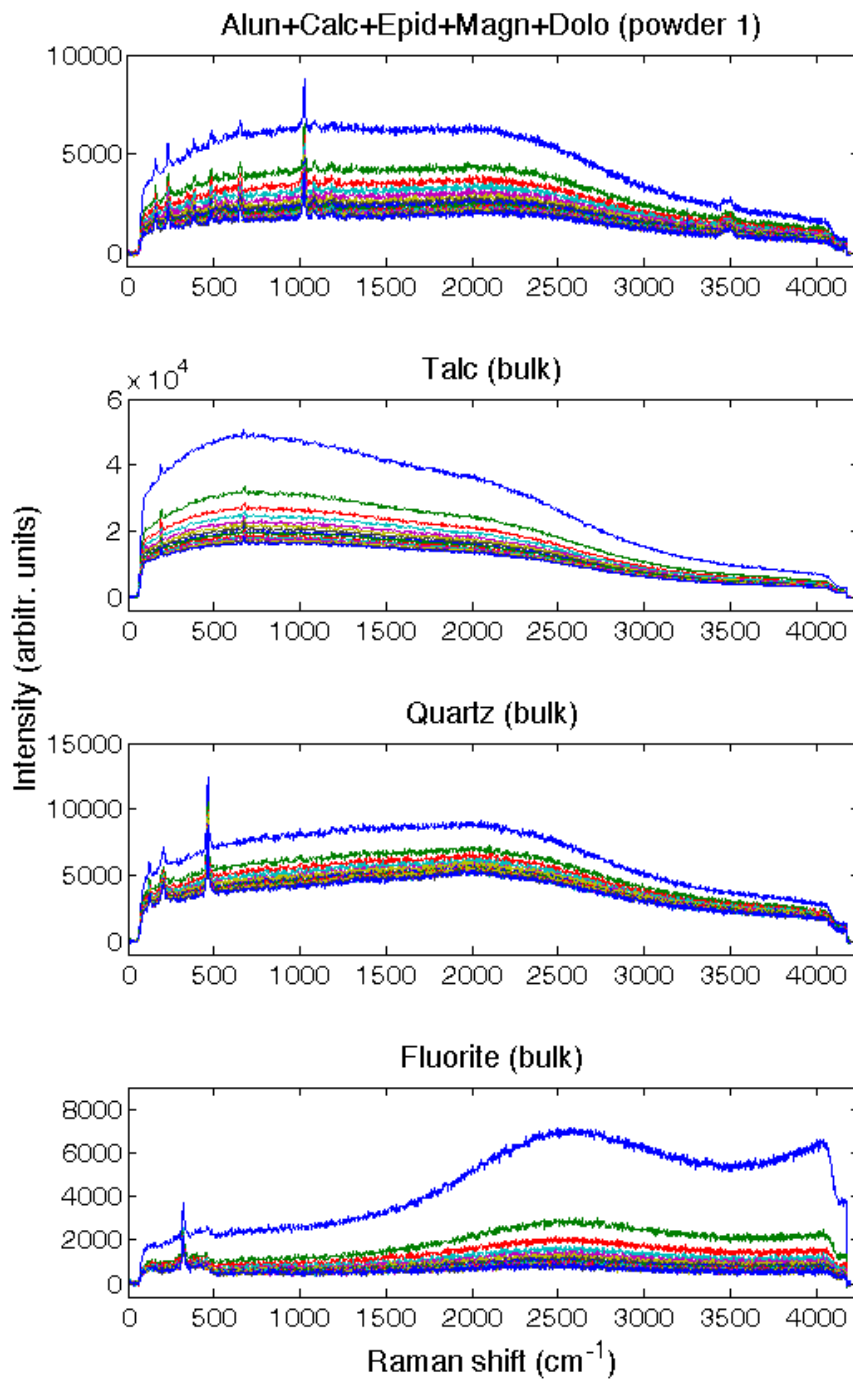
Figure 3-4. Decay evolution for different materials
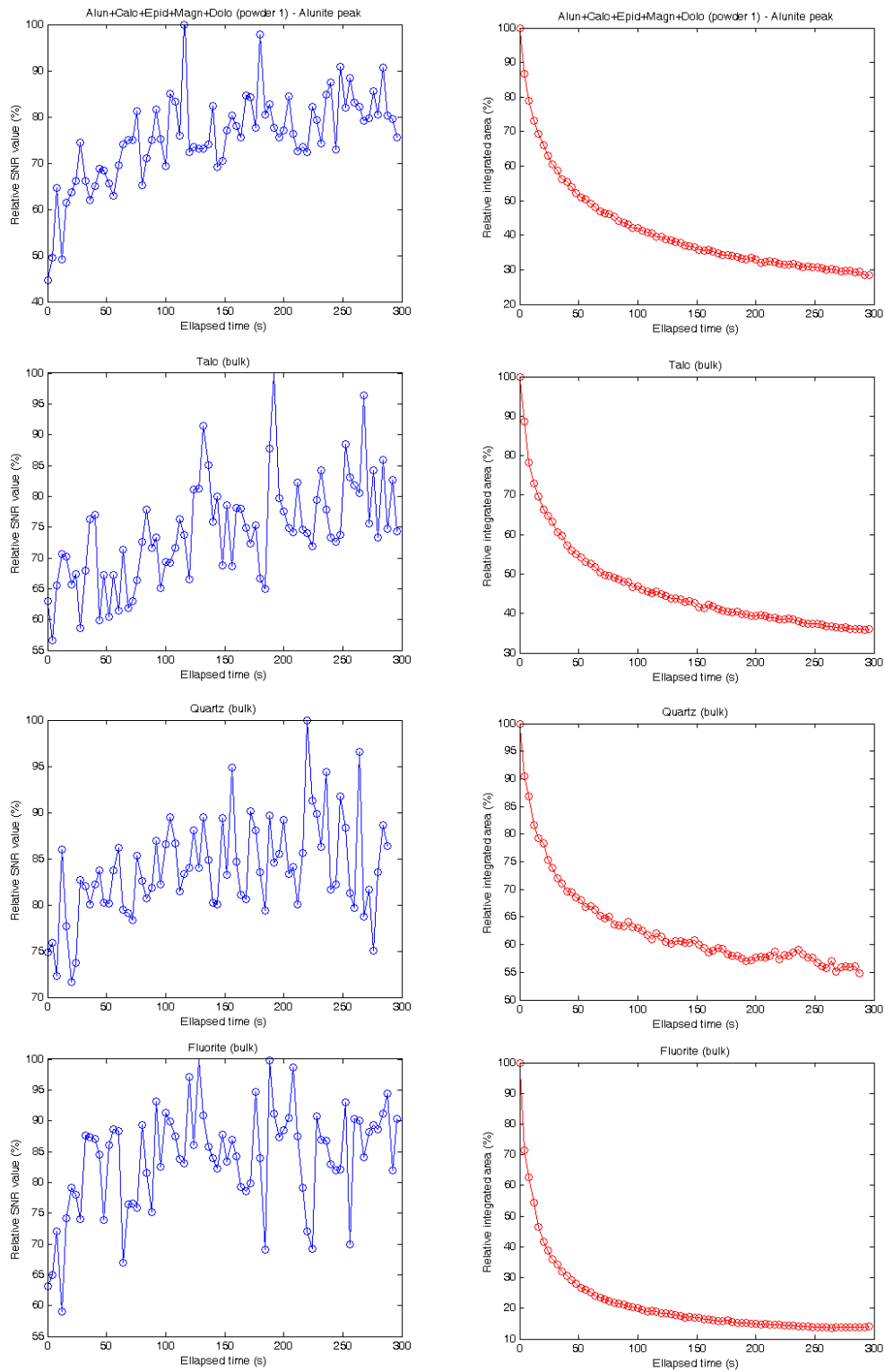
Figure 3-5. SNR and fluorescence background level evolution with excitation time

3.2.3 EXPERIMENT SETUP AND METHODOLOGY

For the development and validation of this algorithm, a selection of fluorescent samples and mixtures with different characteristics and background variation rates was used. This selection, labeled evaluation set of samples, includes several powdered and bulk samples of fluorite, quartz, talc and a mixture of alunite, calcite, epidote, magnesite and dolomite. In addition, several fluorescent natural samples from Tenerife (Las Cañadas, Las Arenas, Ucanca volcano…) were analysed to validate the algorithm response with a wider set of samples. For each of these samples, one second spectra were acquired with the RLS ExoMars simulator every two seconds during a total operation time of 300 seconds, providing a total number of 150 spectra of each material. It has to be noted that the selection of some of the samples for this chapter was based on their Raman response, rather than their relevance with respect to Mars.

3.2.4 FLUORESCENCE DECAY CHARACTERIZATION

As already explained, in general, there is some kind of inverse correlation between the spectrum SNR and the background decay. Given the computational and operational complexity that SNR calculation implies, it was decided to base the algorithm on the background decrease.

In order to characterize the decreasing rate for different samples, we calculated the integrated spectrum area for all the spectra. Figure 3-6 shows the obtained curves for these samples by representing the background integrated area every 20 seconds. As it can be seen, the decreasing rates and the final background levels reached are different for each sample, even between different points of the same sample. Under these conditions, it is complicated to extract, based on the absolute background level, an objective parameter for the algorithm to decide when the background has decreased to an acceptable level.
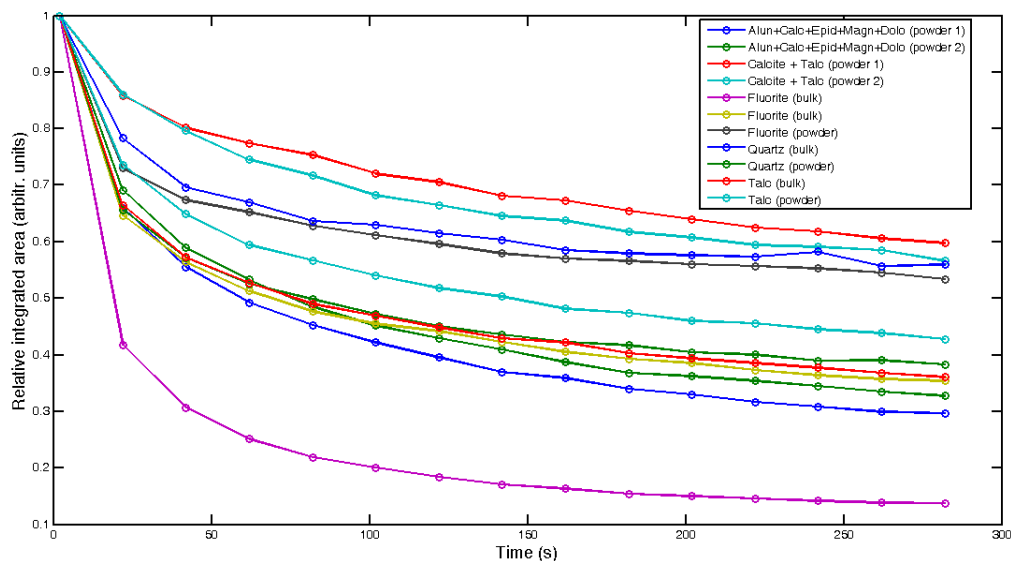
Figure 3-6. Relative integrated area for different materials

However, the background decrease rate (background level with respect to the previous spectrum), represented in Figure 3-7, shows a certain convergence for all the samples. Thus, by using this parameter instead of the relative background level, it is possible to define a general algorithm, as all the samples present a similar behavior with converging values. The proposed algorithm will define a threshold for all the samples which provide a trade-off between time and background decrease, and it will illuminate the sample until the threshold value is reached, when the background level is considered to have reached an acceptable level.

It is important to note that, even if the threshold value is reached, this does not mean that the fluorescence is totally removed. Instead, it means it is reducing at lower rates than the threshold rate. If the sample was to be kept illuminated, the background would still be reduced.
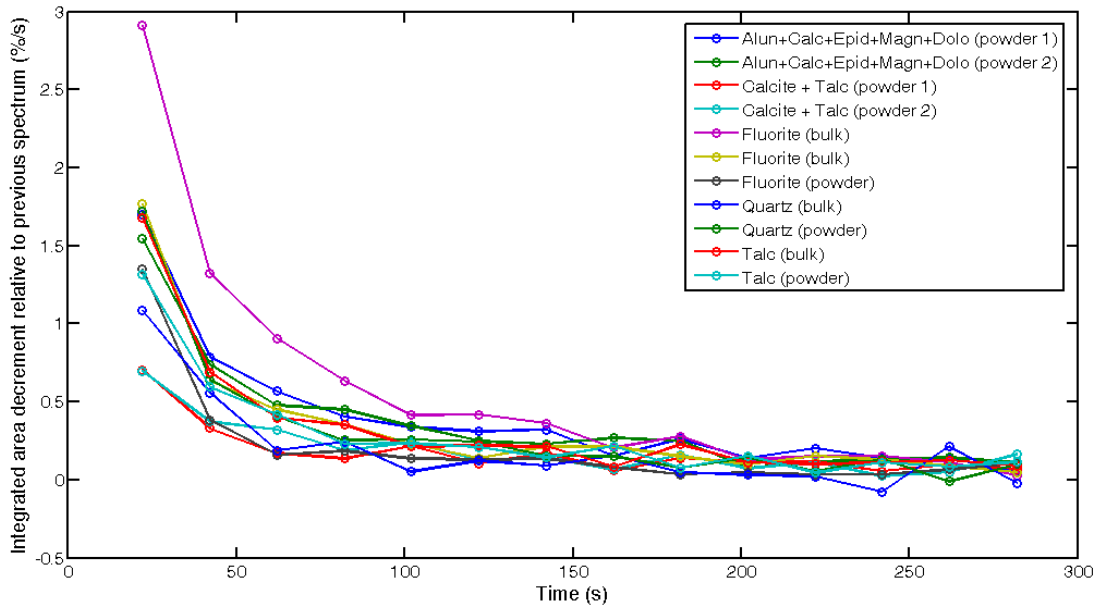
Figure 3-7. Relative area decrement for different materials

3.2.5 ALGORITHM PROPOSAL

3.2.5.1 DESCRIPTION

The total allocated time for the fluorescence quenching time is limited, so there are two possible options to develop the algorithm. The first approach is to keep the sample illuminated during the total available time, while the second approach is to dynamically determine the time depending on the sample. This second option seems to be more suitable, as it will allow saving time and resources during the fluorescence reduction process that can be used for other tasks.

With this baseline, the basic operation of the algorithm consists in reducing fluorescence by comparing the relative integrated area decrement per second [%/s] of

consecutive spectra. If the difference falls below a determined threshold, fluorescence is considered to have been removed (though not completely, as discussed previously). The process also comes to an end if the total time allocated for this operation is reached. The detailed flowchart is displayed in Figure 3-8.
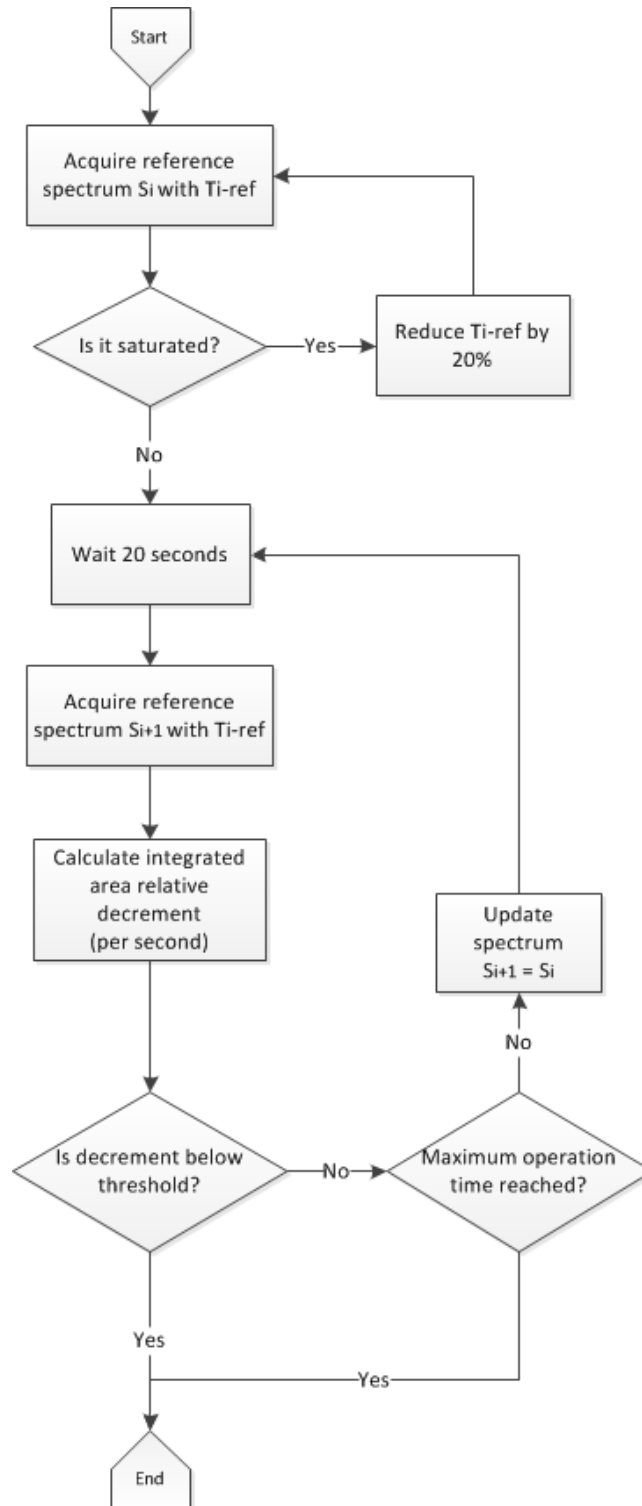
Figure 3-8. Flowchart of the fluorescence reduction algorithm

3.2.5.2   PARAMETERIZATION

The proposed algorithm is constrained by three parameters: the maximum available time, the time elapsed between consecutive acquisitions, and the escape condition (background decrement rate).

The maximum available time is defined at project level, and set from 300 to 600 seconds.

The time elapsed between consecutive acquisitions needs to be decided based on the performed experiments. This value should be short enough not to spend a lot of time for non-fluorescent samples. On the other hand, it should also be long enough to avoid over-constraining the processor with many operations, and also to remove the noise in the background levels that is observed for spectra consecutively acquired in short periods of time (see Figure 2-2). Empirically, it has been observed that acquiring spectra every 20 seconds provides a relatively clean background curve trend, while still providing a relatively short response time. The algorithm will thus acquire spectra every 20 seconds until the escape condition is reached (i.e., no fluorescence is detected, or else, maximum operation time is reached).

To define the escape condition, it is necessary to determine the optimal value to reach a trade-off between SNR and quenching time. To do so, Figure 3-9 represents both the SNR increment and the elapsed time with respect to the escape variable (the background decrement rate per time unit). This figure is obtained by fitting the SNR evolution curves from Figure 3-5 as power functions, and representing the fitted values with respect to each background decrement rate (or escape variable). In addition, the elapsed time to reach a determined background rate is also depicted.

Figure 3-9 shows how, as the background decrement decreases (higher elapsed time), the SNR increment per time unit also tends to decrease. So, the conclusion from this representation is that, to linearly increase the SNR (see regions were the SNR increment value in the graphs is almost constant, marked in Figure 3-9 with blue circles), the necessary time increases exponentially. The definition of the threshold for the background decrement rate is based on these graphs, as a trade-off between SNR increment and elapsed time. Table 3-1 presents the escape values for some samples, obtained visually from the graphs.

We finally propose an escape value that fits all the samples, to provide the best possible robustness to the algorithm. In order to account for possible errors or different behaviors from samples different than those used for this work, a margin of 40% is applied to the escape value. The final threshold value is then defined as per Table 3-2, i.e., a **reduction rate of 0.1 %/s.**

Sample	Escape value (%/s)
Powdered mixture	0.3
Fluorite	0.5
Quartz	0.17
Talc	0.2

Table 3-1. Escape variable for different samples

Escape value (%/s)	
Nominal	0.17
With 40% margin	0.1

Table 3-2. Proposed escape value
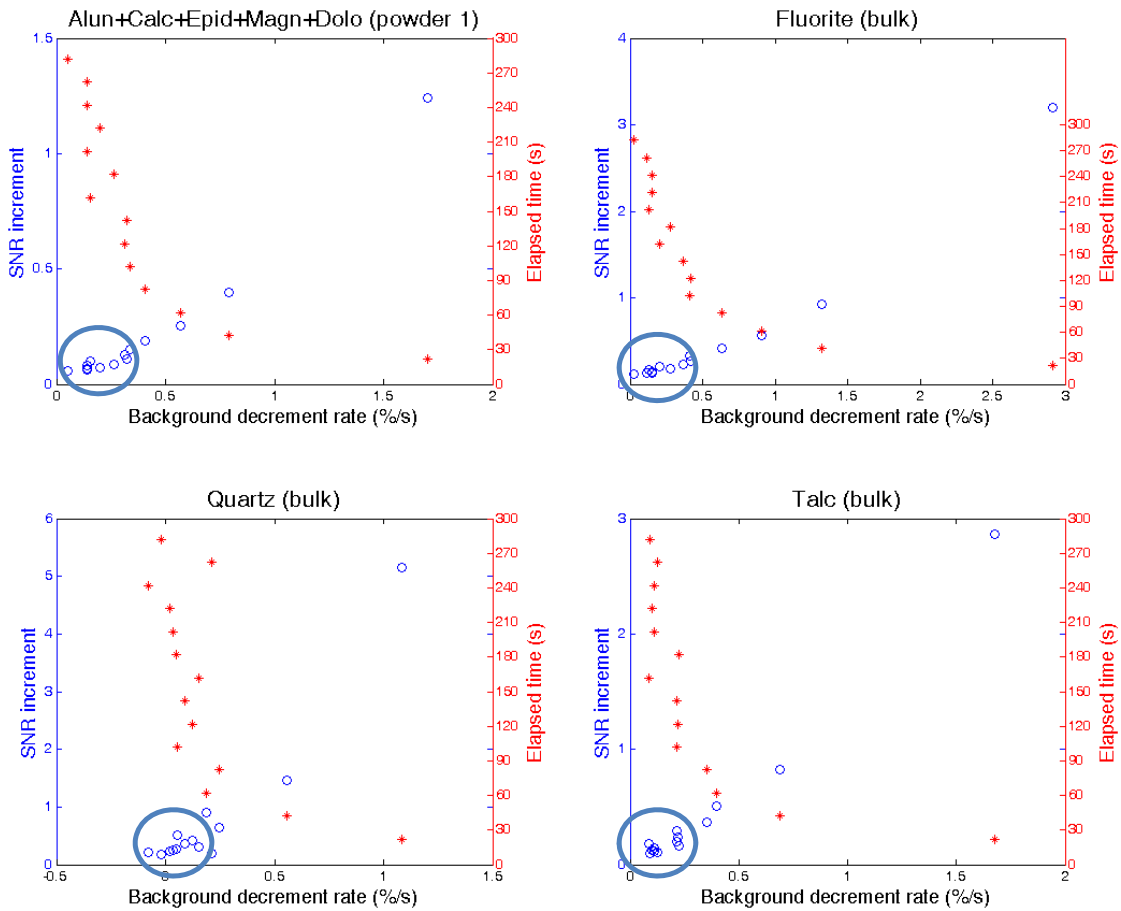
Figure 3-9. SNR increment and elapsed time as a function of background decrement rate.

3.2.5.3   VALIDATION

In order to validate the algorithm, all the samples used for the analysis were tested with the calculated escape parameter of 0.1 %/s. This provided the results in Figure 3-10, where the black rectangle includes the stopping time for the whole set of materials. Just by taking a look to that window it is clear that the aim of the algorithm (saving operation time) is accomplished for several of the samples (as they stop quenching the fluorescence after 100 seconds). In addition, we evaluated the overall saved time from all the samples used for the characterization of the problem (samples in Figure 3-10), providing an average saved time per point of around 120 seconds, assuming a maximum available time of 300 seconds. Thus, in average, the use of this algorithm would save 40% of the time allocated for fluorescence quenching when analyzing fluorescent samples (again, assuming 300 s). Furthermore, for non-fluorescent samples the algorithm would spend only 20 seconds and escape after acquiring the second spectrum, saving most of the time reserved for fluorescence quenching. Table 3-3 summarizes the potential saved time for the different spectra.
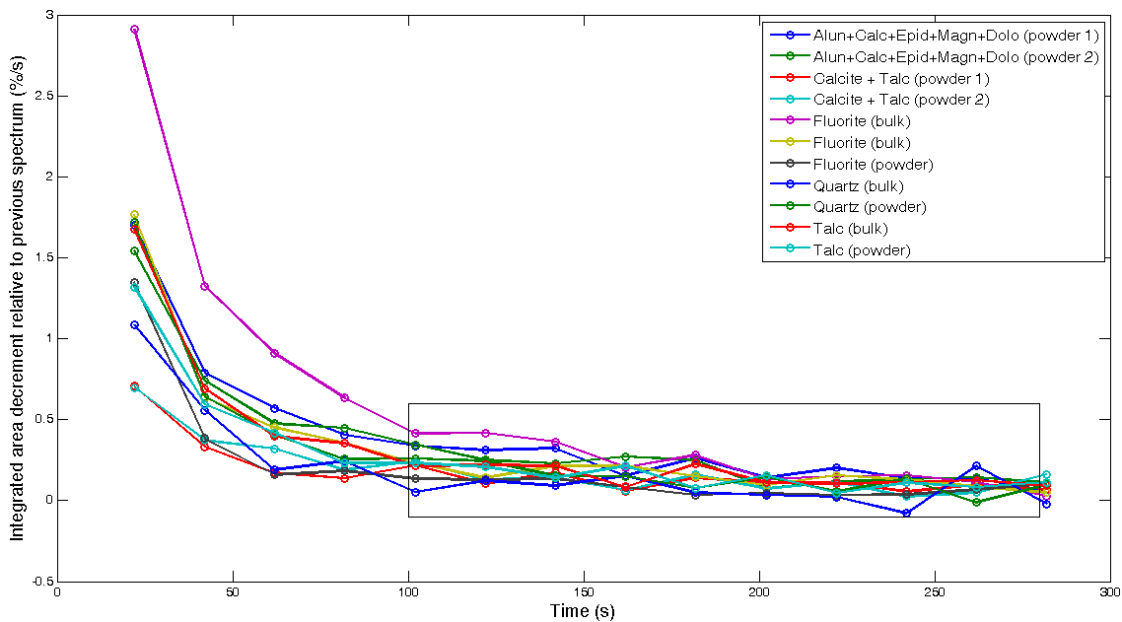
Figure 3-10. Relative area decrement for different materials, with a window showing the algorithm escaping time for the overall samples.

Sample name	Saved time (s) (out of 300 s)
Alun+Calc+Epid+Magn+Dolo (powder 1)	20
Alun+Calc+Epid+Magn+Dolo (powder 2)	100
Calcite + Talc (powder 1)	140
Calcite + Talc (powder 2)	140
Fluorite (bulk)	20
Fluorite (bulk)	100
Fluorite (powder)	140
Quartz (bulk)	200
Quartz (powder)	120
Talc (bulk)	140
Talc (powder)	120
Non-fluorescent sample	280

Table 3-3. Table showing the saved time for all
the samples used for the algorithm characterization

In order to perform a better validation with a higher number of samples, the algorithm was applied to a set of samples from Tenerife, which turned out being fluorescent samples in their majority. A total of 420 different points were analyzed, obtaining an average saved time of 147 seconds, in operations limited to 300 s. Figure 3-11 represents the saved-time histogram for this set of samples. It is interesting to note how there is a relatively high number of samples (about 10% of them) that require more than 300 seconds quenching time (all those with saved time equal to 20 seconds). This result might justify the allocation of 600 seconds for the removal of fluorescence, provided that most of them will not make use of that time (unless the escape parameter is further reduced, of course).

The relative area decrement for all the points acquired for all the samples from Tenerife are depicted in Appendix A. This analysis showed that the algorithm seems to be valid for these samples, as the obtained results (see Figure 3-12, and graphs in the appendix) show similar behavior compared to the training ones. However, it can be seen that some of the points present a weird behavior in some points (increasing values in latter spectra, e.g., blue line in Figure 3-12). These are explained by thermal effects on the sample due to the laser irradiation, which increases the black body radiation of the sample (and thus the background level) temporarily when burnt.
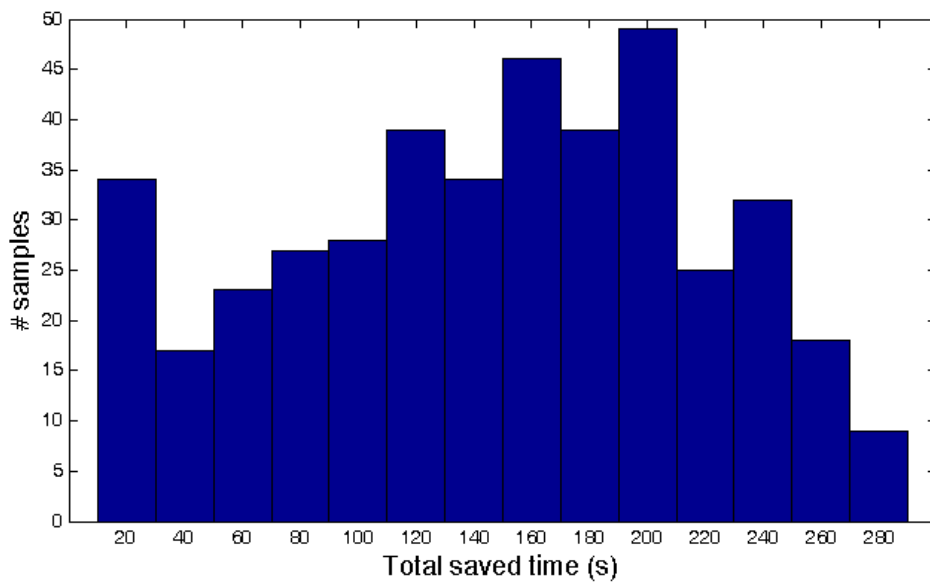
Figure 3-11. Histogram showing the saved time for the validation samples from Tenerife
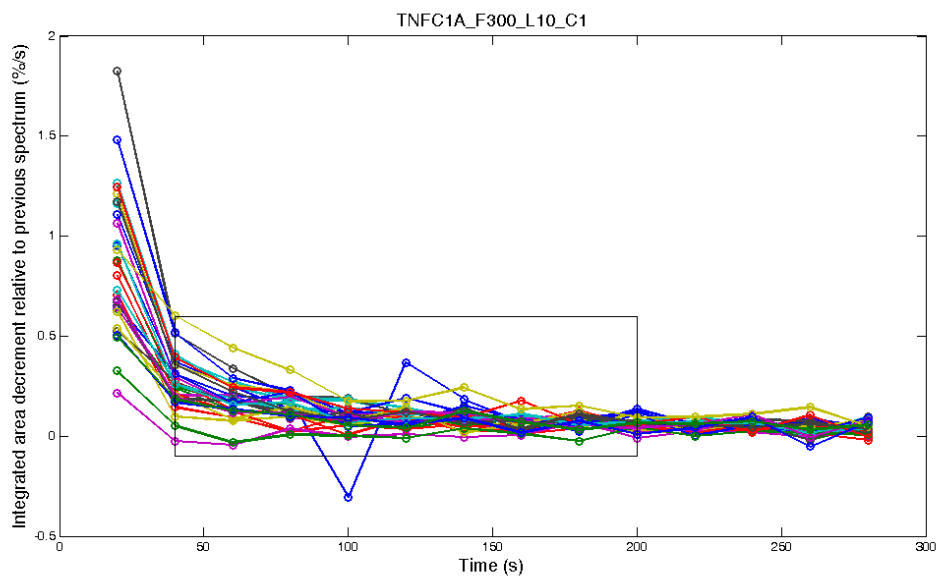
Figure 3-12. Fluorescence decay for different points of sample TNFC1A

3.2.5.4 IMPLEMENTATION

This section details the functions that are expected to be needed for the implementation of this algorithm. It is assumed that the inputs to the algorithm are binned spectra, i.e., one vector of Raman intensities.

calculate_area_spectrum(spectrum) – Calculate the area of the spectrum. The direct addition of all the binned pixels of the spectrum is a perfectly valid approximation to the spectrum area (approximation of Riemann's integration).

calculate_relative_decrement(area_spectrum1, area_spectrum2, time_between_spectra) – Obtain the relative integrated area decrement per second.

compare_decrement_to_threshold(decrement, threshold) – Function to determine whether the decrement is below the threshold or not.

3.3 COSMIC RAY DETECTION AND REMOVAL

3.3.1 COSMIC RAYS AND SPIKES IN RAMAN SPECTRA

Silicon detectors, including CCDs, are quite sensitive to high-energy radiation from local or extraterrestrial sources. So-called cosmic rays can in fact be cosmic rays but may also be background high-energy radiation from the lab or the CCD housing. Such events are generally infrequent but have sufficient energy to generate many electrons in the Si substrate that are stored and analyzed as if they were photoelectrons from Raman scattering. Furthermore, the incidence in Martian conditions is expected to be about 50 times higher than in Earth. The result is a large signal in one or a few pixels, which appears as a spike in the spectrum. An example of cosmic ray in a two dimensional image from the CCD can be seen in Figure 3-13.

Currently, there are two approaches for the implementation of the onboard binning process. One is based on hard-coding the region of interest, i.e., the region of the CCD where the signal is projected. The other one proposes to bin the n maximum values of the whole column, where n is the number of rows to bin. The second option is more robust in terms of inaccuracies or misalignments along the optical path. However, Figure 3-13 shows that, with this approach, the probability of including cosmic rays in the binned spectra is much higher.
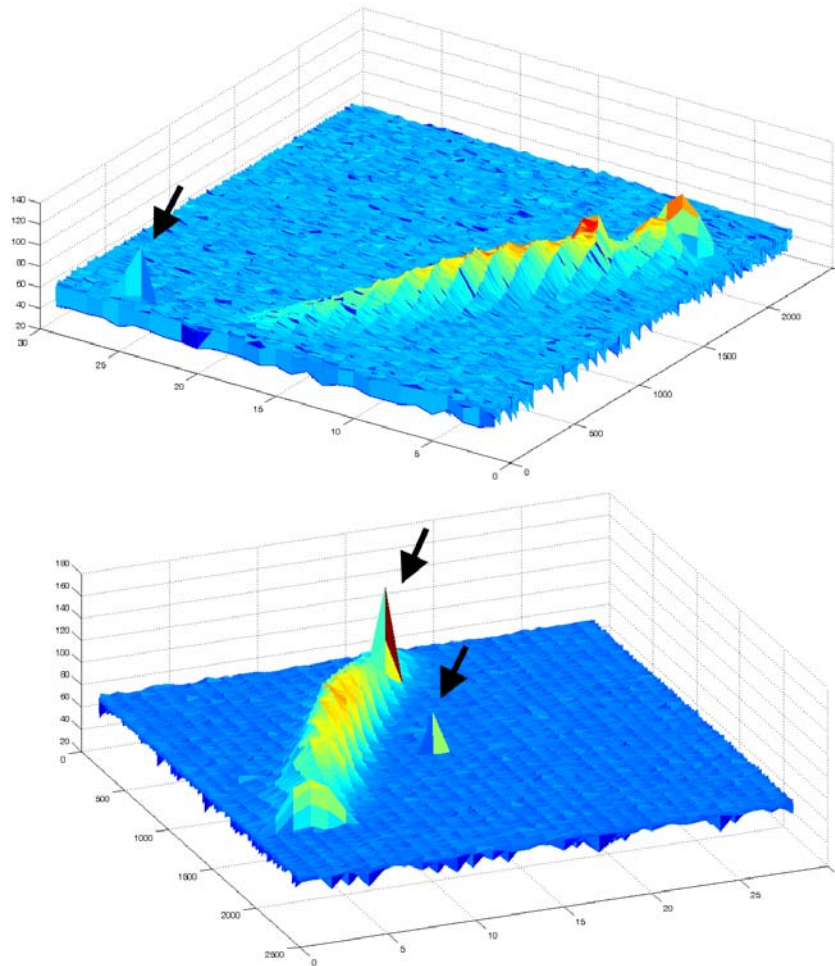
Figure 3-13. Cosmic ray examples in 3D spectra (black arrows).

In addition to positive spikes, it is observed that the CCD present fairly intense negative spikes, somehow as "negative" cosmic rays. These spikes are also required to be detected and removed by the algorithm, as they can affect the baseline calculation.

In the operation cycle of the RLS instrument, this algorithm will be used twice during the adjustment of the acquisition parameters. Firstly, it will be used to remove during the calculation of the final integration time, on reference spectra of 1 second. Final Ti calculation (section 3.5) is based upon the maximum intensity of the spectrum, thus, if the spike happens to be the highest value of the spectrum, it could lead to an inefficient acquisition. Secondly, the algorithm will be applied on the first two spectra acquired with the final Ti, before using these spectra for the calculation of the number of accumulations (as explained in section 3.5). As these spectra will be used during this phase for the calculation of the baseline of the spectrum, it is important to reduce spikes, which can impair the algorithm performance.

Literature proposes several approaches for the removal of cosmic rays [41-45]. In summary, two different paradigms can be used for their elimination: by filtering the spikes out, or by comparing several same-condition acquired spectra. [37] states that detection based on multiple acquisitions perform better than filtering a single spectrum. In the RLS case, several spectra will always be acquired during normal operation. Furthermore, computationally complex filters such as Savitsky-Golay or regression based filters are not foreseen to be included onboard due to their computational cost. Same thing happens with wavelet or polinomial based approaches. Thus, it was decided that a comparing paradigm for the definition of the algorithm would suit better the requirements of the instrument.

Under this baseline, two different algorithms have been proposed. The first one based on comparing several (the more, the better) spectra, taking the minimum values among all the spectra. However, this kind of algorithm does not detect negative spikes. The other one compares only two spectra and defines a threshold over which a point (positive or negative) is considered to be a spike. To provide a more robust performance, our proposal dynamically calculates this threshold as a function of the spectral noise. This last algorithm is the one described in this work, as, in addition to being able to detect negative spikes, it is computationally and conceptually less complex.

3.3.2 Experiment setup and methodology

For the development and validation of this algorithm, the samples acquired during a test campaign with the RLS breadboard were used. These samples are part of the breadboard spectra database, with several spectra from different samples acquired in different conditions that are representative of the operation mode. They include different types of materials such as silicon, calcite, vermiculite, gypsum, alunite, basalt, dolomite, hematite (two different types), clay, jarosite, chert, a mixture of clay and sand, quartz and gossan.

All these spectra were analyzed to search for cosmic rays and negative spikes. Several of them happened to have positive or negative spikes (either due to cosmic rays or to any other effect on the CCD, for example) that are subject to be detected and removed with the algorithm. These spectra form a subset of spectra of calcite, gossan, silicon, hematite, quartz and vermiculite that were used to validate the algorithm. The rest of the samples (those without spikes) were used to characterize the spectra in order to define the parameters (detection threshold) of the algorithm, as explained in the following section.

3.3.3 ALGORITHM PROPOSAL

3.3.3.1   DESCRIPTION

The basic operation of the algorithm consists in comparing two consecutively acquired spectra. To do so, both spectra are subtracted, and their difference is compared to a threshold related to the spectral noise, as exemplified in Figure 3-14. The detailed flowchart is displayed in Figure 3-15 (complemented with Figure 3-16).
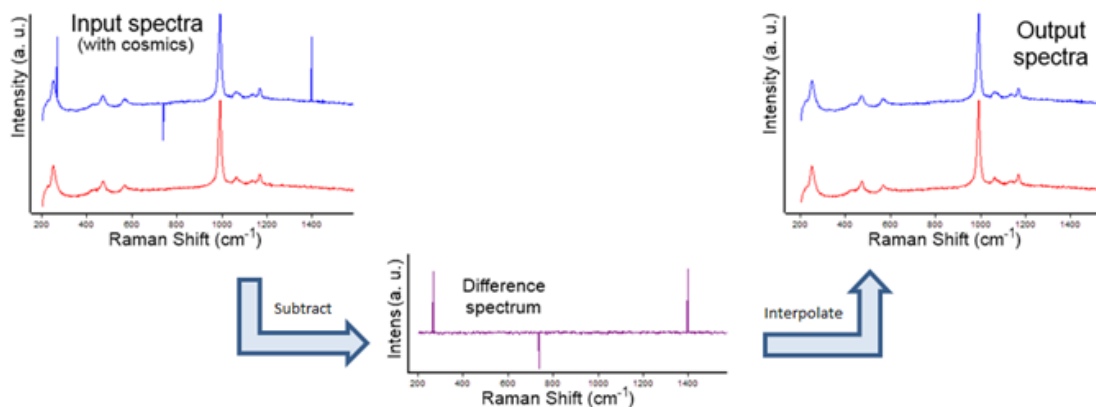
Figure 3-14. Example of cosmic ray detection

The algorithm will use two consecutively acquired spectra (the last spectra acquired during the fluorescence reduction process, explained in section 3.2, or the two first spectra acquired with the final integration time, as explained in section 3.5, depending on the operation phase). These are assumed to have almost similar characteristics, except for the spectral noise, which is then calculated from their difference. There are some issues to take into account which are covered in section 3.3.3.1.1.

Based on the difference spectrum, a decision threshold is defined to detect spectral features on the difference spectrum that are out of bounds, which would indicate the position of spikes or differences between the spectra. Then, if any spectral feature on the difference spectrum is found above (or below, for negative spikes) the threshold, the values of the original spectra in those spectral positions are substituted by the interpolated values between the limits of the peak.

The threshold is defined as a value proportional to the standard deviation of the difference spectrum, i.e., proportional to the spectral noise: a·σ, where "a" is a constant and σ the noise standard deviation. The values for the constant "a" are calculated based on the spectral characteristics of the instrument, as explained in section 3.3.3.2.2.
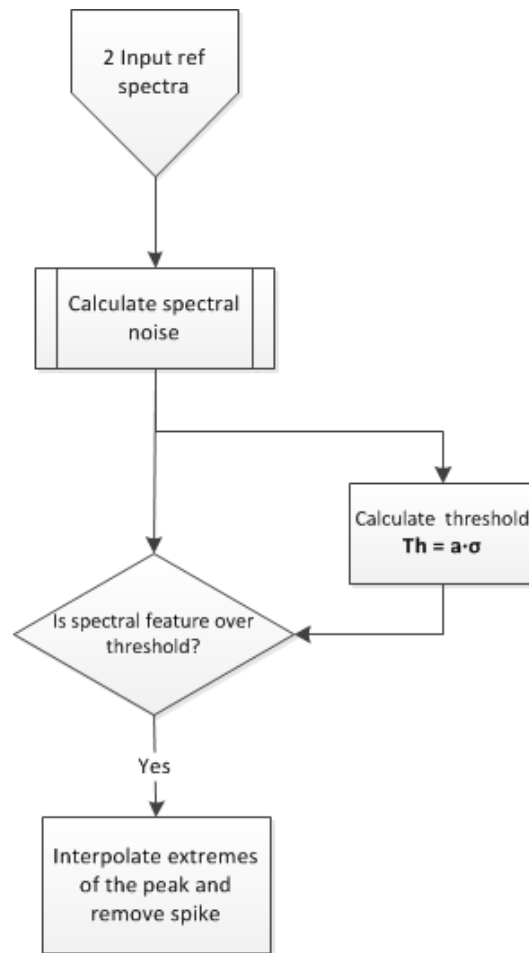
Figure 3-15. Flowchart of the cosmic ray detection algorithm

3.3.3.1.1 SPECTRAL NOISE VALUE CALCULATION

The noise value will be calculated based on the difference of two consecutively acquired spectra (see Figure 3-17). However, for some samples, two consecutively acquired spectra can present slightly different baselines, especially when dealing with fluorescent samples (see example in Figure 3-18). This can falsely affect the noise value, leading to errors that, in some cases, can be very important.

To deal with this issue, an intermediate step is proposed before calculating the noise standard deviation, which is to filter the spectrum to extract the "baseline" of the difference spectrum. By subtracting this "baseline" to the spectrum, the difference spectrum is always centered around zero.

The filter that needs to be applied is a boxcar filter (moving average) with 151 window size. This value is enough to remove most of the variation of the difference spectrum,

providing a good "baseline" of the difference spectrum. By removing this curve, the difference spectrum is unbiased, providing much more accurate results. This approach is therefore always included as a baseline for the noise calculation routine, as per the flowchart in Figure 3-16. An example showing how this subroutine works is displayed in Figure 3-18.
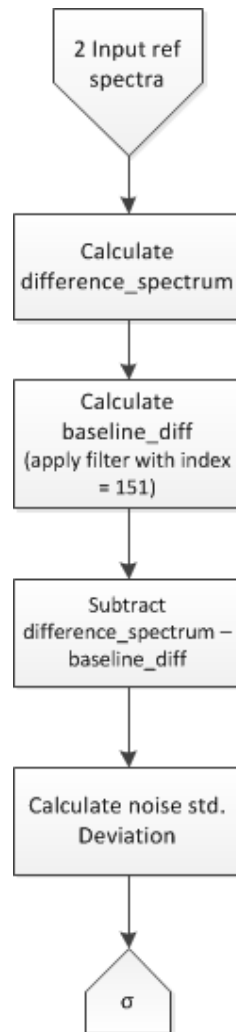
Figure 3-16. Spectral noise calculation flowchart
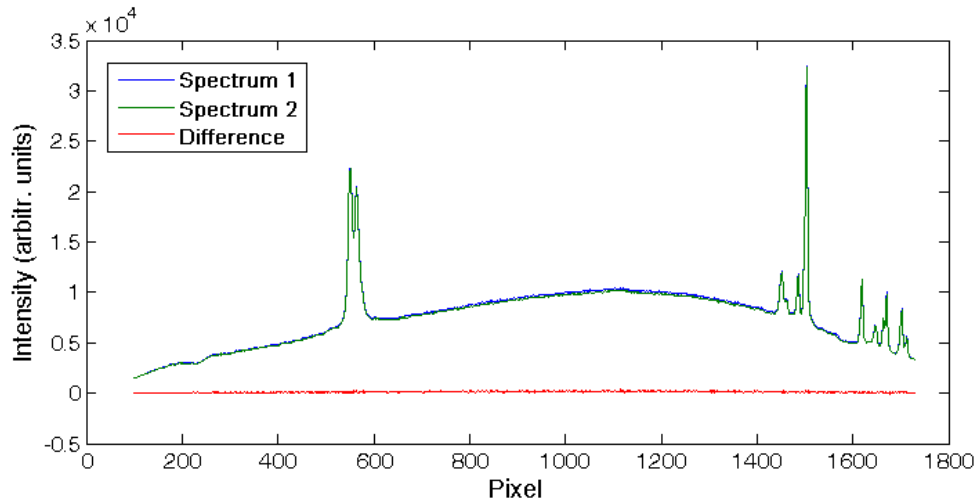
Figure 3-17. Example of difference spectrum calculation (Alunite P1)
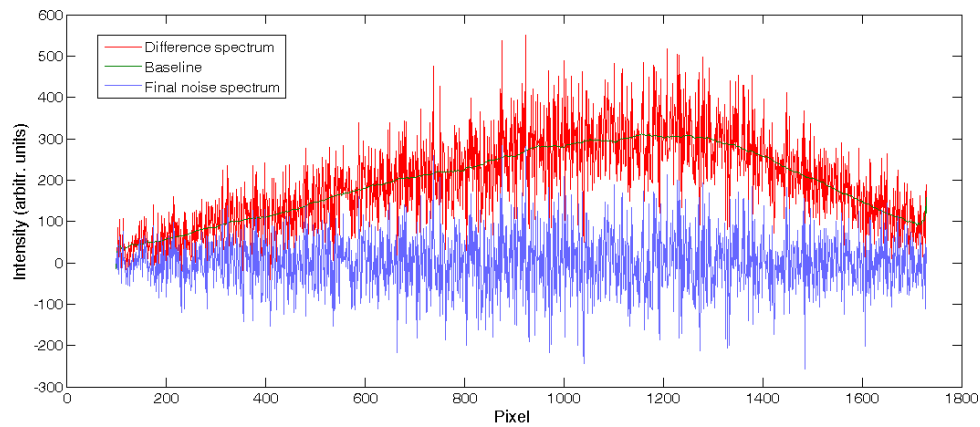
Figure 3-18. Example of offset removal from difference spectrum

3.3.3.2   PARAMETERIZATION

The cosmic ray detection and removal algorithm is based on detecting differences between consecutively acquired spectra. To do so, it is necessary to define thresholds over which the spectral features are considered anomalous, and thus candidates to be removed. As already explained, the definition of these thresholds is based on the noise standard deviation value in the form a·σ, where "a" is a constant and σ the noise standard deviation.

In order to calculate the optimal parameter "a", those samples from the RLS breadboard spectral database that did not present cosmic rays were used to perform a study on their noise values. This information was used to define the threshold.

The cosmic ray detection algorithm will be used twice during the automatic operation of the instrument, once with the 1 second reference spectra and once with the final Ti reference spectra, to avoid errors in the calculation of the optimal integration time, and the SNR (and number of accumulations), respectively. It is expected that these two types of spectra will behave differently in terms of noise. Thus, the threshold is parameterized for both types, providing different values for the parameter "a".

3.3.3.2.1 PROBLEM DESCRIPTION

The analysis of the difference between two spectra provides a relatively simple way to find differences between them, by detecting spectral features on the difference spectrum that are found far from the average values. A cosmic ray or negative peak would create one of these features.

In order to decide whether one of these spectral features corresponds to a spike or else to "regular" noise, the decision threshold is defined as a function of the spectrum noise. To make the algorithm the most robust possible, instead of defining the threshold as $3 \cdot \sigma$, where statistically most of the points should be found, we looked for the absolute maximum values of the difference spectrum, and related that value to the noise standard deviation.

There can be three different factors that affect the maximum difference spectrum values, apart from positive and negative spikes. These are: 1- noise levels in regions without Raman signals; 2- differences on the peak intensity of Raman signals of consecutive spectra due to shot noise, and 3- differences induced by a different background level due to fluorescence effects.

The fluorescence effect can really affect the noise standard deviation, as it provides a non-constant offset to the difference spectrum (the "baseline" of the difference spectrum). However, the influence of fluorescence is almost totally limited by applying the noise calculation method that removes the bias of the difference spectrum.

The evaluation of whether the maximum differences in the spectra are due to noise or differences in the peaks has been assessed in a case-by-case basis, as the spectral position where the difference occurs is known.

3.3.3.2.2 THRESHOLD EVALUATION

In order to calculate the absolute maximum differences among spectra, a selection from the breadboard spectra database was made, selecting only those spectra in which no cosmic rays or spikes (positive or negative) were found. This provided two different sets of samples: 87 spectra of different samples with 1s integration time, and 1542 spectra of different samples with final integration time spectra. The reason why the set of 1 s spectra has so many fewer spectra than the other one is double: on one

hand, fewer spectra are acquired as a reference than for the final acquisition; on the other, the spikes that are produced are much more noticeable in 1 s spectra due to their lower SNR.

Figure 3-19 and Figure 3-20 respectively show the results of the analysis of the noise of both sets of spectra. The boxplot representation shows, for each type of sample, the values of noise, noise absolute maximum difference (maximum limits from the difference spectra) and the magnitude of the quotient between the standard deviation and the maximum absolute difference. This figure represents the minimum value "a" that could be applied to avoid considering regular noise as a spike or cosmic ray.

The results for the 1 second set of spectra provided a minimum threshold value of 6.88σ (~7·σ), while the minimum threshold value for the final Ti set of spectra was 13.57σ (~14·σ). This value (considered the nominal one), with a margin of ~20% provides threshold values of around **8.5·σ for 1 s spectra**, and **17·σ for final Ti spectra**.
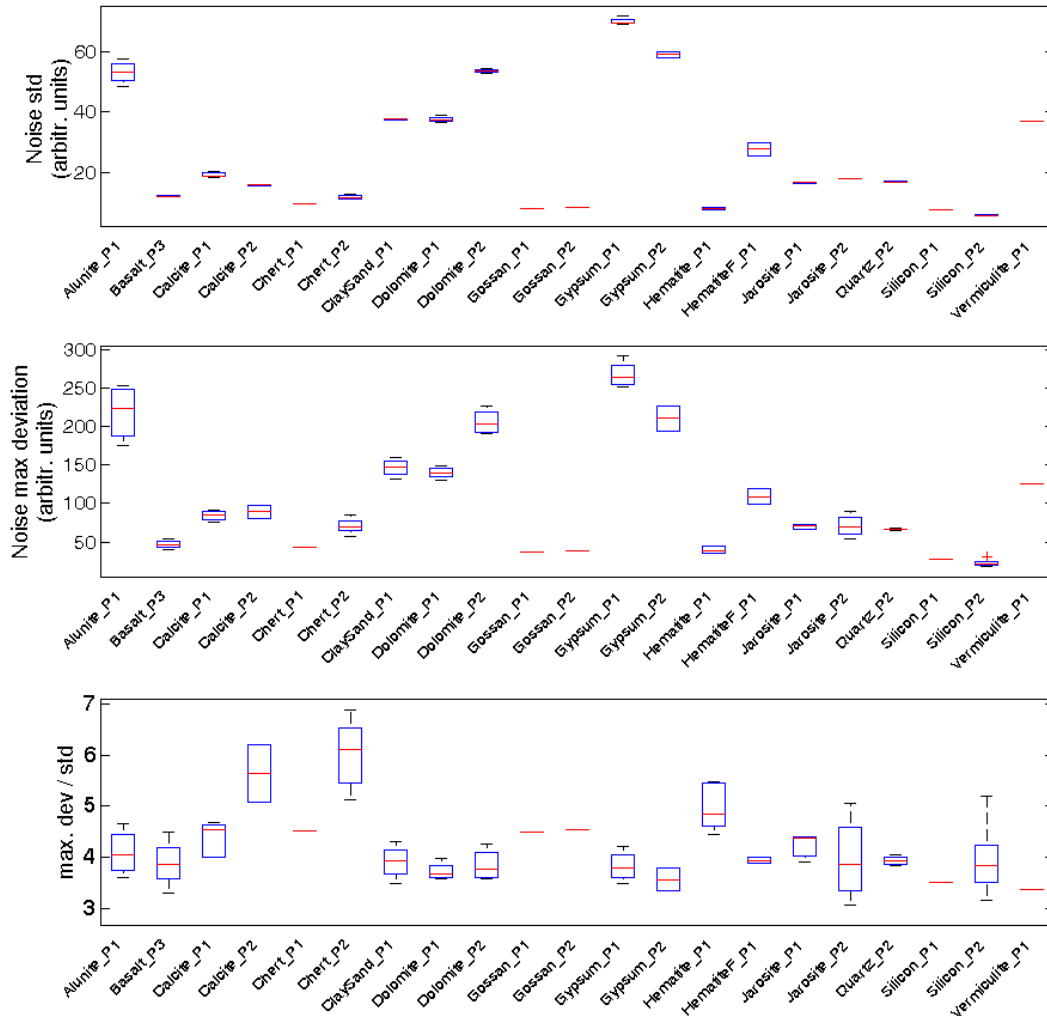
Figure 3-19. Noise standard deviation –std– (top), maximum noise difference –maxdev– (middle) and maxdev/std (bottom) for the 1s Ti set of spectra.
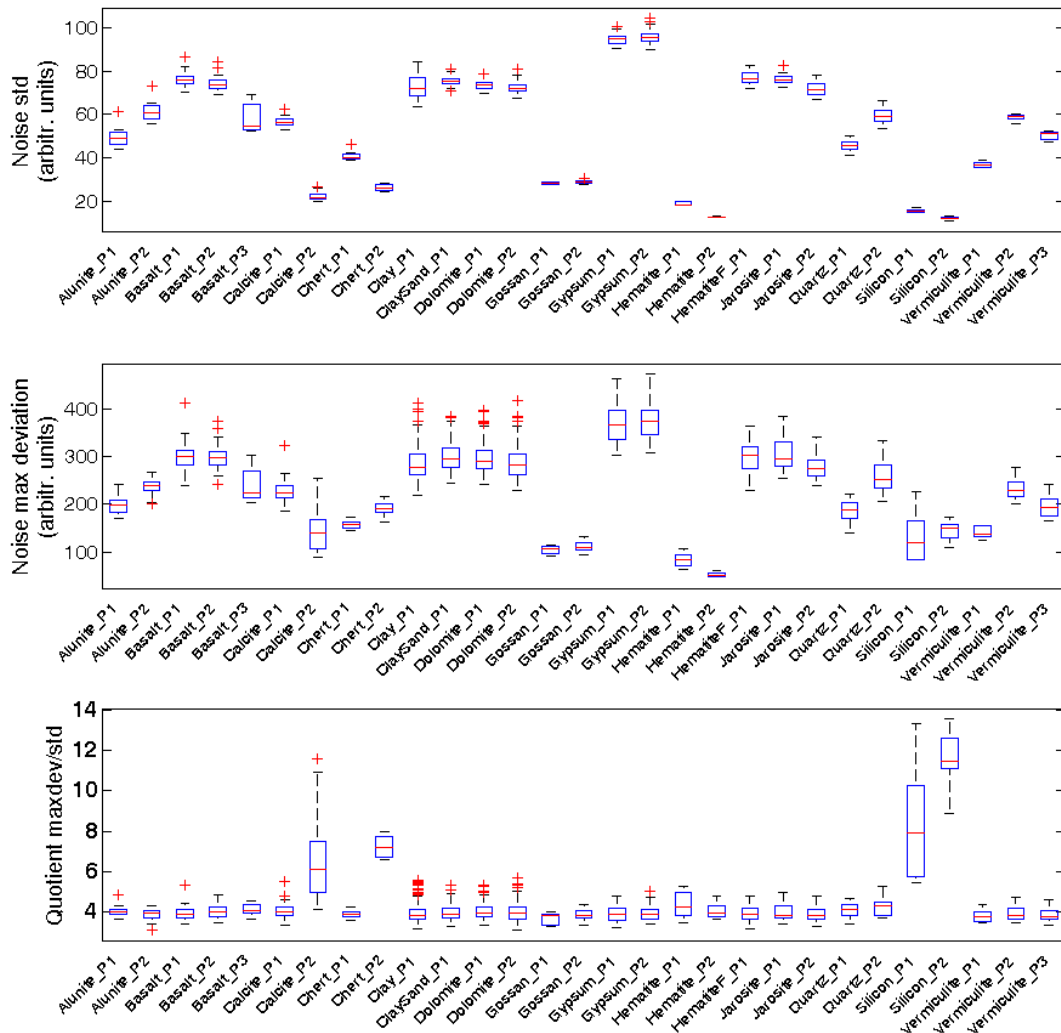
Figure 3-20. Noise standard deviation –std– (top), maximum noise difference –maxdev– (middle) and maxdev/std (bottom) for the final Ti set of spectra.

3.3.3.3 VALIDATION

Though originally this algorithm was intended for the detection of cosmic rays that saturated the spectrum (or that are many times higher than the spectral peaks), we applied the algorithm with the configuration described above to the spectra from the database that presented non-saturating spikes, to test the detection range of the algorithm. An example of the algorithm correctly detecting and removing relatively small spikes can be seen in Figure 3-21. The results show that the most potentially dangerous spikes are removed with this algorithm and the nominal threshold values (20% margin).
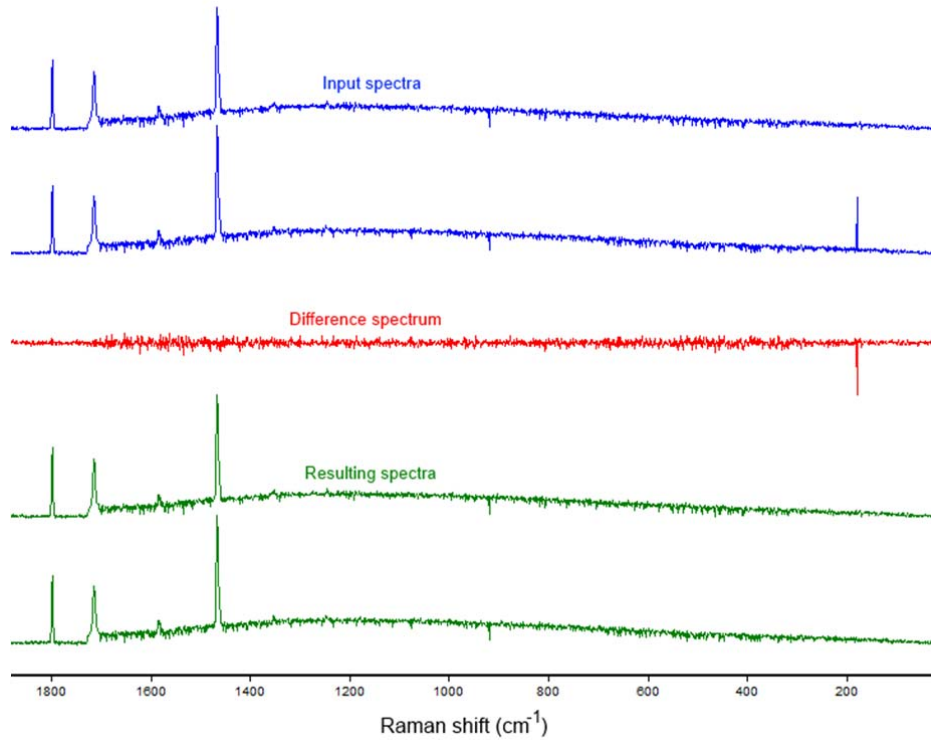
Figure 3-21. Example of spike removal

However, in some cases, some of the spikes are missed, but never with grave consequences, as the spikes that escape the threshold are below the signal value in all cases. In any case, it might be interesting to include a filtering stage (a moving average of a very small window, 3 pixels, for example) to reduce the amplitude of the spikes without really affecting the spectrum. Large windows might affect the behavior of the final Ti calculation algorithm, so it is important to keep this filter size small. The use of this post-processing stage would minimize greatly the risk of errors due to those undetected spikes, while not representatively impacting the following stages of operation.

In order to further validate the algorithm, the 1 second spectra were analyzed with the cosmic ray reduction algorithm with the threshold and final filter as proposed¸ and the results were compared to the original spectra, and also to spectra filtered with a moving average window of size 3 without applying the cosmic ray detection algorithm. The aim is to evaluate the impact of cosmic rays in the final Ti calculation, which is based on the maximum intensity of the reference spectrum.

Table 3-4 shows the results for the raw 1 second reference spectra, the filtered spectra and the results after application of the cosmic ray removal algorithm. Highlighted in the table are the cases where there is difference on the calculated parameter. Several cosmic rays were detected and removed with the algorithm. Figure 3-22 and Figure

3-23 show a couple of examples of the algorithm behavior. In addition, Appendix B shows the results for all the 1 second spectra.

As a general conclusion, and though we can see that only in one case the presence of a cosmic ray importantly affects the final Ti calculation, is that the algorithm is important to improve the acquisition. The reason is that we only have here a small subset of spectra, but also that the cosmic rays are expected to be much more probable in Martian conditions.

Position of spectrum max (pixel)			Value of spectrum max (arbitr. units)			Calculated final Ti (s)		
Original spectrum	Filtered spectrum	Cosmic rays removed	Original spectrum	Filtered spectrum	Cosmic rays removed	Original spectrum	Filtered spectrum	Cosmic rays removed
1484	1484	1484	307	307	307	54	54	54
1484	1484	1484	320	320	320	51	51	51
1484	1484	1484	346	346	346	46	46	46
1484	1484	1484	358	358	358	44	44	44
1484	1484	1484	357	357	357	43	43	43
1484	1484	1484	354	354	354	44	44	44
1484	1484	1484	348	348	348	44	44	44
1484	1484	1484	356	356	356	43	43	43
975	917	917	126	124	124	60	60	60
457	1024	1024	165	118	118	60	60	60
500	499	984	1012	331	124	14	49	60
882	1053	1053	125	120	120	60	60	60
963	904	904	124	127	127	60	60	60
1083	1026	1026	121	119	119	60	60	60
1409	1410	1410	116	114	114	60	60	60
1411	1412	1412	126	112	112	60	60	60
1411	1411	1411	117	117	117	60	60	60
1410	1410	1410	119	119	119	60	60	60
1411	1412	1412	115	117	117	60	60	60
1410	1411	1411	113	113	113	60	60	60
1408	1409	1409	120	117	117	60	60	60
1414	1409	1409	113	116	116	60	60	60
283	1409	1409	349	113	113	46	60	60
1409	1408	1408	117	117	117	60	60	60
1660	1660	1660	239	239	239	60	60	60
1660	1660	1660	241	241	241	60	60	60
1660	1660	1660	243	243	243	60	60	60
1660	1660	1660	245	245	245	60	60	60
1660	1660	1660	238	238	238	60	60	60
1660	1660	1660	241	241	241	60	60	60
1484	1484	1484	307	307	307	54	54	54

Table 3-4. Maximum peak results
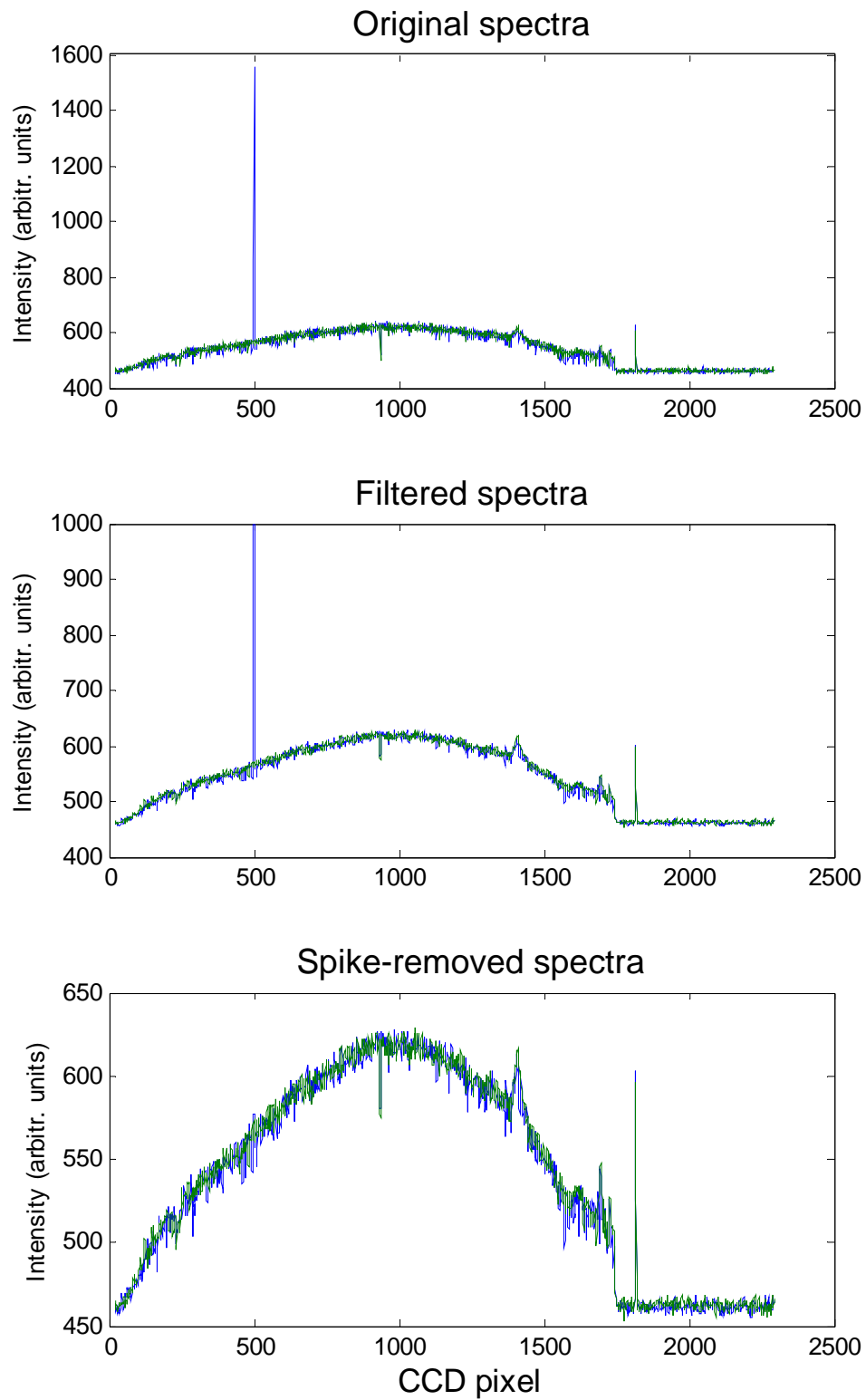
Figure 3-22. Comparative of the raw 1 s reference spectra, the filtered version and the results after application of the cosmic ray removal algorithm for a gossan sample
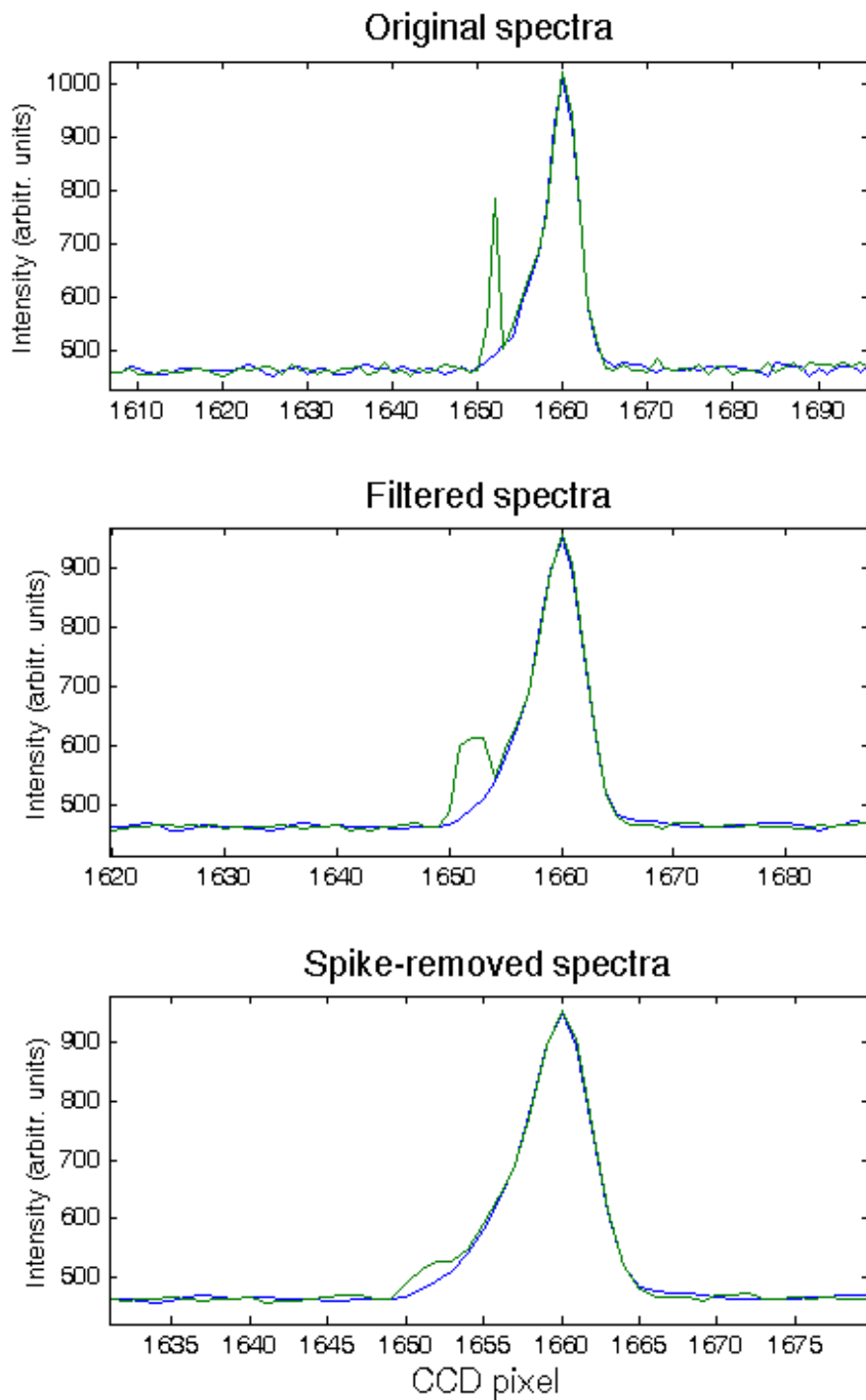
Figure 3-23. Comparative of the raw 1 s reference spectra, the filtered version and the results after application of the cosmic ray removal algorithm for a silicon sample

3.3.3.4   IMPLEMENTATION

This section details the functions that are expected to be needed for the implementation of this algorithm.

cut_spectrum_range(first_pixel, last_pixel) – Function to limit the spectral working range. Needed unless the inputs to the algorithms are already cut to this range. This function is necessary to reduce computation from spectral regions that are of no interest (laser region, for example).

remove_dead_pixels( ) – Function to remove the known dead/hot/damaged pixels from the input spectrum.

calculate_difference_spectrum(spectrum1, spectrum2) – Calculate the difference of two spectra.

smooth_spectrum(spectrum,filter_width) – Apply moving average window to spectrum.

calculate_noise_standard_deviation(range) – Calculate the standard deviation of the input spectrum in a determined range.

compare_spectrum_to_threshold(spectrum, threshold) – Function to select all points in a spectrum that are over a determined spectrum.

interpolate_spectrum( ) – Interpolate values as function of the side points.

3.4 SNR CALCULATION

The RLS instrument will analyze, in each experiment cycle, several points of unknown powdered samples. Different samples behave differently depending on their characteristics, having different scattering efficiencies. Thus, to obtain a determined spectral quality (measured as the SNR of the spectrum), different samples will require different acquisition parameters (integration time, Ti, and number of acquired images, NA). As explained in section 3.5, Ti will be optimized based on the maximum intensity of a reference spectrum, not dealing with quality issues. NA, on the other hand, needs to take into account the spectral quality (SNR) of the acquired spectra. Thus, it is necessary for this algorithm to calculate the SNR of the acquired spectra.

This section presents an algorithm to allow the onboard analysis of the SNR of the acquired spectra, to optimize the operational parameters as a function of the samples under analysis. In order to assess the spectral quality, the algorithm has to face several stages (overall operation is presented in section 3.4.1), being the most demanding one the baseline calculation which requires a deep analysis and optimization of its parameters. The baseline calculation parameterization is highly dependent on the

spectral noise. Thus, an adjustable filter which equalizes the spectral noise for all the input spectra (section 3.4.2) is applied prior to the baseline calculation (section 3.4.3). Finally, the SNR is calculated as described in section 3.4.4.

To develop this algorithm, the set of spectra acquired from different samples with the RLS instrument breadboard was used. Concretely, all the spectra with final integration times were used. Table 3-5 shows the samples, number and integration times of the spectra set.

Sample name	Total # spectra	Ti (s)
Alunite P1	16	33.47
Alunite P2	15	35.95
Basalt P1	45	13.15
Basalt P2	42	13.85
Basalt P3	13	49.7
Calcite P1	43	15
Calcite P2	49	10
Chert P1	10	60
Chert P2	10	58.57
Clay P1	295	2.55
Clay+Sand P1	159	4.14
Dolomite P1	175	3.75
Dolomite P2	301	2.26
Gossan P1	11	60
Gossan P2	10	60
Gypsum P1	80	7.6
Gypsum P2	70	8.8
Hematite P1	10	60
Hematite P2	10	60
Hematite-F P1	19	35.56
Jarosite P1	35	27.09
Jarosite P2	39	19.88
Quartz P1	10	57.29
Quartz P2	32	20.34
Silicon P1	10	60
Silicon P2	11	60
Vermiculite P1	10	60
Vermiculite P2	10	60
Vermiculite P3	10	60

Table 3-5. Spectra in the breadboard spectra database

3.4.1 ALGORITHM HIGH LEVEL DEFINITION

The basic SNR calculation algorithm needs to calculate the signal to noise ratio, i.e, it needs to calculate the intensity of the maximum peak of the spectrum, taking into account the baseline, and the noise standard deviation. To calculate the noise standard deviation, two consecutively acquired spectra are subtracted, using the method described in section 3.3.3.1.1.

Then, to calculate the signal intensity, the background of the spectrum needs to be calculated first, in order to obtain the signal intensity over the spectrum baseline. However, the baseline calculation algorithm is dependent on the input spectral noise, and a pre-processing step is added before to minimize the effect of different noise levels at the input. This pre-processing consists on the equalization of the spectral noise, as described in section 3.4.2. This way, all spectra are expected to have similar noise levels when input to the baseline calculation algorithm, which allows a better parameter optimization. The flowchart of the algorithm is depicted in Figure 3-24.
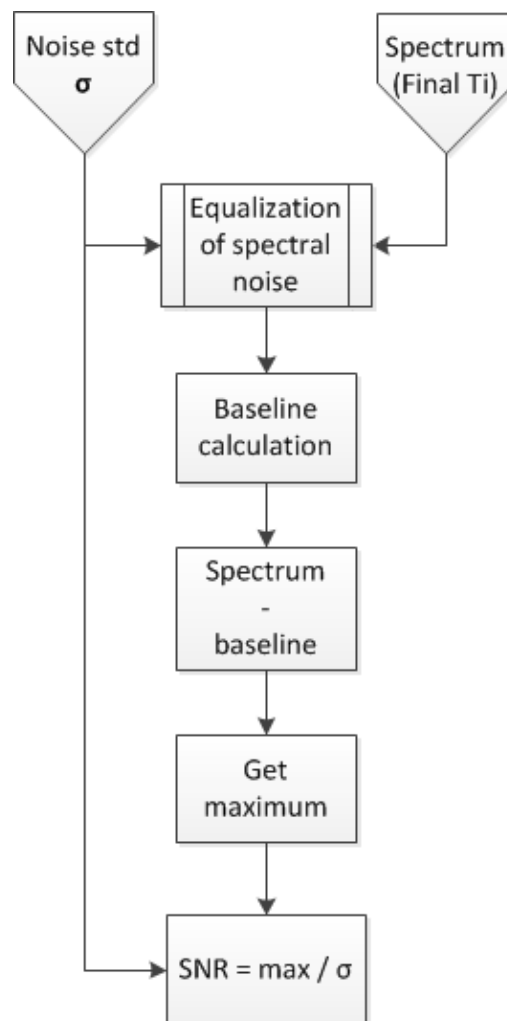
Figure 3-24. Flowchart of the SNR calculation algorithm

85

3.4.2 Equalization of the spectral noise

3.4.2.1 Description

This section presents an algorithm for the equalization of the spectral noise. This kind of equalization is interesting in the framework of the baseline removal algorithm: having the input spectra constant noise ratios, the baseline calculation algorithm will be more robust, as this kind of algorithms is usually affected by variations in the amplitude of the spectral noise.

The algorithm flowchart is displayed in Figure 3-25, and is actually pretty simple. The output spectra are required to have a spectral noise around a determined level (equalization level). This means that, if the input spectrum has a noise value higher than the equalization level, the spectrum will have to be filtered to reach that level. There are two possible ways to implement the filter: one is based on a closed loop where the spectrum is recursively filtered until the noise level is equal to the equalization level. The second one would be to characterize the noise reduction rate of the filter, so the filter is adjusted as a function of the spectral noise in order to provide the adequate noise level to the filtered spectrum. In the RLS case, given that this algorithm is to be executed onboard of the instrument, it was decided to implement the second philosophy: the computational load of the first option would make it incompatible with the requirements of the instrument. With this philosophy, once the input noise level is calculated, the necessary filter index will be obtained from a look-up-table (LUT) mapping the necessary noise decrements to the corresponding filter index.

In order to implement the chosen filtering paradigm, it was necessary to ensure that the behavior of the filters are similar for different samples, so it is possible to obtain a general function for the filter index adjustment. This issue is covered in section 3.4.2.2.1.
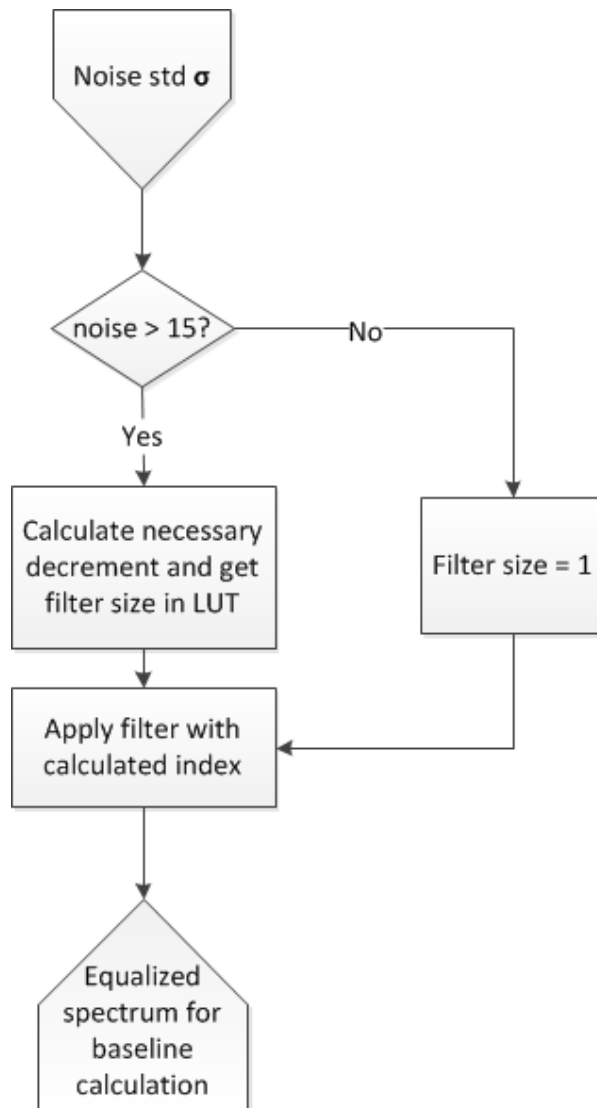
Figure 3-25. Flowchart for the equalization of spectra

3.4.2.2 PARAMETERIZATION

3.4.2.2.1 SAMPLES NOISE CHARACTERIZATION

The first step in the algorithm development consists on checking the noise-sample dependence, i.e., if the calculated noise is dependent on the sample or point under analysis, or if it is related only with the spectrometer characteristics or integration time. This is important in order to ensure that the output spectral noise can be controlled with the filter parameter, no matter the sample or the integration time.

The methodology to characterize the noise was the following: we calculated the noise for all the samples based on the difference of two consecutively acquired spectra (as per section 3.3.3.1.1), in the form 1-2, 2-3, 3-4, etc., providing a total of differences

equal to the number of spectra minus one (e.g., 15 difference spectra for Alunite P1). The difference spectrum is used to calculate the standard deviation of all points between pixels 100 and 1730, to remove the notch filtered region and the final part of the spectrum (see Figure 3-26 for an example of spectrum and the effective range used).
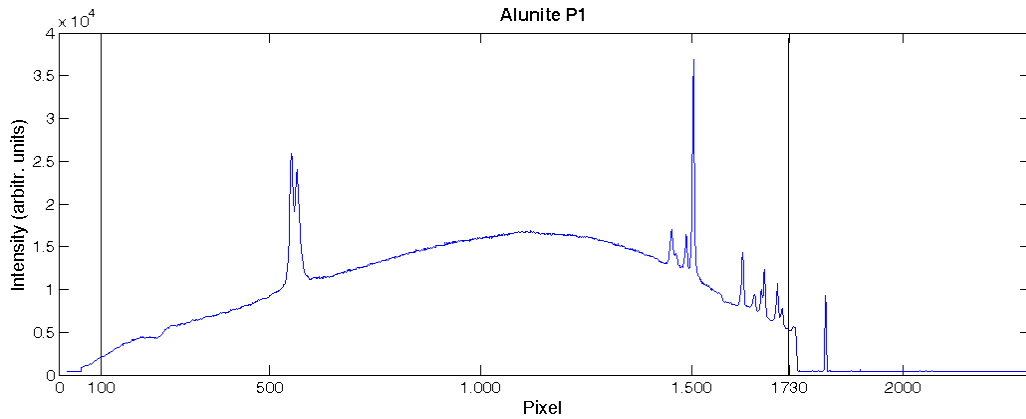
Figure 3-26. Example of spectrum (Alunite P1) and the applied spectral limits (pixel 100 to 1730)

By calculating the noise standard deviation for all the samples, the results shown in Figure 3-27 are obtained. This figure shows that the spectral noise is not only a function of the thermal noise of the instrument, but that it is also dependent on the type of sample and on the point of the sample where the spectrum is acquired. This is shown by the important variations that are found among different samples, and even between different points of the same sample.

Having these results in mind, it will be theoretically possible to equalize the noise in all the spectra by filtering with an adaptive filter which is function of the noise level, i.e., function of the noise standard deviation. The level to which it would be desirable to equalize the noise of the spectra is the minimum average noise level, i.e., somewhere around a standard deviation of the noise of 15 (which is the noise level for Silicon_P2).
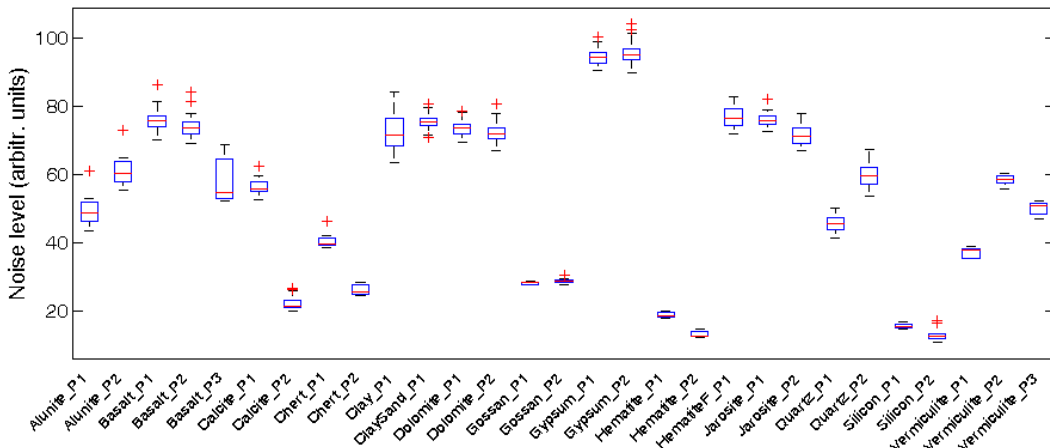
Figure 3-27. Boxplots with the calculated noise for all the spectra in the BB DB.

3.4.2.2.2 FILTER CHARACTERIZATION: NOISE DECAY VS. FILTER INDEX

The chosen filter for the analysis of the spectra was a moving average filter. The reason to use this filter instead of a Savitsky-Golay or a local regression filter is the operational complexity of these filters. The moving average filter has a very low consumption in terms of complexity and operation, without the need to perform any kind of regression or more complex calculations, given that this kind of filter averages all the consecutive points that are included in a window which moves along the spectrum. The higher the index, the lower the noise, but also the lower the definition of the Raman peaks.

To characterize the noise decay vs. the filtering index (window size), all spectra were filtered with several window sizes, providing, for each spectrum, curves of the noise decay with respect to the size of the filter window. The results for the average noise reduction at each point is shown in Figure 3-28 (absolute) and Figure 3-29 (relative).
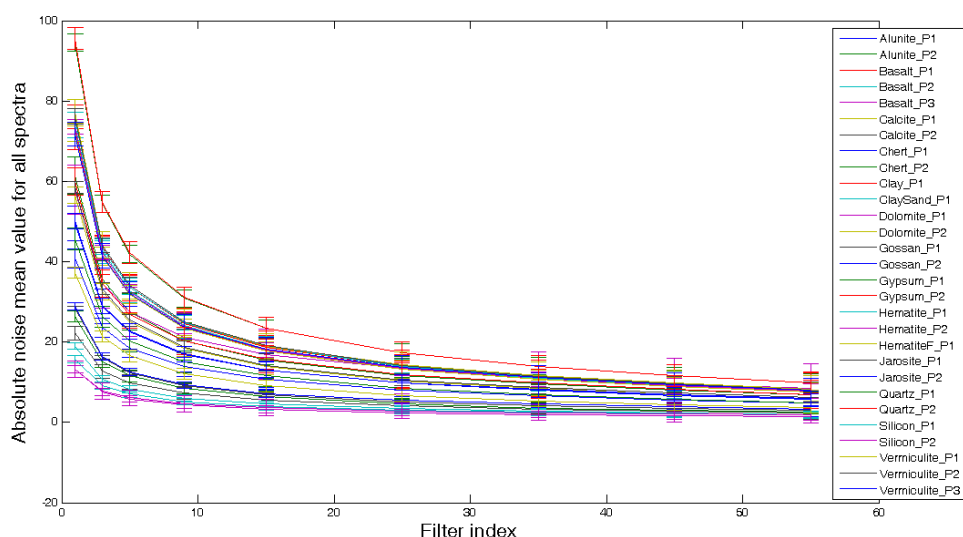
Figure 3-28. Absolute noise decay vs. filter window size for all samples and points
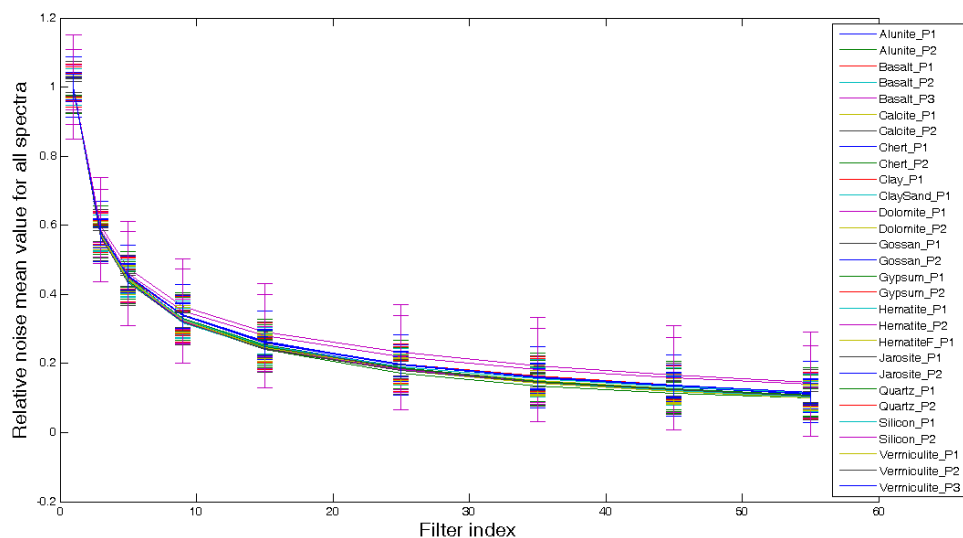
Figure 3-29. Relative noise decay vs. filter window size for all samples and points

89

As shown, the average relative decay with the window size somehow converges among all the different samples. Thus, by averaging all the samples, a global average noise reduction rate (in percentage) can be obtained as a function of the filter window size. By fitting this graph to a power function, the final noise value that a spectrum will have can be extrapolated.

Figure 3-30 shows the fitted curve of noise reduction vs. filter window size. The power function of the form $\Delta\sigma = a \cdot n^b + c$ where n is the filter window size (filter index), $\Delta\sigma$ the noise decrement and $a = -104.8$, $b = -0.4653$ and $c = 104.9$, fitted to the data with a RMSE of 0.2788 (with respect to a maximum value of around 100) and R-square value of 0.9999.
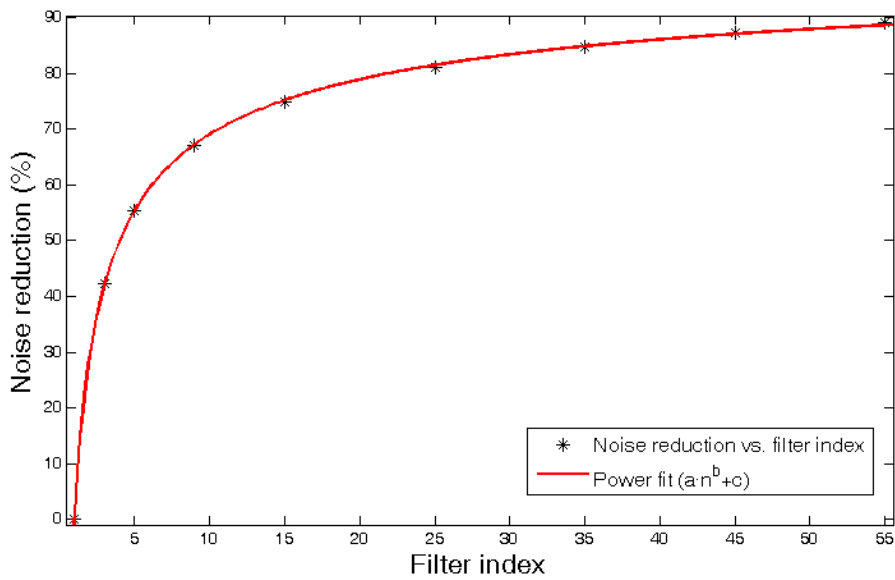
Figure 3-30. Noise reduction vs. Filter window size fit to a power function

Thus, the use of this function for the characterization of the filter provides a very useful tool to equalize the noise values of the different spectra.

The filter implementation needs an odd window size. So, the number of available filter index values will be quantized, and constant window sizes will be applied in determined noise regions (it will not be a continuous function). Table 3-6 defines the look-up table that will be used by the software to determine the filter index. For the breadboard database of spectra, it was estimated that a maximum window size of 35 would be enough to reduce the noise of the spectra up to 80%. Anyway, values are provided up to 55.

Noise maximum filtering (%)	Filter Index
0	1
42.1	3
55.3	5
62.5	7
67.2	9
70.5	11
73.1	13
75.2	15
76.8	17
78.3	19
79.5	21
80.5	23
81.5	25
82.3	27
83.0	29
83.7	31
84.3	33
84.8	35
85.4	37
85.8	39
86.3	41
86.7	43
87.1	45
87.4	47
87.8	49
88.1	51
88.4	53
88.7	55

Table 3-6. LUT for the calculation of the filter window size

3.4.2.3   VALIDATION

To validate the algorithm, all the spectra were analyzed and the noise values calculated. Figure 3-31 represents the results for the analyzed data without filtering and after equalization, showing the variable window sizes applied to the filter as a function of their original noise, as well as the resulting equalized noises. The equalization of the spectra is performed to reach a noise level of 15. This value is only a reference for the equalization level, as the results may vary as a consequence of the limited number of values that the filter can accept. It also provides a desired value, but spectra that have less noise than that are not filtered.
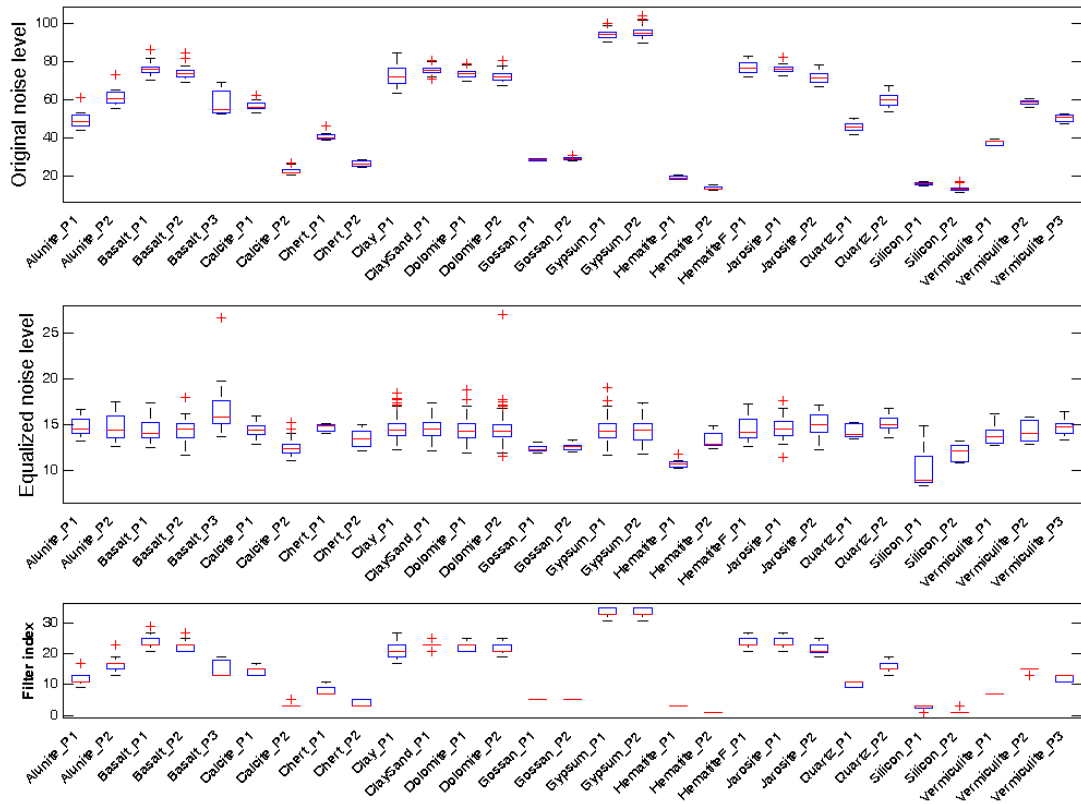
Figure 3-31. Noise level for all samples: original (up), equalized (middle). Filter index (bottom).

3.4.3 BASELINE CALCULATION

The baseline calculation is the key algorithm for the SNR calculation. The correct removal of the baseline will provide accurate results for the signal intensity, while errors might lead to artificially increasing the SNR, thus reducing the spectral quality of the final spectra (given that higher SNR on an individual spectrum might imply a lower number of accumulations). This section describes the algorithm that will be used for the calculation of the baseline, and covers the optimization of its parameters.

The applied methodology for the parameterization and validation of the algorithm was the following: all the baselines for all the spectra were manually calculated. These were considered as the ideal baselines. Then, the baseline for all the samples was automatically calculated with the algorithm, while the parameters of the algorithm were swept along their expected range. Then, the ideal baselines were used for comparison with the automatically calculated ones, with the aim of obtaining those parameters with the overall best performance (minimum error). The error quantification criterion was to minimize the MSE (Mean Squared Error) of the difference of the manually and automatically calculated baselines.

In general, when training models it is interesting to have several sets of training spectra to avoid overfitting. In this case, the results were validated with a second batch of spectra from the same samples to verify that the results were accurate enough.

3.4.3.1 DESCRIPTION

The computational complexity is a key factor when defining onboard algorithmic. Several proposals for the automated calculation of spectra can be found in the literature [12, 46-52]. However, most of them have computational requirements out of the limits of the RLS instrument. Previous work performed by [13, 21, 53] has shown that a very simple baseline approach is quite robust and simple, compared to other proposals. Thus, this is the option that is being proposed for implementation in the application software of the RLS instrument.

This algorithm is a modification of classical approach based on defining, for each point x, a baseline value equal to the minimum of the averages of equidistant points from x, to a certain distance. In other words, the baseline is calculated at each point *p* as the minimum value among the averages from points in positions *p+deltaindex/2* and *p-deltaindex/2*, where *deltaindex* varies from 1 to *delta*. *delta* should be defined as the maximum peak width that is expected to be found in the spectrum. Figure 3-32 presents the graphical illustration of the algorithm behavior. The red dots indicate the possible baseline values, calculated as the average of the extreme points. The final baseline value will be the minimum of all the points.
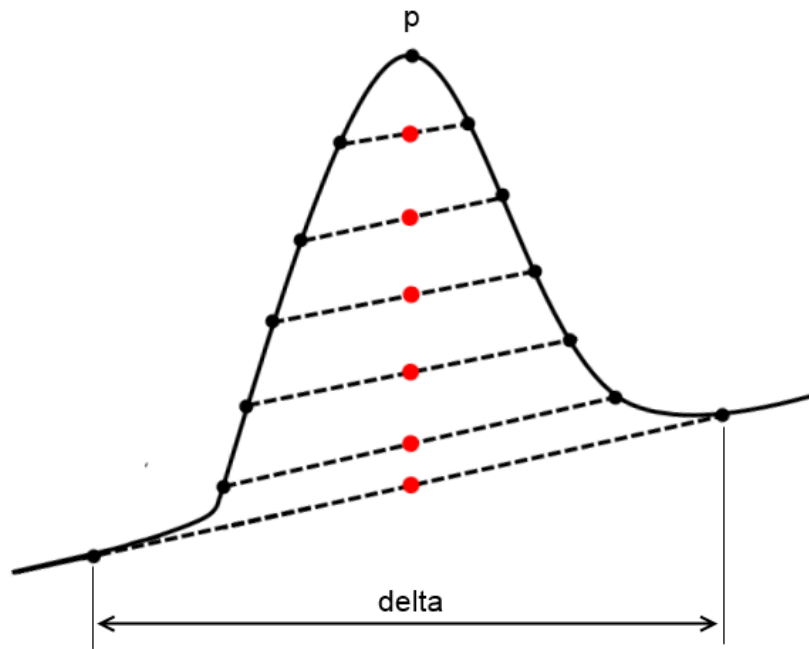
Figure 3-32. Graphical illustration of the baseline algorithm calculation principle

For our implementation, we propose to define two regions, separated at *bandlimit* parameter, to allow defining different *delta* values in each of them (δ_1 and δ_2). This makes the algorithm more robust to work with different types of samples (this way it is possible to deal with different peak widths as a function of the spectral region). The optimization of the algorithm will define the optimal values for δ_1 and δ_2. The pseudocode defining the algorithm is shown below.

```
baseline = int_orig;
pixels = length(rs_orig);

for  pixelindex= 1:pixels
    if rs_orig(pixelindex)>bandLimit
        delta1=delta2;
    end
    for  deltaindex=delta1:-1:1
        pixelindex2 = int32(pixelindex  +  2*delta1 - 2*deltaindex);
        if  (pixelindex2 < pixels +1)
            temp = (baseline(pixelindex) + baseline(pixelindex2))/2;
            spectrumpixel = int32(pixelindex + delta1-deltaindex);
            if (temp < baseline(spectrumpixel))
                baseline(spectrumpixel) = temp;
            end
        end
    end
end
```

3.4.3.2   PARAMETERIZATION

In order to get the best possible results, the algorithm parameters (δ_1, δ_2, and *bandlimit*) were optimized. *bandlimit* was adjusted to a value of 1150 (in pixels). This value is based on the shape of the spectrum, to ensure that the maximum in the response of the spectrometer is excluded from the wide peak region, to avoid identifying the background response as a peak. The spectrometer response was characterized with amorphous zinc borate, which does not have Raman response. This curve is shown in blue in Figure 3-33, and the *bandlimit* value is overlaid, showing the separation between the two spectral regions.
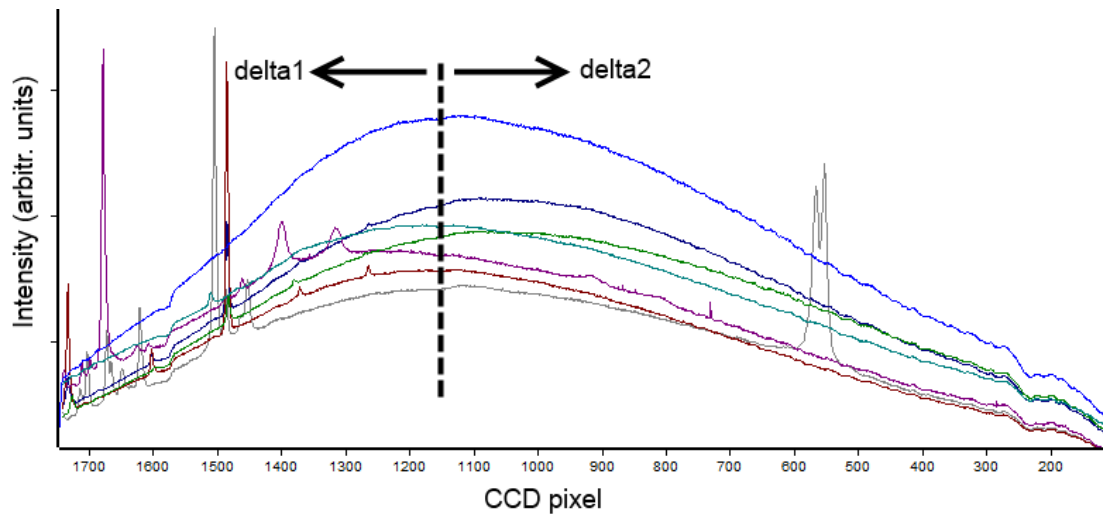
Figure 3-33. Zinc borate spectrum (blue), and spectra from other materials for reference (Alunite – bluish green, calcite - maroon, dolomite - green, chert – purple, gypsum - grey).

The optimization of δ_1 and δ_2 was based on calculating the baseline with different values, for all the spectra in the breadboard database. These were swept 2 by 2 with values ranging from 2 to 80 for δ_1, and from 2 to 100 for δ_2. By comparing each of the automatically calculated baselines with respect to the manually calculated ones, a matrix of MSE values is obtained for each sample. The smaller the error, the closest the calculated baseline to the ideal one. By averaging all the matrixes, the overall MSE matrix value is obtained. This can be used to obtain the best-performing combination of parameters, by taking the minimum values from Figure 3-34 (blue area). Based on this criterion, the most suitable values are δ_1=22 and δ_2=60.

As expected, these values do not fit equally well for all the materials. As shown in Figure 3-35, there are some materials, such as vermiculite or the mixture of clay and sand, whose optimal values for δ_1 and δ_2 are relatively far from the global optimal values (darkest areas in the figures). In these cases there can be an error in the baseline calculation which can provide inaccurate SNR values. However, the NA calculation process was adjusted so that errors in SNR never imply a loss on the final spectrum spectral quality, as explained in section 3.4.4 and 3.5. It can be said that this algorithm, though suboptimal in some specific

cases, is generally robust, and combined with the NA calculation process, guarantees the spectral quality of the final spectra.
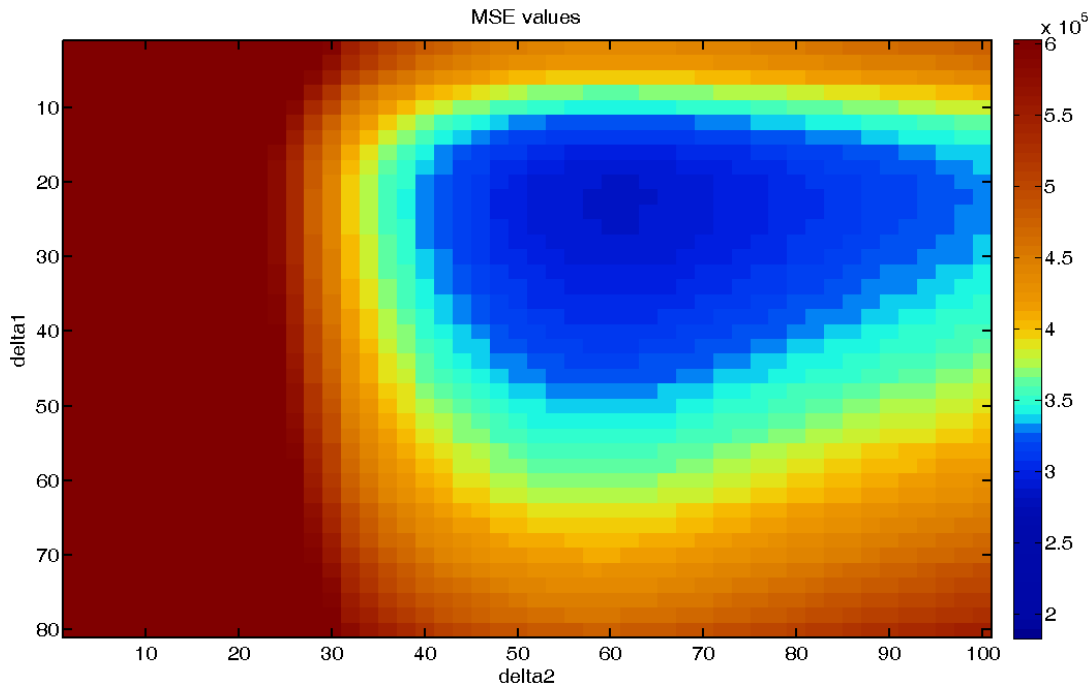
Figure 3-34. MSE values matrix for the two swept parameters.
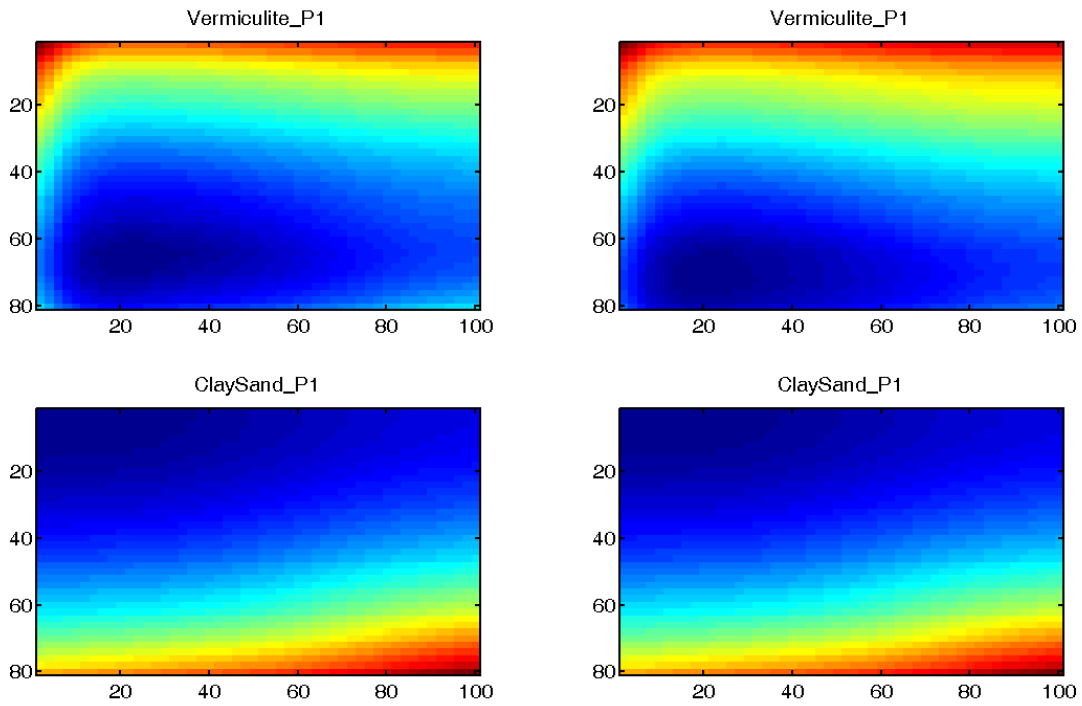
Figure 3-35. MSE values matrix for some individual materials (worst cases).

X axis represents δ_2, and Y axis represents δ_1.

3.4.3.3   V ALIDATION

To ensure that the training process was correctly performed and to ensure that no overfitting occurs with this training method, another batch of spectra from the same materials were used to automatically calculate the baselines. The resulting MSE was only 4% different with respect to the training method. Of course, when the spectral database is updated with new spectra, it will be necessary to re-train the algorithm and to obtain new parameters. In addition, this will need to be done for each new model of the instrument.

3.4.4 SNR C ALCULATION

The SNR value of some spectra during the onboard operation of the RLS instrument is very important in order to optimize the use of the available time, which, as shown in 3.5, can highly reduce the total acquisition time for some samples, potentially increasing the number of points that can be analyzed in a determined sample, providing better science return.

The spectra used for the validation of the algorithm were acquired with the final Ti representative of the sample, as defined in Table 3-5, and then used to manually calculate the baseline of all the spectra, as explained in previous sections. The SNR values obtained when using the manually and automatically calculated baselines were compared to evaluate the performance of the automatic calculation of the SNR. The methodology to calculate the final SNR is as defined in Figure 3-24 (except for the manually calculated baselines, which skip the Baseline Calculation step).

The obtained SNR values are presented in Table 3-7, and scattered in Figure 3-36, showing a very good performance and linearity. However, to study the calculation errors in detail, it is necessary to focus in those spectra with lower SNR values. Figure 3-37 represents the automatically calculated vs. the real SNR in logarithmic scale, providing a better representation of the scattering. This figure shows how, for low SNR values, the automatic estimation of SNR highly diverges from the expected values. The automatic baseline calculation errors induce inaccuracies in the calculation of the SNR, which are especially important when the calculated baseline is lower than the expected one. In these cases, the final SNR can be overweight, which could affect the final quality of the acquired spectra. To avoid these situations, we studied all the spectra to detect the maximum false SNRs that can be obtained as a result of a bad baseline correction. The worst case was around 25. Thus, it was decided (see 3.5) that no reduction of the number of accumulations would be considered for SNR values lower than 30 (25 plus a 20% margin).
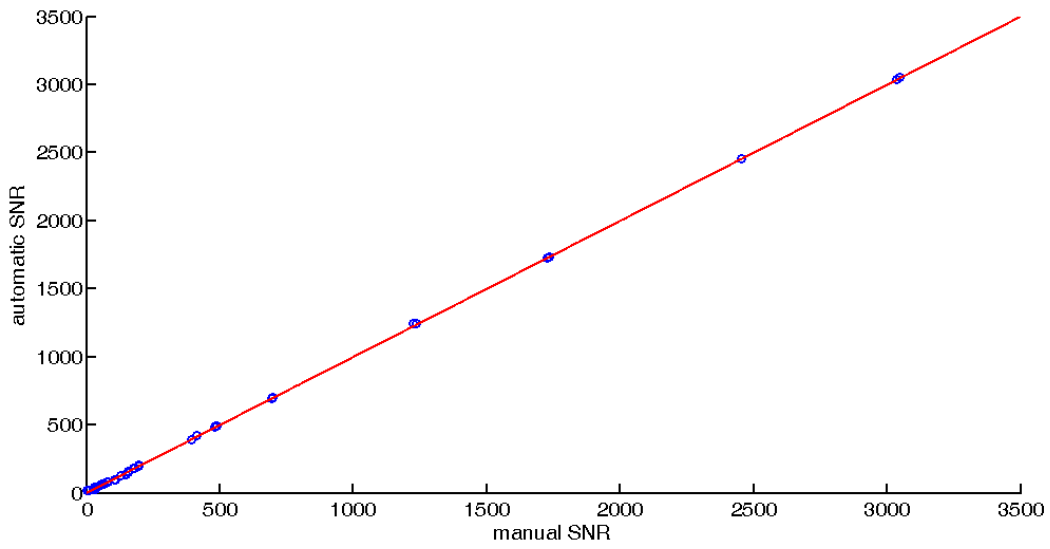
Figure 3-36. SNR values scatter for all samples. Red line is y=x.
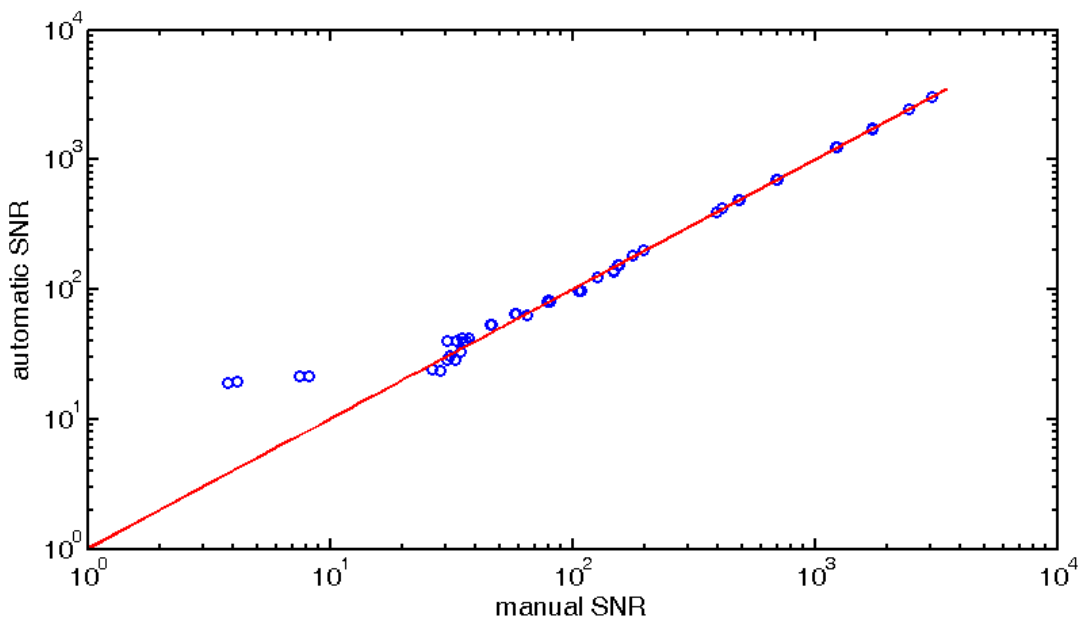
Figure 3-37. SNR values scatter for all samples, in logarithmic scale. Red line is y=x.

Sample	Spectrum number	Manual SNR	Automatic SNR
Alunite_P1	1	484.3	489.3
Alunite_P1	2	482.3	487.3
Alunite_P2	1	413.9	419.4
Alunite_P2	2	415.1	419.7
Basalt_P1	1	177.9	180.8
Basalt_P1	2	178.1	181.5
Basalt_P2	1	197.4	201.5
Basalt_P2	2	197.1	202.0
Basalt_P3	1	36.6	40.5
Basalt_P3	2	35.4	38.9
Calcite_P1	1	487.0	493.8
Calcite_P1	2	485.9	492.7
Calcite_P2	1	1733.9	1735.7
Calcite_P2	2	1725.0	1728.2
Chert_P1	1	391.8	394.3
Chert_P1	2	393.5	395.0
Chert_P2	1	1235.8	1244.3
Chert_P2	2	1225.5	1247.2
ClaySand_P1	1	4.2	19.3
ClaySand_P1	2	3.8	19.1
Dolomite_P1	1	58.4	64.5
Dolomite_P1	2	58.0	64.9
Dolomite_P2	1	46.2	53.0
Dolomite_P2	2	46.6	53.0
Gossan_P1	1	64.8	63.5
Gossan_P1	2	65.0	63.9
Gossan_P2	1	127.0	124.5
Gossan_P2	2	127.1	124.6
Gypsum_P1	1	37.2	42.6
Gypsum_P1	2	34.9	42.0

Sample	Spectrum number	Manual SNR	Automatic SNR
Gypsum_P2	1	30.5	40.0
Gypsum_P2	2	33.3	40.4
Hematite_P1	1	148.1	138.7
Hematite_P1	2	147.6	137.6
Hematite_P2	1	106.4	96.9
Hematite_P2	2	108.4	96.6
HematiteF_P1	1	7.5	21.4
HematiteF_P1	2	8.2	21.5
Jarosite_P1	1	79.7	81.4
Jarosite_P1	2	80.0	81.9
Jarosite_P2	1	78.9	80.6
Jarosite_P2	2	80.4	81.4
Quartz_P1	1	694.9	695.3
Quartz_P1	2	699.2	700.0
Quartz_P2	1	154.9	154.7
Quartz_P2	2	155.5	155.5
Silicon_P1	1	2455.1	2456.2
Silicon_P1	2	2455.1	2456.8
Silicon_P2	1	3034.0	3036.0
Silicon_P2	2	3047.1	3049.9
Vermiculite_P1	1	34.7	32.8
Vermiculite_P1	2	31.0	31.0
Vermiculite_P2	1	32.7	28.7
Vermiculite_P2	2	30.5	29.0
Vermiculite_P3	1	28.4	23.6
Vermiculite_P3	2	26.3	24.0

Table 3-7. SNR values with manually and automatically calculated baselines

The range of SNR in which the number of accumulations will be modified as function of the spectral quality, is defined in section 3.5 for values higher than 30 and lower than a target SNR in the order of 150-200. Figure 3-38 represents the SNRs scatter in that range. The linear fitting to the data shows an almost unitary slope and a positive bias in the SNR calculation of 3.4 (see fitting curve equation in Figure 3-38) that should be subtracted from the SNR values automatically calculated. Under these conditions, the maximum over-estimation error found is around 8-9% (for Dolomite_P1, point 2). Thus, to ensure further robustness to the algorithm, it might be interesting to subtract this error to the SNR automatic calculation values.

The figure shows a scatter plot with "manual SNR" on the x-axis and "automatic SNR" on the y-axis, with the equation $y = 0.97 \cdot x + 3.4$ and a green linear fit line.
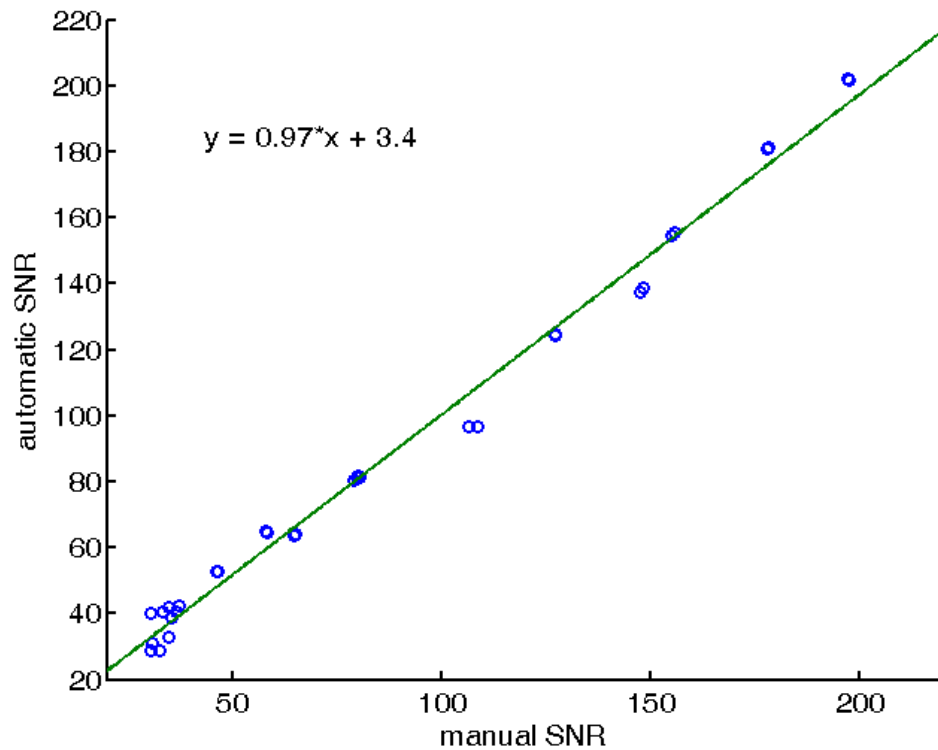
Figure 3-38. Scatter of SNR values that can modify the NA value. Green line is the linear fit.

3.4.5 IMPLEMENTATION

This section details the functions that are expected to be needed for the implementation of this algorithm.

equalized_spectrum(noise_std, spectrum) – Function to equalize the spectral noise based on the spectrum noise standard deviation.

calculate_baseline(rshift, spectrum, bandlimit, delta1, delta2) – Function to obtain the baseline of a spectrum with the proposed algorithm.

subtract_baseline(spectrum, baseline) – Function to get the baseline-corrected spectrum.

get_maximum_of_spectrum(spectrum) – Function to locate the maximum intensity of a spectrum.

calculate_SNR(maximum_value, noise_std) – Function to calculate the SNR value of a spectrum (basically a quotient).

3.5 Calculation of the Integration Time and Number of Accumulations

The Raman scattering efficiency of the materials that will be analyzed in Mars (minerals and rocks in powdered form), are expected to be in the order from 10^{-7} to 10^{-11}. Unhappily, the CCD dynamic range cannot deal with this scattering range, with potential variations of several orders of magnitude. Thus, it is necessary to adjust the acquisition parameters. As explained before, there are two key operational-level parameters that can be adjusted when performing Raman spectroscopy: the integration time (Ti) and the Number of Accumulations (NA).

This section presents a study of the spectral quality of the acquired spectra as a function of these two parameters. In addition, an algorithm is developed and validated for the automated calculation of these parameters, based on the spectral response of the sample under analysis. This way, it will be possible to optimize the acquired spectral quality, while saving operation time for good Raman scatterers, by reducing the final number of accumulations. To do so, the SNR of the spectrum is characterized as a function of NA, in such a way that it will be possible to progressively reduce NA with the SNR increase.

3.5.1 Experiment setup and methodology

For the development and validation of this algorithm, the same set of samples used for the definition of the cosmic ray detection and the SNR calculation algorithm was used, i.e., the RLS breadboard spectra database. As previously detailed, this database includes spectra from different samples acquired in different conditions that are representative of the operation mode, or else, are considered good references: silicon, calcite, vermiculite, gypsum, alunite, basalt, dolomite, hematite (two different types), clay, jarosite, chert, a mixture of clay and sand, quartz and gossan.

All these samples were analyzed with the breadboard, acquiring reference spectra that were used to calculate the optimal integration time, and then, acquiring enough spectra with the final integration time to complete a total operation time of 600 seconds (considered to be the maximum operation time per point for Raman acquisition). In addition, some of them (calcite, silicon and gossan) were analyzed with different integration times and number of accumulations in order to perform studies on the behavior of the SNR with different acquisition parameters.

The SNR calculation is, when possible, based on a method which calculates the noise of the spectrum by subtracting two consecutively acquired spectra (as referenced in section 3.3.3.1.1). This removes the Raman spectral features of the spectrum, providing a noise-like spectrum. However, in some cases, when both spectra were not available, the SNR was calculated by removing the baseline and calculating the noise in a flat region of the baseline-subtracted spectrum. The main difference between both methods is that, being the first one based on calculating the noise from the difference spectrum, the noise in that

case is effectively reduced by a factor $\sqrt{2}$, increasing the resulting SNR by that factor. However, when comparing the SNRs obtained with different methods, this factor is applied throughout the work to make them comparable.

3.5.2 SPECTRAL SNR CHARACTERIZATION AS A FUNCTION OF TI & NA

3.5.2.1   SNR MAXIMIZATION AS FUNCTION OF (TI, NA)

As already stated, the Raman acquisition process can be adjusted by modifying the integration time, Ti, and the number of accumulations, NA. These two parameters directly affect the total operation time, which is calculated as NA*Ti. Total available acquisition time will be from 5 to 10 minutes.

In order to evaluate the spectral quality based on these parameters, several 5 minute spectra were acquired with different (Ti, NA) combinations, for three different materials with different scattering efficiencies (calcite, silicon and gossan). For the three materials, the SNR of each spectrum was calculated. To make the curves comparable in the same graph for all the samples (i.e., to compensate the materials different efficiency), they were normalized and centered by subtracting their average and dividing them by their standard deviation (this graph is represented in Figure 3-39). The results show that, for spectra acquired with equivalent total operation times (300 seconds), the SNR is higher for higher integration times than for higher numbers of accumulations. One of the reasons is that for higher integration times, the CCD needs to be read fewer times, highly reducing the contribution of the readout noise to the spectrum SNR. The implication of this result is quite straightforward: the algorithm will need to **optimize Ti first, and then NA**, while taking into account the total available time.
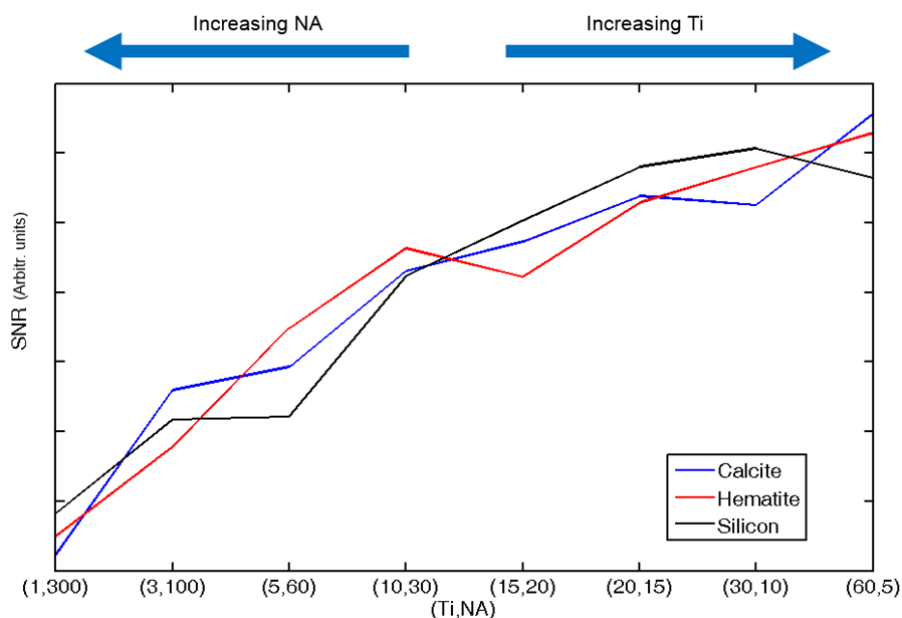
Figure 3-39. SNR evolution for different combinations of acquisition parameters. All spectra were acquired during a total operation time of 300 seconds. Data is normalized and centered for visualization.

3.5.2.2   Ti Characterization

It is well known that, in Raman spectroscopy, the higher the integration time, the higher the acquired signal intensity and SNR. Furthermore, Raman emission is a scattering process which happens with higher probability as a function of the excitation source flux of photons. Given that a CW laser provides a constant flux of photons, it is expected that the number of Raman photons emitted by the sample and received by the CCD will be constant per time unit.

The CCD response to light is quite linear when the saturation level is in the dynamic range of the CCD, as the dark current levels are negligible (<10 electrons per pixel per second for an inverted mode device). This is so, even experimentally, at the working temperature of the CCD installed in the RLS breadboard system used during the spectra acquisition (this might be different if the CCD was a non-inverted mode one). On the other hand, the bias level of the CCD can influence the calculations if not taken into account, especially for poor Raman scatterers. Thus, the **intensity increase is constant with time** when taking into account the effect of the bias level. Figure 3-40 presents the intensity at one pixel with respect to the integration time. It can be seen that the linear adjustment provides an R-square value of 1. Of course, this does not imply that the SNR also increases linearly.

However, given that the integration time needs to be calculated first for the automatic adjustment of the parameters, as justified in section 3.5.2.1, it is the intensity which is more interesting, as the calculation of Ti needs to make sure that the CCD does not saturate. Thus, it is the intensity that will be parameterized with the algorithm, instead of the SNR of the resulting spectra, relieving the processor from the extra computational load that would be necessary to perform these calculations. The criterion to define the integration time is to have the spectral intensity cover 80% of the dynamic range of the CCD.
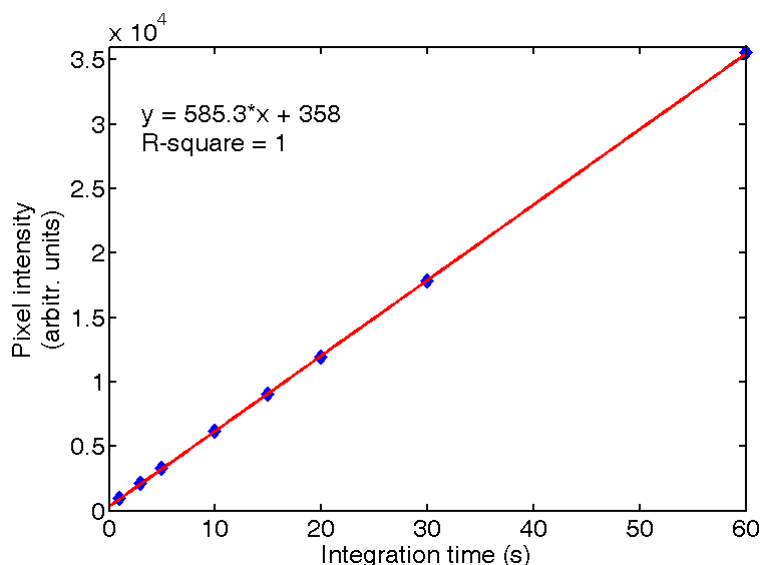
Figure 3-40. Intensity at one CCD pixel vs. Integration time

Guillermo Lopez-Reyes

3.5.2.3   NA CHARACTERIZATION

As it is also well known, the accumulation of several consecutively acquired spectra reduces the noise of the averaged spectrum. However, as seen in section 3.5.2.1, the SNR increase associated with higher numbers of accumulations is lower than the SNR increase provided by higher integration times. Nevertheless, once Ti has been calculated to cover 80% of the dynamic range of the CCD (see section 3.5.2.2), the rest of the available time can be used to acquire and average several spectra to increment the SNR.

In order to characterize the SNR increase with respect to the number of accumulations, several spectra of silicon, calcite and gossan with several integration times were acquired and analysed. By accumulating those spectra with different numbers of accumulations, and calculating the SNR, the SNR evolution with respect to NA was obtained for different materials and integration times. Figure 3-41 shows the curves, normalized for representation purposes. As it can be seen, only up to 150 accumulations are represented for the Ti = 1s spectra, while only 50 for the Ti = 3s spectra, and so on. This is due to the fact that the total operation time during the test was set for these samples to 300 seconds, and the fact that, to calculate the SNR of a spectrum it is necessary to subtract two of them consecutively. Thus, it is only possible to have 50 accumulations of spectra with Ti = 3s, 30 spectra for Ti = 5s, 10 spectra for Ti = 10s, and so on.

This representation shows that the **SNR evolution is similar for different spectra,** not only when acquired with the same integration time, but also with different Ti's. Actually, the correlation value between the curves is higher than 95%.
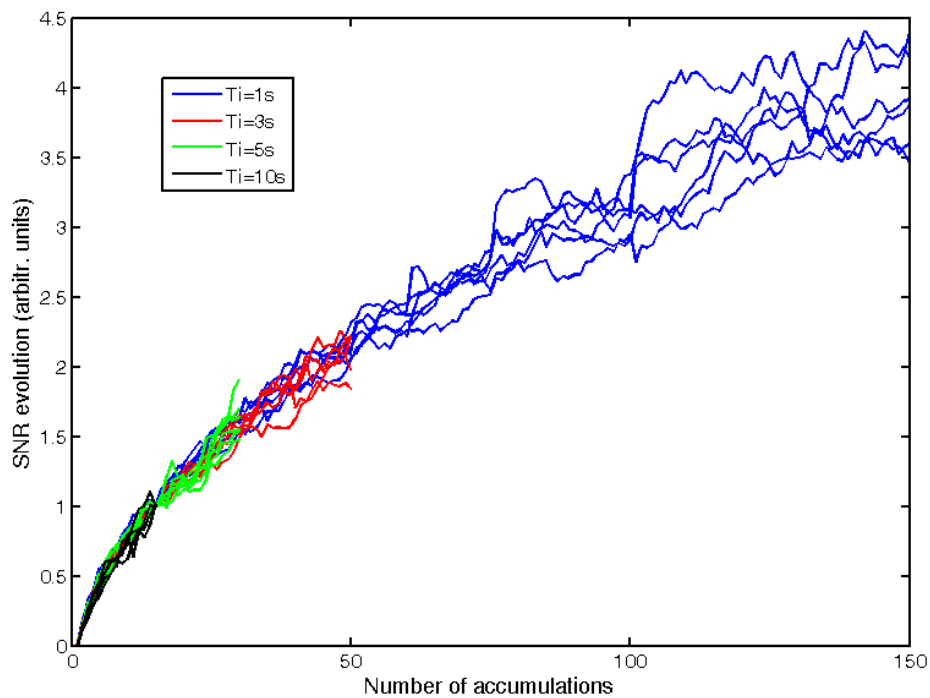
Figure 3-41. SNR evolution for different numbers of accumulations, for different Ti's and samples

104

Given the similarity in the curves for spectra with different integration times, we used the average SNR evolution for the Ti = 1s spectra to fit the SNR evolution to a power function. Figure 3-42 represents this curve, normalized to the SNR of a single (non-accumulated) spectrum. Thus, this figure represents the factor by which the SNR gets multiplied when accumulating different numbers of spectra, or alternatively, the SNR increment in parts per unit. It is interesting to see how the SNR evolution is proportional to the square root of NA with an R-square value very close to 1.

The figure shows a plot with x-axis labeled "Number of accumulations" (0 to 150) and y-axis labeled "SNR increment (parts per unit)" (0 to 12). The fit equation shown is $0.91 \cdot x^{1/2} + 0.106$ with R-square = 0.997.
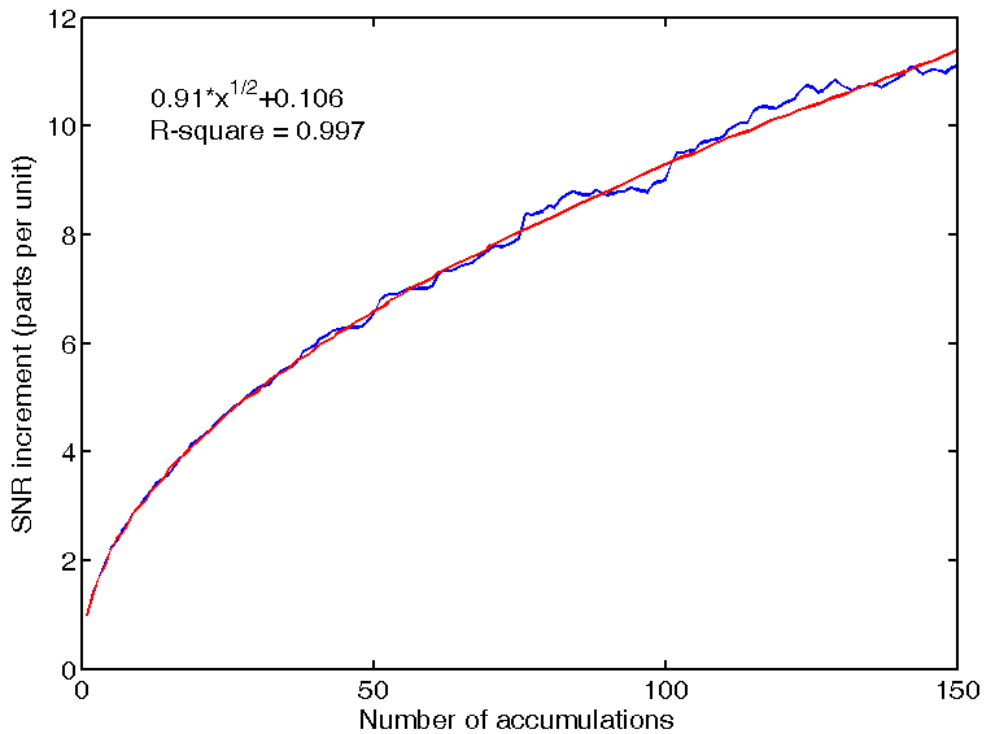
Figure 3-42. SNR multiplying factor for different numbers of accumulations

3.5.3 ALGORITHM PROPOSAL

3.5.3.1 DESCRIPTION

In this section, the detailed flowcharts for the calculation of Ti (Figure 3-43) and NA (Figure 3-44) are presented. As explained in the previous sections, the optimal way to adjust the acquisition parameters is to calculate the final integration time first, and then the number of accumulations, as a function of the available time. In addition, the algorithm proposes to progressively reduce the value of NA with increasing values of SNR (the higher SNR, the lower the necessary NA).

3.5.3.1.1 INTEGRATION TIME CALCULATION

In order to calculate the integration time, it is important to make sure that no pixel of the CCD saturates. However, in order to avoid operational load, the calculation of the pixel of the CCD ROI (Region of interest) that contains the spectrum maximum intensity will be decided based on the binned (1D) spectrum that was used for the previous stages (fluorescence and cosmic ray detection). By using the maximum position of this spectrum it is possible to locate the maximum intensity among all the pixels in the corresponding column in the ROI. However, it is necessary to take into account the list of cosmic rays removed from the 1D spectrum, as these will not be removed from the 2D spectrum. To solve this issue, in case the maximum position coincides with a cosmic ray, the next maximum of the spectrum will be used as the maximum intensity point.

The maximum intensity is the intensity obtained for the reference spectrum integration time (Ti-ref = 1 second). In addition, in order to ensure the linearity of the signal increase with time, it is necessary to subtract the bias level of the CCD. Figure 3-43 shows the detailed flowchart for the calculation of the final integration time.
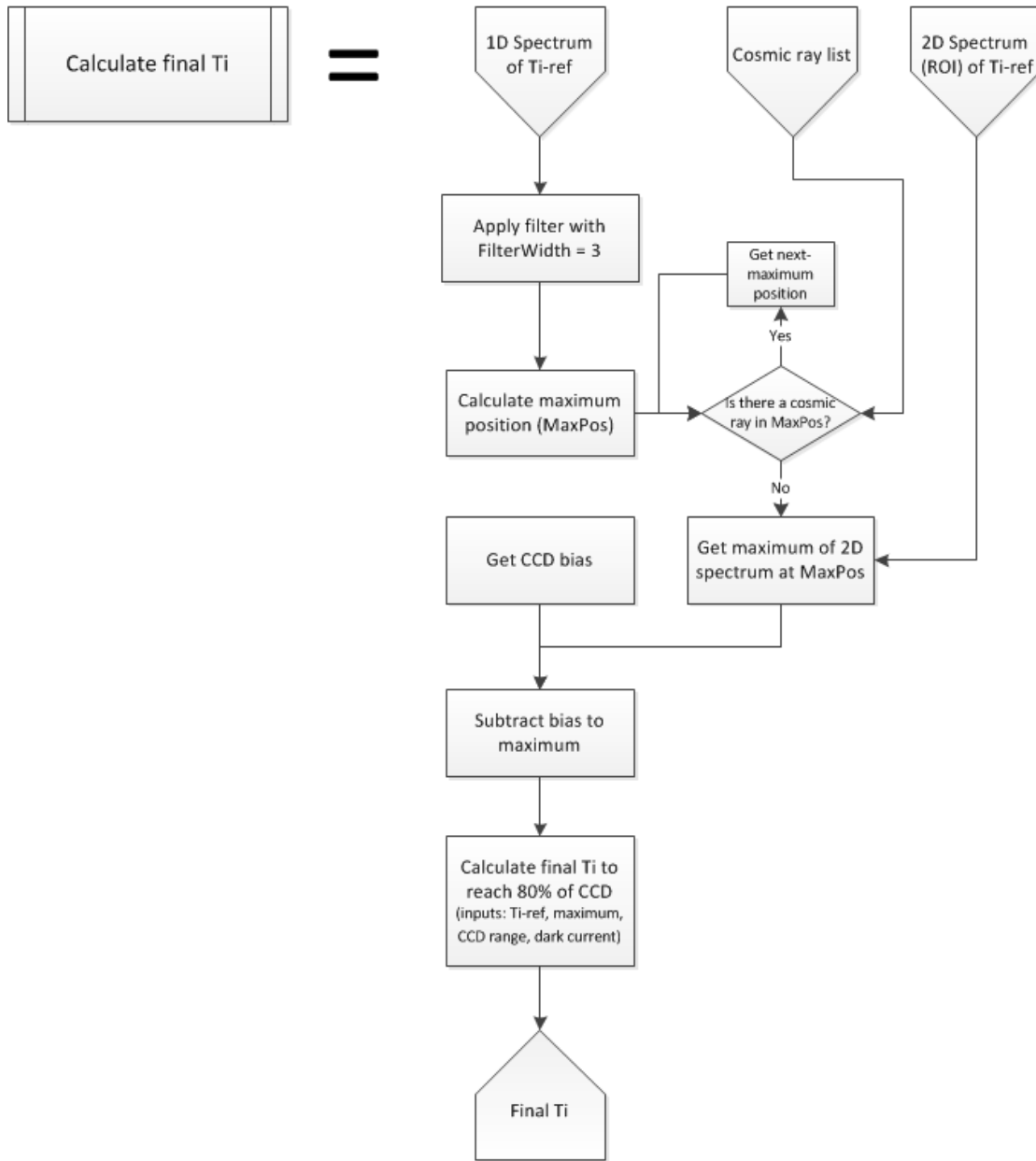
Figure 3-43. Flowchart for the calculation of the integration time

3.5.3.1.2 CALCULATION OF THE NUMBER OF ACCUMULATIONS

The calculation of the number of accumulations is based on the spectral quality of two reference spectra acquired with the final Ti. The reason to have two of them is to reject the possible cosmic rays or spikes that can appear in the spectra, and that can hamper the calculation of the baseline (which is necessary to calculate the SNR, as explained in section 3.4). The calculated SNR value will be used to determine the spectral quality and the necessary NA to increment the SNR as in Figure 3-42, to reach the minimum quality standards required by the scientific team. However, the NA calculation has to fit into the

Guillermo Lopez-Reyes

limits imposed by the operational constraints, such as minimum and maximum number of accumulations, or maximum total operation time (Ti*NA), which cannot be exceeded.
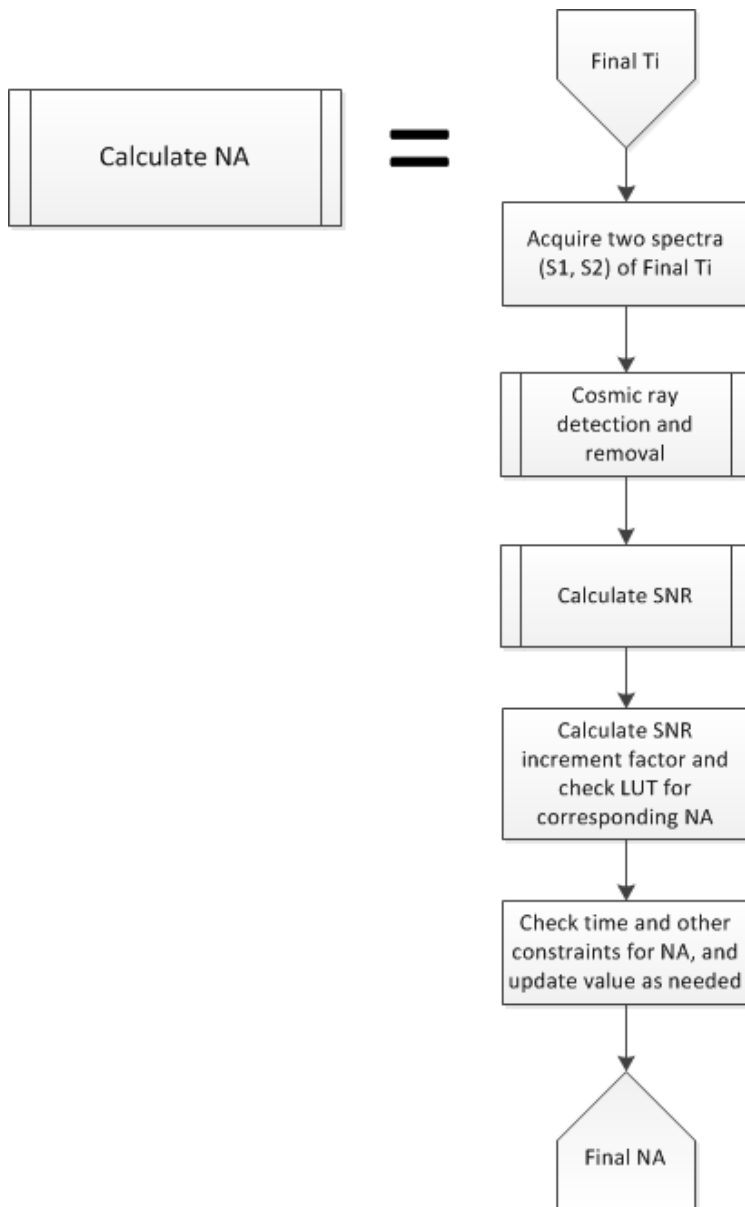
Figure 3-44. Flowchart for the calculation of the number of accumulations

3.5.3.2 PARAMETERIZATION

The calculation of the final integration time is quite straightforward by nature, as it is only limited by the engineering requirement of **filling 80% of the dynamic range of the CCD**. It is important to take the bias level into account to ensure the linear increase of the intensity. By subtracting the bias level (obtained from a dark region of the CCD), the intensity value of the sample scattering itself is obtained, if dark current is negligible, as for an IMO CCD. For other types of CCDs, the dark current would have to be taken into account as well.

Then, to calculate the final optimal integration time, it is quite straightforward based on these parameters, to comply with the following equation:

$$FinalTi = \frac{0.8 \cdot CCD\_dynamic\_range}{maximum - bias - darkCurrent \cdot Ti_{ref}} \cdot Ti_{ref}$$

However, there is a maximum integration time of **Final Ti ≤ 60 s** that limits its value. The reason to choose this value is that it is necessary to acquire dark spectra of Ti. The longer Ti, the higher the time spent without acquiring "relevant" Raman signal. On the other hand, the calculation of NA needs to be parameterized, specifically, establishing limits for SNR and NA. For values below a determined SNR, the maximum possible NA will be acquired. For values over the desired SNR value, the minimum NA will be acquired.

To define the minimum value of SNR, the baseline correction errors of the SNR calculation were considered. As explained in section 3.4, errors in the baseline calculation can lead to considering false peaks that will provide inaccurate SNR values. We performed a study on the baseline correction errors by evaluating all the "false peaks" detected by the algorithm, and calculating their SNR. In this case, the maximum value for all the samples was around 25. Applying a margin of 20%, we thus set the minimum SNR value to consider reducing the number of accumulations to **SNR_{min} = 30**. The SNR value that would be desirable for all the samples (target SNR, or SNR_{max}) is set to **SNR_{max} = 120**, which is twice the SNR obtained with a 10 minutes spectrum of gossan (Ti=60s, NA=10), considered one of the worst scatterers that is expected to be found.

The NA limits were set between 4 and 150 to ensure that at least some accumulations are performed in each case, while providing margin enough to get to high numbers of accumulations:
NA_{min} = 4 and NA_{max} = 150.

In addition to defining the limits for the parameters, it is necessary to parameterize the SNR evolution as a function of NA, as per Figure 3-42. To comply with operational requirements, a look-up table (Table 3-8) will be checked by the system based on the SNR multiplication factor that needs to be applied to reach SNR_{max}. For example, if the SNR of one accumulation was 32, this means it would need to be multiplied by a factor of 120/32 = 3.75. From Table 3-8, the necessary NA would be 16. It is interesting to note that, given that the maximum error in the effective range of SNRs that might reduce the number of accumulations is 8-9% (see section 3.4.4), the LUT defined in Table 3-8 has been limited to NA values which provide an increment of SNR of 10%, higher than the maximum error.

Taking all these issues into consideration, the final calculated value of NA will be defined by the following logic, which calculates values of NA both based on quality as well as on the available time.

Based on SNR:

if $\quad$ SNR<SNR_{min} -> $NA_{aux} = NA_{max}$

else $\;$ if $\;$ SNR >SNR_{max} -> $NA_{aux} = NA_{min}$

else $\quad$ get NA_{aux} from LUT for SNR increment $\geq SNR_{max}$ / SNR

Based on available time:

If $\quad$ (NA_{aux} + 1) * final Ti < max_operation_time $\quad$ then $\quad$ final NA = NA_{aux}

else $\quad$ final NA = floor (max_operation_time / final Ti) - 1

NA	SNR increment		NA	SNR increment		NA	SNR increment
1	1.0		44	6.2		118	10.1
2	1.4		46	6.3		121	10.2
3	1.7		47	6.4		123	10.3
4	1.9		49	6.5		126	10.4
5	2.2		50	6.6		128	10.5
6	2.4		52	6.7		131	10.6
7	2.5		53	6.8		133	10.7
8	2.7		55	6.9		136	10.8
9	2.9		56	7.0		138	10.9
10	3.0		58	7.1		141	11.0
11	3.1		60	7.2		143	11.1
12	3.3		61	7.3		146	11.2
13	3.4		63	7.4		148	11.3
14	3.5		65	7.5			
15	3.7		67	7.6			
16	3.8		69	7.7			
17	3.9		70	7.8			
18	4.0		72	7.9			
19	4.1		74	8.0			
20	4.2		76	8.1			
21	4.3		78	8.2			
22	4.4		80	8.3			
23	4.5		82	8.4			
24	4.6		84	8.5			
25	4.7		86	8.6			
26	4.8		88	8.7			
27	4.9		90	8.8			
28	5.0		92	8.9			
30	5.1		94	9.0			
31	5.2		96	9.1			
32	5.3		98	9.2			
33	5.4		100	9.3			
35	5.5		102	9.4			
36	5.6		105	9.5			
37	5.7		107	9.6			
39	5.8		109	9.7			
40	5.9		111	9.8			
41	6.0		114	9.9			
43	6.1		116	10.0			

Table 3-8. LUT with the SNR factor vs NA

3.5.3.3   VALIDATION

The intent of dynamically adjusting the acquisition parameters is to save operation time when the spectral quality is over a determined threshold. In order to validate the algorithm, all the spectra in the database were acquired with the automatically calculated parameters, and the SNRs of the final spectra were calculated. The parameters calculation results for Ti are shown in Table 3-9. Final-Ti is limited to 60 seconds, thus, for cases where Calculated Ti (the time required to fill the 80% of the CCD range) is higher, the Final-Ti can only be as high as 60 seconds.

Material	Point	Ti-ref	Bias	Ti-ref spectrum max intensity	Calculated Ti (80% CCD)	Final-Ti
Alunite	P1	1	65	457	33.47	33.47
	P2	1	65	430	35.95	35.95
Basalt	P1	5	64.5	5052	13.15	13.15
	P2	5	64.5	4800	13.85	13.85
	P3	1	66	330	49.7	49.70
Calcite	P1	1	65	913	15.47	15.47
	P2	1	64.5	1294	10.67	10.67
Chert	P1	1	66	196	100.92	60.00
	P2	1	65	289	58.57	58.57
Clay+Sand	P1	1	70	3240	4.14	4.14
Dolomite	P1	1	71	3572	3.75	3.75
	P2	1	76	5884	2.26	2.26
Gossan	P1	1	65	105	328	60
	P2	1	65	105	328	60
Gypsum	P1	1	64.5	1790	7.6	7.60
	P2	2	64.5	3045	8.8	8.80
Hematite	P1	1	65	93	468.57	60.00
	P2	1	65.4	80	898.63	60.00
Hematite-F	P1	1	62	431	35.56	35.56
Jarosite	P1	1	65.7	550	27.09	27.09
	P2	1	67	727	19.88	19.88
Quartz	P1	1	65	294	57.29	57.29
	P2	1	65	710	20.34	20.34
Silicon	P1	1	67	272	64.00	60.00
	P2	1	65	267	64.95	60.00
Vermiculite	P1	5	64.5	274	313.13	60.00
	P2	5	64.5	310	267.21	60.00
	P3	5	64.5	352	228.17	60.00

Table 3-9. Acquired and calculated parameters during Ti estimation

The estimation of NA provided the results shown in Table 3-10. The worse scattering samples, that could not acquire as much accumulations as desired (due to their high integration time) present lower final SNRs than expected, as a direct consequence of the limited operation available time (e.g., the vermiculite case). Worst case is the Clay + sand sample (in red in Table 3-10), which, after the whole operation time doesn't have Raman signals and thus SNR is set to 1.

Material	Point	Final Ti	Max NA (time limited)	Final NA	Expected SNR	Final SNR	Saved time
Alunite	P1	33.47	16	4	756	980	433
Alunite	P2	35.95	15	4	658	809	420
Basalt	P1	13.15	44	4	295	280	534
Basalt	P2	13.85	42	4	334	284	531
Basalt	P3	49.70	11	11	110	54	4
Calcite	P1	15.47	39	4	769	967	525
Calcite	P2	10.67	59	4	2568	4673	550
Chert	P1	60.00	9	4	630	611	300
Chert	P2	58.57	9	4	2111	1496	307
Clay+Sand	P1	4.14	143	143	209	1	4
Dolomite	P1	3.75	159	5	126	102	578
Dolomite	P2	2.26	264	7	123	85	582
Gossan	P1	60	9	4	123	126	300
Gossan	P2	60	9	4	235	191	300
Gypsum	P1	7.60	77	10	120	47	516
Gypsum	P2	8.80	67	13	121	42	477
Hematite	P1	60.00	9	4	236	304	300
Hematite	P2	60.00	9	4	159	221	300
Hematite-F	P1	35.56	15	15	70	20	31
Jarosite	P1	27.09	21	4	144	110	465
Jarosite	P2	19.88	29	4	147	92	501
Quartz	P1	57.29	9	4	1170	1310	314
Quartz	P2	20.34	28	4	260	212	498
Silicon	P1	60.00	9	4	4288	2196	300
Silicon	P2	60.00	9	4	5100	1949	300
Vermiculite	P1	60.00	9	9	84	61	0
Vermiculite	P2	60.00	9	9	81	44	0
Vermiculite	P3	60.00	9	9	67	51	0

Table 3-10. Acquired and calculated parameters during NA estimation

On the other hand, for several of the samples not limited by the operation time, the final NA is much lower than the maximum possible value. This implies saving operational time: considering the cases in which time is saved, in average, the algorithm saved 334 seconds per sample, while ensuring SNR values equal or higher to the expected ones (with an average SNR surplus of 20), and which is in most cases higher than the desired SNR value (120). These results are obtained without taking into account the samples with SNR higher than 1400. The reason is that the method used for the SNR calculation from one spectrum is less robust than a method using two spectra to calculate the noise, and small variations in the noise standard deviation imply very high variations on the final SNR value. For this reason, those samples introduced non-representative errors in the data.

Figure 3-45 depicts the correlation between the expected and the final SNR for the rest of the samples. As it can be seen, the correlation is quite linear, with a slope higher than 1, which implies that calculated final SNR is in general higher than the expected one (explaining the average SNR surplus explained above). The fact that the final SNR is higher than the expected ensures the robustness of the algorithm, which guarantees a minimum level of quality, as long as there are time resources available.
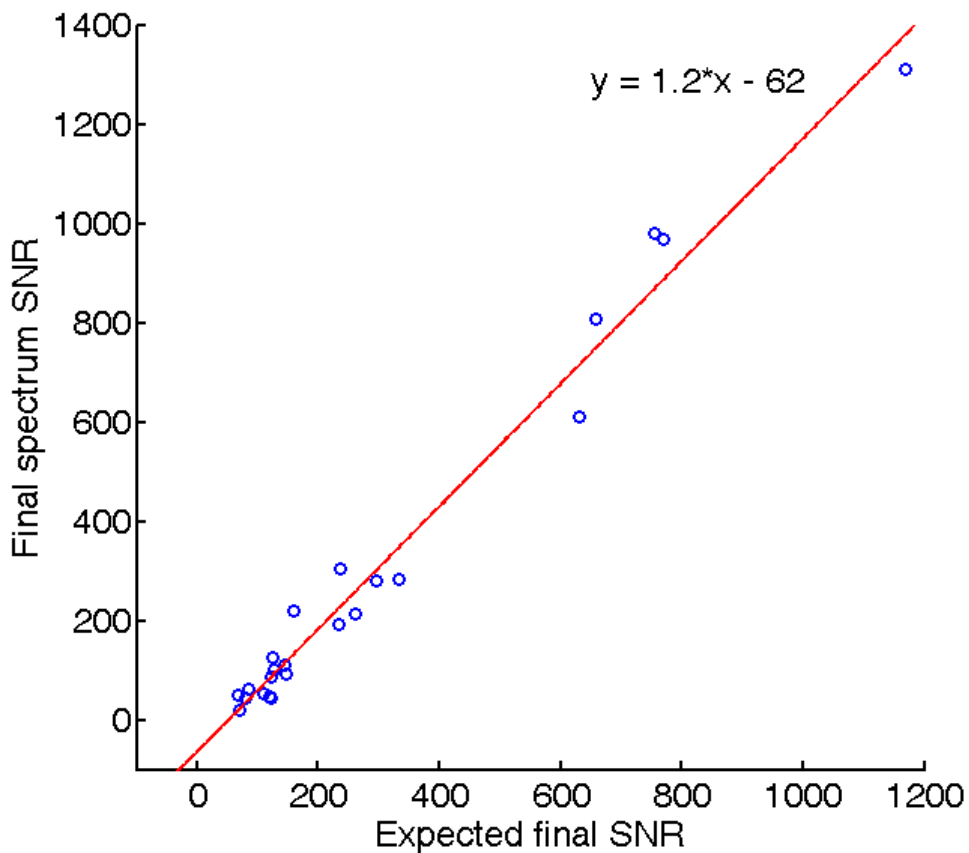
Figure 3-45. Correlation between spectra expected and final SNR

The proposed algorithm for the automated calculation of the acquisition parameters has shown an important potential reduction of the operation time, without compromising the quality standards of the acquired spectra (though it implies not obtaining the highest quality spectra possible). This is proposed and accepted by the scientific team, as the remainder of the time could be used to acquire more samples at different spots, as stated in the operational baseline of the instrument. Of course, there is some limit to the number of points that the instrument will be able to acquire, and the instrument will always comply with the data and time budgets. From the scientific point of view, this possibility is key to obtain the best science return from the instrument, and not performing in such a way would imply a very important loss of capabilities and science scope.

3.5.4 IMPLEMENTATION

This section details the functions that are expected to be needed for the implementation of this algorithm.

cut_spectrum_range(first_pixel, last_pixel) – Function to limit the spectral working range. Needed unless the inputs to the algorithms are already cut to this range.

remove_dead_pixels( ) – Function to remove the known dead/hot/damaged pixels from the input spectrum. Needed unless the inputs to the algorithms are already cut to this range.

smooth_spectrum(spectrum,filter_width) – Apply moving average window to spectrum

Specific for Ti:

max_position = calculate_max_positions_1Dspectrum(spectrum) – Function that returns the position of the maximum intensity on a 1D (vector) spectrum.

calculate_max_value_ROI(max_position) – Function to calculate the maximum intensity from a column in the ROI.

calculate_CCD_bias( ) – Function to calculate the average bias in a (small) non-illuminated spectrum region

calculate_final_Ti(several_inputs) – Function to calculate the final Ti as per equation in section 3.5.3.2.

Specific for NA:

Cosmic ray detection and removal as per section 3.3.3.4**.**

Calculate SNR as per section 3.4.5**.**

NA_aux = calculate_SNR_factor_and_check_LUT(calculated_SNR, expected_SNR) – Function to return the number of accumulations necessary to reach a determined SNR level.

final_NA = verify_NA_value(NA_aux) – Function to check whether the required NA_aux fits into the total available operation time constraints and parameter limits.

Part II

Chapter 4.

Consequences of the sample preparation and distribution procedure onboard the ExoMars rover on the RLS data

Description of the sample preparation and distribution procedure onboard the ExoMars rover regarding crushing and flattening issues, and the consequences on the scientific return of the RLS instrument.

4.1 INTRODUCTION

The operation mode imposed by engineering restrictions in the ExoMars rover implies that the instruments in the Analytical Laboratory Drawer (ALD), such as RLS, will analyze flattened powdered samples. The crushing, delivery and flattening processes will have an impact on the instrument, as there are many factors that can influence it: grain size distribution (especially the lower distributions), flattened surface rugosity, cross-contamination, etc. This chapter covers issues related with the flattening, delivery and crushing issues, both from the sample preparation point of view and its influence on the acquisition results. In addition, it addresses the consequences suffered on the Raman spectra due to the crushing process of the samples, with respect to analyzing bulk materials.

Section 4.2 summarizes the results from a test performed in collaboration with the engineering team of the ALD performed at Kayser-Threde premises with the Sample Preparation and Distribution System breadboard of the rover, integrated with the RLS ExoMars Simulator. This test was performed in order to assess and evaluate the sample preparation and distribution system with samples considered relevant by the RLS team, as well as to assess the feasibility of the collaborative science between RLS and MicrOmega.

Section 4.3 presents the work performed with bulk and powdered samples with different granulometries, and the effects observed on the Raman spectra when analyzing powdered samples instead of bulk samples. It also addresses the issues related to the textural context of the samples during the crushing process, and the consequences for its analytical results.

4.2 SAMPLE FLATTENING AND CRUSHING EVALUATION

The RLS ExoMars Simulator, as described in Chapter 2, is implemented with a flattening blade to simulate the flattening process of the Sample Preparation and Distribution System (SPDS) of the rover. However, the simulator is not as representative as the SPDS itself, and so, in September 2013, a joint end-to-end (E2E) test campaing between the SPDS and RLS instrument was performed at Kayser-Threde premises in Munich. This test had the following objectives:

- Experimental verification of the SPDS capability to prepare and distribute samples in line with the RLS (and MicrOmega) needs.

- Investigation of SPDS capability to support the MicrOmega and RLS combined science requirements (in particular, positioning performance).

- Optimization of SPDS operational procedures to maximize scientific outcomes.

This test campaign constituted the first interactive activity between the SPDS system breadboard and the ALD scientific payload. The setup included a laboratory RLS

breadboard (part of the RLS ExoMars Simulator) with functionality and characteristics similar to the actual system (though non-representative from the engineering point of view). The MicrOmega instrument was emulated with the camera included in the RLS ExoMars Simulator. As explained in Chapter 2, this camera does not represent the MicrOmega characteristics. However, it does from a functional point of view in which a point of interest is detected by this camera in the RLS path, which is then analyzed by the RLS instrument.

Four tests (plus an alignment and calibration tests) were performed during this campaign, in order to address all the objectives. The setup is described in section 4.2.1, while the results are described in sections 4.2.2 to 4.2.6. Section 4.2.7 summarizes the conclusions from this test campaign.

4.2.1 SETUP AND METHODOLOGY

4.2.1.1   SPDS-RLS-MICROMEGA EMULATOR JOINT SETUP

In order to couple the RLS breadboard and MicrOmega emulator to the SPDS system, an interface (I/F) bracket was mechanized by Kyser-Threde (KT) which allowed the manual adjustment of the camera centers of both RLS and MicrOmega cameras on the arc along which the refillable container (RC) was to be placed. The RLS system was accommodated in a position representative of the one occupied in the ALD. The MicrOmega emulator was placed as well in a representative position, though with a smaller angle than the one in the ALD. This is due to set-up constraints, but this is still representative for the test scope and did not jeopardize the test results. A general view of the final setup is shown in Figure 4-1.
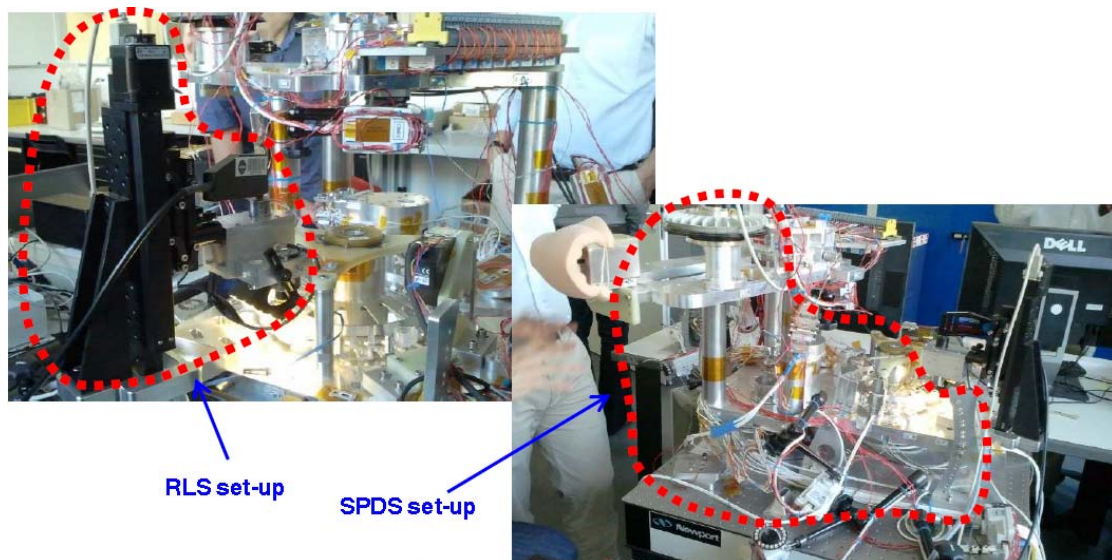
Figure 4-1. SPDS-RLS-MicrOmega emulator joint setup

The analysis with Raman is sensible to ambient light. Thus, to compensate the lack of shutters or blinds at KT premises, KT built a structure to cover the setup providing the necessary low-light environment (see Figure 4-2A).

In order to avoid damage between the SPDS and the RLS/MicrOmega setup, the RLS hard stop was configured to avoid the RLS system getting closer than 9 mm to the RC surface (see Figure 4-2B).
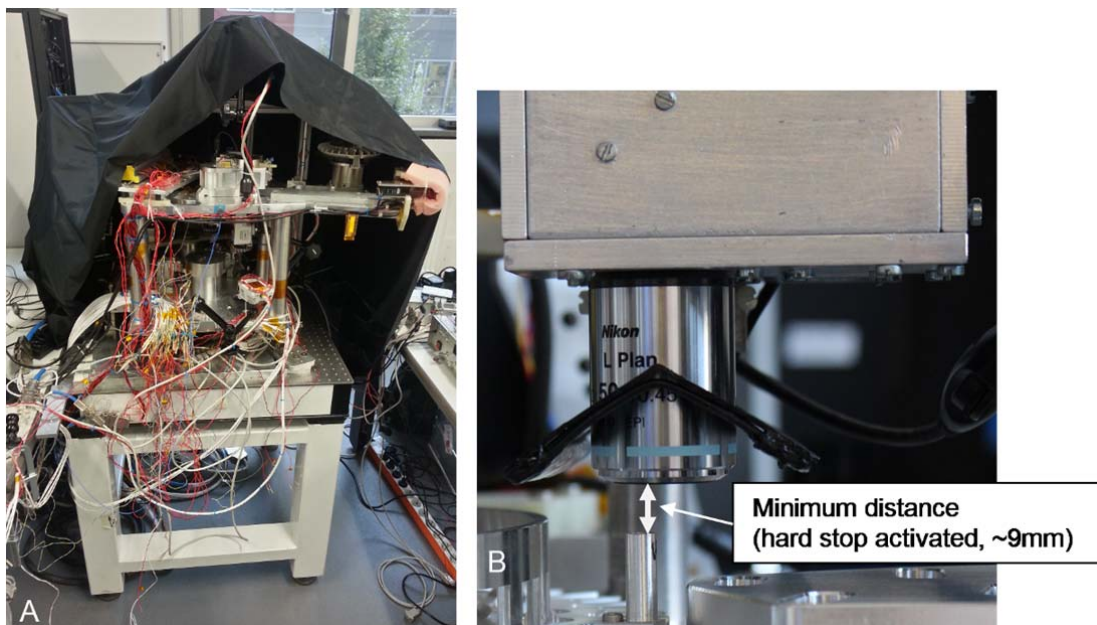
Figure 4-2. Setup darkening structure (A) and RLS-SPDS safe distance (B)

4.2.1.2   SAMPLE SELECTION

For this test campaign, a set of samples was made available by the SPDS (Figure 4-3) and RLS (Figure 4-4) teams. To overcome the uncertainty on the testing time, and to try to keep the testing window limited to one week, this set of samples had a high number of specimens, which were classified depending on priorities in order to be able to optimize the tests to the available time. The selected samples and assigned priorities were agreed by all parties (RLS, SPDS, TAS-I and ESA). Details are shown in Table 4-2 together with the note indicating which of them have been actually tested during the test campaign. Some of the samples provided by the RLS team were provided with different grain sizes distributions: 45-125 microns, 125-250 microns and the so-called SPDS-like. The latter, though not exactly following the grain size distributions described in the crusher reports, were prepared as shown in Table 4-1. This was to try to find a standard grain size distribution more or less representative of the expected samples provided by the crusher.
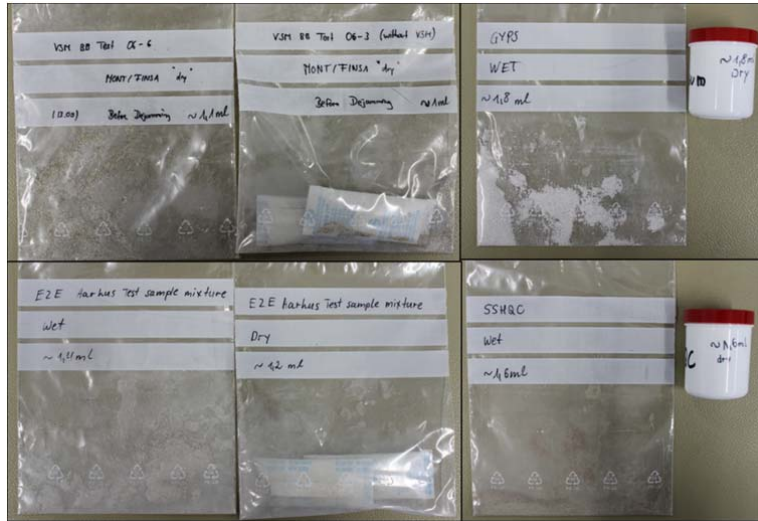
Figure 4-3. **Set of samples provided by KT**

Figure 4-4. Set of samples provided by RLS

Sample	Total weight (g)	Granulometry	Weight (g)
Calcite	1.111 ± 0.001	100% in 45-125um	1.111 ± 0.001
Calcite	1.110 ± 0.001	100% in 125-250um	1.110 ± 0.001
Calcite	1.502 ± 0.006	10% <45um	0.149 ± 0.001
		37% in 45-125um	0.556 ± 0.001
		27% in 125-250um	0.407 ± 0.002
		26% in 250-500um	0.390 ± 0.002
Gypsum	0.916 ± 0.001	100% in 45-125um	0.916 ± 0.001
Gypsum	0.916 ± 0.001	100% in 125-250um	0.916 ± 0.001
Gypsum	1.000 ± 0.004	10% <45um	0.100 ± 0.001
		37% in 45-125um	0.370 ± 0.001
		27% in 125-250um	0.270 ± 0.001
		26% in 250-500um	0.260 ± 0.001
Gossan	1.349 ± 0.001	100% in 45-125um	1.349 ± 0.001
Gossan	1.396 ± 0.001	100% in 125-250um	1.396 ± 0.001
Gossan	1.499 ± 0.004	10% <45um	0.150 ± 0.001
		37% in 45-125um	0.555 ± 0.001
		27% in 125-250um	0.404 ± 0.001
		26% in 250-500um	0.390 ± 0.001
Quartz	1.500 ± 0.004	10% <45um	0.150 ± 0.001
		37% in 45-125um	0.555 ± 0.001
		27% in 125-250um	0.405 ± 0.001
		26% in 250-500um	0.390 ± 0.001
Quartz (85%) + Olivine (15%)	1.499 ± 0.003	60% Quartz in 45-125um	0.900 ± 0.001
		25% Quartz in 125-250um	0.375 ± 0.001
		15% Olivine in 250-500um	0.224 ± 0.001
Basalt	1.502 ± 0.004	10% <45um	0.150 ± 0.001
		37% in 45-125um	0.555 ± 0.001
		27% in 125-250um	0.406 ± 0.001
		26% in 250-500um	0.391 ± 0.001
SVBF09 Svalbard Carbonate AMASE	1.442 ± 0.004	10% <45um	0.144 ± 0.001
		37% in 45-125um	0.534 ± 0.001
		27% in 125-250um	0.389 ± 0.001
		26% in 250-500um	0.375 ± 0.001
Coquimbite	1.030 ± 0.004	10% <45um	0.103 ± 0.001
		37% in 45-125um	0.381 ± 0.001
		27% in 125-250um	0.278 ± 0.001
		26% in 250-500um	0.268 ± 0.001

Table 4-1. Description of the samples prepared by the RLS team

| Test	Sample	Comment	Supplier	Priority	Status
Test 0.1 - Setup Alignment & Calibration	Calibration grid		RLS	1	Done
Test 0.2 - Functional Test (Ref. Samples)	Cyclohexane		RLS	1	Done
	Powdered calcite	TBC, pending results on Cyclohexane	RLS	2	Not done, not necessary, set-up was ok.
	Bulk calcite	TBC, pending results on Cyclohexane and powd. Calcite	RLS	3	Done
Test 1 - Flattening performance + small grain influence	Gypsum diff. grain sizes (SPDS-like)		RLS	1	Done, 1 repetition.
	Gypsum diff. grain sizes (50-125)		RLS	2	Done, 1 repetition.
	Calcite diff. grain sizes (SPDS-like)		RLS	3	Done, 1 repetition.
	Calcite diff. grain sizes (50-125)		RLS	4	Done, 1 repetition.
	S-4 WET (Gypsum/Gypsaarenite)		KT	5	Done, 1 repetition.
	S-1-A DRY (Sandstone high quartz content)		KT	6	Skept
	Gypsum diff. grain sizes (125-250)		RLS	7	Skept
	Calcite diff. grain sizes (125-250)		RLS	8	Skept
	Hematite+Goethite diff. grain sizes (50-125)	Offline	RLS	5	Will be done by KT offline, not necessary to analyze it with RLS BB.
Test 2 - Cross contamination	Hematite+Goethite diff. grain sizes (SPDS-like)		RLS	1	Done, 1 repetition.
	S-4 WET (Gypsum/Gypsaarenite)		KT	1	Done, 1 repetition.
	S-4 WET (Gypsum/Gypsaarenite) cross-cont by Hematite after Hematite and Gypsum dosing	To repeat offline (Test 4)	KT	2	Done. Hematite grains found on the sample surface but not on the RLS spot track. Material mixed in a bag, spectral analysis will be repeated offline on mixed matarial to search hematite grains.
	S-4 WET (Gypsum/Gypsaarenite)		KT	3	Skept
	Hematite+Goethite diff. grain sizes (125-250)		RLS	4	Skept
	S-4 WET (Gypsum/Gypsaarenite)		KT	4	Skept
	Quartz		RLS	5	Skept
Test 3 - Combined science	Quartz (85%) + Olivine (15%) mixture	Some grains of basalt added for visual reference	RLS	1	Done. 1st test no olivine grains in the RLS spot path. 2nd test performed adding basalt some olivine grains in the RLS spot path. Performed 1 ripetition, plus several investigations.
Test 4 - Representative samples	S-7 DRY (Montmorillonite 67%, MgSO4 7H2O 30%, perchlorate 3%)		KT	1	Done. 1 repetition plus one dosing after 1 night with sample in the DS and rel. humidity 53%.
	Basalt from Almagro (Spain)		RLS	2	Done, only 10 points.
	E2E processed material, mix from Mars Env. Test.		KT	4	Done, only 10 points.
	Carbonates from Svalbard / Sulfates from Rio Tinto	TBC	RLS	3	Skept
	Cross contaminated sample of Sandstone and Gypsum (TBC) coming from SPDS E2E test campaign.	Offline	KT	5	Will be done offline by RLS.
Samples to crush	Hematite bulk		RLS	1	Crushed about 1 hour with VSM in not optimal conditions.
	Gypsum bulk		RLS	2	
	Vermiculite, plus other bulk materials provided by RLS.	Offline	RLS	3	Will be done offline by KT, powdered sample will be sent to RLS for analysis. Separate test report will be prepared. Picture of the flattened sample will be taken. Crushing will be performed without cleaning the CS.

Table 4-2. Samples used during the SPDS-RLS E2E test campaign

4.2.2 Test 0: alignment, calibration and setup test

4.2.2.1 Test 0.1: SPDS-RLS alignment and calibration

The aim of this test was the alignment and calibration of the system, which is schemed in Figure 4-5. It was divided in two steps: First step was to correctly setup the RLS+MicrOmega emulator in order to have both systems field of views (FoV) over the carrousel line. To do so, the sample refillable container (RC) was placed in the nominal position of the RLS optical head. The RLS laser was then turned on and the RLS longitudinal positioner adjusted to have the spot in the center of the RC. The spot position was marked on a piece of kapton tape stick to the RC. Then, the RC with the kapton tape mark was placed in the nominal position of the MicrOmega emulator. With the laser still on, the I/F bracket between RLS and SPDS was pivoted with respect to the laser spot in order to place the reference mark under the MicrOmega FoV. Once this was achieved, the SPDS-RLS I/F bracket was fixed to this position.
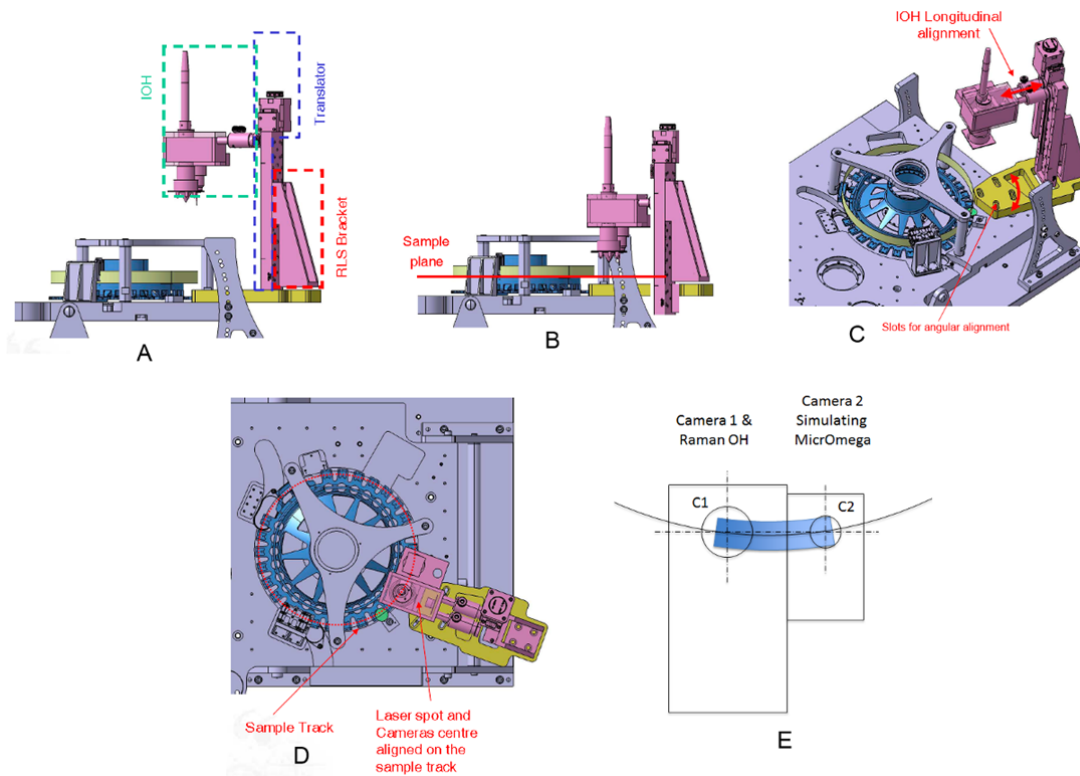
Figure 4-5. Schematics of the SPDS+RLS setup (A, B, D, E) and degrees of freedom for adjustment (C).

Second step was the fine calibration of the angles between RLS and MicrOmega. A calibration grid was available to perform this calibration (Figure 4-6). However, the mark on the kapton tape of step 1 had the right size and very clearly differentiable features. In addition, it also provided higher robustness than the calibration slide in terms of relative movement with respect to the RC due to the movement of the carrousel. Thus, it was decided to use this mark on the kapton tape for the detailed

calibration between RLS and MicrOmega emulator angles (Figure 4-7). The first run of this test consisted on moving the RC under different positions of the MicrOmega emulator FoV, writing down the positions as well as the angles at each position. An image was acquired at each point.
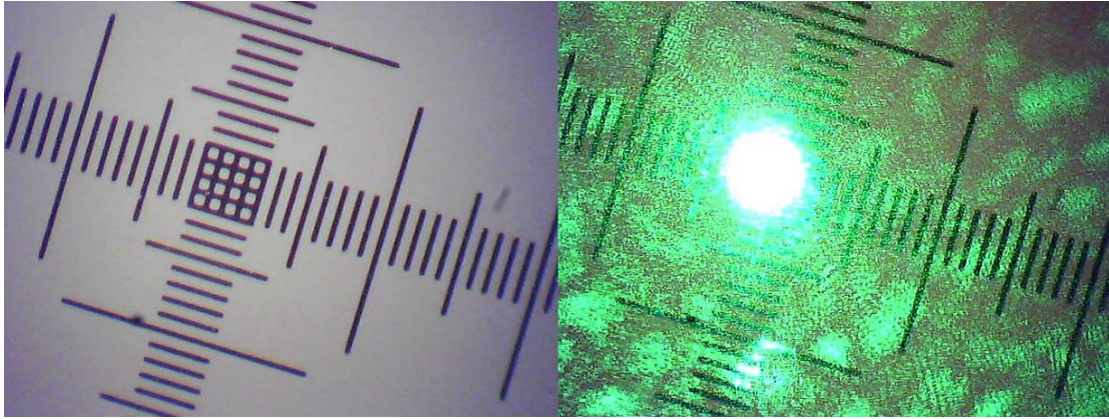
Figure 4-6. Calibration grid and laser spot position in the RLS FoV
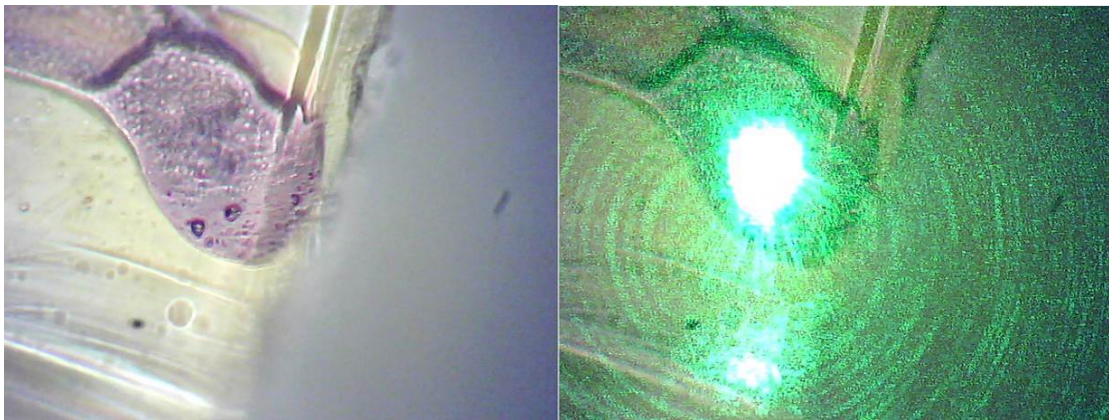
Figure 4-7. Calibration mark and laser spot position in the RLS FoV.

The results of this test provided the region on the MicrOmega FoV that passed below the RLS spot (Figure 4-8), as well as the angle shift between both instruments, taking into account the shift in the MicrOmega FoV with respect to the reference point. Nominal angle between the reference point (furthest point on the MicrOmega FoV) and the RLS spot position was 22.591º.

The correlation between the different positions on the MicrOmega FoV with respect to the angular shift provided that the path could be approximated by a straight line (as the quadratic coefficient was negligible), with 0.0005º shift per micron in the MicrOmega image. This way, any new point selected on the MicrOmega FoV could be placed under the laser spot by modifying the nominal angular shift depending on the distance of the selected point from the reference one.

Figure 4-8. Reference for the calibration of the RLS spot path under the MicrOmega emulator FoV.

The following day a second run of the calibration test (shortened version) was performed to verify that the relative position between the RLS and MicrOmega was still the same.

4.2.2.2 TEST 0.2: FUNCTIONAL TESTS

The aim of this test was to check that the performance of the instruments was correct after the integration of the RLS and SPDS in the same setup. This way, the SPDS checked all the subsystems and found no mal-functioning. The RLS laser power was checked to ensure that no damage had come to it during the trip. The end-to-end behavior of the complete Raman setup was checked with two different samples: bulk calcite in automatic mode, and cyclohexane with a manual acquisition. The end-to-end RLS simulator check was also correct. Laser excitation, spectrometer and detector were working under the nominal expected conditions. As an example, Figure 4-9 shows the spectra acquired during this test.
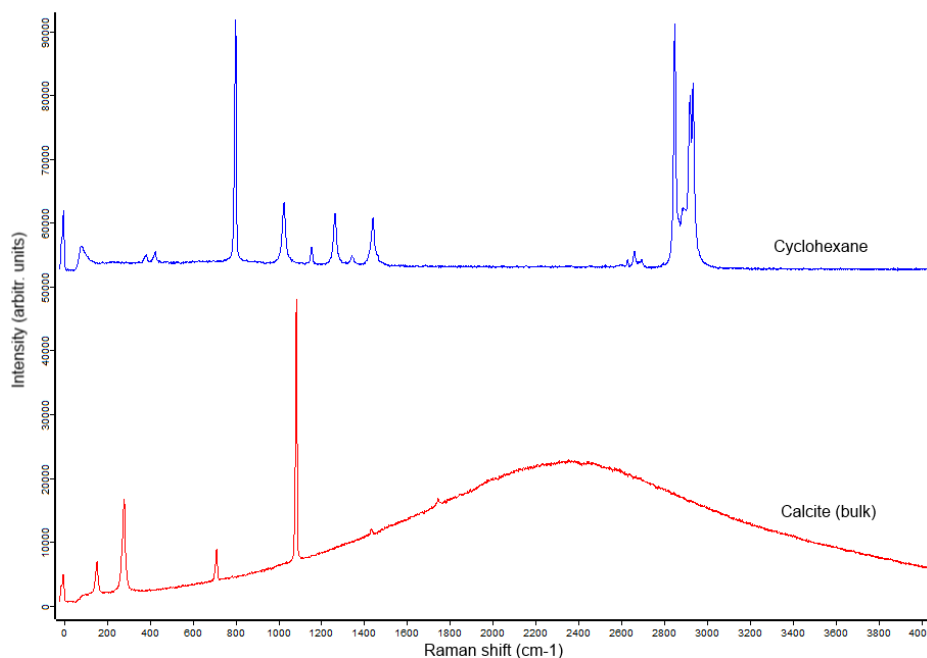
Figure 4-9. Spectra from reference samples in Test 0.2

129

4.2.3 TEST 1: SAMPLE DOSING AND FLATTENING PERFORMANCE AND SMALL GRAIN INFLUENCE

To evaluate the flattening performance of the SPDS and to illustrate the small grain influence in two relevant cases, several samples were used, with different grain distributions. Namely, calcite and gypsum, with 45-125 microns and SPDS-like grain distributions. These samples were dosed and flattened with the SPDS system following the standard procedure, consisting of 3 dosing and flattening cycles, prior to presenting the sample to the instruments. The sample surface was then measured with a 2D profiler installed in the SPDS. The RC was then placed below the MicrOmega emulator camera and several pictures were acquired to complete a whole image of the sample surface. Then, the RC moved to be placed below the RLS instrument, which acquired a total of 25 points of the sample surface, including AF at each point plus the acquisition of spectra at each point (with the acquisition parameters optimized at each point). The points were acquired each 0.2º, equivalent to around 400 microns, covering the 10 mm of sample area.

4.2.3.1 DOSING PERFORMANCE

The complete dosing procedure consists on several doses on the RC, with several flattening steps as well. In order to ensure a good dosing and avoid a "dome" effect on the funnel, a piezoelectric actuator is used. The process is illustrated in Figure 4-10.
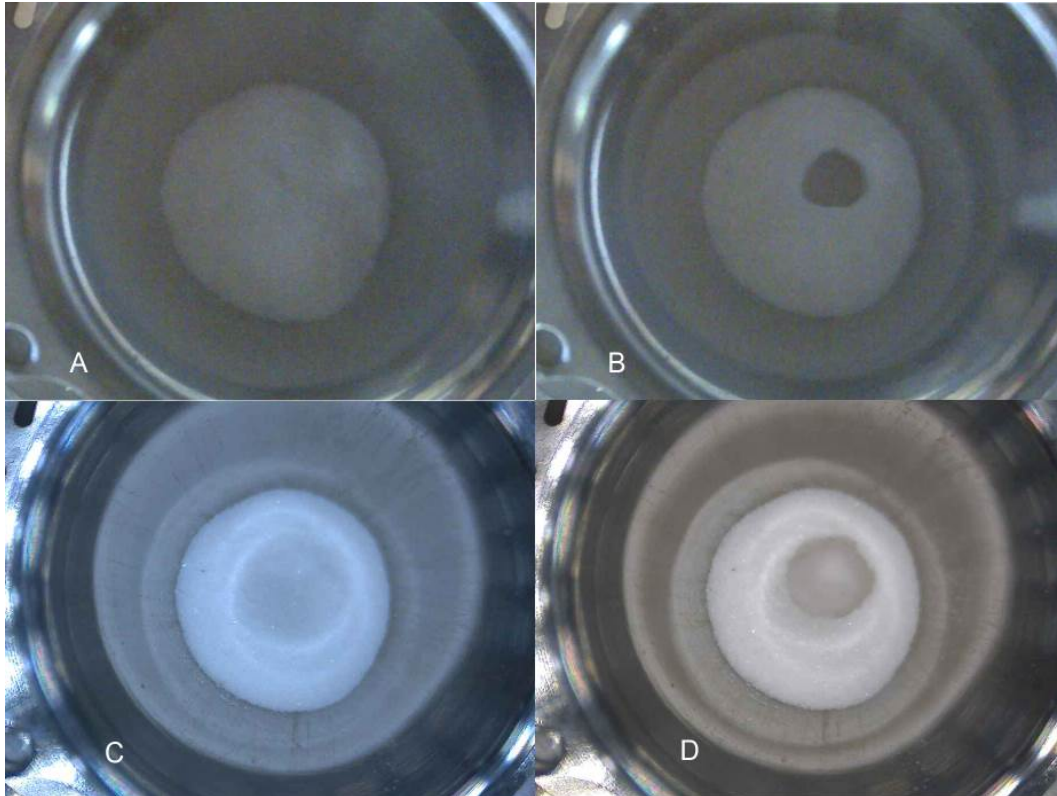
Figure 4-10. Funnel contents after dosing (B, D) and piezoelectric actuation (A, C)

4.2.3.2   FLATTENING PERFORMANCE

The SPDS flattening procedure was performed as follows:

- Dose twice CW on PSDDS/PSHS I/F
- Drive PSHS 2° CCW
- Dose once CW
- Drive PSHS 2° CCW
- Dose once CW
- Flatten sample with flattening speed of 1°/s
- Dose once CCW on PSDDS/PSHS I/F
- Flatten sample with flattening speed of 0.2°/s
- Dose once CCW on PSDDS/PSHS I/F
- Flatten sample with flattening speed of 0.2°/s

A piezoelectric vibration sequence (three vibrations of 5s) was performed after every dosing and before the first dosing (to ensure correct filling and emptying of the dosing cup and to loosen up the sample to avoid cementation).

The resulting sample surface after this procedure is shown in Figure 4-11 for the calcite 45-125 um sample and in Figure 4-12 for the gypsum 45-125 microns sample.

Figure 4-11. Sample surface after the first (A), second (B) and third (C) dosing and flattening steps for the calcite 45-125 um sample. The cross-contamination that can be seen is discussed in section 4.2.4.

Figure 4-12. Sample surface after the first (A), second (B) and third (C) dosing and flattening steps for the gypsum 45-125 um sample.

The pictures taken with the MicrOmega emulator for one of the samples are shown for reference in Figure 4-16. The MicrOmega DoF is quite long, thus, the focusing at each point didn't provide a lot of information, except to confirm when the image were at the extremes of the RC (see lower part in Figure 4-16) and thus is not really a good indicator of the sample surface profile.

The focusing positions at each of the 25 points analyzed with the RLS instrument provide a profile for each sample that are depicted and compared in Figure 4-13 for calcite and Figure 4-14 for gypsum. Figure 4-15 graphically shows the results obtained with the surface 2D mapper incorporated to the test setup. As expected, the lower grain size distributions present flatter profiles than the SPDS-like distributions. The standard deviation over the average and the maximum excursion of the flattened surface obtained from the AF points of the RLS and the 2D mapping of the surface profiler on the SPDS are compared in Table 4-3 and Table 4-4, respectively. As it can be seen, the maximum peak to peak values of RLS measurements are lower than the values with the 2D mapper in the whole surface, which is consistent with the expected results. In addition, there is some kind of correlation between the measurements by RLS and the surface measurements performed by the laser 2D mapper.

The flatness requirement for the SPDS states that "The Flattening Device shall be capable to flatten the top surface of the Sample dosed in the Refillable Container so that 80% of all sample surface measurement points are within ±0.2mm of the expected surface". Table 4-5 shows the percentage of points that are found over +-100 and +-200 microns from the reference plane placed at 200 microns below the sample container surface. The disparity that can be found, especially for the Calcite SPDS-like sample could be explained by the limited number of points evaluated by RLS and the 2D surface analysis: as shown in Figure 4-13, there are several points which are below -400 microns in this sample, which has an important impact on the average, given the small number of analyzed points.
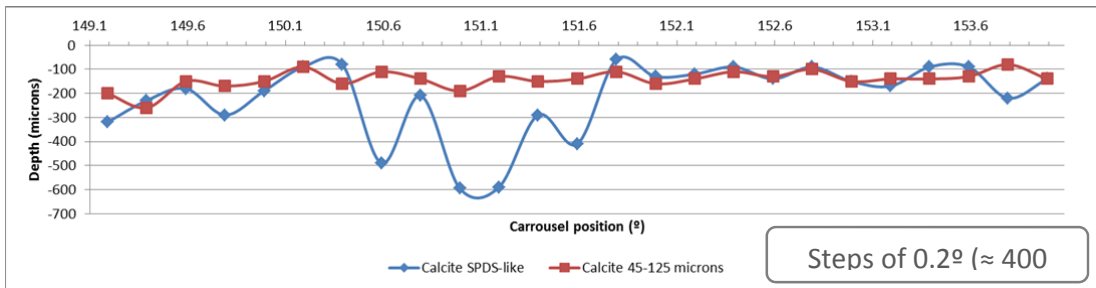
Figure 4-13. RLS AF positions for the calcite samples
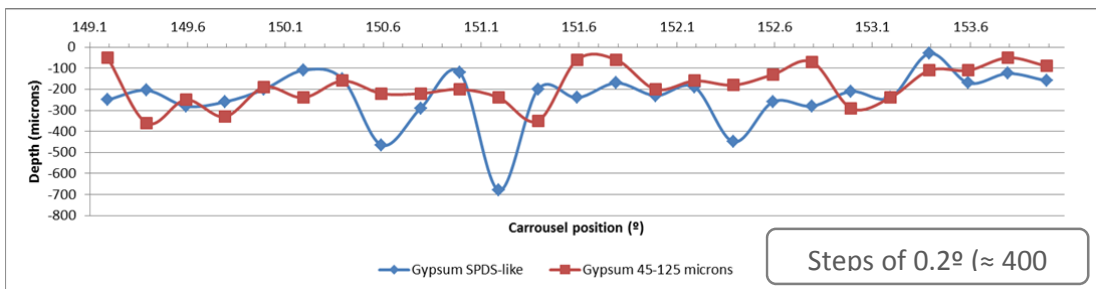
Figure 4-14. RLS AF positions for the gypsum samples

Sample	RLS AF position	SPDS (whole area)	SPDS (one line)
Calcite SPDS-like	155	93	82
Calcite 45-125 microns	37	30	36
Gypsum SPDS-like	132	108	103
Gypsum 45-125 microns	94	51	53

Table 4-3. Standard deviation (in microns) from average

Sample	RLS AF position	SPDS (whole area)	SPDS (one line)
Calcite SPDS-like	535	558	495
Calcite 45-125 microns	180	251	189
Gypsum SPDS-like	650	768	546
Gypsum 45-125 microns	310	522	304

Table 4-4. Maximum peak to peak values (in microns)

Sample	RLS		SPDS (surface)	
	+-200 µm	+-100 µm	+-200 µm	+-100 µm
Calcite SPDS-like	84%	52%	96%	73%
Calcite 45-125 microns	100%	88%	100%	82%
Gypsum SPDS-like	88%	84%	92%	66%
Gypsum 45-125 microns	100%	64%	100%	92%

Table 4-5. Percentage of points within +-200 and +-100 microns over the average surface for RLS and SPDS
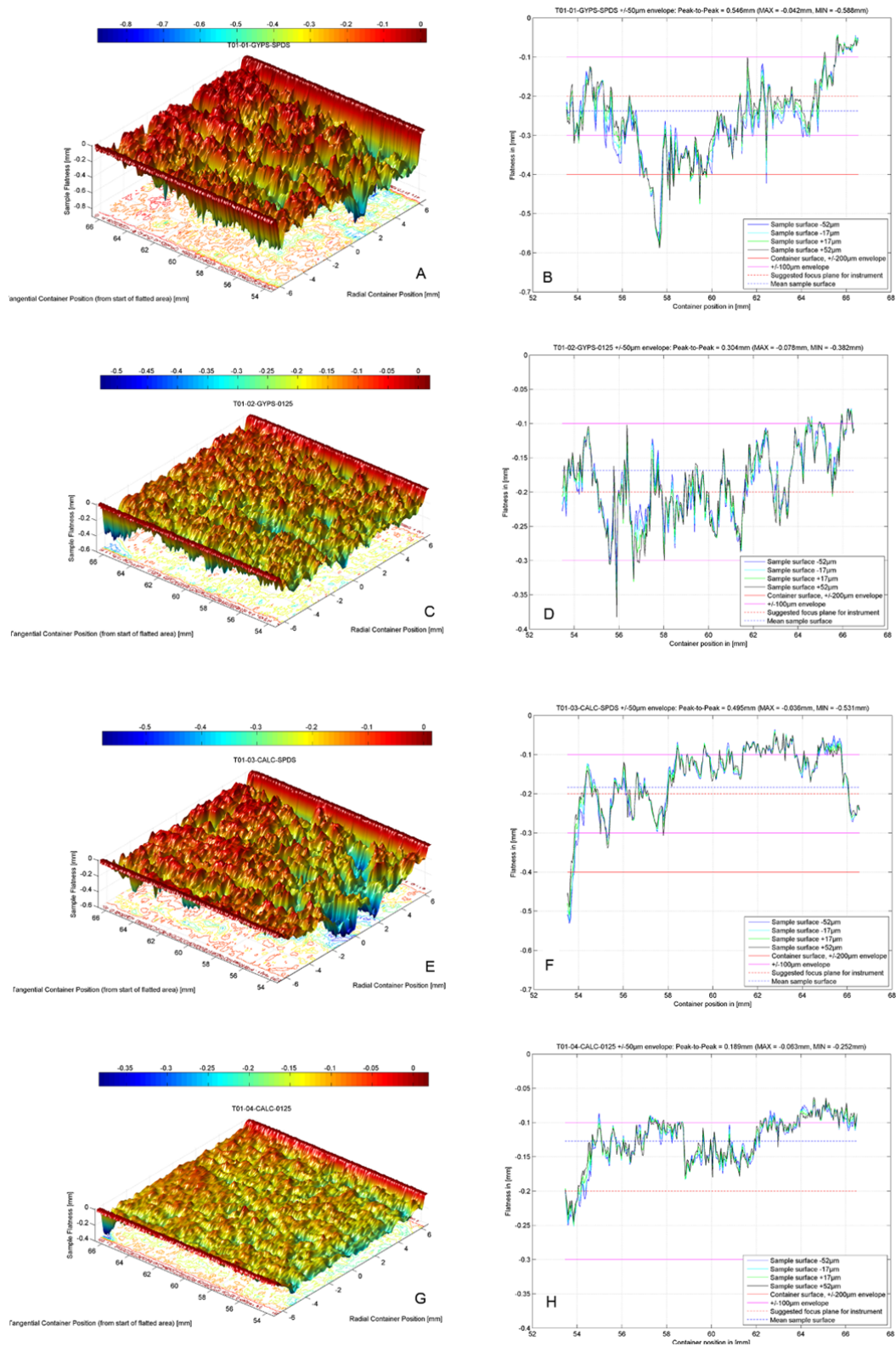
Figure 4-15. Results from the surface mapping performed with the 2D laser profiler
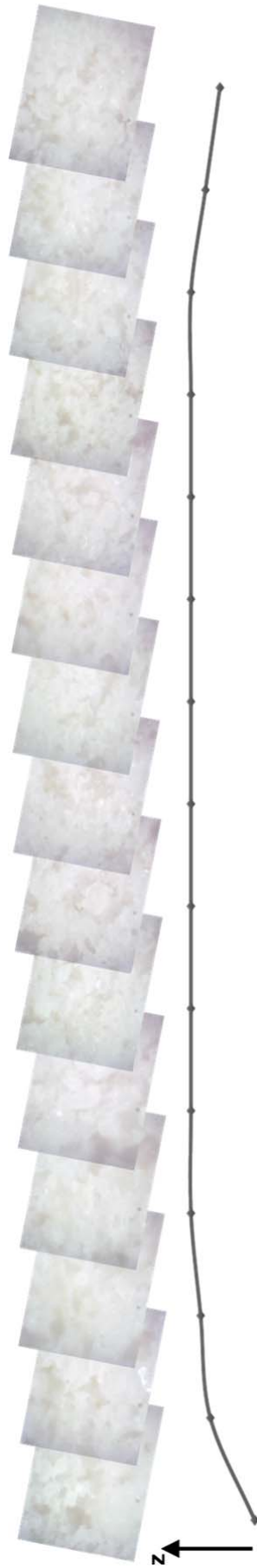
Figure 4-16. MicrOmega emulator camera images of the Gypsum SPDS-like sample during Test 1, and its relative focusing position for each image

Figure 4-17. MicrOmega emulator camera images of the gypsum S-4_WET sample in Test 2. The reddish cross-contamination can be observed throughout the images

135

Guillermo Lopez-Reyes

4.2.3.3   SMALL GRAIN INFLUENCE

It has been reported [54] that the grain size distribution and the crushing process of the sample has an impact on the Raman spectrum of the materials. The crushing process implies an important loss from the geological point of view, as the mapping on a powdered surface will not provide information on the texture and context of the sample, key information for the analysis of alteration and formation processes. In addition, there is usually an increase on the background of the spectrum, which implies a decrease in the spectrum SNR. As a general rule of thumb, the highest grain size, the better performance of the Raman instrument, when spectra are acquired with the same parameters, though other sample-related parameters have also influence on the results. Further development on this issue is presented in section 4.3.

During the execution of this test, the automatic acquisition algorithm for the calculation of the integration time and number of accumulations was used. As shown in Figure 4-18, the averaged spectra of the different samples of calcite show a much lower background level (higher SNR) for the bulk calcite, and increasing background with decreasing grain size.
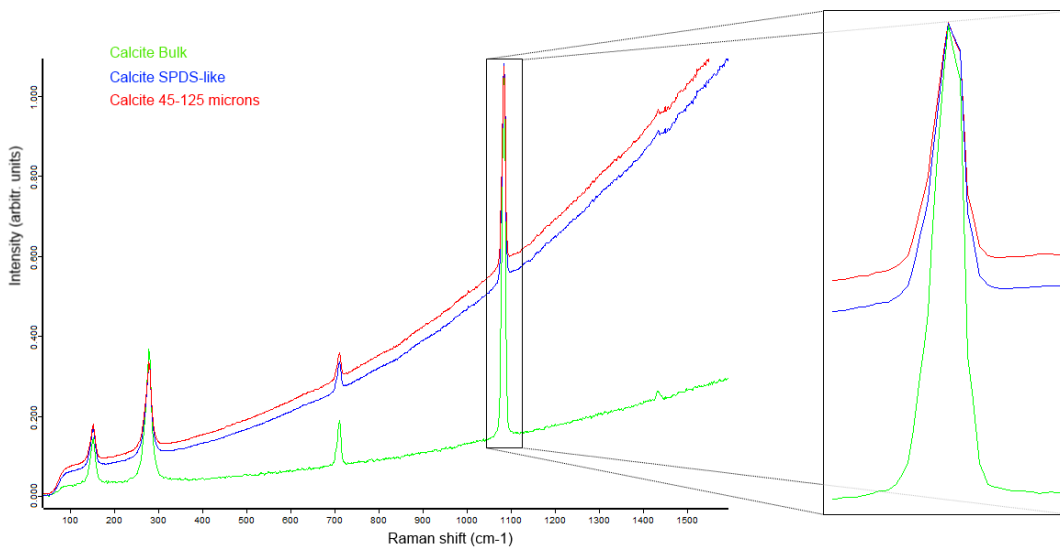
Figure 4-18. Background increase on the Raman spectra of calcite due to different grain size distributions

The analysis of gypsum with different grain sizes, on the other hand, provides a good example of the influence of other inter-dependent parameters (see Figure 4-19), such as the transparency of the sample and the focusing accuracy. With very transparent samples multiple reflections will occur, especially with lower grain sizes. However, if the sample is more opaque, lower grain sizes mean that the laser spot will be better focused, thus providing higher signals. In any case, the comparison with the bulk sample spectrum shows how, in every case, the background increases and the SNR decreases after the crushing of the sample.

136
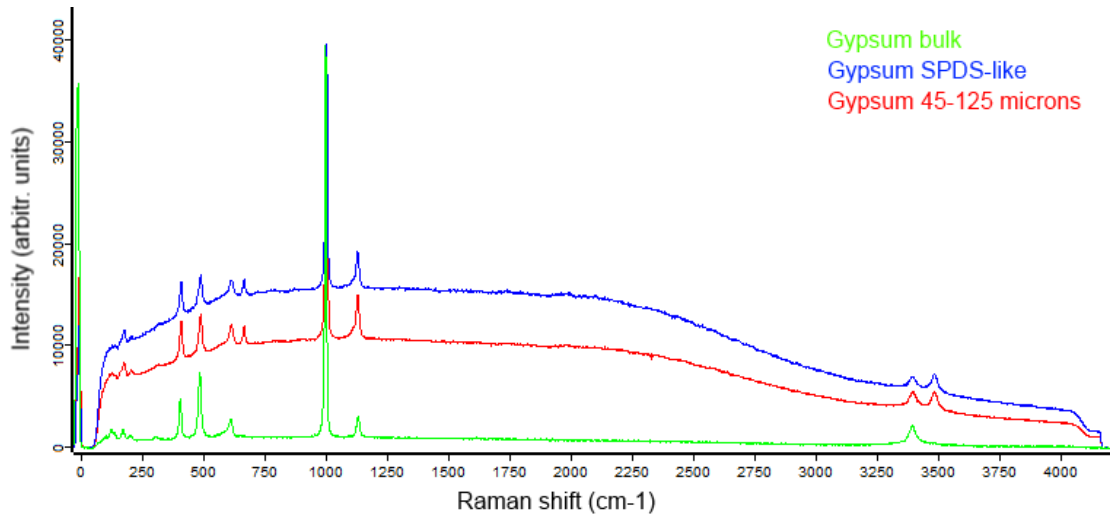
Figure 4-19. Counter-example of background increase. Some samples might present different behavior depending on some sample-related characteristics. A good flattening can somewhat compensate the small grain influence.

As a general conclusion, the crushing process and the presence of small grain sizes induces an increase on the background of the signal with respect to the bulk sample, as it has been reported in the literature, in previous internal studies by the RLS team, and presented here. Exceptions to this general rule may occur between different powdered distributions depending on some features of the samples, and also can be somewhat masked or corrected by the RLS automatic calculation of parameters. SPDS can only influence grain size by changing the gap size of the crusher, and this parameter will be pre-set during the system integration. In addition, if the presence of small grains is unavoidable, then there is an influence on the results of the flattening performance: the better the flattening, (potentially) the better the results, as more material will be placed in focus.

4.2.4 TEST 2: CROSS-CONTAMINATION

This test was mainly aimed at evaluating the cross-contamination performance from the dosing, cleaning and flattening processes (without taking into account the cross-contamination from the crushing station). For this test, the cameras of the SPDS were installed in positions to observe the flattened surface as well as the cleaning performance of the SPDS system.

Two different samples were dosed and flattened during this test. In the first place, the hematite sample with SPDS-like grain size distribution was dosed on the clean RC (though the blade had not been cleaned after the previous test with Gypsum_45-125um), and immediately cleaned after acquiring some pictures with the SPDS cameras and one spectrum with RLS for reference (the process is illustrated in Figure 4-20). The reason this sample was chosen was because the iron oxides are good candidates to produce cross contamination, due to their coloring properties (they are

137

sometimes used for dyeing purposes). In addition, it is a very relevant sample as they are easily found in Mars [55] (these materials provide the red color of the planet).
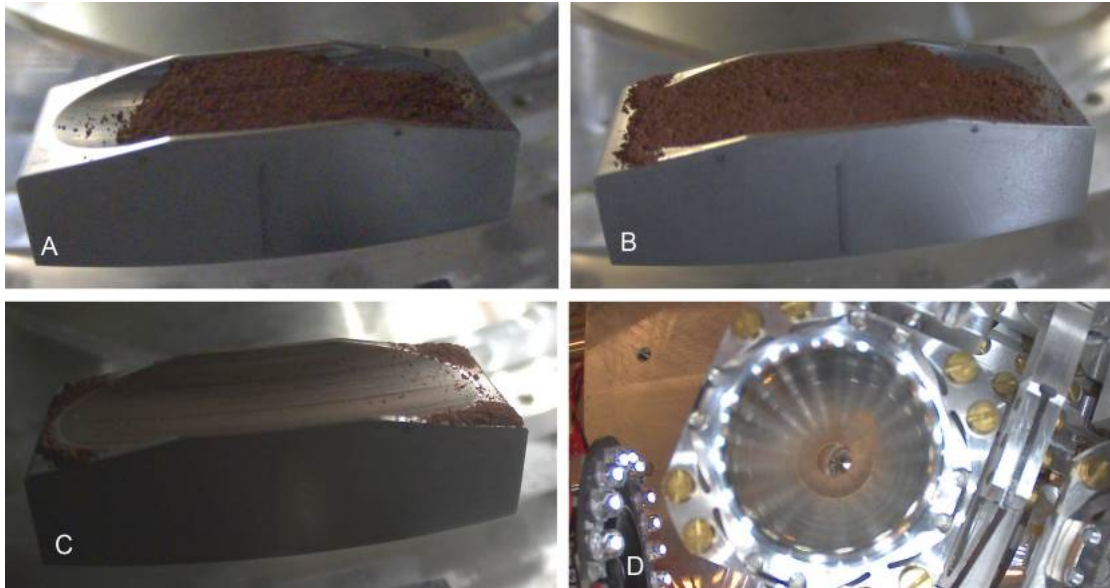
Figure 4-20. Dosing, flattening (A, B) and cleaning (C) processes with the hematite SPDS-like grain size distribution sample. The sample remains in the funnel (D) and RC (C) are the ones that could result in cross-contamination in the next sample.

After dosing and cleaning the hematite sample, a white sample (gypsum, reference S-4_WET) was dosed and flattened (see Figure 4-22). No remarkable differences in terms of visible cross-contamination on the sample surface between dosings were appreciated in the images from the SPDS setup, though a reddish coloration is present in both cases. This was probably due to the remains from the oxide sample in the flattening blade, as shown in Figure 4-21.
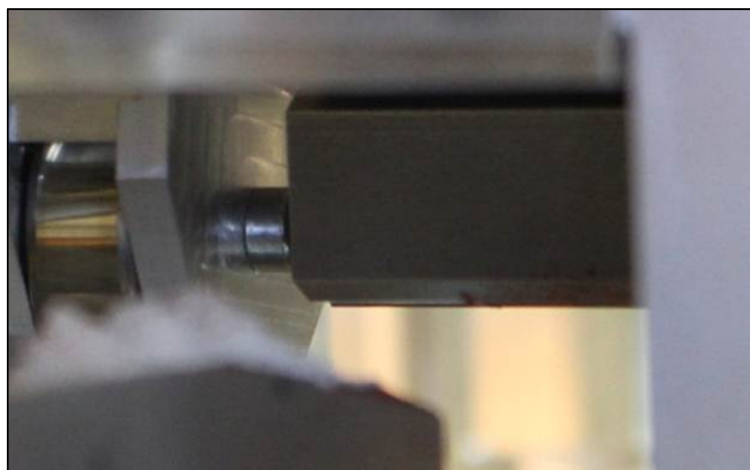
Figure 4-21. Cross-contamination on the Flattening blade
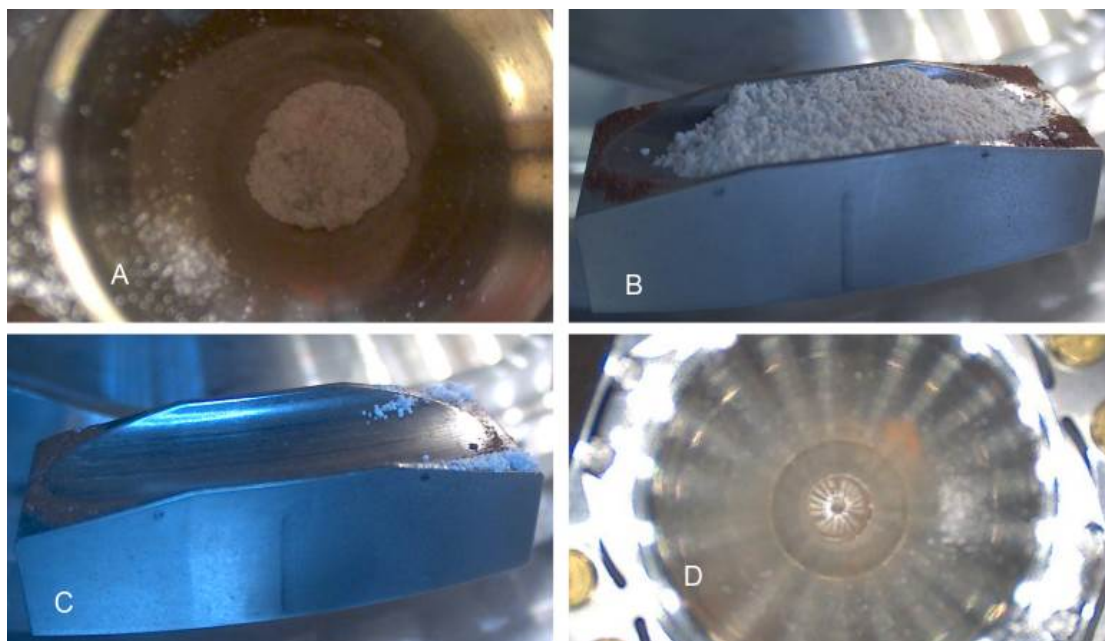
Figure 4-22. Dosing (A), flattening (B) and cleaning processes (C, D) with the gypsum S-4_WET sample. Sample remains from both the hematite and gypsum samples can be observed.

After the flattening process, 36 images were acquired with the MicrOmega emulator camera to show the complete RC (Figure 4-17). These images showed small reddish features corresponding to hematite (that can also be somewhat appreciated in Figure 4-22B). Then the RLS instrument acquired 7 points separated 0.05º at the beginning of the RC, another 7 around the nominal position of the RLS instrument (center of the RC) and another 7 in the final part of the RC. The aim of the test was to evaluate if there was some kind of pattern on the potential cross-contamination of the samples. In addition to these spectra, a visual inspection of the complete line of the RC with RLS camera was performed to look for rests of hematite on the white surface of the S-4 sample.

The results showed that, in spite of the reddish features that can be seen in the MicrOmega FoV (Figure 4-17), no cross-contamination was unequivocally found, either with Raman or visually with the RLS camera on the line (though very small traces of hematite can be seen in RLS FoV in regions which are out of reach of the spot, as shown in Figure 4-23). A test was performed on the mixed sample by automatically acquiring 25 points on the sample which also failed to detect any trace of iron oxides on the S-4_WET sample. Though some very small remains were visible with the RLS camera, they were not detected by RLS, as these were found out of the RLS spot. Thus, the cross-contamination introduced by the dosing and flattening system is an issue to be taken into account and that can't be totally ruled out, though it has very low probabilities: about 20 hematite grains of diameters lower than 10 microns were seen on the 25 RLS images. A gross calculation of the surface covered by hematite in the RLS

images with respect to the total FoV in the 25 images is 0.00026 of the total. Probably, future testing should be focused on the cross-contamination from the crushing station, as it will potentially introduce much more cross-contamination than the dosing and flattening processes.
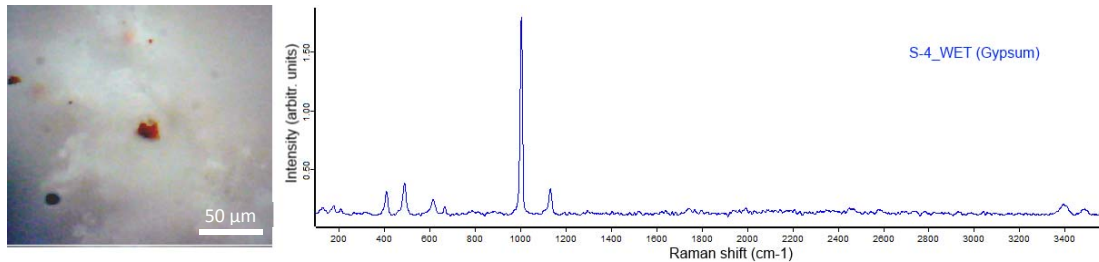
Figure 4-23. Example of undetected cross-contamination, and S-4_WET spectrum.

4.2.4.1   POSSIBLE CROSS-CONTAMINATION FOUND DURING TEST 1

To optimize the available time and resources, the tests were planned in a non-orderly fashion. After the cross-contamination tests, the sample S-7_DRY was dosed and flattened for test 4, and the rests of the sample left overnight in the dosing funnel to check for cementation. The next day, everything seemed to be working fine, without cementation on the funnel. After the analysis of this sample, the Test 1 (dosing and flattening performance) for the calcite SPDS-like and the 45-125 microns samples was performed. During the dosing of the second sample (calcite 45-125 microns), rests of other samples were observed on the flattened surface, as it can be appreciated in Figure 4-11 and Figure 4-24.

Figure 4-24. Cross-contamination on the calcite 45-125 microns simple. Flattened surface (A), rest of sample in the funnel after dosing with reddish features (B), MicrOmega emulator camera (C) and RLS images (D). The red features are iron oxides.

In fact, the Raman analysis of this spot showed that some features not corresponding to calcite can be appreciated in the Raman spectrum, as shown in Figure 4-25. These features might correspond to the iron oxides (goethite) and the sample S-7_DRY analyzed in previous tests (though not in the immediately previous one). This might be explained by some kind of cementation on the funnel when the sample was left overnight, that didn't get released until several activations of the piezoelectric actuator. The relative humidity that day was quite high (53%) due to the weather, which might have had an influence on this issue. In addition, the voltage of the piezoelectric was not set to the nominal value, which would probably have released the cemented sample during the cleaning process. Finally, we have to state that, even if this point was found along the laser spot path, it wasn't one of the prefixed positions but it was found between points 20 and 21. Obviously, this is a matter of luck that doesn't remove importance to the cross-contamination issue.
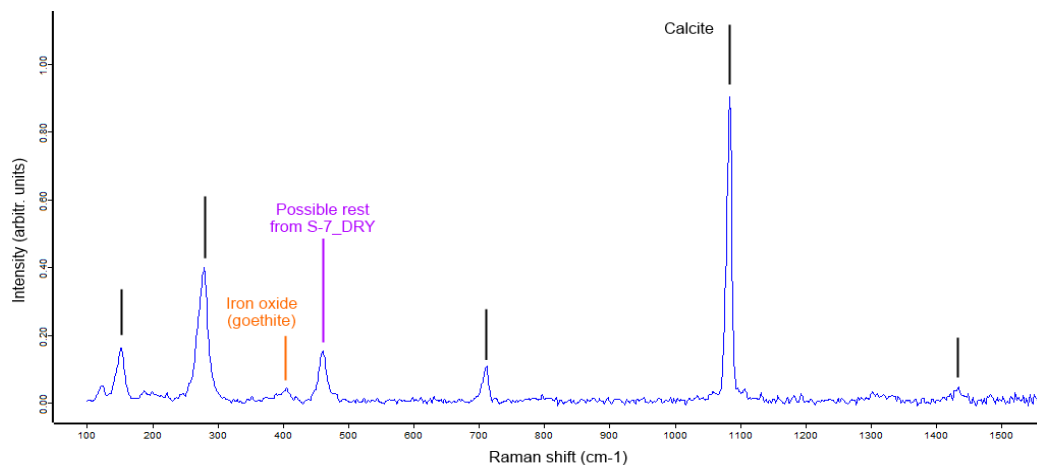
Figure 4-25. Spectrum from calcite 45-125 microns grain size with traces of other materials from previous analysis

4.2.5 TEST 3: MICROMEGA AND RLS COMBINED SCIENCE

This test was aimed at evaluating the possibility of performing combined science between MicrOmega and RLS. This was intended to be performed with a mixed sample of quartz+olivine. However, the imaging contrast of the RLS and MicrOmega emulator cameras provided images of the samples in which it was very difficult to identify and distinguish the olivine grains from the quartz (probably due to the contrast with the LED illumination system implemented in the RLS). This way, it was complicated to find a point to be selected in the MicrOmega emulator FoV that showed differentiable features with respect to its environment. What seemed to be an olivine grain in the MicrOmega emulator FoV didn't provide olivine spectral features in the RLS spectra (and the camera didn't show differences among grains).

It was also tried to find an olivine grain on the RLS path and then move it to the MicrOmega FoV. However, when moving this point to the MicrOmega emulator

camera, it was almost impossible to visually differentiate this grain in the MicrOmega FoV.

To overcome this difficulty, some milligrams of basalt were added to the already-flattened surface of quartz+olivine in order to have very differentiable materials. The sample was re-flattened again with the flattening blade after adding the basalt grains. At this point, several analysis runs were performed.

Run 1: To check the position calibration, Run1 was used to sweep the sample and ensure the correlation between the nominal positions of RLS and MicrOmega by selecting an interesting point in the RLS field of view, and selecting that point in the MicrOmega FoV after moving 22.591 degrees (the RLS-MicrOmega offset). This is depicted in Figure 4-26.
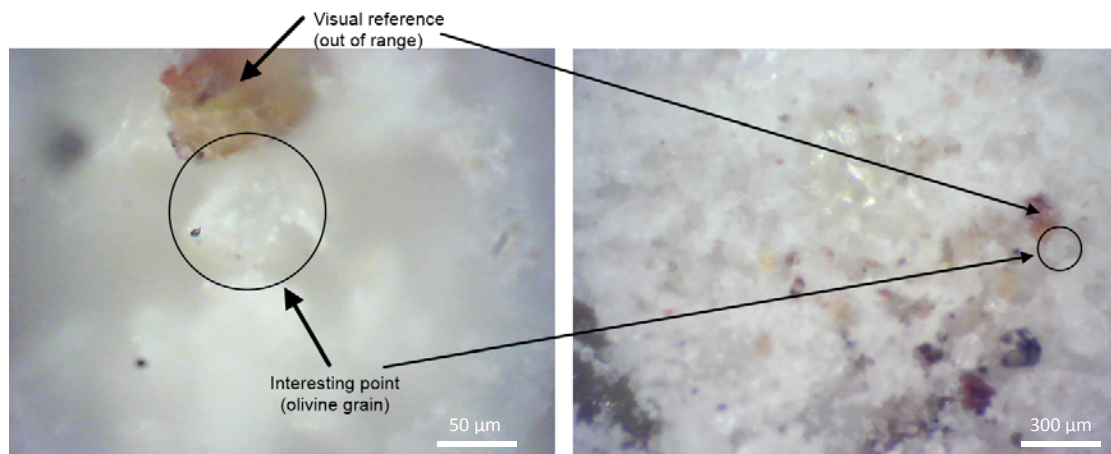
Figure 4-26. Interesting point selected in the MicrOmega emulator FoV (right), and positioned under the RLS spot (left)

Run 2: The SPDS was moved to another position with the interesting point still placed somewhere in the MicrOmega FoV. Then the angular shift to the interesting point was calculated using the procedure explained in the Calibration Test section. Finally, the sample was positioned 0.125 degrees before that point. Then, RLS acquired spectra performing a sweep in steps of 0.01º (~20 microns), to see the spectral differences along the line when the interesting point was found. This evolution can be observed in Figure 4-27. As it can be seen, the interesting point (olivine doublet) is only seen in two of the spectra (it might be almost detected in a third, but with a certain uncertainty), in a very short positioning range (0.2º, around 40 microns).
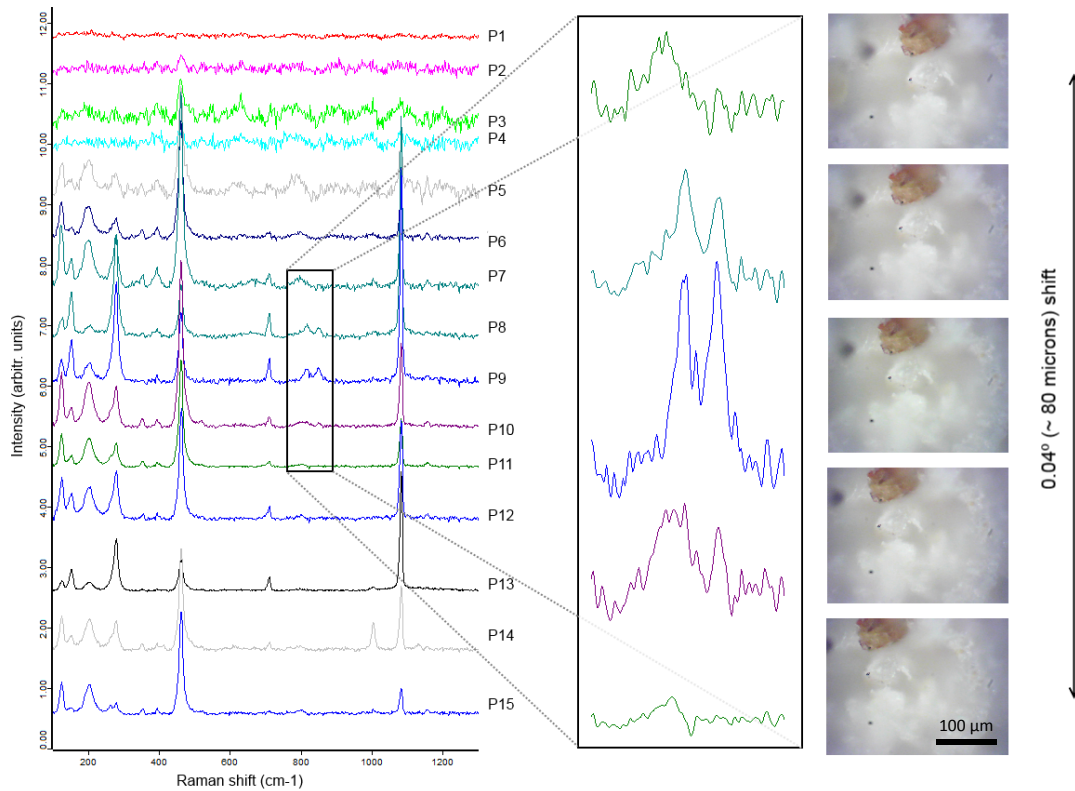
Figure 4-27. Spectral evolution when advancing the sample in steps of 0.01º (~20 microns). The olivine grain is clearly detected in only two of the spectra acquired along the sample surface.

Run 3: The SPDS was zeroed again, and then positioned again to points P6, P9 and P12. The images show a positioning error of around 30 microns, which was enough to lose some of the spectral quality. This displacement was due to the accuracy of the SPDS as well as the sample movement induced by vibrations on the setup (see section 4.2.5.1 for reference on how vibrations on the setup were transmitted to the sample).

This shows that the procedure for analyzing the interesting points selected by MicrOmega (methodology described in Run 2) is adequate to try to overcome the potential positioning inaccuracies. However, the detected positioning accuracy is within the requirement and satisfies, combined with the correct operations (of SPDS and payloads) all scientific needs. Nevertheless, it has to be noted that slight movements of the sample grains due to the settlement of the sample grains can occur and cannot be avoided. Therefore if RLS spectra are not showing the expected spectra difference along the sample scan, the sample should be brought back to MicrOmega for a second analysis to check if the selected grain moved from the previous position or not.

4.2.5.1   INFLUENCE OF VIBRATIONS ON THE RC SAMPLE MORPHOLOGY

A different kind of test was performed, aimed at trying to characterize the positioning accuracy of the positioning system. To do so, a new point was searched and selected

on the MicrOmega FoV and then the following steps, illustrated in Figure 4-28, were executed:

0- Select interesting point on MicrOmega FoV

1- Set RLS spot on the center of the new point (point PC)

2- Move to a point P1 placed at 0.08º (~150 microns distance)

3- Move to a point P2 placed at P1 position + 0.16º (~300 microns distance).

4- Go back to the original point PC

5- Go back to PC again by removing backlash (move 5º back and go forward to PC)

6- Repeat steps 4 and 5 twice (just for visual positioning) – (Steps 4b, 5b, 4c, 5c)

7- Go to MicrOmega backwards

8- Go to MicrOmega removing backlash

9- Go to PC at very low speed of the carrousel to avoid to the maximum the positioning inaccuracies. At this point, some displacement of the sample is noticed in the vertical direction (probably due to vibrations during the autofocus, as explained below).

10- Zeroing of the system, go back to PC at low speed (1º/s) and acquire a spectrum. The reshape of the sample surface was enough to lose the olivine signal.

11- Zeroing of the system, go to MicrOmega nominal position. The sample had moved.

12- Activate piezoelectric actuator to observe movement on the RLS images. Manually position the interesting point on the expected RLS position to show the vertical shift of the point.

13- Go to MicrOmega removing backlash and take a new picture to show the sample morphological repositioning.

The results of the positioning of steps 4 and 5 allowed assessing the following two conclusions:

- The positioning mechanism had a backlash of 100 μm
- The repeatability was very high (+/- 5μm when coming from the same direction)

The high absolute positioning error can be compensated by an offset correction at software level in the frame of the mechanism calibration due to its high repeatability. Please note that the mechanism was not zeroed in this sequence. The zeroing introduces an additional inaccuracy to the absolute positioning performance of the positioning system. However, this should not be required during combined science operation.
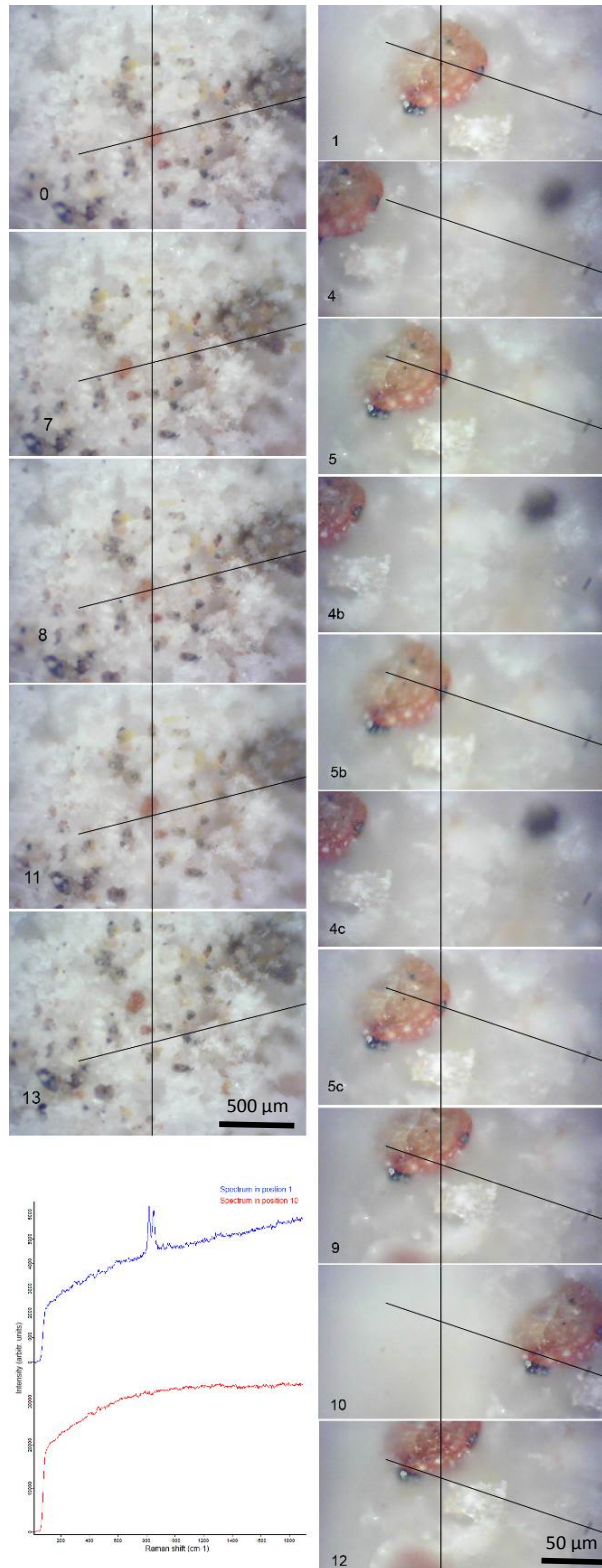
Figure 4-28. Illustration of the positioning during Run 4 (described above). MicrOmega (left) and RLS (right) images. This figure shows the position and trajectory of the sample, as well as the displacement of the interesting point from the RLS spot due to vibrations of the sample. The spectra in the bottom left part show how the olivine grain is not detected at the end of Run 4.

After the movement of the sample, we noticed that vibrations on the system (induced by the RLS autofocus and the SPDS piezoelectric actuator) can induce movements in the sample (see Figure 4-28-11 vs 13, and 5c vs 9 and 12), causing the miss-positioning of the sample under the RLS laser spot. This miss-positioning is enough to place the interesting point out of the laser spot, implying the loss of the spectral features of the interesting point (olivine), as shown in the spectra depicted in Figure 4-28. To further illustrate this issue, and to clarify that it is the sample morphology which is affected, and not the setup hardware, Figure 4-29 shows how the sample surface is reshaped during the vibrations induced by the piezoelectric actuator. These pictures were acquired by the SPDS cameras during the dosing procedure. Of course, a morphological modification is not important during the dosing procedure, and these pictures are only included to provide an example. These effects can be avoided by defining an adequate operational sequence: the positioning system provides the possibility to perform all required science on the RC (repeated analysis by RLS, MicrOmega and MOMA LD-MS) without the need to pass again underneath the flattening blade and without the need of additional dosing. The motion does not induce vibrations that cause any form of alteration of the sample surface (no sample surface alteration detectable). However, this is something that will have to be taken into account during operation, in case any moving part (solar wings?) may cause the reshaping of the sample surface during the collaborative operation of RLS and MicrOmega.

It is also important to note that the methodology for the analysis of potentially interesting points detected by MicrOmega (sweeping the sample in small steps to ensure the coverage of the interesting point) seems to be the correct procedure in order to limit the impact of the potential misalignments.

Finally, if the reshaping of the sample surface with the piezoelectric vibration was characterized and somehow repeatable, it could be theoretically possible to try to place an interesting grain under the RLS instrument range, or else, to move any non-interesting point out of range. The piezo effect on the sample surface will probably be dependent on the sample, making this characterization very difficult. In addition, there would be operational consequences to be taken into account. However, this is something that might be worth studying.
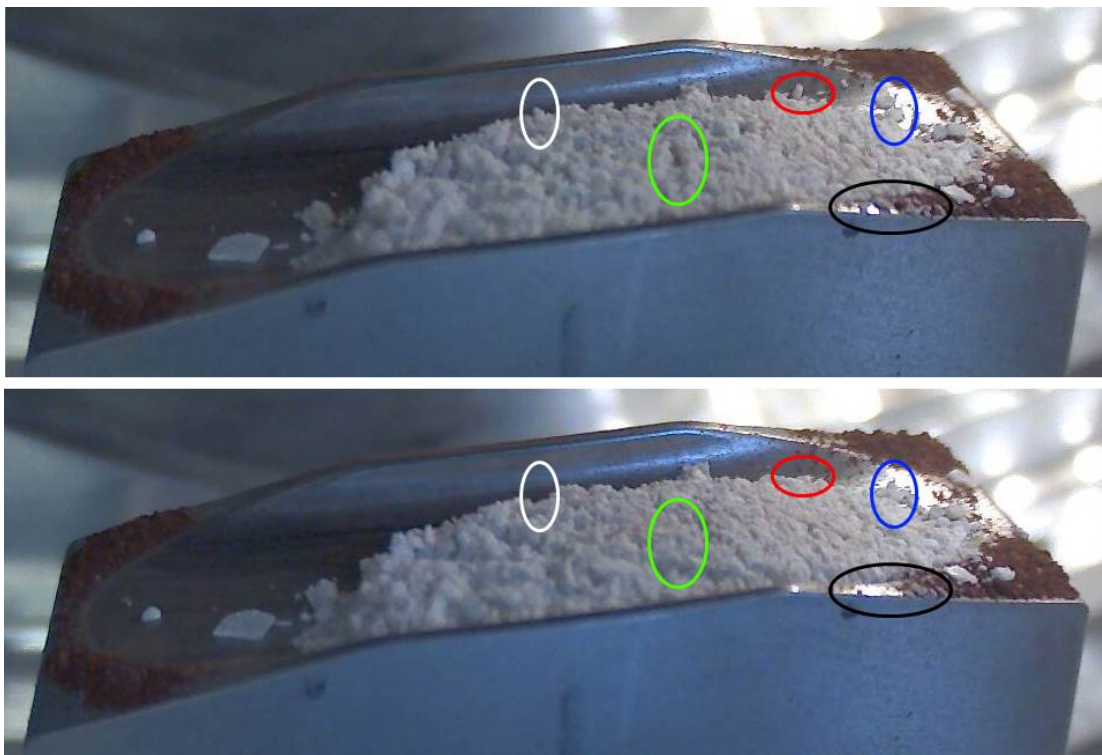
Figure 4-29. Morphological modifications after the vibrations induced by the piezoelectric actuator

4.2.6 TEST 4: ANALYSIS OF REPRESENTATIVE SAMPLES

This test was aimed at performing some experiment cycles on samples considered to be representative by the RLS scientists and principal investigator. The purpose of the test was to evaluate the results on the sample surface provided by the actual SPDS system. The analysis was performed on several points on each sample, to obtain a number of spectra which are enough to characterize the samples. However, due to time constraints, it was decided that the number of acquisitions was to be adjusted during the test to cover as much samples as possible in the available time.

Three samples were analyzed during this test: powdered basalt from Almagro (Spain), Aarhus mixture and S7_DRY.

The basalt from Almagro (Spain) is a relevant sample in the Martian exploration framework, as it is a rock (several conglomerated minerals) formed by volcanic activity. The sample was prepared with a SPDS-like grain size distribution. The 10-point analysis detected several mineral phases including pyroxenes, olivine and even carbonates (calcite), in addition to possible traces of iron oxides (hematite, goethite), as shown in Figure 4-30. Though the analysis with 10 points provided a pretty good characterization of the sample, a complete experiment cycle (20 points at least) is always preferred in order to provide data enough to perform not only qualitative but also quantitative information.
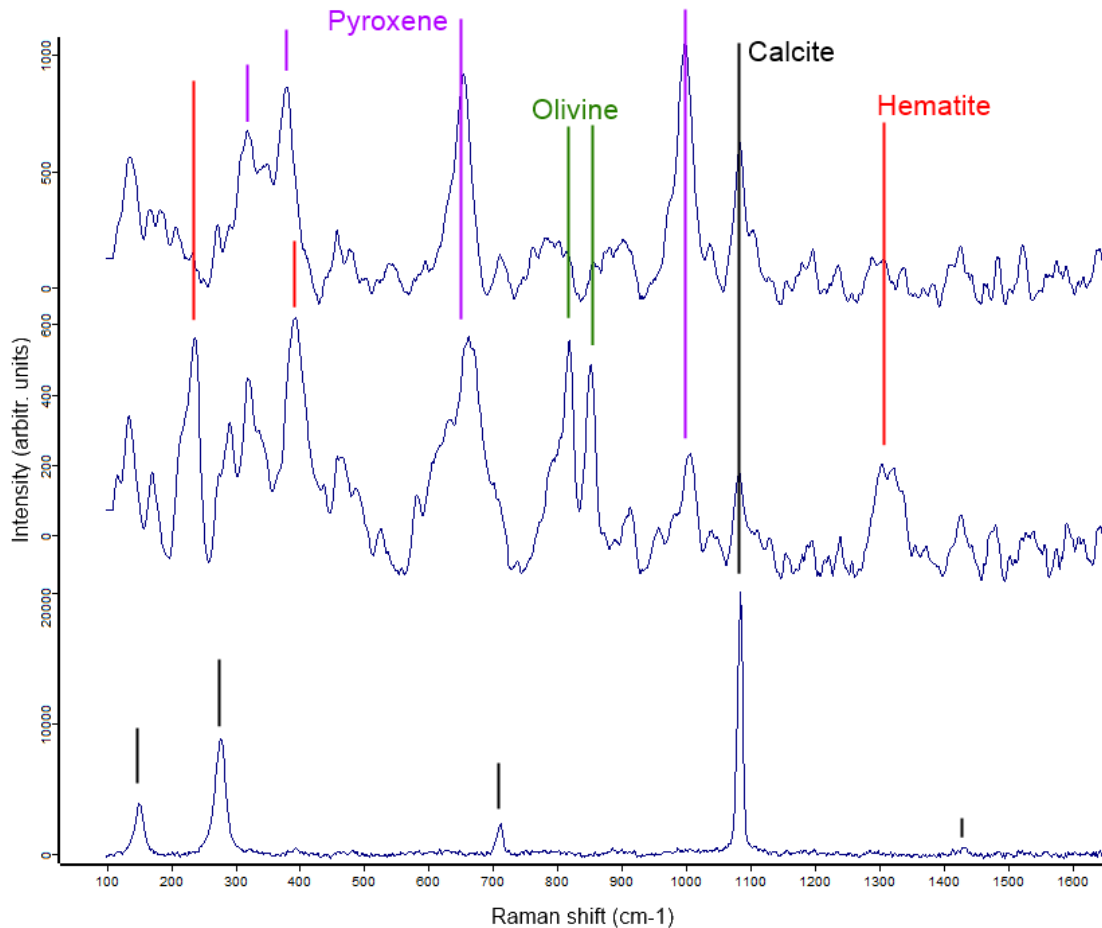
Figure 4-30. Spectra and intrepretation from the basalt sample from Almagro (Spain)

Aarhus mixture is a processed sample that was mixed from the reference materials used by the SPDS team in previous SPDS tests. The Raman analysis showed that this sample was very fluorescent, but the fluorescence removal algorithm wasn't used in order to save time. Still, RLS managed to detect several mineral phases present on the mixture: silicates (quartz) and carbonates (calcite).

The analysis of the S-7_DRY sample Figure 4-31 provided very clear signatures of epsomite ($MgSO_4 \cdot 7H_2O$) with traces of hexahydrite ($MgSO_4 \cdot 6H_2O$) which is not unusually found in epsomite samples. This material was probably formed after the dehydration of the epsomite. No trace of montmorillonite nor perchlorate were found in this sample with the 10 point analysis.
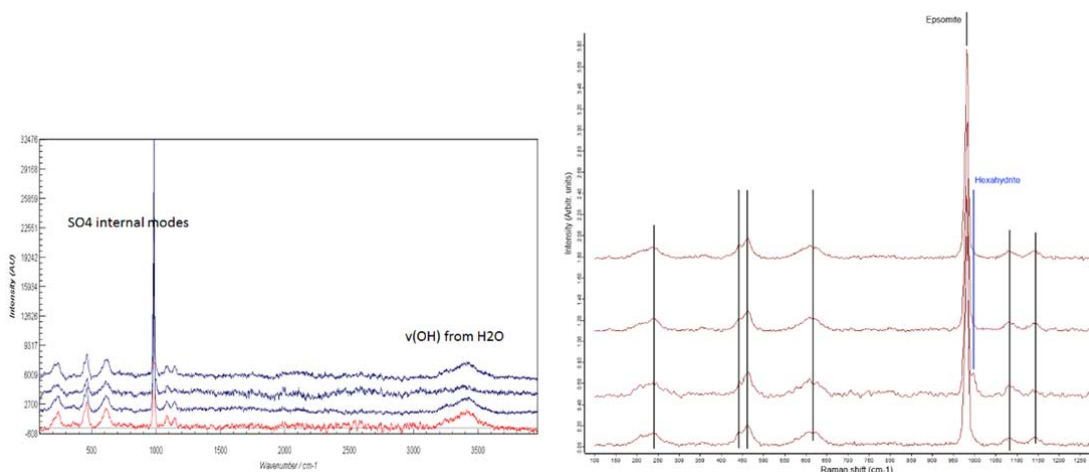
Figure 4-31. Spectra and interpretation from the S7-DRY sample. Vibration modes (left) and materials (right)

4.2.7 CONCLUSIONS

The SPDS-RLS test campaign showed that the SPDS breadboard met the scientific needs regarding dosing, positioning, flattening and cleaning performance, though some cross-contamination issues appeared. The SPDS also showed that it can process non-reference materials achieving very good results.

The flattening and dosing tests showed a good performance in terms of the requirements when assuming the sample reference plane 200μm beneath the container rim. Even samples considered highly critical by RLS regarding flattening performance could be prepared by the SPDS as required by the instrument. The standard deviation and peak to peak distances calculated by the RLS AF and the SPDS 2D surface profiler present some kind of correlation. The existing differences are probably explainable by the different number of points used for the measurement.

The cross contamination tests help conclude that the dosing, flattening and cleaning procedures will not be important sources of cross-contamination unless some kind of cementing occurs in the dosing funnel of the SPDS, which shouldn't occur in terrestrial atmosphere when the piezoelectric voltage is set to nominal value. Though the potential cementing detected during this test was probably influenced by the high relative humidity and the piezoelectric voltage setting, since cross-contamination under Martian conditions is expected to increase, it might be interesting to further test this issue under Martian atmosphere and pressure. More importantly, it will be necessary to study the cross-contamination factor induced by the crushing process, as it will certainly be more problematic than the dosing and flattening system.

The collaborative science test showed that the methodological procedure defined for the analysis of potential interesting points by RLS previously detected by MicrOmega (performing a sweep to ensure that the interesting grain is analyzed) is adequate. This

procedure not only can compensate any potential inaccuracy in the positioning by the SPDS, but also reduces the risk of losing the interesting grain due to vibrations and reshaping of the sample surface (the vibrations produced by the breadboard RLS autofocus system and the piezoelectric actuator of the dosing station induced morphological variations on the sample surface). Any alteration of the sample surface caused by the SPDS can be avoided by operations and will not occur during nominal operation. In addition, the RLS AF in-flight system will not induce vibrations, as it is based on optical reconfigurations rather than the movement of the complete optical head. However, the sample surface reshaping should be taken into consideration to ensure no vibration is induced in the ALD during the RLS-MicrOmega collaborative operation (solar panels?).

Finally, the tests performed on representative samples provided a good impression of the joint operation between SPDS and RLS, though performed with limited time, low number of spectra and non-optimal acquisition conditions.

4.3 POWDERED SAMPLES ANALYSIS: CONSEQUENCES ON THE ANALYTICAL RESULTS

The engineering constraints onboard the ExoMars Rover imply that all the instruments in the analytical laboratory drawer of the Rover will analyze powdered samples. The physical modifications suffered by the core sample during the crushing and distribution process might have an impact on the resulting Raman spectra. Moreover, context information, such as the relationship between the different minerals, is necessary to clearly identifying a rock. The association of certain minerals and textures with carbonaceous matter may also indicate the presence of biosignatures.

In this section, we present the results of the analysis of Raman spectra of powdered samples with different granulometries, with respect to the analysis on bulk samples, summarizing the results from a paper published in the Journal of Raman Spectroscopy, *Effect of grain size distribution on Raman analyses and the consequences for in situ planetary missions*, by Foucher, Lopez-Reyes et al. [54], included in this section for completion. The aim of this study was to quantify the loss of information induced by the crushing process. To do so, several minerals and rocks were analyzed using laboratory equipment and the RLS ExoMars Simulator.

This work was inspired by the experimental observation that, when analyzing powdered samples of minerals, there was a general increase of the background level with respect to the spectra in bulk samples. An example of this issue can be seen in Figure 4-32 and Figure 4-33, for natural samples of calcite and quartz, respectively, with different grain size distributions. It was also observed that, for other materials such as carbon, a decrease in the peak intensity can be observed with lower grain size distributions, as shown in Figure 4-34. In addition, the same analysis on silicon samples showed that the main peak of this material shifts and broadens with lower grain sizes (Figure 4-35).

These results showed that, in addition to the loss of texture due to the crushing process, which is detrimental for the detailed analysis of rocks and traces of life, this process can also modify the Raman signal of some samples in terms of intensity and even spectral position, in such a way that mineral identification can be hampered.

151
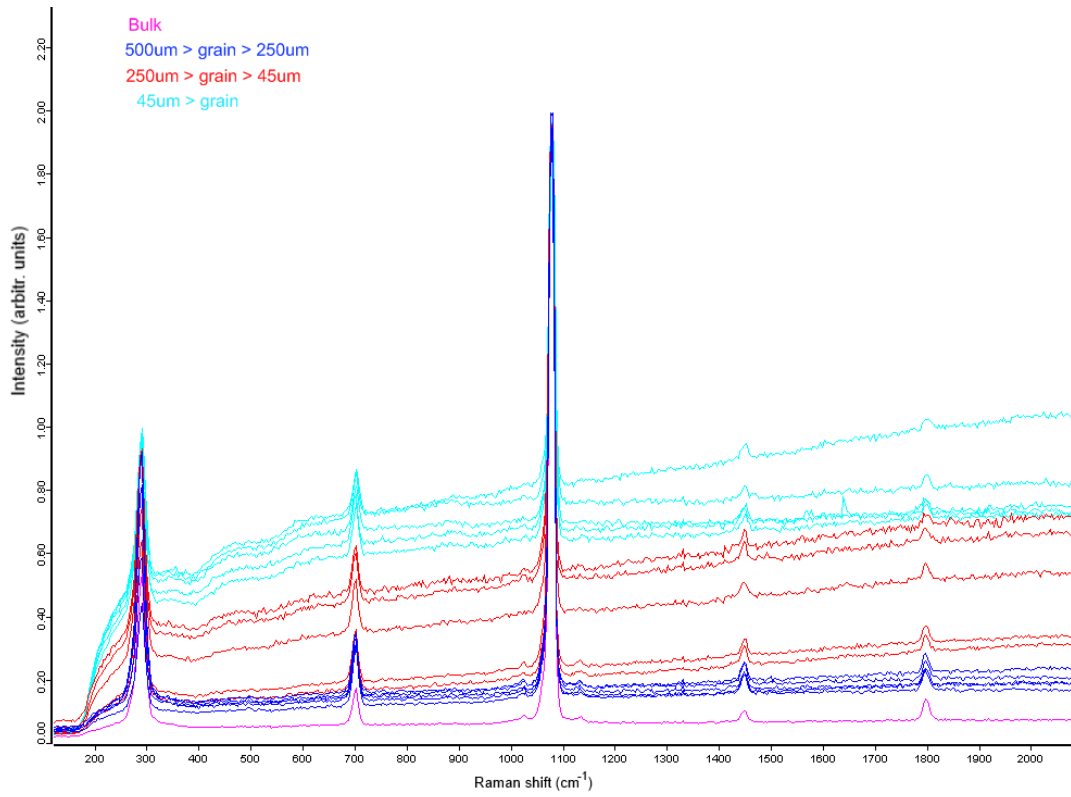
Figure 4-32. Calcite spectra with different grain size distributions.
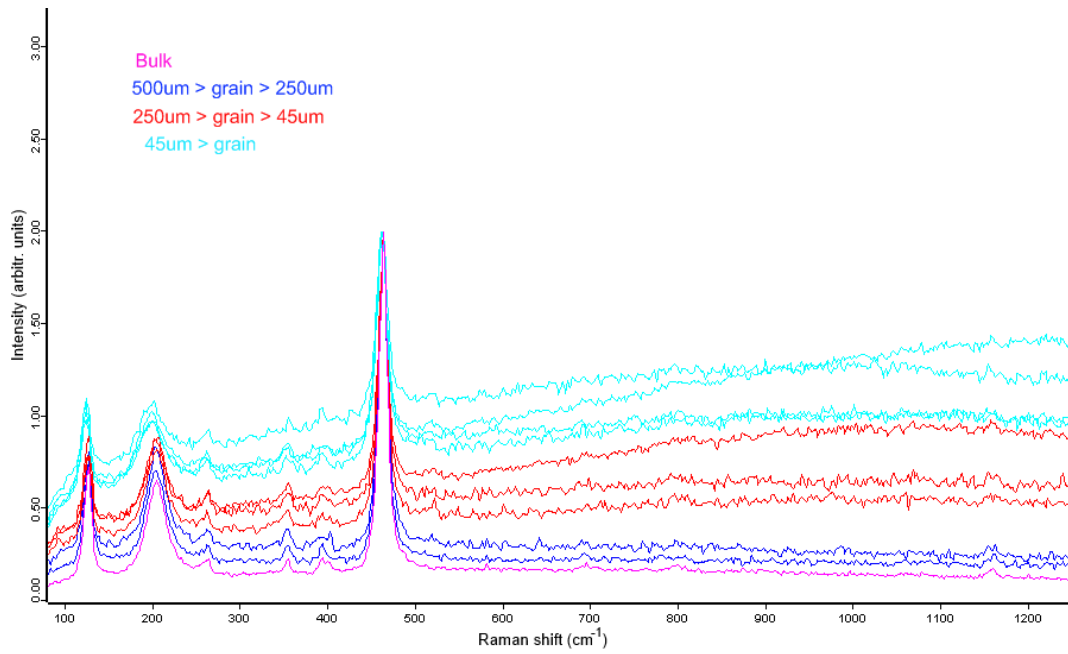
Figure 4-33. Quartz spectra with different grain size distributions.
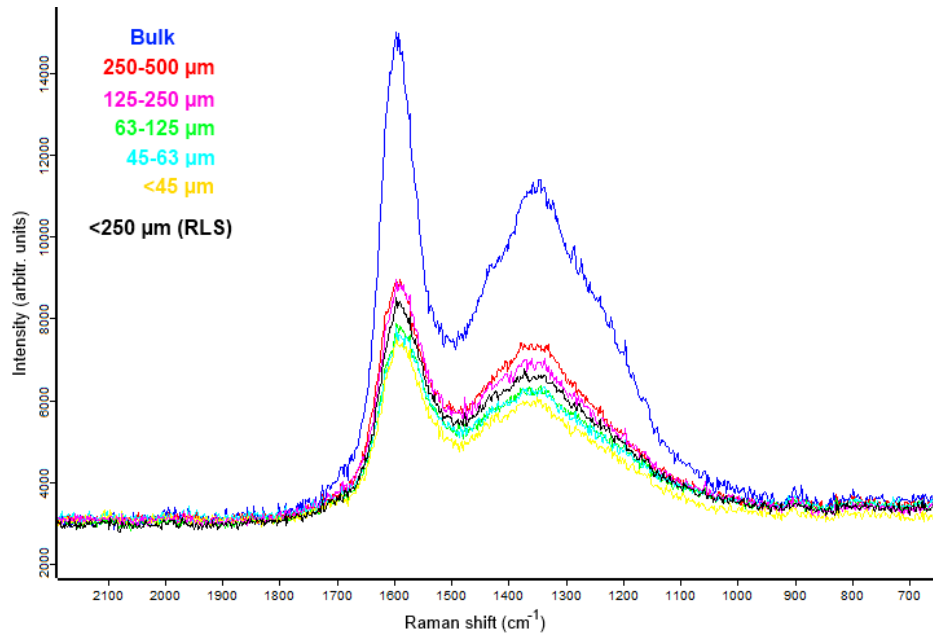
Figure 4-34. Peak intensity decrease for spectra of carbon with different granulometries.
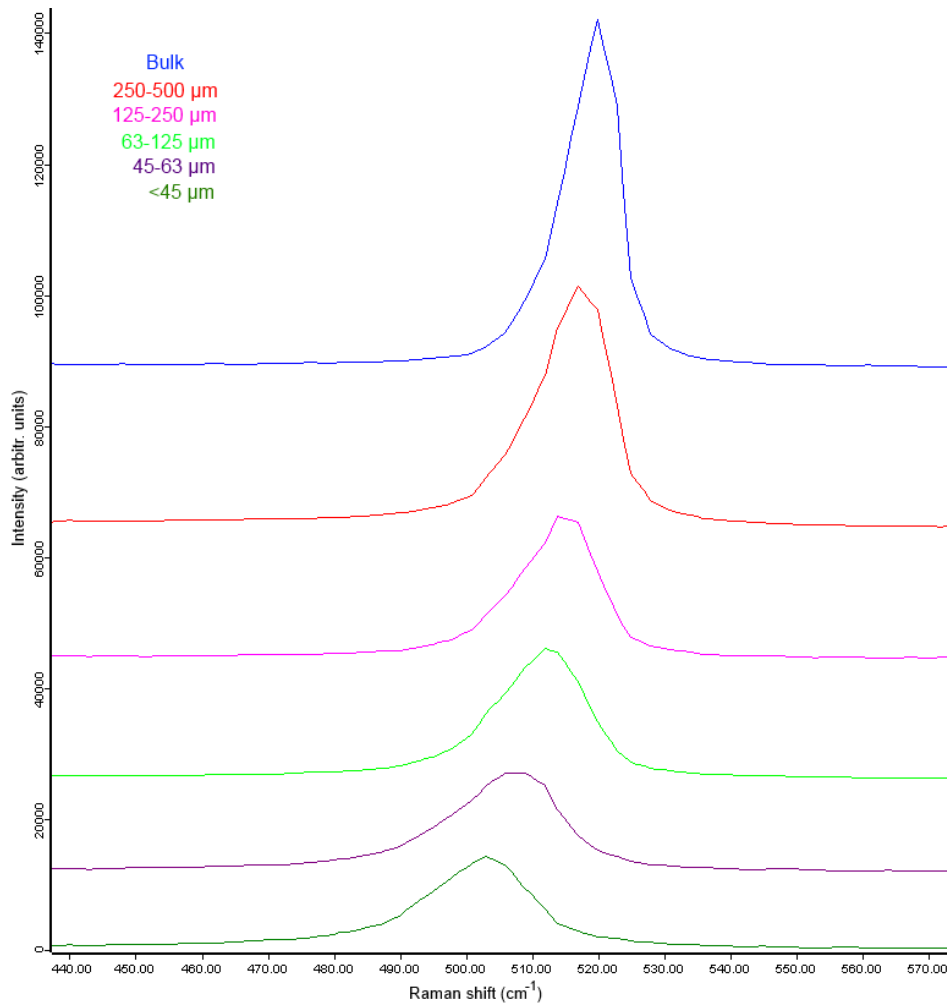
Figure 4-35. Peak shift in silicon samples.

The worst consequences of the powdering issue were observed for the silicon samples. Thus, a stokes-antistokes analysis was performed on silicon in order to evaluate the local temperature reached by the samples. The RLS ExoMars simulator was modified as shown in Figure 4-36, to remove the low pass filter of the Simulator Raman probe and substitute it with a notch filter that allowed observing the antistokes signal. In addition, a high-resolution spectrometer (Horiba iHR320) was used with access to the antistokes signal region, in order to acquire them. This setup also allowed a better quantification of the peak shift.
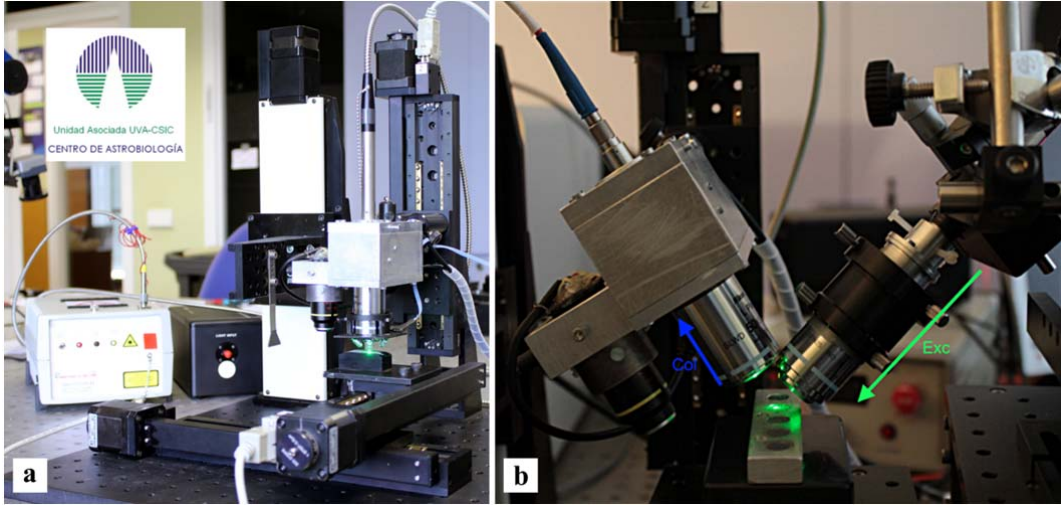
Figure 4-36. Experimental setup: RLS ExoMars Simulator (a) and modification for the Stokes-antistokes analysis (b). This figure is the supplementary material S1 from [54].

Different silicon samples with different granulometries were analyzed with this setup showing the increase in antistokes signal intensity for lower grain sizes (Figure 4-37). The local temperature of the sample depends on the relative intensity between the stokes and antistokes as per equation $T = \frac{hc\nu}{k\left[ln\frac{I_S}{I_{AS}}+4ln\left(\frac{1+\lambda\nu}{1-\lambda\nu}\right)\right]}$, where T is the local temperature of the sample, h is the Planck constant, c the speed of light, k the Boltzmann constant, ν the peak wavenumber, λ the laser wavelength, I_S the stokes and I_{AS} the antistokes peak intensities. Figure 4-37 shows how, for lower grain sizes, the local temperature increases (higher relative antistokes intensity). Furthermore, Figure 4-38 represents the Stokes/Antistokes ratio and the corresponding local temperature value of the sample.
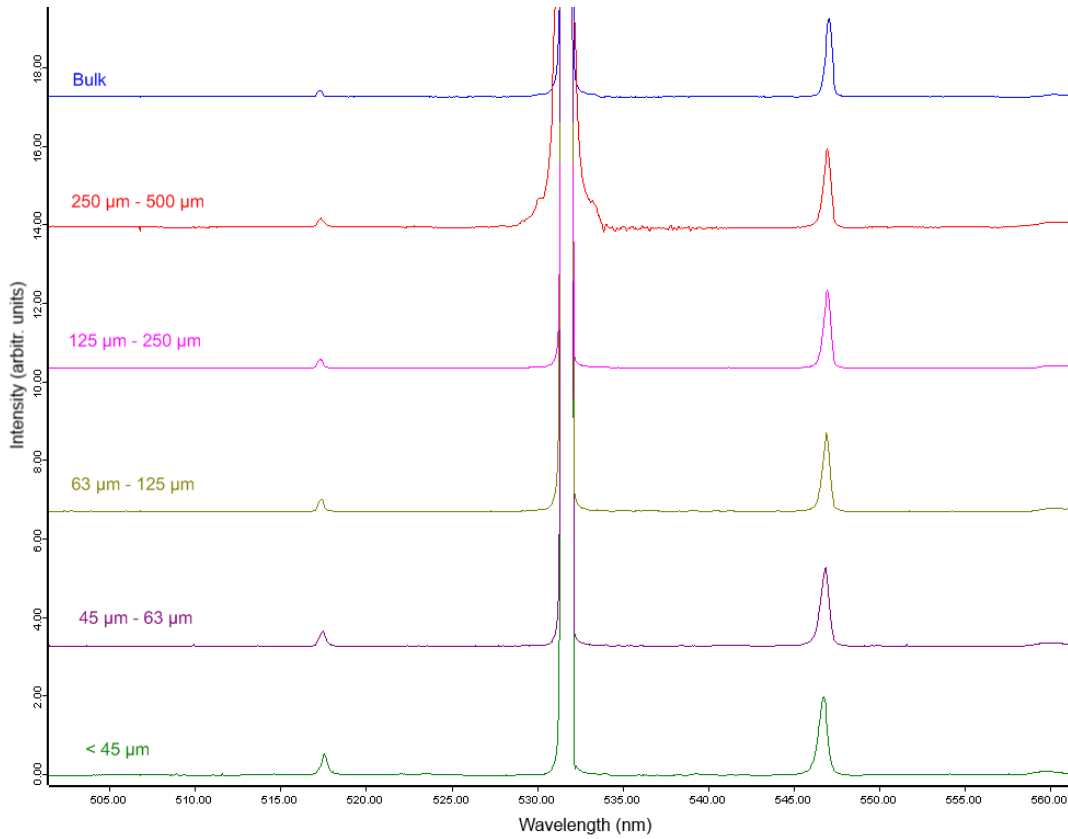
Figure 4-37. Stokes-antistokes evolution for different granulometries of silicon. The antistokes signal increases for lower grain sizes, indicating an increase of the local temperature.
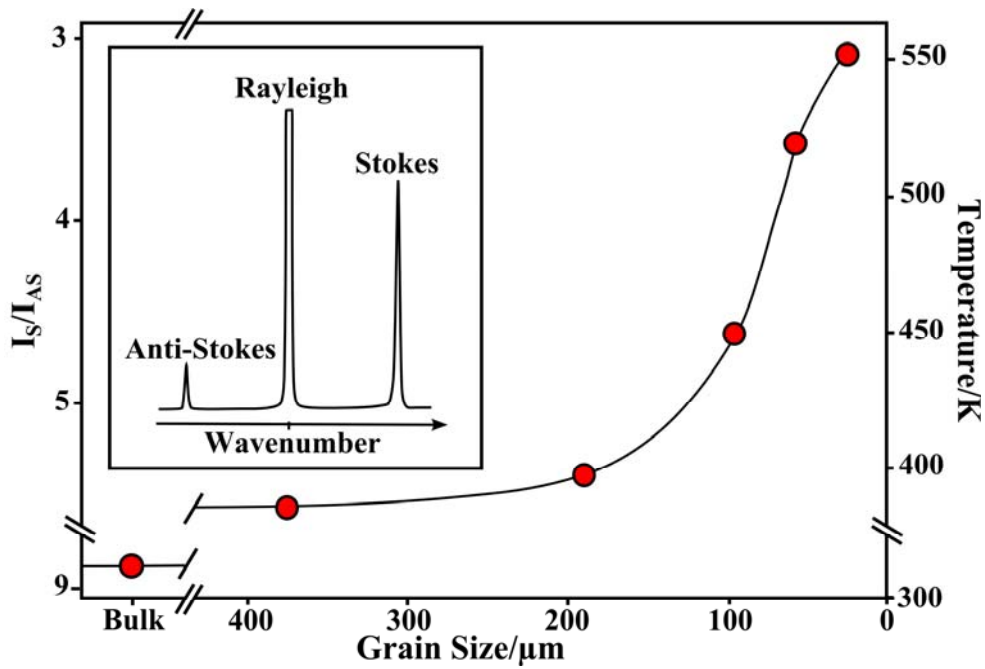
Figure 4-38. Local temperature with respect to the grain size distribution in microns obtained from the Stokes / Antistokes intensity ratio. Extracted from Figure 8 in [54].

Guillermo Lopez-Reyes

From the previous results, it seemed that the observed effects on the silicon samples related to thermal effects. In order to study the impact in a more representative scenario for ExoMars, tests were performed on bulk and powder silicon with the granulometry expected from the ExoMars crusher (powder of 250 microns), and the representative RLS irradiance level. The results showed that the shift of the main peak of silicon was 0.09 nm, which is equivalent to around 3 cm^{-1}, as shown in Figure 4-39. The widening of the peaks was between 0.04 and 0.012 nm (values lower than 4 cm^{-1}).

Given that the RLS instrument minimum resolution is 6 cm^{-1}, this result is encouraging, as the effects of the powdering wouldn't be noticeable under ExoMars / RLS constraints. However, depending on the abundance of the lowest grain sizes, all these effects could be more noticeable, so the grain size distribution provided by the crusher to the instrument is of utmost importance and should be closely studied.
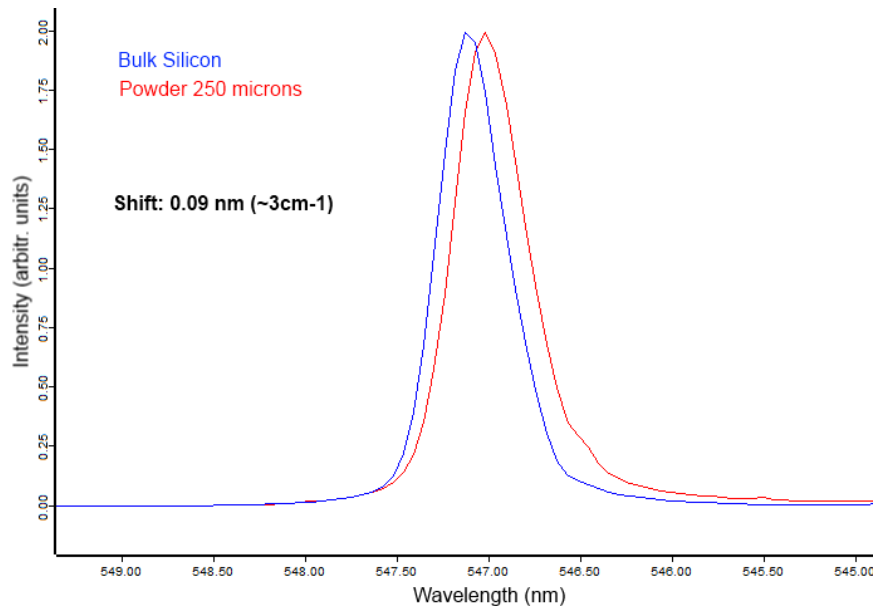
Figure 4-39. Peak shift of silicon at RLS irradiance level

The main conclusion was that most of the observed effects were due to very high local temperatures of the samples. The analysis with the RLS instrument showed that the RLS characteristics (the low irradiance and resolution requirements) minimized most of the undesirable effects found under laboratory conditions, with the exception of the background increase. However, the presence of low size grains on the mixture can potentially hamper the acquired spectra.

Regarding the loss of textural context of the samples, which is unavoidable due to the crushing process, can have its positive counterpart in that the analysis of powdered samples seemed to facilitate the detection of minor phases by random measurement, as shown in the JRS paper attached below [54], as well as in the EJM paper attached as part of Chapter 5 [56], where the detection issue is more deeply studied.

156

Research article

Effect of grain size distribution on Raman analyses and the consequences for *in situ* planetary missions

F. Foucher,[a]* G. Lopez-Reyes,[b] N. Bost,[a,c,d,e] F. Rull-Perez,[b] P. Rüßmann[a,f] and F. Westall[a]

Raman spectroscopy can be used for analysing both mineral and organic phases, thus allowing characterisation of the microbial-scale geological context as well as the search for possible traces of life. This method is therefore very useful for *in situ* planetary exploration missions. Compared with the myriad of sample preparation techniques available in terrestrial laboratories, the possibilities for sample preparation during *in situ* missions on other planetary bodies are extremely limited and are generally restricted to abrasion of rock surfaces or crushing of the target samples. Whereas certain techniques need samples to be prepared in powder form, such as X-ray diffraction, this kind of preparation is not particularly suitable for optical microscopy and/or Raman spectroscopy. In this contribution, we examine the effects of powdering rock and mineral samples on optical observations and Raman analyses. We used a commercial Raman spectrometer, as well as a Raman laser spectrometer that simulates the instrument being developed for the future ExoMars 2018 mission. The commercial Raman spectrometer documents significant modifications to the spectra of the powdered samples, including broadening of the peaks and shifts in their position, as well as the appearance of new peaks. These effects are caused by localised heating of the sample under the laser beam and amplification of nominal surface effects due to the increase in surface area in finer grain sizes. However, most changes observed in the Raman spectra using the Raman laser spectrometer system are negligible because the relatively large (50 μm diameter) laser spot size produces lower irradiance. Furthermore, minor phases were more easily detectable in the powdered samples. Most importantly, however, this sample preparation method results in the loss of the textural features and context, making identification of potential fossilized microbial remains more problematic. Copyright © 2013 John Wiley & Sons, Ltd.

Supporting information may be found in the online version of this article.

Keywords: *in situ* Raman spectroscopy; powder; Mars; minerals; rock

Introduction

Raman spectroscopy can be used for analysing both inorganic (e.g. minerals) and organic molecules, making it a particularly interesting technique for *in situ* planetary science exploration. This is particularly relevant because recent and future missions are focussing on astrobiological investigations searching for complex organic molecules and traces of life. Several studies have focussed on the development and calibration of Raman spectroscopy for science space applications.[1–7] Irrespective of the rocky body explored, one of the primary goals of *in situ* missions is to understand the geological context of the landing sites, and a Raman laser spectrometer (RLS) will be included as part of the Pasteur payload on the upcoming ExoMars rover mission (ESA/Roscosmos, 2018) whose goal is to search for evidence of microbial remains in the Martian soil.[8]

Weight, size and power limitations of spaceflight instrumentation mean that the resolution and sensitivity of analyses made *in situ* are limited compared with those made in terrestrial laboratories. In the same way, only limited sample preparation can be undertaken *in situ*. To standardise the amount of material used for analysis by the mission instruments, samples are crushed into powders. Powdered preparations are the preferred method of sample preparation for techniques, such as X-ray diffraction and gas chromatograph mass spectrometry. Subsurface material

obtained from depths from 0.5 to 2 m will be taken during the ExoMars mission and will be crushed for analysis by a suite of instruments on the rover, including the RLS.[9,10] Note that some indication of the spatial context of different grains could be provided by the close-up imager instrument on board the ExoMars rover, prior to crush the drilled core material.[11]

* Correspondence to: F. Foucher, Centre de Biophysique Moléculaire, UPR CNRS 4301, 45071 Orléans Cedex 2, France.
  E-mail: frederic.foucher@cnrs-orleans.fr

a Centre de Biophysique Moléculaire, UPR CNRS 4301, 45071 Orléans Cedex 2, France

b Universidad de Valladolid, Unidad Asociada Universidad de Valladolid – CSIC – Centro de Astrobiología, Valladolid, Spain

c Université d'Orléans, ISTO, UMR 7327, 45071 Orléans, France

d CNRS/INSU, ISTO, UMR 7327, 45071 Orléans, France

e BRGM, ISTO, UMR 7327, BP 36009, 45060 Orléans, France

f Université d'Orléans, Château de La Source, Avenue du Parc Floral, BP 6749, 45067 Orléans cedex 2, France

Many rocks and targets for *in situ* planetary investigations occur naturally as granular materials with variable grain size distributions (e.g. sands, gravels and clays). Identification of the individual minerals within these samples using optical microscopy can sometimes be difficult, in particular for smaller grain sizes. Raman spectroscopy is particularly useful for such investigations as it can identify each individual mineral phase at the microscopic scale. However, to dissociate the individual spectra of each component, the grains must be larger than the analysed volume.[12,13]

Of specific interest for Martian missions is the search for life and the presence of organic matter in the rocks and soils of the red planet. Whether or not life emerged on Mars, organic molecules should be detected because carbonaceous meteorites and micrometeorites containing abiotic organic material reach the planet surface, even at the present day.[14] Raman spectroscopy is capable of detecting carbonaceous matter in geological materials, whether of biological or abiological origin. Raman spectroscopy has been used to investigate abiogenic carbon in a variety of Martian meteorites[15] and to characterize modern biological carbonaceous matter in rocks.[1,6] Some studies have been able to detect biological carbon molecules (e.g. beta-carotene and amino acids) by using various test conditions including miniaturised and remote Raman spectrometers on powdered minerals and during *in situ* measurements.[6,16–19] In these tests, organics were added to the samples prior to analysis to establish detection limits. In terrestrial materials, the presence of biogenic organic matter produces abundant fluorescence in granular rocks (e.g. clays[20]). Even in very ancient rocks (e.g. ~3.5 billion years old), degraded biological carbon in the form of kerogen can be easily detected by Raman spectroscopy. Although it is not possible to demonstrate a biological origin of the kerogen in ancient rocks, Raman mapping of the organic component in its mineralogical context can be used to aid detection of biosignatures, for example in rocks ~3.5 Ga old.[21,22]

The Raman spectra from crystalline materials may be modified by different parameters and phenomena, as shown in Fig. 1. Because the crushing process can potentially influence all of these conditions, the Raman spectrum obtained from a powdered sample may in fact be quite different from that obtained from the (original) bulk sample. For example, changes in wave number and peak width in the Raman spectra of powdered quartz and olivine powders have previously been observed.[23]
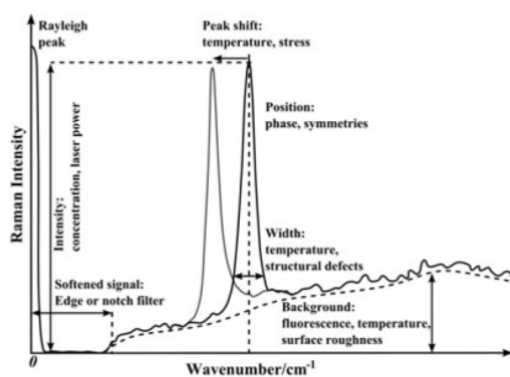
Figure 1. (Illustrative) Annotated Raman spectrum demonstrating the causality of potential modification effects.

The aim of this study is therefore to determine whether any additional Raman spectral parameters change when powdered rocks and minerals are used and whether these changes impact the geological (and astrobiological) interpretations of the data obtained. This investigation is focussed particularly on analytical conditions during future extraterrestrial *in situ* missions. We first studied the effects of crushing on the Raman spectra of different minerals by using a micro-Raman laboratory device. Raman mapping of several grain sizes was used to characterize the fine-scale changes in the spectra. The RLS simulator was used to determine if the observed grain size effects could potentially hamper correct identification of minerals during the ExoMars mission. Additionally, the consequences of crushing on the analyses of igneous rocks and sedimentary rocks (the latter containing fossilised biosignatures) were examined. Finally, the advantages and disadvantages of using crushed rocks are discussed.

Materials

Minerals

The following minerals were selected to study the changes in spectra induced by crushing.

Silicon

Silicon is more an inorganic material than a mineral because it occurs very rarely in nature[24] as small phases associated with gold.[24,25] Thus, although not a representative Martian mineral, silicon was used as a control because it has a diamond structure (O_h^7) and therefore only one first-order Raman-active Si–Si stretching of symmetry F_{2g} located at ~520 cm^{-1}.[26,27] Because its spectrum is formed by only one peak, any changes induced by the sample preparation will be easily detected. The material we used was obtained from the waste dump of the electrometallurgy plant Pechiney in Montricher-Albanne, Haute-Savoie, France.

Quartz

Quartz (SiO_2) is the standard mineral used to study the effects of crushing (i.e. sample preparation) on Raman spectra.[23] We used pure quartz (no traces of organic matter) from a hydrothermal vein associated with a gold deposit in granite from Saint-Yrieix-la-Perche, Haute-Vienne, France. The spatial symmetry of quartz (α-quartz) is D_3^4. Its spectrum consists of several bands corresponding to A_1 Si–O stretching (1085 cm^{-1}), E Si–O stretching (697, 795, 1070 and 1162 cm^{-1}), A_1 O–Si–O bending (356 cm^{-1}), E O–Si–O bending (394 and 450 cm^{-1}), A_1 Si–O stretching/O–Si–O bending (464 cm^{-1}), E O–Si–O bending/Si–O–Si bending (265 cm^{-1}), A_1 Si–O stretching/O–Si–O bending/Si–O torsion (207 cm^{-1}) and E O–Si–O bending/Si–O–Si bending/Si–O torsion (128 cm^{-1}).[28] On Earth, quartz is ubiquitous, but on Mars, it has only recently been detected in the form of amorphous silica associated with hydrothermalism and/or diagenesis.[29] This mineral phase is of particular relevance because it was the lithifying phase of ancient terrestrial volcanic sediments (now largely microcrystalline quartz or chert)[30] and the fossilising medium of the traces of life contained in the sediments .[31–34] Two samples of the ancient sediments ~3.5 Ga-old terrains in Australia (the Pilbara) and South Africa (Barberton) were included in this study and described below. The relevance of these rocks for Mars is that the silicification of the ancient sediments was induced because of the strong influence of silica-rich hydrothermal

fluids on the saturation of silica in the early Archaean seawaters.[35] This is of relevance to early Mars because the planet was as volcanically and hydrothermally active as the early Earth, and water in impact craters, or the hypothesised northern ocean, may also have been saturated in silica. The presence of silica on Mars may thus indicate ancient hydrothermal activity.[36]

Graphite and charcoal

The graphite used for this study comes from the Alibert mine in the Sayan Mountains, Russia. We used commercial charcoal as representative of biogenic carbon. The first-order Raman spectrum of graphite consists of a strong single line (G) at 1582 cm^{-1} and a smaller band at 42 cm^{-1}, which corresponds to the E_{2g2} and E_{2g1} vibrational modes, respectively.[37,38] For disordered carbon, additional bands appear at 1360 (D), 1620 (D′) and 2960 cm^{-1} (D″). Second-order Raman bands are also observed between 2400 and 3200 cm^{-1}. The ratio of intensities and the widths of the D and G peaks provides information about the degree of disorder.[39,40] The latter is very small in graphite but larger in the less well-ordered charcoal. Despite this correlation between Raman spectra and the microstructure of carbonaceous matter, ancient organic matter of biological origin in rocks (known as kerogen) is difficult to distinguish from abiogenic carbon using Raman spectroscopy.[7,41] Moreover, high-grade metamorphism tends to increase the intensity of peak G,[42] whereas abrasion can increase the intensity of peak D.[43,44]

Rocks

A number of rock samples were used to examine the effect of crushing on the detection and identification of mineral phases. In some cases, the samples included kerogenous traces of life.

Basalt

Basalt is the most common rock on Mars.[45] The sample we used was collected from Roca Neyra, Perrier, France[46,47] and is a geochemical and textural analogue of basalts analysed on Mars.[45,48] Rocks are composites of several minerals and thus do not have a specific characteristic Raman spectrum. Basalts consist of olivine, pyroxene, plagioclase, spinels and accessory minerals (e.g. calcite, apatite and zeolite) in a glassy matrix.

Silicified volcanic sediments (cherts)

As noted earlier, the oldest sedimentary rocks on Earth are silicified volcanic sediments (cherts).[30] Here, we use two early Archaean cherts that both contain traces of ancient life: the 3.45 Ga-old Kitty's Gap chert, from Pilbara, Australia, and the 3.3 Ga-old Josefsdal chert from Barberton Greenstone Belt, South Africa.[22,32–34] Volcanic sediments and their alteration products are common on Mars, although, to date, silicified volcanic sediments (cherts) have not yet been found. Nevertheless, silica of possible hydrothermal origin has been identified on Mars.[36] Given the similarities in the microbial-scale environmental conditions of early Mars and early Earth, primitive life forms on both planets (assuming that life appeared on Mars) were likely to have had similar characteristics and lifestyles.[49] Fossilised traces of pre-Noachian/Noachian (~4.5 to ~3.5 Ga) life on Mars may therefore be similar in size and shape to the very primitive forms of life that exist as silicified microfossils in the Josefsdal and Kitty's Gap cherts.[22,34] These sediments are therefore particularly relevant for Mars.

Sample preparation

Bulk samples of the rocks and minerals were analysed. Subsamples of the rocks and minerals were then crushed using an agate mortar and pestle. The crushed samples were sieved into four different grain sizes according to the Wentworth classification system[50]: 160–100, 100–63, 63–40 and <40 μm. The graphite powders tended to flocculate and were therefore washed with deflocculating agent (phosphate-buffered saline solution diluted in ~200 ml of pure water), copiously rinsed and dried. This additional step removed small grains that adhered to the larger graphite grains because of electrostatic forces, an effect not observed with the other samples. The samples were prepared for Raman analysis by placing the different grain size powders from individual samples next to each other on a glass slide and pressing the powder flat to decrease the surface roughness.

Powders of the samples in the 200–250 μm size range were also produced because this is the size of the crushed materials that will be produced by the ExoMars crusher.[51]

Methods

A WITec Alpha500 RA Raman spectrometer with a green laser ($\lambda = 532$ nm, Nd:YAG frequency doubled laser) was used (CBM, Orléans). The laser is connected to the system via a polarized optical fibre and is focussed on the surface by using optical microscope objectives. Depending on the objective used, the spot size can be relatively small, and the laser irradiance on the sample surface, estimated using a calibration grid obtained from the constructor data, can be relatively high (see Table S1). Two gratings were used for the spectrometer: 600 g/mm, with a spectral range of ~4000 cm^{-1} and a resolution of ~3 cm^{-1}, and 1800 g/mm, with a spectral range of ~1300 cm^{-1} and a resolution of ~1 cm^{-1}. The system was optimized for mapping with a piezoelectric positioning system mounted on a motorized moving stage. The samples were scanned over areas ranging from a few square micrometres to up to several square centimetres, and several thousands of spectra were acquired per scan (see Table S1). Large-area scans were made of each glass slide. Very small changes in Raman spectra that would have been impossible to detect using single spectra were highlighted on the Raman maps. In this way, it was possible to observe the variations in the spectra as a function of grain size. The average spectrum for each grain size range was obtained from these maps.

A laboratory model of the ExoMars RLS device (Fig. S1(a)), developed at the Associated Unit University of Valladolid-CSIC-Center of Astrobiology (UVa-CAB), was used to study the effect of crushing on *in situ* measurements on Mars.[9] The laboratory model has an XYZ positioning system with 2.5 μm precision and imaging, autofocus and spectral acquisition capabilities. It can be programmed to work autonomously.[52] The analytical strategy of the RLS instrument is the acquisition of spectra along a linear series of points. At each point, the optical head is focused on the sample, the acquisition parameters (i.e. integration time and number of accumulations) are automatically calculated and the spectrum acquired. The RLS instrument is equipped with a green laser ($\lambda = 532$ nm), a spot size of 50 μm diameter yielding an irradiance level between 0.6 and 1.2 kW/cm^2, with a 6–7 cm^{-1} spectral resolution. Raman spectra were also obtained at the UVa-CAB with a Horiba Jobin Yvon iHR320 spectrometer using a Spex ISA Spectrum One Nitrogen refrigerated CCD with a grating of

Guillermo Lopez-Reyes

1200 g/mm and a green laser ($\lambda = 532$ nm). This system was interfaced with the RLS simulator to acquire the Stokes and anti-Stokes signals (Fig. S1(b)). Analyses with the ExoMars RLS simulator and iHR320 spectrometer were carried out on the same sieved powders as those analysed by the WITec laboratory instrument. The parameters used for the analyses with the RLS simulator and the iHR320 systems are displayed in Tables S2 and S3, respectively.

Results

Minerals

The Raman spectra of the bulk samples are displayed in Fig. S2.

Silicon

The Raman maps made with the WITec instrument shown in Fig. 2 demonstrate that the main peak position has shifted from 521 to 493 cm^{-1} with decreasing grain size. This is associated with the observed increase in full width at half maximum (FWHM), from 6.7 to 27.7 cm^{-1}, an increase in peak asymmetry (broadening on the left of the peak), as well as an increase in the background level. Moreover, new peaks appear in the spectrum at about 280 cm^{-1}. The RLS simulator documented similar effects on the different size classes but did not observe the appearance of new peaks. Peak displacement and broadening, as well as a decrease in intensity, correlate with the reduction of grain size. For the smaller grain size, the peak shifted from 521 to 504 cm^{-1}, and the FWHM increased to 8.2 cm^{-1} with respect to the bulk sample. However, no significant changes were observed in size range of the ExoMars-type powdered sample (200–250 µm). Complementary experiments made with the

iHR320 system on the sample powder showed that the peak displacement was only about 3 cm^{-1} and that the broadening was lower than 4.5 cm^{-1} with respect to the bulk sample. Such variations are within the range of the resolution of the RLS simulator and explain why they were not observed.

Quartz

In contrast to silicon, the quartz spectrum is not as significantly modified, and the variations are small, ~1.4 cm^{-1} for the shift and ~0.8 cm^{-1} for the FWHM, as seen in Fig. 3. A peak located at 520 cm^{-1}, not detected in the bulk sample, increases in intensity with decreasing grain size. Finally, the background levels strongly increase with corresponding decrease in grain size. By using the RLS simulator, a very slight displacement of the main peak can be seen in the spectrum of smallest grain size sample with respect to the bulk sample. A very weak peak at 520 cm^{-1} also appeared. Irrespective of the grain size, an increase of the background level with respect to the bulk sample is observed.

Graphite and charcoal

The Raman spectra of graphite and charcoal powders obtained using the WITec laboratory instrument and the RLS system are shown in Fig. 4. The *G/D* intensity ratio remains unchanged, and the graphite remains clearly distinguishable from charcoal. However, the signal/noise ratio dramatically decreases with corresponding decrease in grain size as a result of a strong increase in the background level. The charcoal sample was also very sensitive to laser power and burned easily, especially at the smallest grain sizes. Finally, the presence of the *D* peak in the Raman spectrum of the bulk graphite (Figs S2(c) and 4(a and b)) can be explained by surface effects induced by the cutting process.[43,44]
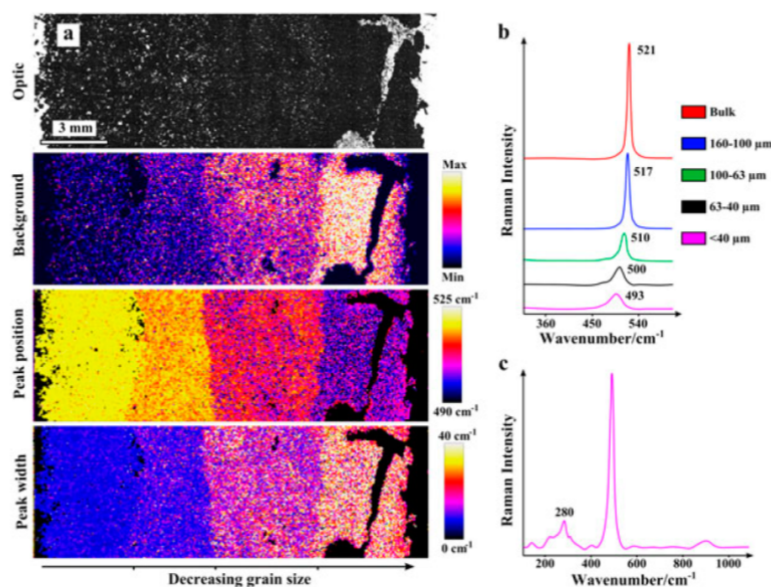
Figure 2. Raman analyses of silicon powders. (a) Photography and large-area Raman maps (decreasing in grain size from left to right) showing background level, position of the main peak and width of the main peak. (b) Average Raman spectra of the main peak for the bulk sample and for each powder. (c) Detailed spectrum of <40 µm grain size powder highlighting the bands in the 280 cm^{-1} region. This figure is available in colour online at wileyonlinelibrary.com/journal/jrs
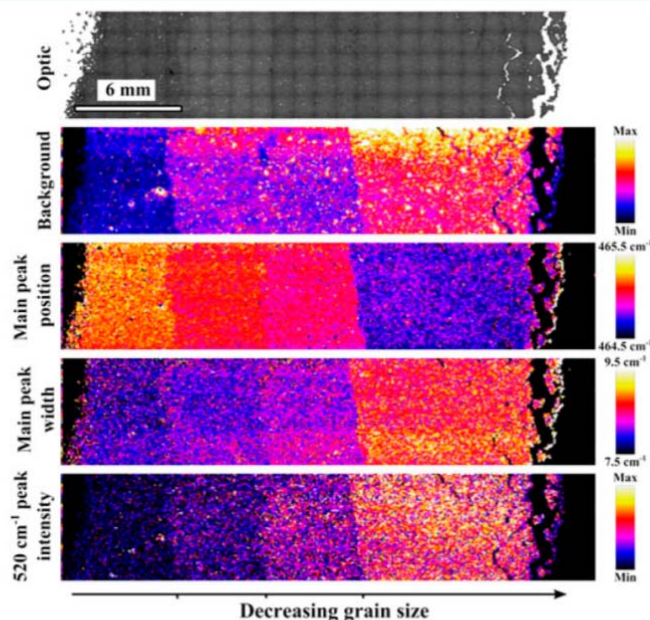
Figure 3. Raman analyses of quartz powders. (a) Photography and large-area Raman maps (decreasing in grain size from left to right) showing background level, position of the main peak, width of the main peak and intensity of the peak at 520 cm^{-1}. This figure is available in colour online at wileyonlinelibrary.com/journal/jrs
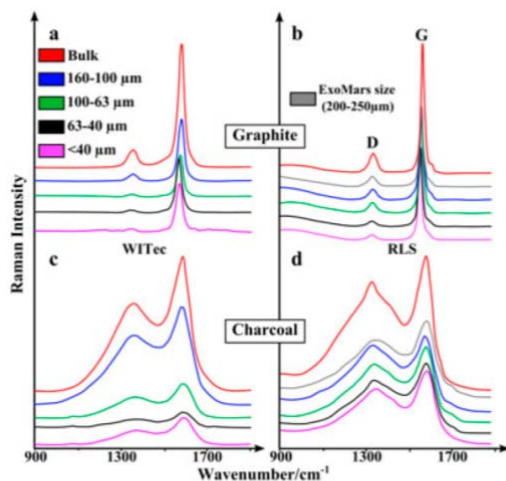
Figure 4. Raman analyses of various graphite and charcoal powders, using the WITec system (a and c, respectively) and the RLS simulator (b and d, respectively). This figure is available in colour online at wileyonlinelibrary.com/journal/jrs

Rocks

Basalt

Rocks consisting of multiple minerals require the acquisition of several Raman spectra to detect each mineral, and thus, spectral mapping is very useful for detecting all the phases within a sample. The Raman maps in Fig. 5 were made for the polished surface of the rock and powdered rock in different grain size using the WITec system. The Raman map made on the polished rock surface of the bulk sample demonstrates its fine crystalline structure and the large number of mineral phases within a small area. Minerals identified include labradorite (plagioclase), augite (pyroxene), forsterite (olivine) and apatite. Trace amounts of elemental carbon were also detected, as well as an unidentified mineral, possibly spinel.

After crushing, recognition of the different minerals becomes more difficult with decreasing grain size. The background level increases strongly, and the empty spaces between the grains diminish the number of spectra per scan. Without any prior knowledge of the studied sample, detection of the different minerals in the scanned areas would be very difficult. Identification of the minerals on the maps was made by a correlation process using the original spectra accumulated on the bulk sample. Even with this technique, not all the minerals observed in the bulk sample could be detected in the Raman maps. The crushing process also leads to the loss of texture, and the pattern of mineral distribution can no longer be used for petrological interpretation of the rock. The value of Raman mapping is significantly reduced on powdered material, and the best way to study these samples was to make a series of single spectra. A line scan of more than 100 single spectra using autofocus was thus made on the larger grain size powder. With this method, it was possible to detect the major phases, as seen Fig. 5(c). However, because of the small spot size (<1.5 μm), the probability of detecting minor phases was very low.

The RLS system uses a similar line scan procedure, but because of its larger spot size (50 μm), it was capable of detecting more phases with fewer measurements (here 30 single spectra). However, the
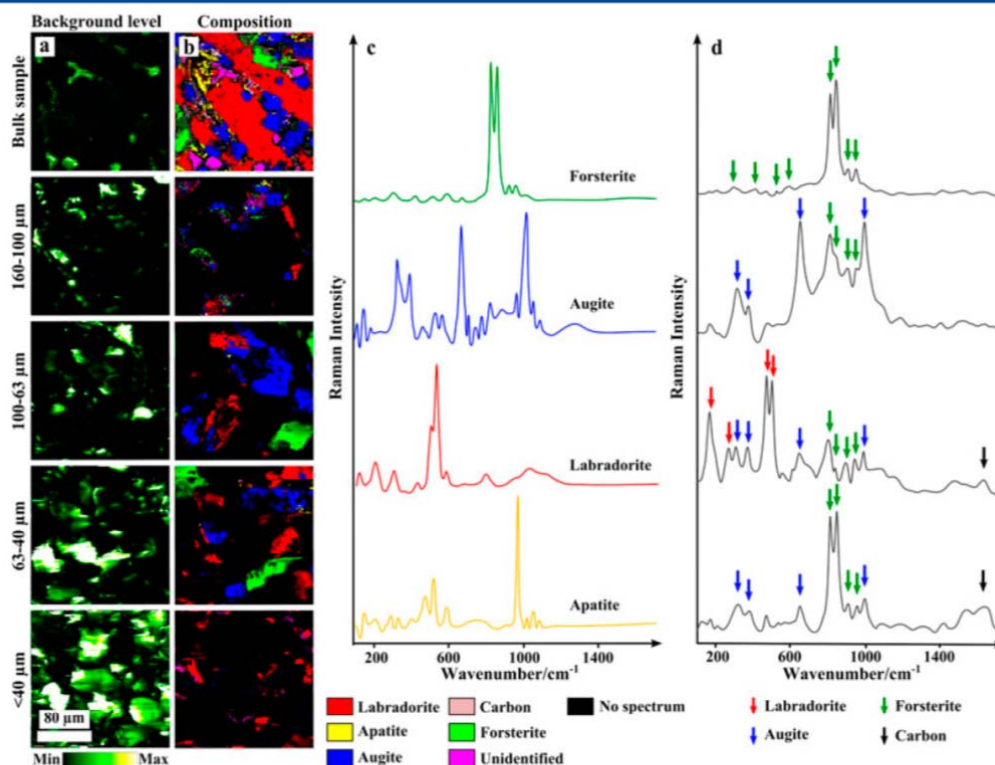
Figure 5. Background level (a) and compositional Raman maps (b) of bulk and powdered basalt samples (scan size 200 × 200 μm²) acquired using the WITec system. Single Raman spectra of 160–100 μm grain-sized basalt powder (c) collected using the line scan method with the WITec system, autofocus on. Single Raman spectra of 200–250 μm grain size (ExoMars type) basalt powder (d), collected with the RLS simulator. This figure is available in colour online at wileyonlinelibrary.com/journal/jrs

individual spectra obtained correspond to a mixture of the spectra of a number of minerals, as can be seen in Fig. 5(d).

Silicified volcanic sediments (cherts)

The silicified volcanic sediment samples were first characterized on a cut surface of the rock (similar to the surface of a rock prepared using the rock abrasion tool of the Mars Exploration Rovers) and in thin section preparations of the rock (Fig. 6). Although the rock contains microfossils, these structures are too small to be observed by optical microscopy, even in thin section. However, the black layers of the Josefsdal chert contain detectable intact layers consisting of *in situ* or detrital fragments of microbial mats and other microbial remains (Fig. 6(d and e)). Likewise, in the Kitty's Gap chert sample, carbonaceous remains of microbial colonies can be observed microscopically on the surface of volcanic grains (Fig. 6(i and j)). The relationship between the texture observed in the optical view and the distribution of the carbonaceous matter detected by Raman spectroscopy is readily apparent in Fig. 6. Although the Raman spectrum of kerogen is not sufficient to demonstrate its biogenicity, the distribution of carbon within the rock and its association with particular minerals is sufficient to, at least, locate interesting areas for further, potential biosignature investigation. Once crushed, the powders of the originally black and white laminated rocks are relatively uniform in colour. Raman mapping of the powder revealed only the presence of quartz. On the other hand, the line scan method, similar to that used for the analysis of the basalt, provided several spectra of carbonaceous matter, as well as quartz. However, the spatial relationship of the carbonaceous matter with respect to particular minerals was lost.

The experiments performed with the RLS simulator showed that this instrument was capable of easily detecting the carbonaceous matter in the sample and carbon peaks occurred in almost all the spectra acquired. In addition, all other minerals detected in the bulk rock were found. As for the basalt, this is due to the larger spot size of the RLS simulator compared with that of the WITec instrument. Figure 7 shows a typical Raman spectrum acquired with the RLS simulator compared with the average spectrum of 80 spectra obtained with the WITec system in which ankerite was also observed.

Discussion

Minerals

Heating induced during measurement with a Raman laser can be very high, in particular with the WITec system where the spot size

Effect of grain size distribution on Raman analyses
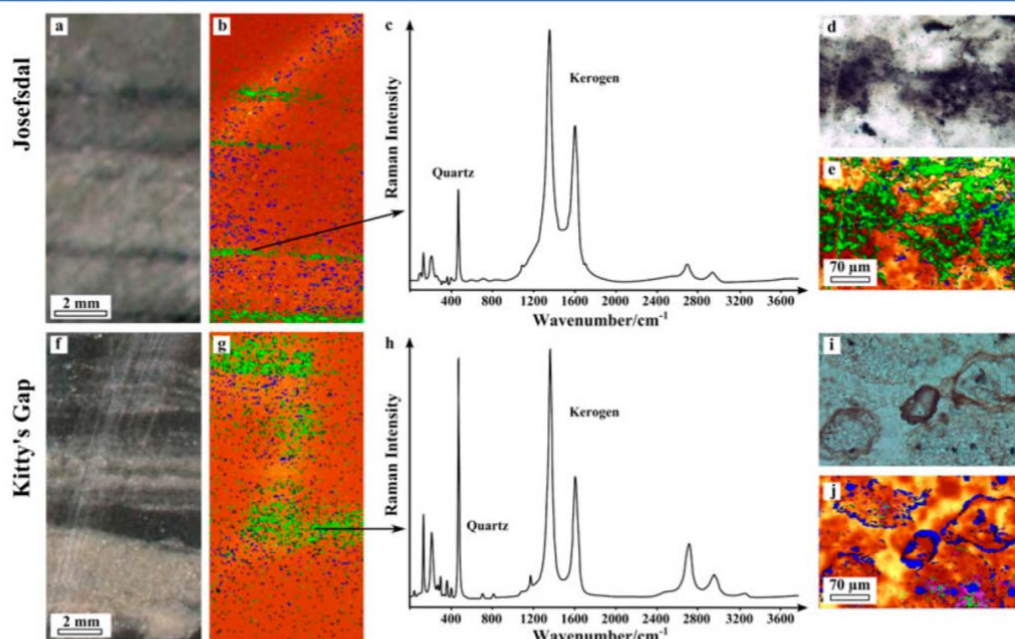
Figure 6. Chert samples from (a–e) Josefsdal, Barberton Greenstone Belt, South Africa, and (f–j) Kitty's Gap, Pilbara, Australia. (a and f) Photography of the cut and polished surfaces and (b and g) corresponding Raman maps, where carbonaceous matter (kerogen) appears green, anatase (TiO$_2$) blue and quartz orange. These Raman maps allow rapid detection of areas of interest, which can be targeted for further analyses with higher resolution spectral (c and h) and imaging (optical (d and i) and Raman (e and j)) techniques. This figure is available in colour online at wileyonlinelibrary.com/journal/jrs
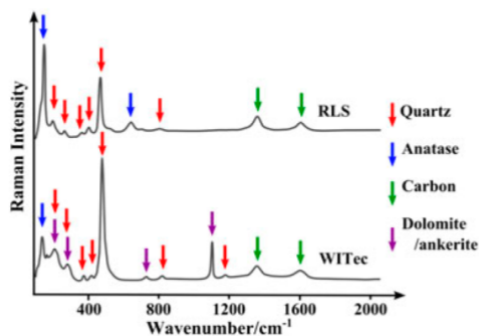
Figure 7. Raman spectra of silicified volcanic sediment powders from Josefsdal, Barberton, South Africa. (a) Typical single spectrum obtained with the RLS instrument and (b) 'equivalent' spectrum obtained by averaging 80 spectra acquired with the WITec system. This figure is available in colour online at wileyonlinelibrary.com/journal/jrs

is only about 1.5 μm in diameter. Thermal effects can thus explain some of the changes observed in the silicon and quartz spectra. The shift and broadening of the main peak of quartz[53] and silicon[54] with temperature are well known. To confirm the thermal origin of the observed changes, complementary investigations were made. The evolution of the Raman spectrum of the finest silicon powder with increasing laser power was studied using the WITec system, and a broadening and shift in the main peak

were observed (Fig. 8(a)). A series of spectra on the same powder were also acquired with the iHR320 system interfaced with the RLS system (see Fig. S2(b)). This configuration allows measurement of the intensity ratio between the Stokes and anti-Stokes signals that is known to be function of the temperature[55,56]:

$$T = \frac{hc\bar{v}}{k\left[\ln\frac{I_S}{I_{AS}} + 4\ln\left(\frac{1+\lambda\bar{v}}{1-\lambda\bar{v}}\right)\right]} \qquad (1)$$

where h is the Planck constant, c the speed of light, k the Boltzmann constant, $\bar{v}$ the peak wavenumber, λ the laser wavelength, I_S the Stokes and I_{AS} the anti-Stokes peak intensities. With this equation, the evolution of the local temperature with the different grain sizes was plotted in Fig. 8(b). The temperature reaches more than 280 °C for the smaller grain size, despite the very low irradiance used (1.2 kW/cm^2). Peak position with respect to temperature has been previously measured on bulk silicon.[54] For a temperature of 280 °C, it is ~519 cm^{-1}. This is in accordance with the iHR320 measurement of 518 cm^{-1}. By using the same model, the temperatures corresponding to ~504 cm^{-1}, as measured with the RLS system, and to ~493 cm^{-1}, as measured with the WITec system on the smallest grain size powder, are 760 and 1200 °C, respectively. The strong increase of the temperatures with decreasing grain size can be explained by the fact that small particles cannot evacuate heat as well as large particles, i.e. the thermal conductivity of the material decreases with grain size. This effect is all the more important because silicon is black. Changes in wavenumber and line width with decreasing grain size have been observed previously and were explained by the
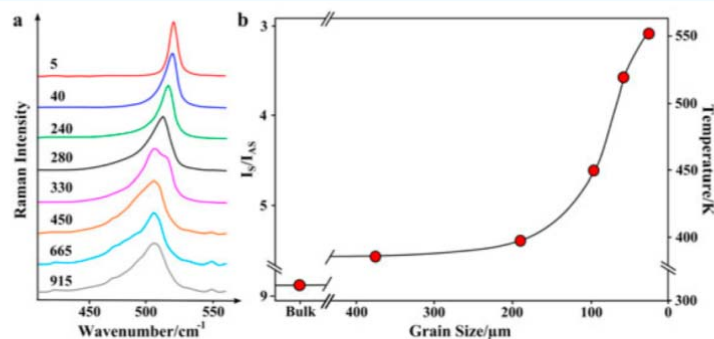
Figure 8. (a) Main peak of silicon corresponding to different laser powers (in kW/cm^2) acquired using the WITec system on the <40 μm grain size powder. (b) Stokes and anti-Stokes intensity ratio and associated local temperature with respect to the grain size distribution of the silicon sample measured using the iHR320 system interfaced with the RLS simulator. This figure is available in colour online at wileyonlinelibrary.com/journal/jrs

contribution of surface modes that are increased in finer grain-sized powder.[23] However, our experiments clearly demonstrate that the major reason for this change is the local increase of temperature. Despite its black colour, this study shows that the spectrum of carbonaceous matter was not as affected by this strong increase in temperature as suggested by previous studies.[41]

Conversely, the appearance of the new peak at 520 cm^{-1} in the quartz spectrum is not a thermal effect associated with the power of the laser. This peak is characteristic of coesite[57] and corresponds to the symmetric stretching of oxygen in the four-membered SiO$_4$ tetrahedra.[58] However, the Raman spectrum of coesite is normally characterised by additional peaks at 204, 270, 355 and 426 cm^{-1}, which were not present in our sample.[57] A similar situation was previously noted in the quartz spectrum in an eclogite from Antartica.[59] In the latter case, it was proposed that the coesite crystalline structure was attained at the nanometric scale within anomalous α-quartz (called cryptic coesite), which developed under high pressure. For this peak to appear experimentally, quartz must be shocked at several tens of GPa.[60] Because this level of pressure was not reached during the crushing of our samples, we therefore propose that the presence of this peak is related to a surface effect. Indeed, the crushing process may create grains with uncommon surface crystalline orientations leading to particular dangling bonds that

have a similar vibrational mode to those of the cryptic coesite. Thus, the probability of observing this peak at 520 cm^{-1} increases with higher grain density in the crushed samples, which agrees with our observations in this study.

To verify this assertion, several Raman analyses were made on the surface of large fragments broken off of the quartz sample. As shown in Fig. 9, the peak at 520 cm^{-1} was observed in spectra made on the surface of concentric rings of conchoidal fractures. The peak disappears when the analysis is made deeper into the sample. These observations on an uncrushed sample are in accordance with a surface effect.

The appearance of new peaks around 280 cm^{-1} in the silicon spectrum is explained by the acoustic phonons of representation $A_{1g} + E_g + F_{2g}$ appearing in a certain polarisation configuration $(z'(y'y')\bar{z}')$.[26,27] This phenomenon is only observable with the WITec system, which uses a linearly polarized laser beam, whereas the RLS simulator does not. In powders, grains are randomly oriented and the density of specifically oriented grains increases with decreasing grain size.

Mineral identification using Raman spectroscopy is made by comparing an acquired spectrum with that of reference spectra from the literature or from compiled databases (e.g. RRUFF). Significant changes in the characteristics of Raman spectra, such as those observed in silicon, can lead to serious misidentification of species.
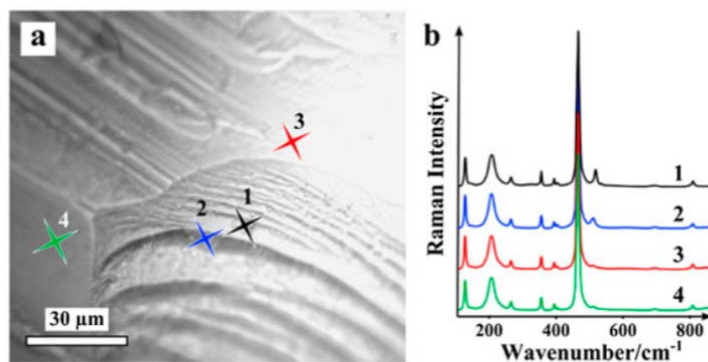
Figure 9. (a) Optical micrograph and (b) associated Raman spectra of a fracture at the surface of the quartz sample. This figure is available in colour online at wileyonlinelibrary.com/journal/jrs

By using the spectrum of the smallest grain size powder of silicon, comparison with the RRUFF database via the CrystalSleuth software (http://rruff.info/) shows a match of only 14% with silicon, the best match (93%) being Paulingite, a rare zeolite mineral with the formula $(K, Na, Ca)_{3-4}(Si, Al)_{21}O_{42}\cdot17-22H_2O$. This experience underlines the potential problems that may arise by Raman analyses of crushed minerals.[19]

Because larger spot size of the ExoMars RLS results in lower irradiance, the changes in spectral parameters described earlier are strongly reduced, and many of the thermal effects are negligible or nonexistent, with the exception of the increase in background level. Changes induced by the structure of the powder itself (e.g. new peaks and surface effects) are relatively limited, and we conclude that the crushing of samples during the future ExoMars mission will not have a necessarily negative effect on identification of the Raman spectra of minerals.

Rocks

By definition, a rock contains several mineral phases that can sometimes be of very small size, such as the basalt analysed in this study. If the Raman spot size is larger than these phases, the resulting spectrum will correspond to a mixture of the individual spectra of several minerals, as demonstrated by the RLS system with its 50 μm spot size. It has previously been shown that the limit of identification of an unknown material is different from the limit of detection of a known material.[19] Indeed, a single peak is most of the time not sufficient to indentify a material, in particular when mixed with other compounds that may exhibit similar Raman bands. The technical parameters such as the wavelength of the laser, the spot size and the CCD camera must be carefully chosen to improve this identification limit.[19] This assertion is all the more true while working on powdered rocks that are characterised by a mixtures of several compounds. However, we have shown that deconvolution of the different spectra can be facilitated by the acquisition of several tens of spectra at different locations on the sample because the probability of always obtaining exactly the same mixture is very low. Most of the minerals present in the rocks used in this experiment, as well as the carbonaceous matter, were thus detected and indentified with the RLS system. Moreover, because of the mixing of the components in the powder, the probability of detecting minor phases within a sample (including those occurring at down to 1% of the mass[52]) is strongly increased. For instance, the RLS system allows detection of carbonaceous matter in the cherts when the probability of detecting it by chance in the bulk sample would have been very low because of its specific distribution. In this case, working on powders is advantageous, except if the detected phase would not be 'representative' of the sample (e.g. inclusion, alteration and contamination).

The most important influence of the crushing process is on the loss of texture, i.e. the spatial relationship between the minerals. Without this textural context, it would be difficult to distinguish between rocks consisted of the same minerals but formed in different contexts (basalts and gabbros for example). This loss of texture will also have consequences on the search for past traces of life. As noted earlier, Raman spectroscopy cannot distinguish between biotic and abiotic carbon. Only the direct correlation of carbonaceous matter with specific structures and/or minerals in the Josefsdal and Kitty's Gap cherts could indicate a potential biosignature. Whereas Raman mapping and optical microscopy on bulk rock allow observation of these very fine details, they

are destroyed once the sample has been crushed, and thus, the detection of biosignatures is strongly compromised.

Our study demonstrates that, despite the limitations of the analytical possibilities of the in situ ExoMars RLS system (i.e. only few tens of spectra per sample and no high-resolution mapping), this system is well adapted to analysing crushed samples and can efficiently identify even minor phases in a sample. For instance, carbonaceous traces might be more easily detected in powdered samples than on the bulk ones. In combination with the high-resolution images of the sample taken by the close-up imager instrument, it might be possible to locate it or infer its mineralogical association. The possibility of preparing polished sections of rocks in situ during a mission would make Raman mapping on such samples very useful, in particular if associated with optical observations made by a close-up imager or a high-resolution camera. Finally, classical analysis combining images of texture and spectral data on a bulk sample in their relative context is preferable when working with laboratory systems.

Conclusion and perspectives

The crushing of minerals may induce changes in Raman spectra in comparison with spectra generated from analyses made on bulk samples. A general increase of the background level, shift in peak position, an increase in peak width and the appearance of new peaks were observed, depending on the material analysed. These effects are induced by a local increase in heat and by the enhancement of crystallographic surface orientation interactions. Because of the relatively large spot size of the ExoMars RLS system, most of these observed spectral changes for powders analysed are negligible or nonexistent. Sample preparation (i.e. crushing) causes a loss of texture unfavourable and potentially destructive of fine-scale features within rocks including any potential traces of life. Thus, detailed optical observations of samples prior to crushing will be essential to ensure acceptable interpretations of data generated by RLS and the other instruments of the Pasteur payload. This study demonstrates that the ExoMars RLS system is suitable for analysing crushed samples and is capable of detecting many phases with only few acquisitions. Furthermore, current work at UVa-CAB demonstrates that the limit of detection is enhanced and that semi-quantification could be possible using powdered samples, thus ensuring the best possible scientific return from the RLS instrument in situ on Mars.

Supporting information

Supporting information may be found in the online version of this article.

References

[1] H. G. M. Edwards, S. E. Jorge-Villar, J. L. Bishop, M. Bloomfield, *J. Raman Spectrosc.* **2004**, *35*, 458.
[2] A. Wang, L. A. Haskin, T. J. Wdowiak, S. W. Squyres, R. J. Wilson, L. E. Hovland, K. S. Manatt, N. Raouf, C. D. Smith, *J. Geophys. Res.* **2003**, *108E1*, 5005.

[3] A. Wang, K. E. Kübler, B. L. Jolliff, L. A. Haskin, *Am. Mineral.* **2004**, *89*, 665.

[4] N. Tarcea, M. Harz, P. Rösch, T. Frosch, M. Schmitt, H. Thiele, R. Hochleitner, J. Popp, *Spectrochim. Acta A* **2007**, *68*, 1029.

[5] J. J. Freeman, A. Wang, K. E. Kübler, B. L. Jolliff, L. A. Haskin, *Can. Mineral.* **2008**, *46*, 1477.

[6] J. Jehlicka, H. G. M. Edwards, P. Vitek, *Planet. Space Sci.* **2009**, *57*, 606.

[7] C. P. Marshall, H. G. M. Edwards, J. Jehlicka, *Astrobiology* **2010**, *10*, 229.

[8] ESA ExoMars website. http://www.esa.int/SPECIALS/ExoMars/SEM10VLPQ5F_0.html

[9] F. Rull, S. Maurice, E. Diaz, C. Tato, A. Pacros, RLS Team, Abstr. 42nd Lunar and Planetary Science Conference, **2011**, 2400.

[10] F. Rull, G. López, A. Catalá, J. Medina, A. Sansano, A. Sanz, F. Sobrón, Abstr. The International Conference: Exploring Mars Habitability, Lisbon, Portugal, June 13–15, **2012**.

[11] J.-L. Josset, F. Westall, B.A. Hofmann, J.G. Spray, C. Cockell, S. Kempe, A.D. Griffiths, M.C. De Sanctis, L. Colangeli, D. Koschny, D. Pullan, K. Föllmi, L. Diamond, M. Josset, E. Javaux, F. Esposito, D. Barnes, *Geophys. Res. Abstr.* **2012**, *14*, 13616.

[12] S. Andò, P. Vignola, E. Garzanti, *Rend. Fis. Acc. Lincei* **2011**, *22*, 327.

[13] E. Garzanti, S. Andò, C. France-Lanord, P. Censi, P. Vignola, V. Galy, M. Lupker, *Earth Planet. Sci. Lett.* **2011**, *302*, 107.

[14] R. E. Summons, J. P. Amend, D. Bish, R. Buick, G. D. Cody, D. J. Des Marais, G. Dromart, J. L. Eigenbrode, A. H. Knoll, D. Y. Sumner, *Astrobiology* **2011**, *11*, 157.

[15] A. Steele, F. M. McCubbin, M. Fries, L. Kater, N. Z. Boctor, M. L. Fogel, P. G. Conrad, M. Glamoclija, M. Spencer, A. L. Morrow, M. R. Hammond, R. N. Zare, E. P. Vicenzi, S. Siljeström, R. Bowden, C. D. K. Herd, B. O. Mysen, S. B. Shirey, H. E. F. Amundsen, A. H. Treiman, E. S. Bullock, A. J. T. Jull, *Sciencexpress* **2012**, *10*, 1120.

[16] J. Jehlicka, H. G. M. Edwards, *Org. Geochem.* **2008**, *39*, 371.

[17] A. Culka, J. Jehlicka, P. Vandenabeele, H. G. M. Edwards, *Spectrochim. Acta A* **2011**, *80*, 8.

[18] A. Culka, J. Jehlicka, L. Strnad, *Spectrochim. Acta A* **2012**, *86*, 347.

[19] P. Vandenabeele, J. Jehlicka, P. Vitek, H. G. M. Edwards, *Planet. Space Sci.* **2012**, *62*, 48.

[20] U. Villanueva, J. C. Raposo, K. Castro, A. de Diego, G. Arana, J. M. Madariaga, *J. Raman Spectrosc.* **2008**, *39*, 1195.

[21] F. Foucher, F. Westall, in *Micro-Raman Spectroscopy and Luminescence Studies in the Earth and Planetary Science*, (Ed.: A. Gucsik), American Institute of Physics, **2009**, *CP 1163*, pp. 47–58.

[22] F. Westall, F. Foucher, B. Cavalazzi, S. T. de Vries, W. Nijman, V. Pearson, J. Watson, A. Verchovsky, I. Wright, J.-N. Rouzaud, D. Marchesini, S. Anne, *Planet. Space Sci.* **2011**, *59*, 1093.

[23] S. K. Sharma, C. H. Chio, P. Deb, P. G. Lucey, N. Domergue-Schmidt, K. A. Horton, XXXIth Lunar and Planetary Science Conference, Houston, USA, March 13–17 **2000**, 2030.

[24] M. E. Back, J. A. Mandarino, *Fleischer's Glossary of Mineral Species* 2008, The Mineralogical Record Inc., Tucson, **2008**, p. 213.

[25] J. L. Jambor, J. Puziewicz, *Am. Mineral.* **1991**, *76*, 665.

[26] P. A. Temple, C. E. Hathaway, *Phys. Rev. B* **1973**, *7*, 3685.

[27] R. Poilblanc, F. Crasnier, *Spectroscopies infrarouge et Raman*, EDP Sciences, Grenoble, **2006**, pp. 466–470.

[28] J. Etchepare, M. Merian, L. Smetankine, *J. Chem. Phys.* **1974**, *60*, 1531.

[29] M. R. Smith, J. L. Bandfield, *J. Geophys. Res.* **2012**, *117*, E06007.

[30] A. P. Nutman, S. J. Mojzsis, C. R. L. Friend, *Geochim. Cosmochim. Acta* **1997**, *61*, 2475.

[31] F. Westall, M. J. de Wit, J. Dann, S. van der Gaast, C. E. J de Ronde, D. Gerneke, *Precambrian Res.* **2001**, *106*, 93.

[32] F. Westall, C. E. J. de Ronde, G. Southam, N. Grassineau, M. Colas, C. Cockell, H. Lammer, *Philos. Trans. R. Soc., B* **2006**, *361*, 18570.

[33] F. Westall, S. T. de Vries, W. Nijman, V. Rouchon, B. Orberger, V. Pearson, J. Watson, A. Verchovsky, I. Wright, J.-N. Rouzaud, D. Marchesini, A. Severine, *Spec. Pap. - Geol. Soc. Am.* **2006**, *405*, 105.

[34] F. Westall, B. Cavalazzi, L. Lemelle, Y. Marrocchi, J.-N. Rouzaud, A. Simionovici, M. Salomé, S. Mostefaoui, C. Andreazza, F. Foucher, J. Toporski, A. Jauss, V. Thiel, G. Southam, L. MacLean, S. Wirick, A. Hofmann, A. Meibom, F. Robert, C. Defarge, *Earth Planet. Sci. Lett.* **2011**, *310*, 468.

[35] A. Hofmann, R. Bolhar, *Astrobiology* **2007**, *7*, 355.

[36] S. W. Squyres, R. E. Arvidson, S. Ruff, R. Gellert, R. V. Morris, D. W. Ming, L. Crumpler, J. D. Farmer, D. J. Des Marais, A. Yen, S. M. McLennan, W. Calvin, J. F. Bell III, B. C. Clark, A. Wang, T. J. McCoy, M. E. Schmidt, P. A. de Souza Jr., *Science* **2008**, *320*, 1063.

[37] F. Tuinstra, J. L. Koenig, *J. Chem. Phys.* **1970**, *53*, 1126.

[38] J. Jehlicka, C. Bény, *J. Mol. Struct.* **1999**, *480–481*, 541.

[39] C. Bény-Bassez, J.-N. Rouzaud, *Scanning Electron. Microsc.* **1985**, *1*, 119.

[40] E. Quirico, G. Montagnac, J.-N. Rouzaud, L. Bonal, M. Bourot-Denise, S. Duber, B. Reynard, *Earth Planet. Sci. Lett.* **2009**, *287*, 185.

[41] J. D. Pasteris, B. Wopenka, *Astrobiology* **2003**, *3*, 727.

[42] O. Beyssac, B. Goffé, C. Chopin, J.-N. Rouzaud, *J. Metamorph. Geol.* **2002**, *20*, 859.

[43] M. R. Ammar, J.-N. Rouzaud, *J. Raman Spectrosc.* **2012**, *43*, 207.

[44] O. A. Maslova, M. R. Ammar, G. Guimbretière, J.-N. Rouzaud, P. Simon, *Phys. Rev. B* **2012**, *86*, 134205-(5).

[45] H. Y. McSween, G. J. Taylor, M. B. Wyatt, *Science* **2009**, *324*, 736.

[46] J.-F. Pastre, A. Chambaudet, J. Couthures, *C. R. Séances Acad. Sci., Vie Acad.* **1983**, *296*, 79.

[47] J.-F. Pastre, *Quaternaire* **2004**, *15*, 87.

[48] H. Y. McSween, R. E. Arvidson, J. F. Bell III, D. Blaney, N. A. Cabrol, P. R. Christensen, B. C. Clark, J. A. Crisp, L. S. Crumpler, D. J. Des Marais, J. D. Farmer, R. Gellert, A. Ghosh, S. Gorevan, T. Graff, J. Grant, L. A. Haskin, K. E. Herkenhoff, J. R. Johnson, B. L. Jolliff, G. Klingelhoefer, A. T. Knudson, S. McLennan, K. A. Milam, J. E. Moersch, R. V. Morris, R. Rieder, S. W. Ruff, P. A. de Souza, S. W. Squyres, H. Wänke, A. Wang, M. B. Wyatt, A. Yen, J. Zipfel, *Science* **2004**, *305*, 842.

[49] F. Westall, *Science* **2005**, *308*, 366.

[50] C. K. Wentworth, *J. Geol.* **1922**, *30*, 377.

[51] A. Sansano, G. Lopez, J. Medina, F. Rull, AMASE 10 team, Abstr. EPSC-DPS **2011**, *6*, 856.

[52] G. Lopez-Reyes, F. Rull, A. Sanz, J. Medina, A. Catala, I. Hermosilla, B. Lafuente, Abstr. GEORAMAN X, Nancy, France, June 11–13, **2012**.

[53] T. M. K. Nedungadi, *Proc. Ind. Acad. Sci., Math. Sci.* **1940**, *11–2*, 86.

[54] T. R. Hart, R. L. Aggarwal, B. Lax, *Phys. Rev. B* **1970**, *1*, 638.

[55] D. A. Long, *Raman Spectroscopy*, McGraw-Hill, New York, **1976**, p. 84.

[56] R. Poilblanc, F. Crasnier, *Spectroscopies infrarouge et Raman*, EDP Sciences, Grenoble, **2006**, p. 319.

[57] T. Kobayashi, T. Hirajima, Y. Hiroi, M. Svojtka, *J. Min. Pet. Sci.* **2008**, *103*, 105.

[58] S. K. Sharma, B. Simons, H. S. Yoder, *Am. Mineral.* **1983**, *68*, 1113.

[59] R. Palmeri, M. L. Frezzotti, G. Godard, R. J. Davies, *J. Metamorph. Geol.* **2009**, *27*, 685.

[60] T. de Rességuier, P. Berterretche, M. Hallouin, *J. Appl. Phys.* **2003**, *93*, 2123.

Chapter 5.

Identification and quantification of minerals from Raman spectra with the RLS instrument

Mineral identification and abundance quantification capabilities based on the automatically acquired spectra from the RLS instrument, by applying different analytical techniques.

5.1 INTRODUCTION

Raman spectroscopy is a powerful tool to perform mineral analysis of samples. It can also easily detect and identify individual mineral phases in complex geological mixtures. This is very interesting, especially in a framework where all the analyzed samples will be in powdered form, as the ExoMars mission. Nevertheless, depending on the present materials, the differences of emission efficiency can vary several orders of magnitude [37]. Under these constraints, it is relatively easy to qualitative identify the mineral phases present in mixtures, as long as the relative efficiencies (related to the molecular cross-section) are in similar orders of magnitude. It becomes complicated to identify minor inefficient materials in good Raman scattering matrix materials. Furthermore, quantifying the abundance of the mineral phases present on complex mixtures becomes even more complicated, especially given that cross-section of samples is not generally available, or easy to calculate.

In this chapter, we present a study of the capabilities of the RLS instrument, from the analytical point of view, using two different analysis paradigms. To do so, we summarize the results published in two papers in the European Journal of Mineralogy (EJM) and the American Mineralogist (AmMin) journals, which are annexed for reference in their corresponding sections.

This work is included in the framework of the thesis as the necessary next step in the development of the RLS instrument. First, the instrument is automatized to provide it with methods for acquiring the best possible spectra, from a purely spectroscopic perspective. Then, the operational context of the instrument onboard the ExoMars rover studied taking into account the engineering constraints imposed by a space rover-driven mission in the resulting analysis. Finally, this chapter addresses the following logical step: study of possible techniques for the analysis of the resulting data obtained from the RLS instrument (under the conditions described in the previous chapters). This is necessary in order to evaluate the capabilities of the instrument in terms of detection thresholds, but also to study if any complementary information can be obtained from sample mappings as a set, instead of analyzing the spectra at different points individually. To do so, several techniques are evaluated with different types of samples including natural rocks, providing information regarding their applicability to the detection of mineral phases, but also for their quantification in relatively simple mixtures.

Section 5.2 summarizes the work published in [56], by Lopez-Reyes, G., et al., (2013) *Analysis of the scientific capabilities of the ExoMars Raman Laser Spectrometer instrument,* European Journal of Mineralogy, 25(5), 721-733. There, it analyzes the reach of the analysis with spectra from the RLS data, acquired with conditions representative of the ExoMars rover from powdered samples of natural rocks and

synthetic samples. This analysis is always based on direct identification of the spectra, or else, on standard statistical and simple techniques to try to assess the mineral abundance of the samples.

Section 5.3 presents the work published in [57], by Lopez-Reyes, G. et al., (2014) *Multivariate analysis of Raman spectra for the identification of sulfates: Implications for ExoMars,* American Mineralogist, 99(8-9), 1570-1579. This paper analyzes and compares several multivariate analysis techniques, in order to assess the validity of this kind of techniques for the qualitative/quantitative detection of minerals in binary mixtures.

5.2 STANDARD ANALYTICAL TECHNIQUES

As explained in previous chapters, the operation mode of the ExoMars rover imposes the instruments in the Analytical Laboratory Drawer (ALD) to analyze powdered samples that are presented to the instruments with a flattened surface. In the RLS instrument case, at least 20 points of each sample will be analyzed, and those spectra will be downloaded to earth.

The crushing process can induce several effects on the acquired Raman spectra, as discussed in detail in Chapter 4 (though with limited impact for an instrument with the characteristics of RLS). Furthermore, it implies the loss of textural context of the samples.

In this work, we analyzed two natural samples from Martian analog sites: a carbonate breccia from Svalbard (Norway) and a silicified volcanic sand sample (chert) from the Barberton Greenstone Belt in South Africa. The qualitative analysis of these samples provided some interesting results: The Raman analysis of the bulk carbonate breccia provided a clear sequence magnesite – dolomite – aragonite in the carbonate crust of the sample (see Fig. 3 in EJM paper). However, the analysis of the powdered sample also detected ankerite. This proved the presence of iron and manganese in the carbonate sequence, which confirmed the results previously reported on this sample by [58].

On the other hand, the analysis of the chert from South Africa, showed that Raman was able to detect the same phases both in bulk and powdered form. However, the relatively sparse carbon presence on the bulk sample was very easily detected in the powdered material (see Fig. 5 in EJM paper), thanks to the crushing process and Raman spectroscopy high sensitivity to carbon. Furthermore, compared to other techniques, Raman spectroscopy provided more information regarding the presence of materials (e.g., XRD failed to detect both anatase and carbon).

The analysis of these two samples shows that, though the texture and the spatial correlation of the detected minerals is lost during the crushing process, the results on

the powder provide relevant information that, in combination with other instruments (e.g., an imaging device as ExoMars CLUPI), can improve the results obtained with the RLS instrument to provide better science return than other techniques, or, in some cases, the analysis on bulk materials.

In addition to the natural samples, experiments were performed in with various powdered samples of binary mixtures, with different proportions varying from 1% - 99% to 99% - 1%. Up to five lines of 30 points each were acquired for each sample. These analyses showed that, for the analyzed materials, the detection threshold for the RLS instrument was as low as 1% with 30 points along a line, and with grain size of 250 microns and lower.

In addition, by performing a very simple statistical calculation it was possible to perform a pseudo-quantification of the mineral abundances on the mixtures. Furthermore, this method did not require the use of the cross-section of the different materials in the mixtures. The reason is that the RLS instrument adjusts the acquisition parameters automatically, which is *de facto* a normalization with respect to the Raman efficiency of the major component of the mixture. Fig. 12 of the EJM paper shows how the RLS instrument adjusts the integration time as a function of the scattering efficiency of the major component of the mixture. This is also the reason why the most unbalanced proportions present less deviation in Figs. 8 and 9 in the EJM paper: the more balanced the mixture proportion, the more noticeable the different Raman cross-sections of the materials.

The results of this section show that the crushing process of the samples in the ExoMars rover seriously hampers the potential science return due to the loss of context of the samples. Nevertheless, there are some results that show that the analysis of the powdered samples might provide, in some cases, better detection capabilities than the bulk sample. Furthermore, it might even allow some kind of quantification of the mineral abundances in binary mixtures, without the need to previously calibrate the mixture.

Below, the complete paper published in the European Journal of Mineralogy is included for reference.

Guillermo Lopez-Reyes

Analysis of the scientific capabilities of the ExoMars Raman Laser Spectrometer instrument

GUILLERMO LOPEZ-REYES[1,*], FERNANDO RULL[1], GLORIA VENEGAS[1], FRANCES WESTALL[2], FRÉDÉRIC FOUCHER[2], NICOLAS BOST[2,3,4,5], AURELIO SANZ[1], ALEJANDRO CATALÁ-ESPÍ[1], ALBERTO VEGAS[6], ISAAC HERMOSILLA[1], ANTONIO SANSANO[1] and JESUS MEDINA[1]

[1] Unidad Asociada UVa – CSIC – Centro de Astrobiologia, Avda. Francisco Valles, 8, 47151 Boecillo, Spain
*Corresponding author, e-mail: guillermo.lopez.reyes@cab.inta-csic.es
[2] Centre de Biophysique Moléculaire, CNRS, Rue Charles Sadron, 45071 Orléans cedex 2, France
[3] Université d'Orléans, ISTO, UMR 7327, 1A rue de la Férollerie, 45071 Orléans, France
[4] CNRS/INSU, ISTO, UMR 7327, 1A rue de la Férollerie, 45071 Orléans, France
[5] BRGM, ISTO, UMR 7327, BP 36009, 45060 Orléans, France
[6] IXION Industry and Aerospace, C/ Julián Camarillo 21B – 3ª planta, 28037 Madrid, Spain

Abstract: The Raman Laser Spectrometer (RLS) is part of the payload of the 2018 ExoMars rover. The Sample Preparation and Distribution System (SPDS) of the rover will crush samples acquired from down to two meters depth under the Martian surface, and provide them to the RLS instrument in the form of flattened powdered samples. The RLS instrument will acquire a minimum of 20 points on the flattened surface of the samples. To be able to obtain the maximum scientific return from the instrument once on Mars, a simulator of the SPDS system has been built to perform a series of experiments in a representative scenario. The crushing process implies the loss of rock structure and texture and, hence, the geological context of the samples. However, qualitative analysis with the RLS simulator on powdered natural samples and rocks showed that the RLS is capable of detecting carbonaceous material occurring in trace amounts in one of the rock samples (a silicified volcanic sand), more easily than with the same analysis on bulk. Furthermore, it is shown that minor phases in carbonate cements that cannot be detected by Raman in the bulk sample can be detected in the powder, thus allowing the identification of all the carbonate phases present in the cement crust.

In order to quantify the detection threshold of the instrument, further analysis on controlled samples were performed. The results with the RLS SPDS simulator showed that the instrument can reach detection thresholds down to 1 % on powdered samples. Furthermore, analysis of controlled mixtures showed that performing a very simple intensity-based statistical analysis of the spectra can provide semi-quantification of the abundance of the mineral species with quite linear calibration curves.

Key-words: ExoMars, RLS, Raman spectroscopy, powdered samples, geological context loss, detection thresholds, quantification.

Introduction

The European Space Agency's future ExoMars mission will send a rover to Mars in 2018 with a payload of scientific instruments including the Raman Laser Spectrometer (RLS) (Pacros *et al.*, 2009). The ExoMars programme has been established by ESA to investigate the Martian environment and try to establish whether life ever existed on Mars (ESA, 2012). To do so, the ExoMars rover is equipped with a drill to obtain samples from the Martian sub-surface down to two meters depth, where the probability of accessing well-preserved organic molecules or biomarkers is much higher than on the surface (Catalá-Espí, 2010). In addition to organic matter, the rover instrument suite will also analyze the mineralogy of rocks and regolith at the landing site.

Raman spectroscopy provides a powerful tool for the definitive identification and characterization of minerals (Sharma *et al.*, 2003; Wang *et al.*, 2006) and biomarkers (Edwards *et al.*, 2003, 2011; Pacros *et al.*, 2009). In addition, technological advances and the miniaturization of the equipment have allowed this spectroscopic technique to be considered as a good candidate for planetary exploration during the last decade (Sharma *et al.*, 2003; Wang *et al.*, 2003; Angel *et al.*, 2012). Thus, the RLS instrument onboard the ExoMars rover will be able to identify the composition and structure of minerals and rocks at the mineral grain scale. Indeed, the Raman spectrometer will be used to identify organic compounds and search for signatures of life, to identify the mineral products and indicators of biological activities, to characterize mineral phases produced by water-related processes and to

172

722   G. Lopez-Reyes *et al.*

characterize igneous minerals and their products resulting from alteration processes (ESA, 2012).

In order to meet the flying robustness requirements of flight instrumentation, as well as to comply with the restrictions imposed by the interaction with other instruments of the payload, the instruments of the ExoMars Analytical Laboratory Drawer (ALD) share a common Sample Preparation and Distribution System (SPDS). The SPDS will provide the instruments in the ALD with a flattened powdered sample obtained by a drill from two meters down the surface of Mars (Vago *et al.*, 2006; Rull, 2012). The RLS, in its automatic mode, will analyze a set of at least 20 points in each sample, with a 50 µm spot size and an irradiance level of 0.6–1.2 kW cm^{-2} with a 532 nm continuous wave (CW) laser (Rull *et al.*, 2011).

The crushing process implies the loss of the geological context within the sample, and it has been seen that the analysis of powdered samples can cause some inherent undesirable effects such as background increase or peak widening with respect to the same analysis performed on bulk samples (Foucher *et al.*, 2013). However, most of these effects are not observed with the RLS instrument due to its technical specifications, while other unexpected advantages can be met (Foucher *et al.*, 2012). In addition, the close-up imager CLUPI on the rover will provide high-resolution images of the sample before crushing. This step is important for recording the spatial structure and texture of the sample and, therefore, its geologic context (Josset *et al.*, 2012). Even though the spatial correlation between the optical images and the Raman mapping on the powder will be lost, the combination of both optical and spectroscopic data can provide information that is vital for aiding interpretation of potential signatures of life (Foucher *et al.*, 2012). In addition, the crushing process provides the instruments with the ability to analyze the interior of the samples. It is important to note that analysis by the instrument in the automatic mode will be performed by randomly mapping the flattened surface of the powdered sample.

In this paper, the ExoMars RLS simulator system is presented. This simulator aims at providing a system in which the Raman scientific team can perform experiments in a rover-like scenario. The ExoMars RLS Simulator has provided the necessary framework for the development of the instrument operation mode, which prepares the instrument to operate autonomously. For example, the algorithm for acquisition parameters adjustment allows the instrument to automatically optimize the integration time and number of accumulations at each point, as a function of intensity and signal-to-noise ratio of a reference spectrum. In addition, the automatic and systematic analysis of samples with the ExoMars RLS Simulator has helped defining the scientific capabilities of the instrument once on Mars, as well as the study of processing techniques which allow obtaining the maximum scientific return from the instrument data products. To address this issue, the results from several qualitative analyses on natural samples (a carbonate cement from Svalbard, Norway, and a chert from Barberton, South Africa) as well as on powdered synthetic mixtures are presented.

Previous studies have shown that, by means of macro techniques as X-ray diffraction (XRD) and Fourier Transform (FT) Raman spectroscopy, it is possible to quantify the presence of mineral species in a mixture by calculating intensity-based calibration curves (Kontoyannis *et al.*, 1997; Kontoyannis & Vagenas, 2000; Dickinson & McGrath, 2001; Vagenas & Kontoyannis, 2003). In this paper the analysis with the RLS instrument of a set of controlled synthetic binary mixtures in different proportions is presented to evaluate this statistical method. The results are compared to those obtained with XRD (InXitu Terra-185; 30 kV–300 µA Cobalt tube; range 5–55° 2θ; resolution 0.25° 2θ) and FT Raman (Bruker RSF 100/S; Laser Coherent Compass 1064 nm, 500 mW Nd:YAG; CCD Bruker D418-T. Acquisition parameters: 450 mW, spot diameter 1 mm, 2048 accumulations).

The ExoMars RLS simulator

The RLS instrument will operate in one of two high-level operational modes (Rull *et al.*, 2011): an automatic scanning and a smart scanning. During the automatic scanning, the rover will place the target following a preconfigured sequence of movements under the RLS optical head. Then RLS will take at least 20 shots per sample at regularly spaced spots on the target. On the other hand, during the smart scanning, MicrOmega IR (Leroi *et al.*, 2009) images will be processed by the rover to determine if there is any target of interest which can be reached by the RLS instrument, and the rover will place such samples under the RLS optical head for analysis.

The Martian subsurface samples obtained with the rover drill will be provided to the Analytical Laboratory Drawer (ALD) instruments by means of the rover Sample Preparation and Distribution System (SPDS). This system will crush the sample to a powder with a grain distribution under 200–250 µm (Sansano *et al.*, 2011), and deposit it in a sample container attached to a carrousel. The carrousel will then move the container below a blade to flatten the surface of the powder before presenting it to the ALD instruments.

In order to achieve the maximum scientific return from the RLS instrument once on Mars, it is necessary to perform experiments under the operation conditions imposed by the rover. We have thus developed the so-called ExoMars RLS Simulator at UVa-CSIC-CAB Associated Unit ERICA (Fig. 1). This Simulator allows the scientific team to perform several scientific key experiments under conditions similar to those provided by the ALD of the ExoMars rover. It also provides the means for the definition and development of the necessary algorithms for the automation of the instrument measurements.

The basic configuration of the ExoMars Simulator consists of an XYZ micrometric positioning system with an optical head attached to the Z axis, and a refillable container attached to the XY positioners (Fig. 1).
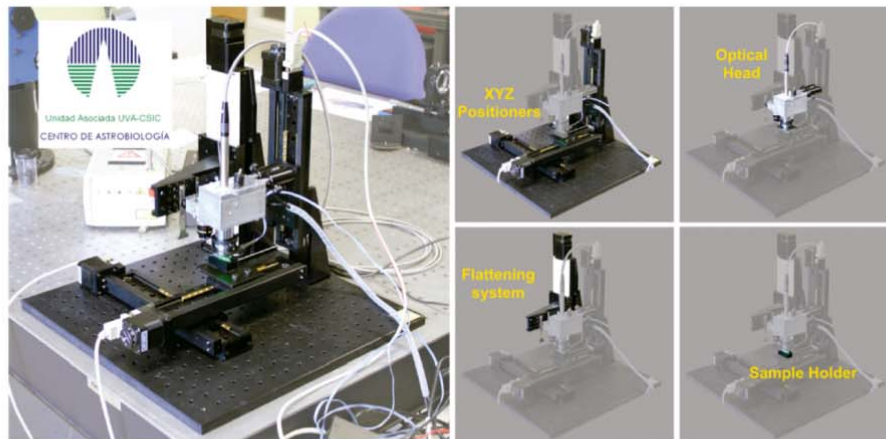
173

Fig. 1. ExoMars RLS Simulator prototype and schematics.

The SPDS carrousel and the focusing system in RLS are emulated with linear positioners (X and Z axis, respectively), assuming that the carrousel radius is big enough to be approached by a straight line. Furthermore, another axis (Y axis) was added to be able to address complementary studies and two-dimensional mapping. The system allows the samples to be analyzed at the mineral grain scale with three-dimensional positioning accuracy of 2.5 μm.

The Raman optical head has a spot size of 50 μm which matches the requirements of the flight instrument inside the rover. An internal camera allows imaging of the sample through the same optical path. In addition, the Raman optical head is coupled to an external camera so that the sample can be imaged using different fields of view. With this coupled device it is possible to locate the position of the laser beam and to conduct analysis on specifically selected locations on the sample, such as making a profile of the flattened surface.

The system is also designed to flatten the sample powder with a SPDS-like design and to position the sample under the instrument optical head, which will automatically focus on the sample surface.

Operation mode

The operation mode of the RLS instrument consists in a series of operations and algorithms that the instrument has to perform autonomously to obtain a Raman spectrum of the highest possible quality in the limited time available during operation on Mars. The definition of the operation mode is a purely scientific task, which shifts the actions that the human operator performs on Earth with the spectrometer (laser focusing on the sample, adjustment of spectrum's acquisition time, minimization of fluorescence effects, *etc.*) to a system able to autonomously decide the

focusing point and what will the optimal acquisition parameters be, *etc.*

The spectral quality of the acquired spectrum depends, for each sample, on two main acquisition parameters: the integration time (T_i) and the number of accumulations (N_a). The developed algorithms are able to estimate these parameters based on the signal-to-noise ratio value of a short-time acquisition during the adjusting phase of the spectral acquisition process.

Mineral analysis

Identification of natural samples

To evaluate the analytical capabilities of the instrument, taking into account the restrictions imposed by the rover-based operation (random mapping of a homogeneously distributed powdered sample), two natural samples from Martian analog sites have been qualitatively analyzed by means of the ExoMars RLS Simulator and the results compared with those obtained with complementary techniques, such as micro-Raman analysis on the bulk sample at interesting points, or other macro techniques, such as XRD and FT Raman. The aim of these analyses is to be able to assess the detection/non-detection capabilities of the instrument under these constraints, as well as to evaluate the scientific information that can be inferred from the Raman instrument data products.

The analyzed samples were provided by the International Space Analogue Rockstore (ISAR, www. isar.cnrs-orleans.fr, see Bost *et al.*, 2013), a collection of Mars-analogue rocks and minerals, or collected *in situ*. The selected samples include a carbonate breccia from Svalbard (Norway), as well as a silicified volcanic sand (chert) containing biosignature traces from the Barberton

G. Lopez-Reyes *et al.*

Greenstone Belt (South Africa). The samples were qualitatively studied in order to assess the strengths and limitations of Raman spectroscopy and the RLS instrument.

Carbonate breccia

A sample of carbonate breccia was collected in Svalbard (Norway) during the expedition AMASE'10 (Steele *et al.*, 2011), from a breccia associated with subglacial volcanism that erupted about 1 million years ago at the Bockfjord Volcanic Complex. These carbonates were deposited by magnesium-rich, hot (probably hydrothermal) waters associated with volcanic activity (Treiman *et al.*, 2002) that cemented basaltic breccia fragments. They are considered to be the closest terrestrial analogs to the Martian carbonates found by the Spirit rover in the Columbia Hills of the Gusev crater in Mars (Morris *et al.*, 2010). Moreover, they have similarities with the carbonates found in fractures of the ALH84001 meteorite (Treiman *et al.*, 2002).

Optical, cathodoluminescence (CL) and electron-microprobe observations performed on thin sections of another sample of the carbonate-cemented breccia (ISAR sample 09SJ15) show that the carbonate breccia crust is layered and that each layer is a different carbonate phase (Bost, 2012) (Fig. 2). The external rings are formed first by

magnesite followed by dolomite (the absence of luminescence in the CL image demonstrates the lack of Mn and Fe). In the second layer, CL confirms the porous texture of the small carbonates grains, which contain Mn (orange-red color in CL associated to the luminescence of the ion Mn^{2+} in the carbonate crystalline lattice, Habermann *et al.*, 2000). The variation in composition of the different carbonate layers was also documented by analyses of these carbonates performed during the AMASE campaign 2010 by the Curiosity rover XRD instrument (CheMin). They showed sequences including magnesite and dolomite or ankeritic dolomite, siderite and minor aragonite (Blake *et al.*, 2011).

Raman analyses of this sample were performed with RLS both in bulk and in powder obtained from a breadboard model of the ExoMars rover crushing station developed by Kayser-Threde.

The analysis on the bulk sample was performed by making a line spectrum to evaluate the carbonate sequence on a flat surface of the carbonate crust (Fig. 3). The results showed a clear sequence: magnesite – dolomite – aragonite in the carbonate crust from the inner part of the crust (in contact with the basalt) to the external face of the sample. These results are in accordance with the petrological analyses and also

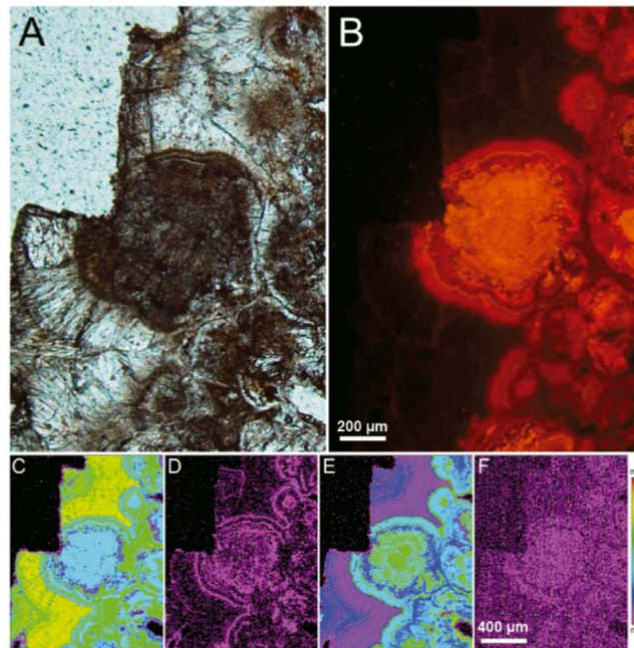Fig. 2. Optical observations and analyses of a carbonate sample from Svalbard (Norway). (A) Optical view in transmitted light. (B) Cathodoluminescence image showing the presence of Mn in the second layer of carbonates. (C-F) Electron microprobe observations: elemental maps corresponding to Mg, Ca, Fe and Mn respectively. Low contents are represented in purple and high contents in red (modified after Bost, 2012).

Fig. 3. Flat surface of a carbonate cement extracted from a volcanic breccia in Svalbard (Norway). The inner part consists of a basaltic core (black), cemented with Mg-Ca carbonates (orange).

Josefsdal Chert

These cherts are hydrothermally-silicified volcanic sediments and are among the oldest sedimentary rocks on Earth (Nutman *et al.*, 1997). They contain the silicified, carbonaceous traces of primitive life forms that can be considered as analogues of primitive life that may have appeared on pre Noachian-Noachian Mars ($\approx$ 4.5–3.7 Ga) (Westall *et al.*, 2006, 2011a and b). Although volcanic sediments abound on Mars, their silicified counterparts have not yet been detected although silica of possible hydrothermal origin has been identified (Squyres *et al.*, 2008; Ruff *et al.*, 2011).

We have studied a sample of the 3.3 Ga-old Josefsdal Chert from Barberton Greenstone Belt in South Africa (ISAR sample 99ZA07) using optical microscopy on a polished cut of the rock surface and on thin sections of the sample, as well as XRD and FT Raman and also automated analysis on a powdered sample with the RLS instrument simulator.

The optical observations focused on laminated structures that appeared dark in transmitted light and probably represent either sedimented carbonaceous matter or laminar microbial mats (observable by Scanning Electron Microscope, Westall *et al.*, 2006, 2011a). Both the microbial mats and the sediments were rapidly lithified by hydrothermal silica (Fig. 4). The individual organisms making up the microbial mats are small, *i.e.* filaments < 0.5 µm in thickness, and below the resolution of the optical microscope, but their collective thickness within the mat-like structure (of the order of $\sim$10 µm) is visible. Thus, it is

with the results of Blake *et al.* (2011), except for the absence of iron and manganese carbonates in our analysis. A strong presence of goethite was found in the carbonate crust, providing the sample its orange coloration (Sansano *et al.*, 2011).

Analysis on the powdered sample was made using random selection of 30 points. This allowed identification of not only all the mineral phases detected in the bulk sample but also ankerite, not found in the bulk analysis, and which confirms the presence of iron and manganese in the carbonate sequence (magnesite – dolomite – ankerite – aragonite), as noted previously (Blake *et al.*, 2011). Thus, even though the spatial context of the powdered sample was lost, the carbonate phases detected with the RLS instrument provide precise information for inferring what the carbonate sequence might have been, especially when complemented with information from other instruments. The results are summarized in Table 1. They demonstrate the usefulness of these kinds of analyses using the RLS on crushed samples.

Fig. 4. Optical image of the Josefsdal chert of Barberton (South Africa). Both the microbial tidal flat mats (thin layered structures) and carbonaceous sediments (dark patches) can be observed.

Table 1. Mineral phases detected within the carbonate from Svalbard (Norway), on the bulk and powdered sample.

	Pyroxene	Olivine	Goethite	Magnesite	Dolomite	Ankerite	Aragonite
Raman on solid	✓	✓	✓	✓	✓		✓
RLS on powder	✓	✓	✓	✓	✓	✓	✓

G. Lopez-Reyes *et al.*

the spatial distribution of the carbonaceous matter within a specific mineralogical context, *i.e.* at the surfaces of the volcanic sediment beds (now silicified) that provides clues about the potential biotic origin of the structures. This information is provided by analysis of the bulk sample using Raman spectroscopy. In addition to quartz and kerogen, anatase related to the alteration of the volcanic grains in the presence of water is common in this chert (Foucher & Westall, 2009).

This sample was analysed with the RLS instrument under operational conditions as those of the rover (powdered samples, limited number of spectra acquired in a limited time) which showed that this instrument was able to detect all the mineral phases present in the sample with only a very limited number of points. Furthermore, random analysis of the powdered sample was able to easily detect the relatively rare carbon that was observed on the bulk sample, as shown in Fig. 5. This is because the crushing process distributes the carbon homogeneously and also due

to the high sensitivity of Raman spectroscopy to carbon compounds (Wang *et al.*, 1989). Macro analysis of the sample with XRD and FT Raman was subject to the same limitations imposed by the crushing of the samples. FT Raman detected all the minerals that were present in the sample while XRD only identified the quartz in the chert. Comparison of the qualitative detection using the different techniques is shown in Table 2.

Table 2. Mineral phases detected by different techniques in the Josefsdal chert from Barberton (South Africa).

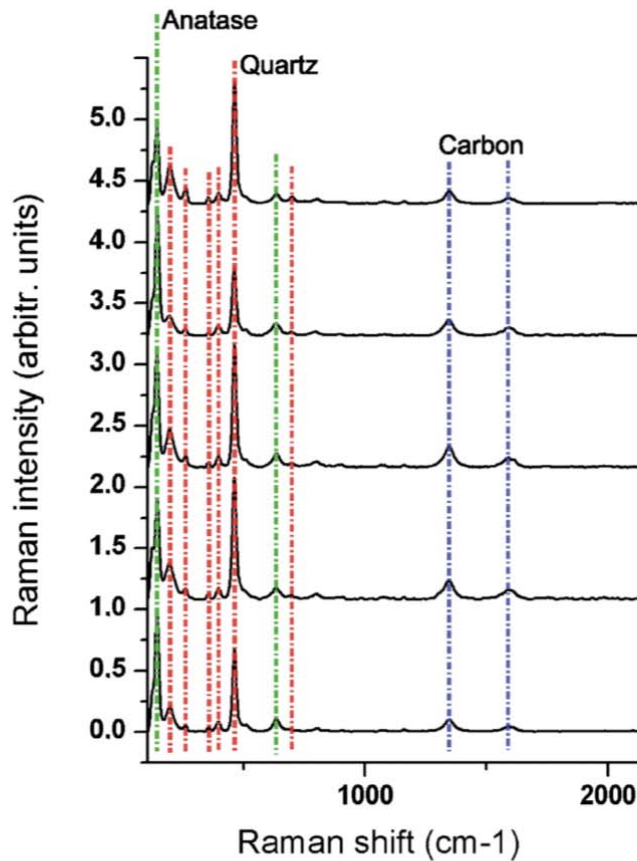|	Quartz	Carbon	Anatase
Raman on solid	✓	✓	✓
RLS on powder	✓	✓	✓
XRD	✓		
FT Raman	✓	✓	✓

Fig. 5. Spectra obtained from Josefsdal chert of Barberton (South Africa) with RLS. Anatase (green), quartz (red) and carbon (blue) were detected with the RLS instrument.

The most important limitation of analyses on powdered samples is the loss of geological texture and structure, in this case hampering the possibility of associating the carbon phase with either specific mineral phases or sedimentary structures (*e.g.* bedding planes), or to actual ancient biotic structures, such as stromatolites. However, as already stated, the ExoMars rover close-up imager CLUPI, with a resolution of 7 μm at 10 cm, will acquire images of the bulk sample before crushing that will complement the RLS data. ExoMars therefore has a suite of instruments with very interesting capabilities for the analysis of *in situ* samples on the red planet.

Synthetic mixtures

Here we present the results of analyzing a series of controlled and well-characterised samples consisting of mixtures of two phases in different proportions. The aim is to establish the detection limits of the instrument when working under the conditions imposed by the rover operation and to compare them to other macro techniques, such as XRD or FT-Raman. In addition, we use a statistical method to try to quantify the abundance of the mineral phases present in the mixtures.

Experiment description

Several sets of synthetic mixtures were prepared with a grain size distribution of ≤ 250 μm. The phases used in the mixtures were chosen on the basis of having clearly separate spectral bands, thus eliminating the necessity of deconvolution and fitting of overlapping peaks. Here we used mixtures of sodium oxalate + potassium nitrate and calcite + gypsum.

For each mixture, a series of five lines of 30 points were analyzed for a set of up to nine proportions ranging from 1 % to 99 % in weight. In total, 1350 points were analyzed (at each point, the optical head has to be focused, the acquisition parameters have to be calculated and the final spectrum has to be acquired). This analysis was feasible thanks to the ExoMars RLS Simulator described in previous sections, which allowed automated analysis.

Statistical analysis

In order to maximize the scientific return obtained from the RLS instrument data products, a statistical procedure previously applied in macro techniques, such as XRD and FT-Raman (Kontoyannis *et al.*, 1997; Kontoyannis & Vagenas, 2000; Vagenas & Kontoyannis, 2003) was applied. However, the spot size of the RLS instrument is only 50 μm and only single or a few grains are detected at the same time. Thus, the statistical analysis has to be performed on the relatively few (20–30) micro-Raman spectra acquired per sample. The detailed point by point procedure is divided into three steps:

(1) Calculate the Raman intensity of the main peak for each phase at each of the *N* points, with respect to the background spectrum level.

(2) Calculate at each point *i* the ratio $r_{material\_i}$ between the peak intensity of each phase and the total intensity of all the phases:

$$r_{material1_i} = \frac{I_1}{I_1 + I_2} ; r_{material2_i} = \frac{I_2}{I_1 + I_2}$$

(3) Calculate the quantification parameter $\bar{r}_{material}$ as the average of the intensity ratio for each phase of the mixture (average of all $r_{material\_i}$):

$$\bar{r}_{material1} = \sum_{i=1}^{N} \frac{r_{material1_i}}{N} ;$$

$$\bar{r}_{material2} = \sum_{i=1}^{N} \frac{r_{material2_i}}{N}$$

Results and discussion

For each mixture, spectra were automatically acquired in five lines of 30 points providing a total of 2700 spectra. Figure 6 shows an example of a few of the spectra obtained along a line in the 50:50 weight proportion of calcite + gypsum mixture. The quantification parameter $\bar{r}_{material}$ was then calculated for each line, phase and proportion, as the average of the $r_{material1\_i}$ and $r_{material2\_i}$ values in the line. The numerical results for some different mixtures can be seen in Tables 3 and 4 (calcite + gypsum) and Tables 5 and 6 (sodium oxalate + potassium nitrate). A representation of the $r_{calcite\_i}$ and r_{gypsum_i} values with different proportions can be seen in Fig. 7. The spectral lines of the different phases used for the analysis are shown in Table 7.

Detection threshold. The results in Tables 3–6 show that all the phases were found in each sample, even when in minor proportions in the 1–99 % mixtures. This would mean that, with a 30 point line, the analyzed phases have a detection threshold down to 1 % with this instrument, and probably lower, depending on the phase. The analysis of the 1–99 % mixtures with XRD and FT Raman showed that these macro techniques were, in most of the cases, incapable of detecting minor quantities of the phases. However, micro Raman was able to detect them, probably due to a trade-off between grain size, spot diameter and the number of acquired points. A very high spot-size/grain-size ratio would allow the detection of several mineral phases at the same time, but would lower the irradiance level and thus the detection threshold. On the other hand, a small ratio allows analysis of the sample at the mineral grain scale but increases the number of necessary points to detect all the mineral phases present in random analysis. The specification of a spot size of 50 μm and a grain size distribution of 250 μm and smaller has proven that this technique is able to detect mineral species simultaneously in many of the points, as well as detecting minor phases
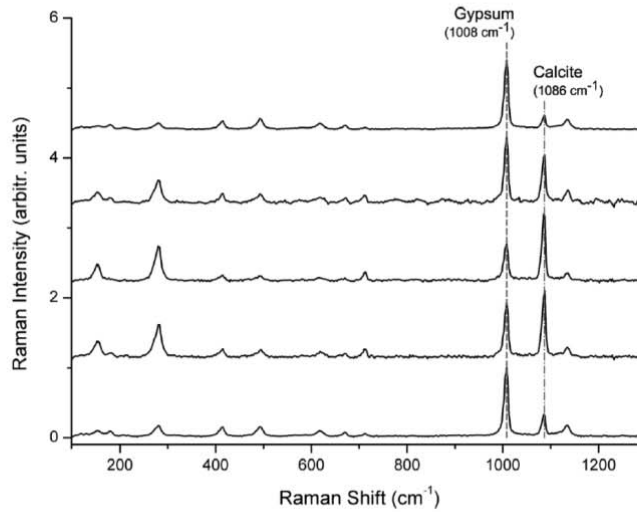
Fig. 6. Some spectra obtained from the 50:50 calcite + gypsum mixture with the RLS instrument. The marked peaks are the ones used for the relative intensity calculation.

Table 3. Quantification parameter $\bar{r}_{calcite}$ values. Calcite (1086 cm^{-1}) + gypsum (1008 cm^{-1}) mixture.

Calcite real proportion (%)	Line 0 (%)	Line 1 (%)	Line 2 (%)	Line 3 (%)	Line 4 (%)	Average (%)	0.95 confidence interval (%)
1	4	1	1	1	0.6	2	1.4
10	11	13	13	9	14	12	1.6
25	28	26	22	16	33	25	5.7
50	60	52	46	51	52	52	4.3
75	75	72	79	76	79	76	2.5
90	92	92	92	89	94	92	1.5
99	97	98	94	97	97	97	1.1

Table 4. Quantification parameter $\bar{r}_{gypsum}$ values. Calcite (1086 cm^{-1}) + gypsum (1008 cm^{-1}) mixture.

Gypsum real proportion (%)	Line 0 (%)	Line 1 (%)	Line 2 (%)	Line 3 (%)	Line 4 (%)	Average (%)	0.95 confidence interval (%)
99	96	99	99	99	99.4	98	1.4
90	89	87	87	91	86	88	1.6
75	72	74	78	84	67	75	5.7
50	40	48	54	49	48	48	4.3
25	25	28	21	24	21	24	2.5
10	8	8	8	11	6	8	1.5
1	3	2	6	3	3	3	1.1

Table 5. Quantification parameter $\bar{r}_{Na-oxalate}$ values for the Na$_2$C$_2$O$_4$ (1456 cm^{-1}) + KNO$_3$ (1046 cm^{-1}) mixture.

Na$_2$C$_2$O$_4$ real proportion (%)	Line 0 (%)	Line 1 (%)	Line 2 (%)	Line 3 (%)	Line 4 (%)	Average (%)	0.95 confidence interval (%)
1	1	1	1	1	0.6	1	0.3
10	12	8	12	12	12	12	1.5
25	25	23	34	27	22	26	4.4
50	55	59	51	48	52	53	3.8
75	81	61	74	72	76	73	6.5
90	92	85	79	92	90	88	4.8
99	98	99	99	99	99	99	0.5

Table 6. Quantification parameter $\bar{r}_{K-Nitrate}$ values for the $Na_2C_2O_4$ (1456 cm^{-1}) + KNO_3 (1046 cm^{-1}) mixture.

KNO_3 real proportion (%)	Line 0 (%)	Line 1 (%)	Line 2 (%)	Line 3 (%)	Line 4 (%)	Average (%)	0.95 confidence interval (%)
99	99	99	99	99	99.4	99	0.3
90	88	92	88	88	88	88	1.5
75	75	77	66	73	78	74	4.4
50	45	41	49	52	48	47	3.8
25	19	39	26	28	24	27	6.5
10	8	15	21	8	10	12	4.8
1	2	1	1	1	1	1	0.5

very intensely at some points in the analysis of about 30 points per mixture (Fig. 7).

It should be noted that the Raman efficiency of individual phases plays an important role as a phase with low Raman scattering efficiency will need longer acquisition times than one with a high Raman scattering efficiency. An extreme case would be a mixture of two materials with very unbalanced Raman scattering efficiencies, in which the most inefficient was at small concentration. In this case, the minor component might not be detected due to the strong Raman response of the other component. This could happen especially during automated RLS operation, where the calculation of the acquisition parameters is based on the signal-to-noise ratio of the highest peak of the reference spectrum.

Quantification of mineral abundance. The results in Tables 3–6 also showed that semi-quantitative analysis can be performed based on the results obtained by RLS, as the quantification parameter $\bar{r}_{material}$ is relatively close to the actual values of the proportion. By representing these values, the calibration curves in Figs. 8 and 9 are obtained. Note the similarity between the different lines analyzed in Tables 3–6. In this case, the average standard deviation of the results of the different lines is 3 % for the gypsum + calcite mixture and 4 % for potassium nitrate + sodium oxalate. This means that the RLS mode of operation is statistically coherent and that one line of 30 points is statistically valid for identifying the mineral phases present in a sample.

When analyzing the samples with macro techniques (FT Raman and XRD), the calibration curves are non-linear (see Figs. 10 and 11 as an example for the calcite-gypsum mixture). This is an expected result due to the different scattering or diffraction efficiencies of the different mineral phases. However, the calibration curves obtained with the RLS, depicted in Fig. 8 and 9 after the representation of the values in Tables 3–6, show a very linear slope of almost unity, with higher deviations for the mixtures with similar abundance (see confidence intervals between the lines in Tables 3–6). The reason for this linear slope is that the RLS instrument analyzes the sample at the mineral grain scale, thus optimizing the integration time at each point to reach a determined level of counts in the CCD. In this way, the differences in Raman efficiency will be compensated by varying the acquisition times at each point. For mixtures containing a high proportion of a phase with lower efficiency, the average integration time is higher than for a mixture with a low proportion of this phase. For example, calcite is a more efficient Raman scatterer than gypsum and Fig. 12 demonstrates that the average total acquisition time is higher for low proportions of calcite than for low proportions of gypsum, and *vice-versa*. To ensure good results, the operation time should be long enough to allow for the compensation of the different Raman efficiencies of the materials. On the other hand, the quantification parameter $\bar{r}_{material}$ in balanced proportions as 50–50 % will show higher deviation from the actual proportion value due to the fact that, in such cases, the most efficient material in the spectrum will generally limit the calculated integration time.

As we have seen, the unitary slope obtained with the RLS instrument could provide this statistical method with a semi-chemometric capability, removing, as a first approach, the need for calibration curves for each possible combination of minerals. This method should also be valid in mixtures in which the main peaks overlap. However, the quantification parameter can be calculated based on other secondary peaks by taking into account the ratio between the main and the secondary peak intensities.

It is important to note that this kind of quantification does not necessarily provide information about the relative abundance of all the minerals present in a sample, especially if there are multiple end-members present. However, the technique is promising and should be further investigated, as it could provide relevant information based only on the data products provided by the RLS instrument.

Conclusions

Raman spectroscopy has the important characteristic of providing non-invasive and non-destructive analyses of many kinds of minerals and biological materials. Long considered as a laboratory technique with a human operator due to technological restrictions, recent developments have permitted the miniaturization of the technique, thus making it applicable for planetary exploration. In order to prepare the system for automated operation, an analogue of the ExoMars rover Sample Preparation and Distribution System has been built at the Unidad Asociada UVa- CSIC- Center of Astrobiology in Valladolid for testing experiments that help understand the possibilities offered within the framework of the ExoMars rover. In particular, the development of the ExoMars RLS simulator has paved
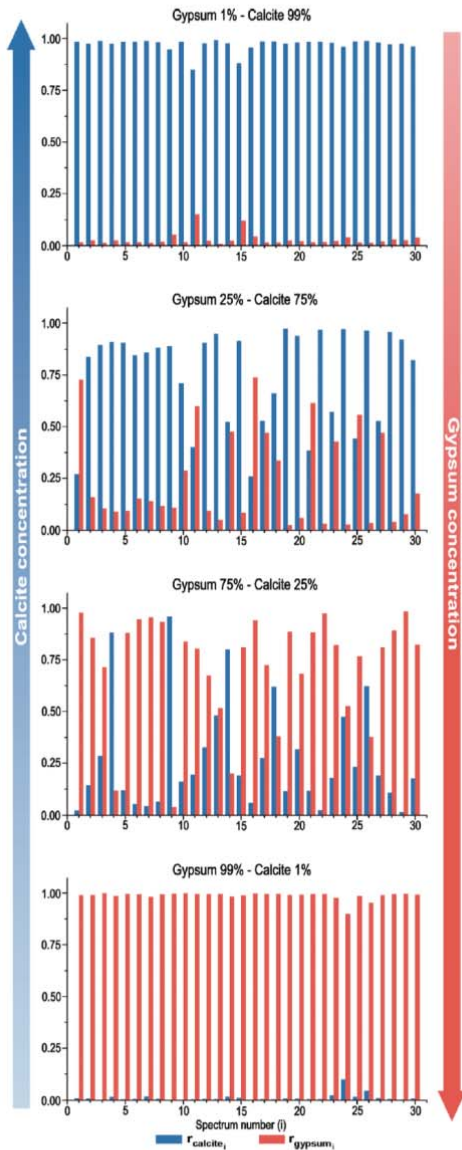
Fig. 7. Quantification parameter $r_{calcite\_i}$ and r_{gypsum_i} values along a line in mixtures with different proportions.

the way for the systematic study of the operation on powdered samples, allowing both the development of the operation mode, and the investigation and understanding of the instrument capabilities at the post-processing state.

The qualitative analysis of natural samples has helped evaluating the performance of the RLS instrument under the conditions imposed by the rover. This instrument is capable of detecting most of the mineral phases present on the sample, even minor ones, with even better detection capabilities than macro techniques such as XRD or FT Raman. Despite the loss of texture and structure due to the crushing process, RLS data products are useful to obtain geologically relevant information, especially when in combination with complementary techniques of the ExoMars rover.

The quantitative analysis of a set of well characterized powdered binary mixtures in different proportions by the RLS instrument has shown that, for these samples, the detection threshold of the instrument goes down to 1 % with 30 random point acquisitions. In addition, a pseudo-chemometry of the abundance of mineral species can be achieved based on the relative intensities between peaks, a method already applied to macro techniques as XRD or FT Raman. The results are promising, as they presented quite linear unitary slopes which provide a good approach for the quantification of the phases present in the studied samples.

Future lines

Despite the level of development achieved with the ExoMars Simulator, this work has to be continued in two directions: the instrument acquisition capabilities to obtain the highest-quality spectra (operation mode), and the development of techniques for obtaining the maximum information from the acquired spectra (data analysis).

The desire to obtain the best possible results highlights the need to develop new algorithms for the operation mode that provide the instrument with a greater intelligence. These new algorithms include the intelligent minimization of fluorescence levels depending on the specific material under study, the detection and treatment of thermolabile samples or the collaborative science with other instruments, so that information obtained by other instruments could be used to improve the performance of the RLS instrument and *vice-versa*.

So far, simulations of the operation mode have been conducted in laboratory conditions. With the addition of a work-in-progress Martian Environment Simulation Cell (Catalá-Espí *et al.*, 2011) to the equipment, the capabilities of the former are being extended and a new and more realistic

Table 7. Spectral lines (main peak corresponding to each material) used for the analysis of the binary mixtures.

	Calcite	Gypsum	Sodium oxalate	Potassium nitrate
Formula	$CaCO_3$	$CaSO_4 \cdot 2H_2O$	$Na_2C_2O_4$	KNO_3
Main peak (cm^{-1})	1086	1008	1456	1046
Reference	(Krishnamurti, 1957)	(Berenblut, 1971)	(Frost, 2004)	(Raju, 1945)

Fig. 8. RLS calibration curves of the calcite + gypsum mixtures, obtained after the representation of the $\bar{r}_{gypsum}$ and $\bar{r}_{calcite}$ values for all the proportions (ranging from 0.01:0.99 to 0.99:0.01) in which both materials were detected.
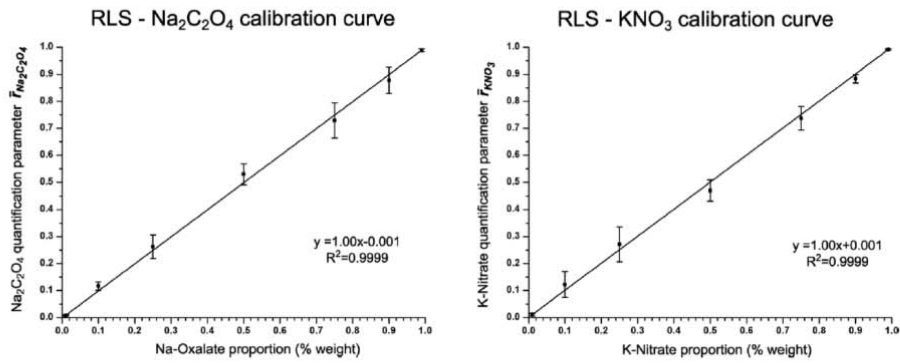
Fig. 9. RLS calibration curves of the potassium nitrate + sodium oxalate mixtures, obtained after the representation of the $\bar{r}_{Na-Oxalate}$ and $\bar{r}_{K-Nitrate}$ values for all the proportions (ranging from 0.01:0.99 to 0.99:0.01) in which both materials were detected.

Fig. 10. XRD calibration curves of the calcite + gypsum mixtures, obtained after the representation of the $\bar{r}_{gypsum}$ and $\bar{r}_{calcite}$ values for all the proportions (ranging from 0.01:0.99 to 0.99:0.01) in which both materials were detected.

732 G. Lopez-Reyes *et al.*
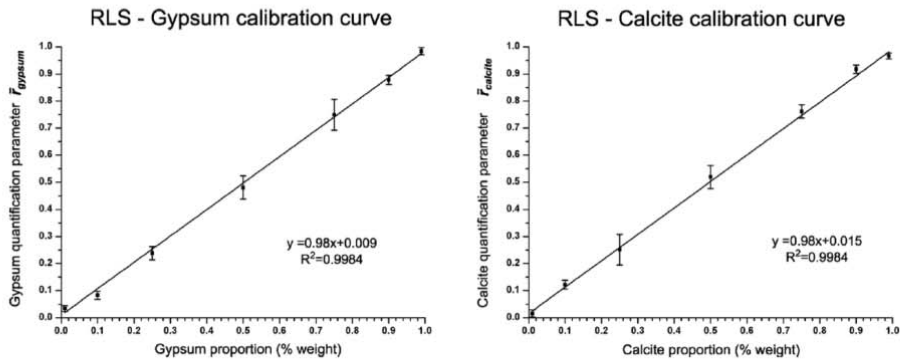

Fig. 11. FT Raman calibration curves of the calcite + gypsum mixtures, obtained after the representation of the $\bar{r}_{gypsum}$ and $\bar{r}_{calcite}$ values for all the proportions (ranging from 0.01:0.99 to 0.99:0.01) in which both materials were detected.
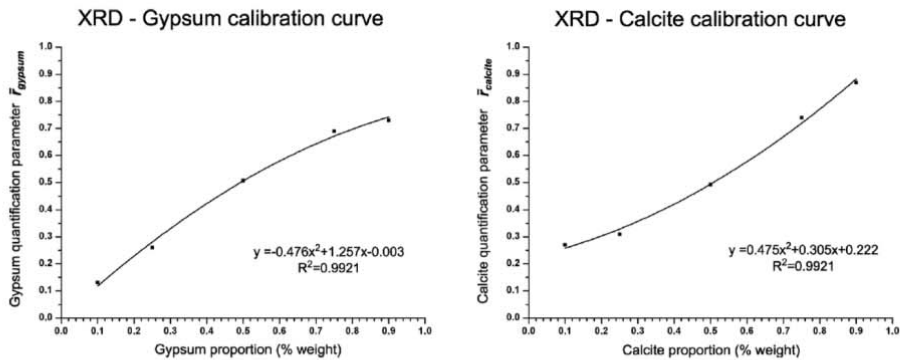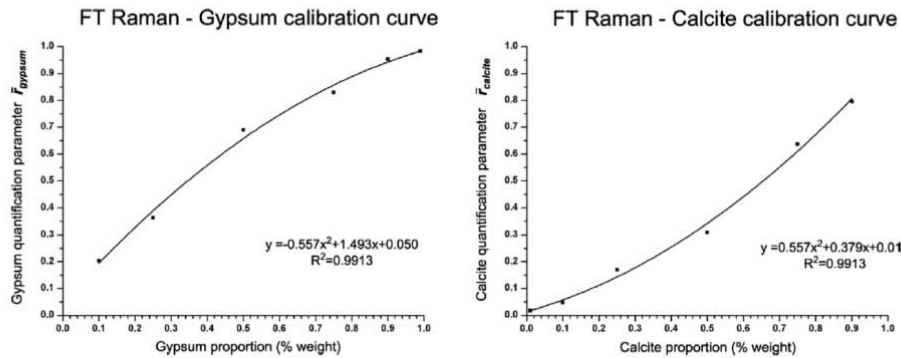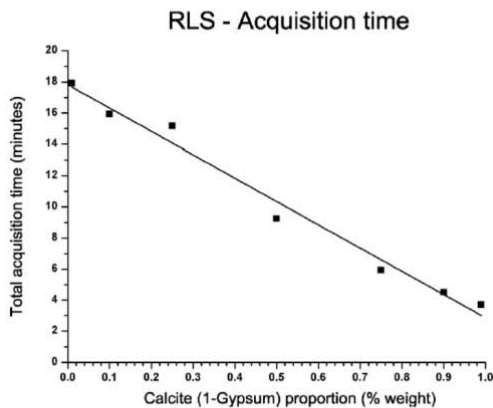

Fig. 12. Total acquisition time for all the spectra in each line of the calcite + gypsum mixture with the RLS instrument.

scientific approach to the autonomous simulation of the operation mode is reached.

Concerning the data analysis, the quantification approach presented in this paper shall be further investigated, especially to apply it on well characterized natural samples. In addition, the application of multivariate techniques such as Partial Least Squares or Artificial Neural Networks is promising for the identification and quantification of minerals and mixtures from Raman spectra. This kind of techniques for the analysis of Raman spectra in the framework of the ExoMars mission is currently under study.

References

Angel, S.M., Gomer, N.R., Sharma, S.K., McKay, C. (2012): Remote Raman spectroscopy for planetary exploration: a review. *Appl. Spectrosc.*, **66**, 137–150.

Berenblut, B.J., Dawson, P., Wilkinson, G.R. (1971): Raman spectrum of gypsum. *Spectrochim. Acta A*, **27**, 1849–1863.

Blake, D.F., Treiman, A.H., Morris, R., Bish, D., Amundsen, H.E.F., Steele, A. (2011): Carbonate cements from the Sverrefjell and Sigurdfjell Volcanos, Svalbard Norway: analogs for Martian Carbonates. *LPSC 42*, 2167.

Bost, N. (2012): Geochemical and mineralogical analysis of Mars analogue materials and the creation of the International Space Analogue Rockstore (ISAR). PhD Thesis, Université d'Orléans, France.

Bost, N., Westall, F., Ramboz, C., Foucher, F., Pullan, D., Meunier, A., Petit, S., Fleischer, I., Klingelhöfer, G., Vago, J.L. (2013): Missions to Mars: Characterisation of Mars analogue rocks for the International Space Analogue Rockstore (ISAR). *Planet. Space Sci*, **82–83**: 113–127.

Catalá-Espí, A. (2010): Operation modes and spectral data treatment in the Raman instrument for ExoMars. Departamento de Astrofisica. Tenerife, Universidad de La Laguna. MSC Astrophysics.

Catalá-Espí, A., Lopez-Reyes, G., Rull, F., Vegas, A. (2011): Design of a small Martian chamber for simulating the RLS-ExoMars operation mode. EPSC Abstracts, **6**, EPSC-DPS2011-708.

Dickinson, S.R. & McGrath, K.M. (2001): Quantitative determination of binary and tertiary calcium carbonate mixtures using powder X-ray diffraction. *Analyst*, **126**(7), 1118–1121.

Edwards, H.G.M., Newton, E.M., Dickensheets, D.L., Wynn-Williams, D.D. (2003): Raman spectroscopic detection of bio-molecular markers from Antarctic materials: evaluation for putative Martian habitats. *Spectrochim. Acta A Mol. Biomol. Spectrosc.*, **59**(10), 2277–2290.

Edwards, H.G.M., Hutchinson, I.B., Ingley, R., Waltham, N.R., Beardsley, S., Dowson, S., Woodward, S. (2011): The search for signatures of early life on Mars: Raman spectroscopy and the Exomars mission. *Spectrosc. Eur.*, **23**(1), 6–15.

ESA, (2012): Exomars project web page at www.esa.int. from http://exploration.esa.int/science-e/www/object/index.cfm?fobjectid=46048.

Foucher, F. & Westall, F. (2009): Investigating the oldest traces of life by AFM/Confocal Raman Spectroscopy: applications for the analysis of Martian Rocks. *in* "Micro-Raman spectroscopy and luminescence studies in the earth and planetary science", A. Gucsik, ed. American Institute of Physics (AIP), New York, CP 1163, 47–58.

Foucher, F., Lopez-Reyes, G., Bost, N., Rull, F., Rüßmann, P., Westall, F. (2012): Effect of the crushing process on Raman analyses: consequences for the Mars 2018 mission. *Geophysical Research Abstracts*, **14**, EGU2012-4318-1.

183

—, —, —, —, —, — (2012b): Effect of grain size distribution on Raman analyses and the consequences for in situ planetary missions. *J. Raman Spectrosc.* (Submitted).

Foucher, F., Lopez-Reyes, G., Bost, N., Rull, F., Rüßmann, P., Westall, F. (2013): Effect of grain size distribution on Raman analyses and the consequences for in situ planetary missions. *J. Raman Spectrosc*, **44**, 916–925.

Frost, R.L. (2004): Raman spectroscopy of natural oxalates. *Anal. Chim. Acta*, **517**, 207–214.

Habermann, D., Neuser, R.D., Richter, D.K. (2000): Quantitative high-resolution spectral analysis of Mn^{2+} in sedimentary calcite. *in* "Cathodoluminescence in the geosciences", M. Pagel, ed. Springer, Berlin, 331–358.

Josset, J.-L., Westall, F., Hofmann, B.A., Spray, J.G., Cockell, C., Kempe, S., Griffiths, A.D., De Sanctis, M.C., Colangeli, L., Koschny, D., Pullan, D., Föllmi, K., Diamond, L., Josset, M., Javaux, E., Esposito, F., Barnes, D. (2012): CLUPI, a high-performance imaging system on the ESA-NASA rover of the 2018 ExoMars mission to discover biofabrics on Mars. *Geophysical Research Abstracts*, **14**, EGU2012–13616.

Kontoyannis, C.G. & Vagenas, N.V. (2000): Calcium carbonate phase analysis using XRD and FT-Raman spectroscopy. *Analyst*, **125**, 251–255.

Kontoyannis, C.G., Orkoula, M., Koutsoukos, P. (1997): Quantitative analysis of sulfated calcium carbonates using Raman spectroscopy and X-ray powder diffraction. *Analyst*, **122**, 33–38.

Krishnamurti, D. (1957): The Raman spectrum of Calcite and its interpretation. *Proc. Indian Natl. Sci. Acad., A*, **46**, 183–202.

Leroi, V., Bibring, J.P., Berthe, M. (2009): Micromega/IR: design and status of a near-infrared spectral microscope for in situ analysis of Mars samples. *Planet. Space Sci.*, **57**, 1068–1075.

Morris, R.V., Ruff, S.W., Gellert, R., Ming, D.W., Arvidson, R.E., Clark, B.C., Golde, D.C., Siebach, K., Klingelhöfer, G., Schröder, C., Fleischer, I., Yen, A.S., Squyres, S.W. (2010): Identification of carbonate-rich outcrops on mars by the spirit rover. *Science*, **329**(5990), 421–424.

Nutman, A.P., Mojzsis, S.J., Friend, C.R.L. (1997): Recognition of ≥3850 Ma water-lain sediments in West Greenland and their significance for the early Archaean Earth. *Geochim. Cosmochim. Acta*, **61**, 2475–2484.

Pacros, A., Haldemann, A., Vago, J., and the ESA ExoMars Project Team. (2009): Overview of the ExoMars payload. *EPSC Abstracts*, **4**, EPSC2009-776.

Raju, M.K. (1945): Raman spectra of mixed crystals. *Proc. Indian Natl. Sci. Acad., A*, **22**, 150–156.

Ruff, S.W., Farmer, J.D., Calvin, W.M., Herkenhoff, K.E., Johnson, J.R., Morris, R.V., Rice, M.S., Arvidson, R.E., Bell, J.F., III, Christensen, P.R., Squyres, S.W. (2011): Characteristics, distribution, origin, and significance of opaline silica observed by the Spirit rover in Gusev crater, Mars. *J. Geophys. Res.*, **116**, E00F23.

Rull, F. (2012): Raman spectroscopy for planetary exploration. Georaman 2012.

Rull, F., Maurice, S., Diaz, E., Tato, C., Pacros, A., and the RLS Team. (2011): The Raman Laser Spectrometer (RLS) on the EXOMARS 2018 Rover Mission. LPSC 42, 2400.

Sansano, A., López, G., Medina, J., Rull, F. (2011): Analysis of Arctic Carbonates Profiles by Raman Spectroscopy using Exomars Raman Laser Spectrometer. EPSC-DPS Joint Meeting 2011. 856 p.

Sharma, S.K., Lucey, P.G., Ghosh, M., Hubble, H.W., Horton, K.A. (2003): Stand-off Raman spectroscopic detection of minerals on planetary surfaces. *Spectrochim. Acta A Mol. Biomol. Spectrosc.*, **59**, 2391–2407.

Squyres, S.W., Arvidson, R.E., Ruff, S., Gellert, R., Morris, R.V., Ming, D.W., Crumpler, L., Farmer, J.D., Marais, D.J., Yen, A., McLennan, S.M., Calvi, W., Vell, J.F., Clark, B.C., Wang, A., McCoy, T.J., Schmidt, M.E., de Souza, P.A. (2008): Detection of silica-rich deposits on mars. *Science*, **320**(5879), 1063–1067.

Steele, A., Amundsen, H.E.F., Fogel, M., Benning, L., Schmitz, N., Conrad, P., Younse, P., Backes, P., AMASE Team. (2011): The Arctic Mars Analogue Svalbard Expedition (AMASE) 2010. LPSC 42, 1588.

Treiman, A.H., Amundsen, H.E.F., Blake, D.F., Bunch, T. (2002): Hydrothermal origin for carbonate globules in Martian meteorite ALH84001: a terrestrial analogue from Spitsbergen (Norway). *Earth Planet. Sci. Lett.*, **204**, 323–332.

Vagenas, N.V. & Kontoyannis, C.G. (2003): A methodology for quantitative determination of minor components in minerals based on FT-Raman spectroscopy – the case of calcite in dolomitic marble. *Vib. Spectrosc.*, **32**, 261–264.

Vago, J., Gardini, B., Baglioni, P., Kminek, G., Gianfiglio, D., and the ExoMars Project Team. (2006): ExoMars: ESA's mission to search for signs of life on the red planet. *LPSC 37*, 1871.

Wang, A., Dhamenincourt, P., Dubessy, J., Guerard, D., Landais, P., Lelaurain, M. (1989): Characterization of graphite alteration in an uranium deposit by micro-Raman spectroscopy, X-ray diffraction, transmission electron microscopy and scanning electron microscopy. *Carbon*, **27**, 209–218.

Wang, A., Haskin, L.A., Lane, A.L., Wdowiak, T.J., Squyres, S.W., Wilson, R.J., Hovland, L.E., Manatt, K.S., Raouf, N., Smith, C.D. (2003): Development of the Mars microbeam Raman spectrometer (MMRS). *J. Geophys. Res.*, **108**, (E1), 5005.

Wang, A., Freeman, J.J., Jolliff, B.L., Chou, I.M. (2006): Sulfates on Mars: a systematic Raman spectroscopic study of hydration states of magnesium sulfates. *Geochim. Cosmochim. Acta*, **70**, 6118–6135.

Westall, F., de Ronde, C.E.J., Southam, G., Grassineau, N., Colas, M., Cockell, C., Lammer, H. (2006): Implications of a 3.472-3.333 Gyr-old subaerial microbial mat from the Barberton greenstone belt, South Africa for the UV environmental conditions on the early Earth. *Philos. Trans. R. Soc. London B*, **361**, 1857–1875.

Westall, F., Cavalazzi, B., Lemelle, L., Marrocchi, Y., Rouzaud, J., Simionovici, A., Salomé, M., Mostefaoui, S., Andreazza, C., Foucher, F., Toporski, J., Jauss, A., Thiel, V., Southam, G., MacLean, L., Wirick, S., Hofmann, A., Meibom, A., Robert, F., Défarge, C. (2011a): Implications of in situ calcification in photosynthesis in a ~3.3 Ga-old microbial biofilm from the Barberton greenstone belt, South Africa. *Earth Planet. Sci. Lett.*, **310**, 468–479.

Westall, F., Foucher, F., Cavalazzi, B., de Vries, S.T., Nijman, W., Pearson, V., Watson, J., Verchovsky, A., Wright, I., Rouzaud, J., Marchesini, D., Anne, S. (2011b): Volcaniclastic habitats for early life on Earth and Mars: a case study from ~3.5 Ga-old rocks from the Pilbara, Australia. *Planet. Space Sci.*, **59**, 1093–1106.

5.3 Multivariate Analysis techniques

The previous section provides a first assessment of the achievable science that can be obtained from the RLS data products, by using standard and simple analytical techniques for the analysis of the spectra. The next logical step is to evaluate more advanced techniques to assess their goodness, and the science return available from them. Basically, applying unsupervised techniques for the analysis of spectra will allow acquiring information and relevant data in an automated and fast way. During operational activities, it will be of paramount importance to obtain scientific relevant data as fast as possible to allow the scientific team to define the next steps in the daily activity plan for the rover itself.

In this section, we summarize the work from [57], paper by Lopez-Reyes et al. in AmMin. This work compares three different multivariate analytical techniques (MVAT), Principal Component Analysis (PCA), Partial Least Squares (PLS) and Artificial Neural Networks (ANN). MVAT, contrary to univariate analysis techniques, work with multiple variables at the same time, providing an overall response of the whole variable system. Concretely, as extracted from the AmMin paper, "PCA and PLS extract latent variables from the system to represent the system in a complexity-reduced variable system. Ideally, the extracted latent variables will respond to physical properties of the model. The difference between PCA and PLS is that, while PCA extracts the variables, PLS also performs a regression on the expected responses of the system for a determined set of inputs. ANNs, on the other hand, is a technique that can model any non-linear function by example-based calibration of computational networks. While ANNs philosophy makes it possible to have direct outputs of the sample presence and abundance, PCA and PLS provide responses based on the latent variables of the system that need to be classified and/or calibrated to extract the mineral phases presence/abundance values". Another difference that differentiates PCA from PLS and ANN is that PCA does not require to train with the 'expected outputs' of the system, which makes it much simpler from the training point of view, removing the risk of over fitting the model.

In order to evaluate these three techniques, we used several sets of spectra from sulfates relevant to Mars, and binary mixtures of those spectra mixed with different proportions. The aim was to quantify the goodness of these techniques, while keeping a relatively controlled scenario regarding the sample mixing.

The results shown in the paper concluded that it will be interesting to further explore the MVAT, as they provided interesting results both for the identification and quantification of the mineral abundances of the binary mixtures under analysis. PCA and PLS showed capability to quantify (and thus identify) to some extent the abundance of mineral phases in the mixtures. ANN provided an identification rate of 100% for pure samples, as well as for both mineral phases in binary mixtures, as long

as the minor was present with in at least 10%. Furthermore, ANN was used to estimate the abundance of the mineral phases in the mixtures with good results. However, this technique requires a more complex training procedure than the other two, plus a pre-processing of the spectra (basically peak detection).

The details on this work are presented in the AmMin paper [57] annexed below.

WHAT LURKS IN THE MARTIAN ROCKS AND SOIL? INVESTIGATIONS OF SULFATES, PHOSPHATES, AND PERCHLORATES

Multivariate analysis of Raman spectra for the identification of sulfates: Implications for ExoMars†

GUILLERMO LOPEZ-REYES[1,*], PABLO SOBRON[2,3,4], CATHERINE LEFEBVRE[2] AND FERNANDO RULL[1]

[1]Unidad Asociada UVA-CSIC-Centro de Astrobiología, C/ Francisco Valles 8, 47151, Boecillo, Spain
[2]Space Science and Technology, Canadian Space Agency, 6767 Rte. de l'Aéroport, St. Hubert, Quebec J3Y 8Y9, Canada
[3]MalaUva Labs, 822 Allen Avenue A, St. Louis, Missouri 63104, U.S.A.
[4]SETI Institute, 189 Bernardo Avenue 100, Mountain View, California 94043, U.S.A.

ABSTRACT

We have built three multivariate analysis mathematical models based on principal component analysis (PCA), partial least squares (PLS), and artificial neural networks (ANNs) to detect sulfate minerals in geological samples from laser Raman spectral data. We have critically assessed the potential of the models to automatically detect and quantify the abundance of selected Ca-, Fe-, Na-, and Mg-sulfates in binary mixtures. Samples were analyzed using a laboratory version of the Raman laser spectrometer (RLS) instrument onboard the European Space Agency 2018 ExoMars mission. Our results show that PCA and PLS, can be used to quantify to some extent the abundance of mineral phases. PCA separated hydrated from dehydrated mixtures and classified mixtures depending on the phase abundances. PLS provided relatively good calibration curves for these mixtures. Upon spectral pre-processing, ANNs provided the most precise qualitative and quantitative results. The detection of mineral phases was 100% accurate for pure samples, as was for binary mixtures where the abundance of mineral phases was >10%. The outputs of the ANN were proportional to the phase abundance of the mixture, thus demonstrating the ability of ANNs to quantify the abundance of different phases without the need for calibration. Taken together, our findings demonstrate that multivariate analysis provides critical qualitative and quantitative information about the studied sulfate minerals.

Keywords: Sulfates, ExoMars, Raman spectroscopy, multivariate analysis, qualitative, quantitative

INTRODUCTION

Laser Raman spectroscopy has been proposed as a powerful tool for the identification of minerals in the context of planetary exploration, including Mars (Sharma et al. 2003; Sobron and Alpers 2013; Sobron et al. 2008; Wang et al. 2003; Wiens et al. 2007), Europa (Angel et al. 2012; Sobron et al. 2013), Venus (Lambert et al. 2010), the Moon (Ling et al. 2009), and asteroids (Kong and Wang 2010). In addition, the feasibility of using laser Raman spectroscopy for the detection of biosignatures in terrestrial analogues to Mars has been demonstrated (e.g., Bower et al. 2013; Dickensheets et al. 2000; Edwards et al. 2003, 2011, 2012; Ellery and Wynn-Williams 2003; Steele et al. 2010; Wynn-Williams and Edwards 2000). A Raman laser spectrometer (RLS) is part of the science instrument payload of the European Space Agency 2018 ExoMars mission; the RLS instrument will target mineralogical and astrobiological investigations on the surface and subsurface of Mars (Rull et al. 2011a, 2011b).

The current concept of operation of the RLS instrument is a raster analysis of crushed drill-core materials (Rull et al. 2011a). In this configuration, the geological and morphological context of the spots analyzed by RLS will be lost, as the crushing stage will preclude correlation between RLS spectra and the imagery acquired by the rover Close-Up Imager CLUPI (Josset et al. 2012). While the synergy between these instruments in the current ExoMars payload configuration has been demonstrated (Lopez-Reyes et al. 2013a), identification of the mineral phases present in the geological targets and quantification of their abundance with RLS will mostly rely on spectral data and not morphology or texture. Therefore, to enable unambiguous identification and quantification of phase abundance, robust spectral processing methods are needed.

Mineral identification using Raman spectroscopy is often performed by comparing acquired spectra to reference spectra available from the literature and different databases, e.g., the RRUFF project database (Downs 2006). Several algorithms have been developed that enable an automated identification of Raman spectra using traditional univariate analysis, i.e., the description of individual variables in a given spectrum (Hermosilla et al. 2013; Kriesten et al. 2008; Perez-Pueyo et al. 2004; Sobron et al. 2008). These algorithms, however, fail to accurately estimate mineral abundance in complex geological samples, although applications for the quantitative analysis of relatively simple mixtures have been proposed (Lopez-Reyes et al. 2013a; Schumacher et al. 2011; Vagenas and Kontoyannis 2003).

Multivariate analysis techniques (MVAT) are statistical techniques that deal with simultaneous measurements on many variables, and aim at understanding the relationships between

* E-mail: guillermo.lopez.reyes@cab.inta-csic.es

these many variables to predict the values of important properties not directly measurable (Johnson and Wichern 2002). Some examples of MVAT applied to the analysis of Raman spectra in the literature show that principal component analysis (PCA) is capable of differentiating mineral species such as carbonates, sulfates, oxides, and silicates in geological samples (Lafuente et al. 2012). Partial least squares (PLS) has been used to determine the quality of biodiesel fuels (Ghesti et al. 2007). Artificial neural networks (ANN) have been designed for the identification and quantification of inorganic salts in water solutions with Raman spectra (Dolenko et al. 2005). Also, combinations and comparisons of these techniques for chemometrical analysis from Raman spectra (mostly qualitative) have been reported (Dorfer et al. 2010; Ishikawa and Gulick 2013; Özbalci et al. 2013).

In this work we evaluate the feasibility of using PCA, PLS, and ANN for the identification and quantification of sulfate salts in binary mixtures through analysis of spectra recorded using a breadboard prototype of the flight Raman laser spectrometer. The use of binary mixtures is a first step to evaluate these techniques, prior to being tested with more complex samples. We utilized a set of pure sulfates synthetized in the laboratory to prepare the binary mixtures. The sulfates we have considered have been proposed as priority targets for astrobiological investigation of Mars (Chou et al. 2013; King and McLennan 2010; Knoll et al. 2005; Wang et al. 2011) because of their association with liquid water on certain locations on Mars and sulfate-reducing bacteria in terrestrial analog environments (Fernandez-Remolar et al. 2012; Nixon et al. 2013; Sanchez-Andrea et al. 2012). Such environments are known to preserve chemical and morphological fossils (Bonny and Jones 2003), thus testifying to the importance of the detection and identification of sulfates and the detailed study of their degree of hydration to address Mars' hydrologic history and potential for habitability.

In terms of identification, Raman spectroscopy is a very powerful technique for the analysis of the mineralogy of terrestrial sulfate samples analog to Mars (Sobron et al. 2009, 2014; Sobron and Alpers 2013), as well as for the detailed characterization of the hydration states of this kind of salts, which is critical for the rigorous interpretation of the hydrologic history of Mars (Chou et al. 2013; Ling et al. 2008; Wang et al. 2006).

MATERIALS AND METHODS

Samples

A set of 17 Raman spectra of sulfates were used as input for calibrating the PCA, PLS, and ANN MVAT models; hereinafter, we will refer to this set of spectra as calibration set. Table 1 lists the materials used to record our set of spectra. The materials were synthesized in the laboratory using standard techniques (Ling et al. 2008)—they were characterized using X-ray diffraction to certify their mineralogical composition. The Raman spectra of these materials were acquired using Raman instrumentation described in (Ling et al. 2008; Wang et al. 2006).

Apart from the calibration set, two more sets of spectra have been defined: the validation and the test sets. A random selection of the three sets of spectra spectra is shown in Figure 1. The calibration set was divided in four subsets: Ca-, Mg-, Fe-, and Na-sulfates. Each of these four groups included sulfates with several degrees of hydration. The validation set consists on a set of binary mixed spectra with different proportions (0.1:0.9, 0.25:0.75, 0.5:0.5, 0.75:0.25, 0.9:0.1) of all mineral pairs of the samples. It was synthetically generated by computing linear combinations of these spectra, parameterized with the expected proportion of the mixture and the cross section of the mixed materials. Randomized noise with realistic amplitude was added to the synthetic spectra to guarantee differentiation.

TABLE 1. Sulfates used for the analysis with multivariate techniques

Hydration state	Mg	Ca	Fe	Na
Anhydrous	Anhydrous Mg-sulfate	Anhydrite	–	Thenardite
1/2 H_2O	–	Bassanite	–	–
1 H_2O	Kieserite	–	Szomolnokite	–
2 H_2O	Sanderite	Gypsum	–	–
3 H_2O	Mg-sulfate tri-hydrate	–	–	–
4 H_2O	Starkeyite	–	Rozenite	–
5 H_2O	Pentahydrate	–	–	–
6 H_2O	Hexahydrate	–	–	–
7 H_2O	Epsomite	–	Melanterite	–
10 H_2O	–	–	–	Glauber's salt
11 H_2O	Meridianiite	–	–	–

Although probably not totally accurate, this set of spectra is expected to behave similarly to weight-proportion mixtures, providing an easy and convenient way to evaluate the models with hundreds of spectra without the actual need of preparing mixtures and acquiring spectra.

The third set of spectra, the test set, included the Raman spectra of powdered mixtures of anhydrite ($CaSO_4$), thenardite ($NaSO_4$), and $MgSO_4$. This set of spectra is used to test the models as well as to assess the goodness of using the computed spectra of the validation set as part of the model calibration/validation procedure. The Raman spectra of anhydrite and $MgSO_4$ show non-overlapping peaks with thenardite, thus facilitating the spectral processing described below. Two subsets of samples (anhydrite + thenardite and thenardite + $MgSO_4$) with a total of 7 samples each were prepared by mixing these materials in the following weight proportions: 0.01:0.99, 0.1:0.9, 0.25:0.75, 0.5:0.5, 0.75:0.25, 0.9:0.1, and 0.99:0.01. These mixtures were analyzed with the RLS instrument, contrary to the calibration and validation spectra. This way, the models independence with respect to the instruments responses could be assessed. Thirty spectra of each
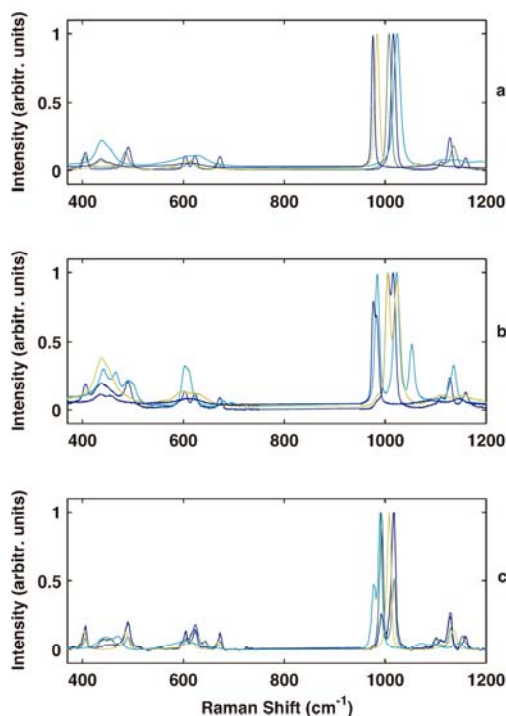
FIGURE 1. Example spectra from the calibration (**a**), validation (**b**), and test sets (**c**).

mixture were acquired and averaged, providing one final spectrum of each mixture. These spectra are representative of the overall behavior of the samples, as the average of different acquisitions of 30 spectra have proved to be almost equal when obtained from an ExoMars-type powdered sample with the RLS instrument (Lopez-Reyes et al. 2013a).

Instrument and Raman spectroscopy

In its current configuration, the ExoMars rover will crush the subsurface-drilled samples and provide the instruments a flattened surface of the powdered sample for analysis. The RLS instrument, in its baseline mode operation, will analyze from 20 to 40 points of each sample, with a 50 μm spot size and an irradiance level of 0.6–1.2 kW/cm² of a 532 nm continuous wave laser (Rull et al. 2011a). To test the analytical capabilities of the instrument in an operation-like environment, including fully automated analysis on powdered samples, a RLS ExoMars Simulator has been developed at the University of Valladolid–CSIC–Center of Astrobiology Associated Unit ERICA (Foucher et al. 2012; Lopez-Reyes et al. 2013a; Rull et al. 2011c). The Simulator was used to acquire the test set spectra from the mixtures with different phase abundances.

To improve the accuracy of the analytical models described below, and given that the aim of the models is to distinguish among different types of sulfates, only spectral regions that are relevant to sulfates were considered. These spectral regions are: (1) the sulfate symmetric stretching v_1 (950 to 1100 cm⁻¹); (2) the sulfate asymmetric stretching v_3 (1100 to 1220 cm⁻¹); (3) the sulfate symmetric and asymmetric bending v_2 and v_4, respectively (100–750 cm⁻¹); (4) the water bending (1600–1700 cm⁻¹); and (5) the water and OH stretching (2800–3800 cm⁻¹). The spectra were baseline-corrected to remove background contributions (e.g., fluorescence), and normalized in intensity in a way such the maximum peak intensity is 1. In our case, the separation between consecutive points of the spectrum is 0.5 cm⁻¹, which provides a total number of variables of about 4000 (each variable corresponding to the intensity at a determined wavenumber of the selected spectral regions), which is the size of the input data to the different models.

MVAT

This section describes the MVAT that have been used in this work: principal component analysis (PCA), partial least-squares regression (PLS), and artificial neural networks (ANNs). PCA and PLS extract latent variables from the system to represent the system in a complexity-reduced variable system. Ideally, the extracted latent variables respond to physical properties of the model. The difference between PCA and PLS is that, while PCA extracts the variables, PLS also performs a regression on the expected responses of the system for a determined set of inputs. ANNs, on the other hand, is a technique that can model any non-linear function by example-based calibration of computational networks. While ANNs philosophy makes it possible to have direct outputs of the sample presence and abundance, PCA and PLS provide responses based on the latent variables of the system that need to be classified and/or calibrated to extract the mineral phases presence/abundance values. The scope of this work does not include the classification step for PCA and PLS responses, but it aims at evaluating the techniques ability to differentiate among such samples. All these techniques are described below in more detail, and were implemented for this work using Mathworks MATLAB.

Principal component analysis (PCA)

PCA is a multivariate technique that computes the variance-covariance structure of a set of variables through a few linear combinations of them, with the objective of reducing the number of data variables (Johnson and Wichern 2002). PCA calculates new variables called principal components (PCs) as linear combinations of the original variables. All the principal components are orthogonal to each other, so the variables in the principal component space do not provide redundant information. The principal components as a whole form an orthogonal basis for the space of the data, which can thus be represented in this new space. There is an infinite number of ways to construct an orthogonal basis to represent the data, so PCA calculates the PCs taking into account that they are uncorrelated among them, while maximizing their variance with coefficient vectors of unit length [as the variance can easily be increased by multiplying by a constant (Johnson and Wichern 2002)]. In other words, the first PC is calculated, among all the possibilities, as the single axis (linear combination of the original variables) in the space that provides the greatest variance by any projection of the data on that axis. The second PC is another single axis that provides the second greatest variance by any projection of the data, and that is orthogonal to the first PC. The third will be also perpendicular to the previous two, and so on. This way, the set of PCs conforms a new set of coordinates which can represent the original data, where most of the

variance of the system can be explained with only a few of the PCs. This is due to the fact that, usually, many of the variables of a system are highly correlated, making it possible to remove some of them and still have a variable set which can account for most of the variability of the system.

The PCA model was trained with the 17 spectra from the pure sulfates (calibration set), and then applied to the validation and test spectra sets.

Partial least squares (PLS)

Partial least squares (PLS) is a common term for a family of multivariate modeling methods that appeared in the 1980s to solve problems in social sciences (e.g., Wold et al. 1983), but that can be applied to a large variety of modeling problems (Martens and Naes 1992), including Raman spectroscopy, as discussed above. The underlying assumption of all PLS methods is, as for PCA, that the observed data are generated by a system which is driven by a small number of latent (not directly observed or measured) variables, called components. In addition, PLS performs a regression of the expected responses for each set of input variables (observations). Thus, this technique is some kind of combination of PCA and linear regression: PLS creates orthogonal score vectors (the latent vectors or components) by maximizing the covariance between two different blocks of variables (Rosipal and Krämer 2006), which correspond to the input variables (predictor variables or observations) and the expected responses (predicted variables or predictions). The higher the number of computed components, the better the model fits the system. In our work, the SIMPLS algorithm (de Jong 1993) has been used, which calculates the PLS factors directly as linear combinations of the original variables. With SIMPLS, the PLS factors are determined such as to maximize the covariance between the predictor variables and the expected responses, while obeying certain orthogonality and normalization restrictions (de Jong 1993).

The responses chosen for the analysis of Raman spectra of sulfates were the weight atomic fraction of the sulfate cation (calcium, magnesium, iron, and sodium), as well as the weight ratio of the water bound to the sulfate molecule, providing a set of five responses. These responses were chosen to reflect the composition of the samples, thus providing the model with a physical basis for the regression of the input variables. Given that the shifts in the Raman bands are produced as a consequence of the frequency shifts associated to the cation-sulfate group oscillators, the cation ratio is the subjacent physical property that we choose as a response in our model. The spectra were baseline-removed, normalized with respect to the maximum peak height, mean centered and scaled to avoid biasing the model.

The model was trained with the spectra from the 17 pure salts (calibration set). As in every model where input variables are regressed to expected responses, there is a risk of over-fitting. To minimize this effect, we optimized our model using a leave-one-out cross-validation method with the validation and test spectra sets, as outlined in the Results and Discussion section.

Artificial neural network

An artificial neural network (ANN) is a mathematical procedure for transforming inputs into desired outputs using highly connected networks of relatively simple processing units called neurons (Johnson and Wichern 2002). Each neuron performs a mathematical operation that produces an output which is a function (usually non-linear) of a series of biased and weighted inputs coming from other neurons. The underlying principle is that neural networks be modeled after the neural activity in the human brain, in which the interconnection of very simple functional units (the neurons) can solve many complex and non-linear problems in a very fast way. Thus, ANNs consist on parallel computational models comprised of densely interconnected adaptive processing units (the neurons). These networks provide a tool for modeling nonlinear static or dynamic systems, making use of their adaptive nature, based on "learning by example." This feature makes this kind of computational models very appealing in applications in which the understanding of the problem is little or incomplete, but where calibration data are readily available (Hassoun 1995).

An ANN can provide very fast outputs, as it only has to compute a limited number of very simple operations that are easily parallelizable. However, the design and calibration of the network can be a hard task. ANNs are set in layers of interconnected networks, in which the first layer has as many neurons as inputs in the system, and the output layer has as many neurons as required outputs. All intermediate layers are called hidden layers, and can have any number of neurons (Johnson and Wichern 2002).

For our design, many different architectures were trained and evaluated to obtain the network with the best performance. In the end, a three-layer network with 33 neurons on the hidden layer was chosen. The input layer was configured with 33 neurons each corresponding to determined spectral positions. The neurons

where configured with log-sigmoid transfer functions. The output layer consisted on 17 outputs, each corresponding to one of the sulfates (where each output should take values between 0 and 1, proportional to the abundance of the sample).

As suggested by (Koujelev et al. 2010), and to improve the ANN model performance, only a selection of spectral positions is fed into the network as input; the v_1 peak positions of all the sulfates, as well as the most intense non-overlapping secondary peaks of the pure samples were selected as inputs. This allows the definition of a predefined set of spectral positions that will be extracted from the input spectra and fed to the network. To determine these peaks, the input spectra to the ANN have to be pre-preprocessed so that only the most intense peaks have non-zero intensities, as depicted in Figure 2: from all the inputs to the model, those below a determined threshold will be set to 0. The definition of this threshold depends on the sample noise and can be decided for each spectrum. In addition, a deconvolution of peaks is needed for mixtures where the peaks are partially overlapped.

The calibration process was performed using a Levenberg-Marquardt back-propagation algorithm (Levenberg 1944; Marquardt 1963) that used spectra of pure sulfates (calibration set) plus some spectra of mixtures with 0.25:0.75, 0.5:0.5, and 0.75:0.25 proportions. The network was trained to provide outputs proportional to the abundance of each sulfate. To avoid over-fitting, the early-stopping technique with the validation set was used. This consists in stopping the iterative calibration process when the output errors for the validation set increase. Finally, to test the network in a more representative scenario, the test set from the RLS instrument was fed to the network.

RESULTS AND DISCUSSION

Principal component analysis (PCA)

The calibration of the PCA model showed that the first three components PC1, PC2, and PC3 explain more than 80% of the variance of the calibration data set, and 90% if the first five components are considered (Fig. 3). This means that 90% of the variance of the system is explained with only five of the calculated orthogonal variables. A dendrogram showing the interconnections for the three PCs model is depicted in Figure 4. Models with higher number of components (up to 10) do not present better separation among different cations, and worsen the discrimination between low- and high-hydration sulfates. In addition, the lower the number of components, the more
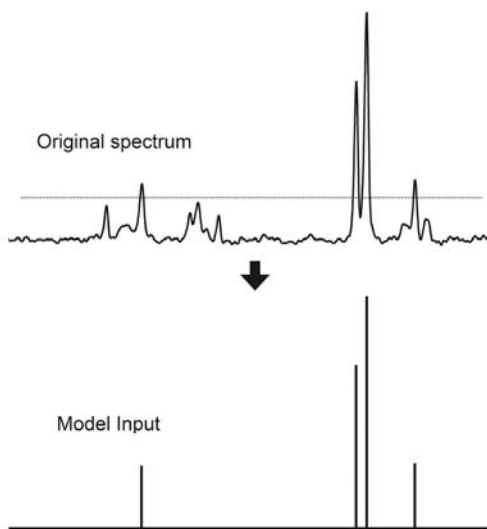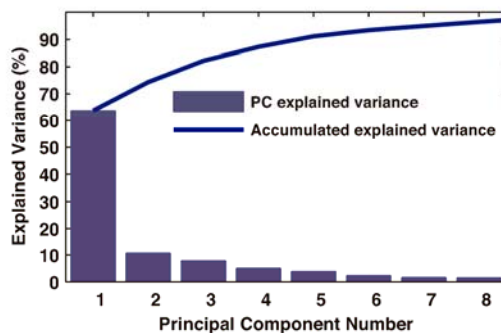

FIGURE 3. Explained variance of the calibration data set (pure spectra of sulfates) after PCA analysis.
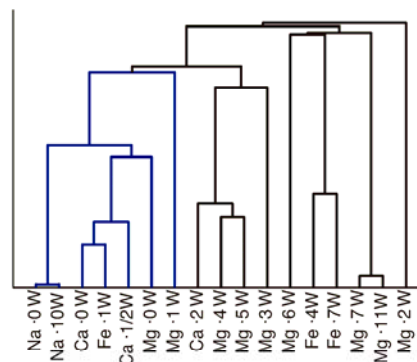

FIGURE 4. Dendrogram showing the classification of the PCA model with 3 PCs. As it can be seen, low-hydrated sulfates are contained in the blue branch of the dendrogram.

general the model can be, so the 3-PC model was selected as the optimal one. The representation of the scores of the calibration samples in the new variable system (with only the first two components, for representation convenience) is displayed in Figure 5 as circles. From the validation set, only the calculated scores for mixed spectra of sulfates of the same cation in 50:50 proportions (worst case) are plotted in Figure 5 as triangles. The mixed spectra contain sulfates with different degree of hydration. The scores in Figure 5 show how PCA succeeds in separating, mostly along PC1, the low-hydration mixtures from the high-hydrated ones, although it fails to distinguish among different types of cations. For example, the model scores the mixture of Fe-sulfates with four and seven water molecules between those two elements, and the same happens with the Ca-sulfates. However, the model fails to correctly separate the Mg-sulfates by hydration level. Therefore, the general conclusion is that PC1 can only be used to separate low hydration from high-hydration sulfates, but not to distinguish among different hydration states of same-cation sulfates.

While no additional direct associations between the principal components and the physical properties of the different molecules
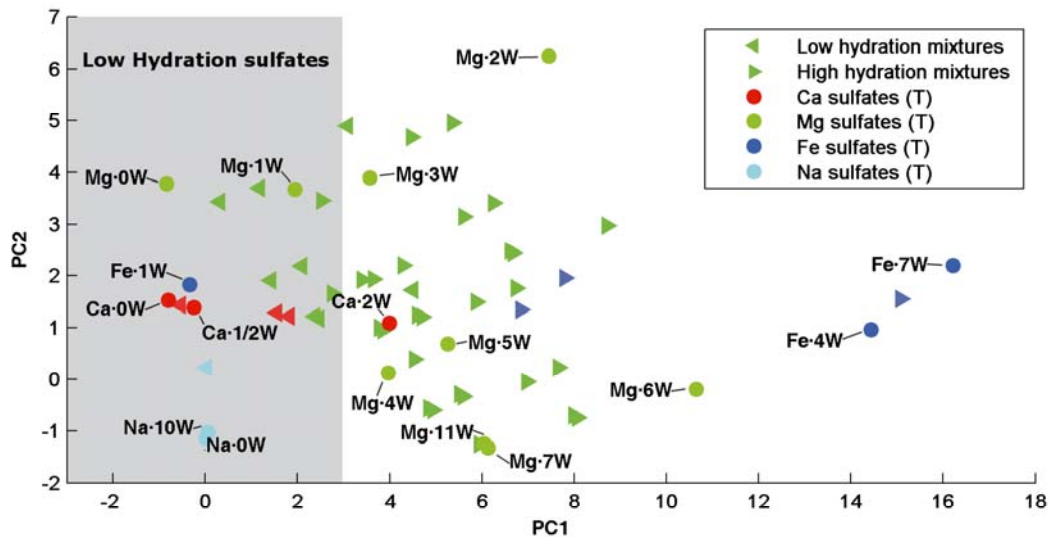

FIGURE 2. Spectrum pre-processing example for the ANN model.

FIGURE 5. Representation of the scores for the calibration set (pure sulfates, circle) and 50:50 mixtures of the sulfates of the same cation from the validation set (triangles). Different colors represent different cations. It can be seen how the low-hydration samples and mixtures are found with lower values of PC1.

(as for example the cation ratio) can be inferred, the PCA analysis of sulfates can provide useful information when representing the PC1–PC2 scores. For example, the representation of the scores of the test spectra set is depicted in Figure 6 (for graphical simplicity, only two PCs are represented). This figure shows how the mixtures are placed between the pure components depending on their relative abundance: the higher the abundance of a sulfate of the mixture, the closer to the corresponding pure sulfate score. This would mean that some kind of quantification could be possible based on the PCA model, even when only trained with pure Raman spectra.
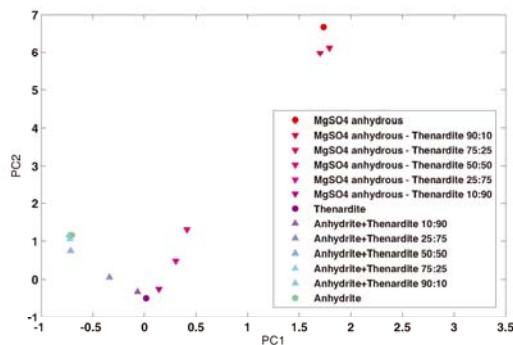

FIGURE 6. Scores for the averaged spectra of the different mixtures acquired with the ExoMars Raman instrument. The scores are placed along lines between the end-members at a distance dependent on the mixture concentration.

Partial least squares (PLS)

The PLS regression of the pure sulfates spectra to their expected weight cation atomic fractions and water presence was performed for different numbers of components. To decide the optimal number of components, the mean squared error (MSE) of the predictions for the calibration, validation, and tests sets was calculated. Figure 7 shows some spectra from the calibration set and the corresponding loadings, where the energies that have the greatest effect on the PLS predictions can be observed. The application of the model to the spectra sets yielded the prediction errors shown in Figure 8a. As expected, the prediction error for the calibration samples (blue line) tends to 0 with increasing number of components, as the model is fitted better. The values of the error for the regression of the test set of spectra (black lines) indicates that the best fitting (minimum prediction error) occurs for a 12-component PLS model. As a general rule, lower numbers of components imply a more general response of the model. Since the spectra of the natural samples were acquired with a different experimental setup than the spectra of the calibration and validation sets, we interpret the 13-component model (where the minimum prediction error for the validation set is found—see red line) as one in which the 13th component accounts for the spectrometer response. Thus, the 12-component model is considered as the optimum for our set of sulfates spectra; more than 98% of the variance of the system can be explained with this model (Fig. 8b), apparently with no over-fitting.

The average absolute prediction error values for the calibration, validation, and test sets are presented in Table 2 for the 12-component model. This table shows how the average error is close to 0 for the calibration and validation sets, while the
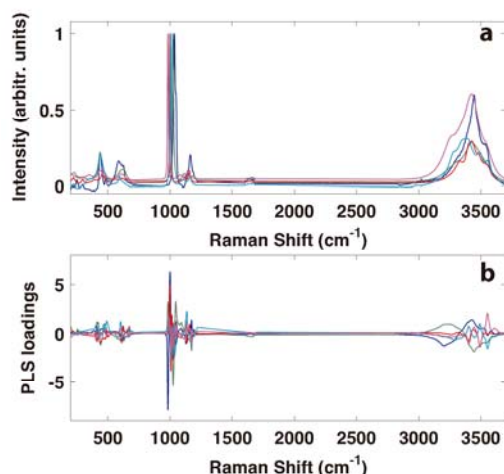
TABLE 2. Average prediction error values with the 12-component PLS model

Response	Calibration set (pure sulfates)	Validation set (Mixed spectra)	Test set (Anhydrite + Thenardite)	Test set (Thenardite + $MgSO_4$)
Hydration ratio	-4.5^{-17}	-0.0087	0.0814	0.0509
Ca-ratio	3.3^{-17}	0.0027	-0.0249	0.0046
Mg-ratio	-1.9^{-17}	-0.0061	-0.0288	-0.0625
Fe-ratio	-1.3^{-17}	0.0021	-0.0737	-0.0211
Na-ratio	-8.1^{-18}	0.0075	0.1123	0.0962
Average	-1.1^{-17}	-0.0004	0.0132	0.0136

TABLE 3. RMSEP values (error variance) with the 12-component PLS model

Response	Calibration set (pure sulfates)	Validation set (Mixed spectra)	Test set (Anhydrite + Thenardite)	Test set (Thenardite + $MgSO_4$)
Hydration ratio	0.0788	0.0821	0.0901	0.0838
Ca-ratio	0.0072	0.0212	0.0327	0.0178
Mg-ratio	0.0109	0.0193	0.0356	0.0639
Fe-ratio	0.0134	0.0227	0.0897	0.0925
Na-ratio	0.0269	0.0236	0.1368	0.1280
Average	0.0274	0.0338	0.0770	0.0772

FIGURE 7. Some spectra from the calibration set (**a**) and corresponding PLS loadings (**b**).

TABLE 4. Correlation values (in %) for the Calculated vs. Expected responses with the 12-component PLS model

Response	Calibration set (pure sulfates)	Validation set (Mixed spectra)	Test set (Anhydrite + Thenardite)	Test set (Thenardite + $MgSO_4$)
Hydration ratio	92.9	89.5	–	–
Ca-ratio	99.7	97.9	98.1	
Mg-ratio	98.7	96.2	–	98.4
Fe-ratio	98.9	97.1	–	–
Na-ratio	93.9	93.3	98.9	98.8
Average	96.8	94.8	98.5	98.6

model prediction is biased in the prediction values for the test set spectra. The root mean square error of prediction (RMSEP) for each test is presented in Table 3. This value represents the prediction error deviation, to compare the prediction accuracy between the different spectra sets. In addition, the predicted vs. expected responses correlations are shown in Table 4, and the responses represented in Figure 9. This figure presents the expected vs. calculated cation ratios for each spectrum. As each spectrum is represented as one individual point in the graph, the
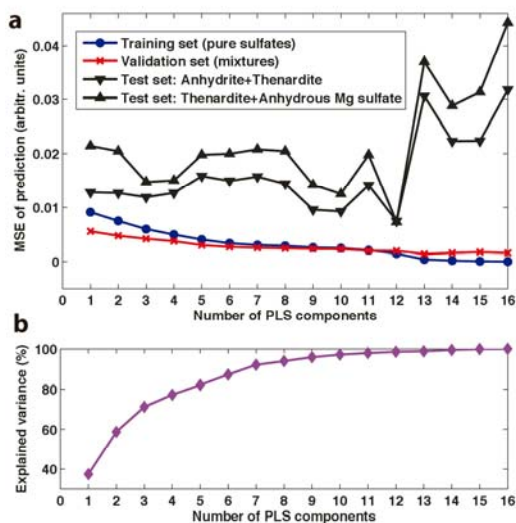
estimation error is defined by the y-axis displacement from the expected value. Thus, the estimation error is 0 when the point is placed on the line with unitary slope ($y = x$). These data show that a correlation between the predicted and the expected responses is present with this model for all the calibration, validation, and test spectra sets.

The RMSEP results for the calibration set in Table 3 show that the model has the lower prediction capabilities for the hydration response, which can be explained by the influence of the many close-to-zero values of these responses, as shown in the plot of the calibration responses in Figure 9a. It is important to note how the results of the test samples provide too high values of the iron cation response, especially in Figure 9d, when it should always be 0.

For the validation spectra, the prediction accuracy is lower (higher RMSEP) than for the calibration spectra, as expected, which is also reflected in its lower correlation and higher prediction bias. The model behavior is worse for well-balanced mixtures (0.5:0.5) than for mixtures with unbalanced proportions, as the model was trained with pure samples only. Most of the outlier points observed in Figure 9b belong to the mixtures in this proportion. The values in Table 4 for the validation set correspond to the averaged correlation for all the proportions, but these improve between 0.5 and 1% when the 0.5:0.5 mixture is not considered.

The results for the test spectra show better correlations than the calibration set in the Na-sulfates case, although the RMSEP
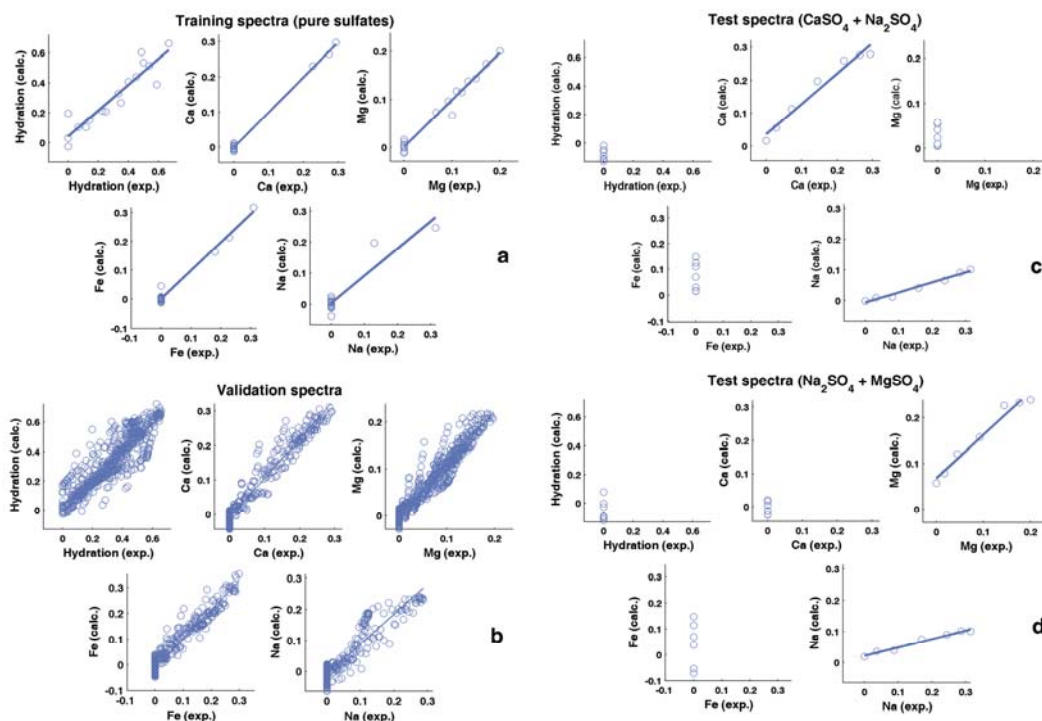

FIGURE 8. MSE of prediction (**a**) and accumulated explained variance (**b**) obtained by the models with different numbers of components.

FIGURE 9. Calculated vs. Expected responses (cation weight ratio) for the calibration (**a**), validation (**b**) and test (**c, d**) spectra for the 12-component PLS model.

value is much higher than for the calibration and validation sets. This can be explained by the fact that the correlation value for the calibration samples is biased with all the spectra from the rest of the sulfates, which are not represented with the test samples. This can be readily observed in Figures 9a and 9b vs. Figures 9c and 9d; the slopes of the curves are more or less unitary for the calibration and validation tests, which indicates a certain accuracy of the model (and which is reflected in similar RMSEP values and low bias), while the test sets show a non-unitary slope linear correlation between the expected and calculated sodium response (with a much higher RMSEP value and prediction bias). This implies that the 12-component PLS model seems to fail to directly predict hydration states and cation abundances for this case. However, it still provides linear calibration curves that could be used to compute these values.

Artificial neural networks (ANNs)

The calibration process of our ANN consisted on providing outputs proportional to the abundance of the materials. As ANNs can model any non-linear function, this seemed to be the most convenient way to do it, contrary to PLS, where the underlying physical principles tried to be modeled.

To evaluate the identification accuracy of the ANN, we established the criterion to consider that a sulfate is detected when the corresponding output is higher than a determined threshold. For

this network, this threshold was set to 0.06. As the output of the network ranges from 0 to 1, this value will be the theoretically lowest detection threshold for this model (i.e., no sulfates will be detected below concentrations of 6%).

Under these premises, the calibration set spectra were detected with 100% accuracy with this ANN. Furthermore, the major phase present in all the samples of validation and test sets was also detected in 100% of the cases. Both phases of the mixtures were detected with 100% accuracy only when the minor phase was present with at least 10% abundance. In other words, our ANN detects 100% of the minerals present in binary mixtures with proportions ranging from 10:90 to 90:10. This implies that the ANN model provides a robust system for the qualitative detection of sulfates in this kind of mixtures, with a detection threshold for minor phases of around 10%, even for spectra acquired with a different hardware setup than the calibration spectra.

The representation of the ANN outputs for spectra of mixtures with different proportions (validation and test sets) shows that the network outputs also provide information on the relative abundance of the materials. As part of the validation set was used to train the model, the results for this set proved very good also in terms of quantification of mineral abundance, as expected. More interesting are the results for the test spectra, which are shown in Figures 10 and 11, where the estimated concentration

193

values from the ANN with respect to the expected ones are represented. These results show that the model accuracy might somehow depend on the samples (e.g., the results for the mixture in Fig. 11 show that the Mg-sulfate concentration tends to be underestimated for mixtures where it is the major component). However, a certain degree of correlation between the modeled values and the actual abundances of the mixtures exists. This is an interesting result, especially bearing in mind that this is true for the test set spectra, while the model was trained with pure spectra obtained with a different spectrometer (calibration set) and computed spectra of mixtures calculated from those samples.

The representation of the maximum and minimum values of the outputs of the ANN, which do not correspond to the minerals present in the spectra can be seen in Figures 10b and 11b. The representation of these values, which should always be 0, is of relevance to show that, with a threshold of 6%, no false identifications are obtained. The conclusion is that the use of ANNs looks promising for providing robust and reliable results under the described premises, not only for identifying the phases present in binary mixtures of sulfates, but also to provide some kind of quantification of their abundance.

MVAT comparative

PCA, PLS, and ANN models have been trained based only in 17 spectra of pure sulfates and mixed spectra computed as linear combinations of those. This procedure has allowed evaluating these analytical techniques without the need to actually prepare all possible combinations of samples for the calibration of the models.

The analysis of selected regions from the Raman spectra implies the analysis of several thousands of variables at the same time. To deal with this amount of information, PCA and PLS calculate new sets of orthogonal variables as linear combinations of the original ones. PLS then regresses these variables to expected responses, while PCA does not. These new variables correspond to latent variables which ideally should be directly related to physical properties of the system (e.g., the degree of hydration of a sulfate). However, this is not always the case. This is probably due to the non-linear nature of the Raman emission, which PCA and PLS try to model with linear processing. ANNs, on the other hand, can provide non-linear transfer functions. Thus, they might be more adequate for the modeling of non-linear effects (as the Raman emission). However, ANNs require a relatively low number of input variables to provide any relevant results. This technique does not perform a reduction to latent variables on its own, so it has been made by only inputting the most relevant spectral positions (corresponding to the wavenumbers of the most representative peaks of the calibration spectra).

We have shown the ability of these MVAT models of providing useful qualitative and even quantitative information for simple binary mixtures, even when the calibration and testing were performed with spectra recorded with different hardware setups. As discussed in the previous sections, PCA separated low- from high-hydrated sulfates, and also somewhat classified samples of mixtures depending on their relative abundance. PLS model outputs presented good correlations to the expected responses. However, although well correlated, the responses in some cases were relatively far from the expected values. The
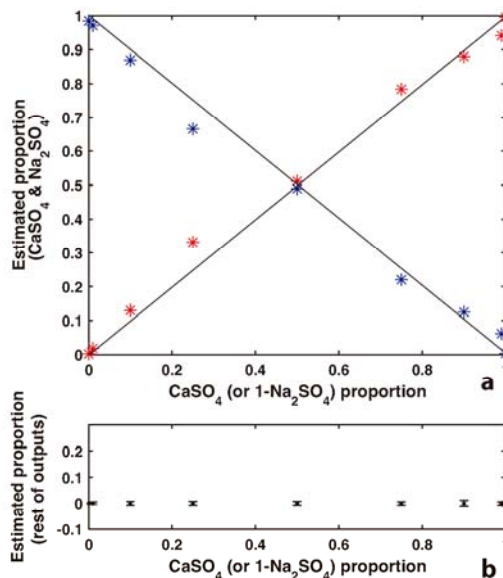

FIGURE 10. ANN outputs for the test set. Outputs corresponding to anhydrite ($CaSO_4$ = red) and thenardite (Na_2SO_4 = blue) (a). Outputs corresponding to the rest of sulfates (b). Vertical axis represents the proportion estimated by the ANN, while horizontal axis represents the expected (known) proportion.

conclusion is that PCA and PLS provided a classification method that needs a previous calibration for the sample under analysis, and a classification method. On the other hand, the ANN model outputs directly provided the abundance of the corresponding salt, in addition to a 100% qualitative detection for mixtures with abundances as low as 10%.

To overcome the various limitations of these MVATs, a synergy between them might be interesting to improve the overall performance of the models. Some classifiers for the qualitative analysis of minerals have been proposed based on integrated PCA and ANN models (Dorfer et al. 2010; Ishikawa and Gulick 2013). Future research will thus focus in developing models for the quantification of mineral abundances that integrate different MVATs.

Implications for ExoMars

ExoMars' RLS instrument will determine the structural and compositional features of materials in rocks and soils at the surface and subsurface of Mars. The ExoMars samples will be collected by a drill, then crushed and delivered to a suite of instruments located in the rover's analytical laboratory, where the RLS instrument sits. A crushing station will provide homogenized powdered samples that will likely feature complex mixtures of mineral phases.

Fast, robust, unsupervised RLS data processing tools able to interpret the intricate spectra that will be obtained from martian samples would benefit ExoMars mission operations in that they
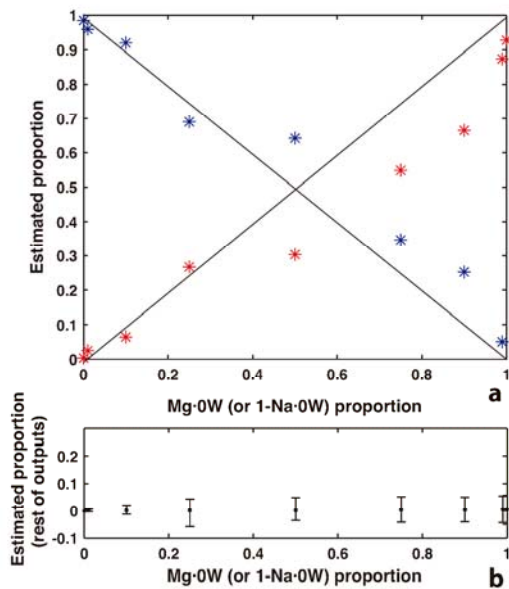
FIGURE 11. ANN outputs for the test set (mixtures in several proportions from RLS). Outputs corresponding to thenardite (Na_2SO_4 = blue) and Mg-sulfate ($MgSO_4$ = red) (a). Outputs corresponding to the rest of sulfates (b). Vertical axis represents the proportion estimated by the ANN, while horizontal axis represents the expected (known) proportion.

may directly support the daily tactical operations of the rover. The different multivariate analysis techniques methodologies—PCA, PLS, and ANN—we have discussed here promise to provide an efficient way to process RLS data during the ExoMars mission.

Future work in this direction will focus on exploring the capabilities of these methodologies to evaluate more complex mixtures; these will include additional synthetic sulfates, oxides, clays, phyllosilicates, carbonates, perchlorates, as well as natural samples.

We will carry out this research using the RLS ExoMars Simulator we have developed at the University of Valladolid-Centro de Astrobiologia and the different RLS models and prototypes that we are developing with the Spanish National Institute for Aerospace Technology (INTA).

REFERENCES CITED

Angel, S.M., Gomer, N.R., Sharma, S.K., and McKay, C. (2012) Remote Raman spectroscopy for planetary exploration: A review. Applied Spectroscopy, 66, 137–150.

Bonny, S., and Jones, B. (2003) Microbes and mineral precipitation, Miette Hot Springs, Jasper National Park, Alberta, Canada. Canadian Journal of Earth Sciences, 40, 1483–1500.

Bower, D.M., Steele, A., Fries, M.D., and Kater, L. (2013) Micro Raman spectroscopy of carbonaceous material in microfossils and meteorites: Improving a method for life detection. Astrobiology, 13, 103–113.

Chou, I.M., Seal, R.R., and Wang, A. (2013) The stability of sulfate and hydrated sulfate minerals near ambient conditions and their significance in environmental and planetary sciences. Journal of Asian Earth Sciences, 62, 734–758.

de Jong, S. (1993) SIMPLS: An alternative approach to partial least squares regression. Chemometrics and Intelligent Laboratory Systems, 18, 251–263.

Dickensheets, D.L., Wynn-Williams, D.D., Edwards, H.G.M., Schoen, C., Crowder, C., and Newton, E.M. (2000) A novel miniature confocal microscope/Raman spectrometer system for biomolecular analysis on future Mars missions after Antarctic trials. Journal of Raman Spectroscopy, 31, 633–635.

Dolenko, T.A., Burikov, S.A., and Sugonjaev, A.V. (2005) Neural network technologies in Raman spectroscopy of water solutions of inorganic salts. In H.J. Byrne, E. Lewis, B.D. MacCraith, E. McGlynn, J.A. McLaughlin, G.D. Osullivan, A.G. Ryder, and J.E. Walsh, Eds., Opto-Ireland 2005: Optical Sensing and Spectroscopy, 5826, 298–305.

Dorfer, T., Schumacher, W., Tarcea, N., Schmitt, M., and Popp, J. (2010) Quantitative mineral analysis using Raman spectroscopy and chemometric techniques. Journal of Raman Spectroscopy, 41, 684–689.

Downs, R.T. (2006) The RRUFF Project: an integrated study of the chemistry, crystallography, Raman and infrared spectroscopy of minerals. Program and Abstracts of the 19th General Meeting of the International Mineralogical Association in Kobe, Japan, O03-13.

Edwards, H.G.M., Newton, E.M., Dickensheets, D.L., and Wynn-Williams, D.D. (2003) Raman spectroscopic detection of biomolecular markers from Antarctic materials: evaluation for putative Martian habitats. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 59, 2277–2290.

Edwards, H.G.M., Hutchinson, I.B., Ingley, R., Waltham, N.R., Beardsley, S., Dowson, S., and Woodward, S. (2011) The search for signatures of early life on Mars: Raman spectroscopy and the Exomars mission. Spectroscopy Europe, 23(1), online.

Edwards, H.G.M., Hutchinson, I., and Ingley, R. (2012) The ExoMars Raman spectrometer and the identification of biogeological spectroscopic signatures using a flight-like prototype. Analytical and Bioanalytical Chemistry, 404, 1723–1731.

Ellery, A., and Wynn-Williams, D. (2003) Why Raman spectroscopy on Mars? – A case of the right tool for the right job. Astrobiology, 3, 565–579.

Fernandez-Remolar, D.C., Preston, L.J., Sanchez-Roman, M., Izawa, M.R.M., Huang, L., Southam, G., Banerjee, N.R., Osinski, G.R., Flemming, R., Gomez-Ortiz, D., Prieto Ballesteros, O., Rodriguez, N., Amils, R., and Dyar, M.D. (2012) Carbonate precipitation under bulk acidic conditions as a potential biosignature for searching life on Mars. Earth and Planetary Science Letters, 351, 13–26.

Foucher, F., Lopez-Reyes, G., Bost, N., Rull, F., Rüßmann, P., and Westall, F. (2013) Effect of grain size distribution on Raman analyses and the consequences for in situ planetary missions. Journal of Raman Spectroscopy, 44, 916–925.

Ghesti, G.F., de Macedo, J.L., Resck, I.S., Dias, J.A., and Dias, S.C.L. (2007) FT-Raman spectroscopy quantification of biodiesel in a progressive soybean oil transesterification reaction and its correlation with H-1 NMR spectroscopy methods. Energy and Fuels, 21, 2475–2480.

Hassoun, M.H. (1995) Fundamentals of Artificial Neural Networks. MIT Press, Cambridge, Massachusetts.

Hermosilla Rodriguez, I., Lopez-Reyes, G., Llanos, D.R., and Rull Perez, F. (2014) Automatic raman spectra processing for Exomars. In E. Pardo-Igúzquiza, C. Guardiola-Albert, J. Heredia, L. Moreno-Merino, J.J. Durán, and J.A. Vargas-Guzmán, Eds., Mathematics of Planet Earth, 127–130. Springer, Berlin.

Ishikawa, S.T., and Gulick, V.C. (2013) An automated mineral classifier using Raman spectra. Computers and Geosciences, 54(0), 259–268.

Johnson, R.A., and Wichern, D.W. (2002) Applied Multivariate Statistical Analysis. Prentice Hall, Upper Saddle River, New Jersey.

Josset, J.-L., Westall, F., Hofmann, B.A., Spray, J.G., Cockell, C., Kempe, S., Griffiths, A.D., De Sanctis, M.C., Colangeli, L., Koschny, D., Pullan, D., Föllmi, K., Diamond, L., Josset, M., Javaux, E., Esposito, F., and Barnes, D. (2012) CLUPI, a high-performance imaging system on the ESA-NASA rover of the 2018 ExoMars mission to discover biofabrics on Mars. EGU General Assembly Conference Abstracts, 14, 13616.

King, P.L., and McLennan, S.M. (2010) Sulfur on Mars. Elements, 6, 107–112.

Knoll, A.H., Carr, M., Clark, B., Farmer, J.D., Fischer, W.W., Grotzinger, J.P., McLennan, S.M., Malin, M., Schroder, C., Squyres, S., Tosca, N.J., and Wdowiak, T. (2005) An astrobiological perspective on Meridiani Planum. Earth and Planetary Science Letters, 240, 179–189.

Kong, W.G., and Wang, A. (2010) Planetary Laser Raman Spectroscopy for surface exploration on C/D-type asteroids: A case study. 41st Lunar and Planetary Science Conference, Abstract 2730. Lunar and Planetary Institute, Houston.

Koujelev, A., Sabsabi, M., Motto-Ros, V., Laville, S., and Lui, S.L. (2010) Laser-

induced breakdown spectroscopy with artificial neural network processing for material identification. Planetary and Space Science, 58, 682–690.

Kriesten, E., Alsmeyer, F., Bardow, A., and Marquardt, W. (2008) Fully automated indirect hard modeling of mixture spectra. Chemometrics and Intelligent Laboratory Systems, 91, 181–193.

Lafuente, B., Sansano, A., Navarro, R., Rull, F., Martínez Frías, J., Medina, J., Lopez, G., and Sobron, P. (2012) Multivariate analysis of Raman spectra for geological classification and identification: Application to Exomars mission. GeoRaman, Nancy.

Lambert, J.L., Morookian, J., Roberts, T., Polk, J., Smrekar, S., Clegg, S.M., Weins, R.C., Dyar, M.D., and Treiman, A. (2010) Standoff LIBS and Raman spectroscopy under venus conditions. 41st Lunar and Planetary Science Conference, Abstract 2608. Lunar and Planetary Institute, Houston.

Levenberg, K. (1944) A method for the solution of certain problems in least squares. Quarterly of Applied Mathematics, 2, 164–168.

Ling, Z.C., Wang, A., Jollif, B.L., Arvidson, R.E., and Xia, H.R. (2008) A systematic Raman, Mid-IR, and Vis-NIR spectroscopic study of ferric sulfates and implications for sulfates on Mars. 39th Lunar and Planetary Science Conference, Abstract 1463. Lunar and Planetary Institute, Houston.

Ling, Z.C., Wang, A., Jolliff, B.L., Li, C., Liu, J., Bian, W., Ren, X., Mu, L.L., and Su, Y. (2009) Raman spectroscopic study of quartz in lunar soils from Apollo 14 and 15 missions. 40th Lunar and Planetary Science Conference, Abstract 1823. Lunar and Planetary Institute, Houston.

Lopez-Reyes, G., Rull, F., Venegas, G., Westall, F., Foucher, F., Bost, N., Sanz, A., Catalá-Espí, A., Vegas, A., Hermosilla, I., Sansano, A., and Medina, J. (2013a) Analysis of the scientific capabilities of the ExoMars Raman Laser Spectrometer instrument. European Journal of Mineralogy, 25, 721–733.

Lopez-Reyes, G., Sobron, P., Lefebvre, C., and Rull, F. (2013b) Application of multivariate analysis techniques for the identification of sulfates from Raman spectra. 15th International Association for Mathematical Geosciences Conference, Madrid.

Marquardt, D.W. (1963) An algorithm for least-squares estimation of nonlinear parameters. Journal of the Society for Industrial and Applied Mathematics, 11, 431–441.

Martens, H., and Naes, T. (1992) Multivariate Calibration. Wiley, New York.

Nixon, S.L., Cockell, C.S., and Cousins, C.R. (2013) Plausible microbial metabolisms on Mars. Astronomy and Geophysics, 54, 13–16.

Özbalci, B., Boyaci, İ.H., Topcu, A., Kadılar, C., and Tamer, U. (2013) Rapid analysis of sugars in honey by processing Raman spectrum using chemometric methods and artificial neural networks. Food Chemistry, 136, 1444–1452.

Perez-Pueyo, R., Soneira, M.J., and Ruiz-Moreno, S. (2004) A fuzzy logic system for band detection in Raman spectroscopy. Journal of Raman Spectroscopy, 35, 808–812.

Rosipal, R., and Krämer, N. (2006) Overview and recent advances in partial least squares. In C. Saunders, M. Grobelnik, S. Gunn, and J. Shawe-Taylor, Eds., Subspace, Latent Structure and Feature Selection, 3940, p. 34–51. Springer Berlin Heidelberg.

Rull, F., Maurice, S., Diaz, E., Tato, C., Pacros, A., and Team., T.R. (2011a) The Raman Laser Spectrometer (RLS) on the ExoMars 2018 Rover Mission. 42nd Lunar and Planetary Science Conference LPSC, Abstract 2400. Lunar and Planetary Institute, Houston.

Rull, F., Sansano, A., Díaz, E., Canora, C.P., Moral, A.G., Tato, C., Colombo, M., Belenguer, T., Fernández, M., Manfredi, J.A.R., and others. (2011b) ExoMars Raman laser spectrometer for Exomars. Society of Photo-Optical Instrumentation Engineers (SPIE), Conference Series, 8152, 12.

Rull, F., Lopez, G., Catala, A., Medina, J., Sansano, A., Sanz, A., and Sobron, F. (2011c) Raman spectroscopy analysis on powdered samples inside the Exomars mission. ESA Congrex Lisbon 2011.

Sanchez-Andrea, I., Rojas-Ojeda, P., Amils, R., and Luis Sanz, J. (2012) Screening of anaerobic activities in sediments of an acidic environment: Tinto River. Extremophiles, 16, 829–839.

Schumacher, W., Kuehnert, M., Roesch, P., and Popp, J. (2011) Identification and classification of organic and inorganic components of particulate matter via Raman spectroscopy and chemometric approaches. Journal of Raman Spectroscopy, 42, 383–392.

Sharma, S.K., Lucey, P.G., Ghosh, M., Hubble, H.W., and Horton, K.A. (2003) Stand-off Raman spectroscopic detection of minerals on planetary surfaces. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 59, 2391–2407.

Sobron, P., and Alpers, C.N. (2013) Raman spectroscopy of efflorescent sulfate salts from iron mountain mine superfund site, California. Astrobiology, 13, 270–278.

Sobron, P., Sobron, F., Sanz, A., and Rull, F. (2008) Raman signal processing software for automated identification of mineral phases and biosignatures on Mars. Applied Spectroscopy, 62, 364–370.

Sobron, P., Sanz, A., Acosta, T., and Rull, F. (2009) A Raman spectral study of stream waters and efflorescent salts in Rio Tinto, Spain. Spectrochimica Acta part A: Molecular and Biomolecular Spectroscopy, 71, 1678–1682.

Sobron, P., Lefebvre, C., Koujelev, A., and Wang, A. (2013) Why Raman and LIBS for exploring icy moons? 44th Lunar and Planetary Science Conference, Abstract 2381. Lunar and Planetary Institute, Houston.

Sobron, P., Bishop, J., Blake, D., Chen, B., and Rull, F. (2014) Natural Fe-bearing oxides and sulfates from the Rio Tinto Mars analogue: Critical assessment of VNIR reflectance spectroscopy, laser Raman spectroscopy, and XRD as mineral identification tools. American Mineralogist, 99, 1199–1205.

Steele, A., McCubbin, F.M., Fries, M., Glamoclija, M., Kater, L., and Nekvasil, H. (2010) Graphite in an Apollo 17 impact melt breccia. Science, 329, 51.

Vagenas, N.V., and Kontoyannis, C.G. (2003) A methodology for quantitative determination of minor components in minerals based on FT-Raman spectroscopy – The case of calcite in dolomitic marble. Vibrational Spectroscopy, 32, 261–264.

Wang, A., Haskin, L.A., Lane, A.L., Wdowiak, T.J., Squyres, S.W., Wilson, R.J., Hovland, L.E., Manatt, K.S., Raouf, N., and Smith, C.D. (2003) Development of the Mars microbeam Raman spectrometer (MMRS). Journal of Geophysical Research, 108, 5005, doi:10.1029/2002JE001902.

Wang, A., Freeman, J.J., Jolliff, B.L., and Chou, I.M. (2006) Sulfates on Mars: A systematic Raman spectroscopic study of hydration states of magnesium sulfates. Geochimica et Cosmochimica Acta, 70, 6118–6135.

Wang, A., Zheng, M.P., Kong, F.J., Ling, Z.C., Kong, W.G., Sobron, P., and Jolliff, B.L. (2011) A Low T, High RH, and potentially life-friendly environment within the martian salt-rich subsurface in equatorial regions. 42nd Lunar and Planetary Science Conference, Abstract 2049. Lunar and Planetary Institute, Houston.

Wiens, R.C., Sharma, S.K., Clegg, S.M., Misra, A.K., and Lucey, P.G. (2007) Combined remote Raman spectroscopy and LIBS instrumentation for Mars astrobiology exploration. Seventh International Conference on Mars, Abstract 3092. Lunar and Planetary Institute, Houston.

Wold, S., Martens, H., and Wold, H. (1983) The multivariate calibration problem in chemistry solved by the PLS method. Lecture Notes in Mathematics, 973, 286–293.

Wynn-Williams, D.D., and Edwards, H.G.M. (2000) Proximal analysis of regolith habitats and protective biomolecules in situ by laser Raman spectroscopy: Overview of terrestrial antarctic habitats and Mars analogs. Icarus, 144, 486–503.

Epilogue

Chapter 6.

SUMMARY, CONCLUSIONS AND FUTURE LINES

6.1 SUMMARY AND CONCLUSIONS

This thesis intended to accomplish two main objectives. The first one is the definition of algorithmic procedures that allow the RLS instrument to autonomously operate and optimize the Raman acquisition and the available resources. The second one, to assess the science return that can be obtained from the RLS instrument, taking into account the limitations imposed by the rover (operation based on powdered samples).

To do so, Chapter 1 introduced Raman spectroscopy and the RLS instrument, while Chapter 2 described the RLS ExoMars Simulator, a system developed in order to be able to accomplish the thesis objectives. Chapter 3 described a set of proposed algorithms for the automatic execution of a complete RLS acquisition cycle, including algorithms for saturation, fluorescence and cosmic ray detection and removal, and for the automatic calculation of the optimal acquisition parameters, based on the quality of spectra of reference. These sections covered the first objective of the thesis.

The following chapters are centered in operation-mode related issues, in a format heavily supported by work published in several journals: Journal of Raman Spectroscopy, European Journal of Mineralogy and American Mineralogist. Chapter 4 studied the behavior of the sample preparation and distribution system of the rover to evaluate the impact on the acquired spectra. In addition, we presented a study of samples with different grain size distributions in order to understand more deeply the effects that the powdering of the sample can induce in Raman spectra. Finally, Chapter 5 presented results for the analysis of spectra acquired with representative conditions of the operation. Both standard and simple statistical techniques, as well as several multivariate analysis techniques were used to assess the performance of these techniques for the analysis of Raman spectra.

6.1.1 OBJECTIVE I: INSTRUMENT AUTOMATION

The RLS ExoMars simulator was developed at the Unidad Asociada UVa-CSIC-CAB in Valladolid in response to the need to operate the RLS instrument in a rover-like scenario from the operational point of view (which includes operating on powdered samples in a point-by-point fashion, positioning the sample and focusing at each point conveniently). It allowed the automatic flattening and positioning of the sample with a spatial resolution of 2.5 microns. The control software provided total flexibility over the system, in order to define any desired operation sequence. This proved to be useful during the data mining process (though in some cases, depending on the hardware used, not all functions would be available, e.g., with the RLS engineering breadboard model). In order to optimize the automatic acquisition of samples, several algorithms were defined, implemented, parameterized, optimized and validated, providing the RLS ExoMars Simulator with the unique capability to optimize the

acquisition parameters at each point of each sample, based on the characteristics of reference spectra.

The definition of the automatic algorithms for the instrument always followed the same development structure for every algorithm: problem description and characterization, algorithm proposal, parameterization (optimization) and validation. Each algorithm was validated unitarily, and the whole procedure applied end-to-end in the RLS ExoMars Simulator. The fluorescence reduction algorithm is one of the most important algorithms to apply in order to save operational time, as it will allow reducing the quenching time as a function of the sample response (based on the background decreasing rate per second). This is very important as the allocated time for fluorescence is relatively high, while not all samples will be fluorescent. Furthermore, a low incidence of fluorescence in Martian samples is expected. However, this is not an issue, as implementing this algorithm will not have impact on the operation with non-fluorescent samples (no time consumption for quenching in this case).

The basic operation of the acquisition parameters adjustment is based upon calculating spectral parameters from reference spectra acquired prior to the final acquisition. These parameters (e.g. the spectral quality) will be used by the algorithms to take decisions on the final acquisition parameters. Thus, spurious peaks or spikes such as cosmic rays might be critical for the correct functioning of the algorithms. In order to minimize the effect of this kind of effects, a dedicated algorithm was implemented for the detection and removal of spikes. This algorithm was based on comparing two spectra and detecting all differential spectral features between them. As a result, many of the spurious detected during our test campaigns were corrected, improving the end-to-end behavior of the system.

The adjustment of the integration time and number of accumulations is based on adjusting Ti to fill 80% of the dynamic range of the CCD, while the number of accumulations is decided based on the spectral characteristics of the reference spectra (the higher the SNR, the lower the necessary number of accumulations). This algorithm is relatively complex, as it requires the calculation of the baseline of the spectrum in order to define its SNR. The parameterization of this algorithm required the characterization of the RLS instrument with different integration times, and with respect to the number of accumulations. Based on this characterization, it was possible to define look-up tables for a fast calculation of the acquisition parameters.

All the algorithmic defined in this thesis is designed to work sequentially, and in the most efficient way possible from the computational point of view (which is critical for the onboard execution). Even if the algorithms are relatively independent, in order to

save computational effort, some common parts among the different blocks are shared and executed only once.

6.1.2 OBJECTIVE II: ACHIEVABLE SCIENCE BASED ON RLS DATA PRODUCTS

The second part of the thesis focused on the analysis of the data products obtained from the RLS instrument, under the operational conditions imposed by the rover, in order to evaluate the science return that can be extracted from RLS. To do so, tests were performed to assess the impact of working on crushed material, and novel analytical techniques applied to the data in order to assess the science exploitation capabilities based on this kind of data.

The tests performed with the sample preparation and distribution system (SPDS) breadboard allowed studying several key aspects from the sample and delivery of the samples: among other things, the rugosity of the flattened surface, the cross-contamination induced by the SPDS and the possibilities of the MicrOmega-RLS collaborative operation, in which the SPDS will place an interesting grain found by MicrOmega for RLS to analyze it. These tests showed that the flattened surface provided by the SPDS was acceptable by RLS, as long as the autofocus mechanism is in the instrument scope. The cross-contamination from the SPDS system, though extant, can be considered negligible compared to the cross-contamination expected from the sample crushing mechanism. Finally, the MicrOmega-RLS collaborative operation showed that it is feasible to improve the overall operation by acquiring spectra at regions of interest detected by other instruments. However, it highlighted some issues to be taken into account during operation. Namely, the sample surface morphology modification due to vibrations in the setup, which could impair the accurate positioning of the samples. In addition to these considerations, this thesis showed how the analysis of powdered samples can induce some effects on the Raman spectra of the samples, as for example a general increase on the background level. Other effects such as peak displacement or broadening have also been addressed. However, the RLS irradiance level and spectral resolution characteristics make most of these effects irrelevant for the RLS instrument.

This thesis also covers several other issues related to the crushing process, such as the sample textural and context loss. This is a major issue, for example when dealing with potentially biotic carbon samples, where the morphology of the bulk sample is key to assess their biotic or abiotic origin. The crushing process removes the mineral context information. However, the crushing process can facilitate the detection of minor traces (which might be found in the interior of the bulk sample) by the RLS instrument, as was the case for carbon in a chert, and for ankerite in a carbonate deposit from Svalbard. In addition, the analyses of mixtures with controlled proportions showed that the detection limit of the instrument could be as low as 1% (and probably lower,

depending on the samples). This performance, in combination with other instruments of the ExoMars rover, such as the Close-Up Imager (CLUPI), can still provide relevant scientific information from the analyzed samples.

The operation on powdered samples will provide a set of spectra from at least 20 points of the same sample. Therefore, it seems logical to exploit the statistical analysis of the data. We have described and applied several techniques for the statistical analysis of the acquired samples in order to evaluate the science return from RLS-type data sets, divided into two types: univariate and multivariate analysis. We applied a univariate analysis technique (as opposed to multivariate) which consisted on calculating the ratio of the main peaks of the samples in binary mixtures. This simple calculation provided a very linear response with the actual abundance of the mineral phases in the mixtures, supported by the automatic adjustment of the acquisition times, which compensates the different cross-sections of the materials (the automatic acquisition *de facto* works as a normalization-during-acquisition procedure).

In addition, we tested and compared several multivariate analysis techniques, again with controlled binary mixtures of samples: Principal Component Analysis (PCA), Partial Least Squares regression (PLS) and Artificial Neural Neworks (ANN). All of these techniques have their own special features, and were described and compared. Furthermore, they were tested in order to assess their capacity to identify and/or quantify the mineral abundance of the samples under analysis.

The results showed that PCA scores gradually separate as a function of the abundance of the endmembers of the mixtures; PLS provided nice calibration curves for the analyzed samples and, finally, models trained with ANN were capable of identifying and quantifying the abundance of mineral phases. The interesting thing about all these statistical techniques is that the trained models seemed to be valid for spectra acquired with different experimental setups and/or different samples of the same type.

All these results showed that, even though the operational mode imposed by the rover for the analysis of samples (several points acquired on a powdered sample) imply the loss of the samples textural context, this might improve the detection limits of the samples, and allow some kind of pseudo-quantification of the mineral abundance in heterogeneous mixtures.

Though the results presented in this thesis project are based on analyses on powdered samples that were carried out in terrestrial ambient conditions, they were not invalidated by analyses executed in Martian conditions. However, all these results, as well as the RLS ExoMars Simulator Martian Simulation Chamber design and construction are not described here, but will be included in a dedicated thesis project.

6.2 FUTURE LINES

Overall, this thesis summarizes a project covering the operation mode of the RLS instrument from bottom to top, from the development of the laboratory equipment necessary for the definition of the operation mode, to the science return that can be obtained from the instrument, with intermediate steps to define the algorithmic for the automated operation, and studies to help understand the effects that the working environment in the ExoMars rover can have on the RLS instrument. Finally, a solution is provided to cover the automated end-to-end operation of the instrument once on Mars, plus possible techniques for the analysis of the returned data.

However, the proposal presented on this thesis will need to be revised and updated during the following years. The optimization of the control algorithmic of the automated operation is based on spectra acquired with a breadboard version of the RLS instrument. It will be necessary to re-optimize the algorithms with the flight version of the instrument, once it is developed. Furthermore, the breadboard spectra database will be completed with new samples that might also be interesting to take into account for the parameterization of the algorithms.

The analytical techniques presented in this work provide a first approach of the use of these techniques for the identification and quantification of Raman spectra. Further analysis should be performed in order to test these techniques with more complex mineral samples (not only with binary mixtures), as well as to improve the models to escalate to a full-operational Raman database. Especially interesting would be to check the behavior of the ratio-based statistical analysis for ternary or more complex samples. Furthermore, the application of ANN seems promising, but has a complex execution. Probably, the application of several multivariate analysis techniques at the same time would provide satisfactory results. Continuing with the development of these analytical techniques will be of paramount importance for the RLS instrument in the framework of the ExoMars mission.

The spectral treatment algorithms for onboard operation here presented are good candidates to be applied for the ground segment during operation (in the Instument Data Analysis Tool of the RLS instrument). These techniques, though limited to fit to the onboard computational resources of the instrument, can set the base for an automated system for the analysis of spectra, especially if combined with multivariate or statistical analyses of the samples and a spectral database. This would provide the scientific team working in the Science Operations Center of the mission with tools to automatically analyze the spectra, obtaining results within minutes of receiving the data from Mars, allowing taking fast tactical decisions during the daily operation of the rover.

Appendix A.

FLUORESCENCE VALIDATION RESULTS

This appendix presents several graphs obtained from the analysis of the fluorescence decay in several natural samples in several areas of Tenerife Island.

The following figures show the validation results of the fluorescence study. These results were obtained with samples obtained from different sites in Tenerife Island. These samples were chosen as most of them happened to be fluorescent, which made them good candidates for validation of the algorithm. These graphs are shown here for reference as appendix to section 3.2, which is where the analytical results and conclusions are presented.
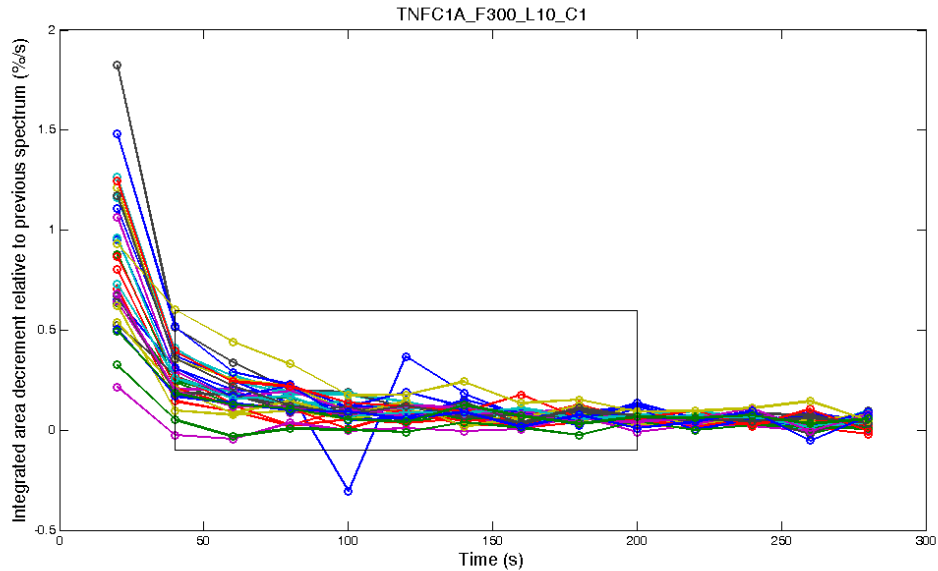
Figure A-1. Fluorescence decay for different points of sample TNFC1A
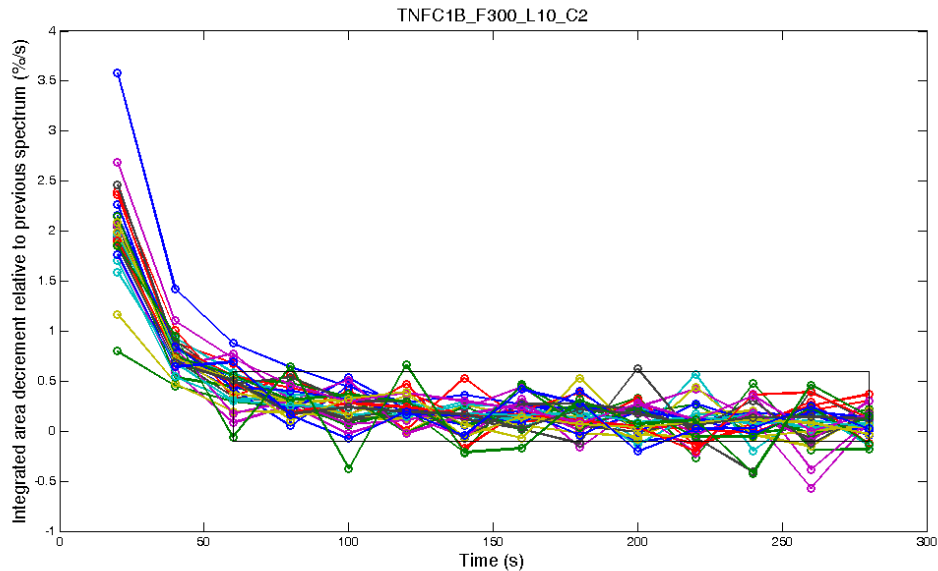
Figure A-2. Fluorescence decay for different points of sample TNFC1B
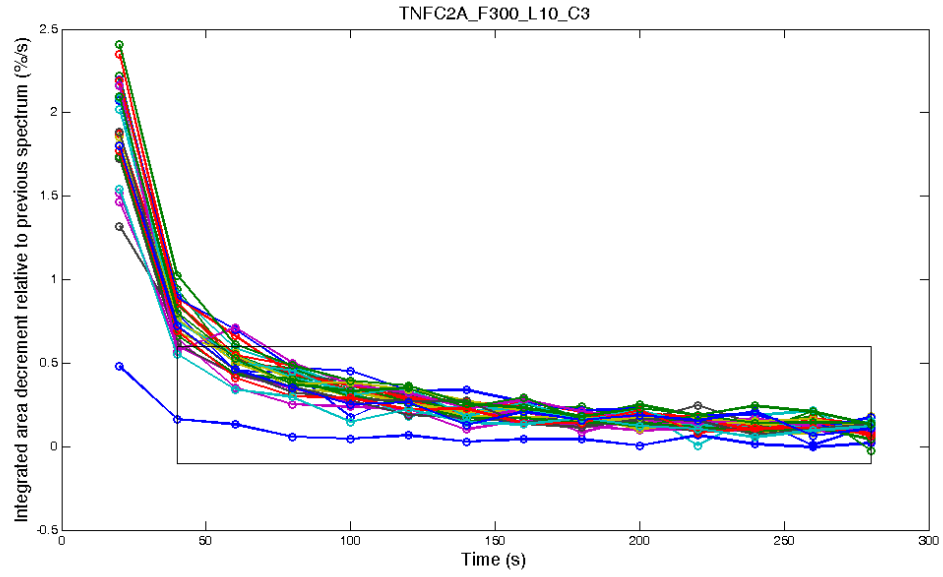
Figure A-3. Fluorescence decay for different points of sampleTNFC2A
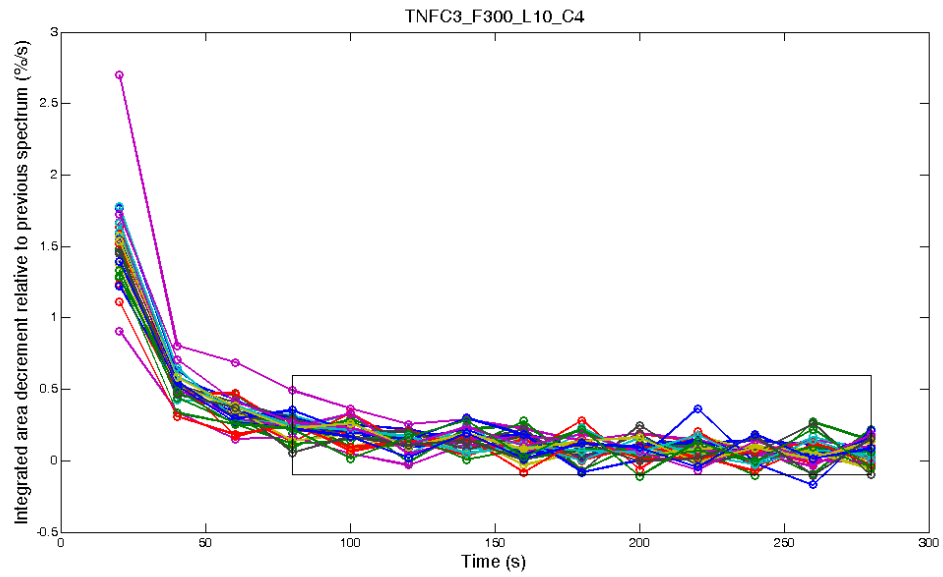
Figure A-4. Fluorescence decay for different points of sampleTNFC3
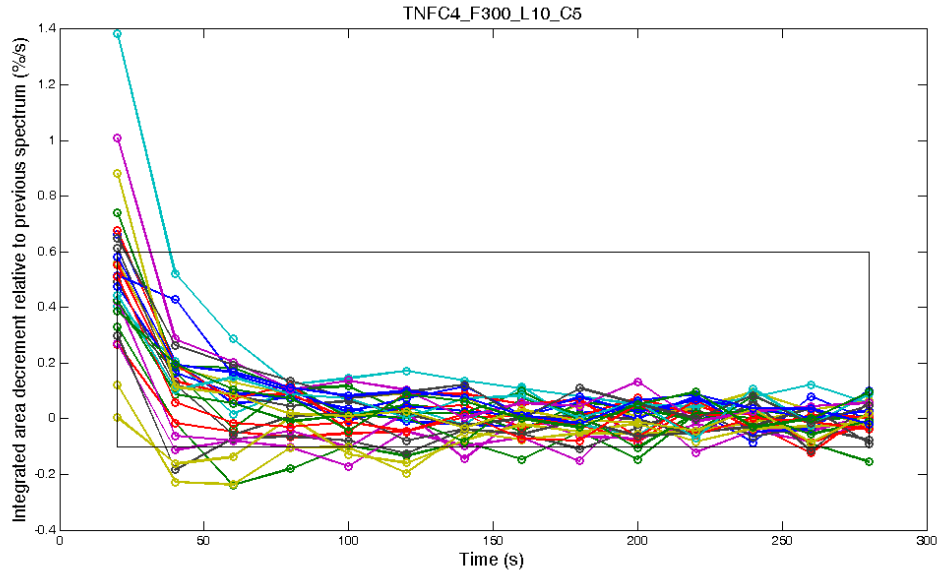
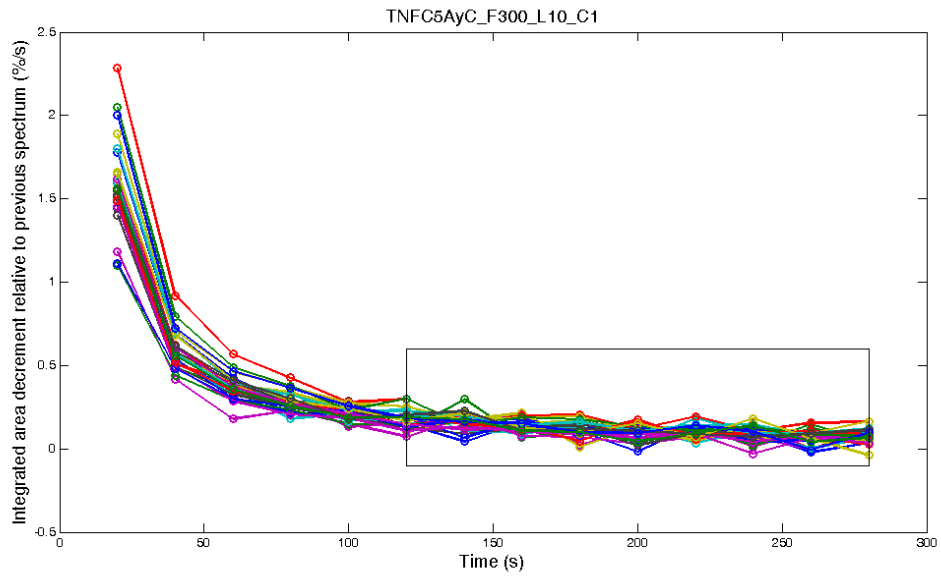**Figure A-5. Fluorescence decay for different points of sampleTNFC4**

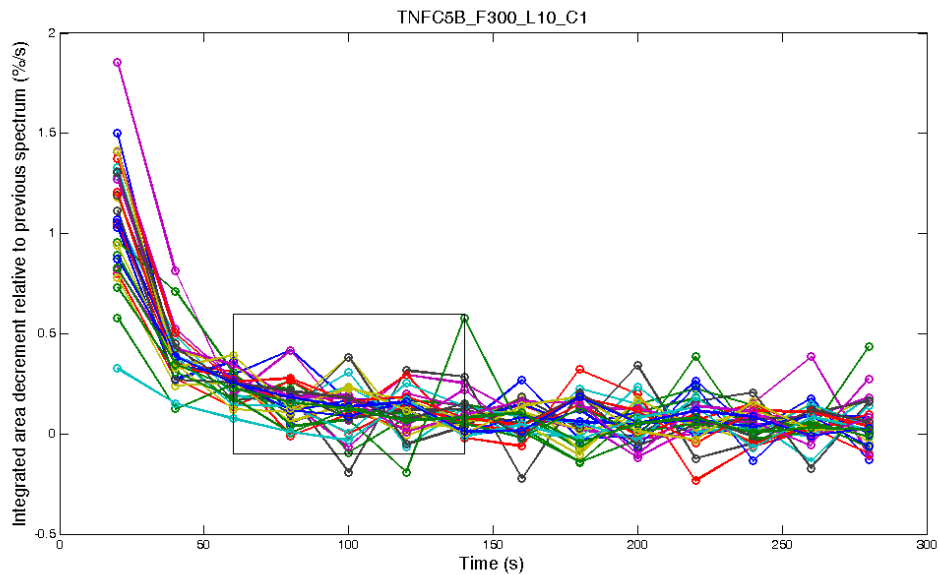**Figure A-6. Fluorescence decay for different points of sampleTNFC5AyC**

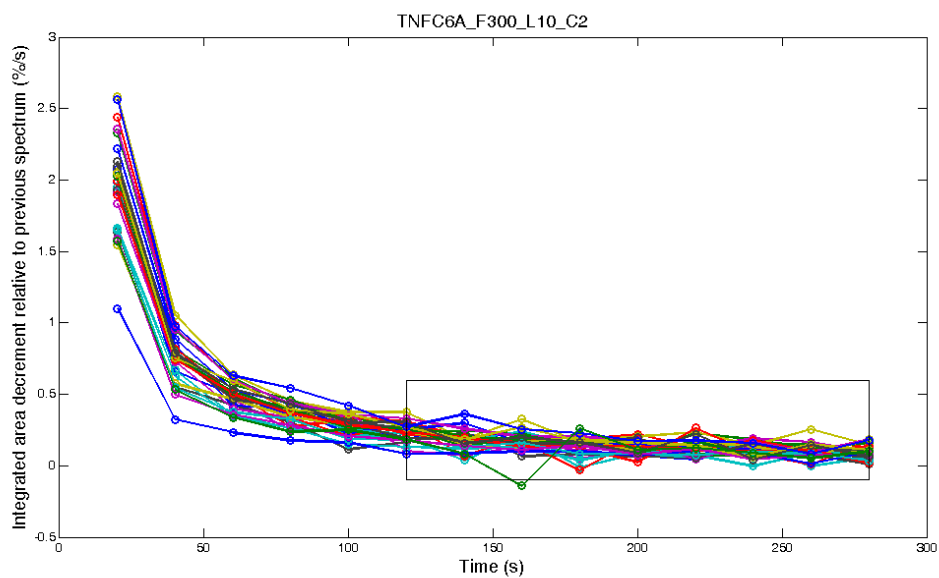**Figure A-7. Fluorescence decay for different points of sampleTNFC5B**

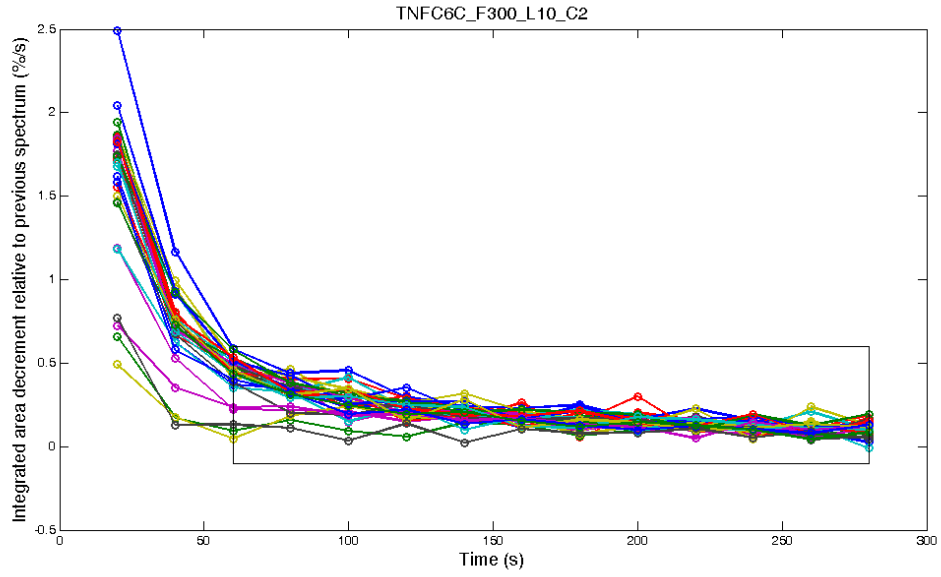**Figure A-8. Fluorescence decay for different points of sampleTNFC6A**

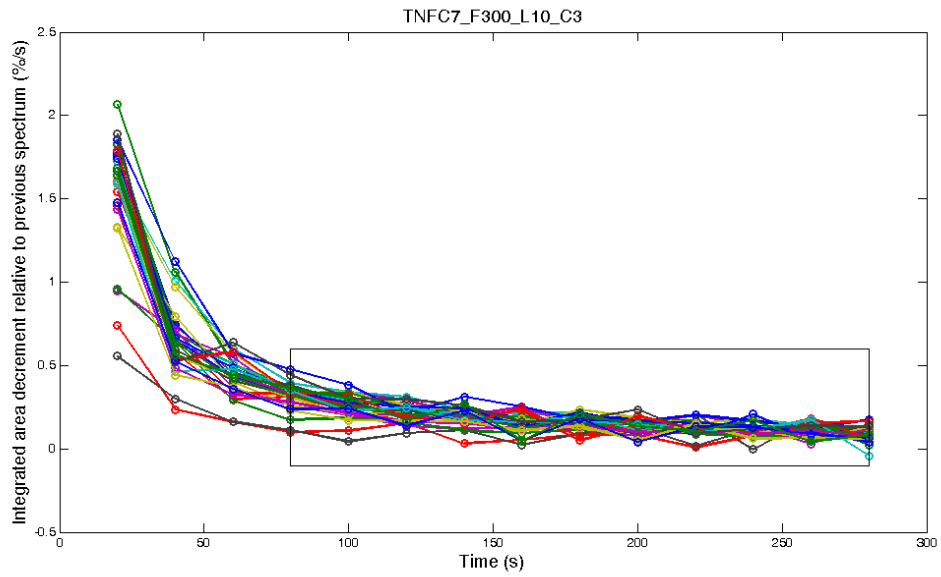**Figure A-9. Fluorescence decay for different points of sampleTNFC6C**

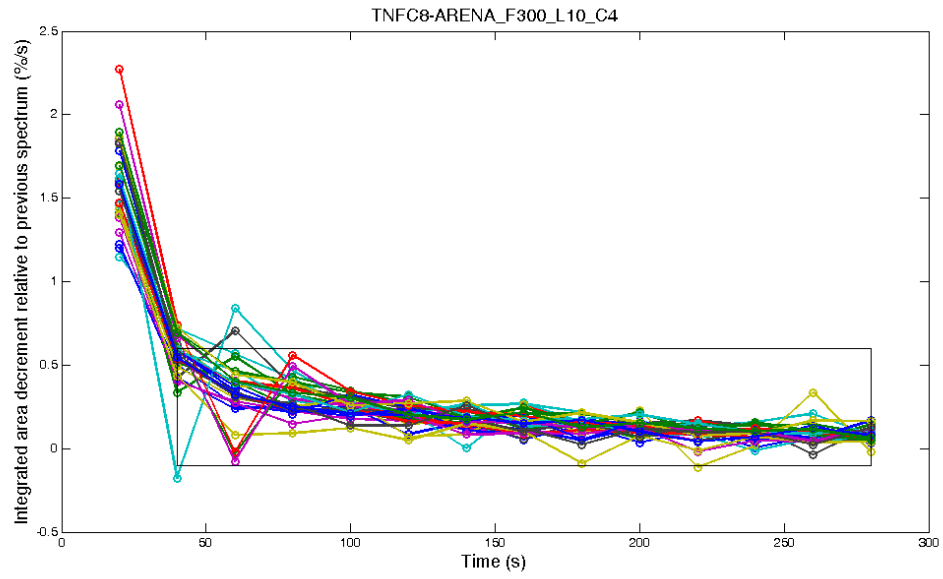**Figure A-10. Fluorescence decay for different points of sampleTNFC7**

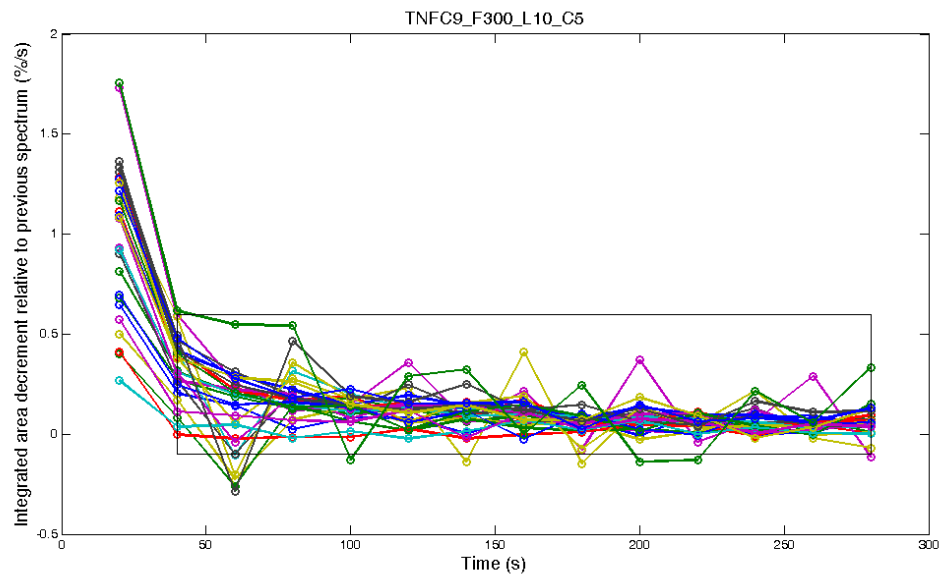**Figure A-11. Fluorescence decay for different points of sampleTNFC8**

Figure A-12. Fluorescence decay for different points of sampleTNFC9
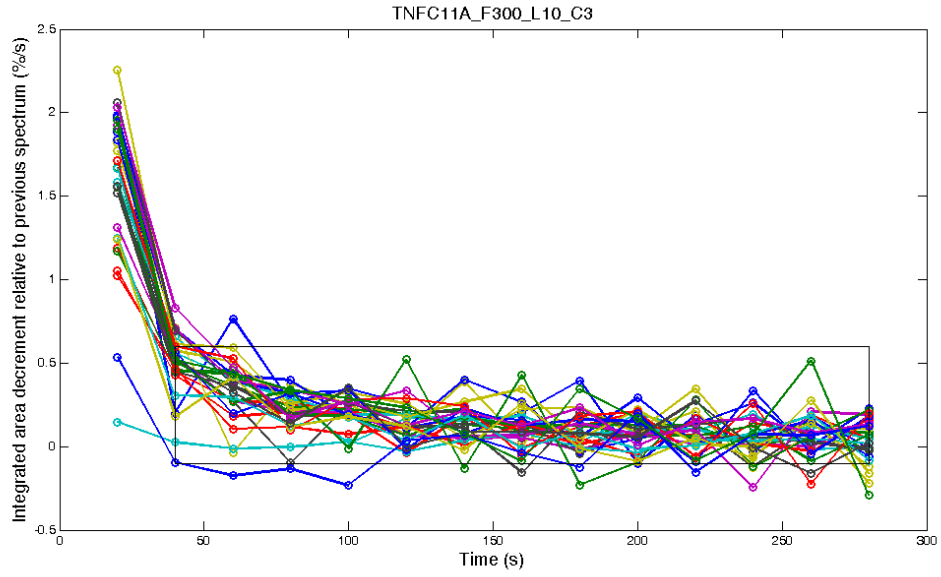
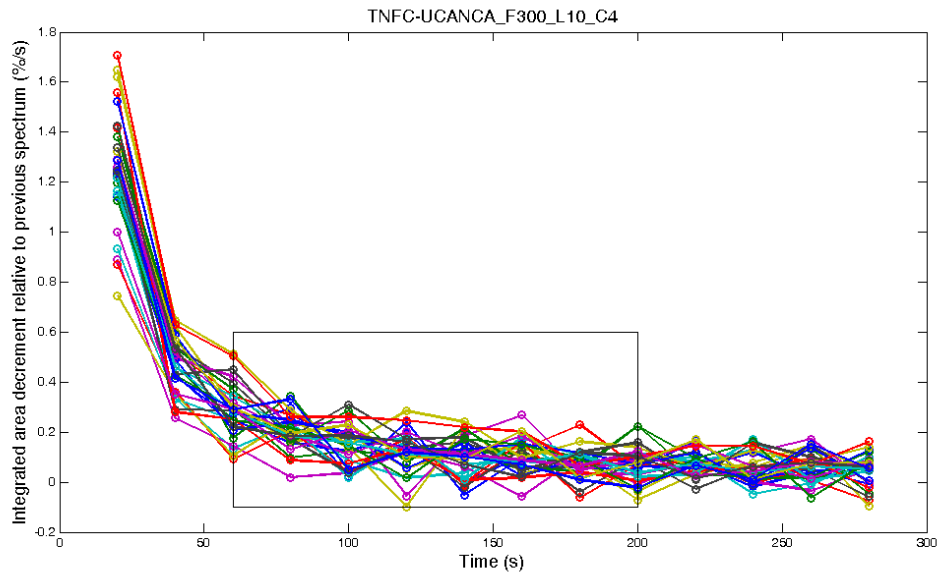**Figure A-13. Fluorescence decay for different points of sampleTNFC11A**

Figure A-14. Fluorescence decay for different points of sampleTNFC-UCANCA

Appendix B.

COSMIC RAY DETECTION VALIDATION RESULTS

This appendix presents several graphs obtained from the analysis of the cosmic rays in several spectra from the RLS instrument breadboard database.

The figures in this appendix present several graphs obtained after the analysis of those spectra from the RLS breadboard database, acquired with 1 second integration time, that present cosmic rays, and how the cosmic ray algorithm deals with them. For reference, the result of the algorithm (bottom graph) is presented with the original spectra (up), and a filtered version of the spectrum (middle), to compare with a filter-based algorithm. These results are presented as appendix to section 3.3.

These figures represent results of the algorithm execution with spectra from samples of calcite, silicon and gossan. For the noisiest spectra, where the spikes are more similar to the noise level, the results after executing the algorithm are equivalent to the filtered version (see Figure B-4, Figure B-6, Figure B-8, Figure B-9, Figure B-10 Figure B-12). However, as a general rule, the algorithm removes the spikes satisfactorily (see Figure B-1, Figure B-2, Figure B-3, Figure B-5, Figure B-7, Figure B-11, Figure B-13, Figure B-14, Figure B-15 and Figure B-16). This is especially important in several cases where not removing them would have implied important errors, due to the very high intensity of the spikes.

Figure B-1. Comparative of the raw 1 s reference spectra, the filtered version and the results after application of the cosmic ray removal algorithm for spectra from sample Silicon_P2.

Figure B-2. Comparative of the raw 1 s reference spectra, the filtered version and the results after application of the cosmic ray removal algorithm for spectra from sample Silicon_P1 – set 1.

Figure B-3. Zoom from Figure B-2.

Figure B-4. Comparative of the raw 1 s reference spectra, the filtered version and the results after application of the cosmic ray removal algorithm for spectra from sample Silicon_P1 – set 2 (no spikes).
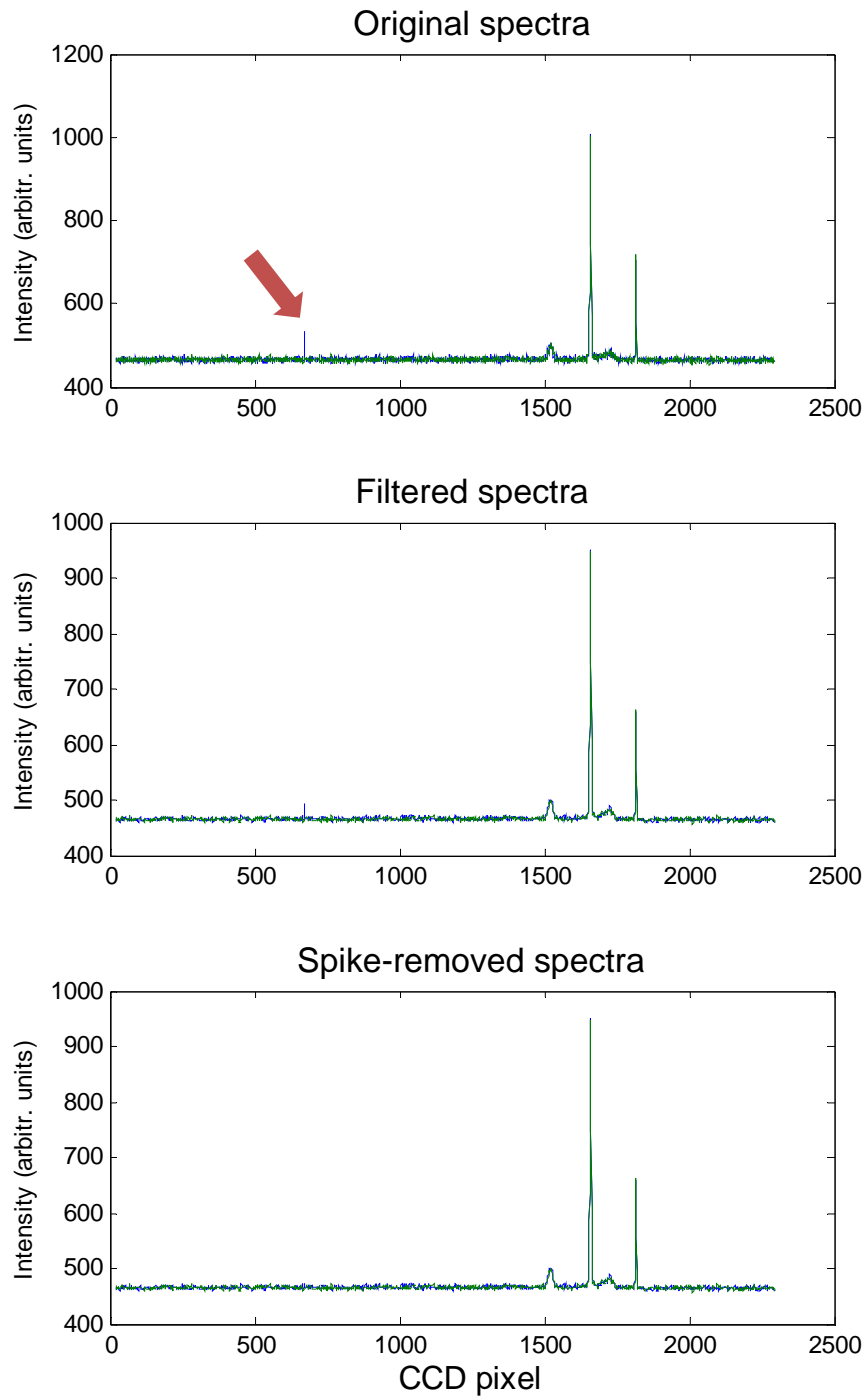
Figure B-5. Comparative of the raw 1 s reference spectra, the filtered version and the results after application of the cosmic ray removal algorithm for spectra from sample Gossan_P2 – set 1.

Figure B-6. Comparative of the raw 1 s reference spectra, the filtered version and the results after application of the cosmic ray removal algorithm for spectra from sample Gossan_P2 – set 2.
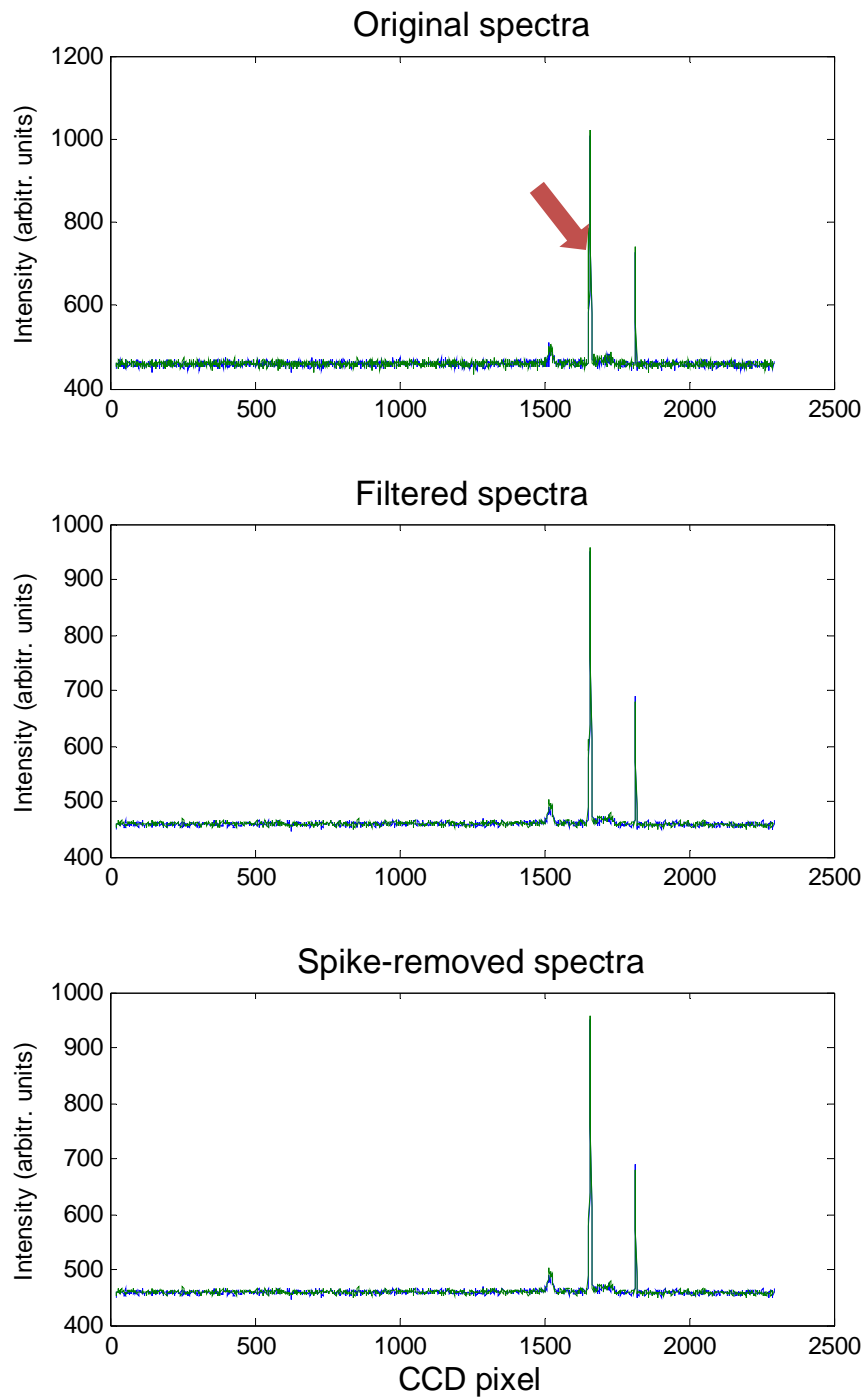
Figure B-7. Comparative of the raw 1 s reference spectra, the filtered version and the results after application of the cosmic ray removal algorithm for spectra from sample Gossan_P2 – set 3.

Figure B-8. Comparative of the raw 1 s reference spectra, the filtered version and the results after application of the cosmic ray removal algorithm for spectra from sample Gossan_P2 – set 4.
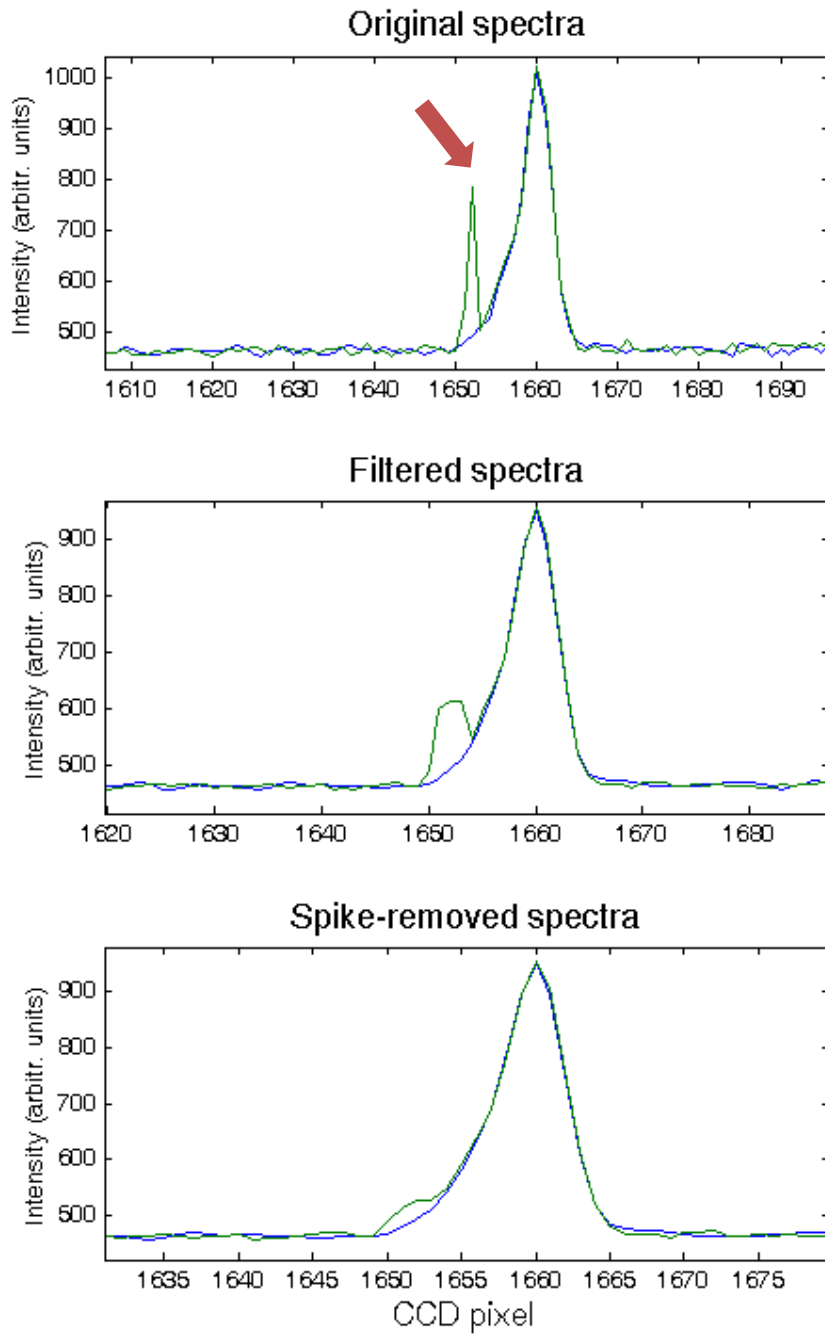
Figure B-9. Comparative of the raw 1 s reference spectra, the filtered version and the results after application of the cosmic ray removal algorithm for spectra from sample Gossan_P2 – set 5.
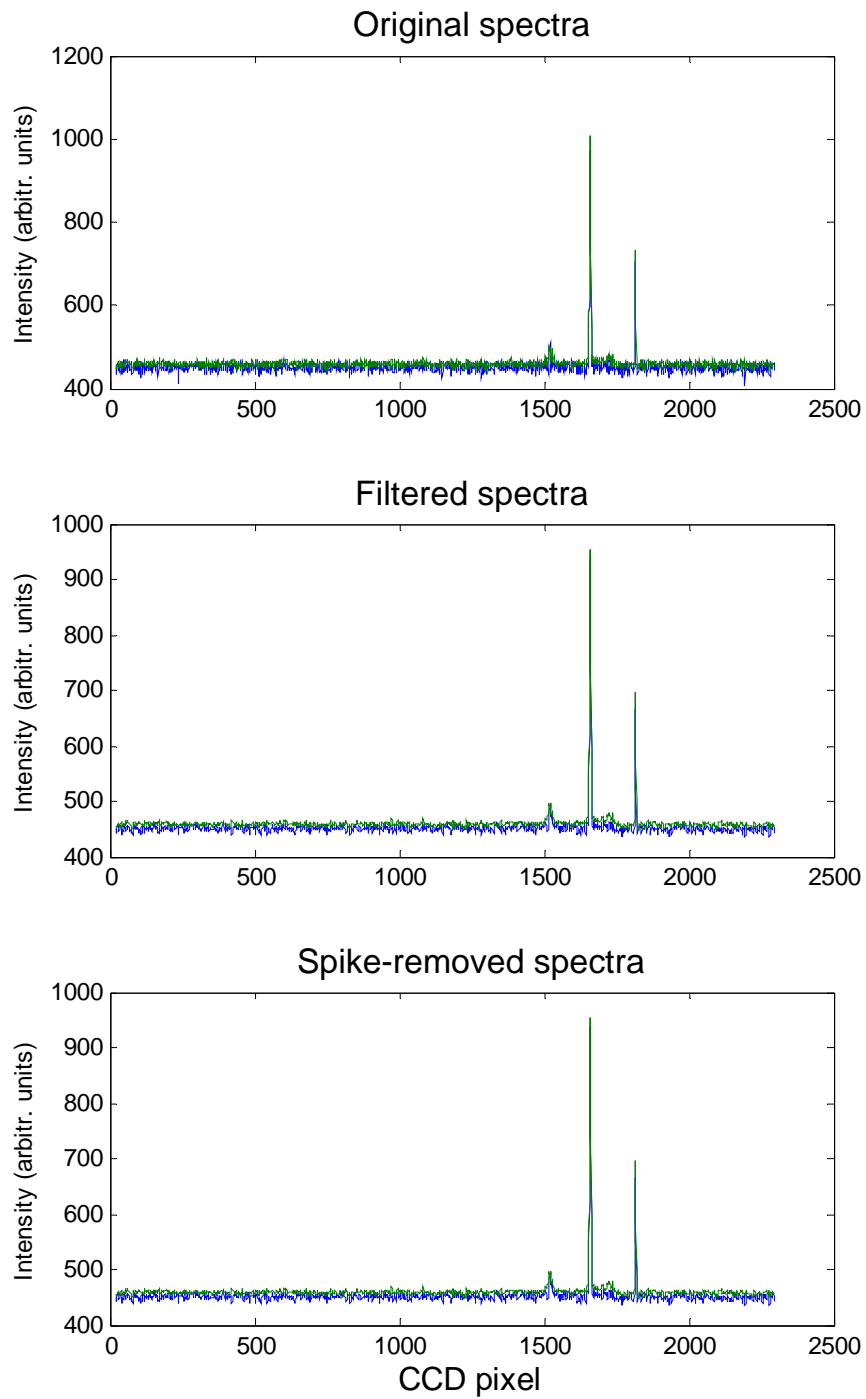
Figure B-10. Comparative of the raw 1 s reference spectra, the filtered version and the results after application of the cosmic ray removal algorithm for spectra from sample Gossan_P1 – set 1.
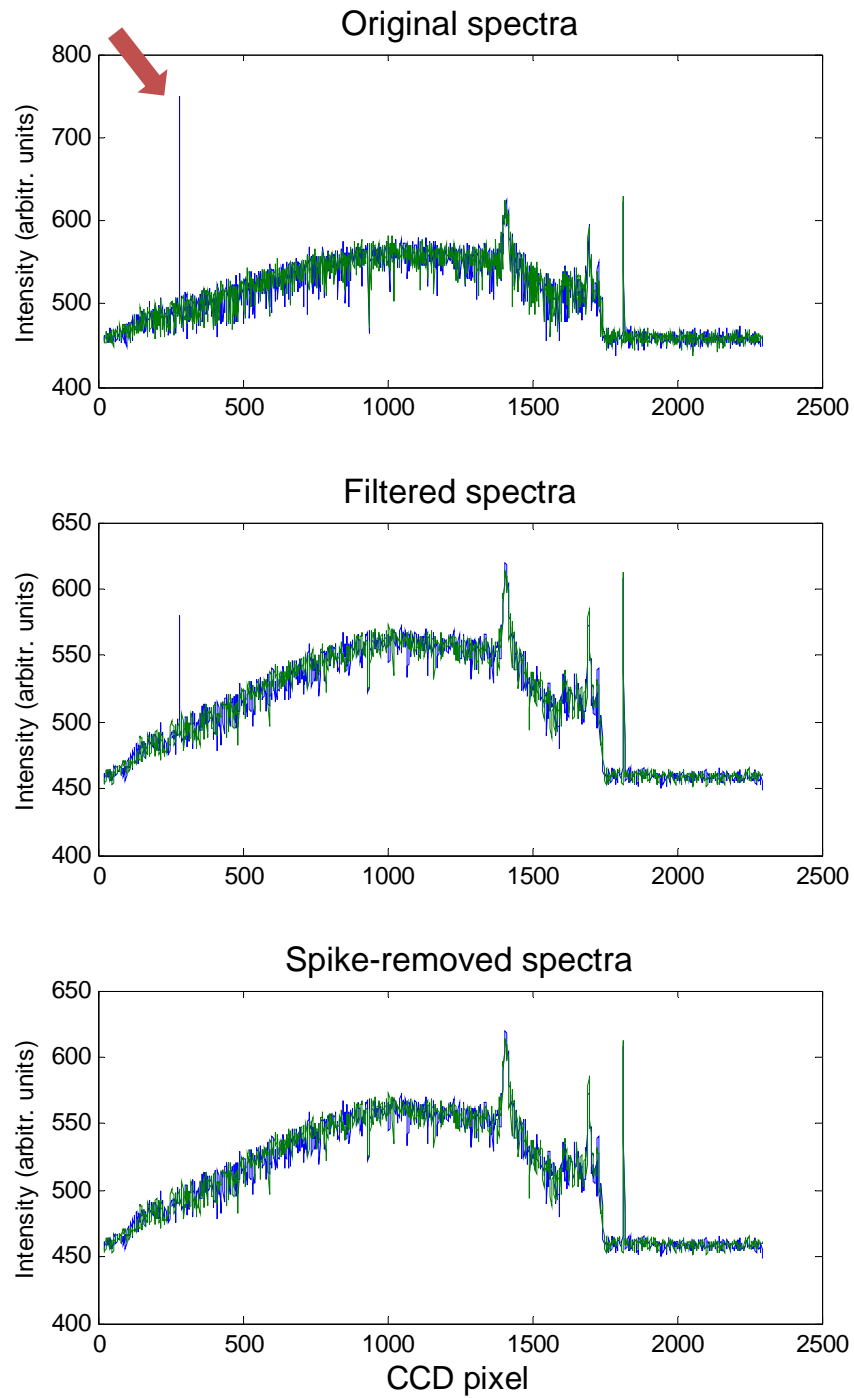
Figure B-11. Comparative of the raw 1 s reference spectra, the filtered version and the results after application of the cosmic ray removal algorithm for spectra from sample Gossan_P1 – set 2.

Figure B-12. Comparative of the raw 1 s reference spectra, the filtered version and the results after application of the cosmic ray removal algorithm for spectra from sample Gossan_P1 – set 3.
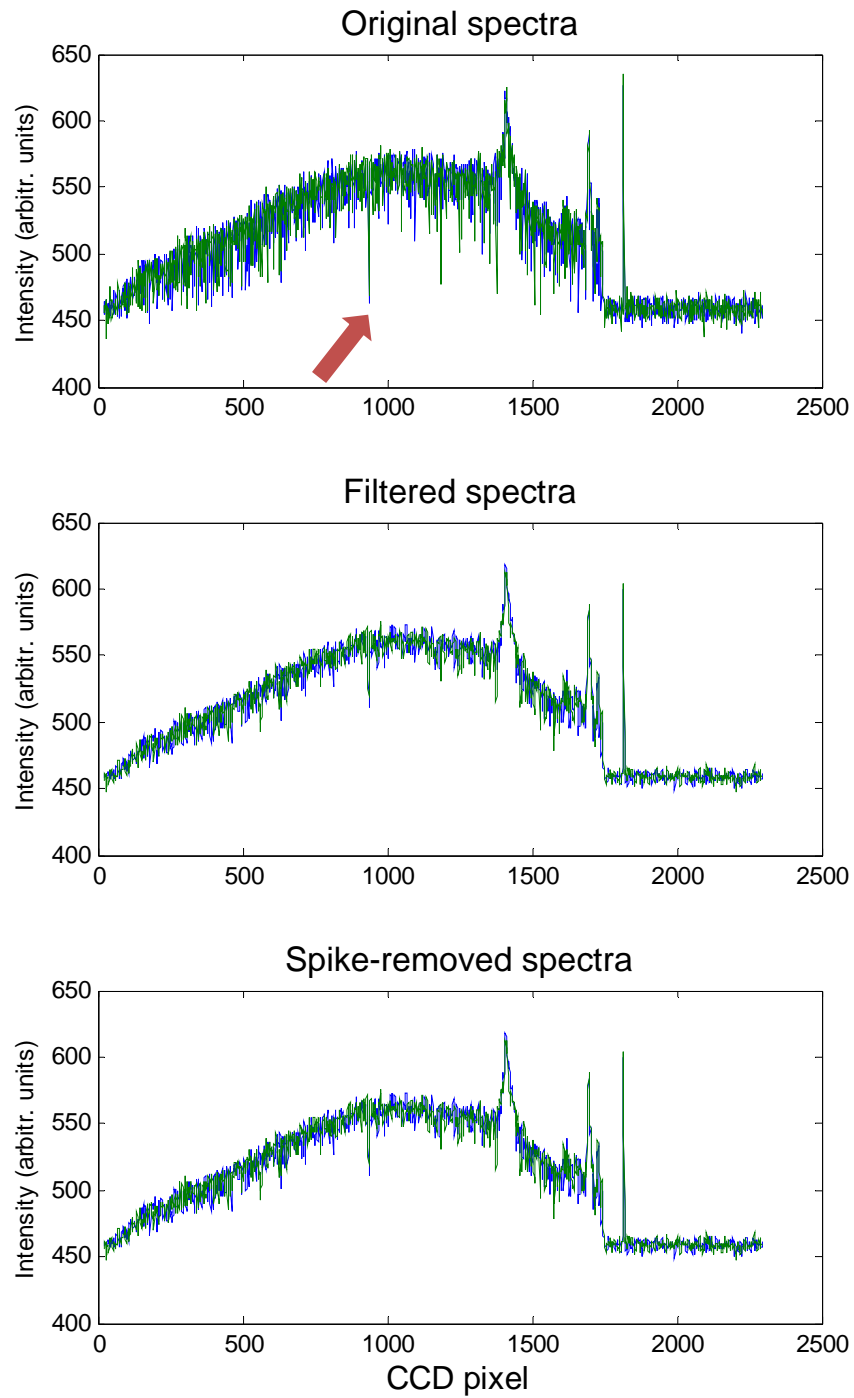
Figure B-13. Comparative of the raw 1 s reference spectra, the filtered version and the results after application of the cosmic ray removal algorithm for spectra from sample Calcite_P2 – set 1.

Figure B-14. Comparative of the raw 1 s reference spectra, the filtered version and the results after application of the cosmic ray removal algorithm for spectra from sample Calcite_P2 – set 2.
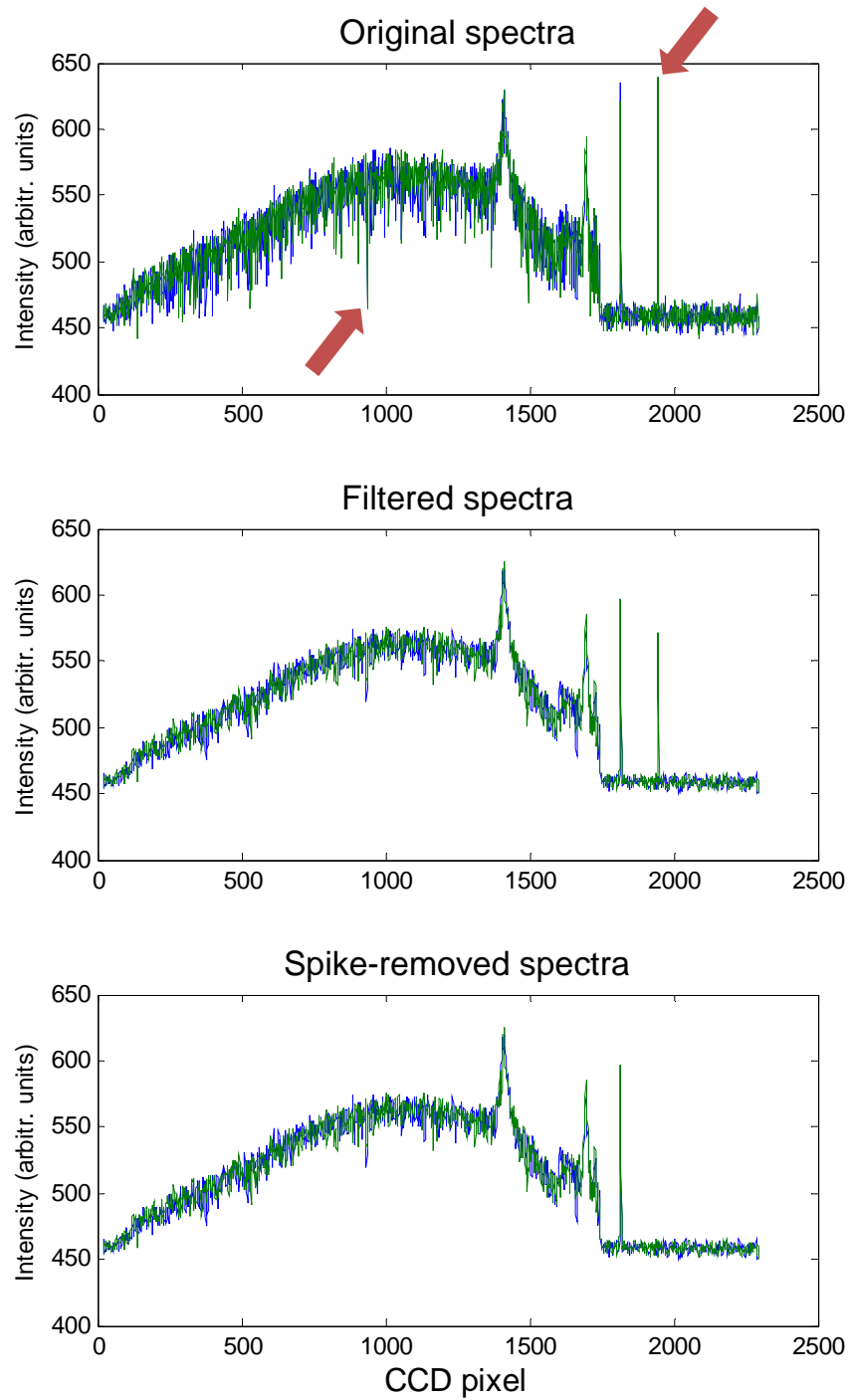
Figure B-15. Comparative of the raw 1 s reference spectra, the filtered version and the results after application of the cosmic ray removal algorithm for spectra from sample Calcite_P2 – set 3.
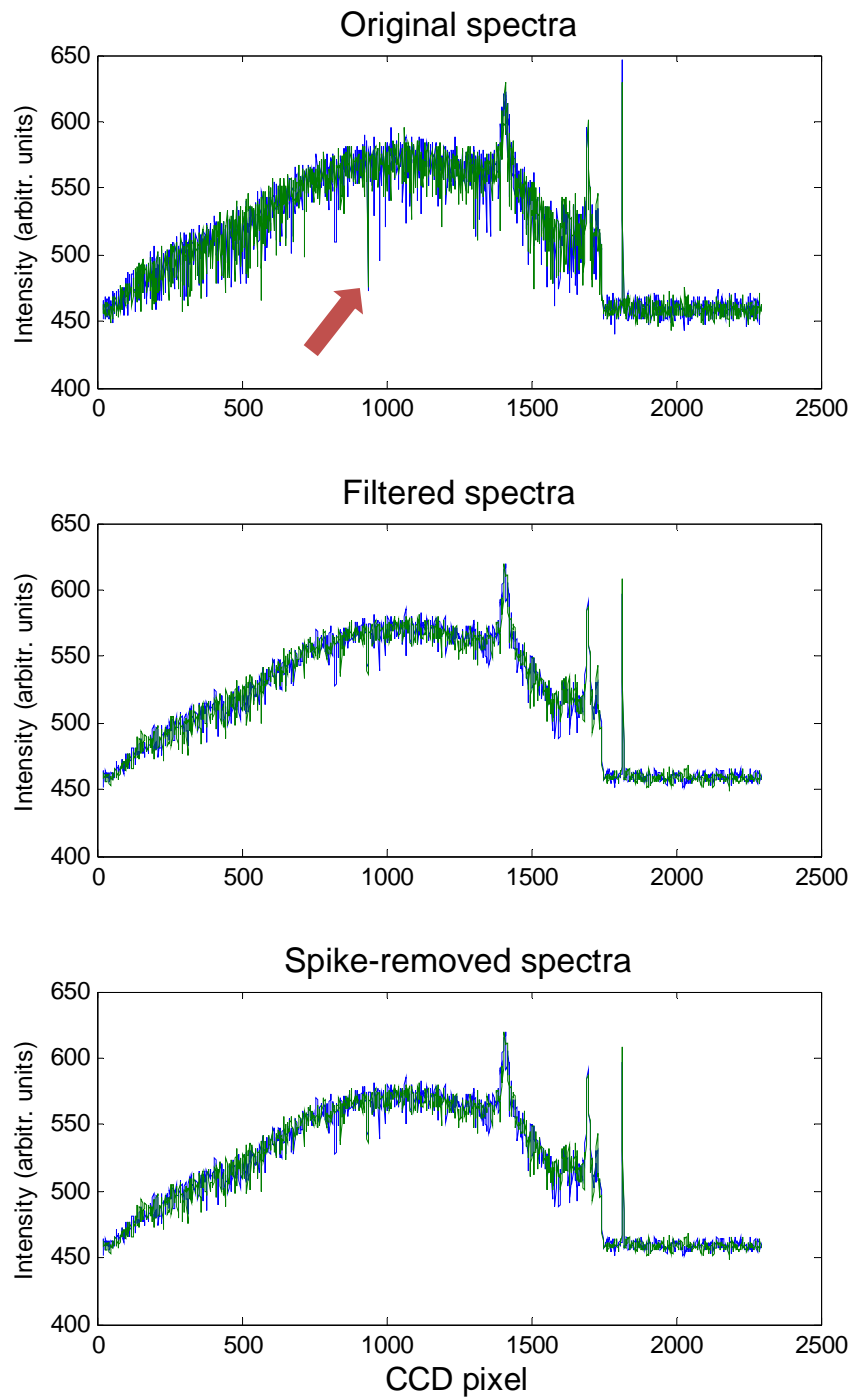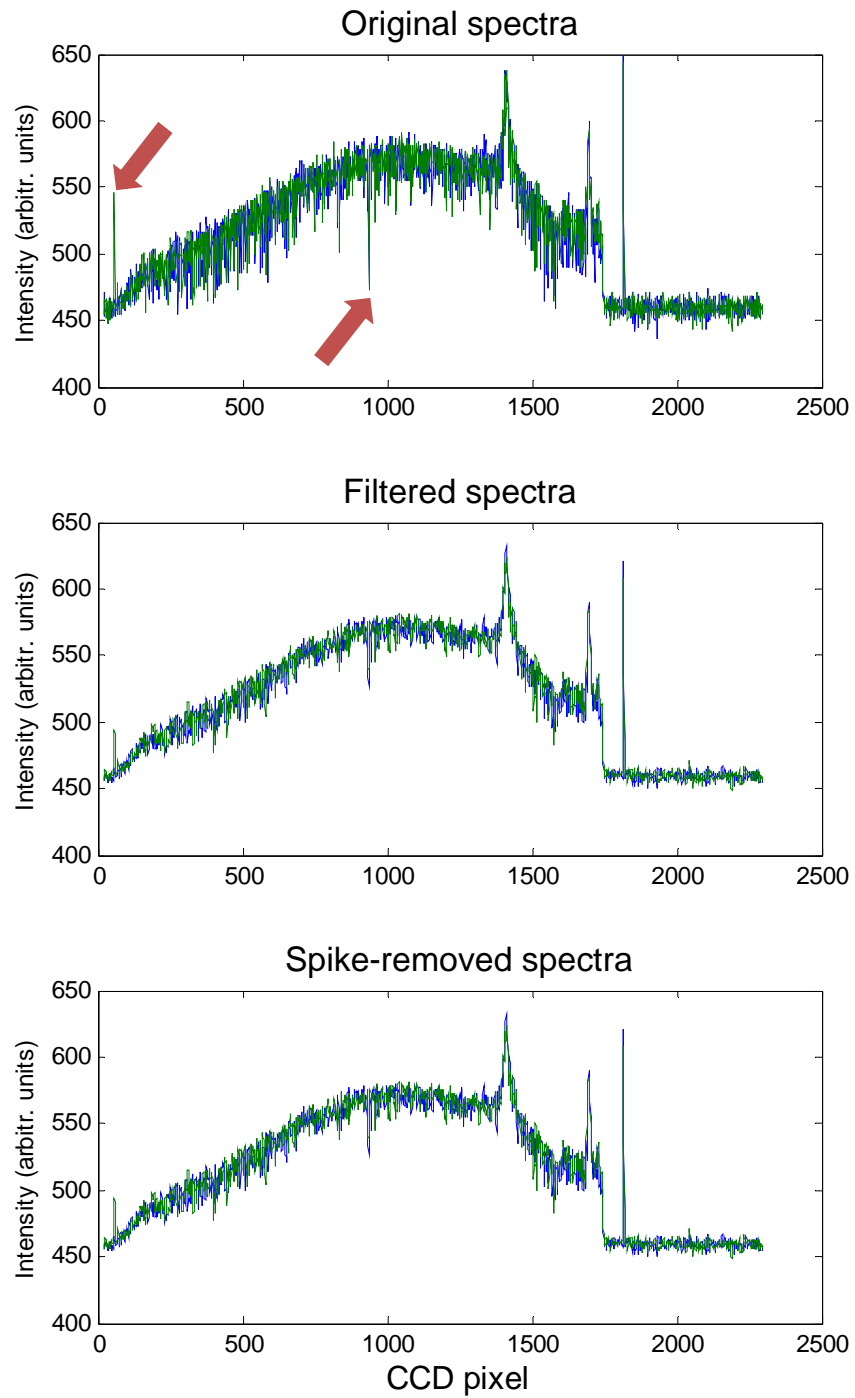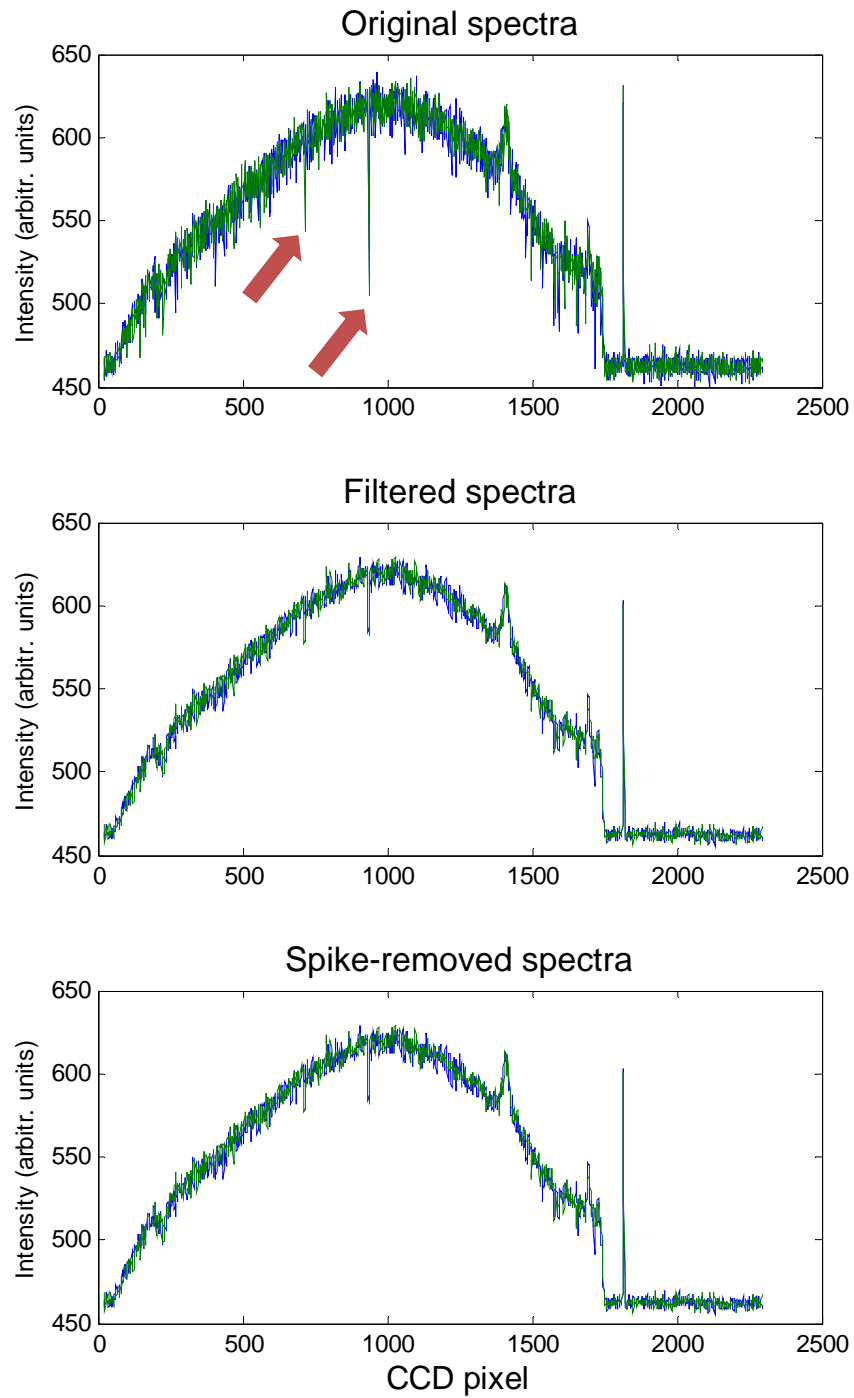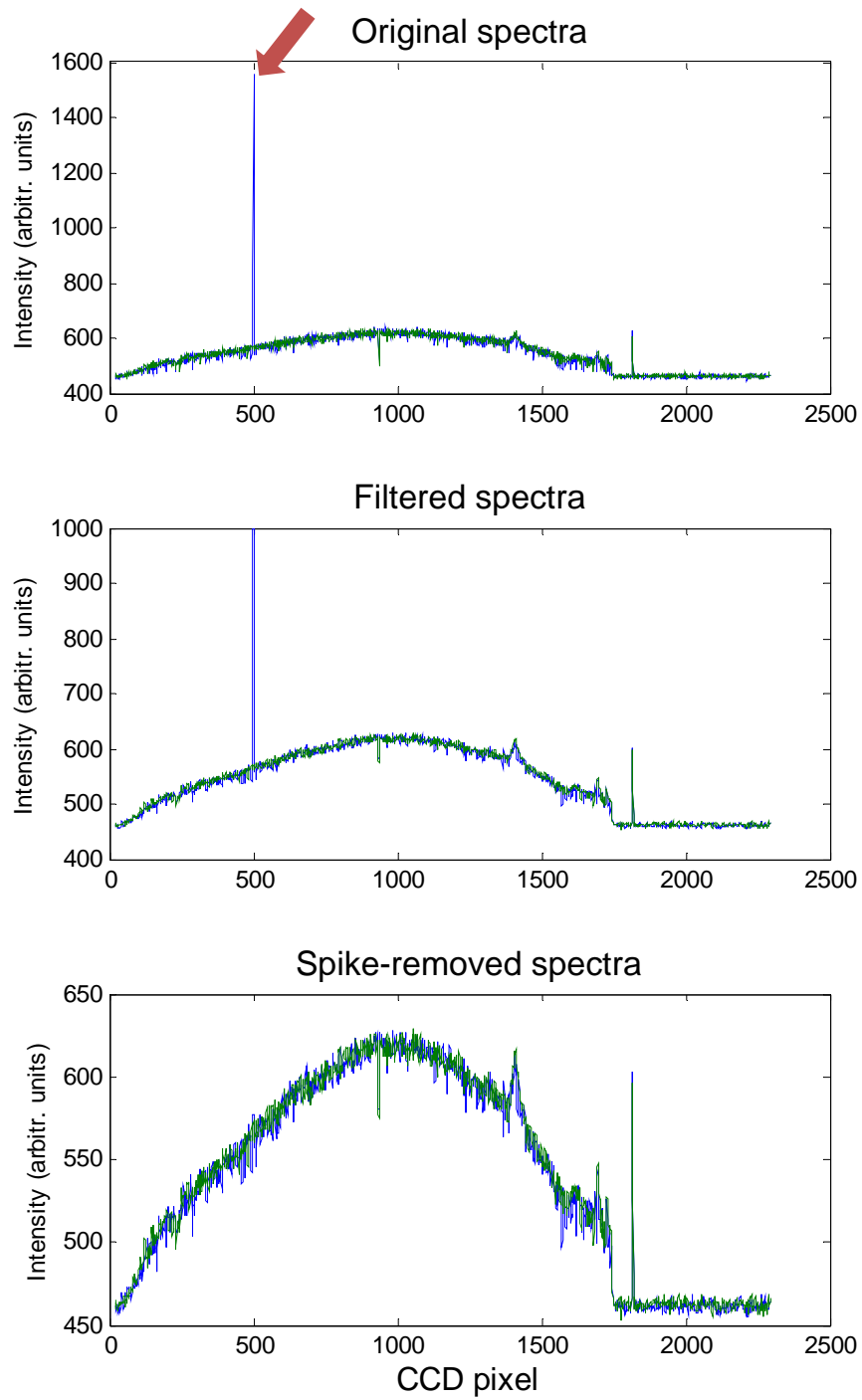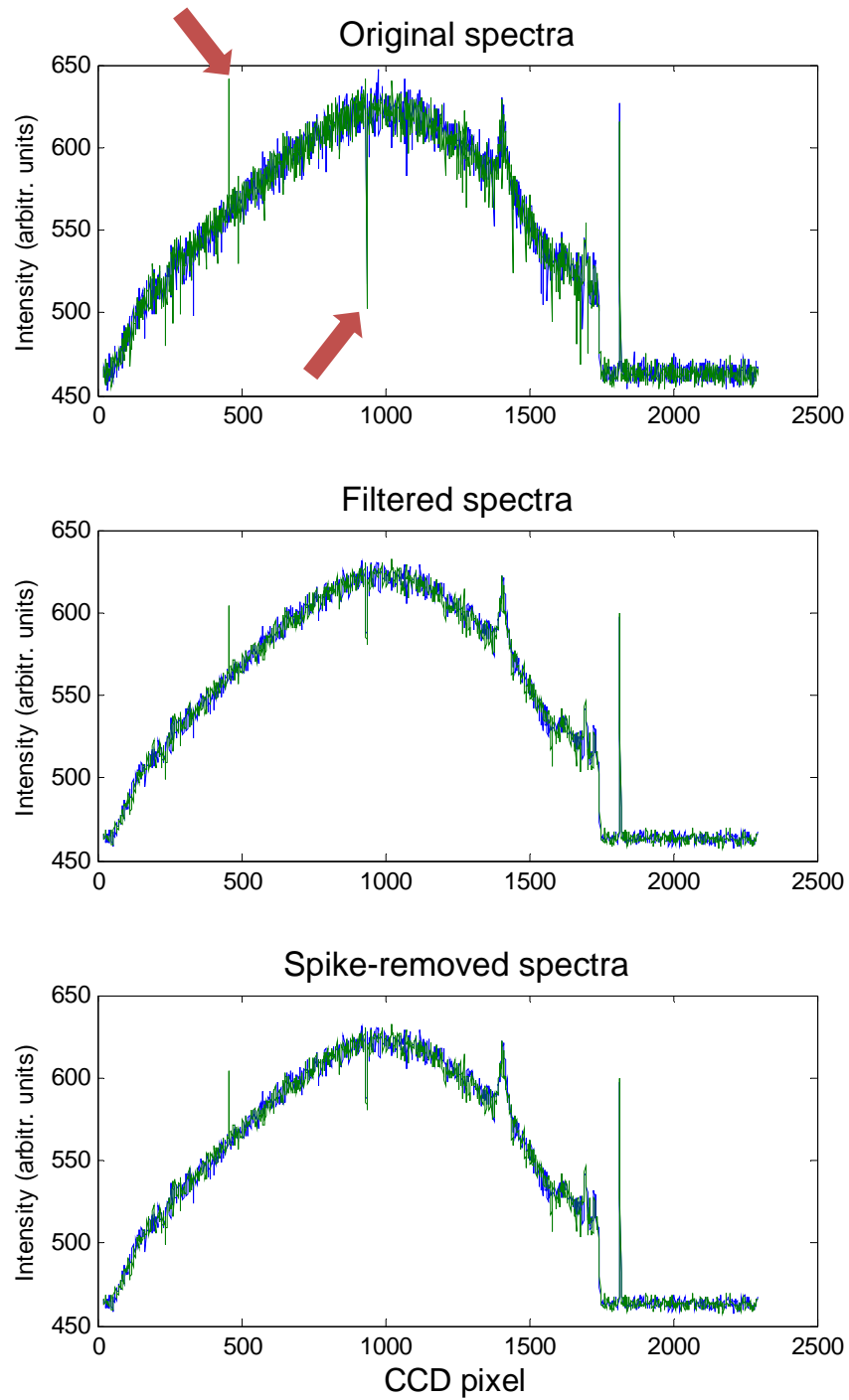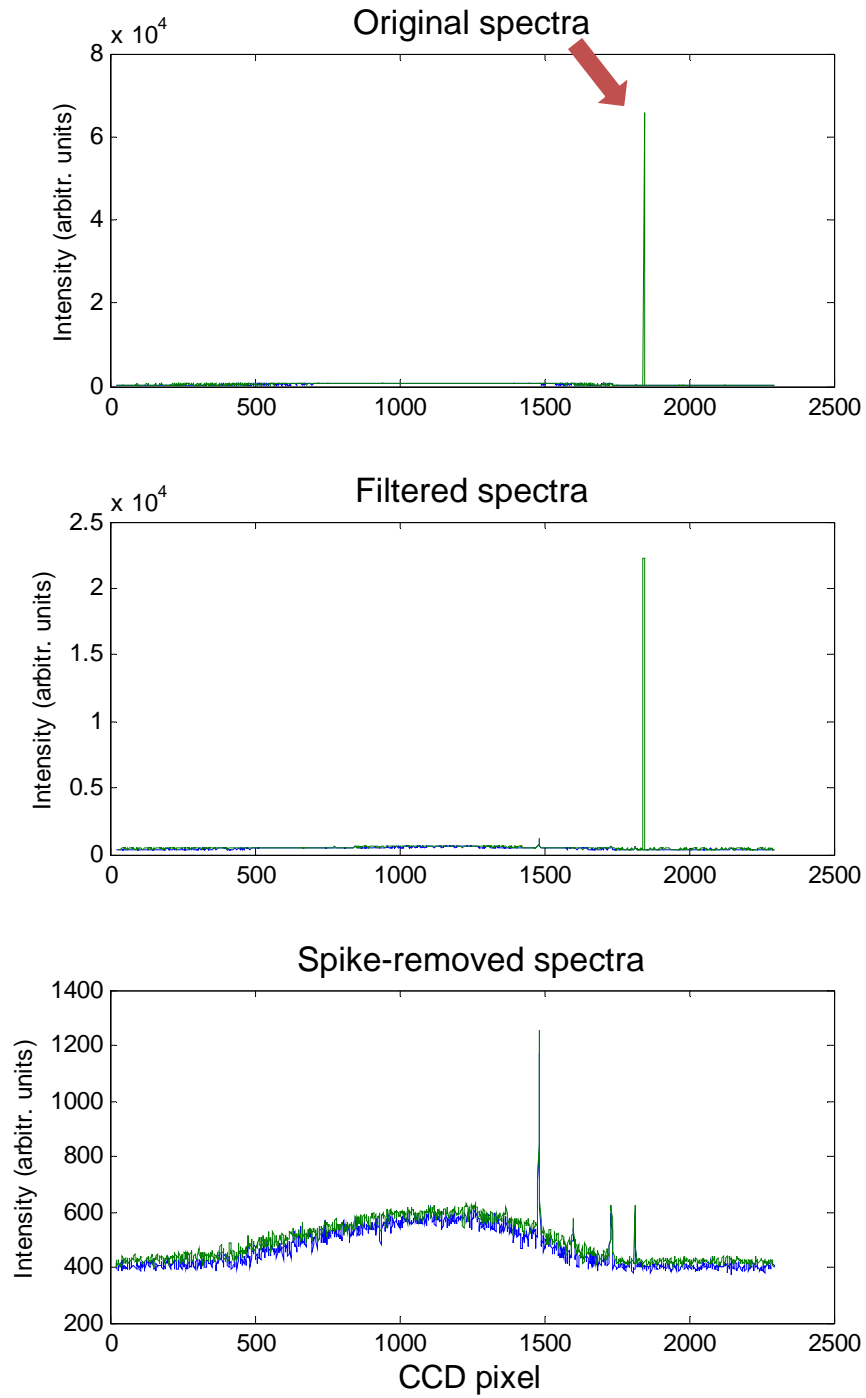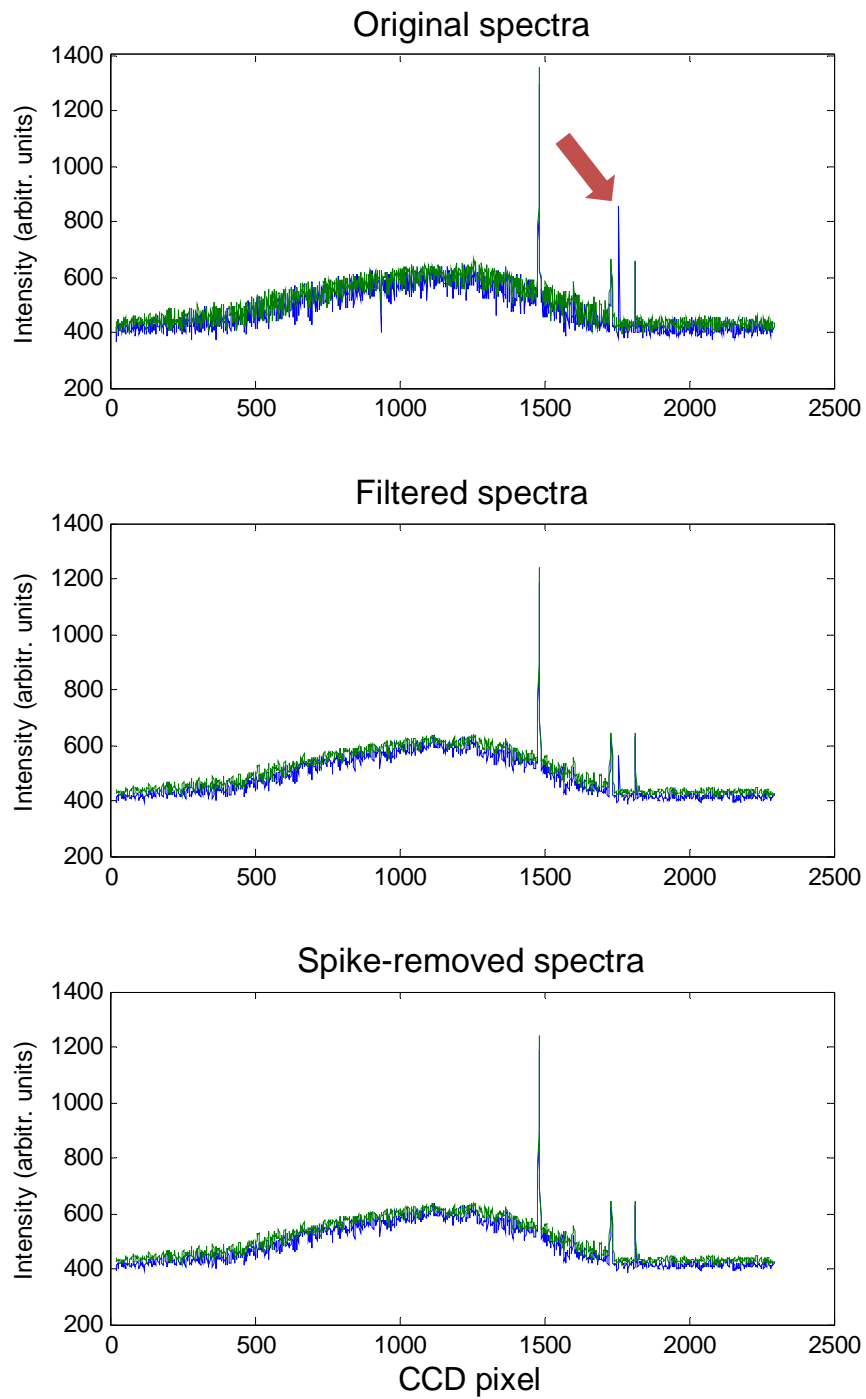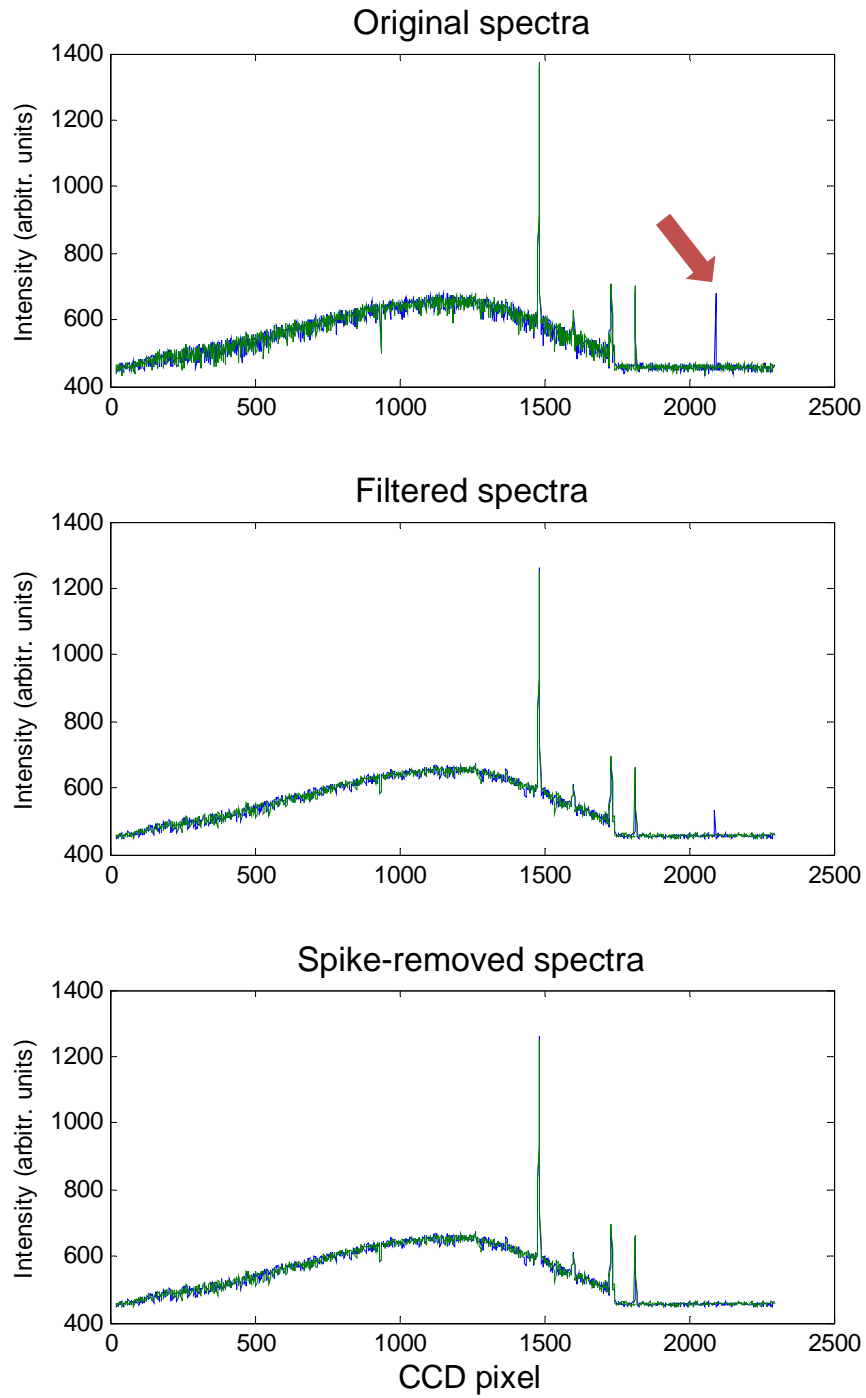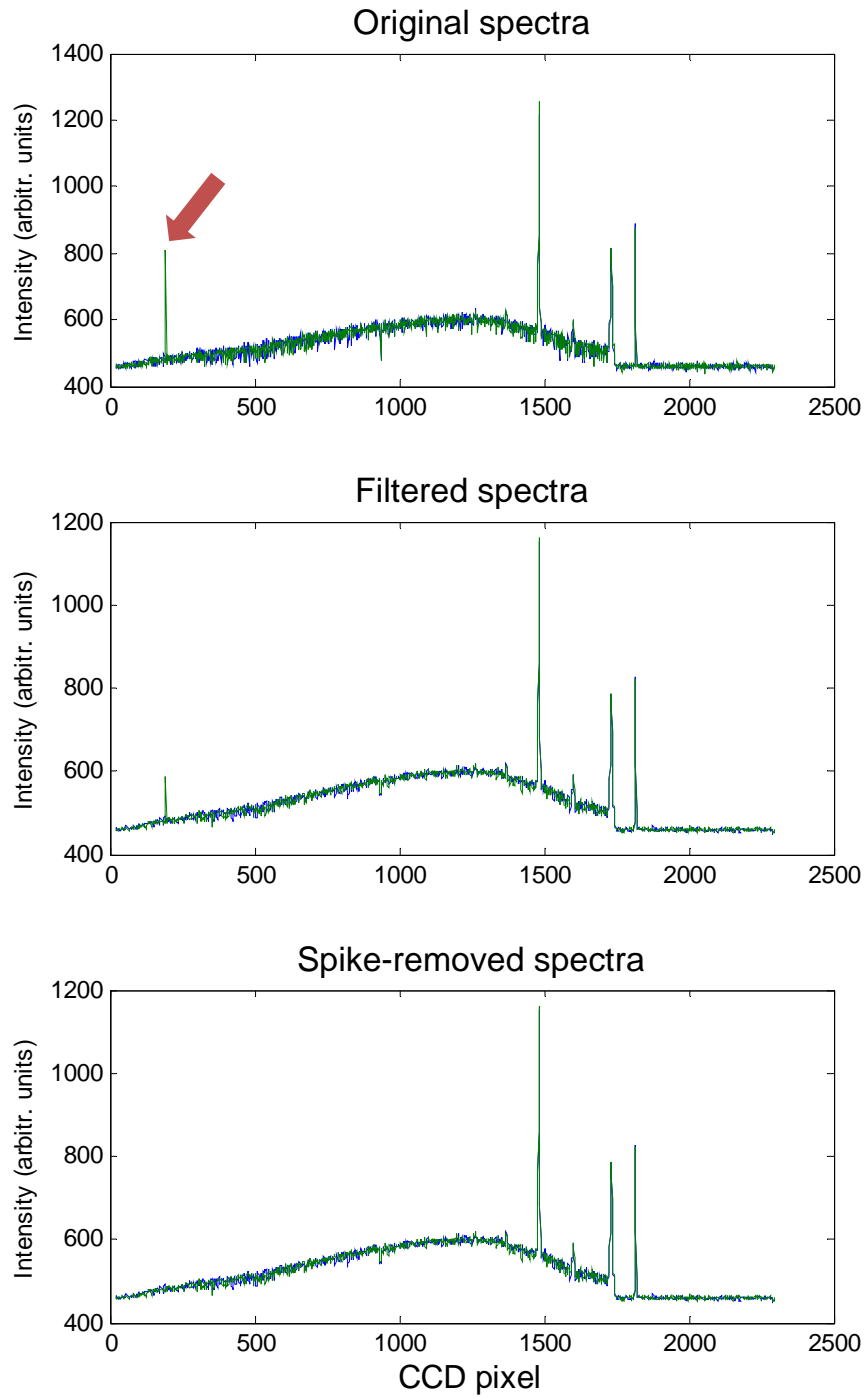
Figure B-16. Comparative of the raw 1 s reference spectra, the filtered version and the results after application of the cosmic ray removal algorithm for spectra from sample Calcite_P1.

Appendix C.

RESUMEN DE LA TESIS EN ESPAÑOL

Este apéndice presenta un resumen de la tesis en el idioma español, describiendo los objetivos, metodología y los principales resultados del trabajo.

La exploración in situ de Marte, junto con el análisis remoto del planeta rojo, es clave para entender su evolución geológica y establecer si el agua y la vida ha existido (o todavía existe) en la superficie marciana. El instrumento *Raman Laser Spectrometer* (RLS), parte de la carga útil de la misión ExoMars de la ESA, será el primer espectrómetro Raman en ser enviado en una misión interplanetaria.

El potencial de la espectroscopia Raman como un procedimiento no invasivo aplicado al análisis de todo tipo de materiales convierte a la espectroscopia Raman en una técnica muy importante en todo tipo de materiales, minerales o análisis biológicos. Esta técnica espectroscópica no tiene precedentes en la exploración planetaria, por lo que ha sido tradicionalmente operada mediante la intervención de un operador humano.

El modo de funcionamiento del instrumento RLS se entiende como la serie de operaciones y algoritmos que el instrumento tiene que realizar de forma autónoma para la obtención de un espectro Raman de la más alta calidad, sin la intermediación de un espectroscopista. Esto no debe confundirse con el desarrollo de los algoritmos necesarios para el instrumento se convierta en un instrumento de vuelo (telemetría, comunicación con el rover, conversión de datos, etc.). La definición del modo de funcionamiento es una tarea puramente científica que traslada las acciones que el operador humano realiza con el instrumento (enfoque del láser en la muestra, el ajuste de tiempo de adquisición del espectro, la eliminación de la fluorescencia...) a un sistema capaz de decidir de forma autónoma cuál es el punto de enfoque, de detectar si la muestra estudiada es fluorescente o no, y por lo tanto requiere un tratamiento especial, de calcular el tiempo óptimo de integración, etc. La definición de todos estos algoritmos se basa en experimentos científicos que ayudan a definir los parámetros objetivos que puedan ser medidos por el instrumento, de forma que pueda llevar a cabo de forma autónoma las mismas tareas que el operador humano realiza en base a su experiencia.

Para el desarrollo y optimización de estas herramientas algorítmicas, ha sido necesario desarrollar una plataforma de ejecución adecuada. Uno de los pasos previos necesarios para llevar a cabo esta tesis fue proporcionar al equipo de ciencia una plataforma en la que llevar a cabo los experimentos en un escenario operacional similar al del *rover*. Es decir, uno en el que se tiene en cuenta cómo el instrumento RLS interactuará con el *rover* durante la misión, sentando las bases para el desarrollo del modo de operación. Para cumplir este pre-objetivo de la tesis, hemos desarrollado el llamado simulador RLS ExoMars, que ha permitido realizar muchos experimentos científicos en condiciones similares a las proporcionadas por el Laboratorio Analítico (ALD) del *rover* ExoMars. Esto incluye la posibilidad de analizar las muestras en condiciones de laboratorio, pero también

en condiciones marcianas cuando trabaja en conjunción con la Cámara de simulación marciana (cuya descripción, diseño y los resultados se describen en detalle en un proyecto de tesis diferente).

Gracias a este simulador ha sido posible definir cómo adquirir espectros de forma autónoma, sin la presencia de un operador humano. Hasta ahora, algunos equipos Raman comerciales son capaces de trabajar de una manera pseudo-automática [3] mediante la realización de una serie de mediciones en diferentes puntos en una muestra, aunque el tratamiento de todos los puntos se hace de igual manera. El uso del simulador RLS ExoMars ha permitido la definición de un sistema de adquisición automático que optimiza los parámetros en cada uno de los puntos, optimizando así el uso del tiempo y los recursos disponibles.

Por otro lado, el uso de este sistema ha permitido la realización de experimentos que han ayudado a entender las posibilidades que el instrumento RLS ofrece en el marco de la exploración planetaria, con el fin de maximizar el retorno científico de los productos resultantes de la operación en Marte. En particular, el desarrollo del simulador RLS ExoMars ha allanado el camino para el estudio sistemático de la operación en muestras en polvo, despejando el camino para el futuro desarrollo de herramientas analíticas.

Estas herramientas formarán parte de la Herramienta de análisis de datos del instrumento (*Instrument Data Analysis* Tool, IDAT), un elemento del segmento de tierra que se utilizará durante la fase operacional de la misión, no sólo para la evaluación de la salud del instrumento, sino también para explotación de ciencia a largo plazo. El uso a corto plazo de IDAT se basará en proporcionar diariamente resultados científicos analizados muy rápidamente y de la forma lo más autónoma posible con el fin de permitir al equipo científico que opera el *rover* el tomar decisiones tácticas para las siguientes actividades del *rover*. La posibilidad de tener esta información analizada en un tiempo muy corto (minutos) después de la recepción de los datos del *rover* es de suma importancia para tomar las mejores decisiones tácticas en cada momento.

En este marco, la presente tesis pretende evaluar las capacidades científicas de instrumento RLS, con **dos objetivos principales**.

En primer lugar, la automatización del instrumento y la optimización de la adquisición de espectros. Para ello, es necesario definir nuevos procedimientos algorítmicos que proporcionan el instrumento con la lógica de decisión autónoma con el fin de optimizar los recursos disponibles, mientras que el aumento de la calidad y cantidad de espectros adquiridos.

En segundo lugar, la explotación de los datos desde el instrumento RLS. Para lograr este objetivo, se han estudiado las consecuencias que tienen sobre los espectros las limitaciones impuestas por la operación en un entorno espacial, y, en concreto, en el *rover* de ExoMars (operación basada en muestras en polvo). Además, se han aplicado y evaluado diferentes técnicas analíticas para el análisis de los espectros adquiridos por el instrumento en estas condiciones operacionales. Esto tiene como objetivo no solamente la identificación de las muestras, sino también la cuantificación de su abundancia en mezclas sencillas, permitiendo así evaluar las capacidades y el alcance de este tipo de técnicas.

Estos dos objetivos están conectados y relacionados con el desarrollo del instrumento RLS sobre la base de la siguiente lógica: primero, es necesario preparar el instrumento RLS para el funcionamiento automatizado y la optimización de adquisición de espectros. Segundo, las limitaciones operacionales impuestas por la operación de espacio / *rover* han de ser tenidas en cuenta y estudiadas para optimizar el proceso general de adquisición. Por último, los datos obtenidos con el instrumento RLS (con sus limitaciones operacionales) se estudian con diferentes técnicas de análisis de datos. La finalidad es analizar y evaluar las posibilidades de explotación de datos, incluyendo análisis estadísticos, los cuales cobran todo el sentido debido a las circunstancias operacionales concretas de RLS, que analizará varios puntos de una misma muestra.

Este trabajo se ha dividido en dos partes, en relación con sus dos objetivos principales. Para ello, el capítulo 1 introduce la espectroscopia Raman y el instrumento RLS, mientras que el capítulo 2 describe el simulador RLS ExoMars, el sistema desarrollado como base para poder lograr los objetivos de la tesis. El capítulo 3 describe un conjunto de algoritmos propuestos para la ejecución automática de un ciclo de adquisición completo del instrumento RLS, incluyendo algoritmos para la detección y eliminación de efectos de saturación, de fluorescencia o rayos cósmicos, así como para el cálculo automático de los parámetros óptimos de adquisición, basándose en la calidad de espectros de referencia. En estas secciones se trata el primer objetivo de la tesis.

Los siguientes capítulos se centran en cuestiones relacionadas con el modo de operación de RLS, en un formato fuertemente apoyado por trabajos publicados en varias revistas como el *Journal of Raman Spectroscopy*, *European Journal of Mineralogy* o *American Mineralogist*. En el capítulo 4 se estudia el comportamiento del sistema de preparación y distribución muestras (SPDS) del *rover* para evaluar el impacto que pudiera tener en los espectros adquiridos. Además, se ha presentado un estudio de muestras con diferentes distribuciones de tamaño de grano con el fin de comprender más profundamente los

efectos que la pulverización de la muestra puede inducir en los espectros Raman. Finalmente, el capítulo 5 presenta resultados obtenidos mediante diferentes técnicas analíticas de los espectros adquiridos en condiciones representativas de la operación. Tanto simples técnicas estadísticas estándar, así como varias técnicas de análisis multivariante se utilizaron para evaluar el comportamiento de estas técnicas para el análisis de los espectros Raman.

Objetivo 1: Automatización del instrumento RLS

El simulador RLS ExoMars se ha desarrollado en la Unidad Asociada UVa-CSIC-CAB en Valladolid, en respuesta a la necesidad de operar el instrumento RLS en un escenario operacional similar al del *rover* (lo que supone analizar varios puntos de la misma muestra, presentada al instrumento en forma de polvo aplanado en superficie). Este simulador permite realizar de forma automática el aplanamiento y posicionamiento de la muestra, con una resolución espacial de 2,5 micras. El software de control proporciona flexibilidad total sobre el sistema, con el fin de definir cualquier secuencia de operación deseada. Esto resultó ser útil durante el proceso de minería de datos (aunque en algunos casos, dependiendo del hardware utilizado, no todas las funciones estarían disponibles, por ejemplo, con el *breadboard* del RLS). Con el fin de optimizar la adquisición automática de muestras se definieron, implementaron, parametrizaron, optimizaron y validaron los algoritmos necesarios para tratar con los fenómenos de fluorescencia, saturación, rayos cósmicos… Esto proporcionó al simulador RLS ExoMars la capacidad de optimizar los parámetros de adquisición en cada punto de cada muestra, basándose en las características espectrales de espectros de referencia.

La definición de los algoritmos automáticos para el instrumento siempre sigue la misma estructura de desarrollo: Primero, descripción, estudio y caracterización del problema físico y su influencia en los espectros Raman. Segundo, definición y propuesta de algoritmo para la resolución del problema. Tercero, parametrización (optimización) basado en datos de referencia y resultados esperados. Y por último, validación con datos representativos. Cada algoritmo se validó unitariamente, y todo el procedimiento de extremo a extremo se aplicó en el simulador RLS ExoMars.

El algoritmo de reducción de la fluorescencia es uno de los algoritmos más importantes para aplicar con el fin de ahorrar tiempo de operación durante la misión, ya que permitirá reducir el tiempo de eliminación de la fluorescencia en función de la respuesta de la muestra (basado en la tasa de decaimiento del fondo espectral en función del tiempo). Tener esta capacidad es muy importante, ya que el tiempo asignado para reducir la fluorescencia de las muestras es relativamente alto, si bien no todas las muestras

presentarán fluorescencia. Por otro lado, se espera una baja incidencia de fluorescencia en muestras marcianas. Sin embargo, esto no es un problema, ya que la implementación de este algoritmo no tendrá impacto en la operación con muestras no fluorescentes (no habrá prácticamente consumo de tiempo en estos casos).

El funcionamiento básico del algoritmo de ajuste de los parámetros de adquisición se basa en el cálculo de los parámetros espectrales de espectros de referencia adquiridos con anterioridad a la adquisición final. Estos parámetros (por ejemplo, la calidad espectral) serán utilizados por los algoritmos para tomar decisiones sobre los parámetros de adquisición finales. Por tanto, espurios u otros picos de alta intensidad no relacionados con la emisión Raman como los rayos cósmicos podrían ser críticos para el correcto funcionamiento de los algoritmos, ya que podrían alterar el valor real de calidad espectral). Con el fin de minimizar el efecto de este tipo de picos, se ha creado un algoritmo específico para la detección y eliminación de espurios. Este algoritmo se basa en la comparación de dos espectros adquiridos consecutivamente, detectando todas las características espectrales que los diferencian. Como resultado, muchos de los espurios detectados durante nuestras campañas de prueba fueron corregidos, mejorando el comportamiento extremo a extremo del sistema.

El ajuste del tiempo de integración (Ti) y el número de acumulaciones (NA) de la adquisición final del espectro se basa ajustar Ti para llenar el 80% del rango dinámico de la CCD, mientras que el número de acumulaciones se decide en función de la calidad espectral de los espectros de referencia (cuanto más alta sea la relación señal ruido, menor será el número necesario de acumulaciones). Este algoritmo es relativamente complejo, ya que requiere el cálculo de la línea base del espectro a fin de definir su SNR. La parametrización de este algoritmo requiere la caracterización del instrumento RLS con diferentes tiempos de integración, y con respecto al número de acumulaciones. Esto demostró que, para diferentes tipos de muestras, el aumento del tiempo de integración incrementa más la SNR que aumentar el número de acumulaciones. En base a esta caracterización, fue posible definir tablas de consulta que permitirán un cálculo muy rápido de los parámetros de adquisición a bordo del *rover*, minimizando en la medida de lo posible los tiempos y recursos de procesamiento.

Toda la algorítmica definida en esta tesis está diseñada para trabajar de forma secuencial, y de la manera más eficiente posible desde el punto de vista computacional (lo cual es fundamental para la ejecución a bordo). Incluso si los algoritmos son relativamente independientes, a fin de ahorrar esfuerzo computacional, algunas partes comunes entre los diferentes bloques se comparten y ejecutan sólo una vez.

Objetivo 2: Rendimiento científico basado en productos de datos de RLS

La segunda parte de la tesis se centra en el análisis de los productos de datos obtenidos con el instrumento de RLS, en las condiciones operacionales impuestas por el *rover*, con el fin de evaluar el alcance de los resultados científicos que se pueden extraer a partir de los datos de RLS. Para ello, se realizaron pruebas para evaluar el impacto de trabajar en material triturado, y se aplicaron técnicas de análisis a los datos con el fin de evaluar las capacidades de explotación científicas. Las pruebas realizadas con el sistema de preparación y distribución de muestras (SPDS) permitió estudiar varios aspectos clave de las muestras y su entrega a los instrumentos: entre otras cosas, la rugosidad de la superficie aplanada, la contaminación cruzada inducida por el SPDS y las posibilidades de operación colaborativa entre MICROMEGA y RLS, durante la cual el SPDS colocaría una zona interesante de la muestra encontrada por MICROMEGA en la posición adecuada para que RLS pudiera analizarlo. Estas pruebas mostraron que la superficie aplanada proporcionada por el SPDS es aceptable por el instrumento RLS, siempre y cuando el mecanismo de enfoque automático de RLS esté disponible. La contaminación cruzada entre diferentes muestras provocada por el SPDS, aunque existente, puede considerarse insignificante en comparación con la contaminación cruzada que se espera obtener del mecanismo de trituración de la muestra. Por último, el test realizado para simular la operación colaborativa entre MICROMEGA y RLS mostró que es factible realizarla. La aplicación de esta clase de operación durante la misión mejoraría el funcionamiento global de ExoMars al adquirir espectros en las regiones de interés detectadas por otros instrumentos. Sin embargo, este test puso de relieve algunas cuestiones que han de ser tenidas en cuenta durante la operación en Marte. A saber, la modificación de la morfología de la superficie de la muestra debido a posibles vibraciones en el sistema, lo que podría poner en peligro la precisión en el posicionamiento de las muestras. Además, esta tesis muestra cómo el análisis de muestras en polvo puede inducir algunos efectos en los espectros Raman de las muestras, como por ejemplo un aumento general en el nivel de fondo. También se han abordado otros efectos en los espectros tales como desplazamiento o ensanchamiento de picos debido al análisis en polvo. Sin embargo, el nivel de irradiancia y las características de resolución espectral de RLS hacen que la mayoría de estos efectos sean despreciables e inobservables por el instrumento RLS.

Esta tesis también abarca otras cuestiones relacionadas con el proceso de trituración de la muestra, como la pérdida de textura y el contexto de la muestra masiva. Este es un asunto de importancia, por ejemplo, cuando se trata de muestras de carbón de origen potencialmente biológico, ya que en un caso así la morfología y textura de la muestra es clave para evaluar si su origen es biológico o no. Sin embargo, el proceso de trituración de

la muestra puede facilitar la detección por parte de RLS de trazas minoritarias (que podrían encontrarse en el interior de la muestra masiva, por ejemplo), como es el caso para el carbón en un *chert,* y para la ankerita detectada en un depósito de carbonatos de Svalbard. Por otro lado, los análisis de mezclas de polvo con proporciones controladas mostraron que el límite de detección del instrumento puede llegar a valores del 1% (e incluso menor, dependiendo de las muestras). Por tanto, a pesar de las limitaciones del análisis en polvo, estos resultados muestran que la combinación de los datos de RLS con los de otros instrumentos del *rover*, como CLUPI, puede proporcionar información científica relevante de las muestras analizadas.

La operación basada en análisis en muestras en polvo proporcionará espectros de al menos 20 puntos de la misma muestra. Por lo tanto, parece lógico realizar análisis estadísticos de los datos. En esta tesis se han descrito y aplicado varias técnicas para el análisis estadístico de los datos de RLS, divididas en dos tipos: el análisis basado en una variable y basado en múltiples variables (multivariante) multivariado. Se aplicó una técnica univariante (en oposición a multivariante) consistente en el cálculo del ratio entre los picos principales de las muestras presentes en mezclas binarias. Los resultados muestran que este cálculo tan simple proporciona una respuesta muy lineal en relación a la abundancia real de los minerales presentes en las mezclas. Esta linealidad se sustenta en el ajuste automático de los tiempos de adquisición, que compensa las diferentes eficiencias de emisión Raman de los distintos materiales (la adquisición automática funciona como una normalización de facto entre todos los espectros durante la adquisición).

Por otro lado, se han probado y comparado varias técnicas de análisis multivariante, basándose en mezclas binarias de muestras controladas. Las técnicas utilizadas fueron *Principal Component Analysis* (PCA), *Partial Least Square Regression* (PLS) y Redes Neuronales (ANN). Todas estas técnicas tienen características diferentes que son descritas y comparadas. Además, se probaron con el fin de evaluar su capacidad para identificar y / o cuantificar la abundancia de minerales de las muestras bajo análisis.

Los resultados mostraron que PCA es capaz de separar gradualmente los componentes de las mezclas en función de su abundancia; PLS también proporciona curvas de calibración de las diferentes mezclas y, por último, los modelos creados con ANN muestran ser capaces de identificar y cuantificar la abundancia de fases minerales con bastante precisión. Lo interesante de todas estas técnicas estadísticas y multivariantes es que los modelos entrenados parecen ser válidos para los espectros adquiridos con diferentes montajes experimentales y/o diferentes muestras del mismo tipo.

Todos estos resultados muestran que, a pesar de la pérdida del contexto y textura de las muestras durante el proceso de trituración, el uso de esta clase de técnicas podría mejorar los límites de detección de RLS, además de permitir la pseudo-cuantificación de la abundancia mineral en mezclas heterogéneas.

Aunque los resultados presentados en este proyecto de tesis se basan en análisis de muestras en polvo adquiridos en condiciones ambientales terrestres, estos no son invalidados por análisis realizados en condiciones marcianas, y por tanto estas técnicas también serán aplicables en dichas condiciones. Sin embargo, los resultados en condiciones de presión y ambiente marcianos, así como el diseño y construcción de la cámara de simulación de ambiente marciano no son parte de los objetivos de esta tesis.

ANNEX.

SCIENTIFIC PRODUCTION

This annex presents the scientific publications (in conferences and journals) that have been produced during the development of this thesis.

Peer-reviewed journals

Lopez-Reyes, G., Sobron, P., Lefebvre, C., and Rull, F. (2014) Multivariate analysis of Raman spectra for the identification of sulfates: Implications for ExoMars. American Mineralogist, 99(8-9), 1570-1579.

Lopez-Reyes, G., Sobron, P., Lefebvre, C., and Rull, F. (2014) Application of Multivariate Analysis Techniques for the Identification of Sulfates From Raman Spectra. In E. Pardo-Igúzquiza, C. Guardiola-Albert, J. Heredia, L. Moreno-Merino, J.J. Durán, and J.A. Vargas-Guzmán, Eds. Mathematics of Planet Earth, p. 135-139. Springer Berlin Heidelberg.

Hermosilla Rodriguez, I., Lopez-Reyes, G., Llanos, D.R., and Rull Perez, F. (2014) Automatic Raman Spectra Processing for Exomars. In E. Pardo-Igúzquiza, C. Guardiola-Albert, J. Heredia, L. Moreno-Merino, J.J. Durán, and J.A. Vargas-Guzmán, Eds. Mathematics of Planet Earth, p. 127-130. Springer Berlin Heidelberg.

Lopez-Reyes, G., Rull, F., Venegas, G., Westall, F., Foucher, F., Bost, N., Sanz, A., Catalá-Espí, A., Vegas, A., Hermosilla, I., Sansano, A., and Medina, J. (2013) Analysis of the scientific capabilities of the ExoMars Raman Laser Spectrometer instrument. European Journal of Mineralogy, 25(5), 721-733.

Foucher, F., Lopez-Reyes, G., Bost, N., Rull-Perez, F., Rüßmann, P., and Westall, F. (2013) Effect of grain size distribution on Raman analyses and the consequences for *in situ* planetary missions. Journal of Raman Spectroscopy, 44(6), 916-925.

Contributions to conferences

Lopez-Reyes, G., Peña Nogales, O., Perez Oliveros, A., Sanz, A., Rull, F. (2014) Automated and self-regulated raman spectra acquisition for space exploration: exomars RLS instrument. GeoRaman XI, GeoRaman book of abstracts, p. 47-48, St. Louis. – **Awarded with the "Best Graduate Student Poster Paper Prize".**

Manrique-Martínez, J.A., Sansano, A., Navarro, R. J., Lopez-Reyes, G., Rull, F., Gomez, J.M. (2014) Remote Raman detection of biomarkers on mineralogical substrates. GeoRaman XI, GeoRaman book of abstracts, p. 85-86, St. Louis.

Foucher, F., Lopez-Reyes, G., Bost, N., Rull-Perez, F., Rüßmann, P., Westall, F., Vago, J., Viso, M. (2014) Strategy for the geological exploration of mars during the exomars mission: effect of sample preparation on raman spectroscopy measurments and payload calibration. . GeoRaman XI, GeoRaman book of abstracts, p. 5-6, St. Louis.

Sansano, A., Navarro, R., Medina, J., Manrique-Martínez, J. A., Lopez-Reyes, G., Rull, F. (2014) Raman study of arctic volcanic Mars analogs. In-situ and laboratory analysis. GeoRaman XI, GeoRaman book of abstracts, p. 51-52, St. Louis.

Sobron, P., Lopez-Reyes, G., Sansano, A., Manrique, J.A., and Rull, F. (2014) Data Fusion in Planetary LIBS+Raman Spectroscopy. 45th Lunar and Planetary Science Conference, p. Abstract #2875. Lunar and Planetary Institute, Houston.

Sansano, A., Navarro, R., López-Reyes, G., and Rull, F. (2014) Development of the Calibration Target for Exomars' Raman Instrument (RLS). 45th Lunar and Planetary Science Conference, p. Abstract #2623. Lunar and Planetary Institute, Houston.

Sobron, P., Lopez-Reyes, G., and Wang, A. (2013) Automated Mineral Identification in Three Mars Analogue Sites Using in-situ NIR Reflectance Spectroscopy and Linear Spectral Unmixing. Analog Sites for Mars Missions II: past, present and future missions to Mars, Washington, DC.

Rull, F., Maurice, S., Diaz, E., Lopez, G., and Catala, A. (2013) Raman Laser Spectrometer (RLS) for ExoMars 2018 Rover Mission: Current Status and Science Operation Mode on Powdered Samples. 44th Lunar and Planetary Science Conference, p. Abstract #3110. Lunar and Planetary Institute, Houston.

Lopez-Reyes, G., Sobron, P., Lefevbre, C., and Rull, F. (2013) Application of Multivariate Analysis Techniques for the Identification of Sulfates from Raman Spectra   Implications for Exomars. 44th Lunar and Planetary Science Conference, p. Abstract #2135. Lunar and Planetary Institute, Houston.

Lopez-Reyes, G., Sobron, P., Lefevbre, C. and Rull, F. (2013) Application of multivariate analysis techniques for the identification of sulfates from Raman spectra. 15th International Association for Mathematical Geosciences Conference, Madrid.

Lopez-Reyes, G. (2013) Application of multivariate analysis techniques for the identification of sulfates from Raman spectra for ExoMars. AbGradCon 2013, Montreal.

Lalla, E., Lopez-Reyes, G., Rull, F., Martinez-Frías, J., Sansano, A., Navarro, R., and Medina, J. (2013) Raman Analysis of Basaltic Samples from Tenerife Island (Cañadas, Azulejos, and Historical Eruptions) with the Exomars RLS Instrument. 44th Lunar and Planetary Science Conference, p. Abstract #2403. Lunar and Planetary Institute, Houston.

Hermosilla, I., Lopez-Reyes, Llanos, D. R., and Rull, F. (2013) Automatic Raman Spectra Processing for Exomars. 15th International Association for Mathematical Geosciences Conference, Madrid.

Diaz, C., Lopez-Reyes, G., Hermosilla, I., Catalá-Espí, A., Sansano, A., Perez, C., Diaz, E., Rull, F., Margolles,R. (2013) Raman Laser Spectrometer Adaptative Operation for Mars Exploration. 64th International Astronautical Congress IAC, Beijing.

Bost, N., Ramboz, C., Le Breton, N., Lopez-Reyes, G., Pilorget, C., De Angelis, S., Foucher, F., and Westall, F. (2013) A Blind Test to test the future ExoMars instruments. European Planetary Science Congress 2013, p. 662.

Rull, F., Maurice, S., Diaz, E., and Lopez-Reyes, G.p.a. (2012) Raman spectroscopy for the 2018 ExoMars mission. European Planetary Science Congress 2012, p. 740.

Lopez-Reyes, G., Sobron, P., Lafuente, B., and Rull, F. (2012) Application of multivariate analysis techniques for the identification of sulphates from Raman spectra - Implication for ExoMars. European Planetary Science Congress 2012, p. 577.

Lopez-Reyes, G., Rull, F., Catala, A., Sanz, A., Medina, J., Hermosilla, I., Lafuente, B. (2012) A simple statistical method for the pseudo-quantification of mineral phases within the ExoMars Raman RLS instrument. GeoRaman Xth, GeoRaman book of abstracts, p. 151-152, Nancy.

Lafuente, B., Sansano, A., Navarro, R., Rull, F., Martínez Frías, J., Medina, J., Lopez, G., Sobron, P. (2012) Multivariate analysis of Raman spectra for geological classification and identification: Application to Exomars Mission. GeoRaman Xth, GeoRaman book of abstracts, p. 149-150, Nancy.

Hermosilla, I., Lopez-Reyes, G., Catalá, A., Sanz, A., Llanos, D.R., and Rull, F. (2012) Raman spectra processing algorithms and database for RLS-ExoMars. European Planetary Science Congress 2012, p. 567.

Foucher, F., Westall, F., Bost, N., Rull, F., Lopez-Reyes, G., and Rüßmann, P. (2012) Effect of the crushing process on Raman analyses: consequences for the Mars 2018 mission. 39th COSPAR Scientific Assembly, 39, p. 555.

Foucher, F., Lopez-Reyes, G., Bost, N., Rull, F., Rüßmann, P. and Westall, F. (2012) Effect of the crushing process on Raman analyses: consequences for the Mars 2018 mission. GeoRaman Xth, GeoRaman book of abstracts, p. 143-144, Nancy.

Foucher, F., Lopez-Reyes, G., Bost, N., Rull, F., Rüßmann, P. and Westall, F. (2012) Effect of the crushing process on Raman analyses: consequences for the Mars 2018 mission. EGU General Assembly 2012, Geophysical Research Abstracts. Vol. 14. EGU2012-4318.

Sansano, A., López, G., Rull, F., and Team, A. (2011) Activities of Exomars Raman Laser Spectrometer Scientific Team During the Campaign AMASE 2010. 42nd Lunar and Planetary Science Conference, p. Abstract #2469. Lunar and Planetary Institute, Houston.

Sansano, A., Lopez-Reyes, G., Medina, J., Rull, F., and AMASE'10 Team. (2011) Analysis of Arctic Carbonates Profiles by Raman Spectroscopy using Exomars's RLS set up European Planetary Science Congress EPSC-DPS, Vol. 6, EPSC-DPS2011-856, Nantes.

Rull, F., Maurice, S., Diaz, E., Tato, C., Pacros, A., and Team., t.R. (2011) The Raman Laser Spectrometer (RLS) on the EXOMARS 2018 Rover Mission. 42nd Lunar and Planetary Science Conference LPSC, p. Abstract #2400. Lunar and Planetary Institute, Houston.

Rull, F., Lopez, G., Catala, A., Medina, J., Sansano, A., Sanz, A., Sobron, F. (2011) Raman Spectroscopy Analysis on Powdered Samples inside the Exomars Mission. The International Conference: Exploring Mars Habitability, Lisbon (Portugal).

Lalla, E., Sansano, A., Sanz, A., Navarro, R., Lopez, G., Venegas, G., Rodriguez Losada, J. A., Medina Garcia, J., Martinez-Frias, J., Rull, F. . (2011) Raman Spectroscopy of Pillow Lavas from the Anaga zone – Tenerife Canary Island. In MACLA, Ed. SEM.

Catalá-Espí, A., Lopez-Reyes, G., Rull, F., Vegas, A. . (2011) Design of a small Martian chamber for simulating the RLS-ExoMars operation mode. European Planetary Science Congress EPSC, EPSC-DPS2011-708.

Rull, F., Lopez, G., Catala A., Sanz A., Sansano, A., Medina J., Martinez-Frias, J. . (2010) Spectroscopic Raman study of powdered mineral samples in automatic mode for ExoMars mission operation simulation. . EGU General Assembly 2010, Geophysical Research Abstracts Vol. 12, .

Rull, F., Catala, A., Lopez, G., Sansano, A., Sanz, A. . (2010) Capabilities of detection on crushed samples by Raman Spectroscopy: Exomars Raman Instrument. . GeoRaman 2010, Book of abstracts.

References

1. Rull, F., et al. *The Raman Laser Spectrometer (RLS) on the EXOMARS 2018 Rover Mission.* in *42nd Lunar and Planetary Science Conference LPSC.* 2011. Houston: Lunar and Planetary Institute.

2. Rull, F.L., G. ; Catala, A. ; Medina, J. ; Sansano, A. ; Sanz, A. ; Sobron, F. . *Raman Spectroscopy Analysis on Powdered Samples inside the Exomars Mission.* in *The International Conference: Exploring Mars Habitability, .* 2011. Lisbon (Portugal).

3. WITec. *WITec Alpha 300R Confocal Raman Microscope. Specifications brochure.* 2012; Available from: http://www.witec.de/en/products/raman/alpha300r/.

4. ESA. *Aurora program web page at www.esa.int.* 2006; Available from: http://www.esa.int/esaMI/Aurora/SEMZOS39ZAD_0.html.

5. ESA. *Exomars project web page at www.esa.int.* 2012; Available from: http://exploration.esa.int/science-e/www/object/index.cfm?fobjectid=46048.

6. Catalá-Espí, A., *Operation modes and spectral data treatment in the Raman instrument for ExoMars*, in *Departamento de Astrofisica.* 2010, Universidad de La Laguna: Tenerife.

7. ESA. *The Exomars Rover Instrument Suite page at www.esa.int.* 2014; Available from: http://exploration.esa.int/mars/45103-rover-instruments/.

8. Pacros, A., Haldemann, A., Vago, J. and the ESA ExoMars Project Team. *Overview of the ExoMars Payload.* in *European Planetary Science Congress EPSC.* 2009.

9. Rull, F., Martinez-Frías, J. , *Raman Spectroscopy goes to Mars. .* Spectroscopy Europe, 2006. **1818**(1).

10. Downs, R.T., *The RRUFF Project: an integrated study of the chemistry, crystallography, Raman and infrared spectroscopy of minerals. .* Program and Abstracts of the 19th General Meeting of the International Mineralogical Association in Kobe, Japan., 2006: p. O03-13.

11. Raman, C.V. and K.S. Krishnan, *A new type of secondary radiation.* Nature, 1928. **121**(3048): p. 501-502.

12. Zhang, Z.-M., et al., *An intelligent background-correction algorithm for highly fluorescent samples in Raman spectroscopy.* Journal of Raman Spectroscopy, 2010. **41**(6): p. 659-669.

13. Quirico, E., et al., *Precursor and metamorphic condition effects on Raman spectra of poorly ordered carbonaceous matter in chondrites and coals.* Earth and Planetary Science Letters, 2009. **287**(1): p. 185-193.

14. Ferrari, A.C., *Raman spectroscopy of graphene and graphite: disorder, electron–phonon coupling, doping and nonadiabatic effects.* Solid State Communications, 2007. **143**(1): p. 47-57.

15. Steele, A., et al., *A reduced organic carbon component in martian basalts.* Science, 2012. **337**(6091): p. 212-215.

16.   Pasteris, J.D. and B. Wopenka, *Necessary, but not sufficient: Raman identification of disordered carbon as a signature of ancient life.* Astrobiology, 2003. **3**(4): p. 727-738.

17.   Marshall, C.P., H.G. Edwards, and J. Jehlicka, *Understanding the application of Raman spectroscopy to the detection of traces of life.* Astrobiology, 2010. **10**(2): p. 229-243.

18.   Schopf, J.W., et al., *Laser-Raman imagery of Earth's earliest fossils.* Nature, 2002. **416**(6876): p. 73-76.

19.   Schopf, J.W., et al., *Raman imagery: a new approach to assess the geochemical maturity and biogenicity of permineralized Precambrian fossils.* Astrobiology, 2005. **5**(3): p. 333-371.

20.   Culka, A., et al., *The detection of biomarkers in evaporite matrices using a portable Raman instrument under Alpine conditions.* Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 2011. **80**(1): p. 8-13.

21.   Culka, A., J. Jehlička, and L. Strnad, *Testing a portable Raman instrument: The detection of biomarkers in gypsum powdered matrix under gypsum crystals.* Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 2012. **86**: p. 347-350.

22.   Edwards, H.G., *Raman spectroscopic protocol for the molecular recognition of key biomarkers in astrobiological exploration.* Origins of Life and Evolution of the Biosphere, 2004. **34**(1-2): p. 3-11.

23.   Edwards, H.G., et al., *Raman spectroscopic analysis of geological and biogeological specimens of relevance to the ExoMars mission.* Astrobiology, 2013. **13**(6): p. 543-549.

24.   Jehlička, J., H. Edwards, and P. Vítek, *Assessment of Raman spectroscopy as a tool for the non-destructive identification of organic minerals and biomolecules for Mars studies.* Planetary and space Science, 2009. **57**(5): p. 606-613.

25.   Vandenabeele, P., et al., *On the definition of Raman spectroscopic detection limits for the analysis of biomarkers in solid matrices.* Planetary and Space Science, 2012. **62**(1): p. 48-54.

26.   Vítek, P., et al., *The miniaturized Raman system and detection of traces of life in halite from the Atacama Desert: some considerations for the search for life signatures on Mars.* Astrobiology, 2012. **12**(12): p. 1095-1099.

27.   Vítek, P., K. Osterrothová, and J. Jehlička, *Beta-carotene—a possible biomarker in the Martian evaporitic environment: Raman micro-spectroscopic study.* Planetary and Space Science, 2009. **57**(4): p. 454-459.

28.   Rull, F. *Raman spectroscopy for planetary exploration*. in *Georaman X 10*. 2012.

29.   Vago, J., Gardini, B., Baglioni, P., Kminek, G., Gianfiglio, D., and the ExoMars Project Team. *ExoMars: ESA's mission to search for signs of life on the red planet*. in *Lunar and Planetary Science Congress LPSC*. 2006.

30.   Sansano, A., Lopez-Reyes, G., Medina, J., Rull, F., and AMASE'10 Team. *Analysis of Arctic Carbonates Profiles by Raman Spectroscopy using Exomars's RLS set up* in *European Planetary Science Congress EPSC-DPS*. 2011. Nantes.

31.   Wang, A., et al., *Development of the Mars microbeam Raman spectrometer (MMRS).* J. Geophys. Res., 2003. **108**(E1): p. 5005.

32.   Wang, A., J.L. Lambert, and P.S. Sobron, *An Instrument Suite for Mineral ID and Biomarker Seeking in Atacama*, in *44th Lunar and Planetary Science Conference*. 2013, Lunar and Planetary Institute: Houston. p. Abstract #2586.

33.   Kong, W.G. and A. Wang, *Planetary Laser Raman Spectroscopy for Surface Exploration on C/D-Type Asteroids   A Case Study*, in *41st Lunar and Planetary Science Conference*. 2010, Lunar and Planetary Institute: Houston. p. Abstract #2730.

34.   Angel, S.M., et al., *Remote Raman Spectroscopy for Planetary Exploration: A Review.* Applied Spectroscopy, 2012. **66**(2): p. 137-150.

35.   Wiens, R.C., et al., *Combined Remote Raman Spectroscopy and LIBS Instrumentation for Mars Astrobiology Exploration*, in *Seventh International Conference on Mars*. 2007, Lunar and Planetary Institute: Houston. p. Abstract #3092.

36.   Sharma, S.K., et al., *Stand-off Raman spectroscopic detection of minerals on planetary surfaces.* Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 2003. **59**(10): p. 2391-2407.

37.   Richard, L.M., *Raman Spectroscopy for Chemical Analysis*. Measurement Science and Technology. Vol. 12. 2001. 653.

38.   Long, D.A., *The Raman Effect: A Unified Treatment of the Theory of Raman Scattering by Molecules*. The Raman Effect. 2002: John Wiley & Sons, Ltd.

39.   Ferraro, J.R., *Introductory Raman Spectroscopy*. 2003: Elsevier Science.

40.   Lakowicz, J., *Principles of Fluorescence Spectroscopy*. Principles of Fluorescence Spectroscopy. 1999, New York, Boston, Dordrecht, London, Moscow: Kluwer Academic/Plenum Publishers.

41.   Schulze, H.G. and R.F. Turner, *A Two-Dimensionally Coincident Second Difference Cosmic Ray Spike Removal Method for the Fully Automated Processing of Raman Spectra.* Applied Spectroscopy, 2014. **68**(2): p. 185-191.

42.   Schulze, H.G. and R.F.B. Turner, *A Fast, Automated, Polynomial-Based Cosmic Ray Spike-Removal Method for the High-Throughput Processing of Raman Spectra.* Applied Spectroscopy, 2013. **67**(4): p. 457-462.

43.   Katsumoto, Y. and Y. Ozaki, *Practical algorithm for reducing convex spike noises on a spectrum.* Applied spectroscopy, 2003. **57**(3): p. 317-322.

44.   Zhao, J., *Image Curvature Correction and Cosmic Removal for High-Throughput Dispersive Raman Spectroscopy.* Applied Spectroscopy, 2003. **57**(11): p. 1368-1375.

45.   Ehrentreich, F. and L. Sümmchen, *Spike removal and denoising of Raman spectra by wavelet transform methods.* Analytical chemistry, 2001. **73**(17): p. 4364-4373.

46.   Schulze, H.G., et al., *A Small-Window Moving Average-Based Fully Automated Baseline Estimation Method for Raman Spectra.* Applied Spectroscopy, 2012. **66**(7): p. 757-764.

47.   Schulze, H.G., et al., *A Model-Free, Fully Automated Baseline-Removal Method for Raman Spectra.* Applied Spectroscopy, 2011. **65**(1): p. 75-84.

48.   Schulze, G., et al., *Investigation of selected baseline removal techniques as candidates for automated implementation.* Applied Spectroscopy, 2005. **59**(5): p. 545-574.

49.   Sobron, P., et al., *Raman signal processing software for automated identification of mineral phases and biosignatures on Mars.* Applied Spectroscopy, 2008. **62**(4): p. 364-370.

50.   Weakley, A.T., P.R. Griffiths, and D.E. Aston, *Automatic Baseline Subtraction of Vibrational Spectra Using Minima Identification and Discrimination via Adaptive, Least-Squares Thresholding.* Applied Spectroscopy, 2012. **66**(5): p. 519-529.

51.   Shao, L. and P.R. Griffiths, *Automatic baseline correction by wavelet transform for quantitative open-path Fourier transform infrared spectroscopy.* Environmental Science & Technology, 2007. **41**(20): p. 7054-7059.

52.   Zhang, Z.-M., S. Chen, and Y.-Z. Liang, *Baseline correction using adaptive iteratively reweighted penalized least squares.* Analyst, 2010. **135**(5): p. 1138-1146.

53.   Lopez-Reyes, G., *Spectral processing for the pseudo-quantification of mineral species in spectra of the Raman Laser Spectrometer instrument of ESA ExoMars mission.* University of Valladolid (Spain), http://uvadoc.uva.es/handle/10324/2653, 2012.

54.   Foucher, F., et al., *Effect of grain size distribution on Raman analyses and the consequences for in situ planetary missions.* Journal of Raman Spectroscopy, 2013. **44**(6): p. 916-925.

55.   Bibring, J.-P., et al., *Mars Surface Diversity as Revealed by the OMEGA/Mars Express Observations.* Science, 2005. **307**(5715): p. 1576-1581.

56.   Lopez-Reyes, G., et al., *Analysis of the scientific capabilities of the ExoMars Raman Laser Spectrometer instrument.* European Journal of Mineralogy, 2013. **25**(5): p. 721-733.

57.   Lopez-Reyes, G., et al., *Multivariate analysis of Raman spectra for the identification of sulfates: Implications for ExoMars.* American Mineralogist, 2014. **99**(8-9): p. 1570-1579.

58.   Blake, D.F., et al., *Carbonate Cements from the Sverrefjell and Sigurdfjell Volcanos, Svalbard Norway: Analogs for Martian Carbonates*, in *42nd Lunar and Planetary Science Conference*. 2011, Lunar and Planetary Institute: Houston. p. Abstract #2167.