Application of Raman Spectroscopy for the characterization of carbon materials.

by

Bruno Guilherme da Fonseca

M. Sc., Universidade Federal de Juiz de Fora, Brazil, 2016

B. Sc., Universidade Federal de Juiz de Fora, Brazil, 2014

A Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree of

DOCTOR OF PHILOSOPHY

In the Department of Chemistry

We acknowledge and respect the lək^{w}əŋən peoples on whose traditional territory the university stands and the Songhees, Esquimalt and W̱SÁNEĆ peoples whose historical relationships with the land continue to this day.

Application of Raman Spectroscopy for the characterization of carbon materials.

by

Bruno Guilherme da Fonseca

M. Sc., Universidade Federal de Juiz de Fora, Brazil, 2016

B. Sc., Universidade Federal de Juiz de Fora, Brazil, 2014

Supervisory Committee

Dr. Alexandre G. Brolo, Supervisor

Department of Chemistry

Dr. Dennis Hore, member

Department of Chemistry

Dr. Jeremy Wulff, member

Department of Chemistry

Dr. Geoffrey Steeves, outside member

Department of Physics

Abstract

This dissertation discusses the applications of Raman spectroscopy for the characterization of carbon materials, through the investigation of heterogeneous hydrogenation, estimation of elemental and organic carbon from diesel emission particulate and automated classification. Carbon materials are present in several aspects of human life. Some illustrative examples come from technological use of graphite, diamond and glassy carbon and the impact of diesel emissions in human health. The general objective of the dissertation is to improve the characterization of carbon materials and show the potential and limitations of methods that combines Raman spectroscopy and chemometrics for carbon materials characterization.

The literature is vast about the utility of Raman spectroscopy for characterizing carbon materials. However, some methods are either not clearly described in terms of reproducibility and implementation, or very specific to a particular type of material. Also, some studies describe their material based in a single spectrum instead of a spatial and statistical characterization.

Chapter 2 demonstrates the usefulness of Raman maps for characterizing carbon films obtained chemically. The study investigated the heterogeneous hydrogenation in the samples and showed the effect of different parameters on the homogeneity of the material. Due to heterogeneous hydrogenation, some regions could be assigned to different types of carbon materials and could lead to different properties. This is highly relevant for potential use in industry as quality control tool. For the scientific community, this chapter showed the importance of the fluorescent background in the Raman spectrum for proper characterization of carbon materials.

Organic carbon and elemental carbon are general classifications for particulate matter present in diesel emissions. These two species present potential harm to the health of workers, monitoring their presence is thus paramount for a safe workplace environment. The literature offers different approaches for estimating elemental carbon and the comparison with different techniques. Most of them ignore the organic fraction and don’t compare with the North American standard technique (NIOSH method). The fluorescent background and the correlation with hydrogenation of carbon materials that were observed in

chapter 2 inspired the use of the background as source of information for estimating organic carbon content. The results from Raman and partial least squares in chapter 3 suggest the background as an important source of information for estimating organic carbon and total carbon. The predictions of total carbon agreed with the values obtained from the standard technique in the range between 60 and 600 µg.

Although Raman spectroscopy is largely used for characterization of carbon materials, it is not rare to find misinterpretation of the spectra. Chapter 4 presents an automated prediction of carbon materials based on their Raman spectra and using principal component analysis followed by linear discriminant analysis. The Raman spectra used in for training the model were obtained from highly oriented pyrolytic graphite, glassy carbon, diamond-like carbon, hydrogenated graphite-like carbon and hydrogenated polymer-like carbon. The testing dataset was based on Raman spectra from the literature and from samples synthesized in the lab. The results showed accuracy of 97 % and some assignments found in the literature could be corrected by the model proposed in this chapter.

Chapter 5 offers a summary of the research projects presented in this dissertation and discuss the possible future projects for developing the use of Raman spectroscopy and machine learning for carbon materials.

Table of Contents

List of Tables

List of Figures

Acknowledgements

I would like to thank:

My supervisor, Dr. Alexandre G. Brolo, for his supervision, help support and patience throughout the course of this thesis.

Dr. Dennis Hore, Dr. Jeremy Wulff, and Dr. Geoffrey Steeves for their assistance and valuable insights for the research projects.

The past Brolo group members for general help and discussions, in special Dr. Regivaldo Gomes, Dr. Antonio Marcos Brito, Dr. Karolina Valente,

All current Brolo group members and friends from the Department of Chemistry.

The glass shop, chemical store, and chemistry department office for their help.

NSERC, Mitacs, UVic, CFI for funding.

1 Introduction

1.1 Motivation and Objectives

1.1.1 Motivation

Carbon materials exhibit diverse and exciting chemical, electrical, and mechanical properties.[1] To illustrate, graphite has high conductivity, whereas diamond serves as an insulator. Amorphous carbon materials can have predominant graphite-like (sp^2) or diamond-like (sp^3) structures, resulting in varying conductive qualities and potential uses as protective coatings.[1-5] These carbon materials, including graphite, diamond, glassy carbon, and diamond-like carbon, find widespread commercial applications in daily life, playing roles in batteries, lubricants, lenses, and in the automotive industry, among others.[1-3, 5]

Raman spectroscopy is an analytical technique that irradiates laser light in the sample and collects the scattered light to observe the vibrational modes of molecules and materials.[6] It can identify molecules (molecular fingerprint), molecular conformation, or crystalline structure.[7-10] The Raman technique has several advantages. Since it is rapid, non-destructive, and portable, Raman spectroscopy has been used for materials characterization and quantification in several areas in chemistry, biology, geology, medicine, and material science.[7, 11-14]

In 1970 Tuinstra and Koenig published the first report of the Raman spectrum of graphite.[15] Since then, Raman spectra of carbon materials have been largely studied, and feature variations correlated with different forms of hybridizations, crystallinity, and allotropes.[16] Carbon materials are also relevant for health and environmental applications.[12] The Raman spectrum of crystalline graphite is characterized by a strong band at 1580 cm^{-1}, called G-band. This band is assigned to the C-C stretching of sp^2 carbon.[16, 17] Increasing disorder in graphite allows the appearance of a second band at 1360 cm^{-1}, called D-band,

assigned to the ring breathing modes of sp^2 carbon.[16, 17] These are the two bands mainly used for the characterization of graphitic materials. Although in graphene, the 2D band at 2700 cm^{-1} and the C-band at 42 cm^{-1} also provide important structural information[18]

In 2000, Ferrari and Robertson published an article proposing a three-stage model for relating the Raman features to the sp^2 structure.[16] Their analysis focuses on G- and D-bands and uses visible lasers for excitation. The model describes the amorphization trajectory from graphite to tetrahedral amorphous carbon and shows the variation of G-band position and the intensity ratio of D- and G-bands. For instance, some disorder is induced in the first stage as the graphite transitions to nanocrystalline graphite. In the Raman spectra, the G-band shifts from 1580 cm^{-1} to 1600 $cm^{-1,}$ and the intensity ratio of D- and G-bands goes from 0 to 2.0. In the second stage, the nanocrystalline graphite changes to amorphous carbon. Then, the G-band position shifts from 1600 cm^{-1} to 1510 cm^{-1}, while the D- and G-band intensity ratio changes from 2.0 to 0.25. The third stage observed the transformation of the dominant hybridization from sp^2 to sp^3, comparing amorphous carbon and tetrahedral amorphous carbon. The G band shifts from 1510 cm^{-1} to 1570 cm^{-1}, and the D- and G-band intensity ratio goes to zero. This paper became a reference for interpreting Raman spectra in the literature, accumulating more than ten thousand citations to date.

In 2005, Casiraghi *et al.* investigated the effects of hydrogenation in the Raman spectra of carbon materials.[19] They controlled the composition of the hydrogenated carbon films by using chemical vapour deposition. They observed a correlation between the hydrogenation and the fluorescent background in the Raman spectra. An equation was proposed to estimate the atomic percentage of hydrogen based on the slope of the fluorescent background and the intensity (height) of the G-band.

Raman spectroscopy's versatility and non-destructive nature make it appealing to material scientists studying electrochemically synthesized carbon films.[20-25] While previous reports on electrochemically prepared carbon thin films often relied on a single spectrum or averaged measurements

from different areas; an approach that suits only uniform materials. The electrochemical deposition could yield thin films with greater variability, and the use of a single or averaged spectrum might not accurately represent the overall composition. Therefore, Raman maps should make the characterization more statistically and spatially relevant.

Toh *et al.* have synthesized and investigated the structure of monolayer amorphous carbon films.[26] They report a synthetic method using chemical vapour deposition and characterization by Raman and X-ray spectroscopy and transmission electron microscopy. The Raman spectra showed broad G and D-band at 1588 and 1340 cm^{-1}, respectively. The intensity ratio of D- and G-bands was 0.82. Their electron microscope images showed regions with crystallites surrounded by continuous random networks. The bond length and bond angle had a broad distribution, varying from 0.9 to 1.8 Å and 90 to 150 °, respectively. For comparison, graphene has a bond length of 1.4 Å and a bond angle of 120 °. This work is very inspiring for the characterization of other films by Raman spectroscopy.

Amorphous carbon materials are also present in the environment, and a common source is the incomplete combustion of diesel.[27] Soot and diesel particles matter are examples of terms used in the literature to describe samples containing amorphous carbon with environmental importance.[28-31] There are several examples of papers reporting Raman spectroscopy for characterizing these types of samples.[27, 32-36] The carbon present in the air can be classified as elemental or organic carbon.[37] The difference between the species is mostly operational. There's a general understanding that elemental carbon is a cluster of carbon nanoparticles composed only of carbon atoms and predominance of sp^2 hybridization.[36, 38] The organic carbon would have the presence of other atoms, such as oxygen and hydrogen, carbon linear chains and the relevant presence of other carbon hybridization.[36]

In a recent publication, Zheng *et al.* investigated the correlation between the Raman intensity with the amount of elemental carbon in diesel particulate matter.[29] They used two different types of NIST (National Institute of Standards and Technology) standard reference materials for creating samples with

different amounts of elemental carbon. The Raman spectra of these samples were obtained using a 785 nm laser, and the elemental carbon was determined by NIOSH method 5040.[39] Two calibration curves were built, one based on the peak height at 1322 cm^{-1} (D-band), and another based on the peak between 800-1800 cm^{-1}. In both cases, the increasing amount of elemental carbon did not correlate linearly with the Raman intensities. They propose a logarithmic function to correlate Raman intensities and elemental carbon. Some testing samples from workplaces were obtained and used to test the proposed calibration. A comparison between NIOSH method 5040 and the Raman method showed errors varying from 4 to 39 %. Other scientific papers exploring the possibility of using Raman for elemental carbon quantification can be found in the literature.[27, 31, 40, 41] Machine learning methods could be tested to improve the accuracy and help with the non-linear response. The use of Raman maps instead of the measurement of just a few points of the sample could lead to a better description of the sample and better performance of the analysis. Another possibility for improvement includes a higher diversity of the samples used to create the calibration curves. Standard samples may not be able to account for variations that occur in samples from the workplace environment. Baseline correction and other preprocessing steps may influence the accuracy of the method. Last, organic carbon has been neglected by several works, but it is another relevant species that need to be quantified.[31, 36, 42, 43]

In addition to advances in carbon materials and Raman techniques, several research groups have reported advances in their fields by using data science tools.[44-49] Such tools include machine learning methods.[13] A recent review by Morelos-Gomez *et al*. describes several works applying data science to carbon materials.[13] Data science tools can be used to predict a property based on the synthesis, but also to predict the best synthetic conditions to optimize a specific property. Some examples in the literature used the synthetic conditions to predict best growth rate of carbon nanotubes.[50, 51] In these cases, the Raman spectra from carbon nanotubes were the metric for growth rate optimization. In other studies, the Raman spectra were the data input used to predict features of the materials.[52-54] They could predict the

number of layers in graphene, and the capacitance of several carbon materials. Other studies have used the carbon nanotubes Raman spectra to detect the type of virus and estimate the mitochondrial consumption of oxygen.[55, 56]

Data science tools have also been applied to characterize amorphous carbon materials.[57, 58] Beckham *et al.* reported the use of machine learning to predict the optimal synthetic conditions of flash graphene.[57] In their study, they transformed amorphous carbon into turbostratic graphene, performed Raman maps for statistical analysis of each reaction and used features of the Raman spectra to classify the carbon materials present in each synthesis. Campos et al. have used Raman spectra and principal component analysis to distinguish sp^2 carbon materials with different levels of amorphization. The principal component analysis allowed them to observed an amorphization trend, which they compared with the amorphization trajectory proposed by Ferrari *et al*..[16] These works contribute for the application of Raman spectroscopy analysis in large scale, but there's still many aspects to be investigated.

1.1.2 General Objective

The general objective of this thesis is to investigate the application of Raman spectroscopy for classifying and quantifying carbon materials. More specifically, the goal is to combine Raman methods with chemometrics to classify different types of carbons even when they are present in the same material (composition heterogeneity).

1.2 Raman scattering

Consider an incident visible light with a specific frequency (ω_1) on a transparent material (gas, liquid or solid). Most of the photons will be transmitted without any change, but some of them will

be scattered. A further investigation of the scattered light will reveal that most of them have the same frequency of incident light (ω_1). A smaller fraction of the scattered light, however, will undergo a frequency change ($\omega_1 \pm \omega_M$). For molecules, ω_M can be related to vibrational transitions.[6]

The scattered light with the same frequency of the incident light is called Rayleigh scattering, while the scattered light with changed frequency is known as Raman scattering. A further classification of the type of scattered light can be associated with the value of its frequency. If the scattered photons have a smaller frequency than the incident light ($\omega_1 - \omega_M$), they are called Stokes scattering (Figure 1-1). While a larger value ($\omega_1 + \omega_M$) of frequency is attributed to anti-Stokes scattering. In summary, the new frequencies observed in the Raman spectrum are Raman bands and they are classified as Stokes or anti-Stokes.

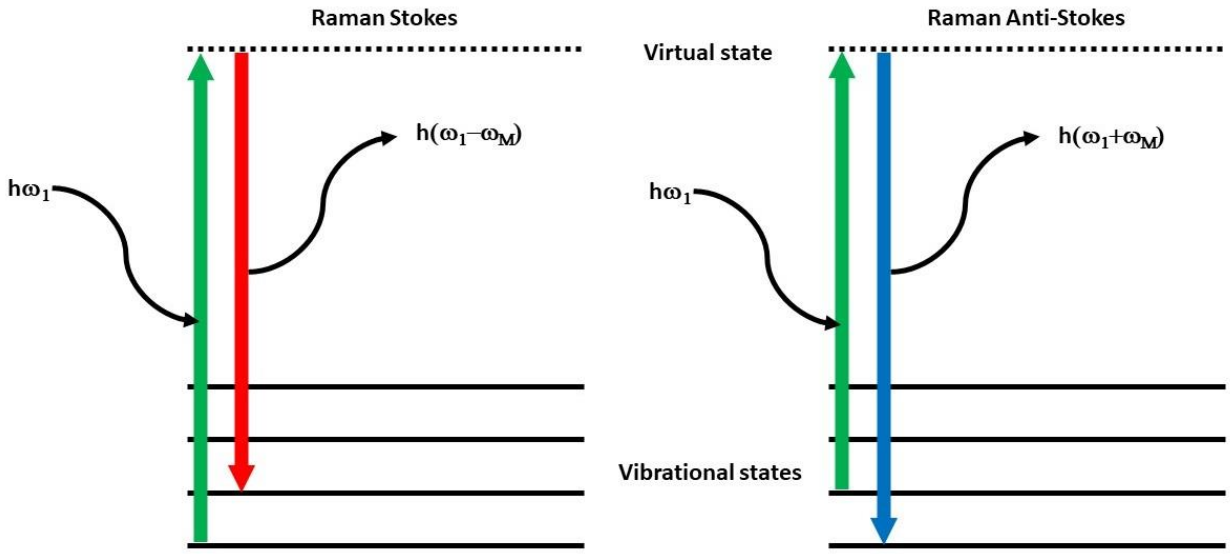

Figure 1-1: Diagram showing Raman Stokes and anti-Stokes process.

Considering that the incident radiation induces an oscillating electric dipole in a molecule, the intensity of light scattered by this molecule (I) will be given by the equation[6]:

$$I = k'_{\omega} l \omega_s^4 p_0^2 sin^2\,\theta \quad (1.1)$$

Where I is the scattered intensity time-averaged power per unit solid angle, ω_s is the frequency of the scattered light, l is the power of the incident light, p_0^2 is the amplitude of the induced electric dipole, θ is the angle between the axis of the dipole and the direction where intensity is being measured.[6] Last, k'_{ω} can be defined as[6]:

$$k'_{\omega} = \frac{1}{32\pi^2 \varepsilon_0 c_0^3} \quad (1.2)$$

While ε_0 is the electric permittivity of vacuum and c_0 being the speed of light in vacuum.

The total induced electric dipole moment vector of a molecule may be written as a series of moment vectors:

$$p = p^{(1)} + p^{(2)} + p^{(3)} + \cdots \quad (1.3)$$

In the expression above, $p^{(1)} >> p^{(2)} >> p^{(3)}$. The first term of this series can be described in Equation 1.4. [6]

$$p^{(1)} = \alpha \cdot E \quad (1.4)$$

Where α is called polarizability. Considering the electric field fluctuating with time can be described by equation 1.5. [6]

$$E = E_0 cos 2\pi \omega_1 t \quad (1.5)$$

Then, the electric dipole moment can be rearranged in Equation 1.6 [6]

$$p^{(1)} = \alpha \cdot E_0 cos2\pi\omega_1 t \quad 1.6$$

For a molecule vibrating with frequency ω_m , the nuclear displacement q can be described in Equation 1.6.[6]

$$q = q_0 cos2\pi\omega_m t \quad 1.7$$

Where the vibrational amplitude is q_0. Considering small amplitudes of vibration, α can be described as linearly dependant on q. Then, we obtain equation 1.8. [6]

$$\alpha = \alpha_0 + \left(\frac{\partial\alpha}{\partial q}\right)_0 q \quad +\ldots \quad 1.8$$

In equation 1.8, α_0 is the polarizability at the equilibrium position, and $\left(\frac{\partial\alpha}{\partial q}\right)_0$ is the rate of variation of α when q is fluctuating. The subscript 0 indicates this rate should be calculated at the equilibrium position. Replacing α (1.8) in equation 1.6, and then q (1.7), it is obtained the expression shown in 1.9. [6]

$$p^{(1)} = \alpha \cdot E_0 cos2\pi\omega_1 t \quad 1.9$$

$$= \alpha_0 \cdot E_0 cos2\pi\omega_1 t + \left(\frac{\partial\alpha}{\partial q}\right)_0 q \cdot E_0 cos2\pi\omega_1 t$$

$$= \alpha_0 \cdot E_0 cos2\pi\omega_1 t + \left(\frac{\partial\alpha}{\partial q}\right)_0 q_0 cos2\pi\omega_m t \; E_0 cos2\pi\omega_1 t$$

The obtained product of cos (1.10) in the final form of equation 1.9 can be used to obtain a new expression (1.11) with valuable physical insights.[6]

$$2\cos x \cos y = \cos(x - y) + \cos(x + y) \quad 1.10$$

$$p^{(1)} = \alpha_0 \cdot E_0 cos2\pi\omega_1 t + \frac{1}{2}\left(\frac{\partial\alpha}{\partial q}\right)_0 q_0 E_0(\cos[2\pi(\omega_1 - \omega_m)t] + \cos[2\pi(\omega_1 + \omega_m)t]) \quad 1.11$$

The first term of 1.11 can be interpreted as an oscillating dipole that radiates light with a frequency equal to the incident light (ω_1). This phenomenon is called Rayleigh scattering. The second and term have light radiated with a frequency related to the incident light and the vibration of the atoms. This phenomenon is called Raman scattering. The sign indicates the type of Raman scattering, stokes ($\omega_1 - \omega_m$) or anti-stokes ($\omega_1 + \omega_m$). Both processes are illustrated in Figure 1-1.

It's also important to emphasize the term ($\left(\frac{\partial\alpha}{\partial q}\right)_0$) multiplying both types of Raman scattering in 1.11. The derivative implies one Raman scattering selection rule. The vibrational modes observed in a Raman spectrum will be the ones with the variation of the polarizability at the equilibrium. Conservation of energy and momentum are the other two selection rules that must be obeyed and will be discussed later in this thesis.

The theory described in the previous paragraphs is the classical theory of Raman scattering. The quantum mechanical theory of Raman scattering considers the discrete states of the molecule, but treats light as electromagnetic waves.[6] The use of time-dependent perturbation theory will lead to a similar expression obtained for the classic theory. In the end, the expression for induced transition electric dipole moment is shown in 1.12.[6]

$$\left(p_\rho^{(1)}\right)_{fi} = \left(\alpha_{\rho\sigma}\right)_{fi} E_{\sigma 0}(\omega_1)\cos\omega_S t \quad 1.12$$

Where the indices i and f refer to initial and final states, respectively. The p refers to the electric dipole moment. The superscript on p indicates the first order induced transition electric dipole moment, while the subscripts ρ and σ relate to the coordinates of the induced electric field. The Greek letter α represent the transition polarizability. E_0 is the electric field amplitude at the equilibrium position, ω_1

and ω_S represent the frequency of the incident and scattered light, respectively, and t is time.[6] The general transition polarizability is displayed in 1.13. [6]

$$\left(\alpha_{\rho\sigma}\right)_{fi} = \frac{1}{\hbar}\sum\left\{\frac{\langle f|\hat{p}_\rho|r\rangle\langle r|\hat{p}_\sigma|i\rangle}{\omega_{ri} - \omega_1 - i\Gamma_r} + \frac{\langle f|\hat{p}_\sigma|r\rangle\langle r|\hat{p}_\rho|i\rangle}{\omega_{ri} + \omega_1 + i\Gamma_r}\right\} \quad 1.13$$

The $\hbar$ is the reduced Planck constant. The letters i, r and f in the brackets refer to the wave function of the initial, intermediate and final states, respectively. $\hat{p}$ is the electric dipole moment operator. The Γ_r relates to the full width of the intermediate state. Finally, the subscript ri on ω indicates the frequency difference between states r and i. It's worth noticing the denominator of the first term can be much smaller if the incident light has the same frequency of the transition between the states r and i. If the denominator is smaller, then the polarizability will be higher, and that is when the resonant effect is observed. The description and discussion are made in much deeper detail in the literature.[6]

1.3 Raman spectra of carbon materials

The Raman spectrum of crystalline graphite under visible excitation is characterized by a strong band at 1580 cm^{-1}, called G-band. From the molecular point of view, this band can be assigned to sp^2 C-C stretching mode.[16] A second band will appear in case of disorder in the material, i.e. atom vacancies, dangling bonds, and sp^3 hybridization.[16] This second peak, called D-band, rises at 1360 cm^{-1} and could be assigned to the ring breathing mode of sp^2 carbon atoms. One structural difference between the two modes is that the G-band doesn't require sixfold rings, so it occurs for any carbon material that contains sp^2 sites.[17]

The loss of crystallinity (disorder) leads to the relaxation of the selection rules, allowing the rise of the D-band. A complete explanation can be found in the literature and requires some knowledge of

solid-state physics (e.g. band structure, density of states).[16, 59] In a few words, the momentum and energy should be conserved. The conservation of energy is illustrated in Figure 1-1. Considering the Raman Stokes scenario, the energy from the incident light ($h\omega_1$) is equal to the sum of energy from scattered light ($h(\omega_1 - \omega_m)$) and the energy associated with the vibrational transition ($h\omega_m$). The momentum conservation works in a very similar way, but considering the momentum of the light (photon) and the vibrational modes. In this case, the momentum of the incident light is very small ($q \sim 0$) compared to the momentum of the vibrational mode. This implies only the vibrational modes close to the center (Γ) of the Brillouin zone are active (G-band), because they are close to $q = 0$.[17] However, the disorder in graphite allows some relaxation in the momentum conservation rule and a mode that is at the edge (K) of the Brillouin zone rises (quasi-selection rule (k=q/2) allowing the D-band).[60] Also, the visible laser excitation is in resonance with π- π* transition, contributing to the rise of the D-band.[17] A much more detailed description can be found in the literature.[16, 59]

Figure 1-2 shows the schematic Raman spectra of different types of carbon materials obtained using visible light. The first spectrum (Figure 1-2A) was obtained from highly oriented pyrolytic graphite. This type of material is very crystalline and only the sharp G-band is observed. The Raman spectrum of glassy carbon is presented in Figure 1-2B. Glassy carbon is some type of amorphous material with a predominance of sp^2 carbon and a 3D structure similar to fullerene. The broadening of the Raman bands can be related to the amorphous character of the material.[16] The D-band is present due to disorder in the material.[59] Last, the Raman spectrum of hydrogenated polymer-like carbon is presented in Figure 1-2C. A fluorescent background can be observed in 1700-1800 cm^{-1}, when compared with region around 1100 cm^{-1}. This fluorescence comes from hydrogenated materials and can prevail over the Raman bands due to the higher cross-section of fluorescence compared to Raman.[19]

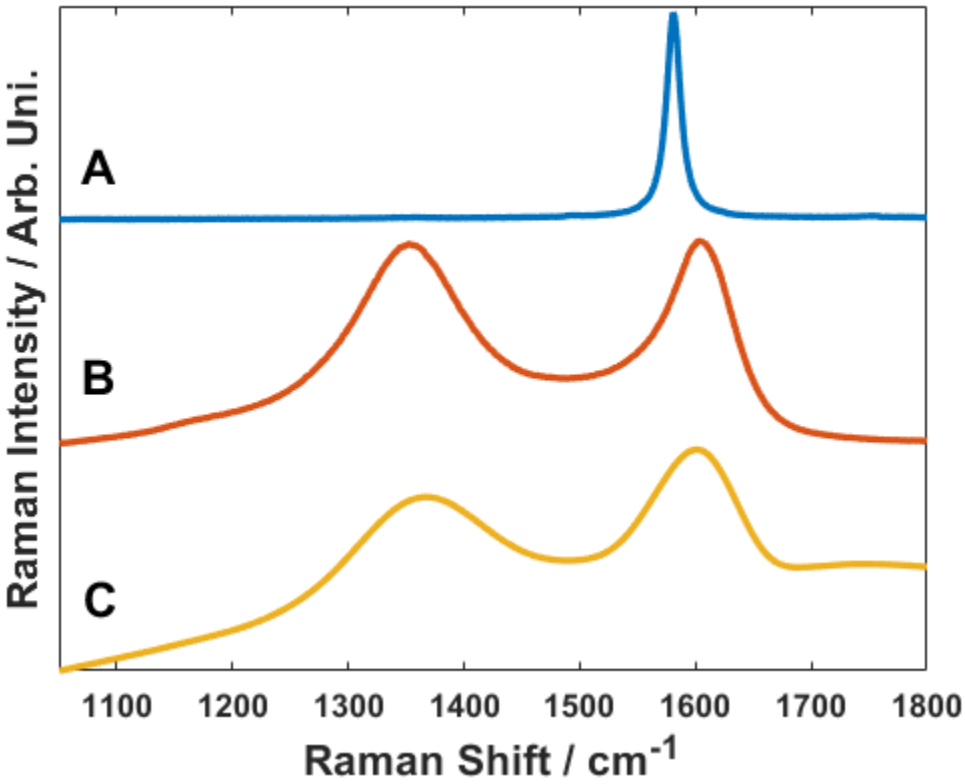

Figure 1-2: A) Raman spectra of highly oriented pyrolytic graphite, B) glassy carbon and C) hydrogenated polymer-like carbon obtained using 532 nm laser, 5s exposure time, 3 accumulations, 50x lens (NA = 0.75).

1.4 Machine learning

1.4.1 Principal Component Analysis

Principal component analysis (PCA) is a method of dimensionality reduction, meaning it reduces the number of variables while keeping most of the data variance.[61, 62] For instance, a Raman spectrum can be thought of as a sequence of data points. Each point corresponds to the Raman intensity at a specific wavenumber. Each wavenumber can be considered as a variable, and the Raman intensities as the values of this variable. Once the PCA is applied, a smaller number of variables will account for a significant amount of information from the data. These new variables (principal components, PCs) are linear combinations of the original Raman intensities.

At the beginning of the PCA, the data is subtracted from the mean, and the covariance matrix is obtained by calculating the covariance of each pair of variables. The number of PCs will be equal to the number of variables and they will be ordered by their eigenvalue, in decreasing order. The vectors obtained can be called loading vector and show how much each original variable contributes to the score of the corresponding PC. The scores and loadings can be used for the investigation of similarities and differences within the data and as input for classification methods. A much more detailed explanation of PCA can be found in the literature.[61-63]

1.4.2 Linear Discriminant Analysis

Linear Discriminant Analysis (LDA) is a technique used to separate two or more groups. It is a similar approach to PCA in a way that it reduces the number of variables, but the emphasis is on finding the direction where the two groups present a larger separation among themselves. Once some of the data has been used to create a model, then it can be used to classify unknown samples. Let's consider we have two groups that need to be separated. LDA will find a new axis where it: 1) Maximizes the distance between the means of two different groups; 2) Minimizes the variation within each group.

Figure 1-3 shows different possibilities. In A, it is the data plotted according to variables X_1 and X_2. In B and C, the projected data shows a mix between three different classes. In D, the classes are clearly separated.

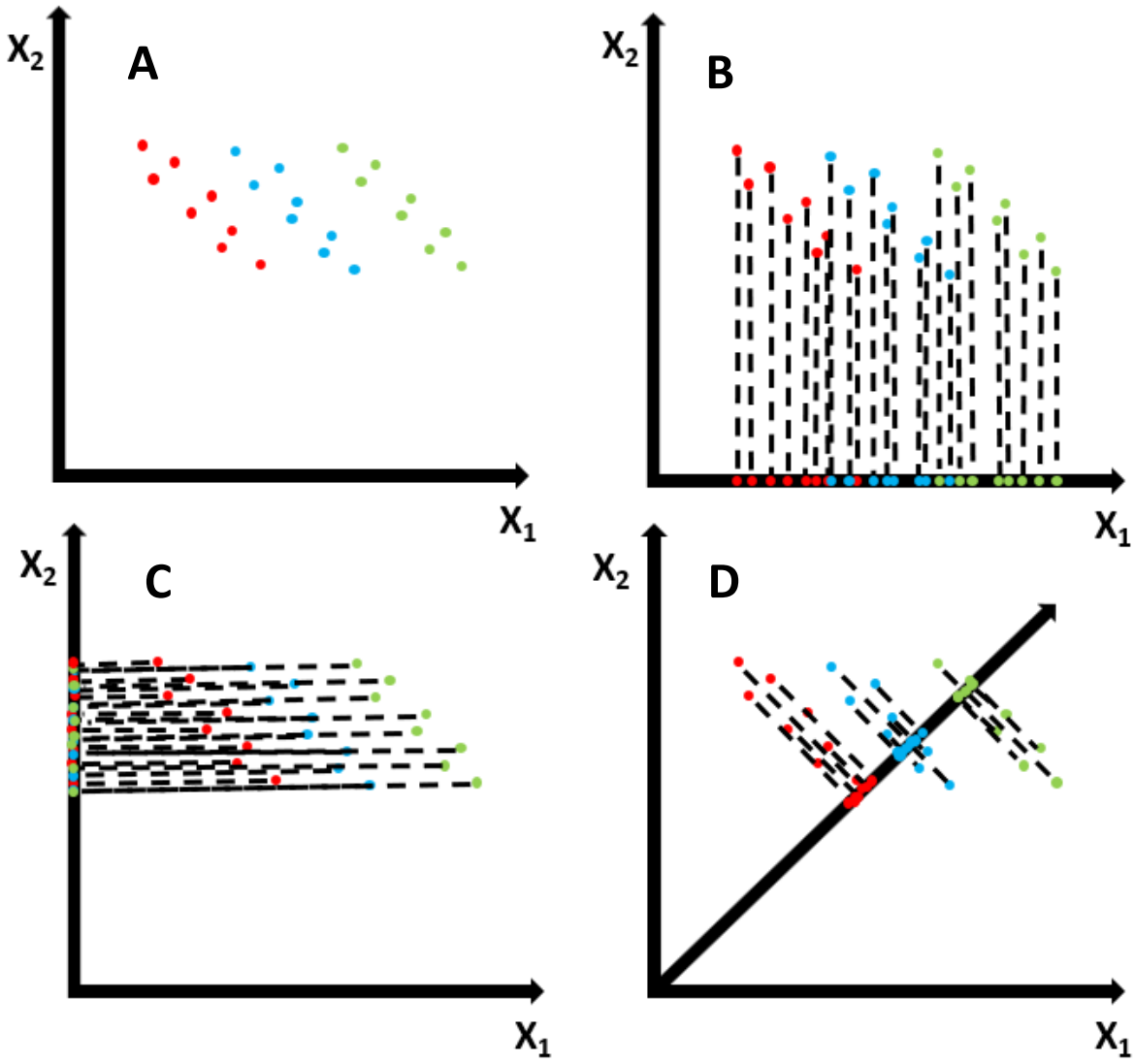

Figure 1-3: Two-dimensional data being projected in a new axis.

Once the model is created, the LDA has to predict the classes of a testing data set. First, it will estimate the probability of a testing example belonging to each class. Then, the predicted class will be the one with the highest probability. Further details about LDA can be found in the literature.[61, 62]

1.4.3 Partial Least Squares Regression

Partial Least Square regression (PLS) is a method for relating a matrix X to a matrix Y. PLS is widely employed in chemometrics. The origin of PLS can be assigned to Herman Wold, who developed a method for the treatment of chains of matrices and applied it to econometrics in 1975.[64] Around 1980, a simpler model involving two matrixes (X and Y) was proposed by Svante Wold and Harald Martens.[65] In a few words, PLS searches for maximum covariance between the score matrices of X and Y. These score matrices will be explained later in the text. Covariance combines high variance of X and high correlation with Y. Also, PLS is a linear method, meaning the final latent variable is a linear combination of the original variables. It can handle strongly collinear and noisy data, and numerous X-variables at the same time it models several response variables (Y). The following steps are involved in PLS:

1 – First PLS component is derived as the latent variable which has maximum covariance between the scores and modelled property y.

2 – The variance of this component is subtracted from the x-data. The residual matrix X_{RES} will have the same number of variables as the original, but different dimensionality, reduced by one.

3 – From the X_{RES}, the next PLS component is calculated, with maximum covariance between the scores and y.

4 – This is repeated until no improvement in modelling is achieved.

The literature discusses in much more detail all the concepts and calculations behind PLS, and presents several methods available.[46, 61, 62, 66-73]

1.5 Organization of the thesis

This thesis is divided into 5 chapters. In the first chapter (Introduction), the theoretical background of Raman spectroscopy, together with the main characteristics of the Raman response of

carbon materials will be presented. The main aspects of the data analytic methods used in the work, principal component analysis, partial least square regression and linear discriminant analysis, will also be presented.

Chapter 2 deals with the use of Raman maps to characterize the compositional heterogeneity in carbon materials synthesized via wet chemistry. The traditional methodology for characterizing carbon materials using Raman spectroscopy involves obtaining one Raman spectrum of a specific point or an average. There is no information regarding variation or compositional heterogeneity in the conventional approach. This work focused on carbon materials obtained electrochemically and showed that the Raman features vary spatially with the synthetic procedures. The most striking variations were observed in the hydrogen content of the samples.

Chapter 3 comprises the use of Raman spectroscopy along with partial least squares (PLS) as a rapid method to characterize exposure to carbon materials in workplace conditions. Elemental (EC), organic (OC) and total carbon (TC) were estimated from samples obtained in mining facilities. Our results have shown the possibility of predicting the percentage of elemental carbon in the range between 50 and 80 % by simple Raman analysis. The developed models could also predict the mass of EC, OC and TC and they revealed the importance of fluorescence background for estimating OC and TC.

Chapter 4 shows the potential of generating a method to automate the classification of carbon materials based on their Raman spectra. 5 different bulk materials have been chosen as training sets for the model: highly oriented pyrolytic graphite (HOPG); glassy carbon (GC); diamond-like carbon (DLC); hydrogenated graphite-like carbon (GLCH); and hydrogenated polymer-like carbon (PLCH). These materials present variations of crystallinity, functionalization, and predominance of sp^2 and sp^3 hybridization. The Raman data used to train the model was collected from both commercial and synthesized samples. The model was tested using Raman spectra obtained from the literature and

materials synthesized in our group. The chemometric methodology included principal component analysis followed by linear discriminant analysis (PCA-LDA). Our results have shown a high agreement between the assignment made by specialists and the predictions of the model. Some observed misclassifications could be avoided by either adding more data to training or by applying some other non-linear model, such as support vector machines, and neural networks. Although, the use of PCA-LDA allowed interpretation and insights regarding how the classification occurred.

Chapter 5 is a summary of the insights obtained in this thesis. The most significant results from chapters 2, 3 and 4 are reviewed and discussed. Finally, some possible perspectives for this work are proposed.

1.6 References

(1) Robertson, J. Comparison of diamond-like carbon to diamond for applications. *Physica Status Solidi a-Applications and Materials Science* **2008**, *205* (9), 2233-2244. DOI: 10.1002/pssa.200879720.

(2) Tyagi, A.; Walia, R. S.; Murtaza, Q.; Pandey, S. M.; Tyagi, P. K.; Bajaj, B. A critical review of diamond like carbon coating for wear resistance applications. *International Journal of Refractory Metals & Hard Materials* **2019**, *78*, 107-122. DOI: 10.1016/j.ijrmhm.2018.09.006.

(3) Caschera, D.; Toro, R. G.; Cortese, B.; Federici, F.; Lombardo, D.; Calandra, P. Diamond-like carbon: a versatile material for developing innovative smart textiles applications. a short review. *Atti Accademia Peloritana Dei Pericolanti-Classe Di Scienze Fisiche Matematiche E Naturali* **2019**, *97, A27*. DOI: 10.1478/aapp.97s2a27.

(4) Jang, S.; Kim, S. H. Distinct effects of endogenous hydrogen content and exogenous hydrogen supply on superlubricity of diamond-like carbon. *Carbon* **2023**, *202*, 61-69. DOI: 10.1016/j.carbon.2022.11.011.

(5) Hainsworth, S. V.; Uhure, N. J. Diamond like carbon coatings for tribology: production techniques, characterisation methods and applications. *International Materials Reviews* **2007**, *52* (3), 153-174. DOI: 10.1179/174328007x160272.

(6) Long, D. A. *The Raman Effect*; John Wiley & Sons Ltd, 2002.

(7) Balan, V.; Mihai, C. T.; Cojocaru, F. D.; Uritu, C. M.; Dodi, G.; Botezat, D.; Gardikiotis, I. Vibrational Spectroscopy Fingerprinting in Medicine: from Molecular to Clinical Practice. *Materials* **2019**, *12* (18), 2884. DOI: 10.3390/ma12182884.

(8) Smith, R.; Wright, K. L.; Ashton, L. Raman spectroscopy: an evolving technique for live cell studies. *Analyst* **2016**, *141* (12), 3590-3600. DOI: 10.1039/c6an00152a.

(9) Peimyoo, N.; Li, J. W.; Shang, J. Z.; Shen, X. N.; Qiu, C. Y.; Xie, L. H.; Huang, W.; Yu, T. Photocontrolled Molecular Structural Transition and Doping in Graphene. *ACS Nano* **2012**, *6* (10), 8878-8886. DOI: 10.1021/nn302876w.

(10) Ma, H. L.; Yang, J. Y.; Dai, Y.; Zhang, Y. B.; Lu, B.; Ma, G. H. Raman study of phase transformation of TiO2 rutile single crystal irradiated by infrared femtosecond laser. *Applied Surface Science* **2007**, *253* (18), 7497-7500. DOI: 10.1016/j.apsusc.2007.03.047.

(11) Solis-Fernandez, P.; Ago, H. Machine Learning Determination of the Twist Angle of Bilayer Graphene by Raman Spectroscopy: Implications for van der Waals Heterostructures. *ACS Applied Nano Materials* **2022**, *5* (1), 1356-1366. DOI: 10.1021/acsanm.1c03928.

(12) Moorchilot, V. S.; Aravind, U. K.; Menacherry, S. P. M.; Aravindakumar, C. T. Single-Particle Analysis of Atmospheric Aerosols: Applications of Raman Spectroscopy. *Atmosphere* **2022**, *13* (11), 1779. DOI: 10.3390/atmos13111779.

(13) Morelos-Gomez, A.; Terrones, M.; Endo, M. Data Science Applied to Carbon Materials: Synthesis, Characterization, and Applications. *Advanced Theory and Simulations* **2022**, *5* (2), 2100205. DOI: 10.1002/adts.202100205.

(14) Clegg, S. M.; Wiens, R.; Misra, A. K.; Sharma, S. K.; Lambert, J.; Bender, S.; Newell, R.; Nowak-Lovato, K.; Smrekar, S.; Dyar, M. D.; et al. Planetary Geochemical Investigations Using Raman and Laser-Induced Breakdown Spectroscopy. *Applied Spectroscopy* **2014**, *68* (9), 925-936. DOI: 10.1366/13-07386.

(15) Tuinstra, F.; Koenig, J. L. RAMAN SPECTRUM OF GRAPHITE. *Journal of Chemical Physics* **1970**, *53* (3), 1126-&, Article. DOI: 10.1063/1.1674108.

(16) Ferrari, A. C.; Robertson, J. Interpretation of Raman spectra of disordered and amorphous carbon. *Physical Review B* **2000**, *61* (20), 14095-14107. DOI: 10.1103/PhysRevB.61.14095.

(17) Ferrari, A. C.; Robertson, J. Resonant Raman spectroscopy of disordered, amorphous, and diamondlike carbon. *Physical Review B* **2001**, *64* (7), 075414. DOI: 10.1103/PhysRevB.64.075414.

(18) Ferrari, A. C.; Basko, D. M. Raman spectroscopy as a versatile tool for studying the properties of graphene. Nature Publishing Group: 2013; Vol. 8, pp 235-246.

(19) Casiraghi, C.; Piazza, F.; Ferrari, A. C.; Grambole, D.; Robertson, J. Bonding in hydrogenated diamond-like carbon by Raman spectroscopy. *Diamond and Related Materials* **2005**, *14* (3-7), 1098-1102, Article; Proceedings Paper. DOI: 10.1016/j.diamond.2004.10.030.

(20) Wu, M. X.; Tang, R. J.; Chen, Y. L.; Wang, S. Y.; Wang, W. C.; Chen, X. H.; Mitsuzaki, N.; Chen, Z. D. Electrochemical reduction of CO2 to carbon films on stainless steel around room temperature. *Electrochemistry Communications* **2020**, *110,* 106606. DOI: 10.1016/j.elecom.2019.106606.

(21) Falcade, T.; Turq, V.; Bonino, J. P.; Malfatti, C. D. Tribological properties of amorphous carbon films obtained by electrodeposition from DMF using 2HEAL protic ionic liquid as dopant. *Diamond and Related Materials* **2017**, *71*, 30-37. DOI: 10.1016/j.diamond.2016.11.014.

(22) Gupta, S.; Rahini, R. L. Electrochemical synthesis and characterization of amorphous hydrogenated carbon (a-C:H) using acetonitrile as electrolyte. *Electrochimica Acta* **2017**, *258*, 1-8, Article. DOI: 10.1016/j.electacta.2017.10.079.

(23) Jiang, J. L.; Zhang, X.; Du, J. F.; Wang, Q.; Dai, J. F.; Wei, Z. Q. Electrochemical Deposition and Field Emission Properties of Graphene/Diamond-Like Carbon Nanocomposite Films. *Acta Physico-Chimica Sinica* **2016**, *32* (7), 1839-1843, Article. DOI: 10.3866/pku.whxb201604063.

(24) Sakita, A. M. P.; Valente, M. A. G.; Della Noce, R.; Fugivara, C. S.; Magnani, M.; Benedetti, A. V. Low-voltage carbon films deposition by electro-exfoliation of graphite into graphene oxide. *RSC Advances* **2016**, *6* (87), 84194-84199, Article. DOI: 10.1039/c6ra16502h.

(25) Zhang, Q.; Wang, Y. R.; Wang, W. C.; Mitsuzak, N.; Chen, Z. D. Low voltage and ambient temperature electrodeposition of uniform carbon films. *Electrochemistry Communications* **2016**, *63*, 22-25, Article. DOI: 10.1016/j.elecom.2015.11.012.

(26) Toh, C. T.; Zhang, H. J.; Lin, J. H.; Mayorov, A. S.; Wang, Y. P.; Orofeo, C. M.; Ferry, D. B.; Andersen, H.; Kakenov, N.; Guo, Z. L.; et al. Synthesis and properties of free-standing monolayer amorphous carbon. *Nature* **2020**, *577* (7789), 199-+. DOI: 10.1038/s41586-019-1871-2.

(27) Ivleva, N. P.; McKeon, U.; Niessner, R.; Poschl, U. Raman microspectroscopic analysis of size-resolved atmospheric aerosol particle samples collected with an ELPI: Soot, humic-like substances, and inorganic compounds. *Aerosol Science and Technology* **2007**, *41* (7), 655-671. DOI: 10.1080/02786820701376391.

(28) Liang, X. Y.; Zhao, B. W.; Wang, K.; Lv, X.; Wang, Y. J.; Liu, J.; Wang, Y. S. Impact of early injection on physicochemical characteristics of diesel soot particles. *Fuel* **2021**, *292,* 120262. DOI: 10.1016/j.fuel.2021.120262.

(29) Zheng, L. A.; Birch, M. E.; Johnson, B.; Breitenstein, M.; Snawder, J.; Kulkarni, P. Correlation between Graphitic Carbon and Elemental Carbon in Diesel Particulate Matter in Workplace Atmospheres. *Analytical Chemistry* 2023, 95, 6, 3283-3290. DOI: 10.1021/acs.analchem.2c04261.

(30) Khan, M. U.; Homan, K. O.; Saki, S. A.; Emad, M. Z.; Raza, M. A. Real-time diesel particulate matter monitoring in underground mines: evolution and applications. *International Journal of Mining Reclamation and Environment* **2021**, *35* (4), 291-305. DOI: 10.1080/17480930.2020.1818937.

(31) Parks, D. A.; Griffiths, P. R.; Weakley, A. T.; Miller, A. L. Quantifying elemental and organic carbon in diesel particulate matter by mid-infrared spectrometry. *Aerosol Science and Technology* **2021**, *55* (9), 1014-1027. DOI: 10.1080/02786826.2021.1917764.

(32) Baldelli, A.; Trivanovic, U.; Sipkens, T. A.; Rogak, S. N. On determining soot maturity: A review of the role of microscopy-and spectroscopy-based techniques. *Chemosphere* **2020**, *252,* 126532. DOI: 10.1016/j.chemosphere.2020.126532.

(33) Ge, H. W.; Ye, Z. P.; He, R. Raman spectroscopy of diesel and gasoline engine-out soot using different laser power. *Journal of Environmental Sciences* **2019**, *79*, 74-80. DOI: 10.1016/j.jes.2018.11.001.

(34) Saffaripour, M.; Tay, L. L.; Thomson, K. A.; Smallwood, G. J.; Brem, B. T.; Durdina, L.; Johnson, M. Raman spectroscopy and TEM characterization of solid particulate matter emitted from soot generators and aircraft turbine engines. *Aerosol Science and Technology* **2017**, *51* (4), 518-531. DOI: 10.1080/02786826.2016.1274368.

(35) Merlen, A.; Buijnsters, J. G.; Pardanaud, C. A Guide to and Review of the Use of Multiwavelength Raman Spectroscopy for Characterizing Defective Aromatic Carbon Solids: from Graphene to Amorphous Carbons. *Coatings* **2017**, *7* (10), 55, 153; Review. DOI: 10.3390/coatings7100153.

(36) Ess, M. N.; Ferry, D.; Kireeva, E. D.; Niessner, R.; Ouf, F. X.; Ivleva, N. P. In situ Raman microspectroscopic analysis of soot samples with different organic carbon content: Structural changes during heating. *Carbon* **2016**, *105*, 572-585. DOI: 10.1016/j.carbon.2016.04.056.

(37) Birch, M. E. Chapter DL: Monitoring Diesel Exhaust in the Workplace. In *NIOSH manual of analytical methods.*, 5th ed.; Andrews, E., O'Connor, P. F. Eds.; US Department of Health And Human Services, Centers for Disease Control and Prevention, National Institute for Occupational Safety and Health, 2016.

(38) Bladt, H.; Ivleva, N. P.; Niessner, R. Internally mixed multicomponent soot: Impact of different salts on soot structure and thermo-chemical properties. *Journal of Aerosol Science* **2014**, *70*, 26-35. DOI: 10.1016/j.jaerosci.2013.11.007.

(39) and, N. I. f. O. S.; (NIOSH), H. ELEMENTAL CARBON (DIESEL PART.) 5040. In *NIOSH Manual of Analytical Methods (NMAM)*, Fourth Edition ed.; 2003.

(40) Zheng, L. N.; Kulkarni, P. Real-Time Measurement of Airborne Carbon Nanotubes in Workplace Atmospheres. *Analytical Chemistry* **2019**, *91* (20), 12713-12723, Article. DOI: 10.1021/acs.analchem.9b02178.

(41) Grafen, M.; Schweiger, G.; Esen, C.; Ostendorf, A. Time-resolved measurement of elemental carbon in urban environment: Comparison of Raman backscattering and aethalometer results. *Journal of Aerosol Science* **2018**, *117*, 34-43. DOI: 10.1016/j.jaerosci.2017.12.003.

(42) Lim, S.; Lee, S.; Ahn, T.; Park, S. Measurement of organic carbon content during the growth of soot particles in propane normal and inverse diffusion flames using a multi-wavelength light extinction method. *Carbon* **2019**, *149*, 519-529. DOI: 10.1016/j.carbon.2019.04.072.

(43) Russo, C.; Apicella, B.; Lighty, J. S.; Ciajolo, A.; Tregrossi, A. Optical properties of organic carbon and soot produced in an inverse diffusion flame. *Carbon* **2017**, *124*, 372-379. DOI: 10.1016/j.carbon.2017.08.073.

(44) Kapoor, S.; Narayanan, A. Leakage and the Reproducibility Crisis in ML-based Science. arXiv: 2022.

(45) Lussier, F.; Thibault, V.; Charron, B.; Wallace, G. Q.; Masson, J. F. Deep learning and artificial intelligence methods for Raman and surface-enhanced Raman scattering. *Trac-Trends in Analytical Chemistry* **2020**, *124*, 15, 115796; Review. DOI: 10.1016/j.trac.2019.115796.

(46) Ramirez, C. M. A.; Greenop, M.; Ashton, L.; Rehman, I. U. Applications of machine learning in spectroscopy. *Applied Spectroscopy Reviews*, **2021**, 56, 733-763, Review; Early Access. DOI: 10.1080/05704928.2020.1859525.

(47) De Bruyne, S.; Speeckaert, M. M.; Van Biesen, W.; Delanghe, J. R. Recent evolutions of machine learning applications in clinical laboratory medicine. *Critical Reviews in Clinical Laboratory Sciences*, **2021**, 58, 131-152; Review; Early Access. DOI: 10.1080/10408363.2020.1828811.

(48) Keith, J. A.; Vassilev-Galindo, V.; Cheng, B. Q.; Chmiela, S.; Gastegger, M.; Mueller, K. R.; Tkatchenko, A. Combining Machine Learning and Computational Chemistry for Predictive Insights Into Chemical Systems. *Chemical Reviews* **2021**, *121* (16), 9816-9872. DOI: 10.1021/acs.chemrev.1c00107.

(49) Smith, M. J.; Geach, J. E. Astronomia ex machina: a history, primer and outlook on neural networks in astronomy. *Royal Society Open Science* **2023**, *10* (5), 221454. DOI: 10.1098/rsos.221454.

(50) Nikolaev, P.; Hooper, D.; Webber, F.; Rao, R.; Decker, K.; Krein, M.; Poleski, J.; Barto, R.; Maruyama, B. Autonomy in materials research: a case study in carbon nanotube growth. *Npj Computational Materials* **2016**, *2, 16031*. DOI: 10.1038/npjcompumats.2016.31.

(51) Chang, J.; Nikolaev, P.; Carpena-Nunez, J.; Rao, R.; Decker, K.; Islam, A. E.; Kim, J.; Pitt, M. A.; Myung, J. I.; Maruyama, B. Efficient Closed-loop Maximization of Carbon Nanotube Growth Rate using Bayesian Optimization. *Scientific Reports* **2020**, *10* (1), 9040. DOI: 10.1038/s41598-020-64397-3.

(52) Wang, J. S.; Li, Z.; Yan, S. C.; Yu, X.; Ma, Y. Q.; Ma, L. Modifying the microstructure of algae-based active carbon and modelling supercapacitors using artificial neural networks. *RSC Advances* **2019**, *9* (26), 14797-14808. DOI: 10.1039/c9ra01255a.

(53) Silva, D. L.; Campos, J. L. E.; Fernandes, T. F. D.; Rocha, J. N.; Machado, L. R. P.; Soares, E. M.; Miquita, D. R.; Miranda, H.; Rabelo, C.; Neto, O. P. V.; et al. Raman spectroscopy analysis of number of layers in mass-produced graphene flakes. *Carbon* **2020**, *161*, 181-189. DOI: 10.1016/j.carbon.2020.01.050.

(54) Su, H. P.; Lin, S.; Deng, S. W.; Lian, C.; Shang, Y. Z.; Liu, H. L. Predicting the capacitance of carbon-based electric double layer capacitors by machine learning. *Nanoscale Advances* **2019**, *1* (6), 2162-2166. DOI: 10.1039/c9na00105k.

(55) Gonzalez-Durruthy, M.; Alberic, L. C.; Curti, C.; Naal, Z.; Atique-Sawazaki, D. T.; Vazquez-Naya, J. M.; Gonzalez-Diaz, H.; Munteanu, C. R. Experimental-Computational Study of Carbon Nanotube Effects on Mitochondrial Respiration: In Silico Nano-QSPR Machine Learning Models Based on New Raman Spectra Transform with Markov-Shannon Entropy Invariants. *Journal of Chemical Information and Modeling* **2017**, *57* (5), 1029-1044. DOI: 10.1021/acs.jcim.6b00458.

(56) Yeh, Y. T.; Gulino, K.; Zhanga, Y. H.; Sabestien, A.; Chou, T. W.; Zhou, B.; Lin, Z.; Albert, I.; Lu, H. G.; Swaminathan, V.; et al. A rapid and label-free platform for virus capture and identification from

clinical samples. *Proceedings of the National Academy of Sciences of the United States of America* **2020**, *117* (2), 895-901. DOI: 10.1073/pnas.1910113117.

(57) Beckham, J. L.; Wyss, K. M.; Xie, Y. C.; McHugh, E. A.; Li, J. T.; Advincula, P. A.; Chen, W. Y.; Lin, J.; Tour, J. M. Machine Learning Guided Synthesis of Flash Graphene. *Advanced Materials* **2022**, *34* (12), 2106506. DOI: 10.1002/adma.202106506.

(58) Campos, J. L. E.; Miranda, H.; Rabelo, C.; Sandoz-Rosado, E.; Pandey, S.; Riikonen, J.; Cano-Marquez, A. G.; Jorio, A. Applications of Raman spectroscopy in graphene-related materials and the development of parameterized PCA for large-scale data analysis. *Journal of Raman Spectroscopy* **2018**, *49* (1), 54-65, Article. DOI: 10.1002/jrs.5225.

(59) Ferrari, A. C.; Basko, D. M. Raman spectroscopy as a versatile tool for studying the properties of graphene. *Nature Nanotechnology* **2013**, *8* (4), 235-246, Review. DOI: 10.1038/nnano.2013.46.

(60) Thomsen, C.; Reich, S. Doable resonant Raman scattering in graphite. *Physical Review Letters* **2000**, *85* (24), 5214-5217. DOI: 10.1103/PhysRevLett.85.5214.

(61) Adams, M. *Chemometrics in Analytical Spectroscopy*; The Royal Society of Chemistry, 2004. DOI: 10.1039/9781847550484.

(62) Sena, M. M.; Almeida, M. R.; Braga, J. W. B.; Poppi, R. J. Multivariate Statistical Analysis and Chemometrics. In *Spectroscopic Methods in Food Analysis*, 1 ed.; Franca, A. S., Nollet, L. M. L. Eds.; CRC Press, Taylor & Francis Group, 2017; pp 273-316.

(63) Bro, R.; Smilde, A. K. Principal component analysis. *Analytical Methods* **2014**, *6* (9), 2812-2831. DOI: 10.1039/c3ay41907j.

(64) Wold, H. Soft modelling: The Basic Design and Some Extensions. In *Systems Under Indirect Observation*, Vol. I and II; 1982.

(65) Wold, S.; Martens, H.; Wold, H. The multivariate calibration problem in chemistry solved by the PLS method. In *Matrix Pencils*, Springer Berlin Heidelberg: 1983; pp 286-293.

(66) Kruger, U.; Zhou, Y.; Wang, X.; Rooney, D.; Thompson, J. Robust partial least squares regression: Part I, algorithmic developments. *Journal of Chemometrics* **2008**, *22* (1-2), 1-13. DOI: 10.1002/cem.1093.

(67) Muratore, M. Raman spectroscopy and partial least squares analysis in discrimination of peripheral cells affected by Huntington's disease. *Analytica Chimica Acta* **2013**, *793*, 1-10. DOI: 10.1016/j.aca.2013.06.012.

(68) Rosipal, R.; Krämer, N. Overview and Recent Advances in Partial Least Squares. Berlin, Heidelberg, 2006; Springer Berlin Heidelberg: pp 34-51.

(69) Wold, S.; Sjostrom, M.; Eriksson, L. PLS-regression: a basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems* **2001**, *58* (2), 109-130. DOI: 10.1016/s0169-7439(01)00155-1.

(70) Andries, J. P. M.; Vander Heyden, Y.; Buydens, L. M. C. Predictive-property-ranked variable reduction in partial least squares modelling with final complexity adapted models: Comparison of properties for ranking. *Analytica Chimica Acta* **2013**, *760*, 34-45. DOI: 10.1016/j.aca.2012.11.012.

(71) Barton, B.; Thomson, J.; Diz, E. L.; Portela, R. Chemometrics for Raman Spectroscopy Harmonization. *Applied Spectroscopy* **2022**, *76* (9), 1021-1041. DOI: 10.1177/00037028221094070.

(72) Höskuldsson, A. PLS regression methods. *Journal of Chemometrics* **1988**, *2* (3), 211-228. DOI: https://doi.org/10.1002/cem.1180020306.

(73) Varmuza, K.; Filzmoser, P. Partial Least Square Regression. In *Introduction to Multivariate Statistical Analysis in Chemometrics*, 1st ed.; CRC Press, 2009; p 336.

2 The application of Raman maps for the spatial characterization of amorphous carbon.

This chapter is published as: B.G. daFonseca; S.S. Thind and A.G. Brolo, Raman maps reveal heterogeneous hydrogenation on carbon materials, J. Raman Spectrosc. 2021, 52, 516-524. Copyright: Reproduced with permission.

All the Raman spectra were obtained by myself. Dr. S.S. Thind synthesized all the carbon films. Dr. A. G. Brolo contributed with suggestions for the design of the experiments. The article was written by myself and edited by Dr. A. G. Brolo.

2.1 Abstract

This work presents an application of Raman spectroscopy as a tool to investigate heterogeneity in the composition of carbon materials obtained electrochemically. A combination of Raman maps and histograms have been used to describe samples synthesized via wet chemistry. The results showed that a simple evaluation of an average spectrum or one single spectrum per sample would have hidden important compositional variations present in the sample. The Raman maps revealed heterogeneous hydrogenation in different areas of the carbon films, while the histograms described the statistical relevance of the classification of the different types of carbon materials. The effect of electrosynthesis parameters on the quality of the films were also investigated. As the deposition time increased, the carbon films showed higher homogeneity in their spatial composition. The nature of the electrolyte led to differences in film functionalization and on the degree of hydrogenation.

2.2 Introduction

Amorphous carbon materials have outstanding properties that can be useful in a variety of applications.[1-3] For instance, amorphous carbon with predominance of tetrahedral (sp^3) structure, known

as diamond-like carbon (DLC), have mechanical characteristics similar to diamond.[4-7] The physical and chemical properties of carbon materials can be tailored by varying the predominant carbon hybridization (sp^2 or sp^3), levels of crystallinity and by chemical modification.[5, 8-10] Carbon materials and coatings play an important role in several modern applications, including as materials for engine components, automobile parts, cutting tools, optical components and biomedical devices.[11-14] The current techniques for the fabrication of high quality amorphous carbon films and materials require complex and expensive equipment.[15] Therefore, there is a large effort in academic and commercial laboratories geared towards simple approaches to generate low cost amorphous carbon films. One approach for obtaining low-cost amorphous carbon films is electrochemical deposition. Different syntheses have been reported in literature varying the applied potential (from 2 V to 2000 V), electrolyte (acetonitrile, DMF, nitromethane, etc), temperature (from 20 °C to 60 °C), time of deposition (1 h to 24 h) and substrate (Si, Ti, ITO-coated glass, Al).[12, 15-27] The only common parameter among all of them is the use of graphite as the anode.

Raman spectroscopy is broadly used as one of the main tools for the characterization of carbonaceous materials. [4, 7, 9, 28, 29] Raman spectra of carbon compounds and films, obtained using visible laser excitation, are generally dominated by two strong features around the ~1360 cm^{-1}-region (D-band) and the 1580 cm^{-1}-region (G-band). Carbon allotropes, such as diamond, graphite and graphene, are easily identifiable by the characteristics of those marker bands.[4] Additional spectral features of the D- and G-bands (relative intensity, band position and full-width of half-maximum) can be used to describe other material properties, including crystallinity, degree of functionalization (e.g. hydrogenation) and even the number of layers in graphene materials.[30-32] A general Raman model to qualitatively distinguish amorphous carbon materials rich in carbon sp^3 (diamond-like carbon or DLC) from samples with high sp^2 carbon content (graphite-like) has been proposed.[7] Such model uses the values of the G-band position , the intensity ratio between D and G bands (I(D)/I(G)) and the full width at half-maximum of the G-band of different forms of standard carbon to describe an amorphization trajectory for carbon materials and films.

In addition, an analysis of the broadband emission background in the Raman spectrum of amorphous carbon can be used to estimate the hydrogen content in the material or film.[4, 8, 9, 33]

The versatility and non-destructive properties of visible Raman spectroscopy is then attractive to material scientists concerned with the electrochemical synthesis of carbon materials and films.[29] In most reports of carbon thin films prepared electrochemically, the Raman characterization was based on a single spectrum or the average from measurements obtained in different areas of the material or film. This approach is obviously appropriate when the synthesized material is relatively homogeneous, such as carbon thick films or materials prepared by physical vapour deposition. However, thin films with a higher degree of variability might be obtained in electrochemical deposition and the use of a single or an average spectrum might not be representative of the overall composition of the material. For instance, amorphous carbon films have been produced by electrochemical reduction of CO_2 on steel electrodes using applied potentials ranging between 3 and 10 V at 50 °C.[16] The resulting amorphous carbon films were analyzed using just one Raman spectrum from each sample, but no information about potential heterogeneities in the film composition was reported. Gupta and Rahini have synthesized hydrogenated amorphous carbon using a solution of acetonitrile and water as electrolyte, under 16 V for 15 h.[15] In all these studies the film composition formed electrochemically is assigned based on a single Raman spectrum.

In this work, the use of Raman mapping is explored for the description of heterogeneities encountered in carbon films obtained electrochemically. This approach should be used to evaluate the effects of fabrication parameters (deposition time, chemical bath, applied potentials) and to infer the quality of the synthesized carbon material or film.

2.3 Experimental

2.3.1 Materials and Electrochemical System

Solutions were prepared using hydrochloric acid (Sigma-Aldrich, 37 %), nitric acid (Sigma-Aldrich, 70 %), and water from a Barnstead Nanopure water purification system (18.2 MΩ cm). Two electrodes electrochemical cell consisting of a glass vial closed with an appropriate rubber lid was used. Both electrodes (Graphite and Ti) were introduced through the rubber lid and connected to a DC power supply. All solutions were deaerated with ultrapure argon (99.999 %) prior to the experiments.

2.3.2 Deposition of Carbon on Ti plates.

Initially, the Ti plate was cleaned by sonication in acetone for 15 min, after which it was etched in 18 % HCl at 85 °C for 10 min. A two-electrode system was used to deposit carbon films on the Ti plate, which served as a cathode while the graphite electrode was chosen as an anode. To drive the current through the electrochemical cell, a power supply was used. The exfoliation of graphite anode was studied in 0.05 M HCl solution, where 3 V was applied for 1 h (sample HCl(1h)) or 2 h (sample HCl(2h)). Similarly, exfoliation of graphite anode was performed in 0.05 M HNO_3 solution, where 3 V was applied for 1 h (HNO_3(1h)) or 2 h (HNO_3(2h)). After the deposition of the exfoliated carbon on the Ti plate, the plate was removed and rinsed in deionized water to remove excess electrolyte solution and loose carbon materials. The resulting films were dried and the amorphous carbon - coated Ti materials were used for the Raman analysis.

2.3.3 Raman measurements and data treatment

Raman spectra were obtained using a Renishaw inVia Raman Microscope with 50x objective lens, backscattering collection, and 532 nm laser excitation. In order to avoid any laser damage or excessive local heating at the sample, a neutral density filter has been used to set the laser power to only 0.5 % of the value at the laser head (nominally, 1.75 mW). The exposure time was 5s and 3 accumulations were acquired for each measurement.[7] Matlab code was used to fit the Raman spectrum in the regions of the D- and G- bands (between 1050 and 1800 cm^{-1}) with two Gaussians and a linear baseline. [34] An example of baseline subtraction and fitting is shown in the Appendix A (Figure A-1). The slope of the linear fluorescent background (*m*) divided by G-band intensity (I(G), *i.e.* peak height) were used to estimate the percentage of hydrogen atoms in the film (*H[at.%]*), according to equation (1) below. This equation was obtained by *Casiraghi et al.* by comparing the Raman spectra of different materials with results from nuclear reaction analysis, and further details can be found in their paper. [8]

$$H[at\ \%] = 21.7 + 16.6 \log\left\{\frac{m}{I(G)}\,[\mu m]\right\} \quad \textbf{(1)}$$

Five Raman maps (containing 196 spectra each) were obtained for every sample. The step size for the maps was 2.5 µm and the total area analyzed in each map was 32.5 µm by 32.5 µm, (~ 1056 μm^2).

2.4 Results and discussion

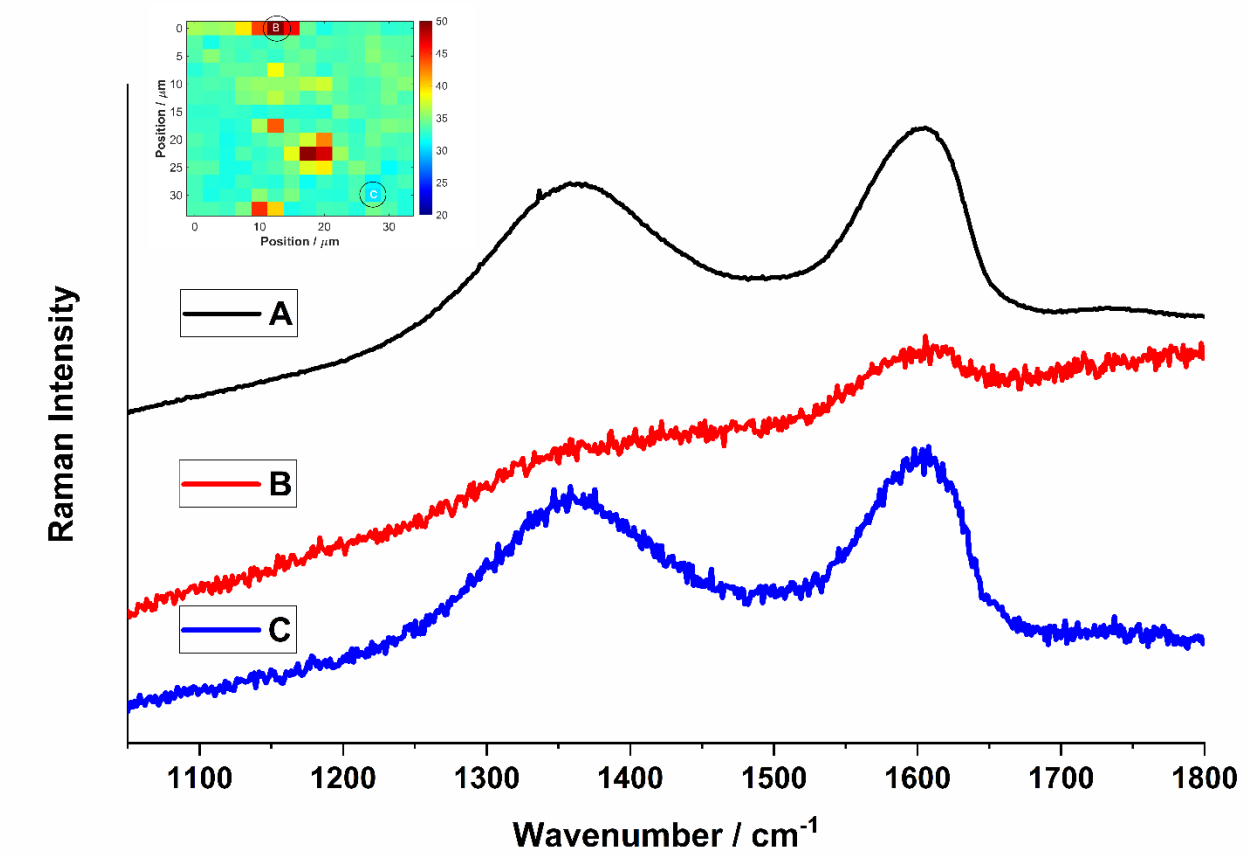

Figure 2-1: Comparison between the average Raman spectrum and two individual spectrum in different spots of the HCl(1h) film. A) The average spectrum from the Raman map shown in the inset (196 spectra). B) The Raman spectrum of the spot with highest H[at. %] in the Raman map. C) The Raman spectrum of spot with the least H[at. %] in the Raman map.

Raman mappings were obtained from electrochemically synthesized carbon materials to demonstrate how the Raman method can unveil spatial variations in the composition of films obtained by wet methods. Figure 2-1 compares the mean spectrum from one of the Raman mappings (32.5 x 32.5 μm^2) to two individual spectra obtained in different locations of the film. Clear differences in the spectral characteristics observed between Figure 2-1A, Figure 2-1B and Figure 2-1C emphasizes a heterogeneous distribution of carbon species throughout the film. The Raman markers of the mean spectrum, G-band at 1595 cm^{-1}, G-band full width at half maximum of 91 cm^{-1} and intensity ratio of D- and G-bands of 0.87

suggest an disordered sp^2 material mainly composed of reduced graphite oxide. This is not surprising due to the synthetic method utilized in this work.[35] The carbon film used in Figure 2-1 (HCl(1h)) was obtained through exfoliation and oxidation of a graphite electrode, followed by reduction on the titanium electrode, in the presence of HCl. The outstanding difference between the average spectrum in Figure 2-1A and a typical graphite oxide spectrum is the fluorescent background observed for the electrochemically produced film. This background is assigned to amorphous carbon materials rich in hydrogen.[8] The hydrogen content estimated for the mean spectrum (Figure 2-1A) was 34.9 % using equation (1).[8] In contrast, the Raman spectrum presented in Figure 2-1B has the highest hydrogen content of the map (H[at.%] = 50 %). Other Raman parameters, such as G-band at 1595 cm^{-1}, G-band full width at half maximum of 105.4 cm^{-1} and intensity ratio of D- and G-bands of 0.86, suggest that the spot in Figure 2-1B could be classified as a hydrogenated polymer-like carbon (PLCH) film. The Raman spectrum in Figure 2-1C (for a different spot in the film, spot 2) contains the least amount of hydrogen (H[at.%] = 30.5 %). The smaller amount of hydrogen (Figure 2-1C) combined to other Raman features, such as G-band at 1595 cm^{-1}, G-band full width at half maximum of 91.6 cm^{-1} and intensity ratio of D- and G-bands of 0.85, suggests that this spot is dominated by hydrogenated reduced graphite oxide (HRGO).[35] Therefore, the average spectrum (Figure 2-1A) could be assigned mainly to HRGO and the Raman mapping showed that compositional ratio of two carbon species (HRGO and PLCH) vary throughout the film, even in the scale of a few micrometers. This simple comparison shows that chemical information might be hidden when the film or material characterization is performed using only one Raman spectrum or only the average. In addition, the compositional variation in space raises questions about the statistical relevance of the classification of different types of films reported previously in the literature.

The spectral region concerning the 2D-band is shown in the Figure A-2. In graphene and graphite the 2D band is strong and shows up at 2700 cm^{-1}. In our sample, only weak bands are observed in the region between 2300 and 3200 cm^{-1}. Such behavior has been observed previously in the literature.[36]

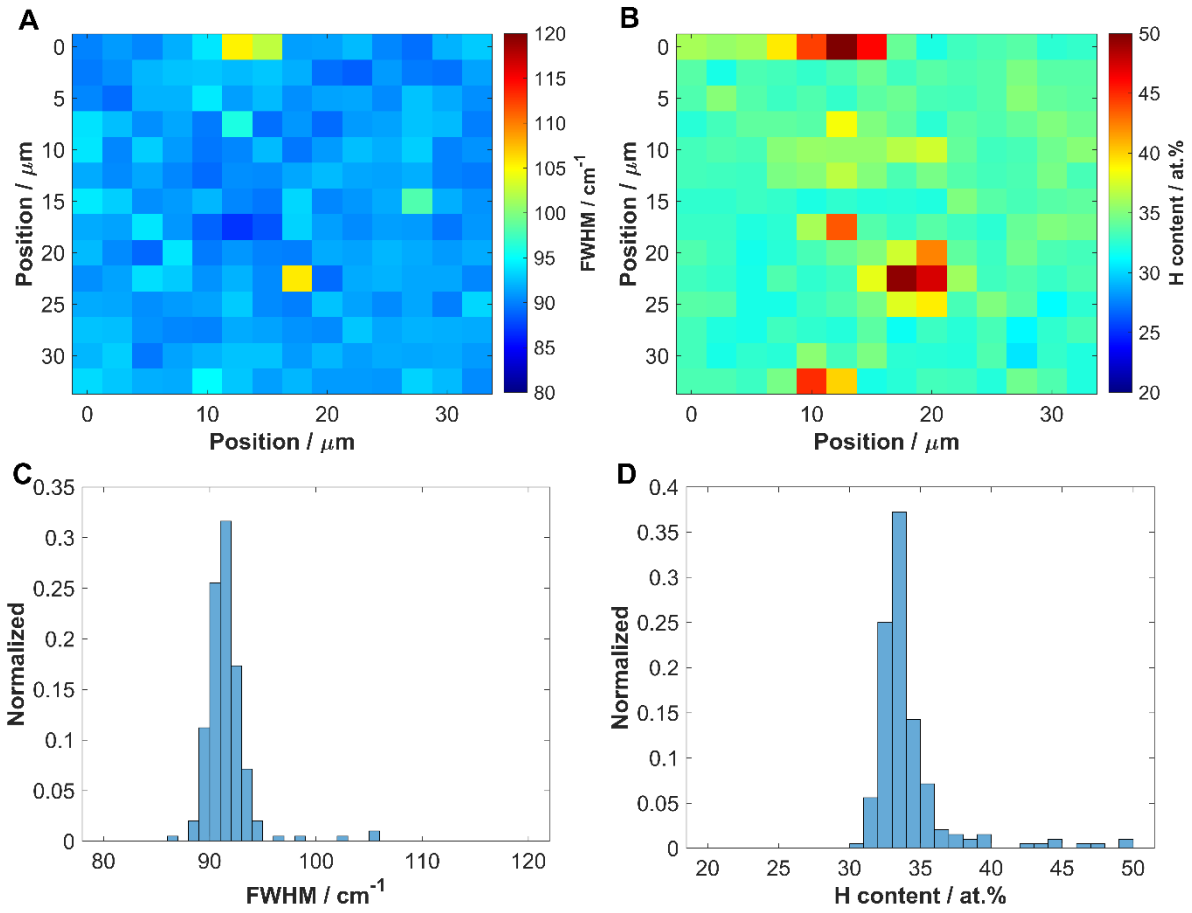

Figure 2-2: Spatial and statistical distribution of hydrogen content and G-band full width at half maximum obtained using Raman maps for the HCl (1h) film. A) Raman map showing the spatial variation of G-band full width at half maximum; B) Raman map showing the spatial variation of hydrogen content; C) Distribution of G-band full width at half maximum values; D) Distribution of hydrogen content values.

Raman maps of the film investigated in Figure 2-1 (HCl (1h) sample) are presented in Figure 2-2A(spatial plot of the G-band full width at half maximum) and Figure 2-2B (spatial plot of the Hydrogen content). Histograms of these quantities (G-band full width at half maximum and Hydrogen content) are shown in Figure 2-2C and Figure 2-2D. The mean and standard deviation of G-band full width at half maximum was 91.6 cm^{-1} ±2.1 cm^{-1}, while individual values ranged from 86 cm^{-1} up to 106 cm^{-1}. The individual histograms, mean and standard deviation values can be found in the support information

(Appendix A). Close to 75 % of the spectra analyzed in Figure 2-2 had G-band full width at half maximum-values between 90 and 93 cm^{-1}, and the map in Figure 2-2A shows that the G-band full width at half maximum-values were distributed throughout the film and not concentrated in only one region. The full width at half maximum is an indicator for structure crystallinity, with lower full width at half maximum-values suggesting more crystalline materials.[7, 28] The amorphous characteristics in carbon films can be correlated to many structural aspects, including the density of sp^3 C-C bonds or degree of hydrogenation.[7, 9] The hydrogen content, presented in Figure 2-2D, showed variation from ~ 30 to 50 %, with average values at 34.2 ± 2.9 %, and no overconcentration of hydrogen in just one region. The contrast between the colors in the maps (Figure 2-2A and 2B) highlight the heterogeneity of the composition in this analyzed region. In terms of hydrogen content, Figure 2-2 indicates that almost 70 % of the spots have values lower than the average (34.2 ± 2.9 at.%), reinforcing the heterogeneity of this sample and the contribution of few spots for higher mean value of hydrogen content. The histograms shown in Figure 2-2C and Figure 2-2D also reveal that the data does not follow a normal distribution. The asymmetry in the histograms is related to the higher content of hydrogen (more than 35 %) in the regions dominated by the PLCH material, corresponding to 17 % of the film, while the remaining 83 % of the film is composed of HRGO.[9, 33]

The maps for other Raman markers (G-band position (mean of 1595 cm^{-1}) and intensity ratio of D- and G-bands (mean 0.84)) are shown in Appendix A (Figure A-7). These quantities (G-band position and intensity ratio of D- and G-bands) presented lower variation than the ones observed for G-band full width at half maximum and hydrogen content, and their respective histograms are much similar to normal distributions. The sp^2 carbon structures are responsible for the G- and D- bands, while the PLCH structures leads to a fluorescence background. The maps and histograms in Figure 2-2 confirm that at two different types of carbon materials were electrochemically coated on the electrode surface. Although the sp^2 signature contributed more homogeneously throughout the film (notice the relative small variations in G-

band position and intensity ratio of D- and G-bands), the spatial changes in hydrogen content suggest various degrees of reduction which can affect the material properties.[7, 28] Comparing the approaches used in Figure 2-1 (mean or only one spectrum) and in Figure 2-2 (maps and histograms), it is clear that, for electrochemically obtained carbon materials, variations in chemical composition can be revealed by Raman maps even in a relative small area.

The applied potential used in this work was 3 V, which is much higher than required for water splitting. Therefore, hydrogen evolution was observed at the titanium electrode during the synthesis of the carbon film.[15, 32] Similar synthesis conditions have been used for the hydrogenation of graphene to produce graphane, or hydrogenated graphene.[37-39] Hence, the similarities observed between Raman features (G-band position, G-band full width at half maximum and intensity ratio of D- and G-bands) of the HCl(1h) sample with reported in the literature for graphite oxide and the presence of hydrogen in the film are not surprising. Reduced graphene oxide has a Raman spectrum that is more similar to the one for graphene oxide than for graphene itself. [40, 41] This is due to the effect of disorder and the remaining oxygen functional groups that are not reduced and become part of the film. [40]

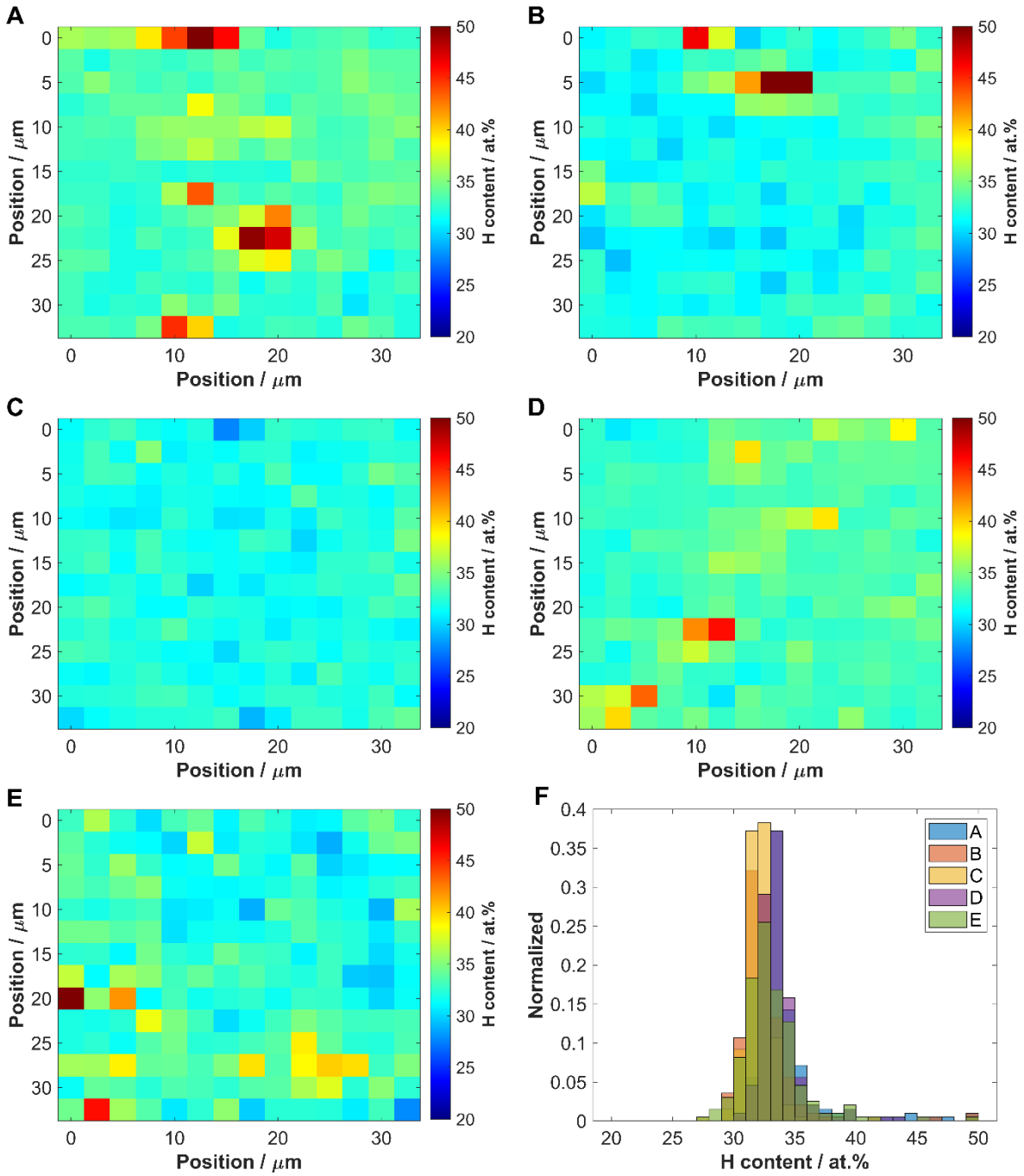

Figure 2-3: Hydrogen content maps at different regions of the HCl(1h) film. A) to E) hydrogen content spatial distribution in random regions of the same HCl(1h) sample. F) Histograms of hydrogen content for all regions (A to E) plotted on the same graph to illustrate the differences in statistical distribution from different regions. The individual histograms and average values are presented in Appendix A.

The results from Figure 2-2 indicated that composition heterogeneities in electrochemically formed films can be identified by Raman, even if they occur within a few μm. However, it is important to

verify if the composition distributions varies in larger spatial scales (mm to cm). Figure 2-3 presents 5 maps of hydrogen content obtained for different random regions in a carbon material synthesized electrochemically on Ti (HCl(1h) sample). The results in Figure 2-3 correspond to almost 1000 Raman spectra obtained at different points of the synthesized carbon material. Although all maps have a similar value for the average hydrogen content, between 32 and 34.2 %, the color contrast in the different maps quickly provides a qualitative picture of the composition distribution throughout the whole film. The spots with high hydrogen content (H[at.%] ≥ 40 %), orange, red and brown shades in the maps, identify the regions rich in PLCH. Quantitatively, the maps in Figure 2-3A, Figure 2-3B, Figure 2-3D and Figure 2-3E contain 17, 5.6, 12 and 12.7 % of PLCH, respectively, while the map in Figure 2-3C contains only 0.5 % of PLCH and it is the most homogeneous region among these five maps. Similarly to Figure 2-2, the regions with high hydrogen content values (dominated by PLCH) are correlated to larger values of G-band full width at half maximum (more disorder).

The distribution of hydrogen content values for each map (Figure 2-3A and Figure 2-3E) are represented together as histograms in Figure 2-3F. In Figure 2-3F, the histograms are overlapped and color coded, but they can be seen individually as separated plots in the support information file (Figure A-7 to Figure A-11). In addition, Figure 2-3F indicates small variations in the distribution of hydrogen content (H[at.%]) obtained from different regions. For instance, broader hydrogen content distributions were recorded for regions A and E compared to region C. In general, the distributions for hydrogen content were not symmetric around the average (i.e. normal or Gaussian distribution), with a tail towards high amounts of hydrogen (Figure 2-3E).

The set of maps in Figure 2-3 indicates that the film is composed by about 9.6 % PLCH, while the remaining of the material can be classified as HRGO.[7-9, 28] In order to achieve a better description of the sample, only one Raman map and its histogram would be better than a single Raman spectrum, although

the use of other maps improved the analysis, showing other variations in composition that can be observed throughout the sample.

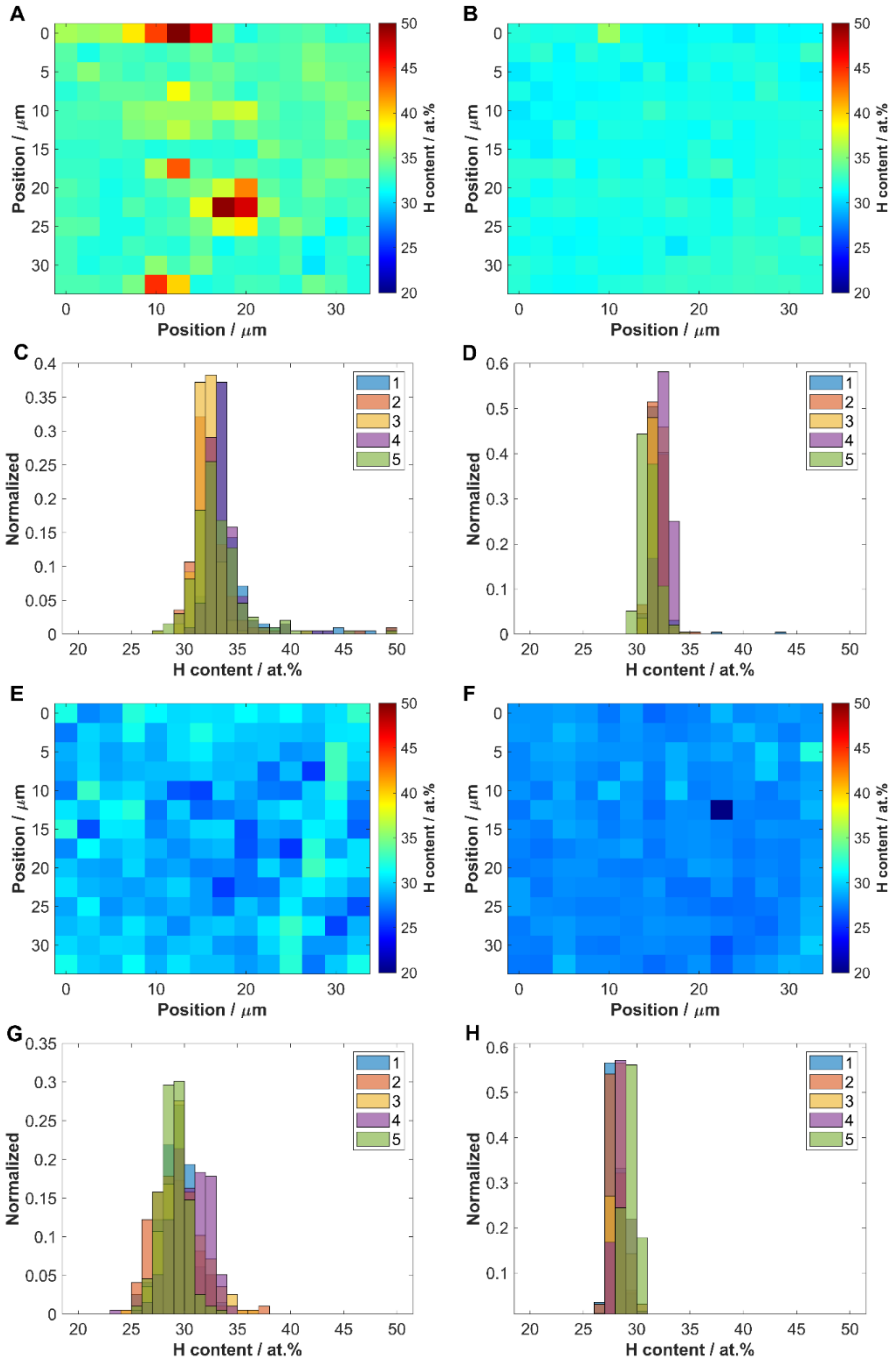

Figure 2-4: Using Raman maps to determine the effect of synthetic electrochemical parameters, electrolyte and coating time, with film composition. A) Hydrogen content map for one region of the HCl(1h) film. B) Hydrogen content map for one region of the HCl(2h) film. C) Five histograms of five different regions of the HCl(1h) film. D) Five histograms of the five regions of the HCl(2h) film. E) Hydrogen content map of one region of the HNO3(1h) film. F) Hydrogen content map of one region of the HNO3(2h) film. G) Five histograms of five regions of a HNO3(2h) film. The individual histograms and average values are presented in the supporting information (Appendix A).

Figure 2-4 illustrates the utilization of Raman mapping to determine the relationship between film composition/homogeneity and electrochemical synthetic parameters. Figure 2-4A and Figure 2-4B present hydrogen content (H[at.%]) maps for electrochemically-synthesized carbon films under the same conditions, in HCl (as indicated in the experimental section), but the deposition times were 1 h for Figure 2-4A (HCl(1h) film) and 2 h for Figure 2-4B (HCl(2h) film). In addition, Figure 2-4E and Figure 2-4F are hydrogen content maps for films prepared in HNO_3 using 1 h (HNO_3(1h)) and 2 h (HNO_3(2h)) electrodeposition times, respectively. Notice that, as in Figure 2-3, Raman maps were obtained in several random regions of the film and the hydrogen content map is only a representative example (all other maps (different regions and other Raman parameters) are available in Appendix A). A quick visual inspection of Figure 2-4A and Figure 2-4B indicates that the HCl(1h) carbon film presented higher hydrogen content (shades of yellow, orange and brown in Figure 2-4A) then the HCl(2h) film (mostly light blue map in Figure 2-4B). The average hydrogen content-value decreased from 34.2 % to 31.9 % and the composition of the film become more homogeneous as the deposition time increased from 1 h (Figure 2-4A) to 2 h (Figure 2-4B). Around 17 % of the spots in Figure 2-4A can be assigned to PLCH, in contrast to only 0.5 % of PLCH in Figure 2-4B. Maps for other Raman parameters (G-band position, intensity ratio of D- and G-bands and G-band full width at half maximum) are presented in the support information. The mean G-band full width at half maximum-values calculated for HCl(1h) and HCl(2h) are similar (90.5 and 91.8 cm^{-1}, respectively). The G-band position and intensity ratio of D- and G-bands maps presented less differences between HCl(1h) and HCl(2h) than what was observed for hydrogen content and G-band full width at half maximum.

The values of hydrogen content maps obtained for HCl(1h) and HCl(2h) samples, including the ones in Figure 2-4A and Figure 2-4B, are summarized as histograms presented in Figure 2-4C and Figure 2-4D. The distribution of hydrogen content was clearly narrower for long electrodepositions. Figure 2-4D also shows that regions with hydrogen content values higher than 35 % are not present for the 2 h film

deposition. All other Raman parameters have narrower distribution for longer deposition (see Appendix A). This set of evidences suggest that longer electrodeposition times leads to more homogeneous and less polymeric films, or, more precisely, to HRGO (with hydrogen content ~ 32 at%) films. Crystalline metallic surfaces are known to induce the formation of graphitic structures due to the epitaxial energy of the metal, epitaxial graphene is something well know in the community. On the other hand, it's been shown the presence of hydrogen can induce other types of films. Also, the border of graphene is described as more reactive and hydrogen evolution is being observed. Then, further investigation of thinner films must be carried out for better understanding.[10, 42-44]

Figure 2-4E and Figure 2-4F shows the hydrogen content maps for carbon films synthesized in HNO_3 medium with different electrodeposition times (1 h and 2 h, respectively). As it was the case for the HCl films, several maps were obtained from different regions of the film and the results in Figure 2-4E and Figure 2-4F are just representative examples. The average hydrogen contents obtained from those maps were hydrogen content = 29.3 ± 1.6 % and hydrogen content = 27.9 ± 0.9 %, for Figure 2-4E and Figure 2-4F, respectively. The color contrast in the maps make it visually clear that the sample obtained after 1 h was more heterogeneous than the sample obtained after 2 h of electrodeposition.

The hydrogen contents obtained from different Raman maps for a particular HNO_3(1h) sample are summarized in histograms presented in Figure 2-4G. The overlap between the histograms in Figure 2-4G was around 30 %, implying composition variations between the different probed regions. The broad distribution of hydrogen content values suggest that this sample could be a mix of HRGO.[19, 35, 40] The histograms for hydrogen content obtained for longer deposition time (HNO_3(2h)), presented in Figure 2-4H, are narrower than the ones presented in Figure 2-4G, reinforcing the composition homogeneity induced by longer coating. The average hydrogen content value is much smaller for the carbon films deposited in HNO_3 than in HCl, probably due to the oxidation characteristics of the nitrate ion. In fact, only a small fraction of the film deposited in HNO_3 could be classified as PLCH.

The average values of intensity ratio of D- and G-bands for the HCl(2h) and HNO_3(2h) are 0.89 ± 0.01 and 0.84 ±0.01, respectively. It has been reported that an increased value of the intensity ratio of D- and G-bands is expected for reduced carbon species. The results from Figure 2-4 and the average intensity ratio of D- and G-bands ratios corroborate the fact that the carbon films formed in HCl film are less oxidized than the ones formed in HNO_3.[35, 40, 41] Although some graphane works interpret the increasing intensity ratio of D- and G-bands as higher disorder, the amorphization trajectory proposed by Ferrari et al indicates that after a certain limit, any augment in disorder leads to a decrease in intensity ratio of D- and G-bands (correlated to an increase in both G-band position and G-band full width at half maximum).[7, 28, 45] The D-band is activated by disorder in the sp^2 lattice, and also assigned to ring breathing modes. Then, the initial defects promote stronger D-bands, while more disorder leads to rings being broken and; therefore, decrease in D-band intensity. In addition, it has been reported in the literature that the hydrogenation reaction also promote the formation of double bonds.[45] In conclusion, the carbon films electrodeposited in HCl media are more hydrogenated and have more aromatic rings than the ones deposited in HNO_3.

The results shown in Figure 2-4 implies that variation in synthetic parameters in electrochemical synthesis of carbon films can have a profound effect in the film composition and homogeneity. Raman maps and histograms helps to guide the synthesis and on the evaluation of the quality of the film.

2.5 Conclusion

Raman maps and histograms of G-band position, full width at half maximum, relative intensity of G- and D-band, and hydrogen content are valuable tools for the analysis of carbon materials and films and should be utilized for a full characterization of electrochemically-prepared materials. Several Raman parameters (G-band position, full width at half maximum, relative intensity of G- and D-band, and

hydrogen content) can be plotted spatially to evaluate local chemical composition and properties. For instance, carbon coatings with different levels of hydrogenation can be correlated to local reduction conditions (hydrogen evolution, presence of oxidizing anions) and different levels of disorder. Variations in film composition seems to be highly dependent on the electrodeposition time, indicating that thicker films tend to be more homogeneous. Electrochemical synthesis provides several advantages that include low cost and easy functionalization. However, due to the potential for local variations in composition, Raman mapping should be used to provide better quantification and characterization of the carbon films produced by this method.

2.6 References

(1) Hainsworth, S. V.; Uhure, N. J. Diamond like carbon coatings for tribology: production techniques, characterisation methods and applications. *International Materials Reviews* **2007**, *52* (3), 153-174. DOI: 10.1179/174328007x160272.

(2) Triroj, N.; Saensak, R.; Porntheeraphat, S.; Paosawatyanyong, B.; Amornkitbamrung, V. Diamond-Like Carbon Thin Film Electrodes for Microfluidic Bioelectrochemical Sensing Platforms. *Analytical Chemistry* **2020**, *92* (5), 3650-3657, Article. DOI: 10.1021/acs.analchem.9b04689.

(3) Palomaki, T.; Wester, N.; Caro, M. A.; Sainio, S.; Protopopova, V.; Koskinen, J.; Laurila, T. Electron transport determines the electrochemical properties of tetrahedral amorphous carbon (ta-C) thin films. *Electrochimica Acta* **2017**, *225*, 1-10, Article. DOI: 10.1016/j.electacta.2016.12.099.

(4) Ferrari, A. C.; Robertson, J. Raman spectroscopy of amorphous, nanostructured, diamond-like carbon, and nanodiamond. *Philosophical Transactions of the Royal Society a-Mathematical Physical and Engineering Sciences* **2004**, *362* (1824), 2477-2512, Article. DOI: 10.1098/rsta.2004.1452.

(5) Ferrari, A. C.; Rodil, S. E.; Robertson, J. Resonant Raman spectra of amorphous carbon nitrides: the G peak dispersion. *Diamond and Related Materials* **2003**, *12* (3-7), 905-910, Article; Proceedings Paper. DOI: 10.1016/s0925-9635(02)00370-9.

(6) Robertson, J. Diamond-like amorphous carbon. *Materials Science & Engineering R-Reports* **2002**, *37* (4-6), 129-281. DOI: 10.1016/s0927-796x(02)00005-0.

(7) Ferrari, A. C.; Robertson, J. Interpretation of Raman spectra of disordered and amorphous carbon. *Physical Review B* **2000**, *61* (20), 14095-14107. DOI: 10.1103/PhysRevB.61.14095.

(8) Casiraghi, C.; Piazza, F.; Ferrari, A. C.; Grambole, D.; Robertson, J. Bonding in hydrogenated diamond-like carbon by Raman spectroscopy. *Diamond and Related Materials* **2005**, *14* (3-7), 1098-1102, Article; Proceedings Paper. DOI: 10.1016/j.diamond.2004.10.030.

(9) Casiraghi, C.; Ferrari, A. C.; Robertson, J. Raman spectroscopy of hydrogenated amorphous carbons. *Physical Review B* **2005**, *72* (8), 085401, Article. DOI: 10.1103/PhysRevB.72.085401.

(10) Irmer, G.; Dorner-Reisel, A. Micro-Raman studies on DLC coatings. *Advanced Engineering Materials* **2005**, *7* (8), 694-705. DOI: 10.1002/adem.200500006.

(11) Tyagi, A.; Walia, R. S.; Murtaza, Q.; Pandey, S. M.; Tyagi, P. K.; Bajaj, B. A critical review of diamond like carbon coating for wear resistance applications. *International Journal of Refractory Metals & Hard Materials* **2019**, *78*, 107-122. DOI: 10.1016/j.ijrmhm.2018.09.006.

(12) Awasthi, S.; Pandey, S. K.; Pandey, C. P.; Balani, K. Progress in Electrochemical and Electrophoretic Deposition of Nickel with Carbonaceous Allotropes: A Review. *Advanced Materials Interfaces* **2020**, *7* (1), 1901096, Review. DOI: 10.1002/admi.201901096.

(13) Tamulevicius, S.; Meskinis, S.; Tamulevicius, T.; Rubahn, H. G. Diamond like carbon nanocomposites with embedded metallic nanoparticles. *Reports on Progress in Physics* **2018**, *81* (2), 024501. DOI: 10.1088/1361-6633/aa966f.

(14) Caschera, D.; Toro, R. G.; Cortese, B.; Federici, F.; Lombardo, D.; Calandra, P. Diamond-like carbon: a versatile material for developing innovative smart textiles applications. A short review. *Atti Accademia Peloritana Dei Pericolanti-Classe Di Scienze Fisiche Matematiche E Naturali* **2019**, *97, A27*. DOI: 10.1478/aapp.97s2a27.

(15) Gupta, S.; Rahini, R. L. Electrochemical synthesis and characterization of amorphous hydrogenated carbon (a-C:H) using acetonitrile as electrolyte. *Electrochimica Acta* **2017**, *258*, 1-8, Article. DOI: 10.1016/j.electacta.2017.10.079.

(16) Wu, M. X.; Tang, R. J.; Chen, Y. L.; Wang, S. Y.; Wang, W. C.; Chen, X. H.; Mitsuzaki, N.; Chen, Z. D. Electrochemical reduction of CO2 to carbon films on stainless steel around room temperature. *Electrochemistry Communications* **2020**, *110,* 106606. DOI: 10.1016/j.elecom.2019.106606.

(17) Falcade, T.; Turq, V.; Bonino, J. P.; Malfatti, C. D. Tribological properties of amorphous carbon films obtained by electrodeposition from DMF using 2HEAL protic ionic liquid as dopant. *Diamond and Related Materials* **2017**, *71*, 30-37. DOI: 10.1016/j.diamond.2016.11.014.

(18) Jiang, N.; Chiang, N.; Madison, L. R.; Pozzi, E. A.; Wasielewski, M. R.; Seideman, T.; Ratner, M. A.; Hersam, M. C.; Schatz, G. C.; Van Duyne, R. P. Nanoscale chemical imaging of a dynamic molecular phase boundary with ultrahigh vacuum tip-enhanced raman spectroscopy. *Nano Letters* **2016**, *16* (6), 3898-3904. DOI: 10.1021/acs.nanolett.6b01405.

(19) Sakita, A. M. P.; Valente, M. A. G.; Della Noce, R.; Fugivara, C. S.; Magnani, M.; Benedetti, A. V. Low-voltage carbon films deposition by electro-exfoliation of graphite into graphene oxide. *RSC Advances* **2016**, *6* (87), 84194-84199, Article. DOI: 10.1039/c6ra16502h.

(20) Zhang, Q.; Wang, Y. R.; Wang, W. C.; Mitsuzak, N.; Chen, Z. D. Low voltage and ambient temperature electrodeposition of uniform carbon films. *Electrochemistry Communications* **2016**, *63*, 22-25, Article. DOI: 10.1016/j.elecom.2015.11.012.

(21) Guo, D.; Cai, K.; Li, L. T.; Zhu, H. S. Preparation of hydrogenated diamond-like carbon films on conductive glass from an organic liquid using pulsed power. *Chemical Physics Letters* **2000**, *325* (5-6), 499-502. DOI: 10.1016/S0009-2614(00)00704-1.

(22) Guo, D.; Cai, K.; Li, L. T.; Huang, Y.; Gui, Z. L.; Zhu, H. S. Evaluation of diamond-like carbon films electrodeposited on an Al substrate from the liquid phase with pulse-modulated power. *Carbon* **2001**, *39* (9), 1395-1398. DOI: 10.1016/S0008-6223(00)00257-8.

(23) Fu, Q.; Jiu, J. T.; Wang, H.; Cao, C. B.; Zhu, H. S. Simultaneous formation of diamond-like carbon and carbon nitride films in the electrodeposition of an organic liquid. *Chemical Physics Letters* **1999**, *301* (1-2), 87-90. DOI: 10.1016/S0009-2614(98)01426-2.

(24) Fu, Q.; Jiu, J. T.; Cao, C. B.; Wang, H.; Zhu, H. S. Electrodeposition of carbon films from various organic liquids. *Surface and Coatings Technology* **2000**, *124* (2-3), 196-200. DOI: 10.1016/S0257-8972(99)00658-1.

(25) Cai, K.; Guo, D.; Huang, Y.; Zhu, H. S. Evaluation of diamond-like carbon films deposited on conductive glass from organic liquids using pulsed current. *Surface and Coatings Technology* **2000**, *130* (2-3), 266-273. DOI: 10.1016/S0257-8972(00)00698-8.

(26) Falcade, T.; Shmitzhaus, T. E.; dos Reis, O. G.; Vargas, A. L. M.; Hubler, R.; Muller, I. L.; Malfatti, C. D. Electrodeposition of diamond-like carbon films on titanium alloy using organic liquids: Corrosion and wear resistance. *Applied Surface Science* **2012**, *263*, 18-24, Article. DOI: 10.1016/j.apsusc.2012.08.052.

(27) Pal, A. K.; Roy, R. K.; Mandal, S. K.; Gupta, S.; Deb, B. Electrodeposited carbon nanotube thin films. *Thin Solid Films* **2005**, *476* (2), 288-294. DOI: 10.1016/j.tsf.2004.09.064.

(28) Ferrari, A. C.; Robertson, J. Resonant Raman spectroscopy of disordered, amorphous, and diamondlike carbon. *Physical Review B* **2001**, *64* (7), 075414. DOI: 10.1103/PhysRevB.64.075414.

(29) Merlen, A.; Buijnsters, J. G.; Pardanaud, C. A Guide to and Review of the Use of Multiwavelength Raman Spectroscopy for Characterizing Defective Aromatic Carbon Solids: from Graphene to Amorphous Carbons. *Coatings* **2017**, *7* (10), 153, Review. DOI: 10.3390/coatings7100153.

(30) Ferrari, A. C.; Basko, D. M. Raman spectroscopy as a versatile tool for studying the properties of graphene. *Nature Nanotechnology* **2013**, *8* (4), 235-246, Review. DOI: 10.1038/nnano.2013.46.

(31) Tan, P. H.; Han, W. P.; Zhao, W. J.; Wu, Z. H.; Chang, K.; Wang, H.; Wang, Y. F.; Bonini, N.; Marzari, N.; Pugno, N.; et al. The shear mode of multilayer graphene. *Nature Materials* **2012**, *11* (4), 294-300, Article. DOI: 10.1038/nmat3245.

(32) Whitener, K. E. Review Article: Hydrogenated graphene: A user's guide. *Journal of Vacuum Science & Technology A* **2018**, *36* (5), 05G401, Review. DOI: 10.1116/1.5034433.

(33) Marchon, B.; Gui, J.; Grannen, K.; Rauch, G. C.; Ager, J. W.; Silva, S. R. P.; Robertson, J. Photoluminescence and Raman spectroscopy in hydrogenated carbon films. *Ieee Transactions on Magnetics* **1997**, *33* (5), 3148-3150. DOI: 10.1109/20.617873.

(34) O'Haver, T. *peakfit.m*. MATLAB Central File Exchange, https://www.mathworks.com/matlabcentral/fileexchange/23611-peakfit-m (accessed 2020 June 4).

(35) Kudin, K. N.; Ozbas, B.; Schniepp, H. C.; Prud'homme, R. K.; Aksay, I. A.; Car, R. Raman spectra of graphite oxide and functionalized graphene sheets. *Nano Letters* **2008**, *8* (1), 36-41. DOI: 10.1021/nl071822y.

(36) Knight, D. S.; White, W. B. Characterization of diamond films by raman-spectroscopy. *Journal of Materials Research* **1989**, *4* (2), 385-393, Article. DOI: 10.1557/jmr.1989.0385.

(37) Zhao, M.; Guo, X. Y.; Ambacher, O.; Nebel, C. E.; Hoffmann, R. Electrochemical generation of hydrogenated graphene flakes. *Carbon* **2015**, *83*, 128-135. DOI: 10.1016/j.carbon.2014.11.033.

(38) Whitener, K. E.; Robinson, J. T.; Sheehan, P. E. Protection from Below: Stabilizing Hydrogenated Graphene Using Graphene Underlayers. *Langmuir* **2017**, *33* (48), 13749-13756. DOI: 10.1021/acs.langmuir.7b03596.

(39) Pumera, M.; Wong, C. H. A. Graphane and hydrogenated graphene. *Chemical Society Reviews* **2013**, *42* (14), 5987-5995. DOI: 10.1039/c3cs60132c.

(40) Muzyka, R.; Drewniak, S.; Pustelny, T.; Chrubasik, M.; Gryglewicz, G. Characterization of Graphite Oxide and Reduced Graphene Oxide Obtained from Different Graphite Precursors and Oxidized by Different Methods Using Raman Spectroscopy. *Materials* **2018**, *11* (7), 1050. DOI: 10.3390/ma11071050.

(41) Moon, I. K.; Lee, J.; Ruoff, R. S.; Lee, H. Reduced graphene oxide by chemical graphitization. *Nature Communications* **2010**, *1*, 73, Article. DOI: 10.1038/ncomms1067.

(42) Meng, W. J.; Gillispie, B. A. Mechanical properties of Ti-containing and W-containing diamond-like carbon coatings. *Journal of Applied Physics* **1998**, *84* (8), 4314-4321. DOI: 10.1063/1.368650.

(43) Wang, Q. H.; Jin, Z.; Kim, K. K.; Hilmer, A. J.; Paulus, G. L. C.; Shih, C. J.; Ham, M. H.; Sanchez-Yamagishi, J. D.; Watanabe, K.; Taniguchi, T.; et al. Understanding and controlling the substrate effect on graphene electron-transfer chemistry via reactivity imprint lithography. *Nature Chemistry* **2012**, *4* (9), 724-732. DOI: 10.1038/nchem.1421.

(44) Stine, R.; Lee, W. K.; Whitener, K. E.; Robinson, J. T.; Sheehan, P. E. Chemical Stability of Graphene Fluoride Produced by Exposure to XeF2. *Nano Letters* **2013**, *13* (9), 4311-4316. DOI: 10.1021/nl4021039.

(45) Ciammaruchi, L.; Bellucci, L.; Castillo, G. C.; Sanchez, G. M. D.; Liu, Q.; Tozzini, V.; Martorell, J. Water splitting of hydrogen chemisorbed in graphene oxide dynamically evolves into a graphane lattice. *Carbon* **2019**, *153*, 234-241, Article. DOI: 10.1016/j.carbon.2019.06.087.

3 Using Raman spectroscopy and partial least squares regression to quantify carbon species in the mining workplace environment.

This chapter is a manuscript in preparation for submission. The authors and the title of the manuscripts are: daFonseca, BG, and Brolo, AG, Rapid Method for the Quantification of Carbon species in Diesel Particulate Matter by Raman Spectroscopy and Partial Least Squares.

This work was a collaboration with D. S. dosSantos and P. Pityn from OSHTech Incorporated. The partial least squares models and tests were performed by myself. All the Raman spectra were obtained by myself and A. Azarakhshi. D. S. dosSantos and P. Pityn performed measurements using the NIOSH 5040 method.

3.1 Abstract

Raman spectroscopy and partial least squares (PLS) algorithm were used for quantification of elemental and organic carbon in air filters from mining operations. The filters collected diesel particulate matter (DPM) at the respirators of mining workers during their shifts (~8 h). The composition of the various types of carbon was independently analyzed using the standard technique for North America (NIOSH 5040). Raman spectra were then collected from the same samples using 532 and 785 nm laser excitation. A linear correlation between the carbon composition determined by NIOSH 5040 and the Raman spectra could be observed within a specific range of values. Therefore, the composition of elemental and organic forms of carbon in a particular filter were well predicted by Raman. The 785 nm laser excitation results have shown a better accuracy than results obtained from 532 nm. This could be due to a more prominent fluorescent background from the organic fraction of the samples under 532 nm

excitation. The model revealed the fluorescence background was an important source of information for predicting the organic fraction. Therefore, the results presented here demonstrate an avenue for quick evaluation of DPM that can be implemented *in loco* for real-time evaluation of the environmental conditions at the workplace even in low resources conditions.

3.2 Introduction

The World Health Organization (WHO) estimates that air pollution kills six million people worldwide every year, and it is linked to increased morbidity resulting from heart disease, chronic obstructive pulmonary disease, lung cancer and acute respiratory infections.[1-6] Air pollution is among the top three global risk factors for disease and death.[7] Combustion of fossil fuels is a main source of air pollution and greenhouse gas emissions, thus it is closely linked to the earth's climate and impacts on ecosystems. Diesel exhaust has solid components, known as diesel particulate matter (DPM), that poses great concerns.[4, 8-12] Polycyclic aromatic hydrocarbons (PAH), nitro-PAH and other organic molecules in DPM are known carcinogens.[13, 14] As well, DPM contains other chemical species including sulfates, nitrates, metals and trace elements.[15]

Exposures to DPM are defined, quantified and regulated in terms of the carbon content, notionally identified as elemental carbon (EC) and organic carbon (OC) species. EC is thought to be primarily pure carbon particles, mostly as sp^2 hybridized graphitic species.[16] It has been shown that the soot nanoparticles rich in EC are more crystalline compared to the soot nanoparticles rich in OC, which are considered to be amorphous. The EC particles have similarities to turbostratic graphite,[17] whereas the OC soot nanoparticles have similarities to large organic molecules (e.g. humic acid), with functional groups and saturated carbon chains.[17] The EC and OC classification is an operational definition based on the standard technique employed for theirs evaluation.[18]

The current method for sampling and analysis of DPM in North America is the NIOSH Method 5040.[19] In this method the sample is volatilized under specific conditions (solvent and temperature) and measured by gas chromatography with thermal-optical analysis using a flame ionization detector. All the carbon content volatilized below 850 °C in the absence of oxygen is said to be OC. After the removal of OC, carbonate and char, the remaining carbonaceous material measured in presence of oxygen is EC.[19] Other methods may also be used for measuring EC and OC and their results show important variations relative to the standard NIOSH method.[20, 21] In general, methods for the determination of EC and OC are time consuming and require significant sample preparation, NIOSH method included.[22, 23] There are examples of attempts to develop fast, real-time methods to estimate carbon.[24-27] They rely on optical methods and usually differentiate the organic fraction only by picking a wavelength in the UV and do not offer a comparison with NIOSH method. Although fast, these methods do not offer much information for chemical speciation. It is then important to investigate alternatives that provides fast and reliable carbon quantification in DPM.[22] A straightforward technology for EC and OC quantification could allow "real-time" monitoring of air pollution in different environments, including low resources settings. This would be useful as a tool to help optimize processes and working conditions and minimize workers exposure to carcinogenic compounds.

Raman spectroscopy is a non-destructive tool for chemical analysis. The Raman spectrum is a particular arrangement of bands that provides a molecular fingerprint for the chemical species present in a sample.[28-33] The Raman technique has been broadly used for investigating carbonaceous materials.[30, 31, 34, 35] The Raman spectra of carbon species is strong and the pattern of peaks is related to both the material structure and composition. For instance, it is easy to differentiate carbon allotropes (e.g. diamond and graphite) by their Raman spectra.[35] A pristine graphite crystal has a characteristic band at 1580 cm^{-1} (known as G-band). Defects (e.g. vacancies, bonded atoms, functional groups), disruption or disorder in

the graphite structure produce a new Raman band at 1350 cm^{-1}, which is called D-band and can be ascribed to benzene-like forms of carbon.[35] The variations in the intensities, widths and positions of the D- and G-bands can be used to derive unique signatures for different forms of carbon.[31, 36]

There are several examples in the literature related to the use of Raman spectroscopic for the analysis of various types of aerosols and different environmentally-relevant chemical compounds.[37-41] There are also several studies about the reactivity of soot probed by Raman parameters.[17, 42, 43] Ivleva *et al.* have proposed the use of the width of D-band as a parameter to estimate EC percentage in a sample.[40] Their method employed a complex fitting routine of the whole Raman spectrum using specific mathematical models for each band. The work involved the use of only 5 reference samples, including 3 soot samples from different sources, 1 sample of humic acid and graphite. This method showed an almost linear negative correlation of the D-band width and the apparent ratio of elemental carbon. Despite the apparent success of this simple correlation to obtain elemental carbon composition, this protocol ignores changes in the bandwidth caused by crystallinity.[35, 44] Also, there was no evidence of the validation of their method with real samples of environmental relevance.

Zheng *et al.* recently reported real-time measurement of carbon nanotubes in workplace environment using Raman spectroscopy. [5] They developed an instrument for collection of particulate matter, acquisition of Raman spectra, followed by data analysis using principal component analysis (PCA) and partial least squares (PLS). They have shown the potential for Raman spectroscopy in quantifying nanotubes and distinguishing them from other carbon materials.

Ess *et al*. investigated the structural changes in soot caused by heating of samples with different OC content. [17] The Raman measurements revealed an increase in the fluorescent background with increasing amount of OC, but the raw data did not show a linear correlation between OC content and fluorescence. High resolution electron microscopy showed that the carbon nanoparticles with lower OC were more crystalline, while the ones with higher OC were more amorphous. Raman analysis of the soot

samples showed an increasing fluorescence background with increasing OC due to the higher amount of fluorescent organic components. The Raman spectra revealed decreasing nanostructural order as OC content increased, indicated by a lower D-band to G-band intensity ratio and an additional shoulder at about 1200 cm^{-1}.

Graphen *et al.* have developed a fully automatized Raman soot spectrometer for quantitative analysis of EC and compared with an optical method (aethalometer).[1] Hourly measurements were obtained through several days in a traffic island in Germany and the values measured by Raman were in agreement with the Black Carbon measured by aethalometer. However, their method requires a relatively high laser power, long exposure time and only estimated EC content in comparison to an optical technique (not a standard method).

In this work, Raman spectroscopy and machine learning (PLS) were used to quantify the EC, OC and total carbon (TC) content in samples obtained from respirators of workers at three different gold mining operations. Raman spectra were collected in different regions of a filter (air sample), then baseline corrected and normalized before PLS was performed. Whereas many previous studies simply removed the fluorescent background from the Raman spectrum, this study instead fitted a horizontal baseline to quantify and verify if that correlates with the OC content.[17] Each set of Raman measurements were obtained directly from the filters, without previous sample preparation, and took less than 3 minutes. The Raman results were compared to the NIOSH Method 5040 which is the standard reference method in North America for the quantification of carbon species. Our results demonstrate that the Raman method has the potential to become a widely used tool for *in loco* analysis of workplace air quality.

3.3 Experimental section

3.3.1 Samples

46 air samples (8.04 cm^2) of DPM were collected on salinized 37 mm quartz fiber filters loaded into 3 piece cassettes. Sampling cassettes were placed near the breathing zone of the worker to measure exposures to DPM contaminated air. Personal sampling pumps were used to draw air through the filters. Samples were collected from individual workers for work-shift durations of 8-12 hours. Workers wore the samplers for an entire shift resulting in volumes of 1 to 1.5 m^3 per air sample. Upon retrieval, the cassettes were capped, sealed and returned to the laboratory for analysis. Standard occupational hygiene sampling was performed in accordance with the specifications of NIOSH Method 5040.[45]

3.3.2 Raman spectroscopy

The Raman spectra of the samples were measured with a Renishaw inVia Raman Microscope with 50x (NA=0.75) objective lens and backscattering collection. The acquisition time was 5 seconds for each of the 3 accumulations measured using lasers with 532 or 785 nm wavelengths. The power of the 532 and 785 nm laser at the sample was 1.75 mW and 1.25 mW, respectively. In both cases, the Raman spectrum was centered in 1500 cm^{-1}, region of the G-and D-bands with 10 Raman spectra collected from the central region of the filter.

3.3.3 Data Processing

For partial least squares analysis using the data collected from the 532 nm, the Raman shift interval was 1050-1800 cm^{-1} (668 points). For 785 nm, the spectral range was 939 – 1995 cm^{-1}, which is equivalent to 1021 points.

Each Raman spectrum was smoothed using a Savitzky-Golay filter. Later, the Raman spectra were baselined using a horizontal line at the minimum of the spectral range. Such choice was made based on a previous study about the Raman of carbon materials.[46] In that case, the fluorescent background was used to estimate the hydrogen content in the sample.[46] Hydrogen content has been shown to correlate well with the organic fraction in carbon materials.[17] Thus, using a horizontal line preserves some information about the fluorescence background in comparison to a simple linear baseline correction. Finally, the spectrum was normalized by the sum of the intensities. A detailed step-by-step account of the spectral manipulation is presented as supplementary information (SI). In this study, ten different PLS models were built to determine the percentage of EC (%EC), mass of EC (mEC), mass of OC (mOC) and mass of TC (mTC) with five models for each laser excitation, 532 nm and 785 nm. Baseline correction, smoothing, normalization and PLS analysis were performed using MATLAB.[47]

PLS is a statistical method for dimensionality reduction and for the determination of a linear regression model. A PLS model creates new latent variables (LV) from the original data, e.g. Raman spectra. These new variables are chosen to maximize the covariance with the target, e.g. the percentage of elemental carbon, %EC. Mathematically, PLS can be described by the equations:

$$X = TP^T + E$$

$$Y = UQ^T + F$$

Where, X is a $n \times m$ matrix (Raman spectra), Y is a $n \times p$ matrix (%EC values); T and U are $n \times l$ score matrices; P and Q are $m \times l$ loading matrices; and E and F are error terms. The score matrices (T and U) are obtained to maximize their covariance, while the P and Q helps us to interpret the model. Further mathematical details can be found in the literature.[48-50]

Variable importance in the projection (VIP) describe the contribution of each variable for the PLS model. [51] The literature suggests the VIP values greater than 1.0 are important predictors for the model.

[52] This can be helpful for interpreting how the model uses the Raman spectra for making predictions. All the plots of the VIP values presented here are in the SI. [52]

In some cases, we plotted the median value in order to reduce the influence of an outlier, making the visualization clearer. For instance, we have predicted 10 %EC values for one sample. The median was obtained by summing the 5th and 6th highest values and then dividing it by two, as in the formula below:

$$median = \frac{mEC_{5th} + mEC_{6th}}{2}$$

All 46 air samples were randomly separated into two equal groups, i.e. training and testing set. The optimum number of the latent variables (LV) was chosen according to the lowest value for estimated mean squared prediction error (EMSPE). Five models were built using Raman spectra obtained for each laser line, 532 nm and 785 nm. The number of components for the models (532 nm) predicting %EC, %EC (for samples with %EC >50 %), mEC, mOC and mTC is 5, 3, 2, 2 and 4, respectively. The number of components for the models (785 nm) predicting %EC, %EC (for samples with %EC >50 %), mEC, mOC and mTC is 3, 4, 3, 3 and 3, respectively. Overfitting was monitored by comparing the error in the testing set with the value obtained from the cross validation.

3.3.4 Quantification of Elemental and Organic Carbon by NIOSH Reference Method 5040

The quartz filters were analyzed by gas chromatography employing thermal-optical analysis for carbon content. The analysis was performed at a commercial ISO 17025 certified laboratory. A small rectangular segment was removed from each filter for analysis. The standard punch has an area of 1.5 cm^2. The amount of organic carbon and elemental carbon were determined, as described previously.[45] The amount of OC quantified in the whole filter ranged from 4.7 to 180 µg/sample. EC content determination ranged from ≤2 to 607 µg/sample. The total carbon content, as the sum of both OC and EC, was in the range of 9 to 790 µg/filter.

3.4 Results and Discussion

3.4.1 Raman spectra of samples with varying percentage of elemental Carbon

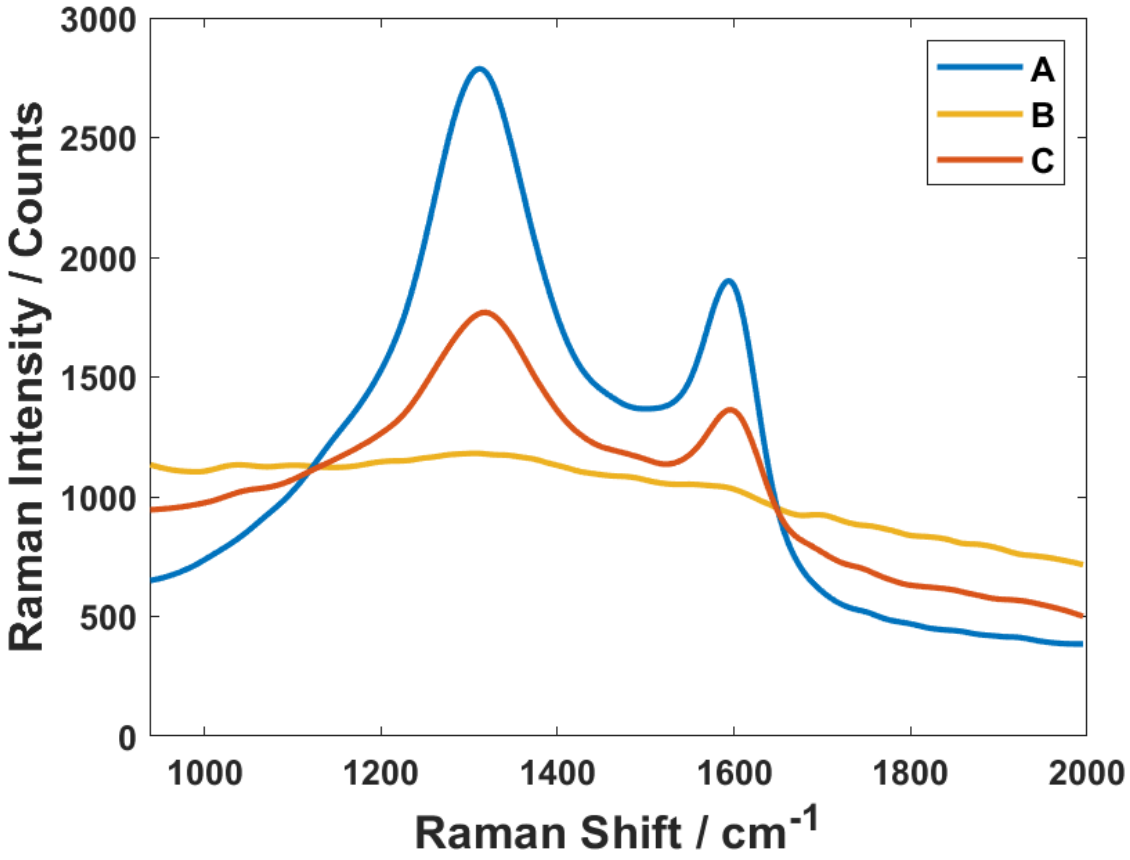

Figure 3-1: Raman spectra from filters with different %EC as determined by the NIOSH Reference Method 5040. The Raman spectra were obtained under 785 nm excitation, 5 s acquisition time and 3 accumulation. A) 79 % EC; B) 3 % EC and C) 51 % EC.

The Raman spectra of three different filters are shown in Figure 3-1. The Raman spectrum (Figure 3-1A) was obtained from the sample with highest (79 %) content of EC as determined by the NIOSH reference method. The features at 1312 cm^{-1} (D-band) and 1596 cm^{-1} (G-band) are characteristic of carbon materials with predominance of sp^2 structures. The spectrum from the sample with highest EC content (Figure 3-1A) presents the least amount of fluorescent background, as expected.[46] The Raman spectrum in Figure 3-1B represents a sample with lowest (3 %) EC content. Figure 3-1B is dominated by a strong

fluorescent background that practically overwhelms the carbon vibrational G- and D-bands. This type of strong fluorescence background in carbon materials has been correlated to hydrogenated-carbon species, described as polymeric.[46] According to Casiraghi *et al.*, it is even possible to estimate the hydrogen content in carbon materials by quantifying the magnitude of this background using green laser (514.5nm) excitation.[46] Humic acid-like material is one example of OC, usually represented as large aromatic molecules rich in hydrogen and oxygen atoms. Therefore, the observation of a fluorescent background in samples with high OC content, as in Figure 3-1B, is not surprising. Figure 3-1C is from a sample with intermediary (51 %) EC content and it presents clearly visible D- and G-Raman bands together with a significant fluorescent background. Overall, Figure 3-1 confirms that the intensity of the D- and the G-bands seems to grow with the EC content, while the fluorescent background increases with the OC content.

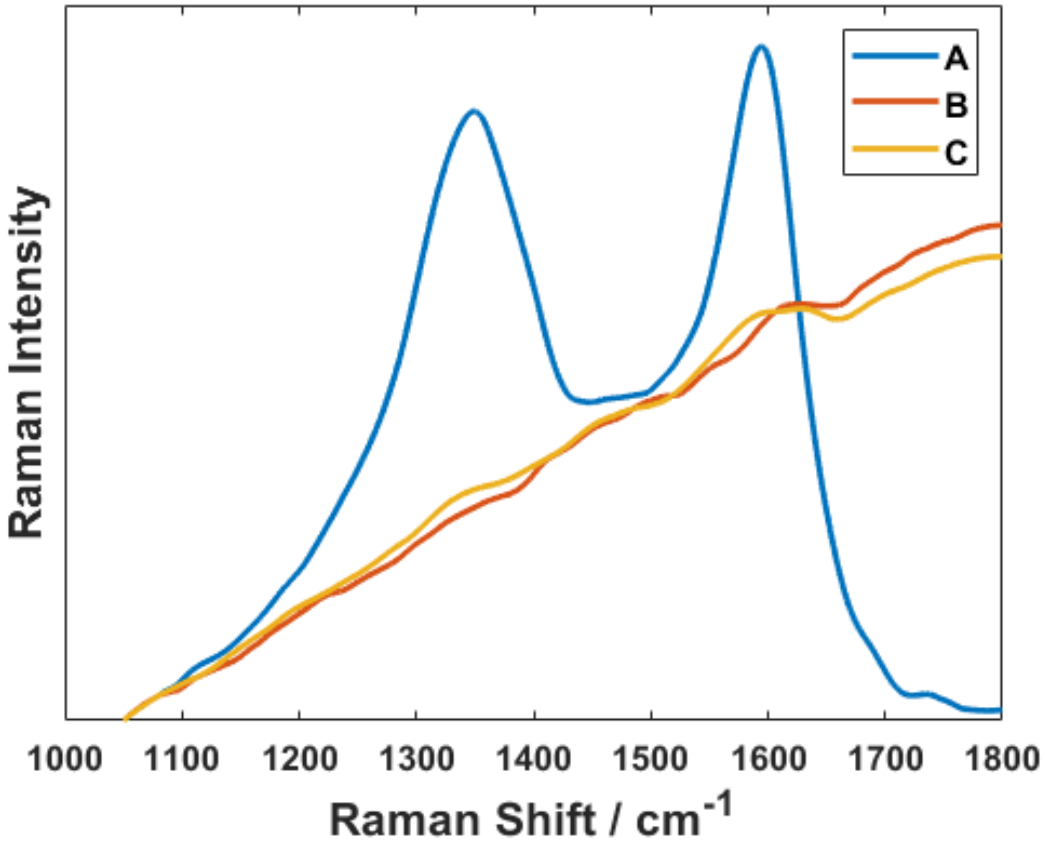

Figure 3-2: Raman spectra from filters with different %EC, obtained under 532 nm excitation. The Raman spectra A, B and C are assigned to filters containing 79 %, 3 % and 51 % of EC, respectively.

Raman spectra for the same samples obtained under 532 nm laser excitation are presented in Figure 3-2. Similarly to Figure 3-1, the 79 % EC sample showed 2 dominant bands (D- and G-), while the Raman spectrum for the sample with the lowest EC content (3 %) was dominated by the fluorescent background under green excitation (532 nm). On the other hand, the 532 nm Raman spectrum of the sample with intermediary EC content (51 %) presented very weak D- and G-bands and had a similar fluorescence background as the 3 % EC sample. The remarkable difference between the Raman spectra obtained under different laser excitations (785 nm in Figure 3-1 and 532 nm in Figure 3-2) is exemplified by the 51 % EC sample. Under 785 nm excitation, the typical carbon D- and G- bands are clearly distinguishable for the 51 % EC sample, while under 532nm they are very weak.

3.4.2 Estimating percentage of elemental Carbon using 532 nm excitation

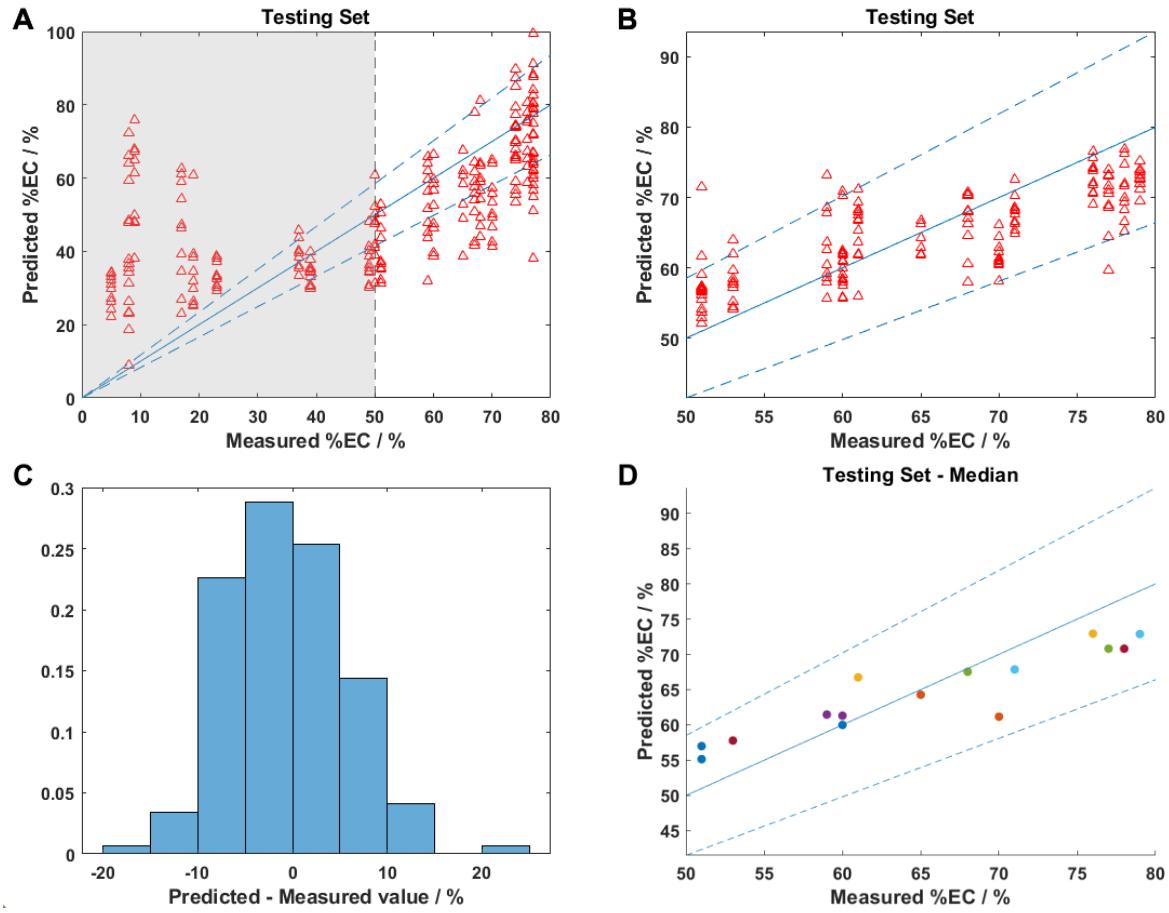

Figure 3-3: A) Scatter plot of the predicted versus measured %EC values. B) New model predicting values in the range between 50-80 %. Dashed lines show the limits of accuracy of the standard technique. C) Histogram of the difference between predicted and the measured values for samples between 50 and 80 %. D) Median of predicted values plotted against the measured values for each sample. All the models were obtained using Raman obtained with 532 nm laser excitation and PLS.

Figure 3-3A compares the predicted (PLS model) values of %EC and the %EC determined by the NIOSH reference method. These predictions were made based on Raman data collected under excitation of 532 nm. The %EC content for these samples ranged from 0 to 80 % according to the NIOSH determination. The values in the gray area (left side of the dashed line at 50 %) in Figure 3-3A show larger deviations. A linear response between the Raman data and the NIOSH values was observed for samples

with %EC higher than 50 % (Figure 3-3A), but not in the range between 0 and 50 % EC. This is related to the high fluorescent background of the organic fraction in the sample, particularly when 532 nm is used as excitation. The fluorescent background can superimpose the Raman bands due to higher cross section and may lead to poor prediction of the model. A new predictive model considering only the samples with 50 to 80 % EC determined by the NIOSH method is presented in Figure 3-3B. Practically all the %EC values predicted by Raman fit in the interval of accuracy for the standard NIOSH method. The histogram in Figure 3-3C shows the absolute error (difference between predicted and measured) for the testing set. More than 65 % of the values predicted have an absolute error of 5 % or smaller. In Figure 3-3D the median predicted values of each sample is plotted, which confirms the linear response for this model compared to the standard technique.

The supporting information (Appendix B) contains plots of the estimated mean squared prediction error (EMSPE) for both models (Figure B-2B and Figure B-3B). When all compositions are taken into consideration (Figure B-2B), the EMSPE is around 300, while for samples with 50 % or more EC (Figure B-3B) the EMSPE value is around 25. The difference in accuracy from the different EC intervals is caused by the strong fluorescence background that originates from the organic fraction in carbon materials. The EMSPE values are in agreement with the absolute error observed in the histograms (Figure B-2E and Figure B-3E).

3.4.3 Estimating percentage of elemental Carbon using 785 nm excitation.

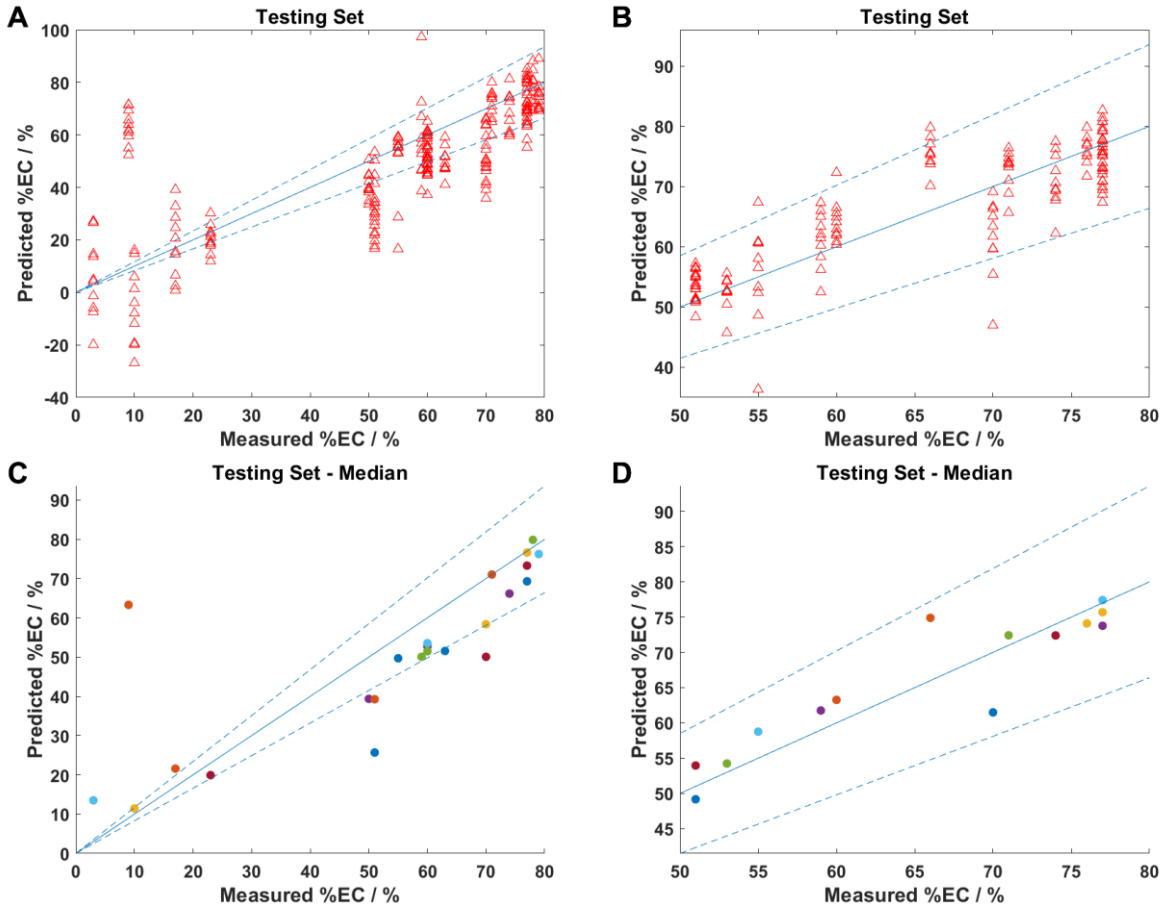

Figure 3-4: A) Scatter plot of the predicted versus measured %EC values. B) New model predicting values in the range between 50-80 %. Dashed lines show the limits of accuracy of the standard technique. C) Median of predicted values plotted against the measured values for each sample. D) Median of predicted values plotted against the measured values for each sample in the range of 50-80 %. All the models were obtained using Raman (785 nm) and PLS.

Figure 3-4A shows the correlation between the %EC predicted from Raman data using 785 nm excitation and the values determined by the NIOSH standard method. The measured values stacked in vertical correspond to the 10 Raman spectra obtained from the same sample. Thus, they have the same measured value. A better agreement between the %EC predicted by Raman and the NIOSH method is observed in Figure 3-4A (using 785 nm excitation) when compared to the results obtained with 532 nm (Figure B- 8). This is related to the smaller intensity of the OC fluorescent background at longer excitation

wavelength, leading to a better agreement between the values predicted by Raman within the accuracy limits of the standard technique in the EC range between 50 and 80 % (Figure 3-4B). Figure 3-4B shows that the Raman predicted values in the EC range between 0-50 % are broadly distributed with higher absolute errors and even negative values. Figure 3-4C presents the median of the %EC predicted by Raman plotted against the values measured by the NIOSH method for each of the 23 testing samples. The medians of 8 samples did not fit within the limits of accuracy of the NIOSH standard technique in Figure 3-4C. The model using samples of all compositions seems to overestimate the values of %EC for samples under 50 % EC, while it underestimates the predicted values for samples above 50 % EC. A new model using only samples with %EC values between 50-80% was implemented and it is shown in Figure 3-4B. Most of the values (93 %) predicted by Raman now fit inside the limits of accuracy when only the specific region of EC composition between 50-80 % is considered. Figure 3-4D shows the median of the Raman predicted values plotted against the EC composition determined by NIOSH. In this case, all values lie inside the limits of accuracy.

3.4.4 Estimating mass of elemental Carbon using 785 nm excitation

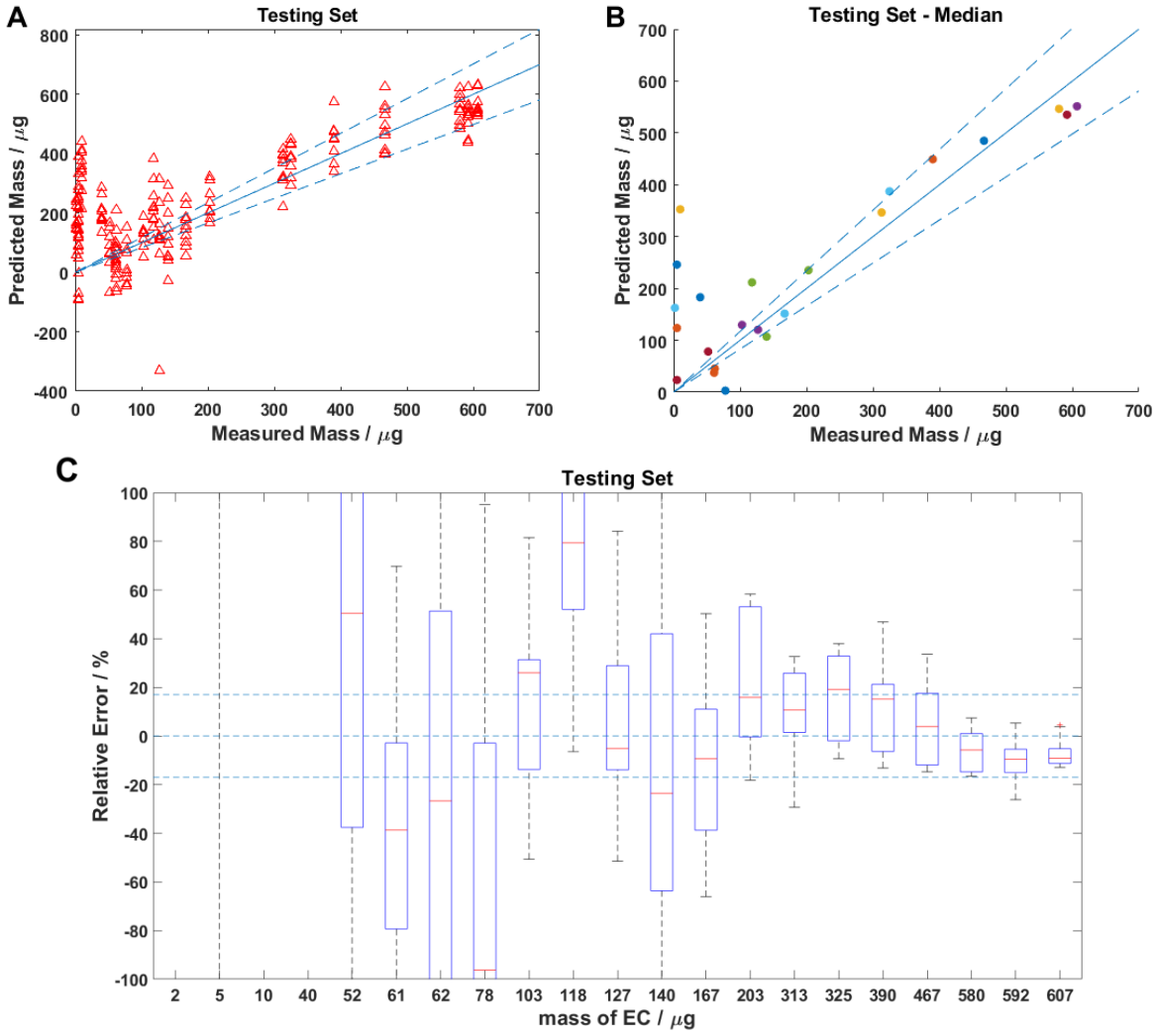

Figure 3-5: A) Plot of the measured mEC against the predicted values. Dashed lines show the limits of accuracy of the standard technique. B) Median of predicted values plotted against the measured values for each sample. Dashed lines show the limits of accuracy of the standard technique. C) Boxplot showing the variation of relative error for each sample. Horizontal dashed lines show the limits of accuracy of the standard technique. The upper and bottom limit of each box usually refer to the 75th and 25th percentile. The whiskers shows the 5 % (lower limit) and 95 % (upper limit) confidence level. The red line shows the median value. The model was created using Raman spectra (785 nm) and PLS regression.

The previous results showed the possibility and limitations of distinguishing between the carbon composition from air samples using Raman spectroscopy. However, some legislations define limits in terms of mass of different types of carbons on a filter collected from a defined volume of air (1 m^3).[53, 54]

Figure 3-5 shows the result of a PLS model developed to quantify the mass of elemental carbon (mEC) presented on filters from the Raman data obtained using 785nm excitation. The mEC values predicted by Raman spectroscopy are plotted against the mEC values measured by the NIOSH method in Figure 3-5A. A much higher variation in the correlation between the Raman and the NIOSH method is found in Figure 3-5A for mEC smaller than 200 µg. Figure 3-5B shows that the median values agree well with the NIOSH method, especially for those higher 200 µg. On the other hand, the Raman prediction overestimates the mEC content when mEC values were smaller than 200 µg. Ten (10) Raman spectra were obtained for each sample in different regions randomly selected. The relative small amount of DPM in the filter was not enough to cover the whole filter which led to overestimated predictions. It is important to emphasize that the laser spot is only about 2 µm diameter. Therefore, the Raman method probes very small areas compared to the standard NIOSH method, which uses all material from the 1.5 cm^2 region of the filter (sample). The 10 Raman spectra presented variations across the sample. To identify these variations, we produced boxplots[55] of mEC predicted by Raman correlated to the NIOSH standard data in Figure 3-5C.[55] The median values for each sample is represented in the boxplot as horizontal red lines. The dashed horizontal lines delineate a limit of ± 17 % tolerance, which is commonly quoted as the typical range of uncertainty for the NIOSH method.[45] The samples with mEC smaller than 50 µg have the largest relative errors and some of them exceed the scale in Figure 3-5C. For mEC values larger than 127 µg, the median and large part of the variation fit the limits of accuracy of the standard technique. In the range between 127 and 607 µg only two of the 11 samples have median values outside of the accuracy range of ± 17 %.

In addition, the limit of quantification for the NIOSH standard technique is 2 µg. Thus, anything smaller than that limit would be assigned as 2 µg and that could also contribute to the large deviations observed in the plot.

3.4.5 Estimating mass of organic carbon using 785 nm excitation

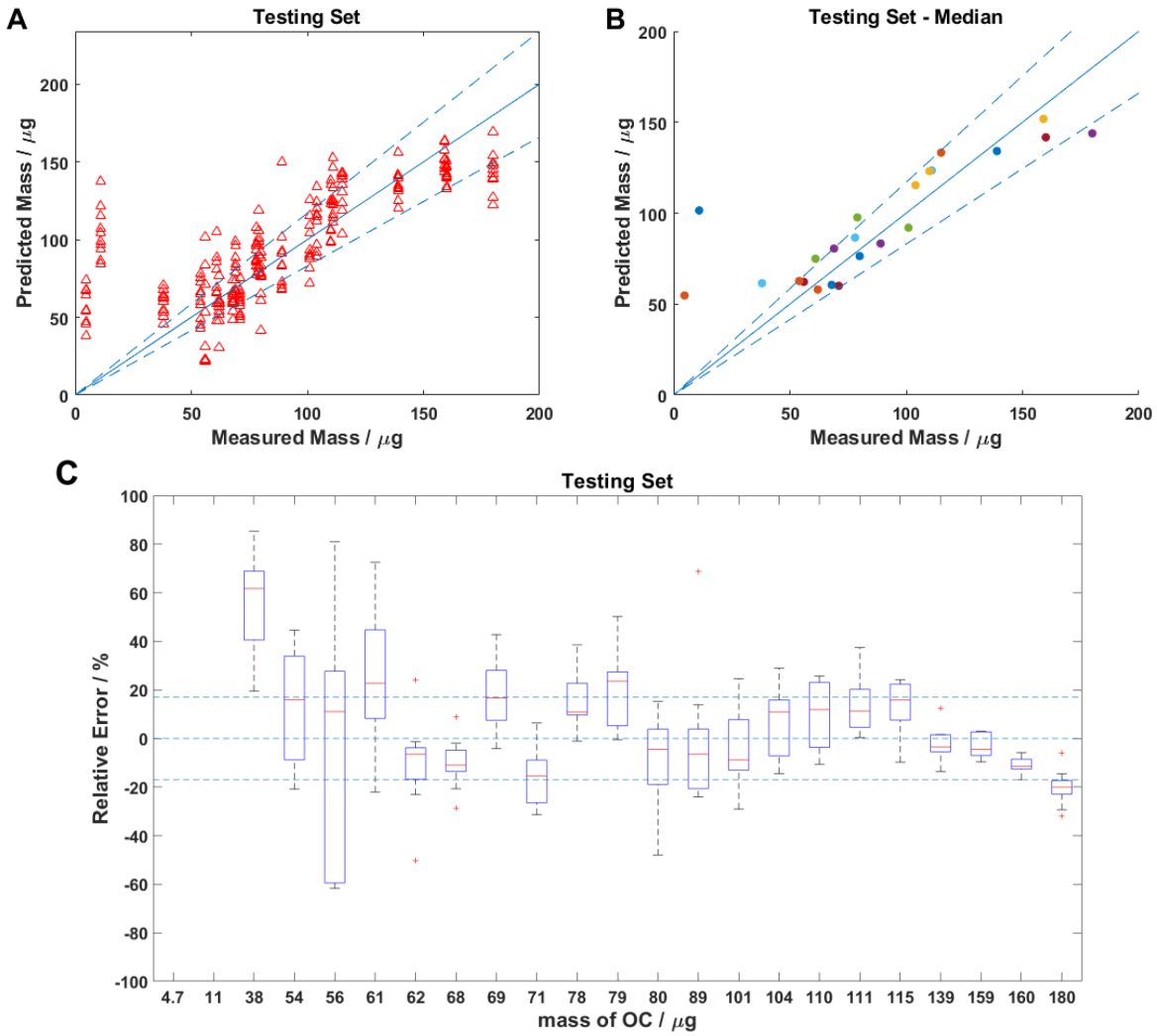

Figure 3-6: A) Plot of mOC (mass of Organic Carbon) obtained from the standard technique against the predicted values using Raman (785 nm) and PLS. Dashed lines show the limits of accuracy of the standard technique. B) Median of predicted values plotted against the measured values for each sample. Dashed lines show the limits of accuracy of the standard technique. C) Boxplot showing the variation of relative error for each sample. Horizontal dashed lines show the limits of accuracy of the standard technique.

Figure 3-6 shows the predicted values of mOC and the variation of relative error for each sample. The predicted values based on each Raman spectrum have been plotted against the value obtained in the NIOSH standard technique and can be seen on Figure 3-6A. In general, the model seems to predict reasonably well the results obtained by NIOSH, considering that most of the predicted Raman data fit inside the limits of accuracy of NIOSH in Figure 3-6A. As observed for %EC and mEC, the largest deviations occur for smaller values of mOC, below 38 μg. Figure 3-6B shows the median of the mOC values predicted by Raman plotted against the measured values by the NIOSH method. It is clear that most of the data in Figure 3-6B fits the limits of accuracy (±17 %). Only 6 samples out of 23 have the median value of the mOC predicted by Raman outside those limits. The boxplot for the variation of the relative error is shown in Figure 3-6C. It is noticeable that the values fit the accuracy limits for most samples between the interval between 54 and 180 μg. The two samples with the lowest amount of mOC presented very large deviations and they fall out of the scale in Figure 3-6C.

3.4.6 Estimating mass of total carbon using 785 nm excitation

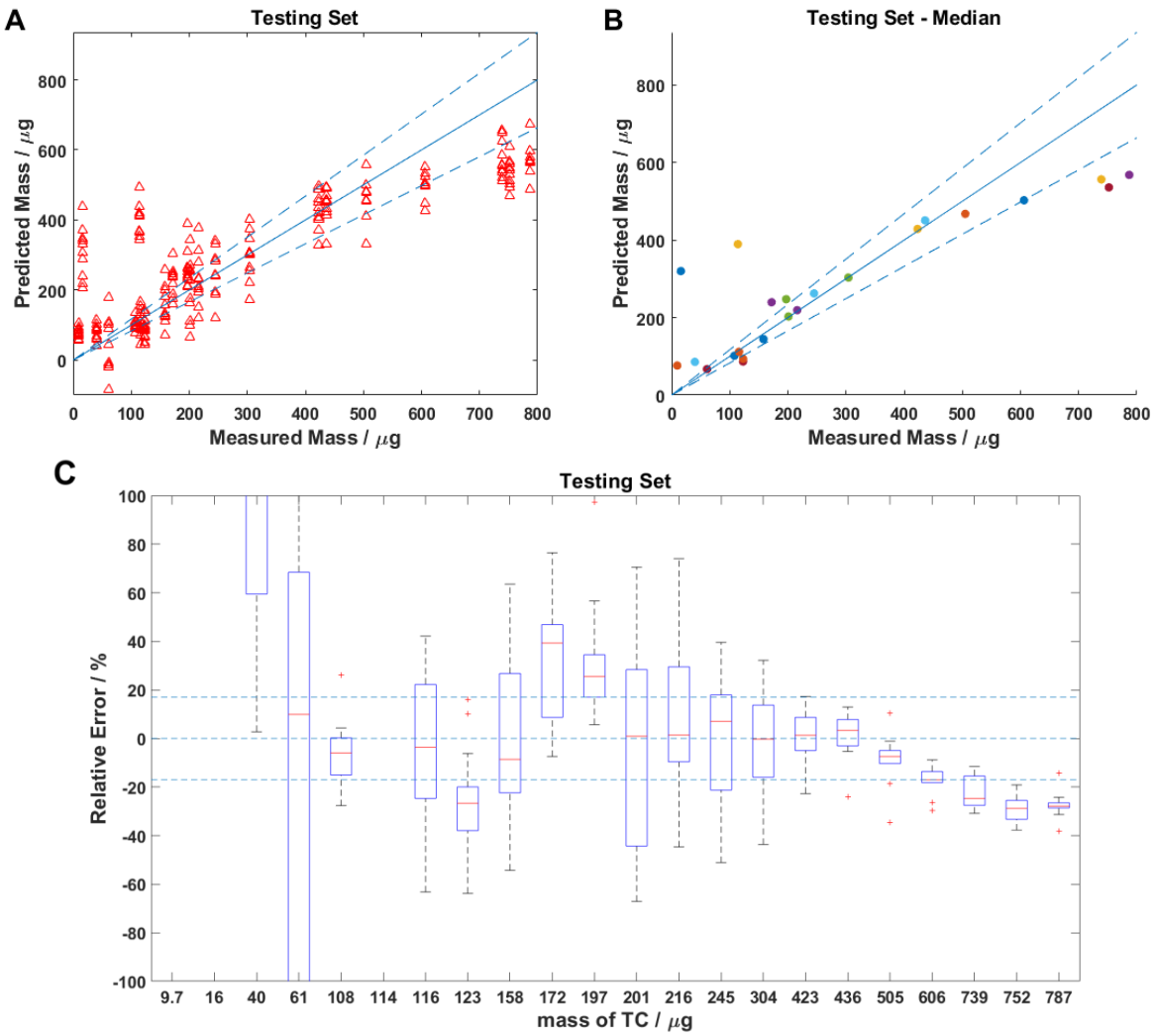

Figure 3-7: A) Plot of of mTC (mass of Total Carbon) obtained from the standard technique against the predicted values using Raman (785 nm) and PLS. Dashed lines show the limits of accuracy of the standard technique. B) Median of predicted values plotted against the measured values for each sample. Dashed lines show the limits of accuracy of the standard technique. C) Boxplot showing the variation of relative error for each sample. Horizontal dashed lines show the limits of accuracy of the standard technique.

Figure 3-7A shows the predicted values of mTC plotted against the measured values. It can be noticed that the model is underestimating the values for samples with mass equal to 600 µg or higher, with the predicted values being smaller than the measured ones. For lower mass, they seem to be more

balanced. The histogram presented in Figure B- 6F shows the difference between the predicted value and the measured is no more than 0.1 µg for 68 % of the measurements. In Figure 3-7B, only the median for each sample is plotted and 61 % of the median values are inside the limits of accuracy of the standard technique. The medians reveal the model is underestimating the mass of mTC for the samples with 739, 752 and 767 µg, while overestimating lower values of mass, from 9.7 to 40 µg. The optimal interval of the model seems to be between 201 and 606 µg, with all the medians fitting the limits of accuracy. It is noteworthy that the samples with mTC of 114 and 116 µg present very different relative error, despite having similar mass. This might be related to the composition of the samples. The lighter sample contains 91 % of OC, while the other sample has 47 % of OC. This large difference in OC can change the Raman spectrum significantly, leading to stronger fluorescent background and degrade prediction.

VIP score can be a helpful tool for interpreting the PLS models. The literature indicates that values above 1.0 can be considered important features for the prediction.[52] The VIP score for the different analytes are presented in the Supporting Information (Appendix B). All exhibit G and D bands as the most important variables. Background has a higher importance for predicting mOC, which agrees with the high level of fluorescence observed in such samples. The mTC seems to be a mix of profiles for mOC and mEC, with high importance to G and D bands and smaller importance to background. This observed behavior makes sense, considering the larger contribution usually comes from the EC. Even the range of masses is very different between the two analytes, from 2 to 600 µg and 4.7 to 180 µg for EC and OC, respectively.

3.5 Conclusion

In conclusion, this work has shown that models using Raman and PLS are able to predict values of quantities that are relevant for the legislation of levels of EC, OC and TC in the workplaces. The method presented here agrees with the reference NIOSH method for certain ranges, particularly for large amounts of carbon compounds. The fluorescent background played a relevant role in the models for OC and TC.

The large variation for estimating EC and TC was observed in samples with smaller amounts of carbon. Such inconvenience could be helped by collecting a larger number of Raman spectra from a particular sample, using, for instance the Raman mapping technique , or using a non-linear model.[56] Also, the TC model has shown some limitations for samples with high OC content. This could mean the model needs more information (samples/spectra) for better training, or there is some non-linearity between the spectra and the sample composition. Finally, we should keep in mind that the NIOSH 5040 has its limitations, such as just providing an average value for a shift of 8 hours and low accuracy. The implementation of time-resolved Raman method for estimating Elemental Carbon has been reported in the literature.[1, 45] Since the definition of these carbon species is operational, rather than directly correlated to absolute chemical compositions, it is common to observe that different methods provide distinct values for those carbon quantity parameters.[20] In contrast to all other methods, the Raman technology, allied to partial least squares regression, could be implemented directly in workplaces using portable, low cost, Raman spectrometers. This would allow for rapid quantification of DPM carbon emissions which should be an useful tool for the optimization of processes and improvement of environmental and working conditions.

3.6 References

(1) Grafen, M.; Schweiger, G.; Esen, C.; Ostendorf, A. Time-resolved measurement of elemental carbon in urban environment: Comparison of Raman backscattering and aethalometer results. *Journal of Aerosol Science* **2018**, *117*, 34-43. DOI: 10.1016/j.jaerosci.2017.12.003.

(2) Attia, A. M. A.; Kulchitskiy, A. R.; Nour, M.; El-Seesy, A. I.; Nada, S. A. The influence of castor biodiesel blending ratio on engine performance including the determined diesel particulate matters composition. *Energy* **2022**, *239*, 121951. DOI: 10.1016/j.energy.2021.121951.

(3) Liang, X. Y.; Zhao, B. W.; Wang, K.; Lv, X.; Wang, Y. J.; Liu, J.; Wang, Y. S. Impact of early injection on physicochemical characteristics of diesel soot particles. *Fuel* **2021**, *292*, 120262. DOI: 10.1016/j.fuel.2021.120262.

(4) Farahani, V. J.; Pirhadi, M.; Sioutas, C. Are standardized diesel exhaust particles (DEP) representative of ambient particles in air pollution toxicological studies? *Science of the Total Environment* **2021**, *788*, 147854. DOI: 10.1016/j.scitotenv.2021.147854.

(5) Zheng, L. N.; Kulkarni, P. Real-Time Measurement of Airborne Carbon Nanotubes in Workplace Atmospheres. *Analytical Chemistry* **2019**, *91* (20), 12713-12723, Article. DOI: 10.1021/acs.analchem.9b02178.
(6) Organization, W. H. *Air Pollution*. World Health Organization, https://www.who.int/health-topics/air-pollution#tab=tab_2 (accessed 2022 January 25th).
(7) World Health, A. Health, environment and climate change: road map for an enhanced global response to the adverse health effects of air pollution: report by the Director-General. World Health Organization: Geneva, 2018.
(8) Apte, J. S.; Brauer, M.; Cohen, A. J.; Ezzati, M.; Pope, C. A. Ambient PM2.5 Reduces Global and Regional Life Expectancy. *Environmental Science & Technology Letters* **2018**, *5* (9), 546-551. DOI: 10.1021/acs.estlett.8b00360.
(9) Delfino, R. J.; Staimer, N.; Tjoa, T.; Gillen, D. L.; Polidori, A.; Arhami, M.; Kleinman, M. T.; Vaziri, N. D.; Longhurst, J.; Sioutas, C. Air Pollution Exposures and Circulating Biomarkers of Effect in a Susceptible Population: Clues to Potential Causal Component Mixtures and Mechanisms. *Environmental Health Perspectives* **2009**, *117* (8), 1232-1238. DOI: 10.1289/ehp.0800194.
(10) Brook, R. D.; Rajagopalan, S.; Pope, C. A.; Brook, J. R.; Bhatnagar, A.; Diez-Roux, A. V.; Holguin, F.; Hong, Y. L.; Luepker, R. V.; Mittleman, M. A.; et al. Particulate Matter Air Pollution and Cardiovascular Disease An Update to the Scientific Statement From the American Heart Association. *Circulation* **2010**, *121* (21), 2331-2378. DOI: 10.1161/CIR.0b013e3181dbece1.
(11) Agency, U. S. E. P. *Health and Environmental Effects of Particulate Matter (PM)*. United States Environmental Protection Agency, https://www.epa.gov/pm-pollution/health-and-environmental-effects-particulate-matter-pm (accessed 2022 January 25th).
(12) Assessment, C. O. o. E. H. H. *Diesel Particulate Matter*. https://oehha.ca.gov/calenviroscreen/indicator/diesel-particulate-matter (accessed Jan 25th 2023).
(13) (IARC), I. A. f. R. i. C. *Outdoor air pollution*; International Agency for Research in Cancer (IARC), 2013.
(14) (IARC), I. A. f. R. o. C. *Diesel and Gasoline Engine Exhausts and Some Nitroarenes*; International Agency for Research on Cancer (IARC), 2012.
(15) Wichmann, H. E. Diesel exhaust particles. *Inhalation Toxicology* **2007**, *19*, 241-244. DOI: 10.1080/08958370701498075.
(16) Mertes, S.; Dippel, B.; Schwarzenbock, A. Quantification of graphitic carbon in atmospheric aerosol particles by Raman spectroscopy and first application for the determination of mass absorption efficiencies. *Journal of Aerosol Science* **2004**, *35* (3), 347-361. DOI: 10.1016/j.jaerosci.2003.10.002.
(17) Ess, M. N.; Ferry, D.; Kireeva, E. D.; Niessner, R.; Ouf, F. X.; Ivleva, N. P. In situ Raman microspectroscopic analysis of soot samples with different organic carbon content: Structural changes during heating. *Carbon* **2016**, *105*, 572-585. DOI: 10.1016/j.carbon.2016.04.056.
(18) Chow, J. C.; Watson, J. G.; Crow, D.; Lowenthal, D. H.; Merrifield, T. Comparison of IMPROVE and NIOSH carbon measurements. *Aerosol Science and Technology* **2001**, *34* (1), 23-34. DOI: 10.1080/027868201300081923.
(19) Birch, M. E. Chapter DL: Monitoring Diesel Exhaust in the Workplace. In *NIOSH manual of analytical methods.*, 5th ed.; Andrews, E., O'Connor, P. F. Eds.; US Department of Health And Human Services, Centers for Disease Control and Prevention, National Institute for Occupational Safety and Health, 2016.
(20) Schmid, H.; Laskus, L.; Abraham, H. J.; Baltensperger, U.; Lavanchy, V.; Bizjak, M.; Burba, P.; Cachier, H.; Crow, D.; Chow, J.; et al. Results of the "carbon conference" international aerosol carbon round robin test stage I. *Atmospheric Environment* **2001**, *35* (12), 2111-2121. DOI: 10.1016/s1352-2310(00)00493-3.

(21) Huebert, B. J.; Charlson, R. J. Uncertainties in data on organic aerosols. *Tellus Series B-Chemical and Physical Meteorology* **2000**, *52* (5), 1249-1255. DOI: 10.1034/j.1600-0889.2000.01146.x.

(22) Parks, D. A.; Griffiths, P. R.; Weakley, A. T.; Miller, A. L. Quantifying elemental and organic carbon in diesel particulate matter by mid-infrared spectrometry. *Aerosol Science and Technology* **2021**, *55* (9), 1014-1027. DOI: 10.1080/02786826.2021.1917764.

(23) Birch, M. E. Chapter DL: Monitoring Diesel Exhaust in the Workplace. In *NIOSH manual of analytical methods.*, 5th ed.; Andrews, E., O'Connor, P. F. Eds.; US Department of Health And Human Services, Centers for Disease Control and Prevention, National Institute for Occupational Safety and Health, 2016.

(24) Drinovec, L.; Mocnik, G.; Zotter, P.; Prevot, A. S. H.; Ruckstuhl, C.; Coz, E.; Rupakheti, M.; Sciare, J.; Muller, T.; Wiedensohler, A.; et al. The "dual-spot" Aethalometer: an improved measurement of aerosol black carbon with real-time loading compensation. *Atmospheric Measurement Techniques* **2015**, *8* (5), 1965-1979. DOI: 10.5194/amt-8-1965-2015.

(25) Khan, M. U.; Homan, K. O.; Saki, S. A.; Emad, M. Z.; Raza, M. A. Real-time diesel particulate matter monitoring in underground mines: evolution and applications. *International Journal of Mining Reclamation and Environment* **2021**, *35* (4), 291-305. DOI: 10.1080/17480930.2020.1818937.

(26) Miller, A. L.; Habjan, M. C.; Park, K. Real-time estimation of elemental carbon emitted from a diesel engine. *Environmental Science & Technology* **2007**, *41* (16), 5783-5788. DOI: 10.1021/es070150a.

(27) Northrop, W. F.; Zarling, D.; Li, X. S.; Asme. Considerations in using photometer instruments for measuring total particulate matter mass concentration in diesel engine exhaust. *Proceedings of the Asme Internal Combustion Engine Fall Technical Conference, 2017, Vol 2* **2017**.

(28) Rosch, P.; Harz, M.; Schmitt, M.; Popp, J. Raman spectroscopic identification of single yeast cells. *Journal of Raman Spectroscopy* **2005**, *36* (5), 377-379. DOI: 10.1002/jrs.1312.

(29) Stone, N.; Kendall, C.; Smith, J.; Crow, P.; Barr, H. Raman spectroscopy for identification of epithelial cancers. *Faraday Discussions* **2004**, *126*, 141-157. DOI: 10.1039/b304992b.

(30) Ferrari, A. C.; Robertson, J. Resonant Raman spectroscopy of disordered, amorphous, and diamondlike carbon. *Physical Review B* **2001**, *64* (7), 075414. DOI: 10.1103/PhysRevB.64.075414.

(31) Ferrari, A. C.; Robertson, J. Interpretation of Raman spectra of disordered and amorphous carbon. *Physical Review B* **2000**, *61* (20), 14095-14107. DOI: 10.1103/PhysRevB.61.14095.

(32) Mazza, T.; Barborini, E.; Piseri, P.; Milani, P.; Cattaneo, D.; Li Bassi, A.; Bottani, C. E.; Ducati, C. Raman spectroscopy characterization of TiO2 rutile nanocrystals. *Physical Review B* **2007**, *75* (4), 045416, Article. DOI: 10.1103/PhysRevB.75.045416.

(33) Mukherjee, R.; Verma, T.; Nandi, D.; Umapathy, S. Understanding the effects of culture conditions in bacterial growth: A biochemical perspective using Raman microscopy. *Journal of Biophotonics* **2020**, *13* (1), e201900233. DOI: 10.1002/jbio.201900233.

(34) Casiraghi, C.; Ferrari, A. C.; Robertson, J. Raman spectroscopy of hydrogenated amorphous carbons. *Physical Review B* **2005**, *72* (8), 085401, Article. DOI: 10.1103/PhysRevB.72.085401.

(35) Ferrari, A. C.; Robertson, J. Raman spectroscopy of amorphous, nanostructured, diamond-like carbon, and nanodiamond. *Philosophical Transactions of the Royal Society a-Mathematical Physical and Engineering Sciences* **2004**, *362* (1824), 2477-2512, Article. DOI: 10.1098/rsta.2004.1452.

(36) Ferrari, A. C.; Basko, D. M. Raman spectroscopy as a versatile tool for studying the properties of graphene. *Nature Nanotechnology* **2013**, *8* (4), 235-246, Review. DOI: 10.1038/nnano.2013.46.

(37) Aggarwal, R. L.; Di Cecca, S.; Farrar, L. W.; Jeys, T. H. Chemical aerosol detection and identification using Raman scattering. *Journal of Raman Spectroscopy* **2014**, *45* (8), 677-679. DOI: 10.1002/jrs.4509.

(38) Stowers, M. A.; Friedlander, S. K. Chemical characterization of flowing polydisperse aerosols by Raman spectroscopy. *Aerosol Science and Technology* **2002**, *36* (1), 48-61. DOI: 10.1080/027868202753339078.

(39) Sinanis, S.; Aleksandrova, M.; Schaber, K. Characterization of Multicomponent Aerosols by Raman Spectroscopy. *Aerosol Science and Technology* **2011**, *45* (6), 751-757. DOI: 10.1080/02786826.2011.559494.

(40) Ivleva, N. P.; McKeon, U.; Niessner, R.; Poschl, U. Raman microspectroscopic analysis of size-resolved atmospheric aerosol particle samples collected with an ELPI: Soot, humic-like substances, and inorganic compounds. *Aerosol Science and Technology* **2007**, *41* (7), 655-671. DOI: 10.1080/02786820701376391.

(41) Sadezky, A.; Muckenhuber, H.; Grothe, H.; Niessner, R.; Poschl, U. Raman micro spectroscopy of soot and related carbonaceous materials: Spectral analysis and structural information. *Carbon* **2005**, *43* (8), 1731-1742, Article. DOI: 10.1016/j.carbon.2005.02.018.

(42) Merlen, A.; Buijnsters, J. G.; Pardanaud, C. A Guide to and Review of the Use of Multiwavelength Raman Spectroscopy for Characterizing Defective Aromatic Carbon Solids: from Graphene to Amorphous Carbons. *Coatings* **2017**, *7* (10), 153, Review. DOI: 10.3390/coatings7100153.

(43) Schmid, J.; Grob, B.; Niessner, R.; Ivleva, N. P. Multiwavelength Raman Microspectroscopy for Rapid Prediction of Soot Oxidation Reactivity. *Analytical Chemistry* **2011**, *83* (4), 1173-1179. DOI: 10.1021/ac102939w.

(44) Ferrari, A. C.; Rodil, S. E.; Robertson, J. Interpretation of infrared and Raman spectra of amorphous carbon nitrides. *Physical Review B* **2003**, *67* (15), 155306. DOI: 10.1103/PhysRevB.67.155306.

(45) and, N. I. f. O. S.; (NIOSH), H. Elemental carbon (diesel part.) 5040. In *NIOSH Manual of Analytical Methods (NMAM)*, Fourth Edition ed.; 2003.

(46) Casiraghi, C.; Piazza, F.; Ferrari, A. C.; Grambole, D.; Robertson, J. Bonding in hydrogenated diamond-like carbon by Raman spectroscopy. *Diamond and Related Materials* **2005**, *14* (3-7), 1098-1102, Article; Proceedings Paper. DOI: 10.1016/j.diamond.2004.10.030.

(47) The Math Works, I. Matlab. Version 2021a. The Math Works, Inc.: Natick, MA, USA, 2021.

(48) Höskuldsson, A. PLS regression methods. *Journal of Chemometrics* **1988**, *2* (3), 211-228. DOI: https://doi.org/10.1002/cem.1180020306.

(49) Andries, J. P. M.; Vander Heyden, Y.; Buydens, L. M. C. Predictive-property-ranked variable reduction in partial least squares modelling with final complexity adapted models: Comparison of properties for ranking. *Analytica Chimica Acta* **2013**, *760*, 34-45. DOI: 10.1016/j.aca.2012.11.012.

(50) Rosipal, R.; Krämer, N. Overview and Recent Advances in Partial Least Squares. Berlin, Heidelberg, 2006; Springer Berlin Heidelberg: pp 34-51.

(51) Chong, I. G.; Jun, C. H. Performance of some variable selection methods when multicollinearity is present. *Chemometrics and Intelligent Laboratory Systems* **2005**, *78* (1-2), 103-112. DOI: 10.1016/j.chemolab.2004.12.011.

(52) Muratore, M. Raman spectroscopy and partial least squares analysis in discrimination of peripheral cells affected by Huntington's disease. *Analytica Chimica Acta* **2013**, *793*, 1-10. DOI: 10.1016/j.aca.2013.06.012.

(53) Brewer, K.; Egyed, M.; Huang, L.; Ling, B.; Matz, C.; Rouleau, M. Human health risk assessment for diesel exhaust. Branch., C. H. C. H. E. a. C. S., Ed.; Health Canada: Ottawa, 2016; p 38.

(54) Government, O. *Sampling for diesel particulate matter in mines*. 1990. https://www.ontario.ca/page/sampling-diesel-particulate-matter-mines (accessed 2022 July 12th, 2022).

(55) Reimann, C.; Filzmoser, P.; Garrett, R. G. Background and threshold: critical comparison of methods of determination. *Science of the Total Environment* **2005**, *346* (1-3), 1-16. DOI: 10.1016/j.scitotenv.2004.11.023.

(56) Gordon, K. C.; McGoverin, C. M. Raman mapping of pharmaceuticals. *International Journal of Pharmaceutics* **2011**, *417* (1-2), 151-162. DOI: 10.1016/j.ijpharm.2010.12.030.

4 Classification of carbon materials using Raman spectroscopy and PCA-LDA.

This chapter is a manuscript in preparation for submission. The authors and the title of the manuscripts are: B. G. daFonseca, S. S. Thind, I. Booth, and A. G. Brolo, A Machine Learning Approach for the Automated Classification of Carbon Materials.

The synthetic experiments were performed by S. S. Thind and I. Booth. The Raman measurements, data processing and modeling were performed by myself.

4.1 Abstract

Principal component analysis (PCA) and linear discriminant analysis were used to classify different types of carbon material based on their Raman spectra. The selected reference materials were highly oriented pyrolytic graphite (HOPG), diamond-like carbon (DLC), glassy carbon (GC), hydrogenated graphite like carbon (PLCH) and hydrogenated polymer-like carbon (PLCH). These materials vary in crystallinity, predominant carbon hybridization, and hydrogen content. The training dataset were Raman spectra collected from both commercial samples (HOPG, DLC, GC) and samples synthesized in our laboratory (GLCH, PLCH). The Raman spectra were collected using 532 nm laser excitation. The classification model revealed that the first principal component (PC1) was the determinant source of information to separate the crystalline from the amorphous carbon samples. PC2 allowed the separation of amorphous material with different levels of hybridization (sp^2 and sp^3). Finally, both PC2 and PC3 contributed to separate materials with different levels of hydrogenation. The classification model was tested using a library of Raman spectra of carbon materials reported in the literature and the results showed a high accuracy

prediction (97 %). The model presented here provide an avenue for automated classification of carbon materials using Raman spectroscopy and machine learning.

4.2 Introduction

Carbon is a very special element due to their electronic structure[1] that supports different hybridizations (sp, sp^2 and sp^3). This allows the formation of several allotropes and the production of different materials with remarkable but contrasting properties.[2, 3] For instance, diamond is composed by sp^3 carbon, making it extremely hard and an electrical insulator. In contrast, graphite is composed by sp^2 carbon, leading to a soft material with high conductivity.

Carbon materials are used for several different applications both in academia and in industry.[4-6] Graphite is used in alkaline batteries as an electrical conductor and as a lubricant for manufacturing tools.[7] Glassy carbon is used for fabrication of camera lenses, solid-state batteries and motor brushes.[8] Diamond-like carbon (DLC) is employed in automotive and cutting tools industries[9].

Raman spectroscopy is a technique broadly employed to identify carbon materials. Characterization by Raman requires minimum sample preparation. It is also a non-destructive and a fast technique. Raman spectroscopy can be used to distinguish different carbon allotropes, to determine the number of layers in graphene, to evaluate the different levels of amorphization, and to estimate hydrogenation in amorphous carbon.[10-12] Raman measurements with visible laser excitation is generally used in the characterization of carbon materials, since it yields strong Raman intensities due to the double resonance effect observed for sp^2 materials.[10] A laser excitation in the UV is more appropriated for sp^3 materials.

The Raman spectrum of diamond (sp^3) is characterized by a band at 1330 cm^{-1}. For crystalline graphite (sp^2), the spectrum has a strong band at 1580 cm^{-1}, which is called G-band.[13] Increasing disorder in graphite leads to the rise of a band at 1360 cm^{-1}, called D-band.[13, 14] Amorphous carbon materials will

have broader bands, and depending on the ratio of sp^3 and sp^2 types of carbon, the G-band may shift to much smaller wavenumber (e.g. 1520 cm^{-1}).[15] The presence of other elements can also cause changes in the Raman spectrum. For instance, hydrogen functionalization leads to a fluorescent background that is proportional to the hydrogenation level in the material.

In a recent review, Morelos-Gomez *et al.* describes the application of data science in several aspects related to carbon materials, including synthesis, characterization, and applications.[4] Combination of Raman and data science have been used to predict properties of carbon nanotubes from different synthetic pathways, the number of layers in graphene, and the capacitance of activated carbon.[16-23]

Campos *et al*. have obtained Raman spectra of materials with different levels of crystallinity and treated their spectral data with principal component analysis (PCA) to follow the different levels of amorphization.[24] This resemble the amorphization trajectory proposed by Ferrari *et al*.[12, 24] The trajectory was also influenced by the type of defect (line or point defect) in the material. In addition, their exploratory analysis emphasized the different Raman features that originates from different carbon materials. Finally, Campos *et al*. also showed the possibility of differentiating graphene from graphite. Although that exploratory analysis show the possibility of differentiation between carbon films by Raman spectroscopy, there was no dataset for testing the models and the figures-of-merit of the model has not been reported.

Silva *et al.* developed two protocols to estimate the number of layers in graphene produced through liquid exfoliation.[16] One protocol was based on the G-band and the other on the 2D-band. The G-band protocol used the intensity ratio between G-band and Si-band (Si was the substrate used to immobilize the graphene samples), and found an empirical relation between the intensity ratio and the number of layers of graphene measured by atomic force microscopy (AFM). Their 2D-band protocol used some neural network procedure for denoising and PCA for differentiating the samples. An attempt for

classification was performed based on a probability density function. The 2D-band method had poorer performance compared to the G-band protocol. The equation used for the G-protocol is not linear, and the method also required an internal standard (Si), while PCA is a linear method for dimensionality reduction.

Kumbhar *et al*. have prepared activated carbon from biomass, called bio-carbon, using several precursors, such as mung bean, black urad bean, black grape seed, and various temperature treatment ranging between 450-750 °C.[25] They obtained Raman spectra for all samples and were able to discriminate between the different temperatures and precursors using PCA. Partial least squares and discriminant analysis (PLS-DA) was also employed to build a classification method, but only cross-validation was performed. Finally, they have built models using neural network algorithms to classify materials by temperature, precursor or temperature and precursor at the same time. According to their validation dataset, they have had some success in the classification.

Villa-Aleman *et al.* have studied detonation soot using Raman spectroscopy. By using PCA, they were able to show differences in the Raman spectra of soot from different types of explosives. That work opened a venue for further studies that combine Raman with machine learning as a tool to predict the origin of explosive based on the Raman spectrum of soot.

Some of the spectroscopic parameters used for carbon material characterization are the G-band position , the full width at half maximum, and the intensity ratio between G- and D-bands (I(D)/I(G)). Those parameters are extracted from the Raman spectrum by fitting the vibrational bands using different functions. The type and number of functions used in the fits can have an effect on the values for the aforementioned parameters and lead to misleading analysis.[26] Baseline corrections are sometimes required to remove any fluorescence background which also affect the obtained values for such parameters. The fact that spectral pre-treatment (baseline correction and curve fitting conditions) has an

effect on the values of the main characterization parameters make an universal robust analysis very difficult. This means that some carbon materials and films are sometimes misclassified in the literature. Moreover, the industrial applications of carbon materials and films suggest that an automated technology for classification could be a very useful tool for quality control.

Here, we propose a machine learning approach that takes advantage of PCA and linear discriminant analysis (LDA) to construct a reliable system for the classification of carbon materials based on their Raman spectrum. This model is a step towards a quick automated classification toolbox for the identification and quality control of carbon materials.

We have obtained Raman spectra from several commercial and our own synthesized samples and used them to build a training dataset. The carbon material training dataset includes Raman spectra of highly oriented pyrolytic graphite (HOPG), glassy carbon (GC), diamond-like carbon (DLC), hydrogenated graphite-like carbon (GLCH) and hydrogenated polymer-like carbon (PLCH). This selection of materials provide a range of crystallinity, compositions, hydrogenation and sp^2/sp^3 content. We then constructed a classification model and subsequently tested using the spectra of carbon material collected from the literature and from other synthetic methods (testing dataset). The testing dataset comprise of Raman spectra reported to be from HOPG, GC, DLC, GLCH, and PLCH. Principal component analysis and linear discriminant analysis (PCA-LDA) have been chosen because it provides easier interpretation compared to other methods. The PCA helps visualize the separation of the data while the LDA finds the best way of making the classification. The carbon materials used here allowed a test of the machine learning capability to segregate crystalline and amorphous forms of carbon materials, distinguishing different predominant hybridization (sp^2 and sp^3), and differentiate materials according to the level of hydrogenation.

4.3 Experimental section

4.3.1 Samples

The Raman spectra from the training dataset were obtained using commercial samples (HOPG, GC and DLC) and materials synthesized in our laboratory (GLCH and PLCH). The sample of HOPG (10x10x1 mm) was bought from Alfa Aesar. The DLC sample was obtained from Crystallume PVD. The GC sample is a glassy carbon rod, 3mm diameter, from Alfa Aesar. The synthesized samples were obtained using an electrochemical method described previously.[27] Briefly, two-electrode electrochemical cell consisting of a glass vial closed and connected to a DC power supply was used. The anode was a graphite rod and the cathode a titanium plate. Different electrolytes, potential, and deposition time was explored to tune the material composition. For the GLCH training set, the 0.5 M HNO_3 was the electrolyte and the resulting hydrogen content in the carbon material was less than 35 at.%.[11] For the PLCH training set, 0.5 M HCl was the electrolyte, producing only carbon samples that contained more than 35 at.% of hydrogen.[11]

The testing dataset for HOPG[28-49], GC[28, 42, 50-72] and DLC[15, 28-30, 73-89] were obtained from the literature. A web tool (WebPlotDigitizer) was used to extract the datapoints from the Raman spectra in the interval between 1050 and 1800 cm^{-1} from the literature.[90]

Testing Raman samples for GLCH and PLCH were synthesized using electrochemical methods. The details for each synthesis can be found in the supporting information (Appendix C). In few words, a setup similar to the training samples has been used, with a graphite as anode and Ti as cathode. The potential applied varied from 5-20 V, deposition time ranged from 1 h to 17 h and several electrolytes were used. This work used different samples for the training and testing datasets to avoid overestimating the machine learning capabilities.

4.3.2 Raman measurements

All Raman measurements were obtained using a Renishaw InVia Raman Microscope with 50x objective lens (NA=0.75), backscattering collection and 532 nm excitation laser at 1.75 mW. The exposure time for each spectrum was 5 s and 3 accumulations were acquired. The spectrometer was equipped with a 2400 l/mm grating centered at 1500 cm^{-1}.

4.3.3 Data Preprocessing

Each acquired spectrum was trimmed to the spectral range between 1050 and 1800 cm^{-1}. This range is the minimum necessary to estimate the hydrogen content in the samples.[11] Then, a horizontal baseline was traced at the minimum value of the spectrum and subtracted (Figure C- 1A). The horizontal baseline was chosen because the information regarding the level of hydrogen content in the carbon materials manifest as a broad fluorescence under the Raman peaks. Last, all spectra were normalized by making the highest intensity in each spectrum equal to one (Figure C- 1B). The SI file contains details of the data processing that allows the visualization of each of these steps.

4.3.4 PCA-LDA method

PCA was performed in all datasets for dimension reduction and the first 3 PCs components accounted for > 95% variance. Then, the linear discrimination analysis (LDA) was the method chosen to create a classification model. LDA finds a dimension that best separate the training data points according to their classification. Once the model was created, it was tested using Raman spectra from the literature and samples synthesized in the lab. The model was built using scikit-learn package on python.[91]

4.4 Results and Discussion

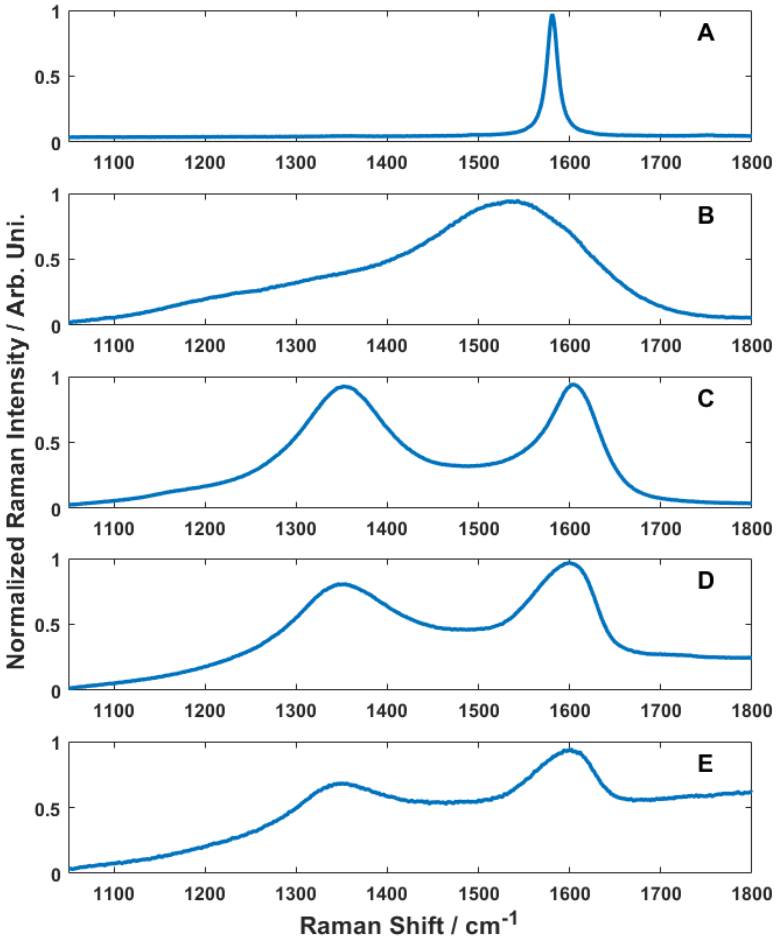

Figure 4-1: Raman spectra of highly oriented pyrolytic graphite (HOPG) (A), diamond-like carbon (DLC) (B), glassy carbon (GC) (C), hydrogenated graphite-like carbon (GLCH) (D) and hydrogenated polymer like carbon (PLCH) (E) were obtained using 532 nm laser, 5s exposure time, 3accumulations, 50x lens (NA = 0.75).

Figure 4-1 shows the mean Raman spectra of 5 different materials used for training the model. The training data set comprises of 588 Raman spectra for each material. The common feature among them is the presence of a G-band between 1540-1600 cm^{-1} and, in most cases, the D-band at 1360 cm^{-1}. The spectrum in Figure 4-1A is assigned to HOPG and it presents a G-band at 1580 cm^{-1}. This band is sharp and intense due to high crystallinity of the material and contributions from resonant Raman.[10, 15] The

spectrum in Figure 4-1B is assigned to DLC. It has a strong and broad band at 1540 cm^{-1} (G-band), which is related to the amorphous structure with low sp^2 carbon. According to the literature this material should contain 70 % sp^3 and 30 % sp^2.[13, 28, 92] The weak D-band (1360 cm^{-1}) could be related to the low number of aromatic rings.[12, 15] Figure 4-1C displays the Raman spectrum of GC. The G-band appears at 1600 cm^{-1}, and the D-band is at 1350 cm^{-1} in this case. Both bands are broader than for HOPG, but weaker and sharper than for DLC, which could be related to the amorphous aspect of the GC material and the predominance of sp^2 carbon. The D-band has higher intensity due to the disorder in the carbon structure and the presence of aromatic rings. The GLCH Raman spectrum shown in Figure 4-1D displays similar features to the GC, but the pronounced background indicates the presence of hydrogen in the composition of the material.[11, 93] It has been shown that it is possible to estimate hydrogen content based on the slope of that background and the G-band intensity.[11] Figure 4-1E shows the spectrum for PLCH, which is characterized by weaker G- and D-bands due to the strong fluorescent background. The higher background is related to a carbon film with hydrogen content above 35 %.[11] This training dataset exemplifies carbon materials with different Raman features (G-band position, G-band full width at half maximum, intensity ratio of D- and G-bands and H content) that can be used for their classification.

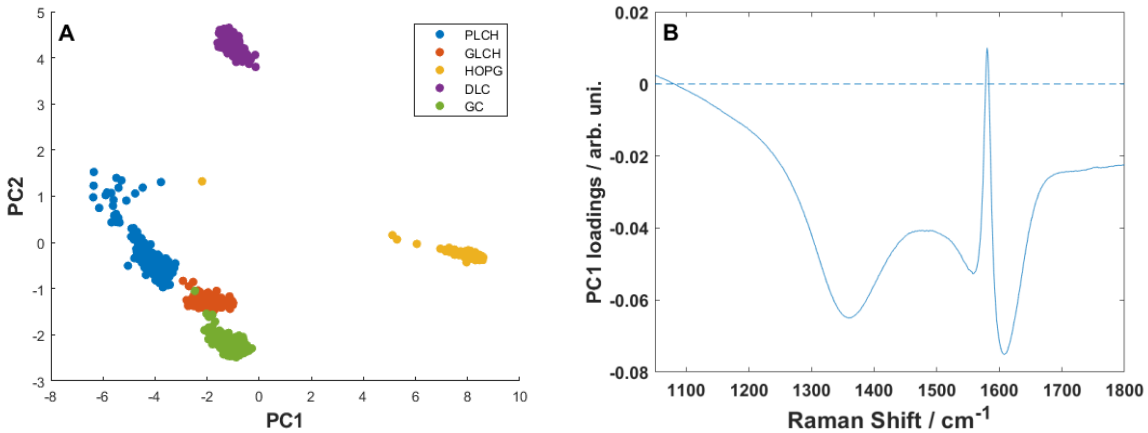

Figure 4-2: A) Plot of PCA scores of PLCH (red), GLCH (yellow), HOPG (green), DLC(blue) and GC (pink) for PC1 and PC2. B) PC1 loadings.

The plot of PCA scores for PC 1 and PC 2 are shown in the Figure 4-2A. Each point in the plot is equivalent to one Raman spectrum and different colors correspond to different types of materials. In the score plot we can notice the yellow (HOPG) points are very distant from the others and have very positive values for PC1, with most values ranging from 6 to 8. All the other four materials appear in the negative region of the PC1, in general between -2 and -4. Figure 4-2B shows the PC1 loadings plot and helps the understanding of how the model separated the Raman spectra from different carbon films. We can notice that each point in Figure 4-2B has a loading value ranging between -0.08 and 0.02. The samples with positive scores for PC1 will be those with high intensity Raman bands at 1580 cm^{-1}. This feature matches the sharp and strong G-band present in the HOPG samples. The samples with negative scores for PC1 will be the ones with broad G and D-bands (negative loadings in Figure 4-2B). This simple rationale explains why the scores of the amorphous materials are all in the negative region of Figure 4-2A, while the very crystalline material presented positive scores. It seems clear that the PC1 component can then distinguish between crystalline and amorphous materials. When compared to the G-band intensity, higher Raman intensities within the region between 1700 and 1800 cm^{-1} can be observed in carbon materials with hydrogen content >20 %. This increased background intensity due to hydrogenation can be seen in the Raman characteristics of PLCH and GLCH presented in Figure 4-2D and Figure 4-2E. The same region (1700 – 1800 cm^{-1}) in the PC1 loadings plot (Figure 4-2B) has negative values, supporting the negative scores observed for PLCH in Figure 4-2A.

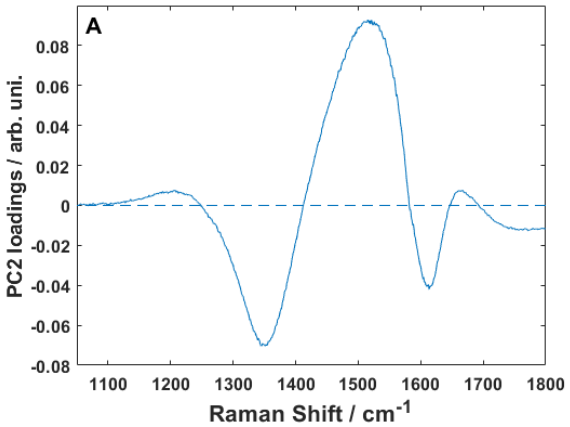

Figure 4-3: PC2 loadings versus Raman shift. This plot reflects the influence of various Raman shifts on the PC2 score.

The PC2 loadings plot is displayed in Figure 4-3 Similarly to the Figure 4-2B discussion above, a qualitative analysis of Figure 4-3 can be used to rationalize the PC2 scores displayed in Figure 4-2A. The positive loading values in Figure 4-3 can be correlated to the broadening and shifting of the G-band towards smaller wavenumbers. Such behaviour is characteristics for the Raman data of DLC (Figure 4-1B) which is the material with the highest scores for PC2 (Figure 4-2A). The high intensity of the D-band (1360 cm^{-1}) and the shift of the G-band to higher wavenumber (1600 cm^{-1}) are the features that contributed the most for the negative PC2 values. These features justify the most negative PC2 score for the GC , followed by GLCH and PLCH. All amorphous materials have similar values for PC1, but showed more variation in PC2 scores. The sp^2 amorphous materials, GC, GLCH and PLCH, have PC2 values smaller than 2, while the amorphous sp^3, DLC, has values around 4. We can then conclude that the PC2 scores contribute to distinguish sp^3 amorphous materials (DLC) from sp^2 amorphous species (GC, GLCH, PLCH).

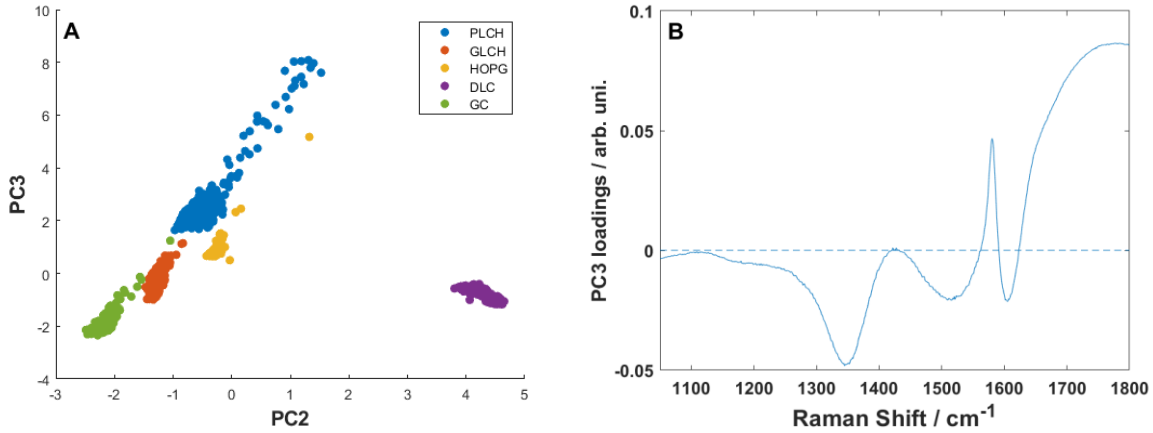

Figure 4-4: A) Plot of PCA scores of PLCH, GLCH, HOPG, DLC and GC for PC2 and PC3. B) PC3 loadings versus Raman shift. This plot reflects the influence of various Raman shifts on the PC3 score.

Figure 4-4A shows the scores plot for PC2 and PC3. The PC3 scores for GC and DLC are negative, while PLCH and HOPG have positive scores. Last, GLCH have samples in both positive and negative regions. Figure 4-4B is the PC3 loadings plot. From the positive region of PC3 (Figure 4-4B), it is possible to infer that the highest contribution is from the fluorescent background between 1700-1800 cm^{-1}. The fluorescence background can be observed in the spectra of hydrogenated materials (GLCH and PLCH), in accordance to the highest PC3 score found for PLCH. The other significant positive feature in Figure 4-4B is a sharp band around 1580 cm^{-1}, which probably contributes to the positive PC3 scores for HOPG observed in Figure 4-4A. The negative region in the PC3 loadings plot (Figure 4-4B) shows features at the D-band (1360 cm^{-1}) and some broadening or shifting of the G-band (from 1580 to 1520 cm^{-1}) region. The lack of a strong background above 1600 cm^{-1} , the presence of a strong D-band and a shifting of the G-band (1600 cm^{-1}) are all features that contribute to the negative PC3 values observed for GC in Figure 4-4A. On the other hand, the amorphous sp^2 materials (GC, GLCH, PLCH) showed a correlation between the degree of hydrogenation in the material and the PC3 score values. The higher positive values associated to the samples with higher hydrogen content suggest that the PC3 component will contribute to the classification of materials with different levels of hydrogenation.

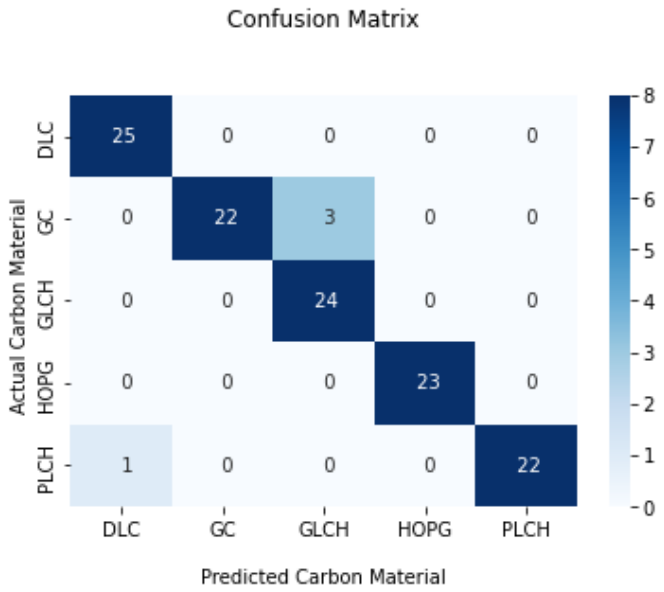

Figure 4-5: Confusion matrix comparing the type of carbon material predicted by PCA-LDA model (horizontal axis) and assignment from the literature (vertical axis).

The construction of a confusion matrix is a common approach to display the testing results in classification methods. The matrix shows how many predictions of the testing data set were correctly classified in the diagonal, while the off-diagonal indicates the cross-classification (confusion) for a particular method. Figure 4-5 shows the confusion matrix constructed from our testing dataset (obtained using the Raman spectra of carbon materials from the literature, as indicated in the experimental session). The horizontal axis of the confusion matrix (Figure 4-5) are the predictions made by the PCA-LDA, while the vertical axis are the classification of the material made in a particular publication. A quick glance to Figure 4-5 indicate that our model correctly classify most (97 %) of the materials reported in the literature. Figure 4-5 also indicates that 3 Raman spectra reported as GC were classified at GLCH by our model. Figure C-2 shows the Raman spectra of the three misidentified testing samples and the mean Raman spectra for GC and GLCH training samples. The region between 1400 and 1600 cm^{-1} seems to present more similarities between the three testing Raman spectra (assigned to CG) with the GLCH spectrum average spectrum of the training dataset. This misclassification is not unexpected, considering that both types of materials are

predominantly amorphous sp^2 and they mostly differ only by the differences in hydrogen content. Since the fluorescent background is related to the level of hydrogenation of the material, it is expected that the region between 1700 and 1800 cm^{-1} would be decisive for the classification, which was not the case from Figure C-2.

The other misclassification in Figure 4-5 was a spectrum assigned to PLCH that was classified as DLC. Figure C-3 shows the Raman spectra for the misclassified PLCH testing sample, the average PLCH spectrum from the training dataset and a DLC reference spectrum. The Raman spectrum of the PLCH testing sample in Figure C-3 is characterized by a very high background and shifted and broader G-band centered at 1550 cm^{-1}. A similar behaviour (broad G-band) is observed in the DLC training set in Figure C-3. However, the intense background is observed only in the PLCH misclassified testing sample. According to Casiraghi *et al*., the strong background should be the most relevant feature for classifying the material as PLCH.[11]

The small misclassification rate of this model (4 out of 116 spectra) could be further improved by increasing the size of the data training sets. For instance, the training set for HOPG, GC and DLC are based on Raman spectra from a single sample. This means that sample-to-sample variation was not completely captured in our model. Moreover, PCA and LDA are both linear methods, and the use of non-linear methods could improve the results. PCA is very efficient for reducing data dimensionality and visualization, but the method does not distinguish between relevant and irrelevant information in the training dataset. This means that some information relevant for classification may not be used.

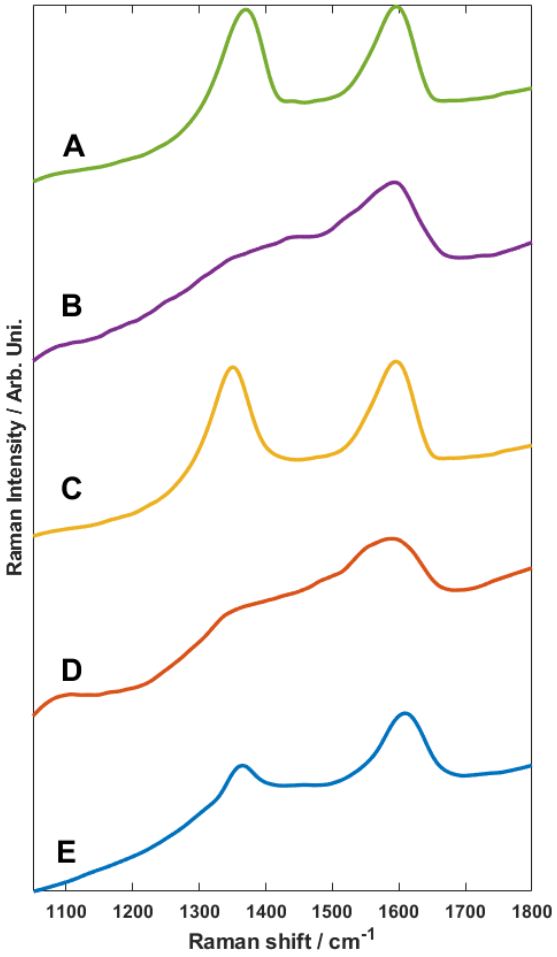

Figure 4-6: Raman spectra reported in the literature classified as DLC.

Figure 4-6 shows five Raman spectra from the literature classified as DLC.[94-96] The spectra are characterized by having a background that overlaps the G- and D- bands. These spectra were submitted to our model and they were all classified as PLCH instead. A quick comparison with Figure 4-1 makes it clear that the broad G-band in the DLC Raman spectra is not present in any of the spectra in Figure 4-6. Also, they all present a background typically related to hydrogenation and observed in samples described as either GLCH or PLCH.

The parameters for the Raman bands of the spectra in Figure 4-6 were fitted using Gaussian functions and they are shown in Table 4-1. The characteristics of the bands in Table 4-1 are useful to understand the PCA-LDA model classification. The five spectra in Figure 4-6 have G-band position between 1574-1602 cm^{-1}. These values are more similar to the mean spectra of PLCH (1596 cm^{-1}) than the DLC (1543 cm^{-1}). Another striking feature is the G-band full width at the half maximum (G-band full width at half maximum). The values observed for the five spectra collected from the literature and presented in Figure 4-6 range from 91-114 cm^{-1}. Those are more similar to the G-band full width at half maximum of the average in the PLCH training set (97 cm^{-1}) than for DLC (177 cm^{-1}). Finally, the hydrogen content (H cont.) was estimated, with values ranging between 40- 46 %. These values are similar to the range assigned to PLCH in the literature.[11]

Table 4-1: Features of Raman spectra from carbon materials obtained using two gaussians for fitting G- and D-band.

Raman Spectrum	G-band position / cm^{-1}	Width (G) / cm^{-1}	I(D)/I(G) / Arb. Uni.	H cont. / % at.
A	1602	97	0.75	44
B	1575	114	0.72	46
C	1574	91	1.1	40
D	1584	97	0.87	44
E	1585	91	1.2	40
DLC (mean)	1543	177	0.42	20
PLCH (mean)	1596	97	0.88	38

Overall, our results show a model elaborated using PCA-LDA can correctly classify carbon materials based on their Raman spectra. However, overestimating the machine learning capabilities is an issue

reported in the literature that must be always be taken into consideration.[97, 98] A larger number of samples generated from different synthetic methods should be useful for better training, testing and to help identifying the limitations of the models. There is a vast literature emphasizing the utility of Raman spectroscopy for carbon materials characterization.[4, 10, 99] The use of machine learning for automated quality control of carbon materials, as suggested here, can be powerful tool to materials scientists and technologists.

4.5 Conclusions

PC1 could separate the crystalline material (HOPG) from the amorphous ones. Among the amorphous materials, PC2 could be used for separating the materials between two classes, predominance of sp^2 (GC, GLCH and PLCH) or sp^3 (DLC) hybridization. The PC3 has shown some correlation with the variation of hydrogen content in the sample, allowing the separation of materials with different levels of hydrogenation. The proposed PCA-LDA model has been successful classifying Raman spectra of Carbon materials, with accuracy of 97 %. Also, the potential of this model has been emphasized by examples of assignment made in the literature that could be improved. The identified materials were HOPG, GC, DLC, GLCH and PLCH. Then, this work shows the promising use of Raman and machine learning for automated classification of Carbon materials.

4.6 References

(1) Speranza, G. The Role of Functionalization in the Applications of Carbon Materials: An Overview. *C-Journal of Carbon Research* **2019**, *5* (4), 84. DOI: 10.3390/c5040084.
(2) Prasanth, R.; Ammini, S.; Ge, L.; Thakur, M.; Thakur, V. K. Carbon Allotropes and Fascinated Nanostructures: The High-Impact Engineering Materials of the Millennium. 2015; pp 2-27.
(3) *Chemical Functionalization of Carbon Nanomaterials: Chemistry and Applications*; CRC Press, 2015. DOI: 10.1201/b18724.

(4) Morelos-Gomez, A.; Terrones, M.; Endo, M. Data Science Applied to Carbon Materials: Synthesis, Characterization, and Applications. *Advanced Theory and Simulations* **2022**, *5* (2), 2100205. DOI: 10.1002/adts.202100205.

(5) Titirici, M. M.; White, R. J.; Brun, N.; Budarin, V. L.; Su, D. S.; del Monte, F.; Clark, J. H.; MacLachlan, M. J. Sustainable carbon materials. *Chemical Society Reviews* **2015**, *44* (1), 250-290. DOI: 10.1039/c4cs00232f.

(6) Harris, P. J. F. Engineering carbon materials with electricity. *Carbon* **2017**, *122*, 504-513. DOI: 10.1016/j.carbon.2017.06.084.

(7) Wissler, M. Graphite and carbon powders for electrochemical applications. *Journal of Power Sources* **2006**, *156* (2), 142-150. DOI: 10.1016/j.jpowsour.2006.02.064.

(8) Sharma, S. Glassy Carbon: A Promising Material for Micro- and Nanomanufacturing. *Materials* **2018**, *11* (10), 1857, . DOI: 10.3390/ma11101857.

(9) Zia, A. W. New generation carbon particles embedded diamond-like carbon coatings for transportation industry. *Advances in Smart Coatings and Thin Films for Future Industrial and Biomedical Engineering Applications* **2020**, 307-332. DOI: 10.1016/b978-0-12-849870-5.00004-5.

(10) Ferrari, A. C.; Basko, D. M. Raman spectroscopy as a versatile tool for studying the properties of graphene. Nature Publishing Group: 2013; Vol. 8, pp 235-246.

(11) Casiraghi, C.; Piazza, F.; Ferrari, A. C.; Grambole, D.; Robertson, J. Bonding in hydrogenated diamond-like carbon by Raman spectroscopy. *Diamond and Related Materials* **2005**, *14* (3-7), 1098-1102, Article; Proceedings Paper. DOI: 10.1016/j.diamond.2004.10.030.

(12) Ferrari, A. C.; Robertson, J. Raman spectroscopy of amorphous, nanostructured, diamond-like carbon, and nanodiamond. *Philosophical Transactions of the Royal Society a-Mathematical Physical and Engineering Sciences* **2004**, *362* (1824), 2477-2512, Article. DOI: 10.1098/rsta.2004.1452.

(13) Ferrari, A. C.; Robertson, J. Interpretation of Raman spectra of disordered and amorphous carbon. *Physical Review B* **2000**, *61* (20), 14095-14107. DOI: 10.1103/PhysRevB.61.14095.

(14) Kuhar, N.; Sil, S.; Verma, T.; Umapathy, S. Challenges in application of Raman spectroscopy to biology and materials. *RSC Advances* **2018**, *8* (46), 25888-25908, Review. DOI: 10.1039/c8ra04491k.

(15) Ferrari, A. C.; Robertson, J. Resonant Raman spectroscopy of disordered, amorphous, and diamondlike carbon. *Physical Review B* **2001**, *64* (7), 075414. DOI: 10.1103/PhysRevB.64.075414.

(16) Silva, D. L.; Campos, J. L. E.; Fernandes, T. F. D.; Rocha, J. N.; Machado, L. R. P.; Soares, E. M.; Miquita, D. R.; Miranda, H.; Rabelo, C.; Neto, O. P. V.; et al. Raman spectroscopy analysis of number of layers in mass-produced graphene flakes. *Carbon* **2020**, *161*, 181-189. DOI: 10.1016/j.carbon.2020.01.050.

(17) Yeh, Y. T.; Gulino, K.; Zhanga, Y. H.; Sabestien, A.; Chou, T. W.; Zhou, B.; Lin, Z.; Albert, I.; Lu, H. G.; Swaminathan, V.; et al. A rapid and label-free platform for virus capture and identification from clinical samples. *Proceedings of the National Academy of Sciences of the United States of America* **2020**, *117* (2), 895-901. DOI: 10.1073/pnas.1910113117.

(18) Gonzalez-Durruthy, M.; Alberic, L. C.; Curti, C.; Naal, Z.; Atique-Sawazaki, D. T.; Vazquez-Naya, J. M.; Gonzalez-Diaz, H.; Munteanu, C. R. Experimental-Computational Study of Carbon Nanotube Effects on Mitochondrial Respiration: In Silico Nano-QSPR Machine Learning Models Based on New Raman Spectra Transform with Markov-Shannon Entropy Invariants. *Journal of Chemical Information and Modeling* **2017**, *57* (5), 1029-1044. DOI: 10.1021/acs.jcim.6b00458.

(19) Su, H. P.; Lin, S.; Deng, S. W.; Lian, C.; Shang, Y. Z.; Liu, H. L. Predicting the capacitance of carbon-based electric double layer capacitors by machine learning. *Nanoscale Advances* **2019**, *1* (6), 2162-2166. DOI: 10.1039/c9na00105k.

(20) Wang, J. S.; Li, Z.; Yan, S. C.; Yu, X.; Ma, Y. Q.; Ma, L. Modifying the microstructure of algae-based active carbon and modelling supercapacitors using artificial neural networks. *RSC Advances* **2019**, *9* (26), 14797-14808. DOI: 10.1039/c9ra01255a.

(21) Nikolaev, P.; Hooper, D.; Webber, F.; Rao, R.; Decker, K.; Krein, M.; Poleski, J.; Barto, R.; Maruyama, B. Autonomy in materials research: a case study in carbon nanotube growth. *Npj Computational Materials* **2016**, *2*, 16031. DOI: 10.1038/npjcompumats.2016.31.

(22) Chang, J.; Nikolaev, P.; Carpena-Nunez, J.; Rao, R.; Decker, K.; Islam, A. E.; Kim, J.; Pitt, M. A.; Myung, J. I.; Maruyama, B. Efficient Closed-loop Maximization of Carbon Nanotube Growth Rate using Bayesian Optimization. *Scientific Reports* **2020**, *10* (1), 9040. DOI: 10.1038/s41598-020-64397-3.

(23) Wahab, H.; Jain, V.; Tyrrell, A. S.; Seas, M. A.; Kotthoff, L.; Johnson, P. A. Machine-learning-assisted fabrication: Bayesian optimization of laser-induced graphene patterning using in-situ Raman analysis. *Carbon* **2020**, *167*, 609-619. DOI: 10.1016/j.carbon.2020.05.087.

(24) Campos, J. L. E.; Miranda, H.; Rabelo, C.; Sandoz-Rosado, E.; Pandey, S.; Riikonen, J.; Cano-Marquez, A. G.; Jorio, A. Applications of Raman spectroscopy in graphene-related materials and the development of parameterized PCA for large-scale data analysis. *Journal of Raman Spectroscopy* **2018**, *49* (1), 54-65, Article. DOI: 10.1002/jrs.5225.

(25) Kumbhar, D.; Palliyarayil, A.; Reghu, D.; Shrungar, D.; Umapathy, S.; Sil, S. Rapid discrimination of porous bio-carbon derived from nitrogen rich biomass using Raman spectroscopy and artificial intelligence methods. *Carbon* **2021**, *178*, 792-802. DOI: 10.1016/j.carbon.2021.03.064.

(26) Tai, F. C.; Lee, S. C.; Chen, J.; Wei, C.; Chang, S. H. Multipeak fitting analysis of Raman spectra on DLCH film. *Journal of Raman Spectroscopy* **2009**, *40* (8), 1055-1059. DOI: 10.1002/jrs.2234.

(27) daFonseca, B. G.; Thind, S. S.; Brolo, A. G. Raman maps reveal heterogeneous hydrogenation on carbon materials. *Journal of Raman Spectroscopy* **2021**, *52* (2), 516-524. DOI: 10.1002/jrs.6042.

(28) Robertson, J. Diamond-like amorphous carbon. *Materials Science & Engineering R-Reports* **2002**, *37* (4-6), 129-281. DOI: 10.1016/s0927-796x(02)00005-0.

(29) Tallant, D. R.; Parmeter, J. E.; Siegal, M. P.; Simpson, R. L. The thermal-stability of diamond-like carbon. *Diamond and Related Materials* **1995**, *4* (3), 191-199. DOI: 10.1016/0925-9635(94)00243-6.

(30) Irmer, G.; Dorner-Reisel, A. Micro-Raman studies on DLC coatings. *Advanced Engineering Materials* **2005**, *7* (8), 694-705. DOI: 10.1002/adem.200500006.

(31) Watanabe, H.; Takahashi, K.; Iwaki, M. Structural characterization of ion-implanted hopg and glass-like carbon by laser raman-spectroscopy. *Nuclear Instruments & Methods in Physics Research Section B-Beam Interactions with Materials and Atoms* **1993**, *80-1*, 1489-1493. DOI: 10.1016/0168-583x(93)90827-s.

(32) Chandran, M.; Shasha, M.; Michaelson, S.; Hoffman, A. Incorporation of low energy activated nitrogen onto HOPG surface: Chemical states and thermal stability studies by in-situ XPS and Raman spectroscopy. *Applied Surface Science* **2016**, *382*, 192-201. DOI: 10.1016/j.apsusc.2016.04.030.

(33) Solin, S. A.; Caswell, N. Raman-scattering from alkali graphite-intercalation compounds. *Journal of Raman Spectroscopy* **1981**, *10* (JAN), 129-135. DOI: 10.1002/jrs.1250100124.

(34) Rantitsch, G.; Lammerer, W.; Fisslthaler, E.; Mitsche, S.; Kaltenbock, H. On the discrimination of semi-graphite and graphite by Raman spectroscopy. *International Journal of Coal Geology* **2016**, *159*, 48-56. DOI: 10.1016/j.coal.2016.04.001.

(35) Li, K.; Liu, Q. F.; Cheng, H. F.; Hu, M. S.; Zhang, S. Classification and carbon structural transformation from anthracite to natural coaly graphite by XRD, Raman spectroscopy, and HRTEM. *Spectrochimica Acta Part a-Molecular and Biomolecular Spectroscopy* **2021**, *249*. DOI: 10.1016/j.saa.2020.119286.

(36) Yuan, L.; Liu, Q. F.; Mathews, J. P.; Zhang, H.; Wu, Y. K. Quantifying the Structural Transitions of Chinese Coal to Coal-Derived Natural Graphite by XRD, Raman Spectroscopy, and HRTEM Image Analyses. *Energy & Fuels* **2021**, *35* (3), 2335-2346. DOI: 10.1021/acs.energyfuels.0c04019.

(37) Inam, A.; Brydson, R.; Edmonds, D. V. Raman spectroscopy study of the crystallinity of graphite formed in an experimental free-machining steel. *Materials Characterization* **2020**, *163*, 110264. DOI: 10.1016/j.matchar.2020.110264.

(38) Moroz, T. N.; Edwards, H. G. M.; Ponomarchuk, V. A.; Pyryaev, A. N.; Palchik, N. A.; Goryainov, S. V. Raman spectra of a graphite-nontronite association in marbles from Oltrek Island (Lake Baikal, Russia). *Journal of Raman Spectroscopy* **2020**, *51* (9), 1885-1893. DOI: 10.1002/jrs.5763.

(39) Miranda, D. A.; Chaves, A. D.; Campello, M. S.; Ramos, S. Origin and thermometry of graphites from Itapecerica supracrustal succession of the southern Sao Francisco Craton by C isotopes, X-ray diffraction, and Raman spectroscopy. *International Geology Review* **2019**, *61* (15), 1864-1875. DOI: 10.1080/00206814.2018.1564073.

(40) Bradley, D. A.; Rozaila, Z. S.; Khandaker, M. U.; Almugren, K. S.; Meevasana, W.; Sani, S. F. A. Raman spectroscopy and X-ray photo-spectroscopy analysis of graphite media irradiated at low doses. *Applied Radiation and Isotopes* **2019**, *147*, 105-112. DOI: 10.1016/j.apradiso.2019.02.016.

(41) Zhang, C.; Santosh, M. Coupled laser Raman spectroscopy and carbon stable isotopes of graphite from the khondalite belt of Kerala, southern India. *Lithos* **2019**, *334*, 245-253. DOI: 10.1016/j.lithos.2019.03.026.

(42) Ray, K.; McCreery, R. L. Spatially resolved Raman spectroscopy of carbon electrode surfaces: Observations of structural and chemical heterogeneity. *Analytical Chemistry* **1997**, *69* (22), 4680-4687. DOI: 10.1021/ac9705531.

(43) Seki, H.; McClelland, G. M.; Bullock, D. C. Raman-spectroscopy of disk coatings in a working magnetic disk drive. *Wear* **1987**, *116* (3), 381-391. DOI: 10.1016/0043-1648(87)90184-0.

(44) Le Guillou, M.; Rouzaud, J. N.; Deldicque, D.; Toulhoat, N.; Pipon, Y.; Moncoffre, N. Structural and nanostructural behavior of deuterium implanted Highly Ordered Pyrolytic Graphite investigated by combined High Resolution Transmission Electron Microscopy, Scanning Electron Microscopy and Raman microspectrometry. *Carbon* **2015**, *94*, 277-284. DOI: 10.1016/j.carbon.2015.06.065.

(45) Krishna, R.; Jones, A. N.; McDermott, L.; Marsden, B. J. Neutron irradiation damage of nuclear graphite studied by high-resolution transmission electron microscopy and Raman spectroscopy. *Journal of Nuclear Materials* **2015**, *467*, 557-565. DOI: 10.1016/j.jnucmat.2015.10.027.

(46) Ammar, M. R.; Galy, N.; Rouzaud, J. N.; Toulhoat, N.; Vaudey, C. E.; Simon, P.; Moncoffre, N. Characterizing various types of defects in nuclear graphite using Raman scattering: Heat treatment, ion irradiation and polishing. *Carbon* **2015**, *95*, 364-373. DOI: 10.1016/j.carbon.2015.07.095.

(47) Yang, D. Q.; Sacher, E. s-p hybridization in highly oriented pyrolytic graphite and its change on surface modification, as studied by X-ray photoelectron and Raman spectroscopies. *Surface Science* **2002**, *504* (1-3), 125-137.

(48) Kimata, T.; Kakitani, K.; Yamamoto, S.; Yamaki, T.; Terai, T.; Nakamura, K. G. Platinum nanoparticles on HOPG surface modified by 380 keV Ar+ irradiation: TEM and Raman studies. *Radiation Effects and Defects in Solids* **2020**, *175* (5-6), 433-439. DOI: 10.1080/10420150.2020.1737866.

(49) Tan, P. H.; Deng, Y. M.; Zhao, Q. Temperature-dependent Raman spectra and anomalous Raman phenomenon of highly oriented pyrolytic graphite. *Physical Review B* **1998**, *58* (9), 5435-5439. DOI: 10.1103/PhysRevB.58.5435.

(50) Madito, M. J.; Ismail, M. Y. A.; Hlatshwayo, T. T.; Mtshali, C. B. The nature of surface defects in Xe ion-implanted glassy carbon annealed at high temperatures: Raman spectroscopy analysis. *Applied Surface Science* **2020**, *506, 145001*. DOI: 10.1016/j.apsusc.2019.145001.

(51) Andrianova, N. N.; Borisov, A. M.; Kazakov, V. A.; Makunin, A. V.; Mashkova, E. S.; Ovchinnikov, M. A. Modification of the Nanoglobular Structure of Glassy Carbon by Heat Treatment and Ion Irradiation. *Journal of Surface Investigation* **2019**, *13* (5), 802-808. DOI: 10.1134/s1027451019050033.

(52) Nathan, M. I.; Smith, J. E.; Tu, K. N. Raman-spectra of glassy carbon. *Journal of Applied Physics* **1974**, *45* (5), 2370-2370. DOI: 10.1063/1.1663599.

(53) Csontos, J.; Papa, Z.; Gardian, A.; Fule, M.; Budai, J.; Toth, Z. Spectroscopic ellipsometric and Raman spectroscopic investigations of pulsed laser treated glassy carbon surfaces. *Applied Surface Science* **2015**, *336*, 343-348. DOI: 10.1016/j.apsusc.2014.12.133.

(54) Ray, K. R.; McCreery, R. L. Characterization of the surface carbonyl and hydroxyl coverage on glassy carbon electrodes using Raman spectroscopy. *Journal of Electroanalytical Chemistry* **1999**, *469* (2), 150-158. DOI: 10.1016/s0022-0728(99)00197-7.

(55) Solopova, N. A.; Dubrovinskaia, N.; Dubrovinsky, L. Raman spectroscopy of glassy carbon up to 60 GPa. *Applied Physics Letters* **2013**, *102* (12), 121909. DOI: 10.1063/1.4798660.

(56) Mayer, P.; Holze, R. Pyridine as a probe molecule for surface enhanced Raman spectroscopy of the silver-modified glassy carbon/solution interface. *Surface Science* **2003**, *522* (1-3), 55-63. DOI: 10.1016/s0039-6028(02)02271-9.

(57) Yi, Y. M.; Weinberg, G.; Prenzel, M.; Greiner, M.; Heumann, S.; Becker, S.; Schlogl, R. Electrochemical corrosion of a glassy carbon electrode. *Catalysis Today* **2017**, *295*, 32-40. DOI: 10.1016/j.cattod.2017.07.013.

(58) Yoshikawa, M.; Nagai, N.; Matsuki, M.; Fukuda, H.; Katagiri, G.; Ishida, H.; Ishitani, A.; Nagai, I. Raman-scattering from sp2 carbon clusters. *Physical Review B* **1992**, *46* (11), 7169-7174. DOI: 10.1103/PhysRevB.46.7169.

(59) Jovanovic, Z.; Kalijadis, A.; Lausevic, M.; Lausevic, Z. The evolution of hydrogen from proton irradiated glassy carbon. *Nuclear Instruments & Methods in Physics Research Section B-Beam Interactions with Materials and Atoms* **2011**, *269* (21), 2578-2583. DOI: 10.1016/j.nimb.2011.07.014.

(60) Lespade, P.; Aljishi, R.; Dresselhaus, M. S. Model for raman-scattering from incompletely graphitized carbons. *Carbon* **1982**, *20* (5), 427-431. DOI: 10.1016/0008-6223(82)90043-4.

(61) Cho, H. J.; Lee, K. W.; Lee, C. E. Room-temperature ferromagnetism in hydrothermally treated glassy carbon. *Current Applied Physics* **2013**, *13* (9), 2055-2058. DOI: 10.1016/j.cap.2013.09.007.

(62) Hlatshwayo, T. T.; Sebitla, L. D.; Njoroge, E. G.; Mlambo, M.; Malherbe, J. B. Annealing effects on the migration of ion-implanted cadmium in glassy carbon. *Nuclear Instruments & Methods in Physics Research Section B-Beam Interactions with Materials and Atoms* **2017**, *395*, 34-38. DOI: 10.1016/j.nimb.2017.01.086.

(63) Ismail, M. Y. A.; Malherbe, J. B.; Odutemowo, O. S.; Njoroge, E. G.; Hlatshwayo, T. T.; Mlambo, M.; Wendler, E. Investigating the effect of heat treatment on the diffusion behaviour of xenon implanted in glassy carbon. *Vacuum* **2018**, *149*, 74-78. DOI: 10.1016/j.vacuum.2017.12.021.

(64) Fang, L.; Lannin, J. S. Disorder induced raman-scattering of nanocrystalline carbon. *Applied Physics Letters* **1992**, *61* (17), 2116-2118.

(65) Malherbe, J. B.; Odutemowo, O. S.; Njoroge, E. G.; Langa, D. F.; Hlatshwayo, T. T.; Theron, C. C. Ion bombardment of glassy carbon. *Vacuum* **2018**, *149*, 19-22. DOI: 10.1016/j.vacuum.2017.11.006.

(66) McCulloch, D. G.; Prawer, S.; Hoffman, A. Structural investigation of xenon-ion-beam-irradiated glassy-carbon. *Physical Review B* **1994**, *50* (9), 5905-5917, Article. DOI: 10.1103/PhysRevB.50.5905.

(67) Odutemowo, O. S.; Malherbe, J. B.; Limbach, R.; Wondraczek, L.; Wendler, E.; Undisz, A.; Njoroge, E. G.; Idisi, D. O.; Dhlamini, M. S. Changes in the mechanical, structural and electrical properties of glassy carbon due to strontium and silver co-implantation and annealing. *Applied Surface Science* **2021**, *537*, 147929. DOI: 10.1016/j.apsusc.2020.147929.

(68) Odutemowo, O. S.; Malherbe, J. B.; Prinsloo, L.; Langa, D. F.; Wendler, E. High temperature annealing studies of strontium ion implanted glassy carbon. *Nuclear Instruments & Methods in Physics Research Section B-Beam Interactions with Materials and Atoms* **2016**, *371*, 332-335. DOI: 10.1016/j.nimb.2015.10.054.

(69) Prawer, S.; Rossouw, C. J. Structural investigation of helium ion-beam-irradiated glassy-carbon. *Journal of Applied Physics* **1988**, *63* (9), 4435-4439. DOI: 10.1063/1.340162.

(70) Saitoh, H.; Shinada, T.; Ohkawara, Y.; Ohshio, S.; Hiraga, H.; Inoue, T. Surface modification of glassy carbon by pulsed laser irradiation with several wavelengths. *Japanese Journal of Applied Physics Part 1-Regular Papers Brief Communications & Review Papers* **2002**, *41* (8), 5359-5366. DOI: 10.1143/jjap.41.5359.

(71) Solopova, N. A.; Dubrovinskaia, N.; Dubrovinsky, L. Synthesis of nanocrystalline diamond from glassy carbon balls. *Journal of Crystal Growth* **2015**, *412*, 54-59. DOI: 10.1016/j.jcrysgro.2014.11.041.

(72) Swain, G. M. The susceptibility to surface corrosion in acidic fluoride media - a comparison of diamond, hopg, and glossy carbon electrodes. *Journal of the Electrochemical Society* **1994**, *141* (12), 3382-3393. DOI: 10.1149/1.2059343.

(73) Kalish, R.; Lifshitz, Y.; Nugent, K.; Prawer, S. Thermal stability and relaxation in diamond-like-carbon. A Raman study of films with different sp(3) fractions (ta-C to a-C). *Applied Physics Letters* **1999**, *74* (20), 2936-2938. DOI: 10.1063/1.123971.

(74) Liu, Y.; Erdemir, A.; Meletis, E. I. A study of the wear mechanism of diamond-like carbon films. *Surface & Coatings Technology* **1996**, *82* (1-2), 48-56. DOI: 10.1016/0257-8972(95)02623-1.

(75) Oguri, K.; Arai, T. Tribological properties and characterization of diamond-like carbon coatings with silicon prepared by plasma-assisted chemical vapor-deposition. *Surface & Coatings Technology* **1991**, *47* (1-3), 710-721. DOI: 10.1016/0257-8972(91)90344-v.

(76) Prawer, S.; Nugent, K. W.; Lifshitz, Y.; Lempert, G. D.; Grossman, E.; Kulik, J.; Avigal, I.; Kalish, R. Systematic variation of the Raman spectra of DLC films as a function of sp(2):sp(3) composition. *Diamond and Related Materials* **1996**, *5* (3-5), 433-438. DOI: 10.1016/0925-9635(95)00363-0.

(77) Tamor, M. A.; Vassell, W. C. Raman fingerprinting of amorphous-carbon films. *Journal of Applied Physics* **1994**, *76* (6), 3823-3830. DOI: 10.1063/1.357385.

(78) Voevodin, A. A.; Zabinski, J. S. Load-adaptive crystalline-amorphous nanocomposites. *Journal of Materials Science* **1998**, *33* (2), 319-327. DOI: 10.1023/a:1004307426887.

(79) Friedmann, T. A.; McCarty, K. F.; Barbour, J. C.; Siegal, M. P.; Dibble, D. C. Thermal stability of amorphous carbon films grown by pulsed laser deposition. *Applied Physics Letters* **1996**, *68* (12), 1643-1645. DOI: 10.1063/1.115891.

(80) Ramsteiner, M.; Wagner, J. Resonant raman-scattering of hydrogenated amorphous-carbon - evidence for pi-bonded carbon clusters. *Applied Physics Letters* **1987**, *51* (17), 1355-1357. DOI: 10.1063/1.98677.

(81) Dillon, R. O.; Woollam, J. A.; Katkanant, V. Use of raman-scattering to investigate disorder and crystallite formation in as-deposited and annealed carbon-films. *Physical Review B* **1984**, *29* (6), 3482-3489. DOI: 10.1103/PhysRevB.29.3482.

(82) Shiao, J.; Hoffman, R. W. Studies of diamond-like and nitrogen-containing diamond-like carbon using laser Raman spectroscopy. *Thin Solid Films* **1996**, *283* (1-2), 145-150. DOI: 10.1016/0040-6090(95)08151-8.

(83) Kozu, T.; Yamaguchi, M.; Kawaguchi, M.; Shima, H.; Kim, J. W.; Matsuoka, M.; Nishida, K.; Yamamoto, T. Evaluating of Diamond Like Carbon Using Deep UV Raman Spectroscopy. *Integrated Ferroelectrics* **2014**, *157* (1), 147-156. DOI: 10.1080/10584587.2014.912908.

(84) Mano, T.; Tanaka, S.; Takagi, M.; Sugiyama, O.; Takai, O. Tribological Characteristics of Diamond-Like Carbon Based Double-Layers Film Prepared by DC-Plasma Ionization Deposition. *Journal of Photopolymer Science and Technology* **2011**, *24* (4), 463-466. DOI: 10.2494/photopolymer.24.463.

(85) Kong, X. H.; Wang, S. A.; Zhao, H. P.; He, Y. D. Preparation of diamond-like carbon films by cathodic micro-arc discharge in aqueous solutions. *Thin Solid Films* **2010**, *518* (15), 4211-4214. DOI: 10.1016/j.tsf.2009.12.078.

(86) Kijaszek, W.; Oleszkiewicz, W. Optimization of radio frequency inductively coupled plasma enhanced chemical vapour deposition process of diamond-like carbon films. *Optica Applicata* **2016**, *46* (2), 167-172. DOI: 10.5277/oa160201.

(87) Sonderby, S.; Berthelsen, A. N.; Almtoft, K. P.; Christensen, B. H.; Nielsen, L. P.; Bottiger, J. Optimization of the mechanical properties of magnetron sputtered diamond-like carbon coatings. *Diamond and Related Materials* **2011**, *20* (5-6), 682-686. DOI: 10.1016/j.diamond.2011.03.016.

(88) Scharf, T. W.; Singer, I. L. Thickness of diamond-like carbon coatings quantified with Raman spectroscopy. *Thin Solid Films* **2003**, *440* (1-2), 138-144. DOI: 10.1016/s0040-6090(03)00703-x.

(89) Burkhard, G.; Tamura, H.; Tanabe, Y.; Sawaoka, A.; Uematsu, K.; Ohmura, S. Formation of diamond-like carbon bonds by pulsed electron-beam annealing of microcrystalline carbon copper bilayer film systems. *Journal of Materials Science Letters* **1994**, *13* (17), 1281-1283. DOI: 10.1007/bf00270961.

(90) Rohatgi, A. *WebPlotDigitizer*. https://apps.automeris.io/wpd/ (accessed 2022 August 18th 2022).

(91) Pedregosa, F.; Gäel, V.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine learning in Python. *Journal of machine learning research* **2011**, *12*, 2825-2830.

(92) Ferrari, A. C. Determination of bonding in diamond-like carbon by Raman spectroscopy. *Diamond and Related Materials* **2002**, *11* (3-6), 1053-1061. DOI: 10.1016/s0925-9635(01)00730-0.

(93) Casiraghi, C.; Ferrari, A. C.; Robertson, J. Raman spectroscopy of hydrogenated amorphous carbons. *Physical Review B* **2005**, *72* (8), 085401, Article. DOI: 10.1103/PhysRevB.72.085401.

(94) Guo, D.; Cai, K.; Li, L. T.; Huang, Y.; Gui, Z. L.; Zhu, H. S. Evaluation of diamond-like carbon films electrodeposited on an Al substrate from the liquid phase with pulse-modulated power. *Carbon* **2001**, *39* (9), 1395-1398. DOI: 10.1016/S0008-6223(00)00257-8.

(95) Cao, C. B.; Zhu, H. S.; Wang, H. Electrodeposition diamond-like carbon films from organic liquids. *Thin Solid Films* **2000**, *368* (2), 203-207. DOI: 10.1016/s0040-6090(00)00765-3.

(96) Ban, C. L.; Huang, B. X.; Wang, C. Z.; Wang, L. P.; Shao, X. Electrodeposition of Diamond-like Carbon (DLC) Films on Mg by Plasma Electrolysis. *Electrochemistry* **2013**, *81* (12), 977-980. DOI: 10.5796/electrochemistry.81.977.

(97) Kapoor, S.; Narayanan, A. Leakage and the Reproducibility Crisis in ML-based Science. arXiv: 2022.

(98) *The Reproducibility Crisis in ML-based Science*. https://sites.google.com/princeton.edu/rep-workshop/home (accessed 2022 August 23rd).

(99) Yin, H.; Sun, Z. H.; Wang, Z.; Tang, D. W.; Pang, C. H.; Yu, X. F.; Barnard, A. S.; Zhao, H. T.; Yin, Z. Y. The data-intensive scientific revolution occurring where two-dimensional materials meet machine learning. *Cell Reports Physical Science* **2021**, *2* (7), 100482. DOI: 10.1016/j.xcrp.2021.100482.

5 Summary and Outlook

5.1 Summary and conclusions

In this work, we have expanded the application of Raman spectroscopy to carbon materials by using Raman mapping and machine learning methods.

In chapter 2, Raman maps and histograms of hydrogenation have been shown as valuable tools for the characterizing carbon materials and should be utilized for a proper description of materials obtained electrochemically. Many Raman features (e.g. hydrogen content, G-band position, G-band full width at half maximum, and intensity ratio between G- and D-band) were plotted spatially to evaluate local chemical composition and properties. For example, carbon coatings with different levels of hydrogenation were found to be correlated to local reduction conditions (hydrogen evolution, presence of oxidizing anions). Time was a sensitive feature for variation in film composition, and thickness was correlated to homogeneity. The synthesis of carbon materials by electrochemical methods has many advantages, but the Raman mapping analysis showed a that materials produced by this route presents heterogeneities that need to be considered for potential applications.

Carbon species can be dangerous for human health, in particular for workers exposed to diesel emissions. Chapter 3 presents the use of Raman and partial least squares for estimating elemental carbon, organic carbon and total carbon in workplace environment. Six models were generated for analysis of three carbon species using two different laser excitations. The results showed the advantage of using 785 nm excitation over 532 nm to minimize strong fluorescent background. Based on the agreement of the values predicted and measured by the standard technique (NIOSH method), this method is promising for future implementation in the evaluation of carbon contamination in workplace environment. In most cases, the median of 10 spectra produced a better prediction than a single spectrum. The models for OC and TC showed the fluorescent background as an important variable for prediction. A significant deviation

in the prediction was observed in samples with smaller amounts of EC or TC. When predicting TC, the samples with high proportion of OC showed larger errors. It's also important to notice the NIOSH 5040 has significant limitations that could contribute for the deviation observed. For instance, the value is an average obtained over 8 hours and has low accuracy.

In chapter 4 is shown the potential of implementing automated classification of carbon materials using Raman spectroscopy and principal component analysis and linear discriminant analysis (PCA-LDA). In this study, the PCA-LDA model had to classify a Raman spectrum as one of the five different materials (highly oriented pyrolytic graphite, glassy carbon, diamond-like carbon, hydrogenated graphite-like carbon or hydrogenated polymer like carbon). The PC1 showed the capability of separating crystalline from amorphous materials. PC2 was helpful for separating the materials according to the predominant hybridization, sp^2 or sp^3. Finally, the PC3 was able to separate the carbon materials with different levels of hydrogenation. The accuracy of 97 % emphasizes the high potential for application of this model. Last, some spectra from the literature assigned to DLC were analyzed and the model offered a different classification. The PCA-LDA model seemed to have made better choices, demonstrating the classification of Raman spectra can be misleading and some help from artificial intelligence can be welcome.

5.2 Outlook and future directions

One possible venue of exploring Raman maps would involve the improvement of the resolution by using tip-enhanced Raman spectroscopy. We have attempted that and unfortunately the data obtained couldn't reveal further details about the sample. This result is in agreement with an article from literature, which showed similar results between the Raman and TER spectra from amorphous carbon. Another possibility is obtaining Raman spectra of the carbon films during the electrochemical deposition. Brolo group has previous experience with electrochemical cells and Raman measurements. This project could

provide better understanding about the carbon structure in different moments of the synthetic process. The formation of bubbles observed during the electrochemical deposition could affect the Raman measurements and be an issue for this investigation.

In chapter 3, the model predicting EC, OC and TC could be complemented by identifying other species present in the filter. The same Raman measurement including a larger spectral range could reveal the bands from other potential analytes, such as silicon oxide. The project would require investigate the relevance of potential analytes and the standard technique used to identify them. The strong fluorescent background that could overlap the Raman signal is a possible limitation. It's known from the literature the organic carbon is more reactive and change structure with time, reducing its fluorescence. It's also known the samples have to travel from the workplace to the lab for the measurements. A larger spectral range, more measurements and non-linear models are examples of possible features to be changed for improving the accuracy of the method. Real time measurements using Raman spectroscopy have been shown in the literature, but the equipment is not commercial available. Thus, an investigation about the capability of portable Raman would be interesting for easy of implementation.

Chapter 4 shows the potential of automated classification of carbon materials. One possible future step is adding other materials to the model. For instance, carbon nanomaterials have been receiving significant attention from the scientific community. However, the 2D band region (2400 - 2900 cm^{-1}) is an important source of information and may require measurements with larger spectral window or data fusion. Another potential spectral region of interest is at 42 cm^{-1}, this region has a band associated with the shear mode of graphite layers (C-band). This band could be used to estimate the number of graphene layers. Argon atmosphere would be necessary to avoid the overlap of the rotational spectrum of nitrogen. Other non-linear models could improve the performance, while interpretable methods could offer more insights about the process of classification.

Appendix A

Appendix A is published as Supporting information to: B.G. daFonseca; S.S. Thind and A.G. Brolo, Raman maps reveal heterogeneous hydrogenation on carbon materials, J. Raman Spectrosc. 2021, 52, 516-524. Copyright: Reproduced with permission.

A.1 Baseline and Fitting.

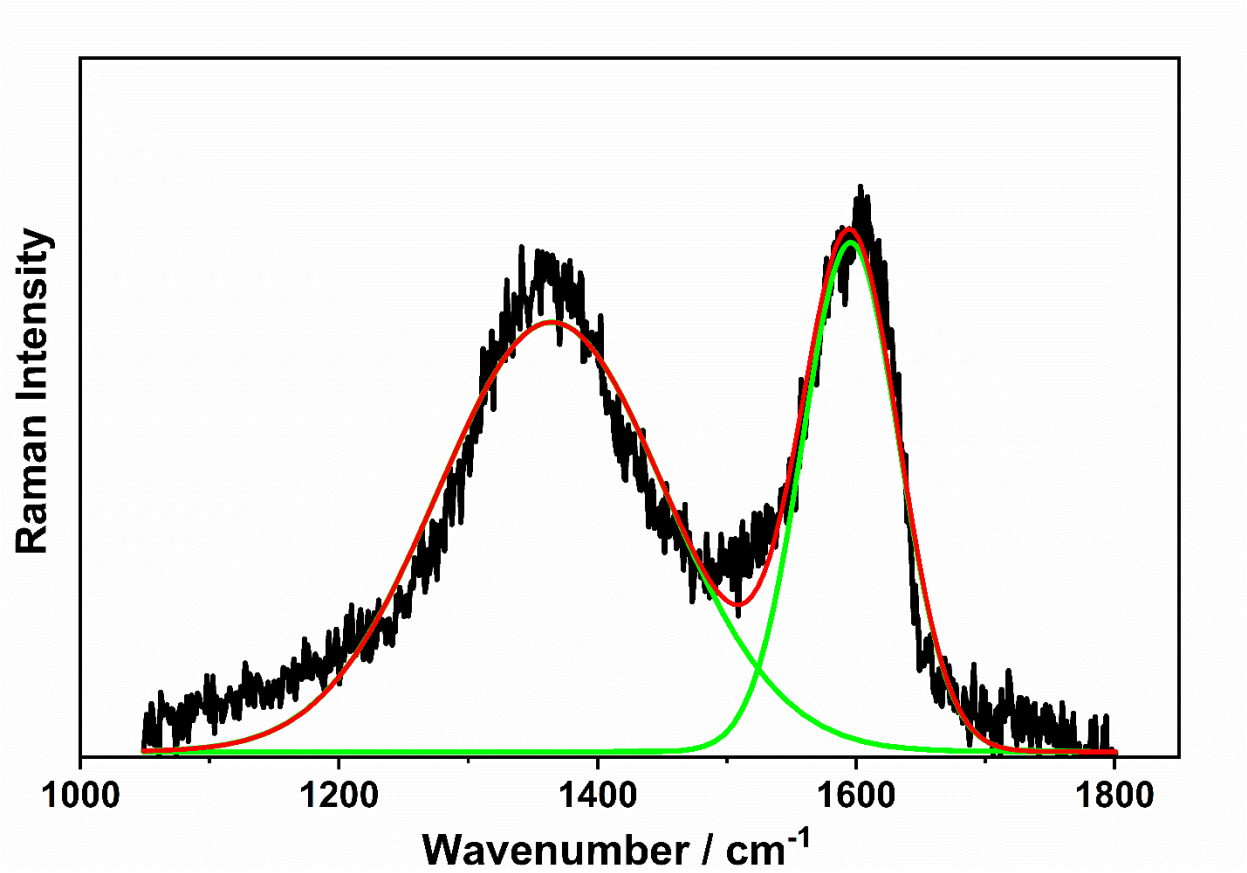

Figure A-1: Raman spectrum from sample HCl (1h) after baseline correction and deconvolution using two Gaussians.

A.2 Raman spectrum including 2D-band.

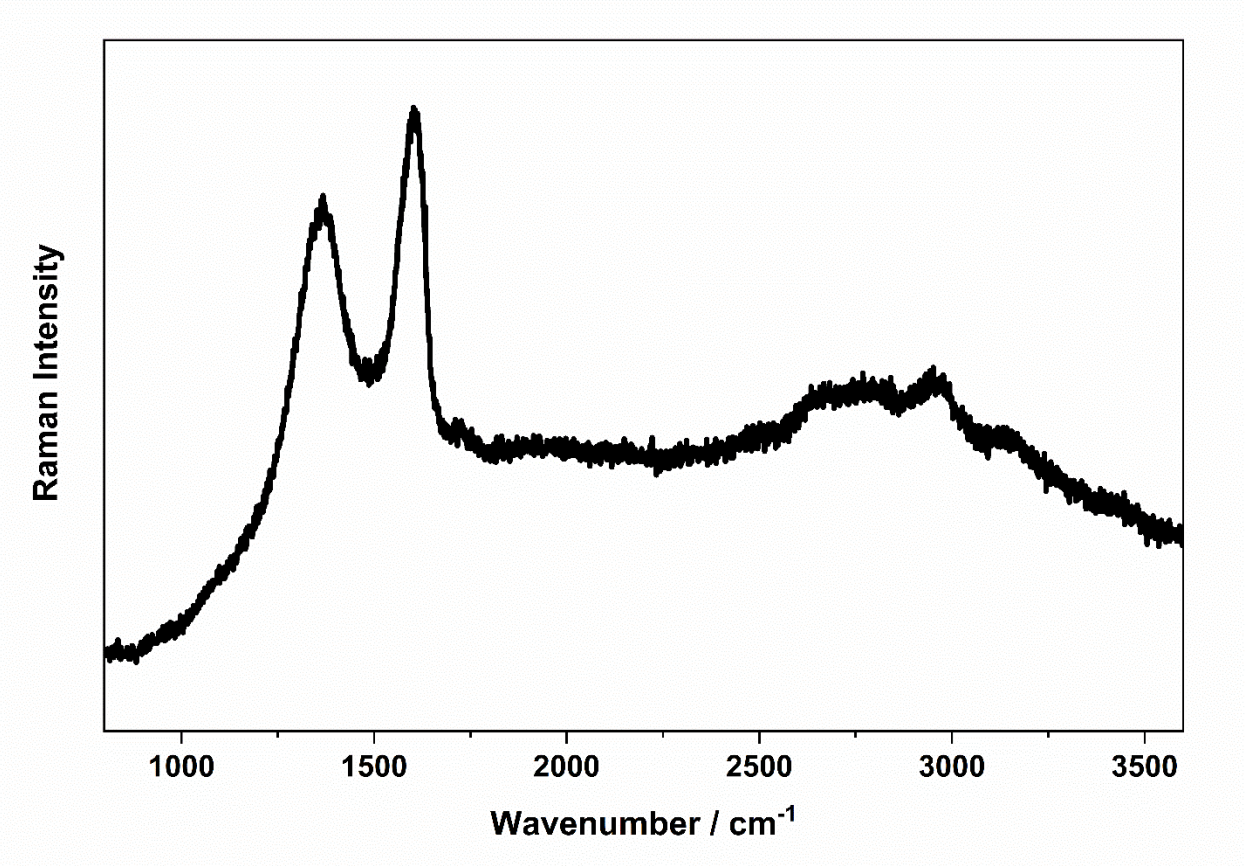

Figure A-2: Raman spectrum of the sample HCl (1h) including the 2D-band.

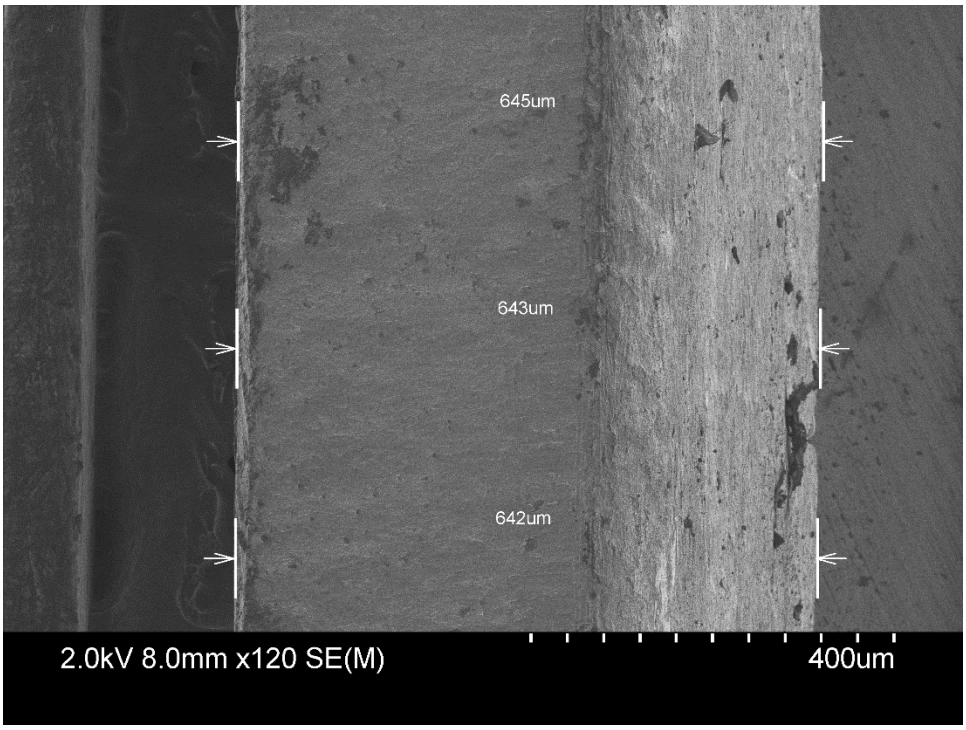

Figure A-3: SEM image of titanium electrode not coated.

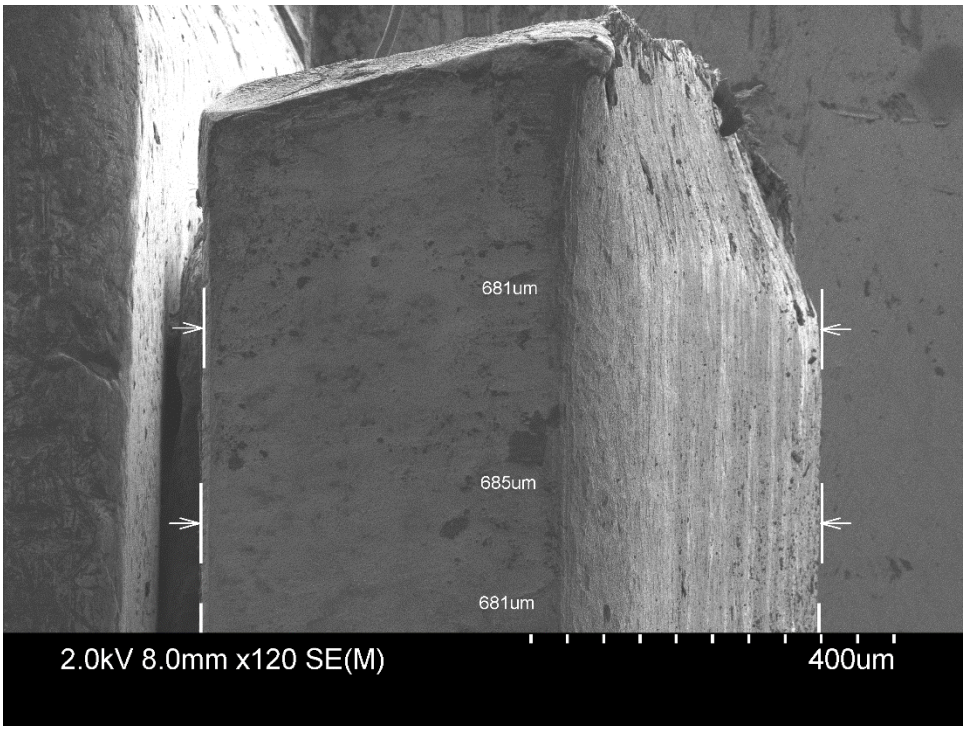

Figure A-4: SEM image of titanium electrode coated with carbon film for 1h. Synthesis HCl(1h).

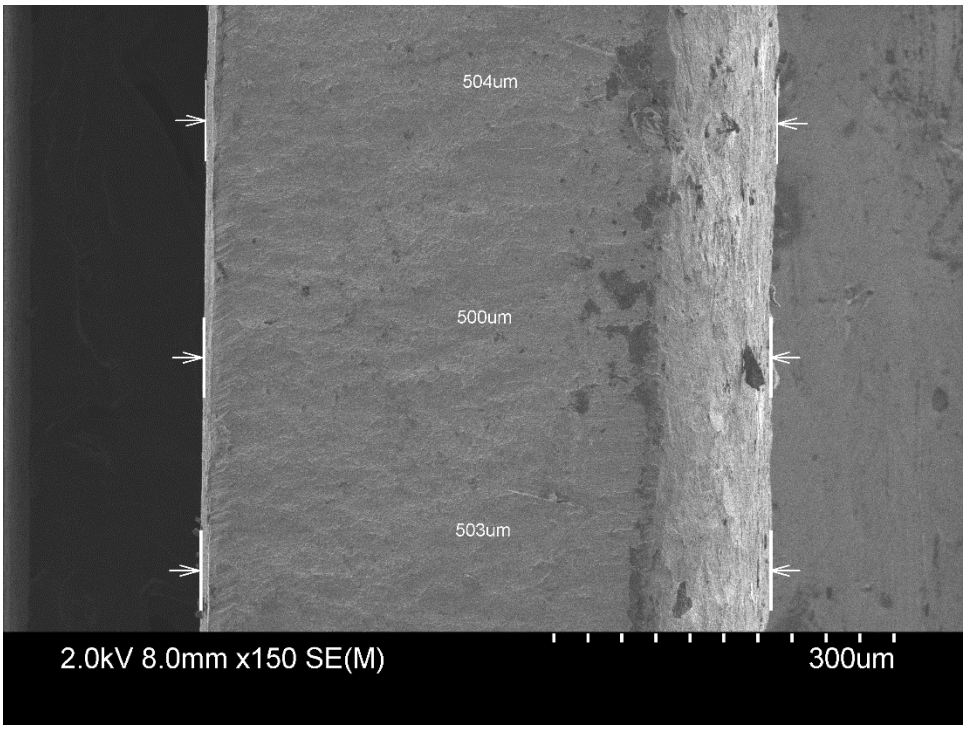

Figure A 5: SEM image of titanium electrode not coated.

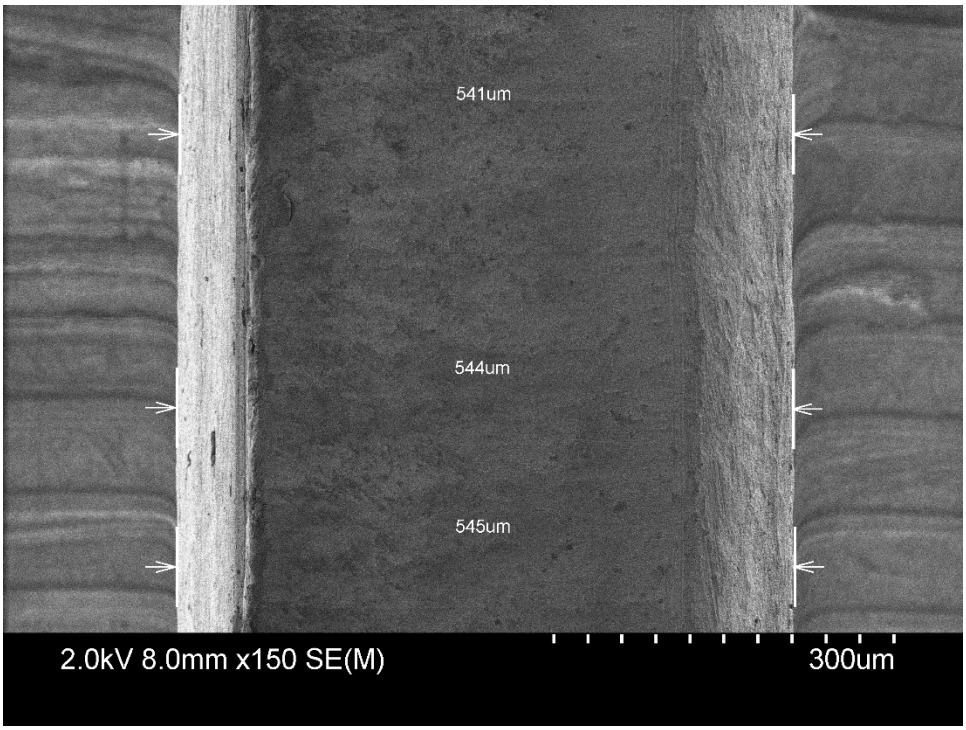

Figure A-6: SEM image of titanium electrode coated with carbon film for 2 h. Synthesis HCl(2h).

Table B- 1: Mean values and standard deviation of G-band position (Pos(G)), G-band full width at half maximum (FWHM(G)), intensity ratio of D- and G-bands (I(D)/I(G)), hydrogen content (H[at.%]) for respective maps.

Sample	Pos(G)	FWHM(G)	I(D)/I(G)	H[at.%]
HCl 1h (1)	1595.0 ± 0.6	91.6 ± 2.1	0.84 ± 0.02	34.2 ± 2.9
HCl 1h (2)	1594.5 ± 0.6	91.0 ± 2.7	0.88 ± 0.02	32.5 ± 2.5
HCl 1h (3)	1594.1 ± 0.5	91.8 ± 1.4	0.91 ± 0.01	32.0 ± 1
HCl 1h (4)	1594.5 ± 0.6	91.4 ± 1.5	0.88 ± 0.01	33.7 ± 1.9
HCl 1h (5)	1593.0 ± 0.7	93.1 ± 2.3	0.90 ± 0.02	33.1 ± 2.6
HCl 2h (1)	1594.9 ± 0.3	90.5 ± 0.8	0.90 ± 0.01	32 ± 1.1
HCl 2h (2)	1595.3 ± 0.2	90.2 ± 0.9	0.89 ± 0.01	31.9 ± 0.6
HCl 2h (3)	1595.2 ± 0.3	90.6 ±0.8	0.89 ± 0.01	32.0 ± 0.5
HCl 2h (4)	1595.3 ± 0.3	90.6 ± 0.9	0.88 ± 0.01	32.6 ± 0.6
HCl 2h (5)	1595.2 ± 0.3	90.5 ± 0.9	0.90 ± 0.01	31.1 ± 0.8
HNO_3 1h (1)	1595.1 ± 0.8	87.5 ± 2.4	0.82 ± 0.02	29.3 ± 1.6
HNO_3 1h (2)	1594.5 ± 0.8	88.9 ± 1.9	0.83 ± 0.02	29.1 ± 2.2
HNO_3 1h (3)	1594.2 ± 0.7	88.7 ± 2.1	0.82 ± 0.02	29.8 ± 1.9
HNO_3 1h (4)	1595.3 ± 0.8	88.0 ± 3.0	0.82 ± 0.02	30.5 ± 1.8
HNO_3 1h (5)	28.9 ± 1.3	88.7 ± 1.8	0.82 ± 0.01	1594.8 ± 0.5
HNO_3 2h (1)	1593.9 ± 0.3	91.0 ± 0.8	0.84 0.01	27.9 ± 0.9
HNO_3 2h (2)	1593.8 ± 0.4	91.1 ± 1.1	0.84 0.01	28.1 ± 1.1
HNO_3 2h (3)	1593.6 ± 0.4	91.5 ± 0.7	0.84 0.01	28.4 ± 0.7
HNO_3 2h (4)	1593.9 ± 0.3	90.7 ± 0.9	0.84 0.01	28.6 ± 0.9
HNO_3 2h (5)	1593.0 ± 0.3	92.0 ± 0.8	0.85 0.01	29.5 ± 0.7

A.3 Sample HCl (1h)

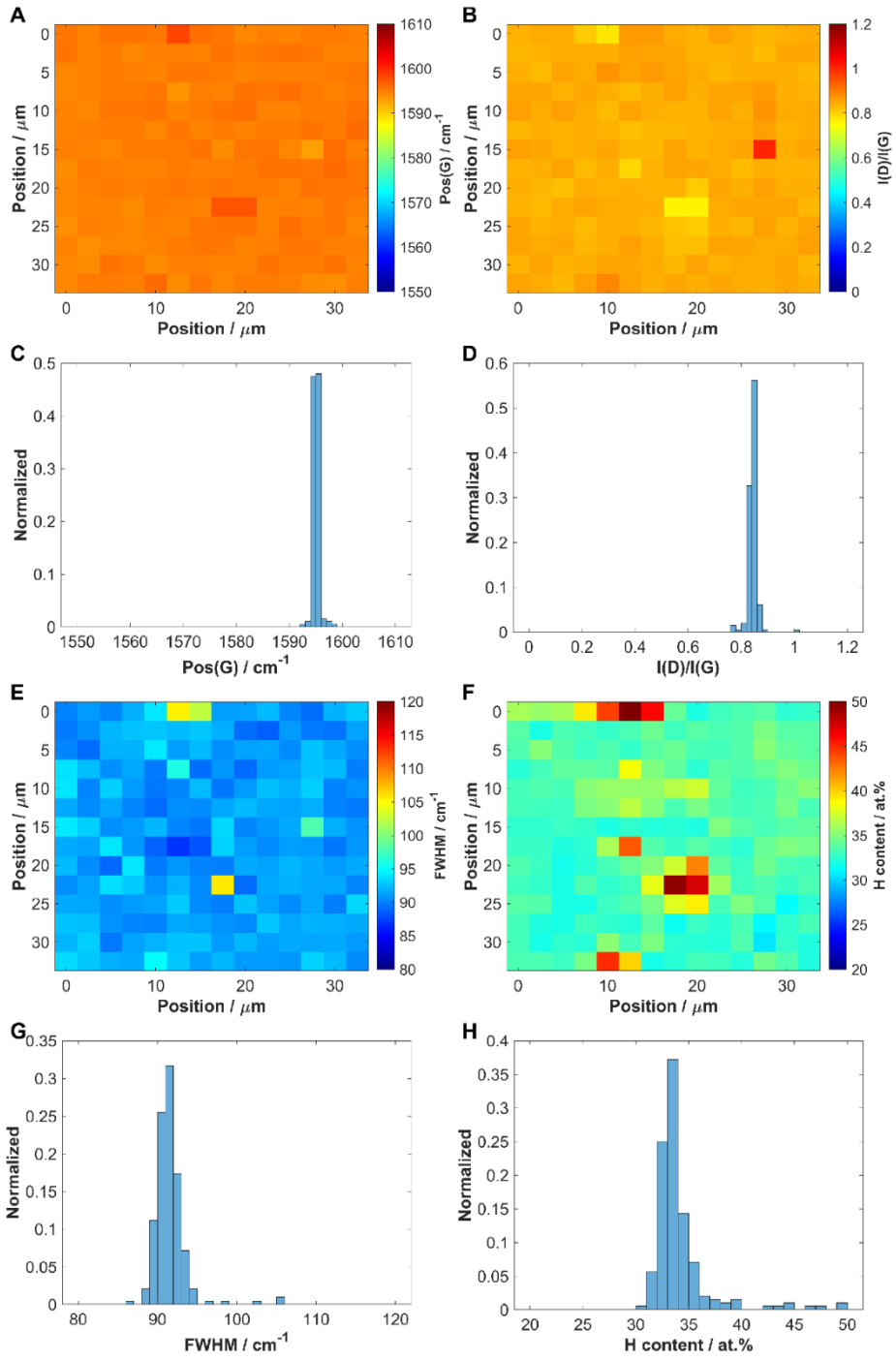

Figure A-7: Raman maps and respective histograms for region 1 of sample HCl(1h) comparing different features. A) G-band position map. B) Intensity ratio of D- and G-bands map. C) G-band position histogram. D) Intensity ratio of D- and G-bands histogram. E) G-band full width at half maximum map. F) Hydrogen content map. G) G-band full width at half maximum histogram. H) Hydrogen content histogram.

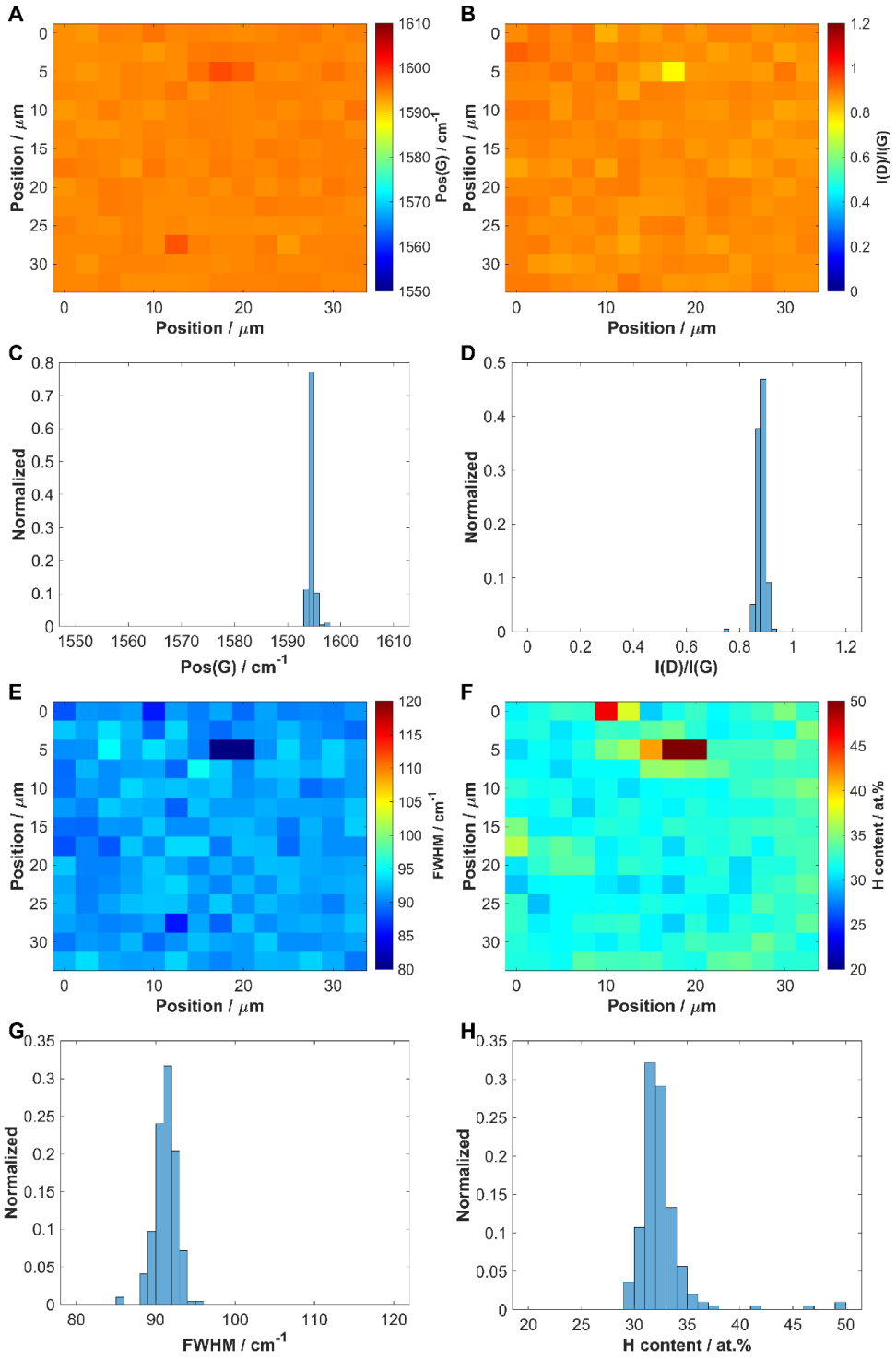

Figure A-8: Raman maps and respective histograms for region 2 of sample HCl(1h) comparing different features. A) G-band position map. B) Intensity ratio of D- and G-bands map. C) G-band position histogram. D) Intensity ratio of D- and G-bands histogram. E) G-band full width at half maximum map. F) Hydrogen content map. G) G-band full width at half maximum histogram. H) Hydrogen content histogram.

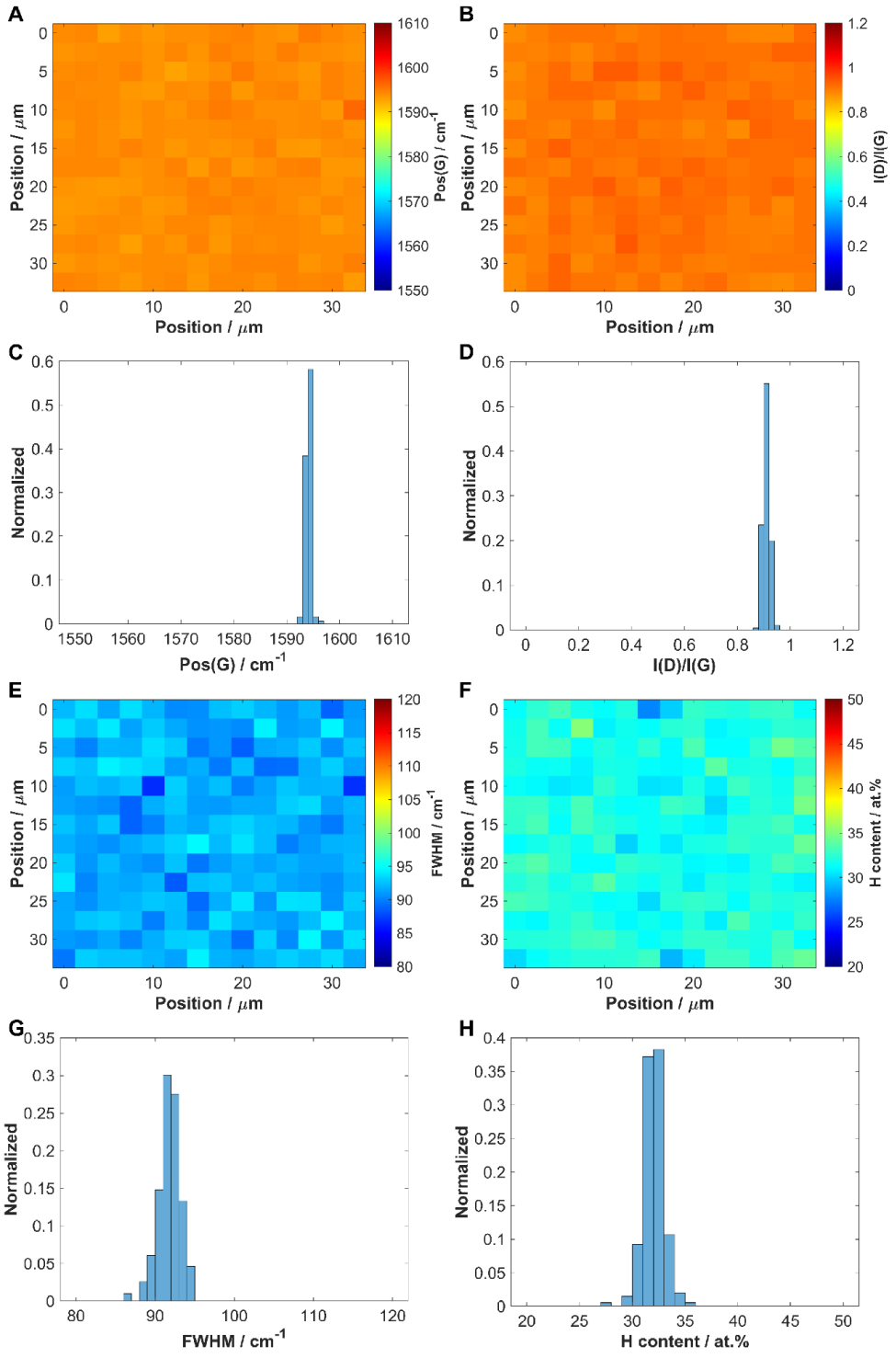

Figure A-9: Raman maps and respective histograms for region 3 of sample HCl(1h) comparing different features. A) G-band position map. B) Intensity ratio of D- and G-bands map. C) G-band position histogram. D) Intensity ratio of D- and G-bands histogram. E) G-band full width at half maximum map. F) Hydrogen content map. G) G-band full width at half maximum histogram. H) Hydrogen content histogram.

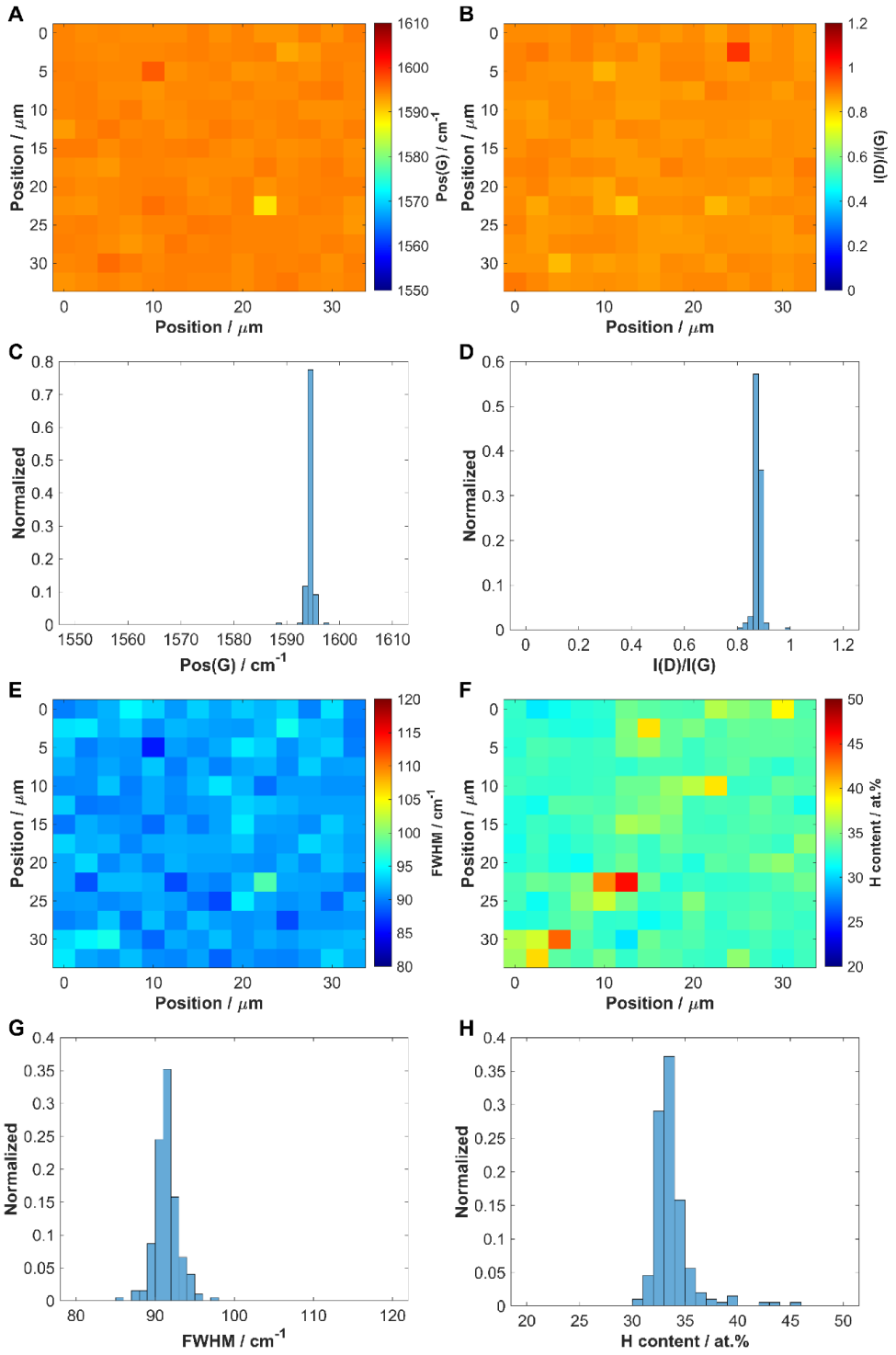

Figure A-10: Raman maps and respective histograms for region 4 of sample HCl(1h) comparing different features. A) G-band position map. B) Intensity ratio of D- and G-bands map. C) G-band position histogram. D) Intensity ratio of D- and G-bands histogram. E) G-band full width at half maximum map. F) Hydrogen content map. G) G-band full width at half maximum histogram. H) Hydrogen content histogram.

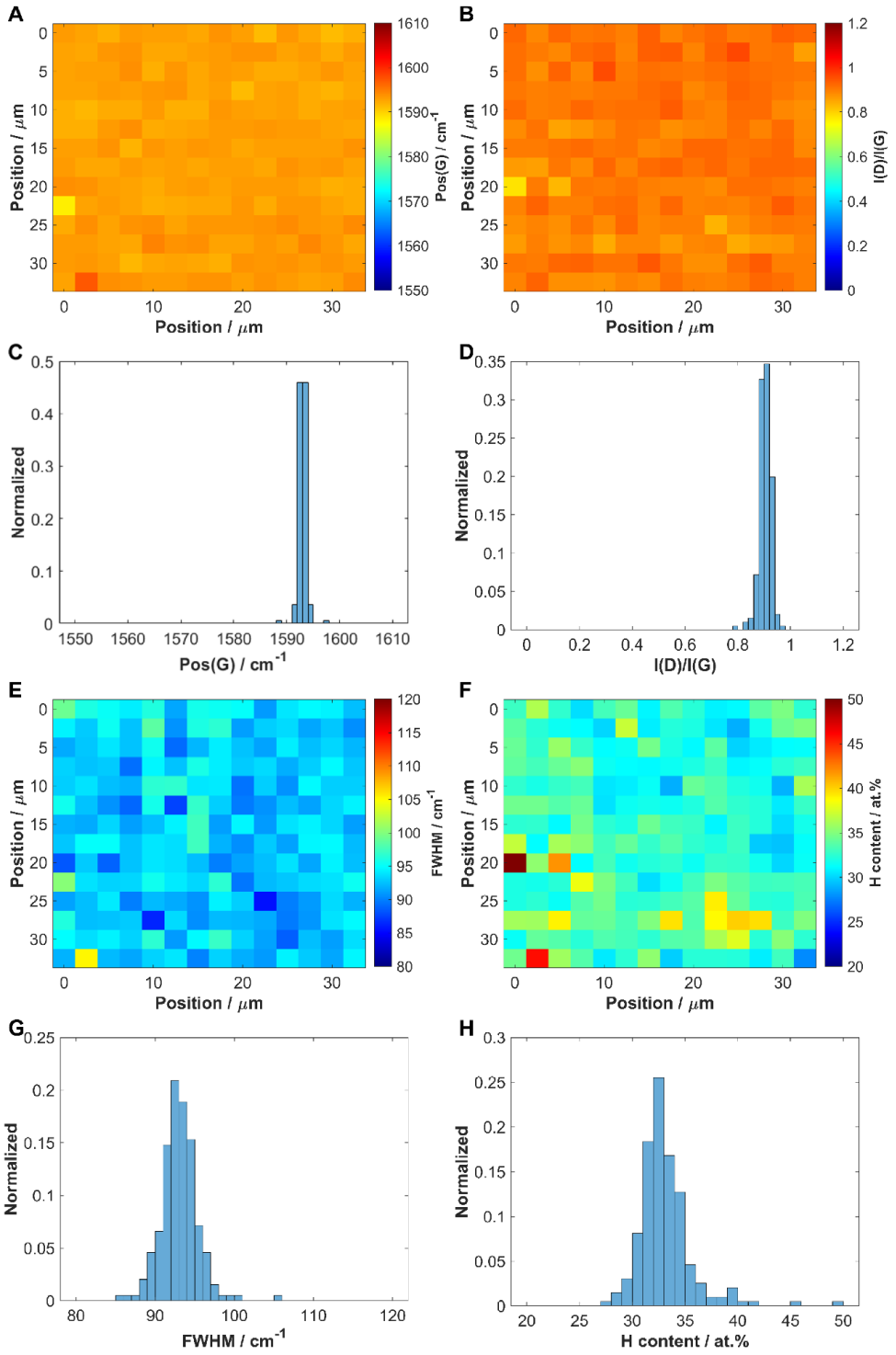

Figure A-11: Raman maps and respective histograms for region 5 of sample HCl(1h) comparing different features. A) G-band position map. B) Intensity ratio of D- and G-bands map. C) G-band position histogram. D) Intensity ratio of D- and G-bands histogram. E) G-band full width at half maximum map. F) Hydrogen content map. G) G-band full width at half maximum histogram. H) Hydrogen content histogram.

A.4 Sample HCl(2h)

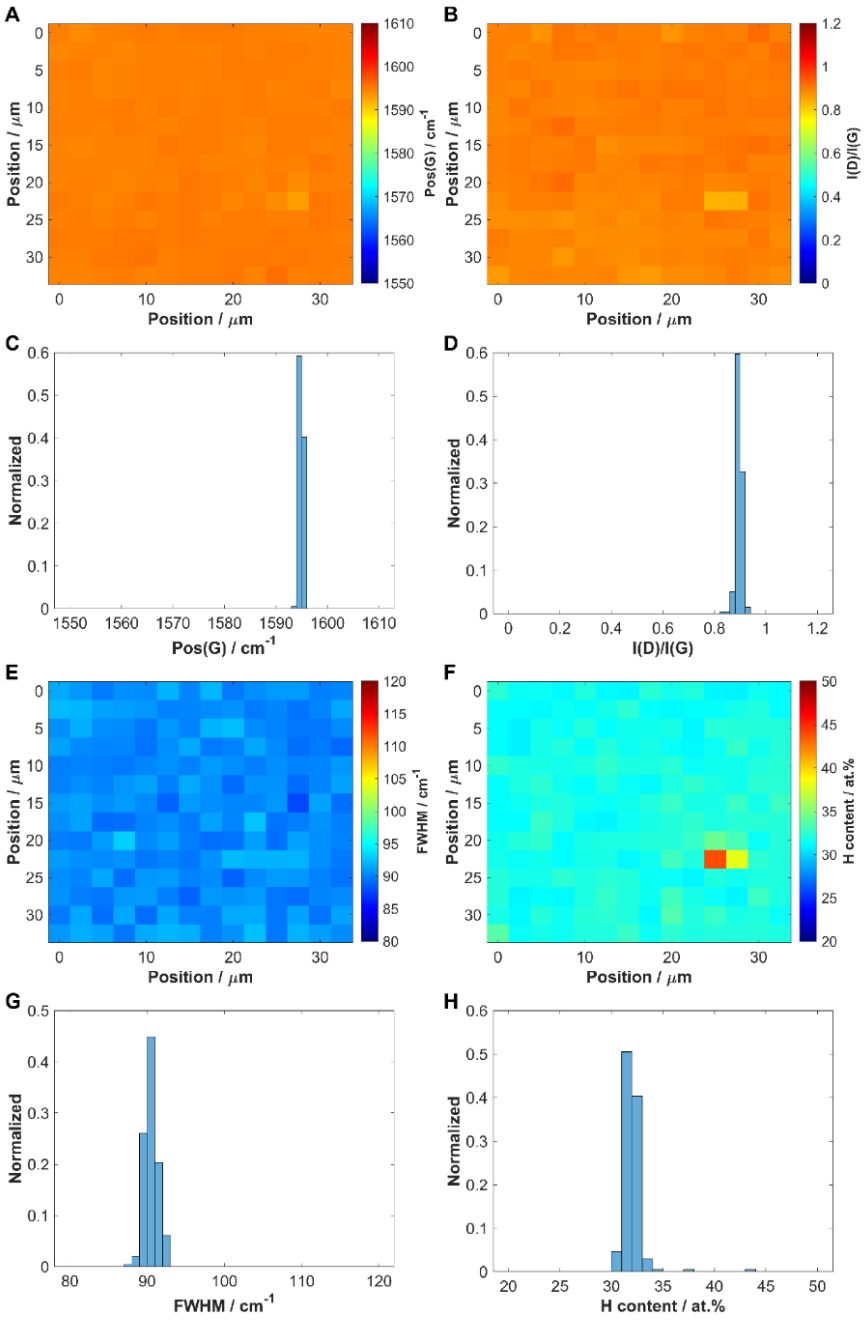

Figure A-12: Raman maps and respective histograms for region 1 of sample HCl(2h) comparing different features. A) G-band position map. B) Intensity ratio of D- and G-bands map. C) G-band position histogram. D) Intensity ratio of D- and G-bands histogram. E) G-band full width at half maximum map. F) Hydrogen content map. G) G-band full width at half maximum histogram. H) Hydrogen content histogram.

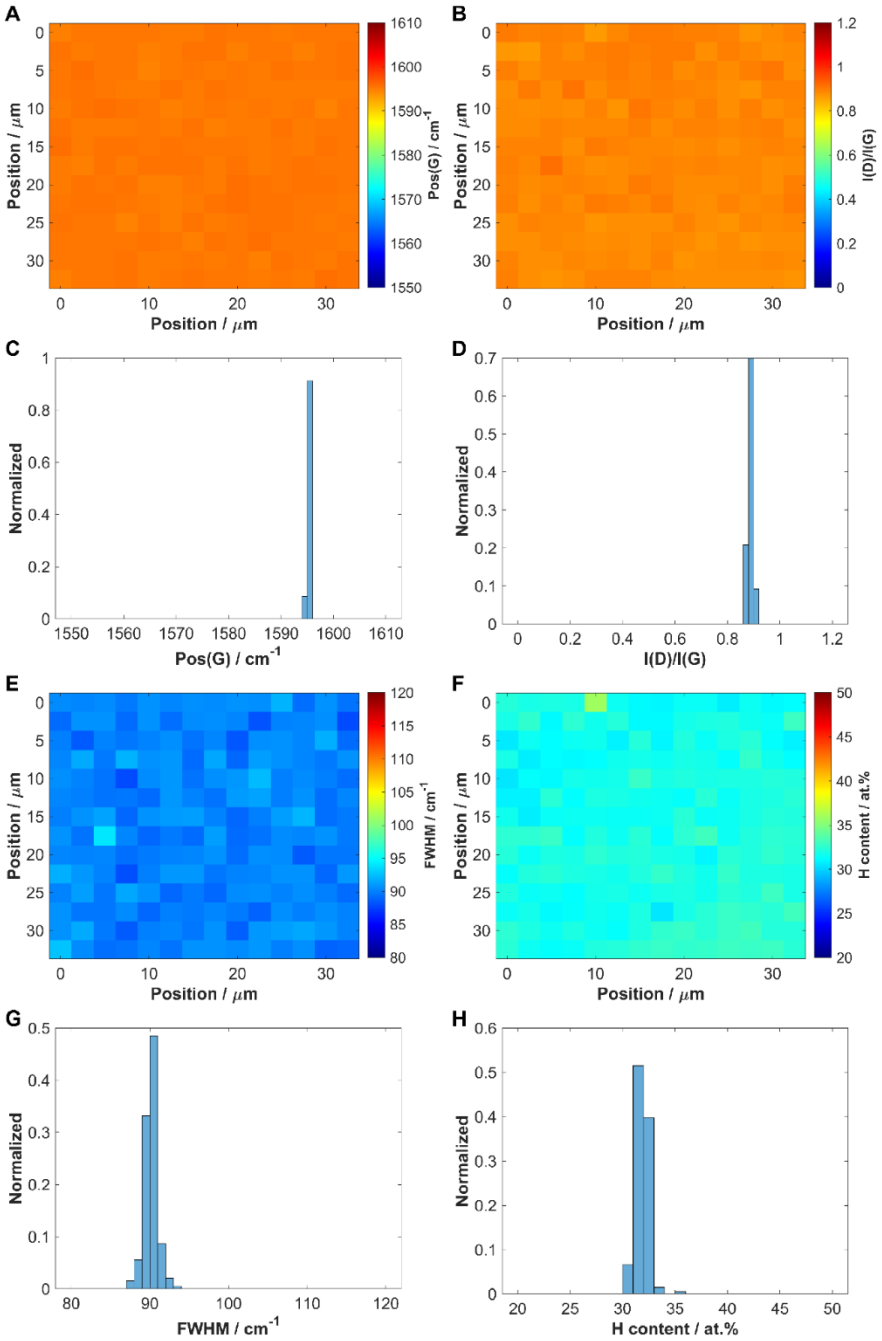

Figure A-13: Raman maps and respective histograms for region 2 of sample HCl(2h) comparing different features. A) G-band position map. B) Intensity ratio of D- and G-bands map. C) G-band position histogram. D) Intensity ratio of D- and G-bands histogram. E) G-band full width at half maximum map. F) Hydrogen content map. G) G-band full width at half maximum histogram. H) Hydrogen content histogram.

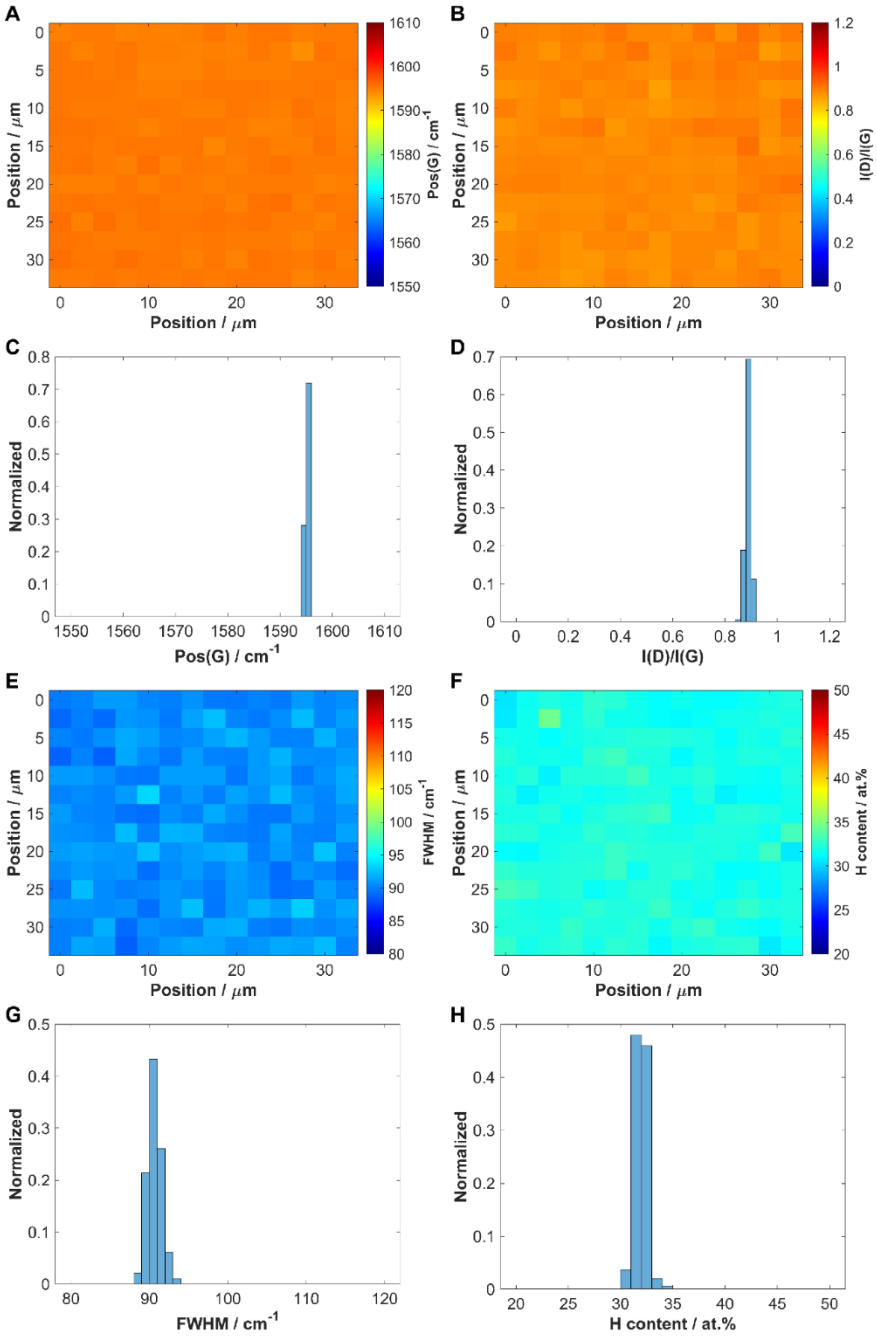

Figure A-14: Raman maps and respective histograms for region 3 of sample HCl(2h) comparing different features. A) G-band position map. B) Intensity ratio of D- and G-bands map. C) G-band position histogram. D) Intensity ratio of D- and G-bands histogram. E) G-band full width at half maximum map. F) Hydrogen content map. G) G-band full width at half maximum histogram. H) Hydrogen content histogram.

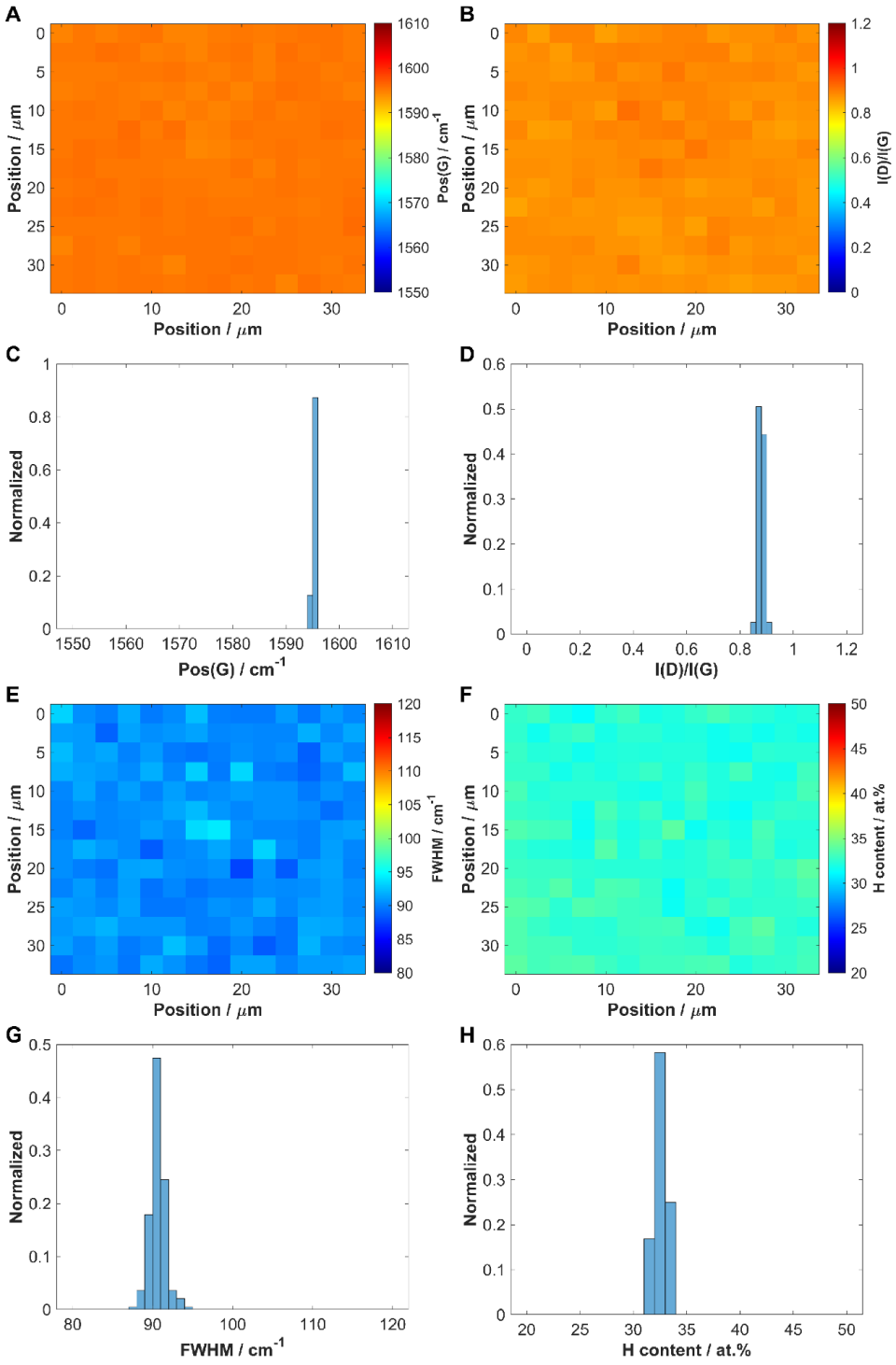

Figure A-15: Raman maps and respective histograms for region 4 of sample HCl(2h) comparing different features. A) G-band position map. B) Intensity ratio of D- and G-bands map. C) G-band position histogram. D) Intensity ratio of D- and G-bands histogram. E) G-band full width at half maximum map. F) Hydrogen content map. G) G-band full width at half maximum histogram. H) Hydrogen content histogram.

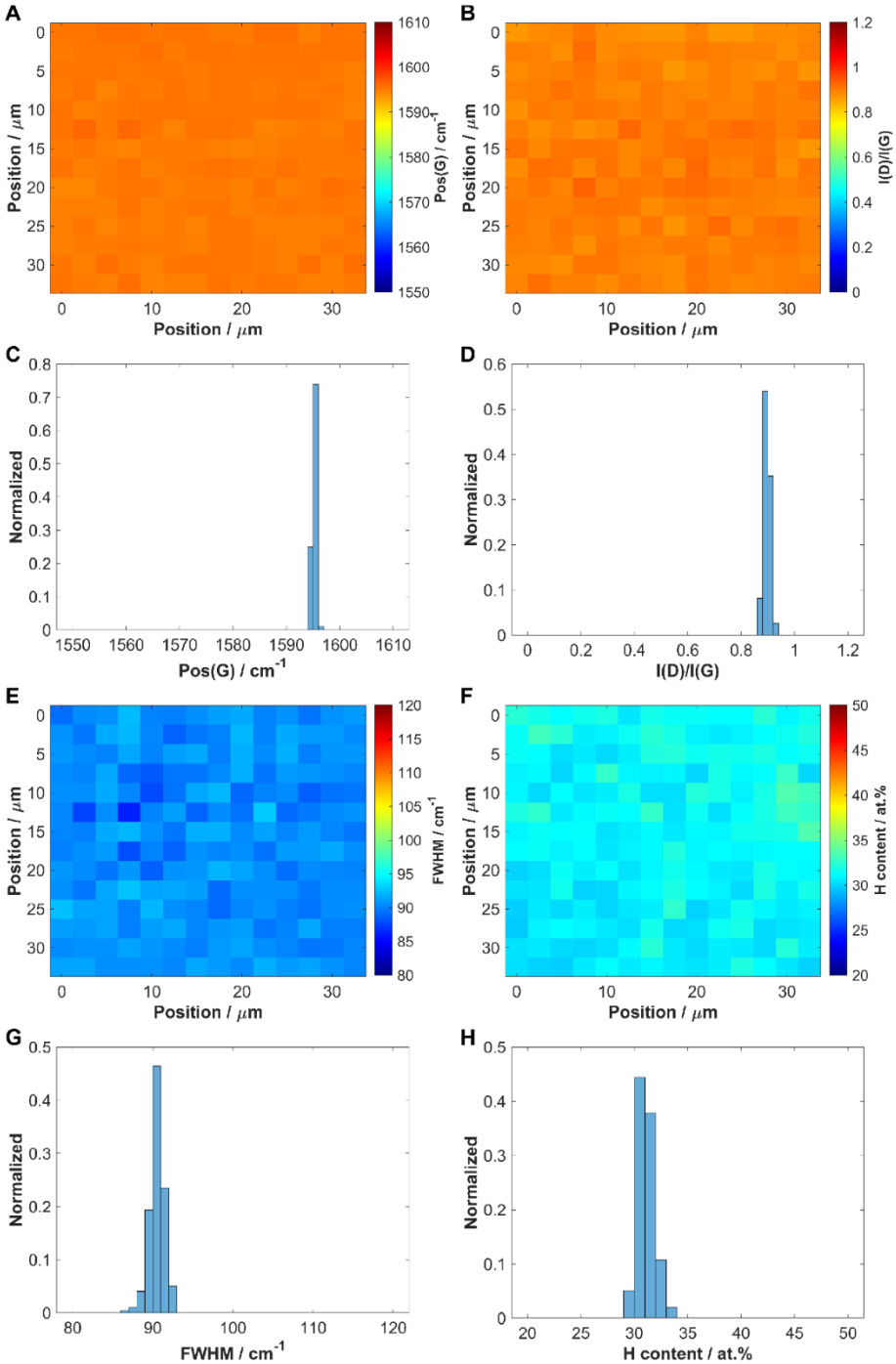

Figure A-16: Raman maps and respective histograms for region 5 of sample HCl(2h) comparing different features. A) G-band position map. B) Intensity ratio of D- and G-bands map. C) G-band position histogram. D) Intensity ratio of D- and G-bands histogram. E) G-band full width at half maximum map. F) Hydrogen content map. G) G-band full width at half maximum histogram. H) Hydrogen content histogram.

A.5 Sample HNO_3(1h)

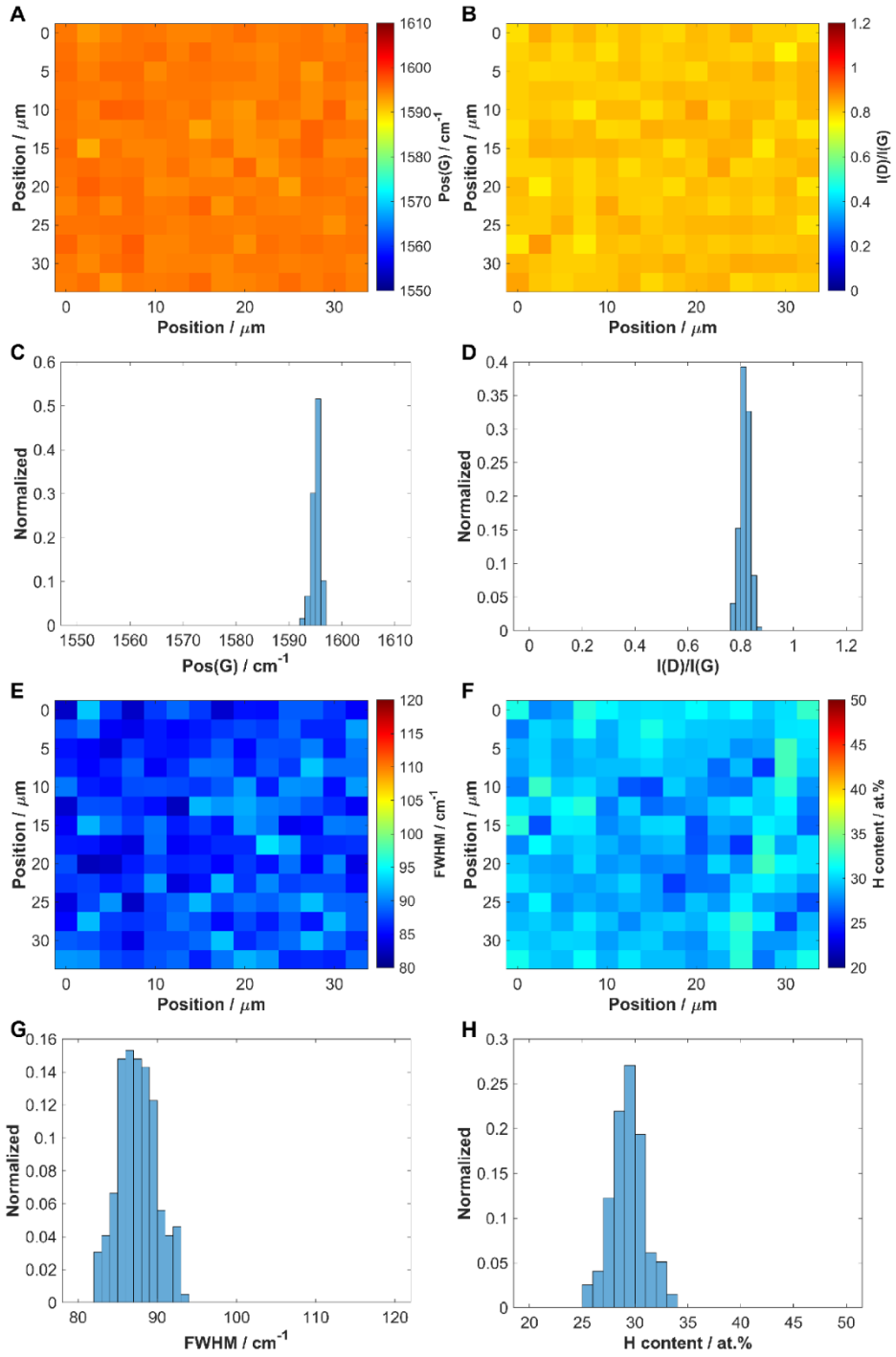

Figure A-17: Raman maps and respective histograms for region 1 of sample HNO3(1h) comparing different features. A) G-band position map. B) Intensity ratio of D- and G-bands map. C) G-band position histogram. D) Intensity ratio of D- and G-bands histogram. E) G-band full width at half maximum map. F) Hydrogen content map. G) G-band full width at half maximum histogram. H) Hydrogen content histogram.

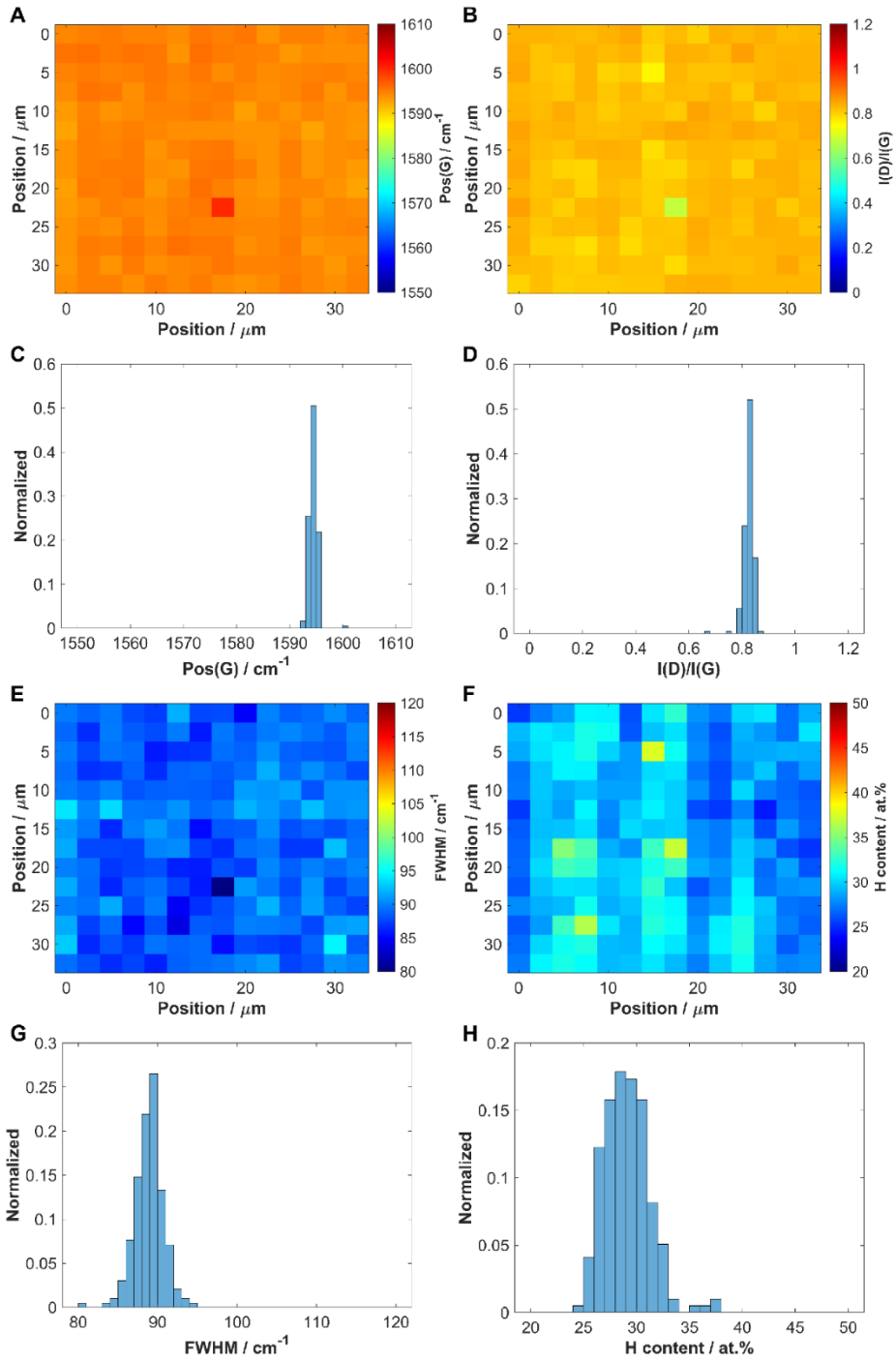

Figure A-18: Raman maps and respective histograms for region 2 of sample HNO_3(1h) comparing different features. A) G-band position map. B) Intensity ratio of D- and G-bands map. C) G-band position histogram. D) Intensity ratio of D- and G-bands histogram. E) G-band full width at half maximum map. F) Hydrogen content map. G) G-band full width at half maximum histogram. H) Hydrogen content histogram.

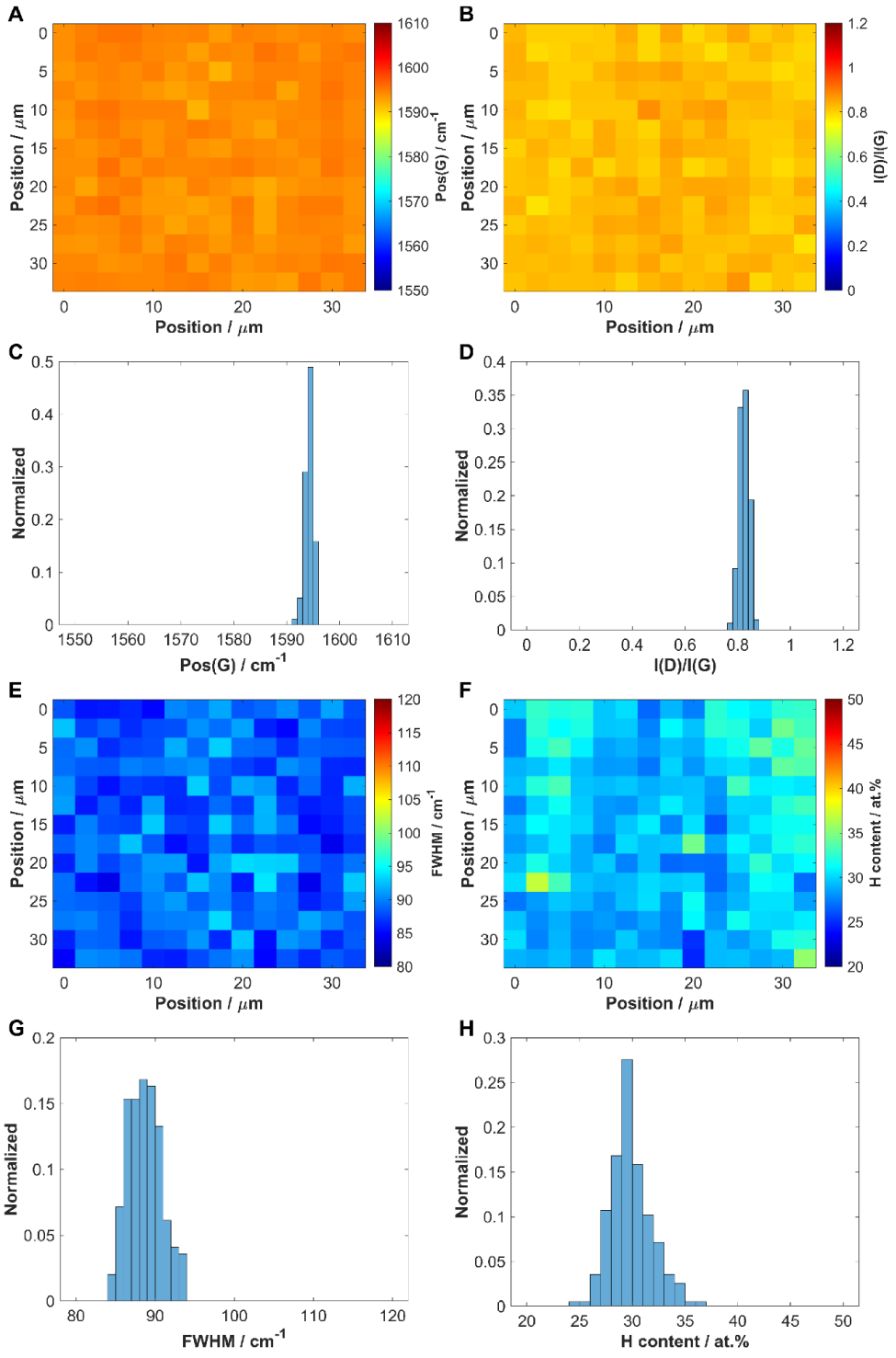

Figure A-19: Raman maps and respective histograms for region 3 of sample HNO_3(1h) comparing different features. A) G-band position map. B) Intensity ratio of D- and G-bands map. C) G-band position histogram. D) Intensity ratio of D- and G-bands histogram. E) G-band full width at half maximum map. F) Hydrogen content map. G) G-band full width at half maximum histogram. H) Hydrogen content histogram.

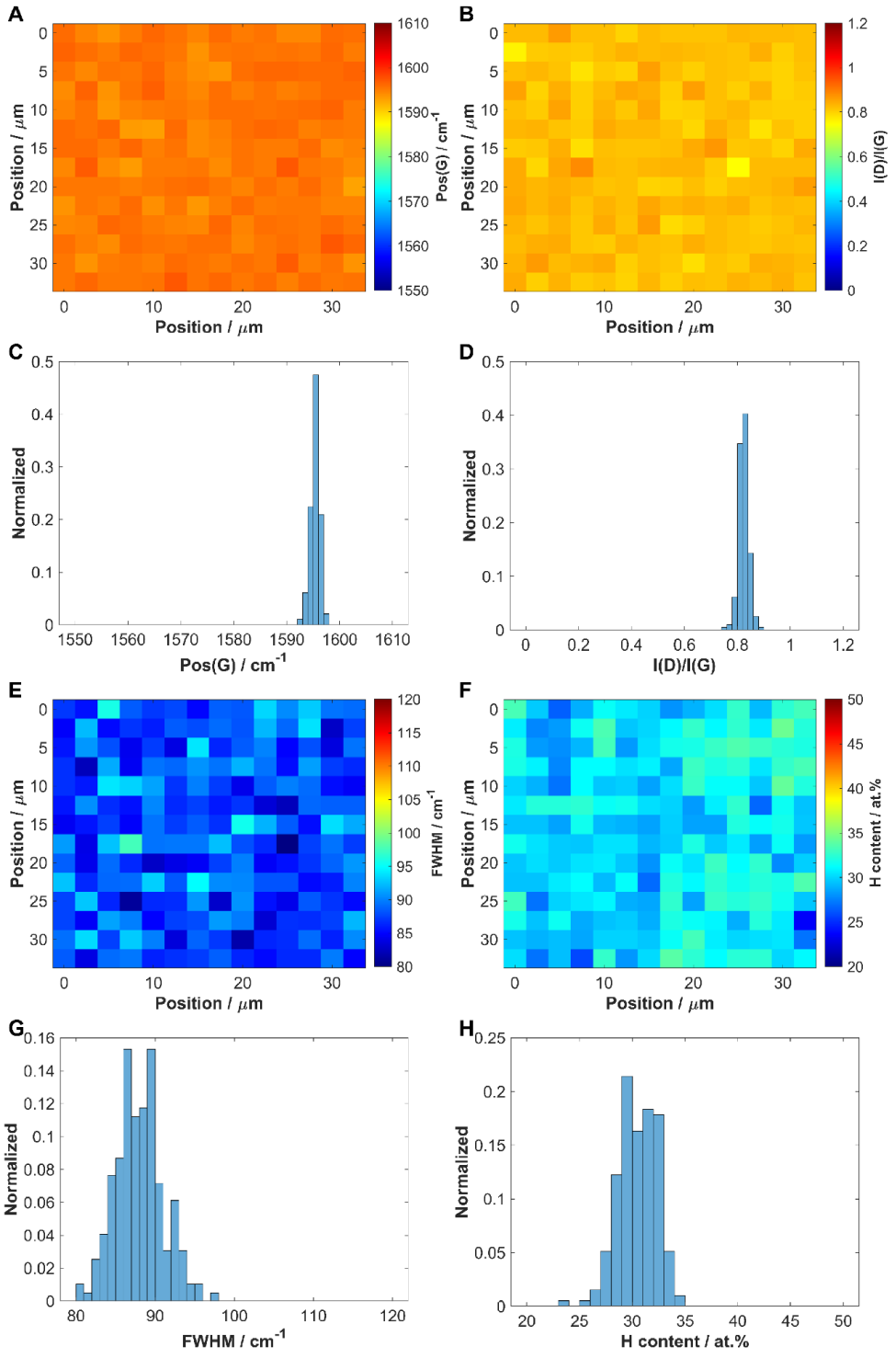

Figure A-20: Raman maps and respective histograms for region 4 of sample HNO_3(1h) comparing different features. A) G-band position map. B) Intensity ratio of D- and G-bands map. C) G-band position histogram. D) Intensity ratio of D- and G-bands histogram. E) G-band full width at half maximum map. F) Hydrogen content map. G) G-band full width at half maximum histogram. H) Hydrogen content histogram.

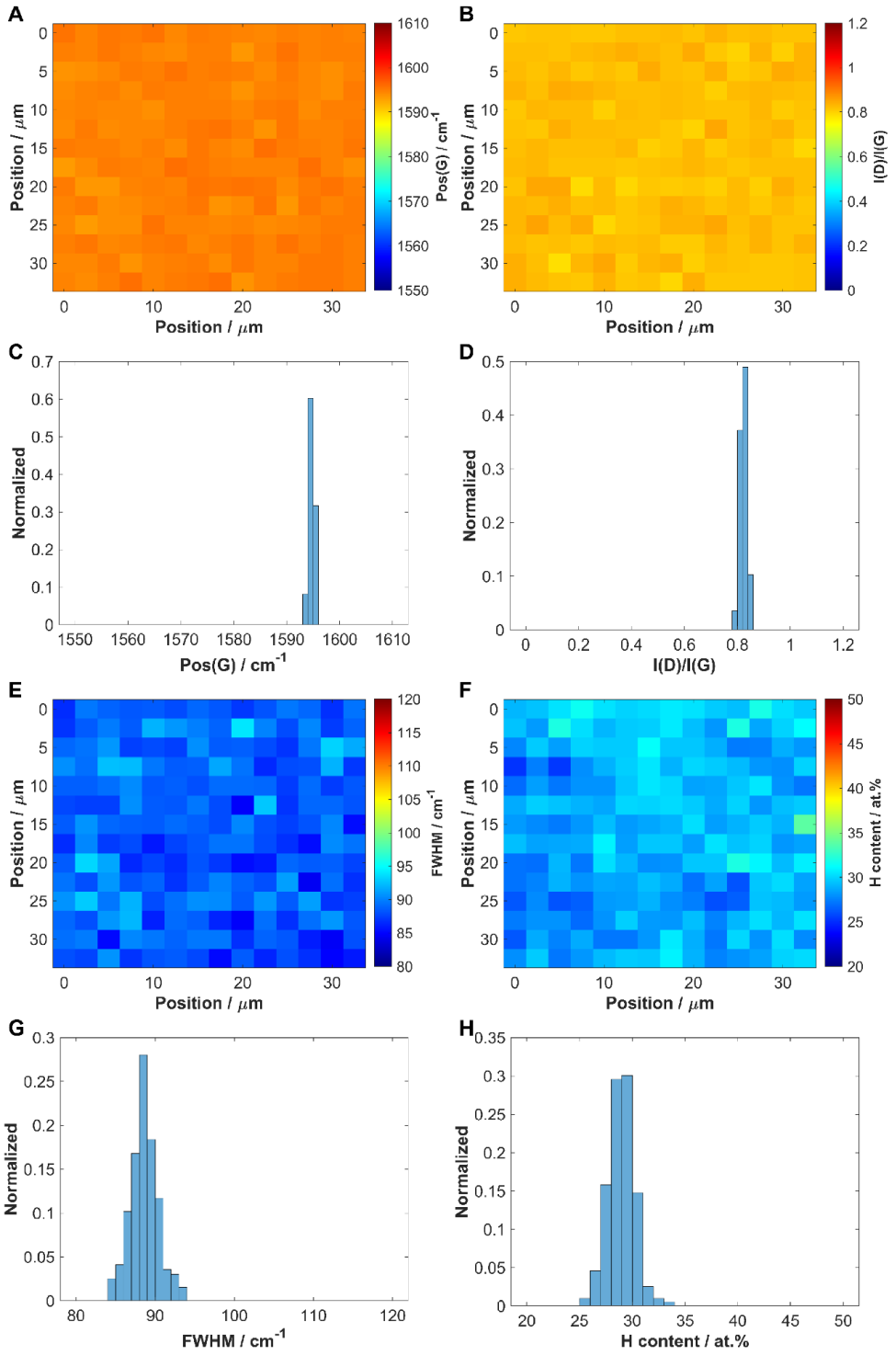

Figure A-21: Raman maps and respective histograms for region 5 of sample HNO_3(1h) comparing different features. A) G-band position map. B) Intensity ratio of D- and G-bands map. C) G-band position histogram. D) Intensity ratio of D- and G-bands histogram. E) G-band full width at half maximum map. F) Hydrogen content map. G) G-band full width at half maximum histogram. H) Hydrogen content histogram.

A.6 Sample HNO_3 (2h)

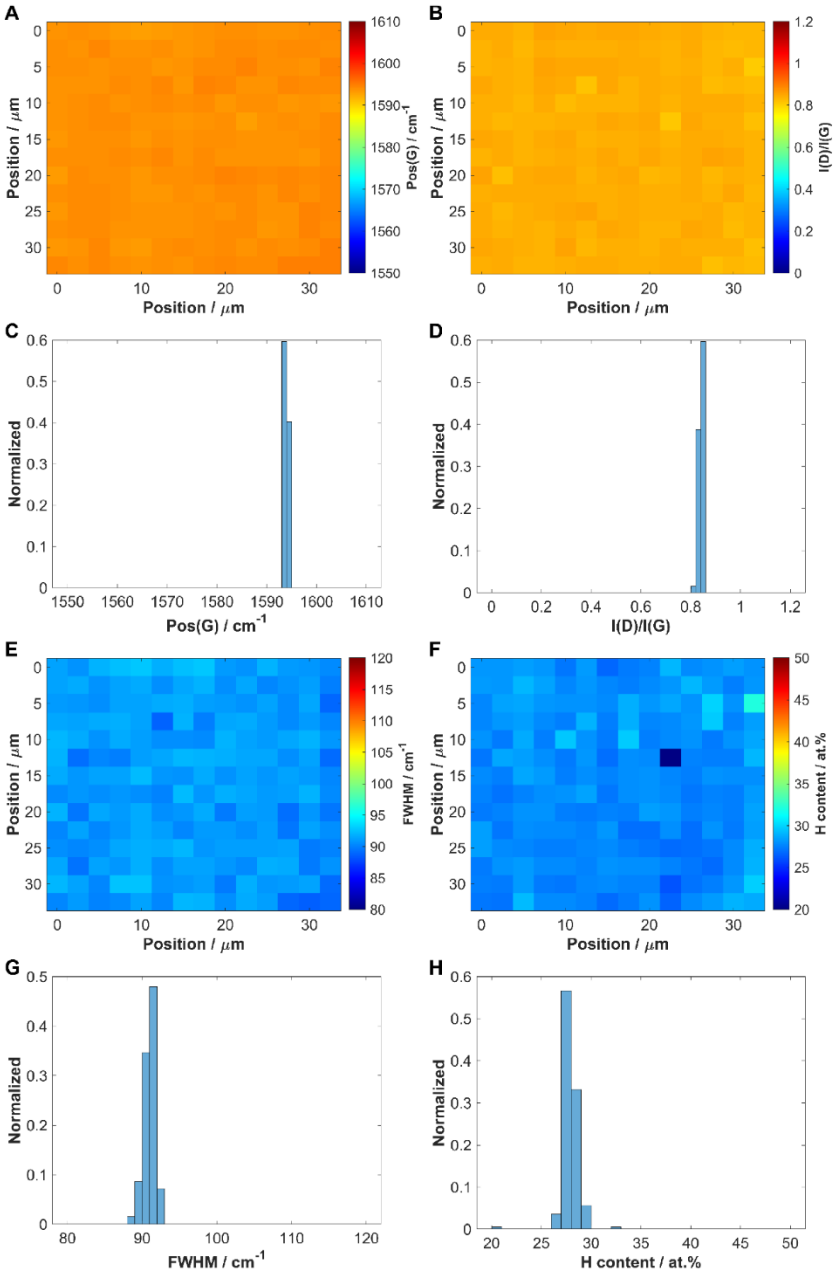

Figure A-22: Raman maps and respective histograms for region 1 of sample HNO_3(2h) comparing different features. A) G-band position map. B) Intensity ratio of D- and G-bands map. C) G-band position histogram. D) Intensity ratio of D- and G-bands histogram. E) G-band full width at half maximum map. F) Hydrogen content map. G) G-band full width at half maximum histogram. H) Hydrogen content histogram.

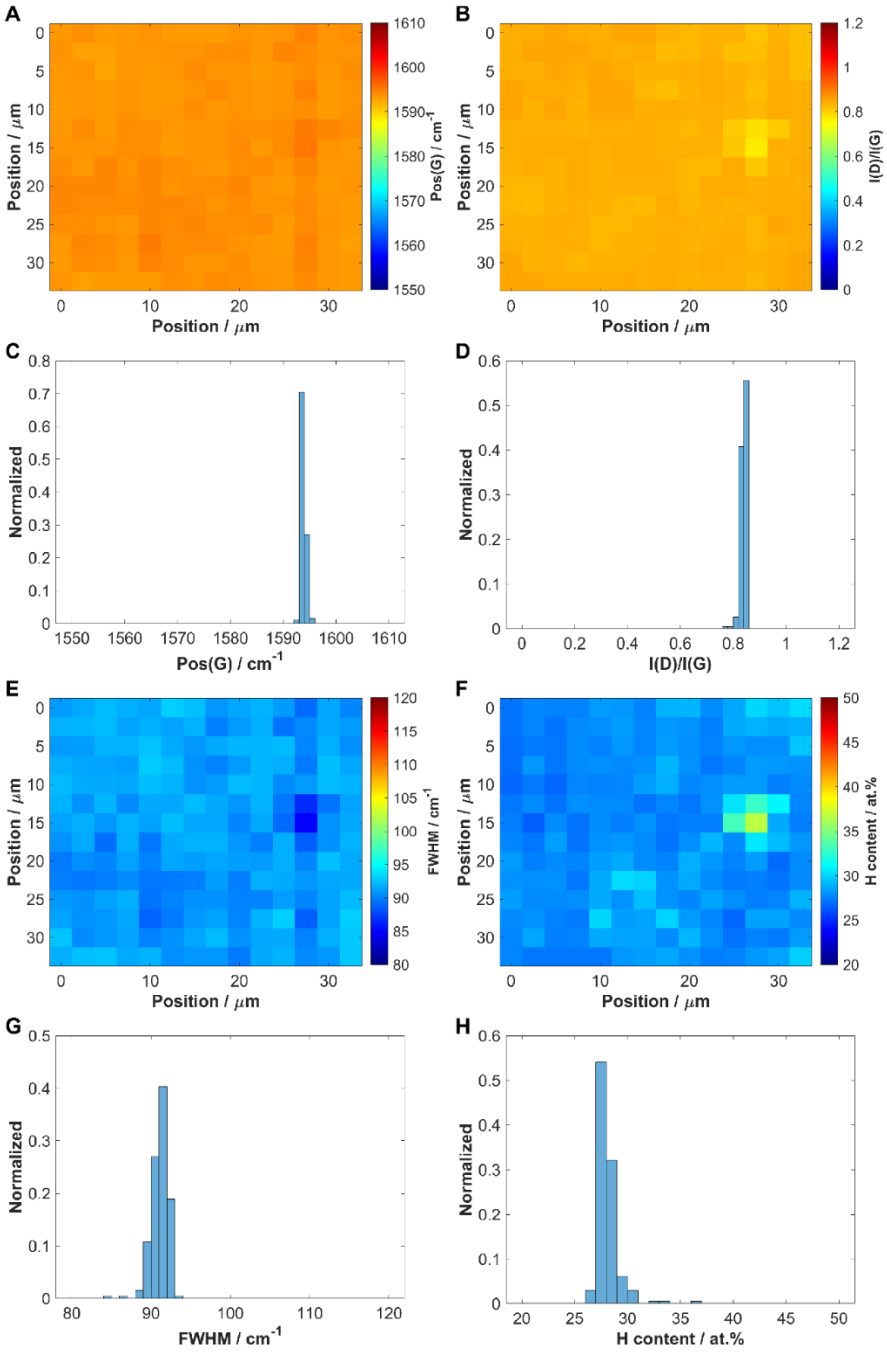

Figure A-23: Raman maps and respective histograms for region 2 of sample HNO_3(2h) comparing different features. A) G-band position map. B) Intensity ratio of D- and G-bands map. C) G-band position histogram. D) Intensity ratio of D- and G-bands histogram. E) G-band full width at half maximum map. F) Hydrogen content map. G) G-band full width at half maximum histogram. H) Hydrogen content histogram.

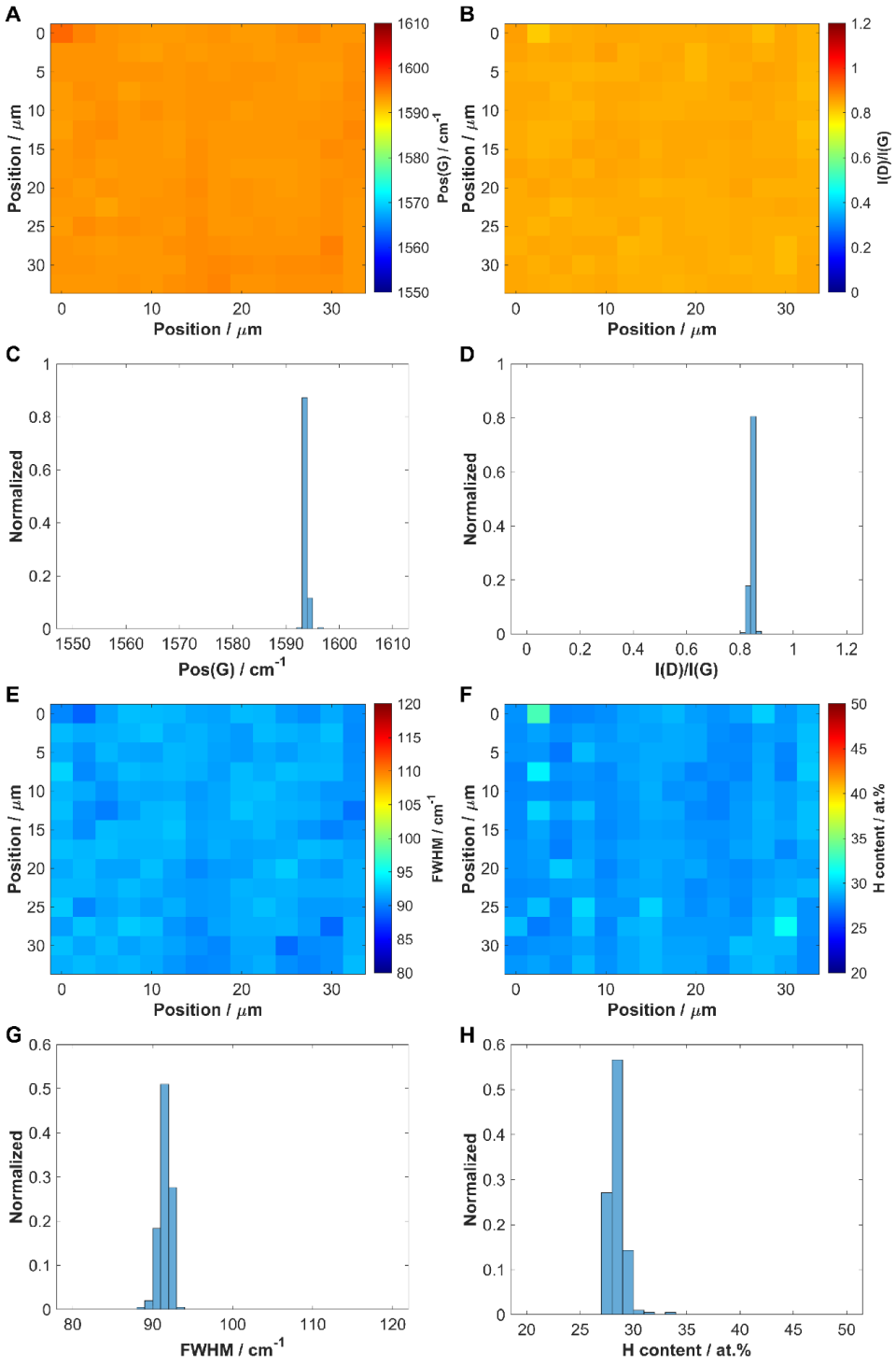

Figure A-24: Raman maps and respective histograms for region 3 of sample HNO_3(2h) comparing different features. A) G-band position map. B) Intensity ratio of D- and G-bands map. C) G-band position histogram. D) Intensity ratio of D- and G-bands histogram. E) G-band full width at half maximum map. F) Hydrogen content map. G) G-band full width at half maximum histogram. H) Hydrogen content histogram.

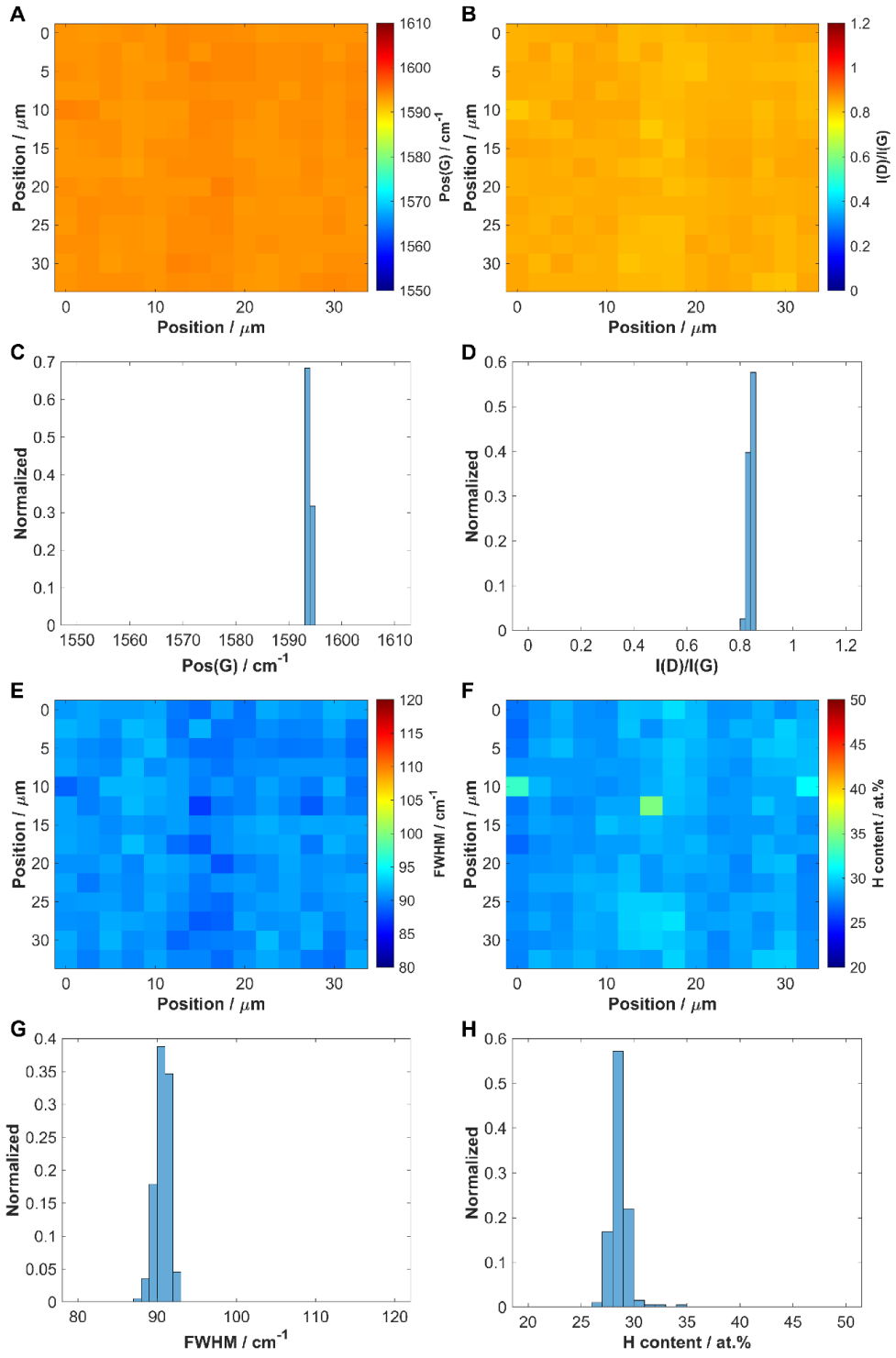

Figure A-25: Raman maps and respective histograms for region 4 of sample HNO_3(2h) comparing different features. A) G-band position map. B) Intensity ratio of D- and G-bands map. C) G-band position histogram. D) Intensity ratio of D- and G-bands histogram. E) G-band full width at half maximum map. F) Hydrogen content map. G) G-band full width at half maximum histogram. H) Hydrogen content histogram.

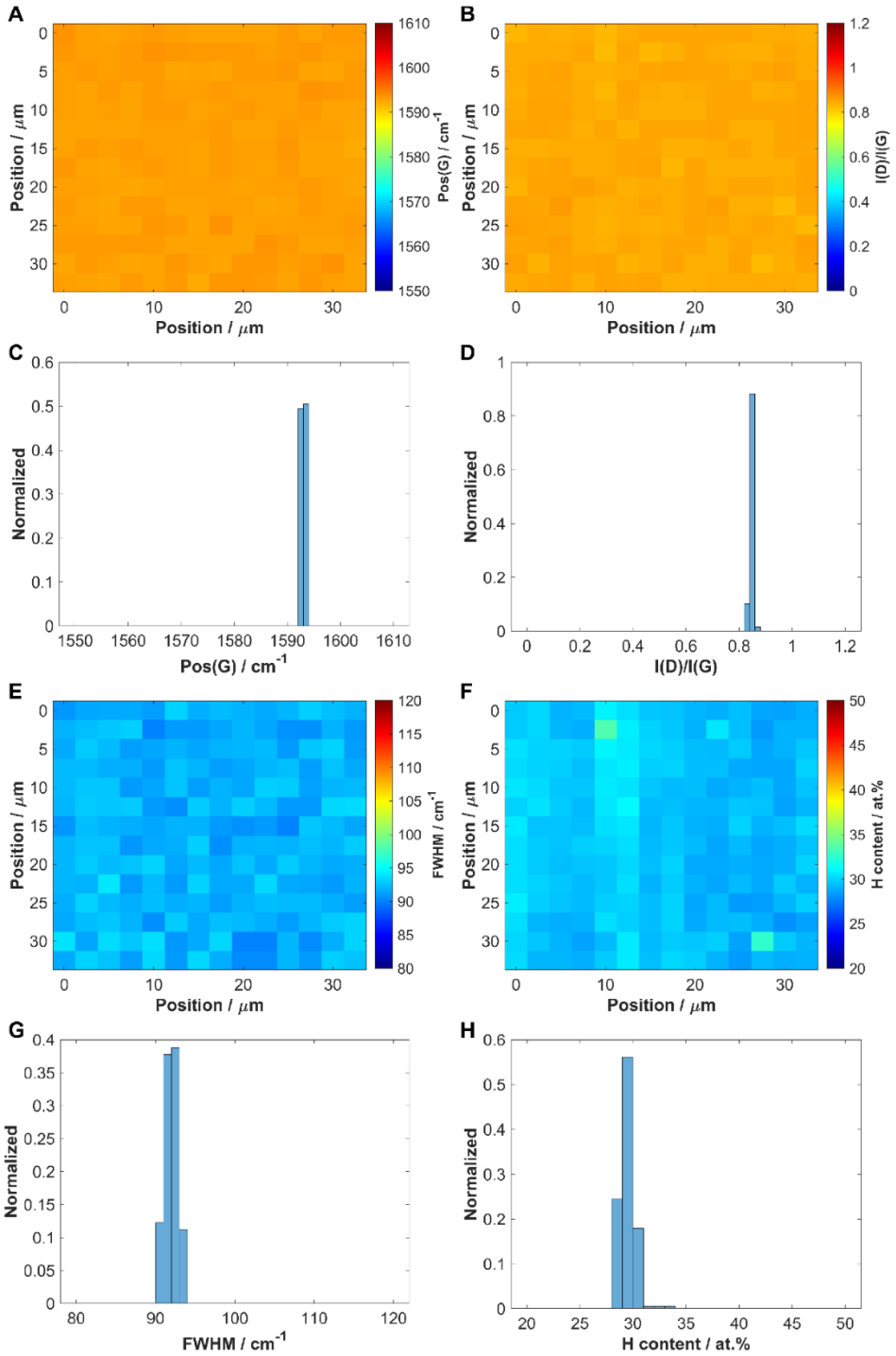

Figure A-26: Raman maps and respective histograms for region 5 of sample HNO_3(2h) comparing different features. A) G-band position map. B) Intensity ratio of D- and G-bands map. C) G-band position histogram. D) Intensity ratio of D- and G-bands histogram. E) G-band full width at half maximum map. F) Hydrogen content map. G) G-band full width at half maximum histogram. H) Hydrogen content histogram.

Appendix B

Appendix B is Support Information for Chapter 3: "Using Raman spectroscopy and partial least squares to quantify carbon species in mining workplace environment."

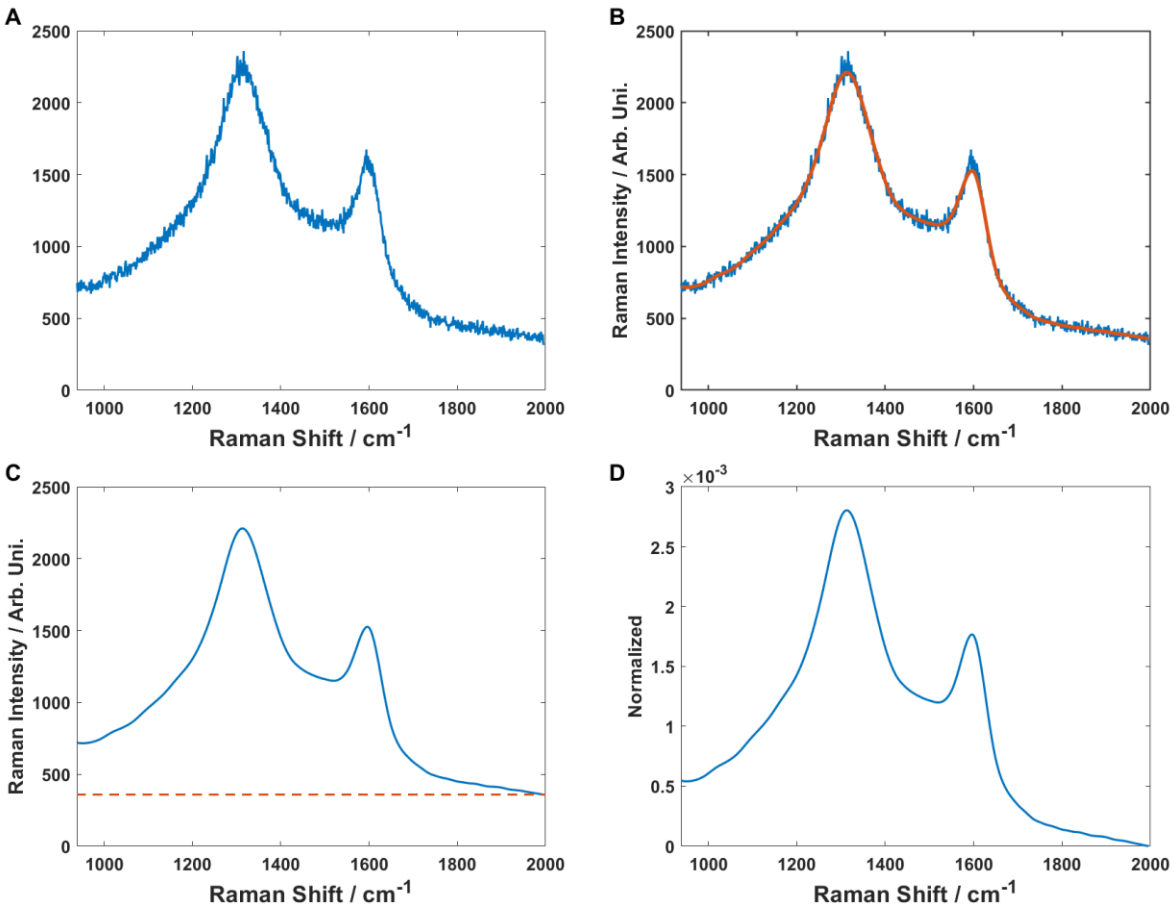

Figure B-1: Preprocessing of Raman spectrum. A) Raw data. B) Raw data (blue) and smoothed (red) spectrum plotted. C) Dashed line showing the horizontal baseline. D) Raman spectrum normalized by the sum of the intensities.

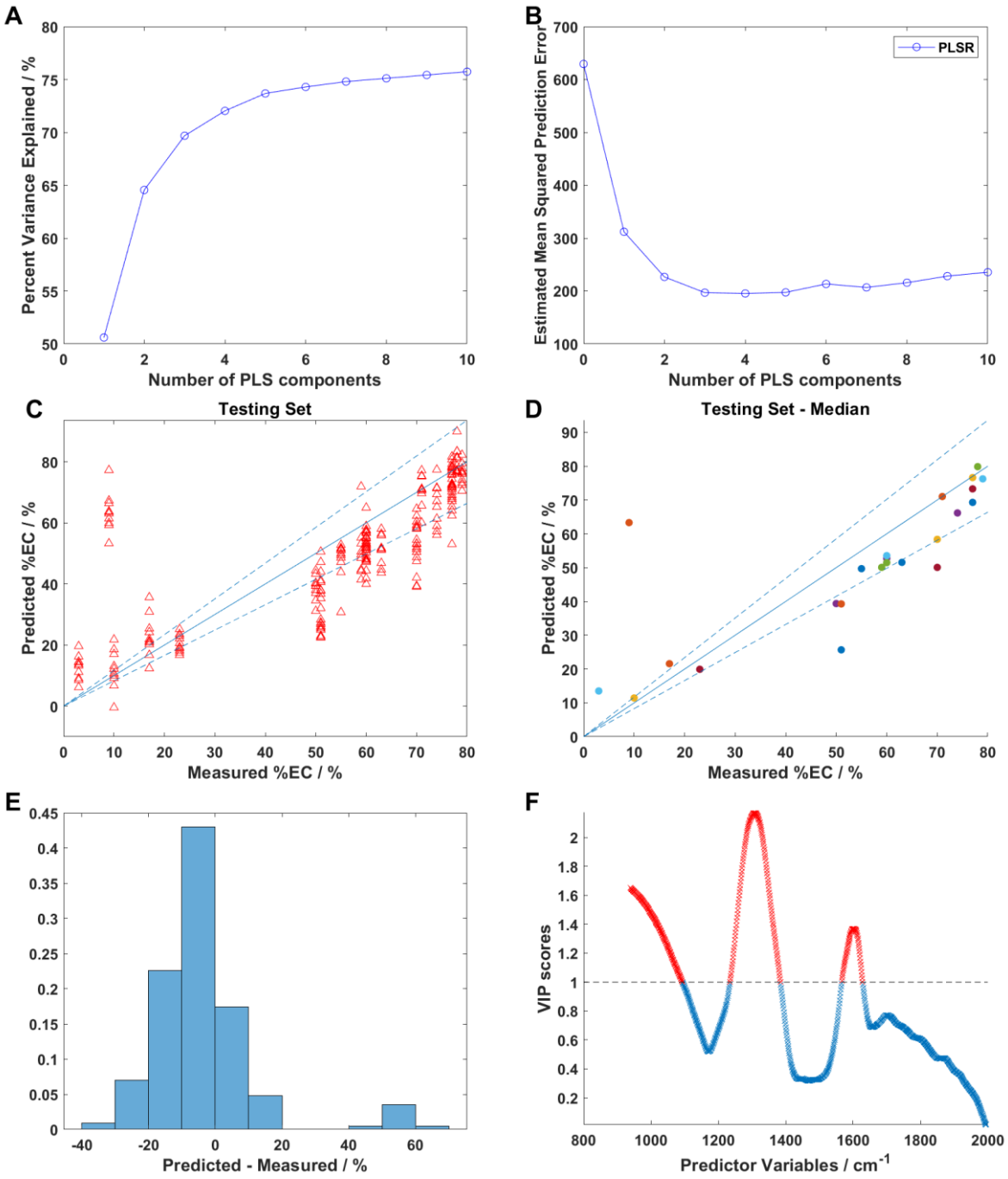

Figure B-2: All the graphs are relative to method estimating %EC, using 785 nm laser. A) Percent variance explained for increasing number of PLS components. B) Estimated mean squared prediction error for increasing number of components. C) Predicted mass plotted against the real ones for testing set. D) Median values of predicted mass for each sample. E) Histogram showing the difference between each predicted and real values. F) Plot of VIP scores showing the importance of each Raman wavenumber for the predictive method.

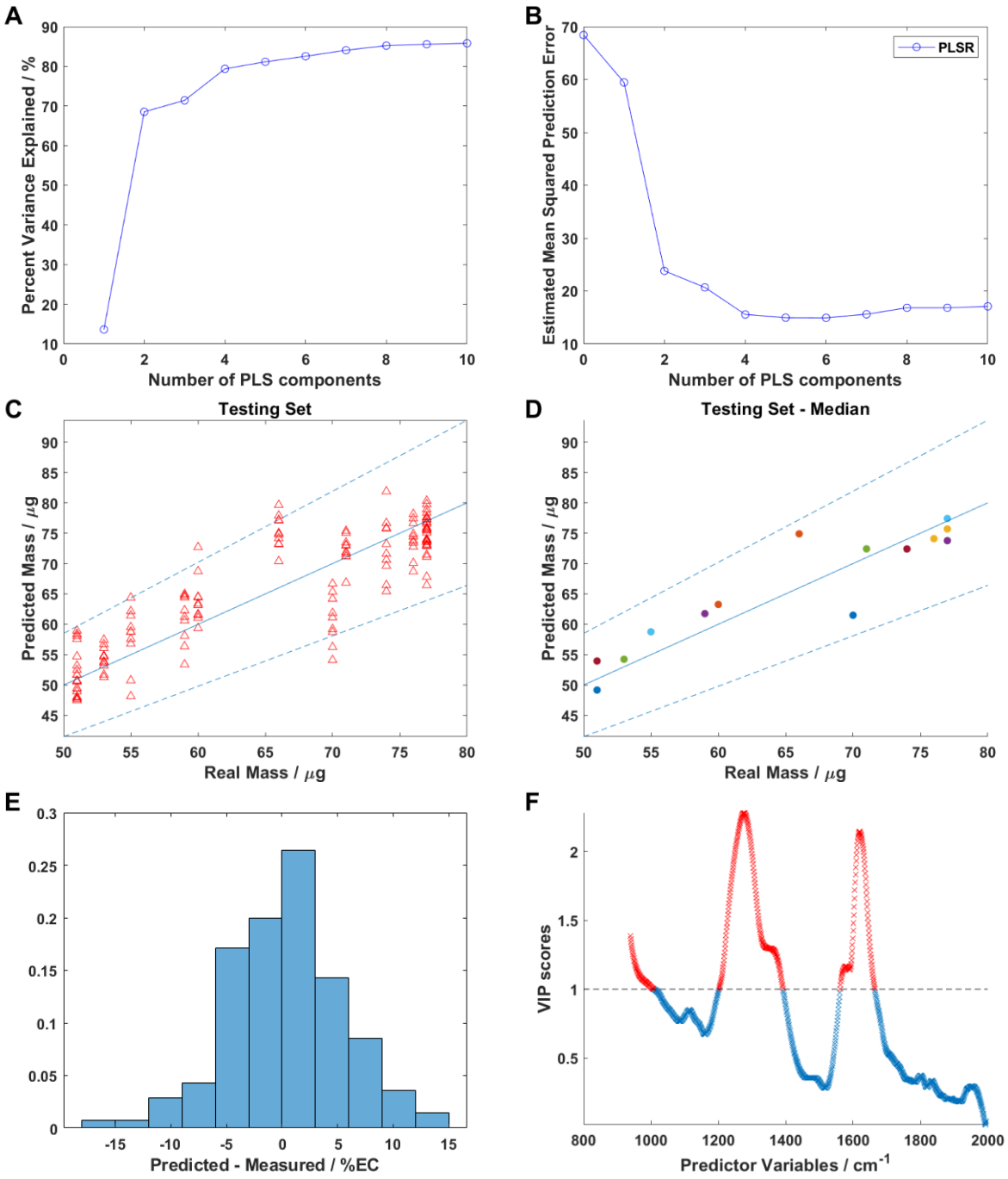

Figure B-3: All the graphs are relative to method estimating %EC, within the range of 50 to 80 %, using 785 nm laser. A) Percent variance explained for increasing number of PLS components. B) Estimated mean squared prediction error for increasing number of components. C) Predicted mass plotted against the real ones for testing set. D) Median values of predicted mass for each sample. E) Histogram showing the difference between each predicted and real values. F) Plot of VIP scores showing the importance of each Raman wavenumber for the predictive method.

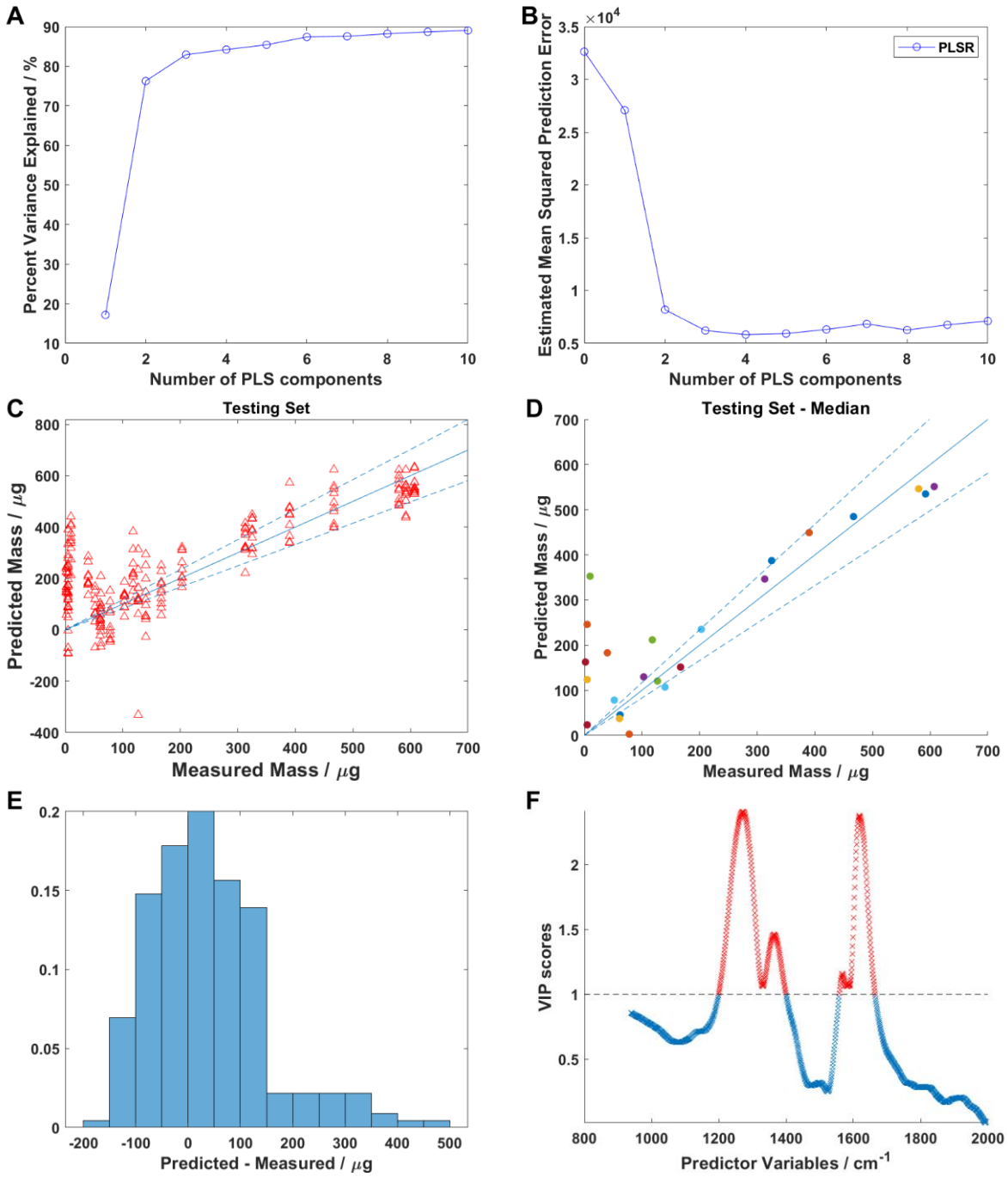

Figure B- 4: All the graphs are relative to method estimating mEC, using 785 nm laser. A) Percent variance explained for increasing number of PLS components. B) Estimated mean squared prediction error for increasing number of components. C) Predicted mass plotted against the real ones for testing set. D) Median values of predicted mass for each sample. E) Histogram showing the difference between each predicted and real values. F) Plot of VIP scores showing the importance of each Raman wavenumber for the predictive method.

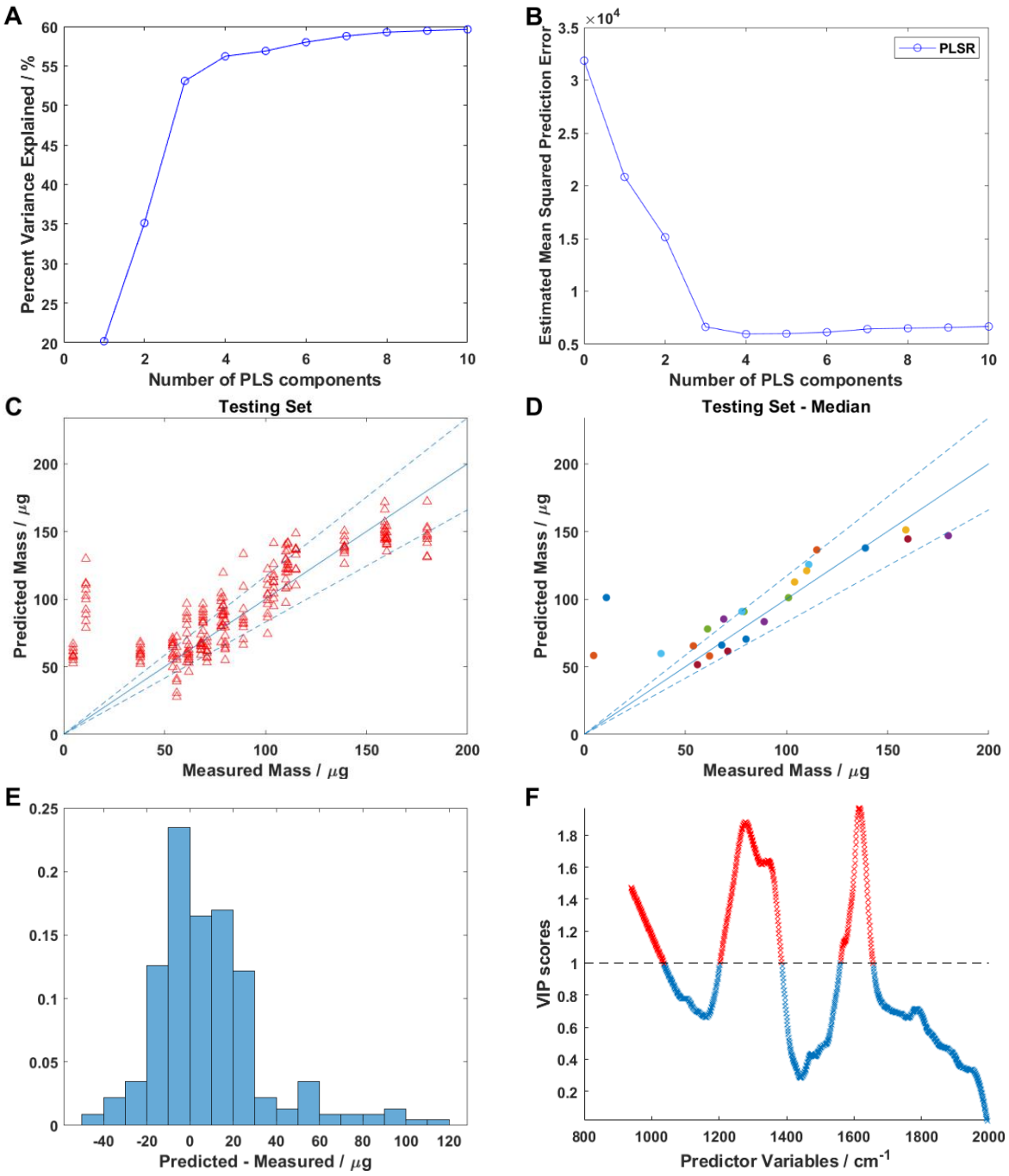

Figure B- 5: All the graphs are relative to method estimating mOC, using 785 nm laser. A) Percent variance explained for increasing number of PLS components. B) Estimated mean squared prediction error for increasing number of components. C) Predicted mass plotted against the real ones for testing set. D) Median values of predicted mass for each sample. E) Histogram showing the difference between each predicted and real values. F) Plot of VIP scores showing the importance of each Raman wavenumber for the predictive method.

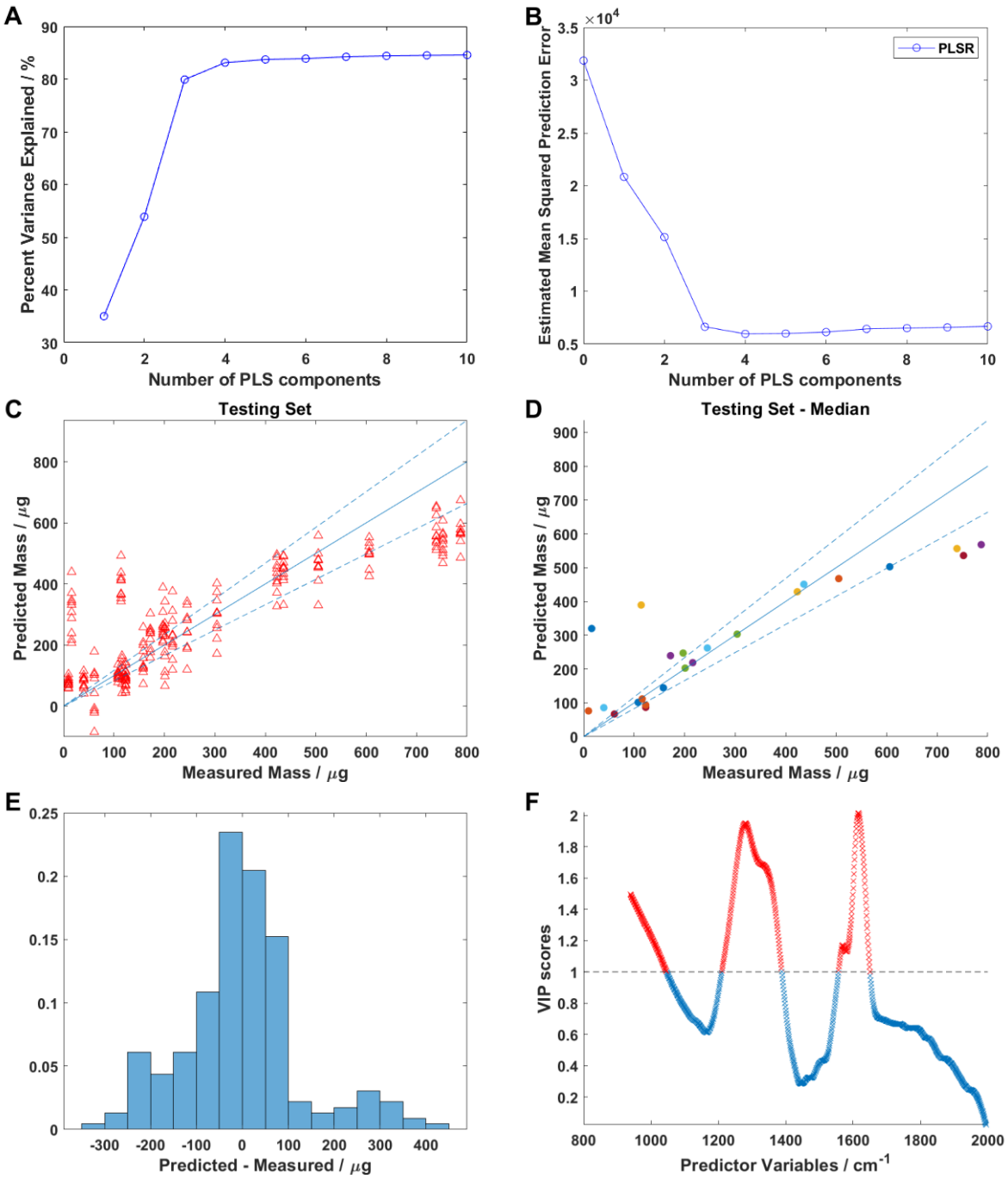

Figure B- 6: All the graphs are relative to method estimating mTC, using 785 nm laser. A) Percent variance explained for increasing number of PLS components. B) Estimated mean squared prediction error for increasing number of components. C) Predicted mass plotted against the real ones for testing set. D) Median values of predicted mass for each sample. E) Histogram showing the difference between each predicted and real values. F) Plot of VIP scores showing the importance of each Raman wavenumber for the predictive method.

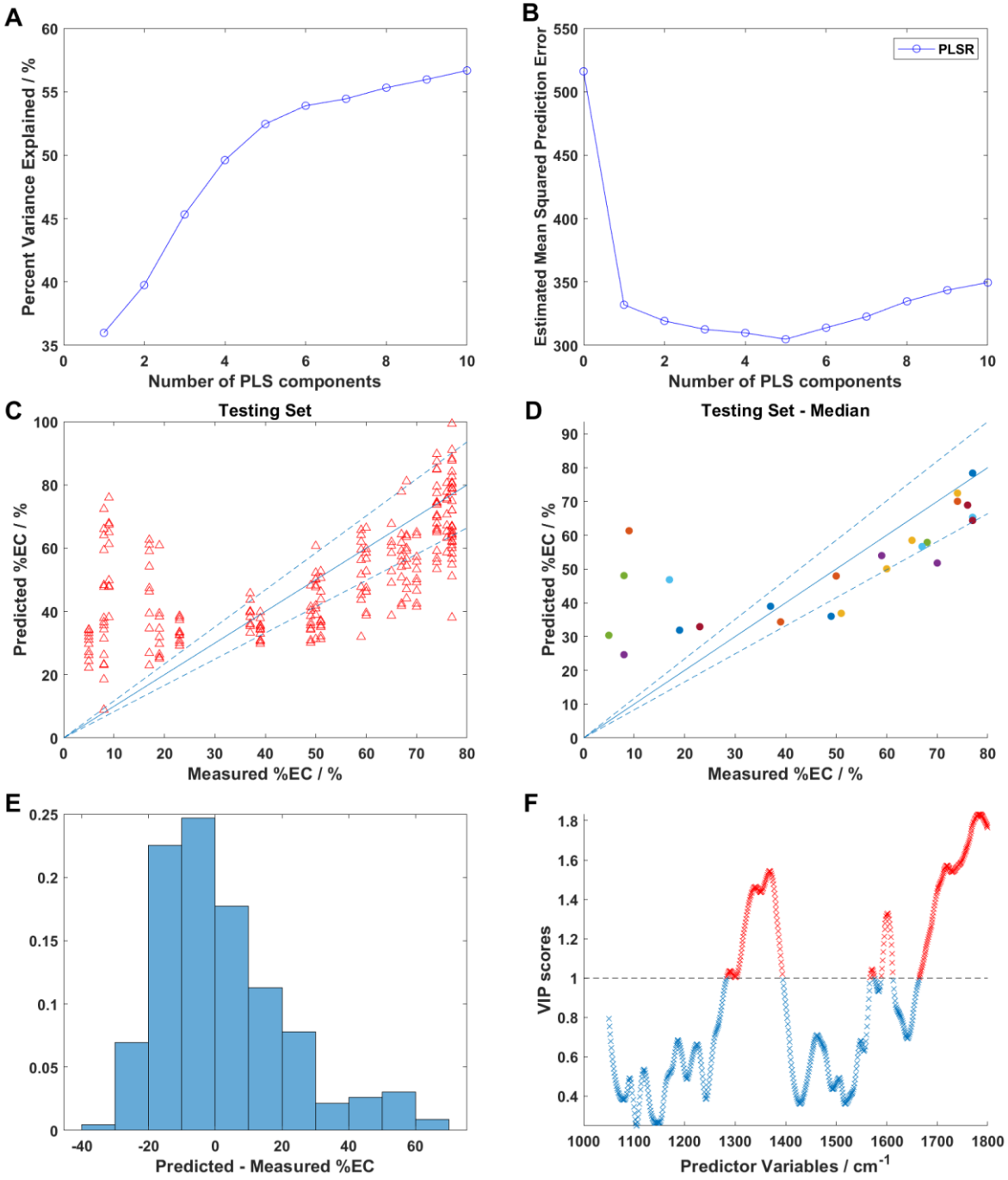

Figure B- 7: All the graphs are relative to the method estimating % of EC, using 532 nm laser. A) Percent variance explained for increasing number of PLS components. B) Estimated mean squared prediction error for increasing number of components. C) Predicted values plotted against the real ones for testing set. D) Median values of predicted %EC for each sample. E) Histogram showing the difference between each predicted and real values. F) Plot of VIP scores showing the importance of each Raman wavenumber for the predictive method.

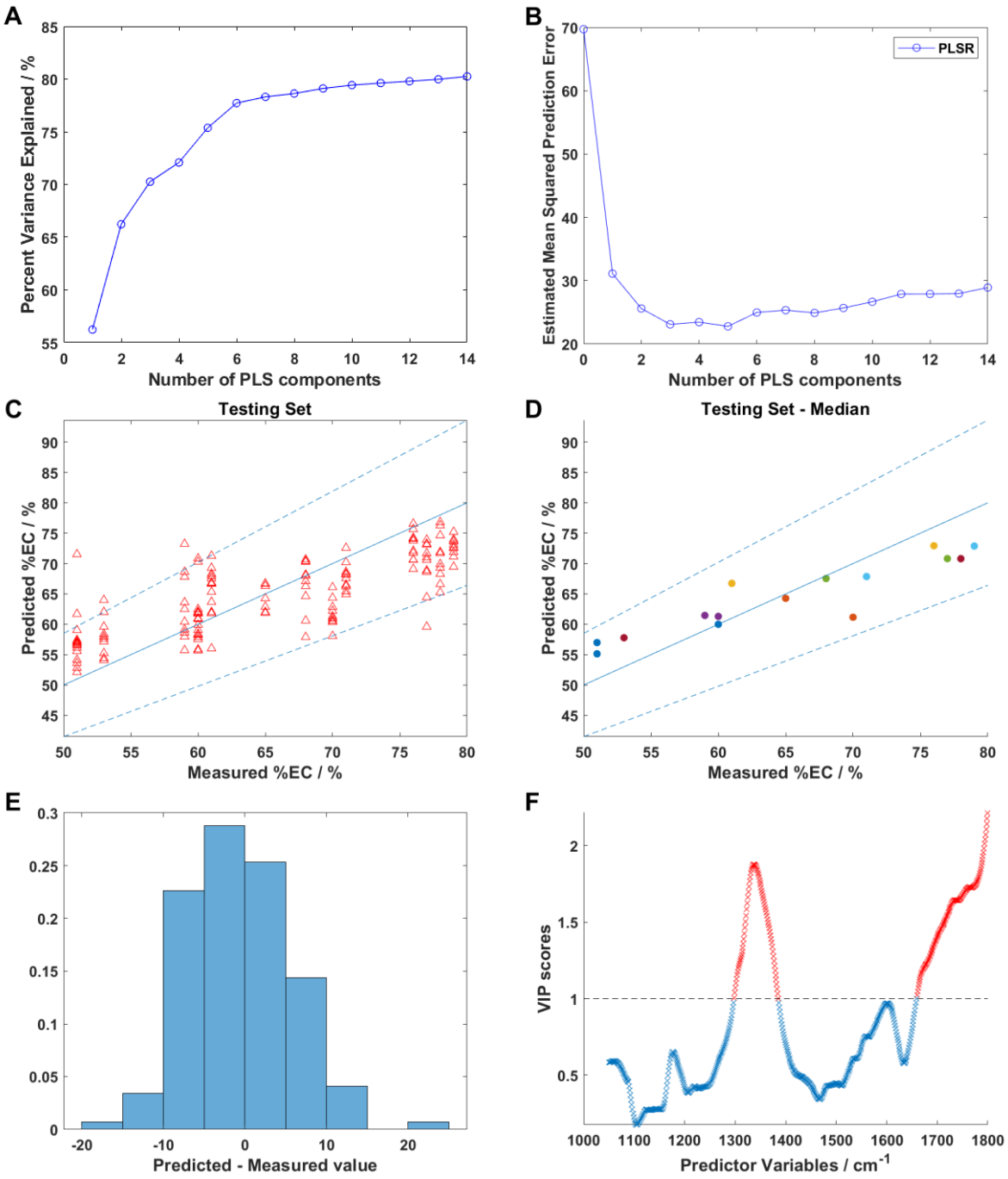

Figure B- 8: All the graphs are relative to method estimating %EC, within the range of 50 to 80 %, using 532 nm laser. A) Percent variance explained for increasing number of PLS components. B) Estimated mean squared prediction error for increasing number of components. C) Predicted mass plotted against the real ones for testing set. D) Median values of predicted mass for each sample. E) Histogram showing the difference between each predicted and real values. F) Plot of VIP scores showing the importance of each Raman wavenumber for the predictive method.

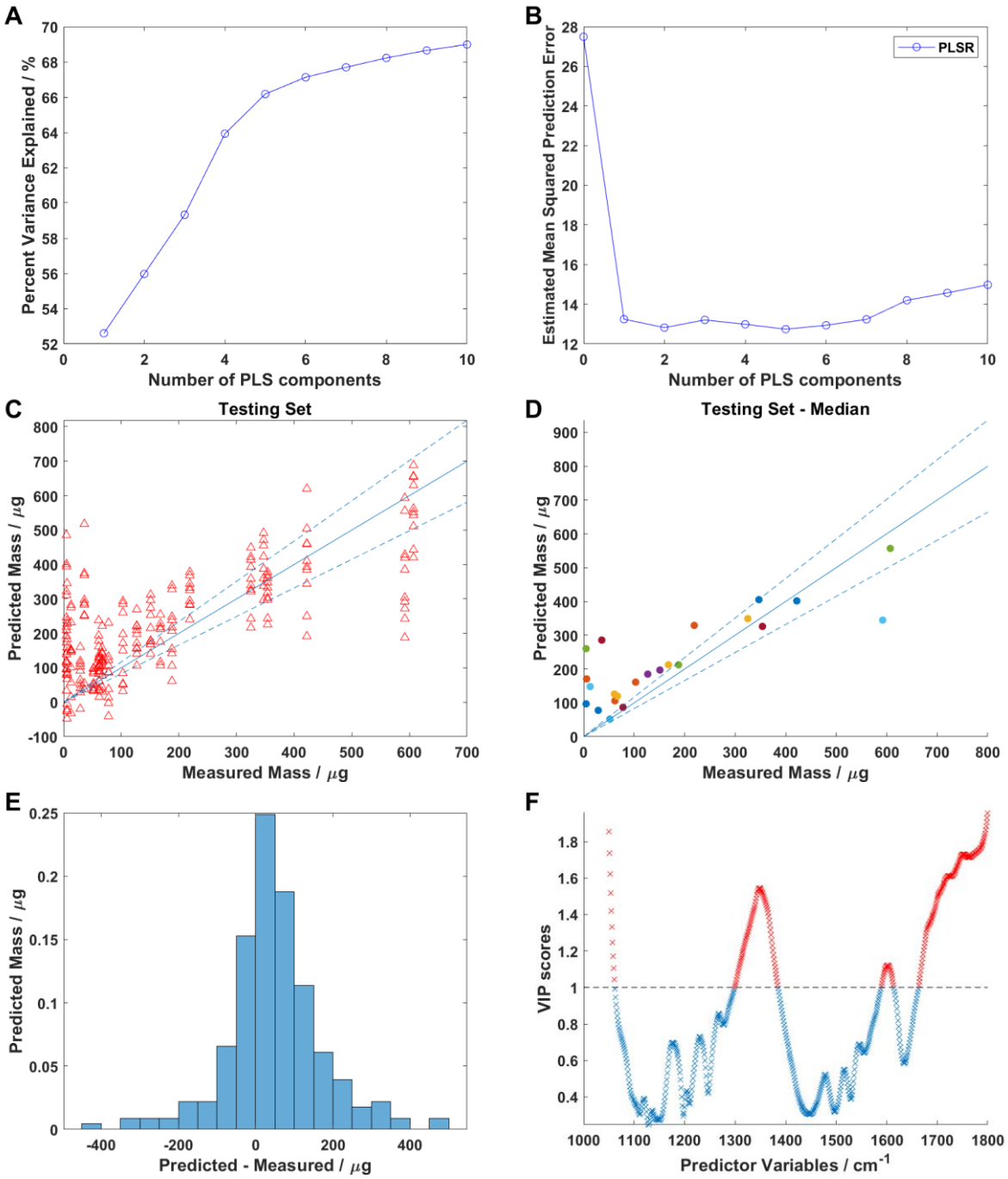

Figure B- 9: All the graphs are relative to method estimating mEC, using 532 nm laser. A) Percent variance explained for increasing number of PLS components. B) Estimated mean squared prediction error for increasing number of components. C) Predicted mass plotted against the real ones for testing set. D) Median values of predicted mass for each sample. E) Histogram showing the difference between each predicted and real values. F) Plot of VIP scores showing the importance of each Raman wavenumber for the predictive method.

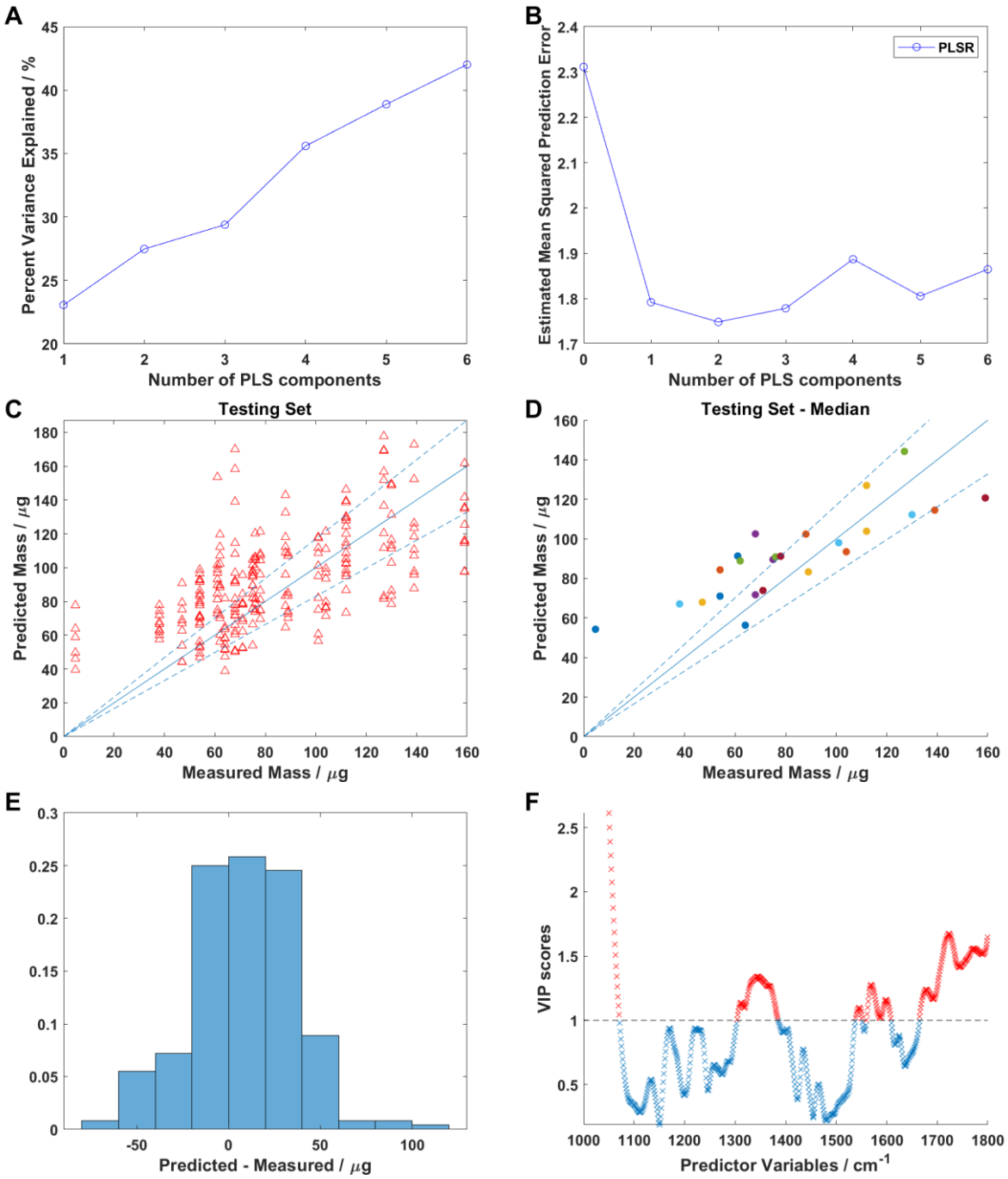

Figure B- 10: All the graphs are relative to method estimating mOC, using 532 nm laser. A) Percent variance explained for increasing number of PLS components. B) Estimated mean squared prediction error for increasing number of components. C) Predicted mass plotted against the real ones for testing set. D) Median values of predicted mass for each sample. E) Histogram showing the difference between each predicted and real values. F) Plot of VIP scores showing the importance of each Raman wavenumber for the predictive method.

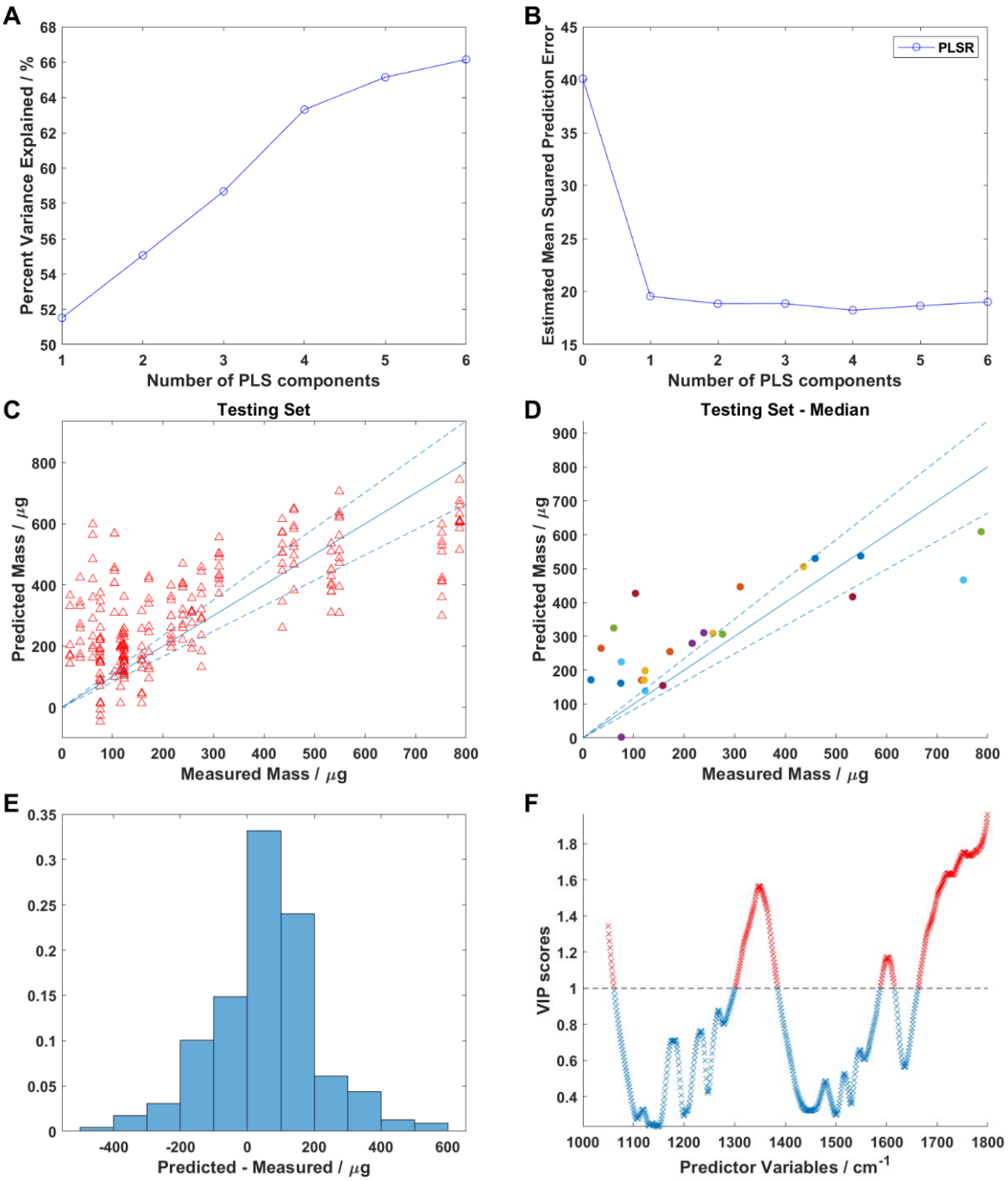

Figure B- 11: All the graphs are relative to method estimating mTC, using 532 nm laser. A) Percent variance explained for increasing number of PLS components. B) Estimated mean squared prediction error for increasing number of components. C) Predicted mass plotted against the real ones for testing set. D) Median values of predicted mass for each sample. E) Histogram showing the difference between each predicted and real values. F) Plot of VIP scores showing the importance of each Raman wavenumber for the predictive method.

Appendix C

Appendix C is Support Information for Chapter 4: "Classification of carbon materials using Raman spectroscopy and PCA-LDA."

Table C-1: Synthetic Parameters.

Electrolyte	Potential	Time	Anode	Cathode
HNO_3	5 V	1 h	Graphite	Titanium
HNO_3	5 V	2 h	Graphite	Titanium
HCl	5 V	1 h	Graphite	Titanium
HCl	5 V	2 h	Graphite	Titanium
ACN AA TBA	5 V	1 h	Platinum	Titanium
ACN AA TBA	5 V	2 h	Platinum	Titanium
ACN AA TBA	5 V	3 h	Platinum	Titanium
ACN AA TBA	5 V	15 h	Platinum	Titanium
ACN	5 V	5 h	Platinum	Titanium
ACN AA Normal TBA	20 V	15 h	Graphite	Titanium
ACN AA High TBA	5 V	15 h	Graphite	Titanium
ACN AA (no TBA)	20 V	17 h	Graphite	Titanium

0.5 M Na_2SO_4	10 V	15 h	Graphite	Titanium
0.5 M H_2SO_4	10 V	15 h	Graphite	Titanium
0.5 M HNO_3	10 V	15 h	Graphite	Titanium
0.5 M HCl	10 V	15 h	Graphite	Titanium
0.5 M HNO_3	5 V	15 h	Graphite	Titanium
0.5 M HCl		Overnight	Graphite	Titanium

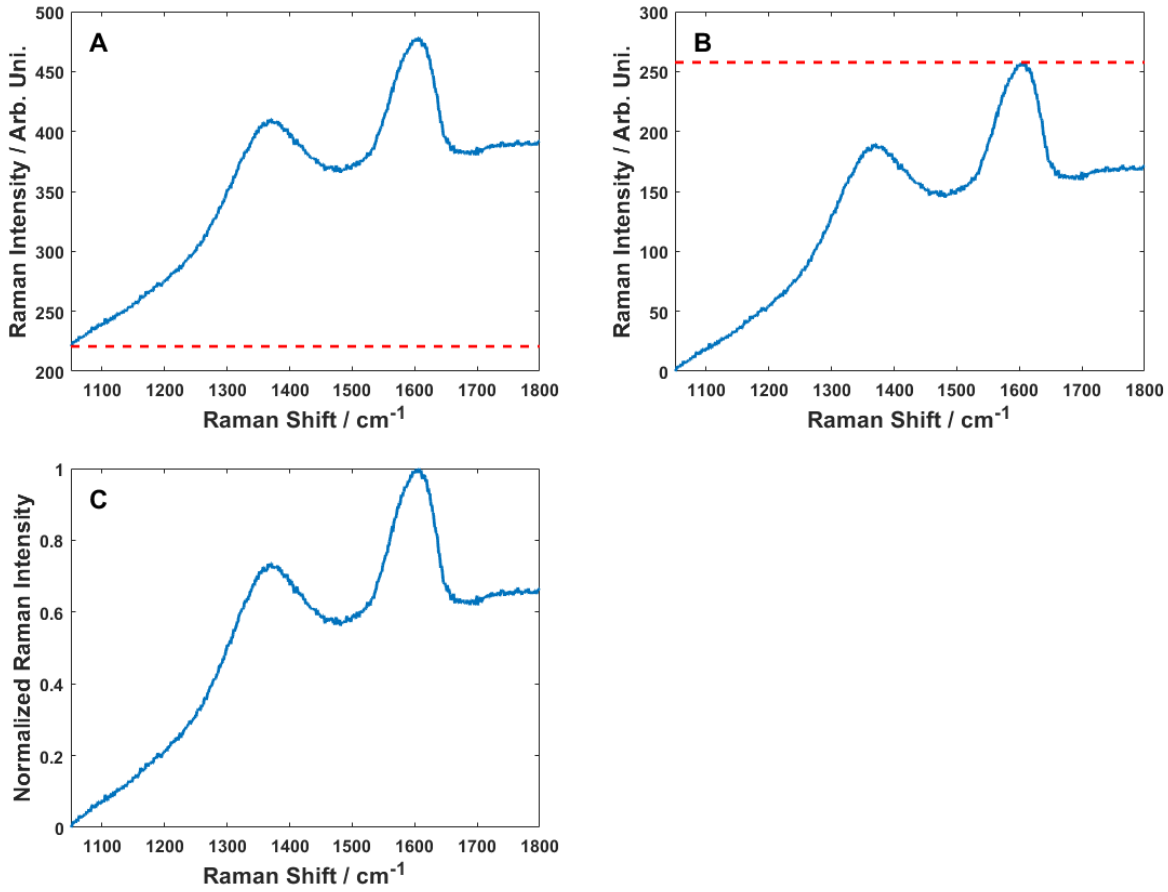

Figure C- 1: Figures illustrating the process of baseline correction and normalization applied to the data. A) Raman spectrum from PLCH (blue line) and minimum intensity (red dashed line) of the spectrum. B) Raman spectrum (blue line) after subtraction of the horizontal baseline at the minimum. The red dashed line represents the maximum that will be used for normalizing the spectrum. C) The final normalized Raman spectrum that will be used for training and testing the machine learning model.

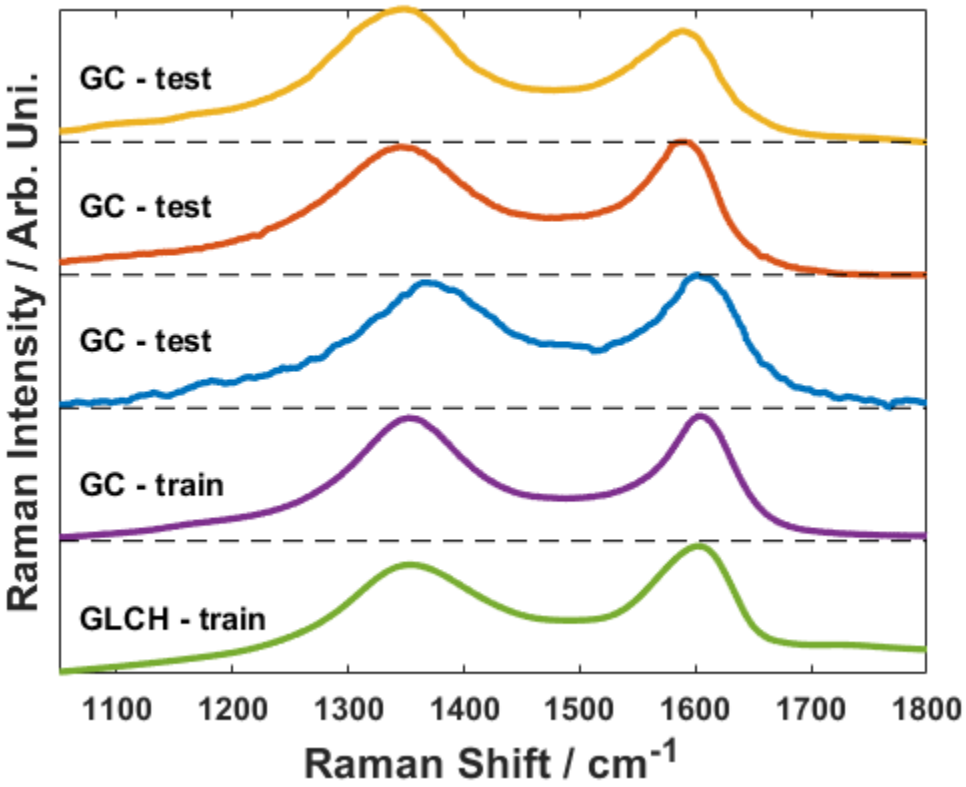

Figure C-2: Raman spectra from misclassified GC test samples and mean spectra from training samples.

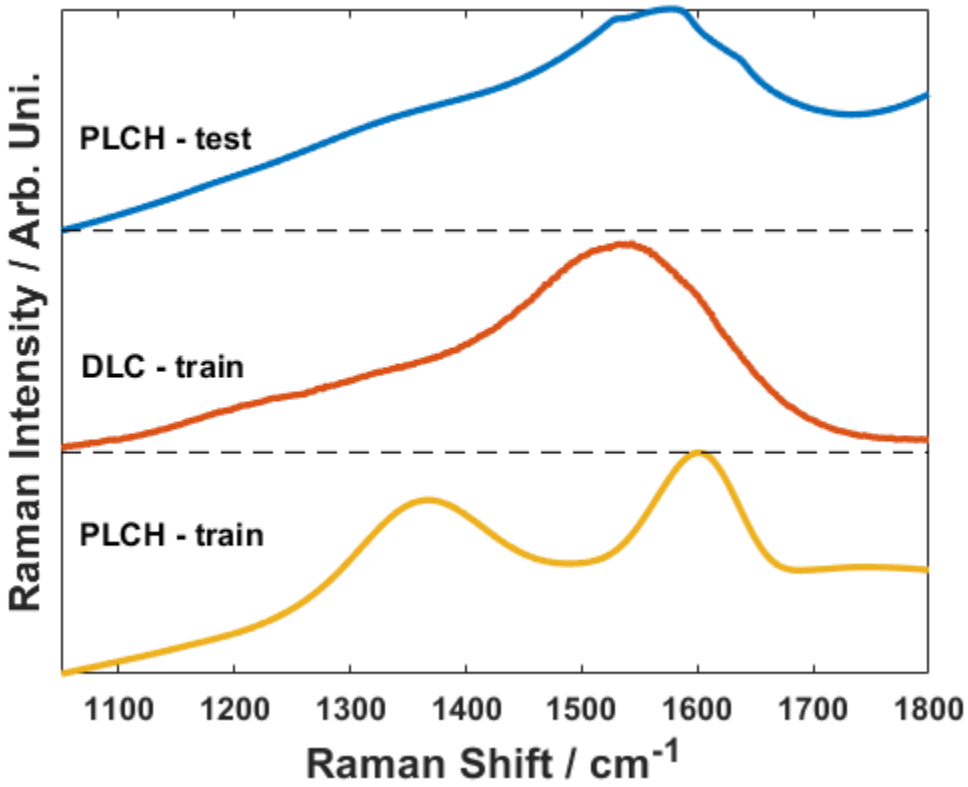

Figure C-3: Raman spectrum from misclassified PLCH test sample and mean spectra from training samples.