RBFTA
Revista Brasileira de Física Tecnológica Aplicada
ISSN: 2358-0089

https://periodicos.utfpr.edu.br/rbfta

Use of computational techniques in Raman spectra analysis

Tulio Santana Ramos
tulip@alumni.usp.br
orcid.org/0000-0002-9613-4542
Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos, São Paulo.

Alfredo A. A. Exposito De Queiroz
alfredo.queiroz@alumni.usp.br
orcid.org/0000-0001-6585-8380
Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, São Paulo
Instituto Tecnológico de Aeronáutica, Campus do CTA, São José dos Campos, São Paulo.

Marcelo B. Andrade
marcelo.barbosa@ufop.edu.br
orcid.org/0000-0001-9137-8831
Departamento de Física, Universidade Federal de Ouro Preto, Ouro Preto, Minas Gerais
Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, São Paulo.

Javier Ellena
javiere@ifsc.usp.br
orcid.org/0000-0002-0676-3098
Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, São Paulo.

ABSTRACT

Raman spectroscopy is a powerful analytical technique widely used across scientific fields due to its ability to generate unique spectral fingerprints for different materials in a non-destructive manner. By analyzing energy shifts caused by molecular vibrations, this method provides valuable insights into the structural and chemical properties of samples. However, raw Raman spectra often contain unwanted background signals, such as fluorescence, which can obscure key spectral features. To address this, preprocessing techniques like baseline correction and peak deconvolution are essential for enhancing spectral resolution and ensuring accurate interpretation. A variety of methods exist for these preprocessing steps, requiring careful selection to optimize spectral analysis. In this study, both visual and numerical comparisons, along with the interpretation of the χ^2 statistical test, were used to evaluate different approaches. Among them, the Asymmetric Least Squares method for baseline correction and the Pseudo-Voigt function for peak deconvolution proved to be the most effective. These techniques significantly improved the quality of Raman spectra, facilitating better identification of vibrational characteristics. By refining spectral preprocessing, this study contributes to more precise and reliable Raman spectroscopy applications, enhancing its use in fields such as materials science, chemistry, and biomedical research.

Keywords: Baseline correction; peak deconvolution; Raman spectroscopy.

INTRODUCTION

Raman spectroscopy is an analytical technique that uses a monochromatic light source (typically a laser) to illuminate a small portion of a sample. During this interaction, a small fraction of the scattered light undergoes a frequency shift due to inelastic scattering. This Raman scattering provides valuable information about the chemical composition, crystal structure, and phase transitions of the material. Specifically, it enables the analysis of vibrational energy levels through the Stokes and anti-Stokes components of the scattered radiation.

In other words, when the scattered light has a different energy than the incident light, this difference can be used to gain insight into atomic vibrations, molecular structure, chemical composition, and intermolecular interactions. As a result, each chemical compound produces a characteristic spectrum (or "fingerprint") that can be used for identification (SANTOS et al., 2019; FARIA; SANTOS; GONÇALVES, 1997; GUEDES; MOREIRA, 2019).

However, the raw spectrum obtained from Raman measurements is not always ideal for analysis due to the presence of fluorescence. This phenomenon results in the emission of absorbed light by the sample, creating a strong background signal that can distort the final spectrum (HORIBA SCIENTIFIC, 2022; CLARKE; OPRYSA, 2004). Therefore, computational techniques are essential for improving visualization and enabling accurate spectral analysis in Raman spectroscopy.

Among these, baseline correction and peak deconvolution play crucial roles in the processing workflow. These steps help standardize spectra and reveal underlying spectral features. This study evaluates different methods for both processes to determine the most suitable approach for Raman spectral analysis. Baseline correction methods assessed include Asymmetric Least Squares, Whittaker smoothing, and Savitzky-Golay filtering. For peak deconvolution, fitting functions such as Gaussian, Lorentzian, Voigt, and Pseudo-Voigt were applied.

ERROR AND STATISTICAL METRICS

To rigorously evaluate the correction and fitting procedures, and the influence of their parameters, a chi-square (χ^2) statistical test was employed. This test measures the difference between observed intensities (O_i) and expected intensities (E_i) at *n* spectral points ($O_0 \ldots O_{n-1}$ and $E_0 \ldots E_{n-1}$), as shown in **Equation 1**.

$$\chi^2 = \sum_{i=0}^{n} \frac{(O_i - E_i)^2}{E_i} \quad (1)$$

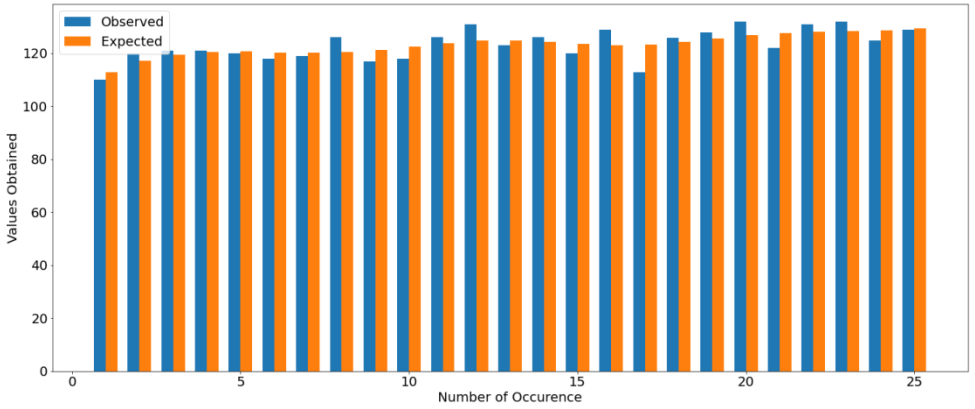
The null hypothesis assumes no difference between the observed and expected values. It is accepted if χ^2 equals 0 and rejected if χ^2 differs from 0. This test allows for a quantitative comparison of the correction and fitting methods, helping identify discrepancies between observed and model-predicted intensities (UGONI; WALKER, 1995).

In baseline correction, the objective is to invalidate the null hypothesis, meaning the χ^2 value must be greater than zero. However, it should not be excessively high, as that would indicate substantial divergence from the original spectrum, potentially distorting or eliminating prominent peaks. For instance,

Figure 1 illustrates the difference between the initial values of the observed and expected spectra used in **Equation 1**. Experimental results showed that when χ^2 exceeded 5570, notable peaks were lost in the corrected spectrum. Additionally, variations in the smoothness parameter began to produce noticeable effects from values of 200 and above.

In contrast, peak deconvolution relies on curve fitting, where the goal is to obtain a χ^2 value as close to zero as possible, indicating minimal deviation between observed data and the fitted model, and thus, a highly accurate fit.

Figure 1 – Comparison between the initial observed and expected spectral values, emphasizing their divergence as the basis for the χ^2 statistical test



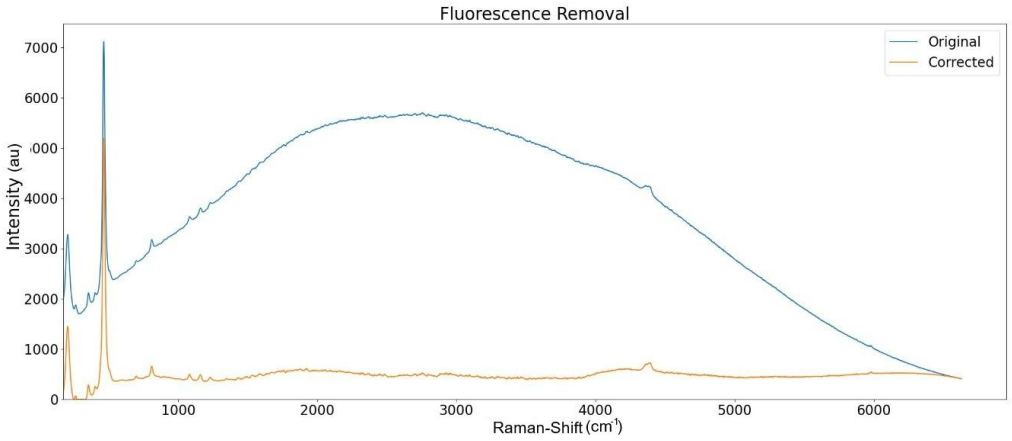
Source: Own authorship.

COMPARATIVE ANALYSIS OF BASELINE CORRECTION METHODS

Raman spectra often exhibit abrupt intensity variations, resulting in distorted curves that are unsuitable for further analysis or processing. Baseline correction serves to standardize these spectra, enabling a more accurate investigation of intrinsic properties and the characteristics of the analyzed sample.

As illustrated in **Figure 2**, a quartz spectrum from the RRUFF Project database (Lafuente et al., 2015; ID R100134), acquired using a 532 nm laser, demonstrates significant distortion due to fluorescence. The corrected version, generated with the CrystalSleuth software (Laetsch & Downs, 2006), shows a clearer signal. Fluorescence can originate not only from the sample itself but also from impurities, ambient conditions (e.g., room temperature), and degradation caused by prolonged laser exposure during data acquisition (Jahn et al., 2021). Such effects introduce noise into the spectrum, producing unfavorable regions characterized by changes in the slope or shape of the background. These artifacts hinder accurate peak identification and quantitative analysis, reinforcing the need for effective baseline correction.

Figure 2 – Fluorescence-affected quartz spectrum from the RRUFF database (ID: R100134) and its corresponding corrected version using CrystalSleuth



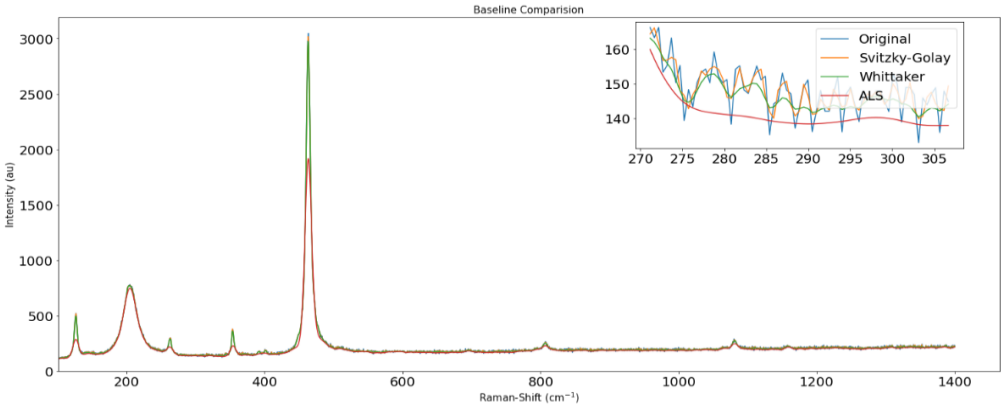
Source: Own authorship.

Due to their versatility and ease of implementation, three baseline correction methods were selected for comparison: Asymmetric Least Squares (ALS), Whittaker Smoothing (WS), and Savitzky-Golay (SG). For the following evaluations, a new Raman spectrum of quartz was collected using a Horiba LabRAM HR Evolution system (200–2200 nm range), with a 532 nm excitation laser.

The ALS method minimizes the sum squared error between the original data and the estimated baseline, adjusting two key parameters: smoothness (λ), which balances baseline flexibility and fit, and penalty (p), which controls asymmetry to tailor the fit to specific spectral features (Exposito de Queiroz and Andrade, 2022). In contrast, WS applies iterative weighting to progressively reduce the second derivative of the signal, smoothing the spectrum while maintaining fidelity to the original data. However, it is subject to constraints related to deviations between the original and corrected curves (Gallagher, 2018). The SG method partitions the dataset into overlapping segments and fits a low-degree polynomial function within each segment to smooth the data. Corrections are then applied based on the differences between the fitted and actual values, enabling localized smoothing.

Figure 3 presents a visual comparison between the original spectrum and the results of each correction method. **Table 1** summarizes selected parameter combinations and their outcomes, where *Pen* denotes the penalty, *λ* the smoothness, *Wind* the window size, and *Asy* the asymmetry level. Full evaluation results are detailed in **Tables 2**, **3**, and **4**.

Figure 3 – Overlay of original and corrected spectra produced by each baseline correction method, highlighting their respective effects



Source: Own authorship.

Table 1 - Effect of parameter variations on the selected methods, including *Pen* (penalty), λ (smoothness), *Wind* (window size), and *Asy* (asymmetry)

WS method			SG method			ALS method		
Pen	Λ	χ^2	Wind	Mode	χ^2	λ	Asy	χ^2
0.01	0.13	1445.64	7	Interp	522.341	10^5	0.05	176099
0.1	0.5	837.965	5	Constant	431.874	10^2	0.2	2866.01

Source: Own authorship.

In the following tables, the first two columns represent the parameters that were varied across executions to evaluate each baseline correction method. The remaining columns present the evaluation metrics: R^2 (coefficient of determination), RMSE (Root Mean Square Error), RSS (Residual Sum of Squares), χ^2 (Chi-square), and Std (Standard Deviation).

In **Table 4**, specifically, the presence of negative R^2 values indicates that the corresponding models performed worse than a simple horizontal line fitted to the original data, highlighting poor baseline estimation in those cases.

Table 2 – Savitzky-Golay parameter effect analysis

Window Length	Mode	R^2	RMSE	RSS	χ^2	Std
4	Interp	0.996963	10.8677	322786	573.35	3.77101
5	Interp	0.999067	6.02383	99171.1	405.973	3.77053
6	Interp	0.996741	11.2577	346367	673.801	3.76995
7	Interp	0.998841	6.71363	123184	522.341	3.76954
8	Interp	0.996611	11.481	360248	735.487	3.76881
5	nearest	0.999066	6.02552	99226.6	406.267	3.77052
5	constant	0.999032	6.13663	102920	431.874	3.77068
5	Mirror	0.999066	6.02786	99303.7	406.702	3.77052
5	Wrap	0.99905	6.07749	100946	417.494	3.77048

Source: Own authorship.

Table 3 – Whittaker parameter effect analysis

Penalty	Optimal Lambda	R^2	RMSE	RSS	χ^2	Std
0.01	0.13	0.996782	11.1866	342008	1445.64	3.71131
0.05	0.32	0.997758	9.33871	238349	1033.3	3.73258
0.1	0.5	0.998154	8.47297	196205	837.965	3.73768
0.2	1.26	0.998413	7.85626	168683	694.469	3.74243
0.3	1.26	0.99871	7.08324	137121	564.879	3.75173
0.4	2	0.998719	7.05711	136111	545.391	3.75642
0.5	2	0.998771	6.91507	130687	525.155	3.76413
0.6	2	0.998733	7.01839	134622	541.055	3.77144

Source: Own authorship.

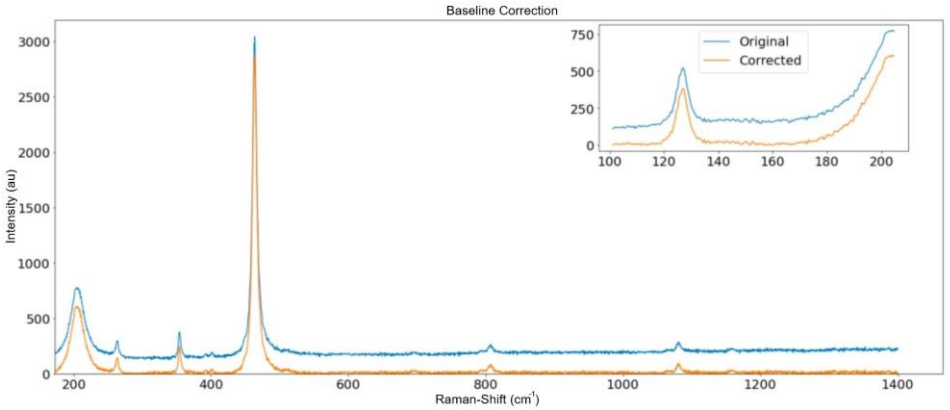
Table 4 – ALS parameter effect analysis

Smoothness	Asymmetry	R^2	RMSE	RSS	χ^2	Std
10^2	0.05	0.91145	58.685	9.41225e+06	8398.42	2.93207
10^3	0.05	0.732942	101.914	2.83865e+07	29603.5	2.23196
10^4	0.05	0.497985	139.73	5.33607e+07	80545.3	1.56001
10^5	0.05	0.281432	167.173	7.63789e+07	176099	1.01378
10^6	0.05	0.115113	185.514	9.40574e+07	320163	0.635159
10^7	0.05	0.0204313	195.187	1.04121e+08	461850	0.397214
10^8	0.05	-0.0429707	201.404	1.10861e+08	591264	0.315423
10^9	0.05	-0.0652516	203.544	1.13229e+08	652125	0.314957
10^2	0.01	0.778661	92.7817	2.35269e+07	24113.4	2.35901
10^2	0.1	0.948218	44.8767	5.50405e+06	4949.4	3.17323
10^2	0.2	0.971246	33.4411	3.05632e+06	2866.01	3.38308
10^2	0.3	0.979767	28.0517	2.15059e+06	2136.21	3.49855
10^2	0.4	0.983742	25.1455	1.72807e+06	1820.77	3.58177
10^2	0.5	0.985419	23.8141	1.54991e+06	1724.2	3.6528
10^2	0.6	0.985328	23.8879	1.55954e+06	1801.84	3.71967

Source: Own authorship.

Alternatively, **Figure 4** presents the correction applied by CrystalSleuth to the same quartz spectrum. While it effectively reduced overall intensity, it did not significantly smooth out noise. The resulting χ^2 value was 654760, notably higher than those obtained with the other methods listed in **Table 1**.

Figure 4 – Correction applied by CrystalSleuth to the quartz spectrum, effectively reducing intensity but with minimal noise smoothing (χ^2 = 654760)



Source: Own authorship.

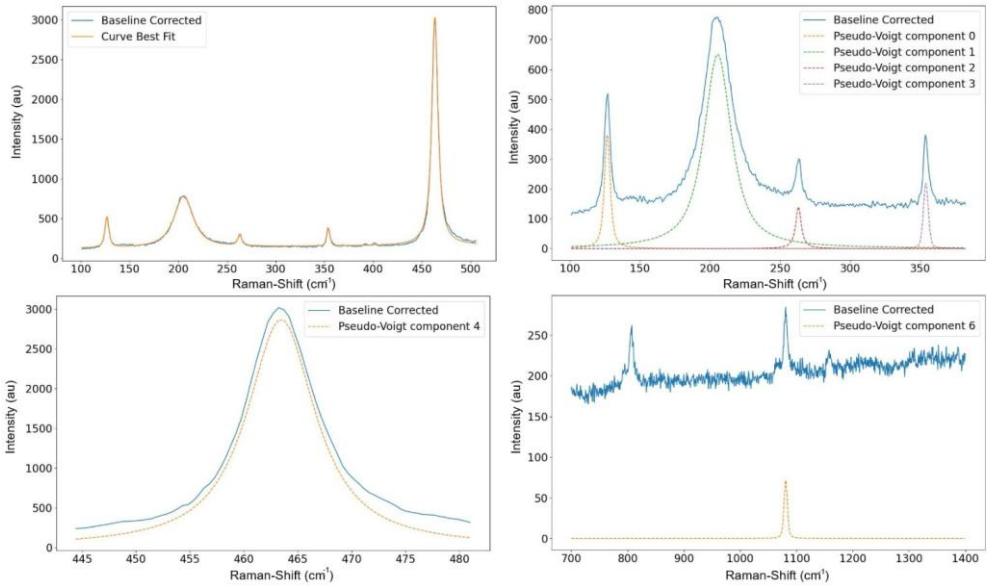
Comparison of **Figure 3** and **Table 1** indicates that the ALS method outperformed the others, yielding a χ^2 value high enough to reject the initial hypothesis while still preserving the integrity of the original spectrum. With appropriately chosen parameters, ALS achieved effective noise reduction without compromising key spectral peaks, which are essential in Raman analysis. This preprocessing step standardizes the data and facilitates more accurate and efficient peak deconvolution.

METHODS OF USING CURVE FITTING FOR PEAK DECONVOLUTION

Peak deconvolution is a computational technique used to separate overlapping peaks in spectra, enabling the detection of hidden features by fitting predefined mathematical functions to the data. Peak overlap may result from limited instrument resolution, which affects spectral line width, or from closely spaced bands with similar energies that interfere with one another (Bi et al., 2015). In Raman spectroscopy, four curve types are commonly used due to their suitability for modeling peak shapes: Gaussian, Lorentzian, Voigt and Pseudo-Voigt functions.

The Gaussian function is based on the normal distribution and is often applied when no prior assumptions are made about the data. The Lorentzian function, derived from the Lorentz distribution, represents the ratio of two independent normally distributed variables with a mean of zero. The Voigt function is a convolution of the Gaussian and Lorentzian curves and is typically used to model asymmetric or broadened peaks (Mani et al., 2022). The Pseudo-Voigt function, by contrast, is a weighted sum of the Gaussian and Lorentzian components, allowing their relative contributions to be adjusted for a better fit (Krishnamurti, 1958). **Figure 5** shows an example of the fitting process, illustrating the best fit achieved using the Pseudo-Voigt function and its individual components.

Figure 5 – Best fit using the Pseudo-Voigt function and its component peaks



Source: Own authorship.

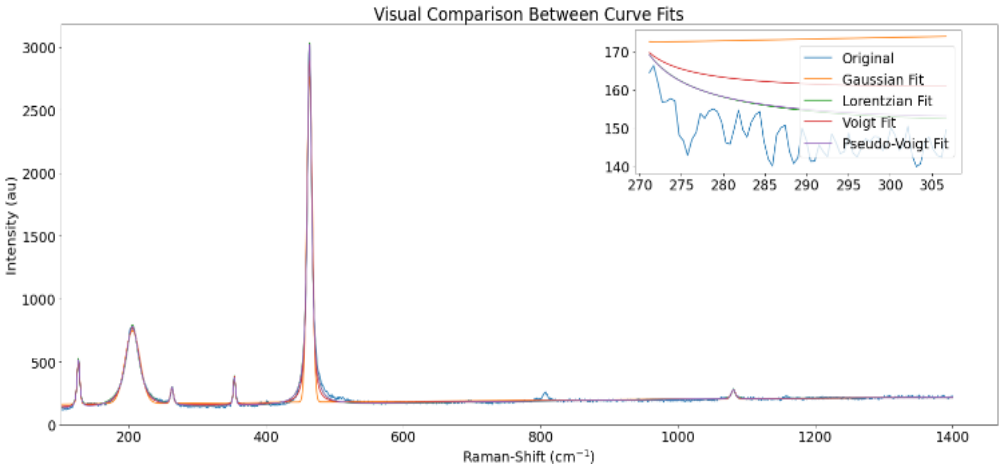
The evaluation combines the visual inspection of the curves in **Figure 6** with the χ^2 values presented in **Table 5**, allowing the Pseudo-Voigt function to yield the best overall performance for this type of material. As previously discussed, lower χ^2 values indicate a better fit, reflecting closer agreement between the spectrum and the model, thereby supporting the initial hypothesis.

However, when the Pseudo-Voigt fit was applied across the entire spectrum, not all known quartz bands were detected. This limitation became evident when comparing the fitted curve with established spectral patterns reported in the literature (Krishnamurti, 1958), which revealed the absence of a characteristic band and an increased χ^2 value of 437200.788. Upon further inspection, the missing peak was found to have been excluded by the algorithm due to its amplitude falling below the predefined detection threshold. This threshold was originally established to prevent the misidentification of random noise as peaks, using amplitude as a filtering criterion to exclude spurious signals. While effective for avoiding overfitting, it can inadvertently suppress genuine low-intensity bands.

To resolve this, a refined parameter optimization strategy was employed. The algorithm's minimum amplitude threshold was reduced to improve sensitivity to subtle but spectrally significant features. Additionally, the initial estimates for peak parameters (such as amplitude, width, and center position) were adjusted based on known quartz features and iterative fitting. The number of initial peak guesses also increased to better account for overlapping or closely spaced bands.

This optimization process allowed for improved retention of key spectral features while maintaining a low χ^2 value, ultimately enhancing the fit's fidelity. As a result, the Pseudo-Voigt function more accurately represented the experimental spectrum, aligning more closely with literature patterns and improving the interpretative reliability of the model.

Figure 6 – Comparison of fitted curves. The Pseudo-Voigt function shows the best overall performance based on visual inspection and χ^2 analysis



Source: Own authorship.

Table 5 – Chi-square (χ^2) values for each curve fitting method, supporting the superior performance of the Pseudo-Voigt function

Curve	χ^2
Gaussian	2929926.98
Lorentzian	470582.904
Voigt	883637.149
Pseudo-Voigt	437200.788

Source: Own authorship.

CHI-SQUARE DISCUSSION IN DETAIL

In both baseline correction and peak deconvolution, the null hypothesis assumes no difference between the observed and expected intensities across the spectrum. However, these two processes interpret the chi-square (χ^2) statistic in opposite ways.

For baseline correction, the goal is to modify the original spectrum to reduce background interference. As such, the null hypothesis should be rejected (reflected by an increase in the χ^2 value) since the data points are intentionally altered to produce a smoother representation. However, excessively high χ^2 values may indicate that the correction has significantly deviated from the original spectrum, potentially distorting prominent peaks that are critical for material identification in Raman analysis.

Therefore, effective baseline correction aims for a χ^2 value that is high enough to reflect meaningful smoothing, but not so high that it compromises key spectral features. Conversely, a very low χ^2 value may suggest that the correction was too conservative and failed to adequately reduce noise or fluorescence. In contrast, peak deconvolution seeks to closely approximate the original spectrum by fitting mathematical functions to overlapping peaks. Here, the null hypothesis should ideally be accepted, meaning that observed and expected values are nearly identical. Thus, lower χ^2 values indicate better performance, while higher

values suggest the selected fitting function did not adequately model the spectral data.

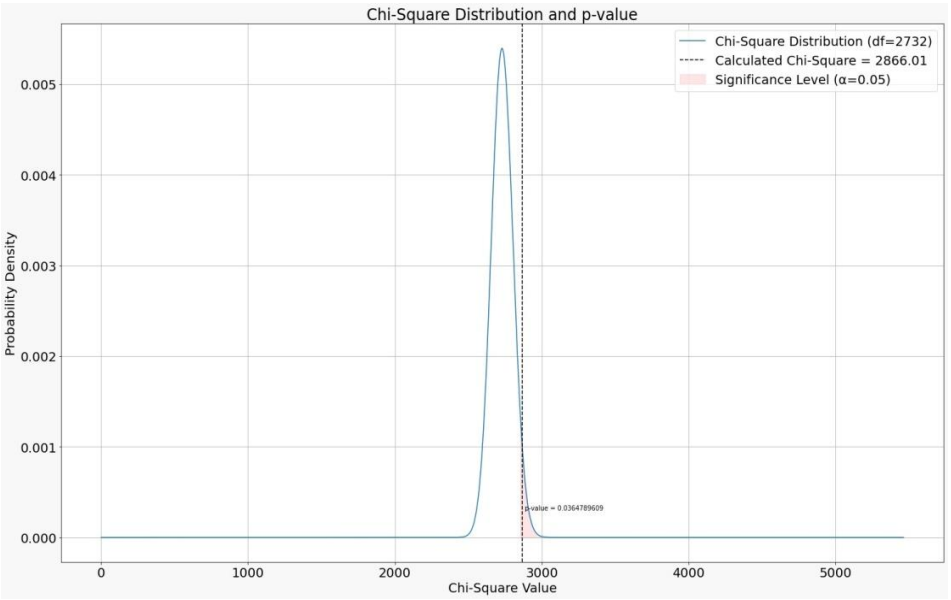
To accurately interpret the chi-square values obtained during testing, several auxiliary components were used. First, the degrees of freedom (df) represent the number of independent comparisons in the analysis. In this context, since two spectra with the same number of points are compared, the degrees of freedom equal the total number of points minus one, resulting in 2732.

This implies that any calculated chi-square value equal to or greater than 2732 will lead to rejection of the null hypothesis. This threshold makes it possible to define acceptable chi-square ranges. For baseline correction, values between 2732 and 5570 were considered valid: the lower bound reflects the degrees of freedom, and the upper bound corresponds to the experimentally determined threshold beyond which peak loss occurred. In contrast, for peak deconvolution, the goal is to accept the null hypothesis, so an ideal chi-square value should be below 2732 (Ugoni and Walker, 1995).

With this foundation, the p-value can be calculated to express the probability of observing a chi-square statistic as extreme as the one obtained. A significance level (α) must also be defined to represent the desired level of confidence. In this study, α was set at 5%, reflecting the high precision required in Raman spectroscopy to distinguish between samples. Moreover, the large degrees of freedom reduce the tail area of the chi-square distribution, making smaller p-values more likely (Ugoni and Walker, 1995; Ferreira & Patino, 2015).

These statistical components are visualized in **Figure 7** using the probability density function (pdf) of the chi-square distribution. The shaded red region denotes the critical region; any calculated value falling within or beyond this region indicates a statistically significant difference. In case of the baseline correction, the calculated p-value was 0.03647, which falls within the critical region, supporting the null hypothesis rejection. For peak deconvolution, the calculated p-value was extremely low (effectively zero) further failing to support the null hypothesis and indicating that the selected fit did not adequately represent the original data.

Figure 7 – Chi-squared distribution illustrating the likelihood of different χ^2 values resulting from the baseline correction methods



Source: Own authorship.

Considering the best results obtained for baseline correction and peak deconvolution, the respective chi-square values were 2866.01 and 437200.788. In both cases, the calculated p-values were lower than the selected significance level ($\alpha = 0.05$), leading to rejection of the null hypotheses and confirming that the observed differences were statistically significant at the 5% level.

Given the intended outcomes for each process, the baseline correction was considered successful, as the rejection of the null hypothesis aligned with expectations. In contrast, the peak deconvolution failed to validate the null hypothesis due to its exceptionally high chi-square value. As previously discussed, this result was directly influenced by the loss of one of the quartz peaks, which significantly impacted the intensity values used in the calculation (Ferreira and Patino, 2015).

CONCLUSION

Baseline correction and peak deconvolution are essential techniques for the analysis and refinement of Raman spectra. By evaluating their performance through graphical comparison and statistical testing, particularly using the chi-square (χ^2) metric, this study identified effective strategies for improving spectral interpretation.

Among the baseline correction methods assessed, Asymmetric Least Squares (ALS) produced the most consistent and accurate results, effectively minimizing fluorescence effects while preserving key spectral features. This preprocessing step created a suitable foundation for subsequent peak deconvolution.

For peak fitting, several curve models were tested to identify hidden spectral features. The Pseudo-Voigt function provided the best overall fit, as indicated by both visual inspection and χ^2 values, suggesting strong alignment with the known peak structure of quartz.

These improvements in spectral accuracy and feature preservation are not only valuable in a methodological sense but also have practical significance across various applied fields. In chemistry, improved peak resolution enables accurate molecular identification and quantification in complex mixtures. In materials science, the ability to detect subtle structural variations and phase distinctions is essential for the design and quality control of advanced materials. In the biomedical field, more reliable spectral interpretation enhances the identification of biomolecular markers, contributing to disease diagnostics, tissue classification, and real-time in vivo analysis.

Overall, the study demonstrated the practical value of combining statistical validation with spectral processing techniques. The findings support the use of ALS for baseline correction and Pseudo-Voigt fitting for peak deconvolution in Raman spectroscopy. These results contribute to more reliable spectral analysis and reinforce the importance of method selection in advancing the accuracy and interpretability of vibrational data.

SUPPLEMENTARY DATA

To ensure reproducibility and facilitate further research, all code and data used in this study are available on GitHub repository (https://github.com/Tulio-Santana-Ramos/Use-of-computational-techniques-in-Raman-spectra-analysis).

Uso de técnicas computacionais na análise de espectros Raman

RESUMO

A espectroscopia de Raman é uma técnica analítica muito eficaz, é amplamente utilizada em diversos campos científicos devido à sua capacidade de gerar assinaturas espectrais únicas para diferentes materiais, além de ser não destrutiva. Ao analisar os deslocamentos de energia causados pelas vibrações moleculares, esse método fornece informações valiosas sobre as propriedades estruturais e químicas das amostras. No entanto, espectros brutos de Raman frequentemente contêm sinais indesejados (ruídos), como a fluorescência, que podem obscurecer características espectrais importantes. Para contornar esse problema, técnicas de pré-processamento, como correção de base e deconvolução de picos, são essenciais para melhorar a resolução espectral e garantir uma interpretação precisa. Diversos métodos estão disponíveis para realizar esses pré-processamentos, exigindo uma seleção cuidadosa de parâmetros para aperfeiçoar a análise espectral. Neste estudo, foram utilizadas comparações visuais e numéricas, junto com a interpretação do teste estatístico χ^2, para avaliar diferentes metodologias. Dentre elas, o método de Mínimos Quadrados Assimétricos para correção de base e a função Pseudo-Voigt para deconvolução de picos se mostraram os mais eficazes. Essas técnicas melhoraram significativamente a qualidade dos espectros de Raman, facilitando a identificação de características vibracionais. Ao aprimorar o pré-processamento espectral, este estudo contribui para aplicações mais precisas e confiáveis da espectroscopia Raman, ampliando seu uso em áreas como ciência dos materiais, química e pesquisa biomédica.

PALAVRAS CHAVE: Correção de linha de base; Deconvolução de picos; Espectroscopia Raman.

Uso de técnicas computacionales en el análisis de espectros Raman

RESUMEN

La espectroscopía Raman es una técnica analítica poderosa, es ampliamente utilizada en diversos campos científicos debido a su capacidad para generar firmas espectrales únicas para diferentes materiales de manera no destructiva. Al analizar los desplazamientos de energía causados por las vibraciones moleculares, este método proporciona información valiosa sobre las propiedades estructurales y químicas de las muestras. No obstante, los espectros Raman en su forma original contienen señales de fondo no deseadas, como la fluorescencia, que pueden oscurecer características espectrales clave. Para abordar esto, las técnicas de preprocesamiento como la corrección de la línea base y la deconvolución de picos son esenciales para mejorar la resolución espectral y garantizar una interpretación precisa. Existen diversos métodos para estos pasos de preprocesamiento, lo que requiere una selección cuidadosa para optimizar el análisis espectral. En este estudio, se utilizaron tanto comparaciones visuales como numéricas, junto con la interpretación de la prueba estadística χ^2, para examinar diferentes metodologías. Entre ellos, el método de Mínimos Cuadrados Asimétricos para la corrección de la línea base y la función Pseudo-Voigt para la deconvolución de picos demostraron ser los más efectivos. Estas técnicas mejoraron significativamente la calidad de los espectros Raman, facilitando una mejor identificación de las características vibracionales. Al refinar el preprocesamiento espectral, este estudio contribuye a aplicaciones más precisas y confiables de la espectroscopía Raman, mejorando su uso en campos como la ciencia de materiales, la química y la investigación biomédica.

PALABRAS CLAVE: Corrección de línea base; Deconvolución de picos; Espectroscopía Raman.

ACKNOWLEDGMENTS

The authors acknowledge the Unified Scholarship Program to Support the Training of Undergraduate Students (PUB-USP), the National Council for Scientific and Technological Development (CNPq, grant number 316927/2023-6), the São Paulo Research Foundation (FAPESP, process numbers 2013/03487-8 and 2024/09091-3), Brasil, for financial support and the Multiuser Structural Crystallography Laboratory at the University of São Paulo São Carlos Institute of Physics (LaMuCrEs-IFSC/USP).

REFERENCES

BI, Y. F. et al. Automatic Recognition of Overlapped Spectral Peaks by Combined Symmetric Zero-Area Conversion and L-M Fitting. *Guang Pu Xue Yu Guang Pu Fen Xi*, v. 35, n. 8, p. 2339-2342, 2015. https://doi.org/10.3964/j.issn.1000-0593(2015)08-2339-04 .

CLARKE, R.; OPRYSA, A. F**luorescence and Light Scattering.** *Journal of Chemical Education*, v. 81, p. 705, 2004. https://doi.org/10.1021/ed081p705.

ENGINEERING STATISTICS HANDBOOK. **Chi-Square Goodness-of-Fit Test**. 2000. Available at: 1.3.5.15. Chi-Square Goodness-of-Fit Test. Accessed on: Feb. 2023.

EXPOSITO DE QUEIROZ, A. A. A.; ANDRADE, M. B. **Prospection of pyrochlore and microlite mineral groups through Raman spectroscopy coupled with artificial neural networks**. *Journal of Raman Spectroscopy*, v. 53, n. 11, p. 1924-1930, 2022. https://doi.org/10.1002/jrs.6433.

FARIA, D. L. A.; SANTOS, L. G. C.; GONÇALVES, N. S. **A demonstration on inelastic light scattering: the Raman experiment revisited.** *Química Nova*, v. 20, n. 3, p. 288-295, 1997. https://doi.org/10.1590/S0100-40421997000300014.

FERREIRA, J. C.; PATINO, C. M. **What does the p value really mean?** *Jornal Brasileiro de Pneumologia*, v. 41, n. 5, p. 485-485, 2015. http://dx.doi.org/10.1590/S1806-37132015000000215.

GALLAGHER, N. **Whittaker Smoother.** 2018. Available at: https://eigenvector.com/wp-content/uploads/2020/01/WhittakerSmoother.pdf. Accessed on: Feb. 2023.

GUEDES, I.; MOREIRA, J. E. **O Efeito Raman**. Universidade Federal do Ceará. 2019. Available at: O Efeito Raman – Seara da Ciência. Accessed on: Feb. 2023.

HORIBA SCIENTIFIC. **What is Fluorescence Spectroscopy?**. 2022**.** Available at: https://www.horiba.com/int/scientific/technologies/fluorescence-spectroscopy/what-is-fluorescence-spectroscopy/. Accessed on: Feb. 2023.

JAHN, I. J.; GRJASNOW, A.; JOHN, H.; WEBER, K.; POPP, J.; HAUSWALD, W. **Noise sources and requirements for confocal Raman spectrometers in biosensor applications**. *Sensors*, v. 21, n. 15, p. 5067, 2021. https://doi.org/10.3390/s21155067.

KRISHNAMURTI, D. **The Raman spectrum of quartz and its interpretation**. *Proceedings of the Indian Academy of Sciences - Section A*, v. 47, n. 5, p. 276-291, 1958. DOI: 10.1007/BF03052811.

LAETSCH, T.; DOWNS, R. **Software for identification and refinement of cell parameters from powder diffraction data of minerals using the RRUFF project and American Mineralogist crystal structure databases.** In: *Abstracts from the 19th General Meeting of the International Mineralogical Association*, Japan, 2006. p. 23-28.

LAFUENTE, B.; DOWNS, R.; YANG, H.; STONE, N. **The power of databases: the RRUFF project**. In: *Highlights in Mineralogical Crystallography*. De Gruyter, 2015. p. 1-30.

MANI, D. et al. **Diffraction Enhanced Imaging Analysis with Pseudo-Voigt Fit Function**. *Journal of Imaging*, v. 8, n. 8, p. 206, 2022. DOI: 10.3390/jimaging8080206.

SANTOS, A. R.; MENEZES, D. B.; ELLENA, J.; ANDRADE, M. B. **Raman Spectroscopy Application in the Characterization of Pertaining Minerals of a Geocollection**. *Química Nova*, v. 42, n. 5, p. 489-496, 2019.

UGONI, A.; WALKER, B. **The Chi square test: an introduction**. *COMSIG Review / COMSIG, Chiropractors and Osteopaths Musculo-Skeletal Interest Group*, v. 4, p. 61-64, 1995.

Recebido: 7 de março de 2025.
Aprovado: 26 de agosto de 2025.

DOI:

Como citar: RAMOS, T S; EXPOSITO DE QUEIROZ, A A A; ANDRADE, M B; ELLENA, J, Use of computational techniques in Raman spectra analysis, **Revista Brasileira de Física Tecnológica Aplicada**, Ponta Grossa, v. 12, n. 2, p. 1-17, dez. 2025.

Contato: Alfredo A. A. Exposito De Queiroz: alfredo.queiroz@alumni.usp.br