Electronic Supplementary Material

For

Novel Application of Silicon Multivacancy Satellite Peaks for Silicate Minerals Analysis in Igneous Rocks Using WD-XRF Coupled with Chemometrics Analysis

Ashok Kumar Maurya*[a]

[a]Geological survey of India, Northern Region, Aliganj Sector E, Lucknow, India-226024.

* Corresponding author's email: ashok.maurya@gsi.gov.in

Spectra Processing

The SUPER Q 6.0 software provided with the WD-XRF instrument can display the recorded X-ray fluorescence spectra, and the spectral data can be exported to a CSV format file. However, this software cannot apply baseline subtraction, deconvolution, fitting, smoothing, etc., to the spectra. Therefore, we used graphic software, OriginPro, for these tasks.

The SUPER Q 6.0 software is designed to display the diffraction angle 2θ values on the abscissa axis. The energy of a fluorescence line is characteristic of an element, while the 2θ value for a fluorescence line is associated with the employed crystal. To investigate chemical shifts on a generalized scale, we re-plotted all spectra with the 2θ values on the abscissa axis replaced by equivalent energy using Bragg's Equation ($n\lambda = 2d \sin\theta$) and Planck's Equation ($E = hc/\lambda$).

After acquiring the spectral data in CSV format, the processing of the spectra was done in the following steps:

(1) Averaging

(2) Smoothing

(3) Baseline subtraction

(4) Lorentz fitting of the spectrum

Averaging of Spectral Data

Each spectrum encompassed 150 data points covering the Si $K\beta^{III}$ and Si $K\beta^{IV}$ satellite peaks. To enhance the accuracy and precision of peak energy and intensity, each spectrum was replicated five times to minimize noise. Outliers within the resulting 5×150 data points were manually discarded. The remaining data points were averaged to generate a smoothed spectrum. An illustrative example of this averaging process is provided in Table S1. A comparison of single scan and five scan average spectra can be seen in Figures S1(a) and S1(b).

Table S1. Illustration of Value Omission and Averaging of Five Scans

Point No.	Energy (eV)	Intensity (cps)					
		Scan 1	Scan 2	Scan 3	Scan 4	Scan 5	Average

57	1862.9	459	446*	461	461	463	461.0
58	1863.2	445	452	476*	447	446	447.5
59	1863.5	453*	440	434	436	435	436.3
60	1863.8	443	445	448	451	452	447.8
61	1864.1	461	459	451	460	458	457.8
----							
* ommited value							

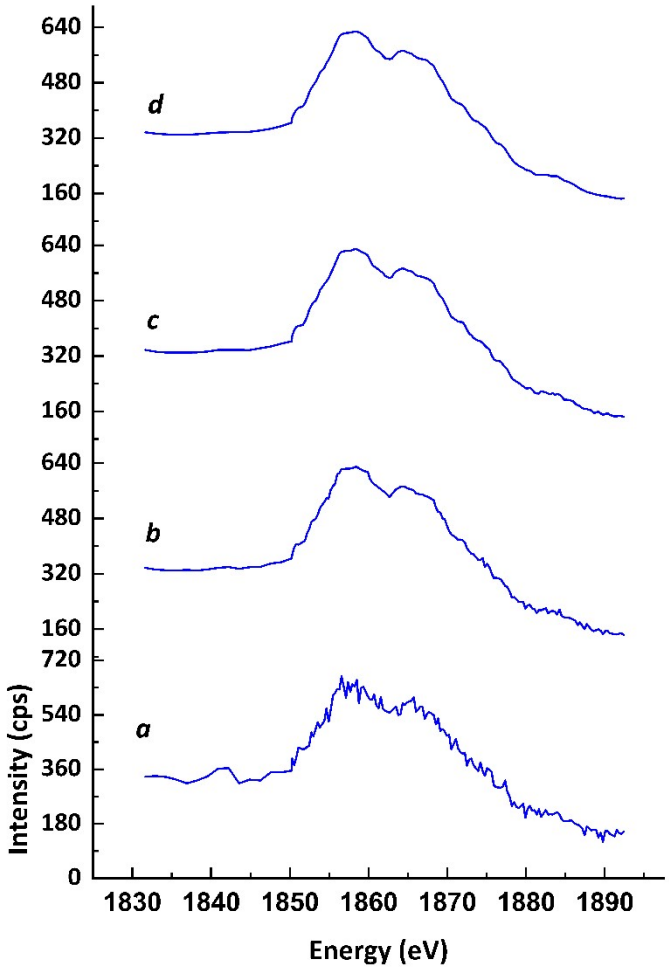

Figure S1. Spectra of Si $K\beta^{III}$ and $K\beta^{IV}$ satellite peaks from CRM GUM-GW. (a) Single scan spectrum; (b) five scans average spectrum, (c) 5 points-second order second-order Savitzky-Golay smoothed spectra and (d) 7 points-second order second-order Savitzky-Golay smoothed spectra.

Smoothing

To enhance the precision in the position and intensity of the Si $K\beta^{III}$ and $K\beta^{IV}$ satellite peaks, we applied two additional successive second-order Savitzky-Golay smoothing processes. This improved the intensity and energy position of the peaks. After applying the Savitzky-Golay smoothing, no shift was observed in either the energy.

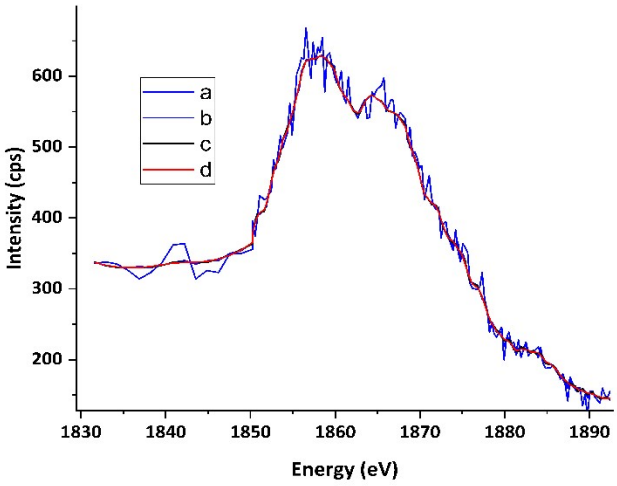

Figure S2. Comparison of energy and intensity positions of spectra of GUM-GW. (a) Single scan spectrum; (b) five scans average spectrum, (c) 5 points-second order second-order Savitzky-Golay smoothed spectra and (d) 7 points-second order second-order Savitzky-Golay smoothed spectra.

or the intensity as can be seen in Figure S2. Comparisons between the smoothed and raw spectra are provided in Figure S1.

Background Subtraction

Background subtraction is a data processing technique used to improve the accuracy of spectral analysis by removing unwanted background signals from the spectra. In X-ray emission spectroscopy, this is particularly important for accurately identifying and quantifying specific peaks, such as the Si $K\beta^{III}$ and $K\beta^{IV}$ satellite peaks.

In OriginPro, the "Subtract Baseline" command automates the background subtraction process:

Baseline Identification: The software identifies the baseline by analyzing regions of the spectrum that do not contain significant peaks.

Fitting and Subtraction: A baseline model is fitted to these regions and subtracted from the entire spectrum, bringing the background to zero.

By using background subtraction, the resultant spectrum more accurately reflects the true intensities of the Si $K\beta^{III}$ and $K\beta^{IV}$ satellite peaks, facilitating more precise analysis and interpretation of the data. A background subtracted spectrum has been depicted in Figure S3.

Lorentz Fit Model in Spectral Analysis

The Lorentz fit model, also known as the Lorentzian function, is a commonly used model for fitting spectral peaks. It is particularly useful for deconvoluting overlapping peaks, such as the Si $K\beta^{III}$ and $K\beta^{IV}$ satellite peaks in X-ray emission spectra.The Lorentzian function describes the shape of a spectral peak based on the Lorentz distribution. It is characterized by its peak position, peak height, and full width at half maximum (FWHM). In OriginPro, applying the Lorentz fit model to the background-subtracted spectrum involves the following steps:

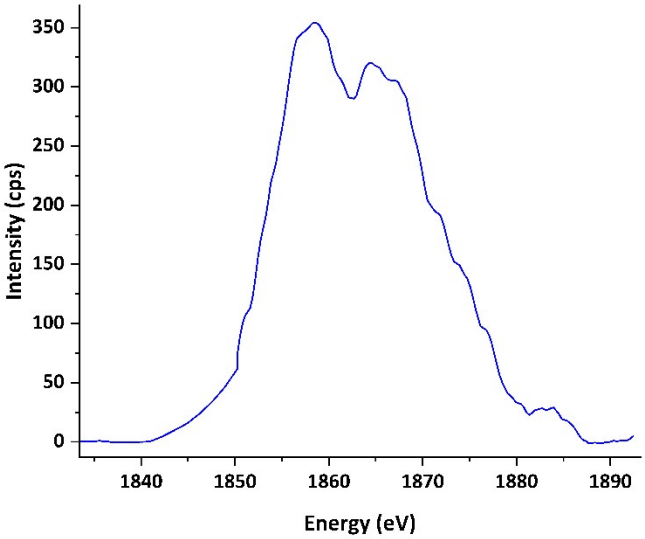

Figure S3. Base line subtracted spectra of GUM-GW.

Select Peak Regions: Identify the regions in the spectrum where the peaks are located.

Apply Lorentz Fit: Use the Lorentz fit function in OriginPro to fit the identified peak regions. The software will provide initial parameter estimates and iteratively refine them to achieve the best fit.

Obtain Deconvoluted Peaks: The software outputs the fitted Lorentzian functions, which represent the deconvoluted peaks. These functions provide the position, height, and FWHM of each peak, allowing for precise analysis of the overlapping Si $K\beta^{III}$ and $K\beta^{IV}$ satellite peaks.

By using the Lorentz fit model, one can accurately resolve overlapping peaks and obtain detailed information about their individual characteristics, enhancing the precision of your spectral analysis. The deconvoluted spectrum has been depicted in Figure S4.

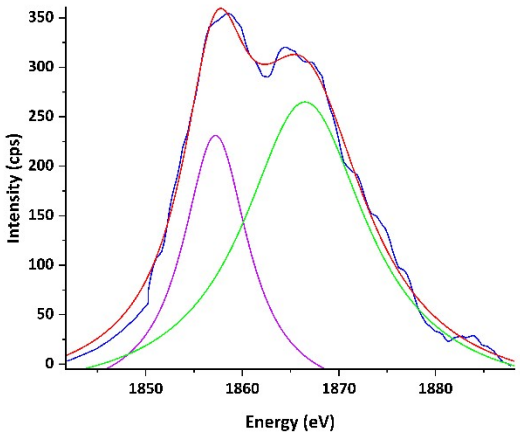

Figure S4. Lorentzian fit of GUM-GW spectra. The blue line represents the baseline-subtracted spectrum, the red line indicates the cumulative fit, the pink line corresponds to the Si $K\beta^{III}$ peak, and the green line represents the Si $K\beta^{IV}$ peak.