Comparative Study of Energy Dispersive X-ray Spectrum Deconvolution: PyMCA and WinAxil

A Thesis Submitted to the Department of Nuclear Sciences & Applications

GRADUATE SCHOOL OF NUCLEAR AND ALLIED SCIENCES

UNIVERSITY OF GHANA

By

NYANTAKYI YAW TWUM BANNOR

(10396011)

In Partial Fulfilment Of The Requirements For The Award

Of

MASTER OF PHILOSOPHY

APPLIED NUCLEAR PHYSICS

JULY, 2015

DECLARATION

I hereby declare that with the exception of references to other people's work which have duly been acknowledged, this thesis is the result of my own research work and no part of it has been presented for another degree in this University or elsewhere.

……………………………………… ………………………………

NYANTAKYI Y. T. BANNOR Date

(10396011)

I hereby declare that the preparation of this project was supervised in accordance with the guidelines of the supervision of thesis work laid down by the University of Ghana.

………………………………… ………………………………

REV. DR. S. AKOTO BAMFORD DR. FRANCIS G. OFOSU

(PRINCIPAL SUPERVISOR) (CO-SUPERVISOR)

………………………………… …………………………………

Date Date

DEDICATION

I dedicate this thesis to my mother and late father, to my sister and brothers, to my benefactors and finally to all teachers past and present.

ACKNOWLEDGEMENT

I am very thankful to God Almighty for giving me the strength and edge to take up this study.

My sincere gratitude goes to my supervisors, Rev. Dr. Samuel Akoto Bamford and Dr. Francis Gorman Ofosu for their tolerance, encouragement, priceless advice and kind supervision.

I am also thankful to Dr. Christian Nuviadenu for his advice and assistance in helping get acquainted with the PyMCA software. Finally, my profound gratitude goes to Dr. V. A. Sole, the developer of the PyMCA software for his prompt response to all my queries regarding the software.

Table of Contents

ABSTRACT

A comparative study of spectrum deconvolution and quantitative analysis between PyMCA & WinAxil has been undertaken. Standard spectra of X-ray Kα lines (Ti, Fe, Co, Ni, Cu, Zn, As, Sr & Y) & Lα lines (Ce, Sm, Dy, Ta, W, Hg & U) were fitted with both software; also spectra of three soil samples were fitted with both software. Intensities of the elements of the standard spectra were obtained from the results of the fitting using both software and comparisons were made separately for both Kα & Lα lines between the two software. Sensitivity calibrations were done for both Kα & Lα lines and comparisons between the two software were also made. The intensities of the spectra of the three soil samples obtained from fitting using both software were also compared. Quantitative analysis of the spectra of the three soil samples was performed and the concentrations of the elements obtained using both software were compared. A simple manual for spectrum fitting and quantitative analysis using PyMCA has been written for first-time users of the software (PyMCA); a flowchart for the manual has also been prepared. Comparison of intensities for Kα elements showed excellent agreement between the two software. The intensities of Lα elements obtained also compared favourably. Comparison of the concentrations of the elements in the spectra of the three soil samples showed excellent agreement. The X-ray spectrum fitting capabilities of PyMCA compare favourably with those of WinAxil. The development of WinAxil has been discontinued; therefore, PyMCA when appropriately adapted can be used in place of WinAxil for qualitative and quantitative analysis for spectra from X-ray tube excitation sources.

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE

INTRODUCTION

1.1 Background

Energy-dispersive X-ray fluorescence spectrometry (EDXRF) is the analytical method of choice for the scientific community because sample preparation is easy and it can detect multiple elements simultaneously. EDXRF is a non-destructive analytical technique and has reasonable sensitivity when compared with other analytical methods (Ofosu, 1995). It has found widespread application in fields such as geology where it is used in the characterization of rocks, ores and soils, in art, archaeology, forensics and has proven to be of value in the analysis of environmental samples. It basically involves qualitative analysis, determination of characteristic peak intensities and quantitative analysis. Qualitative analysis involves the identification of elements present in the sample; the determination of characteristic peak intensities is usually performed using an algorithm within a multichannel analyzer or by means of a software using least-squares fitting procedure whilst quantitative analysis involves the evaluation of intensities and the calculation of elemental concentrations from the evaluated intensities. Qualitative analysis is generally not a problem because every element in a sample is represented by a few characteristic X-ray lines. In particular, qualitative analysis is carried out by using an energy calibration obtained with standards for $K\alpha$ and $L\alpha$ X-ray lines. Before quantitative analysis can be done, the X-ray spectra from the sample must be fitted; this is achieved by using the appropriate software. The core of the software package for EDXRF spectrometric analysis consists of a spectrum analysis program that performs deconvolution of X-ray spectra by means of the optimization of a mathematical model to experimental data. Two software currently available for the determination of characteristic peak intensities are PyMCA and WinAxil.

PyMCA implements most of the needs for data visualisation and analysis of x-ray spectra. The code can be used as a visualisation tool as well as a full featured data analysis program. Spectrum analysis can be performed in a simple mode with automatic peak search or in an advanced mode where the code will search for a predefined list of elements. The code deals with data calibration, different background and peak shape models, theoretical peak ratios, pile-up and escape peak corrections. The code is written in Python but it is available in Unix and Windows. Despite being at its early stages, PyMCA and its fitting engine already implement most of the needs of x-ray fluorescence spectroscopy. It is fast (~1 second per complex spectrum with < 1 GHz processors), portable (it already runs on Solaris, Linux and Windows) and can be freely distributed. Current developments are focused on the implementation of alternative continuum algorithms (Sole, 2015).

WinAxil is written in C++, except for some analysis subroutines, which are written in FORTRAN. WinAxil is a modular software system. This system can read experimental X-ray spectra in IAEA format and old MS-DOS QXAS model files (.INP) as well as write standard results file (.ASR) for further processing with MS-DOS QXAS quantitative analysis package. It is used for spectra which have been produced as a result of sample excitation by X-ray tubes as well radio-isotopes.

1.2 Problem Statement

Most software for spectrum fitting is commercial; commercial software for EDXRF is expensive, mostly equipment-related and not user-friendly. In developing countries, owing to financial constraints, it is difficult to obtain commercial software for x-ray spectrum deconvolution. The only freeware (software that can be obtained free of charge online) that has been extensively used is the IAEA-sponsored software package QXAS which contains the software AXIL. It (AXIL) has proven to be as equally good as commercial software used for x-ray spectrum deconvolution. Development of QXAS has been discontinued and there still exist inherent difficulties and limitations in the software. There is the need to have freeware software that is as good as AXIL and compares well with results obtained from AXIL. Freeware software that has come up recently is PyMCA, which was originally designed for spectrum deconvolution of synchrotron radiation. However, not much information and guidance is available in the public domain for first-time and inexperienced users of PyMCA. This freeware software is platform-independent (Linux, Windows, MacOS X) and is freely available for non-commercial use developed by the European Synchrotron Radiation Facility (ESRF) (Sole et al., 2007).

1.3 Objectives

1.3.1 Main Objective

Since synchrotron radiation sources also produce x-rays, this thesis seeks to determine if the results of PyMCA used for spectrum deconvolution of x-rays from x-ray tubes will compare well with software specifically designed for tube-excited x-ray sources.

1.3.2 Specific Objectives

The specific objectives of this thesis are:

a. to compare spectrum fitting capabilities of PyMCA with WinAxil,

b. to investigate the conditions for adapting PyMCA to other X-ray excitation sources such as X-ray tubes.

1.4 Relevance and Justification

When successfully completed, this work will validate the extension and utilization of PyMCA to other X-ray excitation sources. This means PyMCA can be recommended for use in place of AXIL/QXAS in developing countries where they may be difficulty in acquiring commercial X-ray analysis software.

1.5 Scope and Definition

Before quantitative analysis is performed in EDXRF, spectrum fitting has to be done to extract X-ray peak intensities; this work is to compare the spectrum fitting capabilities of both PyMCA and WinAxil for different X-ray spectra. The excitation source to be considered will be X-ray tubes.

CHAPTER TWO

LITERATURE REVIEW

2.1 Linear Least Squares Fitting

If we have a number of experimental data points (x_i, y_i) and we want to fit these points by the best straight line:

$$f(x_i, a) = f(x_i, a_1, a_2) = a_1 + a_2 x_i \qquad (2.1)$$

least-squares fitting means that we obtain values a_1 and a_2 in such a way that the sum of the square differences between the measured data y_i and the calculated $f(x_i, a)$ is minimal. The objective is to find (a_1, a_2) such that

$$\chi^2(a_1, a_2) = \sum_{\mathrm{i}=1} \left(y_{\mathrm{i}} - f(x_i, a_1, a_2)\right)^2 w_{\mathrm{i}} \qquad (2.2)$$

The weights w_{i} are related to the counting statistics errors in y_i by the relation

$$w_{\mathrm{i}} = \frac{1}{y_i} \qquad (2.3)$$

χ^2 is the chi-square function of the two variables a_1 and a_2.

The minimum of χ^2 can be found by setting its partial derivatives with respect to a_1 and a_2 to zero and calculating the values of a_1 and a_2. So

$$\frac{\partial \chi^2}{\partial a_1} = 0 \ \text{ and } \ \frac{\partial \chi^2}{\partial a_2} = 0. \qquad (2.4)$$

This leads to a system of 2 equations with 2 unknowns

$$\sum_{i=1}(y_{\mathrm{i}} - a_1 - a_2 x_i)\, w_{\mathrm{i}} = 0$$

$$\sum_{i=1}(y_{\mathrm{i}} - a_1 - a_2 x_i) x_i w_{\mathrm{i}} = 0 \qquad (2.5)$$

Solving these two linear equations yields the values of a_1 and a_2. In matrix form, this linear system of order 2 can be written as:

$$\begin{bmatrix} \sum_{i=1} w_{\mathrm{i}} & \sum_{i=1} x_i w_{\mathrm{i}} \\ \sum_{i=1} x_i w_{\mathrm{i}} & \sum_{i=1} x_i x_i w_{\mathrm{i}} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} \sum_{i=1} y_{\mathrm{i}} w_{\mathrm{i}} \\ \sum_{i=1} y_{\mathrm{i}} x_i w_{\mathrm{i}} \end{bmatrix} \qquad (2.6)$$

$$\alpha \qquad a \quad = \qquad \beta$$

In general for any fitting function which is linear in its parameters a_j:

$$f(x, a_1 a_2, . . . , a_n) = a_1 f_1(x) + a_2 f_2(x) + . . . + a_n f_n(x) \qquad (2.7)$$

The optimum parameters for the fitting function to fit best the experimental data y_{i} are obtained by solving the linear system:

$$\alpha \;\; a = \beta$$

$$a = \alpha^{-1} \beta \qquad (2.8)$$

where

$$\beta_k = \sum_{i=1} y_{\mathrm{i}}\, f_{\mathrm{k}}(x_i)\, w_{\mathrm{i}} \qquad (j, k = 1, \;. . . n)$$

and

$\alpha_{j,k} = \sum_{i=1} f_j(x_i)\ f_k(x_i)\, w_{\mathrm{i}} \quad (j, k = 1, \ . \ . \ . n)$

2.2 Non-linear Least-Squares Fitting

If we consider the case when we want to fit a different set of points with a Gaussian plus a background,

$$f(x_i, a_1, \ . \ . \ . a_5) = a_1 + a_2\, x_i + a_3 \exp\left[-\frac{1}{2}\left(\frac{x_i - a_4}{a_5}\right)^2\right] \qquad (2.9)$$

where

a_1 *and* a_2 *are the background parameters*

a_3 *is the height of the peak*

a_4 *is the peak position(centroid)*

a_5 *is the width of the gaussian peak*

then the set of equations resulting from setting the χ^2 partial derivatives to zero with respect to its parameters is non-linear.

Therefore for a non-linear function, χ^2 must be minimized in an iterative way: changing the parameters at each iteration in an intelligent way and reducing χ^2 until reaching the minimum.

The iterative method implemented by WinAxil is based on the Marquardt algorithm. This method combines the linearization of the function with a gradient search to ensure convergence (Tchantchane, 1993).

2.2.1 Function Linearization Method

In the case of function linearization, $f(x_i, a) = f(x_i a_1, \cdots, a_n)$ at the kth iteration is expanded to first order Taylor expansion as a function of the parameters a_j that is:

$$f(x_i a_1, \cdots, a_n) = f\left(x, a_1^k, \cdots, a_n^k\right) + \sum_{j=1} \frac{\partial f(x, a_1^k, \cdots, a_n^k)}{\partial a_j} (a_j - a_j^k) \qquad (2.10)$$

setting $\left(a_j - a_j^k\right) = \delta a_j$

where a^k is the value of the parameter vector at iteration k.

The result is a function which is linear in the parameter increments δa_j $(1, \cdots n)$ such that

$$\chi^2 = \sum_{i=1} \left[\left(y_i - f(x_i, a^k)\right) - \frac{\partial f(x_i, a^k)}{\partial a_1} \delta a_1 - \cdots - \frac{\partial f(x_i, a^k)}{\partial a_n} \delta a_n\right]^2 w_i \qquad (2.11)$$

Following the standard method of linear least squares, the χ^2 can be minimized with respect to each parameter δa_j.

This yields a set of n simultaneous linear equations; in matrix form at the kth iteration, we have:

$$\begin{bmatrix} \sum_{i=1} \frac{\partial f}{\partial a_1} \frac{\partial f}{\partial a_1} w_i \cdots & \sum_{i=1} \frac{\partial f}{\partial a_1} \frac{\partial f}{\partial a_n} w_i \\ \vdots & \vdots \\ \sum_{i=1} \frac{\partial f}{\partial a_n} \frac{\partial f}{\partial a_1} w_i \cdots & \sum_{i=1} \frac{\partial f}{\partial a_n} \frac{\partial f}{\partial a_n} w_i \end{bmatrix} \begin{bmatrix} \delta a_1 \\ \vdots \\ \delta a_n \end{bmatrix} =$$

$$\begin{bmatrix} \sum_{i=1} (y_i - f(x_i, a^k)) \frac{\partial f}{\partial a_1} w_i \\ \vdots \\ \sum_{i=1} (y_i - f(x_i, a^k)) \frac{\partial f}{\partial a_n} w_i \end{bmatrix} \quad (2.12)$$

$$\alpha^k \qquad \delta a^k \qquad = \qquad \beta^k$$

The increments in each parameter δa_j are added to the previous estimation a^k to get a new one (a^{k+1}) at iteration $k+1$:

$$a_j^{k+1} = a_j^k + \delta a_j^k$$

The above procedure is repeated until the last iteration yields a small χ^2 or at some other stopping criteria (setting a maximum number of iteration or when only a very small change in the χ^2 value is obtained).

The disadvantage of this analytical method based on expanding the fitting function $f(x, a)$ with respect to its parameters a_j is that while it converges quite rapidly to the minimum χ^2 from points nearby, it cannot be relied on to approach the minimum with accuracy from an initial guess outside the region where the fitting function can be adequately represented. Therefore, this method is known to diverge if the initial guess of the fitting parameters is chosen far off from the minimum (Tchantchane, 1993).

2.2.2 Gradient Method

In the gradient method, we simply step off from the current trial value in the direction of the negative gradient of χ^2 controlling the step size carefully in order to assure the decreasing behaviour in the χ^2 thus ensuring convergence.

$$-\nabla\chi^2 = -\left(\frac{\partial\chi^2}{\partial a_1}, \cdots \frac{\partial\chi^2}{\partial a_n}\right)^T$$

$$-\nabla\chi^2 = \begin{bmatrix} \sum_{i=1}(y_i - f(x_i, a^k))\frac{\partial f}{\partial a_1} \\ \vdots \\ \sum_{i=1}(y_i - f(x_i, a^k))\frac{\partial f}{\partial a_n} \end{bmatrix}$$

The new trial (a^{k+1}) is obtained from the previous trial using:

$$a_j^{k+1} = a_j^k + \frac{1}{\lambda}\sum_{i=1}(y_i - f(x_i, a^k))\frac{\partial f}{\partial a_j} \qquad j = 1, \cdots, n$$

This method has a very slow convergence property and suffers very much as the search approaches the minimum.

2.2.3 Marquardt Algorithm

This algorithm combines the best features of both the gradient search algorithm and the method of linearizing the fitting function. The broad outline of the algorithm is that at the k^{th} iteration the matrix α and the vector β given above are constructed and the system:

$$\alpha^* \ \delta a \ = \beta$$

where

$$\alpha^* \ = \alpha \ (1 + \lambda)$$

is solved. If λ is small, ($\alpha^* \ \cong \alpha$) this algorithm reduces to the problem of function linearization. In the case of λ being large, the diagonal terms of the matrix α^* dominate which degenerates into n separate equations

$$\delta a_j = \left(\frac{\lambda}{\alpha_{jj}}\right) \beta_j \qquad (j = 1 \cdots n)$$

Thus the factor λ in Marquardt algorithm is used to perform an optimum interpolation between function linearization(as λ gets smaller) and gradient search direction(as λ gets larger): at each iteration the new χ^2 is tested; if it is reduced the iteration is said to be successful otherwise the same iteration is repeated while increasing the parameter λ until the new χ^2 is decreased. After each successful iteration λ is decreased in order to speed up the convergence.

2.3 Function Used for Fitting X-ray Spectra

The fitting function is a model which must represent well the observed spectrum. The result of such a fit is an optimal estimate of a number of parameters (peak areas, energy, resolution, and background). An ideal fitting function must take into account every detail in the spectrum such as the peak shape, the form of the background, deviation from the pure Gaussian shape, peak interference, absorption effects and anomalies in the relative intensities. However, on the other hand, the fitting function describing accurately the spectrum must also be chosen in such a way to minimize the number of variable parameters required to be determined (Tchantchane, 1993).

2.3.1 Peak Shape

The primary contribution to the photo peak in a semiconductor detector is a Gaussian distribution. Therefore the spectral contribution of an element can be considered as composed of a number of Gaussians (nG):

$$Y_i = \sum_{j=1}^{nG} H_j \exp\left[-\frac{1}{2}\left(\frac{x_i - P_j}{\sigma_j}\right)^2\right] = \sum_{j=1}^{nG} H_j G(i, P_j, \sigma_j) \qquad (2.13)$$

where

$Y_i = \ yield\ of an\ element$

$nG\ = number\ of\ X - ray\ lines\ corresponding\ to\ the\ element$

$P_j\ = peak\ position$

$H_j = peak\ height$

$\sigma_j = peak\ width$

It should be noted that three parameters per peak have to be optimized and peak interferences and overlaps cannot be resolved.

A considerable reduction in the number of parameters is obtained by using the relations between energy E on the one hand, and peak position P and width σ on the other hand. The position and width are not determined for every peak separately. Instead, the energy and the resolution calibration constants (c_1 to c_4) are used to determine them. The position and width are determined according to the relations:

$$P_j = c_1 + c_2 E_j \tag{2.14}$$

$$\sigma_j^2 = c_3 + c_4 E_j \tag{2.15}$$

where

E_j *is the characteristic energy corresponding to the line* j

The inclusion of this relation causes a drastic reduction in the number of unknowns and it also brings a substantial increase in the accuracy of the least-squares fit since the optimization of the calibration parameters allow to correct for small variations in energy and calibration.

A further reduction in the number of parameters is obtained by treating a group of lines as a unit; that is, only one net peak area (H) is determined for each group. Lines within a group have fixed intensity ratios during the fit. This is possible because of the relative intensities that exist between K lines of a certain element (K_α and K_β) or between the L lines (L_α,L_β and L_γ).Thus the yield related to an element is best described as:

$$Y_i = H \sum_{j=1}^{nG} R_j \exp\left[-\frac{1}{2}\frac{\left(x_i - c_1 - c_2 E_j\right)^2}{c_3 + c_4 E_j}\right] = H \sum_{j=1}^{nG} R_j \; G(i, E_j) \qquad (2.16)$$

where

R_j *is the relative intensity of the j line(read from the AXIL LIBRARY)*

Further, grouping a series of lines is an optimum way to resolve overlapping between K_α and K_β lines of neighbouring elements or during the interference between L and K lines of heavy and light elements respectively.

2.3.2 Tailing

Deviation from photo peak Gaussian shape can be important, particularly in the lower region (< 4 keV) and may give rise to large errors if not considered. This is due to many factors (Auger transitions, incomplete charge collection in the detector). In AXIL, a numerical correction is added to each channel i, for intense peaks (K_α). For each element, a total array is saved in the AXIL library and added to each channel i. This correction is multiplied by the height of the K_α peak during the fitting process (Tchantchane, 1993).

2.4 Background Component of the Fitting Model

Various approaches have been employed to deal with the continuum background. The first and most common approach is to use a suitable analytical function $B(i)$ with the smallest possible number of free parameters to simulate the background; various such models have been proposed. When only a limited spectral region is being fitted, (< 3 keV wide), a linear polynomial is sufficient to model the background shape:

$$B(i) = \sum_{j=0}^{n} a_j (E_i - E_0)^j \qquad (2.17)$$

where

$B(i)$ *is the background counts in channel i,*

E_i *is the energy of channel i*

E_0 *is a suitable reference energy*

To model the more complex background shape of an entire spectrum, an exponential background polynomial can be employed, optionally modified to account for attenuation effect (Fiori et al, 1981, van Espen et al, 1977):

$$B(i) = a_0 + \exp \sum_{j=0}^{n} a_j (E_i - E_0)^j \qquad (2.18)$$

In this case, some of the parameters can be kept constant while others can be varied. The complete fitting model is then obtained by adding the background model $B(i)$ to the peak model $Y(i)$.

The top-hat filter technique is frequently used in X-ray emission analysis with electron microprobes (Schamber (1979).In this approach, a simple mathematical filtering operation, is applied which effectively removes low-frequency components, i.e. the continuum background from raw spectral data. The resulting spectrum is then fitted with a similarly-filtered model spectrum that contains only peaks of pure element standards.

2.5 Background Estimation of X-ray Spectra

Several approaches are possible for the estimation or modelling of the continuum background in X-ray spectra, depending on the kind of spectra being processed and the kind of spectral information one is interested in. One, the polynomial modelling technique is generally applicable and flexible; however, it has several disadvantages. Since the parameters of the background function are determined simultaneously with the rest of the parameters of the fitting model during the chi-squared-minimalisation procedure, this method of background estimation cannot be employed in situations where fitting of the spectra is unnecessary, unwanted or impossible. Also, as the background shape is the result of a large number of different processes of varying nature(such as scattering, bremsstrahlung, incomplete charge collection, peak tailing, absorption), often smooth mathematical functions such as polynomials are either unsatisfactory to model the complex background shape accurately, resulting in either overestimation or underestimation of some of the peak areas. Moreover when non-linear modelling functions are used, the fitting algorithm must be provided with approximately correct starting values of the non-linear coefficients in order to converge to a meaningful solution.

Many of the limitations cited above do not arise if some iterative mathematical technique is used to remove peaks and progressively reduce the spectrum to the background continuum. The resulting continuum can then be used along with the peak model in conducting the least-squares fit, or it can be subtracted from the spectrum prior to fitting. Recently, two non-polynomial background estimating methods have been proposed

(Ryan et al, 1988) both of which only require the spectral data as input, without a prior knowledge of the type of excitation used or the location of the peaks.

2.5.1 The Iterative Stripping Algorithm

By iteratively suppressing the channels that contain peak information, the shape of the background can be determined.

The algorithm involves three steps. First, the spectrum is compressed, i.e. the number of channels is reduced while all the relevant information concerning the background shape is preserved. Secondly, the actual iterative stripping procedure is executed. Finally, the resulting compressed background is re-expanded, resulting in the background spectrum.

Since the information concerning the continuous background is uniformly distributed throughout a spectrum, whereas data regarding the characteristic peaks are concentrated in a number of selected regions, it is unnecessary to involve all channels of the spectrum in the background estimation process. Since the performance of the stripping algorithm depends on the peak width, a selective spectrum compression is performed prior to the actual peak stripping. For X-ray spectra, the compressed spectrum is typically obtained by selecting every second channel of the original spectrum in the energy region below 2 keV. Between 2 and 5 keV, one in every five channels is retained while above 5 keV, only one out of ten channels is used. More channels are selected in the low-energy regions of the spectrum where the background may be complex in shape.

Near a peak top, the first derivative of a spectrum changes sign abruptly, resulting in a large negative value of the second derivative, while in parts of a spectrum where no peaks occur, only small values of the first and second derivatives are obtained. The stripping algorithm takes advantage of this difference by calculating for each channel i in the spectrum or in a specified region of interest (ROI) the mean channel content M_i:

$$M_i = \frac{Y_{i-1} + Y_{i+1}}{2} \tag{2.19}$$

and by applying the following substitution:

$$\begin{aligned} Y_i^{new} &= M_i \text{ if } Y_i > M_i \\ Y_i^{new} &= Y_i \text{ if } Y_i \leq M_i \end{aligned} \tag{2.20}$$

After the iterative stripping, the actual background spectrum is obtained by copying all the channel contents of the compressed background estimate to their approximate places in the expanded spectrum and calculating the contents of the interlining channels by means of linear interpolation.

2.5.2 Parabolic Envelope Approximation

With this method, the background is modelled as an envelope of a family of parabolas. Since the background is a slowly varying structure, it can be considered to be composed of objects which are several times broader than the X-ray peaks. Moreover, as the amplitudes of X-ray peaks are always positive, the value of the background can nowhere exceed the value of the spectrum itself. Thus, in each channel i of the spectrum, a downwards concave parabola

$$D(j) = H(i)\left[1 - \frac{(i-j)^2}{x^2(u+v\times i)}\right], \tag{2.21}$$

just touching the spectrum from below can be considered. $H(i)$ represents the amplitude of the parabola (i.e. its height) and x, u and v are parameters by which the width of the parabola can be adjusted. Using the condition that for each channel j, the parabola's $D(j)$ can nowhere exceed the spectrum $S(j)$ (which represents the number of counts in channel j of the original spectrum, after elimination of statistical fluctuations), their amplitudes $H(i)$ are determined:

$$H(i) = \min\left[S(j)/\left(1 - \frac{(i-j)^2}{x^2(u+v\times i)}\right)\right] \tag{2.22}$$

and a first approximation to the background $B(i)$ can be obtained as the envelope of these parabolas:

$$B(i) = \max\left[H(i)/\left(1 - \frac{(i-j)^2}{x^2(u+v\times i)}\right)\right] \tag{2.23}$$

If calculated in this way, the background $B(i)$ never exceeds the spectrum, however, under the peaks it sometimes appears to be overestimated, especially if these peaks are situated on a slope (Tchantchane, 1993).

2.6 Background compensation methods

X-ray fluorescence (XRF) is a well-established surface analysis technique (Ang et al., 2003, Schlotz et al., 2000, Janssens et al. 1986). By measuring the fluorescence spectral peak positions and the net peak areas, it is possible to determine which elements are in the sample and the contents of these elements. However, the background can affect the measurement. During analysis, two main sources may produce a background signal in the spectrum. The first is scattering, including Compton scattering and Rayleigh scattering, which is caused by the interactions between the continuous spectrum of the X-ray tube and the sample. The second is the interaction between the characteristic X-rays of the sample and the detecting instrument, including internally generated instrument noise. Especially in the detection of trace elements, the background may have a significant effect on the detection limit and precision (Ang et al., 2003). Therefore, the background must be estimated and removed from the spectrum. This task is generally simple if the amplitudes of the peaks are bigger than the background below them and the latter is slow varying. When these conditions are not respected the removal of the background is more complicated.

Background modelling refers to the description of the background shape by some mathematical function; this function forms part of the fitting model that describes the complete spectrum (i.e. background continuum and photopeaks). During the spectrum evaluation, the parameters of the background function are optimized together with the other parameters in the fitting model in order to obtain the best match between model and spectral data. In contrast, background estimation involves the a priori calculation of an approximate background shape **b** by applying a suitable stripping or filtering algorithm on the original spectral data **y**, which are then fitted with an appropriate fitting model containing the estimated background shape as fixed part of the model. This is equivalent to using the net spectrum **n = y - b** for the peak area determination of the characteristic lines (Vekemans et. al, 1994).

The essential difference between background modelling and estimation procedures is that the latter class of algorithms produces a background shape that is not "fine-tuned" during the subsequent stages of the (iterative) spectrum evaluation procedure.

The background modelling procedures do show this behaviour, which, provided the fitting model is complete (i.e. accounts for all peaks in the spectrum), allows the background shape to be optimally adapted to all non-characteristic information in the spectrum. This feature obviously becomes a disadvantage when the applied model is incomplete, e.g. in the case when escape peaks or characteristic lines have been omitted from the model. In order to maximize the overall goodness-of-fit, the background will be distorted and will cause it in the region of the omitted peak to be higher than in reality. Usually this causes gross underestimation of the weak peaks in the same spectral area. On the other hand, the modelling approach to background compensation can accommodate in a more flexible way small systematic differences between experimental data and the spectrum model such as is the case with tail effects and non-Gaussian peak shapes(Vekemans et. al, 1994).

Spectra generally consist in peaks and noise superimposed on a background. This background, or baseline, which can be either flat, linear with a positive or negative slope, curved, or a combination of all three, is mainly characterised by the fact that it does not vary as quickly as the peaks do. In most software the background is estimated by a least-squares polynomial fitting performed on a user defined subset of points which should belong to the background. Providing that the points are correctly selected, the fitting yields satisfactory results. This can be attributed to the ability of the polynomial model to represent a wide class of backgrounds. The now-classic method of stripping peaks to fit the background consists of comparing a data point with the average of the two adjacent points and replacing the current point by the smaller value. After several iterative calculations, only a smooth background curve remains.

In addition, there are many other good methods, such as orthogonal polynomial decomposition, the least squares method and wavelet analysis. The orthogonal polynomial background estimation method, introduced by Steenstrup, smoothens noise errors. Nevertheless, the above methods all require iterative calculation, and the iterations are different for different spectra. They are not suitable for automatic batched analyzing spectrum. A new background subtraction method for energy dispersive X-ray fluorescence spectra using a cubic spline interpolation has been proposed (Yi et. al., 2015).This method has been tested on spectra measured from a coin, and it has also been tested on an existing sample spectrum. Compared with Fig. 2.1(a), Fig. 2.1(b) shows that the method described above can properly subtract the background. It performs better than stripping peaks.

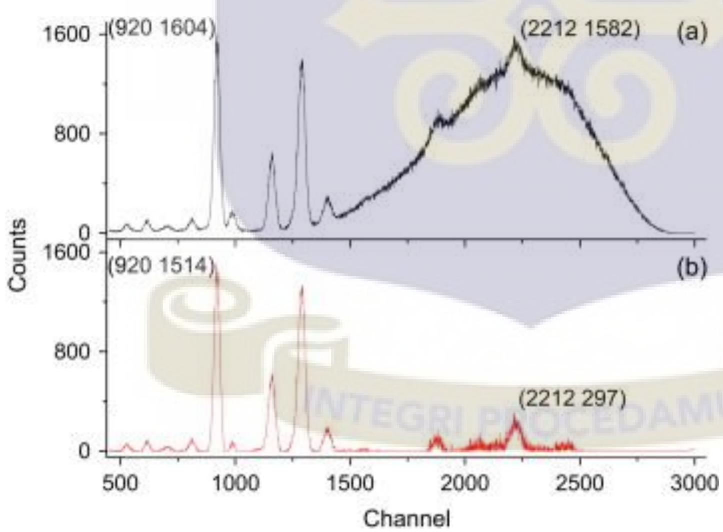


Fig.2.1 (a) Original sample EDXRF spectra (b) Subtracted spectra by the cubic spline interpolation method.

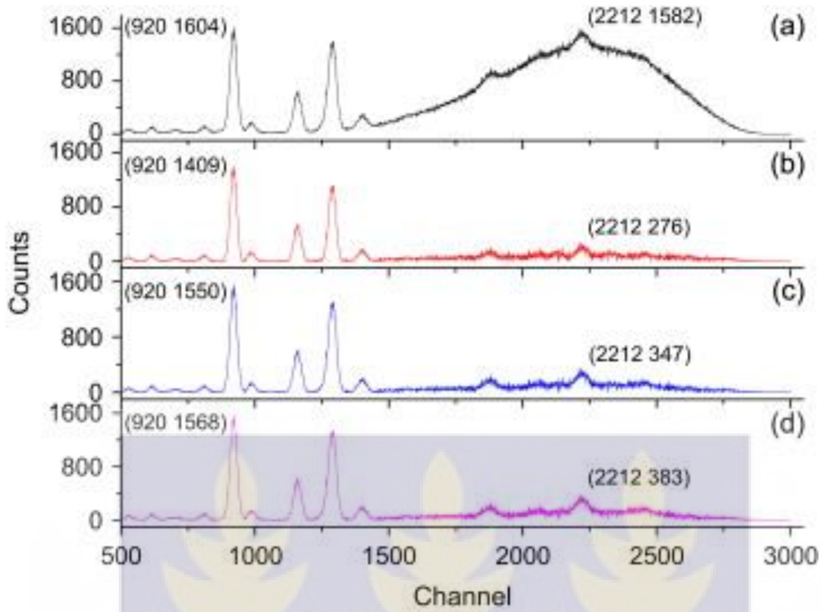


Fig.2.2 (a) Original EDXRF spectrum (same as in Fig. 2.1). (b)– (d) Subtracted spectra by stripping peaks with 1000, 3000 and 5000 times iterative computations.

In addition, the method has been compared with the stripping of peaks. Compared with Fig. 2.2, Fig. 2.1 shows that the effects of the two methods are essentially identical when the number of iterations in the peak stripping method is about 3000. It performs better than stripping peaks.

This method can be applied to analyses that require a short time, such as analyzing an X-ray fluorescence two- dimensional scanning image with 100 x 100 pixels. This method is not suitable if the number of counts is below 10 per channel and in cases where the escape peak has to be considered.

Another method for the subtraction of the background using dual-tree complex wavelet transform has been proposed (Zhao et. al 2014). Compared with real wavelet transform, the proposed method has some attractive properties, including a smooth, non-oscillating, and nearly shift-invariant magnitude with a simple near-linear phase encoding of signal shifts. Therefore, it outperforms real wavelet transform to decompose the background into low-frequency components.

The commonest and probably most appropriate method is the orthogonal polynomial decomposition, used, for example, in the AXIL program (Steenstrup et. al 1981). It fits the background with a linear combination of mutually orthogonal polynomials. With respect to a linear combination of ordinary (non-orthogonal) polynomials, this method presents a higher stability with high-order polynomials, allowing a better approximation of the continuum. It can be expressed as

$$y_i = \sum_{j=0}^{n} a_j p_j(x_i) \qquad (2.24)$$

where y_i is the continuum approximation, p_j the orthogonal polynomials and $a_j$are the proportionality coefficients determined by a least-squares minimization. Although, this method works well, in some cases the quality of background estimation is not that good. To overcome this problem, a different approach has been proposed by Brunetti (2000). It is based on an approximation of the continuum with a projection onto a convex set, the spectrum data. A set is said to be convex if for any two elements of the set, a convex combination of them is again in the set. It is possible to add several constraints to the solution, such as positiveness or boundedness of data and compactly supported data. In this case, all of them can be imposed because, due to finite energy values, a spectrum has always positive bounded values and it presents a compact support. To project an X-ray spectrum onto a convex set is to compare the values estimated with the spectrum data: if the estimation of the continuum has amplitudes higher than the real spectrum values, then they will be locked to the spectrum values. The problems are how to represent the fit and make it converge to the best solution, using only weak constraints or hypothesis on the nature of the signal. A typical approach is to use the information relative to the full-width half-maximum (FWHM) of the peaks. This can be considered a strong constraint because it cannot permit analysis of an unknown spectrum. The results obtained have a quality higher or, at least comparable, to other methods reported in other papers, especially if the spectrum that must be analyzed has very low amplitude peaks compared to the continuum. This method, compared to the well-known orthogonal polynomial method, shows a higher precision for the estimate of the amplitudes of the peaks. Both this method and the polynomial approximation method work under the assumption that the background is varying slower than the peak areas.

Wang et al (2012) have applied the method of Fourier transform to the problem of background subtraction. It is less affected by changes in the background than linear background subtraction.

One algorithm used for background estimation is the Statistics-sensitive Non-linear Iterative Peak-clipping (SNIP) algorithm. The SNIP algorithm is structured into three phases. The first phase applies the low statistics digital filter, and transforms the channel count $y(x)$ in channel x according to $z = G(y)$, where

$G(y) = log\{log(y + I) + I\}$, in order to compress the dynamic range of the channel counts. The second phase is a multipass peak clipping loop which replaces each channel value $z(x)$ with the lesser of $z(x)$ and the mean $Z[x, w(x)]$, where

$Z(x, w) = [z(x + w) + z(x - w)]/2$ and the scanning width $w(x)$ is fixed to twice the X-ray peak FWHM characteristic of channel x. This approach is independent of detector energy calibration. Only 24 passes through this loop are necessary to produce acceptable background spectra. For the final few passes (typically 8) the scanning width is progressively reduced after each pass. This is done to reduce the incidence of oscillations in the background, and eliminates the need to smooth the background afterwards. The final phase transforms the results back to channel counts, $y = G^{-1}(z)$ (Ryan et al, 1988). The basic SNIP algorithm automatically includes the linear component under the peak in the estimated background. A complete spectrum consists of the full energy peak together with escape and sum peaks as well as Compton edges. As such the background of the spectrum is complicated, but the SNIP algorithm can remove these using several clipping filters (Morhac et al, 2008).For example, the linear background is removed by a clipping filter of the second order whilst the Compton edge can be removed by a clipping filter of the fourth order. Unlike other background compensation methods, the SNIP method implicitly determines peak regions and peak-free regions.

Park et. al (2013) have proposed an efficient background correction method for EDXRF spectra. The method has two steps of background modelling and background correction. It is based on the basic concept which differentiates background areas from the peak areas in a spectrum and the SNIP algorithm, one of the popular methods for background removal, is used to enhance the performance. After detecting some points which belong to the background from a spectrum, its background is modeled by a curve fitting method based on them. And then the obtained background model is subtracted from the raw spectrum. This method has been shown to give better results than some of the traditional methods, while working under relatively weak constraints or prior information.

Smoothing the spectrum is a necessary operation to get rid of noise. A Savitzky and Golay algorithm and moving average filter are known methods in the literature (Jenkins et al., 1981, van Grieken and Markowicz, 1993 and Savitzky and Golay,1964). The Savitzky and Golay method performs a local polynomial regression on a distribution of a certain width to determine the smoothed value for each point. Other smoothing methods, such as moving average and Gaussian filters, involve a convolution with the spectrum. The filters have a gain equal to one, and are symmetric in relation to the points to smooth. This produces results from which detection of peaks is more precise.

The strip background is probably PyMCA's most popular background model (Sole). In its simplest implementation it is just an iterative procedure depending on two parameters. These parameters are the strip background width w, and the strip background number of iterations. At each iteration, if the contents of channel i, $y(i)$, is above the average of the contents of the channels at w channels of distance, $y(i-w)$ and $y(i+w)$, $y(i)$ is replaced by the average. At the end of the process we are left with something that resembles a spectrum in which the peaks have been "stripped". This implementation would give a baseline passing by the lower band of the statistical noise associated to the spectrum. This problem can be solved by performing a smoothing prior to calculating the background. The parameter associated with the smoothing is the strip background smoothing width. All these parameters are accessible in the FIT configuration tab of the configuration window (Sole).

The PyMCA strip background enables one to use anchors when backgrounds with a high curvature are encountered such as the one in Fig. 2.3. Anchors are simply channels that will remain fixed during the stripping procedure. That feature is available from PyMCA version 4.0.9 on. Since the energy calibration can change during the fitting procedure, anchors have to be given in channels and not in energy (Sole).

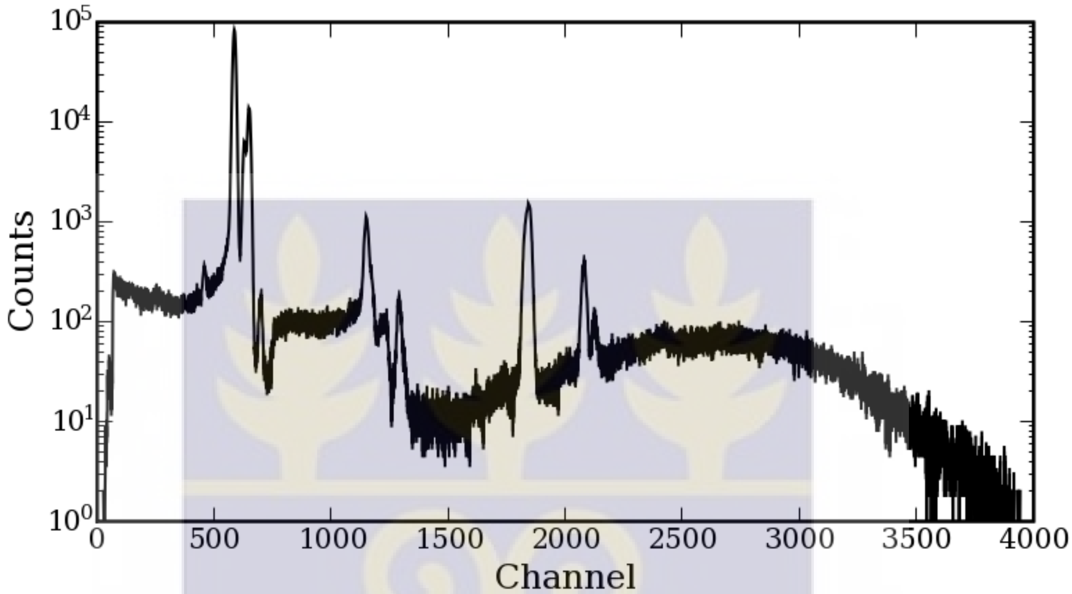


Fig. 2.3 Background with a high curvature

2.7 Spectrum Evaluation

The aim of spectrum evaluation is to obtain unbiased, minimum variance, estimates of the net intensities (peak areas) of the fluorescence lines. Fitting an appropriate mathematical function to the spectral data can do this. The parameters of this function are chosen so that the weighted sum of squared differences between the measured spectral data y_i and the fitting function $y(i)$ are minimal. This object function is known as chi-square:

$$\chi^2 = \sum_{i=n_1}^{n_2} \frac{1}{w_i} [y(i) - y_i]^2 \qquad (2.25)$$

The fitting function generally consists of two parts, one that describes the continuum and one that deals with the characteristic lines (van Espen and Lemberge, 2000):

$$y(i) = y_{cont}(i) + \sum_{i\ Elements} A_j \left[\sum_{k\ lines} R_{jk}\ P(i, E_{jk})\right] \tag{2.26}$$

The inner summation runs over the number of K / L lines of each element with R_{jk} the relative contribution of line k of element j. The outer summation runs over all elements with the linear parameter A_j being the total intensity of element j. $P(i, E_{jk})$ is a function that describes the shape of the x-ray line with energy E_{jk}. A typical implementation of $P(i, E_{jk})$ consists of a Gaussian

$$P(i, E_{jk}) = \frac{Gain}{\sigma_{jk}\sqrt{2\pi}} exp\left[-\frac{1}{2}\left(\frac{E(i) - E_{jk}}{\sigma_{jk}}\right)^2\right] \tag{2.27}$$

in which the width and position obey the energy and resolution calibration functions of the spectrometer:

$$E(i) = Zero + Gain\ \times i \tag{2.28}$$

$$\sigma_{jk} = \left[\left(\frac{Noise}{2\sqrt{2\ln 2}}\right)^2 + \varepsilon\ Fano\ E_{jk}\right]^{1/2} \tag{2.29}$$

The fitting function is linear in the net peak area parameter A_j, but nonlinear in the parameters *Zero*, *Gain*, *Noise* and *Fano* so that a nonlinear least-squares procedure is required. In successive approximations the values of the parameters are altered until the minimum in χ^2 is reached (van Espen and Lemberge, 2000).

To improve the spectrum evaluation procedure, a mathematical model that takes into account the deviation from the Gaussian peak shape must be used. Such a model includes a step and tail function to describe the low energy tailing due to incomplete charge collection in the detector.

$$P(i, E_{jk}) = G(i, E_{jk}) + f_s S(i, E_{jk}) + f_T T(i, E_{jk}) \tag{2.30}$$

in which $G(i, E_{jk})$ is the Gaussian and f_s and f_T are (linear) parameters representing the fraction of photons that gives rise to a step and tail contribution respectively. The step and tail functions can be written as:

$$S(i, E_{jk}) = \frac{Gain}{2E_{jk}} erfc\left[\frac{E(i) - E_{jk}}{\sqrt{2}\sigma}\right] \tag{2.31}$$

$$T(i, E_{jk}) = \frac{Gain}{2\gamma\sigma\, exp\left[-\frac{1}{2\gamma^2}\right]} exp\left[\frac{E(i) - E_{jk}}{\gamma\sigma}\right] erfc\left[\frac{E(i) - E_{jk}}{\sqrt{2}\sigma} + \frac{1}{\sqrt{2}\gamma}\right] \tag{2.32}$$

The functions are normalized, i.e. the integral of the function over the spectrum equals unity. The parameter γ is an extra nonlinear parameter, modelling the width of the tail. To keep the number of parameters to be estimated during the least-squares fit as low as possible, f_s , f_T and γ are expressed as functions of the energy. This functional relation is fitted rather than the value for each peak individually. Spectrum evaluation based on this fitting model gives more reliable results than simple Gaussians (van Espen and Lemberge, 2000).

In general, an XRF spectrum consists of peaks, corresponding to the various elements in the sample, superimposed on a background, which comes from x-ray scatter and detector effects. It is the job of "spectrum processing" to effectively remove the signal (i.e., net peak intensity) from the noise (i.e., the background and artifact peaks). There is an additional complication caused by peak overlap, which usually means that one cannot simply integrate regions around a peak and arrive at an accurate intensity, for a given element (and line).

The number of counts in a particular peak in any EDXRF spectrum is directly proportional to the acquisition time; hence, the relative intensities are dependent only on the relative proportions of the elements in the sample and on their relative detection sensitivities.

Requisite to converting observed peaks into quantitative results accurately is the isolation (deconvolution) of overlapping peaks from one another. Since such overlap alters the observed amplitude and area of the respective peaks, subsequent efforts at correlating peak amplitude or area with the relative concentration of the element may result in gross errors if these peaks are not accurately deconvoluted prior to quantification. Two basic techniques have been used to separate or deconvolute overlapping spectral peaks. These are resolution enhancement techniques (Bertin, 1975) and digital filter-correlation techniques (Schamber, 1976). Some resolution enhancement techniques assume that the observed spectrum is composed of a useful signal plus noise. This observed signal is considered to be the product of the transformed true emission spectrum multiplied against a transformed system response function. The noise can be removed by using an exponential roll-off filter, and the true emission spectrum is then computed by taking the inverse Fourier transform of the transformed observed spectrum divided by the transformed system response function. While the resolution enhancement technique serves to locate and separate partially overlapping spectral peaks, total separation of more severely overlapping peaks must be accomplished by another method since true peak amplitudes and standard deviations may still be obscured. Digital filter-correlation techniques involve the application of a digital filter to both the observed spectrum and to a series of reference files containing peaks derived from real or artificially generated elemental standards for each element present in the observed spectrum. The digital filter supposedly filters out the background and high frequency noise. Deconvolution can then be performed by computing (in the transformed domain) the ratio of the observed intensity of the reference peak to the same peak in the complex spectrum. Peak areas may then be calculated directly from the reference spectrum.

While the digital filter technique is capable of yielding excellent results from ideal reference peaks collected under identical conditions, when such reference files are not available the accuracy of the technique is dependent upon the operator estimates of the amplitude and standard deviation of an "ideal" reference peak. Lam et. al (1979) have applied an optimization procedure called simplex to resolve the true area of overlapping peaks such as those frequently encountered in energy dispersion analysis spectra. This technique overcomes many of the limitations of present deconvolution techniques such as resolution enhancement and digital filter-correlation and has the added advantage of not involving the use of partial derivatives as encountered in gradient-type optimization techniques. This procedure has been applied to a series of artificially generated overlapping Gaussian peaks which are close analogs to actual peaks. The results of this study have shown that this function converges and the correct area of those overlapping peaks can be estimated by this deconvolution.

In energy-dispersive x-ray fluorescence analysis, the estimation of the net area of the peaks is a primary requirement. This task requires a non-linear fitting of the peaks. The most common procedures are based on the Marquardt–Levenberg technique. This technique generally works well only when the spectrum is perfectly known, i.e. all the peaks are recognized. Moreover, it is sometimes difficult to introduce constraints on the fit due to peak shape or other physical properties. Brunetti et. al (2001) have proposed a new technique, based on a set of genetic operators. It works well even when the knowledge of the peaks is incomplete and it also allows one easily to introduce constraints. The results obtained with this algorithm are generally superior with respect to a standard implementation of a Marquardt–Levenberg procedure.

The only drawback is the speed of convergence, which is slower than in the Marquardt–Levenberg technique. Genetic algorithms are probabilistic procedures based on a simulation of life rules. They search for the best-adapted solution with respect to an environment, i.e. the model that can be represented. The genetic algorithms are based on a set called as a population, of possible solutions of the problem. Each member of the population is called a chromosome and can be formed by one or more genes, whose values or states are called alleles. Each gene is a parameter of the model that must be fitted. The evolution of the population is regulated by a set of evolution functions. Across the generations, each member is evaluated by a fitness function that produces an estimate of the adaptation, i.e. of the quality of the solution determining its probability of survival in the next generation.

Bennun et. al (2002) have come up with a new procedure for the identification and quantification of spectral signals based on the method of maximum likelihood is presented. It is particularly appropriate for the analysis of signal interference or cases of a small signal in a sizeable noisy background. The formulation yields precise signal intensity values and has advantages in its estimation of detection limits over established signal processing methods. The method developed is applicable to spectroscopic techniques where net counts are proportional to acquisition time, such as particle induced x-ray emission (PIXE) and other x-ray fluorescence spectroscopic methods.

Liu et al. (2014) have presented a maximum a posteriori (MAP)-based algorithm for the band overlap problem. In the MAP framework, the likelihood probability density function (PDF) is constructed with Gaussian noise assumed, and the prior PDF is constructed with adaptive total variation (ATV) regularization. The split Bregman iteration algorithm is employed to optimize the ATV spectral deconvolution model and accelerate the speed of the spectral deconvolution. The main advantage of this algorithm is that it can obtain peak structure information as well as suppress noise simultaneously. Simulated and real spectra experiments manifest that this algorithm can satisfactorily recover the overlap peaks as well as suppress noise and are robust to the regularization parameter.

There are many different algorithms and approaches to detect peaks in data. The main problem lies in distinguishing true peaks from statistical fluctuations, Compton edges and other undesired spectrum features. The estimation of widths of peaks regions is tightly connected with the localization and identification of peaks in the spectrum. However, a great majority of them is based on convolution operations. The use of the convolution method to locate the peaks is a well-established approach and is utilized in many of the available algorithms. It removes statistical fluctuations and is insensitive to linear background under Gaussians.

Without loss of generality assume that a peak is perfectly Gaussian-shaped given by

$$G(x) = A\exp(-\frac{(x-a)^2}{2\sigma^2}) \qquad (2.33)$$

where x is the channel number and A, a and σ are peak intensity, peak position and standard deviation determining the peak width, respectively. A Gaussian product function that is derived from four successive points has been proposed:

$$p_1(x) = [G(x)G(x-1)]/[G(x-2)G(x+1)] \qquad (2.34)$$
$$= \exp(2/\sigma^2)$$

One can observe that this function depends only on σ. However, it has to be calculated for all possible channel values x. When looking at the function, for peak-free regions $\sigma \rightarrow \infty$ and consequently $p_1(x)$ =1. Otherwise, in peak regions $p_1(x) = constant > 1$.This condition can be utilized for the identification of the peaks and together with the equation immediately above also for the determination of peak width and thus the peak regions (Morhac,2009).

There are a growing number of applications which involve precise measurements of x-ray line intensities. Also, there is growing interest in measurements of the K-series x-ray spectra of heavy elements excited with synchrotron radiation or high-voltage x-ray tubes. The use of K-series x-ray lines of heavy elements instead of L-series radiation may yield better detection limits and reduce the peak overlap in EDXRF spectra. The peak profile of an x-ray line of a low-Z element can be approximated with good accuracy by a Gaussian function. In the case of high-Z elements and the K-series radiation peaks, the Gaussian function does not follow accurately enough the pronounced tail structure on both sides of the peak profile. When such a structure is approximated by a Gaussian profile, it results in a high chi-squared value and a biased peak area due to the discrepancy between the actual peak shape and the profile of the Gaussian function.

Wegrzynek et. al (2001) have proposed the evaluation of the energy-dispersive x-ray spectra of such high-Z elements using Voigt peak shape profiles.

The Voigt profile is obtained by convolution of the natural line profile given by a Lorentz function with the Gaussian detector response function. The Voigt profile reproduces the Gaussian shape in the neighbourhood of the peak centre and follows the Lorentz long tail structures on both sides of the peak. The need to introduce the Voigt peak profile depends on the ratio of the natural line width, Γ, to the energy resolution of the spectrometer at the peak position (represented by the Gaussian standard deviation of the peak, σ),Γ/σ.

CHAPTER THREE

METHODOLOGY

3.1 Software

The software used were WinAxil and PyMCA version 4.7.4.

3.2 WinAxil

The K line peaks of Ti and Mo from IAEA standard spectra of Ti were used for the energy calibration.

The K line peaks of titanium (Ti), iron (Fe), cobalt (Co), nickel (Ni), zinc (Zn), strontium (Sr) and yttrium (Y) were fitted with a linear polynomial of order 3. The following Lα lines were fitted: dysprosium (Dy), samarium(Sm), cerium (Ce), tantalum (Ta), tungsten (W), mercury (Hg), and uranium (U). Dy, Sm, Ce, Ta, W, Hg and U were with a linear polynomial of order 3.

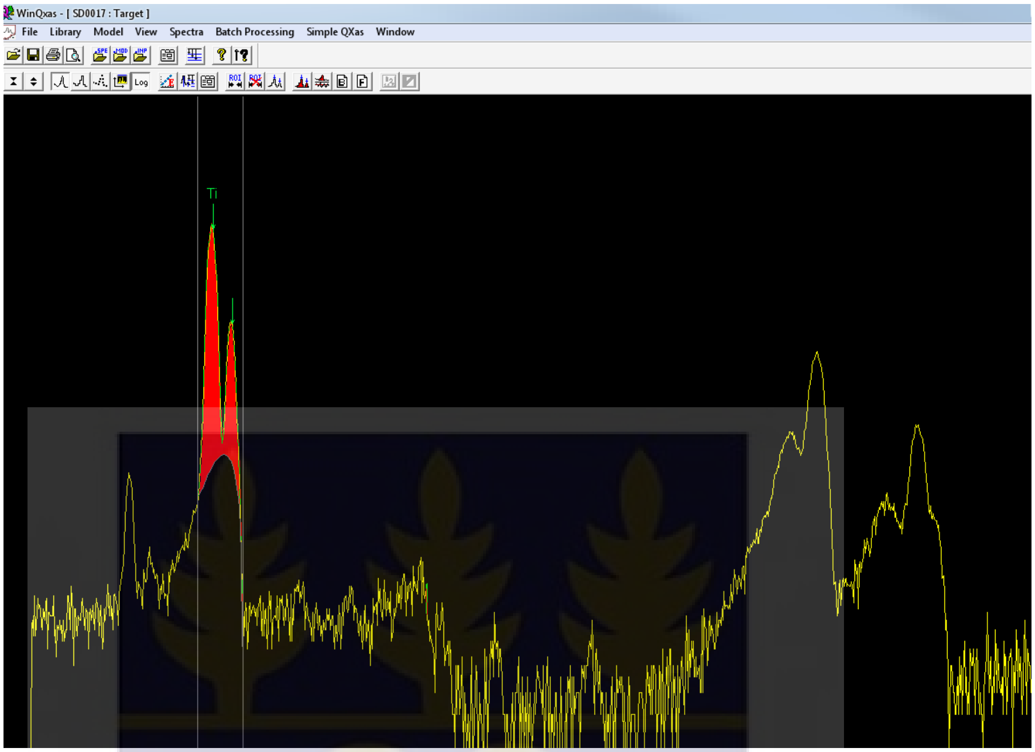


Fig. 3.1 Fitted Ti standard spectrum with WinAxil

3.3 PyMCA

Energy calibration was performed with Ti standard spectrum and saved in a calibration file. This file was then loaded into the software when fitting of the standard spectra was to be done.

Ti was fitted with linear polynomial of order 3. Fe was fitted with a linear polynomial of order 3 with a short tail function. Co was fitted with a linear polynomial of order 3 with a short tail function. Ni was fitted with a linear polynomial of order 3 with a short tail function. Zn was fitted with a linear polynomial of order 3 with a short tail function. Sr and Y were fitted with a linear polynomial of order 3 with a short tail and step functions. Sm was best fitted with a linear polynomial function of order 3 with short and long tails. Dy was best fitted with an exponential polynomial of order 3 with short and long tails. W was best fitted with a linear polynomial of order 3 with a short tail function. Hg was best fitted with a linear polynomial of order 3 giving the best fit. For Ta the best fit was obtained with an exponential polynomial of order 4 with a short tail function. For U the best fit was obtained with an exponential polynomial of order 4.

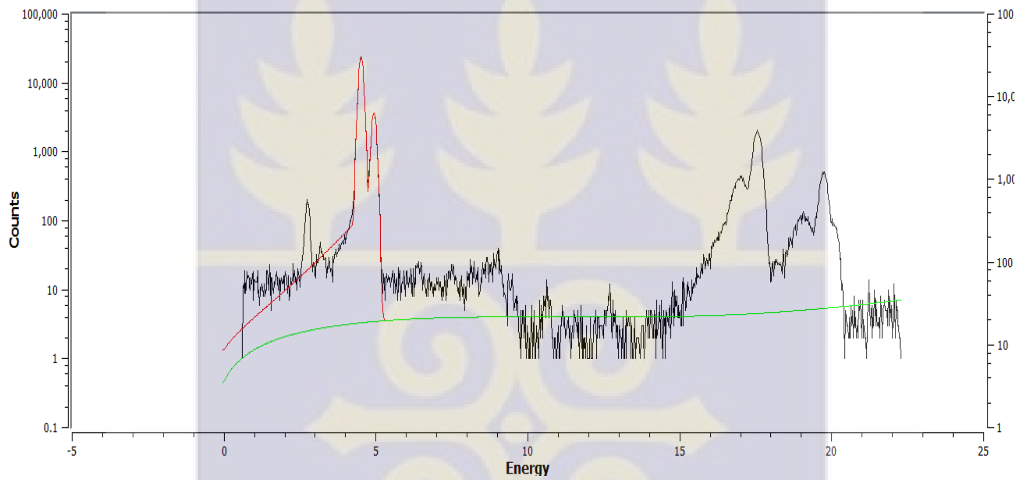


Fig. 3.2 Fitted Ti standard spectrum with PyMCA

3.3.1 Adaptation of PyMCA to X-ray tube excitation sources

The PyMCA software originally written for synchrotron radiation sources can be adapted to X-ray excitation sources by clicking on the 'Beam' tab and clicking 'Open X-ray tube setup'. Here, parameters such as the transmission tube voltage, anode material and window material are input; also needed are the thicknesses of these materials. The photon energy used for the irradiation (primary beam) should also be entered.

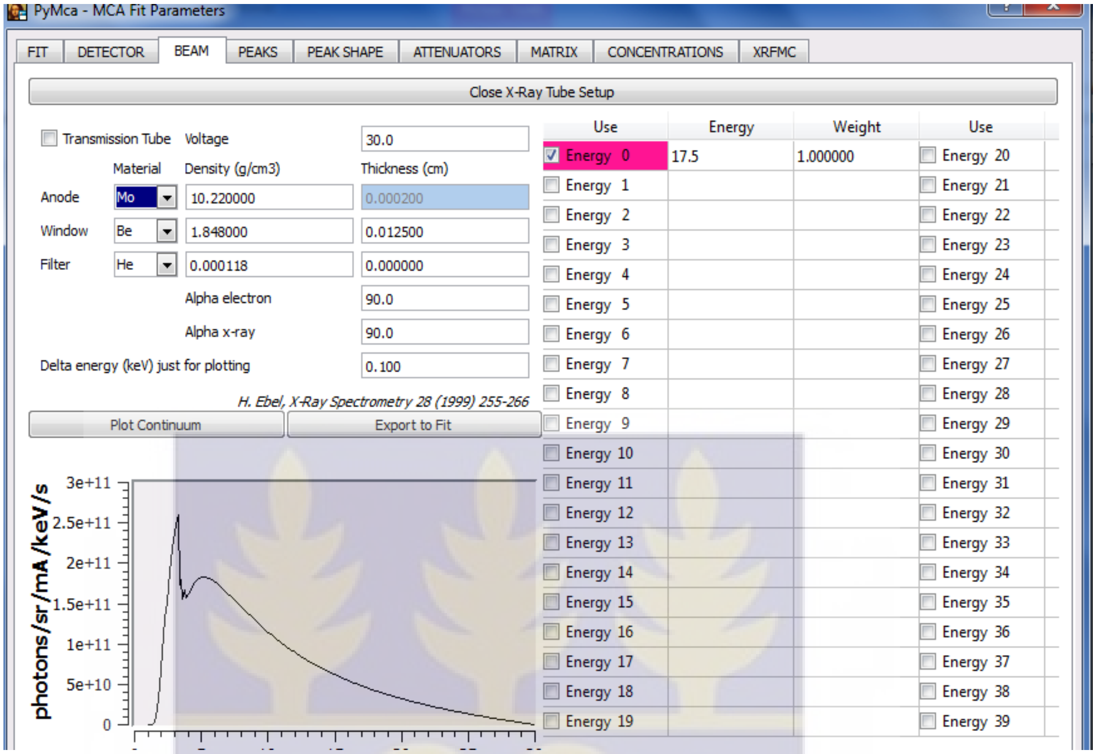


Fig. 3.3 Adaptation of PyMCA to X-ray tube excitation sources

3.4 Peak Intensities

Peak intensities for all the elements were found using the relation

$$I_p = \frac{N_p}{LT \times i} \tag{3.1}$$

where

I_p *is the peak intensity of element* p

N_p *is the peak area of element* p

LT *is the acquisition time* (600 s)

i *is the X* − *ray tube current* (5 mA)

3.5 Sensitivity Calibration

For PyMCA, sensitivity calibration was calculated using the relation

$$S_p = \frac{N_p \times a_p}{LT \times i \times c_p} \tag{3.2}$$

where

S_p *is the sensitivity of element p*

N_p *is the peak area of element p*

a_p *is the absorption factor of the pth element*

LT *is the acquisition time* (600 s)

i *is the X − ray tube current* (5 mA)

c_p *is the mass fraction of the pth element*

For WinAxil, sensitivity calibration was done using the software.

3.6 Quantitative Analysis

Three x-ray fluorescence spectra of soil samples taken from a petrol station (GDGA2-45, GDGA3-45, and GDGA45) were fitted with WinAxil and PyMCA. The soil sample spectra were all fitted with a linear polynomial of order 3 for both WinAxil and PyMCA. The peak areas obtained for the three soil spectra GDGA2-45, GDGA3-45, and GDGA4-45 are shown in tables A9, A10 and A11 respectively Appendix A. The intensities of the fitted elements were obtained using equation (3.2). The acquisition time of 600 s. The X-ray tube current and voltage of 10 mA and 45 kV respectively were input for both software. Detector area and a sample-detector distance were also input for both.

Quantitative analysis was performed for the three soil samples using both WinAxil and PyMCA.

A simplified manual for spectrum fitting and quantitative analysis for PyMCA has been written to serve as a guide for first-time and inexperienced users and is available in Appendix B. A flowchart to make the manual easier to follow is available in Appendix C.

CHAPTER 4

RESULTS AND DISCUSSION

4.1 Intensities

The peak intensities of Kα and Lα lines of the elements fitted with both WinAxil and PyMCA are shown in tables A7 and A8 respectively of Appendix A. The graphical plot of PyMCA intensities against WinAxil intensities for Kα lines is shown in Fig. 4.1 below. Logarithmic scales were used for both axes as the range of values was more than one order of magnitude.

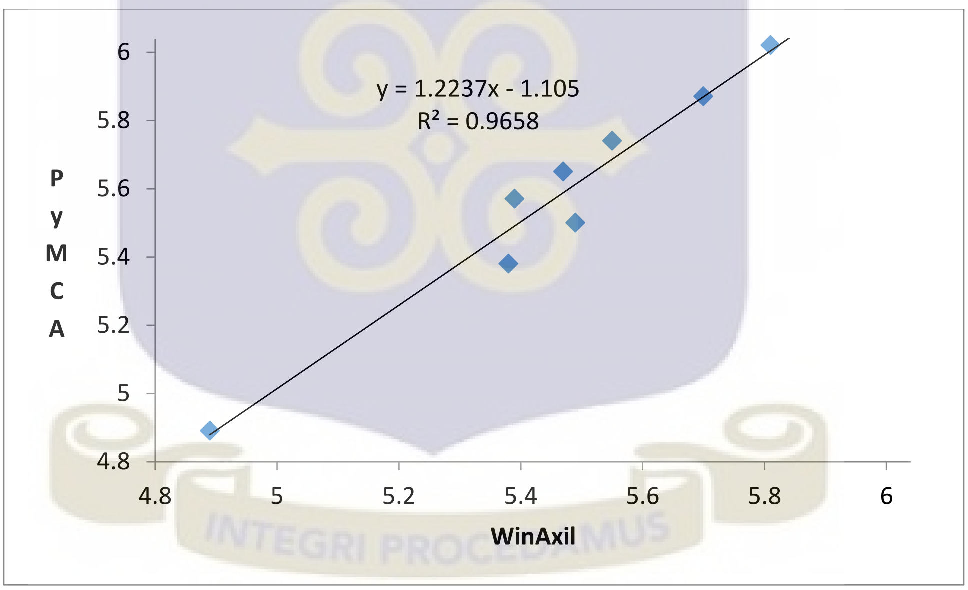


Fig. 4.1 Kα Plot of PyMCA Intensities against WinAxil Intensities

The graphical plot of PyMCA intensities against WinAxil intensities for Lα elements is shown in Fig. 4.2 below. Logarithmic scales were used for both axes as the range of values was more than one order of magnitude.

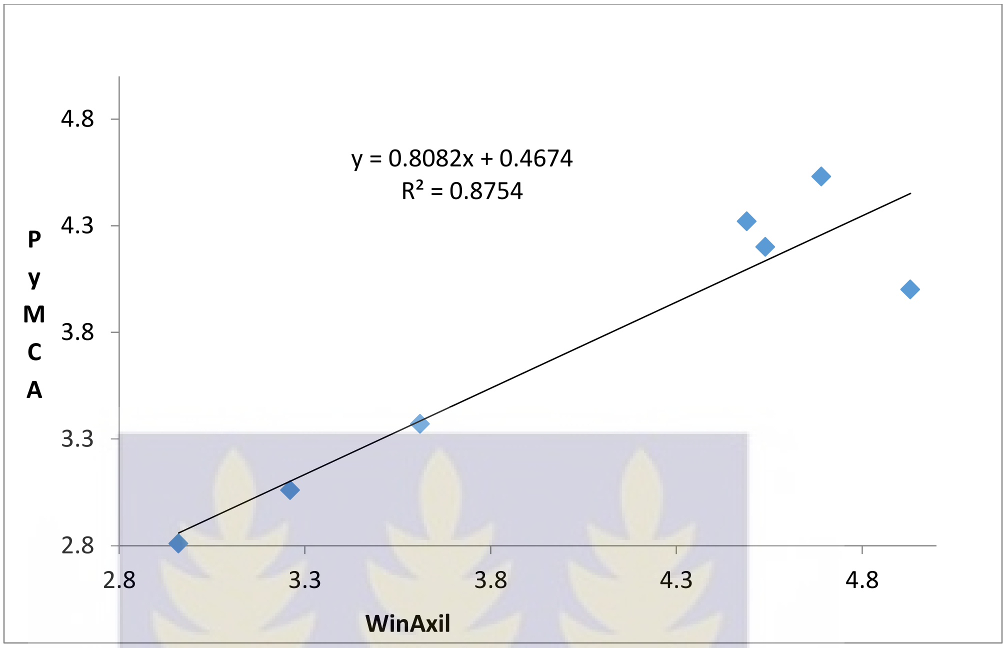


Fig. 4.2 Lα Plot of PyMCA Intensities against WinAxil Intensities

The intensities obtained by fitting Kα lines with both PyMCA and WinAxil show good agreement. This is borne out by the R^2 value of 0.9658. Although the agreement between PyMCA and WinAxil by fitting Lα lines is quite good, it is not as good as that for the Kα elements. Also, the intensities of the soil samples fitted with both PyMCA and WinAxil have R^2 values that indicate excellent agreement between the two software.

4.2 Sensitivity Curves

The sensitivity results for both Kα and Lα lines fitted with PyMCA are in tables A3 and A4 respectively in Appendix A. The Kα and Lα sensitivity values for WinAxil are in tables A5 and A6 respectively in Appendix A.

The sensitivity curves of Kα and Lα lines obtained from the WinAxil and PyMCA software in the previous chapter are shown respectively in Fig. 4.3 and Fig. 4.4 below.

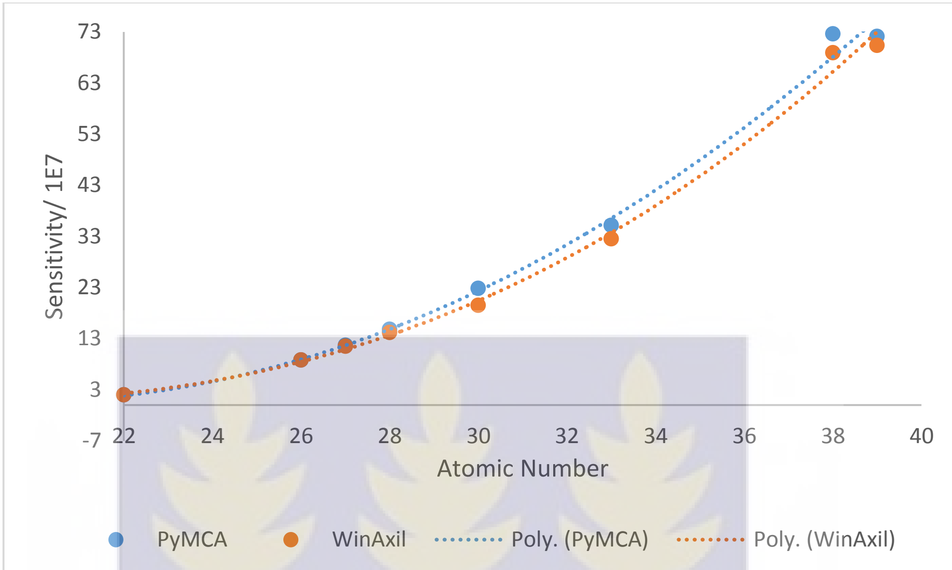


Fig. 4.3 Sensitivity curves of PyMCA & WinAxil for Kα lines

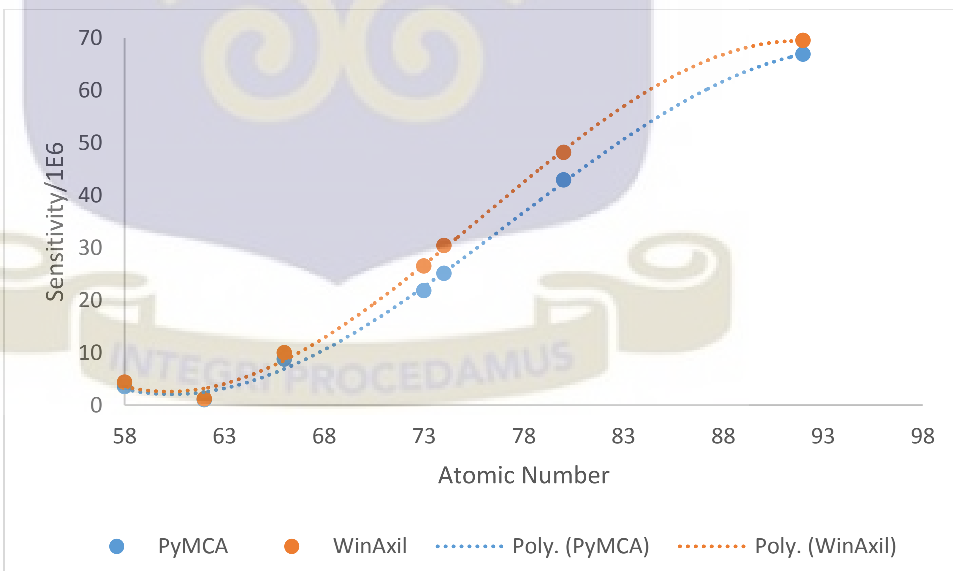


Fig. 4.4 Sensitivity curves of PyMCA & WinAxil for Lα lines

For Kα lines, the sensitivity curve of PyMCA lies above that of WinAxil whilst for Lα lines, it lies below that of WinAxil. For PyMCA, the best fit of the Lα elements was obtained with exponential polynomials whilst linear polynomials were enough to obtain a best fit using WinAxil. The difference between the Lα sensitivity results from the two software arose from the fact that linear polynomials were used to fit Lα lines with WinAxil whilst exponential polynomials were used to fit these same lines with PyMCA. Generally, when fitting L lines, high orders of the polynomial are used to obtain good results than the use of low orders which can be attributed to the complex nature of the L line spectra. The specification of high orders during the fitting process introduces many unknown parameters and this puts a very high burden on the non-linear least squares algorithm that needs to find the optimum values for all these parameters. This burden can cause erroneous values of the parameters to be determined which as a result give high chi-square values indicating that the fitting function could not best fit the experimental data. Thus this could be the reason why the sensitivity values of the L lines for PyMCA are less than those of WinAxil since they (L lines) were fitted with exponential polynomials.

4.3 Quantitative Analysis

4.3.1 Peak Intensities

The intensities obtained for soil sample spectra GDGA2-45, GDGA3-45 & GDGA4-45 are shown in tables A12, A13 and A14 respectively of Appendix A.

The graphical plot of PyMCA intensities against WinAxil intensities for the soil sample spectra GDGA2-45, GDGA3-45 & GDGA4-45 shown in Fig.4.5, Fig. 4.6 & Fig. 4.7

respectively. Logarithmic scales were used for both axes as the range of values was several orders of magnitude.

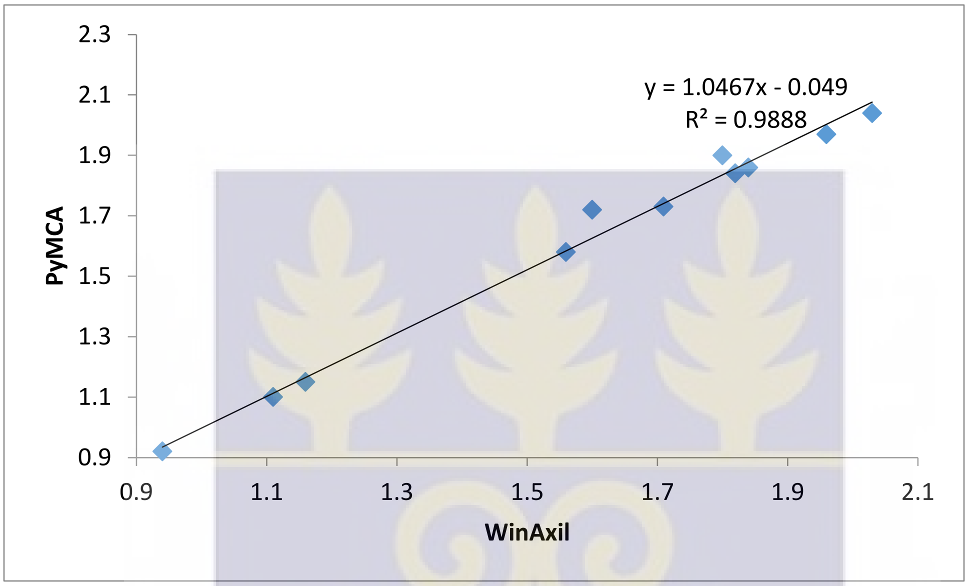


Fig.4.5 Plot of PyMCA Intensities against WinAxil Intensities of GDGA2-45

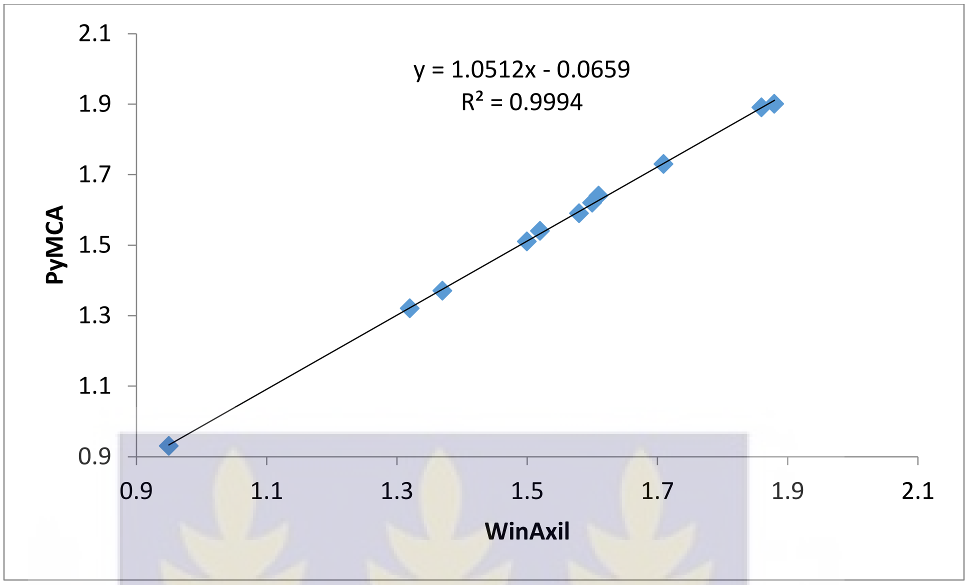


Fig.4.6 Plot of PyMCA Intensities against WinAxil Intensities of GDGA3-45

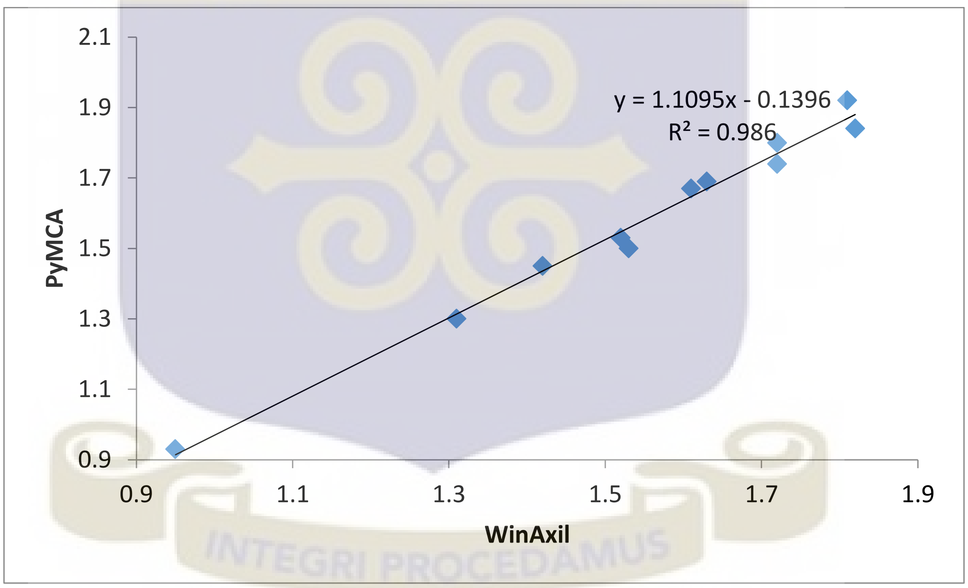


Fig.4.7 Plot of PyMCA Intensities against WinAxil Intensities of GDGA4-45

4.3.2 Elemental Concentrations

The tables of results for the elemental concentrations of the soil samples GDGA2-45, GDGA3-45 and GDGA4-45 are shown below in table A15, table A16 & table A17 respectively of Appendix A.

The graphical plot of elemental concentrations obtained using PyMCA against elemental concentrations obtained using WinAxil for the soil samples GDGA2-45, GDGA3-45 & GDGA4-45 are shown in Fig.4.8, Fig. 4.9 & Fig. 4.10 respectively. Logarithmic scales were used for both axes as the range of values was several orders of magnitude.

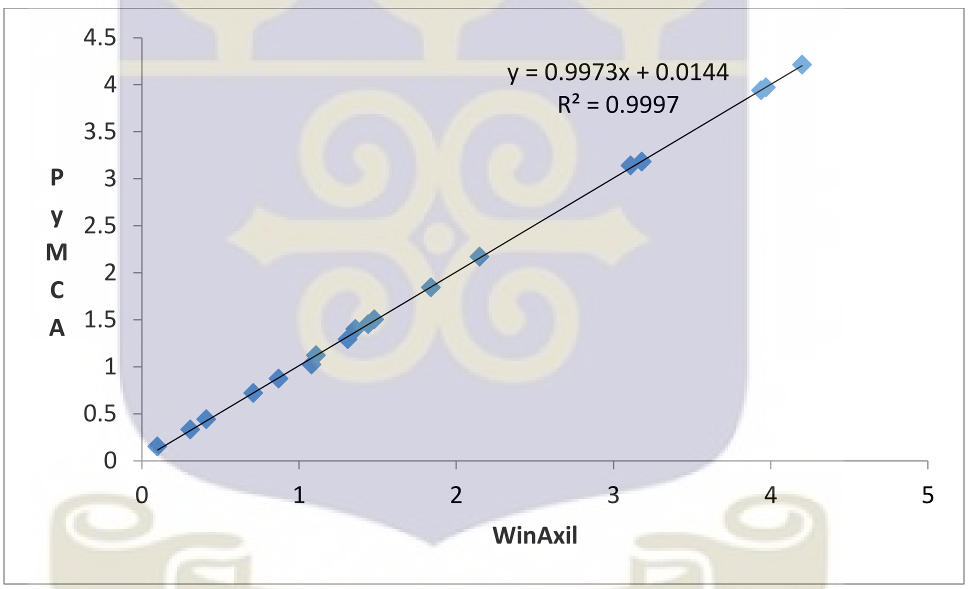


Fig.4.8 Plot of PyMCA Concentration against WinAxil Concentration of GDGA2-45

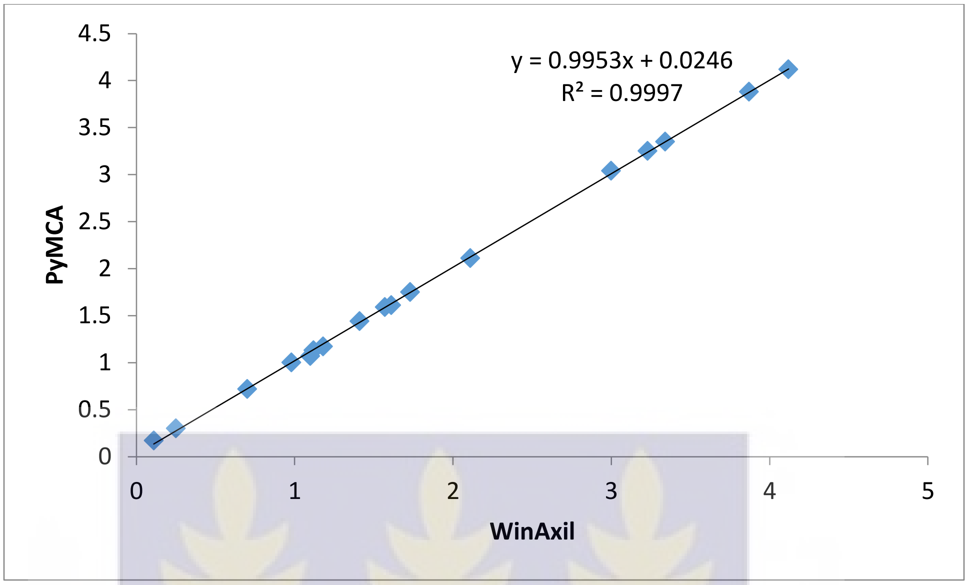


Fig. 4.9 Plot of PyMCA Concentration against WinAxil Concentration of GDGA3-45

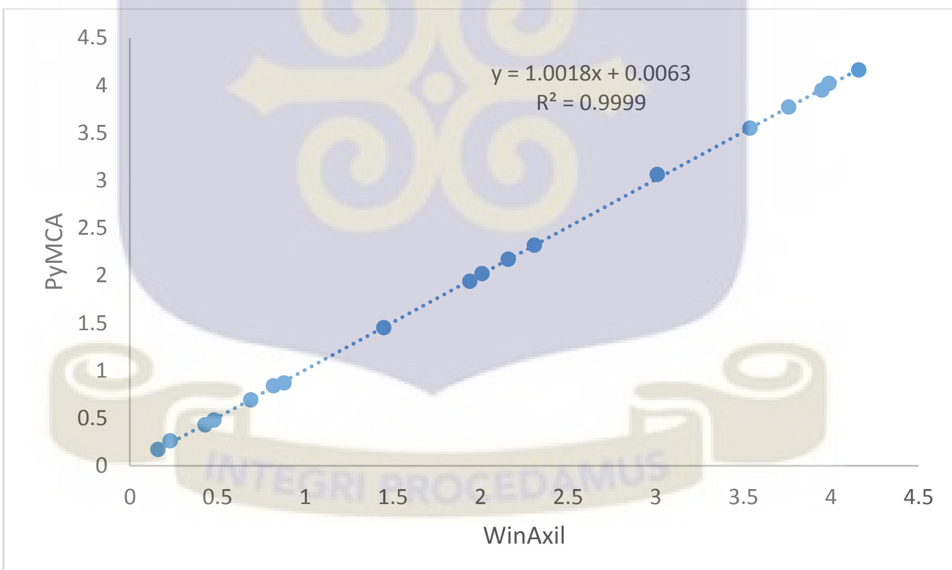


Fig. 4.10 Plot of PyMCA Concentration against WinAxil Concentration of GDGA4-45

The intensities of the elements in the soil sample spectra fitted with both PyMCA and WinAxil show a much better agreement between the two software than for those of the standard spectra. This is because for the standard spectra a much higher background exists since most of the standards are in the pure form, whilst the background in the soil sample spectra is not as high due to the fact that the elements are present in small amounts.

Since intensity is directly proportional to concentration, it is expected that the concentrations of the elements in soil sample spectra fitted with both WinAxil and PyMCA will be in good agreement and this is borne out in fact by the R^2 values of the graphs of the elemental concentrations found in Figs.4.8-4.10.

CHAPTER 5

CONCLUSION & RECOMMENDATIONS

5.0 Conclusion

According to the results obtained, the x-ray peak fitting capabilities using PyMCA compare favourably with those of WinAxil. Since the development of WinAxil has been discontinued, PyMCA can be used in place of WinAxil for qualitative and quantitative analysis with x-ray tube excitation sources. The best fitting results for the linear and exponential background models were obtained by choosing the lower degrees of the polynomials ranging from 1 to 4. Therefore, the conditions for adapting PyMCA for use in the deconvolution of spectra obtained from X-ray tubes involving fitting the background of such spectra with lower degrees of polynomials.

Current documentation on PyMCA where available poses difficulties for inexperienced users; therefore, a simplified version of the manual for spectrum fitting and quantitative analysis with PyMCA has been produced. This will facilitate the use of PyMCA for first-time and inexperienced users; in addition, a flowchart for the manual has also been done.

5.1 Recommendations

The adaptation of PyMCA for the analysis of Proton-Induced X-ray Emission (PIXE) should be studied. Since the Ghana Atomic Energy Commission has acquired a pelletron accelerator which can be used for PIXE analysis, a comparative study of the capabilities of PyMCA and WinAxil should be extended to pelletron accelerators.

The QXAS program, which was written for use in third-world countries under the sponsorship of the IAEA and made freely available because of how expensive commercially-available spectrum deconvolution software are, is no longer being updated; it is recommended that X-ray spectrometry laboratories in the third-world should consider using the PyMCA software since it has proven to give good results in EDXRF analysis.

Care should be taken in the choice of high orders of polynomials since they can cause errors in the results of spectrum fitting.

APPENDIX A

Elements	Peak Areas	
	PyMCA	WinAxil
Ti	195 601.50	192 682.00
Fe	598 993.75	597 734.00
Co	773 057.25	773 057.25
Ni	921 314.75	883 964.25
Zn	1 317 310.00	1 126 646.75
As	1 837 353.25	1 701 253.25
Sr	3 116 458.25	2 587 630.50
Y	2 619 170 .25	1 208 711.25

Table A1 Peak Areas of Kα elements

Elements	Peak Areas	
	PyMCA	WinAxil
Ce	1827.50	2 297.50
Sm	3780.50	4 547.50
Dy	9012.50	10 241..50
Ta	64 015.25	77 816.75
W	71 221.00	86 259.75
Hg	109 347.00	122 856.25
U	206 230.75	214 245.75

Table A2 Peak Areas of Lα elements

Element	Sensitivity(cps $A^{-1}\mu g^{-1}$)
Ti	2.01×10^7
Fe	8.81×10^7
Co	11.5×10^7
Ni	14.8×10^7
Zn	22.8×10^7
As	35.1×10^7
Sr	72.6×10^7
Y	72.1×10^7

Table A3 PyMCA Kα Elemental Sensitivities

Element	Sensitivity (cps $A^{-1}\mu g^{-1}$)
Ce	3.5×10^6
Sm	1.0×10^6
Dy	8.8×10^6
Ta	21.8×10^6
W	25.1×10^6
Hg	42.9×10^6
U	66.9×10^6

Table A4 PyMCA Lα Elemental Sensitivities

Element	Sensitivity (cps $A^{-1}\mu g^{-1}$)
Ti	1.98×10^7
Fe	8.79×10^7
Co	11.5×10^7
Ni	14.2×10^7
Zn	19.5×10^7
As	32.5×10^7
Sr	68.99×10^7
Y	70.33×10^7

Table A5 WinAxil Kα Elemental Sensitivities

Element	Sensitivity (cps $A^{-1}\mu g^{-1}$)
Ce	4.4×10^6
Sm	1.2×10^6
Dy	10×10^6
Ta	26.5×10^6
W	30.4×10^6
Hg	48.2×10^6
U	69.5×10^6

Table A6 WinAxil Lα Elemental Sensitivities

Element	WinAxil		PyMCA	
	Intensity (cps)	σ	Intensity (cps)	σ
Ti	77239.2	276.4	78146.0	278.8
Fe	238918.2	442.8	239610.0	489.4
Co	310400.0	559.5	314457.1	560.4
Ni	247120.0	497.0	375704.7	598.2
Zn	355594.0	592.9	548826.4	740.4
As	500188.0	706.2	735548.0	856.3
Sr	748449.0	865.3	1180114.4	1085.6
Y	641950.0	801.2	1047461.2	1023.2

Table A7 Peak intensities of Kα lines

Element	WinAxil		PyMCA	
	Intensity (cps)	σ	Intensity (cps)	σ
Ce	917.2	31.9	638.7	25.1
Sm	1803.2	42.1	1156.4	33.5
Dy	4092.4	101.1	2342.6	48.3
Ta	31086.0	176.3	20733.9	144.1
W	34536.0	184.6	15912.4	126.5
Hg	49114.8	213.4	34125.2	184.5
U	63731.6	253.2	52386.7	229.3

Table A8 Peak intensities of Lα lines

Element	Peak Areas	
	WinAxil	PyMCA
K	539.8	549.3
Ca	6399.3	6437.4
Ti	2789.3	2797.7
V	83.2	84.5
Cr	159.9	161.0
Mn	1309.8	1335.9
Fe	229879.9	231599.8
Ni	236.3	242.6
Cu	582.5	584.4
Zn	4289.9	4358.9
Ga	172.2	176.9
Br	299.1	319.1
Rb	1698.8	1739.9
Sr	11983.4	12470.0
Y	1195.3	1277.1
Zr	113.9	109.0
Pb	253.7	243.2

Table A9 Peak Areas of soil sample spectrum GDGA2-45

Elements	Peak Areas	
	WinAxil	PyMCA
K	629.8	638.2
Ca	5123.8	5253.1
Ti	2118.9	2217.9
V	128.7	130.8
Cr	229.9	235.2
Mn	1096.7	1164.1
Fe	184977.9	187422.9
Ni	319.2	328.0
Cu	594.8	604.0
Zn	3427.1	3519.1
Ga	161.7	165.8
Br	208.9	210.2
Rb	1679.9	1739.2
Sr	9978.2	10756.1
Y	678.9	684.2
Zr	154.8	155.1
Pb	146.2	145.2

Table A10 Peak Areas of soil sample spectrum GDGA3-45

Elements	Peak Areas	
	WinAxil	PyMCA
K	1253.6	1271.8
Ca	7011.3	7174.8
Ti	2299.6	2358.3
V	276.9	295.6
Cr	1087.3	1162.7
Mn	1289.7	1324.9
Fe	199765.2	208420.1
Ni	221.7	227.1
Cu	9086.4	9233.8
Zn	6499.8	6586.7
Ga	121.2	149.9
Br	301.8	336.4
Rb	1597.9	1633.7
Sr	10456.1	11119.8
Y	1375.9	1419.1
Zr	29.9	28.1
Pb	94.9	93.1

Table A11 Peak Areas of soil sample spectrum GDGA4-45

Element	Peak Intensities (cps)	
	WinAxil	PyMCA
K	90.0	91.5
Ca	1006.6	1072.5
Ti	464.9	466.3
V	13.9	14.1
Cr	26.7	26.8
Mn	218.3	222.7
Fe	38313.3	38600.0
Ni	39.4	40.4
Cu	97.1	97.4
Zn	715.0	726.5
Ga	28.7	29.4
Br	49.9	53.2
Rb	283.1	290.0
Sr	1997.2	2078.3
Y	199.2	212.9
Zr	19.0	18.2
Pb	42.3	40.5

Table A12 Peak Intensities of soil sample spectrum GDGA2-45

Elements	Peak Intensities (cps)	
	WinAxil	PyMCA
K	105.0	106.4
Ca	855.5	875.5
Ti	353.2	369.7
V	21.5	21.8
Cr	38.3	39.2
Mn	182.8	194.0
Fe	30829.7	31237.2
Ni	53.2	54.7
Cu	99.1	100.7
Zn	571.2	586.5
Ga	27.0	27.6
Br	34.8	35.0
Rb	279.9	289.9
Sr	1663.0	1792.7
Y	113.2	114.0
Zr	25.8	25.9
Pb	24.4	24.2

Table A13 Peak Intensities of soil sample spectrum GDGA3-45

Elements	Peak Intensities (cps)	
	WinAxil	PyMCA
K	208.9	212.0
Ca	1168.6	1195.8
Ti	383.3	393.1
V	46.2	49.2
Cr	181.2	193.8
Mn	215.0	220.8
Fe	33294.2	34736.7
Ni	37.0	37.9
Cu	1514.4	1539.0
Zn	1083.3	1097.8
Ga	20.2	25.0
Br	50.3	56.1
Rb	266.3	272.3
Sr	1742.7	1853.3
Y	229.3	236.5
Zr	5.0	4.7
Pb	15.8	15.5

Table A14 Peak Intensities of soil sample spectrum GDGA4-45

Element	Concentration (mg/kg)	
	WinAxil	PyMCA
K	1497	1528
Ca	9279	9305
Ti	1289	1378
V	22.65	24.98
Cr	27.67	28.16
Mn	141.0	149.0
Fe	15987	16240
Ni	7.345	7.460
Cu	12.76	13.07
Zn	68.54	69.25
Ga	2.065	2.120
Br	1.260	1.400
Rb	5.130	5.220
Sr	29.97	31.57
Y	2.562	2.750
Zr	11.89	10.43
Pb	20.56	19.67

Table A15 Elemental concentration of soil sample spectrum GDGA2-45

Element	Concentration(mg/kg)	
	WinAxil	PyMCA
K	1698	1776
Ca	7368	7593
Ti	997.7	1092
V	37.36	38.67
Cr	40.76	41.14
Mn	127.8	129.9
Fe	13034	13140
Ni	9.650	10.09
Cu	13.19	13.51
Zn	53.67	55.90
Ga	1.786	1.989
Br	0.8976	0.9236
Rb	4.987	5.222
Sr	25.98	27.23
Y	1.275	1.472
Zr	14.98	14.83
Pb	12.56	11.75

Table A16 Elemental concentration of soil sample spectra GDGA3-45

Element	Concentration(mg/kg)	
	WinAxil	PyMCA
K	3465	3539
Ca	9865	10370
Ti	8853	8990
V	1032	1161
Cr	86.78	87.39
Mn	145.87	147.8
Fe	14378	14620
Ni	6.543	6.986
Cu	205.8	206.6
Zn	102.9	104.6
Ga	1.698	1.799
Br	1.431	1.478
Rb	4.873	4.905
Sr	27.95	28.16
Y	2.986	3.053
Zr	2.701	2.685
Pb	7.567	7.530

Table A17 Elemental concentration of soil sample spectrum GDGA4-45

APPENDIX B

A SIMPLIFIED PyMCA MANUAL

1. To select a file for fitting, click on the 'SpecFile' tab.

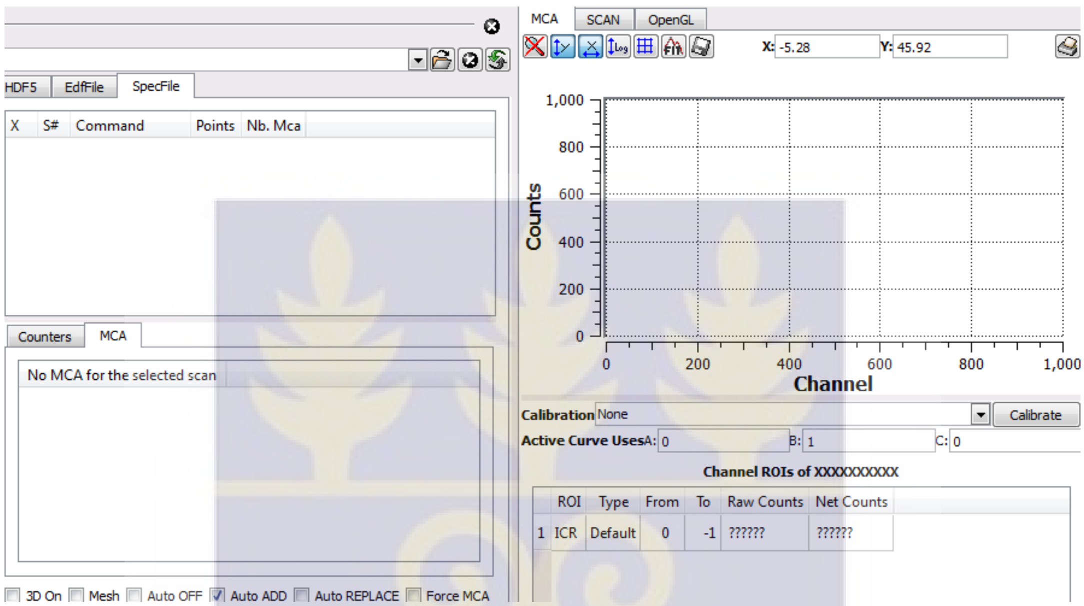


Fig.B1 PyMCA interface

2. Click on 'File' in the menu bar and then click 'Open' and 'Data Source' in the sub menu. Alternatively, file can be loaded by clicking on the open file icon on the top left side below the file menu bar.

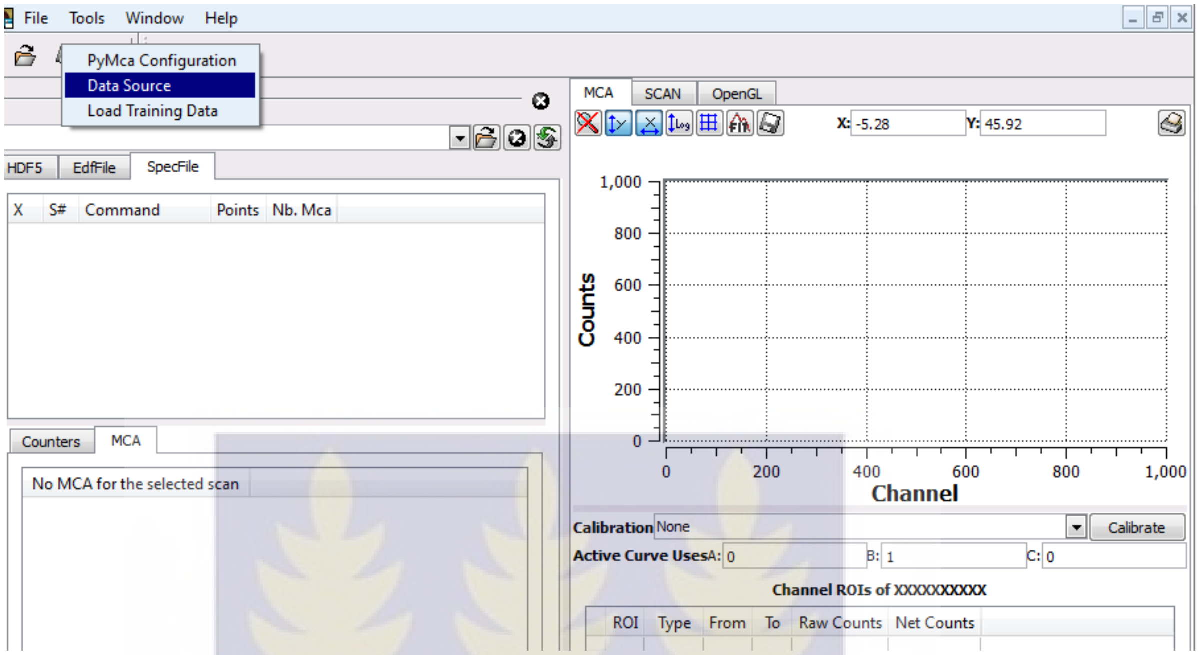


Fig. B2 Opening a spectrum file

A new window opens as below. Click on the downward-pointing arrow at the bottom right and select ‘All files (*)’.

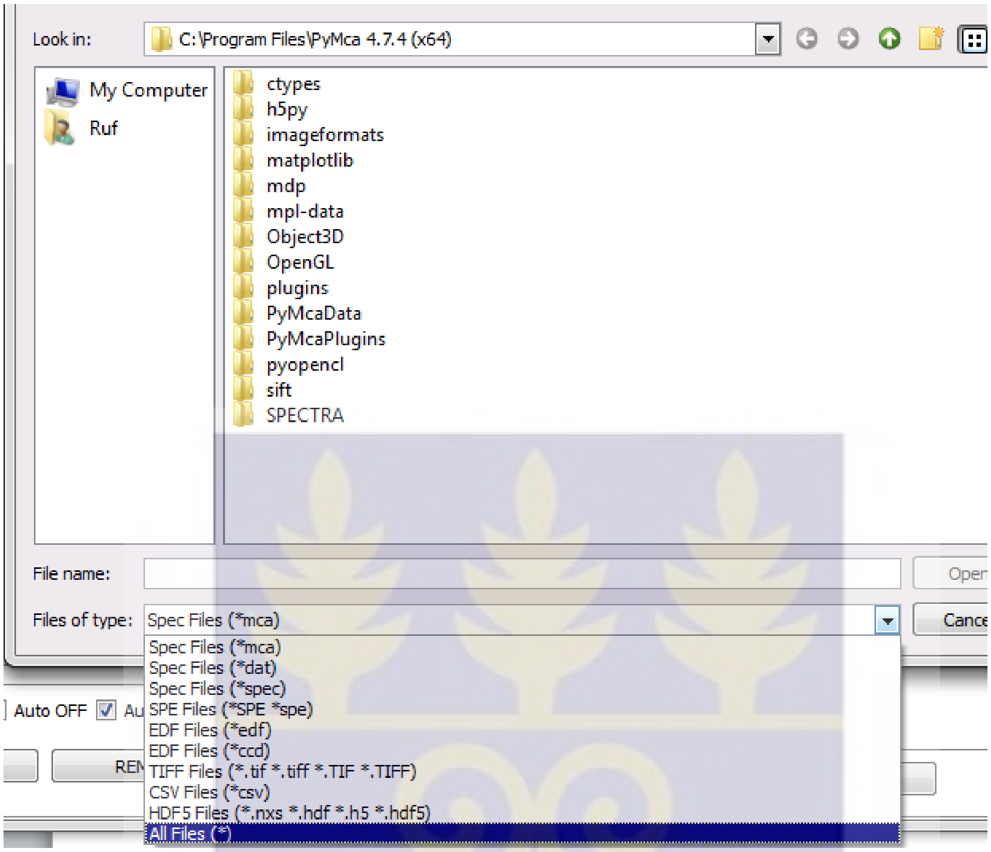


Fig. B3 Opening a spectrum file

3. Select the spectrum file of interest by double-clicking on it.

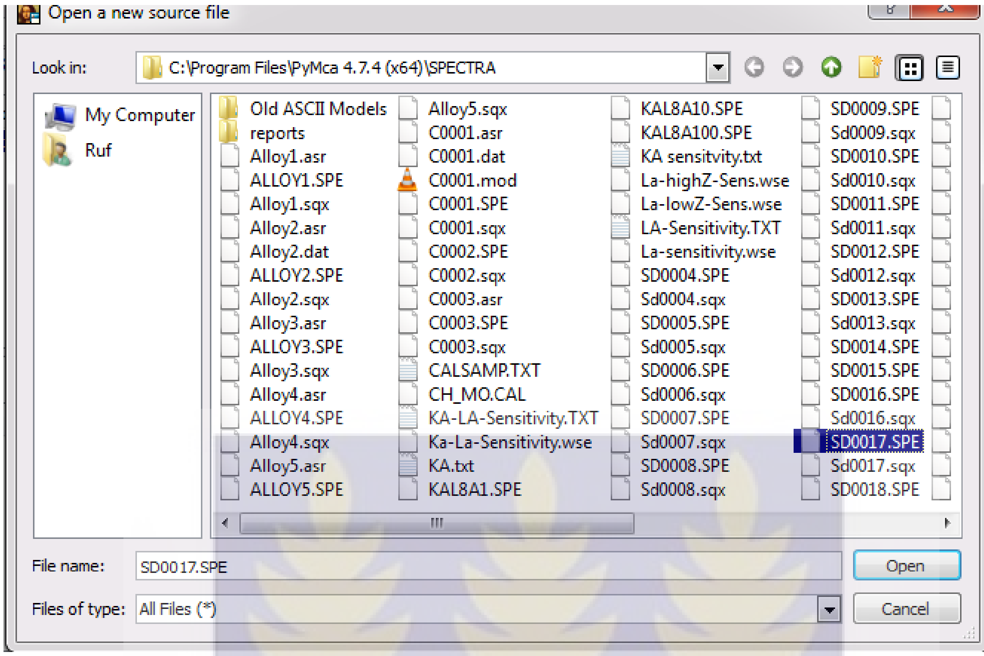


Fig.B4 File selection

If the required spectrum file is on a different drive or a location other than the hard drive, click on that drive, and select the file in that location by double-clicking on it.

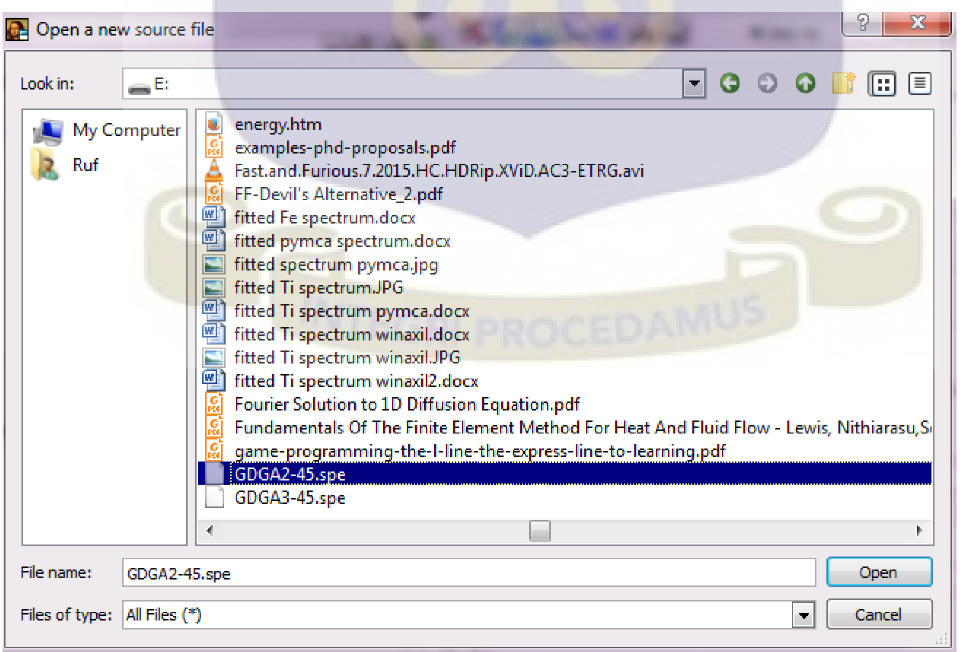


Fig B5. File selection from a different drive.

The user can toggle between a linear and a logarithmic scale by clicking on the ‘log’ tab as shown in Fig. B6a & Fig. B6b respectively.

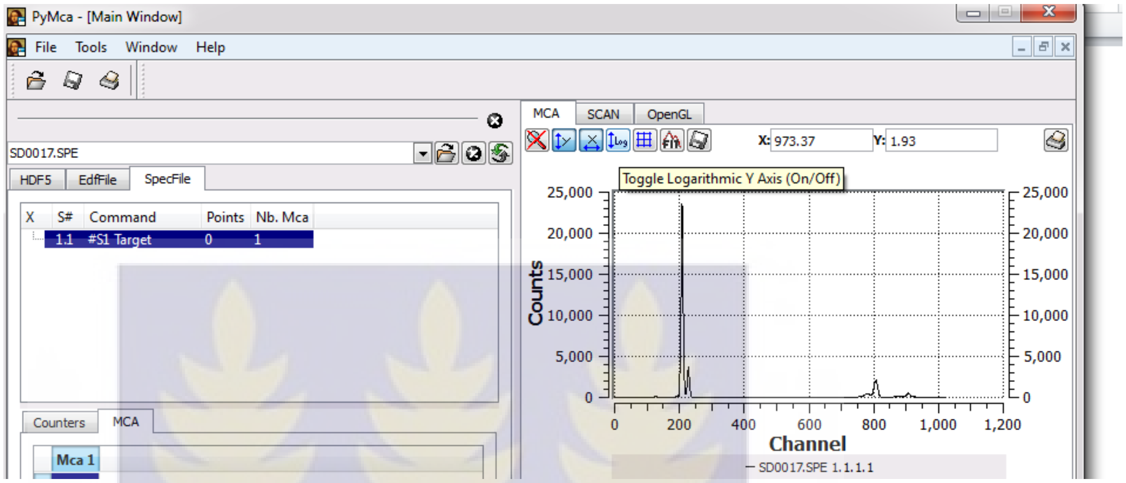


Fig.B6a Linear scale

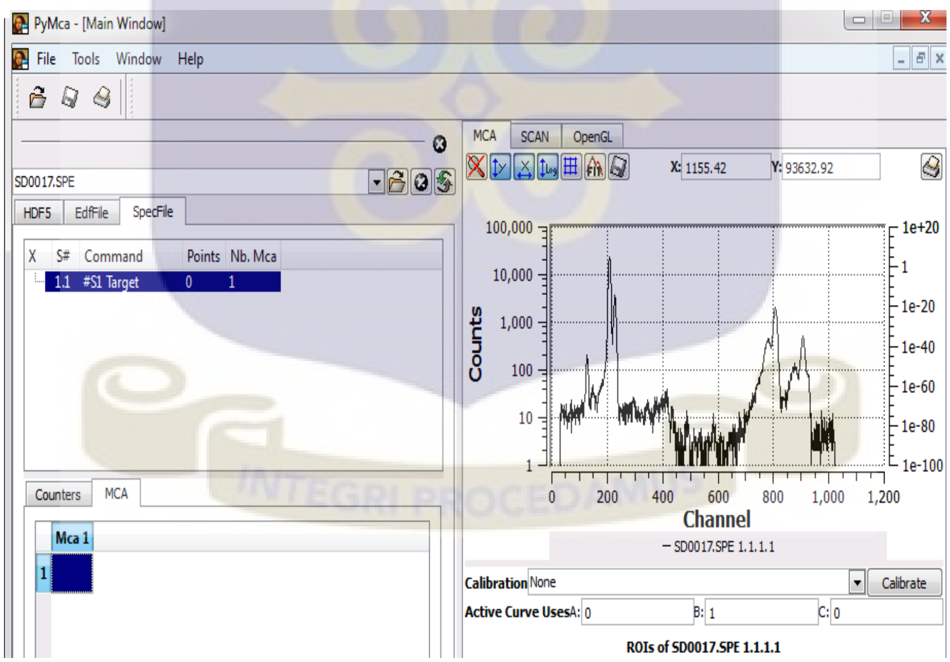


Fig.B6b Logarithmic scale

5. To carry out energy calibration, one creates a new calibration or copies an already existing calibration. To create a new calibration from the existing data, click calibrate at the bottom right, and select compute from the submenu.

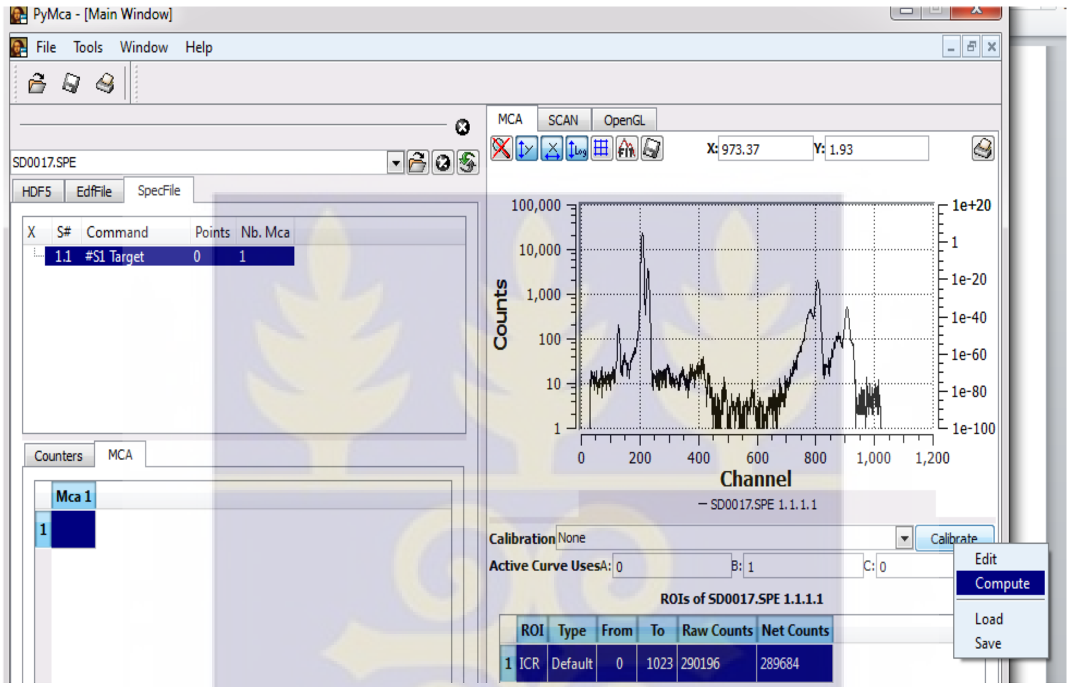


Fig. B7 Beginning energy calibration

To provide inputs for the energy calibration, the program offers you the possibility to search automatically for peaks or to enter a peak with the mouse by selecting the manual search icon. You will be shown a new window with the selected spectrum.

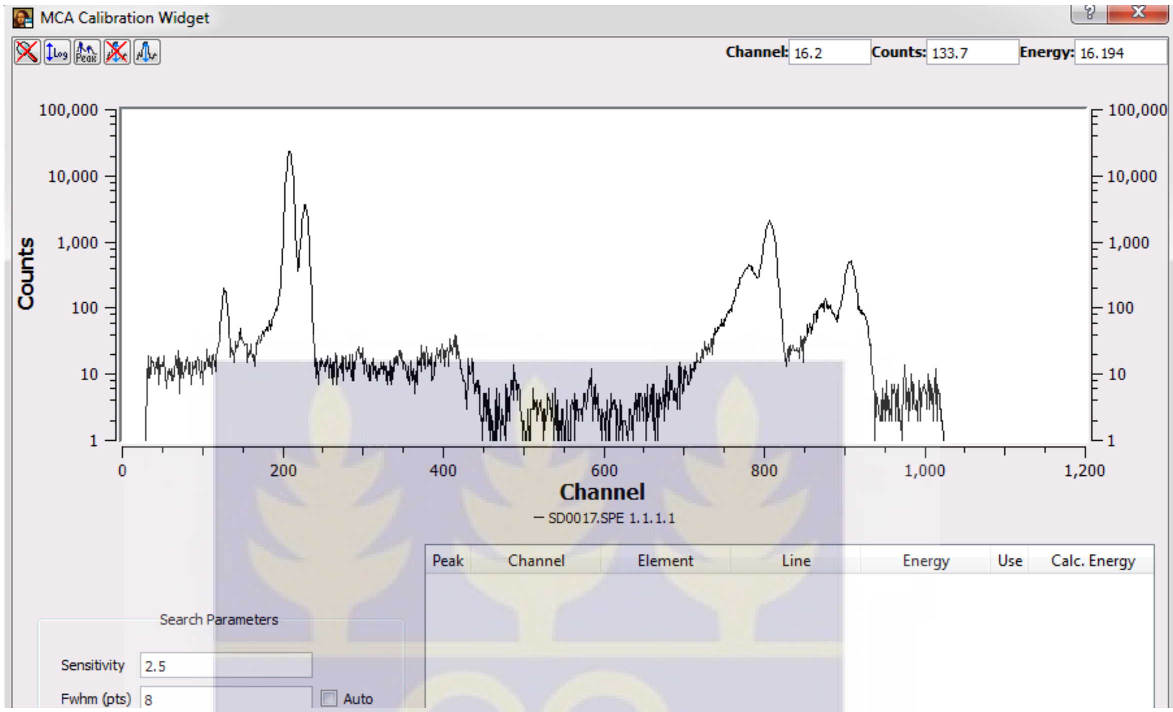


Fig. B8 MCA Calibration Widget interface

6. For automatic selection of the peaks, check the 'auto' boxes at the bottom left, adjacent 'Fwhm (pts)' and 'Yscaling' and click the 'Search' button. The lines that can be used for energy calibration are shown automatically as in Fig B9. Found or entered peaks will be denoted by vertical markers on the graph.

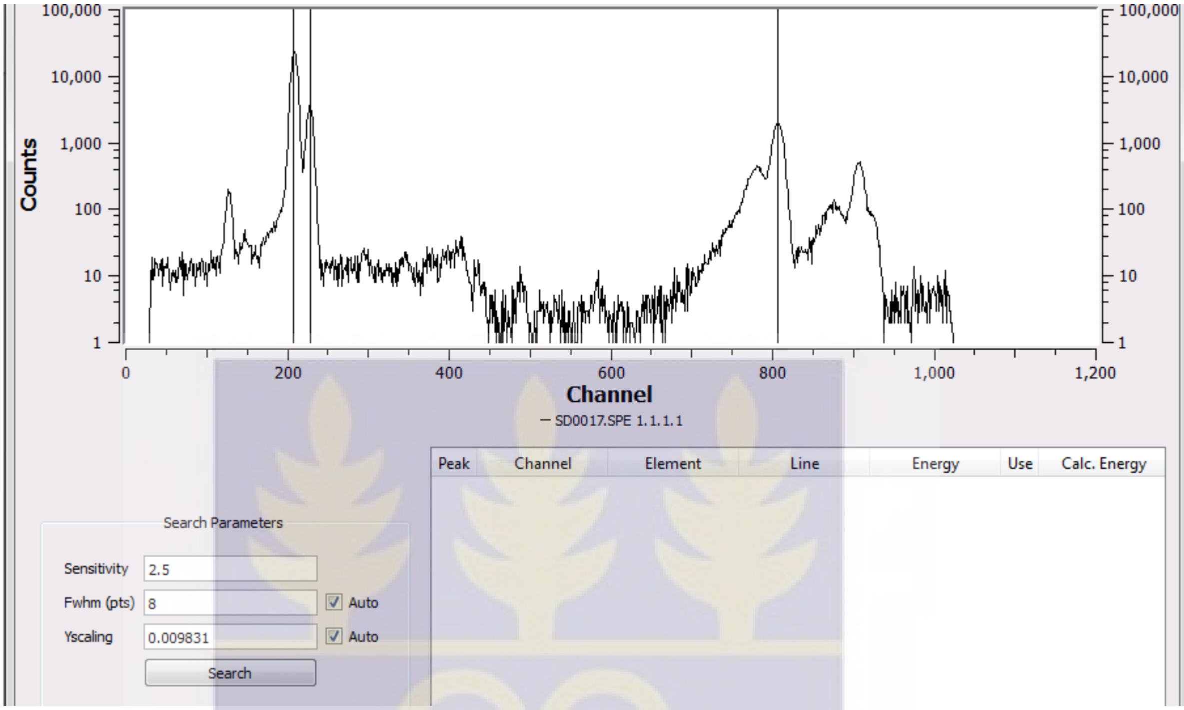


Fig. B9 Peaks to be used for energy calibration

7. To choose the peaks to be used for energy calibration, click on the downwards-pointing arrow opposite 'Order' and select '1st' from the resulting pull-down menu.

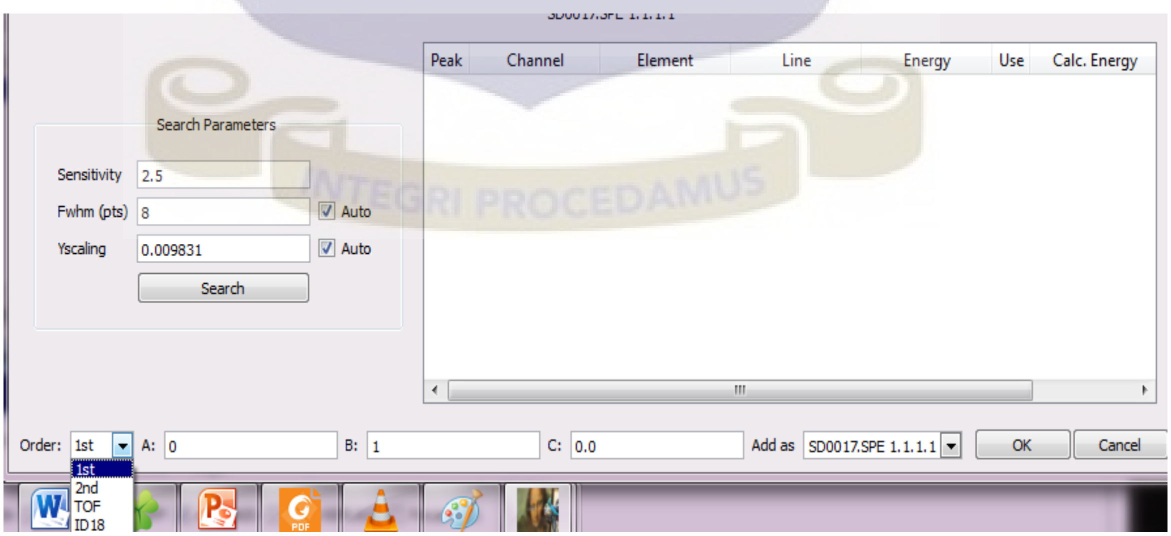


Fig. B10 Selection of first peak for energy calibration

Click on one of the selected peaks (this is the first peak being used in the calibration). A pop-up menu appears as shown:

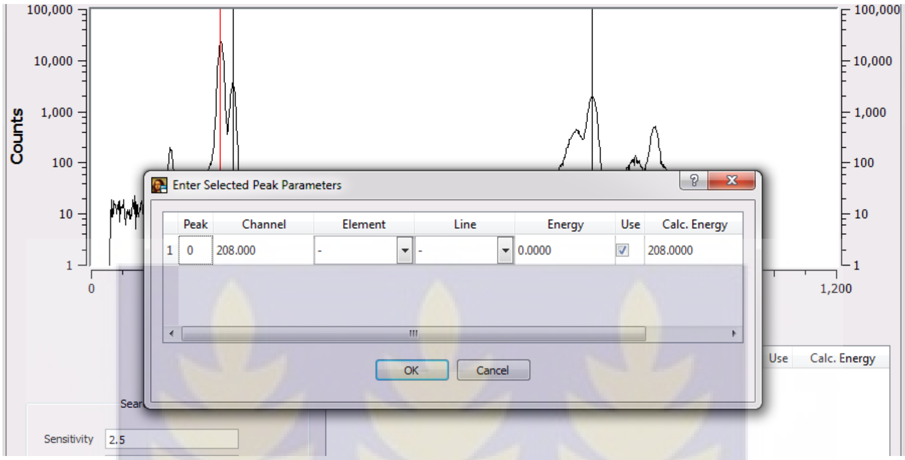


Fig.B11 Pop-up menu for selection of elements & energies

8. Enter the selected peak parameters 'Energy' and 'Line' and click 'OK'. To enter the associated energy to that peak just click on the marker. For Kα elements, the 'KL3' line refers to the Kα line of that particular element.

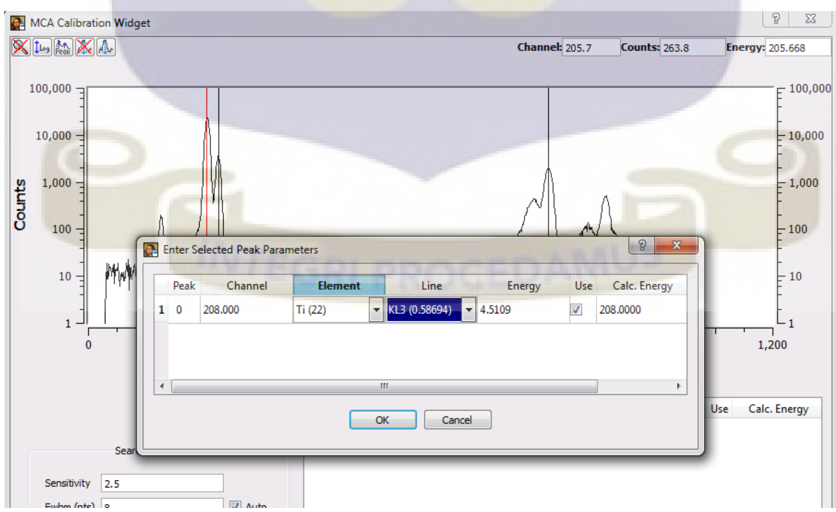


Fig. B12 Selection of first line for energy calibration in pop-up menu

9. Select ‘2nd’ from the ‘Order’ pull-down menu and click on another peak in order to input its selected parameters; then click ‘OK’. This selects the second peak for calibration.

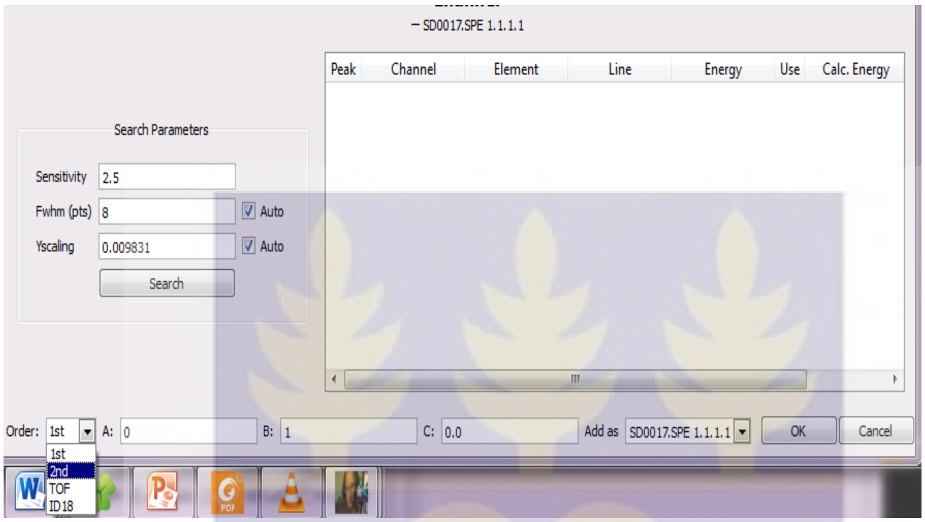


Fig. B13 Selection of second peak for energy calibration

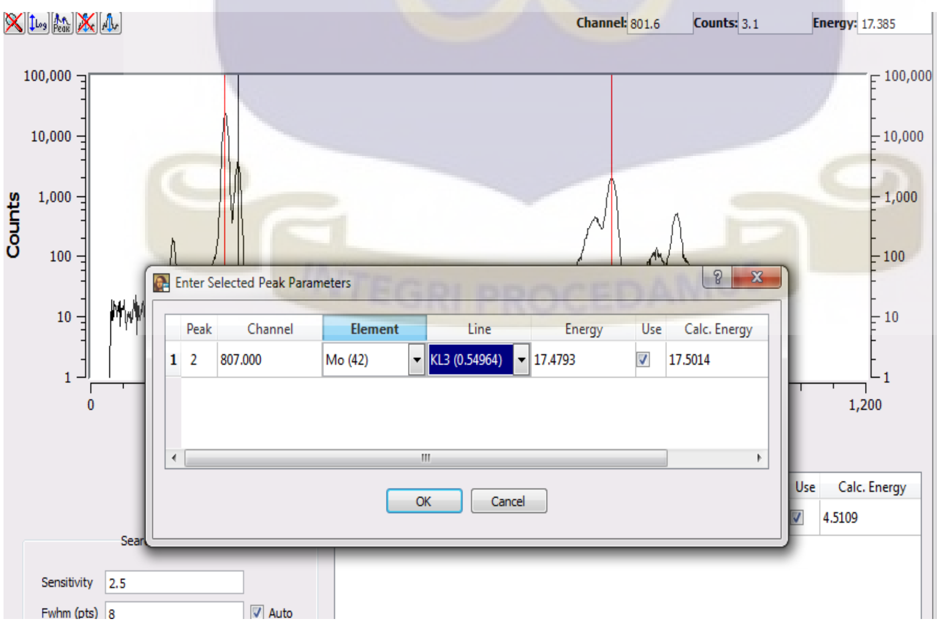


Fig. B14 Entering parameters of second peak

For K elements, the 'KM3' line refers to the Kβ line of that particular element.

For L elements, the L3M5* line represents the Lα line of that particular element. This line together with the Kα line of the coherently scattered line of the x-ray excitation source can be used for energy calibration.

The calibration widget appears as shown:

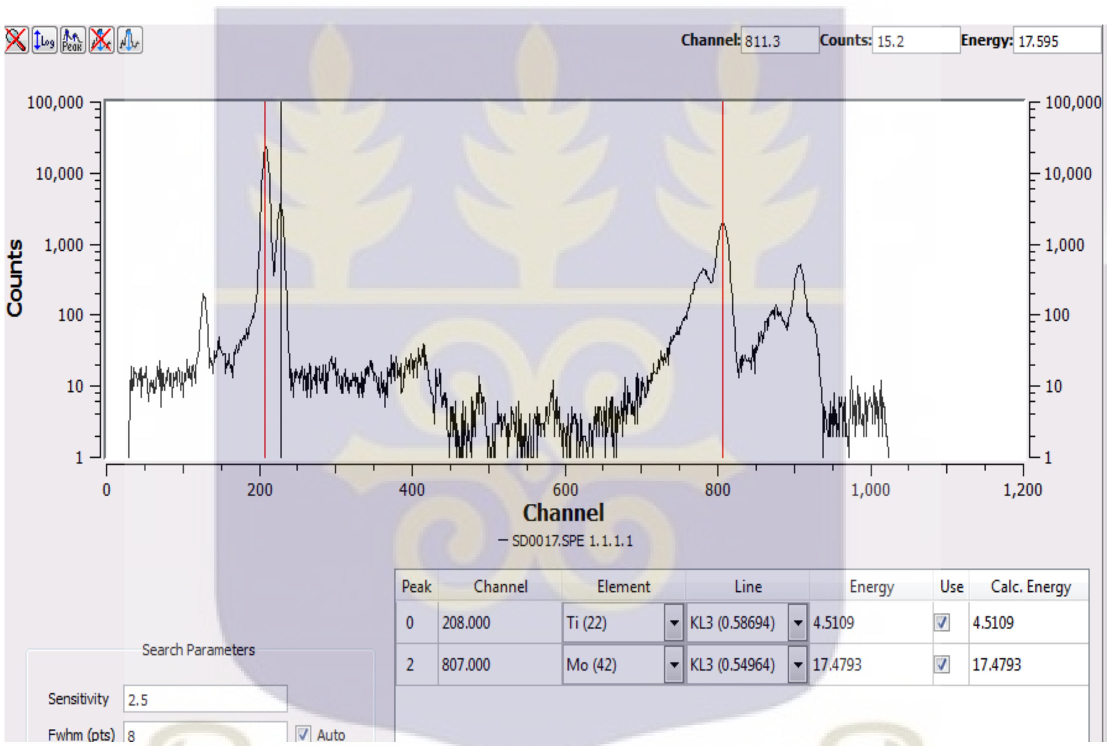


Fig. B15 Peak selection for energy calibration completed

10. To complete inputs for the calibration, click on the 'OK' tab on the bottom right-hand corner of the calibration widget. This exits the energy calibration and registers the spectrum filename in the calibration sub menu as in figure B16.

11. To select this file click on the downwards-pointing arrow opposite 'Calibration' as shown clicking on the filename.

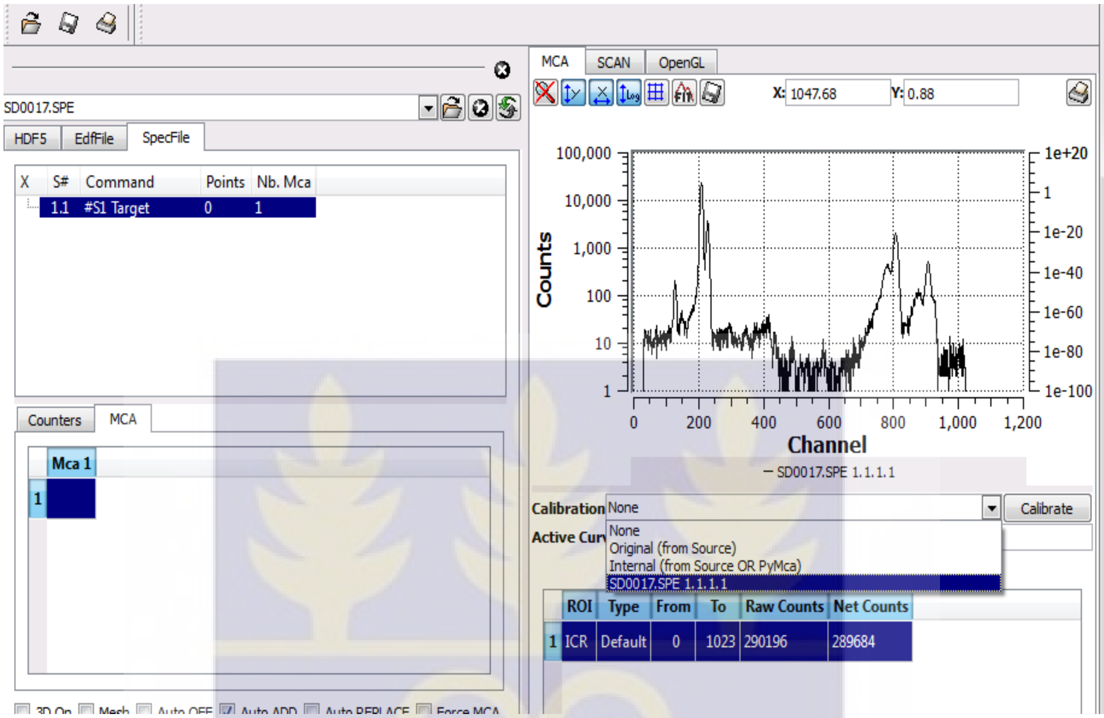


Fig. B16 Activation of energy calibration

The interface should look like below with ‘Channel’ replaced by ‘Energy’:

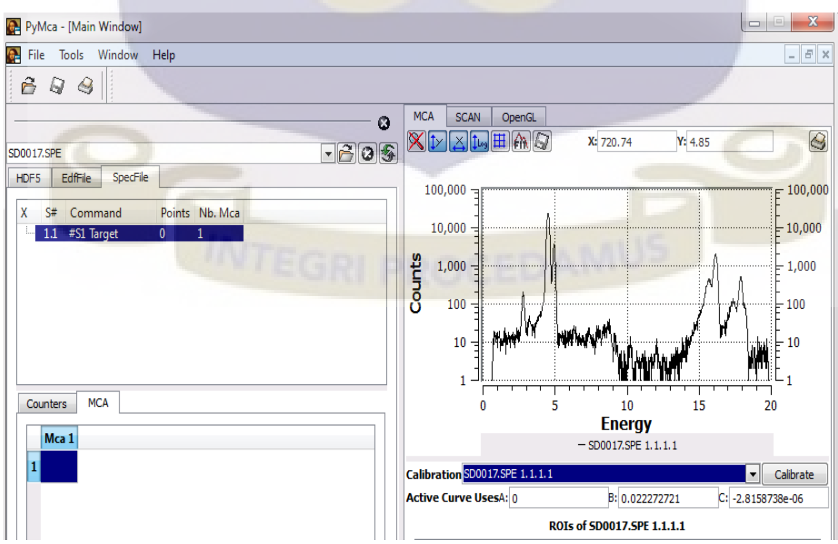


Fig.B17 Channel numbers replaced with energy

Click on 'Calibrate' and on 'Save'. Save the calibration to a location other than the PyMCA root drive.

If the user has internet access with Adobe Flash Player installed, they can view the animated procedure for energy calibration at pymca.sourceforge.net/documentation.

To load an existing calibration, click on 'Calibrate'and then 'Load'. The screen looks as below:

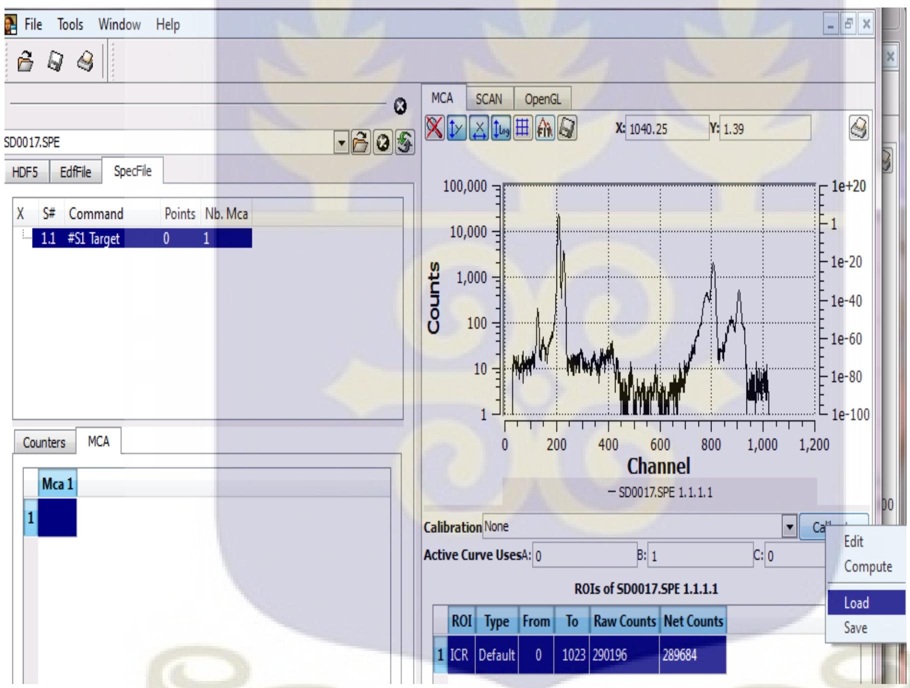


Fig. B18 Loading an existing calibration

Select the appropriate calibration file from the folder it has been saved in and click 'open'. Click on the downward-pointing arrow opposite 'Calibration'; the filename should be present in the list. Clicking it would activate that particular calibration.

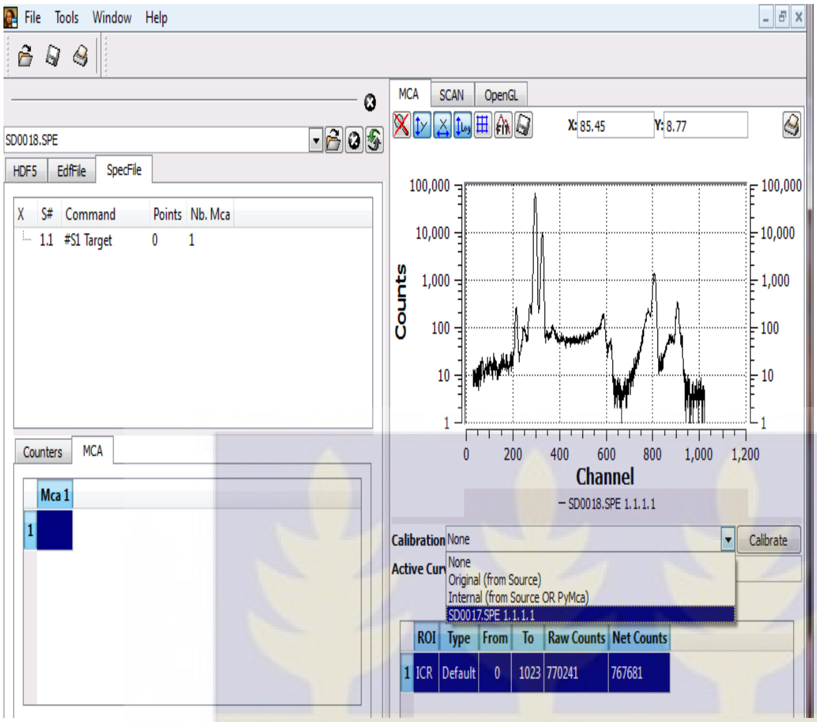


Fig. B19 Activating an existing calibration

12. To perform a fit, load the spectrum file as indicated in steps 1-3. To use an already saved calibration file, click on ‘Calibrate’ and on ‘Load’. A dialogue box opens; go to the location of the calibration file you saved and click on ‘open’. Then, click on the downward pointing arrow opposite ‘Calibration’; the calibration file should be in the options listed. Click on it; that calibration is now active.

Click on the ‘Fit’ tab and select ‘Advanced’ as shown in Fig. B20.

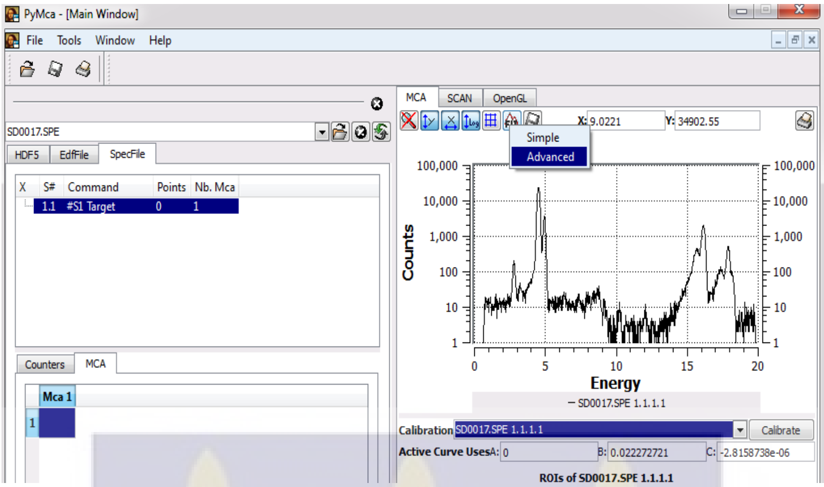


Fig. B20 Selection of Advanced fit option

The screen now appears as shown:

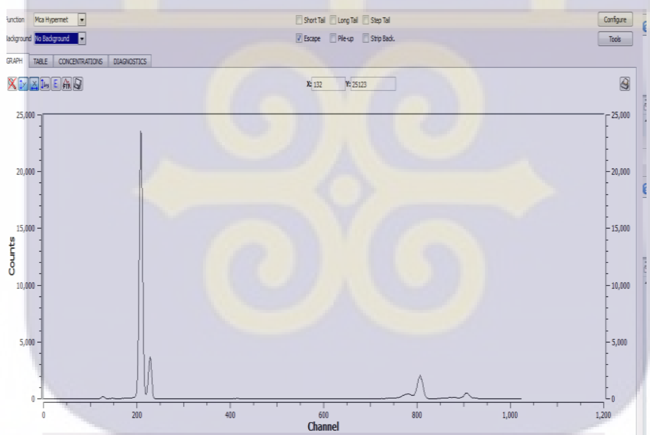


Fig. B21 MCA Advanced Fit screen (Linear view /Channel view)

This is the MCA Advanced Fit screen.

13. Click the 'log' tab to toggle between linear and logarithmic scales and 'E' to toggle between the energy scale and channel numbers as shown in the next two images.

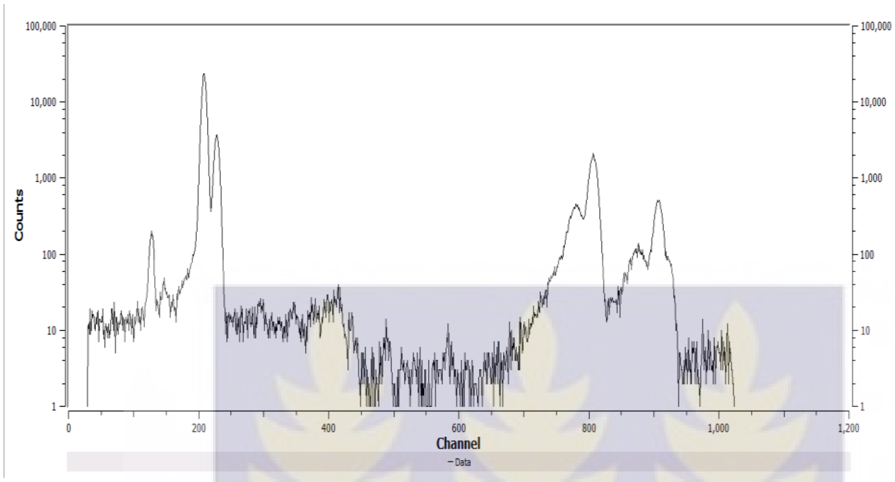


Fig. B22 MCA Advanced Fit screen (Logarithmic/Channel view)

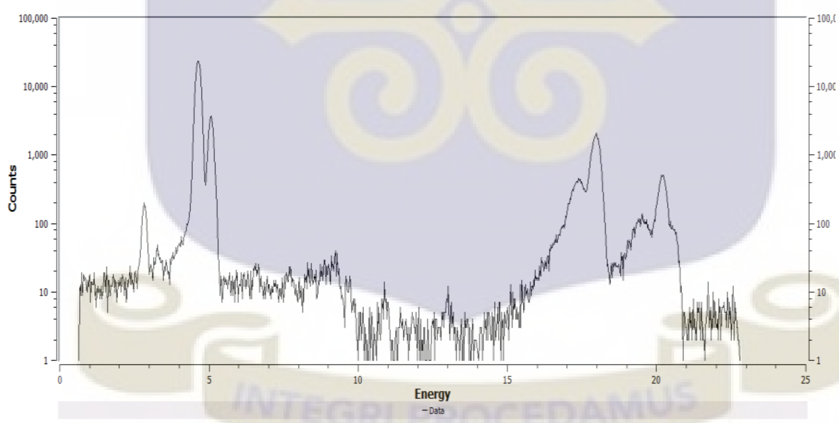


Fig. B23 MCA Advanced Fit screen (Logarithmic/Energy view)

14. Click on the 'Configure' tab to set the background fitting parameters. The sub menu is shown below:

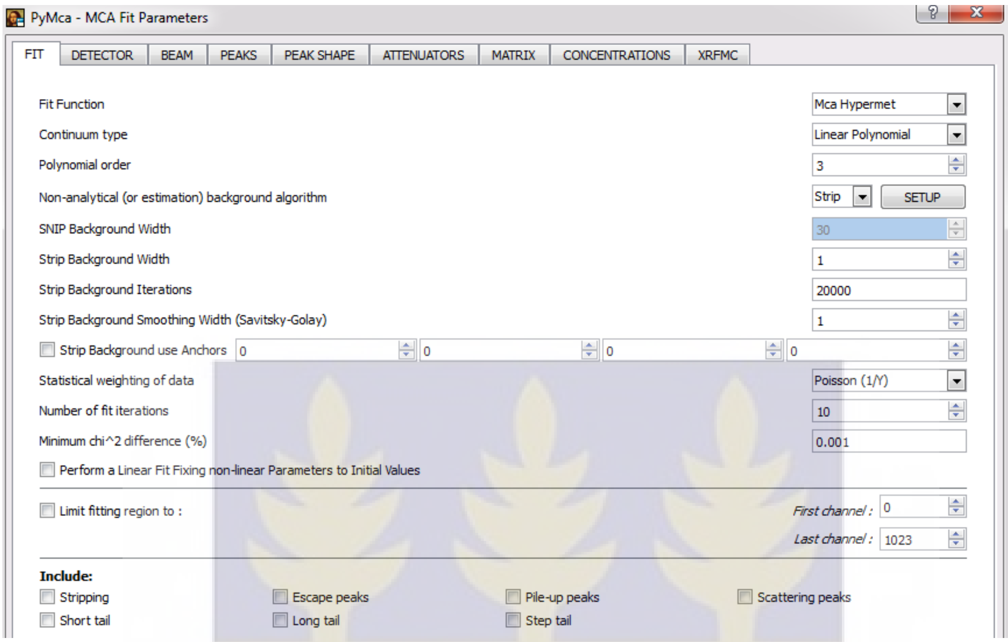


Fig. B24 Selection of background fitting parameters

15. Select the appropriate background fitting parameters(such as ‘Continuum type’ and ‘Polynomial order)’ by clicking on the ‘Fit’ tab. Select the region of interest(ROI) of the spectrum you want to fit by picking the starting & ending channel numbers of the ROI using ‘Limit fitting to:’. To fit the entire spectrum choose ‘First channel:’ to be 0 and ‘Last channel:’ to be 1023.

The experimental parameters under which the spectrum was generated is provided at this stage.

16. To input the x-ray tube voltage and anode target for the x-ray excitation source, click on the ‘Beam’ tab.

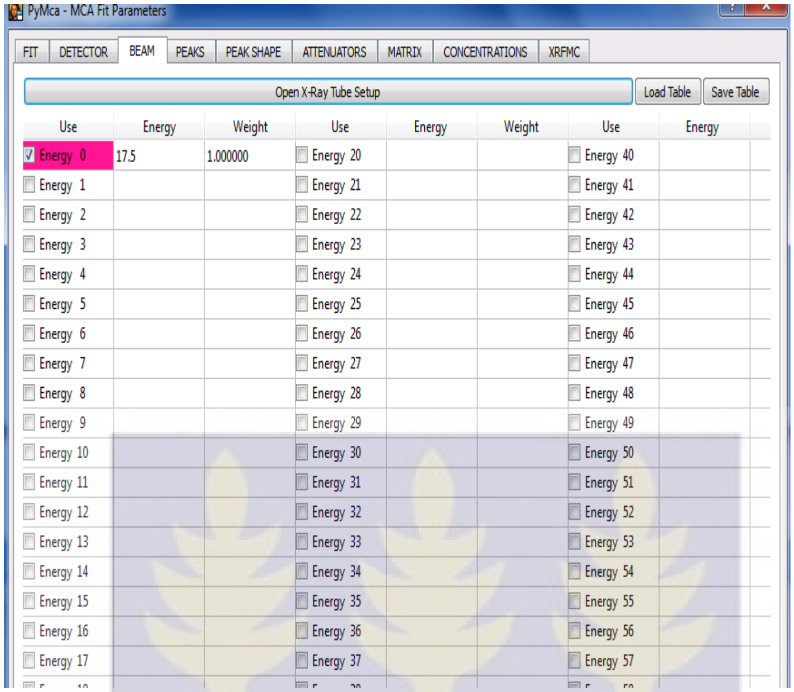


Fig B25 Selection of experimental parameters under which spectrum was generated

Click on ‘Open X-ray Tube Setup’. Input the appropriate x-ray tube voltage (all values are in kV). The user can enter the materials of which the x-ray anode, window and filter are made. This is done by clicking on the downward-pointing arrows opposite. Define the incident energy by double-clicking the top cell in the energy column, selecting the box marked “Energy 0” and inputting the primary energy of characteristic X-radiation being used for irradiation. Make sure the “Weight” is 1.0000.

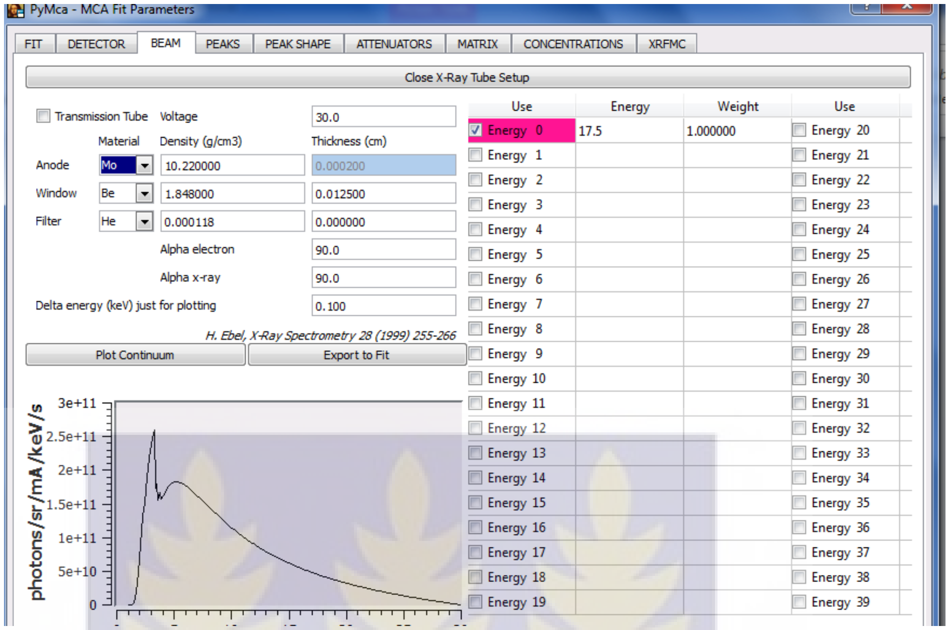


Fig. B26 PyMCA adaptation for X-ray tube excitation sources

16. Choose the elements you want to fit by clicking on the ‘Peaks’ tab and clicking on the appropriate element. Pick the lines to be fitted (whether Kα or Kβ or Lα(L1), Lβ(L2) or Lγ (L3)).

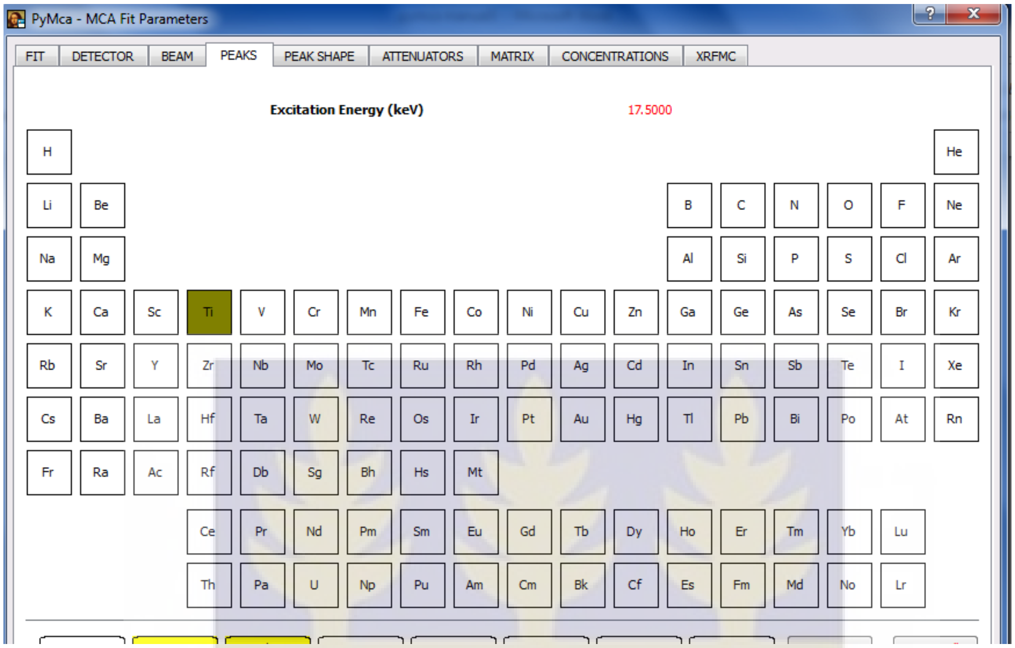


Fig. B27 Selection of an element to be fitted

For a multi-elemental sample, the selection of elements with lines to be fitted is represented in the 'Peaks' tab as:

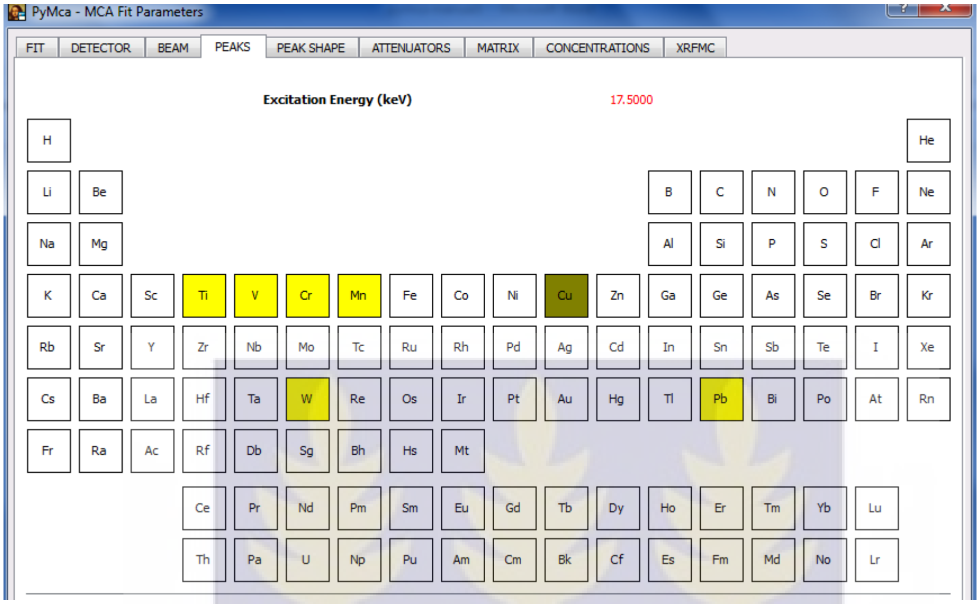


Fig. B28 Selection of multiple elements to be fitted

Clicking on an element highlighted in yellow shows the particular lines of that element (whether K, Ka, Kb, L, L1, L2, L3 or M) that will be fitted.

17. The ATTENUATORS tab allows the user to take into account absorbers that could modify the peak ratios of the elements as seen by the detector. Make sure you define everything that can attenuate the beam between sample and detector-a very important step if quantitative analysis is being done. Here, select ‘MULTILAYER’ as the material for ‘Matrix’.

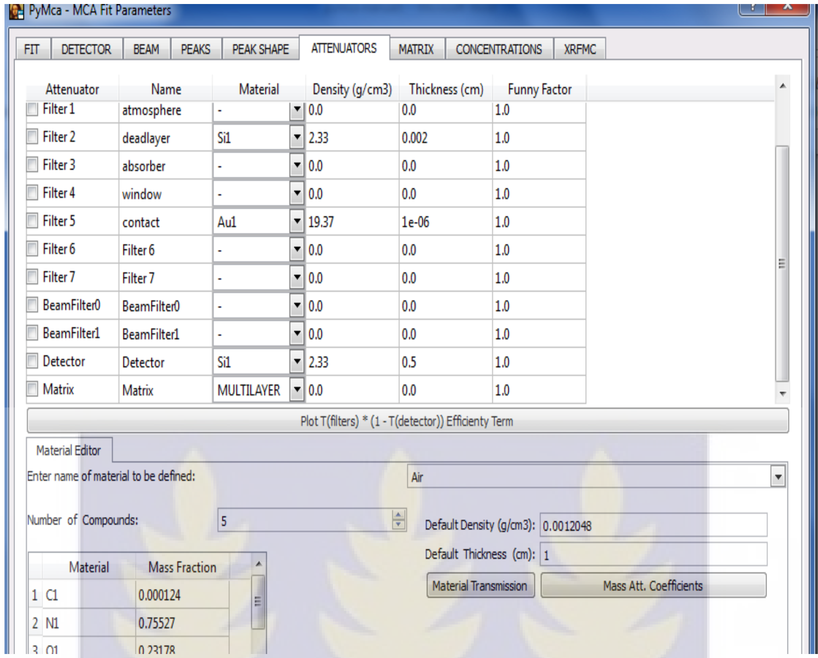


Fig. B29 Selection of attenuators present

18. To enter the source-sample-detector geometry, click on the 'Matrix' tab. This allows the user to enter incident and reflected angles.

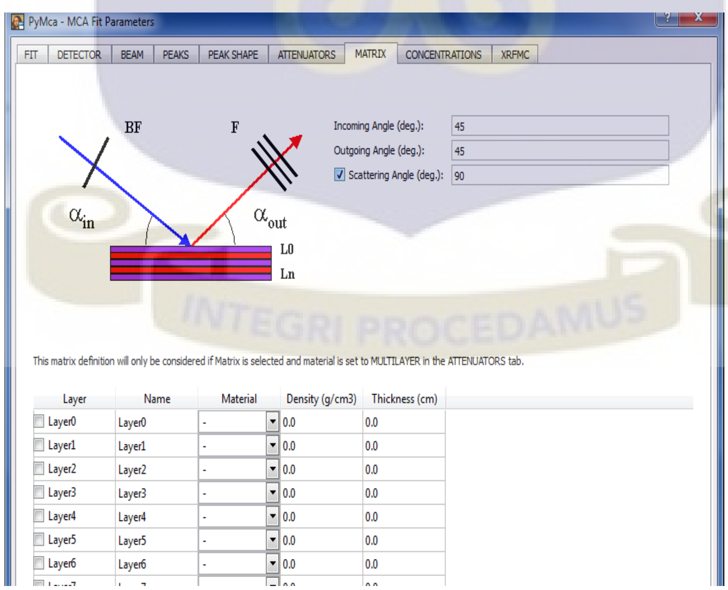


Fig. B30 Entering incident & reflected angles

19. To calculate concentrations, click the ‘Concentrations’ tab; check the ‘From fundamental parameters’ box, enter the photon flux, the time of measurement/irradiation, active area of your detector and the distance between sample and detector.

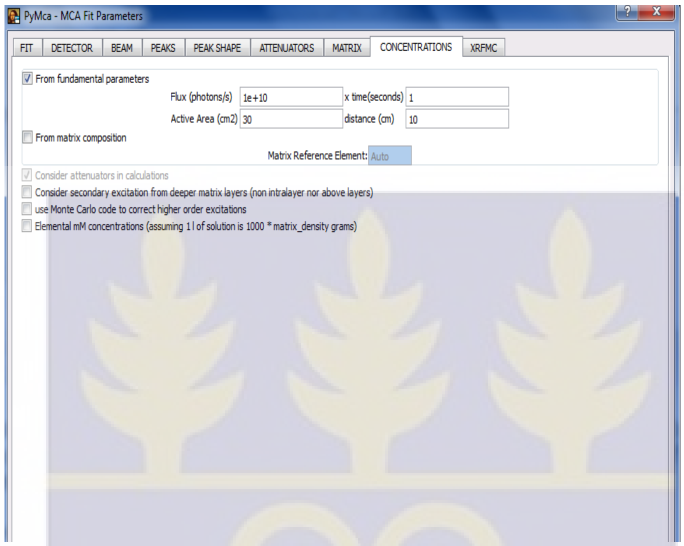


Fig. B31 Entering necessary parameters for concentration calculation

20. Save your configuration to a .cfg file by clicking the ‘Save’ tab. Once again one should save in a location other than the PyMCA root drive preferably on the hard drive. It would be advisable to save in a folder on the hard drive.

21. Click the ‘OK’ tab at the lower right-hand corner to exit. This takes you back to the Mca Advanced Fit screen.

22. Click on the ‘Fit’ icon at the top of the screen to perform the fit.

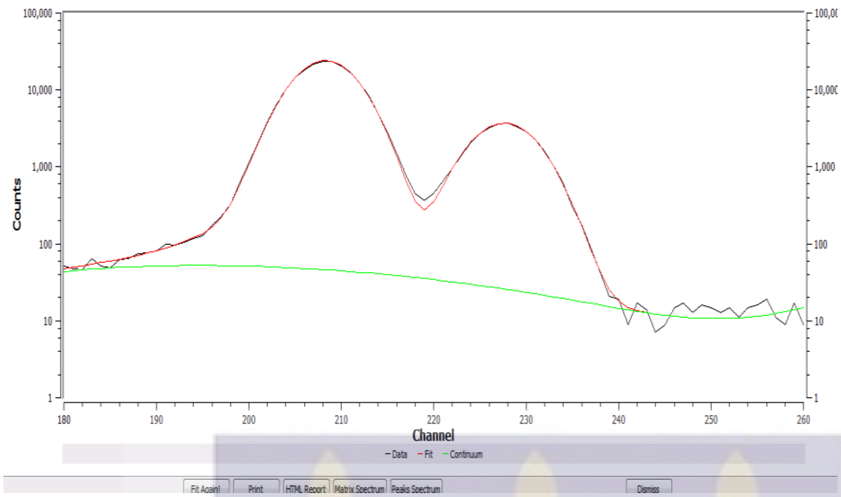
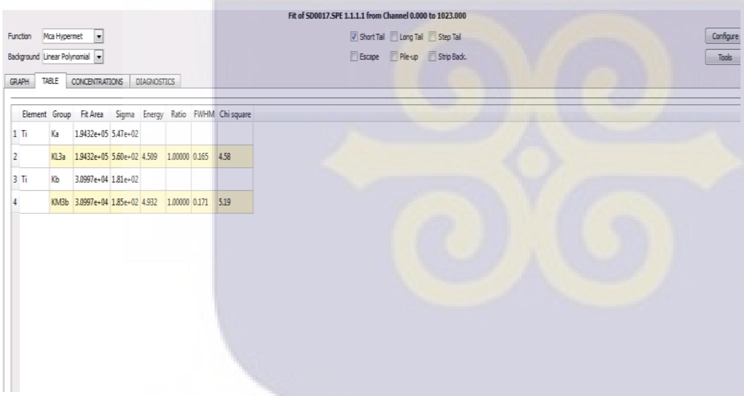


Fig. B32 Fit results

23. Click on the ‘TABLE’ tab to see the results of the fit.



	Element	Group	Fit Area	Sigma	Energy	Ratio	FWHM	Chi square
1	Ti	Ka	1.9432e+05	5.47e+02				
2		KL3a	1.9432e+05	5.60e+02	4.509	1.00000	0.165	4.58
3	Ti	Kb	3.0997e+04	1.81e+02				
4		KM3b	3.0997e+04	1.85e+02	4.932	1.00000	0.171	5.19

Fig. B33 Table of Fit results

24. If a peak in your spectrum has not been fit, you may need to include another element in your model. Place your cursor over the peak and read the X-position in the box above the plot window (make sure you have selected the E button above the graph to get an energy scaled X-axis). Select Tools, Identify Peaks and enter the energy in the top box. The window lists possible candidates.

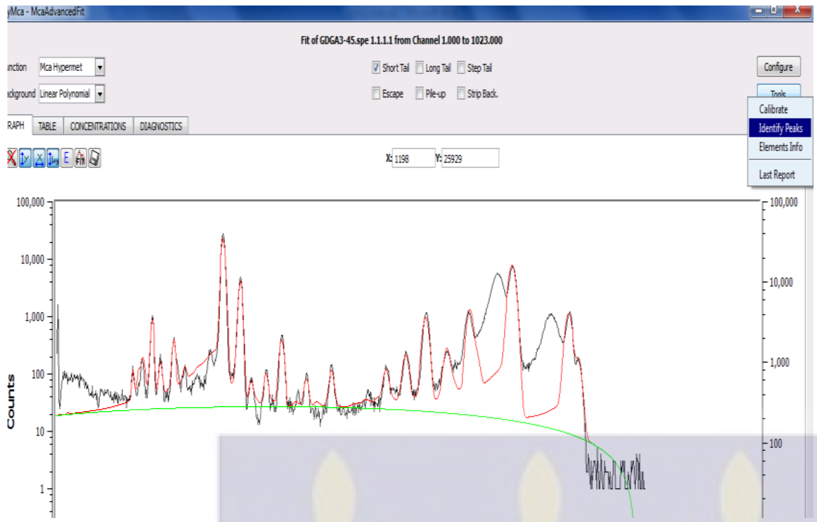


Fig. B34 Identification of missing energies

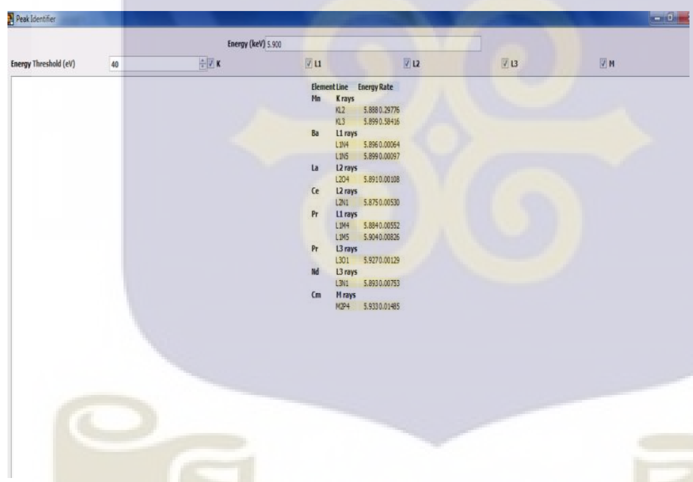
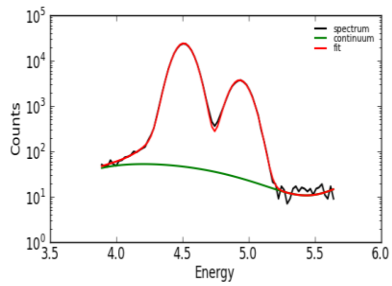


Fig. B35 Peak identifier screen

25. After fitting you can generate an HTML report by clicking on the 'HTML Report' tab at the bottom of the screen. By doing this a pop-up menu appears for you to select the output directory. Clicking on 'Choose' outputs the report in this directory which could be a folder or a drive.

Spectrum, Continuum and Fitted values :



Fit Parameters :

FIT parameters	
Region of Fit	180 - 260
Number of iterations	6
Chi square	2.9212
Last Chi square difference	29.3002 %

Fig. B36 Portion of HTML report

Under the Tools sub-menu are the following options:

Hide Source - Hides the source selection widget allowing the use of a larger graphics display.

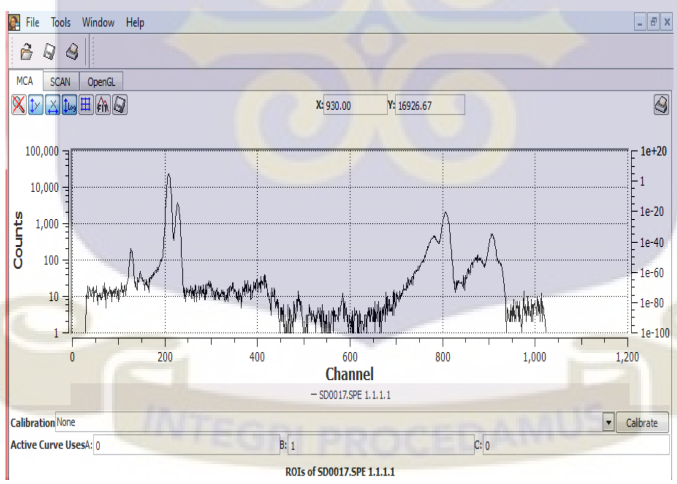


Fig. B37 Viewing a larger graphic display

Show Source - Shows a previously hidden source selection widget.

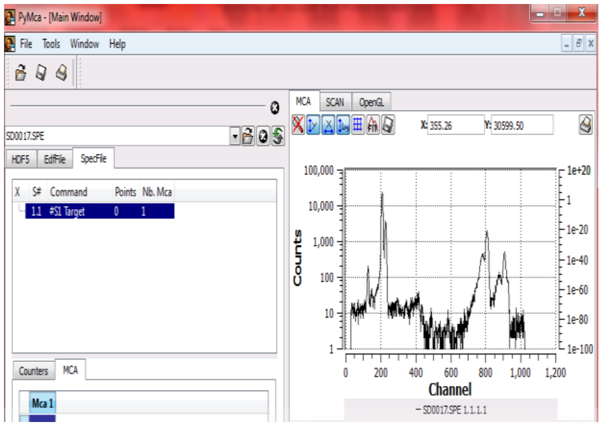


Fig. B38 Normal graphics display

Elements Info - Shows a periodic table to get X-ray properties of the elements.

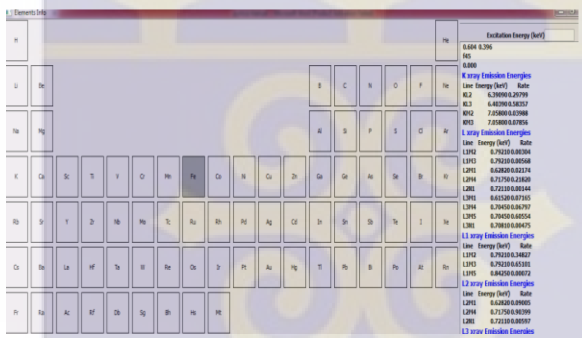
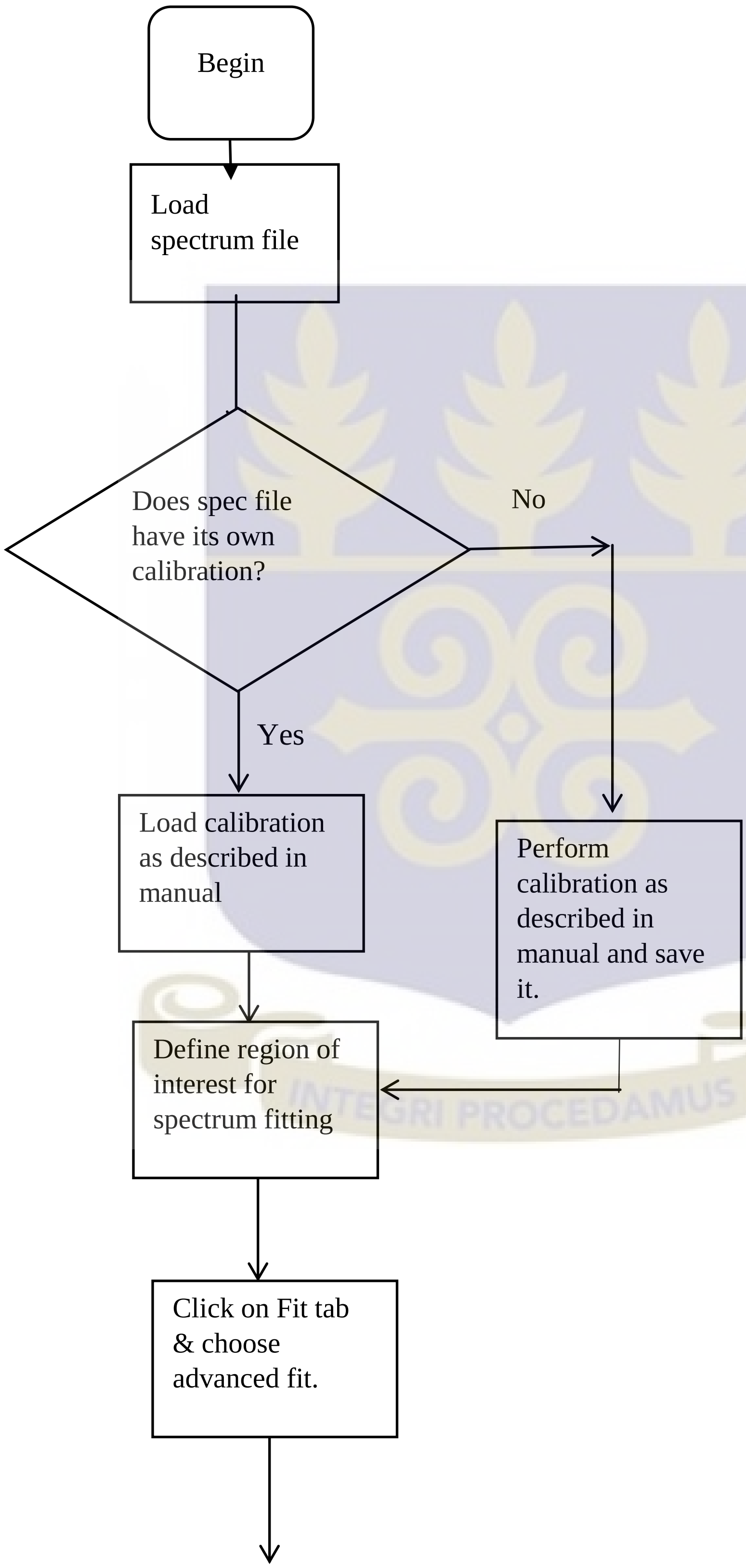


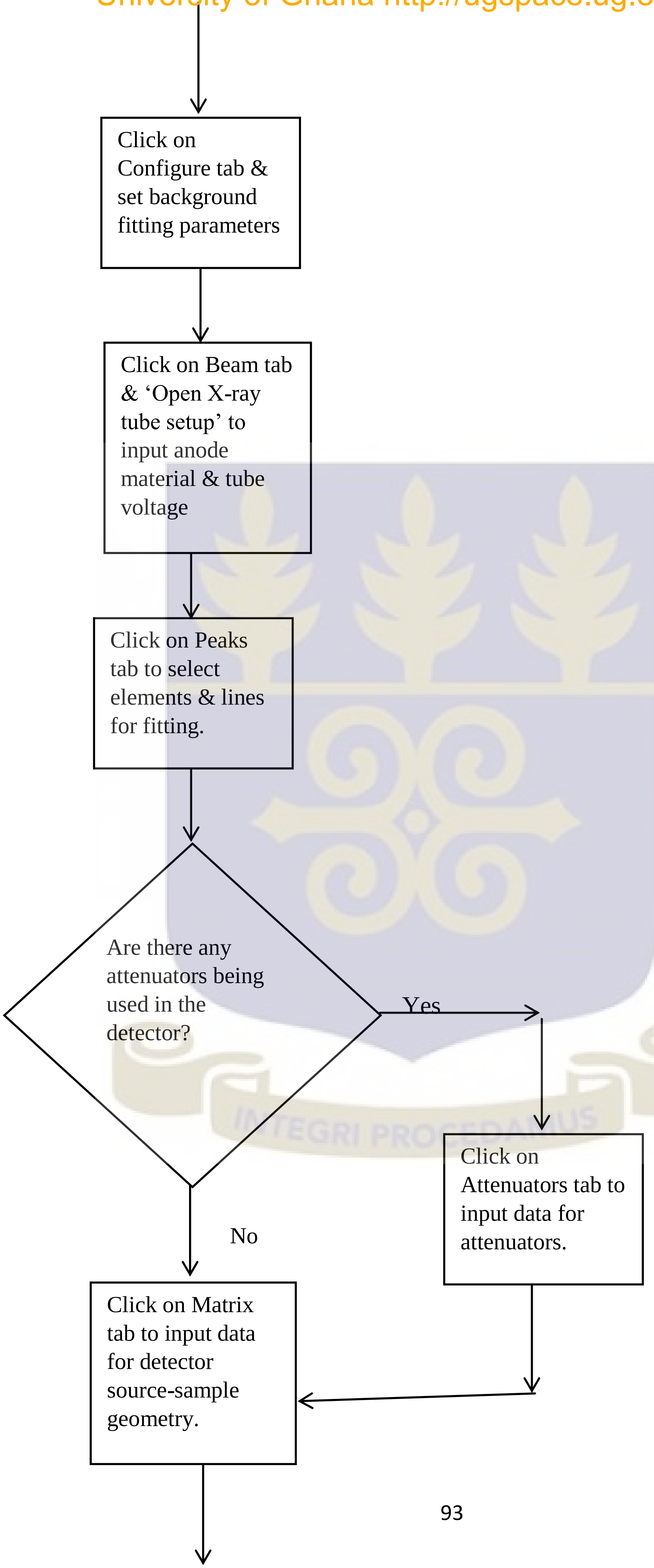
Fig. B39 Periodic table with X-ray properties of indicated element (Fe) to the right.

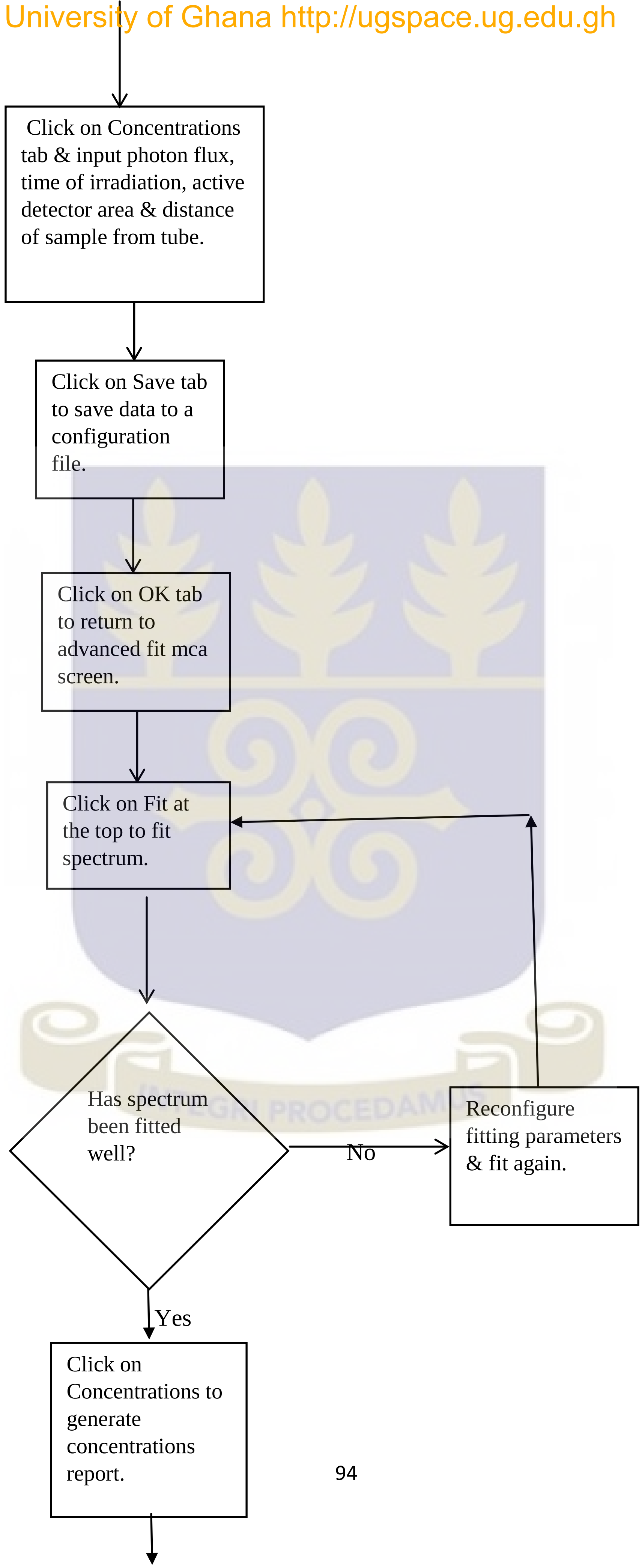
The user should refer to the 'Help' section of the PyMCA software. Some aspects of this manual are taken from that portion of the software; also the pdf file 'PyMCA Chess' was consulted in preparing this manual as well as the page pymca.sourceforge.net/documentation.

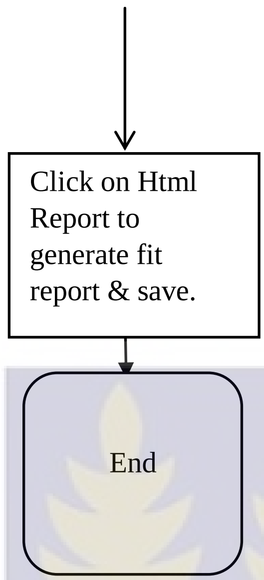
APPENDIX C

FLOWCHART OF SIMPLIFIED PyMCA MANUAL




Click on Configure tab & set background fitting parameters
Click on Beam tab & ‘Open X-ray tube setup’ to input anode material & tube voltage
Click on Peaks tab to select elements & lines for fitting.
Are there any attenuators being used in the detector?
Yes
Click on Attenuators tab to input data for attenuators.
No
Click on Matrix tab to input data for detector source-sample geometry.


Click on Concentrations tab & input photon flux, time of irradiation, active detector area & distance of sample from tube.
Click on Save tab to save data to a configuration file.
Click on OK tab to return to advanced fit mca screen.
Click on Fit at the top to fit spectrum.
Has spectrum been fitted well?
No
Reconfigure fitting parameters & fit again.
Yes
Click on Concentrations to generate concentrations report.


Click on Html
Report to
generate fit
report & save.
End

REFERENCES

Ang, J., Guangyi, T., Shangjun, Z. and Liqiang, L. (2003) X-Ray Fluorescence Spectrometry, Beijing: Science Press, 99.

Bennun, L., Greaves, E.D. and Blostein, J.J. (2002). New procedure for intensity and detection limit determination in spectral trace analysis: application for trace mercury by TXRF. X-Ray Spectrometry 31, 289–295

Bertin, E. P, (1975). Principles and Practice of X-ray Spectrometric Analysis. Plenum Press, New York, New York, 2nd edition.

Brunetti, A and Steger, T.J. (2000). X-Ray spectra background fitting by projection onto convex sets. Nuclear Instrumentation Methods Physics Research, A441, 504.

Brunetti, A. and Golosio, B (2001). Fit of EDXRF spectra with a genetic algorithm X-Ray Spectrometry, 30, 32–36

Fiori, C.E., Myklebust, R.L. and Gorler, K. (1981).Proceedings Workshop on Energy-Dispersive X-ray Spectrometry, editions K.F.J Heinrich et al., National Bureau of Standards, Gaithersburg, MD 20760, March.

Janssens, K and van Espen, P. (1986) Evaluation of energy-dispersive x-ray spectra with the aid of expert systems. Analytica Chimica Acta, 191, 169-180

Jenkins, R., Gould, R.W. and Gedcke, D. (1981). Quantitative X-ray Spectrometry. Marcel Dekker, New York.

Lam, C.F., Forst, A. and Harvey Bank (1979). Simplex: A Method for Spectral Deconvolution Applicable to Energy Dispersion Analysis. Applied Spectroscopy, 33, 3.

Liu, H., Liu, S., Zhang, Z., Sun, J. and Shu, J. (2014). Adaptive total variation-based spectral deconvolution with the split Bregman method. Applied Optics. 53(35):8240-8

Markowicz, A.A. (1993) X-ray physics in Handbook of X-ray Spectrometry, (Editors: R. E. Van Grieken, A. A. Markowicz) New York, NY.

Morhac, M (2009). An algorithm for determination of peak regions and baseline elimination in spectroscopic data. Nuclear Instruments and Methods in Physics Research A600, 2 March. 478-487.

Morhac, M. and Matousek, V. (2008). Peak Clipping Algorithms for Background Estimation in Spectroscopic Data. Applied Spectroscopy, 62, 1.

Ofosu, F.G. (1995) Study of the Effects of the Choice of Different Background Parameters on the Accuracy of X-Ray Intensity Evaluation In Peak Fitting Using the AXIL Software Program MPhil. Thesis, University of Ghana

P. Bevington, Data Reduction and Error Analysis for the Physical Sciences, McGraw-Hill, New York, 1969.

Park ,D.S, Sukanya J, Moonyong ,J and Sook, Y (2013). An efficient background modelling & correction method for EDXRF spectra. Journal of the Institute of Electronics and Information Engineers, 8, 238-244

Peter, B (2006) Theory of XRF, PANalytical BV, Almelo 47.

Ryan, C.G., Clayton, E. Griffin, W.L., Sie, S.H., Cousens, D.R. (1988). SNIP, a statistics-sensitive background treatment for the quantitative analysis of PIXE spectra in geoscience applications. Nuclear Instruments and Methods in Physics Research B34 396-402

Savitzky, A. and Golay, M.J.E. (1964). Smoothing and differentiation of data by simplified least squares procedures. Physical Review.

Schamber, F.H. (1973). A new technique for deconvolution of complex x-ray energy spectra, Proceedings of the Eighth National Conference on Electron Microscope Symposium, 85.

Schamber, F.H. (1976) A modification of the linear least-squares fitting method which provides continuum suppression (Tracor Northern, Inc., Middleton, WI).

Schamber, F.H. (1979) X-ray fluorescence Analysis of Environmental Samples ,edition T.G. Drisbay, Ann Arbor Science Publishers, Michigan.

Schlotz, R. and Uhlig, S. (2000).Introduction to X-ray Fluorescence (XRF). Bruker AXS Inc., Madison, Wisconsin.

Sole, V.A. pymca.sourceforge.net/documentation. Retrieved on 21 May, 2015.

Sole, V.A., Pappillon, E., Cotte, M., Walter, Ph. and Susini, J. (2007). A multiplatform code for the analysis of energy-dispersive X-ray fluorescence spectra, Spectrochimica Acta Part B 62, 63-68

Sole, V. A (2015). Private Communication.

Steenstrup, S (1981). A simple procedure for fitting a background to a certain class of measured spectra. Journal of Applied Crystallography 14, 226.

Tchantchane, A. (1993) Proceedings of Workshop on Trace Element Mobilization and Determination in Different Media by XRF.

van Espen, P., Lemberge, P. (2000) EDXRF spectrum evaluation and quantitative analysis using multivariate and nonlinear techniques. International Centre for Diffraction Data, Advances in X-ray Analysis, 43 van Grieken R.E and Markowicz, A.A. (1993). editors Handbook of X-ray Spectrometry Methods and Techniques, volume 14. New York - Basel - Hong Kong.

Vekemans, B., Janssens, K., Vincze, L., Adams, F. and van Espen, P. (1994) Comparison of several background compensation methods useful for evaluation of energy-dispersive X-ray fluorescence spectra .X-Ray Spectrometry 23, 278.

Wang Z., Ge, L.Q, Zhang, QC. Xu, L, Luo, Y (2012).Application of Fourier transform background subtraction analysis. Nuclear Techniques, 7, 549-551

Wegrzynek, D., Markowicz, A. and Mendoza Cuevas, A. (2001). Evaluation of the energy-dispersive x-ray spectra of high-Z elements using Gaussian and Voigt peak shape profiles X-Ray Spectrometry 30, 403–412

Yi, L., Liu, Z., Kai Wang, Man Chen, Peng, S., Zhao W., He, J. and Zhao, G. (2015) A new background subtraction method for EDXRF using cubic spline interpolation. Nuclear Instrumentation Methods Physics Research, A775, 12

Zhao, F. and Wang, A. (2014). A background subtraction approach based on complex wavelet transforms. X-ray Spectrometry 44, 41-47