# Comparison of Algorithmic and Machine Learning Approaches for the Automatic Fitting of Gaussian Peaks

R. E. Abdel-Aal
Center for Applied Physical Sciences, Research Institute,
King Fahd University of Petroleum and Minerals,
Dhahran 31261, Saudi Arabia

**Abstract**

Fitting gaussian peaks to experimental data is important in many disciplines, including nuclear spectroscopy. Nonlinear least squares fitting methods have been in use for a long time, but these are iterative, computationally intensive, and require user intervention. Machine learning approaches automate and speed up the fitting procedure. However, for a single pure gaussian, there exists a simple and automatic analytical approach based on linearization followed by a weighted linear least squares (LS) fit. This paper compares this algorithmic method with an abductive machine learning approach based on AIM† (Abductory Induction Mechanism). Both techniques are briefly described and their performance compared for analysing simulated and actual spectral peaks. Evaluated on 500 peaks with statistical uncertainties corresponding to a peak count of 100, average absolute errors for the peak height, position, and width are 4.9%, 2.9%, and 4.2% for AIM versus 3.3%, 0.5%, and 7.7% for the LS. AIM is better for the width, while LS is more accurate for the position. LS errors are more biased, underestimating the peak position and overestimating the peak width. Tentative CPU time comparison indicates a 5-fold speed advantage for AIM, which also has a constant execution time while LS time depends on the peak width.

Keywords: Spectral Analysis, Spectroscopy, Peak Fitting, Machine Learning, Gaussian Peaks, Abductive Networks, Least Squares Fit.

Dr. R. E. Abdel-Aal
P. O. Box 1759
KFUPM
Dhahran 31261
Saudi Arabia

Tel.: +966 3 860 4320
Fax: +966 3 860 4281
e-mail : radwan@kfupm.edu.sa

---

† AIM is a Registered Trademark of AbTech Corporation, Charlottesville, VA, USA.

## 1. Introduction

Conventional computational techniques have long been used for identifying and quantifying radioisotopes through the analysis of alpha, gamma-ray, and X-ray spectra, e.g. [1-3]. The measured spectrum is searched for characteristic peaks and iterative procedures are used for matching a set of fitted peaks of various shapes and intensities to the spectrum peaks. The fitting procedure is often carried out interactively and may require considerable user intervention, particularly for multiplets [4]. Initial parameters which influence the result, e.g. multiplicity and tentative estimates of peak locations, need to be determined manually, and therefore the quality of the fit depends on the user [5]. The iterative nature of the procedure makes it computationally intensive, and it may not converge to a solution [5]. All the computational complexities have to be repeated for each individual spectrum analyzed, even though it may be quite similar to a previous one. These limitations need to be overcome in order to achieve the desirable goals of simplifying, speeding up, and automating spectrum analysis.

A recent trend in many areas of applied sciences has been to resort to a machine learning approach when a rigorous algorithmic solution becomes too complex or time consuming. With this approach, a model is automatically developed through training on an adequate number of solved examples. This model can then be used to perform fast predictions of outputs corresponding to new cases. A major advantage of this approach in spectroscopy analysis is that intensive computations are now required only once, i.e. during offline training for model synthesis, rather than being repeated for every sample analyzed during actual use. Using the model to process new spectra becomes a simple and speedy operation which can be implemented in real time using compact and portable apparatus. These are important requirements in applications where speed and simplicity are more important than high accuracy, such as in determining contaminants in waste sites [6]. With the spectrum shape

as input, the problem is reduced to that of automatic pattern recognition which has been applied in many disciplines, e.g. [7]. A machine-learning system can automatically detect the multiplicity of a peak and then determine the peak parameters without requiring user intervention.

A Number of techniques currently exist for the development of machine learning systems [8]. These include statistical pattern recognition methods such as Bayesian classifiers and discriminant functions [9], artificial neural networks [10], as well as methods for the induction of decision trees [11]. These techniques vary in their accuracy, complexity, computational requirements during training, and their ability to provide human-like explanations for their conclusions. Such variations have led to newer techniques combining good features from various methods. An example of such 'hybrids' is the AIM abductive network [12] tool which draws on statistical and multiple regression analysis methods as well as neural networks, resulting in a faster and more automated approach to model synthesis. A mathematical description of the foundations of AIM is given in [13]. Neural network techniques have been proposed as a new approach to automate radiation spectrum analysis [6,14,15]. Trained on a set of reference spectra for isotopes expected to be present in the samples of interest, the network can identify and quantify those isotopes in the spectrum of an unknown mixture. AIM also was used to build models for determining the intensity of radioisotopes in both single- and multiple-isotope samples represented by their gamma-ray spectra collected using an NaI (Tl) scintillation detector [16]. These techniques operate on the spectrum as a whole instead of analyzing each individual peak.

With the conventional computational approach, generalized fitting of spectral peaks with background is usually performed using iterative nonlinear least squares methods (LS) which minimize the weighted sum of the squared errors between the actual data and the data predicted by the combination of fitted functions, i.e.

$$\chi^2 = \sum_{i=1}^{N} w_i \left[ Y_i - y(x_i) \right]^2$$

(1)

where $Y_i$ are the measured counts at the ith channel $x_i$ and $y(x_i)$ are the corresponding computed values from the set of theoretical functions used for the fit, $w_i$ are the weights associated with the data values $Y_i$, and N is the number of channels used for the fit. Optimum parameters for the fitted functions are obtained through an iterative optimization procedure with tentative initial estimates for these parameters being supplied by the user. Iteration is repeated until a convergence criterion is met. In addition to the need for user intervention, such fitting routines are generally complex with large computational requirements. These factors have made it desirable to develop simpler methods for automating the analysis of gaussian peaks using small portable apparatus. A number of machine learning solutions have been described for this purpose. Neural networks were used to identify and evaluate peaks in vapor-phase infrared spectra in the presence of noise [17]. Ref. [18] reports on the use of an analog neural network for fitting the gaussian shape of an ion beam hitting a 16-segment strip detector. The AIM machine learning approach was also used to build models for the automatic identification and analysis of both single and double gaussian peaks in the presence of noise depicting statistical uncertainties in collected spectra [13]. The main peak fitting application considered here is analyzing counting spectra. More recently, other applications have evolved, such as estimating the parameters of gaussian or semi-gaussian shaped nuclear pulses whose waveforms are digitized and acquired using fast (flash) analog to digital converters, including techniques for pulse pileup detection and correction [19].

A pure single gaussian is often adequate for fitting gamma ray spectra from semiconductor detectors [4]. Background is usually negligible in some practical situations including the analysis of alpha spectra [4]. For this special case of a pure gaussian, an analytical non-iterative solution exists for the peak parameters [20,21]. In situations where the statistical errors follow a Poisson distribution, a quadratic iterative fit in the form of Poisson regression [22] can be used to determine the peak parameters. The non-iterative method utilizes a linearization technique for the gaussian data followed by a standard weighted linear least

squares fit. Being fast and automatic, the method competes with the more recent machine learning techniques. This paper compares this linear LS fitting method with the AIM machine learning approach in terms of the accuracy of estimating the peak parameters, the quality of the peak fits, as well as computational complexity and speed. The AIM peak fitting approach was compared elsewhere [13] with the nonlinear LS peak fitting method of Eqn. (1), but the interactive nature of the latter practically limits the comparison to only a few examples and makes it difficult to compare performance on a statistically large evaluation set. Following an overview of the linear LS and the AIM peak fitting methods, data are presented on comparing their performance using both simulated and measured alpha spectral peaks. Statistical data are presented on the accuracy of estimating the peak parameters for a large simulated evaluation sample. Graphical representations and reduced Chi-squared data on example fits are given and results on CPU time measurements for both techniques on a VAX computer are presented.

## 2. AIM Modeling of Single Gaussian Peaks

### 2.1. AIM Overview

AIM is a supervised inductive machine-learning tool for automatically synthesizing abductive network models from a data base of input and output values which represent a training set of example situations. Once synthesized by training on a training data set, the network can be queried with new input data to provide the corresponding predicted output. Abductive networks [12] combine the advantages of the neural network approach with those of advanced statistical methods. A comparison between AIM machine learning and the conventional neural network approach is given in [13].

The work reported here used AIM version 1.0 for the Macintosh computer. AIM models take the form of layered feed-forward abductive networks of functional elements (nodes) [23], see Figure 1. Elements in the first layer operate on various combinations of the independent input variables (x's) and the single element in the final layer produces the predicted output for the dependent variable y. In addition to the functional elements in the main layers of the network, an input layer of normalizers converts the input variables into an internal representation as Z

scores with zero mean and unity variance [23], and an output layer of unitizers restores the results to the original problem space. Both the element type and the combination of inputs to it from all the previous layers are selected automatically for best prediction performance. The used version of AIM supports the following main elements:

(i) A white element which consists of a constant plus the linear weighted sum of all outputs of the previous layer, i.e :

$$\text{"White" Output} = W_0 + W_1X_1 + W_2X_2 + W_3X_3 + .... + W_nX_n \quad (2)$$

where $X_1$, $X_2$,..., $X_n$ are the inputs to the element and $W_0$, $W_1$,..., $W_n$ are the element weights.

(ii) Single, double, and triple elements which implement a third-degree polynomial expression with all possible cross-terms for one, two, and three inputs respectively; for example,

$$\text{"Double" Output} = W_0 + W_1X_1 + W_2X_2 + W_3X_1^2 + W_4X_2^2 + W_5X_1X_2 + W_6X_1^3 + W_7X_2^3 \quad (3)$$

The first step in solving a problem with AIM is preparing a data base of input-output training examples which AIM uses to synthesize the network. AIM synthesizes the model network layer by layer until no further improvement in performance is possible or a preset limit on the number of layers is reached. Within each layer, every element is computed and its performance scored for all combinations of allowed inputs. The best network structure, element types and coefficients, and connectivity are all determined automatically by minimizing the predicted squared error (PSE) criterion [24]. This selects the most accurate model that does not overfit the training data, and therefore strikes a balance between the accuracy of the model in representing the training data and its generality which allows it to fit yet unseen future data. In this way the model is optimized for actual prediction for which it is developed, rather than simply fitting the training data. The user may optionally control this trade-off between accuracy and generality using the complexity penalty multiplier (CPM) parameter [23]. Larger values than the default value of 1 lead to simpler models which are less accurate in fitting the training data but are more likely to generalize well with unseen data during actual use. Lower CPM values produce more complex networks which overfit

the training data and may therefore degrade prediction performance in the presence of noise. AIM's 'Evaluate' utility allows evaluation of the resulting AIM network on an independent set of data and generates a report of the results. To obtain good AIM models, the training set should be a good representation of the problem space. AIM's learning task is also simplified by breaking the problem into smaller and more manageable assignments, and by utilizing any human knowledge on parameters relevant to the model in the choice of input variables to be included in the training data base.

## 2.2. Fitting Gaussian Peaks

AIM models were previously developed [13] for automatically determining the peak height A, position $\mu$, and full-width-half-maximum (FWHM) width $\Gamma$ of single gaussian peaks with no additional background. The models were synthesized by training on 1000 simulated spectra of single peaks having randomly generated parameters. A spectrum consisted of 50 channels, each channel representing an input variable to AIM. This restriction on the spectrum size was imposed by the maximum limit of 50 input variables for the version of AIM used. To make the AIM model suitable for use with a wide range of actual spectral counts, AIM was trained and evaluated on normalized spectra having peak values less than or equal to unity. Such a spectrum is obtained by dividing the counts at each channel of the actual spectrum by an appropriate integer scaling factor, N. Simulated data for a normalized noiseless spectrum having a single gaussian peak with no background were computed as

$$z(i) = A \exp\left[\frac{-(i-\mu)^2}{2\sigma^2}\right] \quad ;i = 1,2,...,50 \tag{4}$$

where the standard deviation width $\sigma$ is $\Gamma$/2.355. The normalized peak height A had a real value in the range 0.1 to 1.0. This normalized spectrum represents an actual spectrum

$$y(i) = P \exp\left[\frac{-(i-\mu)^2}{2\sigma^2}\right] \quad ;i = 1,2,...,50 \tag{5}$$

where y(i) represents the actual count at channel i. z(i) is obtained from y(i) using

$$z(i) = \frac{y(i)}{N} \tag{6}$$

where N is the integer scaling factor chosen to give a maximum value of 1 for the normalized peak A.

To ensure that the randomly generated values for $\mu$, and $\sigma$ produce peaks that are located well within the 50-channel spectra ($\pm 3\sigma$ region), the following algorithm was adopted for selecting the peak parameters for both the training and evaluation sets: (both $\mu$ and $\sigma$ are integers)

1. Select $\mu$ in the range 10 to 40

2. Determine d as min ($\mu$ , 50-$\mu$)

3. Determine the maximum value allowed for $\sigma$ as $\sigma_{max} = d/3$

4. Select $\sigma$ in the range 3 to $\sigma_{max}$

This produces peaks with positions between channels 10 and 40 and FWHM widths between 8 and 20 channels. Peaks positioned closer to the edges of the 50-channel spectrum region have proportionally narrower widths to ensure that at least a region of $\pm 3\sigma$ around the peak center is contained in the 50-channel spectrum. Noise depicting statistical uncertainties in measured counts was introduced as an additional component to the idealized values z(i) of the normalized spectrum of Eqn. (4). Let $\Delta y(i)$ be the standard deviation for the statistical variations in actual counts y(i). In counting experiments where 20 or more counts can be accumulated, statistical variations around the mean count can be assumed to follow the normal distribution [20]. In this case, the standard deviation is given by:

$$\Delta y(i) = \sqrt{y(i)} \tag{7}$$

Relative variations are preserved by the normalization step in Eqn. (6), and therefore the corresponding standard deviation for z(i) is given by:

$$\Delta z(i) = \frac{\sqrt{z(i)}}{\sqrt{N}} \tag{8}$$

Uniformly distributed random variations within ±Δz(i) about z(i) were introduced by adding the following noise component n(i) to the computed value of z(i) in Eqn. (4):

$$n(i) = \Delta z(i)\left[-1 + 2\,\mathrm{rnd}\,(1.0)\right] \tag{9}$$

where rnd(1.0) is a random number between 0 and 1.0. The standard deviation for n(i) is $\Delta z(i)/\sqrt{3}$. We consider two noise levels, a high level with N = 100, and a low level with N = 10 000. For a normalized peak channel having z(i) = 1.0, the maximum percentage noise corresponding to these two noise levels are 10% and 1%, respectively.

AIM training/evaluation database consists of records constructed from the simulated spectra by appending known values for the peak height, position, and width. A typical AIM record for single peak modeling is represented below:

| Inputs : Spectrum Channel Contents | | | | | | | | Outputs : Corresponding Peak parameters | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Ch_1 | Ch_2 | Ch_3 | ... | ... | ... | ... | Ch_50 | Height (A) | Position (μ) | Width (Γ) |

AIM generates a model for each variable declared as output in the training database. Training was performed on noisy data at the high noise level (N = 100). Structure of the model obtained with AIM's default complexity penalty multiplier (CPM = 1) is shown in Figure 2. The peak height is generated by the simplest of the three networks, which consists of two layers with contributions from about 70% of the individual spectrum channels. Increasingly more complex networks with 2 and 3 layers are required to model the peak width and position, respectively. Equations for all the functional elements of the AIM model are given in Figure 3. These determine the peak height (A), position (μ), and width (Γ) automatically from the channel contents, Ch_1, ..., Ch_50, of the 50-channel normalized spectrum representing the peak to be analyzed.

## 3. Linear Least Squares Fitting of a Single Gaussian Peak

Experimental data for a single gaussian peak with no background can be fitted to a pure gaussian to determine the peak parameters using a non-iterative approach. The fitting procedure is simplified through reducing the problem to that of a straight line fit, which is commonly performed using least squares (LS) or maximum likelihood techniques [25]. For this purpose, the spectrum data to be fitted are first transformed by obtaining the ratio between the counts at each two alternate channels and then taking the natural logarithm. For the actual counts spectrum in Eqn. (5) we have

$$Q(i) = \frac{y(i-1)}{y(i+1)} = \exp\left[\frac{2(i-\mu)}{\sigma^2}\right]; \tag{10}$$

and $$\ln Q(i) = \frac{2i}{\sigma^2} - \frac{2\mu}{\sigma^2} \tag{11}$$

The logarithm of the counts ratio exhibits a straight line relationship with the channel number i. Fitting a straight line to the data allows both the peak position $\mu$ and the standard deviation width $\sigma$ to be determined directly from the estimated slope and intercept of the fitted line in Eqn. (11). Substituting the values thus obtained for $\mu$ and $\sigma$ into Eqn. (5) at any data point gives a value for the peak height P. However, a more accurate estimate for P is obtained as a weighted mean for all such values [20,21]. Since statistical uncertainties in the ln Q(i) data used for the linear LS fit vary from one data point to the other, a weighted fit is required which minimizes the sum of the squares of deviations from the fitted straight line at each data point but with each squared deviation weighted such that it has more significance for data points with smaller uncertainties. This is achieved through dividing by the square of the standard deviation of the data at that point.

In practice, the relationship in Eqn. (11) departs from that of a straight line away from the peak center, since peaks in practical spectra often deviate from the gaussian shape in these

regions. For this reason, the fit is usually restricted to a smaller region about the peak center, e.g. over the FWHM width. This means that narrower peaks would be faster to process but the fit quality may suffer if the number of data points contributing to the straight line fit is less than adequate. A complete description of the LS peak fitting algorithm used is shown in Figure 4.

## 4. Performance Comparison

### 4.1 Comparison Using Simulation Data

Both the AIM model and the LS fitting algorithm were evaluated on two simulated sets of 500 records each at two noise levels corresponding to low noise (N = 10,000) and high noise (N = 100). Both sets were new to the AIM model, i.e. previously unseen during AIM training for model development. For each peak, the known true peak parameters were compared with corresponding predictions by the two methods. Predicted values for the position and width by both methods were first rounded to the nearest integer prior to further error analysis. Table 1 summarizes data on the percentage of the evaluation sample having percentage errors within ±2%, ±5%, ±10% for the real values of the normalized peak height, and 0%, ±5%, ±10% for the integer values of the position and width. Included also are the values of the overall mean percentage absolute error (MPAE) over the entire evaluation sample of 500 cases.

As expected, errors in all estimated parameters are larger at the higher noise level for both fitting methods. Among the three peak parameters, both the AIM and the LS methods predict the peak position with the best accuracy. For example, the LS method predicts the position for the whole evaluation sample with 100% accuracy at the low noise level. For all three parameters, LS is clearly superior to AIM at low noise since lower statistical variations in the evaluation data simply improve the accuracy of the LS fits. For the high noise data, AIM is

better than the LS method with peak width, while the latter excels markedly with peak position and slightly with peak height.

Figure 5 shows cumulative percentage plots for the percentage errors in the three peak parameters for the whole evaluation set of 500 cases using both the AIM and LS methods. The plots also indicate the maximum percentage errors encountered. Figure 5(a) shows that peak height errors are reasonably symmetrical for both fitting methods, with no significant tendency towards either positive or negative errors, i.e towards underestimating or overestimating the peak height. Figures 5 (b),(c) show that the LS method tends to underestimate the peak position and overestimate the peak width. More quantitative data on this error bias are given in Table 2 which lists the percentage of both negative errors (underestimates) and positive errors (overestimates) over the evaluation sample of 500 cases at the high noise level as well as separate averages for both the negative and positive errors. In general, average values for both negative and positive errors are not significantly different, and therefore error bias in estimating the peak parameters would be largely due to the difference in error frequencies. AIM has nearly equal percentages of positive and negative errors for all the three peak parameters, indicating no significant bias except for a slight tendency to overestimate the peak height. The LS method behaves similarly only for the peak height, with the bias towards larger peak heights being somewhat more pronounced than for the AIM case. When the LS wrongly estimates the peak position (10% of the cases) the estimates have a 2.6:1 higher chance of being underestimated. Wrong peak width estimates (66% of the cases) have a 4.3:1 higher chance of being overestimated. At the low noise level, the position error bias is not likely to be a problem since position errors are expected to be fairly small as indicated in Table 1. However, since LS width errors are much larger at both noise levels, bias in such errors may have undesirable effects, such as distorting the shape of pulse width spectra of digitized gaussian pulses.

Judging whether the fitting method is intrinsically biased or not should take into account the statistical fluctuations in estimates for the entire error population when derived from only the single 500-case evaluation sample considered. Values of the mean error for such samples are assumed to be normally distributed about the true population mean for the error with a standard deviation equal to the population standard deviation/$\sqrt{n}$, where n is the sample size. With n=500, we can assume that the unknown standard deviation of the entire error population is equal to that of the sample. Table 3 shows data on the mean error and standard deviation for the three peak parameters as calculated from the 500-case evaluation sample. We use the z statistic to test the null hypothesis that the true population mean is 0, i.e. parameter prediction is unbiased. The acceptance limits for a 95% confidence level ($\alpha$ = 0.05) are $-1.96 < z < 1.96$. The hypothesis of being unbiased is rejected for the LS algorithm on all three peak parameters, while it is rejected for AIM only on the peak height. This indicates that biases observed tentatively from results in Table 2 appear to be genuine, since they are too significant to be explained by statistical variations about the true mean for the error population if assumed unbiased. The asymmetric error behavior by the LS method, as compared to AIM, may be related to the fact that LS accuracy is poorer for narrower peaks (smaller number of points available for the linear fit). Larger peak positions may also be assigned narrower peaks to ensure they fit within the 50-channel window. Since AIM learning is neutral to the values of the peak parameters, AIM accuracy appears more symmetrical.

Figure 6 shows spectra for peak height, position, and width estimated using both the AIM and LS methods for one million events of a single peak with a nominal height = 50 counts, position = channel number 25, and width = 19 channels with random statistical uncertainties at the high noise level. Variations in the estimated values are represented by the width of the spectrum peaks. For the gaussian peak height and position, Figure 6(a) and (b), AIM variations are larger than LS variations, while LS variations are larger for the width. This observation is in line with the general trend for the corresponding MPAE error data given

in Table 1.

Twelve out of the 500 simulated evaluation spectra with high noise were selected for comparison between AIM and LS on the basis of fit quality for individual peaks. Table 4 lists the actual peak parameters together with those obtained using the two fitting methods. Listed also are the reduced Chi-Squared ($\chi_r^2$) values for the two fits in each case. Out of the twelve cases, AIM has lower $\chi_r^2$ values than LS in five cases while in two cases the fit quality is nearly identical. At the 90% confidence level and 47 degrees of freedom, the critical value $\chi_r^2{}_{(0.90,47)}$ is 1.272. The computed $\chi_r^2$ values significantly exceed this critical value in only one out of the twelve cases for each of AIM and LS. Figure 7 shows plots of the data as well as AIM and LS fits for cases # 5, 7, and 11. These were chosen out of the twelve cases to demonstrate similar fit quality by both methods, superior fit quality by LS, and superior fit quality by AIM, respectively. It is interesting to note that LS excels with the gaussian having smaller statistical uncertainties, Figure 7(b), while AIM excels with the gaussian having larger uncertainties. This agrees with observations made above in connection with Table 1 that LS is superior to AIM at low noise while the opposite is true at high noise.

4.2. Comparison Using Actual Spectral Data

Both AIM and the linear LS fitting methods were evaluated on actual spectrum data by using them to analyze alpha peaks in the spectrum shown in Figure 8(a). The spectrum, collected using a surface barrier detector, is for alpha particles scattered from a ZnS target sandwiched between two Au layers and deposited on a carbon backing. The spectrum was collected under the following conditions: Energy of the incident alpha particles: 6.5 MeV, Angle of detection: 160°, Detector resolution: 18 keV at 5.486 MeV, ADC used: LeCroy 3511, Thickness of the carbon backing: 40 $\mu g/cm^2$, Thickness of the Au and ZnS layers: 60 A° and

2000 A°, respectively. The peaks, marked Peak 1, Peak 2, and Peak 3 in the figure, are the S($\alpha$,$\alpha$'), S($\alpha$,$\alpha$), and Zn($\alpha$,$\alpha$) peaks, respectively. Normalized spectra of 50 channels each were obtained for the three single peak regions through scaling down the actual counts by factors of 40, 136, and 220, respectively for feeding into the AIM model (Figures 2 and 3). The normalized fit obtained from the predicted peak parameters was multiplied by the corresponding scaling factor for comparison with the actual data. For the LS algorithm (Figure 4), the actual spectra were used directly without normalizing the counts. Comparison between the peak parameters obtained using AIM and LS is shown in Table 5 and the resulting peak fits are shown in Figure 8. In general, alpha peaks are known to have larger tails at the lower end of the peak [4]. This can be seen in the data for the expanded views for the three peaks in Figure 8 (b), (c), and (d), although the effect is not that pronounced. AIM gives superior fits compared to LS in two out of the three fits. The LS method excludes the tail regions to prevent such distortions from affecting the accuracy of determining the peak parameters derived, and the linear fit uses only the central FWHM width around the maximum count. However, peak parameters obtained this way and based on the assumption of an idealized gaussian shape would suffer from large deviations from experimental data in the long tail region which lead to larger value for $\chi_r^2$. The AIM model also is expected to perform poorly with such tailed peaks since all training was performed on idealized symmetrical gaussians. However, training AIM with modified gaussians that more resemble the actual peaks should improve performance. Therefore, improvement is more readily done with AIM as it considers the full peak shape while the LS limits itself to only the central peak region where the peak shape approximately resembles a true gaussian which is required for the linearization transform to be valid.

### 4.3. Comparison of Computational Speeds

CPU time for processing 10,000 peak fits on a VAX Station 3200 was computed for both AIM

and the LS algorithmic solutions shown in Figures 3 and 4, respectively. The peaks constituted different noisy versions of a single gaussian peak having fixed nominal values of 50 counts, 25 channels, and 19 channels for the peak height, position, and FWHM width, respectively. Total CPU time for the LS method amounted to 21.7 s while that for the AIM method was only 4 s. This indicates that AIM fitting is over 5 times faster than LS fitting. It should be noted, however, that no special effort was taken to optimize the LS code for faster execution. While execution time is fixed for the AIM approach regardless of the peak width, it is expected to vary almost linearly with the peak width for the LS linear fitting method, and therefore the LS CPU time requirement would be longer for wider peaks.

## 5. Discussion

We described two non-iterative automatic approaches for the fitting of single gaussian peaks and compared their performance. Both AIM and the linear LS methods have comparable performance for predicting the height of a pure gaussian. However, the LS method excels in predicting the peak position while AIM excels with the peak width. This shows that one of the two methods may be preferred for a given application, depending on the peak parameter of interest. While AIM errors are fairly uniform with nearly equal probability for positive and negative errors, LS tends to underestimate the peak position and overestimate the peak width. This biased behavior for the prediction error with the LS method has been shown to be statistically significant at the 95% confidence level. The effect may introduce distortions in the shape of spectra obtained by histogramming such predicted peak parameters as in the analysis of digitized waveforms of gaussian pulses.

The AIM machine learning approach utilizes the full peak while the LS method is based on a linearizing transform that is restricted to the central region around the peak center to avoid deviations from the linear relationship caused by departures from the true gaussian shape away from the center. Such departures are often encountered in practical spectral peaks,

e.g. the lower energy tailing in alpha spectra. This indicates that AIM is better suited than LS to handle such practical situations, since it can be trained on full records of peak shapes that realistically depict practical peaks, although additional peak parameters may be required in order to model such shape distortions. The LS method is based on the assumption of an idealized gaussian (linear relationship for the transformed data) and therefore attempts to exclude regions of departures from this assumption. The requirement for a weighted linear LS fit makes the LS method more computationally complex and time consuming (about 5 times slower than AIM). Moreover, since the LS method operates on the FWHM width of each individual peak, its execution time depends on the peak width. For AIM, execution time is constant since the AIM model always uses the full 50-channel data record regardless of the width of the peak being analyzed.

One disadvantage of AIM is that it does not readily provide information on uncertainties in the values estimated for the peak parameters. This is overcome through evaluating the model on a number of new test cases, for which the results are known, during the evaluation phase prior to actual use. Performance of the model on such cases provides a degree of confidence in the model and gives the user an idea of expected performance with totally unknown cases during actual use. When a sufficiently large number of cases is available for evaluation, reliable estimates can be determined for the confidence limits.

## Acknowledgements

This work is part of KFUPM/RI Center for Applied Physical Sciences project supported by the Research Institute of King Fahd University of Petroleum and Minerals, Dhahran, Saudi Arabia.

References

[1] Mariscotti MA. A method for automatic identification of peaks in the presence of background and its application to spectrum analysis. Nucl. Instrum. Methods 1967; 50:309-320.

[2] Phillips GW, Marlow KW. Automatic analysis of gamma-ray spectra from Germanium detectors. Nucl. Instrum. Methods 1976; 137:525-536.

[3] Machner H. Automatic analysis of γ-ray spectra including multiplets. Nucl. Instrum. Methods 1987; A258:246-249.

[4] Welsh RB, Gyger F, Jost DT, Von Gunten HR, Krahenbuhl U. NEWFIT, A computer program for the analysis of alpha, X-ray, and gamma-ray spectra. Nucl. Instrum. Methods 1988; A269:615-622.

[5] Madan VK. Processing of digital spectral data, In: S. K. kataria (ed.). Radiation Detection Systems- Lecture notes on pulse processing techniques in radiation detection and spectrometry. Oxford & IBH Publishing Co. PVT Ltd New Delhi,1993, pp. 354-378.

[6] Keller PE, Kangas LJ, Troyer GL, Hashem S, Kouzes RT. Nuclear spectral analysis via artificial neural networks for waste handling. IEEE Trans. Nucl. Sci 1995; 42:709-715.

[7] Vaughan RA (ed.). Pattern recognition and image processing in physics. SUSSP Publications, Edinburgh and Adam Hilger, Bristol, 1991.

[8] Weiss SM, Kulikowski CA. Computer systems that learn. Morgan Kaufmann Publishers, San Mateo, CA, USA, 1991

[9] Duda R, Hart P. Pattern recognition and scene analysis. Wiley, New York, 1973.

[10] Wasserman PD. Neural Computing: Theory and Practice. Van Nostrand Reinhold,

New York, 1989.

[11] Breiman L, Friedman J, Olshen R, Stone C. Classification and regression trees Wadsworth, Belmont, CA, USA, 1984.

[12] Montgomery GJ, Drake KC. Abductive networks, Proceedings of the SPIE Applications of Artificial Neural Networks Conf, Orlando, Florida, 1990, pp. 56-64.

[13] Abdel-Aal RE. Automatic fitting of Gaussian peaks using abductive machine learning IEEE Trans. Nucl. Sci. 1998; 45:1-16.

[14] Olmos P, Diaz JC, Perez JM, Garcia-Belmonte G, Gomez P, Rodellar V. Application of neural network techniques in gamma spectroscopy Nucl. Instrum. Methods 1992; A312:167-173.

[15] Olmos P, Diaz JC, Perez JM, Gomez P, Rodellar V, Aguayo P, Bru A, Garcia-Belmonte G, de Pablos JL. A new approach to automatic radiation spectrum analysis *IEEE Trans. Nucl. Sci.* 1991; 38:971-975.

[16] Abdel-Aal RE and Al-Haddad MN. Determination of radioisotopes in gamma-ray spectroscopy using abductive machine learning, Nucl. Instrum. Methods 1997; A391:275-288.

[17] Wythoff BJ, Levine SP, Tomelline SA. Spectral peak verification and recognition using a multilayered neural network. Anal. Chem. 1990; 62:2702-2709.

[18] Akkila T, Lindblad T, Lund-Jensen B, Szekely G and Eide A. A hardware implementation of an analog neural network for Gaussian peak-fitting Nucl. Instrum. Methods 1993; A327: 573-579.

[19] Raad MW, Abdel-Aal RE, Noras JM and Elguibaly F. Novel digital pulse pile-up

identification and parameter estimation algorithms in nuclear spectrometry, *3rd International Workshop on Advances in Signal Processing for Non Destructive Evaluation of Materials*, 5-8 August 1997, Quebec City, Quebec, Canada.

[20] Knoll G F. Radiation Detection and Measurement. John Wiley New York, 1989.

[21] Mukoyama T. Fitting of gaussian to peaks by non-iterative method. *Nucl. Instrum. Methods* 1975; 125:289-291.

[22] McCullagh P. and Nelder, JA. Generalized Linear Models. London, Chapman and Hall, 1983.

[23] AIM User's Manual. AbTech Corporation, Charlottesville, VA, USA,1990.

[24] Barron AR. Predicted squared error: A criterion for automatic model selection, In: S. J. Farlow (ed.). Self-Organizing Methods in Modeling: GMDH Type Algorithms. Marcel-Dekker, New York, 1984, pp. 87-103.

[25] Bevington PR. Data Reduction and Error Analysis for the Physical Sciences, New York, McGraw-Hill, 1969.

Figure Captions :

Figure 1. A typical AIM network structure showing various types of functional elements.

Figure 2. AIM abductive network model for fitting a single gaussian peak. Training on noisy data with N = 100, CPM =1. A is the normalized peak height.

Figure 3. Equations for the AIM abductive network model of Figure 2 for a single gaussian peak. All variables are indicated on Figure 2. A is the normalized peak height.

Figure 4. The linear LS algorithm for fitting a single gaussian peak. P is the actual peak height.

Figure 5. Cumulative percentage plots for the percentage errors in the three peak parameters for the 500-case evaluation sample used for the results in Table 1. (a) Height, (b) Position, and (c) Width.

Figure 6. Spectra for the 3 peak parameters determined by both AIM and LS for 1 million events of a single gaussian peak at the high noise level. Nominal peak parameters are: Height = 50 counts, Position : channel 25, FWHM Width = 19 channels. (a) Height, (b) Position, and (c) Width.

Figure 7. Comparison between AIM and LS fits for 3 simulation single gaussian peaks. (a) case # 5, (b) case # 7, and (c) Case # 11 in Table 4.

Figure 8. Comparison with actual spectral data. (a) Spectrum of alpha particles scattered from a ZnS target sandwiched between two Au layers on a carbon backing. (b), (c), and (d) AIM and LS fits for the three marked alpha peaks. Data on the fits are given in Table 5.

Table Captions :

Table 1. Error statistics for both AIM and LS methods for fitting single gaussian peaks. The AIM model was developed by training on 1000 cases at the high noise level. Evaluation sample: 500 cases at both the low and high noise levels. Shown are percentages of the evaluation sample having various ranges of percentage error, and data on the overall mean percentage absolute error (MPAE).

Table 2. Data on the percentage frequency and average values for positive and negative prediction errors using both AIM and LS. The data suggests that LS estimates of the peak position and width are more biased than the corresponding AIM estimates.

Table 3. Data on the overall mean error and standard deviation for the 500-case evaluation sample of Table 2 for both the AIM and LS. Shown also are values for the z-statistic used to test the null hypothesis that error is unbiased. n= 500.

Table 4. Comparison between the performance of AIM and LS on 12 simulated single gaussian peaks drawn from the 500-case evaluation sample used for the results in Table 1.

Table 5. Comparison between the performance of AIM and LS on the three single alpha peaks in Figure 8.