Qualitative and quantitative validation of the SINBAD code on complex HPGe gamma-ray spectra

Emmanuel Rohée, Romain Coulon, Frédérick Carrel, Thomas Dautremer, Eric Barat, Thierry Montagu, Stéphane Normand, C. Jammes

Qualitative and quantitative validation of the SINBAD code on complex HPGe gamma-ray spectra

E. Rohée, R. Coulon, F. Carrel, T. Dautremer, E. Barat, T. Montagu, S. Normand and C. Jammes

Abstract–**Radionuclide identification and quantification is a serious concern for many applications as safety or security of nuclear power plant or fuel cycle facility, CBRN risk identification, environmental radioprotection and waste measurements. High resolution gamma-ray spectrometry based on HPGe detectors is a performing solution for all these topics. During last decades, a great number of software has been developed to improve gamma spectra analysis. However, some difficulties remain in the analysis when full energy peaks are folded together with a high ratio between their amplitudes, when the Compton background is much larger compared to the signal of a single peak and when spectra are composed of a great number of peaks. This study deals with the comparison between a conventional analysis method and an innovative approach, called SINBAD ("*Spectrométrie par Inférence Non paramétrique BAyesienne Déconvolutive*"), for radionuclide identification and quantification. For many years, SINBAD has been developed by the CEA LIST for unfolding complex spectra from HPGe detectors. Contrary to the conventional method using fitting procedures, SINBAD uses a probabilistic approach with nonparametric Bayesian inference to process spectrum data. The conventional fitting method founded for instance in Genie 2000 is compared with the nonparametric SINBAD approach regarding some key figures of merit as the peak centroid estimation (identification step) and net peak area determination (quantification step). Complex cases are studied for nuclide detection with closed gamma-rays energies and high full energy peak intensity differences. Tests are performed with spectra from the International Atomic Energy Agency (IAEA) for gamma spectra analysis software benchmark and with spectra acquired in our laboratory. It appears that SINBAD results are better than GENIE 2000 ones in most of the cases even if hard deconvolutions can be achieved thanks to GENIE 2000 at the cost of expert parameters fine tuning which has to be compared with the user-friendly SINBAD operating.**

I. Introduction

GAMMA-ray spectrometry with High Purity Germanium (HPGe) detector is today a recognized solution for radionuclide identification and quantification thanks to its excellent energy resolution (~ 0.2 % at 661.7 keV) [1]. For decades, this instrumentation is used in many fields like in nuclear power plant, experimental reactor, for CBRN risk management, environmental radioprotection, waste measurement or fundamental research [2]. To handle data, spectra analysis software have been developed to identify and quantify isotopes. If these software give satisfactory results in most of cases, convolutions of closed full energy peaks with high net area ratio are globally difficult to process and software often give inaccurate energy and net peak area evaluation without expert user settings [3-4] . A new approach was developed by the CEA LIST for several years through the code called SINBAD (Spectrométrie par Inférence Non paramétrique BAyesienne Déconvolutive) to overcome these difficulties in spectra deconvolution [5]. In this study, SINBAD performances are compared with those obtained using a conventional deconvolution method.

First, the conventional deconvolution method found in reference software like GENIE 2000 [6] and the nonparametric Bayesian estimation method found in SINBAD are exposed. Then, a benchmark between the two codes in difficult convolution cases from IAEA and laboratory spectra is proposed and finally, results are discussed.

E. Rohée, R. Coulon, S. Normand, F. Carrel are with the CEA, LIST, Laboratoire Capteurs et Architectures Électroniques, F-91191 Gif-sur-Yvette, France (e-mail : emmanuel.rohee@cea.fr, romain.coulon@cea.fr, stephane.normand@cea.fr, frederick.carrel@cea.fr).

T. Dautremer, E. Barat, T. Montagu are with the CEA, LIST, Laboratoire Modélisation, Simulation et Systèmes, F-91191 Gif-sur-Yvette, France (e-mail : thomas.dautremer@cea.fr, eric.barat@cea.fr, thierry.montagu@cea.fr).

C. Jammes is with CEA/DEN/SPEx/LDCI, Centre de Cadarache, F-13109 Saint-Paul-lez-Durance, France (e-mail : christian.jammes@cea.fr).

II. Methods

Three main methods for multiplet deconvolution stand out from commercial gamma spectra analysis software. The total-area determination method show very good performances estimating the total area of multiplet but is not suitable when a high deconvolution power is required [7]. The library-oriented method performs well in terms of net peak area and uncertainty determination but is limited to specific radionuclides searches and measurements [7]. The iterative fitting method is more close to the SINBAD philosophy, as it allows to search and measure peaks in a whole spectrum with a significant deconvolution power. Therefore, this is the method used in this comparison study, using the largely commercial software common in laboratories and industries GENIE 2000.

A. Iterative fitting deconvolution method

Peaks are searched with the most commonly used Mariscotti method, also called the second derivative, where the second difference function corresponding to the spectrum data reaches a minimum at centroids [8]. In case of peaks too close in energy, this method can fail, especially if a small peak is drown in a larger one.

To adjust the peak centroid estimation, the following algorithm is used over the range defined by the negative second derivative function.

$$Centroid = \frac{\sum_i c_i i}{\sum_i c_i} \quad (1)$$

Where C_i is the weight of the channel i defined with the number of counts in the channel and coefficients evaluated with the peak width [6]. Then, a significance level is calculated for each peak found and peaks are retained only for value above a sensitivity user-defined threshold.

Peak area measurement algorithms use results from the peak centroid estimation. A least square nonlinear fitting of the spectrum data is performed on the Region Of Interest (ROI), function of the FWHM and centered on the peak position found previously. The residuals from this fitting procedure are calculated. If the latter are higher than the user-defined residual threshold, a peak is added at the highest residual value if the distance to the neighboring peak is above the minimum peak separation distance defined by the user. Then, a new ROI is computed according to the previous ROI. A new fit is performed, new residuals are computed, and a peak is added if necessary. This operation is repeated until there is no more peak to add. Results with this method are strongly dependent on the different threshold values.

Once the best fit for the multiplet is achieved, the net peak area of each individual peak in the multiplet is computed using a numerical integration.

B. SINBAD nonparametric Bayesian estimation method

SINBAD is a gamma spectra deconvolution code developed by CEA LIST. Conventional deconvolution methods involve nonlinear fittings, which require preset parameters as described in the previous section. The SINBAD improvement comes from its Bayesian probabilistic approach, based on the prior that the background is smooth and the peaks are Gaussian.

SINBAD principle is to adjust the gamma spectrum using the model *S*, seen as a probability density.

$$S = aP + (1 - a)B \quad (2)$$

Where a, P and B are random variables. P is the peak model, B the background model and a a scalar between 0 and 1.

The prior chosen for B is a *P*ólya tree mixture [9]. Its smoothness can be controlled to achieve the best background adjustment. Because of the prior that the background is smoother than the peaks, the latest are not adjusted by B. The peaks are described with a Dirichlet processes mixture [10] P which can be seen as a potentially infinite sum of peaks convolved with a Gaussian kernel. The width of the Gaussian kernel is a function of the calibration step.

The algorithm used to adjust the model S to the spectrum data is a Markov Chain Monte Carlo (MCMC) algorithm. Pólya trees and Dirichlet processes are randomly generated. Iteration after iteration, the probability laws of the sampled values converge toward the spectrum data thanks to the knowledge of the previous states. When the algorithm has converged, an estimator of the distribution is determined, so the mean and variance of the peak position and its area are available.

C. Intercomparison

The goal of this study is to compare SINBAD gamma spectra deconvolution performances to those obtained with the conventional method found in the software GENIE 2000 in complex convolution cases. We focus on the capacity of the software to find the correct value for peak position and net peak area regarding the distance between two peaks and their net peak area ratio. The three different tests performed are presented.

1. 1995 IAEA test spectra

The 1995 IAEA test spectra were created by the IAEA as a tool for benchmarks of gamma analysis software. These spectra are composed with ^{226}Ra spectrum with progeny [11]. The spectra are calibrated and peak position and their areas are available. Here, we use spectra of different counting times (2000 s, 667 s, 200 s, 20 s) which are three channels (1.2 keV) shifted and summed with a 2000 s spectrum, and respectively named “ADD1N1”, “ADD3N1”, “ADD10N1” and “ADD100N1”. In this way, peaks overlapping are created with net peak area ratios of 1, 3, 10 and 100.

2. ^{137}Cs test spectra

To access finer net peak area ratio and energy gap between the peaks we want to process, the 1995 IAEA test spectra principle is applied to a ^{137}Cs spectra, acquired with a GL2015 n-type Canberra HPGe detector and a Canberra DSA-1000 analyzer. Spectra of different counting times are shifted and summed to the spectrum measured during 1000 s. The net peak area of the ^{137}Cs peak (661.66 keV) in the latest spectrum is equal to 364484 counts. The gap between peaks goes from 240 eV to 1.2 keV for a net peak area ratio of 1. Net peak area ratios for a gap of 720 eV are 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90 and 100.

3. ^{57}Co – ^{152}Eu two sources test

An intercomparison with real sources is proposed. ^{57}Co and ^{152}Eu have closed gamma-ray energies, respectively 122.0 keV and 121.78 keV. The setup is the same as described in the previous section. First, a spectrum with the source of ^{57}Co solely is acquired, and then a source of ^{152}Eu is added at different distances from the detector. Then, ^{57}Co net full energy peak area with and without ^{152}Eu are compared.

III. Results

We focus first on the difference between the peak position given by GENIE 2000 or SINBAD and the peak position reference. Then, the net peak area gap from the reference $A_{GENIE\ 2000}/A_{ref}$ and A_{SINBAD}/A_{ref} are computed.

1. 1995 IAEA test spectra

In this study, the peak at 609.26 keV (^{214}Bi) is chosen. E_1 is the estimated energy of the peak at 609.26 keV with a net peak area of 73073±438 counts and E^2 is the estimated energy of the same peak shifted by one FWHM (1.2 keV) at 610.46 keV with variable net peak area. No standard deviation on the peak position estimation is available with GENIE 2000 using the residual method, although it is vital in such cases of strong convolution. Consequently, the standard deviation is not taken into account in this peak position estimation study, and we focus only on the mean values.

TABLE I. UNSHIFTED PEAK POSITION BIAS (609.26 KEV)

Net peak area ratio	$E_{1GENIE\ 2000}$-E_{1ref} (keV)	$E_{1SINBAD}$-E_{1ref} (keV)
1	0.16	0.25
3	0.17	0.26
10	0.17	0.26
100	-	0.25

TABLE II. SHIFTED PEAK POSITION BIAS (610.46 KEV)

Net peak area ratio	$E_{2GENIE\ 2000}$-E_{2ref} (keV)	$E_{2SINBAD}$-E_{2ref} (keV)
1	0.16	0.23
3	0.16	0.23
10	0.17	0.27
100	-	0.25

Table I and II show that in this case GENIE 2000 and SINBAD provide mean biases respectively equal to 170 eV and 250 eV. This bias which is larger than the bias observed in the next tests can be explained by the low binning of these spectra with only 0.4 keV/channel, that being a FWHM described by only 3 channels for the energy of interest. Even if SINBAD gives a higher bias of the peak position than GENIE 2000 with the 1995 IAEA test spectra, the second peak is detected for a net peak area ratio of 100, which is not the case with GENIE 2000.

Roughly speaking, Fig. 1 and Fig. 2 show that SINBAD appears more efficient for the determination of the net area of the unshifted peak constant in net area. However there is a drift in estimating the net area of the shifted peak when the net peak area ratio increases. When the net peak area ratio is up to 100, GENIE 2000 is not able to detect the presence of the peak whereas SINBAD detects it.

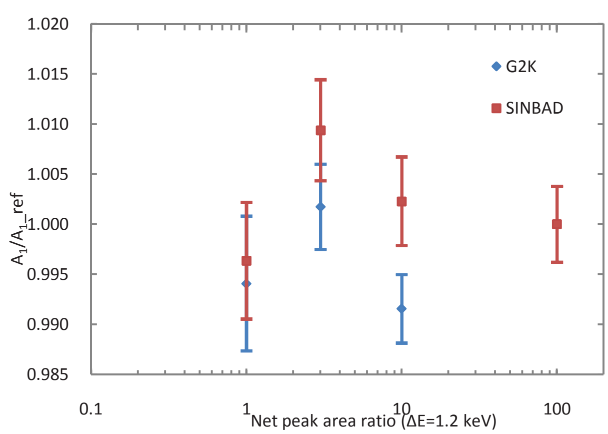

Fig. 1. Net peak area relative gap from the reference of the 609.26 keV peak

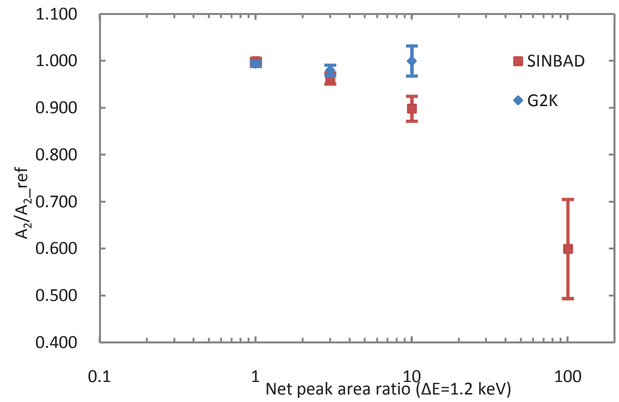

Fig. 2. Net peak area relative gap from the reference of the 610.46 keV peak

2. ^{137}Cs test spectra

The energy bias in case of a gap between peaks from 0.2 to one FWHM (1.20 keV) is shown in Table III for the initial peak and in Table IV for the shifted peak. Fig. 3 and 4 show the net peak area bias. All the results obtained with GENIE 2000 are computed using optimized parameters. Figure 5 is the same graph as Fig. 4 with GENIE 2000 default parameters.

TABLE III. UNSHIFTED PEAK POSITION BIAS (661.66 KEV)

Peaks separation (FWHM)	$E_{1GENIE\ 2000}$-E_{1ref} (keV)	$E_{1SINBAD}$-E_{1ref} (keV)
0.20	-	-0.18
0.24	-	0.01
0.28	-	0.04
0.32	-	-0.04
0.36	-	-0.03
0.40	-	-0.02
0.44	-	-0.02
0.48	0.17	-0.02
0.52	-0.01	-0.01
0.56	-0.01	-0.01
0.60	-0.01	-0.01
0.64	-0.01	-0.01
0.68	-0.01	-0.01
0.72	-0.01	-0.01
0.76	-0.01	-0.01
0.80	-0.01	-0.01
0.84	-0.01	-0.01
0.88	-0.01	0.00
0.92	-0.01	0.00
0.96	-0.01	0.00
1.00	-0.01	0.00

TABLE IV. SHIFTED PEAK POSITION BIAS (661.66 KEV + ΔE)

Peaks separation (FWHM)	$E_{2GENIE\ 2000}$-E_{2ref} (keV)	$E_{2SINBAD}$-E_{2ref} (keV)
0.20	-	-0.08
0.24	-	0.03
0.28	-	0.05
0.32	-	-0.04
0.36	-	-0.04
0.40	-	-0.03
0.44	-	-0.03

0.48	0.57	-0.03
0.52	-0.02	-0.02
0.56	-0.02	-0.02
0.60	-0.02	-0.01
0.64	-0.02	-0.01
0.68	-0.02	-0.01
0.72	-0.01	-0.01
0.76	-0.01	-0.01
0.80	-0.01	-0.01
0.84	-0.01	-0.01
0.88	-0.01	0.00
0.92	-0.01	0.00
0.96	-0.01	0.00
1.00	-0.01	0.00

The peak position location is efficient when the multiplet is found. GENIE 2000 is not able to deconvolute it in all cases. With the optimized parameters and under a gap between peaks of 0.48 FWHM, GENIE 2000 sees the multiplet as a singlet. When the default value of the residual method threshold is used, this limit is 0.60 FWHM. With the default parameters, i.e. only the second derivative is used and the residual method is not activated, the limit is 1.04 FWHM.

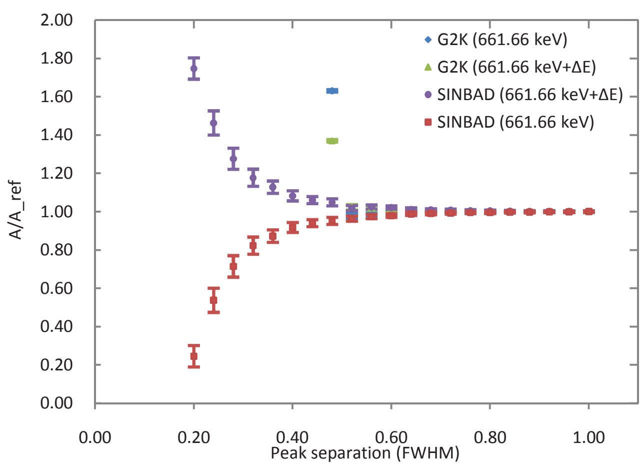

Fig. 3. Net peak area relative gap from the reference on 0 – 1 FWHM

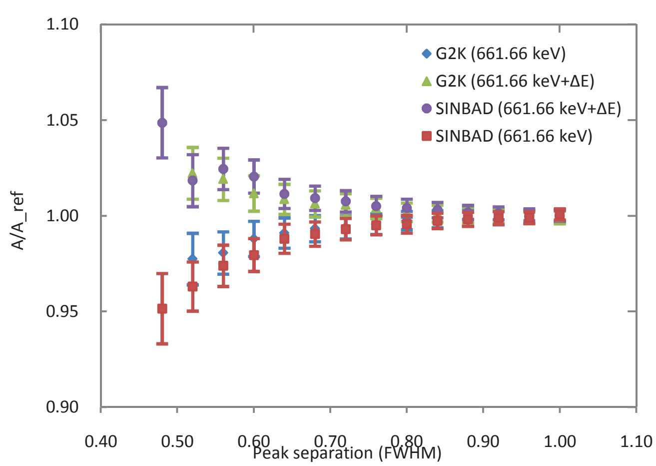

Fig. 4. Net peak area relative gap from the reference on 0.4 – 1 FWHM

SINBAD finds the presence of two peaks with a low peak separation (0.2 FWHM) but the net peak area is highly biased. It converges to the reference value when increasing the distance between the two peaks. The error is less than 5 % from a separation of 0.48 FWHM and uncertainties become compatible with the reference from 0.76 FWHM. As mentioned earlier, the deconvolution with GENIE 2000 only happens above a separation of 0.48 FWHM with the optimized parameters and 0.60 FWHM with the default parameters. The error on the net peak area is less than 5% for a separation of 0.52 FWHM in case of optimized parameters, and this value becomes 0.60 FWHM when using the default value of the residual method threshold. Uncertainties become compatible with the reference for a separation above 0.68 FWHM.

Table V and VI refer to energy bias for a 720 eV gap between peaks with net area ratio from 1 to 100.

TABLE V. UNSHIFTED PEAK POSITION BIAS (661.66 KEV)

Net peak area ratio	$E_{1GENIE\ 2000}$-E_{1ref} (keV)	$E_{1SINBAD}$-E_{1ref} (keV)
1	0.00	0.00
2	-0.01	0.00
3	-0.01	0.00
4	-0.01	0.00
5	-0.01	0.00
6	-0.01	0.00
7	-0.01	0.00
8	-0.01	0.00
9	-0.01	0.00
10	-0.01	0.00
20	-0.01	0.00
30	-0.01	0.00
40	-0.01	-0.01
50	-0.01	0.00
59	-0.01	0.00
71	-0.01	0.00
77	-0.01	0.00
91	-0.01	0.00
100	-0.01	0.00

TABLE VI. SHIFTED PEAK POSITION BIAS (662.38 KEV)

Net peak area ratio	$E_{2GENIE\ 2000}$-E_{2ref} (keV)	$E_{2SINBAD}$-E_{2ref} (keV)
1	0.00	0.01
2	0.00	0.01
3	0.00	0.01
4	0.01	0.02
5	0.01	0.02
6	-0.02	0.00
7	0.00	0.01
8	-0.01	0.01
9	-0.03	-0.01
10	-0.01	0.01
20	-0.04	-0.01
30	-0.08	-0.04
40	-0.13	-0.07
50	-0.12	-0.06
59	-0.20	-0.14
71	-0.18	-0.13
77	-0.19	-0.15
91	-0.25	-0.22

100	-0.45	-0.44

Table VI shows that for the shifted peak and for net area ratios from 10 to 100 the bias of the centroid location is a bit lower for SINBAD compared to GENIE 2000.

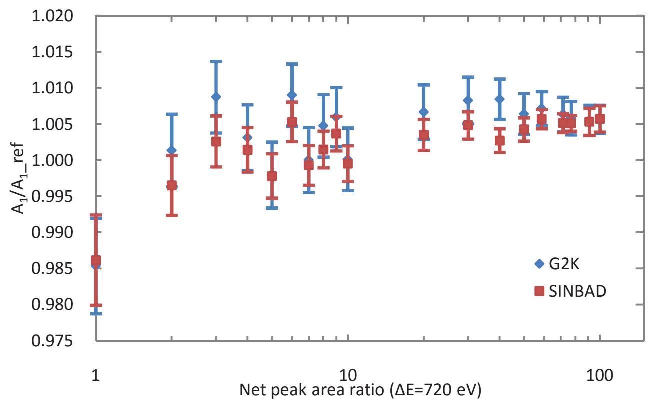

Fig. 5. Net peak area relative gap from the reference of the 661.66 keV peak for a gap between peaks of 720 eV

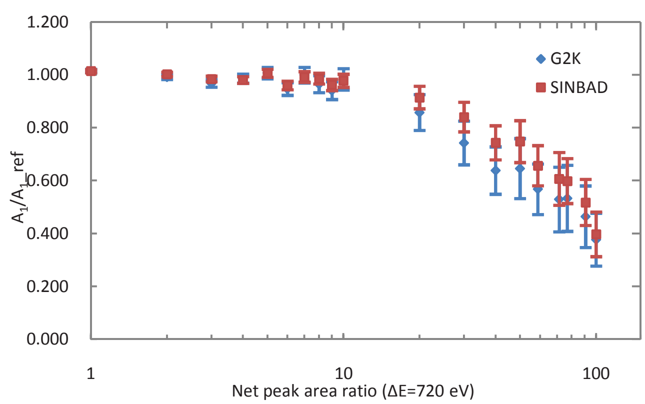

Fig. 6. Net peak area relative gap from the reference of the 662.38 keV peak for a gap between peaks of 720 eV

When the gap between the peaks is fixed at 720 eV and net area ratios vary from 1 to 100, the results for the unshifted peak are consistent with the reference for net area ratios until 10, although the value for the ratio of 1 appears further from the reference. For ratios beyond 10, the net areas obtained by SINBAD are generally closer to the reference. Regarding the shifted peak at 662.38 keV, the drift trend observed with the 1995 IAEA test spectra is confirmed for both codes, but the drift is smaller with SINBAD.

3. ^{57}Co – ^{152}Eu two sources test

Table VII and Fig. 7 show that the centroid position evaluation is more effective with SINBAD than with GENIE 2000 when the two source test is performed since GENIE 2000 is not able to deconvolute the peaks for a net area ratio of 1.4. Moreover, when the deconvolution is achieved, SINBAD shows better performances than GENIE 2000. The ^{57}Co peak is detected in the three cases with net area values closer to the reference than those obtained using GENIE 2000 with optimized parameters.

TABLE VII. ^{57}CO PEAK POSITION BIAS (122.06 KEV)

$A_{Eu152\_ref}/A_{Co57\_ref}$	$E_{GENIE\ 2000}-E_{ref}$ (keV)	$E_{SINBAD}-E_{ref}$ (keV)
4	0.00	-0.02
1.4	-	-0.04
0.7	-0.12	-0.08

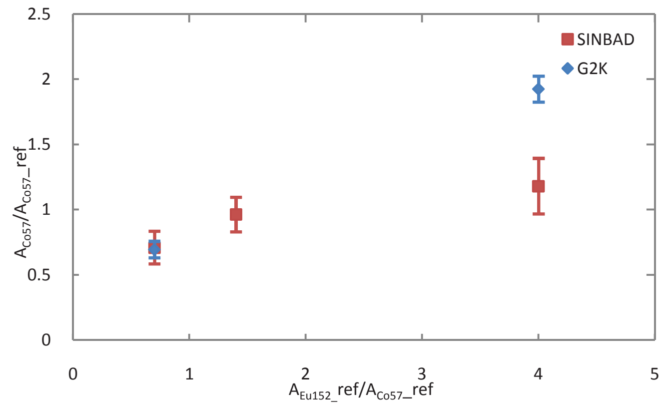

Fig. 7. Net peak area relative gap from the reference of the ^{57}Co peak (122.06 keV)

IV. CONCLUSION

This study shows that in many cases, especially when the distance between the two peaks is large enough, SINBAD and GENIE 2000 offer close performances in terms of estimating the centroid and determining the net area. However, for highly convoluted peaks (gap less than 0.5 FWHM) and despite a biased net area quantification, SINBAD still manages the peak location, which is not the case of GENIE 2000 even with optimized parameters. The nonparametric aspect of SINBAD has to be emphasized, as it makes it a user-friendly program. Indeed, GENIE 2000 needs an expert setting to approach SINBAD performances with a fine adjustment of several thresholds. This fine tuning often implies a trade-off between deconvolution performances and false peak detection in the spectrum.

REFERENCES

[1] G. F. Knoll, *Radiation Detection and Measurement 4th edition*. John Wiley & Sons, 2010.

[2] G. R. Gilmore, *Practical Gamma-ray Spectrometry*. John Wiley & Sons, 2008.

[3] M. Blaauw, "Multiplet deconvolution as a cause of unstable results in gamma-ray spectrometry for INAA," *Nucl. Instruments Methods Phys. Res. Sect. A*, vol. 333, pp. 548–552, 1993.

[4] IAEA, "Intercomparison of gamma ray analysis software packages, IAEA-TECDOC-1011," Vienna, 1998.

[5] E. Barat, T. Dautremer, and T. Montagu, "Nonparametric Bayesian Inference in Nuclear Spectrometry," in *IEEE Nuclear Science Symposium Conference Record*, 2007, pp. 880–887.

[6] Canberra, "Genie TM 2000 Spectroscopy Software - Customization Tools." 2003.

[7] M. Blaauw, R. M. Keyser, and B. Fazekas, "Comparison of alternative methods for multiplet deconvolution in the analysis of gamma-ray spectra," *Nucl. Instruments Methods Phys. Res. Sect. A*, vol. 432, pp. 77–89, 1999.

[8] M. . Mariscotti, "A method for automatic identification of peaks in the presence of background and its application to spectrum analysis," *Nucl. Instruments Methods*, pp. 309–320, 1967.

[9] T. S. Ferguson, “Prior distributions on spaces of probability measures,” *Ann. Stat.*, vol. 2, pp. 615–629, 1974.

[10] T. S. Ferguson, “A bayesian analysis of some nonparametric problems,” *Ann. Stat.*, vol. 1, pp. 209–230, 1973.

[11] M. Blaauw, V. Osorio Fernandez, and W. Westmeier, “IAEA y-ray spectra for testing of spectrum analysis software,” *Nucl. Instruments Methods Phys. Res. Sect. A*, vol. 387, pp. 410–415, 1997.