SGSD: A novel Sequential Gamma-ray Spectrum Deconvolution algorithm

Hadi Shahabinejad, Naser Vosoughi *

Department of Energy Engineering, Sharif University of Technology, Tehran, Iran

ARTICLE INFO

ABSTRACT

A novel approach for analyzing complex gamma-ray spectra using a sequential algorithm is introduced. The developed Sequential Gamma-ray Spectrum Deconvolution (SGSD) algorithm produces a sequence of spectra converging to the best estimation of output spectrum of a gamma-ray detector. In each point of sequence, an isotope of unknown gamma-ray source is identified and the respective response of the detector to unknown source is reconstructed. Effectiveness of the developed algorithm is demonstrated by two empirical and simulation studies. In the case of empirical study, a number of recorded gamma-ray spectra related to a mixed gamma-ray source including different combinations of 5 isotopes (Co-60, Cs-137, Na-22, Eu-152 and Am-241) are analyzed using whole information of spectra. Furthermore, a number of simulated gamma-ray spectra related to a mixed gamma-ray source including different combinations of 30 isotopes are analyzed in simulation study. Both man-made and natural radioisotopes like Ba-133, Co-60, Ir-192, Cs-137, K-40, Th-232 series, U-238 series, Ac-227 series, etc. are used for Monte Carlo simulations. The numerical results of the SGSD algorithm are compared with those of the conventional Non-Negative Least Squares (NNLS) algorithm. Based on the results, the identification procedure of the SGSD algorithm has a remarkable superiority over the NNLS algorithm.

1. Introduction

Identification and quantification of isotopes through the analysis of the gamma-ray spectra recorded in a Multi-Channel Analyzer (MCA) have several applications such as: measuring natural radioactivity (Caciolli et al., 2012; Baldoncini et al., 2018; Morsy et al., 2012; Hendrix et al., 2001), elemental analysis using Prompt Gamma Neutron Activation Analysis (PGNAA) (Meric et al., 2011; Oliveira et al., 1993), measurements in the marine environment for radiological purposes (Androulakaki et al., 2016), etc. The gamma-ray spectra usually consist of several photopeaks rising over a background, Compton continua and other parts related to different physical processes (Knoll, 2000; Yang and Garcia, 2018). Analysis of gamma-ray spectra is usually performed by monitoring the most intense photopeaks of a background-subtracted spectra. In the case of high resolution detectors like Ge spectrometers where the photopeaks are well resolved, the identification and cal-culation of weight fractions of radioisotopes present in the source under study are relatively simple tasks (Olmos et al., 1992). Nevertheless, since the analysis is performed in limited energy windows, only limited spectral information is exploited. Furthermore, the algorithms that work on the basis of photopeak analysis have difficulties in the identification of radioisotopes in low counting rates as well as in the situations where the overlapping of peaks occurs as a consequence of lack of resolution (Caciolli et al., 2012; Alamaniotis and Jevremovic, 2015; Gardner et al., 1997). Significant improvements in gamma-ray spectrum analysis results can be obtained by implementing the Whole Spectrum Analysis (WSA) or Full Spectrum Analysis (FSA) method (Salmon, 1961; Gardner et al., 1997; Hendrix et al., 2001). The WSA method needs much less statistical data or acquisition time than photopeak analysis method to reach the necessary accuracy, since it uses the whole spectral information (Caciolli et al., 2012). In WSA approach, the unknown quantity of jth radioisotope (x_j) to be retrieved by fitting the measured spectrum (detector response) as a linear combination of library spectra, the response matrix that contains the library spectra and couples the count rate in i^{th} channel with the unique concertation of jth component, R_{ij}, and the recorded count rate in the ith energy bin, N_i, are related through equation (1) (Salmon, 1961; Shahabinejad and Vosoughi, 2018).

$$N_i + \varepsilon_i = \sum_j R_{ij} x_j + B_i \tag{1}$$

where, B_i is the recorded counts in the in the ith energy bin due to the background radiation and ε_i is the statistical uncertainty related to the counts recorded in i^{th} channel. Eq. (1) can be written in matrix form as follows:

$$\boldsymbol{N} + \boldsymbol{\varepsilon} = \mathbf{R}\boldsymbol{x} + \mathbf{B} \tag{2}$$

where, $\boldsymbol{N} = (N_1, N_2, \ldots, N_n)^T$, $\boldsymbol{\varepsilon} = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)^T$, $\boldsymbol{B} = (B_1, B_2, \ldots, B_n)^T$, $\boldsymbol{x} = (x_1, x_2, \ldots, x_m)^T$ and $\mathbf{R}$ is the $n \times m$ response matrix. One encounters the so-called over-determined problem (Cichocki and Amari, 2002) in gamma-ray spectroscopy analysis, since the number of channels is often higher than the number of pertinent libraries ($n > m$). Therefore, there is usually no exact solution to this problem.

The WSA approach was first applied for analyzing the gamma-ray spectra obtained using scintillation detectors in 1961 (Salmon, 1961). Salmon used classical Multiple Linear Regression (MLR) method to solve the obtained over-determined problem from WSA approach (Salmon, 1961). Although he achieved appropriate results in his work, the improper results of the classical MLR method like generating negative values for abundance fractions were later reported (Meric et al., 2011; Shahabinejad and Vosoughi, 2018). This difficulty in estimating the correct amounts or abundance fractions of radioisotopes occurs due to either the inverse problem is numerically ill-conditioned (Meric et al., 2012) or the correct library spectra are not given to the method (Shahabinejad and Vosoughi, 2018). Several methods have been proposed to deal with this difficulty like: minimizing the condition number of the covariance matrix (Meric et al., 2011; Meric et al., 2012), applying the Non-Negative Least Square (NNLS) method for FSA of gamma-ray spectra (Baldoncini, 2010; Caciolli et al., 2012; Baldoncini et al., 2018), performing the FSA in several different energy windows (Androulakaki et al., 2016) and controlling the fitting process on gamma-ray spectra by a "comparison module" (Alamaniotis and Jevremovic, 2015; Shahabinejad and Vosoughi, 2018). Artificial Intelligence based techniques like Genetic Algorithms (GAs) (Carlevaro et al., 2008; Alamaniotis et al., 2013), hybrid fuzzy-genetic (Alamaniotis and Jevremovic, 2015), Fireworks Algorithm (Alamaniotis et al., 2015; Alamaniotis and Tsoukalas, 2018) and Particle Swarm Optimization (PSO) (Shahabinejad and Vosoughi, 2018) have also been used to overcome the shortcomings of the classical MLR in gamma-ray spectrum analysis. These methods (in contrast to the traditional linear optimization methods) are able to find the global optima in a large multi-dimensional solution space without suffering the pitfalls of local optima (Mukherjee, 2002; Shahabinejad and Sohrabpour, 2017; Shahabinejad et al., 2016). Furthermore, neural networks have also been applied to analyze gamma-ray spectra as presented in (Yoshida et al., 2003; Kangas et al., 2008). It is required to identify the correct library spectra before applying a fitting process to a gamma-ray spectrum, as the number of pertinent libraries is high in many application and all of them do not contribute in the output spectra. Although Alamaniotis and Jevremovic (2015) have successfully applied a fuzzy system for first identifying the pertinent libraries using photopeaks of a measured spectrum and then fitting on the gamma-ray spectra by genetic algorithm using the identified isotopes, their fuzzy identification method may not produce appropriate results when the number of pertinent libraries increases. As they have reported, there is a possibility for false detection of radioisotopes using developed fuzzy system in low statistics and this possibility increases with increasing the number of pertinent libraries.

In this paper, a novel Sequential Gamma-ray Spectrum Deconvolution (SGSD) algorithm for analyzing gamma-ray spectra is introduced. The SGSD algorithm produces a sequence of spectra converging to the best estimation of spectrum of a gamma-ray detector. In other words, in each step of the algorithm, an isotope of unknown gamma-ray source is identified and the respective response of the detector to unknown source is reconstructed. The SGSD algorithm exploits the advantages of the active set method (Gill et al., 1981) and evolutionary optimization techniques (Shahabinejad and Vosoughi, 2018) to analyze a given gamma-ray spectrum. This approach is used to identify different isotopes of a mixed gamma-ray source and determine their corresponding fractional abundances. Finally, the obtained results of the proposed method are benchmarked against the conventional NNLS algorithm (Lawson and Hanson, 1995; Caciolli et al., 2012; Baldoncini et al., 2018).

2. Non-negative least square algorithm

As already mentioned, one encounters an over-determined problem with no exact solution in gamma-ray spectrum analysis using WSA approach. To solve such a problem, the least square solution is the one that minimizes the sum of weighted squares of residual errors (ε_i) for all measured detector counts (N_i). According to Eq. (1), the NNLS problem can be written as:

$$\hat{x}_j = \underset{x_j \geqslant 0}{\operatorname{argmin}} \ \sum_{i=1}^{n} \left(w_i \left(N_i - \sum_{j=1}^{m} R_{ij} x_j - B_i \right)^2 \right) \tag{3}$$

where, w_i is the weight of residual error related to the ith energy bin. Considering the Poisson distribution for detector counts in each energy bin, the weights are $1/N_i$ in gamma-ray spectroscopy (Phillips, 1978). Since the recorded counts are often zero in some energy bins of measured spectra, the weights can be considered 1 in many applications. Eq. (3) with weights 1 can be written in matrix form as presented in Eq. (4).

$$\hat{\boldsymbol{x}} = \underset{\boldsymbol{x} \geqslant 0}{\operatorname{argmin}} \ ||\boldsymbol{N} - \mathbf{R}\boldsymbol{x} - \boldsymbol{B}||_2^2 \tag{4}$$

Here, the $\| \cdot \|_2$ is the norm 2 of a vector. The problem (4) is a constrained non-linear minimization problem. To solve this type of problem, the NNLS algorithm that is based on the so-called Karush-Kuhn-Tucker (KKT) conditions and active set method (Gill et al., 1981) was first implemented by Lawson and Hanson (Lawson and Hanson, 1995). The KKT conditions are necessary for optimality of a solution to the constrained non-linear minimization problem and can be expressed by Eq. (5) for the problem (4).

$$\begin{aligned} \left(\mathbf{R}^T (\mathbf{R}\boldsymbol{x} + \mathbf{B} - \mathbf{N}) \right)^T \boldsymbol{x} &= 0 \\ \mathbf{R}^T (\mathbf{R}\boldsymbol{x} + \mathbf{B} - \mathbf{N}) &\geqslant 0 \\ \boldsymbol{x} &\geqslant 0 \end{aligned} \tag{5}$$

The active set method uses the fact that only a small subset of constraints are usually satisfied or active at the solution. The jth constraint is said to be active, if the jth regression coefficient is negative, otherwise the constraint is called passive. In this method the active set initially contains all the unknown variables. In each step of the method, some variables are identified and transferred from the active set to the passive set so that the sum of the square of residual errors strictly decreases. The solution to the constrained least squares problem will then be the solution to the unconstrained least square problem using only the variables corresponding to the passive set. The regression coefficients corresponding to the remaining variables in the active set are set to zero (Bro and Jong, 1997).

The NNLS algorithm proposed by Lawson and Hanson is adapted to WSA of gamma-ray spectra and presented in Algorithm

1. In the algorithm, P is the passive set, A is the active set and holds the indices of those isotopes (library spectra) that their quantities are currently fixed at zero ($\boldsymbol{x} = \boldsymbol{0}$), $\mathbf{R}^P$ denotes the matrix containing only the library spectra currently used in the passive set P and the intermediate regression vector $\boldsymbol{z}^P$ contains their optimum fitting parameters estimated using least square method. If the set A is empty or all w_j, $j\epsilon A$, are below a certain low tolerance, the isotopes with positive quantities are identified which their corresponding indices are in set P and thus the problem is solved (stage 1). If any w_j, $j \in A$, is positive, the isotope with the highest w_j is included in the set P (stages 1.2 and 1.3) and the intermediate regression vector $\boldsymbol{z}$ is calculated on this new set (stage 1.4). When a new isotope is included in the passive set P, there is a chance that in the unconstrained solution to the new least squares problem (stage 1.4) some of the regression coefficients will turn negative (stage 1.4.1). Calling the new estimate $\boldsymbol{z}$ and the former $\boldsymbol{x}$, it is possible to adjust the new estimate to satisfy the constraints. The old estimate is feasible but with a worse fit than the new estimate $\boldsymbol{z}$ which, however, is not feasible. Somewhere along the line segmented $\boldsymbol{x}$: $= \boldsymbol{x} + \alpha(\boldsymbol{z}\text{-}\boldsymbol{x})$, $0 \leq \alpha \leq 1$, there is an interval for which the inequalities are not violated. As the fit is strictly decreasing as $\alpha \to 1$, it is possible to find that particular value of α which minimizes the fit and yet retains as many isotopes in the passive set as possible (stage 1.4.2). For the optimal α, one or several quantity of isotopes will become zero and hence they are removed from the set P (step 1.4.4). However, the fit of the model will be improved compared with the prior passive set of isotopes. After adjusting the active and passive sets, the unconstrained solution using the current passive set is calculated (stage 1.4.5). The regression coefficients of isotopes removed from the set P are set to zero (stage 1.4.6). The stages 1.4.1–1.4.6 are repeated until all violating isotopes are moved to the set A. If all regression coefficients in $\boldsymbol{z}$ are positive, the regression vector $\boldsymbol{x}$ is set equal to $\boldsymbol{z}$ (step 1.5) and a new set of $\boldsymbol{w}$ is calculated (step 1.6). The detailed description of this algorithm can also be seen in (Lawson and Hanson, 1995; Bro and Jong, 1997). Lawson and Hanson have proven that the true active set is identified after a finite number of iterations and the solution can be found by applying the standard least squares algorithm for the variables corresponding to the passive set. Although this algorithm works well in many applications, it depends on the $\mathbf{R}^T\mathbf{R}$ which is infeasible for ill-conditioned problems (Meric et al., 2012; Meric et al., 2012).

Algorithm 1: NNLS

Input: $\mathbf{R} \in \mathbb{R}^{n \times m}$, $\boldsymbol{N} \in \mathbb{R}^n$, $\boldsymbol{B} \in \mathbb{R}^n$, tolerance

Output: $\hat{\boldsymbol{x}} \geqslant 0$ such that $\hat{\boldsymbol{x}} = \text{argmin}||\boldsymbol{N} - \mathbf{R}\boldsymbol{x} - \boldsymbol{B}||^2$

Initialization: $P = \varnothing$,

   $A = \{1, 2, \ldots, m\}$, $\boldsymbol{x} = \boldsymbol{0}$, $\boldsymbol{w} = \mathbf{R}^T(\boldsymbol{N} - \mathbf{R}\boldsymbol{x} - \boldsymbol{B})$

Repeat

   1. Proceed if $A \neq \varnothing \wedge (\max_{j \in A}(w_j) > tolerance)$

   1.2. $j = \text{argmax}_{j \in A}(w_j)$

   1.3. Include the index j in P and remove it from A

   1.4. $\boldsymbol{z}^P = \left[(\mathbf{R}^P)^T\mathbf{R}^P\right]^{-1}(\mathbf{R}^P)^T\boldsymbol{N}$

      1.4.1. Proceed if $\min(\boldsymbol{z}^P) \leqslant 0$

      1.4.2. $\alpha = -\min_{j \in P}\left[x_j/(s_j - x_j)\right]$

      1.4.3. $\boldsymbol{x} := \boldsymbol{x} + \alpha(\boldsymbol{z} - \boldsymbol{x})$

      1.4.4. Update A and P

      1.4.5. $\boldsymbol{z}^P = \left[(\mathbf{R}^P)^T\mathbf{R}^P\right]^{-1}(\mathbf{R}^P)^T\boldsymbol{N}$

      1.4.6. $\boldsymbol{z}^A = \boldsymbol{0}$

   1.5. $\boldsymbol{x} = \boldsymbol{z}$

   1.6. $\boldsymbol{w} = \mathbf{R}^T(\boldsymbol{N} - \mathbf{R}\boldsymbol{x} - \boldsymbol{B})$

2. Comment: The computation is completed

---

3. Principle of Sequential Gamma-ray Spectrum Deconvolution algorithm

The SGSD algorithm produces a sequence of spectra converging to the measured spectrum of a gamma-ray detector. In other words, an isotope of unknown gamma-ray source is identified and the respective response of the detector to the unknown source is reconstructed in each step of the algorithm. The SGSD algorithm exploits the advantages of the active set method and evolutionary optimization techniques to analyze the complex gamma-ray spectra. The idea of active and passive sets along with a surplus set is used for identification purposes, while the same method used in (Shahabinejad and Vosoughi, 2018) is applied for fitting and controlling the fitting process in the new algorithm. The evolutionary optimization techniques have an ability to easily put the constraints on the desired variables. Moreover, these methods, in contrast to the NNLS method, do not deal with the inverse of $\mathbf{R}^T\mathbf{R}$ matrix, since they try to find an optimal solution in a direct procedure in the search space. However, the dimensions of the search space increases with increasing the number of library spectra which, in turn, means an increment in the number of iterations or running time to reach the stopping criteria. In other words, these methods suffer from the so called "curse of dimensionality" problem (Chen et al., 2015). For example, suppose that we have given the response matrix of a gamma-ray detector to 700 gamma-ray sources in 1024 energy bins and the response to an unknown mixed gamma-ray source using the same detector. Also, assume that the unknown source is likely to contain 30 gamma-ray isotopes. Without applying an identification method, the iterative algorithms should update 700 variables using a large dimension matrix in each iteration. The SGSD algorithm can overcome the "curse of dimensionality" problem since it uses an identification procedure in the fitting process. PSO is chosen for fitting process in the SGSD algorithm, since it has the advantages of simple implementation, fewer control parameters and better convergence performance among the evolutionary optimization techniques (Gou et al., 2017; Shahabinejad and Vosoughi, 2018).

3.1. Particle swarm optimization method

PSO (Eberhart and Kennedy, 1995) is a branch of artificial intelligence which imitates the social behaviors observed in nature to solve various complex mathematical optimizations (Shahabinejad and Sohrabpour, 2017). Like other evolutionary algorithms, PSO is also a population-based search algorithm that starts with an initial population of individuals, called particles, which are distributed in a random manner over the search space (Eberhart and Kennedy, 1995). A particle represents a candidate solution for a multi-dimensional search problem and has a velocity and a position (solution vector). The set of candidate solutions move through the search space following typical dynamics in search of a global optimum for a cost function (or fitness function). PSO directs the candidate solutions to the global optimum by memorizing both the best position found by all candidate solutions and the best positions found by each candidate solutions in the search process. For the purpose of gamma-ray spectrum analysis, PSO tries to find a global optimum to the following single objective optimization problem (Shahabinejad and Vosoughi, 2018).

$$\underset{\boldsymbol{x}}{Minimize} \left\{\| \boldsymbol{N} - \boldsymbol{N}_{est} \|_2\right\}$$
$$\text{where} \quad \boldsymbol{N}_{est} = \mathbf{R}\boldsymbol{x} + \boldsymbol{B}, \tag{6}$$
$$\text{w.r.t.} \quad 0 \leqslant x_j \leqslant 1$$

where, $\boldsymbol{N}$ vector is the normalized recorded counts of NaI(Tl) detector in all energy bins and $\boldsymbol{N}_{est}$ is the estimated gamma-ray spectra in each iteration of the algorithm with the updated $\boldsymbol{x}$. The detailed

description of gamma-ray spectrum analysis using PSO method can be seen in (Shahabinejad and Vosoughi, 2018). In gamma-ray spectrum analysis using PSO method a "comparison module" introduced by Alamaniotis and Jevremovic (2015) have also been applied for controlling the fitting process. The importance of this module is due to the correlation between the library spectra which leads to a possibility of miscalculation of relative abundances. It consists of two statistical tests (Alamaniotis and Jevremovic, 2015): Significance Of Abundance (SOA) and Theil (1967) tests. In SOA test, the estimated relative abundances are compared to the standard errors of the estimated values to check whether the estimated abundances are significantly different than zero as presented below.

IF $x_j < \sigma(x_j)$
 $x_j = 0;$
ELSE
 $x_j = x_j;$
END

where $\sigma(x_j)$ is the standard error of the estimated value and is computed for isotopes as shown in Eq. (7):

$$\sigma(x_j) = \frac{||\mathbf{N} - \mathbf{N}_{est}||_2}{\sqrt{n - m} \times ||\mathbf{R}_j - \overline{\mathbf{R}}_j||_2}, \quad j = 1, \ldots, m \qquad (7)$$

Here, $\mathbf{R}_j$ is the jth column of the response matrix with dimension of n, m is the number of pertinent libraries and $\overline{R}_j$ is the mean of the jth column of the response matrix.

If one or more coefficients are found to be zero after applying the SOA test, the respective spectra are removed and the second step of the comparison module is activated. In the second step the Theil coefficient is applied to compare the estimated spectrum (i.e. not including rejected spectra) and initial spectrum as follow (Alamaniotis and Jevremovic, 2015; Shahabinejad and Vosoughi, 2018):

$$\text{Theil} = \frac{||\mathbf{N} - \mathbf{N}_{est}||_2}{||\mathbf{N}||_2 + ||\mathbf{N}_{est}||_2} \qquad (8)$$

The Theil coefficient values in Eq. (8) are limited to the interval [0 1], where 0 denotes a perfect estimation and 1 denotes a very poor estimation. The perfect estimation occurs when the estimated spectrum ($\mathbf{N}_{est}$) perfectly matches point to point to the measured spectrum ($\mathbf{N}$), while a very poor estimation occurs when the estimated and the measured spectra do not match at any point. The Theil coefficient is used to control the rejection process of the SOA test in such a way that when the value of this coefficient in current iteration is larger than that of the previous one, the control process of the "comparison module" is completed.

3.2. Sequential Gamma-ray spectrum Deconvolution algorithm

As presented in algorithm 2, inputs to the SGSD algorithm are the detector response matrix (**R**), spectra related to an unknown source in the detector (**N**) and background (**B**), and a predetermined threshold (δ) to check the improvement in Theil coefficient in the evolution process of the algorithm. In this work, the PSO method is applied to minimize both mean absolute error (MAE) and root mean square error (RMSE) (Func in algorithm 2) by solving the following single objective optimization problem.
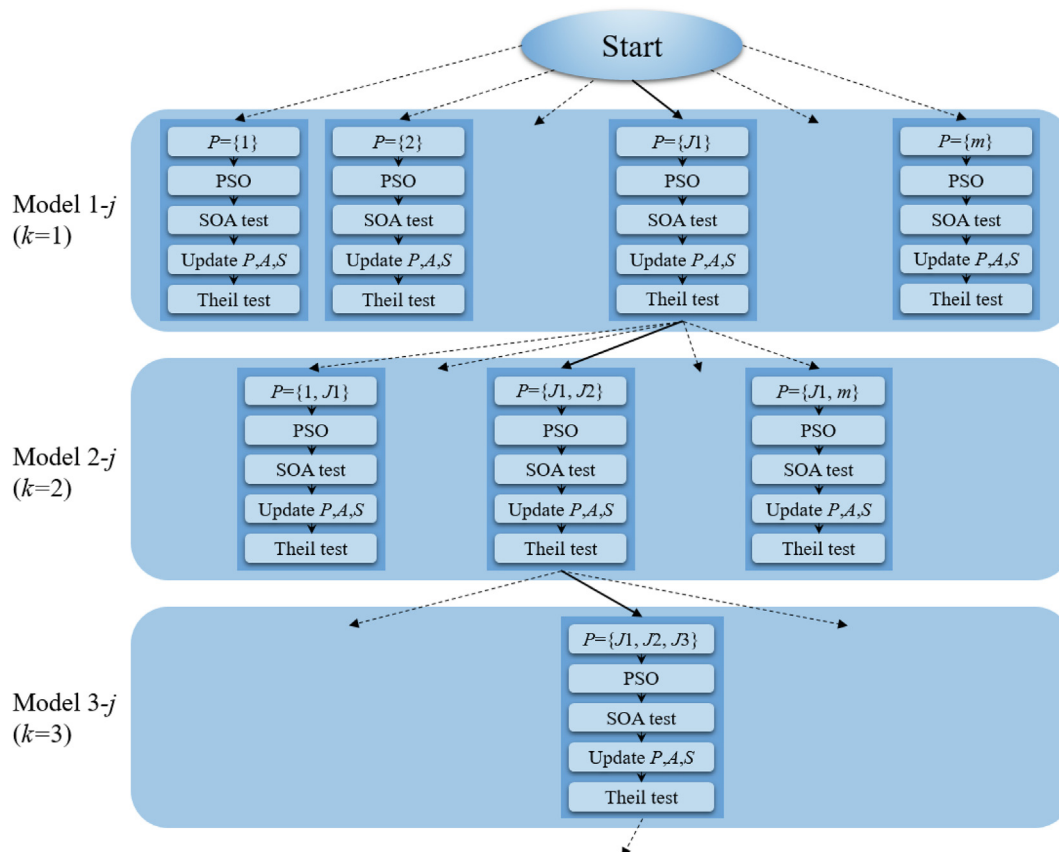

Fig. 1. Choice of library spectra using algorithm 2. Model 1-*J1* is the fittest model between the Models 1-*j* (P={j}, j = 1,2,..,m) while the Models 2-*J2* and 3-*J3* are the fittest models between the Models 2-*j* and 3-*j* respectively.

$$\underset{\boldsymbol{x}}{Minimize}\left\{\sqrt{\left(\frac{MAE}{max\_MAE}\right)^2+\left(\frac{RMSE}{max\_RMSE}\right)^2}\right\}$$

$$\text{where}\quad MAE=\frac{1}{n}\parallel \mathbf{N}-\mathbf{N}_{est}\parallel_1,$$

$$RMSE=\frac{1}{\sqrt{n}}\parallel \mathbf{N}-\mathbf{N}_{est}\parallel_2, \qquad (9)$$

$$\boldsymbol{N}_{est}=\mathbf{R}^P\boldsymbol{x}^P+\boldsymbol{B},$$

$$\text{w.r.t.}\quad 0\leqslant x_j^P\leqslant 1$$

where, n is the number of channels or energy bins, $\parallel\cdot\parallel_2$ and $\parallel\cdot\parallel_1$ are L^2-norm and L^1-norm of a vector respectively, and $\mathbf{R}^P$ and $\boldsymbol{x}^P$ are the response matrix containing only the library spectra currently used in the passive set P and corresponding relative abundances respectively. The scaling of MAE and RMSE are performed using max_MAE and max_RMSE obtained by the comparison of a library spectrum and the measured (or simulated using Monte Carlo

simulations) spectrum. The motivation for selecting the objectives MAE and RMSE is to exploit their properties in order to find a "good" solution in terms of two measures of goodness of fit. The RMSE minimizes bias in the solution, while the MAE measures how close estimated spectra are to the measured (or simulated) ones (Alamaniotis et al., 2013).

The output of the algorithm is also a list of detected radioisotopes with their estimated fractional abundances. In the algorithm, the surplus set S holds the indices of those relative abundances that have zero values ($x_j = 0$), while the passive set P comprises all indices of m relative abundances that have positive values. Furthermore, all the indices of m relative abundances are included in the active set A in initialization step of the algorithm, i.e., all the pertinent library spectra are candidate for reconstructing the detector output spectrum. In addition, $\mathbf{R}^P$ denotes the matrix containing only the library spectra currently used in the passive set P, while the intermediate regression vector $\boldsymbol{z}^P$ contains their optimum fitting parameters estimated using PSO method along with SOA and Theil tests. At the end of algorithm the indices of identified radioisotopes are in the passive set, while the rejected ones are in the surplus set. The problem is solved if the set A is empty, i.e., all the identified isotopes with a positive quantity as well as rejected ones are determined or if the improvement in Theil coefficient is smaller than the small value δ, (stage 1 of algorithm 2). It has to be noted that there is no specific threshold for determining the small value δ and it depends on the experience of the modeler

Table 1
Description of sources and background in experiments.

Source	Activity (Bq)	Net count rate (#/s)
1) ^{60}Co	30565.18	470.04
2) ^{137}Cs	33110.21	228.01
3) ^{22}Na	3944.50	93.62
4) ^{152}Eu	22655.27	322.99
5) ^{241}Am	34304.69	101.43
Background	–	630.55

Table 2
List of 30 radioisotopes used in Monte Carlo simulations. The labels placed on the left side of isotopes are used in the results section.

Isotope	Energy (keV)	Intensity (%)
3) ^{22}Na	511.00, 1274.53	180.00, 99.94
10) ^{24}Na	1368.63, 2754.03	100.00, 99.94
12) ^{40}K	1460.83	11.00
14) ^{46}Sc	889.28, 1120.55	99.98, 99.99
11) ^{51}Cr	320.08	10.00
13) ^{54}Mn	834.85	99.98
6) ^{56}Co	846.77, 977.37, 1037.84, 1175.10, 1238.28, 1360.22, 1771.35, 2015.18, 2034.76, 2598.46	100.00, 1.44, 13.99, 2.28, 67.60, 4.33, 15.69, 3.08, 7.88, 17.28
7) ^{56}Co	122.06, 136.47	85.60, 10.68
8) ^{15}Co	810.78	99.00
1) ^{60}Co	1173.24, 1332.50	99.97, 99.99
15) ^{88}Y	898.04, 1836.06	93.70, 99.20
17) ^{109}Cd	88.04	3.61
19) ^{111}In	171.28, 245.40	90.00, 94.00
16) ^{133}Ba	53.16, 79.61, 81.00, 276.40, 302.85, 356.02, 383.85	2.199, 2.62, 34.06, 7.16, 18.33, 62.05, 8.94
2) ^{137}Ca	661.66	85.10
18) ^{139}Ce	165.86	80.00
4) ^{152}Eu	121.78, 244.70, 344.27, 411.12, 443.97, 778.90, 867.38, 964.08, 1085.87, 1089.74, 1112.07, 1212.95, 1299.14, 1408.01	28.58, 7.58, 26.50, 2.23, 2.82, 12.94, 4.25, 14.61, 10.21, 1.73, 13.64, 1.42, 1.62, 21.01
9) ^{154}Eu	123.07, 247.93, 591.76, 692.42, 723.30, 756.76, 873.19, 996.26, 1004.73, 1274.44, 1596.50	40.79, 6.95, 4.99, 1.80, 20.22, 4.57, 12.27, 10.60, 18.01, 35.19, 1.80
20) ^{192}Ir	205.80, 295.96, 308.46, 316.51, 468.07, 484.58, 588.58, 604.41, 612.47	3.30, 28.67, 30.00, 82.81, 47.83, 3.18, 4.52, 8.23, 5.31
21) Bi	569.70, 1063.66, 1770.24	97.74, 74.5, 6.87
5) ^{241}Am	59.50	35.90
22) ^{238}U series		
^{234m}Pa	1001.03	0.84
^{214}Pb	53.23, 242.00, 295.22, 351.93, 785.96	1.20, 7.43, 19.30, 37.60, 1.07
^{214}Bi	609.31, 665.45, 768.36, 806.17, 934.06, 1120.29, 1155.19, 1238.11, 1280.96, 1377.67, 1401.50, 1407.98, 1509.23, 1661.28, 1729.60, 1764.49, 1847.42, 2118.55, 2204.21, 2447.86	46.1, 1.46, 4.94, 1.22, 3.03, 15.10, 1.63, 5.79, 1.43, 4.00, 1.27, 2.15, 2.11, 1.15, 2.92, 15.40, 2.11, 1.14, 5.08, 1.57
23) ^{232}Th series		
^{228}Ac	99.51, 129.07, 209.25, 270.25, 328.00, 338.32, 409.46, 463.00, 755.32, 772.29, 794.95, 835.71, 911.20, 964.77, 968.97, 1588.19, 1630.63	1.26, 2.42, 3.89, 3.46, 2.95, 11.27, 1.92, 4.40, 1.00, 1.49, 4.25, 1.61, 25.80, 4.99, 15.80, 3.22, 1.51
^{212}Bi	727.33, 785.37, 1620.50	6.58, 1.10, 1.49
^{208}Tl	277.35, 510.77, 583.19, 763.13, 860.56, 2614.53	6.31, 22.60, 84.50, 1.81, 12.42, 99.00
24) ^{227}Ac series		
^{235}U	109.16, 143.76, 163.36, 185.71, 202.11, 205.31	1.54, 10.96, 5.08, 57.2, 1.08, 5.02
^{223}Ra	122.32, 144.23, 154.21, 269.46, 323.87, 338.28, 445.03	1.19, 3.22, 5.62, 13.70, 3.93, 2.79, 1.27
^{227}Th	50.13, 79.72, 93.93, 210.65, 235.97, 256.25, 286.12, 300.00, 304.52, 329.85, 334.38	8.00, 1.89, 1.37, 1.11, 12.3, 7.00, 1.54, 2.32, 1.2, 2.7, 1.05

(Alamaniotis and Jevremovic, 2015). In this work a small value of 1% ($\delta = 0.01$) is considered for predetermined improvement in Theil coefficient.

As presented in Fig. 1, the different parts of the algorithm 2 can be described using the parameter k as following steps:

Algorithm 2: SGSD

Input: $\mathbf{R} \in \mathbb{R}^{n \times m}$, $\mathbf{N} \in \mathbb{R}^n$, $\mathbf{B} \in \mathbb{R}^n$, δ
Output: $\hat{\boldsymbol{x}} > 0$ such that $\hat{\boldsymbol{x}} = \text{argminFunc}$
Initialization: $P = \varnothing$, $A = \{1, 2, \ldots, m\}$, $S = \varnothing$,
   $k = 1, j = 1$, $\text{Theil}_1 = 10^6$, $\text{Theil}_2 = 1$
Repeat
1. Proceed if $A \neq \varnothing \wedge ((\text{Theil}_k - \text{Theil}_{k+1})/\text{Theil}_k \geqslant \delta)$
   1.1. For $j \in A$
      1.1.1. Include the index j in P and remove it from A
      1.1.2. $\boldsymbol{z}^P = \text{argminFunc}$
      1.1.3. Apply SOA test
      1.1.4. Proceed if $\min(\boldsymbol{z}^P) == 0$
         1.1.4.1. Update P and S
         1.1.4.2. $\boldsymbol{z}^P = \text{argminFunc}$
         1.1.4.3. $\boldsymbol{z}^A = \mathbf{0}$
         1.1.4.4. $\boldsymbol{z}^S = \mathbf{0}$
      1.1.5. $\hat{\boldsymbol{x}} = \boldsymbol{z}^P$
      1.1.6. Compute Theil_{k+2}
      1.1.7. Proceed if $\text{Theil}_{k+2} < \text{Theil}_{k+1}$
         1.1.7.2. $\text{Theil}_{k+1} = \text{Theil}_{k+2}$
         1.1.7.2. $J = j$
      1.1.8. Remove the index j from P and include it in A
   1.2. Include the index J in P and remove it from A
   1.3. Proceed if $(\text{Theil}_{k+1} - \text{Theil}_{k+2})/\text{Theil}_{k+1} < \delta$
      1.3.1. Update P and S
      1.3.2. $\boldsymbol{z}^P = \text{argminFunc}$
      1.3.3. $\hat{\boldsymbol{x}} = \boldsymbol{z}^P$
   1.4. $k = k + 1$
2. Comment: The computation is completed

Step I ($k = 1$): The output spectrum is considered as a function of library spectrum $j, j \in A$ (stages 1.1 and 1.1.1). In other words, the passive set has only 1 index related to the jth library spectrum ($P = \{j\}$) in the algorithm. It is called the input-output model 1-j. So, we have m models: 1–1, 1–2, ..., 1-m. All models can be run in parallel, since they are independent.

Step II: The optimum fitting vector $\boldsymbol{z}$ is found using PSO method for each model (stage 1.1.2). When a new isotope is included in the passive set P, there is a chance that in the unconstrained solution to the new optimization problem the regression coefficient of added isotope will turn zero after applying SOA test (stages 1.1.3 and 1.1.4). The quantity of the added isotope will become zero and hence it is removed from the set P and included in the Set S (stage 1.1.4.1). The new estimate of the intermediate coefficient $\boldsymbol{z}$ is then calculated (1.1.4.2) and the regression coefficients of isotopes removed from the set P and those of present in set A are set to zero (stages 1.1.4.3 and 1.1.4.4). The optimum model with the least Theil coefficient is then adopted out of the m models (stages 1.1.5–1.1.8). The index of the identified isotope is removed from the active set and included in the passive set (stage 1.2). If the improvement in Theil coefficient is smaller than the small value δ, the added isotope is removed from the set P and included in the Set S and the intermediate coefficient $\boldsymbol{z}$ is then calculated using the updated P and the regression vector $\hat{\boldsymbol{x}}$ is set equal to $\boldsymbol{z}^P$ (stages 1.3.1 and 1.3.3). It should be noted that the identified isotope in this step ($P = J1$ in Fig. 1) has the most similar spectrum to the spectrum of the unknown source.

Step III ($k = k + 1$): Starting from an optimal model at step I, say model 1-$J1$, where only the library spectrum $\mathbf{R}^{J1}$ ($P = \{J1\}$) appears in the algorithm, take all the combinations $\mathbf{R}^{\{J1,j\}}$ ($j \in A$). Thus, we get m-1 models each of which is named model 2-j. Then finding again an optimum model with the least Theil coefficient just as in step II ($P = \{J1, J2\}$ in Fig. 1). Similar to the step I, the models can be run in parallel.

Step IV: Repeat step III in a similar way by transferring another index from active set to the passive set. For example, the algorithm starts with the optimal model 2-$J2$ as presented in Fig. 1 where only the library spectra $\mathbf{R}^{\{J1,J2\}}$ ($P = \{J1,J2\}$) appears in the algorithm. Then, all the combinations $\mathbf{R}^{\{J1,J2,j\}}$ ($j \in A$) are taken to get m-2 models each of which is named model 3-j. Then finding again an optimum model with the least Theil coefficient just as in step II ($P = \{J1, J2, J3\}$ in Fig. 1).

Step V: The algorithm is stopped when the set A is empty or when the improvement in the Theil coefficient is smaller than δ.

It has to be noted that the presented steps of the algorithm are inspired from Ref. (Takagi and Sugeno, 1985) where the identification and modeling of a complex nonlinear system is performed using a Sugeno fuzzy system. Nevertheless, it is assumed that one deals with a linear system of equations in WSA of gamma-ray spectra after treating its nonlinearities and the simple algorithm presented here can appropriately identify and model the system. The main difference between current work and the method used a fuzzy identification (Alamaniotis and Jevremovic, 2015) is that their fuzzy identification procedure is done based on only photopeaks present in library spectra, while the whole procedure of our algorithm is performed by whole information of the library spectra. Although they have attempted to determine the elements that have not been identified or misidentified in the next steps of their algorithm and modify the identification stage, the identification method presented here can accurately identify the radioisotopes and determine their corresponding fractional abundances.

4. Validation of the proposed algorithm

The implemented SGSD algorithm is applied to analyze both the simulated and experimentally measured gamma-ray spectra obtained using a high efficiency 3×3 inch NaI(Tl) detector. Two example cases including an experimental study and a simulation one are presented for detailed description and comparison of the SGSD and NNLS algorithms in analyzing gamma-ray spectra. Fur-

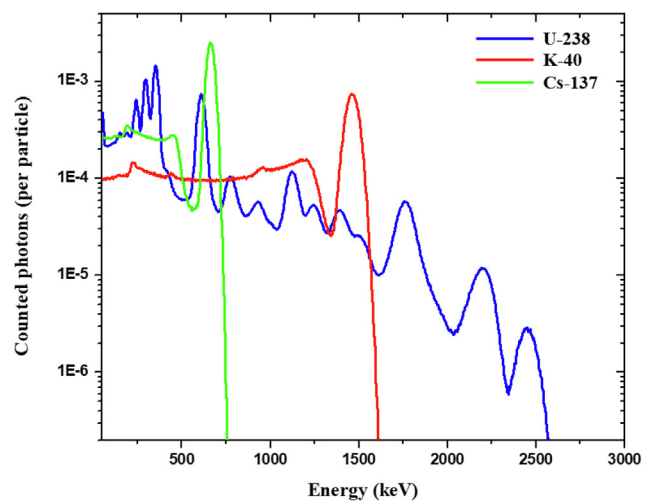**Fig. 2.** Simulated responses of a NaI(Tl) detector to ^{238}U series, ^{40}K and ^{137}Cs gamma-ray sources using MCNPX Monte Carlo code. Note that the values presented on vertical axis are normalized to the number of particle source histories.

thermore, 8 case studies are presented to compare the implemented algorithm against the NNLS algorithm.

In the case of empirical study, the ^{60}Co, ^{137}Cs, ^{22}Na, ^{152}Eu and ^{541}Am point sources with the specifications presented in Table 1 are used to obtain the response matrix (**R**) as well as detector response to different combinations of sources (**N**) and background (**B**). Measurement system includes standard NIM units such as MCA, Spectroscopy Amplifier and related software for data acquisition. The point sources are located in front surface of detector and along its axis. The detector response matrix is measured in 1024 channels of MCA and in 3600 seconds spectrum acquisition with respect to the count rates of isotopes in the background-subtracted full spectrum of the detector (Net count rate) as presented in Table 1. The background count rate in the full spectrum is also included in Table 1. In addition, the energy calibration of the detection system is performed using the ^{241}Am (59.5 keV), ^{152}Eu (121.78, 244.70, 344.27, 778.90, 964.08 keV), ^{22}Na (511.00, 1274.53 keV), ^{137}Cs (661.66 keV) and ^{60}Co (1173.24, 1332.50 keV) sources, resulting in 2.5716 keV per channel. It should be noted that the recorded counts in energies below 50 keV are not considered to analyze the gamma-ray spectra in all algorithms.

In simulation studies, the required detector spectra and response matrix are simulated using MCNPX (Pelowitz, 2008) Monte Carlo code. A 3×3 inch NaI(Tl) detector is modeled by MCNPX computer code to obtain the response matrix and detector spectra related to isotropic point sources located in 5 cm far from the front surface of the detector. It should be noted that only the photon transport mode is used in simulations. The simulated sources are listed in Table 2. As it is shown in Table 2, each of ^{238}U, ^{232}Th and ^{227}Ac series comprise 3 isotopes. Thus, 24 library spectra in the range of 50–3000 keV are simulated using MCNP F8 tally to make the response matrix. A label is assigned to each isotope (library spectrum) placed on the left side of it in Table 2. These labels are used in the results section. To simulate the observed energy broadening in a physical radiation detector, the GEB option is used in the simulations. Using the GEB option in MCNPX code, the tallied energy is broadened by sampling from the Gaussian distribution presented in Eq. (10) (Pelowitz, 2008).

$$f(E) = C\exp\left(-\frac{E - E_0}{A}\right)^2 \tag{10}$$

where, E is the broadened energy, E_0 is the unbroadened energy of the tally, C is a normalization constant and A is the Gaussian width.

The Gaussian width is related to the Full Width Half Maximum (FWHM) by Eq. (11).

$$A = \frac{FWHM}{2\sqrt{\ln 2}} \tag{11}$$

The desired FWHM is specified using Eq. (12) by the user–provided constants, a, b, and c.

$$FWHM = a + b\sqrt{E + cE^2} \tag{12}$$

The used constants in simulations are (Hakimabad et al., 2007): $a = -0.00789$ MeV; $b = 0.06769$ MeV$^{1/2}$; $c = 0.21159$ MeV^{-1}.

It should be noted that the number of particle source histories (NPS) (Pelowitz, 2008) is considered large enough for making the response matrix so that the relative statistical error in each energy bin is less than 5%. The simulated responses of the NaI(Tl) detector to the ^{238}U series, ^{40}K and ^{137}Cs sources obtained using MCNPX Monte Carlo code are illustrated in Fig. 2. Since the output of the MCNPX code is normalized to NPS, the values presented on the vertical axis of Fig. 2 are internally normalized to the number of particle source histories.

The implemented SGSD algorithm is executed on a common PC with Intel Core i7 2.6 GHz processor. To reach the stopping condition criteria an average of about 20 seconds of computer time is required for different scenarios of this work. However, the run time can be significantly reduced by running the different models of algorithm in parallel as mentioned in Step I of algorithm. The developed algorithm is applied to analyze both the simulated and experimentally measured gamma-ray spectra using the obtained response matrices and compared against NNLS algorithm in the next section.

5. Results

In the first example case, both SGSD and NNLS algorithms are applied for WSA of a 10 s long measured gamma-ray spectrum related to mixed ^{137}Cs, ^{22}Na and ^{152}Eu sources. The obtained results are presented in Table 3. The results include the identified radioisotopes (I.R.), their respective estimated relative abundances (R.A.) and standard deviations (S.D.) as well as calculated Theil coefficient between the estimated detector spectra and the measured ones. It should be noted that the standard deviations and Theil coefficient are calculated using Eqs. (7) and (8) respectively,

Table 3
Obtained results of spectrum analysis using SGSD and NNLS algorithms for a 10 s long measured spectrum of a mixed source containing ^{137}Cs, ^{22}Na and ^{152}Eu radioisotopes.

Model k-j	Analysis results							
	SGSD algorithm				NNLS algorithm			
	I.R.	R.A.	S.D.	Theil	I.R.	R.A.	S.D.	Theil
Model 1-j	^{152}Eu	1.00000	0.02875	0.3397970	^{60}Co	0.28436	0.01488	0.4931587
Model 2-j	^{152}Eu	0.93500	0.01518	0.1540650	^{60}Co	0.15211	0.01096	0.3223810
	^{137}Cs	0.46877	0.00957		^{137}Cs	0.57122	0.01859	
Model 3-j	^{152}Eu	0.84419	0.01339	0.1356194	^{60}Co	0.02962	0.00556	0.1515858
	^{137}Cs	0.43668	0.00844		^{137}Cs	0.45274	0.00944	
	^{22}Na	0.15461	0.00898		^{152}Eu	0.89730	0.01497	
Model 4-j	^{152}Eu	0.84151	0.01333	0.1349635	^{60}Co	0.00422	0.00497	0.1347299
	^{137}Cs	0.43572	0.00841		^{137}Cs	0.43294	0.00843	
	^{22}Na	0.15477	0.00895		^{152}Eu	0.85572	0.01338	
	^{241}Am	**0.07248**	**0.02091**		^{22}Na	0.15604	0.00897	
Model 5-j	–	–	–	–	^{60}Co	0.00388	0.00495	0.1341117
	–	–	–	–	^{137}Cs	0.43248	0.00840	
	–	–	–	–	^{152}Eu	0.85379	0.01332	
	–	–	–	–	^{22}Na	0.15557	0.00894	
	–	–	–	–	^{241}Am	0.06245	0.02089	

Table 4
Active (*A*), passive (*P*) and surplus (*S*) sets related to Table 3.

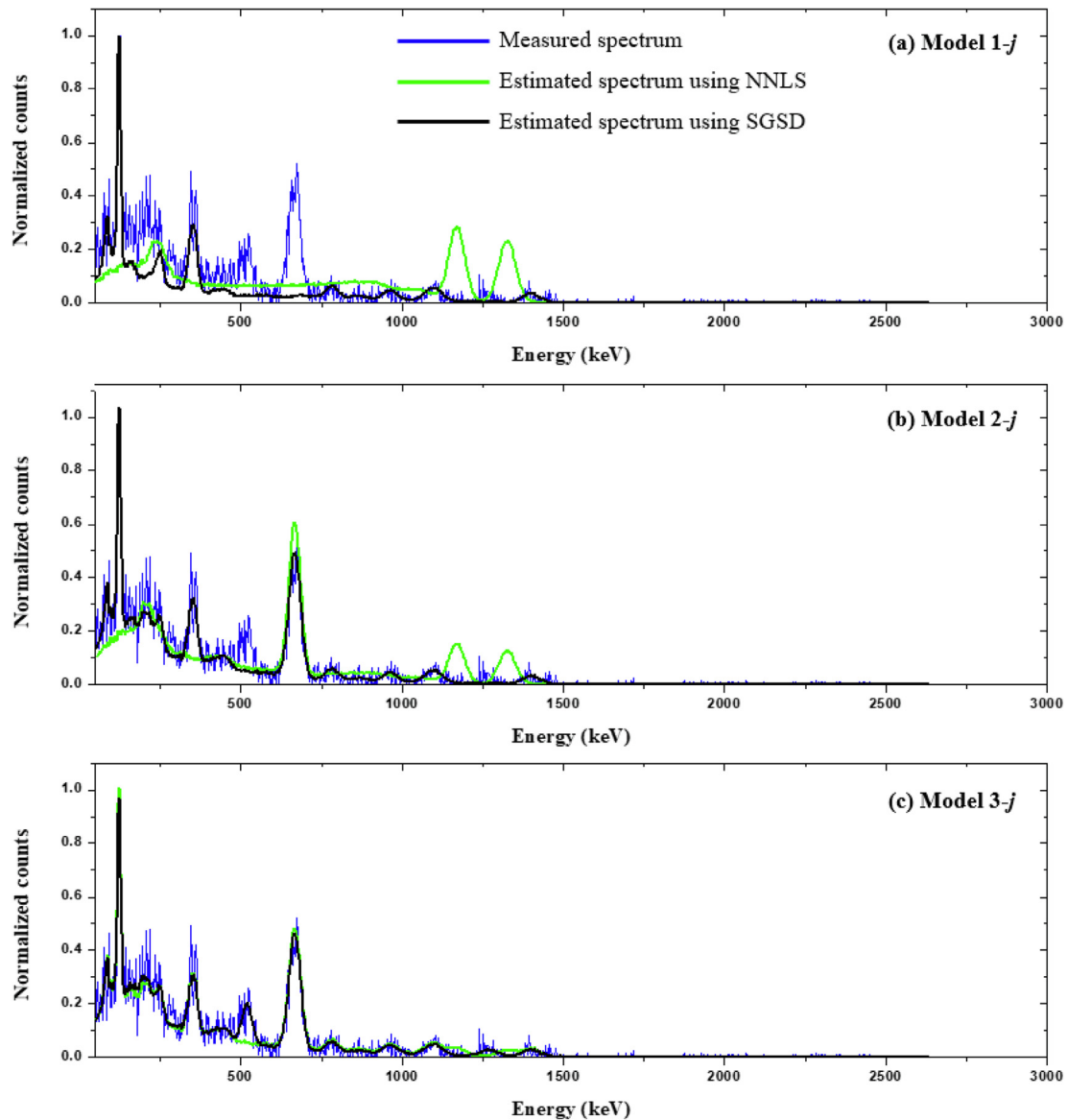| Model *k-j*	Sets				
	SGSD algorithm			NNLS algorithm	
	A	*P*	*S*	*A*	*P*
Model 1- *j*	{1,2,3,5}	{4}	{}	{2,3,4,5}	{1}
Model 2- *j*	{1,3,5}	{4,2}	{}	{3,4,5}	{1,2}
Model 3- *j*	{5}	{4,2,3}	{1}	{3,5}	{1,2,4}
Model 4- *j*	–	{4,2,3}	{1,5}	{5}	{1,2,4,3}
Model 5- *j*	–	–	–	–	{1,2,4,3,5}

Fig. 3. The sequence of reconstructed spectra obtained using NNLS and SGSD algorithms and the measured spectrum for a 10 s long measured spectrum of a mixed gamma-ray source containing ^{137}Cs, [22]Na and [152]Eu radioisotopes.

while relative abundances are calculated by dividing the amount of each isotope by the total amount of all isotopes in the sample.

The models *k-j* shown in the left column of the Table 3 are similar to those described in Section 3.2 (see Fig. 1). As it can be observed in Table 3, the results of NNLS algorithm are also repre-

sented according to these models to perform a better comparison between the identification procedures of SGSD and NNLS algorithms. The Active (*A*), passive (*P*) and surplus (*S*) sets of SGSD algorithm as well as the Active (*A*) and passive (*P*) sets of NNLS algorithm related to Table 3 are presented in Table 4. It should
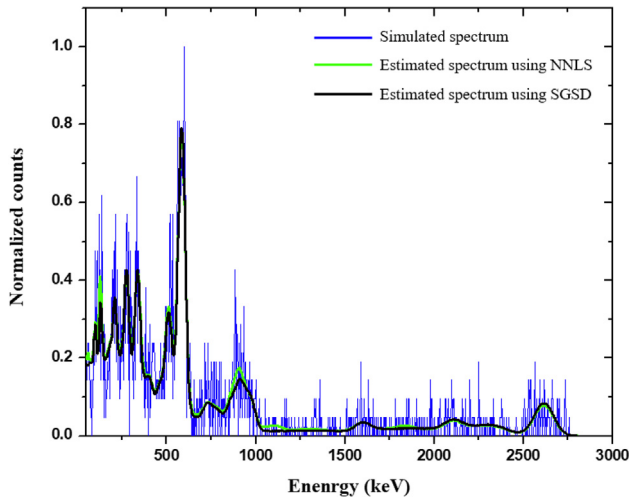
Fig. 4. The estimated spectra obtained using NNLS (Model 11-*j* of Table 5) and SGSD (Model 1-*j* of Table 5) algorithms and the simulated spectrum for a ^{232}Th series gamma-ray source.

be noted that the indices included in sets are similar to those of Table 2. A detailed description of these sets is performed later in this section.

The measured detector spectra are also reconstructed using the obtained relative abundances to find out how well the SGSD and NNLS algorithm can reproduce the input gamma-ray spectra in each model of Table 3. The estimated spectrum of developed SGSD algorithm in comparison with NNLS are calculated using the response matrix in models 1-*j* to 3-*j* where the identification procedure of SGSD algorithm is completed. The sequence of obtained reconstructed spectra is illustrated in Fig. 3.

Since there are 5 library spectra in our experimental studies, both algorithms start with five models in model 1-*j* where *j* is the index of library spectra shown in Table 1. The optimal relative

abundances and corresponding standard deviations are calculated in each model and the fittest model is selected. The fittest models are model 1–4 (^{152}Eu isotope) and model 1–1 (^{60}Co isotope) for SGSD and NNLS algorithms respectively as shown for **Model** 1-*j* in Table 3. The indices 4 and 1 are therefore removed from the active sets and included in the passive sets related to both algorithms as presented for **Model** 1-*j* in Table 4. The reconstructed spectra using the obtained relative abundances are illustrated in Fig. 3(a). As can be seen in Fig. 3(a), the estimated spectrum using SGSD algorithm is more similar to the measured spectrum than that of NNLS algorithm. The calculated Theil coefficients, as a measure of goodness of fit, presented in Table 3 also confirm this. There are only 4 models in model 2-*j*, since an isotope of unknown source has been identified in model 1-*j*. Again, the optimal relative abundances of isotopes and corresponding standard deviations are calculated in each model and the fittest model is chosen. The fittest models are model 2–2 (^{152}Eu-^{137}Cs isotopes) for SGSD algorithm and model 2–2 (^{60}Co-^{137}Cs isotopes) for NNLS as presented for **Model** 2-*j* in Table 3. Thus, the index 2 is transferred from the active sets to the passive sets as presented for **Model** 2-*j* in Table 4. The more similar estimated spectrum of SGSD to the measured spectrum is evident in Fig. 3(b). Improvement in Theil coefficient is about 55% for SGSD algorithm which is more than the stopping condition δ = 1% and therefore the algorithm continues. The identified isotopes in model 3-*j* are ^{22}Na and ^{152}Eu for SGSD and NNLS algorithms respectively as presented in Tables 3 and 4. Since the calculated abundance fraction of ^{60}Co isotope using SGSD algorithm is zero in this step, the isotope is rejected from identification procedure of SGSD and the corresponding index is transferred from active set to surplus set as presented for **Model** 3-*j* in Table 4. As shown in Fig. 3(c), a good estimate of the measured spectrum is calculated using SGSD and the improvement in Theil coefficient is about 12%. The only remaining member of active set related to SGSD algorithm is ^{241}Am, while the isotopes ^{22}Na and ^{241}Am are in the active set of NNLS. The improvement in Theil coefficient in model 4-*j* related to SGSD is about 0.5% that is smaller than the stopping condition. The ^{241}Am isotope is therefore removed from

Table 5

Obtained results of spectrum analysis using SGSD and NNLS algorithms for a simulated spectrum of a source containing only ^{232}Th series.

Model *k-j*	Analysis results							
	SGSD algorithm				NNLS algorithm			
	I.R.	R.A.	S.D.	Theil	I.R.	R.A.	S.D.	Theil
Model 1-*j*	^{232}Th	0.79096	0.01344	0.2115515	^{232}Th	0.80790	0.01343	0.2091902
Model 2-*j*	^{232}Th	0.78222	0.01340	0.2109258	^{232}Th	0.78865	0.01339	0.2084315
	^{152}Eu	**0.08309**	**0.02963**		^{46}Sc	0.03386	0.01246	
Model 3-*j*	–	–	–	–	^{232}Th	0.78155	0.01334	0.2074721
	–	–	–		^{46}Sc	0.03401	0.01241	
	–	–	–		^{57}Co	0.08759	0.01334	
Model 4-*j*	–	–	–	–	^{232}Th	0.77645	0.01334	0.2073891
	–	–	–		^{46}Sc	0.03349	0.01242	
	–	–	–		^{57}Co	0.08779	0.02950	
	–	–	–		^{22}Na	0.01721	0.01839	
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.
Model 11-*j*	–	–	–	–	^{232}Th	0.73939	0.01337	0.2070508
	–	–	–		^{46}Sc	0.01439	0.01244	
	–	–	–		^{57}Co	0.09336	0.02956	
	–	–	–		^{22}Na	0.02417	0.01843	
	–	–	–		^{238}U	0.02119	0.01476	
	–	–	–		^{88}Y	0.02582	0.01443	
	–	–	–		^{207}Bi	0.02983	0.01686	
	–	–	–		^{241}Am	0.02480	0.03996	
	–	–	–		^{58}Co	0.00291	0.01595	
	–	–	–		^{60}Co	0.00087	0.01068	
	–	–	–		^{24}Na	0.00036	0.01465	

Table 6
Active (A), passive (P) and surplus (S) sets related to Table 5. The indices included in sets are similar to those of Table 2.

Model *k*-*j*	Sets				
	SGSD algorithm			NNLS algorithm	
	A	*P*	*S*	*A*	*P*
Model 1-*j*	{1,2,…,22,24}	{23}	{}	{1,2,…,22,24}	{23}
Model 2-*j*	{2,5,8,9,14,22,24}	{23}	{1,3,4,6,7,10,11,12,13,15,…,21}	{1,2,…,13,15, …, 22,24}	{23,14}
Model 3-*j*	–	–	–	{1,2,…,6,8,…,13,15,…, 22,24}	{23,14,7}
Model 4-*j*	–	–	–	{1,2,4,5,6,8,…, 13,15,…, 22,24}	{23,14,7,3}
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
Model 11-*j*	–	–	–	{2,4,6,9,11,12,13,16,…,20,24}	{23,14,7,3,22,15,21,5,8,1,10}

Table 7
Obtained results of spectrum analysis using SGSD and NNLS algorithms for experimentally measured case studies 1–5. The labels included in columns 2 and 5 are similar to those of Table 2.

Case study # a) Isotope-Fractional abundance b) Spectrum acquisition time c) Total counts	Analysis results					
	SGSD algorithm			NNLS algorithm		
	F.A. (%)	S.D. (%)	Error (%)	F.A. (%)	S.D. (%)	Error (%)
Case study 1 a) ^{60}Co −100% b) 1 s c) 524	1) 72.55 4) 27.45	1.94 13.78	27.45 –	1) 67.81 2) 3.67 3) 3.81 4) 24.71	2.06 64.53 66.21 15.20	32.19 – – –
Case study 2 a) ^{60}Co −100% b) 2 s c) 1008	1) 93.60 3) 6.40	1.68 66	6.40 –	1) 91.81 3) 1.81 4) 6.38	1.70 156 65.91	8.19 – –
Case study 3 a) ^{60}Co −48% ^{137}Cs-52% b) 5 s c) 3484	1) 48.44 2) 51.56	1.95 3.11	0.91 0.84	1) 47.82 2) 51.44 3) 0.55 4) 0.03 5) 0.16	1.97 3.11 312 8789 2474	0.38 1.07 – – –
Case study 4 a) ^{137}Cs −55.45% ^{22}Na −6.61% ^{152}Eu-37.94% b) 5 s c) 3378	2) 56.63 3) 8.38 4) 34.99	2.00 14.37 5.13	2.13 26.85 7.78	1) 1.43 2) 55.37 3) 8.60 4) 34.60	46.39 2.04 13.97 5.18	– 0.15 30.23 8.82
Case study 5 a) ^{137}Cs −55.45% ^{22}Na −6.61% ^{152}Eu-37.94% b) 60 s c) 6576	2) 56.00 3) 6.50 4) 37.50	0.62 5.66 1.46	0.99 1.65 1.16	2) 55.58 3) 6.40 4) 37.83 5) 0.19	0.62 5.74 1.45 460	0.24 3.16 0.29 –

the passive set and included in surplus set and the SGSD algorithm is terminated. Thus, the final results of the SGSD algorithm are those presented in **Model** 3-*j* of Table 3. However, the NNLS algorithm continues to model 5-*j*.

The SGSD algorithm is also applied to analyze a simulated gamma-ray spectrum related to a ^{232}Th series source obtained using MCNPX Monte Carlo code and compared against NNLS algorithm in the second example case. The NPS is set 20,000 to have relatively large errors in Monte Carlo calculations, thereby introducing a complex gamma-ray spectrum to analyze using the NNLS and SGSD algorithms (Fig. 4). The obtained results of WSA of simulated gamma-ray spectrum related to the ^{232}Th series including identified radioisotopes (I.R.) and their respective estimated relative abundances (R.A.) and standard deviations (S.D.) as well as calculated Theil coefficient between the estimated detector spectra and simulated ones are presented in Table 5. Furthermore, the Active (*A*), passive (*P*) and surplus (*S*) sets of SGSD algorithm as well as the Active (*A*) and passive (*P*) sets of NNLS algorithm related to Table 5 are presented in Table 6. It should be also noted

that the indices included in sets are similar to those of Table 2. The estimated spectrum of ^{232}Th series gamma-ray source using the developed SGSD algorithm in comparison with NNLS algorithm are calculated using the response matrix and presented in Fig. 4.

In this example case, both algorithms start with 24 models in model 1-*j* where *j* is the index of library spectra shown in Table 2. The optimal relative abundances and corresponding standard deviations are calculated in each model and the fittest model is selected. The fittest model is model 1–23 (^{232}Th series) for both SGSD and NNLS algorithms as presented for **Model** 1-*j* in Table 5. As shown for **Model** 1-*j* in Table 6, the index 23 is therefore removed from the active sets and included in the passive sets related to both algorithms. Since an isotope of unknown source is identified in model 1-*j*, there are only 23 models in model 2-*j*. Again, the optimal relative abundances of isotopes and corresponding standard deviations are calculated in each model and the fittest model is chosen. The fittest models are model 2–4 (^{232}Th-^{152}Eu) for SGSD algorithm and model 2–14 (^{232}Th-^{46}Sc) for NNLS as presented for **Model** 2-*j* in Table 5. In one hand, the improvement in

Table 8
Obtained results of spectrum analysis using SGSD and NNLS algorithms for Monte Carlo simulated case studies 6–8. The labels included in columns 2 and 5 are similar to those of Table 2.

| Case study # a) Isotope-Fractional abundance d) NPS	Analysis results					
	SGSD algorithm			NNLS algorithm		
	F.A. (%)	S.D. (%)	Error (%)	F.A. (%)	S.D. (%)	Error (%)
Case study 6	12) 94.98	0.05	0.02	2) 0.01	479	–
a) ^{40}K-95%	22) 3.96	1.36	0.98	3) 0.04	183	–
^{238}U-4%	23) 1.06	4.62	5.62	12) 94.94	0.05	0.07
^{232}Tu-1%				15) 0.04	123	–
d) 5×10^7				17) 2e-3	7014	–
				21) 0.05	136	–
				22) 3.98	1.35	0.44
				23) 0.94	5.15	5.76
Case study 7	12) 94.89	0.04	0.12	2) 0.02	314	–
a) ^{40}K −95%	22) 4.08	1.08	1.93	3) 0.03	154	–
^{238}U −4%	23) 1.03	3.87	3.48	9) 0.01	692	–
^{232}Tu −1%				12) 94.91	0.04	0.10
d) 5×10^8				15) 0.08	51.82	–
				17) 1e-3	7666	–
				19) 1e-3	6268	–
				21) 5e-3	105	–
				22) 4.00	1.06	0.11
				23) 0.94	4.12	6.38
Case study 8	6) 9.83	0.64	1.75	6) 9.18	0.68	8.21
a) ^{56}Co-10%	8) 14.87	0.48	0.88	8) 14.81	0.47	1.25
^{58}Co-15%	13) 14.03	0.50	0.24	13) 14.43	0.48	3.05
^{54}Mn-14%	14) 39.53	0.14	1.18	14) 39.65	0.14	0.87
^{46}Sc-40%	16) 21.07	0.62	5.34	16) 21.27	0.61	6.32
^{133}Ba-20%	17) 0.68	21.91	32.14	17) 0.66	22.24	33.57
^{109}Cd-1%						
d) 4×10^7						
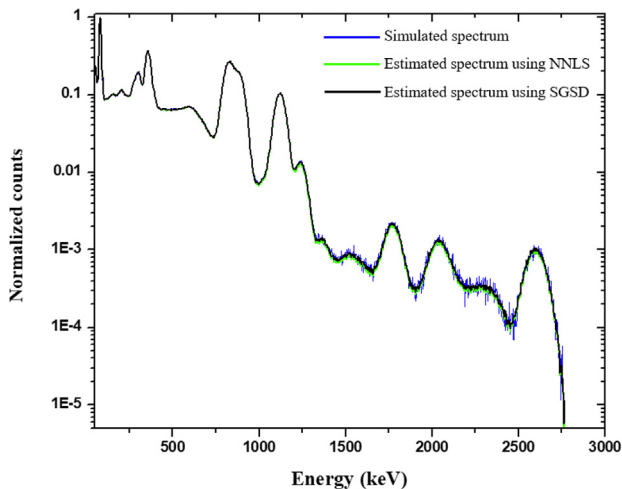
Fig. 5. The estimated spectra obtained using NNLS and SGSD algorithms and the simulated spectrum for a mixed source containing ^{56}Co, ^{58}Co, ^{34}Mn, ^{46}Sc, ^{133}Ba and ^{109}Cd sources.

Theil coefficient in model 2-*j* related to SGSD is about 0.3% which is smaller than the stopping condition. The ^{152}Eu isotope is thus transferred from the passive set to the surplus set and the SGSD algorithm is terminated. The final results of the SGSD algorithm are those presented in **Model** 1-*j* of Table 5. On the other hand, the NNLS algorithm continues to model 11-*j*. It should be noted that in addition to the ^{152}Eu isotope, 15 isotopes are rejected and included in surplus set in model 2-*j* of SGSD algorithm as presented for **Model** 2-*j* in Table 6 since their corresponding relative abundances are zero.

Furthermore, five different empirical case studies are considered to test the implemented SGSD algorithm in comparison with NNLS algorithm as shown in Table 7. These case studies include two measurements of gamma-ray spectra comprised of single ^{60}Co source and 3 measurements consist of different combinations of ^{60}Co, ^{137}Cs, ^{22}Na and ^{152}Eu sources. The fractional abundances of isotopes (F.A. in percent) and their respective fractional standard deviations (S.D. in percent) as well as the errors of obtained abundance fractions with respect to their real fractions (Error in percent) are also calculated and presented in Table 7.

In addition, three different simulated case studies are considered to test the implemented SGSD algorithm in comparison with NNLS algorithm as shown in Table 8. These case studies include two simulations of gamma-ray spectra comprised of mixed ^{40}K, ^{238}U and ^{232}Th series sources and a simulation consist of mixed ^{56}Co, ^{58}Co, ^{54}Mn, ^{46}Sc, ^{133}Ba and ^{109}Cd sources. The fractional abundances of isotopes (F.A. in percent) and their respective fractional standard deviations (S.D. in percent) as well as the errors of obtained abundance fractions with respect to their real fractions (Error in percent) are also calculated and presented in Table 7.

The estimated spectra related to the case study 8 of Table 8 for both SGSD and NNLS algorithms are also calculated and presented in Fig. 5.

6. Discussion and conclusion

As can be seen in Table 3 and Fig. 3, there is a fundamental difference between the identification procedure of the SGSD and NNLS algorithms. In fact, the SGSD algorithm provides a better estimation of the abundance fractions in comparison with NNLS at every step that the algorithms are stopped. Furthermore, the SGSD reject a large number of library spectra in initial steps because they are not involved in the creation of the spectrum of an unknown gamma-ray source based on the information of whole spectrum as discussed in example case 2 and presented in Table 6. Although the final estimated spectrum of NNLS algorithm is more similar to the measured one in the second example case and the Theil value

of the final solution related to the NNLS algorithm (**Model** 11-*j* in Table 5) is also slightly smaller than that of the SGSD, a "good" solution can be obtained using SGSD algorithm. In other words, as it is described in optimization problem (9), the SGSD algorithm tries to minimize both MAE and RMSE in order to find a "good" solution in terms of two measures of goodness of fit. As can be seen in Tables 7 and 8, the identification procedure of the SGSD algorithm is much more efficient than the NNLS algorithm. The SGSD algorithm can identify the correct isotopes of unknown sources in Tables 7 and 8 except for case studies 1 and 2 of Table 7 in which another isotope is identified in addition to ^{60}Co isotope. The reason for this discrepancy is the large statistical errors in recorded counts due to the small acquisition time. However, higher spectrum acquisition time leads to more accurate results as presented in case studies 1, 2, 4 and 5 of Table 7. The same inference can be obtained by comparing the obtained results of case studies 6 and 7 of Table 8. Although the results of the NNLS algorithm can be improved using SOA test, the correct isotopes of an unknown source cannot be identified using this method. For example, the ^{46}Sc and ^{57}Co isotopes are identified in addition to ^{232}Th series in **Model** 3-*j* of Table 5. The standard deviations related to these isotopes are smaller than their relative abundances. Therefore, the SOA test cannot remove the misidentified ^{46}Sc and ^{57}Co isotopes. The better results of SGSD algorithm over NNLS for calculation of fractional abundances are evident in case study 8 of Table 8. Both algorithms have identified the correct isotopes of unknown source but the SGSD algorithm calculates the fractional abundances more accurate than the NNLS algorithm. However, as shown in Fig. 5, the outputs from both algorithms reproduce the entire spectra accurately.

A novel approach for analysis of complex gamma-ray spectra using a sequential algorithm was presented in this article. The implemented SGSD (Sequential Gamma-ray Spectrum Deconvolution) algorithm exploits the advantages of the active set method and an evolutionary optimization technique to analyze the complex gamma-ray spectra. The idea of active set method is used for identification purposes, while the particle swarm optimization is applied for fitting process as presented in (Shahabinejad and Vosoughi, 2018) in the new algorithm. The results of the present algorithm were compared with the results obtained from NNLS (Non-Negative Least Square) algorithm. According to the obtained results, the SGSD algorithm produces more accurate results than NNLS algorithm for both identifying and determining the fractional abundances of isotopes in gamma-ray spectrum analysis.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.anucene.2019.04.057.

References

Alamaniotis, M., Mattingly, J., Tsoukalas, L.H., 2013. Pareto-Optimal gamma spectroscopic radionuclide identification using evolutionary computing. IEEE Trans. Nucl. Sci. 60, 2222–2231.

Alamaniotis, M., Jevremovic, T., 2015. Hybrid fuzzy-genetic approach integrating peak identification and spectrum fitting for complex gamma-ray spectra analysis. IEEE Trans. Nucl. Sci. 62, 1262–1277.

Alamaniotis, M., Choi, C.K., Tsoukalas, L.H., 2015. Application of fireworks algorithm in gamma-ray spectrum fitting for radioisotope identification. Int. J. Swarm Intell. Res. (IJSIR) 6, 102–125.

Alamaniotis, M., Tsoukalas, L.H., 2018. Assessment of Gamma-Ray-Spectra Analysis Method Utilizing the Fireworks Algorithm for Various Error Measures. Critical Developments and Applications of Swarm Intelligence, IGI Global.

Androulakaki, E.G., Kokkoris, M., Tsabaris, C., Elef Theriou, G., Patiris, D.L., Pappa, F. K., Vlastou, R., 2016. In situ γ-ray spectrometry in the marine environment using full spectrum analysis for natural radionuclides. Appl. Radiat. Isot. 14, 76–86.

Baldoncini, M., 2010. Applicazione del metodo Non Negative Least Square alla Full Spectrum Analysis nel processo di calibrazione di uno spettrometro di raggi gamma portatile. graduate thesis. University of Ferrara.

Baldoncini, M., Albéri, M., Bottardi, C., Chiarelli, E., Raptis, K.G.C., Strati, V., Mantovani, F., 2018. Investigating the potentialities of Monte Carlo simulation for assessing soil water content via proximal gamma-ray spectroscopy. J. Environ. Radioact. 192, 105–116.

Bro, R., Jong, S.D., 1997. A fast non-negativity-constrained least squares algorithm. J. Chemom. 11, 393–401.

Caciolli, A., Baldoncini, M., Bezzon, G.P., Broggini, C., Buso, G.P., Callegari, I., Colonna, T., Fiorentini, G., Guastaldi, E., Mantovani, F., Massa, G., Menegazzo, R., Mou, L., Rossi Alvarez, C., Shyti, M., Zanon, A., Xhixha, G., 2012. A new FSA approach for in situ γ ray spectroscopy. Sci. Total Environ. 414, 639–645.

Carlevaro, C.M., Wilkinson, M.V., Barrios, L.A., 2008. A genetic algorithm approach to routine gamma spectra analysis. J. Instrum. 3.

Chen, S., Montgomery, J., Bolufe-Rohler, A., 2015. Measuring the curse of dimensionality and its effects on particle swarm optimization and differential evolution. Appl. Intell. 42, 514–526.

Cichocki, A., Amari, S.I., 2002. Adaptive Blind Signal and Image Processing: Learning Algorithms and Applications. John Wiley & Sons, New York.

Eberhart, R., Kennedy, J., 1995. A new optimizer using particle swarm theory. In: Proceedings of the Sixth International Symposium on Micro Machine and Human Science, pp. 39–43.

Gardner, R.P., Sood, A., Wang, Y.Y., Liu, L., Guo, P., Gehrke, R.J., 1997. Single peak versus library least-squares analysis methods for the PGNAA analysis of vitrified waste. Appl. Radiat. Isot. 48, 1331.

Gill, P.E., Murray, W., Wright, M.H., 1981. Practical Optimization. Academic Press.

Gou, J., Lei, Y.-X., Guo, W.-P., Wang, C., Cai, Y.-Q., Luo, W., 2017. A novel improved particle swarm optimization algorithm based on individual difference evolution. Appl. Soft Comput. 57, 468–481.

Hakimabad, H.M., Panjeh, H., Vejdani-Noghreiyan, A., 2007. Evaluation the nonlinear response function of a 3×3 in NaI scintillation detector for PGNAA applications. Appl. Radiat. Isot. 65, 918–926.

Hendrix, P.H.G.M., Limburg, J., de Meijer, R.J., 2001. Full-spectrum analysis of natural γ-ray spectra. J. Environ. Radioact. 53, 365–380.

Kangas, L.J., Keller, P.E., Siciliano, E.R., Kouzes, R.T., Ely, J.H., 2008. The use of artificial neural networks in PVT-based radiation portal monitors. Nucl. Instrum. Methods Phys. Res. A 587, 398–412.

Knoll, G.F., 2000. Radiation Detection and Measurement. John Wiley & Sons, New York.

Lawson C.L., Hanson R.J., 1995. Solving Least Square Problems, Philadelphia.

Meric, I., Johansen, G.A., Holstad, M.B., Wang, J., Gardner, R.P., 2011. Produced water characterization by prompt gamma-ray neutron activation analysis. Meas. Sci. Technol. 22, 125701.

Meric, I., Johansen, G.A., Holstad, M.B., Wang, J., Gardner, R.P., 2012. On the treatment of ill-conditioned cases in the Monte Carlo library least-squares approach for inverse radiation analyzers. Meas. Sci. Technol. 23, 055603.

Morsy, Z., Abd El-Wahab, M., El-Faramawy, N., 2012. Determination of natural radioactive elements in Abo Zaabal, Egypt by means of gamma spectroscopy. Ann. Nucl. Energy 44, 8–11.

Mukherjee, B., 2002. A high-resolution neutron spectra unfolding method using the Genetic Algorithm technique. Nucl. Instrum. Methods Phys. Res. Sect. A 476, 247–251.

Oliveira, C., Salgado, J., Goncalves, I.F., Carvalho, F.G., Leitao, F., 1993. Prompt gamma neutron activation analysis of cement raw materials. Nucl. Geophys. 7, 431–443.

Olmos, P., Diaz, J.C., Perez, J.M., Garcia-Belmonte, G., Gomez, P., Rodellar, V., 1992. Application of neural network techniques in gamma spectroscopy. Nucl. Instrum. Methods Phys. Res. Sect. A 312, 167–173.

Pelowitz, D.B., 2008. MCNPXTM User's Manual – Version 2.6.0, Los Alamos National Laboratory, LA-CP-07-1473.

Phillips, G.W., 1978. Fitting peaks with very low statistics. Nucl. Instrum. Methods 153, 449–455.

Salmon, L., 1961. Analysis of gamma-ray scintillation spectra by the method of least squares. Nucl. Instrum. Methods Phys. Res. Sect. A 14, 193–199.

Shahabinejad, H., Sohrabpour, M., 2017. A novel neutron energy spectrum unfolding code using particle swarm optimization. Radiat. Phys. Chem. 136, 9–16.

Shahabinejad, H., Hosseini, S.A., Sohrabpour, M., 2016. A new neutron energy spectrum unfolding code using a two steps genetic algorithm. Nucl. Instrum. Methods Phys. Res. Sect. A 811, 82–93.

Shahabinejad, H., Vosoughi, N., 2018. Analysis of complex gamma-ray spectra using particle swarm optimization. Nucl. Instrum. Methods Phys. Res. A 911, 123–130.

Takagi, T., Sugeno, M., 1985. Fuzzy Identification of Systems and Its applications to Modeling and Control. *IEEE* Trans. Syst. Man Cybern. 15, 116–132.

Theil, H., 1967. Economics and Information Theory. Rand McNally, Chicago, IL, USA.

Yang, X., Garcia, R., 2018. Online gamma-ray spectroscopy acquisition. Ann. Nucl. Energy 115, 105–115.

Yoshida, E., Shizuma, K., Endo, S., Oka, T., 2003. Application of neural networks for the analysis of gamma-ray spectra measured with a Ge spectrometer. Nucl. Instrum. Methods Phys. Res. A 484, 557–563.