



Original Article

Full spectrum estimation of helicopter background and cosmic gamma-ray contribution for airborne measurements



Lukáš Kotík*, Marcel Ohera

National Radiation Protection Institute (SÚRO), Bartoškova 28, 140 00, Prague, Czech Republic

ARTICLE INFO

Article history:

Received 20 July 2022

Received in revised form

27 October 2022

Accepted 28 November 2022

Available online 1 December 2022

Keywords:

Airborne gamma-ray spectrometry

Aircraft background

Cosmic radiation

Large NaI(Tl) detectors

Non-negative least square method

Principal component analysis

ABSTRACT

The airborne radiation monitoring has been used in geophysics for more than forty years and now it also has its important role in emergency monitoring. The aircraft background and the cosmic gamma-rays contribute to the measured gamma spectrum on the aircraft board. This adverse effect should be eliminated before the data processing. The paper describes two semiparametric methods to estimate the full spectrum aircraft background and cosmic gamma-ray contribution from spectra measured at altitudes where terrestrial contribution is negligible. The methods only assume to know possible peak positions in spectra and their full width at half maximum, that can be easily obtained e.g. from terrestrial measurement. The methods were applied to real experimental data acquired on Mi-17 and Bell 412 helicopter boards. The IRIS airborne gamma-ray spectrometer, with 4×4 L NaI(Tl) crystals, produced by Pico Envirotec Inc., Canada, was used on helicopters' boards. To obtain valid estimate of the aircraft background and the cosmic contribution, the measurements over sea and large water areas were carried out. However, the satisfactory results over inland were also achieved comparing with those acquired over large water areas.

© 2022 Korean Nuclear Society, Published by Elsevier Korea LLC. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The modern multichannel airborne gamma spectrometers have been used for routine natural radioactivity surveys in geophysics. The basic principles of the airborne gamma spectrometry in geophysical applications were described in many publications [1–3]. In last decades after the Chernobyl disaster, the airborne gamma spectrometry was used for the emergency monitoring of man-made nuclides [4–6]. The Monte Carlo simulations were also applied for the airborne spectrometry [7–9]. Many procedures were summarized in the IAEA recommendations [10,11]. Unmanned aerial vehicles can be successfully applied for small contaminated areas [12,13], but they cannot replace aircrafts or helicopters used for the airborne surveys over large areas. The airborne emergency monitoring is focused on long-term man-made radionuclides, especially ^{137}Cs and ^{134}Cs . The airborne gamma-ray spectrum measured on board is affected by many adverse effects that need to be removed before the data processing. The objective of this paper is to eliminate one of such effects, which are the aircraft background spectrum and cosmic gamma-ray

contribution to the spectra measured on the board. The aircraft background and cosmic gamma-ray contribution is usually eliminated in three spectral windows (K, U, Th) based on the measurements over sea or large water areas [10]. This paper proposes two mathematical methods to estimate the aircraft and cosmic gamma-ray spectra based on experimental data. One method is based on non-negative weighted least squares optimization, the second is based on principal component analysis.

In majority of background measurements, there are not enough data to obtain smooth estimates of the aircraft background and cosmic gamma-ray contribution with no additional information provided. Methods proposed in this paper use artificial basis that is based on physical properties of gamma spectrum and that (approximately) generates the space of the expected value of an aircraft background spectrum and the cosmic gamma-ray contribution.

There are three main effects which contribute to the spectra measured on the aircraft board during airborne survey. The first is radon, the second is the aircraft background and the cosmic gamma-ray contribution and the last is attenuation effect due to flight altitude. The first two effects should be eliminated from the spectrum before the data processing.

* Corresponding author.

E-mail address: lukas.kotik@suro.cz (L. Kotík).

1.1. Radon background correction

Radon (^{222}Rn) is a gas which escapes from soils and rocks. Its decay products can contribute to the spectrum. This has great importance in geophysical surveys. Radon decay products can significantly arise up count rates in the uranium window [10,11]. This can affect the ground uranium activity concentration calculation. The radon removal methods are not simple, some of them are described, for example, in Refs. [14,15]. The effect of daughter products on the spectra measured and methods for estimation of their contribution is not objective of this paper. It is recommended to use data where no radon daughter products are expected to estimate the aircraft background and cosmic gamma-ray contribution.

1.2. Aircraft background and cosmic gamma-ray contribution

The second important aspect is the aircraft background and the cosmic gamma-ray contribution. This contribution should be eliminated in the spectra to avoid biased results in both geophysical surveys and emergency monitoring. The aircraft background spectrum is present on the board due to the on-board radioactivity. It is composed of different components of the radioactivity present in the aircraft (in fuselage material, fuel tank, and in the persons present on the board as the naturally occurring sources of human body) and the intrinsic spectrometer noise. In general, this background component can be considered as constant during the flight.

The cosmic background component arises from the reactions of primary cosmic radiation in the upper atmosphere. This generates the secondary radiation which reacts with the air, aircraft and the detector. The airborne spectrometers usually measure gamma-ray radiation in a range up to 3 MeV. Generally, all gamma-ray radiation higher than 3 MeV is cosmic origin. Counts from energies higher than 3 MeV are detected in the cosmic channel to estimate the cosmic gamma-ray background. The so-called cosmic channel is usually the last energy channel in the spectrum range.

If a spectrum measured on board only consists of the aircraft background and cosmic gamma-ray contribution, it is possible to estimate counts in the individual spectral windows of K, U and Th [10], i.e., the aircraft contribution (constant) and the cosmic gamma-ray contribution (dependent on altitude above sea level). The following procedure has been used for decades in the geophysical airborne surveys. The aircraft and cosmic background contributions in spectral windows have been estimated so far as follows [11]:

$$\text{CR}_{\text{BKGD}} = \text{CR}_{\text{B,Heli}} + S_c \times \text{CR}_{\text{Cos}} \quad (1)$$

where, CR_{BKGD} is the combined aircraft (in our case helicopter) and the cosmic gamma-ray background in each spectral window, $\text{CR}_{\text{B,Heli}}$ is the helicopter background in each spectral window, CR_{Cos} is the cosmic channel count rate, S_c is the cosmic stripping factor for spectral window. In compliance with [10], the coefficients $\text{CR}_{\text{B,Heli}}$ and S_c have to be determined by suitable measurements over large water areas at the altitudes from 1500 m to 3000 m with the minimum air radon concentration. The combined aircraft and cosmic gamma-ray background CR_{BKGD} is linearly related to the count rate in each spectral window, see equation (1). Thus, the linear relationship is described by the cosmic stripping factor S_c (the slope) and helicopter background $\text{CR}_{\text{B,Heli}}$ (the intercept with Y-axis) for each spectral window. Any deviation from linearity means the presence of radon in the air [10,11].

It can be expected that only helicopter background and cosmic gamma contributions are measured at flight altitudes above 1000 m–1500 m above ground level (hereinafter referred to as

AGL). The terrestrial contribution at these altitudes is negligible.

Two types of helicopters are used for the airborne radiation monitoring by the Emergency Preparedness Division at the National Radiation Protection Institute in the Czech Republic. The type of helicopter plays a significant role on the background effect, especially for its size, inner space and fuel tank position. The Czech airborne team uses Mi-17 helicopters (NATO code ID NATO HIP–H) and Bell 412 helicopters. Both helicopter types have their fuel tanks out of the field of view of the airborne spectrometer, which eliminates the variable attenuation effect.

Formula (1) can be extended to the individual energy channels in the full spectrum. However, this approach needs a longer time interval for acquiring the data at the appropriate cruising levels to achieve the smoothed spectrum. Unfortunately, this is usually not possible for many reasons – in particular, helicopter fuel consumption needed for the following airborne survey itself, flight safety, etc. Usually, 5-min spectra can be acquired at each of different altitudes before the airborne survey itself. This gave the idea to use the mathematical methods that incorporates physics knowledge (by applying a physics-based basis used as the regression design matrix) to reduce uncertainty.

2. Material and methods

2.1. Airborne IRIS gamma-ray spectrometer

The experiments were performed with IRIS gamma-ray spectrometer. The spectrometer was provided with four NaI(Tl) crystals with dimensions of $10 \times 10 \times 40$ cm. The spectrum in IRIS was always energy calibrated from 10 keV to 3 MeV and stabilized by the software algorithm according to ^{40}K (1460 keV) and ^{208}Tl (2614 keV) peak positions. The spectrometer has 512 channels. The last channel (channel 512) was the cosmic channel in which all gamma events in the energy range from 3 MeV to approximately 7 MeV were detected.

2.2. Methods for estimation of the aircraft background and cosmic gamma-ray contribution

Two relatively simple to implement methods are proposed. The first method assumes that the multiple spectra are recorded at two or more different altitude levels. The second method does not need such assumption, but it usually provides results with higher uncertainty.

In what follows, the vectors and matrices are written in bold font and arithmetical operations (such as addition, subtraction, division, etc.) applied on them are considered as element-wise.

2.2.1. Basis that generates the space of mean aircraft spectra and cosmic gamma-ray contribution

Suppose that each expected value of each recorded spectrum can be written as linear combination of:

1. Vectors in \mathbb{R}^K that represent (scaled) energy peaks for each radionuclide that can be found in aircraft spectrum. The peaks, sometimes called the primary components, are constructed in the form

$$e^{-(x-\mu)^2/d^2} \quad (2)$$

where x represents the energy or the energy channel ($1, \dots, K$), the parameter $\mu \in \mathbb{R}$ controls the location of the peak and the parameter $d > 0$ controls the width of the peak (full width at half maximum height is approximately $2.355 d$). Elements of such

vectors are mainly close to zero except for the surroundings of the energy (or energy channel) μ . We denote the matrix formed by these columns vectors as $\mathbf{P}_b \in \mathbb{R}^{K \times p_b}$, where p_b is the number of peaks (i.e. the number of vectors that represents peaks). The basis with peaks, primary components, is illustrated in Fig. 1, where $p_b = 8$.

2. Vectors in \mathbb{R}^K that represent energy peaks for cosmic gamma-ray contribution. They are constructed in the same way as the vectors in paragraph 1. The matrix formed by these column vectors is denoted by $\mathbf{P}_c \in \mathbb{R}^{K \times p_c}$. It is usually considered only one peak of cosmic gamma-ray background ($p_c = 1$). It is shown as dashed red peak in Fig. 1. This peak is also considered as the primary component.
3. Vectors in \mathbb{R}^K whose non-negative linear combinations generate the space of possible expected values of aircraft background spectra (respectively cosmic gamma-ray spectra) with primary components in paragraph 1. (respectively 2.) removed. The spectra without primary components are in the following text called secondary components. It is assumed that the secondary components (i.e. Compton scattering) are non-increasing except for the interval with low energies (e.g., energies <117 keV). Hence, only subset of channels (energies) where these vectors are non-increasing (e.g., channels 20–511) are considered. The remaining elements are left as undefined. This basis is constructed from I-spline basis (integral of M-spline) with the intercept included in the basis, [16]. The I-spline basis is formed by non-decreasing functions with the maximum value of one. The basis used to fit the secondary components is then constructed as

$$\mathbf{V} = \mathbf{J}_{K,p} - \mathbf{S}_p \in \mathbb{R}^{K \times p}, \tag{3}$$

where $\mathbf{J}_{K,p} \in \mathbb{R}^{K \times p}$ denotes a matrix of ones and $\mathbf{S}_p \in \mathbb{R}^{K \times p}$ denotes the matrix with I-spline basis with p columns. The basis is illustrated in Fig. 2.

The parameters from equation (2) are chosen such that the location (μ) and the width (d) of a peak corresponds to the data measured in NaI(Tl) crystals of IRIS spectrometer. The results of the paper are calculated for I-spline basis considered for interval from 20th channel to 511th channel (the elements outside this interval are set as undefined), with the polynomial degree of three,

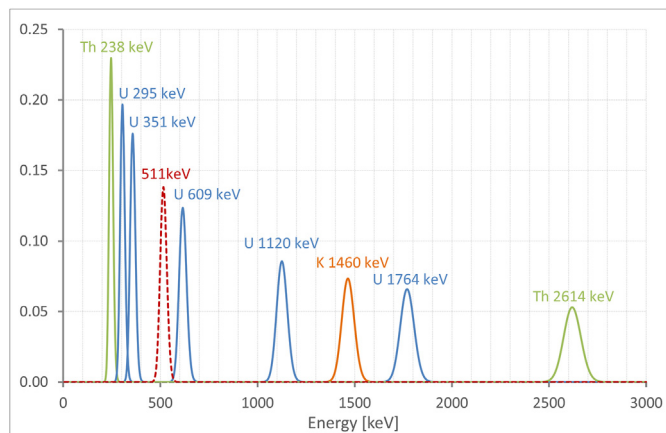


Fig. 1. Basis for primary components – expected peaks in background helicopter spectrum. Dashed peak of 511 keV is the only one peak in the cosmic gamma-ray spectrum. Position (parameter μ) and full width at half maximum (respectively parameter d) corresponds to data measured in NaI(Tl) crystals of IRIS spectrometer.

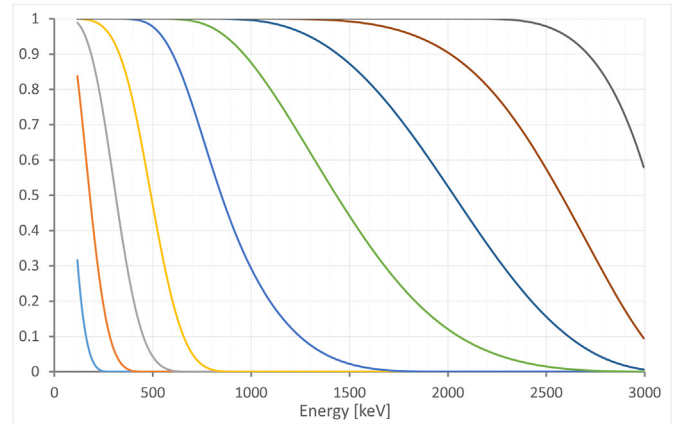


Fig. 2. Basis for secondary components from paragraph 3 in Section 2.2.1 based on I-spline.

boundary knots in channels 10 and 550 (59 and 3220 keV) and internal knots in channels 50, 80, 120, 160, 350 (293, 469, 703, 938 and 2051 keV). It follows that in this particular case $p = 9$ (the number of internal knots + polynomial degree + one for intercept), hence $\mathbf{V} \in \mathbb{R}^{K \times 9}$. This basis is shown in Fig. 2. The optimal selection of knots (its position and its number) is not practically feasible in the most cases because of lack of enough training data. The general advice is to choose a number and the positions of knots such that basis avoid overfitting but still provides enough flexibility to fit the secondary components. Knots in the paper were chosen based on Mi-17 data. Good results could be obtained with 5 internal knots where distance between knots increases with energy. It is important to keep first three internal knots relatively close to each other because there is most significant change in the shape of secondary components. Boundary knots should be chosen outside the range for channels (respectively energies) that are considered to fitting even though it does not need to have physical meaning.

Decomposition of a spectrum to a primary component and to a secondary component is illustrated in Fig. 3.

2.2.2. Non-negative weighted least squares method (NN-WLSQ)

The proposed methods use the non-negative weighted least squares optimization to find estimates of regression parameters. The method produces only non-negative estimates of regression coefficients. This paragraph presents the summary of this method.

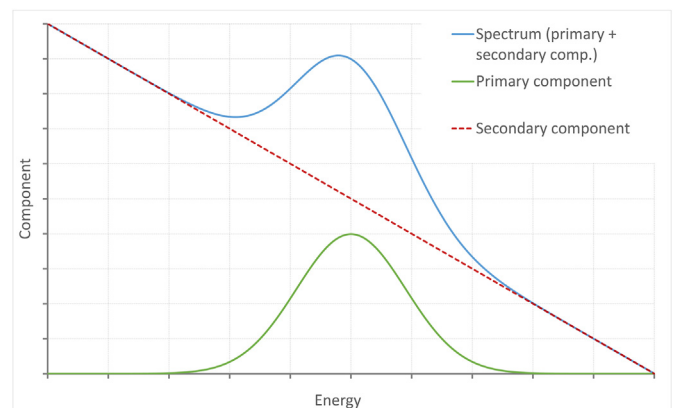


Fig. 3. Illustration of decomposition of a spectrum to primary (Gaussian peaks) and secondary (scattering) components. The spectrum is the sum of primary and secondary components.

Denote by $\mathbf{Y} = (y_1, \dots, y_n)^T$ a vector of dependent variable, assume a matrix with regressors (the design matrix) $\mathbf{X} \in \mathbb{R}^{n \times p}$, denote by \mathbf{x}_i the i^{th} row of matrix \mathbf{X} and denote by $\mathbf{w} = (w_1, \dots, w_n)^T$ a vector with weights, then the non-negative weighted least squares problem is defined as

$$\arg \min_{\beta \geq 0} \sum_i w_i (Y_i - \mathbf{x}_i^T \beta)^2. \quad (4)$$

2.2.3. Notation for method based on non-negative weighted least squares

Suppose that spectra were recorded at $H \geq 2$ different altitudes where no additional gamma ray sources other than aircraft background and cosmic are present. Denote by

1. y_{ijk} recorded number of counts in energy channel $k, k = 1, \dots, K$, of spectrum $j = 1, \dots, n_i$ at altitude $h_i, i = 1, \dots, H$.
2. c_{ij} cosmic channel count rate for spectrum $j = 1, \dots, n_i$ at altitude $h_i, i = 1, \dots, H$.

This method does not need to use each individual spectrum. It only suffices to work with the average spectra at each altitude $h_i, i = 1, \dots, H$:

$$\bar{\mathbf{y}}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} \mathbf{y}_{ij} \in \mathbb{R}^K, \quad (5)$$

where $\mathbf{y}_{ij} \in \mathbb{R}^K$ is j^{th} recorded spectrum at altitude h_i . The average count rates in the cosmic gamma-ray channel for each altitude $h_i, i = 1, \dots, H$, are also used for calculations:

$$\bar{c}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} c_{ij}. \quad (6)$$

2.2.4. Assumptions for method based on non-negative weighted least squares

The proposed method is based on the following assumptions:

1. Recorded counts follow the Poisson distribution, $y_{ijk} \sim Po(\lambda_{ik})$. λ_{ik} is the expected value of count rate in the energetic channel k of a spectrum at altitude h_i .
2. y_{ijk} and y_{lmn} are independent random variables (if $(i, j, k) \neq (l, m, n)$).
3. $c_i, c_j, i \neq j$ are independent random variables that follow the Poisson distribution.
4. The altitudes h_1, \dots, h_H are not close to each other. In other words, the cosmic gamma-ray contributions for each altitude are distinguishable from each other. In ideal case, the differences between expected values of cosmic channel count rates at different altitudes are high and simultaneously variances of calculated means are low. The paper uses the following rules that are based on the coefficients of variations:

$$\frac{1}{\sqrt{n_i \bar{c}_i}} \leq \delta_1, i = 1, \dots, H \quad (7)$$

and

$$\frac{\sqrt{\bar{c}_i/n_i + \bar{c}_l/n_l}}{|\bar{c}_i - \bar{c}_l|} \leq \delta_2, i \neq l. \quad (8)$$

The left side of (7) is an estimate of coefficient of variation of average (and also total) cosmic channel count rates for altitude h_i . The left side of (8) is an estimate of coefficient of variation of difference of averages at altitude h_i and h_l . The parameters δ_1, δ_2 are upper limits for uncertainty. $\delta_1 = 0.005$ and $\delta_2 = 0.03$ could lead to acceptable results (based on the analysis of datasets for two types of helicopters).

5. There is no other systematic contribution to spectra than the aircraft background and the cosmic gamma-ray contribution.

2.2.5. Algorithm for method based on non-negative weighted least squares

The proposed method can be considered as generalization of the spectral windows method described in Section 1.2. Denote by $\lambda_i \in \mathbb{R}^K$ the vector of the expected value of a vector with recorded number of counts at altitude h_i (the expected spectrum at altitude h_i). It holds that:

$$\lambda_i = \mathbf{b} + \gamma_i \mathbf{c}, \quad (9)$$

where:

1. $\mathbf{b} \in \mathbb{R}^K$ is the expected value of the aircraft background (i.e. mean counts for the aircraft background)
2. $\mathbf{c} \in \mathbb{R}^K$ is the expected value of spectrum with cosmic gamma-ray contribution for unit cosmic channel count rate
3. $\gamma_i > 0$ is expected value of cosmic channel count rate at altitude h_i .

It follows that $\gamma_i \mathbf{c}$ is the expected value of cosmic contribution at altitude h_i . The proposed method provides the estimates of vectors \mathbf{b} and \mathbf{c} . Since spectra and the elements of basis from Section 2.2.1 are non-negative, the estimation is done with NN-WLSQ (4).

The vector of dependent variable for NN-WLSQ is:

$$\mathbf{Y} = \begin{pmatrix} \bar{\mathbf{y}}_1 \\ \vdots \\ \bar{\mathbf{y}}_H \end{pmatrix}. \quad (10)$$

The structure of the matrix with regressors (design matrix) is based on the assumption that there exist vectors with non-negative elements $\beta_{Vb}, \beta_{Vc} \in \mathbb{R}^p, \beta_{Pb} \in \mathbb{R}^{p_b}, \beta_{Pc} \in \mathbb{R}^{p_c}$ such that:

$$\mathbf{b} = \mathbf{V} \beta_{Vb} + \mathbf{P}_b \beta_{Pb}, \quad (11)$$

$$\mathbf{c} = \mathbf{V} \beta_{Vc} + \mathbf{P}_c \beta_{Pc}. \quad (12)$$

It follows from (9), (11) and (12) that expected value of vector $\bar{\mathbf{y}}_i$ is equal to:

$$\mathbf{E} \bar{\mathbf{y}}_i = \lambda_i = \mathbf{V} \beta_{Vb} + \mathbf{P}_b \beta_{Pb} + \gamma_i \mathbf{V} \beta_{Vc} + \gamma_i \mathbf{P}_c \beta_{Pc}. \quad (13)$$

To construct the design matrix for NN-WLSQ we replace the unknown parameters γ_i with their estimates \bar{c}_i . The matrix, \mathbf{X} , is therefore in the form:

$$\mathbf{X} = \begin{pmatrix} \mathbf{V} & \mathbf{P}_b & \bar{c}_1 \mathbf{V} & \bar{c}_1 \mathbf{P}_c \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{V} & \mathbf{P}_b & \bar{c}_H \mathbf{V} & \bar{c}_H \mathbf{P}_c \end{pmatrix}. \quad (14)$$

Note that although \mathbf{X} has full column rank it suffers from multicollinearity that increases the variance of the estimate. This effect is attenuated by the data assumption 4 from Section 2.2.4.

The variances of elements of $\bar{\mathbf{y}}_i$ are not the same. There are relatively high differences between these variances. In this case, the

weighted least squares method produces the estimates with lower (the lowest possible if proper weight vector is used) variance. Theoretical vector of weights is in the form:

$$\mathbf{w} = 1/\mathbf{v} \tag{15}$$

where

$$\mathbf{v} = \begin{pmatrix} \lambda_1/n_1 \\ \vdots \\ \lambda_H/n_H \end{pmatrix}. \tag{16}$$

Vectors λ_i are unknown. It can be replaced with the moving average of \mathbf{y}_i or with the fitted values produced by the multistage application of the proposed method: first, the estimation is done with weights determined from moving averages. Then, the estimates are used to calculate fitted (smoothed) spectra and the weights are determined from the fitted spectra, (20). Finally, the method is applied using this weight vector to obtain the final estimate.

Vector \mathbf{v} or its estimate is usually very close to zero for some energy channels. To make method numerically stable, the authors of the paper recommend to add an offset to vector \mathbf{v} . The results in the paper are calculated for offset 0.5, thus for the weight vector $\mathbf{w} = 1/(\mathbf{v}+0.5)$.

The application of the non-negative weighted least square method for the vector of dependent variable \mathbf{Y} , (10), the design matrix \mathbf{X} , (14), and the vector of weights (15) produces the estimate of regression parameters:

$$\hat{\boldsymbol{\beta}} = \begin{pmatrix} \hat{\beta}_{b1} \\ \hat{\beta}_{b2} \\ \hat{\beta}_{c1} \\ \hat{\beta}_{c2} \end{pmatrix} \in \mathbb{R}^{p+p_b+p+p_c}. \tag{17}$$

Each part of this vector corresponds to one column block of the design matrix \mathbf{X} , (14).

The estimates of aircraft background and cosmic contribution for unit cosmic channel count rate are:

$$\hat{\mathbf{b}} = \mathbf{V}\hat{\boldsymbol{\beta}}_{b1} + \mathbf{P}_b\hat{\boldsymbol{\beta}}_{b2}, \tag{18}$$

$$\hat{\mathbf{c}} = \mathbf{V}\hat{\boldsymbol{\beta}}_{c1} + \mathbf{P}_c\hat{\boldsymbol{\beta}}_{c2}. \tag{19}$$

The cosmic gamma-ray contribution at the altitude h_i is estimated as $\bar{c}_i\hat{\mathbf{c}}$. Finally, the fitted (smoothed) spectra for altitude h_i is calculated as:

$$\hat{\lambda}_i = \hat{\mathbf{b}} + \bar{c}_i\hat{\mathbf{c}}. \tag{20}$$

2.2.6. Method based on principal component analysis (PCA)

The main assumption for this method is that spectra were recorded at different altitudes and, if possible, no repeated measurements at the same altitude are realized. An ideal case if spectra are recorded with gradually increasing altitude. This method uses similar concepts, the same notation and the same assumptions as the method based on non-negative least squares, except that there is no need to use notation for altitudes and also there is no need to make assumption about altitudes (data assumptions 3 and 4 in Section 2.2.4). Thus, the notation of altitude h_i can be ignored.

Suppose n recorded spectra $\mathbf{y}_1, \dots, \mathbf{y}_n \in \mathbb{R}^K$ and their corresponding cosmic channel count rates c_1, \dots, c_n . Analogously to the method based on non-negative least squares it is assumed that for

spectrum $i = 1, \dots, n$ it holds $\mathbf{y}_i = (y_{i1}, \dots, y_{iK})^T$, $y_{ik} \sim Po(\lambda_{ik})$ and that y_{ij}, y_{kl} are independent random variables for $(i,j) \neq (k,l)$. Denote by $\lambda_i \in \mathbb{R}^K$ vector $(\lambda_{i1}, \dots, \lambda_{iK})^T$. Further assume $c_i \sim Po(\gamma_i)$, $i = 1, \dots, n$, where γ_i is the expected value of cosmic channel count rate for i^{th} recorded spectrum.

It holds that:

$$\lambda_i = \mathbf{b} + \gamma_i\mathbf{c}, i = 1, \dots, n. \tag{21}$$

The average spectrum:

$$\bar{\mathbf{y}}_n = \frac{1}{n} \sum_{i=1}^n \mathbf{y}_i \tag{22}$$

has the expected value:

$$E\bar{\mathbf{y}}_n = \mathbf{b} + \bar{\gamma}_n\mathbf{c} \tag{23}$$

where $\bar{\gamma}_n$ is the average of $\gamma_1, \dots, \gamma_n$. It follows that variables $\mathbf{z}_i \in \mathbb{R}^K$ defined as:

$$\mathbf{z}_i = \mathbf{y}_i - \bar{\mathbf{y}}_n, i = 1, \dots, n \tag{24}$$

can be expressed in the form $\mathbf{z}_i = (\gamma_i - \bar{\gamma}_n)\mathbf{c} + \mathbf{e}_i$, where \mathbf{e}_i is a vector of residual errors. This error vector has expected value equal to the vector of zeros. Its covariance matrix is approximately equal to $diag\{\lambda_i\}$. The variables \mathbf{z}_i would lie along the line in K dimensional space that goes through the origin in the direction given by vector \mathbf{c} . The deviations from this line are caused by error vectors \mathbf{e}_i . Hypothetically, if $\mathbf{e}_i = \mathbf{0}$ then \mathbf{z}_i lies on this line.

The vector with the cosmic gamma-ray contribution \mathbf{c} for unit cosmic channel count rate can be estimated by the application of the principal component analysis (in geophysics often called NASVD, [17]). Since the variance is not constant throughout energetic channels $1, \dots, K$, it is generally recommended to scale the elements of vectors \mathbf{z}_i with respect to their variance (noise adjustment). The method is thus applied on the data:

$$\mathbf{z}_i / \sqrt{\bar{\mathbf{y}}_n}, i = 1, \dots, n \tag{25}$$

The theoretical first (and only one non-noisy) eigenvector (denote it by $\mathbf{u}, \|\mathbf{u}\| = 1$) after transform $\mathbf{u}\sqrt{\bar{\mathbf{y}}_n}$ (elementwise multiplication) is supposed to be in the form $\alpha_1\mathbf{c}$, where $\alpha_1 \in \mathbb{R}$ is a constant whose absolute value depends on $\|\mathbf{c}\|$.

The estimate of the expected value of spectrum with cosmic gamma-ray contribution for unit cosmic channel count rate is:

$$\hat{\mathbf{c}} = \mathbf{u}\sqrt{\bar{\mathbf{y}}_n}/\alpha_1, \tag{26}$$

where the value of parameter α_1 is estimated from the linear regression problem:

$$c_i = \alpha_0 + \alpha_1 s_i + \xi_i, i = 1, \dots, n \tag{27}$$

where α_0 is the intercept (its least square estimate is approximately equal to average of $c_i, i = 1, \dots, n$), s_i are the principal component scores for the first eigenvector (s_i is the coordinate of $\mathbf{z}_i/\sqrt{\bar{\mathbf{y}}_n}$ with respect to the first eigenvector \mathbf{u}), ξ_i are random independent errors. Value of the parameter α_1 from (27) is estimated using the weighted least squares method with weights $1/c_i, i = 1, \dots, n$.

It follows from (23) that the estimate of the aircraft background is:

$$\hat{\mathbf{b}} = \bar{\mathbf{y}}_n - \bar{c}_n\hat{\mathbf{c}}, \tag{28}$$

where \bar{c}_n is the average of $c_i, i = 1, \dots, n$.

Both estimates (26) and (28) are not smooth enough for usual sample sizes (about 1000 spectra). Smooth estimates can be obtained in a similar way as is used for the first method. The smooth estimate of the aircraft background, $\hat{\mathbf{b}}_s$, is obtained from non-negative least squares regression for the vector of dependent variable $\hat{\mathbf{b}}$ and the design matrix $(\mathbf{V}, \mathbf{P}_b)$ as:

$$\hat{\mathbf{b}}_s = \mathbf{V}\hat{\boldsymbol{\beta}}_{b1} + \mathbf{P}_b\hat{\boldsymbol{\beta}}_{b2}, \quad (29)$$

where $\hat{\boldsymbol{\beta}} = (\hat{\boldsymbol{\beta}}_{b1}^T, \hat{\boldsymbol{\beta}}_{b2}^T)^T$ is the estimate of regression parameters. In this case, no weights are used. The smooth estimate of the cosmic gamma-ray contribution is calculated analogously for the dependent variable $\hat{\mathbf{c}}$ and the design matrix $(\mathbf{V}, \mathbf{P}_c)$.

The better results are obtained if the first energy channels are omitted (e.g., the first 20 channels, i.e. energy <117 keV). To obtain the best results, the spectra should be recorded at the altitudes where the signal from cosmic channel count rate is variable enough (e.g., the coefficient of variation of recorded cosmic channel count rates is at least 20% for the average cosmic channel count rate of 200 cps).

2.2.7. Uncertainty calculation

The covariance matrix of estimates (17) can be estimated using the theory of linear regression modelling. Main assumption is that the covariance matrix of solution to NN-WLSQ can be approximated by the covariance of WLSQ. Once having the covariance matrix of (17) the calculation of standard deviations of many characteristics of the helicopter background and cosmic contribution is straightforward (using property $\text{var}(\mathbf{a}^T \mathbf{X}) = \mathbf{a}^T \text{var}(\mathbf{X}) \mathbf{a}$), because (18), (19) or their sum on an interval are linear combinations of vector (17). This approach ignores the uncertainty of cosmic channel count rates. It is also quite sensitive to cases when the proposed model does not describe the data well (e.g. location of primary components in basis and in data is not the same). Further, because of very small mean count rates for channels with higher energies, the assumption about at least approximative normal error does not hold.

Another way is so called parametric bootstrap – estimates are obtained by simulating data from the model assuming that real values of unknown parameters are equal to their estimates. The spectra are simulated using Poisson distribution. The statistic of interest is then estimated from each simulated dataset. This way the distribution of the statistic of interest can be estimated. Uncertainty of cosmic channel count rates can be incorporated into the analysis. This method can also be used to uncertainty estimates for the PCA based method or to estimate how a deviation from an assumption affects the results.

3. Experimental data

The data from different airborne measurements described below were used to verify the mathematical approach for calculating the aircraft background and the cosmic gamma-ray contribution. All measurements were carried out on the same type of Mi-17 helicopter with the same equipment on the board and nearly the same crew. Table 1 shows the geographical data of the airborne measurements over large water areas and some of the airborne measurements in the Czech inland. The method was also applied on the data acquired with Bell 412 helicopter, however, there is no comparison with the data taken over large water areas. In all cases the measuring times, i.e. the number of 1-s spectra, were 300 at each flight altitude ($n_i = 300$). The sample size is mainly determined by the maximum possible time interval that the aircraft could spent at cruising levels h_i .

3.1. Altitude measurements in France

Five flight levels (300, 500, 750, 1000 and 2000 feet, i.e., 91, 152, 229, 305 and 609 m) were chosen for the measurements over the Mediterranean Sea during the Airborne Gamma Spectrometry Campaign – AGC France in 2019 (Table 1). The measurements were arranged approximately 16 km far from the French coast from the maximum to minimum flight altitude. The distance of 16 km from the coast guaranteed the zero effect of the terrestrial components and the negligible effect of radon. The measuring times on the individual flight levels were at least 5 min while the pilots tried to keep the constant flight altitudes over sea.

3.2. Altitude profile measurements in Switzerland

The International Airborne Radiation Monitoring Campaign was held in Zürich in June 2017 [18], Switzerland. Participants were teams from Switzerland, Germany, France and the Czech Republic. There was a possibility to carry out altitude profile measurements over Zugersee (English: Lake Zug). The Zugersee is a lake in Central Switzerland, situated between Vierwaldstättersee (Lake Luzerne) and Zürichsee (Lake Zürich). The central part of the lake has an area of 3.7 km × 7.0 km. The water level is 417 m above the sea level. This environment gives good conditions for the determination of the helicopter background and cosmic gamma-ray contribution to the spectrum in a range up to 3 MeV. The spectra were taken over the central part of the lake at the altitudes of 90 m, 180 m, 300 m, 600 m, 1200 m 1800 m and 2400 m above the water level (approx. from 500 m to 2800 m above sea level) while only the data of three highest flight altitudes were applicable. Mi-17 helicopter was hovering for 5 min (at least 300 recorded spectra) at each altitude (Table 1).

3.3. Altitude measurements in Czech Republic

Similar experiments were carried out in the Czech Republic over landscape and the results were compared to those acquired in Switzerland (Table 1). Measurements were carried out at the flight altitudes approximately from 2000 m to 3100 m above sea level (ASL). The detailed information is given in Table 1. The experiments were performed with Mi-17 helicopter, ID 0849. The IRIS gamma-ray spectrometer, the crew on board, etc. were identical as in Switzerland. The Mi-17 spectra were calculated in the same way as before.

Two experiments were also carried out with Bell 412 helicopter ID OK BYN near the Vlašim Public Airport, see Table 1. Only two operators and two crew members were present on the helicopter board.

4. Results

First of all, the spectra of cosmic radiation and the Mi-17 helicopter background spectra up to 3 MeV were calculated from the spectra acquired over Zugersee water area based on both mathematical approaches described above. Fig. 4 shows the cosmic gamma-ray contribution S_c calculated from the spectra acquired over Zugersee by both NN-WLSQ and PCA methods. Cosmic gamma-ray contributions S_c were calculated from three altitude datasets as well as two altitude datasets over Zugersee. The results from PCA method are also included. Note that this dataset is not optimal for use of PCA method because majority of spectra were recorded in only three different altitude levels. The Mi-17 helicopter background spectra calculated for the same datasets and by two methods (NN-WLSQ and PCA) are shown in Fig. 5. As seen from figures, the spectra calculated are nearly identical.

Table 1
Basic geographical data of the background measurements in France, Switzerland and Czech Republic.

Location	GPS position	Water (ground) level above sea level	Avgr. altitudes above sea level	Measuring time per altitude
France (Mi-171 ID 9904)	43.2060N, 4.8972E	Mediterranean Sea (sea)	100 up to 700 m	300 s
Switzerland (Mi-17 ID 0834)	47.1504803N, 8.4833119E	417 m (water level); Zugersee (Lake Zug)	507, 597, 717, 1017, 1617, 2217 and 2817 m	300 s
CZ-Vysočina (Mi-17 ID 0839)	49.430764N, 16.055372E 49.432796N, 15.899040E 49.449088N, 15.773195E	approx. 600 m (ground)	2059 m 2490 m 3023 m	240 s
CZ-Brdy (Mi-17 ID 0839)	49.597313N, 13.606682E 49.581171N, 13.420710E 49.564342N, 13.183049E	approx. 450–500 m (ground)	2215 m 2607 m 3167 m	300 s
CZ – Vlašim (Bell-412HP)	49.727143N, 14.880114E	approx. 350–500 m (ground)	2480 m 2782 m 3410 m	300 s
CZ – Vlašim (Bell 412HP)	49.727143N, 14.880114E	approx. 350–500 m (ground)	2177 m 3092 m	300 s

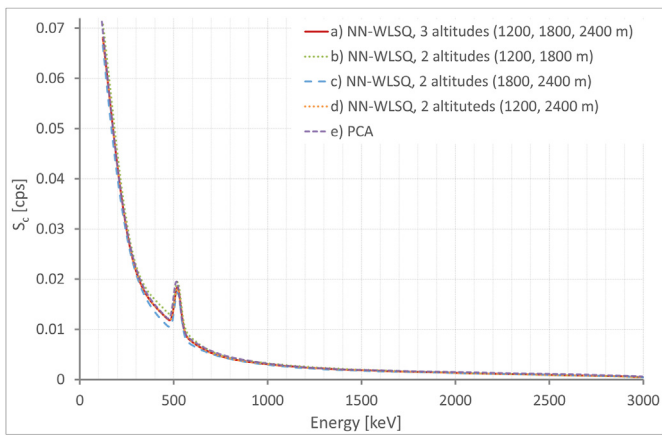


Fig. 4. Cosmic gamma-ray contribution S_c calculated by the algorithms from the spectra acquired over Zugersee (2017) at different altitudes above water level.

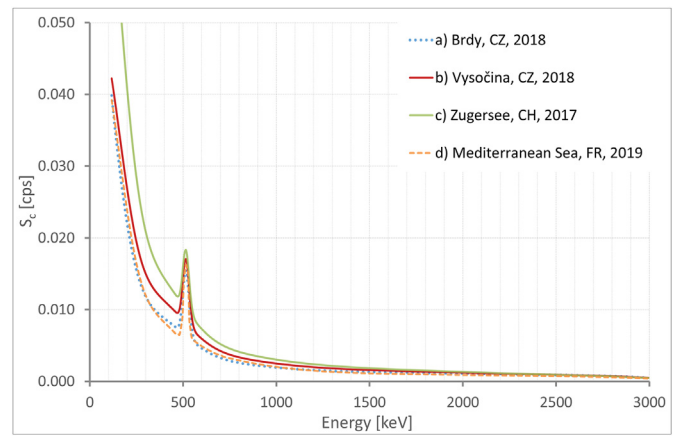


Fig. 6. Comparison of cosmic gamma-ray contributions S_c calculated by the algorithm (NN-WLSQ, 3 altitudes) from the spectra acquired at different locations: a) Brdy, Czech Republic, 2018; b) Vysočina, Czech Republic, 2018; c) Zugersee, Switzerland, 2017; d) France, Mediterranean Sea, 2019.

The same procedures (only NN-WLSQ) were also applied to all high altitude spectra of the different airborne projects in the Czech Republic. Figs. 6 and 7 compare the results from the datasets measured over the landscape in the Czech Republic and over water areas. Also, the cosmic gamma-ray contribution S_c and the Mi-17 helicopter background spectrum calculated from the datasets over the Mediterranean Sea are included. With the exception of the

background helicopter spectrum measured over the Brdy hills, 2018, (see Discussion for more details) the helicopters' spectra are relatively close to each other. Fig. 8 compares the Mi-17 background spectra calculated over Zugersee from three altitudes and Bell-

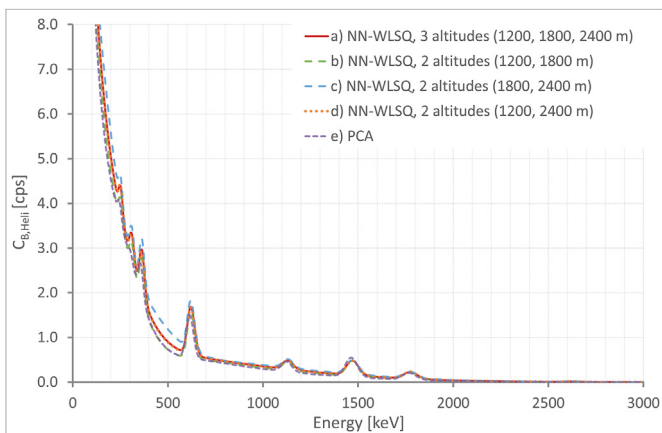


Fig. 5. Mi-17 background spectra calculated by the algorithms from the spectra acquired over Zugersee (2017) at different altitudes above water level.

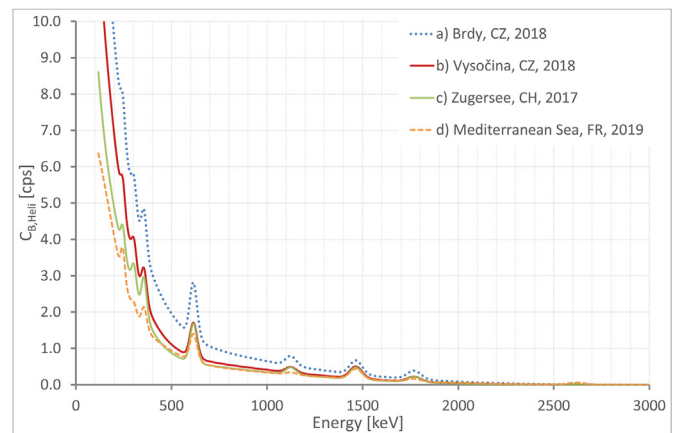


Fig. 7. Comparison of Mi-17 background spectra calculated by the algorithm (NN-WLSQ, 3 altitudes) from the spectra acquired at different locations: a) Brdy, Czech Republic, 2018; b) Vysočina, Czech Republic, 2018; c) Zugersee, Switzerland, 2017; d) France, Mediterranean Sea, 2019.

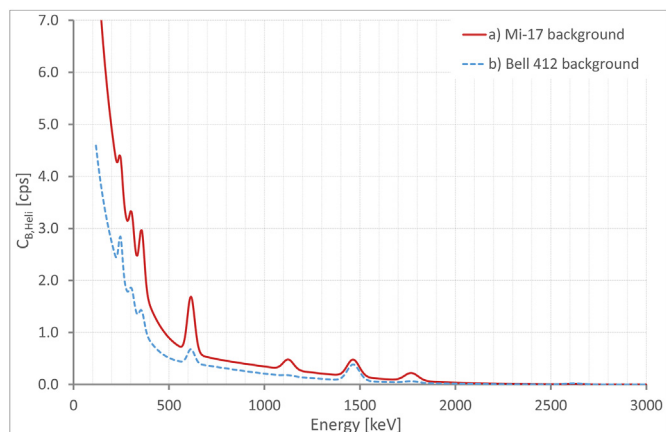


Fig. 8. Comparison of a) Mi-17 and b) Bell 412 background gamma-ray spectra estimated by the NN-WLSQ algorithm (Mi-17 - 3 altitudes, Zugersee; Bell 412 - 3 altitudes Vlasim, Czech Republic).

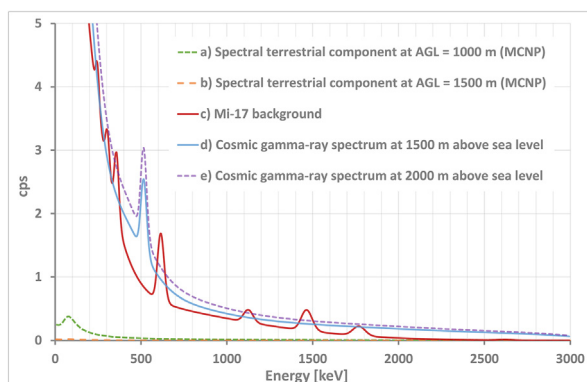


Fig. 9. Monte Carlo simulated terrestrial spectral component at altitudes of 1000 m and 1500 m AGL (1500 m and 2000 m above sea level), (a), b)), compared to Mi-17 estimated background (c), and to cosmic gamma-ray spectrum at 1500 m and 2000 m above sea level (d), e).

412HP helicopter background spectrum also calculated from two and three altitudes (Vlasim, 2020). This helicopter is somewhat smaller compared to Mi-17 which is also seen in the spectral responses. Fig. 9 compares estimated Mi-17 background, cosmic contribution for 1500 m and 2000 m above sea level with MCNP simulated spectra at AGL 1000 m (1500 m above sea level) and AGL 1500 m (2000 m above sea level).

The aircraft background $C_{B,Heli}$ and the cosmic gamma-ray contribution S_c from formula (1) in spectral windows ^{137}Cs , ^{40}K , ^{214}Bi (U-series) and ^{208}Tl (Th-series) were also estimated based on the spectra calculated on the algorithm described above. The

results of the extended window method used mainly in geophysical airborne surveys are shown in Table 2. The standard deviations are calculated using parametric bootstrap (see section 2.2.7). For ^{214}Bi and ^{208}Tl the distribution of error is not symmetric hence classical interpretation of $\pm\sigma$ is not suitable.

Finally, the air kerma (dose) rates as a contribution from the helicopter background and cosmic gamma radiation in a range up to 3 MeV were calculated. The air dose rate calculation was based on the IRIS NaI(Tl) detector calibration using the relative absorbed energy [19]. The Mi-17 helicopter contribution to the air dose rate is about 2.6 nGy h^{-1} while Bell 412 helicopter contributes with approximately 1.5 nGy h^{-1} on the board. The cosmic gamma radiation up to 3 MeV contributes from some 2.0 nGy h^{-1} to 4.3 nGy h^{-1} depending upon the flight altitudes over the Czech Republic (from 200 m to 1800 m AGL).

5. Discussion

Although the estimated background and cosmic spectra are smooth, it does not necessarily mean that there is no uncertainty in the estimates. The differences that are visible in Figs. 6 and 7 are partly caused by uncertainty. Having limited flight hours, the background measurements as well as the airborne survey have to be carried out during the same flight. Limited time to collect all background spectra that can be used for the estimation of the aircraft spectra and the cosmic gamma-ray contribution is about 20 min. This usually represents about 1200 spectra for three flight altitudes. Such sample size is not high enough to obtain smooth estimates. Because of high uncertainty of input data, the estimation with no additional information could easily lead to estimates with false peaks, with wrong shape or wrong position of the peaks. Using the basis with primary components can be considered as providing additional physically based information to obtain smoother estimates with lower uncertainty. On the other side if the wrong basis is used (e.g., the basis with wrong location of peaks) the methods can lead to biased results. Moreover, the other variables can have influence on the measured spectra (e.g., possible presence of radon daughter products). However, the presence of radon decay product was not evaluated during the flights, with only possible exception, that is the survey over the Brdy hills.

The I-spline basis is good for fitting monotonic functions. There is usually no need to estimate the aircraft background and the space contribution for the low energy channels (less than 120 keV). If the energy channels at the beginning are omitted, the secondary component are considered as non-increasing. If there is a need for estimates covering the interval of energy channels at the beginning, the basis for primary components has to be extended with elements that can cover the possible shape of primary components in this interval.

Both methods use the NN-WLSQ (4) to estimate the regression parameters because the basis is constructed in a way that spectra

Table 2

Stripping factors S_c and helicopter background $CR_{B,Heli}$ (resp. estimates of their standard deviation) for extended window method. Values are calculated for Zugersee data (Mi-17) and Vlasim data (Bell412).

Type	Method	$CR_{B,Heli}$			
		^{137}Cs (618–705 keV)	^{40}K (1370–1570 keV)	^{214}Bi (1660–1860 keV)	^{208}Tl (2410–2810 keV)
Mi-17	NN-WLSQ	14.0 (0.41)	9.8 (0.33)	4.9 (0.22)	0.4 (0.17)
Bell412	NN-WLSQ	8.3 (0.48)	8.6 (0.43)	2.0 (0.28)	0.6 (0.17)
Type	Method	S_c			
		^{137}Cs (618–705 keV)	^{40}K (1370–1570 keV)	^{214}Bi (1660–1860 keV)	^{208}Tl (2410–2810 keV)
Mi-17	NN-WLSQ	0.0957 (0.0021)	0.0662 (0.0016)	0.0543 (0.0011)	0.0579 (0.0010)
Bell412	NN-WLSQ	0.0804 (0.0019)	0.0628 (0.0016)	0.0510 (0.0011)	0.0602 (0.0010)

can be composed from only non-negative linear combination of elements in the basis. If there is no access to tool with non-negative least squares method or tool that can solve quadratic programming problem with non-negativity constrains (non-negative least squares problem is equivalent with a quadratic programming problem, [20]), the least squares method can be used. It usually leads to acceptable results.

Although the first method assumes that multiple spectra were recorded at a constant altitude, the estimates are relatively robust to cases when altitude slightly changes during the measurements at one altitude level.

No ideal dataset (spectra recorded with gradually increasing altitude) for PCA method is available to authors of the paper. Based on authors' communication with crew of aircraft the scenario for this method could be often preferred. The method was applied to data for only three altitudes. Even for this data, the results were not far from results of NN-WLSQ method. Application of the method to simulated datasets leads to consistent results.

The background determined over the Brdy hills may be one case of the presence of radon daughter products during the measurement and calculation of the background. Residuals (difference between spectra and "smoothed" spectra) also suggests presence of other source of signal that was not fully explained by the basis with primary and secondary components. The results for Brdy hills are example of results that should not be further used. In such cases further analysis of residuals could provide a clue of what is wrong. Systematic patterns in residuals could be signs of the biased results for a helicopter background or the cosmic contribution.

The uncertainty estimates are only approximate and are only valid when there are no deviations from the theoretical models. Assuming wrong location of a primary component can bias the results (it usually overestimates). Note that in this case the helicopter background and cosmic gamma-ray contribution will be biased and method should not be used.

6. Conclusion

Not always large water areas are available for the helicopter background measurements. Similar experiments as in Switzerland and France were carried out in the Czech Republic over landscape. The background measurements in France over sea, Switzerland over large water areas and in the Czech Republic over landscape were compared. Based on the results presented in this paper we can conclude that good results in determining the helicopter and cosmic spectra were also obtained in the Czech Republic even if the measurements were not carried out over a large water area, but at the altitudes from 1500 to 2900 m (AGL). The mathematical methods using NN-WLSQ method or PCA gave results of the helicopter background and cosmic gamma-ray spectrum that can be used to elimination of helicopter background and cosmic contribution to measured spectra. After the elimination, the variability of results over the reference areas decreased and they were closer to reference values measured on the ground (e.g. by HPGe detector).

Three flight levels for the application of the first method are recommended. If the difference between two altitudes is at least 1000 m or more, only datasets at two flight levels can be used to achieve good results. The acquisition time should be at least 5 min. The longer the measuring time, the better results are achieved. Due to uncertainties in the MCNP response functions for low energies, it was recommended to use only energy range from 120 keV to 3 MeV in all the calculations (both background and measured spectra).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This work was supported by Ministry of Interior of the Czech Republic as part of project Recovery Management Strategy for Affected Areas after Radiation Emergency, VH20172020015 and by the programme of the Ministry of the Interior of the Czech Republic MV-17196-1/OBVV-2022.

References

- [1] R.L. Grasty, B.R.S. Minty, A Guide to the Technical Specifications for Airborne Gamma-Ray Surveys, Australian Geological Survey Organisation, 1995. Record 1995/60.
- [2] B.R.S. Minty, A.P.J. Luyendyk, R.C. Brodie, Calibration and data processing for airborne gamma-ray spectrometry, *AGSO J. Aust. Geol. Geophys.* 17 (2) (1997) 51–66.
- [3] B.R.S. Minty, Fundamentals of airborne gamma-ray spectrometry, *AGSO J. Aust. Geol. Geophys.* 17 (2) (1997) 39–50.
- [4] V.V. Drovnikov, N.Y. Egorov, V.V. Kovalenko, Y.A. Serboulov, Y.A. Zadorozhny, Some results of the airborne high energy resolution gamma-spectrometry application for research of the USSR European territory radioactive contamination in 1986 caused by the Chernobyl accident, *J. Environ. Radioact.* 37 (2) (1997) 223–234.
- [5] B. Bucher, L. Rybach, G. Schwarzc, In-flight, online processing and mapping of airborne gamma spectrometry data, *Nucl. Instrum. Methods Phys. Res.* 540 (2005) 495–501.
- [6] J. Hulka, I. Cespirova, Z. Prouza, Modern data acquisition of the contaminated landscape cover, *Radioprotection* 44 (5) (2009) 619–621.
- [7] J.D. Allyson, D.C.W. Sanderson, Monte Carlo simulation of environmental airborne gamma-spectrometry, *J. Environ. Radioact.* 38 (3) (1998) 259–282.
- [8] J.D. Allyson, D.C.W. Sanderson, Spectral deconvolution and operational use of stripping ratios in airborne radiometrics, *J. Environ. Radioact.* 53 (3) (2001) 351–363.
- [9] J. Klusoň, Environmental monitoring and in situ gamma spectrometry, *Radiat. Phys. Chem.* 61 (2001) 209–216.
- [10] IAEA, Airborne Gamma-Ray Spectrometer Surveying, Technical Report Series No. 322, IAEA, 1991.
- [11] IAEA, Guidelines for Radioelement Mapping Using Gamma-Ray Spectrometry Data, IAEA-TECDOC-1363, IAEA, 2003.
- [12] Y. Sanada, A. Kondo, T. Sugita, Y. Nishizawa, Y. Yuuki, K. Ikeda, Y. Shoji, T. Torii, Radiation monitoring using an unmanned helicopter in the evacuation zone around the Fukushima Daiichi nuclear power plant, *Explor. Geophys.* 45 (1) (2014).
- [13] R. Pöllänen, et al., Radiation surveillance using an unmanned aerial vehicle, *Appl. Radiat. Isot.* 67 (2) (2009) 340–344.
- [14] P. Jurza, I. Campbell, P. Robinson, R. Wackerle R, P. Cunneen P, B. Pavlík, Use of ^{214}Pb photopeaks for radon removal utilizing current airborne gamma-ray spectrometer technology and data processing, *Explor. Geophys.* 36 (2005) 322–328.
- [15] M. Baldoncini, M. Albéri, C. Bottardi, B. Minty, K.G.C. Raptis, V. Strati, F. Mantovani, Exploring atmospheric radon with airborne gamma-ray spectroscopy, *Atmos. Environ.* 170 (2017) 259–268.
- [16] J.O. Ramsay, Monotone regression splines in action, *Stat. Sci.* 3 (4) (1988) 425–441.
- [17] J. Hovgaard, R. Grasty, Reducing statistical noise in airborne gamma-ray data through spectral component analysis, in: A.G. GUBINS (Ed.), Proceedings of Exploration 97: Fourth Decennial Conference on Mineral Exploration, 1998, pp. 753–764.
- [18] G. Butterweck, et al., International intercomparison exercise of airborne gamma-spectrometric systems of the Czech republic, France, Germany and Switzerland in the framework of the Swiss exercise ARM17, report number, 18-04, PSI Bericht Nr (2018).
- [19] M. Ohera, D. Sas, P. Sládek, Calibration of spectrometric detectors for air kerma rate, *Nucl. Technol. Radiat. Protect.* 35 (4) (2020) 323–330.
- [20] V. Franc, V. Hlavac, M. Navara, Sequential Coordinate-Wise Algorithm for the Non-negative Least Squares Problem, pp. 407–414.