DOI: 10.1007/s10967-008-7419-9 *Journal of Radioanalytical and Nuclear Chemistry, Vol. 280, No.3 (2009) 519–531*

Scaling model for prediction of radionuclide activity in contaminated soils using a regression triplet technique

S. Dulanská,[1] M. Meloun,[2] L. Mátel[1]

[1]Department of Nuclear Chemistry, Faculty of Natural Sciences, Comenius University, Mlynska Dolina, 842 15 Bratislava, Slovakia
[2]Department of Analytical Chemistry, Faculty of Chemical Technology, University of Pardubice, Studentská 96, 532 10 Pardubice, Czech Republic

(Received February 4, 2009)

A scaling model was built to calculate the activity of alpha emitting radionuclides in contaminated soil in the lysimeter field. Linear regression can be applied for the evaluation of radioactivity measurement data. Activities of the radionuclides ^{241}Am, ^{238}Pu, ^{239,240}Pu and ^{90}Sr obtained by experiments from real contaminated soils of the experimental lysimeter placed in a nuclear power plant (NPP) in Slovakia were evaluated using linear regression models with the method of least squares. A suitable scaling model for monitoring the ^{241}Am, ^{238}Pu, ^{239,240}Pu alpha radionuclide activity was built using the regression triplet analysis and regression diagnostics. A regular designed scaling model opens the possibilities of long-time activity monitoring of these radionuclides, thus decreasing the number of necessary radiochemical analyses. The Fisher-Snedecor test, however, confirmed that the regression model for ^{90}Sr activity monitoring by ^{241}Am, ^{239,240}Pu activity determination in contaminated soils can not be recommended.

Introduction

Direct determination of hard-to-detect radionuclides (HD-RN), including alpha or beta emitters, in the presence of more energetic radionuclides (RN) is costly and therefore unacceptable for their control.[1] Most transuranic radionuclides belong to such hard-to-detect radionuclides, which are alpha emitters or fission products emitting only beta particles. Alpha spectrometry is normally used for determination of ^{238}Pu, ^{239,240}Pu and ^{241}Am. Due to the similar energies of alpha particles from ^{239}Pu and ^{240}Pu, alpha spectrometry measures their sum activity.[2] The declared content of radionuclides in radioactive waste is linked with prognosis, practical importance, experimental and financial accessibility, and can be realized by various methods. One of the methods is a declaration on the basis of calculation with model and entrance data.[1,4,6] Another method of radionuclide declaration is experimental determination by radiochemical and/or nonradiometric methods. The scaling factor method is used for evaluating radioactivity concentrations of the radioactivity inventory of wastes and is commonly used in nuclear environments for evaluating the radioactivity of HD-RN, i.e. radionuclides from the radioactivity of a key radionuclide.[3,4] International guidelines recommend to use the basic scaling factors methodology, which determines radioactivity of HD-RN.[1,3–9] The scaling factor method is mainly intended for evaluating the radioactivity inventory of wastes.[4] Statistical calculation is a supplemental technique used for the quantitative evaluation of scaling factor method parameters.[1,3,4] The scaling factor, *SF*, is defined as $A_{HD\text{-}RN} = SF \times A_{KN}$, where $A_{HD\text{-}RN}$ and A_{KN} are the activity concentration of the hard-to-detect radionuclides and the key nuclide, and *SF* is the scaling ratio that describes the relationship between radionuclides.[4] The standard guide denotes the geometrical mean and linear regression of logarithm $log(c_{HD\text{-}RN}) = a + b*log(c_{KN})$, where a is constant, R is regression coefficient, which can be applied for the determination of the scaling factor.[4] Some literature sources maintain that the standard guide denotes that the scaling factor method can be applied only when there is a good correlation between KN and HD-RN.[4–10] Experimental data often feature a non-constant spread, a small number of values, asymmetric distribution and exhibit violation of basic assumptions about the sample. These departures should be revealed by exploratory data analysis and therefore particular transformation must be performed if necessary. The so-called retransformed mean should be used as the best mean estimate instead of the geometric mean. Correlation and regression analyses are often confused in evaluations of experimental results. However, the correlation analysis describes the influence of one variable level changes on changes of other variable levels and holds for variables measured quantitatively. It detects the existence and nature of dependencies, measures the goodness-of-fit of an actual model and tests the hypotheses of statistical significance of the model proposed. The y variable does not depend on the x variable but two random variables, x and y, vary jointly. The regression model, however, takes the independent variable x as a non-random variable and the dependent variable y as a random one, in contrast to the correlation model. Scheme and statistical evaluation of suitable scaling models for

* E-mail: dulanska@fns.uniba.sk

0236–5731/USD 20.00

Akadémiai Kiadó, Budapest
Springer, Dordrecht

monitoring the radionuclides ^{241}Am, ^{238}Pu and ^{239,240}Pu investigated in real contaminated soils of an experimental lysimetric field on the basis of experimental results were the objectives of the study.

There are a number of common difficulties associated with real datasets. The first involves the detection and elimination of outliers in the original data. We think of data as being divided into two classes (1) good observations (the majority of data) reflecting population scatter of data and (2) outliers (if any) being a part of the so-called influential fatal points or noise. The goal of any outlier detection is to find this true partition and thus to separate good from outlying observations. Regression diagnostics represents procedures for an examination of the regression triplet (data, model, method) for identification of (a) the data quality for a proposed model; (b) the model quality for a given set of data; (c) fulfillment of all least-square assumptions.

Theoretical

A linear regression model is a model which is formed by a linear combination of explanatory variables **x** or their functions $y = X\beta + \varepsilon$, (Ref. 11). For the determination of statistical properties of the random vectors $\hat{y}_p, \hat{e}$ and σ^2, some basic assumptions are necessary for the method of least squares (LS) to be valid. The random errors ε_i have constant and finite variance, $E(\varepsilon_i^2) = \sigma^2$. The conditional variance σ^2 is also constant and therefore the data are said to be homoscedastic. The random errors ε_i are uncorrelated and therefore $\mathrm{cov}(\varepsilon_i, \varepsilon_i) = E(\varepsilon_i, \varepsilon_i) = 0$. The random errors ε_i have a normal distribution $N(0, \sigma^2)$.

A source of problems in a LS application may be found in the components of the regression triplet, i.e. the data quality for a proposed model, the model quality for a given data set, and the regression method quality when all assumptions used in "classical" LS regression are not fulfilled: regression diagnostics are used because there is no necessity for an alternative hypothesis, but all types of deviations from an ideal regression triplet are discovered.[11–23]

Examination of data quality

Examination of data quality involves detection of the influential point IP (i.e. outliers and high-leverages), which cause many problems in regression analysis by shifting the parameter estimates or increasing the variance of the parameters; a survey is provided in Ref. 22. Outliers are detected by analysis of the various types of residuals, hat matrix elements and related statistics. For the identification of influential points, i.e. outliers and high-leverages, various types of graphs can be used:

(1) The Graph of Predicted Residuals has the high-leverage points easily detected by their locations as they lie outside the line $y = x$ and are located quite far from this line. The outliers are located on the line $y = x$, yet far from its central pattern.[11,22]

(2) The Williams Graph[15] contains two boundary lines drawn, the first for outliers, $y = t_{0.95}\,(n–m–1)$, and the second for high-leverages, $x = 2m/n$. Note that $t_{0.95}(n–m–1)$ is the 95% quantile of the Student distribution with $(n–m–1)$ degrees of freedom.[11,22]

(3) The Pregibon Graph has two different constraining lines drawn, $y = –x + 2(m + 1)/n$ and $y = –x + 3(m + 1)/n$. To distinguish between influential points the following rules are used: (a) a point is strongly influential if it is located above the upper line; (b) a point is influential if it is located between the two lines. The influential point can be either an outlier or a high-leverage point.[11,22]

(4) Gray's L-R Graph[25] has all the points under the hypotenuse of a triangle with a 90° angle in the origin of the two axes. In Gray's L-R graph, contours of the same critical influence are plotted and the locations of individual points are compared with them. It may be determined that the contours are hyperbolic as described by $y = (2x–x^2–1)/(x(1–K)–1)$ where $K = n(n–m–1)/(c^2\,m)$ and c is a constant. For $c = 2$, the constant K corresponds to the limit $2/\sqrt{m/n}$. The constant c is usually equal to 2, 4 or 8.[11,22]

Examination of the model quality

Examination of the regression model quality can be considered directly from the regression scatter plot of y vs. x. Individual parameters are tested for significance using the Student t-test.[11,26]

The Fisher-Snedecor F-test of significance of the regression model proposed is based on the testing criterion

$$F_R = \hat{R}^2(n-m)/[(1-\hat{R}^2)(m-1)]$$

which has the Fisher-Snedecor distribution with $(m–1)$ and $(n–m)$ degrees of freedom, $\hat{R}^2$ means an estimate of the determination coefficient, n is the number of data points and m is the number of parameters for a straight line $m = 2$. With the use of F_R the null hypothesis H_0: $R^2 = 0$ may be tested and concerns a test of significance of all regression parameters β. Various test criteria for a search of regression model quality may be used. One of the most efficient is the mean quadratic error of prediction, MEP, being defined by the Across-validation relationship

$$\mathrm{MEP} = \frac{\sum_{i=1}^{n}(y_i - x_i^T b_{(i)})^2}{n} .$$

where $b_{(i)}$ is the estimate of regression parameters when all points except the ith one were used and $\boldsymbol{x}_i$ is the ith row of matrix **X**. The statistic MEP uses a prediction $\hat{y}_{p,i}$ from an estimate constructed without including the ith point. The MEP can also be used to express the predicted determination coefficient

$$\hat{R}_p^2 = 1 - \frac{n \times \mathrm{MEP}}{\sum_{i=1}^{n} y_i^2 - n\bar{y}^{-2}}$$

Another statistical characteristic, quite generally used, is derived from information theory and entropy, and is known as the Akaike information criterion

$$\mathrm{AIC} = n \times \ln\left(\frac{U(b)}{n}\right) + 2m \quad .$$

The most suitable model is the one which gives the lowest value of the mean quadratic error of prediction MEP and Akaike information criterion AIC and the highest value of the predicted determination coefficient R_p^2.

Examination of the regression method used

Several tests are often performed for the fulfillment of three important assumptions for the least-squares method, namely homoscedasticity, absence of autocorrelation and the normality of random errors. The Cook-Weisberg test of the homoscedasticity of residuals: The test for homoscedasticity is carried out by checking the null hypothesis

H_0: $\lambda = 0$ in relation $\sigma_i^2 = \sigma^2 \exp(\lambda x_i \beta)$ where x_i is the ith concentration. COOK and WEISBERG introduced the test criterion

$$S_f = \frac{\left[\sum_{i=1}^{n} (\hat{y}_i - \bar{y})\hat{e}_i^2\right]^2}{2\hat{\sigma}\sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2}$$

where

$$\bar{y} = (\sum_{i=1}^{n} (\hat{y}_i)/n \quad .$$

When the null hypothesis is valid, the test statistics have approximately an $\chi^2(1)$ distribution with one degree of freedom.[27]

The test of the normality of errors: Firstly, the normality of errors may be simply examined by a rankit quantile–quantile Q–Q plot containing the order statistics of classical residuals $\hat{e}_{(i)}$ in dependence on the quantile of the normalized normal distribution u_{Pi} for $P_i = i/(n+1)$, $i = 1, \ldots, n$. Secondly, the most convenient test for linear models seems to be the Jarque-Berra test, which is based on the criteria of residual skewness and residual kurtosis. When this criterion $L(\hat{e}) > \chi^2_{0.95} = 5.99$, the null hypothesis H_0 about the error normality is rejected.

Violation of some assumptions for the OLS method: The effects of deviations from these basic assumptions for the ordinary least-squares method and methods that correct these effects leading to a more accurate regression model are as follows:

(a) When heteroscedasticity is found in the data, the weighted least squares method (WLS) is used, see pages 102–104 in Ref. 11.

(b) When autocorrelation is found in the data, the generalized least-squares method (GLS) is used, see page 110 in Ref. 11.

Experimental

Soil samples were taken from the lysimeter field Nuclear Power Plant A-1 located at Jaslovske Bohunice (NPP Slovakia), with an unknown amount of ^{90}Sr, ^{241}Am, ^{239,240}Pu, ^{238}Pu radionuclides, which were analyzed. Experimental works on the lysimetric field were defined and performed within the scope of the project "Decommissioning of the NPP A-1" Processing and Treatment of Contaminated Soils and Concrete Rubble".[28] The aim of the proposed lysimetric system was to obtain comparative results on individual radionuclide migration and their leaching in real conditions. The lysimetric field in NPP is a concrete building (9.5 m×4 m) with six plastic tanks. Each tank was filled with 1 m^3 of compressed contaminated soil and had a releasing valve at the bottom. The samples of radioactive material from contaminated soils from tank No. 4 are presented in Table 1.

Tank No. 4 has two sorption layers of natural zeolite made of drainaged gravel layer. It imitates the sandwich configuration of the landfill of contaminated soils. Practical performance of lysimetric experiments began in the course of the year 2002 and the research project lasted for three years.[28] In 2004 cores were drilled from tanks No. 2 and No. 4 for the purpose of control of radionuclides migration in the soil profile. After sampling and determination of samples of soils from Tank No. 4, the results should give the answer on the influence of contaminated soil layers on released radionuclides concentration. The samples were specifically taken by NPP A-1 staff using drilling equipment HILTI DD 250 E.[28]

Table 1. Selected samples of contaminated soils from drill hole in tank[28] No. 4

Depth of layer, cm	Number of soil	a ^{137}Cs, Bq.kg^{-1}	a^{90}Sr, Bq.kg^{-1}
0–6	A	101 000	23 900
9–15	B	210 000	11 100
24–28	C	88 900	4 730
46–50	D	65 900	13 700
77–80	E	61 900	13 100
Zeolite 80–87		527	464

Sample preparation

The concentration of HD-RN is usually determined by destructive assay. The method consists basically of the following: sample dissolution, specific chemical separation process and radiometry depending on chemical and radioactive properties of the nuclides. Before dissolution, the samples were dried to constant weight at 105 °C and ashed at 550 °C for 18 h to remove any carbonaceous material. Ashing temperatures above 550 ºC should be avoided to minimize conversion of Pu to intractable PuO_2. Tracers (^{242}Pu, ^{243}Am) and carriers (Sr^{2+}, Y^{3+}) were added to the samples before digestion with 8M nitric acid and hydrogen peroxide.

1. Digestion in an autoclave: 20 g samples were digested in a mixture of 50 mL 8M HNO_3 + 2.0 mL of 30% H_2O_2 at 150 °C for 8 hours. Samples were cooled and centrifuged. The centrifugate was transferred to a beaker and digestion in an autoclave was repeated twice.

2. Digestion using hot-plate heating: 5 g sample was weighed into a beaker and digested (20 mL 8M HNO_3 + 1 mL H_2O_2) during heating on a hot plate at 70 °C for about 8 hours. The solid and liquid phases were separated, cooled and centrifuged. Leaching was repeated three times. All centrifugates were collected.

3. Microwave digestion procedure (Microwave Laboratory System Milestone Ethos): 0.5 g dry samples and 4 mL conc. HNO_3 and 100 µL H_2O_2 were added to each microwave container.

Separation of plutonium

Tetravalent plutonium (hexanitrate complex $[Pu(NO_3)_6]^{2-}$) was extracted from 8M HNO_3 into a 30% Aliquat-336 in toluene mixture.[29] Aliquat-336 (methyl tricaprylyl ammonium chloride) is a quaternary ammonium salt extracting different species with an anion exchange mechanism. Aliquat-336 extracts tetravalent actinide nitrate complexes. Aliquat-336 (chloride form) was diluted in 30% toluene and converted to the nitrate form by equilibration with 4M and 8M HNO_3. Before separation of plutonium 2–3 g of sodium nitrite were added to the samples to ensure conversion of Pu(III) to Pu(IV). The molarity of the solution has to be 8–8.4M HNO_3.[29]

The remaining water phase containing Am–Cm and Sr was split into two portions and used in further experiments (separation of Am and Sr). Radionuclide impurities, as uranium, were washed out from the organic phase containing plutonium with 8M HNO_3 (twice) and Th with concentrated HCl (four times). The washed phase was discarded. The plutonium was stripped from the Aliquat-336 with two equal volume washes of 0.15M HCl–0.025M $H_2C_2O_4$. The combined plutonium-bearing strip solutions were evaporated to dryness and ashed in a muffle furnace for 30 min at 550 °C. After cooling, 4 mL of concentrated HNO_3 was added and evaporated to dryness. The sample was dissolved using 4 mL concentrated HCl, evaporated and again dissolved with 6 mL 1M HCl and transferred to a plastic tube.[29]

Separation of americium

Americium was separated from the first portion of the water phase. The solution was evaporated and subsequently dissolved in 0.1M HNO_3/4M $NaNO_3$. Americium was separated by liquid-liquid extraction with 0.3M TOPO in toluene.[33] Americium was stripped from TOPO/toluene using 4M HNO_3 and washed with toluene (twice).[30,33] The solution was evaporated and dissolved with 6 mL 1M HCl (do not burn!). The final source for alpha spectrometry of americium and plutonium were prepared by using the micro coprecipitation on neodymium fluoride (NdF_3).[29–31] In this work plutonium and americium were coprecipitated from 1M HCl with 70 µg of NdF_3 by addition of 50 µg Nd carrier and HF.[31] The NdF_3 precipitate was allowed to form for 35 min and was then filtered through a 0.2 µm (25 mm) pore size polysulphone membrane filter (Tuffryn HT-200, Pall Corporation) upon which a substrate of 140 µg of NdF_3 had been formed and thus surface penetration of the $Nd(Pu)F_3$ precipitate into the filter was minimized.[31] The filtered precipitate was washed with 2×3 mL of 4% HF followed by 2×3 mL of 30% ethanol and dried in a vacuum desiccator at room temperature. The activity of plutonium and americium was determined using alpha spectrometry.

A two-chamber α-spectrometer 576A equipped with ULTRA^{TM} ion implanted silicon detectors, 600 mm^2 active area was used for counting alpha radioactivity. The spectra were processed by using the Alpha-vision^{TM} 32–bit emulation software from the company EG&G ORTEC.

Separation of strontium

Strontium was separated from the second portion of the water phase. A strontium fraction was evaporated to dryness and the residue was dissolved in concentrated HNO_3. Strontium-90 was determined by beta-counting

the daughter activity yttrium-90. Yttrium was separated from the fraction, which contained also Am, Sr and another component, by liquid-liquid extraction with TBP (tributyl phosphate). The extractant TBP was conditioned with concentrated HNO_3 in nitrate form (time of separation is to be noted).[31] The organic phase contained yttrium and the aqueous phase contained the Am, Sr fraction. The TBP phase was washed with concentrated HNO_3. Yttrium was eluted from TBP using 15 mL deionized water and 2M HNO_3. Saturated ammonium oxalate solution (30 mL) was added to the beakers with ^{90}Y formed precipitation of yttrium oxalate $Y_2(C_2O_4)_3.9H_2O$. The mixture was further heated on the hot plate (70 °C) with occasional stirring for 15 min. The beakers were cooled to room temperature and the precipitate was filtered (Whatman No. 42 filter paper). The yttrium oxalate precipitate was washed with 25 mL of water, followed by 25 mL of 95% ethyl alcohol.[31] Yttrium oxalate was determined by low level alpha-beta counter Tesla NRR 610 and Tesla NA 6201.

Notice: Yttrium yield was determined from the ratio of the weight of the sample yttrium oxalate to the expected weight of yttrium oxalate as determined from the yttrium carrier standardization.

A separation scheme of ^{241}Am, $^{239,240}Pu$, ^{238}Pu, ^{90}Sr is shown in Figure 1.

Supporting information available

Complete computational procedures, input data specimens and corresponding output in numerical and graphical form for the program QC-EXPERT or ADSTAT (Trilobyte, Pardubice, Czech Republic) are available online at http://meloun.upce.cz in the block *DATA*.

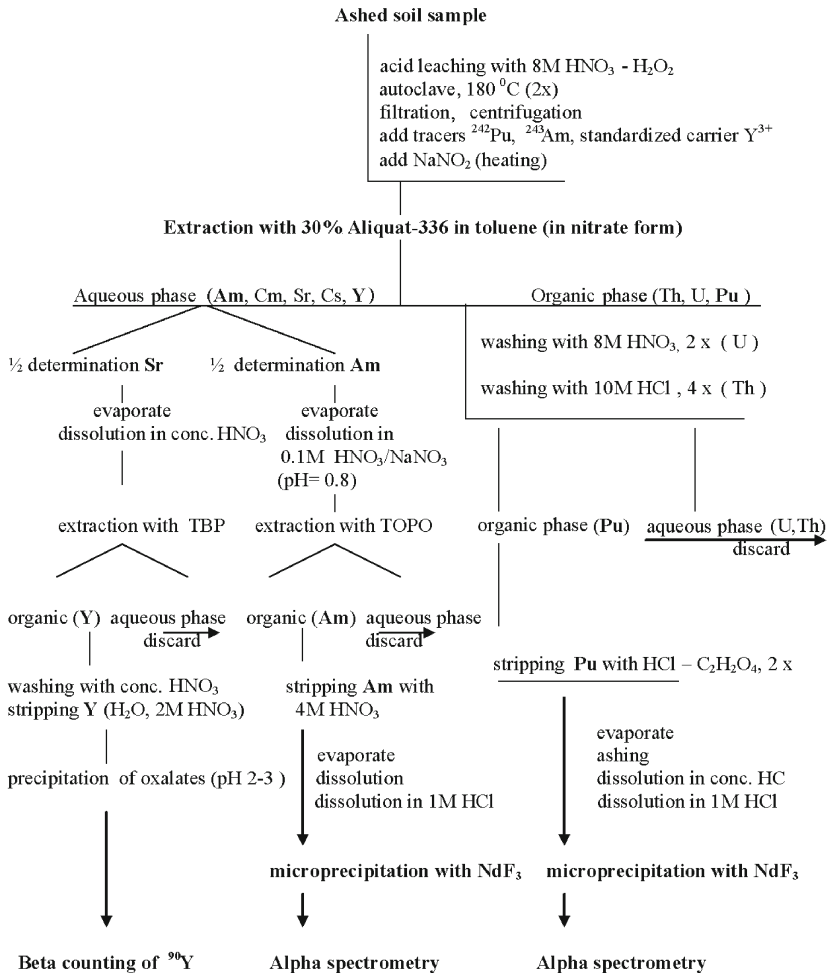

Fig. 1. Scheme for separation of Pu, Am, Sr in contaminated soil samples

Results

Proposal of a scaling model for original data for contaminated soil (B)

Using the original set of data, the ordinary least-squares method (OLS) finds the regression model

$$y = -5.298(1.043, R) - 0.408(0.043, R)\, x_1$$

where standard deviations of the parameters estimated are in brackets and the letter R means that $H_0 : \beta_i = 0$ is rejected and the parameter β_i is statistically significant, while the letter A means that $H_0 : \beta_i = 0$ is accepted and β_i is not statistically significant. The critical quantile $t_{0.975}(27–2) = 2.0595$ of a Student t-test at 5% significance level was used to examine the statistical significance of the individual regression parameters and the estimates of $\hat{\beta}_0$ and $\hat{\beta}_1$ were found to be significant. The scaling model was built with the correlation coefficient $R = 0.8853$ and the determination coefficient $D = 78.38\%$, which also expresses the percentage of points which fulfill the model proposed. The mean error of prediction MEP = 0.2327, the Akaike information criterion AIC = –39.01 and the residual standard deviation $s(e) = 0.469$ were also calculated to prove the scaling model proposed. All these statistics form the resolution criteria for the selection of the best model among several plausible ones.

Detection of influential points

There are many suspicious points in data (10, 20, 22, 24, 25, 11, 6, 12, 16, 14) in Fig. 2 which are located outside the Working-Hotteling confidence bands of the scaling model straight line. Outliers are identified by examination of the residuals.

The scatter plot of regression straight line (Fig. 2), the scatter plot of ordinary residuals against the prediction (Fig. 2a), the index plot of absolute values (Fig. 3b) or squared residuals (Fig. 3c), and the index plot of jackknife residuals may indicate influential points which may be considered suspicious and some testing diagnostics for influential points should be applied. Although the common practice of many programs for statistical analysis of classical residuals is to examine by use of statistical measures of location and spread, such as the residual mean $\overline{e}$, the residual variance $s^2(e)$, the residual skewness $g_1(e)$ and the residual kurtosis $g_2(e)$, these residual statistics do not give a correct indication of the influential points, namely outliers. Diagnostic plots constructed from residuals and hat matrix elements represent a combination of various types of residuals with the diagonal elements of the projection hat matrix H_{ii} and lead to four diagnostic graphs of influential points (the analyzed data set $\{x,y\}$ = {*Pu*-238, *Pu*-239,240} of size $n = 27$, $m = 2$).

The Graph of Predicted Residuals (Fig. 4a), one of the simplest graphs, indicates outliers (6, 24, 16, 20, 22) located far from the central pattern on the line $y = x$.

The Williams Graph (Fig. 4b) has two testing boundary lines, the first line for outliers $y = t_{0.95}(n–m–1)$ detecting three outliers (6, 20, 22), and the second for high-leverage points $x = 2m/n = 0.5$ detecting two high-leverages (8, 19).

The Pregibon Graph (Fig. 4c) is able to distinguish strongly influential points from medium influential points only. The one point (19) was found as medium influential.

Gray's L-R Graph (Fig. 4d) indicates strongly influential points (22, 20, 9, 19, 6) and separates them into outliers (22, 20, 6), points which lie high in the y-axis, and high-leverages (9, 19), which lie in direction of the x-axis.

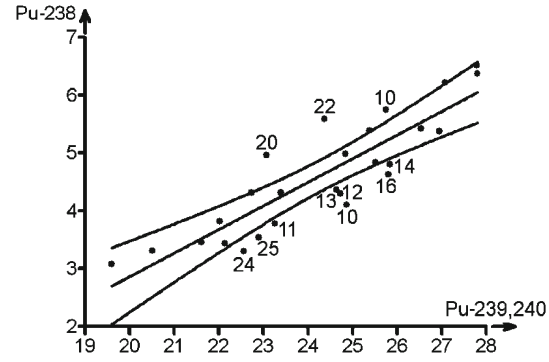

Fig. 2. Scatter plot of the data set $a(^{238}Pu)$ [Bq.kg^{-1}] vs. $a(^{238,239}Pu)$ [Bq.kg^{-1}] presents the regression straight line of radioactivity tested and is constructed with the use of the regression triplet. Influential points which are here suspicious are tested whether they are outliers. Outliers indication shown in Fig. 2 and Fig. 3. Results in Table 2

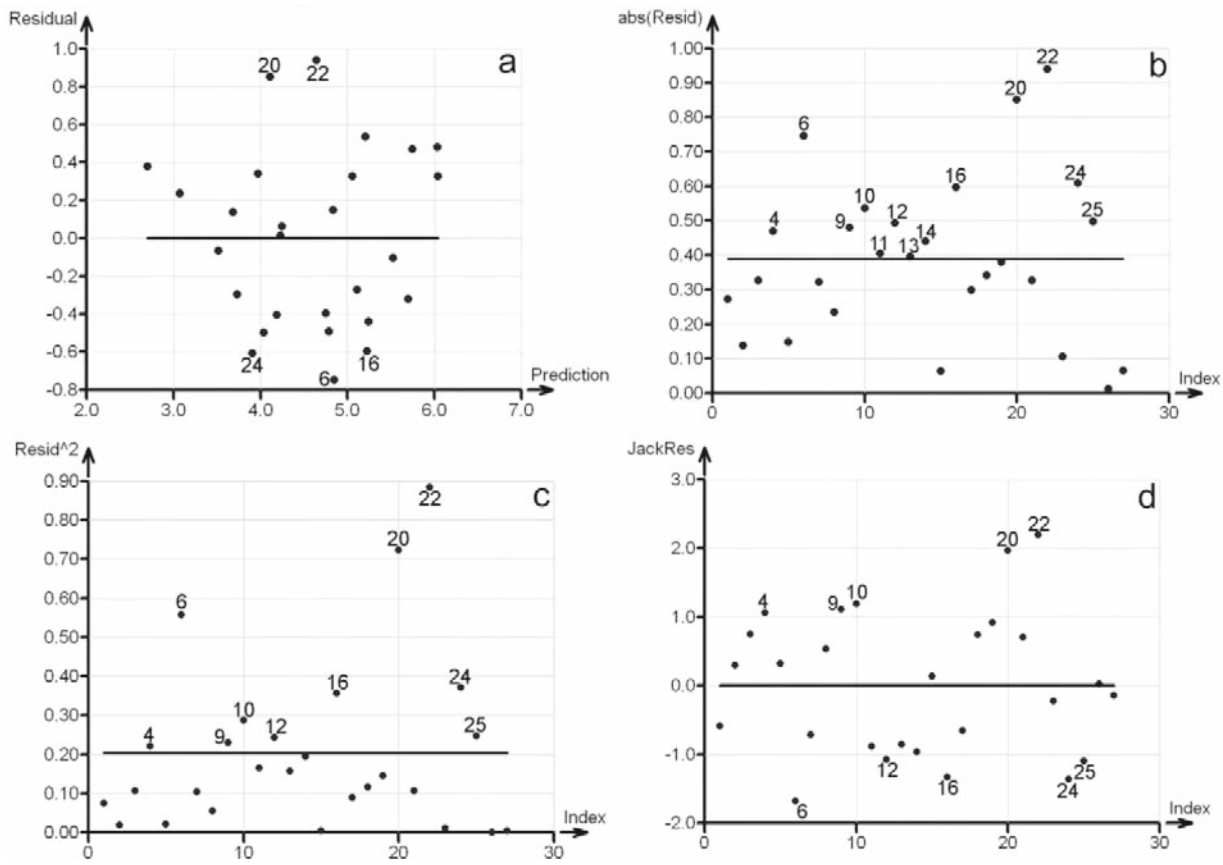

Fig. 3. Prediction graph and three index graphs of various residuals for the data set ^{238}Pu vs. $^{239,240}Pu$: (a) Graph of ordinary residuals vs. prediction; (b) Index graph of the absolute value of residuals; (c) Index graph of squared residuals; (d) Index graph of jackknife residuals

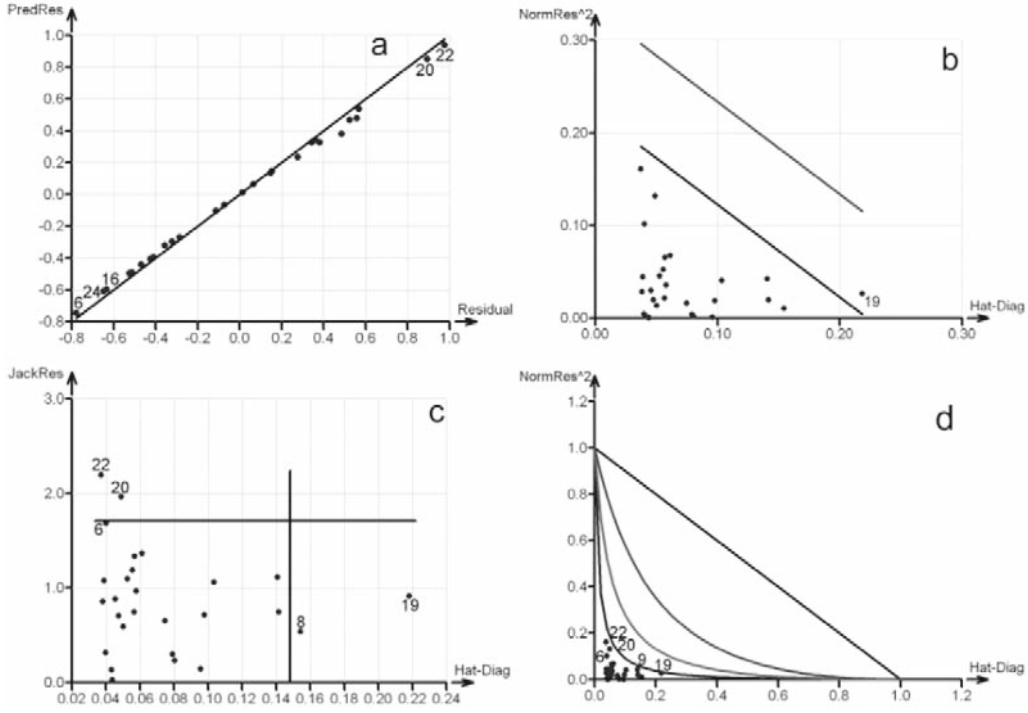

Fig. 4. Diagnostics based on residual plots and hat matrix elements for the data set ^{238}Pu vs. $^{239,240}Pu$: (a) Graph of predicted residuals, (b) Williams graph, (c) Pregibon graph, (d) Gray's L-R graph

Examination of assumptions for least-squares application

Some requirements must be met if the least-squares method is to be applied and to give the best unbiased linear estimates of parameters, the intercept β_0 and the slope β_1:

The Fisher-Snedecor F-test[11] leads to the statistical criterion $F = 139.9$, while the quantile $F(1-\alpha,m-1,n-m) = 4.3$ is lower and therefore the scaling regression model proposed is statistically significant.

The Cook-Weisberg test of heteroscedasticity[11] has the statistical criterion $SC = 0.148$, while the quantile $\chi^2_{1-\alpha,1} = 3.841$ has a higher value and therefore the residuals exhibit homoscedasticity.

The Jarque-Berra test for normality[11] has the statistical criterion $C = 2.167$, while the quantile $\chi^2_{1-\alpha,2} = 5.991$ has a higher value and therefore the residuals exhibit normality.

The Wald test for autocorrelation[11] has the statistical criterion $W = 4.22$, while the quantile $\chi^2_{1-\alpha,1} = 3.841$ has a higher value and therefore the residuals exhibit a slight autocorrelation.

The sign test for dependence trend in residuals[11] has the statistical criterion $S = 1.016$, while the critical value $N = 1.960$ has a higher value and therefore the residuals exhibit no trend.

Construction of a more accurate scaling regression model

The revised scaling model will be further regarded with the intercept term β_0. Since outliers may influence the regression results, they should be treated with care. There are two possible approaches to the data: either to exclude outliers from the data or to use the robust regression method. One of the greatest disadvantages of applying the robust method is a preference for the regression model proposed, here $y = \beta_0 + \beta_1 x$. On the basis of previous graphical and numerical diagnostics of influential points it may be concluded that the three outliers 6, 20 and 22 should be excluded from the original data set, and new parameter estimates should be recalculated:

$$a[^{238}\mathrm{Pu}] = -5.627(0.865) + 0.420(0.035) \times a[^{239,240}\mathrm{Pu}]$$

(in brackets are the estimated standard deviations). The new better scaling model is confirmed as it has been described with the correlation coefficient $R = 0.9296$, the determination coefficient $D = 86.41\%$, thus expressing a percentage of points which fulfill the model proposed; the lower value of the mean error of prediction MEP = 0.1651, the more negative value of the Akaike information criterion AIC = –43.88 and the lower value of the residual standard deviation $s(e) = 0.385$ prove the higher accuracy of the scaling model.

Table 2. Regression model building and testing leading to the straight line model $y = b_0+b_1x$: $^{238}\mathrm{Pu}=b_0+b_1{}^{239,240}\mathrm{Pu}$, $^{241}\mathrm{Am}= b_0+b_1{}^{239,240}\mathrm{Pu}$, $^{90}\mathrm{Sr}=b_0+b_1{}^{239,240}\mathrm{Pu}$, $^{238}\mathrm{Pu}=b_0+b_1{}^{241}\mathrm{Am}$, $^{239,240}\mathrm{Pu} =b_0+b_1{}^{241}\mathrm{Am}$, $^{90}\mathrm{Sr}=b_0+b_1{}^{241}\mathrm{Am}$ for contaminated soil A, where in statistical testing *A* means that H_0 is accepted while *R* means that H_0 is rejected

	Sample A					
Independent variable x	$^{239,240}\mathrm{Pu}$	$^{239,240}\mathrm{Pu}$	$^{239,240}\mathrm{Pu}$	$^{241}\mathrm{Am}$	$^{241}\mathrm{Am}$	$^{241}\mathrm{Am}$
Dependent variable y	$^{238}\mathrm{Pu}$	$^{241}\mathrm{Am}$	$^{90}\mathrm{Sr}$	$^{239,240}\mathrm{Pu}$	$^{238}\mathrm{Pu}$	$^{90}\mathrm{Sr}$
Estimates of two unknown parameters						
Intercept b_0, H_0: $b_0 = 0$	–2.145,*R*	9.559,*R*	10809.967,*R*	–4.536,*A*	–3.384,*R*	7647.692,*R*
Standard deviation $s(b_0)$	0.535	1.908	2279.725	3.637	0.973	3161.304
Slope b_1, H_0: $b_1 = 0$	0.283,*R*	0.647,*R*	312.813,*R*	1.123,*R*	0.311,*R*	420.172,*R*
Standard deviation $s(b_1)$	0.023	0.082	96.825	0.146	0.039	126.450
Goodness-of-fit statistics for regression model building						
Correlation coefficient R	0.949	0.882	0.606	0.870	0.883	0.617
Determination coefficient D	0.901	0.778	0.367	0.756	0.780	0.380
MEP criterion for the best model	0.191	1.684	4434050.074	4.250	0.316	3512759.990
AIC criterion for the best model	–39.706	11.542	295.144	26.296	–22.679	294.724
$s(e)$ for a fitness test	0.335	1.273	1527.344	1.788	0.541	1511.385
F criterion, $F_{crit} = 4.413$, H_0: correlation is significant	154.318, *A*	62.914, *A*	10.438, *A*	59.017, *A*	63.658, *A*	11.041, *A*
Cook-Weisberg test, $\chi^2_{crit} = 3.841$, H_0: homoscedasticity is significant	2.678, *A*	0.638, *A*	4.458, *R*	1.717, *A*	0.061, *A*	7.648, *R*
Jarque-Berra test, $\chi^2_{crit} = 5.991$, H_0: normality is significant	2.684, *A*	0.719, *A*	0.724, *A*	0.760, A	1.372, *A*	0.723, *A*
Wald's test, $\chi^2_{crit} = 3.861$, H_0: autocorrelation is significant	0.214, *A*	1.294, *A*	1.330, *A*	0.030, *A*	0.153, *A*	0.490, *A*
Number of outliers found	6,19	12	0	0	6	0

Table 3. Regression model building and testing leading to the straight line model $y = b_0+b_1x$: $^{238}Pu=b_0+b_1{}^{239,240}Pu$, $^{241}Am= b_0+b_1{}^{239,240}Pu$, $^{90}Sr=b_0+b_1{}^{239,240}Pu$, $^{238}Pu=b_0+b_1{}^{241}Am$, $^{239,240}Pu =b_0+b_1{}^{241}Am$, $^{90}Sr=b_0+b_1{}^{241}Am$ for contaminated soil B, where in statistical testing *A* means that H_0 is accepted while *R* that H_0 is rejected

	Sample B					
Independent variable *x*	$^{239,240}Pu$	$^{239,240}Pu$	$^{239,240}Pu$	^{241}Am	^{241}Am	^{241}Am
Dependent variable *y*	^{238}Pu	^{241}Am	^{90}Sr	$^{239,240}Pu$	^{238}Pu	^{90}Sr
	Estimates of two unknown parameters					
Intercept b_0	–5.627, *R*	5.742, *A*	10467, *A*	9.059, *R*	–2.878, *R*	13354, *R*
Standard deviation $s(b_0)$	0.865	4.053	5530	2.040	–1.008	5023
Slope b_1	0.420, *R*	0.874, *R*	196, *A*	0.580, *R*	0.282, *R*	79, *A*
Standard deviation $s(b_1)$	0.037	0.165	225	0.076	0.004	187
	Goodness-of-fit statistics for a regression model building					
Correlation coefficient *R*	0.930	0.764	0.187	0.863	0.867	0.090
Determination coefficient *D*	0.864	0.584	0.035	0.746	0.751	0.0081
MEP criterion for the best model	0.165	3.626	5604472	1.169	0.282	6.86E+06
AIC criterion for the best model	–43.880	23.752	358.631	2.469	–27.682	378.238
s(e) for a fitness test	0.385	1.643	2332.850	1.013	0.494	2540.544
F criterion, $F_{crit} = 4.300$, H_0: correlation is significant	139.880, *A*	28.050, *A*	0.757, *R*	58.613, *A*	57.257, *A*	0.180, *R*
Cook-Weisberg test, $\chi^2_{crit} = 3.841$, H_0: homoscedasticity is significant	17.830, *R*	27.860, *R*	30.631, *R*	120.860, *R*	169.090, *R*	66.930, *R*
Jarque-Berra test, $\chi^2_{crit} = 5.991$, H_0: normality is significant	2.167, *A*	1.129, *A*	0.544, *A*	0.740, *A*	0.483, *A*	0.197, *A*
Wald's test, $\chi^2_{crit} = 3.841$, H_0: autocorrelation is significant	4.690, *R*	1.340, *A*	0.478, *A*	0.357, *A*	0.050, *A*	0.948, *A*
Number of outliers found	20,22,6	17,23	0	19,23	19,22,23	0

Other data for contaminated soil (B)

The pair correlation coefficient *R* and the Fisher-Snedecor test suggest the statistical significance of the estimated scaling models: $^{241}Am = b_0 + b_1{}^{239,240}$ Pu, $^{238}Pu = b_0 + b_1{}^{241}Am$, $^{239,240}Pu = b_0 + b_1{}^{241}Am$. The mean quadratic error of prediction MEP and the Akaike information criterion AIC served here for optimization among proposed models, as shown in Table 3.

Low values of the determination coefficient (D = 3.5% for the model ^{90}Sr vs. $^{239,240}Pu$ and D = 0.81% for the model ^{90}Sr vs. ^{241}Am) suggest that experimental points do not correspond to the scaling straight line model and such a model for ^{90}Sr activity monitored with ^{241}Am, $^{239,240}Pu$ activity in contaminated soils can neither be recommended nor used.

Suitable scaling models for ^{241}Am, ^{238}Pu and $^{239,240}Pu$ radionuclide monitoring in contaminated soil samples placed in an experimental lysimeter have been demonstrated by regression triplet analysis. Estimates of β_0 and β_1 parameters for all contaminated soils of the experimental lysimeter have been found by the classical least-squares method. The pair correlation coefficient *R* suggests the statistical significance of the $y_i=b_0 + b_1\, a(^{241}Am)$ regression model in which y_i is the activity of examined radionuclides ^{238}Pu and $^{239,240}Pu$. The statistical significance of the scaling regression models has been confirmed with the Fisher-Snedecor test and also with high values of pair-correlation coefficients *R* and *D* for all contaminated soil samples. The results of regression diagnostics for contaminated soils of the experimental lysimeter are in Tables 2 and 4 to 6.

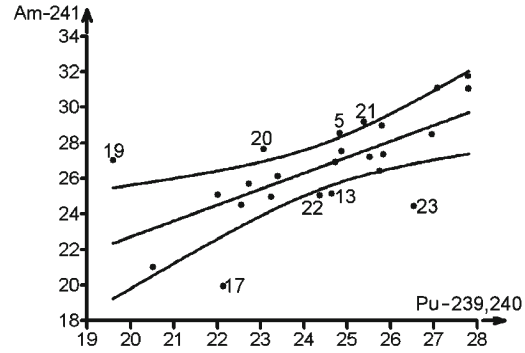

Fig. 5. Scatter plot of the data set $a(^{241}Am)$ [Bq.kg^{-1}] vs. $a(^{239,240}Pu)$ [Bq.kg^{-1}] presents the regression straight line of radioactivity tested and is constructed with the use of the regression triplet

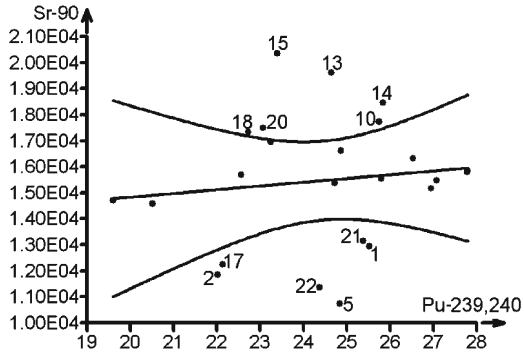

Fig. 6. Scatter plot of the data set $a(^{90}Sr)$ [Bq.kg^{-1}] vs. $a(^{239,240}Pu)$ [Bq.kg^{-1}] presents the regression straight line of radioactivity tested

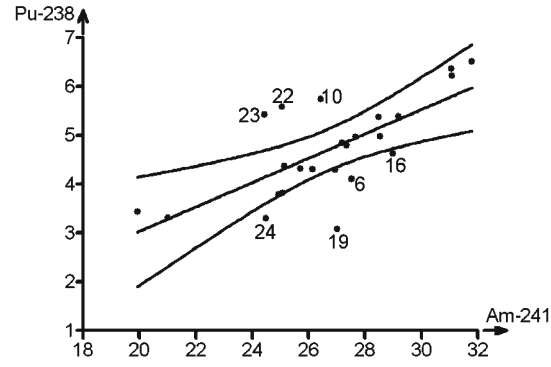

Fig. 8. Scatter plot of the data set $a(^{238}Pu)$ [Bq.kg^{-1}] vs. $a(^{241}Am)$ [Bq.kg^{-1}] presents the regression straight line of radioactivity tested

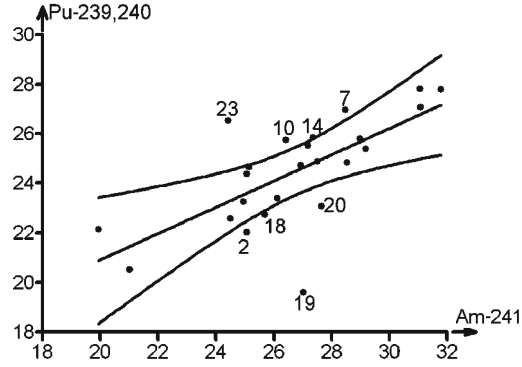

Fig. 7. Scatter plot of the data set $a(^{239,240}Pu)$ [Bq.kg^{-1}] vs. $a(^{241}Am)$ [Bq.kg^{-1}] presents the regression straight line of radioactivity tested

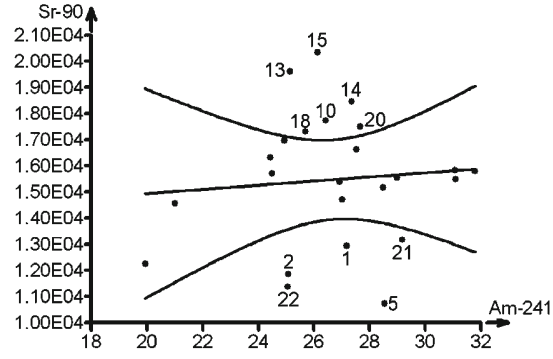

Fig. 9. Scatter plot of the data set $a(^{90}Sr)$ [Bq.kg^{-1}] vs. $a(^{241}Am)$ [Bq.kg^{-1}] presents the regression straight line of radioactivity tested

Table 4. Regression model building and testing leading to the straight line model $y = b_0+b_1x$: $^{238}Pu=b_0+b_1{}^{239,240}Pu$, $^{241}Am= b_0+b_1{}^{239,240}Pu$, $^{90}Sr=b_0+b_1{}^{239,240}Pu$, $^{238}Pu=b_0+b_1{}^{241}Am$, $^{239,240}Pu =b_0+b_1{}^{241}Am$, $^{90}Sr=b_0+b_1{}^{241}Am$ for contaminated soil C, where in statistical testing *A* means that H_0 is accepted while *R* that H_0 is rejected

	Sample C					
Independent variable *x*	$^{239,240}Pu$	$^{239,240}Pu$	$^{239,240}Pu$	^{241}Am	^{241}Am	^{241}Am
Dependent variable *y*	^{238}Pu	^{241}Am	^{90}Sr	$^{239,240}Pu$	^{238}Pu	^{90}Sr
	Estimates of two unknown parameters					
Intercept b_0	–0.428, *A*	4.911, *R*	14586, *R*	–2.652, *A*	0.012, *A*	12445, *R*
Standard deviation $s(b_0)$	0.757	2.213	7088	2.810	0.751	3443
Slope b_1	0.195, *R*	0.850, *R*	289, *A*	1.078, *R*	0.167, *R*	270, *R*
Standard deviation $s(b_1)$	0.024	0.068	220	0.087	0.023	105
	Goodness-of-fit statistics for a regression model building					
Correlation coefficient *R*	0.871	0.958	0.264	0.958	0.890	0.552
Determination coefficient *D*	0.759	0.920	0.070	0.917	0.792	0.305
MEP criterion for the best model	0.279	2.700	29923563	3.462	0.274	5692079
AIC criterion for the best model	–28.524	16.415	430	20.220	–21.522	263.750
s(e) for a fitness test	0.519	1.576	27504798	1.775	0.482	2213.373
F criterion, $F_{crit} = 4.279$, H_0: correlation is significant	66.236, *A*	154.638, *A*	1.727, *R*	154.638, *A*	53.228, *A*	6.574, *A*
Cook-Weisberg test, $\chi^2_{crit} = 3.841$, H_0: homoscedasticity is significant	110.250, *R*	49.194, *R*	132.2, *R*	48.010, *R*	47.900, *R*	55.830, *R*
Jarque-Berra test, $\chi^2_{crit} = 5.992$, H_0: normality is significant	0.153, *A*	0.2400, *A*	3.631, *A*	0.438, *A*	0.884, *A*	1.404, *A*
Wald's test, $\chi^2_{crit} = 3.841$, H_0: autocorrelation is significant	23.310, *R*	0.4796, *A*	34.896, *R*	0.171, *A*	0.425, *A*	1.242, *A*
Number of outliers found	13,9	4	0	4	12	0

Table 5. Regression model building and testing leading to the straight line model $y = b_0+b_1x$: $^{238}Pu=b_0+b_1{}^{239,240}Pu$, $^{241}Am= b_0+b_1{}^{239,240}Pu$, $^{90}Sr=b_0+b_1{}^{239,240}Pu$, $^{238}Pu=b_0+b_1{}^{241}Am$, $^{239,240}Pu =b_0+b_1{}^{241}Am$, $^{90}Sr=b_0+b_1{}^{241}Am$ for contaminated soil D, where in statistical testing *A* means that H_0 is accepted while *R* that H_0 is rejected

	Sample D					
Independent variable *x*	$^{239,240}Pu$	$^{239,240}Pu$	$^{239,240}Pu$	^{241}Am	^{241}Am	^{241}Am
Dependent variable *y*	^{238}Pu	^{241}Am	^{90}Sr	$^{239,240}Pu$	^{238}Pu	^{90}Sr
Estimates of two unknown parameters						
Intercept b_0	–2.874, *R*	19.496, *R*	8440, *R*	–10.620, *A*	–4.943, *R*	3267, *A*
Standard deviation $s(b_0)$	0.459	2.891	1450	6.060	1.900	3439
Slope b_1	0.323, *R*	0.682, *R*	200, *R*	0.964, *R*	0.276, *R*	374, *R*
Standard deviation $s(b_1)$	0.019	0.119	60	0.169	0.053	95
Goodness-of-fit statistics for a regression model building						
Correlation coefficient *R*	0.975	0.811	0.656	0.811	0.794	0.585
Determination coefficient *D*	0.950	0.658	0.430	0.658	0.631	0.342
MEP criterion for the best model	0.137	5.061	1297722	7.071	0.746	1941474
AIC criterion for the best model	–37.351	31.111	239.066	37.678	–6.040	261.323
s(e) for a fitness test	0.316	2.158	1071	2.565	0.803	1348.506
F criterion, $F_{crit} = 4.543$, H_0: correlation is significant	284.914, *A*	32.643, *A*	11.325, *A*	32.643, *A*	27.346, *A*	8.318, *A*
Cook-Weisberg test, $\chi^2_{crit} = 3.841$, H_0: homoscedasticity is significant	73.477, *R*	66.830, *R*	76.820, *R*	28.897, *R*	27.830, *R*	117.8, *R*
Jarque-Berra test, $\chi^2_{crit} = 5.992$, H_0: normality is significant	0.167, *A*	1.781, *A*	0.539, *A*	1.332, *A*	0.630, *A*	0.248, *A*
Wald's test, $\chi^2_{crit} = 3.841$, H_0: autocorrelation is significant	1.262, *A*	1.837, *A*	3.525, *A*	0.003, *A*	1.549, *A*	–0.191, *A*
Number of outliers found	12,16	0	1 2	16	3,12	12

Table 6. Regression model building and testing leading to the straight line model $y = b_0+b_1x$: $^{238}Pu=b_0+b_1{}^{239,240}Pu$, $^{241}Am= b_0+b_1{}^{239,240}Pu$, $^{90}Sr=b_0+b_1{}^{239,240}Pu$, $^{238}Pu=b_0+b_1{}^{241}Am$, $^{239,240}Pu =b_0+b_1{}^{241}Am$, $^{90}Sr=b_0+b_1{}^{241}Am$ for contaminated soil E, where in statistical testing *A* means that H_0 is accepted while *R* that H_0 is rejected

	Sample E					
Independent variable *x*	$^{239,240}Pu$	$^{239,240}Pu$	$^{239,240}Pu$	^{241}Am	^{241}Am	^{241}Am
Dependent variable *y*	^{238}Pu	^{241}Am	^{90}Sr	$^{239,240}Pu$	^{238}Pu	^{90}Sr
Estimates of two unknown parameters						
Intercept b_0	0.748, *A*	9.248, *R*	4945.6, *A*	–1.164, *A*	0.534, *A*	10196, *A*
Standard deviation $s(b_0)$	0.445	1.741	3453	3.361	0.868	4992.1
Slope b_1	0.161, *R*	0.473, *R*	20.610, *A*	1.056, *R*	0.172, *R*	287.74, *A*
Standard deviation $s(b_1)$	0.026	0.103	20.473	0.196	0.051	291.37
Goodness-of-fit statistics for a regression model building						
Correlation coefficient *R*	0.853	0.787	0.026	0.831	0.685	0.255
Determination coefficient *D*	0.727	0.619	0.001	0.691	0.469	0.065
MEP criterion for the best model	0.013	0.228	942200	0.449	0.026	93255
AIC criterion for the best model	–69.168	–22.715	234	–12.75	–54.197	220.22
s(e) for a fitness test	0.101	0.441	922.9	0.615	0.154	919.45
F criterion, $F_{crit} = 4.667$, H_0: correlation is significant	37.359, *A*	21.134, *A*	0.010, *R*	29.068, *A*	11.480, *A*	0.975, *R*
Cook-Weisberg test, $\chi^2_{crit} = 3.841$, H_0: homoscedasticity is significant	14.740, *R*	0.644, *A*	63.390, *R*	30.156, *R*	–0.251, *A*	66.114, *A*
Jarque-Berra test, $\chi^2_{crit} = 5.992$, H_0: normality is significant	0.816, *A*	0.791, *A*	0.250, *A*	0.488, *A*	1.134, *A*	0.611, *A*
Wald's test, $\chi^2_{crit} = 3.841$, H_0: autocorrelation is significant	0.288, *A*	0.488, *A*	10.180, *R*	2.176, *A*	0.022, *A*	11.446, *R*
Number of outliers found	6,17	16	0	13	5	0

The given results confirm theoretical assumptions of a possible correlation between the alpha radionuclides (^{238}Pu, ^{239,240}Pu, ^{241}Am) investigated but do not confirm correlations with strontium (^{90}Sr) in such a complicated matrix as the experimental lysimeter soil. It appears that a combined approach using a sufficient amount of experimental data and adequate statistical treatment is necessary for verification of the hypothetic model. Thus a costly and lenghty radiochemical analysis of one of the alpha radionuclides investigated can be consequently replaced by a statistical factor connecting its content with another alpha radionuclide on using a mathematical model.

Conclusion

Determination of the activity of a radionuclide in contaminated soil may be accomplished by indirect methods, such as the use of a scaling model, relating the inferred activity concentration of one radionuclide to another that is measured. ^{241}Am and ^{239,240}Pu were selected as suitable key nuclides for determination of the parameters of linear scaling model for monitoring radionuclides in the contaminated lysimeter soil. The mechanism ^{239,240}Pu is similar to that of ^{241}Am and ^{238}Pu, which have also similar chemical and physical properties. The scaling model was calculated with ^{239,240}Pu as the indicator variable for the estimated radionuclides ^{238}Pu, ^{241}Am, or with ^{241}Am as the indicator variable for the estimated radionuclides ^{239,240}Pu, ^{238}Pu. The ^{239,240}Pu fraction is the most prominent representative of the alpha-emitters. The Fisher-Snedecor significance test of regression confirmed the linear dependence between alpha radionuclides ^{239,240}Pu, ^{238}Pu and ^{241}Am. The linear dependence between the beta radionuclide ^{90}Sr and alpha radionuclides ^{239,240}Pu, ^{241}Am was not confirmed, Fisher-Snedecor's test did not accept the scheme as significant. The low correlation observed was due to the significant chemical difference and thus no scaling model for ^{90}Sr vs. ^{239,240}Pu, ^{241}Am was used. By fulfilling conditions corresponding with regression triplet components, the least squares method provided for finding a suitable scaling model for ^{241}Am, ^{238}Pu, ^{239,240}Pu alpha radionuclides monitoring in contaminated soils of the experimental lysimetric field. The regular designed scaling model opens possibilities of long-time activity monitoring of these radionuclides.

*

The financial support of the Ministry of Education of the Czech Republic (Grant No MSM0021627502) is gratefully acknowledged. This work was supported by the Grant UK270/2004 and UK 426/2006 Comenius University.

References

1. Strategy and Methodology for Radioactive Waste Characterization, IAEA-TECDOC-1537, ISBN 92-0-100207-6, IAEA, Vienna, 2007.
2. IAEA Technical Report Series Nr. 462, Managing of Low Radioactivity Material from the Decommissioning of Nuclear Facilities, IAEA Vienna, 2008.
3. X. HOU, P. ROOS, Critical Comparison of Radiometric and Mass Spectrometric Methods for the Determination of Radionuclides in Environmental, Biological and Nuclear Waste Samples, Postprint, 2007, http://www.risoe.dk/rispubl/art/2007_333.pdf.
4. K.S. REDUS, et al., Scaling Ratios Produce Misleading Results in D&D Projects, WM'05 Conference, Tucson, February 27-March 3, 2004.
5. ISO 21238:2007 Nuclear Energy-Nuclear Waste-Standard Guide for the Scaling Factor Method to Determine the Radioactivity of Low and Intermediate Level Radioactive Waste Packages Generated at Nuclear Power Plant (LWR), International Standard.
6. Updated Scaling Factors in Low-Level Radwaste, EPRI NP–5077, Project 1557-6, Final Report, 1987.
7. Á. VINCZE, I. GRESITS, S. TÖLGYESI, E. ERDŐS, J. SOLYMOSI P. ORMAI, A. FRITZ, Application of the Scaling Technique for The Characterisation of Different Radioactive Waste at NPP Paks, Radiation Protection in Neighbouring Countries of Central Europe, Prague, Czech Republic, September 8–12, 1997.
8. E. HERTELENDI, Z. SZÜCS, J. CSONGOR, J. GULYÁS, É. SVINGO, P. ORMAI, A. FRITZ, J. SOLYMOSI, I. GRESITS, N. VAJDA, ZS. MOLNÁR, Application of Scaling Technique for Estimation of Radionuclide Inventory in Radioactive Waste. 3rd Regional Meeting Nuclear Energy in Central Europe. Portoroz, Slovenia, 16–19 September, 1996.
9. D. HARKER, Scaling Factor for Waste Activities Measured by G-M Method, ER-WAG7-57, INEL – 95/020.
10. M. NOÉ, W. MÜLLER, R. GENS, M. GILI, A. MORALES, A. YATES, P. ORMAI, Development of methods to provide an inventory of radiologically relevant radionuclides. Analytical methods and correlation of data. Final report (Contracts FI2W-CT90-0034 and FI2W-CT91-0109) EUR 17978, 1998, Euroffice.
11. M. MELOUN, J. MILITKÝ, M. FORINA, Chemometrics for Analytical Chemistry, Vol. 2. PC-Aided Regression and Related Methods, Horwood, Chichester, 1994.
12. D. A. BELSEY, E. KUH, R. E. WELSCH, Regression Diagnostics: Identifying Influential Data and Sources of Collinearity, Wiley, New York, 1980.
13. R. D. COOK, S. WEISBERG, Residuals and Influence in Regression, Chapman & Hall, London, 1982.
14. A. C. ATKINSON, PLOTS, Transformations and Regression: An Introduction to Graphical Methods of Diagnostic Regression Analysis, Claredon Press, Oxford, 1985.
15. S. CHATTERJEE, A. S. HADI, Sensitivity Analysis in Linear Regression, Wiley, New York, 1988.
16. V. BARNETT, T. LEWIS, Outliers in Statistical Data, 2nd Ed., Wiley, New York, 1984.
17. R. E. WELSCH, Linear Regression Diagnostics, Technical Report 923–77, Sloan School of Management, Massachusetts Institute of Technology, 1977.
18. R. E. WELSCH, S. C. PETERS, Finding Influential Subsets of Data in Regression Models, Proc. Eleventh Interface Symp. on Computer Science and Statistics, A. R. GALLANT, T. M. GERIG, (Eds), Raleigh: Institute of Statistics, North Carolina State University, 1978.
19. S. WEISBERG, Applied Linear Regression, Wiley, New York, 1985.

20. P. J. ROUSSEEUW, A. M. LEROY, Robust Regression and Outlier Detection, Wiley, New York, 1987.
21. K. A. BROWNLEE, Statistical Theory and Methodology in Science and Engineering, Wiley, New York, 1965.
22. M. MELOUN, J. MILITKÝ, M. HILL, R. G. BRERETON, The Analyst, 127 (2002) 433.
23. M. MELOUN, J. MILITKÝ, Anal. Chim. Acta. 439 (2001) 169.
24. D. X. WILLIAMS, Applied Statistics, 22 (1973) 407.
25. J. B. GRAY, Graphics for Regression Diagnostics, Proc. of the Statistical Computing Section, Amer. Statist. Assoc., 1985, pp.102–107.
26. ADSTAT (English version), TriloByte Statistical Software, Pardubice, 1999.
27. R. D. COOK, S. WEISBERG, Diagnostics for Heteroscedasticity in Regression, Biometrika, 70, 1983.
28. I. MATUŠEK, KOVACS, Evaluation of the results of lysimetric experiment during the period 2004, Ecosur Report, 2005.
29. L. MÁTEL, V. MIKULAJ, P. RAJEC, J. Radioanal. Nucl. Chem.,175 (1993) 41.
30. C. W. SILL, Nucl.Chem.Waste Managem., 7 (1987) 201.
31. F. D. HINDMAN, Anal.Chim., 55 (1983) 2460.
32. V. MIKULAJ, V. ŠVEC, J. Radioanal. Nucl. Chem., 175 (1993) 317.
33. M. AFSAR, H. SCHUTTELKOPF, Determination of Am-241, Cm-244 in Environmental Samples, KfK 4346, 1988.