Indian Journal of Advanced Chemistry (IJAC)
ISSN: 2582-8975 (Online), Volume-5 Issue-2, October 2025

Introductory Statistical Methods for Radiological Characterization of Radioactive Waste

Rayna Hristova, Rositsa Peycheva, Stefan Simovski, Ilko Mladenov

Abstract: ***This paper provides a comprehensive overview of statistical methods employed in sample-based radiological characterization of radioactive waste (RAW), with a particular focus on the use of nuclide vectors (NVs) and scaling factors (SFs) as applied in commercial RAW characterization projects. These methods are crucial for estimating the activity of difficult-to-measure (DTM) radionuclides by establishing correlations with easy-to-measure (ETM) key nuclides (KNs), thereby minimizing the need for time-consuming and costly radiochemical analyses. A scaling factor (SF) is defined as the ratio of the activity (or specific activity) of a DTM to that of a corresponding KN in a given sample. The applicable standard deviation (SF) is typically determined as the geometric mean of the standard deviations (SDs) calculated from all samples, providing a robust and statistically representative value. The nuclide vector (NV) represents the relative distribution of individual radionuclides within the total activity of a sample or waste stream. NVs are recommended to be derived using the one-sigma concept, which assumes that approximately 68% of all possible values fall within a defined acceptance range, improving statistical confidence. For NVs and SFs to be valid, the underlying datasets must meet several criteria: they must be representative, span a wide range of activity levels, and be statistically homogeneous, meaning they follow a standard or log-normal distribution. Additionally, datasets must be free from significant outliers, typically identified using the Grubbs test, and show adequate correlation between radionuclides, assessed via Pearson or Spearman correlation coefficients. The methodology is demonstrated using data from 10 samples containing Mn-54, Co-60, Nb-94, Fe-55, Ni-63, and Sr-90. Results confirm that the calculated NVs and SFs are statistically valid and representative, supporting their practical application in modern RAW characterization.***

Keywords: ***Radioactive Waste, Statistical Characterization, Nuclide Vectors, Scaling Factors, Correlation Coefficients.***

Nomenclature:

$\bar{A}$	The Average Value of the Numbers of KN in the Sample
$\bar{B}$	the Average Value of the Numbers of DTM in the Sample
$\overline{SF_{AM}}$	Arithmetic Mean of SF
$\overline{SF_{GM}}$	Geometric Mean of SF
$\bar{v}$	Average Value of the Nuclide Vector
$\bar{x}$	Sample Mean of ETM, KN, Bq, Bq/kg
$\bar{y}$	Sample Mean of DTM, Bq, Bq/kg
$\bar{b}_i$	Average Proportion of the i^{th} Nuclide
a, a_i	Activity, Specific Activity of Nuclide i, Bq, Bq/kg
A_i	Rank of KN in the Ordered Samples
$a_{i,n}$	Specific Activity of the i^{th} Nuclide in the n^{th} Sample, Bq, Bq/kg
B_i	Rank of DTM in the Ordered Samples
c	The Proportionality Constant
D	Accepted Level of Difference
G	Grubbs Test Value
H0	Null Hypothesis
Ha	Alternative Hypothesis
M, M_1, M_2	Number of Samples
n	Number of the Current Sample
N	Number of Nuclides in the Sample
r_s	Spearman's Rank Coefficient
r_{xy}	Sample Pearson Correlation Coefficient
s, s_1, s_2	Sample Standard Deviations
s_b^2	Pooled Variance
SF	Scaling Factor
t	Student's t-Distribution
u	Uncertainty
V	Discrete Variable
v_i	Nuclide Vector, Proportion of the i^{th} Nuclide in a Sample
V_i	Averaged Nuclide Vector
X	set of KN Nuclides in the Sample
x_i	Value of KN, Bq, Bq/kg
x_l	Logarithm of KN
Y	set of DTM Nuclides in the Sample
y_i	Value of ETM, KN, Bq, Bq/kg
y_l	Logarithm of DTM

Greek Letters:

$\hat{\beta}_0$	Estimated Intercept
$\hat{\beta}$	Estimated Slope
μ	Mean Value
α	Proportionality Constant
β	Regression Coefficient
β_0	Estimated Slope
σ	Standard Deviation
γ	Gamma Ray

Subscripts:

AM	Arithmetic Mean
GM	Geometric Mean
DTM	Difficult to Measure Nuclide
KN	Key Nuclide
i	Nuclide Index
j	Index of KN
l	Activity Index on A Logarithmic Scale

Manuscript received on 29 May 2025 | First Revised Manuscript received on 16 June 2025 | Second Revised Manuscript received on 17 September 2025 | Manuscript Accepted on 15 October 2025 | Manuscript published on 30 October 2025.

*Correspondence Author(s)
Rayna Hristova*, Department of Radiation Protection, 16A Zlaten Rog Str., fl. Bulgaria. Email ID: r.hristova@atomtoploproekt.com, ORCID ID: 0009-0004-0693-5488
Rositsa Peycheva, Department of Radiation Protection, DIAL Ltd, Bulgaria, Sofia, Bulgaria. Email ID: office@dial-ltd.com
Stefan Simovski, General Manager, 16A Zlaten Rog Str., fl. Bulgaria.
Ilko Mladenov, General Manager, 16A Zlaten Rog Str., fl. Bulgaria.

n	Index of the Sample
s	Index of Spearman's Rank Coefficient
xy	Index of the Sample Pearson Correlation Coefficient
Abbreviations:	
AM	Arithmetic Mean
DTM	Difficult to Measure A Nuclide
ETM, KN	Easy to Measure Nuclide, Key Nuclide
GM	Geometric Mean
LOD	Limit of Detection
NPP	Nuclear Power Plant
NV	Nuclide Vectors
RAW	Radioactive Waste
RAW	Radioactive Waste
RCh	Radiochemistry
SF	Scaling Factor
SRS	Simple Random Sampling
SSC	Systems, Structures, and Components

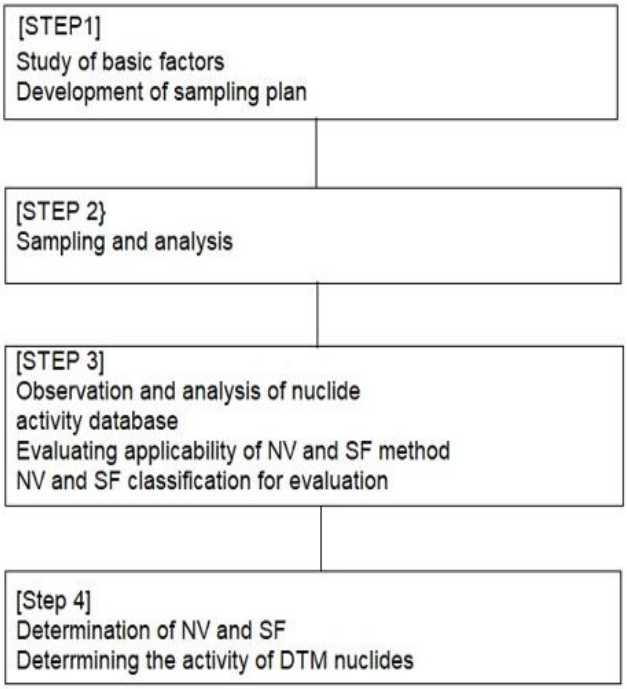
I. INTRODUCTION

Radioactive waste is generated in nuclear facilities during all operational stages of the facility. Methods for the radiological characterization of radioactive waste based on a limited number of samples have been introduced. For this purpose, the generated radioactive waste is divided into streams with similar radiological properties in terms of origin, generation mechanisms, operational history, contamination pathways, types of materials, etc. Nuclide vectors (NV) and scaling factors (SF), determined by statistical methods, are used for further characterization of radioactive waste. In this way, measurements of difficult-to-measure (DTM) nuclides are also greatly reduced. NV represents the activity ratio of a particular nuclide in the nuclide mixture. SF gives the proportion or linear relationship between KN and DTM activity. Their representativeness and validity must be periodically reassessed based on representative sampling. Representative sampling is a form of homogeneity or accumulated sampling. [1]. This article is structured as follows: 1. Brief presentation of the prerequisites for statistical characterization of RAW, including sampling plan and sampling methods; 2. Statistical methods for identifying outliers in the datasets; 3. Statistical methods for assessing the correlation between datasets; 4. Statistical methods for calculation of NVs and SFs; 5. Statistical methods for determining the applicability and representativeness of the resulting NVs and SFs. 6. Demonstration of statistical calculations of NVs and SFs based on measurement data.

II. APPLICATION OF STATISTICAL METHODS FOR THE DETERMINATION OF NV AND SF

A. Prerequisites for the Determination of NV and SF

A flowchart for the application of one of the most important statistical methods for SF estimation is presented in [2]:



[Fig.1: A Flowchart for the Application of One of the Most Important Statistical Methods for SF Estimation Based on [2]]

i. Study of Fundamental Factors

In general, nuclides are grouped according to their generation mechanism: fission products and activation products are dispersed by the coolant, as well as depending on the material properties of the structures, systems, and components (SSC) of a nuclear power plant (NPP). Of importance for deriving NVs and SFs are: historical knowledge and area mapping; selection of radionuclides for investigation; development of a sampling program; sampling measurements; selection of samples; preparation of samples; analysis; calculation of nuclide vectors and SFs; and reassessment of NVs and SFs.

ii. Sampling Plan

Object/stream data is of importance for the development of the Sampling plan. [3], e.g., facility type, material properties, mass, generation mechanism of radionuclides, activity level, homogeneity, and level of preliminary knowledge. The selection of samples also considers the historical understanding of the type of facility or plant, SSC, activated areas, and radionuclide transport routes; incidents; replacement of components; area investigation boundary; physical or chemical treatment, material properties; facility documents and data; interviews; and contamination hypotheses for area boundary establishment.

iii. Sampling and Selection of Samples for Radiochemistry (RCH)

The primary task of sampling is to collect sufficient samples to provide a representative

assessment of the spread of radioactive contamination, as well as the radionuclide contamination of materials. To achieve such representativeness, sampling should encompass possible contamination pathways as well as potential sources of variation in radioactive contamination.

In most practical situations, census data, which is information on all the units in a population, is either impossible or too expensive to collect. Simple random sampling (SRS) or "Hot spot" sampling are basic approaches often used to collect samples to estimate the actual value of a parameter of a population. Typically, the primary number of samples should be large, approximately 200-300. All samples undergo initial gamma spectrometry to be selected as representative samples. that have sufficiently high activity, so that the radiochemical analysis gives meaningful results. The selection of samples must satisfy several criteria. The first requirement is that the selected samples should have sufficiently high activity, such that the probability of obtaining a statistically insignificant pure signal in the radiochemical analysis is minimal. The second criterion is that the samples should be representative, i.e., cover a sufficiently wide range of activities to ensure the derivation of an adequate correlation between ETM and DTM. Lastly, the samples must be of an optimal number to avoid the risks of inadequate estimates of the relationship between the activities of KN and DTM nuclides. In SRS, each member of the population has an equal probability of being included in the sample [4]. This method is suitable for radioactive waste that has been proven to be homogeneous. Other sampling methods can also be found in the literature. Based on the radiological characterisation of selected representative samples, which demonstrate homogeneity, the applicability of NVs and SFs is statistically evaluated, as shown in this document.

B. Statistical Methods for Identifying Outliers in the Data Set

i. ISO-Approved Grubbs Test

Radiochemical data samples should be checked for outliers. The Grubbs test is a test used to detect a single outlier in the data that follows a normal distribution. Ion [5]. The random variable Y ($y_1, y_2, \ldots, y_i, \ldots, yM$) is supposed to be normally (Gaussian) distributed with mean and standard deviation σ.

The hypothesis (H0, Ha) and test statistic (G) for the Grubbs test are defined as follows:

H0: There are no outliers in the dataset

Ha: There is exactly one outlier in the dataset

$$G = \frac{\max_{i=1\ldots M} |y_i - \bar{y}|}{s} \quad \ldots \quad (1)$$

where $\bar{y}$ The mean of nuclide y in M samples, and s is the sample variance.

Grubbs's test detects one outlier at a time. This outlier is expunged from the dataset, and the test is iterated until no outliers are detected.

The sample variance s^2 is:

$$s^2 = \frac{1}{M-1}\sum_{1}^{M}(y_i - \bar{y})^2 \quad \ldots \quad (2)$$

where $\bar{y}$ Is the mean.

If the cause of the outlier can be identified, it should be corrected or removed with careful consideration, or alternative methods should be applied. Values of y_i with relatively minor differences give equal Grubbs's test results.

C. Correlation between Data Sets

Correlation in discrete data sets is assessed, for example, by the Sample Pearson correlation coefficient r_{xy}, which measures the linear correlation between two discrete variables, X and Y.

i. The Sample Pearson Correlation Coefficient (r_{xy})

$$r_{xy} = \frac{\sum_{1}^{M}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{1}^{M}(x_i - \bar{x})^2}\sqrt{\sum_{1}^{M}(y_i - \bar{y})^2}} \quad \ldots \quad (3)$$

The Pearson correlation coefficient r_{xy} has a value between +1 and -1, as determined by the Cauchy-Schwarz inequality, and it reflects the strength of a linear relationship. A value of 0 implies that there is no linear dependency between the variables. Correlations equal to +1 or −1 correspond to data points lying exactly on a line in the case of the sample correlation. To accept a correlation, an r_{xy} of 0.7 or better is required. An *Rxy* of less than 0.7 but better than 0.5 will be accepted if it is demonstrated that no systematic difference is responsible for the poor quality.

ii. Spearman's Rank Coefficient (rs)

Spearman's rank coefficient, rs, is defined as:

$$r_s = \frac{\sum_{1}^{M}(A_i - \bar{A})(B_i - \bar{B})}{\sqrt{\sum_{1}^{M}(A_i - \bar{A})^2}\sqrt{\sum_{1}^{M}(B_i - \bar{B})^2}} \quad \ldots \quad (4)$$

Where A_i is respectively the sequence number (rank) of x_i in the ordered sequence $x_1 < x_2 < \ldots < x_M$, similarly, B_i is the sequence number (rank) of y_i in the ordered sequence $y_1 < y_2 < \ldots < y_M$. The rank coefficient is a non-parametric estimate of the strength of the correlation (when the latter is statistically significant), in the sense that the values of the random variables X and Y are not used to calculate it. The rank correlation coefficient also assesses the significance of the correlation, and the value of r_s has the same meaning as the linear correlation coefficient.

D. Nuclide Vectors

NV, v_i, represents the relative ratios between the activity of a particular nuclide in the mixture and the total activity.

$$v_i = \frac{a_i}{\sum_{1}^{N} a_i} \quad \ldots \quad (5)$$

Where a_i *is* the activity of radionuclide i, vi is the

proportion of nuclide i in the nuclide vector, $_{and}N$ is the number of radionuclides in the current sample.

The determination of NV is based on a primary experimental determination of the correlation of ETM and DTM. Gamma-emitting radionuclides, e.g., C0-60, Cs-137, Sb-124, Ce-144, etc., which are easy to measure using gamma spectrometry, are typically identified as KN in the radionuclide mixture in RAW. KN are usually selected based on the nuclear mechanism of radionuclide generation. For fission products, usually the KN Cs-137, and for activation products, Co-60.

In [3] It is recommended to use the one sigma concept for the determination of nuclide vector, where 68% of all possible samples are in the area defined as the acceptance criterion:

$$(\bar{v} - \sigma) \leq v_i \leq (\bar{v} + \sigma) \quad \ldots \quad (6)$$

Where

$\bar{v}$ The mean value of measurements, v_i; σ is the standard deviation, i.e., standard uncertainty in measurements by considering all uncertainties; v_i is the measured value.

For a discrete random variable V, which takes random values from a finite data set $v_1, v_2, \ldots, v_i \ldots v_M$, each value having the same probability, the standard deviation σ is defined as:

$$\mu = \frac{1}{M}\sum_{1}^{M} v_i \quad \ldots \quad (7)$$

$$\sigma = \sqrt{\frac{1}{M}\sum_{1}^{M}(v_i - \mu)^2} \quad \ldots \quad (8)$$

In this case μ= $\bar{v}$

- About 68% of values drawn from a normal distribution are within one standard deviation, one σ away from the mean;
- about 95% of the values lie within two standard deviations;
- and about 99.7% are within three standard deviations [2].
- This fact is known as the 68-95-99.7 (empirical) rule or the 3-sigma rule.

The mean value is accepted as the nuclide vector for radionuclide i if the applicable acceptance criterion (typically the one sigma concept) is satisfied and representativeness is demonstrated.

E. Approaches to the Composition of Nuclide Vectors

The composition of the NV is based on the results of radiochemical analyses of the samples taken for radiological characterization. There are three main approaches to the compilation of NV:

- Compilation of a covering note;
- Compilation of NV by averaging;
- Compilation of NV based on statistical analysis.

i. Covering NV

When compiling the covering NV, the highest activity fractions of nuclides (v_i) that are not KNs are selected from all analysis results. The rest is assigned to the KNs. This leads to a significant overestimation of difficult-to-measure nuclides (alpha, beta emitters) and an underestimation of key nuclides. In this way, the radiological significance of DTM is emphasised, and the covering NV becomes conservative concerning clearance from regulation.ion [3]. The use of covering NV for clearance from regulation is conservative because it may lead to a significant overestimation of the activity of the materials, which in turn may lead to falsely exceeding regulatory release levels.

ii. NV by Averaging

In this approach, the NV is estimated by averaging the proportions of the corresponding radionuclide over all samples in which it was identified. If v_i denotes the proportion of the i^{th} nuclide in sample n, i.e.:

$$v_i = \frac{a_i}{\sum_1^N a_i} \quad \ldots \quad (9)$$

Where a_i is the specific activity of the i^{th} nuclide in the n^{th} sample, N is the number of significant nuclides in the sample. Then the average proportion of the ith nuclide ($\bar{b}_i$) is given by:

$$\bar{b}_i = \frac{v_i}{\sum_1^M v_i} \quad \ldots \quad (10)$$

Where M is the number of samples in which the i^{th} nuclide was identified. Accordingly, the elements of the NV when compiled by averaging are given by:

$$\mathrm{v}_i = \frac{\bar{b}_i}{\sum_1^N \bar{b}_i} \quad \ldots \quad (11)$$

The specific activities of the nuclides in the samples ($a_{i,n}$) are expected to vary widely, which is primarily driven by the sampling process (striving to sample over a wide range of specific activities), while the proportions of radionuclides in the sample ($b_{i,n}$) is expected to vary over a much narrower interval [3]. The two quantities $a_{i,n}$ and $b_{i,n}$ also have different probability distributions.

NVs determined by averaging are easily calculated. Still, before their evaluation, it is advisable to check that the results to be analyzed are homogeneous (e.g., by a statistical test of the results for belonging to a standard or log-normal distribution). The ratios between the radionuclides in this type of NV reflect the experimentally observed ratios between the radionuclides and reflect the average radionuclide composition of the contamination.

iii. NV based on Statistical Analysis

When compiling NV based on statistical analysis, the proportions (average proportions) of nuclides are allowed to vary within

predetermined limits. In [6], these limits are assumed to be in a 1σ interval around the mean.

In general, by varying the values of *NV,* one aims to maximize several criteria, which are typically set depending on the chosen approach for clearance from regulation [3]. These criteria can be:

- Maximize the amount set by the clearance criterion;
- Maximizing proportions of radionuclides that are difficult to measure (e.g., pure alpha- and/or beta-emitters, as appropriate);
- Maximization of the amount set in the release criterion when a surface activity criterion for release is also foreseen.

The above three criteria often cannot be met simultaneously, as the requirements are sometimes conflicting. For this reason, an optimization algorithm is used to achieve a high proportion of the upper maxima with a nuclide vector by summing the three sub-goals. The sum of the three individual target variables serves as the optimization target value.

By solving the optimization problem, a radionuclide vector can be achieved that is conservative, with a certain level of conservatism [3].

iv. Representative, Covering, and Conservative NVs

Depending on the method of composition, NVs can provide varying degrees of representativeness and conservatism.

- The nuclide vector determined by averaging is representative because it reflects the composition of the contamination and the average proportions of radionuclides in it. However, it may not be conservative since it does not increase the a priori weight of nuclides that have low release levels;
- The covering nuclide vector is conservative because it purposefully increases the weight of DTM nuclides that have low release levels. For this reason, however, it is not representative of the radionuclide composition of the contamination.
- The statistical basis calculation results in nuclide vectors that lie between the two options above. This means that, on the one hand, they do not lead to such a substantial overestimation of difficult-to-measure radionuclides, and on the other hand, they give them a certain additional weight. In this regard, these nuclide vectors can be considered representative and conservative.

The variant of the radionuclide vectors to be used in the procedure for clearance from regulation depends exclusively on the composition and nature of the radioactive contamination. Firstly, this is determined by the radionuclides and their proportions present in the contaminated materials. The number and type of radionuclide vectors that should be calculated for a given facility also depend on the homogeneity of the radionuclide composition of the contamination. Regarding this, the choice of method for evaluating the NV should not be made a priori, but rather estimated after a detailed analysis of the data from the radiochemical analyses.

F. Scaling Factors

As said above, typically, strong γ-emitting radionuclides are selected as KN for NPPs. Half-life and time parameters of the radioactive materials must also be considered. The specific activity of alpha- and beta-emitting, as well as low-energy gamma-emitting, radionuclides, expressed in Bq/kg or Bq/cm² (or activity, Bq), is determined mainly using destructive radiochemical methods on samples collected by wipe tests, electrolytic sampling, and scraping sampling. These radionuclides are DTM.

An SF is introduced for determining the activity-specific proportion of activity between specific KN and DTM.

$$a(DTM_i) = SF_{DTM_i} \times a(KN_j) \quad \ldots \quad (12)$$

The SFs are dimensionless factors for the different DTMs, which enable the determination of total activity by measuring the activity of the respective KN. Once determined, the SF allows determination of the activity of a particular DTM in a batch only by measuring the activity of the corresponding KN.

From a statistical perspective, only radiochemical data above the limit of detection (LOD) should be considered to determine the applicability of SF and to calculate SF. However, in some cases, due to a lack of sufficient radiochemical data, one has no choice but to use the LOD value itself as the accurate radioactivity concentration. Decision-making regarding the use of radiochemical data below the limit of detection (LOD) is required for resampling and radiochemical reanalysis.

The relationship between the radioactivity concentration of DTM nuclides and KN can be more generalized based on the nonlinear relationship as follows. Ows [5]:

$$a_{DTM,i} = c(a_{KN,j})^{\beta} \quad \ldots \quad (13)$$

Where c is the proportionality constant and β is the regression coefficient. In the special case where β equals 1, it becomes a simple linear equation, as mentioned above. If β is not equal to 1, this simple nonlinear model is a simple linear equation on a logarithmic scale.×

$$y_l = \beta_0 + \beta x_l \quad \ldots \quad (14)$$

where $y_l = \log(a_{DTM})$, $x_l = \log(a_{KN})$.

Two parameters, the intercept (β_0) and slope (β), in the simple linear equation are generally estimated by the least-squares method. The least-squares method is a standard approach in regression analysis that minimizes the residual sum of squares. The estimated intercept ($\hat{\beta}_0$) and slope ($\hat{\beta}$) are given as follows:

$$\hat{\beta} = \frac{\sum_{i=1}^{M}(x_{l,i} - \bar{x}_l)(y_{l,i} - \bar{y}_l)}{\sum_{i=1}^{M}(x_{l,i} - \bar{x}_{\ l})^2} \quad \ldots \quad (15)$$

$$\hat{\beta}_0 = \bar{y}_l - \hat{\beta}\bar{x}_l \quad \ldots \quad (16)$$

i. Representative Scaling Factor

The representative SF is calculated as the arithmetic mean (AM) $\overline{SF_{AM}}$ or geometric mean (GM) $\overline{SF_{GM}}$ as follows [7]:

$$\overline{SF_{AM}} = \frac{1}{M}\sum_{1}^{M} SF_i \quad \ldots \quad (17)$$

$$\overline{SF_{GM}} = (\prod_1^M {}^{SF_i} = \log^{-1}(\frac{1}{M}\sum_{1}^{M}\log SF_i) = 10^{(\frac{1}{M}\sum_1^M \log SF_i)} \quad \ldots \quad (18)$$

The radioactivity concentration calculated by the arithmetic mean of SF always yields more conservative values, and the predicted concentration given by the geometric mean is much more severely overestimated in the higher-concentration ranges.

A statistical hypothesis test within the acceptable level of difference (D) between the existing SF_1 and the updated SF_2 can be performed based on the pooled variance (s_b^2). The hypotheses and test statistic (t) under the null hypothesis follow Student's t-distribution.a logarithmic scale with degrees of freedom $M_1 + M_2 - 2$: ($t_{M1+M2-2}$) are defined as follows [5]:

$$H_0: |SF_1 - SF_2| = \log D$$
$$Ha: |SF_1 - SF_2| \neq \log D$$
$$t = \frac{|SF_1 - SF_2| - \log D}{s_b\sqrt{\frac{1}{M_1} + \frac{1}{M_2}}} \approx t_{(M1+M2-2)} \quad \ldots \quad (19)$$

$$s_b^2 = \frac{(M_1 - 1)s^2{}_1 + (M_2 - 1)s_2^2}{M_1 + M_2 - 2} \quad \ldots \quad (20)$$

where SF_1 and SF_2 are geometric means, M_1 and M_2 are the number of samples, and s_1^2 and s_2^2 are the sample variances. D=1 means SF_1=SF_2 because the values of SF are log-transformed [5].

If the null hypothesis (H_0 $\beta_1 \approx 1$, where β_1 is the slope of the SF according to Eq. (13), is true, it is not necessary to update the SF because it cannot be said that the SF has changed over time.

The concept of "factor of 10" is applied. The factor of 10 is defined as:

$$\frac{1}{10}\overline{SF}_{GM} \leq SF_i \leq 10\overline{SF}_{GM} \quad \ldots \quad (21)$$

or

$$\frac{1}{10}a_{DTM,i} \leq \text{~~a~~}_{DTM,i} \leq 10a_{DTM,i} \quad \ldots \quad (22)$$

where $\text{~~a~~}_{DTM,i} = \overline{SF}_{GM} \times a_{KN,i}$ Is the inferred (i.e., calculated) radioactivity concentration, and $a_{DTM,i} = SF \times a_{KN,j}$ is the measured radioactivity concentration [5].

For the application of the factor of 10, the outliers must be identified appropriately.

ii. Uncertainty Assessment

Uncertainty, for example, in the determination of scaling factors determination may be assessed as follows. Ows [8]:

$$u_{SF} = \frac{1}{\sqrt{M}}s(SF) \quad \ldots \quad (23)$$

The sample variance is equal to

$$s(SF) = \sqrt{\frac{1}{M-1}\sum_{1}^{M}(SF_i - \bar{SF})^2} \quad \ldots \quad (24)$$

where $\overline{SF}$ It is the mean value.

This provides a measure of the width of the distribution of mean values that would be expected and is called the standard uncertainty of the mean.

III. RESULTS

A. Determination of Nuclide Vectors

The data for 10 samples with specific activity measurements of Mn54, Co60, Nb94, Fe55, Ni63, and Sr90 are given in Table. After performing the Grubbs test according to Eq. 1, we identified outliers in the last sample. For the correct estimation of NV, the last sample was removed from further estimation. The NVs v_i, the results of the one-sigma concept check, the mean value $\mu=\bar{v}$ and the standard deviation σ of v_i, the average NVs $\bar{b}_l$ V_i are also shown in Table.

The conclusion is that the statistically estimated NVs, vi, and Vi, satisfy the one-sigma concept 100% for all nuclides in a sample, both for v_i and the averaged NVs. All NVs are within the area defined as the acceptance criterion according to Eq. 6 and are applicable, as are their mean values. However, the concept of one sigma is also fulfilled for the last removed sample. The results show that the averaged NVs V_i generally have a smaller span compared to Vi. As stated above, the choice of method for evaluating the NV should not be made a priori, but rather evaluated after a detailed analysis of the radiochemical data has been conducted.

Table-I: Data from Measurement of Radionuclides' Specific Activity

Specific Activity of Radionuclides, Bq/kg						vi According to Eq.5						Value of Eq. 6 for all Nuclides in the Sample
Mn54	**Co60**	**Nb94**	**Fe55**	**Ni63**	**Sr90**	**Mn54**	**Co60**	**Nb94**	**Fe55**	**Ni63**	**Sr90**	
1.07E+02	3.56E+03	2.11E+01	6.14E+03	1.65E+03	5.41E+01	9.27E-03	3.09E-01	1.83E-03	5.32E-01	1.43E-01	4.69E-03	TRUE
2.88E+01	1.86E+03	1.57E+01	2.51E+03	2.13E+02	8.41E+00	6.22E-03	4.02E-01	3.39E-03	5.41E-01	4.60E-02	1.82E-03	TRUE
3.37E+04	1.17E+06	2.82E+02	1.16E+06	5.74E+04	2.05E+02	1.39E-02	4.84E-01	1.17E-04	4.78E-01	2.37E-02	8.47E-05	TRUE
2.24E+04	5.17E+05	2.55E+02	4.84E+05	2.62E+04	3.21E+02	2.13E-02	4.92E-01	2.43E-04	4.61E-01	2.49E-02	3.06E-04	TRUE
2.23E+04	4.11E+05	1.78E+02	5.87E+05	2.97E+04	1.37E+02	2.12E-02	3.91E-01	1.70E-04	5.59E-01	2.83E-02	1.30E-04	TRUE
9.23E+03	2.54E+05	6.68E+02	8.94E+05	1.28E+05	6.06E+02	7.18E-03	1.97E-01	5.19E-04	6.95E-01	9.94E-02	4.71E-04	TRUE
2.20E+05	5.54E+06	2.51E+02	6.91E+06	2.81E+05	1.08E+02	1.70E-02	4.28E-01	1.93E-05	5.34E-01	2.17E-02	8.34E-06	TRUE
2.31E+02	9.46E+03	1.53E+01	1.13E+04	1.84E+03	1.59E+02	1.00E-02	4.11E-01	6.67E-04	4.91E-01	8.00E-02	6.91E-03	TRUE
5.62E+04	2.85E+05	1.83E+03	6.29E+05	1.05E+04	1.03E+02	5.72E-02	2.90E-01	1.86E-03	6.40E-01	1.06E-02	1.05E-04	TRUE
2.50E+06	1.79E+08	2.91E+03	3.05E+08	9.83E+07	6.41E+03							TRUE
$\mu = \vec{v}$ according to Eq.7												
1.82E-02	3.78E-01	9.79E-04	5.48E-01	5.31E-02	1.61E-03							
σ according to Eq. 8												
1.48E-02	9.05E-02	1.08E-03	7.18E-02	4.23E-02	2.36E-03							
$\bar{b}_i$ according to Eq. 10						V_i according to Eq.11						
5.68E-02	9.06E-02	2.08E-01	1.08E-01	3.00E-01	3.23E-01	5.23E-02	8.34E-02	1.91E-01	9.94E-02	2.76E-01	2.97E-01	TRUE
3.81E-02	1.18E-01	3.84E-01	1.10E-01	9.61E-02	1.25E-01	4.37E-02	1.35E-01	4.41E-01	1.26E-01	1.10E-01	1.43E-01	TRUE
8.53E-02	1.42E-01	1.32E-02	9.70E-02	4.96E-02	5.84E-03	2.17E-01	3.61E-01	3.36E-02	2.47E-01	1.26E-01	1.49E-02	TRUE
1.31E-01	1.45E-01	2.76E-02	9.35E-02	5.22E-02	2.10E-02	2.79E-01	3.08E-01	5.87E-02	1.99E-01	1.11E-01	4.47E-02	TRUE
1.30E-01	1.15E-01	1.92E-02	1.13E-01	5.92E-02	8.99E-03	2.92E-01	2.58E-01	4.31E-02	2.54E-01	1.33E-01	2.02E-02	TRUE
4.39E-02	5.80E-02	5.89E-02	1.41E-01	2.08E-01	3.24E-02	8.10E-02	1.07E-01	1.09E-01	2.60E-01	3.84E-01	5.98E-02	TRUE
1.04E-01	1.26E-01	2.20E-03	1.08E-01	4.54E-02	5.74E-04	2.69E-01	3.26E-01	5.70E-03	2.80E-01	1.18E-01	1.49E-03	TRUE
6.15E-02	1.21E-01	7.55E-02	9.96E-02	1.67E-01	4.76E-01	6.15E-02	1.21E-01	7.55E-02	9.95E-02	1.67E-01	4.76E-01	TRUE
3.50E-01	8.52E-02	2.11E-01	1.30E-01	2.23E-02	7.22E-03	4.34E-01	1.06E-01	2.62E-01	1.61E-01	2.77E-02	8.96E-03	TRUE

1.1. Determination of Scaling Factors

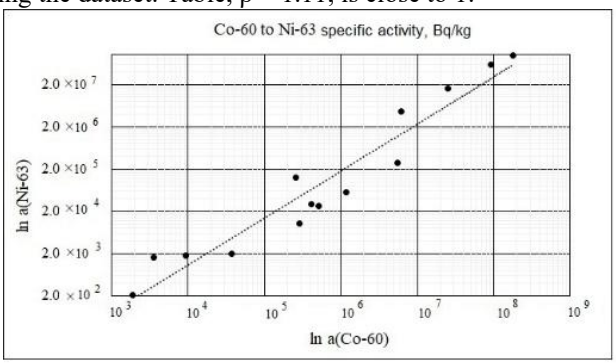
Table gives data from field measurements of the specific activity of DTM Ni-63 and KN Co-60 in 14 samples and provides calculated values for the Sample Pearson correlation coefficient according to Eq. 3, for SF according to Eq. 12, for $\overline{SF}_{GM}$ according to Eq. 18 and for the applicability of the factor of 10 concept for GM, according to Eq. 21.

After evaluating the radiochemical data, the correlation between the DTM and KN measurements was assessed.

Table-II: DTM Ni-63 and KN Co-60 Measured Specific Activity, Sample Pearson Correlation Coefficient, SF, $\overline{SF}_{GM}$ and the Value of the Equation Describing the Factor of 10 Concept

Co60	Ni63	r_{xy} according to Eq. 3	$SF_{Ni-63/Co-60}$ according to Eq. 12	$\overline{SF}_{GM}$ according to Eq. 18	Factor of 10 according to Eq. 21
Bq/kg	Bq/kg	-	-	-	-
3.56E+03	1.65E+03	0.9979	4.65E-01	0.1693	TRUE
1.86E+03	2.13E+02		1.14E-01		TRUE
1.17E+06	5.74E+04		4.90E-02		TRUE
5.17E+05	2.62E+04		5.07E-02		TRUE
4.11E+05	2.97E+04		7.23E-02		TRUE
2.54E+05	1.28E+05		5.03E-01		TRUE
5.54E+06	2.81E+05		5.08E-02		TRUE
9.46E+03	1.84E+03		1.95E-01		TRUE
2.85E+05	1.05E+04		3.67E-02		TRUE
1.79E+08	9.83E+07		5.49E-01		TRUE
9.27E+07	5.73E+07		6.18E-01		TRUE
2.50E+07	1.64E+07		6.56E-01		TRUE
6.10E+06	4.57E+06		7.49E-01		TRUE
3.73E+04	2.03E+03		5.44E-02		TRUE

Studies on radiochemical data from nuclear power plants have shown that the radioactivity concentrations of both DTM nuclides and KN follow a log-normal distribution, with a wide range of radioactivity concentrations spanning several orders of magnitude. 5. The correlation is observed in the scatter diagram in [Fig.2, which uses radiochemical data, with a linear trend line. The regression coefficient in Equation 13 was calculated using the dataset. Table, β = 1.11, is close to 1.



[Fig.2: Log-Normal Distribution of DTM Ni-63 and KN Co-60 Specific Activity on a Scatter Diagram]

Correlation is demonstrated by calculating the Pierson correlation coefficient, as shown in Table.

The results in Table and [Fig.2 show a log-normal distribution between the activity/specific activity of DTM Ni-63 and KN Co-60, and a very good correlation coefficient r_{xy} between DTM and KN. The regression coefficient β is close to 1. The SFs comply with the concept of factor 10. The $\overline{SF}_{GM}$ is. representative and applicable.

IV. DISCUSSION AND CONCLUSION

This article discusses statistical methods for estimating nuclide vectors and scaling factors to characterize radioactive waste radiologically. Once established and deemed representative, NVs and SFs can significantly reduce or eliminate the need for radiological measurements of each waste batch. This is particularly beneficial in avoiding the time-consuming and expensive destructive radiochemical methods used for measuring the activity of radionuclides that are difficult to measure (DTM). However, changes in operational history and practices can lead to variations in radiological composition and waste streams.

Depending on the method of composition, NVs can vary in their representativeness and conservativeness. The NV derived from averaging is generally representative but may not be conservative. On the other hand, a conservative covering NV may not accurately represent the radionuclide composition of the contamination. Statistically calculated NVs, however, can be both representative and conservative.

Using measurements of the specific activity of a range of radionuclides in samples, NVs were statistically calculated. Average NVs were estimated, and outliers were identified and removed. The averaged NVs exhibit a smaller range compared to those without averaging, indicating that the averaged NVs yield closer values across different samples. The results align with the one-sigma concept for both NVs and averaged NVs.

The application of SF is also a contemporary method for characterising raw data. SF represents the ratio between the activity or specific activity of DTM and KN.

This relationship is expected to follow a simple linear model or linear equation on a logarithmic scale, characterized by slope β. If the value of β is close to 1, the relationship between DTM and KN can be described by a simple linear model with a proportionality constant equivalent to SF.

The representative standard deviation (SD) is calculated using either the arithmetic mean (AM) or the geometric mean (GM). The radioactivity concentration calculated by the arithmetic mean of SF tends to yield more conservative values. In contrast, the predicted concentration obtained from the geometric mean is often severely overestimated in higher concentration ranges. The validity of SF is established when there is a significant correlation between DTM and KN of the measured homogenized samples. This includes having a slope β in the linear equation close to 1 and the absence of unexplained outliers.

An example demonstrating the suitability of statistical methods for SF estimation is provided, which includes Table and [Fig.2. The calculated results indicate a log-normal distribution between the activity and specific activity of DTM Nickel-63 and KN Cobalt-60, showing a very good correlation coefficient (r_{xy}), a regression coefficient β close to 1, and compliance with the factor of 10 concept, thereby proving the applicability of the evaluated dataset.

However, it is important to note that other datasets may yield inappropriate results. The waste batch should be assessed for homogeneity, proper sampling, coverage of expected contamination pathways, potential sources of radioactivity variation, and generation mechanism. This necessitates collecting enough samples and measurements.

In summary, the application of statistical methods for estimating NV and SF in the characterisation of RAW is a reliable and modern approach when conducted correctly, and it is time-efficient.

DECLARATION STATEMENT

After aggregating input from all authors, I must verify the accuracy of the following information as the article's author.

- **Conflicts of Interest/Competing Interests: Based on my understanding, this article does not have any** conflicts of interest.
- **Funding Support:** This article has not been funded by any organizations or agencies. This independence ensures that the research is conducted with objectivity and without any external influence.
- **Ethical Approval and Consent to Participate:** The content of this article does not necessitate ethical approval or consent to participate with supporting documentation.
- **Data Access Statement and Material Availability:** The adequate resources of this article are publicly accessible.
- **Author's Contributions:** The authorship of this article is contributed equally to all participating individuals.

REFERENCES

1. ISO 21238, 2007, International standard first edition, nuclear energy-Nuclear Fuel Technology – Scaling Factor Method to Determine the Radioactivity of Low- and Intermediate-Level Radioactive Waste Generated at Nuclear Power Plants, ISO 21238: 2007(E), pp.1-27, https://www.dinmedia.de/en/standard/iso-21238/99669709
2. IAEA-TECDOC-1537 2007, Strategy and Methodology for Radioactive Waste Characterization, Vienna, Austria, pp. 1-178, https://www.iaea.org/publications/7655/strategy-and-methodology-for-radioactive-waste-characterization
3. Bothe, M., 2009, Empfehlungen zur Ermittlung der Repräsentativität von Nuklidvektoren bei Freigabemessungen: Vorhaben 3604S04441, Germany, urn:nbn:de:0221-2009011228, pp. 1-96, https://doris.bfs.de/jspui/bitstream/urn:nbn:de:0221-2009011228/1/BfS_2009_BfS-RESFOR-01-09.pdf
4. Zaffora, B., Magistris, M., Saporta, G. and Paolo La Torre, F., 2016, Statistical sampling applied to the radiological characterization of historical waste, EPJ Nuclear Sciences & Technologies, Volume 2, (34), France, https://doi.org/10.1051/epjn/2016031
5. Kim, T-H., Park, J., Lee, J., Kim, J., Kim, J-Y. and Lim, S., 2020, Statistical Methodologies for Scaling Factor Implementation: Part 1. Overview of Current Scaling Factor Method for Radioactive Waste Characterization, Journal of Nuclear Fuel Cycle and Waste Technology Volume 18 No.4, Korea, https://doi.org/10.7733/jnfcwt.2020.18.4.517, pp. 517-536.
6. Kim, A., Lietzmann, F., 2019, Nuclide Vector for Decommissioning and Release Measurements in Germany, Transactions of the Korean Nuclear Society Autumn Meeting, October 24-25, Korea, pp. 1-4, https://www.kns.org/files/pre_paper/42/19A-032-AndreaMaria.pdf
7. Hwang, K., Lee, S., Kang, S., Lee, K., Jeong, Ch., Ahn, S., Kim, T., Kim. K., Herr, Y., and Song, M., 2004, Development of Radionuclide Inventory Declaration Methods Using Scaling Factors for the Korean NPPs – Scope and Activity Determination Method, Journal of Nuclear Fuel Cycle and Waste Technology, Volume 2, (1), Korea, pp. 77-85, https://doi.org/10.7733/jnfcwt.2020.18.4.517
8. IAEA-TECDOC-1585, 2008, Measurement Uncertainty, A Practical Guide for Secondary Standards Dosimetry Laboratories, IAEA, Vienna, Austria; https://inis.iaea.org/records/9cxam-wkq60/preview/39101877.pdf

AUTHOR'S PROFILE

Rayna Hristova is a nuclear engineer with an MSc in Mechanical Engineering, specializing in Nuclear Engineering, from the Technical University of Sofia, Bulgaria. She began her career at Energoproekt Plc in Sofia, where she worked as a researcher and software developer, focusing on the reliability and safety assessment of nuclear systems. She also completed early training in Artificial Intelligence at the Bulgarian Academy of Sciences, as well as numerous specialised courses in nuclear safety provided by the IAEA, JICA, and other international organisations. Currently, she serves as the Head of the Radiation Protection Department at ATP-Atomtoploproekt Ltd. in Sofia. Her work encompasses nuclear safety analysis and verification, accident analysis, probabilistic safety assessments (PSA), fault tree and HAZOP studies, human factor evaluations, operational data processing, radiation protection, nuclear waste management, and related software development. She is the author of several publications on reliability analysis, human factors, and radiological impact assessment.

Rositsa Peycheva holds a Master's degree in Chemistry and Physics from Sofia University "St. Kliment Ohridski" (2002). With more than 15 years of professional experience in radiochemistry, analytical chemistry, and environmental monitoring, she specialises in the analysis of radioactive and chemical contaminants in industrial and ecological matrices. She currently works as a physicist at DIAL Ltd., where she performs high-level chemical and radiological analyses. Her professional expertise includes method validation and verification, quality assurance by ISO/IEC 17025, and project documentation for radiological assessments. She has contributed to numerous national and international projects involving environmental radioactivity, NORM/TENORM management, and decommissioning of nuclear facilities, including the Kozloduy NPP and the Radiana site. She has authored and co-authored several scientific presentations and peer-reviewed publications related to the rapid determination of radionuclides in water and sediment assessments.

Stefan Simovski is a Bulgarian engineer and project manager with a diverse background in nuclear energy, telecommunications, and business administration. He holds degrees in Mechanical Engineering (with a focus on Nuclear and Thermal Energy) from the Technical University of Sofia, an MSc from the University of National and World Economy, and a degree in Telecommunication Systems and Technologies from New Bulgarian University. Certified by the Bulgarian Chamber of PPP Engineers, Stefan has over two decades of experience in software development, system integration, and managing complex international projects. His work encompasses the design, installation, and commissioning of facilities for treating and conditioning solid radioactive waste, as well as instrumentation and control (I&C) and electrical systems for power plants. He has led multidisciplinary teams and coordinated projects of national and international significance.

Ilko Mladenov is an environmental and radiation protection expert with over 20 years of experience in environmental monitoring, radiological assessment, and management of radioactive materials, including NORM/TENORM. He holds an M.Sc. in Landscape Ecology from Sofia University and a B.Sc. in Environmental Engineering from the University of Mining and Geology. He has led and contributed to national and international projects on uranium mining remediation, site characterization, and regulatory development. As a certified radiation protection specialist, he has developed procedures for the release of materials from regulatory control and conducted environmental risk assessments. Currently Managing Director at DIAL Ltd., he oversees the Dosimetry and Environmental Radiochemistry Laboratory. Mr. Mladenov is an active contributor to IAEA technical meetings and regional workshops and co-author of several publications in radioecology and environmental safety.