Statistical Methodologies for Scaling Factor Implementation: Part 1. Overview of Current Scaling Factor Method for Radioactive Waste Characterization

Tae-Hyeong Kim[1], Junghwan Park[1,2], Jeongmook Lee[1], Junhyuck Kim[1], Jong-Yun Kim[1,3],*, and Sang Ho Lim[1,3],*

[1]*Korea Atomic Energy Research Institute, 111, Daedeok-daero 989beon-gil, Yuseong-gu, Daejeon, Republic of Korea*
[2]*Korea Advanced Institute of Science and Technology, 291, Daehak-ro, Yuseong-gu, Daejeon, Republic of Korea*
[3]*University of Science and Technology, 217, Gajeong-ro, Yuseong-gu, Daejeon, Republic of Korea*

The radionuclide inventory in radioactive waste from nuclear power plants should be determined to secure the safety of final repositories. As an alternative to time-consuming, labor-intensive, and destructive radiochemical analysis, the indirect scaling factor (SF) method has been used to determine the concentrations of difficult-to-measure radionuclides. Despite its long history, the original SF methodology remains almost unchanged and now needs to be improved for advanced SF implementation. Intense public attention and interest have been strongly directed to the reliability of the procedures and data regarding repository safety since the first operation of the low- and intermediate-level radioactive waste disposal facility in Gyeongju, Korea. In this review, statistical methodologies for SF implementation are described and evaluated to achieve reasonable and advanced decision-making. The first part of this review begins with an overview of the current status of the scaling factor method and global experiences, including some specific statistical issues associated with SF implementation. In addition, this review aims to extend the applicability of SF to the characterization of large quantities of waste from the decommissioning of nuclear facilities.

Keywords: Scaling factor, Radioactive waste, Waste characterization, Decision criteria, Statistical methodology

*Corresponding Author.
Jong-Yun Kim, Korea Atomic Energy Research Institute, E-mail: kjy@kaeri.re.kr, Tel: +82-42-868-4736
Sang Ho Lim, Korea Atomic Energy Research Institute, E-mail: slim@kaeri.re.kr, Tel: +82-42-868-2105

ORCID

Tae-Hyeong Kim	http://orcid.org/0000-0001-6483-5902	Junghwan Park	http://orcid.org/0000-0003-2909-3231
Jeongmook Lee	http://orcid.org/0000-0002-3496-5965	Junhyuck Kim	http://orcid.org/0000-0002-5372-0912
Jong-Yun Kim	http://orcid.org/0000-0002-9165-0986	Sang Ho Lim	http://orcid.org/0000-0002-8583-5677

1. Introduction

Radioactive wastes are produced from a variety of industries, research institutes, and even academies. Effective isolation of the radioactive waste at safe and secure disposal sites has been a key issue to ensure there is no harmful effect on public health and natural environment. Disposal methods for waste isolation are determined according to the class of radioactive wastes on the basis of the level of radioactivity, the extent of decay heat, and physicochemical characteristic features [1]. Each class of radioactive waste has its own maximum permissible activity concentration for the repository, and the individual radionuclides of interest present in each class of waste also have their own maximum permissible activity concentration [2-4]. The number and type of radionuclides, as well as the classification of the waste, vary from country to country [5]. For example, in Korea, the radioactive wastes are classified into high-level radioactive waste (HLW), intermediate-level radioactive waste (ILW), low-level radioactive waste (LLW), and very low-level radioactive waste (VLLW) [6] based the recommendation by International Atomic Energy Agency (IAEA) [7,8] and the permissible radioactivity concentrations set by the Nuclear Safety and Security Commission (NSSC) of Korea [9]. The radioactivity concentrations of the 14 specified radionuclides should be identified for the disposal of radioactivates [10]. Accordingly, in practice, up to 31 radionuclides need to be identified for the final disposal [11].

For the determination of radioactivity, various methodologies can be used depending on the type of decay mode. Gamma-ray-emitting radionuclides can be measured easily by direct non-destructive methods and are thus are called easy-to-measured (ETM) nuclides, whereas complex destructive radiochemical assays are usually used for alpha- and beta-ray-emitting radionuclides, including low-energy gamma-ray-emitting radionuclides, which are difficult to measure directly from the outside of the waste packages by non-destructive methods and thus are called difficult-to-

measure (DTM) or hard-to-measure (HTM) nuclides. Because destructive radiochemical assays are time-consuming and labor-intensive because of the long and complicated process of chemical sample pretreatment, radiochemical separation, and radiation detection, it is not practical for a large volume of waste in terms of cost-effectiveness, although the radiochemical method provides the most accurate results. Thus, indirect methods such as the scaling factor (SF) method, mean radioactivity concentration method, dose-to-curie conversion method, representative spectrum, and theoretical calculation methods (activation or burn-up) have been developed and applied around the world in industries, institutes, and academies. Among these techniques, the SF method has been used as the principal method because it provides the most reliable estimation of the radioactivity of DTM nuclides. The SF method predicts the radioactivity of DTM nuclides from the radioactivity of ETM nuclides measured by indirect methods through the correlation between the radioactivity of DTM nuclides and that of ETM nuclides (called key nuclides in the SF method) [12-21]. SF methods rely on statistical evaluation because SF is a mathematical parameter derived from the correlation. Each country applies different statistical methods and guidelines to judge the applicability of the SF method and to determine the optimum SF values. Currently, there is only one international standard of the SF method set by the International Organization for Standardization (ISO) [21], but it is not sufficient for field practice owing to its lack of details. Because the statistical methods applied in the SF method are mostly limited to simple conventional parametric statistical methods [20, 22, 23], there is ample opportunity to apply various statistical methods, such as nonparametric statistics, Bayesian statistics [24], and artificial intelligence [25], for the development of more advanced and more flexible implementation of SF methods. One good example of potential applications is the disposal of waste during decommissioning of nuclear power plants, which are expected to generate approximately 6,200 tons of potential ILLWs per 900-1,300 MWe pressurized water

reactor (PWR) [26]. A total of 6,200 tons of ILLWs corresponds to around 14,000 drums of waste. To dispose of such huge amounts of radioactive wastes, indirect methods such as the SF method must be introduced for the evaluation of radionuclide inventories [27]. Because the SF method has mostly been applied to nuclear power plant operational waste [20, 28-36], it is necessary to review the SF method to be applied for decommissioning wastes [37, 38]

In this review, the statistical methods and criteria applied in the SF method are described throughout the entire process of SF development from sampling to SF implementation. An overview of international experiences with SF development and the usage of several major countries are also presented. Subsequently, some potential issues are derived from the perspective of international guidelines and statistical criteria. In the next issue, more in-depth investigation on potential issues mentioned here, including suggestions and directions as well as solutions.

2. Current status of SF methodologies

The most basic prerequisite assumption of the SF method is the existence of a correlation between DTM nuclides and key nuclides. SF is simply a factor or parameter derived from the mathematical relationship between them. A variety of mathematical models can be proposed from simple linear equations to complicated regression models. Although SF determination is a purely mathematical process, SF methodologies include many technical components, such as planning, analytical procedures, data management, and interpretation of the results. A representative flow diagram proposed by ISO [21] is a good guide for the understanding of SF implementation, as shown in Fig. 1. The specific implementation details for SF methodologies differ from country depending on the policy of the national radioactive waste management, but the procedures can be categorized into four steps: (1) design of experiment, (2) sampling and radiochemical analysis, (3) evaluation of radiochemical data
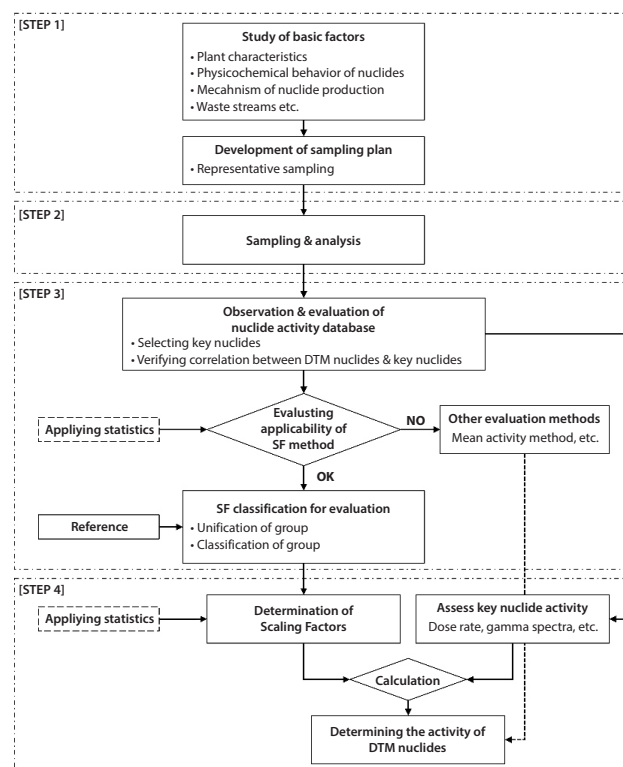

Fig. 1. Flow chart for the presentation of the systematic
SF implementation recommended by ISO [21].

and SF applicability, and (4) determination of SF and the radioactivity of DTM nuclides. In this section, we review the current methodologies of each step in detail.

2.1 Design of experiment

In the design of the experiment, the first task is to identify the factors involved in the SF determination, as depicted in Fig. 1. Correlations between the radioactivity of DTM and key nuclides depend on various factors such as reactor type, reactor component materials, fuel history, production mechanism by which nuclides are generated, variations in reactor coolant chemistry, waste treatment, etc. [20]. These factors should be considered carefully when categorizing wastes into a group representing the average characteristics of the whole waste and cover the categorization of groups

Table 1. Required number of data corresponding to the specific correlation coefficient values proposed by Kashiwagi et al. [39, 40]

Correlation coefficient	Required number of samples
0.60	40
0.70	35
0.80	30
0.90	25
0.95	20

of nuclear power plants, waste streams, and range of radioactivity concentrations of wastes [20, 21]. SF is determined from the radiochemical data of representative samples, which are considered to possess the average characteristics of the waste packages.

The second most important topic in the experimental design is to determine the minimum required number (or size) of samples. Optimization of the entire SF determination process begins with the optimization of the number of samples subjected to radiochemical analysis to determine the radioactivity of DTM nuclides. Because the SF method relies on statistical evaluation, the accuracy of SF values depends on the number of radiochemical data, and therefore, a sufficient number of samples are required to ensure the reliability of the SF method, although the large quantities of samples result in a significant increase in the total cost. However, determination of the optimum sample size is not simple before sampling and radiochemical analysis. To the best of our knowledge, only one approach reported previously by Kashiwagi et al. considered the statistical decision criteria for the required number of samples based on the use of a lower confidence limit for the correlation coefficient [39, 40]. They proposed signs of leveling off the confidence limit values as a suitable decision criterion to determine the number of data instead of using only the correlation coefficient. For example, if the increase rate of the 95% confidence limit is less than 0.005, it is considered as a sign of leveling off. Based on this levelling-off criterion, the required number of data corresponding to the

correlation coefficients is shown in Table 1. As seen in the table, higher correlation coefficients require a smaller amount of data. Nevertheless, decision making in relation to the required sample size requires extra care because the increase rate of the 95% confidence limit also strongly depends on the variance of the data and other factors.

2.2 Sampling and radiochemical analysis

Appropriate sampling is essential to ensure accurate analysis of the samples. Two common practices for representative sampling are homogenized sampling and accumulated sampling. In homogenized sampling, wastes are sufficiently mixed before or during sampling to ensure that the radioactivity of the sample is uniformly distributed. Homogenized sampling endows sufficiently reliable SFs, even with a smaller number of samples. However, it may not be possible for large inhomogeneous samples to be homogenized. In such a case, a sufficient number of sub-samples should be collected to ensure the representativeness of an entire waste package through an accumulated sampling method. In the case of accumulated sampling, it is important to obtain samples with a wide range of radioactivity concentrations to obtain effective correlations between the radioactivity of DTM nuclides and key nuclides.

After sampling, samples were transferred to radiochemical laboratories to perform complex destructive radiochemical analysis for direct measurement. A tedious, time-consuming radiochemical analysis is needed to avoid interference from other nuclides and poor energy resolution due to high self-absorption [41]. It is necessary to ensure that radiochemical analysis is carried out in an appropriate manner in accordance with the characteristics of wastes and radionuclides to be analyzed for accurate correlations.

2.3 Evaluation of radiochemical data and SF applicability

Radiochemical data should be evaluated carefully

before the SF applicability is determined by applying statistics. From a statistical perspective, only radiochemical data above the limit of detection (LOD) should be considered to determine the SF applicability and to determine SF. However, in some cases, lack of sufficient radiochemical data has no choice but to use the LOD value itself as the true radioactivity concentration [20, 42]. More specific case studies are described in Section 3. Decision making for the use of radiochemical data below LOD is required regarding the resampling and radiochemical reanalysis.

Outlier detection is also important in the evaluation of radiochemical data. Statistical methods can be applied to identify outliers in radiochemical data. One applied in the SF method is the ISO-approved Grubbs test [43, 44]. The Grubbs test is a test used to detect a single outlier in the data that follows a normal distribution. The hypothesis (H_0, H_a) and test statistic (G) for the Grubbs test are defined as follows:

$$H_0 : There\ are\ no\ outliers\ in\ the\ data\ set$$
$$H_a : There\ is\ exactly\ one\ outlier\ in\ the\ data\ set$$

$$G = \frac{\max\limits_{i=1,\cdots,n} |y_i - \overline{y}|}{s} \qquad (1)$$

where $\overline{y}$ is the sample mean and s is the sample standard deviation. Another outlier verification for either one or more outliers was proposed by using the normalized fourth central moment [20], kurtosis, which is a measure of the "tailedness" or "peakedness" of the distribution. An outlier test based on kurtosis can be more powerful than the Grubbs test if the number of outliers is unknown [45]. If the cause of the outlier can be identified, the outlier should be corrected or removed. Otherwise, it should not be corrected or removed without careful consideration.

After evaluation of radiochemical data, the applicability of the SF method is determined by the correlation between DTM nuclides and key nuclides. The correlation is observed in the scatter diagram using the radiochemical data. Previous studies on radiochemical data from nuclear power
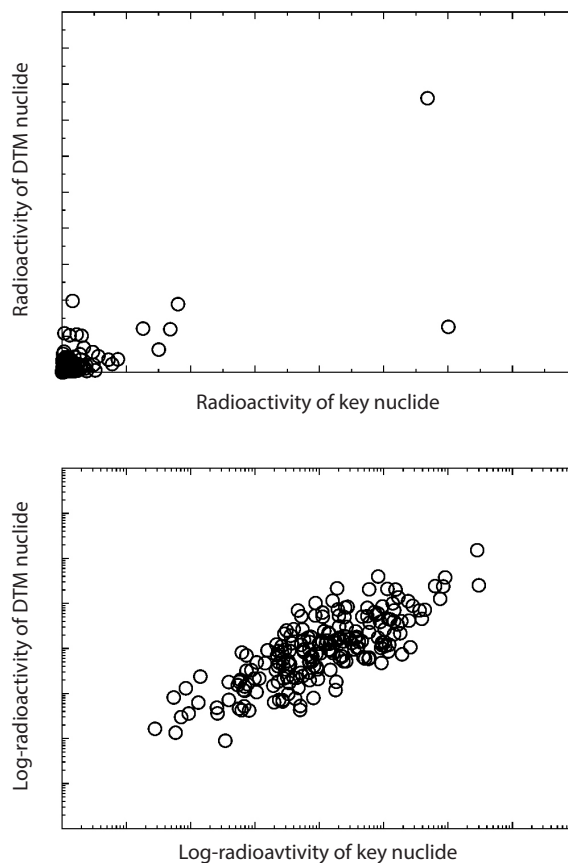

Fig. 2. Scatter plots of 200 random data points following a log-normal distribution as an example. All data generated using the random number generator function and the inverse of lognormal distribution function in Microsoft Excel are presented in linear scale (top) and logarithmic scale (bottom).

plants have shown that the radioactivity concentrations of both DTM nuclides and key nuclides follow a log-normal distribution with a wide range of radioactivity concentrations over several orders of magnitude [20]. In Fig. 2, the characteristic scatter diagram of the simulated radiochemical data (200 points), which follow a log-normal distribution, are depicted on a linear scale and logarithmic scale. As seen in the figure, the correlation is clearer on a logarithmic scale.

The correlation is evaluated by the Pearson product-moment correlation coefficient, which is a measure of linear correlation between two variables [46, 47]. Pearson's correlation coefficient is the covariance of the two

variables divided by the product of their standard deviations. The population Pearson correlation coefficient (ρ_{xy}) and the sample Pearson correlation coefficient (r_{xy}) are given by the following formulas:

$$\rho_{xy} = \frac{\text{cov}(x, y)}{\sigma_x \sigma_y} \quad (2)$$

$$r_{xy} = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \overline{x})^2 \sum_{i=1}^{n}(y_i - \overline{y})^2}} \quad (3)$$

where $\text{cov}(x, y)$ is the covariance, σ_x and σ_y are the population standard deviations, n is the sample size, x_i and y_i are the individual sample points indexed with i, and $\overline{x}$ and $\overline{y}$ are the sample means. The Pearson correlation coefficient has a value between +1 and −1 by the Cauchy-Schwarz inequality, and it reflects the strength of a linear relationship. The specific value of the Pearson correlation coefficient was used as a statistical criterion to determine the applicability of the SF method.

A statistical hypothesis test of the significance of the correlation coefficient was also used to decide whether the population correlation coefficient is significantly different from zero to determine SF applicability. For uncorrelated bivariate normal disturbed data, the hypotheses and test statistic (t), which follows Student's t-distribution with degrees of freedom $n−2$ (t_{n-2}), based on the sample correlation coefficient (r) and sample size (n) are defined as follows:

$$H_0 : \rho = 0$$
$$H_a : \rho \neq 0$$
$$t = r\sqrt{\frac{n-2}{1-r^2}} \sim t_{n-2} \quad (4)$$

The coefficient of determination besides the Pearson correlation coefficient was used to test the applicability of SF method in the case of using linear regression model. The coefficient of determination is a goodness-of-fit measure in a regression model that determines the proportion of variance in the dependent variables that can be explained by the independent variables, defined as follows:

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}} \quad (5)$$

where $SS_{tot} = \sum_{i=1}^{n}(y_i - \overline{y})^2$ is the total sum of squares, $SS_{res} = \sum_{i=1}^{n}(y_i - f_i)^2 = \sum_{i=1}^{n}r_i^2$ is the residual sum of squares, and f_i are the individual sample points indexed with i. The coefficient of determination has a value between 0 and 1. Although the coefficient of determination itself does not indicate correlation, the value of the coefficient of determination was used as a statistical criterion for applicability of the SF method because it is equal to the square of the correlation coefficient in the case of a simple linear regression.

If the correlation between DTM nuclides and key nuclides does not exist, the SF method cannot be applied. Then, the radiochemical data should be reviewed again with consideration of various factors that affect the correlation. Nevertheless, if the correlation cannot be confirmed, other alternative approaches, such as the mean radioactivity concentration method, can be adopted instead of the SF method. Methodologies aside from the SF method are beyond the scope of this review and have been described in detail elsewhere.

2.4 Determination of SF and the radioactivity of DTM nuclides

If the correlations are confirmed using an appropriate statistic, then the SF values are determined using mathematical relationships. The linear relationship between the radioactivity concentration of the DTM nuclide ($a_{D,i}$) and that of the key nuclide ($a_{K,i}$) can be expressed by a simple linear equation as follows:

$$a_{D,i} = SF_i \times a_{K,i} \quad (6)$$

SF_i is a simple proportionality constant of the simple linear equation that passes through the origin. The representative SF ($\overline{SF}$) is calculated using the arithmetic mean (AM) or geometric mean (GM) as follows:
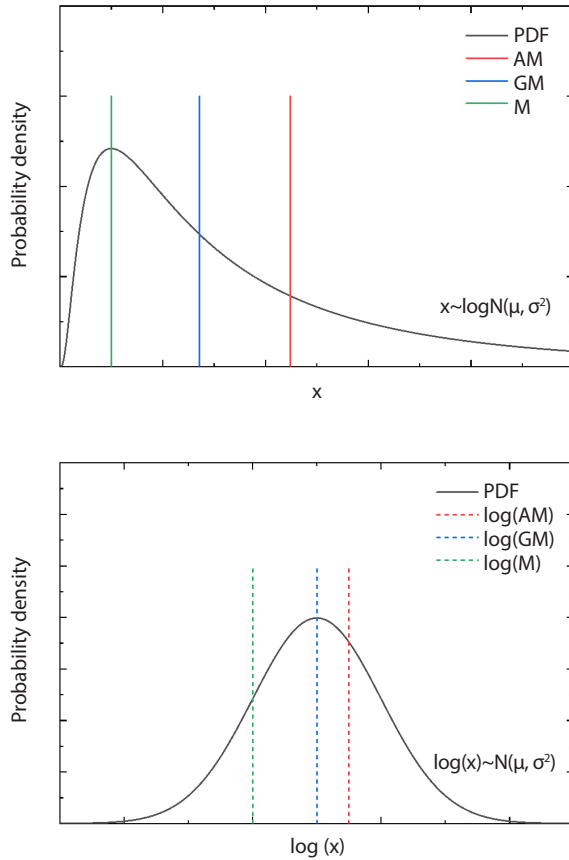
Fig. 3. Probability density of log-normal distribution presented in linear scale (top) and logarithmic scale (bottom), together with arithmetic mean (AM), geometric mean (GM), and mode (M). The position of AM, GM, and M is represented by red, blue, and green solid lines, respectively, and the log-transformed of them are represented by dotted lines.

$$\overline{SF}_{AM} = \frac{1}{n}\sum_{i=1}^{n} SF_i = \frac{1}{n}\sum_{i=1}^{n} \frac{a_{D,i}}{a_{K,i}} \qquad (7)$$

$$\overline{SF}_{GM} = \left(\prod_{i=1}^{n} SF_i\right)^{\frac{1}{n}} = \log^{-1}\left(\frac{1}{n}\sum_{i=1}^{n}\log SF_i\right)$$

$$= \log^{-1}\left(\frac{1}{n}\sum_{i=1}^{n}\log\frac{a_{D,i}}{a_{K,i}}\right) \qquad (8)$$

The geometric mean is the nth root of the product of n numbers or anti-log of the arithmetic mean of log-transformed values. As mentioned before, the radioactivity concentration of DTM nuclides and key nuclides follows a log-normal distribution, the SF represented by their ratio is also known to follow the log-normal distribution [20, 28, 37, 39, 48].

The log-normal distribution is the continuous probability distribution of a random variable whose logarithm is normally distributed, as depicted in Fig. 3, which shows the probability density of log-normal distribution in linear and logarithmic scales. The arithmetical mean is not the mode of the distribution because it is skewed. The geometric mean is the median of the log-normal distribution, and its logarithm is the mean of the normal distribution on a logarithmic scale [49].

Fig. 4 shows three characteristic features of the arithmetic mean and geometric mean for the evaluation of SFs [42]: (1) The arithmetic mean is always greater than or equal to the geometric mean by the inequality of arithmetic and geometric means. (2) The predicted DTM concentration range of both arithmetic and geometric mean is more extended than the actual concentration range of DTM nuclides. (3) The predicted DTM concentration range of the arithmetic mean is shifted toward the higher concentration while that of geometric mean remains almost the same. These three features provide important underlying implications of the respective arithmetic and geometric means for the evaluation of SFs: (1) The radioactivity concentration of DTM nuclides predicted by the geometric mean is underestimated in the low-concentration range, whereas that by the arithmetical mean is overestimated in the high-concentration range, depending on the correlation coefficient. (2) Underestimation in the lower-concentration range has little impact on the estimated inventory of the disposal repository, whereas the overestimation in the higher-concentration range has a much greater impact. (3) Finally, the radioactivity concentration SF calculated by arithmetic mean always yields more conservative values, and the predicted concentration given by arithmetic mean is much more severely overestimated in the higher-concentration ranges.

The relationship between the radioactivity concentration of DTM nuclides and key nuclides can be more generalized based on the nonlinear relationship as follows:
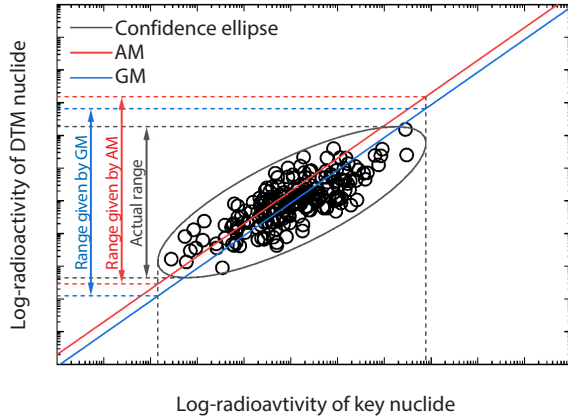
Fig. 4. Evaluation of SF overestimation using the arithmetic mean in comparison with the geometric mean in a log-log plot.

Fig. 5. Evaluation of SF determined using the arithmetic and geometric mean in comparison to linear regression in a log-log plot.

$$a_{D,i} = \alpha (a_{K,i})^\beta \qquad (9)$$

where α is the proportionality constant and β is the regression coefficient. In the special case where β equals 0, it becomes a simple linear equation, as mentioned above. If β is not equal to 0, this simple nonlinear model is a simple linear equation on a logarithmic scale.

$$y = \beta_0 + \beta_1 x \qquad (10)$$

where y is $\log a_{D,i}$ and x is $\log a_{K,i}$. Two parameters, the intercept (β_0) and slope (β_1), 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 ($\hat{\beta}_1$) are given as follows:

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sum_{i=1}^{n}(x_i - \overline{x})^2} \qquad (11)$$

$$\hat{\beta}_0 = \overline{y} - \hat{\beta}_1 \overline{x} \qquad (12)$$

A hypothesis test in simple linear regression can be performed to decide whether the parameter (β_i) is significantly different from the constant (β_i^0). The hypotheses and test statistic (t), which follows Student's t-distribution with
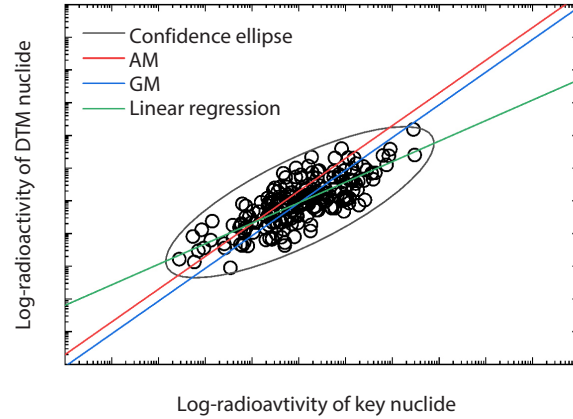
degrees of freedom $n-2$ (t_{n-2}), are defined as follows:

$$H_0 : \beta_i = \beta_i^0$$
$$H_a : \beta_i \neq \beta_i^0$$

$$t = \frac{\hat{\beta}_i - \beta_i^0}{s(\hat{\beta}_i)} \sim t_{n-2} \qquad (13)$$

Simple linear regression analysis on a logarithmic scale can minimize the under- and overestimation of radioactivity concentration, as shown in Fig. 5, compared to the arithmetic and geometric mean techniques, but it is highly affected by outliers and does not match well outside of the range [20].

Two types of post-evaluation of SF have been reported in previous studies for comparison of SFs from different waste streams and periodic updates of SF. Various SFs obtained from different waste streams can be compared to integrate or classify the SFs. A test to compare SFs is a popular two-sample t-test, which is performed to decide whether two SFs are significantly different. A statistical hypothesis test within the acceptable level of difference (D) can be performed based on the pooled variance (s_b^2) [20, 23]. The hypotheses and test statistic (t) under the null hypothesis that follows Student's t-distribution on a logarithmic scale with degrees of freedom $n_1 + n_2 - 2$ ($t_{n_1+n_2-2}$) are defined as follows:

$$H_0 : |SF_1 - SF_2| = \log D$$
$$H_a : |SF_1 - SF_2| \neq \log D$$

$$t = \frac{|SF_1 - SF_2| - \log D}{s_b \sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}} \sim t_{(n_1 + n_2 - 2)} \qquad (14)$$

$$s_b^2 = \frac{(n_1 - 1)\, s_1^2 + (n_2 - 1)\, s_2^2}{n_1 + n_2 - 2} \qquad (15)$$

where SF_1 and SF_2 are geometric means, n_1 and n_2 are the sample sizes, 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. Second, periodic updates of SF are to be considered. Periodic updating has been a critical issue to ensure the long-term stability of the SF values over time. The same hypothetical test as that used in the simple linear regression analysis was used to decide whether periodic updating is required by plotting SFs over time [20, 23, 40]. If the null hypothesis ($H_0 : \beta_1 = 0$, where β_1 is the slope of the SF over time) is true, it is not necessary to update the SF because it cannot be said that the SF has changed over time, even though the slope of SF over time is not exactly zero. If not, the SF should be classified or updated, but the details are not well known.

3. Global experiences of SF implementation

3.1 United States of America

Because the criterion for the radioactivity level of transuranic (TRU) nuclides in low-level radioactive wastes was first offered to the land disposal facility [20, 23], the Electric Power Research Institute (EPRI) initiated an evaluation for radioactivity of TRU and other nuclides in 1976 [50]. They found a correlation between the radioactivity of ^{144}Ce and that of ^{239}Pu in the development of an indirect method to determine TRU nuclide radioactivity. This was the first attempt ever reported regarding the SF method. The concept of SF

was extended to other nuclides, such as beta- and low-energy gamma-emitting nuclides, as published in Title 10 of the US Code of Federal Regulation, Part 61 (10CFR61) by the United States Nuclear Regulatory Commission (US NRC) in 1982 [51]. According to the 10CFR61, the nuclear power plant operator should determine the activity concentration of ^{14}C, ^{60}Co, ^{59}Ni, ^{63}Ni, ^{90}Sr, ^{94}Nb, ^{99}Tc, ^{129}I, and ^{137}Cs, the half-life of which is longer than five years. 10 CFR61 permits indirect approaches, such as the SF method, if the radioactivity is difficult to measure directly. In 1985, after reviewing the radiochemical analysis methods for DTM nuclides, EPRI performed radiochemical analysis on 680 samples of operational radioactive wastes from nuclear power plants to derive the correlation of radioactivity concentration between the nuclides [28]. In 1987, the number of samples was increased to approximately 1,300 to update the US SF [29]. Subsequently, the SF calculation software RADSOURCE was developed with more than 3,000 samples by 1991 [23].

The US NRC and EPRI have led the SF guidance and implementation in the judgment of linear correlation, accuracy of activity concentration, evaluation of SF, etc. The most important fundamental assumption of SF is the linear relationship between the activity concentration of DTM and key nuclides. Nevertheless, to the best of our knowledge, the US has no critical Pearson correlation coefficient values that can be used as a decision criterion, so as to judge whether the geometric mean and SF method is applicable, whereas France and Japan have their own criterion for the correlation coefficient, as described in the following sections. On the other hand, the accuracy of SF and the activity concentrations of DTM nuclides are the next critical concerns. In 1983, the US NRC's branch technical position (BTP) paper on radioactive waste classification recommended the target accuracy of radioactivity concentrations to be within a factor of 10, although the specific details regarding this "factor of 10" were not sufficient to implement it in practice [52]. In 1992, the EPRI compensated for it with log-mean dispersion (LMD) based on the 2σ assumption (i.e., 95% confidence level) defined as follows

[53, 54]:

$$LMD(2\sigma) = \log^{-1}(2\log s_g) \tag{16}$$

$$s_g = \log^{-1}\left(\sqrt{\frac{\sum_{i=1}^{n}(\log SF_i - \log \overline{SF}_{GM})^2}{n-1}}\right) \tag{17}$$

where s_g is the sample geometric standard deviation. If $LMD(2\sigma)$ is less than 10, then at least 95% of the total SF_i data is expected to satisfy the following inequality, which is equivalent to the guideline regarding the accuracy tolerance of a factor of 10:

$$\frac{1}{10}\overline{SF}_{GM} \leq SF_i \leq 10\,\overline{SF}_{GM} \text{ or } \frac{1}{10}a_{D,i} \leq \hat{a}_{D,i} \leq 10a_{D,i} \tag{18}$$

where $\hat{a}_{D,i} = \overline{SF}_{GM} \times a_{K,i}$ is the inferred (i.e., calculated) radioactivity concentration and $a_{D,i} = SF_i \times a_{K,i}$ is the measured radioactivity concentration.

Although various SFs obtained over various timespans from an individual waste stream or mixtures of different waste streams have been evaluated for comparison of such SFs, they continuously assume that the SF data, that is, the radioactivity concentration ratio of DTM to key nuclides, follow a log-normal distribution based on their previous studies. On the basis of log-normality assumption, geometric means and log-mean dispersion have been used to make inferences about the average SF and the variance of SF, respectively, whereas the two-sample Student's t-test with equal variance at the 95% confidence level was performed to identify the difference between the two arithmetic mean values from the log-transformed SF data, as described in detail in the previous section. The SF data are combined to create a representative common SF if the two log-transformed geometric mean values from the two different groups of waste streams are not significantly different at the 95% confidence level. Unlike the accuracy criterion for the activity concentrations, there is no such "factor of 10" criterion in the evaluation of SFs obtained from different waste streams. In the case of temporal trend analysis, SFs were evaluated using a linear regression analysis. In the
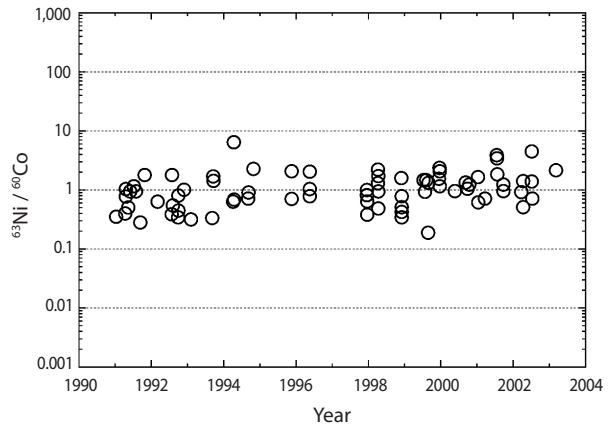

Fig. 6. US experience with various SFs obtained over various timespans for the use of constant ^{63}Ni/^{60}Co SF [23]. Copyright © by WM Symposia. All Rights Reserved. Reprinted.

trend test, if the slope of the log-transformed SF vs. time in days curve is not significantly different from zero at the 95% confidence level, as shown in Fig. 6 [23], the SF is considered to be constant, and the same SF is used continuously. Only a few samples for radiochemical analysis, that is, 6-7 samples per year, are used to evaluate the trend analysis [20].

Despite the danger of severe overestimation, it is notable that LODs as "true activity data" were used between 2002 and 2012 for LILW based on US NRC guidance.

3.2 France

Since 1989, the French Electric Power Corporation (EDF) has carried out two measurement campaigns. In the first campaign, sampling and radiochemical analysis were performed with 10 different nuclear power plants to develop their own SF with the French Alternative Energies and Atomic Energy Commission (CEA). Before sufficient radiochemical data was available, from 1992 to 1995, EDF introduced the international SFs with the agreement of the French Nuclear Waste Agency (ANDRA) for evaluating six DTM nuclides (^{14}C, ^{63}Ni, ^{90}Sr, ^{94}Nb, ^{99}Tc, and ^{129}I) in the operational wastes, regardless of the type of waste. In 1995, the first French SFs were determined from approximately
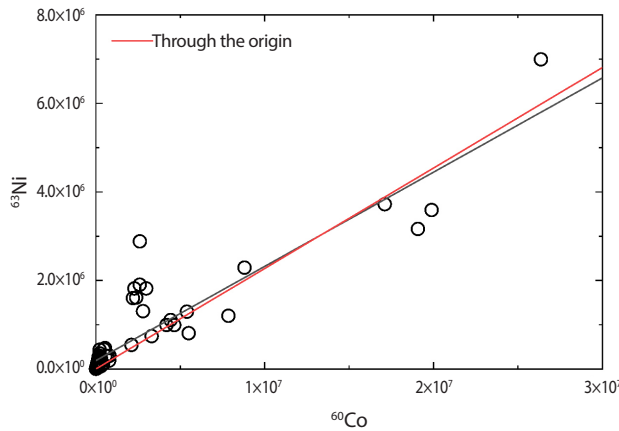
Fig. 7. A representative example of linear regression analysis through origin for the determination of ^{63}Ni/^{60}Co SF adopted in France [20]. The ordinary linear regression and the linear regression through the origin are represented by black and red lines, respectively.

Fig. 8. Decision criterion adopted in Germany for the applicability of log-linear regression analysis for the determination of ^{14}C/^{60}Co SF. Red line represents the tolerance limit of $\alpha_{max}/\alpha = 100$ [20].

500 different results. Interestingly, the French concentrations of SFs were not very different from those of the international SFs implemented from 1992 to 1995, except for ^{129}I, whose values dropped dramatically. The second campaign was carried out over the period 1995 to 1999 to increase the number of samples and determine the SF for ^{99}Tc, which was not determined in the first campaign. Overall, the French SFs were determined from over 1,000 different radiochemical analysis results.

France uses a least-squares linear regression analysis. In the regression analysis fitted through the origin, as shown in Fig. 7, the "true activity" data above the LOD, from which outliers were removed using the Grubbs test, were adopted to determine the SF. Linear regression is applicable where the number of data points is equal to or greater than 5, and the coefficient of determination (R^2) is equal to or greater than 0.7. In case of $0.5 \leq R^2 < 0.7$, the geometric mean is used regardless of the amount of data, whereas an arithmetic mean is used when there are fewer than 5 data points or R^2 is less than 0.5. It is an important criterion that the minimum required number of data is 5 in France, but it is unclear how this criterion is set. They used the same SFs for all the PWRs.
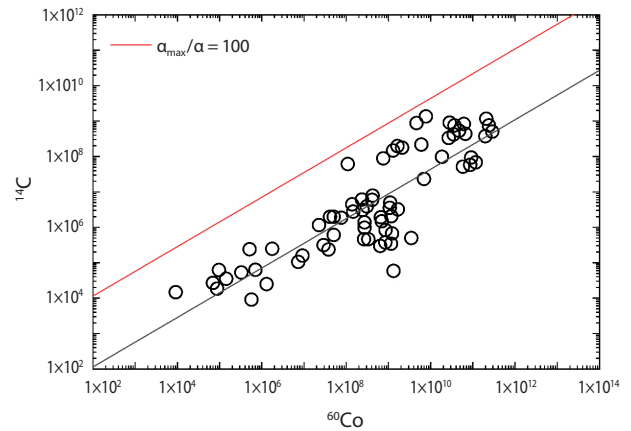
3.3 Germany

As in the case of France, only the data above the LOD are used to determine the SF because the activity concentration measured below the LOD is regarded as meaningless for statistical evaluation [31]. SF is calculated based on the following nonlinear equations between DTM nuclides and key nuclides with two parameters, the proportionality coefficient (α) and the regression coefficient (β).

$$a_{D,i} = \alpha(a_{K,i})^\beta \qquad (19)$$
$$a_{D,i,max} = \alpha_{max}(a_{K,i})^\beta \qquad (20)$$

where α_{max} is defined as the maximum activity concentration ($a_{D,i,max}$) of a DTM nuclide. Two decision criteria for the applicability of the SF method exist: the correlation coefficient and the α ratio. If the correlation coefficient (r) is equal to or greater than 0.7, then the SF method is applicable. It is said that the SF method can be adopted even when $0.5 < r < 0.7$ under special conditions. When the ratio of the maximum α value to the regression-derived α value is greater than 100, as shown in Fig. 8, the correlation is not sufficient due to the excessive variation of data; therefore,

the SF method is not applicable.

3.4 Japan

Since 1992, homogeneous and solidified low-level radioactive wastes from nuclear power plants have been disposed of in Japan's Rokkasho Low-Level Radioactive Waste Disposal Center, where 400,000 disposal drums are permitted. As of 2017, 300,000 drums were disposed of in this center.

In 1992, the SF method was approved by the Nuclear Safety Commission of Japan and has been proposed as a major radioactivity concentration-determination method. The radioactivity concentration of the waste was determined by four approved method, i.e., nondestructive assay, theoretical calculation, SF method, and mean radioactivity concentation method [55]. Interestingly, the radioactivity of DTM nuclides is determined by the SF method, except for ^{3}H and ^{59}Ni, which are determined by the theoretical calculation method and mean radioactivity concentration method, respectively. The mean radioactivity concentration method is used for ^{3}H because it does not have any correlations with key nuclides, whereas the theoretical calculation method is used for ^{59}Ni because the production mechanism and transport behavior are the same as those of its isotope, ^{63}Ni.

It is worth noting that LOD values are regarded as true activity concentrations as of 2001 [20, 42]. For example, the radioactivity concentration of ^{137}Cs, which is a key nuclide and an alpha-emitting fission product in DAW, is too low to detect, and consequently causes an overestimation of the activity concentration of DTM nuclides. However, the impact of overestimation is considered to be negligible compared to the concentration limit of the disposal facility. This is a distinctly different viewpoint from the aforementioned French and German experiences.

The applicability of the SF method is determined by performing a hypothetical test of the correlation coefficient, as described in detail in Section 2. The arithmetic mean is adopted for the SF calculation by the requirement of the regulatory authority, although the nuclear industry sectors in Japan recognize that the geometric mean is more appropriate. They use the generic SF, and the same SF is used if the new SF is not greater than 10 times the existing SF. To this end, annual sampling from every power plant is carried out for radiochemical analysis.

3.5 Korea

In Korea, the radioactivity concentration of 14 nuclides (^{3}H, ^{14}C, ^{55}Fe, ^{58}Co, ^{60}Co, ^{59}Ni, ^{63}Ni, ^{90}Sr, ^{94}Nb, ^{99}Tc, ^{129}I, ^{137}Cs, ^{144}Ce, and total alpha) should be identified to classify the waste for disposal [10]. SF values should be determined conservatively to ensure that the predicted radioactivity is not underestimated. From 2002 to 2005, the first campaign of extensive sampling and radiochemical analysis for 255 samples from 13 different power plants was carried out to develop the first Korean SF [36, 56, 57]. The first Korean SFs were classified by reactor types, plant sites, and six types of waste forms: evaporator bottoms, primary spent resin, secondary spent resin, sludge, spent filter, and DAW. After two years, from 2007 to 2008, the second campaign was conducted to obtain 337 samples from 20 different plants to verify the first Korean SF and to compensate for the insufficient data for some nuclides [57]. To enhance the reliability of SF, the radiochemical data obtained from the first and second campaigns were unified and Korean SF was determined.

Both linear and nonlinear relationships are adopted to determine the SF such that if the correlation coefficient (r) is greater than 0.6, the linear regression on a logarithmic scale is used for calculating SF; otherwise, the geometric mean is adopted [58]. Although the criteria for the "factor of 10" in 10CFR61 are not clear, the concept of "factor of 10" is applied. If the measured radioactivity concentration is more than 10 times greater than the predicted radioactivity concentration, or if the log-mean dispersion is greater than 10, it is considered to be underestimated and the SF is

multiplied by a conservative constant to produce a conservative SF.

4. Notable potential issues for the appropriate decision making and evaluation in relation to SF

Intense public attention and strict regulations focus on the reliability of the SF method since the first operation of the low- and intermediate-level radioactive waste disposal facility in Gyeongju, Korea. The reliability of the methods, procedures, and data analysis is a typical main subject of quality management, such as quality control and quality assurance. Statistical quality control has been proposed and introduced to many field processes, but it is not a strict or rigid system. Instead, it should be regarded as a toolbox to provide flexibility and rationality in any type of decision-making and interpretation. Many statistical tools have been developed and applied to various areas of science and technology. The statistical toolbox can provide many alternatives. Nevertheless, current SF methodologies still rely strongly on old-fashioned statistics, such as Neyman-Pearson's null hypothesis significance testing, which has some obvious vulnerabilities and shortcomings. Therefore, conventional SF methodologies with regard to statistics need to be much improved for advanced implementations in field practices. There have been some cases of misuse and abuse of statistical SF methodologies. In addition, licensees have been struggling with the lack of guidelines for important decisions and rational interpretations of data, including practical implementations. The topics of the present review are limited to the issues in relation to statistical decision-making and data evaluations. This section deals with the potential SF issues as examples regarding the lack of basic and fundamental guidelines briefly in relation to the statistics. Further details for each specific issue will be dealt with later, in separate upcoming review papers.

4.1 Lack of guidelines for the required sample size

Determination of the required sample size should be carried out prior to performing the experiments. The number of samples influences the precision of the estimations and the power of the statistical tests. A smaller sample size may produce inconclusive results, whereas larger sample sizes generally lead to more precise estimation and higher statistical power, but dramatically increase the cost and time of the radiochemical analysis of DTM nuclides in the SF method. Although many statistical formulas are available, there is only one statistical criterion for sample size when it comes to the SF method, as mentioned in Section 2.1, where it is based on the lower confidence limit of the proposed correlation coefficient [39, 40]. However, the required sample size also depends on the sample homogeneity and representative sampling, and therefore the determination of the required sample size has been one of the most difficult tasks in statistics [59, 60]. A large amount of inhomogeneous waste makes it more difficult to suggest specific guidelines for the required sample size. Moreover, if the calculated sample size is too large, the time to acquire a final SF for waste disposal will be too long to meet the licensees' urgent needs. Thus, alternative solutions based on the common sense as reasonable criteria, which are mutually comprehensible, should be given to both licensees and the public. For instance, an interim SF can be used until the finalized SF is obtained when the requirement for the number of data is fulfilled by extensive radiochemical analysis. As mentioned in Section 3, France and Germany have suggested such interim SF values. France had used the US SF values and then replaced them with their own SFs later, after the time-consuming and labor-intensive radiochemical analyses were completed to collect enough data to determine their own SFs. After all, special care must be taken from the beginning to the end of SF implementation to avoid blind reliance on purely mathematical statistics.

4.2 Lack of guidelines for the identification and treatment of outliers

Outliers are data points that differ significantly from others and can cause serious problems (i.e., increasing the variance of the data, reducing the normality of the data and the statistical power of the analysis) during statistical analysis. As mentioned, with the SF method, the Pearson correlation coefficient and the simple linear regression analysis are very sensitive to outliers, while the geometric mean is much less sensitive to outliers. However, it should be very careful to exclude outliers unconditionally from the data because they may not be due to experimental errors but simply appear as large variability of the samples, such as the homogeneity of the samples and the intrinsic large distribution of radioactivity concentrations throughout the samples. Some countries adopted the Grubbs test to reject outliers when developing SF. It should be noted that the Grubbs test is applicable if the population follows a normal distribution and can be used to test the single most extreme value. However, the test was not originally designed to be applied iteratively for the removal of multiple outliers [43]. Thus, alternative statistical methods should be provided, such as the identification of multiple outliers and/or distribution-free techniques. A representative misuse of the Grubbs test is that it is used for the non-normal data. If it follows a non-normal distribution, even the data with extreme values come from the same population. As an educational instance, when the log-mean dispersion based on 2σ assumption equals 10 for the data with a log-normal distribution, at least *ca.* 5% of the total data points are different from the geometric mean more than by a factor of 10, which means that even 10 times higher or lower numbers than the geometric mean value is not extreme.

4.3 Lack of guidelines for the data at concentration below LOD

The guideline for the LOD is a tricky issue. Some DTM nuclides such as ^{99}Tc and ^{129}I in the waste are sometimes undetectable because they are present in very low concentrations in wastes such as decommissioning wastes and/or legacy wastes that have been collected and kept for years [54, 61]. Some countries, such as France and Germany, take into account true radioactivity concentration values only above the LOD, while others like Japan use the LOD values themselves as true radioactivity concentration values when the radioactivity is undetectable [20]. Pure statisticians and ordinary quality managers will not agree with the use of the LOD as true radioactivity concentration values. However, in a manner similar to that proposed in the previous section in this review, the use of the LOD value is an alternative solution based on the conservative common sense. However, the estimation of the radioactivity concentration of DTM nuclides can be too conservative. In the case of decommissioning wastes with extremely low concentrations of DTM nuclides, it may result in a serious problem in terms of the disposal facility's capacity.

4.4 Speculation on type-II errors and power analysis

A statistical hypothesis test as a statistical inference is a method of making a decision that one of two contradictory claims is correct. The two contradictory claims are called the null hypothesis and alternative hypothesis. Two possible decisions are to reject null hypothesis or fail to reject the null hypothesis. The rejection of the null hypothesis means acceptance of the alternative hypothesis. However, failure to reject the null hypothesis does not mean the acceptance of the null hypothesis, which is a representative example of misused hypothesis-testing analysis. Another important factor in making a decision from the hypothesis test is the consideration of type-I and type-II errors. Despite the wide recognition of type-I errors, type-II errors have been ignored in SF implementations. However, besides type-I errors (α), type-II errors (β) are essential for reliable decision making. In the statistical hypothesis test to decide

whether to reject the null hypothesis, the correct decision is not to reject the null hypothesis when it is true and to reject it when it is false (or accept the alternative hypothesis). The rejection of a true null hypothesis and the non-rejection of a false null hypothesis are both incorrect decisions. A suitable range of critical α and β values as the decision criteria can be decided based on the agreement between stakeholders. For example, the LOD increases when α or β decreases. The higher the α value, i.e., the incorrect detection decision, is more publicly acceptable as a conservative decision.

Statistical power (1-β) is an indicator of the capability of a significance test to recognize the difference between means of two data sets. A low power typically means that the sample sizes are insufficient. There are other types of statistical power analysis: criterion, *post hoc*, sensitivity, and *a priori* power analysis. The required α as a decision criterion is derived, provided that power (1-β), effect size, and sample size are given. In *post hoc* analysis, an achieved power is computed provided that α, the effect size, and the sample size are given, whereas the required effect size is computed in the sensitivity analysis provided that α, the power (1-β), and the sample size are given. In *a priori* power analysis, the required sample size before collecting the data can be derived using the input parameters such as the effect size, α, 1-β, and the allocation ratio. In case of a significance test to detect a difference between intercepts or slopes of linear regression, *a priori* power analysis uses the input parameters such as the absolute difference value of intercepts or slopes, α, 1-β, the allocation ratio, the square root of the weighted sum of the residual variances in the two data sets, and the standard deviation of the x-values in each data set so as to calculate the required sample size. In addition to these examples, the power analysis includes many statistical tests such as Fisher's exact test, binomial test, goodness-of-fit test, generic chi-squared test, and logistic regression.

4.5 From conventional parametric statistics into more advanced data science

A hypothesis test involving an estimation with the parameters of the probability distribution from the sample is called the parametric test based on the assumption that the population follows a specific probability distribution, such as a normal distribution. The statistical techniques implemented in the SF method are mostly based on parametric statistics. A common basic assumption for radioactive wastes from nuclear power plants is that the radioactivity concentration population data of wastes follow a normal distribution on a logarithmic scale. Various normality tests, such as Anderson-Darling, Ryan-Joiner, Kologorov-Smirnov, D'Agostino-Pearson, and Shapiro-Wilk, are used to check whether the parametric analyses can be performed. However, each normality test result is sometimes contradictory to each other depending on the test methods. For example, the same data set passes the Anderson-Darling test, but it does not pass the D' Agostino-Pearson test. Unfortunately, there is no objective selection rule that determines the optimal and best normality test for the radioactivity data set in the SF method.

One of the most popular topics in statistical analyses is the *t*-test based on the normality assumption. In the SF method, the parametric *t*-test is applied to determine whether the correlation exists. It is also used to determine whether there is a significant difference between the mean SFs calculated using the data of different waste streams or the data collected at different periods of time. The parametric *t*-test is widely used for small-sample tests that it is quoted in general statistics textbooks, while the *Z*-test is used for the large-sample tests. However, the importance of the fundamental normality assumption is easily overlooked in the analyses using Student's *t*-distribution under the null hypothesis. The Student's *t*-distribution has its own probability density function, $f(t)$, as follows:

$$f(t) = \frac{\Gamma\left(\frac{v+1}{2}\right)}{\sqrt{v\pi}\,\Gamma\left(\frac{v}{2}\right)}\left(1+\frac{t^2}{v}\right)^{-\frac{v+1}{2}} \qquad (21)$$

where v is the degrees of freedom, and Γ is the gamma

function. The probability density function of the Student's *t*-distribution is derived from the definition of the random variable *T* given by

$$T \equiv \frac{Z}{\sqrt{\dfrac{V}{v}}} \tag{22}$$

where Z follows a standard normal distribution with a mean of 0 and a variance of 1, V follows a chi-squared distribution with v degrees of freedom, and Z and V are independent. The probability density function of the Student's *t*-distribution is the joint probability density function of Z and V, which follows the standard normal distribution and the chi-squared distribution, respectively. If the population does not follow the normal distribution, then the random variable T will follow an unknown distribution and will never follow the Student's *t*-distribution. Finally, the normality test should also be performed prior to any *t*-tests. If the data set does not pass the normality tests, alternative nonparametric tests such as Wilcoxon, Mann-Whitney, Kruskal-Wallis, Mood's median, and Friedman tests can be performed instead to test the population location.

Likewise, other statistical analyses such as the Pearson and Spearman correlation test, parametric regression analysis, the Grubbs outlier test, and LOD based on the specified distribution with specified parameters should be used after the normality test. Thus, distribution-free nonparametric statistical techniques are attractive for achieving a robust analysis free from the effect of outliers and for solving problems with small sample sizes [62]. In general, parametric tests are preferred to compare nonparametric tests because of their higher statistical power, which is the probability that the test rejects the null hypothesis when a specific alternative hypothesis is true. However, nonparametric tests would be more appropriate when the sample size is small, when the distribution of population is unknown or cannot be assumed to have an approximately normal distribution. To the best of our knowledge, nonparametric statistical methodologies have never been applied to the SF

method. Big data and data science are now heralding a new era for statistics in radioanalytical chemistry. It is expected that state-of-the-art statistical techniques such as Markov Chain Monte Carlo (MCMC) simulation, Bayesian statistics, artificial intelligence, and chemometric analysis, will play an important role in more advanced and more flexible implementation of SF methods, so as to replace the current SF methodologies depending heavily on the parametric statistics and the correlation analysis of radioactivity concentration between only two DTM and key nuclides.

5. Conclusions

The urgent need for the disposal of large quantities of radioactive wastes and upcoming decommissioning wastes has popularized the SF method as an efficient solution to the problems with the time-consuming and labor-intensive radiochemical analysis. We reviewed four categories of the methodologies in detail and the examples of SF implementation in four major countries: The United States, France, Germany, and Japan. The SF methodologies and implementations differ from country to country. Some of their methodologies are misused infield practice, and they are even contradictory to each other. Although there is an international standard guideline, the level of detail is grossly insufficient to meet licensees' and public needs for rational decision making.

In the long history of SF implementation, statistical thinking has played well throughout all over four categories of SF implementation processes: design of experiment, sampling and radiochemical analysis, evaluation of radiochemical data and SF applicability, and determination of SF and the radioactivity of DTM nuclides. Nevertheless, it still needs to be improved in response to the recent public attention and interest in the reliability of the procedures and data regarding radioactive wastes since the first operation of the low- and intermediate-level radioactive waste disposal facility in Gyeongju, Korea. Now, the nuclear industries are

struggling with the dilemma between cost-effectiveness and public acceptance. Statistical decision-making is a good solution to this dilemma. Flexibility, rationality, and exactness can be given with the help of a new face of statistics and data science, which are already popular and utilized in other areas of science and technology. Bayesian statistics, multivariate analysis, and distribution-free statistical techniques are good examples of statistical decision-making.

As the new era of nuclear decommissioning begins, the development of a advanced concept of SF may be recommended, and some statistical potential issues have been drawn in that context. This critical review is expected to be helpful for the development of advanced SF methodologies that can contribute to the realization of futuristic SFs. We proposed some potential issues to be considered: lack of guidelines for the required sample size, data treatment at concentrations below LOD, identification and treatment of outliers, problematic conventional parametric statistics, speculation on Type-II errors, and power analysis. In our subsequent review, the direction or solution of each specific issue will be discussed in detail based on various statistical approaches, not limiting to the issues mentioned in this review.

Acknowledgements

REFERENCES

[1] International Atomic Energy Agency. Classification of Radioactive Waste, IAEA Safety Standards Series No. GSG-1, International Atomic Energy Agency, Vienna (2006).

[2] International Atomic Energy Agency. Derivation of Activity Limits for the Disposal of Radioactive Waste in Near Surface Disposal Facilities, International Atomic Energy Agency Report, IAEA-TECDOC-1380 (2003).

[3] S.W. Hong, M.S. Kim, K.I. Jung, and J.B. Park, "Determination of Radionuclide Concentration Limit for Low and Intermediate-Level Radioactive Waste Disposal Facility I: Application of IAEA Methodology for Underground Silo Type Disposal Facility", J. Nucl. Fuel Cycle Waste Technol., 15(3), 257-264 (2017).

[4] S.W. Hong, M.S. Kim, K.I. Jung, and J.B. Park, "Determination of Radionuclide Concentration Limit for Low and Intermediate-Level Radioactive Waste Disposal Facility II: Application of Optimization Methodology for Underground Silo Type Disposal Facility", J. Nucl. Fuel Cycle Waste Technol., 15(3), 265-279 (2017).

[5] J.Y. Song, G.H. Song, J.H. Park, and S.W. Ji, "Review Study of Radioactive Waste Classification and Waste Acceptance Criteria for Disposal of VLLW from Decommissioning of Nuclear Facilities", J. Radiat. Ind., 12(3), 239-247 (2018).

[6] Nuclear Safety and Security Commission, Standards for Radiation Protection, etc, Notice No. 2019-10, Nuclear Safety and Security Commission, Seoul (2019).

[7] International Atomic Energy Agency, Classification of Radioactive Waste: A Safety Guide, IAEA Safety Series No. 111-G-1.1, International Atomic Energy Agency, Vienna (1994).

[8] International Atomic Energy Agency, Application of the Concepts of Exclusion, Exemption and Clearance, Safety Guide No. RS-G-1.7, International Atomic Energy Agency, Vienna (2004).

[9] Nuclear Safety and Security Commission, Regulation on the Criteria for the Classification and Clearance of Radioactive Wastes, Notice No. 2017-65, Nuclear Safety and Security Commission, Seoul (2017).

[10] Nuclear Safety and Security Commission, General Acceptance Criteria for Low- and Intermediate-Level Radioactive Waste, Notice No. 2020-11, Nuclear Safety and Security Commission, Seoul (2020).

[11] Korea Hydro & Nuclear Power Co., Standard Operating Procedure: Disposal of Low- and Intermediate-Level Radioactive Waste Packages (in Korean), 3rd ed., No. 8610 (2016).

[12] J.E. Cline, J.R. Noyce, L.J. Coe, and K.W. Wright. Assay of Long-Lived Radionuclides in Low-Level Wastes from Power Reactors, United States Nuclear Regulatory Commission Report, NUREG/CR-4101 (1985).

[13] International Atomic Energy Agency. Characterization of Radioactive Waste Forms and Packages, International Atomic Energy Agency Report, Technical Reports Series No. 383 (1997).

[14] International Atomic Energy Agency. Radiological Characterization of Shut Down Nuclear Reactors for Decommissioning Purposes, International Atomic Energy Agency Report, Technical Reports Series No. 389 (1998).

[15] International Atomic Energy Agency. Review of the Factors Affecting the Selection and Implementation of Waste Management Technologies, International Atomic Energy Agency Report, IAEA-TECDOC-1096 (1999).

[16] International Atomic Energy Agency. Management of Problematic Waste and Material Generated during the Decommissioning of Nuclear Facilities, International Atomic Energy Agency Report, Technical Reports Series No. 441 (2006).

[17] International Atomic Energy Agency. Innovative Waste Treatment and Conditioning Technologies at Nuclear Power Plants, International Atomic Energy Agency Report, IAEA-TECDOC-1504 (2006).

[18] International Atomic Energy Agency. Development of Specifications for Radioactive Waste Packages, International Atomic Energy Agency Report, IAEA-TECDOC-1515 (2006).

[19] International Atomic Energy Agency. Strategy and Methodology for Radioactive Waste Characterization, International Atomic Energy Agency Report, IAEA-TECODC-1537 (2007).

[20] International Atomic Energy Agency. Determination and Use of Scaling Factors for Waste Characterization in Nuclear Power Plants, IAEA Nuclear Energy Series, NW-T-1.18 (2009).

[21] International Organization for Standardization, Nuclear Energy - Nuclear Fuel Technology - Scaling Factor Method to Determine the Radioactivity of Low-and Intermediate-Level Radioactive Waste Packages Generated at Nuclear Power Plants, ISO 21238, International Organization for Standardization, Geneva (2007).

[22] M. Matsui, A. Fujita, and M. Kashiwagi, "Radioactivity Determination and Measurement for Shallow Land Disposal of Low-Level Radioactive Wastes", Proc. of the 12th Sino-Japanese Seminar on Nuclear Safety, December 1-2, Tokyo (1997).

[23] D.W. James, "Impacts of Operational Changes on LLW Scaling Factors", Proc. of Waste Management Symposia (WM2004), February 29-March 4, Tucson (2004).

[24] B. Zaffora, S. Demeyer, M. Magistris, E. Ronchetti, G. Saporta, and C. Theis, "A Bayesian Framework to Update Scaling Factors for Radioactive Waste Characterization", Appl. Radiat. Isot., 159 (2020).

[25] S.C. Lee, K.H. Hwang, S.H. Kang, and K.J. Lee, "A Study on the Improvement of Scaling Factor Determination using Artificial Neural Network", J. Korean Radioact. Waste Soc., 2(1), 35-40 (2004).

[26] International Atomic Energy Agency. Managing Low Radioactivity Material from the Decommissioning of Nuclear Facilities, International Atomic Energy Agency Report, Technical Reports Series No. 462 (2008).

[27] H.S. Jens, "Scaling Factors in Clearance Measurements on Decommissioning Waste at Danish Decommissioning", Proc. of NKS-R Decommissioning Seminar 2013, November 6-7, Halden (2013).

[28] W.T. Best and A.D. Miller. Radionuclide Correlations in Low-Level Radwaste, Electric Power Research Institute Report, EPRI-NP-4037 (1985).

[29] W.T. Best and A.D. Miller. Updated Scaling Factors in Low-Level radwaste, Electric Power Research In-

stitute Report, EPRI-NP-5077 (1987).

[30] W. Wurtinger, "Determination of Activity of NPP wastes on the Basis of Key Nuclides", Proc. of the 2nd International Seminar Radioactive Waste Products, Jülich (1990).

[31] U. Oppermann and W. Müller, "Characterization of Nuclide Inventories in Waste Streams from Nuclear Power Plants", Proc. of the 4th International Conference on Nuclear Waste Management and Environmental Remediation (ICEM1993), September 5-11, Prague (1993).

[32] A. Husain and J.P. Krasznai, "Characterization of Ontario Hydro's Low and Intermediate Level Reactor Waste - Development of Scaling Factors for Difficult-To-Measure Nuclides", Proc. of 5th International Conference on Radioactive Waste Management and Environmental Remediation (ICEM1995), September 3-9, Berlin (1995).

[33] S. Niese and B. Gleisberg, "Determination of Radio-isotopes of Ce, Eu, Pu, Am and Cm in Low-Level-Wastes from Power Reactors", J. Radioanal. Nucl. Chem., 200(1), 31-41 (1995).

[34] S. Niese and B. Gleisberg, "Determination of Radio-isotopes of Ce, Eu, Pu, Am and Cm in Low-Level Wastes from Power Reactors using Low-Level Measuring Techniques", Appl. Radiat. Isot., 47(9/10), 1113-1114 (1996).

[35] T. Kekki and A. Tiitta. Evaluation of the Radioactive Waste Characterisation at the Olkiluoto Nuclear Power Plant, Radiation and Nuclear Safety Authority Report, STUK-YTO-TR162 (2000).

[36] K. Hwang, S. Lee, S. Kang, K.J. Lee, C. Jeong, S. Ahn, T. Kim, K. Kim, Y. Herr, and M. Song, "Development of Radionuclide Inventory Declaration Methods using Scaling Factors for the Korean NPPs - Scope and Activity Determination Method-", J. Korean Radioact. Waste Soc., 2(1), 77-85 (2004).

[37] B. Zaffora, M. Magistris, G. Saporta, and F.P. La Torre, "Statistical Sampling Applied to the Radiological Char-

acterization of Historical Waste", EPJ Nucl. Sci. Technol., 2(34), 1-11 (2016).

[38] B. Zaffora, M. Magistris, J.P. Chevalier, C. Luccioni, G. Saporta, and L. Ulrici, "A New Approach to Characterize Very-Low-Level Radioactive Waste Produced at Hadron Accelerators", Appl. Radiat. Isot., 122, 141-147 (2017).

[39] M. Kashiwagi, W. Müller, and B. Lantès, "Considerations on the Activity Concentration Determination Method for Low-Level Waste Packages and Nuclide Data Comparison between Different Countries", Proc. of International Conference on the Safety of Radioactive Waste Management, March 13-17, Cordoba (2000).

[40] H. Masui, M. Kashiwagi, and H. Ozaki, "Low-Level Radioactive Waste Radioactivity Evaluation Practice in Japan and Suggestion for Rationalization", Proc. of International Workshop on Determination and Declaration of Nuclide Specific Activity Inventories in Radioactive Wastes, September 26-28, Cologne (2001).

[41] IAEA. Radioactive Particles in the Environment: Sources, Particle Characterization and Analytical Techniques, International Atomic Energy Agency Report, IAEA-TECDOC-1663 (2011).

[42] H. Fujihara, H. Masui, M.M. Kashiwagi, M. Motoyama, and H. Ozaki, "Rationalization of Radioactivity Concentration Determination Method for Low-Level Radioactive Waste Generated at Japanese Nuclear Power Plants", Proc. of 8th International Conference on Radioactive Waste Management and Environmental Remediation (ICEM2001), September 30-October 4, Bruges (2001).

[43] F.E. Grubbs, "Procedures for Detecting Outlying Observations in Samples", Technometrics, 11(1), 1-21 (1969).

[44] International Organization for Standardization, Accuracy (trueness and precision) of Measurement Methods and Results. Part 2: Basic Method for the Determination of Repeatability and Reproducibility of a Standard Measurement Method, ISO 5725-2:1994 In-

ternational Organization for Standardization, Geneva (1994).

[45] J.H. Livesey, "Kurtosis Provides a Good Omnibus Test for Outliers in Small Samples", Clin. Biochem., 40(13-14), 1032-1036 (2007).

[46] K. Pearson, "Note on Regression and Inheritance in the Case of Two Parents", Proc. R. Soc. London, 58 (347-352), 240-242 (1895).

[47] A.G. Asuero, A. Sayago, and A.G. González, "The Correlation Coefficient: An Overview", Crit. Rev. Anal. Chem., 36(1), 41-59 (2006).

[48] M. Kashiwagi, H. Masui, Y. Denda, D. James, W. Muller, B. Lantes, M. Garamszeghy, J.L. Leganes, and H. Maxeine, "ISO Standardization of the Scaling Factor Method for Low- and Intermediate-Level Radioactive Wastes Generated at Nuclear Power Plants", Proc. of the 11th International Conference on Environmental Remediation and Radioactive Waste Management (ICEM2007), September 2-6, Bruges (2007).

[49] F.W. Wellmer, The Use of the Lognormal Distribution. In Statistical Evaluations in Exploration for Mineral Deposits, Springer, Berlin (1998).

[50] J.E. Cline, K.L. Wright, and J.W. Hollcroft. Activity Levels of Transuranic Nuclides in Low-Level Solid Waste from US Power Reactors. Final report, Electric Power Research Institute Report, EPRI-NP-1494 (1980).

[51] United States Nuclear Regulatory Commission, Code of Federal Regulation Title 10 Part 61, Low Level Radioactive Waste Scaling Factors, United States Nuclear Regulatory Commission, Washington DC (1986).

[52] United States Nuclear Regulatory Commission, Low-Level Waste Licensing Branch Technical Position on Radioactive Waste Classification, United States Nuclear Regulatory Commission, Washington DC (1983).

[53] J.E. Cline. Utility Use of Constant Scaling Factors, TR-109448 (2006).

[54] B. Cox and P. Saunders. Development of Generic Scaling Factors for Technetium-99 and Iodine 129 in Low and Intermediate Level Waste, Electric Power Research Institute Report, Technical Report 3002005564 (2015).

[55] H. Masui, T. Sakai, T. Funahashi, A. Segi, and M. Kashiwagi, "Radioactive Concentration Determining Method for Dry Active Waste Generated at Japanese Nuclear Power Plants", Proc. of the 7th International Conference on Radioactive Waste Management and Environmental Remediation (ICEM1999), September 26-30, Nagoya (1999).

[56] T. Kim, K. Kang, S.C. Lee, and K.J. Lee, "Development of Operation Program of SF for Intermediate- and Low-Level Radioactive Waste (in Korean)", Proc. of the 2006 Korean Radioactive Waste Society Conference, June 15-16, Seoul (2006).

[57] T. Kim, K. Kang, and J. Ha, "Determination and Verification of the Scaling Factor for the Radionuclide Inventory of the Radioactive Waste from Nuclear Power Plants", J. Nucl. Sci. Technol., 45, 756-757 (2008).

[58] K.Y. Jee, H.J. Ahn, S.C. Sohn, S.H. Han, and K.S. Choi, "Derivation of the Korean Radwaste Scaling Factor", Proc. of the 11th International Conference on Environmental Remediation and Radioactive Waste Management (ICEM2007), September 2-6, Bruges (2007).

[59] L. Festinger, and D. Katz, Research Methods in the Behavioral Sciences, Dryden Press, Orlando (1953).

[60] B. Williams, A Sampler on Sampling, John Wiley & Sons, New York (1978).

[61] T. Kim, "Application of SF Method for Legacy Waste (in Korean)", Proc. of the 12th Nuclear Safety & Security Information Conference, April 5-6, Daejeon (2007).

[62] J.P. Verma and A.S.G. Abdel Salam, Testing Statistical Assumptions in Research, John Wiley & Sons (2019).