Nuclear Engineering and Technology 54 (2022) 3788–3794

Contents lists available at ScienceDirect

Nuclear Engineering and Technology

journal homepage: www.elsevier.com/locate/net

Original Article

Soil sampling plan design of key facilities for denuclearization based on data quality objective process

Yeoryeong Jeon, Yongmin Kim*

Department of Radiological Science, Daegu Catholic University, Gyeongbuk, 38430, Republic of Korea

ARTICLE INFO

Article history:
Received 1 December 2021
Received in revised form
11 April 2022
Accepted 13 May 2022
Available online 17 May 2022

Keywords:
Key facilities for denuclearization
Nuclear activity
Data quality objective (DQO)
Soil sampling
Plan design

ABSTRACT

The possibility of denuclearization of the Korean Peninsula has been continuously debated, and the initiative participation of the Republic of Korea has necessitated preemptive measures against neighboring countries. In this study, we present a proposal for formulating a site survey plan when the amount of site information provided is insufficient and the accuracy of the information is not guaranteed. Considering a case wherein “a soil sample analysis is used to determine the presence or absence of nuclear activity” in a radiochemical laboratory, which is a typical key facility for denuclearization, the optimal soil sample collection plan is designed based on international guidelines and public information. In the event of denuclearization, a scenario that is not based on the expertise of the sample collector is set, and the data quality objective (DQO) process is applied to ensure reality. Consequently, the primary sample collection points can be derived in consideration of accessibility, and the sample collection scale can be adjusted according to the cost. By applying the DQO process to ensure sample representativeness and reality, reliable and resource-efficient soil sample collection can be achieved in radiochemical laboratories and other denuclearization facilities.

1. Introduction

Currently, the denuclearization of the Korean Peninsula is being debated internationally, and the prospect of progressing to the verification stage has necessitated preemptive measures against neighboring countries [1,2]. 38 North, the United States specialized medium for Democratic People's Republic of Korea (DPRK), has indicated that radiochemical laboratories, representing one of the key facilities for denuclearization, have been attracting increasing attention owing to the suspicion of nuclear activity [3]. The radioactive waste in nuclear weapon sites is mainly generated by activities directly linked to nuclear activities such as ore crushing and nuclear fuel production. In general, soil contamination may occur around the nuclear facilities as nuclear activity progresses; therefore, environmental samples can be used to determine the presence and extent of contamination and the radioisotope ratio. Thus, the information obtained through the analysis of environmental samples can be used as a basis for determining the presence or absence of nuclear activity and the exact location of nuclear activity.

Although soil samples can be used, considering that the investigation area is in the target country for denuclearization verification, the integrity of the information provided by the target country cannot be guaranteed, and the amount of information available when designing a local sample collection plan may be insufficient. Therefore, a sample collection plan designed in such a situation may be unrealistic. To compensate for the shortcomings mentioned above, it is necessary to design a plan under a systematic system that enables the required quantity and quality of the collected data to be met. Additionally, the sample collection planning volatility must be minimized and survey efficiency must be improved by clarifying the purpose of sample collection and including the information of the sample collected, such as the location of collection and method of collection, and the samples to be measured. The application of the data quality objective (DQO) process may be considered to satisfy the abovementioned requirements. DQO is used in decommissioning nuclear power plants and various fields such as soil pollution surveys, and it reduces potential mistakes by improving the quality and quantity of information necessary for decision-making [4].

In preparation for the case involving the participation of the Republic of Korea in denuclearization verification in the future, a scenario was established by utilizing the information collected

---

* Corresponding author.
E-mail address: ymkim17@cu.ac.kr (Y. Kim).

https://doi.org/10.1016/j.net.2022.05.012

about a radiochemical laboratory of the Yongbyon nuclear complex in DPRK. The plutonium is reprocessed in a radiochemical laboratory. Besides this, the waste generated by the nuclear activities is stored and treated, and thus, when the samples here are collected and analyzed, the presence or absence of the nuclear activity can be ascertained. In other words, the volume of information that can be utilized and the area where significant results are likely to be obtained after sample analysis are selected as the sample collection area. Based on the established scenario, a countermeasure was proposed by designing and optimizing a soil sample collection plan to determine whether nuclear activity was conducted according to the DQO process to secure the scenario's reality and representativeness of the sample.

2. Design of soil sample collection scenario – DQO process steps 1-5

The DQO process comprises seven steps. If the available information changes or the conditions of the survey site change during the planning stage, an iterative procedure is followed to revise the survey plan to reflect the new information [5]. Steps 1–5 of the DQO process clearly define the problem to be solved by utilizing the collected preliminary information, confirming the information necessary for problem-solving, and defining the spatial and temporal boundaries of the survey plan.

2.1. Prerequisites of the scenario

The DQO process minimizes decision-making errors because it follows systematic procedures during the planning stage of large-scale projects such as US EPA-compliant environmental surveys and decommissioning of nuclear power plants. The established sample collection plans based on statistical methods are likely to vary because of various factors, including decision-makers, experts participating in sample collection, the opinions of regulators, and national conditions. If the volume of information available increases, the plans should be modified, and a repeatedly managing scheme should be established. It is ideal to apply DQO when there the sample's representativeness can be ensured in a constrained way, as in this paper, or when the survey plan fluctuates. Considering that there is no formally known procedure when visiting inspectors for denuclearization verification in cases other than official cases such as verification of unconfirmed nuclear material and notification of International Atomic Energy Agency (IAEA) inspections, prerequisites for setting up an informal scenario are established as below.

- The survey site is an area with potential contamination and no specific information on the contamination distribution.
- There is no specific designation procedure for the sample collection location and quantity (there is a need to determine in advance whether an official procedure is applied).
- A significant amount of uranium or plutonium compounds are present in the collected sample.
- The sample analysis results are the determination of the presence or absence of nuclear activity and its scope, determination of whether there is a change in the nuclear material before and after processing, and determination of whether the ratio of uranium isotopes matches that of the natural state.

The depth of the sample to be collected can be selected in various ways according to the purpose of the analysis. We assumed that surface soil at a sampling depth of 2 mm or less was collected to reflect the prerequisites mentioned above and ensure the ease of sample collection. Only essential tools, such as shovels and scoops, were used for the sample collection. In general, if a sampling site is potentially contaminated, an ideal sampling depth should be collected on the horizon of the expected contaminated soil formation. The samples can be collected based on defined depth intervals, or soil with a certain depth of contamination. The collection of surface soil within 2 mm was empirically used primarily for general environmental analysis in situations wherein the information on the physical properties of the soil in the survey area and the concentration of nuclear material contamination were highly uncertain [6]. Assuming the site is contaminated with uranium and plutonium, the distribution of different pollutants can vary according to the depth. However, because this study targeted the key facility for denuclearization of DPRK, it was considered that the time limit for collecting samples by the method that takes into account various depths of the sample. In addition, because the presence or absence of radionuclide contamination and the contamination level is unclear, a method for rapidly collecting many samples was investigated. The sample collection points were set as adjacent plane and slope near the roads, prioritizing accessibility.

2.2. Key locations of sample collection

First, the environment around the radiochemical laboratory is studied by utilizing the satellite image information provided by 38 North Atlas, and the soil sample collection point is derived [7]. Nuclear fuel burned in a 5 MWe gas-cooled reactor is employed in the radiochemical laboratory. The high-level radioactive waste generated is stored in a facility adjacent to the building, where the primary process proceeds after the reduction [8]. Therefore, as mentioned in the Introduction, direct release of the radioactive waste into the soil around the storage tank can be suspected. In addition, the location of pollution sources such as waste management facilities and waste fuel rod reservoirs can be easily determined. Because there is a high probability of the facility personnel being uncooperative for the sampling, this scenario assumed a case of patrol and visit near the radiochemical laboratory.

Second, the sampling density and pattern must be determined to obtain the sampling point. If one facility is considered as an object of interest, a ‘regional phase’ that collects samples from around 1 to 10 points per square kilometer is considered to be appropriate [6]. In the case of the sampling pattern depicted in Fig. 1, it is not suitable to apply the circular grid pattern because the building arrangement is dense and is located in the middle of a topography consisting of green areas. Therefore, a ‘stratified random sampling pattern’ that randomly collects samples within a divided small grid was applied in the scenario. This pattern was selected based on the assumption that contamination of the surrounding site is almost inevitable. The interval of the sampling pattern was set to 100 m in consideration of the sampling density. The area of interest (indicated by the solid blue line in Fig. 1) included the entire facility and movement path except for the green area surrounding the facility.

Because it is impossible to guarantee the homogeneity of the soil in the survey area, it is reasonable to consider the sampling pattern depicted in Fig. 1, while prioritizing accessibility. This is because even if access to the facility is feasible, the probability of allowing sample collection is extremely low. Four high-priority locations were selected (1–4) as shown in Fig. 1, considering the possible limitations in sample collection.

- 1: A plutonium extraction facility near the vehicle path to the south of the uranium recovery facility
- 2: An area near the radioactive waste treatment facility
- 3: A green space outside the radiochemical laboratory boundary

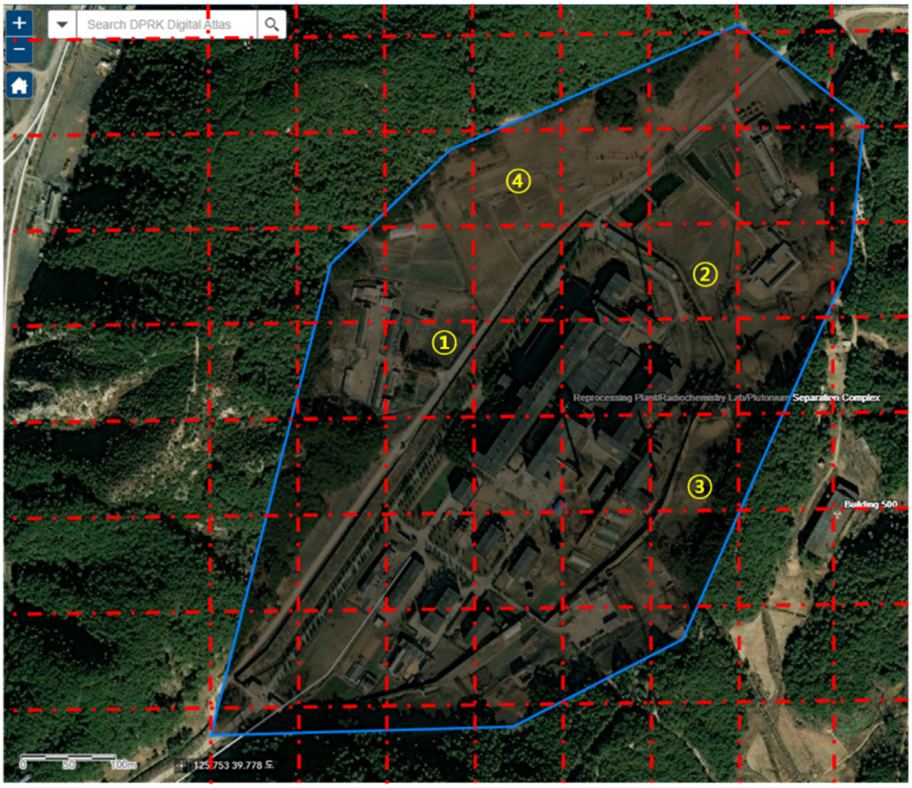

Fig. 1. Survey area of interest and applied sampling pattern in the radiochemical laboratory.

- 4: Flat land outside the radiochemical laboratory boundary

Location 1 is the most accessible location from the facility where plutonium extraction and uranium recovery occur, and is highly likely to be included in the denuclearization verification facility. Location 1 and location 3 exist in green space; therefore, soil samples from these spots can be collected together with moss, which is often used for environmental analysis. Because location 2 is near a radioactive waste treatment facility, it may be contaminated with uranium, plutonium, and other nuclides. If an isotope ratio other than uranium, plutonium, and thorium is designated as a signature, the soil sample collected in location 2 can be used for analysis. The signature referred to here is defined by the IAEA and relates to characteristics that confirm evidence of nuclear activity [9]. Location 3 exists in green space, and although it is slightly far from the facility boundary, it is likely composed of leaf mold, which is the upper layer of the soil composed of the carcasses of animals and plants in the early decomposition stage, including moss. Therefore, the soil sample from location 3 can be used to determine the amount of material artificially added to nature. However, considering the realistic travel route, location 2 and location 3 are somewhat inaccessible. Location 1 is the soil sample collection spot with the highest analysis value and best accessibility. However, if sample collection is not permitted inside and around the facility's boundary, location 4, which is located in the area composed of flat land while being included in the site of the radiochemical laboratory, is preferred.

Table 1 presents the information of the sample collection plan design applied in the scenario establishment for DQO stages 1–5.

3. Optimization of soil sample collection scenario – DQO process steps 6 and 7

Steps 6 and 7 of the DQO process optimize the sample collection plan through probabilistic evaluation. The consequences of making a wrong decision are predicted, acceptable error limits are set, and an approach to collect samples that can meet all the pre-specified requirements is determined. In this study, the scenario was optimized using the Visual Sample Plan (VSP) software to evaluate the sampling design to secure the reality of the established scenario and representativeness of the collected samples [10,11]. Table 2 presents the information of the sample collection plan design applied in the scenario establishment for DQO stages 6 and 7.

3.1. *Specification of acceptable levels of decision error (field variability control)*

Decision-makers of the sampling plan can flexibly manage decision errors by controlling the number of samples to be collected. Fig. 2 depicts the location of the collected samples in the survey area derived using VSP. The points indicated by the red circles in the figure are spots that almost match the main collection spots specified in the scenario. The number of samples required to confirm the presence or absence of contamination was derived in the set scenario as the contamination concentration of the radiochemical laboratory was unknown to the decision-maker. The points depicted in the figure are representative points selected from the population defined using the systematic grid pattern. The number of derived samples is a statistically independent value and is not affected by other factors such as incorrect sample handling, transportation, processing, or storage. When selecting the actual collection point, it may be reasonable to exclude the middle location of the green space that is difficult to access among the spots or the building entrance point if it is difficult to enter the facility.

Table 3 describes the input parameters of VSP used to derive the number of samples.

VSP calculates the number of samples required to reject the null hypothesis in favor of an alternative one, given a selected sampling

Table 1
Design information of scenario regarding DQO steps 1-5.

Information		Contents	DQO step
Sample collection occasion		• Non-official	1
Sample collection objective		• The collection of samples is necessary to determine the presence or absence of nuclear activity in the survey site	
Sample collection agent		• Regulatory agency and unspecified organization or person selected by the nation	
Survey methodology and target number of samples		• Probability-based • The number of samples to be collected to determine the mean and standard deviation values that can represent the survey area should be secured to the maximum possible number considering the limitations.	2
Site information		• Pollution sources: waste treatment facility, spent fuel rod storage (presumption) • Assess route: facility's entrance and the road close to the facility boundary • Topography: concentration of buildings within the survey site and green space centering around the facility	3
Sampling design		• Survey area: 357,616 m^2 • Sampling density: regional phase • Sampling pattern: stratified random sampling • Type of sample: point sample	
Limitation	Restrictions	• Non-cooperative of jurisdiction, unclear budget	4
	Spatial	• Survey area (357,616 m^2)	
	Temporal	• Very high variability. Assuming a short stay (moving a vehicle around the facility and a brief inspection)	
Decision rule	Samples	•Samples collected from a point close to the source of contamination are first designated as the target of analysis. Otherwise, the analysis results of samples containing a significant amount of uranium and plutonium are used.	5
	Analysis	• If there is a significant difference based on the ratio of isotopes in the natural state, nuclear activity is suggested. Otherwise, cross-validation is used to compare the results with those of other analytical institutions.	

Table 2
Design information of scenario regarding DQO step 6 and 7.

Information	Contents	DQO step
The number of samples	• Adjust the permissible parameters of interest (number of samples)	6
Gray region	• Scope to minimize the impact of decision-making errors	
Design optimization	• Suggest a sample collection plan that can meet all the requirements specified in the previous step - Determine the optimal sample collection method - Statistical hypothesis testing	7

approach and inputs to the associated equation. For this sampling design, random point sampling in grids was selected, offering a good balance between providing information regarding the spatial structure of the potential contamination and ensuring all portions of the site are represented (although not as thoroughly as in systematic grid sampling). The equation used to calculate the number of samples is based on a sign test [11]. For this site, the null hypothesis is rejected in favor of an alternative one if the median (mean) is sufficiently smaller than the threshold. The number of samples to be collected is calculated so that if the inputs to the equation are valid, the estimated number of samples causes the null hypothesis to be rejected.

If the secured site information changes, the allowable confidence interval for evaluation is likely to change as well. Therefore, sensitivity analysis was performed using the same input parameters. As presented in Table 4, the number of samples collected based on the change in the contamination criteria and confidence interval was derived.

Once the budget for sample collection has been determined, or when statistically reliable limits have been determined, the number of samples to be collected can be changed, as indicated in Table 4. If more apparent quantitative data such as pollution concentration and pollution criteria are available, the steps can be repeated for new statistical assessments. Furthermore, in step 6, when the factual value is significantly close to the reference value, the gray region can be specified within the range of the values determined for making a decision. The gray region represents the area where the results are relatively insignificant due to decision errors, and the range of gray region can be adjusted from 20% to 95% depending on the confidence level. The gray region must be determined in consideration of the judgment of regulatory agencies and experts, costs, and benefit assessment results.

3.2. *Cost optimization of sample collection plan*

Because this study assumes a scenario in which the specific contamination level of a site is unknown, statistical assessment using the site contamination concentration cannot achieve a meaningful optimization. Additionally, as the value and importance of a sample change depend on the sample collection point, it is necessary to use different sample collection methods for cost-effective sample collection. For example, at the four key locations mentioned in Section 2, the samples collected require sophisticated analysis, thereby increasing the sampling costs and analysis costs. Therefore, scenario optimization was performed through cost evaluation. In general, if the correlation value of the two factors (the number of samples and cost) compared in the statistical calculation result exceeds 0.5, then the two factors are determined to be significantly correlated. Therefore, in this study, the correlation value between using a high budget and a low budget was set to 0.75. The same values were applied for deriving the number of samples and positions of the other inputs.

Table 5 lists the input parameters used for the analysis. The total number of cost-optimized collection samples was 31, among which 8 samples were collected using an expensive approach, and 23 samples were collected using a cost-free (inexpensive) approach. Here, the “expensive approach” means that the type and condition of the sample required differs according to the selected analysis method (equipment), and thus, various items necessary for the sample analysis such as sample reprocessing and equipment

Fig. 2. Twenty-nine sample collection locations (blue points) in the survey area. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Table 3
VSP input parameters for deriving the number of samples and spots.

Parameter	Value
Primary objective of design	Comparison of a site mean or median to a fixed threshold
Type of sampling design	Nonparametric
Sampling pattern	Random sampling in grid (interval of grid is 100 m)
Working (null) hypothesis	The median(mean) value at the site exceeds the threshold
Cost	Total $15,000 (This is the default cost in VSP, which includes the plan and verification cost ($1,000), on-site collection cost ($100), and lab analysis cost ($400))
Specified sampling area	357,616 m^2
Confidence level [%]	90%
p-value	Required percent of the population to be less than the action level

operation are considered. This means that it costs more than a regular sample. The pink points in Fig. 3 indicate the points that need to be collected using an expensive approach, and the blue points are the points that have no cost limit. As summarized in Table 6, the accuracy of location may be lowered if the sample collection method is selected, but the sample collection can be performed at a lower budget.

Table 4
Sensitivity analysis result with respect to the number of samples.

Number of samples					
C.L P	99	97	95	93	91
85	22	22	19	17	15
90	44	34	**29**	26	23
95	90	69	59	52	47

*P: p-value, C.L: confidence level.

Sensitivity analysis was performed to confirm the difference of sample collection method according to the degree of correlation for each sample collection price. The results of the sensitivity analysis were calculated using the statistical formulas built into the VSP [12]. The statistical hypotheses applied to the sensitivity analysis are as follows, and the analysis results are presented in Table 7.

- The relationship between the costs of the expensive and inexpensive methods is linear.
- C_{EX} (expensive cost) is a reasonable price.
- Measured values are typically distributed or approximately follow a normal distribution.
- For a regression line, the variance is uniformly distributed along the entire line.

4. Discussion and conclusion

In this study, a soil sample collection scenario was established to determine the presence or absence of nuclear activity in a radiochemical laboratory, one of the key facilities for denuclearization, and a DQO-based optimization method was proposed. The application of the DQO process can overcome the limitations of the scenario constructed using limited information such as the operation history, including the facility layout, movement route, and surrounding environment based on satellite imagery. Soil samples collected following systematic procedures such as the DQO process demonstrate sufficient analytical value to estimate the history or verify the nuclear activity.

To play a dominant role in the verification of denuclearization in

Table 5
VSP input parameters for cost optimization.

Parameter	Value
Primary objective of design	Comparison of a site mean or median to a fixed threshold
Assumed correlation between two cost method, ρ	0.75
Sample placement method used in the field	Random sampling in grid
Working (null) hypothesis	The median(mean) value at the site exceeds the threshold
Cost (default)	Expensive method: $ 400, Inexpensive method: $ 60
Specified sampling area	357,616 m^2
Is cost efficient compared to simple random sampling doing only expensive analysis measurement?	Yes

Fig. 3. Cost-optimized sample collection spots.

Table 6
Optimized sample collection cost.

Collection method	Number of samples	Price per sample	Cost
Expensive, C_{EX}	23	$400	$1,380
Inexpensive, C_{IEX}	8	$60	$3,200
Total			$4,580

the future, it is necessary to secure sufficient information in advance in case a sample collection opportunity arises. In this study, scenarios were set based on the previously available information, and a sample collection plan based on statistical analysis was established. However, a more reliable and concrete sample collection plan can be established if the budget and information on participating human resources are secured. In addition, on-site

Table 7
Optimized sample collection cost.

	Number of samples: n_{ex}/n_{inx}		
	C_{EX} = $ 200	C_{EX} = $ 400	C_{EX} = $ 600
ρ = 0.65	11/17	10/21	9/25
ρ = 0.7	10/18	9/22	9/26
ρ = 0.75	9/18	**8/23**	8/26
ρ = 0.8	8/18	7/23	7/27
ρ = 0.85	7/18	6/23	6/27

*ρ: coefficient of correlation, n_{ex}: number of samples collected using expensive methods, n_{inx}: number of samples collected using inexpensive methods.

information such as the concentration of the nuclides of interest and contamination concentration of nuclides such as U-238 and Pu-239 in the study area is required.

The sampling plan evaluated using statistical methods is highly variable depending on various factors such as decision-makers, experts participating in the sample collection, opinions of regulatory agencies, and domestic and foreign political situations. When the amount of available information increases and an uncontrollable situation occurs, the sample collection plan needs to be revised; therefore, it is necessary to establish a system to control it.

Declaration of competing interest

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

Acknowledgement

This work was supported by the Nuclear Safety Research Program through the Korea Foundation Of Nuclear Safety (KoFONS) using the financial resource granted by the Nuclear Safety and Security Commission (NSSC) of the Republic of Korea. (No. 2004025).

References

[1] J.H. Moon, B.G. Park, Consideration for the successful verification and dismantlement of North Korea's Nuclear Program, J. Nucl. Fuel Cycle Waste Technol. 7 (2009) 143–151.

[2] Y.H. Park, Case Studies of International Denuclearization and Implications for Denuclearization in the D.P.R.K, Academic Cooperation Foundation Kyung Hee University, Industry, 2019. NSTAR-19PS61-177.

[3] P. Makowsky, F. Pabian, J. Liu, O. Heinonen, North Korea's Yongbyon Nuclear Scientific Research Center: Satellite Update, 2021. https://www.38north.org/.

[4] United States environmental protection Agengy, Guid. Syst. Plan, Using Data Qual. Objectives Process, 2006. EPA QA/G-4.

[5] American National Standard Institute, Characterization in Support of Decommissioning Using the Data Quality Objectives Process, 59, American National Standards Institute/HPS N13, 2008.

[6] B.D. Vivo, H. Belkin, A. Lima, Environmental Geochemistry-Site Characterization, Data Analysis and Case Histories, second ed., Elsevier, 2018.

[7] 38North. http://38northdigitalatlas.org/, Atlas Press.

[8] G.Y. Lee, Review of domestic and foreign technical standard and research trends for the establishment of soil sampling procedure for safeguards inspection, Korea Inst. Nucl. Nonproliferation Control (2019). NSTAR-19PS51-169.

[9] International Atomic Energy Agency, Nuclear Forensics in Support of Investigations, 2015. Nuclear Security Series No. 2-G (Rev.1).

[10] B.D. Matzke, L.L. Newburn, J.E. Hathaway, L.M. Bramer, J.E. Wilson, S.T. Dowson, L.H. Sego, B.A. Pulsipher, Visual Sample Plan Version 7.0 User's Guide, Pacific Northwest National Laboratory, 2014. PNNL-23211.

[11] R.O. Gilbert, J.R. Davidson Jr., J.E. Wilson, B.A. Pulsipher, Visual Sample Plan (VSP) Models and Code Verification, Pacific Northwest National Laboratory, 2001. PNNL-13450.

[12] R.O. Gilbert, Statistical Methods for Environmental Pollution Monitoring, Van Nostrand Reinhold, New York. (Now Published by, Wiley & Sons, New York, 1997), 1987.