Application of Bayesian statistical analysis to on-line seawater radioactivity monitoring

Do Cong Binh [a*], Dong Myung Lee [b]
[a] *Korea Advanced Institute of Science and Technology, 291 Daehak-ro, Yuseong, Daejeon, Republic of Korea, 34141*
[b] *Korea Institute of Nuclear Safety, 62 Gwahak-ro, Yuseong, Daejeon, Republic of Korea, 34142*
[*] *Corresponding author: congbinh83@kaist.ac.kr*

1. Introduction

After the Fukushima nuclear accident in Japan, many countries worried about marine contamination due to artificial radionuclides released from nuclear accidents. Vietnam, in particular, is very interested in an on-line seawater radioactivity monitoring system because we have a vast ocean in our territory. The goal of on-line radiation monitoring is to quickly detect small or abrupt changes in radiation levels due to domestic or especially overseas nuclear accidents early in a nuclear emergency. In cases of low-level radioactivity in a marine environment, it is very difficult to distinguish between a radioactive source and background radiation from natural radionuclides in the seawater.

Generally, the detection decision about whether a radioactive source is present is made based on a specific statistical method. In practice, a proper statistical method is chosen to minimize the false negative rate (β) while holding the false positive rate (α) at desired levels [1]. A conventional radiation monitoring method based on the classical statistics involves setting a decision level (DL) for a given false positive rate, and then a count or count rate is compared to the decision level. If the value of the result is greater than the decision level, the decision that there is activity present above the background is made. The conventional monitoring method can be conducted easily; however, a disadvantage is that only information contained in the latest data point is exploited. Furthermore, the information contained by the entire sequence of data points is disregarded [2].

Unlike classical statistics, Bayesian statistics permits the formal incorporation of prior subjective knowledge, belief, and information beyond that contained in the observed data in the inference process via Bayes' theorem. Bayesian techniques have been applied to reducing false positive rates in low-level radioactivity monitoring [3].

In this study, we apply Bayesian statistics to the analysis of count rate data in the region of interest (ROI) for ^{137}Cs on the spectrum from on-line seawater radioactivity monitoring. In addition, we compare false positive and false negative rates of time series data for classical and Bayesian statistical process control charts.

2. Theory and Methods

Radioactive decay is a random process. Consequently, any measurement based on detecting the radiation emitted in nuclear decay is subject to some degree of statistical fluctuation. The Poisson distribution characterizes the random nature of radioactive decay when the number of nuclei is large and the observation time is short compared with the half-life of the radioactive species. For a Poisson process, the probability P(k) for observing k decays (k =0, 1, 2…) in a time interval t is given by [3]

$$\mathrm{P(k)} = \frac{(rt)^k}{k!}e^{-rt} \tag{1}$$

where r is the mean count rate.

2.1 Bayes' Theorem

The blank and sample measurements in a specific channel are described by the following Poisson distributions respectively [4]

$$\mathrm{P}(b|r_b t) = \frac{(r_b t)^b}{b!}e^{-r_b t} \tag{2}$$

$$\mathrm{P}(s|r_s t) = \frac{(r_s t)^s}{s!}e^{-r_s t} \tag{3}$$

where b is background counts recorded in the specific channel for a count time t and s is sample counts recorded in the same channel for a count time t. The background mean count rate in the channel is given by r_b and the sample mean count rate is given by r_s.

The mathematical form of Bayes' theorem [3] is defined as

$$\mathrm{P}(r|x) = \frac{P(x|r)P(r)}{P(x)} \tag{4}$$

where $\mathrm{P}(r|x)$ is the posterior probability distribution of the unknown parameter r given the data. $\mathrm{P}(x|r)$ defines the probability to obtain an observation x if the process is in the state of r, this is the likelihood function of x given r. The prior probability distribution of r is given as $P(r)$, this is the quantitative description of what we believe about r based on previous experience and knowledge before the experiment is conducted. The $P(x)$ is referred to as the marginal distribution of the data. As a result, Bayesian statistical analysis is used to obtain the posterior probability which summarizes our knowledge of the parameter r give the prior belief and the actual data x.

The posterior probability calculated with Bayes' theorem is a gamma distribution. The gamma distribution is the conjugate prior to the Poisson distribution. This is the reason why we use gamma functions for the prior probability distribution of the background spectrum and sample spectrum. In the case of a time series of independent measurements obtained in radiation monitoring, Bayesian analysis can be conducted sequentially. The general probability density function of the gamma distribution is given by [1]

$$\mathrm{Gamma}(\alpha, \beta) = \frac{\beta^{\alpha}}{\Gamma(\alpha)} r^{\alpha-1} \mathrm{e}^{-\beta r} \qquad (5)$$

where α is the shape parameter, β is the reverse scale parameter. The parameter r is the mean count rate of the system that can be estimated based on the measured count rate.

With the count rate in the ROI for ^{137}Cs on spectrum data, the prior probability distribution of the background and sample are described by the following gamma distributions respectively [4]

$$\mathrm{G}(b_0, t_0) = \frac{(t_0)^{b_0}}{\Gamma(b_0)} r_b^{b_0-1}\, e^{-r_b t_0} \qquad (6)$$

$$\mathrm{G}(s_0, t_0) = \frac{(t_0)^{s_0}}{\Gamma(s_0)} r_s^{s_0-1}\, e^{-r_s t_0} \qquad (7)$$

where we define $\mathrm{b} = \mathrm{b}_0 + \mathrm{b}_1$, as the sum of prior and measured counts for the background measurement in the ROI, similarly $\mathrm{s} = \mathrm{s}_0 + \mathrm{s}_1$, as the sum of prior and measured counts for the sample measurement in the ROI, $\mathrm{t} = \mathrm{t}_0 + \mathrm{t}_1$, as the sum of prior and measured count time.

Similarly, the count information obtained in a time series of fixed-count time, $x = (x_1, x_2, \ldots, x_n)$, the posterior probability of r is given as [1]

$$\mathrm{P}(r|x) \propto P(x|r)P(r) = Gamma(\alpha + \textstyle\sum_n x_i \ \ , \ \beta + n) \quad (8)$$

If we can calculate the probability that the sample count rate (r_s) in the specific ROI is greater than the background count rate (r_b), It gives a decision rule that does not rely on the null hypothesis.

2.2 Experimental methods

The Korea Institute of Nuclear Safety (KINS) developed the network of on-line seawater radioactivity monitoring system and is now test-operating the NaI(Tl) spectrometry systems (SI Detection, Korea). Fig. 1 shows the schematic diagram of the spectrum data acquisition and analysis system used for experimental count rate data. Gamma radiation (E_γ=661.6 keV) from a ^{137}Cs radioactive source (~ 37 kBq, 1-May-2015) was used to experimentally characterize the response of the NaI(Tl) scintillation detector. The ^{137}Cs radioactive source was placed with a distance of 80 cm between the source and the detector. The ROI for ^{137}Cs was set for 661.6 keV full energy peak within 5, 10, 20 channels from side to side. Background count rate data were collected for several days in 15 min. intervals without the source. The performance of statistic methods (Bayesian and classical) was evaluated in terms of average run length (ARL) and detection probability (1-β).

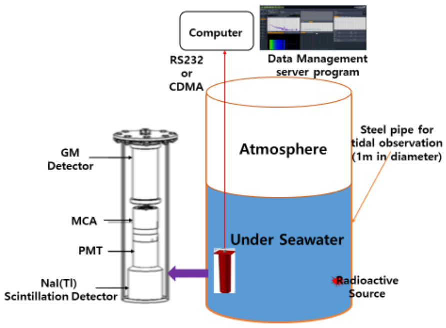
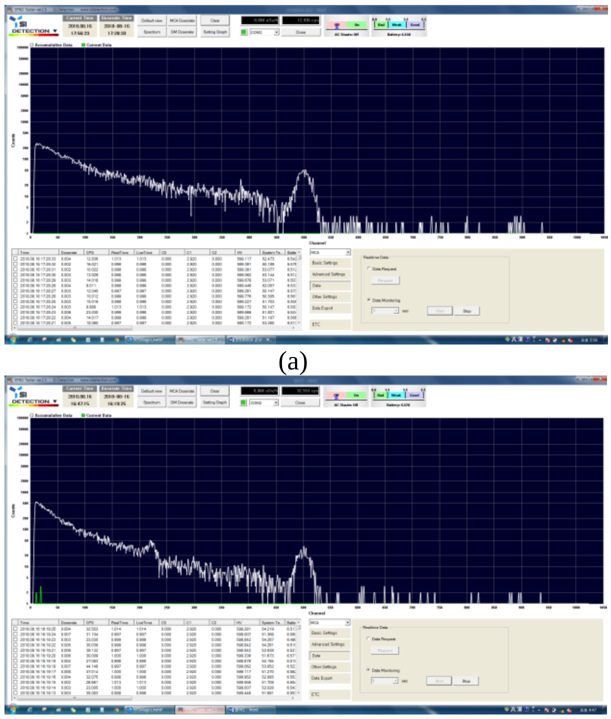


Fig. 1. Schematic diagram of a seawater radioactivity monitoring system.

Time series count rate data were evaluated using the classical 3-σ and cumulative sum(CUSUM) control charts as well as the Bayesian Shiryayev-Roberts(S-R) chart.

3. Results and Discussions

To verify the feasibility of the application of Bayesian statistical analysis, the background pulse height spectrum (PHS) and the source PHS were measured in the tidal observation station for the purpose of off-line experimentation. The typical measured PHSs are shown in Fig. 2.



(a)

(b)

Fig. 2. Measured pulse height spectrum in 15 min under seawater. (a) background (without source) and (b) ^{137}Cs radioactive source (40 cm distance).

In the off-line experiments, the ^{137}Cs radioactive source was positioned at 40 cm, 60 cm, 80 cm distances approximately from scintillation detector. As the results, the count rate in the ROI for ^{137}Cs positioned at 80 cm was nearly similar to that of the background spectrum without any sources. This result implies that our data sets can properly be used to quantify the false positive rate and the ARL.

The average background count rate of 1.05 ± 0.02 cps in the ROI of ^{137}Cs. In this case, the time series count rate data sets of 740 were used to analyze the statistical control charts. Fig. 3 shows the on-line monitoring data during August in 2018.

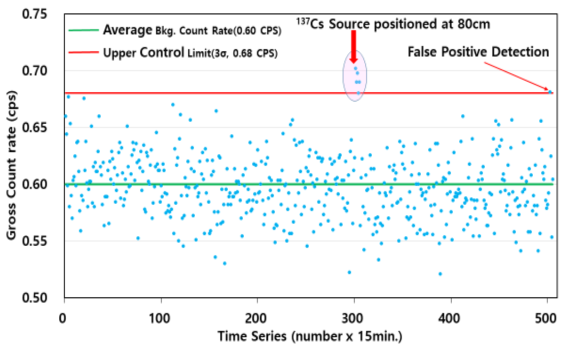


Fig. 3. Monitoring data for 3-σ control chart statistical analysis

When the 3-σ control chart methodology was applied to the monitoring data sets preliminary to full-scale study, for the average background count rate of 0.60 ± 0.027 cps, the corresponding upper control limit (UCL) was calculated to be 0.68 cps. After analyzing the 3-σ results, the false positive detection rate of 0.2% (ARL_0=500) based on 500 data sets, Fig. 3. This value is about 1.5 times the expected value of 0.135% (ARL_0=741).

At present, the CUSUM analysis and S-R analysis are applied to the data of off-line experiment and also the off-line experimental data sets are used to evaluate the ARL and detection probability rate in this study.

4. Conclusions

The use of CUSUM or S-R control charts can be advantageous for low-level radioactivity analysis in marine environments. The count rate probability distribution can be used to calculate the probability that a pulse height spectrum in seawater radioactivity monitoring applications contains a radioactive substance. Also, it can take into account information about the presence of net activity in a batch spectrum by limiting the probability distribution to positive values of net count rate. Continued investigation of the application of Bayesian statistical analysis to on-line seawater radioactivity monitoring will be valuable.

REFERENCES

[1] P. Luo, J. L. Sharp, and T. A. DeVol, Bayesian Analysis of Time-Interval Data for Environmental Radiation Monitoring, IEEE Transactions on Nuclear Science, Vol. 57, p.493, 2010.

[2] T.A. DeVol, A. A. Gohres, and C. L. Williams, Application of classical versus Bayesian statistical control charts to on-line radiological monitoring, Journal of Radio-analytical and Nuclear Chemistry, Vol.282, p.933, 2009.

[3] J. Klumpp and A. Brandl, Bayesian Analysis of Energy and Count Rate Data for Detection of Low-Level Count Rate Radioactive Sources, Health Physics, Vol. 108, p.364, 2015.

[4] F. Lemay, Multichannel Statistical Analysis of Low-Level Radioactivity in the Presence of Background Counts, Health Physics, Vol. 103, p.770, 2012.