Application and improvement of continuous monitoring methods for artificial radionuclides based on Bayesian statistics

Xiang Li[1,2] · Qianhong Huang[1] · Yuxi Xie[3] · Xueyu Gong[1]

Abstract

This paper delves into the Bayesian statistics applications of three preeminent models, Poisson distribution, Gaussian distribution, and Binomial distribution, in the continuous surveillance of artificial radionuclides. It introduces a slide-window method to accelerate the updating of the prior distribution of model parameters and compares the performances of three models before and after utilizing this method. Comparisons among the three models are made before and after using the slide-window. Experimental results demonstrate a marked enhancement in the performances of all models.

Introduction

The Guide to the Expression of Uncertainty in Measurement (GUM) by the International Organization for Standardization (ISO) is considered the de facto standard for evaluating and expressing uncertainty in measurement [1]. The International Organization for Standardization's Guide to the Expression of Uncertainty in Measurement (GUM-ISO) advocates for the representation of the measurement process through a measurement equation $Y = f(X_1, X_2, \ldots X_n)$. Within the context, Xi represents the measurement conditions or inputs, categorized as Type A variables; whereas Y denotes the outcome of the measurement, categorized as a Type B variable. For Type A quantities, the guide suggests their evaluation using conventional statistical methods. This involves characterizing the true value with the sample mean and the uncertainty of Xi with the sample variance or standard deviation. Conversely, for Type B quantities, the ISO-GUM suggests the application of alternative scientific approaches, with Bayesian statistics being particularly prevalent.

According to the Bayesian theory, the estimator Y is treated as a random variable, initially described by a probability distribution with known parameters. To determine these parameters, Bayesian approaches utilize past knowledge to construct a prior distribution. This prior distribution evolves through an iterative process into a posterior distribution as new measurements are obtained. The central tendency (such as the posterior mean or median) and spread of this posterior (such as the posterior standard deviation) then quantitatively describe the measured results of the parameter and its associated standard uncertainty. The discrepancies in the interpretation of uncertainty, thus provoked, have been a recurrent subject of discussion [2–4], with some even supporting for a complete substitution of the ISO-GUM to be grounded entirely in Bayesian statistics in the realm of measurement science [5].

In the application of Bayesian approaches, it is often presumed that the measurement process is directed by a probabilistic model contingent on certain parameters. The determination of these parameters' states is derived from actual data, factoring in any available prior insight. Consequently, the ensuing measurement outcome is typically represented as the expectation of the posterior probability

✉ Xueyu Gong
  gongxueyu@126.com

1 School of Nuclear Science and Technology, University of South China, Changsheng West Road 28, Hengyang 421001, China

2 School of Computer Science and Technology, HengYang Normal University, HengHua Road 16, Hengyang 421001, China

3 School of Physics and Electronic Engineering, HengYang Normal University, HengHua Road 16, Hengyang 421001, China

distribution, with the associated uncertainty depicted by a confidence interval [6]. This leads to an engaging query: In the Bayesian framework, for a particular quantity of interest in the measurement process, would the outcomes vary if different statistical models (likelihood functions) were employed to articulate the quantity?

As with the majority of measurements, the artificial radionuclides monitoring also adheres to the framework of the ISO-GUM, specially concerning online monitoring that necessitates the subtraction of natural background [7, 8], a process that is far from trivial. The principal difficulties are twofold: Firstly, due to the stochastic decay of natural background, the particle flux registered by the detector is subject to a Gaussian or Poisson distribution governed by specific parameters; secondly, the online surveillance of artificial radionuclides occurs in the presence of a natural background, where the sparse counts are easily obscured by background counts. In most cases, the statistics employed to characterized the background (typically the mean count) is deemed to follow a Poisson or Gaussian distribution with defined parameters. Conversely, in select scenarios, such as with delayed coincidence systems, the background level is characterized by the ratio of short-lived progeny counts to the total counts within the natural radon, thorium decay chain, a ratio that can be perceived as a parameter for a binomial distribution [9]. Hence, there are at least three distinct analytical perspectives in estimating the background counts: using Poisson, Gaussian and Binomial distribution respectively. Zabulonov et al. [10] have introduced an expedited method for detecting radioactivity based on the normal distribution model. In their work, Zabulonov treats the natural background counts as normal distributed, and he devises another normal distribution as a prior of the expectation using historical data, and calculates the posterior distribution to evaluate the background radiation. Pyke et al. [11] try using the Gamma distribution as a prior for the Gaussian distribution to tighten the bounds of the confidence interval. Complementarily, Qingpei et al. [12] have employed the normal distribution to depict the characteristic γ peaks of radionuclides to identify various radioactive species by using Bayesian sequential analysis. Concurrently, Luo et al. [13] have depicted the natural background through a Poisson distribution, resorting to the Gamma distribution as a prior for the parameter of the Poisson. Furthermore, Dailibor Nosek have experimented two separate Poisson distributions with distinct parameters for source counts and backgrounds counts by opting for Gamma, uniform, and Jeffrey priors as priors of parameters [14]. Li et al. [9] have captitalized on the ratio of short-lived progeny to total α counts within the radon-thoron decay series to construct a binomial distribution model, employed a Beta distribution as a priori knowledge to infer the posterior

parameters, and facilitated the model for ongoing surveillance of artificial radionuclides contamination.

This paper investigates the applications of Bayesian statistical models in the continuous monitoring of artificial radionuclides. The principal objective is to dissect the disparities and origins in the measurement outcomes when employing Gaussian, Poisson and Binomial distribution models for artificial radionuclides monitoring. The quest is for a universally applicable method within Bayesian modeling that diminishes uncertainty. Accordingly, the main contributions of this paper are: (1) Describe the basic methodologies and principles for continuous surveillance of natural background radiation and artificial radionuclides using the models in detail, including the selection of prior distribution, the calculation of posterior distribution and the construction of the confidence interval; (2) Introducing a slide-window method to speed up the update of prior distribution parameters; (3) Evaluating and comparing the performances of models before and after using slide-window method under varying natural background conditions.

Theory

Bayes' theorem

Consider a measurement process involving a quantity of interest, denoted by θ. Before the measurement start, we construct a distribution $\pi(\theta)$ based on historical data or expert knowledge. With the refinement from measured data x, the posterior distribution of θ, denoted by $\pi(\theta|x)$, can be represented as shown in Eq. 1:

$$\pi(\theta|x) = \frac{P(x|\theta)}{\mathrm{m}(x)}\pi(\theta) \tag{1}$$

Here, P(x|θ) represents the likelihood of observing data x given the parameter θ, and m(x) is the marginal distribution of x, which is independent of θ. Therefore, the posterior distribution of θ can also be written in a form proportional to the product of the likelihood and the prior:

$$\pi(\theta|x) \propto \mathrm{P}(x|\theta)\pi(\theta) \tag{2}$$

It is clear that the posterior distribution's mathematical form is influenced by both the likelihood and the prior distribution. In theory, $\pi(\theta)$ can be any distribution for any parameter. However, an incompatible prior with the likelihood often results in a posterior distribution excessively complex and intractable. Consequently, it is customary to select a prior from the conjugate family of distributions that is compatible with the likelihood's form. Alternatively, we might opt for a uniform distribution as a default description of θ even we have no knowledge about it. References

[15–17] point out that perturbing the prior distribution does not necessarily affect the outcome of Bayesian estimations. Yet, the influence of the likelihood $P(x|\theta)$ on the posterior distribution, despite its significance, is seldom addressed. Given that all inferences in Bayesian statistics are founded on the posterior, we believe that varying likelihood functions will exert distinct impacts on the posterior outcomes.

After obtaining the posterior distribution $\pi(\theta|x)$ for the parameter θ, one can construct a credible interval with a confidence level of 1-α (where α is typically chosen as 5%), denoted as $[\hat{\theta} - k\sigma, \hat{\theta} + k\sigma]$, where k = 3 represents the critical value that defines the bounds of the credible interval, $\hat{\theta}$ and σ represent the expectation and the standard deviation of the posterior distribution. If the expectation falls within this interval, the sample is regarded to be free of issues; otherwise, it is deemed contaminated. In the process of continuous monitoring, the Bayesian approaches calculate the first posterior distribution by choosing a "well-designed" prior and likelihood. Then they sequentially use the previous posterior as the new prior in subsequent measurements, continually updating the posterior in combination with the actual data.

The *Poisson distribution* model

To characterize the natural background using a Poisson distribution, two conditions must be met. The first is that the observation time must be sufficiently short in comparison to the half-life of ^{222}Rn (3.83d). The second is that the number of decay events observed during the monitoring period should be sufficiently high. If we denote the Poisson distribution parameter as θ, then the likelihood of observing x decays (where $x = 0, 1, 2, \ldots$) is given by the probability $P(x|\theta)$, which also serves as the likelihood function with respect to θ:

$$P(x|\theta) = \frac{\theta^x}{x!} e^{-\theta} \qquad (3)$$

The goal of Bayesian analysis is to obtain the posterior distribution of θ, which represents a synthesis of the prior belief about θ and the actual data x. In this study, given that the likelihood function for θ follows a Poisson distribution, a Gamma distribution is employed as the conjugate prior for θ. Thus, the posterior distribution of θ can also be expressed in the form of a Gamma distribution. The probability density density function of a Gamma distribution, denoted as *Gamma(a,b)*, is:

$$\pi(\theta) = \mathrm{Gamma(a,b)} = \frac{b^a}{\Gamma(a)} \theta^{a-1} e^{-b\theta} \qquad (4)$$

where *a* represents the shape parameter, *b* is the rate parameter, and $\Gamma(a)$ is the Gamma function. When there is no

knowledge about θ, one can assign a Jeffrey prior to θ, which is given by $a = 1/2$, $b = 0$ and $\pi(\theta) = \theta^{-1/2}$[14]. Combining with Eqs. (2), (3) and (4), the posterior distribution for θ, can be expressed as:

$$\pi(\theta|x) \propto P(x|\theta)\pi(\theta) = Gamma(a + x, b + 1) \qquad (5)$$

The detailed derivation of the posterior distribution for θ, along with the computation of its posterior expectation and variance, is presented in Appendix A. It is possible to establish a credible interval for natural background level. If the total alpha counts of a new sample is within the interval, the sample is deemed normal, and its information is incorporated into the prior distribution for the next measurement. Conversely, if the alpha counts lie outside this interval, it indicates a likely contribution from artificial radionuclides. Such samples should raise alarms and will not be included in the prior for the subsequent measurement.

The Gaussian distribution model

It is more common and more complex to describe the natural background as a Gaussian distribution [18]. Let the random counts of the background x follow a Gaussian distribution with mean θ and variance σ^2, denoted as $x|\theta \sim N(\theta, \sigma^2)$. The likelihood function for θ is then given by:

$$P(x|\theta) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{ -\frac{(x - \theta)^2}{2\sigma^2} \right\} \qquad (6)$$

where σ^2 is assumed to be known and can be calculated from historical data [8]. Therefore, Gaussian distribution with parameters μ_0 and σ_0 is served as the conjugate prior for θ. This choice ensures that the posterior distribution for θ will also be a Gaussian distribution. The mathematical form of the conjugate prior distribution for θ is:

$$\pi(\theta) = N(\mu_0, \sigma_0^2) = \frac{1}{\sqrt{2\pi}\sigma_0} \exp\left\{ -\frac{(\theta - \mu_0)^2}{2\sigma_0^2} \right\} \qquad (7)$$

When combined with the likelihood function, the posterior distribution for θ can be derived by applying Bayes' theorem. Due to the complexity of the posterior distribution for θ, we provide the expectation α and variance β of the posterior here. For more detailed derivation can be found in Appendix B:

$$\begin{aligned} \alpha &= \left(\frac{\mu_0}{\sigma_0^2} + \frac{n}{\sigma^2}\bar{x} \right) \Big/ \left(\frac{n}{\sigma^2} + \frac{1}{\sigma_0^2} \right) \\ &= \frac{\sigma_n^{-2}}{\sigma_n^{-2} + \sigma_0^{-2}} \bar{x} + \frac{\sigma_0^{-2}}{\sigma_n^{-2} + \sigma_0^{-2}} \mu_0 \end{aligned} \qquad (8)$$

$$\beta = \left(\frac{n}{\sigma^2} + \frac{1}{\sigma_0^2} \right)^{-1} \tag{9}$$

In the Eqs. (8) and (9), n and $\bar{x}$ represent times of observations and the mean counts respectively. If we denote sample mean square error as σ_n^2, which is $\sigma_n^2 = \sigma^2/n$, then Eq. (8) can be written as the second form. On one hand, in the context of artificial radionuclides monitoring, each measurement yields one sample which means $n = 1$ and $\sigma_n^2 = \sigma^2$. On the other hand, the posterior expectation α for θ can be regarded as a weight average of the sample mean x and the prior mean μ_0. One can tell that the weights are inversely proportional to prior variance σ_o^2 and sample variance σ_n^2. If the sample mean's variance σ_n^2 is small, indicating precise measurements, then the weight on the sample mean x is large, and vice versa. The judgement whether a sample is contaminated is similar to that used with the Poisson distribution.

The Binomial distribution model

The natural background radiation composed of radon and thorium progeny, alpha emitting radionuclides originating from radon decay products significantly impact the online monitoring artificial radionuclides. The total alpha count is primarily attributed to ^{218}Po and ^{214}Po, with the ratio of ^{214}Po counts to the total alpha count denoted as θ and regarded as a stable constant [19]. Consequently, the ratio θ together with the total alpha counts from the natural background follow a binomial distribution. With knowledge of the ^{214}Po counts and the ratio θ, one can estimate the approximate range of the total natural background alpha counts. The methodology for calculating the ^{214}Po counts which has been detailed discussed in previous work is identical to this paper [9]. Furthermore, this paper selects the Beta distribution as the prior for θ based on the likelihood function's characteristics. This choice make sure that the posterior distribution for θ retains the same mathematical form. The probability density function of the Beta distribution is given by:

$$Beta(a, b) = \frac{\Gamma(a + b)}{\Gamma(a)\Gamma(b)} \theta^{a-1}(1 - \theta)^{b-1} \tag{10}$$

where a is the shape parameter indicating the number of ^{214}Po counts, and b is the scale parameter representing the total alpha counts except for ^{214}Po. It's worth noting that the uniform distribution $U(0, 1)$ is a special case of the Beta

distribution $Beta(0, 1)$ and can be used as uninformative prior for θ. In the measurement, the relationship between ^{214}Po counts denoted as x and the total alpha counts n can be represented as $\binom{n}{x} \theta^x (1 - \theta)^{n-x}$, which serves as the likelihood function. According to Bayes' theorem, the posterior distribution can be expressed as (See Appendix C for more details):

$$\pi(\theta|x) \propto P(x|\theta)\pi(\theta) = Beta(a + x, b + n - x) \tag{11}$$

The detailed derivation of the posterior distribution for θ, along with the computation of its posterior expectation and variance, is presented in Appendix C. We can similarly construct a credible interval to describe the natural background based on the posterior expectation and variance. Since θ comprises ^{214}Po counts and total alpha counts, it is not appropriate to compare the ratio of a new sample with the interval, as this ratio diminishes the impact of artificial radionuclides contribution. Instead, one should first calculate the posterior expectation of the ratio in the current sample and then compare it with the credible interval from the previous measurement. If the expectation falls within the interval, the sample is adjudged to be normal; conversely, contamination.

Improvement and evaluation of Bayesian models

In the process of background monitoring, the continuous incorporation of natural background samples leads to a progressive accumulation of prior distribution parameters. Consequently, the computation of posterior expectation estimates and decision intervals becomes challenging. Additionally, as the natural background changes constantly due to factors such as atmospheric conditions and wind direction, long ago information contained in the prior might be less valuable for the current assessment. Hence, the timely elimination of outdated information from the prior distribution is crucial and beneficial for both the posterior computation and estimations of the current measurement. This paper introduce a slide-window method to periodically purge the distant historical information form the priors. The principle of the slide-window method is illustrated in Fig. 1, with the following specifics: (1) The algorithm upholds a window of fixed length which can be customized and a window length of 10 is employed in this paper; (2) The windows logs the prior parameters in a first-in-first-out queue, with each data-set entry occupying a

Fig. 1 Illustration of the slide-window method
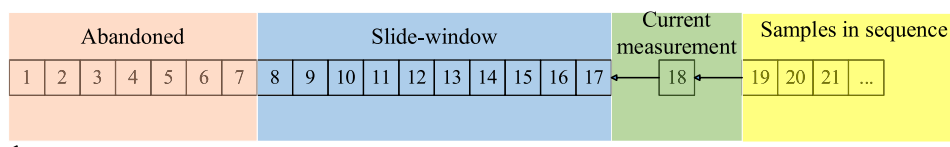
unit of length; (3) Upon the window's record length reaching its capacity, the most ancient record is automatically abandoned.

Upon concluding the analysis of artificial radionuclides monitoring, it is imperative not to ignore the assessment of models' performance. In binary classification tasks such as contamination surveillance, a confusion matrix is often employed to contrast the models' diagnosis with the ground truth, as illustrated in Fig. 2. This comparison enables the computation of specific performance metrics for models [20]. Generally, the primary metric is accuracy which encompasses two key facets of models: the proportion of true positives when artificial radionuclides contamination is present, and the likelihood of avoiding false positives in the absence of contamination. It can be concluded that the accuracy of models is fundamental and still inadequate. Moreover, it's well known that a model would be unsuitable for extensive applications if it refrains from false alarms in natural background situation but falters to sound true alarms promptly upon the occurrence of artificial contamination. Hence, the inclusion of precision and sensitivity becomes essential. Precision denotes the ratio of true positives among all positive alarms; sensitivity denotes the ratio of true positives correctly identified in scenarios of artificial contamination. Both metrics ought to be maximized. Furthermore, the unique property of continuous monitoring should be taken into accounts, specifically, its efficacy being dependent on the varying levels of background radiation.

Simultaneously, the False Positive Rate (FPR), calculated from the confusion matrix and defined as the ratio of false positives to the sum of false positives and true negatives ($FPR = FP/(TN)$), is used as the x-axis. The True Positive Rate (TPR), also known as recall, which represents the proportion of correct alarms when contamination occurs ($TPR = TP/(TP + FN)$), is plotted on the y-axis. These create the Receiver Operating Characteristic (ROC) curve, a widely recognized visual technique based on classification performance [21]. The area beneath the ROC curve, termed AUC (Area Under Curve), serves as a significant metric for model evaluation where a larger AUC value indicates a better model performance [22].
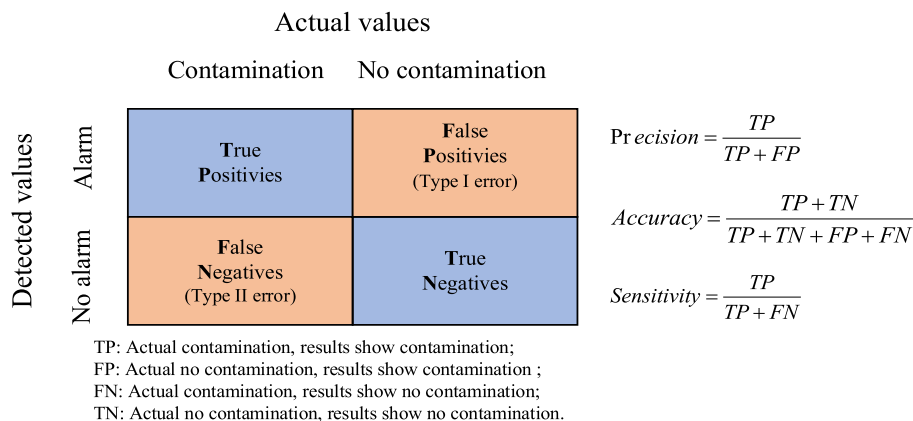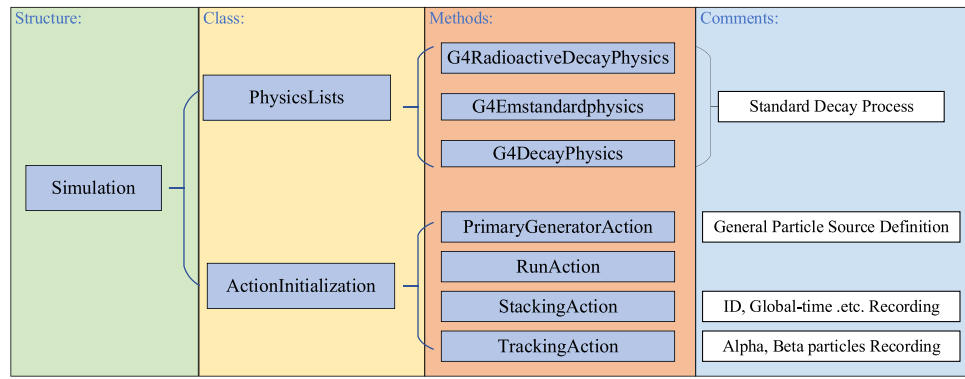
Experimental

Geant4 tool kits are widely regarded as an effective suite for simulating nuclear decay processes, particularly the alpha decay of naturally occurring radon and thorium, which has been highly acclaimed in literature [23, 24]. Its core functionalities include particle trajectory tracking and time recording. In this study, Geant4 software is employed to emulate the physical decay mechanisms of natural radon (^{222}Rn), thorium (^{224}Ra), and artificial alpha-emitting nuclei (^{242}Cm as a representative example). The simulation methodology adheres closely to that of previous investigations [9], with the code architecture illustrated in Fig. 3.

To assess the efficacy of the three models in monitoring natural background, the study simulates an atmospheric environment with a blend of radon and thorium in a 10:1 ratio. The experimental design is as follows: the activity of ^{222}Rn ranges from 0.5 Bq to 10 Bq increasing by 0.5 Bq interval; For ^{224}Ra, the activity spans from 0.05 Bq to 1 Bq with increments of 0.05 Bq. This generates a comprehensive set of 20 data groups, starting from a mix of 0.5 Bq ^{222}Rn with 0.05 Bq ^{224}Ra and ending at a blend of 10 Bq ^{222}Rn with 1 Bq ^{224}Ra. From each data group, a continuous sequence of 50 h of counting and particles' time is extracted as the measurement sample. These sample serve to compare the discrepancies between the estimated natural background levels measured by the three models and the actual total alpha counts, as well as to evaluate the impact of using a slide-window method on the results.

In addition, this study utilizes a 1.5 Bq ^{242}Cm to represent artificial alpha-emitting radionuclides and procures a continuous data set of 50 h as the validation samples. These samples are introduced into the background at the 10th, 20th, 30th, 40th, and 50th hours respectively to assess the accuracy, precision and sensitivity of the models in detecting

Fig. 2 A schematic representation of using a confusion matrix to evaluate model accuracy, precision and sensitivity

Actual values

	Contamination	No contamination
Alarm	True Positives	False Positives (Type I error)
No alarm	False Negatives (Type II error)	True Negatives

Detected values

$$Precision = \frac{TP}{TP + FP}$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$Sensitivity = \frac{TP}{TP + FN}$$

TP: Actual contamination, results show contamination;
FP: Actual no contamination, results show contamination ;
FN: Actual contamination, results show no contamination;
TN: Actual no contamination, results show no contamination.

Fig. 3 The architecture of Geant4 code

Table 1 Data format of one single ^{222}Rn mixed with one single ^{224}Ra

| Name	ID	From	Global time (µs)	Name	ID	From	Global time (µs)
Thread 0 (^{222}Rn)							
Alpha	3	1	2,214,075,205,419	^{222}Rn	1	0	0
Alpha	263	2	2,214,175,036,026	^{218}Po	2	1	2,214,075,205,419
e-	587	262	2,217,706,400,891	^{214}Pb	262	2	2,214,175,036,026
e-	592	588	2,219,265,212,014	^{214}Bi	588	262	2,217,706,400,891
Alpha	606	603	2,219,265,212,115	^{214}Po	603	588	2,219,265,212,014
e-	1079	605	4,292,963,692,692,529	^{210}Pb	605	603	2,219,265,212,115
e-	1082	1077	4,293,090,596,405,458	^{210}Bi	1077	605	4,292,963,692,692,529
Alpha	1088	1080	4,333,136,785,719,434	^{210}Po	1080	1077	4,293,090,596,405,458
				^{206}Pb	1087	1080	4,333,136,785,719,434
Thread 1 (^{224}Ra):							
Alpha	3	1	600,739,858,561	^{224}Ra	1	0	0
Alpha	280	2	600,746,692,087	^{220}Rn	2	1	600,739,858,561
Alpha	625	279	600,747,220,529	^{216}Po	279	2	600,746,692,087
e-	1021	624ara>	637,748,155,007	^{212}Pb	624	279	600,747,220,529
e-	1047	1022	641,698,006,405	^{212}Bi	1022	624	637,748,155,007
Alpha	1050	1045	641,698,006,405	^{212}Po	1045	1022	641,698,006,405
				^{208}Pb	1049	1045	641,698,006,405

artificial contamination against varying background radiation. The Geant4 simulation data format for the mixture of individual radon and thorium nuclei is depicted in Table 1. One can see that each particle has a "from" ID which means its parent nuclei ID, a global time indicating the its arriving time. Through these global time, we are able to calculate the total alpha counts and ^{214}Po counts during a sampling period, which also facilitate comparisons with outcomes derived from the three models.

Results and discussion

The estimation of natural background

Figure 4 depicts the scenario of the natural background monitoring using a Poisson distribution model. The upper surface of the figure illustrates the discrepancy between the
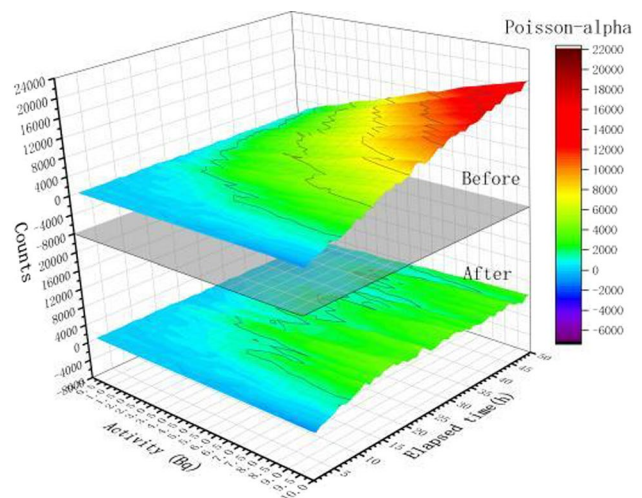
Fig. 4 The discrepancy between the model-estimated total alpha counts and the actual background alpha counts before and after using a slide-window for a Poisson model

model-estimated total alpha counts and the actual background alpha counts. It is evident that as time progresses, there is an increasing divergence between the estimated outcomes and the actual counts across various activities of the natural background. However, it is crucial to note that this does not imply an increase in estimation deviation. Instead, it is due to the accumulation of outdated information in the prior, which causes the estimates to become increasingly conservative and less responsive to the natural fluctuations of the background, resulting in a widening gap. More specifically, the initial monitoring employs a Jeffreys prior for the parameter θ, which is $\pi(\theta) = \theta^{-1/2}$ corresponding to $Gamma(0.5, 0)$. The posterior adheres to a Gamma distribution as stipulated by Eq. (5). The priors utilized in subsequent measurements are inherited from the posteriors of the preceding measurements, with parameter a accumulating the total alpha counts and parameter b tallying the number of measurements. As Eq. (16) shown in Appendix A, the influence of the prior mean on the posterior expectation progressively intensifies, thereby pushing the expected alpha counts further away from the true values.

The key to resolve this issue is to eliminate the outdated information from the prior, which can be accomplished by employing the slide-window method. For instance, a length of 10 slide-window is used to record the parameter a and b of the prior for a Poisson model parameter. When measurements reach the 11th hour (with a measurement taken every hour), the slide-window automatically discards the alpha counts from the first observation from a and parameter b becomes b-1. Then the slide-window adds the current measurement to the end of the queue, and renews the parameter a as $a +$ alpha counts from the current measurement and parameter b as $b + 1$. Specifically, the growth of parameter a is inhibited and the value of parameter b is maintained at 9, which leads a balance between the prior mean and the sample. In this way, the update of the prior becomes twofold: the regular update from the posterior and the elimination records from the parameter a and b. The lower surface of Fig. 4 exhibits the results after using the slide-window. One can see that the surface no longer grows steep over time and background activity for the second reasons, validating the use of the slide-window in background activity estimation.

The application of the Gaussian model for background estimation mirrors that of the Poisson model, yet it diverges in that the Gaussian model inherently comprises two parameters: the expectation and the variance. To met conjugate conditions of the prior, it is important to presume a known overall variance which equals to the variance deduced from a successive series of 50 measurements. The outcomes derived from the Gaussian model before and after using a slide-window are depicted in Fig. 5, wherein the upper surface represents the discrepancy between the estimated total alpha counts and the actual total alpha counts. Notably, this
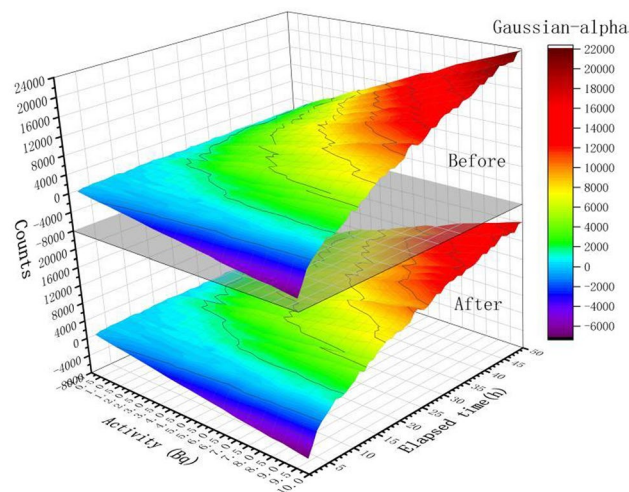
Fig. 5 The discrepancy between the model-estimated total alpha counts and the actual background alpha counts before and after using a slide-window for a Gaussian model

surface is considerably steeper compared with the Poisson model, signifying that the estimated discrepancies magnify with time and activity. This stems from the gradual augmentation of prior distribution's expectation μ_0, which escalates its impact on the posterior expectation as shown in the second form of Eq. (8).

The results of using a slide-window are manifested in the lower portion of Fig. 5. Although this model yields a more moderate surface, the efficacy is less pronounced compared to the Poisson model. The underlying cause of this distinction relates to the disparate methodologies by which the models accumulate prior information: whereas the prior distribution of the Gaussian model's expectation is accumulated from the arithmetic mean of total alpha counts of previous measurements, denoted as $\bar{x}$ in Eq. (8). And removing a outdated record from the arithmetic mean $\bar{x}$ has a negligible impact on the value of $\bar{x}$ itself.

The binomial distribution model diverges from the previously discussed models in its complexity for it requires for supplementary data. It is widely believed that, in the absence of artificial radionuclides contamination, substantial fluctuations in total alpha counts are accompanied with corresponding increase or decrease in the alpha counts of ^{214}Po. Therefore, the ratio of alpha counts of ^{214}Po to the total alpha counts serves as an indicator of the natural background level. The estimation of background using the binomial distribution involves a two-step process: Firstly, the alpha counts of ^{214}Po can be measured for the very short half-life of ^{214}Po (164 µs) through a coincidence system as described in the earlier work. This yields the ratio θ of alpha counts of ^{214}Po to the total alpha counts; Secondly, a Beta distribution with parameters a and b is constructed as the prior for θ through historical measurement data. The

parameters *a* and *b* represents the fractions of the total alpha counts contributed by ^{214}Po ("successful events") and non-^{214}Po ("failures events") respectively. When measuring the natural background using the binomial distribution, one can employ a uniform distribution $U(0,1)$ corresponding to a special case of $Beta(1,1)$ as the prior for θ. With the alpha counts from ^{214}Po, the total alpha counts can then be estimated. The results from the binomial distribution model are illustrated in the upper surface of Fig. 6, revealing the most uniform surface among the three models, closest to the true level of the natural background. Implementation of a slide-window simply requires the subtraction of outdated entries from *a* and *b* to refresh the Beta distribution, resulting in the estimates displayed in the lower half of Fig. 6, where the surface becomes more refined.

Application of artificial radionuclides monitoring

To assess the effectiveness of the three models in artificial radionuclides monitoring, this study conducts two
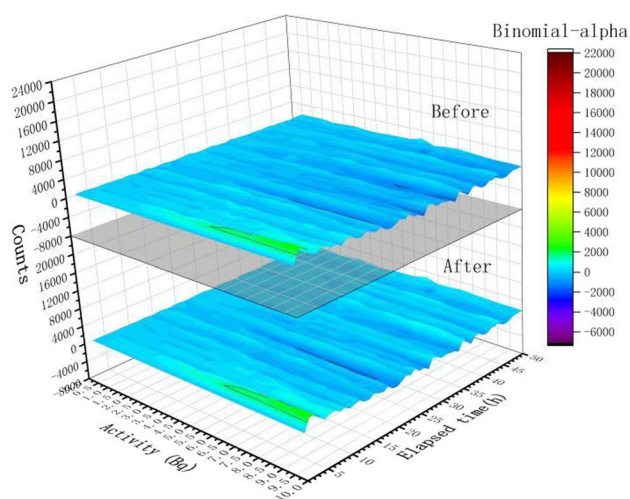
Fig. 6 The discrepancy between the model-estimated total alpha counts and the actual background alpha counts before and after using a slide-window for a Binomial model

Fig. 7 Results of artificial radionuclides monitoring for three models (**a** The slide-window method is not applied; **b** The slide-window method is applied)
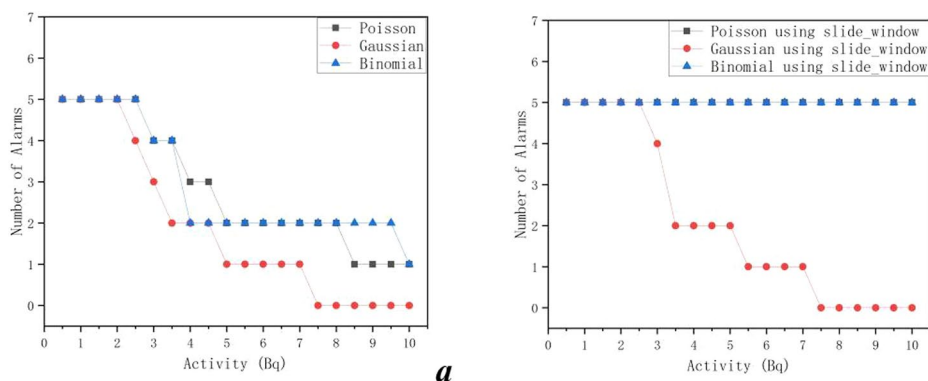
experiments. For experiment 1, five test point of a 1.5 Bq ^{242}Cm source are introduced to background range from 0.5 Bq ^{222}Rn mixed with 0.05 Bq ^{224}Ra to 10 Bq ^{222}Rn mixed with 1 Bq ^{224}Ra at 10th, 20th, 30th, 40th, 50th h in consecutive observations. The results of artificial radionuclides monitoring using the three models are shown in Fig. 7a. It is evident that with lower levels of natural background radiation (Below 3 Bq of ^{222}Rn mixed with 0.3 Bq ^{224}Ra), all three models successfully identify the five points of contamination over the 50 h period. However, as the natural background radiation increases, the ability to discern the artificial radionuclides' alpha counts becomes progressively hindered by the elevated background, leading to a diminished detection performance across all models. Notably, when the background radiation surpasses 7 Bq of ^{222}Rn and 0.7 Bq of ^{224}Ra, the Gaussian model fails to detect any contamination, whereas the binomial and the Poisson distribution models continue to exhibit similar effectiveness of detection.

Experiment 2 is conducted under identical background conditions and contamination sources, comparing the results of the three models after using a slide-window method improvement. The effectiveness of the enhanced triad of models is depicted in Fig. 7b: an enhancement is evident across all three models. Notably, the Poisson and binomial distribution models exhibit the most significant enhancements, with the detection rates for artificial radionuclides attaining almost 100%, while the Gaussian model experiences a modest increase in detection efficiency. This demonstrates that refreshing the model priors using the slide-window method is validated. The less notable improvement in the Gaussian model primarily stems from the fact that historical information in the the prior is stored as a mean value, and influenced by the overall variance. When outdated information is removed, it has a minimal impact on the mean value of the prior.

Model performance analysis

By constructing confusion matrices across varying background levels, we obtain three metrics for the three models

both before and after utilizing a slide-window, as depicted in Fig. 8. The left part of the figure provides a comparison of model performance under normal conditions, while the right part shows the results after applying the slide-window method. Figure 8a, b respectively present the accuracy comparisons for the models. It can be seen that at low background levels (activities less than 3 Bq ^{222}Rn mixed with 0.3 Bq ^{226}Ra), all models exhibit exceptionally high accuracy, effectively identifying points of artificial radionuclides contamination and without triggering false alarms in the absence of such contamination. However, as background level rise, the alpha particles emitted by artificial radionuclides become increasingly obscured by those of background radiation, leading to a decline in the number of contamination detection. Nevertheless, due to the low false alarm rates maintained by the three models, their accuracy remains consistently above 90%. Following the update of model priors with a slide-window, the accuracy of both the Poisson and Binomial models significantly improves, primarily due to the increase of the number of identified artificial contamination as described in the previous section, resulting in an enhancement of the true positives in the confusion matrix. The Gaussian model shows a less significant improvement in accuracy compared to the other two.

The precision metric is commonly believed to reflect the balance between a model's correct and false alarms. A precision below 50% suggests that for every two alarms, there is likely one false alarm, indicating its practical importance in artificial radionuclides monitoring. Figure 8c, d illustrate the differences in precision for the three models before and after using the slide-window method. The precision trends for the three models are similar to the accuracy trends, initially high but subsequently declining. It is important to note that the Gaussian model becomes unquantifiable beyond a background level of 7 Bq because both the true positives and false positives of the confusion matrix are zero. Specifically, the Gaussian model does not issue any false alarm either beyond the background level of 7 Bq. The Poisson model also exhibits deficient performance with a swift decline in precision beyond a 7 Bq background. The decline is primarily due to the combination of a decrease in the true positives and an increase in the false positives. The former derives from the failures of identifying contamination, and the later derives from the emergence of false alarms. The root of false alarms can be resorted to the posterior estimation being influenced by outdated prior. After the adoption of the slide-window method, the three models show a marked rise in identifying contamination, which considerably boost models' precision. Particularly, the Poisson distribution shows a 40% improvement once the natural background attains 7 Bq. In contrast, the Binomial distribution model consistently maintains a precision close to 100%, indicating it virtually never issues false alarms.

The sensitivity metric is a crucial indicator of the models' capability to identify artificial radionuclides contamination. Figure 8e, f respectively depict the differences in sensitivity for the three models before and after employing a slide-window method. It is not surprised to see that sensitivity decreases as the natural background level increases, for alpha particles emitted from artificial radionuclides are more difficult to be discovered as those from background increase. However, through using the slide-window method, the prior information in both the Binomial and Poisson models is updated more rapidly, Estimations of the models are

Fig. 8 Comparison of performances among the three models (**a** the accuracy of models before modification; b: the accuracy of models after modification; **c** the precision of models before modification; d: the precision of models after modification; **e** the sensitivity of models before modification; f: the sensitivity of models after modification)
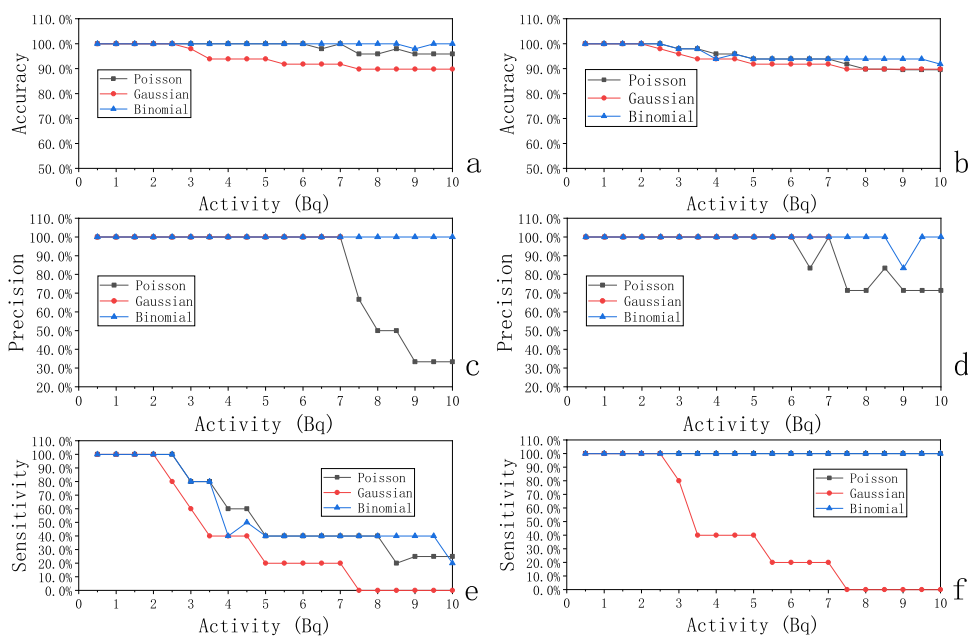
more accurate and thereby enhancing sensitivity. The sensitivity of both the Poisson and Binomial models has been amplified by a factor of 4–5. But for the Gaussian model, the improvement in sensitivity is not significant because the prior uses the mean of historical information.

When all three indicators are considered, we find that the Poisson distribution model performs similarly to the Binomial distribution model. A more intuitive method performs is to calculate the FPR as the a-axis and TPR as the y-axis for the three models, resulting in binary pair (FPR, TPR). At the same time, we add two points, (0, 0) and (1, 1), to create an ROC curve, with the area under the curve representing the AUC value, as shown in Fig. 9. The left image shows the AUC areas for the three models without using the slide-window method. It can be seen that the performance of the Binomial distribution model is slightly higher than the Poisson distribution, both of which significantly outperform the Gaussian model; On the other hand, after improving with the slide-window method, the Binomial and Poisson distribution models almost achieve "perfection" while the Gaussian distribution model also improves about 5%.

Conclusions

This paper provides a detailed description of the specific applications of three common Bayesian models in the field of continuous online monitoring of artificial radionuclides, including the prior selection and posterior calculation of model parameters. Using radon, thorium series decay data and artificial radionuclides decay data simulated with the Geant4 tool kits, the differences between the three models in estimating background are verified. We find that although Bayesian models between the three can accumulate historical experience during the prior updating process, overly outdated information would reduce the accuracy the posterior expectation estimates. Therefore, this paper proposes a slide-window method to quantitatively update the prior distribution of model parameters, and describes the specific process of updating the prior distribution of models' parameters using this method. Finally, this paper compares the accuracy, precision, and sensitivity of the three models in the continuous online monitoring of artificial radionuclides before and after using this method. The analysis results show that: (1) among the three models, the Gaussian model performs moderately, while the Binomial and Poisson distribution models perform comparably; (2) the slide-window method can effectively improve the speed of updating the prior of model parameters, so as to improve the accuracy, precision and sensitivity of models.

However, the implementation of the Binomial distribution model is not only dependent on the total alpha counts but also requires alpha counts from ^{214}Po. The complexity necessitates further research into the measurement of ^{214}Po based on the Bayesian approach, given that the half life of ^{214}Po is 164 μs, where the measurement is easily influenced by the ^{216}Po (164 ms) and ^{212}Po (0.3 μs). Ultimately, the Bayesian models discussed here are not merely applicable to the online monitoring of artificial radionuclides, but equally relevant to other measurement systems that experience background interference such as dose monitoring system.

Availability of date and materials

All data generated or analyzed during this study are included in this published article [and its supplementary information files].
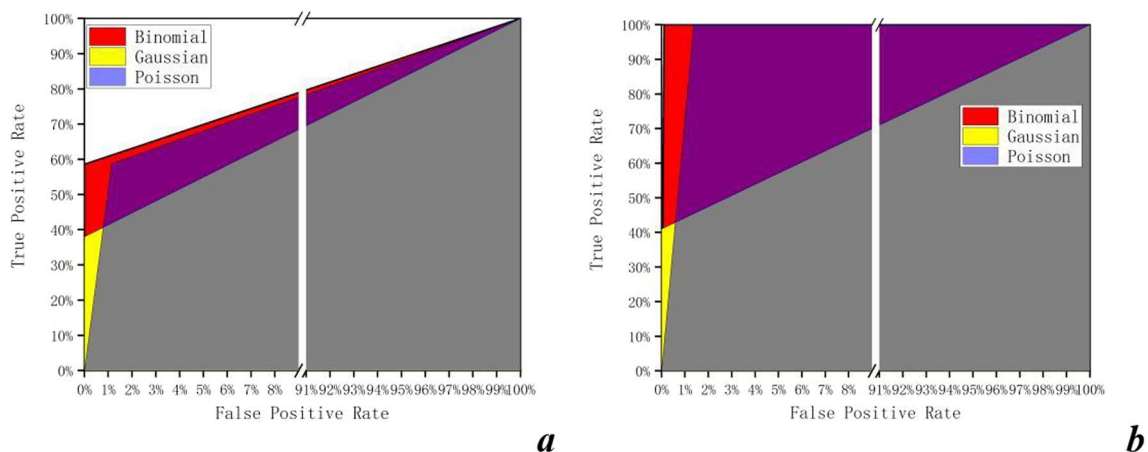
Fig. 9 The AUC area for the three models (**a** The slide-window method is not applied; **b** The slide-window method is applied)

Appendix A

Let $X = (X_1, \ldots, X_n)$ be a random sample from a Poisson distribution $P(\theta)$. The probability density function of the sample X is given by:

$$f(x|\theta) = P(X_1 = x_1, \ldots, X_n = x_n|\theta)$$
$$= \prod_{i=1}^{n} \frac{\theta^{x_i} e^{-\theta}}{x_i!} = \frac{\theta^{n\bar{x}} e^{-n\theta}}{x_1! \ldots x_n!} \propto \theta^{n\bar{x}} \exp\{-n\theta\} \quad (12)$$

where $\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$, The above expression is the likelihood function of θ. Let the prior distribution of θ be a Gamma distribution $\Gamma(a,b)$, with its density function given by

$$\pi(\theta) = \text{Gamma(a,b)} = \frac{b^a}{\Gamma(a)} \theta^{a-1} e^{-b\theta} \propto \theta^{a-1} \exp\{-b\theta\} \quad (13)$$

In the above equation, both a and b are known hyper arguments. Combining Eqs. (12) and (13), we have:

$$\pi(\theta|x) \propto f(x|\theta) \cdot \pi(\theta) \propto \theta^{n\bar{x}+a-1} \exp\{-(n+b)\theta\} \quad (14)$$

The above expression is the kernel of $\Gamma(n\bar{x}+a, n+b)$. After adding a normalizing constant factor, we obtain:

$$\pi(\theta|x) = \frac{(n+b)^{n\bar{x}+a}}{\Gamma(n\bar{x}+a)} \theta^{n\bar{x}+a-1} \exp\{-(n+b)\theta\} \quad (15)$$

Which indicates that the posterior distribution of θ is a Gamma distribution $\Gamma(n\bar{x}+a, n+b)$. In the continuous monitoring of artificial radionuclides, each measurement yields a result, thus $n=1$, $\bar{x} = x$. The posterior expectation is:

$$E(\theta|x) = \overline{\theta_E} = (x+a)/(b+1)$$
$$= \frac{1}{b+1} x + \left(1 - \frac{1}{b+1}\right) \frac{a}{b} \quad (16)$$
$$= \gamma x + (1 - \gamma) \frac{a}{b}$$

In the above equation, a/b is the prior mean value, and x is the sample value. As the number of measurements increase, the weight of the prior mean increase, while the weight of the sample decreases.

Appendix B

Let $X = (X_1, \ldots, X_n)$ be a random sample drawn from a normal distribution $N(\theta, \sigma^2)$. The sample mean $\overline{X}$ is a sufficient statistic for θ and it follows an another normal distribution $N(\theta, \sigma^2/n)$. Therefore, the likelihood function of θ is given by:

$$l(\theta|\bar{x}) = \frac{\sqrt{n}}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{n(\bar{x} - \theta)^2}{2\sigma^2}\right\} \propto \exp\left\{-\frac{n(\bar{x} - \theta)^2}{2\sigma^2}\right\} \quad (17)$$

Let the conjugate prior distribution of θ be a normal distribution $N(\mu_0, \sigma_0^2)$. The likelihood function of θ is given by:

$$\pi(\theta) = \frac{1}{\sqrt{2\pi}\mu_0} \exp\left\{-\frac{(\theta - \mu_0)^2}{2\sigma_0^2}\right\} \propto \exp\left\{-\frac{(\theta - \mu_0)^2}{2\sigma_0^2}\right\} \quad (18)$$

Note $\sigma_n^2 = \sigma^2/n$, the posterior distribution density of θ is:

$$\pi(\theta|\bar{x}) \propto l(\theta|\bar{x})\pi(\theta) \propto \exp\left\{-\frac{1}{2}\left[\frac{(\bar{x} - \theta)^2}{\sigma_n^2} + \frac{(\theta - \mu_0)^2}{\sigma_0^2}\right]\right\}$$
$$\propto \exp\left\{-\frac{1}{2}[A\theta^2 - 2\theta B + C]\right\} \quad (19)$$
$$= \exp\left\{-\frac{A}{2}\left(\theta - \frac{B}{A}\right)^2 - \frac{1}{2}\left(C - \frac{B^2}{A}\right)\right\}$$

where $A = \frac{1}{\sigma_n^2} + \frac{1}{\sigma_0^2}, B = \frac{\bar{x}}{\sigma_n^2} + \frac{\mu_0}{\sigma_0^2}, C = \frac{\bar{x}^2}{\sigma_n^2} + \frac{\mu_0^2}{\sigma_0^2}, C - \frac{B^2}{A} = \frac{(\bar{x} - \mu_0)^2}{\sigma_0^2 + \sigma_n^2}$, which can be further simplified as:

$$\pi(\theta|\bar{x}) \propto \exp\left\{-\frac{A}{2}\left(\theta - \frac{B}{A}\right)^2 - \frac{1}{2}\left(C - \frac{B^2}{A}\right)\right\}$$
$$= \left(\frac{2\pi}{A}\right)^{-1/2} \exp\left\{-\frac{(\theta - B/A)^2}{2/A}\right\} \quad (20)$$

From the above equation, the posterior of θ follows a normal distribution $N(B/A, 1/A)$, with its posterior expectation α and variance β respectively given by:

$$\alpha = B/A = \left(\frac{\mu_0}{\sigma_0^2} + \frac{n}{\sigma^2}\bar{x}\right) \bigg/ \left(\frac{n}{\sigma^2} + \frac{1}{\sigma_0^2}\right) \quad (21)$$

$$\beta = \frac{1}{A} = \left(\frac{n}{\sigma^2} + \frac{1}{\sigma_0^2}\right)^{-1} \quad (22)$$

In the continuous monitoring of artificial radionuclides, each measurement yields a result, thus $n=1$, $\sigma_n^2 = \sigma^2$.

Appendix C

Let X be a random sample drawn from a Binomial distribution $b(n,\theta)$. The probability density function of the sample X is given by:

$$f(x|\theta) = \binom{n}{x} \theta^x (1 - \theta)^{n-x} \quad (23)$$

where the kernel is $\theta^x(1 - \theta)^{n-x}$. Let the prior distribution of θ be a Beta distribution Beta(α, β), the kernel of which is also $\theta^x(1 - \theta)^{n-x}$, the posterior distribution of θ is:

$$\pi(\theta|x) = f(x|\theta)\pi(\theta) \propto \theta^x(1-\theta)^{n-x}\theta^{a-1}(1-\theta)^{b-1}$$
$$= \theta^{a+x-1}(1-\theta)^{n+b-x-1} = \mathrm{Beta}(a+x, n+b-x) \quad (24)$$

The mathematical forms for the posterior expectation and the posterior variance of θ are as follows:

$$\mathrm{E}(\theta|x) = \hat{\theta}_E = \frac{a+x}{a+b+n} = \frac{n}{a+b+n} \cdot \frac{x}{n} + \frac{a+b}{a+b+n} \cdot \frac{a}{a+b}$$
$$= \gamma \cdot \frac{x}{n} + (1-\gamma) \cdot \frac{a}{a+b} \quad (25)$$

$$\mathrm{Var}(\theta|x) = \sigma_\theta^2 = \frac{(a+x)(b+n-x)}{(a+b+n)^2(a+b+n+1)} = \frac{E(\theta|x)[1-E(\theta|x)]}{a+b+n+1}. \quad (26)$$

Authors' contributions All authors contributed to the study conception and design. Experiments, data collection and analysis were performed by XL, QH, YX, and XG. The first draft of the manuscript was written by XL and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript. XL: conceptualization, methodology, software. QH: writing—review and editing, validation. YX: data curation, visualization. XG: supervision, project administration.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Consent for publication Not applicable.

References

1. Guide to the expression of uncertainty in measurement—JCGM 100:2008 (GUM 1995 with minor corrections—Evaluation of measurement data, the International Organization for Standardization. https://www.bipm.org/documents/20126/2071204/JCGM_100_2008_E.pdf. Accessed 14 May 2024
2. Bochud FO, Bailat CJ, Laedermann JP (2007) Bayesian statistics in radionuclide metrology: measurement of a decaying source. Metrologia 44:S95–S101. https://doi.org/10.1088/0026-1394/44/4/S13
3. Kacker RN (2006) Bayesian alternative to the ISO-GUM's use of the Welch–Satterthwaite formula. Metrologia 43:1–11. https://doi.org/10.1088/0026-1394/43/1/001
4. Lira I, Wöger W (2006) Comparison between the conventional and Bayesian approaches to evaluate measurement data. Metrologia 43:S249–S259. https://doi.org/10.1088/0026-1394/43/4/S12
5. Kacker R, Jones A (2003) On use of Bayesian statistics to make the guide to the expression of uncertainty in measurement consistent. Metrologia 40:235–248. https://doi.org/10.1088/0026-1394/40/5/305
6. Weise K, Hübel K, Rose E, Schläger M, Schrammel D, Täschner M, Michel R (2006) Bayesian decision threshold, detection limit and confidence limits in ionising-radiation measurement. Radiat Prot Dosim 121:52–63. https://doi.org/10.1093/rpd/ncl095
7. Laedermann J-P, Valley J-F, Bochud FO (2005) Measurement of radioactive samples: application of the Bayesian statistical decision theory. Metrologia 42:442–448. https://doi.org/10.1088/0026-1394/42/5/015
8. Trassinelli M (2020) An introduction to Bayesian statistics for atomic physicists. In: Journal of Physics: Conference Series 1412. https://doi.org/10.1088/1742-6596/1412/6/062008
9. Li X, Yan YJ, Gong XY (2024) Enhancing artificial radionuclides monitoring: a Bayesian statistical approach combined with the multi time-interval analysis method. J Radioanal Nucl Chem 333:2121–2130. https://doi.org/10.1007/s10967-024-09394-w
10. Zabulonov Y, Burtniak V, Krasnoholovets V (2016) A method of rapid testing of radioactivity of different materials. J Radiat Res Appl Sci 9:370–375. https://doi.org/10.1016/j.jrras.2016.03.001
11. Pyke CK, Hiller PJ, Koma Y, Ohki K (2022) Radioactive waste sampling for characterisation—a Bayesian upgrade. Nucl Eng Technol 54:414–422. https://doi.org/10.1016/j.net.2021.07.042
12. Qingpei X, Dongfeng T, Jianyu Z, Fanhua H, Ge D, Jun Z (2013) Numerical study on the sequential Bayesian approach for radioactive materials detection. Nucl Instrum Methods Phys Res Sect A 697:107–113. https://doi.org/10.1016/j.nima.2012.09.031
13. Luo P, Sharp JL, Devol TA (2013) Bayesian analyses of time-interval data for environmental radiation monitoring. Health Phys 104:15–25. https://doi.org/10.1097/HP.0b013e318260d5f8
14. Nosek D, Nosková J (2016) On Bayesian analysis of on–off measurements. Nucl Instrum Methods Phys Res Sect A 820:23–33. https://doi.org/10.1016/j.nima.2016.02.094
15. Arias-Nicolás JP, Ruggeri F, Suárez-Llorens A (2016) New classes of priors based on stochastic orders and distortion functions. Bayesian Anal. https://doi.org/10.1214/15-ba984
16. Barrera M, Lira I, Sánchez-Sánchez M, Suárez-Llorens A (2019) Bayesian treatment of results from radioanalytical measurements. Effect of prior information modification in the final value of the activity. Radiat Phys Chem 156:266–271. https://doi.org/10.1016/j.radphyschem.2018.11.023
17. Lira I, Grientschnig D (2010) Bayesian assessment of uncertainty in metrology: a tutorial. Metrologia 47:R1–R14. https://doi.org/10.1088/0026-1394/47/3/R01
18. Petraglia A, Sirignano C, Buompane R, D'Onofrio A, Esposito AM, Terrasi F, Sabbarese C (2020) Space-time Bayesian analysis of the environmental impact of a dismissing nuclear power plant. J Environ Radioact 218. ARTN 106241. https://doi.org/10.1016/j.jenvrad.2020.106241
19. Sanada Y, Tanabe Y, Iijima N, Momose T (2011) Development of new analytical method based on beta-alpha coincidence method for selective measurement of Bi–Po-application to dust filter used in radiation management. Radiat Prot Dosim 146:80–83. https://doi.org/10.1093/rpd/ncr116
20. Fawcett T (2006) An introduction to ROC analysis. Pattern Recogn Lett 27:861–874. https://doi.org/10.1016/j.patrec.2005.10.010
21. Valero-Carreras D, Alcaraz J, Landete M (2023) Comparing two SVM models through different metrics based on the confusion matrix. Comput Oper Res 152. ARTN 106131. https://doi.org/10.1016/j.cor.2022.106131

22. Phillips G, Teixeira H, Kelly MG, Salas Herrero F, Várbíró G, Lyche Solheim A, Kolada A, Free G, Poikane S (2024) Setting nutrient boundaries to protect aquatic communities: the importance of comparing observed and predicted classifications using measures derived from a confusion matrix. Sci Total Environ. https://doi.org/10.1016/j.scitotenv.2023.168872

23. Allison J, Amako K, Apostolakis J, Araujo H, Dubois PA, Asai M, Barrand G, Capra R, Chauvie S, Chytracek R, Cirrone GAP, Cooperman G, Cosmo G, Cuttone G, Daquino GG, Donszelmann M, Dressel M, Folger G, Foppiano F, Generowicz J, Grichine V, Guatelli S, Gumplinger P, Heikkinen A, Hrivnacova I, Howard A, Incerti S, Ivanchenko V, Johnson T, Jones F, Koi T, Kokoulin R, Kossov M, Kurashige H, Lara V, Larsson S, Lei F, Link O, Longo F, Maire M, Mantero A, Mascialino B, McLaren I, Lorenzo PM, Minamimoto K, Murakami K, Nieminen P, Pandola L, Parlati S, Peralta L, Perl J, Pfeiffer A, Pia MG, Ribon A, Rodrigues P, Russo G, Sadilov S, Santin G, Sasaki T, Smith D, Starkov N, Tanaka S, Tcherniaev E, Tomé B, Trindade A, Truscott P, Urban L, Verderi M, Walkden A, Wellisch JP, Williams DC, Wright D, Yoshida H (2006) Geant4 developments and applications. IEEE Trans Nucl Sci 53:270–278. https://doi.org/10.1109/Tns.2006.869826

24. Khan AU, DeWerd LA (2021) Evaluation of the GEANT4 transport algorithm and radioactive decay data for alpha particle dosimetry. Appl Radiat Isotopes 176:109849. ARTN 109849 https://doi.org/10.1016/j.apradiso.2021.109849