

# Sequential Probability Ratio Test (SPRT) for Dynamic Radiation Level Determination – Preliminary Assessment

Ding Yuan and Warnick Kernan

Remote Sensing Laboratory-Nellis  
P. O. Box 98521, M/S RSL-11, Las Vegas, NV 89193-8521  
{yuand, kernanwj}@nv.doe.gov

**Abstract.** A Sequential Probability Ratio Test (SPRT) algorithm for reliable and fast determination of a relative radiation level in a field environment has been developed. The background and the radioactive anomaly are assumed to follow the normal and Poisson distributions, respectively. The SPRT formulation has been derived and simplified based on these assumptions. The preliminary evaluation suggests that the algorithm, while offering confident estimations for the log-scaled radiation level, promises the additional advantage of reduction in sampling sizes, particularly in areas with a high radiation level.

## 1 Introduction

Reliable and fast estimation of the local relative radiation level with respect to that of the regional natural background, under prescribed confidence levels, is the central focus of many environmental radiological surveys. A local relative radiation level can be conventionally interpreted as a number of standard deviations away from the background radiation mean. The key issues are reliability and speed. In order to obtain a reliable estimate for a local relative radiation level, the conventional statistical wisdom teaches us that we need more samples. However, a high radiation anomaly may pose a significant health threat to the field surveyors. Conventional health wisdom tells us to leave the high radiation area as soon as possible. Therefore, we need a radiation level estimation algorithm that is both reliable for given confidence levels and fast for high radiation areas.

In our recent research, we tailored Wald's Sequential Probability Ratio Test (SPRT) [1] and developed an algorithm for the dynamic determination of the local relative radiation level. A preliminary experiment in the laboratory environment, using a common industrial radiation source, suggests that the algorithm is promising. It provided a confident estimate for the local radiation level, and reduced the sampling size requirement for high radiation spots.

## 2 Population Assumptions

Reconnaissance environmental radiation surveys are commonly performed by hand-held gamma-detection devices, such as Geiger counters. These gamma-detection devices produce readings of gross gamma ray counts, or the gross count (GC) detected during a given time interval [2].

For a given high radiation anomaly, GC may be assumed to follow a Poisson Distribution [2] with PDF:

$$f_{\mu_1}(x) = P(X = x) = \frac{\mu_1^x e^{-\mu_1}}{x!}, \quad x = 0, 1, 2, \dots, n \tag{1}$$

Where  $\mu_1 = np$ .  $n$  is a large number (number of atoms in the observed physical sample) and  $p$  is the decay rate for an individual particle for a given observation time interval. When  $n$  is large, it can be shown that:

$$f_{\mu_1}(x) \rightarrow N(\mu_1, \mu_1) \tag{2}$$

Where  $\mu_1$  is the mean GC of the physical sample.

For regional radiation background, we assume that GC follows a simple Normal distribution  $N(\mu_0, \sigma_0^2)$  with probability distribution function (PDF):

$$f_{\mu_0, \sigma_0}(x) = \frac{1}{\sqrt{2\pi} \sigma_0} e^{-\frac{(x-\mu_0)^2}{2\sigma_0^2}}, \quad -\infty < x < \infty \tag{3}$$

This different treatment is due to uncertainty about the number of counts in the background sample, and if it is large enough for Poisson-Normal approximation that offers only one parameter for distribution fitting. Normal distribution, on the other hand, has two parameters to choose and should always make a better fit to the background samples.

### 3 Absolute and Discrete Radiation Levels

For measuring the absolute radiation level, the sample mean  $\mu_1$  is a natural choice. In a field survey, we are usually more interested in a discrete measure relative to the background radiation. For this reason, we shall introduce Log Departing Coefficient (LGC) as follows. Suppose that  $b (> 0)$  is a pre-selected log base, and  $\rho (> 0)$  is a pre-selected scaling factor. Define (for  $l \geq 1$ ):

$$LGC = l, \text{ if } \mu_0 + b^{l-1} \rho \sigma_0 < \mu_1 \leq \mu_0 + b^l \rho \sigma_0 \tag{4}$$

For  $LGC \geq 1$ , we have

$$LGC = \text{Ceiling} \left[ \text{Log}_b \left[ \frac{\mu_1 - \mu_0}{\rho \sigma_0} \right] \right] \tag{5}$$

Particularly, for a region with low radiation background, we may take  $b=2$  and  $\rho=1$ , then:

$$LGC = \text{Ceiling} \left[ \text{Log}_2 \left[ \frac{\mu_1 - \mu_0}{\sigma_0} \right] \right] \tag{6}$$

When LGC=0, we can say the sample mean is no more than two standard deviations from the regional mean. When LGC=1, we say that the sample mean is at least two standard deviations from the regional mean, but no more than four standard deviations from the regional mean, etc. For a region with higher radiation background, we may adjust factor  $\rho$ , making  $\rho=0.5$  for compensating the growing speed of LGC. LGC defined this way can be used as a log-scaled discrete measure for the radiation departing level (L) from the regional background.

### 4 Alternative Hypotheses

Let's assume that we collect observations one at a time, using a handheld device such as a Geiger Counter, in an environmental reconnaissance radiological survey situation. We denote  $x_i$  the  $i$ 'th observation, where the observation is the Gamma Gross Counts (GC) for fixed-length time intervals. Essentially for a sample, we have two possibilities:

- (1) That the sample was from the background radiation population – the Null Hypothesis ( $H_0$ ); and
- (2) That the sample was from an anomaly population with higher radiation strength – the Alternative Hypothesis ( $H_1$ ).

If we accept the null hypothesis  $H_0$ , then we continue our field survey. If we accept the alternative hypothesis  $H_1$ , then we mark the area as an anomaly for future detailed work and move on to the next spot.

Let  $\alpha$  and  $\beta$  be the type I (false positive) and type II (false negative) errors associated with the decisions respectively,

$$\alpha = P \{ \text{Accept } H_1 \text{ when } H_0 \text{ is true} \} = P(H_1 | H_0) \tag{7}$$

$$\beta = P \{ \text{Accept } H_0 \text{ when } H_1 \text{ is true} \} = P(H_0 | H_1) \tag{8}$$

then  $1-\alpha$  is the confidence level of accepting  $H_0$  and  $1-\beta$  is the confidence level of accepting  $H_1$ .

### 5 SPRT Basis

Wald's SPRT method [1] has the advantage of handling sequential sampling data. Let  $\{x_1, x_2, x_3, \dots, x_n\}$  be an (independently identically distributed) fresh data set collected since the last decision and a new decision is yet to be made. Assuming  $f_0(x)$  and  $f_1(x)$  are the PDFs for  $H_0$  and  $H_1$  respectively, we can construct a conventional Logarithmic Likelihood Ratio (LLR):

$$\Lambda_n = \text{LLR}(x_1, x_2, \dots, x_n) = \text{Log} \frac{\prod_{i=1}^n f_1(x_i)}{\prod_{i=1}^n f_0(x_i)} = \sum_{i=1}^n \text{Log} \frac{f_1(x_i)}{f_0(x_i)} \tag{9}$$

The Wald's SPRT has the following form (although his original A and B are somehow reversed from what are commonly used today):

- 1) If  $\Lambda_n < A$  then accept  $H_0$ , (and directly move on to next spot).
- 2) If  $A \leq \Lambda_n \leq B$ , request additional measurement made for the same spot.
- 3) If  $\Lambda_n > B$  then accept  $H_1$ , (and mark anomaly etc., move on to the next spot).

Where A and B are two constants satisfying inequalities:

$$\text{Log } \frac{\beta}{1 - \alpha} \leq A < B \leq \text{Log } \frac{1 - \beta}{\alpha} \tag{10}$$

A common practice is simply to take:

$$A = \text{Log } \frac{\beta}{1 - \alpha} \tag{11}$$

$$B = \text{Log } \frac{1 - \beta}{\alpha} \tag{12}$$

It is clear that the parameters A and B are related only to the strengths ( $\alpha, \beta$ ), or confidences ( $1 - \alpha, 1 - \beta$ ), of the test, but are independent of the actual forms of the distribution functions.

## 6 SPRT Formulation

If using the PDFs discussed early for  $H_0$  and  $H_1$  respectively (using normal approximation for Poisson distribution), then

$$\text{Log } \frac{f_1(x)}{f_0(x)} = \text{Log } \frac{\sigma_0}{\sqrt{\mu_1}} + \frac{(x - \mu_0)^2}{2\sigma_0^2} - \frac{(x - \mu_1)^2}{2\mu_1} \tag{13}$$

Substitute these PDFs in the Log-Likelihood Ratio, we have:

$$\begin{aligned} \Lambda_n &= \text{LLR}(x_1, x_2, \dots, x_n) \\ &= \sum_{i=1}^n \left( \text{Log } \frac{\sigma_0}{\sqrt{\mu_1}} + \frac{(x_i - \mu_0)^2}{2\sigma_0^2} - \frac{(x_i - \mu_1)^2}{2\mu_1} \right) \end{aligned} \tag{14}$$

Denote sample mean and sum of squares respectively:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \tag{15}$$

$$S = \sum_{i=1}^n (x_i - \bar{x})^2 \tag{16}$$

Through some simplification, and using sample mean  $\bar{x}$  for estimating  $\mu_1$ , we should finally have:

$$\Lambda_n \triangleq V_n^{(0)} - V_n^{(1)} \tag{17}$$

where

$$V_n^{(0)} = n \text{Log } \sigma_0 + \frac{S + n(\bar{x} - \mu_0)^2}{2 \sigma_0^2} \tag{18}$$

$$V_n^{(1)} = \frac{n}{2} \text{Log } \bar{x} + \frac{S}{2 \bar{x}} \tag{19}$$

### 7 Computational Details

We assume that the regional background population mean  $\mu_0$  and the standard deviations  $\sigma_0$  have been estimated, and also assume that the test strengths  $(\alpha, \beta)$ , or confidences  $(1-\alpha, 1-\beta)$  are given. Subsequently constants A and B can be calculated using formula (11) and (12). For a new spot or physical sample, we take first  $i_{\min}$  independent measures, then compute:

$$x^{(n)} = \frac{1}{n} \sum_{i=1}^n x_i \tag{20}$$

$$\mu_1 = x^{(n)} \tag{21}$$

$$S^{(n)} = \sum_{i=1}^n (x_i - x^{(n)})^2 \tag{22}$$

$$V_n^{(1)} = \frac{n}{2} \text{Log } \mu_1 + \frac{S^{(n)}}{2 \mu_1} \tag{23}$$

$$V_n^{(0)} = n \text{Log } \sigma_0 + \frac{S^{(n)} + n(x^{(n)} - \mu_0)^2}{2 \sigma_0^2} \tag{24}$$

$$\Lambda_n = V_n^{(0)} - V_n^{(1)} \tag{25}$$

The decision and estimation for the local radiation level L are made as follows:

- 1) If  $\Lambda_n < A$  then accept  $H_0$  and assign  $L = 0$ , output both  $\mu_1$  and  $L$  ;
- 2) If  $A \leq \Lambda_n \leq B$ , request additional measurement made for the same spot.
- 3) If  $\Lambda_n > B$  then accept  $H_1$  and compute:

$$L = \text{LGC} = \text{Ceiling} \left[ \text{Log}_b \left[ \frac{\mu_1 - \mu_0}{\rho \sigma_0} \right] \right] \tag{26}$$

Now output both  $\mu_1$  and  $L$ , and move to the next physical sample or spot.

Although SPRT sampling sequence terminates with probability 1, it does require a stop limit. For case 2, we adapted Wald' s stop limit[1]: if  $n = i_{\max}$  and we still have  $A \leq \Lambda_n \leq B$ , then:

$$\text{if } \Lambda_{i_{\max}} \leq \frac{A+B}{2}, \text{ we accept } H_0 \text{ and assign } L = 0; \tag{27}$$

$$\text{if } \Lambda_{i_{\max}} > \frac{A+B}{2}, \text{ we accept } H_1 \text{ and compute } L = LGC \tag{28}$$

Apparently, all local radiation level Ls computed this way had confidence levels  $(1-\alpha, 1-\beta)$ .

This algorithm was implemented in Mathematica[3] for quick evaluation. In the implementation, we set  $i_{\min} = 5$ ,  $i_{\max} = 11$ ,  $\alpha=0.05$  and  $\beta=0.05$ .

### 8 Experimental Data

Experimental data were collected in a laboratory environment using a typical industrial Cs137 source. A device similar to a Geiger counter was used for measuring gamma GC. The counting interval of the device was set at half of a second. Figure 1 shows the half-second background gamma GC data measure of the lab ground. Background radiation has a mean  $\mu_0 = 773$  and standard deviation  $\sigma_0 = 33$ . (This also suggests  $33 = \sigma_0 \neq \sqrt{\mu_0} = 28$  for background radiation, and therefore anomaly source and background should be handled by different types of distributions.)

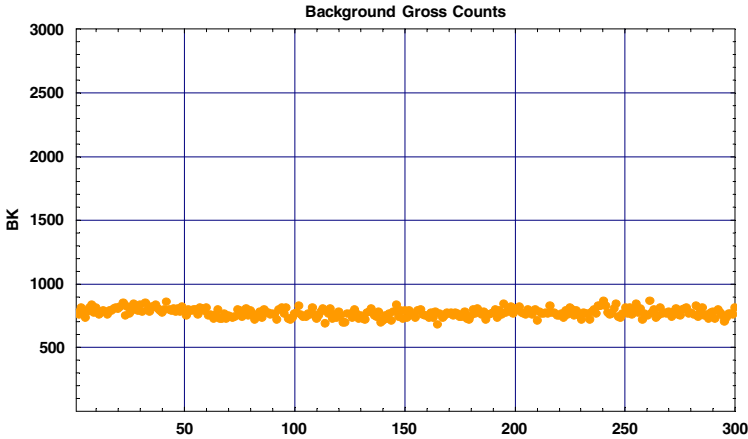


Fig. 1. Background gamma GC data, where the x-axis is the data-point position in the sequence

During the experiment, a common industrial source, Cs137, was placed on top of a table in the lab. The source gamma strength was measured at different distances from 1 to 50 feet for simulating sources of different strengths (Figure 2).

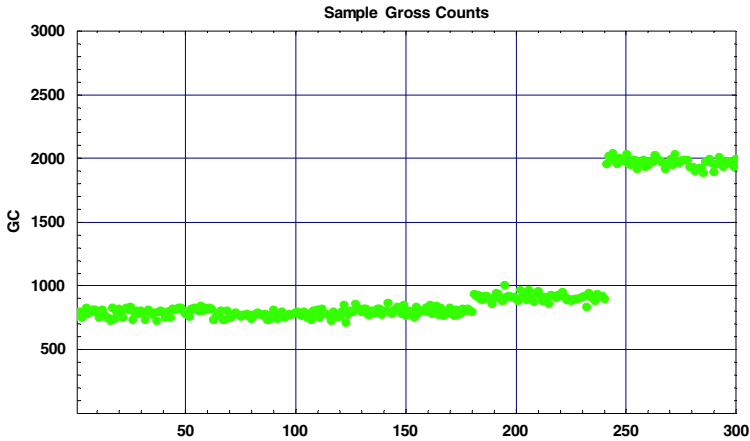


Fig. 2. Physical sample gamma GC data measured at different distances, where the x-axis is the data-point position in the sequence

### 9 Results and Analysis

The SPRT derived radiation levels (LGC) of the experimental data at different strengths (or distances) are shown in Figure 3. Comparing Figure 2 and 3, it is clear that the SPRT radiation level or LGC is exponentially correlated to the relative strength of the anomaly with respect to the background.

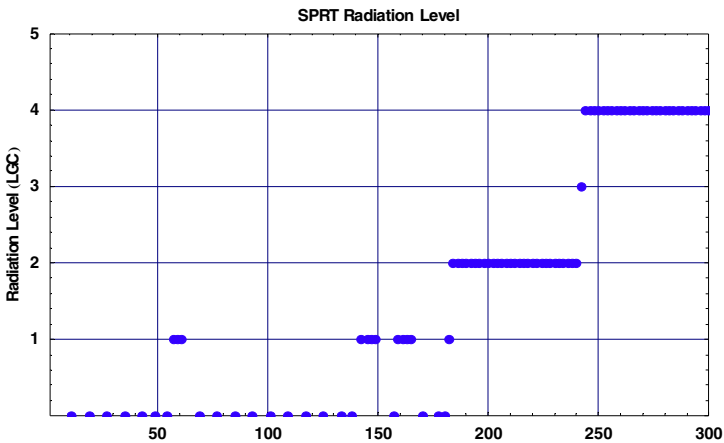
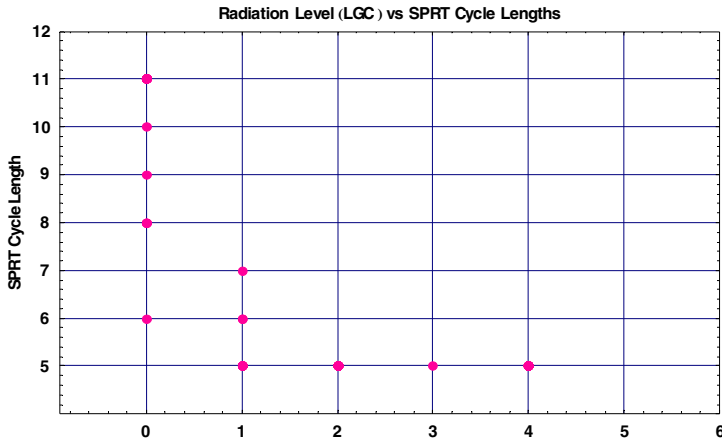


Fig. 3. SPRT derived anomaly radiation level (LGC) from the experimental data



**Fig. 4.** Plot of the SPRT estimated radiation levels (LGC) vs. the SPRT cycle length (the number of observations needed for deriving a LGC estimate) of the experimental data

Further, the relationship between the estimated radiation level (LGC) and the SPRT cycle length, (i.e., the number of observations needed for deriving an estimate for radiation level) is analyze in Figure 4.

It is clear that for high-radiation anomaly ( $LGC \geq 2$ ), the SPRT needs the minimum sample size  $i_{\min} = 5$  to derive the needed LGC with the given confidence levels. For low-radiation anomaly ( $LGC = 1$ ), the SPRT needs no more than seven observations. For background confirmation ( $LGC = 0$ ), the SPRT may need a longer observation cycle, anywhere from 6 to 11 observations. This is precisely a property we would like to have for a field radiation instrument, minimizing the time needed for high radiation anomaly sampling.

## 10 Summary

Preliminary assessment suggests that the SPRT method described in this paper is a promising algorithm for quick determination of field radiation levels. While maintaining prescribed confidence levels for its estimations, it reduces the sample sizes for high radiation regions or spots. This implies that it could reduce the exposure time for the field radiation surveyors.

## Acknowledgement

This manuscript has been authored by Bechtel Nevada under Contract No. DE-AC08-96NV11718 with the U.S. Department of Energy. The United States Government retains and the publisher, by accepting the article for publication, acknowledge that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes.



## References

1. Wald, A., 2004, *Sequential Analysis*, Dover Ed., Dover Publications, Mineola, NY, USA.
2. Knoll, G. F., 1999, *Radiation Detection and Measurement*, 3<sup>rd</sup> Ed., John Wiley & Sons, Hoboken, NJ, USA.
3. Stephen Wolfram, 2003, *The Mathematica Book 5ed.*, Wolfram Media, Champaign, IL, USA.