Pacific Northwest
NATIONAL LABORATORY

PNNL-34211, Rev. 1

Overview of a Methodology for Calculating the A Priori Scan Minimum Detectable Concentration for Post-Processed Radiological Surveys

FINAL REPORT

June 2023

Dan Fortin
Lisa Newburn
Deb Fagan
Jan Irvahn
Jen Huckett

Prepared for the U.S. Nuclear Regulatory Commission
Office of Nuclear Regulatory Research
Under Contract DE-AC05-76RL01830
Interagency Agreement: 31310019N0001
Task Order Number: 31310021F0022

PACIFIC NORTHWEST NATIONAL LABORATORY
operated by
BATTELLE
for the
UNITED STATES DEPARTMENT OF ENERGY
under Contract DE-AC05-76RL01830

Printed in the United States of America

Available to DOE and DOE contractors from
the Office of Scientific and Technical Information,
P.O. Box 62, Oak Ridge, TN 37831-0062
www.osti.gov
ph: (865) 576-8401
fox: (865) 576-5728
email: reports@osti.gov

Available to the public from the National Technical Information Service
5301 Shawnee Rd., Alexandria, VA 22312
ph: (800) 553-NTIS (6847)
or (703) 605-6000
email: info@ntis.gov
Online ordering: http://www.ntis.gov

Overview of a Methodology to Calculate the A Priori Scan Minimum Detectable Concentration for Post-Processed Radiological Surveys

FINAL REPORT

June 2023

Dan Fortin
Lisa Newburn
Debbie Fagan
Jan Irvahn
Jen Huckett

Prepared for the U.S. Nuclear Regulatory Commission
Office of Nuclear Regulatory Research
Under Contract DE-AC05-76RL01830
Interagency Agreement: 31310019N0001
Task Order Number: 31310021F0022

Pacific Northwest National Laboratory
Richland, Washington 99354

Summary

Increased continuous data collection using automated data loggers and autonomous radiological survey devices or vehicles has introduced a need for corresponding guidance and statistical techniques for data that are collected without surveyor vigilance. This report presents a method for calculating the *a priori* scan minimum detectable concentrations (MDCs) for surveys performed without vigilance similar to methods described in the Multi-Agency Radiation Survey and Site Investigation Manual, NUREG-1507, and NUREG/CR-6364. *A priori* scan MDCs are calculated during survey planning to ensure that survey parameters (e.g., scanning speed, scanning altitude, detector geometry) will lead to collecting data in which potentially contaminated areas can be detected when data are processed after the survey, within acceptable statistical error probabilities.

The *a priori* scan MDC calculations detailed in NUREG-1507 assume surveyor vigilance (i.e., pausing or stopping to investigate further when audio click data from a ratemeter indicates potential areas of concern). MARSSIM explains that ratemeters are the most common recording or display device used for portable radiation measurement systems; providing a display that represents the number of events occurring over a period (e.g., counts per minute [cpm]). Whereas the number of events can be accumulated over a period using a digital scaling device, resulting in information about the total number of events that occurred over a fixed period, ratemeter displays vary over time to provide short-term averages (NRC 2020).

This report expands upon NUREG-1507 to provide *a priori* scan MDC calculations assuming surveys will be completed without vigilance based on audio click data, providing an incremental advance in the *a priori* scan MDC methodology by moving from a with-vigilance to a without-vigilance surveying paradigm. A case study is provided to demonstrate the *a priori* scan MDC calculation. It is known, however, that there are differences between audio click used data in surveying with vigilance and data collected by the ratemeter used when surveys are conducted without vigilance. The magnitude of differences between the audio and ratemeter data, and their subsequent impacts on scan MDC calculations is yet unknown, to our knowledge. Future research and development will be required to assess such differences and determine the efficacy and applicability of the method developed in this report when applied to logged ratemeter data rather than audio data.

Varying background radiation levels have also posed challenges for scanning surveys. When surveys are conducted with vigilance, implicit data processing takes place in real time while surveyors notice gradual changes in audible background. When surveys are conducted without vigilance, new methods are required for the purposes of analyzing continuously collected survey data. The lag-k method is well-suited to a varying background. It performs hypothesis testing by treating survey data as a linear time series and assessing differences between k consecutively collected data points before and after each observation to account for potentially changing background. Simulation studies show that the lag-k method is superior to traditional approaches in terms of meeting acceptable error probabilities when optimal values of k are selected and contamination is localized (as opposed to wide-spread). Several case studies are provided to demonstrate the lag-k method.

Acronyms and Abbreviations

cpm	counts per minute
DCGL	derived concentration guideline levels
MARSSIM	Multi-Agency Radiation Survey and Site Investigation Manual
MCNPX	Monte Carlo N-particle Extended
MDC	minimum detectable concentration
MDCR	minimum detectable net count rate
NRC	U.S. Nuclear Regulatory Commission

Contents

Figures

Tables

1.0 Introduction

Radiological surveys to support the decommissioning of sites and facilities that may contain radioactive contamination involve both static measurements and scan measurements (scanning). Static measurements are aimed at characterizing an overall mean level of residual contamination but are unlikely to detect small areas of elevated activity (hotspots) because of the relatively limited area over which they are collected. Radiological scanning, which allows for greater spatial coverage over a comparable study period, can offer a higher probability of detecting an area of elevated radiological activity if it exists on a site.

Techniques for scanning have traditionally involved surveyors moving instruments over surface areas or land areas and responding to audio output from the instrument or surveying with vigilance. For planning purposes, an *a priori* scan minimum detectable concentration (MDC) is calculated prior to survey execution to ensure that survey parameters (e.g., scanning speed, scanning altitude, detector geometry) will lead to collecting data in which potentially contaminated areas can be detected by the surveyor during the scanning process within acceptable statistical error probabilities that meet regulatory requirements.

New technology has allowed scanning instruments to be coupled with data-loggers and positional information such as GPS, providing the possibility of scan surveys that collect continuous data without surveyor vigilance (i.e., where the surveyor is not responding to instrument output [audio click data] in real time). The resulting continuously collected scan data is available for post-processing analysis after the survey is conducted. While NUREG-1507 provides details on calculating an *a priori* scan MDC for the with-vigilance paradigm, no such guidance exists for planning purposes in the without-vigilance paradigm.

1. The report focuses on calculating an appropriate *a priori* scan MDC during the planning effort for surveys that will be conducted without vigilance. It reviews the current U.S. Nuclear Regulatory Commission (NRC) NUREG-1507 guidance and methodology for with-vigilance surveys, highlights limitations in the without-vigilance case, and develops a solution. This report extends Section 3 of NUREG-1507 to determine an *a priori* scan MDC when surveying without vigilance is planned, and Section 6 of NUREG-1507 which discusses post-processing continuously collected data.
2. This report also introduces a hypothesis testing procedure called lag-k. The approach is designed to account for a varying level of background radiation by estimating background from locations that are adjacent to the location of interest but not too close.

The following sections of this report describe survey paradigms (Section 2.0), existing NRC guidance for calculating *a priori* MDCs and our extension (Section 3.0), lag-k as a potential hypothesis testing method (Section 4.0), considerations when calculating *a priori* scan MDC for surveys without vigilance (Section 5.0), and a case study using the lag-k method (Section 6.0). Section 7.0 provides a history of literature and methods leading up to recent scan MDC methods and developments. Section 8.0 provides references.

2.0 Radiological Survey Scanning Paradigms

Radiological scan surveys are performed to identify potential areas of elevated radiation relative to a background level that should be investigated further. Regardless of the specific scan technique, the performance of a scan survey can be evaluated using a signal detection theory framework wherein an ideal observer is modeled using the detection decision probability distributions and, if necessary, human factors are incorporated to predict actual performance.

Two general classes of survey scanning paradigms are discussed: with- and without-surveyor vigilance. Historically, scanning was conducted with real-time surveyor vigilance, where the person conducting the survey continually monitored the audible detector response and decided whether that signal represented residual contamination in excess of background that should be investigated further. This paradigm formed the basis for the *a priori* scan MDC detailed in NUREG-1507 (NRC 2020). However, as survey technology has evolved to incorporate advances in data storage and mapping technology, surveys are more often being conducted without any monitoring or decision-making during the scan survey, thus eliminating the human factors related to surveyor efficacy. To put the discussion in context, the use of human factors is briefly discussed in Section 2.1, followed by a discussion of the considerations for the two sampling paradigms in Sections 2.2 and 2.3.

2.1 The Ideal Observer and Human Factors in Scanning MDC Calculations

As noted in NUREG-1507 (NRC 2020), scan survey efficacy is based on a number of human factors, some of which are present in scanning whether or not one is incorporating surveyor response to audio output to identify possibly elevated areas. We use the term "with-vigilance" to identify scanning activities that use the surveyor's audio response to alter the speed and/or coverage of possibly elevated areas. We use the term "without-vigilance" when scanning data are collected in the absence of surveyor response to audio output of the detector.

Scanning with-vigilance includes the influence of rewards for correctly identifying contamination and penalties for making both false-positive and false-negative errors. In addition, human limitations such as the ability to match the planned-for scanning speeds and techniques also influence decision errors.

When scanning without-vigilance, the surveyor is simply the vehicle by which the detector is moved around the survey unit. Human factors associated with real-time decision-making are eliminated but the human limitations of maintaining constant, ideal scan speed and distance remain.

In both cases, the minimum detectable count rate (MDCR) is adjusted based on surveyor efficiency, a value between one and zero that captures the decrease in detectability based on human factors. NUREG-1507 suggests that for scanning with-vigilance, surveyor efficiency ranges between 0.5 and 0.75, based on simulations and field studies that included a population of surveyors and a suite of scenarios and included both surveyor decision-making as well as surveyor scan parameters (speed, distance, etc.).

Up to this point, work has not been done to estimate the human factors in a survey without-vigilance paradigm. A follow-on activity to study the effect of surveyor scan parameters using scanning technology coupled with the human surveyor should be considered to evaluate various gross-counting instruments applied to sources, and source scenarios to vary speed, distance, type of source (distributed vs. point source), and shielding should be considered to understand the impacts of these variables on survey results. Further, technology factors (e.g., unmanned ground or aerial vehicles) and human-technology teaming factors (e.g., remotely manned vehicles) should also be considered for surveys that are without vigilance.

2.2 Scanning with Real-Time Surveyor Vigilance

Scanning with real-time surveyor vigilance is described in detail in NUREG/CR-6364 (NRC 1998) and NUREG-1507 (NRC 2020). To summarize, the surveyor conducts a scan moving at a consistent speed, listening to the clicks from a detector, and pauses for additional observation upon noticing a significant increase in audible output that could represent an area of radiological contamination in excess of background or a previously determined investigation level. This method does not rely on electronic recording and storage of detector readings and locations, and the pause upon hearing increased detector output helps make sure the peak count rate at a given location is accurately captured. However, this paradigm is affected by additional human factors that can affect detection capability. Because the surveyor decides when to pause for additional observation and flagging, the surveyor's prior beliefs about the likelihood and intensity of a contaminated area, along with the expected cost of either missing a source or flagging a false positive, will affect the actual detection performance of this survey method.

2.3 Scanning with Limited to No Real-Time Surveyor Vigilance

Since the development of the *a priori* scan MDC equation in NUREG-1507 (NRC 2020), survey technology has evolved to using GPS instrumentation along with data acquisition software to automatically record radiation measurements along with corresponding geospatial coordinates. The post-processed map can then be evaluated to characterize contamination and determine whether further investigation is needed. Several considerations for how this paradigm changes the influence of human factors are introduced here and drive the discussion in the rest of this report.

- Instead of relying on a surveyor to listen and decide whether to respond to increased counts, a survey of this type will automatically record average count rates over fixed counting intervals as discrete point estimates. The distance between discrete points will vary depending on scan speed, and the accuracy of each location should be considered in light of data acquisition timing and GPS logging and accuracy.
- Because this type of survey presents an alternate scanning paradigm wherein the scan results are not evaluated or responded to in real time, the human factor of surveyor efficiency accounting for error in deciding whether contamination is present is eliminated. However, the opportunity to immediately pause to further investigate potential areas of contamination no longer exists.
- An additional significant difference in this survey paradigm is that it results in dense, spatially registered data that can be used to more accurately and completely characterize the survey results. The availability of complete survey data also opens avenues for retrospective analysis of the survey technique, efficiencies, and scan speeds.

In Section 6.3 of NUREG-1507, the authors state, "Furthermore, without pausing, the electronically captured count rate data may not accurately reflect the peak count rate at a given location as the ratemeter may not reach full scale" (NRC 2020). This concern poses a real barrier to developing a method for calculating a scan MDC in a without-vigilance paradigm. The statistical methodology underlying the limit of detection calculation is based on the data stream produced by the audio clicks as heard by an actively listening surveyor.

The scan MDC for post-processed data presented in this document relies on the captured data being a binned version of the audio click data stream. Ratemeter counts that do not satisfy this requirement will need an additional method to account for any non-linear transformation of the audio click data stream to a ratemeter count data stream.

3.0 Scan MDC

The process of calculating the scan MDC is complicated and relies heavily on statistical decision theory, health physics expertise, and access to sophisticated modeling software. In particular, this report focuses on a statistical approach to calculating the MDCR, a critical element of the *a priori* scan MDC calculation. This report uses a familiar classical statistical framework, but recent developments in Bayesian models for signal detection of gross count measurements (Brogan and Brandl 2019) may provide another useful signal detection framework in the post-processing context. Additional work will be needed to compare this model to traditional and some non-traditional approaches for *a priori* MDC calculations for survey planning purposes.

NRC guidance addressing the fundamental concepts of MDCs for both static and scan surveys is described in NUREG-1507 (NRC 2020). In addition to the MDC concepts and statistical foundations, NUREG-1507 focuses on the practical considerations of variables affecting MDCs, such as the type of instrumentation used for surveying and the field conditions. For scanning surveys, characterizing surveyor efficiency plays an important role in calculating the scan MDC.

Figure 3.1 shows a conceptual algorithm with requirements for calculating an *a priori* scan MDC. Field conditions, instrument information, and human performance inputs capture surveyor, surface, instrument efficiencies, count-to-exposure-rate ratio and/or exposure-rate-to-concentration factors. The conversion from MDCR (count rate) to MDC (concentration) is necessary so that MDC can be compared to regulatory limits such as the derived concentration guideline levels (DCGLs) (NUREG-1575, Rev. 1 [NRC/EPA/DOE 2000]) to ensure these quantities can be detected using survey parameters during planning.

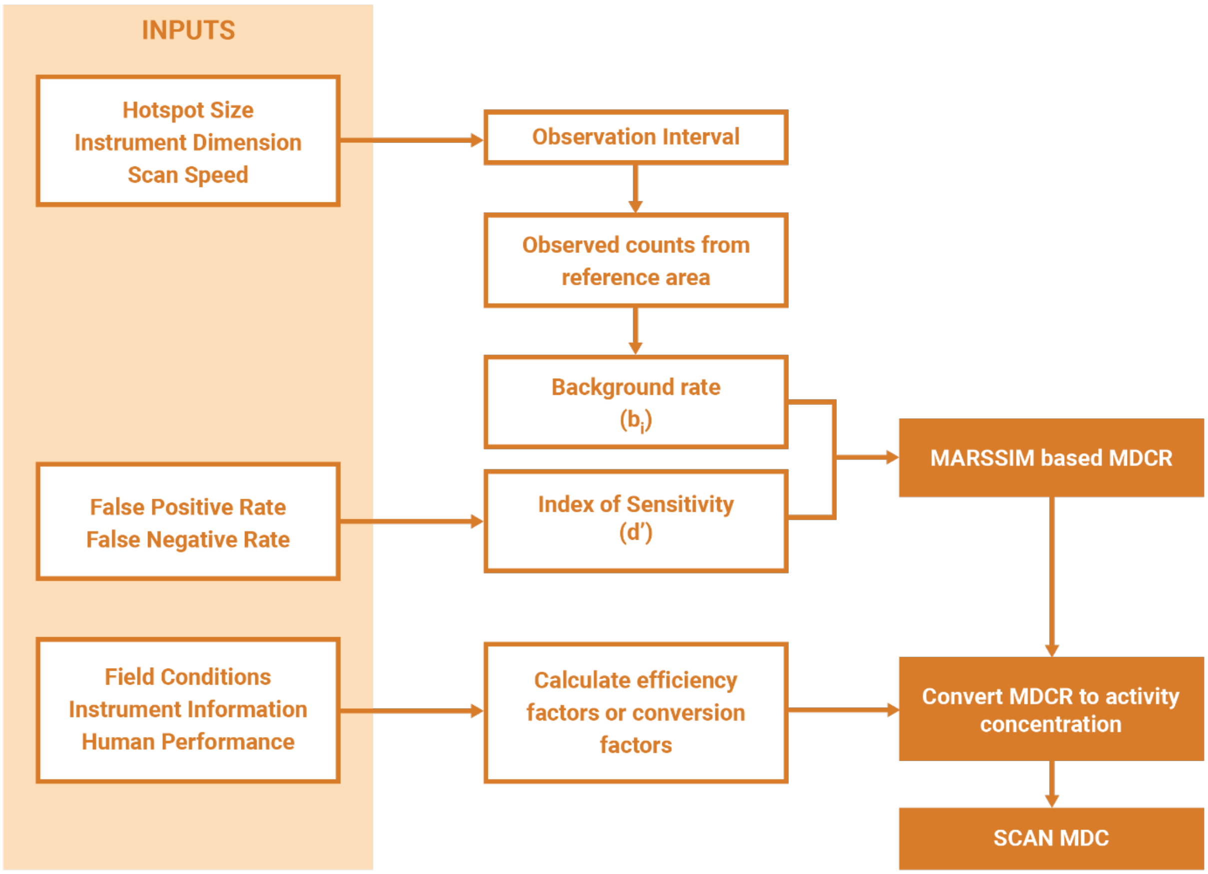

Figure 3.1. Flowchart showing the requirements for calculating a scan MDC.

There are multiple approaches for calculating scan MDC that are based on different assumptions concerning the data generating process. In Section 3.1 we detail the current guidance in NUREG-1507 for calculating scan MDC based on signal detection theory (found in Section 6 of NUREG-1507). In Section 3.3 we detail an alternative approach based on the limit of detection that follows the approach referenced in Section 3 of NUREG-1507. We demonstrate the connection between the limit of detection and d' (details in Section 3.1.3). Both approaches calculate the MDCR and then convert the MDCR to a scan MDC using software such as MicroShield. The key difference is that the existing guidance computes the MDCR based on d' while the alternative computes the MDCR based on the limit of detection.

3.1 Statistical Framework Using Signal Detection

The objective of a scan survey is to determine whether observed measurements represent background radiological levels or whether they reflect residual contamination signals in addition to the background (NUREG/CR-6364 [NRC 1998]). Figure 3.2 illustrates two corresponding statistical probability distributions, where the net background (noise) distribution is centered at zero and the background plus signal (noise + signal) distribution is centered at two standard deviations above zero. It is assumed that each is well-approximated by a normal distribution with standard deviation σ (NUREG-1507 [NRC 2020]). The difference between these distributions is captured by the difference in the means.

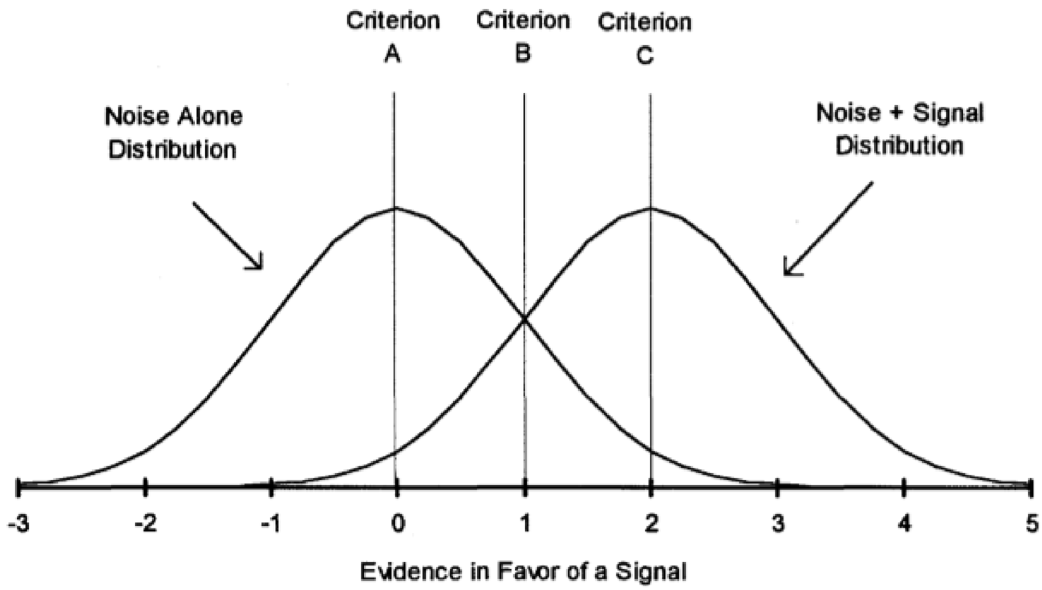

Figure 3.2. The signal detection theory view of signals in noise and decision criteria (from NUREG-CR-6364 [NRC 1998]).

3.1.1 Decision Criteria

A decision criterion is typically established to determine whether a signal is present at any given location. Figure 3.2 shows three different criteria (Criteria A, B, and C).

- Samples from the distribution centered on Criterion A are obtained by subtracting an estimate of the mean background from the observed gross background observations. Criterion A is the mean of the net background (noise only) distribution.
- Samples from the distribution centered on Criterion C are obtained by subtracting the same estimate of the mean background from the observed gross signal + noise count. Criterion C is the mean of the net signal + noise distribution.
- Criterion B is the intersection of the noise-only and signal + noise distributions. Assuming the distributions have a common standard deviation, Criterion B is the midpoint between the two means, A and C.

Each criterion will result in different true-positive and false-positive rates. For example, using Criterion A would flag most of the observations from the noise + signal distribution (true positives), it would also flag 50% of the observations from the background or noise-only distribution (false positives). Using Criterion C would result in fewer false positives but would increase the number of false negatives by flagging only 50% of the observations from the signal + noise distribution. Criterion B leads to equal false positive and false negative rates in this case.

3.1.2 Ideal Observer and Surveyor Efficiencies

During survey planning, the selection of decision criteria is influenced by several factors, including *a priori* probabilities of residual contamination and costs associated with outcomes due to potential false negative and false positive errors. If the distributions describing the detection scenario can be specified similar to Figure 3.2, then the error rates can be controlled by selecting an optimal decision criterion, independent of human factors. This theoretical optimal detection criterion and subsequent decision-making capability is referred to as an ***ideal observer***. “The ideal observer makes optimal use of the available information to maximize the percentage of correct responses, providing an effective upper bound for comparisons with actual surveyors” (NUREG-1507 [NRC 2020]). An MDCR is determined for an ideal observer based on the index of sensitivity d', minimum detectable number of net source counts in the observation interval, and background counts in the observation interval, where the observation interval is determined based on scan speed and area of extent of the contamination.

When scanning surveys are completed with vigilance, survey practitioners respond in real time to audio output, making their best judgment about whether a signal above background is present or not at numerous locations along the scanning transect and then pausing to collect additional data where their judgment indicates further data collection is necessary. Relying on the surveyor to respond introduces inefficiencies when compared to an ideal observer. Experiments have shown surveyor performance/efficiency varies depending on multiple factors (scan speed, detector height, variable background (noise) distributions within a site, etc.) (NUREG-1507 [NRC 2020]). These are accounted for in NUREG-1507 by applying a “surveyor efficiency factor” to the MDCR from the ideal observer to calculate the scan MDC value (along with instrument and source efficiency factors). The MDCR and scan MDC are intended for planning and detection in the first phase of data collection only, and this report does not describe additional details regarding the second phase conducted during the pause, although such details are provided in NUREG-1507 (NRC 2020).

When surveys are conducted without vigilance, this same surveyor efficiency is not applicable because the surveyor does not respond to audio output in real time during the survey but after the survey is complete by post-processing collected ratemeter data. There are, however, other conditions and inefficiencies that must be considered. First, collected ratemeter data (reported in cpm) is not equivalent to a binned version of the audio click output from the detector. Next, since the ratemeter data are collected without pausing “the electronically captured count rate data may not accurately reflect the peak count rate at any given location, as the ratemeter may not reach full scale if the observation interval over an area of elevated direct radiation is less than two to four seconds” (NUREG-1507 [NRC 2020]), although this is a concern in the first phase of with-vigilance surveys too. Finally, while some technology exists that includes integrated sensor and detector technology (e.g., the RS-700 mounted radiation detection system), others combine components (e.g., GPS, altimeter, scanner) from different equipment manufacturers to collect each data stream individually. Reasons for combining the components from different manufacturers or producers can include cost, availability, and the desire for customization to address site-specific needs.

When combining various components, collected data streams may not be aligned in time with the ratemeter cpm observations, and so additional data processing is required to align them with the cpm data, potentially introducing additional variation or uncertainty that translate to additional inefficiency. The consequence of these factors is that the resulting cpm and supporting data distributions are likely different from the Poisson distribution used to model the audio output, and so, as suggested in NUREG-1507, the corresponding *a priori* scan MDC calculations for with-vigilance surveys are likely inaccurate for without-vigilance surveys (NUREG-1507 [NRC 2020]).

Differences between the ratemeter cpm data and the audio data have not been well studied. As an incremental step toward producing a method to calculate an *a priori* scan MDC for without-vigilance surveys, this report makes the simplifying assumption that ratemeter cpm data are a binned or integrated version of the audio data. Using this assumption, this report develops an *a priori* scan MDC. This development moves the needle, although not fully, to the goal of developing an *a priori* scan MDC for surveys planned without vigilance. To achieve that goal fully, research will be required to understand the implications of collecting cpm data without vigilance to determine the following and develop a scan MDC that accounts for these factors.

- The extent to which the cpm data distribution diverges from the normal approximation to the Poisson assumed for the audio data.
- The effect of post-processing required to align other data streams with the ratemeter cpm data.
- Based on these findings, whether or not the NUREG-1507 *a priori* scan MDC calculations can be calibrated, say, by applying a factor to account for vigilance, and/or whether a new derivation will be required for the without-vigilance case.

Additional challenges due to variation in background (noise distribution) within a site are present for surveys conducted both with and without vigilance. These are addressed among various other topics in NUREG-1507 (NRC 2020) using a post-processing approach. For the purposes of planning and calculating *a priori* scan MDCs, however, alternative methods are required. This paper proposes the lag-k method as one approach, and it is described in detail in Section 4.0.

3.1.3 Statistical Decision Theory and MDC Equations

This section discusses fundamental statistical concepts of statistical decision theory and the hypothesis testing framework used to develop MDC equations for both static measurements (static MDCs) and for scan surveys (scan MDCs). An inconsistency in NUREG-1507 notation and citations for MDC equations is addressed here by showing a direct connection between the detection limit, L_D, used in the static MDC context, and the index of sensitivity, d', used in the scan MDC context.

NUREG-1507 (NRC 2020) describes the fundamental concepts of statistical decision theory for MDCs using a hypothesis testing framework developed by Currie (1968) for the purposes of making a decision about the presence of activity at a site. The null (H_0) and alternative (H_A) hypotheses are framed as follows.

H_0: No net activity due to residual contamination is present at the site.

H_A: Net activity due to residual contamination is present at the site.

Data distributions associated with each hypothesis are determined by making the following assumptions (NUREG-1507 [NRC 2020]).

- Data are assumed to follow Poisson distributions, which are adequately approximated by normal distributions in both the null and alternative hypotheses.
- The data distribution under the null hypothesis (H_0) is normal and centered at zero, representing the net noise distribution when no net activity is present.
- The data distribution under the alternative hypothesis (H_A) is normal and centered at a point greater than zero, representing the net signal distribution when net activity is present.

Signal detection theory is used to derive MDC calculations (NUREG-1507 [NRC 2020]) as follows. Originally developed for naval radar applications, where humans were tasked with responding to audible radar pulses, the index of sensitivity d' represented the distance between mean number of pulses in the background (noise only) and the mean in the signal + noise distribution, in units of the standard deviation σ. Given the simplifying assumption that these distributions share a common standard deviation σ, and the standard normal distribution quantiles corresponding to false positive rate $1-\beta$, $Z(false\ positive) = k_{1-\beta}$, and true positive rate α, $Z(true\ positive) = k_{\alpha}$, d' is calculated as follows.

$$d' = Z(false\ positive) – Z(true\ positive) = k_{\alpha} – k_{1-\beta} = k_{\alpha} + k_{\beta}$$

Here, $1-\beta$ is the desired true-positive rate, α is the desired false-positive rate, and k_p ($p = \alpha$ or β) is the $(1-p)^{th}$ percentile of the standard normal distribution. The motivation for using signal detection theory is its ability to characterize surveyor performance via this d' statistic (Abelquist 2014). For example, if the false negative rate is $\beta = 0.2$, then the true positive rate is $1-\beta = 1-0.2 = 0.8$ and the corresponding quantile is $k_{\beta} = \mathrm{Z}(1-\beta) = 0.84$. If the false positive rate is $\alpha = 0.05$, then that quantile is $k_{\alpha} = Z(1-\alpha)$ = 1.645. So, $k_{1-\beta}$ is smaller than k_{α}, resulting in a positive value for $k_{\alpha} - k_{1-\beta} = 1.645 - -0.84 = 2.49$, or the distance d'. Since $k_{1-\beta}$ = $-k_{\beta}$, that means $k_{\alpha} - k_{1-\beta}$ can be rewritten as $k_{\alpha} + k_{\beta}$ as shown on the right-hand-side of the equation.

Static MDC calculations in NUREG-1507 (NRC 2020) make use of critical levels and detection limits based on acceptable false positive and false negative rates specified during the planning phase in place of d', although the two are related. The critical level L_C depends on an acceptable false positive rate with respect to the null hypothesis (noise only distribution), illustrated in Figure 3.3 for $\alpha = 0.05$. The detection limit L_D depends on an acceptable false negative rate with respect to the alternative hypothesis (signal + noise distribution), shown in Figure 3.4 for $\beta = 0.20$. Aligning the regions bounded by L_C and L_D, the relative position under null and alternative hypotheses can be seen in Figure 3.5, demonstrating the connection between L_C and L_D and the index of sensitivity d'. The relationship can be written mathematically as follows.

$$L_D = k_{\alpha}\sigma + k_{\beta}\sigma = \left(k_{\alpha} + k_{\beta}\right)\sigma = d'\sigma$$

The purpose of this detail is to highlight that the methods developed for static and scan MDC rely on similar methodology despite differences in notation. The emphasis on d' in the scan MDC context is driven by a focus on human/surveyor performance.

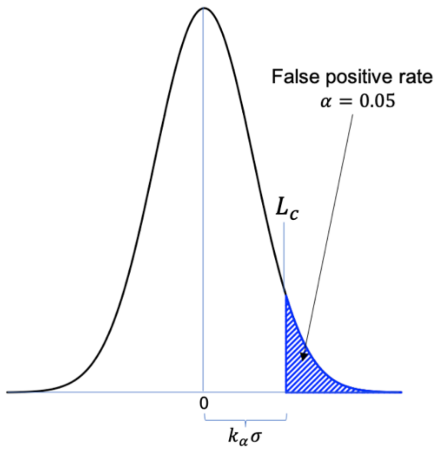

Figure 3.3. False-positive rate, α, and associated critical level, L_C, for a distribution of net counts under the null hypothesis of no net activity.

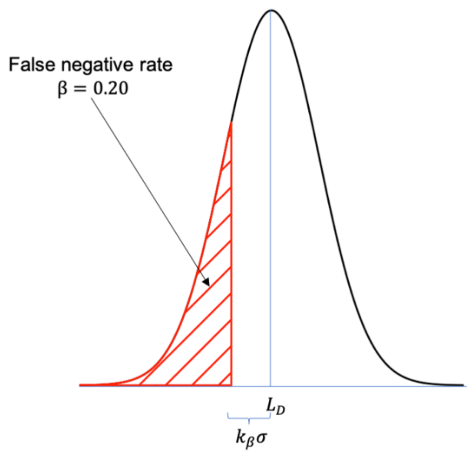

Figure 3.4. False-negative error rate, β, and its relationship to the detection limit, L_D.

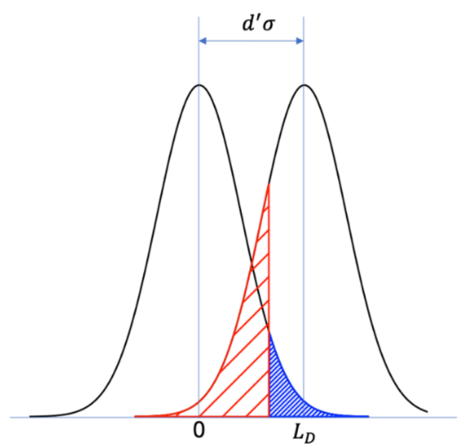

Figure 3.5. Distributions of the null (centered at zero) and alternative (centered at L_D) hypotheses and the relationship between the detection limit, L_D, and sensitivity index, d'.

3.1.4 Development of the Minimum Detectable Count Rate

The MDCR is an integral component of the MDC calculation. It is a function of the minimum detectable number of net source counts in an observation interval, the index of sensitivity, the background counts in the observation interval, and the observation interval length (in seconds), which depends on the scan speed and areal extent of contamination.

In the statistical decision-making framework outlined above, d' is a sensitivity index that adjusts the background counts in an observation interval situation based on the acceptable false-positive (Type I error) and false-negative (Type II error) detection rates specified in the data quality objectives, as described in NUREG-1575, Rev.1 (NRC/EPA/DOE 2000). The sensitivity index, d', is defined as the difference between the means of the noise and signal + noise distributions in units of the common standard deviation. Figure 3.2 illustrates an example of a noise and signal + noise distribution on a standardized scale. Here, using the notation from Figure 3.2 and 3.5, $d' = C - A = 2$, indicating that the mean of the signal + noise distribution is two standard deviations greater than the mean of the noise distribution. NUREG-1507 (NRC 2020) notes that for a true-positive rate of 95% and a false-positive rate of 5%, d' is equal to 3.29.

The MDCR for an ideal observer is given by the following equation.

$$MDCR = d' \times \sqrt{b_i} \times \frac{60}{i} \qquad (3.1)$$

Here,

- MDCR = minimum detectable (net) count rate in counts per minute;
- b_i = background counts for an observation time interval that lasts i seconds;
- d' = index of sensitivity based on false-positive and false-negative error rates; and
- i = temporal extent of the observational interval (in seconds), based on the scan speed and the areal extent of the contamination hotspot.

To understand the MDCR Equation (3.1), it is helpful to understand the statistical properties of the Poisson distribution, often used to describe counting processes that take place within some finite time interval, such as the number of observed audible clicks in an observation interval. It is characterized by a single parameter that represents both the mean and variance of the distribution. This implies that an observed count is the best estimate for the site's mean count and variance, assuming a Poisson distribution for counts. As usual, the standard deviation is equal to the square root of the variance, so the observed mean count provides an estimate of the site standard deviation. In Equation (3.1), $\sqrt{b_i}$ is an estimate of the background distribution's standard deviation.

3.1.5 Conversion to Scan MDC

Once the hypothesis testing framework controlling for specified error rates has been used to calculate the MDCR, it can then be converted to a scan MDC. Determining the scan MDC requires applying of conversion factors to the MDCR to translate from the net count rate to the measurable surface activity or soil concentration. The steps outlined by Abelquist (2014) in the process of calculating scan MDCs for land areas are as follows:

1. Calculate the MDCR (in counts per minute) for a given background, observation interval, and performance level (Type I and Type II decision errors).

2. Convert the MDCR to a surveyor MDCR using surveyor efficiency p.
3. Translate the MDCR to the minimum detectable exposure rate using the relationship of net count rate to net exposure rate for a particular survey instrument.
4. Translate the minimum detectable exposure rate to the scan MDC using a model such as MicroShield for specific conditions.

Under the survey paradigm with no real-time surveyor vigilance, Step 2 can be effectively eliminated because error rates are not based on the human factors related to surveyor response. The remaining steps can be followed as-is. Methods and computer codes (e.g., MicroShield) used to establish these relationships are described in NRC (2020) and Abelquist (2014).

3.2 Scan MDC Based on Limit of Detection

Section 6 of NUREG-1507 calculates scan MDC from the MDCR, which is derived from d'. In this section we show how to calculate a scan MDC based on the limit of detection rather than d', the connection between d' and the limit of detection having been established in Section 3.1.3. The limit of detection approach was previously employed in Section 3 of NUREG-1507 and follows Currie's 1968 method.

In order to calculate a limit of detection, we must first calculate the critical limit as described in Section 3.2.1. The critical limit is the threshold of net counts at which follow-up action will be taken. Net counts below the critical limit are not surprising under the null hypothesis of no contamination. Net counts above the critical limit are suggestive of contamination, and there is evidence that the alternative hypothesis is true.

3.2.1 Hypothesis Test Based on Critical Level

Let $\mathbf{X}(s_i)$ represent measurements (i.e., gross counts) recorded at location s_i, where $i = 1, \ldots, N$, during a time interval t_i (e.g., $t_i = 1\ second$) of observation. A scanning instrument is assumed to move continuously at a fixed rate in meters/second, and s_i is assumed to be the spatial midpoint of observation interval t_i. For each location s_i, the observed counts are assumed to follow a Poisson distribution, written as follows.

$$\mathrm{X}(s_i) \sim Poisson(\lambda_i) \tag{3.2}$$

Here, the parameter λ_i represents the mean and variance of the Poisson distribution. When the mean is sufficiently large, the Poisson is well-approximated by a normal distribution and written as follows, where $\mu_i = \sigma_i^2 = \lambda_i$.

$$X(s_i) \sim N\left(\mu_i, \sigma_i^2\right) \tag{3.3}$$

The test statistic for a hypothesis test for the difference between mean count at location i and the mean background count is calculated based on the difference between gross counts observed at location i and the average measurements observed at N background locations. The test statistic $D(s_i)$ is then compared to a standard normal distribution to determine a p-value and is calculated as follows, where i and j are location indices. The background measurements must be obtained in the same way as site measurements.

$$D(s_i) = X(s_i) - \frac{1}{N} \sum_{j \in BackgroundLocations} X(s_j)$$

Under the null hypothesis, the mean measurement $X(s_i)$ is equal to the background mean, so the expected value of $D(s_i)$, written $E[D(s_i)]$, is zero.

$$E[D(s_i)] = 0$$

Under the null hypothesis, $X(s_i)$ also has the same variance at all locations, σ_B^2. Under the null the variance of $D(s_i)$ is

$$Var(D(s_i)) = Var(X(s_i)) + \frac{1}{N} Var\left(X(s_j)\right),$$

$$= \left(1 + \frac{1}{N}\right) Var(X(s_i)),$$

$$= \left(1 + \frac{1}{N}\right) \sigma_B^2.$$

The critical level is:

$$L_c = k_\alpha \times \sqrt{\left(1 + \frac{1}{N}\right) \sigma_B^2}.$$

So, L_c is k_α standard deviations of $D(s_i)$ above the mean of $D(s_i)$ (which is 0). This critical level is the hypothesis testing threshold for net counts. Locations where net counts are above the critical level require follow-up. Locations where net counts are below the critical level do not require further action.

3.2.2 Limit of Detection Based on Critical Level

Once we have computed the critical level, we compute the limit of detection. Under the alternative hypothesis, at location s_j specifically, $X(s_j)$ still follows a Poisson distribution,

$$X(s_j) \sim Poisson(\lambda_j + \lambda_S),$$

where λ_j is the mean of the background distribution at location s_j and λ_S is the mean of the signal at location s_j. When $\lambda_j + \lambda_S$ is large, this is well approximated by a normal distribution,

$$X(s_j) \sim N(\mu_j, \sigma_j^2),$$

where $\mu_j = \sigma_j^2 = \lambda_j + \lambda_S$. Under the alternative hypothesis. our test statistic no longer has mean 0,

$$E[D(s_j)] = \lambda_S.$$

Under the alternative hypothesis, the variance of our test statistic at location s_j is

$$Var\left(D(s_j)\right) = Var[X(s_j)] + \frac{1}{N}Var[X(s_i)],$$

$$= \lambda_j + \lambda_S + \frac{1}{N}\lambda_j,$$

$$= \left(1 + \frac{1}{N}\right)\lambda_j + \lambda_S,$$

$$= \left(1 + \frac{1}{N}\right)\sigma_B^2 + \lambda_S,$$

$$= \sigma_D^2,$$

assuming independence between all measurements.

The limit of detection is:

$$L_D = L_C + \sigma_D k_\beta,$$

$$= \sqrt{\left(1 + \frac{1}{N}\right)\sigma_B^2}k_\alpha + \sqrt{\left(1 + \frac{1}{N}\right)\sigma_B^2 + \lambda_S}k_\beta.$$

Of course, λ_S *is* the limit of detection, so we have a quadratic equation to solve,

$$L_D^2 - L_D\left[2\sqrt{\left(1 + \frac{1}{N}\right)\sigma_B^2}k_\alpha + k_\beta^2\right] + \left(k_\alpha^2 - k_\beta^2\right)\left(1 + \frac{1}{N}\right)\sigma_B^2 = 0.$$

As noted in the literature, this simplifies when $\alpha = \beta$,

$$L_D = 2\sqrt{\left(1 + \frac{1}{N}\right)\sigma_B^2}k_\alpha + k_\alpha^2,$$

$$= 2L_C + k_\alpha^2.$$

The MDCR for an ideal observer is $L_D \times 60/i$, where i is the length of the observation interval in seconds. Note that L_D takes the place of $d' \times \sqrt{b_i}$ in the signal detection theory approach to calculating MDCR. The two terms, L_D and $d' \times \sqrt{b_i}$, are the same when we assume both normal distributions have the same variance. As shown, they do not have the same variance, and the equal variance assumption is not needed. The conversion to scan MDC from MDCR proceeds as described in Section 3.1.5.

3.3 Considerations for Using NUREG-1507 Scan MDC Calculations

NUREG-1507 (NRC 2020) refers to a counting observer as one who makes a decision regarding the presence of contamination based on a number of counts observed in a given time interval (see Section 6.5 in NUREG-1507) in real time (i.e., the surveyor vigilance approach). Data collected using a GPS-based system that records counts at a regular interval, typically one second for GPS-based gamma walk-over scans (i.e., no surveyor vigilance), can be assessed in a similar fashion post-survey. However, there are several important considerations.

- Without vigilance, the lack of a real-time surveyor response implies that an observer cannot adjust for local background fluctuations.
- A constant background mean across the survey unit must be assumed, but the actual background can often be highly variable (in both with and without vigilance surveys).
- The observation interval is different for with versus without vigilance surveys—with vigilance, the human surveyor implicitly defines the observation interval in real time whereas without vigilance, the observation interval is defined by the programming of the data logging equipment (including scan speed).
- There is a lack of defined methods for assessing survey results and guiding further investigation, thereby resulting in non-uniform implementation.

As GPS-based surveys with no surveyor vigilance have become more widespread, the need to develop guidance to make sure these considerations are addressed has increased. This section evaluates existing scan MDC methods applied when surveying without-vigilance for a counting observer described in NUREG/CR-6364 (NRC 1998) and NUREG-1507 (NRC 2020) in the context of this type of a survey.

3.3.1 Calculation of MDCR

Section 6.7 of NUREG-1507 develops the steps for calculating the MDCR using the ideal observer, arguments related to counting statistics, and the notion of surveyor efficiency. The number of counts in the MDCR equation is achieved for a given background by arguing that the detectability index D is asymptotic to the sensitivity index d' of an ideal observer, where d' is determined via a statistical hypothesis testing framework (see Section 2.0 above). The resulting number of counts is then adjusted by a surveyor-specific efficiency term to arrive at an MDCR, which is then used to determine in real time where contamination is present. The distinction between that and what is advocated for here is that the contamination decision is made during post-processing rather than in real time. The sensitivity index d' based on acceptable decision errors is still applicable to this paradigm and can be selected and used in the calculation of the post-survey MDCR.

3.3.2 Limitations of Existing Scan MDC Approach Applied to Surveying Without-Vigilance

One key difference between scan MDC calculations based on the survey with-vigilance paradigm and scan surveys using automatic data-loggers in post-survey decision-making is the definition of the observation interval. In the with-vigilance approach, a human surveyor is listening and responding in real time—the observation interval is defined as the time that the detector can respond to the contamination source. The observation interval can be increased directly by surveying at a slower pace, effectively decreasing the calculated scan MDC. In the without-vigilance approach, the observations are automatically recorded at specific time intervals (typically 1 second for GPS-enabled detectors). Further consideration is therefore required regarding the assumption that the detector perfectly traverses the hotspot during the 1-second interval. Both the speed of the detector and the distance from the center of the source have significant impacts on detector efficiency (Hart 2003).

Details outlining how to incorporate detector speed and distance using GPS-based gamma radiation surveys are described by Alecksen and Whicker (2016). Alecksen and Whicker (2016) present an approach for producing scan MDCs based on detector efficiencies modeled with probabilistic Monte Carlo N-Particle Extended software simulations.

Another issue associated with the scan MDC is that the mean number of background counts is assumed to be constant sitewide, when it is well-known that the background mean varies spatially. The hypothesis testing framework considers each point of measurement individually and is therefore agnostic to the variable yet spatially correlated counts realized in a scan survey. Existing approaches to mean background estimation implicitly assume a constant mean across the survey unit and create vulnerability to missing areas of elevated contamination depending on the level of variability in background conditions. Because the shift to the without-vigilance paradigm means that an observer is no longer responding to an increased number of counts, the opportunity for any kind of localization or adjustment to local background fluctuations is lost. This is a major concern related to using the with-vigilance approach in the without-vigilance decision context. Section 4.0 describes a statistical approach that has statistical foundation and the potential to mitigate the effects of background variability.

An additional limitation of the without-vigilance paradigm is in the lack of established methods the with-vigilance approach provides for survey planning. Because the measured count rates and locations are all recorded prior to any decision-making, there is an opportunity to use those data for retrospective evaluation of the achieved scan MDC. However, the traditional scan MDC approach described in the Multi-Agency Radiation Survey and Site Investigation Manual (MARSSIM) (NRC/EPA/DOE 2000) does not include specified methods for such assessment.

Finally, the with-vigilance approach requires further investigation in any area in which readings exceed the critical level. This is done easily by including a real-time pause in the with-vigilance paradigm during which surveyors can err on the side of conservatism, but a without-vigilance evaluation requires development of a post-survey investigation level that will satisfy the required detection sensitivity. Practitioners should consider increasing the false-positive tolerance in the without-vigilance paradigm to mimic the “conservatism” in the with-vigilance paradigm. Subsequent revisiting, rescanning, and/or sampling and lab analysis in identified areas will be required to evaluate whether the location is in or near a hotspot.

4.0 Approach to Variable Background

A statistical approach for calculating the MDCR for scan MDC calculations makes use of the availability of georeferenced measurements in post-processed data from surveys conducted without vigilance. While appropriate scan MDC calculations have been developed for post-processing cases, they are focused on detector efficiency calculations and computer codes for converting observed count rates to contaminant concentrations (Alecksen and Whicker 2016; Falkner and Marianno 2019; King et al. 2012). Previous work has provided methods relevant to the bottom box of conversion factor inputs in Figure 3.1. This section is focused on methods for the top two input boxes (i.e., hotspot size, instrument dimension, scan speed, and statistical parameters) to inform statistical calculations of the MDCR.

One challenge with scanning to identify areas of contamination relative to background is that background radiation levels can vary dramatically. Therefore, an appropriate model for background radiation is not a constant value characterized by an average, but a two-dimensional surface with peaks, valleys, and possibly step-function “jumps” across the site. When an underlying trend is a nuisance to an analysis and not the focus, one way to mitigate its effect is to perform local differencing (computing differences between observations and their neighbor[s]). This approach is taken here to compute net counts using spatially localized average background rather than subtracting an overall background count rate. The distance between spatially localized neighbors is determined by the site conceptual model and observed variation in background such that neighbors are assumed to be independent (uncorrelated).

This method is called the lag-k approach, where k indicates the distance between independent neighbors. In lag-k, net counts are estimated using differences between observations at least k units apart. The following sections provide statistical properties of the distribution of such differences and explain how they are used in a hypothesis testing framework. This approach is robust relative to fluctuations in background and uses a familiar hypothesis testing framework for the scan MDC where the main inputs are still the desired false-positive and false-negative error rates familiar to MARSSIM users for calculating the scan MDC.

Lag-k is appropriate when the conceptual model for a site indicates that background radiation levels vary across the site, but residual contamination hotspots are present in small areas. It may not be suitable for sites with large areas of distributed residual contamination near the regulatory limit.

4.1 Lag-k Applied to Scan MDC

4.1.1 Lag-k Approach to Estimating the Background

The lag-k approach is a hypothesis testing procedure for localized sources intended to account for variability in background. A site with non-localized contamination (i.e., residual contamination spread uniformly across the site) is not a suitable candidate for the lag-k approach.

Though the mean value of counts, $E[\mathbf{X}(s_i)]$, will vary in space across a decision unit because of naturally occurring background variation, within close spatial proximity of a given location, the distribution of background counts is assumed to be similar (the notation $E[\mathbf{X}]$ represents the expected value of the random variable $\mathbf{X}$). The lag-k method requires $E[\mathbf{X}(s_i)] = E[\mathbf{X}(s_j)]$ and $Var[\mathbf{X}(s_i)] = Var[\mathbf{X}(s_j)] = \sigma_B^2$ when $|i - j| \leq k$ if no contamination is present. This requirement means that local measured counts observed in “close” proximity can be assumed to have nearly the same underlying count distribution.

This formulation of the lag-k method assumes that data is collected along a transect via a detector moving at a uniform speed recording counts in uniform time intervals. The variables i and j index observations; the first observation, $X(s_1)$, is taken at location s_1. The variable k denotes the difference in the index between two observations. That means k is a measure of distance in index space. Because the detector is moving at a uniform speed and taking measurements at uniform time intervals, the real-world distance between two observation locations is determined by the distance in index space between the two measurements. Distance in index space is a linear transformation of distance in real-world space.

Limitations of this approach include that is appropriate for collecting data along a transect. Other methods could be considered that use location information if/when assessment of background variability in all directions is a concern. Additional methods are covered in PNNL Task 1b report (PNNL 2022) and include geospatial and geostatistical techniques such as heterogeneity assessments, variogram analyses in multiple directions. Generalized least squares regression; local indicator of spatial association (LISA), also known as the Local Moran’s I statistic; spatially explicit Bayesian regression models; few shot machine learning; variogram tomography should be considered when location data and/or data collected on additional variables can be used to model spatial heterogeneity in background and/or areas of concern. Further, kriging can provide a way to determine if and when variability can be detected using collected data. Kriging methods appropriate for variability in one or more direction include multi-Gaussian, generalized linear models, fixed rank kriging, Geospatial Extension to MARSSIM (GEM) using geostatistical simulation, Bayesian kriging, and artificial intelligence or machine learning. See references and details on these methods in Huckett et al. (2022). While that report is focused on 3D or subsurface applications, these same methods are applicable for variable background and area of concern, as these conditions are expected in the subsurface and were addressed therein.

4.1.2 Calculation of Detection Limit and MDCR

The critical level, detection limit, and MDCR can be derived using a hypothesis testing framework similar to the framework used for determining detection limits for static measurements (Currie 1968). In this section, the index of sensitivity (d') is not used in deriving the MDCR as is done in NUREG-1507 (NRC 2020) and MARSSIM (NRC/EPA/DOE 2000). This is due to both a choice in notation and an emphasis on classical hypothesis testing notation over the signal detection theory framework that emphasizes metrics of human performance. Inherent in the d' notation is the idea of a shared standard deviation between the distribution that describes background only observations and the distribution that describes background plus signal observations. These distributions do not actually have a shared standard deviation and we do not need to make such a simplifying assumption. The signal detection theory framework and notation are discussed in detail in Section 3.1.

Statistical hypothesis testing frameworks can be used for *a priori* calculations when the distributional family (e.g., normal) of the random variables of interest can be established *a priori*. This is because false-positive and false-negative error rates correspond to quantiles of the null and alternative hypothesis distributions and need to be calculated numerically to produce the scan MDC. As described in the previous section, the Poisson count distribution is often well-approximated by a Normal distribution; moreover, it is known that linear combinations of independent Normal distributions are also normally distributed.[1] This provides flexibility in how to define the random variable of interest for a hypothesis test while still maintaining an assumption of normality (e.g., this fact is used below in Equation 4.1, which can be expressed as a linear combination).

Let the statistic $\mathbf{D_k}$ be equal to the difference in measurement between an observation and the average of its neighbors along the transect (at lag distance k). A hypothesis test using $\mathbf{D_k}$ is derived from the distribution of $\mathbf{D_k}$ under the null and alternative hypotheses. Note that this is only possible in a context where scan measurements and their spatial locations are recorded so that both neighboring locations and neighboring values will be available for data analysis. This formulation is novel to the post-processing (i.e., without-surveyor vigilance) context.

The conceptual motivation for formulating a hypothesis test based on a lagged-difference distribution came from the traditional setting of a surveyor responding to audible output from the detector. One way to model the human thought process involved in scanning is to presume that the surveyor is responding to noticeable increases in audible detector response relative to the audible response a few steps before and/or a few steps after passing over a contaminated location. That is to say, the surveyor is responding to an increase in audible response relative to recent neighboring locations. The framework developed here can be thought of as a quantification of this type of surveyor decision process.

Consider the following hypotheses.

H_0: No net activity is present at location s.

H_A: Net activity is present at location s.

Fundamentally, the goal is to construct a net count distribution, which requires characterizing background values. Instead of using the average of measurements over a reference area, observations local to the sample location, s, are used while making the simplifying assumptions described above. These assumptions imply that observations at a distance k from the sample location can be used to characterize the appropriate background distribution for use characterizing the appropriate background distribution to decide whether net activity is present at location s.

The distance k should satisfy the requirements of preserving large-scale background variation and small-scale spatial autocorrelation. That is, k should be large enough that the lag-k neighbors are not also exposed to the hotspot contamination. At the same time, it should be small enough that the background distributions at location s and its lag-k neighbors are approximately the same.

Because of the nature of scan surveys, where constant speed is regulated but constant transect width is not as regulated or controlled, the spatial nature of the data can be simplified by considering only the distance along the scan transect instead of considering an omnidirectional distance. For pre-planning purposes, this simplification may be necessary so that the number of neighbors is known *a priori*, whereas for an omnidirectional distance, the number of neighbors

[1] If X and Y are two independent, normally distributed random variables with mean μ_1 and μ_2 and standard deviation σ_1 and σ_2, respectively, then the random variable $V=X+Y$ is normally distributed with mean equal to $\mu_1 + \mu_2$ and standard deviation $\sqrt{\sigma_1^2 + \sigma_2^2}$.

might not be known. Therefore, the statistical analysis only considers the two lag-k neighbors along the transect when characterizing the net count at location s.

The test statistic for the hypothesis test is based on the difference between the counts at a given location and the average of the counts observed at the two lag-k neighbors:

$$\mathbf{D_k}(s_i) = X(s_i) - \frac{X(s_{i-k}) + X(s_{i+k})}{2} \tag{4.1}$$

Assuming statistical independence among the count distribution $\mathbf{X}(\cdot)$, the mean of $\mathbf{D_k}(s_i)$ is given by

$$E[\mathbf{D_k}(s_i)] = E[\mathbf{X}(s_i)] - \frac{1}{2}(E[\mathbf{X}(s_{i-k})] + E[\mathbf{X}(s_{i+k})]) = \mu_i - \frac{1}{2}(\mu_{i-k} + \mu_{i+k}) = 0 \tag{4.2}$$

and the variance of $\mathbf{D_k}(s_i)$ is given by

$$Var[\mathbf{D_k}(s_i)] = Var[\mathbf{X}(s_i)] + (1/2)^2(Var[\mathbf{X}(s_{i-k})] + Var[\mathbf{X}(s_{i+k})]) = \frac{3}{2}\sigma_B^2. \tag{4.3}$$

When the Poisson count distributions have a sufficiently high count rate, they are approximately normally distributed, and thus, the distribution of the test statistic $\mathbf{D_k}$, under the null hypothesis, is normally distributed such that

$$\mathbf{D_k}(s_i) \sim N\left(0, \frac{3}{2}\sigma_B^2\right). \tag{4.4}$$

The critical level, L_c, is defined as the $1 - \alpha$ percentile of the $\mathbf{D_k}$ distribution,

$$L_c = k_\alpha \sqrt{\frac{3}{2}\sigma_B^2}, \tag{4.5}$$

where $\Pr\left(\mathbf{D_k} < k_\alpha \sqrt{\frac{3}{2}\sigma_B^2}\right) = 1 - \alpha$.

To derive a detection limit, L_D, consider the distribution of $\mathbf{D_k}(s_i)$ when contamination is present at location s_i. Let λ_S denote the count rate corresponding to point-source contamination at location s_i. The count rate distribution at location s_i has a Poisson distribution,

$$\mathbf{X}(s_i) \sim Poisson(\lambda_i + \lambda_S), \tag{4.6}$$

but the contribution of counts from the contamination source is assumed to be negligible for measured counts more than (or equal to) a distance k for location s_i. Under this assumption, the test statistic $\mathbf{D_k}$ is normally distributed with mean

$$E[\mathbf{D_k}(s_i)] = E[\mathbf{X}(s_i)] - \frac{1}{2}(E[\mathbf{X}(s_{i-k})] + E[\mathbf{X}(s_{i+k})]) = \lambda_i + \lambda_S - (\tfrac{1}{2})(\lambda_{i-k} + \lambda_{i+k}) = \lambda_S \tag{4.7}$$

and variance

$$Var[\mathbf{D_k}(s_i)] = Var[\mathbf{X}(s_i)] + (\tfrac{1}{2})^2(Var[\mathbf{X}(s_{i-k})] + Var[\mathbf{X}(s_{i+k})]) = \lambda_S + \frac{3}{2}\sigma_B^2. \tag{4.8}$$

The detection limit is defined to be the mean of the alternative hypothesis distribution such that the probability of failing to reject the null hypothesis will happen with probability β, the false-negative error rate. Because failing to reject the null hypothesis happens when the observed test statistic is less than the critical value L_c, the detection limit satisfies

$$\beta = Pr(\boldsymbol{D_k} < L_k) = Pr\left(\boldsymbol{D_k} < k_\alpha\sqrt{\tfrac{3}{2}\sigma_B^2}\right) \tag{4.9}$$

$$= Pr\left(\frac{\boldsymbol{D_k} - \lambda_S}{\sqrt{\lambda_S + \frac{3}{2}\sigma_B^2}} < \frac{k_\alpha\sqrt{\frac{3}{2}\sigma_B^2} - \lambda_S}{\sqrt{\lambda_S + \frac{3}{2}\sigma_B^2}}\right)$$

$$= Pr\left(\boldsymbol{Z} < \frac{k_\alpha\sqrt{\frac{3}{2}\sigma_B^2} - \lambda_S}{\sqrt{\lambda_S + \frac{3}{2}\sigma_B^2}}\right),$$

where $\mathbf{Z}$ has distribution $N(0,1)$. Because λ_S is the mean of the alternative hypothesis distribution, it follows that the detection limit is the value of λ_S, such that

$$\frac{k_\alpha\sqrt{\frac{3}{2}\sigma_B^2} - L_D}{\sqrt{L_D + \frac{3}{2}\sigma_B^2}} = -k_\beta, \tag{4.10}$$

where k_β is the $1-\beta$ percentile of the standard Normal distribution. By rearranging the previous formula to

$$\frac{L_D - k_\alpha\sqrt{\frac{3}{2}\sigma_B^2}}{k_\beta} = \sqrt{L_D + \tfrac{3}{2}\sigma_B^2}, \tag{4.11}$$

it is clear that this equation has the same form as that derived by Abelquist (2014, p. 231). In Abelquist's derivation, the notation σ_0 is used to denote the standard deviation of the null hypothesis distribution. Thus, by defining

$$\sigma^* = \sqrt{\tfrac{3}{2}\sigma_B^2} \tag{4.12}$$

and using Equation 9.6 from Abelquist, the detection limit is given by

$$L_D = k_\alpha\sigma^* + \frac{k_\beta^2}{2}\left(1 + \sqrt{\left(1 + \frac{2\sigma^*}{k_\beta}\right)^2 + \frac{4\sigma^*}{k_\beta}\left(\frac{k_\alpha}{k_\beta} - 1\right)}\right). \tag{4.13}$$

With the simplifying assumption that alpha and beta are both equal, this equation reduces to the simplified form

$$L_D = 2L_C + {k_\alpha}^2. \tag{4.14}$$

The detection limit can be in terms of counts or a rate. For scanning, typically a rate (e.g., cpm) is used. When the detection limit represents a rate, it is equivalent to the MDCR.

The formulas derived here are similar to those derived by Abelquist (2014) because both approaches are based on differences in count distributions. Abelquist notes that for paired observations of the background and sample, $\sigma_0 = \sqrt{2}s_B$, under the assumption that the standard deviation of the blank distribution and the background distribution are the same (where σ_0 is the standard deviation of the net count distribution under the null hypothesis). In the approach presented here, the standard deviation under the null hypothesis is $\sigma^* = \sqrt{\frac{3}{2}}\sigma_B$, which can be thought of as paired observations between the sample and the average of the two lag-k neighbors.

4.1.3 Conversion to Scan MDC

Once the MDCR has been calculated, it must be converted to a scan MDC by applying conversion factors to translate from the net count rate to the measurable surface activity or soil concentration. The steps and requirement for that conversion will be the same as those for the process described in Section 3.1.5. A detailed illustration of the entire process of calculating a scan MDC for planning a post-processed survey is included in Section 5.0.

For land area scan MDCs, MARSSIM (NRC/EPA/DOE 2000) recommends the use of MicroShield to model soil concentration and gamma exposure rate. This, together with the determination of the count rate to exposure rate, is used to convert the MDCR to the scan MDC. For GPS-based surveys, Alecksen (2016) presents an alternative method using the probabilistic Monte Carlo N-particle Extended (MCNPX) transport simulation code. This method does not assume that the contamination source is centered in the observation interval, which is more realistic for GPS-based surveys. Note that MCNPX is no longer being updated regularly. Starting with MCNP6, Los Alamos National Laboratory has put the alpha particle and heavy ion transport features into the MCNP6.x versions.

4.2 Advantages and Limitations of Approach

The primary advantage of this method of post-processing surveys is that it is robust relative to fluctuation in background levels. For areas with variable background, it can be difficult to identify reasonable reference units and summarize the background population with only a single median level and standard deviation. The nature of the comparison between differences allows the method to automatically adjust for local variation. This ensures that areas in which readings are slightly higher because of historical use or geological properties are not overly likely to trigger further investigation. It also improves the ability to detect and flag for follow-up the areas of elevated contamination within regions that feature lower than average background levels.

One limitation of this approach is the introduction of the additional lag parameter, k, that must be understood and determined by the end user to be used in the post-processing analysis. This parameter significantly affects the detection performance of this method. However, the expected hotspot size can be used for determining a meaningful physical basis for the lag size. The choice of the lag value depends on the following considerations, described in more detail in Section 4.4.

- k must be large enough that the lag-k neighbors are not also exposed to the hotspot contamination, given a hotspot exists at a particular location.

- k must not be so large that variability in background mean affects the analysis.

- k could also be determined based on observed temporal correlation structures.

A consequence of this robustness to variable background is that gradual increases in contamination levels across a site may not be detected during data analysis by this method. For example, if a site has a slowly increasing gradient of contamination in a north-south direction across the site, the local differencing will filter out that signal. However, this concern should not be overemphasized because scanning is aimed at identifying small areas of contamination, and large area persistent trends can be identified by visual inspection as part of the data analysis process. The implied conceptual model for a survey site is one in which background radiation levels vary across the site and contamination occurs mainly in small areas. If a site has large areas of distributed contamination consistently near the regulatory limit, this method will likely underestimate the scan MDC.

4.3 Example Calculations

This example investigates the effect of non-constant background on the probability of detecting a point source. Consider the case of a scan survey where measured counts are recorded at 1,000 evenly spaced locations along a transect. At each location, the number of counts is generated by a Poisson probability distribution, but the mean of the distribution is given by the function $f(x) = 1000 + A\sin(2\pi x/1000)$. The location along the transect, x, ranges between 0 (where the first measured count is recorded) and 1000 (where the final measured count is recorded). The units of distance are notional. The parameter, A, controls the amplitude of the sign function: when A = 0, the background mean is constant; when A is large, the background will have a peak and a valley. Figure 4.1 shows an example of a simulated process using A = 50.

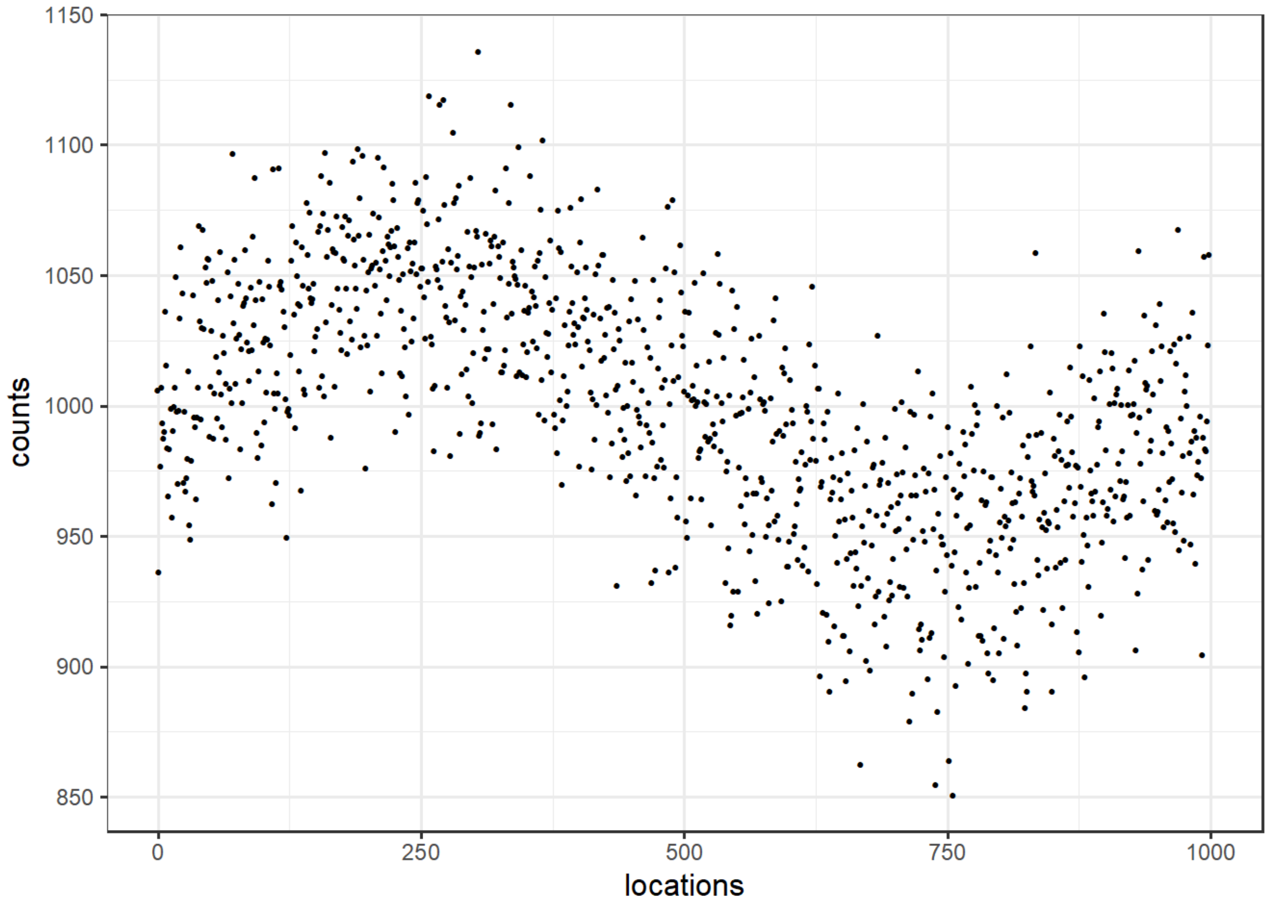

Figure 4.1. Simulated background counts along a transect using a background mean function $f(x) = 1000 + 50\sin(2 * pi * x/1000)$, where units of distance between locations are notional.

For comparison purposes, a non-localized approach as well as the difference in lag-k neighbor approach described in this report are performed on the same sets of simulated count data. The non-localized approach aggregates all background data and calculates a sitewide upper quantile to use as a critical value. For this example, it is clear that point-source contamination may be challenging to detect using such a non-localized method if point-source location coincides with the "valley" in background values. To illustrate this problem, consider a single point source whose location coincides with the lowest mean background level (location 750 in this example). In our simulation framework, the point-source contamination is achieved by adding another Poisson random variable, X, to the counts at location 750. The mean of the random variable, X, can be chosen to represent either low or high contamination sources.

The detection limit is defined to be the mean value that will result in contamination being detected a high percentage of the time, say 95%. The detection probability is estimated by computing 1,000 simulated runs for a given mean source count rate and computing the proportion runs with successful detection. This was done for point-source contamination count rate values ranging from 0 to 250 cpm. A lag of 10 was used for the lag-k difference method.

4.3.1 Calculating the Detection Limit

The detection limit is given by $L_D = 2L_C + k_\alpha{}^2$. For the non-localized method, the critical level is given by $L_c = k_\alpha\sqrt{2}s_B$, where s_B is the standard deviation estimated using reference area background measurements. For the lag-k difference method, the null hypothesis standard deviation σ^* is estimated using the observed lag-k differences from the reference area. Denote the sample standard deviation of the differenced values as s_B^*. Thus, for the lag-k difference method, the critical level is given by $L_c = k_\alpha s_B^*$.

To estimate the standard deviations in this example, a separate reference area dataset was generated and used to calculate s_B and s_B^*. The two cases are worked out in detail below to demonstrate the performance of the two methods using a constant mean background (Case 1) and a variable mean background (Case 2).

Case 1: Constant mean background: $A = 0$

A constant mean background is generated by setting the amplitude parameter, A, to zero, resulting in the mean function $f(x) = 1000$. An example simulated dataset is shown in Figure 4.2.

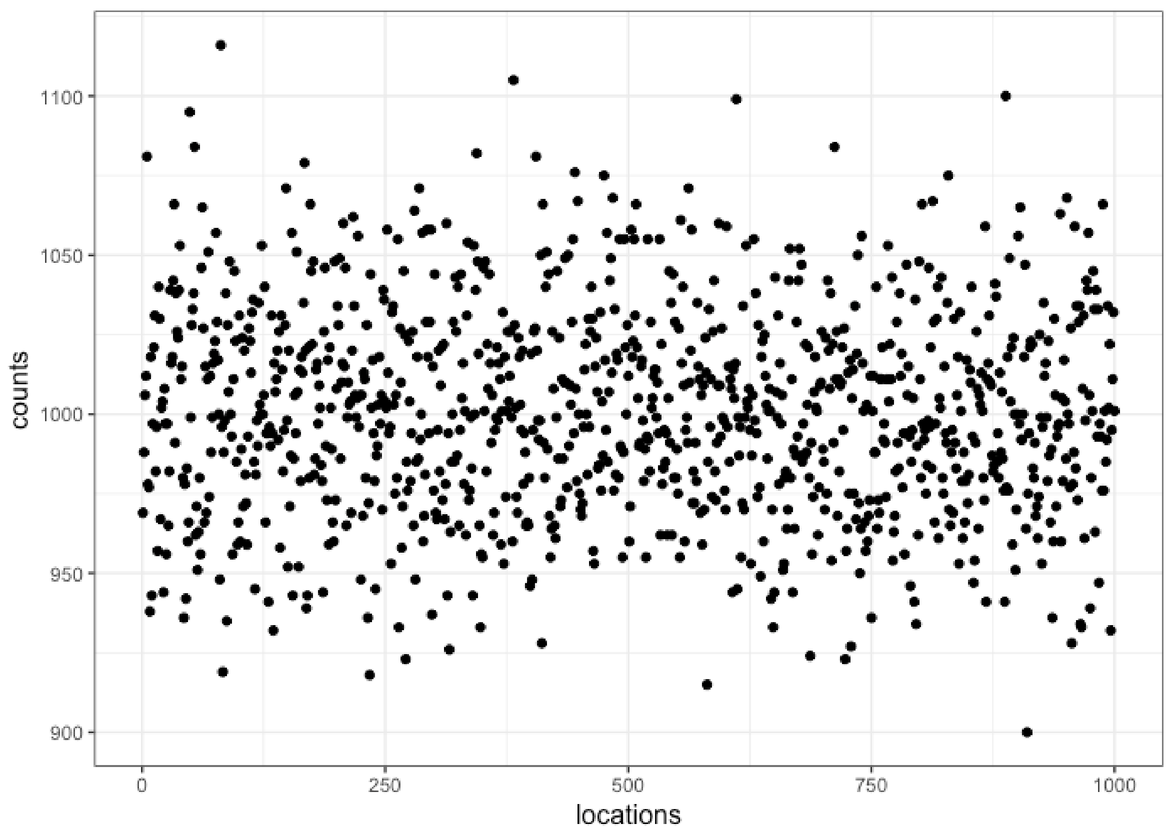

Figure 4.2. Example of simulated data using constant background.

Using a constant mean of 1000, the background standard deviation is equal to $\sqrt{1000} = 31.6$, and the lag-k distribution standard deviation is given by $s_B^* = \sqrt{3/2} * 31.6 = 38.7$. For the simulation study, these values are estimated empirically by generating a reference area dataset and computing the respective empirical standard deviations. For example, the dataset in Figure 4.3 results in the following values.

	$A = 0$
Average	999
s_B	30.5
s_B^*	35.6

These values are used to compute the critical levels and limits of detection as would be done in practice. For the non-localized method,

$$L_D = 2(k_\alpha \sqrt{2} s_B) + k_\alpha^2 = 2(1.645 * \sqrt{2} * 30.5) + (1.645)^2 = 144.6.$$

While for the lag-k method,

$$L_D = 2(k_\alpha s_B^*) + k_\alpha^2 = 2(1.645 * 35.6) + (1.645)\text{^}2 = 119.8.$$

We simulated 1,000 count datasets randomly for each of a range of mean point-source (representing hotspot) contamination values, ranging from 0 to 250 cpm. For each mean, the probability of detection was determined based on the proportion of times the hotspot was detected by each hypothesis test (one non-localized hypothesis test and one lag-k hypothesis test was conducted for each simulated dataset). For example, consider a mean hotspot contamination level of 50 cpm. 1,000 datasets were simulated based on random draws from a Poisson distribution with mean set to 50 cpm and then, for each data set, non-localized hypothesis test and one lag-k hypothesis test was conducted and resulted in a true positive or false negative detection of the hotspot. The non-localized hypothesis test resulted in 250 out of 1,000 true positive detections, or a detection rate of 0.25, and the lag-k hypothesis test resulted in a little over 375 out of 1,000 true positive detections, or a detection rate of 0.375. These points are shown in Figure 4.3 where 50 cpm on the x-axis intersect the curves at 0.25 and 0.375 on the y-axis. The 1,000 simulations and hypothesis tests were repeated for each mean on the x-axis to investigate how detection rates for each hypothesis testing method changed as the mean contamination level varied between 0 cpm and 250 cpm, represented by the curves in Figure 4.3.

Lag-k hypothesis testing resulted in slightly higher detection rates than the non-localized method, shown by the black line starting and staying higher than the red line in Figure 4.3. Although, the detection rates vary only by roughly 12.5% at the most and get closer as the mean hotspot contamination increases to over 100 cpm or so. Further, Figure 4.3 shows that both methods result in roughly the nominal error rate ($\alpha = 0.05$) at their limits of detection, $L_D = 144.6$ and $L_D = 119.8$. This is shown where the vertical dashed lines at each limit of detection on the x-axis intersect the curves near $1 - \alpha = 1 - 0.05 = 0.95$, or 95% probability of detection. This exercise demonstrates that when the background distribution has a constant mean, these two hypothesis testing methods perform similarly and have probabilities of detection close to the nominal levels. In the next case, we perform a similar exercise but when the background mean varies rather than remains constant.

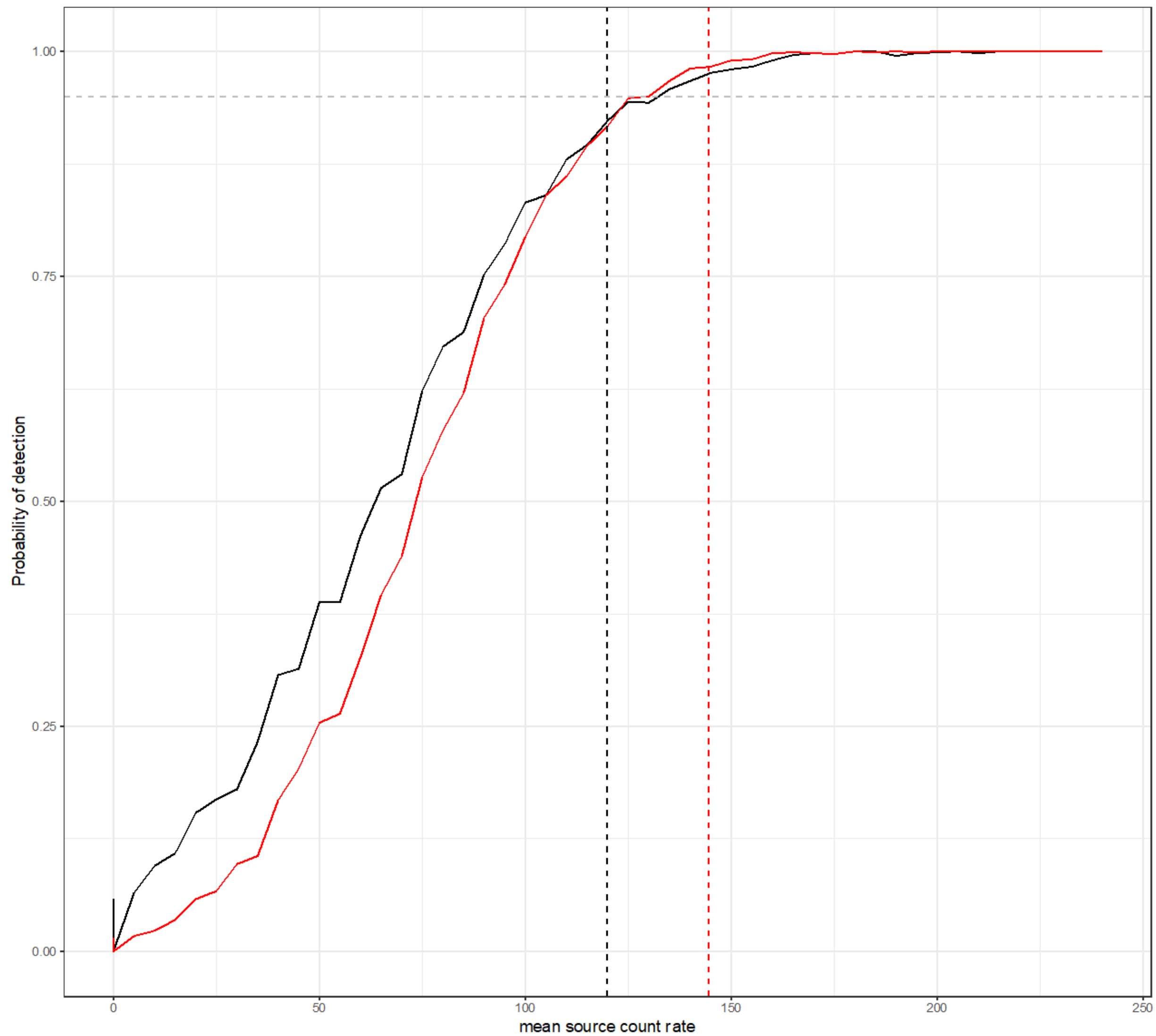

Figure 4.3. Probability of detecting a source term as a function of mean source count rate. Solid lines represent calculated probabilities of detection using the non-localized hypothesis test (red) and lag-k hypothesis method (black). Each point on the line represents the observed probability of detection (y-axis) for a given mean source count rate (x-axis). The vertical dashed lines are the calculated detection limit for each method. These results are based on 1,000 simulations from a process with non-constant background ($A = 0$).

Case 2: Use a variable background mean by setting the amplitude parameter: A = 50

To simulate background with variable mean, the amplitude parameter A is set to a value of 50. This results in simulated datasets similar to those discussed in Case 1, but the relationship between the two standard deviations cannot be established using the formula because they do not have constant means. Using the simulated data, the estimated standard deviations and detection limits are given in the following table.

Table 1. Calculated standard deviation and critical value for the lag-k method (k = 5, 10, 20) when amplitude A = 50.

Statistic	**Value**
Average	1000
s_B	48.8
s_B^*	38.6

For the non-localized method,

$$L_D = 2(k_\alpha\sqrt{2}s_B) + k_\alpha^2 = 2(1.645 * \sqrt{2} * 48.8) + (1.645)^2 = 229.7.$$

For the lag-k method,

$$L_D = 2(k_\alpha s_B^*) + k_\alpha^2 = 2(1.645 * 38.6) + (1.645)^2 = 129.7.$$

The difference between the detection limits is much greater in this case than the previous because variability due to the non-constant background is filtered out by the differencing used in the lag-k method but is absorbed in the non-localized standard deviation estimate. Repeating the simulations and hypothesis testing on the simulated datasets for each mean value, as described in Case 1, produces the probabilities of detection shown in Figure 4.4.

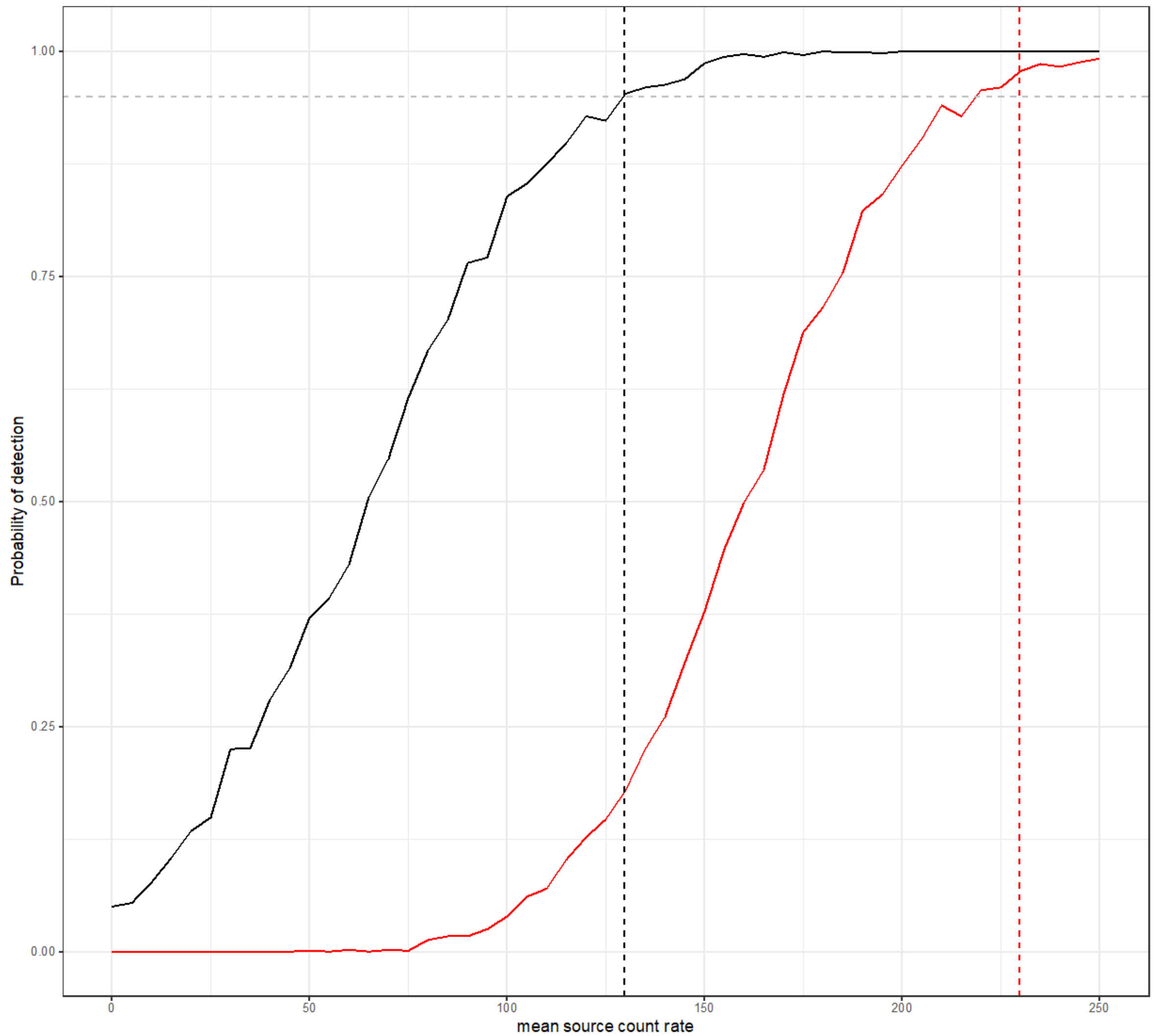

Figure 4.4. Probability of detecting a point source located at the lowest mean background level (location 750 on Figure 4.1) as a function of mean source count rate with variable background. Solid lines represent calculated probabilities of detection using the non-localized hypothesis test (red) and lag-k hypothesis method (black). Each point on the line represents the observed probability of detection (y-axis) for a given mean source count rate (x-axis). The vertical dashed lines are the calculated detection limit for each method. These results are based on 1,000 simulations from a process with non-constant background ($A = 50$).

This example highlights a problem with using a non-localized approach and single reference area mean to account for background at every location in the decision unit. When background values vary considerably across a decision unit, detection limits (and hence MDCs) are large because the single reference area standard deviation captures the variability of the trend in background mean over the decision unit. In this case, the proposed lag-k difference method is robust relative to such fluctuations in background, which results in much lower detection limits when the background mean in the decision unit is not approximately constant.

4.4 Implementation Considerations

Suppose an NRC licensee is tasked with planning a scan survey for a land area and the licensee wants to use the lag-k difference method to estimate the scan MDC. The primary parameter needed for planning is an estimate of the lag-k difference standard deviation s_B^*. This is best achieved by performing some initial scans along transects in an appropriate reference area so that the lag-k difference can be computed for this dataset and the resulting differences can be used to calculate the standard deviation. If scan values are not available but the background reference area scan standard deviation has been estimated, then the relationship $s_B^* = \sqrt{\frac{3}{2} s_B^2}$ can be used to provide an estimate of the lag-k difference standard deviation. Note that knowledge of the background count rate, and hence background count standard deviation, is needed in the non-localized method, so the proposed method can be used without any additional sampling effort from the user.

The choice of the lag value depends on a few considerations. The primary consideration is the assumption that if hotspot contamination exists at a particular location, then k should be large enough that the lag-k neighbors are not also exposed to the hotspot contamination. In this study, when investigating whether a hotspot exists or not, the size of the hotspot is assumed to be small (e.g., $0.25\ \mathrm{m}^2$ areal dimension for land area scans and $100\ \mathrm{cm}^2$ for building surface scans); thus, the value of k does not need to be large. In general, the expectations about hotspot size should be determined and documented as part of the DQO process. The chosen lag value should not be so large that variability in background mean values affects the analysis. That is, the expectation is that background values are approximately constant within distance k of a measured location. If the actual hotspot size is larger than the assumed hotspot size, it is more likely to be identified during post-processing analysis of the observed data, though simulation studies using larger hotspot areas would need to be conducted to quantify the detection performance. As the size of the hotspot increases, the likelihood of detecting the hot spot during sampling or direct measurement also increases, although the primary purpose of sampling and direct measurement is to calculate area-wide concentration for comparison to area-wide DCGL.

To clarify the concept of large-scale background variation and small-scale spatial autocorrelation, consider a simple statistical model for a process with the large-scale trend and spatial dependence given by $Y = m(x) + e(x)$, where the error term e has mean 0 and variance-covariance matrix E with elements $E[i,j]$ that are a function of the distance between point locations, $d(si, sj)$. For the statistical independence assumptions in the scan MDC model to be valid, $m(x)$ must be smooth in the sense that it is approximately constant locally and the small-scale spatial dependence decays to a negligible level quickly (<10–20 m).

Another approach to determining the value of k would be to use the observed temporal correlation structure in the data. If a scan survey is conducted in a reference area, then an autocorrelation plot can be used to display how the correlation decays as a function of lag distance. The value of k can be chosen such that correlations beyond this distance are negligible, thereby satisfying the statistical independence assumption in the model.

Once the lag-k is chosen and s_B^* is estimated using reference area scan survey data (or scoping survey scan survey data, conducted *a priori*), the MDCR can be calculated. No adjustment needs to be made for surveyor efficiency because no surveyor judgment is being used in this method.

In site data analysis, reference area data might not be very representative of background values within the decision unit. In this case, it might be better to use portions of the decision unit data to characterize the background and inform investigation levels. This can be done even for areas in which some contamination is present within a decision unit because some portions of the decision unit will be uncontaminated and data from these areas can be used to characterize background distribution values (King and Vitkus 2015).

5.0 Case Studies

5.1 Case Study 1

This case study illustrates the application of the proposed alternative scan MDC methodology to plan a scan survey for a large land area using a NaI scintillation detector without surveyor vigilance. Scan data collected in two different reference areas and within the decision unit were provided by the NRC and used in the calculation of the MDCR. MicroShield was used to convert the MDCR to the scan MDC. In this example, the reference area data are used to calculate the *a priori* scan MDC for planning purposes. Section 5.2 demonstrates how a critical level based on reference area data can be used to flag locations that may require further investigation using the lag-k method.

The case study includes two reference areas as well as a large site area that is subdivided into sub areas with high density scans. The reference areas are shown in Figures 5.1 and 5.2. Scan paths in the reference areas are single, closed-loop paths rather than traditional transects. Reference Area 2 has two such paths, Reference Area 1 has one. The site survey map is shown in Figure 5.3 and is composed of sub-areas based on a triangular grid within which traditional back-and-forth transects are traversed and connected by a single scan path. The discussion about identifying locally elevated areas focuses on the sub-area identified within the circle in Figure 5.3 rather than the entire site. Note from the scale in Figure 5.3 that this sub-area is among those with the highest observed counts.

Histograms of the reference and site areas are shown in Figure 5.4. The representativeness of the reference areas for characterizing site background is questionable. Note that the reference area distributions are shifted right of the site area, with Reference Area 1 not only being shifted relative to Reference Area 2 but also having a larger standard deviation. If the critical level were determined by simply using an upper quantile of the reference area(s) distribution, it is unlikely that any points within the site area would be flagged as elevated. However, the lag-k approach does identify areas within the site sub-area that may need further investigation, as they indicate areas that exceed local background.

The case study is not a site assessment—the site conceptual model has not been formulated, DCGLs were not calculated, and post-processing of site data was not completed. In a site assessment, the site data would be post-processed following MARSSIM guidance. The case study is simply intended to demonstrate how to calculate the *a priori* scan MDC for planning purposes using the lag-k method to ensure scanning will identify elevated areas relative to background within the prescribed error rates.

Additionally, the case study demonstrates that the lag-k method is somewhat invariant to the selection of k for this particular study and is therefore promising as a site assessment tool. Future work should include a comprehensive study of the lag-k method for post-processing by (1) applying it to a variety of simulated and real site contamination scenarios to determine the importance of k selection in general and (2) comparing it with current MARSSIM approaches as well as other methods, such as geostatistical analysis, that are being used for site assessments.

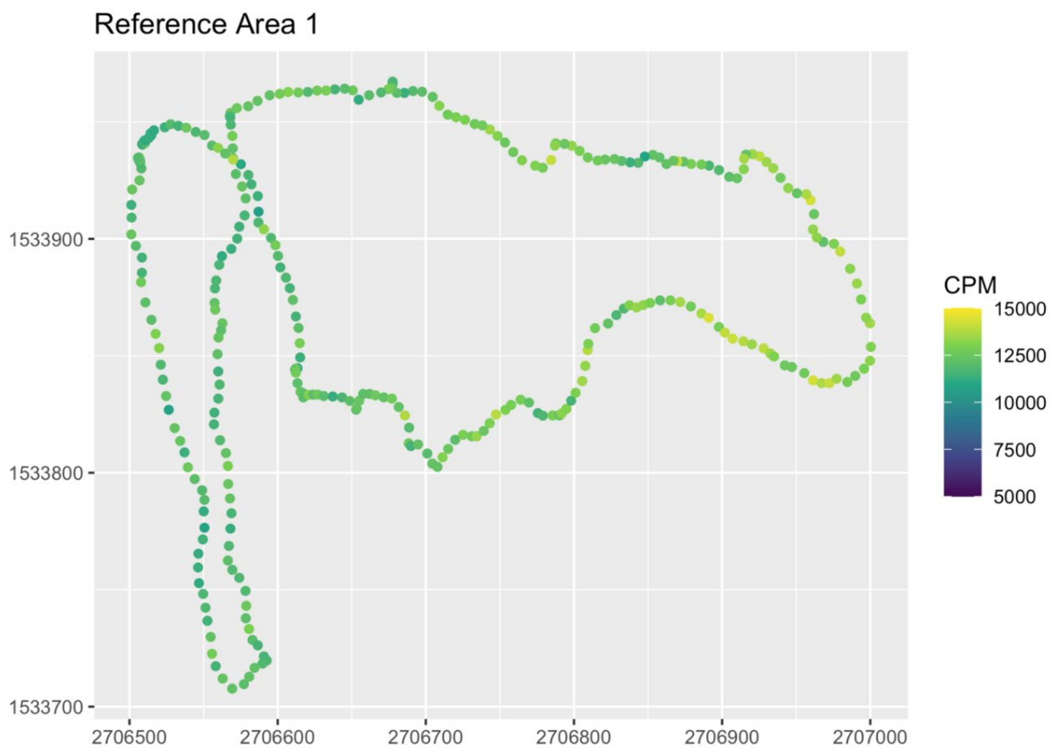

Figure 5.1. Scan survey data from Reference Area 1 with easting on the x-axis and northing on the y-axis.

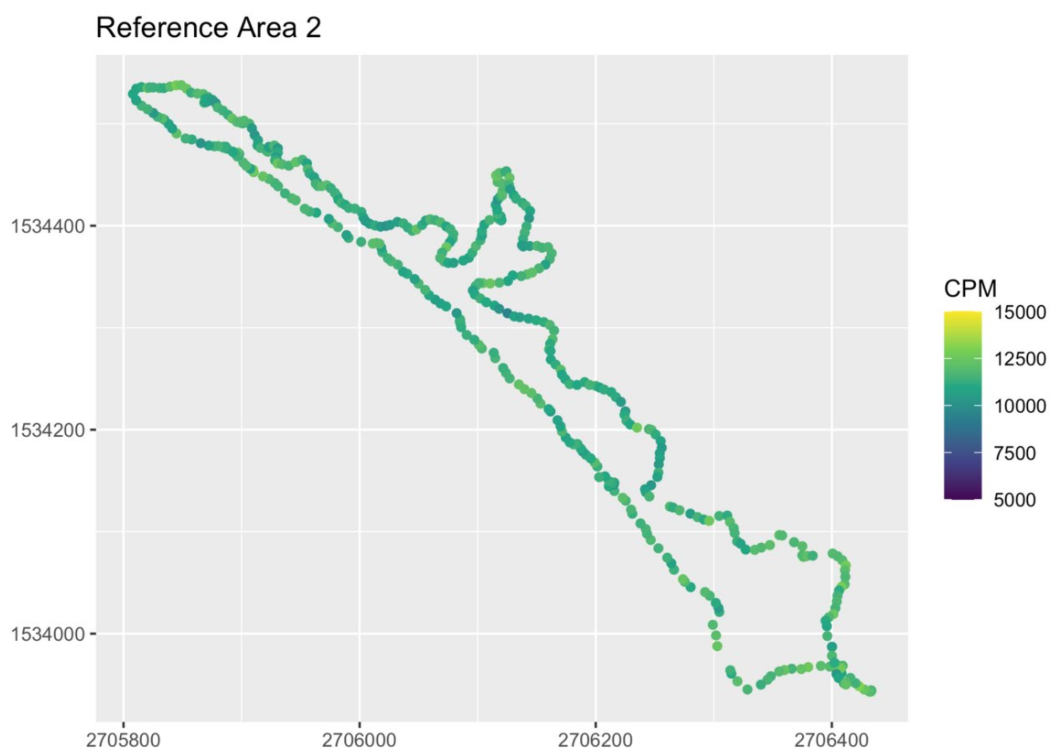

Figure 5.2. Scan survey data from Reference Area 2 with easting on the x-axis and northing on the y-axis.

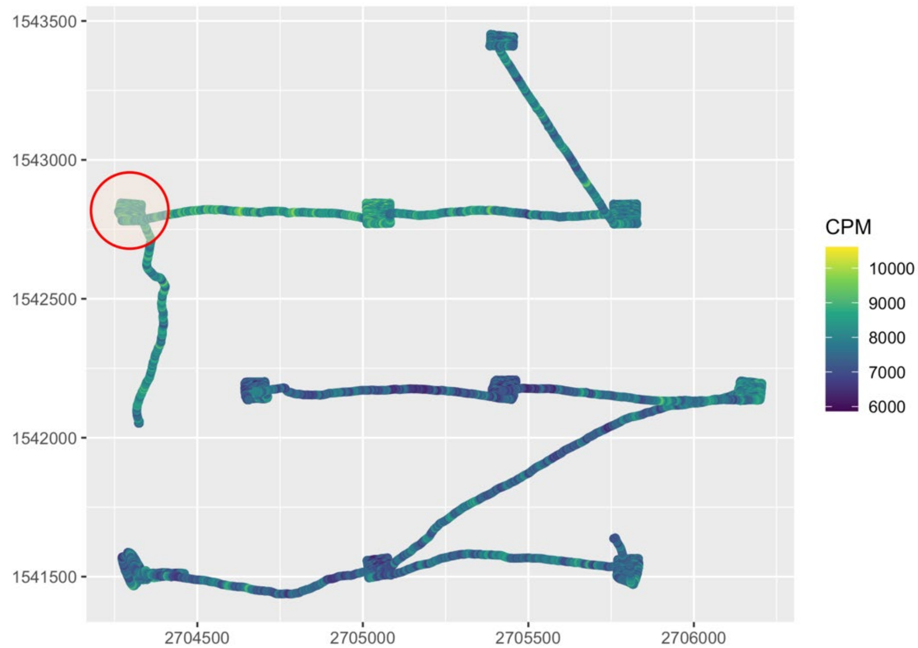

Figure 5.3. Survey unit scan data (in cpm) with easting on the x-axis and northing on the y-axis. The area within the red circle is the sub-area shown in Figure 5.6.

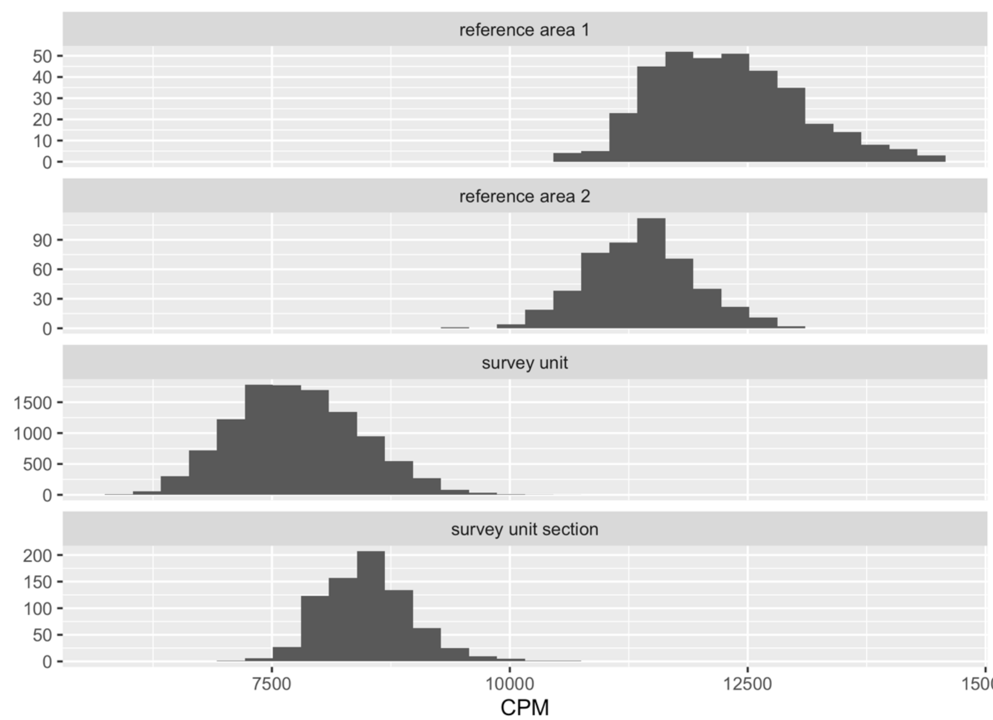

Figure 5.4. Histograms of survey unit and reference area data (with counts on the y-axis and cpm on the x-axis). Reference Areas 1 and 2 correspond to Figures 5.1and 5.2, respectively. The survey unit and survey unit sections correspond to Figures 5.3 and 5.6, respectively. Note that the reference areas are shifted right of the survey data. Reference area 1 has a higher standard deviation than any other area.

5.1.1 Application of the Proposed Alternative *A Priori* Scan MDC Methodology for Planning Purposes

For this case study, assume that the surveyor used a GPS-based logger capturing scan measurements at 1-second intervals as well as spatial coordinates on the same interval. Site scan locations and values for reference areas are based on an actual NRC-regulated site.

The acceptable Type I and Type II error rates are specified to be 0.05 each, and the hotspot size of concern is assumed to be a circular footprint of 0.25 m^2 with uniform contamination 15 cm deep in the soil. The detector used is a 1.5 in. by 1.25 in. NaI scintillation detector (Victoreen Model 489-55), and the reference area survey scans are available for calculating the *a priori* scan MDC.

Similar to the steps outlined in Section 4.1, the steps to calculate scan MDCs for land areas are:

1. Calculate the MDCR (in cpm) for a given background and performance level (Type I and Type II decision errors) using Equation 4.14.
2. Translate the MDCR to minimum detectable exposure rate using the relationship of net count rate to net exposure rate for a particular survey instrument.
3. Translate the minimum detectable exposure rate to scan MDC using a model such as MicroShield for specific conditions.

Note that, because the is survey to be conducted with no real-time surveyor vigilance, the MDCR is not adjusted for surveyor efficiency.

Histograms of the two reference areas are shown in Figure 5.4, and the scan survey data from the reference areas are shown in Figure 5.1 and Figure 5.2. Scan data for the survey unit are presented in Section 5.1.4.

For this example, the fact that data from two reference areas are available means that either Reference Area 1, Reference Area 2, or a combination of both could be used to characterize background for planning purposes. To simplify this example, only data from Reference Area 1 are used to characterize background activity for survey planning purposes because the greater variability in this reference area will result in a conservative estimate of the MDCR.

5.1.2 Calculating the MDCR

The MDCR can be calculated using Equation 5.1:

$$MDCR = 2(k_\alpha s_B^*) + k_\alpha^2 \tag{5.1}$$

where

s_B^* = the standard deviation of the expected distribution of lag-k differences;
k_α = the $1 - \alpha$ percentile of the standard Normal distribution; and
α = the desired Type I and Type II error rate (when they are assumed equal).

With $\alpha = 0.05$, k_α is equal to 1.645. To obtain s_B^*, the prior data from the reference area can be evaluated to determine the value for k, reflecting the distance that should be used for calculating the localized differences, and to estimate an appropriate value for the standard deviation of the lag-k difference distribution. Considerations for selecting k are outlined in

Section 4.4. For this case study, a value of k = 20 is used, resulting in a lag-k difference distribution with a standard deviation equal to 758 cpm (Figure 5.5).

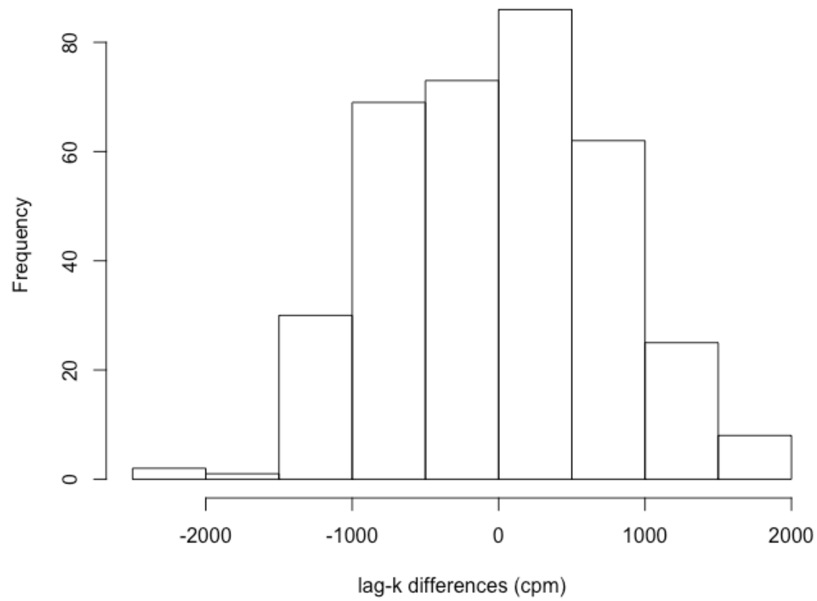

Figure 5.5. Histogram of the lag-k differences using k = 20 for the background reference area data.

Now that the necessary inputs have been defined, the MDCR is calculated as

$$\text{MDCR} = 2(k_\alpha s_B^*) + k_\alpha^2 = 2(1.645 * 758) + (1.645)\hat{}2 = 2496.5 \text{ cpm.} \tag{5.2}$$

5.1.3 Translating MDCR to MDC

The MDCR can be converted to a minimum detectable exposure rate, and the exposure rate can be converted to the scan MDC. To perform this conversion, it is necessary to specify the characteristics of the instrument, the hotspot, and the soil. For the 1.5 in. by 1.25 in. NaI scintillation detector, the ratio of the counting rate to the exposure rate specified by the detector manufacturer is 350 cpm/uR/h for Cs-137. Thus, the minimum detectable exposure rate is given by 2496.5/350 = 7.13 uR/h.

The specified hotspot of concern has a circular footprint size of 0.25 m^2 and uniform contamination down to 15 cm in the soil. A scan rate of 0.5 m/s will provide an observation interval of approximately 1 second. Using these parameters, the MicroShield modeling code determined an exposure rate of 1.31 uR/h based on an arbitrary concentration of 5 pCi/g.

Using these values, the scan MDC can be calculated as follows:

$$\text{Scan MDC} = (5 \text{ pCi/g})*(7.13/1.31) = 27.2 \text{ pCi/g} \tag{5.3}$$

This value can be compared with scan MDC requirements (e.g., $DCGL_{EMC}$) calculated based on dose modeling consistent with regulatory guidance to make sure the equipment and survey technique are adequate for detecting the radiation levels of concern.

5.1.4 Apply Lag-k to Post-Processed Data Collected Without Vigilance

Now, assume the survey was conducted and lag-k was used to do the hypothesis testing. The survey unit scan data are analyzed using the lag-k method to flag locations that exceed the critical level calculated using background reference area data. Again, this example is focused on the sub-area identified in the circle in Figure 5.3, shown in higher resolution in Figure 5.6 below to represent the entire scan footprint within that specific survey unit. Recall the histograms in Figure 5.4 showed that the reference areas tended to have higher counts than the survey unit, potentially of concern when using them to characterize background. However, the lag-k method characterizes the distribution of local differences within a survey area which does not rely on the background measurements from reference areas, pointing to an advantage of the lag-k method.

The lag-k ($k = 20$) histogram distributions for each of the survey areas are shown in Figure 5.7. The means of these distributions are centered at zero and the lag-k reference area distributions can be used to derive critical values (Equation (4.5)) to be used to compare against the lag-k values within the survey unit. Results are shown in Figure 5.8, Figure 5.9, and Figure 5.10 for scenarios using Reference Area 1 to characterize background using the lag-k method (with k = 5, 10, 20) and using Reference Area 2 to characterize background using the lag-k method (k = 5, 10, 20). Table 2 shows the estimated standard deviation and critical value of the lag-k distributions for each scenario. Note that the values in this table are based on the distribution of lag-k differences; therefore, the units are in cpm. However, the critical values represent a difference between an observation and its neighboring values not the observation itself. Thus, the critical value equal to 1247.43 cpm in Table 2 corresponding to Reference Area 1 using k = 20 has the following interpretation: any location in the survey unit whose value is greater than 1247.43 cpm more than the average of its two lag-20 neighbors is flagged as red in the corresponding plot in Figure 5.10.

The consistency of flagged locations across choices of k indicate that analysis is somewhat insensitive to this choice and that reasonable choices of k based on the conceptual site model should be sufficient, though further work on data-driven choices of k would be useful to users.

Figure 5.11 identifies the areas in Figure 5.6 that exceed local background using a lag-k with k = 20 and either Reference Area 1 or Reference Area 2 critical values (last row in Table 2). The reference areas return similar elevated values, with Reference Area 2 identifying two additional values. In a site assessment, these areas could be identified as needing additional investigation.

It is advisable to analyze data using multiple values similar to what is shown here to help distinguish between locations of concern (which should be flagged for most values of k) from noise/false positives (which may be indicated by only being flagged for a small proportion of the values of k used in the analysis). Future work will investigate the possibility of flagging values based on integrating over a reasonable range of values for k in order to both remove the requirement for a user-specified k and reduce the false positive rate. Future work should study differences in average values of points within a certain distance from each observation and the observation and compare that to only using the two points that distance before and after it.

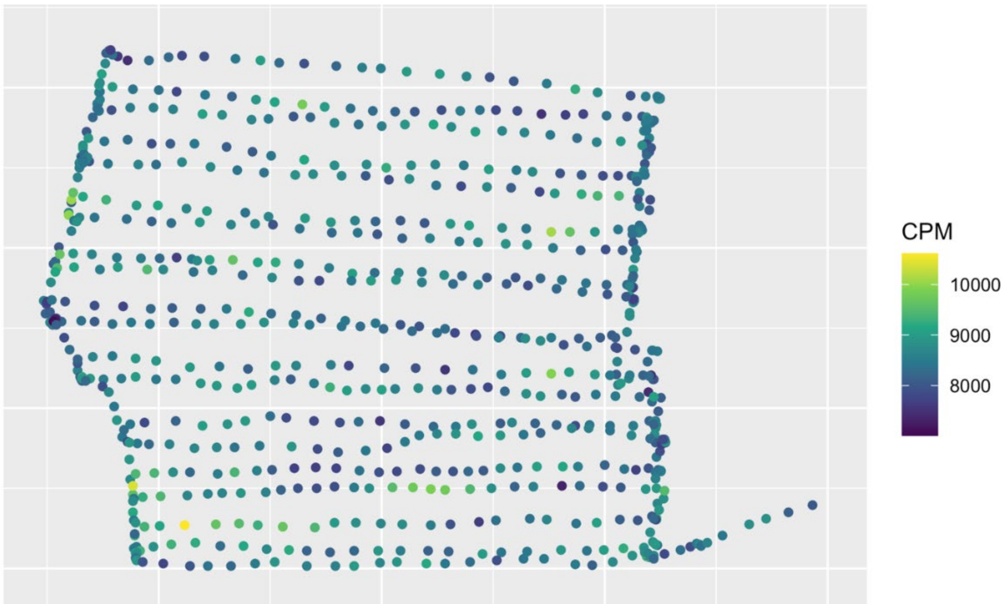

Figure 5.6. One continuously scanned section of the survey unit data (in cpm) with easting on the x-axis and northing on the y-axis. This portion of the survey unit is located in the upper-left region of the survey unit within the red circle in Figure 5.3.

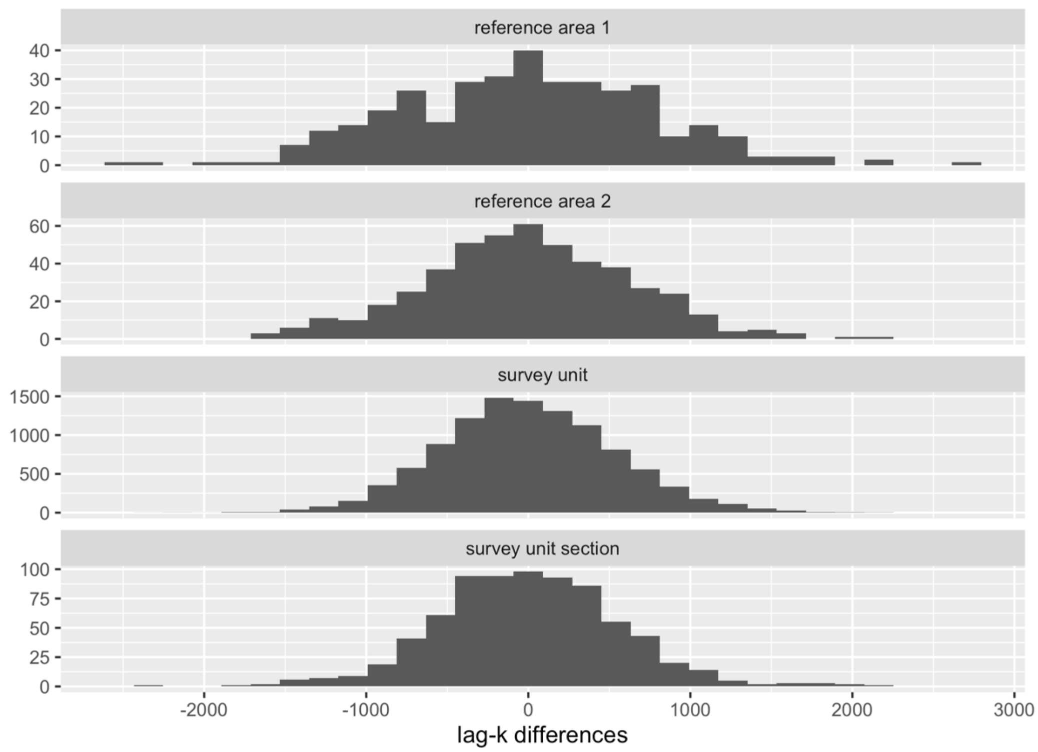

Figure 5.7. Histograms of the lag-k (k = 20) differences from background reference area and survey unit. Reference Areas 1 and 2 correspond to Figures 5.1 and 5.2, respectively. The survey unit and survey unit sections correspond to Figures 5.3 and 5.6, respectively. Note that the histograms are centered at zero, regardless of the distribution location (e.g., mean) in Figures 5.1 through 5.6.

Table 2. Calculated standard deviation and critical value for the lag-k method (k = 5, 10, 20) using reference area one data and reference area two data. Note that the values in this table are based on the distribution of lag-k differences; therefore, the units are in cpm. However, the critical values represent a difference between an observation and its neighboring values not an observed count rate itself.

	Ref Area 1: s_B^*	Ref Area 1: L_c	Ref Area 2: s_B^*	Ref Area 2: L_c
Lag 5	774.27	1273.56	635.72	1045.67
Lag 10	697.98	1148.07	692.55	1139.15
Lag 20	758.4	1247.43	735.62	1209.99

Figure 5.8. Flagged locations (red) indicate observations exceeding the critical level for the reference area 1 (top) and 2 (bottom) and lag-k (k =5) value, with easting on the x-axis and northing on the y-axis. The critical values used for each plot are listed in Table 2.

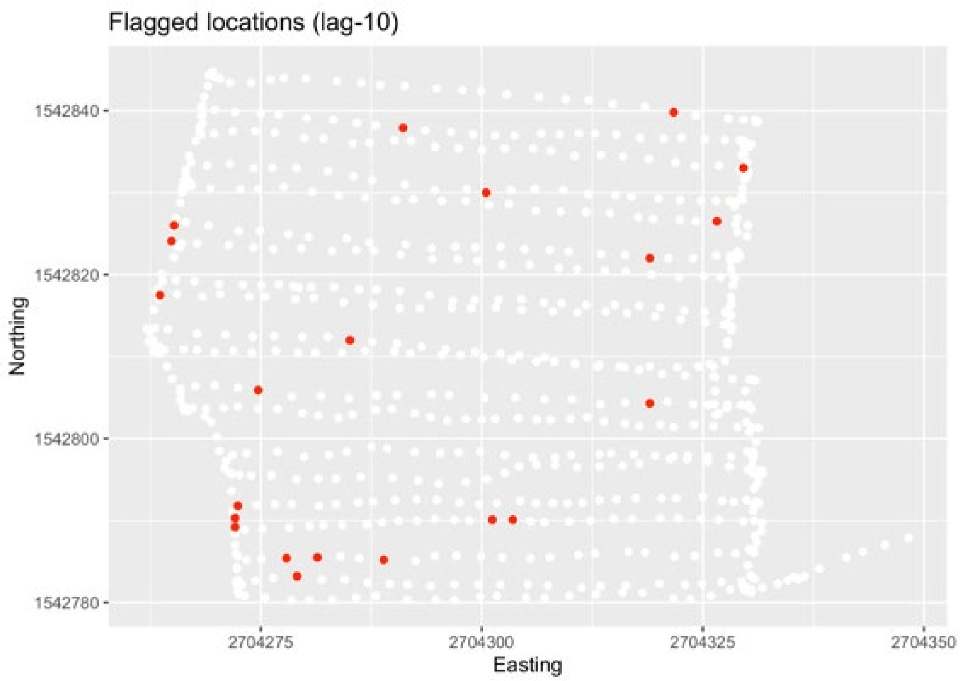

Figure 5.9. Flagged locations (red) indicate observations exceeding the critical level for the reference area 1 (top) and 2 (bottom) and lag-k (k =10) value, with easting on the x-axis and northing on the y-axis. The critical values used for each plot are listed in Table 2.

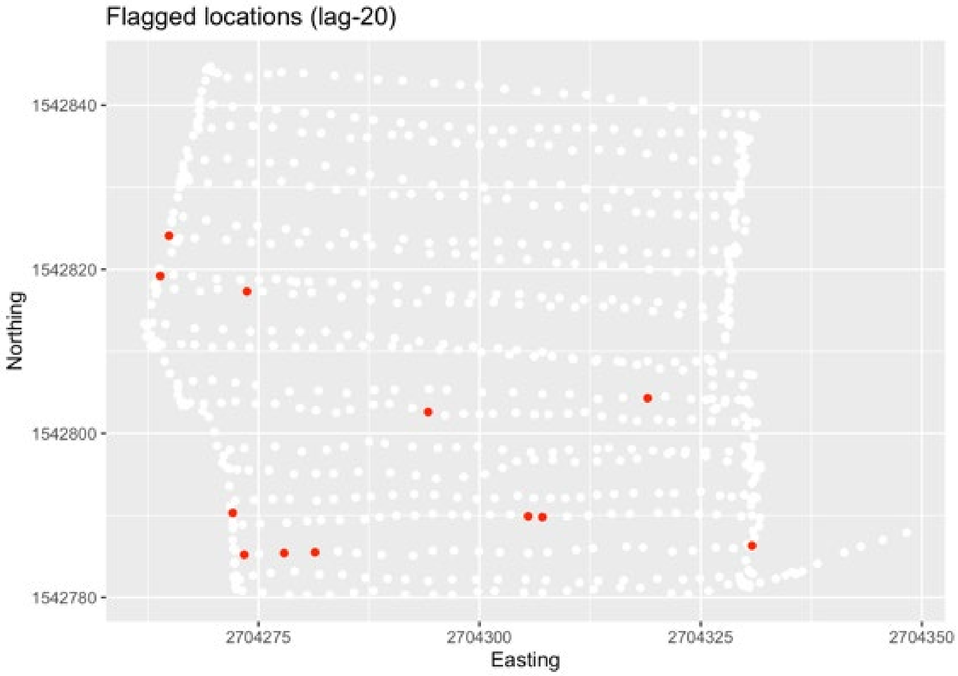

Figure 5.10. Flagged locations (red) indicate observations exceeding the critical level for the reference area 1 (top) and 2 (bottom) and lag-k (k =20) value, with easting on the x-axis and northing on the y-axis. The critical values used for each plot are listed in Table 2.

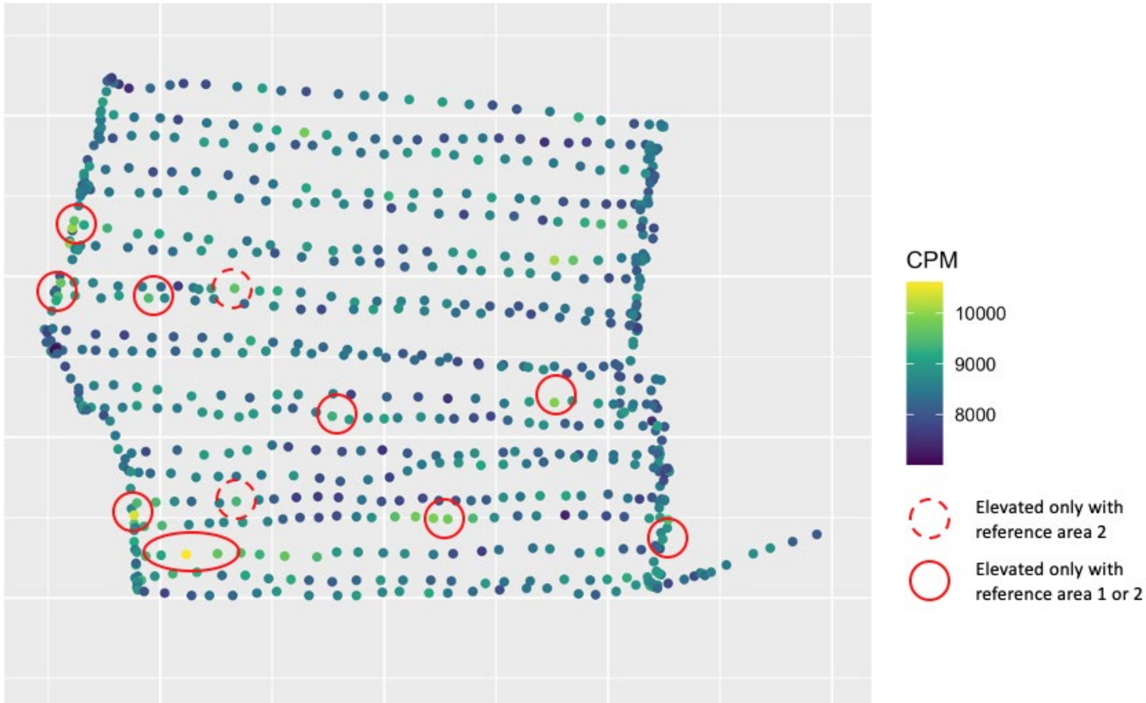

Figure 5.11. Sub-area of survey unit with easting on the x-axis and northing on the y-axis and points identified as greater than local background using lag-k ($k = 20$) identified in the red circles (also shown in Figure 5.10). Solid red circles indicate points flagged using Reference Area 1 or 2. Dashed red circles indicate points flagged using Reference Area 2 but not Reference Area 1.

5.2 Case Study 2: Evaluating Competing Analysis Methods

Existing frameworks, including the lag-k method, perform hypothesis tests on an observation-by-observation basis. Each observation is either flagged for follow-up investigation or not. When the true nature of each observation is known, these types of binary classification procedures are well described with a two-by-two confusion table. A confusion table provides the necessary information to calculate false positive and false negative rates. The false positive and false negative rates are key metrics that should be used to evaluate analysis methods (positive predictive power and negative predictive power may also be of interest).

Table 3. Notional confusion table.

	Predicted Positive	Predicted Negative	Total Actual
Actual Positives	True Positives=68	False Negative=17	Positives=85
Actual Negatives	False Positives=145	True Negatives=253	Negatives=398
Total Predicted	Positives=213	Negatives=270	

To create a dataset with known signal, we started with the observed counts shown in Figure 5.2. Additional counts were added to 85 observations toward the north of the reference area. The size of these signals ranged between 25 additional cpm and 10,000 additional cpm.

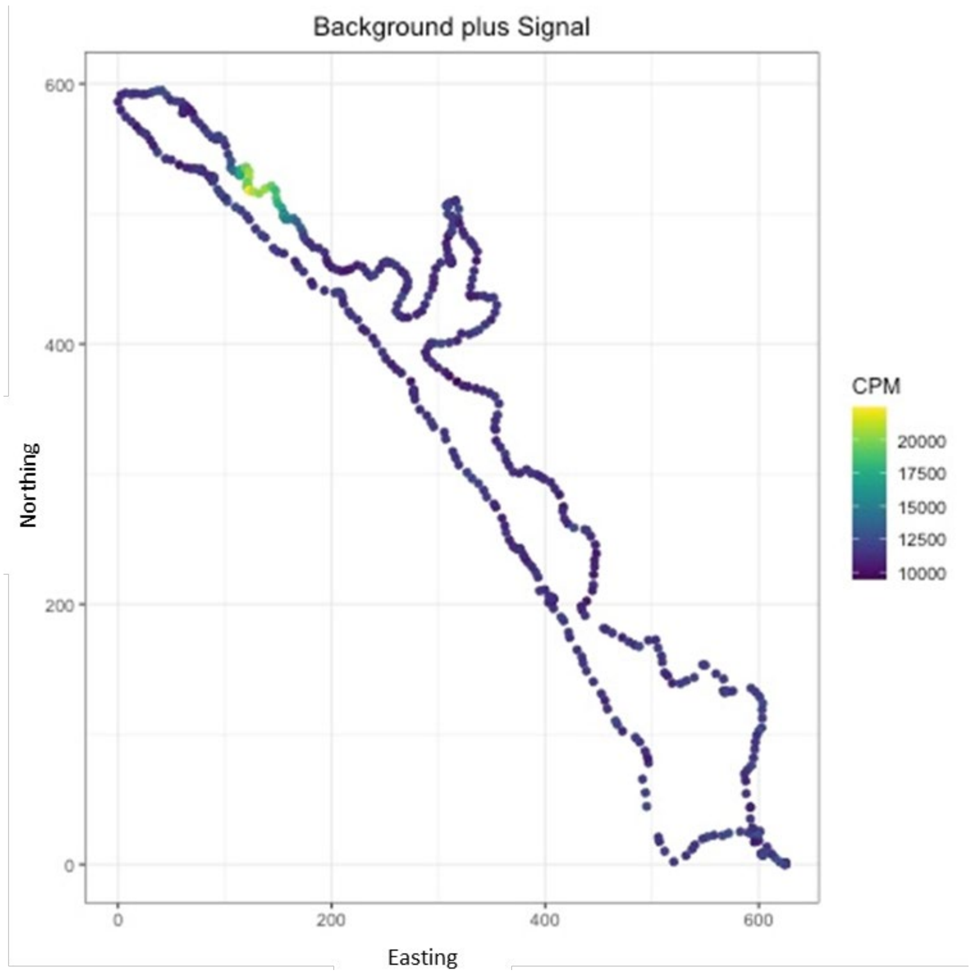

Figure 5.12. Heatmap of observed counts per minute with known inserted signal for Reference Area 2, with easting on the x-axis and northing on the y-axis.

When calculating an *a priori* critical level, it is important to base that calculation on the actual counts associated with a single observation. For this example, suppose our best *a priori* estimate of the background count rate is 12,000 cpm and suppose the individual observations are counts obtained over a full minute long time window (actual time windows are much shorter—this assumption keeps the bookkeeping math to a minimum). Because of the Poisson assumption, we expect a standard deviation of 110 counts. The critical level calculation depends on α, the acceptable probability of false positives, and N, the number of observations that will be used to estimate the background. The lag-k approach uses N=2. If we set $\alpha = 0.05$, we can calculate the critical level, $L_c = \sqrt{(1+1/N)}\sigma_B k_{0.05} = \sqrt{3/2} \times 110 \times 1.645 = \ 222$ net counts per 60 second time window.

For this example, a lag of k=53 performs well, evidenced by producing a relatively high true positive rate and a relatively low false positive rate. Using the lag-k critical level of 222 net cpm, we can perform 483−2×53=377 hypothesis tests, one for each observation where a lag-53 background can be calculated. (It may be prudent to add some buffer measurements during the study design, since numerous points in the scan would not have sufficient data on either side to calculate the lag-53 background. The impact that such omission would have on results is not addressed here but should be included in future research.) The null hypothesis is that an observation's counts come from a background distribution, and the alternative hypothesis is that an observation's counts come from a background plus signal distribution. All observed counts that are more than 222 cpm above their specific lag-k background are flagged for follow-up. The results are shown in the confusion table in Table 4. Out of 85 observations with known signal, 72 were flagged as exceeding the background for a true positive rate of 85%. Out of 292 observations with no known signal, 59 were flagged as exceeding background for a false positive rate of 20%.

Table 4. Case Study 2: lag-k confusion table.

Statistic	Predicted Positive	Predicted Negative	Total Actual
Actual Positives	True Positives=72	False Negatives=13	Positives=85
Actual Negatives	False Positives=59	True Negatives=233	Negatives=292
Total Predicted	Positives=131	Negatives=246	

For the sake of comparison, consider an approach that relies on a global estimate of the background, perhaps based on a reference area where 500 observations will be obtained. Again, our *a priori* estimate of the background count rate is 12,000 cpm but now N is 500, not 2 as is the case for the lag-k approach. For $\alpha = 0.05$, the critical level is:

$$L_c = \sqrt{(1 + 1/N)}\sigma_B k_{0.05}$$

$$= \sqrt{\left(1 + \frac{1}{500}\right)} \times 110 \times 1.645$$

$$= 180 \text{ net counts per 60 second time window}$$

We can perform all 483 hypothesis tests since every observation can be compared to the observed mean background of, say, 11,374 cpm. Every observed count above 11,374+180=11,554 gets flagged for follow-up. Out of 85 observations with known signal, 68 were flagged as exceeding background for a true positive rate of 80%. Out of 398 observations with no known signal, 145 were flagged as exceeding background for a false positive rate of 36%.

Table 5. Case Study 2: reference area with 500 observations confusion table.

Statistic	Predicted Positive	Predicted Negative	Total Actual
Actual Positives	True Positives=68	False Negatives=17	Positives=85
Actual Negatives	False Positives=145	True Negatives=253	Negatives=398
Total Predicted	Positives=213	Negatives=270	

Critical level formulas, such as those used in NUREG-1507, are often cited in the literature in a paired setting, which means N=1 and is likely not applicable to scan surveys. For this example, this leads to a critical level of 255 net counts per 60 second time window. If that single background measurement were 11,374 cpm, then every observed count above 11,374+255=11,629 would get flagged for follow-up. Out of 85 observations with known signal, 66 were flagged as exceeding background for a true positive rate of 78%. Out of 398 observations with no known signal, 124 were flagged as exceeding background for a false positive rate of 31%.

Table 6. Case Study 2: reference area with one observation confusion table.

Statistic	Predicted Positive	Predicted Negative	Total Actual
Actual Positives	True Positives=66	False Negatives=19	Positives=85
Actual Negatives	False Positives=124	True Negatives=274	Negatives=398
Total Predicted	Positives=190	Negatives=293	

Finally, we compare the three approaches via true positive and false positive rates:

- Lag-k (for $k = 53$) at 85% true positive rate
- Reference area with 500 observations at 80% true positive rate
- Reference area with 1 observation at 78% true positive rate

False positive rates yield a different ranking with:

- Lag-k (for $k = 53$) at 20% false positive rate
- Reference background with 1 observation at 31% false positive rate
- Reference background with 500 observations at 36% false positive rate

Please note that these comparisons apply to a single example and any serious methodological comparison would include a sophisticated simulation study. Appropriate loss functions can be used to consolidate the information in a confusion table to produce a single ranking, but one needs to carefully consider the tradeoffs between, for example, the false positive and false negative rates.

6.0 Future Work

Several topics should be considered to further advance method development. Below, we provide recommendations focused on additional literature review, data for testing and investigation, parameter and method performance evaluations; and outline a simulation study to determine lag-k method performance.

6.1 Recommendations

- The Science Advisory Board agrees that current MARSSIM guidance does not adequately address modern scanning surveys. Arising from significant technological advancement over the past two decades, newer scanning instruments and mobile systems represent attractive options for consideration and assessment. In addition to the literature reviewed throughout this report, the following literature should be reviewed in the context of continuous data collection and statistical analysis.
  - Quantitative measurements with various example systems are described in the scientific literature (Marques et al. 2021; Peeva, 2021; Ji et al. 2020; Rahman et al. 2020; Ji et al. 2019; Lee and Kim, 2019; Sanada et al. 2019; Azami et al. 2018; Falciglia et al. 2018; Wilhelm et al. 2017; Sinclair et al. 2016; Sanada and Torii, 2015; Kock et al. 2014; Sanderson, 2013; Tanigaki et al. 2013; Kock and Samuelsson, 2011).
  - Detection efficiency and minimum detectable activity for mobile scanning have been investigated regarding scanning speed and signal processing (Falkner and Marianno, 2021; Marianno, 2015).
  - The SAB does not endorse specific detection systems or commercial equipment but does emphasize the importance of detection system calibration to yield measurement quantification with uncertainties that can support defensible final survey results.
- Expand the library of available site and reference area datasets on which to demonstrate scan MDC and hypothesis testing (e.g., lag-k) methods.
- Evaluate performance of these methods compared to the traditional with-vigilance MDC calculation method.
- Consider the impact that omissions of “edge” or “fringe” locations would have when points in the scan would not have sufficient data on either side to calculate the lag-k background (e.g., in the lag-53 case), and whether (and how much) buffer should be added to mitigate such effects.
- Study the impacts of surveyor scan parameters (e.g., using sleds mounted both inside and outside) with various gross-counting instruments, sources, and source scenarios. Vary speed, distance, type of source (distributed vs. point source), and shielding to understand impacts of these variables on without vigilance survey results.
- Survey other existing signal detection frameworks for scan data (Brogan and Brandl 2019) and investigate applications to *a priori* MDCR calculation.
- Further investigate methods described by Alecksen and Whicker (2016), Alecksen and Whicker (2023), and additional related resources to determine how MCNPX code could be used to convert a calculated MDCR to a scan MDC.
- Use simulation and field studies to evaluate the ability of hypothesis testing methods (e.g., lag-k) to detect elevated areas in data collected via without-vigilance surveys for various

instrument configurations and radiological sources, concentrations, spatial contamination areas, and distributions.

- As a result of the above studies, develop recommendations on the following:
  - Approaches to implementation, including software tools and needs for licensees and permit holders to implement the lag-k method.
  - Methods to select an optimal k value for the lag-k method.
  - Flagging values based on integration over a reasonable range of values for k to remove the requirement for a user-specified k and/or reduce the false positive rate.
- Identify the limitations of the lag-k method by outlining site conceptual models for which this approach is/is not suitable.
- Extend lag-k to two dimensions when scanning transects are close to one another.
- Generate test datasets that would be available for licensees and permit holders to learn how to post-process scan data collected without vigilance and apply the lag-k method.

One of the key concerns with producing a scan MDC value for a without-vigilance survey is the difference between the audio click data stream and the logged ratemeter display data stream. NUREG-1507 warns against calculating a scan MDC if logged ratemeter observations will be used because there are concerns the ratemeter will not reach full scale. However, some technology could facilitate simultaneous and synchronized data collection (e.g., Aleckson and Whicker describe using scalar counting mode output to calculate scan MDC). Future work should review the data recommended in Alecksen and Whicker (2023) in greater detail to determine if it provides a viable solution; catalogue and review available technology; and address discrepancies between data streams, potentially building a model to translate the ratemeter display data stream to a binned audio click data stream. This work would require taking field data and pairing logged ratemeter counts with the true audio data stream (possibly via an audio recording with time stamps). After collection, the data would need to be processed, producing paired counts of the two data streams. The paired data streams would then be analyzed to identify and quantify biases between the data capture techniques. Further mathematical modeling work would then be required to determine whether a scan MDC could be calculated for without vigilance surveys that intend to use ratemeter display data.

6.2 Simulation Study for Lag-k

Considerable work needs to be done to characterize the conditions under which the lag-k method performs well and to determine when it performs poorly. The key metric to evaluate performance is a confusion table that contains true positive, true negative, false positive, and false negative counts for lag-k hypothesis tests. To determine these counts, and thus estimate the true positive, true negative, false positive, and false negative rates that can describe the performance of the lag-k method, a known ground truth is required. Such a ground truth can be known by setting it and generating simulated data sets, and then evaluating the method. By doing this for a range of ground truth conditions, the results will lead to recommendations about when and where to use the lag-k method. Performing field tests with known sources is also advisable, but preferable after the simulation study informs specific areas where these more costly studies would be most effective.

Simulation studies can cover many topics of interest including variation in the data generation process, variation in background estimation, variation in analysis methods, and even variation in the metrics of evaluation. The data generation process encompasses both background and source distributions. The background distributions could vary spatially and may extend beyond Poisson. The source distributions could range between a single point source to many point sources to elevated regions of varying shape and size.

The estimation of background is always a critical component in detection methodologies. Different background estimation approaches could be included in the simulation study, enabling the comparison of a variety of analysis methods, including lag-k, Currie's 1968 single paired measurement approach, and the signal detection theory approach given in NUREG-1507. The lag-k approach itself can be expanded to encompass a range of k values or weighted combinations of k values. The lag-k method could be extended to a two-dimensional distance metric expanding beyond the transect-based approach. Backgrounds could be estimated from the entire survey region, from subsets of the survey region, from reference areas that share the same background distribution as the simulated site of interest, or from reference areas that have different background distributions from the simulated site of interest.

Current testing methodology is built around individual location-based hypothesis tests with no multiple testing correction, so the natural scale for method evaluation is on the same individual location-based scale. However, some consideration should be given to a coarser grid for counting true positives and false negatives. A simulation study could provide helpful guidance concerning when a surveyor can expect to find sources of interest and when analysis methods are likely to fail.

7.0 Literature Review

PNNL reviewed numerous articles during research and development of these methods. This section summarizes the relevant literature and provides a timeline of when each article was published and what it contributed to the calculation and application of scan MDCs.

7.1 Seminal Research Paper

The MDC (scanning or not) specifies a level of radioactivity beyond background that can be practically detected by a measurement process. For surface activity measurements, MDC is often reported in units of disintegrations per minute per 100 square centimeters (dpm/100 cm^2). Currie (1968) presented a statistical foundation for scan MDC that is in use today.

Some of the key ideas presented in Currie (1968) include the following:

> *"The statistics of detection and determination apply directly to observations rather than to the underlying physical quantity and, therefore, the following discussions deal specifically with the observed (or observable) signal (meter reading) and its associated random fluctuations."*
> Currie (1968)

- Currie focused on describing the statistical properties of observed counts and used Poisson counting statistics to model radioactive decay.
- For large counts a Poisson distribution can be approximated by a normal distribution.
- The mean and variance of a Poisson distribution are the same so distributions with larger means also have larger variances.
- Currie produced formulas for the limit of detection, L_D, based on α, the probability of a false positive, and β, the probability of a false negative.
- Currie produced formulas for L_C, the critical level, which maps to the decision criteria yielding the desired error rate, α.

Currie (1968) also explained the distinction between gross signal (total counts) and net signal, arrived at by removing the mean in background. This naturally led to questions about the best way to estimate background so as to remove it. Currie investigated one background estimation approach, where the sample of interest is paired to a blank sample "identical, in principle, to the sample of interest, except that the [contaminant] substance sought is absent.

In a later publication, Currie (1984) summarized ideas surrounding the lower limit of detection for the NRC. This work included references to earlier work on lower limits of detection and presaged future publications by noting that the normal assumption must be replaced with an exact Poisson treatment for low count levels. Currie (1984) also considered differences between sample and blank (background) counting times with a model that allowed for measurements from multiple blanks.

7.2 Scaling and Efficiency

Brodsky and Gallaghar (1991) provided an examination of the minimum detectable amount, MDA, in the context of smear survey practices. During smear surveys, a designated area is selected, and samples are taken by wiping or swiping a specific surface using a suitable collection medium, such as filter paper or a swab. The collection medium is then analyzed to determine the presence and level of radioactive contamination. The purpose of smear surveys is to ensure that radioactive materials are properly contained and do not pose a risk to workers, the public, or the environment. The standard two normal distribution approach with unequal variances was used to model net background counts and net sample counts. Gallaghar [1991) used the notation of K for a calibration factor and T for the counting time, which was the same between a sample and its paired blank. Brodsky and Gallaghar [1991) also introduced a low background rate update for one of Currie (1984)'s parameters from 2.71 to 3. Brodsky (1992) provided a more thorough update to Currie (1968)'s L_D formula that rounds the constant 2.71 up to 3. This change accounts for when the blank has minimal activity (low counts) so the normal approximation to a Poisson is suspect.

Brodsky (1992) also included a formula for MDA, which is a scaled version of L_D that includes K and T where K is a calibration factor that converts counts per minute to counts per minute per becquerel and where T is the total counting time of sample assuming a paired blank. Building on Brodsky (1992), Strom and Stansbury (1992) adjusted the MDA for different counting times between sample and blank.

Commission et al. (1997), also known as NUREG-1575 and the Multi-Agency Radiation Survey and Site Investigation manual (MARSSIM) was developed collaboratively among DOD, DOE, EPA, and NRC to provide standardized guidance document for investigating radioactively contaminated sites. NUREG-1507 (NRC 1998 and 2020) is an NRC document that provides guidance to licensees for performing surveys, including an introduction to MDC (under the unequal variances framework) and development of scan MDC using signal detection theory and an equal variance approximation obtained via the index of sensitivity, d′. In NUREG-1507, MDCR (minimum detectable count rate) is calculated first and then transformed into Scan MDC via scaling factors. NRC (1998, 2020) documented a two stage with vigilance approach, noting that Scan MDC is generally computed in the first stage.

Brown and Abelquist (1998) or NUREG/CR-6364 explored human performance factors, signal detection theory, two stage approach, vigilance, and uses $\sqrt{p}$ as a scaling factor for the efficiency of the surveyor. Hart et al. (2003) investigated the scan detection efficiency for detectors to inform corresponding scaling factors. Brandl and Jimenez (2008) elucidated continuously recorded measurements and explored alarm levels conditional on known backgrounds, count rates, and sample counting times. King et al. (2012) accounted for the swing of a detector during walking surveys.

Brandl (2013) investigated low signal-to-background ratios to lower decision thresholds for continuous scanning measurements. This work assumed negative binomial distributions for the sequential count data, extending beyond the analysis of each data point individually.

Abelquist (2014) provided an excellent source explaining and defining relevant statistics including L_C, L_D, K, and T in the context of scan MDC via signal detection theory. This work included different collection times for background and samples surveyor efficiency and alpha scan MDC.

Alecksen and Whicker (2016) focused on the conversion of MDCR to scan MDC via Monte Carlo N-particle Extended (MCNPX) to establish efficiency scaling factors. Note that MCNPX is no longer being updated regularly. Starting with MCNP6, Los Alamos National Laboratory has put the alpha particle and heavy ion transport features into the MCNP6.x versions.

Brogan and Brandl (2018) improved hypothesis testing and handled low count rate data when background is not perfectly known by adding data points together and extending collection times. This method does not account for varying background conditions. It uses a variation on the negative binomial distribution to calculate false positive rates when *n* out of *N* sequential measurements exceed a threshold level.

Watson et al. (2018) performed an experiment for alpha/beta radiation using a Shiryaev-Roberts control chart methodology.

Brown and Abelquist provided information for radiation exposure including dose exposure formulas in NUREG/CR-6364 (NRC 1998).

Falkner and Marianno (2019) developed a well-defined relationship to predict MDA as a function of detector speed. Detector efficiency can be improved by slowing the speed of travel; this work quantifies the relationship between detector speed and the minimum detectable activity. The work is agnostic to the with-vigilance versus without-vigilance dichotomy.

Justus (2019) developed decision levels from the Poisson model without using a normal approximation.

Alecksen and Whicker (2023) covered similar topics to those included in PNNL's Scan MDC report. Their research is in agreement that scan MDC can be calculated for without vigilance surveys using the existing (NUREG-1507) approach, "Whether based on the approximate MDCR calculation or the exact MDCR expression, the probabilistic method described by Alecksen and Whicker (2016) for determining scan MDCs for GPS-based gamma surveys is a statistically valid application of the MDCR concept, provided that the gamma instruments used for the scan system are operated in scaler counting mode rather than ratemeter mode."

Additionally, Alecksen and Whicker (2023) performed field tests to compare unequal variance analysis to simpler equal variance analysis. They found that the unequal variance approach produced error rates closer to nominal but also stated that the equal variance approximation was not problematic and may be preferred as a simpler method.

Alecksen and Whicker (2023) also considered alternate method evaluation procedures. They investigated the grouping of individual observations to improve sensitivity and specificity at the expense of localization, an idea also included in the PNNL report.

7.3 Bayesian Methods

Klumpp (2013) explored a Bayesian approach for modeling low count rate radioactive sources via always-on detection system. This approach used a gamma prior on the rate parameter, combined the prior with data, and created a posterior distribution for the rate. Their version of a hypothesis test involved extracting the probability that the true count rate was above a background from the posterior distribution.

Tandon et al. (2016) implemented a Bayesian aggregation of data from mobile spectrometers. Their Bayesian approach to hypothesis testing extended beyond one observation per hypothesis test. Their method relies on spectrometry data in addition to count data to help separate background radiation from source radiation. They also leverage information in the spatial correlations of their observations to help localize sources.

Brogan and Brandl (2019) developed a Bayesian interaction model to analyze gross count measurements. They used the Bayesian linear regression model to study the relationship between gross count measurements and their standard deviation observed in the "current" and previous four measurements at fixed time intervals. They incorporated a predictor (explanatory variable) in the model to indicate whether data originated from background or measurements and used its estimate to develop a decision rule to identify elevated areas. This method requires previously recorded background gross count data.

The following articles detail case studies that use Bayesian methods:

- Kim et al. (2019) found that the maximum detectable depths for weakly active radioactive sources was 21 cm using a low-resolution NAI(TI) detector and Bayesian inference.

- Arahmane et al. (2021) combined spectrometry results from HPGe detectors with Bayesian methodologies to estimate surface activities of low-activity uranium contamination.

- Michaud et al. (2021) described a hierarchical Bayesian model to localize radiation sources amidst a varying background.

- Arahmane et al. (2022) used Bayes factors to simulate the detection of weak uranium signals on concrete.

7.4 Counts to Dose Conversion Factors

To varying degrees NUREG-1507, MARSSIM, Abelquist (2014), Alecksen and Whicker (2016), and Brown (2018) discussed components of the counts to dose conversion factors.

7.5 Varying Background*

King and Vitkus (2015) gave an example of an applied scan MDC process starting with an MDCR analysis through to radiological survey data. They demonstrated how deeply important the estimation of background is. Michaud (2021) explicitly accounted for a variable background in their Bayesian modeling work.

7.6 Additional Sources

The following research papers may be pertinent to these topics but reviewing them was outside this report's efforts:

- Abd Rahman et al. (2020)
- Azami et al. (2018)
- Falciglia et al. (2018)
- Ji et al. (2019)
- Ji et al. (2020)
- Kock and Samuelsson (2011)
- Kock et al. (2014)
- Lee and Kim (2019)
- Marques et al. (2021)
- Peeva (2021)

- Sanada and Torii (2015)
- Sanada et al. (2019)
- Sinclair et al. (2016)
- Tanigaki et al. (2013)
- Wilhelm et al. (2017)

8.0 References

Abd Rahman, N.A., Sahari, K.S.M., Jalal, M.F.A., Rahman, A.A., Abd Adziz, M.I. and Hassan, M.Z. 2020, April. “Mobile robot for radiation mapping in indoor environment.” *In IOP Conference Series: Materials Science and Engineering* (Vol. 785, No. 1, p. 012021). IOP Publishing.

Abelquist, E. 2014. *Decommissioning Health Physics: A Handbook for MARSSIM Users.* Boca Raton, FL: CRC Press, Taylor and Francis Group.

Alecksen, T. and R. Whicker. 2016. “Scan MDCs for GPS-Based Gamma Radiation Surveys.” *Health Physics* 111 (2): S123–S132.

Aleksen, T. and R. Whicker. 2023. “Retrospective detection sensitivity for GPS-based gamma radiation surveys.” *Health Physics* 124(6): 451–461.

Arahmane, H., J. Dumazert, E. Barat, T. Dautremer, N. Dufour, F. Carrel, M. Michel, and F. Lainé. 2021. “A reliable absolute and relative Bayesian method for nuclear decommissioning: Low-level radioactivity detection with gamma-ray spectrometry.” *IEEE Transactions on Instrumentation and Measurement* 70: 1–18.

Arahmane, H., J. Dumazert, E. Barat, T. Dautremer, F. Carrel, N. Dufour, and M. Michel. 2022. “Bayesian inference based on a bivariate gamma distribution of kibble for low-level radioactivity detection in nuclear decommissioning operations.” *Process Safety and Environmental Protection* 163: 727–742.

Azami, K., Ootagaki, T., Ishida, M. and Sanada, Y. 2018. “Characteristics of radiocesium contamination of dry riverbeds due to the Fukushima Daiichi Nuclear Power Plant accident assessed by airborne radiation monitoring.” *Landscape and ecological engineering* 14: 3-15.

Brandl, A. 2013. “Statistical considerations for improved signal identification from repeated measurements at low signal-to-background ratios.” *Health physics* 104(3): 256–263.

Brandl, A. and Herrera Jimenez, A.D. 2008. “Statistical criteria to set alarm levels for continuous measurements of ground contamination.” *Health Physics* 95(2): S128–S132.

Brodsky, A. 1992. “Exact calculation of probabilities of false positives and false negatives for low background counting.” *Health Physics* 63 (2): 198–204.

Brogan, J. and A. Brandl. 2018. “Enhancing test statistics by utilizing data patterns in sequential measurement strings in radiation detection.” *Health Physics* 115(6): 698–704.

Brogan, J. and A. Brandl. 2019. “Developing Detection Decisions on the Absence or Presence of a Radiological Source Using a Bayesian Interaction Model.” *Health Physics* 117 (6): 637–647.

Brown, S.H., R. Edge, J. Elmer, and M. McDonald. 2018. “Establishing radiological screening levels for defense-related uranium mine (drum) sites on BLM land using a recreational future-use scenario.” *Health Physics* 114(6): 588–601.

Currie, L. A. 1968. “Limits for qualitative detection and quantitative determination. Application to radiochemistry.” *Analytical Chemistry* 40 (3): 586–593.

Falkner, J. and C. Marianno. 2019. “Modeling minimum detectable activity as a function of detector speed.” *Radiation Detection Technology and Methods* 3 (3): 25.

Currie, L.A. 1984. *Lower Limit of Detection: Definition and Elaboration of a Proposed Position for Radiological Effluent and Environmental Measurements.* Technical report, National Bureau of Standards, Washington, DC (USA).

Falciglia, P., L. Biondi, R. Catalano, G. Immè, S. Romano, and F. Vagliasindi. 2018. “Preliminary investigation for a quasi-quantitative characterization of soils contaminated with 241Am and 152Eu by low-altitude unmanned aerial vehicles (UAVs) equipped with small size γ-ray spectrometer: detection efficiency and minimum detectable activity (MDA) concentration assessment.” *Journal of Soils and Sediments* 18: 2399-2409.

Falkner, J. and C. Marianno. 2019. “Modeling Minimum Detectable Activity as a Function of Detector Speed.” *Radiation Detection Technology and Methods* 3: 1–8.

Hart, K., W. Duffy, K. Higley, C. Marianno, and C. Moss. 2003. “Predicting instrument detection efficiency when scanning point and small area radiation sources.” *Health Physics* 84 (5): 616–625.

Huckett, J., D. Fagan, Z. Weller, M. Obiri, L. Newburn, F. Day-Lewis, and D. Peeler. 2022. “Subsurface Radiological Survey Design and Geospatial Analysis Tool Recommendations.” PNNL-33647.

Ji, Y.Y., Lim, T., Choi, H.Y., Chung, K.H. and Kang, M.J. 2019. “Development and Performance of a Multipurpose System for the Environmental Radiation Survey Based on a LaBr 3 (Ce) Detector.” *IEEE Transactions on Nuclear Science* 66(12): 2422-2429.

Ji, Y.Y., Lim, T., Hitomi, K. and Yajima, T. 2020. “Spectrometric Estimation of Dose Rate Induced from Radioactive Cesium in the Ground Using a Mobile Gamma-Ray Spectrometry Based on a LaBr3 (Ce) Detector.” *Health physics* 118(2): 215-225.

Justus, A.L. 2019. “A Purely Poisson-based Approach to Estimating Audible Scan Survey Sensitivities.” *Health Physics* 116(1): 27–41.

Kim, J., K.T. Lim, K. Park, and G. Cho. 2019. “A Bayesian Approach for Remote Depth Estimation of Buried Low-level Radioactive Waste with a NaI(TI) Detector.” *Sensors* 19(24): 5365.

King, D. A., N. Altic, and C. Greer. 2012. “Minimum Detectable Concentration as a Function of Gamma Walkover Survey Technique.” *Health Physics* 102 (2): S22–S27.

King, D. A. and T. Vitkus. 2015. “Lessons learned on the presentation of scan data.” *Health Physics* 109 (3): S212–S218.

Klumpp, J. and Brandl, A. 2015. “Bayesian Analysis of Energy and Count Rate Data for Detection of Low Count Rate Radioactive Sources.” *Health Physics* 108(3):364-370.

Kock, P. and Samuelsson, C. 2011. “Comparison of airborne and terrestrial gamma spectrometry measurements-evaluation of three areas in southern Sweden.” *Journal of Environmental Radioactivity* 102(6): 605-613.

Kock, P., Rääf, C. and Samuelsson, C. 2014. “On background radiation gradients–the use of airborne surveys when searching for orphan sources using mobile gamma-ray spectrometry.” *Journal of environmental radioactivity* 128: 84-90.

Lee, C. and Kim, H.R. 2019. “Optimizing UAV-based radiation sensor systems for aerial surveys.” *Journal of environmental radioactivity* 204: 76-85.

Marianno, C.M. 2015. “Signal processing and its effect on scanning efficiencies for a field instrument for detecting low-energy radiation.” *Health Physics* Vol. 109: 78-83.

Marques, L., A. Vale, P. Vaz. 2021. “State-of-the-Art Mobile Detection Systems for Different Scenarios.” *Sensors* 21(4): 1051.

Michaud, I.J., K. Schmidt, R.C. Smith, and J. Mattingly. 2021. “A Hierarchical Bayesian Model for Background Variation in Radiation Source Localization.” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 1002: 165288.

NRC (U.S. Nuclear Regulatory Commission). 1998. *Human Performance in Radiological Survey Scanning*. NUREG/CR-6364, BNL-NUREG-52474, Washington, D.C.

NRC (U.S. Nuclear Regulatory Commission). 1998. *Minimum Detectable Concentrations With Typical Radiation Survey Instruments for Various Contaminants and Field Conditions*. NUREG-1507, Washington, D.C.

NRC/EPA/DOE (U.S. Nuclear Regulatory Commission, U.S. Environmental Protection Agency, U.S. Department of Energy). 2020. *Multi-Agency Radiation Survey and Site Investigation Manual (MARSSIM)*. NUREG-1575, Rev. 2, EPA 402-P-20-001, DOE/AU-0002. Washington, D.C.

Peeva, A. 2021. “Now Available: New Drone Technology for Radiological Monitoring in Emergency Situations.” *IAEA (International Atomic Energy Agency)*.

Rahman, N.A.A., K.S.M. Sahari, M.F.A. Jalal, A.A. Rahman, M.I.A. Adziz, M.Z. Hassan. 2020. “Mobile robot for radiation mapping in indoor environment.” IOP Conference Series: Materials Science and Engineering, Vol. 785, 012021.

Sanada, Y. and Torii, T. 2015. “Aerial radiation monitoring around the Fukushima Dai-ichi nuclear power plant using an unmanned helicopter.” *Journal of environmental radioactivity* 139: 294-299.

Sanada, Y., Urabe, Y., Sasaki, M., Ochi, K. and Torii, T. 2019. “Evaluation of ecological half-life of dose rate based on airborne radiation monitoring following the Fukushima Dai-ichi nuclear power plant accident.” *Journal of Environmental Radioactivity* 210: 105816.

Sanderson, D. 2013. “Measuring regional scale distribution of radiocaesium.” Presented at the Caesium workshop: Fukushima recovery–understanding, modelling and managing radiocaesium decontamination, COEASSE. Fukushima, Japan. September 30–October 3, 2013.

Sinclair, L.E., Fortin, R., Buckle, J.L., Coyle, M.J., Van Brabant, R.A., Harvey, B.J., Seywerd, H.C. and McCurdy, M.W. 2016. “Aerial mobile radiation survey following detonation of a radiological dispersal device.” *Health physics* 110(5): 458-470.

Strom, D.J. and P.S. Stansbury. 1992. “Minimum Detectable Activity when Background is Counted Longer than the Sample.” *Health Physics* 63(3): 360–361.

Tandon, P., P. Huggins, R. Maclachlan, A. Dubrawski, K. Nelson, and S. Labov. 2016. “Detection of Radioactive Sources in Urban Scenes Using Bayesian Aggregation of Data from Mobile Spectrometers.” *Information Systems* 57: 195–206.

Tanigaki, M., Okumura, R., Takamiya, K., Sato, N., Yoshino, H. and Yamana, H. 2013. “Development of a car-borne γ-ray survey system, KURAMA.” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 726: 162-168.

Watson, M.M., A.F. Seliman, V.N. Bliznyuk, and T.A. DeVol. 2019. “Evaluation of Shiryaev-Roberts Procedure for On-line Environmental Radiation Monitoring.” *Journal of Environmental Radioactivity* 192: 587–591.

Wilhelm, E., Gutierrez, S., Ménard, S. and Nourreddine, A.M. 2017. “A method for determining Am-241 activity for large area contamination.” *Applied Radiation and Isotopes* 119: 86-93.

Pacific Northwest National Laboratory

902 Battelle Boulevard
P.O. Box 999
Richland, WA 99354
1-888-375-PNNL (7665)

www.pnnl.gov* | *www.nrc.gov