3-2022

Analysis of Task Performance during Radiological Surveillance by Means of Discrete Event Simulation

Michael H. Ames

ANALYSIS OF TASK PERFORMANCE DURING RADIOLOGICAL SURVEILLANCE BY MEANS OF DISCRETE EVENT SIMULATION

THESIS

Michael H. Ames, MSgt, USAF

AFIT-ENV-MS-22-M-175

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

AFIT-ENV-MS-22-M-175

ANALYSIS OF TASK PERFORMANCE DURING RADIOLOGICAL SURVEILLANCE BY MEANS OF DISCRETE EVENT SIMULATION

THESIS

Presented to the Faculty

Department of Systems Engineering and Management

Graduate School of Engineering and Management

Air Force Institute of Technology

Air University

Air Education and Training Command

In Partial Fulfillment of the Requirements for the

Degree of Master of Science in Environmental Engineering and Science

Degree of Master of Science in Industrial Hygiene

Michael H. Ames, BS

Master Sergeant, USAF

March 2022

AFIT-ENV-MS-22-M-175

ANALYSIS OF TASK PERFORMANCE DURING RADIOLOGICAL SURVEILLANCE BY MEANS OF DISCRETE EVENT SIMULATION

Michael H. Ames, BS

Master Sergeant, USAF

Committee Membership:

Dr. J. M. Slagley, PhD
Chair

Maj J. P. Kristbaum, PhD
Member

Dr. J. R. Cezeaux, PhD
Member

AFIT-ENV-MS-22-M-175

Abstract

The surveillance and detection of radioactive contamination on surfaces and in the environment are commonly investigated by surveyors utilizing portable detection equipment. The availability of Discrete Event Simulation (DES) and Human Performance Modeling (HPM) allows for the analysis of physical and cognitive processes associated with these operations, as well as the effect that external environmental factors have on surveyor performance. This research uses the Improved Performance Research Integration Tool (IMPRINT) to approximate the performance of a radiological detection task informed by the observation of six surveyors. The effects of chemical Individual Protective Equipment (IPE) use is evaluated along with the effects of elevated ambient environmental temperatures. Along with the development of a novel human performance model for the surveillance task, results of this study predict up to a 33% increase in survey completion time when chemical IPE is worn and up to a 50% decrease in surveyor efficiency from the effects of elevated ambient temperatures. Overall, this study represents the novel use of a DES to model the cognitive and physical tasks associated with radiological surveillance activities and the impacts from key physical and environmental stressors.

Acknowledgments

I would like to thank my advisor, Dr. Jeremy Slagley, for his guidance and mentorship during this research and throughout my academic journey. I would also like to thank my committee members Major Joseph Kristbaum and Dr. Jason Cezeaux, whose insights and recommendations have made this work possible. Additionally, I would like to express my appreciation to Dr. William Brown and Dr. Eric Abelquist, whose groundbreaking work laid the foundation for this research and whom have been invaluable. Most importantly, I would like to thank my wife and son for their love and sacrifice during this effort.

Michael H. Ames

Table of Contents

Page

List of Figures

Page

List of Tables

Page

ANALYSIS OF TASK PERFORMANCE DURING RADIOLOGICAL SURVEILLANCE BY MEANS OF DISCRETE EVENT SIMULATION

I. Introduction

General Issue

Portable radiological survey equipment is commonly used to determine the presence and extent of residual contamination on surfaces and in the environment. Personnel tasked with performing these surveys may do so as part of an emergency response scenario or may do so to comply with specific requirements for licensees of regulated radioactive material. In either case, surveys are performed using scanning and evaluation techniques that rely on a technician's ability to correctly interpret the output of the instrument and make appropriate decisions.

One of the most important measures applicable to a surveyor is termed the Minimum Detectable Concentration (MDC). The MDC for a scanning survey depends on many of the same factors that influence the detection under static (fixed) conditions. These include the level of background measured by the instrument, the types of radiation monitored, intrinsic characteristics of the detector, source and measurement geometry, and the desired level of confidence [1]. Equally important are factors which depend on the technique of the surveyor and the ability to discriminate between background count rates and count rates due to areas of elevated activity. Under certain physical and environmental stressors, however, the performance of a surveyor and the ability to discriminate between elevated radiological count rates and background count rates is reduced while the amount of time required to perform a radiological survey is increased. The Minimum Detectable Count Rate (MDCR) of a surveyor, closely related to the MDC

for a scan survey, is determined to demonstrate that the equipment and techniques employed are sensitive enough to alert a surveyor to a potential area of concern. An MDCR several times higher than observed background count rates may result in missed areas of contamination on a surface or materials, leading to an incorrect decision regarding status and compliance with applicable standards. If background count rates are also elevated, as might be encountered in a radiological or nuclear incident, failure to consider the MDCR may also lead to the potential for increased radiological health risk, non-compliance with applicable regulatory standards, contamination of otherwise "clean" surfaces and materials, and the cost (monetary or otherwise) of additional investigational and decontamination resources.

While the direct observation of radiological surveillance activities is preferred in most cases to simulations, several disadvantages of direct observation should be considered. Modern computational tools and software packages can analyze a task or human-computer interface as discrete events, providing evaluators with a powerful tool from which to estimate changes in task performance. Used correctly, these tools can provide information on workload, stressors, and impacts to performance.

Problem Statement

The amount of quantitative data available which describes the impacts from specific environmental stressors on radiological survey performance is limited. The effects of heat stress and wear of chemical protective equipment on radiological surveyor performance is not well known and has not been well studied or observed.

Research Objectives

The objective of this research is to develop a model using Discrete Event Simulation (DES) software which can approximate the sensitivity and discriminability of a radiation surveyor to perform a surveillance task using portable equipment. Using available data on the degradation of task performance from environmental and physical stressors (Heat and chemical protective equipment), predict the impact that these stressors have on surveyor performance, survey task completion times, and the MDCR.

Research Focus

This research focused on using discrete event simulation software to provide quantitative estimations of the changes to task performance during a radiological surveillance task involving portable equipment.

Investigative Questions

1. Can a radiological detection task be modeled as a series of discrete events using computer simulation software?
2. Does the use of personal protective equipment modify the amount of time required to perform a radiological detection task?
3. Do external environmental conditions, such as elevated ambient air temperatures, affect the performance of a surveyor performing a radiological detection task?

Assumptions/Limitations

Models approximating human performance are often simplified to include enough detail to sufficiently describe the tasks while balancing the available time and resource constraints. Several global assumptions were made for this study. First, the radiological surveyor data used for this study was collected over 20 years ago during a previous effort to describe radiological surveyor performance. While every attempt was made to ensure the accuracy and completeness of the dataset, the information available is assumed to accurately describe the performance of the radiological task. Second, SME estimates used to assign VACP values are assumed to be accurate, based on the author's prior experience conducting radiological surveillance tasks. A complete set of assumptions specific to each model can be found in Appendix A.

Implications

Research into the human factors associated with radiological surveillance is extremely limited. Although observation of field performance during radiological detection tasks is preferred in many cases, modeling allows for the study of effects from individual stressors, both intrinsic and extrinsic to the surveyor. Factors such as heat, cold, fatigue, personal protective equipment, etc., can be applied to individual surveyor tasks, thereby providing a quantitative estimation of impact on surveyor performance. Discrete Event Simulation allows for the separation of tasks and subtasks to identify the cognitive or psychomotor process of an individual over time. Armed with these tools, an estimation of the impact different environmental changes has on a surveyor can be made.

This can be done without the need for experimentation involving actual personnel and the expenditure of resources to engineer or redesign human-machine interfaces.

II. Literature Review

Chapter Overview

This chapter will introduce the underlying framework of signal detection theory for single factor discrimination, the use of discrete event simulation software for task and human performance modeling, and the Minimum Detectable Concentration (MDC) as it relates to radiological surveys.

Signal Detection Theory Framework

Signal Detection Theory (SDT) is a framework for organizing and processing the perception of signals based on a stimulus, either auditory or visually, within a background. The origination of SDT dates back to 1860 by Gustav Fechner and concerned the discrimination of two similar stimuli in a noisy background [2]. Fechner is commonly regarded as one of the first experimental psychologists, researchers who apply the scientific process to understand human cognition and thought. Fechner is credited with discovering the first experimental type of SDT referred to as Two-Alternative Forced-Choice (2AFC). 2AFC studies are considered comparison designs; they allow for the measurement of sensitivity to two stimuli where an observer must choose which one is “correct”. An example may be an experiment designed to answer which one of two lights are perceived to be brighter, or which one of two weights are perceived to be “heavier” by an observer. 2AFC study designs are often preferred to a less binary

"yes/no" study design in that they reduce the impacts of observer bias, thereby leading to a better estimation of sensitivity where the impact of bias on the overall result is of concern [3]. If a researcher wishes to ensure that the impact of any bias on the results is reduced, a 2AFC type of study design is often preferred. As will be discussed later, sensitivity and bias are two key measures used to describe the performance of an observer in a one-interval (single factor) experiment. The advent of what experimental psychologists now regard as modern SDT began with a series of publications in the mid-1950s by several pioneering psychologists, including Peterson, Birdsall, Neyman, Pearson, Tanner, and Swets. The first application of SDT to the auditory domain is from work by Smith and Wilson (1953) [4].

Experiments using one interval (yes-no or present-absent) approaches collect and record four different types of responses from the observer: hits, misses, false alarms, and correct rejections. The number of hits (H) correspond to the number of responses made when a condition of interest (in general, a signal) is present and an observer indicates a response of "yes". A "miss" is recorded when an observer fails to respond (or responds no) when a signal is present and an observer indicates a response of "no". A False Alarm (FA) is recorded when a signal is not present, but the observer responds "yes". Correspondingly, a "correct rejection" is annotated when an observer correctly decides that the signal is not present, and "correctly" responds "no". Table 1 shows the four different types of responses in a one-interval study.

Table 1. Four possible responses in a one-interval study

Condition	Observer	
	Response "Yes"	**Response "No"**
Signal Present	*Hit (H)*	*Miss*
Signal Absent	*False Alarm (FA)*	*Correct Rejection*

Using the response types above, empirical formulas can be derived which represent the Hit Rate (HR), or the number of observations to which an observer responded "yes" when a signal was present over the total number of "signal present" conditions. Written empirically:

$$Hit\ Rate = P(\text{"yes"} \mid \text{Signal Present}) \quad (1)$$

Where:

$P = the\ proportion\ of\ responses$

The HR, therefore, can be determined by the proportion of "yes" responses when a given condition (or stimulus) is present. Modern day SDT also refers to this value as 1-β (not to be confused with the likelihood ratio discussed below) [2]

Similarly, the false alarm (FA) rate (also annotated as α), or the number of observations where an observer responded "yes" over the total number of "signal absent" conditions, can be calculated as:

$$False\ Alarm\ Rate = P(\text{"yes"} \mid \text{Signal Absent}) \quad (2)$$

Where:

$P = the\ proportion\ of\ responses$

So called “one interval” designs are simple but effective study designs within the modern SDT framework. Under this design, the ability of an observer to discriminate between two stimuli, often simplified to a “signal” and the combination of signal plus noise is the standardized distance between two normal gaussian distributions with approximately equal variances (one representing the noise, the other the combination of signal + noise) [5]. The use of the terms “signal” and “noise” are typical in the field of psychophysics and used in many published works utilizing SDT. The common measure of distance between the means of two stimuli is known as *d’* (pronounced dee-prime), also referred to as the *index of sensitivity*. The criterion (a measure of bias) is a point of decision by an observer and can be associated with the act of an observer adopting either a conservative or liberal approach when reporting observations. The placement of the criterion will depend on the *a priori* probability of signals and the costs (such as time) of an outcome as judged by an observer [6],[1].

Figure 1 shows the statistical model indicating the relationship between the two stimuli, the location of the criterion, and the relative measure of *d’*.

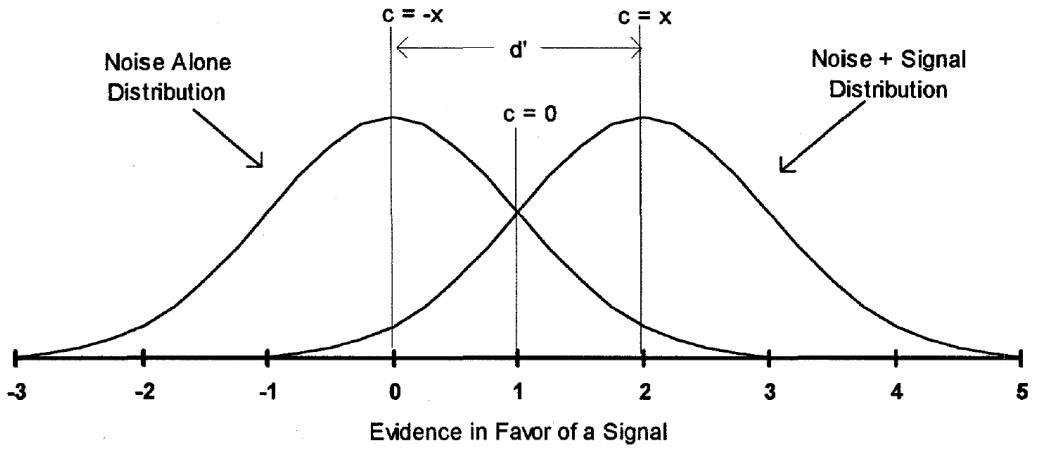


Figure 1. SDT Measures of Sensitivity and Criterion with Gaussian Distributions [1]

Work by Tanner and Swets (1954) introduced the concept of the decision criterion, often abbreviated "*c*" [7]. When *c* equals zero, an observer is neither conservative nor liberal in their decision, commonly viewed as ideal or nonbiased. Negative values of *c* indicate the willingness of an observer to respond "yes" to the presence of a signal relative to an ideal observer and is therefore more liberal. Positive values of *c* indicate the willingness of an observer to respond "no" relative to an ideal observer, also referred to as conservative.

While several measures exist which describe the sensitivity of an observer to a given set of stimuli (for example, the use of probability correct *(p(c))*, the most widely used in SDT is *d'*[3]. It is defined as the difference between values which correspond to the inverse of the normal cumulative distribution function (CDF or Φ^{-1}). The index of sensitivity can be written empirically as:

$$d' = z(H) - z(FA) \tag{3}$$

Where:

$d' = index\ of\ sensitivity$

$z(H) = Inverse\ of\ the\ normal\ CDF\ associated\ with\ the\ Hit\ Rate\ (\Phi^{-1})$

$z(F) = Inverse\ of\ the\ normal\ CDF\ associated\ with\ the\ False\ Alarm\ Rate\ (\Phi^{-1})$

The location of the criterion (*c*), in the same units of standard deviation associated with the z-normal CDF, can be calculated as:

$$c = \frac{1}{2}[z(H) + z(FA)] \tag{4}$$

Where:

$c = criterion\ value$

$z(H) = Inverse\ of\ the\ normal\ CDF\ associated\ with\ the\ Hit\ Rate\ (\Phi^{-1})$

$z(F) = Inverse\ of\ the\ normal\ CDF\ associated\ with\ the\ False\ Alarm\ Rate\ (\Phi^{-1})$

Several other measures of bias exist according to SDT for the normal (gaussian) model, including the relative criterion (*c'*) and the likelihood ratio (β). The relative criterion simply scales the criterion location relative to sensitivity ($\frac{c}{d'}$). The likelihood ratio (β) is a ratio of two continuous density functions corresponding to the Signal and Signal + Noise [3]. This can be rewritten as:

$$\beta = e^{cd'} \tag{5}$$

Where:

$\beta = Likelihood\ Ratio$

$e = base\ of\ the\ natural\ log$

$c = criterion\ value$

$d' = index\ of\ sensitivity$

While not directly applied to this research, the use of β as a measure of bias is important in SDT and may lead to a different set of conclusions about an observer's behaviors than when *d'* and *c* are used alone. An example of this is given in Hautus et al (2022), where the measure of bias of an observer after training may be indicated using β, but not when *c* is used alone. This information may be of use to researchers looking at the

impact of detection sensitivity *before* and *after* some amount of training or instruction is provided to a study participant [3].

Values associated with the measurement of sensitivity (*d'*) and bias (*c*) can be plotted on a graph where the rate of true positives (i.e. the hit rate) and the rate of false positives (i.e. the false alarm rate) are plotted on the y and x axis, respectively. This type of graph was first described by Peterson and Birdsall in 1954 [8]. The curve produced by plotting d' is referred to as the Receiver Operating Characteristic, or ROC [9]. The use of ROC curves is not limited solely to the display of psychoacoustic data, but rather broadly used as a diagnostic tool in biology, psychology, radiology, and the medical sciences to compare the accuracy of a test or model [10]. An idealized ROC chart showing several measures of *d'* on both linear and z-transformed scales can be seen in Figure 2.

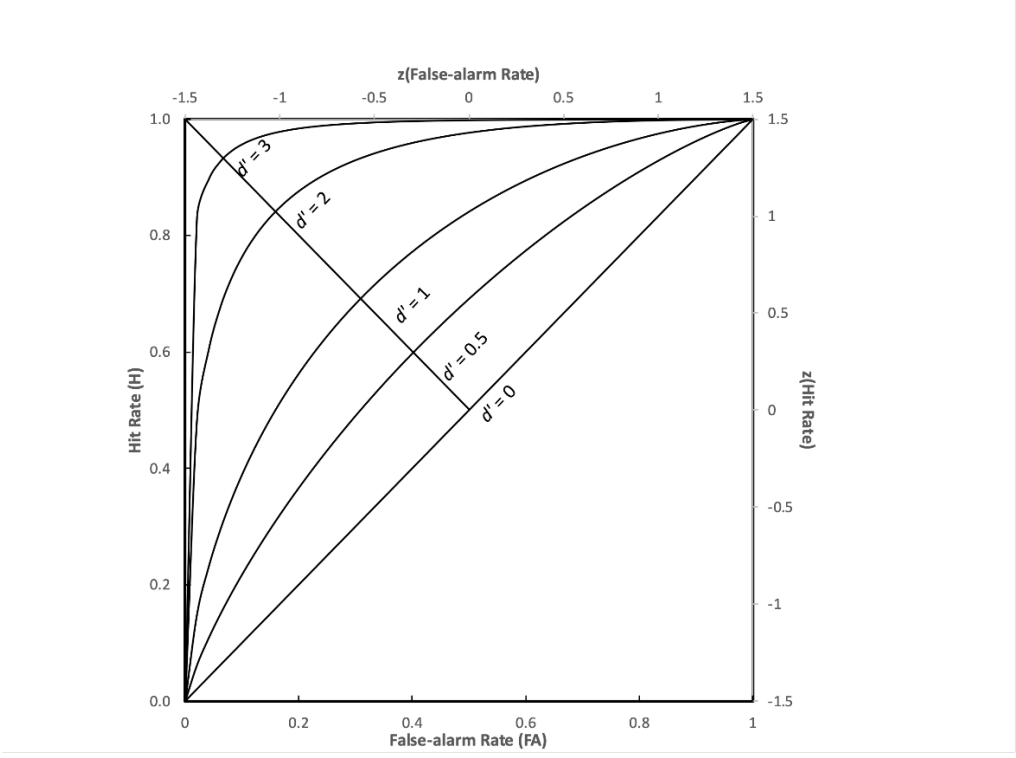


Figure 2. ROC Curves of *d'* with both linear and z-transformed scales

For a given measure of *d'*, there are several combinations of H and F that can result. Points along a given measure of discriminability (indicating changes to HR and FA rate) change only based on the placement of the criterion of the observer. In each case, *d'* remains static and the location of the criterion (*c*) changes. This relationship can be seen in Figure 3. Two different values of *c* (labeled C_1 and C_2) can be seen along a fixed value of *d'* plotted on the ROC chart. These same values are seen on the corresponding probability density graphs (again with the same value of *d'*), each marked by a dashed vertical line. The Gaussian curve highlighted in red corresponds to signal + noise, whereas the curve in blue corresponds to noise alone. This illustrates an important relationship between measures of discriminability (*d'*) and bias (*c*) of an observer,

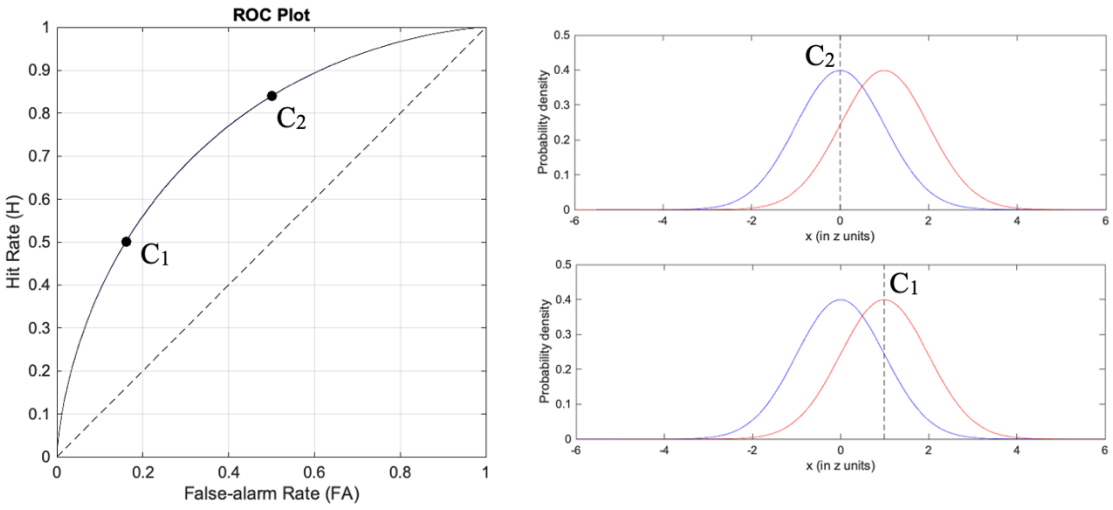


Figure 3. Signal detection interpretation of ROC data with different criterion [11]

The depiction of a line across the major diagonal at *d'*=0 is sometimes referred to as the "chance line", where the surveyor's ability to discriminate the signal from noise

falls to zero. At increasing levels of sensitivity (discriminability), the resulting curve shifts up and to the left (describing an increase in hit rate and corresponding decrease in FA rate).

If the goal of a surveyor is to maximize discriminability (maximizing the HR and reducing the FA rate), then the concept of the "ideal" observer becomes important in order to describe how "efficient" an observer is at maximizing the probability of a "hit" while reducing the probability of a "false alarm". Theoretically, there is no maximum value of *d'*. As the HR approaches a value of 1 and the FA rate is reduced to a closer approximation of 0, the value of *d'* will increase dramatically (limited only by the number of significant figures used). Practically, values of d' are rarely given above 5 (corresponding to a HR of 0.99 and FA rate of 0.01) in most experiments, and many are specifically designed to avoid situations where the HR or FA rate are equal to exactly 0 or 1 to avoid errors when determining a z value from the standard normal (z) distribution. When encountered, several standard practices have been used when analyzing operator results (such as adding 0.5 to values of 0) to compensate for these errors at the cost of slightly modifying the overall value of *d'*. Under experimental conditions an observer is far from error-free. The distribution of values used to determine the ideal observer has a Poisson, rather than a normal distribution, which is found to describe the absolute thresholds at which a surveyor could discern changes in audio tones over a period of time, incorporated into early ideal-observer models of performance [5].

If the performance of an observer can be measured via the index of sensitivity (*d'*), and the performance of an ideal surveyor can also be defined in similar terms, then

the efficiency of an observer can be calculated based on the square of the ratio, given as [3]:

$$e_{surveyor} = \left[\frac{d'_{surveyor}}{d'_{ideal}}\right]^2 \quad (6)$$

Where:

$e_{surveyor} = the\ overall\ efficiency\ of\ a\ surveyor$

$d'_{surveyor} = Index\ of\ sensitivity\ determined\ by\ the\ difference\ in\ H\ and\ FA$

$d'_{ideal} = Index\ of\ sensitivity\ determined\ for\ the\ ideal\ observer$

SDT and Radiological Surveillance

The perception of audio signals, the ability to correctly discriminate "signal" from "noise", and the changes to criterion based on the bias of an observer are important factors during radiological surveillance using portable instrumentation. As indicated by Brown and Abelquist (1998), a surveyor's sensory processes (i.e., the ability to discriminate a "hit") and a surveyor's decision process (expectations of contamination *a priori* and weighing of costs) are part of the scanning survey process leading to the identification of contamination [1]. A surveyor typically uses both audible and visual cues to be alerted to the presence of contamination. Surveyors must rely on the perception of instrument-generated clicks, ticks, or visual readings to drive decisions regarding the presence of radioactivity. Within the SDT framework, the perception of a signal in background noise is directly analogous to the detection of radioactivity given a certain amount of natural background (registered as identical clicks or ticks), which is a function of both the instrument used and the environment [1].

After extended monitoring periods, changes in surveyor performance may be attributed to what has been termed “vigilance decrement”. Vigilance and vigilance decrement has been the focus of several studies since the early 1980s. In many cases, it is difficult to determine if changes to sensitivity are the result of changes from a vigilance decrement or simply changes to an observer’s criterion following extended monitoring periods. As was discussed within the previous section on SDT, changes to the criterion are essentially reflective on the willingness (or lack thereof) of a surveyor to respond “yes” to the presence of a signal. After extended periods, a surveyor may simply form conclusions (in the form of biases) regarding the presence of contamination. In such a circumstance, failure to identify of signal may decrease the expectations of further signals being encountered, decreasing sensitivity over successively longer periods of time. Although some studies have produced positive findings, these have been associated only with tasks involving high event rates and successive determinations. In other studies, a vigilance decrement fails to appear. Regardless, changes to the probability of correct determinations are possible, but are subject to several factors, including a surveyor’s expectations for contamination, training, and experience. To date, no vigilance studies have been conducted which utilize similar output to a radiological survey instrument [1].

The decision process of a surveyor is derived from research generating a model describing an “alerted monitor”. This model utilizes a two-stage decision process for a radiation surveyor, which uses information from one detection process to aid in the decisions encompassing a second, subsequent process. This model describes a two-stage decision process of a surveyor as well, where continuous movement of a detector and the perception of signals corresponds to the first stage. As described by Brown & Abelquist

(1998), sensitivity is usually relatively low. However, information from this stage is used to influence decisions for the second stage, corresponding to longer pauses over an area of interest, and therefore, higher sensitivity and reduced false positives [1].

An ideal observer is one which makes optimal use of the information presented during an experiment to maximize a goal (in most cases, maximizing the number of hits). To determine the performance of an ideal counting observer for an audio detection task, the presentation of information (the signal and signal plus noise) is assumed to have a Poisson distribution, with each represented by two overlapping distributions. Using probabilities which correspond to the cumulative Poisson distribution [12] and the corresponding means of the two distributions (based on the associated observation intervals and number of counts presented to an observer), an ROC chart indicating the performance of the ideal observer can be created. The ROC chart for an ideal observer given a 60 Counts Per Minute (CPM) background and 180 CPM source counts is provided in Figure 4, recreated from the procedure and example values used in Brown & Abelquist (1998) [1]. The difference in curves associated with the first and second stages are due to differences in the observation interval, which correspond to an increased number of counts registered by an observer (within 1 and 4 seconds, respectively). Observation intervals of the first stage are typically on the order of one to two seconds, whereas the interval of the second stage is described as being several seconds long [13].

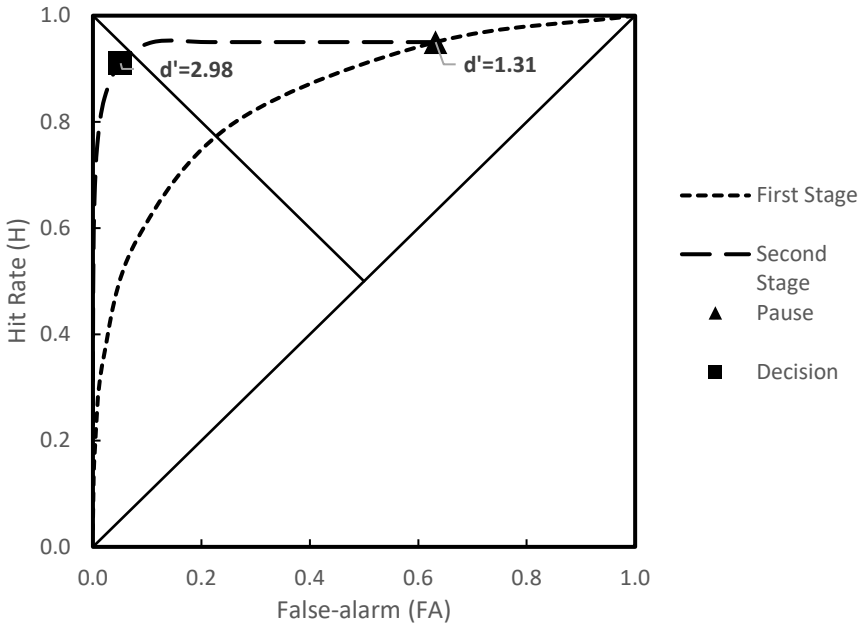


Figure 4. ROC Chart of the Ideal Observer

The *d'* value for the ideal observer under the conditions given is 2.98, corresponding to a surveyor's sensitivity during the final decision at the specified count rates. While the proportion of H is nearly identical between the two stages, the second stage is characterized by a significantly lower rate of false alarms [1].

Simulation of Human Processes and Discrete Event Simulation

Discrete Event Simulation (DES) is a method of modeling a system as it evolves over time. A system, defined broadly, is simply a collection of people or machines that interact. Events are occurrences that serve to change the state of that system [14]. DES can be applied to nearly any system with tasks that can be represented as instantaneous, discrete events. DES has been applied to a range of industries including healthcare [15], defense [16], and manufacturing [17]. The advantages of DES simulation in modeling include reduced time and cost compared to real-world, physical representations.

Additionally, DES allows for the evaluation of the current state of a system and the ability to evaluate future changes by modifying certain parameters. DES allows systems engineers to make changes to a system to optimize or meet a desired metric of performance or evaluate the impact of automation or augmentation on human performance. Depending on the application, however, DES models can require a large amount of data to construct, and even more importantly, do not apply to every circumstance [15]. Thus, DES models are often purpose built for specific applications and with specific outputs in mind at the time of construction.

An alternate type of simulation, Agent-Based Simulation (ABS), is considered a special type of DES [14]. In ABS, the interactive effects between entities are considered. This is important since in a human-machine system, the humans may interact with each other. ABS models are often applied to matters where human behaviors are considered major factors[18]. However, relatively simple models of human behavior, where the individual attributes and behavioral characteristics of each individual are not of concern for the process being modeled, can be described sufficiently by DES [14].

DES models are well suited for modeling human performance and for measuring workload [18], [19], [20], [21]. While "workload" can refer to several different physical processes, it is mental workload that is often of interest. Indeed, where human productivity is of concern, so should be the consideration for mental workload that is devoted by a person at any particular time. Multiple resource theory (MRT), developed by Wickens (1984), introduces four primary components describing resource processing: Visual, Auditory, Cognitive, and Psychomotor (VACP). For any task performed, there is some demand from one or more of these available resources [22]. Rating scales for each

component have also been developed to aid in the standardization of applying MRT to a wide range of tasks.

The assignment of VACP values to a task is best done with the help of Subject Matter Experts (SMEs) familiar with the task and process being modeled. SME's also provide valuable information about the amount of time a particular task may take, the associated distribution of time or accuracy, and several other important parameters which may assist an analyst when constructing the model [20]. Simulation models built to monitor workload provide maximum values which, when exceeded, should be analyzed further. An exceedance of allowed resources in any one channel may lead to task degradation or failure. Several workload studies have suggested an overall VACP value of 40 as the threshold for human performance [23]. This threshold is typically only exceeded if two tasks run concurrently and require the same resource (i.e., two tasks requiring a shared visual resource are performed simultaneously). An alternate assessment of workload developed by the National Aeronautics and Space Administration (NASA), known as the Task Load Index (TLX), has also be used for measuring and predicting human performance [23]. A comparison between the two methods of monitoring workload is outside the scope of this research, but has been studied previously [24].

The United States Army Research Laboratory funded the development of a DES with workload and human performance modeling capabilities known as the Improved Performance Research Integration Tool (IMPRINT) [25]. IMPRINT was designed to incorporate the VACP process developed from MRT and closely integrate the basic elements of a standard DES, including the ability to monitor task times, incorporate

probability and logic based task pathways, as well as incorporate the impact of various stressors of interest to the military on task performance [26]. IMPRINT divides the VACP method into 7 channels (visual, auditory, cognitive, fine motor, gross motor, speech, and tactile). These values are shown in Appendix B – VACP Values and Descriptors. IMPRINT is particularly focused on military operations, incorporating branch-specific occupational series codes and terminology and stressors of particular interest in military operating environments. It has been used to model driver workloads using 13 volunteer subjects for model input [23] along with understanding how task shifting between personnel in a large, multi-shift operations center occurs with multiple operator types [27]. In both studies, the importance of applying good quality input to the model was highlighted. Most importantly, both studies highlighted how additional data should be gathered to better inform the model to make additional predictions of outcomes for future work (rather than applying existing data to drive new conclusions).

Several studies have evaluated the complex task of human-machine automation optimization in DES models. Operator workload is a key element to consider when incorporating automation strategies, where some workload associated with a task is offloaded to an automation device, and then brought back to an operator at some later time as would be the case with machine augmentation, (e.g., an aircraft autopilot system).

Rusnock and Geiger (2017) utilized an IMPRINT threat detection task model to analyze VACP workload values and determine the best strategy for dividing the work of a particular task while considering the situational awareness of the operator [19]. The researchers found that depending on the goal, different strategies need to be employed to optimize workload or task times.

The validity of a simulation model is determined by how closely it is to being representative of an actual system being studied. Constraints on resources limit the extent of model validation, and several methods and resources can be used to help a model designer decide if a model is valid, and to the extent required. A simple model of a well observed set of tasks may not require the same amount of validation as a more complex task. Verification is a term used to describe whether the model's design assumptions have been sufficiently incorporated into the simulation model [14].

The Minimum Detectable Concentration (MDC)

The basic definition of the MDC was given in 1968 by Dr. Lloyd Currie in the work titled "Limits for Qualitative Detection and Quantitative Determination" [28]. This work was primarily driven by the complexity and roughly equivalent meaning of several other quantities, including the lower limit of detection (LLD) and minimum detectable true activity. At the time, there was some inconsistency with which these units were applied and exactly what they referred to. The related quantity *Minimum Detectable Amount* (MDA) is also used nearly interchangeably but has several key differences. The MDC, defined differently from the detection limit, is now defined as the minimum activity concentration on a surface or within a material volume that an instrument is expected to detect [13]. This value is meant to be determined *a priori*, or before measurements have been collected.

To extract every significant use of the MDC in literature would be an arduous task. The continued confusion of nomenclature by researchers, industry, and government sources makes that task even more difficult. In the U.S, the U.S. Nuclear Regulatory

Commission has published several guidance documents which serve to standardize the way in which MDC is applied. Nuclear Regulatory Guide (NUREG) 1507, *Minimum Detectable Concentrations with Typical Radiation Survey for Instruments for Various Contaminants and Field Conditions,* serves as the best resource for determining the process of assigning an MDC value to a particular survey and instrument combination. The need for such standardization was driven from changes to the regulations promulgated by the U.S. Nuclear Regulatory Commission governing the decommissioning of facilities by licensees [13].

A statistical framework is important when describing the basic MDC equation. For a response to radiation, two distributions are presented to an observer (as was the case in the discussion of SDT). A blank (or background distribution) can be subtracted from a measurement of a sample to determine the net counts of a sample of interest. Assuming the underlying distribution of this count frequency is normal (or approximates the standard normal (z) distribution), the associated mean (μ) and standard deviation (σ) can be determined. This procedure determines the so-called critical level, or L_C, used to drive the conclusions regarding the surveyor's acceptance of risk of committing a Type 1 (False Positive) error. A Type 1 error is analogous to the term "False Alarm", used earlier when defining the framework of single factor SDT.

When added to the quantity corresponding to the probability of committing a Type II error (or a Miss in SDT terminology) and the standard deviation associated with the net count, the value of the detection limit (L_D) is determined. Figures 5 and 6, reproduced from NUREG 1507, describe the relationship between L_C, L_D, and the two types of error.

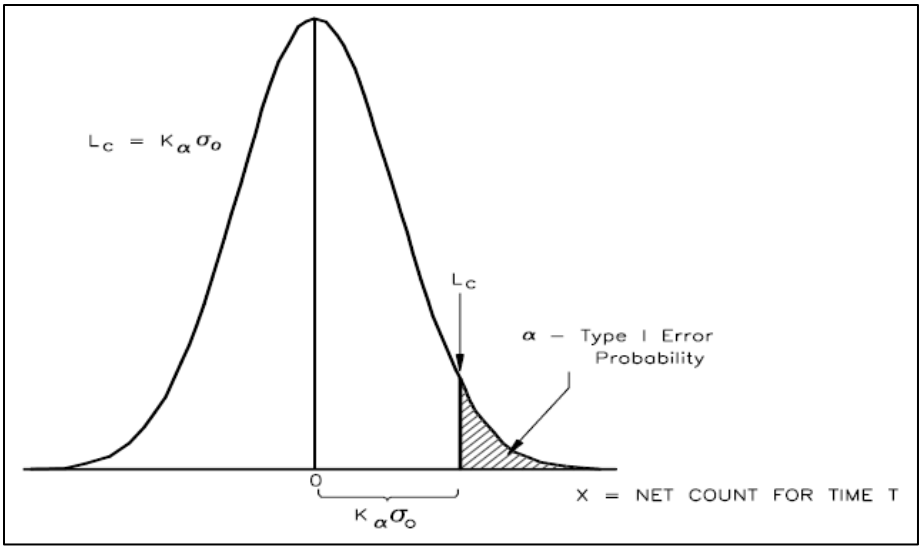


Figure 5. The Critical Level, L_C [13]

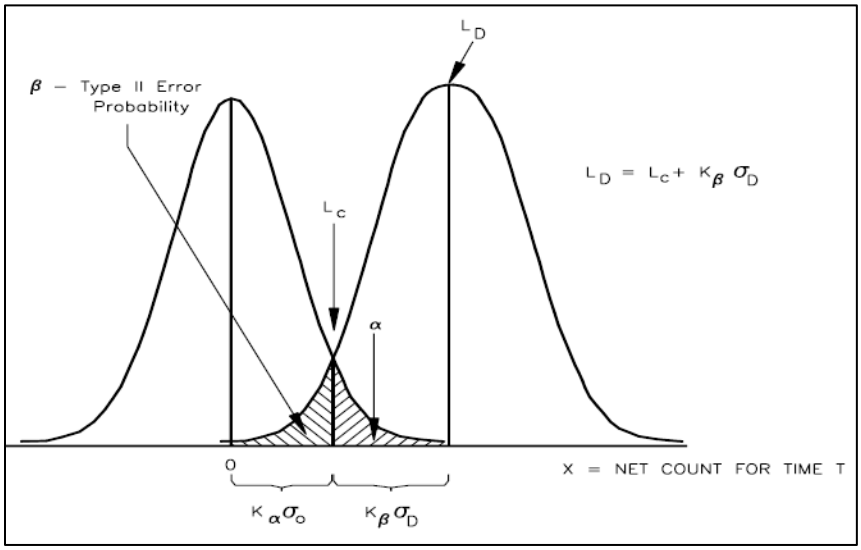


Figure 6. The Detection Limit, L_D [13]

As shown in Figure 5, the critical level (L_c) is equal to the quantity K_α, or the value of the standard normal deviate associated with the 1 tailed probability of $1 - \alpha$, multiplied by the standard deviation of the count data distribution (σ_0). In many cases, the value of K_α is set equal to 1.645, corresponding to an α value of 0.05. This value

corresponds to a 5 percent probability of incorrectly concluding that a sample contains material when it in fact does not (a Type 1 error). Figure 6 describes the detection limit (L_D), equal to that of the critical level plus the quantity of the product of K_β (the value of the standard normal deviate associated with the 1 tailed probability of $1-\beta$) and the standard deviation of the net sample counts (σ_D). Like K_α, the value of K_β is also often set to 1.645 but can be adjusted depending on survey objectives.

For the purposes of this research, the general MDC equation for static counts is therefore defined as:

$$MDC = \frac{3 + 4.65\sqrt{C_B}}{KT} \tag{7}$$

Where:

$$\sqrt{C_B} = number\ of\ background\ counts\ over\ time\ T$$

$$K = Proportionality\ constant$$

$$T = Count\ time$$

The quantities of 3 and 4.65 are constants represented by the Type 1 and Type 2 error probabilities. The value of (α) is customarily set at 0.05, corresponding to a 95% confidence level, but may be set lower. While the above equation is technically correct, several other factors can influence the value of the MDC in the field. These factors, including geometry, surface type, and the various types of efficiency fractions related to both the source and the detector should be considered. The MDC of a scanning, or

moving, detector includes additional values related to a surveyor's technique and accounts for the fraction of time the sensitive area of a detector is placed in "view" of a source [13]. As was the case during static conditions, the scan MDC is an *a priori* quantity; it is meant to be determined before measurement begins to demonstrate that the combination of instrument and techniques can meet the predetermined data quality objectives (DQOs) of the survey. The term DQO is used extensively to support decommissioning and license termination surveys in the U.S but can be interpreted more broadly to apply to any survey conducted with a specific purpose and set of goals. Since virtually all surveys are conducted with some sort of predefined set of goals, the MDC is considered (whether formally or not) when implementing an appropriate combination of instrument and surveyor techniques.

For a scanning survey, the Minimum Detectable Count Rate (MDCR), is the minimum detectable count rate (net) for a surveyor, given as:

$$MDCR = d' * \sqrt{b_i} * \frac{60}{i} \tag{8}$$

Where:

$d' = index\ of\ sensitivity$

$b_i = number\ of\ background\ counts\ in\ the\ observation\ interval$

$i = observation\ interval$

The MDCR equation given above can be divided by an additional term, ρ, to account for the performance of a less than ideal observer.

$$\rho = \left[\frac{d'_{actual}}{d'_{ideal}}\right]^2 \quad (9)$$

Where the two *d'* terms are identical to those discussed previously in the review of SDT. The quantity of ρ, a value less than 1, accounts for the inefficiency of an average surveyor relative to the ideal. NUREG-1507, along with NUREG/CR-6364, suggest using an *a priori* ρ value of between 0.5 and 0.75 [13]. This value is chosen by survey planners based on several factors, including the prior experience of those involved. However, the assignment of this value is largely subjective in cases where empirical data on the performance of surveyors relative to the ideal is not known. Brown & Abelquist (1999) suggest not using ρ values over 0.5 when surveys are to be performed in field conditions [1]. Field conditions refer to instances where extrinsic factors, such as the external environment, terrain, and Personal Protective Equipment (PPE), may impact a surveyor's ability to maintain sufficient detectability levels relative to the ideal case. The square root of ρ is used in the final MDCR equation, accounting for the inefficiency of a surveyor:

$$MDCR_{Surveyor} = \frac{d' * \sqrt{b_i} * \frac{60}{i}}{\sqrt{\rho}} \quad (10)$$

Where:

$d' = index\ of\ sensitivity$

$b_i = number\ of\ background\ counts\ in\ the\ observation\ interval$

$i = observation\ interval$

$\sqrt{\rho} = square\ root\ of\ surveyor\ efficiency$

Lastly, the complete scanning MDC equation is given as:

$$Scan\ MDC = \frac{d' * \sqrt{b_i} * \frac{60}{i}}{\sqrt{\rho} * \varepsilon_{inst} * \varepsilon_{src}} \quad (11)$$

Where:

$$\varepsilon_{inst} = Instrument\ efficiency$$

$$\varepsilon_{src} = source\ efficiency$$

The two terms represented by ε account for the inherent properties of both the detector and the source. In the case of the detector, it, like a surveyor, does not register every counting event with perfect accuracy. Source efficiency is related to the physical properties of the material being surveyed and can vary significantly (wood, concrete, soil, etc.) [13].

Radiological Surveillance and the MDC

Few studies have been performed previously which analyzed the performance of a surveyor to discriminate audible tones within a radiological context. Brown and Abelquist (1998) conducted a series of trials in both controlled and field settings to observe the performance of radiological surveyors. The results of this report form a portion of the methodologies adapted in several joint-federal publications which standardize radiological decision making for licensees in the U.S, including Nuclear Regulatory Guide (NUREG) Report 1575, *Multi-Agency Radiation Survey and Site Investigation Manual (MARSSIM)* and NUREG Report 1507, *Minimum Detectable Concentrations with Typical Radiation Survey for Instruments for Various Contaminants and Field Conditions* [13],[29].

Tzelgov et al (1987) evaluated operator performance when utilizing hand-held radiation monitors. While the primary intent was to determine whether performance was impacted by audible or visual indications (or a combination of the two), the results only indicated better surveyor performance when using the audible display of an instrument during a search (rather than detection) task. The researchers concluded that visual dominance plays a role when both visual and audible indicators are used simultaneously [30].

Casey (1991) evaluated radiological survey and sorting processes during 3 studies at two nuclear power plants. Workers screened and sorted through contaminated items, while their relative performance, in terms of hit rate and false alarm, was measured. Sensitivity (in terms of *d'*) was documented amongst the surveyors. Results of the study indicated a relatively high percentage of items correctly classified, at the cost of many false alarms (0.44). Measures of *d'* ranged from 1.58 to 1.67, despite achieving true positive proportions of nearly 1 in several studies. Thus, the choice of a surveyor's criterion during a survey can significantly affect the conclusions made. A surveyor may be more willing to declare an item to be contaminated (a false alarm), since the cost of this decision (more time) is balanced with the risk of mischaracterizing the material in question [31].

King et al (2012) evaluated the changes in MDC during environmental radiological surveys using common portable equipment when subjected to "pendulum-like" motion common during walk over surveys. Ideally, the detector is to be held in a low, flat trajectory perpendicular to the ground, to best approximate the geometry used during calibration and maximize efficiency. However, certain patterns (including motion

and probe velocity) were found to have a significantly negative impact on the MDC (up to 47% under worst case conditions). Results of the study highlight the importance of surveyor techniques and the significant variability associated with surveyors who may not account for the impact of these movements. The study specifically evaluated environmental surveys which employ portable gamma scintillator detectors primarily intended to measure small amounts of radioactivity in soil or deposited on the ground, and did not employ human subjects for comparison to modeled results [32].

Falkner and Marianno (2021) evaluated the relationship between detector speed and the Minimum Detectable Activity (MDA) as it relates to detection efficiency. A model was validated which provided a more accurate representation of detector efficiency while moving (as would be typical for surveillance tasks). A robot was used to perform scans of actual sources, and the results used to determine if the predictive model provided a more accurate estimation of the MDA than classical methods (which assume static counts). Results indicated agreement with the predictive model, but did not present results within a statistical context, but rather qualitatively described the shape of the curves produced. However, the recent study verifies the impact of velocity on detection performance, and is analogous to the change in the scanning minimum detectable concentration of a surveyor under different observation intervals. [33]

Summary

This chapter presented the underlying framework used to describe a typical radiological detection task, found to be explained by a simple one interval signal detection task and associated quantities. Basic quantities in signal detection are well-

suited for describing the choices made by a radiological surveyor, describing both bias (as a criterion value) and overall sensitivity (measured as *d'*), arguably two of the most important factors to consider when describing surveyor performance. Also summarized was the use of discrete event simulation software to analyze human performance and associated workload, as well as previous use to model highly complex tasks, including those that involve human-machine automation techniques. Finally, the background necessary to describe the concept of the Minimum Detectable Concentration (MDC) was presented, including the MDCR and Scan MDC equations, incorporating values of *d', i,* and ρ. Previous research evaluating the performance of radiological surveyors was summarized, including those that sought to evaluate how performance can be decreased under certain operator conditions. Few studies have evaluated changes to the MDC by describing the individual stressors that affect a surveyor during a radiological detection task.

III. Methodology

Chapter Overview

The purpose of this chapter is to describe the various tools used during this research and to present the methods of data analysis. Both commercially available and limited export-controlled software tools were employed to analyze human performance data during radiological surveillance tasks to construct a model using DES which simulates changes in performance under certain environmental conditions. The results of

this output are related back to the radiological MDC paradigm to quantitatively estimate the impact on a surveyor under these conditions.

Test Subjects

Human performance data from research conducted as part of Brown and Abelquist (1998) and reported in NUREG CR/6364, *Human Performance in Radiological Survey Scanning* was used in this research. Results of surveyor performance during a radiological detection task using a common Geiger-Mueller based survey instrument was obtained by the author. In total, the performance of 6 volunteer participants, employed with the Oak Ridge Institute for Science and Education (ORISE) and conducted in March of 1995, was analyzed. Observed performance data was used to drive the probabilistic path logic of the IMPRINT simulation models.

Scanning and Measurement Methodology

Data was provided in terms of common observation measures, i.e., number of hits, misses, false alarms, and correct rejections, for each participant. As described in NUREG/CR-6364, a 1.5 meter by 5-meter test surface was created on a wall which covered a total of 16 sources corresponding to 9 separate areas of contamination. Survey participants were instructed to scan the area slowly, but otherwise received few instructions prior to performing the task. The prior experience of the surveyors to perform radiological surveys ranged from zero to several years. Locations and number of pauses and decision points were recorded by an independent observer. These observations and corresponding measures of sensitivity are provided in Appendix C.

While observations corresponding to the common SDT measures were collected, limited information concerning length of pauses and probe velocity is available. At the time of observation, tools were not available which allowed for the collection of information which would make detailed analysis of this task possible in the future. Brown (2000) described how practical constraints (such as time allotted) might affect radiological survey performance [34], however for the portable Geiger-Mueller survey, no time limit was prescribed. The absence of this data in units of time makes estimation of changes in total survey time (a desired measure of surveyor performance) impossible. Correspondingly, the time a surveyor took to indicate a hit or false alarm would not be known. Fortunately, the observational data includes sufficient information about the total survey area, probe width, number of pauses, and prescribed velocity from which the length of pause in each stage can be estimated for each of the 6 surveyors.

To do so, several assumptions must be made. The total survey area is given as 1.5 x 5 meters, or 75,000 square centimeters. The G-M survey probe used (an Eberline HP-260) has a probe width (housing width) of 2.75 inches, or 38.32 square centimeters. Assuming the surveyors moved the probe across the survey area at a velocity of 1 probe width per second *consistently* (as indicated in NUREG/CR-6364) and did not overlap or resurvey previously scanned areas, the length of time the probe is *not stationary* is equal to the total survey area divided by the probe width, or about 32.67 minutes. Knowing the total survey time, the amount of time the probe *is stationary* can be inferred by the difference between the two quantities.

An assumption also needs to be made based on the observed criterion and the process of flagging a source due to the longer observation interval of the second (final)

stage. Therefore, it is assumed that the fraction of time spent on pauses in the first stage is 0.25 (25%) of the total. Similarly, the fraction of time spent deciding a source is present in the second stage is 0.75, or 75% of the total time per pause. A table of results for each surveyor can be seen in Table 2. Results are provided in units of seconds unless otherwise indicated.

Table 2. Time and Pause Length Calculations

Time and Pause Length Calculations							
Number of Pauses		71	140	104	156	91	91
Time (min)		50	54	47	72	70	63
Survey Area (cm2)	75000						
Probe Width (cm2)	38.26						
Probe Velocity (pw/s)	1						
# of sec based on Probe Velocity	1960.27						
Length of time probe is moving (min)	32.67						
Length of time Probe is stationary (min)		17.33	21.33	14.33	39.33	37.33	30.33
Length of Pauses (sec)		14.64	9.14	8.27	15.13	24.61	20.00
Pause Length (1st Stage)	0.25	3.66	2.29	2.07	3.78	6.15	5.00
	Mean	3.82					
	St Dev	1.57					
	95% CI	1.65					
Pause Length (2nd Stage)	0.75	10.98	6.86	6.20	11.34	18.46	15.00
	Mean	11.47					
	St Dev	4.70					
	95% CI	4.94					

Surveyor Data

The SDT framework discussed in Chapter 2 serves to put both observations of surveyor performance and the results obtained following simulations using IMPRINT in terms of common units so potential impacts to the MDC can be determined. Figure 7 contains both theoretical values of *d'* and *c*, with the observations from the surveyors plotted according to their HR and FA rate. The area and points in green correspond to the initial (pause) stage; the area shaded in red and red points corresponds to the detection (final) stage. Generally, good agreement is found between the different surveyors in

terms of their overall HR/FA rate, regardless of prior experience level (which was not well known) and variations in total survey time. However, each of the surveyors adopted varying choices of criterion at each stage, within about 0.5 standardized units. Surveyors tended to adopt a liberal detection strategy in the first stage of scanning, tending to respond "yes" or "hit" more often than the ideal (unbiased) observer. Correspondingly, all the surveyors adopted a conservative approach to the final stage, leading to a much higher measure of discriminability. This is likely due to the presence of prior information from which to make an enhanced judgement and a surveyor benefiting from a longer observation interval and therefore more counting information. This behavior is also described by Brown and Abelquist (1998).

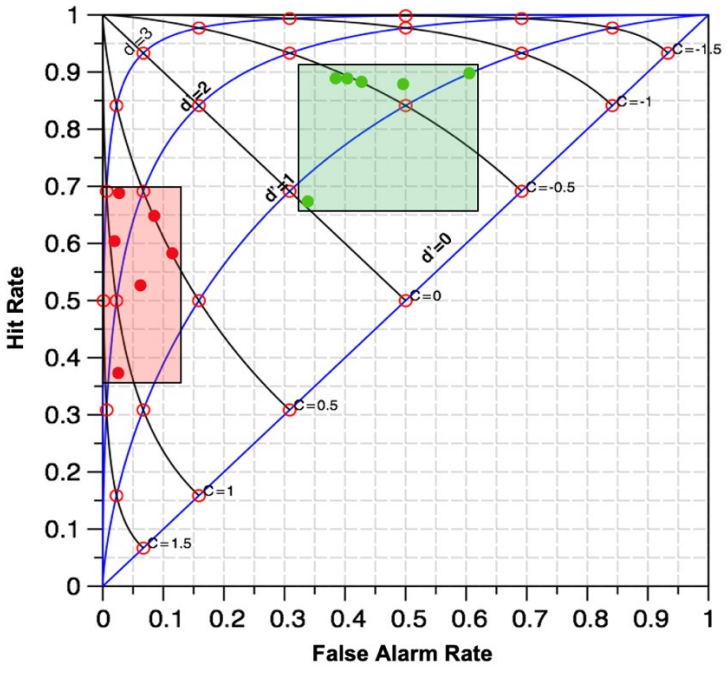


Figure 7. ROC Graph of surveyor observations with sensitivity and bias curves [11]

According to the framework in modern SDT, for each surveyor and the associated measure of sensitivity to a signal, there exists a range of possible criterion values. This describes the behavior identified by Brown and Abelquist (1998), who similarly described several conditions with different HR/FA rates corresponding to equal values of *d'*. While sensitivity remains constant, the location of the criterion will determine which responses are reported and which are not [1]. Figure 7 explains this graphically. It may be expected, therefore, for two surveyors to be equally sensitive in registering the presence of contamination, but each may choose to adopt a different criterion which affects their response.

Calculation of common SDT Measures

MATLAB 2021b (The MathWorks Inc, Natick, Massachusetts)[35] was used to run a standalone application which calculates the common SDT values of interest, including *d'*, β, and *c*. This software was used to generate ROC and probability plots corresponding to each of the surveyors, as well as to inform plots and graphs corresponding to model results from IMPRINT [11]. ROC plots were also generated manually using Microsoft Excel for Mac version 16.57 (Microsoft Corp, Redmond, Washington). Plots and graphs from each of the surveyors can be seen in Appendix D.

Modeling Human Performance using DES

Human Performance Modeling (HPM) can be used to inform the design of a system to better optimize the human-machine interface and enhance performance. For this research, the Improved Performance Research Integration Tool (IMPRINT) Pro version 4.6.60 (Alion Science and Technology, Boulder, Colorado), developed for the

Army Research Laboratory (ARL) Human Research and Engineering Directorate (HRED), was used to build a simulation model which closely approximates the actual performance of a surveyor during a radiological detection task. The model uses probabilistic pathways informed by the Hit and False Alarm rates of a surveyor over a two-stage detection task. The model and task times are informed by a combination of observations (observed time for each surveyor to complete the exercise) and the duration of time the probe is held stationary (either in the pause or decision stages).

The model consists of 18 task (event) nodes each corresponding to the decision process a surveyor undergoes during a basic radiological detection task involving portable survey equipment. Several nodes exist to align the process with the larger detection task. For example, "Orient Probe" does not have any task time or probability associated with it and serves only to indicate the logical sequence of events in terms more familiar to those familiar with radiological detection. Several nodes (highlighted in red) are physical tasks performed by the surveyor (pause, stop). These nodes have corresponding task completion times (on the order of several seconds each) that drive the probabilistic path logic that follows it.

Input values for the probabilistic pathways are informed by the observational data collected by Brown & Abelquist (1999) for a detection task using a portable Geiger-Mueller based survey instrument. The values for each of the surveyors, as well as a theorized "average" surveyor based on the average performance of all six surveyors, are provided in Appendix E. A full list of data collection requirements by task node can be seen in Appendix F.

Figure 8 displays the layout of the IMPRINT model for the initial "Pause" stage. The model begins with a surveyor moving the probe across the surface to be scanned, which has an associated time duration defined according to the survey and detection methodology discussed in the previous section. The next task node, "Pause", corresponds to the initial decision point of a surveyor. Events from this node flow to two possible outcomes: Pause on a source or Pause in Background. Both are assumed to be equally likely random events if the surveyor receives no other information about the location of contamination. This is a reasonable assumption under most conditions but may not be applicable in every circumstance. To account for the equally random events, the preceding "Pause" node is of the "Multiple" type, indicated by an "M" besides the task. With this type, both task nodes that follow ("Pause on a Source" and "[Pause] In Background" both run simultaneously and with a probability of 1. The subsequent events are then associated with a probabilistic outcome, defined according to the basic outcomes described in SDT outlined in Table 1 (i.e., a hit, miss, false alarm, or correct rejection).

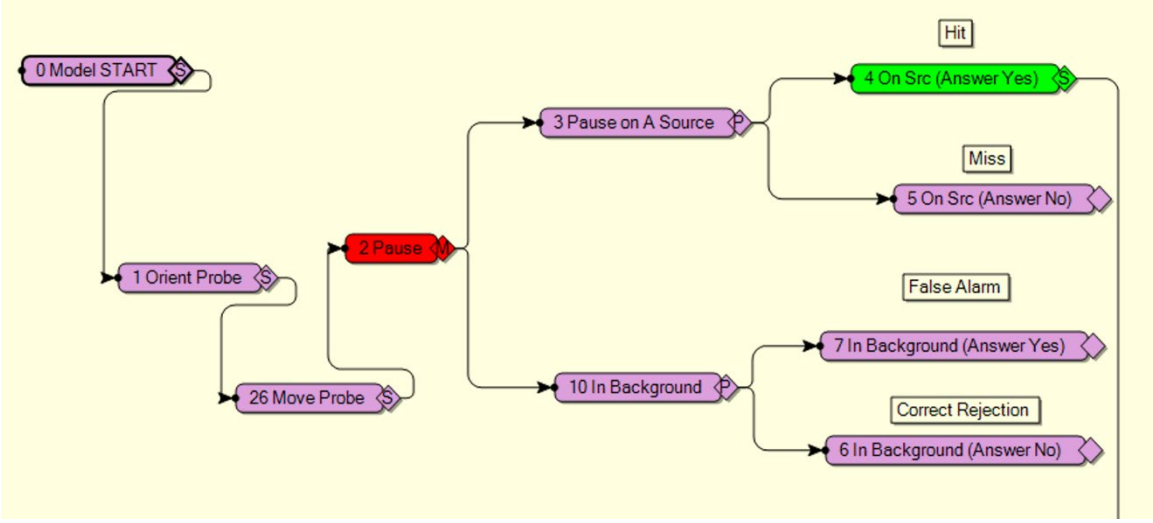


Figure 8. IMPRINT Network Diagram - Pause Stage

The second stage "detection" follows if, and only if, a positive "hit" is made from the pause stage and can be seen in Figure 9. This represents the concept of the alerted monitor and two-stage decision process described by Brown & Abelquist (1999). The flow of logic and probabilistic pathways is identical to the pause stage. In most cases, however, the rate of false alarms is significantly lower than in the first stage. The hit rate for all six observers in the second stage is always equal to or lower than the first, revealing an upper limit based on the sensitivity and decision choices of the first stage.

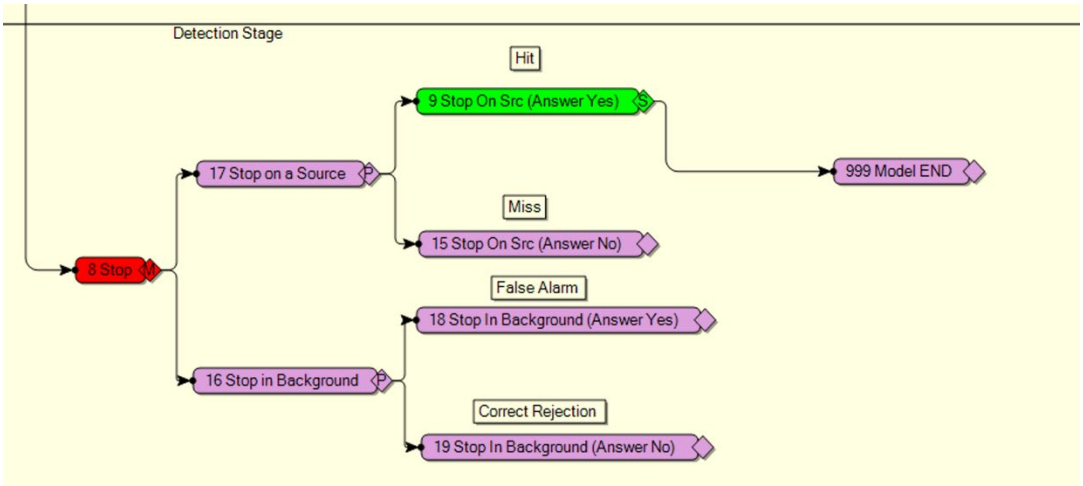


Figure 9. IMPRINT Network Diagram - Detection Stage

IMPRINT can accept task durations, accuracy, and path logic in a number of forms, including based on those approximating a particular parametric distribution (i.e. normal, lognormal, triangular, uniform). The simulation model for each surveyor, however, is based on direct observations of times and probabilities associated with hits and false alarms. From such a limited data set, it is difficult to ascertain information about an underlying distribution of the data, nor could one reliably be inferred. For this reason, model inputs are thought to be observed quantities, which eases the complexity of

model validation. The goal, therefore, is to obtain output values consistent with results observed from the surveyors using input parameters resulting from direct observation.

The output from IMPRINT is based on the number of times each task node was performed. Tasks associated with a higher probabilistic rate of occurrence are also those which run more often than those with a lower probabilistic rate of occurrence. To put these values in terms of *d'*, the hit and false alarm rates for each stage were manually calculated from the resulting task performance report generated following each simulation run. Table 3 describes the calculations performed using Microsoft Excel to calculate the values of *d'* from each stage. Values referenced correspond to task node numbers shown in Figures 8 and 9, except where noted (see notes 1, and 2).

Table 3. Calculation of d' from IMPRINT results

		Pause Stage	**Decision Stage**
Hit Rate	=	=4/(4+5)	=9/(9+15)
False Alarm Rate	=	=7/140[1]	=IF(18>0, 18/(7+18), 0.5[2]/(7+18))
z(H)	=	=NORMSINV(HR)	=NORMSINV(HR)
z(F)	=	=NORMSINV(FA Rate)	=NORMSINV(FA Rate)
d'	=	=z(H)-z(F)	=z(H)-z(F)

[1] Fixed value corresponding to the number of opportunities for false alarms in the first (continuous) stage. This value is derived and described by Abelquist & Brown (1998) and used for this study.

[2] As described by Hautus et al (2020), False alarm rates of either 0 or 1 imply an infinite value of *d'* and yield an erroneous *d'* value when calculated according to the procedure above. To compensate for this, when the number of false alarms in the second stage equals exactly zero, a value of 0.5 is used as the dividend to produce a false alarm rate slightly greater than 0. This technique has been reviewed and results in FA rates just slightly over 0, negligibly impacting the accuracy of the calculated value of *d'*.

Nodes which have either estimated duration or inform probabilistic decision logic have VACP values corresponding to Table 8 (Appendix B) associated with them. In IMPRINT, VACP values are assigned by resource according to the rating scale. Figure 10 is an example of the workload input for a detection task. Note that "CrewStation" is a

default interface in IMPRINT and cannot be removed. It can, however, be deselected and excluded from analysis (as was the case during this research).

Interface	Used Here	RI Pair Demand Values						
		Auditory	Cognitive	Fine Motor	Gross Motor	Speech	Tactile	Visual
{}Signaling Device{}	☐							
1. GM Survey Instrument	☑	7.00	4.60	4.60				6.00
CrewStation	☐							

Figure 10. Workload Demand Tab in IMPRINT

A key difference between IMPRINT and many other DES software packages is the ability to model human performance and include factors which approximate their effect on task execution. IMPRINT does this by employing moderators which modify the task by either changing the accuracy with which a task is performed or the amount of time a particular task took to execute. A variety of default Performance Shaping Factors (PSFs) are available which modify tasks based on the environment (cold, heat, noise, vibration, etc.) or the physical strain endured (MOPP, Level A, or weight load). The data used to drive these PSFs are based on existing research and studies of human performance [26]. Multiple PSFs can be applied at once, and are weighted based on the VACP rating assigned [26]. When summed, measures of VACP by task can provide a workload profile that captures variation over time. Workload monitoring is one of the primary ways that IMPRINT has been utilized in literature. The information gathered from an analysis of workload according to the VACP scale can provide information on personnel utilization, idle times, and effective strategies for distributing workload at various times. While this functionality is included in IMPRINT, a limited amount of task duration information from which to apply to the model was collected in the original

observational work. Additionally, the simulation model is of extremely short duration (typically less than 30 seconds each), which serves to model the task from detection to detection (or pause to pause). Workload levels within such a short duration are difficult to interpret in a meaningful context and are better suited for more complex tasks with more detailed time duration inputs over longer durations.

To accurately reproduce the performance of an actual surveyor, the model simulation is run equal to the number of times the surveyor paused. This was performed in IMPRINT to create baseline models used for validation (discussed below) and for the creation of alternate models, from which PSFs were applied to simulate the effects of a stressor of interest. To evaluate surveyor performance in alternate environmental conditions, MOPP PSFs were applied to task nodes 1, 2, 8, and 26, which correspond to decision tasks by the surveyor. For this model, the independent variable is the MOPP level. The dependent, or response variable is the survey completion time. Figure 11 indicates the VACP value choices for the affected tasks. To apply the heat PSF, tasks which have probabilistic logic associated with them were assigned VACP values and mapped according to the 7-channel VACP descriptors, which are termed “taxons” in IMPRINT [26]. For this model, the independent variable is the ambient environmental air temperature. The dependent, or response variable is the measure of surveyor discriminability, as measured by the quantity *d’.* The assignment of VACP values to the alternate model evaluating the effects of heat on a surveyor can be seen in Figure 12. Additionally, a sensitivity analysis was conducted to determine the level of effect alternate VACP values had on model output. The results of this analysis can be seen in Appendix G.

Task ID	Function	Task	Task Demand Values				
			R: Auditory GM Survey Instru	R: Cognitive GM Survey Instru	R: Fine Motor GM Survey Instru	R: Visual {}Signaling Devici	R: Visual GM Survey Instru
0	(Root)	Model START					
1	(Root)	Orient Probe	1.00	4.60	2.60		4.40
10	(Root)	In Background					
15	(Root)	Stop On Src (Answer N					
16	(Root)	Stop in Background					
17	(Root)	Stop on a Source					
18	(Root)	Stop In Background (A					
19	(Root)	Stop In Background (A					
2	(Root)	Pause	7.00	4.60	4.60		6.00
26	(Root)	Move Probe	1.00	4.60	4.60		6.00
3	(Root)	Pause on A Source					
4	(Root)	On Src (Answer Yes)					
5	(Root)	On Src (Answer No)					
6	(Root)	In Background (Answe					
7	(Root)	In Background (Answe					
8	(Root)	Stop	7.00	4.60	4.60		6.00
9	(Root)	Stop On Src (Answer Y					
999	(Root)	Model END					

Figure 11. VACP Values assigned for MOPP

Task ID	Function	Task	Task Demand Values				
			R: Auditory GM Survey Instru	R: Cognitive GM Survey Instru	R: Fine Motor GM Survey Instru	R: Visual {}Signaling Devici	R: Visual GM Survey Instru
0	(Root)	Model START					
1	(Root)	Orient Probe					
10	(Root)	In Background					
15	(Root)	Stop On Src (Answer N	7.00	4.60	4.60		6.00
16	(Root)	Stop in Background					
17	(Root)	Stop on a Source					
18	(Root)	Stop In Background (A	7.00	4.60	4.60		6.00
19	(Root)	Stop In Background (A	7.00	4.60	4.60		6.00
2	(Root)	Pause					
26	(Root)	Move Probe					
3	(Root)	Pause on A Source					
4	(Root)	On Src (Answer Yes)	7.00	4.60	4.60		6.00
5	(Root)	On Src (Answer No)	7.00	4.60	4.60		6.00
6	(Root)	In Background (Answe	7.00	4.60	4.60		6.00
7	(Root)	In Background (Answe	7.00	4.60	4.60		6.00
8	(Root)	Stop					
9	(Root)	Stop On Src (Answer Y	7.00	4.60	4.60		6.00
999	(Root)	Model END					

Figure 12. VACP Values assigned for Heat

Assignment of Performance Shaping Factors

The wearing of Chemical Individual Protective Equipment (IPE), otherwise known as Mission Oriented Protective Posture (MOPP) gear, can be applied to relevant

task nodes. There are a total of 5 levels of MOPP (Including MOPP level 0, where chemical IPE is available, but not worn), each of which are individually selectable for each task. IMPRINT assigns a degradation factor originating from research conducted by the U.S. Army Ballistic Research Laboratory. In generally, performance degradation factors range from 1 to 1.7 for all task types, with a mean of 1.5 over more than 700 different tasks evaluated specifically for MOPP Level 4 [36]. Degradation factors in IMPRINT are applied as multipliers against the time to complete a task, and do not affect the accuracy of task performance. For any task where MOPP 4 is selected, the time duration is multiplied by 1.7 [26]. For other MOPP Levels, degradation factors are assigned reduced values, with MOPP Level 1 equal to about 1 [37]. These values are derived from research conducted by the U.S. Army Ballistic Research Laboratory, which found that the accuracy and performance of a task is unaffected by increasing levels of MOPP. Only the time required to complete a specific task is affected, described as degradation factors applied to the time required to a complete a specific task. A wide variety of tasks (over 700 in total) have been evaluated and maintained in the U.S. Army P2NBC2 database [36].

The heat stressor taxon in IMPRINT uses a combination of the dry bulb temperature and relative humidity to determine a heat degradation factor, which affects the accuracy of task completion only, rather than the amount of time to complete it [26]. IMPRINT has several settings to manage the result of task failure or determine alternate path logic in the event task accuracy falls below predefined thresholds. For this research, since tasks of interest correspond to points of decision by a surveyor to pause or stop over

a source, the failure (or less than 100% accuracy) of that task is registered as a "miss" and manually subtracted from the model output corresponding to a "hit".

Heat degradation factors in IMPRINT are derived from studies of task performance and accuracy conducted in the late 1960s and 1970s. While dated, there is clear evidence of the effect of extreme heat on the performance of sedentary type tasks.

Figure 13, from the *Bioastronautics Data Book* (1973), describes the effect on the number of mistakes over time when performing sedentary tasks over a range of temperatures. IMPRINT includes values from this research applied to the visual, numerical, cognitive, fine motor, and communication domains to derive the proportion of time a task may fail to be conducted accurately.

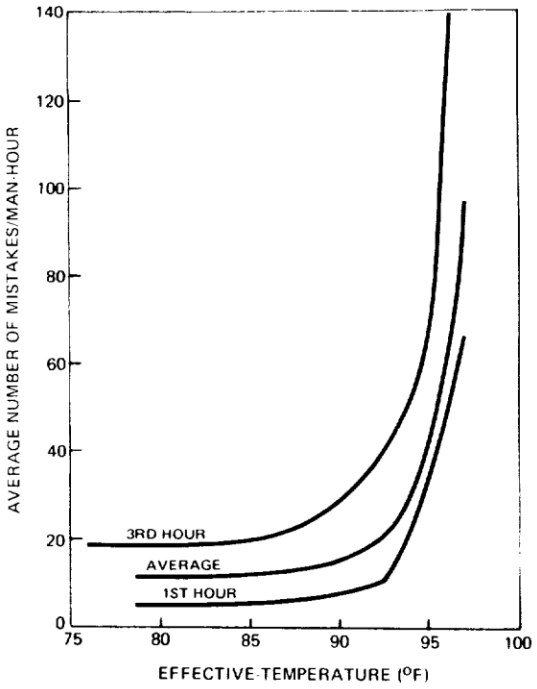


Figure 13. Number of Mistakes Per Hour Performing Tasks in Extreme Heat [38]

For the simulation of the effect of heat on a surveyor, dry bulb temperatures of 94 to 102 degrees and 103 to 111 degrees at 61 to 70% relative humidity were selected based on results shown in Figure 13. The relationship between heat and cognitive performance is well studied [39] and of particular interest to the military community, since operations can take place in any locale and temperature environment. The selected values represent a reasonable range of temperatures and humidity values where radiological surveys are likely to take place in support of various operations, including Deactivation & Decommissioning (D&D) [40] and Chemical, Biological, Radiological, and Nuclear (CBRN) operations [41]. As indicated in Figure 13, after extended periods of time, increased ambient temperature has a markedly increased impact on the number of mistakes made per man-hour.

Baseline IMPRINT Model Verification

Verification and Validation (V&V) of simulation models is a critical step in human performance modeling and discrete event simulation. Validation is achieved when the real-world performance of a surveyor is approximated by the results of the model. The goal, however, is not complete accuracy. Rather, the goal of verification and validation (V&V) is to increase the confidence in the model's output and evaluate its ability to approximate the system being investigated.

For this research, a conceptual model was created based on the descriptions of surveyor behavior and observations of their decision-making process. A two-stage decision model was devised typical for portable Geiger-Mueller based survey instruments as described in Brown & Abelquist (1998) [1]. Both subjective and objective tests were

applied to the verification process. The researcher's prior experience in radiological surveillance and detection tasks was used to subjectively analyze the conceptual model and determine its general accuracy. Objective tests (hypothesis test using confidence intervals defined by the student's t statistic) are used to assess the model's ability to accurately predict an outcome (in this case, *the index of sensitivity*, *d'*). The goal of the test is to determine if accurate measures of a surveyor's performance can be approximated for a radiological surveillance and detection task. The detection performance of each surveyor was used for model validation input. Surveyor 2 was found to have the highest observed value of *d'* amongst the participants and whose confidence interval test is produced below. A 90% ($\alpha = 0.10$) two-sided confidence interval was created to determine if the modeled surveyor sensitivity (as measured by *d'*) is within 0.15 standard units of the observed (expected) value for each surveyor. The value of 0.15 *d'* was chosen by the researcher to be a reasonably close measure of sensitivity while accounting for the model's early stage of development. Within the context of this research, a *d'* value of 0.15 corresponds to a surveyor efficiency value (represented as ρ) of about 8% (using an ideal surveyor *d'* value of 2.98). The test was performed according to equations published in the *Handbook of Tables for Probability and Statistics*, Beyer (2018) [12].

Tables 4 and 5 display the results of 10 independent runs of the model and the mean and standard deviation of both stages. Only the final (decision) stage is used to verify the performance of an observer, as it is in this stage that a surveyor applies all of the information gathered to make a decision. This is standard convention when describing the detectability of a surveyor [1]. The mission was run a number of times equal to the

number of reported pauses by each surveyor with each run utilizing a different random number seed. The random number seed is used during the generation of random numbers used for the probabilistic pathway logic [26].

Table 4. IMPRINT Results - Pause Stage Surveyor 2

Run	Hit Rate	False Alarm Rate	Z Adjust	*d'*
1	0.89	0.51	1.2, 0.02	1.19
2	0.89	0.51	1.2, 0.03	1.17
3	0.89	0.50	1.2, 0	1.20
4	0.89	0.50	1.24, 0	1.24
5	0.89	0.51	1.2, 0.03	1.17
6	0.89	0.50	1.24, 0	1.24
7	0.90	0.51	1.28, 0.02	1.26
8	0.90	0.51	1.28, 0.04	1.25
9	0.9	0.50	1.28, -0.02	1.30
10	0.89	0.51	1.20, 0.02	1.19
Mean	0.893	0.506		**1.221**
Standard Deviation	0.0048	0.0052		**0.0433**

Table 5. IMPRINT Results - Decision Stage Surveyor 2

Run	Hit Rate	False Alarm Rate	Z Adjust	*d'*
1	0.51	0.007	0.02, -2.46	2.48
2	0.51	0.007	0.02, -2.46	2.48
3	0.51	0.007	0.02, -2.45	2.47

4	0.51	0.007	0.03, -2.45	2.48
5	0.51	0.007	0.02, -2.46	2.48
6	0.52	0.007	0.05, -2.45	2.50
7	0.52	0.007	0.04, -2.46	2.49
8	0.52	0.007	0.06, -2.46	2.51
9	0.52	0.007	0.04, -2.44	2.48
10	0.52	0.007	0.04, -2.45	2.49
Mean	0.515	0.007		**2.486**
Standard Deviation	0.005	8.67362E-19		**0.0111**

If $\bar{Y}$ equals modeled output and μ_0 equals the expected (Observed), then:

$$\bar{Y} \pm t_{\alpha/2,,n-1}\frac{s}{\sqrt{n}} \quad (12)$$

Where degrees of freedom (df) is equal to *n*-1, s is equal to standard deviation of the modeled output over n replications, ε is the boundary to be evaluated, and *n* is equal to the number of model replications (10). Then,

$$\bar{Y} \pm t_{\frac{0.010}{2},9}\frac{s}{\sqrt{n}}$$

$$2.486 \pm 1.833\frac{0.011}{\sqrt{10}} = 2.480, 2.492$$

$$\mu_0 = 2.600$$

$$\varepsilon = 0.15$$

$$Upper\ boundary: |2.492 - 2.600| = \mathbf{0.108}\ (Less\ than\ 0.15)$$

$$Lower\ boundary: |2.480 - 2.600| = \mathbf{0.12}\ (Less\ than\ 0.15)$$

Both the upper and lower boundaries of the confidence interval for Surveyor 2 are less than the selected value of ε (0.15). When calculating the boundaries for the other surveyors however, the model fails to approximate values of *d'* within 0.15 units reliably. The greatest difference between observed and modeled results was found to be 0.47 units of *d'*, with an average difference of 0.30 units of *d'* between the observed and upper boundaries generated by the model. Figure 14 displays the observed values of *d'* for each surveyor (indicated by the solid line) along with the upper and lower boundary values of *d'* predicted by the model.

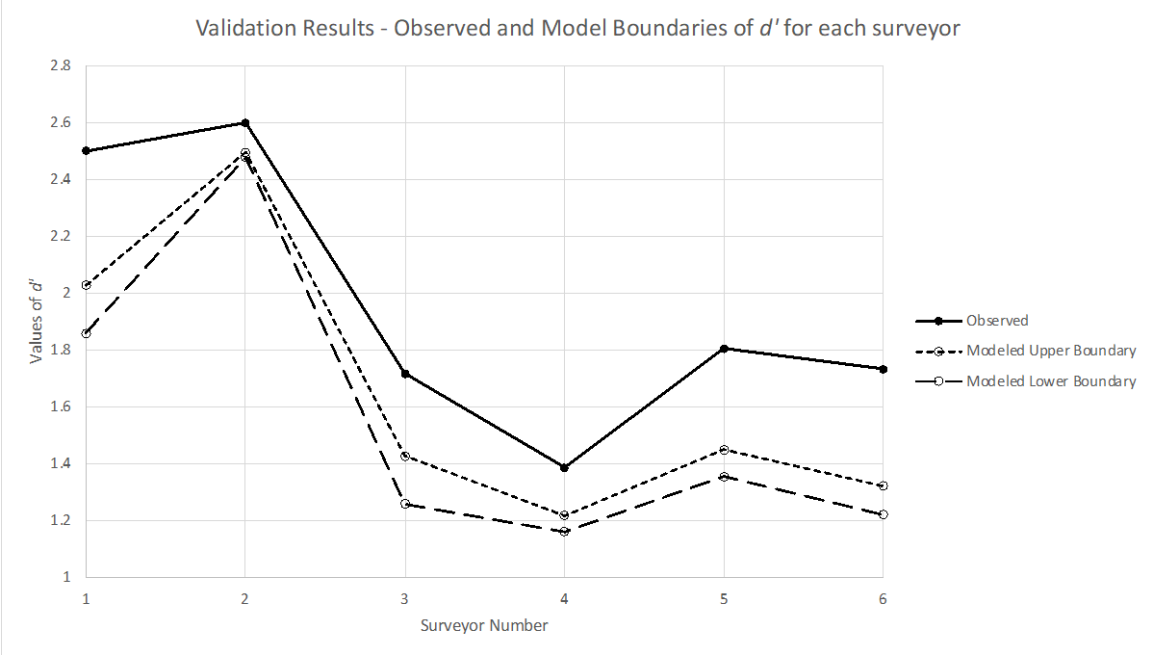


Figure 14. Observed and Modeled Validation Boundary Results

The result of the model generally tends to underestimate the expected value of *d'* by about 20%. While the model can approximate the observed surveyor performance to within 0.15 units of *d'* at the 90% confidence level (as was the case for Surveyor #2), it

fails to do so for the others. In general, the model is found to be effective to within approximately 0.3 units of *d'* amongst all six surveyors. A table of results for the upper and lower boundaries of the model results can be seen in Appendix H. A larger sample set is needed to determine the cause of the variability, and to identify areas of future improvement and model refinement.

Summary

This section described the various tools and techniques used during this research to generate results from the model simulation. The methods used for analyzing the performance of the surveyors used during this research were also explored and summarized. In addition, the IMPRINT Network diagrams were presented, along with the procedure used to adapt the results of the simulation to values of *d'* and the selection of VACP stressor values associated with both MOPP and heat models. Lastly, the baseline model was validated to accurately approximate (at the 90% confidence level) the observed performance of a surveyor to within approximately 0.30 standard units of *d'*.

IV. Analysis and Results

Chapter Overview

This chapter presents the results of the alternate IMPRINT models which incorporate PSFs to account for the wear of various levels of MOPP IPE by a surveyor, as well as the effects of extreme heat on the values of *d'* and surveyor efficiency. The alternate IMPRINT models are based on the performance of an average surveyor,

determined from the performance of six actual radiological surveyors. The effects of MOPP are presented as increases in the amount of time to perform the same radiological survey task in the various MOPP levels. The effects of heat on surveyor performance are put in terms of the scan MDC for comparison to changes in discriminability at elevated ambient temperatures.

Modeling Results – Effects of MOPP

To determine the effects on surveyor performance at increasing levels of MOPP and therefore reflect changes due to increased amounts of chemical IPE worn, alternate models were created in IMPRINT incorporating estimates of task times associated with the "Pause", "Move Probe", and "Stop" task nodes. Previously presented in Table 2, estimated task durations for each surveyor, as well as the theoretical "average" surveyor, were presented for each of these tasks. Each level of MOPP includes a prescribed combination of IPE to be worn, including a chemical protective overgarment, chemical resistant gloves and boots, and a protective mask with CBRN-certified filtration cartridges. These combinations for each MOPP level can be seen in Figure 15. Additional layers and equipment increase the physical strain on an individual wearer, making task execution more difficult. Increased heat and mental stress, loss of visual and tactile acuity, and reduced hearing are characteristic stressors on individuals wearing this equipment [42].

Figure 15. Mission Oriented Protective Postures (MOPP) [42]

The process of estimating the duration of pauses for each stage, as well as the amount of time the probe was *not* stationary was previously described in Section 3. To determine if the data approximated a known parametric distribution, SAS JMP Pro Version 15.2.1 (SAS Institute, Cary, NC) was used to apply continuous best fit curves to the data. Due to the small size of the dataset (n=6), results from the Goodness of Fit tests failed to produce a reliable estimate of the underlying distribution nor was a parametric transform found to reliably produce meaningful results. Calculated task durations for "Pause", "Stop", and "Move Probe" are assigned based on a weighting factor (0.25 for "Pause", 0.75 for "Stop") or assigned based on a consistent probe velocity shared amongst all surveyors. Since a relationship between variables is known, the triangular distribution is the most appropriate approximation to apply to the data in order to assess variability. IMPRINT accepts task times in a variety of distribution formats, including the

triangular, as well as many others depending on the amount of information known about the timing info.

The triangular distribution is characterized by three quantities: a known minimum, maximum, and estimated central value. These quantities, represented as *a*, *b,* and *c* below, can be visualized as a probability density graph which forms a triangle, with a width equal the minimum and maximum and peak approximating the measure of central tendency. Minimum and maximum values were observed directly from Table 2. The value of *c* was calculated using the formula for the mean (μ) of a triangular distribution, given according to equation 12 [43]:

$$\mu = \frac{a + b + c}{3} \tag{13}$$

The mean task times calculated from the six surveyors were substituted for μ, and solved for the quantity *c*. The results can be seen in Table 6.

Table 6. Triangular Distribution Values for MOPP Analysis

Task	a	b	c
Pause	2.07	6.15	3.24
Decision	6.2	18.46	9.75
Move Probe	12.57	27.61	17.87

These values were used as input variables for the corresponding task nodes in IMPRINT, as seen in Figure 16.

Task ID	Function	Task	Time							
			Time Requirement	Time Calculation	Task Duration	Time Expression	Distribution Type	Parameter 1 Value	Parameter 2 Value	Parameter 3 Value
0	(Root)	Model START	00:00:00.00	Value	00:00:00.00					
1	(Root)	Orient Probe	00:00:00.00	Value	00:00:00.00					
10	(Root)	In Background	00:00:00.00	Value	00:00:00.00					
15	(Root)	Stop On Src (Answ	00:00:00.00	Value	00:00:00.00					
16	(Root)	Stop in Backgroun	00:00:00.00	Value	00:00:00.00					
17	(Root)	Stop on a Source	00:00:00.00	Value	00:00:00.00					
18	(Root)	Stop In Backgroun	00:00:00.00	Value	00:00:00.00					
19	(Root)	Stop In Backgroun	00:00:00.00	Value	00:00:00.00					
2	(Root)	Pause	00:00:00.00	Distribution			Triangular	00:00:03.24	00:00:02.07	00:00:06.05
26	(Root)	Move Probe	00:00:00.00	Distribution			Triangular	00:00:17.87	00:00:12.56	00:00:27.61
3	(Root)	Pause on A Source	00:00:00.00	Value	00:00:00.00					
4	(Root)	On Src (Answer Ye	00:00:00.00	Value	00:00:00.00					
5	(Root)	On Src (Answer No	00:00:00.00	Value	00:00:00.00					
6	(Root)	In Background (An	00:00:00.00	Value	00:00:00.00					
7	(Root)	In Background (An	00:00:00.00	Value	00:00:00.00					
8	(Root)	Stop	00:00:00.00	Distribution			Triangular	00:00:09.75	00:00:06.20	00:00:18.46
9	(Root)	Stop On Src (Answ	00:00:00.00	Value	00:00:00.00					
999	(Root)	Model END	00:00:00.00	Value	00:00:00.00					

Figure 16. Task Duration Inputs Values for MOPP Analysis in IMPRINT

Separate models for each level of MOPP were then created, with each simulation set to run 1000 times. A new random number seed was used for each model run. Figure 17 shows the results of the simulated output, with the minimum, maximum, and mean values for each level of MOPP. Horizontal lines are also shown corresponding to $\pm$ 1 standard deviation of the mean. As the model reflects a surveyor's decision to pause and investigate an area of interest, the minimum, maximum, and average task durations from the simulation output are multiplied by 109, or the number of times the average surveyor paused. As can be seen in the figure, the resulting mean task times for MOPP 0, where the effect of MOPP is removed, is nearly identical (within 1 minute) of the mean of the observed task times for the six survey participants.

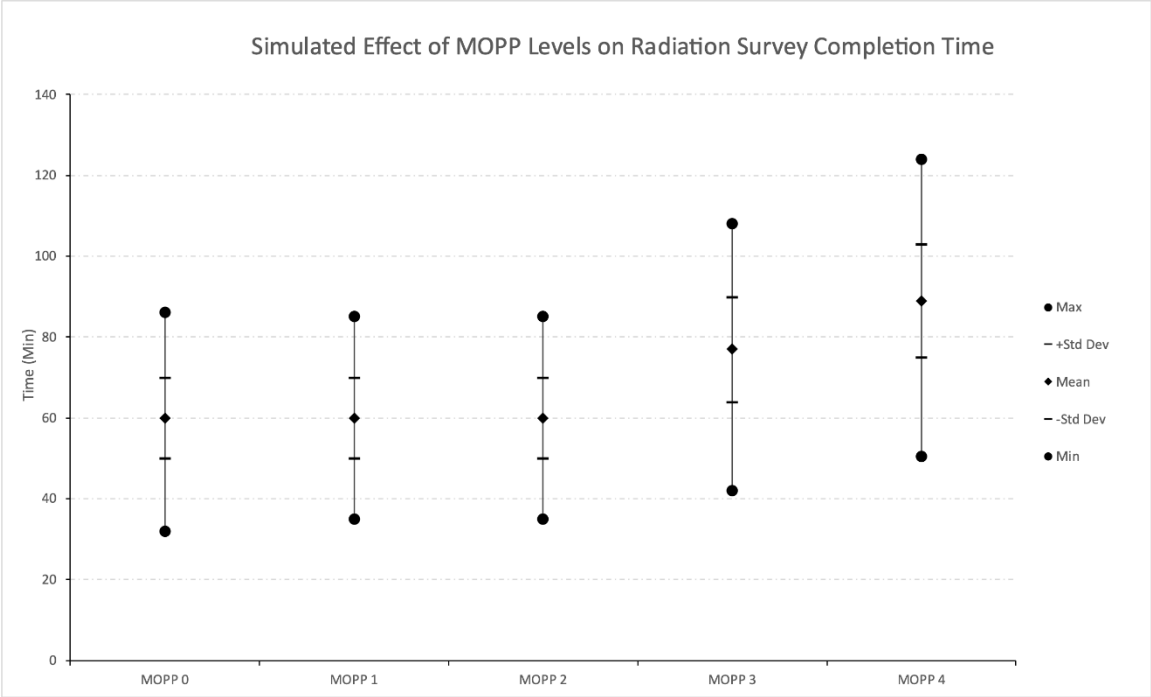


Figure 17. Effect of Different MOPP Levels on Survey Completion Times

Results of the model simulations yield a survey duration of 60 minutes for MOPP levels 0 through 2, between which little difference is observed. IMPRINT applies little weighting (referred to as degradation factors in the relevant literature) against these MOPP levels, which is reflected in the simulation results at these levels. At MOPP levels 3 and 4, however, more significant task degradation is seen, and is applied as a weighted average affected most significantly by the "oral communication" taxon [26]. Taxons for each task are mapped according to the 7-channel VACP assignments shown in Figure 11. Values are weighted automatically using the default IMPRINT assignment methodology. For both the "Pause" and "Stop" nodes, taxon weighting of 34% (Visual Recognition), 26% (Information Processing/Problem Solving), and 40% (Oral Communication) were assigned. These weights directly impact the total task times produced by the model.

For the simulated average surveyor, a mean difference of 29 minutes is indicated when MOPP 4 is worn compared to MOPP 0, with a maximum estimated difference of 38 minutes. However, it should be noted that the variability associated with each MOPP level is significant and is largely due to a lack of sufficient data to use as model input and a lack of additional information about the possible underlying distribution of the timing information. Few of the total task nodes have any timing information and this information is derived from estimates of timing not directly observed or recorded during data collection. Overall, the output values are consistent with published degradation factors, which range from 1.5 to 1.7 for MOPP 4 depending on the type of task [36]. For this model, the observed degradation factor between MOPP 0 and MOPP 4 is 1.5, consistent with earlier research by the U.S. Army. Additional observational data is needed to effectively reduce the large amount of variability seen in the model output and further validate the ability of the model to simulate human performance while wearing MOPP.

Modeling Results – Effects of Heat

To determine the effects of heat stress on the average surveyor, VACP values were assigned to the task nodes as outlined in Figure 12. Probabilistic path logic was updated to reflect that of the average surveyor, using the arithmetic mean of the number of hits, misses, false alarms, and correct rejections associated with all six surveyors. The values for Hit Rate and False Alarm rate were calculated based on the methodology outlined in the previous chapter. The discriminability and bias measures associated with the average surveyor can be seen in Appendix E.

The effects of heat on surveyor performance are evaluated based on the degradation of task accuracy, produced by the output reports of task performance from IMPRINT. For each temperature and humidity evaluated, the simulation was run a total of 109 times, corresponding to the number of times the average surveyor paused. A more significant amount of post-processing is required, however, in order to put the model output in terms of the desired metric (*d'* and ρ). IMPRINT output indicating the number of failures associated with a task is manually subtracted from the number of times that task would have run, since task failures due to the effects of heat are not automatically subtracted by IMPRINT. This convention is acceptable, since the effects of not producing a "hit" due to heat can be subtracted from the number of times the corresponding task failed to run. Similarly, the result of a false alarm not occurring is subtracted from the number of times the corresponding task node was executed by the simulation. The result can be processed similarly to the baseline model and put in terms of *d'* and ρ to evaluate changes in discriminability by a surveyor and the overall efficiency relative to an ideal surveyor presented with similar count rate information.

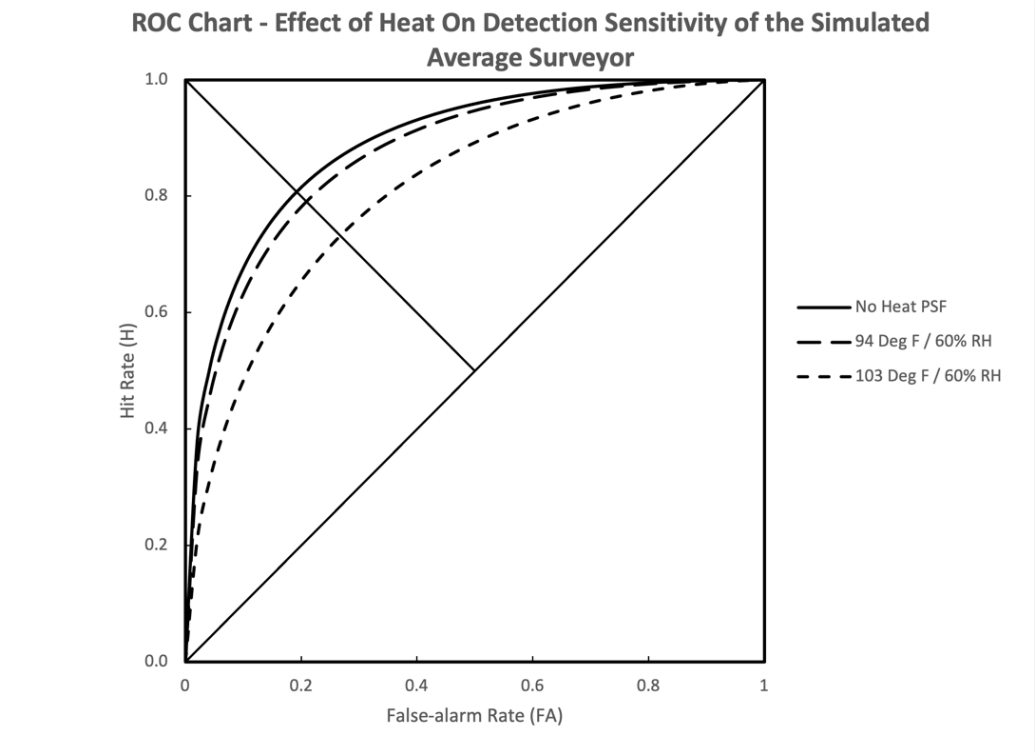


Figure 18. Simulated Effect of Heat Stress on d'

Figure 18 demonstrates the effects of heat on the simulated average surveyor on the traditional ROC chart. The solid line, corresponding to a *d'* value of 1.74, describes the performance of a surveyor where the effect of heat is not considered, and is similar to the *d'* calculated for the average surveyor in Appendix E. Recall that while the ability of a surveyor to accurately discriminate a source from background may be described by this curve, the choice of the criterion (*c)* value may in fact be anywhere along this line, with the minor diagonal line indicating where $c = 0$. Simulated discriminability at two different ambient environmental temperatures are shown, indicated by two separate dashed lines of differing weight. At 94 degrees Fahrenheit (F) and 60% Relative Humidity (RH), a *d'* of 1.62 is produced. Additionally, a simulated *d'* of 1.24 is produced by the model when evaluating the effects of an ambient air temperature of 103 degrees F

and 60% RH. These values of *d'* correspond to overall surveyor efficiency values of 0.34, 0.29, and 0.17, respectively.

Impact to the Surveyor MDCR

The quantities of *d'* , ρ, and *i* in the surveyor's MDCR equation are referred to as human factor inputs [13]. Thus, other than the number of background counts within the evaluated observational interval, all the terms of the MDCR equation depend on values which vary, at times significantly, due to human variability and error. The choice of *d'* in the numerator of the Surveyor MDCR equation is made during the survey planning process by considering what percentage of true positives are desired and the percentage of false positives that can be tolerated (or willing to be accepted to meet survey objectives). As it applies to the MDCR, the value of *d'* in the numerator is thought of as a performance criterion to be met, rather than an observed or measured quantity. A table of values of *d'* for most common true and false positive proportions of interest used for selecting this performance criterion can be seen in Appendix I. ρ, the ratio of the surveyor's efficiency relative to the ideal observer, has historically been assigned a value of either 0.5 or 0.75, depending on the information available to survey planners, and is independent of the value of *d'* selected as performance criteria [1]. However, recall that the value of ρ, according to equation 9, is defined by the measure of sensitivity of an actual surveyor relative to the ideal surveyor. While the results of several experimental trials support these values, they also fail to consider factors of which a surveyor has little control, namely the effects of equipment used to protect the surveyor from radioactive contamination (e.g., chemical protective IPE with MOPP) and the ambient environmental

temperatures at which a survey is performed. While several surveyors observed during the detection task had ρ values of 0.70, in general most of the surveyors had ρ values of ~0.30.

The observation interval, i, is a function of scan speed; slower (or static) movement of a survey probe results in longer observation intervals which serve to provide more count rate information to the observer.

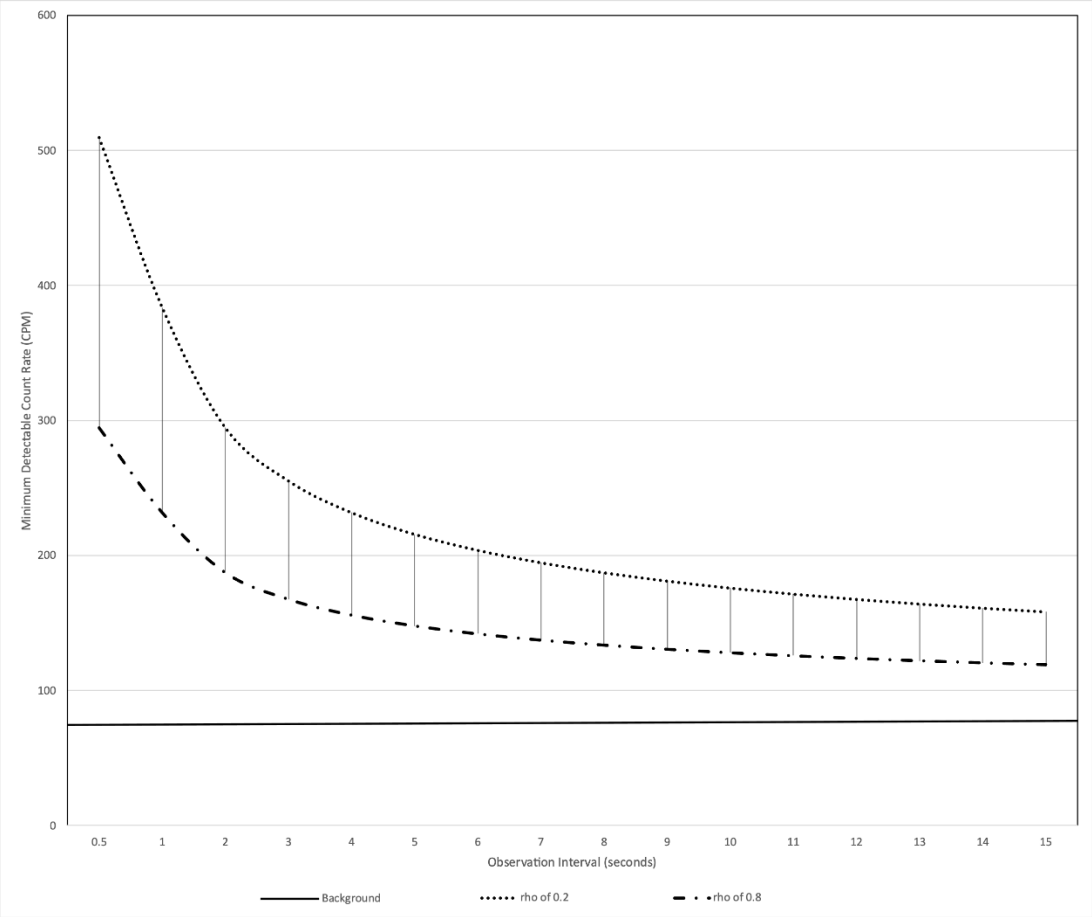


Figure 19. Surveyor MDCR as a Function of Observation Interval

The surveyor MDCR equation, reproduced below, can be evaluated over increased observation intervals while the quantities of d' and $\sqrt{b_i}$ are held static. Under

these conditions, the MDCR can be graphed to produce curves shown in Figure 19. The two dashed lines correspond to lower and upper boundary values of ρ at 0.2 and 0.8, respectively, in a background of 80 cpm and with a selected *d'* of 1.96 (corresponding to a desired true positive proportion of 0.9 and false positive proportion of 0.25)[1]. In practice, values of ρ used in the MDCR equation will fall between these two limits, however the overall shape of the resulting curve will remain the same. At increasing background count rates, the same basic relationship is observed, however the net count rates (in CPM) are increased accordingly. As seen in Figure 19, increased values of ρ lead to an increase in the minimum count rate that can reliably be detected by a surveyor (based on the value chosen for *d'* in the numerator).

Results from Brown & Abelquist (1998) provide clear evidence on the importance of evaluating the components of the MDCR to determine if a chosen survey strategy will provide the sensitivity required to meet the goals of a survey. Recommendations for the use of ρ values no greater than 0.5 under field conditions is based on a limited amount of direct observation, but is a reasonable assumption, nonetheless. As noted by Brown & Abelquist (1998), the performance of the six surveyors used during this research was not reliably correlated to their experience with performing radiological surveys.

Recommendations regarding using ρ values between 0.5 and 0.75 are based on the results of a series of simulations and experimental trials in addition to the Geiger-Mueller survey task evaluated in this report. Efficiencies for 4 out of the 6 survey participants fall below this range, despite all surveyors adopting a similar decision strategy applying the appropriate criterion over the first and second scanning stages.

When applied to the simulated average surveyor, values of ρ, describing the efficiency relative to the ideal surveyor, are reduced between 14.7 and 50 percent at temperatures ranging from 94 deg F to 111 deg F and relative humidity values between 61 to 70%. When values of the simulated output are converted into quantities of ρ (using equation 9) and are applied to the surveyor MDCR equation, the following estimates of the MDCR are created.

$$MDCR_{Surveyor} = \frac{d' * \sqrt{b_i} * \frac{60}{i}}{\sqrt{\rho}}$$

Where:

$d' = index\ of\ sensitivity$

$b_i = number\ of\ background\ counts\ in\ the\ observation\ interval$

$i = observation\ interval$

$\sqrt{\rho} = square\ root\ of\ surveyor\ efficiency$

Assuming a *d'* of 1.96 (a desired 90% true positive rate and accepted 25% false positive rate), background count rate of 80 cpm, short 1 second observation interval corresponding to the initial pause stage, and in an environment with air temperatures between 94 and 102 degrees F; the net MDCR for a given surveyor becomes:

$$MDCR_{Surveyor} = \frac{1.96 * \sqrt{1.33} * \frac{60}{1}}{\sqrt{0.34}} = 232\ cpm\ (no\ heat)$$

$$MDCR_{Surveyor} = \frac{1.96 * \sqrt{1.33} * \frac{60}{1}}{\sqrt{0.29}} = 252\ cpm\ (with\ heat)$$

And second (decision stage assuming a 4 sec observation interval):

$$MDCR_{Surveyor} = \frac{1.96 * \sqrt{5.33} * \frac{60}{4}}{\sqrt{0.34}} = 116\ cpm\ (no\ heat)$$

$$MDCR_{Surveyor} = \frac{1.96 * \sqrt{5.33} * \frac{60}{4}}{\sqrt{0.29}} = 126\ cpm\ (with\ heat)$$

The effect on the MDCR is more evident when the model results using the more extreme heat PSF (above 103 degrees F) are applied to the MDCR:

$$MDCR_{Surveyor} = \frac{1.96 * \sqrt{1.33} * \frac{60}{1}}{\sqrt{0.34}} = 232\ cpm\ (no\ heat)$$

$$MDCR_{Surveyor} = \frac{1.96 * \sqrt{1.33} * \frac{60}{1}}{\sqrt{0.17}} = 329\ cpm\ (with\ heat)$$

And second (decision stage assuming a 4 sec observation interval):

$$MDCR = \frac{1.96 * \sqrt{5.33} * \frac{60}{4}}{\sqrt{0.34}} = 116\ cpm\ (no\ heat)$$

$$MDCR = \frac{1.96 * \sqrt{5.33} * \frac{60}{4}}{\sqrt{0.17}} = 165\ cpm\ (with\ heat)$$

The example above is in terms of the net (source) counts, therefore the *a priori* MDCR which would warrant additional investigation using the more extreme heat (above

103 deg F) value would be equal to 329 + 80 cpm, or 409 cpm gross. This count rate is several times (5x) more than the background count rate and indicates a potential source of risk. The above would represent that under simulated conditions, a surveyor would investigate a count rate of 312 cpm (gross) where temperature effects are not applied but would fail to investigate a count rate below 409 cpm in an elevated temperature environment. Surveyor planners would need to either adopt a slower scan speed (thereby increasing the observation interval) or reconsider the chosen value of *d'* and be willing to accept either a reduced rate of true positives, or an increased rate of false positives reported. Both are viable options depending on the goals and constraints of the survey.

As previously analyzed, increasing levels of MOPP only affects the time required to perform a task, rather than decreasing the accuracy or measures of discriminability of a surveyor. However, research into how a surveyor uses the available time and this effect on discriminability is limited. As has been described by Brown (2000), how a surveyor utilizes the time available may directly impact overall sensitivity, however this relationship has not been researched [34].

Since human variability is one of the key aspects to increase the surveyor's MDCR, it is conceivable that one method in reducing the MDCR is by minimizing the detrimental contributions that the quantities of *d'*, ρ, and *i* have when evaluating surfaces for radiological contamination. Removing the human element from radiological surveys is one method of accomplishing this, and drones and robotic technologies have been used in recent history with great success, with several commercially available currently. However, in many instances radiological surveys involving field personnel and portable detection equipment are the only option, depending on the availability of resources. The

detection of particulate contamination on surfaces using small unmanned aerial systems is also difficult and currently unsupported using many such systems [44], [45].

There are limited commercial solutions for surveyors, however, wishing to improve one or more of the measures of performance discussed in this research. As previously proven, the observation interval (*i*) is a key quantity, alerting a surveyor to potential areas of contamination requiring further investigation independent of a surveyor's ability to accurately discriminate source counts. At longer observation intervals, the gross MDCR approaches a value approximately one and a half times the background count rate. Gross surveyor MDCRs more than twice the background count rate are generally deemed unacceptable for stringent decommissioning surveys, since at this count rate, a surveyor may be unable to discriminate net (source) counts from background counts. However, in other situations, survey planners and decision makers may choose to accept MDCRs several times the background count rates, at the risk of potentially failing to identify areas of potential interest (e.g., a discrete radiological source or elevated measurement area).

Assuming that the efficiency of a surveyor is not a quantity that can be modified and substitution or replacement with a surveyor with a known higher sensitivity is not feasible, the only quantity that can be reasonably modified is often the observation interval. As was seen in Figure 19, increasing the observation interval serves to lower the Gross surveyor MDCR, especially when observation intervals are initially very low (less than 1 second). Given a known background count rate, a recommended observation interval can be prescribed which results in a sufficiently low MDCR to meet survey objectives.

Technologies which serve to monitor the velocity and movement of a radiological survey probe and provide feedback to the surveyor when velocity across a surface exceeds a predefined value may be of particular interest to survey planners. To ensure a surveyor moves a survey probe with a velocity sufficiently slow enough to realize a desired observation interval, visual feedback could be presented alerting the surveyor when a preset velocity is exceeded.

For a Geiger-Mueller type survey probe with a probe housing width of 2.75 inches, a common recommendation for field surveyors is to keep probe velocities at or below 1 probe width per second, which corresponds to an observation interval of 1 second. As can be observed from Figure 19, this is likely the maximum recommended observation interval and probe velocity. Slowing to a velocity of 0.5 probe widths per second doubles the observation interval, decreasing the MDCR by ~20% (based on Figure 19 and assuming all other quantities remain unchanged).

Investigative Questions Answered

All three investigative questions conducted as part of this research were answered. The use of IMPRINT as a DES was successfully employed to simulate a radiological detection task utilizing portable equipment. A task network representing the physical and cognitive elements employed during a radiological detection task was created and validated to accurately simulate observed human performance. It was found that the use of personal protective equipment (in the form of chemical IPE) increases the amount of time it takes to complete a simulated survey task involving the use of portable equipment by an average surveyor. The impact that the external environment has on the performance

of a surveyor was evaluated by simulating the effect of heat stress. When applied to the simulated average surveyor, surveyor efficiency (ρ) is reduced between 14.7 and 50% at temperatures ranging from 94 degrees Fahrenheit to 111 degrees Fahrenheit and relative humidity values between 61 to 70%.

V. Conclusions and Recommendations

Conclusions of Research

The major effort this research undertook was the development of a simulation model which could approximate the human performance of a radiological detection task. The model developed used probabilistic pathway logic and the application of performance shaping factors which simulate the effect that chemical IPE and extreme heat has on survey completion.

Significance of Research

IMPRINT has seldom been used to model short duration or cognitive tasks, and instead employed to model larger, complex tasks or functions on the scale of minutes, hours, or days [46]. Additionally, this research represents the sole use of IMPRINT to model a process fully described within the signal detection framework for the perception of audio tones. The success of the model to accurately represent the performance on a surveyor is augmented by IMPRINT's built-in performance shaping factors, ability to map to the VACP framework outlined according to Multiple Resource Theory and produce reports and outputs that could easily be put in terms of overall task duration or common measures of sensitivity found in SDT. Finally, the model output highlights the

impact of heat stress, which is related back to surveyor efficiency (ρ) and applied to an evaluation in terms of the surveyor MDCR.

Recommendations for Future Research

Continued development and refinement of models used in DES is common and expected. While the radiation surveyor model developed for this research approximates some measures of performance by a surveyor, it is not complex enough to predict all factors which influence a surveyor. The performance shaping factors included in IMPRINT for heat and personal protective equipment are dated, stemming from research conducted over 40 years ago. While some findings may still be valid, additional tools and technologies are now available which may change or modify the degradation factors applied. Unfortunately, as of 29 September 2019, IMPRINT is no longer funded or supported by the U.S. Army Research Laboratory (ARL) and U.S. Army Combat Capabilities Development Command (CCDC). Access to support resources, including the SharePoint page and technical assistance, is no longer available. Continued use is contingent based on software compatibility of the final version released.

An investigation of the results from the alternate models compared to real-world surveyors in personal protective equipment and under elevated temperature conditions would provide valuable insight into the additional factors impacting a surveyor's performance to perform radiological detection tasks.

Summary

The creation of models which approximate human performance using discrete event simulation software offers a powerful tool to predict the impact on performance from one or more environmental stressors. The number and types of stressors available for application is limited only by the availability of accurate observational performance data with which to associate to a set of tasks. Radiological surveillance activities are well suited to being modeled by DES, as the physical and cognitive processes have been extensively researched and can be observed with minimal difficulty in the physical environment. Depending on the task being modeled, DES can allow for the evaluation of potential modifications to improve efficiency, safety, or resource utilization. The widespread availability of computational tools and personal computers make evaluations or modifications to the human system in a digital environment easier to evaluate or modify than they would be in the physical environment.

Although some smaller studies have evaluated human performance during radiological detection tasks, the work by Brown & Abelquist (1998) is the largest and most comprehensive to date, employing both field and controlled experimental trials to understand detection performance using most of the common types of radiological detection equipment. However, little observational information has been collected which fully records and describes the amount of time associated with pause and decision stages. In some cases, experimental trials in NUREG/CR-6364 utilized a very small number of survey participants, which makes meaningful interpretation difficult. An increased number of survey participants per experimental trial, with a more robust investigation of individual parameters (such as experience), would also be advised. Furthermore,

advances in radiological survey equipment as well as technology used to support recordkeeping and documentation have not been evaluated and should be incorporated into an updated report on human performance during radiological survey activities.

This research demonstrates that surveyors are negatively affected by the impact of chemical Individual Protective Equipment (IPE) wear and their external environment under elevated ambient temperature conditions. An increase in the duration of time to complete a radiological monitoring task and decrease in sensitivity to indications of elevated radiological count rates due to elevated temperature conditions are predicted for the average surveyor performing radiological surveillance employing portable Geiger-Mueller based equipment. Recommendations in *Human Performance in Radiological Survey Scanning, NUREG/CR-6364* suggest values of ρ no greater than 0.5 under field conditions, which should be interpreted as an uppermost boundary when evaluating the MDCR for field surveys performed where elevated environmental temperatures are known or suspected to exist. Instead, survey planners should consider using ρ values of 0.25 to 0.425 (50 and 85 percent of 0.5, respectively) based on the reduction of ρ indicated by the results of the model. This recommendation should be investigated when comparing the model results to observations of surveyors in the physical environment.

Appendix A. Assumptions and Justifications

Table 7. List of Assumptions

Assumption	Type	Justification
Data recorded accurately describes the observations of surveyor performance	Data	A complete set of observations was obtained from the previous study's original author and data collector. Recorded information was unaltered, clearly legible, and authentic.
VACP values assigned for tasks accurately describes workload by channel	Task ID 1, 2, 8, and 26	SME input using selection matrix in IMPRINT
Surveyors moved instrument probe at a consistent rate (1 probe width per second) across the surface	Data	A reasonable average based on instructions provided to each surveyor, in the absence of discrete probe velocity information
Surveyors conducted the survey in a manner which did not overlap previously scanned areas	Data	Surveyors are not likely to repeat scans over previously surveyed areas, and were aware of which areas had been previously surveyed
Of the total time spent when a surveyor is stationary, 25% of the total is spent during the 1st stage, and 75% of the total is spent during the second	Task ID 2, 8, and 26	Of the total time spent stationary, a surveyor is expected to stop during the second stage 4 times longer than the pause associated with the first stage. This expectation is based on SME input and experience, in the absence of more specific information
A surveyor performs no other tasks during the completion of a survey	General	A surveyor is only concerned with completing the radiological survey task assigned in the shortest time possible
Temperature and humidty is asusmed to be invariable	Model	For the Heat model, the temperature and humdity experienced by the average surveyor does not change or fluctuate from the values used

Appendix B. VACP Values and Descriptors

Table 8 presents the standardized VACP values used in IMPRINT [25]

Table 8. 7-Channel VACP Scales

	Visual Benchmarks
Value	**Benchmark**
0.0	Nothing
1.0	Register/Detect (Detect Occurrence of Image)
3.0	Inspect/Check (Discrete Inspection/Static Condition)
4.0	Locate/Align (Selective Orientations)
4.4	Track/Follow (Maintain Orientation)
5.0	Discriminate (Detect Visual Differences)
5.1	Read (Symbol)
6.0	Scan/Search Monitor (Continuous/Serial Inspection)
	Auditory Benchmarks
Value	**Benchmark**
0.0	Nothing
1.0	Detect/Register Sound (Detect Occurrence of Sound)
2.0	Orient to Sound (General Orientation/Attention)
3.0	Interpret Semantic Content (Speech) Simple (1-2 Words)
4.2	Orient to Sound (Selective Orientation/Attention)
4.3	Verify Auditory Feedback (Detect Occurrence of Anticipated Sound)
6.0	Interpret Semantic Content (Speech) Complex (Sentence)
6.6	Discriminate Sound Characteristics (Detect Auditory Difference)
7.0	Interpret Sound Patterns (Pulse rates etc.)
	Cognitive Benchmarks
Value	**Benchmark**
0.0	Nothing
1.0	Automatic (Simple Association)
1.2	Alternative Selection
4.6	Evaluation/Judgement (Consider Single Aspect)
5.0	Rehearsal
5.3	Encoding/Decoding, Recall
6.8	Evaluation/Judgement (Consider Several Aspect)
7.0	Estimation, Calculation, Conversion

	Fine Motor Benchmarks
Value	**Benchmark**
0.0	Nothing
2.2	Discrete Actuation (Button, Toggle, Trigger)
2.6	Continuous Adjustive (Flight Control, Sensor Control)
4.6	Manual (Tracking)
5.5	Discrete Adjustment (Rotary, Vertical Thumb Wheel, Lever Position)
6.5	Symbolic Production (Writing)
7.0	Serial Discrete Manipulation (Keyboard)
	Gross Motor Benchmarks
Value	**Benchmark**
0.0	Nothing
1.0	Walking on Level Terrain
2.0	Walking on uneven terrain
3.0	Jogging on Level Terrain
3.5	Heavy Lifting
5.0	Jogging on Uneven Terrain
6.0	Complex Climbing
	Speech Benchmarks
Value	**Benchmark**
0.0	Nothing
2.0	Simple (1-2 words)
4.0	Complex (Sentence)
	Tactile Benchmarks
Value	**Benchmark**
0.0	No tactile activity
1.0	Alerting
2.0	Simple discrimination
4.0	Complex symbolic information

Appendix C. Observations of Surveyor Performance

Data from Brown & Albequist (1998), *Human Performance in Radiological Survey Scanning*, NUREG/CR-6364 [1]

GM

[redacted]						
# Pause	71				0.36	0.67
# on Src	8	6	# of Src		0.007143	0.666667
# not Src	63					
# FA	51					
# Opp	140	9	# Sources			
FAR #1	0.36	0.67	HR #1	0.78		
# False +	1	6	# True +			
FAR #2	0.02	0.92	HR # 2	3.44		
	0.007143	0.666667		2.88		

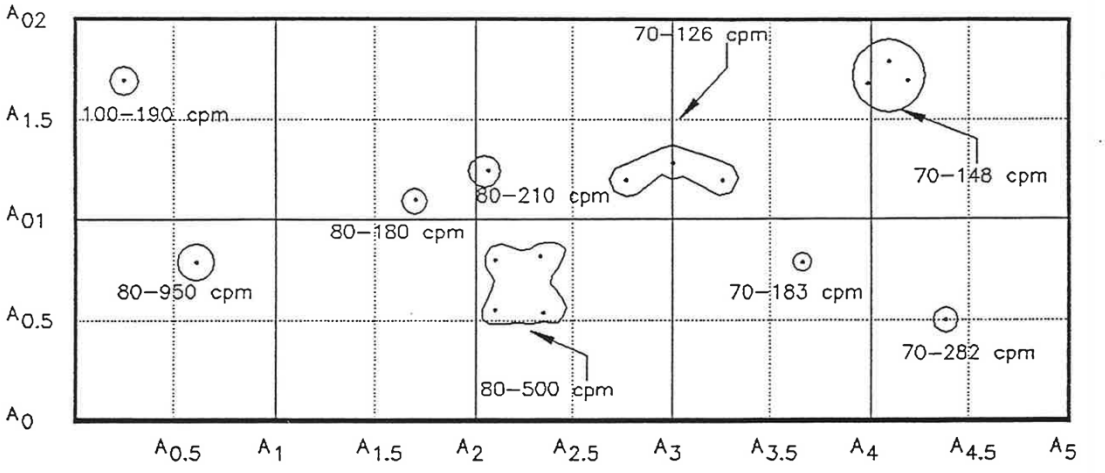
[redacted]						
# Pause	140				0.58	0.89
# on Src	32	8	# of Src		0.004	0.555556
# not Src	108					
# FA	72					
# Opp	125	9	# Sources			
FAR #1	0.58	0.89	HR #1	1.03	0.16	0.85
# False +	0	5	# True +			
FAR #2	0.01	0.63	HR # 2	2.78		
	0.004	0.555556		2.79		

[redacted]						
# Pause	104				0.44	0.89
# on Src	23	8	# of Src		0.007143	0.333333
# not Src	81					
# FA	62					
# Opp	140	9	# Sources			
FAR #1	0.44	0.89	HR #1	1.36		
# False +	1	3	# True +			
FAR #2	0.02	0.38	HR # 2	1.82		
	0.007143	0.333333		2.02		

[redacted]						
# Pause	156				0.60	0.89
# on Src	28	8	# of Src		0.071429	0.555556
# not Src	128					
# FA	84					
# Opp	140	9	# Sources			
FAR #1	0.60	0.89	HR #1	0.97		
# False +	10	5	# True +			
FAR #2	0.12	0.63	HR # 2	1.50		
	0.071429	0.555556		1.60		

[redacted]						
# Pause	91				0.39	0.89
# on Src	23	8	# of Src		0.035714	0.666667
# not Src	68					
# FA	54					
# Opp	140	9	# Sources			
FAR #1	0.39	0.89	HR #1	1.51		
# False +	5	6	# True +			
FAR #2	0.09	0.75	HR # 2	2.00		
	0.035714	0.666667		2.23		

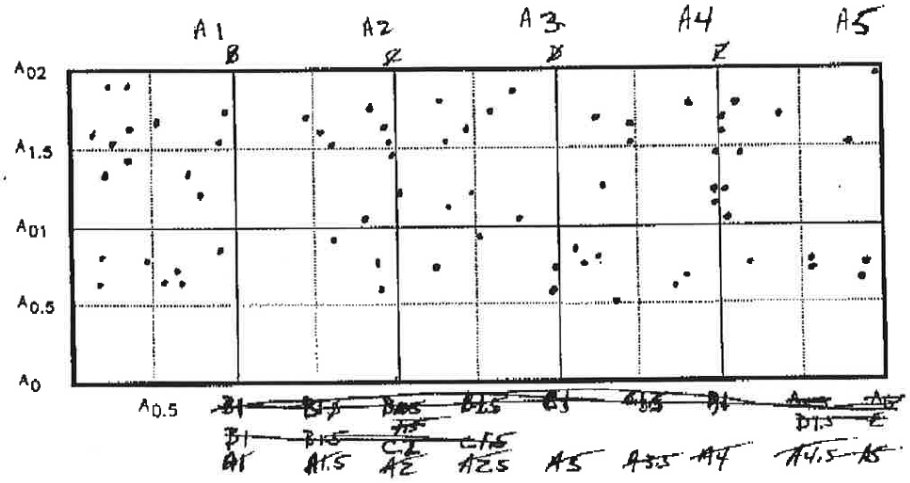
# Pause	91			0.36	0.89
# on Src	28	8 # of Src		0.021429	0.555556
# not Src	63				
# FA	51				
# Opp	140	9 # Sources			
FAR #1	0.36	0.89 HR #1	1.57		
# False +	3	5 # True +			
FAR #2	0.06	0.63 HR # 2	1.88		
	0.021429	0.555556	2.16		



NAME: [redacted] DATE: 3/9/95

BKG: _____ (cpm)

≈ 50 min to finish [redacted]



● - STOP TO EVALUATE POTENTIAL HOT SPOT

(#) ACTUAL cpm OF SPOT IDENTIFIED

FEET 0 3

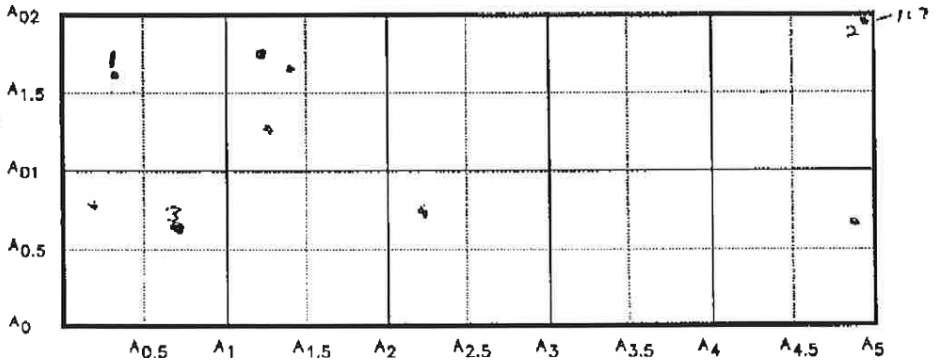
METERS 0 1

71 PAUSES

NRCB (x)

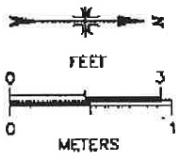
NAME: [redacted] DATE: 3/9/95

BKG: 25-45 (cpm)



● - STOP TO EVALUATE POTENTIAL HOT SPOT

- ACTUAL cpm IF SPOT IDENTIFIED



(1) 153
(2) 103
(3) 500

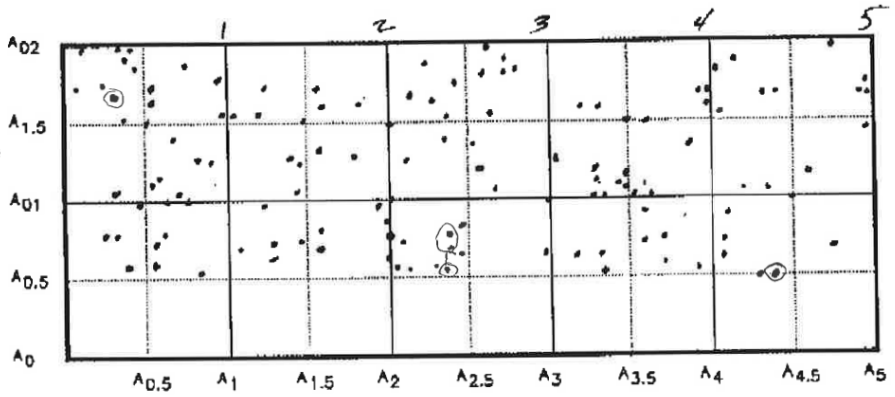
2 CORRECT HITS

NAME: [redacted] DATE: 3/9/95

BKG: _____ (cpm)

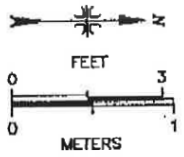
[redacted]

$\bar{x}$ 54.07 min



● - STOP TO EVALUATE POTENTIAL HOT SPOT

- ACTUAL cpm IF SPOT IDENTIFIED

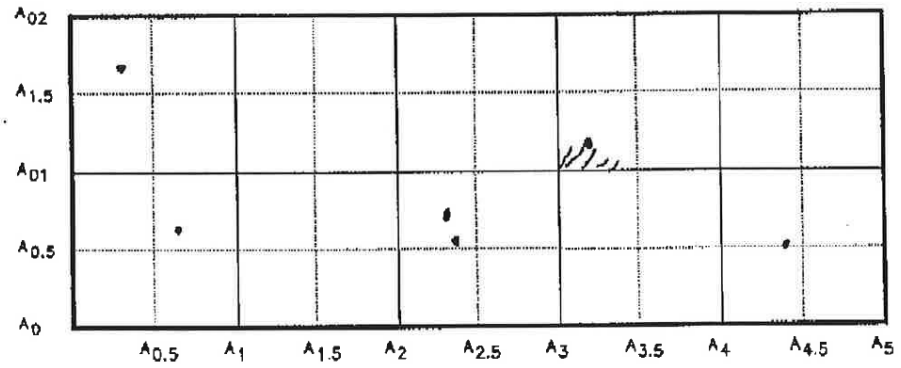


Scan speed was slower than [redacted] it was more through.

NRDS (6c)

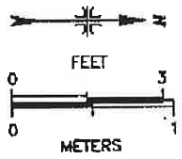
NAME: [redacted] DATE: 3/9/95

BKG: _____ (cpm)



● - STOP TO EVALUATE POTENTIAL HOT SPOT

- ACTUAL cpm IF SPOT IDENTIFIED



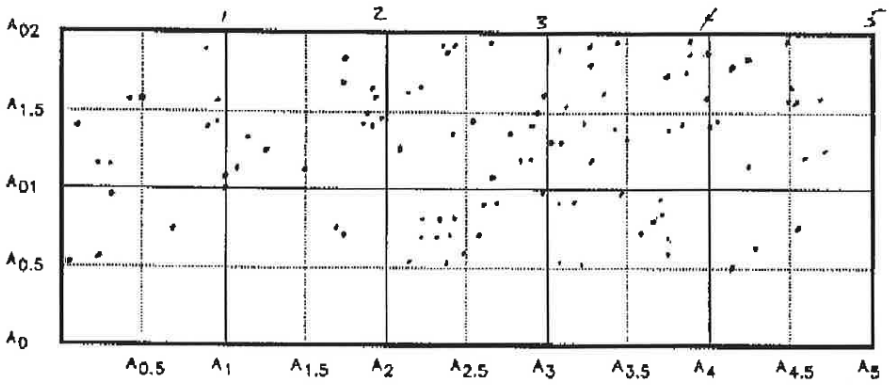
5 CORRECT HITS

NRDB (+)

NAME: [redacted] DATE: 3/8/95

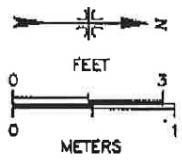
BKG: _____ (cpm)

≈ 47 min.



● - STOP TO EVALUATE POTENTIAL HOT SPOT

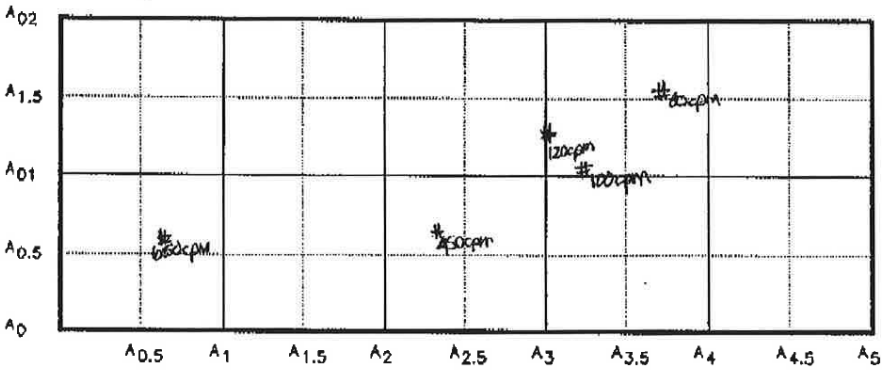
- ACTUAL cpm IF SPOT IDENTIFIED



NRCB (h)

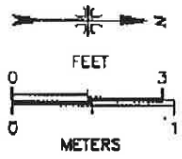
NAME: [redacted] DATE: 3/9/95

BKG: _____ (cpm)



● — STOP TO EVALUATE POTENTIAL HOT SPOT

- ACTUAL cpm IF SPOT IDENTIFIED

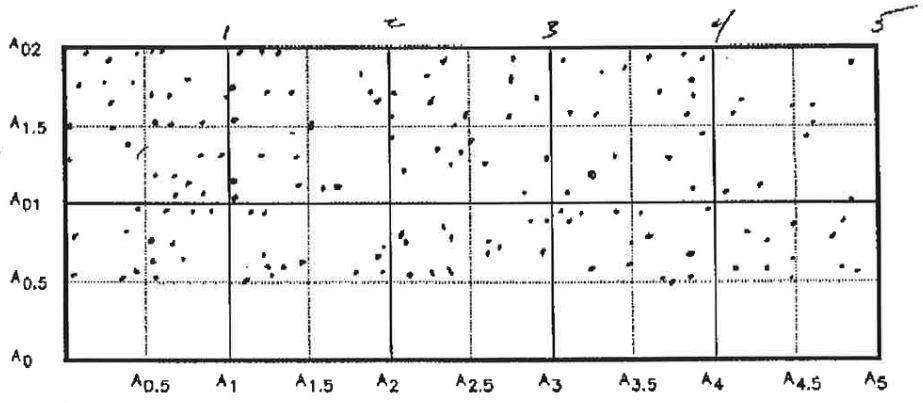


NRCM (4)

NAME: [redacted] DATE: 3/9/95

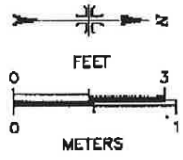
BKG: ____ (cpm) NO EXPOSURE

22 min
+50
72 min.



● - STOP TO EVALUATE POTENTIAL HIT SPOT

- ACTUAL cpm IF SPOT IDENTIFIED

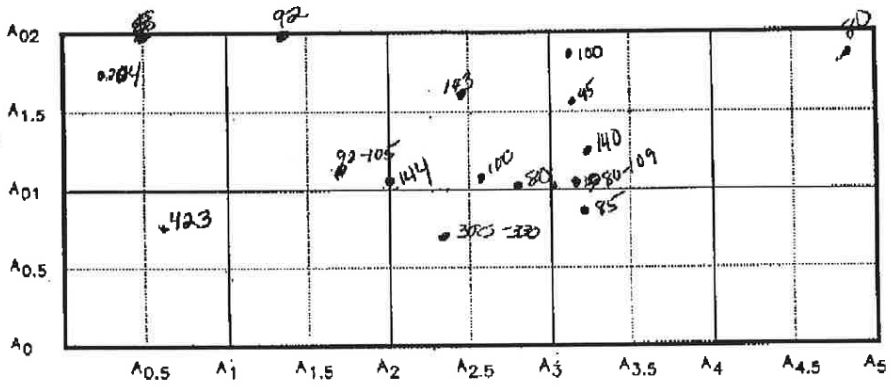


Users have a slower scan speed.
May not understand the fast response of scale mode.

NAME: [redacted] DATE: 3/9/95

BKG: _____ (cpm)

NO EXPENSE



● - STOP TO EVALUATE POTENTIAL HIT SPOT

- ACTUAL CPM IF SPOT IDENTIFIED

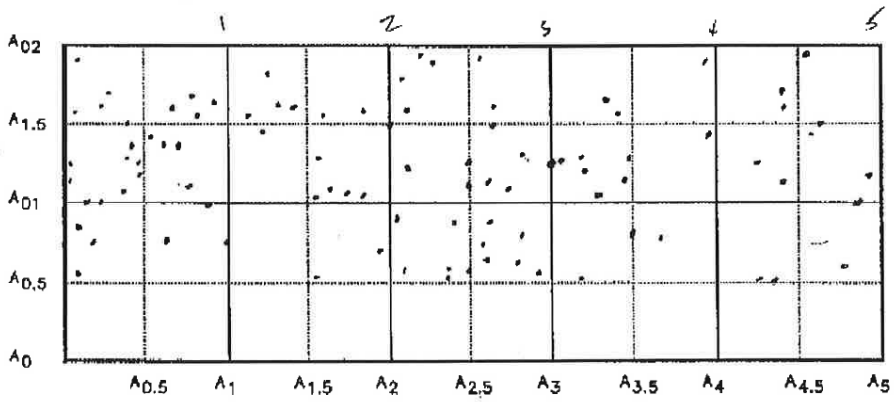
FEET
0 3
0 1
METERS

NRQR (x)

Name: [redacted] Date 3/10/95

Bkg: _____ cpm

≈ 40 min the 1st 1/2 of Board.
≈ 30 min the last 1/2 of Board
70 min.



- Actual cpm if Spot is identified.

. # of stops made

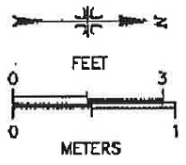
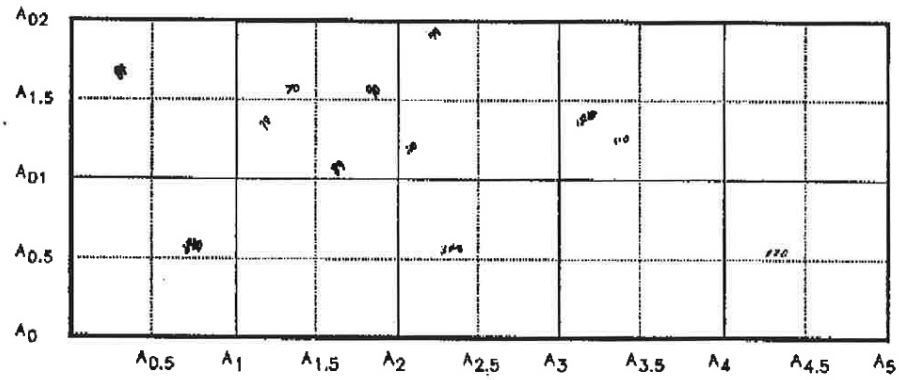


FIGURE 6-1: Indoor Wall Map Used For Recording Results of Scan Survey

NRC8(1)

Name: [redacted] Date 3/10/95

Bkg: _____ cpm



- Actual cpm if Spot is identified.

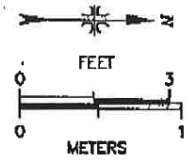


FIGURE 6-1: Indoor Wall Map Used For Recording Results of Scan Survey

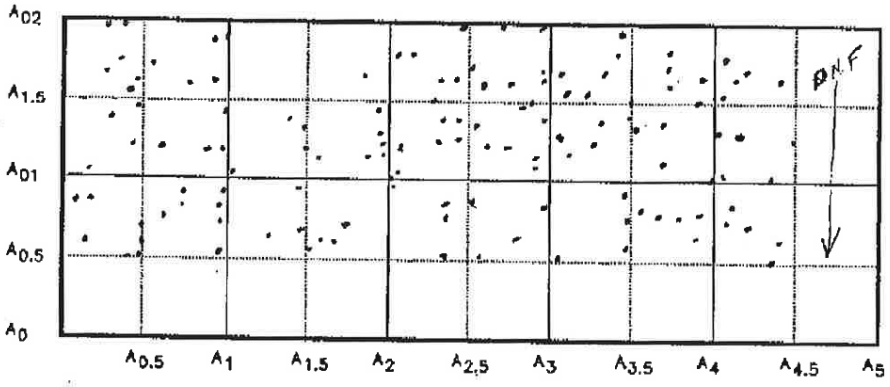
NRC8 (1)

Name: [redacted] Date 3/10/95

Bkg: _____ cpm

± 1.05 hrs.

proj 70



- Actual cpm if Spot is identified.

Good 100% coverage.

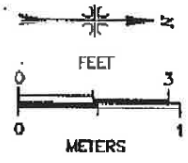
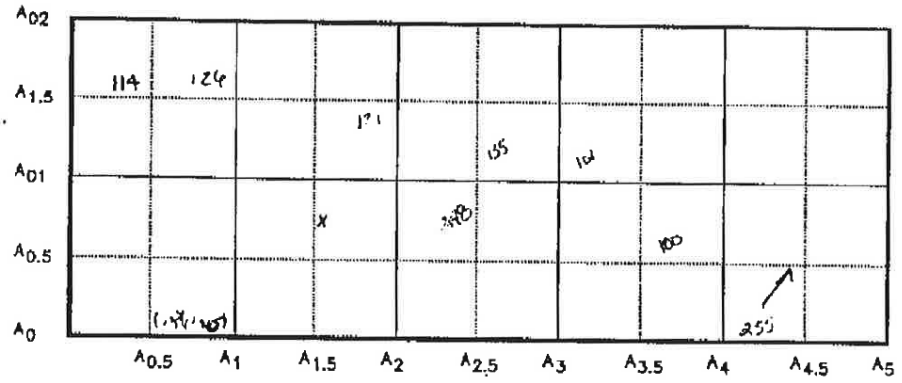


FIGURE 6-1: Indoor Wall Map Used For Recording Results of Scan Survey

NRCS(1)

Name: [redacted] Date 3/10/95

Bkg: _____ cpm



- Actual cpm if Spot is identified.

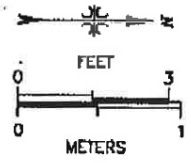


FIGURE 6-1: Indoor Wall Map Used For Recording Results of Scan Survey

NRC6(1)

Appendix D. Measures of Surveyor Performance

Probability density function graphs and ROC charts which display both the pause and detection stages for each surveyor used in this research are given below.

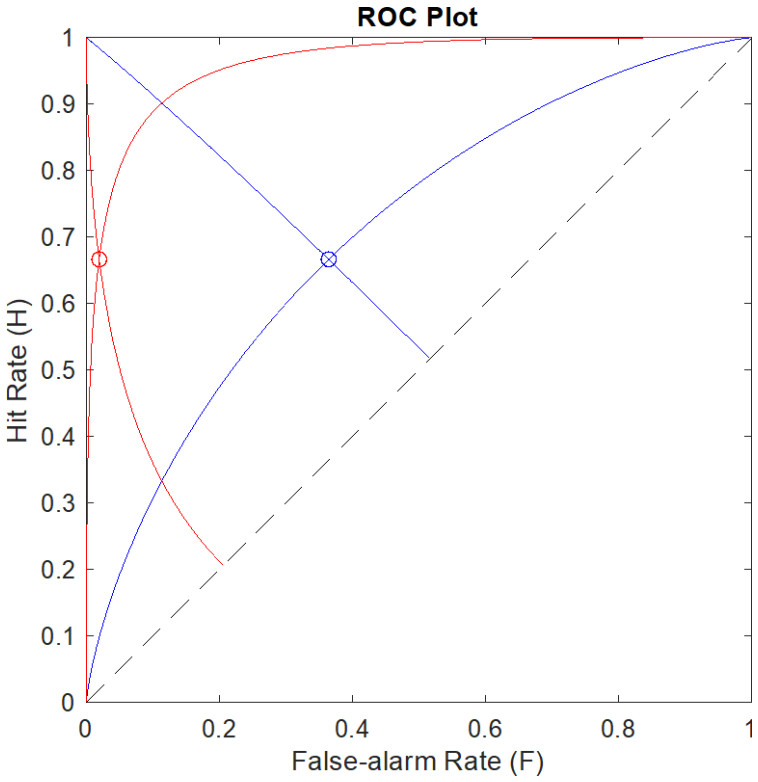


Figure 20. ROC Plot – Surveyor 1

In the ROC graph, the blue lines indicate the value of d' and the location of the criterion for the first (pause) stage of detection. The red lines correspond to the values of d' and c in the final (detection) stage. The intersections (and circles) can be related back to the HR and FAR found during the analysis of surveyor performance (Given in Appendices C and E)

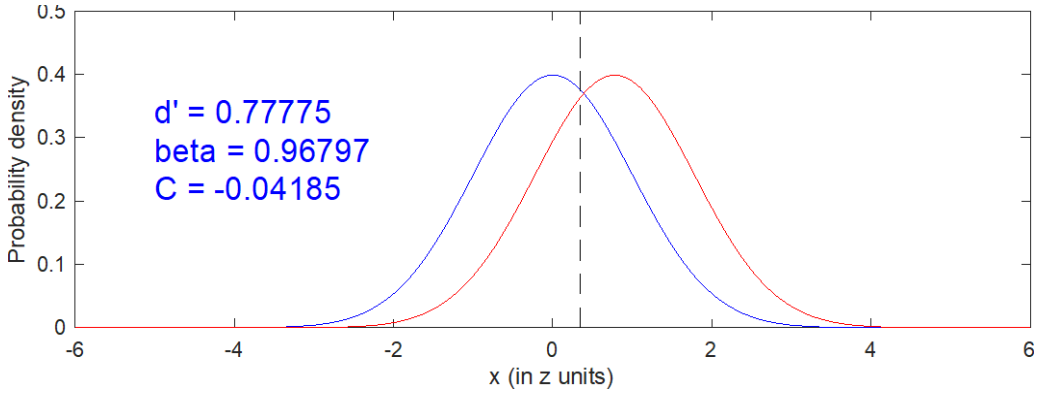


Figure 21. Probability density graph (S/S+N) – Surveyor 1 - Pause Stage

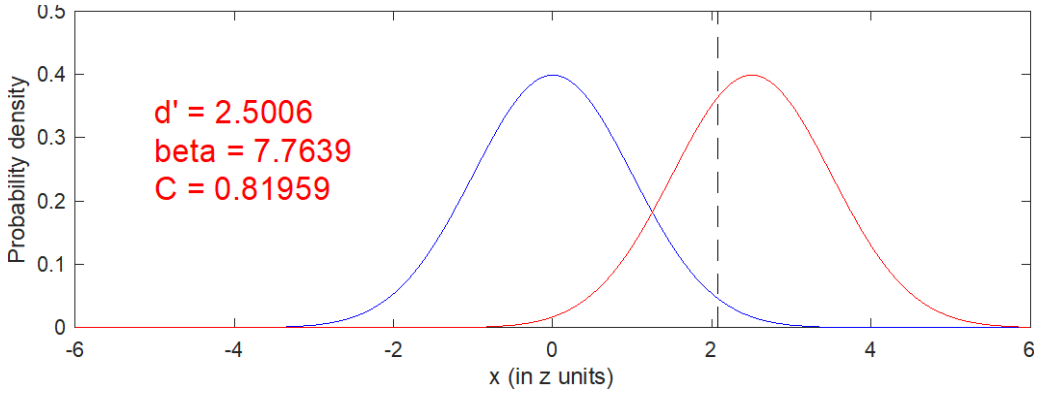


Figure 22. Probability density graph (S/S+N) – Surveyor 1 - Detection Stage

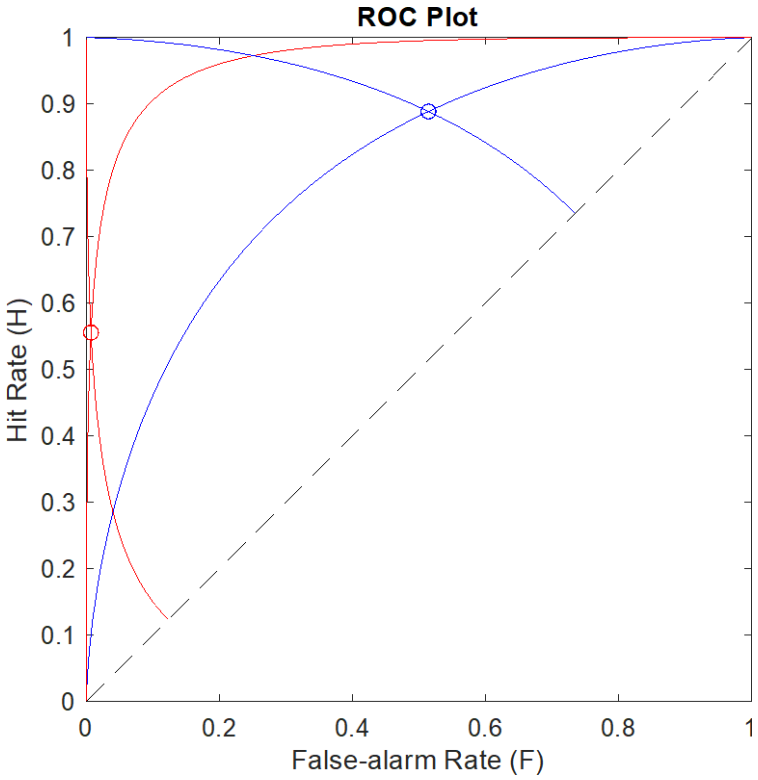


Figure 23. ROC Plot – Surveyor 2

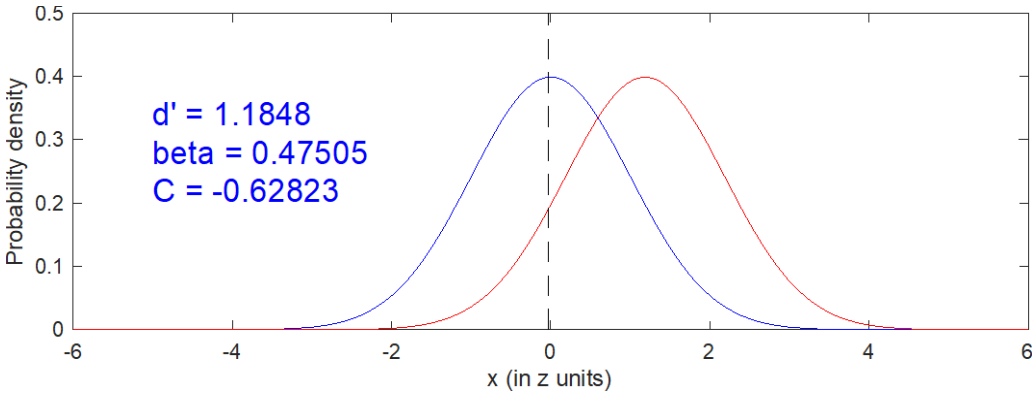


Figure 24. Probability density graph (S/S+N) – Surveyor 2 - Pause Stage

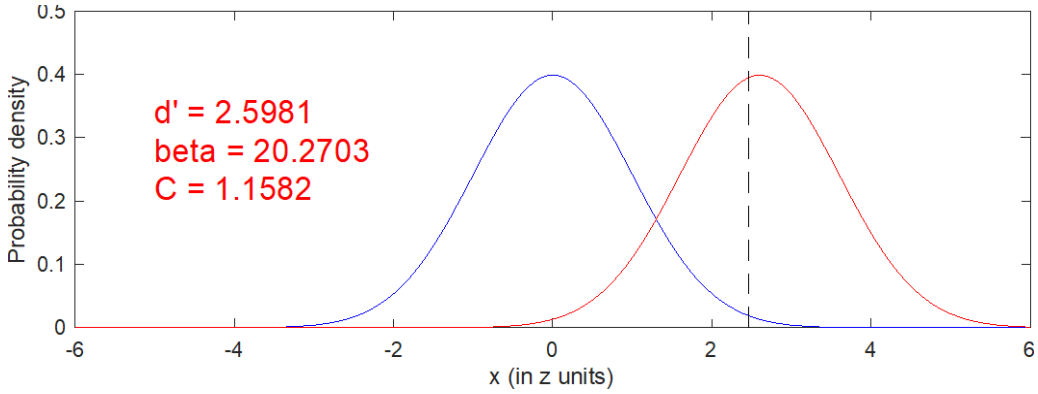


Figure 25. Probability density graph (S/S+N) – Surveyor 2 - Detection Stage

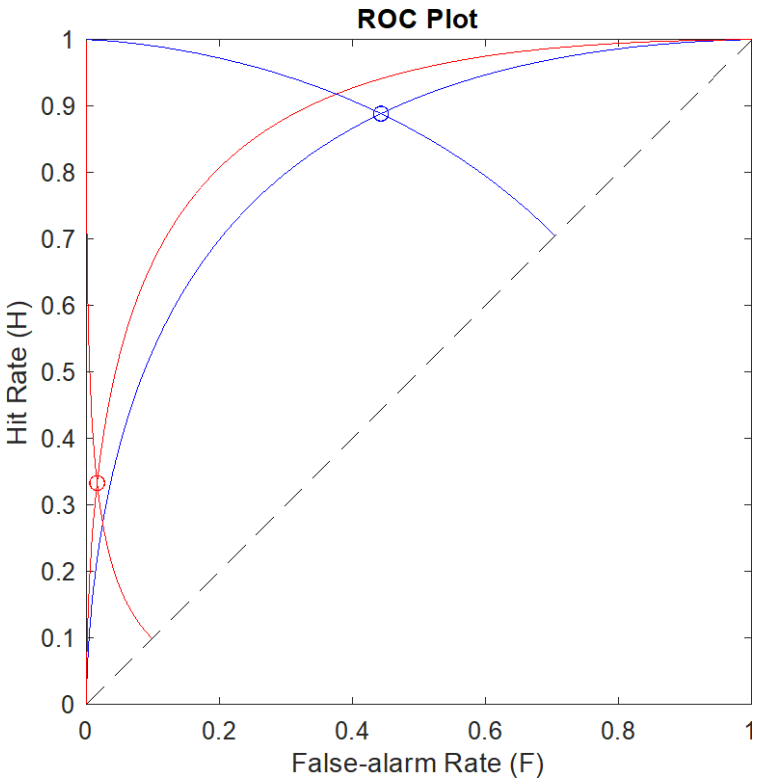


Figure 26. ROC Plot – Surveyor 3

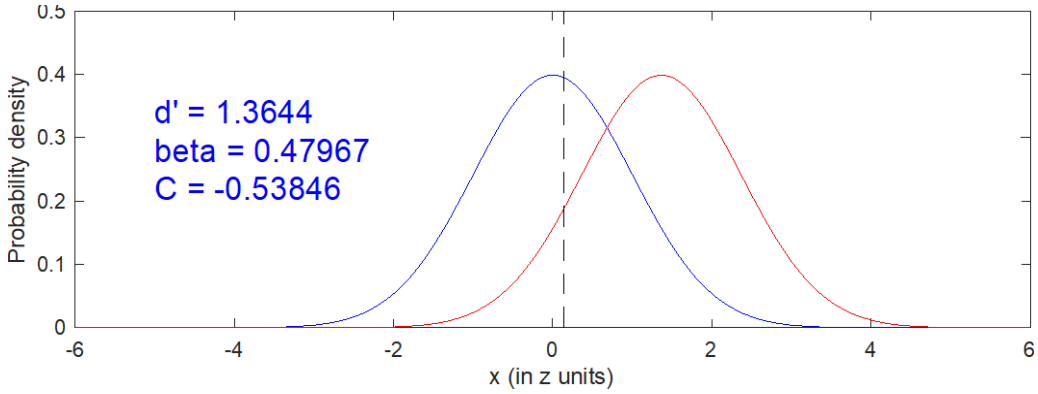


Figure 27. Probability density graph (S/S+N) – Surveyor 3 - Pause Stage

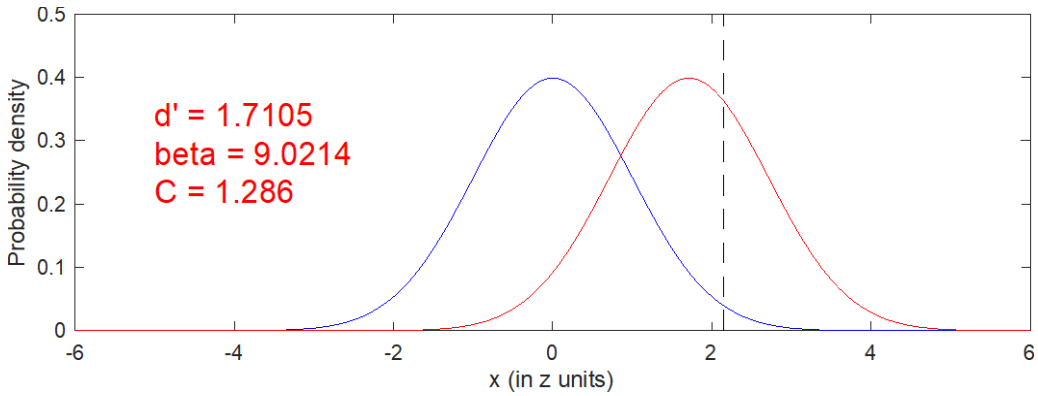


Figure 28. Probability density graph (S/S+N) – Surveyor 3 - Detection Stage

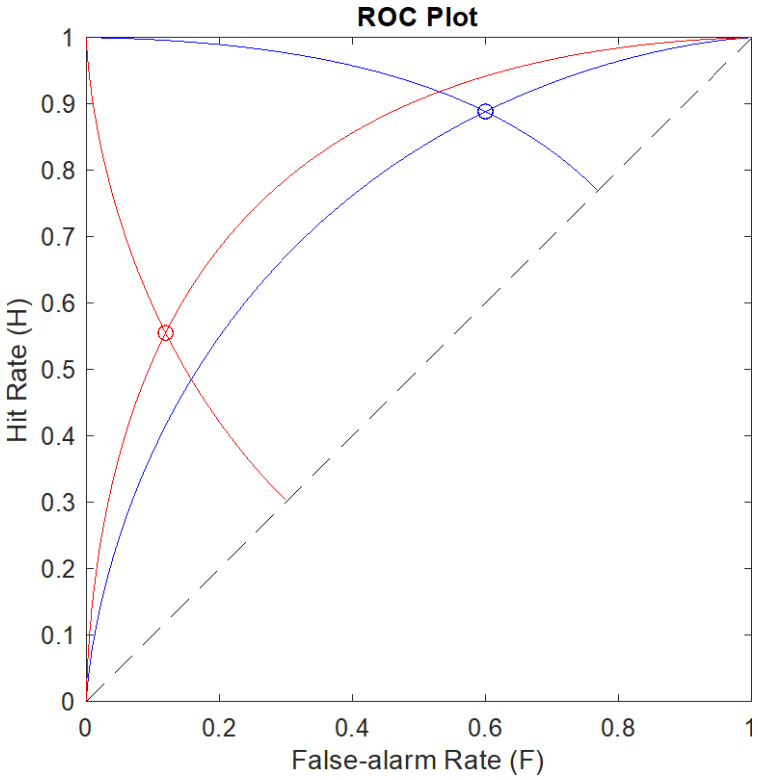


Figure 29. ROC Plot – Surveyor 4

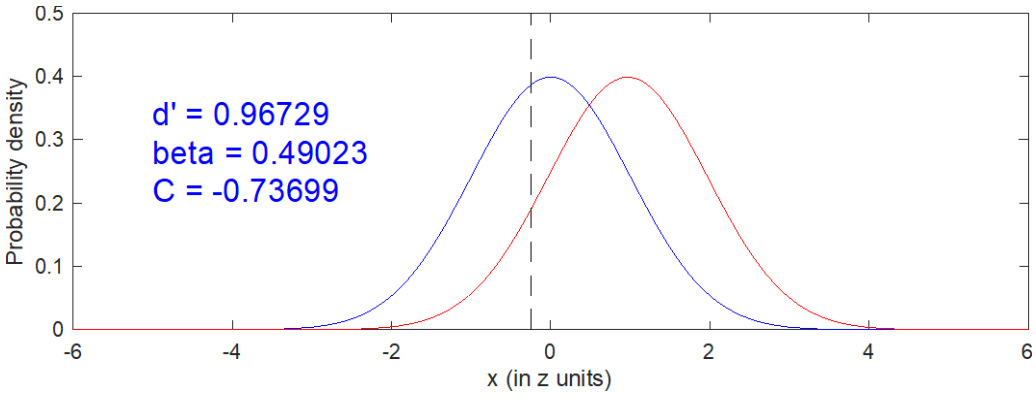


Figure 30. Probability density graph (S/S+N) – Surveyor 4 - Pause Stage

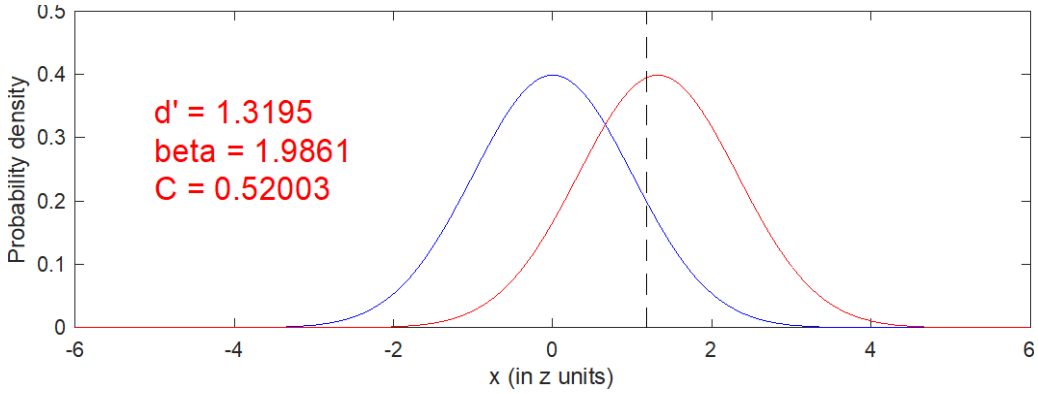


Figure 31. Probability density graph (S/S+N) – Surveyor 4 - Detection Stage

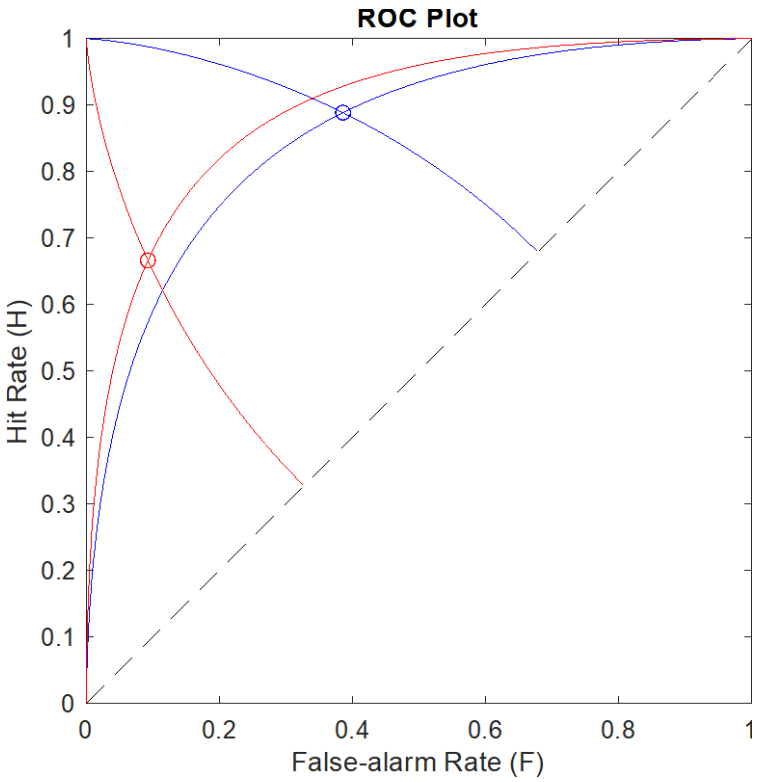


Figure 32. ROC Plot – Surveyor 5

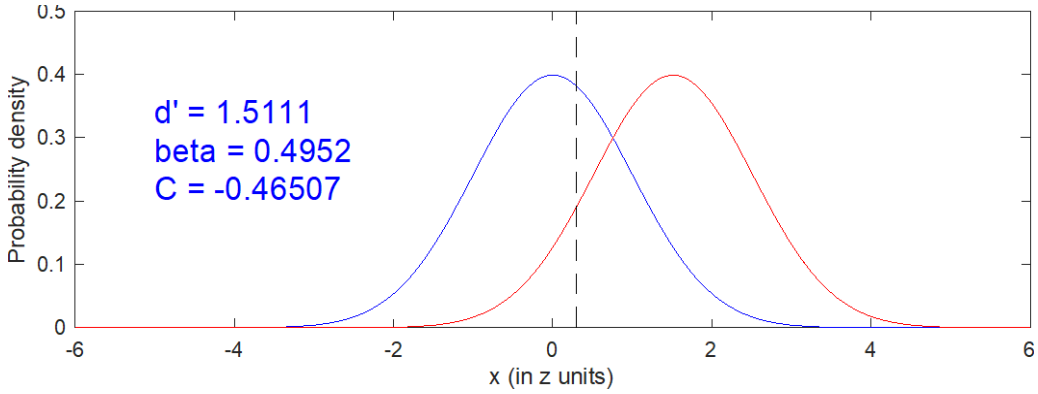


Figure 33. Probability density graph (S/S+N) – Surveyor 5 - Pause Stage

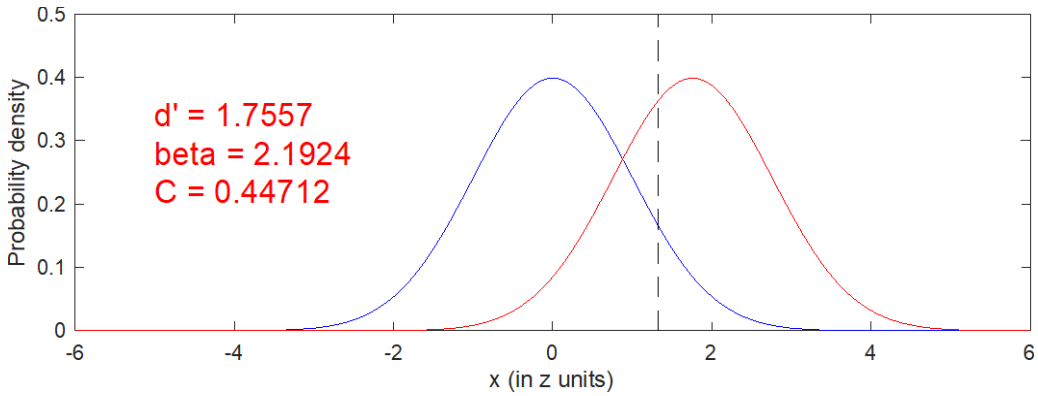


Figure 34. Probability density graph (S/S+N) – Surveyor 5 - Detection Stage

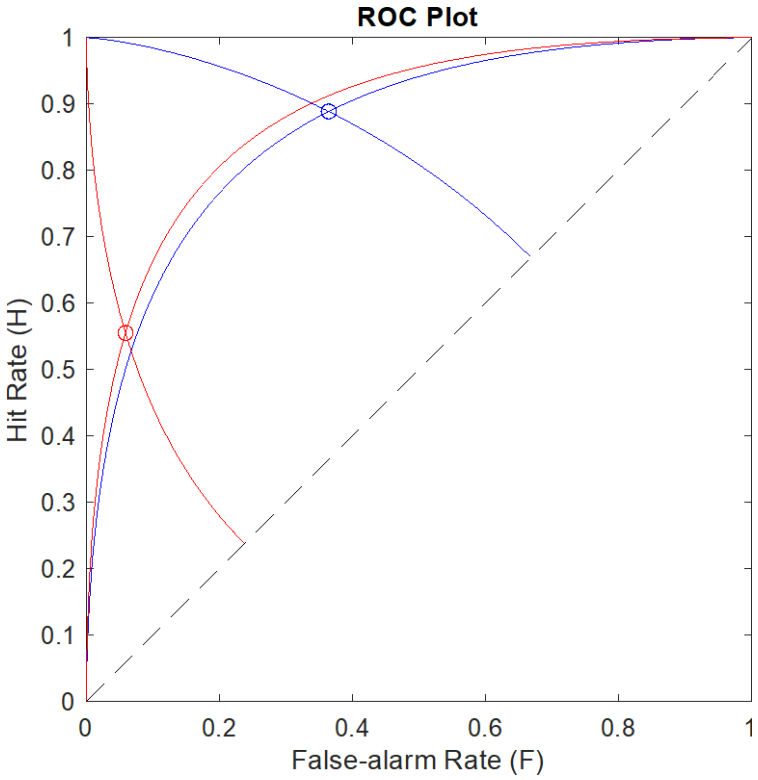


Figure 35. ROC Plot – Surveyor 6

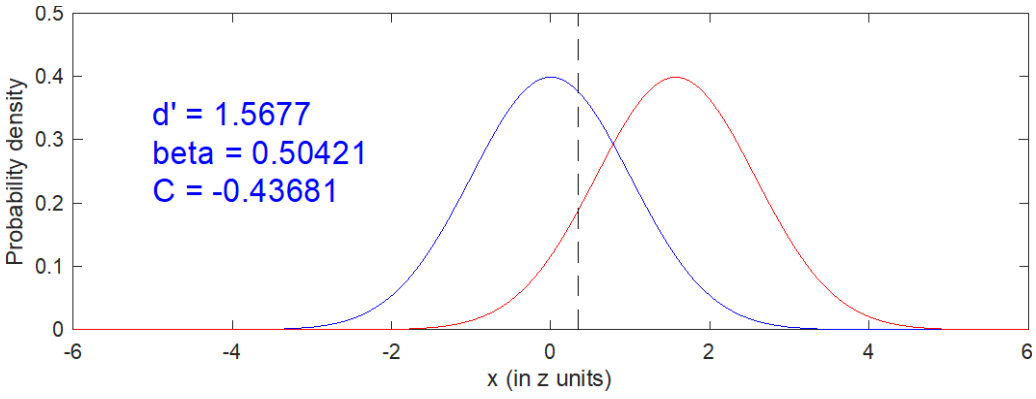


Figure 36. Probability density graph (S/S+N) – Surveyor 6 - Pause Stage

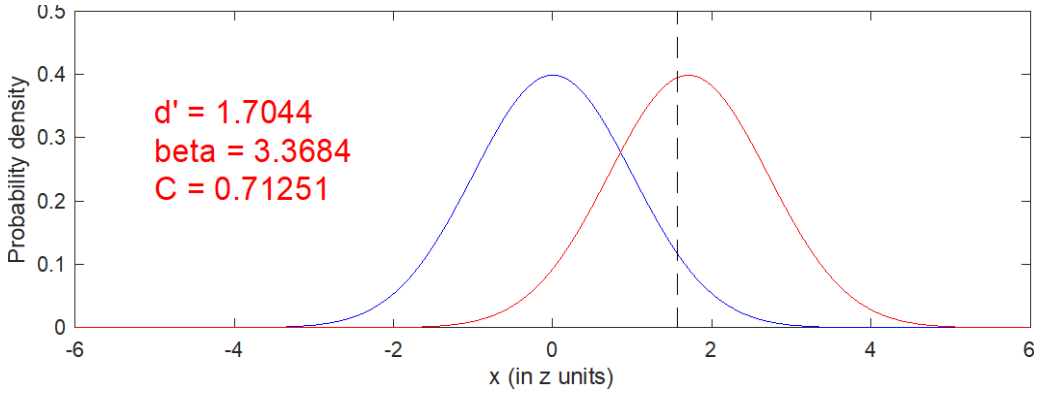


Figure 37. Probability density graph (S/S+N) – Surveyor 6 - Detection Stage

Appendix E. Surveyor Performance Calculations

Table 9. Surveyor Performance Calculations

	Surveyor 1		Surveyor 2		Surveyor 3		Surveyor 4		Surveyor 5		Surveyor 6		Avg	
	Pause	Decision	Pause	Decision	Pause	Decision	Pause	Decision	Pause	Decision	Pause	Decision	Pause	Decision
# hits	6	6	8	5	8	3	8	5	8	6	8	5	8	5
# misses	3	3	1	4	1	6	1	4	1	3	1	4	1	4
# false Alarms	51	1	72	0	62	1	84	10	54	5	51	3	62	3
# correction rejections	89	50	68	72	78	61	56	74	86	49	89	48	78	59
H (hit rate)	0.667	0.667	0.889	0.556	0.889	0.333	0.889	0.556	0.889	0.667	0.889	0.556	0.852	0.556
F (false-alarm rate)	0.364	0.019	0.514	0.007	0.443	0.016	0.600	0.106	0.386	0.085	0.364	0.056	0.445	0.051
z(H)	0.431	0.431	1.221	0.140	1.221	-0.431	1.221	0.140	1.221	0.431	1.221	0.140	1.044	0.140
z(F)	-0.347	-2.070	0.036	-2.460	-0.144	-2.148	0.253	-1.246	-0.291	-1.374	-0.347	-1.593	-0.138	-1.638
d'	0.778	2.501	1.185	2.600	1.364	1.717	0.967	1.386	1.511	1.805	1.568	1.733	1.182	1.777
c	-0.042	0.820	-0.628	1.160	-0.538	1.289	-0.737	0.553	-0.465	0.472	-0.437	0.727	-0.453	0.749
Beta	0.968	7.764	0.475	20.417	0.480	9.146	0.490	2.152	0.495	2.342	0.504	3.523	0.585	3.785
ln Beta	-0.033	2.049	-0.744	3.016	-0.735	2.213	-0.713	0.766	-0.703	0.851	-0.685	1.259	-0.536	1.331
Surveyor Eff (p)	0.068575875	**0.70890114**	0.15914558	**0.76626423**	0.21103338	**0.33416441**	0.10607265	**0.21768541**	0.25888125	**0.36917562**	0.27860868	**0.34044636**	0.158	**0.358**

Appendix F. Model Tasks and Data Collection Methodology

The data collection requirements for both the alternate MOPP and Heat stress models can be seen in Tables 10 and 11 below.

Table 10. Data Collection Requirements for MOPP Model

Task ID	Task Name	Data Requirements/Methods
0	Model START	None
1	Orient Probe	VACP Value = 12.60
2	Pause	Calculated estimate of the amount of time (in seconds) that the surveyor's probe was stationary divided by the total number of pauses, and multiplied by a factor of 0.25. VACP Value = 22.20
3	Pause on A Source	Probability of path logic execution to Task ID 4 & 5. Corresponds to the probability of the surveyor registering a "Hit" or a "Miss"
4	On Src (Answer Yes)	None
5	On Src (Answer No)	None
6	In Background (Answer No)	None
7	In Background (Answer Yes)	None
8	Stop	Calculated estimate of the amount of time (in seconds) that the surveyor's probe was stationary divided by the total number of pauses, and multiplied by a factor of 0.75. VACP Value = 22.20
9	Stop On Src (Answer Yes)	None
10	In Background	Probability of path logic execution to Task ID 6 & 7. Corresponds to the probability of the surveyor registering a "False Alarm" or a "Correct Rejection"
15	Stop On Src (Answer No)	None
16	Stop in Background	Probability of path logic execution to Task ID 18 & 19. Corresponds to the probability of the surveyor registering a "False Alarm" or a "Correct Rejection"
17	Stop on a Source	Probability of path logic execution to Task ID 9 & 15. Corresponds to the probability of the surveyor registering a "Hit" or a "Miss"
19	Stop In Background (Answer No)	None
26	Move Probe	Calculated estimate of the amount of time (in seconds) that the surveyor's probe was moving (Not stationary). VACP Value = 16.20
18	Stop In Background (Answer Yes)	None
999	Model END	None

Table 11. Data Collection Requirements for Heat Model

Task ID	Task Name	Data Requriements/Methods
0	Model START	None
1	Orient Probe	None
2	Pause	None
3	Pause on A Source	Probability of path logic execution to Task ID 4 & 5. Corresponds to the probability of the surveyor registering a "Hit" or a "Miss"
4	On Src (Answer Yes)	None
5	On Src (Answer No)	None
6	In Background (Answer No)	None
7	In Background (Answer Yes)	None
8	Stop	None
9	Stop On Src (Answer Yes)	None
10	In Background	Probability of path logic execution to Task ID 6 & 7. Corresponds to the probability of the surveyor registering a "False Alarm" or a "Correct Rejection"
15	Stop On Src (Answer No)	None
16	Stop in Background	Probability of path logic execution to Task ID 18 & 19. Corresponds to the probability of the surveyor registering a "False Alarm" or a "Correct Rejection"
17	Stop on a Source	Probability of path logic execution to Task ID 9 & 15. Corresponds to the probability of the surveyor registering a "Hit" or a "Miss"
19	Stop In Background (Answer No)	None
26	Move Probe	None
18	Stop In Background (Answer Yes)	None
999	Model END	None

Appendix G. VACP Sensitivity Analysis

The choice of VACP benchmark values are based upon SME input and thus, are considered subjective in nature. Workload demand values are automatically mapped to taxons, using the three taxons with the highest VACP benchmark scores to determine the proportional weighting assigned for a given task [26]. To determine the effect, if any, other benchmark values have on the overall taxon weighting assigned, a sensitivity analysis was conducted for the alternate IMPRINT models investigating the effects of MOPP and heat stress. Figures 11 and 12 indicate the VACP values chosen based on SME input. Table 8 in Appendix B lists the standardized descriptor info used to select the most appropriate value for the radiological detection task by the author.

Table 12 describes the effect of varying one or more VACP descriptor values on the overall measures of task time from MOPP wear.

Table 12. MOPP Sensitivity Analysis – Task Time

	Task Demand Values								
Task ID	Visual	Auditory	Cognitive	Fine Motor					
2	6.00	7.00	4.60	4.60		Min	Max	Mean	SD
8	6.00	7.00	4.60	4.60		51.00	124.00	89.00	14.00
26	6.00	1.00	4.60	4.60					
2	6.00	**6.60**	4.60	4.60		Min	Max	Mean	SD
8	6.00	**6.60**	4.60	4.60		50.31	124.00	88.00	14.23
26	6.00	1.00	4.60	4.60	RPD	**1.36215576**	**0**	**1.1299435**	**1.62947219**
2	**5.10**	**6.00**	4.60	4.60		Min	Max	Mean	SD
8	**5.10**	**6.00**	4.60	4.60		43	106	75	12.17
26	**5.10**	1.00	4.60	4.60	RPD	**17.0212766**	**15.6521739**	**17.0731707**	**13.9854796**
2	**5.10**	**6.00**	**1.20**	4.60		Min	Max	Mean	SD
8	**5.10**	**6.00**	**1.20**	4.60		46.00	117.00	83.00	14.00
26	**5.10**	1.00	**1.20**	4.60	RPD	**10.3092784**	**5.80912863**	**6.97674419**	**0**
2	**5.10**	**6.00**	**1.20**	**2.60**		Min	Max	Mean	SD
8	**5.10**	**6.00**	**1.20**	**2.60**		36.00	97.00	68.00	12.00
26	**5.10**	1.00	**1.20**	**2.60**	RPD	**34.4827586**	**24.4343891**	**26.7515924**	**15.3846154**

Varying the auditory channel by 0.4 (the next lowest increment available) results in a mean Relative Percent Difference (RPD) in task time of 1.1% compared to the

default SME input, with no change in the modeled maximum. When both the visual and auditory channels are modified by 0.9 and 0.4, respectively however, a mean RPD of 17% is found. Varying visual, auditory, and cognitive channels by 0.9, 0.4, and 3.4 (the next lowest selectable value) results in a mean RPD of 7%, due in part to a shift in the taxon weighting, according to the schema discussed previously. When all benchmarks are modified by the next lowest selectable value, an RPD of up to 27% for the mean task time is observed.

Similar results are obtained when evaluating the sensitivity of VACP input to other values within the heat stress model. The heat stress PSF uses the same VACP channels as MOPP, however the percent change can be associated with changes in accuracy, thus affecting the value of *d'*. At most, changes in VACP values across all channels result in a mean difference of about 26%, or about 0.3 units of *d'*. These results highlight the importance of selecting appropriate VACP values representative of the task based on experienced SME input.

Appendix H. Validation Boundary Results

*All values in units of *d'*

Surveyor	Source	Value of d'
1	Observed	2.501
2	Observed	2.6
3	Observed	1.717
4	Observed	1.386
5	Observed	1.805
6	Observed	1.733

		Run 1	Run 2	Run 3	Run 4	Run 5	Run 6	Run 7	Run 8	Run 9	Run 10		Mean	Standard dev	Upper	Lower	UPB	LB
1	Model	2.04	1.98	2.05	2	2.02	1.66	1.68	1.97	2.01	2.02		1.943	0.145987062	2.02762074	1.85837926	0.47337926	0.64262074
2	Model	2.5	2.5	2.49	2.49	2.51	2.49	2.46	2.47	2.48	2.48		2.487	0.014944341	2.49566242	2.47833758	0.10433758	0.12166242
3	Model	1.49	1.19	1.24	1.24	1.24	1.24	1.24	1.52	1.53	1.5		1.343	0.144917601	1.42700083	1.25899917	0.28999917	0.45800083
4	Model	1.26	1.27	1.2	1.19	1.19	1.2	1.16	1.17	1.13	1.12		1.189	0.048636977	1.2171922	1.1608078	0.1688078	0.2251922
5	Model	1.55	1.52	1.38	1.4	1.42	1.28	1.37	1.38	1.4	1.32		1.402	0.081486195	1.44923311	1.35476689	0.35576689	0.45023311
6	Model	1.41	1.41	1.25	1.15	1.28	1.15	1.24	1.27	1.28	1.27		1.271	0.087869853	1.32193336	1.22006664	0.41106664	0.51293336

Appendix I. Values of *d'* for Selected True and False Positive Proportions

Reproduced from Brown & Albequist (1998), *Human Performance in Radiological Survey Scanning*, NUREG/CR-6364 [1]

Table 13. Values of *d'* with associated proportions

False Positive Proportion	True Positive Proportion							
	0.6	**0.65**	**0.7**	**0.75**	**0.8**	**0.85**	**0.9**	**0.95**
0.05	1.90	2.02	2.16	2.32	2.48	2.68	2.92	3.28
0.10	1.54	1.66	1.80	1.96	2.12	2.32	2.56	2.92
0.15	1.30	1.42	1.56	1.72	1.88	2.08	2.32	2.68
0.20	1.10	1.22	1.36	1.52	1.68	1.88	2.12	2.48
0.25	0.93	1.06	1.20	1.35	1.52	1.72	1.96	2.32
0.30	0.78	0.91	1.05	1.20	1.36	1.56	1.80	2.16
0.35	0.64	0.77	0.91	1.06	1.22	1.42	1.66	2.02
0.40	0.51	0.64	0.78	0.93	1.10	1.30	1.54	1.90
0.45	0.38	0.52	0.66	0.80	0.97	1.17	1.41	1.77
0.50	0.26	0.38	0.52	0.68	0.84	1.04	1.28	1.64
0.55	0.12	0.26	0.40	0.54	0.71	0.91	1.15	1.51
0.60	0.00	0.13	0.27	0.42	0.58	0.82	1.02	1.38

Bibliography

[1] W. S. Brown and E. W. Abelquist, “Human performance in radiological survey scanning,” *NUREG/CR-6364*, Mar. 1998, doi: 10.2172/573244.

[2] J. T. Wixted, “The Forgotten History of Signal Detection Theory,” *J. Exp. Psychol. Learn. Mem. Cogn.*, vol. 46, no. 2, pp. 201–233, 2019, doi: 10.1037/xlm0000732.

[3] C. D. Hautus, Michael J.; Macmillan, Neil A.; Creelman, *Detection Theory: A User’s Guide: 3rd edition*, 3rd ed. 2022.

[4] M. Smith and E. A. Wilson, “A model of the auditory threshold and its application to the problem of the multiple observer.,” *Psychol. Monogr. Gen. Appl.*, vol. 67, no. 9, pp. 1–35, 1953, doi: 10.1037/h0093654.

[5] D. M. Green and J. A. Swets, *Signal Detection Theory and Psychophysics*. Wiley, 1966.

[6] A. I. Schulman and G. Z. Greenberg, “Operating characteristics and a priori probability of the signal.,” *Percept. Psychophys.*, vol. 8, no. 5-A, pp. 317–320, 1970, doi: 10.3758/BF03212600.

[7] W. P. Tanner Jr. and J. A. Swets, “A decision-making theory of visual detection.,” *Psychological Review*, vol. 61, no. 6. American Psychological Association, US, pp. 401–409, 1954, doi: 10.1037/h0058700.

[8] W. Peterson, T. Birdsall, and W. Fox, “The theory of signal detectability,” *Trans. IRE Prof. Gr. Inf. Theory*, vol. 4, no. 4, pp. 171–212, 1954, doi: 10.1109/TIT.1954.1057460.

[9] J. A. Swets, “The Relative Operating Characteristic in Psychology,” *Science (80-.).*, vol. 182, no. 4116, pp. 990–1000, Dec. 1973, doi: 10.1126/science.182.4116.990.

[10] H. Liu, G. Li, W. G. Cumberland, and T. Wu, “Testing Statistical Significance of the Area under a Receiving Operating Characteristics Curve for Repeated Measures Design with Bootstrapping,” *J. Data Sci.*, vol. 3, no. 3, pp. 257–278, 2021, doi: 10.6339/jds.2005.03(3).206.

[11] I. Serrano-Pedraza, “Matlab software for sound and image (2D and 3D) processing and presentation.” Teaching Innovation Project, 2018, [Online]. Available: https://www.ucm.es/serranopedrazalab/software-1.

[12] W. H. Beyer, *Handbook of Tables for Probability and Statistics*, 2nd ed. CRC Press, 2018.

[13] U.S. Nuclear Regulatory Commission, “Minimum Detectable Concentrations with Typical Radiation Survey Instruments for Various Contaminants and Field Conditions,” vol. NUREG 1507, 2020.

[14] A. M. Law, *Simulation Modeling and Analysis*, 5th ed. New York, NY, USA: McGraw-Hill, 2015.

[15] M. Gunal, M; Pidd, “Discrete event simulation for performance modeling in healthcare: a review of the literature,” *J. Simul.*, vol. 4, no. 1, pp. 42–51, 2010.

[16] R. R. Hill, J. O. Miller, and G. A. McIntyre, “Applications of discrete event simulation modeling to military problems,” in *Proceeding of the 2001 Winter Simulation Conference (Cat. No.01CH37304)*, 2001, vol. 1, pp. 780–788 vol.1,

doi: 10.1109/WSC.2001.977367.

[17] S. Mason, T. Baines, J. M. Kay, and J. Ladbrook, “Improving the design process for factories: Modeling human performance variation,” *J. Manuf. Syst.*, vol. 24, no. 1, pp. 47–54, 2005, doi: https://doi.org/10.1016/S0278-6125(05)80006-8.

[18] A. Greasley, “Methods of modelling people using discrete-event simulation,” *SIMULTECH 2016 - Proc. 6th Int. Conf. Simul. Model. Methodol. Technol. Appl.*, no. Simultech, pp. 312–317, 2016, doi: 10.5220/0006005803120317.

[19] C. F. Rusnock and C. D. Geiger, “Simulation-Based Evaluation of Adaptive Automation Revoking Strategies on Cognitive Workload and Situation Awareness,” *IEEE Trans. Human-Machine Syst.*, vol. 47, no. 6, pp. 927–938, 2017, doi: 10.1109/THMS.2016.2618004.

[20] J. Keller, “Human performance modeling for discrete-event simulation: Workload,” *Winter Simul. Conf. Proc.*, vol. 1, pp. 157–162, 2002, doi: 10.1109/wsc.2002.1172879.

[21] J. Hugo and D. Gertman, “The use of computational human performance modeling as task analysis tool,” *8th Int. Top. Meet. Nucl. Plant Instrumentation, Control. Human-Machine Interface Technol. 2012, NPIC HMIT 2012 Enabling Futur. Nucl. Energy*, vol. 1, pp. 90–101, 2012.

[22] C. D. Wickens, *Engineering psychology and human performance*. New York, NY: HarperCollins Publishers, 1992.

[23] Army Research Laboratory, *Validation of a Task Network Human Performance Model of Driving*, no. ARL-RP-0174. 2007.

[24] D. Boeke, “Exploring Individual Differences in Workload Assessment,” 2014.

[25] Alion Science and Technology, “IMPRINT PRO.” Boulder, 2022.

[26] Alion Science and Technology, *IMPRINT Pro User Guide*. 2018.

[27] D. Powers and A. Marquis Gacy, “Using IMPRINT to model operator staffing and workload considerations in a 24/7 full-service mission operations center,” *15th Int. Conf. Sp. Oper. 2018*, no. June, pp. 1–12, 2018, doi: 10.2514/6.2018-2394.

[28] L. A. Currie, “Limits for qualitative detection and quantitative determination. Application to radiochemistry,” *Anal. Chem.*, vol. 40, no. 3, pp. 586–593, Mar. 1968, doi: 10.1021/ac60259a007.

[29] United States Government, *NUREG 1575 - Multi-Agency Radiation Survey and Site Investigation Manual (MARSSIM)*, vol. 2000, no. NUREG-1575, Rev. 1. 2000.

[30] J. Tzelgov, R. Srebro, A. Henik, and A. Kushelevsky, “Radiation Search and Detection by Ear and by Eye,” *Hum. Factors*, vol. 29, no. 1, pp. 87–95, Feb. 1987, doi: 10.1177/001872088702900110.

[31] S. M. Casey, “Inspection of Low-Level Radioactive Waste,” *Hum. Factors*, vol. 33, no. 1, pp. 1–15, Feb. 1991, doi: 10.1177/001872089103300101.

[32] D. A. King, N. Altic, and C. Greer, “Minimum detectable concentration as a function of gamma walkover survey technique,” *Health Phys.*, vol. 102, no. 2 SUPPL. 1, 2012, doi: 10.1097/HP.0b013e318237e757.

[33] J. T. Falkner and C. M. Marianno, “Validating a Methodology That Associates Minimum Detectable Activity with Detector Velocity,” *Health Phys.*, vol. 121, no.

1, pp. 30–37, 2021, doi: 10.1097/HP.0000000000001406.

[34] W. S. BROWN, “A TWO-STAGE MODEL OF RADIOLOGICAL INSPECTION: SPENDING TIME,” 2000, Accessed: Jan. 23, 2022. [Online]. Available: http://inis.iaea.org/Search/search.aspx?orig_q=RN:32003044.

[35] The MathWorks Inc, “MATLAB.” Natick, Massachusetts, 2021.

[36] H. M. Davis, Edward G.; Wick, Charles H.; Salvi, Lucia; Kash, “Soldier Performance of Military Operational Tasks Conducted While Wearing Chemical Individual Protective Equipment (IPE): Data Analysis In Support of the Revision of the U.S. Army Field Manual on NBC Protection (FM 3-4),” no. April, 1990.

[37] H. L. Taylor and J. Orlansky, “THE EFFECTS OF WEARING PROTECTIVE CHEMICAL WARFARE COMBAT CLOTHING ON HUMAN PERFORMANCE,” 1991.

[38] J. F. Parker Jr. and V. R. West, *Bioastronautics Data Book: Second Edition. NASA SP-3006.*, vol. 3006. 1973.

[39] N. Gaoua, S. Racinais, J. Grantham, and F. El Massioui, “Alterations in cognitive performance during passive hyperthermia are task dependent,” *https://doi.org/10.3109/02656736.2010.516305*, vol. 27, no. 1, pp. 1–9, Feb. 2011, doi: 10.3109/02656736.2010.516305.

[40] R. Frost, S; Mogridge, “Physiological safety of airfed suit use during nuclear decommissioning,” 2008.

[41] I. T. Parsons, M. J. Stacey, and D. R. Woods, “Heat Adaptation in Military Personnel: Mitigating Risk, Maximizing Performance,” *Front. Physiol.*, vol. 10, p.

1485, Dec. 2019, doi: 10.3389/FPHYS.2019.01485/BIBTEX.

[42] Department of the Air Force, *Air Force Manual 10-2503, Operations In A Chemical, Biological, Radiological, and Nuclear (CBRN) Environment*. 2019.

[43] E. W. Weisstein, "Triangular Distribution," Accessed: Feb. 14, 2022. [Online]. Available: https://mathworld.wolfram.com/.

[44] Kromek, "UAV Radiation Mapping Drone." https://www.kromek.com/product/aerial-radiation-mapping-drone/ (accessed Jan. 26, 2022).

[45] L. R. Pinto *et al.*, "Radiological scouting, monitoring and inspection using drones," *Sensors*, vol. 21, no. 9, p. 3143, May 2021, doi: 10.3390/s21093143.

[46] C. J. Buck-Gengler, W. D. Raymond, A. F. Healy, and L. E. Bourne Jr., "Modeling cognitive tasks in IMPRINT.," in *Training cognition: Optimizing efficiency, durability, and generalizability.*, New York, NY, US: Psychology Press, 2012, pp. 201–224.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From – To)*
24-03-2022	Master's Thesis	September 2020 – March 2022

TITLE AND SUBTITLE

Analysis of Task Performance During Radiological Surveillance by Means of Discrete Event Simulation

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Ames, Michael H., MSgt, USAF

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(S)
Air Force Institute of Technology
Graduate School of Engineering and Management (AFIT/EN)
2950 Hobson Way, Building 640
WPAFB OH 45433-8865

8. PERFORMING ORGANIZATION REPORT NUMBER

AFIT-ENV-MS-22-M-175

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
United States Air Force School of Aerospace Medicine
2510 5th Street, Bldg. 840
WPAFB OH 45433-7913

10. SPONSOR/MONITOR'S ACRONYM(S)

USAFSAM/OE

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
DISTRUBTION STATEMENT A. APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

13. SUPPLEMENTARY NOTES
This material is declared a work of the U.S. Government and is not subject to copyright protection in the United States.

14. ABSTRACT
The surveillance and detection of radioactive contamination on surfaces and in the environment are commonly investigated by surveyors utilizing portable detection equipment. The availability of Discrete Event Simulation (DES) and Human Performance Modeling (HPM) allows for the analysis of physical and cognitive processes associated with these operations, as well as the effect that external environmental factors have on surveyor performance. This research uses the Improved Performance Research Integration Tool (IMPRINT) to approximate the performance of a radiological detection task informed by the observation of six surveyors. The effects of chemical Individual Protective Equipment (IPE) use is evaluated along with the effects of elevated ambient environmental temperatures. Along with the development of a novel human performance model for the surveillance task, results of this study predict up to a 33% increase in survey completion time when chemical IPE is worn and up to a 50% decrease in surveyor efficiency from the effects of elevated ambient temperatures. Overall, this study represents the novel use of a DES to model the cognitive and physical tasks associated with radiological surveillance activities and the impacts from key physical and environmental stressors.

15. SUBJECT TERMS
Discrete Event Simulation, Radiological Detection, Human Performance Modeling, IMPRINT

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Jeremy M. Slagley, AFIT/ENV
U	U	U	UU	123	19b. TELEPHONE NUMBER *(Include area code)* (937) 255-3636, ext 4632 Jeremy.slagley@afit.edu

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39-18