

To the Graduate Council:

I am submitting herewith a dissertation written by Cameron Wright Coates entitled "System Effectiveness Model Formulation With Application To Nuclear Safeguards Systems." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Industrial Engineering

Denise Ford Jackson, Major Professor

We have read this dissertation
and recommend its acceptance:

Dr. Xueping Li

Dr. Charles Noon

Dr. Charles H. Aikens

Acceptance for the Council:

Carolyn R. Hodges
Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records)

System Effectiveness Model Formulation With Application To Nuclear Safeguards Systems

**A Dissertation
Presented for the
Doctor of Philosophy Degree
The University of Tennessee, Knoxville**

**Cameron Wright Coates
May 2009**

Dedication

I dedicate this dissertation to my father, Herbert Wright Coates born February 6, 1921. A farm boy from a very rural town; Gould Station, Quebec, Canada, he volunteered for the Royal Canadian Air Force during World War II. He was unable to qualify as a pilot due to color blindness and became an airplane mechanic, which was a natural progression from life as a self reliant farmer. After the war ended he was able to attend Bishop's University in Lennoxville, Quebec with the help of the Canadian GI Bill benefits. As a married vet attending university to become a chemist, life was no doubt difficult, but he, with the help of his wife, my mother Kathleen, persevered and eventually went to work at the Chalk River Nuclear Power Plant near Deep River Ontario, Canada, where I was born.

Herb eventually immigrated with his family in tow to Charlotte, North Carolina in the USA. He left the nuclear industry behind and started a new career in textiles working for Celanese Inc. and Fibers Industries Inc. While I have some memories of living in Canada, I grew up from the age of three to adulthood as a southern boy. Herb always counseled me to become an engineer as he was "just" a chemist. Nevertheless he had a distinguished career as a chemist with numerous patents in an era when the pay from your employer for a patent was a single dollar.

Thanks Dad.

Acknowledgments

I want to acknowledge several groups and an individual for their support to me during my work to obtain a PhD in Industrial Engineering.

First, my family has been very supportive in every way that counts. They have offered encouragement, love and support during this journey. Second, my employer (UT-Battelle) has been outstanding in supporting my efforts financially and through rearrangement of my work schedule. Third, the INMM has supplied encouragement and support for my research. Fourth, my graduate committee whose guidance has been spot on and was essential to the dissertation. Last but certainly not least, my graduate advisor, Dr. Denise Ford Jackson has been with me every step of the way.

Thanks to you all.

Abstract

Evaluation of a given system's effectiveness has numerous pitfalls. System objectives may be poorly defined, may shift during the system life, or may be hard to quantify. Further, individual perceptions of the quantifications may differ. Whatever the cause, system effectiveness has been an elusive term to quantitatively define. This research posits a quantitative system effectiveness model and establishes a utilitarian approach for use with an illustrative application to an operating nuclear safeguards system.

The Department of Energy (DOE) defines domestic safeguards, which are applied to nuclear material as; "an integrated system of physical protection, material accounting, and material control measures designed to deter, prevent, detect, and respond to unauthorized possession, use, or sabotage of nuclear materials."

This research includes the investigation of the utility coefficients and simulation of a domestic nuclear safeguards system, as well as simulation of an airport passenger screening system consisting of: an identification screening system; an X-ray system for checking bags and computers; and a walk through metal detector.

Expert judgment was used to determine the relative importance (utility) of the individual subsystems through a statistically analyzed web survey. The survey population is nuclear material protection, control, accounting, and plant management experts. The mean utility coefficients determined during the survey were applied to the system components developed assigned randomly generated values of component effectiveness and combined to produce an overall system effectiveness. Simulated Type I and Type II error rates are used for illustration of the probabilistic methodology currently used by DOE (calculating protection effectiveness) and the posited and heuristically based methodology (system effectiveness). Use of the heuristically based system effectiveness methodology illustrates an approach that combines the subsystem components of plant management, physical protection, material accounting, and material control for a domestic safeguards system. The system effectiveness methodology is complimentary to and more robust than the protection effectiveness calculation and can offer opportunities for cost savings during the system lifecycle.

Table of Contents

Chapter 1	Introduction.....	1
1.1	Background	1
1.2	Significance of Safeguard Systems	5
1.4	Research Questions and Hypotheses	9
1.4	Summary and Introduction of Remaining Chapters	10
Chapter 2	Literature Review.....	11
2.1	Systems Theory.....	11
2.2	Specific Hypothesis	17
2.3	Applicable Theory.....	19
2.4	Approaches	28
2.5	Prior Work Related to Effectiveness	28
2.6	Area of Application	29
2.7	Utility Theory	31
2.8	Facility Scenarios	34
Chapter 3	Research Methodology.....	39
3.1	Introduction To System Effectiveness Function.....	41
3.2	Function Development and Use	44
3.3	Case Study	49
3.4	Safeguards System Effectiveness	54
3.5	Description of the Research Survey Design	56
3.6	Description of the Surveyed Population.....	56
3.7	Description of the Sample	57
3.8	Instruments.....	57
3.9	Procedures.....	58
3.10	Survey Data Analysis	58
3.11	System Simulations	62
3.12	Methodology Summary.....	67
Chapter 4	Analysis	70
4.1	Survey Data.....	70
4.2	Simulation Data	75
4.3	Validating the Methodology With Airport Simulation	81
4.4	Validating the Methodology With Safeguards Simulation	83
4.5	Verifying The Results	88
Chapter 5	Conclusions and Recommendations	89
5.1	Summary	89
5.2	Conclusions.....	93
5.3	Recommendations for Future Work.....	94
References		95
APPENDICES		100
Appendix I	Safeguards Survey	101
Appendix II	Survey Results.....	112
Vita		150

List of Tables

Table 1 System Intentions.....	14
Table 2 Type I and II Errors.....	19
Table 3 Type I and II Errors as Alarms	24
Table 4 MAUT Worksheet.....	33
Table 5 Comparison of Methods	36
Table 6 Example Calculation Of Individual Normalized Coefficients	61
Table 7 Example Subsystem Effectiveness Calculation.....	71
Table 8 Mean Response Tabulation.....	73
Table 9 Example Overall Effectiveness Calculation	76
Table 10 Output Data from Simulation	78
Table 11 Comparison of Effectiveness Calculation with Probabilistic Method.....	82
Table 12 Theoretical Alarm Rates for Safeguards Simulation.....	84
Table 13 Simulated Error Rates	85

List of Figures

Figure 1 Image of DOE Nuclear Materials Table.....	7
Figure 2 Image of DOE Graded Safeguards Table	8
Figure 3 Example of Graded Access.....	9
Figure 4 General System Structure	13
Figure 5 Risk, Consequences, Threats and Vulnerabilities	20
Figure 6 Hypothesis Testing Errors	23
Figure 7 Proposed Basis of Effectiveness.....	26
Figure 8 ROC Applet Output 1	27
Figure 9 ROC Applet Output 2	27
Figure 10 Example of Three Layer Facility	35
Figure 11 Three Dependent Layers.....	35
Figure 12 Three Independent Layers	36
Figure 13 Generic System Lifecycle.....	38
Figure 14 Effectiveness as a Simple Function of Confidence Level and Power.....	45
Figure 15 First Iteration Solution Space with $r = 2$	46
Figure 16 First Iteration Solution Space $r = 200$	46
Figure 17 Second Iteration Solution Space with $r = 2$	48
Figure 18 Second Iteration Solution Space with $r = 200$	48
Figure 19 System Effectiveness Function Solution Space $r = 2, c = 1$	50
Figure 20 System Effectiveness Function Solution Space $r = 200, c = 1$	50
Figure 21 System Effectiveness Function Solution Space $r = 200, c = 5$	51
Figure 22 System Stress Factor Curves.....	51
Figure 23 Airport Passenger Security Screening	52
Figure 24 Example Baseline Airport Solution Space	53
Figure 25 Example Airport Security Under Stress	53
Figure 26 Example Airport Security Under Stress With Compensation.....	54
Figure 27 Sample Safeguards Survey.....	59
Figure 28 Question As The Respondent Would See It.....	60
Figure 29 Example Plant Management Component Rankings	61
Figure 30 Airport Security Check Point Flow Chart	63
Figure 31 Simple Domestic Safeguards Simulation Modules	64
Figure 32 Accounting Sub-model Modules.....	65
Figure 33 Security Sub-model Modules	66
Figure 34 Decisions Probabilities and Error Rates for Process Lines	68
Figure 35 Decision Probabilities and Error Rates for Security Line	69
Figure 36 Full Utility Model Calculations with Random Performance Levels.....	72
Figure 37 Simulation of Airport Passenger Screening	77
Figure 38 Passenger Screening Simulation Output for 100,000 Travelers.....	80
Figure 39 Convergence of Simulated Airport Model to Theoretical Value	86
Figure 40 Convergence of Simulated Safeguards System Model to Theoretical Value	87

Nomenclature

α	Type I error rate
β	Type II error rate
c	The stress level to the system
r	The unstressed tolerance of false alarms to missed alarms (α_0/β_0)
BD	Bag Denied
BTIE	Bag Type I Error
BTIIE	Bag Type II Error
CD	Computer Denied
CSV	Excel comma-separated-values
CTIE	Computer Type I Error
CTIIE	Computer Type II Error
DB	Denied Boarding from Computer, Bag, or Metal
DOE	Department of Energy
FMEA	Failure Modes And Effects Analysis
FNF	False Negative Fraction (a.k.a. Type II error),
FPF	False Positive Fraction (a.k.a. Type I error)
FTIE	Frisk Type I Error
FTIIE	Frisk Type II Error
IAEA	International Atomic Energy Agency
IDD	Identification Denied
IDTIE	Identification Type I Error
IDTIIE	Identification Type II Error
INCOSE	International Council on Systems Engineering
INMM	Institute of Nuclear Material Management
K_a	$(r)^{1/c}$ for $r > 1$; otherwise, $K_a = 1$
K_b	$(r)^{-1/c}$ for $r < 1$; otherwise, $K_b = 1$
MA	Material Accounting
MArank	Material Accounting Component Rank
MC	Material Control
MC&A	Material Control and Accounting
MCrank	Material Control Component Rank
MDIIE	Walk Through Metal Detector Type II Error
MDOK	Metal OK
MetDen	Metal Denied
NRC	Nuclear Regulatory Commission
PM	Plant Management
PMrank	Plant Management Component Rank

<i>PP</i>	Physical Protection
<i>PPrank</i>	Physical Protection Component Rank
<i>ROC</i>	Receiver (or relative) operating characteristic
<i>SCC</i>	UT Statistical Consulting Center
<i>SNM</i>	Special Nuclear Material
<i>TNF</i>	True Negative Fraction
<i>Total TI</i>	Total Type I Errors
<i>Total TII</i>	Total Type II Errors
<i>TPF</i>	True Positive Fraction
<i>TSA</i>	Transportation Security Administration
<i>u_a</i>	The utility of the material accounting subsystem
<i>u_c</i>	The utility of the material control subsystem
<i>u_{ij}</i>	The utility of the j^{th} component of the i^{th} subsystem
<i>u_m</i>	The utility of the management subsystem
<i>UN</i>	United Nations
<i>UNSCR</i>	United Nations Security Council Resolution
<i>u_p</i>	The utility of the physical protection subsystem
<i>WTMD</i>	Walk Through Metal Detector

Chapter 1 Introduction

1.1 Background

Assessing the effectiveness of existing systems is essential for making decisions relative to the allocation of resources needed to ensure and support the efforts of the system. Greater demands are being placed on systems to respond to user expectations in a cost-efficient manner. Even greater demands are placed on the information produced and used in making decisions which often have long-term consequences. This is especially true of systems whose purpose is detection. This is truer still when detection is part of a Protection/Security system.

In the current era of world-wide terrorism, effectiveness of Protection/Security systems takes on increased importance. Since September 11, 2001, the American public has become acutely aware and concerned about security from terrorist acts; and they expect the government to provide that security. Today's Protection/Security systems are expected to perform effectively against all types of threats (e.g. materials, chemicals, physical forces, electronic attack).

A looming threat is the use of nuclear materials to cause massive destruction; and this calls for measures to safeguard the public's well-being. The Department of Energy (DOE) defines domestic safeguards, which are applied to nuclear material as; "an integrated system of physical protection, material accounting, and material control measures designed to deter, prevent, detect, and respond to unauthorized possession, use, or sabotage of nuclear materials." The public expects this system to be effective.

System Effectiveness

Assessing the success and effectiveness of today's complex systems is an increasingly challenging problem. *"Demands for increased performance, lower system life cycle costs, longer operating capacities and improved productivity and efficiency must be balanced against limited resources, scant and sometimes unknown data, the identification and resolution of conflicts and problems, and resource allocation."* [1] That was true twenty years ago when Habayeb made that statement, and it is still true today. Thus, what is needed is a systematic approach for identifying problem areas and assessing system effectiveness.

In the current environment of national security, a system's level of effectiveness takes on increased importance. Given the recent incidents of terrorism abroad, international security needs highlight the importance in assessing the effectiveness of systems that are required by both national and international bodies. The United Nations (UN) Security Council Resolution (UNSCR) 1540 states that *inter alia* States (Countries) shall: *"take and enforce effective measures to establish domestic controls to prevent the proliferation of nuclear, chemical, or biological weapons and their means of delivery,*

including by establishing appropriate controls over related materials and to this end shall: (a) Develop and maintain appropriate effective measures to account for and secure such items in production, use, storage or transport; and (b) Develop and maintain appropriate effective physical protection measures;” [2] A central theme of this resolution is “effective measures”. Having measures is necessary; but they are not sufficient for providing security unless they are effective. That sufficiency creates confidence in the international safeguards concerning nuclear activities and promotes world peace. Without confidence, the commitment to non-proliferation may be weakened by fear and mistrust.

The burden of designing, building, and maintaining systems with effective measures is placed on the member States of the UN; however, there is no mutually agreed upon list of what these measures should be or how their effectiveness should be determined. Systems engineers, then, are needed to assist in such efforts. Systems engineers will be charged with evaluating the effectiveness of these, often complex and integrated, systems and to begin that activity, they need a definition for effectiveness.

Lack of Overall System Effectiveness Methodology

Given the level of concern about having effective security systems in this post- 9/11 environment and the usefulness of information they provide in decision making, it is surprising to find a lack of cohesive and accepted methodologies used to address their overall system effectiveness in the open literature. Soban and Mavris detail the lack of a System Effectiveness Methodology and much of the following section closely corresponds to their analysis [3]. There are a multitude of systems modeling tools, and the creation, use, and improvement of these tools is a flourishing endeavor [4]. In addition, many decision makers and analysts use these tools in their own individual way. However, finding information specifically detailing overall *methodologies* is difficult. There are several possible reasons for this lack of obvious resources. These reasons, which correspond to those described by Soban and Mavris [3] are detailed below.

Methodology or Framework

Merriam-Webster defines a methodology as: “a body of methods, rules, and postulates employed by a discipline; or: a particular procedure or set of procedures” and framework as “a basic conceptional structure (as of ideas).” [5] From these definitions, a framework is more conceptual and a methodology is more defined and has more rigor. The Microsoft Corporation in discussing the “Best Practices For IT Solutions Success”, describes a framework as being similar to a compass that can give guidance along a journey and a methodology as a specific set of directions to a specific location. [6]

As part of the Global Text Project (<http://globaltext.org/>), the information systems (IS) community has looked at IS methodologies and notes that a methodology is usually defined as “a holistic approach to the problem-solving process and the word ‘method’ is a subset of a methodology. Methodologies in general are viewed as problem-solving

processes, to solve mainly ill-structured problems. Problems where the solution and the outcome is not easily predictable i.e. having high uncertainty. Methodologies provide users with the means to reflect on reality, determine the problems and solving them. To sum up, methodologies provide users with ways of thinking, doing and learning. Methodologies provide users with ways to approach the problem and controlling the process of solution development. They provide an appreciation of standards and they enforce a more disciplined and consistent approach to planning, developing and managing.” [7]

Difficulty Assessing Government and Classified Material

Originally, system effectiveness studies were confined to military and space systems. Agencies of the US Government such as the Department of Defense (DOD) or the National Aeronautics and Space Administration (NASA) were the ultimate customers. With these agencies as customers, the available literature on system effectiveness and the accompanying models were published primarily as technical reports, but rarely appear in widely published journals. [8] Today’s analysts appear to have new interest in system effectiveness studies using campaign modeling, especially in the area of technology infusions. However, much of this work is classified or proprietary, limiting accessible publications and information. Finally, those non-government agencies that do make advances in theater modeling and system effectiveness may find it necessary to keep their in-house methods proprietary in order to retain their competitive edge. Because of these restrictions, some fundamental contributions to this field are not available to appear in this body of research.

System of Systems or Family of Systems Approach

In order to successfully formulate a system effectiveness methodology, it is imperative to clearly define the system and its components. Given that most systems are part of larger systems (making them subsystems), the question becomes the relationship between those subsystems. This represents an expanding progression of what is considered the system. Two fundamental types of aggregate systems are: a System of Systems; and a Family of Systems. The definition of these types of aggregate systems does not have consistency across the industry. Of particular importance in the analysis of the system is independence and interoperability.

The Office of the Under Secretary of Defense for Acquisition, Technology and Logistics (DoD-ATL) defines a system of systems as:

“A key principal to the understanding of a “system-of-systems” is the notion that a system performs a function not possible with any of the individual parts acting alone. Thus, a system can be viewed as any organized assembly of resources and procedures united and regulated by interaction or interdependence to accomplish a set of specific functions. In this context, a “system-of-systems” can be viewed as a set or arrangement of interdependent systems that are related or connected to provide a given capability. The loss of any part of the system will degrade the performance or capabilities of the whole.”[9]

This is in contrast to their definition of a family of systems, which is:

“a set or arrangement of independent (not interdependent) systems that can be arranged or interconnected in various ways to provide different capabilities. The mix of systems can be tailored to provide desired capabilities dependent on the situation. Thus, interoperability of the independent platforms is a key consideration in the ad hoc deployment of a ‘family-of-systems’”[9]

The above definitions are in contrast to others who have indicated a system of systems as being independent but interoperable. [10] However, the key point in this research is the assumption that the subsystems of the example system are independent and they are interoperable. Using the DoD-ATL definition the assumption is that a safeguards system is a Family of Systems.

Quantification

The key word in the definition of system effectiveness is “quantification”. In order for the decision maker or designer to evaluate the system, the results of the analysis must be presented as quantifiable metrics. This involves restating a research goal or design decision into a question that can be answered quantitatively. Dixon [11] states this explicitly: *“An engineering analyst must begin by defining quantitatively answerable questions”*. Mathematical methods, thus, become primary tools in system analysis because of their ability to rapidly provide these calculable (quantifiable) metrics. In addition, mathematical modeling allows the user to understand and make informed decisions at various levels within the system hierarchy. With the “system of systems” or “family of systems” concepts comes an appreciation of the potential complexities and interactions involved. Mathematical modeling offers significant benefits: “There are many interrelated elements that must be integrated as a system and not treated on an individual basis. The mathematical model makes it possible to deal with the problem as an entity and allows consideration of all major variables of the problems on a simultaneous basis”. The challenge is in determining what elements to model and how to model the relationships among them.

Thus, what is needed is a methodology that addresses the above – one that

- (a) Provides a clear definition of a system and system effectiveness that is broad across many applications;
- (b) Complements the probabilistic approach (DOE approach of P_E)
- (c) Adds to the open literature concerning systems effectiveness measurement; and
- (d) Recognizes the complexities of system structures; and

It is the mission of this research to do just that, and show how it applies to Safeguards Systems for nuclear materials. The framework of the work is that of a family of systems.

Safeguards Goals

The document that was developed to detail the safeguards obligations of states that are party to the Nuclear Non-Proliferation Treaty (NPT), INFCIRC/153, provides a technical definition of the safeguards objective, namely “the timely detection of the diversion of significant quantities of nuclear materials from peaceful activities...and deterrence of

such diversion by the risk of early detection." The key terms of this objective were not defined in INFCIRC/153; this task was given to the Standing Advisory Group on Safeguards Implementation (SAGSI) of the IAEA, an advisory group of technical safeguards experts. SAGSI considered the problem of quantifying the safeguards objective for several years. It identified four terms appearing either explicitly or implicitly in the statement of the objective just quoted as in need of quantitative expression. These were: significant quantities, timely detection, risk of detection, and the probability of raising a false alarm. It defined the associated numerical parameters (significant quantity, detection time, detection probability, and false alarm probability) as detection goals. Where:

A significant quantity (SQ) was defined as "the approximate quantity of nuclear material in respect of which, taking into account any conversion process involved, the possibility of manufacturing a nuclear explosive device cannot be excluded."

The detection time (the maximum time that should elapse between a diversion and its detection) should be of the same order of magnitude as conversion time, defined as the time required to convert different forms of nuclear material to the components of a nuclear explosive device.

On the basis of common statistical practice, SAGSI recommended a detection probability of 90-95%, and a false-alarm probability of less than 5%.

It seems rational that the detection goals should be operational criteria for safeguards effectiveness. In particular, a diversion of less than a significant quantity would not provide enough material for a nuclear explosive, and with regard to timely warning, it would obviously be advantageous to know about a diversion in time to do something about it, that is, before the diverter could assemble a weapon from the diverted material. [12]

1.2 Significance of Safeguard Systems

The Nuclear Safeguards community in the world is divided into two communities or communities of practice (CoP). Those CoPs are focused either on domestic safeguards or international safeguards.

The domestic safeguards CoP can be represented by the United States Department of Energy (DOE) which defines safeguards in numerous DOE orders [13] as; *"an integrated system of physical protection, material accounting, and material control measures designed to deter, prevent, detect, and respond to unauthorized possession, use, or sabotage of nuclear materials."* The Nuclear Regulatory Commission (NRC) of the United States includes two major areas (Physical Protection and Material Control & Accounting) in its description of domestic safeguards [14]. Therefore, the NRC definition is consistent with the DOE definition. Typically organizations involved with domestic safeguards are legally responsible for the control and accounting of the nuclear material as well as the physical protection of the material.

The international safeguards CoP can be represented by the International Atomic Energy Agency (IAEA) which is a subsidiary organization to the UN. The IAEA defines Safeguards as *“a set of activities by which the IAEA seeks to verify that a State is living up to its international undertakings not to use nuclear programs for nuclear weapons purposes. The safeguards system is based on assessment of the correctness and completeness of the State’s declarations to the IAEA concerning nuclear material and nuclear-related activities.”* [15] It should be noted that the physical protection component is absent from this definition, so at its heart, international safeguards consist of verification of appropriate accounting and control for nuclear material.

The IAEA guidelines for effective safeguards are that the IAEA can detect the diversion of a significant quantity (e.g. 8 kg plutonium of 25 kg of highly enriched uranium), with a 90% probability, within a specified 'conversion time' (related to the time required to convert different forms to metal) and with a false-alarm rate of no more than 5%. [16]

The international safeguards community verifies a State’s declarations of nuclear material inventories. Under these circumstances, effectiveness has a different emphasis and is not further developed in this research; however, the domestic Safeguards Systems effectiveness techniques developed would be applicable (with appropriate subsystems and components to the subsystems) to international use.

The modeling of domestic safeguards is assumed the most useful to the system manager and, as such the safeguards definition chosen for this research is for the domestic safeguards system, and the unmodified term “safeguards” hereon refers only to domestic safeguards.

In the specific case of nuclear safeguards, the system studied is one designed to safeguard nuclear material stored in a secure location (hereon referred to simply as a safeguards system).

Figure 1 is an image of the table listing the nuclear materials identified by DOE [17]. Similarly the identification of nuclear materials as special nuclear material (SNM) as defined by NRC is: *(1) plutonium, uranium 233, uranium enriched in the isotope 233 or in the isotope 235, and any other material which the Commission, pursuant to the provisions of section 51 of the act, determines to be special nuclear material, but does not include source material; or (2) any material artificially enriched by any of the foregoing but does not include source material.* [14]

Physical protection is required of SNM at all times, but the level of physical protection measures (such as guards, gates, fences, etc.) required is determined using a graded approach that is dependent on the category (I, II, III, IV) and attractiveness (A, B, C, D, E), in this case as determined by DOE in Figure 2 [17].

Table I-1. Nuclear Materials.

Material Type	SNM, Source, or Other	Reportable Quantity*	Weight Field Used for Element	Weight Field Used for Isotope	Material Type Code
Depleted Uranium (U)	source	kilogram	total U	U-235	10
Enriched Uranium ¹	SNM	gram	total U	U-235	20
Normal Uranium	source	kilogram	total U	--	81
Uranium-233	SNM	gram	total U	U-233	70
Plutonium-242 ² (Pu)	SNM	gram	total Pu	Pu-242	40
Plutonium-239-241	SNM	gram	total Pu	Pu-239 + Pu-241	50
Plutonium-238 ³	SNM	tenth of a gram	total Pu	Pu-238	83
Americium-241 ⁴ (Am)	other	gram	total Am	Am-241	44
Americium-243 ⁴	other	gram	total Am	Am-243	45
Berkelium (Bk)	other	microgram	--	Bk-249	47
Californium-252 (Cf)	other	microgram	--	Cf-252	48
Curium (Cm)	other	gram	total Cm	Cm-246	46
Deuterium ⁵ (D)	other	tenth of a kilogram	D ₂ O	D ₂	86
Lithium-6 (Li)	other	kilogram	total Li	Li-6	60
Neptunium-237 (Np)	other	gram	total Np	--	82
Thorium (Th)	source	kilogram	total Th	--	88
Tritium ⁶ (H-3)	other	hundredth of a gram	total H-3	--	87

* Materials are reported to the nearest whole unit, except for Pu-238, deuterium, and tritium.

¹ Uranium in cascades is treated as enriched uranium. Uranium in cascades should be reported as material type 89.

² Report as Pu-242 if the contained Pu-242 is 20 percent or greater of total plutonium by weight; otherwise, report as Pu-239-241.

³ Report as Pu-238 if the contained Pu-238 is 10 percent or greater of the total by weight plutonium; otherwise, report as Pu-239-241.

⁴ Americium only reportable when separated from plutonium

⁵ For deuterium in the form of heavy water, both the element and isotope weight fields will be used; otherwise, report isotope weight only.

⁶ Tritium contained in water (H₂O or D₂O) used as a moderator in a nuclear reactor is not an accountable material.

Figure 1 Image of DOE Nuclear Materials Table

11-22-00
DOE M 474.1-1A

Table I-4. Graded Safeguards.

	Attractiveness Level	Pu/U-233 Category (quantities in kgs)				Contained U-235 Category (quantities in kgs)				All E Materials Category IV
		I	II	III	IV ¹	I	II	III	IV ¹	
WEAPONS Assembled weapons and test devices	A	All	N/A	N/A	N/A	All	N/A	N/A	N/A	
PURE PRODUCTS Pits, major components, button ingots, recastable metal, directly convertible materials	B	≥2	≥0.4<2	≥0.2<0.4	<0.2	≥5	≥1<5	≥0.4<1	<0.4	
HIGH-GRADE MATERIALS Carbides, oxides, solutions (≥25g/L) nitrates, etc., fuel elements and assemblies, alloys and mixtures, UF ₆ or UF ₄ (≥50% enriched)	C	≥6	≥2<6	≥0.4<2	<0.4	≥20	≥6<20	≥2<6	<2	
LOW-GRADE MATERIALS Solutions (1 to 25g/L), process residues requiring extensive reprocessing, moderately irradiated material, Pu-238 (except waste), UF ₆ or UF ₄ (≥20% < 50% enriched)	D	N/A	≥16	≥3<16	<3	N/A	≥50	≥8<50	<8	
ALL OTHER MATERIALS Highly irradiated forms, solutions (<1 g/L), uranium containing < 20% U-235 (any form, any quantity)	E									Reportable Quantities

¹ The lower limit for Category IV is equal to reportable quantities in this Manual.

I-7

Figure 2 Image of DOE Graded Safeguards Table

Areas with different levels of access are illustrated by an example facility and can be seen in Figure 3. In this case there are four levels of access which are consistent with DOE terminology. The lowest level access in this example is the property protection area (PPA) and access is gained through pedestrian portal 1. The next level of access in this example is the limited area (LA) and access is gained through pedestrian portal 2. Both of these pedestrian portals are equipped with keypad entry devices which can automatically check an individual's credentials and allow access. Another level of access can be a protected area (PA), with access in this case through guard station 3, where a guard can verify access credentials. And finally there is a material access area (MAA) with access through guard station 4. Each level of access has a discrete access list allowing only those with a recognized need to enter. The specific security measures for each of these types of areas is determined by the attractiveness and category of the material in the MAA with category 1 material having the most stringent security measures. The access control system is a vital system for safeguards effectiveness.

Regarding accounting, a facility has an accounting system for tracking SNM inventories; documenting SNM transactions; issuing periodic reports; and assisting with the detection of unauthorized system access, data falsification, and material gains or losses. The accounting system provides a complete audit trail for all SNM from receipt through disposition.

Thus, the effectiveness measure of safeguards subsystems is vital information for both domestic and international communities to ensure global safety.

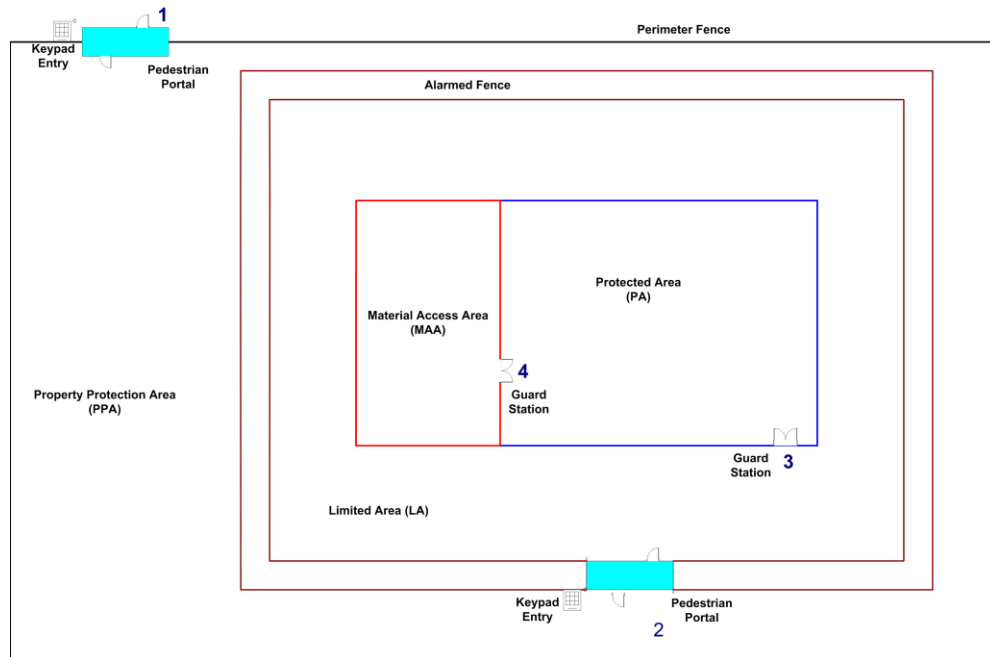


Figure 3 Example of Graded Access

1.4 Research Questions and Hypotheses

In summary, there exists a need in the Safeguards CoP for an integrated and systematic methodology that can assess an operating safeguards system's effectiveness. This methodology will be used to aid the decision maker in resource allocation, trade studies between system components, and requirements definition. The inherent presence of uncertainty in such a process has in the past necessitated the use of a detailed probabilistic analysis environment, facilitated through the use of mathematical modeling. A heuristic methodology that takes established subsystem performance rates and combines them in a way to establish a single system effectiveness measure would be a useful advancement in systems analysis. The heuristic measures of effectiveness proposed for this research are Type I (false positive) and Type II (false negative) errors. These needs give rise to specific research questions A and B and Hypotheses 1 and 2:

- A. Can a methodology be formulated and used to analyze the effect of incorporating measures of effectiveness at the component-subsystem level, yet assess the overall effectiveness at the system level?

Hypothesis 1: *The Multi-Attribute Utility Theory (MAUT) methodology is a valid approach to incorporate safeguards measures of effectiveness at the component-subsystem level and assess them at the safeguards system level.*

- B. Can uncertainties, such as incomplete data, rapidly changing technologies, and uncertainties in the modeling environment be addressed through heuristics in a system effectiveness methodology?

Hypothesis 2: Calculating safeguards system effectiveness using Type I and Type II error rates is a valid modeling approach upon which to base an effectiveness methodology with incomplete data, rapidly changing technologies, and uncertainties.

1.4 Summary and Introduction of Remaining Chapters

Measuring a system's effectiveness to date has not been a straight forward process. Up to the present time, one has to develop the framework in which the evaluation will be conducted and determine what type of measures of effectiveness will be used with no agreed or common performance-based methodology for how to determine an overall system effectiveness. Domestic nuclear safeguards systems are particularly important systems which need continuous evaluation of performance. A methodology to determine a system's effectiveness in near real time, with modern computing power, will give facility and systems managers a greater capability to mitigate the risks associated with threats to materials and essential services.

The remainder of this research addresses the concerns previously stated and follows a structured approach in addressing them. Structurally, the relevant literature with respect to system effectiveness and safeguards systems in particular are presented in Chapter 2. Chapter 3 contains the research methodology: 1) development of the function or calculating system effectiveness; 2) an explanation of the web-based survey design to find the utility coefficients; and 3) the construction of the simulations used to validate the effectiveness calculation. The research results are presented and analyzed in Chapter 4 and conclusions and recommendations are presented in Chapter 5.

Chapter 2 Literature Review

This chapter presents a comprehensive review of the literature relevant to the research undertaken and discusses the research questions and hypotheses. It includes basic system theory, approaches to effectiveness and applicable effectiveness theory, prior work related to effectiveness, the specific safeguards application issues including a discussion of risk, a discussion on utility theory, and finally safeguards facility scenarios.

2.1 Systems Theory

There are many possible definitions of a system; however, there is general agreement across fields and disciplines as to what constitutes a system. The following definition is representative of this agreement:

*A **system** may be considered as constituting a nucleus of elements combined in such a manner as to accomplish a function in response to an identified need...A system must have a **functional** purpose, may include a mix of products and processes, and may be contained within some form of hierarchy [18]*

Currently the International Council on Systems Engineering (INCOSE) defines a **system** as: *“a combination of interacting elements organized to achieve one or more stated purposes.”*[19] This research uses the INCOSE definition as a working definition.

There is also much discussion in the literature concerning what is meant by **effectiveness**. The MIT Engineering Systems Division defines **effectiveness** as the ratio of functions achieved to the totality of functions desired and **efficiency** as the ratio of functions achieved to the resources used. [20] Using this definition the common term “**cost effectiveness**” is shown to be an **efficiency**.

The definitions of **system effectiveness** vary widely and are often application dependent. Some examples that illustrate the diversity of **system effectiveness** definitions include:

“The probability that the system can successfully meet an operational demand within a given time when operated under specified conditions” [21]

“A measure of the degree to which an item can be expected to achieve a set of specific mission requirements, and which may be expressed as a function of availability, dependability and capability” [22]

Tillman, et. al. in an annotated bibliography on system effectiveness models in 1980 concluded “A wide range of definitions, and measures of system effectiveness are used without strong guiding logic.” [8].

NASA defines effectiveness as “a quantitative measure of the degree to which the system’s purpose is achieved”. [23]

Noel Sproles [24] says that effectiveness is the answer to the question of ***“Does this meet my need?”*** He then defined Measures of Effectiveness (MoE) as ***“standards against which the capability of a solution to meet the needs of a problem may be judged. The standards are specific properties that any potential solution must exhibit to some extent. MoEs are independent of any solution and do not specify performance or criteria”***. [25]

Finally, **system effectiveness** holds different meanings for different communities and applications. Some organizations tailor their definitions and methods to apply to very specific problems. In addition, **system effectiveness** is often associated with other concepts, such as “operations research”, “industrial engineering”, and “systems analysis”. However, even these have specific approaches and definitions; hence, finding “the” definition of system effectiveness is difficult. A new, consistent definition for system effectiveness, therefore, is necessary and must be justified by identifying key elements crucial to a useful definition. First, the term “effectiveness” implies that some sort of *quantification* needs to occur. This quantification must necessarily be the result of systematic *analysis* of variables and metrics that represent the system performing its function. In addition, in order to perform the quantification, an *intended or expected effect* needs to be identified in order to properly model the results of the system performance. Combined, these concepts result in the following definition for use in formulating the assessment of system effectiveness:

System effectiveness is a quantified measure of achievement, through functional analysis , of the expected system results.[3]

The above will be the working definition of system effectiveness used for the remainder of this dissertation and subsequent research.

Another useful definition is: “A system is a purposeful collection of interrelated components that work together to achieve some objective” [26], which is an earlier variant of the INCOSE definition of “a combination of interacting elements organized to achieve one or more stated purposes.” [19]. These definitions are basic to systems theory. A generalized representation of a system is shown in Figure 4 and as shown a system exists to respond to a demand placed on it by an entity in its environment (the input). It is the response that entity which usually decides the degree to which its demand was met; that is, whether the system was “effective” in meeting its intended purpose. Earlier investigations [25] [27] [28] revealed various issues which are still to

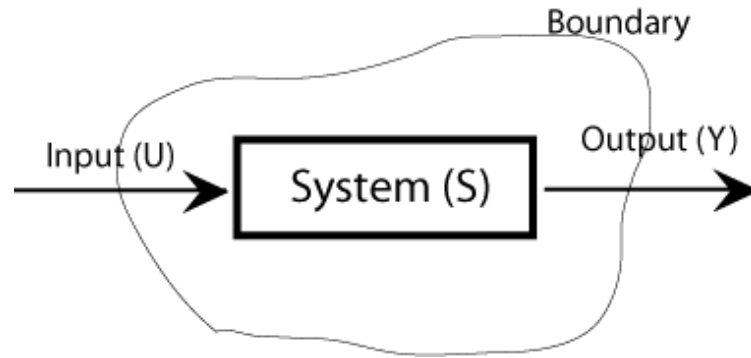


Figure 4 General System Structure

be resolved when attempting to define and measure system effectiveness. Some of them are:

- dealing with a system as part of a system (or family) of systems;
- the need to combine both quantitative and qualitative measures; and
- the relationship between performance measures and effectiveness measures and how to aggregate them for the entire system.

According to Smith and Clark, [28] many approaches for mapping performance measures to effectiveness measures are inadequate, especially as systems become more networked and complex in behavior. To overcome the shortcomings of traditional approaches to measuring effectiveness, it is critical to measure system effectiveness relative to the system domain and purpose and to provide a comprehensive value. The purpose of this research is to respond to that need.

Systems theory grew as a reaction to the generally accepted practice in the early part of the 20th century of analyzing systems only as the sum of their parts. General systems theory evolved by considering the characteristics of whole systems and their overall outcomes. [29] In every case a specific system is described specifically and has discrete boundaries and interactions with the system's environment. A system can be broken into sub-systems each with a discrete function (e.g. a radio that has a receiver subsystem, a signal processing subsystem and a speaker subsystem) or aggregated into super systems (a system of systems or a family of systems) to produce a combined effort to supply a solution to a more complex need (e.g. a transportation system: that includes a baggage handling system, a passenger handling system, a food delivery system, and an overall system management system). Systems allow for actions caused by actors that may come from outside the system (or from within the system) but can have interactions with the system. The actors are users of the system and often called entities. The level of interactions among the internal and external entities impacts the complexity of the system. According to Rouse [30], the complexity is a function of the purpose, or intention, of the system. In Table 1 a categorization of some possible system intentions is shown.

Table 1 System Intentions

Intention	Example
Classification	"It's an instance of type S."
Explanation	"It's type S because"
Prediction	"It's future output will be Y."
Control	"If input is U, it's output will be Y."
Detection	"It's output is not Y, but should be."
Diagnosis	"It's output is not Y because ..."

This research will look at systems whose intention is to detect anomalies. Such systems specifically, safeguard systems are addressed; and a case study of a related security system (airport security system) is presented.

According to Smith and Clark [28], measures of effectiveness (MoEs) should have the following properties:

1. The measure needs to **increase as effectiveness increases**;
2. The measure needs to be **bounded above by an ideal system and bounded below by zero for non-compliance**;
3. To manage **complexity and allow for system decomposition**, any measure needs to represent and support system decomposition and aggregation (for equivalent systems aggregate measures must be equivalent regardless of level of decomposition).
4. To facilitate comparisons between systems (which may have different internal characteristics and differing primary purposes) it is necessary to normalize the final effectiveness scores. The range [20] is chosen (with 0 denoting an ineffective system and 1 denoting a perfectly effective "ideal" system)
5. Ideally the measures should be **ratio scales** which means that they have a natural zero point and numbers which are multiples of each other directly indicate their value. (For example, a system with an effectiveness measure of 0.8 is twice as effectiveness as a system with a measure of 0.4). It should be noted that ratio scales directly support the achievement of properties 1 to 4.

All of these properties impact the choice of approaches to define MoEs. Two approaches from Decision Science meet these mathematical requirements and are considered as candidates for measuring effectiveness:

1. Multi-Attribute Utility Theory (MAUT) [31] and.
2. The probabilistic modeling technique of Bayesian Networking (BN), [32] [33] [34].

Both these approaches deal with measures between 0 and 1 with MAUT using a utility coefficient and BN using probability.

Expected Utility expresses the worth of a consequence. Expected utility is the probabilistically weighted sum of the utilities. In Decision Science under uncertainty, it is considered “rational” to choose between alternatives (a' and a'') based on the value of expected utility. MAUT provides a basis for measuring expected utility but the formulation of the utility function is dependent on many complex independence conditions being established. These independence conditions are difficult to establish and verify and consequently MAUT is often considered unsuitable for measuring effectiveness in highly complex domains. In the case of this research, the decomposed system elements will be assumed independent and the MoEs will be measureable. Thus, the alternative often used is to provide a mechanism that can aggregate measures that account for relationships. Such an approach is Bayesian Inference (commonly called Bayesian Networks, BN). In common with MAUT, BN provides a well-grounded, consistent mathematical framework which (in addition) supports the forward and backwards propagation of evidence. [33] So it is able to answer the questions:

If I observe something:

1. What may have caused this?
2. What outcomes will this influence?

Bayes law states that the probability of **a** given **b** equals the probability of **b** times the probability of **b** given **a** divided by the probability of **a**, which is stated notationally by the following:

$$Pr(a|b) = Pr(b) \times Pr(b|a) / Pr(a).$$

The notation **$Pr(a|b)$** means the probability of **a** given **b**.

And by simple rearrangement: $Pr(b|a) = Pr(a) \times Pr(a|b) / Pr(b)$. This justifies forward and backward propagation of evidence.

So within an effectiveness context, this rule can be used to answer the question “what is the effectiveness of a system given the effectiveness of another system?” This feature is important because many complex systems function as members of larger systems and are influenced by other member systems. Therefore, a complex system may be viewed as a network. A BN is able to update the probabilities in uncertain nodes (using Bayes rule) given evidence obtained from related nodes. This property and the intuitive way BN model complex relationships among nodes make them a suitable technique for building causative models [33], [34]. There is evidence [35] to suggest that the predictive value of BN is robust against incorrect estimates of the probability values populating the nodes as long as the causative links are correct and have appropriate weighting.

For military systems, and particularly networked systems based on new technology, there is often insufficient data (or operational experience) to quantify a system's effectiveness. So the recourse is often to guesstimate using expert judgment. In addition, to classical statements of effectiveness, such subjective (qualitative) judgment needs to be handled. Cox's work [36] [37] is accepted as the justification for the use of subjective probability within a Bayesian framework [34]. Cox derived Bayes' rule (and other probabilistic rules) from the rules of logic and two axioms without reference to the frequentist definition of probability. He thus claimed that probability theory, in essence, has involved two ideas: "the idea of frequency in an ensemble and the idea of reasonable expectation". Reasonable expectation is the probability of an event which is not based on extensive trials but more on subjective judgment and expert opinion. It provides a "measure of the reasonable expectation of an event in a single trial". These axioms [37] are:

"the probability of an inference on given evidence determines the probability of its contradictory on the same evidence"; and

"the probability on given evidence that both of two inferences are true is determined by their separate probabilities, one on the given evidence, and the other on this evidence with the additional assumption that the first inference is true". that subjective probability is equally valid for modeling causal relationships under uncertainty [34].

Both of these "axioms" can be considered to be valid in the domain of measuring effectiveness; that is, effectiveness and its inverse are related and the effectiveness of two systems combined is dependent on the effectiveness of the first system, if the systems are causally related.

So by measuring effectiveness and using values between zero and one, it is possible to aggregate their effects as long as their causative relations can be established. Thus, it is possible to build a BN to model effectiveness in such a way that total system effectiveness can be inferred from subsystem effectiveness [33]. This assessment of effectiveness can be performed in both a "forward" and "reverse" direction; that is, given subsystem effectiveness total system effectiveness can be determined or if a system is effective, measures of required subsystem effectiveness can be inferred. The same inference that total system effectiveness can be determined given subsystem effectiveness follows for MAUT under the family of systems approach (independence conditions) and true MoEs (according to Smith and Clark's criteria).

Measures of Effectiveness

If performance measures can be causatively mapped to an effectiveness measure then an approach can be also used to calculate this measure but this mapping should be done within a subsidiary model. Based on Sproles' distinction between Measures of Performance (MoP) and MoE, a MoP measures the internal characteristics of a solution while a MoE measures external parameters that are independent of the solution and are measurements of how well the problem has been solved. According to Sproles a MoE

is: “A measure of the ability of a system to meet its specified needs (or requirements) from a particular viewpoint(s). This measure may be quantitative or qualitative and it allows comparable systems to be ranked. These effectiveness measures are defined in the problem-space. Implicit in the meeting of problem requirements is that threshold values must be exceeded”. [25]

A Security Risk is the probability of sustaining a loss of a specific magnitude during a specific time period due to a failure of security systems. The intended MoEs are the primary variables for the determination of any given system’s effectiveness under the proposed heuristic approach and will be the error rates for the different subsystem components (rates of Type I (α) and Type II (β) error). These will be used as the fundamental input variables for the proposed model. Every functioning system should have an established testing procedure and records on “false positive or false alarms” (Type I errors) and “false negative or missed alarms” (Type II errors). There are such things as nuisance alarms (such as a small animal tripping a motion sensor) which depending on context and focus can be treated as either false alarms (Type I error) or true alarms (no error). These error rates have the distinct advantage of being directly measurable in many cases and will facilitate the determination of system effectiveness for specific systems that can be tracked over time. The missed alarms are the inherently the most difficult rate to determine. Yu-Sung Wu et. al. developed an estimation procedure for intrusion containment in large scale distributed systems to reduce error propagation, thus helping the system remain functioning. [38]

2.2 Specific Hypothesis

Given the research questions A and B as previously stated we can answer them as follows:

- A. *Can a methodology be formulated and used to analyze the effect of incorporating measures of effectiveness concepts at the component-system level, yet assess them at the system level?*

Hypothesis 1: The Multi-attribute Utility Theory (MAUT) methodology is a valid approach to incorporate safeguards measures of effectiveness at the component-subsystem level and assess them at the safeguards system level.

As Druzzdel established, by measuring effectiveness and using values between zero and one, it is possible to aggregate their effects, that is, given subsystem effectiveness, total system effectiveness can be determined. The method under consideration in this research is MAUT whose fault is the inability to determine independence of subsystems in highly complex systems and an expected utility when experience with the system is low. By focusing on Families of Systems which are independent and interoperable, we can use the independence of subsystems allowing the MAUT algebra to be used and, further, the intended effectiveness evaluation is of a well defined and operating system that provides historical MoE

data on each subsystem allowing this question to be answered in the affirmative for MAUT. And thus answers Research question A. and affirms Hypothesis 1:

- B. Can uncertainties, such as incomplete data, rapidly changing technologies, and uncertainties in the modeling environment be addressed through heuristics in a system effectiveness methodology?

Hypothesis 2: Calculating safeguards system effectiveness using Type I and Type II error rates is a valid modeling approach upon which to base an effectiveness methodology with incomplete data, rapidly changing technologies, and uncertainties.

This research proposes a shift from a probabilistic approach that analyzes the likelihood of events (by whatever means) and determines effectiveness based on these probabilities to a heuristic approach which uses the experience of measured error rates on a real-time sample (Type I and Type II errors previously mention and seen in Table 2). Tversky and Kahneman describe three heuristics that are employed in making judgments under uncertainty:

- 1) representativeness, which is usually employed when people are asked to judge the probability that an object or event A belongs to class or process B;
- 2) availability of instances or scenarios, which is often employed when people are asked to assess the frequency of a class or the plausibility of a particular development; and
- 3) adjustment from an anchor, which is usually employed in numerical prediction when a relevant value is available.

They indicate that these heuristics are highly economical and usually effective, but they lead to systematic and predictable errors. [39] Dewhurst, et. al. indicate that research in cognitive psychology indicates that individuals at all educational levels use heuristics especially when making a decision within a limited amount of time. [40] Therefore, it seems reasonable to move toward heuristics when making decisions in real time and on limited data. In the case of personal observations, factors such as accuracy of one's memory and uniqueness of one's prior experiences, may affect the validity of answers arrived at through heuristics. [40]. In the case of using a measured rate of errors, as proposed in this research, the more data gathered over time (producing an increased sample size) the closer the heuristic methodology is to the probabilistic likelihood of events methodology. When in practice, the intended effectiveness evaluation described in this research, is of a well defined and operating system that provides historical MoE data on each subsystem allows this question to be answered in the affirmative and affirms Hypothesis 2.

To validate Hypotheses 1 and 2, the posited methodology and calculation approach will be compared to an existing probabilistic calculation methodology using theoretical values of system effectiveness.

Table 2 Type I and II Errors

Statistical Decision	True State of the Null Hypothesis (H_0)	
	H_0 True	H_0 False (H_1 True)
Accept H_0	Correct ($1-\alpha$) TPF	Type II error (β) FNF
Reject H_0	Type I error (α) FPF	Correct ($1-\beta$) TNF

The International Council on Systems Engineering (INCOSE) states that regarding model validity *“It is crucial to prove that the model is trustworthy and suitably represents reality, particularly in cases where a feel for system behavior is absent, or when serious consequences can result from inaccuracy. Models can be validated by: (1) Experience with application of similar models in similar circumstances; (2) Analysis showing that the elements of the model are of necessity correct and are correctly integrated; (3) Comparison with test cases in the form of independent models of proven validity or actual test data; and (4) The modeling schema itself can be validated by using small scale models.”*[19]]

2.3 Applicable Theory

Baker, et. al. show the interactions between Risk, Consequences, Threats and Vulnerabilities under a homeland security framework, which can be seen in Figure 5. This leads to a risk relationship seen as follows [41]:

$$R = C \times T \times V,$$

where:

R = Risk associated with an adversary attack and/or system/asset failure

C = Consequence(s), the negative outcomes associated with degradation or failure of the system or asset(s). Consequences of an attack can be measured by loss of life, economic impact, loss of public confidence, or other metrics

T = Threat, the probability or likelihood that a given attack scenario with the potential to disrupt systems or assets and cause undesirable consequences will occur. Threats are characterized by their means and likelihood of occurrence

V = Vulnerability, a weakness in the system or asset, or supporting systems or assets (e.g., security systems, etc.) to the threat (**T**) that would cause degradation or failure

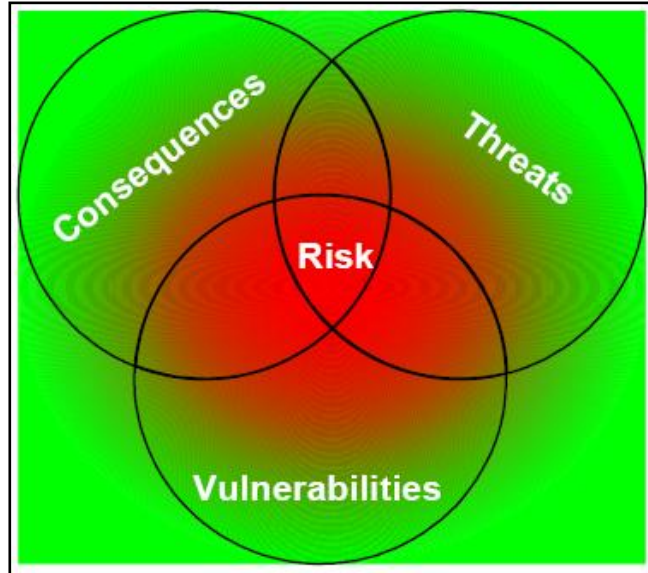


Figure 5 Risk, Consequences, Threats and Vulnerabilities

Baker, et. al. ultimately quantify risk with the following equation [41]:

$$R = P_A * (1 - P_E) * C$$

Where:

R = risk associated with adversary attack

P_A = likelihood of the attack

P_E = probability security system is effective against the attack

(1 – P_E) = probability that the adversary attack is successful (also the probability that security system is not effective against the attack)

C = consequence of the loss from the attack.

A key question becomes: What is the likelihood of the attack (**P_A**)? It is a general rule of thumb to assume the probability of attack =1 and this term drops out of the equation. The next Key Question becomes what is the probability security system is effective against the attack (**P_E**)? Estimates for **P_E** are taken from past experience (personal experience or testing) or industry data (which again supports the heuristic methodology). In determining the consequence value (**C**) any number of factors can be involved however, in this research the values assumed are:

0.0 = No impact/consequence

1.0 = Very significant impact

Thus the only consequence of importance to the system is significant which requires action and the risk equation reduces to:

$$R = 1 - P_E$$

The focus for the effectiveness of a security system becomes the probability security system is effective against the attack (P_E , or in a layered system the layers effectiveness or P_{EL}).

The DOE defines a **system performance effectiveness equation** for a layered system of protection in DOE M 470.4-1, Attachment 2, Part 1, Section E, Appendix 4, (Page 4-1) as follows [42] Shown here:

$$P_{EL} = P_{IL} \times P_{NL} = P_{DL} \times P_{NL} = P_{AL} \times P_{SL} \times P_{NL}$$

Where:

P_{EL} = the system effectiveness contribution for layer L;

P_{IL} = Probability of Interruption given first detection at layer L, $P_{IL} = P_{DL}$ if detection on layer L is timely, and is equal to 0 ($P_{IL} = 0$) if detection is not timely;

P_{DL} = Probability of Detection at layer L, $P_{DL} = P_{SL} \times P_{AL}$ on layer L. P_{DL} is the probability of first detection at layer L, given that detection has not occurred at an earlier layer, multiplied by the probability of sensing at an earlier layer, multiplied by the probability of sensing at layer L (P_{SL}) and the probability of assessment at layer L (P_{AL});

P_{SL} = Probability of Sensing on layer L;

P_{AL} = Probability of Assessment on layer L; and

P_{NL} = Probability of Neutralization given first detection at layer L.

L = the number of detection layers in the system before the critical detection point (CDP) in the adversary path(s). Detection after the CDP must be counted.

P_{EL} = the system effectiveness of the layer. The system effectiveness of the layer is the product of the probability of interruption of the layer and the probability of neutralization given that detection occurred at that layer ($P_{IL} \times P_{NL}$). The probability of neutralization is determined discretely for each layer given detection at the layer. The neutralization determination is made if detection (regardless of the extent) takes place at the layer in question. Neutralization will occur sometime past the detection point and would be valid for the probability of neutralization of that specific layer.

P_{DL} = the product of the probability of sensing and the probability of assessment of the layer ($P_{SL} \times P_{AL}$). Note that detection and assessment will be different between the

elements of the layer and between layers. P_{IL} of the layer is defined as $P_{IL} = P_{DL}$ if detection on layer L is timely, and is equal to 0 ($P_{IL} = 0$) if detection is not timely.

The contributions of each layer along the adversary pathway are then combined to determine the overall system effectiveness, where the overall system effectiveness is provided by the sum of the contributions of each layer (only those encountered along the adversary pathway) to the system effectiveness. An example of the system effectiveness equations for a three-layer system protecting SNM would be as follows:

$$P_E = P_{E1} + \{(1 - P_{I1}) \times P_{E2}\} + \{1 - (P_{I1} + \{(1 - P_{I1}) \times P_{I2}\}) \times P_{E3}\}$$

This equation uses a probabilistic focus for the effectiveness calculation and establishes the procedure for calculating effectiveness of subsequent layers. In the case of the example facility with graded access, the layers could correspond to the four layers indentified (PPA, LA, PA, and MAA) in Figure 3. The DOE **system performance effectiveness equation** is the most similar in scope to the methodology used in this research. Facility scenarios will be explored to compare the DOE method to the proposed heuristic method for a comparison of the techniques and indicate their complimentary nature. Further, in Chapter 4, a direct comparison (on a one layer system) of this probabilistic methodology to the posited heuristic utility methodology will be made in order to validate the research Hypothesis 1: **“The Multi-attribute Utility Theory (MAUT) methodology is a valid approach to incorporate safeguards measures of effectiveness at the component-subsystem level and assess them at the safeguards system level.”**

For this research the probability of neutralization is assumed to be very close to 1 reducing the effectiveness equation for a single layer to:

$$P_E = P_D$$

In many cases the performance of a **single layer** detection system can be characterized by two primary parameters: the detection probability, P_D , and the false alarm or false positive rate (α or **Type I error**). A good detector has a very high detection probability (as close to 100 percent as possible), while still maintaining a very low rate of α . These two parameters are coupled, primarily through the detection threshold (t) as seen in Figure 6. The more sensitive the detection threshold is, the higher the false alarm rate. As the threshold is moved to the left, i.e., the device is adjusted to detect lesser incidences of noise plus signal and the detection probability increases but a larger area of the signals is included in the uncertain population, representing a higher false alarm rate.

Any detection scheme depends on a separation of the two peaks: the probability distribution of the noise and that of the noise plus signal and as the signal to be detected decreases, the two curves move together (i.e., the distribution of signals will be shifted to the left), making discrimination more difficult.

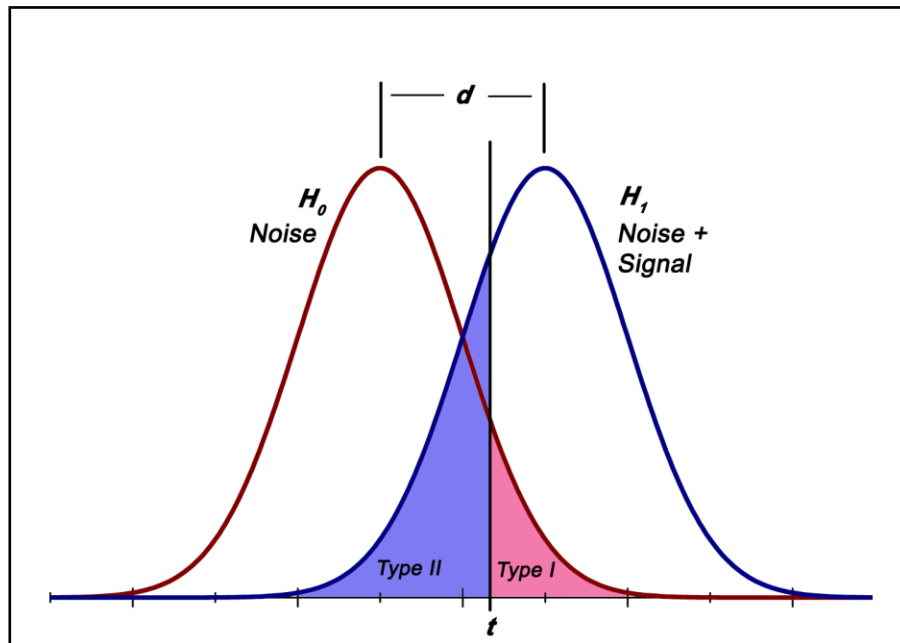


Figure 6 Hypothesis Testing Errors

The true combined detection probabilities and false alarm rates can only be determined by measurements in an operational environment.

Detection Criteria

What actually constitutes acceptable detection probability and false alarm rate is not easily determined. The former is a question of acceptable risk while the latter is an operational problem. Setting the minimum acceptable detection probability is a subjective issue. If there were 10 attempted bombings per year (out of 40 million international enplanements), a detection probability of 0.90 would allow on average one expected dangerous situation (and sometimes more) to go undetected per year. If there were only one bomb attempt per year, the statistical expectation would be for one to go undetected about every 10 years. Would a terrorist be deterred by these odds and would the flying public accept them as “safe”? The operational part of the problem can be analyzed reasonably objectively, yet it, too, is difficult to specify precisely.

When we consider the effectiveness of procedures for predicting the system effectiveness, we find that present methods for measuring the power of such a predictor procedure are less than satisfactory. Typical measures involve (a) developing a statistic that serves to predict the effectiveness, (b) selecting a cut-off value of that statistic, and (c) counting the number of correct classifications and misclassifications when this method is applied to a set of data prediction examples. We find this to be unsatisfactory

because it does not take into account the fact that misclassifications of the two different kinds may bear different costs for the users of the system. In this setting a tool, as described by Swets, used for evaluating the effectiveness of a prediction procedure is the “Receiver Operating Characteristic” (ROC) of the prediction method itself [43] [44]. The ROC is a curve that represents the fraction of all effective cases that would be classified correctly, plotted as function of the fraction of all cases in which the system is ineffective, but is falsely predicted to be effective. In other words, it measures the ability of the predictive scheme to correctly *detect* effective decisions, while keeping track of the degree to which the scheme generates “false alarms”.

The importance of the ROC lies in two facts: (1) it permits users of the procedure to select thresholds for classification that are appropriate to their own assessments of the relative cost of missing cases as opposed to false alarms in which they incorrectly predict effectiveness; and (2) The relative strength of two procedures for predicting the effectiveness can often be read easily from the ROC, since a prediction procedure whose ROC curve lies always above that of another procedure will be superior to it no matter what the user’s specific estimates of costs and values may be.

Safeguards systems are fundamentally alarm systems and the basic system response is to create an alarm when a condition is not as it should be. There are numerous situations that could be considered failures or errors under these circumstances; however, some background in errors will be useful. In particular, the issue of the type of error needs to be developed. Miller tracks the history of the terms **TYPE I ERROR** and **TYPE II ERROR** [45]. In their first joint paper Neyman and Pearson in 1928 referred to “the first source of error” and “the second source of error” [46]. Neyman and Pearson progress to use the term “*Errors of first and second kind*” in 1933 [47]. The first use of the terms *Type I error* and *Type II error* is then made by Neyman and Pearson in 1933 in their work “The Testing of Statistical Hypotheses in Relation to Probabilities A Priori” [48]. A table depicting those errors was seen in Table 1 and is now modified with respect to alarms and presented as Table 3.

Table 3 Type I and II Errors as Alarms

<i>Decision</i>	True State	
	No Alarm	Alarm
<i>No Alarm</i>	Correct ($1-\alpha$) TNF	Type II error Missed Alarm (β) FNF
<i>Alarm</i>	Type I error False Alarm (α) FPF	Correct ($1-\beta$) TPF

Neyman and Pearson formed the basis of statistical hypothesis testing, and Figure 6 shows a common graphic that illustrates the principles involved. The fundamental approach is to fix a probability of a Type I error arbitrarily (called the significance level usually 0.05 or 0.01) and then choose a criterion so that the probability of a Type II error is minimized. Neyman and Pearson showed that the best test was in terms of the likelihood ratio. Under their rule the alternative hypothesis (H_1) is accepted when the likelihood ratio is greater than t where t is the test threshold and is chosen to produce the desired significance level of a Type I error. Often, instead of considering the probability of a Type II error, the quantity 1 minus the probability of a Type II error (which is the probability of accepting H_1 when H_1 is true) or, the power of the test becomes the focus. Table 3 also includes another terminology used in diagnostic analysis (TPF - true positive fraction, FPF - false positive fraction, TNF - true negative fraction, FNF - false negative fraction)

Sage and Rouse formulated categories of errors in an assessment and response framework such that errors could be: (1) errors in the detection of a problem, (2) errors in diagnosis of a problem, and (3) errors in planning and executing actions [49]. This work only looks at the Type I and Type II error rates in the system and does not use the taxonomy that Sage and Rouse mention.

Villemeur describes Failure Modes and Effects Analysis (FMEA) in which FMEA is an inductive analysis used to systematically study causes and effects likely to affect the performance of a system. There are four main steps in performing an FMEA: “(1) *Definition of the system, its function and components*; (2) *Identification of the component failure modes and their cause*; (3) *Study of the failure mode effects*; and (4) *Conclusions and recommendations*” [50]

FMEA is focused on finding significant failure modes and making changes in the system that correct the failures. While many of the activities in an FMEA are useful and essential in determining how to keep a system running, the focus is not to determine a *quantified measure of achievement, through functional analysis, of the expected system results* (effectiveness) of the system. In a number of cases Type I and Type II errors may not be “failures” of the system. Failure being defined in FMEA as: “*The event in which any part of an item does not perform as required by its specification. The failure may occur at a value in excess of the minimum required in the specification, i.e., past design limits or beyond the margin of safety.*”[19] The system may have performed as designed and still created an error.

Figure 7 shows the proposed basis for the effectiveness calculation which uses the rates of achieving the desired outcome. The populations in Figure 6 and Figure 7 are assumed to be normally distributed with equal variance.

Therefore in the context of a radar system Type I errors were false alarms and Type II errors were deemed misses. ROC curves were developed in the 1950's when

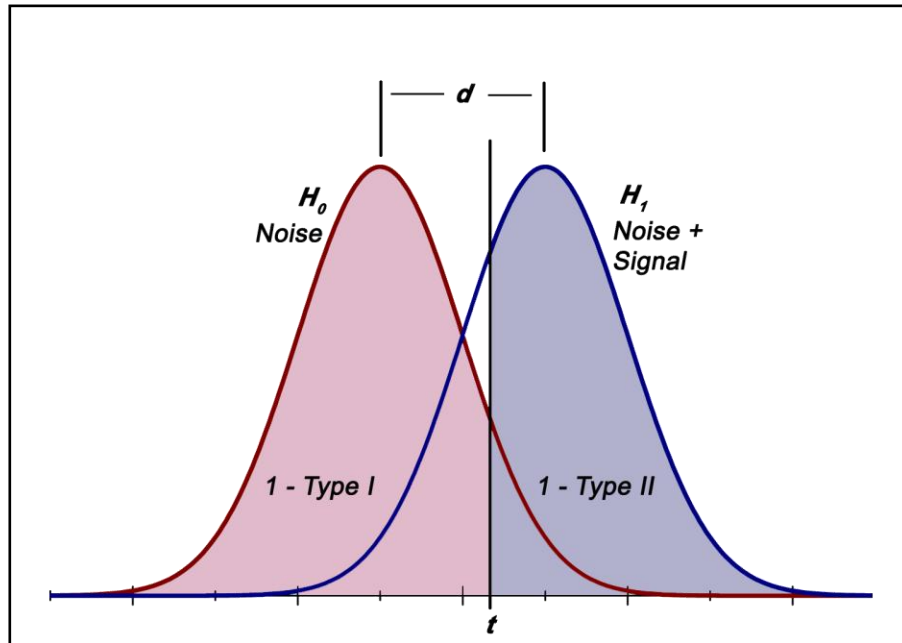


Figure 7 Proposed Basis of Effectiveness

attempting to make sense of radio signals contaminated by noise and today is commonly used in medical diagnostics. J. van Schalkwyk uses applets in an online internet resource that takes user input to create example ROCs [51]. Figure 8 and Figure 9 show the affect of shifting the test threshold (t) on Type I and Type II error as well as the resulting ROC curve. Figure 8 shows TPF, FPF (Type I error), FNF (Type II error), and TNF which in this context is related to a test for disease and a patient's true diagnosis. The ROC curve plots TPF ($1-\beta$) against the FPF (α).

In Figure 8 the ROC is seen on the right hand side (the curved portion). The graph plots the TPF vs. the FPF and the straight line shows the condition of a 50% chance of being either true positive or false positive. This is the case when the two normal distributions are not separated (have the same mean). In the case shown in Figure 9 there is a greater curve separation. As can be seen in Figure 9, the larger curve separation creates a more severe curvature of the ROC. The shape of the ROC can vary depending on the various combinations of distributions for the noise vs. signal. Figures 5 and 6 depict normal curves but ROCs can be constructed from any number of distributions leading to skewed curves in one direction or the other. ROCs are useful in evaluating the response of a system and have been in use for many years. This affirms the hypothesis: ***“Calculating safeguards system effectiveness using Type I and Type II error rates is a valid modeling approach upon which to base an effectiveness methodology with incomplete data, rapidly changing technologies, and uncertainties.”*** as ROCs evaluate operating systems using the TPF ($1-\beta$) and FPF (α). This demonstrates that it is valid to use the Type I and Type II statistical error rates

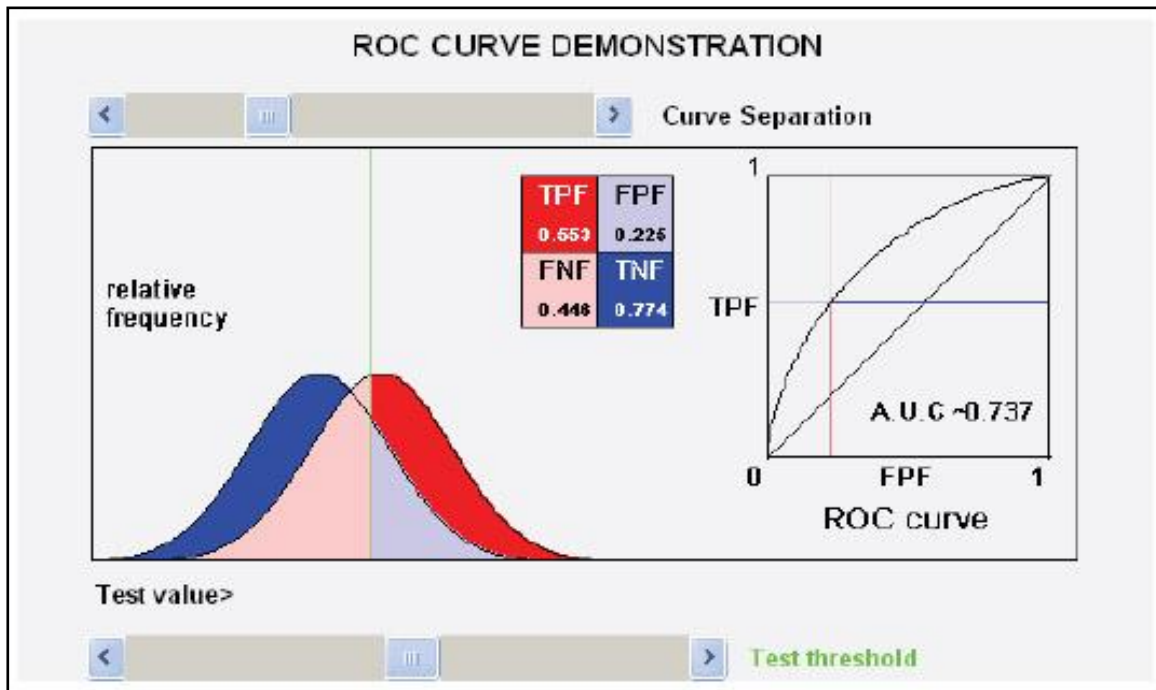


Figure 8 ROC Applet Output 1

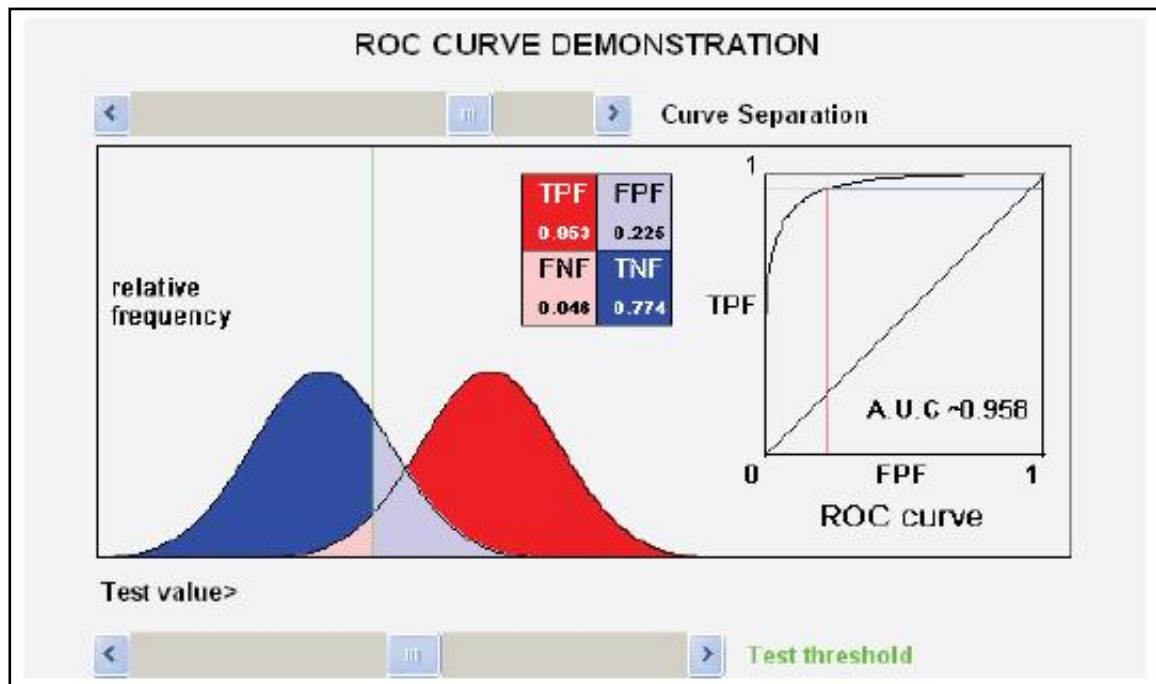


Figure 9 ROC Applet Output 2

to evaluate a system, ROCs do not directly calculate the effectiveness of a given system but yield a curve that allows the system manager to interpret effectiveness, however, with the modification proposed in this research, effectiveness can be directly calculated.

2.4 Approaches

There are two fundamental approaches to evaluating system effectiveness: a systems resource approach and a goal centered approach. [52] Under a **resource approach**, system effectiveness is determined in terms of resource availability instead of specific task objectives (in the MIT taxonomy this is an efficiency). Under a **goal centered approach**, system effectiveness is determined by comparing the systems performance against specific objectives (in the MIT taxonomy this is effectiveness). Under either grading approach, there will always be interest in improving the system.

This research is based on the goal centered approach to systems effectiveness evaluations and complies with the MIT definition of an effectiveness. This is particularly important in systems that protect assets (safeguard systems) because if the asset is not safeguarded, at some point in time the probability of losing the asset to a malicious actor becomes unacceptable.

2.5 Prior Work Related to Effectiveness

Hamilton and Chervany look at evaluating management information systems effectiveness and as such provide detail for the two frameworks for system effectiveness.[53] The goal centered view first determines the objectives of the system and then develops criterion measures to assess the system. Under this framework the system's effectiveness is determined by comparing system performance to system objectives. Under the system-resource framework effectiveness is conceptualized in terms of resources and their use (an efficiency). The systems-resource view recognizes that there are considerations external to accomplishment of system objectives which need to be weighed in the analysis of a given system. Hamilton and Chervany summarize problems in evaluating system effectiveness as: (1) Objectives and measures of accomplishments (MoEs in this context) are often inadequately defined initially; (2) Efficiency-oriented and easily quantified objectives and MoEs are typically employed; (3) Objectives and MoEs used to evaluate the system are not the same as those defined initially; and (4) Individual perceptions may differ on what the objectives and MoEs are; and further summarize that realistic mutual agreement, concerning the definition of appropriate objectives and measures of accomplishment, is typically not reached at the outset and this lack of agreement makes evaluation of system effectiveness difficult.

Soban and Mavris recognize the need for an integrated and efficient framework that can rapidly assess system effectiveness for today's complex systems. They recognize that the definitions of system effectiveness vary widely and are often application dependent.

They quote some of the more common definitions such as: (a) *“The overall capability of a system to accomplish its intended mission”* [8]; (b) *“The probability that the system can successfully meet an operational demand within a given time when operated under specified conditions”* [4]; (c) *“A measure of the degree to which an item can be expected to achieve a set of specific mission requirements, and which may be expressed as a function of availability, dependability and capability”* [54]; and (d) Soban and Mavris also present a new definition for system effectiveness as the term “effectiveness” implies an element of quantification. The following is the working definition of system effectiveness for this research. **“System effectiveness is a quantified measure of achievement, through functional analysis, of the expected system results”** [55]

Soban and Mavris’ purpose was to evaluate systems involved in air superiority in the theater environment, and they used a Response Surface Methodology which comprised two steps, facilitated by a common statistical evaluation program (JMP). The first step was effect screening, which creates a linear model used to determine the sensitivity of a response to various inputs and to screen out, using a Pareto analysis, those variables that did not contribute significantly to the variability of the response. The second step was surface fitting, yielding a polynomial representation that gives the response as a function of the most important input parameters. Soban and Mavris’ output table had a listing for each variable in their model and how it influenced their effectiveness function. The methodology is useful in the specific application and uses system level metrics but looked at individual response terms and was not used to create an overall effectiveness.

2.6 Area of Application

The area of specific interest in this research is that of domestic safeguards systems consisting of Physical Protection, Material Control, Material Accounting, and Management subsystems.

Bennett, et. al. look at a general methodology for comparative evaluation of physical protection system effectiveness. The Bennett (et al) methodology considers the interrelations of physical protection system elements and provides a framework for the integration of each element with: (1) A definition of what can be done to cause the undesired event (Fault Tree Study); (2) A physical description of the facility (Plant Physical Layout); (3) Detail of the security system (Security System Description) and (4) Characteristics of the adversary (Adversary Attributes). [56]

The Bennett (et al) methodology involves a very detailed analysis that is time consuming to create, and while this type of analysis is essential during the design of an adequate physical protection system, ongoing evaluations of effectiveness become unwieldy. Further, the methodology did not integrate a complete set of domestic safeguards subsystems and components.

Wilkey et. al. worked on defining a set of quantifiable metrics to allow consistent analysis of Material Control and Accounting (MC&A) effectiveness. Their intent was to

model the effect of changes to the systems used and to quantify the extent to which these changes improve the effectiveness and efficiency of the systems. [57] Their effort focused on: (1) Analysis of MC&A systems and identification of important activities; (2) Identification of MC&A activities that are amenable to quantified evaluation; (3) Selection of a set of MC&A activities from step (2) that are sufficiently representative of the MC&A program to allow their metrics to be used to evaluate the MC&A program as a whole; and (4) Development of quantitative metrics for the MC&A activities selected in step (3) and a function for using the metrics to calculate the overall system effectiveness.

Wilkey's (et al) effort produced a breakdown of the MC&A Program in order for metrics to be established. Detailed areas of metrics were presented. The work focused on the areas of Material Control, Material Accounting, Measurement, and System Design; however, it did not produce a function to determine overall effectiveness for MC&A systems.

Al-Ayad and Judd specifically look at a framework for evaluating the effectiveness of nuclear safeguards systems and describe an analytical tool for that evaluation. The model evaluates probabilistic input data to evaluate performance using the probability and time to detect material diversion attempts and prevent them [24]. Their approach is to aggregate performance indices over all threats. Their *Aggregated Systems Model* (ASM) uses a two step process to aggregate performance. First each adversary type is assumed to select a strategy with the greatest chance of success. Second weighted averages are calculated based on a subjectively assigned likelihood for each adversary scenario. While this is useful during design, no ongoing performance data is contemplated or used in the analysis.

Sicherman, Fortney, and Patenaude use a database model for evaluating material accounting effectiveness. This model was implemented in a Protracted Insider Module added to the Analytic System and Software for Evaluating Safeguards and Security (ASSESS). It recognizes that an insider is likely to make a protracted effort to divert material. The Protracted Insider module has four general steps in the material accounting effectiveness evaluation: (1) Specify activities capable of detecting anomalies for a target of interest, and how often they are performed; (2) Model each detection activity in terms of how data is generated, transmitted, processed, etc.; (3) For each activity, specify personnel with access to each stage in the activity affecting the integrity of material accountability information (e.g., people who measure, enter/transmit and process MC&A data, and those who are involved in calibration or other procedures that could affect the validity of accounting information); and (4) Perform an analysis using input from the previous steps to estimate probabilities of detection for each MC&A activity, given particular insider methods for subverting the activity. [58] This methodology integrates all of the aspects of a safeguards system but, it is closer to FMEA and remains probabilistic with no performance data used.

2.7 Utility Theory

Peter C. Fishburn, a noted pioneer in decision theory, describes utility theory as being concerned with people's choices and decisions. He further elaborates that utility theory deals with "*people's preferences and judgments of preferability, worth, value, goodness or any number of similar concepts.*" [59] Fishburn lists the two classifications of utility theory as ***predictive*** and ***prescriptive***. The predictive approach is concerned with the ability to predict choices or behavior. The prescriptive approach attempts to indicate how an individual should make a decision. In the case of effectiveness calculations the concept adapts to real-time decision making regarding how well a system is working. In this case the effectiveness calculation is prescriptive in that the utilities of individual factors are those that the decision makers are determining as their preferences for importance in a multidimensional environment.

Assumptions in utility theory include the primary proposition that utility numbers $u(x)$, $u(y)$ can be assigned so that if x and y are in sample space X then:

$$x \preceq y \text{ if and only if } u(x) \leq u(y)$$

This is read:

x is not preferred to y if and only if the utility of x is less than or equal to the utility of y .

Fishburn makes a simplifying assumption of independence for the individual utility values. In the case described for this research for a "family of Systems" it is assumed systems are decomposed to a point of independence. There are two forms for multidimensional preferences:

a) **additive** forms where

If $x = (x_{i1}, x_{i2}, \dots, x_{ij}, \dots, x_{in})$ then:

$$u(x_{i1}, x_{i2}, \dots, x_{ij}, \dots, x_{in}) = u_{i1} + u_{i2} + \dots u_{ij} + \dots u_{in}$$

b) **lexicographic** forms which include one factor that is overwhelmingly more important than others. [60]

Fishburn indicates that applications of lexicographic utilities appear limited because of the over-whelming importance notion. For an example, if a prisoner of war is unwilling to reveal vital information regardless of the torture he may undergo, his decision is based on an unwillingness to divulge the information and thus is the overwhelming factor.

Utility theory is often used in selection of alternatives. In this case the result is a choice between alternatives based on specific criteria or considerations. Diane Halpern

presents a method and worksheet for selection of alternatives. A summary of the process consists of [61] [62]:

1. **Frame the Decision.** Make a precise statement of the problem that will help to narrow it. Give clear and careful thought to this first step. The way in which the problem is defined will determine the character of all the following steps.
2. **Generate Alternatives.** Think of all the possible alternatives that could solve the problem.
3. **List the Considerations (Aspects).** Write down all the variables (considerations) that affect the decision.
4. **Weight the Considerations.** Give each consideration an importance score that reflects its relative importance to you. Use a 5-point scale ranging from 1 = of slight importance to me and 5 = of great importance to me. Some considerations will probably have the same weights, but your considerations should represent a range of importance ratings.
5. **Weight the Alternatives.** Take the alternatives one at a time. For each alternative, decide how well it satisfies each consideration. Rate the alternative on each consideration, using a scale of -2, -1, 0, +1, +2. Positive numbers indicate the alternative is favorable or compatible (or “pro”) on that consideration; +2 indicates highly compatible, +1 indicates somewhat compatible. Negative numbers indicate the alternative is incompatible or unfavorable (“con” on that alternative; -2 indicates highly incompatible, -1 indicates somewhat incompatible. A rating of 0 indicates neither favorable nor unfavorable to the decision. In the column for that alternative, write the weighting next to each alternative.
6. **Calculate the Decision.** Do each alternative separately. Multiply the assigned weight for each consideration (1 to 5) by the satisfaction weighting for the alternative (-2 to +2). Write the product in the extra column next to that consideration. To find the total assessment score, add all the products for that alternative.

Table 4 shows a partially completed worksheet to illustrate the technique. In this case alternate 1 (Alt 1) and alternate 2 (Alt 2) have been scored against considerations 1 – 7 (Cons 1-7). On the Cons 1 line which is weighted (5) and Alt 1 is weighted -2 yields a score for Alt 1 on Cons 1 of -10. This continues for all alternatives in Table 4 and the total scores are for **Alt 1 = -2** and for **Alt 2 = 10**. In this case Alt 2 is preferred to Alt 1.

Turban, and Metersky develop a utility theory approach (MAUT) to multivariable system effectiveness for the Naval Air Development Center. They attempt to reduce subjective evaluation of the system's performance by developing a procedure, based on decision and utility theories, so that the evaluation can be calculated objectively. Turban and Metersky analyze a system by finding the value $U(A_i)$ they call system effectiveness.

Table 4 MAUT Worksheet

	Alt 1		Alt 2		Alt 3		Alt 4		Alt 5	
Cons 1 (5)	-2	-10	2	10						
Cons 2 (5)	-1	-5	1	5						
Cons 3 (4)	0	0	0	0						
Cons 4 (3)	1	3	-1	-3						
Cons 5 (3)	2	6	-2	-6						
Cons 6 (2)	1	2	1	2						
Cons 7 (1)	2	2	2	2						
	Total	-2	Total	10	Total		Total		Total	

where:

$$U(A_i) = U(x_{i1}, x_{i2}, \dots, x_{ij}, \dots, x_{im}), \quad i = 1, 2, \dots, m$$

And in looking at different solutions for a given system design one tries to maximize $U(A_i)$. Their approach is based on the additively assumption mentioned above which can be stated as System Effectiveness is the sum of the values of the measures of effectiveness (MoEs).[63] The value of the MoE (u_i) for each course of action is determined by the product of the utility of the MoE $u(M_{ij})$ times the utility of the performance level x_{ij} of the MoE [63].

$$u_{ij} = u(M_{ij}) \bullet x_{ij}$$

Turban and Metersky modified the Churchman-Ackoff method in which the subjects were to give a 1.0 value to the most important MoE and values between 0 to 1.0 to the other MoEs [64]. According their modified method the subjects were asked to allocate 100% of importance among all MoEs.

The approach taken by Turban, and Metersky involved: (a) Use the modified Churchman-Ackoff approach to obtain rough estimates of the relative values of the performance variables $u(M_{ij})$; (b) Use the consistency test suggested by Churchman and Ackoff to adjust and improve the estimation; (c) Employ the above method with several dozen independent judges (group judgment); (d) Employ a statistical test (Coefficient of Concordance) to find the degree of agreement among judges; (e) Reconcile differences among judges; (f) Find the utility of the output coefficients (x_{ij}); (g) Compute u_{ij} 's and the system effectiveness (which is the sum of the values of the u_{ij} 's) and (h) select the alternative course of action with the highest value

The Churchman-Ackoff utility theory approach as used by Turban, and Metersky shows significant promise in developing a methodology (using today's technology i.e. a statistically analyzed web-based survey) for finding the correct proportional combination

of dissimilar and sometimes conflicting subsystems and subsystem elements into a combined measure of safeguards system effectiveness.

Turban and Metersky assert that the additivity assumption seems to be reasonable in their case since they deal with narrow ranges of $U(A_i)$, and with several u_{ij} 's, none of which is significantly large (over 30% of $U(A_i)$). They note that their procedure is valid in the range of one decision situation.

This research uses a systems approach of defining and decomposing the top level system into units with measurable Type I and Type II errors (assumed to be independent) and the determining the utility factors through a web-based survey of experts in the domestic nuclear safeguards community. An additive utility model will be used in this research but that does not preclude the use of a lexicographic model in future work. This research is consistent with the assumptions of Fishburn in utility theory as well as the boundaries expressed by Turban and Metersky in the application to system. Turban and Metersky have established a methodology that calculates the system effectiveness of design alternatives for systems in a static case. The extension from the static analysis of Turban and Metersky to the dynamic analysis in this research is reasonable.

2.8 Facility Scenarios

A safeguarded facility could contain any number subsystems or layered defenses for protecting the material. Figure 10 shows a situation under which the defenses are seen in three distinct layers. The layers can be seen as dependent (sequential) or independent. The DOE **system performance effectiveness equation** (here seen when $P_{Ei} = P_{di}$) would model this situation with the layers considered sequential and dependent as $P_E = \text{protection effectiveness}$.

$$P_E = P_{d1} + \{(1 - P_{d1}) \times P_{d2}\} + \{1 - (P_{d1} + \{(1 - P_{d1}) \times P_{d2}\}) \times P_{d3}\}$$

This approach scenario dependent and is based on an analysis of the attacker's pathway and predetermining it to be a linear attack (Layer 1 then Layer 2 then Layer 3). Figure 11 shows a graphic that includes three layers subject to linear attack. In each layer 70% of the attacks are detected/neutralized ($P_d = 70\%$ for each layer can be seen in Figure 11) resulting in a $P_E = 0.97$ for this specific situation as seen in Table 5. The heuristic system effectiveness approach would model a three facility as not being scenario dependent and the layers would be independent of attacker pathway as seen in Figure 12 and the equation used to calculate system effectiveness for a family of systems using MAUT would be:

$$e_s = \text{System Effectiveness} = \sum u_i e_i$$

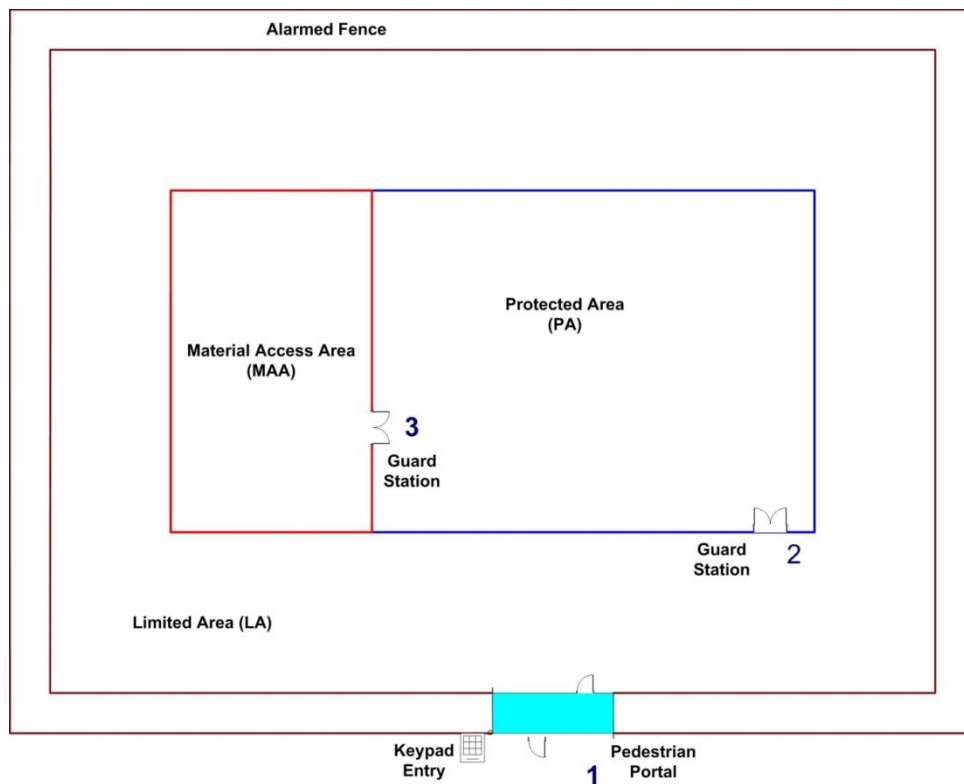


Figure 10 Example of Three Layer Facility

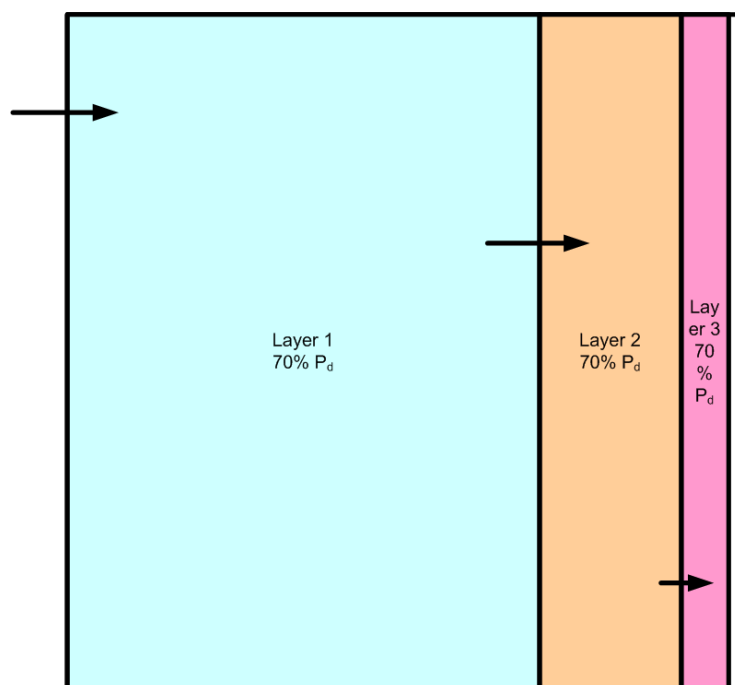


Figure 11 Three Dependent Layers

Table 5 Comparison of Methods

	e_i	u_i	e_i	u_i
P_{d1}	0.70	0.33	0.70	0.33
P_{d2}	0.70	0.33	0.70	0.33
P_{d3}	0.70	0.33	0.00	0.00
P_E	0.97		0.91	
e_s		0.70		0.47

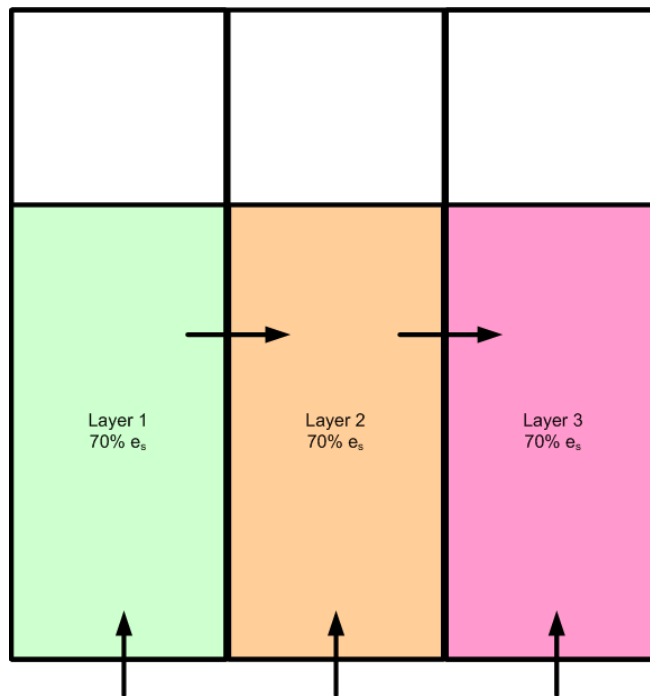


Figure 12 Three Independent Layers

Subject to: $0 \leq e_i \leq 1 \quad \sum u_i = 1$

Where:

u_i = utility of the i th subsystem

e_i = the effectiveness of the i th subsystem

And with equally distributed utilities (33%), and 70% subsystem effectiveness and overall system effectiveness would be $e_s = 70\%$. This generic attack analysis would include non linear attacks such as an insider which has access to the facility and would not need to penetrate layers one and two, but could begin the attack inside layer three.

Comparing the DOE method to the system effectiveness equation shows that for a specific dependent layered pathway analysis, the protection effectiveness can be very high ($P_E = 0.97$) while the system effectiveness would still be 70%. Also seen in, if one of the layers loses effectiveness the $P_E = 0.91$, which in the specific case of the linear attack, may still be acceptable for that pathway, however, the system effectiveness is reduced significantly to $e_s = 0.47$, which would likely be an unacceptable situation for the generic non-directional attack.

The two approaches are not measuring the same phenomenon but are complimentary. During the design and initial vulnerability assessments (VAs) of the nuclear safeguarded facility, all the possible attack pathways are analyzed. These initial VAs can require 5 technical experts for up to a year doing full time analysis. [65] During operation of the facility however, the system effectiveness methodology would be used to determine near real time operational system effectiveness performance (e_s). This relationship to a generic system lifecycle can be seen in Figure 13. There are, on an approximately annual basis, periodic VAs, which require 5 technical experts for approximately 3 weeks (or more on complex facilities). [65] A system effectiveness calculation can be included in the design of the system and provide ongoing performance information. In this case a statistician already on staff could provide weekly (or daily) updates of system performance with very limited effort. Should the ongoing system effectiveness measurement give the system owners enough information and confidence to extend the time between periodic VAs or shorten the time required to perform the VA, significant cost savings could occur.

Summarizing the two methods, the system effectiveness calculation represents a more robust (capable of performing without failure under a wider range of conditions) measure than the scenario dependent protection effectiveness calculation. This means that as e_s increases so does the robustness of the system. As the protection effectiveness (P_E) increases the system is only improving with respect to a specific attack scenario. In the case where $e_s = P_E$ the system is more robust against attack.

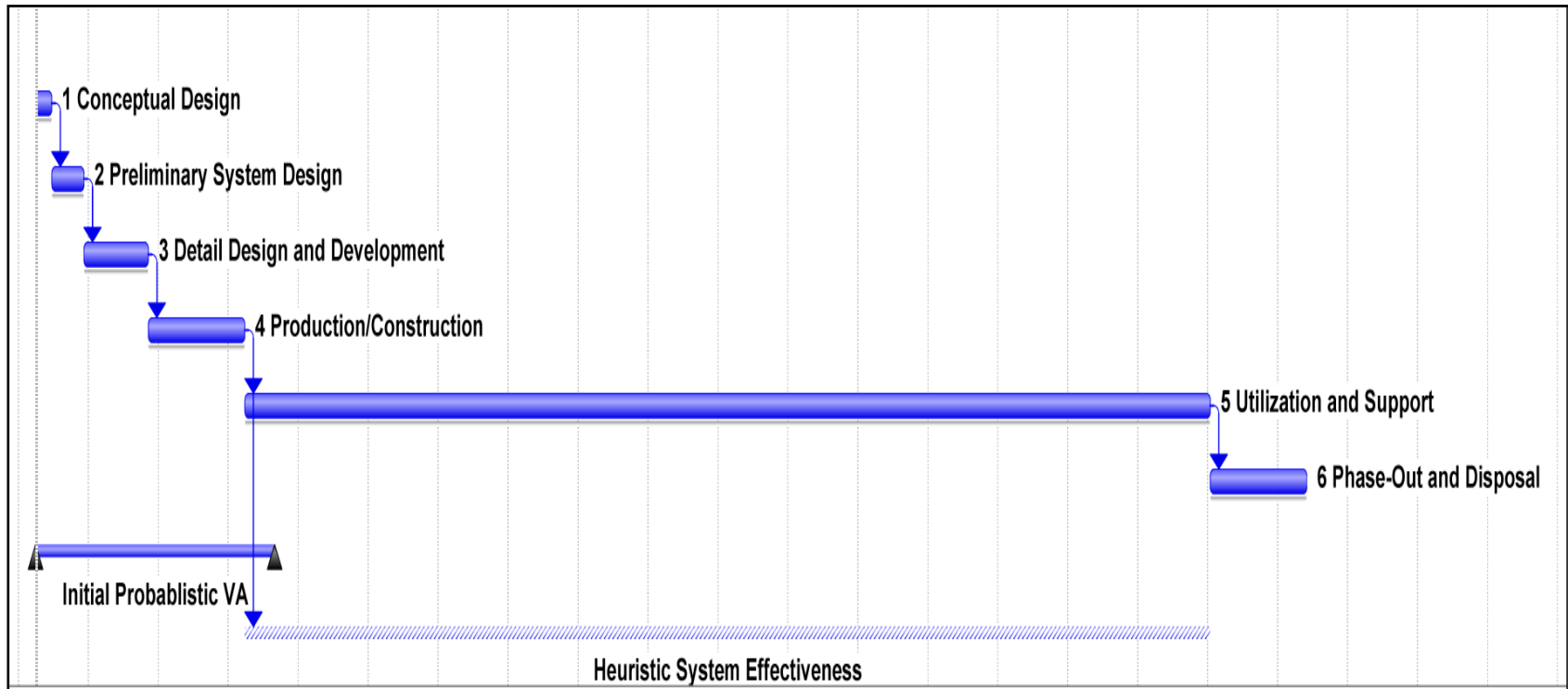


Figure 13 Generic System Lifecycle

Chapter 3 Research Methodology

Chapter three presents the “research methodology” for answering the research questions and confirming the hypotheses. In Chapter 1 the research was determined to be important and needed in the domestic safeguards CoP. Chapter 2 provided a comprehensive review of the relevant literature. Chapter 3 will present the “*methodology for determining system effectiveness*” (System Effectiveness Methodology) in summary but focuses on the “research methodology” for answering the research questions and confirming the hypotheses. To address the research questions (A. and B.) and hypotheses 1 and 2), a checklist of questions must be addressed:

- A. *Can a methodology be formulated and used to analyze the effect of incorporating measures of effectiveness concepts at the component-subsystem level, yet assess them at the system level?*

Hypothesis 1: *The Multi-Attribute Utility Theory (MAUT) methodology is a valid approach to incorporate safeguards measures of effectiveness at the component-subsystem level and assess them at the safeguards system level.*

1. Is the proposed solution a Methodology?
 - To be shown through synthesis of literature definitions
 2. Are the selected inputs measures of effectiveness?
 - To be shown through a survey users to determine measures and synthesize results
 3. Can the measures of effectiveness be aggregated at the subsystem/component level to produce an overall effectiveness measure?
 - Develop a function and heuristic to support aggregation
 4. Can the System Effectiveness Methodology be compared to a validated methodology?
 - The System Effectiveness Methodology will be validated through Synthesis of valid approaches from literature review, application of same data to all and through simulation
- B. *Can uncertainties, such as incomplete data, rapidly changing technologies, and uncertainties in the modeling environment be addressed through heuristics in a system effectiveness methodology?*

Hypothesis 2: *Calculating safeguards system effectiveness using Type I and Type II error rates is a valid modeling approach upon which to base an effectiveness methodology with incomplete data, rapidly changing technologies, and uncertainties*

1. Is the System Effectiveness Methodology robust than the DOE Protection Effectiveness Model?
Show conditions under which the methodology has broader application
2. As the level of experience increases how are the heuristics affected?
Show that with increased experience and data there is more confidence in the heuristics
3. What relationship do incomplete data, rapidly changing technologies, and uncertainties in the modeling environment have with the heuristics?
Show how incomplete data, rapidly changing technologies, and uncertainties affect the heuristics.

To develop the System Effectiveness Methodology, firstly a suitable MoE will be determined and evaluated, Secondly the “*methodology for determining system effectiveness*” will be defined and specific goals for the model/methodology and finally the research questions and hypotheses will be addressed.

The MoEs selected are based on the error rates α and β . The MoEs will be developed and the MoE validity established based on Smith and Clark’s definition. [28]

Goals of the System Effectiveness Methodology are:

- That it uses a clear definition of a system and system effectiveness that is broad across many applications
- That it complements the probabilistic approach (DOE approach of P_E);
- That it add to the open literature concerning systems effectiveness measurement and
- That it recognizes the complexities of system structures.

The System Effectiveness Methodology: will be based on heuristic rather than probabilistic principles although probabilistic elements are compatible; is to be a generalization in order to have as broad of application as possible; will provides an ongoing single measure of effectiveness quantification based on system performance for a working system in near real-time. A summary of the System Effectiveness Methodology would have the following steps:

1. Define the family of systems to be evaluated.
2. Decompose the system into subsystems (to the point that error rates can be measured).
3. Determine the utilities (u_i) of each subsystem.
4. Measure the error rates α and β for each subsystem.
5. Calculate the subsystem effectiveness (e_i) for each subsystem (function addressed later in this chapter).
6. Calculate e_s using the following **Family of Systems Effectiveness Equation**:

$$e_s = \text{System Effectiveness} = \sum u_i e_i$$

3.1 Introduction To System Effectiveness Function

Hamilton and Chervany summarize problems in evaluating system effectiveness as: (1) Objectives and measures of accomplishments are often inadequately defined initially; (2) Efficiency-oriented and easily quantified objectives and measures are typically employed; (3) Objectives and measures used to evaluate the system are not the same as those defined initially; and (4) Individual perceptions may differ on what the objectives and measures are. Hamilton and Chervany summarize that realistic mutual agreement concerning the definition of appropriate objectives and measures of accomplishment is typically not reached at the outset and this makes evaluation of system effectiveness difficult.[53]. Soban and Mavris recognize the need for an integrated and efficient framework that can rapidly assess system effectiveness for today's complex systems. They recognize that the definitions of system effectiveness vary widely and are often application dependent.[55]

Hamilton and Chervany indicate that there are two fundamental approaches to evaluating system effectiveness: a systems resource approach and a goal centered approach. [52] Under a **resource approach**, system effectiveness is determined in terms of resource availability instead of specific task objectives which is seen, according to MIT, as an **efficiency** [20]. The term cost-effective captures the essence of the resource approach but also leads many to confuse efficiency with true effectiveness measures.. The term cost-effective is usually modified by either "more" or "less" indicating a relative evaluation and not usually described in terms of percent effective. Under a **goal centered approach**, effectiveness is determined by comparing performance against specific objectives which agrees with the MIT definition of an **effectiveness**. Often under a goal centered approach, the effectiveness is defined in terms of percent effective and usually indicates a more absolute measure of obtaining the objective. Under either evaluation approach, there will always be interest in improvement.

This research is based on the goal centered approach to systems effectiveness evaluations. As was discussed the resource based approach is a measure of efficiency (MIT Definitions). [20] For system effectiveness measurement, the concept is simple, know the system goal and know, with a measure of precision, how often the system goal is achieved. This is particularly important in systems that protect assets (safeguard systems) because if the asset is not safeguarded, at some point in time the probability of losing the asset to a malicious actor becomes unacceptable. A standard definition of system effectiveness is made by Blanchard and Fabrycky "*the probability that a system can successfully meet an overall operational demand, within a given time, when operated under specific conditions.*"[21] This captures the nature of systems effectiveness but fails to account for types of errors. Using this definition, the measure of effectiveness becomes how often the system functioned correctly (1-Type II error

rate). What is lost is the rate that the system failed to function when it should have (Type II error rate)

Type I and Type II errors are defined in statistics as errors based on testing the null hypothesis (H_0) and the alternative hypothesis (H_1). [48] This is sometimes called a truth table and can be seen in Table 2 and Table 3.

In a detection system, a Type I error, also called a false alarm, happens when the alarm is activated when it should not have been; a Type II error, also called a missed alarm, happens when the alarm is not activated when it should have been. The probability of each of these wrong decisions is denoted by:

α = probability of false alarm, that is, the probability of having no threshold (t) exceedance but alarm activation.

β = probability of missed alarm, that is, the probability of having critical threshold exceedance but no alarm activation.

The tolerance of a type I or II error is related to a trade-off between the benefits of a correct decision and the costs of a wrong decision and it could vary substantially, depending on the relative consequences of possible missed and false alarms. For example, the cost of a false alarm at the airport for metal detection is a few seconds of the passengers and screeners time while the cost of a missed alarm could be as high as a hijacked plane.

In general, the effectiveness of a system is based on its ability to make correct decisions that minimize the probability of false and missed alarms while improving the cost-benefit ratio. Because the probability of a wrong decision is primarily due to having only partial knowledge, any subsequent predictions are based on uncertainty.

Conceptually a Type I error in relation to a system could be likened to the times the system acted when it should not have acted and the Type II error as those times the system did not act when it should have acted. The terms α and β are the probability of Type I (false positive or false alarm) and Type II (false negative or missed alarm) errors, respectively. Other conceptual statements relating to these quantities are: $1-\alpha$, which is the probability the system “did it right” and is often referred to as the **confidence level** of a test; and $1-\beta$, which is the probability the system “did not do it wrong” and is often referred to as the **power** of a test.

Using the quantities α and β , a generic mathematical model for system effectiveness can be developed. The posited relationship of system effectiveness (e_s) for a generic system can be constructed as follows:

$$e_s = f(\alpha, \beta)$$

In developing the functions for this relationship, another conceptual construct will be useful. The quantities α and β can be thought of in safeguard systems as representing the probability (or rate) of either:

- 1) false alarms (α) also called the false positive fraction (FPF), the converse of which ($1-\alpha$) is the true positive fraction (TPF); or
- 2) missed alarms (β) also called the false negative fraction (FNF), the converse of which ($1-\beta$) is the true negative fraction (TNF).

As previously indicated there is always a tradeoff between false alarms and missed alarms based on the consequence of each. In searching for a function that expresses the effectiveness in terms of α and β which represent the error rates, the compliments of these values become of interest which indicate proper functioning. Further, recognizing that the quantities α , $1-\alpha$, β , and $1-\beta$ are always valued from 0-1, a relationship for system effectiveness needs to appropriately emphasize the false alarm term or missed alarm term when needed and as such, a penalty (discount) on effectiveness based on these quantities needs to be developed. Defining the baseline acceptable error ratio as $r = \alpha_0/\beta_0$ this becomes an indication of the system manager's tolerance of false alarms to missed alarms, where α_0 is the acceptable rate of Type I errors (false alarms) and β_0 is the acceptable rate of Type II errors (missed true alarms), each term can be adjusted based on this ratio.

This leads to the posited function for system effectiveness which increases the effect of α or β appropriately to the overall effectiveness and is seen below and which will be developed further in section 3.2.

$$e = (1 - \alpha^{K_a}) \cdot (1 - \beta^{K_b})$$

Where:

$$K_a = (r)^{1/c} \text{ for } r > 1 ; \text{ otherwise } K_a = 1$$

$$K_b = (r)^{-1/c} \text{ for } r < 1 ; \text{ otherwise } K_b = 1$$

The two parameters “ r ” and “ c ” are described as follows:

Parameter “ r ” is the relative tolerance of false alarms to missed alarms (α_0/β_0) in a non stressed environment.

Parameter “ c ” is the stress level to the system. This can be cost, threat level, increased consequence of failure, or other local stresses. When $c = 1$, stress is neutral. Increasing c means increasing stress levels.

The fundamental calculation for determining system effectiveness in this research is twofold:

- 1) develop a relationship for Type I and Type II errors as they relate to system effectiveness and to posit this relationship as a function for calculation of a system's effectiveness (given the rates of these error types) and
- 2) use an MAUT approach for combining the elements of a system into an overall effectiveness calculation.

Following the ability to calculate a system's effectiveness where the overall errors rates can be determined, there are cases where the overall error rates for the full system are not available and it is necessary to decompose the system into subsystems and components that can have their error rates determined. The utility theory, whereby the utility of each sub element is determined and the effectiveness for that element is measurable yields a methodology for combining the different values into one overall system effectiveness.

3.2 Function Development and Use

The starting point is the development of a relationship with which to calculate the effectiveness of a given whole system. In this case, whole system, means a generic system that does "something" of value and that doing "the something" wrong is negatively valued and the error rates are known. The starting point in attempting to develop a relationship calculating system effectiveness in terms of Type I error and Type II error is to posit that system effectiveness is a function of the types of error as seen below.

$$e_s = f(\alpha, \beta)$$

Taking an intuitive approach the basic function was to simply posit that the effectiveness was the rate of true negative alarms (TNF or $1-\alpha$) times the rate of true alarms (TPF or $1-\beta$). This could be also be described as the **confidence level** times the **power** as seen the initial equation below and Figure 14 (generated from a MATLAB program to graph a solution mesh of the algorithm)

$$e_s = (1-\alpha) \bullet (1-\beta)$$

This simple function would likely be adequate if the relative consequence of false alarms and missed alarms were equal, and in a number of applications this may be the case, however, in real world nuclear safeguards applications this is rarely the case. Using the above equation as a starting point, the difference in relative importance of false alarms to missed alarms needs to be added. Good design practice for a system will include a specific set of tolerances to false alarms (baseline tolerance of false alarms = $\alpha_0 < 0$) and missed alarms (baseline tolerance of missed alarms = $\beta_0 < 0$). The ratio of these tolerances will be defined as " r ", or the **baseline tolerated error ratio** ($r = \alpha_0/\beta_0$), which by definition is positive and can be used to further develop the algorithm.

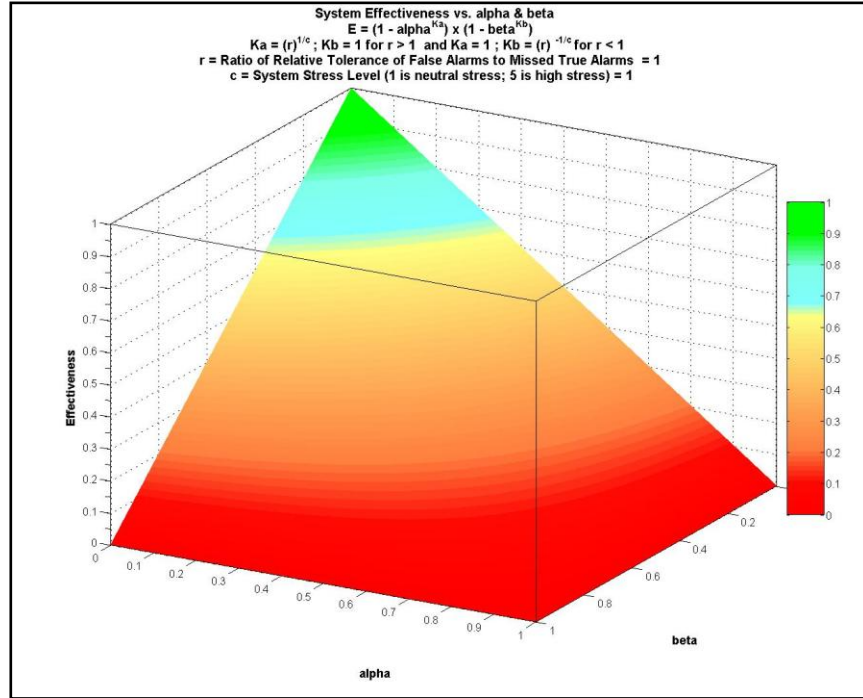


Figure 14 Effectiveness as a Simple Function of Confidence Level and Power

With “ r ” defined, the next step is to determine where this effect should be added in the calculation. Initially an inverse exponent relationship of the $(1-\alpha)$ and the $(1-\beta)$ terms was chosen with the intention that with a preference for one type of error the other type of error should be discounted. This creates the first iteration to the basic function as seen below.

$$e_s = (1 - \alpha)^r \bullet (1 - \beta)^{1/r}$$

Figure 15 also generated from a MATLAB program to graph a solution mesh of the function shows the solution space (calculated effectiveness for values of α and β between 0 and 1) with $r = 2$ (as a convenience r is chosen as 2 where two false alarms are tolerated for every missed alarm). There appears to be an opposite effect of that which is desired as the expected result would have tolerated more false alarms for the effectiveness value calculated. Figure 15 shows the effectiveness falling at a higher rate verses false alarms than the rate of effectiveness decrease verses missed alarms and in fact shows tolerance for missed alarms. Intuitively, when a tolerance for a given type of alarm is present, the rate of that type of alarm should be discounted and not enhanced. This effect becomes more clear at the more extreme values such as shown in Figure 16. In this case, there is a preference to allow 200 false alarms (α_0) to 1 missed alarm (β_0). Figure 16 shows that a small increase in the false alarm rate lowers the effectiveness to almost zero while a large increase in the missed alarm rate has very little effect. This is the reverse of the situation needed to discount the effect of a

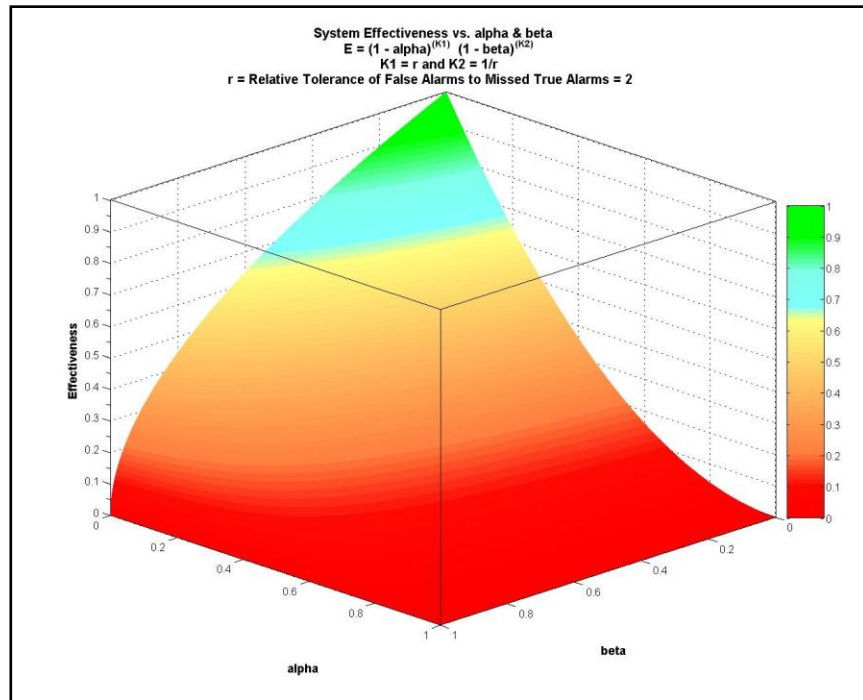


Figure 15 First Iteration Solution Space with $r = 2$

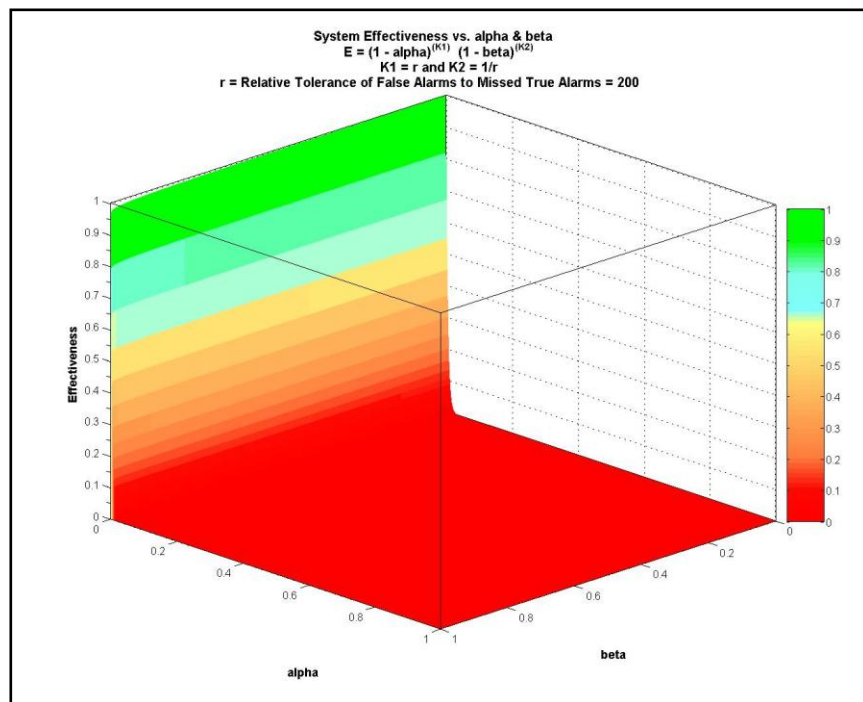


Figure 16 First Iteration Solution Space $r = 200$

false alarm. Another factor in the function structure is in the concept of discounting (using parameter r) the correct responses of the system ($1-\alpha$ and $1-\beta$) seems intuitively backwards and the discount should act on the rate of errors not the rate of correct responses. In order to correct the situation a second iteration to the function was made and can be seen in the following equation where the discount operates on the error rates.

$$e_s = (1 - \alpha^r) \bullet (1 - \beta^{1/r})$$

Figure 17 shows the solution space for the above second iteration equation with $r = 2$. Again in this case two false alarms would be tolerated for every missed alarm. This has the effect of discounting the error rate terms, however this applies the parameter r to both the α and β terms. Intuitively, when a tolerance for a given type of alarm is present then only the rate of that type of alarm should be discounted. Figure 17 shows that with a tolerance for false alarms the effectiveness calculation has a tolerance for false alarms, however missed alarms seem to have an enhanced affect. This can be seen at the extreme values as seen in Figure 18, where false alarms (α) have almost no affect on the calculation of system effectiveness and a very small increase in missed alarm rates (β) causes a very significant decrease in effectiveness.

Conceptually, the calculation function should affect (discount) only the term where there is a tolerance. This leads to the creation of K-factors (K_a and K_b) which have separate domains. The correction can be seen in the system effectiveness function below which also contains a further modification to the calculation and recognizes that outside influences/conditions can affect a system's effectiveness. This term is introduced signifying increased consequences of failure or increased stress on the system (parameter c) that are changes from the baseline condition. This idea of stress and a discount leads to a two parameter relationship. The final form of the effectiveness equation can be seen below as the system effectiveness function.

$$e = (1 - \alpha^{K_a}) \bullet (1 - \beta^{K_b})$$

Where $K_a = (r)^{1/c}$ for $r > 1$; otherwise, $K_a = 1$

Where $K_b = (r)^{-1/c}$ for $r < 1$; otherwise, $K_b = 1$

The two parameters are described as:

Parameter " r " is the relative tolerance of false alarms to missed alarms (α_0/β_0) in a non stressed environment (normal operations).

Parameter " c " is the stress level to the system. This stress can be (inter alia) threat level, increased consequence of failure, or other local stresses. When parameter $c = 1$, stress is neutral and increasing c means increasing stress levels.

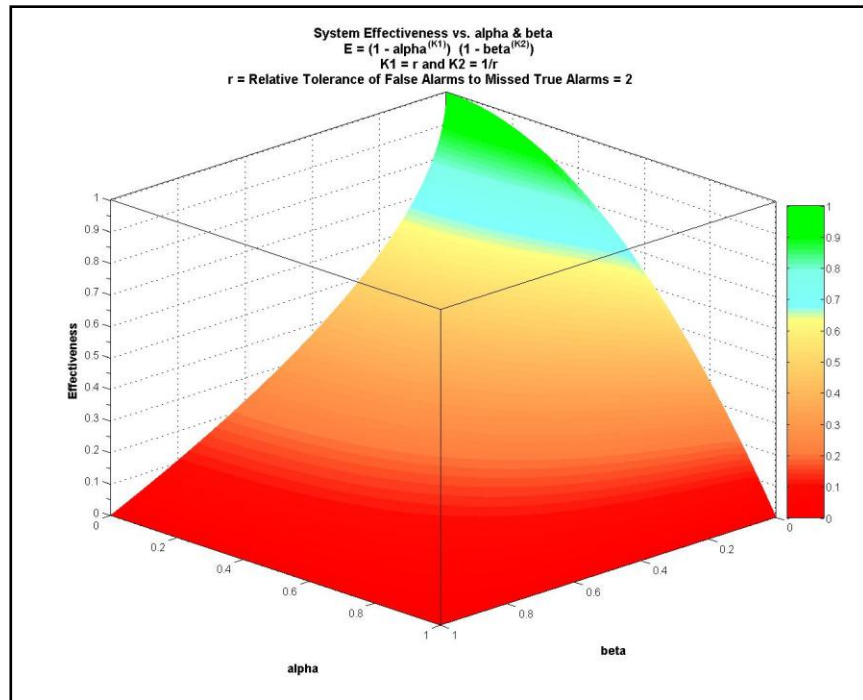


Figure 17 Second Iteration Solution Space with $r = 2$

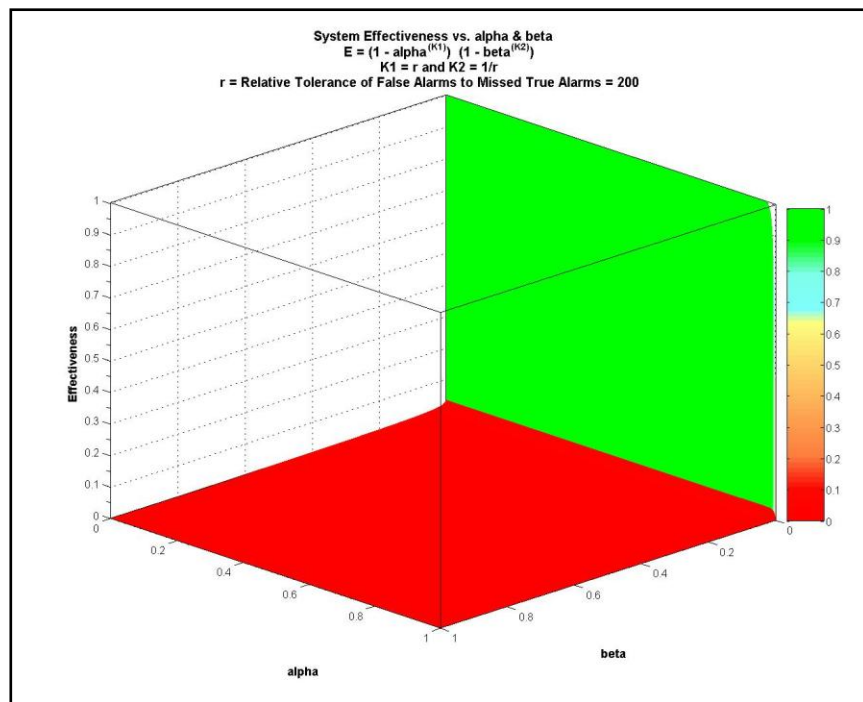


Figure 18 Second Iteration Solution Space with $r = 200$

Figure 19 shows the solution space for the system effectiveness function with $r = 2$. Again in this case two false alarms would be tolerated for every missed alarm. This has the effect of discounting the overall effectiveness based only on the tolerance for the specific type of alarm. The relationship operates as expected (discounting a false alarm but not discounting a missed alarm even at extreme values as can be seen in Figure 20. The effect of parameter c can be seen in Figure 21. Here the consequence/stress factor operates on the r parameter to lessen the discount.

When comparing Figure 20 to Figure 21, it is shown that under the same error rates (α and β), a lower system effectiveness is determined by the function based on a higher stress or consequence of failure (stress factor - parameter c). Figure 22 explores the affects of the stress factor (parameter c) on the discount. Example values are used in Figure 22 which is a graph of the alpha term $(1 - \alpha^{K_a})$ verses the Type I error (α) and a ratio $r = 20$. This curve was generated in Microsoft Excel where $K_a = (r)^{1/c}$ for $r = 20$ and $c = \{1, 1.5, 2, 5, 10, 20\}$. The $(1 - \alpha^{K_a})$ curve remains flat for high α values showing a large tolerance for false alarms before the effectiveness based on false alarms is reduced. As parameter “ c ” increases above 1 the tolerance for false alarms (and thus the discount of the α error rate) is reduced. This equates to increased stress on the system and changing the effectiveness based on the stressed conditions. Stress or consequence conditions could include: cost of a false alarm; consequence of a false alarm; increased stress levels to the systems environment or other similar stress/consequences.

With a high enough stress, the tolerance for false alarms could be canceled to a large extent. A reasonable range for c has not been established in this research; however, the range could be from 1 to 20, 1 to 10, 1 to 5, or 1 to 2 depending on the application and the system owner’s requirements.

3.3 Case Study

The example of airport passenger screening in Figure 23 identifies the basic system elements of a passenger screening system. [66] In this example the Walk Through Metal Detector (WTMD) identified as C in Figure 23 will be the system boundaries. In the WTMD case, a false alarm (α) will be defined as alarming when a small amount of metal in a non hazardous form (some analysts might call this a nuisance alarm since metal is detected but it is not a threat). A missed alarm (β) would be a threatening form of metal making it through the WTMD. In this example under non stressed conditions airport security may be able to accept 100 false alarms to 1 missed alarm ($\alpha_0/\beta_0 = r = 100$). The solution space for this situation would appear as in Figure 24.

If there is a stressed condition such as an specific threat to a given airport, then the effectiveness solution space changes in relation to the threat. In this case the example of the stress factor is assumed to increase to $c=5$ is shown in Figure 25. In Figure 25 the area of acceptable effectiveness (whatever the cutoff level of acceptable

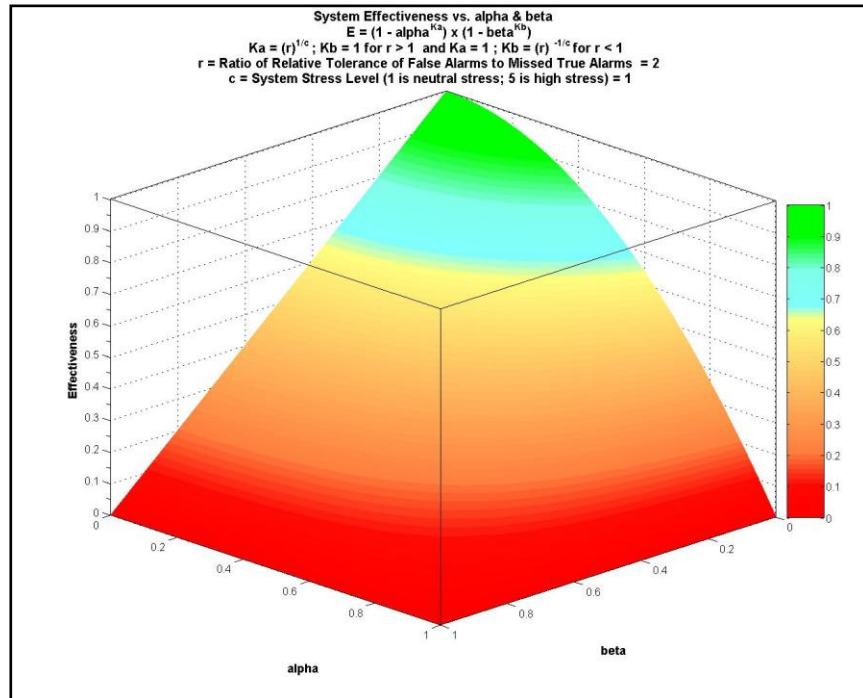


Figure 19 System Effectiveness Function Solution Space $r = 2, c = 1$

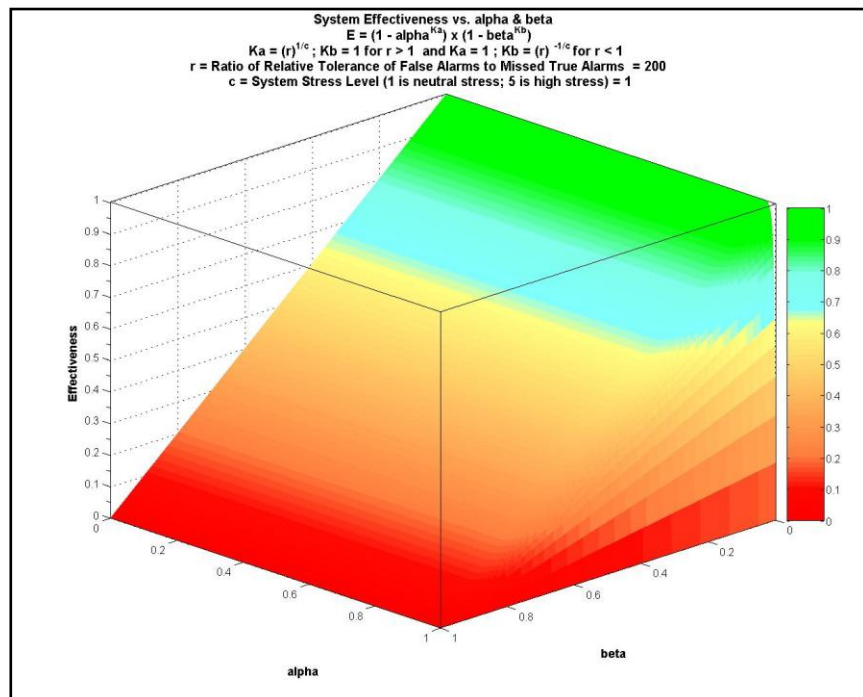


Figure 20 System Effectiveness Function Solution Space $r = 200, c = 1$

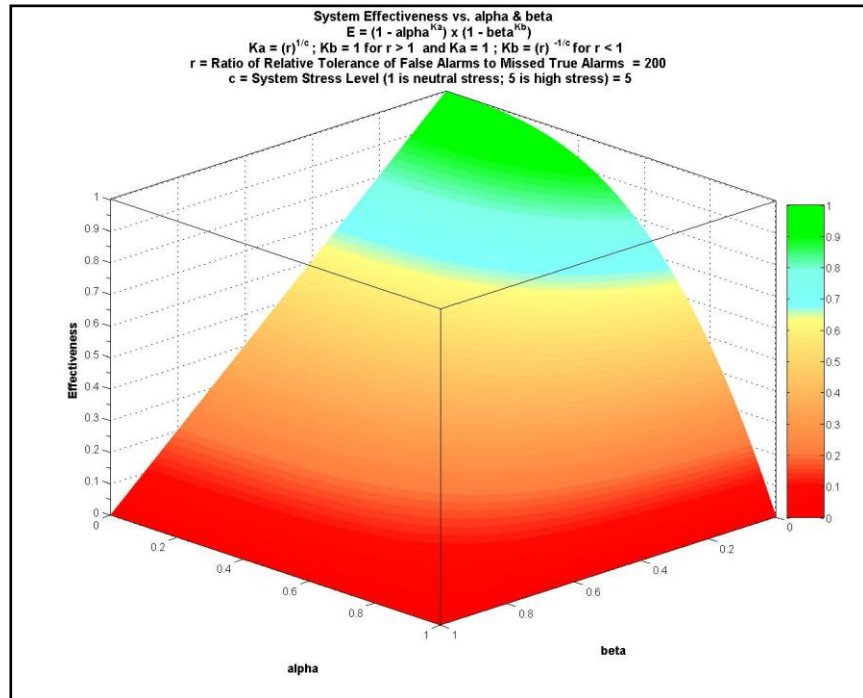


Figure 21 System Effectiveness Function Solution Space $r = 200$, $c = 5$

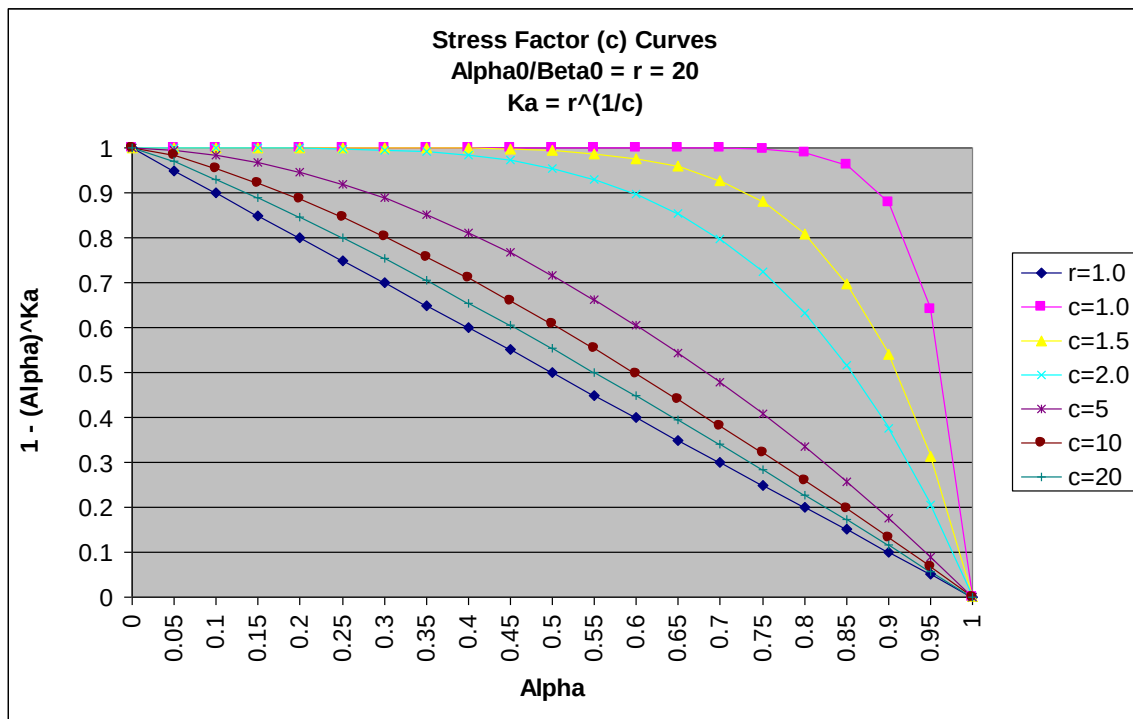


Figure 22 System Stress Factor Curves

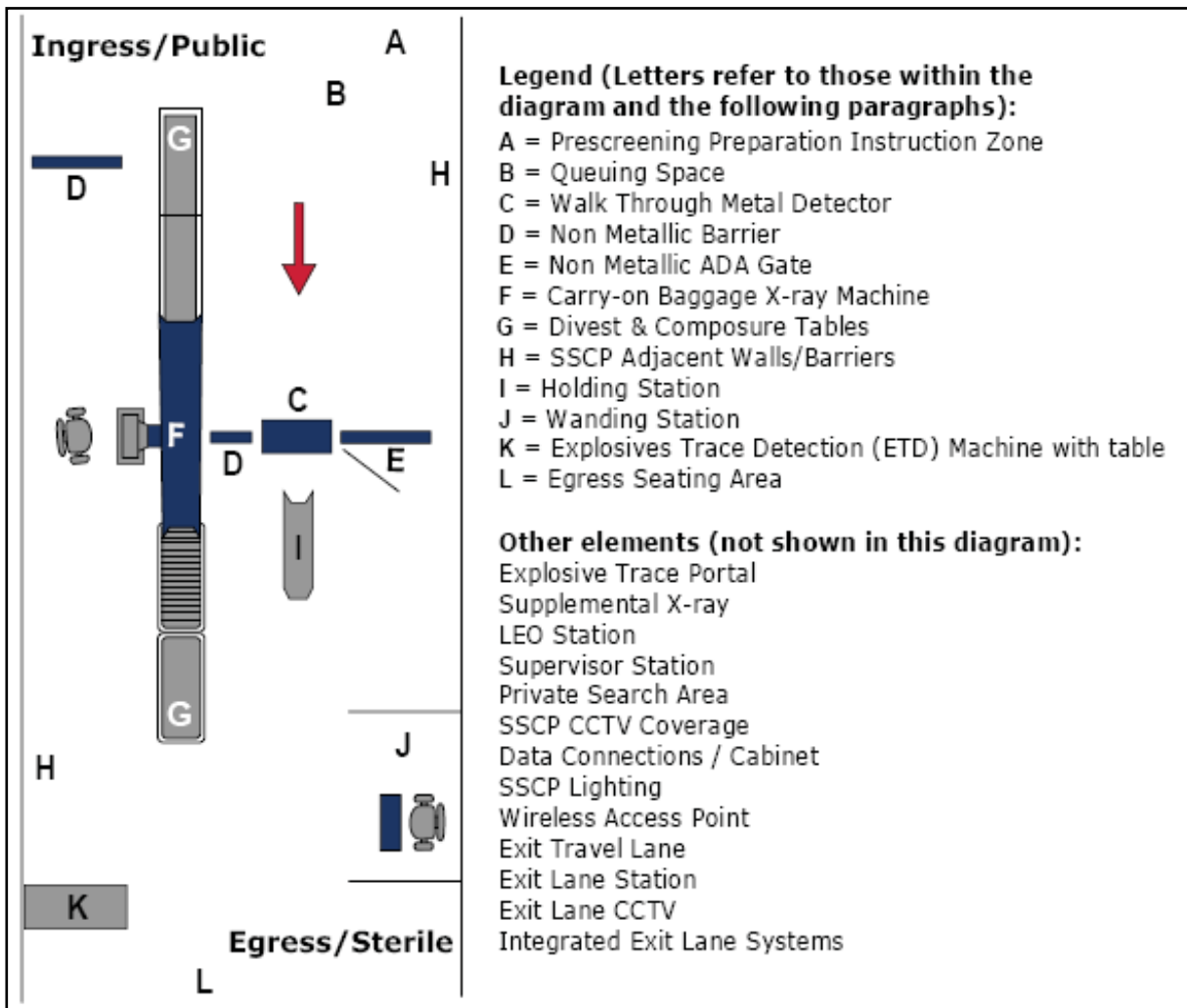


Figure 23 Airport Passenger Security Screening

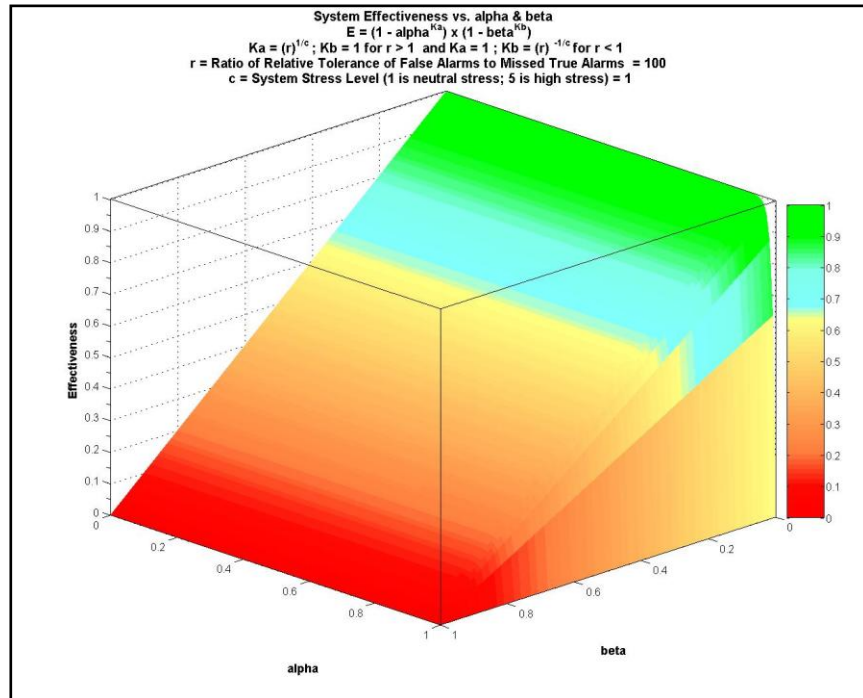


Figure 24 Example Baseline Airport Solution Space

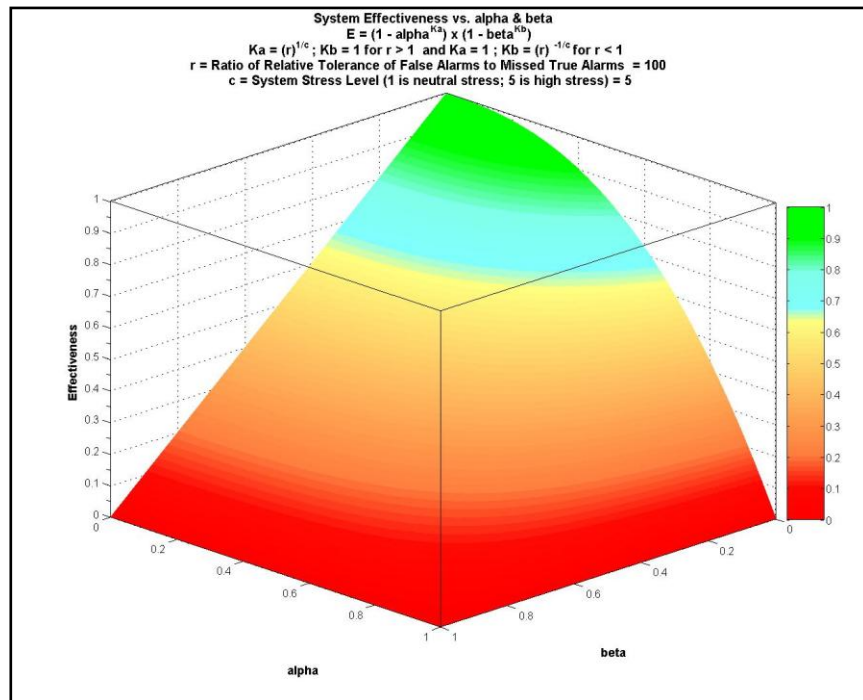


Figure 25 Example Airport Security Under Stress

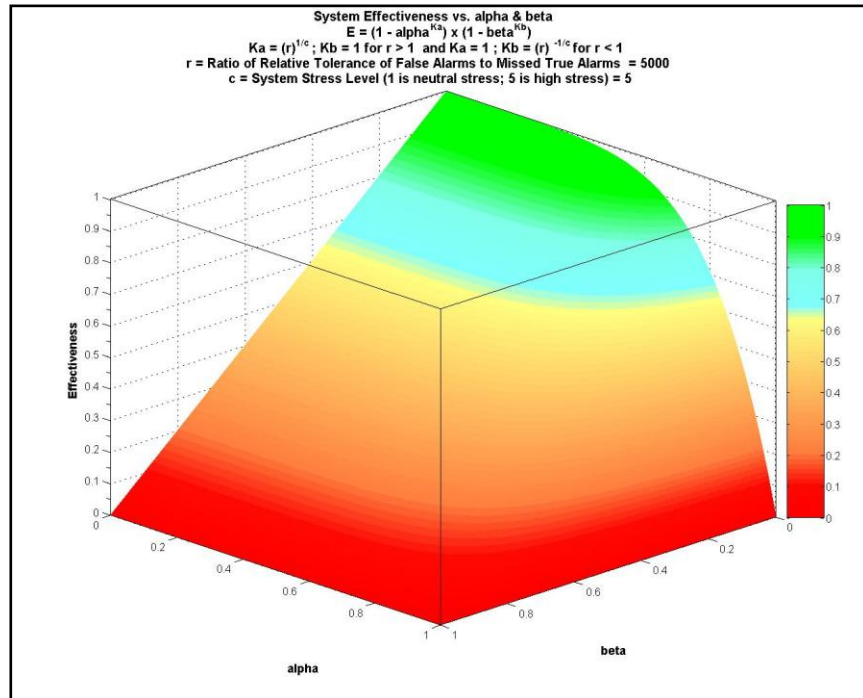


Figure 26 Example Airport Security Under Stress With Compensation

effectiveness may be) has decreased significantly. If the system owner wanted to make a change to the solution space to gain back a logical change would be to greatly increase the tolerance to false alarms. Such that $r = 5000$ and this can be seen in Figure 26.

Figure 26 shows that the compensatory measure of increasing the sensitivity on the WTMD by allowing 5000 false alarms to one missed alarm has indeed increased the solution space for system effectiveness but it may not be considered enough and in that event some either some other compensatory measure would be required in order to regain baseline effectiveness requirements or additional risk would be accepted by the system owner.

3.4 Safeguards System Effectiveness

The safeguards system requirements are generally related to the definition of safeguards. The DOE definition of a Safeguards System specifically lists three types of subsystems and implies another (integration can be correlated to management). Management is often not included in a system definition; however, this analysis adds management as an explicit subsystem. A given safeguards system can have a variety of components. The final form of a model will vary depending on the individual subsystems as well as the importance for each of the subsystems and its components.

In any system, subsystems are identified to handle different systems requirements. These subsystems can be in layers and/or compartments. The distinction is that subsystems in layers could cover multiple requirements leading to potential interactions between subsystems while compartmentalized subsystems would not have interactions leading to simpler mathematical terms. The model chosen for this research is that of four independent compartments (physical protection, accounting, control, and management). This is an idealized case; however, the case can be made that compartmentalized subsystems is the worst case for effectiveness. In the absence of interaction, overlapping subsystems represent redundancy and increase the overall effectiveness for the overlapped portion.

In this utilitarian approach, the physical protection, material accounting, material control, and management subsystems are assumed to be independent and compartmented. This leads to application of the **Family of Systems Effectiveness Equation** which is a relationship for several independent compartmentalized sub-systems:

$$\text{System Effectiveness} = e_s = \sum u_i e_i$$

Where:

u_i = the utility the i^{th} subsystem
 e_i = the effectiveness of the i^{th} subsystem

Such that for: Physical Protection $i = p$; Material Control $i = c$; Accountability $i = a$; and Management $i = m$; also, the effectiveness of a Safeguards System (e_s) can be represented in a utility model as a linear combination the effectiveness of the Physical Protection (e_p), the Material Control (e_c), and the Material Accounting (e_a) and Management (e_m) subsystems leading to Equation 7 which can be expanded in the domestic safeguards case such that we can:

Let

$$e_s = u_p e_p + u_c e_c + u_a e_a + u_m e_m$$

subject to: (s.t.)

$$0 \leq e_i \leq 1 \quad \sum u_i = 1$$

Where:

u_p = the relative importance of the physical protection subsystem
 u_c = the relative importance of the material control subsystem
 u_a = the relative importance of the material accounting subsystem
 u_m = the relative importance of the management subsystem

Given the defining relationship in the system effectiveness function and model in the family of systems effectiveness equation, subject to the relationships defined for each

subsystem, a repeatable and quantifiable value of system effectiveness for a given safeguards system based on the rates of Type I and Type II errors becomes workable.

And for specific sub-systems with individual error rates:

$$e_p = (1 - \alpha_p^{K_{ap}}) \bullet (1 - \beta_p^{K_{bp}})$$

$$e_a = (1 - \alpha_a^{K_{aa}}) \bullet (1 - \beta_a^{K_{ba}})$$

$$e_c = (1 - \alpha_c^{K_{ac}}) \bullet (1 - \beta_c^{K_{bc}})$$

$$e_m = (1 - \alpha_m^{K_{am}}) \bullet (1 - \beta_m^{K_{bm}})$$

With K_{ap} and K_{bp} related to r_p the tolerated ratio and stress level for the physical protection subsystem

With K_{aa} and K_{ba} related to r_a the tolerated ratio and stress level for the material accounting subsystem.

With K_{ac} and K_{bc} related to r_c the tolerated ratio and stress level for the material control subsystem.

With K_{am} and K_{bm} related to r_m the tolerated ratio and stress level for the management subsystem.

This type of model and methodology also lends itself to a linear systems approach and subsequent optimization. [67] What remains is to determine the utility of the separate subsystems to define a solution space for calculating an example safeguards system's effectiveness and simulate an example environment to illustrate the utility of the methodology.

3.5 Description of the Research Survey Design

The fundamental approach used in this work for determining the relative importance of each safeguards system's subsystems was a web based survey voluntarily completed by Material Protection, Accountability, and Control practitioners, encompassing all areas of the safeguards field, in such a way that facilitated statistical analysis yielding reasonable subsystem component utilities (u_{ij}) which are estimations of the utility of the j^{th} component in the i^{th} subsystem.

3.6 Description of the Surveyed Population

The surveyed population in this case will be defined as members of the Institute of Nuclear Material Management (INMM) and those that attend the INMM annual Meetings. The Institute of Nuclear Materials Management was formed in 1958 [68] to encourage: (a) *The advancement of nuclear materials management in all its aspects;* (b) *The promotion of research in the field of nuclear materials management.* (c) *The*

establishment of standards, consistent with existing professional norms; (d) The improvement of the qualifications of those engaged in nuclear materials management and safeguards through high standards of professional ethics, education, and attainments, and the recognition of those who meet such standards; and (e) The increase and dissemination of information through meetings, professional contacts, reports, papers, discussions, and publications. The INMM currently has over 1000 members worldwide and holds annual technical meetings where approximately 900 individuals, both members as well as interested and experienced non-members attend.

3.7 Description of the Sample

The sampling method was self selection. The survey was voluntary and completed by those in attendance at the INMM 49th Annual Meeting held in Nashville TN. The attendees were given a notice of the survey and the web link to the survey page and encouraged to take the survey during the week of the annual meeting or shortly thereafter. The INMM holds an exhibit at each annual meeting with a high percentage of the attendees browsing through the exhibit area. A booth with two laptop computers was constructed to allow exhibit attendees from the Annual Meeting the opportunity to participate in the survey during exhibit hours. On numerous occasions both laptops were occupied and potential participants were asked to return and take the survey. No records were kept regarding the actual number of potential participants which were turned away and never returned. During the week of the exhibit, electronic mail was sent to each of the attendees that pre-registered indicating the booth number for the survey as well as giving the web link for individuals to take the survey online in their room, at work, or at home.

There were 926 registered attendees at the conference (including those that registered on-site and therefore not included in the electronic mail notice). A total of 102 completed the survey. This leads to a 95% confidence level of +/- 9.1% error. This is the first such survey recalled by many of the participants and there are no records concerning how many of the 926 actually attended the exhibit so taking the 926 population size accounts for the worst case survey completion rate. While the rate of response leads to a fairly high error bar, cooperation from the INMM executive committee was helpful in achieving this rate of response and is a good first effort. The INMM mission is to support research; therefore the executive committee supplied trade show booth space at a significantly reduced cost. Small trade show gift items were given to attendees as inducements to take the survey and the verbal response from exhibit attendees indicated the gift items were very popular and indeed an inducement.

3.8 Instruments

The basic data collecting instrument was a web survey created with the assistance of the University of Tennessee (UT) Statistical Consulting Center (SCC) whose mission is to “*help UT students, faculty, and staff enhance the quality of their research by working together to effectively apply analytical methods, especially statistics.*” [69] The building

tool for the web survey was the “SPSS DimensionNet Interview Builder” which gathered response data and offered an option of output in an Excel comma-separated-values (CSV) data file. Interview Builder also allows for creating and testing the survey prior to final distribution of the web link. Figure 27 shows a graphic of the web survey under construction (in build mode) using the Interview Builder with introduction questions. The web survey was beta tested by 15 volunteers with an MPC&A background. During the beta test, questions were asked and an opportunity was given for feedback and to suggest additional components for the various subsystems. The beta test assured more comprehensive subsystems and ultimately produced a higher quality survey instrument. The questions in the final instrument were developed from the responses of the beta testers and through consultation with the SCC.

The final complete survey tool is presented in Appendix I, however, a sample question for the rating of coefficients can be seen in Figure 28 as the respondent would see it on the screen.

3.9 Procedures

The target window of data collection was the INMM 49th Annual Meeting in July 2008. The meeting was held in Nashville TN and readily accessible for travel from Knoxville. An email was sent during the meeting to inform the members and attendees of the existence of the survey. Support from the INMM executive committee was solicited to “advertise” the survey and encourage individuals to take the online survey. The survey link remained open for a 2-week period after the annual meeting. During the survey, the individuals were solicited for availability of cleansed data regarding the safeguards systems of which they were familiar. However, no sanitized data were available.

3.10 Survey Data Analysis

The CSV data file downloaded at the end of the survey period had a record of 102 responses. Each of the respondents’ u_{ij} values was reported on a scale from 0-100 with 0 being “low” and 100 being “high.” Every individual could have different total raw score; therefore, uniform scaling was needed for comparison purposes. One could: a) take the mean of the raw responses and then normalize; or b) normalize the respondents’ raw scores before taking the mean. In the case of a), there was no assurance that a 50 raw score from a given respondent meant the same as a raw score of 50 from any other respondent. So in this case in order to normalize the responses of an individual the value was expressed as a fraction of the total responses for the individual in that subsystem. Therefore, an individual’s responses are normalized according to the following equation.

$$\text{Normalized } u_{ij} = \frac{u_{ij}}{\sum_j u_{ij}}$$

Section A Plant Management

(A2) Please indicate the level of importance of the following items on a scale of 0 (low) to 100 (high) to a Plant Management Subsystem.

	Component Rank (0 - 100)					
Personnel Reliability						
Budgeting						
Knowledge of Material Protection, Control, and Accounting principals						
System Oversight (audits, surveillance, etc)						
System Documentation						
Configuration Management						
Independent Assessment						
Organizational Communication						
Organizational Coordination						
Personnel Training						

Figure 28 Question As The Respondent Would See It

An example of the calculation can be seen in Table 6 for an individual response on utility coefficients for their Plant Management importance responses. In this case the individual rated Personnel Reliability within the Plant Management subsystem (the Personnel Reliability PMRank) as having an importance of 50 out of 100. The total of all of the individuals' responses was 720 leading to a normalized Personnel Reliability PMRank of 0.0694.

These normalized values were loaded into a statistical calculation program (JMP) which is a product in the SAS family of statistical software. The graphical distributions and tabular moments of the responses for each question are the results of processing the full set of responses in JMP and the results of two example coefficients are listed in Figure 29. In Figure 29 the straight red lines on the quantile plot and the curved red lines on the histogram indicate a fitted normal curve showing a visually good fit, however the normal quantile plot shows a good deal of skewness and the Shapiro-Wilk W Test (for less than 2000 data points) indicates the data are not normally distributed; however, the values are close to the mean and for the purposes of this methodology will suffice.

Table 6 Example Calculation Of Individual Normalized Coefficients

Item	Personnel Reliability PMRank	Budgeting PMRank	Knowledge of MPC&A PMRank	System Oversight PMRank	System Documentation PMRank	Configuration Management PMRank	Independent Assessment PMRank	Org Communication PMRank	Org Coordination PMRank	Personnel Training PMRank	Sum
Raw Score	50	60	80	90	90	90	80	70	60	50	720
Normalized	0.0694	0.0833	0.1111	0.1250	0.1250	0.1250	0.1111	0.0972	0.0833	0.0694	1.0000

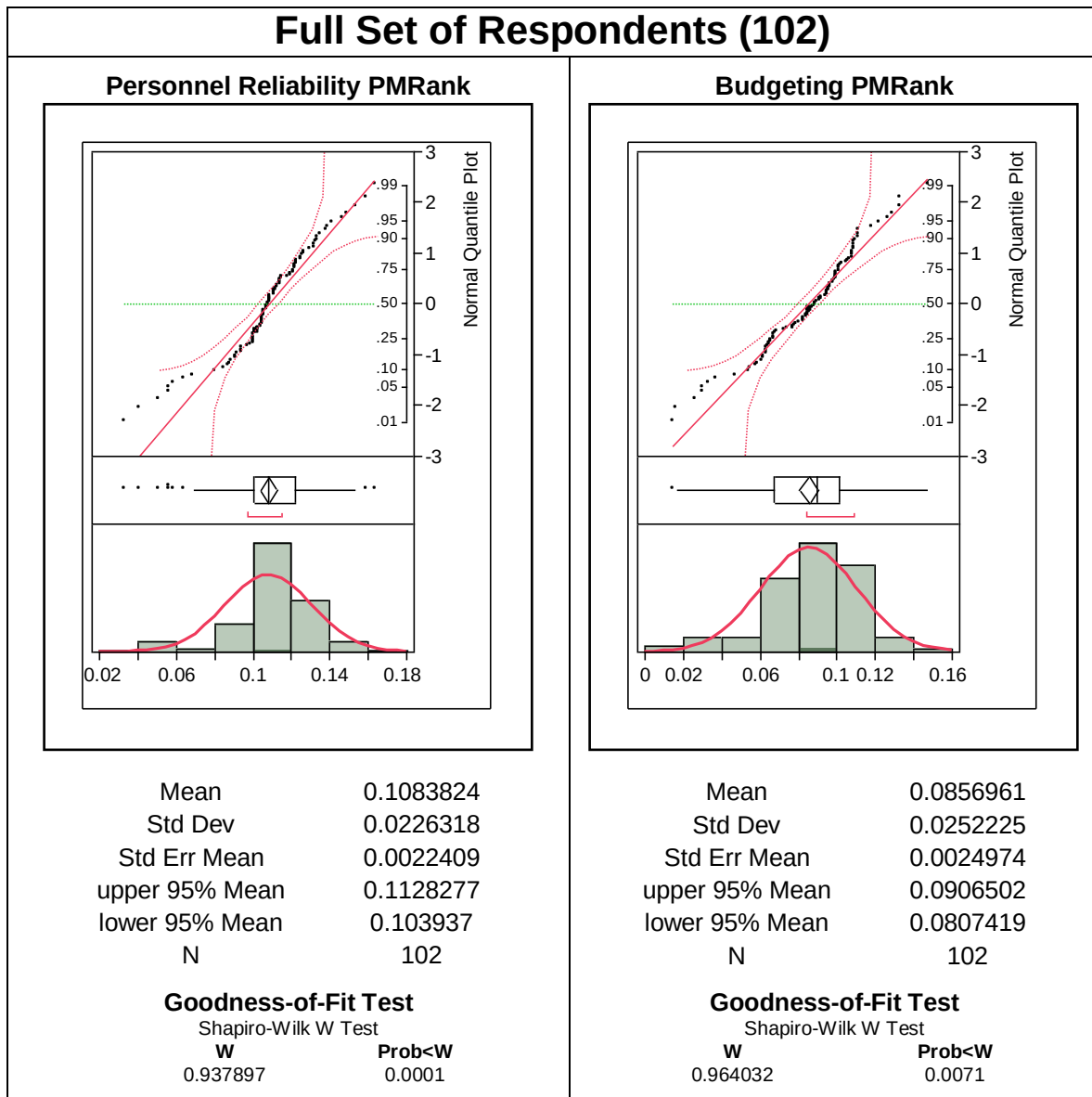


Figure 29 Example Plant Management Component Rankings

3.11 System Simulations

Simulation refers to methods and applications that mimic the behavior of real systems. [70] With the advancement in computing power available computer simulation is more popular and powerful. Computer simulation refers to methods for studying a wide variety of models using numerical evaluation using software designed to imitate the system's operations or characteristics. [70] Concurrent with the survey development, data gathering and analysis, simulations with appropriate subsystem components, were developed using a commercially available simulation software package (Arena 7.0). Historical error rate data was not available for analysis and validation so the simulations provide estimates of Type I and Type II error rates. There were two simulations created: 1) an airport passenger screening system used to screen airline passengers as shown in Figure 23, provided by the Transportation Security Administration (TSA); and 2) a simple domestic safeguards system modified from earlier work by Coates. [71] In both simulations the system components can be modeled for error production and the results used to calculate an overall effectiveness.

In the airport passenger screening simulation the elements were the identification screening system, a combined X-ray/explosives system for checking bags and computers, and the walk through metal detector.

A flow chart for this system to be simulated is shown in Figure 30. At each decision point in the flow chart a Type I or Type II error can occur. Figure 30 has a number of decision points and error rates built into the simulation such as: "Determine if ID is Acceptable" has a 99% OK rate; "Check Computer" has a 99% OK rate; "Check Bag" has a 99% OK rate; "Check Passenger for Metal" has a 99% OK rate; and "Hand Wand Positives" has a 99% OK rate. After each decision there are assumed Type I and Type II error rates (α and β rates = 1%), however, the fact that the decision rates are 99% and the error rates are 1% are only for convenience and one does not determine the other. The airport simulation provided simulated error rates for comparison purposes between the DOE probabilistic evaluation as adapted for the particular system.

The simple domestic safeguards system simulation modules at the top level can be seen in Figure 31. There are three processing lines: "Insider Accounting"; "Insider Control"; and "Insider Management" and the "Create Security Condition" line which does not have an insider. Each of the processing lines is constructed in a similar fashion and using the example of the Accounting Sub-model the simulation modules can be seen in Figure 32 and the Security Sub-model modules can be seen in Figure 33. Inherent to this simulation is the ability of the security condition to affect whether an

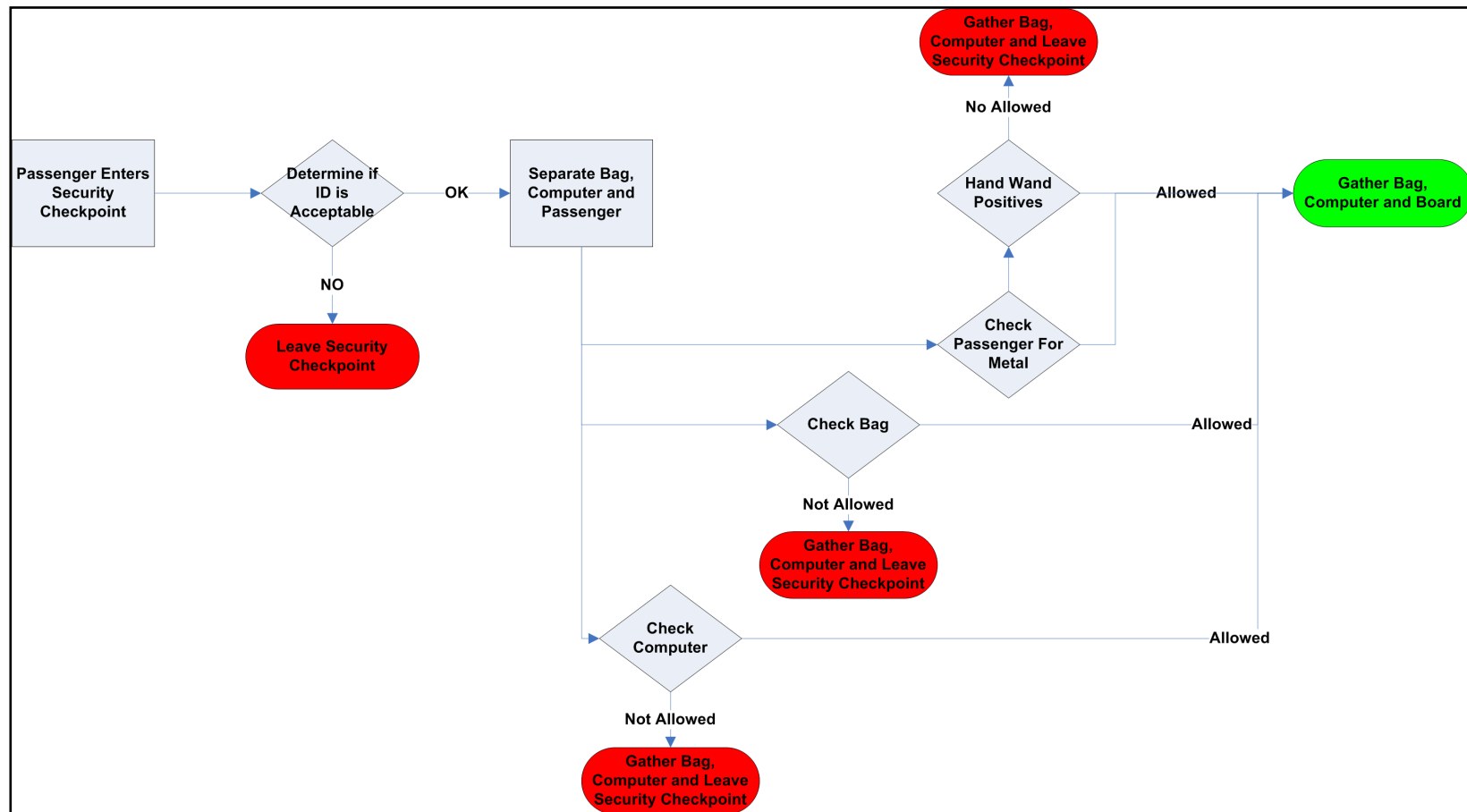


Figure 30 Airport Security Check Point Flow Chart

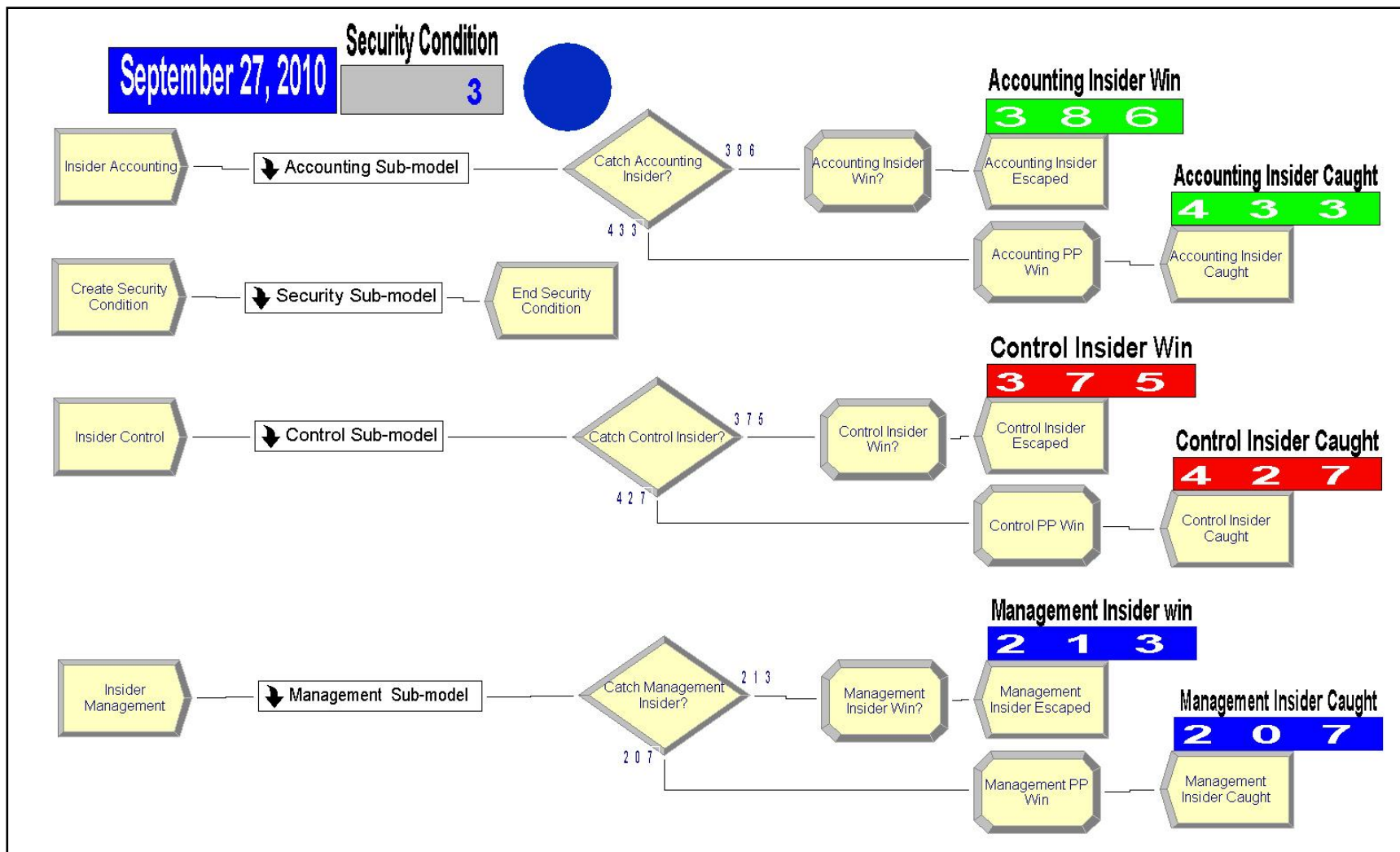


Figure 31 Simple Domestic Safeguards Simulation Modules

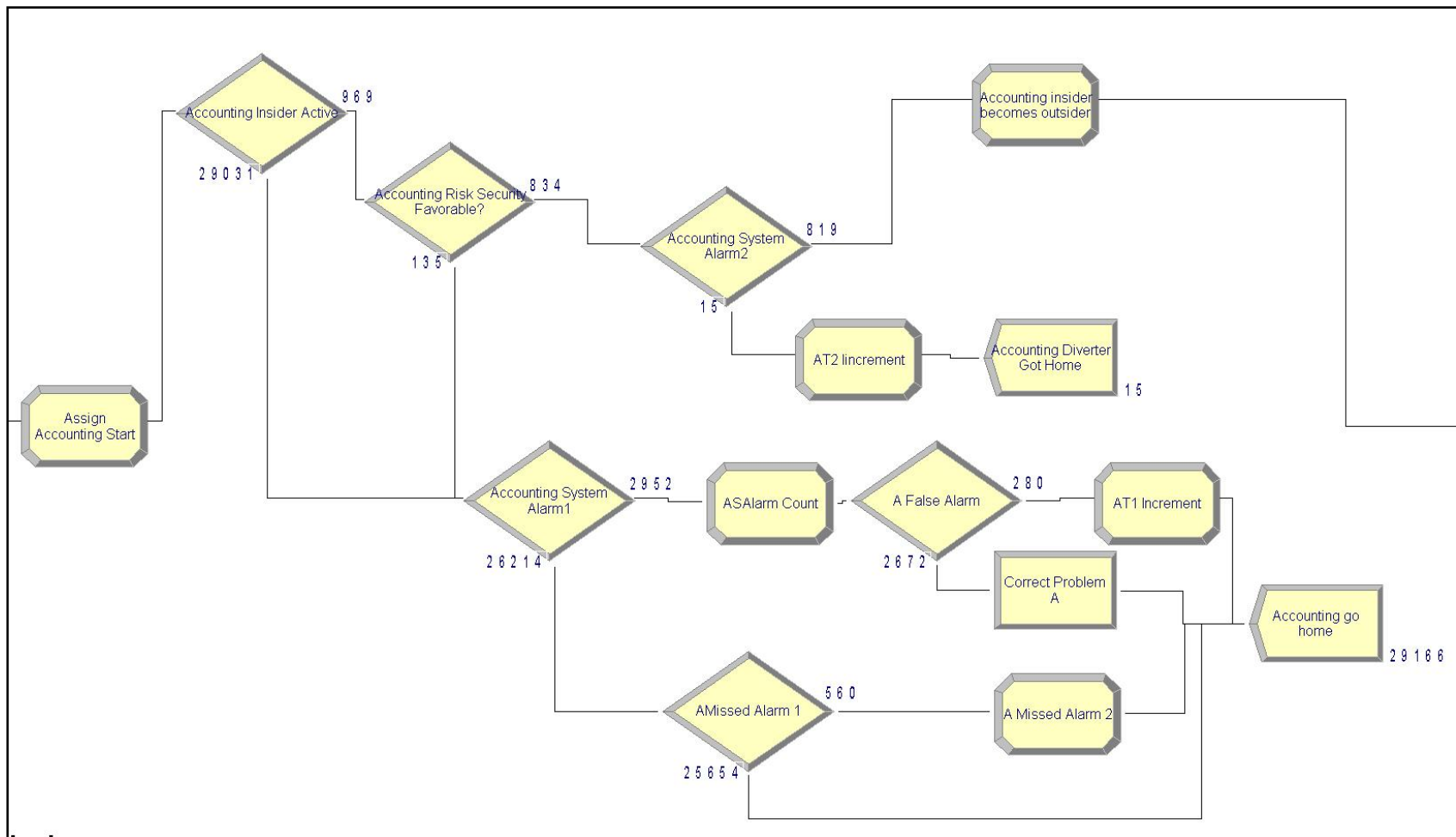


Figure 32 Accounting Sub-model Modules

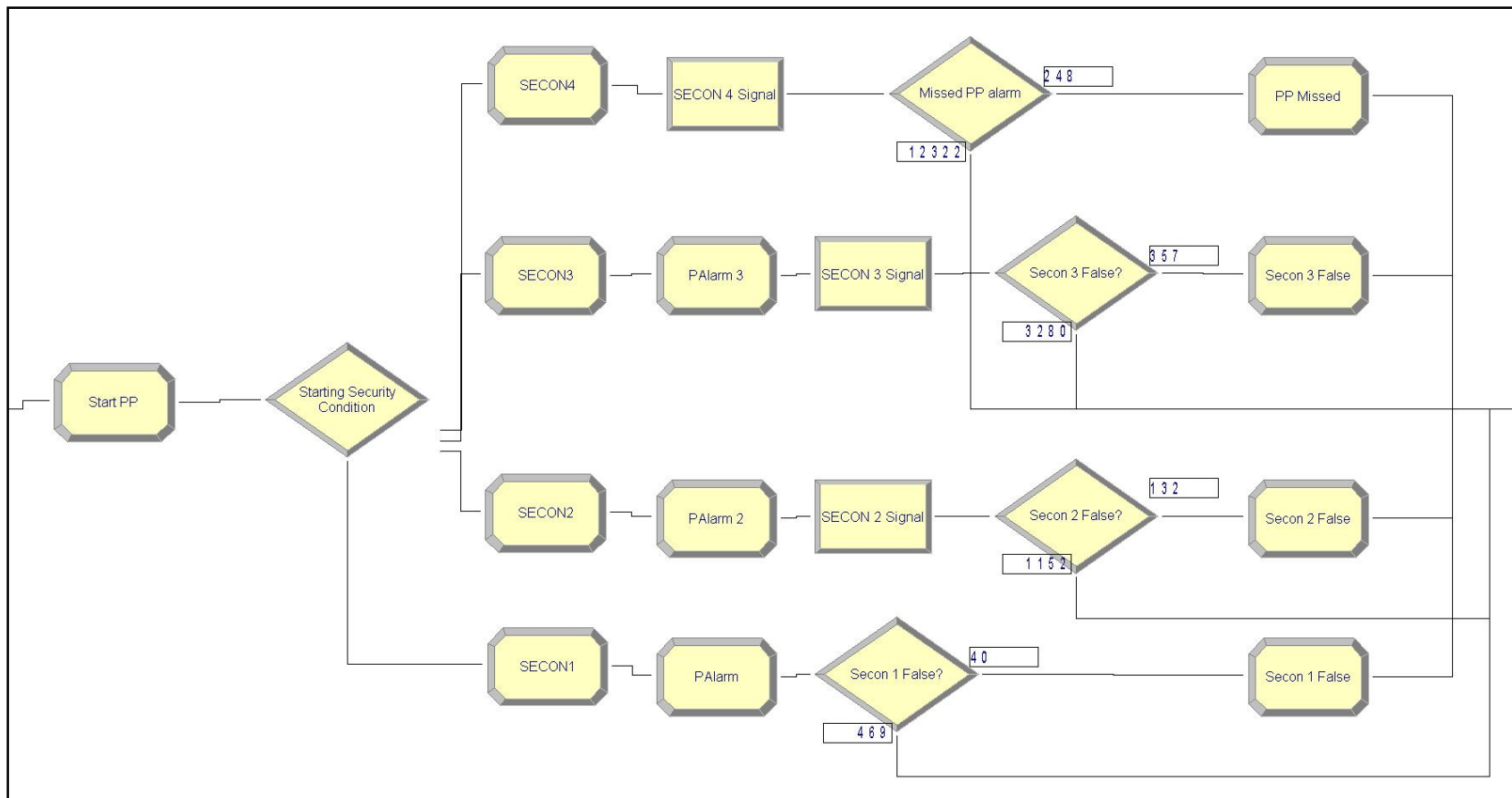


Figure 33 Security Sub-model Modules

insider is active or inactive. The probabilities of the processing lines decisions for insider activity, alarm frequency and error rates can be seen in Figure 34. The Security line sets the alarm level or security condition (SECON). There are four security conditions: normal (SECON 4); slightly elevated (SECON 3); significantly elevated (SECON 2) and high risk (SECON 1). The insider is active only when the security condition is either SECON 3 or SECON 4. The probabilities of the security condition alarms can be seen in Figure 35. These theoretical values are programmed into the simulation and used to calculate the theoretical system effectiveness for a test of convergence on the simulated values.

3.12 Methodology Summary

An additive MAUT approach has been chosen as a valid methodology for applying the posited system effectiveness function for determining effectiveness based on Type I and Type II error rates with the conditions for this calculation comparable to the validity claim of Turban and Metersky [63]. The application of Turban and Metersky was to determine the best design alternative solution of several designs a system. This research advances that approach through the determination of effectiveness of an individual solution over time. This research determines the utility of components of a simple domestic safeguards system through a statistically analyzed web survey of 102 nuclear material experts. Further the validity of the function will be determined through comparison of the results using the probabilistic methodology validated by the DOE in their system performance effectiveness equation. Simulation of the Type I and Type II error rates will be used to compare the two methodologies and calculate a performance level of individual subsystems and components so that the overall system effectiveness can be determined.

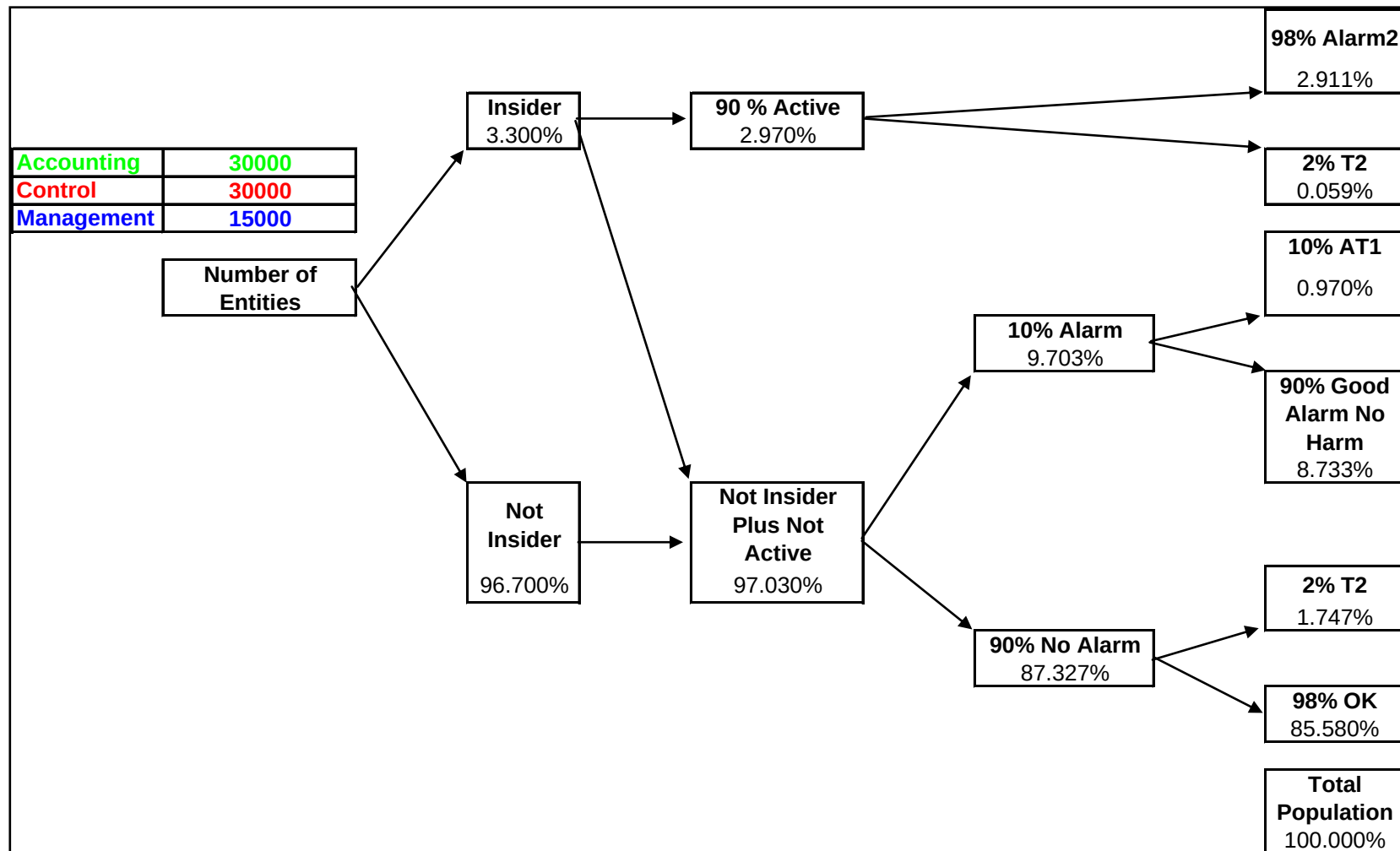


Figure 34 Decisions Probabilities and Error Rates for Process Lines

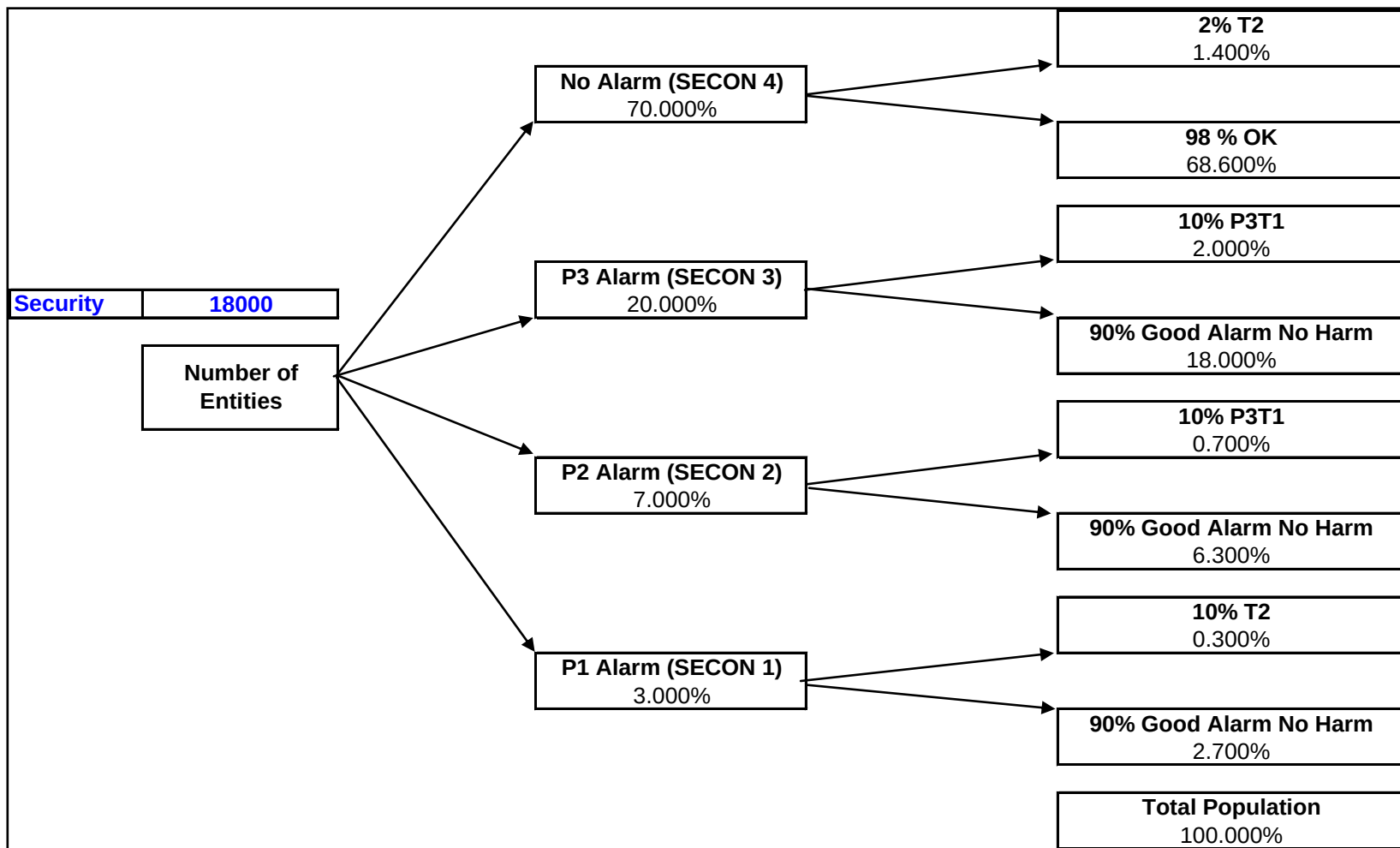


Figure 35 Decision Probabilities and Error Rates for Security Line

Chapter 4 Analysis

Chapter 4 will take the results of the research and analyze their meaning by starting with the survey and simulation data. Next this chapter will validate the System Effectiveness Methodology using an Airport Simulation and a Safeguards Simulation.

4.1 Survey Data

In analyzing the responses, the mean for each component's utility was chosen for inclusion in the overall effectiveness calculation. An illustration of how the data were organized can be seen in Table 7 in row B (with numbers in blue). Row C (numbers in red). In row A of Table 7, labeled "Average Value From Full Survey" the mean utility values from the survey with all the components are shown (however not all columns are shown in Figure 36 which illustrates how the full model would be constructed). When these utilities are summed across the line they are not constrained to sum to 1 (row A sums to 1.0001 in Table 7 and other rows in Figure 36 sum to 0.9999, 0.9998, 1.0003 etc) but the methodology requires that the mean utilities sum to one, therefore the values from Row A of Table 7 are (as seen in row B with numbers in blue). Row C (numbers in red) is generated by a random number generator in Microsoft Excel between 0 and 1 and represents a hypothetical effectiveness of the given component. The subsystem effectiveness represents the sum of the products of the component effectiveness times the normalized mean utility for that component, as illustrated in row D of Table 7 leading to a subsystem effectiveness of 0.7241.

The normalized mean utilities received from the web-based survey have been tabulated in Table 8. Table 8 contains both responses from the full data set (all 102 responses seen in Column A) and columns B-E show the responses for the filtered data. The filtered data set includes the responses for respondents that have greater than 5 years of experience in their field. A respondent with greater than 5 years experience in a field was deemed to have greater expertise in the area. Therefore the Plant Management (PM) filtered column (column B) has responses only from those with greater than 5 years of experience in Plant Management; the Material Control (MC) column (C) has responses from only those with greater than 5 years of experience in Material Control and so forth for Material Accounting (MA) and Physical Protection (PP). Totaling the number of respondents of the filtered columns indicates that some individuals had greater than 5 years of experience in more than one area. Therefore in the first row of mean coefficients (Personnel Training PMrank) the filtered mean value (from those with > 5 years of experience in plant management) was 0.36% less than the mean value of the full set of respondents (which can be seen in column F).

In many cases the difference between the full data set and the filtered data set had very little difference and well below the error rate (95% confidence interval of +/- 9.1%) for the survey. Those coefficients that showed differences greater than the error rate of the

Table 7 Example Subsystem Effectiveness Calculation

Row	Item	Personnel Reliability PMRank	Budgeting PMRank	Knowledge of MPC&A PMRank	System Oversight PMRank	System Documentation PMRank	Configuration Management PMRank	Independent Assessment PMRank	Org Communication PMRank	Org Coordination PMRank	Personnel Training PMRank	Sum
A	Average Value From Full Survey	0.1097	0.1090	0.1084	0.1075	0.1028	0.0983	0.0959	0.0940	0.0889	0.0857	1.0001
B	Normalized to Sum to 1.0	0.1097	0.1089	0.1084	0.1075	0.1028	0.0983	0.0959	0.0940	0.0888	0.0857	1.0000
C	Full Survey Random Hypothetical (0.5 to 1.0) Element Effectiveness Value	0.5317	0.8034	0.7671	0.7277	0.6490	0.7627	0.8647	0.6012	0.8026	0.7524	
D	(Effectiveness)*(Normalized utility)	0.0583	0.0875	0.0831	0.0782	0.0667	0.0750	0.0829	0.0565	0.0713	0.0645	0.7241

Plant Management Elements		Personnel Training PMRank	Org Communicatio in PMRank	Independent Assessment PMRank	Budgeting PMRank										
Average Values From Full Survey		0.1097	0.1075	0.0889	0.0857										
Normalized to Sum to 1.0		0.1097	0.1075	0.0888	0.0857										
Full Survey Random Hypothetical (0.5 to 1.0) Element Effectiveness Value		0.5317	0.7277	0.8026	0.7524										
Average Values From Filtered Survey PM Experience >5		0.1093	0.1141	0.0848	0.0949										
Normalized to Sum to 1.0		0.1093	0.1141	0.0848	0.0949										
PM Filtered Survey Random Hypothetical (0.5 to 1.0) Element Effectiveness Value		0.5317	0.7277	0.8026	0.7524										

Figure 36 Full Utility Model Calculations with Random Performance Levels

Table 8 Mean Response Tabulation

	Full	Filter				
Safeguards Subsystems and Components		PM	MC	MA	PP	Delta
Basis	102	15	36	39	23	11
						(Filter-Full) / Full
Plant Management Component Rank						
Personnel Training PMrank	0.1097	0.1093				-0.36%
Knowledge of MPC&A PMrank	0.1089	0.0989				-9.19%
Personnel Reliability PMrank	0.1084	0.0995				-8.15%
Org Communication PMrank	0.1075	0.1141				6.16%
Org Coordination PMrank	0.1028	0.1016				-1.18%
System Documentation PMrank	0.0983	0.1006				2.34%
Configuration Management PMrank	0.0959	0.1009				5.28%
System Oversight PMrank	0.0940	0.0953				1.40%
Independent Assessment PMrank	0.0888	0.0848				-4.54%
Budgeting PMrank	0.0857	0.0949				10.72%
Physical Protection Component Rank						
Alarm Response PPrank	0.0633				0.0623	-1.52%
Target ID PPrank	0.0623				0.0654	4.91%
System Objectives PPrank	0.0615				0.0615	-0.03%
Threat Definition PPrank	0.0612				0.0615	0.54%
Alarm Assessment PPrank	0.0605				0.0618	2.19%
Personnel Reliability PPrank	0.0602				0.0582	-3.29%
Vulnerability Assessment PPrank	0.0598				0.0608	1.74%
Facility Characterization PPrank	0.0591				0.0620	4.85%
Material Control System PPrank	0.0591				0.0590	-0.22%
Access Delay PPrank	0.0585				0.0566	-3.22%
Secure Comm PPrank	0.0585				0.0572	-2.19%
PP System Testing PPrank	0.0575				0.0589	2.34%
Exterior Intrusion Sensors PPrank	0.0573				0.0572	-0.12%
Special Response Teams PPrank	0.0568				0.0567	-0.31%
Internal Guard Force PPrank	0.0564				0.0556	-1.41%
Interior Intrusion Sensors PPrank	0.0554				0.0555	0.15%
Independent System Assessment PPrank	0.0524				0.0497	-5.28%
Material Control Component Rank						
MC Containment MCrank	0.0770		0.0754			-2.04%
Entry Access Control MCrank	0.0764		0.0786			2.85%
MC Item Tracking MCrank	0.0751		0.0745			-0.78%
MC Surveillance MCrank	0.0743		0.0729			-1.85%
	A	B	C	D	E	F

Table 8 (Cont.) Mean Response Table

	Full	Filter				
Safeguards Subsystems and Components		PM	MC	MA	PP	Delta
Basis	102	15	36	39	23	11
						(Filter-Full) / Full
MC PP MCrank	0.0733		0.0762			3.99%
MC Process Control MCrank	0.0730		0.0732			0.38%
MC Storage Monitoring MCrank	0.0723		0.0747			3.36%
MC Procedures MCrank	0.0717		0.0745			3.90%
Nuclear Material Portal Monitors MCrank	0.0713		0.0735			3.04%
Tamper Indicating MCrank	0.0702		0.0636			-9.40%
MC Waste Monitoring MCrank	0.0696		0.0705			1.24%
MC Attribute monitors MCrank	0.0679		0.0652			-3.96%
MC Oversight MCrank	0.0665		0.0665			-0.04%
MC Daily Admin Checks MCrank	0.0614		0.0606			-1.29%
Material Accounting Component Rank						
Accounting System MArank	0.0747			0.0768		2.81%
Measurement Accuracy MArank	0.0717			0.0739		3.10%
Physical Inventory MArank	0.0715			0.0743		3.96%
Inventory Detection and Resolution MArank	0.0705			0.0720		2.18%
Shipper Receiver Difference MArank	0.0704			0.0684		-2.80%
Material Transfer Monitoring MArank	0.0700			0.0699		-0.22%
Anomaly Detection MArank	0.0695			0.0695		-0.02%
Measurement Control MArank	0.0684			0.0691		1.04%
MA Statistical Evaluation MArank	0.0650			0.0670		3.17%
MA Process Monitoring MArank	0.0647			0.0619		-4.46%
MA Procedures MArank	0.0640			0.0640		0.01%
Real Time Accounting MArank	0.0625			0.0582		-6.95%
MA Holdup MArank	0.0609			0.0617		1.33%
MA Independent Assessment MArank	0.0582			0.0573		-1.52%
Oversight MArank	0.0580			0.0560		-3.48%
Safeguards Rank						
Plant Management	0.2241	0.2095				-6.53%
Physical Protection	0.2600				0.2578	-0.83%
Material Control	0.2602		0.2666			2.49%
Material Accounting	0.2557			0.2660		4.04%
	A	B	C	D	E	F

survey included: Knowledge of MPC&A PMrank; Budgeting PMrank; and Tamper Indicating MCrank. These coefficients are highlighted in yellow in Table 8. The coefficients from the unfiltered and filtered data sets are used in an example overall system effectiveness calculation and the results are shown in Figure 36.

The calculations in Figure 36 show two results for each subsystem: 1) using the utilities from the full data set; and 2) the utilities using the filtered data for each area of expertise. The method of calculation in the last three columns of Figure 36 are illustrated in Table 9. Each of the subsystem effectiveness calculations (shown in Figure 36 and column A of Table 9) are combined by multiplying the subsystem's utility using both the raw subsystem utility from the survey and those normalized using the family of systems effectiveness equation (Table 9 column B and C respectively) times the subsystem's calculated effectiveness (column A). This is done using both the full survey data and the filtered data while using the same randomly generated component effectiveness values. The results do not show significant overall effectiveness differences from the two sets of calculations (76.76% with the full set utilities and 76.91% for the filtered data utilities). Further, these data do not account for is any component that could have an overriding affect on the system effectiveness as this was not part of the survey.

4.2 Simulation Data

A graphic of the airport passenger screening simulation model can be seen in Figure 37. This model was created using Arena 7.0. The simulation model generated Type I and Type II errors based on predetermined percentages as noted in Chapter 3. In this case each decision was set to 99% true (99% effective) and the rate of each type of error was set at 1 percent. Daniel Morgan, of the Congressional Research Service of The Library of Congress, reported that screeners failed to detect items 20% of the time in 1987. [72].

Current error rates are likely to be closely held, therefore, for this simulation an assumed effectiveness number of 99% was chosen for illustration purposes. Table 10 shows the values of the output variables from the simulation as seen in Figure 37 with three runs for each for 1000, 10,000 and 100,000 travelers. Table 10 column A indicates the name of the variable in alphabetical order with column B indicating the notation used in the output files. In the columns D, E, and F, there are sub cells with values that indicate individual runs which are identified in column C. In the case on Identification Denied," travelers were identified as having faulty identification (failed the ID screening) and did not proceed through the rest of screening. Of the travelers that passed the ID screening, those that had either "Bag Denied," "Computer Denied," or "Metal Denied" were denied boarding. In the case of run 1 with 100,000 travelers there were: 1006 that failed the ID screening; 932 with "Denied Bag", 999 with "Computer Denied" and 10 with "Metal Denied" for a total of 1941 denied states. An example of the simulation output can be seen in Figure 38 which was replicated 3 times with 1000 travelers.

Table 9 Example Overall Effectiveness Calculation

	Hypothetical Line Effectiveness Value	Factor from Survey	Normalized To Sum to 1.0
Full Survey Plant Management Elements	0.7241	0.2241	0.2241
Filtered Survey Plant Management Elements	0.7237	0.2067	0.2095
Full Survey Plant Management Elements	0.7114	0.2600	0.2600
Filtered Survey Plant Management Elements	0.7136	0.2632	0.2666
Full Survey Material Accounting Elements	0.8440	0.2601	0.2602
Filtered Survey Material Accounting Elements	0.8438	0.2626	0.2660
Full Survey Physical Protection Elements	0.7851	0.2557	0.2557
Filtered Survey Physical Protection Elements	0.7864	0.2545	0.2578
Full Survey Total System Effectiveness	0.7676	$\Sigma = 0.9999$	$\Sigma = 1.0000$
Filtered Survey Total System Effectiveness	0.7691 A	$\Sigma = 0.9869$ B	$\Sigma = 1.0000$ C

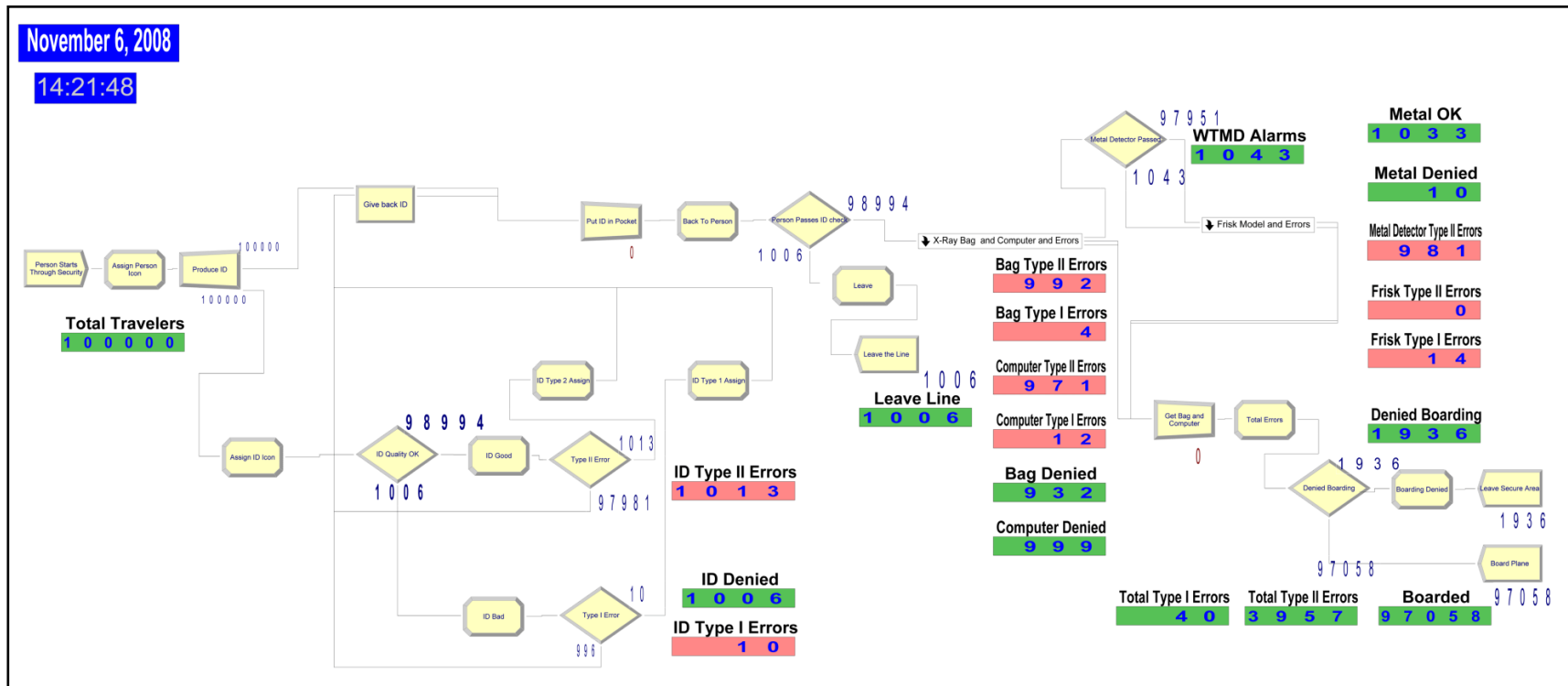


Figure 37 Simulation of Airport Passenger Screening

Table 10 Output Data from Simulation

Output	Notation	Run	1000 Travelers	10,000 Travelers	100,000 Travelers
Bag Denied	BD	1	11	112	932
		2	12	86	972
		3	10	116	1026
Bag Type I Error	BTIE	1	0	1	4
		2	0	0	9
		3	0	0	7
Bag Type II Error	BTIIE	1	5	88	992
		2	8	90	1022
		3	5	85	959
Computer Denied	CD	1	16	106	999
		2	12	110	1017
		3	17	107	1019
Count Total	Count Total	1	1000	10000	100000
		2	1000	10000	100000
		3	1000	10000	100000
Computer Type I Error	CTIE	1	1	2	12
		2	0	2	11
		3	0	3	12
Computer Type II Error	CTIIE	1	9	87	971
		2	6	89	926
		3	13	113	968
Denied Boarding from Computer, Bag, or Metal	DB	1	27	218	1936
		2	24	195	1991
		3	26	222	2040
Denied	Denied	1	1	1	2
		2	1	2	2
		3	2	2	2
Frisk Type I Error	FTIE	1	0	1	14
		2	0	0	3
		3	0	1	6
A	B	C	D	E	F

Table 10 (cont.) Output Data from Simulation

Output	Notation	Run	1000 Travelers	10,000 Travelers	100,000 Travelers
Frisk Type II Error	FTIIE	1	0	0	0
		2	0	0	0
		3	0	0	0
Identification Denied	IDD	1	8	87	1006
		2	7	87	1014
		3	8	94	963
Identification Type I Error	IDTIE		1	2	10
			1	2	9
			0	2	7
Identification Type II Error	IDTIIIE	1	10	93	1013
		2	7	100	1028
		3	9	78	981
Walk Through Metal Detector Type II Error	MDIIE	1	9	79	981
		2	12	100	989
		3	7	101	995
Metal OK	MDOK	1	6	117	1033
		2	12	84	987
		3	6	101	1012
Metal Denied	MetDen	1	0	0	10
		2	0	0	11
		3	0	0	9
Total Type I Errors	Total TI	1	2	6	40
		2	1	4	32
		3	0	6	32
Total Type II Errors	Total TII	1	33	347	3957
		2	33	379	3965
		3	34	377	3903
Walk Through Metal Detector Alarm	WTMD Alarm	1	6	117	1043
		2	12	84	998
		3	6	101	1021
Walk Through Metal Detector Passed	WTMD passed	1	986	9796	97951
		2	981	9829	97988
		3	986	9805	9806
A	B	C	D	E	F

Airport Security

Replications: 3

Replication 1

Start Time:

0.00

Stop Time:

9,981.81

Time Units: Minutes

Time Persistent

Variable	Average	Half Width	Minimum	Maximum
BD	479.28	(Correlated)	0	932.00
BTIE	1.9820	(Insufficient)	0	4.0000
BTIIE	481.69	(Correlated)	0	992.00
CD	511.85	(Correlated)	0	999.00
Count Total	49,851.61	(Correlated)	0	100,000.00
CTIE	6.4540	(Insufficient)	0	12.0000
CTIIE	468.28	(Correlated)	0	971.00
DB	992.63	(Correlated)	0	1,936.00
Denied	0.01977973	0.001164391	0	2.0000
FTIE	5.9167	(Insufficient)	0	14.0000
FTIIE	0	(Insufficient)	0	0
IDD	504.18	(Correlated)	0	1,006.00
IDTIE	5.8675	(Insufficient)	0	10.0000
IDTIIE	487.55	(Correlated)	0	1,013.00
MDIIE	482.79	(Correlated)	0	981.00
MDOK	523.53	(Correlated)	0	1,033.00
MetDen	3.4042	(Insufficient)	0	10.0000
Total TI	20.2200	(Insufficient)	0	40.0000
Total TII	1,920.30	(Correlated)	0	3,957.00
WTMD Alarm	526.93	(Correlated)	0	1,043.00
WTMD passed	48,820.50	(Correlated)	0	97,951.00

Figure 38 Passenger Screening Simulation Output for 100,000 Travelers

4.3 Validating the Methodology With Airport Simulation

The utility theory of combining individual components of a decision into a single overall decision has been validated through the work of Fishburn [59] [60] as was specifically detailed in Chapter 2. In the case of systems effectiveness, each component of the decision becomes the effectiveness of the system component. Claims for validity of utility theory in calculating system effectiveness are made by Turban and Metersky [63]. Turban and Metersky use the technique to evaluate the design alternatives that maximize the system effectiveness based on a specific utility for each measure of effectiveness. Different design alternatives are evaluated and Turban and Metersky assume there is no overwhelmingly important factor in the effectiveness calculation (no one factor >30% of the total utility which equals 1) and their assumption is born out in this research where individual utilities for components ranged from 5% to 11% of the subsystem's total utility and no subsystem had greater than 26% of the total system utility (each of the four subsystems ranged from 22% to 26%).

This work extends Turban and Metersky's work by applying their methodology to a near-real-time decision of a working system's effectiveness. What remains is a comparison of the values using the system effectiveness calculation to a recognized approach. A probabilistic approach for calculating system effectiveness has been validated by the Department of Energy in their system performance effectiveness equation [42].

$$P_{EL} = P_{IL} \times P_{NL} = P_{DL} \times P_{NL}$$

Where:

P_{EL} = the system effectiveness contribution for layer L;

P_{IL} = Probability of Interruption given first detection at layer L, $P_{IL} = P_{DL}$ if detection on layer L is timely, and is equal to 0 ($P_{IL} = 0$) if detection is not timely;

P_{NL} = Probability of Neutralization given first detection at layer L.

Adapting the equation to an airport passenger screening system requires the assumptions of: 1) the screening system is a single layer; 2) The detection is timely leading to $P_{IL} = P_{DL}$; and 3) $P_{NL} = 1$. For a passenger screening system these are reasonable assumptions. This leads to a simplified equation for this one layer system or $P_E = P_D$. The calculation of the $P_E = e_s = (\text{system alarms}) / (\text{system alarms} + \text{missed alarms})$. The data from the airport simulation are used to compare the two methodologies. In Table 11 three quantities are used in the calculation: A) System Alarms (shaded yellow) or alarms the system made during the simulation run; B) Missed Alarms (shaded tan) based on the simulation parameters and; C) False Alarms (shaded green) also identified by the simulation parameters. The one layer assumption also leads to a simplified effectiveness calculation using the posited function as all System Alarms, Missed Alarms, and False Alarms are used in the one layer which becomes one system.

Table 11 Comparison of Effectiveness Calculation with Probabilistic Method

Output Is The Average from 3 Simulation Runs of 100,000 Travelers Each	
Bag Denied	976.7
Bag Type I Error (False Alarm)	6.7
Bag Type II Error (Missed Alarm)	991.0
Computer Denied	1011.7
Computer Type I Error (False Alarm)	11.7
Computer Type II Error (Missed Alarm)	955.0
Frisk Type I Error (False Alarm)	7.7
Frisk Type II Error (Missed Alarm)	0.0
Identification Denied	994.3
Identification Type I Error (False Alarm)	8.7
Identification Type II Error (Missed Alarm)	1007.3
Walk Through Metal Detector Type II Error	988.3
Metal OK	1010.7
Metal Denied	10.0
Walk Through Metal Detector Alarm	1020.7
System Alarms	4013.3
Missed Alarms	3941.7
False Alarms	1045.3
α = False Alarm Rate = False Alarms / System Alarms	0.260
β = Missed Alarm Rate = Missed Alarms / (Missed Alarms + System Alarms)	0.495
Overall Effectiveness Calculation With False Alarms of Equal Importance	37.3%
Overall Effectiveness Calculation With False Alarms of Zero Importance	50.5%
Overall Effectiveness Calculation With $r = 2$ (False Alarms/Missed Alarms under no stress)	47.0%
Overall Effectiveness Calculation With $r = 5$ (False Alarms/Missed Alarms under no stress)	50.4%
<i>DOE system performance effectiveness equation reduces to $P_E = P_D \approx \text{System Alarms} / (\text{System Alarms} + \text{Missed Alarms})$</i>	50.5%

As seen in Table 11 the probabilistic model (*system performance effectiveness equation*) used by the DOE, calculates the system effectiveness as 50.5% while the effectiveness function posited in this research showing equal base tolerance of false and missed alarms ($r = 1$) the effectiveness is calculated to be 37.3%. When r is increased to 2 (base tolerance of two false alarms to 1 missed alarm) the effectiveness is calculated as 47.0% and when r is raised to 5, the effectiveness is calculated as 50.4% and in fact converges on the DOE value.

4.4 Validating the Methodology With Safeguards Simulation

In order to tie this research together, a simple domestic safeguards system was simulated. Figure 31 through Figure 35 detail the design parameters of this simulation. The assumptions are that the four sub-systems are independent with relation to effectiveness. The simulation does have a component of when the insider is only active based on SECON level but these do not affect the error rates. Recalling that we:

Let:

$$e_s = u_p e_p + u_c e_c + u_a e_a + u_m e_m$$

s.t.

$$0 \leq e_i \leq 1 \quad \sum u_i = 1$$

Where:

- u_p = the utility the physical protection subsystem
- u_c = the utility the material control subsystem
- u_a = the utility the material accounting subsystem
- u_m = the utility the management subsystem

The utilities (u_i) were determined through the survey of experts and leads to the equation used in the effectiveness calculation for this simulation.

$$e_s = 0.2578e_p + 0.2666e_c + 0.2660e_a + 0.2905e_m$$

The simulation will only cover the first layer of subsystems and the utilities of subsystem elements will not be used for simplicity. The simulation will produce error rates based on the theoretical rates from Figure 34 and Figure 35. These theoretical rates calculated on the count basis, as seen in Table 12, will be used to calculate both the theoretical effectiveness using the posited function and the DOE system performance effectiveness equation for comparison. The simulated data for the longest runs can be seen in Table 12.

Five repetitions under each condition were conducted and data saved to an excel file and are listed in Table 13. The conditions were proportional to the Accounting entities and were 300, 3000 or 30000 Accounting entities. The results of the function calculations verses both the number of entities and the theoretical value for both the function and the DOE methodology can be seen in Figure 39 and Figure 40. As can be

Table 12 Theoretical Alarm Rates for Safeguards Simulation

	A	C	M	P
Count	30000	30000	15000	18000
SAlarm2	873.18	873.18	436.59	
AMissed	541.78	541.78	270.89	
SAlarm1	2911	2911	1455	5400
AT1	291	291	146	540
AT2	18	18	9	252
Total System Alarms	3,784	3,784	1,892	5,400
Total T1	291	291	146	540
Total T2	560	560	280	252

Table 13 Simulated Error Rates

Entities = Count				November 2, 2008	
Variable	REP 1	REP 2	REP 3	REP 4	REP 5
A Count	30,000	30,000	30,000	30,000	30,000
AAlarm2	809	808	810	873	819
AMissed	529	573	525	548	560
ASAlarm1	2,951	2,901	2,892	2,817	2,952
AT1	307	311	272	309	280
AT2	16	15	19	18	15
Total A System Alarms	3,760	3,709	3,702	3,690	3,771
Total A T1	307	311	272	309	280
Total A T2	545	588	544	566	575
C Count	30,000	30,000	30,000	30,000	30,000
CAlarm2	746	820	781	815	802
CMissed	491	522	492	507	504
CSAlarm1	2,829	2,802	3,004	2,939	2,931
CT1	285	267	317	301	286
CT2	13	16	14	10	18
Total C System Alarms	3,575	3,622	3,785	3,754	3,733
Total C T1	285	267	317	301	286
Total C T2	504	538	506	517	522
M Count	15,000	15,000	15,000	15,000	15,000
MAlarm2	449	416	437	451	420
MMissed	270	249	264	282	258
MSAlarm1	1,397	1,455	1,488	1,405	1,412
MT1	144	156	148	110	136
MT2	8	12	8	9	8
Total M System Alarms	1,846	1,871	1,925	1,856	1,832
Total M T1	144	156	148	110	136
Total M T2	278	261	272	291	266
P Count	18,000	18,000	18,000	18,000	18,000
PAlarm1	561	535	538	518	509
PAlarm2	1,255	1,274	1,272	1,242	1,284
PAlarm3	3,470	3,592	3,464	3,665	3,637
PP1T1	59	56	60	51	40
PP2T1	147	130	145	124	132
PP3T1	353	392	350	375	357
PPT2	242	232	236	241	248
Total P System Alarms	5,286	5,401	5,274	5,425	5,430
Total P T1	559	578	555	550	529
Total P T2	301	288	296	292	288

Run	Number of Travelers	ID Denied	IDTIE	IDTIE	Bag Denied	BTIE	BTIE	Computer Denied	CTIE	CTIE	WTMD Alarm	MDIE	MDOK	FTIE	FTIE	System Alarms	Missed Alarms	False Alarms	Simulation Effectiveness
1a	1000	8	1	10	11	0	5	16	1	9	6	9	6	0	0	41	33	8	0.4459
1b	1000	7	1	7	12	0	8	12	0	6	12	12	12	0	0	43	33	13	0.3947
1c	1000	8	0	9	10	1	5	17	0	13	6	7	6	0	0	41	34	7	0.4533
2a	10000	87	2	93	112	1	88	106	2	87	117	79	115	1	0	422	347	121	0.3914
2b	10000	87	2	100	86	0	90	110	2	89	84	100	82	0	0	367	379	86	0.3767
2c	10000	94	2	78	116	0	85	107	3	113	101	101	100	1	0	418	377	106	0.3925
3a	100000	1006	10	1013	932	4	992	999	12	971	1043	981	1030	14	0	3980	3957	1070	0.3666
3b	100000	1014	9	1028	972	9	1022	1017	11	926	998	989	983	3	0	4001	3965	1015	0.3748
3c	100000	963	7	981	1026	7	959	1019	12	968	1021	995	1006	6	0	4029	3903	1038	0.3771

Theoretical	100000	1000	10	990	990.1	9.901	980.199	990.1	9.901	980.199	990.1	980.199	990	9.901	9.901	3970.3	3940.498	1029.7	0.3717
-------------	--------	------	----	-----	-------	-------	---------	-------	-------	---------	-------	---------	-----	-------	-------	--------	----------	--------	--------

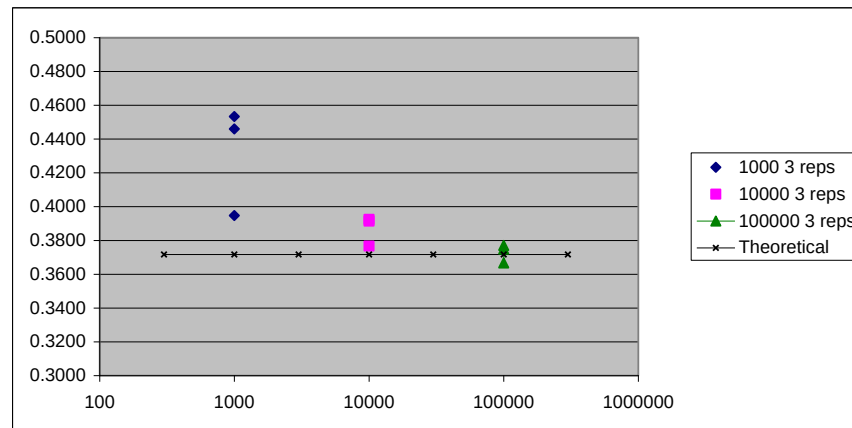


Figure 39 Convergence of Simulated Airport Model to Theoretical Value

Run	Number of A Entities	Number of C Entities	Number of M Entities	Number of P Entities	A Alarms	AT1	AT2	A e _s	C Alarms	CT1	CT2	C e _s	M Alarms	MT1	MT2	M e _s	P Alarms	PT1	PT2	P e _s	Simulation Effectiveness		
1a	300	300	150	180	40	1	3	0.9297	33	1	5	0.8678	22	2	5	0.8103	53	3	5	0.9113	0.8834		
1b	300	300	150	180	33	5	2	0.9236	36	0	5	0.8780	18	2	3	0.8494	55	9	5	0.8960	0.8887		
1c	300	300	150	180	38	1	8	0.8257	36	4	5	0.8697	17	2	1	0.9328	63	9	6	0.8975	0.8783		
1d	300	300	150	180	28	1	9	0.7562	46	4	9	0.8319	23	1	5	0.8204	49	6	5	0.8962	0.8259		
1e	300	300	150	180	43	6	9	0.8159	37	3	10	0.7840	21	3	1	0.9368	58	3	3	0.9485	0.8668		
2a	3000	3000	1500	1800	381	37	50	0.8775	359	28	46	0.8822	200	10	27	0.8793	525	47	29	0.9408	0.8954		
2b	3000	3000	1500	1800	361	29	47	0.8803	346	14	49	0.8748	164	12	28	0.8508	553	60	31	0.9369	0.8872		
2c	3000	3000	1500	1800	397	24	64	0.8588	362	37	50	0.8716	171	14	39	0.8107	549	55	24	0.9493	0.8754		
2d	3000	3000	1500	1800	362	24	62	0.8510	349	24	49	0.8737	208	14	33	0.8602	536	57	27	0.9423	0.8824		
2e	3000	3000	1500	1800	412	39	58	0.8706	358	25	66	0.8414	191	16	33	0.8483	542	54	34	0.9327	0.8741		
3a	30000	30000	15000	18000	3760	307	545	0.8690	3575	285	504	0.8722	1846	144	278	0.8651	5286	559	301	0.9367	0.8864		
3b	30000	30000	15000	18000	3709	311	588	0.8586	3622	267	538	0.8671	1871	156	261	0.8729	5401	578	288	0.9396	0.8847		
3c	30000	30000	15000	18000	3702	272	544	0.8683	3785	317	506	0.8773	1925	148	272	0.8722	5274	555	296	0.9375	0.8893		
3d	30000	30000	15000	18000	3690	309	566	0.8624	3754	301	517	0.8746	1856	110	291	0.8622	5425	550	292	0.9401	0.8856		
3e	30000	30000	15000	18000	3771	280	575	0.8641	3733	286	522	0.8734	1832	136	266	0.8695	5430	529	288	0.9415	0.8876		
r = 2						u _a = 0.2660						u _c = 0.2666					u _m = 0.2095				u _p = 0.2578		
Theoretical	30000	30000	15000	18000	3784	291	560	0.8673	3784	291	560	0.8673	1892	146	280	0.8673	5400	540	252	0.9467	0.8876		
DOE System Performance Effectiveness Equation P _E =P _D					3784	291	560	0.8712	3784	291	560	0.8712	1892	146	280	0.8712	5400	540	252	0.9554	0.8928		

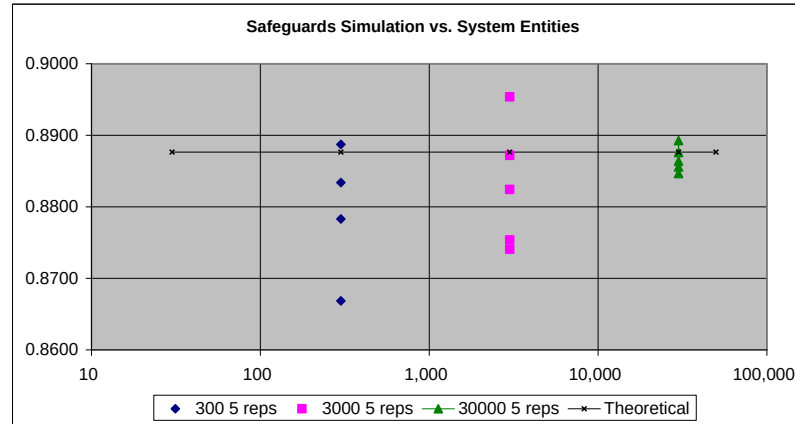


Figure 40 Convergence of Simulated Safeguards System Model to Theoretical Value

seen, the simulation values show convergence to a theoretical value which indicates simulation validity. The converged and theoretical values show agreement with DOE probabilistic method which again shows calculation validity.

4.5 Verifying The Results

The verification of the survey results can be seen in Table 8 that shows the mean response from the survey. A cross check of the table indicating the overall coefficients are reasonable can be seen by looking at the difference of the filtered data to the full survey data for any given coefficient. There were only three values that were outside of the error bars of the survey indicating very good internal agreement and excellent overall agreement between those with greater expertise in a given area verses the entire population. Much of the time the data were not normally distributed and these visually showed the responses being closer to the mean than would be expected in a normal distribution. Again this may indicate that the utilities determined in this survey had good internal agreement, however larger sample sizes would be required to draw this conclusion.

Verifying the airport simulation takes two approaches. The first is checking for internal consistency in the data and as can be seen in Table 11 the “Denied Boarding” value of 1936 (shaded green) for run 1 with 100,00 travelers does not equal (“Bag Denied” + “Computer Denied” + “Metal Denied”) in this case ($932 + 999 + 10 = 1941$). This indicates that there were some travelers after ID screening which had failed more than one of the additional screenings (“Bag Denied” or “Computer Denied” or “Metal Denied”) which agrees with the fact that the maximum “Denied” value was 2 (shaded yellow). “Denied” was a condition variable used to deny boarding if the value was equal to or greater than 1 when any detection was made. If the value for an entity only reached 1 then there were no duplicates and thus “Denied Boarding” would equal (“Bag Denied” + “Computer Denied” + “Metal Denied”) which was the case.

In further verifying the simulation one needs to look for convergence to a theoretical value as the run parameters increase the number of travelers. Figure 39 and Figure 40 show convergence of the output for the simulation runs and plots the effectiveness calculation verses a theoretical value.

Chapter 5 Conclusions and Recommendations

Chapter 5 pulls together the research and summarizes the findings. The checklists and criteria developed during the earlier chapters will be summarized and specifically addressed.

5.1 Summary

In summarizing this research, recall the MIT definition of effectiveness as “the ratio of functions achieved to the totality of functions desired,” [20] which corresponds to the **goal centered approach**, of evaluating system effectiveness which is determined by comparing the systems performance against specific objectives. [52] Under any evaluation approach, there will always be interest in improving system effectiveness.

The System Effectiveness Methodology provides an ongoing single measure of effectiveness quantification based on system performance for a working system in near real-time. In comparing the two calculation methodologies directly the following can be summarized:

The DOE Protection Effectiveness Model

With the overall activity described by DOE Vulnerability Assessment (VA) Requirements which include the description of the physical and operational characteristics of a safeguards and security (S&S) system, assigning values such as delay and detection, and analyzing the results to determine the relative effectiveness in conjunction with the adversary’s capabilities as identified in the Design Basis Threat (DBT). [17]

- *Accounts only for sequential attack (analysis may include other modes – insider threat)*
- *Requires VA practitioners must be specifically trained for VAs*
- *Requires that Initial Surveys be comprehensive*
- *Repeats with periodic surveys usually every 12 months*
- *Requires VA teams to be led by a Federal employee (likely has 5 technical specialists working up to 12 months initially and 3 weeks annually thereafter)[65]*

The Heuristic System Effectiveness Model

- *Accounts for non linear attack (Insider Threat)*
- *Can be created to provide error data automatically*
- *Provides near real time quantification of system performance*
- *Routine values can be compiled and reported by a single statistician with a periodicity deemed reasonable (daily/weekly/monthly)*

Therefore, the Heuristic System Effectiveness Model, is more robust than and compliments the DOE Protection Effectiveness Model

Are $(1-\alpha^x)$ and $(1-\beta^y)$ [in this case $(1-\alpha^{K_a})$ and $(1-\beta^{K_b})$] MoEs ?

Using the definition of Smith and Clark [28], Do the MoEs) should:

1. Increase as effectiveness increases;

Yes (error rate decreases)

2. Be bounded above by an ideal system and below by zero for non-compliance;

Yes (performance rates $(1-\alpha^{K_a})$ and $(1-\beta^{K_b})$ range from 0 to 1)

3. Manage complexity and allow for system decomposition,

Yes (error rates for decomposed system elements used)

4. Facilitate comparisons between systems it is necessary to normalize the final effectiveness scores.

Yes The range chosen (with 0 denoting an ineffective system and 1 denoting a perfectly effective “ideal” system) allows similar systems to be compared

5. Should be ratio scales which means that they have a natural zero point and numbers which are multiples of each other directly indicate their value.

Yes. The measures are modified ratio scales in that a system that performs correctly $(1-\alpha^{K_a})$, or $(1-\beta^{K_b})$ 80% of the time has performed correctly twice as many times as a system that performs correctly 40% of the time or is discounted based on the known tolerance for α or β .

Therefore the selected measures are suitable MoEs.

Does the methodology?:

- Provide a clear definition of a system and system effectiveness that is broad across many applications?;

The definitions are accepted by the community and the effectiveness has been show to be robust.

- That it complements the probabilistic approach (DOE approach of P_E);

The P_E approach is used by the DOE in the design phase and for periodic evaluations.

The e_s approach is used for a near real-time assessment.

- Add to the open literature concerning systems effectiveness measurement?;

The methodology, simulations, and survey have been presented in open forums and included in the proceedings.[67] [71] [73] [74]

- Recognize the complexities of system structures?

Decomposition and aggregation of system components are built into the methodology

Therefore the Goals of the system effectiveness methodology are met.

Are the Hypotheses Affirmed?

When looking at the hypotheses and the checklists developed for each in Chapter 3 the hypotheses can be addressed as follows:

Hypothesis 1: *The Multi-Attribute Utility Theory (MAUT) methodology is a valid approach to incorporate safeguards measures of effectiveness at the component-subsystem level and assess them at the safeguards system level.*

1. Is the proposed solution a Methodology?

- The system effectiveness equation developed and applied with the MAUT is a methodology as a body of methods, rules, and postulates employed by a discipline; [5] and a specific set of directions to a specific location. [6]

2. Are the selected inputs measures of effectiveness?

Using the error rates to determine the correct performance inputs used in the system effectiveness function, i.e. $[(1-\alpha^{K_a})$ and $(1-\beta^{K_b})]$ are MoEs based on Smith and Clark [28] criteria

- a. The measures increase as effectiveness increases;
- b. The measures are bounded above by an ideal system and bounded below by zero for non-compliance;
- c. The measures plus MAUT allow for complexity and allow for system decomposition, and aggregation.
- d. The measures are normalized with the range [20] chosen (with 0 denoting an ineffective system and 1 denoting a perfectly effective “ideal” system)
- e. Are ratio scales which means that they have a natural zero point and numbers which are multiples of each other directly indicate their value.

3. Can the measures of effectiveness be aggregated at the subsystem/component level to produce an overall effectiveness measure?

- The system effectiveness equation has built in the capability to decompose to a level where the MoE can be measured and then aggregated to produce a single measure.

4. Can the methodology be compared to a validated methodology?

- The methodology is validated based on comparison to the probabilistic methodology used by the DOE and found to be complimentary but more conservative in that a wider range of attacks are allowed. When a single layer system is compared using data from simulations for the two methodologies they result in the same value as the tolerance for false alarms is raised to infinity.

Therefore, Hypothesis 1 is affirmed.

Hypothesis 2: Calculating safeguards system effectiveness using Type I and Type II error rates is a valid modeling approach upon which to base an effectiveness methodology with incomplete data, rapidly changing technologies, and uncertainties

1. Is the system effectiveness methodology robust than the DOE Protection Effectiveness Model?

The system effectiveness equation applied using MAUT (the family of systems effectiveness equation), has broader application than the protection effectiveness equation used by DOE in that it provides for the non-linear attacks (e.g. insider threat) and attacks that don't depend on specific scenarios. This produces a more robust measure.

2. As the level of experience increases how are the heuristics affected?

During the lifecycle of a system, specifically the operational phase, the amount of error rate data increases to improve the overall accuracy of and confidence in the MoEs.

3. What relationship do incomplete data, rapidly changing technologies, and uncertainties in the modeling environment have with the heuristics?

With less or incomplete data the heuristics will garner less confidence until operational experience is gained or through specific testing to determine performance. As technologies change creating newer or more accurate methods of determining the error rates, the heuristic approach will have greater validity.

Therefore, Hypothesis 2 is affirmed.

Regarding the research questions:

- A. *“Can a methodology be formulated and used to analyze the effect of incorporating measures of effectiveness concepts at the component-subsystem level, yet assess them at the system level?”*

As Druzdzel established, by measuring effectiveness and using values between zero and one, it is possible to aggregate their effects, that is, given subsystem effectiveness total system effectiveness can be determined

And thus research question A is answered in the affirmative

- B. *“Can uncertainties, such as incomplete data, rapidly changing technologies, and uncertainties in the modeling environment be addressed through heuristics in a system effectiveness methodology?”*

This research shifted from a probabilistic approach that analyzes the likelihood of events and determining effectiveness based on those probabilities to a heuristic approach which uses the experience of measured error rates on a real-time sample (Type I and Type II errors previously mentioned and seen in Table 2). Heuristics have been shown to be highly economical and usually effective. [39] Dewhurst et. al. indicate that research in cognitive psychology indicates that individuals at all educational levels use heuristics especially when making a decision within a limited amount of time. [40] In the case of using a measured rate of errors, as in this research, the more data gathered over time (producing an increased sample size) the more confidence these measures garner.

And thus research question B is answered in the affirmative

5.2 Conclusions

A valid and robust methodology for calculating system effectiveness of a subsystem in near real time has been developed and examples of use presented with the system effectiveness function.

$$e = (1 - \alpha^{K_a}) \cdot (1 - \beta^{K_b})$$

Where $K_a = (r)^{1/c}$ for $r > 1$; otherwise $K_a = 1$

Where $K_b = (r)^{-1/c}$ for $r < 1$; otherwise $K_b = 1$

This allows e to be calculated for any system/subsystem in a dynamic and ongoing basis. When the system is decomposed into subsystems an MAUT methodology is applied using the family of systems effectiveness equation.

$$e_s = \text{System Effectiveness} = \sum u_i e_i$$

$$\text{Subject to: } 0 \leq e_i \leq 1 \quad \sum u_i = 1$$

The example of a domestic safeguards system illustrates the full methodology in use and shows that the system can be decomposed to the level of subsystems to the extent that the α and β rates can be measured.

Once a baseline for a specific system is measured, comparisons over time will allow an ongoing and consistent effectiveness evaluation.

Should conditions (c = consequence/stress factor) change the system assumptions can be modified (change the ratio r) to gain acceptable solution space.

5.3 Recommendations for Future Work

The system effectiveness equation and MAUT methodology offer a repeatable and quantifiable application in evaluating generic system effectiveness and in particular within the domestic safeguards community. In this specific instance, the subcomponents need refinement. The survey instrument could easily be revised based on this experience and be repeated at the next (or any subsequent) INMM Annual Meeting. Once the survey instrument is refined and attendees are familiar with the concept and time requirements for taking the survey, a greater proportion of the population would likely respond leading to a smaller error band in the results.

For each individual system, the definition and decomposition of the system components is the key to the applicability to the real world. Once a system manager has determined the worth of the methodology, the utilities need to be developed based on manager and/or expert input, thus yielding a valid method for an ongoing evaluation of any given system's effectiveness.

This work opens an area of research that allows the consistent ongoing evaluation of a given systems effectiveness. The subsystems are defined and error rates measured giving a significantly increased quantification and consistency in system effectiveness determinations. Methodologies for determining the parameters r and c need to be developed as guidance for the system manager.

Specific areas of potential future work to apply this system effectiveness methodology could include inter alia such areas as the medical field, quality assurance (QA) applications, and justice system evaluations. In each of these cases measuring the rates of error are much more accessible. In particular the missed alarm (β) rate would have a much more direct measurement. In the case of the medical field testing for presence of a particular disease would have a direct measurement of false positives and false negatives. In the area of QA processes for determining quality could be judged again based on α and β rates. In the justice applications false positives against individuals might have much lower tolerance than in a number of other applications and the effectiveness calculations adjusted accordingly.

References

References

1. Habayeb, A.R., *System Effectiveness*. 1987, Headington Hill Hall, Oxford, England: Pergamon Press.
2. United Nations Security Council. *Resolution 1540 (2004)*. 2004 [cited 11/11/2008; Available from: <http://www.un.org/sc/1540/>].
3. Soban, D.S. and D.N. Mavris. *The Need For A Military System Effectiveness Framework: The System Of Systems Approach*. AIAA-2001-5226 2001 11/29/2008]; Available from: http://smartech.gatech.edu/dspace/bitstream/1853/25365/1/Soban-2001_1.pdf.
4. ARINC Res. Corp, *Reliability Engineering*. 1964, Englewood Cliffs, NJ: Prentice Hall
5. Merriam-Webster. *Merriam-Webster Online*. 2009 [cited 2009 2/15/2009]; Available from: <http://www.merriam-webster.com/>.
6. Korzenowski, K. *Microsoft Solutions Framework, Executive Overview*. 2002 [cited 2009 2/14/2009]; Available from: www.incose.org/northstar/2002Slides/MSF%20Executive%20Overview%20-%20kylek.ppt
7. Abdallah, S. *Information Systems Methodologies*. 2007 [cited 2009 02/14/2009]; Available from: <http://ocp.uni-passau.de/drupal/?q=node/293>.
8. Tillman, F.A., C.L. Hwang, and K. Way, *System Effectiveness Models: An Annotated Bibliography*. IEEE Transactions of Reliability, 1980. **Vol. R-29**(No. 4): p. 295-304.
9. DoD-ATL. *Frequently Asked Questions*. 2009 [cited 2009 2/15/2009]; Available from: www.acq.osd.mil/dpap/Docs/FAQs%20--%20SoS%20&%20FoS.doc.
10. Jackson, D., G. Sedrick, and A. Badiru. *Functional Development of Performance Metrics for System of Systems*. in *2009 Industrial Engineering Research Conference*. 2009.
11. Dixon, J.R., *Design Engineering: Inventiveness, Analysis, and Decision Making*. 1966, New York: McGraw-Hill Book Company.
12. Miller, M.M. *Are IAEA Safeguards on Plutonium Bulk-Handling Facilities Effective?* 1990 11/29/2008]; Available from: <http://www.nci.org/k-m/mmsggrds.htm>.
13. DOE-SA. *DOE M 470.4-7 Safeguards And Security Program References*. 2005 12/2007]; Available from: <http://www.directives.doe.gov>.
14. NRC. *10 CFR PART 70--DOMESTIC LICENSING OF SPECIAL NUCLEAR MATERIAL*. 2007 12/2007]; Available from: <http://www.nrc.gov/reading-rm/doc-collections/cfr/part070/full-text.html#part070-0004>.
15. IAEA. *International Atomic Energy Agency, IAEA Safeguards: Stemming the Spread of Nuclear Weapons*. 2007 [cited 2007; Available from: <http://www.iaea.org/>].
16. IAEA, *IAEA Safeguards Glossary*. International Nuclear Verification Series. 2001.
17. DOE SA. *DOE M 470.4-1, SAFEGUARDS AND SECURITY PROGRAM PLANNING AND MANAGEMENT*. 2005; Available from: <http://www.directives.doe.gov>.
18. Blanchard, B.S., *Logisitics Engineering and Management*. 5th ed. 1998: Prentice Hall Inc.
19. International Council on Systems Engineering (INCOSSE). *Systems Engineering Handbook v3.1*. 2007 12/2007]; Version 3.1:[Available from: <http://www.incose.org/ProductsPubs/products/sehandbook.aspx>].

20. Allen, T., et al. *ESD Terms and Definitions*. Massachusetts Institute of Technology, Engineering Systems Division, Working Paper Series 2001 [cited 2009; Version 12:[Available from: <http://esd.mit.edu/WPS/esd-wp-2002-01.pdf>.
21. Blanchard, B.S. and W.J. Fabrycky, *Systems Engineering and Analysis*. 3rd ed. 1998, Upper Saddle River, New Jersey: Prentice Hall.
22. MIL-STD-721B, *Definitions of Effectiveness Terms for Reliability, Maintainability, Human Factors, and Safety*, US Department of Defense, Editor. 1966.
23. Hoban, F., et al. *NASA Systems Engineering Handbook*. 1995 12/2007]; Available from: <http://ntrs.nasa.gov/search.jsp?R=824878&id=6&q=N%3D4294799254>.
24. Al-Ayat, R.A. and B.R. Judd, *Framework for Evaluating the Effectiveness of Nuclear-Safeguards Systems. [Aggregated Systems Model (ASM)]*, in UCRL-85711; CONF-811012-11; Other: ON: DE82002102. 1981, Lawrence Livermore National Laboratory. p. Size: Pages: 4.
25. Sproles, N., *Formulating Measures of Effectiveness*. Systems Engineering, 2002. 5(4): p. 253-263.
26. Sommerville, I., *Software Engineering*. 6th ed. 2001, Boston: Addison-Wesley.
27. Smith, N. and T. Clark. *An Exploration of C2 Effectiveness – A Holistic Approach*. in *2004 Command and Control Research and Technology Symposium*. 2004. San Diego, CA.
28. Smith, N. and T. Clark. *A Framework to Model and Measure System Effectiveness*. 2005 11/11/2008]; Available from: http://www.dodccrp.org/events/11th_ICCRTS/html/papers/054.pdf.
29. Institute for Distance Education. *About Systems Theory*. 2003 [cited 11/11/2008; Available from: <http://ide.ed.psu.edu/change/theory/systems.htm>.
30. Rouse, W.B. *Complex Systems Phenomena, Characteristics & Research Questions* 2005 11/30/2008]; Available from: http://hcs.ucla.edu/lake-arrowhead-2007/Paper1_Rouse.ppt.
31. Keeney, R.L. and H. Raiffa, *Decisions with Multiple Objectives: Preferences and Value Tradeoffs* 1993, Cambridge, United Kingdom: Cambridge University Press.
32. Smith, N. *A Proposed Framework to Formulate and Measure Effectiveness*. in *DORS Conference*,. 2005. Melbourne, Australia.
33. Druzdzel, M.J. *Explanation in Probabilistic Systems: Is It Feasible? Will It Work?* in *Intelligent Information Systems V Proceedings*. 1996. Deblin, Poland.
34. Agosta, J. and M. Gardos, *Bayes Network “Smart” Diagnostics*. Intel Technology Journal, 2004. **08**(04).
35. Pradhan, M., et al., *The sensitivity of belief networks to imprecise probabilities: An experimental investigation*. Artificial Intelligence Journal, 1996. **84**(1-2): p. 357.
36. Cox, R.T., *Probability, Frequency and reasonable Expectation*. American Journal of Physics, 1946. **14**(1).
37. Cox, R.T., *The algebra of probable inference*. 1961, Baltimore, Mayland: The Johns Hopkins Press.
38. Wu, Y.-S., et al. *Automated Adaptive Intrusion Containment in Systems of Interacting Services*. [Electrical and Computer Engineering - ECE Technical Reports] 2005 [cited 2009 04/13/2009]; Available from: <http://docs.lib.purdue.edu/ecetr/68>.

39. Tversky, A. and D. Kahneman, *Judgment Under Uncertainty: Heuristics and Biases*. Science, New Series, 1974. **185** (4157): p. 1124-1131.
40. Dewhurst, M.E., et al., *Probability Biases in Genetic Problem Solving: A Comparison of Undergraduates, Genetic Counseling Graduate Students, and Genetic Counselors*. Journal of Genetic Counseling, 2007. **16**(2).
41. Baker, A.B., et al., *A Scalable Systems Approach for Critical Infrastructure Security*, in *Other Information: PBD: 1 Apr 2002*. 2002. p. Size: 53 pages.
42. Office of Security and Safety Performance Assurance, *Safeguards and Security Management*, in *DOE M 470.4-1*, Department of Energy, Editor. 2005.
43. Swets, J.A., *The Relative Operating Characteristic in Psychology*. Science, 1973. **182**(4116): p. 990-1000.
44. Swets, J.A., *Signal Detection Theory And ROC Analysis In Psychology And Diagnostics: Collected Papers*. 1996, Mahwah, New Jersey: Lawrence Erlbaum Associates, Publishers.
45. Miller, J. *Earliest Known Uses of Some of the Words of Mathematics*. 2007 [cited 12/2007; Available from: <http://jeff560.tripod.com/mathword.html>].
46. Neyman, J. and E.S. Pearson, *On the Use of Certain Test Criteria for Purposes of Statistical Inference, Part I*. Biometrika, 1928. **20A**: p. 175-240
47. Neyman, J. and E.S. Pearson, *On the Problem of the Most Efficient Tests of Statistical Hypotheses*. Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character, 1933. **231**: p. 289-337.
48. Neyman, J. and E.S. Pearson, *The Testing of Statistical Hypotheses in Relation to Probabilities A Priori*. Proceedings of the Cambridge Philosophical Society, 1933. **24** p. 492-510.
49. Sage, A.P. and W.B. Rouse, *Handbook of Systems Engineering and Management*. 1999, New York: John C. Wiley & Sons, INC.
50. Villemeur, A., *Reliability, Availability, Maintainability, and Safety Assessment*. Vol. 1. 1992, West Sussex, England: John Wiley and Sons Ltd. 363.
51. van Schalkwyk, J. *Receiver Operating Characteristic Curves: an Introduction*. 1998-2006 April 6, 2007 12/2007]; Available from: <http://www.anaesthetist.com/mnm/stats/roc/Findex.htm>.
52. Hamilton, S. and N. Chervany, *Evaluating Information Systems Effectiveness - Part 1 Comparing Evaluation Approaches*. MIS Quarterly, 1981. **5**(3): p. 55-69.
53. Hamilton, S. and N.L. Chervany, *Evaluating Information System Effectiveness - Part I: Comparing Evaluation Approaches*. MIS Quarterly, 1981. **5**(3): p. 55-69.
54. DoD, *Definitions of Effectiveness Terms for Reliability, Maintainability, Human Factors, and Safety*, US Department of Defense, Editor. 1966.
55. Soban, D.S. and D.N. Mavris. *Formulation of a Methodology for the Probabilistic Assessment of System Effectiveness*. [Presented at the AIAA 2000 Missile Sciences Conference, Monterey CA, November 7-9, 2000.] 2000 10/31/2008]; Available from: <http://smartech.gatech.edu/bitstream/1853/6327/1/AIAA-MSC-2000-DS.pdf>.
56. Bennett, H.A., et al., *Safeguards System Effectiveness Modeling*, in SAND-76-5234; CONF-760615-6. 1976, Sandia National Laboratory. p. Size: Pages: 18.
57. Wilkey, D.D., et al., *Evaluation of Effectiveness for MC&A Programs*, in LA-UR-04-4452. 2004, Los Alamos National Laboratory.

58. Sicherman, A., D.S. Fortney, and C.J. Patenaude, *A Database Model for Evaluating Material Accountability Safeguards Effectiveness Against Protracted Theft*, in UCRL-JC-114597; CONF-930749--54. 1993, Lawrence Livermore National Laboratory. p. Size: 8 P.
59. Fishburn, P.C., *Utility Theory*. Management Science, 1968. **14**(5): p. 335-378.
60. Fishburn, P.C., *Methods of Estimating Additive Utilities*. Management Science, 1967. **13**(7): p. 435-453.
61. Halpern, D.F., *Thought and knowledge: An introduction to critical thinking*. 4th ed. 2003, Mahwah, NJ: Erlbaum.
62. Jackson, D.F., *Summary provided in an E-mail from Dr. D. F. Jackson*, C.W. Coates, Editor. 2009: Knoxville.
63. Turban, E. and L.M. Morton, *Utility Theory Applied to Multivariable System Effectiveness Evaluation*. Management Science, 1971. **17**(12): p. B817-B828.
64. Churchman, C.W., R.L. Ackoff, and N.M. Smith Jr, *An Approximate Measure of Value*. Journal of the Operations Research Society of America, 1954. **2**(2): p. 172-187.
65. Lambert, L.D., *Discussion of VA Requirements*, C.W. Coates, Editor. 2009: Oak Ridge. p. 1.
66. TSA, *Recommended Security Guidelines for Airport Planning, Design and Construction revised June 14, 2006*, Transportation Security Administration, Editor. 2006.
67. Coates, C. and C.H. Aikens. *Safeguards System Effectiveness Optimization*. in *47th INMM Annual Meeting*. 2006. Nashville TN: Institute of Nuclear Material Management.
68. Institute of Nuclear Materials Management (INMM). *INMM Home Page*. 2007 12/2007]; Available from: <http://www.inmm.org/>.
69. UTK-OIT. *Statistical Consulting Center*. 2007 12/2007]; Available from: <http://oit.utk.edu/scc/>.
70. Kelton, W.D., R.P. Sadowski, and D.A. Sadowski, *Simulation with Arena Second Edition*, ed. K.E. Case and P.M. Wolfe. 2002, New York: McGraw Hill.
71. Coates, C.W. *Simulation of a Simple Domestic Safeguards System*. in *INMM Annual Meeting*. 2007. Phoenix, AZ: Institute of Nuclear Material Management Proceedings.
72. Morgan, D., *Aviation Security Technologies and Procedures: Screening Passengers and Baggage*, C.R. Service, Editor. 2001, The Library of Congress: Washington DC. p. 12.
73. Coates, C.W. and D.F. Jackson. *System Effectiveness Model Formulation*. in *INMM 49th Annual Meeting*. 2008. Nashville, TN: Institute of Nuclear Material Management.
74. Coates, C.W. *Domestic Safeguards System Effectiveness Survey Results (Accepted)*. in *INMM 50th Annual Meeting*. 2009. Tucson, AZ: Institute of Nuclear Material Management Proceedings.

APPENDICES

Appendix I Safeguards Survey

This appendix contains a printout of the final survey tool with all questions in the order given. The respondent was asked to rate subsystem components in all areas however the data was filtered in some of the analyses to those with greater than 5 years experience in the area for comparison to the full data set.

Domestic Safeguards Effectiveness Survey

The purpose of the survey is to gauge the safeguards community's evaluation of, and interaction between, domestic safeguards subsystems and how they contribute to overall system effectiveness. Definition of a System The International Council on Systems Engineering (INCOSE) defines a system as *"a combination of interacting elements organized to achieve one or more stated purposes."* Approaches to System Effectiveness There are two fundamental approaches to evaluating system effectiveness: a resource approach and a goal centered approach. Under a resource approach, system effectiveness is determined in terms of resource availability instead of specific task objectives. The term cost-effective captures the essence of the resource approach. The term is usually modified by either "more" or "less" cost-effective indicating a relative evaluation. Under a goal centered approach, system effectiveness is determined by comparing the system's performance against specific objectives. Often under a goal centered approach, the effectiveness is defined in terms of percent effective and usually indicates a more absolute measure of obtaining the objective. Under either grading approach, there will always be interest in improving system effectiveness; however, this analysis assumes a goal centered approach. This Survey's Definition of Domestic Safeguards The U.S. Department of Energy (DOE) has defined Safeguards as; *"an integrated system of physical protection, material accounting, and material control measures designed to deter, prevent, detect, and respond to unauthorized possession, use, or sabotage of nuclear materials."* United Nations Security council resolution (UNSCR) 1540 decides that, inter alia, States (Countries) shall: *"take and enforce effective measures to establish domestic controls to prevent the proliferation of nuclear, chemical, or biological weapons and their means of delivery, including by establishing appropriate controls over related materials."* The following survey is based on a generic domestic safeguards system, containing a category 1 (it has more than 200 kg of highly enriched uranium in the form of metal) amount of nuclear materials, that is comprised of four subsystems. Those subsystems are Physical Protection, Material Control, Material Accounting, and Plant Management. Each of these subsystems

will be further broken down into somewhat more specific components. It is hoped this survey will gauge your opinion on how these subsystems and components interact and their relative importance. The survey should take less than 15 minutes of your time. Thank You

INTRODUCTION_2

Please indicate your highest education level.

- ☐ No Diploma
☐ High School
☐ Bachelors
☐ Masters
☐ PhD

EXPERIENCE

In which of the following areas do you have experience?

- ☐ Physical Protection
☐ Material Accounting
☐ Material Control
☐ Plant Management
☐ Domestic Safeguards Policy

INTRODUCTION_3

Please list your years of experience in the following domestic safeguards technical areas:

Years (0 - 100)

Physical Protection					.		
Material Accounting					.		
Material Control					.		
Plant Management					.		
Domestic Safeguards Policy					.		

SECA_1_

Section A Plant Management

(A1); An effective Plant Management subsystem is an important part of a domestic safeguards system?

- ☐ True
☐ False

SECA_2

Section A Plant Management

(A2) Please indicate the level of importance of the following items on a scale of 0 (low) to 100 (high) to a Plant Management Subsystem.

	Component Rank (0 - 100)					
Personnel Reliability						
Budgeting						
Knowledge of Material Protection, Control, and Accounting principals						
System Oversight (audits, surveillance, etc)						
System Documentation						
Configuration Management						
Independent Assessment						
Organizational Communication						
Organizational Coordination						
Personnel Training						

SECA_4

Section A Plant Management (A3) Hypothetical: The subsystem covering Plant Management has alarms designed to indicate when negative conditions exist. In general for evaluation of a Plant Management subsystem, missed alarms should have ____ times (0.00 to 50.00) the weight of false alarms. (1 signifies equal weight)

(0 - 50)

----------------------	----------------------	----------------------	----------------------	----------------------	----------------------

SECA_5

Section A Plant Management (A4) Hypothetical: A "personnel reliability" subsystem has a 20% rate of rejecting reliable employees and a 5% rate of accepting unreliable employees. What is your guess for how effective (percent) the personnel reliability subsystem is?

(0 - 100)

--	--	--	--	--	--	--	--

SECA_6

Section A Plant Management

(A5) Plant Management subsystems, with the components as indicated in this survey, function independently from which of the following subsystems:

- ☐ Physical Protection
☐ Material Control
☐ Material Accounting
☐ No Answer

SECB_1

Section B - Physical Protection

(B1) An effective Physical Protection subsystem is an important part of a domestic safeguards system?

- ☐ True
☐ False

SECB_2

Section B - Physical Protection

(B2) Please indicate the level of importance of the following items on a scale of 0 (low) to 100 (high) to a Physical Protection Subsystem.

Component Rank (0 - 100)

Determining System Objectives	<table border="1"><tr><td></td><td></td><td></td><td></td><td></td><td></td></tr></table>						
Facility Characterization	<table border="1"><tr><td></td><td></td><td></td><td></td><td></td><td></td></tr></table>						
Threat Definition	<table border="1"><tr><td></td><td></td><td></td><td></td><td></td><td></td></tr></table>						
Target Identification							

					.		
Exterior Intrusion Sensors					.		
Interior Intrusion Sensors					.		
Alarm Assessment					.		
Access Delay (Barriers, Fences, Gates, etc)					.		
Secure Communications (Radios etc,)					.		
Alarm Response					.		
Material Control System					.		
Physical Protection System Testing					.		
Personnel Reliability					.		
Independent System Assessment					.		
Vulnerability Assessment					.		
Internal Guard Force					.		
Special Response Teams					.		

SecB_4

Section B - Physical Protection (B3) Hypothetical: The subsystem covering Physical Protection has alarms designed to indicate when negative conditions exist. In general for evaluation of a physical protection subsystem, missed alarms should have ____ times (0.00 to 50.00) the weight of false alarms. (1 signifies equal weight)

(0 - 50)

				.		
----------------------	----------------------	----------------------	----------------------	---	----------------------	----------------------

SecB_5

Section B - Physical Protection (B4) Hypothetical: An intrusion alarm system has a 10% false alarm rate and a 2% missed alarm rate. What is your guess for how effective (percent) the intrusion

alarm system is?

(0 - 100)

--	--	--	--	--	--	--	--

SECB_6

Section B - Physical Protection

(B5) Physical Protection subsystems, with the components as indicated in this survey, function independently from which of the following subsystems:

- ☐ Plant Management
- ☐ Material Control
- ☐ Material Accounting
- ☐ No Answer

SECC_1

Section C - Material Control

(C1) An effective Material Control subsystem is an important part of a domestic safeguards system?

- ☐ True
- ☐ False

SECC_2

Section C - Material Control

(C2) Please indicate the level of importance of the following items on a scale of 0 (low) to 100 (high) to a Material Control Subsystem.

Component Rank (0 - 100)

Entry/Access Control (Badges, Biometrics, Locks, etc)	<table border="1"><tr><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr></table>								
Tamper Indicating Devices	<table border="1"><tr><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr></table>								
Nuclear Material Portal Monitors	<table border="1"><tr><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr></table>								
Physical Protection	<table border="1"><tr><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr></table>								
Procedures	<table border="1"><tr><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr></table>								
Item/Lot Tracking	<table border="1"><tr><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr></table>								

Oversight (audits, tests, etc)	☐	☐	☐	☐	.	☐	☐
Material Surveillance	☐	☐	☐	☐	.	☐	☐
Material Containment	☐	☐	☐	☐	.	☐	☐
Daily Administrative Checks	☐	☐	☐	☐	.	☐	☐
Attribute (Presence) Monitors	☐	☐	☐	☐	.	☐	☐
Process Control	☐	☐	☐	☐	.	☐	☐
Waste Monitoring	☐	☐	☐	☐	.	☐	☐
Storage Monitoring	☐	☐	☐	☐	.	☐	☐

SECC_4

Section C - Material Control (C3) Hypothetical: The subsystem covering Material Control has alarms designed to indicate when negative conditions exist. In general for evaluation of a Material Control subsystem, missed alarms should have ____ times (0.00 to 50.00) the weight of false alarms. (1 signifies equal weight)

(0 - 50)

(0 - 50)							
						.	

SECC_5

Section C - Material Control (C4) Hypothetical: A nuclear material portal monitor system has a 50% false alarm rate and a 10% missed alarm rate. What is your guess for how effective (percent) the portal monitor system is?

(0 - 100)

--	--	--	--	--	--

SECC_6

Section C - Material Control

(C5) Material Control subsystems, with the components as indicated in this survey, function independently from which of the following subsystems:

- ☐ Plant Management
- ☐ Physical Protection

- ☐ Material Accounting
☐ No Answer

SecD_1

Section D - Material Accounting

(D1) An effective Material Accounting subsystem is an important part of a domestic safeguards system?

- ☐ True
☐ False

SecD_2

Section D - Material Accounting

(D2) Please indicate the level of importance of the following items on a scale of 0 (low) to 100 (high) to a Material Accounting Subsystem

	Component Rank (0 - 100)									
Accounting system						.				
Physical inventory						.				
Measurement (mass) accuracy/uncertainty						.				
Material transfer monitoring						.				
Real-time accounting						.				
Measurement Control and Calibration						.				
Oversight (audits, etc)						.				
Anomaly detection and resolution						.				
Inventory difference analysis						.				
Shipper/receiver difference analysis						.				
Process monitoring						.				
Statistical evaluation						.				
Procedures						.				
Material hold-up						.				

Independent
assessment

--	--	--	--	--	--	--	--

SecD_4

Section D - Material Accounting (D3) Hypothetical: The subsystem covering Material Accounting has alarms designed to indicate when negative conditions exist. In general for evaluation of a Material Accounting subsystem, missed alarms should have ____ times (0.00 to 50.00) the weight of false alarms. (1 signifies equal weight)

(0 - 50)

--	--	--	--	--	--	--	--

SecD_5

Section D - Material Accounting (D4) Hypothetical: A nuclear material bar code system has a 8% rate of rejecting a true reading and a 4% rate of accepting a false reading. What is your guess for how effective (percent) the bar code system is?

(0 - 100)

--	--	--	--	--	--	--	--

SecD_6

Section D - Material Accounting

(D5) Material Accounting subsystems, with the components as indicated in this survey, function independently from which of the following subsystems:

- ☐ Plant Management
- ☐ Physical Protection
- ☐ Material Control
- ☐ No Answer

SecE_1

Section E - Summary

(E1) Given the domestic safeguards system as described in this survey, please indicate the level of importance on a scale of 0 (no importance) to 100 (high importance) of the following subsystems for determining overall system effectiveness:

Component Rank (0 - 100)

Plant Management					.		
Material Control					.		
Material Accounting					.		
Physical Protection					.		

SEC_E_2

Section E – Summary (E2) Hypothetical (Take your best shot!) The domestic safeguards system as described in this survey has the following evaluations: Plant Management - 40% effective Material Control - 60% effective Material Accounting - 10% effective Physical Protection - 90% effective What is your guess at the overall system effectiveness?

(0 - 100)

				.		
----------------------	----------------------	----------------------	----------------------	---	----------------------	----------------------

SEC_E_3

Section E – Summary (E3) Hypothetical (Take your best Shot!) The domestic safeguards system as described in this survey has the following evaluations: Plant Management - 50% effective Material Control - 50% effective Material Accounting - 50% effective Physical Protection - 50% effective What is your guess at the overall system effectiveness?

(0 - 100)

				.		
----------------------	----------------------	----------------------	----------------------	---	----------------------	----------------------

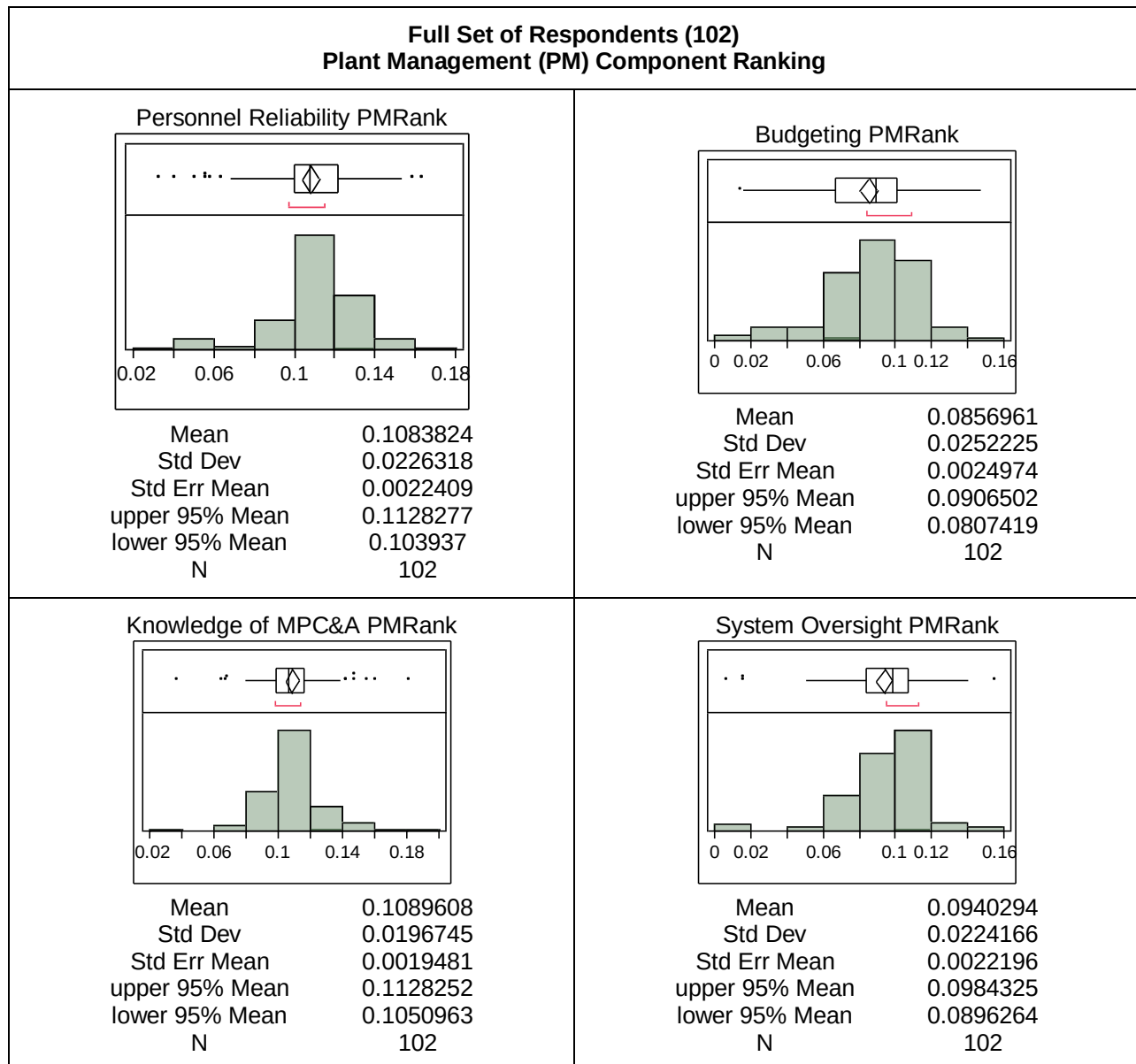
COMMENTS

Please add any additional comments you would like to include and THANK YOU for your participation!

Appendix II Survey Results

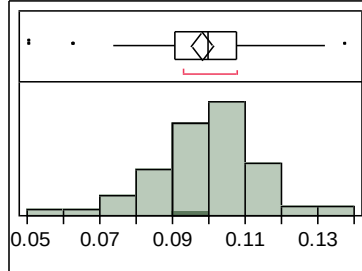
This Appendix presents the JMP program output for the responses in the Safeguards Survey in graphical and Tabular form. Appendix II is separated into sections. Section 1 contains the distributions for the Full Set of respondents. Section 2 contains the distributions for the Plant Management filtered responses. Section 3 contains the Material Control filtered responses. Section 4 contains Material Accounting filtered responses. Section 5 contains the Physical Protection filtered responses.

Section 1 Full Set of Respondents



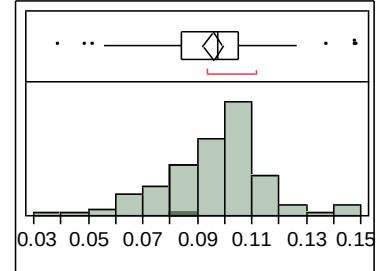
Full Set of Respondents (102)
Plant Management (PM) Component Ranking

System Documentation PMRank



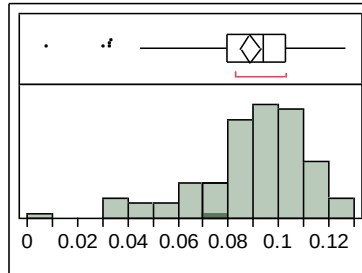
Mean	0.0983137
Std Dev	0.015442
Std Err Mean	0.001529
upper 95% Mean	0.1013468
lower 95% Mean	0.0952806
N	102

Configuration Management PMRank



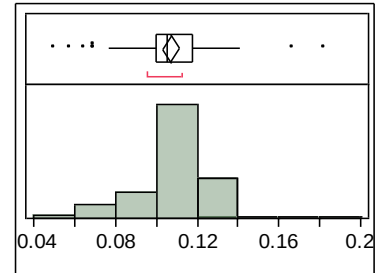
Mean	0.0958922
Std Dev	0.0192405
Std Err Mean	0.0019051
upper 95% Mean	0.0996713
lower 95% Mean	0.092113
N	102

Independent Assessment PMRank



Mean	0.0888529
Std Dev	0.0222472
Std Err Mean	0.0022028
upper 95% Mean	0.0932227
lower 95% Mean	0.0844832
N	102

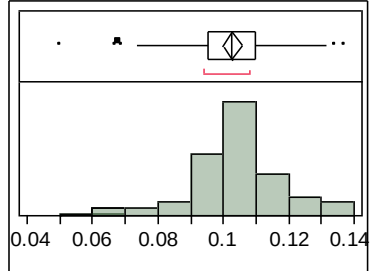
Org Communication PMRank



Mean	0.1074706
Std Dev	0.0196417
Std Err Mean	0.0019448
upper 95% Mean	0.1113286
lower 95% Mean	0.1036126
N	102

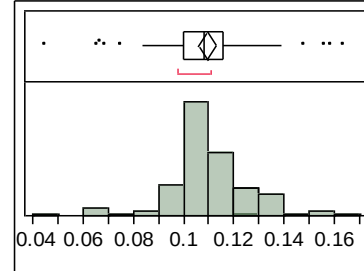
Full Set of Respondents (102)
Plant Management (PM) Component Ranking

Org Coordination PMRank



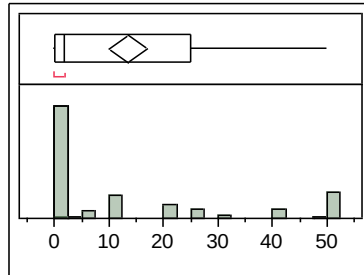
Mean	0.1028333
Std Dev	0.0146159
Std Err Mean	0.0014472
upper 95% Mean	0.1057042
lower 95% Mean	0.0999625
N	102

Personnel Training PMRank



Mean	0.1096765
Std Dev	0.0176198
Std Err Mean	0.0017446
upper 95% Mean	0.1131373
lower 95% Mean	0.1062156
N	102

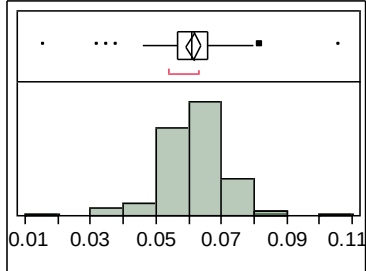
Missed Alarms / False Alarms



Mean	13.59902
Std Dev	18.077887
Std Err Mean	1.7899775
upper 95% Mean	17.149854
lower 95% Mean	10.048186
N	102

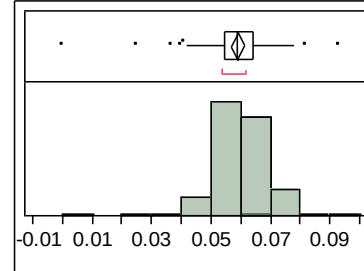
Full Set of Respondents (102)
Physical Protection (PP) Component Ranking

System Objectives PPrank



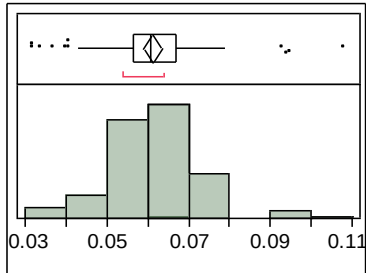
Mean	0.0615098
Std Dev	0.0107921
Std Err Mean	0.0010686
upper 95% Mean	0.0636296
lower 95% Mean	0.05939
N	102

Facility Characterization PPrank



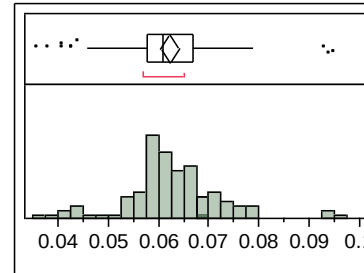
Mean	0.0591863
Std Dev	0.0109406
Std Err Mean	0.0010833
upper 95% Mean	0.0613352
lower 95% Mean	0.0570373
N	102

Threat Definition PPrank



Mean	0.0612059
Std Dev	0.0119763
Std Err Mean	0.0011858
upper 95% Mean	0.0635583
lower 95% Mean	0.0588535
N	102

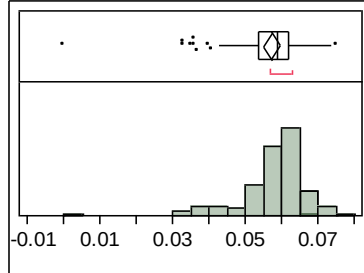
Target ID PPrank



Mean	0.0623431
Std Dev	0.0101499
Std Err Mean	0.001005
upper 95% Mean	0.0643368
lower 95% Mean	0.0603495
N	102

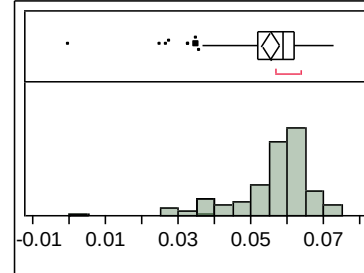
Full Set of Respondents (102)
Physical Protection (PP) Component Ranking

Exterior Intrusion Sensors PPrank



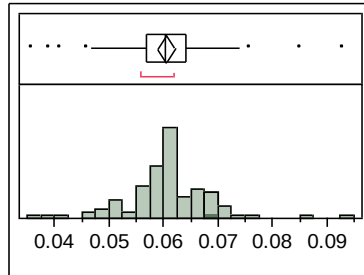
Mean	0.0573431
Std Dev	0.0103402
Std Err Mean	0.0010238
upper 95% Mean	0.0593742
lower 95% Mean	0.0553121
N	102

Interior Intrusion Sensors PPrank



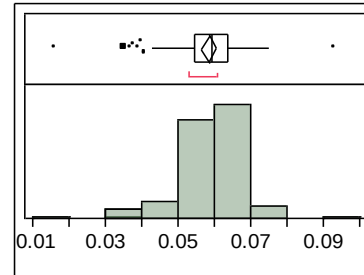
Mean	0.055451
Std Dev	0.0114681
Std Err Mean	0.0011355
upper 95% Mean	0.0577035
lower 95% Mean	0.0531984
N	102

Alarm Assessment PPrank



Mean	0.0605588
Std Dev	0.0080341
Std Err Mean	0.0007955
upper 95% Mean	0.0621369
lower 95% Mean	0.0589808
N	102

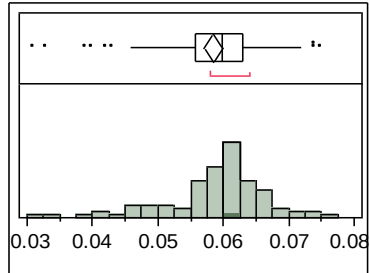
Access Delay PPrank



Mean	0.058549
Std Dev	0.0096428
Std Err Mean	0.0009548
upper 95% Mean	0.060443
lower 95% Mean	0.056655
N	102

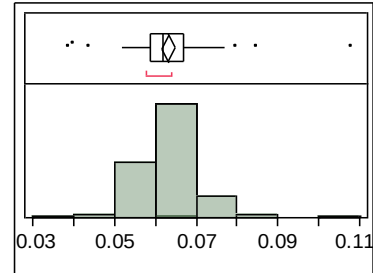
Full Set of Respondents (102)
Physical Protection (PP) Component Ranking

Secure Comm PPrank



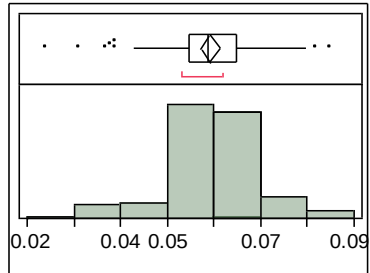
Mean	0.0585098
Std Dev	0.0078118
Std Err Mean	0.0007735
upper 95% Mean	0.0600442
lower 95% Mean	0.0569754
N	102

Alarm Response PPrank



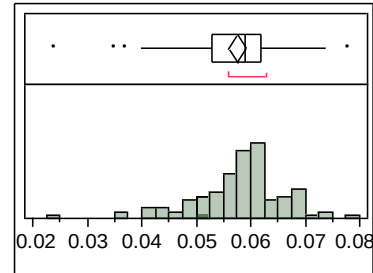
Mean	0.0633235
Std Dev	0.0081825
Std Err Mean	0.0008102
upper 95% Mean	0.0649307
lower 95% Mean	0.0617163
N	102

Material Control System PPrank



Mean	0.0591863
Std Dev	0.0098169
Std Err Mean	0.000972
upper 95% Mean	0.0611145
lower 95% Mean	0.0572581
N	102

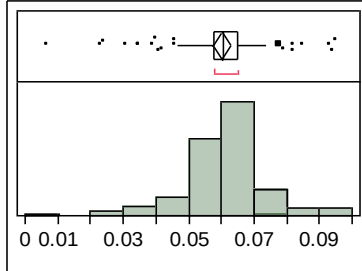
PP System Testing PPrank



Mean	0.0575784
Std Dev	0.0086074
Std Err Mean	0.0008523
upper 95% Mean	0.0592691
lower 95% Mean	0.0558878
N	102

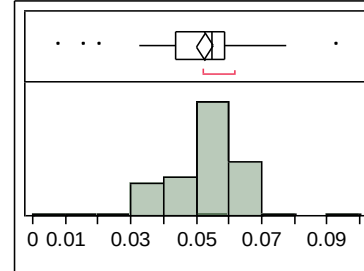
Full Set of Respondents (102)
Physical Protection (PP) Component Ranking

Personnel Reliability PPrank



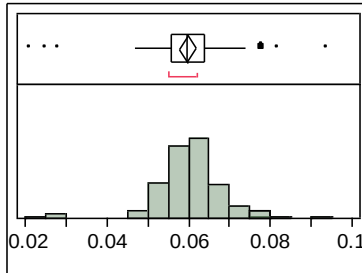
Mean	0.0602549
Std Dev	0.0132633
Std Err Mean	0.0013133
upper 95% Mean	0.0628601
lower 95% Mean	0.0576497
N	102

Independent System Assessment PPrank



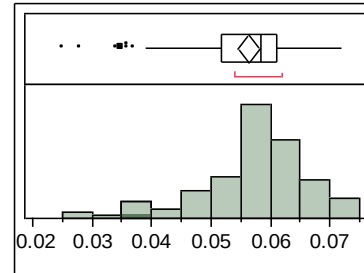
Mean	0.0524804
Std Dev	0.0121726
Std Err Mean	0.0012053
upper 95% Mean	0.0548713
lower 95% Mean	0.0500895
N	102

Vulnerability Assessment PPrank

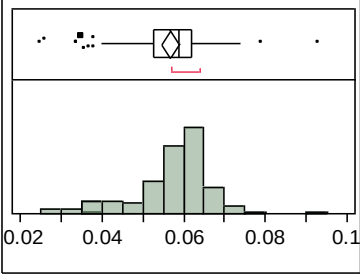
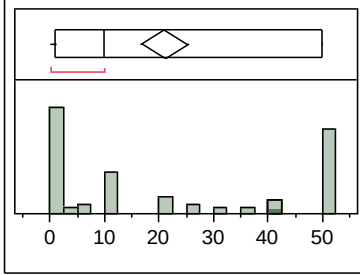


Mean	0.0598431
Std Dev	0.0095396
Std Err Mean	0.0009446
upper 95% Mean	0.0617169
lower 95% Mean	0.0579694
N	102

Internal Guard Force PPrank

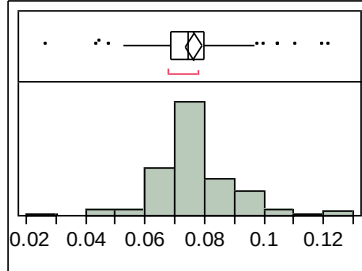


Mean	0.0564608
Std Dev	0.009411
Std Err Mean	0.0009318
upper 95% Mean	0.0583093
lower 95% Mean	0.0546123
N	102

Full Set of Respondents (102) Physical Protection (PP) Component Ranking	
Special Response Teams PPrank  Mean0.0568824 Std Dev0.010461 Std Err Mean0.0010358 upper 95% Mean0.0589371 lower 95% Mean0.0548276 N102	Missed Alarms / False Alarms PP  Mean21.081609 Std Dev20.794742 Std Err Mean2.2294303 upper 95% Mean25.51357 lower 95% Mean16.649648 N87

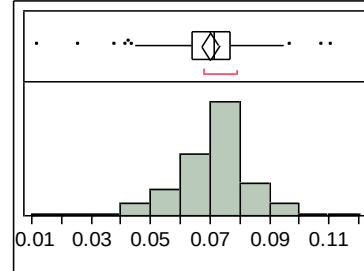
Full Set of Respondents (102)
Material Control (MC) Component Ranking

Entry Access Control MCrank



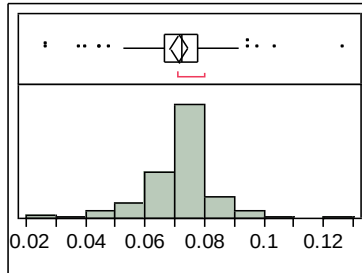
Mean	0.0764118
Std Dev	0.0142581
Std Err Mean	0.0014118
upper 95% Mean	0.0792123
lower 95% Mean	0.0736112
N	102

Tamper Indicating MCrank



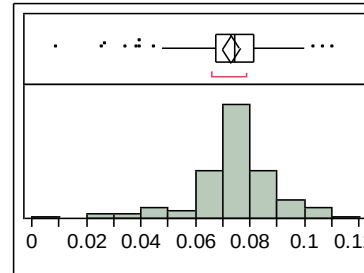
Mean	0.0702157
Std Dev	0.0147753
Std Err Mean	0.001463
upper 95% Mean	0.0731178
lower 95% Mean	0.0673135
N	102

Nuclear Material Portal Monitors MCrank



Mean	0.0712843
Std Dev	0.0135657
Std Err Mean	0.0013432
upper 95% Mean	0.0739489
lower 95% Mean	0.0686198
N	102

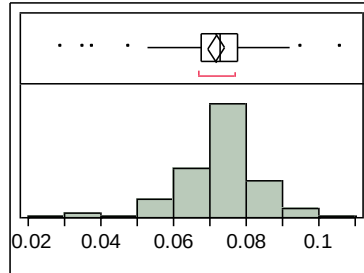
MC PP MCrank



Mean	0.0733039
Std Dev	0.0164496
Std Err Mean	0.0016288
upper 95% Mean	0.0765349
lower 95% Mean	0.0700729
N	102

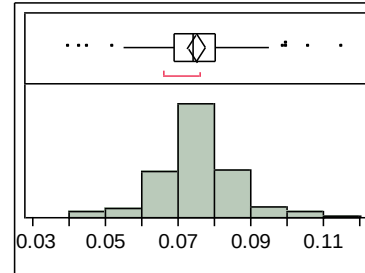
Full Set of Respondents (102)
Material Control (MC) Component Ranking

MC Procedures MCrank



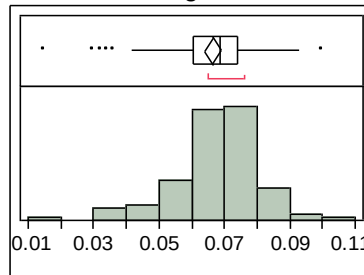
Mean	0.0717353
Std Dev	0.0114675
Std Err Mean	0.0011355
upper 95% Mean	0.0739877
lower 95% Mean	0.0694829
N	102

MC Item Tracking MCrank



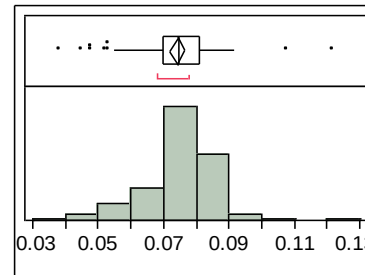
Mean	0.0750392
Std Dev	0.011727
Std Err Mean	0.0011611
upper 95% Mean	0.0773426
lower 95% Mean	0.0727358
N	102

MC Oversight MCrank



Mean	0.0665294
Std Dev	0.0132655
Std Err Mean	0.0013135
upper 95% Mean	0.069135
lower 95% Mean	0.0639238
N	102

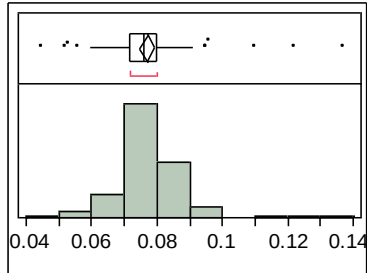
MC Surveillance MCrank



Mean	0.0742647
Std Dev	0.0119396
Std Err Mean	0.0011822
upper 95% Mean	0.0766099
lower 95% Mean	0.0719196
N	102

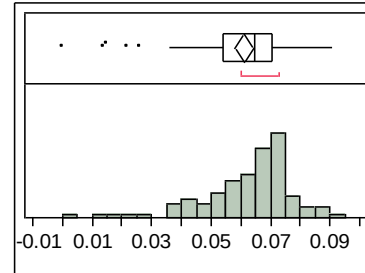
Full Set of Respondents (102)
Material Control (MC) Component Ranking

MC Containment MCrank



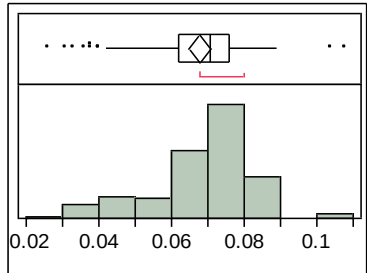
Mean	0.0769902
Std Dev	0.0117932
Std Err Mean	0.0011677
upper 95% Mean	0.0793066
lower 95% Mean	0.0746738
N	102

MC Daily Admin Checks MCrank



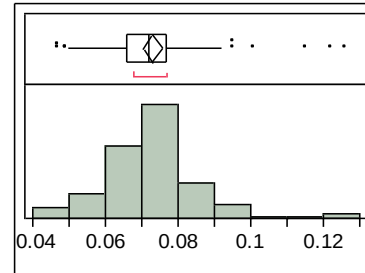
Mean	0.0613725
Std Dev	0.0158557
Std Err Mean	0.00157
upper 95% Mean	0.0644869
lower 95% Mean	0.0582582
N	102

MC Attribute monitors MCrank



Mean	0.0678529
Std Dev	0.0146331
Std Err Mean	0.0014489
upper 95% Mean	0.0707271
lower 95% Mean	0.0649787
N	102

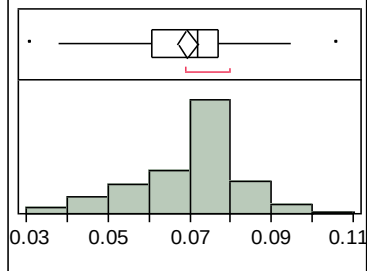
MC Process Control MCrank



Mean	0.072951
Std Dev	0.013562
Std Err Mean	0.0013428
upper 95% Mean	0.0756148
lower 95% Mean	0.0702871
N	102

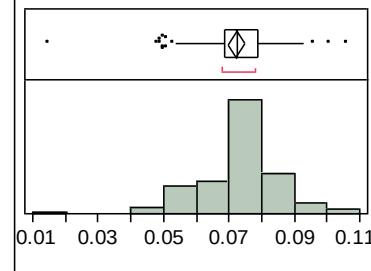
Full Set of Respondents (102)
Material Control (MC) Component Ranking

MC Waste Monitoring MCrank



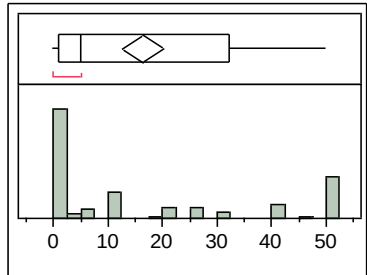
Mean	0.0695882
Std Dev	0.0132787
Std Err Mean	0.0013148
upper 95% Mean	0.0721964
lower 95% Mean	0.0669801
N	102

MC Storage Monitoring MCrank



Mean	0.0722745
Std Dev	0.0128211
Std Err Mean	0.0012695
upper 95% Mean	0.0747928
lower 95% Mean	0.0697562
N	102

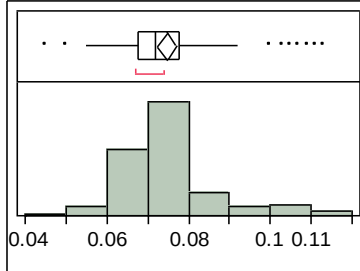
Missed Alarms / False Alarms 3



Mean	16.520588
Std Dev	19.454992
Std Err Mean	1.9263312
upper 95% Mean	20.341911
lower 95% Mean	12.699265
N	102

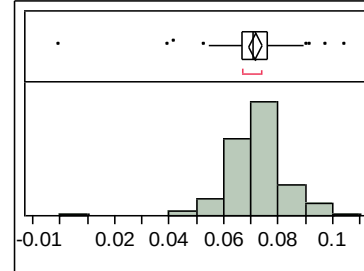
Full Set of Respondents (102)
Material Accounting (MA) Component Ranking

Accounting System MArank



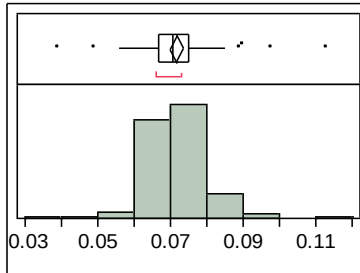
Mean	0.0747549
Std Dev	0.0117688
Std Err Mean	0.0011653
upper 95% Mean	0.0770665
lower 95% Mean	0.0724433
N	102

Physical inventory MArank



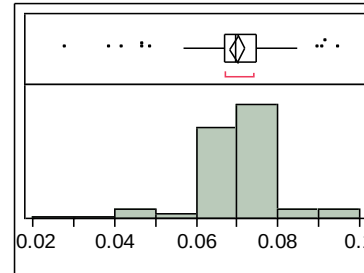
Mean	0.0715098
Std Dev	0.0121098
Std Err Mean	0.0011991
upper 95% Mean	0.0738884
lower 95% Mean	0.0691312
N	102

Measurement Accuracy MArank



Mean	0.0716863
Std Dev	0.0088999
Std Err Mean	0.0008812
upper 95% Mean	0.0734344
lower 95% Mean	0.0699382
N	102

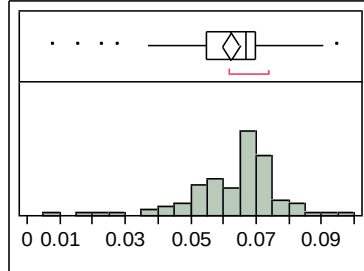
Material Transfer Monitoring MArank



Mean	0.0700588
Std Dev	0.0097115
Std Err Mean	0.0009616
upper 95% Mean	0.0719663
lower 95% Mean	0.0681513
N	102

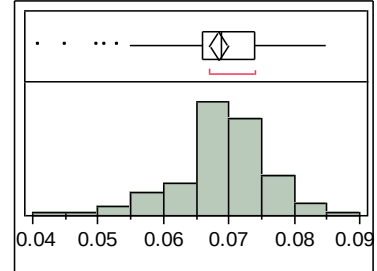
Full Set of Respondents (102)
Material Accounting (MA) Component Ranking

Real Time Accounting MArank



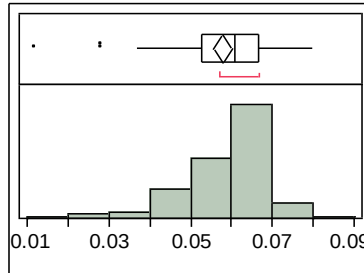
Mean	0.0625294
Std Dev	0.0142104
Std Err Mean	0.001407
upper 95% Mean	0.0653206
lower 95% Mean	0.0597382
N	102

Measurement Control MArank



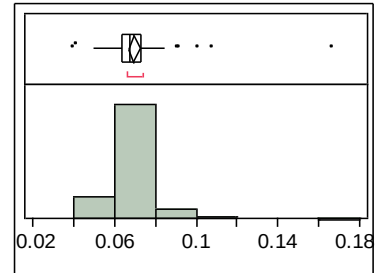
Mean	0.068451
Std Dev	0.0076865
Std Err Mean	0.0007611
upper 95% Mean	0.0699607
lower 95% Mean	0.0669412
N	102

Oversight MArank



Mean	0.058
Std Dev	0.011521
Std Err Mean	0.0011407
upper 95% Mean	0.0602629
lower 95% Mean	0.0557371
N	102

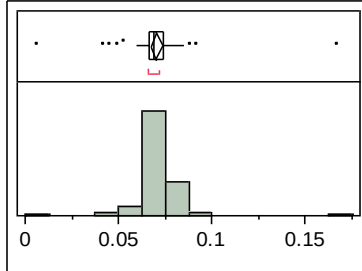
Anomaly Detection MArank



Mean	0.0695098
Std Dev	0.0138584
Std Err Mean	0.0013722
upper 95% Mean	0.0722318
lower 95% Mean	0.0667878
N	102

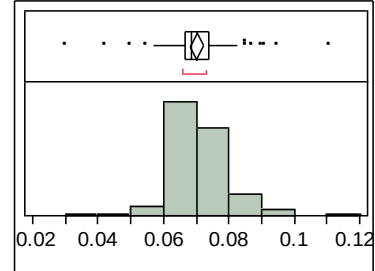
Full Set of Respondents (102)
Material Accounting (MA) Component Ranking

Inventory Detection and Resolution MArank



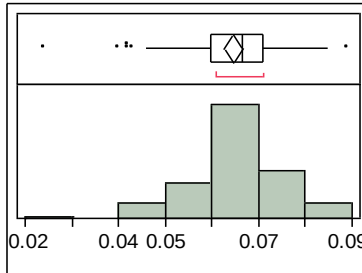
Mean	0.0704706
Std Dev	0.0137263
Std Err Mean	0.0013591
upper 95% Mean	0.0731667
lower 95% Mean	0.0677745
N	102

Shipper Receiver Difference MArank



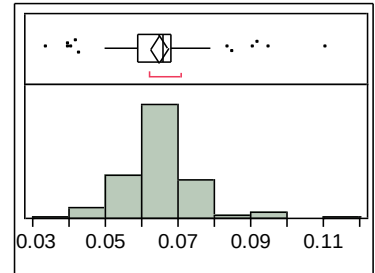
Mean	0.0704118
Std Dev	0.0098175
Std Err Mean	0.0009721
upper 95% Mean	0.0723401
lower 95% Mean	0.0684834
N	102

MA Process Monitoring MArank



Mean	0.0647647
Std Dev	0.0103686
Std Err Mean	0.0010266
upper 95% Mean	0.0668013
lower 95% Mean	0.0627281
N	102

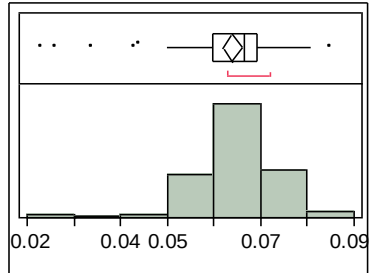
MA Statistical Evaluation MArank



Mean	0.065
Std Dev	0.0111008
Std Err Mean	0.0010991
upper 95% Mean	0.0671804
lower 95% Mean	0.0628196
N	102

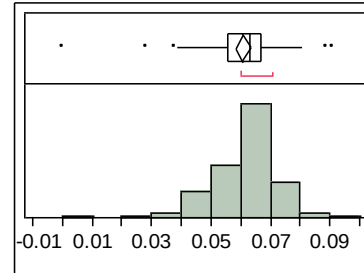
Full Set of Respondents (102)
Material Accounting (MA) Component Ranking

MA Procedures MArank



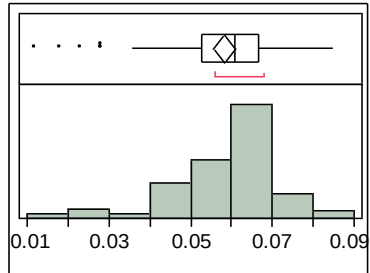
Mean	0.064
Std Dev	0.0099851
Std Err Mean	0.0009887
upper 95% Mean	0.0659613
lower 95% Mean	0.0620387
N	102

MA Holdup MArank



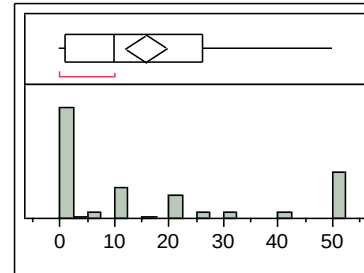
Mean	0.0609412
Std Dev	0.012497
Std Err Mean	0.0012374
upper 95% Mean	0.0633958
lower 95% Mean	0.0584865
N	102

MA Independent Assessment MArank



Mean	0.0581961
Std Dev	0.0135698
Std Err Mean	0.0013436
upper 95% Mean	0.0608614
lower 95% Mean	0.0555307
N	102

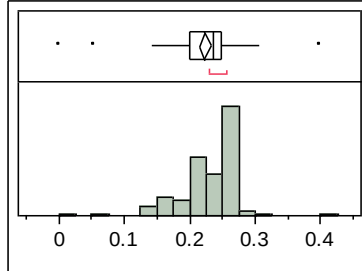
Missed Alarms / False Alarms MA



Mean	15.932353
Std Dev	19.041732
Std Err Mean	1.8854124
upper 95% Mean	19.672504
lower 95% Mean	12.192202
N	102

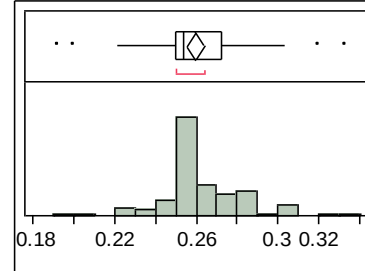
**Full Set of Respondents (102)
Summary Subsystem Ranking**

Plant Management SGrank



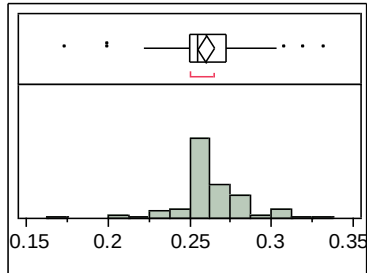
Mean	0.224098
Std Dev	0.0472448
Std Err Mean	0.0046779
upper 95% Mean	0.2333778
lower 95% Mean	0.2148183
N	102

Material Control SGrank



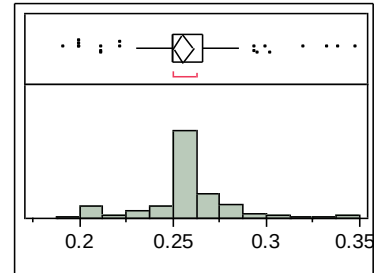
Mean	0.2599902
Std Dev	0.0215749
Std Err Mean	0.0021362
upper 95% Mean	0.2642279
lower 95% Mean	0.2557525
N	102

Material Accounting SGrank



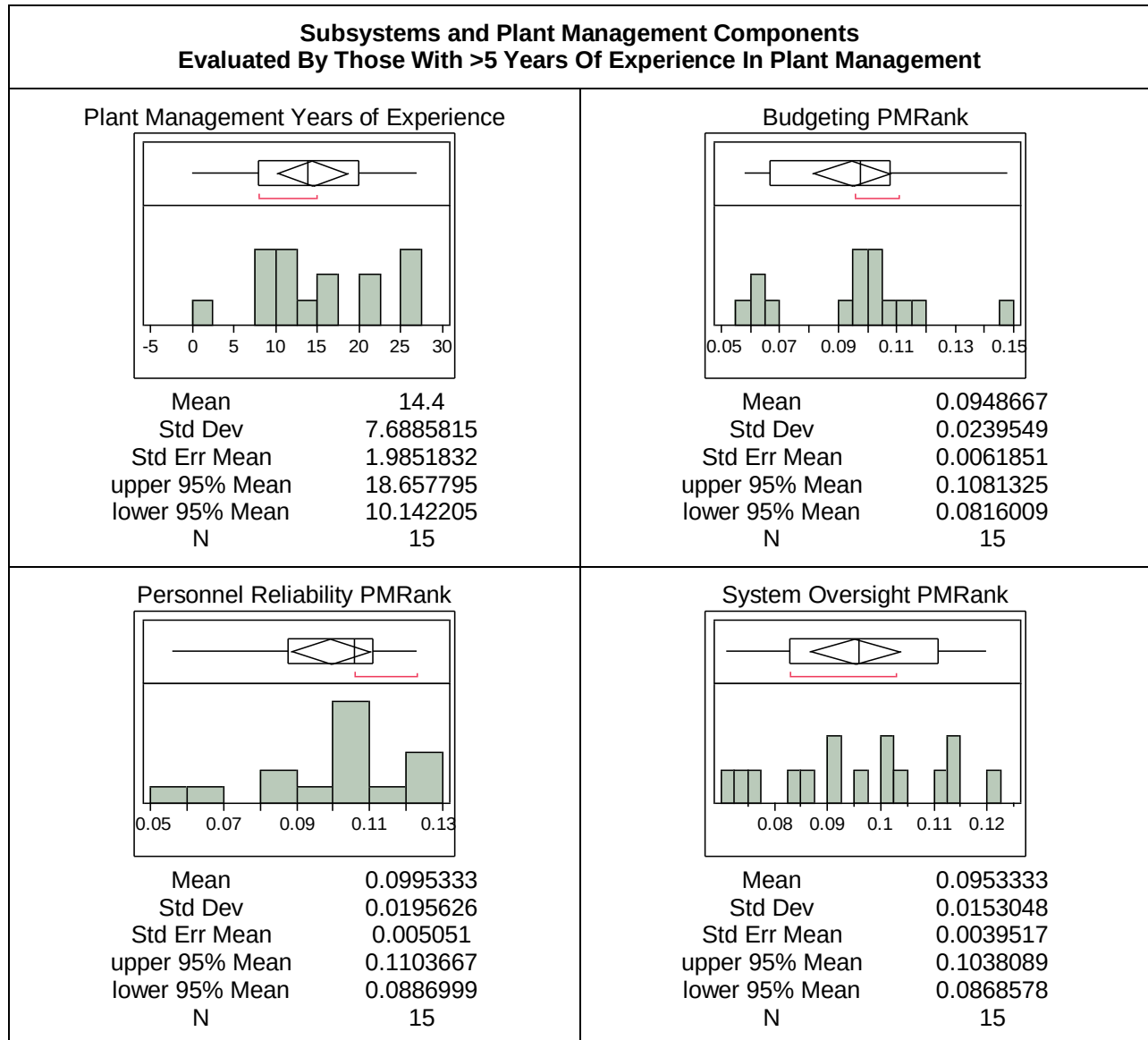
Mean	0.2601471
Std Dev	0.0234573
Std Err Mean	0.0023226
upper 95% Mean	0.2647545
lower 95% Mean	0.2555396
N	102

Physical Protection SGrank



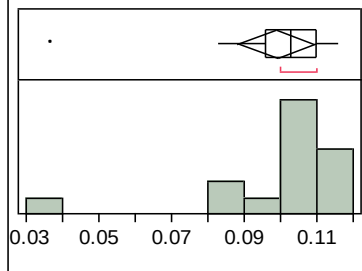
Mean	0.2556961
Std Dev	0.0265281
Std Err Mean	0.0026267
upper 95% Mean	0.2609067
lower 95% Mean	0.2504855
N	102

Section 2 Plant Management Filtered Results



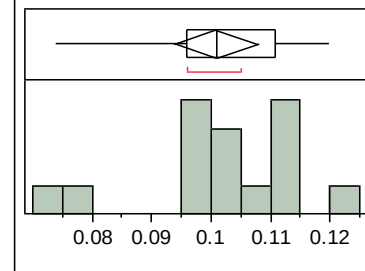
**Subsystems and Plant Management Components
Evaluated By Those With >5 Years Of Experience In Plant Management**

Knowledge of MPC&A PMRank



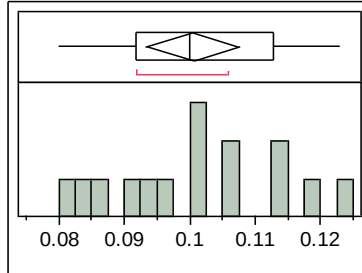
Mean	0.0989333
Std Dev	0.0194035
Std Err Mean	0.00501
upper 95% Mean	0.1096786
lower 95% Mean	0.088188
N	15

Configuration Management PMRank



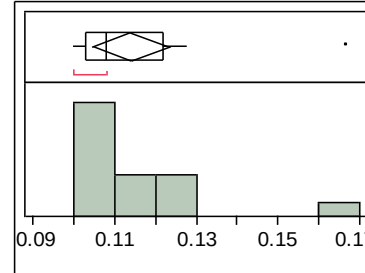
Mean	0.1009333
Std Dev	0.0126122
Std Err Mean	0.0032564
upper 95% Mean	0.1079177
lower 95% Mean	0.093949
N	15

System Documentation PMRank



Mean	0.1006
Std Dev	0.0128663
Std Err Mean	0.0033221
upper 95% Mean	0.1077251
lower 95% Mean	0.0934749
N	15

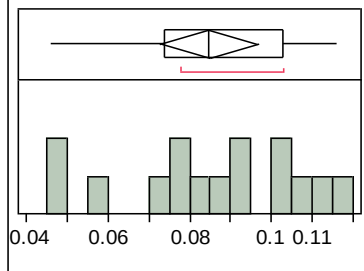
Org Communication PMRank



Mean	0.1140667
Std Dev	0.0170732
Std Err Mean	0.0044083
upper 95% Mean	0.1235215
lower 95% Mean	0.1046118
N	15

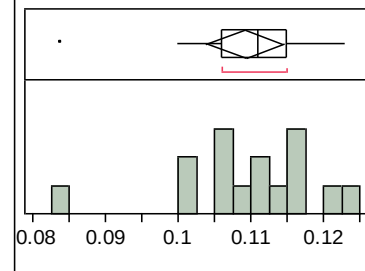
**Subsystems and Plant Management Components
Evaluated By Those With >5 Years Of Experience In Plant Management**

Independent Assessment PMRank



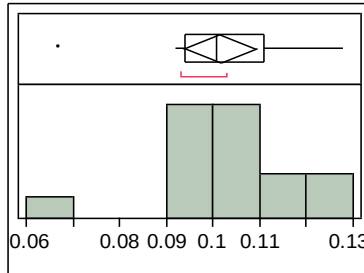
Mean	0.0848
Std Dev	0.0214982
Std Err Mean	0.0055508
upper 95% Mean	0.0967053
lower 95% Mean	0.0728947
N	15

Personnel Training PMRank



Mean	0.1092667
Std Dev	0.0096397
Std Err Mean	0.002489
upper 95% Mean	0.114605
lower 95% Mean	0.1039284
N	15

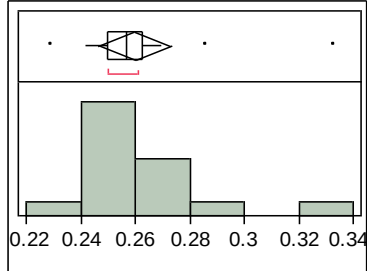
Org Coordination PMRank



Mean	0.1016
Std Dev	0.0142267
Std Err Mean	0.0036733
upper 95% Mean	0.1094785
lower 95% Mean	0.0937215
N	15

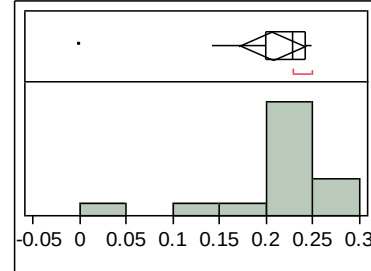
**Subsystems and Plant Management Components
Evaluated By Those With >5 Years Of Experience In Plant Management**

Material Control SGrank



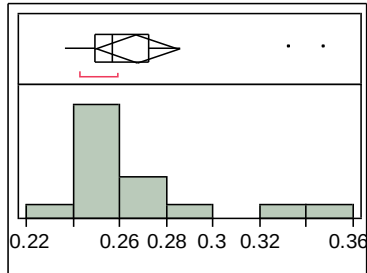
Mean	0.2602
Std Dev	0.0240511
Std Err Mean	0.00621
upper 95% Mean	0.2735191
lower 95% Mean	0.2468809
N	15

Plant Management SGrank



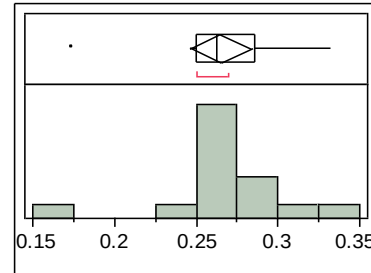
Mean	0.2067333
Std Dev	0.0640373
Std Err Mean	0.0165344
upper 95% Mean	0.242196
lower 95% Mean	0.1712706
N	15

Physical Protection SGrank



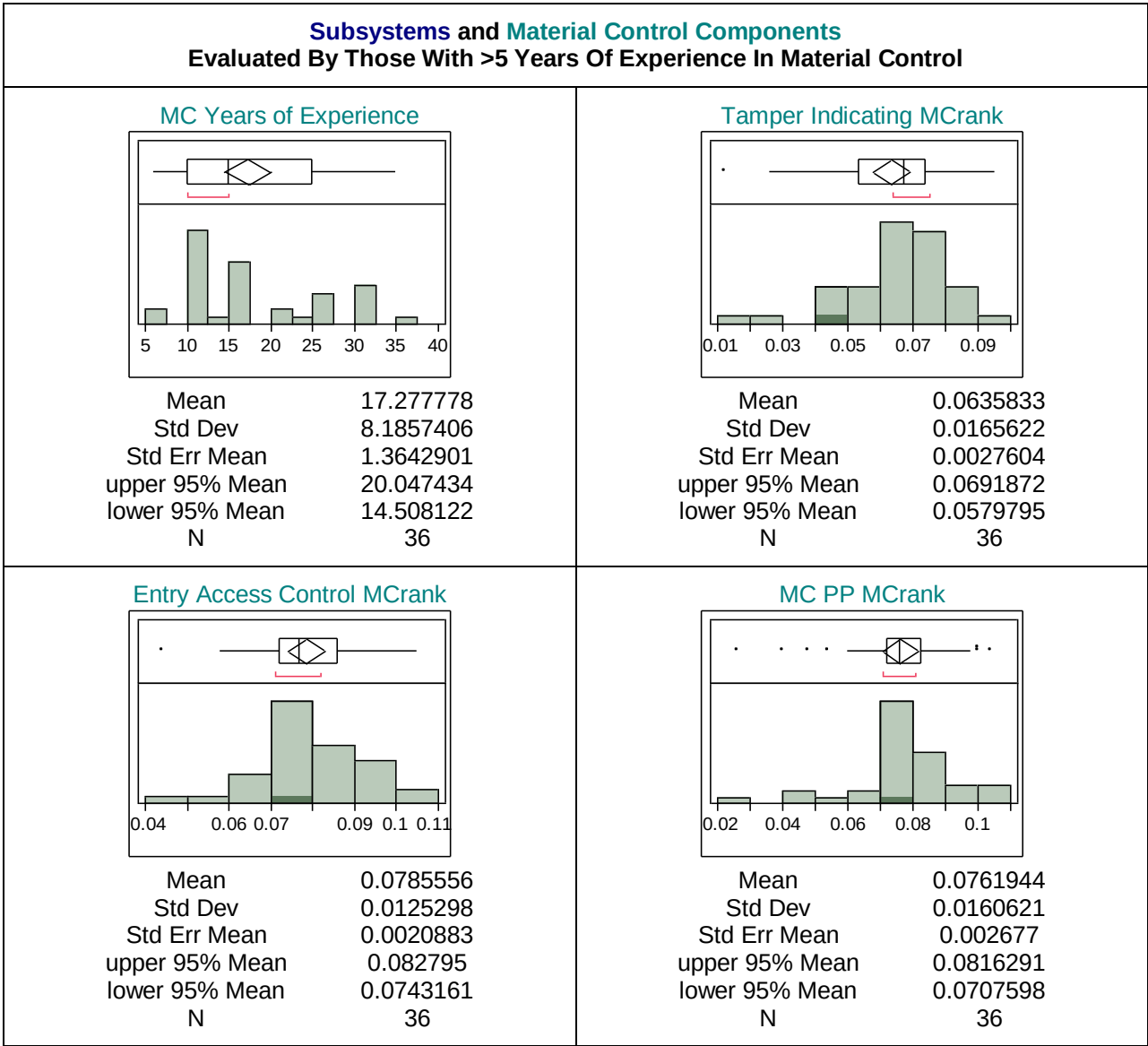
Mean	0.2675333
Std Dev	0.0322288
Std Err Mean	0.0083214
upper 95% Mean	0.285381
lower 95% Mean	0.2496856
N	15

Material Accounting SGrank



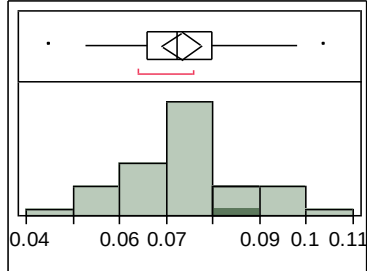
Mean	0.2654
Std Dev	0.0345436
Std Err Mean	0.0089191
upper 95% Mean	0.2845296
lower 95% Mean	0.2462704
N	15

Section 3 Material Control Filtered Results



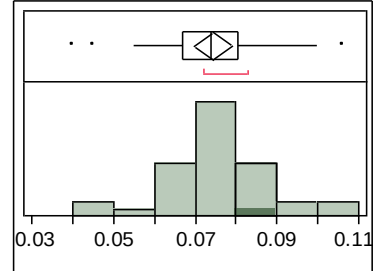
Subsystems and Material Control Components
Evaluated By Those With >5 Years Of Experience In Material Control

Nuclear Material Portal Monitors MCrank



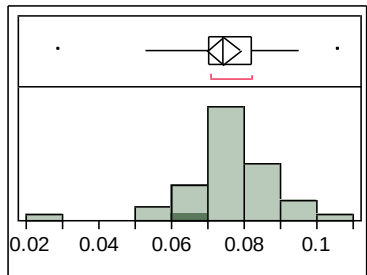
Mean	0.0734167
Std Dev	0.01265
Std Err Mean	0.0021083
upper 95% Mean	0.0776968
lower 95% Mean	0.0691365
N	36

MC Item Tracking MCrank



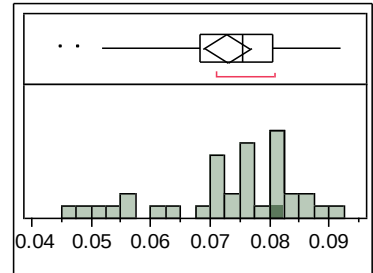
Mean	0.0744167
Std Dev	0.0134449
Std Err Mean	0.0022408
upper 95% Mean	0.0789658
lower 95% Mean	0.0698676
N	36

MC Procedures MCrank



Mean	0.0745
Std Dev	0.0131138
Std Err Mean	0.0021856
upper 95% Mean	0.0789371
lower 95% Mean	0.0700629
N	36

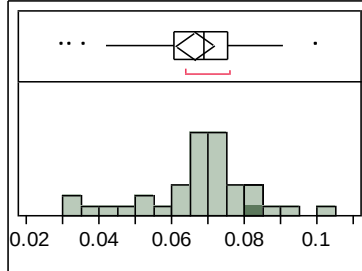
MC Surveillance MCrank



Mean	0.0728611
Std Dev	0.0117331
Std Err Mean	0.0019555
upper 95% Mean	0.076831
lower 95% Mean	0.0688912
N	36

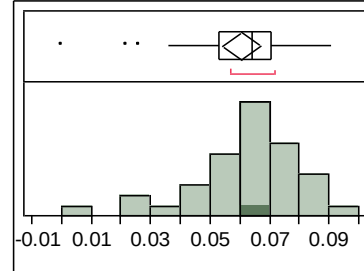
Subsystems and Material Control Components
Evaluated By Those With >5 Years Of Experience In Material Control

MC Oversight MCrank



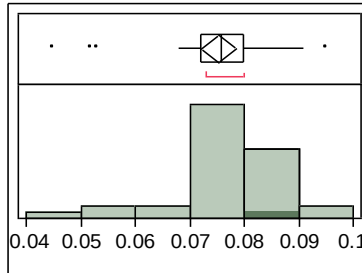
Mean	0.0664722
Std Dev	0.0154541
Std Err Mean	0.0025757
upper 95% Mean	0.0717011
lower 95% Mean	0.0612433
N	36

MC Daily Admin Checks MCrank



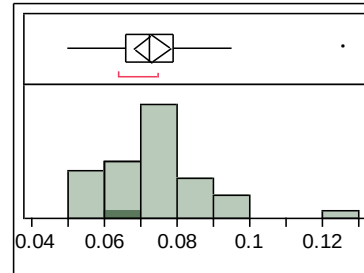
Mean	0.0605556
Std Dev	0.0185679
Std Err Mean	0.0030947
upper 95% Mean	0.066838
lower 95% Mean	0.0542731
N	36

MC Containment MCrank



Mean	0.0753889
Std Dev	0.0097491
Std Err Mean	0.0016248
upper 95% Mean	0.0786875
lower 95% Mean	0.0720903
N	36

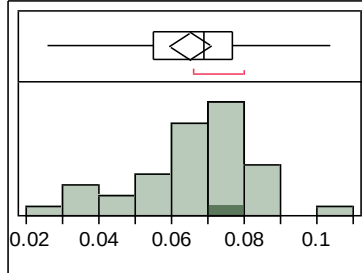
MC Process Control MCrank



Mean	0.0731944
Std Dev	0.0143543
Std Err Mean	0.0023924
upper 95% Mean	0.0780513
lower 95% Mean	0.0683376
N	36

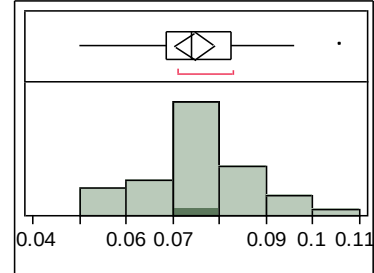
Subsystems and Material Control Components
Evaluated By Those With >5 Years Of Experience In Material Control

MC Attribute monitors MCrank



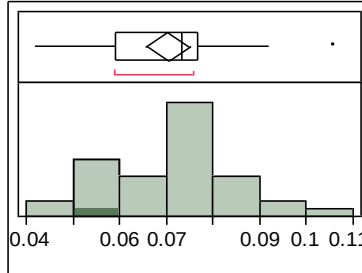
Mean	0.0651389
Std Dev	0.0163588
Std Err Mean	0.0027265
upper 95% Mean	0.0706739
lower 95% Mean	0.0596039
N	36

MC Storage Monitoring MCrank



Mean	0.0746667
Std Dev	0.0122147
Std Err Mean	0.0020358
upper 95% Mean	0.0787995
lower 95% Mean	0.0705338
N	36

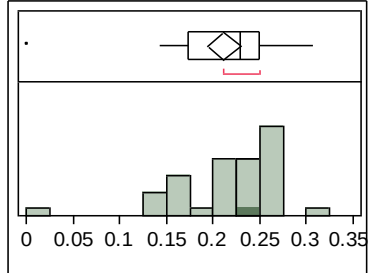
MC Waste Monitoring MCrank



Mean	0.0704167
Std Dev	0.0141973
Std Err Mean	0.0023662
upper 95% Mean	0.0752204
lower 95% Mean	0.065613
N	36

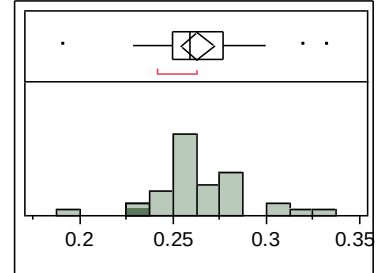
Subsystems and Material Control Components
Evaluated By Those With >5 Years Of Experience In Material Control

Plant Management SGrank



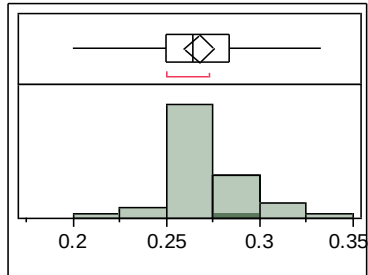
Mean	0.2121944
Std Dev	0.0544677
Std Err Mean	0.009078
upper 95% Mean	0.2306237
lower 95% Mean	0.1937652
N	36

Material Control SGrank



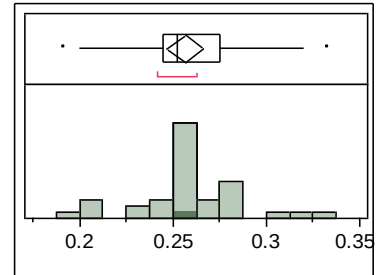
Mean	0.2631667
Std Dev	0.0261233
Std Err Mean	0.0043539
upper 95% Mean	0.2720055
lower 95% Mean	0.2543278
N	36

Material Accounting SGrank



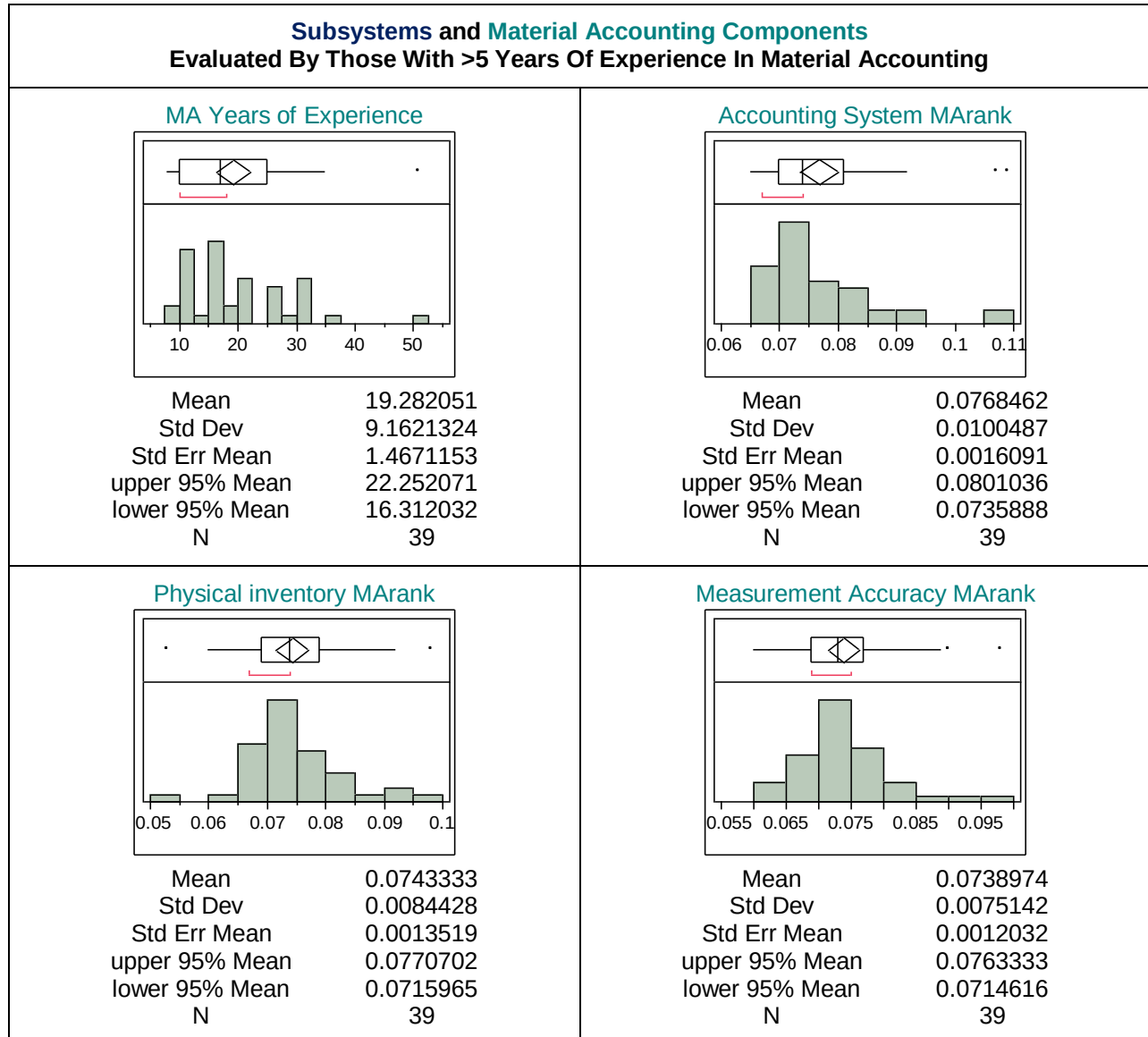
Mean	0.2676944
Std Dev	0.0236387
Std Err Mean	0.0039398
upper 95% Mean	0.2756926
lower 95% Mean	0.2596962
N	36

Physical Protection SGrank

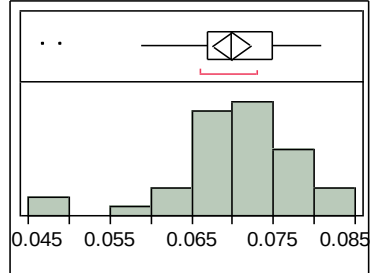


Mean	0.2569167
Std Dev	0.0288596
Std Err Mean	0.0048099
upper 95% Mean	0.2666814
lower 95% Mean	0.247152
N	36

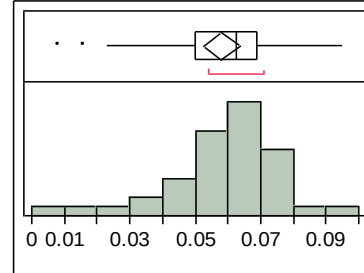
Section 4 Material Accounting Filtered Results



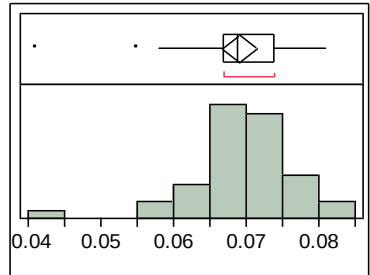
Subsystems and Material Accounting Components
Evaluated By Those With >5 Years Of Experience In Material Accounting

Material Transfer Monitoring MArank

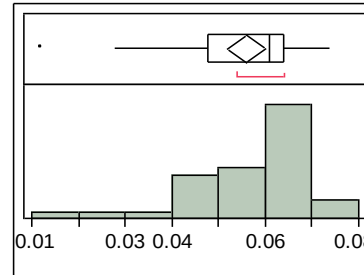
Mean	0.0698974
Std Dev	0.007337
Std Err Mean	0.0011749
upper 95% Mean	0.0722758
lower 95% Mean	0.0675191
N	39

Real Time Accounting MArank

Mean	0.0581795
Std Dev	0.0175287
Std Err Mean	0.0028068
upper 95% Mean	0.0638616
lower 95% Mean	0.0524973
N	39

Measurement Control MArank

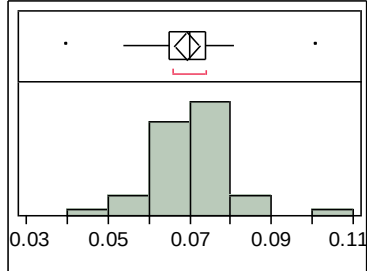
Mean	0.0691538
Std Dev	0.0074782
Std Err Mean	0.0011975
upper 95% Mean	0.071578
lower 95% Mean	0.0667297
N	39

Oversight MArank

Mean	0.0559744
Std Dev	0.0129462
Std Err Mean	0.0020731
upper 95% Mean	0.060171
lower 95% Mean	0.0517777
N	39

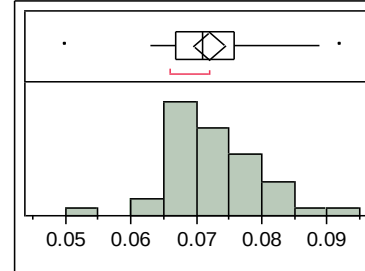
Subsystems and Material Accounting Components
Evaluated By Those With >5 Years Of Experience In Material Accounting

Anomaly Detection MArank



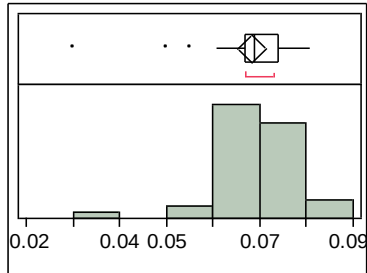
Mean	0.0694872
Std Dev	0.0093383
Std Err Mean	0.0014953
upper 95% Mean	0.0725143
lower 95% Mean	0.0664601
N	39

Inventory Detection and resolution MArank



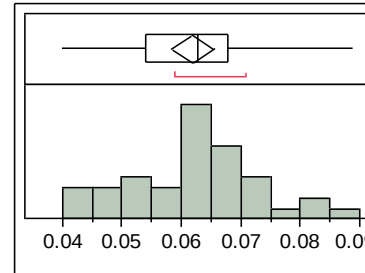
Mean	0.072
Std Dev	0.0075044
Std Err Mean	0.0012017
upper 95% Mean	0.0744326
lower 95% Mean	0.0695674
N	39

Shipper Receiver Difference MArank



Mean	0.0684359
Std Dev	0.0091473
Std Err Mean	0.0014647
upper 95% Mean	0.0714011
lower 95% Mean	0.0654707
N	39

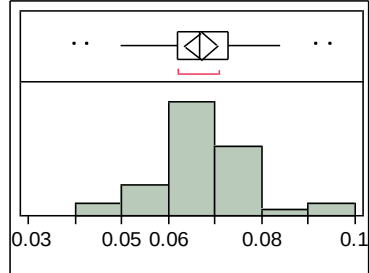
MA Process Monitoring MArank



Mean	0.0618718
Std Dev	0.0110696
Std Err Mean	0.0017726
upper 95% Mean	0.0654601
lower 95% Mean	0.0582834
N	39

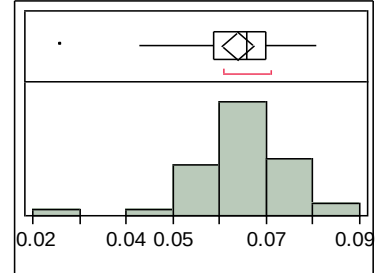
Subsystems and Material Accounting Components
Evaluated By Those With >5 Years Of Experience In Material Accounting

MA Statistical Evaluation MArank



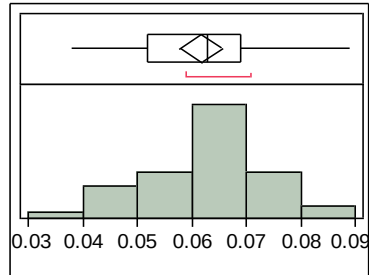
Mean	0.0670513
Std Dev	0.0108748
Std Err Mean	0.0017414
upper 95% Mean	0.0705765
lower 95% Mean	0.0635261
N	39

MA Procedures MArank



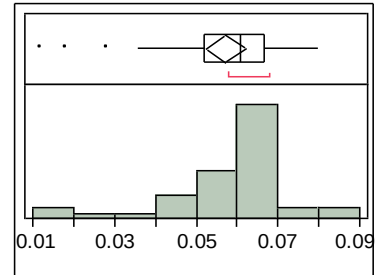
Mean	0.064
Std Dev	0.0103084
Std Err Mean	0.0016507
upper 95% Mean	0.0673416
lower 95% Mean	0.0606584
N	39

MA Holdup MArank



Mean	0.0617436
Std Dev	0.0116703
Std Err Mean	0.0018687
upper 95% Mean	0.0655267
lower 95% Mean	0.0579605
N	39

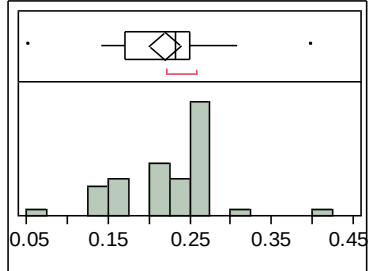
MA Independent Assessment MArank



Mean	0.0573077
Std Dev	0.0151695
Std Err Mean	0.0024291
upper 95% Mean	0.0622251
lower 95% Mean	0.0523903
N	39

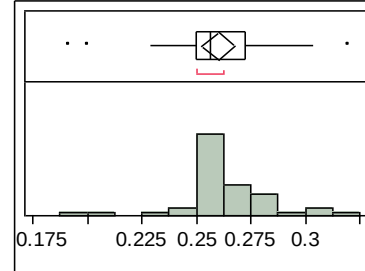
Subsystems and Material Accounting Components
Evaluated By Those With >5 Years Of Experience In Material Accounting

Plant Management SGrank



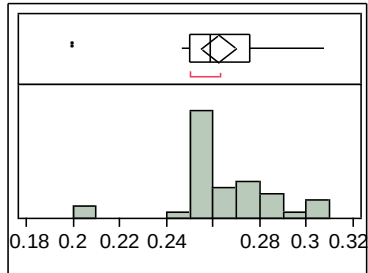
Mean	0.2201538
Std Dev	0.0573395
Std Err Mean	0.0091817
upper 95% Mean	0.2387412
lower 95% Mean	0.2015665
N	39

Material Control SGrank



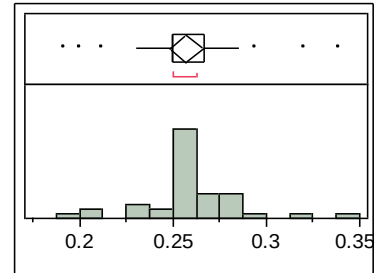
Mean	0.2601795
Std Dev	0.0241551
Std Err Mean	0.0038679
upper 95% Mean	0.2680097
lower 95% Mean	0.2523493
N	39

Material Accounting SGrank



Mean	0.2625641
Std Dev	0.0225456
Std Err Mean	0.0036102
upper 95% Mean	0.2698725
lower 95% Mean	0.2552557
N	39

Physical Protection SGrank

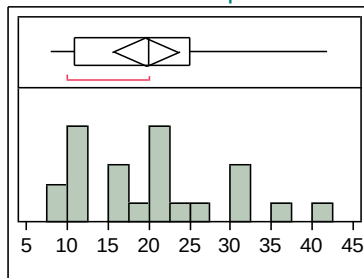


Mean	0.2571538
Std Dev	0.0269859
Std Err Mean	0.0043212
upper 95% Mean	0.2659017
lower 95% Mean	0.248406
N	39

Section 5 Physical Protection Filtered Results

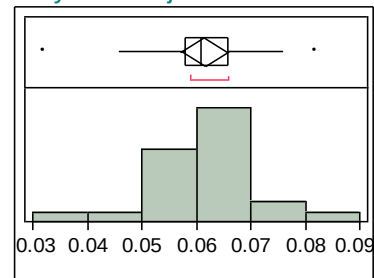
Subsystems and Physical Protection Components Evaluated By Those With >5 Years Of Experience In Physical Protection

PP Years of Experience



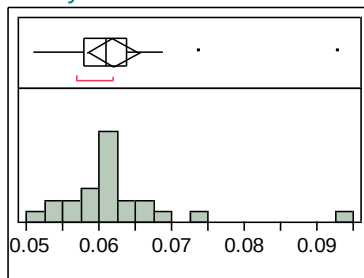
Mean	19.652174
Std Dev	9.3303449
Std Err Mean	1.9455114
upper 95% Mean	23.686918
lower 95% Mean	15.61743
N	23

System Objectives PPrank



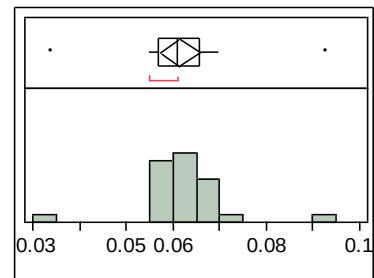
Mean	0.0614783
Std Dev	0.0100266
Std Err Mean	0.0020907
upper 95% Mean	0.0658141
lower 95% Mean	0.0571424
N	23

Facility Characterization PPrank



Mean	0.0620435
Std Dev	0.0084932
Std Err Mean	0.001771
upper 95% Mean	0.0657162
lower 95% Mean	0.0583707
N	23

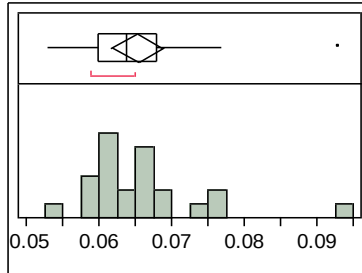
Threat Definition PPrank



Mean	0.0615217
Std Dev	0.0098667
Std Err Mean	0.0020573
upper 95% Mean	0.0657884
lower 95% Mean	0.0572551
N	23

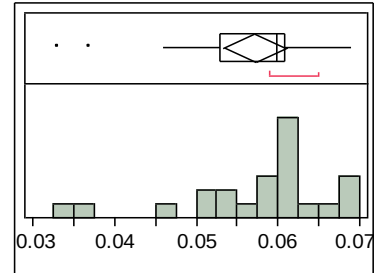
Subsystems and Physical Protection Components
Evaluated By Those With >5 Years Of Experience In Physical Protection

Target ID PPrank



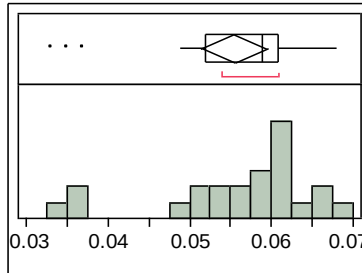
Mean	0.0653913
Std Dev	0.0083161
Std Err Mean	0.001734
upper 95% Mean	0.0689875
lower 95% Mean	0.0617951
N	23

Exterior Intrusion Sensors PPrank



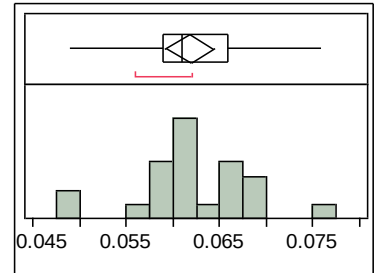
Mean	0.0572609
Std Dev	0.0090263
Std Err Mean	0.0018821
upper 95% Mean	0.0611641
lower 95% Mean	0.0533576
N	23

Interior Intrusion Sensors PPrank



Mean	0.0555217
Std Dev	0.0094958
Std Err Mean	0.00198
upper 95% Mean	0.059628
lower 95% Mean	0.0514155
N	23

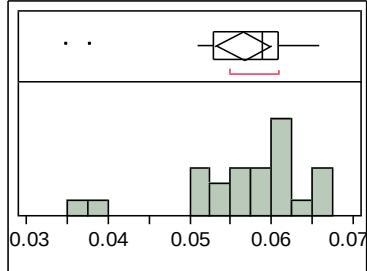
Alarm Assessment PPrank



Mean	0.0618696
Std Dev	0.0060925
Std Err Mean	0.0012704
upper 95% Mean	0.0645042
lower 95% Mean	0.059235
N	23

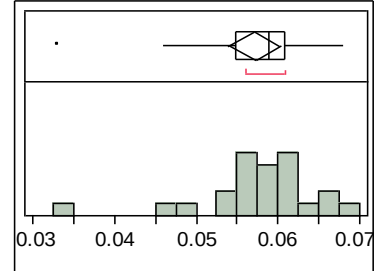
Subsystems and Physical Protection Components
Evaluated By Those With >5 Years Of Experience In Physical Protection

Access Delay PPrank



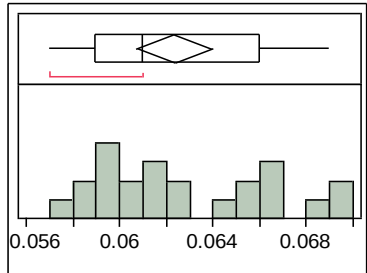
Mean	0.0566522
Std Dev	0.0078196
Std Err Mean	0.0016305
upper 95% Mean	0.0600336
lower 95% Mean	0.0532707
N	23

Secure Comm PPrank



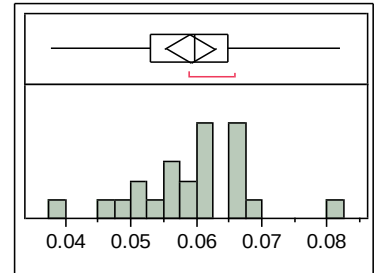
Mean	0.0572174
Std Dev	0.0074098
Std Err Mean	0.0015451
upper 95% Mean	0.0604216
lower 95% Mean	0.0540132
N	23

Alarm Response PPrank



Mean	0.0623478
Std Dev	0.0037247
Std Err Mean	0.0007767
upper 95% Mean	0.0639585
lower 95% Mean	0.0607371
N	23

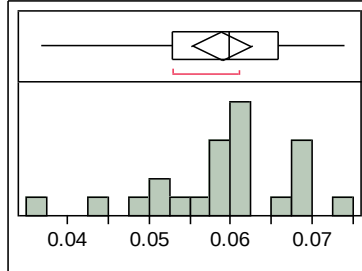
Material Control System PPrank



Mean	0.0590435
Std Dev	0.0089264
Std Err Mean	0.0018613
upper 95% Mean	0.0629035
lower 95% Mean	0.0551834
N	23

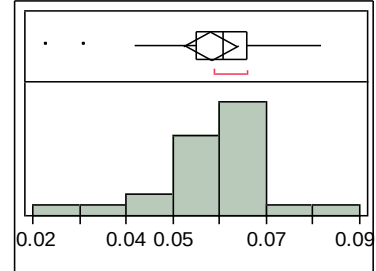
Subsystems and Physical Protection Components
Evaluated By Those With >5 Years Of Experience In Physical Protection

PP System Testing PPrank



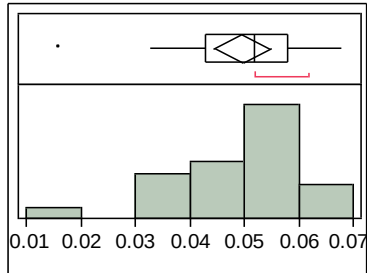
Mean	0.058913
Std Dev	0.0085648
Std Err Mean	0.0017859
upper 95% Mean	0.0626167
lower 95% Mean	0.0552093
N	23

Personnel Reliability PPrank



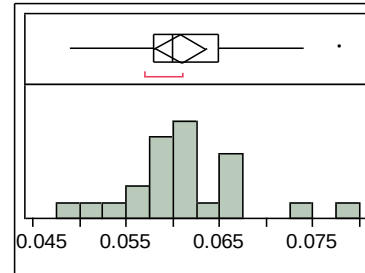
Mean	0.0582609
Std Dev	0.0131433
Std Err Mean	0.0027406
upper 95% Mean	0.0639445
lower 95% Mean	0.0525773
N	23

Independent System Assessment PPrank



Mean	0.0496957
Std Dev	0.0118338
Std Err Mean	0.0024675
upper 95% Mean	0.054813
lower 95% Mean	0.0445783
N	23

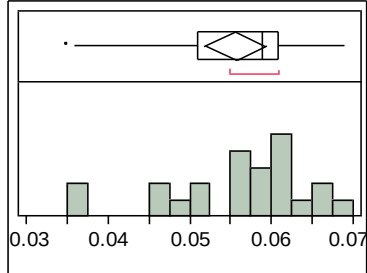
Vulnerability Assessment PPrank



Mean	0.0608696
Std Dev	0.0064899
Std Err Mean	0.0013532
upper 95% Mean	0.063676
lower 95% Mean	0.0580631
N	23

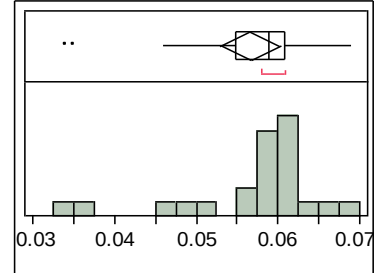
Subsystems and Physical Protection Components
Evaluated By Those With >5 Years Of Experience In Physical Protection

Internal Guard Force PPrank



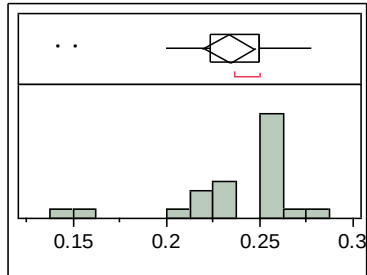
Mean	0.0556522
Std Dev	0.0087158
Std Err Mean	0.0018174
upper 95% Mean	0.0594212
lower 95% Mean	0.0518832
N	23

Special Response Teams PPrank



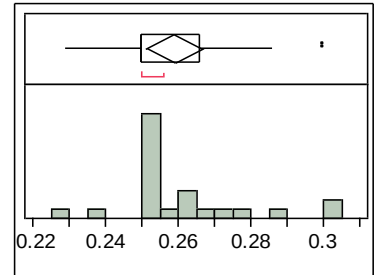
Mean	0.0566957
Std Dev	0.0085729
Std Err Mean	0.0017876
upper 95% Mean	0.0604028
lower 95% Mean	0.0529885
N	23

Plant Management SGrank



Mean	0.2336522
Std Dev	0.0319852
Std Err Mean	0.0066694
upper 95% Mean	0.2474836
lower 95% Mean	0.2198207
N	23

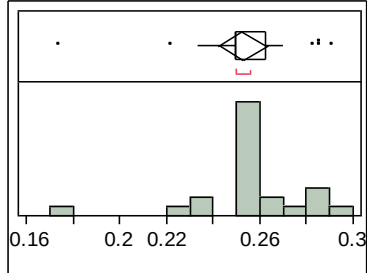
Material Control SGrank



Mean	0.259087
Std Dev	0.0177634
Std Err Mean	0.0037039
upper 95% Mean	0.2667684
lower 95% Mean	0.2514055
N	23

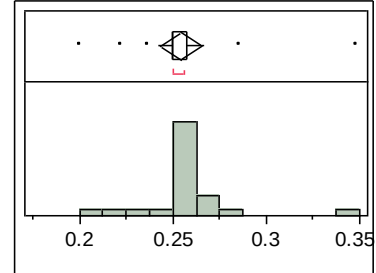
Subsystems and Physical Protection Components
Evaluated By Those With >5 Years Of Experience In Physical Protection

Material Accounting SGrank



Mean	0.2528261
Std Dev	0.0242499
Std Err Mean	0.0050565
upper 95% Mean	0.2633125
lower 95% Mean	0.2423396
N	23

Physical Protection SGrank



Mean	0.2544783
Std Dev	0.0259403
Std Err Mean	0.0054089
upper 95% Mean	0.2656957
lower 95% Mean	0.2432609
N	23

Vita

Cameron W. Coates has worked in the Department of Energy (DOE) Oak Ridge complex since 1980 and is currently assigned to the Global Nuclear Security Technology Division of the Oak Ridge National Laboratory. Mr. Coates received a Bachelor of Science in Chemical Engineering from the University of Virginia and a Master of Science in Systems Engineering from The George Washington University. He completed a four year assignment to the DOE National Nuclear Security Administration Office of International Safeguards in 2004. He is a Senior Member of the Institute of Nuclear Material Management and is an associate editor of Material Control and Accountability for the Journal of Nuclear Material Management. This dissertation is undertaken in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Industrial Engineering.