

SANDIA REPORT SAND2006-7588 Unlimited Release Printed December 2006

Modeling Threat Assessments Of Water Supply Systems Using Markov Latent Effects Methodology

Consuelo J. Silva

Prepared by Sandia National Laboratories Albuquerque, New Mexico 87185 and Livermore, California 94550

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under Contract DE-AC04-94AL85000.

Approved for public release; further dissemination unlimited.

Issued by Sandia National Laboratories, operated for the United States Department of Energy by Sandia Corporation.

NOTICE: This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government, nor any agency thereof, nor any of their employees, nor any of their contractors, subcontractors, or their employees, make any warranty, express or implied, or assume any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represent that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government, any agency thereof, or any of their contractors or subcontractors. The views and opinions expressed herein do not necessarily state or reflect those of the United States Government, any agency thereof, or any of their contractors.

Printed in the United States of America. This report has been reproduced directly from the best available copy.

Available to DOE and DOE contractors from

 U.S. Department of Energy Office of Scientific and Technical Information P.O. Box 62 Oak Ridge, TN 37831

Available to the public from

 U.S. Department of Commerce National Technical Information Service 5285 Port Royal Rd. Springfield, VA 22161

SAND2006-7588 Unlimited Release Printed December 2006

Modeling Threat Assessments of Water Supply Systems Using Markov Latent Effects Methodology

Consuelo J. Silva Corporate Projects Sandia National Laboratories P.O. Box 5800 Albuquerque, New Mexico 87185-MS0937

Abstract

Recent amendments to the Safe Drinking Water Act emphasize efforts toward safeguarding our nation's water supplies against attack and contamination. Specifically, the Public Health Security and Bioterrorism Preparedness and Response Act of 2002 established requirements for each community water system serving more than 3300 people to conduct an assessment of the vulnerability of its system to a terrorist attack or other intentional acts. Integral to evaluating system vulnerability is the threat assessment, which is the process by which the credibility of a threat is quantified. Unfortunately, full probabilistic assessment is generally not feasible, as there is insufficient experience and/or data to quantify the associated probabilities. For this reason, an alternative approach is proposed based on Markov Latent Effects (MLE) modeling, which provides a framework for quantifying imprecise subjective metrics through possibilistic or fuzzy mathematics. Here, an MLE model for water systems is developed and demonstrated to determine threat assessments for different scenarios identified by the assailant, asset, and means. Scenario assailants include terrorists, insiders, and vandals. Assets include a water treatment plant, water storage tank, node, pipeline, well, and a pump station. Means used in attacks include contamination (onsite chemicals, biological and chemical), explosives and vandalism. Results demonstrated highest threats are vandalism events and least likely events are those performed by a terrorist.

ACKNOWLEDGMENTS

I sincerely appreciate the time and efforts that were devoted to this project by many people. Specifically, I thank Vince Tidwell for his guidance throughout this project and Ray Finley for his enthusiastic support of this work. I also thank Troy Grossarth and Nathanael Rackley for all of their efforts in developing the MLE modeling interface. Additionally, I thank my UNM thesis committee members Kerry Howe and Bruce Thomson for accepting this research idea and for their support as the project developed. Lastly, I thank my Sandia co-workers and managers for their encouragement and support while I juggled my education and work responsibilities throughout my degree program.

Funding for this project was provided by Sandia National Laboratories, Laboratory Directed Research and Development (LDRD) Program and SNL's University Part Time Program. Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

CONTENTS

FIGURES

TABLES

1. INTRODUCTION

1.1. Background

There are approximately 160,000 public drinking water systems in the United States. Each of these systems supplies water to at least 15 connections or 25 people. Most people in the US (268 million) get their water from a community water system. There are approximately 54,000 community water systems but seven percent (3,797 systems) serve 81 percent of the people (EPA Factoids, 2004). The ability of these utilities to deliver safe, clean water is essential to public health, economic growth, and quality of life.

In 1996, an executive order (E.O. 13010) on critical infrastructure protection included water supply systems as one of eight national infrastructures vital to the security of the United States. In 1997, the Presidential Commission on Critical Infrastructure Protection identified vulnerabilities in the drinking water sector and identified three attributes crucial to water supply users: water must be available on demand, it must be delivered at sufficient pressure, and it must be safe for use (The President's Commission on Critical Infrastructure Protection, 1997). As a result of these findings, the Environmental Protection Agency (EPA), other federal agencies, water utilities, and state and local governments have taken steps to improve the security of water systems. In 1998, President Clinton issued Presidential Decision Directive (PDD) 63 that established a private/public partnership to put in place prevention, response, and recovery methods to ensure the security of the nation's critical infrastructures against criminal or terrorist attacks. This directive assigned the responsibility for improving preparedness and increasing security of drinking water systems and supplies to the Environmental Protection Agency (The Clinton Administration's Policy on Critical Infrastructure, 1998).

The events of September 11, 2001 had the effect of broadening and accelerating efforts to safeguard the nation's water utilities against terrorism and other threats. In particular, President Bush signed into law the Public Health Security and Bioterrorism Preparedness and Response Act of 2002 (HR3448 Public Law 107-188). This law added new drinking water security and safety requirements and required all community water systems serving more than 3300 people to complete vulnerability assessments and prepare and/or update their emergency response plans. The vulnerability assessments intended to help water utilities evaluate the risk posed by potential threats and identify corrective actions that could reduce or mitigate the consequences of these adversarial actions (Scharfenaker, 2002). These assessments were to serve as a guide to the water utility by providing a prioritized plan for security upgrades modifications of operational procedures, and/or policy changes to mitigate risks. The assessments were intended to be a dynamic process in which the utilities review their vulnerability assessments periodically to account for changing threats or changes to the water system. Furthermore, in December of 2003, President Bush issued the Homeland Security Presidential Directive /HSPD-7, which established a national policy for the federal government to identify, prioritize and protect critical infrastructures as part of homeland security (Homeland Security Presidential Directive, 2003).

1.2. Problem

Water is vital to society. It is required for transporting goods, growing food, and maintaining life. Since water is such a critical asset, it has been used as a means of harming others in times of war or conflict. It has been reported that as far back as the sixth century BC, Assyrians implemented bioterrorism into their war strategy by poisoning enemy wells with rye ergot, a fungus that causes convulsions if ingested (Eitzen 1997). It has also been reported that in 1993 Saddam Hussein of Iraq poisoned and drained the water supplies of Shiite Muslims as a military tool to suppress their opposition to his government (Gleick 1993).

The American Water Works Association Research Foundation completed a study in February 2003 to assemble a database of security incidents, threats and hoaxes involving water systems. This study yielded 264 events; where 193 occurred in North America. (Welter et. al., 2003). Although these instances have occurred over time, it was not until the terrorist attacks in the United States on September 11 2001, that the security of water distribution systems fell under scrutiny. Moreover, in the January 2001 report of the President's Commission on Critical Infrastructure Protection, ongoing water research was characterized as a small effort that left a number of gaps and shortfalls relative to U.S. water supplies (Critical Infrastructure Assurance Office, 2001). It also went on to state, "that gaps exist in four major areas

- Threat/vulnerability risk analysis
- Identification and characterization of biological and chemical agents
- A need to establish a center of excellence to support communities in conducting vulnerability and risk assessment, and
- Application of information assurance techniques to computerized systems used by water utilities, as well as the oil, gas, and electric sectors for operational data and control operations."

Since EPA is the lead agency for protecting the nation's drinking and wastewater infrastructure, they have developed the "Water Security Research and Technical Support Action Plan" (EPA, 2004) that addresses drinking water supply, water treatment, water storage, and drinking water distribution infrastructure. The key research and technical support needs listed in the plan to enhance protection of water infrastructure are:

- "Identify and characterize threats that could be used to disrupt water systems
- Develop methods for detecting and monitoring contaminants in water
- Create rapid screening technologies for the identification of unknown contaminants
- Improve detectors and early warning systems for water distribution and collection systems
- Enhance models for contaminant transport in pipes and distribution systems
- Test and evaluate the performance of sensors and biomonitors
- Refine fate and transport information for contaminants in water
- Develop treatment or inactivation techniques for water contaminants
- Evaluate and improve decontamination and disposal techniques for contaminated materials and equipment
- Establish contingency planning and infrastructure backup procedures
- Improve methods for assessing risks to the public from water contamination
- Enhance risk communication and information sharing among individuals and organizations dealing with a threat of attack
- Provide training and exercises that enhance preparedness, response, and mitigation to water system threats or attacks"

Thus, defining the threat to water systems was not only listed in the Critical Infrastructure Commission as a major issue, it was also the first action item listed for the EPA in its plan to make water systems safer.

Furthermore, there are currently no federal standards or agreed upon industry best practices within the water infrastructure sector to govern readiness, response to security incidents, or recovery. EPA is not authorized to require water infrastructure systems to implement specific security improvements or meet particular security standards, although efforts to develop voluntary protocols and tools have been ongoing (Copeland & Cody, 2005).

There have been some efforts to develop methods to determine vulnerabilities of water utilities; however, the actual threat to water utilities has not been determined. Threat assessment is a procedure for determining the credibility or seriousness of a threat. It attempts to determine the likelihood that a threat will be carried out and the likelihood that the system/asset being attacked is able to defend against the attack (Skiba & Peterson, 2003). Conversely, a vulnerability assessment evaluates the susceptibility to potential threats and identifies corrective actions that can reduce or mitigate the risks of serious consequences from adversarial actions and it considers risks posed to the surrounding community because of the attack (EPA, 2002). The following section provides information on the current approaches to these areas.

1.3. Current Approaches to Threat Assessment

Following the presidential commission's findings, Haimes and others (1998) reviewed needs and opportunities to reduce the vulnerability of public water systems to willful attack. They developed a hierarchical holographic model (Haimes, 1981) to better understand the complexity and interconnectedness that characterizes the security of water distribution systems. This model was also used to explore different approaches to hardening (in terms of security, robustness, resilience, and redundancy) water distribution systems against attack. Subsequent to this work, Ezell and others introduced an infrastructure risk analysis model (Ezell et al., 2000a) and applied it within the context of a municipal water distribution system (Ezell, 2000b). The model provides an analytical methodology for quantifying risk that involves decomposition of utility operations along the dimensions of function, component, structure, state, and vulnerability. Potential threats were identified through scenario modeling, while conditional and expected losses for each scenario are calculated via Asbeck and Haimes's (1984) partitioned multiobjective risk method.

Other models that have been developed include Haestad Methods contaminant transport model WaterSAFE[®] (Haestad Methods, 2005) and the US Air Force H₂OMAP (Boulos, 2002). These models focused on contaminant transport, identifying customer impacts, determining

contaminant origination, flushing techniques, fire flow availability, and methods for isolating the contaminant. These models do not determine threat assessments, but how to respond to events that might occur. Moreover, since these models only considered the actual system components of a water utility they do not take into account other factors that can lead up to a successful attack. For example, how aware (or unaware) the water utility personnel are with respect to the security of the utility. Finally, the models depend on probabilities of events where there is little data associated with events, therefore making actual probabilities difficult to quantify.

To assist the water utility industry in conducting the vulnerability assessments that were required by the Public Health Security and Bioterrorism Preparedness and Response Act of 2002, EPA in partnership with the American Water Works Association Research Foundation and Sandia National Laboratories developed the Risk Assessment Methodology for Water Utilities (RAM- W^{TM}) (AWWA, 2002). This performance-based, consequence-driven, risk–management tool aids in identifying system vulnerabilities, identifying critical facilities and assets, and determining the level of protection to which the security system should be designed. This approach is performance-based meaning that it evaluates risk based on the effectiveness of the security system against malevolent acts. Central to the methodology is the evaluation of risk,

$$
R = P_a (1 - P_e) C \tag{1}
$$

where *R* is risk, P_a is the probability of attack, P_e is the probability of system effectiveness, and *C* is the consequence. Operationally, risk is determined through a coordinated consequence analysis (to define *C*) and threat assessment (for quantifying the term $P_a(1-P_e)$). According to this approach, risk assessment is performed primarily through a process of expert elicitation in which values for each term in the risk equation are quantified according to a structured and defined on a scale of high, medium or low. In this way, $RAM-W^{TM}$ is designed to facilitate comparative analyses relying on relative rankings determined from the risk equation. However, in RAM-W, the P_a value is always set to one, therefore the overall equation is driven only by consequence and system effectiveness and not the actual likelihood of an attack. This is important to note, since this methodology assumes that an attack will always occur.

Even before the events of September 11, 2001, water utilities were performing risk-based analyses to quantify vulnerabilities to natural events such as earthquakes, floods and other natural disasters. Various methodologies are available for performing these risk assessments (e.g., Killuru, 1996; Levitt, 1997). Recently, the American Water Works Association Research Foundation conducted a comprehensive review of the experience base with water utility disasters and offered guidance for risk management and analysis (Grigg, 2003). "The project showed that available methods for risk analysis are in limited use by utilities because of a lack of data on threats and vulnerabilities and a lack of training and priority within utilities." Similar experiences have been encountered by water utilities performing vulnerability assessments, in which relevant threat data has been difficult to obtain. In turn, the utilities are faced with decisions concerning risk reduction programs, potentially costing millions of dollars, driven by this highly ambiguous data (Danneels and Finley, 2004).

These results underscore the difficulty in conducting the threat assessment part of any risk-based analysis. Although the governing relation for threat assessment is quite simple, (*Pa(1-Pe*) term in Equation 1), defining the associated terms is difficult. The problem is insufficient experience and data for quantifying the associated probabilities; particularly in the case of willful attacks.

1.4. Alternative Approach to Risk Assessment

As discussed in the previous section, one of the primary concerns with the current methods of conducting vulnerability assessments is that there is little data available to determine actual event probabilities. For this reason, these models assume P_a , the probability of attack value, is one. That is, the event will occur. Furthermore, these models do not take into account outside effects that can contribute to whether an attack is successful.

This research presents a *possibilistic* approach to threat assessment using Markov Latent Effects (MLE) as an alternative to the traditional probabilistic modeling methodologies. MLE modeling is a convenient threat assessment framework for analyzing subjective metrics within a quantitative, repeatable, and defensible process. Moreover, MLE modeling provides a means of calculating P_a . This type of modeling can help water utilities determine what threat level they want to protect to and what assets have the highest threats. This type of modeling may also help determine best practices and industry standards for water utilities to follow for protection of their systems.

1.5. Objectives

The purpose of this research is to develop Markov Latent Effects (MLE) modeling as an approach to conduct threat assessments for water systems. Using a Markov Latent Effects model to analyze threats allows one to visualize the network of events that must coincide for a threat to achieve its intended purpose. Since there is not enough data to determine absolute probabilities, the MLE method is a convenient threat assessment framework for analyzing subjective metrics within a quantitative, repeatable, and defensible process. These qualities of MLE modeling contribute to a defensible decision making process that can be used by water utilities to determine where to spend limited funds on infrastructure upgrades, security improvements, etc. Furthermore, MLE modeling allows for the term P_a (Possibility of Attack) to be calculated rather than setting it directly to one. This gives a more realistic threat assessment value that can be used in the overall risk assessment equation, which enables risk to not only be driven by the consequence of the attack but also the other defining terms.

First the fundamental underpinnings of MLE modeling will be described. An MLE model is then developed within the context of a water distribution system and presented. Finally, application of this approach is demonstrated on a U.S. water utility.

2. METHODS OF APPROACH

The possibilistic framework adopted for water supply system threat assessment is Markov Latent Effects modeling (Cooper, 2004). The following sections provide the theoretical background underpinning the MLE model.

2.1. MLE Background

The Markov Latent Effects approach had its genesis in the mid-1990s, following a series of accident and security-breach investigations. These investigations revealed that there were strong cultural (e.g., lack of personal responsibility, poor communication, failure to address problems) and environmental (poor working conditions, time constraints, pressure to "look good") factors contributing to the events. Since these factors were generally present at a time well before the occurrence of the event, they were termed "latent effects" (J. Reason, 1997). Although the accident and security-breach investigations identified the need for mathematical treatment of these latent effects, their subjective nature defied conventional analysis. To address this need, Cooper (2004) combined the concept of latent effects with a chained subjective analysis methodology to formulate the "Markov Latent Effects" modeling framework, named after A. A. Markov, who explored (in the late $19th$ century) the formal mathematical role of a chain of occurrences in determining subsequent events (Feller, 1957). MLE was originally developed for assessing the effects of organizational and operational factors on high consequence system safety for the Federal Aviation Administration (Cooper, 1999).

In order to give a systemic structure to the approach, a top-down mathematical decomposition strategy is employed. This strategy enables determination of the most appropriate items to be measured and aggregated as imprecise subjective numbers through possibilistic mathematics. The resulting metrics provide a reference point for assessment and for aiding management decisions. Having a mathematical model to run test cases facilitates predictive exercises and helps convey the top-down system perspective, while mathematically portraying the results.

2.2. Decomposition Factors

The obvious focus in any operation is on the results of the operation itself. However, lying behind these operations are modes in which the operation was set up and implemented, the management of the operation, and the environment. A "top-down" approach inverts the focus to make sure that potentially important hidden factors are identified and not neglected. This is best accomplished according to the following steps:

- 1. Recognize environmental constraints and threats
- 2. Determine the management philosophy that is established within that environment
- 3. Evaluate the working conditions
- 4. Assess the risks inherent in actual operation. (Cooper 2001).

This approach helps better identify risks, but more importantly, helps determine *why* problems might arise (a forward-looking approach) and helps identify what can be done to improve overall integrity (similar to an in-depth root cause and correction analysis, but also accommodating hypothetical events).

The MLE model has a performance-focused approach built around the timing of latent effects. These latent effects are identified through decomposition of the complex threat system into subsystems that are more manageable or decision elements that trace a particular threat from its inception to the point of consequence (Figure 1).

Figure 1. MLE Decomposition.

In this way, the decomposition helps visualize the network of events that must coincide for a threat to achieve its intended consequence. For example, consider a contamination event targeting a treated water storage tank that requires the following set of events: utility budget cuts limit security upgrades; failure of utility to protect sensitive system information; poor maintenance practices by utility; assailant acquires critical system information; assailant training and mission preparation; defeat of the limited physical security protecting the tank; and, undetected breach of the tank hatch due to inoperable alarm. In this way, decomposition provides a basis for identifying credible threats and for visualizing all the latent effects that contribute to the success of a threat event.

A latent effect is an occurrence, condition, or behavior that does not necessarily cause an immediate problem, but can combine later with other occurrences, conditions, or behaviors. Latent effects are represented sequentially because of the chained nature of the decomposition. As in the example above, a breach of the water supply tank was made possible by a sequence of latent effects; specifically, failure of the utility to make security upgrades, protect sensitive information and maintenance of the intrusion alarm system.

2.3. Decision Element Structure

Each decision element identified in the decomposition process represents a single factor that influences the likelihood that a threat will occur. Each decision element produces an output subject to a set of inputs (Figure 1). These inputs include direct effects and/or latent effects.

Direct effects are inputs unique to that decision element. Each direct effect is subjectively assigned an attribute value. Attribute values are qualitative choices, mapped on a scale of 0 (very weak) to 1 (very strong) that reflect the strength of the relationship between the direct effect and the decision element. Latent effects represent the influence that one decision element imposes on another. Their attribute values are simply represented by the output value of the "latent" decision element.

Along with an attribute value, each input is assigned a weighting factor. These weighting factors reflect the ability of an input, whether a direct or latent effect, to influence the decision element relative to all other attributes contributing to that element. As such, the aggregated sum of the weighting factors contributing to an individual decision element must equal one.

2.4. Data Aggregation

Data aggregation allows attributes to be combined to derive various sorts of information, such as combined ratings of subsystems or the entire system, and trends information. The strategy extends naturally to decision analysis, where decision aids must be developed for assessing system integrity, determining the need for operational restrictions, and selecting among alternative approaches or forensic hypotheses. Since measurements can lead to assessments when compared to norms or acceptance margins, subsequent analytical methodology and information presentation can contribute to a structured approach for defensible decision-making.

According to the discussion above, care must be given to the selection of the prescribed data aggregation scheme; that is, the manner by which the observed metrics (i.e., attribute values) are combined to yield quantitative information toward the decision of interest, in this case, the information defining the credibility of a particular threat. For each decision element, input values are aggregated to provide an assessment score, and this process continues for each decision element until an overall threat assessment score is derived. There are varieties of weighted sums that can be employed, each having a different influence on the outcome. The choice of the weighed sum depends on one's confidence in the data and the nature of the decision process. One form that has been used is termed "soft aggregation." An example of a soft aggregation weighted sum is:

$$
WS_j = \frac{1}{1 + e^{-a(\sum_{i=1}^{n} w_i x_i - b)}}
$$

(2)

where WS_i is the weighted sum for the *j*th decision element, *x* is the attribute value for direct

inputs or latent effects, *w* is the weighting factor for direct input or latent effect,
\n(where
$$
\sum_{i=1}^{n} w_i = 1
$$
)

= *i* 1

n is the number of direct inputs and latent effects contributing to the decision element. The "*a*" is the slope parameter the shift parameter is "*b*".

Figure 2 depicts the soft aggregation scheme of Equation 2, using weights (*wi*) equal to 0.5 and attribute values of zero and one, $x_1=0$ and $x_2=1$. The shift parameter (*b*) in this graph is 0.5 and the slope parameters (*a*) range from 5.5 to 30. These values are specific to those used in the model and will be explained in further detail in the following sections.

Figure 2. Soft Aggregation Scheme

An alternative to soft aggregation is the simple linear weighted sum where the variables are the same as above.

$$
WS_j = \sum_{i=1}^n w_i x_i \tag{3}
$$

Figure 3. Soft Aggregation vs. Linear Aggregation

2.5. Model Implementation

With the MLE model, assessments are performed for specific threats by inputting attribute values that are specific to the threat. As attribute values are assigned, primarily through expert elicitation, care must be taken to reduce subjective variation in the attribute value to a minimum. That is, the process must be repeatable; it must be assured that similarly qualified people seeing the same situation would record similar, but not necessarily exactly the same, results. Repeatability is accomplished by specifically defining the numerical representation of the qualitative information during the elicitation process. This is accomplished by using elicitation guides (e.g., Figure 4) during the elicitation process.

Direct Input Variable: On Site Security Elicitation Question: What is the level of physical security monitoring at a site? *Enter any number or range of numbers between 0 and 1 to indicate a qualitative judgment (or range of possible judgments) of the strength of the relationship. 0 represents extremely weak, and 1 represents extremely strong Attribute Value* 0.0 to 0.3 No direct monitoring of the site. Site is remote and is not visited by security personnel. 0.3 to 0.5 Limited monitoring occurs at the site. Security personnel visit site at random intervals. 0.5 to 0.7 There is limited direct monitoring of site. Security checks site regularly during off-hours. 0.7 to 0.9 Security personnel on site 24 hours a day but is only stationed at the entrance of facility, not at site. 0.9 to 1.0 Direct monitoring of the site occurs around the clock by security personnel. Security personnel are on site 24-hours and are constantly patrolling facility.

Figure 4. Sample Elicitation Guide

Issues of uncertainty can be addressed within the MLE modeling framework. Uncertainty inherent to the Markov inputs arises from two sources. One is that a person entering an input may be unsure of the value. Another is that a collection of people collaborating on the input may not agree on exactly the same value regardless of the amount of guidance provided. To capture this uncertainty, an attribute value can be represented by a range of numbers rather than a single deterministic value. This suggests that the resulting output will be in the form of a range of numbers rather than a single value. However, using Microsoft Excel as the interface for the current MLE model, ranges of values cannot be entered within one scenario. Therefore, ranges of values can be entered by running scenarios multiple times.

Other important features of MLE modeling are the Importance and Sensitivity metrics. The Importance metric allows a user to identify those features that most significantly contribute to the success of a threat. The Importance metric is simply calculated by deriving the difference between the output value with the input as entered and the output value when one of the input values has been set to zero. By repeating this process the input with the greatest influence on the output can be found. Similarly, the Sensitivity is calculated as the difference between the output

value with the input as entered and the output value when one of the input values has been set to one. The Sensitivity metric measures the potential for improvement in that input to result in a measurable improvement in the total result.

To summarize, an MLE threat assessment model consists of a network of decision elements and inputs as defined through the decomposition process. The model is populated with a set of fixed weighting factors and a prescribed data aggregation scheme. Assigning corresponding attribute values for direct effects (with help from the elicitation guides, e.g., Figure 4) and aggregating the values over all the decision elements provides an overall threat assessment score. The resulting assessment score provides a measure of the credibility of that specific threat scenario. By consistently following the same analysis scheme, assessment scores calculated for different threat scenarios can be compared and ranked to identify the most credible threat. The threat assessment scores range between 0 and 1. A threat assessment score of 1, for example, would mean that the system has the highest likelihood of being attacked and the system effectiveness against the attack is nonexistent. Conversely, a threat assessment value of 0 means that the system can defeat the threat and the possibility of successful attack is zero.

3. MLE WATER MODEL

3.1. Purpose

To define the MLE model in more specific terms, specific application to threat assessment of water distribution systems is made. The purpose of this model is to calculate in a quantitative, repeatable, and defensible process, threat assessment values for a water utility system based on different scenarios. Each scenario consists of three parts, the assailant, the weapon, and the asset. For example, an insider (the assailant) who wants to contaminate (weapon) a water storage tank (asset).

The MLE water model is based on literature reviews of other vulnerability assessment-type models, such as RAM-W, Yacoiv and Haimes models and the Cooper MLE aviation model. It was also developed with the help of individuals from the EPA, and multiple organizations within Sandia National Laboratories. A literature review was not only conducted on vulnerability assessment models, but also on legislation before and after the September 11, 2001 terrorist attacks in the US. Furthermore, literature on human behavior threat assessment models developed by the Department of Justice and United States Secret Service for assassins, weapons, and other targeted violence was researched. After reviewing the literature, a broader understanding of assailants, water systems, and weapons was obtained. This aided in determining the decision elements needed in the model and their significance in where to place them within the model framework.

3.2. Overview

The MLE threat assessment model for water distribution systems is formulated within a fully generic context; that is, it has direct application to almost any water utility. The model is adapted to a particular utility by the utility specific attribute values input into the model by water utility personnel. The MLE model structure is presented in Figure 5 and a description of each direct input variable is given in Table 1. Additional discussion regarding each of the elements and their location in the model is discussed in the following sections.

Grey boxes represent Direct Effects/Direct Inputs and white boxes represent Decision Elements. Weighting factors for each are shown by corresponding color.

Table 1. Description of the direct inputs to the MLE model

 \mathbf{r}

3.3. MLE Water Model Structure

The MLE model is a top down approach based on the threat assessment equation, $P_a(1-P_e)$, (from Equation 1) at the highest level. Accordingly, at the top level, the model is divided into two decision elements; the Possibility of Attack (P_a) and the System Effectiveness (P_e) . The overall model is shown in Figure 5.

The Possibility of Attack (as opposed to the probability of attack due to lack of probabilistic data) side of the model is a measure of the overall possibility that a specific attack will take place. This side of the model is formulated as a balance between Assailant Perceived Consequence of the attack and the Assailant Efforts that go into planning and implementing the attack. The Assailant Efforts and the Assailant Perceived Consequence are balanced against each other to determine a value for P_a. The balance between Assailant Perceived Consequence and Assailant Efforts recognizes that the level of effort required by the attack must be compatible with the perceived payoff. For example, an assailant will have to put a significant amount of effort into attacking a water tank with a biological agent, these efforts may be deemed as acceptable to the assailant if there is a high-perceived consequence.

The System Effectiveness (P_e) side of the model is also structured around a balance. However, this balance is between the Assailant's Capability and the overall security of an asset, Asset Security. These two elements are balanced against each other to determine a value for P_e. That is, the capability of an assailant offsets the level of asset security. For example, an incapable assailant would not be able to overcome a highly secured asset.

From these two secondary levels of the model, Possibility of Attack and System Effectiveness, the rest of the model is developed as described in the following sections.

3.3.1. Possibility of Attack

As described above the Possibility of Attack is a balance between Assailant Perceived Consequence and Assailant Efforts. These are described in detail in the following sections. The aggregation for Possibility of Attack uses a soft aggregation scheme (Equation 2) of the decision elements Assailant Perceived Consequence and Assailant Efforts. Soft aggregation is used to yield a relatively slow response in the tails of the distribution and rapid changes in the central regions. Specifically, the case where the assailant is faced with a selectively high level of effort that results in limited consequence then the aggregate score will remain small. High aggregate scores occur when assailant efforts are relatively low and attack consequences are high. As a result, resolution in the aggregate scores is needed when the efforts and the consequences are more equally balanced, which should yield intermediate scores.

In addition to using a soft aggregation scheme to balance Assailant Perceived Consequence and Assailant Efforts, an "IF" statement is used to calculate Pa.

$$
WS_j=0.01 \text{ IF Agent Accessibility} \le 0.01 \text{ otherwise}
$$

$$
WS_j = \frac{1}{1+e^{\int_{t=1}^{u} w_j x_j - b_j}}
$$
(4)

The slope parameter (*a*) in this calculation is 5.5 and the shift parameter (*b*) is 0.5 in order to achieve the aforementioned reasons. The reason that an IF statement is used for the Possibility of Attack, is so that the assailant's access to a weapon is a significant driver of whether or not an attack can occur. For example, a vandal does not typically have access to a biological or chemical Agent; therefore without access to a weapon the overall possibility of that particular attack occurring is little to none.

3.3.1.1. Assailant Perceived Consequence

The Assailant Perceived Consequence portion of the model is the measure of the payoff or return that the assailant perceives will occur because of the attack. Five main consequences are identified based on the work of Fein and Vossekuil (1998) and the RAM- W^{TM} model (AwwaRF, 2002). The first four decision elements are: Attention Goals, Health Goals, Economic Goals, and Disruption Goals. These latent decision elements are weighted 0.3, 0.2, 0.2, and 0.3, respectively. Attention Goals and Disruption Goals are weighted slightly higher since an assailant may be somewhat more concerned about the disruption that an attack causes and attention that it receives rather than the heath or economic impacts. The fifth decision element, Media Attention, is not a direct input into Assailant Perceived Consequence but is a latent effect to Attention Goals since the amount of media attention an attack receives can correlate with the overall attention goals. The aggregate Assailant Perceived Consequence score is calculated as a linear weighted sum of each of the four latent decision elements (Equation 3).

The decision element Attention Goals is a measure of the Media Attention that occurs because of the attack balanced against the attention that the assailant desires. The direct inputs that feed into Attention Goals are Desired Attention and Media Attention. Media Attention a latent effect of Asset Importance, Health Impact, Economic Losses, and Social Disruption, with weights of 0.2, 0.3, 0.3, and 0.2. The aggregation for Media Attention is a linear weighted sum (Equation 3) of the direct inputs. The decision element Attention Goals is a balance of the actual Media Attention versus the Desired Attention of the assailant. Therefore, this decision element employs a form of the soft aggregation scheme (Equation 2) of the decision element Media Attention and the direct input Desired Attention. The aggregation scheme used is:

$$
WS_j = \frac{1}{1 + e^{(-a((x/y) - b)))}}
$$
(5)

where *x* is the Actual Attention received (Media Attention) and *y* is the Desired Attention. The slope parameter (*a*) equals 15.5 and the shift parameter (*b*) is set to 0.8. This aggregation scheme and parameters are used with the intention that if the assailants' desires are met (actual Media Attention greater than Desired Attention) the goal is successful therefore returning a score of 1 when the Media Attention score is close to or larger than the Desired Attention and zero otherwise.

The decision element Health Goals is a measure of the Health Impact that balanced against the Desired Health Impact that the assailant would like to occur. The direct inputs that feed into Health Goals are Desired Health Impact and Health Impact. The decision element Health Impact is made up of two direct inputs, Deaths and Illnesses with weights of 0.6 and 0.4. Number of deaths is weighted more heavily since a death is more significant than an illness. The aggregation for Health Impact is a linear weighted sum (Equation 3) of the direct inputs Deaths and Illnesses. The aggregation for the decision element Health Goals uses Equation 6.

$$
WS_j = 1.0 \text{ IF } y=0 \text{ and } x<0.05 \text{ otherwise};
$$

\n
$$
WS_j = 0.0 \text{ IF } x/y>1.2 \text{ otherwise};
$$

\n
$$
WS_j = \frac{1}{1 + e^{(-a((x/y) - b))}}
$$
 (6)

Variables include *x* as the latent effect Health Impact and *y* as the Desired Health Impact. In the soft aggregation equation used for this decision element, the slope parameter (*a*) is 30 and shift parameter (*b*) is 0.8, all other variables are as previously defined. The reason this equation includes multiple IF statements is so that if the actual impacts of an attack are greater than the desired impacts the assailant is penalized. For example consider a case where an Insider only wants to bring to light vulnerabilities in the utility by dumping household chemicals into a storage tank, however, many people become ill. This was not his intention; therefore, the Health Goals decision element is set to zero.

The decision element Economic Goals is a measure of the Economic Losses that occur because of the attack balanced against the Desired Economic Impact that the assailant would like to occur. Economic Losses can be considered a motivator of an assailant based on the consequences that the September 11, 2001 event had on the airline industry. The direct inputs that feed into Economic Goals are Desired Economic Impact and Economic Losses. The decision element Economic Losses is made up of three direct inputs, Repair Costs, Protection Costs and Economic Disruption with weights of 0.3, 0.1, and 0.3. The decision element Health Impact is also a latent effect of Economic Losses with a weight of 0.3. Health impact is a latent effect of Economic Losses because a health outbreak includes economic losses. The aggregation for Economic Losses is a linear weighted sum (Equation 3) of the direct inputs Repair Costs, Protection Costs and Economic Disruption and the latent effect Health Impact. The aggregation for Economic Goals employs Equation 6 where *x* is the latent effect Economic Losses and *y* is the direct input Desired Economic Impact for the same reasons as mentioned previously.

The decision element Disruption Goals is a measure of the Social Disruption that occurs because of the attack balanced against the Desired Level of Disruption that the assailant would like to occur. The direct inputs that feed into Disruption Goals are Desired Level of Disruption and Social Disruption. The decision element Social Disruption is made up of two direct inputs, Duration of Impact and Number of People Impacted each weighted at 0.5. Social Disruption is also a latent effect to Media Attention because the degree of disruption promotes attention. The aggregation for Social Disruption is a linear weighted sum (Equation 3) of the direct inputs Duration of Impact and Number of People Impacted. The aggregation for Disruption Goals employs Equation 6 where *x* is the latent effect Social Disruption and y is the direct input Desired Level of Disruption for the same reasons as mentioned previously.

3.3.1.2. Assailant Efforts

The Assailant Efforts portion of the model measures the effort an assailant will have to devote to the attack. Assailant Efforts is comprised of decision elements based on the work of Fein et. al. (1999) and the RAM-WTM model (AwwaRF, 2002). Three decision elements are structured as direct latent effects of Assailant Efforts. These decision elements are: Asset Locality, Chance of Assailant Being Caught, and Assailant Capability. The latent decision elements are weighted 0.2, 0.3, and 0.5 respectively. Assailant Capability is weighted as half of Assailant Efforts since the capability of an assailant greatly influences the amount of effort an assailant must put forth in order to carry out an attack. For example, an insider has high capabilities with respect to access to the utility, therefore reducing his overall efforts. The aggregate Assailant Efforts score is calculated as a linear weighted sum (Equation 3) and subtracted from 1. The weighted sum is subtracted from one to reverse the sense of the score to make it compatible with Assailant Perceived Consequence, with which it is aggregated to determine the Possibility of Attack. That is, the Possibility of Attack will increase as Assailant Perceived Consequences increase and/or Assailant Efforts decrease. For purposes of consistency in inputting values all direct inputs are defined in such a way that their scores increase.

The decision element Asset Locality is a measure of the overall ease that the assailant has in accessing the asset and the familiarity that the assailant has with the location of the asset. The direct inputs that feed Asset Locality are Acclimation Needed, Travel Restrictions, Language Barrier, and Time of Travel, with weights of 0.3, 0.3, 0.2, and 0.2. These inputs attempt to describe how an assailant might select a target (Borum et. al, 1999). The aggregation for Asset Locality is a linear weighted sum of the inputs (Equation 3).

The next decision element, Chance of Assailant Being Caught, is a measure of how likely an assailant can plan and complete an attack without being caught. The direct inputs that feed into Chance of Assailant Being Caught are Strength of Police Force, Federal Level of Involvement, Profiled Group and Political Will, with weights of 0.3, 0.3, 0.2, and 0.2. The equation used to determine the value for Chance of Assailant Being Caught is Equation 3, a linear weighted sum of the inputs.

Finally, the last latent effect to Assailant Efforts is Assailant Capability. This decision element recognizes all of the assailant's capabilities, with respect to the attack, as described within the System Effectiveness portion of the model.

3.3.2. System Effectiveness

As explained previously, the System Effectiveness is a balance between the Assailant's Capability and Asset Security. The aggregation for System Effectiveness uses a soft aggregation scheme (Equation 2) of the decision elements Assailant Capability and Asset Security. This aggregation assumes that the latent effects are equally weighted at 0.5, the slope parameter (a) equals 7.5 and the shift parameter is set to 0.5. This equation balances Assailant Capability against Asset Security. Where Asset Security is high and Assailant Capability is low the System Effectiveness is high, while if Asset Security is low and Assailant Capability is high System Effectiveness is low. Alternatively when Assailant Capability and Asset Security are relatively equally matched, the weighted soft aggregation scores change quickly.

3.3.2.1. Assailant Capability

The Assailant Capability portion of the model is made up of decision elements that relate to the overall capabilities of the assailant such as the assailants qualifications, motivations, amount of funding available, technical abilities, personnel and weapons available, and ability to access the asset. Eight main efforts have been identified based on the work of Fein et. al. (1998, 1999, 2000, 2002), Borum et. al. (1999), Pynchon (1999), Vossekuil (2001) and the RAM- W^{TM} model (AwwaRF, 2002). These efforts are defined by decision elements and all are latent effects of Assailant Capability. The decision elements are: Qualifications, Funding, Motivation, Access, Preparation, Technology Capability, Weapon Dissemination, and Weapon Accessibility. These latent decision elements are weighted 0.1, 0.1, 0.25, 0.25, 0.1, 0.1, 0.05, and 0.05, respectively. As noted by the weighting factors, Motivation and Access are more heavily weighted than the other decision elements indicating that these two elements make up half of the weighting for Assailant Capability. Thus, if an assailant has access to an asset the capability of that assailant increases significantly. Moreover, if the assailant is motivated they will have more desire to carryout the attack, therefore increasing their overall capability based on their willingness to accomplish the task at any cost. Assailant Capability is aggregated using Equation 3 and subtracted from one.

The first latent effect of Assailant Capability is Qualifications. This is a measure of whether or not the assailant has the qualifications needed to carry out an attack. The direct inputs that feed into Qualifications are Assailant Training and Special Expertise. These inputs are a balance of one another as Assailant Training is the training that the assailant has and Special Expertise is the training that the assailant needs in order to carry out the attack. The equation used to aggregate this decision element is Equation 5 using a slope parameter of 30 and shift parameter is 0.8. This methodology is used so that the function behaves like a step function with the intention that the assailant's abilities are balanced against the required abilities e.g. an assailant either has the ability to carry out an attack (in which case the score is 1) or he doesn't (score of 0).

The decision element Funding is a measure of the whether or not the assailant has the funds needed to carry out the attack. The direct inputs that feed into Funding are a balance of one another and are Cost of Attack and Available Funds. The aggregation for this decision element also uses Equation 5 to balance whether or not the assailant has the required funds.

The next decision element in Assailant Capability is Motivation. Motivation is a measure of the drive the assailant has to carry out the attack. The direct inputs that feed into Motivation are Peer Pressure, Hatred, Cause and Outlook, each with a weighting of 0.25. These four attributes are common motivators of both assassins and insiders according to Fein and Vossekuil (1999) and the Insider Threat Study for the Banking and Finance Sector studies (Randazzo, et. al, 2004). The assailant's current lifestyle can have a significant influence on their ability and likelihood of attack. Moreover, a group's influence over an individual can strongly impact their behavior and actions (Pynchon and Borum, 1999). The aggregation scheme for this decision element is Equation 3, a linear weighted sum of the inputs.

The next decision element that feeds into Assailant Capability is Access. Access is a measure of how accessible the system is to the assailant. This decision element has three direct inputs, System Knowledge, System Accessibility and Recognition with weights of 0.4, 0.3, and 0.3. These three elements are attributes of assailants as indicated in $RAM-W^{TM}$ (AwwaRF, 2002) and the Insider Threat Study for the Banking and Finance Sector (Randazzo et. al., 2004). The aggregation for this decision element is linear weighted sum of the inputs.

The next decision element that feeds into Assailant Capability is Preparation. This decision element is a balance between the amount of planning that goes into an attack versus the required planning to carry out that attack. The direct inputs into Preparation are: Degree of Planning, Network, and Required Planning. The input value Required Planning is a common component of attacks, noted by the Banking and Finance Sector Insider Threat Study, where the findings indicated that most incidents required little planning or networking, most of the attacks took place from the individuals' workstation during regular business hours (Randazzo et. al. 2004). The aggregation for this decision element balances Degree of Planning and Network versus Required Planning with weights for Degree of Planning and Network being 0.4 and 0.6 balanced against Required Planning using a form of Equation 5:

$$
WS_j = \frac{1}{1 + e^{-a((\sum_{i=1}^{n} w_i x_i / y) - b)}}\tag{7}
$$

where x_i is the amount of planning that goes into an attack, w_i is the corresponding weight, *n* is the number of attribute values contributing to the decision element Preparation and *y* is Required Planning for the attack. The slope parameter used in this case is 30 and the shift parameter is 0.8. These parameters are used so that when an assailant has a plan that meets or exceeds what is required, the decision element score yields a one, otherwise the score is zero.

The next decision element that feeds into Assailant Capability is Technology Capability. Technology Capability is a measure of how much technology is needed for an attack versus the amount of technology available for the attack. The direct inputs that feed into Technology Capability are Available Technology and Technology Needed and the latent effect Preparation. This decision element is a balance between Available Technology and Preparation versus the Technology Needed. The aggregation for this decision element balances Available Technology and Preparation versus the Technology Needed (Equation 7). The slope parameter used in this case is 30 and the shift parameter is 0.8 for the same reasons as mentioned previously.

The next decision element that feeds into Assailant Capability is Means Dissemination. Means Dissemination is a measure of how difficult it is for an assailant to deliver an agent or attack means versus his ability to disseminate it. The direct inputs into Means Dissemination are Difficulty of Dissemination versus Ability to Disseminate Means and are based on the Sandia Report 2003-0031 (Teter et. al. 2003). This decision element is a balance between these two inputs. The aggregation for this decision element uses a soft aggregation scheme (Equation 5) where the slope parameter used is 30 and the shift parameter is 0.8 for the same reasons as mentioned previously.

The final decision element that feeds into Assailant Capability is Weapon Means Accessibility. Means Accessibility is a measure of how easily the assailant can access the weapon in sufficient quantity to accomplish the attack. The direct inputs into Weapon Accessibility are Limitation to Weapon Availability and Weapon Accessibility. This decision element is a balance between these two inputs. The aggregation for this decision element uses Equation 5, with a slope parameter of 30 and the shift parameter of 0.8.

3.3.2.2. Asset Security

The decision elements that make up the Asset Security side of the model take into account the response capability of the water utility and community, the delay functions of the asset, the detection abilities of the asset and the culture of the water utility. All of the decision elements are aggregated to obtain an overall value for System Effectiveness, Pe.

The Asset Security portion of the model is composed of the three basic decision elements that make up a Physical Protection System (PPS): Detection, Delay, and Response, in addition to the culture of the water utility. A Physical Protection System is the overall security system of an asset. These three elements are the basis of many critical asset assessments and underpin the RAM-WTM process. Detection is intrusion sensing, alarm communication, and alarm assessment. Delay is made up of the barriers protecting an asset and Response is the communication to the response force, and the deployment of the response force with the ultimate goal being to stop/prevent an attack. (Garcia, M.L., 2001) The elements Delay and Response are direct latent effects of the output decision element, Asset Security and are equally weighted as 0.5. However, Detection is not a direct latent effect of Asset Security, but feeds into the decision elements contributing to Response with a weight of 0.5. This choice of structure is deliberate in that detection has no value unless there is an associated response.

Furthermore, the decision element Security Environment, which represents the security culture of the utility, is a latent effect to the Detect, Delay, and Response decision elements since the culture of the water utility can have an effect on all aspects of security. The aggregation for Asset Security uses Equation 2 and assumes that the direct latent effects feeding into the decision element are equally weighted at 0.5, the slope parameter (a) equals 5.5 and the shift parameter is set to 0.5. This equal weighting is intentional since the amount of time it takes to respond to an event, in order to try to mitigate it, should be faster than the time it takes an assailant to overcome the delay features of an asset.

3.3.2.2.1. Security Environment

The Security Environment decision element is a measure of the concern and awareness that utility employees have toward system security. The decision elements that make up Security Environment are Employees and Management. These are weighted 0.4 and 0.4, respectively. The decision element Employees is further broken down into Employee Culture and Employee Training each with weights of 0.5. The decision element Management is also broken down into these two categories in addition to the direct input Security Administration. These are weighted 0.5, 0.3, and 0.2. These inputs are based on Cooper's development of the MLE model (2001). An additional direct input to Security Environment is Intelligence which is weighted as 0.2. The aggregation for the decision element Security Environment is a linear weighted sum of the three

decision elements Employees, Intelligence, and Management using Equation 3. Furthermore, the attitude of the personnel influences the likelihood and speed at which they will detect and respond to abnormal operations which is why Security Environment is a latent effect of the Detection, Delay and Response elements.

3.3.2.2.2. Detection

The first required function of a security system is the detection of an assailant's actions. The Detection decision element relates to the different methods of detecting breaches and/or contamination of a water distribution system. Five decision elements form latent effects to the Detection element. The decision elements are: Remote Monitoring, System Monitoring, Physical Monitoring, Public Surveillance and the Security Environment. These latent effects are weighted 0.2, 0.1, 0.4, 0.1, and 0.2. These weights represent how likely utility personnel are to respond to the alarm based on the type of monitoring. All of these are part of the intrusion sensing, alarm communication, and alarm assessment categories in $RAM-W^{TM}(AwwaRF, 2002)$. The aggregation for the decision element Detection is Equation 3, a linear weighted sum of the decision elements.

The first direct input into Detection is Remote Monitoring. This element is made up of inputs that measure confidence in the detection system including Alarms and Video Surveillance with weights of 0.4 and 0.6. The aggregation for the decision element Remote Monitoring is Equation 3, a linear weighted sum of the decision elements. The Alarm decision element is further resolved into the direct input Alarm Surveillance of the detector, and the latent effect Alarm Reliability. Alarm Surveillance has a weighting factor of 0.4 and Alarm Reliability has a factor of 0.6. This weighting represents that reliably of the system is more important than the surveillance, since if the system is reliable, it is more likely that it will be monitored by utility personnel. The Alarm decision element is aggregated by Equation 3, a linear weighted sum.

The Alarm Reliability decision element is further broken into direct inputs of Quality, Maintenance, and Sensitivity with weights of 0.4, 0.3, and 0.3. The equation used to determine the value for Alarm Reliability is Equation 3, a linear weighted sum of the inputs.

The Video decision element is broken down into the direct input Video Surveillance, and the latent effect Video Reliability with weights of 0.4 and 0.6. This weighting considers the reliably of the video system more important than the surveillance. The Video Reliability decision element is further broken into direct inputs of Quality and Maintenance. Both Video Maintenance and Video Quality have weighting factors of 0.5 into Video Reliability. The equation used to determine the value for Video Reliability is Equation 3, a linear weighted sum of the inputs. The Alarms and Video decision elements are aggregated into the overall decision element Remote Monitoring by a linear weighted sum.

The second decision element that is a latent effect of Detection is System Monitoring. This decision element addresses the ability of inline sensors (e.g., pressure, flow, chlorine) to detect changes resulting from an attack. The elements that make up this decision element are Sensor Reliability and Sensor Location. These are weighted 0.6 and 0.4 using the same reasoning as before in that reliability is more important than surveillance. The equation used to determine the value for System Monitoring is Equation 3, a linear weighted sum of the inputs.

The Sensor Reliability decision element is further broken down by direct inputs of Sensor Quality, Sensor Maintenance and Sensor Sensitivity with weights of 0.4, 0.3, and 0.3. The equation used to determine the value for Sensor Reliability is Equation 3, a linear weighted sum of the inputs.

The third decision element that makes up Detection is Physical Monitoring. Physical Monitoring is a measure of the physical monitoring that occurs on site at a water utility that could aid in detecting an issue. Physical Monitoring has two direct inputs, Onsite Security and Onsite Personnel with weights of 0.6 and 0.4. The equation used to determine the value for Physical Monitoring is Equation 3, a linear weighted sum of the Onsite Security and Onsite Personnel inputs.

The next decision element to make up Detection is Public Surveillance. Public Surveillance is a measure of how likely it is that a citizen would report a suspicious activity to officials. Public Surveillance is made up of three direct inputs, Likelihood of Reporting, Location, and Visibility with weights of 0.4, 0.3, and 0.3. The equation used to determine the value for Public Surveillance is Equation 3, a linear weighted sum of the direct inputs.

In addition to the four decision elements that make up Detection, the decision element Security Environment is also a latent effect to Detection. Security Environment has a weighting factor of 0.2 into the Detection decision element. The reason that Security Environment is a latent effect to Detection is because if a water utility has an environment that does not take monitoring their system seriously, detection of events is unlikely.

3.3.2.2.3. Delay

The Delay decision element quantifies the balance between the amount of time for an attack to reach its full effect and the time before a response occurs. This decision element is a latent effect to Asset Security. Three decision elements are latent effects to Delay. They are Critical Attack Time, Time to Respond, and Physical Security. These have weights of 0.25, 0.50, and 0.25. Note that the weighting values for Critical Attack Time, Time to Respond and Physical Security are scaled to represent equivalent times. Thus, the Time to Respond weight is balanced against the Critical Attack Time plus the Physical Security time. Therefore, a high Delay value represents the case where a response is likely to occur well before the assailants could accomplish their attack, while an intermediate value represents the case where the response and attack time are nearly equal. The aggregation used for the decision element Delay uses Equation 2, where the slope parameter (a) equals 10.5 and the shift parameter is set to 0.5.

The decision element Critical Attack Time is a measure of the time it takes an assailant to carry out the attack. This decision element is made up of two direct inputs, Time to Implement Attack and Time for Attack to Impact. These have weights of 0.2 and 0.8. The Time for Attack to Impact is weighted more heavily since this time impact is generally much longer (i.e., 4 times) than the time to implement the attack. Critical Attack Time is aggregated using a liner sum of the direct inputs (Equation 3).
The next decision element that feeds into Delay is Time to Respond. This decision element is a measure of the time required to respond to an attack. Direct inputs to this decision element include Response Time, Detection Time, and Decision Time with weights of 0.1, 0.3 and 0.3. Furthermore, Security Environment is a latent effect to Time to Respond with a weighting factor of 0.3, since the culture of the water utility can effect the time it takes for personnel to respond to an event. In the model, these direct inputs are represented as 1-Response Time, 1- Detection Time, and 1-Decision Time. This is deliberate since in the elicitation guides an immediate response, detection, or decision yields a zero value and an infinite time yields a value of one (i.e. they increase with time). However, the structure of the model is such that short decision times improve security; that is, a short time to respond and long critical attack time leads to high delay scores, while long time to respond and short critical attack time lead to low delay times. The decision element Time to Respond is aggregated using a liner sum of the direct inputs (Equation 3).

The last decision element that feeds into Delay is Physical Security. This element represents the delay measures that an assailant would have to breach in order to carry out a successful attack. This decision element has four direct inputs. These represent the overall physical security of the site and are: Asset Barriers, Distance to Asset Access Limitations, and Difficulty to Deliver Weapon. These have weights of 0.2, 0.1, 0.4, and 0.2 corresponding to the relative time delay associated with each security element. In addition to these direct inputs, Detection is a latent effect to Delay with a weighting factor of 0.1. Detection is a latent effect to Delay because the degree of sophistication of the detection network represents a physical security delay to the assailant. Time to Respond is aggregated using a liner sum of the direct inputs (Equation 3).

3.3.2.2.4. Response

Unlike Detection, both the Delay and Response elements are latent effects of Asset Security. This structure reflects the relationship that the overall asset security of a system depends on whether or not a response will occur and that the response will occur before the attack reaches its full effect. For example, a high score for Asset Security represents the case where a mitigating response is likely to occur and will occur well before the attack takes effect.

The Response decision element measures both the type of response and the likelihood of a response. Response also involves actions taken by a security force (guards, police or other law enforcement) to prevent an assailants' success (AwwaRF, 2002). The decision element Response is composed of the two decision elements Response Strength and Likelihood of Response with weights of 0.4 and 0.6. Note that the input Likelihood of Response is weighted more heavily, 0.6, than the Response Strength, 0.4, given that the strength of the response does not matter if no response occurs. Response is aggregated using a liner sum of these direct inputs (Equation 3).

The decision element Response Strength is defined by the decision elements Team Qualifications, Response Force, and Response Weapons. These are weighted 0.4, 0.3, and 0.3 and are aggregated using a linear weighted sum. This weighting signifies that it is more important to have well qualified people who are highly capable rather than a large response force.

The decision element Team Qualifications is a measure of the training that both the assailant and the responder have. This decision element is made up of direct inputs Assailant Training and Responders Training. These inputs are a balance of each other and the decision element Team Qualifications is aggregated using Equation 5 where the slope parameter (*a*) equals 30 and the shift parameter (*b*) is set to 0.8. This methodology is used so that the function behaves like a step function with the intention that the assailant's training is balanced against the responder's training. For example, if the responder's training is better than the assailant's training a value of one is yielded therefore increasing the overall System Effectiveness value.

The next decision element that makes up Response is Response Force. Response Force is a measure of the number of assailants versus the number of responders for an event. The direct inputs into Response Force are Number of Assailants and Number of Responders. These inputs are also a balance of each other and the decision element Response Force is aggregated using Equation 5 where the slope parameter (*a*) equals 30 and the shift parameter (*b*) is set to 0.8. This methodology is also used in this case so that the function behaves like a step function with the intention that the number of assailants is balanced against the number of responders.

The final decision element that feeds into Response Strength is Response Weapons. This measures the weapons that the assailant has versus the weapons that the responders have. The direct inputs into Response Weapons are Assailant Weapons and Responders Weapons. These inputs are also a balance of each other where the decision element Response Weapons is aggregated using Equation 5 where the slope parameter (*a*) equals 30 and the shift parameter (*b*) is set to 0.8.

The last decision element that feeds into Response is Likelihood of Response. This decision element is a measure of how likely it is that a response will occur. Likelihood of Response is a function of the direct inputs Response Procedures, Dedicated Response Personnel and the latent effects Detection and Security Environment. These inputs have weights of 0.1, 0.2, 0.5, and 0.2. Note that Detection is half the weight of this decision element because without detection it is unlikely that any type of response will take place. Moreover, the decision element Security Environment is a latent effect because the environment that the utility personnel work in can have a significant influence over how likely it is that a response occurs. Likelihood of Response is aggregated using a liner sum of the direct inputs (Equation 3).

3.4. Software Environment

The MLE model presented here is implemented within the commercial spreadsheet package, Microsoft Excel. Interactive interfaces have been developed using Visual Basic. These interfaces provide the user with a set of menu driven instructions that act as guides through the analysis process. The user has the options to add, modify, or create from existing databases assailants, weapons, assets, and security environments. Specifically, to initially add any of these elements, the user queries each of the elicitation guides makes a selection and saves the inputs. These values are then uploaded into the model and a variety of scenarios can be developed. Once scenarios are determined, the accompanying calculations are performed and output into the overall MLE matrix. This allows the user to inspect how individual input values affect the

overall results. Results of each scenario can be saved, organized and ranked for inspection by the user. Figure 7 depicts the main menu of the interface and the buttons to add or modify assailants, weapons, assets and security environments.

Figure 6. Example of one of the MLE Model Interfaces

3.5. MLE Model Calibration

This MLE model was calibrated using two methods. The first method was collaboration with subject matter experts. Select individuals from the EPA, water utilities, and local water experts reviewed the Asset Security side of the model. These individuals reviewed the model structure, elicitation guides, and/or weighting factors. Their comments and suggestions were used when developing the final version of the Asset Security side of the model. This collaboration helped in validating the input values, weights, and elicitation guides. For the Possibility of Attack and Assailant Capability portions of the model, collaboration with the Security Department at Sandia National Laboratories was conducted to review the inputs and determine from a high level perspective if the attribute values were appropriate. However, specific elements of the Assailant Capability part of the model relied mainly on literature reviews of assailants and threat assessment studies.

The second calibration method was based on a comparison between actual occurred events and the MLE generated threat assessment value. For this calibration, data from the AwwaRF study (Welter, G.J., 2003) was used. This study documents an assembly of over 250 security incidents that have direct relevance to a water system that occurred, were planned, or were threatened. Only occurred and planned events were used in the calibration as credibility of the threatened (hoax) events was often suspect. Since the AwwaRF report is classified as confidential and proprietary, specific events cannot be identified. However, in order to create scenarios to calibrate the MLE model, multiple groupings of these attacks were developed based on assailant, weapon, and asset. Therefore, based on these three items, scenarios were developed and modeled using the MLE model.

From these scenarios, values for assets, assailants, and weapons were input into the MLE model. To the extent available, values are based on the information provided in the report and from general information available in the open literature. Asset security related inputs were generally assumed to follow normal conditions found within US water utilities. Model structure, weighting values and uncertain input values were then adjusted in a self-consistent fashion until a reasonable fit between the MLE model scores and the AwwaRF data was achieved. Representation of this data is shown in the following tables. The specifics are eliminated due to the classified nature of the data.

The following tables represent the number of occurred scenarios (as a percentage of the total) compared to the threat assessment value that the MLE model generated. The scenarios are grouped by assailant in these tables.

			AwwaRF
			Occurred $&$
AWWARF COMBINED	MLE Threat	AWWARF COMBINED	Threatened
RESULTS OF OCCURRED:	Assessment	RESULTS OF OCCURRED:	Events/Total
THREATENED: PLANNED	Score	THREATENED: PLANNED	Events
Terrorist; Tank; Contamination		Terrorist; Tank; Contamination	
(Chem)	0.05	(Chem)	0.01
Terrorist; Water Tank; Bomb	0.02	Terrorist; Water Tank; Bomb	0.00
Terrorist; Transmission Lines;		Terrorist; Transmission Lines;	
Bomb	0.02	Bomb	0.01

Table 2. AwwaRF vs. MLE Comparison Assailant=Terrorist

Table 3. AwwaRF vs. MLE Comparison Assailant=Insider

Table 4. AwwaRF vs. MLE Comparison Assailant=Vandal

As depicted in Table 2, where a terrorist is the assailant, overall MLE threat scores are very low and the actual events that have occurred are essentially zero. This demonstrates that the MLE threat assessment values are reasonable with respect to terrorism events. Tables 3 and 4 represent the Insider and Vandal scenarios. The disparity in the scenarios for these assailants is potentially due to lack of reporting occurred events due to the nature of the attacks (AwwaRF, 2002). However, the MLE threat assessment values depict that these are the most likely events to occur but because of the type of attack, e.g. vandalism, water utilities may not document the event due to its lack of severity.

Thus, by reviewing the data and the correlation between the scenarios and understanding the reasons for the differences in the Vandal and Insider events, this calibration, demonstrates that the equations, weightings, and inputs used in the MLE model are suitable for determining overall threat assessment scores. Additional event data is needed to calibrate the model further.

4. MLE WATER MODEL CASE STUDY

A case study was conducted to demonstrate utilization of the MLE model. Specifically, the model is used to show how it can be used in the threat assessment process to demonstrate the type of results that are obtained, how they can be interpreted in the context of a threat assessment, and how the MLE model can be used to identify mitigation strategies.

4.1. Water Utility Description

The MLE model application was performed for a Generic municipal water distribution system of a size that would commonly serve a community of approximately 500,000 customers. The data source for this analysis is the vulnerability assessment performed for the community in 2002 using the RAM-W methodology. The primary source of water for this city is groundwater.

4.2. Facility Prioritization

Following the RAM- W^{TM} process (AwwaRF, 2002) water utility assets were prioritized to determine which facilities are critical to operation. Although a utility may have a number of assets, there may be duplication of their functions, therefore potentially eliminating the need to evaluate all assets. Thus, using the pairwise methodology, outlined in the RAM- \tilde{W}^{TM} process, critical assets are defined based on water utility mission objectives. Typical objectives for a water utility are providing sufficient water to meet fire fighting flows, providing service to critical customers, and distributing potable water. Based on these objectives, twelve critical assets were identified. However, because some assets had similar functions, six assets were evaluated in the MLE model.

4.3. Utility Assets Characterization

The assets identified for analysis in the MLE modeling were: a water treatment plant, a water storage tank, buried pipeline, pump station, node, and a well. These assets were then characterized by collecting all of the existing security and general features of each asset. These items can include fences, locks, barriers, alarms, accessibility, sensors, asset location, etc. As part of the site characterization, the water utility must also collect information about its security policies, procedures, and overall security training at the site. This includes items such as evaluating emergency response plans, access control, and the security culture of the management and staff personnel of the facility, e.g. are alarms taken seriously, is security training conducted, etc.

The asset "water treatment plant" is a conventional filtration plant. The most commonly used chemicals at this plant are stored onsite in capacities over 100,000-gallons. The perimeter of this facility has a six-foot tall industrial chain link fence with a 3-strand barbed wire outrigger. All gates are sensored and cameras are located at entrance gates so visitors can request access to the facility. An uncovered reservoir is on site. The control building houses the control room for plant operations and all doors are equipped with sensors. One operator is on duty during off hours. This plant is located outside the city but is accessible by a maintained roadway system.

The asset "water storage tank" is a completely enclosed metal roof reservoir. The entrance hatch is sensored and secured with a pad lock. The perimeter of the facility is secured with a six-foot tall chain link fence with a three-strand barbed wire outrigger and gates are sensored. The control and pump rooms are in a block building onsite. All doors into this building are sensored and there is a communication link. This tank is located in a sparsely populated area. Water utility personnel are not stationed at this location. These characteristics are typical to all the water tanks onsite

The asset "buried pipeline" is considered the main pipeline from the water treatment facility to the distribution system. This pipeline is assumed buried to a three-foot minimum below grade. The line is detectable by conventional methods and does not have a contaminant monitoring system associated with it. The pipeline is located within the street right-of-way. Disruption (e.g. physical damage) to this water line would be noticeable in the control system. Neither utility nor security personnel monitor this length of pipe. The barriers associated with this asset are the inherent properties of the pipe, i.e. material type, type of joints, burial depth, etc. These characteristics are typical to all the pipes within the distribution system.

The asset "pump station" is the main booster pump station for the city. It is located some distance away and is accessible by a maintained road system. The site does not have personnel on site on a regular basis but utility employees occasionally conduct work at this site. The pumps are open to the environment. There is a six-foot high perimeter fence with three strand barbed wire outriggers and sensored gates, but no cameras for alarm assessment. The block building located at this facility houses the controllers and SCADA system. All doors into this building are sensored and there is a communication link. These characteristics are typical to all the pump stations within the distribution system.

The asset "node" is a point in the system where a contaminant can be injected. These are located throughout the distribution system and do not have security or utility personal located at them. These are open to the environment and do not have any security features associated with them. A typical node is a home connection to the distribution system.

The asset "well" is a typical well with an onsite pump with SCADA controls. The building located on site houses the SCADA and pump controls. There is a liquid chlorinator onsite for water treatment. The perimeter consists of a six-foot tall chain link fence with three strand barbed wire outriggers. Gates are sensored and padlocked. A communication link is also available.

4.4. Utility Culture Characterization

The utility culture characterization differentiates between a water utility that is highly conscious of security aspects of their system and those that have little regard for security measures. Our generic city has some policies and procedures in place related to security, however they are not threat-specific nor do they provide enough detail to be effective if an attack on the system occurred. Furthermore, system-wide security training is not conducted. Employees are granted access to all facilities and no background checks are performed. Separation policies in terms of returning keys, badges, changing passwords, etc., do not exist and there is no policy or procedure

to confront trespassers. Alarms are monitored and security policies state that a response should take place to investigate all alarms but operators do not always follow these requirements and typically dispatch a response only if multiple alarms are received. Policies do not exist for controlling keys, vehicle access, alarm logging, security incidents or unauthorized intrusions. This utility has an emergency response plan that is based on the Y2K event, however, it does not identify how to manage intentional attacks or provide training on how to respond. Therefore, the security culture of this city is considered low.

4.5. Threat Analysis and Characterization

A threat analysis identifies and describes the types of assailants that may disrupt a water utility from performing its missions. This information is used to analyze how an assailant can attack the utility's assets and can help develop a baseline for the utility to determine the level at which they should protect their assets. To determine the threat for specific attacks, several steps must occur.

According to Garcia (2001) the methodology for developing threat definitions consists of three basic parts:

- 1. List the information needed to define the threat.
- 2. Collect the threat information on the potential threat.
- 3. Organize the information to make it useful.

Listing the threat information includes identifying the type(s) of assailants, that is, defining their motivations, goals, tactics, weapons, numbers, and capabilities. These assailants may include insiders, vandals, terrorists, or any type of individual assailant or combination of assailants, as long as each is clearly defined.

Sources of threat information can include, but is not limited to, items such as criminal reports, intelligence reports, historical data, employee conflicts, and/or expressed threats that are related to the water utility. Defining the potential means that each assailant has available for their use is an important part of the threat information since the means can effect the overall outcome of an attack. After this information is collected, it must be organized in a form that is useful to define an assailant threat. The information gathered in this process is based on the RAM-W vulnerability assessment and is used to determine input values into the MLE model using the elicitation guides. For this case study, three assailants are defined: terrorist, insider, and vandal. The following paragraphs describe each of these assailants.

A terrorist is defined as an organized, highly motivated group of up to five outsiders equipped with sophisticated tools, explosives (larger than backpack quantities e.g. truck bombs), and chemical, biological and radiological agents. All equipment is person portable and easily obtainable. They have extensive knowledge of the security system and of the water operations. They can work as a single unit or in teams to achieve their goals. Their goal is to use their materials with the intent to cause massive deaths/illnesses and disrupt water distribution.

An insider is defined as a single motivated disgruntled employee or authorized contractor working unaccompanied with authorized access, possessing extensive knowledge of the water system, the security system and security procedures. The insider has access to hand and power tools and authorization to access the chemicals available at the water utility. The types of agents that are available to an insider are those on the open market. The insider's goal is to inhibit delivery of water by damaging or manipulating assets or to introduce substances (primarily onsite chemicals) into the water supply to disrupt utility operations but with limited physical harm to the general public.

A vandal is defined as one or two outsiders, with no authorized access or inside information, using portable hand tools with the intent to inflict physical damage to the water utility facility or theft of water utility property or equipment. The types of weapons that are available to a vandal are those on the open market. These outsiders do not intend to cause physical harm to water utility employees or to end-users.

The means that will be used in this case study for each of the assailants are: bombs, contamination, on site chemicals, biological contamination and chemical contamination.. The act of "vandalism" (e.g. damage to an asset with spray paint) will also be evaluated.

4.6. Threat Scenarios

Once the water utility has defined critical assets, assailants, attack means, and its security culture, it must decide what threat assessment scenarios to evaluate. A scenario consists of an assailant, a attack means, an asset and a utility environment. For example, a vandal with a can of spray paint intended for use on a water storage tank in a culture that does not have much concern for security. The scenarios used in this research are composed of the combination of the previously defined assets (water treatment plant, water storage tank, buried pipeline, pump station, node, and well), assailants (terrorist, insider, and vandal), and means (bomb, on site chemical contamination, biological contamination, chemical contamination, and vandalism) in a security environment where the concern for security is low. This research also evaluated these scenarios in a high security environment to note how the latent effects of the culture of the utility can affect the overall threat assessment value. Therefore, this study identifies 72 varying scenarios to evaluate and rank based on the overall threat assessment score. Note that each of the scenarios is modeled twice, first using a low security environment and then again using a high security environment, resulting in 144 combined scenarios. Scenarios are listed in Table 5.

Table 5. List of Scenarios

4.6.1. Input Values

To model a particular threat scenario, each of the 79 attribute values are quantified and input into the model. These input values are determined with the aid of elicitation guides (e.g., Figure 4). All elicitation guides used in this MLE model are included in the appendix of this document for reference. Attribute values is determined by comparing the vulnerability assessment documentation to the criteria given in each elicitation guide. The team of individuals that provided vulnerability assessment data consisted of a mixture of Sandia National Laboratories security, operational, and risk assessment methodology personnel, and personnel from water utilities that possessed a clear and detailed understanding of the history and operations of a utility facility. Attribute values used in the assessment of each of the threat scenarios are given in Tables 6 to 10.

With respect to each attribute associated with the assailant's capabilities and efforts, higher values represent that the assailant is more capable, has stronger desires, or his efforts are greater. Scores are also dependant on the type of weapon used by each assailant. In general, scores are typically high for a terrorist and low for an insider or a vandal. See Table 6 for assailant input values.

Table 6. Assailant Input Values

Attribute values for individual Assets vary according to the Detection and Delay features of each asset. For example, values tend to be higher at the treatment plant where there are multiple forms of monitoring (e.g., on site security and personnel, video/alarm systems, in-line sensors) and much lower for a node or a pipe. Moreover, these input values are dependent on the type of threat scenario being modeled. For example, since sensors cannot detect bombs or vandalism their values are set to zero, whereas for contamination events, which be detected, these values range from 0.2 to 0.8. The attribute value Detection is similar because a bomb is immediately detected whereas contaminants in the water can take some time to be detected in the system. See Table 7 for asset input values.

Table 7. Asset Input Values

Input values for the biological and chemical weapons are based on Teter's report (Teter, D.M., 2003). The input values for a bomb, on-site contamination, and vandalism are based on the elicitation guides. For each attribute, higher values represent more difficulty in obtaining the weapon, greater risk to the assailant using the weapon, and more difficultly in disseminating and delivering the weapon. In general, values are higher for a contamination event than for a bomb or vandalism. See Table 8 for weapon input values.

Table 8. Means Input Values

Attribute values for the Utility Environment elements for a Low Security Environment are constant because these elements deal with characteristics associated with the utility as a whole. However, the level of federal involvement would be higher for a terrorist than for an insider or vandal attack, so this value differs. In the High Security scenario, all input values are constant for all assailants. In general, values are low for the Low Security scenario because the utility has a relatively poor respect for security-related issues and is poorly prepared to respond to a threat event. See Table 9 for input values.

Table 9. Utility Environment Values

Attribute values used for the Final Scenario values are dependent on all aspects of the scenario, e.g. assailant, means, and asset. These values also represent the consequence of the attack, e.g. the cost to repair an asset if damaged by a bomb or the number of people affected by an attack. Values are high in cases, for example, where the cost of attack is high, or many people are affected. See Table 10 for input values for each scenario. Note that these values are constant in the high and low utility environments.

Table 10. Final Scenario Values

	Deaths	Illnesses	Repair Costs	Protection Costs	Economic Disruption	Duration of Impact	Number of People Impacted	Cost of Attack	Technology Needed	Required Planning	Ability to Disseminate Means	Means Accessibility	Time to Implement an Attack	Time for Attack to Impact	Decision Time	Sensor Sensitivity
Scenario Name	Input Values															
Terrorist/Treatment Plant/Bomb		$[0.0 0.0 0.7 0.3 0.5 0.3 0.7 0.2 0.6 0.8 1.0 0.7 0.4 0.1 0.3 0.0$														
Insider/Treatment Plant/Bomb		$[0.0 0.0 0.7 0.3 0.5 0.3 0.7 0.2 0.6 0.5 0.3 0.2 0.4 0.1 0.3 0.0]$														
Vandal/Treatment Plant/Bomb		$0.0 0.0 0.7 0.3 0.5 0.3 0.7 0.2 0.6 0.5 0.3 0.2 0.4 0.1 0.3 0.0$														
Terrorist/Treatment Plant/Contamination		$0.0 0.5 0.3 0.3 0.5 0.3 0.7 0.1 0.5 0.8 0.5 0.7 0.2 0.7 0.5 0.7 $														
Insider/Treatment Plant/ Contamination		$[0.0 0.5 0.3 0.3 0.5 0.3 0.7 0.1 0.5 0.3 0.7 0.7 0.2 0.7 0.5 0.7]$														
Vandal/Treatment Plant/ Contamination		$[0.0 0.5 0.3 0.3 0.5 0.3 0.7 0.1 0.5 0.8 0.3 0.1 0.2 0.7 0.5 0.7]$														
Terrorist/Treatment Plant/Contamination (Bio)		$1.0 0.7 0.5 0.3 0.7 0.6 0.7 0.4 0.5 0.8 0.1 0.5 0.2 0.7 0.5 0.7 $														
Insider/Treatment Plant/ Contamination(Bio)		$1.0 0.7 0.5 0.3 0.7 0.6 0.7 0.4 0.5 0.8 0.1 0.1 0.2 0.7 0.5 0.7 $														
Vandal/Treatment Plant/ Contamination (Bio)		$1.0 0.7 0.5 0.3 0.7 0.6 0.7 0.4 0.5 0.8 0.1 0.1 0.2 0.7 0.5 0.7 $														
Terrorist/Treatment Plant/Contamination (Chem)		$1.0 0.7 0.5 0.3 0.7 0.6 0.7 0.4 0.5 0.8 0.1 0.3 0.2 0.7 0.5 0.7 $														
Insider/Treatment Plant/ Contamination (Chem)		$1.0 0.7 0.5 0.3 0.7 0.6 0.7 0.4 0.5 0.8 0.1 0.1 0.2 0.7 0.5 0.7 $														
Vandal/Treatment Plant/ Contamination (Chem)																$1.0 0.7 0.5 0.3 0.7 0.6 0.7 0.4 0.5 0.8 0.1 0.1 0.2 0.7 0.5 0.7 $

4.8. Case Study Results

Utilizing all values within each scenario described in the previous tables, threat assessment values are calculated using the aggregation schemes described in the previous sections. To organize the results of this case study, the model is broken down by its two parts, Possibility of Attack and System Effectiveness. Breaking down the results into these categories allows the significance of these portions of the model to be recognized. Following the analysis of these two elements, overall threat assessment scores are discussed.

4.8.1. Possibility of Attack Decision Element Values

As discussed previously, the Possibility of Attack side of the model is a measure of the overall possibility that a specific attack will take place. Since this side of the model is a balance between Assailant Perceived Consequence and Assailant Efforts, it recognizes that the level of effort required by the attack must be compatible with the perceived payoff.

4.8.1.1. Assailant Perceived Consequence

The Assailant Perceived Consequence portion of the model is the measure of the payoff or return that the assailant perceives will occur because of the attack. This part of the model is based on the decision elements Attention Goals, Health Goals, Economic Goals, and Disruption Goals. The values listed in Table 11 reflect the aggregation of these decision elements for each scenario evaluated.

Table 11. Assailant Perceived Consequence Values

As shown in Table 11, there tends to be groupings of particular types of scenarios with respect to the Assailant Perceived Consequence values. The first grouping in the range of 0.69 and 0.50 are those scenarios conducted by a vandal using a bomb or vandalism. One of the reasons that all of the scenarios ranked in the top category is because vandals have very low desires. Referring to Table 6, Assailant Input Values, the vandal has the lowest desires for any particular attack scenario. Therefore, due to the aggregation applied for these decision elements, when a goal is met, e.g. the perceived attack impact meets the assailant's goals, as in vandalism events, the overall value for Assailant Perceived Consequence increases.

The next grouping of values includes the scenarios associated with primarily insider or vandal contamination events at multiple assets. The reason the values in this range, 0.45-0.22, are higher than other types of events is again because when goals are met based on the assailant's desires, Assailant Perceived Consequence increases. Also, this category includes many of the biological and chemical agent attacks. The reason these are grouped within this range due to the number of illnesses or deaths that occur if these agents are used. This component of Heath Effects then increases therefore increasing Assailant Perceived Consequence.

The next group of values includes those scenarios with an Assailant Perceived Consequence value of 0.20. These scenarios include the Insider scenarios associated with a bomb or vandalism as a weapon. The reason these are ranked as only a 0.2 in terms of consequence, is because an Insider has a greater ability to disrupt the water system by more sophisticated means than vandalism or a bomb, e.g. contamination.

The next group of values includes those scenarios with an Assailant Perceived Consequence value of 0.18-0.04. For the most part, these scenarios include Terrorist scenarios associated with a biological or chemical agents. The reason these are ranked in this category, is because a Terrorist typically wants to disrupt the water system in a broader range, e.g. via a treatment plant that serves a larger population.

The last group of values includes those scenarios with an Assailant Perceived Consequence value of 0.0. For the most part, these scenarios include Terrorist scenarios associated with vandalism, bombs or onsite contamination. The reason these are ranked so low (zero perceived consequence), is also because Terrorists typically want to disrupt the water system in a broader range, e.g. via biological or chemical weapons.

All of these Assailant Perceived Consequence values will be balanced against the Assailant Effort values to determine the Possibility of Attack Value.

4.8.1.2 Assailant Efforts

The Assailant Efforts portion of the model measures the amount of effort that an assailant will have to devote to the attack. Assailant Efforts is comprised of the three decision elements Asset Locality, Chance of Assailant Being Caught, Assailant Capability. The values in Table 12 reflect the aggregation of these decision elements. Note, the values shown are 1-Assailant Efforts, therefore indicating that the higher the value the *easier* the attack is for the assailant, e.g. efforts are low.

Table 12. Assailant Efforts Values

As was the case in Assailant Perceived Consequence, there are also distinct groupings within the Assailant Efforts portion of the model. However, the assailant and the weapon more specifically determine these groupings. The first grouping includes values 0.91-0.88, which consists of vandalism scenarios conducted by an insider and on-site chemical contamination events by an insider. These represent the scenarios that take the least amount of effort for this assailant. Therefore, an Insider has a low value for Asset Locality (don't have restrictions or a distance to travel, learn a language, become acclimated, etc.), has little chance of being caught, and a high capability to carry out that type of attack.

The next grouping of scenarios has the value of 0.76. This grouping includes all of the insider events where a biological or chemical contaminant is the associated weapon. The reason these are grouped at a somewhat low effort, is because if an insider had access to these types of weapons, his ability to distribute them into the water system would not take as much effort as a vandal or a terrorist since an insider has the access, qualifications, and other means to perform this task.

The next group in this category includes those scenarios associated predominantly with a vandal, an insider and a terrorist with bombs and on-site contamination in the range of 0.74-0.49. The range of attacks becomes more difficult for the individual assailants as the weapon becomes more difficult, e.g. a vandal attack with vandalism has lower efforts (0.74) than a vandal attack with a bomb (0.49). The terrorist scenarios also fall with in this range since he still has to manage items such as becoming familiar with the site, potentially learning a language, being part of a profiled group, and all the other decision elements within the Assailant Efforts portion of the model, even though the weapon is easy for a terrorist.

The last group of scenarios includes those associated with a vandal or a terrorist using a biological or chemical weapon with Assailant Effort scores of 0.47-0.46. The reason these are grouped at the highest effort, is because if a vandal or terrorist had access to these types of weapons, his ability to distribute them into the water system would be more difficult than that of an insider. However, access to the weapon is key in these types of events.

4.8.2 Overall Possibility of Attack (Pa) Values

The Possibility of Attack value is determined by balance of Assailant Perceived Consequence and Assailant Efforts. Therefore, Table 13 shows all of the scenarios and their ranks with respect to the possibility of how likely each particular scenario will occur.

Table 13. Possibility of Attack Values (Pa)

As shown in Table 13, there are clear groupings of scenarios. These groupings illustrate the overall possibility of attack with respect to the assailant and the means used. The first grouping includes the scenarios with a vandal and an insider as assailants. These values, 0.97-0.70, are expected for a Possibility of Attack value since the vandal and the insider have high perceived consequence values and require the least amount of efforts for these particular attacks. It is also demonstrated within this grouping the most likely means to be used by their relative order of ease. In all of the groupings, the means become more difficult as the P_a value decreases. For example, the Pa value is higher in the case where the assailant is using vandalism but is lower when he is using Chem/Bio contamination which may be more difficult to obtain.

The second grouping includes those scenarios that involve a terrorist as the assailant. These scenarios range in values of 0.15-0.02. Within these scenarios, contamination events with biological or chemical weapons are the most likely for a terrorist since they have the highest perceived consequence whereas a vandalism attack does not produce the consequence a terrorist desires.

The last grouping in this category includes the scenarios associated with a vandal and an insider as the assailant using a biological or chemical contamination. The score of 0.01 represents that these types of attacks are the least likely to occur because these assailants typically do not have access to these type of agent. The reason this value is not set to zero is to indicate that there can be a *very* minor possibility that these assailants could have access to this type of agent, therefore not completely eliminating these events from occurring.

4.8.3. System Effectiveness Decision Element Values

The System Effectiveness side of the model is a balance between the Assailants Capability and Asset Security. Since this side of the model is a balance between these two elements, it recognizes the overall balance between assailant capability and asset security. For example, a highly secured asset balanced against an unqualified assailant would result in a high System Effectiveness score, which would then decrease the overall Threat Assessment score.

4.8.3.1. Assailant Capability

The Assailant Capability portion of the model measures how capable the assailant is with respect to an attack. The decision elements that relate to the overall capabilities of the assailant are items such as the assailants' qualifications, motivations, amount of funding available, technical abilities, personnel and means available, and ability to access the asset. The values in Table 14 reflect the aggregation of these decision elements. Note, the values shown are 1-Assailant Capability, therefore indicating that the *larger* the value the less capable the assailant is to carry out an attack. Reasoning for the reverse of these values is discussed in prior sections.

Table 14. Assailant Capability Values

The groupings in this category are dependent on the Means used in the attack and the assailant. The first grouping includes values 0.88-0.78, which consists of the scenarios conducted by a vandal with biological and chemical agents, on-site contamination and bombs. Vandalism does not rank in this grouping because it is a much easier attack than an attack with a bomb or contamination. Within this grouping the high values indicate those attacks were the assailant is the least capable. This ranking is determined by the means used and the asset attacked, e.g. a water treatment plant has more security features than a tank and injecting a biological or chemical agent at this location is the most difficult attack for a vandal.

The next grouping of scenarios has values from 0.61-0.41. This grouping includes insider scenarios except vandalism and on-site contamination as weapons. Vandalism and on site contamination do not fall in this grouping for an insider since an insider's capability is greater in these areas.

The next grouping in this category includes those scenarios associated with a vandal using vandalism as a means at a score of 0.38. These rank as the easiest attack for a vandal since this type of attack is in line with their area of expertise.

The next group of scenarios includes the terrorist scenarios using biological and chemical agents, bombs, onsite contamination and vandalism. The values for these scenarios range from 0.30- 0.25, ranking biological and chemical attacks the most difficult and vandalism and bomb events as the easiest for a terrorist.

The final grouping includes the insider vandalism and onsite contamination scenarios. The values associated with this group range from 0.16-0.11. This represents the insider being the most highly qualified when using onsite contamination and vandalism as attack means. Some of the reasons that an insider is the most capable in these events is because he has access to the facility and chemicals, does not need any particular training or specialized weapons, already has training associated with injecting chemicals, does not need much funding, and can be very motivated.

4.8.3.2 Asset Security

Asset Security is a measure of how secure an asset is with respect to an attack. The decision elements that make up the Asset Security side of the model take into account the detection abilities of the asset (Detect), the delay functions of the asset (Delay) the response capabilities of the water utility and community (Response) and the culture of the water utility (Security Environment). Large aggregate scores reflect good security. The values in Table 15 reflect the ranked aggregation of these decision elements.

Table 15. Asset Security Values

The groupings for Asset Security are dependent on the asset, the assailant and the means used in the attack. The higher the Asset Security value the more secure the asset is to the associated attack. Therefore, in general the least protected assets are those that are in conjunction with a terrorist. This is due to the terrorist having more training than the response force, therefore reducing the Response decision element, and the low security environment. Moreover, assets that do not have detection instruments associated with them or are located in areas where response is slow, for example water tanks or wells, can contribute to lower Asset Security values. Conversely, the most protected asset is typically the water treatment plant. This is partially due to the Detection (video, sensors, on site security/personnel), Delay (physical security of the site), and Response attribute values associated with this asset. Furthermore, the contamination events have higher asset security values, again, due to the Detection, Delay and Response attributes of these scenarios. Overall, the vandalism scenarios rank low for Asset Security. Since the assets, wells, tanks, and pump stations, typically don't have any detection or response methods for this type of attack these scenarios rank second lowest, next to a terrorist in Asset Security. Lastly, the terrorist scenarios rank as the least secure with respect to asset security for the reasons mentioned because of the capabilities of the terrorist.

As part of Asset Security the utilities Security Environment is a latent effect to the three decision elements, Detect, Delay and Response. Although all of the scenarios evaluated up to this point have used a Low Security environment, changing the Security Environment to High can have an affect on not only the Asset Security values but also on the overall threat assessment scores. This will be discussed later in this analysis.

4.8.4 Overall System Effectiveness (Pe) Values

The System Effectiveness value is a balance between the Assailants Capability and Asset Security. Therefore, Table 16 shows a ranking of each scenario with respect to their overall System Effectiveness.

Table 16. System Effectiveness Values (Pe)

The groupings for System Effectiveness are dependent on the asset, the assailant, and the weapon used in the attack. The higher the System Effectiveness value better protected the asset is to that attack. Therefore, the first dominant grouping includes the scenarios with the vandal as the assailant and the means being contamination and a bomb because a vandal is not very capable in these types of scenarios nor does he have access or skills using a bomb or a contaminant. These scenarios have values for System Effectiveness ranging from 0.82 to 0.64.

The next grouping in this category includes mainly those events associated with an insider with contamination and bomb as the weapon. The values associated with these scenarios are 0.59- 0.20.

The subsequent grouping in this category includes the scenarios where a vandal is the assailant and vandalism is the means. This ranking, 0.20-0.17, represents that the asset is not very well protected to these types of events

The next category of events includes most of the terrorist scenarios ranking from 0.16 to 0.10. This is a low system effectiveness value indicating that most assets in a water system are not well protected to a terrorist.

The last categories of scenarios include those with an insider using onsite chemicals for contamination and vandalism. This signifies that water utility assets are most susceptible to insider contamination or vandalism events than any other type of attack.

4.8.5 Threat Assessment

To reiterate the definition of threat assessment, the MLE model is a top down approach based on the equation, $P_a(1-P_e)$. At the top level, the MLE model is divided into two sections; the Possibility of Attack (P_a) and System Effectiveness (P_e) . Therefore the previously calculated values in Tables 13 and 16 are those used in the threat assessment equation. The following table (Table 17) ranks the overall Threat Assessment for each scenario.

Table 17. Threat Assessment Values Pa(1-Pe)

The overall Threat Assessment values range from 0.80 to 0.00 and are predominately segregated by the assailant. The higher the Threat Assessment value the greater the threat of that scenario occurring.

The first dominant grouping includes the scenarios where the vandal is the assailant using vandalism and the insider is the assailant using onsite chemicals and vandalism. The values in his category range from 0.82 to 0.56. Since these scenarios are associated with a vandal and an insider utilizing their strongest skill sets (high Assailant Capability), on assets that are easily accessible to them (Asset Security), this demonstrates that a water utility is most vulnerable to these types of attacks. Specifically water utilities should be concerned with vandalism, as these are the most likely events to occur.

The second grouping contains primarily the scenarios where an insider is the assailant and the means are contamination and a bomb. The values associated with these scenarios are 0.37-0.35. These values represent that these scenarios have a low Threat Assessment value but also depict that they are more likely to occur than a terrorist attack or a vandal with a bomb.

The third grouping of scores includes those scenarios associated primarily with a terrorist as the assailant using biological or chemical agents, onsite contamination, bombs and vandalism. These values range from 0.13-0.02. These scenarios can be further resolved into subcategories noting the specific means used in each scenario, e.g. a terrorist is most likely to contaminate a water treatment plant with a biological or chemical agent before he will vandalize an asset or use onsite chemicals for contamination.

This last group of scenarios includes those associated with an insider using biological or chemical agents and a vandal using any type of contamination agents. The reason this is the lowest threat is because these assailants typically do not have access to these means, therefore the likelihood of a threat associated with these means is very low. Although these values are very low with respect to the other scenarios, these types of attacks can still occur and water utilities should still consider them to determine appropriate protection measures for their assets.

4.8.6. Discussion

In the previous sections, a MLE model for water utility threat assessment is developed and applied to a generic municipal water utility for specific scenarios. For the most part, as depicted in Tables 11 to 17, the overall results of the modeled scenarios reflect the most logical outcomes for each of the scenarios. For example in Table 14, the four most likely threats to a water utility is a vandalism of an asset by a vandal. This is consistent with current problems that many water utilities encounter. Furthermore, the least likely threats are those attacks associated with a terrorist and then those where biological and chemical agents are used by an insider or vandal.

Comparing the case study results using the MLE model to the RAM- W^{TM} vulnerability assessment is difficult since within the vulnerability assessment the Possibility of Attack value is always set at one. Therefore, the assessments are driven solely by system effectiveness and consequence. However comparisons to the system effectiveness values for some scenarios can be drawn. Table 18 lists the scenarios that can be compared based on the system effectiveness scores in the MLE model and the RAM-W model. Not all scenarios can be compared since the RAM-W vulnerability assessment does not consider many of the events that the MLE model evaluates. For example, the RAM-W model does not consider a terrorist event since the RAM-W study determined that no facility or asset is capable of operating under this level of threat. Also Vandalism events in the RAM-W model were considered to damage assets with hand tools, and not just spray paint or vandalize the asset. Furthermore, a contamination event with onsite chemicals is considered to be "ineffective" because the quantity of chemicals stored at this utility is not large enough to significantly contaminate the system. Bomb events were also not used in the scenarios including a low level vandal or by an insider.

However, the following table of scenarios does indicate moderate changes to the overall threat assessment when the Possibility of Attack is calculated rather than assumed to be one. As depicted in Table 18, the overall threat assessment score decreases when Possibility of Attack (P_a) is not set to one but is calculated by the MLE methodology. This is an important feature that can help water utilities determine what assets to protect and where the greatest threats may occur. If a larger comparison of varying scenarios could have been determined this reduction in Pa would provide a better sense of the significance of calculating P_a .

	MLE	MLE	**RAM-W	RAM-W	MLE
Scenario	(Pa)	(Pe)	(Pe)	$1*(1-Pe)$	$Pa*(1-Pe)$
Low Security Insider					
Pump Station Vandalism	0.6918	0.0716	0.1000	0.9000	0.6423
Low Security Vandal					
Pipeline Contamination	0.0100	0.6775	0.9000	0.1000	0.0032
Low Security Vandal					
Treatment Plant					
Contamination	0.0100	0.7622	0.9000	0.1000	0.0024
Low Security Vandal					
Water Storage Tank					
Contamination	0.0100	0.6538	0.9000	0.1000	0.0035
Low Security Vandal					
Well Field					
Contamination	0.0100	0.6482	0.9000	0.1000	0.0035

Table 18. MLE vs. RAM-W System Effectiveness Scores

**RAM-W only uses a relative ranking for System Effectiveness where 0.1=Low, 0.5=Medium and 0.9=High

4.8.7 Sensitivity and Importance Analysis

The sensitivity and importance $(S \& I)$ analysis feature of this MLE model can help determine how sensitive and important specific attributes are to attacks. The Importance metric allows a user to identify those features that most significantly contribute to the success of a threat. The Importance metric is simply calculated by deriving the difference between the output value with the input based on actual utility conditions, and the output value when one of the input values has been set to zero. By repeating this process for each input value, the input with the greatest influence on the output is determined. Similarly, the Sensitivity is calculated as the difference between the output value with the input based on actual utility conditions, and the output value when one of the input values has been set to one. The Sensitivity metric measures the potential for improvement in that input to result in a measurable improvement in the total result. Using the sensitivity and importance analysis feature embedded in the model, the user can determine the most effective methods for decreasing the likelihood of attack for particular scenarios. The user can model various utility environments to establish methods for improving the security culture of the utility to determine areas that can be improved to achieve the most return on investment.

Utilizing these tools, the user can determine the best places within the system to devote time, funds, and efforts so that investments are not made in areas where little or no return on investment is gained from the improvements. Thus, from an economic standpoint, these tools can help water utilities invest wisely in their assets, personnel and training.

Using these calculation methods, a Sensitivity and Importance assessment was performed on the 72 Low Security scenarios. In the Appendix section of this report, all of the sensitivity analysis data is presented. Each table is grouped by assailant and weapon and then by asset. The first number in each column is the threat assessment score for that particular scenario with all the inputs as originally entered. The attributes listed, e.g. Duration of Impact, Deaths, etc., are the top three attributes in each category (sensitivity or importance) that had the most effect on the threat assessment score when that particular attribute was calculated as a zero or a one. Beneath each of the decision elements is the threat assessment score when that decision element was set to zero or one. This value can be compared to the original threat assessment score to determine how much of an effect that particular attribute had on the overall threat assessment value.

Based on this analysis it is evident that within each category of assailant and weapons there are similarities within each attack regardless of the asset. In the scenarios where a terrorist is the assailant and bomb is the weapon, the most sensitive elements are the decision elements Duration of Impact, System Accessibility, and Recognition. The areas of most importance are: Desired Attention, Federal Level of Involvement and Profiled Group. In the scenarios where an insider is the assailant and bomb is a weapon, the most sensitive elements are the decision elements Assailant Training, Network, and Detection Time. The areas of most importance are: Required Planning, Special Expertise and Technology Needed. In the scenarios where a vandal is the assailant and bomb is a weapon, the most sensitive elements are the decision elements Network, Degree of Planning, and Assailant Training. The areas of most importance are: Required Planning, Special Expertise and Technology Needed.

In the scenarios where a terrorist is the assailant and onsite contamination is the weapon, the most sensitive elements are the decision elements Duration of Impact, Deaths, and System Accessibility. The areas of most importance are: Desired Attention, Federal Level of Involvement and Profiled Group. In the scenarios where an insider is the assailant and an onsite chemical is the weapon, the most sensitive elements are the decision elements Deaths, Repair Costs, and Economic Disruption. The areas of most importance are: Desired Attention and Illnesses. In the scenarios where a vandal is the assailant and an onsite chemical is the weapon, the most sensitive elements are the decision elements Means Accessibility, Assailant Training, and Decision Time. The areas of most importance are: Limitation to means Availability, Required Planning and Time for Attack to Impact.

In the scenarios where a terrorist is the assailant and contamination with a biological or Chemical agent is the means, the most sensitive elements are the decision elements Duration of Impact, Number of People Impacted, and Repair Costs. The areas of most importance are: Desired Attention, Federal Level of Involvement and Profiled Group. In the scenarios where an insider is the assailant and contamination with a biological or chemical agent is the means, the most sensitive elements are the decision elements Means Accessibility, Decision Time, and Network. The areas of most importance are: Limitation to Means Availability, and Time for Attack. In the scenarios where a vandal is the assailant and contamination with Anthrax or Cyanide is the weapon, the most sensitive elements are the decision elements Means Accessibility, Assailant Training, and Available Funds. The areas of most importance are: Limitation to Means Availability, Required Planning and Time for Attack to Impact.

In the scenarios where a terrorist is the assailant and vandalism is the weapon, the most sensitive elements are the decision elements System Accessibility, Recognition, and System Knowledge. The areas of most importance are: Desired Attention, Federal Level of Involvement and Profiled Group. In the scenarios where an insider is the assailant and vandalism is the weapon, the most sensitive elements are the decision elements Duration of Impact, Number of People Impacted, and Peer Pressure. The areas of most importance are: Desired Attention, Federal Level of Involvement, and Political Will. In the scenarios where a vandal is the assailant and vandalism is the weapon, the most sensitive elements are the decision elements System Knowledge, Recognition, and System Accessibility. The areas of most importance are: Responders Training, Number of Responders and Responders Weapons.

4.8.8 Mitigation Strategy

Based on the sensitivity and importance analysis, and the values that are most dominant in each category, strategies to reduce the overall threat assessment are considered. However many of these sensitive and important decision elements are assailant dependent, e.g. the desires of the assailant, the Network they are associated with, the Training or Funds they have in addition to other assailant specific decision elements. Unfortunately, these decision elements cannot be controlled since they are related to the characteristics of individual assailant. However, some important and sensitive measures were determined that can be controlled. For example, System Accessibility, Detection & Decision Time, Duration of Impact, and Number of People Impacted. As Accessibility becomes easier (e.g. set to a value of 1) the overall threat assessment score increases. This is also the case for Detection & Decision Time as these values increase (e.g. a long detection or decision time) the overall threat assessment score increases. Lastly, as the Duration of Impact and Number of People Impacted increase, the overall threat assessment also increases. Therefore to reduce the threat assessment these attributes associated with the asset can be improved. Some improvements can include: hardening the system to reduce the ability for assailant to gain access to the asset, accelerate the detection and decision time of responders, and build in redundancies into the system so that if one asset is attacked, e.g. a tank, it can be bypassed so that the duration of impact and number of people impacted is minimal.

Finally, an overarching strategy to reduce the threat of attack to an asset is to increase the Security Environment of the water utility. This includes the following items:

- Developing relationships with the local police force and federal government in the area so that appropriate measures are taken in the event of an attack,
- Developing response procedures that establish protocols on how to manage attacks
- Train all personnel including first responders, management and employees
- Identify appropriate weapons for response personnel in the event of an attack
- Assign dedicated response personnel to critical assets
- Change the attitude of the management and employees of the utility to consider security a priority
- Obtain and utilize intelligence regarding threats appropriately.

According to a recent article in Opflow (Murphy and Kirmeyer, 2005) improving these types of activities can begin to improve system security within the existing budgets, policies, and procedures which can then lay the foundation for implementing other more sophisticated security systems. Therefore, an example of how these can improve the threat assessment score, Table 19 compares a low security environment vs. a high security environment.

	High Utility Environment: Threat	Low Utility Environment: Threat
Scenario Name	Assessment	Assessment
Low Security Vandal Water Storage Tank Vandalism	0.62	0.80
Low Security Vandal Pump Station Vandalism	0.61	0.80
Low Security Vandal Well Field Vandalism	0.61	0.80
Low Security Vandal Treatment Plant Vandalism	0.56	0.78
Low Security Insider Water Storage Tank Vandalism	0.34	0.65
Low Security Insider Pump Station Vandalism	0.34	0.65
Low Security Insider Well Field Vandalism	0.34	0.65
Low Security Insider Treatment Plant Contamination	0.33	0.73
Low Security Insider Treatment Plant Vandalism	0.33	0.73
Low Security Insider Well Field Contamination	0.19	0.56
Low Security Insider Water Storage Tank Contamination	0.19	0.56
Low Security Insider Treatment Plant Bomb	0.10	0.37
Low Security Insider Pump Station Bomb	0.07	0.34
Low Security Insider Home Contamination	0.06	0.26
Low Security Insider Well Field Bomb	0.06	0.31
Low Security Insider Pipeline Contamination	0.06	0.25
Low Security Insider Water Storage Tank Bomb	0.04	0.26
Low Security Terrorist Treatment Plant Contamination Bio	0.03	0.13
Low Security Terrorist Treatment Plant Contamination Chem	0.03	0.13
Low Security Vandal Pump Station Bomb	0.02	0.13
Low Security Vandal Water Storage Tank Bomb	0.02	0.12
Low Security Vandal Well Field Bomb	0.01	0.12
Low Security Insider Transmission Pipeline Bomb	0.01	0.10
Low Security Vandal Treatment Plant Bomb	0.01	0.10
Low Security Terrorist Well Field Contamination Bio	0.01	0.05
Low Security Terrorist Well Field Contamination Chem	0.01	0.05
Low Security Terrorist Water Storage Tank Contamination Bio	0.01	0.05
Low Security Terrorist Water Storage Tank Contamination Chem	0.01	0.05

Table 19. Comparison of Threats for a High Security vs. Low Security Utility

As indicated in the comparison between a High Security Environment and a Low Security Environment, the threat assessment scores for many of the scenarios decrease considerably. The ranking order of the attacks remains fairly constant, e.g. a vandal vandalizing a water tank is still the most likely attack to occur, however, the change from a 0.80 score to a 0.62 represents the attack is less likely to take place at a utility where the culture of the personnel at a utility considers security a priority. Some of the most significant changes in threat assessment are those scenarios where the insider is the assailant. This comparison especially demonstrates how the latent effects of the Security Environment contribute to the overall outcome of the threat assessment scenarios.

5. CONCLUSION

Water utilities are faced with the daunting task of assessing the vulnerability of their utilities to disruption by natural and malevolent acts. Integral to such effort is threat assessment, which is complicated by the lack of data and experience to conduct a full probabilistic assessment. Markov Latent Effects modeling is a methodology that avoids these pitfalls by providing a possibilistic framework to threat assessment.

There are a number of advantages to MLE modeling. MLE modeling provides a framework for utilizing both qualitative and quantitative data to score threat scenarios. Further, these calculations are performed within a spreadsheet environment, supported by an interactive, menudriven user interface, which makes the model operationally manageable for utility staff and management. MLE modeling also provides a consistent, quantitative basis for evaluating threat scenarios. This allows the head-to-head comparison of results and ultimate ranking of threats. Finally, and most importantly, MLE modeling provides a structured framework for integrating information from a wide range of disparate sources. This is important because security cannot be assessed from a handful of disjointed methods, rather, security must be assessed from a systems perspective. This is demonstrated by the MLE model (Figure 6). Specifically, threat assessment does not only depend on the security features of the asset alone, but on the assailant, the weapon, and the security environment. Furthermore, by decomposing the entire water system and developing a methodology to determine a value for Pa (possibility of attack), it allows one to fully evaluate all of the elements that make up a threat assessment. In this way, MLE modeling captures and structures these disparate information sources and aggregates the data to produce a truly integrated threat score.

Benefits of using this particular MLE modeling scheme to determine threat assessments for water utilities is that it can be used for any type of asset that a water utility owns (i.e. elicitation guides and attribute values are not asset specific). One of the main benefits of using this MLE modeling methodology as opposed to other methods is not having to determine actual probabilities. This is key since lack of data makes calculating probabilities difficult, especially when trying to determine threat assessments for willful attacks to an asset. MLE modeling also allows the user to calculate the Possibility of Attack (P_a) value rather than always assume a worst-case situation. This also provides a more realistic approach to threat assessment and presents an overall threat assessment value that is based on both the System Effectiveness and the Possibility of Attack. Furthermore, when this threat assessment value is input into the overall risk equation (Equation 1), Risk now becomes a function of all three inputs, Possibility of Attack, System Effectiveness, and Consequence, therefore risk is no longer solely driven by the consequence of the attack.

MLE modeling is effective as demonstrated by using an existing vulnerability assessment that used the RAM-W methodology and comparing those results to the MLE model results. However, limitations exist regarding the variety of comparisons since the RAM-W and MLE model did not exactly compare the same scenarios of attack and weapons used. Future work in this area is required. The MLE model also depicted a correlation between actual events that have occurred based on the AwwaRF study (Welter, 2003) as shown in the calibration graphs discussed previously.

MLE modeling is not limited to water distribution systems, but is applicable to any type of threat assessment. This methodology is a valuable tool for being able to determine the possibility of attack since most current methods for calculating P_a use a worst-case scenario setting P_a equal to one. Therefore, when using the risk equation (Equation 1), the overall risk is consequence driven. Since using this methodology is very conservative, there can be consequences to this in terms of increased costs for protection methods that may not be necessary if the threat of it actually occurring is very low.

Future work for continual development of the MLE model for water systems includes adding additional attributes to the model. Specifically an important addition to this model would be the addition of SCADA operations. Moreover, further resolving the Security Environment attributes may be beneficial in order to determine a more all-encompassing assessment of the Security Environment at a water utility. For example, additional input values regarding how water utility funds are spent, e.g. on security upgrades or other site necessities. Other inputs can be items like whether or not background checks are conducted on utility personnel, and more specific inputs to the overall utility culture of the site.

There are still some limitations regarding the MLE model, in particular within the Assailant Efforts side of the model regarding how much effort an assailant has to devote to an attack based on the weapon used. This portion of the model is ranking events such as an insider contaminating a tank with a biological or chemical agent as 0.76, not very difficult. However, the weapon that is used in this scenario *is* difficult to access; therefore, the Assailant Efforts should be higher in cases where the weapon is not easy to access. Possible solutions to this problem are influencing the Assailant Efforts equation based on the means, similar to the way the Possibility of Attack is aggregated.

Additional future work to the MLE model should include other types of calibration methods. One method in particular that would be useful is collaborating with water utility professionals to work through the list of 72 scenarios to rank them in their professional judgment on what the highest to lowest threats are, in addition to ranking of the possibilities of these attacks. This type of calibration would then allow a comparison between the outputs of the MLE model vs. water utility professional's opinions to determine if the MLE is ranking threats in a reasonable order. Supplementary to this collaboration, the MLE methodology could be piloted with water utility personnel to determine, from a water utility operator's viewpoint, the overall ease, and use of the MLE model.

Finally, other work that can be conducted is to develop standard input values for the assailant and weapon portions of model so that users are only concerned with the asset. Most of this has been completed, however, it may need some refinement to include additional assailants with varying weapons, capabilities and goals. Weapon values for most biological and chemical contaminants have already been determined from the Teter report (2003). Lastly, scenarios that include teaming of assailants should be considered, e.g. an insider teamed with a terrorist. These types of teams can have significant impacts on the success of an attack.

6. REFERENCES

- 1. Asbeck, E., and Haimes, Y.Y. (1984). The partitioned multiobjective risk method, Large Scale Sys., vol. 6, no. 1, pp. 13-38.
- 2. AWWA Research Foundation and Sandia National Laboratories. (2002). Risk Assessment Methodology for Water Utilities (RAM-WTM). Awwa Research Foundation, Denver, Colorado.
- 3. Borum, R., Fein, R., Vossekuil, B., and Berglund, J. (1999). Threat Assessment: Defining an Approach for Evaluating Risk of Targeted Violence., Behavioral Sciences and the Law, no. 17, pp. 323-337. John Wiley & Sons, New York.
- 4. Boulos, P. F., (2002). Integrating GeoBase with Advanced Geospatial Networks., United States Air Force MWH Soft, Inc.
- 5. Cooper, A.J., (1999). Markov Modeling for Aviation Safety Analysis., Sandia Report, SAND99-1661C
- 6. Cooper, A.J., (2001). The Markov Latent Effects Approach to Safety Assessment and Decision Making., Sandia Report, SAND2001-2229
- 7. Cooper, A.J. (2004). Soft Markov chain relations for modeling organizational behavior., Risk Decision and Policy, no. 9, pp. 1-12.
- 8. Copeland, C., Cody, B., (updated January 5, 2005). Terrorism and Security Issues Facing the Water Infrastructure Sector. Congressional Research Service., Order Code RL32189.
- 9. Critical Infrastructure Assurance Office. (2001). Report of the President of the United States on the Status of Federal Critical Infrastructure Protection Activities. http://www.iwar.org.uk/cip/resources/ciao/final-ciao.pdf
- 10. Danneels, J.J., and Finley, R.E. (2004). Assessing the vulnerabilities of U.S. drinking water systems., J. of Contemporary Water Research and Education, no. 129, pp. 8-12.
- 11. Eitzen, E.M., Takafuji, E.T. (1997). Historical Overview of Biological Warfare, Textbook of Military Medicine, Medical Aspects of Chemical and Biological Warfare, Office of the Surgeon General, Department of the Army, USA. 415-424.
- 12. Ezell, B.C., Farr, J.V., and Wiese, I. (2000a). Infrastructure risk analysis model, J. Infrastruct. Syst., vol. 6, no 3, pp. 114-117.
- 13. Ezell, B.C., Farr, J.V., and Wiese, I. (2000b). Infrastructure risk analysis of municipal water distribution systems, J. Infrastruct. Syst, vol. 6. no. 3, pp. 118-122.
- 14. Fein R.A., Vossekuil B., (1998). Protective Intelligence Threat Assessment Investigations: A Guide for State and Local Law Enforcement Officials., NIJ/OJP/DOJ Publication., no. 170612. Washington, DC: U.S. Department of Justice
- 15. Fein R.A., Vossekuil B., (1999). Assassination in the United States: An operational study of recent assassins, attackers, and near lethal approachers. Journal of Forensic Sciences*,* no. 50, pp. 321-333*.*
- 16. Fein R.A., Vossekuil B., (2000). Protective Intelligence and Threat Assessment Investigations: A Guide for State and Local Law Enforcement Officials. United States Department of Justice.
- 17. Fein R.A., Vossekuil B., Pollack W.S., Borum, R., Modzeleski, W., Reddy, M., (2002). Threat Assessment in Schools: A Guide to Managing Threatening Situations and to Creating Safe School Climates. United States Secret Service and United States Department of Education.
- 18. Feller, W. (1957). Introduction to Probability Theory and its Applications, John Wiley & Sons, New York.
- 19. Garcia, M.L. (2001). The Design and Evaluation of Physical Protection Systems. Butterworth-Heinmann Publishers, Boston.
- 20. Gleick, P.H., (1993). Water and conflict: Fresh water resources and international security. International Security 18 vol. 1, pp. 79-112.
- 21. Grigg, N.S. (2003). Water utility security: Multiple hazards and multiple barriers., J. Infrastruct. Syst., vol. 9, no. 2, pp. 81-88.
- 22. Haestad Methods (2005). WaterSAFE® Model. http://www.haestad.com/software/watersafe/default.asp
- 23. Haimes, Y.Y. (1981). Hierarchical holographic modeling, IEEE Trans. On Sys., Man and Cybernetics, vol. SMC-11, no. 9, pp. 606-617.
- 24. Haimes, Y.Y., Matalas, N.C., Lambert, J.H., Jackson, B.A, and Fellows, J.F.R. (1998). Reducing vulnerability of water supply systems to attack. J. Infrastruct. Syst., vol. 4, no. 4, pp. 164-177.
- 25. Killuru, R.V. (1996). Risk assessment for management, a unified approach. Risk management handbook for environmental, health, and safety professional, McGraw-Hill, New York.
- 26. Levitt, A.M. (1997). Disaster planning and recovery: a guide for facility professionals. Wiley, New York.
- 27. Murphy, B., Kirmeyer, G., (Oct. 2005), Unfinished Business Improving Distribution System Security, Opflow, pp. 18-23.
- 28. Pynchon, M.R., Borum, R., (1999). Assessing Threats of Targeted Group Violence: Contributions from Social Psychology. Behavioral Sciences and the Law*,* no. 17, pp. 339- 355.
- 29. Randazzo, M.R., Keeney, M.M., Kowalski, E.F., Cappelli, D.M., Moore, A.P., (2004). Insider Threat Study: Illicit Cyber Activity in the Banking and Finance Sector. U.S. Secret Service and CERT Coordination Center
- 30. Reason, J. (1997). Managing the Risks of Organizational Accidents. Ashgate, New York.
- 31. Scharfenaker, M.A. (July, 2002). Water suppliers, USEPA, Congress addressing terrorism threat, Journal of AWWA, pp. 16-27.
- 32. Skiba, R., Peterson, R.L., (2003). Threat Assessment Safe and Responsive Schools., http://www.unl.edu/srs
- 33. Teter, D.M., (2003). Biological and Chemical Sabotage of Public Water Systems: A Threat Analysis., Sandia Report, SAND2003-0031, OUO and ECI
- 34. The Clinton Administration's Policy on Critical Infrastructure Protection: Presidential Decision Directive 63. (May, 1998). http://www.fas/org/irp/offdocs/paper598.htm
- 35. The Presidents Commission on Critical Infrastructure Protection. (October, 1997). Critical Foundations Protecting America's Infrastructures. Report of the President's Commission on Critical Infrastructure Protection., Sector Summary Reports. A45*.* Appendix A,
- 36. The White House, (December 17, 2003) Homeland Security Presidential Directive/HSPD-7, Critical Infrastructure Identification, Prioritization and Protection.
- 37. United States Environmental Protection Agency (2004). EPA's Role in Water Security Research. Office of Research and Development National Homeland Security Research Center, EPA/600/R-04/-37.
- 38. United States Environmental Protection Agency, (2002). Vulnerability Assessment Fact sheet., Office of Water EPA 816-F-02-025, (4601M). www.epa.gov/ogwdw/securit/index.html
- 39. United States Environmental Protection Agency, (March, 2004). Water Security Research and Technical Support Action Plan., Office of Water EPA/600/R-04/063
- 40. United States Environmental Protection Agency (January, 2004). Factoids: Drinking Water and Ground Water Statistics for 2003., Office of Ground Water and Drinking Water (4606M) 816-K-03-001
- 41. Vossekuil, B., Borum, R., Fein, R., Reddy, M., (2001). Preventing Targeted Violence Against Judicial Officials and Courts., The Annals of the American Academy, no. 576, pp. 78-90.
- 42. Welter, G.J., Rest, G.B., Moran, K.L., (2003). Actual and Threatened Security Events at Water Utilities. AWWA Research Foundation and the Environmental Protection Agency. Report classified as confidential and proprietary.

APPENDIX A: MLE THREAT ASSESSMENT MODEL & ELICITATION GUIDES

The MLE excel model file with all associated elicitation guides can be requested from the author. The elicitation guides are located within the individual worksheets of the Excel file named "FinalThreatModel13Nov2005.xls". Note: Macros must be enabled in order to use the GUI interface to input and output scenarios.

APPENDIX B: THREAT ASSESSMENT DATA & SENSITIVITY AND IMPORTANT ANALYSIS

This section includes all data used to determine the threat assessment of a scenario. It is on a CD that can be requested from the author. This workbook includes all the input data and output data for all scenarios using a high and low utility environment. It also includes all data associated with the MLE sensitivity and importance analysis.

The file name is AllHighLowAsmtsDataAndSenImpAnalysis.xls.

DISTRIBUTION

2 Department of Civil Engineering University of New Mexico Attn: Kerry Howe MSC01 1070 Albuquerque, NM 87131-0001

