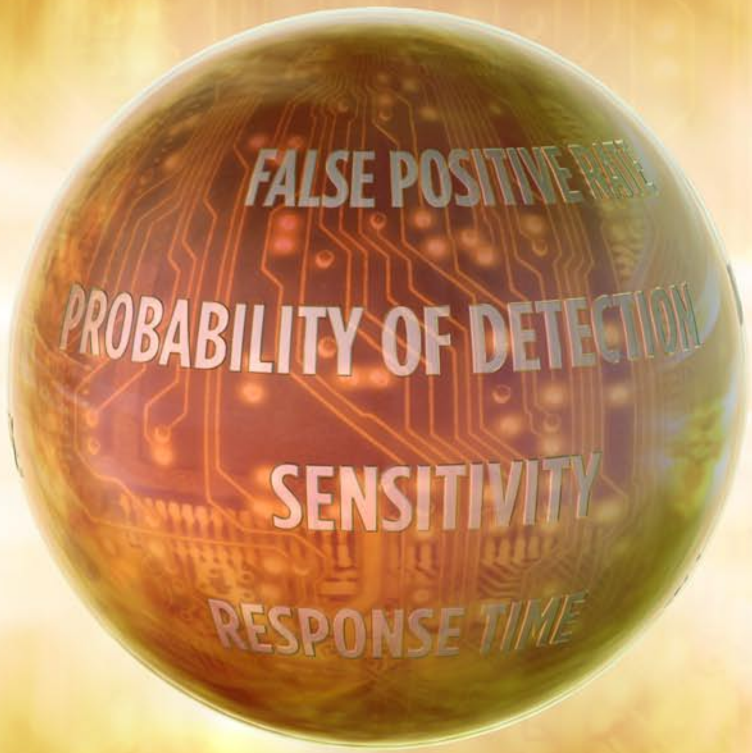
Chemical and Biological Sensor Standards Study



Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **2005**	2. REPORT TYPE **N/A**	3. DATES COVERED **-**

4. TITLE AND SUBTITLE **Chemical and Biological Sensor Standards Study**	5a. CONTRACT NUMBER 5b. GRANT NUMBER 5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER 5e. TASK NUMBER 5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **DARPA Arlington, VA**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S) 11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release, distribution unlimited

13. SUPPLEMENTARY NOTES
The original document contains color images.

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**SAR**	**36**	

Chemical and Biological Sensor Standards Study

Defense Advanced Research Projects Agency

LTC John Carrano, Ph. D.
Study Chair
Program Manager, Microsystems Technology Office

Study Panel

Dr. Tom Jeys (Study Co-Chair)
Dr. Dan Cousins
Dr. Jay Eversole
Dr. Jim Gillespie
Dr. Dennis Healy
COL (ret) Nate Licata
Mr. William Loerop
LTC Mike O'Keefe, Ph. D.
Dr. Alan Samuels
Dr. John Schultz
Dr. Mike Walter
Dr. Ngai Wong
Dr. Bill Billotte
Ms. Maureen Munley
Dr. Erin Reich
Dr. Jason Roos

Table of Contents

Executive Summary

The Department of Defense (DoD) has an urgent need for the development of sensors for early warning and protection of military forces against potential chemical and biological (C/B) attacks. Unfortunately, the current standards and protocols used to evaluate C/B sensors are not adequate; and, this inadequacy hampers the development of new sensors as well as the proper evaluation of current sensors. In order to address these shortcomings, there is a need for a set of sensor metrics and measurement protocols by which the efficacy of C/B sensors can be properly judged.

The performance of a C/B agent sensor is most properly characterized by four key parameters: sensitivity, probability of correct detection, false positive rate, and response time. The sensor's Receiver Operating Characteristic (ROC) explicitly captures the performance trade-off between sensitivity, probability of correct detection, and false positive rate; and, it implicitly captures the performance tradeoff with regard to response time. By specifying the sensitivity of a sensor system, there is an implicit associated detection confidence and false positive rate as dictated by the ROC curve. It is meaningless to specify sensitivity without indicating the probability of detection and the false positive rate. The sensor sensitivity, the detection confidence, the false positive rate, and response time are all related, and all depend on the sensor's operating environment. The relation of the key sensor metrics is shown in Figure A.

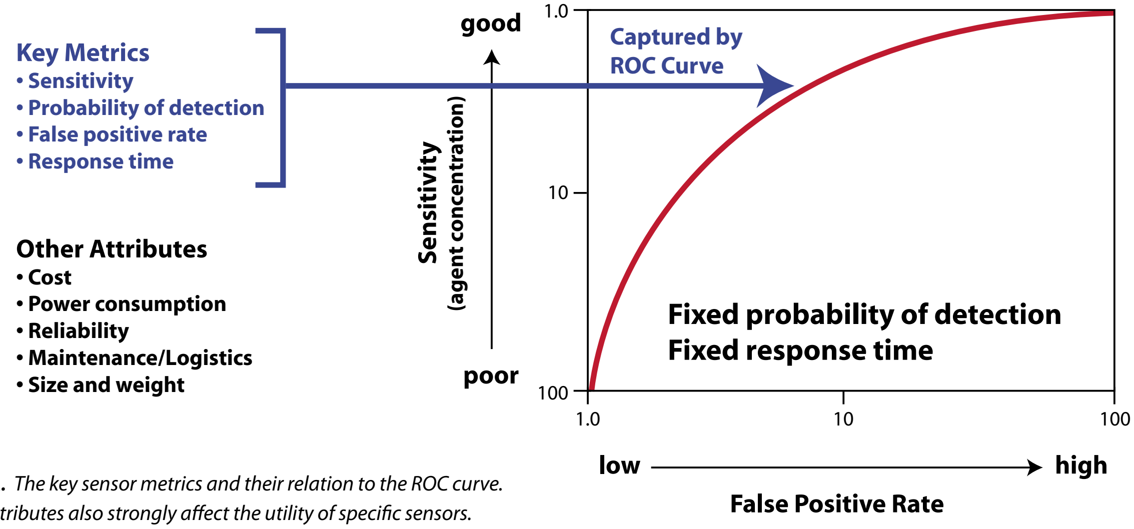


Figure A. The key sensor metrics and their relation to the ROC curve. Other attributes also strongly affect the utility of specific sensors.

Following a simplified vulnerability and threat analysis, the Study found that there is an extremely large uncertainty in the required sensitivity for C/B sensors. This uncertainty in the required detection sensitivity strongly suggests that there is not a single ideal sensor sensitivity for C/B agent detection and that, depending on the acceptable false positive rates and other sensor attributes, different sensors with widely different sensitivity capabilities may have useful roles in the defense against C/B agents.

The impact of the large uncertainty in required sensor sensitivity can be mitigated by the evaluation of sensors over a wide range of sensitivities and false positive rates. Such an evaluation, along with the construction of ROC curves, will enable the operation of sensors at different sensitivities and different false positive rates depending on the perceived threat levels, false positive rate tolerance and mission objectives. The ambiguity in sensor requirements further suggests that, depending on the acceptable false positive rates and other sensor attributes, different sensors with widely different sensitivity capabilities may have useful roles in the defense against C/B agents.

Given this ambiguity in the threat, it is preferable to have sensors capable of switching from one operating mode to another based on the mission. For example, in a low C/B threat condition the user would have the ability to adjust a sensor to operate with a very low false positive rate at the expense of sensitivity. Whereas, in a high C/B threat situation, the user would have the capability to adjust the sensor to operate with higher sensitivity and accept the accompanying higher false positive rate. It is also likely that users will want to increase the sensor sensitivity immediately after an attack so as to have greater protection from another potential attack. Figure B illustrates the utility of multiple and adjustable sensor sensitivities. Regardless of the C/B threat, it is essential that the sensor ROC curve be established so that appropriate and predictable performance adjustments are made possible.

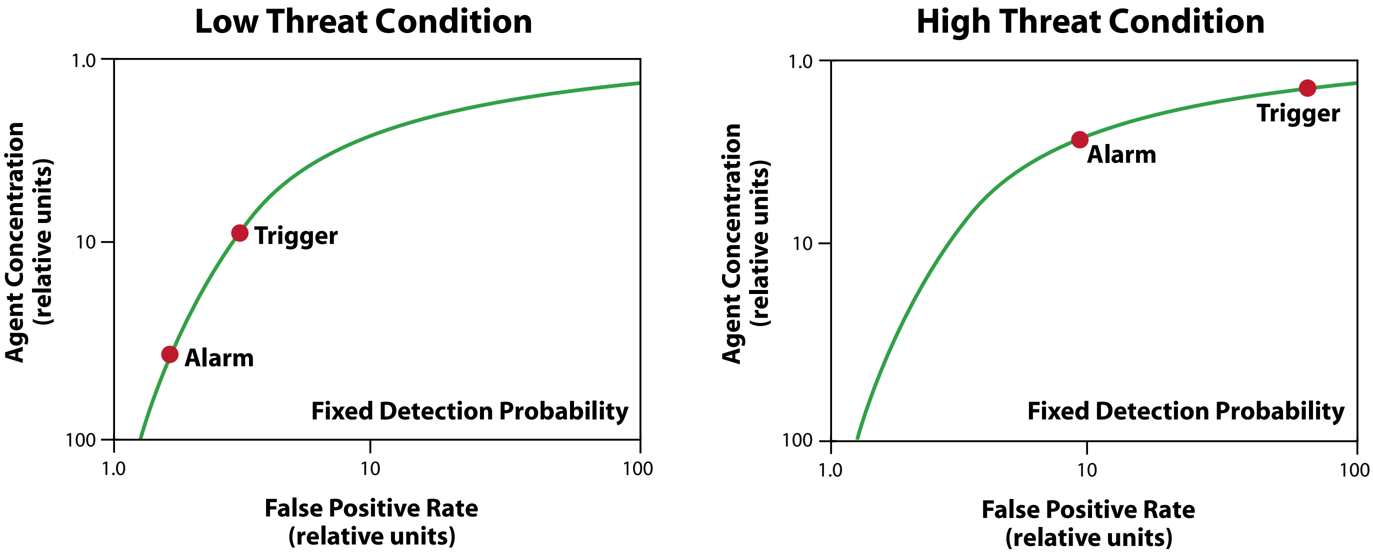


Figure B. *Dependent upon the threat condition, the sensor trigger and alarm thresholds should be adjustable along the sensor ROC curve. In the context of this report, an alarm refers to an event that results in a high disruption action, and a trigger is an event that results in a low disruption action.*

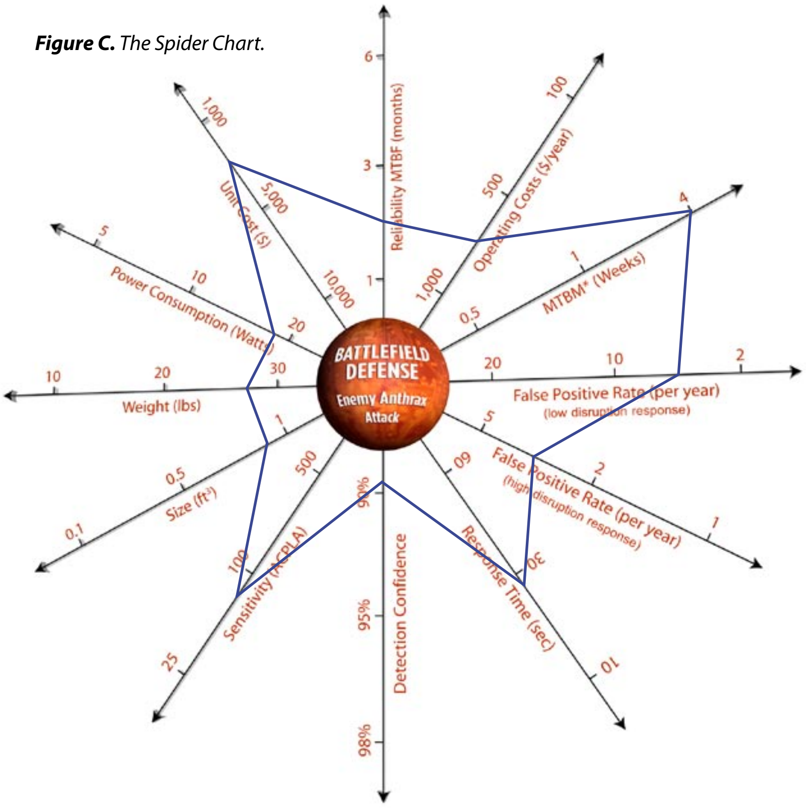
In order to capture the overall performance of a C/B sensor, we employ a graphical technique known as the "the spider chart," as shown in Figure C. Each of the sensor metrics is assigned a "leg" on the chart, with "better" performance moving out from the center. One can then plot the performance of a sensor, as shown by the blue line in Figure C. The spider chart can then be used as a means of comparison between sensors, or simply as a means to judge the overall efficacy of a given sensor.

In addition to identifying C/B sensor metrics and measurement protocols, the Study Panel produced five key findings, which are:

1. **Sensor testing and characterization is inadequate.**
2. **Potential threats span a very wide dynamic range (> 10^6 in concentration).**
3. **ROC curves are essential for sensor development, testing and evaluation.**
4. **Sensors should allow for multiple operating modes.**
5. **Sensor requirements are not well defined.**

The report details these key findings and makes subsequent recommendations.

Figure C. *The Spider Chart.*



INTRODUCTION

As the threat of a chemical or biological attack on the United States homeland and military forces abroad continues to grow, the Department of Defense's (DoD) sense of urgency to develop effective chemical and biological (C/B) sensors to mitigate this threat also grows. Presently, the standards and protocols used to evaluate C/B sensors are antiquated and inadequate for today's expanding requirements. This inadequacy hinders the development of new sensors and the proper evaluation of current sensors for use throughout the DoD community. In order to address these shortcomings, a set of sensor metrics and measurement protocols need to be developed thereby allowing for the evaluation of C/B sensor efficacy, and affording a common language to communicate ideas, developments, failures and successes. As a direct result of this need, the Defense Advanced Research Projects Agency (DARPA) commissioned the Chemical and Biological Sensor Standards Study (CBS3) by assembling a panel of C/B detection experts to evaluate this problem and develop concrete, viable solutions to begin its resolution.

Throughout the study, the panel kept the user (i.e. the Warfighter) at the forefront of the discussion. It is paramount that the sensor standards and evaluation protocols align with the DoD mission needs and requirements. These needs include C/B sensors that are: highly performing, efficient, reliable, easy-to-use and readily integrated into the Warfighter concept of operations. In order to facilitate the development of sensors that meet these needs, guidelines must be established that detail both real world performance metrics for developers to build to, and measurment protocols that the testing community can employ to evaluate C/B sensors. While simple in concept, this task becomes exceedingly difficult when considering the vast number of possible mission scenarios and technology applications.

In order to limit the scope of this effort, the panel focused a set of scenarios employing **detect-to-warn** sensor technology with **aerosol dissemination** of the C/B threat.

The CBS3 Panel[1] identified five key findings, which are:

1. The potential agent threat spans a dynamic concentration range greater than 10^6.
2. Requirements for sensors are not well defined.
3. Receiver Operating Characteristic (ROC) curves are essential for the development, testing and evaluation of C/B sensors.
4. Sensors must allow for multiple operating modes allowing for adjustment of sensitivity and false positive rate to meet operational requirements.
5. C/B sensor testing and characterization is outdated and inadequate.

This report details these key findings and makes subsequent recommendations.

Four distinct scenarios were developed that represent both the Force Protection (i.e., Movement to Contact and Defensive Scenario) and the Fixed Facility situation (i.e., External Attack on a Fixed Facility and Internal Attack on a Fixed Facility). The panel recognizes that these scenarios do not include all plausible scenarios, sensor architectures, or dissemination techniques; however, they do provide an exemplary setting by which to describe parameter values in terms of key sensor metrics to address present concerns of the DoD for various scenarios. Further discussion of the selected scenarios can be found in Appendix C.

It was the intent of the study to remain technology agnostic so that the recommendations from this study may apply to all sensors. Ultimately, it is the intent of this report to stimulate thought and galvanize the detection community such that the appropriate emphasis will be given to developing C/B sensor standards and testing guidelines that will ensure sensors meet the needs of the Warfighter. It is the Panel's expectation that this report will help to focus future C/B development and testing efforts.

Author's Note

For the purpose of this report, there are many different possibilities and scenarios against which any one sensor could be evaluated. These scenarios are not meant to represent the entire spectrum of the C/B challenge, rather, they will evolve as the mission changes, as the threat is better understood, and as they are adopted for other purposes.

The ultimate goal is to substantially reduce the Warfighter's exposure to C/B agents, and as such the focus is on detect-to-warn systems, which by definition provide ample warning to personnel before encountering a C/B threat such that countermeasures can be taken to avoid exposure. In principle, a detect-to-warn system should prevent exposure and the need for treatment. Depending on the threat condition, two types of protective actions are considered when estimating the acceptable false positive rate. The first is a *high disruption action* and the second is a *low disruption action*.

A *high disruption action* (initiated by a sensor alarm event) has consequences that result in a significant interruption in the normal or planned activities. For instance, a high disruption action would be a significant change in the tactics and procedures impacting operations on the battlefield. Similarly, in consideration of a fixed facility, a high disruption action would shutdown normal operations at a Command Center, for example, resulting in evacuation or shelter-in-place. A *low disruption action* (initiated by a sensor *trigger* event) has consequences that result in no significant interruption in normal or planned activities. An example of a low disruption action on the battlefield might be initiating NBC reconnaissance and donning of protective equipment. Additional sample acquisition and testing, or alterations in building airflow without a perceived difference by personnel are also representative of low disruption actions.

Footnotes

[1] The Study Panel is composed of respected scientific and technical subject matter experts from Defense Advanced Research Projects Agency, Defense Threat Reduction Agency, Pacific Northwest National Laboratories, Massachusetts Institute of Technology Lincoln Laboratories, Booz Allen Hamilton, DoD Joint Requirements Office, DoD Research, Development and Engineering Command, DoD Joint Program Executive Office-Chemical and Biological Defense, US Naval Research Laboratory, and the US Army Research Laboratory. The Panel also received invaluable input from the US Army Dugway Proving Ground and from the Honorable Dr. Richard Danzig, former Secretary of the Navy.

SENSOR METRICS

Receiver Operating Characteristic (ROC)

The performance of a C/B agent sensor is most properly characterized by a number of interrelated parameters such as sensitivity, probability of correct detection, false positive rate and response time. In the operation of a sensor, it is generally possible and often useful to vary one or more of these parameters in order to optimize sensor performance for specific applications. For instance, in some circumstances one might be willing to accept a higher false positive rate in order to obtain a better sensitivity. The sensor's Receiver Operating Characteristic (ROC) quantitatively captures the performance trade-off between sensitivity, probability of correct detection and false positive rate (see Figure 1). There are several variants, but, for continuously operating C/B detection sensors, it is useful to consider ROC curves that relate sensor sensitivity rate to false positive at a given detection confidence.

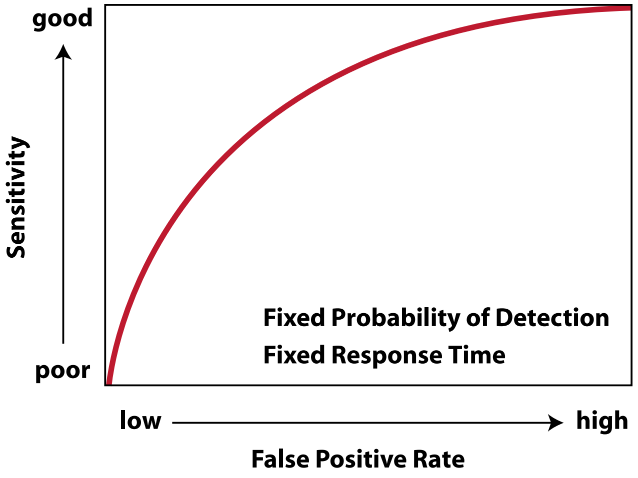


Figure 1. *Receiver Operating Characteristic (ROC) curve. The ROC curve directly relates two important sensor performance metrics: the sensitivity and false positive rate. In the example shown above, the curve is drawn for a fixed probability of detection; and, because the response time affects the probability of detection and the false positive rate, it is implicitly represented in the ROC curve. As the sensor is adjusted from poor to good sensitivity the false positive rate increases. Since sensors must operate in both low and high threat conditions and with varying tolerance for false positives, it is important to understand the tradeoff between various operating points of the sensor. The ROC curve captures this tradeoff.*

The ***sensitivity*** of the sensor system is generally understood to be the minimal detectable agent concentration and is typically given in units of particles or mass per unit air volume (particles/liter or mg/m^3). By itself, sensitivity is a rather incomplete specification of sensor performance. Due to sensor noise and the confounding effects of interferents (i.e., clutter) in the environment, detecting an agent is not assured during field operation, and in fact is not assured even in the laboratory from one identically prepared experiment to the next. Situations in which masking effects due to interferents or noise cause the sensor to fail to detect the agent is an error referred to as a ***missed detection*** (also known as a ***false negative***). A second form of error is the ***mistaken detection*** (also known as a ***false positive***), in which the sensor mistakes noise and clutter signals as an indication of the presence of an agent of interest, when in fact none is present.

By specifying the sensitivity of a sensor system, there is an inherent associated detection confidence and false positive rate as dictated by the ROC curve. The ***detection confidence*** is the probability of detecting an agent when in fact one is present at or above the specified concentration. The ***false positive rate*** is the rate (events per unit time) at which the sensor mistakenly detects the presence of an agent when in fact none is present. It is meaningless to specify sensitivity without indicating the probability of detection and the false positive rate. For example, one could improve the apparent sensitivity of confident detection by setting the detector to alarm at anything. Furthermore, the sensitivity and associated probability of detection are meaningful only when generated in the context of a real-world environment. The sensor sensitivity, the detection confidence, the false positive rate, and the response time are all related, and all depend on the sensor's operating environment.

A characterization of a notional sensor's (detection) performance consists of a family of ROC curves as shown in Figures 2 and 3. Figure 2 shows a family of ROC curves for different levels of detection confidence in a single environment. These curves are generated, in principle, by repeatedly adjusting the sensor detection threshold and measuring the detectable agent concentration with a fixed detection confidence and false positive rate in a given background. Figure 3 shows an example of ROC curves for a fixed detection confidence in various environments. The false positive rate generally depends upon the environment in which the sensor is operating, and thus the ROC curves also depend on the environment.

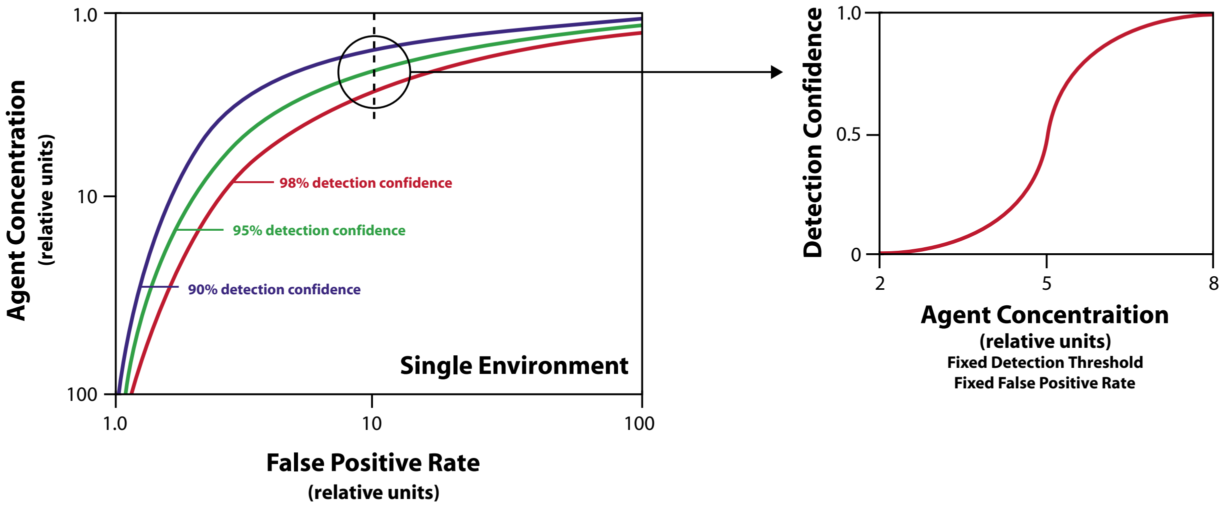


Figure 2. *Receiver Operating Characteristic (ROC) curves for different levels of detection confidence in a single environment, with a fixed reponse time.*

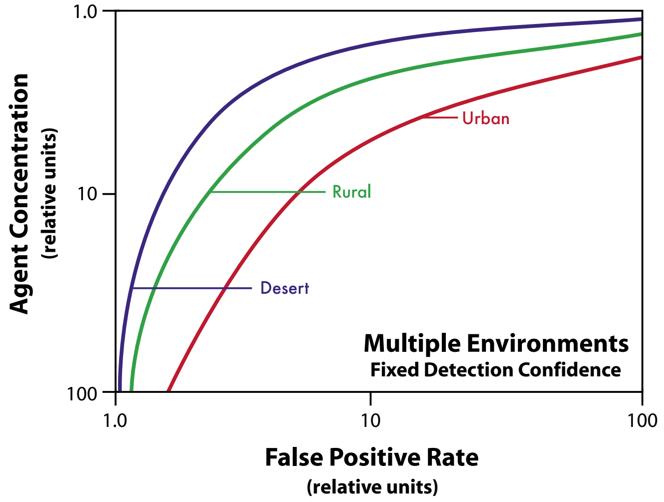


Figure 3. *Receiver Operating Characteristic (ROC) curves for different environments at a fixed detection confidence and fixed response time. Because the level of background clutter varies from one environment to another so do the ROC curves. Typically an urban environment has a much wider range of clutter than does a desert environment, and at a fixed sensitivity a sensor operating in an urban environment will have a higher false positive rate than a sensor operating in a desert environment.*

The ROC curves illustrated in Figures 2 and 3 characterize the performance of a sensor to the extent that:

- The environments in which the sensor has been tested represent the environments in which it will be used, and
- The ranges of ambient conditions and activities that impact the sensor have been explored in each environment.

Measurement of ROC curves

Ideally the ROC curves for a given sensor are constructed by operating the sensor in various environments; challenging the sensor with various agents, agent concentrations, agent temporal profiles and measuring the agent detection probability and false positive rate as a function of the sensor's concentration threshold setting. Unfortunately, this ideal approach has three major flaws. First, it is generally not possible to release agents (and often, even agent simulants) into environments of interest. Second, once an agent has been released into an area/environment, it is difficult to quantify how much was actually present in each sensor sample. Third, this ideal approach requires a very large amount of field data, which is difficult and expensive to aquire.

To discuss practical methods of determining the ROC curves for a given sensor, it is necessary to discuss the nature of real environments, the physical causes of variation in sensor responses, and the process by which ROC curves can be determined from distributions of sensor responses. Methods of measuring sensor response distributions and practical limitations on testing C/B agent sensors to identify useful sensor testing protocols can then be considered.

The atmosphere contains a huge array of trace biological molecules, chemicals, and particulates resulting from ecological processes, animal activities, human activities, weather, fires, and myriad other sources. At any given time, there is a wide range of materials present in the air, including minerals, pollen, dust, mold, bacteria, trace chemical vapors and so forth. For example, over one thousand chemical species have been found in diesel exhaust and their concentrations vary significantly from engine to engine, with fuel/air/engine temperatures, fuel types and many other factors. The ambient make-up of air at a given location is constantly changing due to wind, weather events, reactions among chemicals, reactions of chemicals to light, condensation, evaporation, precipitation, lofting, changes in lighting conditions and the chemical sources, etc. Every air sample contains a different and vastly complicated mix of biologicals, chemicals, and particulates. Together, the physical location and the distribution of natural conditions and activities at that location define a measurement environment.

A sensor's response to a given air sample consists of three major components: noise, clutter, and signal. Noise is fluctuation in sensor responses due to factors that are internal to the sensor and independent of the measurement environment. Clutter is the sensor response to all factors associated with the measurement environment other than the agent. Signal is the sensor response to the agent. Due to the variations in the measurement environment, the clutter component of a sensor response is different in different environments. While the background in a given environment is generally far too complex to measure or model, sensor responses to various environments are relatively easy to measure and can be assembled into distributions that can be used to characterize sensor performance in that measurement environment. The distribution of sensor responses for a given measurement environment is fixed, to the extent that the likely range of weather and activity conditions has been sampled and no major new activities or events take place close enough to affect the ambient materials found there.

In view of these considerations, and the complexity of trace C/B materials in most environments, the best practical method of evaluating sensor performance is to:

- Acquire raw sensor responses for a variety of agents and agent concentrations in laboratory settings.
- Acquire raw sensor responses in a variety of environments.
- Appropriately combine the raw agent and environmental response data.
- Pass the combined data through the sensor detection algorithm.
- Record the detection probability and false positive rate for various algorithm sensitivity thresholds.

Then for each challenge agent concentration, the false positive rate that corresponds to a predetermined detection probability (or confidence) is measured. Figure 4 illustrates the process of acquiring a sensor's ROC curves.

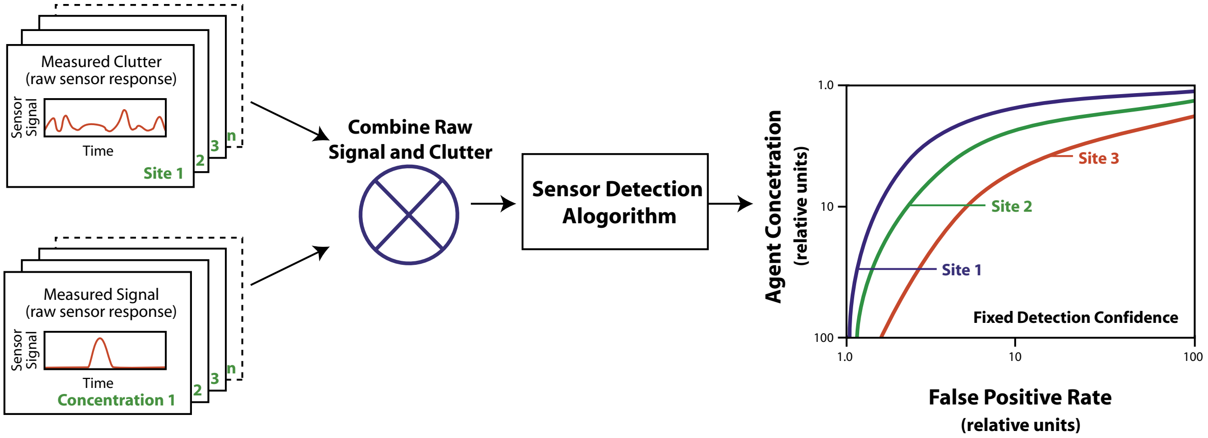


Figure 4. *Illustration of the process for generating a sensor's ROC curves. Raw sensor responses are acquired for many different environments (sites) and for many different agent concentration challenges. These raw responses are appropriately combined and processed by the sensor detection algorithm. At a fixed detection confidence the false positive rate is recorded for various algorithm threshold settings corresponding to different concentration sensitivities.*

Recommendations

- Sensitivity should always be stated with the probability of detection, the false positive rate and the response time.
- Sensor testing must occur in environments in which the sensor will operate and must be generated with different levels of detection confidence in a single environment.
- ROC curves must be developed for sensors at each stage of the development to determine readiness for the next developmental stage.

Response Time

Another important figure of merit for continuously operating sensors is the ***sensor response time***, meaning the time interval between the arrival of the target agent concentration and the sensor detection declaration. The response time is illustrated in Figure 5. The required response time for a given sensor is dependent on the intended mission. This report focuses on detect-to-warn missions in which it is imperative that the sensor response time is less than the time necessary to take protective actions.

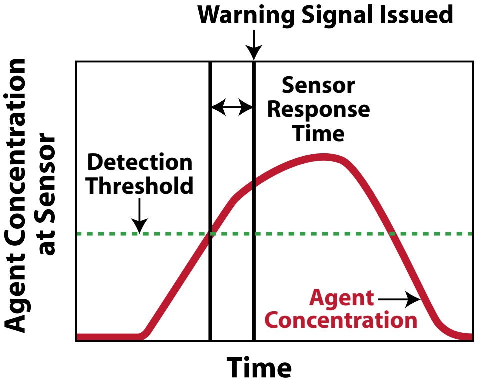


Figure 5. *The sensor response time is the time interval between when the agent concentration reaches the sensor's specified sensitivity (or detection threshold) and the time that the sensor issues a detection signal.*

Vulnerability Analysis

The purpose of a C/B agent sensor, when operated in a detect-to-warn mode, is to prevent or to reduce exposure to an agent. Inhalation exposure is measured as a dosage, which is the time-integrated accumulation of agent in the human lung. Dosage has units of mass for chemical agents and biological toxins, and units of number of Colony Forming Units (CFU) or Plaque Forming Units (PFU) for bacterial or viral biological agents.

A major problem that must be confronted when considering the evaluation of C/B agent sensors is the fact that these sensors do not actually measure dosage, but rather they measure concentration. The exact relationship between agent dosage and detection concentration is poorly defined. An estimate of the required detection concentration for a given dose is based on the range of many variables, including temporal profile of the agent, particle size distribution, breathing rate, exposure time, etc., resulting in a dynamic range of greater than 10^6 for the required detection concentration. (see inset for detailed calculation). This extremely large uncertainty in the required detection sensitivity strongly suggests that there is not a single "magical" ideal sensor sensitivity for C/B agents.

The impact of this uncertainty can be mitigated by the evaluation of sensors over a wide range of sensitivities and false positive rates. Such an evaluation, along with the construction of ROC curves, will enable the operation of sensors at different

Calculating Agent Concentration from Agent Dose.

Ignoring the time variation of the agent concentration and the distribution of agent particle size, the required sensor concentration sensitivity for a given agent dosage is as follows:

$$C = \frac{D}{\frac{4\pi}{3} r^3 \rho f B \eta T}$$

See the table for the variable designation.

Using Bacillus anthracis as an example and the values indicated in the table below, the required detection concentration varies from 0.3 particles/liter to 1 x 10^5 particles/liter. **The values for the required detection concentration assume the extremes range for each parameter to roughly account for all possibilities.*

Table 1. *Variables, estimated values and estimated value ranges for calculating agent concentration from agent dosage*

Variable	Estimated Value	Estimated Value Range
Dosage (D) for LD_{50}	10,000 cells	20,000-5,000 cells
Particle radius (r)	1.5 μm	0.4-2.5 μm
Agent density (ρ)	1 cell/(0.8 μm)^3 =0.5 cells/μm^3	0.5-1.0 cells/μm^3
Fill factor (f)	0.6	0.3-0.7
Breathing rate (B)	15 liters/min	10-20 liters/min
Lung retention efficiency (η)	0.5	0.3-0.7
Concentration (C)	63 ppl	10^5-0.3 ppl

sensitivities and different false positive rates depending on the perceived threat levels, false positive rate tolerance and mission objectives. The ambiguity in sensor requirements further suggests that, depending on the acceptable false positive rates and other sensor attributes, different sensors with widely different sensitivity capabilities may have useful roles in the defense against C/B agents.

Adjustable Operating Modes for Sensors

Given the ambiguity in the C/B threat, it is desirable to have sensors capable of switching from one operating mode to another based on the mission. For example, in a low C/B threat condition the user would have the ability to adjust a sensor to operate with a very low false positive rate at the expense of sensitivity. Whereas, in a high C/B threat situation, the user would have the capability to adjust the sensor to operate with higher sensitivity and accept the accompanying higher false positive rate. It is also likely that users will want to increase the sensor sensitivity immediately after an attack so as to have greater protection from another potential attack. This is consistent with the concept of "reload" as described by Richard Danzig in his seminal paper "Catastrophic Bioterrorism: What is to be done."[2] Figure 6 illustrates the utilization of ROC curves to set sensor trigger and alarm thresholds. It is paramount that a sensor has well characterized ROC curves such that intelligent adjustments in operational modes can be achieved. How and when a user chooses to adjust modes is dependent on operational conditions. Clearly, affording the opportunity to adjust the sensor performance is predicated on **good** training of the user or decision maker, in the operation of the sensor. Figure 6 also depicts the operation of a sensor with simultaneous dual thresholds. The "Alarm" threshold indicates a lower false positive rate at the expense of reduced sensitivity. This "Alarm" level is used to initiate a "high disruption" action (as defined on page 7 of this report). The "Trigger" level provides improved sensitivity, but with the concomitant increase in the false positive rate. The "Trigger" level is used to initiate "low disruption" actions. By employing both the trigger and alarm levels simultaneously, the sensor can achieve excellent sensitivity without generating an unacceptable number of high disruption actions.

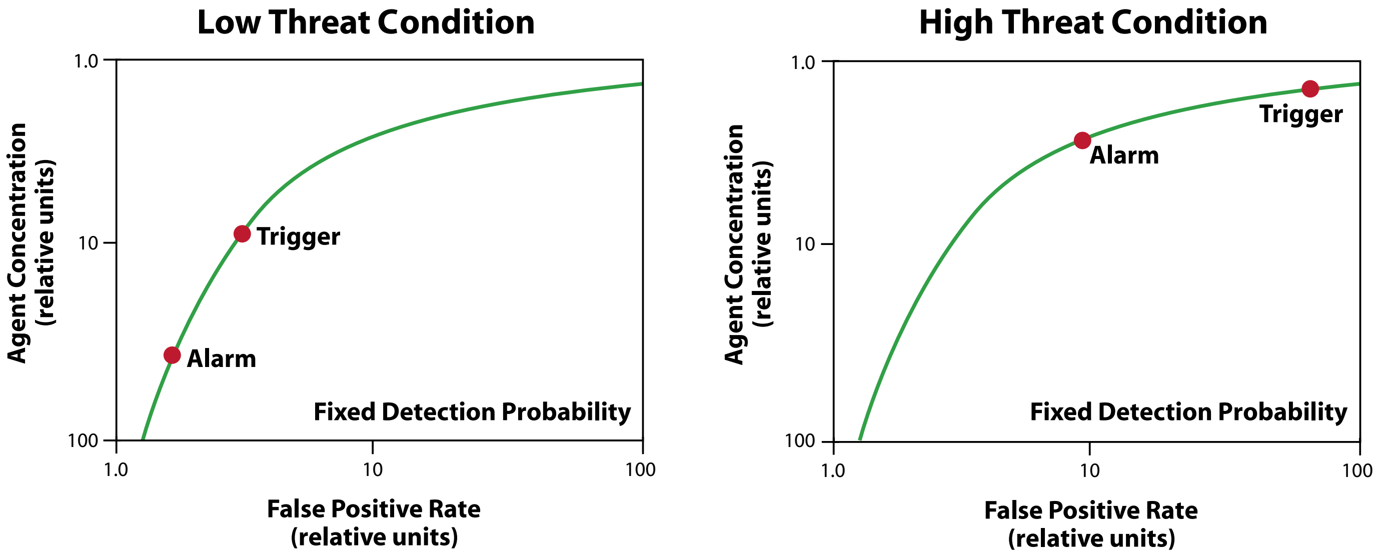


Figure 6. *Dependent upon the threat condition, the sensor trigger and alarm thresholds should be adjustable along the sensor ROC curve. In the context of this report, an alarm refers to an event that results in a high disruption action, and a trigger is an event that results in a low disruption action.*

Footnotes

[2] R. Danzig. August 2003. *Catastrophic Bioterrorism - What Is To Be Done?* Center for Technology and National Security Policy, National Defense University.

Other Sensor Attributes

In addition to the sensor performance metrics there are several other important sensor attributes that strongly affect the sensors utility for various missions. We specifically differentiate between "metrics" and "attributes" owing to the relative ease with which one can both define and measure the attributes vice the more complex metrics. For example, the size and weight of a sensor are easily quantified whereas its ROC curve is much more difficult to generate.

The sensor ***initial cost*** affects how a sensor is employed and the numbers of sensors employed. Disposable sensors should be very inexpensive. While non-disposable sensors deployed with units on the battlefield could cost significantly more. In contrast, a single sensor unit for protecting a facility from external attack could be quite expensive, whereas multiple sensors employed for internal attacks might need to cost less. Depending on performance and mission requirements, sensor costs could change dramatically. Another related aspect is the ***operating cost***, which is comprised of any costs incurred after the initial acquisition cost. This includes both logistic and maintenance costs, consumable supplies, repair parts and operator training. Operating cost can range from very low for disposable sensors, to lifetime costs that greatly exceed the initial cost of the sensor for more paramount sensors. If only one sensor is to be maintained, a higher operating cost may be more tolerable than in a situation where large numbers of sensors are deployed.

Power consumption is another important sensor attribute. For force protection roles, the sensor should typically be battery powered. A disposable sensor should have a very low power requirement. For building sensor systems, an AC line would be available to provide power. The power consumption must also be considered in light of the ***mission duration***.

Maintenance consists of the actions taken to keep materiel in a serviceable condition or restore it to serviceability. Although maintenance may be carefully followed, sensor failures will still occur based on the ***reliability*** of the system.

Reliability is the probability that an item will perform to its intended function for a specified interval under stated conditions. The longer the sensor performs without experiencing an unexpected failure (i.e., the mean time between failure or MTBF), the better the reliability. We assume here that stated routine maintenance requirements are met.

Ruggedness is the ability to withstand shock, vibration, exposure to harsh weather conditions and even some effects of enemy nuclear weapons (e.g., electromagnetic pulse or EMP).

The ***form factor***, i.e., the size, weight and shape of the sensor, is of particular concern in the battlefield role where sensors are frequently moved. Man portable, small sensors are highly desirable in this role. Small form factor is normally less critical from the facility standpoint because sensors will usually remain in place.

The final attribute, ***environmental considerations***, is the set of guidelines meant to protect the environment, the military and noncombatant civilian populations. These include issues such as safe disposal of reagents and used consumables to excessive noise and laser eye-safety. These can have a serious impact on sensor acceptance.

Thus, the overall performance of a C/B sensor can be characterized by a small number of interrelated key metrics: sensitivity, probability of detection, false positive rate and response time. However, other sensor attributes (i.e. cost, power consumption, reliability, size and weight, maintenance, ruggedness and environmental/safety issues) can impact sensor acceptance and fielding. The exact requirement for each of these metrics is, however, dependent on the concept of operations for the mission.

Recommendations

- **The sensor requirements must be determined from operational scenarios.**
- **Sensors must be developed with different performance characteristics for different missions, i.e., one "size" sensor does not fit all applications.**
- **A sensor "knob" is needed to switch from low false positive rate (lower sensitivity) to high sensitivity (higher false positive rate) depending on perceived threat level.**
- **The sensor should be designed such that simultaneous dual thresholds are possible to mitigate high disruption actions.**

Figure 7. *The Spider Chart.*

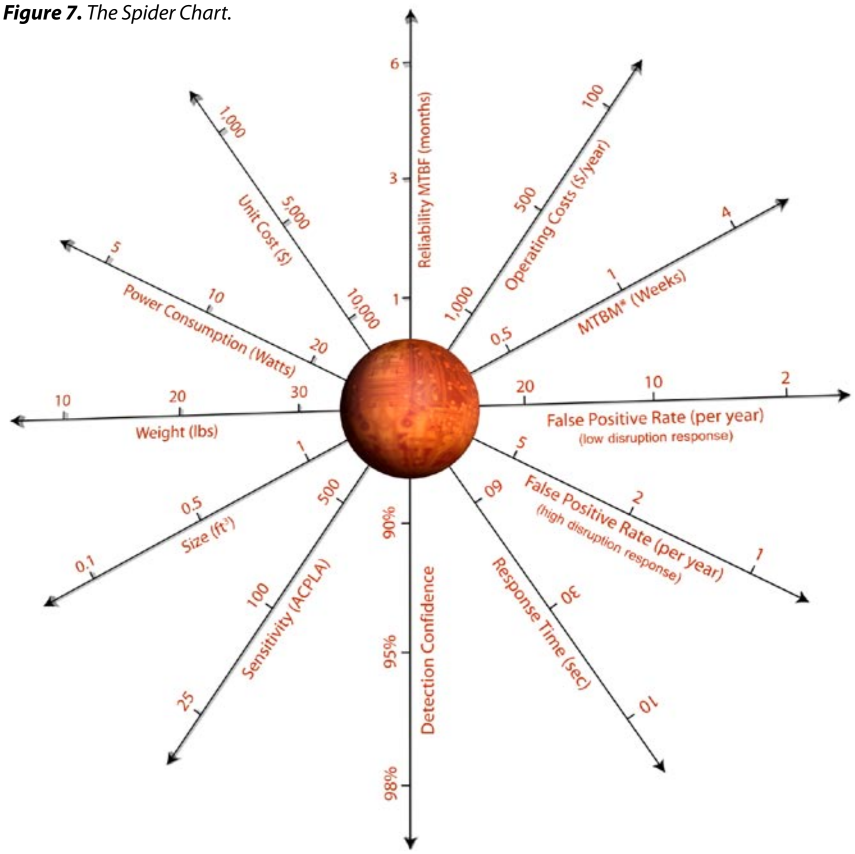
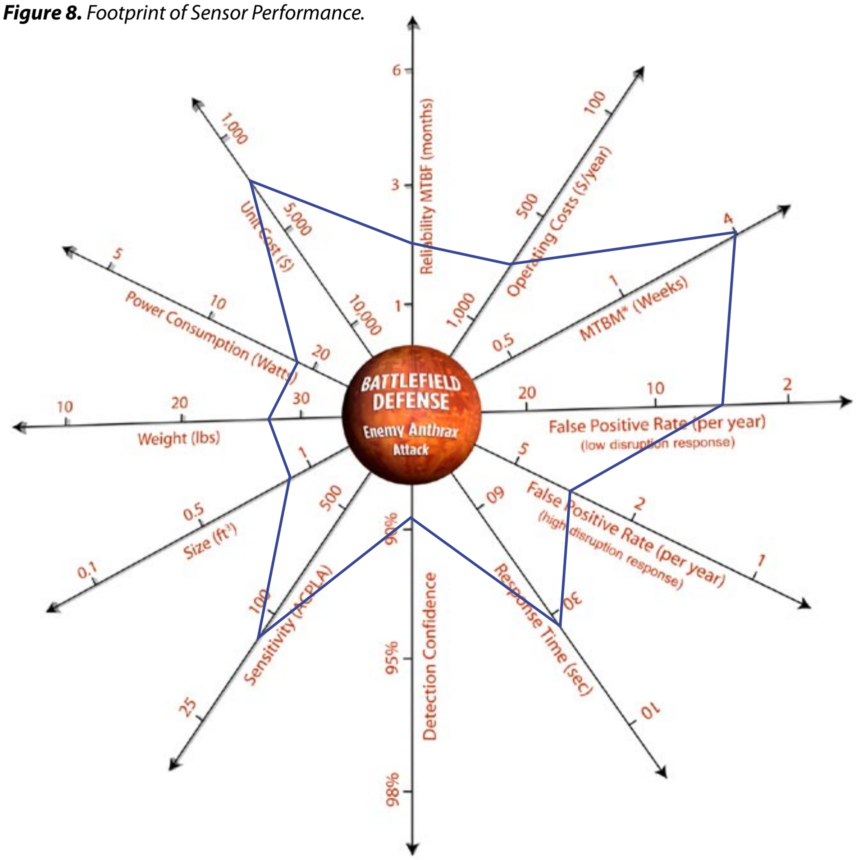


Figure 8. *Footprint of Sensor Performance.*



The Spider Chart

To capture the overall performance of a sensor we use a graphical technique called "the spider chart." The spider chart pictured in Figure 7 integrates all of the sensor metrics into one visual chart, which can be utilized to evaluate the overall requirements and performance of a sensor. Each of the previously defined sensor metrics and attributes are assigned to a leg of the spider chart. The tick marks on each leg represent measured or predicted values that are associated with the respective metric and improve in value as they radiate out from the center of the chart. The values depicted on the chart for each metric and attribute are designed to reflect the requirements of the stated mission.

On any given leg of the spider chart, we find plotted three values. The center value represents an average acceptable value for that particular metric. The inner and outer values can be thought of as the "error bar" associated with that metric. The innermost value may be considered the minimally acceptable performance, and the outermost value the point at which one reaches diminishing returns. As shown by the blue line in Figure 8, connecting the metric values provides a "footprint" of the sensor performance with regard to mission requirements. The larger the area encompassed, the better the sensor for the designated mission. This footprint is a valuable tool for comparison of multiple sensors with common mission requirements, as well as comparison of sensors to a well-defined set of metrics for a concept of operations.

The spider chart is an effective tool throughout the entire developmental cycle. In the early stages of R&D, the spider chart plot may represent the expected performance of any given sensor, based on design predictions and preliminary emperical data. In the later stages, the spider chart plot may represent the result of extensive prototype testing.

SENSOR TESTING

Performance testing of C/B sensors is an elaborate process that is essential throughout the sensor development life cycle. Ideally, a testing protocol will mimic "real life" conditions; however, this notion is impractical given the numerous variables (e.g., wind speed, temperature, humidity, cloud concentration, agent concentration, etc.) involved in a vapor or an aerosol release of a C/B weapon. As such, sensor testing generally starts in a controlled setting such as a laboratory or chamber and moves to open air test areas as data is collected analyzed and understood. When in a laboratory setting, live C/B agents can be employed to evaluate sensor performance; however, as the testing moves to the open-air environment, for legal and practical reasons, live agent is not used. Instead, simulants of C/B agents are used at designated and approved testing areas.

A ***simulant*** is a benign chemical compound or a biological material that mimics a chemical warfare agent (CWA) or a biological warfare agent (BWA) that might be used in an attack. An ideal chemical simulant will mimic structural and physical properties (e.g., size, aerosol dynamics, spectral properties, volatility, density, etc.) of the agent or agents of interest. For example diphenyl chlorophosphate is a commonly used chemical that is utilized in lieu of nerve agents (e.g., Tabun, Sarin, and Soman) in testing chemical agent sensors. A biological simulant will mimic the structure, physical properties and pathogenic properties (e.g., genetic similarity, aerosol dynamics, size, shape, etc.) of the agent of interest. For example, *Bacillus globigii* is a widely used bacterium that is utilized in lieu of *Bacillus anthracis* in testing biological agent sensors; both are gram positive, spore forming bacteria that have similar structural and genetic properties.

Additionally, to mimic "real life" conditions it is important to incorporate environmental influences or interferents that may be present during a particular mission and affect the sensor's performance. In the context of CWA and BWA sensing, an ***interferent*** is any material present in the environment that retards and/or inhibits a detector's ability to accurately detect the agent or agents for which it is programmed, or causes a detector to false alarm or malfunction. Although the entire spectrum of interferents is unknown, some examples of common interferents are: fuel products, engine exhaust, pollen, smoke, dust, dirt and human skin or dandruff flakes.

Presently, the testing protocols are not adequate to evaluate sensors and extrapolate their performance to the real world environment. Thus, it is vital to both the Warfighters and sensor developers to create well-defined test protocols that can be used to evaluate sensors.

The fundamental purpose for testing and evaluation is to assess the performance of sensors. One of the most frustrating properties that Warfighters experience with C/B sensors is a high false alarm rate. Currently, the existing protocols are designed to establish the performance of a C/B sensor by testing the lower detection limit of a system against a variety of contrived backgrounds, which attempt to replicate battlefield conditions. Regrettably, this protocol currently fails to quantitatively assess the operational false alarm rate. In order to make a statistically sound assessment of the actual performance of a sensor, an experiment must be designed such that an objective assessment of the sensor's efficacy in both agent detection and false alarm rejection is captured.

Referee equipment used to support both chamber and field trials is also limited, especially in light of advances in detector technology. Current referee equipment often fails to monitor the simulant aerosols at adequate temporal and spatial resolution or at adequate collection rate to provide truth data for developmental detectors. Delivery systems in use do not provide adequate control over metered injection into the chamber or field release to afford an adequate span over the range of concentrations that represents the expected threat.

Generally, the measured properties of an environment should be as complete as possible while preserving a focus on collecting such background data with the actual sensor under development. In order to develop an effective system performance model, it is important to record statistically significant information on the sensor response in an operational environment.

The aforementioned challenges need to be solved before adequate standards and protocols for the evaluation of chemical and biological sensors can be implemented. While it may be that the current test community can solve these problems if given adequate resources, it is the panel's recommendation that an independent and dedicated oversight authority be established to do this. A dedicated panel may be better equipped to solve these challenges and to identify new challenges as they occur. It is envisaged that the oversight authority would focus on the key issues that the testing community encounters such as:

- Increasing the number of simulants employed to evaluate developmental sensor performance.
- Identifying simulants whose properties more closely represent the agents they are meant to emulate in more than one variable (e.g., physical properties and chemical composition). When possible, nonvirulent forms of a pathogen species can be employed in lieu of the distantly related microbes commonly used in current tests.
- Improving the validation metrics and measurement equipment to better quantify test aerosols (spatially and temporally).
- Presenting aerosol challenges with independently controlled concentration and particle size.
- Performing studies of the ambient environment over a wide variety of operational environments to statistically quantify the types and quantities of interferents found.
- Producing time-resolved background measurement sets in several standardized operational environments or with standardized ambient environment challenges (accelerated false positive tests) to assess the frequency of false positives.
- Developing a comprehensive and standardized signal database using target (live) agents.
- Developing test plans under experiment design theory to assure a statistically significant number of trials to construct the ROC curve.

Recommendations

- **Establish an independent and dedicated oversight authority to improve standardized testing methodologies and equipment.**
- **Research and development on improved testing methods appropriate to sensor development stages.**
- **Update and modernize test equipment.**
- **Develop a detailed understanding of threat agent characteristics.**

FINDINGS AND RECOMMENDATIONS

The Panel highlighted five major findings from the Study. A review of each of the findings, along with corresponding recommendations is listed below:

1. ROC curves are essential for the development, testing and evaluation of C/B sensors.

Recommendations:

- Sensitivity should always be stated with the probability of detection, false positive rate and response time.
- Sensor testing must occur in environments in which the sensor will operate and ROC curves must be generated with different levels of detection confidence.
- ROC curves must be developed for sensors at each stage of development to determine readiness for the next development stage.

2. Threats span a dynamic concentration range greater than 10^6. This extremely large uncertainty in the required detection sensitivity strongly suggests that there is not ideal sensor sensitivity for C/B agents.

Recommendations:

- The sensor requirements must be determined from operational scenarios.
- Sensors must be developed with different performance characteristics for different missions, i.e., one 'size' sensor does not fit all applications.
- In order to cover a range of operational scenarios, a suite of C/B sensors must be available.

3. Sensor should allow for multiple operating modes.

Recommendations:

- A sensor "knob" is needed to switch from low false positive rate (lower sensitivity) to high sensitivity (higher false positive rate) depending on perceived threat level.
- Sensors should be designed such that dual thresholds are possible for high and low disruption actions (alarm and trigger respectively).

4. C/B sensor testing and characterization is outdated and inadequate.

Recommendations:

- Establish oversight authority for improved testing methodologies and equipment.
- Provide a detailed understanding of threat biological-agent characteristics.

5. Sensor requirements are not well defined.

Recommendations:

- Need to develop "realistic" scenarios for modeling sensor requirements.
- Need to do empirical testing on these "realistic" scenarios.
- Need to characterize "real" environments.

While this study has produced significant findings, the Panel recognizes that there are a number of other aspects of sensor evaluation that should be considered. Specifically the following subjects should be studied:

- Study the effect of networking C/B sensors as an option to reduce the false positive rate and increase detection confidence.
- Study the feasibility of simulation tools to model preliminary design and performance of C/B sensors.
- Investigate how ROC curves should be compared with one another (e.g. area under the curve/accuracy).
- Perform an error analysis on the threat scenarios to determine the impact of assumptions on outcomes.

Appendix A: Vulnerability based assessment of biological sensor sensitivity requirements

The purpose of a C/B agent sensor is to prevent or to reduce exposure to an agent. Inhalation exposure is measured as a dosage, which is the time-integrated accumulation of agent in the human lung. Dosage has units of mass for chemical agents or biological toxins and units of numbers for biological agents.

A major problem that must be confronted when considering the evaluation of C/B agent sensors is the fact that most of these sensors do not actually measure ***dosage*** but rather concentration. The exact relationship between agent dosage and detection concentration is poorly defined. In the case of biological agent detection, an estimate of the required detection concentration for a given dose is based on a range of many variables, including temporal profile of the agent, particle size distribution, particle fill factor, agent density, breathing rate, lung retention efficiency, and exposure time. Considering the variation in each of these parameters the variation in the required detection concentration spans a range of greater than 10^6 particles per liter (see Table A-1 and detailed calculation below). This extremely large dynamic range and uncertainty in the required detection sensitivity strongly suggests that there is not easily defined ideal sensor sensitivity for chemical and biological agents.

Agent Dosage

The product of the agent concentration, the human breathing rate, the lung retention efficiency, and the exposure time gives the accumulated inhalation dosage of agent.

$$D = C'B\eta\, T \qquad \textbf{EQUATION 1}$$

where;

D is the dosage in units of CFUs or PFUs for bacterial and viral agents and µg for bio-toxins or chemical agents, C' is the agent concentration in "cells" per liter of air for biological agents or µg per liter of air for chemical agents or bio-toxins, B is the human breathing rate in liters per minute, η is the lung retention efficiency, and T is the exposure time.

For aerosolized agents, it is useful to express the agent concentration C' ("cells"/liter, or µg/liter) in terms of a particle concentration C (particles/liter)[3].

$$C' = V\rho f C \qquad \textbf{EQUATION 2}$$

where;

V is the aerosol particle volume (µm^3), ρ is the agent density per unit volume ("cells"/µm^3 or µg/µm^3), f is the agent particle fill factor (fraction of particle containing agent), and C is the aerosol particle concentration (particles per liter of air).

Substituting equation 2 into equation 1 and expressing the particle volume in terms of the particle radius the following is obtained.

Footnote

[3] Most detectors of aerosolized agents measure the aerosol particle concentration (number of particles per liter of air).

$$D = \frac{4\pi}{3} r^3 \rho f C B \eta T$$ **EQUATION 3**

where;

r is the aerosol particle radius.

For an aerosol with a distribution of particle sizes and a temporally dependent particle concentration then the dosage is given below.

$$D = \int_{t=-\infty}^{\infty} \int_{r=r_0}^{\infty} \frac{4\pi}{3} r^3 \rho f C(r,t) B \eta(r)\, dr\, dt$$ **EQUATION 4**

where;

r_o is the smallest possible agent particle diameter, $C(r,t)$ is the aerosol concentration as a function of r and t, $\eta(r)$ is the agent retention efficiency of the lung as a function of particle size.

Agent Concentration

For a given agent dosage, the required detection concentration is not well defined. There are many ways to generate the same dosage with different combinations of concentration, agent temporal profile, agent particle size distribution, etc. Ignoring the time variation of the agent concentration and the distribution of agent particle size (equation 4), the required sensor concentration sensitivity for a given agent dosage can be obtained by solving equation 3 for C.

$$C = \frac{D}{\frac{4\pi}{3} r^3 \rho f B \eta T}$$ **EQUATION 5**

If, for example, one desires to prevent a biological agent dosage which would result in the death of 50% of an untreated population (LD50) and if the variables in equation 5 are assumed to have the estimated values shown in Table 1 then the biological agent sensor must detect an agent concentration of 63 particles per liter or better. Of course, this calculation is based on estimates and it is useful to consider the possible errors in these estimates and the effects on the sensor requirements. A quick way to measure the uncertainty in the required detection concentration is to use the extreme values for each variable in equation 5. This approach is not strictly correct since it assumes that the variable variations are positively correlated. Except for possibly the breathing rate, lung retention efficiency, and particle diameter the variables are not correlated. Nonetheless, if the uncertainty in the required detection concentration based on the range of variation shown in Table 1 is estimated, a required detection concentration, which varies from 0.3 particles per liter to 105 particles per liter, is obtained.

***Table 1.** Variables, estimated values and estimated value ranges for calculating agent concentration from agent dosage*

Variable	Estimated Value	Estimated Value Range
Dosage (D) for LD$_{50}$	10,000 cells	20,000-5,000 cells
Particle radius (r)	1.5 μm	0.4-2.5 μm
Agent density (ρ)	1 cell/(0.8 μm)^3 =0.5 cells/μm^3	0.5-1.0 cells/μm^3
Fill factor (f)	0.6	0.3-0.7
Breathing rate (B)	15 liters/min	10-20 liters/min
Lung retention efficiency (η)	0.5	0.3-0.7
Concentration (C)	63 ppl	10^5-0.3 ppl

Agent temporal profiles

The required sensitivity of a chemical or biological agent concentration sensor is dependent on the temporal profile of the agent vapor or aerosol concentration. This is true even when the dose and agent duration (FWHM) are fixed. Figure A-1 shows the agent concentration at a particular location as a function of time for three different agent temporal profiles: step, Gaussian, and exponential. Each of these profiles has the same integrated dose ($\int_{-\infty}^{\infty} Cdt = 1$) and time duration (full width half maximum = 1). The functional forms of these profiles are given in equations 6.

$$C_{step} = 1 : for\ -0.5 \le t \le 0.5,\ C_{step} = 0 \text{ otherwise}$$

EQUATION 6

$$C_{Gaussian} = 2\sqrt{\frac{Ln(2)}{\pi}}\, e^{-(2\,Ln(2)t)^2} : for -\infty \le t \le \infty$$

$$C_{\exp onential} = Ln(2)\, e^{-Ln(2)(t+0.5)} : for\ t \ge -0.5,\ C_{\exp onential} = 0 \text{ otherwise}$$

For equal dosage and time duration agent releases, an instantaneous concentration sensor must be 6% more sensitive to detect the Gaussian profile agent release as compared to the step profile agent release and 31% more sensitive to detect the exponential profile as compared to the step profile. For example, an instantaneous concentration sensor, that is capable of detecting a step release of 25 particles per liter, must detect 24 particles per liter for a Gaussian release and 17 particles per liter for an exponential release. If the sensor does not respond to the instantaneous concentration but instead responds to a short term average then the difference between required sensitivities increases. For example, if the sensor averages the concentration over one time unit (equal to the agent time duration FWHM), then the sensor must be 24% more sensitive for a Gaussian release as compared to a step release and 50% more sensitive for an exponential release as compared to a step release. In this case, a concentration sensor that is capable of detecting a step release of 25 particles per liter must detect 19 particles per liter for a Gaussian release and 12.5 particles per liter for an exponential release.

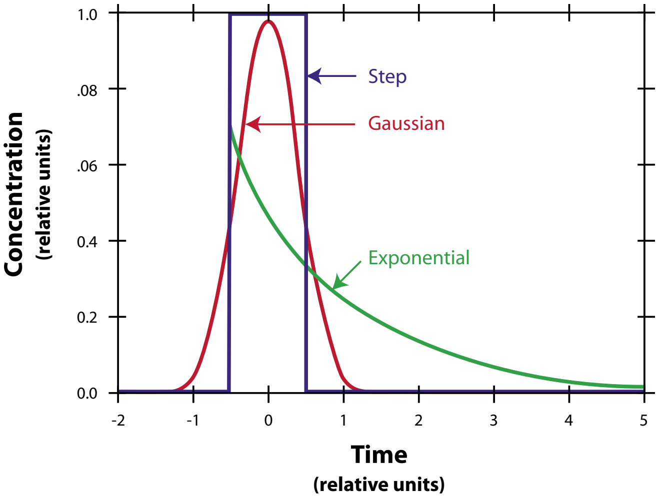
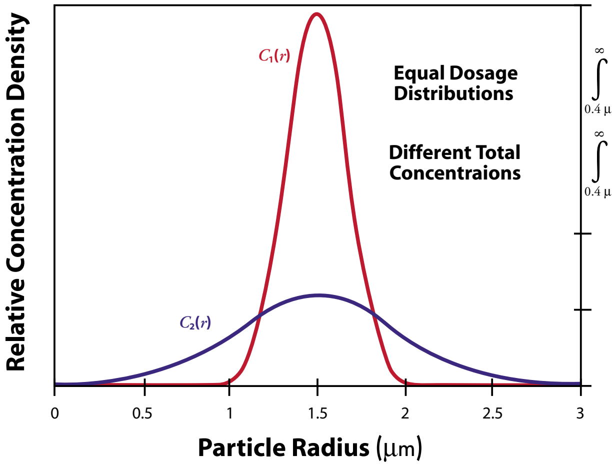


Figure A-1. *Agent concentration temporal profiles with equal dosage but unequal peak or short time averaged concentrations.*

Agent particle size distribution

As indicated in equation 4, the agent dosage is dependent on the particle size distribution for aerosol threats. This in turn means that the required sensitivity of a chemical or biological agent aerosol concentration sensor is also dependent on the particle size distribution of the agent aerosol. Figure A-2 shows two agent particle size distributions each with the same number mean particle size, but with different distribution widths. These distributions also have the same dosage ($\int_{0.4\,\mu m}^{\infty} r^3 C_1(r)\,dr = \int_{0.4\,\mu m}^{\infty} r^3 C_2(r)\,dr$) but they have different total concentrations ($\int_{0.4\,\mu m}^{\infty} C_2(r)\,dr = 0.76 \int_{0.4\,\mu m}^{\infty} C_1(r)\,dr$). Thus to generate equal warnings for each of these agent particle size distributions a sensor must be 24% more sensitive for distribution $C_2(r)$ as compared with distribution $C_1(r)$.



$$\int_{0.4\,\mu m}^{\infty} r^3 C_1(r)\,dr = \int_{0.4\,\mu m}^{\infty} r^3 C_2(r)\,dr$$

$$\int_{0.4\,\mu m}^{\infty} C_2(r)\,dr = 0.76 \int_{0.4\,\mu m}^{\infty} C_1(r)\,dr$$

Figure A-2. *Agent particle size distributions corresponding to equal dosage but unequal total concentration.*

This vulnerability-based analysis indicates that sensor requirements are very uncertain. The impact of this uncertainty can be mitigated by the evaluation of sensors over a wide range of sensitivities and false positive rates. Such an evaluation (and the construction of ROC curves) would enable the operation of sensors at different sensitivities and different false positive rates depending on the perceived threat levels, false alert rate tolerance, and mission objectives.

The uncertainty in requirements also suggests that, depending on false positive rates, sensor size, sensor cost, and other sensor attributes, different sensors with widely different sensitivity capabilities may have useful roles in the defense against chemical and biological agents.

Appendix B: Detection Theory

The air and surfaces of any real location contain a huge array of trace biological molecules, chemicals and particulates resulting from ecological processes, animal activities, human activities, weather, fires and a myriad other sources. At any given time, there are considerable amounts of materials present in the air, including minerals, pollen, dust, mold, and bacteria. Over one thousand chemical species have been found in diesel exhaust, and their concentrations vary significantly from engine to engine, with fuel/air/engine temperatures, fuel types and many other factors. The ambient make-up of air at a given location is constantly changing due to wind, weather events, reactions among chemicals, reactions of chemicals to light, condensation, evaporation, precipitation, lofting, changes in lighting conditions and the chemical sources, etc. Every air sample contains a different and vastly complicated mix of biologics, chemicals and particulates. Together, the physical location and the distribution of natural conditions and activities at that location define a measurement environment.

A sensor's response to a given air sample consists of three major components. Noise is fluctuation in sensor responses due to factors that are independent of the measurement environment. Clutter is the sensor response to all factors associated with the measurement environment other than the agent. Signal is the sensor's response to the agent. Due to the variations in the measurement environment, the clutter component of the sensor's response to each sample is different, as is the noise component of its response. While the background in a given sample is generally far too complex to measure or model, sensor responses to samples are easily measured and can be assembled into distributions that can be used to characterize sensor performance in that measurement environment. The distribution of sensor responses for a given measurement environment is consistent, to the extent that the likely range of weather and activity conditions has been sampled and no major new activities or events take place close enough to affect the ambient materials found there. The distributions of sensor responses to samples from a given environment can be plotted as histograms. Figure B1 contains a background histogram plotted from sensor responses to background samples that do not contain the agents of interest, and an agent histogram plotted from agent samples from the same environment, to each of which we have added a known quantity of the agent. The agent histogram has a distribution of responses because it contains noise and clutter in addition to signal. However, the signal changes the shape and location of the agent sample histogram, as illustrated by the hypothetical agent histogram for an agent concentration C1.

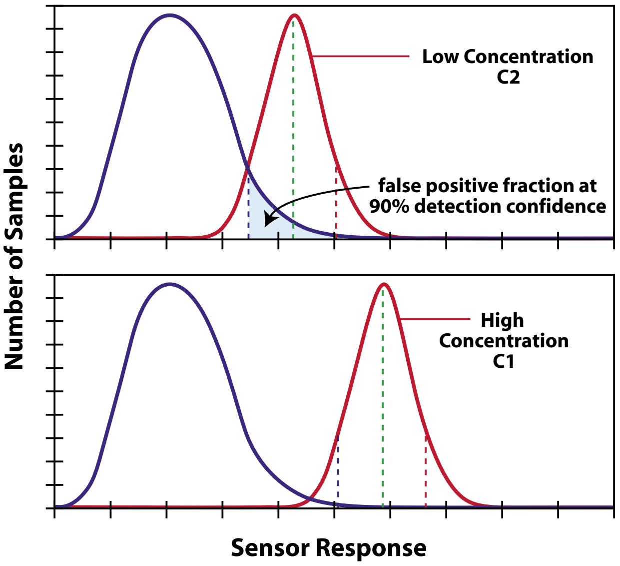


Figure B1. *Sensor background and agent signals showing the 10%, 50% and 90% detection probability thresholds for two different magnitude agent challenges, C2 and C1.*

By analyzing the agent histogram, we can determine a ***threshold*** sensor response, above which a desired percentage of the sensor responses to agent samples (the ***detection confidence***) lie. For example, 90 percent of the sensor responses to positive samples lie above the threshold labeled T90, so we say that the detection threshold for 90 percent detection confidence to an agent concentration of C1 is T_{90} (lower plot in figure B1). However, a fraction of the sensor responses to background samples are also above T90. This fraction is the ***false positive fraction*** for detecting C1 with 90 percent confidence at this location. From these two distributions, then, we have just determined a point (agent concentration C1, false positive fraction FPF_{90}) on this sensor's ROC curve for 90 percent detection confidence in this environment. Using the same two histograms, we can determine points on the sensor's ROC curves for other detection confidences by choosing different thresholds (T_{xx}). Note that one can decrease the false positive fraction by accepting a decrease in the confidence with which a concentration of C1 can be detected.

To determine the complete ROC curves for a given measurement environment, agent histograms are measured for additional concentrations until the desired range of agent concentrations has been explored. The upper plot in figure B1 illustrates the background histogram with the agent histogram for a second, lower concentration C2. Note that because C2 is lower, the agent histogram has moved closer to the background histogram. A lower detection threshold T_{90} is therefore required to achieve 90% detection confidence for a concentration of C2, which results in a larger false positive fraction.

Note, again, that these curves are unique to a given environment because the clutter components of the sensor responses are unique to the measurement environments. Measurement and analysis of the background histogram and the desired range of agent histograms must therefore be repeated in each of the desired measurement environments to determine the ROC curves in that environment.

The methodology for evaluating sensors outlined above implies the need for a tremendous number of careful measurements in a wide range of locations under operational conditions, and working with many samples of agents in those environments. In practice, the number of measurements that can be made will be severely limited by practical constraints on cost and time. Safety considerations prohibit working with agents outside test chambers at designated locations, or working with any purposely-released chemicals in most operating environments. Use of simulants is of dubious value for most types of sensors because there generally isn't a simulant that duplicates the relevant physical characteristics of the agent, as well as sensor response to the agent. Most simulants are also hazardous to at least some extent.

In view of these considerations and the complexity of the environments, the best practical method of evaluating sensor performance is probably to measure background histograms in real environments, measure signals (the component of sensor response due to agent) and noise in test chambers, and use numerical analysis to synthesize the agent histograms needed to determine ROC curves. The rationale for this method of sensor evaluation is the following.

There isn't any effective alternative to measuring at least background histograms in real environments because the factors that affect sensors are too complex, highly variable, and hard to measure. The complexity and variability make it difficult (in most cases) to predict what is actually interfering with the sensor. Simulating real environments by adding interferents or challenges to sealed testing chambers where agents might be used is generally ineffective because it is difficult to duplicate the complexity and to select the relevant interferents as they are often undefined. It is generally even more difficult to numerically model real environments from first principles. Simulating and modeling complex environments is also very expensive. While it might be possible to take air samples in a real measurement environment and bring them to a sealed test chamber, the content of such samples usually changes over time.

To control the expense of measuring background histograms in real environments, it is important to match levels of effort to the development/deployment status of the sensor. In early stages of development, for example, background histogram measurements might be made in just a few environments that are broadly representative of envisioned operational use, and selected to contain any specific environmental factors that are likely to cause problems for the sensor. As development of a given sensor progresses and the resources devoted to it increase, the range of environments can be steadily increased to provide more complete performance assessments. After the sensor is deployed, it would be highly beneficial to record background measurements during training exercises and force deployments. This would enable assessment of sensor performance in the broadest possible range of environments, and provide much of the information needed to improve performance by modifying the sensors, tactics for using them, or training.

In general, sensor signals and noise are relatively easy to measure in the laboratory, and physics-based numerical models that have been vetted with careful experiments are quite accurate. If the sensor system (to include both the instrument and the data analysis algorithm) response to noise, clutter, and signal, collectively, is linear (meaning that these three components of the response can simply be added, with appropriate weighting, to determine the total response), numerical synthesis of the agent distributions is straightforward: a fixed signal corresponding to a given agent concentration is simply added to each background sample measurement. If the physics and chemistry of a sensor are understood, the functional form (linear or otherwise) of instrument-level (raw signal) response is known. However, it can be much more difficult to determine whether the data analysis algorithm used to convert raw signals to sensor responses is linear. The safest numerical approach is therefore to combine noise, clutter, and signal in the correct way at the raw signal level, and then apply the specific data analysis algorithms to be used in the sensor to generate the synthetic histograms of sensor response to positive samples.

Appendix C: Scenarios

As previously discussed, there are many different possibilities and scenarios and there is no way that any one sensor could be evaluated against all of the possible mission scenarios, environments and threats. In order to bound the discussion and provide metrics for the purposes of this report, the focus is on four scenarios that are applicable to ongoing missions within DoD. The scenarios are as follows: Movement to Contact, Deliberate Defense, External Attack on Fixed Facility, Internal Attack on Fixed Facility. These scenarios are not meant to represent the entire spectrum of the challenge, they are not static, instead, they will evolve as the mission changes, the threat is better understood and as they are adopted for other purposes.

In these scenarios, the goal is to substantially reduce exposure to the Warfighter as such the focus is on detect-to-warn systems, which by definition provide ample warning to personnel before encountering a C/B threat such that countermeasures can be taken to avoid exposure. In principle, a detect-to-warn system should prevent exposure and the need for treatment. The alarm or trigger of a detect-to-warn sensor can result in either a high or low disruption action. A high disruption action has consequences that result in a significant interruption in the normal or planned activities. For instance, a high disruption action would be a significant change in the tactics and procedures impacting operations at the force level. Similarly, in consideration of a fixed facility, a high disruption action would shutdown normal operations at a Command Center resulting in evacuation or shelter-in-place. A low disruption action has consequences that result in no significant interruption in normal or planned activities. An example of a low disruption action in the field would present minor interruptions at a local level, i.e., initiating reconnaissance and donning of protective posture. Additional sample acquisition and testing or alterations in airflow without a perceived difference by personnel are representative of low disruption actions.

Force Protection Scenarios

Consider the battlefield force protection scenario in which the enemy quickly releases a quantity of agent near ground level, relying on favorable meteorological conditions to transport the agent onto military forces. In this outdoor case, the enemy is assumed to release 1 kg of anthrax or a 55-gallon drum (225 kg) of Sarin.

Movement to Contact

In a Movement to Contact, the enemy is attempting to disrupt transportation along the Main Supply Route (MSR) by releasing agent alongside the MSR and relying on winds to carry the agent across the MSR. Sensors are placed along the outskirts of the protected MSR corridor a distance of 500 m apart, as shown in Figure C-1. They could be deployed quickly by reconnaissance forces, similar to how a highway crew puts down cones before road construction. As a consequence, these sensors must be rugged, battery powered and have wireless communications. The ability to network these sensors can be used to reduce overall false positive rate. They would also possess the ability to geolocate and when networked to provide directional information as to the likely release location. After receiving a detection warning, personnel in convoys on the MSR would be directed to don masks and proceed to the decontamination point, where forces, vehicles and cargo would be decontaminated with chlorinated water. By acting on the direction information, attack helicopters can be rapidly directed to locate and possibly retaliate against the enemy.

Deliberate Defense

In a Deliberate Defense situation, we are protecting forces in defensive position behind Forward Edge of Battle Area (FEBA). Assume the same release masses as in the Movement to Contact scenario, except here the distance between release, sensor, and protected forces is smaller. Sensors are located at observation posts 500 m forward of the FEBA, as shown in Figure C-2. Since the release is closer, the warning time is shorter and the potential exposure to forces is higher. To account for this event, the sensors must be faster and more sensitive. In response to a trigger, forces behind the FEBA are directed to don masks. Because the sensors are within a secured area, they also may be more readily maintained and consume more power. Because of the fixed positions, a standoff sensor could also be used to monitor the upwind direction. However, a point sensor with 30-second response time provides adequate warning to forces behind the FEBA.

Fixed Facility Protection

The defense of fixed facilities is critical to the overall success of DoD missions. Most buildings can support full-time particulate filtration at moderate levels, which is helpful against biological attack. However, the conversion to High Efficiency Particulate Air (HEPA)/activated carbon filtration (>99.99 percent removal for C/B agents) involves major costly renovations and has substantial operations and maintenance costs due to added energy usage and filter maintenance. Since most modern buildings have electronically controlled heating, ventilation, and air-conditioning (HVAC) systems, the addition of C/B agent sensors makes them good candidates for detect-to-warn defense strategies using active HVAC air routing techniques. Since the sensors are at or within the building, there is little or no "upwind" warning time as is described in the Force Protection scenarios. As a result, the ability to reduce exposure is based mostly on preventing contaminated air from reaching occupants.

In an External and Internal Fixed Facility Attack Scenarios, consider a military administration center operating in a small-sized office building. We assume moderate filtration (e.g. 99 percent removal of respirable sized particles, but no chemical vapor filtration) and a typical level of overpressure (e.g. 0.04 inches of water gauge). Personnel in the building are not expected to don masks.

External Attack on Fixed Facility

In an External Attack on a Fixed Facility the enemy releases agent from 1 km upwind of the building, as shown in Figure C-3. A single biological and chemical point sensor is situated on the rooftop near the fresh air intakes. When the plume impacts the building, the sensor begins to detect contaminant and after a 10 second response time, issues a trigger signal for the HVAC system to shut down. It is assumed that this shutdown is completed in an additional 10 seconds. Much of the contamination enters the building within the first minute. If the HVAC shutdown was accidentally delayed until after the cloud has passed, the triggered action can actually increase the exposure to occupants by trapping contaminant inside. This action must be avoided.

Once the HVAC is shut down, contaminated air is no longer actively drawn into the building. The shutdown also causes the building overpressure to drop and allows leakage through exterior openings, particularly through the upwind facing walls. Once the cloud has passed, the sensor must also be capable of testing for residual external contamination.

Internal Attack on Fixed Facility

In an Internal Attack on a Fixed Facility a person inside the building releases a small amount of agent. As shown in Figure C-4, the release takes place in Zone A of the building and due to the HVAC system the agent will spread quickly throughout the building, contaminating the building occupants.

To support a detect-to-warn defense, sensors are placed in the return air ducts throughout the building. Sensors trigger an active HVAC response to rebalance the building. The HVAC system first isolates the release area from the rest of the building by overpressure, and then begins a full building flush with fresh outside air. Assume sensors have a 30 second response time and the HVAC system can be rebalanced in an additional 60 seconds. This operating mode would continue until an identification of the contaminant can be completed.

A summary of requirements for sensor performance is shown in Table C-1. In addition to these requirements there are several other sensor attributes that strongly affect the sensors utility for various missions, these attributes are shown in Table C-2. It is important to remember that the requirements of sensors are very scenario dependent.

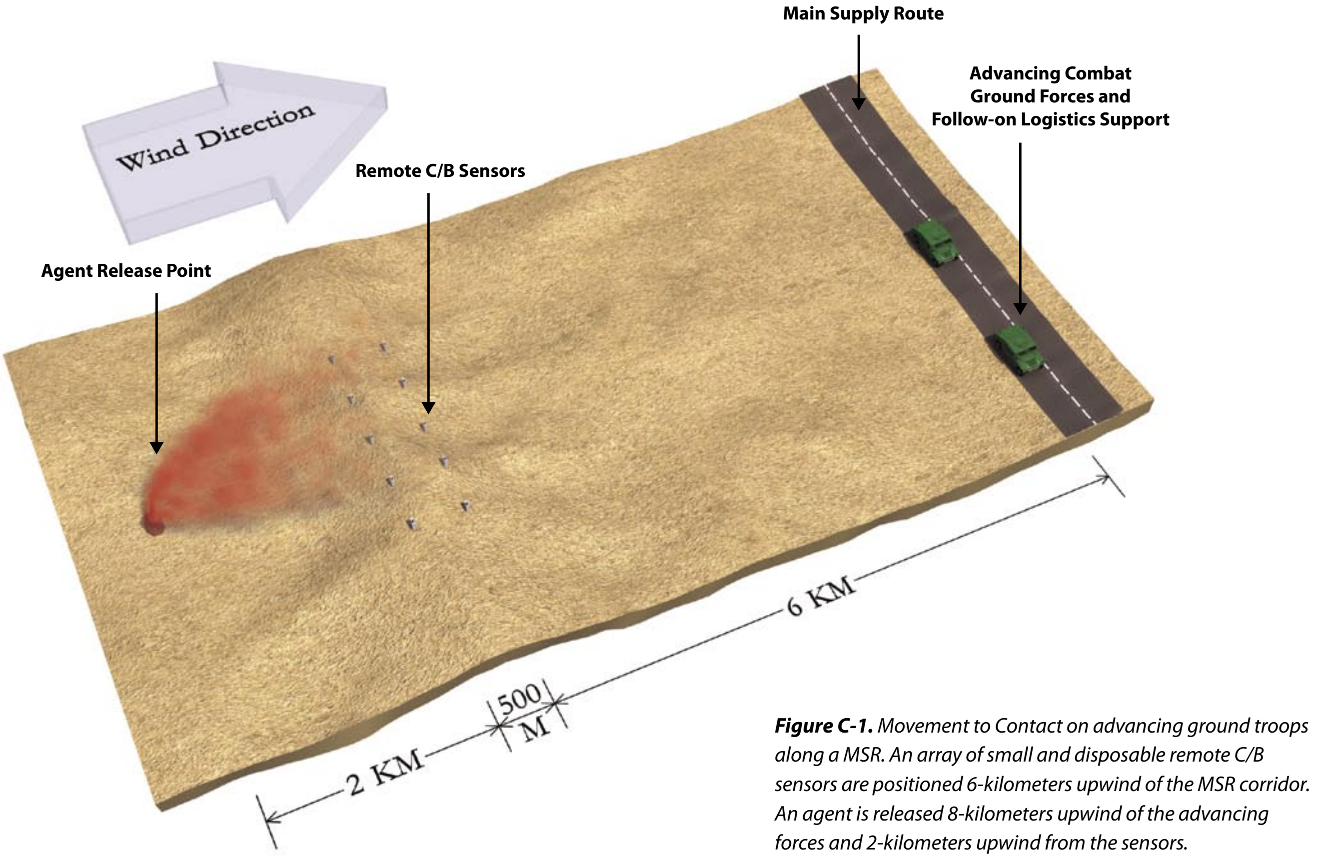


Figure C-1. *Movement to Contact on advancing ground troops along a MSR. An array of small and disposable remote C/B sensors are positioned 6-kilometers upwind of the MSR corridor. An agent is released 8-kilometers upwind of the advancing forces and 2-kilometers upwind from the sensors.*

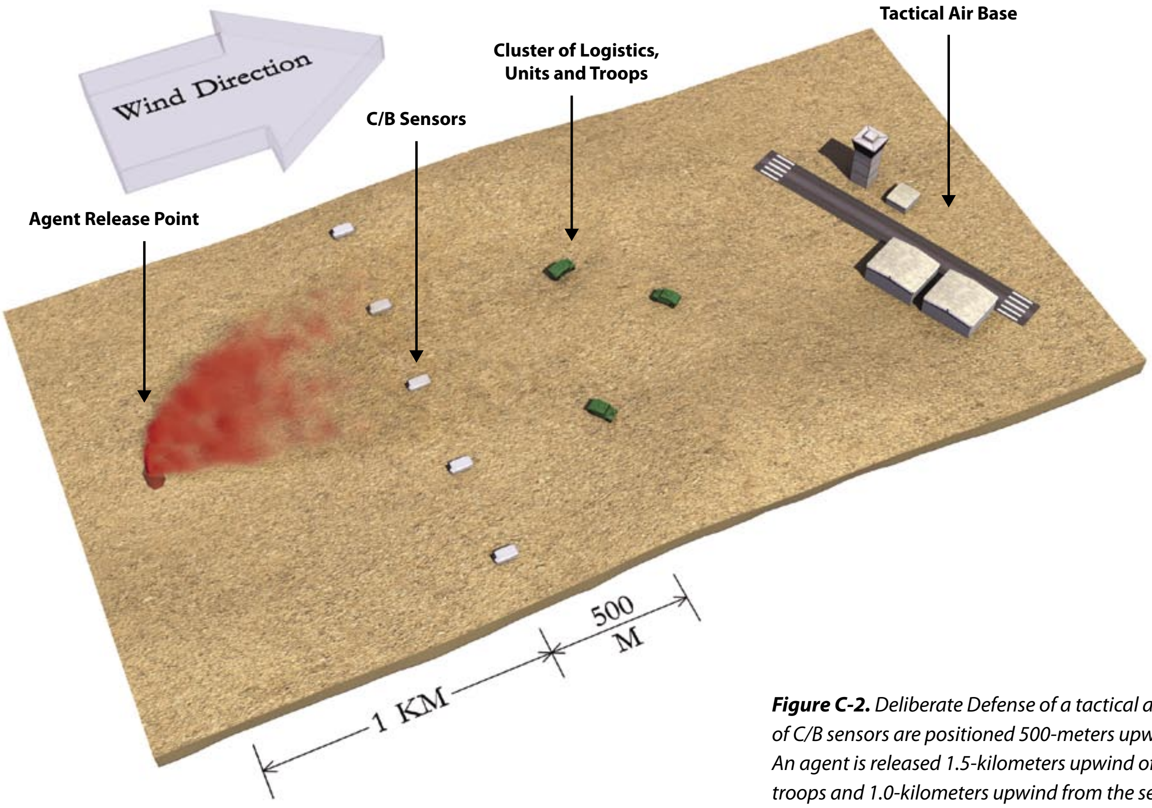


Figure C-2. *Deliberate Defense of a tactical air base. An array of C/B sensors are positioned 500-meters upwind of the FEBA. An agent is released 1.5-kilometers upwind of logistics units and troops and 1.0-kilometers upwind from the sensors.*

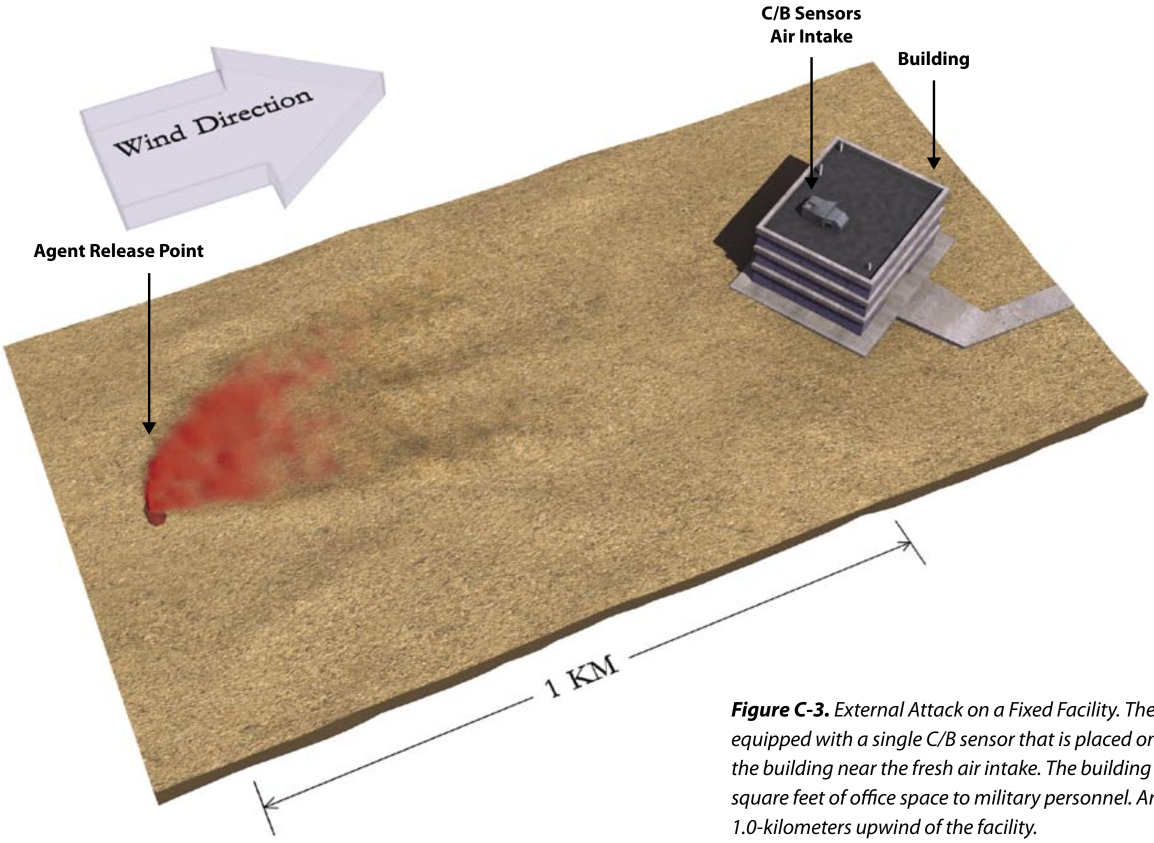


Figure C-3. *External Attack on a Fixed Facility. The facility is equipped with a single C/B sensor that is placed on the rooftop of the building near the fresh air intake. The building provides 200,000 square feet of office space to military personnel. An agent is released 1.0-kilometers upwind of the facility.*

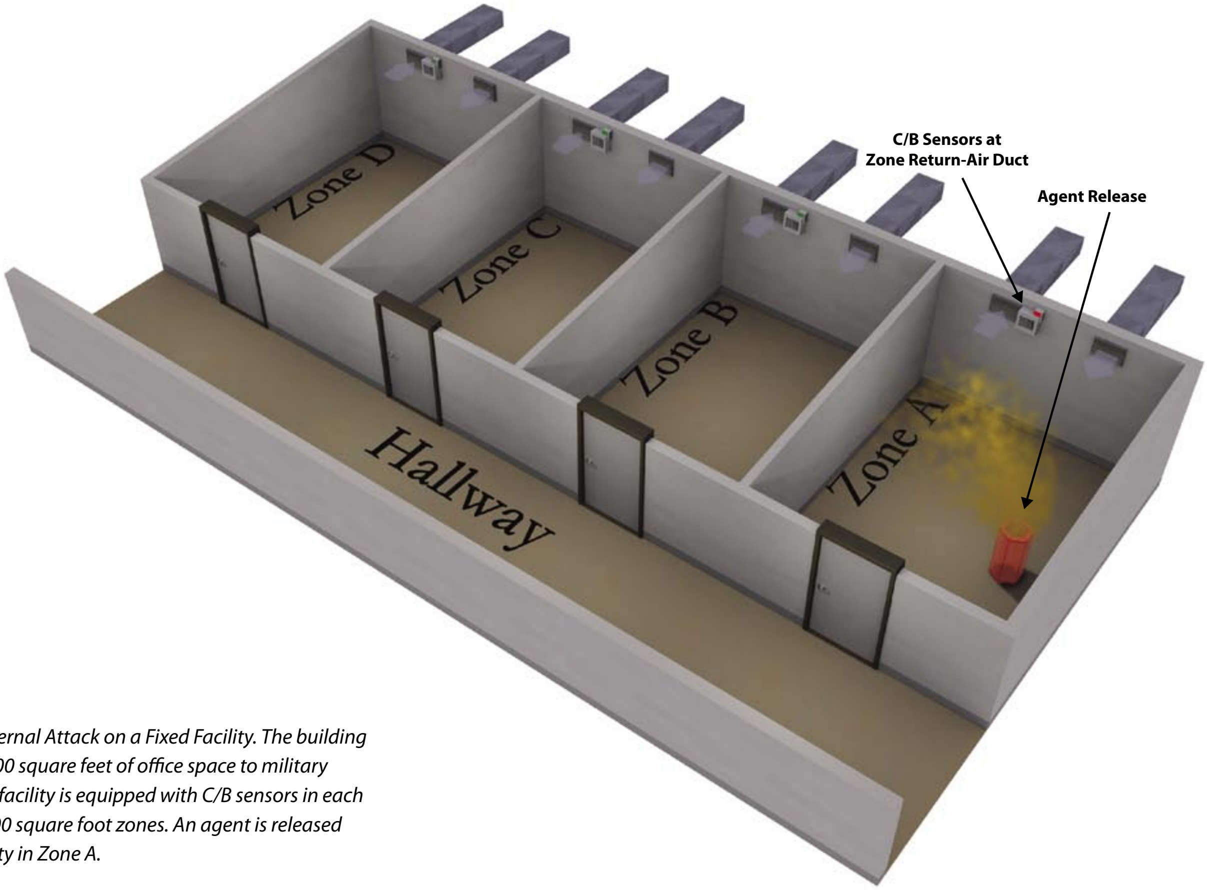


Figure C-4. *Internal Attack on a Fixed Facility. The building provides 200,000 square feet of office space to military personnel. The facility is equipped with C/B sensors in each of its four 50,000 square foot zones. An agent is released inside the facility in Zone A.*

Table C-1. *A Threat Analysis of Sensor Performance and Protective Benefit.*

	Scenario 1 Battlefield (Movement to Contact)	Scenario 2 Battlefield (Defense)	Scenario 3 Building (External Attack)	Scenario 4 Building (Internal Attack)
Sensor Sensitivity				
• *B. anthracis* (ppl)	1000	100	50	1000
• Sarin (mg/m^3)	50	5	2	100
Response Time (sec)	60	30	10	30
Detection Confidence Pd (%)	98	98	98	98
False Positive Rate				
• Low disruption	1/week	1/month	1/week	1/month
• High disruption	1/month	2/year	2/year	2/year
Target Dose				
• No bio-defense	2 x LCt50	800 x LCt50	600 x LCt50	850 x LCt50
• No chem-defense	<0.05 x LCt50	1.4 x LCt50	1x LCt50	6.4xLCt50
Exposure Reduction with Defense	>1000 x	>1000 x	15 x	850 x

Table C-2. *A Threat Analysis of Other Sensor Attributes.*

	Scenario 1 Battlefield (Movement to Contact)	Scenario 2 Battlefield (Defense)	Scenario 3 Building (External Attack)	Scenario 4 Building (Internal Attack)
Unit Production Cost	$100	$1,000-$10,000	$10,000-$50,000	$2,000-$4,000
Power Source	<1 Watt (Battery)	<10 Watt (Battery)	<5 Kilowatt (AC line)	<50 Watt (AC line)
Mission Duration	Days-Weeks	Weeks-Months	Years	Years
Maintenance (MTBM)	None	Week	3 Months	3 Months
Reliability (MTBF)	>Weeks	>Quarter	1 Year	1 Year
Size and Weight	Soda Can	Briefcase	Residential AC Unit	Breadbox
Sensor Density	4/km	1/forward unit	1/building	1/building zone
Environent (example clutter)	Battlefield Outdoors (smoke, diesel, dirt)	Battlefield Outdoors (smoke, diesel, dirt)	Outdoors (urban, rural, desert)	Indoors (cleaners, dust)

Acknowledgements

The panel thanks Dugway Proving Ground, U.S. Army Edgewood Chemical Biological Center, Chamber Corporation and the Office of the Secretary of the Navy for their presentations.

The panel especially thanks the Booz Allen Hamilton team for contributing to Panel discussions and for their efforts to help write this report.