NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

OPTIMIZING SYSTEMS OF THRESHOLD DETECTION SENSORS

by

David C. Banschbach

March 2008

Thesis Advisor:	Ronald D. Fricker, Jr.
Second Reader:	W. Matthew Carlyle

Approved for public release; distribution is unlimited

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington DC 20503.			
1. AGENCY USE ONLY ***(Leave blank)***	**2. REPORT DATE** March 2008	**3. REPORT TYPE AND DATES COVERED** Master's Thesis	
4. TITLE AND SUBTITLE Optimizing Systems of Threshold Detection Sensors		**5. FUNDING NUMBERS**	
6. AUTHOR(S) David C. Banschbach			
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943-5000		**8. PERFORMING ORGANIZATION REPORT NUMBER**	
9. SPONSORING /MONITORING AGENCY NAME(S) AND ADDRESS(ES) N/A		**10. SPONSORING/MONITORING AGENCY REPORT NUMBER**	
11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.			
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution is unlimited		**12b. DISTRIBUTION CODE** A	

13. ABSTRACT (maximum 200 words)

When implementing a system of sensors, one of the biggest challenges is to establish a threshold at which a signal is generated. All signals that exceed this detection threshold are then investigated to determine whether the signal was due to an "event of interest," or whether the signal is due simply to noise. Below the threshold all signals are ignored.

We develop a mathematical model for setting individual sensor thresholds to obtain optimal probability of detecting a significant event, given a limit on the total number of false positives allowed in any given time period.

A large number of false signals can consume an excessive amount of resources and could undermine confidence in the system's credibility. One motivation for this problem is that it allows decision makers to explicitly optimize system detection performance while ensuring it meets organizational resource constraints.

Our simulations demonstrate the methodology's performance for various sizes of sensor networks, from ten up to thousands of sensors. Such systems apply to a wide variety of homeland security and national defense problems, from biosurveillance to more classical military sensor applications.

14. SUBJECT TERMS Threshold Detection, Shewhart Chart, Statistical Process Control, Epidemiologic Surveillance, Syndromic Surveillance.			**15. NUMBER OF PAGES** 86
			16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT Unclassified	**18. SECURITY CLASSIFICATION OF THIS PAGE** Unclassified	**19. SECURITY CLASSIFICATION OF ABSTRACT** Unclassified	**20. LIMITATION OF ABSTRACT** UU

NSN 7540-01-280-5500 Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release; distribution is unlimited

OPTIMIZING SYSTEMS OF THRESHOLD DETECTION SENSORS

David C. Banschbach
Lieutenant, United States Navy
B.A., Seattle University, 2001

Submitted in partial fulfillment of the
requirements for the degree of

MASTER'S OF SCIENCE IN OPERATIONS RESEARCH

from the

NAVAL POSTGRADUATE SCHOOL
March 2008

Author: David C. Banschbach

Approved by: Ronald D. Fricker, Jr.
Thesis Advisor

W. Matthew Carlyle
Second Reader

James N. Eagle
Chairman, Department of Operations Research

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

When implementing a system of sensors, one of the biggest challenges is to establish a threshold at which a signal is generated. All signals that exceed this detection threshold are then investigated to determine whether the signal was due to an "event of interest," or whether the signal is due simply to noise. Below the threshold all signals are ignored.

We develop a mathematical model for setting individual sensor thresholds to obtain optimal probability of detecting a significant event, given a limit on the total number of false positives allowed in any given time period.

A large number of false signals can consume an excessive amount of resources and could undermine confidence in the system's credibility. One motivation for this problem is that it allows decision makers to explicitly optimize system detection performance while ensuring it meets organizational resource constraints.

Our simulations demonstrate the methodology's performance for various sizes of sensor networks, from ten up to thousands of sensors. Such systems apply to a wide variety of homeland security and national defense problems, from biosurveillance to more classical military sensor applications.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

LIST OF FIGURES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

EXECUTIVE SUMMARY

Statistical process control (SPC) has been used for monitoring and quality control in manufacturing since its pioneer Walter Shewhart first implemented the control chart that bears his name. Throughout the last eighty-plus years SPC has evolved greatly through the development of other monitoring tools such as cumulative sum (CUSUM) and exponentially weighted moving average (EWMA) charts that take past data into account. Importantly, the use of SPC techniques has spread to non-manufacturing disciplines, which recognize the benefits of SPC practices for monitoring processes in various fields.

A constant theme throughout the literature of SPC and detection theory is the balancing of the probabilities of detection and false alarm. Because the impact of each of these events is very disparate, and varies with each application, there is no consensus on exactly how this should be done.

It would be easy to maximize the probability of detection when there are unlimited resources to investigate signals for validity. In reality, however, there are always constraints on resources. As such, we must find a good balance between our chances of detection and false alarms. Thus, we learn that we can set our detection threshold low enough so that we have a good opportunity of detecting a true signal from the underlying noise, without exceeding the constraints on our investigative capability. How to find this point is challenging, however, when we have many sensors working simultaneously to sift through potentially massive amounts of data.

The motivation for this study is the proliferation of biological surveillance systems that have been implemented since the terrorist attacks of September 11, 2001, and the anthrax-laced letters received in government offices in the months that followed. Such systems, which monitor health care data for early indications of man-made or naturally occurring disease epidemics, must constantly balance the inherent trade-off that arises between sensitivity, signal timeliness, and rate of false alarms (specificity), when setting detection thresholds.

Our methodology assigns probabilities of experiencing some “event of interest” to individual sensor coverage areas. In practice these probabilities can be derived from any data that is available, including actual expected probability based on intelligence, or some feature of the area such as population, volume of traffic or commerce, or land or sea area.

Given these prior probabilities of an event of interest, manifesting itself as a shift in the mean of the "no event" distribution, we use mathematical nonlinear programming techniques to determine appropriate individual thresholds to maximize the probability of detecting an event of given magnitude, subject to a constraint on the expected number of false alarms per time period. In fact, this methodology can be applied to any multi-sensor data-monitoring problem where we are interested in assessing if a shift in the underlying distribution has occurred in one or more of the sensors, without exceeding investigative resources.

In the biosurveillance context, it has often been assumed by policymakers, that adding more sensors provides a better chance of detecting a manmade or naturally occurring epidemic. Using this technique for models of various numbers of sensors, we examine the performance implications of policy alternatives. All of this work was done with Microsoft Excel, one of the most widely available analytical tools in use today.

One lesson we demonstrate is that adding marginal sensors to an existing system of sensors, while keeping the expected rate of false alarms constant, lowers the system-wide probability of detection and detracts from the probabilities of detection for each of the existing sensors. The only way to add sensors without diminishing overall detection performance is to allow for more false alarms.

We first apply our method to the two hundred most-populous counties in the United States, using population as a proxy for probability of experiencing a major biological event of interest, which would manifest itself in data as an increased number of patients seeking treatment. We provide an analysis of our method’s false alarm-constrained detection performance, showing that the most-populous counties (or those assigned the largest probability weight, in the general sense) receive the lowest thresholds. We assume that this result is desirable, as those areas with the highest

probability of event will also have the highest probability of detecting it, although with the highest probability of false alarm as well. We then compare the results of similar models using the same population data and resource constraints to evaluate the performance of systems consisting of various numbers of county-based sensors. We find that while holding the acceptable number of false alarms at a constant, individual sensor and system-wide detection performance improves for systems with less sensor coverage, as investigative resources are not spread so thin.

We perform several excursions meant to simulate implementation of various policies. The first is a mandate to shift all detection thresholds down by the same given percentage. This approach might be done in an effort to improve detection performance system-wide. As the thresholds are lowered, the detection performance rises although at a decreasing rate, and the number of false alarms rises at an increasing rate. We find that small percentage decreases in all thresholds (less than five to ten percent) result in modest nonlinear percentage increases in detection performance (again, less than ten percent). However, we also find that caution is warranted in following such an approach because the cost of false alarms can quickly outstrip the benefits of improved detection.

We also explore the possibility of allowing each sensor to vary its threshold according to local policy. This could be due to the availability of (or lack of) local investigative resources. Through a simulation of random perturbations about each sensor's optimal threshold, we find that there is negligible benefit from this policy in terms of detection probability, and a major cost in terms of false alarms. As the amount of allowed variability at each sensor increases, the number of false alarms increases at a rapid rate. We conclude that dispersal of control over detection thresholds is not a good policy unless very tight controls on the bounds of such movement are implemented.

Next, we adapt a tool of health care surveillance implemented several times in the last few years. Drop-in surveillance has been used to monitor populations near major events such as the Super Bowl and major political conventions. Looking at Denver County, Colorado, host of this year's Democratic National Convention, we examine what happens when we force a particular sensor's detection threshold to be lower than its optimal value in order to improve detection performance in a given area. Performing a

sensitivity analysis, we demonstrate that it is possible to dramatically improve detection performance in a single sensor area without significantly degrading overall system performance. We note that although it is possible to tailor constraints to achieve certain collateral goals (those that are separate from the overall goal of maximizing system-wide probability of detection), we must be careful not to over-constrain the optimization system. Doing so will severely degrade the optimal solution and defeats the purpose of using these techniques.

Finally, taking advantage of some properties derived from our statistical assumptions, we derive a single-variable analytical solution that allows us to solve problems of virtually unlimited size. We use this solution to present results from a system of sensors covering all 3,141 counties in the United States (see Figure E1 below). We find that scaling up from 200 to 3,141 counties results in a small decrease in overall probability of detection, lower detection thresholds (with correspondingly higher probabilities of detection) for a small handful of the largest counties, and higher thresholds (resulting in lower probabilities of detection) for the remaining smaller counties. In a system of this size, we realize that the very large number of small counties, in relation to the handful of largest counties, results in well over one thousand sensors with no probability of detection of the given distributional shift. This large-scale model should provide some insight to biosurveillance policymakers as to the reasonable limits of system expansion.

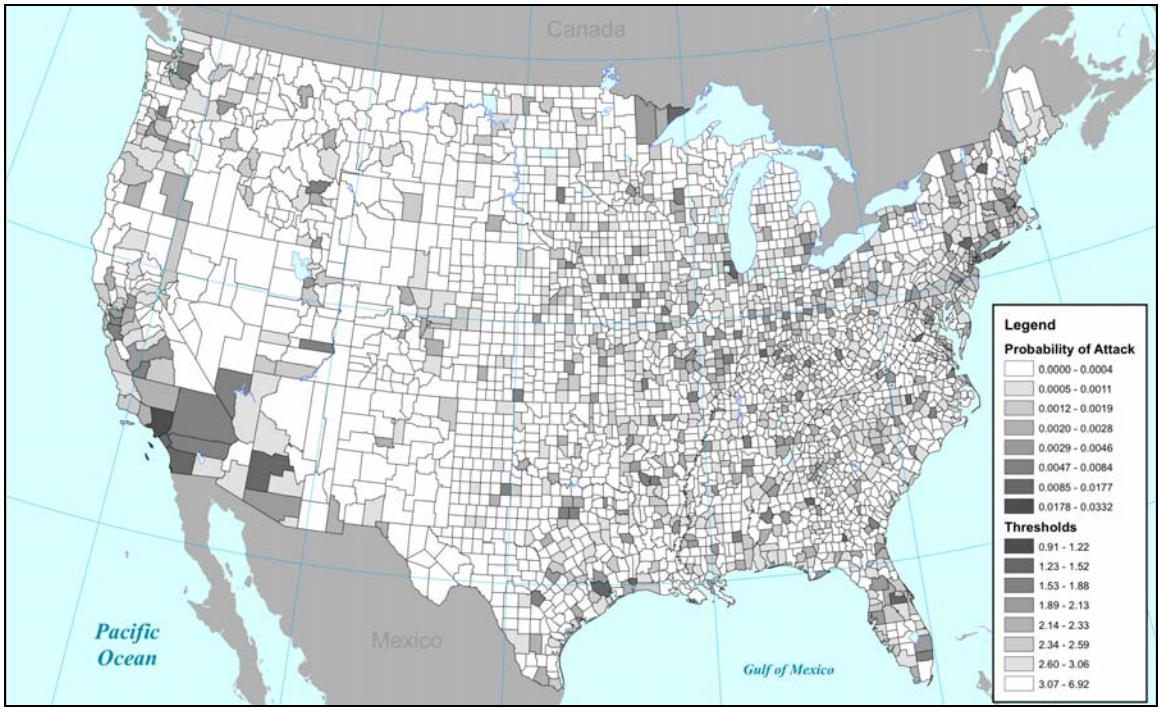

Figure E1: Output from optimizing 3,141 counties

THIS PAGE INTENTIONALLY LEFT BLANK

ACKNOWLEDGMENTS

I would like to thank my wife Catherine, at whose urging I attended the Naval Postgraduate School. I would not have made it through the challenging Operations Research curriculum without her constant support and encouragement.

I would also like to thank Professor Ron Fricker for his guidance throughout this past year while we developed this thesis. Although trained as a statistician, he guided me though the thesis even as it relied more and more on optimization and other operations research tools. But more importantly, he has been a friend who has helped me navigate the academic world here at NPS and given me perspective on the role of the academic in national defense.

THIS PAGE INTENTIONALLY LEFT BLANK

I. SURVEILLANCE AND STATISTICAL PROCESS CONTROL

A. INTRODUCTION

Following the terrorist attacks of September 11, 2001, and the anthrax-laced letters received in government offices in the months that followed, the federal government, along with state and local public health departments nationwide, has installed biological surveillance systems across the country to provide the earliest possible indication that a bio-terror attack is underway. Such systems, which monitor health care data for signs of man-made or naturally occurring disease epidemics, must constantly balance the inherent trade-off that arises between sensitivity, reporting timeliness, and rate of false alarms (specificity), when setting detection thresholds. These detection thresholds are simply the signal intensity or magnitude above which an alarm is generated to notify of an abnormal condition.

The problem of optimizing a system consisting of multiple threshold detection sensors, in terms of setting individual sensor thresholds to maximize the probability of detecting an event of interest, in the form of a shift in the underlying distribution, has not been solved. We turn to the industrial engineering field for tools to help us maximize the probability of detection of an event of a given magnitude, while limiting the false alarm rate to some manageable level. In this thesis, we develop a model using nonlinear mathematical programming techniques to determine appropriate individual thresholds at each location in a distributed set of sensors, taking into account the likelihood of some event of interest in each sensor's coverage area, and accounting explicitly for constraints on resources available for investigation of false positives.

B. STATISTICAL PROCESS CONTROL

With the expanding use of mass machine production in the early twentieth century, techniques were needed to ensure the quality of the items produced without conducting inspections on every single item. In the 1920's, Dr. Walter A. Shewhart of Bell Telephone Laboratories pioneered statistical process control (SPC) through his

development of the control chart that bears his name. In effect, these control charts are sensors that provide a signal when the monitored parameter exceeds a particular threshold value.

In recent years, the use of statistical process control techniques has expanded to other, non-manufacturing applications. MacCarthy and Wasusri (2001) and Montgomery (2001) each cite an extensive list of SPC applications outside the manufacturing industry, from monitoring and evaluating to planning and forecasting. Some areas of study include water quality and chemical monitoring, predictive and preventative maintenance, customer service and satisfaction, trends in patient mortality, crime rates, and food industry hygiene.

In the manufacturing sector, the monitored process has a desired state, for example achieving components with particular measurements. In contrast, many non-manufacturing processes have objectives and measurements that must be defined by the user. The sensors might also be tasked to monitor similar processes in many locations at once. In this case, using the same detection threshold in all locations might not be a good solution because the data in each location may come from distributions with different statistical properties, or have different performance or levels of importance.

In manufacturing, there is always some amount of natural variability that cannot be avoided. Each item that comes off an automated assembly line must be within some tolerance of a target value, even if it is that value is not exactly satisfied. With only these "chance" causes of variation present, a process is said to be in statistical control. When some factor external to the process is introduced, such as a defect in material, operator error, or an inadvertent or incorrect adjustment to a machine, these "assignable causes" place the system in a state known as out of control. (Montgomery, 2000) Once the process goes out of control, manufacturing must be stopped and the cause investigated and corrected. SPC aims to detect as early as possible any assignable causes in order to minimize the costs of producing defective items.

The Shewhart chart is a relatively simple tool used for monitoring a process. An example is shown in Figure 1. The observations that are plotted are the measurements of

individual samples taken from the production line. The desired value of the process in this chart, e.g. the diameter of a circular object such as a washer, or the length of a bolt, is 5.00 units. The lines at 3.00 and 7.00 are the lower control limit (LCL) and the upper control limit (UCL), respectively; these are the extreme values of the distributional parameter being monitored (often the mean μ) that are permissible. While the process is in control, nearly all samples taken will have measured values within the control limits; these limits are specified so that the items produced in the process are highly likely to be within specification limits so long as the process is in control. Beyond these specification limit values, the items are not acceptable because they are either too big or too small. Warning limits are often added as well, placed between the target value and the control limits, to indicate that the process might be drifting toward one of its control limits, and that some action should be taken to prevent production of defective material. If the process is in control, a random pattern should appear on the control chart.

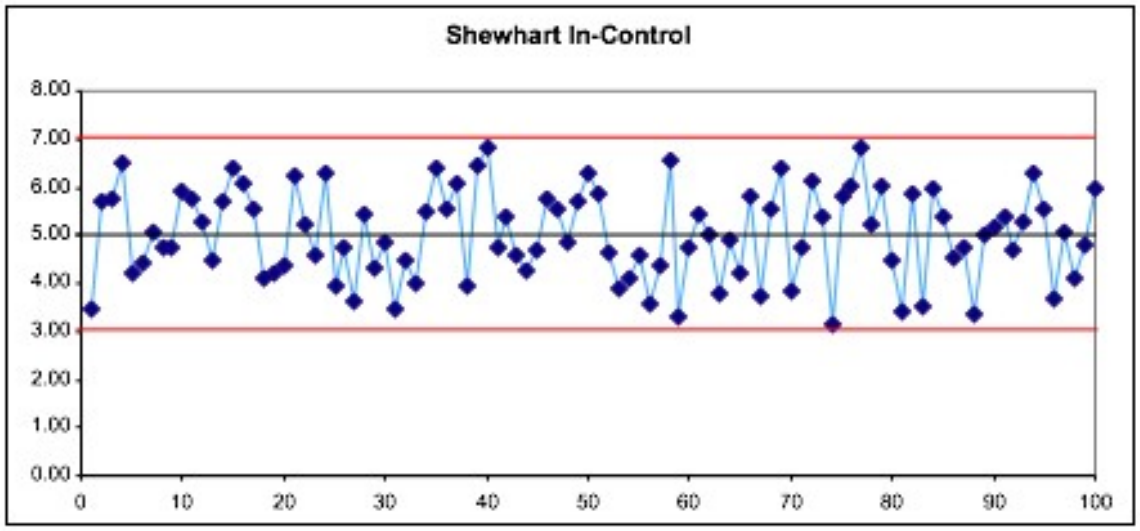

Figure 1: In-control Shewhart chart.
Process mean $\mu = 5$, $\sigma = 1$; control limits set at $+/-2\sigma$.

The upper and lower control limits can be chosen to be particular values, such as $+/-$ 0.1 millimeters from the target value, or they can be given in terms of a number *k* of standard deviations, $+/-k\sigma$.

The most common Shewhart chart is the $\bar{x}$ (x-bar) chart, which is used to monitor the mean of the in-control distribution. This form of the Shewhart chart assumes that samples are taken at some regular interval, and that the sample mean $\bar{x}$ is normally distributed by the central limit theorem. (Montgomery, 2000) However, the Shewhart chart is not restricted to monitoring production statistics that are normally distributed. Rather, the main idea of the Shewhart chart is to monitor when observations fall far out in the tail of the statistic's distribution. This idea can be applied to any distribution so long as the distribution can be specified and the control limits can be calculated as the appropriate quantiles of the distribution. Stoumbos and Reynolds (2000) provide analysis of the performance of Shewhart control charts against several heavy-tailed symmetric and asymmetric distributions. They demonstrate that a Shewhart chart with control limits $+/-k\sigma$ based on the normal distribution will detect shifts of other distributions, although with higher false alarm rates as evidenced by shorter average times to signal (ATS), due to the higher densities in the tail regions of non-normal distributions.

The Shewhart control chart in Figure 1 represents a normally distributed process with 100 observations drawn from a normal distribution. The process has a mean $\mu = 5.00$ and standard deviation $\sigma = 1$, with 2σ control limits. All observations fall between these control limits, so this process is in-control. Figure 2 is another Shewhart control chart using exactly the same observations, but with an upward shift in the distribution mean by 0.5σ for all observations starting at observation #40. As it turns out, this observation's value of 7.33 exceeds the control limit value ("threshold") of 7.00. At this point, the process should be stopped and investigated to find out whether there has been a change in the underlying distribution.

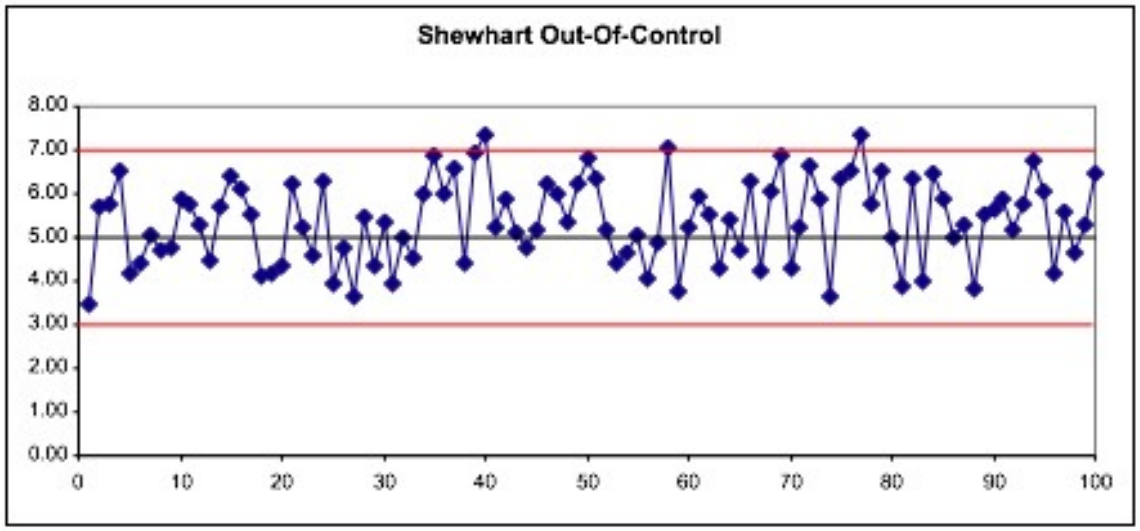

Figure 2: Out-of-control Shewhart chart.
$+0.5\sigma$ shift in process mean occurs at observation #40.

It should be noted that in the industrial quality control literature there is a great amount of discussion on efficient means of sampling an industrial process. While larger samples and those taken at more frequent intervals are generally more representative of the product being produced, the process of sampling itself, as well as the possibly destructive testing, each has its costs. There are, however, individuals control charts that directly plot individual measurements of every item produced instead of taking samples. These are used when it is not feasible to have a sample size larger than $n = 1$. Individual measurements are discussed in Montgomery (2001) and Smith (1998). For our purposes, and motivated by the biosurveillance problem we will discuss in Chapter III, we will assume that each item of data is being plotted on an individuals Shewhart chart, and not a sample mean from a multi-unit sample.

Statistical process control methods have evolved greatly since the introduction of the Shewhart chart. The cumulative sum chart (CUSUM) uses accumulated deviance from the target value over time to indicate departure from the in-control condition. In a similar manner, the exponentially weighted moving average chart (EWMA) accumulates evidence from historical observations, assigning greater weight to more recent observations. Because our study will rely on a method based on the Shewhart control

chart, details on the CUSUM and EWMA are not included here. See Box and Luceño (1997), Montgomery (2001), and Stoumbos and Reynolds (2000) for more details.

C. PREVIOUS STUDIES OF SPC IN NON-INDUSTRIAL SETTINGS

There have been several studies in recent years that attempt to use SPC techniques to monitor data in settings outside of the manufacturing industry. Montgomery (2001) points out that once we get past the idea of not having precise specification limits, we can apply SPC techniques to any process where we can take measurements that demonstrate quality or performance. A brief summary of several of these non-industrial applications of SPC follows.

Montgomery (2001) cites the use of sampling and testing using control charts by the finance division of a company to reduce the flow time of checks through the division, thereby reducing the percentage of invoices paid late and resulting in substantial savings to the company. Bamford and Greatbanks (2005) examine the use of basic quality tools for data collection, display, and statistical analysis for several routine processes that occur regularly in everyday life, such as time spent waiting for and receiving services or completing various tasks.

Gordon and Pollack (1994) developed a non-parametric statistical system using individual observations to detect departures from a symmetric distribution centered on zero, similar to the conditions in our study. They set a constraint on the false alarm rate for the single sensor by fixing the same average run length for each simulation. The model achieves performance comparable to a standard CUSUM for relatively longer runs from simulation start to the shift.

Most uses of SPC in health care surveillance have used variations on CUSUM control charts because of this tool's ability to detect small deviations over time. Statistical process control is well suited to use in manufacturing because while a system is in control, variability is caused only by random sources. In comparison, even in-control processes in health care monitoring experience non-random sources of variation known as overdispersion. Overdispersion occurs when the variability in the data exceeds that predicted by using an implicit relationship between the mean and the variance. (Hinde

and Dimétrio, 1998) This could be due to a multitude of factors that often cannot be controlled and may be difficult to account for in statistical models. This additional variability must be either modeled and removed or taken into consideration when setting detection thresholds in order to prevent excessive false alarm rates.

There have been many studies that have imposed simulated disease outbreaks overlaid on top of actual in-control health care data to test various detection algorithms. Mohtashemi, et al. (2006) used a Shewhart method to detect simulated outbreaks of influenza using the mean disease incidence level and a 2σ detection threshold, with a false-alarm rate constrained to be 3.3 percent annually. Fricker (2007a) converts existing multivariate methods into directionally sensitive univariate Shewhart and modified CUSUM methods, and successfully tests them against multivariate data. Olson, et al. (2005) considered the spatial distribution of emergency room patients with respiratory symptoms to detect clusters of disease outbreak.

Woodall (2006) weighs the advantages and disadvantages of the use of control charts in health care surveillance, and discusses some specific issues and techniques found by public health officials in their implementations that can be adapted by traditional SPC users in industry. Woodall concedes that the problem of overdispersion might only be solved by increasing control limit widths to reduce false alarm rates.

Aylin et al. (2003) apply a CUSUM control chart model to monitor deaths of patients under the care of primary care physicians in England. They attempt to estimate overdispersion in the healthcare practices they examined by using data from other practices that were assumed to be in-control. They point out a general lack of models that monitor multiple data sources over time. Marshall et al. (2004) uses a CUSUM for prospective monitoring of deaths of all types in England in district general hospitals. They also handle the issue of overdispersion, as well as attempting to handle multiple testing over time by controlling false alarm rate. Both studies use the framework of a shift of *k* standard deviations from a standard normal in-control distribution similar to the models in this thesis.

The Marshall model differs, however, from the models presented in this thesis in that all "sensors" in the system have their thresholds h set at an equal number of standard deviations from the in-control distribution, with $h \in \{2\sigma, 3\sigma, 4\sigma\}$. Marshall examines the percentages of sensors that signal true detections and false alarms for all combinations of distributional shifts $k\sigma$ (a CUSUM parameter) and h and for various numbers of sensors that have experienced actual shifts in the mean. In contrast, our model determines appropriate individual thresholds h_i for each sensor i to achieve a maximum overall probability of detection of an actual distributional shift, while limiting the probability of false alarm to a specified overall rate.

In their discussion of issues related to statistical epidemiologic surveillance, Wilson et al. (2006) weigh in on a specific trade-off that is inherent in any type of statistical monitoring, that of limiting type I and type II errors. (A type I error is a false alarm, or saying a distributional shift has occurred when none has; a type II error is a missed detection, or saying that no shift in the underlying distribution has occurred when in fact it has.) They conclude that while excessive false alarms may reduce readiness in the health community, the general method of setting limits on false alarms is inappropriate in this context because it increases the probability of missing an actual event (type II error) where lives would be lost and time wasted. Similarly, Washburn (2002) points out that in many sensor applications, the costs associated with false alarms and missed detections are so disparate that it is impractical to quantify both on the same scale.

D. ISSUES INVOLVED IN USE OF SPC TECHNIQUES IN NON-INDUSTRIAL APPLICATIONS

Because SPC techniques are concerned with identifying increases in the incidence of events in the presence of underlying noise, as well as for making an explicit trade-off between false signals and missed detections, they have broad applicability in many surveillance applications.

In a threshold detection sensor system, we are interested in sifting through the background noise of naturally-occurring incidents to identify a relatively small increase

in the number of cases above the background noise, which indicates a shift in the underlying distribution, while limiting type I and type II errors.

A major issue with statistical monitoring is systematic effects. The parameter being monitored can change over time, often due to factors that are not of interest to the model. Reis et al. (2003), Fricker et al. (2007a), and Fricker et al. (2007b) discuss statistical models that eliminate systematic components of the data, such as day-of-week and seasonal effects. This idea will be discussed further in Chapter II as we develop our model.

The design of control charts has evolved to include economic factors, since reducing the cost of manufacturing is a key objective of SPC. Some of these factors in the industrial sector are sampling costs, losses from producing sub-standard product, and the costs incurred from investigating false alarms. (Montgomery, 2000) In non-industrial applications of SPC, these factors are often not an issue because all data points must be examined in order to have sufficient coverage. Determining the sample size and frequency at which we sample is therefore not a concern. Because we are looking for an event of small magnitude, with the goal of quick detection, and because data in many applications are collected automatically from an electronic source, every data point and every interval of time is considered. These are then analyzed using individuals control charts, discussed in Montgomery (2001).

Finally, we should consider the type of control chart that is appropriate for the particular surveillance application. In some applications of SPC, any deviation—whether positive or negative—from the desired target level is considered an event of interest. In such a case, both directions should be monitored, using either a two-sided control chart or two one-sided control charts. In other applications, such as the biosurveillance problem discussed in Chapter III, only changes in one direction are of interest and thus the application of a single one-sided control chart will provide greater statistical power. For example, assume a Shewhart chart with an in-control distribution that is a standard normal $X \sim N(0,1)$ distribution, and a constraint on probability of false alarm set at five percent. For a two-sided test, the right- and left-hand tail areas will each have an area of

2.5 percent, resulting in z values of $+/-1.96$. Following a shift in the distribution mean of magnitude $\delta = 1$, the area to the left of +1.96 on the resulting normal distribution $X \sim N(1,1)$ will be 0.831. The area to the right of this threshold corresponds to the probability of detection, $P_d = 1 - 0.831 = 0.169$. If the test is one-sided, the right-hand tail will have an area of five percent and a z value of +1.64; the area to the right of +1.64 is 0.739. With the same shift in the distribution mean, the probability of detection will be $P_d = 1 - 0.739 = 0.261$. Table 1 demonstrates probabilities of detection for shifts of magnitude $\delta \in \{1\sigma, 2\sigma, 3\sigma\}$ for both one- and two-sided hypothesis tests. Clearly when we are only interested in detecting an increase in the mean, the one-sided hypothesis test performs better.

Magnitude of Shift	Two-Sided Test (P_d)	One-Sided Test (P_d)
$\delta = 1$	0.169	0.261
$\delta = 2$	0.516	0.641
$\delta = 3$	0.851	0.913

Table 1: Comparison of performance for one- and two-sided hypothesis tests. Assumes a standard normal $X \sim N(0,1)$ in-control distribution.

II. SURVEILLANCE THROUGH A SYSTEM OF CONTROL CHARTS

A. DESIGNING A SYSTEM OF CONTROL CHARTS

In industrial SPC, individual univariate control chart applications are designed and implemented using one of three approaches. Each of these design approaches is concerned not only with setting the detection threshold(s), but also with determining sampling frequency and sample size, both of which involve real production costs that reduce a commercial organization's profit. As discussed previously, our use of individuals control charts (i.e., control charts that plot every observation individually) allows us to ignore that portion of the SPC literature concerned with sampling and issues related to minimizing sampling costs.

The first approach for developing control charts is the statistically designed control chart, where the width of the control limit and the power of the test are pre-selected based on the average time between false signals (referred to as the in-control average run length or ARL). For a two-sided Shewhart chart with 3σ limits, where the statistic is normally distributed, there is a 0.0027 probability of committing a type I error (signaling that the process is out-of-control when in fact it is in-control). This means that 99.73 percent of observations taken while the process is in-control will fall between the upper and lower control limits and that a false signal will only be generated on average once every 370 observations. (Montgomery, 2000)

A second approach is the economic design of control charts, where thresholds along with sampling criteria are chosen in order to minimize the costs associated with quality control. These include the costs associated with monitoring production, the cost of false alarms, and investigation costs to find the assignable causes when production standards are not met.

A third approach is an economic-statistical design which seeks to minimize lost economic costs, while setting an upper limit on the probability of type I error and a lower

limit on the power, or probability of detection. Saniga (1989) also sets an upper limit on the average time to signal (ATS) for a given shift in the mean.

Unlike these approaches, here we seek not to design a single control chart, but a system of control charts. In so doing we use an essentially statistical approach, both for its simplicity, and because determining the costs of type I and type II errors in a surveillance setting is non-trivial. However, we note that if such costs are available, the optimization model could be modified to account for economic considerations as well. The model seeks to explore how to optimally set detection thresholds at each location to maximize overall probability of detection of a distributional shift of a specified magnitude, subject to constraints on the expected number of false alarms per time period. We note that this constraint is an implicit economic constraint, in the sense that false alarms must be investigated and such investigations consume resources.

B. MODELING ASSUMPTIONS

There are several major assumptions that we rely upon for the development of the model in this study. We first lay out the assumptions that abstract the system we are looking at, and then relate the mathematical and statistical assumptions used in the development and testing of our model.

The first simplifying assumption we make is that our system of sensors is designed, or at least has set its threshold levels, in order to detect increases in the background levels of the measure of interest. Several examples were cited in the previous chapter where SPC was used to detect deviations from an in-control state, each of which was interested in an increase in the distributional mean level of the statistic of interest. The previous chapter also discussed the statistical benefit of increased sensitivity that results from use of a one-sided control chart.

Because in this problem we assume that each observation is being monitored, we have designed the model around individuals Shewhart charts. Essentially, we are assuming that monitoring timeliness is so important that each observation will be monitored individually as it arrives. That is, instead of taking samples from a population, or waiting to aggregate a series of individual observations into a combined measurement,

we are examining every data point and comparing it to our threshold. Montgomery (2001) discusses the use of individuals control charts.

Another assumption that simplifies the problem for our purposes is that each sensor provides complete coverage of its assigned area, without any overlap with adjacent sensors. This allows for each sensor to be independent of all other sensors. In the real world, however, it will be difficult and perhaps even undesirable to implement a system of sensors where there is no overlap in coverage; it is better to have two sensors detect an event than to have a possibility that the one sensor where the event occurs will miss the event.

The final simplifying assumption we make is that at most one event (distributional shift) occurs per time period within our sensor network. Using Y_i as a binary variable to indicate the presence or absence of an event in sub-region *i*, we can express this assumption as:

$$\sum_i Y_i \leq 1 \quad ; \quad Y_i \in [0,1] \tag{1}$$

Allowing for more than one event per time period complicates study of the central problem. The problem as we have laid it out, however, is conservative in its estimate of the probability of detection: it will be easier to detect two independent events than just one. Looking at the problem of $\sum_i Y_i \leq 1$ then is a lower bound on the probability of detection for one or more events in the region per unit of time.

Using the law of total probability we briefly consider the mathematics of studying multiple events in a single time period. If we have *n* sub-regions, each with a particular probability of experiencing the event in a given time period, the sums of their individual probabilities must be one. If we randomly select one sub-region for the first event to occur, we must remove that sub-region from consideration for all other events in that time period. This gives us a new problem with $n-1$ sub-regions; all of the probabilities of event for these $n-1$ sub-regions must then be re-normalized to sum to one. This process can be repeated for any number of events for any time period. It is easy to see that this would best be modeled using simulation to examine the effects of different

numbers of expected events on our detection and false alarm performance. However, by assuming that at most one event occurs per time period, we can look at the probabilities of event and the probability of detection as a snapshot rather than as a sequential probability game. This greatly simplifies our analysis and its presentation in this thesis.

The first of the statistical assumptions we make is that the observations we are monitoring are independent and identically normally distributed. We justify this assumption by assuming that we are monitoring the standardized residuals from a model used to account for and remove systematic trends in the data. See Fricker et al. (2007a) and Fricker et al. (2007b) for discussion on the use of adaptive regression models to remove systematic trends from biosurveillance data, such as day-of-week or seasonal effects.

In addition, we assume that the event of interest will be manifested as a jump increase in the mean of the in-control distribution. While this may seem a broadly simplifying assumption, in statistical process control the control limits are established at values that are the minimum shift that is of interest to the production process. Shifts of smaller magnitude are presumably not interesting; conversely, much larger shifts in the mean are trivial to detect with a threshold set below the expected shift. For example, assume an in-control distribution $X \sim N(0,1)$, and our control limits set at $h = \mu + 2\sigma$. This will result in a relatively easy detection of a 3σ increase in the mean, and a relatively low probability of detecting a much smaller single standard deviation shift in the distribution. These scenarios are depicted in Table 1 in the previous chapter.

Thus we can represent the out-of-control state as a shift in the mean of the in-control distribution. In particular, if an observation from the in-control distribution is denoted as $X \sim N(0,\sigma^2)$, then an observation from the out-of-control distribution can be denoted as $X' \sim N(\delta,\sigma^2)$ where $\delta = k\sigma$ for some $k > 0$. Thus, if the in-control observations come from a standard normal distribution, an event of magnitude 2σ results in subsequent observations coming from $X' \sim N(2,1)$. (Montgomery et al., 2006)

The Shewhart methodology assumes that samples are independent of one another. However, the independence assumption is unrealistic in many applications, where

processes generate data that is autocorrelated over time. (Stoumbos and Reynolds, 2000) Such temporal change will significantly affect the performance of a Shewhart $\bar{x}$ chart; as previously described, we have assumed that temporal effects can be accounted for and removed from the data using an approach such as adaptive regression prior to implementing our model. Adapting our model for temporal methods such as CUSUM and EWMA, whose statistics are autocorrelated, is an area for future research.

C. MODEL OBJECTIVE

The objective of the model developed in this thesis is to maximize the probability of detecting a distributional shift (event) of a certain magnitude within a region, subject to a constraint on the expected number of false alarms per unit of time. The region is divided into multiple, disjoint sub-regions, each of which is being monitored for an event, and each of which has some probability of having the event occur within that sub-region. Recall from the previous section our assumptions of independence and full sensor coverage.

While the goal is to "tune" the sensor detection thresholds in each sub-region, such that the probability of detecting an event of particular magnitude somewhere in the region is maximized, there is a trade-off inherent in setting the detection thresholds, both in industrial SPC and in threshold detection applications. As a threshold is lowered toward the background noise of the data, the sensor becomes more sensitive to the natural variation in the data and is more likely to generate false alarms. As a result, detection thresholds must be set as low as possible in order to maximize the probability of detection, but not so low that the rate of resulting false positive signals is intolerable. Hence, the goal is to develop an optimization framework to determine individual thresholds for each sub-region that will maximize the overall probability of detecting an event, subject to an upper limit on the expected false alarm rate over the entire region.

To determine the optimal set of thresholds, we first need to determine the probability of an event occurring in each of the sub-regions. These probabilities can be determined or estimated in a variety of ways. The algorithm we use to estimate this probability of event is as follows. First, we assign each sub-region a numerical weight.

The weights can be any number on any scale, with larger weights indicating a higher likelihood of experiencing the event. These individual weights are then normalized by dividing each individual weight by the sum of all weights assigned across the region. This gives us each sub-region's proportion of the total weight, and thus its estimated probability of experiencing the event. This model is described mathematically in the following section.

D. MODEL FORMULATION

Assume a region has been divided up into N non-overlapping sub-regions, with each sub-region containing one sensor. The output from each sensor i, $i = 1,\ldots, n$, is expressed as X_i. When there is no event of interest, the system is in control, with the X_i's being independent and identically distributed. We express this as $X_{it} \sim F_0$ for all i and for $t = 1, 2, 3, \ldots$. If an event of interest occurs at time τ, then for one i $X_{it} \sim F_1$ for $t \geq \tau$. We are interested in determining a detection threshold h_i for each sub-region i to detect this distributional shift. A signal is generated at any time t when any sensor output X_{it} exceeds its threshold h_i.

Recall that in Chapter I we discussed the rationale for limiting the number of events per unit of time to no more than one. Equation (1) expressed this constraint using a binary variable Y_i. Thus, given some information about each sub-region and each sensor, we calculate or estimate the:

- Probability of an event in sub-region i: $\Pr(\text{event in sub-region } i)$;
- Probability of detecting an event should one occur in sub-region i: $\Pr(\text{detect} \mid \text{event in sub-region } i)$; and
- Expected number of false alarms in sub-region i per period: $E(\text{false alarms for sub-region } i)$.

Given $\Pr(\text{detect} \mid \text{event in subregion } i)$ and $\Pr(\text{event in subregion } i)$ for each sub-region, the probability the system detects an event in the whole region is

$$\Pr(\text{detect event on region}) = \sum_{i=1}^{n} \Pr(\text{detect} \mid \text{event in sub-region } i) \Pr(\text{event in sub-region } i) \qquad (2)$$

As previously described, we use the weight of a particular sub-region m_i as a fraction of the total regional weight $M = \sum_{i=1}^{n} m_i$ as a proxy for the probability of an event in that sub-region, with $\Pr(\text{event in region } i) = \frac{m_i}{M}$.

The probability of a false signal in sub-region *i* for a given threshold h_i is thus

$$p_i(h_i) = \int_{x=h_i}^{\infty} f_0(x)dx = 1 - F_0(h_i), \tag{3}$$

where *f(x)* is the pdf of the distribution of the statistic being monitored, under the assumption of no event, evaluated at *x*.

Using this same notation, the expected number of false alarms in sub-region *i* per time period is $E(\text{false alarms for sub-region } i) = 1 \times p_i(h_i) + 0 \times [1 - p_i(h_i)] = p_i(h_i)$. Assuming that the regions are independent, the expected number of false alarms for the region per time period is thus $\sum_{i=1}^{n} p_i(h_i)$.

Assuming that the event of interest manifests itself as a shift in the mean of the no-event distribution, so that $f_1(x) = f_0(x) + \delta$, $\delta > 0$, the probability of a true signal can be denoted as

$$p_i(h_i - \delta) = \int_{x=h_i}^{\infty} f_1(x)\,dx = \int_{x=h_i-\delta}^{\infty} f_0(x)\,dx = 1 - F_0(h_i - \delta) \tag{4}$$

Given the preceding, we can then express the problem of maximizing the region-wide probability of detection subject to a constraint (κ) on the expected number of false signals as a non-linear programming model (NLP), which we will refer to as STOPT (Sensor Threshold Optimization) later in the thesis.

Indices

i	Sub-region within area of interest

Data

m_i	Relative weight assigned to sub-region i
κ	Max allowed false alarms in region
δ	Magnitude of shift in mean incidence level

Variables (5)

h_i	Detection threshold in sub-region i

Formulation

$$\underset{\bar{h}}{\text{maximize}}\ p_d \equiv \sum_i \frac{m_i}{M}\left[1 - F_0(h_i - \delta)\right]$$

$$\text{subject to} \sum_{i=1}^{N}\left[1 - F_0(h_i)\right] \le \kappa,$$

Due to the potential for high false alarm rates in sub-regions that have a higher probability of detecting an event, one possible variation would be to add a second constraint to limit the probability that a given sub-region will experience false alarms. Alternatively, we could specify a minimum acceptable probability of detection for a particular sensor, by setting a lower bound on the probability of detection. These respective constraints take the form of Equations (6) and (7).

$$1 - F_0(h_i) \le \alpha_u \quad ; \quad 0 \le \alpha_u \le 1 \tag{6}$$

$$1 - F_1(h_i) \ge \rho_l \quad ; \quad 0 \le \rho_l \le 1 \tag{7}$$

The inequality in (6) places an upper limit α_u on the probability of signaling a false alarm for a particular location, and (7) places a lower limit ρ_l on probability of detection by a particular sensor. The effect of (6) is to force the thresholds for the constrained sub-regions to move higher, forcing them away from the underlying noise to lessen the probability of false alarm. While this is desirable in terms of reducing resources spent on investigating false alarms, the constraint also decreases the probability of detection for those sub-regions. Equation (7) does the opposite, pulling thresholds down to increase sensitivity, with the trade-off that the probability of false alarm increases.

The STOPT model in (5) is designed so that we can control the probability of detection and the number of region-wide false alarms by changing the value on the right hand side of the constraint. This model has the potential for enhancement by tailoring the constraints in (5), (6), and (7) to meet the objectives of the system. Expressed as a nonlinear program, and assuming that the in-control distribution at each sensor is $X_i \sim N(0,1)$ with a shift of magnitude δ in the mean incidence level, our model becomes:

Indices

i	Sub-region within area of interest

Data

m_i	Relative weight assigned to sub-region i
κ	Max allowed false alarms in region
δ	Magnitude of shift in mean incidence level

Variables

h_i	Detection threshold in sub-region i

Formulation

$$\underset{\vec{h}}{\text{maximize}} \sum_i \frac{m_i}{\sum m_i}[1-\Phi(h_i-\delta)]$$

$$\text{subject to} \sum_i [1-\Phi(h_i)] \le \kappa$$

$\Phi(h_i)$ is the standard normal cumulative distribution function (CDF) evaluated at the threshold h_i. So the term $1-\Phi(h_i)$ in the constraint is the area to the right of the detection threshold h_i, representing the probability of committing a type I error.

$\Phi(h_i-\delta)$ is the standard normal CDF evaluated on the out-of-control distribution, which in the case of our models is simply the in-control distribution $X_i \sim N(0,1)$ shifted rightward by some number of sample standard deviations $k\sigma = \delta$, with a resulting distribution of $X_i' \sim N(\delta,1)$. So the term $1-\Phi(h_i-\delta)$ is the area to the right of the detection threshold h_i, representing the probability of detecting an actual event.

Because both our objective function and constraint are nonlinear, we need to determine whether a globally optimal solution exists. It is possible in nonlinear programming problems for solver algorithms to converge to a point that is locally optimal, despite the fact that at other areas on the function or space of interest there are areas where we can do much better.

It is relatively easy to demonstrate that the objective function in (5) is strongly quasiconvex over the constraint regions, which in a minimization problem would give a globally optimal solution. Unfortunately, because this is a maximization problem, a globally optimal solution is not guaranteed on the basis of quasiconvexity of the objective function alone. (Bazaraa et al., 1993) However, under the previously specified assumptions, and as described in Fricker and Banschbach (2008), a globally optimal solution does exist.

Overall the results from testing of the model in (5) show us that areas assigned a higher probability of event will have thresholds set closer to the mean incidence level to achieve higher probabilities of detection, with accompanying higher probabilities of false alarm; areas with lower assigned probabilities of event will have higher thresholds to keep false alarm rates low, with lower probabilities of detection. The overall expected number of false alarms system-wide must still be no greater than some level κ that is specified as reasonable. We discuss detailed results in the following chapter.

III. SETTING OPTIMAL THRESHOLDS

A. TOOLS FOR DESIGNING, TESTING, AND IMPLEMENTING MODELS

The models in this thesis were implemented in Microsoft Excel, using Frontline's Premium Solver (student version) to solve the nonlinear optimization problems. The main advantages of Excel are its familiarity to many users, and the ease with which model inputs can be changed and the effects of the change recognized immediately. The Premium Solver uses the Generalized Reduced Gradient (GRG2) nonlinear optimization algorithm, described in Lasdon, et al. (1978)

Excel's solver, however, limited the size of problem that could be examined. Because our problem is nonlinear, there is a limit of 200 variables over which this nonlinear solver can optimize. Problems with more variables can be solved using more advanced versions of Frontline's Premium Solver or a more powerful optimization package such as the General Algebraic Modeling System (GAMS). Appendix A contains GAMS code for a small sample problem that can be expanded to larger models.

The remaining sections of this chapter seek to apply our algorithm to the intended problem of biosurveillance through the use of various examples and excursions. We conclude with Section E, which describes a model that takes advantage of the normality assumption to reduce the *n*-variable optimization problem to a one-variable optimization. The model is derived in Appendix B and described in Fricker and Banschbach (2008), and is limited only by the dimensions of the Excel spreadsheet.

B. A PRACTICAL APPLICATION OF A THRESHOLD-BASED DETECTION SYSTEM TO EPIDEMIOLOGICAL SURVEILLANCE

The problem that motivated this thesis is that of improving the performance of epidemiological surveillance systems, sometimes also called syndromic surveillance. (A syndrome is a grouping of diseases, each of which presents similar symptoms. See Fricker (2007b) for a detailed overview of syndromic surveillance.) Such public health monitoring systems have operated at the federal, state, and local levels for several years.

Fricker (2007b), Fricker and Rolka (2006), Stoto et al. (2004), and Pavlin et al. (2003) each describe epidemiological or syndromic surveillance, and discuss the issues of sensitivity and specificity in the area of health care surveillance. Bravata et al. (2004), Sebastiani and Mandl (2004), Heffernan et al. (2004a and 2004b), Loonsk (2004), and Wagner et al. (2004), each describe several implementations of syndromic surveillance, and Lober et al. (2002) compares several systems that have been used in recent years.

In epidemiological surveillance, three factors are crucial to success: reporting timeliness; analysis of public health data; and high sensitivity and specificity, where high specificity is analogous to a low rate of type I errors and therefore an ability to distinguish between noise and actual events. Bravata et al. (2004) point to the lack of study on sensitivity and specificity in routine health care surveillance, a major shortcoming that we will address explicitly in the implementation of our model. Additionally, many implementations of epidemiologic surveillance systems are seemingly "ad-hoc," with the common belief that adding more sensors will improve detection performance. It is often soon realized following implementation that having more sensors leads to higher false alarm rates and no improvement in detection of even significant naturally occurring health events.

To avoid complications from applying our model to an existing system, we will apply it to an abstract system that captures the essence of national epidemiological surveillance, so as to not get overly specific with the details of any particular system. We will assume a centrally monitored system of geographically distributed statistical sensors examining health care data on hospital admissions for a single syndrome. We will also assume that our system is using a simple Shewhart-type control chart of individual observations (say patient admission counts, for example) to examine data in a single period of time with systematic effects already removed. As previously discussed, temporal effects generally cause some autocorrelation between time periods, but which we assume can be removed via appropriate modeling. Future research is needed to apply our optimization techniques to other temporal methods such as the CUSUM and EWMA, whose statistics are autocorrelated. All other assumptions about independent sensors with

non-overlapping coverage still apply. Finally we assume that our control charts are one-sided to detect only increases in the number of cases.

The sub-regions we will consider are individual counties; it is assumed that a single sensor provides coverage for each county. To determine the regional event probability assigned to each county, and in the absence of other information, we use the fraction of the population in a particular county (out of the total population of all counties present in the model) as a proxy for probability of an outbreak or biological terror attack in that county. While likely overly simplistic, in an actual implementation of this method these probabilities of attack can be determined by any method desired. That said, using population as a surrogate is not necessarily unreasonable. For example, areas such as cities that have large, dense populations are more likely than sparsely populated areas to experience severe outbreaks of naturally occurring contagious diseases. In terms of bioterrorism, it is also not unreasonable to assume the objective of an attack would be to kill or sicken as many people as possible and consequently larger cities have greater appeal as bioterrorist targets.

We will now present some models using population to estimate the probability of detecting a biological event of interest, such as a bioterrorist attack or natural outbreak, which manifests itself as an increase in the distribution underlying the number of cases of a particular syndrome in one sensor location. As we develop from simple models of only ten to twenty counties to the use of population figures for the two hundred most populous counties in the U.S., we examine the policy implications that come along with the decision to add more sensors. Often it is assumed that having more sensors will lead to improved probability of detection by additional coverage area, but our model reveals that there are both costs and benefits of such system expansions.

Note that throughout our example problems we hold constant the parameters of the models we directly compare to each other. These include the magnitude of expected distribution shift $k\sigma = 2$ and the constraint on expected number of false alarms κ per time period ($\kappa = 1$ for the first two small hypothetical models; $\kappa = 4$ for the remaining models using actual population values). Despite setting these parameters somewhat

arbitrarily in our examples, they can easily be changed to explore different scenarios as well as the sensitivity of any system to different modeling assumptions.

C. POPULATION-BASED MODELS

1. Model Using Ten Counties (STOPT10)

The model described in Equation (5) is first implemented using fictional population data for a set of ten counties (STOPT10). The population of each county is the term m_i, which we use as the relative weight for each county. Using this relative weighting system, each county is then assigned a probability of experiencing an event, given that an event occurs, which is its proportion m_i/M of the total population M.

The sensor in each county is assumed to be independent of those in all other counties. The populations and the resulting individual thresholds, probabilities of detection (p_i), and false alarm (α_i), for each sensor, are shown in Table 2 below. The in-control model assumes the sensor readings are $X_i \sim N(0,1)$ and the out-of-control condition manifests as a 2σ shift in the mean of the in-control state, resulting in $X_i \sim N(2,1)$ for one county.

i	m_i	m_i/M	h_i	$p_i = \Phi(h_i - \delta)$	$\alpha_i = \Phi(h_i)$
1	1,000,000	0.169	1.009	0.839	0.156
2	800,000	0.136	1.121	0.810	0.131
3	700,000	0.119	1.187	0.792	0.118
4	600,000	0.102	1.264	0.769	0.103
5	600,000	0.102	1.264	0.769	0.103
6	500,000	0.085	1.356	0.740	0.088
7	500,000	0.085	1.356	0.740	0.088
8	400,000	0.068	1.467	0.703	0.071
9	400,000	0.068	1.467	0.703	0.071
10	400,000	0.068	1.467	0.703	0.071

Table 2: Model results from ten counties (STOPT10). $P_d = 0.771$. Assumes an $X_i \sim N(0,1)$ in-control distribution, shift in mean level of 2σ, and false alarm rate limited to $\kappa = 1$.

The overall probability of detecting an event of magnitude 2σ somewhere in the region, given that we are willing to accept one false alarm, is 0.771. When we allow for

two false alarms, all individual thresholds are lowered and the overall probability increases to 0.881. The two rightmost columns (p_i and α_i) are the individual probabilities of detection and false alarm, respectively, for each individual county *i*, for the case of one false alarm per time period. As we lower individual thresholds to improve p_i, the individual false alarm probabilities α_i increase.

One factor to keep in mind throughout the discussion of population-based models is the fact that the distribution of the total population is skewed rather heavily toward a relatively small number of areas. It is likely to be the case in other implementations that large percentages of population will be concentrated in a few areas such as major cities, with the remainder spread over a large number of smaller rural areas.

2. Model Using Twenty Counties (STOPT20): Effects of Adding More Sensors to the System

Expanding the population-based model STOPT10 depicted in Table 2 to twenty counties provides the model output STOPT20 in Table 3. Counties #1 through #10 are the same as in STOPT10. Allowing for one false alarm per time period, the overall probability of detection of a 2σ event for STOPT20 is 0.662. As before, allowing for more false alarms allows all thresholds h_i to be lowered, resulting in higher probabilities of detection p_i. Allowing for two false alarms the probability of detection of the same 2σ event is 0.782. The detection thresholds h_i for each of the ten largest counties are higher in STOPT20 than for the same scenario in STOPT10, resulting in lower individual probabilities of detection p_i but also lower probabilities of false alarm α_i.

The effect of scaling-up the problem from ten to twenty counties is clearly negative for the ten largest counties due to lower individual probabilities of detection p_i; yet the benefit to counties #11 through #20 is obvious because in the second model they each have some level of sensor coverage, where previously they had none.

A key area for statistical and policy analysis is as follows. Adding additional sensors to cover less-populous areas detracts from the more-populous areas and from the overall probability of detecting an event of a given magnitude (given the event occurs in

one of the more-populous areas), but there is significant individual benefit to the counties that are added. Does the additional benefit gained by providing coverage to more of the population outweigh the lower probability of detection for the most populous counties?

i	m_i	m_i/M	h_i	$p_i = \Phi(h_i - \delta)$	$\alpha_i = \Phi(h_i)$
1	1,000,000	0.119	1.200	0.788	0.115
2	800,000	0.095	1.311	0.755	0.095
3	700,000	0.083	1.377	0.733	0.084
4	600,000	0.071	1.456	0.707	0.073
5	600,000	0.071	1.456	0.707	0.073
6	500,000	0.060	1.546	0.675	0.061
7	500,000	0.060	1.546	0.675	0.061
8	400,000	0.048	1.657	0.634	0.049
9	400,000	0.048	1.657	0.634	0.049
10	400,000	0.048	1.657	0.634	0.049
11	300,000	0.036	1.801	0.579	0.036
12	300,000	0.036	1.801	0.579	0.036
13	300,000	0.036	1.801	0.579	0.036
14	300,000	0.036	1.801	0.579	0.036
15	250,000	0.030	1.893	0.543	0.029
16	250,000	0.030	1.893	0.543	0.029
17	250,000	0.030	1.893	0.543	0.029
18	200,000	0.024	2.003	0.499	0.023
19	200,000	0.024	2.003	0.499	0.023
20	150,000	0.018	2.147	0.441	0.016

Table 3: Model results from twenty counties (STOPT20). $P_d = 0.662$. Assumes an $X_i \sim N(0,1)$ in-control distribution, shift in mean level of 2σ, and false alarm rate limited to $\kappa = 1$.

3. Model Using 200 Counties (STOPT200): A Practical Application of the Model to the U.S.

The models in the preceding two sub-sections are now expanded to a representative system of sensors to monitor the populations of the two hundred largest counties in the United States (STOPT200). This is the largest single model that can be tested using the standard Excel solver applied to an *n*-variable optimization problem. For STOPT200, the in-control situation is assumed to be a standard normal distribution $X_i \sim N(0,1)$. Assuming a limit on false alarms to an arbitrarily chosen rate of four per time period, the probabilities of detecting an event with a given shift from the in-control

mean are given in Table 4. A table showing complete results and individual counties' probabilities of detection for a 2σ shift is shown in Appendix C; the following sections will expand analysis upon this model to explore different assumptions and inputs to the model.

Magnitude of Shift	Probability of Detection
$\sigma = 1$	0.230
$\sigma = 2$	0.537
$\sigma = 3$	0.851

Table 4: Comparison of probabilities of detection for various magnitudes of shift using STOPT200.
Assumes an $X_i \sim N(0,1)$ in-control mean, with an expected value of $\kappa = 4$ false alarms per time period.

Expanding upon this analysis, we now compare the results of our approach with the results achieved from naively setting equal thresholds for all sensors to achieve the same probabilities of detection. These results are summarized in Table 5 for various magnitudes of distributional shift in the background incidence levels. For each magnitude of shift, we set performance as a constant in terms of probability of detection, with our optimization model maintaining a false alarm rate of four per time period as before. For each magnitude of shift in the mean, we solve for the set of equal thresholds h_i that will give us the desired probability of detection. Note that in order to achieve comparable detection performance against each shift, the false alarm rates incurred by the naïve method increase dramatically. This is a result of setting thresholds too high for a relatively small number of large counties and too low for the large number of small counties. For a 2σ shift, the optimized thresholds for Los Angeles County (#1) and Lorain County, OH (#200), were 0.812 and 2.560, respectively. In a practical sense, the naïve system of equal thresholds is a misallocation of resources, where the locations with more population ("weight" in a generic sense) receive too little sensitivity and those with less population receive too much.

Magnitude of Shift	Probability of Detection	Thresholds (Naïve)	# False Alarms (Optimized)	# False Alarms (Naïve)	Percentage Increase in False Alarm Rate
$\sigma = 1$	0.230	1.739	4	8.206	105.2%
$\sigma = 2$	0.537	1.907	4	5.651	41.3%
$\sigma = 3$	0.851	1.959	4	5.008	25.2%

Table 5: Comparison of false alarm rates between naïve and optimized methods using STOPT200.

4. Comparison of the Performance of the U.S. Counties Models for 10, 20, 40, 100, and 200 Counties

A key issue that must be discussed by policymakers is that of how many sensors to employ. As mentioned previously when we looked at ten- and twenty-county models, the overall probability of detection of an event decreased when we increased the number of counties. However, the additional counties that previously had no coverage benefited a great deal by adding sensors.

In the national epidemiologic surveillance context, there likely will be conflict between the local and national public health authorities on this issue. Those in heavily-populated cities and counties will advocate having fewer sensors located in only the most-populous areas in order to provide the highest probabilities of detection in those localities, which because of their large populations present themselves as obvious targets to bioterrorists. National authorities may argue either way. One argument is similar to that just discussed, that of concentrating national surveillance efforts in the most-populous areas; they would use the higher probability of detection as an argument to prevent expansion of the system. The other argument is that by leaving large numbers of counties and therefore a large proportion of the population uncovered—possibly many smaller counties whose total populations are larger than that of the handful of largest counties—that other populous areas are left unprotected and are therefore attractive targets. Table 6 and Figure 3 demonstrate that the overall probability of detection for events of given magnitudes decrease as more sensors are added to the system.

# Counties:	10	20	40	100	200
$\sigma = 1$	0.791	0.614	0.459	0.305	0.230
$\sigma = 2$	0.964	0.890	0.790	0.641	0.537
$\sigma = 3$	0.997	0.986	0.962	0.907	0.851

Table 6: Overall detection performance (P_d) for various numbers of sensors versus a 2σ shift.
Assumes false alarm rate of $\kappa = 4$ per time period.

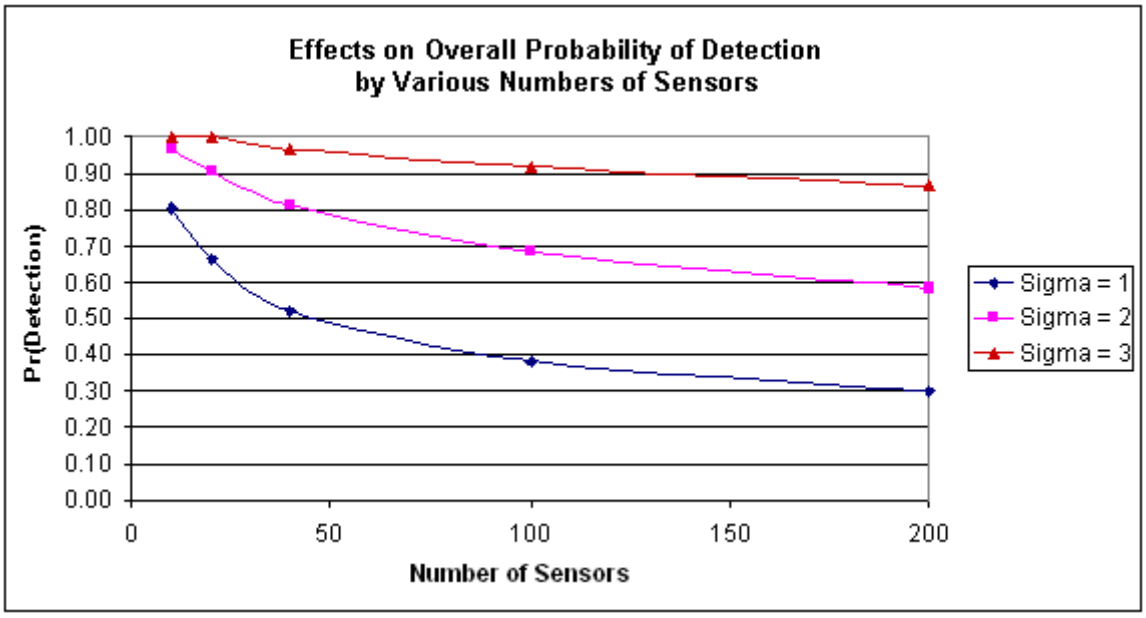

Figure 3: Overall detection performance (P_d) for various numbers of sensors versus a 2σ shift.
Assumes false alarm rate of $\kappa = 4$ per time period.

This thesis has already demonstrated that adding more sensors to a system has both costs and benefits. The marginal sensor adds detection capability for the county where it is installed, increasing the probability of detection from zero to some appreciable level. But this addition detracts from counties with sensors already installed, particularly those with larger populations, and decreases the probability of detection for the entire system.

D. EXCURSIONS FROM STOPT200

1. Effect of Placing Additional Constraints on Individual Thresholds

In implementing a model like STOPT200, we may want to add additional constraints to tailor the model to specific needs. Using constraints similar to those in Equations (6) and (7), we will examine two cases. The first is the case where we desire to artificially set a threshold higher than the optimal h_i^* from STOPT200 to drive down false alarm probability at that location; in the second case we want to set the threshold lower to improve detection probability in a certain location.

The first case is that of setting an upper limit on the probability of false alarm for a particular sensor. In the output of STOPT200 (two hundred counties) shown in Appendix C, we notice that in Los Angeles County (#1), if no event occurs, there will be a 0.209 probability of experiencing a false alarm. If this is felt to be excessive, we can force this probability (or any other α_i) to be lower by setting a constraint on the maximum allowed probability of false alarm. Adding the constraint $\alpha_1 \leq 0.15$ results in a higher threshold h_1 and lower probability of detection in the county, down from 0.883 to 0.832. Solving STOPT200 with the additional constraint results in an overall probability of detection of 0.537, unchanged from the previous optimal solution. The impact on other counties is minimal, with only very minor changes in probabilities of detection and false alarm (all less than about six percent, and most less than one percent).

The other case is that of ensuring that the probability of detection at a particular sensor meets a specified minimum probability of detection. This could be due to a high-profile event such as a major political or trade convention or sporting event where large crowds could be targeted. This situation of tuning sensors already in-place is a variation on the use of drop-in surveillance discussed in Fricker (2007b), Hutwagner (2003), Toprani et al. (2005), and Sebastiani and Mandl (2004). Alternatively this method could be used to compensate an area that has a lower assigned probability of attack than its importance or perception as a potential target. For example, Washington, D.C., is the 23rd–most populous city in the U.S., yet because of the concentration of federal

government facilities and workers, it should be given a "weight" comparable to that of the most-populous city, New York. For our example, we will look at Denver County, Colorado. With approximately 567,000 people in 2006, it was the 105th-most populous county, and will be the site of the 2008 Democratic National Convention.

In this second case, we add a constraint of the form in Equation (7) to establish a minimum probability of detection in that location. Again referring back to the optimized output of STOPT200 in Appendix C, we see that the probability of detection in Denver County is only about forty percent (0.404). Our new constraint on STOPT200 to require a minimum probability of detection in Denver County takes the form $\rho_{105} \geq \rho_{\min}$. Table 7 shows the results of several minimum probabilities of detection for Denver, with $\rho_{\min}$ ranging from 0.500 to 0.900. As the requirement for detection probability becomes more stringent, we notice the thresholds h_{105} going down which also results in higher false alarm rates α_{105}.

Solution	$p_{\min}$	h_{105}	α_{105}	P_d
1	0.404	2.244	0.012	0.537
2	0.500	2.000	0.023	0.537
3	0.600	1.747	0.040	0.537
4	0.700	1.476	0.070	0.536
5	0.800	1.158	0.123	0.534
6	0.900	0.718	0.236	0.530

Table 7: Effects of various threshold constraints for Denver County, CO. Solution 1 denotes previous optimal solution.

This scenario of the national political convention motivated us to change the requirements of the STOPT200 model. Denver County's threshold h_{105} was lowered to increase its probability of detecting an event of magnitude 2σ. If we desire a probability of detection in this county as high as ninety percent, we will have to lower Denver's threshold to 0.718, roughly the same as the fifth-largest county (h_5 = 0.714). The system-wide effect of changing the threshold in just one county by this large amount, however, is minimal, with the overall probability of detection falling by just 0.7 percentage points. In contrast to the uniform downward shift of all thresholds h_i as discussed earlier in this section, we are only shifting one threshold to improve detection performance in that location. It is possible to set constraints on multiple sensors, but care must be given not to

over-constrain the model; doing so could severely degrade overall performance by moving the system too far from the optimal combination.

Note that it is necessary to re-solve the optimization after setting these additional constraints. If instead we only adjust the desired threshold h_{105} without re-solving, the false alarm constraint would be exceeded; for h_1 the expected number of false alarms would be less than the constraint, and we would have an artificially low probability of detection.

2. Effect of a Uniform Downward Shift in Individual Thresholds h_i by All Sensors

Given the optimal solutions generated in the previous chapter, one scenario that comes to mind is that of implementing a uniform shift of all thresholds h_i. The motivation for this could be an increased likelihood of an event somewhere in the region, and a willingness to handle additional false alarms during a particular, although limited, time frame. Thus, after optimizing STOPT200 as before, we lower all h_i's by the same percentage to increase the probability of detection until the number of expected false alarms reaches an upper limit deemed acceptable. The converse situation of raising all h_i's seems less likely to occur. Note that a one hundred percent decrease in all h_i's results in all thresholds becoming equal to zero. A threshold of $h_i = 0$ results in a probability of detection equal to 0.977 for a shift of magnitude 2σ, but there will be a very high number of false alarms as each piece of random noise results in a false alarm.

The effect that we expect from gradually lowering all h_i's is to see an increasing probability of detection and an increasing number of false alarms. Figure 4 demonstrates that the rate of increase of false alarms outstrips the rate of increase in probability of detection. We notice that on the left-hand side of the graph the "P(Detection)" curve is relatively steep and the "# False Alarms" curve is relatively flat; the converse is true on the right-hand side of the graph where the probability of detection is asymptotic to 1.00 and the number of false alarms is increasing rapidly. The effect is that of diminishing marginal returns from decreasing the set of h_i's. We conclude that a small system-wide percentage decrease in h_i's is appropriate to increase overall probability of detection,

keeping in mind that the percentage increase in false alarms will be much greater than the percentage increase in probability of detection, particularly for larger percentage shifts in h_i.

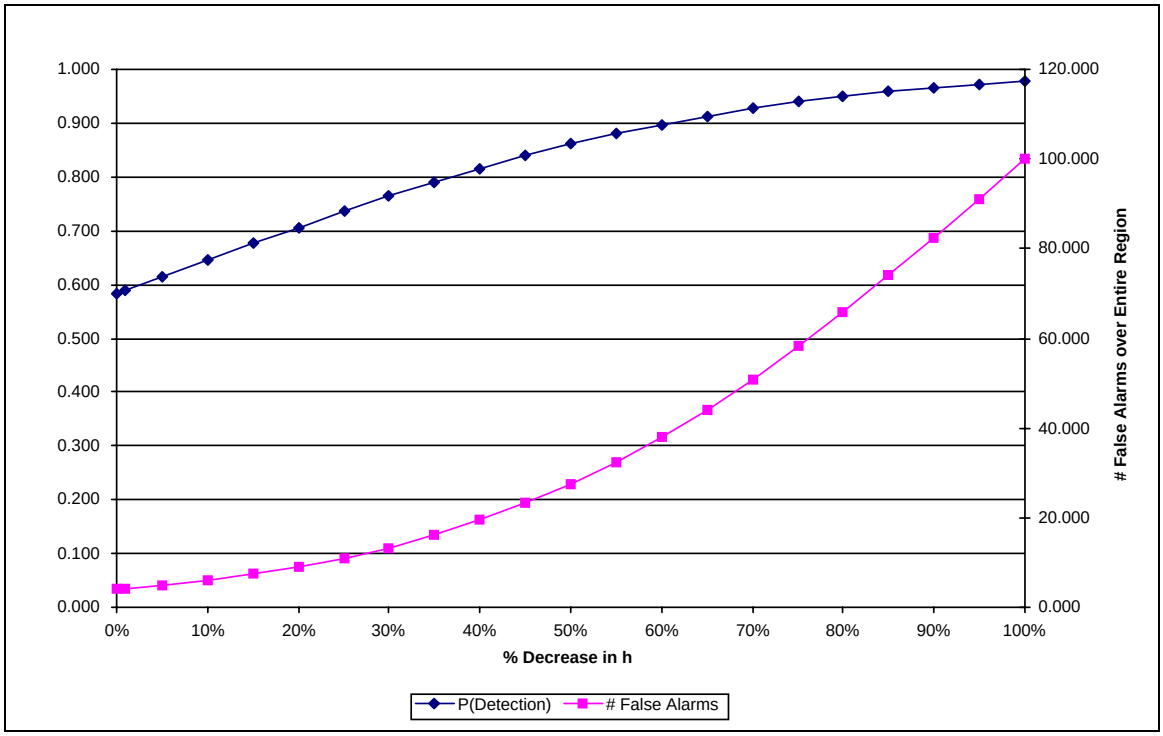

Figure 4: Effects on performance (P_d) from uniform percentage decrease in all thresholds.
Assumes a 2σ shift in mean disease incidence level from $X_i \sim N(0,1)$ in-control mean.

3. Effect of Independent Shifts in Thresholds h_i by All Sensors

Given an optimal solution for appropriate detection thresholds from STOPT200, one policy possibility is to give each locality some amount of autonomy to adjust its sensor's threshold h_i up or down. The threshold decision by each sensor would be arrived at by examining a variety of factors for its area, including scarcity of resources to investigate additional false alarms (forces a higher h_i), or a desire to increase probability of detection even at risk of higher false alarms (forces a lower h_i). It may be decided to allow each locality to raise or lower its threshold by up to a certain percentage.

We simulate these perturbations of each sensor's detection threshold about the optimal h_i found in STOPT200 to generate a new set of h_i's. Each new threshold is a normal random variable with its mean equal to the original h_i and standard deviation of $p \cdot h_i$ where p is the percentage of fluctuation that we allow for the sensor. Thus, the new h_i is a random variable $Y_i \sim N(h_i, p \bullet h_i)$. This vector of new random variables results in more variability, and therefore more freedom, to areas with smaller populations or weights and less variability to areas with larger populations, which have a larger effect on the overall system performance.

For the simulation, samples of one hundred sets of random h_i's were generated for each allowed percentage of fluctuation. Again, there are two measures that are of interest to us as we examine this possibility: the average probability of detection and number of false alarms at each level. Figures 4 and 5 show a typical results for probability of detection and number of false alarms for percentages ranging from five to two hundred percent of the h_i values. As the percentage of fluctuation increases, the mean probability of detection stays relatively constant (Figure 5). This indicates that even as we increase the amount of shift in h_i, the probability of detection does not change noticeably.

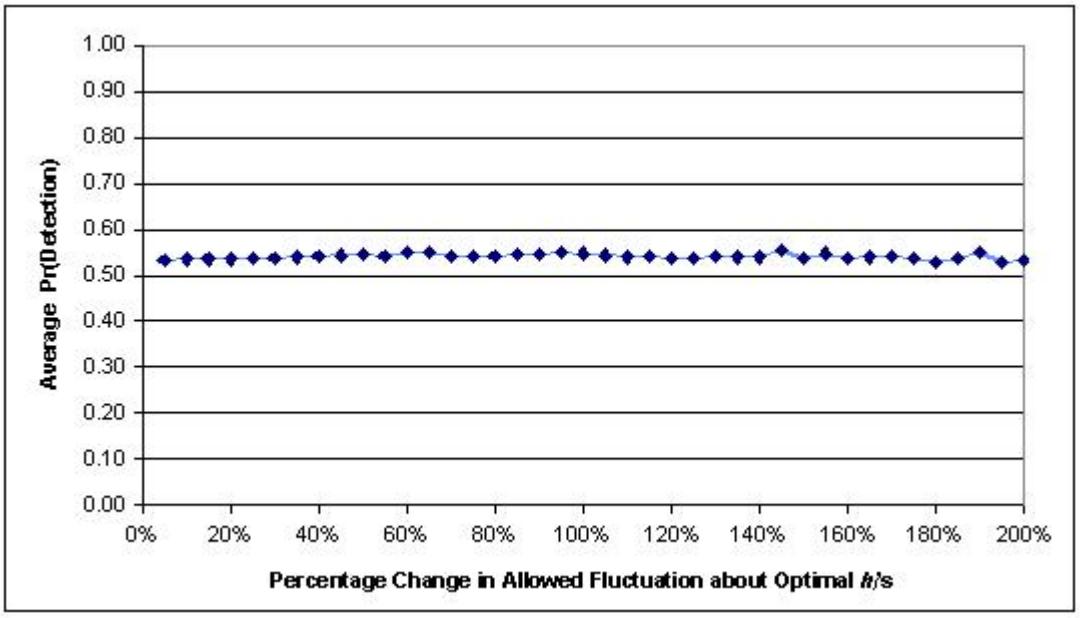

Figure 5: Effect on overall performance (P_d) of STOPT200 by random shifts in individual thresholds.

However, as the percentage of fluctuation in h_i increases, the number of false alarms increases significantly (Figure 6). Despite causing very little change in probability of detection, an increase in the percentage of allowable change in h_i by two hundred percent raises the average number of false alarms by nearly sixteen hundred percent, from four false alarms to sixty-two per period. Even a thirty percent change in h_i causes a doubling of the rate of false alarms.

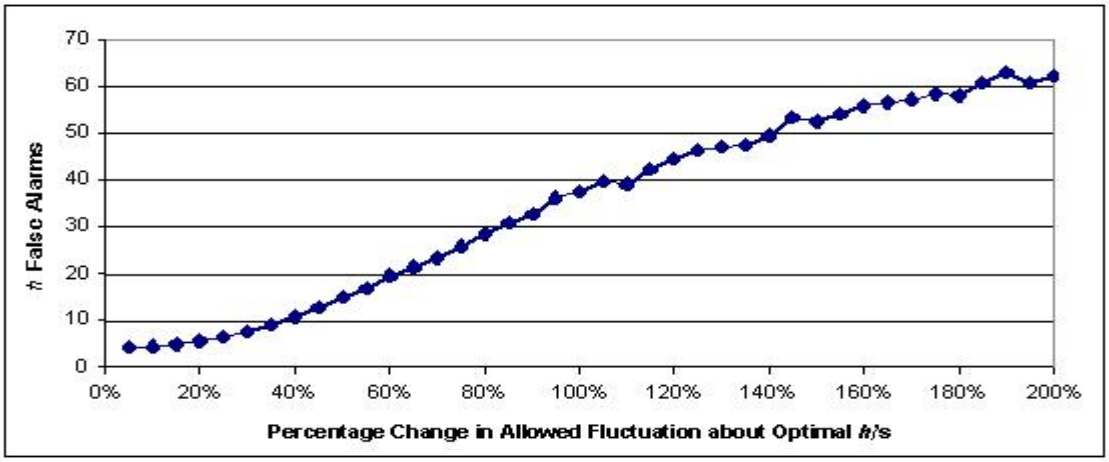

Figure 6: Effect on overall false alarms of STOPT200 by random shifts in individual thresholds.

Clearly when individual sub-regions are allowed to vary their thresholds widely, Figure 5 shows there is no additional benefit to the system overall, while Figure 6 shows the false alarms rise dramatically. But if the percentage of fluctuation allowed is kept under twenty percent or so, the false alarm rate stays relatively constant at or below five per time period.

E. ANALYTICAL SOLUTIONS

The assumption of normality leads us to a simplified optimization solution that reduces the *n*-variable problem of Equation (5) to a one-variable optimization problem. Appendix B and Fricker and Banschbach (2008) discuss the derivation, which takes advantage of some useful properties of the normal distribution. The result is a log-linear relationship between the individual county's probability of a biologic event of interest somewhere in the country, and the set of optimal thresholds h_i.

We apply this new, reduced model to more easily solve the problem of setting thresholds for two hundred counties with the same outcome as presented in Section C.3 of this chapter. Because the single-variable nonlinear optimization problem solves much faster than the *n*-dimensional problems we have already shown, we are able to quickly solve problems of practically unlimited size, without experiencing problems due to number of variables. Below we discuss the impact of placing one threshold sensor in each of the 3,141 counties in the United States, and compare the results with that of the model for only two hundred counties.

For these examples we keep the same parameters as in previous sections of this chapter, with the limit on expected number of false alarms maintained at four per time period ($\kappa = 4$). The overall probability of detection falls from 0.537 in STOPT200 to 0.518 for 3,141 counties (STOPT3141). Table 8 shows an excerpt from the output of STOPT3141. (A larger excerpt showing more counties is included in Appendix D.) The ten largest counties are included, with other selected counties representing every ten percent decrease in individual probability of detection. There is a positive effect for the largest counties in terms of probability of detection, but clearly a negligible benefit for smaller counties as the thresholds set for counties #1325 (in rank by population) and smaller are more than double the magnitude of the shift for which we are monitoring. In fact, 4σ and higher thresholds are assigned beginning at county #884. Even following a 2σ shift, Jackson County, Ohio (#1325), will be in the right-hand tail region of the out-of-control distribution, 2.34σ beyond the new mean $\mu_1 = 2$.

Figure 7 is a map showing all 3,141 counties, with shading to indicate their relative probabilities of event (here, "probability of attack" in the bioterrorism context) and their assigned optimal thresholds h_i. An example of a system of distributed statistical sensors on the same magnitude as this large model is the BioSense system operated by the Centers for Disease Control. This implementation of national biological surveillance has over 1600 sensors distributed throughout the United States. Figure 8 shows sensor locations for BioSense-equipped civilian hospitals and Department of Defense and Veterans Administration medical facilities superimposed on a population density map of the United States. (CDC)

However, the smaller counties receive very little coverage because of their small populations. We see from STOPT3141 that there is less than a one percent probability of detection of an event of magnitude 2σ from county #1325 through #3141. For county #368 and smaller we see less than a ten percent probability of detection of a 2σ event. All other counties between #1 and #200 that are presented can be compared directly to the results from the two hundred counties model using Table 9.

i	*County, State*	m_i	m_i/M	h_i	$p(d_i)$	$p(\alpha_i)$
1	Los Angeles County, CA	9,948,081	0.0332	0.45	0.939	0.326
2	Cook County, IL	5,288,655	0.0177	0.88	0.868	0.188
3	Harris County, TX	3,886,207	0.0130	1.09	0.817	0.137
4	Maricopa County, AZ	3,768,123	0.0126	1.12	0.812	0.132
5	Orange County, CA	3,002,048	0.0100	1.27	0.767	0.102
6	San Diego County, CA	2,941,454	0.0098	1.28	0.763	0.099
7	Kings County, NY	2,508,820	0.0084	1.39	0.728	0.082
8	Miami-Dade County, FL	2,402,208	0.0080	1.42	0.718	0.077
9	Dallas County, TX	2,345,815	0.0078	1.44	0.712	0.075
10	Queens County, NY	2,255,175	0.0075	1.47	0.703	0.071
21	Suffolk County, NY	1,469,715	0.0049	1.76	0.595	0.039
36	Contra Costa Co., CA	1,024,319	0.0034	2.01	0.498	0.022
77	Lake County, IL	713,076	0.0024	2.25	0.400	0.012
132	Chester County, PA	482,112	0.0016	2.52	0.301	0.006
200	Lorain County, OH	301,993	0.0010	2.84	0.200	0.002
368	Brazos County, TX	159,006	0.0005	3.28	0.100	0.001
1325	Jackson County, OH	33,543	0.0001	4.34	0.010	0.000

Table 8: Excerpt from model setting thresholds for 3,141 counties (STOPT3141). Overall probability of detection 0.515. Assumes a 2σ shift and false alarm limit $\kappa = 4$.

Because we have previously shown that adding more sensors reduces detection performance of the most populous counties, we now explore what happens if we shrink this system back down to the point where all sensors have at least a ten percent probability of detection. If we remove all counties below #368 and re-solve the problem (STOPT368) to keep within our false alarm constraint, the probability of detection increases slightly from 0.518 to 0.523 for $\kappa = 4.000$ false alarms. The overall probability of detection increases by 0.5 percentage points and the individual probabilities of detection for the largest counties increase, while we can presume that there are significant cost savings by not installing or operating these 2,700-plus sensors.

Thus far in this section we have treated each county's share m_i/M of the total population M as a probability of experiencing a biological event of interest, either a terrorist attack or a natural disease outbreak. By optimizing using the model STOPT in Equation (5) we achieve a set of optimal thresholds h_i that achieve a relatively high probability of detection (more than half) for the magnitude shift we have assumed throughout the use of population models in this thesis. However, some might insist on using the same detection threshold in each county, in order to give each county an equal probability of detection of the expected shift.

		Threshold h_i		$p(d_i)$		$p(\alpha_i)$	
i	*County, State*	3,141	200	3,141	200	3,141	200
1	Los Angeles County, CA	0.45	0.81	0.939	0.883	0.326	0.209
2	Cook County, IL	0.88	1.13	0.868	0.809	0.188	0.130
3	Harris County, TX	1.09	1.28	0.817	0.764	0.137	0.100
4	Maricopa County, AZ	1.12	1.30	0.812	0.759	0.132	0.097
5	Orange County, CA	1.27	1.41	0.767	0.722	0.102	0.079
6	San Diego County, CA	1.28	1.42	0.763	0.719	0.099	0.078
7	Kings County, NY	1.39	1.50	0.728	0.691	0.082	0.067
8	Miami-Dade County, FL	1.42	1.52	0.718	0.684	0.077	0.064
9	Dallas County, TX	1.44	1.53	0.712	0.679	0.075	0.063
10	Queens County, NY	1.47	1.55	0.703	0.672	0.071	0.060
21	Suffolk County, NY	1.76	1.77	0.595	0.592	0.039	0.039
36	Contra Costa Co., CA	2.01	1.95	0.498	0.521	0.022	0.026
77	Lake County, IL	2.25	2.13	0.400	0.449	0.012	0.017
132	Chester County, PA	2.52	2.33	0.301	0.373	0.006	0.010
200	Lorain County, OH	2.84	2.56	0.200	0.288	0.002	0.005

Table 9: Comparison of results for 3,141 counties vs. 200 counties (Selected counties). Assumes a 2σ shift and limit on false alarms $\kappa = 4.000$.

If we take this approach of setting equal thresholds for all 3,141 counties in an attempt to detect a 2σ shift somewhere in the country, the performance of our system suffers greatly. All thresholds will be set at $h = 3.02$ to achieve a false alarm rate of four per time period. Note that this threshold is fifty percent above the expected magnitude of the shift. The false alarm rate will be less than one percent (0.001) in all locations. The system-side probability of detection will only be 0.154, equal to the probabilities of each of the counties.

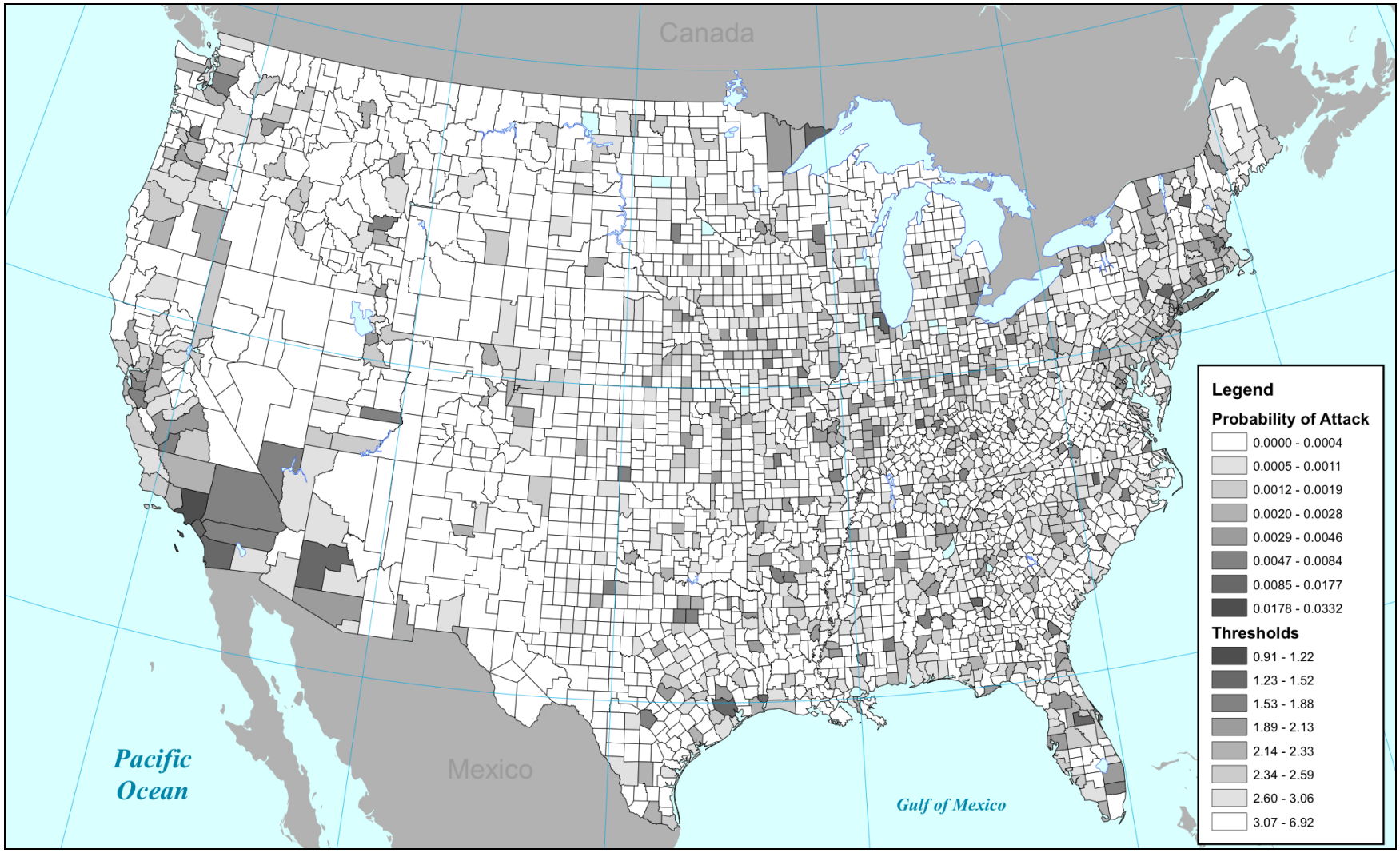

Figure 7: Output from optimizing 3,141 counties (STOPT3141).

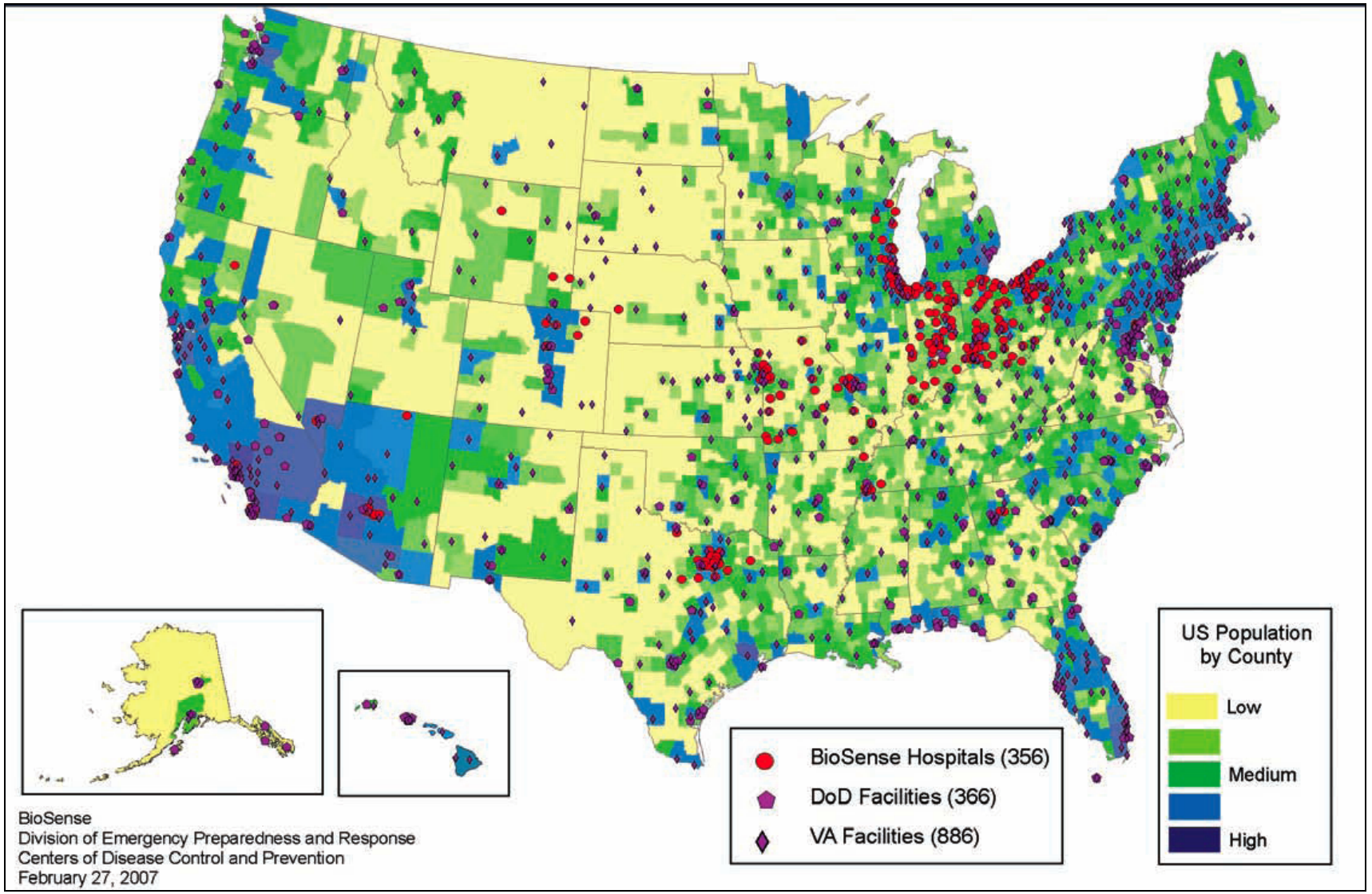

Figure 8: Centers for Disease Control (CDC) map of BioSense sensor locations superimposed on county population density.

IV. CONCLUSIONS AND FUTURE RESEARCH

A. CONCLUSIONS

In this thesis we have presented a model using mathematical optimization tools and software to improve the performance of a system of threshold sensors, given an existing set of sensors, an expected minimum expected magnitude of shift in the distributional mean at a single location, and some constraint on the acceptable number or rate of false alarms. We have formally presented an example that demonstrates that a system is not necessarily improved by adding more sensors. In fact, doing so will cause negative effects on the overall system's probability of detection, holding all else constant, as well as on the detection performance in those locations that already have sensor coverage.

The implications for policymakers are clear. It is not sufficient to simply add sensors to an existing sensor network in an informal fashion and expect improved performance. While additional coverage may be gained, to the benefit of those areas receiving new sensors, the overall detection performance of the system may be degraded. Additionally, to meet the defined limit on the expected number of false alarms, the individual thresholds of all sensors must be adjusted. Adding one new sensor without adjusting thresholds will cause an increase in the system-wide false alarm rate; adding many sensors will cause significant increases in the system-wide false alarm rate.

B. AREAS FOR FUTURE RESEARCH

This research has focused on the problem of simultaneously setting detection thresholds in multiple locations. We have assumed away the problem of adjusting thresholds for temporal factors such as day-of-week and seasonal effects, which have been studied extensively in other applications of SPC to detection problems involving single sensors. Future work should develop optimization models for detection systems based on the CUSUM and EWMA techniques, which can be more sensitive to detecting small shifts.

There are many areas where our models can be applied. Any system of distributed sensors that monitors for an increase in a statistic over a network with a large number of sensors should benefit from this work. These include many of the previous non-industrial applications of statistical process control discussed in Chapter I, including service performance, finance, health care monitoring, outbreak detection, water quality monitoring, and maintenance. There may even be applications for industry—despite having production quality measurements based on engineering requirements, it may be desirable to set optimal warning or control limits on machines and processes that are based on criticality, likelihood to drift to an out-of-control state, or to experience other types of failures.

In our discussion, we have made the assumption of one independent sensor per sub-region. There are practical issues that arise from such an assumption, namely the adequacy of coverage for a large sub-region. It is unlikely to be the case that areas farther away from the sensor will receive equal coverage as those near the sensor itself. For example, a sensor covering a single county may provide good coverage to the city in which it is located, but rural areas far away from that city are not likely to receive equal coverage. This has a practical application to epidemiologic surveillance, in that it is possible for an outbreak to start in a rural area and spread to the city. If the increase in cases can be detected before the spread to the city occurs, steps can be taken to mitigate the effects of the outbreak.

This suggests that regions are likely to employ multiple sensors and an area for future study is how to integrate multiple sensors providing coverage for a single area. How should data from multiple sensors in a sub-region be compiled and analyzed? We suggest two options for future consideration. The first uses sensor fusion to compile data received at multiple sensors within a sub-region; the data from these sensors are then treated in aggregate as though received at a single sensor, with a threshold assigned by the solution to the optimization problem for the entire region. The second option is to solve the large regional problem first using our optimization model; subsequently we apply this model to solve a subordinate optimization problem for the multiple sensors

within a given sub-region to set thresholds and allocate the expected number of false positives according to the weights assigned or populations covered by each sensor.

A final statement should be made as to the adoption of optimization techniques to improve the performance of a system. As in previous efforts to use statistical process control for monitoring non-industrial processes, it is important to systematically implement such a system and to trust the results that the model provides. Training must be provided throughout the organization to ensure that the process itself and the benefits derived from it are understood and appreciated by all members. We previously referred to the “ad-hoc” nature of sensor implementation in the biosurveillance context; such lack of strategy often leads to piecemeal implementation, which can result in sub-optimal system performance and even complete failure due to loss of confidence in the system resulting from problems such as rampant false signaling.

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX A. GAMS CODE FOR SAMPLE PROBLEM

10 fictional sub-regions, based on an $X \sim N(0,1)$ in-control distribution, a 2σ shift in the mean incidence level, and false alarm limit $\kappa = 1$.

```
OPTIONS
  NLP=MINOS
  optca = 0
  optcr = 0;

SET
  i        Sub-region index/County01*County10/;

SCALARS
  K        Max allowed false alarms in region /1/
  delta    Magnitude of shift in mean level  /2/
  total    Total weight of region;

PARAMETERS
  m(i)     Relative weight of sub-region i

/
County01   1000000
County02   800000
County03   700000
County04   600000
County05   600000
County06   500000
County07   500000
County08   400000
County09   400000
County10   400000
/
;

PARAMETER totalWeight;
     totalWeight=sum(i,m(i));

PARAMETER pa(i);
     pa(i) = m(i)/totalWeight;

VARIABLE h(i)  Detection threshold in sub-region i;
```

```
VARIABLE PD    Overall probability of detection given event
occurs in region;

EQUATIONS
  OBJ          Overall probability of detection given event
occurs in region
  FA           Overall number of false alarms must be less
than K
;

  OBJ..    PD=E=sum(i,pa(i)*(1-errorf(h(i)-delta)));
  FA..     sum(i,1-errorf(h(i)))=L=K;

MODEL Regions/ALL/;

SOLVE Regions USING NLP MAXIMIZING PD
```

APPENDIX B. DERIVATION OF ONE-DIMENSIONAL OPTIMIZATION MODEL

We have assumed that our in-control distribution F_0 is a standard normal distribution, $F_0 \sim N(0,1)$, and that the signal manifests itself as a shift in the mean to a new distribution $F_1 \sim N(\delta,1)$, $\delta > 0$. Using the log-linear relationship derived in this section, the model of Equation (5) can be expressed as a one-dimensional optimization problem, which is much easier and faster to solve than the *n*-dimensional problems solved previously.

By observing that maximizing the area to the right of h_i on a normal distribution is equivalent to minimizing the area to the left of h_i, and that our optimal solution is located on the constraint boundary (see Fricker and Banschbach, 2008), we can re-express the maximization problem in (5) in the following way:

$$\begin{aligned} &\underset{\bar{h}}{\text{minimize}} \sum_{i=1}^{n} \Phi(h_i - \delta) p_i \\ &\text{subject to} \sum_{i=1}^{n} \Phi(h_i) = n - \kappa \end{aligned} \tag{8}$$

Working on the constraint first, we separate the first term of the summation and re-express this constraint as

$$\Phi(h_1) + \sum_{i=2}^{n} \Phi(h_i) = n - \kappa \tag{9}$$

Solving (9) for h_1 we get

$$h_1 = \Phi^{-1}\left(n - \kappa - \sum_{i=2}^{n} \Phi(h_i) \right) \tag{10}$$

Following a similar pattern for the objective function we can separate the first term of the summation h_1 and express (8) as

$$\min_{\bar{h}} \; f = \Phi[h_1 - \delta] p_1 + \sum_{i=2}^{n} \Phi(h_i - \delta) p_i \tag{11}$$

Substituting the expression for h_1 (10) into the objective function (11) gives us an unconstrained optimization problem:

$$\min_{\tilde{h}} f = \Phi\left[\Phi^{-1}\left(n-\kappa-\sum_{i=2}^{n}\Phi(h_i)\right)-\delta\right]+\sum_{i=2}^{n}\Phi(h_i-\delta)p_i \tag{12}$$

To find the set of h_i's at which f is minimized, we first differentiate (12) with respect to each threshold h_i and set each equation equal to zero. We then simultaneously solve the set of equations for each h_i. The solution for $n = 2$ is as follows. After solving for h_1 in the constraint and substituting into the now-unconstrained objective function, the partial derivative with respect to h_2 is

$$\frac{\partial f}{\partial h_2} = \frac{\exp\left(-\frac{{h_2}^2+\delta^2}{2}\right)\left(p_2\exp[h_2\delta]-p_1\exp\left[\sqrt{2}\delta Erf^{-1}\left\{2-2\kappa-Erf\left[\frac{h_2}{\sqrt{2}}\right]\right\}\right]\right)}{\sqrt{2\pi}} \tag{13}$$

We can easily see that (13) will equal zero only if

$$p_2\exp[h_2\delta] = p_1\exp\left[\sqrt{2}\delta Erf^{-1}\left\{2-2\kappa-Erf\left[\frac{h_2}{\sqrt{2}}\right]\right\}\right] \tag{14}$$

Taking the natural log of both sides of (14) and simplifying yields

$$Erf\left[\frac{h_2+\frac{1}{\delta}\left(\ln(p_2)-\ln(p_1)\right)}{\sqrt{2}}\right] = 2-2\kappa-Erf\left(\frac{h_2}{\sqrt{2}}\right) \tag{15}$$

We know that $Erf\left(z/\sqrt{2}\right) = 2\Phi(z)-1$, so we can further simplify (15) to

$$\Phi\left(h_2+\frac{1}{\delta}\ln(p_2)-\frac{1}{\delta}\ln(p_1)\right)+\Phi(h_2) = 2-\kappa \tag{16}$$

Making the substitution $\mu = h_2+\frac{1}{\delta}\ln(p_2)$, we get

$$\sum_{i=1}^{2}\Phi\left[\mu-\frac{1}{\delta}\ln(p_i)\right] = 2-\kappa \tag{17}$$

The expression in (17) is equivalent to

$$\sum_{i=1}^{2} \Pr\left(Z \le \mu - \frac{1}{\delta} \ln(p_i) \right) = 2 - \kappa \tag{18}$$

Here Z is a standard normal random variate. Using the expression in (18), we see that the constrained minimization problem in (8) reduces to finding μ such that the sum of the left tails of n normal distributions is equal to $n - \kappa$, where $0 \le n - \kappa \le n$.

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX C. EXCEL RESULTS FOR STOPT200

Terms: m_i, Population in county i; m_i/M, county share m_i of total population M; h_i, threshold level set in county i; $p(d_i)$, probability of detection in county i; $p(\alpha_i)$, probability of detection in county i.

Based on an $X_i \sim N(0,1)$ in-control distribution and a 2σ shift in the mean incidence level. Overall $P_d = 0.537$, E(*false alarms*) = 4.

i	*County, State*	m_i	m_i/M	h_i	$p(d_i)$	$p(\alpha_i)$
1	Los Angeles County, CA	9,948,081	0.0587	0.812	0.883	0.209
2	Cook County, IL	5,288,655	0.0312	1.127	0.809	0.130
3	Harris County, TX	3,886,207	0.0229	1.282	0.764	0.100
4	Maricopa County, AZ	3,768,123	0.0222	1.297	0.759	0.097
5	Orange County, CA	3,002,048	0.0177	1.411	0.722	0.079
6	San Diego County, CA	2,941,454	0.0174	1.421	0.719	0.078
7	Kings County, NY	2,508,820	0.0148	1.500	0.691	0.067
8	Miami-Dade County, FL	2,402,208	0.0142	1.522	0.684	0.064
9	Dallas County, TX	2,345,815	0.0138	1.534	0.679	0.063
10	Queens County, NY	2,255,175	0.0133	1.553	0.672	0.060
11	Riverside County, CA	2,026,803	0.0120	1.607	0.653	0.054
12	San Bernardino County, CA	1,999,332	0.0118	1.614	0.650	0.053
13	Wayne County, MI	1,971,853	0.0116	1.621	0.648	0.052
14	King County, WA	1,826,732	0.0108	1.659	0.633	0.049
15	Broward County, FL	1,787,636	0.0105	1.670	0.629	0.047
16	Clark County, NV	1,777,539	0.0105	1.673	0.628	0.047
17	Santa Clara County, CA	1,731,281	0.0102	1.686	0.623	0.046
18	Tarrant County, TX	1,671,295	0.0099	1.703	0.617	0.044
19	New York County, NY	1,611,581	0.0095	1.721	0.610	0.043
20	Bexar County, TX	1,555,592	0.0092	1.739	0.603	0.041
21	Suffolk County, NY	1,469,715	0.0087	1.767	0.592	0.039
22	Middlesex County, MA	1,467,016	0.0087	1.768	0.592	0.039
23	Alameda County, CA	1,457,426	0.0086	1.771	0.590	0.038
24	Philadelphia County, PA	1,448,394	0.0085	1.775	0.589	0.038
25	Sacramento County, CA	1,374,724	0.0081	1.801	0.579	0.036
26	Bronx County, NY	1,361,473	0.0080	1.806	0.577	0.035
27	Nassau County, NY	1,325,662	0.0078	1.819	0.572	0.034
28	Cuyahoga County, OH	1,314,241	0.0078	1.824	0.570	0.034
29	Palm Beach County, FL	1,274,013	0.0075	1.839	0.564	0.033
30	Allegheny County, PA	1,223,411	0.0072	1.860	0.556	0.031
31	Oakland County, MI	1,214,255	0.0072	1.864	0.554	0.031

i	*County, State*	m_i	m_i/M	h_i	$p(d_i)$	$p(\alpha_i)$
32	Hillsborough County, FL	1,157,738	0.0068	1.888	0.545	0.030
33	Hennepin County, MN	1,122,093	0.0066	1.903	0.539	0.029
34	Franklin County, OH	1,095,662	0.0065	1.915	0.534	0.028
35	Orange County, FL	1,043,500	0.0062	1.939	0.524	0.026
36	Contra Costa County, CA	1,024,319	0.0060	1.948	0.521	0.026
37	Fairfax County, VA	1,010,443	0.0060	1.955	0.518	0.025
38	St. Louis County, MO	1,000,510	0.0059	1.960	0.516	0.025
39	Salt Lake County, UT	978,701	0.0058	1.971	0.512	0.024
40	Fulton County, GA	960,009	0.0057	1.980	0.508	0.024
41	Westchester County, NY	949,355	0.0056	1.986	0.506	0.024
42	Pima County, AZ	946,362	0.0056	1.988	0.505	0.023
43	DuPage County, IL	932,670	0.0055	1.995	0.502	0.023
44	Montgomery County, MD	932,131	0.0055	1.995	0.502	0.023
45	Pinellas County, FL	924,413	0.0055	1.999	0.500	0.023
46	Erie County, NY	921,390	0.0054	2.001	0.500	0.023
47	Travis County, TX	921,006	0.0054	2.001	0.500	0.023
48	Milwaukee County, WI	915,097	0.0054	2.004	0.498	0.023
49	Shelby County, TN	911,438	0.0054	2.006	0.497	0.022
50	Honolulu County, HI	909,863	0.0054	2.007	0.497	0.022
51	Bergen County, NJ	904,037	0.0053	2.010	0.496	0.022
52	Fairfield County, CT	900,440	0.0053	2.012	0.495	0.022
53	Fresno County, CA	891,756	0.0053	2.017	0.493	0.022
54	Hartford County, CT	876,927	0.0052	2.026	0.490	0.021
55	Marion County, IN	865,504	0.0051	2.032	0.487	0.021
56	New Haven County, CT	845,244	0.0050	2.044	0.482	0.020
57	Prince George's County, MD	841,315	0.0050	2.046	0.482	0.020
58	Duval County, FL	837,964	0.0049	2.048	0.481	0.020
59	Macomb County, MI	832,861	0.0049	2.051	0.479	0.020
60	Mecklenburg County, NC	827,445	0.0049	2.055	0.478	0.020
61	Hamilton County, OH	822,596	0.0049	2.058	0.477	0.020
62	Ventura County, CA	799,720	0.0047	2.072	0.471	0.019
63	Baltimore County, MD	787,384	0.0046	2.080	0.468	0.019
64	Middlesex County, NJ	786,971	0.0046	2.080	0.468	0.019
65	Wake County, NC	786,522	0.0046	2.080	0.468	0.019
66	Essex County, NJ	786,147	0.0046	2.080	0.468	0.019
67	Worcester County, MA	784,992	0.0046	2.081	0.468	0.019
68	Kern County, CA	780,117	0.0046	2.084	0.466	0.019
69	Montgomery County, PA	775,688	0.0046	2.087	0.465	0.018
70	Pierce County, WA	766,878	0.0045	2.093	0.463	0.018
71	Gwinnett County, GA	757,104	0.0045	2.099	0.460	0.018
72	San Francisco County, CA	744,041	0.0044	2.108	0.457	0.018
73	El Paso County, TX	736,310	0.0043	2.113	0.455	0.017

i	*County, State*	m_i	m_i/M	h_i	$p(d_i)$	$p(\alpha_i)$
74	Essex County, MA	735,958	0.0043	2.114	0.455	0.017
75	Monroe County, NY	730,807	0.0043	2.117	0.453	0.017
76	DeKalb County, GA	723,602	0.0043	2.122	0.451	0.017
77	Lake County, IL	713,076	0.0042	2.129	0.449	0.017
78	San Mateo County, CA	705,499	0.0042	2.135	0.446	0.016
79	Jefferson County, KY	701,500	0.0041	2.138	0.445	0.016
80	Hidalgo County, TX	700,634	0.0041	2.138	0.445	0.016
81	Collin County, TX	698,851	0.0041	2.140	0.445	0.016
82	Oklahoma County, OK	691,266	0.0041	2.145	0.442	0.016
83	Suffolk County, MA	687,610	0.0041	2.148	0.441	0.016
84	Multnomah County, OR	681,454	0.0040	2.152	0.440	0.016
85	Cobb County, GA	679,325	0.0040	2.154	0.439	0.016
86	San Joaquin County, CA	673,170	0.0040	2.158	0.437	0.015
87	Snohomish County, WA	669,887	0.0040	2.161	0.436	0.015
88	Will County, IL	668,217	0.0039	2.162	0.436	0.015
89	Jackson County, MO	664,078	0.0039	2.165	0.434	0.015
90	Jefferson County, AL	656,700	0.0039	2.171	0.432	0.015
91	Norfolk County, MA	654,753	0.0039	2.172	0.432	0.015
92	Providence County, RI	635,596	0.0037	2.187	0.426	0.014
93	Monmouth County, NJ	635,285	0.0037	2.187	0.426	0.014
94	Baltimore city, MD	631,366	0.0037	2.190	0.425	0.014
95	Bucks County, PA	623,205	0.0037	2.197	0.422	0.014
96	Bernalillo County, NM	615,099	0.0036	2.203	0.419	0.014
97	Hudson County, NJ	601,146	0.0035	2.215	0.415	0.013
98	Kent County, MI	599,524	0.0035	2.216	0.414	0.013
99	Denton County, TX	584,238	0.0034	2.229	0.409	0.013
100	District of Columbia, DC	581,530	0.0034	2.231	0.409	0.013
101	Davidson County, TN	578,698	0.0034	2.234	0.408	0.013
102	Tulsa County, OK	577,795	0.0034	2.234	0.407	0.013
103	El Paso County, CO	576,884	0.0034	2.235	0.407	0.013
104	Lee County, FL	571,344	0.0034	2.240	0.405	0.013
105	Denver County, CO	566,974	0.0033	2.244	0.404	0.012
106	Ocean County, NJ	562,335	0.0033	2.248	0.402	0.012
107	Polk County, FL	561,606	0.0033	2.249	0.402	0.012
108	Delaware County, PA	555,996	0.0033	2.254	0.400	0.012
109	Summit County, OH	545,931	0.0032	2.263	0.396	0.012
110	Bristol County, MA	545,379	0.0032	2.263	0.396	0.012
111	Montgomery County, OH	542,237	0.0032	2.266	0.395	0.012
112	Arapahoe County, CO	537,197	0.0032	2.271	0.393	0.012
113	Brevard County, FL	534,359	0.0032	2.273	0.392	0.012
114	Union County, NJ	531,088	0.0031	2.276	0.391	0.011
115	Jefferson County, CO	526,994	0.0031	2.280	0.390	0.011
116	New Castle County, DE	525,587	0.0031	2.282	0.389	0.011

i	*County, State*	m_i	m_i/M	h_i	$p(d_i)$	$p(\alpha_i)$
117	Camden County, NJ	517,001	0.0030	2.290	0.386	0.011
118	Johnson County, KS	516,731	0.0030	2.290	0.386	0.011
119	Washington County, OR	514,269	0.0030	2.292	0.385	0.011
120	Stanislaus County, CA	512,138	0.0030	2.294	0.384	0.011
121	Anne Arundel County, MD	509,300	0.0030	2.297	0.383	0.011
122	Passaic County, NJ	497,093	0.0029	2.309	0.379	0.010
123	Volusia County, FL	496,575	0.0029	2.310	0.378	0.010
124	Lancaster County, PA	494,486	0.0029	2.312	0.378	0.010
125	Lake County, IN	494,202	0.0029	2.312	0.377	0.010
126	Kane County, IL	493,735	0.0029	2.313	0.377	0.010
127	Plymouth County, MA	493,623	0.0029	2.313	0.377	0.010
128	Ramsey County, MN	493,215	0.0029	2.313	0.377	0.010
129	Fort Bend County, TX	493,187	0.0029	2.313	0.377	0.010
130	Morris County, NJ	493,160	0.0029	2.313	0.377	0.010
131	Douglas County, NE	492,003	0.0029	2.314	0.377	0.010
132	Chester County, PA	482,112	0.0028	2.325	0.373	0.010
133	Richmond County, NY	477,377	0.0028	2.330	0.371	0.010
134	Sedgwick County, KS	470,895	0.0028	2.336	0.368	0.010
135	Sonoma County, CA	466,891	0.0028	2.341	0.367	0.010
136	Utah County, UT	464,760	0.0027	2.343	0.366	0.010
137	Dane County, WI	463,826	0.0027	2.344	0.365	0.010
138	Hampden County, MA	460,520	0.0027	2.348	0.364	0.009
139	Onondaga County, NY	456,777	0.0027	2.352	0.363	0.009
140	Guilford County, NC	451,905	0.0027	2.357	0.361	0.009
141	Burlington County, NJ	450,627	0.0027	2.358	0.360	0.009
142	Pasco County, FL	450,171	0.0027	2.359	0.360	0.009
143	Spokane County, WA	446,706	0.0026	2.363	0.358	0.009
144	Lucas County, OH	445,281	0.0026	2.364	0.358	0.009
145	Genesee County, MI	441,966	0.0026	2.368	0.356	0.009
146	Virginia Beach city, VA	435,619	0.0026	2.376	0.354	0.009
147	Jefferson Parish, LA	431,361	0.0025	2.380	0.352	0.009
148	East Baton Rouge Parish, LA	429,073	0.0025	2.383	0.351	0.009
149	Tulare County, CA	419,909	0.0025	2.394	0.347	0.008
150	Greenville County, SC	417,166	0.0025	2.397	0.346	0.008
151	York County, PA	416,322	0.0025	2.398	0.345	0.008
152	Adams County, CO	414,338	0.0024	2.401	0.344	0.008
153	Clark County, WA	412,938	0.0024	2.403	0.344	0.008
154	Knox County, TN	411,967	0.0024	2.404	0.343	0.008
155	Solano County, CA	411,680	0.0024	2.404	0.343	0.008
156	Monterey County, CA	410,206	0.0024	2.406	0.342	0.008
157	Polk County, IA	408,888	0.0024	2.407	0.342	0.008
158	Seminole County, FL	406,875	0.0024	2.410	0.341	0.008
159	Mobile County, AL	404,157	0.0024	2.413	0.340	0.008

i	*County, State*	m_i	m_i/M	h_i	$p(d_i)$	$p(\alpha_i)$
160	Hillsborough County, NH	402,789	0.0024	2.415	0.339	0.008
161	Berks County, PA	401,149	0.0024	2.417	0.338	0.008
162	Santa Barbara County, CA	400,335	0.0024	2.418	0.338	0.008
163	Montgomery County, TX	398,290	0.0023	2.421	0.337	0.008
164	Washoe County, NV	396,428	0.0023	2.423	0.336	0.008
165	Dakota County, MN	388,001	0.0023	2.434	0.332	0.007
166	Cameron County, TX	387,717	0.0023	2.434	0.332	0.007
167	Waukesha County, WI	380,985	0.0022	2.443	0.329	0.007
168	Stark County, OH	380,575	0.0022	2.444	0.329	0.007
169	Orange County, NY	376,392	0.0022	2.449	0.327	0.007
170	Clackamas County, OR	374,230	0.0022	2.452	0.326	0.007
171	Sarasota County, FL	369,535	0.0022	2.459	0.323	0.007
172	Mercer County, NJ	367,605	0.0022	2.461	0.322	0.007
173	Pulaski County, AR	367,319	0.0022	2.462	0.322	0.007
174	Westmoreland County, PA	366,440	0.0022	2.463	0.322	0.007
175	Ada County, ID	359,035	0.0021	2.473	0.318	0.007
176	Prince William County, VA	357,503	0.0021	2.475	0.317	0.007
177	Butler County, OH	354,992	0.0021	2.479	0.316	0.007
178	Williamson County, TX	353,830	0.0021	2.481	0.315	0.007
179	Richland County, SC	348,226	0.0021	2.489	0.313	0.006
180	Allen County, IN	347,316	0.0020	2.490	0.312	0.006
181	St. Louis city, MO	347,181	0.0020	2.490	0.312	0.006
182	Washtenaw County, MI	344,047	0.0020	2.495	0.310	0.006
183	St. Charles County, MO	338,719	0.0020	2.503	0.308	0.006
184	Lane County, OR	337,870	0.0020	2.504	0.307	0.006
185	Lehigh County, PA	335,544	0.0020	2.507	0.306	0.006
186	Forsyth County, NC	332,355	0.0020	2.512	0.304	0.006
187	Charleston County, SC	331,917	0.0020	2.513	0.304	0.006
188	Anoka County, MN	327,005	0.0019	2.520	0.301	0.006
189	Placer County, CA	326,242	0.0019	2.521	0.301	0.006
190	Somerset County, NJ	324,186	0.0019	2.524	0.300	0.006
191	Nueces County, TX	321,457	0.0019	2.529	0.299	0.006
192	Marion County, FL	316,183	0.0019	2.537	0.296	0.006
193	Collier County, FL	314,649	0.0019	2.539	0.295	0.006
194	Manatee County, FL	313,298	0.0018	2.542	0.294	0.006
195	Luzerne County, PA	313,020	0.0018	2.542	0.294	0.006
196	Hamilton County, TN	312,905	0.0018	2.542	0.294	0.006
197	McHenry County, IL	312,373	0.0018	2.543	0.294	0.005
198	Marion County, OR	311,304	0.0018	2.545	0.293	0.005
199	Madison County, AL	304,307	0.0018	2.556	0.289	0.005
200	Lorain County, OH	301,993	0.0018	2.560	0.288	0.005

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX D. SELECTED EXCEL RESULTS FOR STOPT3141

Terms: m_i, Population in county i; m_i/M, county share m_i of total population M; h_i, threshold level set in county i; $p(d_i)$, probability of detection in county i; $p(\alpha_i)$, probability of detection in county i.

Based on an $X_i \sim N(0,1)$ in-control distribution and a 2σ shift in the mean incidence level. Overall $P_d = 0.515$, E(*false alarms*) = 4.

i	*County, State*	m_i	m_i/M	h_i	$p(d_i)$	$p(\alpha_i)$
1	Los Angeles County, CA	9,948,081	0.0587	0.452	0.939	0.326
2	Cook County, IL	5,288,655	0.0312	0.884	0.868	0.188
3	Harris County, TX	3,886,207	0.0130	1.094	0.817	0.137
4	Maricopa County, AZ	3,768,123	0.0126	1.115	0.812	0.132
5	Orange County, CA	3,002,048	0.0100	1.271	0.767	0.102
6	San Diego County, CA	2,941,454	0.0098	1.285	0.763	0.099
7	Kings County, NY	2,508,820	0.0084	1.394	0.728	0.082
8	Miami-Dade County, FL	2,402,208	0.0080	1.423	0.718	0.077
9	Dallas County, TX	2,345,815	0.0078	1.440	0.712	0.075
10	Queens County, NY	2,255,175	0.0075	1.466	0.703	0.071
11	Riverside County, CA	2,026,803	0.0068	1.540	0.677	0.062
12	San Bernardino County, CA	1,999,332	0.0067	1.549	0.674	0.061
13	Wayne County, MI	1,971,853	0.0066	1.558	0.671	0.060
14	King County, WA	1,826,732	0.0061	1.611	0.652	0.054
15	Broward County, FL	1,787,636	0.0060	1.625	0.646	0.052
16	Clark County, NV	1,777,539	0.0059	1.629	0.645	0.052
17	Santa Clara County, CA	1,731,281	0.0058	1.647	0.638	0.050
18	Tarrant County, TX	1,671,295	0.0056	1.671	0.629	0.047
19	New York County, NY	1,611,581	0.0054	1.696	0.619	0.045
20	Bexar County, TX	1,555,592	0.0052	1.720	0.610	0.043
21	Suffolk County, NY	1,469,715	0.0049	1.759	0.595	0.039
22	Middlesex County, MA	1,467,016	0.0049	1.761	0.595	0.039
23	Alameda County, CA	1,457,426	0.0049	1.765	0.593	0.039
24	Philadelphia County, PA	1,448,394	0.0048	1.769	0.591	0.038
25	Sacramento County, CA	1,374,724	0.0046	1.805	0.577	0.036
26	Bronx County, NY	1,361,473	0.0045	1.812	0.575	0.035
27	Nassau County, NY	1,325,662	0.0044	1.830	0.568	0.034
28	Cuyahoga County, OH	1,314,241	0.0044	1.836	0.565	0.033
29	Palm Beach County, FL	1,274,013	0.0043	1.857	0.557	0.032
30	Allegheny County, PA	1,223,411	0.0041	1.885	0.546	0.030

i	*County, State*	m_i	m_i/M	h_i	$p(d_i)$	$p(\alpha_i)$
31	Oakland County, MI	1,214,255	0.0041	1.890	0.544	0.029
32	Hillsborough County, FL	1,157,738	0.0039	1.922	0.531	0.027
33	Hennepin County, MN	1,122,093	0.0037	1.944	0.522	0.026
34	Franklin County, OH	1,095,662	0.0037	1.960	0.516	0.025
35	Orange County, FL	1,043,500	0.0035	1.993	0.503	0.023
36	Contra Costa County, CA	1,024,319	0.0034	2.006	0.498	0.022
37	Fairfax County, VA	1,010,443	0.0034	2.015	0.494	0.022
38	St. Louis County, MO	1,000,510	0.0033	2.022	0.491	0.022
39	Salt Lake County, UT	978,701	0.0033	2.037	0.485	0.021
40	Fulton County, GA	960,009	0.0032	2.050	0.480	0.020
	●●●●●●					
55	Marion County, IN	865,504	0.0029	2.121	0.452	0.017
77	Lake County, IL	713,076	0.0024	2.254	0.400	0.012
99	Denton County, TX	584,238	0.0020	2.390	0.348	0.008
105	Denver County, CO	566,974	0.0019	2.411	0.341	0.008
132	Chester County, PA	482,112	0.0016	2.521	0.301	0.006
165	Dakota County, MN	388,001	0.0013	2.670	0.251	0.004
185	Lehigh County, PA	335,544	0.0011	2.769	0.221	0.003
200	Lorain County, OH	301,993	0.0010	2.841	0.200	0.002
270	Alachua County, FL	227,120	0.0008	3.036	0.150	0.001
368	Brazos County, TX	159,006	0.0005	3.280	0.100	0.001
592	Cabell County, WV	93,904	0.0003	3.640	0.050	0.000
1325	Jackson County, OH	33,543	0.0001	4.344	0.010	0.000
2209	Amite County, MS	13,466	0.0000	4.968	0.001	0.000
2568	Conejos County, CO	8,406	0.0000	5.290	0.001	0.000
2569	Woods County, OK	8,385	0.0000	5.292	0.000	0.000
2800	Cameron County, PA	5,489	0.0000	5.582	0.000	0.000
3000	Graham County, KS	2,677	0.0000	6.073	0.000	0.000
3141	Loving County, TX	60	0.0000	8.670	0.000	0.000

LIST OF REFERENCES

Aylin, P., N. Best, A. Bottle, C. Marshall. 2003. Following Shipman: a pilot system for monitoring mortality rates in primary care. *Lancet* 362: 485-91.

Bamford, D.R., R.W. Greatbanks. 2005. The use of quality management tools and techniques: a study of application in everyday situations. *International Journal of Quality & Reliability Management* 22(4): 376-392.

Bazaraa, M.S., H.D. Sherali, C.M. Shetty. 2006. *Nonlinear Programming: Theory and Algorithms*. New York: Wiley-Interscience.

Bradley, C.A., H. Rolka, D. Walker, J. Loonsk. 2005. BioSense: Implementation of a National Early Event Detection and Situational Awareness System. *Morbidity and Mortality Weekly Report* 54(Suppl): 11-19.

Bravata, D.M., K.M. McDonald, W.M. Smith, C. Rydzak, H. Szeto, D.L. Buckeridge, C. Haberland, D.K. Owens. 2004. Systematic Review: Surveillance Systems for Early Detection of Bioterrorism-Related Diseases. *Annals of Internal Medicine* 140: 910-922.

Box, G., A. Luceño. 1997. *Statistical Control by Monitoring and Feedback Adjustment*. New York: John Wiley & Sons.

Centers for Disease Control and Prevention (CDC). URL www.cdc.gov. Accessed May 14, 2007.

Devore, J.L. 2004. *Probability and Statistics for Engineering and the Sciences*. Belmont, CA: Brooks/Cole.

Fricker, R.D. 2007a. Directionally Sensitive Multivariate Statistical Process Control Methods with Application to Syndromic Surveillance. *Advances in Disease Surveillance* 3(1): 1-17.

Fricker, R.D. 2007b. Syndromic Surveillance. *Encyclopedia for Quantitative Risk Assessment* (to appear).

Fricker, R.D., H.R. Rolka. 2006. Protecting Against Biological Terrorism: Statistical Issues in Electronic Biosurveillance. *Chance* 19: 4-13.

Fricker, R.D., B.L. Hegler, D.A. Dunfee. 2007a. Assessing the Performance of the Early Aberration Reporting System (EARS) Syndromic Surveillance Algorithms. *Statistics in Medicine* (to appear).

Fricker, R.D., M.C. Knitt, C.X. Hu. 2007b. Comparing Directionally Sensitive MCUSUM and MEWMA Procedures with Application to BioSurveillance. *Quality Engineering* (in submission).

Frontline Systems, Inc. URL www.solver.com. Accessed August 18, 2007.

GAMS Homepage. URL www.gams.com. Accessed September 25, 2007.

Gordon, L., M. Pollak. 1994. An Efficient Sequential Nonparametric Scheme for Detecting a Change of Distribution. *The Annals of Statistics* 22(2): 763-804.

Heffernan, R., F. Mostashari, D. Das, M. Besculides, C. Rodriguez, J. Greenko, L. Steiner-Sichel, S. Balter, A. Karpati, P. Thomas, M. Phillips, J. Ackelsberg, E. Lee, J. Leng, J. Hartman, K. Metzger, R. Rosselli, D. Weiss. 2004a. New York City Syndromic Surveillance Systems. *Morbidity and Mortality Weekly Report* 53(Suppl): 25-27.

Heffernan, R., F. Mostashari, D. Das, A. Karpati, M. Kulldorff, D. Weiss. 2004b. Syndromic surveillance in public health practice: New York City. *Emerging Infectious Diseases* 10: 25-27.

Hinde, J., C.G.B. Demétrio. 1998. Overdispersion: Models and estimation. *Computational Statistics & Data Analysis* 27: 151-170.

Hutwagner, L., W. Thompson, G.M. Seeman, T. Treadwell. 2003. The bioterrorism preparedness and response Early Aberration Reporting System (EARS). *Journal of Urban Health: Bulletin of the New York Academy of Medicine* 80 (No. 2, Supplement 1): 89i–96i.

Kleinman, K., R. Lazarus, R. Platt. 2003. A Generalized Linear Mixed Models Approach for Detecting Incident Clusters of Disease in Small Areas, with an Application to Biological Terrorism. *American Journal of Epidemiology* 3: 217.

Lasdon, L.S., A.D. Waren, A. Jain, M. Ratner. Design and Testing of a Generalized Reduced Gradient Code for Nonlinear Programming. *ACM Transactions on Mathematical Software* 4: 34-50.

Lober, W.B., B.T. Karras, M.M. Wagner, J.M. Overhage, A.J. Davidson, H. Fraser, L.J. Trigg, K.D. Mandl, J.U. Espino, F.C. Tsui. 2002. Roundtable on bioterrorism detection: information system-based surveillance. *Journal of the American Medical Informatics Association* 9: 105-115.

Loonsk, John W. 2004. BioSense — A National Initiative for Early Detection and Quantification of Public Health Emergencies. *Morbidity and Mortality Weekly Report* 53(Suppl): 53-55.

MacCarthy, B.L., T. Wasusri. 2002. A review of non-standard applications of statistical process control (SPC) charts. *International Journal of Quality & Reliability Management* 19(3): 295-320.

Marshall, C., N. Best, A. Bottle, P. Aylin. 2004. Statistical issues in the prospective monitoring of health outcomes across multiple units. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 167(3): 541–559.

Montgomery, D.C. 2001. *Introduction to Statistical Quality Control*. New York: John Wiley & Sons.

Montgomery, D.C., E.A. Peck, G.G. Vining. 2006. *Introduction to Linear Regression Analysis*. New York: Wiley.

Mohtashemi, M., P. Szolovits, J. Dunyak, K.D. Mandl. 2006. A susceptible-infected model of early detection of respiratory infection outbreaks on a background of influenza. *Journal of Theoretical Biology* 241(4): 954-963.

Pavlin, J.A., F. Mostashari, M.G. Kortepeter, N.A. Hynes, R.A. Chotani, Y.B. Mikol, M.A.K. Ryan, J.S. Neville, D.T. Gantz, J.V. Writer, J.E. Florance, R.C. Culpepper, F.M. Henretig, P.W. Kelley. 2003. Innovative surveillance methods for rapid detection of disease outbreaks and bioterrorism: results of an interagency workshop on health indicator surveillance. *American Journal of Public Health* 93:1230-1235.

Reis, B.Y., M. Pagano, K.D. Mandl. 2003. Using temporal context to improve biosurveillance. *Proceedings of the National Academy of Sciences* 100: 1961-1965.

Saniga, E.M. 1989. Economic Statistical Control-Chart Designs With an Application to $\overline{X}$ and R charts. *Technometrics* 31(3): 313-320.

Sebastiani, P., K. Mandl. 2004. Biosurveillance and Outbreak Detection. In *Data Mining: Next Generation Challenges and Future Directions*, ed. H. Kargupta, A. Joshi, K. Sivakumar, and Y. Yesha, 185-198. Boston: MIT Press.

Smith, G. 1998. *Statistical process control and quality improvement*. Upper Saddle River, New Jersey: Prentice-Hall, Inc.

Stoto, M.A., M. Schonlau, L.T. Mariano. 2004. Syndromic Surveillance: Is It Worth the Effort? *Chance* 17(1): 19-24.

Stoumbos, Z.G., M.R. Reynolds, Jr. 2000. Robustness to non-normality and autocorrelation of individuals control charts. *Journal of Statistical Computation & Simulation* 66(2): 145-187.

Toprani, A., R. Ratard, S. Straif-Bourgeois, T. Sokol, F. Averhoff, J. Brady, D. Staten, M. Sullivan, J.T. Brooks, A.K. Rowe, K. Johnson, P. Vranken, E. Sergienko. Surveillance in hurricane evacuation centers - Louisiana, September-October 2005. *Morbidity and Mortality Weekly Report* 55: 32–35.

U.S. Census Bureau, Population Division. 2007a. *Annual Estimates of the Population for Incorporated Places over 100,000, Ranked by July 1, 2006 Population* (SUB-EST2006-01). Release Date: June 28, 2007.

U.S. Census Bureau, Population Division. 2007b. *County total population, population change and estimated components of population change: April 1, 2000 to July 1, 2006* (CO-EST2006-alldata). Release Date: March 22, 2007.

Wagner, M.M., J. Espino, F.C. Tsui, P. Gesteland, W. Chapman, O. Ivanov, A. Moore, W. Wong, J. Dowling, J. Hutman. 2004. Syndrome and Outbreak Detection Using Chief-Complaint Data — Experience of the Real-Time Outbreak and Disease Surveillance Project. *Morbidity and Mortality Weekly Report* 53(Suppl): 28-31.

Washburn, A.R. 2002. *Search and Detection.* Topics in Operations Research Series. Linthicum, MD: Institute for Operations Research and the Management Sciences (INFORMS).

Wilson, A.G., G.D. Wilson, D.H. Olwell. 2006. *Statistical Methods in Counterterrorism: Game Theory, Modeling, Syndromic Surveillance, and Biometric Authentication.* New York: Springer.

Woodall, W.H. 2006. The Use of Control Charts in Health-Care and Public-Health Surveillance. *Journal of Quality Technology* 38(2): 88-103.

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, VA

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, CA

3. Associate Professor Ronald D. Fricker, Jr.
   Operations Research Department
   Naval Postgraduate School
   Monterey, CA

4. Jerome I. Tokars, MD, MPH
   Centers for Disease Control and Prevention
   Atlanta, GA

5. Henry Rolka, MPS, MS
   Centers for Disease Control and Prevention
   Atlanta, GA

6. Lori Hutwagner, MS
   Centers for Disease Control and Prevention
   Atlanta, GA

7. John Copeland, MS
   Centers for Disease Control and Prevention
   Atlanta, GA

8. Howard Burkom, PhD
   The Johns Hopkins University
   Applied Physics Laboratory
   Laurel, MD

9. Jeffrey Kline, CAPT, USN (Ret.)
   Operations Research Department
   Naval Postgraduate School
   Monterey, CA

10. CAPT Scott Miller, USN
    Information Sciences Department
    Naval Postgraduate School
    Monterey, CA

11. LAU Hoong Chuin
Associate Professor
School of Information Systems
Singapore Management University
Singapore