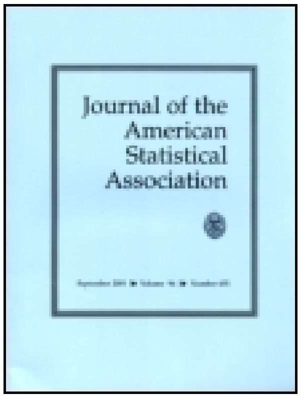


Journal of the American Statistical Association

Journal of the American Statistical Association
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/uasa20

A Latent Source Model to Detect Multiple Spatial Clusters With Application in a Mobile Sensor Network for Surveillance of Nuclear Materials

Jerry Q. Cheng [a] , Minge Xie [b] , Rong Chen [b] [c] & Fred Roberts [d]

[a] Robert Wood Johnson Medical School , Rutgers University , New Brunswick , NJ , 08901

[b] Department of Statistics and Biostatistics , Rutgers University , Piscataway , NJ , 08854

[c] School of Statistics , Central University of Economics and Finance , Beijing , 100081 , China

[d] Control, and Interoperability Center for Advanced Data Analysis (CCICADA), and Emeritus Director of the Center for Discrete Mathematics and Theoretical Computer Science (DIMACS) , Rutgers University , Piscataway , NJ , 08854

Accepted author version posted online: 20 Jun 2013.Published online: 27 Sep 2013.

To cite this article: Jerry Q. Cheng , Minge Xie , Rong Chen & Fred Roberts (2013) A Latent Source Model to Detect Multiple Spatial Clusters With Application in a Mobile Sensor Network for Surveillance of Nuclear Materials, Journal of the American Statistical Association, 108:503, 902-913, DOI: 10.1080/01621459.2013.808945

To link to this article: http://dx.doi.org/10.1080/01621459.2013.808945

A Latent Source Model to Detect Multiple Spatial Clusters With Application in a Mobile Sensor Network for Surveillance of Nuclear Materials

Jerry Q. CHENG, Minge XIE, Rong CHEN, and Fred ROBERTS

Potential nuclear attacks are among the most devastating terrorist attacks, with severe loss of human lives as well as damage to infrastructure. To deter such threats, it becomes increasingly vital to have sophisticated nuclear surveillance and detection systems deployed in major cities in the United States, such as New York City. In this article, we design a mobile sensor network and develop statistical algorithms and models to provide consistent and pervasive surveillance of nuclear materials in major cities. The network consists of a large number of vehicles on which nuclear sensors and Global Position System (GPS) tracking devices are installed. Real time sensor readings and GPS information are transmitted to and processed at a central surveillance center. Mathematical and statistical analyses are performed, in which we mimic a signal-generating process and develop a latent source modeling framework to detect multiple spatial clusters. A Monte Carlo expectation-maximization algorithm is developed to estimate model parameters, detect significant clusters, and identify their locations and sizes. We also determine the number of clusters using a modified Akaike Information Criterion/Bayesian Information Criterion. Simulation studies to evaluate the effectiveness and detection power of such a network are described.

KEY WORDS: AIC and BIC criteria; Cluster detection; EM algorithm; Likelihood inference; MCMC algorithm; Nuclear detection and surveillance.

1. INTRODUCTION

Since the attacks of September 11, 2001, homeland security has garnered increased attention of ordinary people and it has become one of the top priorities of the United States government. Among all the possible attacks by terrorists, nuclear attack is potentially the most devastating, and the global proliferation of nuclear weapon technology has made the threat increasingly serious. The U.S. government has made significant efforts to curb nuclear proliferation. In spite of many accomplishments, no effort can give full assurance against a clandestine delivery of a nuclear weapon for a terrorist attack. The graveness of such a cataclysmic possibility is apparent. As part of the effort, the Domestic Nuclear Detection Office (DNDO) within the Department of Homeland Security (DHS) was established in 2005 to improve the nation's capability to detect and collect information on unauthorized attempts to import, possess, store, develop, or transport nuclear or radiological material for use against the United States. The DNDO, in partnership with the National Science Foundation (NSF), has supported the Academic Research Initiative (ARI) program in frontier research at academic institutions focusing on detection systems, individual sensors, or other research that is potentially relevant to the detection of nuclear weapons, special nuclear material, radiation dispersal devices, and related threats. The Center for Discrete Mathematics and Theoretical Computer Science (DIMACS) of Rutgers University, through the involvement of the DHS University Center of Excellence, for Command, Control, and Interoperability (CCICADA), which is based at DIMACS, has led a multi-institution research project on nuclear detection supported by DNDO and this article describes one of the research thrusts of this project.

This article focuses on one of the aspects of detecting nuclear materials using a fleet of mobile radiation sensors in metropolitan areas. Major cities are attractive targets for terrorists because of their dense population and economic importance. However, a major city spreads throughout a large geographic area and is difficult to monitor and patrol. It is important to develop effective detection and surveillance methods to overcome or mitigate such difficulties. Throughout the article, we will use the generic term "nuclear detection" to include detection of any radiation-emitting materials of concern. Often, the nuclear and radiological materials of particular concern are radiation dispersion devices (RDDs), more commonly known as dirty bombs, and special nuclear materials (particularly highly enriched uranium and weapons-grade plutonium) that could provide the fissile material for a nuclear weapon.

When a hidden stationary or moving nuclear source emits radioactive energy to its immediately surrounding area, a sensor

Jerry Q. Cheng is Assistant Professor, Robert Wood Johnson Medical School, Rutgers University, New Brunswick, NJ 08901 (E-mail: *jcheng1@rwjms.rutgers.edu*). Minge Xie is Professor of Statistics, Department of Statistics and Biostatistics, Rutgers University, Piscataway, NJ 08854 (E-mail: *mxie@stat.rutgers.edu*). Rong Chen is Professor of Statistics, Department of Statistics and Biostatistics, Rutgers University, Piscataway, NJ 08854 (E-mail: *rongchen@stat.rutgers.edu*) and School of Statistics, Central University of Economics and Finance, Beijing 100081, China. Fred Roberts is Professor of Mathematics, Director of the Command, Control, and Interoperability Center for Advanced Data Analysis (CCICADA), and Emeritus Director of the Center for Discrete Mathematics and Theoretical Computer Science (DIMACS), Rutgers University, Piscataway, NJ 08854 (E-mail: *froberts@dimacs.rutgers.edu*).
This research was supported by research grants from NSF (CBET 0736134) and DHS (2008-DN-077-ARI012). Xie's research is also supported in part by NSF (DMS-0915139, SES-0851521, DMS1107012). Chen's research is also supported in part by NSF (DMS-0800183, DMS-0905763, DMS-0915139). The authors thank our colleagues in DIMACS and CCICADA, as well as Drs. Simon Labov and Karl Nelson of Lawrence Livermore National Laboratory (LLNL), for their comments and suggestions. The authors also thank Mr. Dong Dai for his help in developing the computing code for heteroscedastic sensors and Tsvetan Asamov for his help in developing a Visual Basic tool used in simulation. The research developed in this article is the authors' own and it does not reflect the policy or view of any sponsor agencies. The authors thank the editor, the associate editor, and two referees for the constructive comments and suggestions that help improve the quality of this article greatly.

Journal of the American Statistical Association
September 2013, Vol. 108, No. 503, Applications and Case Studies
DOI: 10.1080/01621459.2013.808945

for nuclear detection within a certain range of the source would be activated and send out warning signals (though it is also possible for a sensor to send a false alarm). If there are multiple sensors nearby, a cluster of activated sensors will be formed around the source. The cluster is often visible when plotted on a map, and it can in turn help reveal the location of the hidden source with high accuracy. This consideration translates the hidden source detection problem into a visible cluster detection problem. Statistically, we inspect the entire region of interest, and test whether one or more spatial clusters exist in the region and whether or not the spatial clusters are statistically significant against random false alarms and noise in the signals. The center of the detected cluster is often used as an estimate of the location of the nuclear source.

Traditionally *scan statistics* are used to detect a cluster of events in spatial data (Glaz and Balakrishnan 1999; Balakrishnan and Koutras 2001; Glaz, Naus, and Wallenstein 2001; Fu and Lou 2003). The most commonly used scan statistic is the maximum number of events within a fixed size window that scans through the study area. Based on scan statistics, a generalized likelihood ratio test has been developed to test the null hypothesis that all signals are uniformly distributed in the area (the case of no cluster) (Naus 1966). Other scan statistics and related likelihood-based tests for localized temporal or spatial clustering have been developed, often using a range of fixed window sizes or a range of fixed number of cases (Dembo and Karlin 1992; Kulldorff and Nagarwalla 1995; Su, Wallenstein, and Bishop 2001; Naus and Wallenstein 2004). Scan statistics methods have also been developed under the Bayesian framework (Lawson 1995; Gangnon and Clayton 2000; Denison and Holmes 2001; Gangnon and Clayton 2003).

Scan statistics procedures are very effective in detecting a single significant cluster, and they also have had some success in detecting multiple clusters of fixed sizes. But they are not particularly suitable for detecting multiple clusters of varying sizes (see, e.g., Xie, Sun, and Naus 2009). In recent years, several model-based procedures have been developed to detect multiple clusters of varying sizes. For instance, Demattei, Molinari, and Daures (2006, 2007) proposed a stepwise regression model combined with model selection procedures to locate and determine the number of clusters. This method relies on a weighted least-square formulation, although the response variable (gaps between incidents) is typically non-Gaussian. Xie, Sun, and Naus (2009) developed a latent source cluster model for temporal data, which uses standard likelihood inference for detecting multiple clusters. Sun (2008) extended the approach to spatial data. Based on likelihood inference, the latent source model approach is more efficient in detecting clusters of varying sizes than the weighted least-square approaches.

Another class of statistical methods for detecting a cluster of events is the “disease mapping” type of approaches both in Bayesian and frequentist paradigms; see, for example, Lawson (1995), Waller et al. (1997), Ghosh et al. (1999), Knorr-Held and RaBer (2000), and Diggle, Rowlingson, and Sun (2005). These approaches consider intensity functions (i.e., local density estimates of intensities) and they do not directly make inference about clusters. To detect clusters, the “disease mapping” approaches rely on a certain choice of threshold, which is often subjective. In addition, Dirichlet process (DP) models (Ferguson 1973; Lo 1984; Escobar and West 1995; Neal 2000; Teh et al. 2006; McCullagh and Yang 2008) have been developed and applied in clustering problems. As a Bayesian nonparametric density estimation method, DP is another potentially powerful approach for cluster detection and would be an interesting research topic for further study. Our approach is more parametric in nature.

In this article, we design a mobile sensor network for surveillance of nuclear materials in a metropolitan area. The network consists of a large number of moving vehicles, such as taxicabs, police cars, fire trucks, and other participating private and public vehicles, on which nuclear sensors and Global Position System (GPS) tracking devices are installed. Real time sensor readings are processed at a central surveillance center, where the data are analyzed to detect significant clusters of signals that might indicate the locations of nuclear sources. We further develop and expand the latent source model method developed by Xie et al. (2009) and Sun (2008) and use likelihood inference to detect multiple clusters simultaneously in a region. Instead of only considering positive signals and treating the negative ones as background, our model uses the information from both positive and negative signals and achieves higher power in detecting sources. This method is suitable to analyze signals from the proposed mobile sensor framework and reduce false alarms.

Note that it is often difficult to model accurately the movements of the nuclear sources and the sensors, since it is difficult to model or track the destination and intention of each driver or terrorist. In our data analysis, we do not consider motion models for either the nuclear sources (i.e., terrorist movement) or the sensors (i.e., the movements of taxicabs, etc.), and the statistical analysis is performed based on data collected at each fixed time point. A sequence of data analyses in consecutive times (e.g., every 30 sec or every 1 min, say) can then form a dynamic surveillance—just like the movement in a movie, which is formed from stationary film frames. This approach is robust against potential model misspecification of the movements of the nuclear sources and the sensors. It is applicable to detect both stationary and moving sources.

The rest of the article is arranged as follows. Section 2 begins with a prototype of a mobile sensor network and considers related models for nuclear intensity, sensor reading, and detection. Section 3 provides a latent source model, a likelihood-inference-based methodology, and an EM/Markov chain Monte Carlo (MCMC) algorithm to detect clusters and make inference based on information (sensor signals and their locations) collected at each fixed time point. The section concludes with an additional development in detection using mixed-type sensors. Section 4 describes simulation studies on several practical scenarios, where detection power of the network is estimated under various sets of parameters. In this section, to generate data that simulate city traffic, we introduce a movement model with a constraint that vehicles can only travel on street grids, though the data analysis is based on the development in Sections 2 and 3 and does not depend on the movement model. Section 5 concludes the article with discussion and future research directions.

2. NUCLEAR DETECTION USING A MOBILE SENSOR NETWORK

2.1 Mobile Sensor Network Prototype

Proactive monitoring and detection via pervasive surveillance is crucial to detect and thwart malicious attacks in major cities. Here, we propose a prototype of a mobile sensor network as follows:

(i) Inexpensive nuclear sensors and GPS tracking devices are mounted on a large number of vehicles, such as police cars, taxicabs, fire trucks, buses, and other participating vehicles.
(ii) The sensors and GPS devices constantly send detection signals and location information to a central command center. These signals are marked onto a map of a metropolitan area under surveillance.
(iii) Real time analysis is performed at the command center using sophisticated statistical algorithms including the latent source modeling method discussed in this article to detect and pinpoint nuclear sources.
(iv) Sensors and the control center are periodically serviced and calibrated to ensure validity and accuracy.

The advantage of a mobile network over a static surveillance network is multifold. First, due to mobility, the failure of a portion of the sensors will not have significant impact on the surveillance coverage, while any sensor failure in a static network will create blind spots. Second, the locations of blind spots in a static network are unknown since a failed sensor often would send in negative signals. On the other hand, the blind spots of a mobile network (where vehicles cannot reach) can be covered with a small number of static sensors. Third, mobile sensors may not need to be highly reliable and of large range, due to their mobility, while static sensors would need to be more reliable and powerful to maintain the coverage. Since our sensors are vehicle mounted, their size and power requirements will not be high. Hence, inexpensive sensors could be deployed. Fourth, mobile sensors can be inspected and calibrated during vehicle inspection and routine maintenance, while static sensors can only be serviced by visits of maintenance personnel to their physical locations. Fifth, it is almost impossible to tamper with a mobile sensor network, while a static network can be an easy target.

Due to these attractive characteristics, there have been many studies and applications of sensor networks in military and civil applications, including surveillance, smart homes, and remote environment monitoring (Akyildiz et al. 2002a, 2002b). Much of the research is devoted to sensor placement, sensor reorganization, and communications. In the area of radiation detection, the idea of using a network of mobile sensors has been adopted and tested by the Radiation Laboratory at Purdue University 2008. They used a network of cell phones with GPS capabilities to detect and track radiation. The noise and false positive detection problems were tackled by setting and tuning solid-state devices. A multisensor nuclear threat detection problem was studied in Hochbaum (2009) using a combinatorial network flow algorithm, in which the main research objective is the development of a fast algorithm for detection with mobile sensors (such as taxicabs) instead of the efficient use of data information. In this article, we use a latent source modeling method for source detection. This statistical approach is effective at detecting true signals against random errors and thus reducing false alarms.

2.2 Nuclear Intensity, Sensor Reading, and Detection Models

In this article, we consider the source as a portable nuclear device transported by an individual via vehicles or bags (FEMA 2008). As nuclear radiation emits from a source, its total energy denoted as E remains constant due to the Energy Conservation Law. Following Wein et al. (2006) and many others, we assume that radiation travels in spherical waves. Let $z(\rho)$ be the nuclear energy intensity at distance ρ from the source, therefore,

$$z(\rho) = \frac{E}{4\pi\rho^2} \equiv \frac{c}{\rho^2},$$

where $c \equiv E/4\pi$ is a constant factor related to the total energy of the source. The radiation intensity decreases by the inverse square of distance to the source.

In the following, we establish the behavior of a signal detection sensor based on the above physical principle. We assume all sensors are independent, following an identical model related to their locations. Suppose a nuclear detection sensor is triggered when the radiation intensity it receives exceeds a certain threshold. For low-cost sensors that can be economically distributed in large quantity, numerical readings are often not very accurate. Instead a binary signal mechanism is cheaper to design with low cost of maintenance. As a result, a binary sensor outcome model obtained by thresholding (Stroud and Saeger 2003; Boros et al. 2009; Elsayed et al. 2009) is adopted here. Specifically, let R denote the reading status of the sensor with a value of 1 for a positive reading and 0 otherwise. We describe R with a threshold model:

$$R = \mathbf{1}_{\{z(\rho)\geq h\}} = \mathbf{1}_{\{c/\rho^2\geq h\}}, \tag{1}$$

where h is a threshold for detection and $\mathbf{1}_{\{.\}}$ is the indicator function. That is, if the intensity $z(\rho)$ at the sensor location is greater than or equal to the threshold h, the sensor reports a detection; otherwise the sensor reports a negative reading.

In the case when there are multiple nuclear sources, their energy levels and positions will jointly determine the reading status of a sensor. Let Γ be the number of sources and c_γ be the energy factor of the γth source. When sources have different energy spectra, they will activate a sensor independently. In this case, the threshold model is

$$R = \mathbf{1}_{\{\max_{\gamma\in\{1,\ldots,\Gamma\}} z_\gamma(\rho_\gamma)\geq h\}} = \mathbf{1}_{\{\max_{\gamma\in\{1,\ldots,\Gamma\}} c_\gamma/\rho_\gamma^2\geq h\}}, \tag{2}$$

where ρ_γ is the distance of the γth source to the sensor and $z_\gamma(\rho_\gamma)$ is its corresponding intensity. When the sources assume the same energy spectrum, intensities from all sources at the sensor location are aggregated: $z_{\text{total}} = \sum_{\gamma=1}^{\Gamma} c_\gamma/\rho_\gamma^2$. Then the reading R can be determined by

$$R = \mathbf{1}_{\{z_{\text{total}}\geq h\}} = \mathbf{1}_{\{\sum_{\gamma=1}^{\Gamma} c_\gamma/\rho_\gamma^2\geq h\}}. \tag{3}$$

As with any detection device, a nuclear sensor may not be 100% accurate. A sensor might display positive readings when there is no nuclear source nearby (false positive), or fail to detect a real source (false negative). This is unavoidable even for expensive and highly accurate sensors and in particular for our study with a massive number of relatively inexpensive sensors. The sensor errors can be from the variability in the manufacturing process, routine wear and tear, missing scheduled maintenance and calibrations, and other malfunctions. In addition, random traces of weak environmental nuclear signals can also trigger false alerts. Furthermore, wireless signals from a mobile sensor to the control center may incur transmission errors. We use two parameters, *sensitivity* η and *specificity* ζ, to measure the average performance of a sensor device.

Specially, let D be the binary indicator of a sensor signal, with a value of 1 for a positive detection and 0 otherwise. We have

$$\eta = P(D=1|R=1) \quad \text{and} \quad \zeta = P(D=0|R=0). \tag{4}$$

The terms $1-\eta$ and $1-\zeta$ correspond to false negative rate (FNR) and false positive rate (FPR), which are quality characteristics of a sensor. Then the probability of a positive detection is

$$\begin{aligned} P(D=1) &= P(D=1|R=1)P(R=1) \\ &\quad + P(D=1|R=0)P(R=0) \\ &= \eta P(R=1) + (1-\zeta)(1-P(R=1)) \\ &= (1-\zeta) + (\zeta+\eta-1)P(R=1). \end{aligned} \tag{5}$$

Under the perfect scenario where both η and ζ are 1, $P(D=1)$ is the same as $P(R=1)$. Realistically, both of them are less than 1. It is noted that η and ζ may depend on the source and its energy spectra. Furthermore, both of them are fixed and often unknown constant parameters for any given nuclear source.

2.3 Detection of Nuclear Sources and the Statistical Problem of Cluster Detection

The sensor physical reading models (1)–(5) about sensor readings in Section 2.2 can be connected to the latent source cluster detection problem in statistics. For instance, with a single source, the threshold model (1) can be expressed as $R = \mathbf{1}_{\{I\}}$, where $I = \{\rho \le (c/h)^{1/2}\}$ is a sphere, or a circle on a two-dimensional map, centering at the nuclear source and with a radius $(c/h)^{1/2}$. From (4) the ratio of the probabilities of a positive reading inside and outside the set I is

$$\alpha = \frac{P(D=1|I)}{P(D=1|\bar{I})} = \frac{P(D=1|R=1)}{P(D=1|R=0)} = \frac{\eta}{1-\zeta}. \tag{6}$$

Assume that $\eta + \zeta > 1$ (which is a condition satisfied by almost all commercially available detection devices), we have $\alpha > 1$. Thus, formula (6) states that a sensor is α times more likely to report a positive signal ($D = 1$) inside I than outside I. This statement matches exactly the underlying concept of a spatial cluster in many statistics developments, which is defined as an area within which an incident of interest is more likely to happen (i.e., with a higher probability of happening per unit of area) than outside of the area (Kulldorff and Nagarwalla 1995; Gangnon and Clayton 2000; Xie et al. 2009). Here, an incident of interest is an alert signal with $D = 1$. Thus, the sensor reading models presented in this section link us to a statistical cluster detection problem.

A similar connection can be explored in the presence of multiple nuclear sources at different locations. The motivation for simultaneous detection of multiple sources is the fact that it is possible for terrorists to plant multiple nuclear materials in an urban region at the same time but at different places. Such a plot may induce more severe damage and make prevention and rescue efforts more difficult.

3. LATENT SOURCE MODELING PROCEDURES FOR SOURCE DETECTION

Assume that there are k stationary or moving nuclear sources in a given two-dimensional rectangular region, $\mathcal{I} = (0, L) \times (0, W)$, where L is the length and W is the width. Suppose that the jth source ($j = 1, \ldots, k$) is located at $\mathbf{o}_j$ and has a strength that can be effectively detected by a sensor within a distance of r_j. If the distance of a sensor to $\mathbf{o}_j$ is less than r_j, the sensor would generate a positive reading (subject to false negative errors). Denote by cluster I_j the circle centering at $\mathbf{o}_j$ with radius r_j. Also define a $2 \times k$ matrix $\mathbf{O} = (\mathbf{o}_1, \ldots, \mathbf{o}_k)$ and a vector $\mathbf{r} = (r_1, \ldots, r_k)$ as the collection of locations and sizes (radii) of clusters. For simplicity, we assume in this article that these k clusters are not overlapping; see sec. 6 and chap. 5 of Cheng (2010) for discussions of overlapping cases. Outside the k clusters, a sensor is far away from nuclear sources and would have a negative reading (subject to false positive errors).

We treat $\mathbf{O}$ and $\mathbf{r}$ as latent random variables and assume that cluster centers follow a density function $\psi_o(\cdot; \lambda_o)$ and radii a density function $\psi_r(\cdot; \lambda_r)$. Here, λ_o and λ_r are unknown parameters jointly denoted by $\boldsymbol{\lambda} = (\lambda_o, \lambda_r)$, while ψ_o and ψ_r are their density functions, respectively. We use a uniform distribution on $\mathcal{I}$ for $\psi_o(\cdot; \lambda_o)$ and a truncated exponential distribution for $\psi_r(\cdot; \lambda_r)$. Other choices for $\psi_r(\cdot; \lambda_r)$ are inverse Gamma or log-normal distributions (see, e.g., Sun 2008). Let Ω_k be the set of $(\mathbf{O}, \mathbf{r})$ such that the k clusters specified by $(\mathbf{O}, \mathbf{r})$ are nonoverlapping. Given that there are k nonoverlapping clusters, the joint conditional likelihood function of $(\mathbf{O}, \mathbf{r})$ is

$$f_{\boldsymbol{\lambda}}(\mathbf{O}, \mathbf{r}|k) = \frac{\prod_{j=1}^{k}\{\psi_o(\mathbf{o}_j; \lambda_o)\psi_r(r_j; \lambda_r)\}\mathbf{1}_{\{(\mathbf{O},\mathbf{r})\in\Omega_k\}}}{C_{\boldsymbol{\lambda},k}}, \tag{7}$$

where $C_{\boldsymbol{\lambda},k}$ is the normalization constant for the truncated density.

Assume that there are N sensors. Let $\mathbf{y}_i$, a two-dimensional vector, be the location of the ith sensor ($i = 1, \ldots, N$), and δ_i be its detection status indicator with value of 1 if the sensor sends a positive signal and 0 a negative one. Let $\mathbf{Y} = \{\mathbf{y}_1, \ldots, \mathbf{y}_N\}$ and $\boldsymbol{\delta} = \{\delta_1, \ldots, \delta_N\}$. Also, extending (4) to multiple sources, we define the sensitivity related to the jth cluster η_j for $j = 1, \ldots, k$ and specificity ζ as

$$\eta_j = P(\delta_i = 1|\mathbf{y}_i \in I_j) \quad \text{and} \quad \zeta = P\left(\delta_i = 0|\mathbf{y}_i \notin \bigcup_{j=1}^{k} I_j\right) \quad \text{for } i = 1, \ldots, N. \tag{8}$$

Let $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_k)$. As in the case of a single source, we treat $\boldsymbol{\eta}$ and ζ as unknown parameters to be estimated. Here, each of the k elements in $\boldsymbol{\eta}$ reflects a sensitivity related to a particular source, while we retain a single parameter ζ because it relates to the region outside of all the k clusters. Under (8) and extending (6), the ratio of the probabilities of a positive reading inside I_j and outside of all clusters (i.e., the background) is $\alpha_j = \eta_j/(1-\zeta)$.

Given the locations $\mathbf{O}$ and the sizes (radii) $\mathbf{r}$ of the clusters corresponding to the k sources, $(\mathbf{y}_1, \delta_1), \dots, (\mathbf{y}_N, \delta_N)$ are an independently identically distributed (iid) sample of size N from a density function,

$$\begin{aligned} f_{\boldsymbol{\eta},\zeta}(\mathbf{y}, \delta|\mathbf{O}, \mathbf{r}, k) &= f_{\boldsymbol{\eta},\zeta}(\delta|\mathbf{y}, \mathbf{O}, \mathbf{r}, k) f(\mathbf{y}|\mathbf{O}, \mathbf{r}, k) \\ &= f_{\boldsymbol{\eta},\zeta}(\delta|\mathbf{y}, \mathbf{O}, \mathbf{r}, k) f(\mathbf{y}) \\ &= [\zeta^{1-\delta}(1-\zeta)^{\delta}]^{\mathbf{1}\{\mathbf{y}\notin\bigcup_{j=1}^{k} I_j\}} \\ &\quad \times \prod_{j=1}^{k} \left[\eta_j^{\delta}(1-\eta_j)^{1-\delta}\right]^{\mathbf{1}\{\mathbf{y}\in I_j\}} f(\mathbf{y}), \quad (9) \end{aligned}$$

where $f(\mathbf{y})$ is the density function of the sensor location, which is assumed to be unrelated to the sources. We consider a hypothesis testing problem—H_0: $\eta_1 = \cdots = \eta_k = 1 - \zeta$ versus H_1: at least one $\eta_j \neq 1 - \zeta$. Note that under the null hypothesis H_0, the joint density function in (9) is reduced to $\zeta^{1-\delta}(1-\zeta)^{\delta} f(\mathbf{y})$. In this case, the sensor detection indicator is independent of the sensor's location, implying that there is no source in the region.

For given k clusters, the proposed latent source model can be expressed using a hierarchical structure,

$$\begin{aligned} &\text{Level 1:} \quad \mathbf{y}_i, \delta_i|\mathbf{O}, \mathbf{r} \sim \text{Model (9)}, \quad \text{for } i = 1, 2, \dots, N; \\ &\text{Level 2:} \quad \mathbf{o}_j|\lambda_o \sim \psi_o(\cdot) \text{ and } r_j|\lambda_r \sim \psi_r(\cdot), \\ &\qquad\qquad \text{for } j = 1, 2, \dots, k \text{ and } (\mathbf{O}, \mathbf{r}) \in \Omega_k. \quad (10) \end{aligned}$$

If further imposing prior distributions on the parameters $\boldsymbol{\eta}$, ζ, and $\boldsymbol{\lambda}$, we obtain a corresponding Bayesian model with an additional third level to (10):

$$\begin{aligned} &\text{Level 3:} \quad \zeta \sim \pi_{\zeta}(\cdot), \;\; \lambda_o \sim \pi_{\lambda_o}(\cdot), \;\; \lambda_r \sim \pi_{\lambda_r}(\cdot), \\ &\qquad\qquad \text{and} \quad \eta_j \sim \pi_{\eta}(\cdot), \quad \text{for } j = 1, 2, \dots, k, \quad (11) \end{aligned}$$

where π_{η}, π_{ζ}, π_{λ_o}, and π_{λ_r} are the priors for η_j, ζ, λ_o, and λ_r, respectively. If flat priors are used, the results from the likelihood inference using hierarchical model (10) are often the same as or similar to the ones using the Bayesian hierarchical model (11). We also note that there are other formulations for the same detection problem, for example, those using Poisson assumptions or Bayesian state-space models. Here, we use the current formulation to turn the nuclear detection problem into a formal statistical testing problem without involving thresholding.

To simplify our inference, we first treat the number of clusters k as given and known; Section 3.2.2 provides modified Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) to determine k when it is unknown. Alternatively in the Bayesian approach, some authors such as Rodriguez, Sunson, and Gelfand (2008) directly used a nested DP prior to place a collection of distributions on different clusters. The approach allows for simultaneously clustering groups and observations within groups, and the groups are clustered by their entire distribution, rather than by particular features of the distribution. Its outcomes may be sensitive to a particular choice of the Dirichlet prior and caution should be applied. This is an interesting approach and its potential use in our problem requires further study.

The proposed latent source model is connected to the models considered in the scan statistic literature, though ours is significantly different. Our setting is very different from the typical epidemiology setting for the use of scan statistic, such as Glaz, Naus, and Wallenstein (2001) and Sun (2008), where the population density is high and the disease is rare (comparing to population). In their case, even in the center of the disease cluster, the ratio of positives to the population density is extremely small. Hence, the distribution of negatives is essentially the same inside or outside the cluster. Under such a setting, the negative signals can be ignored. In our case, we do not have a dense population of sensors. Within the detection region of the source, the FNR is not very high, so there is a significant difference, in terms of negative signals, inside or outside the detection region. As a result, we cannot ignore the negative signals in our study.

We develop our inference framework based on the hierarchical model (10) with known k. From this model, the conditional joint distribution function of $(\mathbf{y}_1, \delta_1), \dots, (\mathbf{y}_N, \delta_N)$, when given k, $\mathbf{O}$, $\mathbf{r}$, and $(\mathbf{O}, \mathbf{r}) \in \Omega_k$, is

$$\begin{aligned} &f_{\boldsymbol{\eta},\zeta}(\mathbf{Y}, \boldsymbol{\delta}|\mathbf{O}, \mathbf{r}, k) \\ &\quad = \prod_{i=1}^{N} f_{\boldsymbol{\eta},\zeta}(\mathbf{y}_i, \delta_i|\mathbf{O}, \mathbf{r}, k) \\ &\quad = \exp\left\{\sum_{j=1}^{k} Z_j \log \eta_j + \sum_{j=1}^{k} Z_j^* \log(1-\eta_j) \right. \\ &\qquad + \left(n^* - \sum_{j=1}^{k} Z_j^*\right) \log\zeta + \left(n - \sum_{j=1}^{k} Z_j\right) \log(1-\zeta) \\ &\qquad \left. + \sum_{i=1}^{N} \log f(\mathbf{y_i})\right\}, \quad (12) \end{aligned}$$

where $Z_j = \sum_{i=1}^{N} \delta_i \mathbf{1}_{\{\mathbf{y}_i \in I_j\}}$, $Z_j^* = \sum_{i=1}^{N} (1-\delta_i) \mathbf{1}_{\{\mathbf{y}_i \in I_j\}}$, $n = \sum_{i=1}^{N} \delta_i$, and $n^* = \sum_{i=1}^{N} (1-\delta_i)$ are the number of positive readings inside the jth cluster, the number of negative readings inside the jth cluster, the total number of positive readings, and the total number of negative readings, respectively. With observing $(\mathbf{Y}, \boldsymbol{\delta}, k)$, the observed likelihood function of the above model is

$$l(\boldsymbol{\theta}) = f_{\boldsymbol{\theta}}(\mathbf{Y}, \boldsymbol{\delta}, k) = \int \dots \int f_{\boldsymbol{\theta}}(\mathbf{Y}, \boldsymbol{\delta}, \mathbf{O}, \mathbf{r}, k) d\mathbf{O} d\mathbf{r}, \quad (13)$$

where $\boldsymbol{\theta} = (\boldsymbol{\eta}, \zeta, \boldsymbol{\lambda})$ and the integrand is the joint density of the complete data $(\mathbf{Y}, \boldsymbol{\delta}, \mathbf{O}, \mathbf{r}, k)$:

$$\begin{aligned} &f_{\boldsymbol{\theta}}(\mathbf{Y}, \boldsymbol{\delta}, \mathbf{O}, \mathbf{r}, k) \\ &\quad = f_{\boldsymbol{\eta},\zeta}(\mathbf{Y}, \boldsymbol{\delta}|\mathbf{O}, \mathbf{r}, k) f_{\boldsymbol{\lambda}}(\mathbf{O}, \mathbf{r}, k) \\ &\quad = \exp\left\{\sum_{j=1}^{k} Z_j \log \eta_j + \sum_{j=1}^{k} Z_j^* \log(1-\eta_j) \right. \\ &\qquad + \left(n^* - \sum_{j=1}^{k} Z_j^*\right) \log\zeta + \left(n - \sum_{j=1}^{k} Z_j\right) \log(1-\zeta) \\ &\qquad \left. + \sum_{i=1}^{N} \log f(\mathbf{y_i})\right\} \prod_{j=1}^{k} \{\psi_o(\mathbf{o}_j) \psi_r(r_j)\} \mathbf{1}_{\{(\mathbf{O},\mathbf{r}) \in \Omega_k\}}. \end{aligned}$$

3.1 Estimation of Model Parameters and Locations of Clusters

3.1.1 Monte Carlo EM Algorithm for Model Estimation. Since the log-likelihood $l(\boldsymbol{\theta})$ from (13) involves multiple integrations, it is difficult to compute its value and its first and second derivatives directly. We instead implement a Monte Carlo expectation-maximization (MC-EM) algorithm (see, e.g., Tanner 1993, sec. 4.5) where we treat $(\mathbf{Y}, \boldsymbol{\delta}, \mathbf{O}, \mathbf{r}, k)$ as the complete data and $(\mathbf{Y}, \boldsymbol{\delta}, k)$ as the observed data. The steps of the algorithm are as follows.

Step 0. Set $s = 0$ and select a set of starting parameter values $\boldsymbol{\theta}^{(0)} = (\boldsymbol{\eta}^{(0)}, \zeta^{(0)}, \boldsymbol{\lambda}^{(0)})$.

Step 1 (E-step). Calculate the conditional expectation of the complete log-likelihood function, given the observed observations and $\boldsymbol{\theta} = \boldsymbol{\theta}^{(s)}$:

$$Q\left(\boldsymbol{\theta}|\boldsymbol{\theta}^{(s)}\right) = Q_1\left(\boldsymbol{\eta}, \zeta\,|\boldsymbol{\eta}^{(s)}, \zeta^{(s)}\right) + Q_2\left(\boldsymbol{\lambda}|\boldsymbol{\lambda}^{(s)}\right), \tag{14}$$

where

$$\begin{aligned}
&Q_1\left(\boldsymbol{\eta}, \zeta\,|\boldsymbol{\eta}^{(s)}, \zeta^{(s)}\right)\\
&\quad= \sum_{j=1}^{k} E\left\{Z_j|\mathbf{Y}, \boldsymbol{\delta}, k, \boldsymbol{\eta}^{(s)}, \zeta^{(s)}\right\}\log \eta_j\\
&\qquad+ \sum_{j=1}^{k} E\left\{Z_j^*|\mathbf{Y}, \boldsymbol{\delta}, k, \boldsymbol{\eta}^{(s)}, \zeta^{(s)}\right\}\log(1-\eta_j)\\
&\qquad+ \left(n^* - \sum_{j=1}^{k} E\left\{Z_j^*|\mathbf{Y}, \boldsymbol{\delta}, k, \boldsymbol{\eta}^{(s)}, \zeta^{(s)}\right\}\right)\log\zeta\\
&\qquad+ \left(n - \sum_{j=1}^{k} E\left\{Z_j|\mathbf{Y}, \boldsymbol{\delta}, k, \boldsymbol{\eta}^{(s)}, \zeta^{(s)}\right\}\right)\\
&\qquad\times \log(1-\zeta),\\
&Q_2\left(\boldsymbol{\lambda}|\boldsymbol{\lambda}^{(\mathbf{s})}\right)\\
&\quad= \sum_{j=1}^{k} E\left\{\log\psi_o(\mathbf{o}_j)|\mathbf{Y}, \boldsymbol{\delta}, k, \lambda_o^{(s)}\right\}\\
&\qquad+ \sum_{j=1}^{k} E\left\{\log\psi_r(r_j)|\mathbf{Y}, \boldsymbol{\delta}, k, \lambda_r^{(s)}\right\}.
\end{aligned}$$

Step 2 (M-step). Update the parameter estimates: $\boldsymbol{\theta}^{(s+1)} = (\boldsymbol{\eta}^{(s+1)}, \zeta^{(s+1)}, \boldsymbol{\lambda}^{(s+1)})'$, by maximizing the $Q_1(\boldsymbol{\eta}, \zeta\,|\boldsymbol{\eta}^{(s)}, \zeta^{(s)})$, and $Q_2(\boldsymbol{\lambda}|\boldsymbol{\lambda}^{(\mathbf{s})})$ functions: $(\boldsymbol{\eta}^{(s+1)}, \zeta^{(s+1)}) = \operatorname{argmax} Q_1(\boldsymbol{\eta}, \zeta\,|\boldsymbol{\eta}^{(s)}, \zeta^{(s)})$, and $\boldsymbol{\lambda}^{(s+1)} = \operatorname{argmax} Q_2(\boldsymbol{\lambda}|\boldsymbol{\lambda}^{(s)})$.

Stop if $\|\boldsymbol{\theta}^{(s+1)} - \boldsymbol{\theta}^{(s)}\|$ is very small (the criterion that we have used is less than $\epsilon(\|\boldsymbol{\theta}^{s}\| + \epsilon)$ with $\epsilon = 10^{-3}$).

Step 3. Increase s by 1 and go back to Step 1.

The conditional expectations in the E-step of the EM algorithm do not usually have explicit form. To solve this problem, we use a Monte Carlo method. In particular, we use Gibbs and importance sampling methods to generate clusters one at a time. That is, we generate samples of the jth cluster from its full conditional distribution given the rest of the clusters:

$$\begin{aligned}
&f_{\boldsymbol{\theta}}((\mathbf{o}_j, r_j)|(\mathbf{o}_l, r_l), l = 1, \ldots, k, l \neq j, \mathbf{Y}, \boldsymbol{\delta}, k)\\
&\quad\propto f_{\boldsymbol{\theta}}(\mathbf{Y}, \boldsymbol{\delta}, \mathbf{O}, \mathbf{r}, k)\\
&\quad\propto \left(\frac{\eta_j}{1-\zeta}\right)^{Z_j}\left(\frac{1-\eta_j}{\zeta}\right)^{Z_j^*}\psi_{\lambda_o}(\mathbf{o}_j)\psi_{\lambda_r}(r_j)\mathbf{1}_{\{(\mathbf{O},\mathbf{r})\in\Omega_k\}}.
\end{aligned} \tag{15}$$

Since it is still difficult to directly generate samples from (15), we adopt an importance sampling approach. With that, it is easy to simulate a $(\mathbf{o}_j, r_j)$ from $\varphi_{\lambda_o}(\mathbf{o}_j)\varphi_{\lambda_r}(r_j)$. Specifically, we simulate a large number, say S, of random deviates $(\mathbf{o}_j, r_j)^{[s]}(s = 1, \ldots, S)$ from the prior distribution $\varphi_{\lambda_o}(\mathbf{o}_j)\varphi_{\lambda_r}(r_j)$. For each s, compute weight

$$w_s = \left(\frac{\eta_j}{1-\zeta}\right)^{Z_j^{[s]}}\left(\frac{1-\eta_j}{\zeta}\right)^{Z_j^{*[s]}}\mathbf{1}_{\{(\mathbf{O}^{[s]},\mathbf{r}^{[s]})\in\Omega_k\}},$$

where $Z_j^{[s]}$, $Z_j^{*[s]}$, and $\mathbf{1}_{\{(\mathbf{O}^{[s]},\mathbf{r}^{[s]})\in\Omega_k\}}$ are calculated with $(\mathbf{O}^{[s]}, \mathbf{r}^{[s]}) = \{(\mathbf{o}_j, \mathbf{r}_j)^{[s]}, (\mathbf{o}_l, r_l), l = 1, \ldots, k, l \neq j\}$. We then choose one $(\mathbf{o}_j, r_j)$ among the S sets $(\mathbf{o}_j, r_j)^{[1]}$,...,$(\mathbf{o}_j, r_j)^{[S]}$ with respective probabilities $(p_1, \ldots, p_S)$, where $p_s = w_s/\sum_{s=1}^{S} w_s$. The sample set $\{(\mathbf{o}_j, r_j) : j = 1, \ldots, k\}$ forms a Gibbs sample that is used to calculate the conditional expectations in the E-step.

3.1.2 Identification of Cluster Regions. In nuclear detection, a primary goal is to identify regions of the sources determined by the centers and radii: $\{(\mathbf{o}_j, r_j) : j = 1, \ldots, k\}$. Their conditional expectations given $(\mathbf{Y}, \boldsymbol{\delta}, k)$ (similar to "posterior means" in the context of Bayesian statistics) are $E\{\mathbf{o}_j|\mathbf{Y}, \boldsymbol{\delta}, k\}|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}}$ and $E\{r_j|\mathbf{Y}, \boldsymbol{\delta}, k\}|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}}$, where $\widehat{\boldsymbol{\theta}}$ is obtained from the MC-EM algorithm. They can be estimated by their sample means or medians if their distribution is not symmetric.

3.2 Likelihood Ratio Test and Determination of the Unknown Number of Clusters

3.2.1 Likelihood Ratio Test of Significant Clusters. For a fixed k, we obtain k clusters from the aforementioned estimation procedures. An interesting problem is to check whether any of the k sources is significant, or equivalently, to determine whether any of the detected sources is due to random noise in the context of nuclear detection. In particular, we test a hypothesis about the parameter set $(\boldsymbol{\eta}, \zeta)$, that is, H_0: $\eta_1 = \cdots = \eta_k = 1-\zeta$ versus H_1: at least one $\eta_j \neq 1-\zeta$. We propose to use the log-likelihood ratio test statistic

$$\begin{aligned}
&\Lambda(\mathbf{Y}, \boldsymbol{\delta}, k)\\
&\quad= \log\left\{\frac{\max_{H_1\cup H_0} f_{\boldsymbol{\theta}}(\mathbf{Y}, \boldsymbol{\delta}, k)}{\max_{H_0} f_{\boldsymbol{\theta}}(\mathbf{Y}, \boldsymbol{\delta}, k)}\right\}\\
&\quad= \log\iint f_{\widehat{\boldsymbol{\eta}},\widehat{\zeta}}(\mathbf{Y}, \boldsymbol{\delta}|\mathbf{O}, \mathbf{r}, k) f_{\widehat{\boldsymbol{\lambda}}}(\mathbf{O}, \mathbf{r}|k)\,d\mathbf{O}\,d\mathbf{r} + \log C_{\widehat{\boldsymbol{\lambda}},k}\\
&\qquad- \Bigg[n\log n + n^*\log n^* - (n+n^*)\log(n+n^*)\\
&\qquad+ \sum_{i=1}^{N}\log f(\mathbf{y_i})\Bigg],
\end{aligned}$$

where $\widehat{\boldsymbol{\theta}} = (\widehat{\boldsymbol{\eta}}, \widehat{\zeta}, \widehat{\boldsymbol{\lambda}})^T$ are the nonrestricted maximum-likelihood estimations (MLEs; under $H_1 \bigcup H_0$) estimated from the aforementioned MC-EM algorithm, and $C_{\widehat{\boldsymbol{\lambda}},k}$ is the normalization constant in the truncated density (7) under the MLE $\widehat{\boldsymbol{\lambda}}$. This test statistic also involves multiple integrations and it is difficult to evaluate. We again use a Monte Carlo method. In particular, we note that $f_{\boldsymbol{\lambda}}(\mathbf{O}, \mathbf{r}|k)/f_{\boldsymbol{\lambda}}(\mathbf{O}, \mathbf{r}, k) = 1/C_{\boldsymbol{\lambda},k}$ and, given $\boldsymbol{\lambda}$, $f_{\boldsymbol{\lambda}}(\mathbf{O}, \mathbf{r}|k)$ is easy to simulate from. We use a rejection sampling approach to simulate M sets of samples, say, $(\mathbf{O}^{(l)}, \mathbf{r}^{(l)})$, $l = 1, \ldots, M$, from $f_{\boldsymbol{\lambda}}(\mathbf{O}, \mathbf{r}|k)$, with $f_{\boldsymbol{\lambda}}(\mathbf{O}, \mathbf{r}, k)$ being the candidate distribution. The acceptance rate of this rejection sampling method is $C_{\boldsymbol{\lambda},k}$, and its empirical acceptance rate, say r, can be used to estimate $C_{\boldsymbol{\lambda},k}$. Thus, we approximate Λ by

$$\Lambda^* = \log\left[\frac{1}{M}\sum_{l=1}^{M} f_{\widehat{\boldsymbol{\eta}},\widehat{\zeta}}(Y, \boldsymbol{\delta}|\mathbf{O}^{(l)}, \mathbf{r}^{(l)}, k)\right] + \log r$$
$$- \left[n\log n + n^*\log n^* - (n+n^*)\log(n+n^*)\right.$$
$$\left. + \sum_{i=1}^{N}\log f(\mathbf{y_i})\right]. \qquad (16)$$

Since the parameters $\boldsymbol{\lambda} = (\lambda_o, \lambda_r)$ are nuisance parameters in the test and they only exist under the alternative hypothesis H_1, the likelihood test statistic does not follow the asymptotic chi-squared distribution, and the usual chi-squared test is no longer valid (Davies 1977, 1987). Following the suggestion of (Davies 1987), we consider a simulation-based Monte Carlo testing approach (Dwass 1957). Specifically, we sample L sets of k clusters under the null hypothesis (no true sources in the study region) and compute for each set the test statistic according to (16). When L is large, the empirical distribution of these values provides a good approximation to the theoretical distribution of the test statistics Λ^* under the null hypothesis. We thus obtain the critical value from the L values for our simulation-based Monte Carlo test.

3.2.2 Determination of the Unknown Number of Clusters. The previous estimation and testing procedures are for a given number of sources k. In reality, k is unknown and needs to be determined from the observed data. We treat this as a model selection problem, and use the AIC criterion (Akaike 1974) and BIC criterion (Schwarz 1978), two widely used approaches, for such a purpose. However, when a model selection problem involves missing (or latent) variables, a direct application of the usual AIC or BIC criterion can be problematic; see Claeskens and Consentino (2008). Our numerical studies (not reported in the article) also confirm this observation. Claeskens and Consentino (2008) proposed a modified AIC criterion to overcome the problem and demonstrated its superior performance in the case involving missing (or latent) variables. In particular, using the Kullback–Leibler distance to measure the distance between the true data generating density and the model density used for describing the data, they derived a new criterion based on the Q function in the EM algorithm. This requires no additional effort in our case since the Q function is available from the EM algorithm in the model estimation step.

Following their approach, we use the modified AIC and BIC criteria as

$$\mathrm{AIC}(k)_{\mathrm{mod}} = -2Q_1(\widehat{\boldsymbol{\eta}}, \widehat{\zeta}|\widehat{\boldsymbol{\eta}}, \widehat{\zeta}) + 2(k+3), \qquad (17)$$
$$\mathrm{BIC}(k)_{\mathrm{mod}} = -2Q_1(\widehat{\boldsymbol{\eta}}, \widehat{\zeta}|\widehat{\boldsymbol{\eta}}, \widehat{\zeta}) + (k+3)\log(N), \qquad (18)$$

where $Q_1(\widehat{\boldsymbol{\eta}}, \widehat{\zeta}|\widehat{\boldsymbol{\eta}}, \widehat{\zeta})$ is the Q function in (15) at the value of the MLE $(\widehat{\boldsymbol{\eta}}, \widehat{\zeta})$ from the EM algorithm, $k+3$ is the number of parameters to be estimated. Given a dataset, several competing models may be ranked according to their AIC or BIC values.

With the aforementioned estimation, testing, and model selection steps, we summarize the approach for detecting clusters. Denote by $\mathcal{K}$ a preselected set of k's, which is computationally manageable but large enough to cover all potential choices of the correct number of clusters. For each k in $\mathcal{K}$, we apply the MC-EM algorithm to get the parameter estimates and calculate the modified AIC or BIC values. With the modified AIC or BIC values, we determine the number of clusters k. For the k chosen by AIC or BIC, the source locations and impact ranges are estimated and testings are performed.

3.3 Detection With Mixed-Type Sensors

The above method is based on the assumption that all sensors have the same characteristics, such as error rates, detection ranges, etc. In practice, we may use different types of sensors in a mobile sensor network with different characteristics. For example, conversations with law enforcement officials suggested a possible combination of a large number of lower quality sensors on a fleet of taxicabs and a limited number of higher quality sensors on police vehicles. If there are m types of sensors, the impact range of the jth nuclear source on the lth type sensors is assumed to be

$$\rho_j^{(l)} = c^{(l)}\rho_j^{(1)}, \quad l = 2, \ldots, m; \quad j = 1, \ldots, k,$$

where $\rho_j^{(1)}$ is the reference, $c^{(l)}$ is the ratio of the source impact range for the lth type sensors over that for the reference type sensors. We assume that $c^{(l)}$ is given and can be obtained from manufacturer's specification. The location and reading status data from the m types of sensors are $(\mathbf{Y}, \Delta) = \{(\mathbf{Y}^{(1)}, \ldots, \mathbf{Y}^{(m)}), (\boldsymbol{\delta}^{(1)}, \ldots, \boldsymbol{\delta}^{(m)})\}$, where $\mathbf{Y}^{(l)} = (\mathbf{y}_1^{(l)}, \ldots, \mathbf{y}_{N_l}^{(l)})$ is a $2 \times N_l$ matrix denoting the set of N_l positions from the lth type sensors and the vector $\boldsymbol{\delta}^{(l)} = (\delta_1^{(l)}, \ldots, \delta_{N_l}^{(l)})$ denoting the reading status for those sensors. For the lth type sensors, the density function from (9) is modified as

$$f_{\boldsymbol{\eta}^{(l)},\zeta^{(l)}}(\mathbf{y}, \delta|\mathbf{O}, \mathbf{r}, k)$$
$$= \left[\left(\zeta^{(l)}\right)^{1-\delta}\left(1-\zeta^{(l)}\right)^{\delta}\right]^{\mathbf{1}\{\mathbf{y}\notin\bigcup_{j=1}^{k} I_j\}}$$
$$\times \prod_{j=1}^{k}\left[\left(\eta_j^{(l)}\right)^{\delta}\left(1-\eta_j^{(l)}\right)^{1-\delta}\right]^{\mathbf{1}\{\mathbf{y}\in I_j\}} f(\mathbf{y}), \qquad (19)$$

where $\boldsymbol{\eta}^{(l)} = (\eta_1^{(l)}, \ldots, \eta_k^{(l)})$. Hence for each type of sensor, there is a set of parameters to be estimated.

The conditional joint density function of $\mathbf{Y}$, Δ given $\mathbf{O}$, $\mathbf{r}$ and k from (12) is modified as

$$f_{\boldsymbol{\eta},\boldsymbol{\zeta}}(\mathbf{Y}, \Delta|\mathbf{O}, \mathbf{r}, k)$$
$$= \exp\left\{\sum_{l=1}^{m}\left[\sum_{j=1}^{k} Z_j^{(l)}\log\eta_j^{(l)} + \sum_{j=1}^{k} Z_j^{*(l)}\log\left(1-\eta_j^{(l)}\right)\right.\right.$$
$$+ \left(n^{*(l)} - \sum_{j=1}^{k} Z_j^{*(l)}\right)\log\zeta^{(l)} + \left(n^{(l)} - \sum_{j=1}^{k} Z_j^{(l)}\right)$$
$$\left.\left.\times\ \log\left(1-\zeta^{(l)}\right) + \sum_{i=1}^{N_l}\log f\left(\mathbf{y_i}^{(l)}\right)\right]\right\}, \qquad (20)$$

where $\boldsymbol{\eta} = (\boldsymbol{\eta}^{(1)}, \ldots, \boldsymbol{\eta}^{(m)})$, $\boldsymbol{\zeta} = (\zeta^{(1)}, \ldots, \zeta^{(m)})$, $Z_j^{(l)} = \sum_{i=1}^{N_l} \delta_i^{(l)} \mathbf{1}_{\{\mathbf{y}_i^{(l)} \in I_j\}}$, and $Z_j^{*(l)} = \sum_{i=1}^{N_l} (1 - \delta_i^{(l)}) \mathbf{1}_{\{\mathbf{y}_i^{(l)} \in I_j\}}$ are the number of positive and negative readings from the lth type sensors inside the jth cluster, respectively, $n^{(l)} = \sum_{i=1}^{N_l} \delta_i^{(l)}$ and $n^{*(l)} = \sum_{i=1}^{N_l} (1 - \delta_i^{(l)})$ are the total number of positive and negative readings from the lth type sensors, respectively.

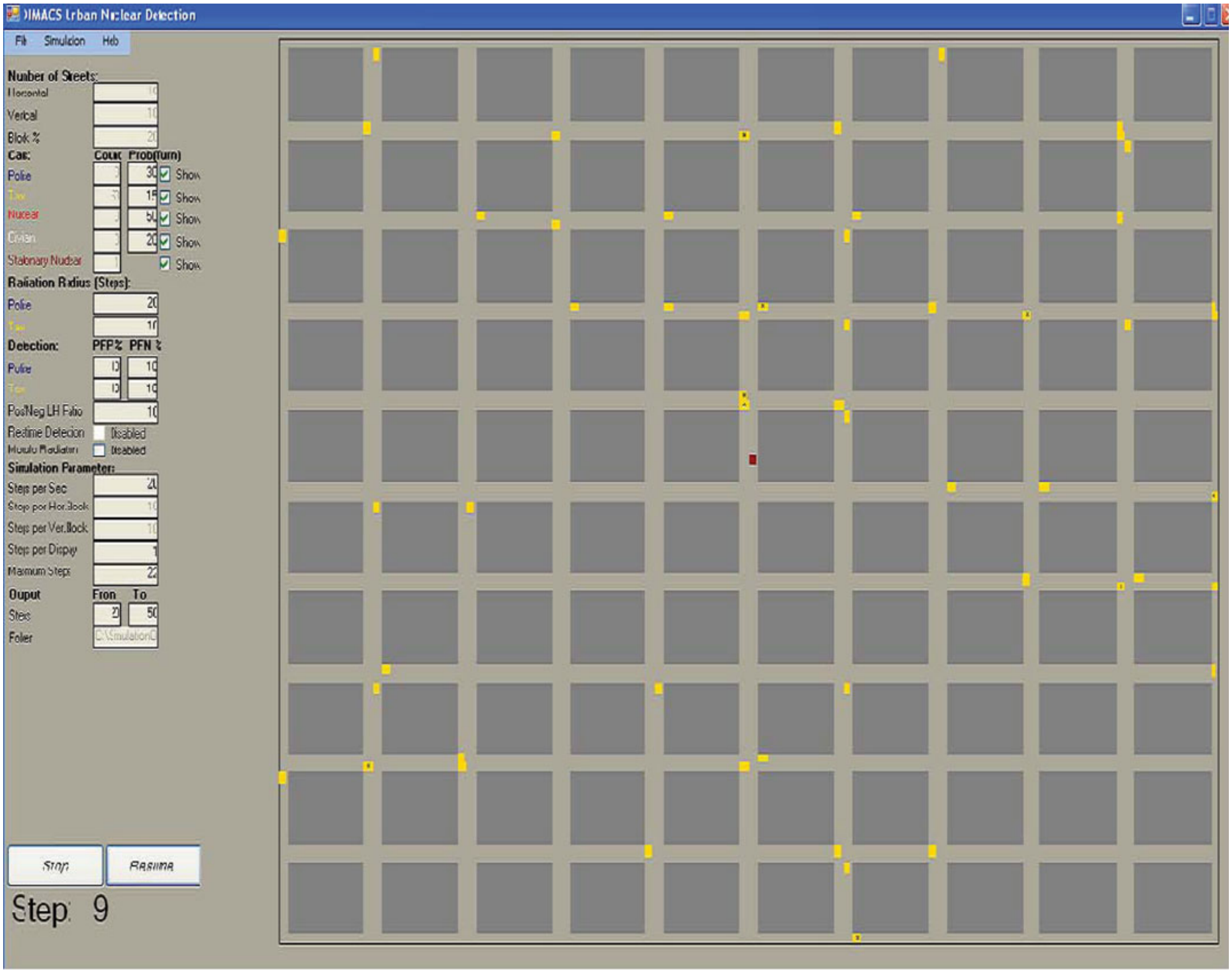
The cluster detection steps using this model are similar to those in Section 3 for the model with a single sensor type (9). Details are omitted in this article and are available upon request.

4. SIMULATION STUDIES

In this section, we present simulation studies to demonstrate that a mobile sensor network with analysis procedures using the proposed latent source modeling method can effectively detect either single or multiple nuclear sources. Intuitively, factors, such as range, error rates, and number of sensors will directly affect how well nuclear sources can be detected. These factors form a set of network parameters, and we study how they relate to the method's performance in detecting nuclear sources. In Section 4.1, we describe a Visual Basic graphical simulation tool that we used to generate our data. In Section 4.2, we apply the cluster detection methods developed in Section 3 to study the signal data simulated using the graphical tool under various scenarios.

4.1 Simulation Design and Evaluation

We designed a mobile sensor network in an area of similar size as downtown Manhattan in New York City. The study region was set to an area of 25 by 25 blocks. Each block was a square with side length of 200, which represents 200 feet in real distance. To achieve this, a street grid-based graphical tool was developed using Visual Basic Version 9.03 (Microsoft, Inc.) for visualization of simulated traffic patterns in a metropolitan area. The tool supports multiple control parameters such as numbers of streets in either horizontal or vertical direction, size of street blocks and types, quantities and speed of vehicles, a probability of turning upon reaching an intersection, etc. Figure 1 is a snapshot of the simulation tool. Since the number of registered taxicabs is more than 13,000 in New York City (Schaller Consulting 2006), we assume that there are at least 1500 participating vehicles (such as taxicabs, police cars, fire trucks,



)IMACS Urban Nuclear Detection

Number of Steets:
Horizontal 1(
Vertical 1(
Blok % 2(
Car: Count Prob(Turn)
Polie 3(30 ☑ Show
Tax 5(15 ☑ Show
Nucear 5(50 ☑ Show
Civian 2(20 ☑ Show
Stationary Nuclear 1 ☑ Show
Radiation Radius (Steps):
Polie 2(
Tax 1(
Detection: PFP % PFN %
Pulie 1(1(
Tax 1(1(
PosNeg LH Ratio 1(
Realtime Detecion □ Disabled
Moulu Radiation □ Disabled
Simulation Parameter:
Steps per Sec 2(
Steps per Hor.Block 1(
Steps per Ver.Block 1(
Steps per Display 1
Maximum Steps 2(
Ouput From To
Stes 2(5(
Foler C:\SimulationD

Step 9

Figure 1. Snapshot of the simulation tool.

security cars, etc.) in the downtown area. In the simulation tool, all vehicles are confined within the street grids.

In our simulation, the set of network parameters was selected as follows. The source impact range is set to 150 or 200 feet, roughly corresponding to 0.75 and 1.0 of a block length, respectively. Note that current technology does not achieve such high source impact range numbers, but we are seeking to understand the impact of potential future technology. The current technology provides false alarm rates that are relatively low and future technology can be expected to make them even lower. Medalia (2010) reported that the number of false alarms in 2005 from the radiation sensors at the Lincoln Tunnel between New York City and New Jersey were 50–60 out of 43 million vehicles. And it is expected that the available detecting devices will have less than 3% false alarm rate. Based on this, sensitivity and specificity are set to $(\eta, \zeta) = (0.95, 0.95)$ or $(0.98, 0.98)$, and the number of participating vehicles mounted with sensors is set from 500 to 2500 with increments of 500. In summary, there are two impact ranges, two error rates, and five sensor numbers, from which we can construct $2 \times 2 \times 5 = 20$ sets of network parameters. For each set of parameters, we generated random positions (on the street grids) of participating vehicles and randomly placed a single or multiple sources anywhere in the study region. For multiple sources, we assumed that they had the same energy spectrum. Based on models (3) and (5), we assigned a probability of positive detection for each sensor and activated it accordingly. The positions of the sensors and their nuclear detection signals, either positive or negative, were collected. These geographic positions and the detection signals formed an observed dataset. We then applied the proposed latent source modeling approach to detect clusters. Simulations were repeated 500 times in each setting of network parameters.

To measure performance in our simulation study, we consider two criteria: (i) at least one of the clusters is statistically significant at a given level (5% in our case); (ii) the detected cluster regions cover the true nuclear sources. The first criterion is related to the "hypothesis testing" power (abbreviated as "testing power" in this article), which is computed as the percentage of times that at least one of the clusters is statistically significant at 5% level for the test described in Section 3.2.1. The empirical critical value (obtained via the Monte Carlo testing method) is used. We also computed the "detection power" as the percentage of times that the proposed algorithm resulted in a correct detection, satisfying both criteria above. The latter is a more stringent performance measure than the testing power alone.

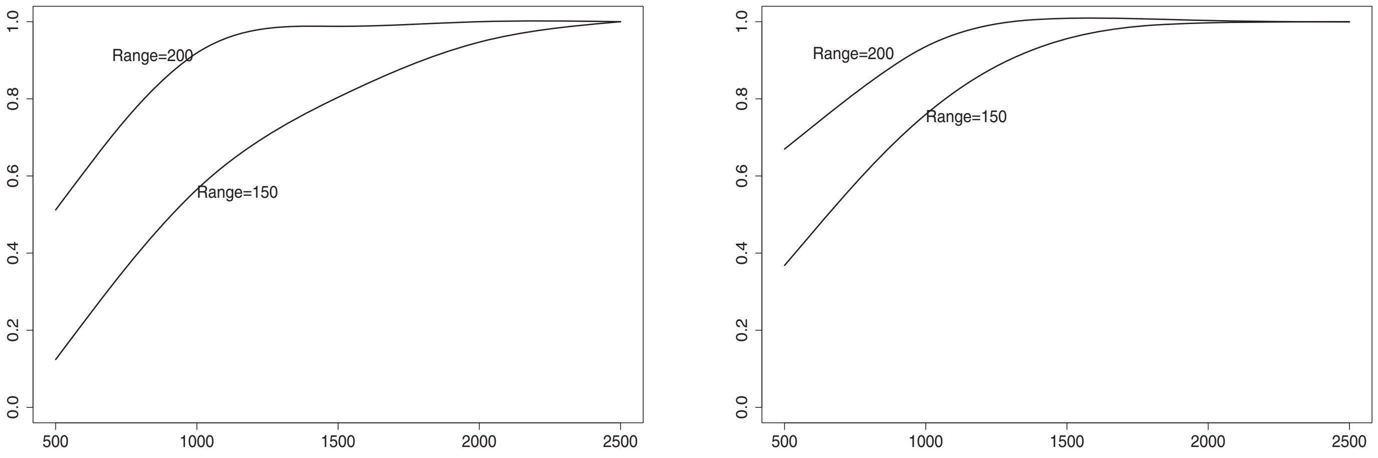
4.2 Performance of Mobile Sensor Networks

4.2.1 Power for Detecting One Source. We first studied the effect of sensors' impact range and the number of sensors for a single source randomly placed in the study region. Sensitivity and specificity were set at (0.95, 0.95). With 5% level tests, we plot detection power in Figure 2(a). The result shows that we need 1000 200-foot sensors to achieve 92% detection power, while 1000 150-foot sensors yield about 57%. When the number of sensors reaches 1500 or above for 200-foot sensors and 2500 and above for 150-foot sensors, respectively, the detection power is close to 100%.

We repeated the same simulations with sensitivity and specificity of (0.98, 0.98) and obtained the power curves in Figure 2(b). This network of higher quality sensors achieved much higher power. For example, the detection power increased from 57% to 76% when both networks have 1000 sensors of 150-feet range.

The first half of Table 1 presents performance measures under four network settings (two error rates and two source impact ranges), with the number of sensors fixed at 1500, for a single randomly placed source. We use 1500 sensors here because it is feasible for a city as large as New York City to equip this many taxicabs with nuclear detection sensors under a reasonable budget. In the table, we report detection power, testing power, and sizes of hypothesis tests using the Monte Carlo method.

We also compared our detection algorithm with a varying-window-size scan statistic approach from the free SaTScan software (*www.satscan.org*) and a stepwise regression method (Demattei, Molinari, and Daures 2006, 2007). Using the same settings, we list the detection power, testing power, and sizes from the two comparative methods in the top half of Table 1.



(a) $\eta = 59.0$, $\zeta = 0.95$ (b) $\eta = 0.98$, $\zeta = 0.98$

Figure 2. Detection power of a mobile sensor network with one source from one simulation time slice for New York City. X-axis represents the numbers of sensors and Y-axis the detection power. (a) Two different sensor ranges with assumed $\eta = 0.95$, $\zeta = 0.95$. (b) Two different sensor ranges with assumed $\eta = 0.98$, $\zeta = 0.98$.

Table 1. Summary of detection powers, hypothesis testing powers, and sizes in various cases when true $k = 1$ and 2 and the number of sensors is 1500

			Latent source model		SaTScan		Step regression	
	(η, ζ)	Range	Detection	Testing	Detection	Testing	Detection	Testing
True $k = 1$								
Power	(0.95,0.95)	150	80.4%	85.2%	5.2%	8.5%	49.0%	66.1%
		200	98.8%	99.2%	48.6%	51.0%	85.0%	92.5%
Size		–	–	5.4%	–	6.6%	–	33.0%
Power	(0.98,0.98)	150	96.7%	95.2%	5.4%	13.5%	80.9%	94.0%
		200	100.0%	100.0%	95.0%	98.4%	95.5%	97.8%
Size		–	–	5.6%	–	4.8%	–	35.0%
True $k = 2$								
Power	(0.95,0.95)	150	65.3%	95.8%	1.0%	8.4%	–	41.8%
		200	97.2%	100.0%	32.4%	52.6%	–	61.8%
Size		–	–	5.9%	–	5.8%	–	35.6%
Power	(0.98,0.98)	150	82.5%	99.6%	2.5%	14.0%	–	83.2%
		200	100.0%	100.0%	88.5%	99.6%	–	98.2%
Size		–	–	5.2%	–	5.0%	–	37.5%

Our algorithm produces better results in all cases, especially in the case of shorter sensor range (i.e., weaker sensors). Our method and SaTScan can control the Type I error around 5%, while the stepwise regression method makes about 30% false alarm errors.

4.2.2 Power for Detecting Two Sources. To demonstrate the performance of a mobile sensor network when there is more than one source, we studied the same network settings as in Section 4.2.1 but with two sources. Here, the two sources were placed randomly in the study region. We report the detection performance in the bottom half of Table 1. We observe when we have two clusters to detect, the detection power is lower than that for only one cluster under the same network setting since the detection is required to cover both clusters, while the testing power is higher since the testing is significant as long as one of the clusters is significant from the testing process.

In contrast to the simultaneous detection of multiple sources in our method, the scan statistics uses a sequential approach. When it detects a significant primary cluster, it replaces the data inside the primary cluster with the average outside of it, then does the detection to pick up a secondary cluster. As a result, we observe from Table 1 that SaTScan achieves much lower power for detecting two sources than our method. For the stepwise regression method, we are not able to calculate the detection power for the multiple cluster case since the numeric summary of the locations and sizes of detections is not available. Also the stepwise regression method produces a high level of false alarm errors as in the one cluster case.

4.2.3 Power for Detection Using Mixed-Type Sensors. In what follows, we selected the setting of 1500 sensors with detection range 150 feet and 95% for both η and ζ as a benchmark in further studies. This setting achieved detection powers of 80.4% in one cluster and 63.4% in two clusters. To assess the value of using higher quality sensors, we substituted a small number (50, 75, 100) of sensors with higher quality ones of both sensitivity and specificity at 0.98. The source impact ranges of the new sensors were set at 200, 250, or 300 feet.

The detection results for the combination of the two types of sensors for one cluster are presented in Figure 3. We observed a significant improvement in detection power as compared to the results in Table 1. For example, with a substitute of 50 high-quality sensors (200-foot detection range and 2% error rates), the detection power increases from 80.4% to 85.6% (a 6.5% relative increase). We can observe that substituting with higher quality sensors increase detection power significantly (i.e., with a simple t-test or some other approaches). Among the various cases with high-quality sensors, the improved detection power may not differ significantly. For example, the detection power increases marginally from 85.2% to 86.9% with the two settings for 50 high-quality sensors—one with 200-foot range and the other with 300 foot. The reason for this is, we speculate, that the increase of the detection power levels off and is not sensitive to some small changes.

4.2.4 Determining the Number of Sources. To evaluate the proposed AIC and BIC criteria for determining the number of clusters, we used the benchmark setting of the mobile sensor network described at the beginning of Section 4.2.3 and put 0–4 sources in the study region. Then we followed the procedures in Section 3.2.2 and defined $\mathcal{K}$={0, 1, 2, 3, 4} as the preselected set of k's. For each $k \in \mathcal{K}$, we calculated the values of the modified version $\mathrm{AIC(k)}_{\mathrm{mod}}$ and $\mathrm{BIC(k)}_{\mathrm{mod}}$ from (17) and (18), and estimated the number of clusters accordingly.

Table 2 summarizes the model selection results. The modified AIC and BIC methods seem to work well. The correct number of sources was always selected most often in all settings. In addition, BIC, with a larger penalty term, appears to perform slightly better than AIC. For those wrongly chosen cluster numbers, the majority tends to be one smaller than the true number (i.e., $k - 1$). These findings are similar to the model selection results reported in Pan (2001), Sun (2008), and Xie et al. (2009).

5. DISCUSSION

In this article, we propose a robust mobile sensor network and develop a statistical algorithm to provide consistent and

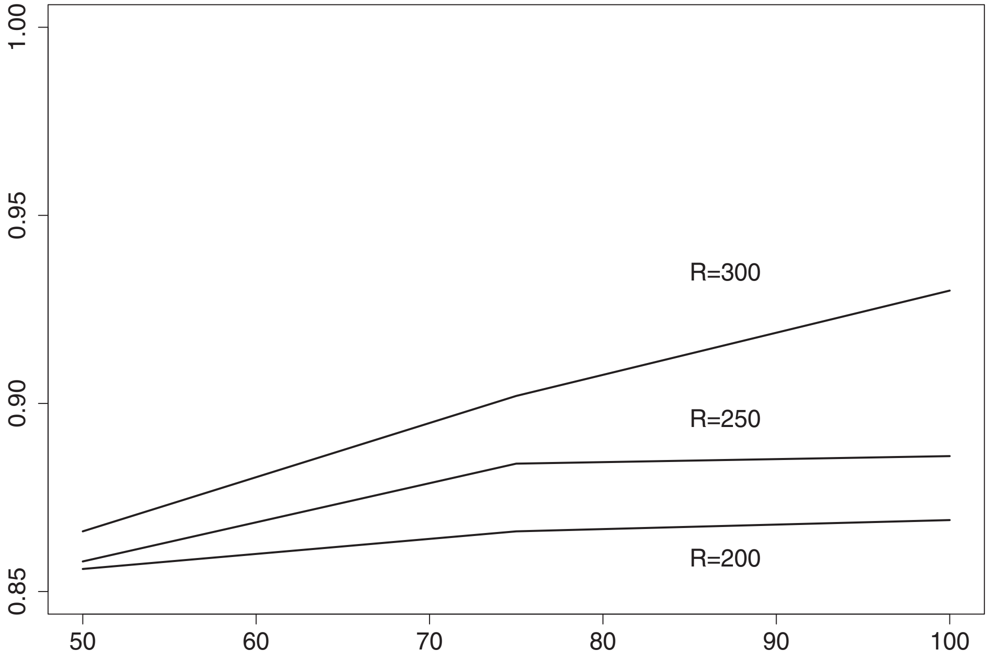


Figure 3. Power improvement with substitution of high-quality sensors at the benchmark setting of 1500 sensors with 150-foot detection range and 5% error rates. The X-axis represents the number of high-quality sensors with 2% error rates and different detection ranges. The Y-axis represents the detection power.

pervasive surveillance for nuclear materials in major cities. Simulation studies under the settings of Section 4 suggest that the proposed network and statistical methods can provide an effective tool to detect nuclear sources placed in a spatial region. This study aims to provide a forward-looking design and implementation of a detection capability from a sensor management point of view.

In this study, we assume that the clusters are nonoverlapping. In practice, when two nuclear sources are close enough, a sensor can be activated by either or both of them. The clusters formed around the sources might be overlapping. Chap. 5 of Cheng (2010) relaxed the nonoverlapping cluster assumption and studied detection of overlapping clusters. Though the detection power increases slightly when the two or more sources are nearby, the computational cost is much higher. In practice, it seems reasonable to use one cluster to represent multiple adjacent sources.

The mobile sensor network proposed here can be supplemented by static sensors. In fact, in most cases such a supplement is necessary to cover locations with sparse or zero traffic, such as a large park in the city. Our detection algorithm can be extended to handle both types of networks since the static sensors' positions and states can be merged and processed with those of mobile sensors. Our models have only considered two-dimensional regions and have disregarded the possibility that a stationary source might be located above the ground in a building. In practice, the source can be treated with slightly weaker energy intensity placed at the ground level.

Although the nuclear detection algorithm developed here is computationally intensive due to multiple levels of Monte Carlo and EM iterations, it is feasible for real world application. It takes less than 1 min on average on a personal desktop computer with 2.6 GHz processors. Thus, a slightly higher performance computing device should be able to provide meaningful real time nuclear detection.

Finally, we do not consider in the article the possibility of shielding of the nuclear energy by the buildings or other materials. The shielding issue is very complex. It depends on the geographic distribution and shapes of the buildings and other large objects in the region. It also depends on the packing

Table 2. Model selection evaluation. Reported in the table are the percentages of times that the AIC or BIC criterion selects the specified number of clusters. The bold entry indicates the most frequently selected number of clusters in each case

	AIC Estimated					BIC Estimated				
True k	$k = 0$	$k = 1$	$k = 2$	$k = 3$	$k = 4$	$k = 0$	$k = 1$	$k = 2$	$k = 3$	$k = 4$
0	**100.0**	0.0	0.0	0.0	0.0	**100.0**	0.0	0.0	0.0	0.0
1	26.2	**64.5**	6.8	2.5	0.0	25.4	**68.2**	3.4	3.0	0.0
2	6.8	19.0	**48.2**	25.6	0.4	5.4	20.2	**65.2**	9.0	0.2
3	1.8	2.4	14.8	**48.8**	32.2	1.8	2.4	24.6	**61.8**	9.4
4	2.0	7.6	25.2	23.2	**42.0**	2.2	3.0	11.2	20.4	**62.2**

materials used to transport the nuclear materials. Nelson and Sokkappa (2008) provided a detailed report and developed a generic packaging model to study the impact of shielding for several nuclear and packing materials. Further study is needed to understand the impact of shielding, especially in metropolitan areas. It is an interesting and challenging topic for future research.

[Received June 2011. Revised March 2013.]

REFERENCES

Akaike, H. (1974), "A New Look at the Statistical Model Identification," *IEEE Transactions on Automatic Control*, 19, 716–723. [908]

Akyildiz, I., Su, W., Sankarasubramaniam, Y. E., and Cayiric, E. (2002a), "Survey on Sensor Networks," *IEEE Communications Magazine*, 40, 102–114. [904]

——— (2002b), "Wireless Sensor Networks: A Survey," *Computer Networks*, 38, 393–422. [904]

Balakrishnan, N., and Koutras, M. (2001), *Runs and Scans With Applications*, New York: Wiley. [903]

Boros, E., Fedzhora, L., Kantor, P. B., Saeger, K., and Stroud, P. (2009), "A Large-Scale Linear Programming Model for Finding Optimal Container Inspection Strategies," *Naval Research Logistics*, 56, 404–420. [904]

Cheng, J. Q. (2010), "Bayesian Methods for Non-Standard Missing Data Problems," Ph.D. dissertation, Department of Statistics, Rutgers University. [905,912]

Claeskens, G., and Consentino, F. (2008), "Variable Selection With Incomplete Covariate Data," *Biometrics*, 64, 1062–1069. [908]

Davies, R. B. (1977), "Hypothesis Testing When a Nuisance Parameter is Present Only Under the Alternative," *Biometrika*, 64, 247–254. [908]

——— (1987), "Hypothesis Testing When a Nuisance Parameter is Present Only Under the Alternative," *Biometrika*, 74, 33–43. [908]

Demattei, C., Molinari, N., and Daures, J. (2006), "Spatclas: An R Package for Arbitrarily Shaped Multiple Spatial Cluster Detection for Case Event Data," *Computer Methods and Programs in Biomedicine*, 84, 42–49. [903,910]

——— (2007), "Arbitrary Shaped Multiple Spatial Cluster Detection for Case Event Data," *Computational Statistics and Data Analysis*, 51, 3931–3945. [903,910]

Dembo, A., and Karlin, S. (1992), "Poisson Approximation for *r*-Scan Processes," *Annals of Applied Probability*, 2, 329–357. [903]

Denison, D., and Holmes, C. (2001), "Bayesian Partitioning for Estimating Disease Risk," *Biometrics*, 57, 143–149. [903]

Diggle, P., Rowlingson, B., and Sun, T. L. (2005), "Point Processes Methodology for On-Line Spatio-Temporal Disease Surveillance," *Environmetrics*, 16, 423–434. [903]

Dwass, M. (1957), "Modified Randomization Tests for Nonparametric Hypothesis," *Annals of Mathematical Statistics*, 28, 181–187. [908]

Elsayed, E. A., Young, C. M., Xie, M., Zhang, H., and Zhang, Y. (2009), "Port-of-Entry Inspection: Sensor Deployment Policy Optimization," *IEEE Transactions on Automation Science and Engineering*, 6, 265–276. [904]

Escobar, M. D., and West, M. (1995), "Bayesian Density Estimation and Inference Using Mixtures," *Journal of the American Statistical Association*, 9, 577–588. [903]

FEMA (2008), "Are You Ready? Guide," *An In-Depth Guide to Citizen Preparedness*, IS-22. Available at *http://www.fema.gov/areyouready/nuclear_blast.shtm*. [904]

Ferguson, T. S. (1973), "A Bayesian Analysis of Some Nonparametric Problems," *The Annals of Statistics*, 1, 209–230. [903]

Fu, J. C., and Lou, W. Y. (2003), *Distribution Theory of Runs and Patterns and Its Applications: A Finite Markov Chain Embedding Approach*, Singapore: World Scientific. [903]

Gangnon, R., and Clayton, M. (2000), "Bayesian Detection and Modeling of Spatial Disease Maps," *Biometrics*, 56, 922–935. [903,905]

——— (2003), "A Hierarchical Model for Spatially Clustered Disease Rates," *Statistics in Medicine*, 22, 3213–3228. [903]

Ghosh, M., Natarajan, K., Waller, L. A., and Kim, D. (1999), "Hierarchical Bayes GLMs for the Analysis of Spatial Data: An Application to Disease Mapping," *Journal of Statistical Planning and Inference*, 75, 305–318. [903]

Glaz, J., and Balakrishnan, N. (1999), *Scan Statistics and Applications*, Boston, MA: Birkhauser. [903]

Glaz, J., Naus, J., and Wallenstein, S. (2001), *Scan Statistics and Applications*, New York: Springer. [903,906]

Hochbaum, D. (2009), "The Multi-Sensor Nuclear Threat Detection Problem," in *Proceedings of the Eleventh INFORMS Computing Society (ICS) Conference*, eds. J. Chinneck, B. Kristjansson, and M. Saltzman, New York: Springer, pp. 389–399. [904]

Knorr-Held, L., and RaBer, G. (2000), "Bayesian Detection of Clusters and Discontinuities in Disease Maps," *Biometrics*, 56, 13–21. [903]

Kulldorff, M., and Nagarwalla, N. (1995), "Spatial Disease Clusters: Detection and Infection," *Statistics in Medicine*, 14, 799–810. [903,905]

Lawson, A. (1995), "Markov Chain Monte Carlo Methods for Putative Pollution Source Problems," *Environmental Epidemiology*, 14, 2473–2486. [903]

Lo, A. Y. (1984), "On a Class of Bayesian Nonparametric Estimates: I. Density Estimates," *The Annals of Statistics*, 12, 351–357. [903]

McCullagh, P., and Yang, J. (2008), "How Many Clusters," *Bayesian Analysis*, 3, 101–120. [903]

Medalia, J. (2010), "Detection of Nuclear Weapons and Materials: Science, Technologies, Observations," *Congressional Research Service*, R40154. Available at *http://www.fas.org/sgp/crs/nuke/R40154.pdf*. [910]

Naus, J. (1966), "A Power Comparison of Two Sets of Non-Random Clustering," *Technometrics*, 8, 493–517. [903]

Naus, J., and Wallenstein, S. (2004), "Multiple Window and Cluster Size Scan Procedures," *Methodology and Computing in Applied Probability*, 6, 389–400. [903]

Neal, R. M. (2000), "Markov Chain Sampling Methods for Dirichlet Process Mixture Models," *Journal of Computational and Graphical Statistics*, 9, 249–265. [903]

Nelson, K., and Sokkappa, P. (2008), "A Statistical Model for Generating a Population of Unclassified Objects and Radiation Signatures Spanning Nuclear Threats," Technical Report LLNL-TR-408407, Lawrence Livermore National Laboratory (LLNL). Available at *http://www.osti.gov/energycitations/product.biblio.jsp?query_id=1&page=0&osti_id=947761*. [913]

Pan, W. (2001), "Akaike Information Criterion in Generalized Estimation Equations," *Biometrics*, 57, 120–125. [911]

Purdue University (2008, January 22), "Cell Phone Sensors Detect Radiation to Thwart Nuclear Terrorism," *Purdue University News Service*. Available at *http://news.uns.purdue.edu/x/2008a/080122FischbachNuclear.html*. [904]

Rodriguez, A., Sunson, D. B., and Gelfand, A. E. (2008), "The Nested Dirichlet Process" (with discussion), *Journal of the American Statistical Association*, 103, 1131–1144. [906]

Schaller Consulting (2006, March), "The New York City Taxicab Fact Book," *NYC Taxi and Livery Issues*. Available at *http://www.schallerconsult.com/taxi*. [909]

Schwarz, G. (1978), "Estimating the Dimension of a Model," *The Annals of Statistics*, 6, 461–464. [908]

Stroud, P. D., and Saeger, K. J. (2003), "Enumeration of Increasing Boolean Expressions and Alternative Digraph Implementations for Diagnostic Application," in *Proceedings of Computer, Communication and Control Technologies* (Vol. 5), eds. H. Chu, J. Ferrer, T. Nguyen, and Y. Yu, pp. 328–333. [904]

Su, X., Wallenstein, S., and Bishop, D. (2001), "Non-Overlapping Clusters: Approximation Distribution and Application to Molecular Biology," *Biometrics*, 57, 420–426. [903]

Sun, Q. L. (2008), "Statistical Modeling and Inference for Multiple Temporal or Spatial Cluster Detection," Ph.D. dissertation, Department of Statistics, Rutgers University. [903,905,906,911]

Tanner, M. (1993), *Tools for Statistical Inference*, New York: Springer-Verlag. [907]

Teh, Y. W., Jordan, M. I., Beal, M. J., and Blei, D. M. (2006), "Hierarchical Dirichlet Processes," *Journal of American Statistical Association*, 101, 1566–1581. [903]

Waller, L. A., Carlin, B. P., Xia, H., and Gelfand, A. E. (1997), "Hierarchical Spatio-Temporal Mapping of Disease Rates," *Journal of the American Statistical Association*, 92, 607–617. [903]

Wein, L., Wilkins, A., Baveja, M., and Flynn, S. (2006), "Preventing the Importation of Illicit Nuclear Materials in Shipping Containers," *Risk Analysis*, 26, 1377–1393. [904]

Xie, M., Sun, Q., and Naus, J. (2009), "A Latent Model to Detect Multiple Temporal Clusters," *Biometrics*, 65, 1011–1020. [903,905,911]