Chapter 4

4-1 Introduction[8]

Spatial surveillance is a key component of monitoring programs which provide an early detection capability of disease and pest incursions as well as informing assessments of plant and animal health status for trade purposes. International standards for phytosanitary measures and guidelines for surveillance have been established under the International Plant Protection Convention (FAO 1998). These guidelines distinguish between two broad classes of surveillance: specific surveys in which information is obtained on a particular pest over a relatively narrowly defined spatial-temporal extent; and general surveillance activities in which information is gathered on one or more pests over a wider area and from many sources - including specific surveys (Pheloung 2004). Some examples of foreign organisms that are of concern include: Siam weed (*Chromolaena odorata*); papaya fruit fly (*Bactrocera papayae*); red imported fire ant (*Solenopsis invicta*); branched broomrape (*Orobanche ramosa*) and kochia (*Bassia scoparia*) (Pheloung 2005).

Plant pests take a variety of forms including insects, weeds, fungi, bacteria, viruses and other harmful organisms and usually find their way into the country via trade and travel (Pheloung 2004). Of particular concern is Australia's vulnerability to fruit fly and since 2006 there has been renewed interest in developing a strategy to underpin a national approach to the management of this pest (http://www.planthealthaustralia.com.au/fruitfly/public.asp?pageID=243). This national fruit fly strategy (NFFS) builds upon the successful national fruit fly trapping program which targets exotic fruit fly pests (*Bactrocera spp.*) entering through international pathways at ports (Pheloung 2005).

[8] *The illustrative examples and motivating context used in this chapter relate to the detection of an equine influenza outbreak. It is acknowledged that the methodology is potentially better suited to the detection of plant pest and disease outbreaks. Unfortunately our requests to access the Australian Plant Pest Data Base were unsuccessful as were attempts to interface with the CRC Plant Biosecurity Surveillance research program.*

With respect to animal diseases, a number of potential and serious risks exist including: Avian Influenza (or Bird Flu); Bovine Spongiform Encephalopathy (BSE); Foot and Mouth Disease (FMD); Equine Influenza (EI); Rabies; and Varroa mite. Australia has developed a number of emergency response plans as well as a spatial and textual, web based software application tool called BioSIRT (Biosecurity Surveillance Incident Response and Tracing).

While incident response plans and tools are vital components of a combative strategy, it has been noted that by the time an incursion is detected, the prospects for eradication are very poor and prohibitively expensive (Pheloung 2004). Nairn et al. (1996) and others have long advocated strategies based on avoidance rather than eradication. Fox et al. (2009) noted that surveillance programs for monitoring invasive plants were expensive yet budgets allocated for this purpose were invariably "highly constrained". Under such circumstances there is a clear need to allocate scarce monitoring resources in the most effective way possible. Fox et al. (2009) also observed that previous attempts at 'optimisation' utilised economic tools that did not have any spatial or temporal representation.

In the following sections of this chapter we outline a mathematical programming approach to the identification of an 'optimal' configuration of 'sensors' as part of a general biosecurity surveillance program. A sensor in this context is defined broadly as any instrument, method, procedure, or device that acquires information or samples related to the biosecurity threat under investigation. We acknowledge that our (artificial) example of sensor network optimisation for detecting an EI outbreak may not be the most suitable candidate for our methodology. However, due to our inability to access realistic plant data (see footnote on previous page) and the ready availability of EI and demographic data, we used the latter for illustrative purposes. The suitability or otherwise of the example does not diminish the integrity of the proposed strategy.

4-2 A surveillance network for EI

On August 24, 2007 a disease strategy for equine influenza (EI) was released by Animal Health Australia (2007). The AHA report noted that "there has been no occurrence of EI in Australia … and vaccination is not practised." Regrettably, that situation changed with the first detection of EI in the same month in the Sydney area. The disease spread rapidly through northern NSW into Queensland where it concentrated in the Brisbane region (DPI 2008). The NSW situation by September 12, 2007 is shown in Figure 1.

Equine influenza (EI) is an acute, highly contagious (having an almost 100% infection rate), viral disease which spreads rapidly in horses and other equine species (NSW DPI 2008). Humans are not affected by this virus although they can be responsible for its spread. Most animals exposed to the virus will show signs within a period of 1-5 days. Typically, an infected animal will develop a fever, a dry hacking cough and have a suppressed appetite. Recovery usually takes 2 to 3 weeks. Being a virus, there is no effective treatment and the risk of secondary infections, such as pneumonia is high.

Figure 50. Detection of EI in NSW as at September 12, 2007.
Source: http://www.dpi.nsw.gov.au/__data/assets/pdf_file/0009/179406/IP-equine-influenza-map-nsw-12-sept-07.pdf

With respect to biosecurity, responsible agencies in Australia divide their operations into pre-border, border, and post-border monitoring and surveillance activities. A recurring and important issue is where to place limited resources and effort so as to maximize the effectiveness of the surveillance program. For example, given the map of Figure 50 together with other ancillary information about the population at risk (size, geographic extent, spatial aggregation, susceptibility etc.) what configuration (numbers, types, and placement) of monitoring activities delivers the 'best' surveillance outcome? Clearly, terms like 'best' need to be defined as does a metric of surveillance utility. One optimization criterion might, for example, be the maximization of the probability of detection. However for our problem formulation, we make the following assumptions:

- detection probability is a non-decreasing function of monitoring effort;
- detection probability is heterogeneous in space and time;
- the cost of monitoring is directly proportional to the amount of resources devoted to the monitoring program;
- the total cost of monitoring is constrained.

We consider a formulation of the general surveillance design problem (ie. the optimal placement of sensors in a distributed network) as a constrained integer linear programming problem (ILP). The objectives are to:

i. Recast a conceptual understanding of the monitoring network for EI as a constrained optimisation problem;
ii. Provide realistic models of 'risk' over a 2D space;
iii. Code the problem for solution using ILP;
iv. Solve an artificial example problem.

In the following sections we provide details associated with items (i) to (iii) above and in particular, demonstrate the feasibility of item (iv). To this extent, we regard (i) to (iv) as forming the basis of a prototype 'system' which we has been developed to a proof-of-concept stage. The challenge that lies ahead is to implement this system on a suitably calibrated real problem.

4-3 The Maximal Covering Location Problem (MCLP)

The optimal sensor location problem for building an EI surveillance network is a variant of a problem known in the Operations Research (OR) literature as either the maximal covering problem (MCP) or the maximal covering location problem (MCLP)

(Church and ReVelle 1974, Underhill 1994, Downs and Camm 1996, Chakrabarty et al. 2002). Numerous implementations of the MCLP exist including the placement of guards in an art gallery (O'Rourke 1987), the siting of regional health facilities and services (Eaton et al. 1977, Pirkul and Schilling 1988), optimal placement of ambulance services (Chuang and Lin 2007) and conservation reserve site selection (Önal 2003). More recently the omnipotent threat to homeland security has focussed the attention of some researchers on optimal network designs to detect terrorist activities (Anderson et al. 2007).

The objective of an MCP/MCLP is to determine the configuration of sensors of varying types (in terms of cost, monitoring range, detection capabilities) that achieves mandated levels of surveillance accuracy at minimum total cost. When sensors can be located anywhere on a plane the problem is referred to as the planar maximal covering location problem (PMLCP) (Church 1984). Church and ReVelle (1974) provided the first mathematical formulation of the MCLP for the problem of siting public facilities. In this context a candidate facility site 'covers' a demand node if it is within some maximum distance from that node. Mathematically, we have

Maximize: $$\mathbb{Z} = \sum_{i=1}^{m} c_i \, y_i \tag{4.1}$$

s.t. $$\sum_{j=1}^{n} a_{ij} x_j \geq y_i \qquad \forall\, i \in I \tag{4.2}$$

$$\sum_{j=1}^{n} x_j = k \tag{4.3}$$

$$y_i = \{0,1\} \qquad \forall\, i \in I \tag{4.4}$$

$$x_j = \{0,1\} \qquad \forall\, j \in J \tag{4.5}$$

In this formulation, $J = \{ j \mid j = 1,\ldots,n \}$ represents the set of candidate facility sites and $I = \{ i \mid i = 1,\ldots,m \}$ denotes the set of demand nodes; c_i is the population to be served at demand node *i*. Our (binary) decision variables are x_j such that $x_j = 1$ if site *j* is chosen for a facility and $x_j = 0$ otherwise. The constraint represented by equation 4.3 requires that *k* facilities / sites are to be selected. The indicator variables y_i assume a value of unity if demand node *i* is covered by some facility *j* and assume a value of zero otherwise. Similarly, variable a_{ij} is unity if the shortest distance between demand node *i* and facility *j* does not exceed some prescribed maximum permissible distance.

Mathematicians refer to MCLPs as being NP-hard. While not wishing to go into the complexities of mathematical algorithms, a problem is assigned to the NP (nondeterministic polynomial time) class if it is solvable in polynomial time. A problem is NP-hard if an algorithm for solving it can be translated into one for solving any NP-problem. NP-hard therefore means 'at least as hard as any NP-problem', although it might in fact be harder (Weisstein, *undated*). Practical approaches to solving the MCLP problem include mathematical programming methods, genetic algorithms, graph theory, complete enumeration and heuristics. Genetic algorithms (GAs) have become popular in recent years as a way of solving large optimisation problems such as the MCLP (Arakaki and Lorena 2001, Buczak et al. 2001). A Genetic Algorithm is an adaptive heuristic search algorithm that embodies the evolutionary concept of 'survival of the fittest'. The success of GAs is due to an intelligent exploitation of a problem's solution space.

Perhaps one of the most common solution techniques (possibly due to the accessibility of 'off-the-shelf' solvers) is *integer linear programming* (ILP). Integer linear programming is a variant of linear programming (LP) in which the decision-variables assume integer values only – in the case of the MCLP, the decision-variables are binary (0/1). In the formulation above, we typically have $m \gg n$ and Downs and Camm (1996) note that direct ILP approaches to solving equations 4.1 to 4.5 above frequently suffer from a high degree of both primal and dual degeneracy. Replacing y_i with $h_i = 1 - y_i$ in the MCLP formulation leads to an equivalent minimisation (of *uncovered* nodes) problem which, it is claimed, identifies an optimal solution more quickly and reduces primal degeneracy (Downs and Camm 1996).

In the next section we describe a more general version of the MCLP in which each sensor type is characterised by a spatially-explicit weighting function (representing say, a detection probability or 'depth of feel') rather than a deterministic cut-off range. The distinction is that the weighting-function can provide an anisotropic representation of monitoring 'effectiveness' (which itself may be sensor-specific) that assigns higher weight when the target is close and less weight when the target is far away. This 'envelope of effectiveness' is then combined with a 2D 'risk' surface to produce a final set of weightings (the a_{ij} of equation 4.2). We are unaware of any similar formulation of the MCLP and to that extent, we believe that this represents a new and original contribution to the literature on covering problems. We have termed this the generalised MCLP or *g*-MCLP.

4-4 The generalised MCLP (g-MCLP)

Our problem formulation is based on subdividing a region of interest into a regular grid of appropriate dimensions. What constitutes 'appropriate' will be problem specific but will be based on considerations of: (i) computational resources; (ii) mathematical tractability; and (iii) a context-specific decision unit. By (iii) we mean the dimensions of a grid *cell* that are commensurate with the scale of the range over which sensors are effective and the geographic extent of the surveillance space. Thus, for example, if a particular monitoring device or activity is effective up to 50-80 km and surveillance is over 250,000 km^2 an appropriate grid cell might, for example be 20 km x 20 km. The motivation for discretising the problem is two-fold: (i) a discrete representation of the surveillance space facilitates codifying and solving the problem as a MCLP; and (ii) it is consistent with the way in which management agencies characterise regional risk (Figure 37). The generic situation is depicted in Figure 39.

The overall network configuration is defined by the number, type, and spatial location of individual sensors. For each grid cell we define the following binary decision variable:

$$x_{i,j,k} = \begin{cases} 1 & \text{if cell } \{i,j\} \text{ has a sensor of type } k \\ 0 & \text{otherwise} \end{cases} \tag{4.6}$$

The concept of ‘envelope of effectiveness’ is illustrated in Figure 51 which shows ellipses centred at three locations (cells). The orientation and spatial extent of an ellipse is assumed to be a function of the monitoring instrument, device or procedure. Furthermore, it is assumed that the information for cells close to the sensor location will be more relevant or accurate than for cells further away. This implies a spatial gradient or weighting of ‘accuracy or ‘relevancy’. Note, this assumption may not apply in all instances in which case an equal weighting scheme can be used.

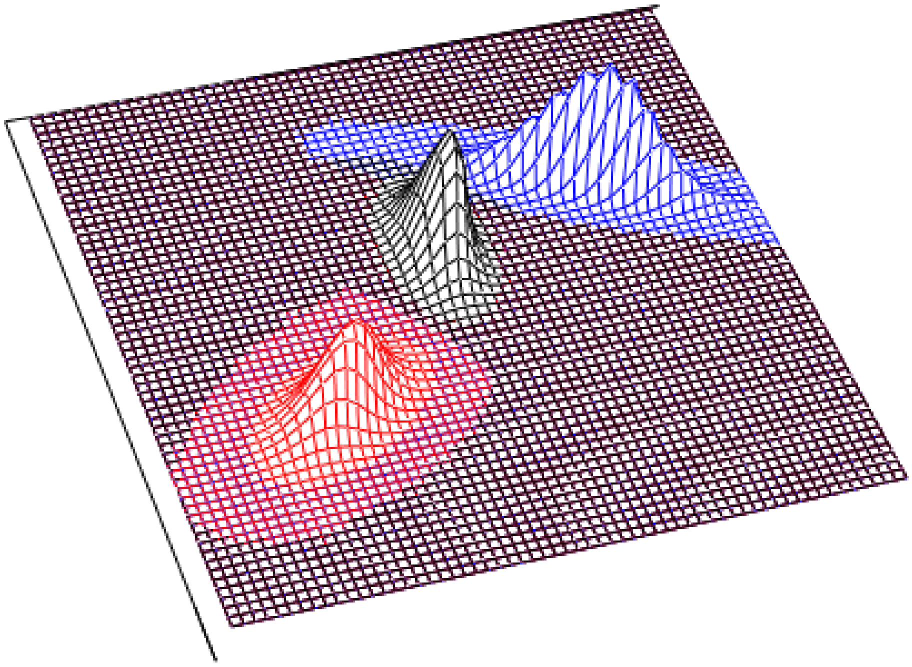

Figure 51. Illustration of 'envelope of effectiveness' for three sensor types. Note that range and orientation can be different for each sensor.

In order to formulate a mathematical objective function, we need first to define a measure of utility. Obvious candidates include risk (however defined), cost, and detection probability. Clearly, the first two choices result in minimisation problems while use of the third results in a maximisation problem. For illustrative purposes, we will use a combination of both risk (defined as the likelihood of spread of infection) and detection probability. The objective function will be the maximisation of the overall probability of detecting the spread of EI. Note, detection probability is a characteristic of the sensors while the risk of spread is a function of uncontrollable (external) factors. However, the overall probability of detecting the spread of EI is a product of both terms. Thus we have:

$$P\left[\text{detection at } \underline{x} \text{ with sensor located at } \underline{y} \middle| \text{disease present in vicinity of } \underline{x}\right] = w(h) \quad (4.7)$$

$$P\left[\text{present at location } \underline{x}\right] = E(\underline{x}) \quad (4.8)$$

$$P\left[\text{detection at } \underline{x} \text{ with sensor located at } \underline{y}\right] = E\left(\underline{x}\right) w\left(h\right) \quad (4.9)$$

where $h = \left\|\underline{x} - \underline{y}\right\|$ is the distance between the site of interest and the sensor location. Furthermore, we impose the requirements that $w\left(0\right) = 1$ and $w\left(\infty\right) = 0$.

4-4-1 A logistic model for disease probability

In the absence of hard data, we have used the following logistic model to compute the probabilities in equation 4.8:

$$E\left(\underline{x}\right) = \frac{\exp\left\{\beta_0 + \beta_1 EI\left(\underline{x}\right) + \beta_2 \text{log\_density}\left(\underline{x}\right) + \beta_3 EI\left(\underline{x}\right) \cdot \text{log\_density}\left(\underline{x}\right)\right\}}{1 + \exp\left\{\beta_0 + \beta_1 EI\left(\underline{x}\right) + \beta_2 \text{log\_density}\left(\underline{x}\right) + \beta_3 EI\left(\underline{x}\right) \cdot \text{log\_density}\left(\underline{x}\right)\right\}} \quad (4.10)$$

where the β terms are model parameters; *EI* is an indicator variable having value 1 if EI has been detected at $\underline{x}$; and log_density is the natural logarithm of population density at $\underline{x}$. For positive values of β_1, β_2 and β_3, risk is an increasing function of population density and will tend to increase rapidly if there has been at least one reported incidence of EI at $\underline{x}$.

The discrete version of the surveillance problem involving a single sensor is shown in Figure 52.

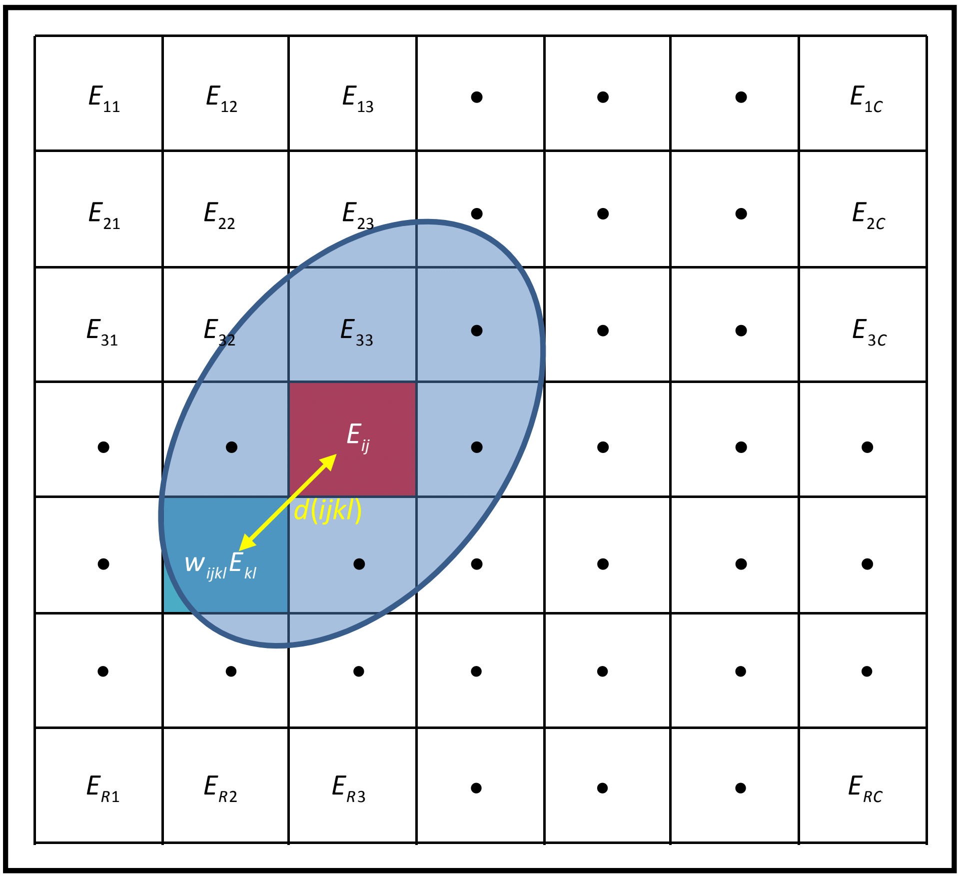

Figure 52. Representation of single sensor network. Sensor is located location in cell {*i,j*}. The probability of detection in cell {*k,l*} is a product of the risk/likelihood for cell {*k,l*} (denoted E_{kl}) multiplied by the sensor's detection probability w_{ijkl} which is a function of distance (d_{ijkl}) between cells {*i,j*} and {*k,l*}*).*

In reality the situation is more complex than indicated by Figure 52 since there will be multiple sensors and it is conceivable that some grid cells will fall within the envelope of effectiveness of more than one sensor. The more general situation is shown in Figure 53.

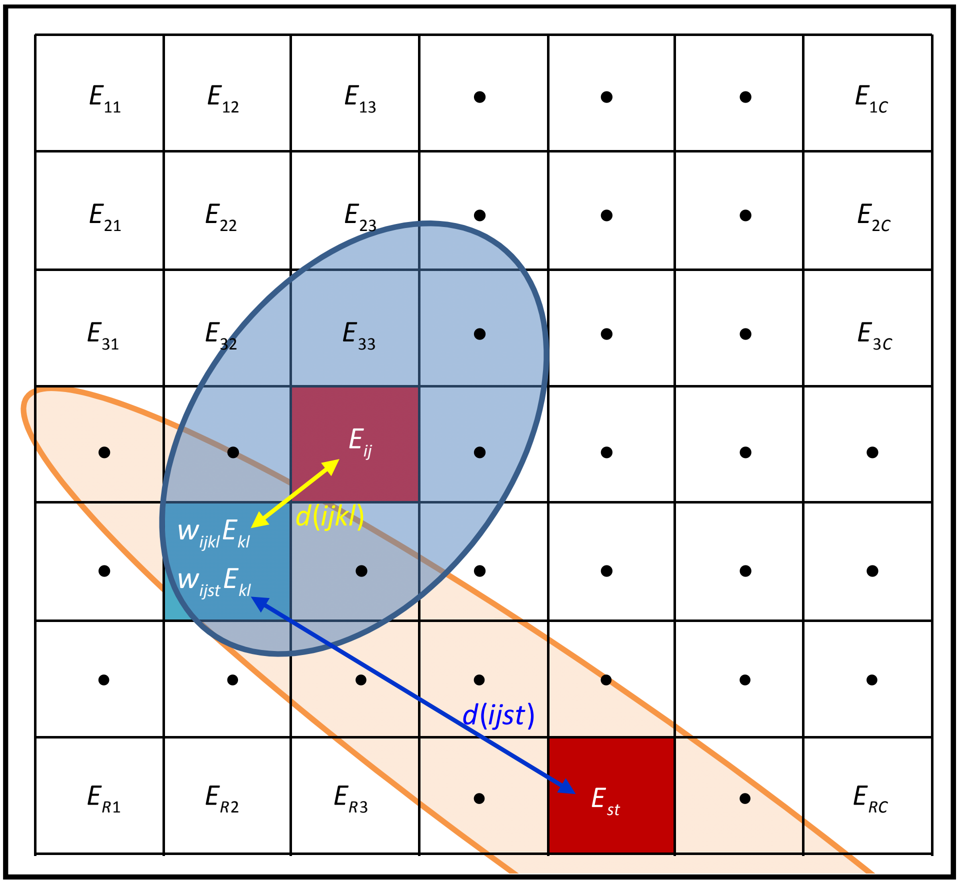

Figure 53. Generalisation of previous figure for multiple sensor network.

4-5 Problem formulation

We consider the problem of optimally locating m sensors on an $r \times c$ grid. It will be convenient to use matrix notation to define the objective function and in order to simplify this expression it is helpful to 'unpack' the row-column data associated with the sampling grid. It doesn't matter how this is done although we have chosen to do this row-by-row to form the following three matrices:

Decision matrix: X

This is a $(mrc\ x\ m)$ block-diagonal matrix with $X = diag\left(\underline{X}^{(1)},\ldots,\underline{X}^{(m)}\right)$ where $\underline{X}^{(p)}$ is a $(rc\ x\ 1)$ column vector having elements defined by equation 3.1 for each of the *rc* grid cells for the p^{th} sensor.

Weight matrix: W

This is a $(rc\ x\ mrc)$ matrix having structure: $W = \left[W^{(1)}\middle|W^{(2)}\middle|\cdots\middle|W^{(m)}\right]$ where $W^{(p)}$ is a $(rc\ x\ rc)$ matrix of detection probabilities associated with the p^{th} sensor. Let $\underline{w}_j^{(p)}$ be the j^{th} $(rc\ x\ 1)$ column of $W^{(p)}$. The elements of $\underline{w}_j^{(p)}$ are the detection probabilities for all *rc* grid cells obtained when sensor *p* is placed within the j^{th} grid cell.

Risk vector: E

E is a $(1\ x\ rc)$ row vector whose entries are the unpacked *E*'s of Figure 52.

Objective function

The objective function is:

$$\text{Maximize} \quad \mathbb{Z} = E.W.X.\underline{1}^T \tag{4.11}$$

where $\underline{1}^T$ is an $(m\ x\ 1)$ vector of ones.

Constraints

A complete set of constraints will require elicitation. However, by way of example we could have constraints on cost; number of sensor sites; and sub-region representation (to accommodate subjective requirements for monitoring in certain regions).

Number of sensors of type j:

$$\underline{1}^T\ \underline{X}^{(j)} \le n_j \tag{4.12}$$

where $\underline{1}$ is a $(rc\ x\ 1)$ vector of ones.

Total number of sensors:

$$\underline{J}^T\ \underline{X}\ \underline{1}^T \le N \tag{4.13}$$

where $\underline{J}^T$ is an $(1\ x\ mrc)$ vector of ones and $\underline{1}^T$ is a $(m\ x\ 1)$ vector of ones.

Total Cost: $$C^T \underline{X}\, \underline{1}^T \le \mathbb{C} \qquad (4.14)$$

where C^T is a $(1\, x\, mrc)$ vector of costs and $\underline{1}^T$ is a $(m\, x\, 1)$ vector of ones.

The structure of C^T is $\left[\underline{C}^{(1)} \middle| \underline{C}^{(2)} \middle| \cdots \middle| \underline{C}^{(m)}\right]$ where component $\underline{C}^{(j)}$ is a $(1\, x\, rc)$ vector of costs associated with locating a sensor of type *j* within each of the *rc* grid cells. Note, that this formulation is sufficiently general in that it acknowledges that establishment an operating costs can vary spatially even for the same type of equipment.

Representitveness $$\sum_{j=1}^{m} L_S^T \underline{X}^{(j)} \ge R_S \qquad (4.15)$$

where L_S^T is a $(1\, x\, rc)$ vector of ones and zeros such that a 1 indicates that the relevant grid cell is a member of sub-region *S* and R_S is a scalar lower bound on the number of sensor locations that *must* be placed within sub-region *S*.

Minimum spacing $$\underline{X}^{(i)} \left[\underline{X}^{(j)}\right]^T \ge S\, \underline{1} \quad \forall \ \{i, j\} \qquad (4.16)$$

where S (a scalar) is the minimum separation between any two sensors.

Note that equation 4.16 is non-linear in the decision-variables. Equation 4.16 can be linearised with the imposition of additional constraints. This is achieved as follows: suppose *u* and *v* are two binary variables. Then their product *uv* can be replaced by a new binary variable η with the additional constraints: (i) $u + v \ge 2\eta$; and (ii) $u + v - 1 \le \eta$.

4-6 Example

Important note: *This example is for illustrative purposes only and uses artificially constructed data to represent EI risk.*

We demonstrate the approach outlined in the previous section by examining a number of scenarios for optimal sensor location over the state of NSW under various constraints. A 14 x 14 grid has been used corresponding to a grid cell size of approximately 68 km (north-south) x 144 km (east-west). Demographic data for local government areas has been taken from publicly available information (see Appendix F). A surface/contour plot of the population density is shown in Figure 54 while Figure 55 shows population density contours with EI detection locations superimposed.

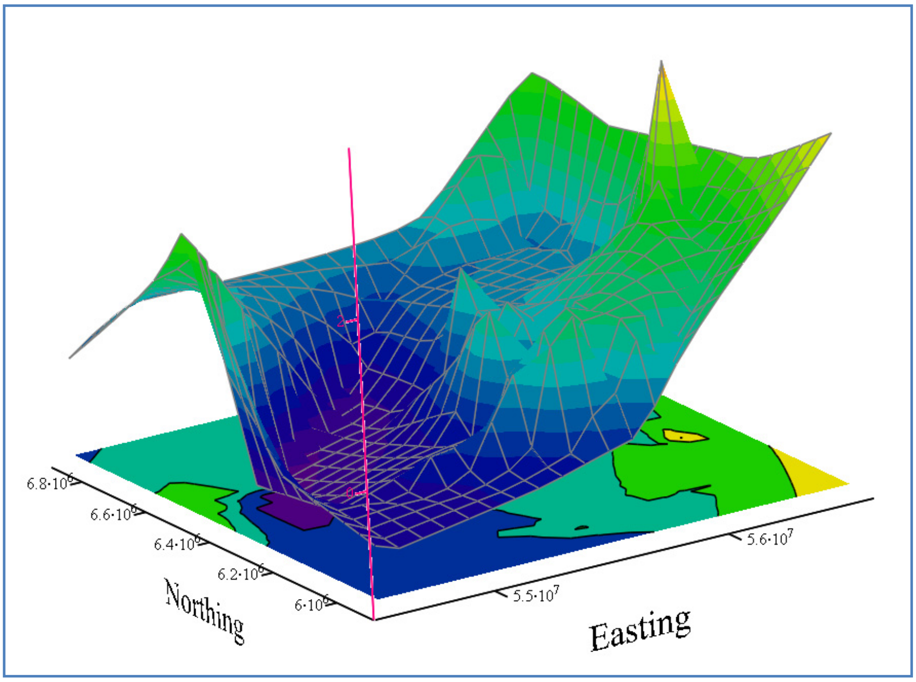

Figure 54. 3D Surface/contour plot of NSW population density. (Raw data given in Appendix F).

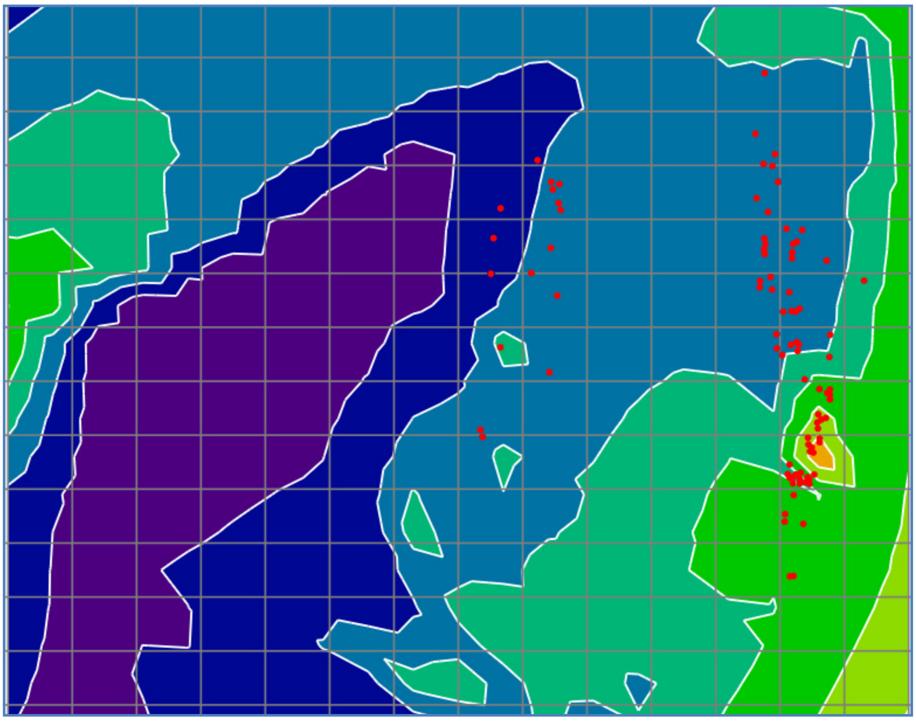

Figure 55. NSW population density contours. Solid red circles correspond to EI locations as shown in Figure 1.

The following parameter values have been used in equation 4.10: $\beta_0 = -1.0023$; $\beta_1 = 1.6554$; $\beta_2 = 0.23728$; and $\beta_3 = 1.16869$. A plot of the 3D risk surface is shown in Figure 56 and again as a contour plot with superimposed EI detection locations in Figure 57.

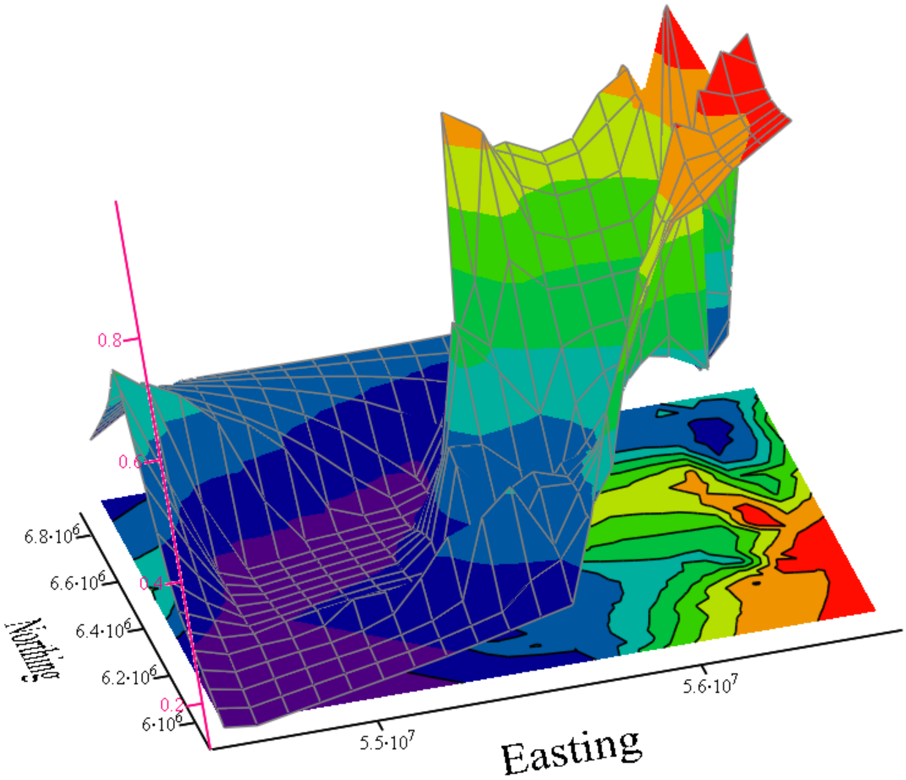

Figure 56. 3D surface/contour plot of EI risk over NSW.

Because of the discreteness of the problem formulation our decision variables are individual *cells* - not infinitesimally small points on the map. For this reason we need a measure of risk for a cell rather than a point. To this end, the continuum of risk has been averaged over each grid cell to obtain an average cell risk (Figure 58).

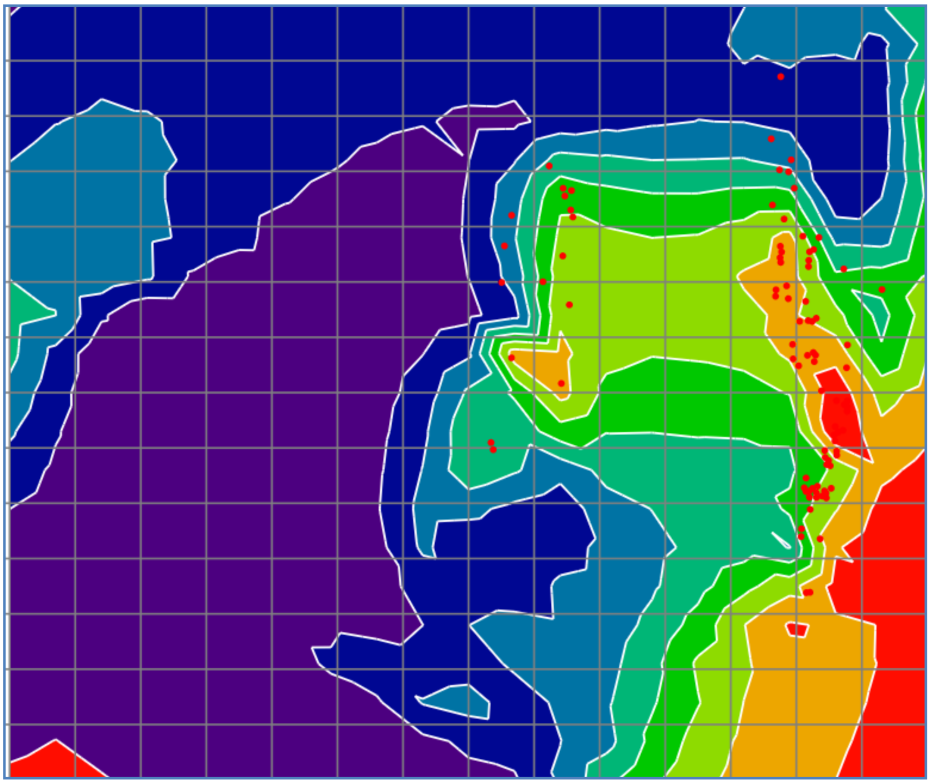

Figure 57. Contour plot of EI risk in NSW. Solid red circles are EI locations from Figure 1.

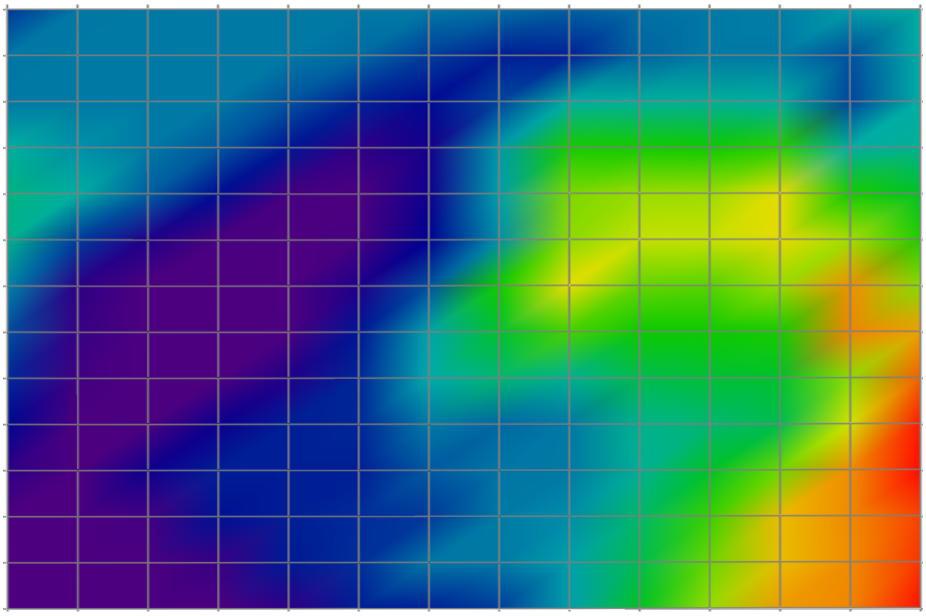

Figure 58. Block-averaged risk probabilities.

4-6-1 Implementation

Pre-processing of input data (creation of *W* matrix and risk vector *E*) and various other manipulations were performed in Mathcad 14 (Parametric Technology Corporation). The objective function is given by equation 4.11. The ILP solution was implemented using extended *LINGO 11* (Lindo Systems 2007). The LINGO code appears in Appendix G. Elements of the vector *E* and cost coefficient data appear between the "data" and "enddata" lines of the LINGO program in Appendix G.

Scenario evaluation

A number of scenarios were devised to illustrate the impact on the optimal sensor network as a result of altering one or more constraints. The system is sufficiently flexible to allow the incorporation of as many constraints as desired as well as the ability to relax or tighten constraints individually or simultaneously. Combinations of different surveillance modes can be readily accommodated by specifying the number and type of 'sensors' available. Individual sensor effectiveness is reflected in the matrix of weights, *W* defined earlier. For the purpose of illustration, we assume we have available three different surveillance methods or sensors. The number and characteristics of each of these is given in Table 4-1.

Table 4-1 Sensor characteristics.

Sensor Type	No. Sensors available	Major range	Major range	angle
1	1	50	25	340^0
2	2	30	25	60^0
3	2	35	20	0^0

A graphical depiction of the sensor characteristics corresponding to the parameters in Table 4-1 is shown in Figure 59.

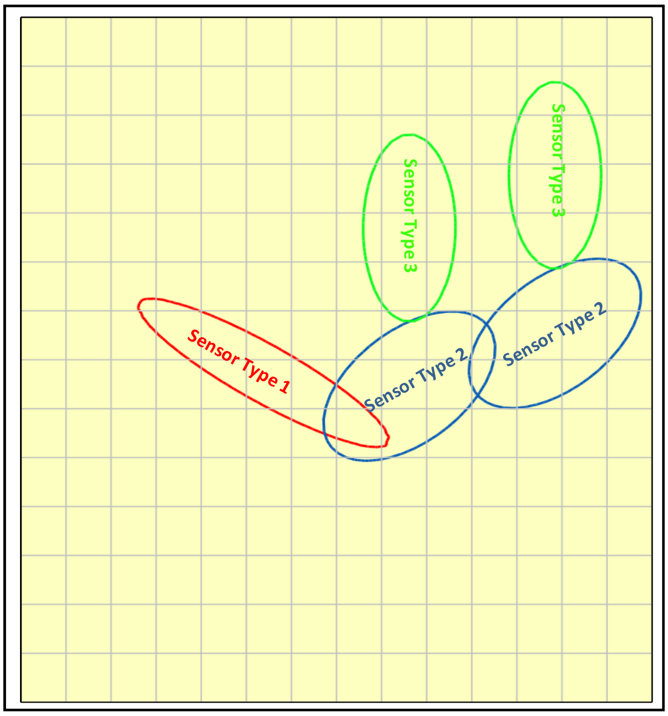

Figure 59. generic representation of three sensors having characteristics given in Table 4-1.

Six scenarios showing the impact on the optimal solution as a result of changing the minimum separation and cost constraints are listed in Table 4-2. An arbitrary function was used to generate a cost surface over the region of interest (Figure 60). By the same reasoning that was applied to the risk metric, costs at individual points need to be averaged over each cell (Figure 61).

Table 4-2. Constraint data.

Scenario	Max. number of sensors			Minimum sensor spacing ($\leq$ km)	Cost constraint ($\leq$ \$)
	Type 1	Type 2	Type 3		
1	1	2	2	30	∞
2	1	2	2	35	∞
3	1	2	2	45	∞
4	1	2	2	60	∞
5	0	2	2	0	20000
6	0	2	2	35	20000

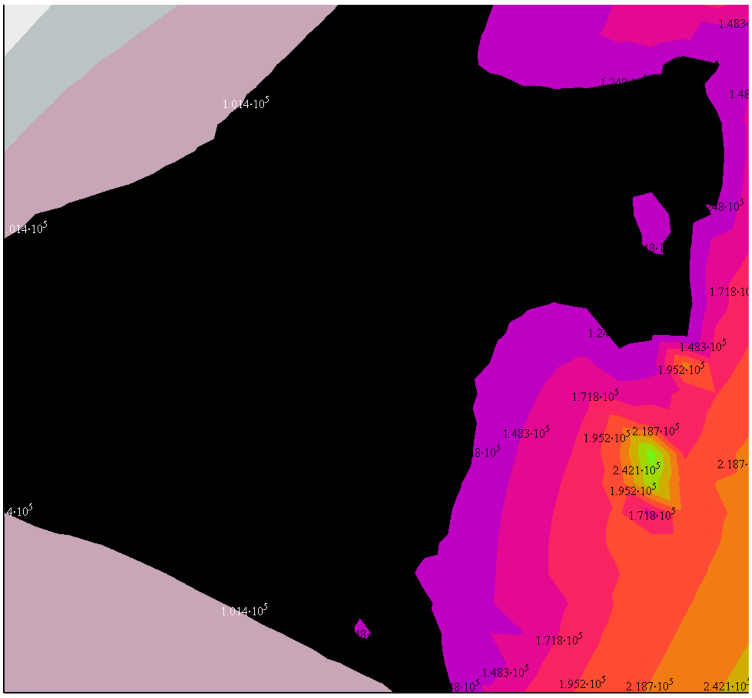

Figure 60. Cost-contours for surveillance monitoring.

The optimal sensor configuration for each scenario has been identified (Figures 62 to 67[9]). Figure 68 shows a solution configuration overlaid on satellite imagery. While the results in Figures 62 to 67 are intuitively sensible (solutions targeting areas of high risk while honouring the minimum separation and cost constraints) their identification requires a considerable amount of computation. Nevertheless, the use of highly optimised numerical algorithms such as those found in LINGO® meant that a solution was generally found within a few minutes when run on a standard desktop computing platform.

[9] The contour lines in Figures 62 to 67 have no meaning and are an anomaly of the software used to produce these figures.

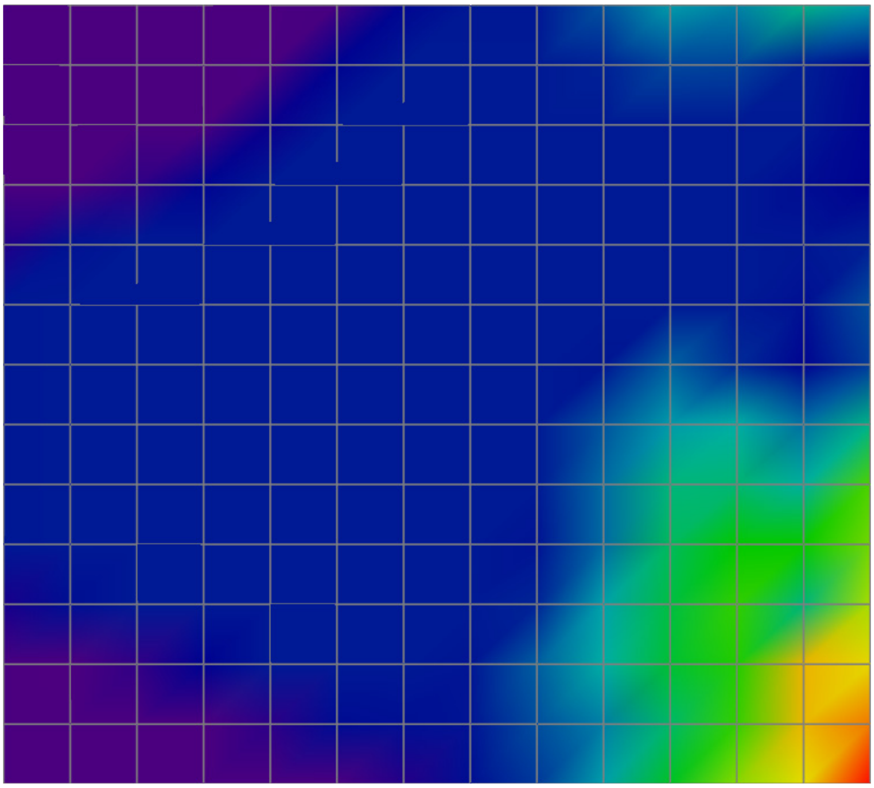

Figure 61. Block-averaged surveillance monitoring costs.

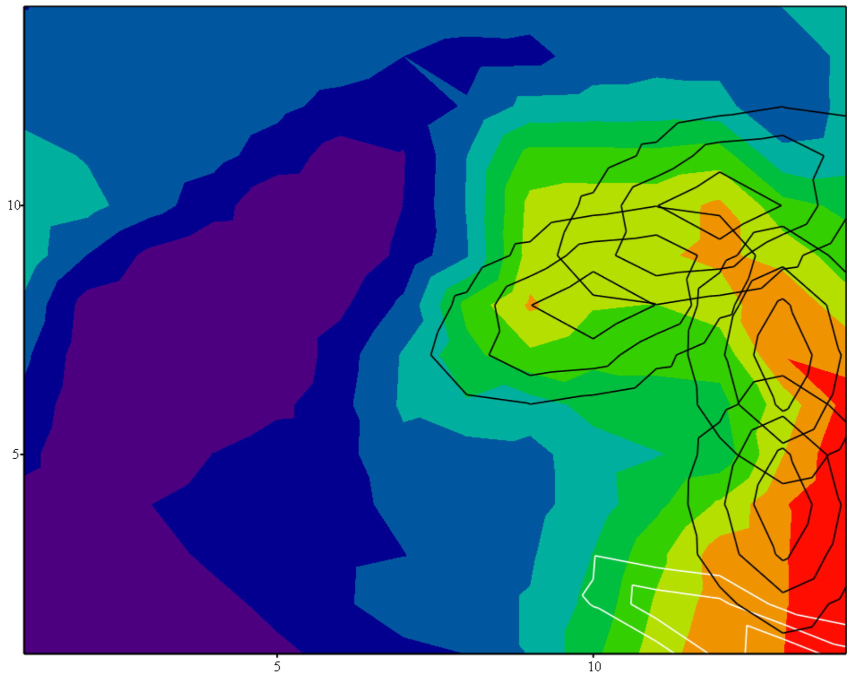

Figure 62. Optimal solution for scenario #1: 5 sensors; min spacing 30 km; no cost constraint.

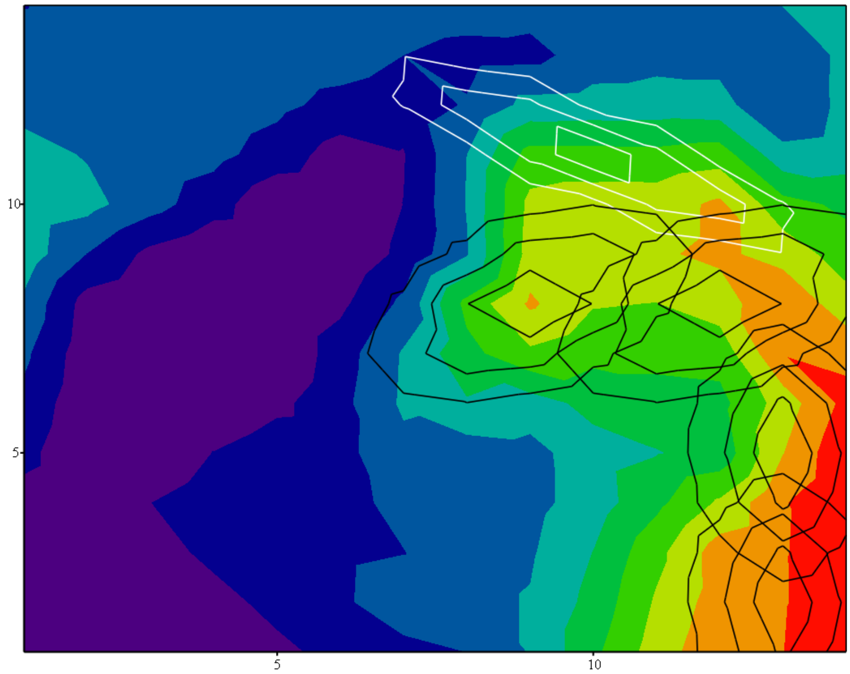

Figure 63. Optimal solution for scenario #2: 5 sensors; min spacing=35 km; no cost constraint.

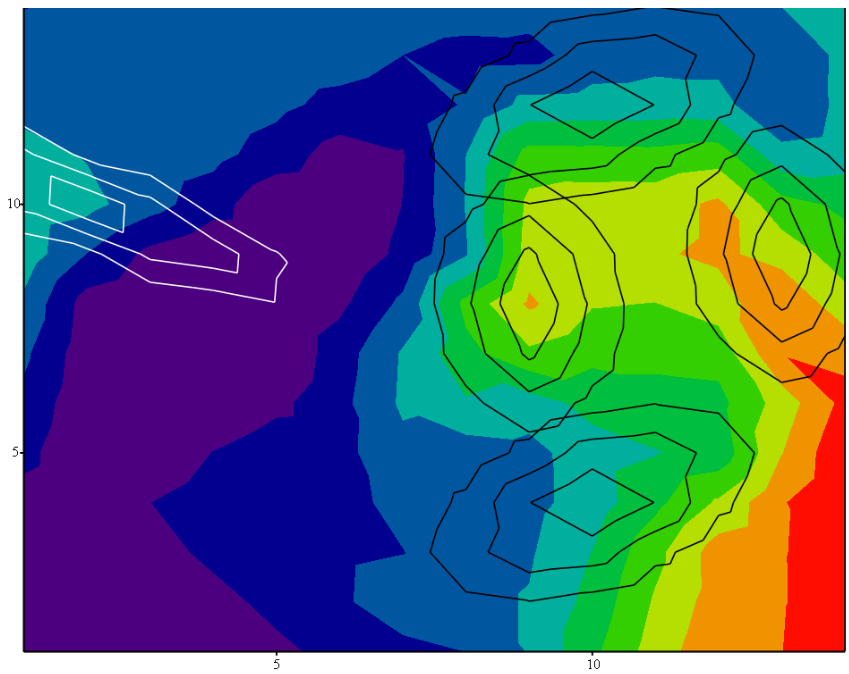

Figure 64. Optimal solution for scenario #3: 5 sensors; min spacing=45 km; no cost constraint.

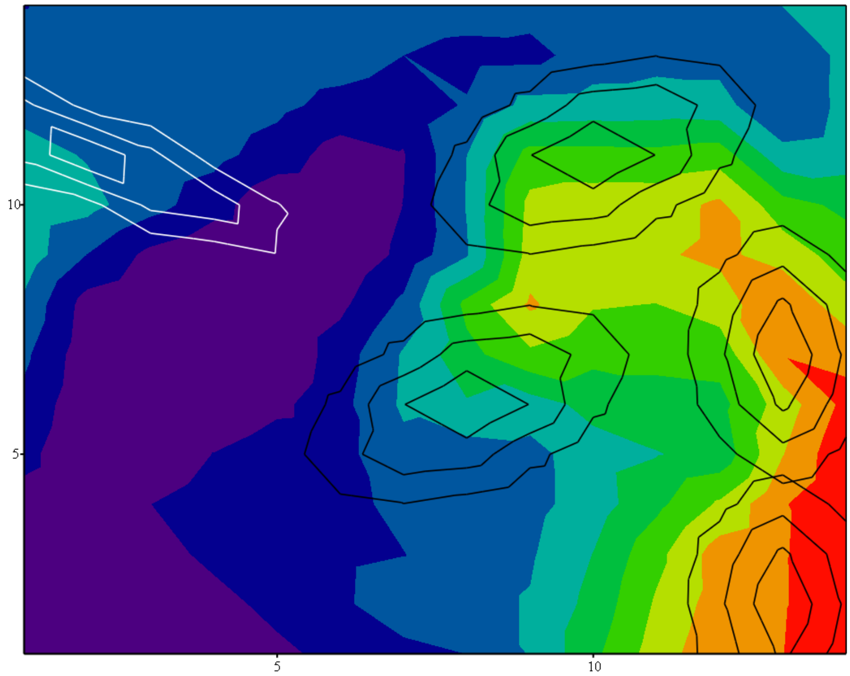

Figure 65. Optimal solution for scenario #4: 5 sensors; min spacing=60 km; no cost constraint.

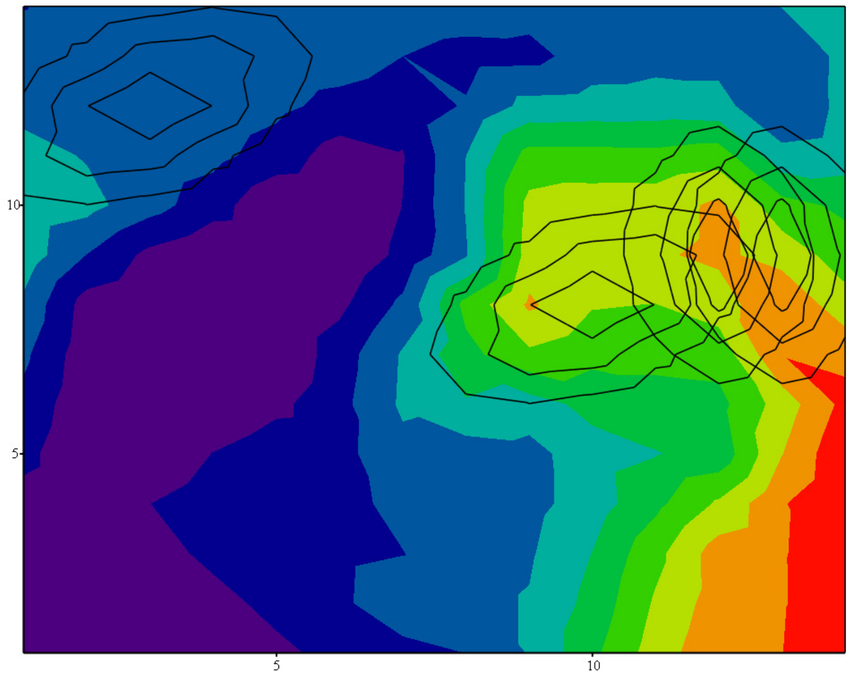

Figure 66. Optimal solution for scenario #5: 4 sensors; no min spacing; cost <=$20,000.

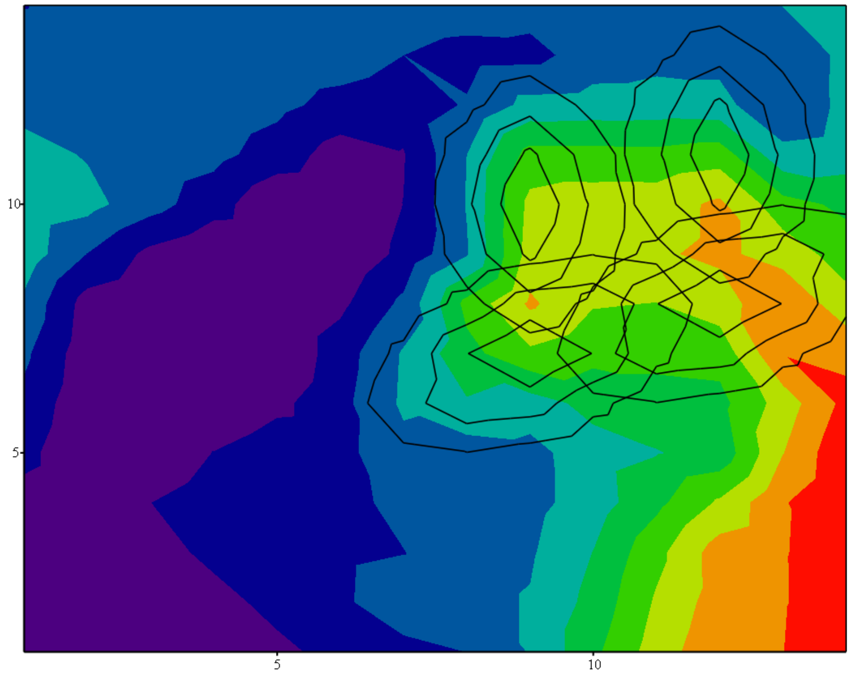

Figure 67. Optimal solution for scenario #6: 4 sensors; min spacing 35 km; cost <=$20,000

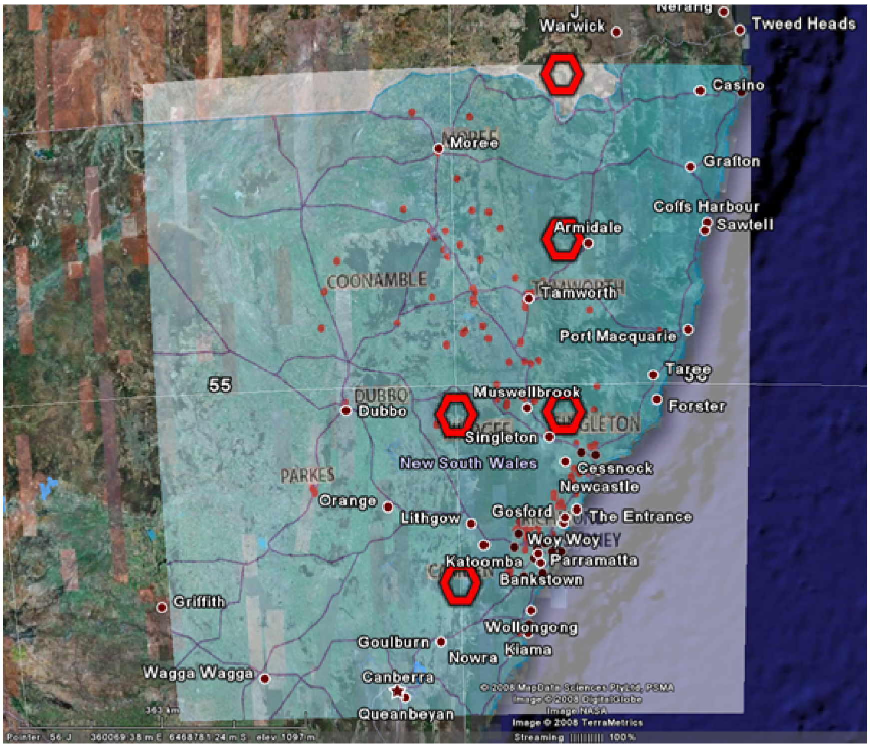

Figure 68. Optimal monitoring locations corresponding to Figure 62 overlaid on geographic map. (*Note: Apparent positional discrepancies due to different mapping projections).*

4-7 Discussion

The use of mathematical programming techniques to optimise network design problems is indeed not new. Applications of mathematical programming techniques to the optimisation of sparse sensor networks have been associated with air quality monitoring (Fox et al. 2009, McElroy et al., 1986), water supply security (Chakrabarty, & Iyengar, 2002) and computer network integrity (Noel and Jajodia, 2007). We have uncovered only limited evidence of mathematical optimisation methods applied to the identification of optimal network designs for biosecurity surveillance.

Alternatives to the conventional mathematical programming methods listed above include genetic algorithms and heuristic algorithms (Kanaroglou et al., 20005; Liu, et al., 1986). These are briefly discussed below.

Genetic algorithms (GAs)

These commence with a group of possible solutions; coded as binary strings and then the solutions with greater utility are selected, perturbed and re-combined to produce even better solutions. Genetic algorithms are an example of an evolutionally algorithm which is generally completed either for a set number of generations or until it does not seem that any further improvement is possible. GAs provide efficient location of near optimal solutions without any necessary previous understanding of the search space.

Greedy searches

These are a type of heuristic algorithm whereby actions are chose on the basis of their improvement in the objective function *for that step*. They are computationally efficient and can allow for a reasonable solution to very complex problems. Noel and Jajodia (2007) use a modified greedy search. They are also scalable to constraints to a greater degree than other approximate algorithms. Krause et al. (2006) discusses the problem with the performance of greedy searches in a situation where communication costs are considered critical. In terms of computational effort greedy search algorithms tend to perform significantly worse than other stochastic algorithms.

Gradient ascent

Gradient ascent is a non-linear optimisation algorithm which attempts to find extrema by moving in the direction of steepest gradient of the objective function. It was used for example by Vickers et al. (2006) to guide the optimal placement of sensors on the base of information on the state of the environment (specifically water flow) for a defensive military application. Gradient ascent methods will identify *local* extrema but can get 'trapped' and fail to reach a global optimum. To circumvent this problem, gradient ascent programs are often run as ensembles generated with randomised initial conditions. The best performing member of the ensemble is chosen as an optimum and generally provides a more robust estimation of the global maximum.

The issue of surveillance network design is as important as the surveillance activities and data analysis methods themselves. A sub-optimal monitoring network design is not only wasteful of precious monitoring resources but compromises statistical power – that is the ability to identify disease outbreaks, quarantine threats, or (bio) security violations when they have in fact occurred. As noted by Patil et al. (2006) "surveillance geoinformatics of spatial and spatiotemporal hotspot detection and prioritization is a critical need for the 21st century".

Although remote sensing will continue to provide an important capability in plant protection and monitoring, the need for ground-based surveillance systems will remain. To be effective, remote sensing needs to be able to resolve zones of infection that are as small as 5m in diameter – that is, of the order of a single pixel of information generated by present day satellites (Patil et al. 2006).

While the siting of biosurveillance 'sensors' has been recognised as an important consideration in monitoring program design, most efforts in this regard have been driven largely by logistical considerations using heuristic algorithms. For example, in response to the September 11 2001 terrorist attacks the United States government, through its Department of Homeland security, deployed the BioWatch Program to provide early warning of a mass pathogen release. Although exact details of the location of BioWatch monitoring sites is unknown, it is thought that these may have been co-located with EPA air quality monitoring sites "on the basis of cost and ease of access" (Shea and Lister, 2003).

In this chapter we have presented a design tool/methodology which compliments the temporal monitoring of Chapter 2 and the spatial-temporal modelling of Chapter 3. Together, these three chapters provide new and novel methods for the design, analysis, and prediction of disease/pest movement in space and time.

We regard these methods as a starting point for further exploration and development. In particular, they need to be thoroughly 'road-tested' on a suite of comprehensive scenarios based on actual biosecurity monitoring and surveillance case studies.

Chapter 5

5-1 INTRODUCTION

Since the events of September 11 2001, there has been increased emphasis on monitoring and surveillance to detect and prevent further terrorist attacks. While significant resources have been devoted to the mechanics of screening, far less attention has been paid to quantifying the efficacy of these surveillance programs.

Given the high volumes of passengers, containers, mail items, and various other inter-continental 'movements', it is pertinent to examine the effectiveness of the inspection regimes. From a statistical perspective, the answer is partly provided by the theory and methods of statistical process control (SPC) and elementary probability theory (*see chapter 1*). However, bio-surveillance, syndromic surveillance, and counter-terrorism surveillance are fundamentally different from monitoring activities for quality assurance in an industrial manufacturing process. Unlike the industrial setting where there is generally good information on the performance of the manufacturing process (eg. percent defective; proportion of non-conforming or 'out-of-spec' items), monitoring in the context of bio / homeland security is characterised by extreme uncertainty. For example, because of the nefarious activities of terrorist organisations, security and intelligence organisations do not always know what it is they're looking for. As terrorists become increasingly sophisticated in their modes of attack, our uncertainty in both the likelihood of an attack and which sentinels to monitor increases. By way of example, prior to 9/11 there was little or no surveillance at general aviation (GA) airports or monitoring of student pilot training. This lack of information regarding what to monitor is a case of Shackle-Popper Indeterminism (SPI) (Ben-Haim, 2006). The distinguishing features of SPI are:

Intelligence:	What people know influences how they behave.
Discovery:	What will be discovered tomorrow cannot be known today.
Indeterminism:	Tomorrow's behaviour cannot be modelled completely today.

In contrast to traditional methods of (constrained) optimisation, Ben-Haim (2006) developed Info-Gap theory to identify robust solutions to decision-making problems under extreme uncertainty. Info-gap theory has recently been applied to assessing the performance of counter-terrorism surveillance programs (Moffitt et al. 2005) and the identification of robust strategies to deal with bioterrorism attacks (Yoffe and Ben-Haim, 2006). Thompson (unpublished) examined the general sampling problem associated with inspecting a random sample of n items (containers, flights, people, etc.) from a finite population of N such items in a biosecurity context using an info-gap approach. The basic situation considered by Thompson is that there is a probability $p(n)$ of a catastrophic outcome (eg. terrorist attack) given that n events / items out of N have been inspected. The info-gap formulation of the problem permitted the identification of a sample size n such that $p(n)$ did not exceed a nominal threshold, π_c when severe uncertainty about $p(n)$ existed. Implicit in this formulation was the assumption that the detection probability (ie. the probability of detecting a weapon, adverse event, anomalous behaviour etc.) once having observed or inspected the relevant item / event / behaviour was unity. In the context of counter-terrorism, our uncertainties (or info-gaps) will most certainly extend to a lack of certitude in detection. This is a consequence of the 'discovery' aspect of the SPI phenomenon identified above. Consider the following examples:

1. Quarantine authorities invoke various levels of inspection ranging from a cursory examination (eg. opening the door of a container and visually the contents that are accessible) to full inspection (eg. removal and inspection of the entire contents). Clearly, the probability of detecting a prohibited import will be much lower in the first case than the second even though the container is regarded as having been inspected in both cases;

2. Indeterminism means that security agencies do not always know what they're looking for. For example, liquids are currently banned from hand-luggage on certain international flights. This is because explosive devices can be assembled from innocuous constituents carried separately by more than one passenger.

Prior to this 'discovery', the detection probability for an acetone-peroxide based explosive would have been small.

In the following sections we describe the general surveillance problem for which the probability of detection is less than unity. We then provide an info-gap formulation to help identify sampling strategies that are robust to multiple sources of uncertainty – including the detection probability.

5-2 Surveillance with imperfect detection

Following Thompson (unpublished), we assume that there is a finite population of N objects, events, people, or behaviours that are potentially subject to inspection. From this population of N 'objects' a random sample of size n is to be inspected. We define the following events:

I – the event that an object is inspected;

W – the event that an object is a security threat (eg. the object is a weapon, the person is a terrorist, the behaviour is indicative of malicious intent);

D – the event that the security breach is identified / detected.

Furthermore, we assume that only inspected objects are classified as either belonging to D or $\overline{D}$. We thus have $\{I\} = \{D\} \cup \{\overline{D}\}$ and hence

$$P[I] = P[D] + P[\overline{D}] \tag{5.1}$$

Furthermore, $\{I\} = \{D \cap W\} \cup \{\overline{D} \cap W\} \cup \{\overline{D} \cap \overline{W}\} \cup \{D \cap \overline{W}\}$ and thus

$$P[I] = P[D \cap W] + P[\overline{D} \cap W] + P[\overline{D} \cap \overline{W}] + P[D \cap \overline{W}] \tag{5.2}$$

In a biosecurity / counter-terrorism context, arguably, the most important probability is *not* $P[W]$ (the probability of a security threat) but rather it is the *conditional probability* $P[W|\overline{D}]$ ie. the probability of a security threat *given* that no breach of security was detected.

The lack of detection of a security breach is due to: (i) the absence of a security threat; and/or (ii) imperfections of the detection equipment / method/ process. Our inability to distinguish between (i) and (ii) is an info-gap.

5-2-1 Problem formulation

From elementary probability theory:

$$P\left[W\left|\overline{D}\right.\right]=\frac{P\left[W\cap\overline{D}\right]}{P\left[\overline{D}\right]} \tag{5.3}$$

From equation 5.2 we have:

$$P\left[W\cap\overline{D}\right]=P[I]-P[D\cap W]-P\left[\overline{D}\cap\overline{W}\right]-P\left[D\cap\overline{W}\right] \tag{5.4}$$

Each of the joint probabilities in Equation (4) can be expressed in terms of relevant conditional probabilities *viz:*

$$P\left[W\cap\overline{D}\right]=P[I]-P\left[D\left|W\right.\right]P[W]-P\left[\overline{D}\left|\overline{W}\right.\right]P\left[\overline{W}\right]-P\left[D\left|\overline{W}\right.\right]P\left[\overline{W}\right] \tag{5.5}$$

Note that $P\left[\overline{D}\left|\overline{W}\right.\right]=1$ and $P\left[D\left|\overline{W}\right.\right]=0$.

We next define the *detection efficiency*, θ as $P\left[D\left|W\right.\right]$ i.e. the probability that a security breach will be detected given a threat actually exists. Furthermore, we let $\phi=P\left[W\right]$ be the *unconditional* probability that an object is a security threat and $\lambda=P\left[I\right]={}^{n}\!/\!_{N}$ the inspection fraction or probability. Hence, equation 5.3 can be written as

$$P\left[W\left|\overline{D}\right.\right]=\frac{\phi(1-\theta\lambda)}{P\left[\overline{D}\right]}$$

$$\begin{aligned}&=\frac{\phi(1-\theta\lambda)}{(1-\theta)\phi\lambda+\phi(1-\lambda)+(1-\theta)\lambda+(1-\phi)(1-\lambda)}\\&=\frac{\phi(1-\theta\lambda)}{1-\theta\lambda\phi}\end{aligned} \tag{5.6}$$

Next, observe that

$$\begin{aligned} P[W] &= P\left[W \cap \overline{I}\right] + P[W \cap I] \\ &= P\left[W \cap \overline{I}\right] + P[D \cap W] + P\left[\overline{D} \cap W\right] \end{aligned}$$

Therefore

$$\begin{aligned} P\left[\overline{D} \cap W\right] &= P[W] - P\left[W \cap \overline{I}\right] + P[D \cap W] \\ &= P[W] - P\left[W \middle| \overline{I}\right] P\left[\overline{I}\right] - P[D|W] P[W] \\ &= \phi - (1-\lambda) P\left[W \middle| \overline{I}\right] - \theta\phi\lambda \\ &= \phi(1-\theta) - (1-\lambda) P\left[W \middle| \overline{I}\right] \qquad (5.7) \end{aligned}$$

But, $\{W\}$ and $\{I\}$ are *independent* events and therefore $\{W\}$ and $\{\overline{I}\}$ are also independent which means $P\left[W \middle| \overline{I}\right] = P[W]$ and thus equation 5.7 becomes

$$P\left[\overline{D} \cap W\right] = \phi(1-\theta\lambda) \qquad (5.8)$$

Notice that for the probability in equation 5.8 to be non-negative $\lambda \geq \theta$ i.e the sampling fraction must be at least as large as the detection efficiency. Substituting equation 5.8 into equation 5.6 gives

$$P\left[W \middle| \overline{D}\right] = \frac{\lambda + (1-\theta)\phi - 1}{\phi(\lambda-\theta) + P\left[\overline{D} \cap \overline{W}\right]}$$

But $P\left[\overline{D} \cap \overline{W}\right] = P\left[\overline{D} \middle| \overline{W}\right] P\left[\overline{W}\right]$ and since $P\left[\overline{D} \middle| \overline{W}\right] = 1,$ this becomes $P\left[\overline{D} \cap \overline{W}\right] = P\left[\overline{W}\right] = (1-\phi)$. Thus,

$$P\left[W \middle| \overline{D}\right] = \frac{\phi(1-\lambda\theta)}{1-\phi\theta\lambda} = p(\lambda,\theta,\phi) \qquad (5.9)$$

Note, that when 100% inspections are performed, the conditional probability in equation 5.9 becomes

$$P\left[W\middle|\overline{D}\right]=\frac{\phi(1-\theta)}{1-\theta\phi}=p(1,\theta,\phi) \qquad (5.10)$$

and under these conditions, this probability is only zero when the detection efficiency is 100%. For 0% detection efficiency $p(1,0,\phi)$ is ϕ - the unconditional probability that the object is a security threat. Furthermore, whenever the inspection rate is $\leq 100\%$, $p(\lambda,\theta,\phi)$ *exceeds* $p(1,\theta,\phi)$. This increase in 'risk' may be regarded as the 'cost' associated with less than complete inspection. We thus define our performance criterion Ψ to be the ratio $\frac{p(\lambda,\theta,\phi)}{p(1,\theta,\phi)}$, thus

$$\Psi(\lambda,\theta,\phi)=\frac{\phi(1-\lambda\theta)}{1-\phi\theta\lambda}\cdot\frac{1-\theta\phi}{\phi(1-\theta)} \qquad (5.11)$$

We next consider an info-gap formulation to assess the effects of uncertainty in key parameters (namely θ and ϕ) on the performance criterion given by equation 5.11.

5-3 An Info-Gap model for surveillance performance

Information-gap (hereafter referred to as info-gap) theory is a recent development designed to assist decision makers faced with severe uncertainty (Ben-Haim 2006, Regan et al. 2005, Carmel and Ben-Haim 2005). Info-gap theory aims to address the "robustness" of decision making under uncertainty. It asks the question: how wrong can a model and its parameters be without jeopardising the quality of decisions made on the basis of this model?

Info-gap theory derives its robustness functions from three elements: a performance measure, a process model and a non-probabilistic model of uncertainty. The performance measure is a statistical, economic or bio-physical metric of value to the decision maker. The decision maker may wish to increase the performance measure (e.g. dollar value of a share portfolio) or reduce it (e.g. probability of not detecting a terrorist attack). In each case there is often a critical performance value which defines a change in decision. In our case, the performance measure is Ψ - effectively the reduction in surveillance efficacy when less than 100% inspection is employed.

The process model is a mathematical summary of the system in question. It describes the relationship between the performance measure and the important characteristics of the system in question. In this example the performance threshold is the maximum tolerable reduction in surveillance efficacy and the process model is given by equation 5.11.

The info-gap model of uncertainty for the uncertain quantities Θ in the process model is the unbounded family of nested sets $U\left(\alpha,\widetilde{\Theta}\right)$ of possible realisations Θ, where α represents the unknown "horizon of uncertainty" and $\widetilde{\Theta}$ our best or initial estimate of Θ. This model satisfies two axioms:

contraction: $$U\left(0,\widetilde{\Theta}\right)=\left\{\widetilde{\Theta}\right\} \quad (5.12)$$

nesting: $$\alpha<\alpha' \Rightarrow U\left(\alpha,\widetilde{\Theta}\right)\subset U\left(\alpha',\widetilde{\Theta}\right) \quad (5.13)$$

The contraction axiom states that in the absence of uncertainty $(\alpha=0)$, our best estimate $\widetilde{\Theta}$ is correct, while the nesting axiom states that the range of uncertain variation increases as the horizon of uncertainty increases. In all cases α is unknown and unbounded with $\alpha \geq 0$. In this example the uncertain quantities are the detection efficiency θ and ϕ, the probability that an object is a security threat. Thus, $\Theta=(\theta,\phi)$ and our initial or best estimate of these parameters is denoted $\widetilde{\Theta}=\left\{\tilde{\theta},\tilde{\phi}\right\}$.

In this section we consider uncertain parameter values – the detection efficiency θ and the probability that an object is a security threat, ϕ. The fractional errors $\left|\left(\theta-\tilde{\theta}\right)/\tilde{\theta}\right|$ and $\left|\left(\phi-\tilde{\phi}\right)/\tilde{\phi}\right|$ are unknown. With this prior information we formulate the following fractional-error info-gap model:

$$U\left(\alpha,\tilde{\theta},\tilde{\phi}\right)=\left\{\begin{matrix}(\theta,\phi): & \max\left[0,(1-\alpha)\tilde{\theta}\right]\le\theta\le\min\left[1,(1+\alpha)\tilde{\theta}\right] \\ & \max\left[0,(1-\alpha)\tilde{\phi}\right]\le\phi\le\min\left[1,(1+\alpha)\tilde{\phi}\right]\end{matrix}\right\}, \quad \alpha\ge 0 \quad (5.14)$$

This is a bounded family of nested sets of $\left\{\tilde{\theta},\tilde{\phi}\right\}$ values with the sets becoming more inclusive as the horizon of uncertainty, α increases.

The definition of the performance measure, process model and uncertainty model(s) completes the specification of the formulation of the info-gap analysis. We now turn to the derivation of the robustness function. In info-gap parlance "robustness" is defined as the greatest horizon of uncertainty, across all uncertain model components, such that the performance measure still meets the pre-defined requirement. In our application the robustness of a surveillance regime in which λ x100% of the target population is inspected, is the greatest horizon of uncertainty $\widehat{\alpha}$ for which all combinations of the uncertain parameters $\widetilde{\Theta}=\left\{\tilde{\theta},\tilde{\phi}\right\}$ the minimum required inspection performance is achieved, that is

$$\widehat{\alpha}\left(\lambda,\gamma_d\right)=\max\left\{\alpha:\left(\min_{(\theta,\phi)\in U\left(\alpha,\tilde{\theta},\tilde{\phi}\right)}\Psi\left(\lambda,\theta,\phi\right)\ge\gamma_d\right)\right\} \quad (5.15)$$

where γ_d is the required value of Ψ. Equation 5.15 is the robustness function for this application of the info-gap model. The strategy of robust-satisficing (Ben-Haim 2006) is to attempt to guarantee an adequate level of surveillance performance, by choosing a value of λ which is highly robust to uncertainty. Thus, for any inspection fraction λ, the robustness function indicates the confidence in attaining the minimum performance requirement with that λ.

Examination of the process model (equation 5.11) reveals that it is a monotonic decreasing function with respect to θ and a monotonic increasing function with respect to ϕ. Combining this observation with the uncertainty model (equation 5.14) allows us to write the inner minimum of the robustness function (equation 5.15) as

$$h\left(\alpha,\lambda,\theta,\phi\right)\ge\gamma_d \quad (5.16)$$

where

$$h(\alpha,\lambda,\theta,\phi)=\frac{(1-\alpha)\phi\left[1-(1+\alpha)\lambda\theta\right]}{1-\left(1-\alpha^2\right)\phi\theta\lambda}\cdot\frac{1-\left(1-\alpha^2\right)\theta\phi}{(1-\alpha)\phi\left[1-(1+\alpha)\theta\right]} \quad (5.17)$$

5-4 Illustrative Example

Suppose new intelligence suggested that a clandestine operation had been planned to smuggle native fauna out of the country and although the exact mode of export is unknown, it is thought to rely on secret cavities sown into a passenger's clothing. Airport security and quarantine staff thus have no clear idea what they are looking for except that they have been instructed to closely monitor the appearance, texture, and integrity of passengers' clothes. Our best guess of the parameters $\tilde{\Theta}=\left\{\tilde{\theta},\tilde{\phi}\right\}$ is $\tilde{\phi}=0.7$ and $\tilde{\theta}=0.05$ although considerable uncertainty exists around these figures. Figure 1 plots the performance function $\Psi(\lambda,\theta,\phi)$ as a function of robustness for a range of λ values.

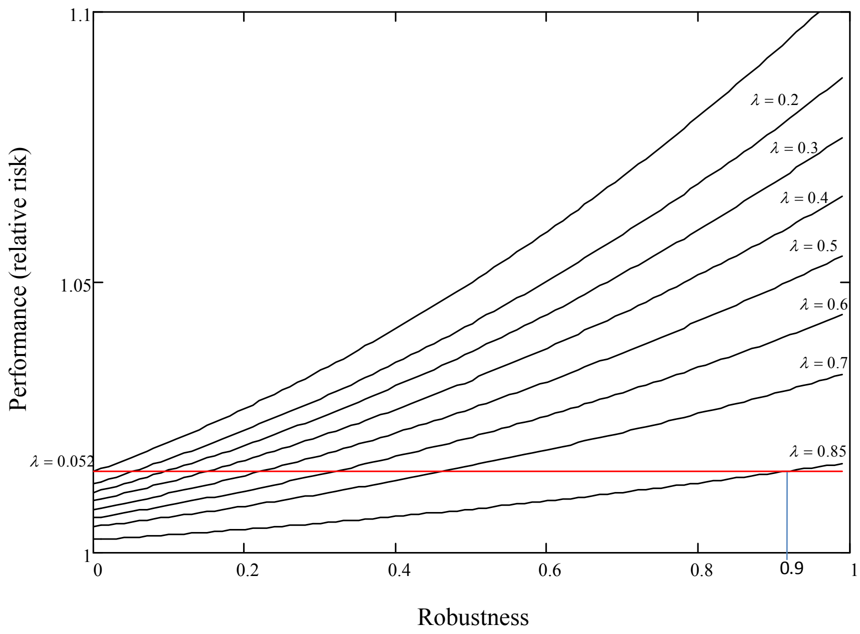

Figure 69. Robustness of surveillance performance for various sampling fractions (lambda).

In recognition that 100% detailed inspection of all passengers is not feasible, a reduced level of surveillance will be tolerated provided the increased risk (of an undetected threat)is no more than 1.5% (relative to complete inspection). The dashed (red) horizontal line in Figure 69 is thus our maximum tolerable relative risk. To meet this performance requirement a minimum of around 5% of passengers will have to be screened. At this level of screening, the robustness to uncertainty is zero and hence, if our initial estimates of the probability of a quarantine threat or of the detection probability are wrong, the performance requirement will not be met. Increasing the surveillance rate to 50% results in about 20% robustness, while an inspection rate of 85% will guarantee the performance requirement is met even if our initial guesses for the parameters are in error by 90%.

5-4-1 Comparison with a Bayesian Approach

The previous example has been modelled using the *WinBUGS* software. The directed acyclic graph is shown in Figure 70. Beta priors have been placed on θ and ϕ. In particular:

$$\theta \sim unif(0.05, 0.15)$$

and

$$\phi \sim unif(0.1, 0.8)$$

(Note: ϕ has a truncated distribution). The respective prior densities for θ and ϕ are shown in Figures 71 and 72. The *WinBUGS* code is shown in Figure 73.

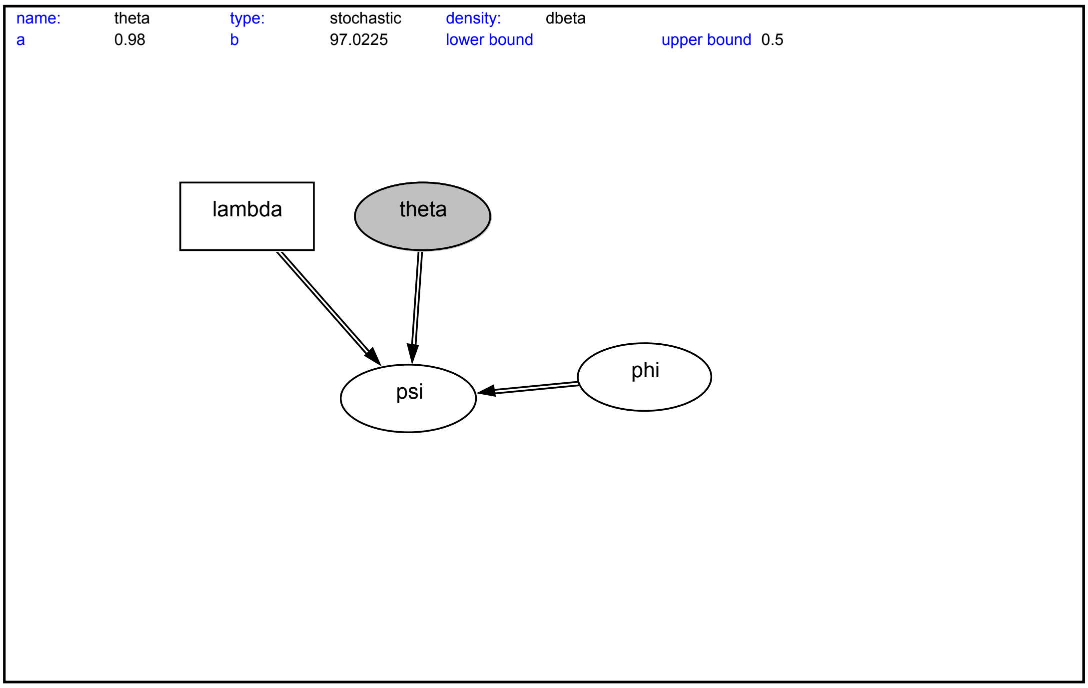

Figure 70. Directed acyclic graph for the biosurveillance example.

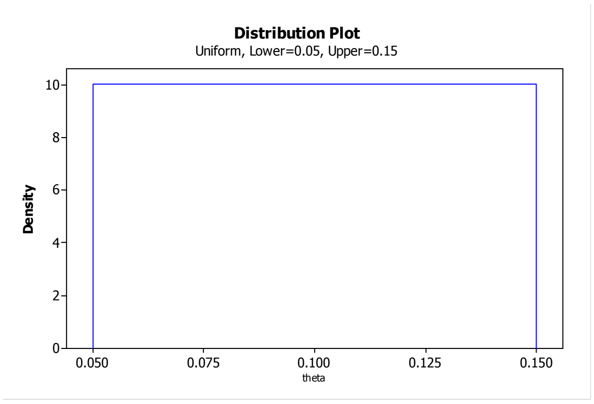

Figure 71. Prior density for theta.

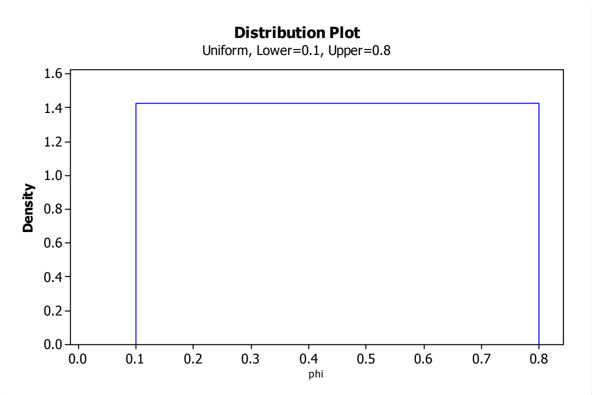

Figure 72. Prior density for phi.

```
model;
{
theta~dunif(0.05,0.15)
phi~dunif(0.1,0.8)
psi<-phi*(1-lambda*theta)*(1-theta*phi)/(1-phi*theta*lambda)/(phi*(1-theta))
}

list(lambda=0.85)
```

Figure 73. *WinBugs* code for biosurveillance example

Empirical *cdfs* have been plotted for a variety of λ values (Figure 74). Using a notional minimum performance criterion of $\Psi = 1.015$, we see that with a 60% inspection rate, there is about a 25% probability of meeting this target. This probability increases to around 90% for sampling inspection rates of at least 85%. These results accord well with the IG analysis.

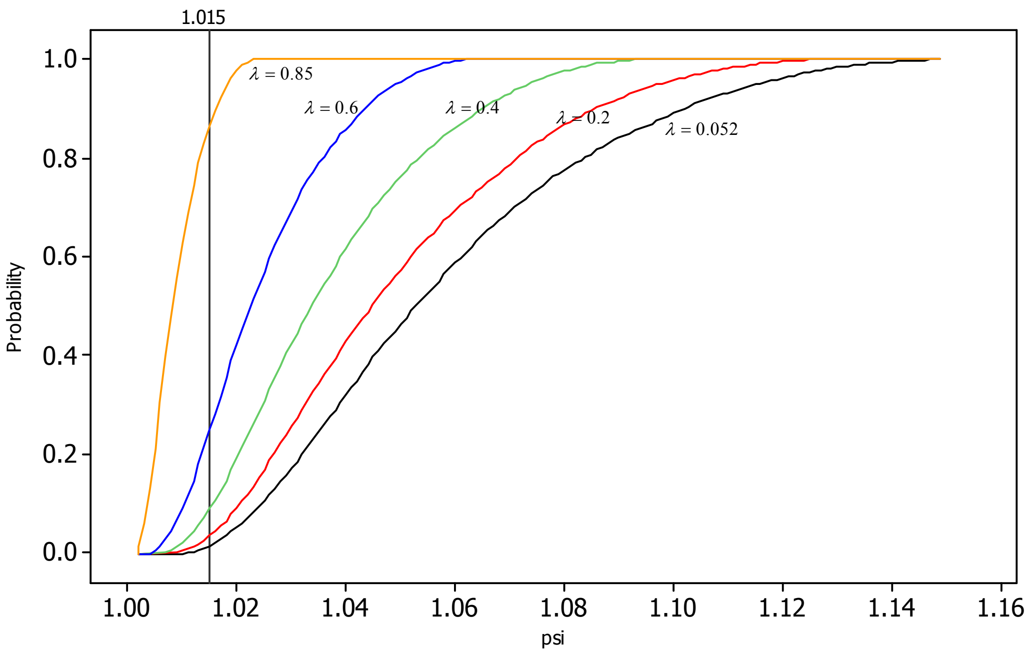

Figure 74. Emprical cdfs for performance measure of equation 5.11 based on 10,000 simulated results

5-5 Discussion

Physical and biosecurity is maintained and enhanced by a combination of activities and strategies, not least of which are border inspections of people and containers. A characteristic linking both bio-terrorism and biosecurity are the "unknown unknowns"[1] – that is, we often don't know what it is we're looking for. Monitoring strategies for the detection of invasive plant or animal species are particularly problematic due to the general absence of economic considerations and the climate of severe uncertainty about the likelihood of species introductions and successful detection (Moffitt et al. 2008).

Moffitt et al. (2008) used Info-Gap theory (Ben-Haim, 2006) to deal with the inherent uncertainty in biosecurity monitoring while Thompson and Fox (2008) independently used the same approach in the context of bioterrorism surveillance. In this chapter we have followed the approach outlined in Thompson and Fox (2008) and also compared this with the results of a Bayesian approach to the problem of determining an appropriate level of monitoring effort. With the choice of priors used in the Bayesian analysis to reflect reasonable assumptions about our knowledge (or ignorance) of detection and threat probabilities, it was found that both the IG and Bayseian methods resulted in similar inspection rates. We suspect this is more coincidence than a convergence of paradigms. The IG approach is, despite its appearance, relatively unsophisticated in its treatment of uncertainty. Whereas Bayesian methods characterise and manipulate uncertainty via probability functions or probability density functions, the info-gap method is deterministic and is essentially a sensitivity analysis on selected model parameters. While there is nothing inherently wrong with this, we suspect that the use of realistic prior distributions coupled with informative probability models for inspection and detection are likely to yield more informative results than the IG method. One of the difficulties with the IG approach is the interpretation of the *robustness* metric that is central to the IG paradigm (Fox, 2008). Other issues with the IG philosophy have been raised by Sniedovich (2007). Irrespective of the method used, a recurring message from these types of analyses is that the invariably small, flat-rate inspection policies that are in widespread use at present are unlikely to provide

[1] This term has been attributed to former US Defence Secretary Donald Rumsfeld who used it during a press briefing on Afghanistan on February 12, 2002.

adequate levels of immunity to the severe uncertainty that characterises biosecurity and biosurveillance. Given the importance and level of interest in robust decision-making for biosecurity, considerably more work needs to be done in this area.

6 References

Anderson, L.B., Atwell, R.J., Barnett, D.S. and Bovey, R.L. (2007) Application of the Maximum Flow Problem to Sensor Placement on Urban Road Networks for Homeland Security. *Homeland Security Affairs*, **3(3)**, 1-15.

Animal Health Australia (2007). Disease strategy: Equine influenza (Version 3.0). Australian Veterinary Emergency Plan (AUSVETPLAN), Edition 3, Primary Industries Ministerial Council, Canberra, ACT.

Alexandersen, S., Oleksiewicz, M. and Donaldson, A. (2001) The Early Pathogenesis of foot and mouth disease in pigs infected by contact: a quantitative time-course study using TaqMan RT-PCR. *J. Gen. Virol.* **82**, 747-755.

Arakaki, R.G.I. and Lorena, L.A.N. (2001) A Constructive Genetic Algorithm for the Maximal Covering Location Problem. *Proc. MIC'2001- 4th Metaheuristics International Conference, Porto Portugal, July 16-20, 2001.*

AusVet and CSIRO (2005) Review of the Potential Impacts of New Technologies on Australia's Foot and Mouth Disease (FMD) Planning and Policies – Report to Australian Government Department of Agriculture, Fisheries and Forestry. DAFF, Canberra. (*Available at:* http://www.daff.gov.au/__data/assets/pdf_file/0004/146875/ausvet_fmd_technology_fullreport_jul06.pdf)

Baron, M.I. (2001) Bayes Stopping Rules in a Change-Poiny Model with a Random Hazard Rate. *Sequential Analysis*, **20(3)**, 147-163.

Ben-Haim, Y. (2006). *Information-gap decision theory: Decisions under severe uncertainty.* 2nd. edition, Academic Press, San Diego.

Benneyan, J.C. (2001a) Number-Between g-type statistical quality control charts for monitoring adverse events. *Health Care Management Science*, **4**, 305-318.

Benneyan, J.C. (2001b) Performance of Number-Between g-type statistical quality control charts for monitoring adverse events. *Health Care Management Science*, **4**, 319-336.

Benyoussef, A., Boccara, N. Chakib, H., and Ez-Zahraouy (1999) Lattice three-species models of the spatial spread of rabies among foxes. *Int. J. Modern Phys. C*, **22**, 1025-1038.

Buczak, A.L., Wang, H., Darabi, H. and Jafari, M.A. (2001) Genetic algorithim convergence study for sensor network optimization. *Information Sciences*, **133**, 267-282.

Burkhom, H.S., Murphy, S.P. and Shmueli, G. (2007) Automated Time Series Forecasting for Biosurveillance. *Statistics in Medicine*, **26**,4202–4218.

Callinan, I. (2008) Equine influenza - The August 2007 outbreak in Australia Report of the Equine Influenza Inquiry. The Hon. Ian Callinan AC April 2008 Commonwealth of Australia

Cannon, R.M. and Garner, M.G. (1999) Assessing the risk of wind-borne spread of foot-and-mouth disease in Australia. *Environment International*, **25(6/7**), 713-723.

Carpenter, T.E. (2001) Methods to investigate spatial and temporal clustering in veterinary epidemiology. *Preventative Veterinary Medicine*, **48**, 303-320.

Carmel, Y. and Ben-Haim, Y. (2005) Info-gap robust-satisficing model of foraging behavior: Do foragers optimize or satisfice? *American Naturalist*, **166**, 633-641.

Chakrabarty, S. and Iyengar, S.S. (2002) Sensor placement for grid coverage under imprecise detections. *Proceedings of the Fifth International Conference on Information Fusion* 1581-158.

Chakrabarty, K., Iyengar, S.S., Qi, H. and Cho, E. (2002) Grid coverage for surveillance and target location in distributed sensor networks. *IEEE Transactions on Computers*, **51(12)**, 1448-1453.

Chuang, C. and Lin, R. (2007) A maximum Expected Covering Model for an Ambulance Location Problem. *J. Chinese Inst. Indust. Eng.* **24(6)**, 468-474.

Church, R. and ReVelle, C. (1974) The Maximal Covering Location Problem. *Papers of the Regional Science Association*, **32**, 101-118.

Church, R. (1984) The Planar Maximal Covering Location Problem. *J. Regional Science*, **24(2)**, 185-201.

Colosimo, B. M. and del Castillo, E. eds. (2007) *Bayesian Process Monitoring, Control and Optimization.* Chapman and Hall / CRC, Boca Raton.

Commonwealth of Australia (2002) Meat hygiene assessment: Objective methods for the monitoring of processes and products. 2nd. Edition. Canberra, Australia. http://www.daff.gov.au/corporate_docs/publications/pdf/quarantine/mid/msqa2ndedition.pdf

Doherr, M.G. and Audigé, L. (2001) Monitoring and surveillance for rare health-related events: a review from the veterinary perspective. *Philos.Trans.R.Soc.London Ser.B,* **356**, 1097–1106.

Doran, R.J. and Laffan, S.W. (2005) Simulating the spatial dynamics of foot and mouth disease outbreaks in feral pigs and livestock in Queensland, Australia using a susceptible-infected-recovered cellular automata model. *Preventive Veterinary Medicine*, **70**, 133-152.

Donaldsen, A.I. (1983) Quantitive data on airborne FMD virus; its production, carriage and deposition. *Phil. Trans. Roy. Soc. Lond. Ser B – Biol. Sci.,* **302**,529-534.

Dooley, K., and R. Flor (1998) Perceptions of Success and Failure in Total Quality Management Initiatives. *J. of Quality Management*, **3(2)**, 157-174.

Department of Primary Industries Victoria (2008) Equine Influenza in Australia 2007: Victoria Situation Report Date: Friday 25 January 2008 – 1200 hrs. http://www.polocrossevic.org.au/DocumentDownload/SITREP_%20INDUSTRY_EISDCHQ_20080125.pdf

Department of Primary Industries NSW (2008) Equine influenza FAQs - About the disease and symptoms. http://www.dpi.nsw.gov.au/agriculture/livestock/horse/influenza/faqs/disease

Downs, B.T., and Camm, J.D. (1996) An Exact Algorithm for the Maximal Covering Problem. *Naval Research Logistics*, **43**, 435-461.

Eaton, D.J., Church, R.L. and ReVelle (1977) Location Analysis: A New Tool for Health Planners. Methodological Working Document No. 53, Sector Analysis Division, Bureau for Latin America, Agency for International development.

FAO (1998) Guidelines for Surveillance, International Standards for Phytosanitary Measures Publication No. 6 (http://www.ippc.int/IPP/En/ispm.htm)

FAO (2005) Regional Standards for Phytosanitary Measures requirements for the establishment and maintenance of Pest Free areas for Tephritid Fruit Flies. RAP Publication 2005/26. The Asia and Pacific Plant Protection Commission (APPPC)

Fox, D.R. (2008) . To IG or not to IG? – that is the question. Decision-making under uncertainty. *Decision Point* 24 10-11 (Available at: http://www.aeda.edu.au/docs/Newsletters/DPoint_24.pdf).

Fox, D.R., Bisignanesi, V., Joynt, B., and McGuire, R. (2009) Optimising Victoria's Air Quality Monitoring Network. The Australian Centre for Environmetrics, January 2009, University of Melbourne, Parkville, Australia.

Fox, J.C., Buckley, Y.M., Panetta, F.D., Bourgoin, J., and Pullar, D. (2009) Surveillance protocols for management of invasive plants: modelling Chilean needle grass (Nassella neesiana) in Australia. *Diversity and Distributions*, *(in press).*

Fricker, R.D. (2008) Syndromic Surveillance. In *Encyclopedia of Quantitative Risk Assessment and Analysis*, Melnick, E. and Everitt, B. (eds). John Wiley & Sons Ltd, Chichester, UK.

Fuentes, M.A. and Kuperman, M.N. (1999) Cellular Automata and Epidemiological Models with Spatial Dependence. *Physica A*, **267**, 471-486.

Garner, M.G. and Cannon, R.M. (1995) Potential for Wind-Borne Spread of Foot and Mouth Disease Virus in Australia - A report prepared for the Australian Meat

Research Corporation. *(available at:* http://www.daff.gov.au/__data/assets/pdf_file/0007/159541/fmdwind.pdf)

Hall, J.A. and Golding, L. (1998) Standard methods for whole effluent toxicity testing: Development and Application. NIWA Client report MfE80205, National Institute of Water and Atmospheric Research, Hamilton, New Zealand.

Hamada, M. (2002) Bayesian Tolerance Interval Control Limits for Attributes. *Qual. Reliab. Engng. Int.* **18**, 45-52.

Hogan, W.R., Cooper, G.F., Wallstrom, G.L., Wagner, M.M. and Depinay, J. (2007) The Bayesian aerosol release detector: An algorithm for detecting and characterizing outbreaks caused by an atmospheric release of *Bacillus anthracis*. *Statistics in Medicine*, **26**, 5225-5252.

Höhle, M. and Paul, M. (2008) Count data regression charts for the monitoring of surveillance time series. *Computational Statistics and Data Analysis*, **52**, 4357-4368.

Kanaroglou, P.S., Jerrett, M., Morrison, J., Beckerman, B., Arain, M.A., Gilbert, N.L., and Brook, J.R. (2005) Establishing an air pollution monitoring network for intra-urban population exposure assessment: A location-allocation approach. *Atmospheric Environment* **39**, 2399-2409.

Korth, W. (1997) Determination and application of recovery factors in chemical residue testing of food products by Australian laboratories associated with the national residue survey. http://www.daff.gov.au/corporate_docs/publications/pdf/animalplanthealth/nrs/recovery_factors_korth_1997.pdf

Krause, A., Guestrin, C., Gupta, A., and Kleinberg, J. (2006) Near-optimal sensor placements: maximizing information while minimizing communication cost. *Proceedings of the fifth international conference on Information processing in sensor networks* 2-10.

Kulldorff, M. (1997) A spatial scan statistic. *Communications in Statistics: Theory and Methods*, **26**:1481-1496.

Kulldorff, M. (2001) Prospective time-periodic geographical disease surveillance using a scan statistic. *Journal of the Royal Statistical Society* (Series A), **164**, 61-72.

Lindo Systems (2007) *Lingo User's Guide*. Lindo Systems Inc., Chicago.

Liu, M.K., Avrin, J., Pollack, R.I., Behar, J.V., and McElroy, J.L. (1986) Methodology for designing air quality monitoring networks: I. Theoretical aspects. *Environmental Monitoring and Assessment* **6**, 1-11.

Marshall, C., Best, N., Bottle, A. and Aylin, P. (2004) Statistical Issues in the prospective monitoring of health outcomes across multiple units. *J. Royal Statist. Soc. (A),* **167(3)**, 541-559.

Menzefricke, U. (2002) On the Evaluation of Control Chart Limits Based on Predictive Distributions. *Commun. Statist. – Theory Meth.*, **31(8)**, 1423-1440.

Moffitt, L.J., Stranlund, J.K., and Field, B.C. (2005) Inspections to Avert Terrorism: Robustness Under Severe Uncertainty. *Journal of Homeland security and Emergency Management*, **2(3)**, 1-17.

Moffitt, L.J., Stranlund, J.K. and Osteen, C.D. (2008) Robust detection protocols for uncertain introductions of invasive species. *J. Environmental Management*, **89**, 293-299.

Mostashari, F. and Hartman, J. (2003) Syndromic Surveillance: a Local Perspective. *Journal of Urban Health: Bulletin of the New York Academy of Medicine* **80(2)**, Supplement 1.

Nairn, M.E., Allen, P.G., Inglis, A.R. and Tanner, C. (1996) *Australian quarantine : a shared responsibility* . Canberra, A.C.T. : Dept. of Primary Industries and Energy, 1996.

Noel, S. and Jajodia, S. (2007) Attack Graphs for Sensor Placement, Alert Prioritization, and Attack Response. *Presented at the Cyberspace Research Workshop - Air Force Cyberspace Symposium.*

Önal, H. (2003) First-best, second-best, and heuristic solutions in conservation reserve site selection. *Biological Conservation*, **115**, 55-62.

O'Rourke (1987) *Art Gallery Theorems and Algorithms.* Oxford University Press, New York, NY.

Patil, G.P. Acharaya, R., Glasmeier, A., Myers, W., Phoha, S. and Rathbun, S. (2006) Hotspot Detection and Prioritization GeoInformatics for Digital Governance. Penn State University, Center for Statistical Ecology and Environmental Statistics, Department of Technical Report Number 2006-0516.

Pheloung, P. (2004) Plant pest surveillance in Australia, *Australian Journal of Emergency Management,* **19(3)**, 13-16.

Pheloung, P. (2005) Contingency planning for plant pest incursions in Australia In*: Identification of risks and management of invasive alien species using the IPPC framework. Proceedings of a workshop in Braunschweig, Germany 22–26 September 2003* . FAO (2005).

Pirkul, H. and Schilling, D. (1988) The Siting of Emergency service Facilities with Workload capacities and Backup services. *Management Science*, **34(8)**, 896-908.

Radaelli, G. (1998) Planning time-between-events Shewhart control charts. *Total Quality Management*, **9(1**), 133-140.

Regan, H.M., Ben-Haim, Y., Langford, W., Wilson, W.G., Lundberg, P., Andelman, S.J. and Burgman, M.A. (2005) Robust decision making under severe uncertainty for conservation management, *Ecological Applications*, **15(4)**, 1471-1477.

Riley, S. (2007) Large-Scale Spatial-Transmission Models of Infectious Disease. *Science*, **316**, 1 June 2007, 1298-1301.

Robinson, A., Burgman, M., Atkinson, W., Cannon, R., Miller, C., and Immonen, H., (2009) Import Clearance Data Framework. ACERA Project 0804 Final Report. Australian centre of Excellence for Risk Analysis, University of Melbourne.

Shea, D.A. and Lister, S.A. (2003) The BioWatch Program: Detection of Bioterrorism Congressional Research Service Report No. RL 32152 November 19, 2003. Available at: http://www.fas.org/sgp/crs/terror/RL32152.html

Shmueli, G., and Burkom, H.S. (in press) Statistical Challenges Facing Early Outbreak Detection in Biosurveillance", *Technometrics (Special Issue on Anomaly Detection*).

Sniedovich, M. (2007) A Critique of Info-Gap: Myths and Facts. Second SRA Conference August 20-21, 2007 Hobart, TAS, Australia. (Available at: http://www.acera.unimelb.edu.au/materials/papers%2007/Moshe_Sniedovich_SRA.pdf).

Sonesson, C. and Bock, D. (2003) A Review and Discussion of Prospective Statistical Surveillance in Public Health. *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, **166(1)**, 5-21.

Sørensen, J.H., Mackay, D., Jensen, C.O., and Donaldson, A. (2000) An integrated model to predict the atmospheric spread of foot and mouth disease virus. *Epidemiol. Infect.* **26**, 93-97.

Stark, K.D.C., Regula, G., Hernandez, J., Knopf, L., Fuchs, K., Morris, R.S. and Davies, P. (2006) Concepts for risk-based surveillance in the field of veterinary medicine and veterinary public health: Review of current approaches BMC Health Services Research **6(20).** Available at http://www.biomedcentral.com/1472-6963/6/20

Stephens. M.A. (1974) EDF Statistics for Goodness of Fit and Some Comparisons, *Journal of the American Statistical Association*, **69(347)**, 730-737

Thompson, C.J. (unpublished) Modelling Risk and Uncertainty in Bio- and Homeland security. Presented at Australian and new Zealand Society for Risk Analysis Conference, University of Melbourne, July 17-19, 2006.

Thompson, C.J. and Fox, D.R. (2008) Sampling and Inspection for Monitoring Threats to Homeland Security, in *Encyclopedia of Quantitative Risk Assessment and Analysis*, Melnick, E. and Everitt, B. (eds). John Wiley & Sons Ltd, Chichester, UK, pp 1600-1603.

Tsiamyrtzis, P., and Hawkins D.M. (2007) *Bayesian Process Monitoring, Control and Optimization.* Chapman & Hall/CRC Press London/Boca Raton, FL.

Underwood, A.J. (1994) On beyond BACI: Sampling designs that might reliably detect environmental disturbance. *Ecological Applications*, **4(1),** 3-15.

Vickers, L., Stolkin, R. and Nickerson, J.V. (2006) Computational Environmental Models Aid Sensor Placement Optimization. *Military Communications Conference, 2006. MILCOM 2006* 1-6.

Wagner, M.M, Aryel, R.M., and Moore, A.W. (2006) *Handbook of Biosurveillance*, Elsevier.

Weisstein, E. W. (*undated*) NP-Hard Problem. From *MathWorld*--A Wolfram Web Resource. http://mathworld.wolfram.com/NP-HardProblem.html

Wong, W., Cooper, G., Dash, D., Levander, J., Dowling, J., Hogan, W., and Wagner, M. (2005). Use of Multiple Data Streams to Conduct Bayesian Biologic Surveillance, *Morbidity and Mortality Weekly Report*, **54** (supplemental), pp. 63-69.

Yoffe, A. and Ben-Haim, Y. (2006) An Info-Gap Approach to Policy Selection for Bio-Terror Response. IEEE International Conference on Intelligence and Security Informatics, ISI 2006, San Diego, CA, USA May 23-24 2006, 554-559.

APPENDIX A : DATA USED IN CHAPTER 1 EXAMPLE

*	5.966	23.313	20.971	18.324	6.164	8.349	1.108	19.483	9.941
3.563	3.743	1.552	9.793	2.808	22.799	2.789	0.015	0.104	2.83
21.088	0.702	9.899	21.599	1.16	0.965	0.195	21.147	0.423	5.451
16.495	2.707	1.047	5.013	5.712	7.103	4.145	5.057	8.668	19.926
6.384	2.808	15.013	7.363	0.349	1.98	4.654	6.911	2.904	0.424
2.257	3.409	9.08	10.489	14.868	0.123	18.668	1.915	0.968	4.609
6.428	12.826	11.432	13.102	21.721	1.839	8.508	6.479	10.587	10.6
10.602	1.344	5.558	6.039	1.375	7.862	7.242	3.396	19.974	32.228
4.928	8.224	4.206	0.734	4.386	0.462	0.597	21.667	7.259	2.426
6.356	10.23	2.183	0.959	2.917	9.976	5.189	0.72	5.679	10.163
5.903	7.936	0.701	11.881	11.151	14.205	6.05	0.779	8.724	7.169
13.771	0.61	1.1	3.945	1.053	0.779	1.41	3.454	1.356	0.061
19.92	31.168	7.937	12.486	22.916	1.578	2.46	3.691	9.384	0.44
6.278	0.113	4.232	2.348	4.667	4.774	5.353	2.536	7.808	7.331
10.9	16.738	2.292	3.082	13.185	1.235	2.471	17.38	1.407	10.346
17.379	5.669	3.893	1.8	2.041	3.875	15.387	17.414	1.693	6.714
34.181	6.708	2.954	6.668	6.965	0.528	0.756	4.081	4.881	2.697
5.011	2.028	11.68	15.126	2.506	7.047	5.783	1.436	3.053	5.868
9.69	1.605	23.962	5.221	4.929	2.158	0.27	12.947	16.342	0.436
5.259	2.887	1.46	3.164	1.665	36.372	15.661	1.749	0.763	2.477
3.188	2.569	5.877	11.601	0.448	7.424	0.594	13.119	9.131	2.206
0.585	4.692	3.905	3.994	31.585	3.071	13.219	8.659	0.326	13.719
5.003	15.532	1.169	3.618	3.98	1.698	0.319	4.415	2.548	2.379
25.935	15.764	1.957	5.437	34.542	5.101	7.958	13.953	5.546	0.96
0.576	11.084	1.708	6.436	0.195	1.679	12.284	25.663	0.964	7.051
12.53	4.02	22.363	108.031						

Table entries are days

APPENDIX B: Derivation of Equation 2-5

Equation 2-5 is given in the text as:

$$p\left(y\middle|a,b\right)=\frac{N^y}{y!}\prod_{j=0}^{y-1}\left(\frac{j+a}{j+a+b}\right)\left\{1+\sum_{m=1}^{\infty}\left[\prod_{r=0}^{m-1}\frac{y+a+r}{y+a+b+r}\right]\frac{\left(-N\right)^m}{m!}\right\};$$
$$y=\left\{0,1,\ldots N\right\},\ a,b>0 \qquad \text{(A1)}$$

The proof follows.

We have $p\left(y\middle|a,b\right)=\int_0^1\left[\frac{e^{-N\theta}\left(N\theta\right)^y}{y!}\right]\frac{1}{\beta(a,b)}\theta^{a-1}\left(1-\theta\right)^{b-1}\,d\theta$ which can be re-written as

$$\left[\frac{N^y\beta(y+a,b)}{y!\beta(a,b)}\right]\int_0^1 e^{-N\theta}\,\frac{1}{\beta(y+a,b)}\theta^{y+a-1}\left(1-\theta\right)^{b-1}\,d\theta \qquad \text{(A2)}$$

The integrand in equation A2 is the expected value of $e^{-N\theta}$ with respect to the $\beta(y+a,b)$ density.

The moment-generating function for a $\beta(y+a,b)$ density is

$$M_\theta\left(t\right)=E\left[e^{t\theta}\right]=1+\sum_{m=1}^{\infty}\left[\prod_{r=0}^{m-1}\frac{y+a+r}{y+a+b+r}\right]\frac{t^m}{m!} \qquad \text{(A3)}$$

And thus the integral in equation A2 is simply $M_\theta\left(-N\right)$ and hence

$$\left[\frac{N^y\beta(y+a,b)}{y!\beta(a,b)}\right]\int_0^1 e^{-N\theta}\,\frac{1}{\beta(y+a,b)}\theta^{y+a-1}\left(1-\theta\right)^{b-1}\,d\theta$$

$$=\left[\frac{N^y\beta(y+a,b)}{y!\beta(a,b)}\right]\left\{1+\sum_{m=1}^{\infty}\left[\prod_{r=0}^{m-1}\frac{y+a+r}{y+a+b+r}\right]\frac{(-N)^m}{m!}\right\} \qquad \text{(A4)}$$

It is relatively straightforward to show that the ratio of the two beta terms in equation A4 can be expressed as $\prod_{j=0}^{y-1}\left(\frac{j+a}{j+a+b}\right)$, thus completing the proof.

APPENDIX C: Derivation of Equation 2-9

Equation 2-9 is given in the text as

$$p\left[X_{t+1}\middle|y_t\right]=\binom{n_{t+1}}{x_{t+1}}\frac{\beta\left(x_{t+1}+y_t+a,n_{t+1}-x_{t+1}+b\right)}{\beta\left(y_t+a,b\right)}\frac{\left\{1+\sum_{m=1}^{\infty}\left[\prod_{r=0}^{m-1}\frac{x_{t+1}+y_t+a+r}{n_{t+1}+y_t+a+b+r}\right]\frac{\left(-N_t\right)^m}{m!}\right\}}{\left\{1+\sum_{m=1}^{\infty}\left[\prod_{r=0}^{m-1}\frac{y_t+a+r}{y_t+a+b+r}\right]\frac{\left(-N_t\right)^m}{m!}\right\}} \tag{B.1}$$

The proof follows.

As given in the text, $N=\sum_{i=1}^{t}n_i$ is the total number of sampled units and Y is the total number of failures at time *t* i.e. $Y=\sum_{i=1}^{t}X_i$. Now, $p\left[X_{t+1}\middle|y_t\right]=\int_0^1 f\left(X_{t+1}\middle|\theta\right)p\left(\theta\middle|y_t\right)d\theta$ with $p\left(\theta\middle|y_t\right)=\dfrac{p\left(y_t\middle|\theta\right)p\left(\theta\right)}{p\left(y_t\right)}$.

Furthermore, $p\left(y_t\right)=\int_0^1 p\left(y_t\middle|\theta\right)p\left(\theta\right)d\theta$ with $p\left(\theta\right)=\dfrac{1}{\beta\left(a,b\right)}\theta^{a-1}\left(1-\theta\right)^{b-1}$.

Now, $Y\middle|\theta\sim bin\left(N,\theta\right)$ but since N is large and θ small it is reasonable to assume $Y\middle|\theta\sim Poisson\left(N\theta\right)$ i.e. $p\left(Y\middle|\theta\right)=\dfrac{e^{-N\theta}\left(N\theta\right)^y}{y!}$. Hence we obtain the *pdf* for y_t as

$$\begin{aligned}p\left(y_t\right)=\int_0^1 p\left(y_t\middle|\theta\right)p\left(\theta\right)d\theta&=\int_0^1\left\{\frac{e^{-N\theta}\left(N\theta\right)^{y_t}}{y_t!}\right\}\frac{1}{\beta\left(a,b\right)}\theta^{a-1}\left(1-\theta\right)^{b-1}d\theta\\&=\frac{N^{y_t}}{y_t\,\beta\left(a,b\right)}\int_0^1\left[e^{-N\theta}\right]\theta^{y_t+a-1}\left(1-\theta\right)^{b-1}d\theta\\&=\frac{N^{y_t}}{y_t\,\beta\left(a,b\right)}\beta\left(y_t+a,b\right)\int_0^1 e^{-N\theta}\frac{1}{\beta\left(y_t+a,b\right)}\theta^{y_t+a-1}\left(1-\theta\right)^{b-1}d\theta\end{aligned}$$

$$= \frac{N^{y_t}}{y_t\, \beta(a,b)} \beta(y_t + a, b)\, E_\theta\left[e^{-N\theta}\right] \tag{B.2}$$

where $E_\theta\left[e^{-N\theta}\right]$ is the expectation of $-N\theta$ with respect to the *pdf* for θ [i.e. the $\beta(y_t + a, b)$ density].

From mathematical distribution theory, it is known that the *moment generating function* for the beta density is defined as:

$$\begin{aligned} M_\theta(t) &= E\left[e^{t\theta}\right] \\ &= 1 + \sum_{m=1}^{\infty} \left[\prod_{r=0}^{m-1} \frac{y_t + a + r}{y_t + a + b + r} \right] \frac{t^m}{m!} \end{aligned} \tag{B.3}$$

Given that $E_\theta\left[e^{-N\theta}\right] = M_\theta\left[-N\right]$ it follows that

$E_\theta\left[e^{-N\theta}\right] = 1 + \sum_{m=1}^{\infty} \left[\prod_{r=0}^{m-1} \frac{y_t + a + r}{y_t + a + b + r} \right] \frac{(-N)^m}{m!}$ and hence:

$$p(y_t) = \frac{N^{y_t}}{y_t\, \beta(a,b)} \beta(y_t + a, b) \left\{ 1 + \sum_{m=1}^{\infty} \left[\prod_{r=0}^{m-1} \frac{y_t + a + r}{y_t + a + b + r} \right] \frac{(-N)^m}{m!} \right\} \tag{B.4}$$

Substituting equation B.4 into $p(\theta | y_t) = \frac{p(y_t | \theta)\, p(\theta)}{p(y_t)}$ we obtain:

$$p(\theta | y_t) = \frac{\dfrac{e^{-N\theta} (N\theta)^{y_t}}{y_t!} \dfrac{1}{\beta(a,b)} \theta^{a-1} (1-\theta)^{b-1}}{\dfrac{N^{y_t}}{y_t\, \beta(a,b)} \beta(y_t + a, b) \left\{ 1 + \sum_{m=1}^{\infty} \left[\prod_{r=0}^{m-1} \dfrac{y_t + a + r}{y_t + a + b + r} \right] \dfrac{(-N)^m}{m!} \right\}}$$

$$= \frac{e^{-N\theta}\, \theta^{y_t + a - 1} (1-\theta)^{b-1}}{\beta(y_t + a, b) \left\{ 1 + \sum_{m=1}^{\infty} \left[\prod_{r=0}^{m-1} \dfrac{y_t + a + r}{y_t + a + b + r} \right] \dfrac{(-N)^m}{m!} \right\}} \tag{B.5}$$

Returning to $p\left[X_{t+1}\left|y_t\right.\right]$ we can write $p\left[X_{t+1}\left|y_t\right.\right]=\int_0^1 f\left(x_{t+1}\left|\theta\right.\right)p\left(\theta\left|y_t\right.\right)d\theta$. Now,

$$\int_0^1 f\left(x_{t+1}\left|\theta\right.\right)p\left(\theta\left|y_t\right.\right)d\theta=$$

$$\int_0^1\left[\binom{n_{t+1}}{x_{t+1}}\theta^{x_{t+1}}\left(1-\theta\right)^{n_{t+1}-x_{t+1}}\right]\frac{e^{-N\theta}\,\theta^{y_t+a-1}\left(1-\theta\right)^{b-1}}{\beta\left(y_t+a,b\right)\left\{1+\sum_{m=1}^{\infty}\left[\prod_{r=0}^{m-1}\frac{y_t+a+r}{y_t+a+b+r}\right]\frac{\left(-N\right)^m}{m!}\right\}}d\theta$$

$$=\frac{\binom{n_{t+1}}{x_{t+1}}\int_0^1\left[e^{-N\theta}\,\theta^{x_{t+1}+y_t+a-1}\left(1-\theta\right)^{n_{t+1}-x_{t+1}+b-1}\right]d\theta}{\beta\left(y_t+a,b\right)\left\{1+\sum_{m=1}^{\infty}\left[\prod_{r=0}^{m-1}\frac{y_t+a+r}{y_t+a+b+r}\right]\frac{\left(-N\right)^m}{m!}\right\}}$$

$$=\frac{\binom{n_{t+1}}{x_{t+1}}\beta\left(x_{t+1}+y_t+a,n_{t+1}-x_{t+1}+b\right)}{\beta\left(y_t+a,b\right)\left\{1+\sum_{m=1}^{\infty}\left[\prod_{r=0}^{m-1}\frac{y_t+a+r}{y_t+a+b+r}\right]\frac{\left(-N\right)^m}{m!}\right\}}\int_0^1\left[e^{-N\theta}\frac{\theta^{x_{t+1}+y_t+a-1}\left(1-\theta\right)^{n_{t+1}-x_{t+1}+b-1}}{\beta\left(x_{t+1}+y_t+a,n_{t+1}-x_{t+1}+b\right)}\right]d\theta \tag{B.6}$$

Noting that the integrand in equation B.6 is a beta *pdf*, we have:

$$p\left[X_{t+1}\left|y_t\right.\right]=\frac{\binom{n_{t+1}}{x_{t+1}}\beta\left(x_{t+1}+y_t+a,n_{t+1}-x_{t+1}+b\right)}{\beta\left(y_t+a,b\right)\left\{1+\sum_{m=1}^{\infty}\left[\prod_{r=0}^{m-1}\frac{y_t+a+r}{y_t+a+b+r}\right]\frac{\left(-N\right)^m}{m!}\right\}}E\left[e^{-N\theta}\right]$$

$$= \frac{\binom{n_{t+1}}{x_{t+1}} \beta\left(x_{t+1}+y_t+a, n_{t+1}-x_{t+1}+b\right)}{\beta\left(y_t+a, b\right)\left\{1+\sum_{m=1}^{\infty}\left[\prod_{r=0}^{m-1} \frac{y_t+a+r}{y_t+a+b+r}\right] \frac{(-N)^m}{m!}\right\}}\left\{1+\sum_{m=1}^{\infty}\left[\prod_{r=0}^{m-1} \frac{x_{t+1}+y_t+a+r}{n_{t+1}+y_t+a+b+r}\right] \frac{(-N)^m}{m!}\right\}$$

thus completing the proof.

APPENDIX D : DATA USED IN CHAPTER 3 EXAMPLE

Row	Column	x	n	Row	Column	x	n
20	13	0	10	24	20	0	9
22	26	0	12	12	7	0	13
3	16	0	9	20	15	0	8
8	23	1	13	25	24	0	9
1	24	0	4	24	17	0	8
2	19	0	11	13	17	4	5
14	14	1	7	17	18	0	11
24	24	0	17	18	22	4	12
16	16	0	10	26	5	0	7
26	21	0	14	10	8	1	9
19	20	0	3	20	6	0	17
16	13	0	6	25	7	0	14
12	11	2	10	12	2	0	8
13	2	0	13	20	3	0	14
16	1	0	8	9	6	2	8
2	17	0	15	25	3	0	15
10	1	0	11	14	7	0	9
15	19	6	8	7	21	0	6
19	17	0	12	25	22	0	12
18	5	0	10	3	13	1	15
17	4	0	6	12	9	0	6
7	26	0	10	23	8	0	17
15	8	0	6	21	24	1	13
12	24	3	14	14	2	0	15
6	22	0	8	14	15	2	14
26	22	0	12	21	6	0	9
4	7	7	11	8	16	6	8
4	16	0	13	2	5	3	11
12	18	10	10	22	22	0	11
4	10	9	11	24	7	0	5
23	25	0	11	8	18	1	6
5	14	3	11	3	11	1	8
16	14	0	5	16	23	5	6
9	5	0	9	17	15	0	14
1	7	0	4	3	23	0	7
6	8	12	13	2	24	0	15
4	22	0	6	25	16	0	6
14	4	0	12	6	11	6	10
23	18	0	11	21	9	0	10
22	15	0	11	6	12	7	14
26	9	0	9	1	3	3	11

Row	Column	x	n	Row	Column	x	n
26	8	0	13	22	9	0	12
18	9	0	11	11	11	3	9
8	22	0	10	18	15	0	9
4	6	7	9	20	26	0	8
12	1	0	10	17	19	1	9
13	15	2	7	13	6	0	8
12	19	6	7	13	19	12	13
19	15	0	14	5	3	2	7
2	9	2	9	1	6	1	5
1	17	0	10	16	20	3	8
10	21	2	13	4	13	0	8
10	5	0	11	25	17	0	6
2	6	3	8	9	19	3	7
2	3	5	12	18	16	0	11
22	14	0	14	12	4	0	11
16	25	5	13	7	9	5	7
13	18	10	11	19	10	0	14
21	5	0	4	3	5	5	14
5	23	0	12	1	21	0	10
18	17	0	6	22	11	0	9
13	11	0	14	26	18	0	11
26	17	0	9	26	14	0	7
18	8	0	9	7	24	0	12
7	17	2	8	9	17	5	9
1	13	0	13	5	6	6	8
15	1	0	7	14	17	5	9
17	8	0	9	6	2	0	12
6	25	0	15	16	2	0	15
4	24	0	15	25	26	0	10
24	5	0	9	26	16	0	8
25	21	0	5	7	12	10	12
2	4	6	12	3	3	5	10

APPENDIX E : MATHCAD CODE FOR CELLULAR AUTOMATA MODEL

Establish grid dimensions

$R := 26 \qquad C := 26 \qquad i := 1..R \qquad j := 1..C$

Define spatial probability model(s)

$a := 0.0025 \qquad b := .0055$

$$p(r1,c1,r2,c2) := \exp\left[-0.5 \cdot \frac{\left[(r1-r2)^2 + (c1-c2)^2\right]}{.12}\right]$$

$$p1(r1,c1,r2,c2) := \exp\left[-1 \cdot \left[\frac{(r1-r2)^2}{a} + \frac{-2(r1-r2)\cdot(c1-c2)}{\sqrt{a \cdot b}} + \frac{(c1-c2)^2}{b}\right]\right]$$

Data Entry - observed cases and locations

Data matrix D: col1=row index; col2=col index; col3=r cases; col4=sample size N

$D :=$

$g(a,b) := 0 \qquad K := rows(D) - 1$

$I := matrix(R,C,g) \qquad L := 0..K$

$$P_{\left(D_{L,0}\right), D_{L,1}} := \frac{D_{L,2}}{D_{L,3}}$$

P is matrix of observed proportions

Initial probabilities

```
GG(r0,c0) := | for  i ∈ 1..R
             |   for  j ∈ 1..C
             |     g[i,j] ← p1(r0,c0,i,j)
             | M ← g
             | return  M
```

Update probabilities

```
M(A) := | B ← A
        | for ii ∈ R − 1 .. 1
        |   for jj ∈ C − 1 .. 1
        |     | P ← 1
        |     | for is ∈ (ii − 1) .. (ii + 1)
        |     |   for js ∈ (jj − 1) .. (jj + 1)
        |     |     P ← P·[A_is,js·(1 − p(is, js, ii, jj)) + (1 − A_is,js)]  if is ≠ ii ∨ js ≠ jj
        |     | B_ii,jj ← A_ii,jj + (1 − A_ii,jj)·(1 − P)
        | B
```

Likelihood Function

```
L1(G, tt) := | for it ∈ 0 .. tt
             |   | A ← G_it
             |   | sum ← 0
             |   | for il ∈ 0 .. K
             |   |   | θ ← if[A_(D_il,0), D_il,1 ≠ 0, A_(D_il,0), D_il,1, 10^−30]
             |   |   | θ ← if(θ ≠ 1, θ, 0.999999)
             |   |   | sum ← sum + [D_il,2·log(θ) + (D_il,3 − D_il,2)·log(1 − θ)]
             |   | BB_it ← sum
             | return BB
```

Iterative step

```
Lik(τ) := | for r0 ∈ 1..R
          |   for c0 ∈ 1..C
          |     | G_0 ← GG(r0, c0)
          |     | for T ∈ 1..τ
          |     |   G_T ← M(G_(T−1))
          |     | LL ← L1(G, τ)
          |     | for t ∈ 0..(rows(LL) − 1)
          |     |   x_t ← t
          |     | vs ← cspline(x, LL)
          |     | fit(xx) ← interp(vs, x, LL, xx)
          |     | dfit(xx) ← d/dxx fit(xx)
          |     | xx ← τ/2
          |     | τ  on error t0 ← root(dfit(xx), xx)
          |     | Tmax_(r0, c0) ← t0
          |     | Lmax_(r0, c0) ← fit(t0)
          | result ← stack(Tmax, Lmax)
          | return result
```

Execute Spatial-Temporal Likelihood Routine

Q := Lik(200) Call Likelihood routine and evaluate for 200 time increments

Qt := submatrix(Q, 3, 23, 3, 23)

Ql := submatrix(Q, 28, 53, 1, 26)

For each grid cell extract max. likelihood value and time

APPENDIX F : DATA USED IN CHAPTER 4 EXAMPLE

Source: NSW Department of Local Government Comparative Information on New South Wales:
http://www.dlg.nsw.gov.au/Files/Comparatives/0506data.xls (2005/06 population data).
Eastings and Northings were separately digitised and generally correspond to the location of the shire office or the main shire town.

LGA	Area	Popn	Density	East	North
Albury City Council	313	47247	150.9489	55489544.19	6006937.56
Armidale Dumaresq Council	4235	24611	5.811334	56372288.10	6623485.37
The Council of the Municipality of Ashfield	8	40018	5002.25	56326691.92	6248511.47
Auburn Council	32	64209	2006.531	56318056.99	6252811.05
Ballina Shire Council	484	39953	82.54752	56554531.31	6806290.42
Balranald Shire Council	21699	2730	0.125812	54734757.03	6164111.33
Bankstown City Council	77	177000	2298.701	56320515.48	6246228.39
Bathurst Regional Council	3820	37001	9.686126	55737652.31	6300358.92
The Council of the Shire of Baulkham Hills	401	161068	401.6658	56313143.65	6262799.59
Bega Valley Shire Council	6280	32431	5.164172	55754085.51	5937494.66
Bellingen Shire Council	1602	12758	7.963795	56490221.37	6631080.26
Berrigan Shire Council	2067	8289	4.01016	55392522.77	6053397.24
Blacktown City Council	240	283458	1181.075	56306319.45	6261359.58
Bland Shire Council	8560	6530	0.76285	55562644.59	6238218.91
Blayney Shire Council	1525	6773	4.441311	55709424.31	6287420.05
Blue Mountains City Council	1432	76511	53.42947	56245749.74	6253862.08
Bogan Shire Council	14611	3105	0.212511	55543782.77	6465771.25
Bombala Council	3944	2534	0.642495	55699752.38	5912814.16
Boorowa Council	2579	2495	0.967429	55657688.30	6187875.51
The Council of the City of Botany Bay	22	37074	1685.182	56333335.12	6242502.82
Bourke Shire Council	41679	3906	0.093716	55397636.12	6670893.85
Brewarrina Shire Council	19188	2168	0.112987	55486395.53	6685470.28
Broken Hill City Council	170	20203	118.8412	54544093.91	6463992.40
Burwood Council	7	31158	4451.143	56324615.79	6249816.81
Byron Shire Council	567	30827	54.36861	56559834.71	6831369.76
Cabonne Shire Council	6026	12703	2.108032	55674404.14	6336963.95
Camden Council	201	51367	255.5572	56287315.55	6229394.42
Campbelltown City Council	312	150216	481.4615	56298132.52	6228207.01
City of Canada Bay Council	20	67261	3363.05	56325763.09	6251416.42
Canterbury City Council	34	134126	3944.882	56325862.77	6247335.13
Carrathool Shire Council	18940	3274	0.172862	55355775.11	6191578.81
Central Darling Shire Council	53511	2406	0.044963	54769399.77	6473160.35
Cessnock City Council	1966	48502	24.6704	56343973.20	6366566.04
Clarence Valley Council	10441	49538	4.744565	56493548.28	6715426.76
Cobar Shire Council	45606	5013	0.10992	55389904.90	6514598.89
Coffs Harbour City Council	1175	67442	57.39745	56512367.25	6649928.18
Conargo Shire Council	8751	1782	0.203634	55334636.58	6091907.25
Coolamon Shire Council	2433	4127	1.69626	55518298.98	6147394.41
Cooma-Monaro Shire Council	5229	9792	1.872633	55691001.82	5987716.22
Coonamble Shire Council	9926	4714	0.474914	55632618.98	6574634.93
Cootamundra Shire Council	1524	7623	5.001969	55593913.40	6166502.05
Corowa Shire Council	2324	11058	4.758176	55445129.82	6016107.36
Cowra Shire Council	2810	13185	4.692171	55656488.52	6254887.97
Deniliquin Council	130	8169	62.83846	55314945.72	6066425.01
Dubbo City Council	3428	39263	11.45362	55651130.34	6431257.08
Dungog Shire Council	2251	8440	3.749445	56383121.71	6414117.90
Eurobodalla Shire Council	3422	36389	10.63384	55768055.96	5996288.26
Fairfield City Council	102	187790	1841.078	56310890.27	6250540.33
Forbes Shire Council	4720	9974	2.113136	55593744.79	6305538.82
Gilgandra Shire Council	4836	4660	0.963606	55657591.12	6490315.33
Glen Innes Severn Council	5487	8735	1.591945	56493539.67	6715131.84
Gloucester Shire Council	2952	4917	1.66565	56405622.61	6433457.25
Gosford City Council	940	163304	173.7277	56345843.68	6300152.02
Goulburn Mulwaree Council	3220	27112	8.419876	55748774.02	6150722.05
Great Lakes Council	3376	34695	10.27696	56453974.41	6439552.50
Greater Hume Shire Council	5746	10510	1.829099	55503364.04	6052899.50

LGA	Area	Popn	Density	East	North
Greater Taree City Council	3730	46986	12.59678	56453923.26	6469651.07
Griffith City Council	1640	25140	15.32927	55411983.84	6205380.69
Gundagai Shire Council	2458	3764	1.531326	55600985.37	6119128.65
Gunnedah Shire Council	4994	12074	2.417701	56237879.00	6569537.96
Guyra Shire Council	4395	4460	1.01479	56372299.30	6656404.23
Gwydir Shire Council	9453	5530	0.584999	56346675.62	6661081.79
Harden Shire Council	1869	3773	2.018727	55625658.56	6175448.30
Hawkesbury City Council	2776	63824	22.99135	56297510.80	6279081.23
Hay Shire Council	11328	3534	0.31197	55302157.44	6180238.52
Holroyd City Council	40	91941	2298.525	56314305.51	6254900.40
The Council of the Shire of Hornsby	462	157204	340.2684	56232801.04	6268957.55
The Council of the Municipality of Hunters Hill	6	13928	2321.333	56328936.41	6254666.14
Hurstville City Council	23	76036	3305.913	56324566.30	6239997.91
Inverell Shire Council	8606	15794	1.835231	56317571.48	6704579.90
Jerilderie Shire Council	3375	1871	0.55437	55384546.63	6086686.12
Junee Shire Council	2031	5922	2.915805	55553436.85	6141106.73
Kempsey Shire Council	3380	28742	8.50355	56484928.10	6561452.45
The Council of the Municipality of Kiama	258	20357	78.9031	56303626.32	6161275.82
Kogarah Municipal Council	16	55800	3487.5	56327543.22	6240364.98
Ku-ring-gai Council	86	108697	1263.919	56327503.08	6257453.72
Kyogle Council	3589	9630	2.683199	56500401.96	6833849.82
Lachlan Shire Council	14973	7360	0.491551	56502628.97	6836653.29
Lake Macquarie City Council	644	190320	295.528	56367738.45	6340389.58
Lane Cove Municipal Council	11	32326	2938.727	56330507.85	6256892.92
Leeton Shire Council	1167	12026	10.30506	55445543.27	6176454.44
Leichhardt Municipal Council	11	51142	4649.273	56328303.11	6249671.53
Lismore City Council	1290	43628	33.82016	56524662.57	6813735.01
City of Lithgow Council	4567	20889	4.5739	56235915.36	6291875.75
Liverpool City Council	305	170192	558.0066	56308135.68	6243967.18
Liverpool Plains Shire Council	5086	7852	1.543846	56279679.29	6511748.58
Lockhart Shire Council	2895	3520	1.215889	55474185.61	6102402.15
Maitland City Council	392	61517	156.9311	56364940.41	6377203.27
Manly Council	15	38886	2592.4	56341520.33	6259018.68
Marrickville Council	17	75114	4418.471	56329631.61	6246191.78
Mid-Western Regional Council	8737	22141	2.534165	55742729.20	6391132.55
Moree Plains Shire Council	17928	15936	0.888889	55775586.20	6737355.89
Mosman Municipal Council	9	28363	3151.444	56337511.71	6255408.96
Murray Shire Council	4345	6729	1.548677	55310506.69	6034804.42
Murrumbidgee Shire Council	3505	2620	0.747504	55397800.91	6147943.88
Muswellbrook Shire Council	3406	15149	4.447739	56301140.37	6428420.07
Nambucca Shire Council	1491	18755	12.57881	56489126.94	6604929.58
Narrabri Shire Council	13031	14172	1.08756	55767571.32	6641936.98
Narrandera Shire Council	4117	6582	1.598737	55459446.51	6155188.87
Narromine Shire Council	5264	7033	1.336056	55616795.69	6433183.33
Newcastle City Council	183	146967	803.0984	56386048.62	6356218.24
North Sydney Council	11	60944	5540.364	56334048.79	6254330.01
Oberon Council	3627	5447	1.501792	55764651.74	6267017.81
Orange City Council	285	37791	132.6	55695559.24	6315302.28
Palerang Council	5134	11470	2.234125	55722018.65	6095827.43
Parkes Shire Council	5958	15034	2.52333	55609517.33	6332764.55
Parramatta City Council	61	151860	2489.508	56315214.15	6256346.91
Penrith City Council	405	177955	439.3951	56286574.28	6263835.68
Pittwater Council	91	57354	630.2637	56342396.61	6278492.66
Port Macquarie-Hastings Council	3687	70581	19.14321	56485350.65	6525726.25
Port Stephens Council	858	63579	74.1014	56412603.20	6383659.01
Queanbeyan City Council	172	37169	216.0988	55702816.16	6085362.79

LGA	Area	Popn	Density	East	North
Randwick City Council	36	126034	3500.944	56337477.15	6246080.77
Richmond Valley Council	3051	20913	6.854474	56504462.10	6807112.80
Rockdale City Council	28	95341	3405.036	56326583.34	6242267.69
Ryde City Council	41	99550	2428.049	56322072.18	6258394.28
Shellharbour City Council	147	63124	429.415	56301532.58	6173077.36
Shoalhaven City Council	4568	93615	20.49365	56293860.94	6141111.81
Singleton Shire Council	4896	22270	4.548611	56328047.51	6395563.94
Snowy River Shire Council	6030	7293	1.209453	55664173.05	5973508.95
Strathfield Municipal Council	14	31624	2258.857	56323712.34	6250189.48
Sutherland Shire Council	335	215053	641.9493	56320710.43	6232658.02
Council of the City of Sydney	27	148367	5495.074	56334163.20	6251129.02
Tamworth Regional Council	9713	54522	5.613302	56302632.83	6558389.66
Temora Shire Council	2802	6337	2.261599	55549030.58	6188166.31
Tenterfield Shire Council	7332	6805	0.928123	56404713.87	6787004.06
Tumbarumba Shire Council	4392	3613	0.822632	55591544.06	6040261.05
Tumut Shire Council	4566	11347	2.485107	55611372.19	6092928.90
Tweed Shire Council	1309	80935	61.82964	56538815.48	6866575.02
Upper Hunter Shire Council	8071	13424	1.663239	56250573.49	6441047.34
Upper Lachlan Shire Council	7102	7328	1.031822	55726919.69	6184282.75
Uralla Shire Council	3230	6075	1.880805	56356507.93	6609024.65
Urana Shire Council	3357	1389	0.413762	55433261.52	6090174.20
Wagga Wagga City Council	4824	58055	12.03462	55537764.94	6114710.87
The Council of the Shire of Wakool	7520	4836	0.643085	55263694.54	6071647.71
Walcha Council	6267	3283	0.523855	56365571.22	6570426.76
Walgett Shire Council	22336	8031	0.359554	55607681.74	6678300.78
Warren Shire Council	10760	3273	0.304182	55579346.64	6492580.42
Warringah Council	150	139626	930.84	56340392.61	6264446.51
Warrumbungle Shire Council	12380	10508	0.848788	55686315.27	6541687.52
Waverley Council	9	61611	6845.667	56338382.16	6247691.36
Weddin Shire Council	3410	3848	1.128446	55607364.63	6248881.02
Wellington Council	4113	8599	2.090688	55682594.43	6396278.51
Wentworth Shire Council	26269	7300	0.277894	54584782.43	6225613.24
Willoughby City Council	23	63959	2780.826	56333248.08	6258396.24
Wingecarribee Shire Council	2689	44670	16.61212	56262880.05	6181838.81
Wollondilly Shire Council	2557	41463	16.21549	56279419.07	6214683.49
Wollongong City Council	684	192402	281.2895	56307010.27	6188621.72
Woollahra Municipal Council	12	52747	4395.583	56337665.28	6250073.97
Wyong Shire Council	745	143393	192.4738	56353093.20	6313175.65
Yass Valley Council	3999	12936	3.234809	55674748.32	6142796.17
Young Shire Council	2694	12035	4.467335	55619421.95	6202330.27

APPENDIX G: LINGO® CODE FOR OPTIMAL SENSOR CONFIGURATION

```
model:
sets:
row/1..14/:;
col/1..14/:;


type/1..3/:N;
grid(row,col,type):E,X;

G1(row,col,row,col):D,W;

G2(row,col):Y,cost;

endsets

@for(G1(i1,j1,i2,j2):D(i1,j1,i2,j2)=@if(i1 #eq# i2 #and# j1 #eq#
j2,999,@sqrt((i1-i2)^2+(j1-j2)^2)));


!max=@sum(grid:E*X);

min=@sum(grid:E*(1-X));

@for(type(k):@sum(grid(i,j,k):x(i,j,k))<=N(k));

@for(grid(i,j,k)| j #eq# 14:x(i,j,k)=0);

@for(grid(i,j,k)|i #ge# 7 #and# j #eq# 13 :x(i,j,k)=0);

@for(grid(i,j,k)|i #ge# 9 #and# j #eq# 12 :x(i,j,k)=0);

@for(grid(i,j,k)|i #ge# 10 #and# j #eq# 11 :x(i,j,k)=0);

@for(grid(i,j,k)|i #ge# 12 #and# j #eq# 10 :x(i,j,k)=0);


@for(type(k): x(14,9,k)=0);

@for(G2(i,j):@sum(type(k):x(i,j,k))<=1);

@for(G2(i,j):Y(i,j)=@sum(type(k):x(i,j,k)));

Total_cost=@sum(grid(i,j,k):x(i,j,k)*cost(i,j));

Total_cost<=20000;

@for(row(i1):
@for(col(j1):
@for(row(i2):
@for(col(j2):
```

```
Y(i1,j1)+Y(i2,j2)>=2*W(i1,j1,i2,j2);
Y(i1,j1)+Y(i2,j2)-1<=W(i1,j1,i2,j2)))));


@for(G1:999*(1-W)+D>2);


!@for(grid:@sum(grid(i,j,k) | i #LE# 3 #AND# j #LE# 3:
      x(i,j,k))<=1);

!@for(grid:@sum(grid(i,j,k) |i #GE# 4 #AND# i #LE# 6 #AND# j #LE# 3:
      x(i,j,k))<=1);

!@for(grid:@sum(grid(i,j,k) |i #GE# 7 #AND# i #LE# 9 #AND# j #LE# 3:
      x(i,j,k))<=1);

!@for(grid:@sum(grid(i,j,k) |i #LE# 3 #AND# j #GE# 4 #AND# j #LE# 6:
      x(i,j,k))<=1);

!@for(grid:@sum(grid(i,j,k) |i #GE# 4 #AND# i #LE# 6 #AND# j #GE# 4
#AND# j #LE# 6:
      x(i,j,k))<=1);

!@for(grid:@sum(grid(i,j,k) |i #GE# 7 #AND# i #LE# 9  #AND# j #GE# 4
#AND# j #LE# 6:
      x(i,j,k))<=1);

!@for(grid:@sum(grid(i,j,k) |i #LE# 3 #AND# j #GE# 7 #AND# j #LE# 9:
      x(i,j,k))<=1);

!@for(grid:@sum(grid(i,j,k) |i #GE# 4 #AND# i #LE# 6 #AND# j #GE# 7
#AND# j #LE# 9:
      x(i,j,k))<=1);

!@for(grid:@sum(grid(i,j,k) |i #GE# 7 #AND# i #LE# 9  #AND# j #GE# 7
#AND# j #LE# 9:
      x(i,j,k))<=1);


@for(grid:@BIN(X));
@for(G1:@BIN(W));


data:

E=
0.09     0.136    0.166
0.128    0.188    0.224
0.14     0.212    0.231
0.142    0.217    0.228
0.142    0.214    0.22
0.14     0.208    0.209
0.137    0.202    0.202
0.133    0.2      0.207
0.132    0.204    0.224
0.136    0.213    0.242
0.141    0.22     0.251
0.145    0.221    0.25
```

0.145	0.205	0.242
0.113	0.156	0.195
0.112	0.191	0.232
0.152	0.259	0.304
0.161	0.286	0.307
0.16	0.287	0.296
0.157	0.276	0.277
0.148	0.263	0.256
0.136	0.257	0.253
0.131	0.266	0.282
0.137	0.286	0.328
0.152	0.306	0.359
0.161	0.316	0.37
0.161	0.31	0.361
0.154	0.28	0.337
0.122	0.21	0.266
0.116	0.213	0.275
0.155	0.282	0.348
0.16	0.3	0.338
0.154	0.288	0.312
0.139	0.264	0.277
0.123	0.249	0.251
0.122	0.26	0.266
0.146	0.302	0.339
0.177	0.355	0.429
0.195	0.392	0.474
0.197	0.401	0.486
0.185	0.383	0.47
0.164	0.331	0.421
0.124	0.24	0.321
0.129	0.224	0.297
0.164	0.285	0.358
0.157	0.287	0.326
0.136	0.256	0.283
0.109	0.221	0.238
0.093	0.212	0.217
0.115	0.253	0.262
0.176	0.341	0.386
0.24	0.435	0.524
0.269	0.494	0.583
0.272	0.509	0.595
0.255	0.481	0.578
0.212	0.408	0.512
0.146	0.286	0.376
0.136	0.22	0.297
0.164	0.266	0.336
0.14	0.249	0.283
0.104	0.205	0.23
0.074	0.174	0.189
0.071	0.184	0.189
0.111	0.255	0.268
0.191	0.376	0.432
0.274	0.499	0.6
0.317	0.576	0.664
0.329	0.601	0.677
0.32	0.577	0.671
0.277	0.495	0.61
0.189	0.347	0.441
0.118	0.192	0.276
0.125	0.219	0.289
0.091	0.191	0.223
0.064	0.155	0.177
0.056	0.144	0.156

0.071	0.182	0.187
0.12	0.276	0.295
0.202	0.41	0.474
0.286	0.535	0.641
0.33	0.61	0.698
0.343	0.638	0.709
0.342	0.628	0.722
0.312	0.558	0.694
0.218	0.401	0.51
0.09	0.155	0.241
0.086	0.169	0.234
0.061	0.144	0.171
0.051	0.126	0.143
0.06	0.142	0.149
0.094	0.205	0.209
0.161	0.312	0.332
0.247	0.439	0.5
0.306	0.54	0.637
0.317	0.595	0.678
0.319	0.62	0.69
0.335	0.63	0.725
0.334	0.586	0.746
0.246	0.439	0.572
0.076	0.128	0.204
0.073	0.138	0.189
0.055	0.124	0.141
0.052	0.124	0.136
0.07	0.158	0.164
0.115	0.231	0.24
0.177	0.33	0.36
0.236	0.43	0.493
0.272	0.501	0.589
0.279	0.539	0.622
0.284	0.568	0.642
0.31	0.598	0.698
0.335	0.586	0.768
0.266	0.459	0.62
0.066	0.11	0.17
0.066	0.125	0.16
0.055	0.125	0.135
0.061	0.14	0.153
0.085	0.181	0.193
0.125	0.245	0.265
0.169	0.319	0.365
0.2	0.385	0.456
0.216	0.433	0.517
0.231	0.47	0.557
0.248	0.513	0.598
0.278	0.563	0.673
0.313	0.575	0.777
0.265	0.467	0.654
0.056	0.096	0.14
0.061	0.12	0.142
0.065	0.137	0.142
0.084	0.165	0.178
0.107	0.204	0.221
0.128	0.25	0.278
0.149	0.297	0.352
0.163	0.339	0.41
0.177	0.379	0.454
0.201	0.428	0.512
0.229	0.491	0.582
0.273	0.562	0.679

```
0.322     0.585     0.793
0.277     0.48      0.677
0.044     0.082     0.114
0.061     0.114     0.129
0.08      0.145     0.146
0.102     0.18      0.194
0.114     0.215     0.237
0.125     0.249     0.279
0.142     0.283     0.333
0.157     0.32      0.376
0.175     0.365     0.418
0.208     0.431     0.496
0.257     0.516     0.597
0.315     0.6       0.711
0.354     0.619     0.816
0.29      0.499     0.688
0.035     0.069     0.092
0.05      0.102     0.114
0.068     0.134     0.137
0.093     0.172     0.184
0.112     0.21      0.229
0.125     0.244     0.267
0.14      0.278     0.31
0.159     0.318     0.347
0.185     0.373     0.397
0.23      0.455     0.49
0.291     0.555     0.612
0.349     0.64      0.731
0.369     0.645     0.812
0.289     0.509     0.671
0.033     0.059     0.073
0.045     0.085     0.094
0.054     0.11      0.112
0.073     0.142     0.149
0.1       0.181     0.193
0.125     0.22      0.232
0.145     0.257     0.267
0.164     0.299     0.301
0.192     0.361     0.355
0.243     0.449     0.451
0.306     0.55      0.572
0.358     0.626     0.679
0.372     0.621     0.735
0.291     0.484     0.598
0.029     0.042     0.051
0.039     0.059     0.068
0.042     0.073     0.077
0.05      0.093     0.1
0.07      0.121     0.135
0.093     0.153     0.169
0.11      0.183     0.195
0.13      0.218     0.222
0.168     0.27      0.268
0.222     0.341     0.348
0.28      0.418     0.443
0.322     0.473     0.521
0.332     0.466     0.556
0.26      0.362     0.45
;
N=1,2,2;

cost=
4280.094
```

4280.094
4280.094
4280.094
4280.094
4345.848
5101.356
5252.13
5252.13
5710.304
6228.32
6079.8
6631.55
6243.155
4280.094
4280.094
4280.094
4280.094
4300.452
5076.084
5267.808
5267.808
5267.808
5332.088
5625.14
5601.17
5316.081
5118.956
4280.094
4280.094
4280.094
4276.428
4945.356
5267.808
5267.808
5267.808
5267.808
5267.808
5267.808
5267.808
5267.808
5092.954
4280.094
4285.632
4578.132
5091.372
5267.808
5267.808
5267.808
5267.808
5267.808
5267.808
5267.808
5267.808
5168.849
5070.597
4742.322
5204.862
5267.808
5267.808
5267.808
5267.808
5267.808
5267.808

5267.808
5267.808
5267.808
5267.808
5248.097
5335.754
5253.066
5267.808
5267.808
5267.808
5267.808
5267.808
5267.808
5267.808
5267.808
5267.808
5267.808
5267.808
5193.122
5679.003
5253.066
5267.808
5267.808
5267.808
5267.808
5267.808
5267.808
5267.808
5267.808
5243.472
5822.574
5408.995
5086.388
5553.474
5253.066
5267.808
5267.808
5267.808
5267.808
5267.808
5267.808
5267.808
5267.808
5735.617
6229.293
6403.78
5950.831
6664.698
5253.066
5267.808
5267.808
5267.808
5267.808
5267.808
5267.808
5267.808
5267.808
5899.752
6753.56
6823.062
6462.877
8108.99
5253.066

5267.808
5267.808
5267.808
5267.808
5267.808
5267.808
5267.808
5219.353
5902.19
6986.068
7652.653
7672.996
8323.233
4827.732
5173.74
5267.808
5267.808
5267.808
5267.808
5267.808
5267.808
5357.372
6275.522
7181.484
7945.927
6766.541
8616.6
4280.094
4267.068
4560.426
5070.702
5267.808
5267.808
5267.808
5237.856
5804.34
6404.603
7280.268
7856.119
9231.327
9019.89
4280.094
4280.094
4280.094
4272.918
4619.082
5153.85
5183.193
5223.22
5684.175
6045.716
6912.198
7961.793
9211.644
9576.162
4280.094
4280.094
4280.094
4280.094
4280.094
4292.184
4697.551
5176.938

```
5524.386
6256.168
7381.806
8446.131
8983.71
10161.406
;

enddata

end
```