

---

# Statistics and Biosecurity

**David R. Fox**

Director, Australian Centre for Environmetrics, University of Melbourne

**The “risk” and “results” perspectives of biosurveillance rely on maths, probability and statistics.**

---

Australia is free of the world’s worst animal diseases, such as foot-and-mouth disease and bird flu (avian influenza, H5N1), although the list of potential threats is long (<http://www.daff.gov.au/animal-plant-health/pests-diseases-weeds/animal>). There are good reasons for taking whatever steps are necessary to ensure that this status is maintained.

The 2001 foot-and-mouth disease outbreak in the UK had disastrous consequences with the slaughter of more than 4.2 million animals and substantial economic loss. Four outbreaks of what is thought to be foot-and-mouth disease occurred in Australia in the 19th century; however, there have been no reported outbreaks for over 100 years.

Equine influenza is another highly contagious disease afflicting horses, donkeys, mules and zebras. Shortly after Animal Health Australia released its disease strategy for equine influenza, an equine influenza outbreak was detected in the Sydney area. The disease spread rapidly through northern NSW into Queensland, where it concentrated in the Brisbane region. It wasn’t until Christmas Day 2008 that Australia was officially declared free of equine influenza.

In handing down his findings, the Hon. Ian Callinan AC highlighted shortcomings in the government’s monitoring and surveillance protocols for biosecurity threats. A 2005 AusVet–CSIRO report suggested that Australia review its data requirements and mathematical modelling in order to understand the quantitative aspects of animal disease outbreaks. Similarly, an analysis of the 2001 foot-and-mouth disease outbreak in the UK revealed serious shortcomings in data

collection, processing and analysis activities in the initial stages of the outbreak.

The Department of Agriculture, Fisheries and Forestry (DAFF) defines biosecurity as procedures and policies designed to protect the economy, environment and people’s health from pests and disease. A related concept is bioterrorism, which the US Centers for Disease Control and Prevention defines as “the deliberate release of viruses, bacteria or other germs (agents) used to cause illness or death in people, animals, or plants”.

A characteristic linking both bioterrorism and biosecurity is the “unknown unknowns”\* – that is, we often don’t know what (or who) it is we’re looking for. In any event, the process of “looking” (i.e. biosurveillance) provides an early warning capability for detecting threats, whether they be intentional or unintentional.

Thus biosurveillance has both retrospective and prospective components. The retrospective aspect is concerned with activities that detect, monitor and analyse patterns of disease outbreaks – in other words, after the event. Prospective biosurveillance, on the other hand, is concerned with identifying and assessing the risk of a disease outbreak or a bioterrorist attack before it has occurred.

In either case it is evident that the tools of mathematics, probability and statistics have a critical role to play in both modes of biosurveillance.

## Role of Mathematics and Statistics

### Monitoring in Time

Early work on developing statistical tools for

biosurveillance for the most part represented a reworking or adaptation of conventional statistical methodologies. While these methods are certainly applicable, recognition is growing that the data and processes underpinning modern biosecurity and biosurveillance deviate from the contexts in which they were originally developed.

Traditional statistical tools struggle with the nuances of biosecurity data, which invariably exhibit “curses” such as data paucity, non-normality and over-dispersion (i.e. variation in the data that is much greater than is predicted by the models and/or theory). Other issues such as an inability to deal with data from multiple sources, and a focus on natural/physical processes rather than “choice processes”, are also cited as reasons for the failure of traditional methods of monitoring and analysis.

The peculiarities of biosurveillance systems demand “new” statistical approaches to both data acquisition and analysis.

Techniques that have been successfully applied to the analysis of syndromic and climatic data are candidates for biosurveillance. Syndromic surveillance is underpinned by a belief that signals of an emerging “syndrome”, such as a flu outbreak, can be identified by an analysis of multiple time-series of ancillary variables such as absenteeism records and sales of non-prescription cold and flu medications together with an analysis of spatial clustering of outbreaks.

A major barrier at present is the difficulty in “proving” that any of these new systems have made a difference or even do what they’re meant to. For example, no syndromic surveillance system has provided early warning

## SA Scientists Solve Seagrass Mystery

Release of a major study into the mysterious decline of Adelaide’s coastal seagrasses is a testament to the collaboration of more than 60 researchers from across Australia according to CSIRO, which coordinated the study.

Adelaide Coastal Waters Study Director, Professor David Fox, says that six research teams from a range of local research organisations have pieced together the complex story behind the loss of more than 5000 hectares of shallow sub-tidal seagrass along Adelaide’s metropolitan coast since the mid-1930s and 1940s.

Professor Fox says that the final report released by the South Australian government is the culmination of a major research undertaking that included the release of 20 technical reports investigating the causes of water quality decline and seagrass loss in Gulf St Vincent, adjacent to Adelaide’s shoreline.

The work, initiated in 2001, provides an integrated understanding of the Adelaide coastal system from which to guide future management actions, Professor Fox says.

The study found that many years of near-continuous inputs of nutrient-rich, turbid and coloured water, and wastewater, have resulted in significant changes to and degradation of Adelaide’s coastal marine environment.

“This study has focused on the loss of seagrass, mainly *Amphibolis* and *Posidonia*, seabed instability, and water quality degradation,” he says.

“Seagrass meadows are primary producers at the bottom of the food chain, and they provide natural habitat for many species of fish, crustaceans and other marine animals.



The River Torrens discharging stormwater to Adelaide’s coastal waters on 25 October 2005. Image credit: Simon Bryars

“Taking the seagrasses out of the system causes a ‘domino effect’, where the seafloor becomes less stable and hence promotes a further loss of seagrass.

The study found that water quality improvement plans over the past 10 years, coupled with reduced volumes discharged to the sea, have made a positive difference.

However, Professor Fox says that it may take at least 20 years and up to 100 years for seagrasses to regrow, and large-scale recovery of seagrass meadows requires continued, lasting reductions in coastal inputs and a replanting effort.

Source: CSIRO

of bioterrorism, and no large-scale bioterrorist attack has occurred since existing systems were instituted.

While the use of syndromic surveillance for counter-terrorism (see, for example, <http://www.bt.cdc.gov/surveillance/ears>) is a recent development, similar systems have been used for some time now to detect outbreaks, patterns and trends in diseases and epidemics (see, for example, <http://www.satscan.org>). These techniques do not appear to have had any appreciable uptake in Australia or elsewhere around the world in quarantine inspection and biosecurity, although control charting has been recommended for detecting temporal clusters in veterinary monitoring programs

As part of a larger research project undertaken for the Australian Centre of Excellence for Risk Analysis (ACERA) at the University of Melbourne, I developed a Bayesian framework – where observations are used to update the probability that a hypothesis may be true – for biosurveillance using control charting methods. This allowed expert opinion and/or prior belief about the monitored process to be incorporated as well as providing other enhancements. For example, “ignorance” about a new or previously undetected threat is readily accommodated, and the intrinsic updating of prior information means that these methods are evolutionary, learning and adaptive. Although Bayesian methods have not been widely used in biosecurity/biosurveillance applications, a number of papers have recently appeared that indicate a growing awareness of the potential utility of this statistical paradigm.

### Monitoring in Space

Spatial surveillance is a key component of monitoring programs that 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. 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 spatio-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.

Foreign organisms of concern include Siam weed (*Chromolaena odorata*), papaya fruit fly (*Bactrocera papayae*), red imported fire ant (*Solenopsis invicta*), branched broomrape (*Orobancha ramosa*) and kochia (*Bassia scoparia*).

With respect to animal diseases, a number of potential and serious risks exist including avian influenza (or bird flu), bovine spongiform encephalopathy (“mad cow” disease), foot-and-mouth disease, equine influenza, 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).

A recent review of infectious disease outbreaks noted a number of interesting phenomena in the spatial dynamics of disease propagation in human and animal populations. Examples of these included “spatial waves of infection” and the tendency of disease incidence to occur in spatial clusters.

The phenomenon of epidemic travelling waves is not new, with historical examples provided by the European plague in the Middle Ages, the influenza pandemic in the early 20th century and the spread of cholera in Asia and Eastern Europe during the 1960s.

A number of spatial and spatio-temporal modelling tools for biosurveillance have used conventional statistical modelling approaches. For example, the BioSense program run by the US Centers for Disease Control (<http://www.cdc.gov/BioSense/>) uses small area regression and testing (SMART) to enhance early detection and situational awareness of possible biologic terrorism attacks. However, the method in BioSense only uses spatial

**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 also compromises statistical power ...**

information to bin data into separate time series, and is thus not strictly a spatial model.

### **Space–Time Predictions: Cellular Automata Models**

Mathematical models describing population dynamics usually use either differential or difference equations depending on whether “time” is treated as a continuous or discrete variable. An alternative to this approach are cellular automata methods, having discrete time increments and a matrix representation of a geographical network. A set of rules governs the evolution of the automata such that the state of an element at each time step is expressed in terms of its own state and those of its neighbours at earlier time steps.

Cellular automata methods enjoy a number of advantages over conventional differential–difference equation approaches, such as considerably faster computational speeds and the ease with which certain epidemiological features can be incorporated, as well as local and seasonal effects. Cellular automata models have been used to model a range of problems associated with pathogen and disease spread, including rabies in fox populations and foot-and-mouth disease in feral pigs in Queensland.

## **Biosurveillance Monitoring Design – Optimal Resource Allocation**

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. A number of commentators have long advocated strategies based on avoidance rather than eradication, noting 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. Previous attempts at “optimisation” used economic tools that did not have any spatial or temporal representation.

The use of mathematical programming

techniques to optimise network design problems is not new. Applications of mathematical programming techniques to the optimisation of sparse sensor networks have been associated with air quality monitoring, water supply security and computer network integrity. There is only limited evidence that mathematical optimisation methods have been used to help design and “optimise” monitoring networks for biosecurity surveillance.

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 also compromises statistical power – that is, the ability to identify disease outbreaks, quarantine threats or (bio)security violations when they have occurred. It has been claimed that surveillance geoinformatics of spatial and spatiotemporal hotspot detection and prioritisation 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 five metres in diameter – that is, of the order of a single pixel of information generated by present-day satellites.

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 11 September 2001 terrorist attacks the US 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”.

## **The Future**

While the opportunities to use statistical science for biosurveillance are great, so too are the challenges, not least of which is an ability to

demonstrate the effectiveness of a monitoring program. One of the biggest conundrums for both scientists and managers working in this domain is the problem of zero data. Data streams generated by counting processes, such as the number of disease outbreaks or number of anthrax attacks per unit time, invariably result in sequences that are all zeroes. Statisticians make a distinction between structural zeroes and sampling zeroes. Structural zeroes arise because an event cannot happen (e.g. male pregnancy). Sampling zeroes are an artefact of the incompleteness of the sampling process and arise not because the phenomenon didn't (or couldn't) exist but because we simply failed to observe it.

In the biosecurity context it is difficult (if not impossible) to know if the mostly zeroes in our data are sampling or structural. The critics of syndromic surveillance exploit this anomaly to argue that the millions of dollars that the US government spends on this type of early-warning monitoring for a bioterrorist attack are a waste of money because no such attack has occurred since 2001.

Interestingly, the supporters of syndromic surveillance appeal to the same set of results as "proof" that the system works because would-be attackers are deterred by the surveillance and the inferred risk of getting caught. This debate is not dissimilar from the Y2K computer issue that consumed us all back in 1999.

While mathematics and statistics cannot resolve all the debates and difficulties, this discipline certainly has a role to play in assisting managers, politicians and scientists to make informed decisions about what to monitor, where to monitor and when to monitor.

---

This article has been adapted from material and references in Report 0605 of the Australian Centre of Excellence for Risk Analysis (ACERA). In preparing this article, the author acknowledges the financial and other support provided by the Department of Agriculture, Fisheries and Forestry (DAFF), the University of Melbourne, Australian Mathematical Sciences Institute (AMSI) and Australian Research Centre for Urban Ecology (ARCUE). The views expressed in this article are the author's and do not necessarily reflect those of the aforementioned organisations.

---

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

## Repatriating Australia's Antarctic Waste

Years of scientific expeditions to the Antarctic region have left their mark in the form of tonnes of historic waste, ranging from rusty railway tracks to old tins of milk powder and even dead huskies. Under the Madrid Protocol - the protocol to the Antarctic Treaty on environmental protection - this waste must be cleaned up and returned to its host countries.

Biosecurity Australia, the Australian Quarantine and Inspection Services (AQIS) and the Australian Antarctic Division of the Department of the Environment, Water, Heritage and the Arts are collaborating to develop measures to enable the repatriation of Australia's historic waste, including its adherent soil. Waste that is currently generated by Australia is either incinerated or returned to Australia for recycling or disposal.

---

Source: AQIS



Drums and contaminated soil at Wilkes Station, Wilkes. Photograph by Dixon, Grant, Australian Antarctic Division. © Commonwealth of Australia