Veterinaria Italiana, **43** (3), 595-604

Simulating disease spread within a geographic information system environment

Samuel Beckett & M. Graeme Garner

Summary

Simulation modelling is a tool that can be used to investigate the effectiveness and efficiency of exotic disease control, eradication and surveillance strategies. The Australian Government Department of Agriculture, Fisheries and Forestry (DAFF) has been involved with disease simulation modelling for more than 10 years. Although the focus has been on foot and mouth disease, models are now being developed for avian influenza, classical swine fever and other diseases. Recent models are spatially explicit, and incorporate a range of animal species and production types. The models also encompass a range of disease transmission pathways, including farm-to-farm animal movements, movements through saleyards, windborne spread, spread by feral animals and the less well-defined phenomenon of local spread. The DAFF spatial models are unique in that they are developed within the environment of a geographic information system (GIS) – MapBasic®/MapInfo®. This simplifies the spatial elements of their code and improves their ability to handle spatial data layers. Such layers vary, but may include the following: farm locations or boundaries; masks identifying grazing; cropping and non-agricultural land; water bodies and waterways; population centres, administrative boundaries and roadways; vegetation and other land cover masks; and, where relevant, elevation. The GIS environment also provides immediate access to sophisticated maps and tabular outputs.

Keywords

Australia, Geographic information system, MapBasic, MapInfo, Modelling, Simulation.

Simulazione della diffusione delle malattie all'interno di un sistema informativo geografico

Riassunto

Il modello di simulazione è uno strumento che può essere utilizzato per verificare l'efficacia e l'efficienza delle strategie di controllo, eradicazione e sorveglianza delle malattie esotiche. Il Dipartimento dell'Agricoltura, Pesca e Foreste del governo australiano (DAFF) utilizza i modelli di simulazione delle malattie da più di dieci anni. Sebbene l'attenzione si sia concentrata perlopiù sullo studio dell'afta, si stanno sviluppando altri modelli per l'influenza aviaria, per la peste suina classica ed altre patologie. I recenti modelli sono dal punto di vista spaziale molto precisi e comprendono una vasta gamma di specie animali e tipologie di allevamento. I modelli comprendono anche una gamma di possibili vie di trasmissione delle malattie, inclusi gli spostamenti degli animali da una azienda all'altra, movimenti di compravendita, trasmissione tramite il vento, per mezzo di animali selvatici o inselvatichiti e altri meno ben definiti fenomeni di trasmissione localizzata. I modelli spaziali DAFF sono unici poiché sviluppati all'interno di un sistema informativo geografico (GIS)-MapBasic®/MapInfo®. Ciò semplifica gli elementi spaziali del loro codice rendendoli più facilmente utilizzabili sui "data layers" spaziali. Questi layers variano, ma possono includere l'ubicazione delle aziende o i loro confini, maschere che identificano pascoli, coltivazioni e terre non coltivate, corsi d'acqua e bacini, centri abitati, confini amministrativi e strade, vegetazione e altre zone territoriali e, ove rilevante, l'altitudine. L'ambiente GIS fornisce inoltre un accesso

Department of Agriculture, Fisheries and Forestry, GPO Box 858, Canberra, ACT 2601, Australia
Beckett@Broadleaf.com.au

immediato a mappe molto sofisticate e relative tabelle.

Parole chiave
Australia, MapBasic, MapInfo, Modelli, Simulazione, Sistema informativo geografico.

Introduction

Preparedness for an incursion of an exotic animal disease is of key importance to government, industry, producers and the general public. Australia's exotic disease preparedness is based on emergency management principles (7). These include the development of surveillance, monitoring and early warning systems, the compilation and maintenance of Australia's veterinary emergency plans (AUSVETPLAN) (1) and the conduct of training and awareness programmes.

A good understanding of the likely behaviour of exotic diseases under Australian conditions is a necessary component of effective preparedness and response planning. Recent experience with outbreaks of foot and mouth disease (FMD), for example, in previously free countries, such as Japan, the Republic of Korea, the United Kingdom, France and the Netherlands, has highlighted the importance of well-considered response strategies. In the absence of contemporary Australian experience with such diseases, disease modelling is a tool that can be used pre-emptively to investigate the likely spread of disease under different outbreak scenarios and the effectiveness and cost-efficiency of eradication strategies. The increasing recognition of spatially relevant factors that are likely to affect the character and extent of spread, and the effectiveness of spatially-targeted strategies, such as emergency ring vaccination or contiguous slaughter, mean that models that take into account spatial relationships are becoming increasingly important (5).

The Australian Government Department of Agriculture, Fisheries and Forestry (DAFF) has been involved with disease modelling for over 10 years (2, 3, 4, 5). The objectives of this work have been to identify geographic regions, sub-populations and production systems that might be at greater risk, to evaluate the effectiveness and efficiency of different control, eradication and surveillance strategies, to underpin economic impact studies and to provide realistic scenarios for preparedness or training exercises. This philosophy concurs with conclusions of Taylor, who states in his seminal *Review of the use of models in informing disease control policy* (8) that the most appropriate use of models is as tools in 'peacetime' to aid retrospective analysis of real epidemics. Hypothetical scenarios can then be modelled to develop insights into the relative merits of different strategies in different situations.

DAFF is developing models that operate on a range of scales, including the farm, regional and national levels. Of these, the regional model, AusSpread, has undergone the longest and most intensive period of development and is currently the cornerstone of the DAFF modelling initiative. In this context, a 'region' denotes a part of Australia delimited by natural or geopolitical boundaries and characterised by reasonably homogenous animal production industries and systems. For a continent such as Australia, with diverse environmental and production systems, regions represent the most appropriate scale on which to assess control, eradication and surveillance strategies.

The Australian AusSpread model is unique amongst spatial simulation models in that it is developed within the environment of a geographic information system (GIS) – specifically, MapBasic®/MapInfo® (MapInfo Corporation, Troy, New York). MapBasic® is a simple, intuitive and 'complete' development environment with more than 300 statements and functions and a syntax and program structure similar to Borland Pascal (Borland Software Corporation, Scotts Valley, California) or Microsoft© Visual Basic (Microsoft Corporation, Redmond, Washington). MapBasic® applications are run in MapInfo®, and, thus, take advantages of all of the MapInfo® data management and geographic capabilities. MapBasic® applications can also access a large number of GIS functions and statements, which greatly simplifies many aspects of the spatial disease spread code, and can utilise the mapping and reporting facilities provided by MapInfo®. On balance, these features of the language mean that extremely sophisticated spatial simulation models can be developed relatively quickly and easily.

In this paper, we examine the spatial determinants that will commonly be included in a simulation model, and explain how this process can be facilitated by development within a GIS environment. We illustrate these principles using the AusSpread regional model for FMD. The paper concludes with information on future directions for DAFF spatial simulation modelling.

Spatial determinants of disease spread and control

Two groups of spatial determinants are incorporated into most modern disease simulation models, namely:

- the spatial determinants of disease epidemiology
- the spatial determinants of disease control, eradication and surveillance.

Both are important in determining the characteristics of a simulated outbreak, as well as the cost and efficiency of simulated response strategies. The two groups of determinants are not, however, mutually exclusive, as many spatial determinants of disease epidemiology – such as the distance between farms – will also be relevant to disease response. For the purpose of discussion, the following are considered:

- farm location and management
- spatial characteristics of wild (native or feral) animal populations
- location of relevant infrastructure and administrative boundaries
- spatial heterogeneity in terrain, environment and climate
- spatial determinants of the cost and logistics of response strategies.

Farm location and management

The emphasis of most DAFF modelling is on animal or zoonotic disease in production animal populations. Consequently, substantial effort is invested in developing robust spatial farm datasets. In some cases, farm locations or farm boundary files are available through governments, industry and other sources. In other cases, farm locations are synthesised from agricultural statistics, topographic data layers and land cover and land use information. Wherever possible, the methods for synthesising farm locations are verified by way of comparative statistics.

However, farm location is only one part of the farm-level spatial information required for disease simulation modelling. Equally important are spatially relevant aspects of farm management; typically, purchasing and selling patterns and other movement of livestock, and any off-farm and onto-farm practices that might be relevant to indirect disease transmission. In this context, the term 'farm' encompasses a range of different enterprises, each of which is implemented separately in most DAFF models and requires separate parameters. Beef cattle feedlots, for example, are a special sort of enterprise and exhibit very different spatial management characteristics from (for example) extensive beef breeding farms or sheep farms. Also relevant are the various animal movement hubs – such as saleyards (auction sites), abattoirs, show grounds and live export ports – each of which has a spatial location and exhibits a range of spatially relevant behaviours. The map in Figure 1 shows the region from which a particular beef cattle feedlot sources its animals. This region has a distinct size and elliptical shape, and a distinct spatial orientation, each of which will be important to include if simulating feedlot purchases.

Another example is given in Figure 2 which shows the disparity in movements of beef cattle within Australia in different seasons. Accurate representations of these movement patterns by season will be essential when modelling a disease of beef cattle at a national scale.

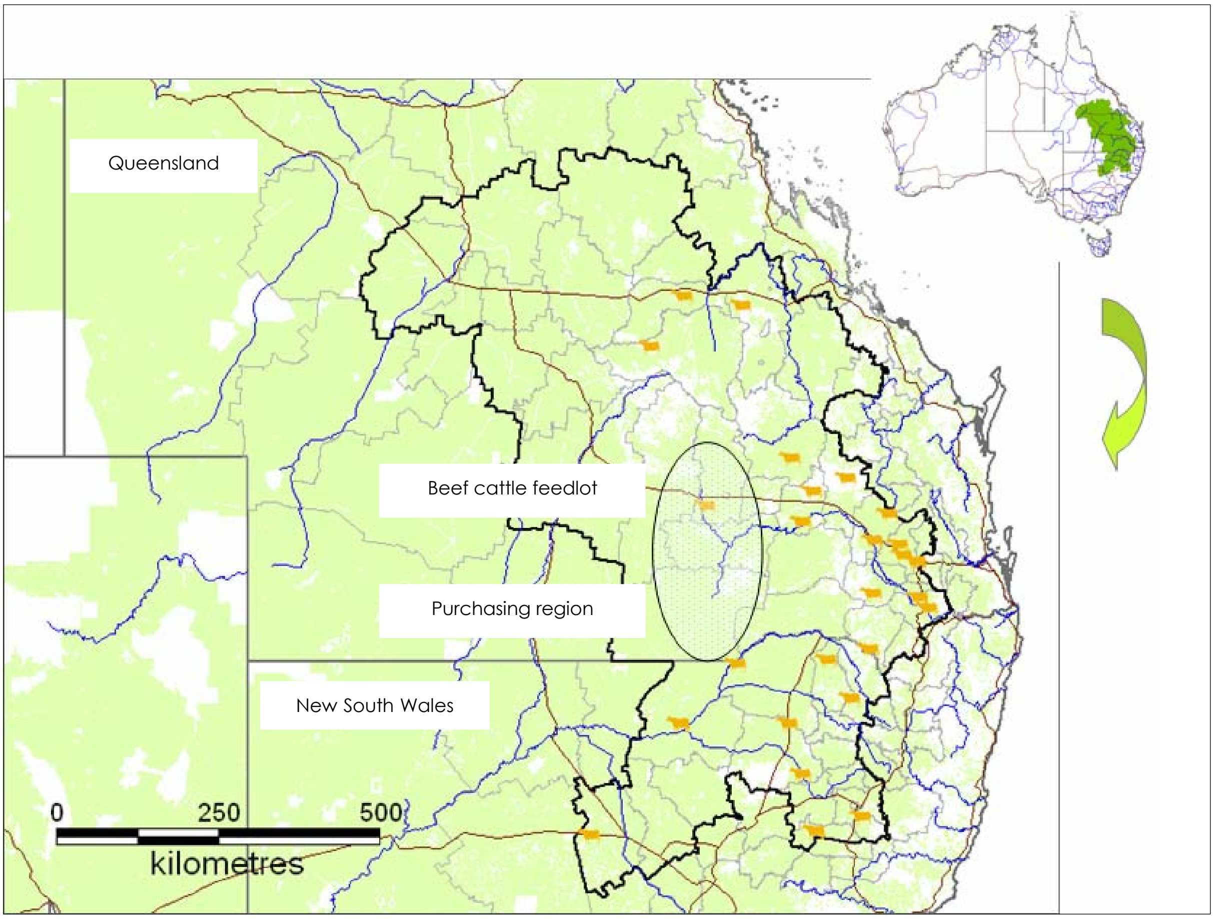


Figure 1
Purchasing region for a beef cattle feedlot in the southern region of the state of Queensland, Australia

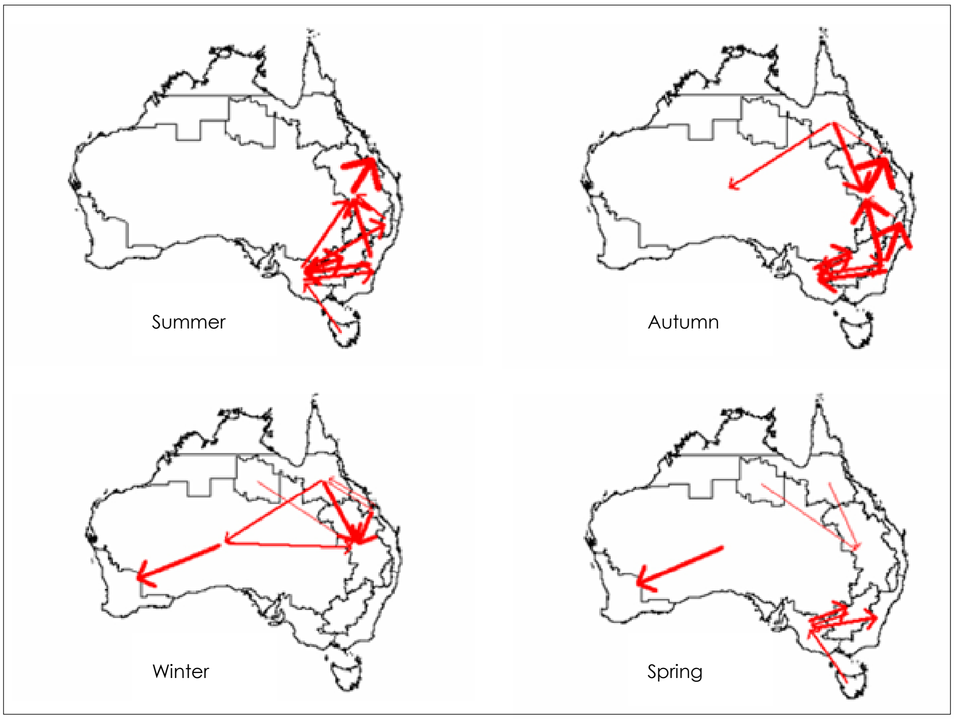


Figure 2
Beef cattle movements within Australia by season

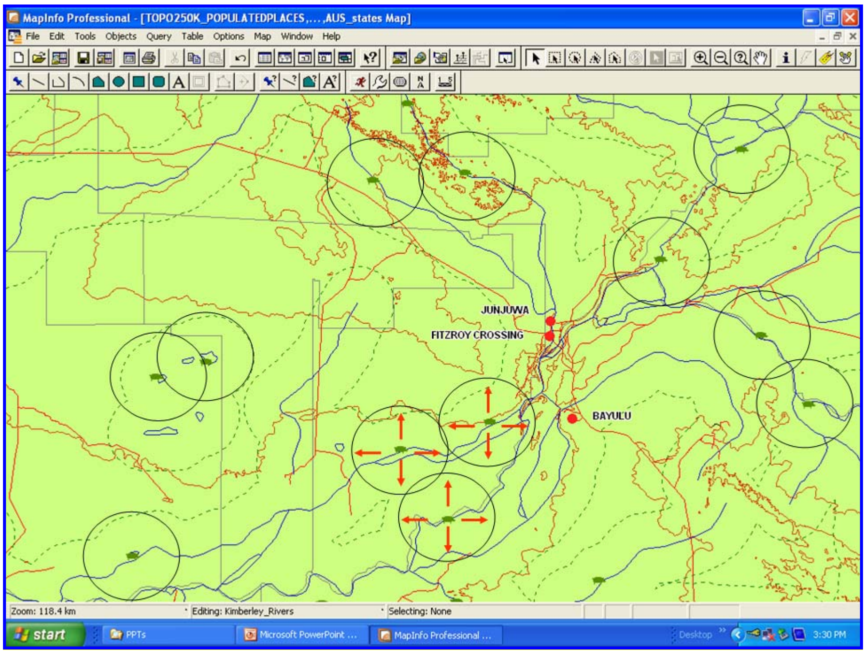
Wild (native or feral) animal populations

Few spatial disease simulation models incorporate both production animals and wildlife – whether native wildlife or feral animals. The AusSpread regional model does this, as illustrated in recent research into the possible role of feral pigs in the spread of FMD within the Kimberley region of Western Australia (6). In this study, it was important to capture the distribution and density of feral pigs in this remote part of Australia, as well as spatially relevant aspects of feral pig behaviour, such as seasonal home ranges and daily movements, breeding behaviour, migration patterns and interactions with livestock and other wildlife. Farm boundaries and cattle management and marketing practices in the Kimberley were also required.

A screen shot from a simulation using the AusSpread Kimberley model is given in Figure 3. In this model, feral pig herds were moved daily within a daily range, interacting with livestock if livestock are present. At then end of each day, the herds were moved to a new daily range, but stayed within a seasonally defined 'home range'. This was a confluent surface common to all feral pig herds in the population and was dictated by the availability of surface water and vegetation. Artificial stock dams and areas of improved pasture are also considered. As the season changes from wet to dry in this tropical region of northern Australia, the feral pig home ranges contract towards remaining surface water (whether rivers or stock dams) and it is here that there is the greatest potential for interactions with livestock (Fig. 4).

Source: B. Madin, 2004

Figure 4
Feral pigs and beef cattle interacting at a stock dam in the Kimberley region of the state of Western Australia



Zoom: 118.4 km
Editing: Kimberley_Rivers
Selecting: None

Figure 3
AusSpread screen shot showing the daily movements and home ranges of feral pigs in the Kimberley region of the state of Western Australia

This simple example of feral pigs in the Kimberley region of northern Australia demonstrates the importance of spatial determinants to wildlife simulation modelling. The principle is not, however, unique to this scenario. Similar determinants were required when considering the inclusion of other feral species (principally goats and deer) in other regional modelling studies. Likewise, many vector-borne disease agents have wildlife hosts and are associated with epidemics that can be characterised by changes in the spatial distribution or density of these hosts, or their interactions with production animals or humans.

Location of relevant infrastructure and administrative boundaries

Infrastructure and administrative boundaries are included in simulation models as required. Saleyards (auction sites), abattoirs and live export ports have been mentioned and will often be important to the spread or management of animal and zoonotic diseases. Roads and transport networks are less commonly considered, although may be relevant to the passage of, for example, milk tankers or heavily laden stock trucks. Human population centres are clearly relevant to zoonotic disease models and are denoted as areas of non-agricultural land in most animal disease models. Administrative boundaries are generally important to disease response. In Australia, for example, jurisdiction for animal disease response rests with the seven States and Territories and will commonly be configured according to the boundaries of affected local government areas (LGAs).

Spatial heterogeneity in terrain, environment and climate

This group of determinants includes the following:

- elevation or gradient
- rivers, lakes and other water bodies
- deserts; vegetation and land use classifications
- the identification of areas with distinct climates.

Collectively, these factors can be relevant to the behaviour of the 'actors' of the model (principally livestock or wild animals) or to disease agent survival or dispersal. FMD virus (FMDV), for example, is relatively sensitive to extremes of sunlight, heat and low humidity, and will be less likely to survive in the environment in parts of the country where these are the norm. Other diseases, such as Japanese encephalitis, rely on insect vectors and these in turn will be distributed according to climate and the environment. The importance of proximity to surface water to the ecology of feral pigs has already been mentioned and will have ramifications for all diseases for which they are a host.

Spatial determinants of the logistics and efficacy of response strategies

This group encompasses a very broad range of spatially defined factors that can determine the practicability and effectiveness of practices relating to animal tracing, surveillance, slaughter, vaccination, wild animal control, zoning and regionalisation. Some of these, such as the distance between farms, will have other ramifications (above) and are likely to have been considered. Others, such as the logistics of feral animal baiting strategies in remote parts of Australia, will be specific to particular response practices.

Modelling in a geographic information system

From the discussion above it can be seen that a large number of spatial determinants of disease epidemiology, control, eradication and surveillance will be important to most disease simulation modelling studies. These can be accommodated by way of spatial data 'layers', as in the example of farm locations; or by way of spatially explicit coded logic, as in the example of the daily movements and interactions of feral pig herds. By developing within a GIS environment, we have access to both sorts of spatial function – specifically, data layers are encapsulated as MapInfo® (vector) data files; and spatially explicit logic is encapsulated in coded references to GIS statements and functions. Access to GIS map and tabular outputs is an ancillary benefit. The manner in which GIS handles spatial data files will be well known to the readers of this journal, and need not be expanded upon. Programming in a GIS language, and accessing GIS outputs, will be discussed in turn.

Programming in geographic information systems

MapBasic® is considered a 'complete' development language, in that it allows the user to create a graphical user interface (GUI), manage data files, populate arrays and other variables at a global or local level, loop through coded logic to represent events performed during each time step, and display the output or write it to files. Further loops can then be added to accommodate stochastic variability. Collectively, these actions encompass a typical simulation model.

The difference between MapBasic® and non-GIS languages is that it offers two key advantages, namely:

- access to sophisticated queries and logical tests built on spatial functions and statements
- access to spatial 'object' variables (noting that the term 'object' is not used here in the sense of 'object orientated programming paradigm').

These two advantages will be discussed in turn.

Spatial functions and statements commonly used in the AusSpread model include object buffering, and distance, area and bearing calculations. Spatial queries and logical tests are also commonly based on object contiguity, joins or merges. For example, saleyard point locations might be buffered to create drawing or dispersal zones; or watercourses might be buffered to create feral pig home ranges. Buffers thus created can be stored in 'object' variables (below) or tables, and can be used in queries or logical tests. Continuing the examples, farms that lie within the buffered saleyard drawing region can be selected from the broader population of farms as potential sellers and pig herds can be moved each day within the buffered watercourse. Collectively, these complex spatial queries and logical tests can be achieved in several lines of GIS code. Distance and bearing calculations are also commonly used. The closest weather station to an infected pig farm, for example, can be found with a simple series of GIS statements and used subsequently to parameterise a windborne spread routine. Likewise, the farms that lie in the direction of the prevailing wind can be found and flagged as possible exposures to a windborne virus plume. Area calculations are most commonly used for the generation of simulation results. Examples include the area under movement restrictions, the area encompassed by affected LGAs or the total area encompassed by a ring vaccination programme. Contiguity is most obviously relevant to the implementation of a contiguous cull policy, and can be coded in this case within a single statement. Another example is the selection of LGAs contiguous to an affected area and subsequent flagging for movement controls. Object joins and overlaps are also very powerful. It might be important, for example, to determine the overlap between a feral pig herd home range and land grazed by beef cattle. Alternatively, it might be helpful to temporarily merge the spatial objects that delineate the boundaries of affected farms to calculate the total area of affected farmland.

These examples illustrate the range of sophisticated queries and logical tests that can be built on spatial functions and statements. The ease of coding has been mentioned and, to illustrate this, a brief code segment has been inserted below (Fig. 5). In this ostensibly simple example, two concentric ring buffers are created around a saleyard. The first (arbitrarily termed 'near_sales') describes the region from which most sellers will be sourced; while the second (termed 'far_sales') describes the wider region from which the balance will be sourced. The radius of the first buffer (the distance from the saleyard to the buffer rim) is given by the value of a variable termed 'near_dist_sellers'; and the radius of the second (outer) buffer is given by 'max_dist_sellers'. In literal terms, the first select statement (the third line of code) then reads, 'select all (*) fields (columns) from the

```
near_sales=buffer(this_sale_obj, near_dist_sellers, "km")
far_sales=buffer(this_sale_obj, max_dist_sellers, "km")

select * from dataset
     where dataset.obj within near_sales
     into near_farms

select * from dataset
     where dataset.obj within far_sales and
        not(dataset.obj within near_sales)
     into far_farms
```

Figure 5
MapBasic® code segment

dataset table for those farms whose object (in this case, the farm centroid) lies within the near_sales buffer and place these records in a temporary table called near_farms'. The second select statement for the outer concentric buffer is similar, although care is taken not to re-select any farms that lie within the smaller inner buffer. So now we have a temporary table containing those farms in the population that will contribute most animals to the sale and another containing those that will contribute the balance. From here it is only a matter of randomly selecting the required number of sellers from each table.

Collectively, these very simple statements describe quite a complex spatial operation; and one that would require a great deal more code and computation were it not for easy access to the GIS buffer function.

The second key advantage of programming in a GIS environment is that it provides access to spatial 'object variables'. In this context, object variables are memory parcels that can hold points, lines and region objects. The advantage of object variables is that the computer can process them very much faster than tabular vector data. This means that by populating an object variable with the spatial information contained in a table record (for example, a farm boundary object) , and performing functions on this variable rather than on the original table object, a great deal of processing time can be saved.

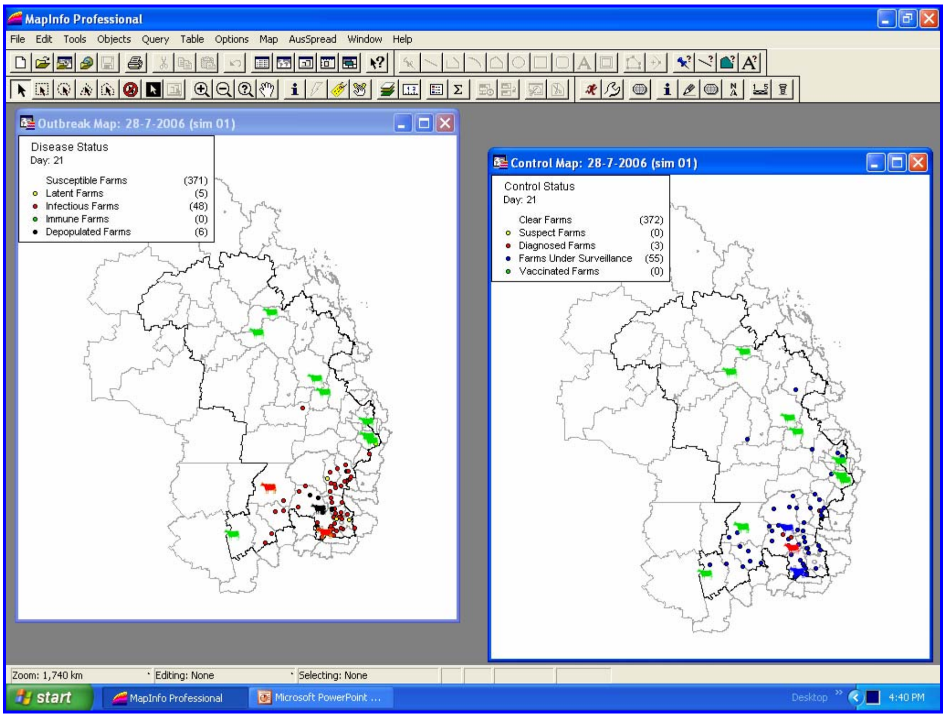
This function is used throughout the code for AusSpread, as well as in many of the smaller DAFF subprograms that have been written in MapBasic® to perform routine data manipulation tasks. An example of the latter is a module that is used to create a synthetic farm dataset from agricultural statistics and land use information. This module places farms at points in a landscape by repeatedly testing landscape 'suitability' at those points – that is, does it lie over a populated area, a river or lake, a desert, another farm, etc.? Each of these queries relates one object to another and, if each is performed on object variables, then the collective operation will take a fraction of the computing time that would otherwise be required for repeated table-based spatial queries. Object variables do, however, have two key constraints, as follows:

- they cannot be mapped and displayed visually
- they are purged (or cleared) on completion of a subroutine (if local variables) or the main program (if global variables).

These constraints limit the application of object variables to processing and calculation (as described).

Accessing geographic information system outputs

The advantages offered by a GIS-based programming language have been described above. A separate feature of MapBasic® is that its compiled programs run within MapInfo® GIS. This means that the programmer does not have to make separate 'calls' to MapInfo® (or another GIS) to perform mapping functions; or code map and other outputs from first principles. Note that mapping in this context can be as simple as the display of tabulated vector data, for example, farm boundaries. In the case of AusSpread, however, complex multi-layered thematic maps are used to illustrate disease outbreaks on a daily basis, or, as required, at the end of the simulation. It can also be helpful to display different sorts of maps concurrently. One might, for example, illustrate the progression of the disease in the population while the other illustrates what is known to controllers (Fig. 5). This form of display has proved to be very useful in disease preparedness exercises. Other standard GIS display options, such as tables and plots (Fig. 6), can also be accessed directly by way of simple coded statements. By adding an



Disease Status
Day: 21

	Susceptible Farms	(371)
○	Latent Farms	(5)
●	Infectious Farms	(48)
○	Immune Farms	(0)
●	Depopulated Farms	(6)

Control Status
Day: 21

	Clear Farms	(372)
○	Suspect Farms	(0)
●	Diagnosed Farms	(3)
●	Farms Under Surveillance	(55)
○	Vaccinated Farms	(0)

Figure 6
AusSpread map output showing disease events and control status

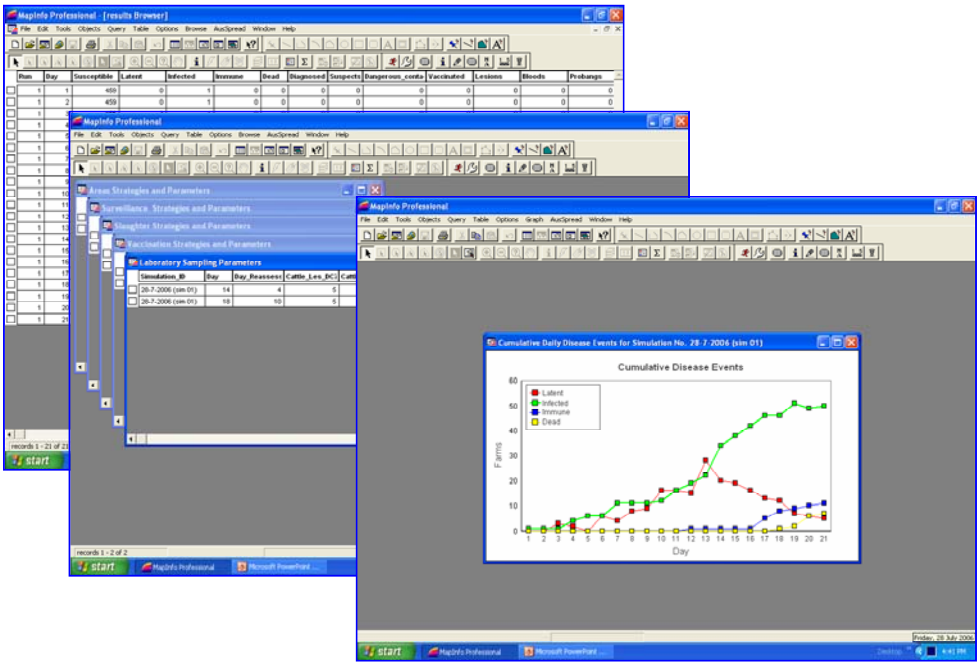


Figure 7
AusSpread tabular and graphic outputs

AusSpread menu to the MapInfo® menu bar, and by linking items on this menu to coded segments that map or otherwise display or export key outputs, the user can generate high-quality outputs as desired and with minimal knowledge of GIS.

Next steps

AusSpread is a model 'shell' that provides a framework for the rapid development of regionally specific spatial simulation models for FMD. Using the shell, DAFF will continue to conduct detailed modelling studies, including computable general equilibrium (CGE) economic modelling studies, for each of 12 identified Australian regional environments. As relevant, these studies will require enhancement of some of the model's components, such as the feral pig transmission module. DAFF is also using the AusSpread model shell to look at other animal and zoonotic diseases, including classical swine fever (hog cholera), avian influenza and pandemic influenza. For the reasons outlined in this paper, each of these modelling studies will continue to be performed using the MapBasic®/MapInfo® development environment.

References

1. Animal Health Australia 2002. Disease strategy: foot-and-mouth disease (Version 3.1). Australian veterinary emergency plan (AUSVETPLAN) Edition 3. Animal Health Australia, Canberra, 68 pp.
2. Garner G. 2004. Using epidemiological modelling to assist FMD preparedness in Australia. *Aust J Emerg Manage*, **19** (3), 9-12.
3. Garner M.G. & Lack M.B. 1995. An evaluation of alternate control strategies for foot-and-mouth disease in Australia: a regional approach. *Prev Vet Med*, **23**, 9-32.
4. Garner M.G. & Lack M.B. 1995. Modelling the potential impact of exotic diseases on regional Australia. *Aust Vet J*, **72**, 81-87.
5. Garner M.G. & Beckett S.D. 2005. Modelling the spread of foot-and-mouth disease in Australia. *Aust Vet J*, **83**, 30-38.
6. Madin B. 2004. Modelling foot-and-mouth disease in the Kimberley region of Western Australia. Unpublished thesis, Faculty of Veterinary Science, University of Sydney, Camden, NSW, Australia
7. Murray G. & McCutcheon S. 1999. Model framework and principles of emergency management. *In* Management of animal health emergencies (J.G. Murray & P.M. Thornber, eds). *Rev Sci Tech*, **18**, 15-20.
8. Taylor N. 2003. Review of the use of models in informing disease control policy development and adjustment. A report for DEFRA. Veterinary Epidemiology and Economics Research Unit, Reading, 98 pp.