Modelling the spread of livestock disease on a national scale: the case for a hybrid approach

R.A. Bradhurst [a], **S.E. Roche** [b], **M.G. Garner** [b], **A.S.M. Sajeev** [a] **and P. Kwan** [a]

[a] *School of Science and Technology, University of New England, Armidale, New South Wales, Australia*
[b] *Animal Health Epidemiology Program, Animal Health Policy Branch, Department of Agriculture,
Canberra, Australian Capital Territory, Australia*
Email: *rbradhur@myune.edu.au*

Abstract: An epidemic of exotic disease in a livestock population can lead to substantial economic losses. For example, the projected cost of a foot-and-mouth disease (FMD) epidemic in Australia is in the billions of dollars. This includes the direct cost of eradicating the disease (e.g., movement restrictions, culling and vaccination), and the impact to export markets from the loss of Australia's FMD-free status.

Epidemics can be difficult to study empirically, particularly if a pathogen is dangerous, rare, or simply not present in a country. In these circumstances a model of disease spread can be a valuable epidemiological tool. When responding to an epidemic, animal health personnel might be restricted to enacting existing policies that leave little scope for the trialing of new control strategies. Computational modelling compensates for the limited opportunities an epidemiologist has to experiment in the field.

Models of disease spread typically employ population-level approaches such as equation-based modelling, or individual-level approaches such as agent-based modelling. Population-level models can be concise and computationally efficient, but they do not isolate individual contributions to an epidemic. The finer granularity of individual-level models can introduce a computational overhead. In the case of a very large-scale model, an individual-level approach can require a highly parallel platform such as a high-performance computing cluster in order to function efficiently.

Epidemics are dynamically shaped by the complex interplay between host, pathogen and the environment. Modelling livestock disease spread on a national scale presents unique challenges due to large populations, varying herd types and farming practices, and regional and geopolitical differences. An alternative to pure population-level and individual-level modelling is a fusion of the two approaches into a hybrid model. This tactic is employed in the Australian Animal Disease Spread (*AADIS*) model, currently under development. The spread of disease within a herd is modelled from the top down by a system of ordinary differential equations. The spread of disease between herds is modelled from the bottom up by a spatially-aware agent-based model. Homogeneity is a reasonable abstraction for a herd of domestic animals and thus intra-herd spread of disease is well suited to equation-based modelling. The national set of herds is however, heterogeneous, making inter-herd spread of disease well suited to agent-based modelling.

AADIS models the transfer of disease from an infectious herd to a susceptible herd by five stochastic spread pathways: direct contact, indirect contact, local spread, airborne transmission and spread through saleyards. Herds can be viewed abstractly as autonomous nodes in a network. Over discrete time steps of one day, the disease spread pathways generate the network topology. Network paths can subsequently be traversed forward to assess the downstream impact of an infected herd, or backward to trace the historical infection route. The network topology thus captures the spatiotemporal history of the simulated epidemic.

AADIS is implemented in Java and employs open-source products such as PostgreSQL, PostGIS and OpenMap. It has an asynchronous object-oriented architecture that takes advantage of the inexpensive parallelism available on a multi-core x64 target.

Keywords: AADIS, epidemiological model, hybrid model, spatiotemporal model, livestock disease spread.
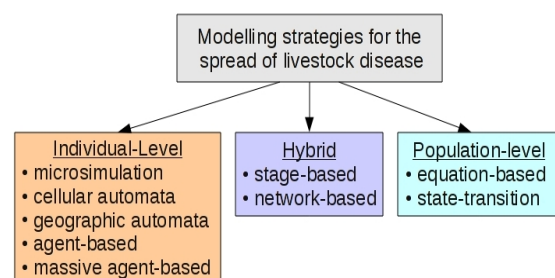
INTRODUCTION

An epidemic of exotic disease in a livestock population can lead to substantial economic losses. An example is foot-and-mouth disease (FMD), a highly contagious disease of cloven-hoofed animals, not present in Australia. The impact to Australia's Gross Domestic Product from a 12-month FMD epidemic is estimated to be $16 billion over ten years (Matthews, 2011). This includes the direct cost of controlling the disease (e.g., culling of infected animals and vaccination), and longer-term impacts such as the suspension of export markets due to the loss of Australia's FMD-free status. In the absence of within-country experience of a disease, computational modelling is an important tool for veterinary epidemiologists and disease managers to study the potential spread and impact of a disease, and evaluate control strategies.

The Australian Animal Disease Spread *(AADIS)* model, currently under development, simulates the spatiotemporal spread of livestock disease across Australia. Whilst the model architecture accommodates any pathogen, the initial test-case disease is FMD. Modelling on a national scale is complex due to large populations, heterogeneous herd types and farming practices, and regional differences in animal health policy. Some examples of the varied influences on the spread of disease in livestock include:

- direct contact between animals (e.g., stock movements between farms, saleyards and abattoirs),

- indirect contact between animals (e.g., via fomites transferred on vehicles, personnel and/or equipment),

- viability of virus in the environment (influenced by relative humidity, temperature, UV radiation, etc.),

- presence of vectors and conditions suitable for the establishment of vector-borne diseases,

- pathogen characteristics (e.g., some viruses are species-specific while others span species),

- climatic and seasonal factors (e.g., stock movement patterns may vary according to the time of year),

- livestock management and market practices (can vary with species, farm type and region),

- detection and reporting of disease (how long an epidemic progresses undetected is influenced by the willingness and timeliness of case reporting by owners, as well as the presence of surveillance programs),

- disease control policies and the effectiveness of control measures (e.g., biosecurity practices, movement restrictions, vaccination, culling, treatment),

- availability of resources to respond to an epidemic (e.g., animal health personnel, equipment, vaccine).

Equation-based models are concise and well-established predictors for systems that are dominated by physical laws (Parunak et al., 1998). Epidemics however, can be dynamically reshaped by irregular factors such as climate, economics, geography, psychology, sociology and jurisdiction-dependent animal health policies. The complex and variable environment in which a disease propagates can be difficult to quantify mathematically. This is reflected in the adoption of individual-level models by veterinary epidemiologists, for example, microsimulations (Morris et al., 2002) and agent-based systems (Ward et al., 2011). Individual-level modelling is a paradigm shift away from top-down predictive mathematical algorithms. A population is viewed from the bottom up as comprising individuals with innate goals, logic and state. Interactions between individuals, and between individuals and the environment generate the model outcome in the form of emergent behavior (Macal and North, 2010). The effect of variation, uncertainty and chance can be incorporated through stochastic techniques such as Monte-Carlo sampling (Garner and Beckett, 2005). Hybrid models combine population-level modelling techniques with individual-level techniques (Bobashev et al., 2007).

Models can be distinguished by how they handle time, space, variability, chance and uncertainty (Taylor, 2003). However, these criteria do not lend themselves to a taxonomy. For example, a stochastic model may or may not be spatial, a deterministic model may view time discretely or continuously, etc. A simple classification scheme is thus proposed whereby models are grouped as population-level, individual-level or a hybridisation of both approaches (Figure 1). A hybrid approach is the cornerstone of the *AADIS* model of disease spread in livestock.

Figure 1. Modelling strategies.

1. POPULATION-LEVEL MODELLING

A population can be dynamically disaggregated into logical, non-spatial compartments according to disease state. A typical set of compartments is Susceptible, Exposed, Infectious and Recovered (SEIR). Mathematical or statistical rules quantify how the population transitions between states over time. The cardinality of the compartments change as an epidemic waxes and wanes. This can also be viewed as population 'flow' between compartments.

Mathematical models of the spread of disease can be traced at least back to Bernoulli in the 18th century (Dietz and Heesterbeek, 2002). Contributions from the early 20th century include systems of ordinary differential equations (ODEs) (Earn, 2008) and the Reed Frost model (Fine, 1977). Equation-based models (EBMs) can provide a concise and computationally efficient representation of the spread of disease. For example, solving an SEIR ODE system yields compartment ratios at any point in time over the solution interval. Unfortunately the implied assumption that model parameters are fixed over the solution interval is not realistic. Garner and Hamilton (2011) note that irregular environmental events dynamically reshape an epidemic. An EBM can become complex, unwieldy and harder to solve as more variables are factored into the mathematical abstraction (Parunak et al., 1998).

The flow of a population between compartments can also be dictated by probabilities derived from historical data or expert opinion. Miller (1976) employs a Markov chain-based state-transition model to predict the spread of FMD across the USA. He notes that variable force of infection and control measures are more easily incorporated into the probabilities governing state transitions, than augmenting and resolving the ODE system of an EBM.

Individuals are not discernible within the compartments of a population-level model. For this reason the technique is best suited to homogeneous populations that are well-mixed; indeed homogeneity is a typical assumption for such models. This is a limitation if the underlying physical population is heterogeneous.

2. INDIVIDUAL-LEVEL MODELLING

D'Souza et al. (2009) note that reducing biological systems to a set of linear models does not necessarily reflect their inherently multi-scale, coupled, non-linear and heterogeneous characteristics. Instead of prescribing how a population behaves from the top-down, individual-level modelling allows population behaviour to emerge from the bottom-up. This is done by simulating the natural behaviour of individuals and then observing them interact with each other, and with the environment. The notion of a population of synergistic individuals is flexible, e.g., buyers and sellers in a market (Heppenstall et al., 2005), bacteria and immune cells in tissue (D'Souza et al., 2009) or the occupants of a city (Colizza and Vespignani (2010). A vivid example of emergent behaviour is the flocking of birds such as starlings (Feder 2007). Individual birds autonomously carry out simple rules such as maintaining a heading/speed that is compatible with immediate neighbours, whilst watching out for environmental obstacles and predators. At a population level, complex and dynamic 3D flocking patterns emerge with no centralised coordination.

Individuals have autonomous state and can be modelled with passive software entities (e.g., data structures) or active (e.g., threaded objects). In an epidemiological context, susceptible individuals coexist with infectious individuals within the confines of an environment. The interpretation of an individual is model specific. For example, one model might simulate the spread of disease between animals, while another might simulate the spread of disease between farms. An individual-level model is not constrained by the assumption that the population is homogeneous and well-mixed. Individual-level model types include microsimulations, cellular automata, geographic automata and agent-based models.

2.1. Microsimulations

Early examples of the individual-level approach can be found in socio-economic 'micro-analytic' models from the 1960s (Boman and Holm, 2004). These models contrasted with macroeconomic models of the time by focusing on the role of individuals in an economy, rather than the traditional top-down economic sector-based approach. Individuals in early microsimulations were represented by passive software entities and the models were generally non-spatial. Microsimulations have since evolved to include spatial-awareness, stochasticity and/or intelligent individuals. An example of a sophisticated microsimulation is *AusSpread* which stochastically models the spatiotemporal spread of disease between individual farms using Markov chain Monte Carlo methods (MCMC) (Garner and Beckett, 2005). Model parameters that are subject to natural variability, chance and/or uncertainty, are randomised with a statistical bias by sampling from probability distributions. For a given scenario, a stochastic model typically produces a unique outcome for

each simulation run, however trends can emerge out of the probabilistic mechanisms by re-running a simulation scenario many times.

2.2. Cellular and geographic automata

Whereas microsimulations can be spatial or non-spatial, and have passive or active individuals, agent-based systems tend to be spatial with active individuals. Agents are autonomous with respect to goals, logic and state and can sometimes evolve their behaviour over time. Cellular automata (CA) are agent-based systems where the environment is spatially discretised into a lattice of cells. Agents may be static and confined to a home cell, or mobile across the environment. Time is also discretised in that systems of rules are triggered *en masse* at periodic intervals. CA can model micro-environments such as bacteria within tissue (D'Souza et al., 2009), right up to macro-environments such as the spread of virus across a population (Doran and Laffan, 2005). Geographic automata (GA) extend the CA lattice to a geographic environment. For example, contact between agents might be a function of geographical distance rather than based on the Moore neighbourhood of cells (Ward et al., 2011).

2.3. Agent-based models

An agent-based model (ABM) is a continuation along the evolutionary path of individual-level models. Whereas a CA environment is constrained to a lattice of cells, and a GA environment is constrained by geography, an ABM environment is generalised. Agents may interact over a variety of topologies, including Euclidean space, contact networks, GIS and aspatial 'soup' (Macal and North, 2010). Agent-based modelling is well suited to heterogeneous populations, as each type of individual can have a customised agent type. A massive agent-based model (MABM) is an expansion of an ABM to handle large populations. The distinction between an ABM and a MABM is somewhat arbitrary. However, for illustrative purposes, let's suppose that an ABM has up to thousands, or perhaps hundreds of thousands of agents, while a MABM may run into the millions, or even billions of agents. Large numbers of agents can impose a considerable computational burden, and MABMs rarely employ 'off the shelf' ABM frameworks. MABMs typically involve custom software implementations (Parker and Epstein, 2011), and highly parallel platforms such as high-performance computing clusters (Carley et al., 2006) or general purpose computing on graphics processing units (D'Souza et al., 2009).

3. HYBRID MODELLING

A hybrid model blends population-level and individual-level modelling approaches. This provides a compromise between the computational economies of say an EBM, and the individual-level granularity of say an ABM. Colizza and Vespignani (2010) note that hybrid models scale well for systems with large populations, such as pandemics. Two interesting hybrid approaches are stage-based and network-based.

3.1. Stage-based hybrid model

A stage-based hybrid model is modal in that simulation is conducted at any given time by either a population-level model (such as an EBM), *or* an individual-level model (such as an ABM). An ABM can capture the subtle interactions between individuals that shape the early stages of an epidemic, however, as the epidemic progresses, increasing levels of agent activity may impede model performance. A stage-based approach involves a dynamic switchover from the ABM to an EBM at a threshold number of cases. Although the granularity of modelling decreases, overall model performance is maintained by the computationally efficient EBM. As the epidemic wanes and the number of cases falls back below a threshold value, it is possible to switch from the EBM back to the ABM. Stage-based hybrid modelling is employed by Bobashev et al. (2007) in a global study of human influenza. When the cases threshold is reached for a particular city, the ABM is halted and a snapshot of agent states used as initial conditions for an EBM. The EBM progresses the epidemic within the city efficiently albeit without the granularity of the individual.

3.2. Network-based hybrid model

A metapopulation is a collection of spatially distinct groups of the same species (i.e., sub-populations), that can interact to some degree. A network-based model abstracts a sub-population as a node in a network. Disease spread between sub-populations is abstracted as some form of network topology. For example, the movement of individuals (and potentially disease) between sub-populations defines a contact-network (or mobility network). Balcan et al. (2009) and Colizza and Vespignani (2010) model the spread of influenza-like diseases between human sub-populations via a contact network of global scale. A network-based model

is termed hybrid when it employs a blend of population-level *and* individual-level modelling approaches. Vincenot and Moriya (2011) for example, model the spread of disease within a sub-population with a system dynamics-based EBM, and the spread of disease between sub-populations with a contact-based ABM. *AADIS* extends the Vincenot and Moriya (2011) network-based hybrid model from single-species to multi-species, and from a contact-based topology to a multi-layer topology.

4.   THE AADIS HYBRID APPROACH

4.1.   Modelling unit of interest

AADIS disaggregates the Australian livestock population into herds based on species and farming practices. Herd types include: intensive beef cattle, extensive beef cattle, dairy cattle, sheep, large piggeries, small piggeries, feedlots and small-holders. This results in a national population of approximately 250,000 herds. Farms can have one or more herds. The choice of herd as the unit of interest rather than the farm captures heterogeneity within multi-species farms. For example, farms that manage sheep and cattle can be modelled as two independent herds with distinct disease dynamics and animal management practices. The more granular choice of an animal as the modelling unit of interest implies a population size in the order of 100,000,000. Apart from the computational burden, this choice does not make epidemiological sense in a livestock model of national scale. As a counter-example, consider the spread of disease between humans. Each person has their own contact network of family, friends, classmates, colleagues, team mates, fellow commuters, etc. (Colizza and Vespignani, 2010). An infectious individual can set off a chain reaction of infection rippling across their personal contact network. Livestock, on the other hand, are typically managed as herds and effectively share a single contact network whilst on a farm.

A herd has type, size, location, and species/pathogen-specific characteristics of how disease spreads. *AADIS* views a herd as homogeneous with respect to both species and farming practices, and well-mixed. This implies that any one member of a herd has the same likelihood of contracting a disease as any other member. This follows from the assumption that animals in a herd share a single contact network whilst on a property. Further, the rate of animals joining a herd via births and transfers in, is assumed equal to the rate of animals leaving a herd via deaths and transfers out. The homogeneous, well-mixed and constant herd size assumptions, although not quantifiable, are reasonable concessions when modelling on a large-scale.

4.2.   Within-herd spread

The assumption that a herd is homogeneous and well-mixed suits the strengths of population-level modelling. *AADIS* models within-herd spread of disease with a non-spatial, deterministic, pathogen-specific and species-specific EBM. Each herd instance has its own EBM customised for the herd type and size (Figure 2). The EBM is implemented as an SEIR ODE system that is solved numerically via a 4th-order Runge Kutta method (Cash and Karp, 1990). When a herd becomes infected, the EBM is started and generates compartment ratios over the solution interval. In the absence of asynchronous events triggering change within the herd (e.g., a *cull* message), the solution remains in place for the duration of the simulation run. The EBM yields herd prevalence and the appearance of clinical signs as a function of time.
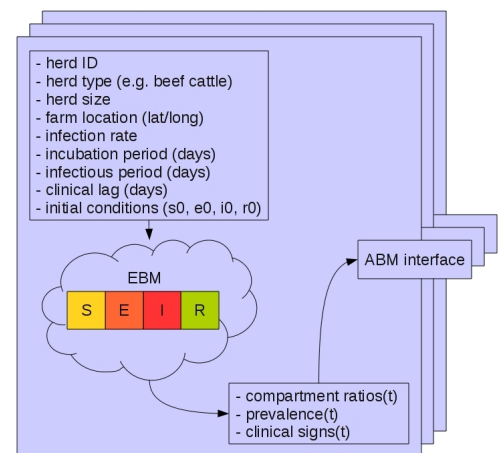

Figure 2. Herd instances

4.3.   Between-herd spread

AADIS models the transfer of disease between herds over five independent and concurrent 'spread pathways':

• direct contact – from the movement of animals between premises,

• indirect contact – for example via fomites transferred by vehicles, equipment or personnel,

• local spread – proximity-based spread of disease, e.g., contact over a fence shared by adjoining premises,

• airborne transmission – virus excreted by animals in aerosol form that remains viable in the air,

• saleyard spread – from animal movements in and out of saleyards.

Each pathway has a set of rules as to whether disease will transfer between an infectious herd and a susceptible herd on any given day. The direct, indirect and saleyard pathways are 'target' oriented in that for each infected herd a suitable candidate herd is sought. For example, for disease to spread over the direct pathway a livestock movement must emanate from an infected premises, the consignment must include infected animals, and a susceptible herd of a particular type must exist at a suitable location. The local and airborne pathways are 'sweep' oriented in that for each infected herd, *all* eligible candidates in a search area are identified. For example, for disease to spread over the airborne pathway an infected herd must be shedding virus, environmental conditions such as wind direction, humidity and temperature must be conducive to virus remaining viable, and susceptible herds must exist within a sector emanating from the infected herd. Decisions in all five pathways are made stochastically by sampling from probability distribution functions. The spread of disease between
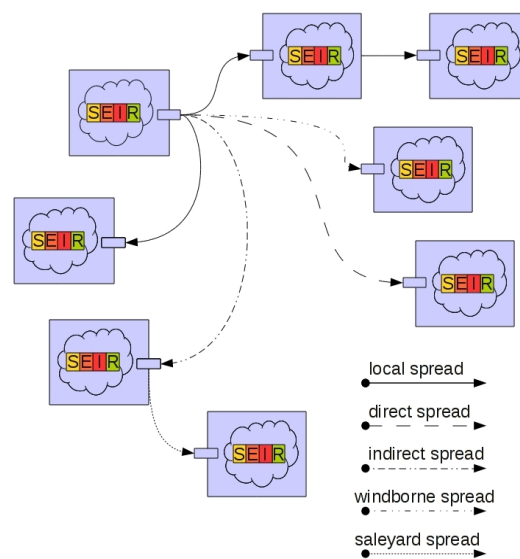

Figure 3. Connectivity between herds

heterogeneous entities is well suited to agent-based modelling as agents themselves need not be homogeneous. Each herd is an agent in the ABM. The disease spread pathways define the ABM environment. The national set of herds can be viewed abstractly as nodes in a network (Dube et al., 2011). If all of the rules for a pathway fire then a relationship forms between the herds (Figure 3). Each spread pathway provides a means for nodes to 'connect' and thus contributes a layer of network topology. Over time as a simulation runs, the spread pathways generate the topology of the herd network. The topology takes the form of a directed acyclic graph, until such time as recovered herds lose their immunity. Network paths can subsequently be traversed forward to determine the downstream impact of an infected herd, and backward to trace the historical infection route. The network topology thus captures the spatiotemporal history of the simulated epidemic.

5. CONCLUSIONS

Models of disease spread are important tools when studying the epidemiology and control of livestock diseases. In countries such as Australia, livestock herds are spread across vast and diverse regions, and produced under different management and marketing systems leading to potentially quite complex models. Efficiency is thus an important consideration when implementing models on a national scale. A population-level modelling approach is concise and computationally efficient but the assumption that the population is homogeneous and well-mixed has limitations from a real world perspective. An individual-level modelling approach accommodates heterogeneity in a population but can be a computationally challenging to implement on a large scale. We present a case for modelling the spread of livestock disease on a national scale by fusing population-level and individual-level modelling paradigms into a hybrid model. *AADIS* models the spread of disease within a herd from the top down with a deterministic EBM, and the spread of disease between herds from the bottom up with a spatially-aware stochastic ABM.

REFERENCES

Balcan D., Colizza V., Goncalves B., Hu J., Ramasco J.J. and Vespignani A. (2009). Multiscale mobility networks and the spatial spreading of infectious diseases. *Proceedings of the National Academy of Sciences of the USA*, 106(51), 21484-21489.

Bobashev G.V., Goedecke D.M., Yu F. and Epstein J.M. (2007). A hybrid epidemic model: Combining the advantages of agent-based and equation-based approaches. *Proceedings of the 2007 Winter Simulation Conference*, pp. 1532-1537.

Boman M. and Holm E. (2004). *System Approaches and Their Application.* Springer, Netherlands (2004), 95-118.

Carley K.M., Fridsma D.B., Casman E., Yahja A., Altman N., Chen L., Kaminsky B. and Nave D. (2006). BioWar: Scalable agent-based model of bioattacks. *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans*, 36(2), 252-265.

Cash J.R. and Karp A.R. (1990). A variable order Runge-Kutta method for initial value problems with rapidly varying right-hand sides. *ACM Transactions on Mathematical Software*, 16(3), 201-222.

Colizza V. and Vespgnani A. (2010). The Flu Fighters. *Physics World*, 23(2), 26–30.

Dietz K. and Heesterbeek J.A.P. (2002). Daniel Bernoulli's epidemiological model revisited. *Mathematical Biosciences*, 180(2002), 1-21.

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. *Preventative Veterinary Medicine*, 70(1-2), 133-152.

D'Souza R.M., Marino S. and Kirschner D. (2009). Data-Parallel Algorithms for Agent-Based Model Simulation of Tuberculosis On Graphics Processing Units. *Proceedings of the 2009 Agent-Directed Simulation Symposium,* The Society for Modeling and Simulation International, March 2009.

Dube C., Ribble C., Kelton D. and McNab B. (2011). Introduction to network analysis and its implications for animal disease modelling. *Review Scientific et Technique*, 30(2), 425-436.

Earn D.J.D. (2008). A light introduction to modelling recurrent epidemics. *Mathematical Epidemiology*, eds F. Brauer, P. Van den Driessche and J. Wu., Springer-Verlag Berlin, pp. 3-17.

Feder T. (2007). Statistical physics is for the birds. *Physics Today*, 60(10), 28-30.

Fine P.E.M. (1977). A commentary on the mechanical analogue to the Reed-Frost epidemic model. *American Journal of Epidemiology*, 106(2), 87-100.

Garner M.G. and Beckett S.D. (2005). Modelling the spread of foot-and-mouth disease in Australia. *Australian Veterinary Journal*, 83(12), 758-766.

Garner M.G. and Hamilton S.A. (2011). Principles of Epidemiological Modelling. *Review Scientific et Technique*, 30(2), 407-416.

Heppenstall A.J., Evans A.J. and Birkin M.H. (2005). A hybrid multi-agent/spatial interaction model system for petrol price setting. *Transactions in GIS*, 9(1), 35-51.

Macal C.M. and North M.J. (2010). Tutorial on agent-based modeling and simulation. *Journal of Simulation*, 4(3), 151-162.

Matthews K. (2011). A review of Australia's preparedness for the threat of foot-and-mouth disease'. Australian Government Department of Agriculture, Fisheries and Forestry, Canberra.

Miller W.M. (1976). A state-transition model of epidemic foot-and-mouth disease. New Techniques in Veterinary Epidemiology and Economics, 1976, 56-72.

Morris R.S., Sanson R.L., Stern M.W., Stevenson M. and Wilesmith J.W. (2002). Decision-support tools for foot and mouth disease control. *Review Scientific et Technique*, 21(3), 557-567.

Parker J. and Epstein J.M. (2011). A distributed platform for global-scale agent-based models of disease transmission. *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, 22(1).

Parunak H.V.D., Savit R. and Riolo R.L. (1998). Agent-based modeling vs. equation-based modeling: a case study and users' guide. *Proceedings of Multi-agent systems and Agent-based Simulation* (MABS'98), 10-25.

Taylor N. (2003). Review of the use of models in informing disease control policy development and adjustment. School of Agriculture, Policy and Development, The University of Reading.

Vincenot C.E. and Moriya K. (2011). Impact of the topology of metapopulations on the resurgence of epidemics rendered by a new multiscale hybrid modeling approach. *Ecological Informatics*, 6, 177-186.

Ward M.P., Laffan S.W. & Highfield L.D. (2011). Disease spread models in wild and feral animal populations: application of artificial life models, *Review Scientific et Technique*, 30(2), 437–446.