Application of a global sensitivity analysis technique to the New Zealand Standard Model of foot-and-mouth disease

A dissertation presented
in partial fulfilment of the requirements
for the degree of Master of Veterinary Studies
at Massey University

L. Katharine A. Owen
2012

"I have been, and remain, entirely committed to the idea that modeling is the essence of science and the habitat of all epistemology."

- Robert Rosen, Essays on Life Itself

Abstract

"All models are wrong, but some are useful" George Edward Pelham Box (Box, 1987)

Unfortunately, often people don't realise that models are wrong, or alternatively do not appreciate that they are not supposed to be 'right'. A model, whether a map, a regression model or a stochastic model that creates data, is merely a representation of reality. It is a tool and is bound by assumptions, the quality of the data used to create it as well as the code itself. This dissertation attempts to provide a simple guide to help non-modellers to understand what a model is and the uncertainties that exist in the modelling process, the methods available for sensitivity analysis and describes the role of this analysis in corroborating a model.

In New Zealand we are fortunate not to have had an outbreak of foot-and-mouth disease (FMD), but recognise the devastating consequences of an outbreak and the need to develop appropriate controls for eradication of the virus. Given the complexity of the interactions over time and space between the virus, the unique environment of New Zealand, its livestock industries and the controls available, it is unlikely that these issues could be explored adequately and with sufficient rigour without the use of modelling.

This thesis is divided into three main sections. In the first section a review of the literature related to simulation modelling and sensitivity analysis is provided. In New Zealand we are fortunate to have a well corroborated platform for disease modelling – InterSpread Plus. This has allowed development of a model scenario for exploring FMD spread within a set of InterSpread Plus parameters termed the New Zealand Standard model (NZSM). The NZSM is a complex epidemiological model and has been, at the time of writing, under-utilised. An element of the under-utility is due to the complexity and computational demands that are inherent to an epidemiological model. The second section of this thesis provides a description of the New Zealand standard model in an attempt to explain its logical framework.

In the third section a sensitivity analysis of the NZSM is carried out in an effort to identify those settings in the model that had the greatest influence on the predicted number of infected premises in a simulated outbreak of FMD in New Zealand.

The outcomes of the sensitivity analysis have increased the understanding of the NZSM model itself and have provided initial insight into ways the model may be improved and refined. If heeded, this can result in an increased ability to interpret and communicate model outputs. This, in turn, will increase confidence that the NZSM parameter set provides an appropriate indication of the way FMD might spread if it were introduced into the farm animal population in New Zealand, and therefore better preparation for any future FMD outbreak.

Contents

Introduction

Foot-and-mouth disease virus outbreaks in countries that are ‘normally’ free are rare events with devastating consequences. This makes modelling of outbreaks in the country of interest a valuable tool in planning and assessing the difficult questions of disease control. However, the outcome of an outbreak and the effect of available interventions can be complex with many interdependent variables and considerable uncertainty.

In New Zealand a parameter set termed ‘the New Zealand Standard Model’ has been developed to run within the InterSpread Plus platform. This parameter set incorporates knowledge of the current predominant foot-and-mouth virus strain, Pan-Asia O, New Zealand farm data and current knowledge of patterns of movements between farms and/or saleyards, and the likely impacts of controls that are utilised to eradicate the disease.

To make a model useful the challenge is to understand and communicate the questions that need to be answered by the decision makers to the modellers, but also to communicate and understand the meaning and limitations of the model outputs. This results in confidence in, but not misuse of, the model. Sensitivity analysis is a tool to aid this understanding.

This dissertation is presented in five main chapters and attempts to describe the value of a global sensitivity analysis technique (that is, a technique in which multiple inputs are varied simultaneously) in improving the understanding of the New Zealand Standard Model for foot-and-mouth disease.

Recognising the need to apply an appropriate sensitivity analysis method to a complex stochastic model, Chapter 1 attempts to explain the features of such a model, and outlines the origins of uncertainty in such a model. The intent is to provide a guide for an informed non-modeller. Chapter 2 then discusses and reviews available sensitivity analysis techniques.

Chapter 3 is a description of the New Zealand Standard model based on the unpublished work of Robert Sanson and Graham Mackereth in 2006. This provides an understanding of the model that is used in New Zealand, introducing the uncertainties in data and the complexity of the model.

The application of a sensitivity analysis technique to the New Zealand standard model is described in Chapter 4. It presents a screening approach using partial rank correlation coefficients. This analysis should be seen as a first step in a process to refine the standard model and an aid to increasing confidence in its use as a tool to aid decision making.

Chapter 5 provides a discussion of the future application of sensitivity analysis in improving the New Zealand Standard model and its fitness for purpose to compare disease control scenarios for foot-and-mouth disease in New Zealand.

List of Figures

List of Tables

Definitions of terms

Term	Definition
Chance	The unknown and unpredictable element in happenings that seems to have no assignable cause.
Dynamic	Account is taken of time varying interactions among variables.
Factor	General term for model input(s) and parameters.
Global technique	A technique in which all factors of interest are allowed to vary simultaneously over their entire range.
Local technique	A technique in which the input space is only explored in a small area so the assumption of linearity is plausible.
Non-linear	Involving relationships such as equations which are not of the first degree
Qualitative	Based on qualities rather than measurable quantities.
Quantitative	Capable of being measured. Quantitative techniques generally provide information on the amount of variance in the output that is explained by each factor.
Uncertainty	The lack of certainty, A state of having limited knowledge where it is impossible to exactly describe existing state or future outcome, more than one possible outcome.
Variability	The quality, state or degree of being variable.

Chapter 1: The characteristics of a simulation model

What is a simulation model?

A model is a representation, rather than a replication, of an existing or new system of interest. Models aim to represent real world systems using a quantitative framework, allowing low cost and quick experimentation with system characteristics (Guitian & Pfeiffer, 2006). There is a balance between complexity and oversimplification to ensure a model provides information on the system under investigation and incorporates its salient features, without being so complex that it cannot be understood.

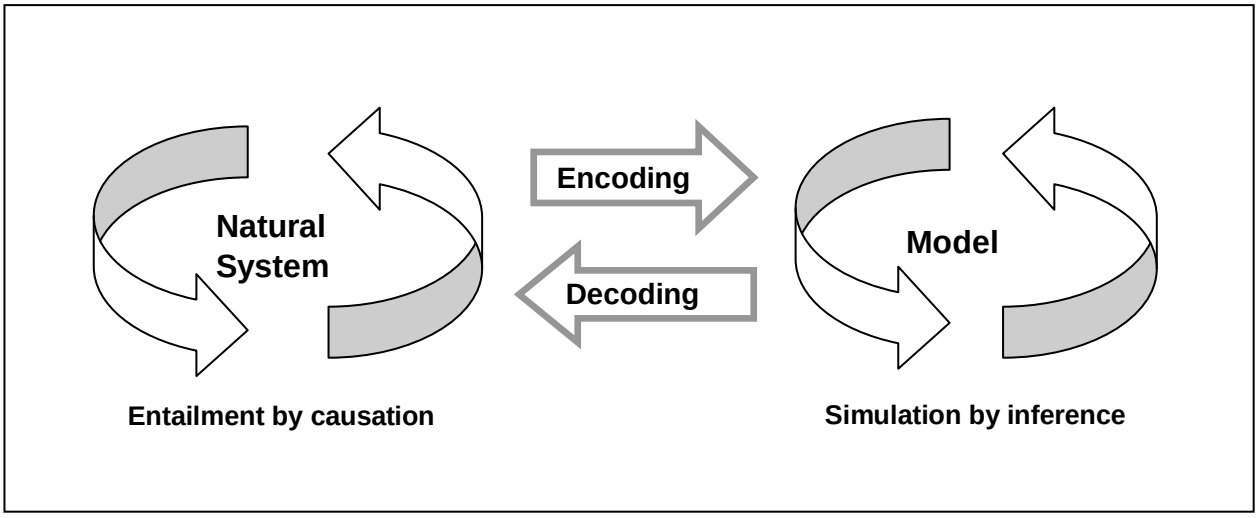


Figure 1: Diagram of the relationship between a natural system and a model (after Rosen, 1991).

Rosen (1991) describes the relationship between a model and a natural system: *encoding* is essentially the describing a system by assigning labels such as numbers to the natural phenomena, *decoding* is the process of taking the outputs of a model and using them to predict the behaviour of the system. If a model is successful, when the data created from the natural system are run through the inferential processes of the model predictions are generated, which will agree (when verified) with the causal entailment of actual phenomena in the natural system.

It is important to note that encoding and decoding involve generalisation and add subjectivity, for example decisions around:

- selection of which processes should be modelled;
- equations used to represent and link the processes; and
- units of spatial and temporal aggregation (De Pauw et. al., 2008).

Simulation is the operation of a model of a system (Figure 1). Broadly it may be thought of as a tool to evaluate the performance of a system under different specifications and time periods and therefore predict the effect of changes to the system. A major benefit of modelling is that it can be used to infer the values of parameters or behaviour that cannot be evaluated experimentally or are too expensive to implement. However, it is important to remember that models represent the natural system but are never true to reality (Kao, 2002). Models encourage discovery of solutions by a process of trial and error: They do not provide <<MS: Yes?>> answers, but should change answers.

Models can be considered according to how they treat the following three key elements (Taylor, 2003):

• variability, chance and uncertainty;

• time; and

• space.

Variability, chance and uncertainty

The Council for Regulatory Environmental Modeling in the USA (2009) summarise the differences between deterministic and stochastic modelling. "*Deterministic modelling employs classical statistics and is useful for capturing the most likely or 'average' conditions observed in a given system. This is known as the 'frequentist' approach to summarising model input data. Frequentist statistics rely on measures of central tendency (median, mode, mean values) and represent uncertainty as the deviation from these measures.*'

Although a deterministic model may be used to simulate extreme or pre-defined values, these models commonly assign average or most likely values to all parameters and model the average

or ‘expected’ outcome (Green and Medley, 2002). These models produce a single result for each set of input parameters, although sensitivity analysis can be used to test the effects of a range of input values.

In contrast, a stochastic model contains and produces stochastic uncertainty (random variability) (Green and Medley, 2002). This is done by adding an appropriate probability distribution to the average value used in the input or by sampling from a distribution or table of values. The output of a stochastic model is a distribution that represents a range of effects, which is a probability interval between which the true value is likely to lie (Green and Medley, 2002).

Time

Models can also vary in the way they treat time — either as continuous or in discrete intervals, e.g. days, weeks, etc. Continuous time models use differential equations to represent changes over time (Cowled and Garner, 2008). This approach is mathematically and computationally elegant and widely used because it enables processes to be represented by relatively simple sets of equations. However, the approach may not adequately reflect processes that are not continuous in time. An alternate approach is to treat time as discrete convenient units with changes taking place being simulated repeatedly over intervals as required (Taylor, 2003). Although this approach is computationally more complex, it is quite simple to incorporate processes that vary over time (e.g. seasonal effects).

Space

Models may be spatial or non-spatial. Non-spatial models assume that spatial proximity has no influence on disease transmission probability (French and White, 2004). In these models, only the size (or density) of the population (or sub-populations) is important and not the spatial relationship between individuals. Spatial models incorporate spatial attributes such as individual locations and contact structures between individuals to determine who is at risk of infection. Like time, space can be treated as either a continuous or discrete variable (French and White, 2004).

Generally, the most appropriate type of model to use in a given situation will depend on the purpose of the modelling exercise. For example, whilst deterministic models may be useful for understanding basic infection dynamics, they are of limited use as a predictive tool since any one epidemic is unique and unlikely to follow an ‘average’ pattern (Green and Medley, 2002). Whilst

stochastic models cannot predict what will happen, they permit a distribution of likely outcomes under different scenarios to be studied (Cowled and Garner, 2008).

Environmental modelling has many similarities to disease simulation modelling; in particular the problems to be addressed are complex and involve technological, environmental, economic and societal aspects. There is ambiguity in defining the problem, uncertainty in the data and uncertainty in the model itself (Refsgaard, et. al. 2007). These ambiguities and uncertainties must be considered when model outputs are used for decision making. If the error is not understood one cannot be certain that the model or system represents the true state. It is likely that better communication and understanding of the uncertainties of disease model outputs will increase a decision maker's confidence in the use of disease models and will prevent inappropriate implementation, such as the controversial use of models for prediction with unreliable data in the epidemic of FMD that occurred in the United Kingdom in 2001 (Garner et. al. (2007).

The dilemma of complexity *vs* simplicity

Models should be parsimonious – as simple as possible, but no simpler. If over complex, the resulting model will be compromised by the uncertainties of the input variables. Even if a model is parsimonious, it will have elements of incomplete knowledge, inaccurate measurement, ambiguity and contradiction (Marino et. al., 2008). Mathematical models may be complex, but epidemiologic models are inherently so.

Complex models, generally those with more parameters, that are stochastic and spatially explicit, aim to provide detailed effects and more quantified outcomes. Complex models are based on fundamental processes and, as such, are more conceptually straightforward but offer no compact mathematical description and so are computationally demanding. They also present challenges for communication and understanding of assumptions and constraints. As a model becomes more complex it needs more accurate data that may require the involvement of multi-disciplinary teams. Green and Medley (2002) provide an example:

'Again the level of complexity of spatial modelling is an art and introducing factors such as physical geography is not a small task. Modern weather forecasting has essentially the same problem, although somewhat more complex because air can move in three dimensions. However, weather forecasters have a considerable advantage because of their access to vast quantities of accurate data from satellite, weather stations, radar and the like.

Likewise, the modellers needed detailed and accurate data on the spatial distribution of farms and livestock, and the current epidemic situation. Unfortunately,[in the UK 2001 foot-and-mouth disease outbreak] these data were not always reliable or known. This created a serious problem for predictions from these models. It is like asking a forecaster what the weather will be tomorrow, without providing information on both the weather today and the position in the world to be forecast.'

The value of a model depends on the question to be answered, the data available and the communication of the constraints, assumptions and outcomes of the simulations. Both types of models have the ability to include necessary spatial structure and heterogeneity. Validation procedures and sensitivity analysis help to provide confidence in the robustness of a model, and comparison of model predictions with subsequent field experience will also be valuable (Morris et. al., 2001). The paper by Keeling and Danon (2009) advocates the use of models to quantify uncertainty of predictions and the need for good quality of data and appropriate level of detail.

Characteristics of disease models

In general, 'epidemiological' models are composed of a series of mathematical and logical relationships that are analysed by numerical methods rather than the analytical methods of pure mathematical models (Bagni et. al., 2002). With this approach, output variables are predicted as a consequence of decision rules (inference) rather than through algebraic solutions. These types of models are often more flexible and able to incorporate a high degree of detail and can be useful as tactical models. They also tend to have a more transparent logical framework that is easier for non-mathematicians to follow (Taylor, 2003) but tend to be more complex to construct and time consuming to run compared to mathematical models (Carson, 2005). Epidemiological models are generally stochastic and dynamic. They may be complex with many factors, and so are likely to be non-linear and have multiple interactions.

In disease modelling, mathematical and epidemiological models are different approaches to explore questions about a given disease scenario. Both approaches are legitimate and useful when used appropriately. To explore the two approaches this discussion will use descriptions of the attributes of both types of model in the context of their use during the epidemic of foot-and-mouth disease in the United Kingdom in 2001.

There are many forms of mathematical models, but in the modelling of infectious diseases the term is usually applied to a state transition model utilising differential equations to describe the transmission process from state to state. The most influential process in such a mathematical disease simulation model is infection, the measurement of which focuses on the reproductive number (R_0), the product of the contact rate and the probability of a successful transmission. Complex mathematical models are usually analysed by means of numerical solutions to differential equations, or from numerous stochastic simulations (Britton and Lindenstrand, 2009).

InterSpread Plus (Stevenson et al., 2012), the North American Animal Disease Spread Model, NAADSM (Harvey et al., 2007) and Auspread (Garner & Beckett, 2005) can be classified as epidemiological models. Epidemiological models utilise mathematical principles but also sample on statistical distributions to represent biological processes (the host/pathogen/environment interactions of the transmission processes and effects of control measures) and their inherent variability (Morris et al., 2001). For example, mathematical models of infectious disease often focus on infection rather than the clinical expression of the infection (disease) whereas epidemiological models may attempt to capture these additional complexities (Green and Medley, 2002). An epidemiological model attempts to represent the interactions that produce the characteristic behaviour of the disease (Morris et al., 2001). Such models are, of course, still representations with associated assumptions and constraints, and both types of model can be useful if well constructed, used in the correct context with good communication between those involved (Kao, 2002).

Both mathematical and epidemiological models describe the dynamics of infection through time and/or space. Both types of models define the basic unit of interest. Although state transition models have been used and applied to FMD situations at the individual animal level (Woolhouse et. al., 1996; Bates et al.2003; Tsutsui et al., 2003; Heuer et al., 2007; Carpenter et al., 2011) in FMD epidemic models the unit of interest is the generally the farm rather than the individual animal. A set of mutually exclusive states (susceptible, latent, infectious, recovered/immune) will then be defined that an individual unit can have, and the transition between these states is determined by mathematical functions in the case of a mathematical model or by the outcomes of the sampling processes in epidemiological modelling. The values used in both processes are derived in similar ways: from published information, from experimental infection studies, expert

elicitation and analysis of available data. The parameter estimates derived for both types of model require rigorous validation of parameter estimates (Kao, 2002).

Deterministic vs stochastic models in disease simulation

Deterministic models use a single value, usually an average, for each input value and therefore the outcome is a single value representing the average pattern of the outcome of interest. Deterministic models cannot provide a quantified outcome in a particular situation as no event is truly average. However, deterministic models can be valuable to increase understanding and knowledge, to raise hypotheses for further investigation (Green and Medley, 2002). An example of a deterministic model applied to foot-and-mouth disease is the differential equation approach of the 'Ferguson Model' used in the 2001 UK foot-and-mouth disease outbreak (Ferguson et. al., 2001 quoted in Kao, 2002). The mathematical models developed by Keeling et al., (2001) and Woolhouse et al., (2001) are stochastic.

Heuer et al., (2007) utilised both a deterministic and stochastic mathematical models in their assessment of likelihood of detection of foot-and-mouth disease. Both were SLIR (susceptible-latent-infectious-resistant) models linked to a clinical state by the rate of development of clinical signs. Whereas the deterministic model used a single value, the stochastic model assumed that individuals moved from one state to another using a Poisson-distributed average. In their conclusion Heuer et al., (2007) summarises the values of the different approaches:

'Whereas deterministic modelling is useful for evaluating the relative merits of various components such as detection sensitivity, dissemination rate, or sample size, they lack the ability to give an accurate account of start and finish dates for surveillance teams in the field in the event that good prior information about the day of the incursion is available. Insights into the variability of the assumed parameters and of chance events, such as whether or not infection will take place during a time interval, was more realistically approached using the stochastic process, that effectively allowed inter-event time to vary. As would be expected in real life, this resulted in various proportions of no-outbreak or fadeout runs, and by moderate differences in the outcome between median stochastic and average deterministic models. Moreover, the stochastic outcome was not symmetrically distributed around the deterministic result.'

Stochasticity is important when a model feature depends on assumptions about individual heterogeneity as specifying and using random distributions allows individuals to behave

differently (Britton and Lindenstrand, 2009). Britton and Lindenstrand (2009) used data from the H1N1 pandemic to illustrate that estimating a value of R_0 from the initial phase of an epidemic is hard without additional knowledge about the distributions of the infectious and latent periods.

Heterogeneity: spatial structure in disease modelling

Although mathematical and epidemiological models may have no spatial context, spatial structure is often important in the transmission of disease. In infectious disease local spread, windborne spread and the contact network structure are very important at different stages of an epidemic. In a situation where a non quantified outcome is wanted or the model is being utilised for simple exploration and where the proportion of infected premises is low the simple 'mean field' approximation is adequate. In this case any one unit is equally likely to infect any other. Once the proportion of infected premises increases, clustering is significant and the status of local premises is important. When this occurs it is appropriate to consider spatial structure (Kao, 2002). The work of Keeling et al., (2001) and Dangerfield et al., (2009) showed that at equilibrium their models incorporating spatial structure always displayed greater variation about the mean than their mean field model, although the differences were generally small unless the prevalence of infection was low. By contrast, during the early epidemic growth phase when the level of infection is increasing exponentially, the mean field model generally showed more variation.

The complex epidemiological models for foot-and-mouth disease consider spatial structure as this is essential for understanding host-pathogen-environment interactions. Techniques for considering spatial structure can be conceptually simple using known locations of farms (where the appropriate data are available) or can be generated using census and land information. This method is used by the epidemiological models and also the mathematical model described by Woolhouse (2001). Alternatives are conceptually more demanding and possibly simpler to implement such as the 'moment closure' technique used by Ferguson (2001) and described by Kao (2002) or the pairwise approach described by Dangerfield et al., (2009). Similar principles apply for other types of heterogeneity, such as regional variability in the intensity of control measures.

The addition of parameters

Addition of parameters to a model can increase accuracy and precision of the outcomes of interest. This increased ability to predict comes at a cost of needing better data to ensure accuracy and also increased computational expense. Not including an important parameter may equally invalidate results.

Uncertainty in the modelling process

Uncertainty is a component of imperfect knowledge. Uncertainties may arise from poor understanding of an issue, whether because of less than perfect data or an inability to encode a natural process, as well as different expectations and values: If uncertainties are characterised the evidence is more defensible and transparent. Specifically this would refer to the appropriateness of the distributions used to describe a model input, both in shape and boundaries. For deterministic factors this would be the stochastic and subjective uncertainties that lead to an inaccurate average value. Perfect knowledge of a stochastic variable results in a perfectly characterised variability. Brown (2004) has characterised imperfect knowledge in more detail and distinguishes imperfect knowledge about the real system from ignorance – the lack of awareness that knowledge is not perfect.

Characterisation of uncertainty

Uncertainty is categorised into two main groups:

1. **Stochastic** uncertainty (also known as aleatory uncertainty, random uncertainty, objective uncertainty, variability, irreducible or type A uncertainty) is the uncertainty of the 'true' value that is due to the natural variation of the population or randomness within the real system. These uncertainties are unable to be removed. They are best described by probability distributions and are therefore modelled using probabilistic techniques. Stochastic uncertainty cannot be reduced by gathering information through research or data collection. Perfect knowledge of a stochastic phenomenon does not provide the 'true' value, but provides a perfectly characterised variability.

2. **Epistemic** uncertainty (also known as subjective or type B uncertainty) is uncertainty due to lack of knowledge of the true value. Subjective uncertainties reflect the lack of

knowledge about the problem of interest and can, in theory, be reduced by increased data collection efforts (Helton and Davis, 2000; Saltelli et al., 2000). This category includes uncertainties about the model used to describe the reality, its boundary and operation conditions and mathematical errors (Beyer and Sendhoff, 2007). Epistemic uncertainty can be reduced by gathering information through research or data collection.

Sources of uncertainty in disease modelling

'*Uncertainties in model predictions arise from uncertainties in the values assumed by the model parameters (parameter uncertainty) and the uncertainties and errors associated with the structure of the model (model uncertainty). When assessing uncertainty one is interested in identifying, at some level of confidence, the range of possible and then probable values of the unknown of interest. All sources of uncertainty and variability need to be considered. Although parameter uncertainty assessment has been extensively discussed in the literature, model uncertainty is a relatively new topic of discussion by the scientific community, despite being often the major contributor to the overall uncertainty'* (Droguett and Mosleh, 2008).

In modelling uncertainty can arise from:

1) stochasticity in the population;

2) uncertainty due to model structure and factors (this includes model inputs and parameters);

3) ambiguity in defining the problem; and

4) framing assumptions.

An additional source of uncertainty is the accuracy of the recorded output data that is in turn used to measure simulation accuracy (Butts et al, 2004). This source of uncertainty is important but is not always relevant in disease modelling. For example, for outbreaks of foot-and-mouth disease the accuracy of the simulation cannot be assessed directly except when data is available for the true disease event. An example is the uncertainty in the number of infected premises, which may be based on a variety of often subjective definitions rather than data based on standard laboratory testing post outbreak (Savill et al., 2007; Tildesley et al., 2008).

Stochasticity in the population cannot be changed or decreased by improved measurement. The shape and limits of the distribution around the outputs from a disease model represents the uncertainty in the inference, and includes this element. The appropriate distribution can be produced from analysis of population data such as movement records from a database such as the Animal Movement Licensing System (AMLS2) in the UK analysis of event information from previous outbreaks in the country or overseas, analysis of scientific literature, Delphi conference, or other consultation with experts in the field (Sanson et al., 2006).

Uncertainty in the model structure is included in the concept of sensitivity analysis addressed in this thesis, but this uncertainty cannot be wholly addressed by sensitivity analysis, but use of model comparisons and real epidemic data can help to untangle these uncertainties.

Ambiguity in defining the problem and framing assumptions can only be remedied by gaining a shared understanding of the problem through iterative dialogue between the modellers, policy makers and experts (Butts et al., 2004; Jakeman et al., 2006; Refsgaard et al., 2007).

In practice uncertainty may be thought of as the degree of confidence that a decision maker has about possible outcomes and/or probabilities of these outcomes. Uncertainty must be differentiated from wrong judgement. However, if the uncertainties of model output are not well understood then the decision maker is not able to have confidence in the model and will be open to misjudging the output. Uncertainty must be understood in order to properly interpret the model outcome and must be differentiated from ignorance. Uncertainty is a known degree of unreliability, whereas ignorance is a lack of awareness that knowledge is wrong or imperfect. It is particularly important, but difficult, to consider uncertainty in large models. Therefore, the important thing is that evidence, including that provided by models is transparent in its assumptions and incorporates differing views, value judgements and framing assumptions. The mapping of model assumptions, factors, laws and structures onto model inferences is uncertainty analysis and sensitivity analysis is a component of this process.

The role of sensitivity analysis in supporting evidence

Regardless of the discipline, models aim to be accurate representations of a system of interest to aid decision making, improve understanding of the real system or as conservative ‘screening models’ to broadly examine situations or pathways of interest (Frey and Patil 2002). Although

models are significantly different for different disciplines there are many common principles, allowing techniques developed for one discipline to be applied in other areas.

Scientists need to provide evidence that is defensible and transparent in its assumptions. Decisions that have large social, economic or political impacts will invariably come under scrutiny. This creates demands for evidence as *ad hoc* informal consideration of information is insufficient. The best available data must be used to provide this evidence (Caro 2000). When models are used for policy analysis they provide evidence based on incomplete knowledge, sometimes in the form of probability, in situations where '*facts are uncertain, values in dispute, stakes high and decisions urgent'* Saltelli (2006).

Scepticism about the use of models is apparent in disease modelling, but is equally encountered in environmental and engineering fields and is not confined to academia (Saltelli 2006). Van der Sluijs et al., (2007) describe a debate concerning the use of environmental models in the absence of proper and reliable quality criteria for the assessment of model uncertainties.

Controversy occurs when there is uncertainty. It is easy for a poorly informed non-modeller to remain unaware of limitations, uncertainties, omissions and subjective choices in models. If this occurs, there is a risk that:

- too much is read into the outputs and/or predictions of the model, creating confidence in inappropriately constructed models and a reliance on models as a substitute for analysis of field data (Kitching 2005); or
- the model is ignored and intuition or other inferior tools relied upon; or
- a model is used for purposes different from those intended, making invalid conclusions very likely (Jakeman 2006; Letcher 2006).

As discussed, a model is considered a heuristic tool that can be used to provide evidence for decision makers (Oreskes 1994). However, the relationship between the inputs (factors, assumptions and structures) and outputs is not intuitive. Satelli (2006) summarises: '*The uncertainty associated with encoding is extremely difficult to quantify and cannot be scrutinised directly against evidence, therefore a justification of the encoding process (what was excluded and included) would be considered an important element of the [sensitivity] analysis. If the*

framing assumptions and their values are not recognised studies appear more factual and value-neutral than is warranted.'

Or, in the opinion of Hornberger and Spear (1991):

'... most simulation models will be complex, with many parameters, state variables and non-linear relations. Under the best circumstances, such models have many degrees of freedom and, with judicious fiddling, can be made to produce virtually any desired behaviour, often with both plausible structure and parameter values.'
Or, more concisely:

'Cynics say that models can be made to conclude anything provided that suitable assumptions are fed into them.' (The Economist (1998) cited in Saltelli 2000).

The only way to mitigate these risks is to increase awareness of what the whole modelling process entails, what choices are made, what constitutes good practice for testing and applying models, how the results of using models should be viewed, and what sorts of questions users should be asking of modellers. This amounts to specifying good model practice, in terms of development, reporting and critical review of models (Jakeman 2006).

Despite good practices that are well established in chemical engineering, biostatistics, risk analysis and environmental modelling (Saltelli 2006; Jakeman 2006; Refsgaard 2007) there is little evidence of application of these practices in the wider scientific community. Sensitivity and uncertainty analysis are important components of good practice to obtain credible results and valuable information from a modelling process (Campolongo 2007).

Chapter 2: Sensitivity analysis and techniques

Introduction

In the last ten years there have been several notable reviews of sensitivity analysis methods in the scientific literature: Frey and Patil (2002) considering application to food safety risk assessment, the comprehensive review of Helton et al., (2006) with a focus on determination of results, and Saltelli et al., (2006) reviewed published articles using sensitivity analysis as a key word. Cariboni et al. looked at the use of sensitivity analysis in ecological modelling and Ratto et al. (2007) focussed on sensitivity analysis used in mechanistic modelling in hydrology. This is in addition to the widely referenced books by Saltelli (2000), Saltelli et al. (2004) and the chapter discussing sampling based methods by Helton and Davies (2000).

Sensitivity analysis (SA) was initially intended to simply deal with the uncertainties in the input variables and the model parameters, but has been extended to include model conceptual uncertainty (uncertainty in model structure, assumptions and specifications). *'It is an important element of the judgement for the corroboration, or falsification, of the scientific hypotheses embedded in a model.'* Saltelli (2000).

The aim of this chapter is to describe published methods for this overall process of sensitivity analysis in a manner that will increase understanding of the application of the process to a complex stochastic, spatially explicit disease model.

This will be achieved by providing:

1. A simple guide to uncertainty and sensitivity analysis. Using a simple model adapted and simplified from Saltelli et al. (2000) to describe the assumptions and concepts associated with, and used to describe, sensitivity analysis methods.

2. A literature review of selected approaches to sensitivity analysis.

3. A discussion of considerations to be taken into account when choosing a sensitivity analysis method.

Uncertainty analysis and sensitivity analysis

Uncertainty analysis focuses on how uncertainty in the model input propagates through the model and affects the model output. Sensitivity analysis considers how much each individual source of uncertainty contributes to the output uncertainty (Crosetto 2001). This can be expanded to include uncertainty in the problem and the model itself.

A model can be represented as:

$$y = f(x_1, x_2, ..., x_k) \quad or \quad y = f(X), x \in \Delta$$

Where y is the model output, x are the input factors, some of which are uncertain, k is the number of factors whose variation is of interest, f is a deterministic function that represents the model structure and code, and X is a vector of the k input arguments from a defined domain Δ.

As uncertainty can be propagated through different structures, *f* does not have to be assumed to be constant. Each value of *x* may be a surrogate for a more complex underlying process. Each is assumed to have a true value, imprecisely known because of lack of knowledge, or in the case of a stochastic uncertainty, it is considered to have a 'true' value in order to assess its importance relative to other factors.

The underlying idea is that uncertain analysis results are functions of uncertain analysis inputs. In turn uncertainty in the inputs results in a corresponding uncertainty in the results.

Therefore, uncertainty analysis answers the question: what is the uncertainty in the output given the uncertainty in the input? (Heldon and Davis 2000; Saltelli 2000). In the broader sense this may be expressed as: what is the degree of confidence I can have in the output from this model?

This process produces a distribution of values for the output that incorporates all uncertainties, including model assumptions.

In the event that uncertainty analysis assumes that the appropriate model has been selected (bias is negligible), the effect of the uncertainty associated with the factor is studied in the uncertainty

analysis. The outcome of the uncertainty analysis would be a characterised distribution of uncertainty around the outcome of interest (Crosetto et al., 2001).

SA is a component of the overall uncertainty analysis around the question: how important are the individual elements of the inputs with respect to the uncertainty in the output of interest (Heldon and Davis 2000; Saltelli 2000). Importance is the term used to capture how much or what kind of influence each factor has on the output(s) of interest (Morris 2006).

For our model:

$$y = f(\underset{\sim}{x}), \underset{\sim}{x} \in \Delta$$

A particular input x_i would be important if:

- *$\delta y / \delta x_i$ is large in at least some regions of Δ: i.e.y varies substantially as the value of x changes; or*
- *y is relatively complex as a function of x*

Saltelli (2000) provides the following definition of sensitivity analysis 's*ensitivity analysis studies the relationships between information flowing in and out of the model'.*

To be useful a sensitivity analysis quantitatively apportions the uncertainty, or at least indicates the relative effect each factor has on the output distribution. '*The scope of SA is not only to quantify and rank in order of importance the sources of prediction uncertainty, but, also to identify the elements (parameters, assumptions, structures, etc.) that are mostly responsible for the model realisations in the acceptable range'* (Saltelli et al., 2006).

Assumptions and concepts associated with, and used to describe, sensitivity analysis methods

Many different strategies and methods are available to perform sensitivity analysis. Terminology around corroboration of models is still controversial (Kleijnen 1995; Rykiel 1996; Oreskes 1998). The terminology around sensitivity analysis can therefore be confusing.

The term 'sensitivity analysis' is traditionally applied to the apportioning of the uncertainty of the output. However, more recently the term is used for a process that incorporates both a

determination of uncertainty and the apportioning of that uncertainty. Although the 'sensitivity analysis method' is the technique used to produce the measure of sensitivity, it is in fact an overall approach based on an objective. Sensitivity analysis is described in the literature in relation to the following characteristics which will be discussed later:

1) The purpose of the analysis, for example *screening techniques.*

2) The area of the input space that is explored, for example *global* or *local* (also known as *one at a time* methods).

3) The method of obtaining the data for analysis, for example *sampling based*, *importance measures*, or *differential analysis*.

    a. The sampling technique itself where applicable, for example *Latin Hypercube sampling*.

4) The method of analysing the importance of factors, for example *variance based techniques, response surface methodology* or *differential analysis.* These methods may be *qualitative* or *quantitative* and may also be referred to by the sensitivity measure used, for example *regression, partial rank correlation coefficients*.

'Mathematical' methods for sensitivity analysis are described by (Frey and Patil, 2002). Cacuci and Ionescu-Bujor (2005) and Ionescu-Bujor et al. (2005) differentiate deterministic from statistical methods. Campolongo et al. (2000) uses an approach based on the objective of the analysis: differential analysis: (local), Monte Carlo analysis (global), importance measures including the Fourier Amplitude Sensitivity Test (FAST) and response surface methodology. Each of these descriptors are not mutually exclusive, for example differential analysis is a local technique, the Morris method is a global screening method, and variance based measures can be used in conjunction with a sampling technique. The overall approach may involve a combination of techniques, for example a *screening method* followed by a *quantitative technique*.

As there is such a difference in descriptors it is useful to consider the terminology, concepts and assumptions associated with the methods.

In a model the factors (inputs and parameters) form a dynamic matrix called the input space. The input space contains all possible combinations of values of these variables. At any one point in

the input space each factor has a specific value, somewhere in its range of distribution. The current location in the input space changes over time as different values of inputs are selected between or within a model run.

The relationships between factors affect the choice of technique, the choice of evaluation method, and number of individual points that need to be evaluated to determine that effect of the factor on the uncertainty of the outcome.

Major assumptions of some sensitivity analysis methods are additivity and linearity. Each factor has a direct or main effect on the uncertainty of the output that is independent of the effects of another variable. This may be the only effect on output uncertainty that is constant for all values of the factor and does not change when combined with other factors. If this is the case, measuring the influence of the factor at any point in the input space represents the true influence on the output variance. This will not be the case if the model is non-linear or not additive. Before an uncertainty analysis is undertaken it should be checked that the assumptions of statistical independence of the input parameters are satisfied by analysing the available data (Blower and Dowlatabadi 1994).

Additivity

In an additive model the effects of two factors are added together to obtain their joint effect. If the effects of simultaneous changes in one or more factors are not equivalent to the sum of the individual effects interactions exist between these factors. Therefore, the main effect will not represent the total effect of the factor on the uncertainty of the output. The total effect is the main effect plus or minus the effect of the interaction, and can only be measured if the technique allows the interacting variables to change simultaneously.

For our model:

$$y = f(\underset{\sim}{x}), \underset{\sim}{x} \in \Delta$$

For a system with k factors there may be interaction terms up to the order k.

Linearity

As a factor is varied across its range it may have a non-linear effect on the uncertainty of an output. That is, it may only be influential on the output uncertainty at certain values of its range, or the influence changes, in a non-linear fashion, across the range.

The true effect of the factor can only be determined if it is allowed to vary across its whole range, and if a technique can reflect the variation that exists.

The assumption of linearity '*effectively boils the behaviour of the model down to a single slope parameter for every factor'* (Morris 2006).

For our model:

$$y = f(\underset{\sim}{x}), \underset{\sim}{x} \in \Delta$$

The linearity assumption implies that $\delta y / \delta x_i$ *takes the same value everywhere in* Δ.

Monotonicity

If a relationship is not linear it may still be monotonic. In a monotonic relationship, if the factor's value changes the effect on the output will be in a consistent direction. The incremental change may differ but not the direction. For example:

1) As the factor's value increases the effect on the output may increase by a differing increment but it always increases.

2) As the factor's value decreases the effect on the output may increase by a differing increment but it always increases.

For our model:

$$y = f(\underset{\sim}{x}), \underset{\sim}{x} \in \Delta$$

The monotonicity assumption implies that it may be acceptable to assume that the slope of y with respect x_i *averaged over* Δ *is an acceptable measure of input importance.*

Global

In a global technique, all factors of interest are allowed to vary simultaneously over their entire range. This allows the entire input space to be explored and any calculation of the effect of a factor will take into account the relationships between variables as well as the main effect of the variable itself. The sensitivity analysis output should represent the total effect of a factor provided the technique's assumptions are fulfilled and the technique has been appropriately applied. Global techniques can identify interactions in non-linear and non-additive models, and as they do not require assumptions of additivity or linearity, they are model-independent.

Local

When a local technique is used the input space is only explored in a small area. Local SA techniques effectively reduce the input space to a small enough size that an assumption of approximate linearity is plausible (Morris 2006). Results should not be extrapolated to other areas of the space.

For our model:

$$y = f(\underset{\sim}{x}), \underset{\sim}{x} \in \Delta$$

In a local technique Δ is, in effect, defined to be small enough that an assumption of approximate linearity is plausible.

In a one at a time technique only one factor is varied, whilst all others are held at a nominal value, such as their most likely value. As only one factor is changing, the effect on the outcome variation of interactions cannot be considered. One at a time techniques are not appropriate if interactions exist for the factors of interest, but they may be suitable for non linear effects if the range that the factor is varied across is sufficient.

For our model:

$$y = f(\underset{\sim}{x}), \underset{\sim}{x} \in \Delta$$

The values of (k – 1) parameters are fixed and the value of the k^{th} parameter is varied over a specified range.

Only one parameter may be varied, only a small region of the space is explored and the values of the (k -1) parameters should be estimated with a very high degree of precision.

Qualitative versus quantitative

Qualitative methods do not give any information on the relative difference of importance. They are most suited to screening. Quantitative techniques generally provide information on the amount of variance in the output that is explained by each factor.

Number of runs

The number of runs can be reduced by increasing the complexity of the assumptions about the model. Local methods are the least computationally demanding as they have an assumption of linearity so they are able to produce derivative based measures with simulations much lower in number than the number of derivatives to be estimated.

Morris (2006) explores the number of evaluations required for screening methods.

For our model:

$$y = f(\underset{\sim}{x}), \underset{\sim}{x} \in \Delta$$

Where k is large the number of individual points that are needed to characterise y as a factor of x with sufficient resolution without extensive knowledge or assumptions about the nature of f is large. As the assumptions about the nature of f become more complex the number of individual points decreases.

The number of runs can be reduced if the following can correctly be assumed:

1) If **linearity** can be assumed, the number of runs can be reduced as evaluations in one region of the input space can be extrapolated.

2) If one can assume that most inputs have little or no effect on the output then strategies can be applied to identify important inputs using smaller numbers of runs.

3) **Monotonicity**. If linearity cannot be assumed, but *y* is monotonic for the arguments of interest a relatively modest number of runs is required to evaluate the input space. The number is higher than with the linearity assumption and is less easily determined.

Selected approaches to sensitivity analysis

This section seeks to outline some of the key types of techniques used in sensitivity analysis of a range of models. A number of techniques are covered in summary or in tabular form as they are bound by assumptions and would not be considered for use in a complex model. An attempt is made to clarify some of the groupings that are used in the literature.

Sensitivity analysis methods can be classified in a number of ways to help understand the suitability of a method for a model and its objective. Frey and Patel (2002) classify methods as mathematical, statistical or graphical as they focus on the analysis as an 'add on' to a modelling technique.

Mathematical methods are suitable for deterministic models and '*typically involve calculating the output for a few values of an input that represent the possible range of the input'* and these methods typically *'assess the impact of range of variation in the input values on the output'.* (Frey and Patel 2002).

Statistical methods are simulation-based and typically assess the effect of variance in inputs on the output distribution. One or more inputs can be varied at a time and the effects of interactions can be captured.

Graphical methods represent sensitivity in the form of graphs, charts or surfaces. They give a visual indication of how an output is affected by input variation and can be useful as a screening method or in combination with mathematical or statistical methods.

Local methods

Local methods are bound by assumptions of linearity and are unable to account for interactions or stochastic models. They include the 'mathematical methods' in Frey and Patel (2002)'s classification. A brief summary of the methods is presented in Table 2.

Local methods cannot consider the density function of the inputs, and are not model independent. Generally, all local methods are one at a time, but all one at a time methods are not local (Cariboni et al. 2007). The distribution of the output is also typically not captured; rather the mean is used to represent the output. It has been suggested that using the variance as well as the

mean of the output could provide some insight when local measures are used with stochastic models (Bar Massada and Carmel 2008).

Local methods often involve calculating a derivative which may be normalised by its standard deviation where the output and factor of interest have different ranges of uncertainty. Recent examples of their application are discussed by Bar Massada and Carmel (2008).

Recent papers compare local methods, global methods and combination of global methods with local measures, such as the work of Antonio and Hoffbauer (2008).

Table 1: One at a time local methods for sensitivivity analysis. Adapted from Frey and Patil (2002) and Helton and Davis (2003).

Name	Description	Assumptions	Advantages	Disadvantages	References
Nominal range sensitivity analysis **One at a time**	One of the model inputs is varied across its range of plausible values while holding all other inputs at nominal values.	Linearity. Accuracy of range for each sensitive input. Inputs not correlated. Deterministic model. Can use variance as well as mean to attempt to capture stochastic factors.	Simple. Easily applied.	Small portion of the space is explored. Interactions are not considered. Combined effects due to small changes in several inputs are not considered.	Bar Massada (2008); Saltelli et al. (2006).
Difference in Log-odds ratio	Application of nominal range sensitivity analysis when output is a probability.	Output is a probability. Linearity. Accuracy of range for each sensitive input. Inputs not correlated. Deterministic model.	Simple. Easily applied.	Small portion of the space is explored. Interactions are not considered. Combined effects due to small changes in several inputs are not considered.	Frey and Patil (2002).
Break- even analysis	Combinations of input values are identified for which a decision maker is indifferent for the decision options, or where equivalent levels of risk reduction exist. The range of the inputs is considered to see if all options are plausible. A line or curve is applied to represent the indifference or 'break even' points. If the range of uncertainty for an input encloses the break-even point the input will be important in making the decision.		Can be used to direct further research, modelling and elicitation.	Complex if more than two decisions. No clear way to rank the relative importance of the sensitive inputs.	

Name	Description	Assumptions	Advantages	Disadvantages	References
Automatic differentiation	Technique to perform local SA. Instructions are added to the code to compute derivatives automatically	Linearity necessary to order key inputs. Accuracy of range for each sensitive input. Inputs not correlated. Deterministic model. Partial derivatives must be able to be evaluated locally.	Good for models with complex numerical differentiation calculations. Do not need to have detailed knowledge of the model algorithm to interpret output. Time saving compared with traditional methods.	Small portion of the space is explored. Interactions are not considered. Combined effects due to small changes in several inputs are not considered. The possible range of the input values is not considered.	Frey and Patil (2002).

Sampling-based techniques

Many of the techniques utilised are sampling-based. In their review Helton et al. (2006) describe sampling-based approaches as involving '*the generation and exploration of a mapping from uncertain analysis inputs to uncertain analysis results.'* Components of the implementation are described:

1) Select the input factors.

2) Define the distribution.

3) Generate the sample.

4) Analysis to produce a mapping from inputs to outputs.

5) Presentation of results of the analysis, that is approximating the distributions of the elements of the output constructed from the corresponding elements of the factors.

6) Determination of sensitivity analysis results.

Select the input factors

Selection of which input factors to include in your analysis can include defining trigger parameters allowing sampling across model structures or hypotheses. Parameters should be defined to represent factors when multi-dimensional maps of factors are used.

Define the distributions

In probabilistic methods this involves the definition of distributions. In a stochastic model used to predict disease spread all our knowledge about model input is coded into the joint probability distribution of the input factors (Saltelli 2000).

For our model:

$$y = f(\underset{\sim}{x}), \underset{\sim}{x} \in \Delta$$

This involves the definition of distributions that characterise the epistemic uncertainty in the elements $x_1, x_2, \ldots, x_k$ of X.

Sometimes the probability distribution for the factor is not known. In this situation it may be beneficial to determine the effect of the probability distribution of the input parameters on the values of the outcome variables. Blower and Dowlatabadi (1994) suggest to conduct multiple sensitivity analyses, each with a different set of probability distributions for input factors. A weighting scheme is described to investigate distribution effects without the need for additional computer simulations (Iman & Conover, 1980 cited by Blower and Dowlatabadi, 1994).

The distribution function for each factor can be (Helton et al. 2006):

1) Taken from the literature.

2) Derived from data by fitting an empirical distribution function.

3) Based on expert opinion.

4) Chosen to be a truncated normal distribution, where truncation serves to avoid sampling outliers.

5) Based on a Russian roulette system or a weighted system for different trigger values where a factor is a trigger for inclusion of a model structure or input file.

6) A defined correlation structure between input factors.

Helton et al. (2006) describe the characterisation of the subjective (epistemic) uncertainty distributions as '*the most important part of a sampling based analysis, as these distributions determine both the uncertainty in y and the sensitivity or the elements of y to the elements of x'*.

Stochastic uncertainty of factors is already captured in the distribution allocated to the factor in a stochastic model, and the subjective uncertainty is represented by altering these distributions to include the subjective components (has the correct shape/bounds been allocated to the factor to capture the 'true' value). This may be done by a single analysis or by a team of independent experts or something in the middle. This could provide a valuable opportunity for collaboration between modellers, field experts and policy makers. Helton provides some useful 'do nots' for the characterisation of the distributions. He recommends specifying quartiles of the corresponding cumulative distribution function rather than by specifying defining parameters for a distribution type. This is for the clarity and flexibility 'closer contact with the original sources

of the information or insight'. He also recommends averaging values for distributions from multiple experts. In a stochastic model such as the New Zealand Standard Model (NZSM) the distributions allocated to factors already contain some of the subjective uncertainty and may have been derived using a similar process. It is important to be aware of the source and justification for parameters.

Helton et al. (2006) also suggests '*A possible analysis strategy is to perform an initial exploratory analysis with rather crude definitions to identify the most important factors, then resources can be carried out with these improved uncertainty characteristics.'*

If the probability distribution function of the factor is unknown or uncertain the effect of the probability distribution function of the input parameters on the values of the outcome variables can be evaluated. These are known as *distributional effects*. To do this multiple sensitivity analyses could be carried out, each with a different set of probability distributions (Blower and Dowlatabadi 1994). A weighting scheme is available to investigate distribution effects without the need for additional computer simulations (Iman and Conover 1980).

Generation of the sample

Samples are taken as a full factorial sampling design is not possible as this involves using every value of each parameter to form every combination of parameter values. The entire parameter space is explored, but this is time consuming and impractical. Monte Carlo sampling techniques are generally utilised, and those described are random sampling, importance sampling and stratified methods such as Latin Hypercube sampling.

Simple random sampling draws random values from the probability density function for each factor by generating a random number from a uniform distribution. The sample may therefore not represent all subsets of the distribution of the factor (Helton et al. 2006; Xu et al. 2005). Latin Hypercube sampling is a type of stratified Monte Carlo sampling that was introduced to address this problem (McKay et al. 1979). This sampling regime stratifies the distribution of the factor into a number of equal-probability intervals. Random values are drawn from each of these intervals, increasing the likelihood that the sample represents all subsets.

Where Monte Carlo sampling is used there can be efficiencies in utilising a stratified sampling method, such as Latin Hypercube sampling rather than a simple random method. This is

discussed by Helton and Davis (2003), Helton et al. (2005) Helton et al. (2006), Manache and Melching (2008), and Sallaberry et al. (2008). Xu et al. (2005) reported that Latin Hypercube sampling captured more variability than a simple random technique, particularly when the sample size was small.

Helton et al. (2005) compared sensitivity analysis results using three independent random samples of 100 each and three Latin Hypercube samples of 100 each. They found that the results of both sampling techniques were robust and similar. Sample size did not appear to contribute to poor performance with either technique.

Blower and Dowlatabadi (1994) describe the technique of Latin Hypercube sampling in more detail: '*The estimation uncertainty for each input parameter is modelled by treating each input parameter as a random variable. Probability density functions are defined for each parameter, each of the marginal distributions are stratified and the value of each input parameter is the n randomly chosen. LHS is efficient as each value of each parameter is only used once in the analysis. The random samples of the input parameters form an input vector for each simulation of the model. The model is run N times. Distributions of the outcome variables can be directly derived and so LHS enables the results of a deterministic model to be interpreted within a statistical framework. Descriptive statistics are used to characterise the distributions and the SA can then be performed by calculating partial rank correlation coefficients [or another suitable method] for each input parameter and outcome variable.*

If the outcome variable is a monotonic function of each of the input parameters the LHS is the most efficient design for estimating the mean value and the population cumulative distribution function. If sample sizes are large, LHS is the most efficient design (variance of the estimate of the expectation of the function of the outcome variable is less than if simple random sampling is used) even if the monotonicity assumption does not hold. Handcock (1989) found that LHS can be at least an order of magnitude more efficient than simple random sampling.'

The design was originally proposed for statistically independent factors, but now the method has been extended to incorporate statistical dependencies.

The Blower and Dowlatabadi (1994) method for Latin Hypercube sampling for determination of sensitivity using partial rank correlation coefficients involves:

1. Determination of appropriate probability distributions.

2. Calculation of the required number of simulations. Latin Hypercube sampling involves sampling without replacement, therefore if only K draws are made (K = the number of uncertain variables) the K^{th} draw would be predetermined. The lower limit to the number of simulations (N) should be at least K + 1. The appropriate sample size for a specific analysis should also be determined by the desired significance level for the sensitivity measure.

3. Divide the range of each of the k parameters into N equi-probable intervals.

4. Generate a N × k matrix. N sampling indices of the first variable are paired randomly with N indices of the second variable, these pairs are then randomly paired with N values of the third variable and so on until all k input variables are included and the N × k matrix is complete.

The Latin Hypercube sampling technique involves sampling without replacement, therefore if only *k* draws are made (*k* = the number of uncertain variables) the k^{th} draw would be predetermined. The lower limit to the number of simulations (*N*) should be at least *k* + 1. The appropriate sample size for a specific analysis should also be determined by the desired significance level for the sensitivity measure. In their sensitivity analysis using Latin Hypercube sampling and partial ranked correlation coefficients Blower and Dowlatabadi (1994) used 100 runs.

Marino et al. (2008) discusses the issue of sample size. There is no exact rule for determining the sample size for most methods. A minimum value may be known, for example calculation of partial ranked correlation coefficients needs at least N = *k* + 1 samples, where *k* is the number of factors of interest (those that are varied) and *N* is the required sample size. These authors suggest that a good approach is to systematically increase the sample size and check if the sensitivity indices are consistently captured and ranked between two or more consecutive experiments. The principle is that there is no benefit to increasing the sample size beyond this point as the sensitivity analysis conclusions will not change.

Marino et al. (2008) describe how formal measurement of the correlation between experiments can be achieved by the top-down coefficient of concordance (TDCC). The TDCC measure is most sensitive to agreement on the top rankings from a set of rankings. If two rankings are compared the measure is the correlation coefficient based on Savage scores, if more than two are

compared the Kendall's coefficient of concordance is used. The TDCC asymptotically follows a normal distribution. The TDCC can be used to compare the optimal sample sizes between methods. An alternative method to assess the adequacy of sample size in Latin Hypercube sampling is based on the use of *t*-distribution with replicated sampling (see Sections 6 and 7 in Helton and Davis, 2000).

Replicated sampling is a good idea as the replicates will contain different factor combinations diluting any effects due to unlikely combinations and the outputs from the replicates can be used to obtain standard errors (Blower and Dowlatabadi 1994). Marino et al. (2008) used averaged replicates to reduce the effects of aleatory uncertainty, and found that there was an advantage to this approach despite the consequent loss in statistical power. Averaging three replicates for 100 samples (a total 300 simulations) appeared equivalent to utilising 1000 simulations with no replicates.

A previously performed LHS cannot generally be increased in size without altering the structure of the sample. Sallaberry et al. (2008) describes a procedure for extending the size of a LHS that results in a new LHS sample with a correlation structure similar to the original.

Propagation of sample, presentation and determination of results

There are numerous analytical methods that can used to determine measures of sensitivity. A number of sampling based methods are summarised in Table 2 but not discussed further. References to descriptive papers are provided for the reader, should further information be of interest.

Table 2 Approaches to determine results from sampling based sensitivity analysis. After Helton et al. (2006) and Campolongo (2000).

SA method	Uses	Outputs	Disadvantages	References
Scatterplots	• Can reveal the relationship between input and prediction. • Can reveal non-linear or other unexpected relationships. • Starting point in a complex analysis including other techniques. • Can provide understanding of relationships between factors and output uncertainty. • Scatterplots can be used to seek non monotonic patterns.	• Plots of the output *vs* the factor. • May be two or three dimensional. • Some formal procedures for identifying patterns in scatterplots have been described based on determining if some measure of central tendency in the outcome variable is a function of individual independent variables.	• May require initial knowledge of variable interactions.	Helton (2006), Kleijnen and Helton (1999a), Kleijnen and Helton (1999b).
Correlation	• Measures strength of the linear relationship between a factor and the output. • Can be non-linear	• Pearson sample correlation coefficient (CC). • Values between -1 and 1 (positive value positive correlation, negative value negative correlation, linear relationship is nil at 0, strong approaches 1 or -1).	• Although non-linear correlation coefficients can be used, generally only linear relationship is measured – no correlation does not preclude the existence of a well-defined non-linear relationship.	Helton (2006), Kleijnen and Helton (1999a), Kleijnen and Helton (1999b).
Regression analysis	If factors are independent: • The absolute values of the standardised regression coefficient (SRC) are useful comparative measure of importance. • The sign of the SRC indicates if the factor and the output vary in the same direction.	• Algebraic representation of the relationships between the output and one or more factor. • Usually linear least-squares models. • SRC represents factor importance based on the effect of moving the factor away from its expected value by a fixed fraction of its standard deviation on the output relative to the	• Dependencies/correlations can be included, but generally independence is assumed. • Although regression can be linear it can be difficult to determine a suitable form for nonlinear models.	Manache and Melching (2008), Kleijnen and Helton (1999a), Kleijnen and Helton (1999b), Helton et al. (2006).

SA method	Uses	Outputs	Disadvantages	References
	• Should be constructed in a stepwise manner starting with the most influential variable from bivariate analyses. • Can be non linear regression analysis.	standard deviation of the output. • Value between -1 and 1 (1 = perfect positive linear relationship, -1 perfect negative linear relationship).		
Partial correlation	• PCCs provide a measure of the linear relationship between the output and the individual inputs after correction for the linear effects on the output of the other factors. • It is related to a SRC but is less influenced by the distribution assigned to the input variable, the effects of the other factors and the magnitude of the uncertainty in the factor. • PCCs tend to be larger than CCs and SRCs. A large PCC does not necessarily imply that the corresponding input variable makes a large contribution to the uncertainty in the output variable under consideration. When the sample variables are orthogonal the CC SRC and PCCs will produce identical rankings.	• Partial correlation coefficients (PCCs) are obtained from a sequence of regression models. • It characterises the effect on the output that results from varying the input by a fixed fraction of its standard deviation with the linear effects of other variables removed.	• Assumes independence. • Highly correlated factors can cancel the other's effect. • Do not rely on statistical assumptions.	Manache and Melching (2008), Kleijnen and Helton (1999a), Kleijnen and Helton (1999b) and Helton et al. (2006).
Rank transformations	• Non parametric techniques. • Rank transformations convert a non linear, monotonic relationship between an factor and an output into a linear relationship, improving the resolution	• Values for factor and output are replaced by their corresponding rank. • Smallest value has a rank of 1 then upwards to largest. • Tied values are given their average	• Assumes monotonicity. • Do not rely on statistical assumptions. • It may not improve the analysis if more complex	Manache and Melching, (2008).

SA method	Uses	Outputs	Disadvantages	References
	of the SA. • Scatterplots can be used to check for monotonicity.	rank. • Rank correlation coefficients (RCC). • Rank regression. • Standardised rank regression coefficients (SRRCs). • Partial rank correlation coefficients (PRCCs).	relationships exist.	
Two sample tests (Smirnov, Cramér-von Mises test, Mann-Whitney test, two-sample t-test statistic	• The sample for the factor under consideration is partitioned into two subsamples based on the quantiles of the output distribution.	• If the distribution of the factor in the subsamples is different the factor is considered influential, e.g. t-test of $>90^{th}$ quantile *vs* the rest is significant.	• Not robust. • Qualitative. • Results depend on the choice of the quantile for splitting the sample.	Conover (1980), Saltelli (2000).
Stepwise regression analysis	• Alternative form of regression modelling. • A sequence of regression models is constructed: o First model: the most influential factor. o Second next most influential (given the first factor). o Third next most influential (given the first and second factor). o Test interactions o Continue until no further significant model improvement from adding variables indicates that no further variance is explained.	• The model coefficients of determination. • SRCs in individual models indicate variable importance. When the input variables are uncorrelated R^2_y = SRC = PCC.	• Care required when interpreting regression analyses when correlation is present as the coefficient may be influenced by correlation not just importance.	Campologno (2000).

SA method	Uses	Outputs	Disadvantages	References
	• Can check if regression coefficients for individual variables are significantly different from zero to aid when to stop construction of the models.			
Monte Carlo Filtering extended to Bayesian generalised likelihood estimation (GLUE) , generalised sensitivity analysis	• Factors are mapped onto the output space then factors are censored if *Y* is unacceptable. • In GLUE many combinations of factor and alternative model structure can be compared with evidence.	• Used for factor mapping.		Young (1996), Hornberger and Spear (1981)
Measures of importance	• Includes the measure of importance, the measure of Hora and Iman (1996), Iman and Hora (1990), Sobol and FAST, the importance measure of Ishigami and Homma (HIM). • Sobol and FAST are discussed elsewhere, but are generalisations of these concepts and are truly model free sensitivity measures.	• Variance correlation expectation (VCE) and the correlation ratio.[1] • The numerator is the *variance correlation expectation* (VCE) and the ratio the *correlation ratio.* • The measure of Hora and Iman (1986) is the square root of the VCE. • Iman and Hora (1990) proposed a measure based on regression, for which the rank based measures can be used • HIM is high for an influential variable,	• Computationally demanding. • Measure of importance not robust. • Rank based measures are robust but conclusions are hard to interpret back to the original models. • Influenced by outliers. • HIM is not robust.	Iman and Hora (1990), Ishigami and Homma (1990).

[1] $\frac{Var_{x_j}[E(Y|X_j = x_j]}{Var(Y)}$ Where $E(Y|X_j = x_j$ denotes the expectation of Y conditional on a fixed value of X_j, *and*, Var_{xj} stands for the variance over all possible values of X_j.

SA method	Uses	Outputs	Disadvantages	References
		smaller for a non-influential variable.		

Selected techniques for sensitivity analysis

The following global methods will be summarised in some more detail:

1) Monte Carlo sampling with partial ranked correlation coefficients.
2) The revised Morris method (Campolongo et al., 2007).
3) The variance based methods.
4) Response surface methodology

Monte Carlo sampling with partial ranked correlation coefficients

Blower and Dowlatabadi (1994) describe the technique: "*The N observations of each outcome variable may be used to assess the sensitivity of the outcome variables to the estimation uncertainty in the input parameters. The* [probability density functions] *of the input variables are rarely normally distributed and the outcome variables are generally non- linear functions of the input variables; hence non-parametric tests of ranked data are necessary. In a LHS scheme all of the parameters are varied simultaneously and the input parameters are often interdependent. Calculating partial ranked correlation coefficients (PRCC) enables that determination of the statistical relationships between each input parameter and each outcome variable while keeping all of the other parameters constant at their expected value. This procedure enables the independent effects of each parameter to be determined, even when the input parameters are correlated. A PRCC indicates the degree of monotonicity between a specific input variable and a particular outcome variable. Only outcome variables that are monotonically related to the input parameters should be used. Monotonicity can be assessed by examining scatterplots of each input variable against each outcome variable.*

The sign of the PRCC indicates the qualitative relationship between the input variable and the output variable. The magnitude indicates the importance of the uncertainty in estimating the value of the input variable in contributing to the imprecision in predicting the value of the outcome variable. The relative importance of the input variables can be directly evaluated by comparing the values of the PRCC."

Blower and Dowlatabadi describe the application of a Latin Hypercube sampling and partial ranked correlation coefficients (PRCC) technique to a deterministic HIV transmission model. The

model was complex and there was a high degree of uncertainty in the estimates of model factors. The authors describe 'complex interdependencies for several of the factors'. The uncertainty analysis described is based on the assumption that the input parameters are statistically independent. If some of the input parameters are dependant, certain combinations of parameters are more likely to occur than others and sampling should be from the appropriate joint distribution functions. LHS is also efficient in stochastic models.

The technique can be extended by using a stepwise or rank regression to determine how much of the variation in the outcome variables is due to each of the key input variables that were identified by their partial ranked correlation coefficients. Monotonicity is an important assumption of the partial ranked correlation coefficients. Any non-monotonic factor may have a low partial ranked correlation coefficients but be important (Blower and Dowlatabadi 1994).

The revised Morris method (Campolongo et al., 2007)

The Morris method was described by Morris (1991) and was extended by better defining the measure and the sampling strategy by Campolongo et al. (2007). The goal of Morris design is the screening of unimportant factors, or more specifically to determine which input parameters have effects that are:

1) negligible;

2) linear and additive; and

3) nonlinear or involved in interactions with other inputs.

The method requires a sample of the input space: in effect individually randomised one at a time experiments. Between 10 and 50 random starting points are chosen and each trajectory is generated by moving the factors of interest one at a time in a random order. For each input a number of incremental ratios, the Elementary Effects (EE), are calculated and then used to compute basic statistics and derive sensitivity information. An EE for each factor and each trajectory is calculated by subtracting the function evaluated at the random starting point from that evaluated after incrementing that factor. A sensitivity measure (μ), which assesses the overall influence of the factor on the output is effectively the mean of the distribution of these elementary effects for a factor across all trajectories. The revised sensitivity measure (μ^*) is instead calculated from an estimate of the mean of the distribution of the absolute values of the

elementary effects. The use of μ* thus addresses the problem of the effects of opposite signs which occurs when the model is non-monotonic. Unfortunately it also results in a loss of information on the sign of the effect.

A second sensitivity measure σ is effectively the standard deviation of the distribution of EEs for a factor across the trajectories. This measure estimates the higher order effects, which may be non-linear effects and/or be due to interactions with other factors.

This information can be recovered by generating both μ and μ* and comparing the values. The ranking of the factors is done by considering μ, μ* or μ and σ at the same time. If μ and μ* are both high the output factor is monotonic with respect to the input factor and the input factor has influence. Where an output factor is not monotonic with respect to an input factor the different signs will effectively cancel each other out and μ would be low while μ* or σ is high.

The computational cost of the method is considered to be r (k + 1). The design is such that r trajectories are each k + 1 points (one for each factor increment). Each trajectory provides k EEs, and one value of sensitivity per factor.

In order to provide a better coverage of the sample space without increasing computational efficiency the random starting points for the trajectories could be generated by random sampling. However a different approach has been described by Campolongo et al. (2007) that also provides better coverage of the input space without increasing the number of runs. This is done by generating a large number of random Morris trajectories and then selecting a smaller number of trajectories with the widest spread for analysis.

The Morris method requires a low sample size and therefore a relatively low computational cost when compared with quantitative techniques, but there may be considerable variability for the sensitivity measures. Bootstrap techniques can be used to estimate the variability. Although the Morris technique is global it does not resolve non-linear from interaction effects (Campolongo et al., 2007).

Despite the lower computational cost the technique performs well in their comparison with variance based techniques. The method can also be used to asses the importance of groups. A group of factors can be changed simultaneously and a combined elemental effect calculated for the group. Another enhancement of the work of Campolongo et al. (2007) is that the revised

measure allows the individual factors in a group to be changed in different directions, rather than in the same direction.

The Morris method is an efficient method and is therefore suited to a complex model where the number of uncertain factors is high and/or the model is expensive to compute. The computational cost is the key limiting factor when considering sensitivity analysis, and so screening methods are most suitable. The Morris method is global, computationally efficient and '*should be implemented as a first preliminary analysis when the execution time of the model is several hours or days'* (Marino et al., 2008).

This method may be the most appropriate method for screening of factors in a disease simulation model as the combination of Latin Hypercube Sampling and the PRCC technique described above, is trivial to code and resilient to type II errors.

'Using EE it is very easy to produce trajectories moving one coordinate at a time preventing waste of simulations if the model crashes. With as few as four or possibly only two such trajectories (each costing k + 1 runs with k number of factors) you can get a very good picture of what is going on. If you think of these trajectories as strings of k + 1 points in a k-dimensional space, it is evident that you can get even better results by selecting these strings as far apart from one another as possible' (Saltelli, pers comm.).

Variance based methods

This category includes the FAST and Sobol methods described by Saltelli et al. (2000). They are currently not applicable to disease simulation models as they are computationally too demanding, and do not cope well with stochastic factors.

The FAST approach to sensitivity analysis is based on evaluating the computer model at input vectors that lie along a parametric curve in a *k*-dimensional input space. Coordinates of the curve, in each dimension, are defined by a periodic function, generally a transformation of a sine or cosine function. When estimation of variance of conditional expectation (VCE) indices is of interest, the periods of these input functions are set to different values. This can be done is such as way as to have the curve cover the input space as thoroughly as possible, but more important, to separate contributions of output variation associated with each input. The FAST estimates of VCEs are essentially equivalent to sums of squared coefficient estimates obtained from a linear

regression of output on trigonometric functions of the inputs. The critical implied assumption is that the computer model is a reasonably simple function of the inputs (Morris et al., 2006).

A recent paper describes difficulties with FAST in handling stochastic uncertainty: this was inappropriately partitioned to the total order sensitivity index. The first order index was unaffected. The authors attempted to remedy the problem by utilising replication and averaging to limit the stochastic uncertainty. The paper also indicated the computational demands of the method. A total of 53,456 simulations were necessary for Fourier amplitude sensitivity test (FAST) compared to 300 for partial ranked correlation coefficients. FAST can identify non-monotonic sensitivities, although the presence of these may not be known *a priori* (Marino et al., 2008).

In a comparison of FAST and PRCC for an agent based model both techniques identified a similar set of important parameters. The variance based measure detected a smaller subset of important factors than the PRCC, and there were marked differences between the rankings obtained by the different techniques (Marino et al., 2008).

Sampling plans for variance based measures

In their paper Morris et al. (2006) discuss three sampling methods for variance decomposition measures of importance: the first based on substituted columns, the plan based on permuted columns described by McKay (1995) and proposes a third balanced incomplete block design. All four designs are constructed so groups of runs differ in the values of all inputs except one. The permuted column design is the most efficient, however the balanced incomplete block design eliminates estimator bias due to unwanted matches and is more efficient than the substituted columns designs.

These sampling plans are intended to support estimation of VCE indices under minimal assumptions about the computer model. If further assumptions are reasonable, the FAST and an approach based on the use of spatial Gaussian processes described by Oakley and O'Hagan, (2004) are alternatives.

Response surface methodology

Campolongo (2000) provides a concise outline of the technique:

'*The procedure is based on developing a response surface approximation to* [a metamodel of] *the model under consideration. This approximation is then used as a surrogate for the original model in uncertainty and sensitivity analysis. The analysis can be divided into six steps:*

1) *Selection of ranges and distributions for each variable.*
2) *Development of an experimental design defining the combinations of variable values for which model evaluation will be performed.*
3) *Evaluations of the models.*
4) *Construction of a response surface approximation to the original model (usually using least squares approaches).*
5) *Uncertainty analysis.*
6) *Sensitivity analysis.*

Different types of experimental designs are available to select the design points at which the model will be evaluated. The one ultimately selected will depend on many factors: the number of independent variables under consideration, the possible presence of quadratic or higher order effects, the possible importance of variable interactions and the computational effort required to evaluate the model.

There is an important distinction from Monte Carlo analysis, the points are selected by classical experimental design ensuring that a specified structure exists between the values of individual factors but no probabilistic weight is assigned to the factors. In a Monte Carlo analysis the weighting can be useful in the construction of estimated means and variances for the output.'

LHS/PRCC can also be a first step to construct response surfaces (key variables can be used as the best subset of predictor variables to determine the relationship between the independent and dependent variables (Blower and Dowlatabadi 1994).

In a Bayesian modelling framework, the posterior distribution includes any functions of the computer model, including the VCEs. This approach has substantial practical advantages relative to the sampling-based experiments. For example, any experimental design can be used for this purpose (Morris 2006).

Any method that explicitly utilises the selected input values in estimating VCEs has the potential to be more efficient than sampling-based techniques. The price for this efficiency is that explicit or tacit assumptions must be made about the link between input values and output values. Or, as Morris et al. (2006) summarise: '*Selection of the most appropriate approach is largely dependent on what can be safely assumed about the behaviour of y as a function of x.'*

Recent applications of RSM can be found in robotics (Kewlani and Iagnemma 2008) and envirionmental modelling (Iooss, Van Dorpe, and Devictor 2006). In the work of Marrel et al. (2008) a Gaussian process model is developed and applied to a complex hydrogeological computer code.

Bayesian analysis is also used to account for model selection uncertainty where several alternative models are available. Bayesian analysis is usually implemented by the so-called Bayesian model averaging (BMA) technique described by Hoeting et al. (1999). BMA may be useful in the approach to sensitivity analysis in which all possible sources of uncertainty are taken into account (model structures, model parameters and data uncertainty). The technique involves averaging of the posterior distributions of all applicable models weighted by their posterior probability, providing a better average predictive value. In their paper Bishop and Shanley (2008) discuss the problem in adding stochastic perturbations to estimates that are realistic.

Bayesian model averaging can be combined with sensitivity analysis. When this is done the sensitivity analysis helps to characterise the properties of the posterior distribution, allowing the identification of the input factors or combinations of factors that are mostly controlled by data and hence are mostly responsible for good model behaviour (Saltelli et al., 2006). An example where a combination of BMA and sensitivity analysis applied to time series modelling is provided in the paper.

Other non sampling based methods include evidence theory, possibility theory, fuzzy set theory, interval analysis, FORM and SORM. These methods will not be discussed further.

Choosing a sensitivity analysis method

Despite the development of good practice and guidelines for sensitivity analysis in a recent review of published methods in *Science*, one at a time methods were found, mostly unjustified and illicit '*The Monte Carlo approach allowing for the simultaneous propagation of the entire input distributions is used only for uncertainty analysis purposes, while for SA, the methods applied are the local derivatives or the one-at-a-time approach (OAT), which is sometimes wrongly applied also for uncertainty analysis purposes'* (Saltelli et al., 2006). The only global method found was the use of partial ranked correlation coefficients and a sampling based approach in Blower et al. (2000).

The overall sensitivity analysis approach will depend on the reason for the evaluation, the characteristics of the model and the computational constraints. In his 2006 paper Morris discusses how as the complexity of a model increases there is a need to use a more sophisticated method for evaluating the uncertainty of the output as fewer assumptions can be made.

Reasons for performing sensitivity analysis

In their book Satelli et al (2000) emphasise that any sensitivity analysis has to begin with a clear goal, and explains what can be achieved by sensitivity analysis In addition, the chosen goal influences the appropriate method for analysis as establishing the goal of your analysis defines the form of the output function needed to answer the question(s).

The goal of an analysis should be a high level statement from which outcomes can be derived to implement the analysis. Usually the output uncertainty is described in terms of its variance, but it may be another measure such as shift in central tendency of an output (Saltelli and Tarantola 2002). In disease modelling goals may be to determine which factors cause the model output to vary the most or to find out the factor(s) that do or don't have influence on the output of interest. This can, in itself, have a higher goal of simplifying the model to reduce complexity and improve communication or reduce computational demand. In other disciplines the goal may simply be to partition the mean or to determine the 'best' input values to fix to obtain an appropriate output value.

Establishing the goal of your analysis and the form of the output function that answers the question(s) is the first step in sensitivity analysis. The goal should be a high level statement rather

than the model output as is. Although sensitivity analysis can be used for any outcome decided by the problem owner, such as to corroborate or falsify the scientific hypotheses embedded into the model, to provide a local measure of the effect of a given input on a given output or to gain a qualitative assessment of the uncertainty around some best estimate value of Y (Saltelli 2004).

Saltelli et al. (2000) suggests some possibilities for outcomes that may be applicable at different stages of the modelling process:

1) Factor prioritisation.

2) Factor fixing.

3) Variance cutting.

4) Factor mapping.

Factor prioritisation (where should research effort be directed?)

Factor prioritisation is when an analysis aims to identify the most important factors, those that if fixed to its true, but unknown value would lead to the greatest reduction in variance of the output Y. Other factors may be ranked in their order of importance. This method can be used to prioritise research to important factors that would have the most effect on output variance if measured more accurately. Cariboni et al. (2007) explains that this will be a probabilistic prediction as the true value is generally not known before the measurement for sensitivity, for example *'factor X_j seems to be the one that- on average, once fixed, would reduce most the output variance.'*

Factor prioritisation assumes that all uncertain factors are able to be determined at the same cost per factor. The ideal use of this outcome is to inform research.

Factor fixing (screening)

The aim of a screening process is to identify the factor or subset of input factors that can be fixed at any given value over their range of uncertainty without significantly reducing their output variance. Once identified the remaining factors explain most of the output variance. This has application in simplifying complex models, and can also help in proving or disproving a given model representation (Cariboni et al. 2007).

The concept of screening methods in sensitivity analysis is the same as that in diagnostic testing: screening methods are relatively economical and aim to provide insights into which factors are less important and then can be followed by a more computationally demanding quantitative method focussed on the important factors. In the quantitative method factors of little influence detected by the screening technique can be fixed at their nominal value without losing significant information in the model. Screening techniques are relatively economical in computational demand and the output may be a ranking of importance. Most commonly, sampling-based techniques are used as screening methods (Campolongo et al., 2007; De Pauw et al., 2008; Mokhtari et al., 2006; Morris 2006).

Mokhtari et al. (2006) evaluated sampling based (screening) methods against qualitative methods and found that while screening methods were able to identify less important inputs, they did not perform as well as the variance-based qualitative techniques in evaluating the sensitivity of the important inputs. De Pauw et al. (2008) used a combination of a screening method, the extended Morris, and computationally demanding quantitative technique, the FAST, to explore how the individual parameters interact and contribute to the output uncertainty. Morris (2006) focuses on the selection of appropriate techniques, and discusses aspects of the efficiency of screening designs and the compromise between complexity/assumptions and number of evaluations required. Screening methods should be global.

Variance cutting

In this setting, sensitivity analysis aims to reduce the variance of the output Y from its unconditional value V(Y) to a lower, pre-established threshold value by simultaneously fixing the smallest number of factors. The key difference between variance cutting and screening is that in screening the factors are considered individually, in variance cutting it is a group of factors that are fixed. This may be applicable in risk assessment models where the variance of Y may need to be equal or smaller than a given target variance.

Factor mapping

In factor mapping the aim is to determine which values of which factors give a realisation in a certain range of output space. This can allow the simulation outputs to be categorised into those acceptable or not acceptable according to evidence or opinion. ‘Which factor is most responsible

for producing realisations of Y in the region of interest?' or 'Which factor is most responsible for determining an acceptable or unacceptable realisation of Y?'

Consideration of model characteristics

Morris (2006) discusses how as the complexity of a model increases there is a need to use a more sophisticated method for evaluating the uncertainty of the output as fewer assumptions can be made.

Saltelli et al. (2000) describe four desirable properties of sensitivity analysis techniques:

1) The ability to cope with the influence of scale and shape. The influence of the input should incorporate the effect of the range of input variation and the form of its probability density function (range and parameters). Working with groups of inputs is useful for a model with many factors as this can markedly reduce the number of runs required for the analysis. However, this does result in losing information on the individual inputs in a group.

2) A global method should evaluate the effect of a factor while all others are also varying: 'multidimensional averaging'.

3) It should be model independent, that is work regardless of the additivity or linearity of the model. Interaction should be appreciated.

4) Grouped factors should be able to be treated as if they were single factors.

For evaluation of a complex disease simulation model all four properties would be necessary. Marino et al. (2008) worked with an agent based model, and refer to three other papers that attempt similar work (Lempert et al., 2002; Riggs et al., 2008; Segovia-Juarez et al., 2004).

Mokhtari and Frey (2005) introduce a decision framework for method selection for microbial food safety risk models based on the objective of the analysis, characteristics of the model under study, the amount of detail expected and the characteristics of the method. Cariboni et al. (2007) also present a decision tree, based on methods characteristics and application. Manache and Melching (2008) classified regression and correlation based techniques and provided a decision tree for method selection and inference.

Table 3 compares key model types against the ideal characteristics of sensitivity analysis approaches described by Saltelli (2000). It can be seen that only variance based methods and Monte Carlo filtering methods satisfy all four characteristics.

Table 3: Methods assessed for the 'ideal' characteristics of a sensitivity analysis method Adapted from Saltelli et al. (2000).

Method	Scale and shape	Multi dimensional averaging	Model independence	Grouping of factors
Local	No	No	No	Yes
Regression	Yes	Yes	No	No
Revised Morris	To some extent	Yes	Yes	Yes
Variance-based methods	Yes	Yes	Yes	Yes
Monte Carlo sampling	Yes	Yes	Yes	Yes

The flow diagram presented in Figure 2 explores the choice of methods taking account of their inherent assumptions.

In a disease simulation model where linearity and additivity cannot be assumed choice of methods is limited to model independent approaches and the use of rank correlation coefficients. The New Zealand standard model is computationally demanding. Variance based methods and Monte Carlo filtering can therefore be excluded as these techniques are computationally too demanding.. This suggests that suitable methods for analysis of the NZSM would be screening methods: the Morris method and a sampling based method utilising partial rank correlation coefficients.

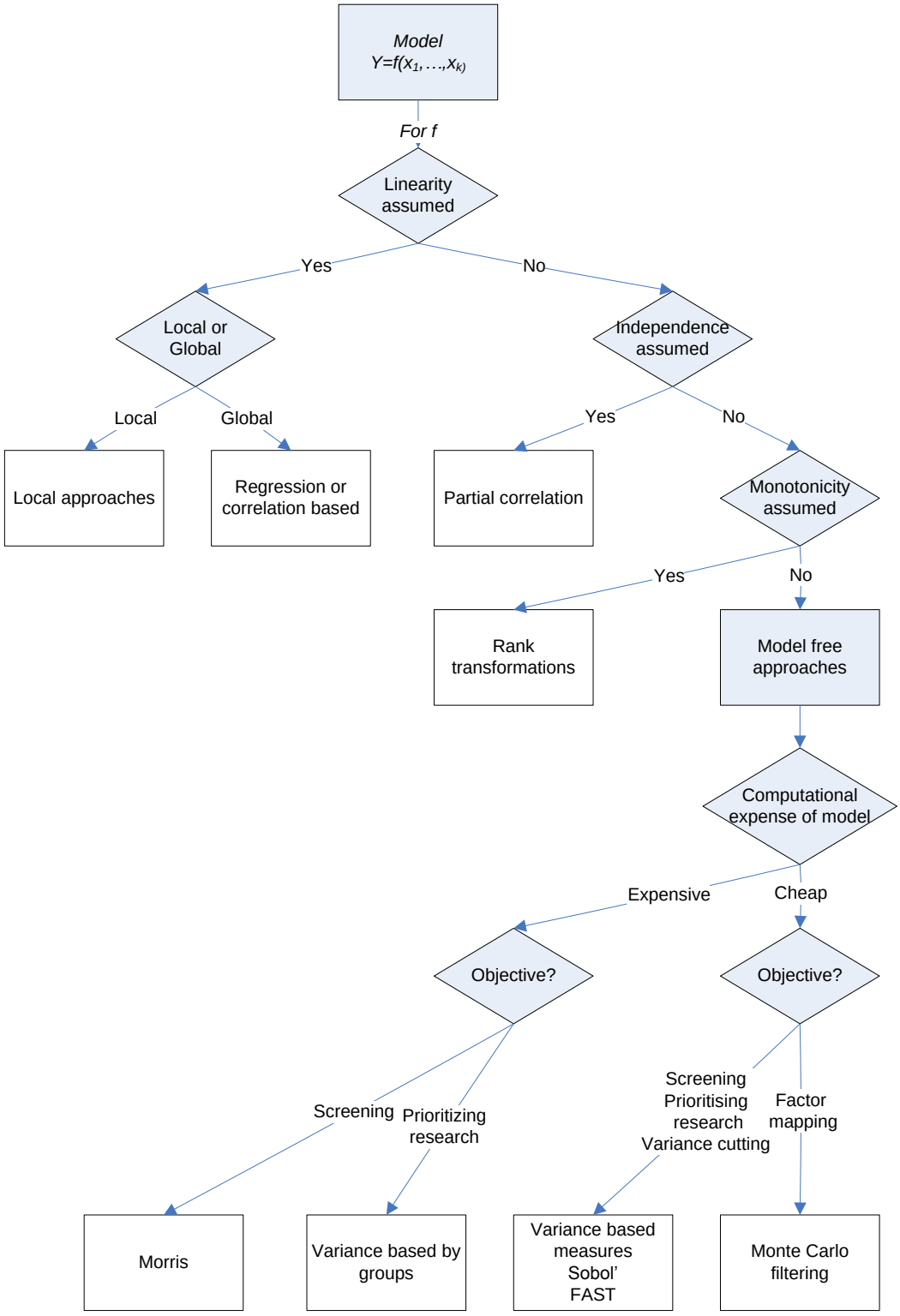


Figure 2: Selection of method for sensitivity analysis based on model assumptions. Flow diagram derived from descriptions of available methods. Model-free methods can be applied in a more assumption bound setting, but inappropriate to use in situations where this assumption is invalid.

Chapter 3: The New Zealand Standard Model

Introduction

New Zealand has long recognised its economic vulnerability to foot-and-mouth disease (FMD), and the exclusion of this disease has always been a primary focus of the competent authority. Systems to meet that end were designed over many years in the full knowledge that the virus remains endemic in Asia, Africa, the Middle East and parts of South America and it is difficult to predict the impact that any new strain will have on world animal health (Pharo 2002). The emergence of the current circulating strain of FMD type O virus in Asia has shown similarity to the emergence of the Pandemic Pan-Asia strain of type O FMD virus between 1990 and 2001. This highlights the continued threat posed by FMD in the region and illustrates that it is possible for new and significantly different FMD viruses to appear (Gibbens 2010; Paton 2010).

Modelling provides the opportunity to explore the effects of both an incursion of FMD and the control mechanisms available for its eradication. The structure of InterSpread Plus (IS+) allows for the definition of either relatively simple spread models with few parameters, for instance a single local spread mechanism utilising a radial transmission kernel, or complex models with large numbers of diverse spread mechanisms, which permit the representation of a high degree of heterogeneity, such as movement patterns between different farm classes. It also permits a wide range of control mechanisms to be introduced including resource-constrained depopulation, surveillance, movement controls, tracing activities and vaccination.

The New Zealand Standard model (NZSM) is a set of IS+ parameters defining how an epidemic of FMD might behave in New Zealand. Its purpose is to provide decision support. This may be to compare the size, duration or economic impacts of an outbreak when different disease control mechanisms are applied, to determine the benefits of different surveillance systems or to target those systems to the areas of highest risk or impact. Due to the variability of FMD virus strains and limited data available, the NZSM should be considered an exploratory tool based on the

'most likely' outbreak scenario. It utilises the characteristics of Pan-Asia O virus responsible for the epidemic of FMD that occurred in the UK in 2001, designated O UKG 2001 and current knowledge of the movement patterns of between farms and/or saleyards, and the impacts of controls that are utilised to eradicate the disease.

Materials and methods

InterSpread Plus is described in detail by Stevenson et al. (2012). IS+ is a spatially explicit state transition model of infectious disease. The units of interest are the farms (rather than animals) and farms can exist in one of several states at any time. IS+ differs from 'classic' state transition models in that farm locations are individually defined and simulation determines the transition of a farm from one state to another. This approach provides flexibility as the behaviour of population units can be modified on the basis of any number of combinations of class, state, location and/or time. In addition, to simulate the transfer of animals to a different location, population units can be divided and regrouped throughout the course of a simulation. In order to utilise the attributes of an epidemiologic model it is necessary to define the population at risk. This includes provision of details describing the locations of farms, the numbers of animals present on each farm, the frequency and type of movements or contacts that occur in the population and the surveillance systems that are in place.

Within IS+ the characteristics of the virus and its host interaction are defined by a series of epidemiological parameters that govern disease behaviour (for example, definitions of days from infection to onset of clinical signs). The model allows these factors to vary by farm class, location and/or time, if necessary.

The New Zealand standard model is parameterised to mimic best knowledge of FMD spread. To do this, the model imitates the known transmission mechanisms for foot-and-mouth disease. Three main categories of transmission are parameterised independently:

1) Spread by defined mechanisms- direct movements of animals, their products, or people and other fomites.

2) "Local spread" the undefined spread where there is no clear linkage other than vicinity.

3) Spread by wind - "airborne spread".

The NZSM simulates a series of pathways representing disease spread. Each step in the pathway is an independent probability assessment based on distributions that represent best knowledge of FMD behaviour. In the NZSM the unit of time for simulation is a single day and the number of time periods that are simulated can be varied to fit the purpose of the simulation.

Initially, the model is seeded with "virus" and there is a period often described as "silent spread" (Anderson 2002) where virus transmission occurs before clinical disease has been detected. Detection can occur on a set date after introduction, or may be entirely driven by the background surveillance parameters.

Once the disease is detected a series of parameters that define response policies are activated. The intent is to disrupt disease transmission by applying movement restrictions, intensifying surveillance activities and interventions such as depopulation and/or vaccination.

Parameterisation of the NZSM

To define the characteristics of movements among the farm population three techniques were utilised. Firstly, a diary based study was carried out by Sanson (2005). This provided information on the frequency of animal movements, the destinations to which animals were moved and the distances travelled. In this study 500 herd/flock managers were selected representing the beef, dairy, grazing/dairy heifer rearing, sheep, and mixed sheep and beef sectors. These individuals were then asked to record the movements of all animals, products, people, vehicles and equipment onto or off their farms during two separate 3-week periods, representing relatively 'busy' and 'quiet' times of the year with respect to livestock production in New Zealand. Movements were geo-coded where possible, allowing movement distances to be calculated. To reflect movements on a worst case scenario, for the NZSM 'busy' time movements were used. A total of 193 herd/flock managers returned 'busy period' diaries representing a total of 12,052 individual movement events. The risk of transfer of FMD virus was assessed assuming disease presence on the study farm at the time of the movement. The data were then analysed to establish movement frequencies and distributions of distances travelled, by each of the different pastoral livestock sectors.

The second information gathering technique involved a two-stage Delphi conference (Linstone and Turoff 1975). Opinions were solicited from a group of epidemiologists knowledgeable on biosecurity matters in New Zealand. In particular, as this area was not well covered in the diary

study, details were sought on the characteristics of the two pig farm classes used in the model and their infection characteristics.

Analysis of data from historical FMD outbreaks, in particular the 1967 and 2001 epidemics in the United Kingdom comprised the third source of information for model parameter development (Sanson 2000; Sanson 2006a). As FMD viruses are so diverse the transmission characteristics of the model are based on the 'most likely' scenario of the Pan-Asia O serotype virus.

Information to parameterise control measures is derived from the New Zealand government response procedures to stamp out FMD should it occur (unpublished). In short, the response procedures involve a policy of 'stamping-out' slaughter of all affected and in-contact susceptible animals on farm premises, followed by the disposal of carcasses by burial, burning or rendering. Affected premises are disinfected and not re-stocked for a defined period. 'Stamping-out' may also involve the pre-emptive slaughter of herds that have a high likelihood of incubating FMD due to their proximity to an infected herd or receipt of conveyors of disease. The policy recommends an initial national standstill of movements of susceptible species, the use of zones around infected premises and intensive surveillance.

Distinct movement restrictions are defined in the policies: (1) a national animal movement standstill and (2) restriction of movement within controlled areas related to detected, infected farms.

The national animal movement standstill refers to the wide-scale imposition of movement controls following the initial discovery of FMD. It is generally applied to the whole country, although in some cases it could be applied to one of the major islands. It is restrictive on high risk movements such as live animals, but may also include constraints on animal products and vehicles involved in live animal transport. The imposition remains in force generally long enough for a complete epidemiological appraisal of the national situation. In a New Zealand context, it is expected to last up to 21 days, depending on the extent of spread identified at the time of disease detection.

In the NZSM a 10 kilometre surveillance zone is implemented around each detected, infected farm and lasts for 90 days (3 months). Farms in this area are under restrictions of movement and are surveillance patrol visits are regular. In the model, similar movement restrictions are applied to farms in direct or indirect contact with known infected farms that have been detected by

tracing of movements. These latter restrictions however only apply until a full incubation period has elapsed. This is generally 14 days for FMD virus (Sanson 2006).

Once the initial tracing activities and epidemiological investigations have been completed the national standstill is contracted to an area that includes all farms at risk. Within this area high risk movements remain heavily curtailed, and movements out of the zone that risk transmission of virus are prohibited. All livestock markets in this area will remain closed, and only direct movements of animals to slaughter will be permitted.

The population at risk

Details of the farm population at risk for the NZSM are derived from AgriBase, a New Zealand farm database (Sanson and Pearson 1997). Initially, the locations of the 73 operational livestock markets that trade in susceptible species were included: 46 in the North Island and 27 in the South Island. At the time of parameterisation there were approximately 105,000 farms registered in New Zealand. In 2011 81,759 farms had location details and held at least one cattle-beast, sheep, goat deer or pig (Sanson 2006). Camelids are not included on the basis of the conclusions of a review paper by Wernery and Kaaden (2004). The data in the model are regularly populated with the most recently available farm data.

Farm State

A farm state is the FMD infection and control status of a farm and is the mechanism used to restrict or allow activities or movements related to that farm. As well as the intrinsic states-infected, clinical signs and detected, a farm receives one or more states when control measures are applied. The meanings of these user-defined states are outlined in Table 4: List of user-defined states applied to farms in the NZSM

Farm type

Farms are grouped to reflect recognised differences in expression and transmission of FMD in different management systems and species. Farm classes used for parameterisation of the NZSM are dairy, pastoral livestock, drystock, lifestyle farmers, breeder pigs and finisher pigs. Delimiters for farm files are shown in Table 5.

Results

Background surveillance

Background surveillance represents the probability of detection of an incursion by a farmer or veterinarian, the reporting of that suspicion to the New Zealand Ministry for Primary Industries (MPI) and the confirmation of the presence of FMDV by an MPI laboratory. Background surveillance is set to operate from the start of a simulation and ceases after detection of the index farm. Once disease has been detected the background surveillance parameter is replaced by a *farmer self-reporting* parameter, which reflects a higher probability of detection given that FMDV has now been confirmed in the country.

If the date of detection is not set, the interval between the day of incursion and the day of detection depends on how frequently [stock on] a farm is observed and the probability of detection of disease, if disease is present on an observed farm. In the NZSM all farms are observed daily for all farm types, but the probability of detection is lower on extensive farms. Farm state will change to detected on the day of detection. The daily probability of detection is shown in Figure 3.

Onset of clinical signs

In the NZSM the appearance of clinical signs in a herd is relative to the date of infection. The probability of clinical signs becoming apparent in a herd (after detection of the index case) is defined on the basis of a lookup table. Daily estimates of the probability of clinical signs are derived from analyses of the UK FMD data archive (Defra 2004). These probabilities are shown in Figure 5. It should be noted that the appearance of clinical signs does not mean that the presence of disease in a herd is automatically detected; rather it is an event that increases the probability of detection.

Movements

In the NZSM there are 11 movement types specified (Table 6) including those for lifestyle enterprises. Each movement has a set frequency of occurrence, an associated set of destination farm types and a probability distribution defining the range of distances over which the movement event will occur. Every day each applicable movement type is applied to each infected farm. Once a destination farm type is calculated, IS+ will randomly select a destination farm of

that type within a selected distance band. A transmission probability determines if infection of the recipient farm occurs.

Movement types

Sanson et al. (2006) defined three movement risk categories.

High risk movements involve direct animal-to-animal contact and represent the movement of possibly infected animals directly from farm to farm or via saleyards. These movements are important mechanism of transmission during the silent spread phase of an epidemic.

Medium risk movements are indirect contacts, the movement of people that have been in contact with infected animals (farmers, veterinarians, artificial inseminators, livestock handlers, trucks and some fomites and animal products).

Low risk movements are non-animal products, people who have not had direct contact with susceptible animals, other vehicles and fomites. This category also includes non-susceptible animals.

In the NZSM high risk movement are further characterised by species and farm type. Medium and low risk movements to farm are not further characterised. The model also allows for simulation of the movement of a dairy tanker.

Probability of transmission

A probability of disease transmission is associated with each movement risk category. The probability of transmission is temporally related to either the onset of clinical signs or day of infection of the source farm.

For high risk movements, transmission is set to occur from the date of infection of the source farm. The probability of transmission used for each farm at each time interval is derived from a lookup table and differs by farm type (Table 7).

The probability of transmission via medium risk movements is the same for all farm types and is set at approximately ten-fold less than that for high risk movements, based on expert elicitation. Transmission via medium risk movements begins one day before the onset of clinical signs for the farm as 5% of exposures rising to 50% of exposures over a 16-day period. The probability of

transmission for low risk movements increases from 2% on day 1 to 10% on day 16 after onset of clinical signs.

Distance

For each movement from an infected property the distance moved is determined from a lookup table with the associated probability of the movement occurring in each defined distance band.

Dairy tanker movements

Within the NZSM it is assumed that dairy tanker-related spread would only occur until enhanced biosecurity measures were put in place following initial detection of disease. FMDv may be detectable in milk as early as 2.5 days after herd exposure (Thurmond et al 2006). Using information from the FMD outbreak that occurred on the Island of Funen in Denmark in 1982 the probability of transmission 4 days prior to onset of clinical signs was in the order of one in 1000. This increased to three in 1000 two days prior to the onset of signs up until the time enhanced biosecurity measures are activated (Westergaard 1982).

Local spread

Local spread is a term used to refer to short distance (generally 10 kilometre or less) spread of virus between livestock units when there is no clear linkage other than geographical proximity (Sanson 1994). Local spread parameters in the NZSM were largely derived from an analysis of local spread reports from the UK 2001 FMD epidemic (Sanson et al., 2006a) and in developing parameters for an international model comparison exercise (Dube et. al., 2007; Sanson et. al., 2011). Local spread transmission from a farm occurs from the day of onset of clinical signs and ends once the farm is depopulated. The probability of disease transmission by this route is derived from a lookup table of probabilities as a function of distance and time, as shown in Table 8. Different species have different relative susceptibility to infection by local spread- cattle 1, sheep goats and deer 0.9 and pigs 0.8.Where farms have been detected and depopulated the probability of transmission is halved until the farm is cleared.

Airborne spread

Airborne spread is not parameterised in the NZSM as it is not an important epidemiological feature of the O UKG 2001 virus strain. However, transmission parameters have been derived by

extending the work of Sanson et al. (2000) who reported on an analysis into the start of the UK 1967-1968 FMD epidemic. These could be used to explore scenarios that include airborne spread where necessary.

Within-herd spread

Within-herd spread of FMDv on a farm depends on the incubation period (the distribution of time periods from infection to the onset of clinical signs, often termed the 'latent period') and the length of time that clinical signs are likely to be present and the waxing and waning of infectiousness as the virus spreads in the herd. Where a farm is not detected, the onset of natural immunity and eventually the reversion of the farm to a fully susceptible state occurs (Sanson et al., 2006).

Increasing infectiousness is captured by the transmission parameters associated with movement types. The waning of animal numbers in an infectious state in a farm that is not detected is parameterised using a modifier derived from a user-defined lookup table (Figure 4). This modifier is applied as a multiplier to the probability of transmission value for the relevant day for the particular spread mechanism. Waning of farm infectivity is not differentiated by type, status or the type of animals present.The Delphi conference participants believed that maximal infectiousness would be achieved by days 16 post infection under New Zealand farming conditions and this level of infectiousness would decline over a further 17 days. This was represented as a daily decline in a linear fashion applied from day 17 onwards, reaching 0 at day 33.

Onset of natural immunity is not modelled in the NZSM. Due to the surveillance activities under the implemented stamping-out policy herds are unlikely to remain undetected long enough to reach a state of being immune to disease (Sanson et al., 2006).

Methods to control disease

Movement restrictions

Nine movement controls are defined to restrict high, medium and low risk movements off farms in controlled areas (Table **Error! Reference source not found.**). These areas are defined as either explicit regional boundaries or radial zones around detected premises.

The initial standstill is in place for 14 days and other controls start at detection of the infected premises and persist for the duration of the epidemic.

As a small proportion of farmers may unwittingly or knowingly flout movement regulations in a controlled area, compliance with movement controls for controlled areas is parameterised by a proportional reduction in the baseline movement numbers. These proportions are listed in and were elicited from the Delphi conference (Sanson 2006).

Tracing

In a FMD response tracing of movements is key to identify high risk properties for surveillance visits and therefore identify infected farms not already detected. In the NZSM, forward (off) tracing is applied to infected farms of all types once they have been detected. Backward (on) tracing is applied for high risk and medium risk movements once detected.

Based on a consensus of opinion from the Delphi conference it was estimated that approximately 10% of high and medium risk movements on and off farms would be forgotten or not be able to be completed. Once a trace has detected a movement the donor or recipient farm has its status changed to trigger at-risk surveillance visits. Farms may not be added to the surveillance lists until a specified delay has occurred to reflect the time taken to trace movements. Values utilised are shown in Table 10.

Surveillance

Surveillance after detection of the index farm is parameterised to represent: (1) the visits by patrol veterinarians after a farm has been identified as at-risk as a result of backward or forward tracing; (2) surveillance that occurs in the zone around each infected and detected farm; and (3) post detection self-reporting: this background surveillance incorporates an element of enhanced awareness that occurs in an outbreak.

Patrol visits occur when a farm is put on the surveillance list following a tracing event. High risk traces are visited every day, and once visited disease will be detected. Ninety percent of farms will be visited on the day they are placed on the surveillance list and the remaining ten percent will be visited the next day. There is no difference for different farm types. Ninety percent of medium risk traces and only fifty percent of low risk traces are put on the surveillance list, and

remain on this list for 17 days. Visits to medium risk traces are most likely to occur the day after being added to the list and all will have been visited by the end of the second day. Once visited, repeat visits occur every two days. Low risk traces are visited by the third day and visits repeated every three days. Disease is always detected once the farm is visited if the farm has cattle or pigs, but detection on farms with only sheep or goats is drawn from a logistic (0.25, 0.8, 0.74, 1.7) function.[2]

Farms within 3 kilometres of each detected, infected farm are visited every 2 days. The first visit will occur on the second day of establishment of the zone, and farms will remain on the surveillance list for 17 days. Disease is always detected once a farm has been visited if cattle or pigs are present. Again, detection on farms with only sheep or goats is drawn from a logistic (0.25, 0.8, 0.74, 1.7) function.

All farms are able to self- report disease, and detection probability is higher if a farm has cattle or pigs. This reflects the intensive nature of these farm types. Detection probability is related to the time from onset of clinical signs, and the values used to define detection probability as a function of time from onset of clinical signs are provided in Figure 6.

Vaccination zones

The NZSM allows for radial vaccination zones of 1.5, 3 or 5 kilometre radius. As farms are modelled as centroids 267 metres are added to these distances to mimic the average farm radius calculated from Agribase.

Resources

An epidemic of FMD will place significant demands on the human resources available for response, especially at the early stages when surveillance, tracing, depopulation and vaccination activities (if used) occur concurrently. Time is needed to create and train these teams, and teams are limited in terms of the amount of work they can undertake in each time period. In the NZSM resource constraints can be specified to limit the numbers of animals slaughtered per day. Based on an assumption that a single slaughter team is comprised of two slaughter men, one armourer and two labourers, it is estimated that the following numbers of animals can be processed per day: 800 sheep, 400 beef cattle, 800 dairy cattle, 640 pigs, 480 deer, and 400 goats.

[2] Logistic function $y = a + c / \{1 + \exp[-b(x - m)]\}$ with parameters a, b, c, and m.

The number of resource teams that might be brought into action and by when was elicited using the Delphi conference with additional input from an operational expert. Most respondents believed 3-5 slaughter teams could be mobilised immediately, and a further 15-30 teams mobilised within about 9 days after initial detection (Sanson 2006). Following the Quads model comparison project described in Dube et. al., (2007) and Sanson et al., (2011) parameterisation of the resource for depopulation was amended to also restrict initial resource availability. One farm/day could be depopulated on day 1 increasing to 4 on the fourth day following first detection.

It should be noted that although resource intensive, the current model does not limit personnel for surveillance and tracing. Work is underway to better understand the available resources and develop suitable parameters for exploring the priorities for surveillance, tracing and vaccination should this measure be used.

Discussion

This paper has described the model of how FMD might spread in New Zealand based on a knowledge of the country's major livestock sectors and the epidemiology of the virus. Infectious disease models can be used to provide evidence for decision support in novel scenarios (Taylor 2003). Epidemiological models such as the NZSM attempt to represent the interactions that produce the characteristic behaviour of the disease within a given geographical area. To do this they utilise mathematical principles but also sample from statistical distributions to represent biological processes (the host/pathogen/environment interactions of the transmission processes and effects of control measures) and their inherent variability (Morris et al., 2001).

Although conceptually relatively intuitive, epidemiological models require a large amount of data for parameterisation and can have problems with transparency due to the large number of parameters and their complex relationships (Green and Medley 2001; Kao 2002). The NZSM utilises details drawn from spatial farm databases, direct data analysis and the results of both observational and experimental studies to define model parameters. The validity of the parameters that are used in the complex representation of the biological processes is vital, as is a sound understanding of the assumptions and limitations of the model when interpreting the results.

The epidemiology of the virus

Each of the seven serotypes of FMD virus has a different epidemiological profile and it is difficult to predict the impact that any new strain will have on world animal health (Pharo 2002). Although it is possible for new and significantly different FMD viruses to appear as FMD remains endemic in Asia, Africa, the Middle East and parts of South America, the prevalence of type O and volumes of trade and travel means it remains the major threat to FMD free countries such as New Zealand (Pharo 2002; Gibbens 2010; Paton 2010). Much of the world's information on FMD in a free country is based on details collected during the epidemic of FMD that occurred in the UK in 2001. Concerted efforts by UK authorities to record detailed information about case and non-case farm holdings throughout the course of the epidemic continue to allow detailed epidemiological analyses. It is reasonable to parameterise the NZSM on characteristics of the PanAsia O serotype responsible for the epidemic derived from this analysis.

Spread through movement

Although New Zealand does have a comprehensive database of farm location details it does not, at the time of writing, have an electronic animal identification system that allows movement information to be captured 'real time'. Until this is information is available the most appropriate method for collection of data relating to movement frequencies off dairy, beef and sheep farms remains a combination of Delphi conference and the notebook study. This can be used to estimate movements and their transmission probabilities for animal groups with a reasonable degree of certainty. A further study on pig movements has been completed by Pearson (2008) which has informed an update to pig movement parameters, however, at the present time the availability of data on deer movements is somewhat limited and further studies are required to elucidate these. In New Zealand a partnership between industry and government is overseeing the development of the National Animal Identification and Tracing system (NAIT), which was mandatory from 1 July 2012 for cattle, with deer to follow on 1 March 2013 (MPI 2012). The system will link people, property and animals and provide New Zealand livestock owners, processors and government with timely and quality information on the current location, movement history and other related attributes associated with livestock.

There is a distinct difference in spread mechanisms and movement patterns between the time that disease is first detected and movement controls are instigated. Movements of infected livestock

are of greatest significance in the period prior to initial detection. Once FMD has been detected and extensive movement controls have been instigated, the contribution of live animals to ongoing spread should be extremely limited. Continuing medium and low risk movements become significant.

There are gaps in our knowledge on the probabilities of spread associated with medium and low risk movements. During a real epidemic, the use of a comprehensive epidemic data management system such as EpiMAN (Sanson et al., 1999) may permit actual transmission probabilities to be measured and collect the appropriate data to inform future scenario exploration. It should be noted however, that FMD data is specific to the population and the viral serotype, and high quality data from a very large number of infected farms have to be collected before one can derive precise estimates. For example, even though there were approximately 2000 infected places in the 2001 UK epidemic, this still didn't provide enough data to estimate local spread probabilities with a high degree of precision.

Other spread mechanisms

Local spread is clearly an important mechanism that can operate throughout an epidemic despite the levels of movement controls implemented. This was witnessed during the 1967 and 2001 UK epidemics, although how much infection attributed to local spread was due to forgotten or non-traced medium and low-risk short distance movements is unknown. The parameters defined in the NZSM have attempted to tease out actual daily probabilities of transmission by local spread, after accounting for other forms of spread such as traceable contacts. The real-time farmer-diary movement study by Sanson (2005) accounted for all off-farm movements and informed the medium and low risk movement parameters. It is therefore possible that, by using local spread parameters derived from the UK 2001 epidemic as well as medium and low risk movement frequencies generated from New Zealand data, we are inflating the overall rate of spread. This question can only be fully resolved by careful analysis of detailed records obtained during the course of an actual outbreak.

The role that airborne spread plays in an epidemic appears to depend on the type of virus introduced, whether there is pig involvement or not, and the prevailing weather conditions (Alexandersen and Donaldson 2002; Mikkelsen et al., 2003). Airborne spread has clearly occurred in previous FMD outbreaks and, almost without exception, it has involved transmission

off an infected pig farm to susceptible cattle or sheep farms downwind. A number of studies have shown that there are large strain differences in viral excretion rates (Sellers and Parker 1969; Alexandersen et al., 2003). Although the NZSM does not represent airborne spread, transmission could be included based on transmission probabilities calculated from data from the start of the UK 1967-1968 FMD epidemic, one of the more spectacular spread events in recent history. The Isle of Wight incident in 1981 (Donaldson et al., 1982) also showed high transmission potential.

Dairy tanker spread of FMDv is unlikely to be significant, although it cannot be totally ruled out. It is only likely to occur before the initial detection, or in regions outside of the defined controlled areas, as enhanced biosecurity measures adopted by tanker operators and farmers should eliminate any risk.

<<MS: I don't think you need this paragraph because the next chapter covers sensitivity analysis in a lot of detail.>>

In conclusion, the NZSM is based on best available data to represent current knowledge of the epidemiology of FMDV Pan Asia O strain, movement patterns between farms and/or saleyards in New Zealand, and the impact of MAF policies to eradicate the disease. It is important to remember that the purpose of the model is to explore scenarios and control options, rather than to provide any definitive answer. To do this the NZSM is fit for purpose providing there is good communication of associated assumptions and constraints between those involved (Kao 2002). It is hoped that this paper helps with this understanding.

Table 4: List of user-defined states applied to farms in the NZSM (note: disease states generated by the model and are not user defined)

State	Description of user-defined state
Milking	Dairy herd in milk during the simulation
Tracing	Farm identified by tracing activities but has not yet been added to a list for contact visit surveillance
Waiting	An infected farm that has been detected, but is waiting to be depopulated and processed
Processing	An infected farm where the depopulation is underway
Completed	An infected farm once depopulation is completed
Standstill	All farms in the controlled area of the initial standstill
In controlled area x/y/z…	All farms in a specific controlled area
Surveillance zone	All farms within a controlled area associated with an infected, detected farm
High risk surveillance	A farm on the high risk surveillance list. The state is removed once the duration of the surveillance is finished. If a farm is in two high risk surveillance lists at the same time, and both surveillance lists apply this state, the state will not be reset when the farm leaves the first list.
Medium risk surveillance	A farm on the medium risk surveillance list. The state is removed once the duration of the surveillance is finished. If a farm is in two medium risk surveillance lists at the same time, and both surveillance lists apply this state, the state will not be reset when the farm leaves the first list.

Table 5: Farm file variables used in the NZSM.

Column	Details
1	Unique farm identifier.
2	Agribase farm identifier.
3	Farm type.
4	Count of beef cattle present.
5	Count of dairy cattle
6	Count of pigs present.
7	Count of sheep present.
8	Count of deer present
9	Count of goats present
10	Coordinates of the farm centroid.

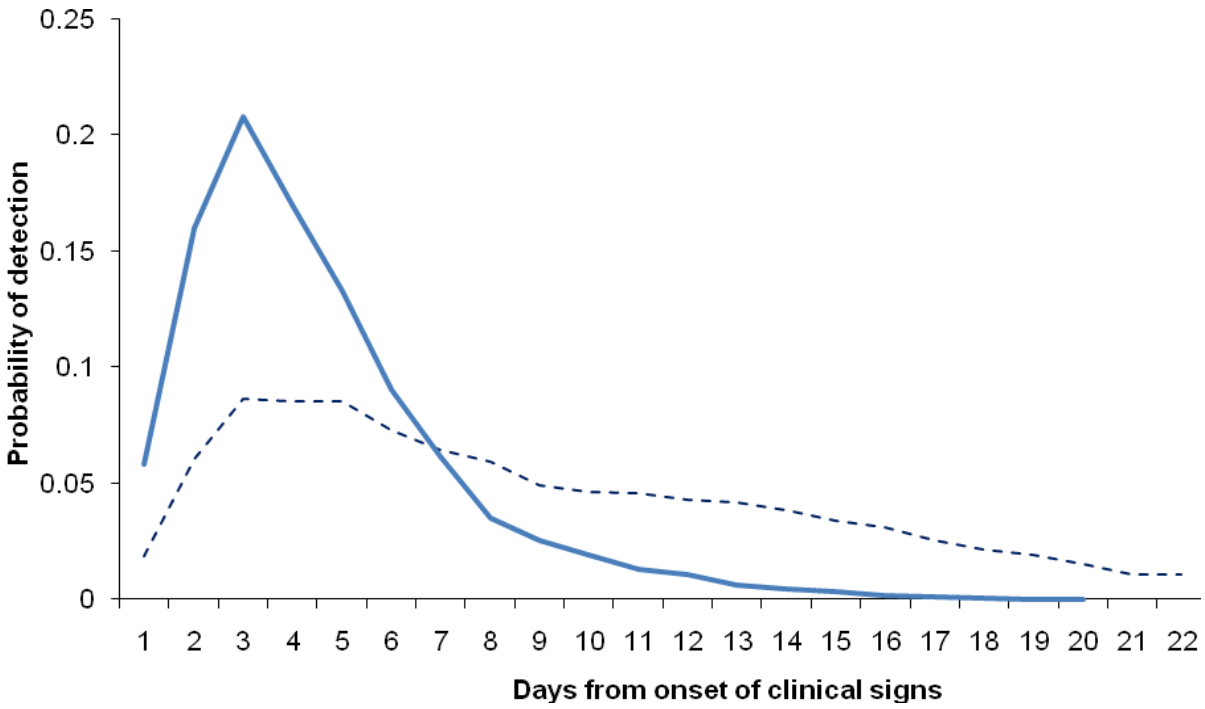


Figure 3: Detection probability for background surveillance in the NZSM. The probability of detection in a dairy, pig breeder or pig finisher farm (intensive farming type) is shown as a solid line, detection probability on a drystock or pastoral livestock farm is shown as a dashed line.

Table 6: Summary characteristics of movements to farms defined in the NZSM. Each movement results in one direct contact unless stated otherwise.

Movement category	Movement	Restrictions	Differences by farm type of destination	Distance	Frequency	Transmission
High risk to another farm	**High risk movement of pastoral livestock farm to farm**	If disease has been detected on the farm, the movement does not occur.	75% go to another pastoral farm, 23% to dairy, low probability of movement to other farm types.	Probability of a movement decreases with distance. Range 0-1000km	Number per time period Poisson 0.03	Transmission is relative to infection and increases to 0.525 at 6d PI and 1 at 16d PI
High risk to another farm	**High risk movement of dairy cattle to farm**	If disease has been detected on the farm, the movement does not occur.	Most likely to go 57% go to another dairy farm. 25% to a pastoral livestock and 20% to a drystock farm.	Probability of a movement decreases with distance. Range 0-1000km	Number per time period Poisson 0.042	Transmission is relative to infection and increases to .62 at 6d PI and 1 at 16d
High risk to another farm	**High risk movement of drystock to farm**	If disease has been detected on the farm, the movement does not occur.	80% go to a dairy farm. Low probability of movement to a pastoral or drystock farm but no movement to pig farms.	Probability of a movement decreases with distance. Range 0-1000km	Number per time period Poisson 0.1125	Transmission is relative to infection and increases to 0.458 at 6d PI and 1 at 16d PI
High risk to another farm	**High risk movement of Breeder pigs to farm**	If disease has been detected on the farm, the movement does not occur.	80% of movements go to a finisher farm, and 15% to another breeder farm. low probability of movement to a dairy, pastoral or drystock farm.	Probability of a movement decreases with distance. Range 0-1000km	Number per time period Poisson 0.1096 from pig farm Poisson 0.008 from lifestyle farm.	Transmission is relative to infection and increases to 0.458 at 6d PI and 1 at 16d PI
High risk to another farm	**High risk movement of stock to lifestyle farm**	If disease has been detected on the farm, the movement does	Based on pastoral livestock but movements 1/10 of that class	Probability of a movement decreases with distance. Range 0-1000km	Number Per Time Period Poisson 0.003	Transmission is relative to infection and increases to 0.525 at 6d PI and .9

Movement category	Movement	Restrictions	Differences by farm type of destination	Distance	Frequency	Transmission
		not occur.				at 9d PI
High risk from a saleyard	**Movment to farm from saleyard**	No restriction	Movements can occur to any farm type with different probabilities according to the origin farm	Movements decrease in probability with distance in range 0-1241km	Movements only if destination has animals	Transmission probabilities increase with time from infection up to 1 at 16 days PI
High risk to a saleyard	**Movement of pastoral livestock to saleyard**	Movement does not occur if the origin farm is in the controlled area and detected.	The number of movements/time period is qualified by origin farm type using Poisson distributions.	Movements decrease in probability with distance in range 0-1241km	No movements from pig farms. The destination of the secondary contacts is determined by Movt 06: FromSaleyard	These movements generate secondary contacts. The number farms receiving animals is determined by a Poisson distribution, lambda 1.942
High risk to a saleyard	**Movement of pigs to saleyard**	Movement does not occur if in a controlled area AND detected	Movements only from pig breeder farms and lifestyle farms. No movement from other farm types	Movement probability decreases with distance, range 0-904 km. 95% of movements up to 80km.	Number per time period from pig breeders Poisson 0.036. Lifestyle movements Poisson 0.008	These movements generate secondary contacts. The number of secondary contacts is determined by a Poisson distribution
Medium risk to another farm	**Indirect contacts to farm**	Movement does not occur if disease on the source farm is detected	Number of movements differs by farm type and is determined using Poisson distributions.	Probability decreases with distance. 80% from 0-20km, to 0.7% at 1000km	No delay to infection, Number per time period Poisson 0.264, lifestyle Poisson 0.1429	Transmission relative to clinical signs increasing from 5% at 6d post clinical signs to 50% 16d post clinical signs
Low risk to another farm	**Low risk movements to farm**	Movement does not occur if disease on the source farm is	Number is determined by a Poisson distribution. Low probability of movement. No differences for different	Probability of movement decreases by distance 90% at 0-20km, 0.2% at	1 day delay to infection.	Probability of transmission relative to clinical signs increasing from 2%

Movement category	Movement	Restrictions	Differences by farm type of destination	Distance	Frequency	Transmission
		detected	farm classes	1000km		on day 1 to 10% on day 16 post clinical signs
Dairy tanker route	**Movement of dairy tanker**	Movement does not occur if it occurs in a controlled area AND disease is detected on the farm to be visited, so transmission stops once disease is detected and the farm in a controlled area.	The movement can only go to a dairy farm.	Within 40km	Driven as a secondary contact by fixed route	Transmission probabilities are low <0.03 and start from 3 days before clinical signs.

Table 7: Probability of transmission for high risk stock movements to farm used in the NZSM.

Movement	Days from infection		
	6	11	16
Pastoral livestock high risk to farm	0.525	0.800	1.000
Dairy high risk to farm	0.620	0.800	1.000
Grazing dry high risk to farm	0.620	0.800	1.000
Pig breeding high risk to farm	0.046	0.800	1.000
From saleyard	0.046	0.776	1.000

Table 8: Probability of local spread as a function of days relative to the onset of clinical signs on the source farm and distance from infected source farm. The probability values of the last row in the table apply until the end of the simulation.

Day relative to onset of clinical signs	Distance (km)			
	1	**2**	**3**	**4**
-1	0.000	0.000	0.000	0.000
0	0.007	0.002	0.000	0.000
+1	0.012	0.003	0.001	0.000
+2	0.012	0.004	0.001	0.000
+3 onwards	0.009	0.004	0.001	0.000

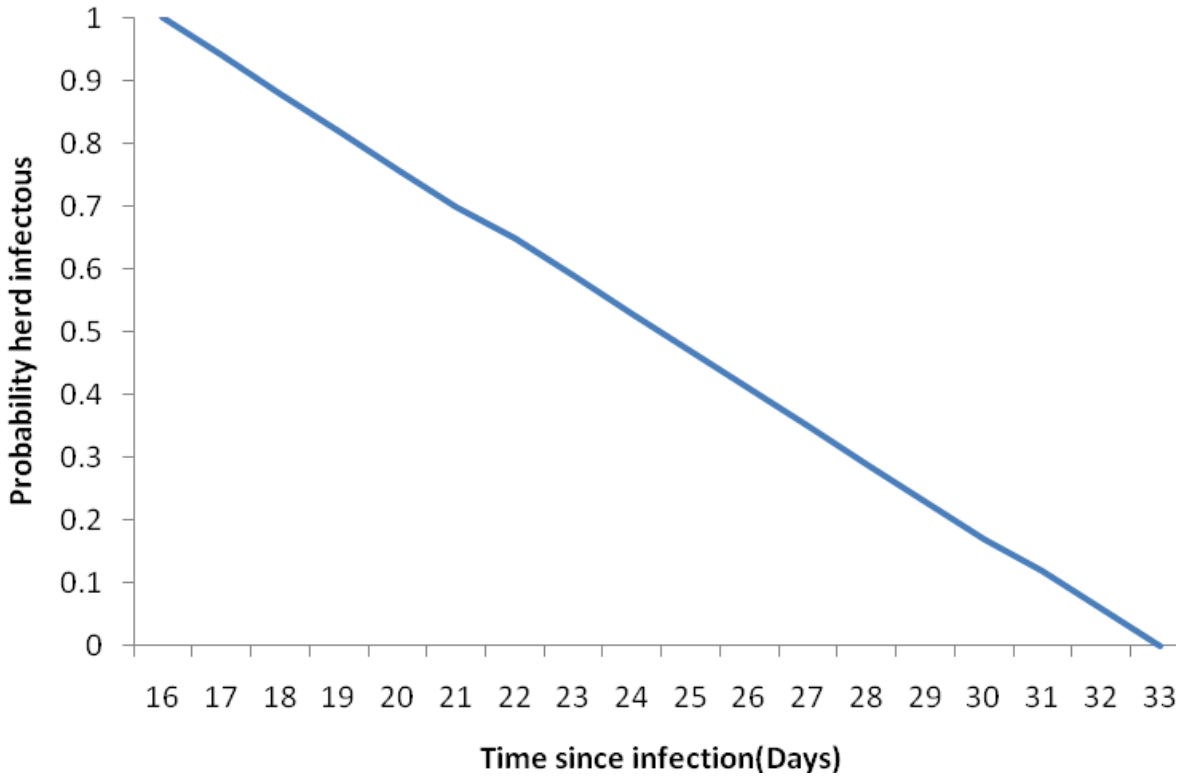


Figure 4; Within herd dynamics-probability that a herd is infectious by days since infection. A herd is infectious from day 1 to 16, after day 16 the probability a herd is infectious decreases daily until it is no longer infectious by day 33. This probability is used as a modifier to decrease the appropriate probability of transmission for movements from the herd.

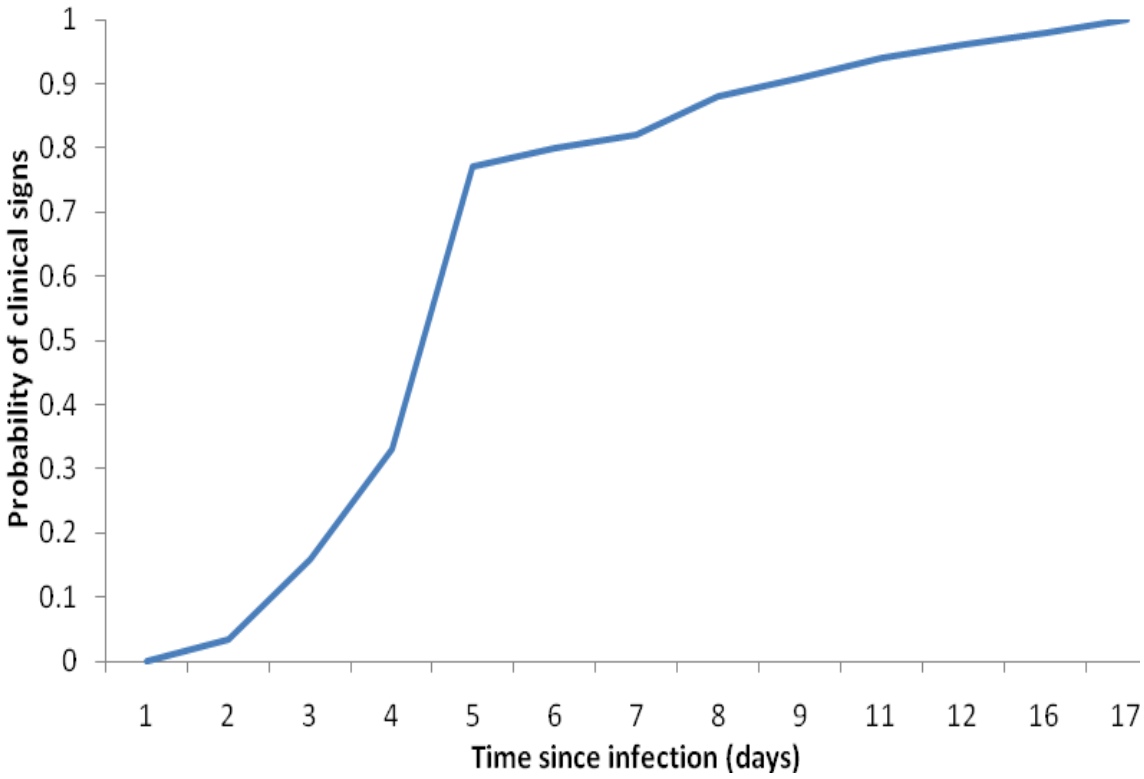


Figure 5: Within herd dynamics-Probability of clinical signs appearing in a time period from infection. The probability increases rapidly from 2-5 days post infection. Clinical signs are apparent in all herds from day 17 onwards, but this does not mean they are detected

Table 9: Settings that determine effectiveness of tracing when a movement has occurred to or from an infected farm. If tracing is required for a movement type the probability multiplier reflects if the movement has been forgotten. If the movement is not forgotten and the trace has been detected the recipient farm is added to the appropriate surveillance list after the applicable delay. Forward indicates a forward trace i.e. off an infected farm, backward indicates a backward trace i.e. onto an infected farm.

Movement type	Multiplier - Probability movement forgotten	Tracing delay
High risk to another farm (forward)	0.11	BetaPert 1 2 3
High risk on to farm (backward)	0.082	BetaPert 1 2 3
Medium risk to another farm (forward)	0.212	BetaPert 3 4 5
Medium risk on to farm (backward)	0.194	BetaPert 2 3 5
Low risk to another farm (forward)	0.36	BetaPert 4 5 7
High risk to a saleyard (forward)	0.063	BetaPert 1 2 3
High risk from a saleyard (backward)	0.058	BetaPert 1 2 3
Dairy tanker	0.014	Constant 1

Table 10: Movement modifiers reflecting control policies. On initiation of a controlled area movements are suppressed by reducing movements within the area by the proportion given in the table.

Movement type	Control area			
	National standstill	Infected zone	10km surveillance zone	In controlled area, outside surveillance and infected zone
High risk	0.914	0.942	0.951	0.951
Medium risk	0.604	0.804	0.850	0.850
Low risk	0.238	0.390	0.520	0.520

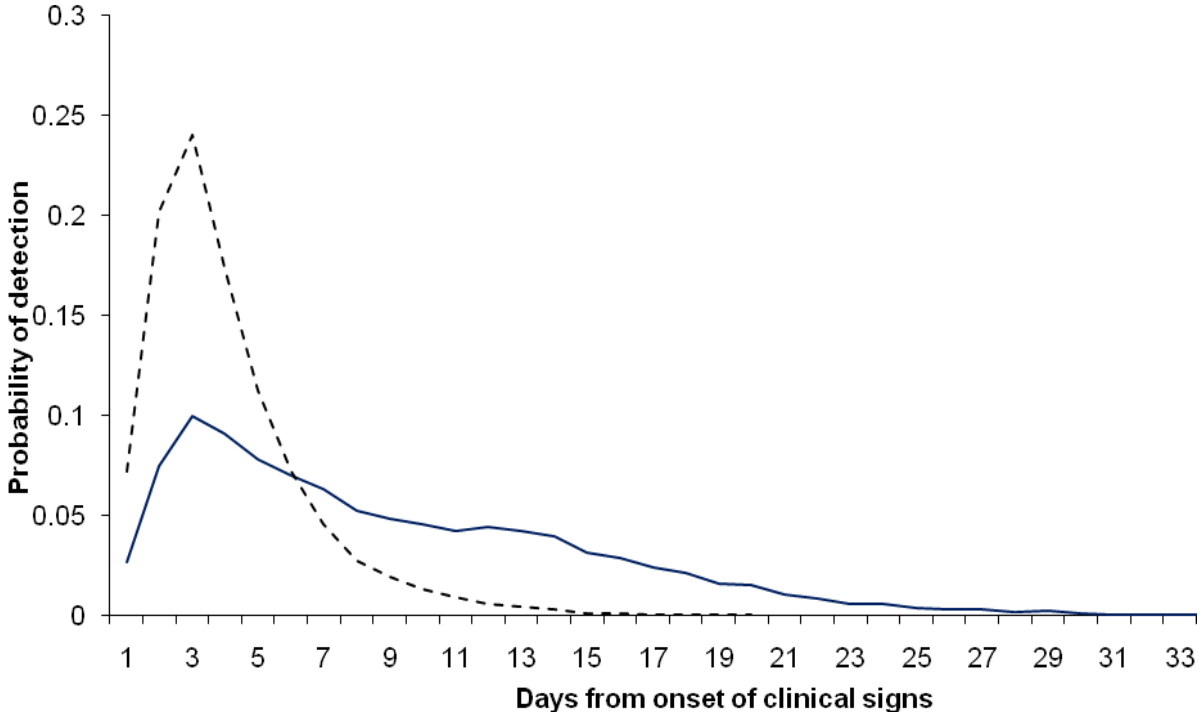


Figure 6: Detection probability for self reporting in the NZSM post detection of the index case. The probability of detection in a dairy, pig breeder or pig finisher farm is shown as a dashed line, detection probability on a drystock or pastoral livestock farm is shown as a solid line.

Chapter 4: A sensitivity analysis of the New Zealand standard model of foot-and-mouth disease

Owen, K., Stevenson, M.A., Sanson, R.L. (2011) A sensitivity analysis of the New Zealand standard model of foot-and-mouth disease. *Revue Scientifique et Technique de l'Office International des Epizooties 30(2): 513-526.*

Abstract

Disease simulation models can be a valuable tool for planning a response to exotic disease incursions, as they provide a fast, low-cost mechanism for identifying the likely outcomes of a range of outbreak scenarios and disease control strategies. To use these tools effectively and with confidence, decision-makers must understand the simplifications and framing assumptions that underlie a model's structure. Sensitivity analysis, the analytical process of identifying which input variables are the key drivers of the model's output, is a crucial process in developing this understanding.

This paper describes the application of a sampling-based sensitivity analysis to the New Zealand standard model (NZSM). This model is a parameter set developed for the InterSpread Plus model platform to allow the exploration of different outbreak scenarios for an epidemic of foot-and-mouth disease in New Zealand. Based on 200 iterations of the NZSM, run for a simulation period of 60 days, settings related to farm-to-saleyard movements and the detection of disease during the active surveillance phase of the epidemic had the greatest influence on the predicted number of infected premises. A small number of counter-intuitive findings indicated areas of model design, implementation and/or parameterisation that should be investigated further. A potentially useful result from this work would be information to aid the grouping or elimination of non-influential model settings. This would go some way towards reducing the overall complexity of the NZSM, while still allowing it to remain fit for purpose.

Introduction

In countries with good biosecurity controls at their borders, incursions of exotic diseases such as foot-and-mouth disease (FMD) are rare. Thus, predictions of the likely outcomes of a given outbreak scenario are difficult because animal health authorities generally have little experience of disease behaviour, given the (often) unique geographical distribution of susceptible livestock species and the way in which farm enterprises interact with each other. During an incursion of an exotic infectious disease, such as FMD, a range of strategies may be applied, including various combinations of culling infected herds, pre-emptive culling of herds at risk and mass vaccination. Depending on the speed with which it is implemented, each strategy is typically accompanied by a range of positive and negative consequences or 'knock-on effects' for various participants in the agricultural sector. Timely and informed decisions must be made about which control and eradication strategies should be adopted at a time of crisis. In this environment, it is important that the evidence used to inform decision-making is transparent in its assumptions and that decisions taken about a particular course of action are able to incorporate differing views, value judgements and framing assumptions (Saltelli et. al. 2006).

Regardless of discipline, the fundamental objective of modelling is to provide an accurate representation (as opposed to replication) of a system of interest (Saltelli et. al. 2006). A model that meets these objectives provides a low-cost and quick mechanism for identifying the likely outcome of a range of complex situations and scenarios. This, in turn, improves understanding of the system as a whole and can be used as an aid for decision-making (Frey & Patil 2002; Guitian & Pfeiffer 2006).

In animal health, infectious disease models have the potential to combine knowledge of the population at risk, epidemiological characteristics of the infectious agent, and the logistics of control efforts and their economic consequences, making them a valuable tool for supporting decision-making (Taylor 2003). This said, a lack of transparency in the way that models work and their framing assumptions can result in decision-makers losing confidence in their outputs and, consequently, not using them to their full potential. On the other hand, decision-makers may ignore or be unaware of the key simplifications inherent in a model, and may place too much confidence in its outputs, resulting in inappropriate ('risky') decision-making (Kitching et. al. 2005). The only way to mitigate these potential problems is to increase the decision-maker's awareness of:

- what the whole modelling process entails
- what constitutes good practice for using models
- how the results of models should be viewed
- what sorts of questions users should be asking of modellers.

This amounts to specifying good model practice in terms of development, reporting and critical review (Jakeman et. al. 2006). Sensitivity analysis, the analytical process of identifying which input variables are key drivers of the model's output, should be regarded as a key component of good model practice.

InterSpread Plus (IS+) (Sanson 1993; Stevenson et. al. 2012) is a simulation model of infectious disease designed for use with domestic animal populations. Within the IS+ framework, the unit of interest is the farm: a defined location in space containing one or more of the animal species susceptible to the disease of interest. InterSpread Plus is a state-transition model (Dietz 1966, Kermack & McKendrick 1927), with a set of defined states in which farms may be at a given point in time:

- susceptible
- infected
- clinical
- detected
- immune.

The structure of IS+ allows for a range of model definitions, from relatively simple spread models with few parameters (for instance, a single, local spread mechanism using a radial transmission kernel) to more complex models, with a range of spread mechanisms (e.g. local, airborne, and direct- and indirect-contact transmission pathways). It also provides the ability to apply a range of control strategies, including: resource-constrained depopulation, surveillance, movement controls, tracing activities and vaccination. The settings used to define each of the parameters needed to drive an IS+ model vary but, in general, require either numeric values declared as point estimates, defined distributions and/or look-up tables.

In 2005, the New Zealand Ministry of Agriculture and Forestry commissioned the development of a set of IS+ parameters to best represent the behaviour of an FMD epidemic if the virus entered the country, causing an outbreak. The intention was that this parameter set, termed the

'New Zealand standard model' (NZSM) (Sanson et. al. 2006), would be used to provide decision support before, and at the time of, an epidemic of FMD. The NZSM incorporates the known epidemiology of the disease with current knowledge of animal movement patterns between farms and/or saleyards (animal markets) in New Zealand. This allows researchers to explore different outbreak scenarios to compare size, duration or economic impacts under different control and surveillance strategies.

This paper describes the application of a sampling-based sensitivity analysis technique to the NZSM. The authors' aim was to contribute to the corroboration of the NZSM by identifying those settings in the model that had the greatest influence on the predicted number of infected premises in a simulated outbreak of FMD in New Zealand.

Materials and methods

The settings used in the NZSM model can be placed into two broad categories:

- those settings defining how disease spreads from one location to another
- settings defining how the disease will be controlled, once it has been detected.

The settings defining disease spread include details of:

- off-farm movement events (their frequency and the distance over which they occur)
- local spread (the probability of infection occurring on destination premises at given space-time separations from an infected source)
- characteristics of the FMD virus being modelled (e.g. the number of days from infection to the onset of clinical signs, and the number of days from infection to the onset of infectiousness).

The settings defining disease control include:

- details of the intensity of surveillance
- the timing, extent and effectiveness of movement restrictions, tracing activities and depopulation of farm premises.

Three distinct movement restrictions are defined within the NZSM:

- a national animal movement standstill for 14 days

- an infected-zone standstill (covering the affected region of the country)
- a 10-km surveillance zone around detected infected premises.
- movement controls around detected infected premises.

In total, the NZSM is composed of 107 individual settings within 51 parameters. Details of these parameters and the settings within each parameter are provided in Tables 11-14.

The approach adopted for the sensitivity analysis described in this paper closely follows the methodology used by Blower and Dowlatabadi (1994). In their 1994 paper, Blower and Dowlatabadi conducted a sensitivity analysis of a deterministic model of human immunodeficiency virus. Their model comprised 34 differential equations containing 20 parameters. These authors assigned a probability density function to each of the 20 parameters and used Latin Hypercube sampling (Iman & Helton 1988, McKay et. al. 1979) to sample from each distribution, ensuring that the entire range of possible values in the distribution was represented. The authors took a slightly different approach, since many of the input parameters in the NZSM were themselves defined as probability distributions. For the authors' analyses, the lower and upper bounds of the range of biologically plausible settings for each parameter of each probability distribution defined within the NZSM were specified. These bounds were then used to define the lower and upper bounds of a uniform distribution. For example, if the number of off-farm movements per day from a dairy farm was parameterised using a Poisson distribution with mean $\lambda = 0.04$ (equivalent to, on average, one off-farm movement event every 25 days), the authors specified the plausible range of values for λ as 0.01 to 0.1. That is, they believe that a single movement from a dairy farm might occur as infrequently as every 100 days ($\lambda = 0.01$) or as frequently as every 10 days ($\lambda = 0.1$). Settings defined as empirical distribution functions were entered into the model as look-up tables, and a set of three alternative candidate table definitions were defined. As an example, the probability that disease will be detected on a farm as a function of the number of days since the onset of clinical signs and two candidate distributions is shown in Figure 6.

To produce a set of data suitable for sensitivity analysis, the authors made a random draw from each uniform distribution to generate appropriate settings for each of the 107 settings of interest. For settings defined as look-up tables, a number between one and four was selected at random and the details for the corresponding look-up table were selected. A vector of length k was generated (composed of samples for each of the $k = 107$ input settings in the NZSM) and these

values were then used as the settings for a single model run. At the conclusion of the single model run, the total predicted number of infected premises after a simulation period of 60 days was calculated and stored. This process was repeated 200 times, generating a matrix comprising 108 columns (the settings for the 107 input settings plus the single numeric value representing the predicted number of infected premises) and 200 rows (the number of model runs).

Sensitivity analyses were performed by calculating partial rank correlation coefficients (PRCCs) for each input parameter and the outcome variable, using the approach described by Iman and Conover (1980), Iman and Helton (1988) and Iman et al. (1981). The significance of a non-zero PRCC value was tested by computing a *t* test statistic, which approximated a Student's *t* distribution with N-2 degrees of freedom, where N equalled the number of model runs.

Since PRCCs indicate the degree of monotonicity between two variables, care was taken to ensure that only those settings monotonically related to the output variable were used. A monotonic relationship is one in which an outcome variable moves in only one direction (up or down) as an explanatory variable increases, but the relationship is not necessarily (but can be) linear. Plots of the number of infected premises as a function of the simulated setting values were generated to identify those settings where the monotonicity assumption was satisfied (Iman and Helton 1988).

Partial rank correlation coefficients provide two useful pieces of information. First, the sign of the PRCC indicates the qualitative relationship between the input setting and the output: positive PRCCs arise when increases in the value of an input setting result in increases in the output variable; negative PRCCs arise when increases in the value of an input setting result in decreases in the output variable. Secondly, the magnitude of the PRCC indicates the importance of the input setting in contributing to the value of the outcome variable. The further the PRCC from zero, the greater the influence of the variable on the outcome. Thus, the relative importance of each of the input settings can be directly evaluated by comparing their PRCC values.

Results

For these analyses, the population of interest comprised farms located in the North Island of New Zealand. Each outbreak was initiated by seeding infection into a single farm located in the lower half of the North Island. The median predicted number of infected premises (based on 200

iterations) after 60 days was 7 (minimum 1; maximum 99). The median outbreak duration was 22 days (minimum 1; maximum 60).

Scatterplots of the predicted number of infected premises as a function of the simulated values for each setting showed that the assumption of monotonicity held for all 107 settings evaluated, and that the sampling technique provided a set of candidate values that were adequately distributed across the plausible range of values for a given setting (results not presented).

Details of settings within the parameters defining farm-to-farm and farm-to-saleyard movements, surveillance before and after detection of the outbreak, tracing and movement restrictions are shown in Tables 11, 12, 13 and 14, respectively. Table 14 also provides details of the parameters defining resources available for depopulation. Partial rank correlation coefficient values for settings related to movement, surveillance, tracing and movement restrictions, and depopulation resources are shown in Figures 8, 9 10 and 11 respectively. In Figures 8 to 11, PRCC values significantly greater or less than zero are indicated by solid circles.

Of all the movement settings used in the NZSM, farm-to-saleyard movements collectively had the greatest influence on the predicted number of infected places at 60 days (Figure 7). The settings defining the frequency of movement events off pastoral livestock and pig-breeding farms to saleyards per time period; the number of secondary contacts generated from movements of pastoral livestock, dairy, dry-grazing and pig-breeding farms to saleyards; and the probability of disease transmission from the movements of pastoral livestock, dairy, dry- grazing and pig-breeding farms had PRCC values that were positive and statistically significant at the alpha level of 0.05. Other movement types with significant PRCCs included the frequency of high-risk movements off dairy and dry-grazing farms; the frequency of medium-risk movements (off all farm types), and the probability of transmission after high-risk movements off dry-grazing and pig-breeding farms.

In the NZSM, the ‘background’ surveillance setting defined the degree of pre-epidemic surveillance for FMD that would ultimately result in detection of the first infected premises and initiation of control activities. Increases in the probability of detection during background surveillance on dry and pastoral livestock enterprises decreased the predicted number of infected premises (Figure 8). Increases in the probability of detection on dry-grazing and pig-breeding farms that self-report the presence of disease significantly decreased the number of infected

premises. Increases in the probability of detection in enterprises involving beef cattle, pigs and deer, which received low- and medium-risk contacts, decreased the predicted number of infected premises. Increases in the probability of detection on deer farms receiving low-risk contacts were associated with a significant increase in the predicted number of infected premises.

Partial rank correlation coefficient values for each of the monitored tracing parameters are shown in Figure 9. Increases in the probability of forgetting movement events off and onto pastoral livestock farms were associated with a decrease in predicted epidemic size. This result was counterintuitive. Increases in the delay in tracing high-risk movements onto dairy and pig-breeding farms were also associated with a decrease in predicted epidemic size. Increases in the probability of forgetting off-farm low- risk movements and on-farm dairy tanker movements were associated with an increase in predicted epidemic size.

For control activities, increases in the proportion of restricted high-risk movements inside the infected zone; high-risk movements out of the control area, and low-risk movements within defined surveillance areas significantly decreased the predicted number of infected premises (Figure 10). Increases in the probability of medium-risk movements being restricted during the initial standstill period, and increases in the probability of medium-risk movements being restricted outside the control area, were associated with an increase in the predicted number of infected premises, again a counterintuitive result.

Discussion

On the whole, the authors' findings made biological sense and provided indirect confidence that the NZSM parameter set provides an appropriate indication of the way FMD might spread if it were introduced into the farm animal population in New Zealand. Collectively, the settings defining farm-to-saleyard animal movements had the greatest influence on the predicted number of infected premises (Figure 7): a finding consistent with analyses of the data from the FMD outbreak that occurred in the United Kingdom in 2001 (Gibbens et. al. 2001, Morris et. al. 2001, Thrusfield et. al. 2005). This implies that efforts taken to accurately record the frequency of farm-to-saleyard movements, the number of secondary contacts and estimates of the probability of disease transmission following a movement event should enhance the accuracy of NZSM predictions.

To the best of the authors' knowledge, the work of Sanson (2005) is the only study to document details of farm-to-saleyard movements of livestock in New Zealand. Given the impact of farm-to-saleyard movement patterns on model output, it is essential that the frequency and distance estimates provided by studies of this type are updated regularly, since the propensity of livestock owners to shift animals to saleyards will vary over time and depend on the slaughter value of individual animals, as well as the costs of grazing, transport and seasonal conditions. Implementation of the National Animal Identification and Tracing System (www.nait.co.nz) and routine analysis of data recorded by this system would partly meet this requirement. Additional studies would still be required, however, to provide an estimate of disease transmission probabilities when a movement takes place.

Nine of the ten detection probability surveillance settings had significant PRCC values that were negative. This means that increases in the probability of detection were associated with a decrease in the predicted number of infected premises. A single setting, the probability of detection on premises with deer after a low-risk contact, had a positive PRCC (Figure 8). This finding was counter-intuitive. Detailed analyses of the model's behaviour – i.e. following the step-by-step sequence of infection events following low-risk movement events onto farms with deer – would be an obvious approach for investigating this anomaly further. This is an example of another benefit of the sensitivity analysis process. By identifying counter-intuitive model behaviour, a sensitivity analysis allows us to identify specific areas of the model that should be investigated in detail for possible errors in design, implementation and/or parameterisation.

Sensitivity analysis of the tracing parameters related to high-risk pastoral livestock and dairy-farm movements also presented findings that were counter-intuitive. Analyses to clarify the mechanism of these effects, using an approach similar to that described above, are required to investigate these anomalies further. An additional explanation is that the number of simulation days specified (n = 60) was insufficient to allow the full effect of changes in tracing efficacy to be reflected in model output.

Increases in the probability of restricting high-risk movements inside the infected zone and outside the control area were associated with a decrease in the predicted number of infected premises (Figure 9). This finding is consistent with the known biology of FMD (Sanson 1994). Increases in the probability of restricting medium-risk movements during the initial standstill period and outside the control area were associated with an increase in the predicted number of

infected premises. This was yet another finding that was counter-intuitive. Further analyses are required to investigate this.

If a model is non-linear or non-additive, the influence of a variable will change at different points in the input space due to interactions with other variables. In these situations, where linearity and additivity cannot be assumed, local sensitivity methods are inappropriate. Global approaches that are independent of the model, or at least assume monotonicity rather than linearity or independence, should be used (Morris et. al. 2006, Salell et. al. 2000). Global methods involve simultaneous adjustments, which allow the entire parameter domain, or at least a substantial area of the domain, to be analysed. A range of global techniques have been described to explore the behaviour of models used in economics, engineering, chemistry and physics. These techniques include 'elementary effects' methods (for example, that of Campolongo et al. (2007), based on the so-called 'Morris' method (Morris 1991); variance-based methods, and sampling-based methods, using parametric tests of ranked data of the type described in this study (Blower and Dowlatabadi 1994). Of this group, the most suitable approaches for complex disease simulation models include the elementary effects methods and sampling-based methods using parametric tests of ranked data.

Type II errors (the failure to identify a factor of considerable influence on the model) are recognised as potential problems when using parametric tests of ranked data (A. Saltelli, personal communication). An alternative would be to use an elementary effects approach, such as the adapted Morris method (Campolongo et. al. 2007). The Morris method is no more computationally demanding than the method described here and has the advantage of being more resilient to type II errors. A benefit of applying multiple sensitivity analysis techniques to the same model is that the combined knowledge provides a more detailed picture of how the parameters interact and contribute to model output uncertainty, and thus results in deeper insight into the model's behaviour (De Pauw et. al. 2008).

Although good practices are well established for sensitivity analysis of models used in chemical engineering, biostatistics and risk analysis (Jakeman et. al. 2006, 20), the uptake of these techniques appears to be relatively poor in the wider scientific community. Sensitivity analysis is an important component of good scientific practice and should be regarded as an integral part of model development, rather than as an additional and non-essential set of analyses (Refsgaard et al. 2007). The approach described in this paper should be seen as one element that contributes to

the corroboration of IS+ and the NZSM, in the context of a specific problem and particular management scenarios. Other approaches that are currently being applied are the multiple model comparisons of similar outbreak scenarios (Dube et. al. 2007, Sanson et. al. 2011) and continuous seeking of expert opinion.

Simulation models of disease in human and animal populations are typically composed of a series of logical processes that allow a response (usually the presence or absence of disease at a given location) to be predicted as a function of a set of defined decision rules (Bagni et. al. 2002). These models can be tactically useful as they follow a logical, biologically valid process that is flexible and can incorporate a high degree of detail (Taylor 2003). Although the flow of logic in disease simulation models tends to be straightforward, attempts to incorporate a high level of detail make it difficult for developers to provide a concise description of a model's overall design to non-technical personnel. This is particularly the case with 'generic' simulation models (i.e. those designed to simulate a range of infectious disease conditions, such as IS+), since these often incorporate settings that may not be directly applicable to a given disease scenario of interest. Thus, a balance needs to be struck between complexity and simplification to ensure that simulation models provide sufficient information about the system under investigation without being so complex that they cannot be widely understood.

The analyses presented in this paper represent the first of a number of steps that may be applied to refine the NZSM. A potentially useful result of this work would be information that informs the grouping of non-influential settings (e.g. the low- and medium-risk movement parameters). This would go some way towards reducing the overall complexity of the NZSM, while still allowing it to remain fit for its purpose. This simplified model would potentially offer greater transparency to decision-makers but retain the benefits of the parent model's complexity. Results from the simplified model could be compared with the fully parameterised version for validation.

Several other possibilities should be considered for further work. In particular, it is important to carry out sensitivity analyses at various times during the simulation; for example, at the time the disease is first detected, then at regular intervals throughout the control and eradication phase of the epidemic. This process would identify how the sensitivity of the model changes during the simulation period, quantifying the way in which prediction precision changes over time and the effect of time on both the values of the PRCC and their relative rankings.

..

Table 9: Details of the ten parameters defining farm-to-farm and farm-to-market movements within the New Zealand Standard Model (NZSM). Also shown are the settings used in the NZSM and candidate settings for the sensitivity analysis.

Parameter	Setting in NZSM	Candidate settings
1. Pastoral livestock, high risk to farm:		
Number per time period	Poisson ($\lambda = 0.03$)	λ = uniform (0, 0.1)
Number of direct contacts	Constant $n = 1$	n = uniform (0, 5)
Probability of transmission	Table (6,11,16; 0.525,0.8,1)[a]	6,11,16; 0.12,0.52,1 6,11,16; 0.25,0.62,1 6,11,16; 0.525,0.8,1 6,11,16; 0.7,0.88,1
2. Dairy, high risk to farm:		
Number per time period	Poisson ($\lambda = 0.042$)	λ = uniform (0, 0.1)
Number of direct contacts	Constant $n = 1$	n = uniform (0, 5)
Probability of transmission	Table (6,11,16; 0.62,0.8,1)	6,11,16; 0.12,0.52,1 6,11,16; 0.25,0.62,1 6,11,16; 0.525,0.8,1 6,11,16; 0.7,0.88,1
3. Dry grazing, high risk to farm:		
Number per time period	Poisson ($\lambda = 0.1152$)	λ = uniform (0,1)
Number of direct contacts	Constant $n = 1$	n = uniform (0, 5)
Probability of transmission	Table (6,11,16; 0.673,0.8,1)	6,11,16; 0.12,0.52,1 6,11,16; 0.25,0.62,1 6,11,16; 0.525,0.8,1 6,11,16; 0.7,0.88,1
4. Pig breeding, high risk to farm:		
Number per time period	Poisson ($\lambda = 0.131$)	λ = uniform (0, 1)
Number of direct contacts	Constant $n = 1$	n = uniform (0, 5)
Probability of transmission	Table (6,11,16; 0.458,0.8,1)	6,11,16; 0.12,0.52,1 6,11,16; 0.25,0.62,1 6,11,16; 0.525,0.8,1 6,11,16; 0.7,0.88,1
5. Medium risk to farm:		
Number per time period	Poisson ($\lambda = 0.4743$)	λ = uniform (0,1)
Number of direct contacts	Constant $n = 1$	n = uniform (0, 5)
Probability of transmission	Constant $n = 0.05$	n = uniform (0, 0.1)
6. Low risk to farm:		
Number per time period	Poisson ($\lambda = 0.0595$)	λ = uniform (0,0.1)
Number of direct contacts	Constant $n = 1$	n = uniform (0, 5)
Probability of transmission	Constant $n = 0.01$	n=uniform (0, 0.1)
7. Pastoral livestock to saleyard:		
Number per time period	Poisson ($\lambda = 0.0135$)	λ = uniform (0, 0.1)

Parameter	Setting in NZSM	Candidate settings
Number of secondary contacts	Poisson (λ = 1.942)	λ = uniform (0, 5)
Probability of transmission	Table (6,11,16; 0.458,0.776,1)	6,11,16; 0.12,0.52,1 6,11,16; 0.25,0.62,1 6,11,16; 0.525,0.8,1 6,11,16; 0.7,0.88,1
8. Dairy to saleyard:		
Number per time period	Poisson (λ = 0.005)	λ = uniform (0, 0.1)
Number of secondary contacts	Poisson (λ = 1.942)	λ = uniform (0, 5)
Probability of transmission	Table (6,11,16; 0.458,0.776,1)	6,11,16; 0.12,0.52,1 6,11,16; 0.25,0.62,1 6,11,16; 0.525,0.8,1 6,11,16; 0.7,0.88,1
9. Dry grazing to saleyard:		
Number per time period	Poisson (λ = 0.003)	λ = uniform (0, 0.01)
Number of secondary contacts	Poisson (λ = 1.942)	λ = uniform (0, 5)
Probability of transmission	Table (6,11,16; 0.458,0.776,1)	6,11,16; 0.12,0.52,1 6,11,16; 0.25,0.62,1 6,11,16; 0.525,0.8,1 6,11,16; 0.7,0.88,1
10. Pig breeding to saleyard:		
Number per time period	Poisson (λ = 0.036)	λ = uniform (0, 0.1)
Number of secondary contacts	Poisson (λ = 1.942)	λ = uniform (0, 5)
Probability of transmission	Table (6, 11, 16; 0.458, 0.776, 1)	6,11,16; 0.12,0.52,1 6,11,16; 0.25,0.62,1 6,11,16; 0.525,0.8,1 6,11,16; 0.7,0.88,1

NZSM: New Zealand standard model
a) Table (6,11,16; 0.525,0.8,1) is interpreted as:

6	11	16
0.525	0.81	1

This specifies the probability that a destination farm will be infected, given the difference in the number of days between the onset of clinical signs on the source farm and the time when the movement occurs. In the above example, if an off-farm movement occurs from an infected farm six days after the onset of clinical signs, the probability that transmission will occur is 0.525

Table 10: Details of the six parameters defining surveillance before and after detection of the outbreak within the New Zealand standard model. Also shown are the settings used in the New Zealand standard model and candidate settings for the sensitivity analysis.

Parameter	Setting in NZSM	Candidate settings
1. Background surveillance:		
All farm types selection probability	Constant $n = 1$	n = uniform (0, 1)
Pastoral livestock detection probability	Table[(a)]	
Dairy detection probability	Table[(a)]	
Dry grazing detection probability	Table[(a)]	
Pig detection probability	Table[(a)]	
2. Self report surveillance:		
All farm types selection probability	Constant $n = 1$	n = uniform (0, 1)
Pastoral livestock detection probability	Table[(a)]	
Dairy detection probability	Table[(a)]	
Dry grazing detection probability	Table[(a)]	
Pig detection probability	Table[(a)]	
3. Surveillance following HR contact:		
All farm types selection probability	Constant $n = 1$	n = uniform (0, 1)
Beef cattle detection probability	Constant $n = 1$	n = uniform (0, 1)
Dairy cattle detection probability	Constant $n = 1$	n = uniform (0, 1)
Deer detection probability	Constant $n = 1$	n = uniform (0, 1)
Goats detection probability	Constant $n = 1$	n = uniform (0, 1)
Pigs detection probability	Constant $n = 1$	n = uniform (0, 1)
Sheep detection probability	Constant $n = 1$	n = uniform (0, 1)
4. Surveillance following MR contact:		
All farm types selection probability	Constant $n = 0.9$	n = uniform (0, 1)
Beef cattle detection probability	Constant $n = 1$	n = uniform (0, 1)
Dairy cattle detection probability	Constant $n = 1$	n = uniform (0, 1)
Deer detection probability	Constant $n = 1$	n = uniform (0, 1)
Goats detection probability	Logistic (0.25 0.8 0.74 1.7)[(b)]	0.25, 0.2, 0.74, 1.7 0.25, 0.4, 0.74, 1.7 0.25, 0.6, 0.74, 1.7 0.25, 0.8, 0.74, 1.7 0.25, 1.0, 0.74, 1.7
Pigs detection probability	Constant $n = 1$	n = uniform (0, 1)
Sheep detection probability	Logistic (0.25 0.8 0.74 1.7)[(b)]	0.25, 0.2, 0.74, 1.7 0.25, 0.4, 0.74, 1.7 0.25, 0.6, 0.74, 1.7 0.25, 0.8, 0.74, 1.7 0.25, 1.0, 0.74, 1.7
5. Surveillance following LR contact:		
All farm types selection probability	Constant $n = 0.5$	n = uniform (0, 1)

Beef cattle detection probability	Constant $n = 1$	n = uniform (0, 1)
Dairy cattle detection probability	Constant $n = 1$	n = uniform (0, 1)
Deer detection probability	Constant $n = 1$	n = uniform (0, 1)
Goats detection probability	Logistic (0.25, 0.8, 0.74, 1.7)[(b)]	0.25, 0.2, 0.74, 1.7 0.25, 0.4, 0.74, 1.7 0.25, 0.6, 0.74, 1.7 0.25, 0.8, 0.74, 1.7 0.25, 1.0, 0.74, 1.7
Pigs detection probability	Constant $n = 1$	n = uniform (0, 1)
Sheep detection probability	Logistic (0.25, 0.8, 0.74, 1.7)[(b)]	0.25, 0.2, 0.74, 1.7 0.25, 0.4, 0.74, 1.7 0.25, 0.6, 0.74, 1.7 0.25, 0.8, 0.74, 1.7 0.25, 1.0, 0.74, 1.7
6. Surveillance following patrol visit:		
All farm types selection probability	Constant $n = 1$	n = uniform (0, 1)
Beef cattle detection probability	Constant $n = 1$	n = uniform (0, 1)
Dairy cattle detection probability	Constant $n = 1$	n = uniform (0, 1)
Deer detection probability	Constant $n = 1$	n = uniform (0, 1)
Goats detection probability	Logistic (0.25, 0.8, 0.74, 1.7)[(b)]	0.25, 0.2, 0.74, 1.7 0.25, 0.4, 0.74, 1.7 0.25, 0.6, 0.74, 1.7 0.25, 0.8, 0.74, 1.7 0.25, 1.0, 0.74, 1.7
Pigs detection probability	Constant $n = 1$	n = uniform (0, 1)
Sheep detection probability	Logistic (0.25, 0.8, 0.74, 1.7)[(b)]	0.25, 0.2, 0.74, 1.7 0.25, 0.4, 0.74, 1.7 0.25, 0.6, 0.74, 1.7 0.25, 0.8, 0.74, 1.7 0.25, 1.0, 0.74, 1.7

a) See Figure 7 for details

b) 

HR: high risk
LR: low risk
MR: medium risk
NZSM: New Zealand standard model

Table 11: Details of the eight parameters defining tracing efficiency within the New Zealand standard model. Also shown are the settings used in the New Zealand standard model and candidate settings for the sensitivity analysis.

Parameter	Setting in NZSM	Candidate settings
1. Pastoral livestock, high risk:		
Probability of forgetting a movement off the property	Constant $n = 0.11$	n = uniform (0, 1)
Delays in tracing a movement off the property	Table (0.5,1; 0,1) [(a)]	0.5, 1; 0, 1 0.25, 0.75, 1.00; 0, 1, 2 0.25, 0.50, 0.75, 1; 0, 1, 2, 3
Probability of forgetting a movement onto the property	Constant $n = 0.082$	n = uniform (0, 1)
Delays in tracing a movement onto the property	Table (0.5, 1; 0, 1)	0.5, 1; 0, 1 0.25, 0.75,1.00; 0, 1, 2 0.25, 0.50, 0.75, 1; 0, 1, 2, 3
2. Dairy high risk:		
Probability of forgetting a movement off the property	Constant $n = 0.11$	n = uniform (0, 1)
Delays in tracing a movement off the property	Table (0.5, 1; 0, 1)	0.5, 1; 0, 1 0.25, 0.75, 1.00; 0, 1, 2 0.25, 0.50, 0.75, 1; 0, 1, 2, 3
Probability of forgetting a movement onto the property	Constant $n = 0.082$	n = uniform (0, 1)
Delays in tracing a movement onto the property	Table (0.5, 1; 0, 1)	0.5, 1; 0, 1 0.25, 0.75, 1.00; 0, 1, 2 0.25, 0.50, 0.75, 1; 0, 1, 2, 3
3. Dry grazing:		
Probability of forgetting a movement off the property	Constant $n = 0.11$	n = uniform (0, 1)
Delays in tracing a movement off the property	Table (0.5, 1; 0, 1)	0.5, 1; 0, 1 0.25, 0.75, 1.00; 0, 1, 2 0.25, 0.50, 0.75, 1; 0, 1, 2, 3
Probability of forgetting a movement onto the property	Constant $n = 0.082$	n = uniform (0, 1)
Delays in tracing a movement onto the property	Table (0.5,1; 0,1)	0.5, 1; 0, 1 0.25, 0.75, 1.00; 0, 1, 2 0.25, 0.50, 0.75, 1; 0, 1, 2, 3
4. Pig breeding:		
Probability of forgetting a movement off the property	Constant $n = 0.11$	n = uniform (0, 1)
Delays in tracing a movement off the property	Table (0.5, 1; 0, 1)	0.5, 1; 0, 1 0.25, 0.75, 1.00; 0, 1, 2

		0.25, 0.50, 0.75, 1; 0, 1, 2, 3
Probability of forgetting a movement onto the property	Constant n = 0.082	n = uniform (0, 1)
Delays in tracing a movement onto the property	Table (0.5, 1; 0, 1)	0.5, 1; 0, 1 0.25, 0.75, 1.00; 0, 1, 2 0.25, 0.50, 0.75, 1; 0, 1, 2, 3
5. Medium risk:		
Probability of forgetting a movement off the property	Constant n = 0.212	n = uniform (0, 1)
Delays in tracing a movement off the property	BetaPert (a = 1, b = 2, c = 3)	b = uniform (1, 3)
Probability of forgetting a movement onto the property	Constant n = 0.194	n = uniform (0, 1)
Delays in tracing a movement onto the property	BetaPert (a = 1, b = 2, c = 3)	b = uniform (1, 3)
6. Low risk:		
Probability of forgetting a movement off the property	Constant n = 0.36	n = uniform (0, 1)
Delays in tracing a movement off the property	BetaPert (a = 2, b = 3, c = 4)	b = uniform (2, 4)
7. Dairy tanker:		
Probability of forgetting a movement off the property	Constant n = 0.014	n = uniform (0, 1)
Delays in tracing a movement off the property	Table (0.5, 1; 0, 1)	0.5, 1; 0, 1 0.25, 0.75, 1.00; 0, 1, 2 0.25, 0.50, 0.75, 1; 0, 1, 2, 3
Probability of forgetting a movement onto the property	Constant n = 0.014	n = uniform (0, 1)
Delays in tracing a movement onto the property	Table (0.5, 1; 0, 1)	0.5, 1; 0, 1 0.25, 0.75, 1.00; 0, 1, 2 0.25, 0.50, 0.75, 1; 0, 1, 2, 3
8. Saleyard, high risk:		
Probability of forgetting a movement off the property	Constant n = 0.063	n = uniform (0, 1)
Delays in tracing a movement off the property	Table (0.5, 1; 0, 1)	0.5, 1; 0, 1 0.25, 0.75, 1.00; 0, 1, 2 0.25, 0.50, 0.75, 1; 0, 1, 2, 3
Probability of forgetting a movement onto the property	Constant n = 0.058	n = uniform (0, 1)
Delays in tracing a movement onto the property	Table (0.5, 1; 0, 1)	0.5, 1; 0, 1 0.25, 0.75, 1.00; 0, 1, 2 0.25, 0.50, 0.75, 1; 0, 1, 2, 3

a) Table (0.5, 1; 0, 1) is interpreted as:

0.5	1
0	1

NZSM: New Zealand standard model
This specifies the number of time periods it takes to trace the specified movement type in the specified direction. In the above example, 50% of movements will be traced on the same day as the day of detection and 100% will be traced within one day of detection

Table 12: Details of the three parameters defining the efficacy of movement restrictions and the single parameter defining resources available for depopulation within the New Zealand standard model. Also shown are the settings used in the New Zealand standard model and candidate settings for the sensitivity analysis.

Parameter	Setting in NZSM	Candidate settings
1. Probability restriction HR movements:		
Initial standstill	Constant $n = 0.914$	n = uniform (0, 1)
Inside infected zone	Constant $n = 0.942$	n = uniform (0, 1)
Inside surveillance zone	Constant $n = 0.951$	n = uniform (0, 1)
Outside control area	Constant $n = 0.951$	n = uniform (0, 1)
2. Probability restriction MR movements:		
Initial standstill	Constant $n = 0.604$	n = uniform (0, 1)
Inside infected zone	Constant $n = 0.804$	n = uniform (0, 1)
Inside surveillance zone	Constant $n = 0.850$	n = uniform (0, 1)
Outside control area	Constant $n = 0.850$	n = uniform (0, 1)
3. Probability restriction LR movements:		
Initial standstill	Constant $n = 0.238$	n = uniform (0, 1)
Inside infected zone	Constant $n = 0.390$	n = uniform (0, 1)
Inside surveillance zone	Constant $n = 0.520$	n = uniform (0, 1)
Outside control area	Constant $n = 0.520$	n = uniform (0, 1)
Number of farms depopulated per team per time period:		
Pastoral livestock	Triangular ($a = 0$, $b = 0$, $c = 5$)	b = uniform (0, 5)
Dairy	Triangular ($a = 0$, $b = 1$, $c = 3$)	b = uniform (0, 3)
Dry grazing	Triangular ($a = 0$, $b = 0$, $c = 3$)	b = uniform (0, 3)
Pig	Triangular ($a = 0$, $b = 0$, $c = 3$)	b = uniform (0, 3)

HR: high risk
MR: medium risk
LR: low risk

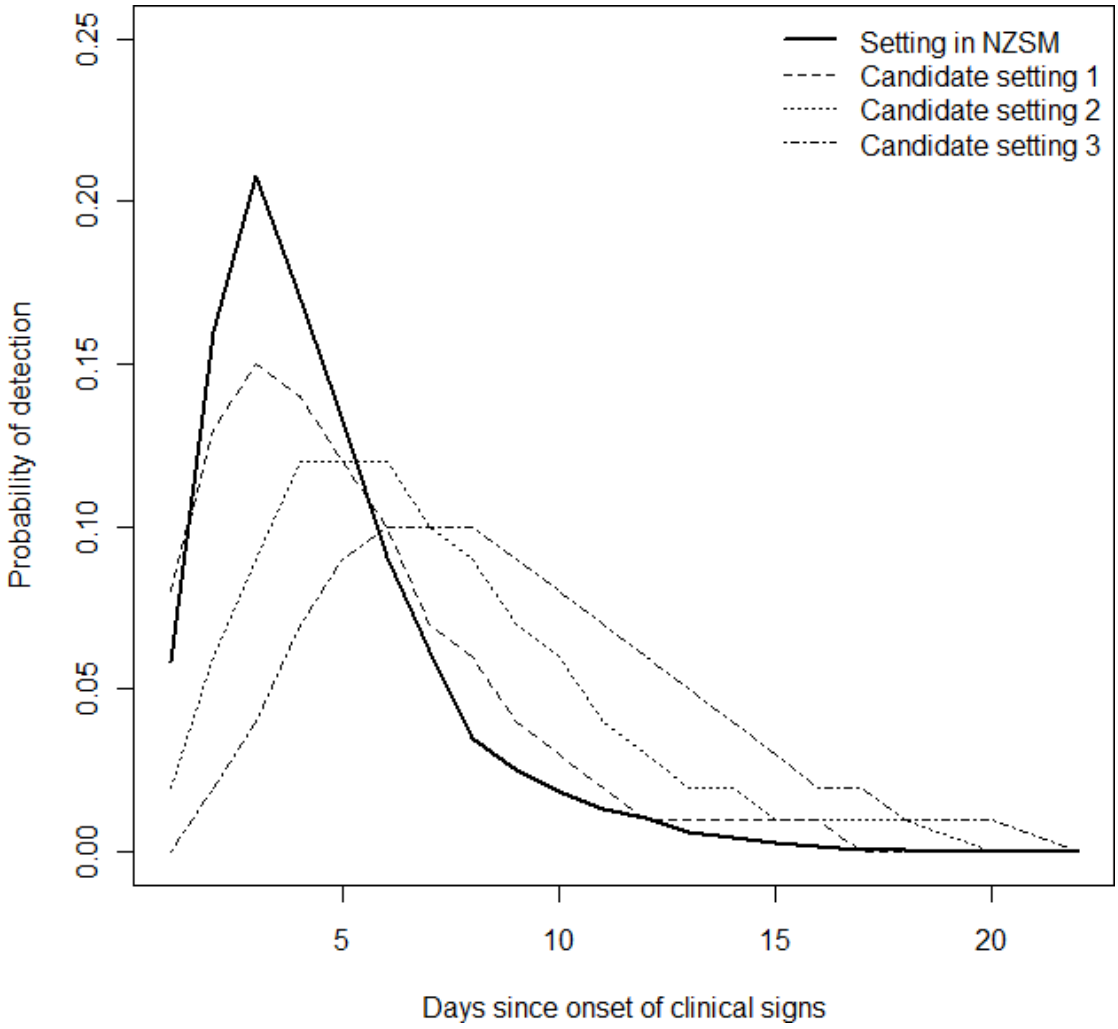
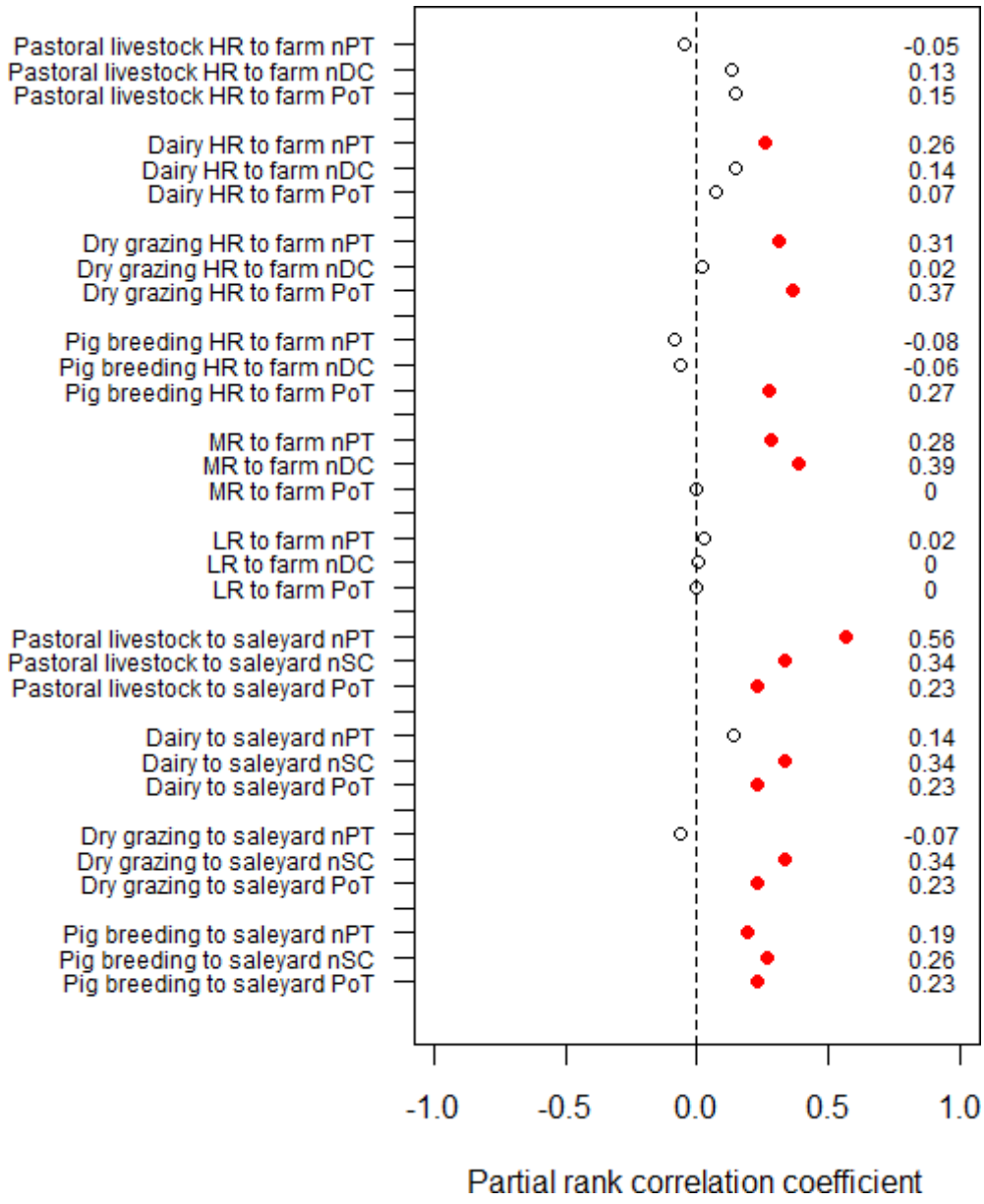
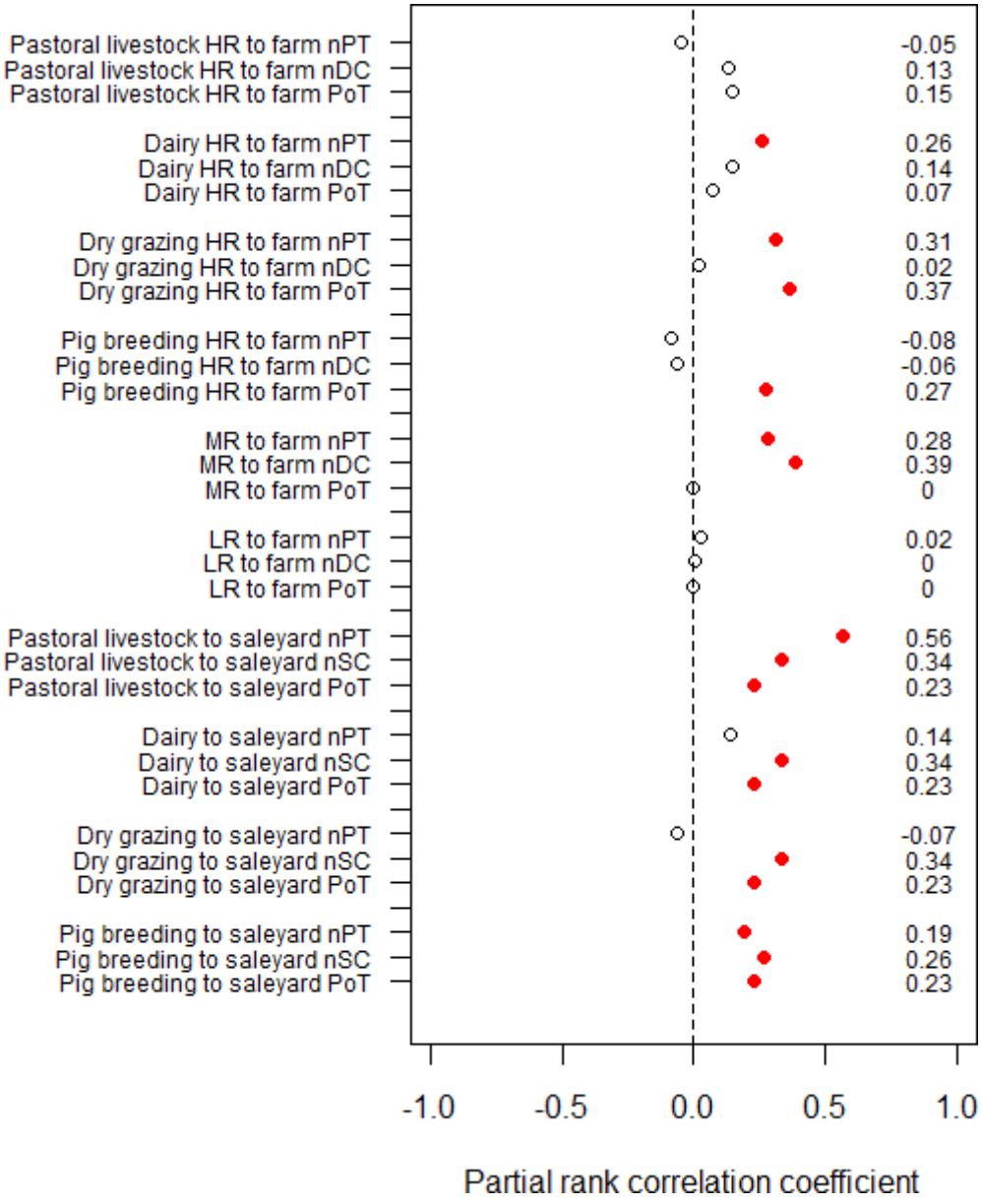


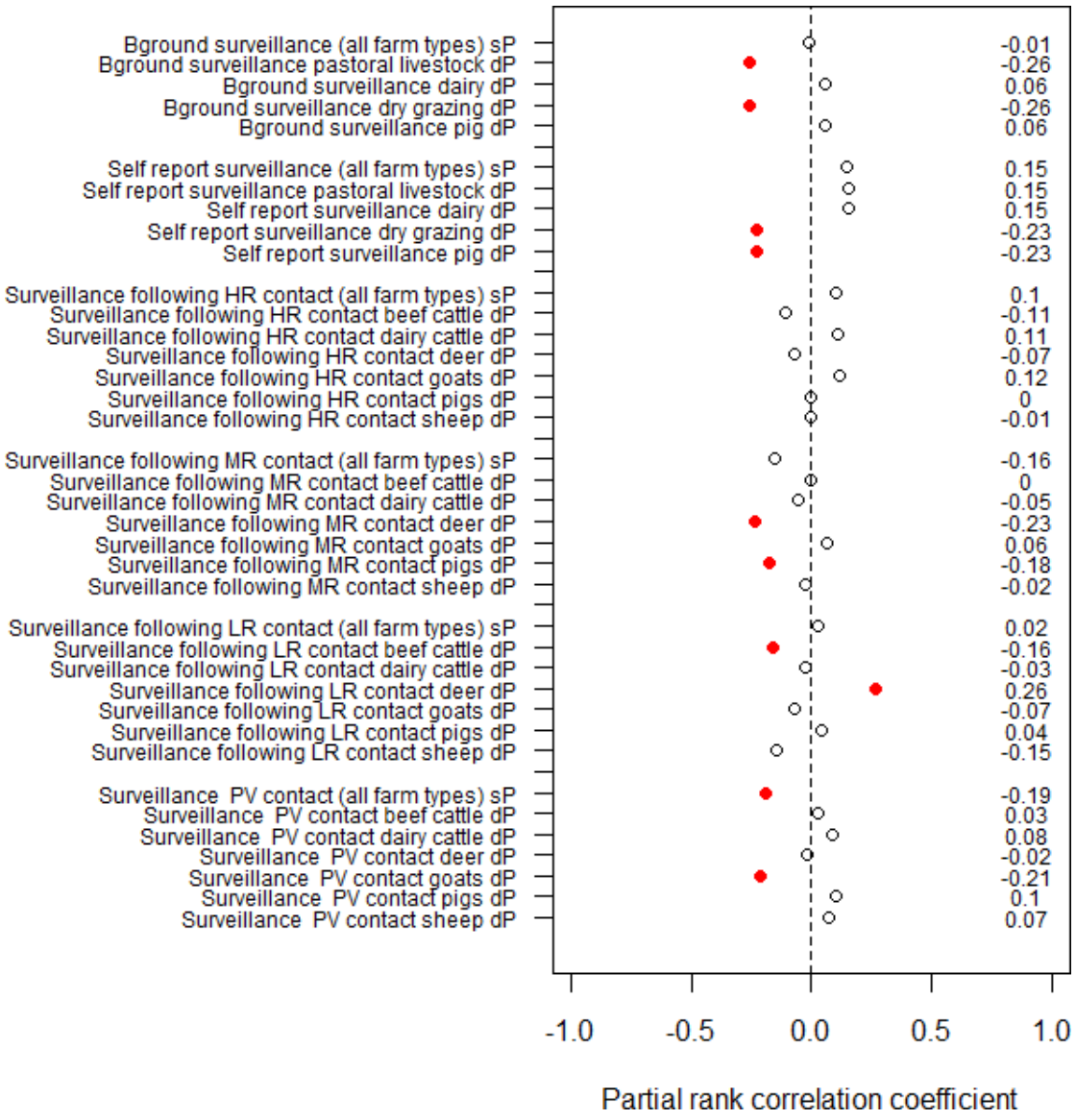
Figure 7: **Line plot showing the probability that a pig or dairy farm will be detected as positive for the disease, as a function of the number of days since the onset of clinical signs. The solid line shows the settings used in the New Zealand standard model (NZSM). The dashed lines show the three candidate settings used in the sensitivity analyses. The candidate settings represent increasingly lower likelihood of detection but increasingly longer times for detection to occur.**


Pastoral livestock HR to farm nPT
Pastoral livestock HR to farm nDC
Pastoral livestock HR to farm PoT
Dairy HR to farm nPT
Dairy HR to farm nDC
Dairy HR to farm PoT
Dry grazing HR to farm nPT
Dry grazing HR to farm nDC
Dry grazing HR to farm PoT
Pig breeding HR to farm nPT
Pig breeding HR to farm nDC
Pig breeding HR to farm PoT
MR to farm nPT
MR to farm nDC
MR to farm PoT
LR to farm nPT
LR to farm nDC
LR to farm PoT
Pastoral livestock to saleyard nPT
Pastoral livestock to saleyard nSC
Pastoral livestock to saleyard PoT
Dairy to saleyard nPT
Dairy to saleyard nSC
Dairy to saleyard PoT
Dry grazing to saleyard nPT
Dry grazing to saleyard nSC
Dry grazing to saleyard PoT
Pig breeding to saleyard nPT
Pig breeding to saleyard nSC
Pig breeding to saleyard PoT
-0.05
0.13
0.15
0.26
0.14
0.07
0.31
0.02
0.37
-0.08
-0.06
0.27
0.28
0.39
0
0.02
0
0
0.56
0.34
0.23
0.14
0.34
0.23
-0.07
0.34
0.23
0.19
0.26
0.23
-1.0
-0.5
0.0
0.5
1.0
Partial rank correlation coefficient



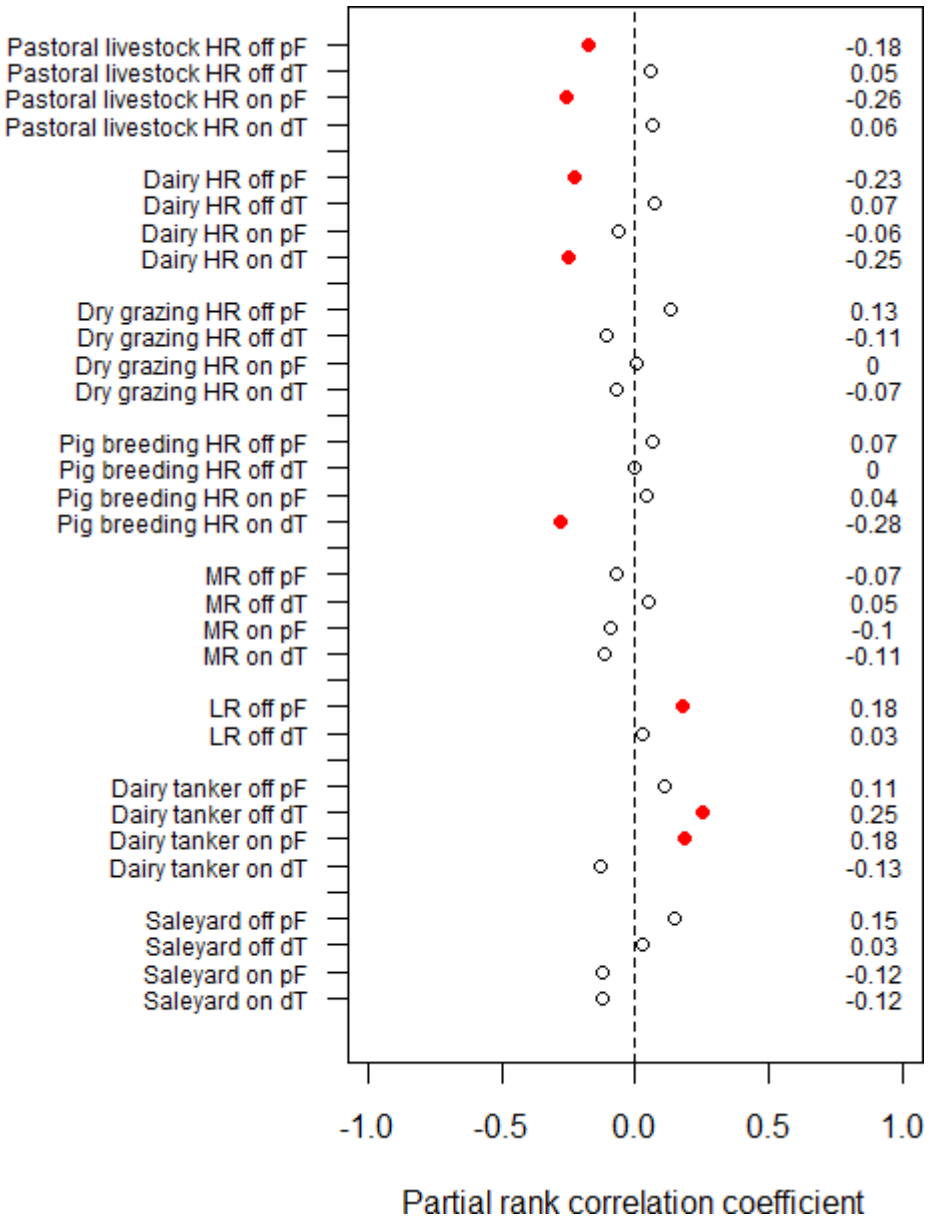
HR: high risk
nPT: number per time period
nDC: number of direct contacts
PoT: probability of transmission
MR: medium risk
LR: low risk
nSC: number of secondary contacts

Figure 8: Partial rank correlation coefficients for settings within the ten parameters defining farm-to-farm and farm-to-saleyard movements within the New Zealand standard model. Solid circles (•) identify settings whose partial rank correlation coefficient values were significant at the alpha level of 0.05.



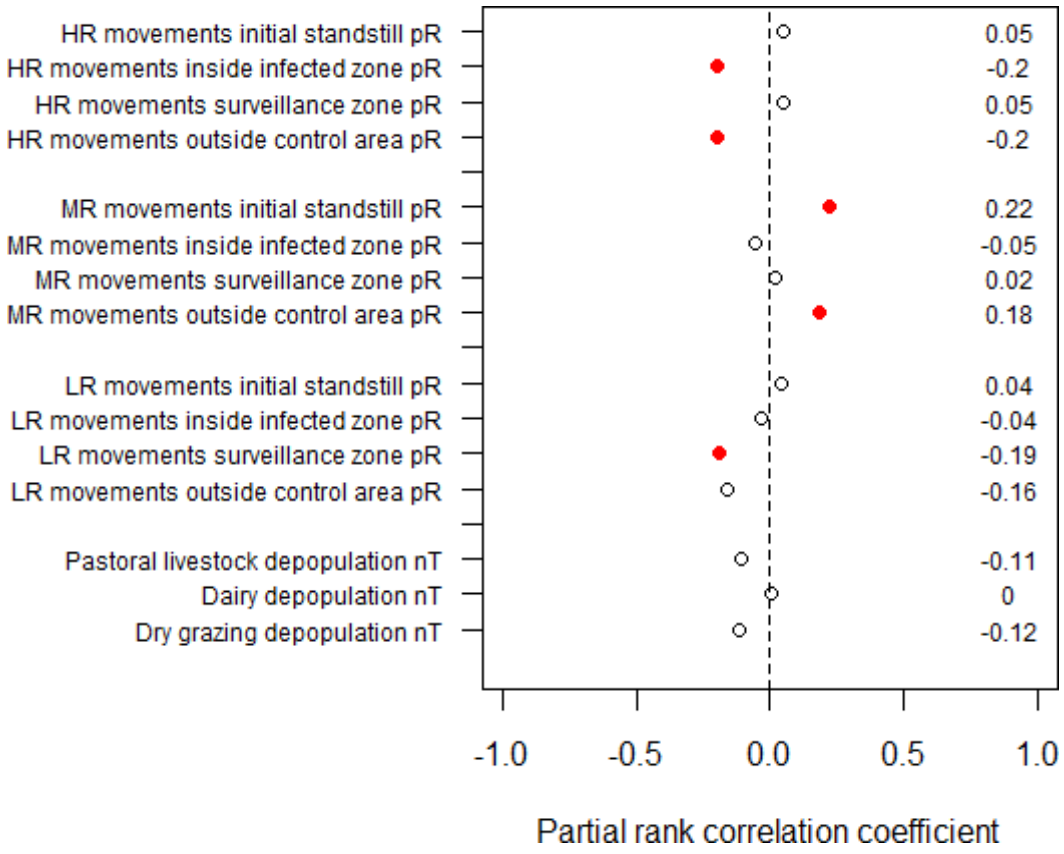
Bground: background
sP: selection probability
dP: detection probability
HR: high risk
MR: medium risk
LR: low risk
PV: patrol visit

Figure 9: Partial rank correlation coefficients for settings within the six parameters defining surveillance within the New Zealand standard model. Solid circles (•) identify those settings whose partial rank correlation coefficient values were significant at the alpha level of 0.05.



HR: high risk
pF: probability of forgetting an on- or off-farm movement
dT: tracing delay (days)
MR: medium risk
LR: low risk

Figure 10: Partial rank correlation coefficients for settings within the eight parameters defininig efficacy of tracing within the New Zealand standard model. Solid circles (•) identify those settings whose partial rank correlation coefficient values were significant at the alpha level of 0.05.



HR: high risk
pR: probability of restriction
MR: medium risk
LR: low risk
nT: number of herds able to be processed per time period (days)

Figure 11: Partial rank correlation coefficients for settings within the three parameters defining efficacy of movement restrictions, and the single parameter defining the resources required for depopulation, within the New Zealand standard model. Solid circles (•) identify those settings whose partial rank correlation coefficient values were significant at the alpha level of 0.05.

Chapter 5: Looking to the future

Introduction

Ultimately the goal for modelling in New Zealand is to best manage an incursion of exotic disease by better understanding how the disease is likely to behave in the population. Policy decisions in novel scenarios are improved by increasing the value, utilisation and understanding of the models available.

The outcomes of the sensitivity analysis have already increased the understanding of the model itself and have provided initial insight into ways the model may be improved and refined. If heeded, this can only result in an increased ability to interpret and communicate model outputs. This, in turn, increases confidence that the NZSM parameter set provides an appropriate indication of the way FMD might spread if it were introduced into the farm animal population in New Zealand, and therefore better preparation for any future FMD outbreak. The initial work described can be built upon by:

- Improving the sensitivity analysis techniques and their application.
- Improving the model using the results of the analysis: decreasing complexity and increasing investment in the areas with greater influence.
- Developing and applying the principles of good practice in model development to include appropriate sensitivity analysis as well as other methods for model corroboration.
- Utilising sensitivity analysis methods as optimisation tools.

Improvement in sensitivity analysis techniques

The complexity of epidemiological models and the number of parameters mean that sensitivity analysis is computationally demanding. These characteristics severely limit the choice and application of the techniques and therefore its usefulness. Ideally sensitivity analysis would be

hard-wired into the model program generating outputs that can be examined as the model runs, producing real time indication of the key drivers of variation in model output.

Currently, few alternative approaches are available for such complex stochastic models, but possibilities do exist such as the elementary effects approach. Techniques applied together to the same model are likely to provide a more detailed picture of how the parameters interact and contribute to model output uncertainty, and deeper insight into the model's behaviour.

More techniques may become available and it is important to continue to follow advances in disciplines such as hydrology, environmental and meteorological modelling and apply these techniques or combinations to InterSpread+ derivatives.

It is likely that different parameters are influential at different outbreak stages. Whilst real-time sensitivity analysis may be very much in the future, an intermediate step could be to carry out sensitivity analyses at different times following simulation start, for example at the time the disease is first detected then at regular intervals throughout the control and eradication phase of the epidemic. This process would identify how the sensitivity of the model changes during the simulation period and provide a more accurate picture of influence.

Model improvement

Model improvement is a continuous process and requires ongoing and justified investment in research to generate data to improve the parameters as well as to incorporate data from sources such as NAIT which improve understanding of the livestock population and movement. The results of the sensitivity analysis help to justify the prioritisation of our limited resources to research or data management for the parameters of most influence such as movements.

Counterintuitive results in this preliminary sensitivity analysis may be a factor of the difficulty in applying a sensitivity analysis to such a complex model. However, these findings may lead to discovery of errors in the model and effort must focus on exploring the reasons for these results.

Ultimately different models of different complexity should be applied in different settings. Broad brush analysis of a control strategy to communicate issues and alternatives may be more effective if a model is simpler. A very detailed model is likely to be most appropriate to inform resourcing questions or exploration of specific control issues. The models we have available currently are complex epidemiological models with the associated advantages and disadvantages (see Chapter 1). The analyses presented in this paper represent the first of a number of steps that

may be applied to refine the NZSM, reducing the overall complexity of the model while still allowing it to remain fit-for-purpose. This simplified model would potentially offer greater transparency to decision makers but would retain benefits of the parent model's complexity. Results from the simplified model could be compared with the fully parameterised version for validation.

Continued development and implementation of good model practice in New Zealand

It is important to note that a sensitivity analysis does not stand on its own. It should be coupled with other techniques to corroborate a model for its intended purpose. Also, the sensitivity analysis itself is an iterative process and the conclusions from one round of analysis are likely to inform a new iteration of the analysis, whether with the same or a different technique.

In good practice for model development, such as described by Jakeman et al. (2006) uncertainty and sensitivity analysis should be used iteratively during model development, before the model is used, and when the model is used. Figure 10 illustrates a possible process for model development using the InterSpread Plus platform using elements of good practice for model development. Techniques such as multiple model comparison are already utilised in the New Zealand, for example in the QUADS comparison projects described by Dube et al. (2007) and Sanson et. al. (2011).

Considering the benefits of each type of uncertainty assessment it can be seen that these methods are relevant to the development of a stochastic, spatially explicit simulation model such as those based on InterSpread Plus. The development process in Figure 10 could be appropriate for models developed within the InterSpread Plus framework: in this process the New Zealand Standard Model is an example of a disease specific model. The figure attempts to capture the interactions between the key players (the policy maker) and the modelling group (epidemiologists, modellers and programmers) and the techniques that can be applied at each stage.

The process is intended to suggest a structured approach that enables understanding by all parties of the state of confidence in any model. The process is initiated from a disease management problem where evidence is needed to explore options for control or eradication of an organism. Although this point of identification is placed with the policy maker, it could equally be signalled by a modeller. The policy maker and the modelling group would then

consider the problem and determine if InterSpread Plus may provide a suitable platform for model scenarios to explore the issue.

New applications of the techniques

Exploration of scenarios to optimise control methods

Sensitivity analysis techniques could be applied as optimisation tools, providing an effective way to explore scenarios for control of animal diseases.

The question of how to manage an epidemic to minimise cost is extremely complex. A traditional method of trial and error requires a very large number of scenarios which is impractical and computationally demanding. A possibility is to treat the problem as a linear programming issue[3] using a sensitivity analysis approach to vary appropriate parameters simultaneously. A logistic regression model could then be developed to guide the combination of control measures that has the biggest influence on the outputs of interest. This technique may be applied in New Zealand in the near future to explore vaccination scenarios.

Conclusion

Sensitivity analysis is a tool in our toolkit to increase knowledge of how useful our model is: increasing confidence in model outputs, and ultimately disease management planning for exotic diseases. Even the initial work described in Chapter 4 has increased the understanding of the NZSM and has provided insight into ways the model may be improved and refined. If heeded, this can only result in an increased ability to interpret and communicate model outputs. This, in turn, increases confidence that the NZSM parameter set provides an appropriate indication of the way FMD might spread if it were introduced into the farm animal population in New Zealand, and therefore better preparation for any future FMD outbreak.

Like any tool, it is better if the right one is purchased and it is maintained, used regularly and upgraded as necessary rather than sitting, slowly rusting in the tool box in the shed. It is important to continue work in this area.

[3] Linear programming is a mathematical method for determining a way to achieve the best outcome (such as maximum profit or lowest cost) in a given mathematical model for some list of requirements represented as linear relationships.

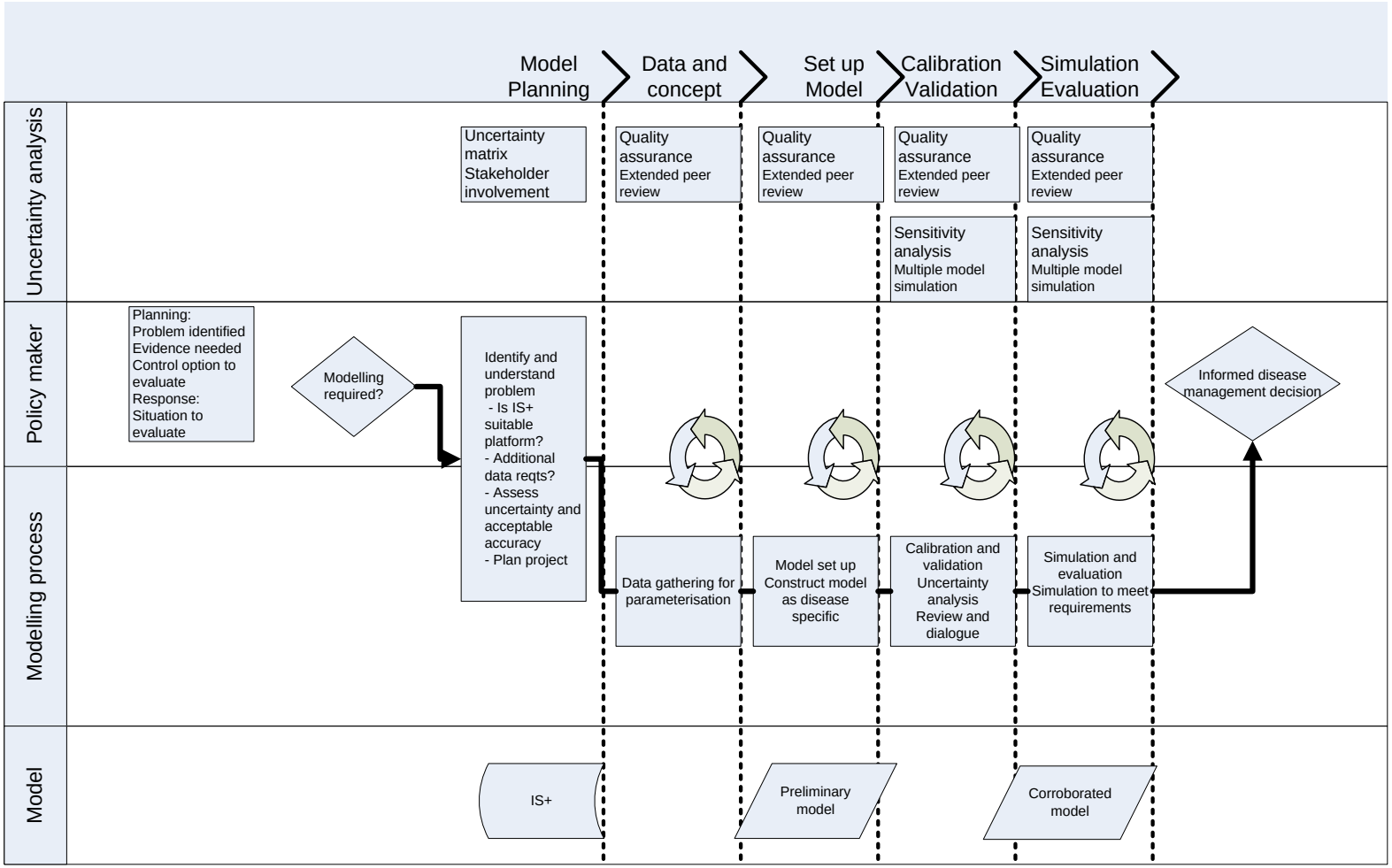


Figure 12: Proposed process for model development for disease management decisions using InterSpread Plus. Interactions between the modeller and policy maker during model development are depicted by circular arrows.

..

References

AMLS2 Contacts Database Application. n.d. Department for Environment, Food and Rural Affairs, UK. Accessed online 2/11/2011. http://amls.defra.gov.uk/

Antonio C C & Hoffbauer L. 2008. From local to global importance measures of uncertainty propagation in composite structures. Composite Structures, 85(3), 213-225.

Bagni R, Berchi R & Cariello P. 2002. A comparison of simulation models applied to epidemics. Journal of Artificial Societies and Social Simulation, 5 (3) Accessed online 2/8/2010. http://jasss.soc.surrey.ac.uk/5/3/5.html

Bar Massada A& Carmel Y. 200. Incorporating output variance in local sensitivity analysis for stochastic models. Ecological Modelling, 213(3-4), 463-467

Bates TW, Carpenter TE, Thurmond MC: 2003, Results of epidemic simulation modeling to evaluate strategies to control an outbreak of foot-and-mouth disease. American Journal of Veterinary Research 64, 205–210

Beyer H G & Sendhoff B. 2007. Robust optimization - A comprehensive survey. Computer Methods in Applied Mechanics and Engineering, 196(33-34), 3190-3218

Bishop C H & Shanley K T. 2008. Bayesian Model Averaging's problematic treatment of extreme weather and a paradigm shift that fixes it. Monthly Weather Review 136, 4641–46.

Blower S M & Dowlatabadi H.1994. Sensitivity and uncertainty analysis of complex-models of disease transmission - an hiv model, as an example International Statistical Review, 62(2), 229-243

Box G E P, Draper N R.1987. Empirical Model-Building and Response Surfaces, John Wiley & Sons Ltd, New York.

Britton T & Lindenstrand D. 2009. Epidemic modelling: Aspects where stochasticity matters. Mathematical Biosciences, 222(2), 109-116.

Butts M B, Payne J T, Kristensen M, Madsen H. 2004. An evaluation of the impact of model structure on hydrological modelling uncertainty for streamflow simulation. Journal of Hydrology, 298 (2004) 242–266

Cacuci D G & Ionescu-Bujor. 2005. Deterministic local sensitivity analysis of augmented systems - I: Theory Nuclear Science and Engineering, 151(1), 55-66.

Campolongo F , Saltelli A, Sorensen T & Tarantola Y. 2000. Hitchhiker's guide to sensitivity analysis. In Saltelli A, Chan K, & Scott M (Eds.). Sensitivity Analysis. John Wiley & Sons Ltd, New York.

Campolongo F, Cariboni, J & Saltelli A. 2007. An effective screening design for sensitivity analysis of large models. Environmental Modelling & Software, 22(10), 1509-1518.

Cariboni J, Gatelli D, Liska R & Saltelli A. 2007. The role of sensitivity analysis in ecological modelling. Ecological Modelling, 203(1-2), 167-182

Caro J J. 2000. Disease-simulation models and health care decisions. Canadian Medical Association Journal, 162(7), 1001-1002.

Carpenter TE, O'Brien JM, Hagerman AD and McCarl BA 2011. Epidemic and economic impacts of delayed detection of foot-and-mouth disease: A case study of a simulated outbreak in California. Journal of Veterinary Diagnostic Investigation. Accessed online 22/8/2013 http://vdi.sagepub.com/content/23/1/26

Carson J S. 2005. Introduction to modelling and simulation. In Proceedings of the 37th Winter Simulation Conference, December 4-11th, Orlando, Florida.

Conover W J. 1980. Practical Nonparametric Statistics (2nd ed), John Wiley & Sons Ltd, New York.

CREM- The Council for Regulatory Environmental Modeling in the US EPA (US Environmental Protection Agency). 2009. Guidance on the Development, Evaluation, and Application of Environmental Models. EPA/100/K-09, Environmental Protection Agency, Washington, DC: USA.

Crosetto M, Moreno Ruiz J A & Crippa B. 2001. Uncertainty propagation in models driven by remotely sensed data. Remote Sensing of Environment, 76(3), 373-385.

Dangerfield C E, Ross J V & Keeling M J. 2009. Integrating stochasticity and network structure into an epidemic model. Journal of the Royal Society Interface, 6(38), 761-774.

De Pauw DJW, Steppe K & De Baets B. 2008. Unravelling the output uncertainty of a tree water flow and storage model using several global sensitivity analysis methods. Biosystems Engineering, 101(1), 87-99.

Defra: FMD Data Archive. 2004. Accessed online 10/11/2011. http://footandmouth.csl.gov.uk/

Dietz K. 1966. On the model of Weiss for the spread of epidemics by carriers. Journal of Applied Probability, 3 (2), 375-382.

Droguett E L & Mosleh A. 2008. Bayesian methodology for model uncertainty using model performance data. Risk Analysis, 28(5), 1457-1476.

Dubé C, Stevenson M A, Garner M G, Sanson R L, Corso B A, Harvey N, Griffin J, Wilesmith J W & Estrada C. 2007. A comparison of predictions made by three simulation models of foot-and-mouth disease. New Zealand Veterinary Journal, 55 (6), 280-288.

French N R & White PC L. 2004. The use of GIS in modelling the spatial and temporal spread of animal diseases. In Durr P, Gatrell A C (Eds.), GIS and Spatial Analysis in Veterinary Science. CABI Publishing, Oxford.

Frey H C & Patil S R. 2002. Identification and review of sensitivity analysis methods. Risk Analysis, 22(3), 553-578.

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

Garner M G, Dube C, Stevenson M, Sanson R, Estrada C & Griffin J. 2007. Evaluating alternative approaches to managing animal disease outbreaks – the role of modelling in policy formulation. Veterinaria Italiana 43(2), 285-298.

Gibbens J C , Sharpe C E , Wilesmith J W, Mansley L M, Michalopoulou E, Ryan J B & Hudson M. 2001. Descriptive epidemiology of the 2001 foot-and-mouth disease epidemic in Great Britain: the first five months. Vet. Rec., 149 (24), 729-743.

Green L E & Medley G F. 2002. Mathematical modelling of the foot-and-mouth disease epidemic of 2001: Strengths and weaknesses. Research in Veterinary Science, 73(3), 201-205.

Guitian J & Pfeiffer D. 2006. Should we use models to inform policy development? The Veterinary Journal. 172 (3), 393-395.

Harvey N, Reeves A, Schoenbaum M A, Zagmutt-Vergara F J, Dube C, Hill A E et. al. 2007. The North American animal disease spread model: A simulation model to assist decision making in evaluating animal disease incursions. Preventive Veterinary Medicine, 82(3-4), 176-197.

Helton J C, Davis F J. 2000. Sampling Based Methods. In Saltelli A, Chan K, & Scott M (Eds.). Sensitivity Analysis. John Wiley & Sons Ltd, New York.

Helton J C & Davis F J. 2003. Latin Hypercube sampling and the propagation of uncertainty in analyses of complex systems Reliability Engineering & System Safety, 81(1), 23-69.

Helton J C, Davis F J & Johnson J D. 2005. A comparison of uncertainty and sensitivity analysis results obtained with random and Latin Hypercube sampling. Reliability Engineering & System Safety, 89(3), 305-330.

Helton J C, Johnson J D, Sallaberry C J & Storlie C B. 2006. Survey of sampling-based methods for uncertainty and sensitivity analysis. Reliability Engineering & System Safety, 91(10-11), 1175-1209.

Heuer C, French N P, Jackson R, & Mackereth G F. 2007. Application of modelling to determine the absence of foot-and-mouth disease in the face of a suspected incursion. New Zealand Veterinary Journal, 55(6), 289-296.

Iman R L & Conover W J. 1980. Small sample sensitivity analysis techniques for computer models, with an application to risk assessment. Communication in Statistics -Theory and Methods, 9 (17), 1749-1842.

Iman R L & Helton J C. 1988. An investigation of uncertainty and sensitivity analysis techniques for computer models. Risk Analysis, 8 (1), 71-90.

Iman R L, Helton J C & Campbell J E. 1981. An approach to sensitivity analysis of computer models: part 1 – introduction, input variable selection and preliminary variable assessment. Journal of Quality Technology, 13 (3), 174-183.

Iooss B, Van Dorpe F & Devictor N. 2006. Response surfaces and sensitivity analyses for an environmental model of dose calculations. Reliability Engineering & System Safety, 91(10-11), 1241-1251.

Iman R L & Hora S C. 1990. A robust measure of uncertainty importance for use in fault tree system analysis. Risk Analysis, 10(3): 401-406.

Ishigami T, Homma, T. 1990. An importance quantification technique in uncertainty analysis for computer models. In Proceedings of the1st International Symposium on Uncertainty Modelling and Analysis (ISUMA), 398-403

Jakeman A J, Letcher R A & Norton J P. 2006. Ten iterative steps in development and evaluation of environmental models. Environmental Modelling and Software, 21 (5), 602-614.

Kao R R. 2002. The role of mathematical modelling in the control of the 2001 FMD epidemic in the UK Trends in Microbiology, 10(6), 279-286.

Keeling M J, & Danon L. 2009. Mathematical modelling of infectious diseases. British Medical Bulletin, 92(1), 33-42

Keeling MJ, Woolhouse M E J, Shaw DJ, Matthews L, Chase-Topping M, Haydon DT et al. 2001. Dynamics of the 2001 UK foot-and-mouth epidemic: Stochastic dispersal in a heterogeneous landscape. Science, 294(5543), 813-817.

Kermack W O & McKendrick A G. 1927. A contribution to the mathematical theory of epidemics. Proceedings of the Royal Society A, 115, 700-721.

Kewlani G & Iagnemma K. 2008. A stochastic response surface approach to statistical prediction of mobile robot mobility. In Proceedings of IEEE/RSJ International Conference on Robots and Intelligent Systems. Accessed online 2/8/2010. http://web.mit.edu/mobility/publications/IROS_Final.pdf

Kitching R P, Hutber A M & Thrusfield M V. 2005. A review of foot-and-mouth disease with special consideration for the clinical and epidemiological factors relevant to predictive modelling of the disease. The Veterinary Journal, 169 (2), 197-209.

Kleijnen J P C & Helton J C. 1999a. Statistical analyses of scatterplots to identify important factors in large-scale simulations, 1: Review and comparison of techniques. Reliability Engineering & System Safety, 65(2), 147-185.

Kleijnen J P C & Helton J C. 1999b. Statistical analyses of scatterplots to identify important factors in large-scale simulations, 2: Robustness of techniques. Reliability Engineering & System Safety, 65(2), 187-197.

Lempert R, Popper S, Bankes S. 2002. Confronting surprise. Social Science Computer Review, 20 (4), 420–440.

Letcher R A, Croke B F W, Merritt W S & Jakeman A J. 2006. An integrated modelling toolbox for water resources assessment and management in highland catchments: Sensitivity analysis and testing. Agricultural Systems, 89(1), 132-164.

McKay M, Conover W & Beckman R. 1979. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. Technometrics, 21, 239-245.

McKay M. 1995. Evaluating prediction uncertainty. NUREG/CR-6311, U S Nuclear Regulatory Commission and Los Alamos National Laboratory Report.

Manache G & Melching C S. 2008. Identification of reliable regression- and correlation-based sensitivity measures for importance ranking of water-quality model parameters. Environmental Modelling & Software, 23(5), 549-562.

Marino S, Hogue I B, Ray C J & Kirschner D E. 2008. A methodology for performing global uncertainty and sensitivity analysis in systems biology. Journal of Theoretical Biology, 254(1), 178-196.

Marrel A, Iooss B, Van Dorpe F & Volkova E. 2008. An efficient methodology for modeling complex computer codes with Gaussian processes. Computational Statistics & Data Analysis, 52(10), 4731-4744.

Mokhtari A & Frey H C. 2005. Recommended practice regarding selection of sensitivity analysis methods applied to microbial food safety process risk models. Human and Ecological Risk Assessment, 11(3), 591-605.

Mokhtari A, Frey H C& Zheng J Y. 2006. Evaluation and recommendation of sensitivity analysis methods for application to stochastic human exposure and dose simulation models. Journal of Exposure Science and Environmental Epidemiology, 16(6), 491-506.

Morris M D. 1991. Factorial sampling plans for preliminary computational experiments Technometrics, 33 (2), 161-174.

Morris M D. 2006. Input screening: Finding the important model inputs on a budget. Reliability Engineering & System Safety, 91(10-11), 1252-1256.

Morris M D, Moore L M & McKay M D. 2006. Sampling plans based on balanced incomplete block designs for evaluating the importance of computer model inputs. Journal of Statistical Planning and Inference, 136, 3203-3220.

Morris R S, Wilesmith J W, Stern M W, Sanson R L & Stevenson M A. 2001. Predictive spatial modelling of alternative control strategies for the foot-and-mouth disease epidemic in Great Britain. Veterinary Record, 149 (5), 137-144.

Ministry for Primary Industries, New Zealand. National Animal Identification and Tracing (NAIT). Accessed online 2/5/12. http://www.mpi.govt.nz/agriculture/rural-communities/national-animal-identification-and-tracing-nait

Oreskes N, Shraderfrechette K & Belitz K. 1994. Verification, validation, and confirmation of numerical-models in the earth-sciences. Science, 263(5147), 641-646.

Patterson J, Meek M E B, Strawson J E, Liteplo R G. 2007. Engaging expert peers in the development of risk assessments. Risk Analysis, 27, 6, 1609-1621.

Patton D E, Olin S S. 2006. Scientific peer review to inform regulatory decision making: Leadership responsibilities and cautions Risk Analysis, 26, 1, 5-16.

Pearson, AB - 2008. Pig Movements Transmission of Exotic Diseases. Final Report for Project 07/ 10/475, Ministry of Primary Industries, New Zealand. Accessed 2/09 2010.

http://www.biosecurity.govt.nz/files/pests/salt-freshwater/022009-pig-movements-transmission-of-exotic-diseases.pdf

Ratto M, Young P C, Romanowicz R, Pappenberger F, Saltelli A & Pagano A. 2007. Uncertainty, sensitivity analysis and the role of data based mechanistic modeling in hydrology. Hydrology and Earth System Sciences, 11(4), 1249-1266.

Refsgaard J C, van der Sluijs JP, Højberg AL & Vanrolleghem PA. 2007. Uncertainty in the environmental modelling process – a framework and guidance. Environmental Modelling and Software, 22 (11), 1543-1556.

Riggs T, Walts A, Perry N, Bickle L, Lynch J N, Myers A, Flynn J, Linderman J J, Miller M J & Kirschner D E. 2008. A comparison of random vs chemotaxisdriven contacts of T cells with dendritic cells during repertoire scanning. Journal of Theoretical Biology, 250 (4), 732–751.

Rosen R. 1991. Life Itself: A Comprehensive Inquiry into the Nature, Origin, and Fabrication of Life, Columbia University Press, pp60-61.

Sallaberry C J, Helton J C & Hora S C. 2008. Extension of Latin Hypercube samples with correlated variables. Reliability Engineering & System Safety, 93(7), 1047-1059.

Saltelli A . 2004. Sensitivity Analysis in Practice: A Guide to Assessing Scientific Models. John Wiley and Sons Ltd., New York

Saltelli A, Chan K & Scott E. 2000. Sensitivity Analysis. John Wiley & Sons Ltd., New York.

Saltelli A & Tarantola S. 2002. On the relative importance of input factors in mathematical models: Safety assessment for nuclear waste disposal. Journal of the American Statistical Association, 97 (459), 702-709.

Saltelli A, Ratto M, Tarantola S & Campolongo F. 2006. Sensitivity analysis practices. Strategies for model-based inference. Reliability Engineering and Systems Safety, 91, 1109-1125.

Sanson R. 1993. The development of a decision support system for an animal disease emergency. PhD thesis submitted to Massey University, Palmerston North, New Zealand.

Sanson R L. 1994. The epidemiology of foot-and-mouth disease: implications for New Zealand. New Zealand Veterinary Journal, 42 (2), 41-53.

Sanson R L. 2005. A survey to investigate movements off sheep and cattle farms in New Zealand, with reference to the potential transmission of foot-and-mouth disease. New Zealand Veterinary Journal, 53 (4), 223-233.

Sanson R L., Harvey N, Garner M G , Stevenson M A , Davies T M, Hazelton M L, O'Connor JO, Dubé C, Forde-Folle K N & Owen K. 2011. Foot-and-mouth disease model verification and 'relative validation' through a formal model comparison. Revue Scientifique et Technique (International Office of Epizootics). 30 (2), 527-540

Sanson R, Stevenson M, Mackereth G & Moles-Benfell N. 2006. The development of an InterSpread Plus parameter set to simulate the spread of FMD in New Zealand. In Prococeedings of the 11th International Symposium on Veterinary Epidemiology and Economics (ISVEE), 6-11 August, Cairns, Australia.

Savill N J, Shaw DJ, Deardon R, Tildesley M J, Keeling M J, Woolhouse MEJ, Brooks, SP & Grenfell BT. 2007. Effect of data quality on estimates of farm infectiousness trends in the UK 2001 foot-and-mouth disease epidemic. Journal of the Royal Society Interface, 4(13), 235-241.

Segovia-Juarez J L, Ganguli S & Kirschner D. 2004. Identifying control mechanisms of granuloma formation during M. tuberculosis infection using an agent-based model. Journal of Theoretical Biology, 231 (3), 357–376.

Slocum N. 2005. Participatory Methods Toolkit: A practitioner's manual. Method: Delphi. The King Baudouin Foundation and the Flemish Institute for Science and Technology Assessment (viWTA). Accessed online 2/2/2010: http://www.viwta.be/files/ToolkitDelphi.pdf.

Stevenson M, Sanson R, Stern M, O'Leary B, Mackereth G, Sujau M, Moles-Benfell N & Morris R. 2006. InterSpread Plus: a spatial and stochastic simulation model of disease in animal populations. In Technical paper for research project BER-60-2004. Biosecurity New Zealand, Wellington, New Zealand.

Stevenson M, Sanson R, Stern M, O'Leary B, Sujau M, Molles-Benfell N & Morris R. 2012. InterSpread Plus: a spatial and stochastic simulation model of disease in animal populations

Preventive Veterinary Medicine. Accessed online 18/9/2012 DOI: http://dx.doi.org/10.1016/j.prevetmed.2012.08.015

Taylor I. 2003. Policy on the hoof: the handling of the foot-and-mouth disease outbreak in the UK 2001. Policy and Politcs., 31 (4), 535-546.

Taylor N. 2003. Review of the use of models in informing disease control policy development and adjustment. A report for the Defra Veterinary Epidemiology and Economics Research Unit (VEERU), Reading, 94 pp. Accessed on line 2/2/2010. http://epicentre.massey.ac.nz/resources/acvsc_grp/docs/Taylor_2003.pdf

Thrusfield M, Mansley L, Dunlop P, Taylor J, Pawson A & Stringer L. 2005. The foot-and-mouth disease epidemic in Dumfries and Galloway, 2001. 1: Characteristics and control. Veterinary Record, 156 (8), 229-252.

Thurmond MC, Perez AM. 2006. Modeled detection time for surveillance for foot-and-mouth disease virus in bulk tank milk. American Journal of Veterinary Research, 67(12), 2017-2024

Tildesley M J, Deardon R, Savill N J, Bessell P R, Brooks S P, Woolhouse M E J, Grenfell BT & Keeling MJ. 2008. Accuracy of models for the 2001 foot-and-mouth epidemic. Proceedings of the Royal Society B-Biological Sciences, 275(1641), 1459-1468.

Tsutsui T. Minami N, Koiwai M, Hamaoka T, Yamane I & Shimura K. 2003. A stochastic-modeling evaluation of the foot-and-mouth-disease survey conducted after the outbreak in Miyazaki, Japan in 2000. Preventive Veterinary Medicine, 61(1), 45-58.

Van der Sluijs J P. 2007. Uncertainty and precaution in environmental management: Insights from the UPEM conference. Environmental Modelling & Software, 22(5), 590-598.

Woolhouse M E J, Haydon D T, Pearson A & Kitching R P. 1996. Failure of vaccination to prevent outbreaks of foot-and-mouth disease. Epidemiology and Infection, 116(3), 363-371.

Xu C, He H S, Hu Y, Chang Y, Li X & Bu R. 2005. Latin Hypercube sampling and geostatistical modeling of spatial uncertainty in a spatially explicit forest landscape model simulation. Ecological Modelling, 185(2-4), 255-269.

Young P, Parkinson S, & Lees M. 1996. Simplicity out of complexity in environmental modelling: Occam's razor revisited. Journal of Applied Statistics, 1996, 23 (2-3), 165-210.