Stochastic and Deterministic Models for Agricultural Production Networks

P. Bai[1,a], H.T. Banks[2,b], S. Dediu[2,3,c], A.Y. Govan[2,d], M. Last[4,e], A Lloyd[2,f], H.K. Nguyen[2,5,g], M.S. Olufsen[2,h], G. Rempala[6,i], and B.D. Slenning[7,j]

February 22, 2007

1. Department of Statistics, University of North Carolina, Chapel Hill, NC
2. Department of Mathematics and Center for Research in Scientific Computation, North Carolina State University, Raleigh, NC
3. Statistical and Applied Mathematics Institute, Research Triangle Park, NC
4. National Institute of Statistical Sciences, Research Triangle Park, NC
5. Center for Naval Analyses, Alexandria, VA
6. Department of Mathematics and Statistics, University of Louisville, KY
7. Department of Population Health and Pathobiology, College of Veterinary Medicine, North Carolina State University, Raleigh, NC

Abstract

An approach to modeling the impact of disturbances in an agricultural production network is presented. A stochastic model and its approximate deterministic model for averages over sample paths of the stochastic system are developed. Simulations, sensitivity and generalized sensitivity analyses are given. Finally, it is shown how diseases may be introduced into the network and corresponding simulations are discussed.

Keywords: Agricultural production networks, stochastic and deterministic models, sensitivity and generalized sensitivity functions, foot and mouth disease

[a]pbai@email.unc.edu, [b]htbanks@ncsu.edu, [c]sdediu@ncsu.edu, [d]aygovan@ncsu.edu, [e]mlast@niss.org, [f]alun_lloyd@ncsu.edu, [g]nguyenhk@cna.org, [h]msolufse@ncsu.edu, [i]g.rempala@louisville.edu, [j]barrett_slenning@ncsu.edu

1 Introduction

The current production methods for livestock follow the just-in-time philosophy of manufacturing industries. Feedstock and animals are grown in different areas. Animals are moved from one farm to another, depending on their age. Unfortunately, shocks propagate rapidly through such systems; an interruption to the feed supply has a much larger impact when farms have minimal surplus supplies in-stock than when they have large reserves. The just-in-time movement of animals between farms serves as another vulnerability. Stopping movement of animals to and from a farm with animals infected by a disease will have effects that quickly spread through the system. Nurseries supplying the farm will have nowhere to send their animals as they grow up. Finishers and slaughterhouses supplied by the farm will have their supply interrupted.

The devastating Foot and Mouth Disease (FMD) that hit the United Kingdom (UK) in 2001 lead to the slaughter of millions of animals. The outbreak shook many western nations as they watched a nation with an advanced animal health surveillance and response system fail to get FMD under control, in part because they were unable to mount a rapid response in the face of modern agricultural marketing systems [15]. In an effort to eradicate the disease, the marketing channels were stopped, putting uninfected producers in the situation of having no income to maintain their livestock and no means to move them to locations where feed, shelter, and other support were available. As a result, between six million and ten million animals were destroyed in the UK over seven months, with over one-third of those animals being destroyed for welfare reasons [41]. Two years after the outbreak, animal agriculture in the UK was still declining, a chilling postscript to the widespread infrastructure damage FMD had wrought on the nation [37].

More recently, the world has witnessed the apparent failure of widespread national and international plans for using animal destruction to stem the spread of the highly pathogenic H5N1 strain of avian influenza. In a process frighteningly reminiscent of the UK FMD experience, the programs have also allowed domestic markets within and beyond affected countries to suffer. Global consumption of poultry has dropped sufficiently to cause US domestic producers (e.g., Tyson, Pilgrims Pride, et al.) to absorb decreased demand and decreased prices. This drop has translated to decreases, as well, in non-poultry markets, exacerbating the market effects of a disease not even present in the western hemisphere [19]. It has become painfully apparent that in the large-scale, interdependent, and highly mobile animal agriculture industry of the USA, the unintended consequences and market ripple effects of a disease incursion into our system could be even more severe than what was witnessed in the UK in 2001 and across Europe in 2005-6, and could induce decision makers to call for even more draconian measures than previously seen. What is needed is a new view of how our emergency response programs might affect modern animal agriculture, a view that allows workers to assess the potential for other prevention strategies and responses. The view should also allow analysts to identify bottlenecks in the food and feed supply chain, and to test potential mitigation tools, procedures, and/or practices to increase the resilience of animal agriculture to catastrophic animal diseases.

This paper presents initial statistical and mathematical modeling ideas to address the above issues, using the North Carolina swine industry's potential response to FMD as an example. We focused our attention on the North Carolina swine industry because it is both the second largest swine industry in the United States, and is local to us. Our goal was to develop a model that could be used to investigate how small perturbations to the agricultural supply system would affect its overall performance. A hurricane that throttles inter-farm transportation for a short duration, or a disease outbreak that spreads through distribution channels are example causes of the perturbations of interest. In the former case, the just-in-time delivery systems may not provide enough slack to absorb the shock. In the latter case, strategies that involve destruction of all livestock in an infected branch of the system may be overly harsh; a more moderate response may be as effective without the high toll on the infrastructure.

We model a simplified swine production network in North Carolina containing four levels of production nodes: growers/sows (*Node* 1), nurseries (*Node* 2), finishers (*Node* 3), and processing plants/slaughter houses (*Node* 4). At grower or sow farms (*Node* 1), the new piglets are born and typically weaned three weeks after birth. The three-week old piglets are then moved to the nursery farms (*Node* 2) to mature for another seven weeks. They are then transferred to the finisher farms (*Node* 3), where they grow to full market size, which takes about twenty weeks. Once they reach market weight, the matured pigs are moved to the processors (slaughter houses) (*Node* 4). Pork products then continue through wholesalers to consumers. There are also several inputs to the system which we will not consider, such as food, typically corn grown in the midwest. There are several types of breeding farms where pure-bred stock are raised; these are typically crossed to produce hybrid strains for pork production.

Our paper is organized as follows. In Section 2, we formulate a nonlinear stochastic model for our agricultural network and show how it can be converted to an equivalent (in a sense made precise below) deterministic differential equation model. This deterministic model readily lends itself to simulations and sensitivity analysis techniques. In Section 3 we present numerical simulations of the production model (without perturbations such as infectious disease), and carry out a sensitivity analysis of the model. Simulations of the model in the presence of an infectious disease are presented in Section 4. Finally, in Section 5 we give our conclusions and remarks for future work.

In addition to the development of models for a typical production network permitting perturbations, a significant contribution in this paper is the demonstration of stochastic, mathematical and computational *methodology* that is available to domain scientists, statisticians and applied mathematicians working in a concerted team effort on complex problems of the type exemplified here. The co-authors of this paper constituted such a team organized under the auspices and with the support of the Statistical and Applied Mathematical Sciences Institute (SAMSI) as a year long working group in its recent research program on National Defense and Homeland Security.

2 Modeling

We consider stochastic models to track an agricultural network. We are interested in how the parameters used in the model affect the overall capacity of a network, and how one discerns the existence and location of any bottlenecks. With deterministic models, one can answer the first question with a sensitivity analysis. Thus, after developing a typical stochastic production model, we also show how to obtain its deterministic approximation. We then demonstrate how to superimpose a simple contagious disease model on the production model that allows simulation of dynamics and spread of FMD through a production chain.

2.1 Basic Model

We consider a simplified swine production network with four aggregated nodes: sows (*Node* 1), nurseries (*Node* 2), finishers (*Node* 3), and slaughter houses (*Node* 4). Our goal is to study the effects of perturbations within the network. This can be done either by affecting the nodes or the transitions between nodes directly or indirectly. For instance, a problem with the breeding farms would result in a reduction of sows available for producing new piglets. This would result in a reduced rate of transition from *Node* 1 to *Node* 2, since we could not grow as many piglets. We could then track the effect of this through our network.

Although unavoidable in actual production processes, we assume in our example that there are no net losses in the network (i.e., the total number of pigs in the network remains constant) and that the only deaths occur at the slaughter houses. Thus we assume that the number of processed pigs per day at the slaughter houses is equal to the number of newborn piglets per day at the growers. We can model reduced birth-rates by reducing the rate at which piglets move to the nurseries. This leads us to deal with a *closed network*. We note that this approximation is realistic when the network is efficient and operates at or near full capacity (i.e, when the number of animals removed from the chain are immediately replaced by new production/growth, avoiding significant idle times). Our closed network model for the swine production is summarized schematically in Figure 1.

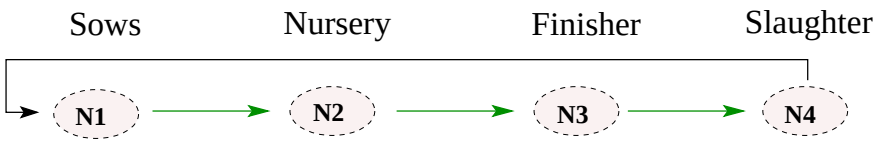

Figure 1: Aggregated agricultural network model.

Each node with corresponding population number N_i, $i = 1, \ldots, 4$, in Figure 1 represents an aggregation of all the production units corresponding to that level in the production network. Given a specific production network, any of the four levels of the chain may be broken into its constituent units (e.g., farms), and analyzed in detail as a separate sub-network. The directed edges between the nodes represent the movement of the pigs through the network. The rate is determined by the pigs' residence time, the number of pigs at each

node, and the capacity constraints at the corresponding nodes. Let L_i denote the *capacity constraint* at node i, $i = 1, \ldots, 4$. Since we have a closed network, it is assumed that there is no capacity constraint at *Node* 1 and therefore we take $L_1 = \infty$. We also define S_m to be the maximum exit rate at *Node* 4, i.e., the maximum killing capacity at the slaughter house. The residence times at each node, together with the capacity constraints and the slaughterhouse killing capacity, based on very rough estimates of swine production in North Carolina [1], are given in Table 1.

Name	Sows	Nursery	Finisher	Slaughter
Node	N1	N2	N3	N4
Piglet residence time (days)	$21_{(N1\to N2)}$	$49_{(N2\to N3)}$	$140_{(N3\to N4)}$	$1_{(N4\to N1)}$
Assumed capacity (in thousands)	∞	825	2,300	20

Table 1: Network parameters based on swine production in NC.

2.2 Stochastic and Deterministic Models

We model the evolution of the food production network shown in Figure 1 as a continuous time discrete state density dependent jump Markov Chain (MC) [3, 21] with a discrete state space embedded in an $\mathbb{R}^4$ non-negative integer lattice $\mathcal{L}$. The state of this MC at time t is denoted by $\mathbf{X}(t) = (X_1(t), \ldots, X_4(t))$, where $X_i(t)$ is the number of pigs at node i at time t, $i = 1, \ldots, 4$.

The rates of transition of $\mathbf{X}(t)$ are nonlinear functions $\lambda_i : \mathcal{L} \to [0, \infty)$ for $i = 1, \ldots, 4$, and for $\mathbf{x} \in \mathcal{L}$ are given by:

$$\begin{aligned}
\lambda_1(\mathbf{x}) &:= q_1(x_1 - 1, x_2 + 1, x_3, x_4) &&= k_1\, x_1\, (L_2 - x_2)_+ \\
\lambda_2(\mathbf{x}) &:= q_2(x_1, x_2 - 1, x_3 + 1, x_4) &&= k_2\, x_2\, (L_3 - x_3)_+ \\
\lambda_3(\mathbf{x}) &:= q_3(x_1, x_2, x_3 - 1, x_4 + 1) &&= k_3\, x_3\, (L_4 - x_4)_+ \\
\lambda_4(\mathbf{x}) &:= q_4(x_1 + 1, x_2, x_3, x_4 - 1) &&= k_4\, \min(x_4, S_m)
\end{aligned} \tag{1}$$

where $k_i, i = 1, \ldots, 4$, is proportional to the service rate at node i; $L_i, i = 2, 3, 4$, is the buffer size (capacity constraint) at node i and S_m is the slaughter capacity at node 4 as discussed above. For any real z, the symbol $(z)_+$ is defined as the non-negative part of z, i.e., $(z)_+ = \max(z, 0)$. Then $q_1(x_1 - 1, x_2 + 1, x_3, x_4)$ is given by

$$q_1(x_1 - 1, x_2 + 1, x_3, x_4) \equiv \lim_{h \to 0+} \frac{\Pr[\mathbf{X}(t + h) = (x_1 - 1, x_2 + 1, x_3, x_4) | \mathbf{X}(t) = (x_1, x_2, x_3, x_4)]}{h}.$$

The other q_i are given similarly.

Let $R_i(t)$ $i = 1, \ldots, 4$, denote the number of times that the i'th transition occurs by time t. Then R_i is a counting process with intensity $\lambda_i(\mathbf{X}(t))$, and the corresponding stochastic process can be defined by

$$R_i(t) = Y_i\Big(\int_0^t \lambda_i(\mathbf{X}(s))ds\Big), \quad i = 1 \ldots, 4, \tag{2}$$

where the Y_i are independent unit Poisson processes. That is, sample paths $r_i(t)$ of $R_i(t)$ are given in terms of sample paths $\mathbf{x}(t)$ of $\mathbf{X}(t)$ by

$$r_i(t) = Y_i\Big(\int_0^t \lambda_i(\mathbf{x}(s))ds\Big), \quad i = 1 \ldots, 4. \tag{3}$$

We write R_i in this form to illustrate that λ_i is a rate of the corresponding counting process.

Let $\mathbf{e}_i, i = 1, \ldots, 4$, be standard basis vectors of $\mathbb{R}^4$ and define, for i incremented by one modulo 4, the vectors

$$\boldsymbol{\nu}_i = \mathbf{e}_{(i+1)(\mathrm{mod}4)} - \mathbf{e}_i \quad i = 1, 2, \ldots, 4,$$

which represent the vector of changes in system counts at ith transition. We write the state of the system at time t as

$$\mathbf{X}(t) = \mathbf{X}(0) + \sum_i R_i(t)\boldsymbol{\nu}_i = \mathbf{X}(0) + \boldsymbol{\nu}\boldsymbol{R}(t), \tag{4}$$

where $\boldsymbol{\nu}$ is the matrix with rows given by the $\boldsymbol{\nu}_i$, and $\mathbf{R}(t)$ is the (column) vector with components $R_i(t)$. In the chemical literature, the matrix $\boldsymbol{\nu}^T$ is often referred to as *the stoichiometric matrix* [29]. More specifically, we have

$$\begin{aligned} X_1(t) &= X_1(0) - R_1(t) + R_4(t) \\ X_2(t) &= X_2(0) + R_1(t) - R_2(t) \\ X_3(t) &= X_3(0) + R_2(t) - R_3(t) \\ X_4(t) &= X_4(0) + R_3(t) - R_4(t). \end{aligned} \tag{5}$$

The above system typically cannot be solved for a stationary distribution and an empirical approach based on the so-called Gillespie algorithm [29] can be used to investigate the long term behavior of the system (see Section 3.2). The approximate large population behavior of an appropriately scaled system may be also analyzed via macroscopic deterministic rate equations as we shall explain next (the original theory is due to Kurtz and is discussed in [21] and the references therein).

Let N be the total network or population size. If N is known we may consider the animal *units per system size* or the units *concentration* in the stochastic process $\mathbf{C}^N(t) = \mathbf{X}(t)/N$ with sample paths $\mathbf{c}^N(t)$. For large systems this approach leads to a *deterministic* approximation (obtained as solutions to the *system rate equation* defined below) to the stochastic equation (4), in terms of $\mathbf{c}(t)$, the *large sample size average* over sample paths or trajectories $\mathbf{c}^N(t)$ of $\mathbf{C}^N(t)$.

We rescale the rate constants k_i, L_i and S_m as follows:

$$\begin{aligned} &\kappa_4 = k_4, \quad \kappa_i = Nk_i, \ i = 1, 2 \text{ or } 3, \\ &s_m = S_m/N, \quad l_i = L_i/N. \end{aligned} \tag{6}$$

According to Equation (1), this rescaling implies that

$$\lambda_i(\mathbf{x}) = \kappa_i \, x_i(L_{i+1} - x_i)_+/N = N\kappa_i \, c_i^N(l_{i+1} - c_i^N)_+ \quad i = 1, 2, 3,$$

and

$$\lambda_4(\mathbf{x}) = \kappa_4 \, \min(x_4, S_m) = N\kappa_4 \, \min(c_4^N, s_m).$$

Recall that for large N the *Strong Law of Large Numbers* (SLLN) for the Poisson Process Y implies $Y(Nu)/N \approx u$ [30]. One can use this fact, along with the rescaling of the constants as given above, to argue that sample paths $r_i(t)$ for the counting process (2) defined in terms of the sample paths $\mathbf{x}(t)$ or $\mathbf{c}^N(t) = \mathbf{x}(t)/N$ may be approximated for large N in terms of the deterministic variables $\mathbf{c}(t)$, the averages over sample paths or trajectories $\mathbf{c}^N(t)$ of $\mathbf{C}^N(t)$, by

$$\begin{aligned} r_i^{(N)}(t) = \frac{1}{N} r_i(t) &= \frac{1}{N} \, Y_i\big(\int_0^t \lambda_i(\mathbf{x}(s)) ds \big) \\ &= \frac{1}{N} \, Y_i\big(N \int_0^t \kappa_i c_i^N(s)(l_{i+1} - c_{i+1}^N(s))_+ \, ds \big) \\ &\approx \int_0^t \kappa_i \, c_i(s)(l_{i+1} - c_{i+1}(s))_+ \, ds \quad \text{for } i = 1, 2, 3, \end{aligned} \tag{7}$$

and similarly

$$r_4^{(N)}(t) = \frac{1}{N} r_4(t) \approx \int_0^t \kappa_4 \min(c_4(s), s_m) \, ds.$$

For a full and rigorous discussion of this "approximation in mean", see Chapters 6.4 and 11 of [21] and Chapter 5 of [3]. The averages $\mathbf{c}(t)$ satisfy a system of deterministic ordinary differential equations which can be heuristically derived by beginning with Equation (5). Upon dividing both sides of each equation by N and applying the above, we obtain the *rate equations*, i.e., the system of integral equations approximating via the SLLN the original

stochastic system, as follows:

$$\begin{aligned}
c_1^N(t) &= c_1(0) - r_1^{(N)}(t) + r_4^{(N)}(t) \\
&\approx c_1(0) - \int_0^t \kappa_1 c_1(s)(l_2 - c_2(s))_+ \, ds + \kappa_4 \min(c_4(t), s_m) \\
c_2^N(t) &= c_2(0) + r_1^{(N)}(t) - r_2^{(N)}(t) \\
&\approx c_2(0) - \int_0^t \kappa_2 c_2(s)(l_3 - c_3(s))_+ \, ds + \int_0^t \kappa_1 c_1(s)(l_2 - c_2(s))_+ \, ds \\
c_3^N(t) &= c_3(0) + r_2^{(N)}(t) - r_3^{(N)}(t) \\
&\approx c_3(0) - \int_0^t \kappa_3 c_3(s)(l_4 - c_4(s))_+ \, ds + \int_0^t \kappa_2 c_2(s)(l_3 - c_3(s))_+ \, ds \\
c_4^N(t) &= c_4(0) + r_3^{(N)}(t) - r_4^{(N)}(t) \\
&\approx c_4(0) + \int_0^t \kappa_3 c_3(s)(l_4 - c_4(s))_+ \, ds - \kappa_4 \min(c_4(t), s_m). \qquad (8)
\end{aligned}$$

Upon approximating the $c_i^N(t)$ on the left above by the $c_i(t)$ and differentiating the resulting equations, we find that the integral equation system is equivalent to a system of ordinary differential equations for $\mathbf{c}(t) \in \mathbb{R}^4$ given by

$$\begin{aligned}
\frac{dc_1(t)}{dt} &= -\kappa_1 c_1(t)(l_2 - c_2(t))_+ + \kappa_4 \min(c_4(t), s_m) \\
\frac{dc_2(t)}{dt} &= -\kappa_2 c_2(t)(l_3 - c_3(t))_+ + \kappa_1 c_1(t)(l_2 - c_2(t))_+ \\
\frac{dc_3(t)}{dt} &= -\kappa_3 c_3(t)(l_4 - c_4(t))_+ + \kappa_2 c_2(t)(l_3 - c_3(t))_+ \\
\frac{dc_4(t)}{dt} &= -\kappa_4 \min(c_4(t), s_m) + \kappa_3 c_3(t)(l_4 - c_4(t))_+ \qquad (9)
\end{aligned}$$

with the initial conditions $\mathbf{c}(0) = \mathbf{c}_0$. As we shall see in the next section, solutions of these equations yield quite good approximations to the sample paths of the stochastic system.

We remark that the product nonlinearities $x_i \, (L_{i+1} - x_{i+1})_+$ of (1) where transportation occurs more rapidly the further the node level is from capacity (i.e., the system reacts more rapidly to larger perturbations from capacity) are only one possible form for these terms. One could also reasonably argue for alternative terms of the form $x_i \, \chi_{i+1}$ where χ_{i+1} is the characteristic function for the set $\{(L_{i+1} - x_{i+1}) > 0\}$ so that the transportation rate from a node depends only on the number available at that node so long as capacity at the next node has not been reached. We remark that in this case the sensitivity analyses below are more difficult due to a lack of continuity of the dynamics in the system equations.

3 Computations and Model Comparison

3.1 Model Parameter Values

In order to carry out numerical simulations and to compare the results of the stochastic and deterministic models (equations (5) and (9), respectively), we must choose reasonable values for all model parameters. We note that our paper focuses on methodological issues and, for confidentiality and proprietary reasons, only limited information on the swine production network was available to us. Thus some of these parameter values may be only rough approximations to those that might be obtained using inverse problem techniques with data from actual production networks [1]. Consequently, the subsequent discussions in this paper are in no way an attempt to validate the above models. Nonetheless, we believe that the order-of-magnitude approximate parameter values we are using here are sufficient to allow us to develop and demonstrate effective use of methods and techniques which could be used with actual production network based parameters.

The parameters of the stochastic model, with the exception of the transition rate constants k_i, are given in Table 1. From the expressions for the transition rates (1), we see that the residence times, t_i, that pigs spend at node i are given by

$$t_i = \frac{1}{k_i} \frac{1}{(L_{i+1} - X_{i+1}(t))} \quad \text{for } i = 1, 2 \text{ or } 3, \tag{10}$$

and $t_4 = 1/k_4 = 1$. As discussed above, the nonlinear form of the transition rates (1) means that the residence time at a given node depends on how far the following node is below its capacity. Consequently, we determine the k_i by assuming that the given residence times pertain to the network in its equilibrium state.

Considering the deterministic model equations (9), we see that, if there is to be a flow through the system, the equilibrium population sizes $N_i^* = Nc_i^*$ at each node must be less than the capacities of the nodes. It is then straightforward to see that the equilibrium numbers of individuals at each of the first three nodes are proportional to the t_i. This makes intuitive sense since there is no loss as individuals move between nodes, and so, at equilibrium, the relative residence times must equal the relative numbers of individuals at the nodes. This argument need not apply to the slaughter node, however, since individuals will spend longer there than the specified one day residence time if the equilibrium value N_4^* is greater than S_m. The flow rate from node four back to node one is the smaller of N_4^* and S_m and so we have that

$$(N_1^*, N_2^*, N_3^*, N_4^*) = \begin{cases} (t_1 N_4^*, t_2 N_4^*, t_3 N_4^*, N_4^*) & \text{if } N_4^* \leq S_m \\ (t_1 S_m, t_2 S_m, t_3 S_m, N_4^*) & \text{otherwise.} \end{cases} \tag{11}$$

Notice that solving for the equilibrium of the deterministic model does not give us the value of N_4^*: since the network is closed, the total size of the population is equal to its value at the initial time. The values of the k_i, for $i = 1, 2$ and 3, are then given by

$$k_i = \frac{1}{t_i \left(L_{i+1} - N_{i+1}^*\right)}. \tag{12}$$

The parameter values and the initial states for the system (5) are tabulated in Table 2.

Parameters	Definition	Values	Units
k_1	transition rate at node 1	$1/(90 \cdot 21)$	1/(pigs·days)
k_2	transition rate at node 2	$1/(200 \cdot 49)$	1/(pigs·days)
k_3	transition rate at node 3	$1/(5 \cdot 140)$	1/(pigs·days)
k_4	transition rate at node 4	1	1/days
L_1	capacity node 1	∞	pigs
L_2	capacity node 2	825	pigs
L_3	capacity node 3	2300	pigs
L_4	capacity node 4	20	pigs
S_m	slaughter capacity	440	pigs
$X_1(t_0)$	initial condition	800	pigs
$X_2(t_0)$	initial condition	700	pigs
$X_3(t_0)$	initial condition	1500	pigs
$X_4(t_0)$	initial condition	165	pigs

Table 2: Aggregated agricultural network model: Parameters for stochastic simulations, together with our chosen initial conditions. All numbers of pigs are given in thousands here.

To obtain the parameters we use in our deterministic simulations we simply rescale the parameters in Table 2 by the total network size $N = X_1(t_0) + X_2(t_0) + X_3(t_0) + X_4(t_0) = 3,165,000$, using equation (6) and $c_i(t_0) = X_i(t_0)/N$ for $i = 1, \ldots, 4$. The results are given below in Table 3.

Parameters	Definition	Values	Units
κ_1	scaled rate at node 1	1.674	1/days
κ_2	scaled rate at node 2	0.323	1/days
κ_3	scaled rate at node 3	4.521	1/days
κ_4	scaled rate at node 4	1	1/days
l_2	scaled capacity at node 2	$2.607 \cdot 10^{-1}$	dimensionless
l_3	scaled capacity at node 3	$7.267 \cdot 10^{-1}$	dimensionless
l_4	scaled capacity at node 4	$6.3 \cdot 10^{-3}$	dimensionless
s_m	scaled slaughter capacity	$1.390 \cdot 10^{-1}$	dimensionless
$c_1(0)$	scaled initial condition	$2.528 \cdot 10^{-1}$	dimensionless
$c_2(0)$	scaled initial condition	$2.212 \cdot 10^{-1}$	dimensionless
$c_3(0)$	scaled initial condition	$4.739 \cdot 10^{-1}$	dimensionless
$c_4(0)$	scaled initial condition	$5.21 \cdot 10^{-2}$	dimensionless

Table 3: Aggregated agricultural network model: Rescaled parameter values and initial conditions for the deterministic model.

3.2 Stochastic Simulations

The standard method for the stochastic simulation of the discrete state continuous time Markov Chain of the type considered here is based on a standard Monte Carlo algorithm, also known as the *Gillespie algorithm* [29]. This algorithm is described below:

1. For a given state of the system $\mathbf{x}$, compute $\lambda_i(\mathbf{x})$ for $i = 1 \ldots, M$ (in our case $M = 4$).
2. Calculate the summation of the rates $\lambda = \sum_{i=1}^{M} \lambda_i(\mathbf{x})$ and simulate the time until the next transition by drawing from an exponential distribution with mean $1/\lambda$.
3. Simulate the transition type $\mathcal{R}_{\mathbf{X}} \in \{1, \ldots, 4\}$ by drawing from the discrete distribution with $P(\mathcal{R}_{\mathbf{X}} = i) = \lambda_i(\mathbf{x})/\lambda$.
4. Update the system state $\mathbf{x}$ and repeat.

Using the above algorithm implemented in the statistical software R [40], we carried out numerous simulations for the model (5) with the initial conditions and values for parameters $\mathbf{q}^* = (k_1 \ldots, k_4, S_m, L_2, L_3, L_4)$ given in Table 2.

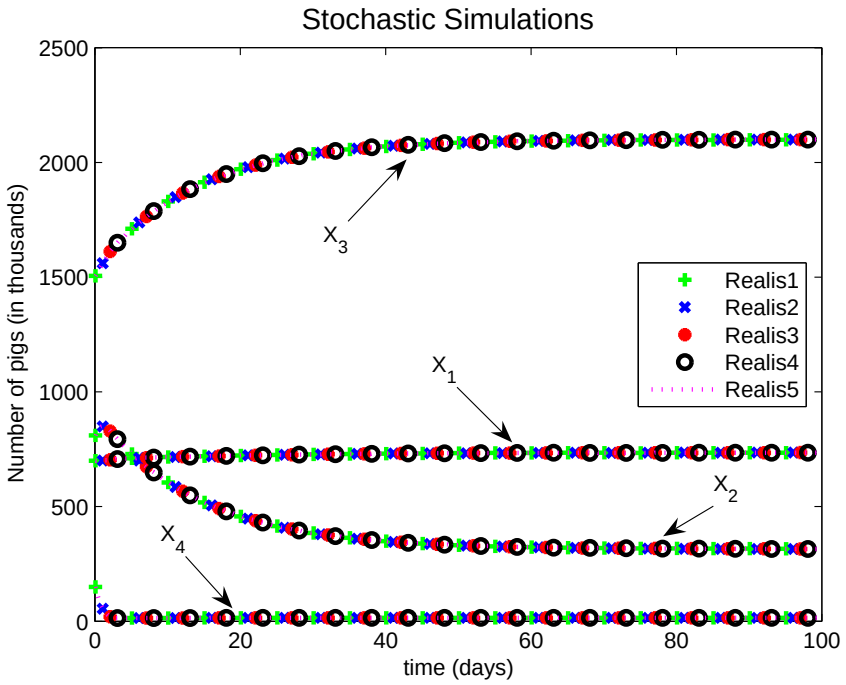

Figure 2: Simulation of the stochastic model (5) of the food production chain via the Gillespie algorithm. Parameter values are as given in Table 2; N=3,165,000.

A sample of five realizations is plotted in Figure 2. Note that the realizations exhibit very little visible differences. However, when one carries out the simulations for a smaller

system ($N = 3,165$ pigs with the parameters in Table 2 scaled accordingly), the variations are readily visible as can be seen in Figure 3. We also remark that these two figures offer graphic depictions of the approximation theory discussed in Section 2.2 where in the case of very large N one can cannot distinguish between the stochastic simulations and the corresponding deterministic simulations for the sample path averages.

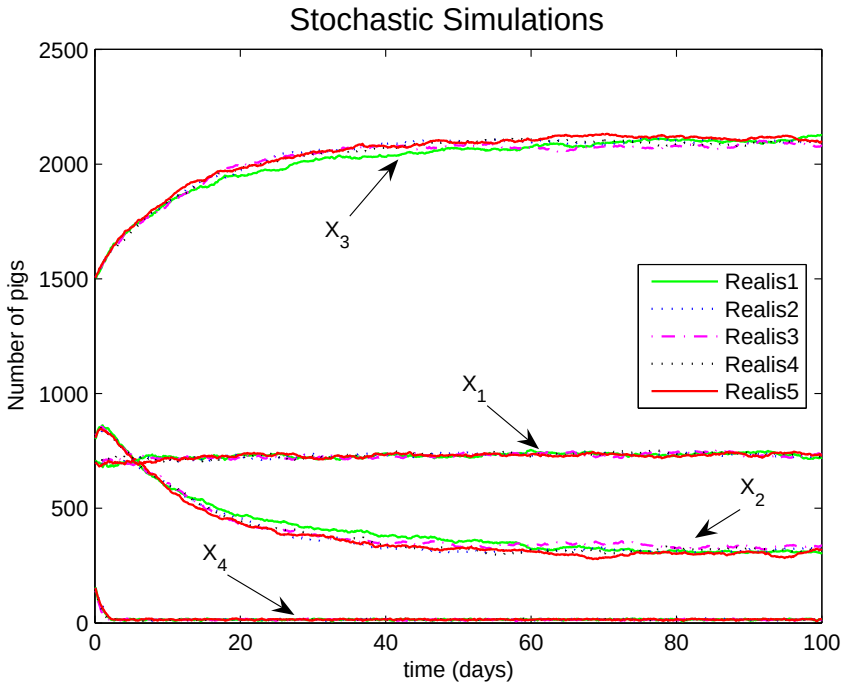

Figure 3: Simulation of the stochastic model (5) with N=3,165 and parameter values from Table 2 scaled accordingly.

3.3 Deterministic Simulation

We numerically solved the ODE system (9) using the *ode15s* solver in Matlab. We fixed the parameters $\mathbf{q} = \mathbf{q}^*$ to be the same as in the stochastic simulation above (but scaled as in Table 3) and graph the solution of the rate equations (9) for $t \in [0, 100]$ in Figure 4. In Figure 4(left), we plot the numerical solution of the concentrations in system (9). To facilitate comparison with the MC realizations plotted in Figure 2, we also depict the rescaled quantities $N_i(t) = Nc_i(t)$, which provide approximations to averages over sample paths of $X_i(t)$ in Figure 4(right). As expected, we find that the stochastic and deterministic computations provide similar numerical results with the realizations fluctuating about the solution of the deterministic system for the averages as predicted by the theory.

In order to observe the finer dynamics in the network, we plot the solution of the network on a smaller time scale in Figure 5(left). We find that our model solutions approach the

steady states rather quickly. All components remain stable thereafter suggesting that the steady states are (at least locally) asymptotically stable (this can be verified with analytical arguments). We believe that this behavior of the model describes the food production network realistically when it is uninterrupted by external events. In Figures 5(right), 6(left) and 6(right) we observe similar behavior in the food production chain for different values of the parameters L_4 and S_m. As one can observe in these figures, when the value of S_m is sufficiently large, the state N_4, which is related to the replenishment of the network, will never reach S_m, making the slaughter capacity constraint inactive in the production system. Only when S_m is smaller than a certain critical value will it be active and in this case we observe accumulation of animals in the slaughter house (e.g., see Figure 6(right)). These calculations along with numerous others we carried out suggest reasonable stability properties of the production chain in the absence of any interventions such as FMD (we will investigate such disturbances below).

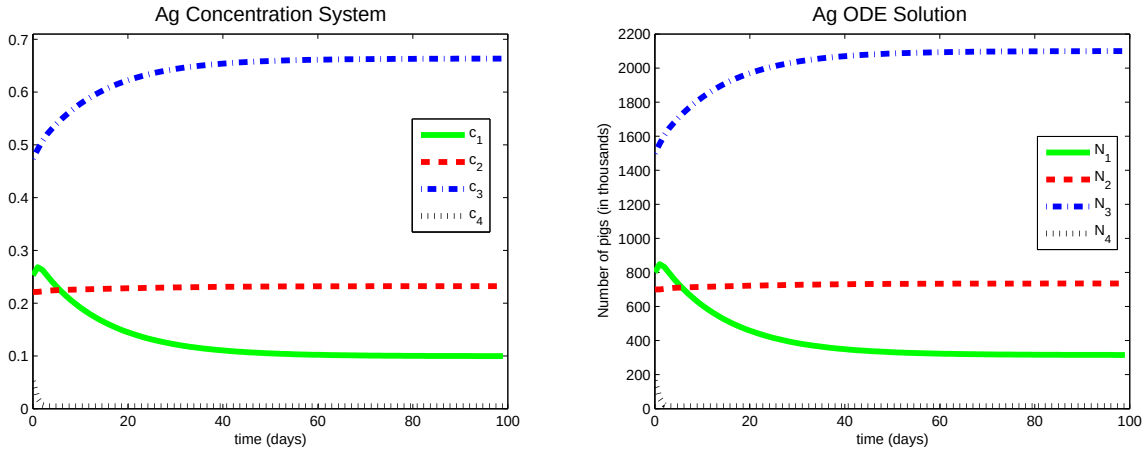

Figure 4: Numerical solution of the deterministic (rate equation) system (9) for $t \in [0, 100]$. Parameter values are as given in Table 3.

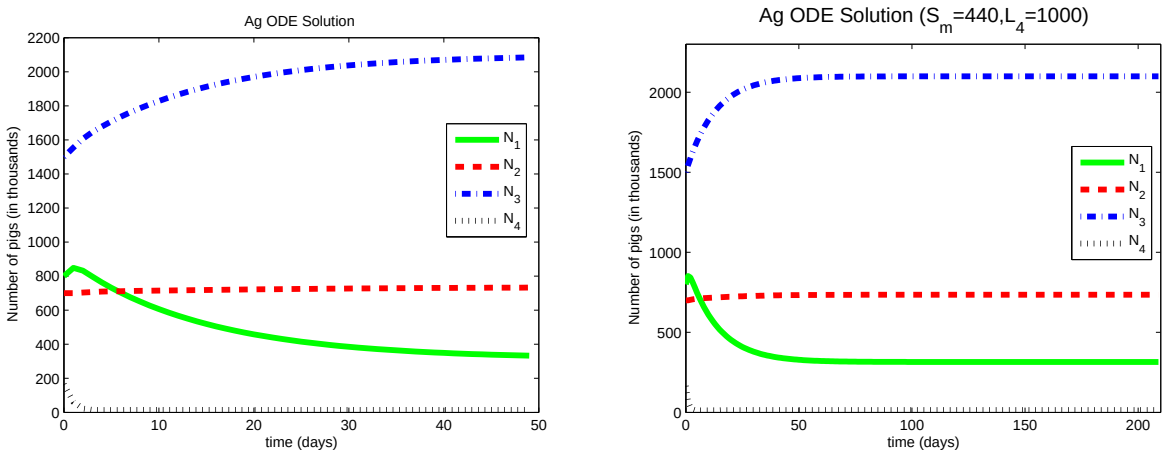

Figure 5: (left) Numerical solution of the corresponding deterministic system for $t \in [0, 50]$. (right) Numerical solution of the deterministic system (9) for $t \in [0, 210]$ with $L_4 = 1000$ and $S_m = 440$. All other parameter values are as given in Table 3.

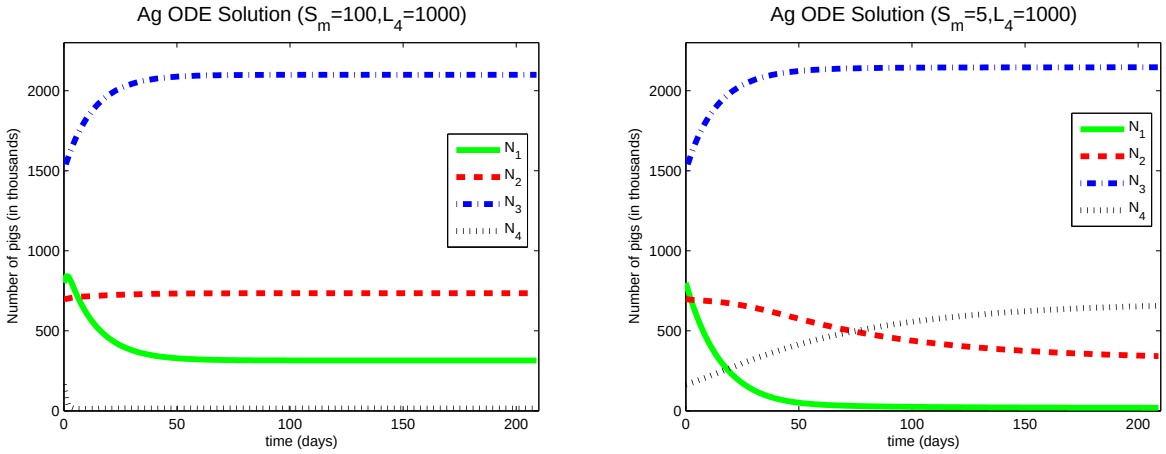

Figure 6: (left) Numerical solution of the deterministic system (9) for $t \in [0, 210]$ with $L_4 = 1000$ and $S_m = 100$. (right) Numerical solution of the deterministic system (9) for $t \in [0, 210]$ with $L_4 = 1000$ and $S_m = 5$. All other parameter values are as given in Table 3.

3.4 Sensitivity Analysis

In this section, we perform a sensitivity analysis of the deterministic model (9), investigating how much the solution of the system changes when the rates κ_i, the capacities l_i, or the initial conditions c_{0i}, $i = 1, \ldots, 4$ change. This analysis will be used to identify the parameters and the initial conditions to which the system is the most and least sensitive.

A second issue we address here, which is of great interest for inverse or parameter esti-

mation problems in a typical nonlinear regression model, is the sensitivity of the parameter estimates with respect to the data measurements. We carry out this analysis by means of the *generalized sensitivity functions* (GSF) recently introduced by Thomaseth and Cobelli [44]; these are specifically designed for input-output identification experiments. GSF are based on information theoretical criteria (the Fisher information matrix) and, when used in conjunction with the traditional sensitivity functions, give a more accurate picture on the time distribution of the information content of measured outputs with respect to individual model parameters.

In order to use the well-developed sensitivity analysis for the theory of dynamical systems for our purposes, we begin by writing the system (9) in vector form. We introduce the notation

$$\mathbf{c}(t) = (c_1(t), c_2(t), c_3(t), c_4(t))^T, \quad \mathbf{q} = (\kappa_1, ..., \kappa_4, s_m, l_2, ..., l_4), \quad \mathbf{c_0} = (c_1(0), ..., c_4(0))^T,$$

and denote by $\mathbf{F} = (f_1, f_2, f_3, f_4)^T$ the vector function whose entries are given by the expressions in the right side of (9). Then $\mathbf{F} : \mathbb{R}^4 \times \mathbb{R}^8 \to \mathbb{R}^4$, and we can write our ODE system in the general vector form

$$\begin{aligned} \frac{d\mathbf{c}}{dt}(t) &= \mathbf{F}(\mathbf{c}, \mathbf{q}), \\ \mathbf{c}(0) &= \mathbf{c_0}. \end{aligned} \tag{13}$$

In order to quantify the variation in the state variable $\mathbf{c}(t)$ with respect to changes in the parameters q_j, $j = 1, \ldots, 8$ and the initial conditions c_{0k}, $k = 1, \ldots, 4$, we are naturally led to consider the *sensitivity matrices*

$$\mathbf{Y} = \{y_{ij}\}_{\substack{i=1,...,4 \\ j=1,...,8}} = \left\{ \frac{\partial c_i}{\partial q_j} \right\}_{\substack{i=1,...,4 \\ j=1,...,8}}, \tag{14}$$

and

$$\mathbf{Z} = \{z_{ik}\}_{\substack{i=1,...,4 \\ k=1,...,4}} = \left\{ \frac{\partial c_i}{\partial c_{0k}} \right\}_{\substack{i=1,...,4 \\ k=1,...,4}}. \tag{15}$$

We note that since our function $\mathbf{F}$ is sufficiently regular, the solutions c_i are differentiable with respect to q_j and c_{0k} and therefore our sensitivity matrices $\mathbf{Y}$ and $\mathbf{Z}$ are well defined. The physical interpretation of the sensitivity matrices is obvious. Similar to the partial derivatives through which they are defined, they have a *local character* (in time and parameters). If, for example, the entry $y_{ij} = \partial c_i / \partial q_j$ of the matrix $\mathbf{Y}$ takes values very close to zero in a certain time subinterval, then the state variable c_i is *insensitive* to the parameter q_j on that particular subinterval. The same entry y_{ij} can take large values on a different subinterval, indicating that in this time subinterval, the state variable c_i is *very sensitive* to the parameter q_j.

From sensitivity analysis theory for dynamical systems [10, 20, 25, 36, 42], we also know that $\mathbf{Y}(t)$ is a 4×8 matrix that satisfies the ODE system

$$\begin{aligned} \dot{\mathbf{Y}}(t) &= \mathbf{F_c}(\mathbf{c}, \mathbf{q})\mathbf{Y}(t) + \mathbf{F_q}(\mathbf{c}, \mathbf{q}), \\ \mathbf{Y}(0) &= \mathbf{0}_{4\times 8}, \end{aligned} \tag{16}$$

and $\mathbf{Z}(t)$ is a 4×4 matrix that satisfies

$$\begin{aligned} \dot{\mathbf{Z}}(t) &= \mathbf{F_c}(\mathbf{c}, \mathbf{q})\mathbf{Z}(t), \\ \mathbf{Z}(0) &= \mathbf{I}_{4\times 4}. \end{aligned} \tag{17}$$

Here we have used the notation $\mathbf{F_c} = \partial\mathbf{F}/\partial\mathbf{c}$ and $\mathbf{F_q} = \partial\mathbf{F}/\partial\mathbf{q}$ for the 4×4 and the 4×8 Jacobian matrices of $\mathbf{F}$ with respect to $\mathbf{c}$ and $\mathbf{q}$, respectively, while $\mathbf{0}$ and $\mathbf{I}$ are the zero and the identity matrices with appropriate dimensions. Note that while equations (16), (17) are linear in $\mathbf{Y}$ and $\mathbf{Z}$, they must be solved in tandem with equation (13), which is nonlinear. Consequently, the sensitivity analysis involves the solution of a set of nonlinear equations.

We will compute the sensitivity of the system (13) with respect to $\mathbf{q}$ and $\mathbf{c_0}$ when the solutions are essentially at *steady state*. We carry this out by numerically solving the systems (16) and (17) for the same values of the parameters $\mathbf{q} = \mathbf{q}^*$ and initial conditions $\mathbf{c} = \mathbf{c_0^*}$ as used in the stochastic simulations presented above (i.e., those given in Table 3) and by evaluating the solution at the fixed time $t = 210$ (arbitrarily chosen, but sufficiently large for our system to closely approach its steady state). Due to the nature of our problem, in which the parameters have different units and the state variables vary widely over many orders of magnitude, it is appropriate to consider the *relative sensitivities* S_{c_i,q_j} defined as the limit of the relative change in c_i divided by the relative change in q_j when the relative change in q_j goes to zero, i.e.,

$$S_{c_i,q_j} = \lim_{\Delta q_j \to 0} \frac{\Delta c_i/c_i}{\Delta q_j/q_j}. \tag{18}$$

A simple analysis of the definition above (assuming that both c_i and q_j are nonzero) yields that the relative sensitivity S_{c_i,q_j} can be obtained by normalizing the usual sensitivities $\partial c_i/\partial q_j$ such that

$$S_{c_i,q_j} = \frac{\partial c_i}{\partial q_j} \cdot \frac{q_j}{c_i}. \tag{19}$$

We note that the S_{c_i,q_j} are dimensionless variables, invariant with respect to changes in units for c_i and q_j, which we can utilize to compare the degree of sensitivity of the state variables with respect to different parameters. In Table 4, we tabulate the relative sensitivities at time $t = 210$ of each state variable c_i with respect to each parameter q_j and each initial condition c_{0k}. For any fixed parameter/initial condition, we also tabulate the *sensitivity of the system*, cumulatively defined as the Euclidean norm of the relative sensitivities of the four state variables with respect to that parameter/initial condition. In other words, the sensitivity of the system with respect to q_j is given by

$$S_{q_j} = \Big[\sum_{i=1}^{4} S_{c_i,q_j}^2\Big]^{1/2}. \tag{20}$$

For the particular choice of the parameters $\mathbf{q} = \mathbf{q}^*$ and for the particular initial condition $\mathbf{c_0} = \mathbf{c_0^*}$, the data displayed in the last column of Table 4 reveals that near the steady-state ($t = 210$) the system (9) is *most sensitive* (in decreasing order) to l_3, c_{03} and l_2 and *least*

	c_1	c_2	c_3	c_4	System
κ_1	-2.65×10^{-1}	8.97×10^{-2}	8.34×10^{-3}	2.08×10^{-3}	2.80×10^{-1}
κ_2	-4.50×10^{-1}	-5.78×10^{-2}	8.76×10^{-2}	2.19×10^{-2}	4.62×10^{-1}
κ_3	1.70×10^{-1}	-9.05×10^{-3}	-2.40×10^{-2}	2.43×10^{-1}	2.98×10^{-1}
κ_4	5.45×10^{-1}	-2.28×10^{-2}	-7.18×10^{-2}	-2.67×10^{-1}	6.12×10^{-1}
s_m	0	0	0	0	0
l_2	-2.42×10^{0}	8.22×10^{-1}	7.65×10^{-2}	1.91×10^{-2}	2.56×10^{0}
l_3	-5.17×10^{0}	-6.65×10^{-1}	1.00×10^{0}	2.51×10^{-1}	5.32×10^{0}
l_4	6.80×10^{-1}	-3.61×10^{-2}	-9.63×10^{-2}	9.75×10^{-1}	1.19×10^{0}
c_{01}	1.86×10^{0}	2.27×10^{-1}	2.11×10^{-2}	5.29×10^{-3}	1.88×10^{0}
c_{02}	1.63×10^{0}	1.99×10^{-1}	1.85×10^{-2}	4.63×10^{-3}	1.64×10^{0}
c_{03}	3.49×10^{0}	4.27×10^{-1}	3.97×10^{-2}	9.93×10^{-3}	3.52×10^{0}
c_{04}	3.84×10^{-1}	4.69×10^{-2}	4.37×10^{-3}	1.09×10^{-3}	3.87×10^{-1}

Table 4: Relative sensitivities of the state variables/ system with respect to parameters q_j and the initial conditions c_{0k} near steady state ($t = 210$). Baseline parameter values ($\mathbf{q} = \mathbf{q}^*$) and initial conditions $\mathbf{c_0} = \mathbf{c_0^*}$ are as in Table 3.

sensitive to s_m. One interesting outcome we obtain from the analysis of the sensitivity data presented in Tables 4 and 5 is that the state variable c_1 is more sensitive to the parameter l_3 and the initial condition c_{03} as compared to the parameters κ_1, κ_4, s_m and l_4 and the initial condition c_{01} on which this state variable depends directly. Without this sensitivity analysis, we could not infer this behavior simply by looking at the system (9) since the parameters l_3 and c_{03} do not appear in the right side of the equation which defines dc_1/dt. We find similar results for c_2 and c_4 which are more sensitive to the non-direct initial condition c_{03} as compared to their direct initial conditions c_{02} and c_{04}.

As we pointed out previously, the sensitivity functions by definition have a local character both in the time domain and in the parameter domain, which implies that the results displayed in Table 4 characterizes the sensitivity of the system only for $\mathbf{q} = \mathbf{q}^*$, $\mathbf{c_0} = \mathbf{c_0^*}$ and $t = 210$. However, to obtain a broader picture of the sensitivity map for our system (9) near steady state ($t = 210$) or for the whole time interval, one can compute the relative sensitivities (19) and (20) with respect to q_j and c_{0k} on a uniform grid in a parameter neighborhood around the central values q_j^* and c_{0k}^* and analyze the results thus obtained. (Although we carried out such analyses, the results are not presented here.)

The sensitivity analysis we have performed so far is usually encountered in simulation studies (direct problems) where we need to quantify the effects of parameter variations on the trajectories of model outputs. Unlike simulations, in identification studies (inverse problems) one typically wants to estimate model parameters from data measurements and one question of interest is to determine at which time points the measurements are most informative for the estimation of specific parameters. In addition one may also desire a

c_1	c_2	c_3	c_4	System
s_m	s_m	s_m	s_m	s_m
κ_3	κ_3	c_{04}	c_{04}	κ_1
κ_1	κ_4	κ_1	κ_1	κ_3
c_{04}	l_4	c_{02}	c_{02}	c_{04}
κ_2	c_{04}	c_{01}	c_{01}	κ_2
κ_4	κ_2	κ_3	c_{03}	κ_4
l_4	κ_1	c_{03}	l_2	l_4
c_{02}	c_{02}	κ_4	κ_2	c_{02}
c_{01}	c_{01}	l_2	κ_3	c_{01}
l_2	c_{03}	κ_2	l_3	l_2
c_{03}	l_3	l_4	κ_4	c_{03}
l_3	l_2	l_3	l_4	l_3

Table 5: Summary of the sensitivity analysis presented in Table 4. The columns rank the parameters and initial conditions in order of increasing sensitivity for the four state variables, c_i, and the system.

priori information about the degree of correlation between model parameters which the TSF functions of (14) and (15) used alone cannot provide. To address these questions, Thomaseth and Cobelli [44] introduced the *generalized sensitivity functions* (GSF) which provide information on the relevance of measurements of output variables of a system for the identification of certain parameters as well as on parameter correlation. More precisely, the generalized sensitivity functions describe the sensitivity of the parameter estimates with respect to data measurements. We illustrate below the utility of these functions to provide a better understanding of our network model, by performing a sensitivity analysis for the inverse problem of estimating the parameters κ's of the system (9) through an ordinary least squares procedure.

For a single-output model $f(t, \theta)$ (e.g., the case where one has longitudinal observations of one component c_i of the system (13)) with discrete time measurements

$$y(t_k) = f(t_k, \theta_0) + \epsilon_k, \qquad k = 1, \ldots, M, \tag{21}$$

where θ_0 is the "true" parameter values (assumed to exist in most statistical inference and information content formulations–see [8, 11, 43]), the generalized sensitivity functions (GSF) are defined by

$$\mathbf{gs}(t_k) = \sum_{i=1}^{k} \Big\{ \Big(\Big[\sum_{j=1}^{M} \frac{1}{\sigma^2(t_j)} \nabla_\theta f(t_j, \theta_0) \nabla_\theta f(t_j, \theta_0)^T \Big]^{-1} \times \frac{\nabla_\theta f(t_i, \theta_0)}{\sigma^2(t_i)} \Big) \bullet \nabla_\theta f(t_i, \theta_0) \Big\}, \tag{22}$$

where "$\bullet$" represents element-by-element multiplication (see [44] and the Appendix of [7] for motivation and details). The measurement errors ϵ_k in (21) are assumed to be independently

and identically distributed with zero mean and known variance $\sigma^2(t_k)$ and the nonlinear model function $f(t,\theta)$ is assumed to be differentiable with respect to θ. Moreover, for simplicity, it is assumed that the observations $y(t_k)$, as represented in (21), are used to estimate θ_0 by minimizing the weighted sum of squares (WSS)

$$WSS(y,\theta) = \sum_{i=1}^{M} \frac{[y(t_i) - f(t_i,\theta)]^2}{\sigma^2(t_i)}. \tag{23}$$

In our case, the nonlinear model function f is replaced by a vector–valued function which is the solution of the system (9) and the generalized sensitivity functions (GSF) are given by

$$\mathbf{gs}(t_k) = \sum_{i=1}^{k}\sum_{l=1}^{4}\Big\{\Big(\Big[\sum_{j=1}^{M}\sum_{l=1}^{4}\frac{1}{\sigma_l^2(t_j)}\nabla_\theta c_l(t_j,\theta_0)\nabla_\theta c_l(t_j,\theta_0)^T\Big]^{-1} \times \frac{\nabla_\theta c_l(t_i,\theta_0)}{\sigma_l^2(t_i)}\Big) \bullet \nabla_\theta c_l(t_i,\theta_0)\Big\}. \tag{24}$$

Here $\nabla_\theta c_l$ represents the gradient of the state variable c_l with respect to θ, where $\theta \in \mathbb{R}^P$ is a vector including all (or just a subset of) the parameters κ's and l's and the initial conditions c_0's.

We note that the generalized sensitivity functions (24) are vector-valued functions having the same dimension P as θ and defined only at the discrete time points $t_k, k = 1,\ldots,M$. They are cumulative functions, at each time point t_k taking into account only the contributions of the measurements up to t_k, thus representing the influence of longitudinal measurements in contributing to the parameter estimates.

From (24) it follows that all the components of **gs** are one at the end of the experiment, i.e., $\mathbf{gs}(t_M) = \mathbf{1}$. If one defines $\mathbf{gs}(t) = \mathbf{0}$ for $t < t_1$ (**gs** is zero when no measurement is collected) and interpolates **gs** continuously between observation times, then each component gs_p of **gs** varies continuously from 0 to 1 during the experiment. As we will see in the example below, this transition is not necessarily monotonic ($gs_p, p = 1,\ldots,P$ may have oscillations) nor is it restricted to values in $[0,1]$ (i.e., gs_p may take values outside $[0,1]$) if large correlations between parameter estimates exist. As discussed in [44], the time subinterval during which this transition has the *sharpest increase* corresponds to measurements which provide the most information on possible variations in the corresponding true model parameters.

Since the GSF theory is developed in the context of estimation problems, we considered next an estimation problem using simulated "data". The data used for inversion was generated by first numerically solving the system (9) for the parameter values given in Table 3 and then adding 5% Gaussian white noise to the solution obtained. We consider the problem of using this data to estimate the parameters κ_1, κ_2, κ_3 and κ_4 in an ordinary least squares procedure (with the other parameters and initial conditions fixed at the values from Table 3). The true values for the κ's are $\bar{\kappa} = (1.674, 0.322, 4.521, 1)^T$. For $\theta_0 = \bar{\kappa}$, the generalized sensitivity functions (24) along with the traditional sensitivity functions for the system (9) are presented in Figure 7. In both figures we note a very well defined time subinterval, from $t = 0$ to about $t = 60$, where both GSF and TSF plots exhibit sharp increases/decreases. After this, the TSFs reach very quickly a steady state and the GSFs are forced to approach

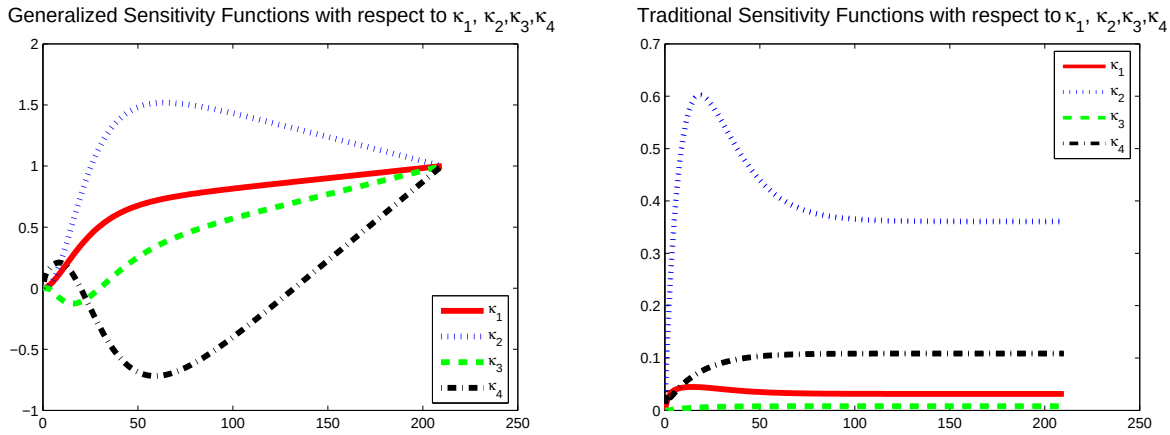

Figure 7: Generalized and Traditional Sensitivity Functions for κ_1, κ_2, κ_3, κ_4, employing a simulated data set with $M = 210$, generated as discussed in the text. Underlying parameter values are $\kappa_1 = 1.674$, $\kappa_2 = 0.322$, $\kappa_4 = 4.521$, $\kappa_4 = 1$, while all other parameters are as given in Table 3.

one. According to the theory, the interval $[0, 60]$ is the time region where measurements are most informative for estimating the true parameters $\bar{\kappa}$. So at least intuitively, sampling more data points in this region would result in more information about the parameters $\bar{\kappa}$ and therefore more accurate estimates for them. By computing the correlation matrix whose elements are given by standard formulas in least squares theory [9, 43], one can also observe that strong correlations exist between estimates for κ_3 and κ_4. In fact, the correlation matrix for these parameters is given by

$Corr$	κ_1	κ_2	κ_3	κ_4
κ_1	1.0000	0.4941	0.0004	0.2209
κ_2	0.4941	1.0000	0.1388	0.1497
κ_3	0.0004	0.1388	1.0000	-0.9502
κ_4	0.2209	0.1497	-0.9502	1.0000

which is in agreement with the dynamics of the curves shown in Figure 7. Positive correlation between κ_1 and κ_2 is clearly indicated as the corresponding gsf graphs increase together while the negative correlation between κ_3 and κ_4 is evidenced by the early opposite slope behavior in their corresponding gsf graphs.

Ordinary least squares inverse problems carried out with different sets of data points illustrate and support the theory. We first performed the least squares minimization for a data set DS_1 consisting of a total of 15 observations, of which 8 were taken at equidistant points in the interval $[0, 60]$ and 7 taken at equidistant points in the interval $[80, 210]$ (for simplicity, we exclude the transition interval $[60, 80]$ from our analysis here). The estimates for κ_1, κ_2, κ_3 and κ_4 along with the $\mathbb{R}^4$ Euclidian norm of the error are displayed in Table 6.

	Data Points in			Estimates for				
Data Set	[0,60]	[80,210]	Total	κ_1	κ_2	κ_3	κ_4	error
DS_1	8	7	15	1.638	0.329	1.775	2.028	2.932
DS_2	15	7	22	1.579	0.305	4.341	0.966	0.207
DS_3	30	7	37	1.708	0.302	2.728	1.178	1.803
DS_4	8	14	22	1.625	0.322	1.830	1.786	2.804
DS_5	8	0	8	1.686	0.337	2.379	1.597	2.224
DS_6	15	0	15	1.736	0.330	6.378	1.064	1.859
DS_7	30	0	30	1.724	0.298	3.549	1.013	0.973
DS_8	0	17	17	12.540	1.966	14.587	9.814	17.314
DS_9	0	33	33	45.554	7.588	69.183	32.589	84.600
DS_{10}	0	66	66	51.161	9.842	92.137	36.002	106.964
True Parameter Values				**1.674**	**0.322**	**4.521**	**1**	

Table 6: Parameter estimates for the transition rates κ_1, κ_2, κ_3 and κ_4 with different data sets.

If we increase the number of data points in the interval $[0, 60]$ from 8 to 15 and keep the number of data points in $[80, 210]$ the same (see DS_2 entry, same table), we observe a significant decrease in the Euclidian norm of the error from 2.932 to 0.207 which represents a significant increase in the accuracy of the parameter estimates. A similar decrease from 2.224 to 1.859 and then to 0.973 is observed when we solve the least squares problem only with data from $[0, 60]$ (see DS_5, DS_6 and DS_7 entries). Thus, numerical calculations support the fact that increasing the number of data points in the region $[0, 60]$ yields more accurate estimates for the parameter $\bar{\kappa}$, in agreement with the theoretical expectations from TSF and GSF.

A totally different outcome is obtained when we carry out numerical estimation with an increasing number of data points in the interval $[80, 210]$. As one can see by comparing the entries DS_1 and DS_4 in Table 6, only a small gain is obtained in the accuracy of the parameter estimates is gained by doubling the number of data points in the interval $[80, 210]$. Moreover, when we try to estimate the parameters $\bar{\kappa}$ with data from the interval $[80, 210]$ alone (see DS_8, DS_9 and DS_{10}), we obtain very large errors which increase in magnitude as the number of sample points increases. Although puzzling at first view, this phenomenon is not surprising at all from the perspective of the theory presented above and that presented in [6]. Indeed, by blindly sampling more data points from the region where the generalized sensitivity functions exhibit the so called "forced-to-one" artifact and the traditional sensitivity curves are flat, we simply introduce redundancy in the sensitivity matrix, thus increasing the condition number of the Fisher information matrix for our problem. For an illustration and discussion of this phenomenon, see [6]. By the Cramér-Rao inequality, the consequence is that the variance of the unbiased estimator (and the corresponding standard errors) will be huge, making our estimates less useful. The same phenomenon (introduction of redundancy in the sensitivity

matrix) is responsible for the poorer results which are obtained when we estimate $\bar{\kappa}$ using data set DS_3 (the Euclidian norm of the error increased as compared to DS_2, where we doubled the number of points in $[0, 60]$). We observe an important drawback of the GSF: while they specify the most informative regions with respect to the estimation of parameters, they do not provide any information about the necessary number of data points to be used in those regions.

We remark that in this section we have illustrated a methodology to quantify the sensitivity with respect to parameters and initial conditions of a large complex system. This methodology is a fundamental tool in identifying those parameters in a model which require further, more in-depth investigation as well as the basis of a statistical analysis for quantifying uncertainty in estimators (e.g., see [5, 6, 9, 11, 43]) as well as information content based model selection techniques [8].

4 Foot-and-Mouth Disease

Having established a basic model for the movement of animals (pigs) in the agricultural system (the pork industry of North Carolina), we are now ready to model the spread of an infectious disease in the food production network. In this section we describe the incorporation of an SIR-type infection into the system and present simulations to illustrate the spread of Foot and Mouth Disease throughout the aggregated agricultural network.

We describe the infection by an SIR process [2, 12]. It is assumed that a population can be partitioned into three groups: susceptible (S), infectious (I) and removed (R). In many settings the removed class represents individuals who have recovered from the infection. Individuals move between these classes as they become infected and recover from infection. Recovery is assumed to confer permanent immunity to infection and the demography of the population (i.e., births and non disease-related deaths) is ignored. In a well-mixed population the epidemiological model can be described by the flowchart of Figure 8 and equations (25).

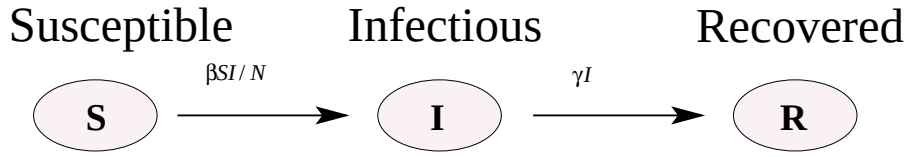

Figure 8: Flow diagram of the simple SIR model.

$$\begin{aligned} \frac{dS}{dt} &= -\frac{\beta SI}{N} \\ \frac{dI}{dt} &= \frac{\beta SI}{N} - \gamma I \\ \frac{dR}{dt} &= \gamma I. \end{aligned} \tag{25}$$

Here, S, I and R denote the numbers of susceptible, infective and removed individuals, respectively. The transmission parameter is β: this parameter, when combined with the rate at which individuals meet each other and the probability that an infective would infect a susceptible during any one such meeting, yields the transmission rate. It is assumed that recovery occurs at constant rate γ, so that $1/\gamma$ is the average duration of infection. The population size is denoted by N, and we have in this case that $N = S + I + R$.

The behavior of the simple SIR model is governed by the basic reproductive number, R_0 [2, 12, 31, 14]. This quantity equals the number of secondary infections caused by the introduction of a single infectious individual into an otherwise completely susceptible population. In terms of model parameters, the basic reproductive number is given by

$$R_0 = \beta/\gamma.$$

An outbreak of infection can only ever occur if R_0 is greater than one, otherwise the number of infectives can never increase.

We now combine the agricultural network model and the SIR model to produce a description of the potential spread of an SIR-type infection in our agricultural system. It is most convenient for us to work with the numbers of individuals of each type found at the various nodes of the network, and so we convert the concentrations of animals of the network model (9) into numbers. We write the number of individuals found at node i as N_i, and so we have that $N_i(t) = Nc_i(t)$.

We expand the first three nodes of the network, i.e., those describing the growers/sows, nurseries and finishers, by including an SIR model within each of them. The infection statuses of the animals at the ith node are tracked by the quantities S_i, I_i and R_i. Since N_i is defined to be the total number of animals at this node we have that $N_i = S_i + I_i + R_i$.

We make the following set of assumptions about the infection process and its impact upon the agricultural system:

(1) Pigs are born healthy, but susceptible. Piglets are introduced into the network at the first node.

(2) There is no infection or recovery during transport between the nodes. This is based on the idea that, in most cases, transportation takes no more than a couple of days, which is relatively short compared to the amount of time the pigs spend at each node.

(3) There is no infection in the slaughter node. In the absence of human intervention this does not mean that the infectious animals are not processed: the assumption is that the infection is not propagated after the animal's death.

(4) Infected pigs in node i recover at rate γ_i. These recovery rates may differ between the nodes and, since the individuals found at different nodes will have different ages, these parameters depict age-dependent recovery rates.

(5) Recovered pigs have temporary immunity, i.e., pigs do not immediately become susceptible after recovering from FMD. The rates at which recovered pigs become susceptible

again are ρ_i ($i = 1, 2$ or 3). While some diseases afford permanent immunity to the recovered animal, immunity acquired after FMD infection wanes in a matter of months. If vaccination of animals were considered there would be an accumulation of animals in the recovered groups of the appropriate nodes.

(6) We assume that the FMD-related death rate is small enough that we can ignore such deaths. Control strategies such as culling infective or susceptible animals could be modeled by including death terms.

(7) Since our network model assumes that the system is closed, we assume that any deaths are replenished by the introduction of piglets into the network. As already mentioned, these introductions occur into the first node. Consequently, as illustrated in the flowchart, Figure 9, the slaughtered animals that leave node four effectively return to the S_1 class. Similarly, if animal culling were being considered, the model would include flows from the appropriate classes back to S_1.

(8) No human intervention. In the model as presented here, we assume that humans do not make any adjustments to operation of the agricultural system in response to the infection: animal movement and processing continues as normal. Of course, one of the main aims of creating a model such as this is to enable the consideration of control strategies. In this study, we do not do this as we wish to first establish the baseline (no-control) behavior of the system.

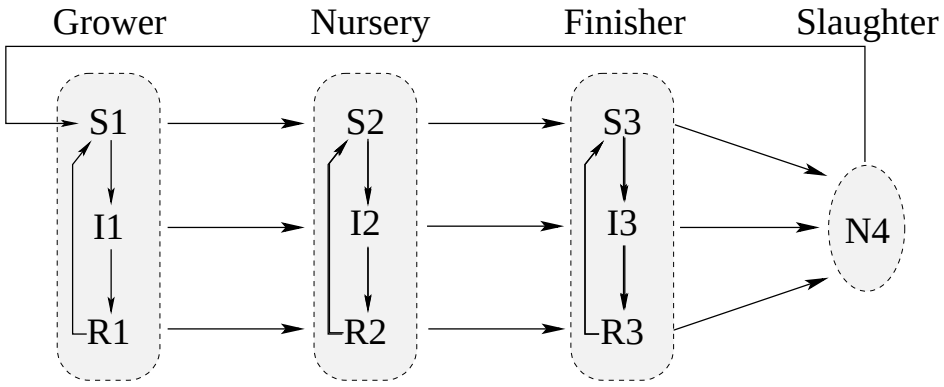

Figure 9: Flow diagram of the aggregated agricultural network model with SIR infection.

With this set of assumptions we obtain the following system of equations for the deterministic model of the aggregated agricultural network model with an SIR infection.

$$
\begin{aligned}
\frac{dS_1}{dt} &= -\beta_1 S_1 I_1/N_1 + \rho_1 R_1 - k_1 S_1 (L_2 - N_2)_+ + k_4 \min(N_4, S_{\max}) \\
\frac{dI_1}{dt} &= \beta_1 S_1 I_1/N_1 - \gamma_1 I_1 - k_1 I_1 (L_2 - N_2)_+ \\
\frac{dR_1}{dt} &= \gamma_1 I_1 - \rho_1 R_1 - k_1 R_1 (L_2 - N_2)_+ \\
\\
\frac{dS_2}{dt} &= -\beta_2 S_2 I_2/N_2 + \rho_2 R_2 - k_2 S_2 (L_3 - N_3)_+ + k_1 S_1 (L_2 - N_2)_+ \\
\frac{dI_2}{dt} &= \beta_2 S_2 I_2/N_2 - \gamma_2 I_2 - k_2 I_2 (L_3 - N_3)_+ + k_1 I_1 (L_2 - N_2)_+ \\
\frac{dR_2}{dt} &= \gamma_2 I_2 - \rho_2 R_2 - k_2 R_2 (L_3 - N_3)_+ + k_1 R_1 (L_2 - N_2)_+ \\
\\
\frac{dS_3}{dt} &= -\beta_3 S_3 I_3/N_3 + \rho_3 R_3 - k_3 S_3 (L_4 - N_4)_+ + k_2 S_2 (L_3 - N_3)_+ \\
\frac{dI_3}{dt} &= \beta_3 S_3 I_3/N_3 - \gamma_3 I_3 - k_3 I_3 (L_4 - N_4)_+ + k_2 I_2 (L_3 - N_3)_+ \\
\frac{dR_3}{dt} &= \gamma_3 I_3 - \rho_3 R_3 - k_3 R_3 (L_4 - N_4)_+ + k_2 R_2 (L_3 - N_3)_+ \\
\\
\frac{dN_4}{dt} &= -k_4 \min(N_4, S_{\max}) + k_3 N_3 (L_4 - N_4)_+ .
\end{aligned}
$$

In many ways, this model resembles a standard multi-group epidemiological model and so we might hope to be able to find the basic reproductive number of the system using standard multi-group methodology [12, 13]. It is straightforward to find the next generation matrix, whose entries, t_{ij}, give the average numbers of secondary infections that result in node i from the introduction of one infectious individual into node j when all individuals at node i are susceptible. For within-patch transmission, we have

$$
\begin{aligned}
t_{33} &= \frac{\beta_3}{\gamma_3 + k_3 (L_4 - N_4)_+} \\
t_{22} &= \frac{\beta_2}{\gamma_2 + k_2 (L_3 - N_3)_+} \\
t_{11} &= \frac{\beta_1}{\gamma_1 + k_1 (L_2 - N_2)_+} .
\end{aligned}
$$

Here, the term $1/(\gamma_i + k_i (L_{i+1} - N_{i+1})_+)$ is the average duration of infection for infectives in node i, corrected for their departure on account of transportation. Movement of individuals between nodes reduces the average number of within-node secondary infections, with this effect being most noticeable if the node residence time is short compared to the duration of

infection. For between-node transmission we have

$$\begin{aligned} t_{32} &= \frac{k_2(L_3-N_3)_+}{\gamma_2+k_2(L_3-N_3)_+}t_{33} \\ t_{21} &= \frac{k_1(L_2-N_2)_+}{\gamma_1+k_1(L_2-N_2)_+}t_{22} \\ t_{31} &= \frac{k_1(L_2-N_2)_+}{\gamma_1+k_1(L_2-N_2)_+}\frac{k_2(L_3-N_3)_+}{\gamma_2+k_2(L_3-N_3)_+}t_{33} \\ t_{12} &= t_{23} = t_{13} = 0. \end{aligned}$$

Here, the term $k_i(L_{i+1}-N_{i+1})_+/(\gamma_i+k_i(L_{i+1}-N_{i+1})_+)$ gives the probability that an infectious individual in node i is transported to node $i+1$ before it recovers.

Multi-group theory reveals that, for an irreducible system (i.e., one in which infection can travel between any pair of nodes, possibly via intermediate nodes), the basic reproductive number is given by the dominant eigenvalue of this matrix. Our system is not irreducible, however, since infected animals can only move from a node to the following node: infection can spread to succeeding nodes but not preceding ones. Consequently, the standard definition of R_0 for multi-group systems is not helpful to us. From the t_{ij} we can, however, easily calculate the average number of secondary infections caused by the introduction of a single infective into node j when all other individuals in the population are susceptible. These quantities equal $t_{11}+t_{21}+t_{31}$, $t_{22}+t_{32}$ and t_{33}, for nodes 1, 2 and 3, respectively.

We now turn to numerical simulation of the model. Unfortunately, the existing studies in the literature do not provide us with an appropriate set of parameter values to use. Epidemiological parameters for the spread of FMD between various types of animals, including pigs, have been quantified [17, 18, 22, 23, 24, 32, 33, 34, 38, 39, 46] but on spatial scales that are quite different from our aggregated network description. On a large spatial scale, transmission between farms has been described, for instance during the 1967/68 and 2001 outbreaks in the UK, and parameter estimates obtained [22, 23, 24, 32, 33, 34, 46]. Large-scale studies, however, take the individual unit of the model to be farms and so do not consider transmission between individual animals.

On a small spatial scale, detailed transmission experiments [17, 18, 39, 38] have examined the spread of infection between small numbers of closely housed animals, either within or between pens. (Given the earlier comment regarding age-dependent epidemiological parameters, it is interesting to note that some of these experiments have been carried out on pigs of different ages.) These experiments, which typically involve placing one or more infected animals in close contact with a number of susceptible animals, demonstrate the high degree of infectiousness of FMD. In many instances every susceptible animal became infected [17]. Instances in which all animals become infected provide less informative estimates of R_0 than might be hoped since the statistical methodology employed gives an infinite estimate for R_0. An alternative statistical approach [18] accounts for the time dependence in the experiment and provides estimates of the transmission parameter β.

The literature provides us with satisfactory estimates for the average durations of infectiousness and immunity [28], but not R_0 (or, equivalently, β). For illustrative purposes, we

shall take R_0 to equal 10 at each level of the network in the absence of transportation.

Parameters for the Aggregated Agricultural Network Model with SIR				
Parameters	Definition	Values	Units	Reference
R_0	basic reproductive number	10		
$\gamma_1, \gamma_2, \gamma_3$	recovery rate	1/31,1/31,1/31	1/day	[28]
ρ_1, ρ_2, ρ_3	rate of loss of immunity	1/180,1/180,1/180	1/day	[28]

All the simulations that follow assume that the network is initially in its demographic steady state. The epidemiological assumptions that we have made lead to this holding for all future times. (This would not be the case if we had deaths at the nursery or finisher nodes.) We choose to introduce infection by infecting a certain number of individuals in the initial node.

Our first simulation (Figure 10) illustrates the movement of a cohort of infectious individuals through the network in the absence of ongoing transmission. We set $\beta_1 = \beta_2 = \beta_3 = 0$, and introduce infection by placing $50,000$ infected piglets in the grower node.

In this simulation infected animals either recover or get transported to the next node. In the grower node, the number of infected animals simply decreases exponentially, while the number of recovered animals increases and then declines. Within a fairly short time period, the grower population is entirely replaced by susceptible individuals, reflecting the rapid turnover of the grower node. We see the appearance of infection first at the nursery and then at the finisher node. Since the system is at demographic equilibrium, we see no change at the slaughter node.

For our second simulation (Figure 11), we assume that, in the absence of transportation, the infection would have basic reproductive number equal to 10 at each node. Consequently, we set $\beta_i/\gamma_i = 10$. As before, we take the initial population to be in demographic equilibrium, but we now introduce just 200 infective piglets into the grower node.

The system rapidly approaches an equilibrium in which infectious animals are present at each node. The fraction of animals that are susceptible at equilibrium is much higher at the grower node (24.8%) than at either the nursery (6.9%) or finisher (7.9%). This should be expected since all animals arriving at the grower are susceptible, while the infection statuses of those entering the nursery or finisher reflect the composition of the preceding node (i.e., grower and nursery nodes, respectively). For example, only 25% of the arrivees at the nursery are susceptible. The susceptible fraction at the finisher is slightly higher than at the nursery due to the replenishment of susceptibles by loss of immunity: this relatively slow process has a higher chance of occurring at the finisher since an animal's average residence time there is longer than at the nursery.

The differing susceptible fractions between the nodes mean that, relative to the size of the population of each node, there is less ongoing transmission at either the nursery or finisher node than at the grower. The equilibrium infective fraction decreases as the supply chain is traversed (46.8%, 31.5% and 16.2% at the grower, nursery and finisher, respectively), while the recovered fraction increases (28.4%, 61.5% and 75.9%).

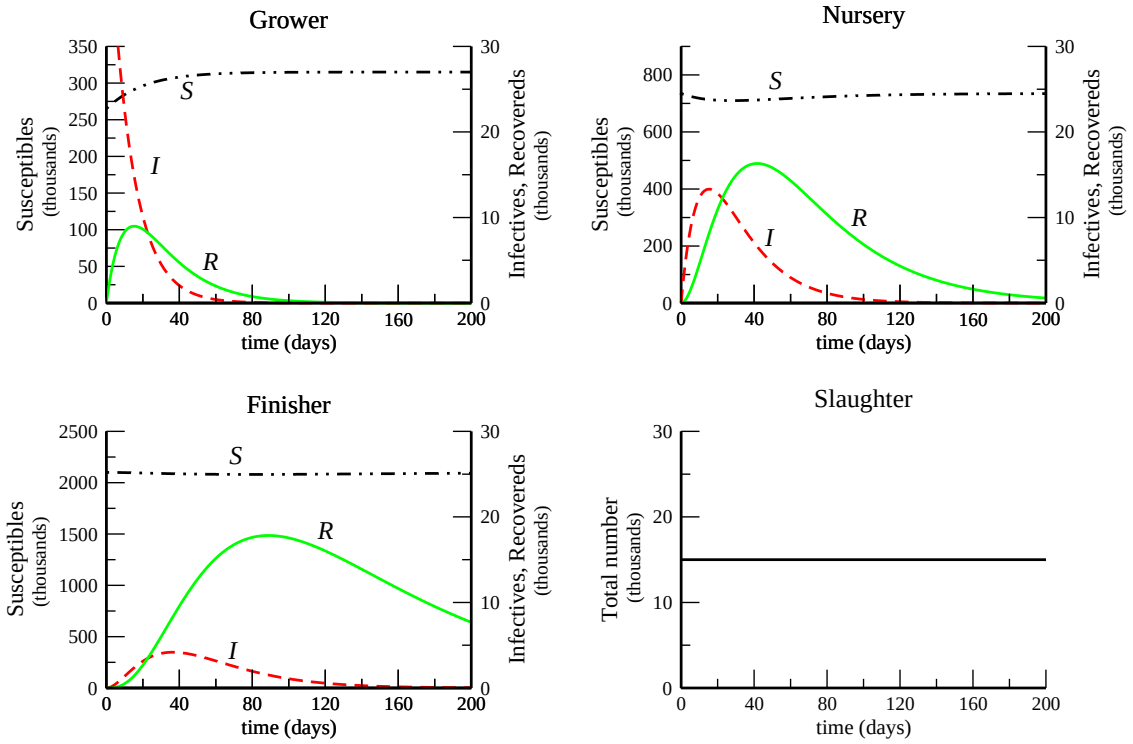

Figure 10: Simulation I: Recovery and movement of infectious individuals through the network, in the absence of transmission. Notice that susceptible numbers are plotted on different scales (left vertical axis on each panel) from infective and recovered numbers (right vertical axis on each panel). Parameter values are given in the text: each β_i is set equal to zero, infection is assumed to last 31 days and immunity lasts 180 days. The initial conditions are $N_1 = 315,000$, $N_2 = 735,000$, $N_3 = 2,100,000$ and $N_4 = 15,000$. Of the piglets at the first node, $50,000$ are initially taken to be infectious, while the remainder of the population is susceptible.

The impact of age-dependent transmission rates and the differing residence times is explored in our third simulation (Figure 12). Again only 200 infectives were introduced but now we assume that younger animals are less infectious than older animals. The basic reproductive numbers, in the absence of transportation, at the grower, nursery and finisher nodes are taken to equal two, ten and fifteen, respectively. (We exaggerate the age-dependent differences in R_0 for illustrative purposes.)

Transportation has a major impact on the dynamics of the infection: the movement of individuals out of the grower node is sufficiently rapid that the infection cannot persist in this node. The number of infectives falls roughly exponentially, as does the number of recovereds, although the latter only occurs after an initial rise (which largely reflects the recovery of the initial pool of infectives). The dynamics in this node are somewhat reminiscent of those seen in the first simulation (in which there was no transmission), although they play out on a slower timescale since there is some ongoing transmission.

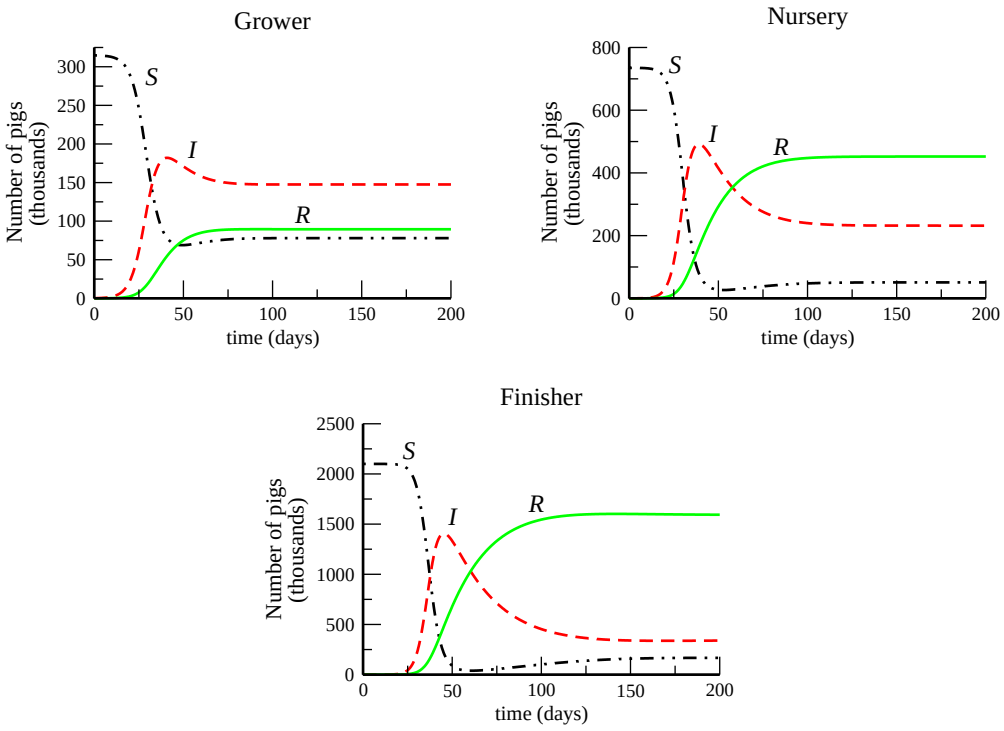

Figure 11: Simulation II: Dynamics following the introduction of infection, with $R_0 = 10$. This value of the basic reproductive number determines the value of the transmission parameters (β_i). Initial conditions and other parameter values are as in the previous figure, except that only 200 of the 315,000 individuals at the grower node are infective at the initial time.

Disease transmission at the nursery and finisher nodes occurs sufficiently rapidly that the prevalence of infection approaches a positive, endemic, equilibrium at both. Observe that, even though the transmission parameter in the nursery node is the same as it was in the previous simulation, the equilibrium numbers of susceptibles and infectives are higher here than they were in the previous simulation. This reflects the differences between the compositions of the populations entering the nursery node in the two simulations, with more susceptibles arriving in the age-dependent situation.

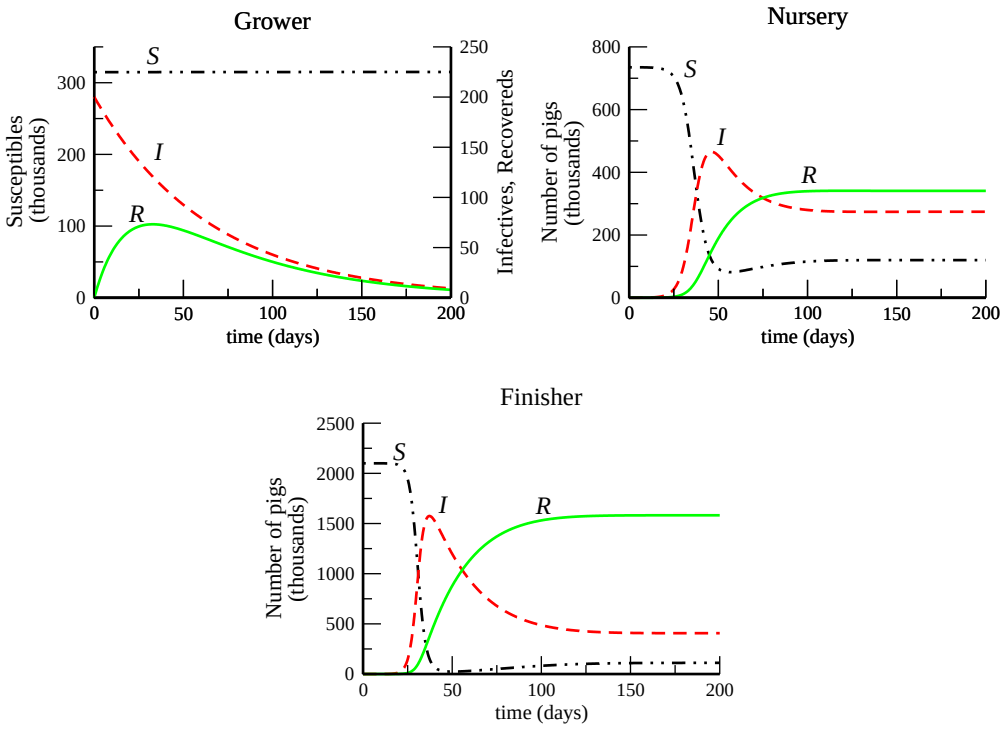

Figure 12: Simulation III. Age-dependent transmission of infection. If there were no transportation of animals, the basic reproductive numbers at the grower, nursery and finisher nodes would equal 2, 10 and 15, respectively. These values of R_0 determine the parameters β_1, β_2 and β_3. All other parameter values and initial conditions are as in the previous figure. For the panel depicting the grower, note that susceptible numbers are plotted on a different scale (left vertical axis) from the infective and recovered numbers (right vertical axis), and these infective and recovered numbers are *not* in thousands. Also note that infection goes extinct in the grower node while a positive equilibrium level of infection is achieved at the nursery and finisher nodes.

5 Concluding Remarks

In this paper we have demonstrated a methodological approach to the investigation of production networks and their vulnerability to disturbances such as diseases. The stochastic model and the resulting approximate deterministic system we employ were shown to agree well, but are not validated. Rather, we carry out simulations and sensitivity analyses with parameter values that are only loosely based on a swine network. We use the deterministic model to show how to determine the parameters to which the model at these parameter values exhibits the most sensitivity. Finally, we demonstrate how disease can be introduced and the resulting network vulnerabilities analyzed. An interesting next step would involve obtaining experimental data to validate and perhaps improve the model for a specific production network. This would require using inverse problem algorithms with the data to obtain estimates along with measures of associated uncertainty (e.g., standard errors [9, 43]) for the underlying transition rates k_i.

Other obvious questions for further investigation involve the introduction of stochasticity in the transmission dynamics of the infection. It is well known, for example, that random effects can have a major impact on the invasion of an infection into a population. The use of constant rates of recovery and loss of immunity for the infection should also be questioned. These assumptions, which are biologically unrealistic for many infections, could be important in cases such as the one presented above where the life span of the animals is comparable in length to the duration of infection and immunity.

The randomness seen in the stochastic network model originates from the random movement of discrete individuals from node to node. The analysis of Sections 2.2 and 3.2 shows that, due to an averaging effect, these random effects become less important as the system size N increases. Application of the stochastic transportation model to describe a real-world situation should, therefore, account for the size of the groups in which pigs are transported between nodes. If, for example, one thousand pigs were moved at a time, the appropriate notion of an "individual" within the model would be a thousand pigs. Treating each group of a thousand animals as a unit would lead to a marked increase in the magnitude of stochastic fluctuations seen at the population level. Consequently, even though the system size in our simulations is on the order of millions of pigs, it might be that the resulting stochastic fluctuations in a more realistic model of the production system are closer to those shown in Figure 3 than to those of Figure 2.

The approach outlined in this paper has rather obvious potential for application to a wide range of problems. These include the investigation of the spread of diseases through spatially or structurally distributed dynamic populations (e.g., avian flu through migrating bird populations, contagious infections through highly mobile and/or age-structured human and animal populations). In some of these cases the natural nodal structure would be a continuum, requiring stochastic and deterministic models with a continuum of spatial/structural heterogeneities, leading to partial differential equation systems. Such applications would undoubtedly motivate the development of interesting new stochastic and deterministic mathematical and computational methodologies.

We also note that the approach and methodology presented here are useful for investigation of a wide range of perturbations other than disease (e.g., loss of capacity at a given node such as a factory being shut down for some reason) in supply networks. In particular they would be useful in the assessment of risk of a food-borne pathogen (e.g., salmonella, listeria, etc.) entering the food chain [27] due to contamination (either accidental or deliberate) at some stage of the supply chain.

Acknowledgements

This research was supported in part by the Statistical and Applied Mathematical Sciences Institute, which is funded by NSF under grant DMS-0112069, in part by the National Institute of Allergy and Infectious Disease under grant 9R01AI071915-05, and in part by the U.S. Air Force Office of Scientific Research under grant AFOSR-FA9550-04-1-0220.

References

[1] G. Almond (College of Veterinary Medicine, NC State University), R. Baker (College of Veterinary Medicine, Iowa State University), M. Battrell (Murphy Farms, NC), J. Tickel (Emergency Programs, NC Dept of Agriculture and Consumer Services), personal communications.

[2] R. M. Anderson and R. M. May, *Infectious Diseases of Humans: Dynamics and Control*, Oxford University Press, Oxford, 1991.

[3] H. Andersson and T. Britton, *Stochastic Epidemic Models and Their Statistical Analysis*, Lec Notes in Statistics, **151**, Springer Verlag, New York, 2000.

[4] H. T. Banks, S. Dediu and S.L. Ernstberger, Sensitivity functions and their use in inverse problems, to appear.

[5] H. T. Banks, S. Dediu and H.K. Nguyen, Sensitivity of dynamical systems to parameters in a convex subset of a topological vector space, Center for Research in Scientific Computation Report, CRSC-TR06-25, September, 2006, North Carolina State University; *Math. Biosci. and Engineering,* submitted.

[6] H. T. Banks, S.L. Ernstberger and S.L. Grove, Standard errors and confidence intervals in inverse problems: sensitivity and associated pitfalls, CRSC Tech Rep. CRSC-TR06-10, N.C. State University, March, 2006; *J. Inverse and Ill-posed Problems*, to appear.

[7] J. J. Batzel, F. Kappel, D. Schneditz and H.T. Tran, *Cardiovascular and Respiratory Systems: Modeling, Analysis and Control*, SIAM Frontiers in Applied Math, SIAM, Philadelphia, 2006.

[8] K. P. Burnham and D. R. Anderson, *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*, Springer-Verlag, New York, 2002.

[9] G. Casella and R. L. Berger, *Statistical Inference*, Duxbury, Pacific Grove, CA, 2002.

[10] J.B. Cruz, ed., *System Sensitivity Analysis*, Dowden, Hutchinson & Ross, Inc., Stroudsberg, PA, 1973.

[11] M. Davidian and D. Giltinan, *Nonlinear Models for Repeated Measurement Data*, Chapman & Hall, London, 1998.

[12] O. Diekmann and J.A.P. Heesterbeek, *Mathematical Epidemiology of Infectious Diseases*, John Wiley & Sons, Chichester, 2000.

[13] O. Diekmann, J.A.P. Heesterbeek and J.A.J. Metz, On the definition and computation of the basic reproduction ratio R_0 in models of infectious diseases in heterogeneous populations, *J. Math. Biol.*, **28** (1990), 365–382.

[14] K. Dietz, The estimation of the basic reproductive number for infectious diseases, *Stat. Meth. Med. Res.*, **2** (1993), 23–41.

[15] W. Dörfer-Kreissl, Report of measures to control Foot and Mouth Disease in the European Union in 2001 and future measures to prevent and control animal diseases in the European Union, *European Parliament Session Document A5-0405/2002*, 28Nov2002, p.45 (para 1,5), p.52 (para 1).

[16] R. Durrett, *Essentials of Stochastic Processes*, Springer, New York, 1999.

[17] P.L. Eble, A. Bouma, M.G. de Bruin, F. van Hemert-Kluitenberg, J.T. van Oirschot and A. Dekker, Vaccination of pigs two weeks before infection significantly reduces transmission of foot-and-mouth disease virus, *Vaccine*, **22** (2004), 1372–1378.

[18] P.L. Eble, A. de Koeijer, A. Bouma, A. Stegeman and A. Dekker, Quantification of within-and between-pen transmission of Foot-and-Mouth disease virus in pigs, *Vet. Res.*, **37** (2006), 647–654.

[19] A. El Amin, Round up: Denmark hit by deadly strain of bird flu, FoodProductionDaily.Com— Europe; News Headlines-Supply Chain, 17Mar2006, http://www.foodproductiondaily.com/news/ng.asp?id=66483-bird-flu-avian-france

[20] M. Eslami, *Theory of Sensitivity in Dynamic Systems: An Introduction*, Springer-Verlag, Berlin, 1994.

[21] S. N. Ethier and T. G. Kurtz, *Markov Processes: Characterization and Convergence*, J. Wiley & Sons, New York, 1986.

[22] N. M. Ferguson, C. A. Donnelly and R. M. Anderson, The foot and mouth epidemic in Great Britian: Pattern of spread and impact of interventions, *Science* **292** (2001), 1155–1160.

[23] N. M. Ferguson, C. A. Donnelly and R. M. Anderson, Transmission intensity and impact of control policies on the foot and mouth epidemic in Great Britain, *Nature* **413** (2001), 542–548.

[24] M. J. Ferrari, O. N. Bjørnstad, A. P. Dobson, Estimation and inference of R_0 of an infectious pathogen by a removal method, *Math. Biosci.*, **198** (2005), 14–26.

[25] P.M. Frank, *Introduction to System Sensitivity Theory*, Academic Press, Inc., New York, NY, 1978.

[26] C. Fraser, S. Riley, R. M. Anderson and N. M. Ferguson, Factors that make an infectious disease outbreak controllable, *Proc. Natl. Acad. Sci. USA* **101** (2004), 6146–6151.

[27] M. A. van der Gaag, F. Vos, H. W. Saatkamp, M. van Boven, P. van Beek and R.B.M. Huirne, A state-transition simulation model for the spread of Salmonella in the pork supply chain, *European J. Operational Res.*, **156** (2004), 782–798.

[28] M. G. Garner and M. B. Lack, An evaluation of alternate control strategies for foot-and-mouth disease in Australia: A regional approach, *Prev. Vet. Med.*, **23** (1995), 9–32.

[29] D. T. Gillespie, The chemical Langevin equation, *J. Chemical Physics* **113** (2000), 297–306.

[30] G. R. Grimmett and D. R. Stirzaker, *Probability and Random Processes*, Clarendon Press, Oxford, 1982.

[31] J. A. P. Heesterbeek, A brief history of R_0 and a reciepy for its calculation, *Acta Biotheor.*, **50** (2002), 189–204.

[32] D.T. Haydon, M.E.J. Woolhouse and R.P. Kitching, An analysis of foot-and-mouth-disease epidemics in the UK, *IMA J. Math. Appl. Med. Biol.*, **14** (1997), 1–9.

[33] S.C. Howard and C.A. Donnelly, The importance of immediate destruction in epidemics of foot and mouth disease, *Res. Vet. Sci.*, **69** (2000), 189–196.

[34] M.J. Keeling, M.E.J. Woolhouse, D.J. Shaw, L. Matthews, M. Chase-Topping, D.T. Haydon, S.J. Cornell, J. Kappey, J. Wilesmith and B.T. Grenfell, Dynamics of the 2001 UK foot and mouth epidemic: Stochastic dispersal in a heterogeneous landscape, *Science*, **294** (2001), 813–817.

[35] W. O. Kermack and A. G. McKendrick, A contribution to the mathematical theory of epidemics, *Proc. Roy. Soc. Lond.*, **A 115** (1927), 700–721.

[36] M. Kleiber, H. Antunez, T.D. Hien and P. Kowalczyk, *Parameter Sensitivity in Nonlinear Mechanics: Theory and Finite Element Computations*, John Wiley & Sons, New York, NY, 1997.

[37] *Living Countryside*, AgriStats - UK Agriculture - farming statistics, http://www.ukagriculture.com/statistics/farming statistics.cfm.

[38] K. Orsel, A. Dekker, A. Bouma, J.A. Stegeman and M.C.M. de Jong, Vaccination against foot and mouth disease reduces virus transmission in groups of calves, *Vaccine*, **23** (2005), 4887–4894.

[39] K. Orsel, M.C.M. de Jong, A. Bouma, J.A. Stegeman and A. Dekker, The effect of vaccination on foot and mouth disease virus transmission among dairy cows, *Vaccine*, **25** (2007), 327–335.

[40] *R: A Language and Environment for Statistical Computing*, Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0.

[41] J. Rushton, T. Willmore, A. Shaw and A. James, Economic analysis of vaccination strategies for foot and mouth disease in the UK, *Royal Society Inquiry into Infectious Diseases in Livestock*, 2002, p21, Tables 3 and 4, http://www.royalsoc.ac.uk/inquiry/index/566.pdf.

[42] A. Saltelli, K. Chan and E.M. Scott, eds., *Sensitivity Analysis*, Wiley Series in Probability and Statistics, John Wiley & Sons, New York, NY, 2000.

[43] G. A. F. Seber and C. J. Wild, *Nonlinear Regression,* John Wiley & Sons, Inc., New York, 1989.

[44] K. Thomaseth and C. Cobelli, Generalized sensitivity functions in physiological system identification, *Ann. Biomed Eng.*, **27** (1999), 607–616.

[45] D. Thompson, P. Muriel, D. Russell, P. Osborne, A. Bromley, M. Rowland, S. Creigh-Tyte and C. Brown, Economic costs of the foot and mouth disease outbreak in the United Kingdom in 2001, *Rev. Sci. Tech. Off. Int. Epiz.*, **21** (2002), 675–687.

[46] M. Woolhouse, M. Chase-Topping, D. Haydon, J. Friar, L. Matthews, G. Hughes, D. Shaw, J. Wilesmith, A. Donaldson, S. Cornell, M. Keeling, and B. Grenfell, Foot-and-mouth disease under control in the UK, *Nature* **411** (2001), 258–259.