

Pieter Trapman · Ronald Meester · Hans Heesterbeek

A branching model for the spread of infectious animal diseases in varying environments

Received: 14 July 2003 / Revised version: 14 November 2003 /
Published online: 3 March 2004 – © Springer-Verlag 2004

Abstract. This paper is concerned with a stochastic model, describing outbreaks of infectious diseases that have potentially great animal or human health consequences, and which can result in such severe economic losses that immediate sets of measures need to be taken to curb the spread. During an outbreak of such a disease, the environment that the infectious agent experiences is therefore changing due to the subsequent control measures taken. In our model, we introduce a general branching process in a changing (but not random) environment. With this branching process, we estimate the probability of extinction and the expected number of infected individuals for different control measures. We also use this branching process to calculate the generating function of the number of infected individuals at any given moment. The model and methods are designed using important infections of farmed animals, such as classical swine fever, foot-and-mouth disease and avian influenza as motivating examples, but have a wider application, for example to emerging human infections that lead to strict quarantine of cases and suspected cases (e.g. SARS) and contact and movement restrictions.

1. Introduction

Recent outbreaks of infectious diseases of animals (e.g. classical swine fever (CSF), foot and mouth disease (FMD) and avian influenza (AI)) in Western Europe have had great impact on the economy, public life and animal health and welfare in the countries involved. During such an outbreak one would like to be able to compare the effectiveness of proposed control measures in, for example, their ability to reduce the expected final size and the expected duration of the outbreak. Typical for the strategies aimed at stopping outbreaks of important diseases of farm animals, is that infected herds are removed from the population by culling upon detection. A second characteristic is that due to increasing quantity and quality of the imposed control measures, the environment that the infectious agent experiences, is changing. By this we mean that consecutive measures can make, for example, contact opportunities between herds different in different phases of the outbreak, or can

J. P. Trapman, J. A. P. Heesterbeek: Faculty of Veterinary Medicine, Utrecht University, Yalelaan 7, 3584 CL Utrecht, The Netherlands. e-mail: trapman@vet.uu.nl

J. P. Trapman, R. W. J. Meester: Division of Mathematics, Vrije Universiteit Amsterdam, De Boelelaan 1081a, 1081 HV Amsterdam, The Netherlands

Key words or phrases: Stochastic epidemic – Branching process – Varying environments – Iterative method

make the infectious period, or rate with which infectivity is produced, differ for farms infected at different times.

In most cases, mathematical methods for computing outbreak characteristics such as expected final size and expected duration, assume a constant environment in that the control measures are not compounded in time and do not lead to changes in the rates that govern epidemic spread (see e.g. [2]). In this paper we aim to develop stochastic methods, based on branching processes, which allow us to compare the effectiveness of control strategies during such outbreaks in situations where the environment is varying because of changes in subsequent measures of control.

Much work has already been done to describe the spread of classical swine fever (see e.g. [10, 11, 1]) and foot and mouth disease (see e.g. [5, 6, 9, 8]). In this paper, we will model the spread of infections in a much more analytic way than is done in earlier models [5, 6, 9, 1]. We use an iterative method that computes properties of the spread, like the probability of a major outbreak of the infection and the final size of an epidemic (i.e. the total number of infected herds) very efficiently. Furthermore we can derive some properties of the duration of the epidemic. We allow for different types of herd.

In our model, it is essential that once the infection in a herd is detected, the whole herd will be culled. Therefore, our main interest is the number of infected herds, but the number of infected animals in an infective herd is important for determining the infectivity of a herd and the distribution of the detection times. We therefore model the spread of the infection on two levels, namely the spread of the infections *within* a herd and the spread of infection *between* herds; both are described by a stochastic process.

We use a special branching process to describe the spread of the infection among herds. The parameters of this branching process depend on the time since the infection of the herd and on the environment, which is determined by the real time. Using branching processes to describe epidemics is of course not new (see e.g. [4]), but no theory exists that gives short-term predictions (as opposed to asymptotics) for general branching processes in varying environments, with an age-dependent birth rate.

For computations, it is necessary that after a certain moment the environment is constant. To achieve this we assume that after some time no new measures will be taken and the effects of all measures taken in the past will either be constant or absent. In other words, although the values of the parameters may differ from those before certain measures were taken, the values are assumed to be constant after a given moment in time.

Our model can be used to predict the effects of various control measures and strategies during an ongoing outbreak. Meester et al. gave a method to estimate the parameters from the data available during an outbreak, [10]. We consider measures like single vaccination of all herds of a certain type, a total transport ban or killing of animals just after birth. Some of these measures cause a varying environment. The fraction susceptible animals in a herd or the fraction susceptible herds of the total number of herds may be varying and so the infection rate may vary in time as well. In fighting outbreaks, additional measures will be implemented as soon as present measures turn out to be insufficient. A change in measure will change

the values of the parameters. We assume that the changes in the environment are *deterministic*.

We develop the theory using a classical swine fever outbreak in herds of pigs as motivating example throughout. In our paper we use the same input data as Klinkenberg et al. [1].

2. Spread of the infection within a herd

2.1. The model

As mentioned in the introduction, we first need to model the spread of the infection within one herd, since the infectivity of an infective herd, and also the time at which the infection is detected in a certain herd, depend on the number of infected animals in that herd. We use t for the time elapsed since the first measures were implemented. The variable τ is used for describing the spread of the infection within the herds, it is the time since infection of a particular herd and therefore relative. The τ -clock starts ticking at the moment the first animal in the herd becomes infected. From now on, we will call τ the “infection-age” or “age” of the herd.

We use only four types of disease-related parameters:

- μ : The recovery rate of individual animals in a herd.
- λ : The infection rate of individual animals within a herd.
- α : The per capita detection rate of infected animals within a herd.
- β : The rate at which one infected animal infects susceptible herds.

We assume that as soon as an infection at a particular herd is detected, the whole herd will be culled instantaneously. Further we assume that the rate of detection is proportional to the number of infected animals at infection-age τ , $I(\tau)$, i.e. the detection rate is $\alpha I(\tau)$, where α does not depend on the infection-age. We assume that the infection in a herd develops as an autonomous process until detection. The number of infected animals in a herd therefore depends only on the infection-age τ , and not on the absolute time t .

We describe the number of infective animals by an ordinary birth and death process, writing $p_i(\tau)$ for the probability of i infective animals in a herd at infection-age τ and δ_{ij} for the Kronecker Delta function. We have, (see e.g. [3]):

$$\begin{aligned}
 p_i(0) &= \delta_{i1}, \\
 \frac{dp_0(\tau)}{d\tau} &= \mu p_1(\tau), \\
 \frac{dp_1(\tau)}{d\tau} &= -(\lambda + \mu)p_1(\tau) + 2\mu p_2(\tau), \\
 \frac{dp_i(\tau)}{d\tau} &= -i(\lambda + \mu)p_i(\tau) + (i - 1)\lambda p_{i-1}(\tau) + (i + 1)\mu p_{i+1}(\tau) \quad \forall i \geq 2.
 \end{aligned}$$

Solving these differential equations leads to:

$$\begin{aligned}
 p_0(\tau) &= \frac{e^{r\tau} - 1}{Re^{r\tau} - 1}, \\
 p_i(\tau) &= (1 - p_0(\tau))(1 - Rp_0(\tau))(Rp_0(\tau))^{i-1} \quad \forall i \geq 1,
 \end{aligned}$$

where $r = \lambda - \mu$ and $R = \frac{\lambda}{\mu}$ is the reproduction ratio. We assume $\lambda > \mu$, otherwise the infection will typically only cause a minor outbreak within a herd, and will therefore typically not infect other herds.

Hence, conditioned on the event that the epidemic in a herd does not go extinct before age τ , i.e. $I(\tau) > 0$, $I(\tau)$ is geometrically distributed with parameter $(1 - Rp_0(\tau)) = \frac{R-1}{Re^{r\tau}-1}$, which is small for large τ .

A geometric random variable with small parameter can be approximated by an exponential random variable with the same parameter. Therefore, for large τ , $I(\tau)$ can be approximated by $\tilde{H}(\tau)$, where $\tilde{H}(\tau)$ is an exponential random variable with parameter $\frac{R-1}{Re^{r\tau}-1}$. Further, if X is exponentially distributed with parameter x , then cX is exponentially distributed with parameter x/c . Therefore, $\tilde{H}(\tau)$ is distributed as $H(e^{r\tau} - \frac{1}{R})$, where H is an exponential random variable with parameter $\frac{R-1}{R}$. For large τ , the term $\frac{1}{R}$ is negligible compared to $e^{r\tau}$. So we use the approximation

$$I(\tau) = He^{r\tau}, \quad (1)$$

where the equality is in the distributional sense.

We can interpret this approximation as follows. The random variable H represents the random character of the start of the outbreak in a herd, when only a few animals are infective. If the disease does not go extinct, then after the initial phase, there are many infected animals in the herd, each of which causes an independent number of new infections per time unit. According to the law of large numbers, this means that the growth rate is eventually almost deterministic. We also use this approximation for small τ .

We use the word *infective* for animals or herds that are able to spread the infection. We use the notation $I(\tau; h)$ for the number of infective animals in a particular herd, with $H = h$ given.

2.2. Discussion of the within-herd model

1. We assume that the infection and recovery rate of individual infected animals are independent of time and age.
2. It is possible to use alternatives for the birth and death process to describe the spread within a herd, for example the contact process. The contact process is appropriate when we have a situation where all animals are positioned in a row and do not change position. Each animal can only infect its two nearest neighbours. We assume that recovered animals with two infective neighbours are re-infected immediately. Therefore, we only consider the animals at the edge of a row of infective animals. We still assume that the recovery rate is μ . An infective animal infects each of its susceptible neighbours with rate λ . We have that

$$\begin{aligned} p_i(0) &= \delta_{i1}, \\ \frac{dp_0(\tau)}{d\tau} &= \mu p_1(\tau), \end{aligned}$$

$$\begin{aligned}\frac{dp_1(\tau)}{d\tau} &= -(2\lambda + \mu)p_1(\tau) + 2\mu p_2(\tau), \\ \frac{dp_i(\tau)}{d\tau} &= 2(-(\lambda + \mu)p_i(\tau) + \lambda p_{i-1}(\tau) + \mu p_{i+1}(\tau)) \quad \forall i \geq 2.\end{aligned}$$

Let the number of infectives be $I(\tau)$. Then $\mathbb{E}(I(\tau)|I(\tau) > 0) = 2r\tau + o(\tau)$. Furthermore, we can show that $\text{Var}(I(\tau)|I(\tau) > 0) = o(\tau^2)$. This implies that for $I(\tau) > 0$ we have $\frac{I(\tau)}{\mathbb{E}(I(\tau)|I(\tau) > 0)} \rightarrow 1$. So, in contrast to the birth and death model, we may approximate the random variable $I(\tau)$ for large τ by a deterministic variable $2r\tau$. This makes computations much easier in this case.

3. In the original model, the number of animals in one herd is assumed to be very large compared to the number of infected animals. Therefore, we assume that the contact rate between infected and susceptible animals is constant, and hence so is the birth rate. From data of outbreaks of CSF in the past, we can see that the number of infected animals until detection of the infection within the herd, is small compared to the total number of animals in the herd, so these assumptions seem justified in this case [11].
4. The within-herd infection and recovery parameters λ and μ can be measured experimentally or from data of past or on-going outbreaks. The parameter α depends on the development of symptoms of infected animals and how attentive the farmers are. Therefore, in reality this α will change at the first detection of the infection in the country (or in neighbouring countries), due to higher awareness of farmers and veterinarians. It is very difficult to estimate α for the period before the time of the first detection. For the time after the first detection, we can estimate α from data of past outbreaks in the same area or try to estimate this parameter during the ongoing epidemic. Using data from past outbreaks is dangerous, because the characteristics of the virus and of the farming practice may have changed. Estimation of the parameters, during an on-going outbreak is done by Meester et al. [10]. This method has some problems, e.g. the time necessary to get enough data for a reliable estimate. Another problem is that in [10] the infectivity of a herd does not depend on the ‘‘age’’ of the herd. This independence of age is essential for the estimations made. For our model it is not necessary that α is constant in time. It is possible to extend the model and use $\alpha(t)$ instead of α .
5. We assume that culling is the only measure which influences the within-herd spread of the infection. Vaccination is assumed to show no effect on the spread in the herd. For CSF this assumption is justified by the fact that vaccination will lead to immunity only after two weeks. Therefore, during this first two weeks the spread of the infection is not affected by this measure. For the time after these two weeks, we assume that we can use the same speed of propagation of the infection, for computational reasons. The approximation leads to results that are conservative, that is, too pessimistic.
6. We do not take characteristics of individual animals into account that might cause the individuals to differ in infectivity, susceptibility or contact pattern. Often age and type of species can have a substantial influence. For CSF this

implies we do not distinguish between the age of animals in one herd. The detection rate and infectivity of herds with many young animals does not significantly differ from the detection rate and infectivity of herds with especially older animals [11, 1].

7. We also use the approximation $I(\tau) = He^{r\tau}$ for small τ . This is not correct, but due to the small number of infective animals at small τ , the probability that the disease is detected at small τ is small too. We will see later that the number of infections in that period is as well, so the overall influence of the events while the herd is ‘young’ is probably not so large.

3. Spread of the disease between herds

3.1. The model for non-varying environments

In this section, we consider classical swine fever as a concrete example.

We distinguish between two types of farms: *multipliers* (m) are, roughly speaking, farms where young piglets are born, and *finishers* (f) are farms that buy piglets and fatten them. p_m denotes the fraction multipliers of the total number of herds and p_f is the fraction finishers. We assume that within either of these types, a birth and death model (with the same parameters for both types) describes the within-herd spread. The infectivity per non-transport contact of both types of herds develops in the same way too. However, transport contacts are only allowed from multipliers to finishers. Therefore, we take a larger infectivity rate for contacts from multipliers to finishers.

We define $A_\xi(t, \tau; h)$ as the infectivity of a herd at time t , while the herd was infected τ time units ago and with $H = h$, where ξ is a two-dimensional vector, denoting the two types of herds involved in the contact, so that ξ can be ff , fm , mm or mf . When no measures are implemented, $A_\xi(t, \tau; h)$, with h a given realisation of H , is proportional to the number of infected animals in the herd.

$$A_\xi(t, \tau; h) = \beta_\xi I(\tau; h)$$

Here β_ξ is a constant depending only on ξ . Because the non-transport contacts all happen at the same rate we can define $\beta_m := \beta_{fm} = \beta_{mm}$. Further we write $\beta_f := \beta_{ff}$ and β_{mf} is β_f plus some additional term for infections caused by transport of piglets from multipliers to finishers. Because all non-transport infections happen at the same rate the ratio β_m/β_f is exactly the ratio of the multipliers to the finishers. Note that the infectivity does not depend on the *absolute* time t .

We define β by $\beta := \beta_m + \beta_f$, hence $\beta_m = p_m\beta$ and $\beta_f = p_f\beta$. In order to consider transport contacts we also define $\beta_{mf} = p_f\beta + \beta_{tr}$, where β_{tr} is the proportionality factor of the part of the infection rate that is due to transport contacts. There is no term p_f in front of β_{tr} , because all transport contacts are from multipliers to finishers. We assume that the total number of herds is very large; in our computations, we assume it to be infinite.

We already know that the detection-rate is given by $\alpha I(\tau; h) = \alpha h e^{r\tau}$. From this we can deduce, for given h , the probability $p_{nd}(\tau; h)$ that an infected herd of age τ is not yet detected is

$$p_{nd}(\tau; h) = e^{-\int_0^\tau \alpha h e^{rs} ds} = e^{-\frac{\alpha}{r} h (e^{r\tau} - 1)} \tag{2}$$

In the case where no measures are implemented, the expected number of infected multipliers by one infective multiplier up to age τ , for given h , $\mu_{mm}(\tau; h)$, is given by:

$$\mu_{mm}(\tau; h) = \frac{\beta_m}{\alpha} (1 - e^{-\frac{\alpha h}{r} (e^{r\tau} - 1)}). \tag{3}$$

To see this, we first consider only one type of herd. We prove the following proposition:

Proposition 1. *Suppose that the infection and detection rate are proportional to the number of infective animals in a herd and that the environment is non-varying. Then the distribution of the size of the progeny of an infective herd does not depend on the number of infective animals in that herd at the start of the process.*

Proof. The infection and the detection rate are proportional to the number of infective animals in a herd. The ratio of infection rate and detection rate is given by β/α . We call the detection of the herd and the infections by this herd *events*. The probability that an event is an infection is $\frac{\beta}{\alpha + \beta}$ and a detection $\frac{\alpha}{\alpha + \beta}$. Therefore, the number of events, including detection, is described by a geometric random variable with parameter $\frac{\alpha}{\alpha + \beta}$. Hence the direct offspring distribution does not depend on the size at the start of the process. Moreover, the same holds for the offspring of this direct offspring. \square

In the same way we can prove that the size of the future offspring of an infective herd is independent of its age.

To prove (3), consider two types of herds. We note that 1 added to the number of multipliers infected by one infective multiplier is given by a geometric random variable with parameter $\frac{\alpha}{\alpha + \beta_m}$. (1 is added because the final event will be the only detection, and the number of events is described by a geometric random variable.) So the expected number of future infections of an infective multiplier is $\frac{\beta_m}{\alpha}$, at all times. The probability that an infective herd is not yet detected at age τ is given by $e^{-\frac{\alpha}{r} h (e^{r\tau} - 1)}$. From this and Proposition 1 we deduce that the expected number of infections after age τ is given by $\frac{\beta_m}{\alpha} e^{-\frac{\alpha}{r} h (e^{r\tau} - 1)}$. By subtracting the expected number of infections after age τ from the expected total number of infections by one herd we get the expected number of infections until age τ .

In the same way we can deduce $\mu_{mf} = \frac{\beta_{mf}}{\alpha} (1 - e^{-\frac{\alpha h}{r} (e^{r\tau} - 1)})$, $\mu_{fm} = \frac{\beta_{fm}}{\alpha} (1 - e^{-\frac{\alpha h}{r} (e^{r\tau} - 1)})$ and $\mu_{ff} = \frac{\beta_{ff}}{\alpha} (1 - e^{-\frac{\alpha h}{r} (e^{r\tau} - 1)})$.

Now consider the probability p_{kl}^f that a finisher infects k multipliers and l finishers. All events (infections of multipliers, infections of finishers and detection of an infected herd) happen at a rate proportional to the number of infective animals

in the infective herd. Therefore, the rates are all proportional to each other. First, we only consider infections and detections, and we do not yet consider the different types of herds infected. Detection occurs with rate $\alpha h e^{r\tau}$ and infection occurs with rate $\beta_m h e^{r\tau} + \beta_f h e^{r\tau} = \beta h e^{r\tau}$. As in the proof of Proposition 1 we can describe the total number of events, \bar{D} say, by an ordinary geometric random variable with parameter $\frac{\alpha}{\alpha + \beta}$. So the probability that $n + 1$ events occur, i.e. n herds are infected by one finisher, is

$$\mathbb{P}(\bar{D} = n + 1) = \left(\frac{\alpha}{\alpha + \beta}\right) \left(\frac{\beta}{\alpha + \beta}\right)^n$$

If in total n herds are infected by one finisher, we know by the lack of memory property that the number of infected multipliers, N_{fm} , is binomially distributed with parameters n and $\frac{\beta_m}{\beta}$. That is,

$$\mathbb{P}(N_{fm} = k | \bar{D} = n + 1) = \binom{n}{k} \left(\frac{\beta_m}{\beta}\right)^k \left(\frac{\beta_f}{\beta}\right)^{n-k}.$$

Note that $p_{kl}^f = \mathbb{P}(N_{fm} = k, N_{ff} = l) = \mathbb{P}(N_{fm} = k | \bar{D} = k + l + 1) \mathbb{P}(\bar{D} = k + l + 1) = \binom{k+l}{k} \left(\frac{\beta_m}{\beta}\right)^k \left(\frac{\beta_f}{\beta}\right)^l \left(\frac{\alpha}{\alpha + \beta}\right) \left(\frac{\beta}{\alpha + \beta}\right)^{k+l}$.

The generating function of p_{kl}^f is now given by

$$g^f(s_1, s_2) = \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} p_{kl}^f s_1^k s_2^l = \frac{\alpha}{\alpha + \beta - \beta_m s_1 - \beta_f s_2}.$$

We can deduce

$$g^m(s_1, s_2) = \frac{\alpha}{\alpha + \beta + \beta_{tr} - \beta_m s_1 - (\beta_f + \beta_{tr}) s_2}$$

in the same way. Note that we did not need the distribution of the random variable H , we only needed the proportionality factors of the parameters.

In the special case where infection and detection rates are proportional and these rates are known for the time before the first detection, we can also give the distribution function of the number of infective herds at the time of the first detection. This distribution function is the same as the distribution function of the number of direct infections by one herd, because it still holds that an event is a detection with probability $\frac{\alpha}{\alpha + \beta}$ and the probabilities of infections of finishers and multipliers are respectively $\frac{\beta_f}{\alpha + \beta}$ and $\frac{\beta_m}{\alpha + \beta}$.

We are also interested in the probability that a herd infects k multipliers and l finishers, before the herd reaches age τ . We consider a finisher. We write $\mathbb{P}(N_{fm}(\tau; h) = k, N_{ff}(\tau; h) = l)$ for this probability when the detection age τ_d and $H = h$ are given. The infections before age τ_d occur independently of each other and have the lack of memory property. Therefore, for $\tau \leq \tau_d$ the times of infections of the different types of herds are described by an inhomogeneous Poisson-process with rate $\beta_\xi h e^{r\tau}$. So the number of infections until age τ is Poisson distributed with parameter $h\beta_\xi \int_0^\tau e^{rs} ds = h\beta_\xi (e^{r\tau} - 1)$. Now we see:

$$\mathbb{P}(N_{fm}(\tau; h) = k, N_{ff}(\tau; h) = l) = \frac{(h \frac{\beta_m}{r} (e^{r\tau} - 1))^k}{k!} \frac{(h \frac{\beta_f}{r} (e^{r\tau} - 1))^l}{l!} e^{-h \frac{\beta}{r} (e^{r\tau} - 1)}.$$

Note that now we know the life length distribution, the expected number of infections by a herd up to age τ and the underlying Galton-Watson process of the branching process, we can use the general theory of branching processes to determine some other properties, like the expected duration of the outbreak. ([4], chapters 2 and 6).

3.2. Discussion

1. We use a branching idea to describe the spread of the infection among different herds. For this approach, we assume that one herd has contacts with many other herds. In our model we do not take spatial distribution of the herds into account. As long as there are many herds in an area, the local exhaustion of susceptible herds can be ignored. If the outbreak is in an advanced state, however, local depletion of susceptible herds (either by the epidemic progression or by so-called pre-emptive culling) will certainly play a role. As long as there are relatively few infected herds in a neighbourhood, (i.e. a group of herds that have contacts with each other) we assume that the infection rate does not directly depend on this number of infected herds. Measures like ring-culling (culling of all herds within a certain distance of an infected herd) and ring vaccination cannot be considered in our model.
2. We distinguish between two types of contacts, transport contacts and indirect contacts. We have assumed that the indirect contacts are the same for all herds. Transport contacts are only from multipliers to finishers. Indirect contacts include transport of the virus by wind, by visits from an infective herd to a susceptible herd etc. Transport contacts are transports of infected animals from a multiplier to a finisher.
3. We simplify the model by assuming that the measures are implemented at the time of the introduction of the virus in the population (or, in other words, that the infection is detected immediately). It would be desirable to implement the measures from the moment of first detection. It is difficult, however, to estimate parameters like α and β for the time before the first detection, when awareness is not yet heightened by announcement of the outbreak, and when increased hygienic measures on farms are not yet taken. If we know the value of the parameters of the model for the period between introduction and first detection, we are able to estimate the probability of extinction and the expected final size of such a situation. However, these calculations still require that we use the information from our simplified model, where measures are started directly upon introduction.
4. The function $\mu_{mm}(\tau; h)$ is not required for our calculations of the final size and the probability of extinction. The reason why we include this function in our paper is that this function is interesting in its own right. It gives the expected offspring τ time units after an infective herd is infected itself. In practical applications, it is possible that by contact tracing some herd is suspected of

being infected τ time-units ago. This $\mu_{mm}(\tau; h)$ is then the expected number of infected multipliers by the suspected herd, if that suspected herd is a multiplier and if it is really infected.

5. The proportionality of the infection rate to the detection rate is essential in this section. Due to this property, we can find an underlying ordinary Galton-Watson process. Without this proportionality, we may ask whether it is possible to give a nice expression for the generating function. We also lose the independence of h in the generating function, which gives some computational problems.
6. Knowing the generating function of the number of infected herds at the first detection is important because if in some way it is possible to estimate parameters for the time before the first detection, we can use the distribution function of the number of infective herds at the time measures are implemented. We do not have the ‘ages’ of the infective herds at the first detection, but we can find a worst-case-scenario, by finding what age of a herd at $t = 0$ leads to the biggest offspring. If the detection and infection rate does not depend in the same way on the number of infected animals, we cannot deduce the distribution function at the time of the first detection in this way.

4. The model for varying environments

4.1. The approach

In the previous section, we considered the model for the spread of an infectious disease in non-varying environments. We heavily used the proportionality of the infection and detection rate of herds. This proportionality does not hold in a varying environment. Therefore, we need a different approach.

We consider only one type of herd; the multi-type model is a straightforward generalisation of this. We assume that the spread within a herd is not influenced by the state of the environment. The detection rate only depends on the number of infected animals in a herd and is written as $\alpha I(\tau; h)$, where $I(\tau; h)$ is the number of infected animals as described in Section 2. In this section, we assume that the within-herd spread is described by a birth-and-death process. So $I(\tau; h) = he^{r\tau}$. We can easily deal with other descriptions of the within-herd spread.

The infection rate may change due to control measures. Some measures lead to an infection rate that only changes finitely many times, other measures cause a continuously varying infection rate. We assume that in either case the environment is constant after some given time, $t = T$. That is the moment that measures have no more added value and therefore do not lead to new changes in the values of the model parameters. The infection rate is given by

$$A(t, \tau; h) = \beta\phi(t)I(\tau; h),$$

where $\phi(t)$ describes the effects of the measures implemented. We assume that $\phi(t)$ is deterministic. Because the environment does not influence the within-herd spread, $\phi(t)$ does not depend on the age τ . For the rest of this section it is assumed that $\phi(t) = 1$ for $t \geq T$.

The assumption of constant environment after some given time is often realistic. With a vaccination for instance, we know the moment that all vaccinated animals no longer live. In other control measures, like a transport ban, we can vary the time T and compute the effects.

We are looking for the probability of extinction of the infection, the expected final size and a generating function for the number of infected herds at a certain moment. To do this we use a discrete approximation of $\phi(t)$, so we have only a finite number of changes. Because the environment is non-varying after a certain moment, we can use the ordinary theory of branching processes to get all interesting properties for herds infected *after* $t = T$. By using backward iteration we will find the relevant properties of the epidemic for a herd infected in another interval, because these properties only depend on what happens in the intervals after the interval of infection. We can compute these properties for the herd infected at $t = 0$.

4.2. The probability of extinction

Because an infected herd will almost surely be detected in finite time, the probability of extinction of the ‘progeny’ of an infected herd x , is equal to the probability of extinction of the progeny of all the herds infected by x . The probability of extinction of the ‘progeny’ of a herd infected at time t , is denoted by $q(t)$. So:

$$q(t) = \mathbb{E}\left(\prod_i q(t_i)\right) \quad (4)$$

Here, the times t_i are the times that herds are infected by the herd infected at time t ; the expectation is over these random times of infection. The empty product is defined as 1.

In order to make computations possible we use a discrete time approximation. We divide the positive real line into $N + 1$ intervals, labeled $1, 2, \dots, N + 1$. The interval $(0, T]$ is divided into N intervals of equal length, where t is in interval i , if $t \in ((i - 1)\frac{T}{N}, i\frac{T}{N}]$, $1 \leq i \leq N$. The final interval $N + 1$ is (T, ∞) ; in this final interval all the model parameters are constant. The time $t = 0$, the moment of the first infection, is not included in one of the intervals, we treat this point in our notation as interval 0.

If the function $\phi(t)$ is discontinuous we may choose another discretisation, so that the discontinuities are on the boundaries of intervals. It is not essential that all intervals have the same length.

For t in interval i , $1 \leq i \leq N$, $q(t)$ and $\phi(t)$ are approximately constant. In our ‘discrete time model’ we will write $q(i)$ and $\phi(i)$ for the value of q and ϕ in interval i . For instance, we can take $\phi(i)$ to be the value of $\phi(t)$ at the midpoint of interval i . Because we accept discontinuities in $\phi(t)$, this function may differ significantly for neighbouring intervals.

Now, with $n(i, j)$, the number of infections in interval j , due to one herd infected in interval i , we can write (4) as:

$$q(i) = \mathbb{E}\left(q(i)^{n(i,i)} q(i+1)^{n(i,i+1)} \dots q(N+1)^{n(i,N+1)}\right), \quad (5)$$

where the expectation is over the numbers of infections in each interval. $q(0)$ is the probability of extinction of the progeny of the herd infected at time $t = 0$. We assume that all infections and detections take place on the midpoint of the interval, except of course for the events in the final interval, because there we can compute everything explicitly.

Given that a herd is not yet detected and $H = h$, the probability of infection in a certain interval by that herd does not depend on the number of infections in previous intervals. Consider a herd infected in interval i , denote by $D(i; h) = k$ the event that this particular herd is detected in interval k , given $H = h$. After detection no further infections occur. Define $n(i, l, k; h)$ as the number of infections in interval l due to a particular herd infected in interval i , while $H = h$ and the interval of detection k , are given. Now, by using independence we have (writing P_H for the distribution function of H)

$$\begin{aligned}
 q(i) &= \int \sum_{k=i}^{N+1} \mathbb{E}(q(i)^{n(i,i,k;h)} q(i+1)^{n(i,i+1,k;h)} \dots \\
 &\quad \dots q(k)^{n(i,k,k;h)} | D(i; h) = k) \mathbb{P}(D(i; h) = k) dP_H \\
 &= \int \sum_{k=i}^{N+1} \mathbb{E} \left(\prod_{l=i}^k (q(l)^{n(i,l,k;h)} | D(i; h) = k) \right) \mathbb{P}(D(i; h) = k) dP_H \\
 &= \int \sum_{k=i}^{N+1} \prod_{l=i}^k \mathbb{E}(q(l)^{n(i,l,k;h)} | D(i; h) = k) \mathbb{P}(D(i; h) = k) dP_H.
 \end{aligned}$$

Here the expectation inside the integral depends on h , contrary to the non-varying environment case. Note that the number of infections in a certain interval is not independent of the interval of detection. It does not make any difference for the number of infections whether a herd is detected shortly after the considered interval or very long after that, but it does make a difference whether the detection is in the considered interval itself. So $p(n; i, l, k; h) := \mathbb{P}(n(i, l, k; h) = n)$ depends on the detection interval k . For computational reasons we suppose in this discrete model that for $i \leq N$, $p(0; i, i, k; h) = 1$. We write $p_{det}(i, k; h)$ for $\mathbb{P}(D(i; h) = k)$. We are interested in $q := q(0)$. We can easily compute all these probabilities in our model.

Because we assume a birth-and-death process for the within-herd spread, we have to take the random character of H into account. Doing this leads to the following formulae:

$$\begin{aligned}
 q(i) &= \sum_{k=i}^{N+1} \int_0^\infty \frac{R-1}{R} e^{-h \frac{R-1}{R}} (\mathbb{E}(q(i)^{n(i,i,k;h)}) \\
 &\quad \dots \mathbb{E}(q(k)^{n(i,k,k;h)}) p_{det}(i, k, h) dh.
 \end{aligned}$$

Here we have conditioned on the time of detection and on h . We can rewrite this formula as:

$$q(i) = \sum_{k=i}^{N+1} \int_0^\infty \frac{R-1}{R} e^{-h \frac{R-1}{R}} \left(\prod_{l=i+1}^k \left(\sum_{n=0}^\infty p(n; i, l, k; h) q(l)^n \right) \right) p_{det}(i, k; h) dh.$$

Here $q(i)$ depends on $q(l)$ for $l > i$. As mentioned before, we can compute $q(N + 1)$ by the ordinary theory of branching processes. We use backward iteration to compute $q(0)$.

Note that we also know the probability of extinction for herds infected after time $t = 0$. If we estimate (for example by contact tracing) the moment of infection of a certain herd, we can use this information to improve predictions.

4.3. The expected final size of the epidemic

In almost the same way we can calculate the expected final size of the epidemic i.e. the expected total number G of infected herds. This will also be done for only one type of herd.

If $q < 1$, there is a positive probability for the final size to be infinite and therefore the expected final size will always be infinite; so to have the possibility of a finite expected final size we assume $q = 1$. We denote by $G(t)$ the size of the progeny of a particular herd, infected at time t (including this ancestor) and write $G(0) = G$. The expected number of herds in the progeny of herd X , including X itself, is 1 plus the sum of the expected size of the progenies of all herds infected by this X , i.e.

$$\mathbb{E}(G(t)) = 1 + \sum_i \mathbb{E}(G(t_i))$$

Where the t_i 's are again the random times at which infections by the element, infected on time t , occur. The empty sum is defined to be equal to 0.

In the same way as we deduced the formulae for $q(i)$ we deduce the formulae for $\mathbb{E}(G(i))$ in the discrete approximation. We write $\bar{G}(i)$ for $\mathbb{E}(G(i))$:

$$\begin{aligned} \bar{G}(i) &= 1 + \sum_{k=i}^{N+1} \int_0^\infty \frac{R-1}{R} e^{-h \frac{R-1}{R}} \mathbb{E}(n(i, i, k; h)) \bar{G}(i) + \dots \\ &\quad \dots + \mathbb{E}(n(i, k, k; h)) \bar{G}(k) p_{det}(i, k; h) dh \\ &= 1 + \int_0^\infty \frac{R-1}{R} e^{-h \frac{R-1}{R}} \\ &\quad \sum_{k=i}^{N+1} \left(\sum_{l=i}^k \left(\sum_{n=0}^\infty n p(n; i, l, k; h) \bar{G}(l) \right) \right) p_{det}(i, k; h) dh \end{aligned}$$

For $\alpha > \beta$, $\bar{G}(N + 1)$ is known from the ordinary theory of branching processes (see e.g. [4]): for $\phi(N + 1) = 1$, $\bar{G}(N + 1) = \frac{\alpha}{\alpha - \beta}$.

4.4. A generating function for the number of infective herds

We will give the generating function for the number of infective and infected herds at any given moment. Here we will consider the generating function for the infective and infected herds at time $t = T$, the time after which the parameter values are considered to be constant.

Remember that the age of an infective herd only influences the number of infective animals in that herd. From Proposition 1 in Section 3.1, we know that the expected direct offspring of a herd, born after $t = T$, is independent of the age of that herd, given that the herd is not yet detected at that time. Furthermore we know that the size of the offspring (infected after time $t = T$) of a herd, infective at that time, does not depend on the offspring of other herds that are infective at time $t = T$. So for the distribution of the number of infections after time $t = T$, only the distribution of the number of infective herds at that time is important.

By using Proposition 1 and the theory for ordinary branching processes, we can compute everything we want to know. We will determine the distribution of X , the number of infective herds at time $t = T$. X_i is the number of infective herds at time T , from the progeny of one particular herd infected in interval i . (Here the herd itself is a part of its own progeny.) We denote X_0 by X . We also use Y_i for the number of infected herds in the progeny of a particular herd infected in interval i , that is detected before time $t = T$. The generating function of the distribution of X_i and Y_i is

$$\tilde{g}_i(s_1, s_2) = \mathbb{E}(s_1^{X_i} s_2^{Y_i}) = \int_0^\infty \frac{R-1}{R} e^{-h \frac{R-1}{R}} (\mathbb{E}(s_1^{X_i} s_2^{Y_i} | h)) dh.$$

We write $\tilde{g}_i(s_1, s_2; h) = \mathbb{E}(s_1^{X_i} s_2^{Y_i} | h)$.

We again assume that a herd does not infect other herds in the interval wherein it becomes infected itself. So:

$$\begin{aligned} \tilde{g}_i(s_1, s_2; h) &= \mathbb{E}(s_1^{X_i} s_2^{Y_i} | h) \\ &= \sum_{k=i}^{N+1} p_{det}(i, k; h) \mathbb{E}(s_1^{X_i} s_2^{Y_i} | D(i, k; h), h) \\ &= s_2 \sum_{k=i}^N p_{det}(i, k; h) \mathbb{E}\left(\prod_{l=i}^k \tilde{g}_l(s_1, s_2)^{n(i,l,k;h)}\right) \\ &\quad + s_1 p_{det}(i, N+1; h) \mathbb{E}\left(\prod_{l=i}^N \tilde{g}_l(s_1, s_2)^{n(i,l,N+1;h)}\right) \\ &= s_2 \sum_{k=i}^N p_{det}(i, k; h) \prod_{l=i}^k \mathbb{E}(\tilde{g}_l(s_1, s_2)^{n(i,l,k;h)}) \\ &\quad + s_1 p_{det}(i, N+1; h) \prod_{l=i}^N \mathbb{E}(\tilde{g}_l(s_1, s_2)^{n(i,l,N+1;h)}) \end{aligned}$$

$$\begin{aligned}
 &= s_2 \sum_{k=i}^N p_{det}(i, k; h) \prod_{l=i+1}^k \sum_{n=0}^{\infty} p(n; i, j, k; h) (\tilde{g}_l(s_1, s_2))^n \\
 &\quad + s_1 p_{det}(i, N + 1; h) \prod_{l=i+1}^N \left(\sum_{n=0}^{\infty} p(n; i, l, N + 1; h) (\tilde{g}_l(s_1, s_2))^n \right).
 \end{aligned}$$

For interval N we have:

$$\tilde{g}_N(s_1, s_2) = s_2 p_{det}(N, N) + s_1 p_{det}(N, N + 1)$$

Now we can determine $\tilde{g}_0(s_1, s_2)$ point wise. Note that we only need to compute $\tilde{g}_0(s_1, s_2)$ in $m + 1$ points to give a good approximation of the first m derivatives of this function for a certain point. With this derivatives we are able to compute the first m moments of the size at time $t = T$ or to approximate $\mathbb{P}(X = n)$ for all $n \leq m$.

With this generating function we can also compute the probability of extinction and the expected final size of the epidemic. We can use the same reasoning as before. The progenies of all herds, infective at time $t = T$ have to go extinct. This occurs with probability q^{X_0} . So the probability of extinction is $\sum_{k=0}^{\infty} \mathbb{P}(X_0 = k)(q(N + 1))^k = \tilde{g}_0(q(N + 1), 1)$. Note that with these s_1 and s_2 the model is exactly the same as the model in Section 4.2. For the expected final size we add the expected number of infected herds, already detected, to the expected size of the progeny of all the infective herds at time $t = T$ (again including the herds infective at that time). This is $\frac{d}{ds_1} g_0(s_1, 1)|_{s_1=1} \tilde{G}(N + 1) + \frac{d}{ds_2} g_0(1, s_2)|_{s_2=1}$. We can only use this property if after time $t = T$ the infection and detection rate are proportional to each other, because otherwise we need to know the ages of the infective herds at time $t = T$.

In a non-varying environment, we know for $q = 1$ the “speed” at which $\mathbb{P}(Z(t) > 0)$ decreases, for $t \rightarrow \infty$, where $Z(t)$ is the number of infective individuals at time t , descending from one individual infected at time $t = 0$ [4]. So we can give an upper bound for the probability the disease is already extinct at a certain time. We do this by assuming that all infective herds at time $t = T$ are infected at that time. So all herd at time T have age $\tau = 0$.

Now with the notation S_t for the probability that the progeny of a herd infected at time T , will not survive until time $T + t$ and with $Z(t)$ for the number of infected herds at time t , we have that:

$$\mathbb{P}(Z(T + t) = 0) = \sum_{k=0}^{\infty} \mathbb{P}(Z(T) = k)(S_t)^k = \tilde{g}(S_t, 1).$$

4.5. Discussion

1. By using this iterative method, we can compute the expected final size of an outbreak very fast, but for computing any higher moments of this final size, we need substantially more computational effort. By using simulations, it is possible to estimate these higher moments too. (see [1])

2. The lack of memory property is very important for our computations. We used it to write the formula of $q(i)$ in a ‘convenient’ form, with expectations in front of every $q(l)^n$. The speed of computations also heavily depends on the lack of memory property of the infection rate. In each interval the number of infections by one herd is Poisson distributed. We can easily integrate h out of the formula for $q(i)$, for Poisson distributed numbers of infections. (With the given distribution of H).

5. Results

In this section we use the Dutch classical swine fever epidemic of 1997 as example. We use the same data as Klinkenberg et al. in [1]. However, Klinkenberg et al. used simulations and for every simulation the parameters were (pseudo-)randomly chosen from the distributions of the parameters, while we used only estimations of parameters for our computations.

We consider the following set of control measures:

- A Total transport prohibition,
- B Killing of young piglets, in combination with a breeding ban,
- C Vaccination of all piglets (not sows) at multiplier herds, followed by recurrent vaccination of newborn piglets,
- D Single vaccination of all pigs at finishing herds,
- E Vaccination of piglets on arrival at finishing herds.

Choosing a (combination of) measure(s) is called a *control strategy* or *scenario*. The effects of different scenarios on the fraction of infective animals in a multiplier, in a finisher and the possibility of transport-infections ($\phi_m(t)$, $\phi_f(t)$ and $\phi_{tr}(t)$ respectively) are given in Table 1 (This table is adapted from [1]). For example, the 0 for ϕ_{tr} in strategy A means that infection is not possible by transport contacts. As another example, $\phi_f(t) = t/100$ for $t \leq 100$ in strategy D means that at time t a fraction $t/100$ of the animals in a finisher is infective. Note that the first 12 strategies are in a constant environment. Therefore, for those strategies we can compute properties of the spread of the disease directly. The probability of extinction is given in Table 2.

We also compute the expected final size of an epidemic; see Table 3. We only need to compute this for the strategies with almost sure extinction. The expected number of infected multipliers, while the initially infected herd was a finisher is denoted by G_{fm} . In a similar way we define G_{mm} , G_{mf} , G_{ff} . The last column of Table 3 is the expected number of infected herds, when initially five multipliers and five finishers were infected. Note that we cannot compare the expected final size with the results of Klinkenberg et al. (Table 6 in [1]), because they use the median of 1000 simulations, not the mean. The results of Klinkenberg et al. are given in Table 4. The 95% confidence intervals of the final size are also taken from [1] and were estimated by simulation.

We used the generating function for the number of infected and infective herds at time $t = T$, the point in time after which the parameter values are assumed to be constant, to estimate the size at time T (which varies for different strategies).

Table 1. The effect of different strategies on $\phi_f(t)$, $\phi_m(t)$ and $\phi_{rr}(t)$ ¹.

Strategy	Time interval	$\phi_f(t)$	$\phi_m(t)$	$\phi_{rr}(t)$
none	$t > 0$	1	1	1
ABCD	$t > 0$	0	365/1975	0
ABC	$t > 0$	1	365/1975	0
ABD	$t > 0$	0	1	0
AB	$t > 0$	1	1	0
ACD	$t > 0$	0	1009/1975	0
AC	$t > 0$	1	1009/1975	0
AD	$t > 0$	0	1	0
A	$t > 0$	1	1	0
BCD	$t > 0$	0	365/1975	0
CD	$t > 0$	0	1009/1975	0
DE	$t > 0$	0	1	1
BC	$0 < t \leq 100$	$1 - t/100$	365/1975	0
	$t > 100$	0	365/1975	0
BDE	$0 < t \leq 70$	0	$1 - 23t/1975$	1
	$t > 70$	0	365/1975	0
BD	$0 < t \leq 70$	$t/100$	$1 - 23t/1975$	1
	$70 < t \leq 100$	0.7	365/1975	0
	$100 < t \leq 170$	$1.7 - t/100$	365/1975	0
	$t > 170$	0	365/1975	0
BE	$0 < t \leq 70$	$1 - t/100$	$1 - 23t/1975$	1
	$70 < t \leq 100$	$1 - t/100$	365/1975	0
	$t > 100$	0	365/1975	0
B	$0 < t \leq 70$	1	$1 - 23t/1975$	1
	$70 < t \leq 170$	$1.7 - t/100$	365/1975	0
	$t > 170$	0	365/1975	0
C	$0 < t \leq 100$	$1 - t/100$	1009/1975	0
	$t > 100$	0	1009/1975	0
D	$0 < t \leq 100$	$t/100$	1	1
	$t > 100$	1	1	1
E	$0 < t \leq 100$	$1 - t/100$	1	1
	$t > 100$	0	1	1

Table 2. Probability of extinction for various strategies.

Strategy	$q_m(0)$	$q_f(0)$
none	0.32	0.51
ABCD, ABC, ABD	1	1
ACD, AD, BCD, CD		
AB	0.61	0.61
AC	0.80	0.80
A	0.61	0.61
DE	0.69	0.46
BC, BDE, BD, BE, B, C	1	1
D	0.38	0.61
E	0.39	0.59

Table 3. Expected final size for various strategies.

Strategy	G_{mm}	G_{mf}	G_{fm}	G_{ff}	$G_{5m+5f;f+m}$
ABCD	1.2	0	0.2	1	11.8
ABC	7.5	35.4	6.5	36.4	429.4
ABD	5.7	0	4.7	1	56.9
ACD	1.7	0	0.7	1	17.3
AD	5.7	0	4.7	1	56.9
BCD	1.2	0	0.2	1	11.8
CD	1.7	0	0.7	1	17.3
BC	1.1	0.8	0.3	1.8	21.6
BDE	2.7	3.6	0.9	2.0	45.4
BD	3.6	8.5	1.3	4.0	87.2
BE	3.2	5.8	1.2	3.2	66.6
B	4.5	13.2	1.9	6.8	132.3
C	2.4	1.0	1.4	2.0	33.8

Table 4. Median and 95% confidence interval of the final size¹.

Strategy	median of final size	95% CI
ABCD	11	10 – 16
ABD	42	13 – 452
ACD	16	10 – 35
AD	42	13 – 452
BCD	11	10 – 16
CD	16	10 – 35
BC	21	12 – 38
BDE	35	15 – 69
BD	67	20 – 157
BE	53	22 – 110
B	108.5	37 – 238
C	33	15 – 73

Table 5. Expected sizes at time T .

Strategy	initially infected type	undetected m	detected m	undetected f	detected f
B	multiplier	0.39	3.98	0.47	12.57
B	finisher	0.18	1.61	0.22	6.50
C	multiplier	0.35	1.72	0.14	0.82
C	finisher	0.35	0.72	0.14	1.82
D	multiplier	5.26	4.24	9.93	6.12
D	finisher	2.39	1.72	4.57	3.28
E	multiplier	5.34	4.55	6.88	6.27
E	finisher	2.64	2.00	3.45	3.65

We estimate the first two moments of the size at time T for the four pure strategies in a varying environment B, C, D, E. (Table 5) Especially the results of B and C are of interest because they will give an idea of the variance of the final size of a herd.

Next we give covariance matrices (for different strategies and different initially infected herds) of the number of infected multipliers not yet detected, the number of infected multipliers already detected, the number of infected finishers not yet detected and the number of finishers already detected, respectively. We denote by $Var_i(J)$ the covariance matrix for an initially infected herd of type i , while the strategy is J . By comparing the values in these matrices to the results in Table 5, we know how much trust we may put in the expected sizes.

$$Var_m(B) = \begin{pmatrix} 0.85 & 2.19 & 0.52 & 8.43 \\ 2.19 & 16.97 & 2.63 & 56.60 \\ 0.52 & 2.63 & 1.06 & 10.21 \\ 8.43 & 56.60 & 10.21 & 237.42 \end{pmatrix}$$

$$Var_f(B) = \begin{pmatrix} 0.38 & 0.98 & 0.23 & 3.43 \\ 0.98 & 8.02 & 1.17 & 23.84 \\ 0.23 & 1.17 & 0.48 & 4.13 \\ 3.43 & 23.84 & 4.13 & 90.10 \end{pmatrix}$$

$$Var_m(C) = \begin{pmatrix} 0.92 & 0.66 & 0.20 & 0.68 \\ 0.66 & 1.68 & 0.25 & 1.00 \\ 0.20 & 0.25 & 0.21 & 0.26 \\ 0.68 & 1.00 & 0.26 & 1.87 \end{pmatrix}$$

$$Var_f(C) = \begin{pmatrix} 0.92 & 0.67 & 0.20 & 0.67 \\ 0.67 & 1.67 & 0.25 & 1.00 \\ 0.20 & 0.25 & 0.21 & 0.26 \\ 0.67 & 1.00 & 0.26 & 1.87 \end{pmatrix}$$

$$Var_m(D) = \begin{pmatrix} 56.4 & 27.5 & 97.8 & 48.7 \\ 27.5 & 18.4 & 53.2 & 24.2 \\ 97.8 & 53.2 & 197.9 & 90.6 \\ 48.7 & 24.2 & 90.6 & 54.5 \end{pmatrix}$$

$$Var_f(D) = \begin{pmatrix} 23.0 & 11.4 & 39.5 & 17.6 \\ 11.4 & 8.1 & 22.2 & 10.5 \\ 39.5 & 22.2 & 80.9 & 32.5 \\ 17.6 & 10.5 & 32.5 & 18.0 \end{pmatrix}$$

$$Var_m(E) = \begin{pmatrix} 55.0 & 29.0 & 65.3 & 46.8 \\ 29.0 & 21.0 & 38.9 & 29.7 \\ 65.3 & 38.9 & 95.5 & 59.4 \\ 46.8 & 29.7 & 59.4 & 53.0 \end{pmatrix}$$

$$\text{Var}_f(E) = \begin{pmatrix} 24.3 & 12.9 & 28.3 & 18.3 \\ 12.9 & 9.9 & 17.2 & 12.0 \\ 28.3 & 17.2 & 41.9 & 22.7 \\ 18.3 & 12.0 & 22.7 & 19.0 \end{pmatrix}$$

Up to now we used exact parameters. In reality we cannot estimate the parameters exactly. In order to get some insight into how the computed properties depend on the values of the different parameters, we varied one parameter while keeping the other parameters constant. The results for the strategies B and D are given in Figures 1 and 2. Note that the scales on the axes differ for the different varying parameters. On the x -axes we put the ratio between the value of the parameter and the point estimator of that parameter. We see that for these strategies, the parameters r and R have relatively little influence on the computed quantities, while α and β do have significant influence.

In a constant environment the only thing that matters is the ratio $\frac{\beta}{\alpha}$. For varying environments we varied α , while keeping $\frac{\beta}{\alpha}$ and $\frac{\beta r}{\alpha}$ constant. The results are varying, but not very much.

5.1. Discussion

1. We cannot estimate the parameters exactly. Because the expected final size, the probability of extinction and the generating function for the number of infective herds at a certain moment heavily depend on some of the parameters, we cannot really use the computed values in a quantitative way. But we can use them to *compare* different scenarios, while using the same parameter values.
2. We computed the probability of extinction for the progeny of one infective herd, but in reality we may have more than one infected herd at the moment the first measures are implemented, so the probability of extinction may be much less than the computed q . We assumed that different infected herds infect other herds independently of each other. If we simplify the model by assuming that all herds at time $t = 0$ have age 0, we can estimate the expected final size by the number of initially infected herds times the estimated final size of one infected herd. In the same way we can estimate the probability of extinction by the computed probability for one herd, to the power of the number of initially infected herds. From previous outbreaks of FMD, CSF and AI in the Netherlands, one can see that as a rule several farms are already infected at the moment the first case is suspected or confirmed.
3. The results are very much the same as the results of Klinkenberg et al. but we do not need simulations to estimate the properties. Also, our method is much faster than using simulations. In Table 6 of [1] it is suggested that the expected final size, with initially five infected multipliers and five infected finishers, is estimated by simulations. In reality the median of 1000 simulations is used in that paper. This explains why our results differ on that point.
4. Note that strategy ABC gives an extremely high expected final size. Klinkenberg et al.[1] did not even give this final size. This is because for the given parameters the process is very near to a process with a positive probability of surviving i.e. a positive probability of an infinite final size.

Method B

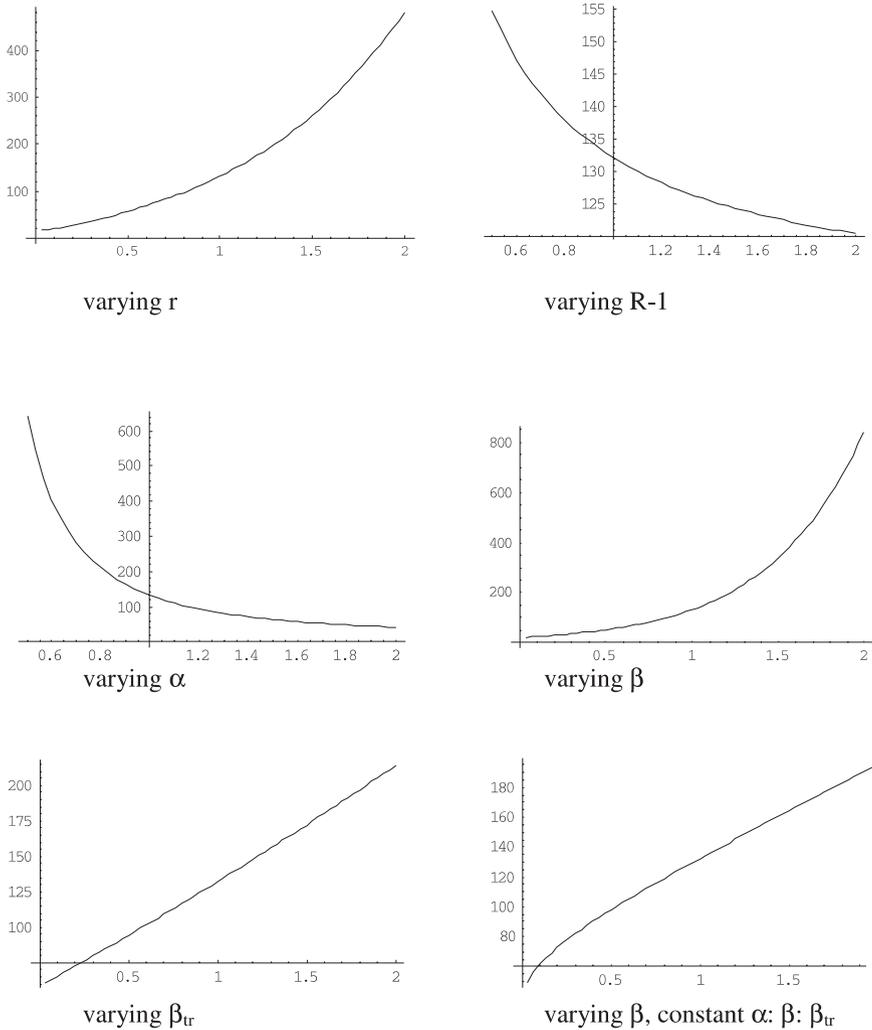


Fig. 1. The expected final size of an epidemic of classical swine fever initiated by 5 multipliers and 5 finishers, for varying parameters is indicated at the vertical axis. The ratio between the varying parameter and the point estimator of this parameter is given at the horizontal axis.

5. Using the expected size at $t = T$ and the variance of this size, we see that the expected final size is not very informative in some scenarios. Large variance may imply that a large set of “numbers of infected herds” have a not-ignorable probability, even if we know the exact parameters.
6. The computed covariance matrices do have values of different order in them. Consider for instance strategy B. The variance of the number of finishers and

Method D

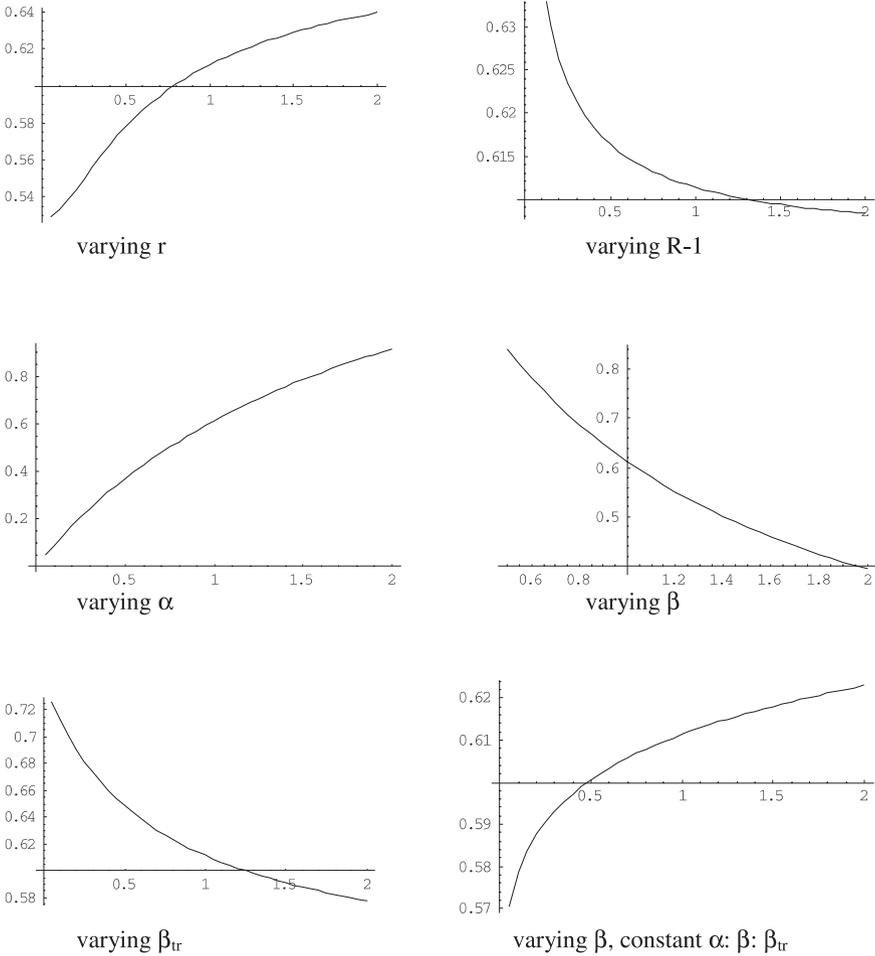


Fig. 2. Probability of extinction of an epidemic of classical swine fever initiated by a finisher, for varying parameters is indicated at the vertical axis. The ratio between the varying parameter and the point estimator of this parameter is given at the horizontal axis.

multipliers already detected is much higher than the variance of the number of finishers and multipliers that are still alive at time $t = T$. This is because the infection rate decreases for this strategy. At the start of the process one infective herd may infect several other herds with a relatively high probability, while after some time infections become rare. This is why the expected number of infective herds at $t = T$ is much less than the expected number of infected herds.

7. For constant environments we are interested in the variance of the final size. Note that in all of the strategies with almost sure extinction in a constant environment, only one type of herd can be infected, so we only have to deal with

one type of herd. We use the underlying Galton-Watson process to deduce the variance of the final size (see Section 2.11 of [4]). If we take m for the expected number of direct infections by one infective herd, the expected final size is $\frac{1}{1-m}$ and the variance is $\frac{m(1+m)}{(1-m)^2}$. Note that for a large expected final size, the variance will be of the square order of the final size. Due to this large variance, we cannot give exact quantitative predictions about the final size of an epidemic. This also indicates that there is a large intrinsic uncertainty in the problem.

8. If the infection rates are increasing in time, the expected final size will decrease for increasing r , while for decreasing infection rates the expected final size will increase for increasing r . This is because r can be seen as a parameter describing the speed of the process in a constant environment. Large values of r correspond to short generation lengths compared to small r . So, if the infection rate is increasing in time, the n -th generation for small values of r will have a larger infection rate than the n -th generation for large r .

The effect of decreasing R , the parameter describing the random effects at the start of the within-herd spread, is less clear from formulae and definitions. From the figures we can only see that R has relative small effect on the expected final size and the probability of extinction.

9. Some parameters heavily influence the expected final size and the probability of extinction. We have already seen that in a constant environment r and R do not influence these quantities. For predictions the ratios $\beta : \beta_{tr} : \alpha$ are most important. So the estimation effort is best devoted towards estimating these ratios.

6. Conclusion and final remarks

By using a stochastic model, we could estimate the probability of extinction and the expectation of the final size of an epidemic in a varying environment. We only need that the environment is not varying anymore after some given time. We use a branching process in varying environments, depending on the age of the particles. In the constant environment the probability of extinction and the final size are known. By using an iterative process, we computed this probability and size for the time the environment is still varying.

By using generating functions, we can compute many important properties, like the moments of the final size, and a lower bound for the probability that the process has gone extinct after a given time. These generating functions are very useful especially if in a constant environment the expected number of infections after a certain age does not depend on that age. This only holds if the infection and detection rate are in direct proportion, for the non-varying environment.

It is difficult to estimate the parameters for the computations. Some of the properties computed in this model, like the probability of extinction heavily depend on the parameters α and β , describing the infection and detection rate. The model presented in this paper may be useful to *compare* the effect of different measures, but it is very dangerous to use this model for absolute quantitative predictions.

We used the model to describe the spread of classical swine fever. It is worth investigating if the same model, with other parameters, may be used to describe

the spread of other animal diseases, with culling at detection, or even human diseases, which lead to strict quarantine of detected infected individuals (and suspected cases) and contact and movement restrictions. Emerging infections such as SARS are possible examples of this.

References

- [1] Klinkenberg, D., Everts-van der Wind, A., Graat, E.A.M., De Jong, M.C.M.: Quantification of the Effects of Control Strategies on Classical Swine Fever Epidemics. *Math. Biosci.* **186** (2), 145–173 (2003)
- [2] Diekmann, O., Heesterbeek, J.A.P.: *Mathematical Epidemiology of Infectious Diseases*. Chichester: John Wiley & Son, 2000
- [3] Grimmett, G.R., Stirzaker, D.R.: *Probability and Random Processes*. New York: Oxford University Press Inc., 1992
- [4] Jagers, P.: *Branching Processes with Biological Applications*. London: John Wiley & Son, 1975
- [5] Ferguson, N.M., Donnelly, C.A., Anderson, R.M.: The Foot-and-Mouth Epidemic in Great Britain: Pattern of Spread and Impact of Interventions. *Science* **292**, 1155–1160 (2001)
- [6] Ferguson, N.M., Donnelly, C.A., Anderson, R.M.: Transmission Intensity and Impact of Control Policies on the Foot and Mouth Epidemic in Great-Britain. *Nature* **413**, 542–547 (2001)
- [7] Jagers, P., Lu, Z.W.: Branching Processes with Deteriorating Random Environments. *J. Appl. Probab.* **39** (2), 395–401 (2002)
- [8] Kao, R.R.: The Role of Mathematical Modelling in the Control of the 2001 FMD Epidemic in the UK. *Trends in Microbiology* **10** (6), 279–286 (2002)
- [9] Keeling, M.J., Woolhouse, M.E.J., Shaw, D.J., Matthews, L., Chase-Topping, M., Haydon, D.T., Cornell, S.J., Kappey, J., Wilesmith, J., Grenfell, B.T.: Dynamics of the 2001 UK Foot and Mouth Epidemic - Dispersal in a Heterogeneous Landscape. *Science* **294**, 813–817 (2001)
- [10] Meester, R., de Koning, J., de Jong, M.C.M., Diekmann, O.: Modelling and Prediction of Classical Swine Fever Epidemics. *Biometrics* **41**, 178–184 (2002)
- [11] Stegeman, A., Elbers, A.R.W., Smak, J., de Jong, M.C.M.: Quantification of the Transmission of Classical Swine Fever Virus between Herds during the 1997-1998 Epidemic in The Netherlands. *Preventive Veterinary Medicine* **42**, 219–234 (1999)

¹ Reprinted from *Mathematical Biosciences*, 186 (2), KLINKENBERG, D.; EVERTS-VAN DER WIND, A.; GRAAT, E.A.M. AND DE JONG, M.C.M., Quantification of the Effects of Control Strategies on Classical Swine Fever Epidemics, 145-173 , Copyright (2003), with permission from Elsevier.