An epidemiological-based model for disease spread within the real options framework: The impact on the optimal timing of treatment

C.E. Dangerfield[a], A.E. Whalley[b], N. Hanley[c], M.F. Macpherson[d], A. Kleczkowski[d], J. Healey[e] and C.A. Gilligan[a]

[a] Department of Plant Sciences, University of Cambridge, Downing Street, Cambridge, CB2 3EA, UK
[b] Warwick Business School, The University of Warwick, Coventry, CV4 7AL, UK
[c] School of Geography & Geosciences, Irvine Building, University of St Andrews, North Street, St Andrews, Fife, KY16 9AL, UK
[d] Computing Science and Mathematics, School of Natural Sciences, Cottrell Building, University of Stirling, Stirling, FK9 4LA, UK
[e] School of Environment, Natural Resources and Geography, Thoday Building, Deiniol Road, Bangor, Gwynedd, LL57 2UW, UK

1. Introduction

Invasive pests and pathogens are increasing worldwide posing a threat to agriculture and forestry production (Gilligan 2008; Strange and Scott 2005). Deciding whether or not to implement expensive control strategies is complicated by uncertainties about the future spread of a pest or pathogen combined with the irreversible nature of many control strategies. Therefore it may be best not to apply treatment immediately, but to wait and see how the disease progresses (Sims and Finnoff 2013).

The real options approach has previously been used to investigate the impact of uncertainty on the timing of control actions for disease outbreaks, as it provides a convenient way to couple uncertainty with economic analysis (Sims and Finnoff 2013). Viewing disease control as an option which can be exercised to reduce damage to the host species of a pest or pathogen, the real options framework can be used to determine the optimal timing of control (Ndeffo Mbah et al. 2010; Saphores 2000). That is, when is it optimal to exercise the option to control the disease? How long should we delay actions in order to learn more?

To incorporate uncertainty into the decision making process, the progress in the level of infection is described by a stochastic process. Epidemiological modelling is a convenient approach for determining the form of such a stochastic process since it describes the evolution in the level of infection within a host species based on the characteristics of disease spread, (Keeling and Pejman 2008). Furthermore, epidemiological modelling frameworks are well established and have become an important tool in understanding the invasion and persistence of pathogens (Gilligan and van den Bosch 2008). Such an approach compartmentalises the population based on infection status (Keeling and Pejman 2008). For example, the simplest model assumes there are only two health states, either

susceptible (S) or infected (I) and is termed the SI-model. Individuals move from the susceptible to infected compartment at a rate that is proportional to the current level of infection. Since susceptibles are explicitly included within the modelling framework, the change in infection level will also depend on the current level of susceptible individuals. This leads to a logistic-type term in the deterministic equation describing the evolution in the level of infection over time, which agrees with results from experimental studies (Large, Beer, and Patterson 1946). The deterministic equation can then be extended to include future uncertainty by incorporating a noise term that represents the random fluctuations caused by variability in environmental conditions, (Keeling and Pejman 2008).

The real options approach, however, traditionally neglects the characteristics of disease spread in the formulation of the stochastic process describing uncertainty in infection levels. Instead the progress of disease spread is assumed to follow a geometric Brownian motion (GBM), (Saphores 2000; Sims and Finnoff 2012). Such a model ignores the impact of the size of the susceptible population on disease spread and essentially assumes that the mean level of infection grows exponentially. While this holds for the early part of the epidemic, it does not capture the longer term limiting behaviour of disease spread due to the finite number of susceptible hosts. The advantage of GBM is that it is simple and allows for closed-from solutions to the real options problem. However, in this paper we show that sacrificing epidemiological accuracy for computational ease can significantly affect the conclusions from the real options approach.

The principle objective of this paper is to show how epidemiological based modelling approaches can be used within the real options framework. By, contrasting the standard approach (GBM) with two alternative, more realistic, descriptions of the disease process, we show that using an inappropriate stochastic process to describe future uncertainty in the level of infection can lead to sub-optimal timing of control measures. The first alternative description incorporates logistic-type behaviour into the drift term, but has the same diffusion term as GBM. The second description of the disease process incorporates a logistic-type noise term into the diffusion term as well as the drift term. Both alternate formulations to GBM arise directly from the *SI*-type epidemiological model, according to different assumptions surrounding the incorporation of the noise term into the deterministic equation.

Our results suggest that not using an epidemiological-basis to describe future uncertainty in infection can lead to over-valuation of the option to treat and ultimately postponed deployment of control. Furthermore, ignoring the natural upper boundary of the system when there is a fixed size host population can result in control never being deployed, leading to losses due to disease damage. In particular, if treatment is deployed at the 'wrong' time then a portion of the option value will

never be realised. The structure of this paper is as follows. In Section 2 we review the existing literature on the control of invasive pests and pathogens. Section 3 motivates the new SDEs proposed to model the spread of infection and derives the associated real options model. The results are presented in Section 4 and finally Section 5 discusses the implications of the results and concludes.

2. Literature Review

Economists have investigated the effects of "natural risks" on the management of renewable resources such as forests as well as agricultural systems. Here, we focus on the control of invasive pests and pathogens since the main application of the model presented in this paper is to the deployment of treatment to minimise environmental damage due to an epidemic outbreak. Given the ecological uncertainty which characterises future disease spread, and the irreversibility of certain actions which can be taken in response to such a risk, then a real options approach is attractive. The first application of real options to pests and diseases is (Saphores 2000) who considers the optimal timing of pesticide application under future uncertainty in a pest population. Varying the level of uncertainty in the pest population dynamics, Saphores (2000) showed that greater uncertainty in future pest densities increases the threshold pest density at which it is optimal to spray, since the probability the pest population will become small is larger. This gives rise to the "wait and see" approach with regards to dealing with invasive species, the idea being that when there is great uncertainty in the future dynamics of the invasive species there is value in waiting to learn more before investing in controls.

Sims and Finnoff (2012, 2013) have since extended this initial work in a number of ways. (Sims and Finnoff 2013) consider the implications of the reversibility of the control strategy on (i) how long a regulator should wait to take action and (ii) the severity of the action taken (extent of control measure). If control measures are partly reversible, then it is optimal always to act as soon as possible, i.e. never to adopt a wait-and-see policy. This emphasises the fact that reversibility of actions is key, and that there is a trade-off between the speed and the severity (i.e. reversibility) of actions taken. (Sims and Finnoff 2012) consider the impact of a spatial boundary on timing of control by treating the maximum area that can be infected as an upper barrier within the real options framework. The spatial scale can impact the timing and stringency of control strategies, and so incorporating an upper bound is important in planning measures to minimise environmental damage (Sims and Finnoff 2012).

The real options approach has also been used to investigate the timing of specific control measures to minimise damage from disease or invasive species. (Sims 2011) considers the optimal timing of salvage harvest to recoup timber values following a disease outbreak in a forest crop. Sims (2011) finds that slower rates of forest growth delay the optimal timing of salvage harvest, while large timber and non-timber values suggest more immediate action is optimal. Multiple interacting control options, namely chemical and biological control, are considered in (Marten and Moore 2011). They find that biological control is sufficient to manage the pest, so long as infestation can be detected and controlled without substantial delay. However, if the pest reaches high levels before controls can be employed a more costly combined strategy is optimal for pest management.

In most studies, the value of the damage caused by the disease or pest is considered, and aim of the control is to minimise the damage cost within the real options framework. In (Ndeffo Mbah et al. 2010), they take a different approach by directly considering the value added by applying control (namely treatment for a disease or pest) in terms of the monetary gain per unit of infection prevented. Furthermore they incorporate a logistic term into the drift coefficient of the stochastic process describing disease spread, and so the mean growth of the process is limited by a parameter that represents the carrying capacity. In particular they find the new SDE leads to a difference in the optimal time to treat when compared with the standard GBM assumption, highlighting the dependence of the real options approach on the formulation of the underlying uncertainty in the disease dynamics. However, the disadvantage of the approach taken in (Ndeffo Mbah et al. 2010) is that the stochastic process can increase above the carrying capacity (Sarkar 2009). While in certain applications (e.g. harvesting fish) it is reasonable for the population to increase above the carrying capacity, in the study of disease spread this parameter represents a physical boundary such as the fixed number of trees or plants within a fixed area that can be infected. Therefore it is impossible for the area infected to reach a value greater than this bound and so trajectories of the stochastic process that go beyond this point do not have any applicable meaning in this context.

Although the impact of uncertainty in disease spread on the optimal timing of control measures has been extensively studied within the economics literature, there seems to have been a separation between traditional epidemiological models and those used within the real options framework. Therefore a key contribution of this paper is to show how the uncertainty in disease spread can be formulated directly from basic epidemiological principles. This leads to two different SDEs: the first of which has previously been studied in (Ndeffo Mbah et al. 2010) and the second of which incorporates a logistic type term into the diffusion as well as the drift coefficient. To the best of our knowledge this is the first time that such a process has been used within the real options framework to determine the optimal timing of control. In particular this new formulation provides a natural way

to incorporate the upper bound in the level of infection directly into the model. The complexity of this new stochastic process means that the real options model no longer permits closed form solutions, therefore we frame the problem in a similar manner to (Ndeffo Mbah et al. 2010) by assuming that treatment eradicates infection and so damage due to disease is not permanent.

3. Real Options Model

3.1 Setup

Consider an agricultural or forest disease outbreak in a particular crop or tree species at the landscape scale for which there is treatment available that would eradicate current levels of infection. In particular we consider that the number of trees or plants remains fixed. This is typically the case in agriculture or even-aged forest stands where typically crops or trees are initially planted and then harvested at some fixed time. Treatment can be applied at any time for a one-off fixed cost C, and this treatment (e.g. removal of infected trees or plants) is completely irreversible. Furthermore there is uncertainty in the future levels of infection due to environmental and demographic noise associated with the disease transmission process. The decision-making authority is faced with the following choice: should treatment be administered immediately or should they wait to learn more about the progression of the disease? In particular, waiting allows the decision maker to determine whether the level of infection gets worse or better over time.

Traditional net-present-value (NPV) analysis would advocate undertaking treatment providing the value of the investment (i.e. the application of treatment and the associated savings in economic losses), is greater than the cost C. However, due to uncertainty in disease dynamics combined with the irreversibility of the decision to treat, there is value in delaying treatment so as to learn more about the progress of the disease (Dixit and Pindyck 1994). That is, there is a value associated with the option to treat.

To include uncertainty into the decision making approach, we assume that the level of infection, I, can be described by a stochastic process. Traditionally the increase in the level of infection is assumed to follow geometric Brownian motion (GBM) and so the dynamics of the level of infection is given by the following SDE (Saphores 2000; Sims and Finnoff 2012; Sims and Finnoff 2013)

$$dI = \beta I dt + \sigma I dW, \qquad (1)$$

where β is the rate of transmission, σ, which we term the volatility, is a parameter that scales the amount of the uncertainty and I is the current level of infection. In the limit as $\sigma \to 0$ equation (1) is

equivalent to assuming deterministic exponential growth in infected area. The advantage of using the GBM is that the logarithm of the infected area follows a Brownian motion, and so analytic solutions to the real options model can be obtained. However, it does not have any epidemiological-basis, other than arguably during the very early stages of an epidemic, questioning it's applicability within this context.

3.2 Epidemiologically-Based Model of Uncertainty in Disease Spread

Traditional epidemiological models of disease spread compartmentalise the population based on infection status. The simplest model assumes there are just two infection states, susceptible (S) or infectious (I). Susceptibles move to the infected compartment upon infection. Therefore, the increase in the number of infected individuals is the per capita rate at which a susceptible contracts infection multiplied by the number of susceptible individuals. The rate at which a susceptible contracts infection is the rate of transmission per infected contact, β, multiplied by the probability of contact with an infectious individual, I/I_{max} where I_{max} is the maximum number of potential infected individuals that there can be. Assuming the total population remains constant, I_{max} is equivalent to the total population size and so $S = I_{max} - I$. Initially ignoring uncertainty in disease spread, the evolution in the level of infection is therefore given by the following ordinary differential equation (ODE)

$$\frac{dI}{dt} = \beta I\left(1 - \frac{I}{I_{max}}\right)$$

This model is referred to as the SI model within the epidemiological literature (Keeling and Pejman 2008). Uncertainty in future disease spread is incorporated by assuming there is variability in the transmission parameter, β, due to external forces. For example, fluctuations in temperature and climate have been shown to modify the infection rate (Sturrock et al. 2011). There are two different approaches for doing this.

Firstly the 'corrected apparent infection rate' is perturbed, leading to $\beta(1 - I/I_{max}) \rightarrow \beta(1 - I/I_{max}) + \sigma\xi$ (Marcus 1991), where ξ is white noise and σ is a constant that controls the magnitude of the perturbation. The uncertain evolution of future disease spread is described by the following SDE

$$dI = \beta I\left(1 - \frac{I}{I_{max}}\right)dt + \sigma I dW \qquad (2)$$

This SDE, which we refer to as the mean-reverting SDE, has been used within the real options framework in previous studies to describe the increase in infected area (Ndeffo Mbah et al. 2010) as

well as the growth in pest populations (Marten and Moore 2011). When the level of infection reaches I_{max} the magnitude of the diffusion term is non-zero and so there is a positive probability that the level of infection will exceed I_{max}. Since I_{max} represents the total number of trees or crops, which we assume remains fixed, such a boundary cannot be exceeded in reality, questioning the applicability of this stochastic process to the problem at hand.

Alternatively the transmission rate itself is perturbed, leading to $\beta \to \beta + \sigma\xi$ and so the evolution in the level of infection is given by the following SDE

$$dI = \beta I\left(1 - \frac{I}{I_{max}}\right)dt + \sigma I\left(1 - \frac{I}{I_{max}}\right)dW \qquad (3)$$

We refer to the above equation as the logistic SDE. As the level of infection, I, reaches I_{max}, both the drift and diffusion term approach 0 and so the trajectories of the SDE remain within the interval $[0, I_{max}]$, unlike for the previous approach. Therefore the physical upper boundary of the total population size is preserved directly within the dynamics of the SDE.

This approach provides a way of relating the uncertainty in future levels of infection to the randomness of the transmission process due to environmental factors. Therefore it provides an epidemiological-based approach to incorporating uncertainty into the decision problem. Furthermore this approach can be extended to more complex epidemiological models, for example in the case of diseases where there is an additional recovery state (termed the SIR model).

3.3 The Decision Problem

We assume that the only effect of treatment is to eradicate infection and for simplicity we assume treatment is applied instantaneously. Treatment could, for example, involve removal of infected material or application of a pesticide or biocontrol agent. We only consider the gain in economic value from the timber or crop saved and not from the wider environmental damage that may be reduced. Therefore the value of applying treatment at time t is simply:

$$V(t) = pI(t), \qquad (4)$$

where p is the gain in yield per unit of infected area treated, which is assumed to be constant over time. Viewing the application of treatment as an investment with value $V(t)$, the decision problem can be viewed as a real option (Dixit and Pindyck 1994), which analogously to a financial option (Black and Scholes 1973) is the right but not the obligation to make an investment for a fixed price in the future. The payoff for applying treatment at time t is $V(t) - C$ and so we want to maximise the expected present value,

$$F(V,0) = \max \mathbb{E}\left[(V(t_*) - C)e^{-r(t_*)} \right]. \quad (5)$$

$F(V,0)$ is the initial value of the option to apply treatment at some point in the future, t_* is the time in the future at which the decision is made, r is the discount rate and $\mathbb{E}$ denotes the expectation. The expectation must be taken since $I(t)$ (and therefore also $V(t)$) is a stochastic process.

Using standard methods from dynamic programming, the value of the option, $F(V,t)$, must satisfy the following Bellman equation, (Dixit and Pindyck 1994) (see appendix for details)

$$\frac{\partial F}{\partial t} + \frac{1}{2} b(V,t)^2 \frac{\partial^2 F}{\partial V^2} + a(V,t) \frac{\partial F}{\partial V} - rF = 0. \quad (6)$$

The functions $a(V,t)$ and $b(V,t)$ for each stochastic process used to describe the level of infection are given in Table 1.

Stochastic Process	$\boldsymbol{a(V,t)}$	$\boldsymbol{b(V,t)}$
Geometric Brownian Motion	$\beta \mathrm{V}$	$\sigma \mathrm{V}$
Mean-Reverting SDE	$\beta \mathrm{V}\left(1 - \frac{\mathrm{V}}{\mathrm{pI_{max}}}\right)$	$\sigma \mathrm{V}$
Logistic SDE	$\beta \mathrm{V}\left(1 - \frac{\mathrm{V}}{\mathrm{pI_{max}}}\right)$	$\sigma \mathrm{V}\left(1 - \frac{\mathrm{V}}{\mathrm{pI_{max}}}\right)$

Table 1. Form of the functions in the Bellman equation for each infection process

$F(V,t)$ must also satisfy the following boundary conditions

$$F(0,t) = 0$$

$$F(V_*,t) = V_* - C$$

$$\frac{\partial F}{\partial V}(V_*,t) = 1.$$

V_* is the value at which treatment should be applied immediately. That is it represents the boundary between the *continuation region* and the *exercise region* (region in which treatment applied). The first condition follows from the fact that if the value of applying treatment goes to 0 it remains at zero, i.e. infection cannot be re-introduced from an outside source. The second condition is called the *value matching condition,* (Dixit and Pindyck 1994), which states that when treatment is undertaken immediately the net gain is $V_* - C$. Finally the last condition is the *smooth pasting condition* which ensures optimality of the choice of V_*, since if F were not continuous at V_* then one could do better by investing at a different point (see (Dixit and Pindyck 1994) for further discussion). We note that this is a *free-boundary problem* since the location of the boundary is unknown and

must be determined as part of the solution. This threshold in the value of treatment can easily be converted to a threshold in the level of infected area (I_*) by dividing through by p.

Due to the complexity of the logistic SDE, the problem no longer permits a closed form solution. Therefore PDE given by (6) with corresponding boundary conditions is solved numerically in MATLAB using a finite difference method; see appendix for details. Baseline parameter values, and ranges for those parameters that are varied are given in Table 2.

Model Parameter	***Description***	***Base case (Range)***
β	Disease transmission rate	0.05 ([0.05, 0.8])
σ	Volatility	0.5 ([0.1, 0.9])
$\mathrm{I}_{\max}$	Carrying Capacity	100
C	Cost of treatment	20
r	Risk-free discount rate	0.1
p	Gain in yield per unit of infected area treated	1

Table 2. Parameter values used in numerical simulations.

4. Results

The solution to the free boundary problem provides the value of the option to treat as a function of the treatment value (V). Providing the value of the option to treat, $F(V)$, is greater than the NPV of immediate treatment, $F(V) > V - C$, there is value in retaining the option to treat, and so it is beneficial to wait. When $F(V) = V - C$, there is no additional gain in waiting and so treatment should be applied immediately. The value of treatment at which $F(V)$ first equals the NPV is termed the threshold value of treatment, V^*. This quantity is of interest since it determines the boundary between the waiting region and the region in which treatment should be applied immediately. We investigate how V^* varies according to different characterisations of the future uncertainty. We also explore the sensitivity of V^* to the epidemiological parameters σ and β (note that, as in (Sims and Finnoff 2013), we do not consider uncertainty over the economic parameters such as the price of timber). In Figure 1 we present the results directly in terms of the value of treatment V. In subsequent figures, the threshold value V^* is converted to the *proportion* of infected area (by rearranging equation (4) and rescaling by I_{max}). This is to emphasise the proportion of infection at which to apply treatment for the different processes.

4.1 Impact of epidemiologically-based models on the value of treatment threshold

Using an epidemiological-based SDE to describe uncertainty in disease spread (mean-reverting or logistic SDE) decreases the threshold value and so treatment is deployed at a lower proportion of infected area, compared with the standard GBM assumption. This can be seen in Figure 1 which shows the value of the option as a function of the treatment value for the three different SDEs. The effects hold for a wide range of parameter values. In particular, the threshold value is closer to the zero NPV (purple dotted line, Figure 2) under the mean-reverting and logistic models. As the level of infection becomes large both the drift and diffusion terms for GBM are greater than for the other two SDEs. There is thus a greater probability that the value of treatment, V, becomes large in the future which means the expected return from waiting is greater under the assumption of GBM. On the other hand, in the case of the logistic SDE as the level of infection approaches I_{max} the magnitude of both the drift and diffusion terms tend to zero and so the value of treatment cannot go above some maximal threshold $V_{max} = pI_{max}$. Hence there is a lower value to be obtained from waiting, and so it is optimal to apply treatment at a lower threshold level of infection (i.e. at a lower value of treatment). The threshold at which to act for the mean-reverting SDE lies between GBM and the logistic SDE. This reflects the trade-off between the growth in the drift term, which becomes dampened as the level of infection approaches I_{max}, as for the logistic SDE, and the diffusion term which becomes large with increasing levels of infection, as for GBM. Therefore the expected return from waiting is greater than for the logistic SDE but less than for GBM, and so the threshold value lies between the two.

The relative magnitudes of the infection rate (β) and the discount rate r are important. If $\beta = r$, the option value under GBM never intersects the standard NPV (red dashed line) leading to the conclusion that treatment should never be applied. In fact, such a situation is equivalent to an American call in financial options and it has been shown that it is never optimal to exercise the option early, i.e. treatment should not be applied (Wilmott, Howison, and Dewynne 1995). This is typically why in previous studies the assumption is made that $\beta < r$, (Sims and Finnoff 2012) and indeed this is why we investigate the behaviour of GBM over a subset of the parameter values used for the other two processes. For the mean-reverting and logistic SDEs a finite threshold value still exists if $\beta \geq r$ (Figure 2). Therefore, using an epidemiological-based approach to modelling disease uncertainty ensures there are no restriction on the range of transmission rate (β) and discount rate (r) parameters investigated, as is the case with GBM.

4.2 Impact of uncertainty on the timing of treatment and value of waiting

For all processes the threshold value is a monotonic increasing function of volatility, which is consistent with options theory (Dixit and Pindyck 1994). Increasing the volatility increases the option

value of waiting by increasing the probability of more extreme outcomes including those where the net benefit of treatment is high, whilst leaving the value of immediate treatment unchanged. The threshold, V^*, represents the treatment level at which the benefit of immediate treatment, $V^* - C$, exactly equals the cost of immediate treatment in the form of the value of retaining the option to treat at a later date, which is forgone once treatment is applied.

The rate of increase in the threshold value (V^*) with increasing volatility (σ) is greater for GBM and the mean-reverting SDE than for the logistic equation (Figure 2). Therefore, the difference between the optimal time to treat between the three processes is greatest when future uncertainty (σ) is large. Furthermore, the threshold at which to treat increases above the maximum area that can be infected (I_{max}) when $\sigma > 0.4$ in the case of GBM and when $\sigma > 0.6$ in the case of the mean-reverting SDE (Figure 2). Therefore, for large volatility, the threshold at which to apply treatment obtained from GBM or the mean-reverting equation is unattainable and so it is never optimal to apply treatment. This is not the case for the logistic equation, where the threshold always remains below I_{max}, and so is attainable (Figure 2). The reason for this fundamental difference in the conclusions from the GBM and mean-reverting equation compared with the logistic equation is as follows. As volatility increases, the probability that the treatment value is large increases more for the mean-reverting or GBM-type SDE, whereas for the logistic SDE there is an upper bound to the value of treatment at $V_{max} = pI_{max}$. Therefore, as volatility increases, the threshold at which to act remains bounded since the logistic SDE cannot reach large values above $V_{max} = pI_{max}$ in the future, no matter how large the volatility becomes.

Similarly, as the transmission rate increases, the threshold at which to act increases for both the mean-reverting and logistic SDEs. The threshold also increases for GBM since when $\beta \geq r$ the threshold is essentially infinite. That is, as β increases above the discount rate the threshold at which to act increases to a level that is unattainable in finite time. Since we assume the value of treatment is proportional to the level of infection, treatment is most valuable when the level of infection is large. As the transmission rate increases, the level of infection is growing faster and so it is beneficial to wait longer. For very fast spreading diseases ($\beta > 0.2$) the threshold level for the mean-reverting SDE can go above I_{max}, as with the case for large uncertainty, Figure 2. However, once again the logistic equation threshold remains within an attainable region (i.e. $V^* < V_{max}$), Figure 2. In particular, the threshold levels off close to V_{max} for large β (shown in Figure 2b).

The implications of the different models for decision makers can be seen in the policy plots in Figure 3. Under the logistic equation the waiting region is smallest, and so treatment will be applied earlier in the course of the epidemic, compared with assumptions based upon the GBM and the mean-

reverting equation. Furthermore, when there is great uncertainty (σ large) or the disease is fast spreading (β large) there is no region in which treatment should be deployed under GBM or the mean-reverting equation. In such cases the GBM and mean-reverting equation lead to significantly different conclusions for policy makers: GBM and the mean-reverting equation imply the decision maker should never apply treatment, while under the logistic equation treatment should be applied as soon as the proportion of infected area becomes large. Since the final size of the epidemic can be close to I_{max}, not applying treatment could lead to large losses. Therefore, when uncertainty is large or the disease is fast spreading, the logistic equation assumption is more appropriate than GBM or the mean-reverting equation since it provides a realistically attainable threshold at which to apply treatment.

4.3 Loss in value from delaying treatment too long

To quantify the potential economic impact from inaccurate assumptions regarding the uncertainty in future disease spread, we consider the value of the option to treat if the evolution of infection follows the logistic equation but the 'wrong' threshold, derived from either the mean-reverting or GBM model[1]. We find that treating at the wrong threshold leads to a loss in value, Figure 4. This loss arises because treatment thresholds are higher under mean-reversion or GBM than under the logistic assumption, and hence sub-optimally high when infection actually evolves according to the logistic equation. This reduces the value of treating at a sub-optimally high threshold (Figure 4, red and green lines), which fall below the standard NPV (Figure 4, dashed black line). The loss in value is greatest for GBM where 26% of the optimal value is lost by exercising sub-optimally late, i.e. at too high a threshold, while treating at the lower mean-reverting threshold leads to a lower loss of 5%[2]. Therefore, if the model used does not appropriately capture uncertainty in disease spread, then the excessive delay before treatment also implies that the full value of the option to treat is not realised. Figure 5 shows that the proportional loss from using the wrong assumptions increases with volatility for both the mean-reverting and GBM models. This is consistent with our earlier finding that the difference between the thresholds for the logistic and both the mean-reverting and GBM is greatest for large uncertainty. Also the proportional loss from using the mean-reverting thresholds is smaller than using the GBM thresholds since the mean-reverting threshold is closer to threshold for the logistic equation (Figure 1).

[1] This is calculated by solving the differential equation associated with the logistic equation but imposing the threshold values obtained from either the mean-reverting or the GBM models (see Appendix for details).
[2] As for standard GBM, the proportion of the optimal value obtained if the treatment occurs at the wrong threshold is independent of the value of treatment V.

5. Conclusion

The real options approach has been extensively applied within the literature to investigate the optimal timing of strategies to control disease and pest outbreaks. Uncertainty in disease spread is typically assumed to follow Geometric Brownian Motion due to its simplicity, rather than any epidemiological basis. However, ignoring the basic characteristics of disease transmission, namely logistic growth in the level of infection, in the SDE describing disease spread may result in treatment being deployed too late, and indeed not at all if the level of uncertainty is high. This has significant implications for the formulation of the real options approach used to inform disease control policy.

We have shown that the stochastic process describing uncertainty in disease spread can be derived directly from basic epidemiological principles. Comparing the mean-reverting and logistic models with the standard assumption of GBM we find the following key differences:

1) When uncertainty is large or the disease is fast spreading then the threshold value for GBM and the mean-reverting equation is unattainable, implying treatment should never be applied. This is not the case for the logistic equation, where the threshold at which to treat always lies in a realistic range.
2) The greatest threshold value at which to act is for GBM and is least for the logistic SDE. Therefore delaying treatment based on the results from the GBM or mean-reverting formulation may result in the late application of treatment.
3) The differences in the threshold at which to treat between the three models increases with increasing volatility and transmission rate. Therefore the implications of the different models will be most disparate when there is great uncertainty and the disease is fast spreading.

Describing the uncertainty in disease dynamics in terms of the logistic SDE leads to qualitatively similar implications for decision makers as GBM or the mean-reverting equation, namely that there is value to be gained from waiting to apply treatment. However, the important difference is that the logistic SDE implies that treatment should be applied earlier (i.e. at a lower level of total infected area) and that it is always optimal to apply treatment before the whole area becomes infected, even if the uncertainty is very large. In particular, if treatment is excessively delayed, by using thresholds from the GBM or mean-reverting models when uncertainty in the level of infection actually follows the logistic, then only a portion of the optimal value is obtained. As uncertainty, and hence the difference between the thresholds, increases this proportion decreases, and so when uncertainty is

high using predictions from the mean-reverting or GBM models leads to greater losses due to implementing treatment at the wrong threshold.

A further disadvantage of the GBM approach is that it restricts the range of parameter regimes that can be investigated since if $\beta \geq r$ then the threshold value at which to apply treatment is not finite. Since the time horizon for ecological management, such as forestry, is long, typically 40-60 years, the discount rate is usually taken to be less than the standard discount rate[3] (Pindyck 2006). Therefore, the assumption of GBM to characterise the uncertainty in the level of infection restricts application of the real options approach to slower spreading diseases.

Since the main aim of this paper is to investigate the effect of more realistic characterisations of disease uncertainty on the conclusions of the real options model, we have used the simplest formulation of the value gained from treatment application. That is we assume that the value of treatment is in the recovery of currently infected hosts, which would be applicable, for example, in the case of a pesticide or fungicide that kills the pest or pathogen upon application, allowing complete recovery of the previously infected hosts. However, often there is an additional value to applying treatment, namely that by reducing the level of infected hosts future infections are averted. That is, there is a sunk benefit associated with early application of treatment due to the reduction in potential future damage as a result of disease (Dixit and Pindyck 1994, 412–418). The decision problem could also be extended by including economic uncertainty (market risks) which is potentially correlated with disease risk as well as considering the interaction effects between economic and ecological uncertainty. We leave these for future work.

One key issue with using the real options framework to assess the impact of uncertainty on the optimal timing of disease control measures is that the form of the process that best represents uncertainty in disease spread is rarely discussed. We have shown here that assumptions surrounding disease spread uncertainty can have significant impacts on the conclusions of the real options model. Therefore it is important to justify the stochastic process used to describe the uncertainty associated with disease spread. The approach described here provides a convenient method for formulating the future uncertainty since it is derived directly from epidemiological principles of pathogen transmission. Epidemiological modelling has become a well-established field which uses mathematical models to describe the evolution of a disease outbreak based on epidemiological mechanisms of the spread of an infection. Such models have been successfully deployed to inform

[3] What the appropriate standard discount rate should be is itself a frequently discussed topic, (Gollier 2011), with previous studies using discount rate values of 8% (Sims and Finnoff 2012) and 10% (Ndeffo Mbah et al. 2010).

the structure of control measures within human (Choi et al. 2010), animal (Brooks-Pollock, Roberts, and Keeling 2014) and plant (Cunniffe et al. 2014) health. However, to date such modelling approaches have been largely ignored within the real options literature on the optimal timing of disease control.

This paper represents the first attempt to incorporate traditional epidemiological models of disease spread into a real options framework. Since this leads to greater complexity in the stochastic process describing disease spread, closed form solutions are no longer attainable and the problem must be solved numerically. However, we have shown that ignoring the epidemiological principles of disease spread can result in late application of treatment, or indeed may suggest treatment should never be applied. This result cautions against over-simplifying the description of uncertainty in disease dynamics within the real options approach to the optimal timing of disease control.

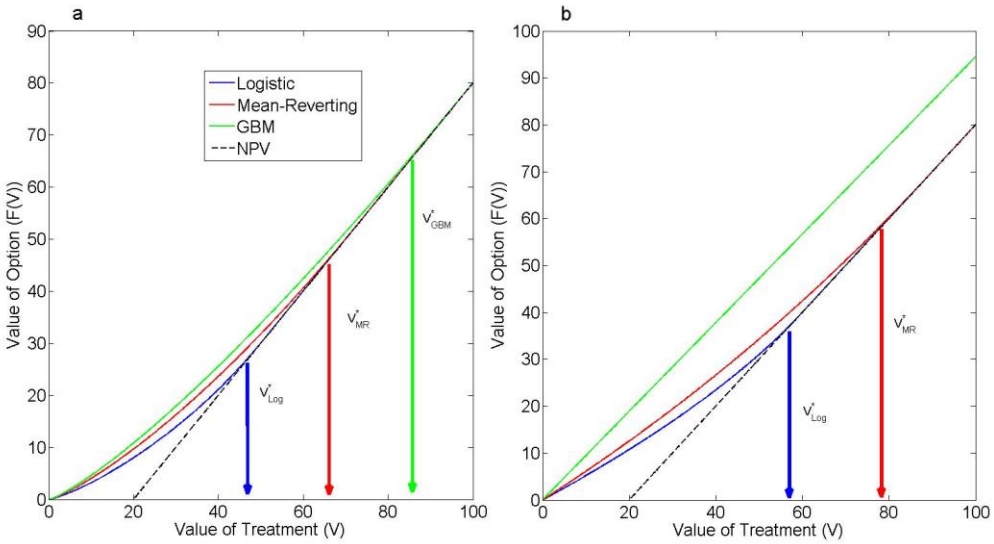

Figure 1: Value of the option to treat as a function of the value of treatment for the three different stochastic process assumptions. The standard NPV is shown as a dashed line. Plot (a) is for the case $\beta = 0.05$ and $r = 0.1$ and so $\beta < r$. Plot (b) is the case when $\beta = r = 0.1$, and it can be seen in this case that GBM never intersects the standard NPV, showing it is never optimal to apply treatment. The volatility is taken to be $\sigma = 0.5$, and the other parameter values are given in Table 1.

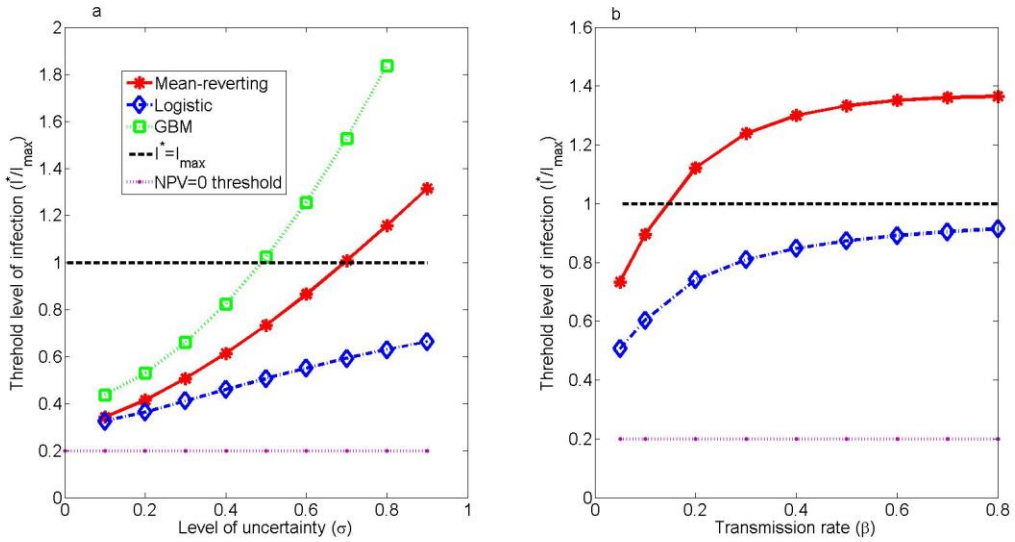

Figure 2: Threshold of level of infection (I^*/I_{max}): (a) as a function of volatility (σ) for all three stochastic processes and (b) as a function of the transmission rate (β) for the mean-reverting and logistic SDEs only. The dashed black line shows when $I^* = I_{max}$, and so it is clear that the mean-reverting and GBM can go above this natural boundary while this is not the case for the logistic SDE. Other parameter values are given in Table 1.

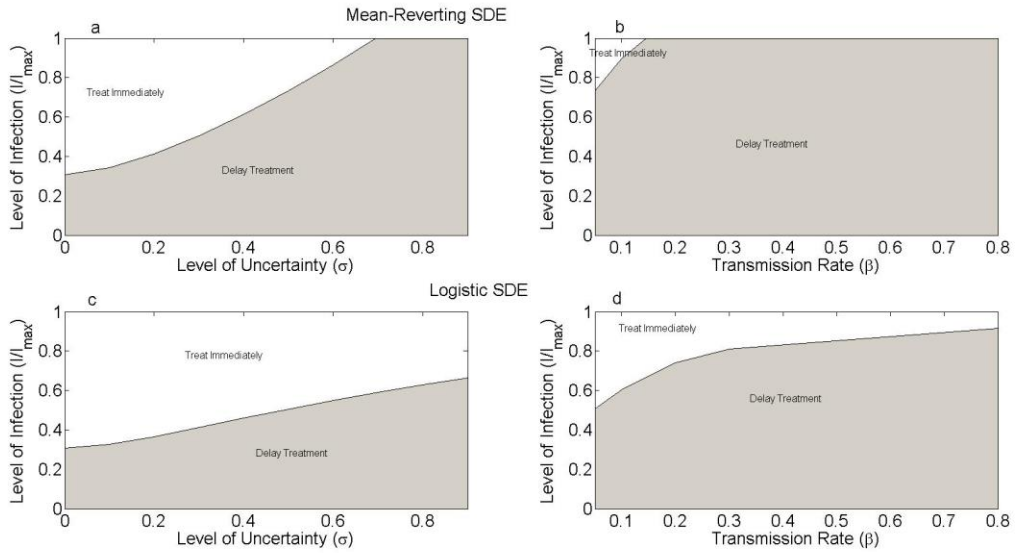

Figure 3: Policy plots showing the region in which treatment should be applied immediately and where treatment should be delayed for the mean-reverting SDE ((a) and (b)) and the logistic SDE ((c) and (d)) for different levels of uncertainty ((a) and (c)) and transmission rates ((b) and (d)). Other parameter values are given in Table 2.

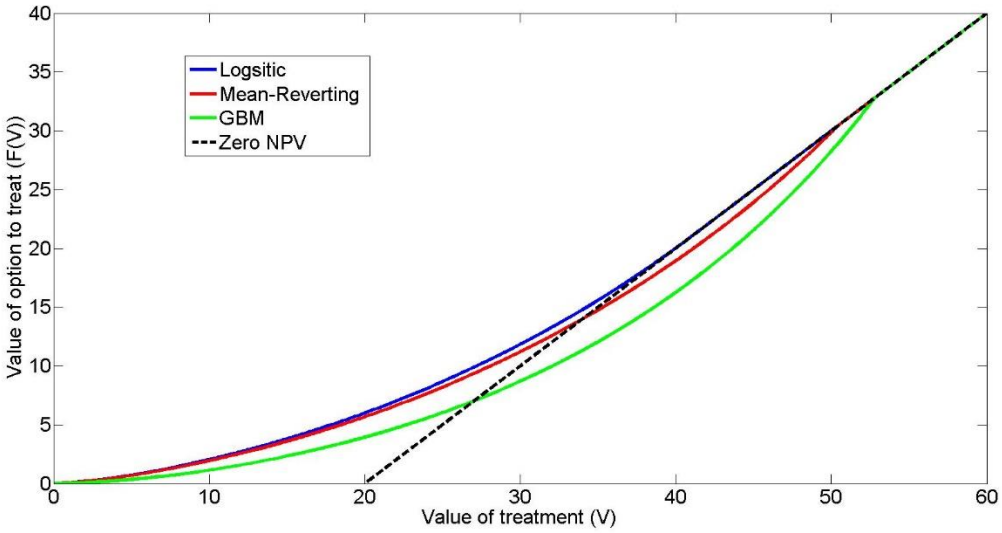

Figure 4: Value of the option to treat ($F(V)$) as a function of the treatment value (V) for the logistic equation when treatment is applied at the wrong threshold, namely the threshold obtained from the mean-reverting model (red line) and the GBM model (green line). Also shown are the value of the option when treatment is applied at the optimal threshold (blue line) and the standard NPV (black line), i.e. the value of the option when NPV is zero.

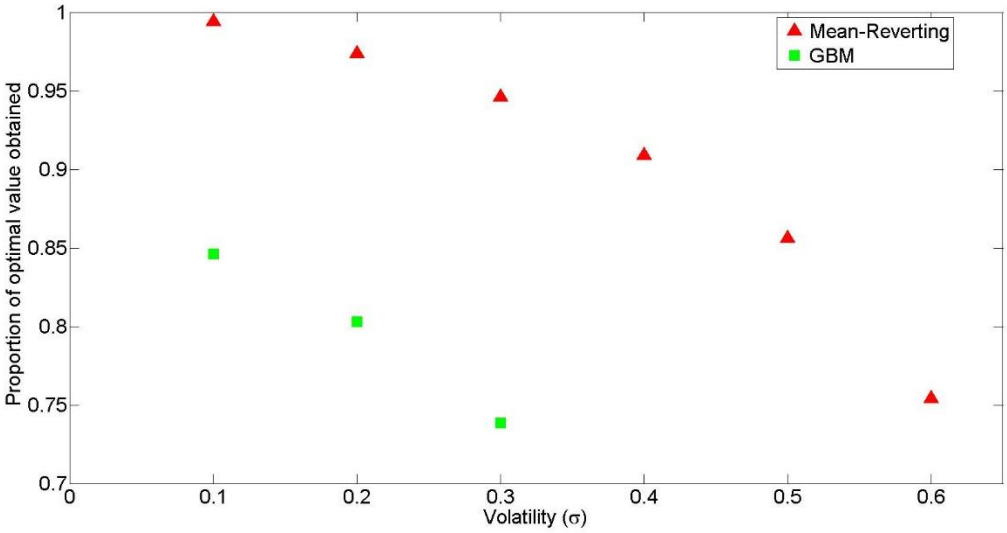

Figure 5: The proportion of the optimal value that is obtained under the logistic model when the treatment threshold from the mean-reverting (red triangles) and the GBM (green squares) models are enforced. The proportion of value obtained is shown as a function of the volatility (σ).

Bibliography

Black, F, and M Scholes. 1973. "The Pricing of Options and Corporate Liabilities." *Journal of Political Economy* 81 (3): 637–54.

Brooks-Pollock, Ellen, Gareth O. Roberts, and Matt J. Keeling. 2014. “A Dynamic Model of Bovine Tuberculosis Spread and Control in Great Britain.” *Nature* 511 (7508). Nature Publishing Group: 228–31. doi:10.1038/nature13529.

Choi, Yoon Hong, Mark Jit, Nigel Gay, Andrew Cox, Geoff P. Garnett, and William John Edmunds. 2010. “Transmission Dynamic Modelling of the Impact of Human Papillomavirus Vaccination in the United Kingdom.” *Vaccine* 28 (24). Elsevier Ltd: 4091–4102. doi:10.1016/j.vaccine.2009.09.125.

Cunniffe, Nik J., Francisco F. Laranjeira, Franco M. Neri, R. Erik DeSimone, and Christopher a. Gilligan. 2014. “Cost-Effective Control of Plant Disease When Epidemiological Knowledge Is Incomplete: Modelling Bahia Bark Scaling of Citrus.” *PLoS Computational Biology* 10 (8): e1003753. doi:10.1371/journal.pcbi.1003753.

Dixit, Avinash, and Robert Pindyck. 1994. *Investment Under Uncertainty*. Princeton University Press.

Gilligan, Christopher a. 2008. “Sustainable Agriculture and Plant Diseases: An Epidemiological Perspective.” *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 363 (1492): 741–59. doi:10.1098/rstb.2007.2181.

Gilligan, Christopher a, and Frank van den Bosch. 2008. “Epidemiological Models for Invasion and Persistence of Pathogens.” *Annual Review of Phytopathology* 46 (May): 385–418. doi:10.1146/annurev.phyto.45.062806.094357.

Gollier, Christian. 2011. “Pricing the Future : The Economics of Discounting and Sustainable Development.” *Princeton University Press*, no. 230589: 1–210.

Insley, Margaret. 2002. “A Real Options Approach to the Valuation of a Forestry Investment 1,” no. June 2000.

Keeling, Matt J, and Rohani Pejman. 2008. *Modeling Infectious Diseases In Humans and Animals*. Princeton University Press.

Large, E C, W J Beer, and J B E Patterson. 1946. “Field Trials of Copper Fungicides for the Control of Potato Blight; Spray Retention.” *The Annals of Applied Biology* 33 (1): 54–63.

Marcus, Ruth. 1991. “Deterministic and Stochastic Logistic Models for Describing Increase of Plant Diseases.” *Crop Protection* 10 (2): 155–59. doi:10.1016/0261-2194(91)90065-Y.

Marten, Alex L., and Christopher C. Moore. 2011. “An Options Based Bioeconomic Model for Biological and Chemical Control of Invasive Species.” *Ecological Economics* 70 (11). Elsevier B.V.: 2050–61. doi:10.1016/j.ecolecon.2011.05.022.

Ndeffo Mbah, Martial L, Graeme a Forster, Justus H Wesseler, and Christopher a Gilligan. 2010. “Economically Optimal Timing for Crop Disease Control under Uncertainty: An Options Approach.” *Journal of the Royal Society, Interface / the Royal Society* 7 (51): 1421–28. doi:10.1098/rsif.2010.0056.

Pindyck, Robert S. 2006. “Uncertianty in Environmental Economics.” *NBER Working Paper Series*.

Saphores, J. 2000. “Pest Population : A Real Options Approach” 82 (August): 541–55.

Sarkar, Sudipto. 2009. “Optimal Fishery Harvesting Rules under Uncertainty.” *Resource and Energy Economics* 31 (4): 272–86. doi:10.1016/j.reseneeco.2009.05.002.

Sims, Charles. 2011. “RESPONSE TO A STOCHASTIC INFESTATION,” no. June: 383–408.

Sims, Charles, and David Finnoff. 2012. “The Role of Spatial Scale in the Timing of Uncertain Environmental Policy.” *Journal of Economic Dynamics and Control* 36 (3). Elsevier: 369–82. doi:10.1016/j.jedc.2011.09.001.

Sims, Charles, and David Finnoff. 2013. “When Is a ‘wait and See’ Approach to Invasive Species Justified?” *Resource and Energy Economics* 35 (3). Elsevier B.V.: 235–55. doi:10.1016/j.reseneeco.2013.02.001.

Strange, Richard N, and Peter R Scott. 2005. “Plant Disease: A Threat to Global Food Security.” *Annual Review of Phytopathology* 43 (Figure 1): 83–116. doi:10.1146/annurev.phyto.43.113004.133839.

Sturrock, R. N., S. J. Frankel, a. V. Brown, P. E. Hennon, J. T. Kliejunas, K. J. Lewis, J. J. Worrall, and a. J. Woods. 2011. “Climate Change and Forest Diseases.” *Plant Pathology* 60 (1): 133–49. doi:10.1111/j.1365-3059.2010.02406.x.

Wilmott, P, S Howison, and J Dewynne. 1995. *The Mathematics of Financial Derivatives*. Cambridge University Press.

Appendix

Formulation of the Bellman Equation

The evolution in the value of treatment is given by the SDE,

$$dV = a(V,t)dt + b(V,t)dW,$$

Where the functions $a(V,t)$ and $b(V,t)$ are obtained using Ito’s Lemma.

At a given time t the decision authority either applies treatment (i.e. invests in treatment) and obtains the net gain $V(t) - C$, or they wait and obtain the expected change in the value of the option. Therefore, we can formulate the decision of whether or not to apply treatment as an optimal stopping problem, described by the following

$$F(V,t\,) = \max[(\,\mathrm{V} - C\,), e^{-r\,dt}\mathbb{E}[F(V + dV, t + dt)|V]].$$

The first term on the right is the return the decision authority would obtain if treatment were applied immediately while the second term is the value in retaining the option to treat in the future. Therefore, in the region where it is optimal to wait and not treat immediately, referred to as the *continuation region*, the second term on the right hand side of the equation will be larger. It follows that in the continuation region

$$rF(V,t\,)dt = \mathbb{E}[dF].$$

Since $V(t)$ is a stochastic process, then using Ito’s Lemma we can evaluate dF as follows,

$$dF = \frac{\partial F}{\partial t} + \frac{1}{2}\frac{\partial^2 F}{\partial V^2}(dV)^2 + \frac{\partial F}{\partial V}dV.$$

Substituting in the equation for dI, and noting that $\mathbb{E}[dW] = 0$, we obtain the following

$$\mathbb{E}[dF] = \frac{\partial F}{\partial t}dt + \frac{1}{2}b(V,t)^2\frac{\partial^2 F}{\partial V^2}dt + a(V,t)\frac{\partial F}{\partial V}dt,$$

where $a(V,t)$ and $b(V,t)$ are the drift and diffusion coefficients, respectively, of $V(t)$ (given in Table 1 in the main text).

Using the fact that $rFdt = \mathbb{E}[dF]$, we arrive at the following partial differential equation (PDE)

$$\frac{\partial F}{\partial t} + \frac{1}{2}b(V,t)^2\frac{\partial^2 F}{\partial V^2} + a(V,t)\frac{\partial F}{\partial V} - rF = 0,$$

with corresponding boundary conditions,

$$F(0,t) = 0$$

$$F(V_*,t) = V_* - C$$

$$\frac{\partial F}{\partial V}(V_*,t) = 1.$$

V_* is the value at which treatment should be applied immediately.

Simulation Methods

We solve the system of equations backwards in time from the end time T to the present day ($t = 0$), using a dynamic programming approach. We take T to be large ($T = 100$) so that we reach an equilibrium value for the threshold value, i.e. V_* is independent of time. So essentially we consider the problem over an infinite time horizon. At each time step the linear complimentary or free-boundary problem is solved using the Crank-Nicolson finite-difference method where successive-over-relaxation (SOR) is implemented to obtain a solution to the discretisation problem (Wilmott, Howison, and Dewynne 1995). In order to solve the problem numerically, an upper boundary condition must be stipulated in the case where it is not worthwhile treating immediately.

We assume that as V becomes large ${\partial^2 F}/{\partial V^2} = 0$ which is used to obtain the finite difference scheme at the upper boundary (Insley 2002). The values of model parameters and parameters within the numerical method used in simulations are given in Table 1.

Exercising the option at the wrong threshold

In order to determine the value of the option to treat if the wrong threshold is implemented in the logistic model we solve the PDE given by (6) again using the Crank-Nicolson finite-difference. However, we no longer have a free boundary problem, since we know the upper boundary at the threshold value (where the threshold value is taken from that obtained for the mean-reverting or GBM model). Let V^W be the threshold value obtained from the mean-reverting or GBM model. We set

$$F(V) = V - C, \qquad V \geq V^W.$$

Therefore we no longer need to set the second derivative with respect to V equal to zero since we now have a defined upper boundary condition.