Running Head: Approximate optimization of insect pests

Management of a stage-structured insect pest: an application of approximate optimization

Sean C. Hackett^{1*} and Michael B. Bonsall^{1,2}

¹ Mathematical Ecology Research Group, Department of Zoology University of Oxford, South Parks Road, Oxford, OX1 3PS

² St. Peter's College, New Inn Hall Street, Oxford, OX1 2DL, UK

* Correspondence author. E-mail: sean.hackett@zoo.ox.ac.uk

Abstract

Ecological decision problems frequently require the optimization of a sequence of actions over time where actions may have both immediate and downstream effects. Dynamic programming can solve such problems only if the problem dimensionality is sufficiently low. Approximate dynamic programming (ADP) provides a suite of methods applicable to problems of arbitrary complexity at the expense of guaranteed optimality. The most easily generalised method is the lookahead policy: a brute-force algorithm which identifies reasonable actions by constructing and solving a series of temporally truncated approximations of the full problem over a defined planning horizon. We develop and apply this approach to a pest management problem inspired by the Mediterranean fruit fly, *Ceratitidis capitata*. The model aims to minimise the cumulative costs of management actions and medfly-induced losses over a single 16-week season. The medfly population is stage-structured and grows continuously while management decisions are made at discrete, weekly intervals. For each week, the model chooses between inaction, insecticide application or one of six sterile insect release ratios. Lookahead policy performance is evaluated over a range of planning horizons, two levels of crop susceptibility to medfly and three levels of pesticide persistence. In all cases, the actions proposed by the lookahead policy are contrasted to those of a myopic policy which minimises costs over only the current week. We find that lookahead policies always out-performed a myopic policy and decision quality is sensitive to the temporal distribution of costs relative to the planning horizon: it is beneficial to extend the planning horizon when it excludes pertinent costs. However, longer planning horizons may reduce decision quality when major costs are resolved imminently. ADP methods such as the lookahead policy based approach developed here render questions intractable to dynamic programming amenable to inference but should be applied carefully as their flexibility comes at

the expense of guaranteed optimality. However, given the complexity of many ecological management problems, the capacity to propose a strategy that is “good enough” using a more representative problem formulation may be preferable to an optimal strategy derived from a simplified model.

Key-words: dynamic programming, approximate dynamic programming, lookahead policy, sequential decision problem, continuous state space, stage structure, sterile insect release,

Ceratitis capitata

Introduction

Ecological management problems span a range of disciplines from conservation to public health and are frequently presented as a sequential decision problem in which decisions must be made recurrently and actions taken now can influence the actions we may take in the future. Problems of this type can be represented as a dynamic program, a sequential decision problem (which may be stochastic) in which the system evolves over time and that satisfies key technical assumptions (Powell 2014). The system is summarised by a set of state variables that describe features -which may be qualitative or quantitative- required to model changes in the system induced by the passage of time and management actions and to evaluate decisions. For example, a simple single-species conservation problem could use the abundance of the focal species and remaining funds as state variables. In solving a dynamic program, the goal is to identify an optimal strategy for each state. That is, to find for each state the action that, when applied to that state produces the best outcome with respect to the specified time horizon, which may be finite or infinite (Powell 2011, 2014, Marescot et al. 2013). Thus, the chosen action must be optimal, not only with respect to the current state, but also the sequence of actions and states that will follow from this state as a consequence of the action. Classically, dynamic programs in ecology have been approached using dynamic programming, a method which partitions a complex, multi-period, decision problem into a sequential series of simpler sub-problems (Clark and Mangel 2000, Powell 2011, Marescot et al. 2013). The combined solutions for each sub-problem describe the optimal strategy for the full problem (Clark and Mangel 2000, Powell 2011, Marescot et al. 2013).

Traditionally, dynamic programming problems in ecology have been solved using Bellman's equation (also known as the dynamic programming equation) which characterises the optimal value V of a system in state S at time t by solving an expression of the form

$$V_t(S_t) = \max_a \{C_t(S_t, a_t) + V_{t+1}(S_{t+1})\} \quad [1]$$

to identify the action a which maximises the current $C_t(S_t, a_t)$ and future value $V_{t+1}(S_{t+1})$ of the decision (which assumes optimal behaviour from $t + 1$ onwards and may be discounted when appropriate). The Bellman equation defines an iterative algorithm which identifies the best mapping between states and actions by recursively identifying the action that optimises the sum of both current and future benefits for each combination of state, action and time. As equation [1] defines the optimal value of a state recursively in terms of both immediate and long term rewards, the downstream effects of a decision made now are accommodated automatically (Clark and Mangel 2000, Powell 2011, Marescot et al. 2013). Thus, the selected action will be optimal with respect to the time horizon of interest. Equation [1] is typically solved within a backwards recursion algorithm (for finite horizon problems) or as part of a value or policy iteration algorithm (for infinite horizon problems) all of which readily accommodate the stochasticity and non-linearity inherent to many ecological questions (Marescot et al. 2013).

However, irrespective of the chosen horizon, to solve a dynamic program using Bellman's equation requires that the problem be of sufficiently low dimensionality that the solution for each state, at each time (for finite horizon problems) can be compiled in a lookup table (Clark and Mangel 2000, Powell 2011, Marescot et al. 2013). Three curses of dimensionality that limit the potential applications of dynamic programming arise from this condition: should any one of (i) the state space, (ii) the action space or (iii) (for stochastic models) the outcome space grow

sufficiently large the problem becomes intractable to solution by traditional dynamic programming algorithms (Powell 2011). In some instances it is possible to reduce the dimensionality of the problem, e.g. by discretising a continuous state variable and then using linear interpolation (Clark and Mangel 2000), such that traditional algorithms may still be used (Shoemaker 1982). However, in reducing the dimensionality of the problem, we sacrifice realism which may be of use when making predictions in a management context and constrain the range of questions we can explore and exclude problems where dimensionality reduction is not possible.

In circumventing the curses of dimensionality, it is useful to distinguish between a dynamic program and the method used to solve it. Bellman's equation and dynamic programming are frequently conflated which constrains how dynamic programs are designed and solved (Marescot et al. 2013, Powell 2014). A dynamic program is solved using a policy, which is any rule for mapping feasible actions onto states (Powell 2014). Thus, Bellman's equation is not itself a dynamic program, but is both a means of describing an optimal policy (actions which satisfy equation [1] will always be optimal) and a valid algorithm for the solution of a dynamic program of sufficiently low dimensionality (Powell 2014). However, alternative policies for decision-making exist which are more permissive of high dimensional problems and capable of accommodating complications such as continuous state variables or continuous random outcomes. Such policies are collected together under the umbrella term 'approximate dynamic programming' (ADP). At the expense of guaranteed optimality, ADP methods can accommodate problems of arbitrary complexity (Powell 2011). In the context of complex ecological management problems, the trade of sub-optimal decisions against the ability to

consider a broader range of questions, and to incorporate greater realism, may be a fair one (e.g. Nicol *et al.* 2010; Nicol & Chadès 2011).

Approximate dynamic programming policies can be broadly subdivided into four policy classes based upon the type of approximation used and the feature of the problem approximated (Powell 2014). The simplest policy type is the cost function approximation (CFA) which expands upon a myopic policy, a policy which optimises only with respect to the current benefits of an action ($C_t(S_t, a_t)$ in equation [1]) by introducing a correction term (which may contain tuneable parameters). The correction term is designed to promote decisions deemed “good” from the perspective of the current objective and to increase the robustness of the myopic policy to uncertainty (Powell 2014, Powell and Meisel 2016). A CFA is neither approximating nor explicitly modelling the downstream effects of an action: it remains a myopic policy, concerned only with the single period effect of an action. However, myopic policies can prove surprisingly effective for some problems (Wilson *et al.* 2006) and a well-designed CFA may be of most utility where a myopic policy already produces satisfactory results (Powell and Meisel 2016).

The policy type most commonly associated with ADP is the value function approximation (VFA). A VFA employs either parametric or non-parametric methods to define a function approximating the future value of downstream states resulting from an action we take now (the term $V_{t+1}(S_{t+1})$ in equation [1]) (Powell 2011, 2014, Powell and Meisel 2016). The dynamic program is then solved and actions suggested as if the VFA holds true. VFAs are extremely versatile and can be tailored to a diverse range of problems however, they are most useful when the value of the future is easily approximated and we can be confident of the fidelity of the approximation as a poorly specified VFA can significantly underperform (Powell 2014, Powell and Meisel 2016). Policy function approximations (PFAs) also utilise functional approximations

but, unlike VFAs, they do not attempt to approximate an actions downstream value. Instead, PFAs define a function representing the policy itself and then tune this function until it produces satisfactory outcomes (Powell 2014, Powell and Meisel 2016). Thus, while both PFAs and VFAs utilise functional approximations, they differ significantly in how they approach the optimization problem. VFAs explicitly solve the optimization problem as if the constructed VFA provides a true representation of future costs and benefits. That is, a VFA identifies an optimal strategy with respect to the chosen functional approximation. PFAs do not solve the original optimization problem, they are not searching for an optimal strategy as the decision rule at for each period is fixed by the architecture of the PFA. Instead, a PFA is searching for the parameter values which produce the best long-run behaviour. As such, PFAs are most usefully applied to problems where the relationship between state and action is clearly defined and easily observable (for example, determining how much water to add to a reservoir) or where a good decision rule can be easily conceived (for example, top up the reservoir by X litres whenever it falls below level Y) (Powell 2011, 2014, Powell and Meisel 2016). However, it follows from their comparative rigidity that PFAs struggle to accommodate non-stationary processes and parameters and are best suited to stationary applications (Powell 2014, Powell and Meisel 2016).

The fourth policy type is the lookahead policy, also referred to as the rolling horizon procedure and model predictive control. Lookahead policies identify actions by solving a series of truncated approximations of the full problem (Nicol et al. 2010, Powell 2011, 2014, Powell and Meisel 2016). For example, the optimal action that could be taken during the current period t for a system in state S_t is that which maximises the benefits accrued from the current period t to the terminal period T where, for the purposes of this example, the values of t and T are arbitrary (with the caveat that $t < T$). Where the curses of dimensionality render this full problem

intractable it may still be possible to identify good actions by optimising between the current period t and some truncated planning horizon $H < T$ (Powell 2014, Powell and Meisel 2016). Thus, lookahead policies do not depend upon functional approximations and instead essentially approximate the problem itself. Consequently, while the actions suggested by a lookahead policy are unlikely to be optimal with respect to the full horizon of interest, lookahead policies are extremely versatile. The runtime of a lookahead policy is independent of the magnitude of the state space (Nicol et al. 2010, Powell 2011), they readily accommodate non-stationarity (Powell 2014, Powell and Meisel 2016) and they can be applied to problems where the value of a state is a complex function or is subject to sufficient uncertainty that it is difficult to reliably approximate (Powell 2014, Powell and Meisel 2016). Thus we follow the example of Nicol *et al.* (2010) and Nicol & Chadès (2011) in suggesting that lookahead policies are an excellent candidate for application to ecological dynamic programs that are otherwise intractable to exact solution methods.

In what follows we more fully explain the operation of a lookahead policy and then construct and apply one to the intraseasonal management of a stage structured pest population inspired by the Mediterranean fruit fly (or medfly), *Ceratitis capitata*. The medfly is a widely distributed highly invasive fruit pest of substantive economic significance (Diamantidis et al. 2009, 2011, Szyniszewska and Tatem 2014). Medfly can be long-lived (Carey 2011, Diamantidis et al. 2011), reproduce continuously and exhibit overlapping generations (Barclay 2005, 2016), properties which are not conducive to traditional dynamic programming methods. At present, medfly management is highly reliant upon insecticidal sprays (Vontas et al. 2011) but sterile insect releases – in which pest suppression results from sustained mass-releases of sterilised male pests - have also provided effective areawide suppression in some regions (Enkerlin 2005). Thus, we

consider a lookahead policy seeking to integrate these tools to suppress the simulated medfly population. We find that lookahead policies are able to produce reasonable action sequences which effectively integrate both bi-sex lethal insect releases and insecticidal sprays for the management of Medfly. Furthermore, they offer significant potential for future extension to accommodate more complex problems such as those presented by the recent diversification of sterile insect methods, facilitated by developments in genetic engineering (Alphey et al. 2010, Bourtzis et al. 2016, Leftwich et al. 2016, Harvey-Samuel et al. 2017) .

Methods

Overview

We present a sequential decision model for the combined application of self-limiting insect releases (male-only releases, the progeny of which do not survive to the reproductive stage) and insecticidal sprays to the intraseasonal management of a polyvoltinous pest inspired by the medfly. As medfly populations typically reproduce continuously and exhibit overlapping generations (Barclay 2005, 2016), we represent the medfly population with a set of stage-structured delay differential equations. This representation creates a large state space with multiple continuously valued state-variables, rendering the problem intractable to traditional dynamic programming methods. Thus, we approach this problem using a lookahead policy, a variant of approximate dynamic programming that identifies good (if sub-optimal) actions by solving the optimization problem over a truncated time horizon (Nicol et al. 2010, Powell 2011, 2014).

We assume a manager seeking to minimise the cumulative costs of medfly induced losses and medfly management over the course of the season. A season is sub-divided into W discrete

weeks. At the start of each week w , a decision is made, based upon the current state, and applied to the medfly population instantaneously. Note that we adopt the convention of Powell (2016) in defining the state of the system to be the minimally dimensioned function of history required to model the transitions of the system from a given point in time onwards. Using this definition, the state can include historical values from prior time points (Powell 2014), which enables us to represent the medfly population with delay equations. While actions are taken at discrete, one-week intervals, their effects and the dynamics of the medfly population are resolved in continuous time. Each week spans t_w time increments (hence, a day spans $\frac{t_w}{7}$ time increments), giving a total season duration of $t_w W$ time increments. That is, the state after t_w time increments informs the decision for the subsequent week.

For any given week w the manager may elect to do nothing or to choose from one of two alternative actions: insecticidal foliar sprays and self-limiting releases with an early-acting bisex lethal effect. The decision to spray or to perform a self-limiting release instantaneously updates the relevant control variable and incurs an immediate point cost. Spraying sets φ the instantaneous insecticide-induced mortality rate, to its maximum value, $0 < \varphi_{max} < 1$ (note that this implementation prevents spraying eradicating the population, there will always be survivors) from which it subsequently decays with time at instantaneous rate α . Spraying is assumed to be uniform across the entire landscape; there are no refugia. For each spraying action the manager incurs a constant point cost c^{spray} . If the decision is made to release self-limiting insects, then the manager may choose from a range of permissible release ratios. Thus, the costs of self-limiting releases scale with both the magnitude of the medfly population and the release ratio. If the basic cost of a single self-limiting insect is c^{ste} then the cost of releasing m^{ste} self-limiting males is $c^{ste} m^{ste}$. Irrespective of the release ratio, released insects are assumed to be distributed

uniformly across the landscape. In addition to management costs, feeding by larval medfly continuously devalues the crop from one week to the next at instantaneous rate γ which broadly reflects the susceptibility of the crop to damage. Thus, large values of γ portray a more vulnerable crop while smaller values indicate a greater tolerance to medfly.

The medfly model

We consider a closed medfly population with a 1:1 sex ratio. We assume that male and female medfly exhibit equivalent parameter values and so we explicitly track only female numbers. The medfly population is subdivided into the following sequential stages: eggs e , larvae l , pupae p and pre-ovipositional females f^{pre} which then enter one of two reproductive stages, f^{wild} or f^{ste} based upon whether they are mated to a wild-type male (superscript *wild*) or a self-limiting male (superscript *ste*). Neither sprays nor self-limiting releases can directly affect the abundance of the juvenile stages. Thus, in the absence of competition and other exogenous sources of mortality, eggs, larvae and pupae are subject only to recruitment, maturation and natural mortality (which reflects viability as opposed to competition or predation). For any given life stage i , μ_i denotes the instantaneous natural mortality rate and τ_i defines the duration of the stage in days. The proportion of individuals surviving stage i to enter stage $i + 1$ is $\sigma_i = \exp(-\mu_i \tau_i)$. If wild-mated females oviposit at an age-invariant rate of λ eggs per female per day then the instantaneous changes in egg, larval and pupal abundance are found as

$$\frac{de}{dt} = \lambda f^{wild}(w) - \mu_e e(t) - \lambda f^{wild}(t - \tau_e) \sigma_e \quad [2]$$

$$\frac{dl}{dt} = \lambda f^{wild}(t - \tau_e) \sigma_e - \mu_l l(t) - \lambda f^{wild}(t - \tau_e - \tau_l) \sigma_e \sigma_l \quad [3]$$

$$\frac{dp}{dt} = \lambda f^{wild}(t - \tau_e - \tau_l) \sigma_e \sigma_l - \mu_p p(t) - \lambda f^{wild}(t - \tau_e - \tau_l - \tau_p) \sigma_e \sigma_l \sigma_p \quad [4].$$

Changes in adult abundance are more complex. First, adults are subject to additional exogenous mortality from insecticide φ and second, pre-ovipositional adults may be recruited into either the wild-mated or sterile-mated stage. Pre-ovipositional females attain reproductive maturity after τ_{fpre} days at which point they are instantaneously mated and, if mated to a wild-type male, immediately commence oviposition. While female re-mating by female medfly is possible (Kraaijeveld and Chapman 2004) we assume that released and wild-type males are equivalent with respect to post-copulatory traits (Barclay 2005). However, insecticide-induced mortality complicates the determination of pre-ovipositional stage survivorship σ_{fpre} as mortality over the duration of this stage will not be constant in the presence of insecticide. Recruitment into the two mated stages must reflect this potential for variation in pre-ovipositional survivorship. Therefore, we let $M_a(t)$ denote the total instantaneous mortality rate experienced by all adult medfly at time t which is found as

$$M_a(t) = \mu_a + \varphi(t) \quad [5]$$

where μ_a is the baseline instantaneous natural mortality rate for all adult stages and $\varphi(t)$ is the instantaneous rate of insecticide-induced mortality at time t . We then use $M_a(t)$ to determine the change in pre-ovipositional survivorship over the duration of the stage

$$\frac{d\sigma_{fpre}}{dt} = \sigma_{fpre}(t) \left(M_a(t - \tau_{fpre}) - M_a(t) \right) \quad [6].$$

Using equations [5] and [6] the instantaneous rates of change for pre-ovipositional females, wild-mated females and sterile-mated females are found as

$$\begin{aligned} \frac{df^{pre}}{dt} = & \lambda f^{wild}(t - \tau_e - \tau_l - \tau_p) \sigma_e \sigma_l \sigma_p - M_a(t) f^{pre}(t) \\ & - \lambda f^{wild}(t - \tau_e - \tau_l - \tau_p - \tau_{f^{pre}}) \sigma_e \sigma_l \sigma_p \sigma_{f^{pre}}(t) \quad [7] \end{aligned}$$

$$\begin{aligned} \frac{df^{wild}}{dt} = & \left(\frac{m^{wild}(t)}{m^{wild}(t) + m^{ste}(t)} \right) \lambda f^{wild}(t - \tau_e - \tau_l - \tau_p - \tau_{f^{pre}}) \sigma_e \sigma_l \sigma_p \sigma_{f^{pre}}(t) \\ & - M_a(t) f^{wild}(t) \quad [8] \end{aligned}$$

$$\begin{aligned} \frac{df^{ste}}{st} = & \left(\frac{m^{ste}(t)}{m^{wild}(t) + m^{ste}(t)} \right) \lambda f^{wild}(t - \tau_e - \tau_l - \tau_p - \tau_{f^{pre}}) \sigma_e \sigma_l \sigma_p \sigma_{f^{pre}}(t) \\ & - M_a(t) f^{ste}(t) \quad [9] \end{aligned}$$

where m^{wild} denotes the number of wild-type males and m^{ste} is the number of self-limiting males. At any given instant, the number of wild type males is simply found as $m^{wild} = f^{pre} + f^{wild} + f^{ste}$.

The control variables, m^{ste} and φ , are not static between decision periods; insecticide toxicity declines with time and self-limiting insects are removed from the population by either natural mortality or by exposure to insecticide residues. We capture the change in insecticide mortality as a simple exponential decay

$$\frac{d\varphi}{dt} = -\alpha\varphi(t) \quad [10]$$

where α determines the persistence of the toxin. Self-limiting male abundance declines according to

$$\frac{dm^{ste}}{dt} = -M_{m^{ste}}(t) m^{ste}(t) \quad [11]$$

where $M_{m^{ste}}$ denotes the self-limiting equivalent of M_a and is found at a given instant in time as

$$M_{m^{ste}}(t) = \mu_{m^{ste}} + \varphi(t) \quad [12]$$

where $\mu_{m^{ste}}$ is the baseline self-limiting male mortality rate. Note that equation [12] assumes that self-limiting males do not differ from wild-type adults in their susceptibility to insecticide. From one week to the next larval medfly erode the value of the crop at rate

$$\frac{dc}{dt} = \gamma(2l(t)) \quad [13]$$

where γ is the per capita larval feeding rate and the doubling of the larval population accounts for damage caused by male larvae.

Equations [2-13] define the transition function for our model. That is they determine how the state of the system evolves over time and in response to management actions (Powell 2014, Powell and Meisel 2016). Given that we define the state of a system S_t to be the minimally dimensioned function of history required to model the transitions of that system from time t onwards, then the state for this model, at the start of any given week w is summarised as

$$\begin{aligned} S_w = & \left(e(t), l(t), p(t), f^{pre}(t), f^{wild}(t), f^{ste}(t), f^{wild}(t - \tau_e), f^{wild}(t - \tau_e - \tau_l), f^{wild}(t \right. \\ & - \tau_e - \tau_l - \tau_p), f^{wild}(t - \tau_e - \tau_l - \tau_p - \tau_{fpre}), M_a(t), M_a(t \\ & \left. - \tau_e) \sigma_{fpre}(t), \varphi(t), m^{ste}(t), M_s(t), c(t) \right) \quad [14]. \end{aligned}$$

The manager's objective is to minimise losses from larval medfly feeding and the cumulative cost of medfly suppression over the duration of the season. Thus, the optimal decision for any week a_w^* is that which satisfies

$$a_w^* = \arg \min_{a_w} \{ C(S_w, a_w) + \sum_{w'=w+1}^W (C(S_{w'}, a_{w'})) \} \quad [15]$$

where $C(S_w, a_w)$ is a contribution function determining the full cost accrued between week w and week $w + 1$ which is found as

$$C(S_w, a_w) = c_a + \int_{t'=t_0}^{t'=t_w} \gamma(2l(t')) \quad [16]$$

where c_a denotes the point cost for the action implemented at the start of week w . The value of S_{w+1} is found via equations [2-13].

The Lookahead Policy

For a sufficiently small state space equation [15] could be solved exactly via backwards recursion (Clark and Mangel 2000, Marescot et al. 2013) however, the state space for this problem is complex and continuously valued. Thus, we approach this problem using a lookahead policy, a variant of approximate dynamic programming that uses a series of computationally tractable approximations of the model to suggest actions (Nicol et al. 2010, Powell 2011, 2014). We adopt the terminology of Powell (2014) and Powell & Meisel (2015) and define equation [15] to be the base model, the full problem that we seek to solve. Solution of the base model for all combinations of state, action and time would return a series of decisions which are optimal with respect to both their immediate and future costs but this is unattainable for a problem of this magnitude.

Rather than address the base model directly, lookahead policies proceed by constructing a series of truncated approximations of equation [15], termed lookahead models (Powell 2014, Powell and Meisel 2016). For example, to identify the action to be taken in week w given state S_w we construct the deterministic lookahead model

$$a_w = \arg \min_{a_w} \{C(S_w, a_w) + \sum_{w'=w+1}^H (C(\tilde{S}_{ww'}, \tilde{a}_{ww'}))\} \quad [17]$$

where $H < W$ is a suitably short planning horizon chosen to capture important behaviours while maintaining computational tractability. Tildes distinguish between variables from the base model from those in the lookahead model while the ww' subscript denotes a variable encountered in week w' within the lookahead model when solving for the action to be taken in week w (Powell 2014). Thus, $\tilde{a}_{ww'}$ is the action implemented in week w' within the lookahead model when solving for a_w . This notation emphasises the distinction between the lookahead model (equation [17]) and the base model (equation [15]). Solving a lookahead model is not equivalent to solving the base model but by solving the lookahead model over a sufficiently long horizon, good, if sub-optimal, actions will be taken that will outperform those suggested by a purely myopic policy (Nicol et al. 2010, Powell 2011, 2014, Powell et al. 2012, Powell and Meisel 2016).

For the current application, we note that the immediate contribution from a given state-action pair is resolved over a one-week interval. Thus, a myopic policy applied to this problem would simulate the effects of each of the A available actions from S_w to S_{w+1} and produce a decision based upon which of the A alternative downstream states had the least cost. By repeating the process of iteratively applying each action to each state for up to H periods beyond $w + 1$, a lookahead model may attain a better representation of an actions value. In optimising over the interval $(w, \dots, w + H)$ a lookahead model generates a tree of decisions and states. By identifying the terminal state with the least cost and then solving back towards the original input state at the root of the tree the model isolates the action which is optimal with respect to the limited horizon of the lookahead model (Nicol et al. 2010, Powell 2011, 2014). The returned decision is then applied to state S_w and the model increments to week $w + 1$ where a new lookahead model is

constructed and solved. While a lookahead model will return an optimal sequence of decisions over the full lookahead horizon H only the first of these decisions is implemented; the downstream decisions in the sequence were made only to improve the decision to be made now and are discarded (Powell 2011, 2014, Powell and Meisel 2016).

Parameterisation

The demographic model is parameterised using data presented in Diamantidis *et al.* 2009 (for adult parameters) and Diamantidis *et al.* 2011 (for juvenile parameters). Specifically, we utilise parameters derived from a Hawaiian medfly population which Diamantidis *et al.* (2011) would classify as short-lived. Stage-specific instantaneous mortality rates μ_i for the pre-adult stages (eggs, larvae and pupae) were obtained from the stage duration and survivorship data presented by Diamantidis *et al.* (2011) as

$$\mu_i = -\frac{\log(\sigma_i)}{\tau_i} \quad [18].$$

The baseline natural instantaneous mortality rate for adult medfly μ_a was specified as the inverse of the longevity of an adult female where longevity refers to the average number of days a female medfly is expected to live post-eclosion. This was taken to be 50 days, based upon Figure 3 of Diamantidis *et al.* (2009). A constant value of 5 female eggs per female per day was taken to be a reasonable representation of daily egg production λ and the duration of the pre-ovipositional stage was taken to be 10 days (Diamantidis *et al.* 2009). Costs are unitless and intended only to provide a reasonable approximation of the relative costs of releases versus foliar sprays. We specify the point cost of spraying as $c^{spray} = 4000$ and the cost per sterile insect released as $c^{ste} = 0.003$. Thus, the cost of spraying is constant, irrespective of medfly abundance, while the absolute cost of a sterile release depends upon the total number of insects released which will

vary with the chosen release ratio and the medfly population density. As this model only considers management over a single season, we do not consider the discounting of future costs (Zhang and Swinton 2009). We assume that the medfly population is highly susceptible to foliar spraying (Vontas et al. 2011) and so specify the maximum mortality rate from insecticide exposure as $\varphi_{max} = 0.9$. We assume that self-limiting males are equivalent to wild-type insects with respect to their natural mortality rate and their susceptibility to insecticide. The baseline values for all model parameters are summarised in Table 1.

Simulations

All simulations were performed using code written in R version 3.3.3 and were initiated with a population of 20,000 medfly; 10,000 mated females and 10,000 adult males. A season lasts for $W = 16$ weeks. The continuous time dynamics separating each decision period were approximated using an RK4 solver adapted to deal with time delays with a resolution of $t_w = 700$ time increments per week. The performance of three alternative lookahead policies with progressively longer planning horizons, $H = \{1,2,3\}$, were evaluated relative to one another and to a myopic policy, $H = 0$. We chose the progression $H = \{1,2,3\}$ to provide a simple illustration of the benefits and limitations of lookahead policies. In practice, the planning horizon would be determined by practical constraints such as the accuracy of available forecasts and the time available for computation. Costs from larval medfly infestation accumulated at one of two rates, $\gamma = 0.001$ or $\gamma = 0.01$, where the lower value represents a comparatively resilient crop in which large larval infestations are required before substantive losses occur and the larger value depicts a more susceptible crop for which the inverse is true. Once applied, insecticides decayed exponentially at one of three exponential rates, $\alpha = \{0.175, 0.35, 0.7\}$ where $\alpha = 0.35$ is taken to be the default value and depicts an insecticide which decays to low levels of toxicity ($\varphi < 0.1$)

within approximately 1 week while $\alpha = 0.7$ and $\alpha = 0.175$, respectively reflect insecticides with greater and lesser persistence. We focus upon the parameters α and γ to demonstrate how the actions suggested by each policy change to accommodate variation in the urgency of medfly management (as determined by α , the susceptibility of the crop) and the extent to which each policy is able to plan actions in a way which effectively accounts for pesticide persistence (determined by γ).

Results

Effect of planning horizon (H) length under different damage regimes

For any given scenario, a myopic policy ($H = 0$) is always outperformed by a lookahead policy ($H \geq 1$). Even a minimal lookahead policy which evaluates actions over only one additional time period ($H = 1$) relative to a myopic policy confers substantial benefits in terms of cost reduction (Fig.1a). By evaluating an action over only a single week the myopic policy overemphasises short-term costs which biases it towards the application of insecticide and inaction (Fig.1b). Sterile releases are employed less frequently and only at low-mid release ratios (Fig.1b,e). The lookahead policies more frequently utilise self-limiting releases, often implementing larger release ratios than the myopic policy, and less frequently elect to perform no action (Fig.1c,d,f,g).

The rate at which costs inflicted by the larval stage accrue interacts with the planning horizon H to influence action selection. Figure 2 illustrates cost trajectories, action frequencies and release ratios for a medfly population feeding on a resilient crop with a damage accrual rate of $\gamma = 0.001$. Counterintuitively, a reduction in the damage rate does not guarantee a substantive reduction in costs relative to a scenario with a more damaging pest (compare Fig.1a and Fig.2a).

When pest-induced costs accumulate slowly, the negative effects of a poor decision are temporally delayed and a longer planning horizon is needed to incorporate these costs into decision-making. A myopic policy is now strongly biased against self-limiting releases and more frequently elects to perform no action, incurring significant costs in doing so (Fig.2a,b,e). A lookahead policy improves the quality of decisions, prompting a more proactive approach with fewer periods of inaction (Fig.2c,d), but releases are still employed less frequently relative to a scenario with a higher damage rate (Fig.1c,d) and smaller release ratios are more likely to be selected (Fig.2f,g).

The interaction between the extent of the planning horizon and the threat posed by the pest influence the temporal sequence of actions in addition to their frequency. Figure 3 depicts the chronological sequence of actions implemented over the course of the 16-week season for the policies illustrated in Figure 1 ($\gamma = 0.01$, the right-hand column) and Figure 2 ($\gamma = 0.001$, the left-hand column). The set of alternate actions is listed on the y-axis. For $\gamma = 0.01$ (Fig.2d-f) the initial action is always to spray, irrespective of the policy. Spraying provides rapid knockdown of the founding population (which is comprised entirely of reproductive adults), reducing both the initial rate of population growth and the cost of larger release ratios. In contrast, when $\gamma = 0.001$, this initial spray is not identified as cost-effective until the planning horizon of the lookahead policy is extended to $H = 2$ (Fig.2c).

Taken together, the results shown in Figures 1-3 illustrate how lookahead policies reduce costs by employing proactive and diverse action sequences relative to a myopic policy. Thus, while still inevitably sub-optimal, lookahead policies can make interventions that prevent the medfly population from growing to highly damaging densities while a myopic policy will only intervene once the pest has attained a density for which significant damage is imminent. The extent to

which this is possible depends upon the value of γ , the rate at which larvae induced costs accumulate relative to H , the length of the planning horizon. As the value of γ is increased, costs resulting from a poor decision are increasingly experienced in the near-term future. This enables even a myopic policy to suggest actions which will, to an extent, pre-empt costs before they occur (Fig.3a,d). However, lookahead policies still provide significant cost reductions (Fig.1a) by accounting for the possibility of damage over a longer time span. When the value of γ is low the inverse applies and the consequences of a poor decision may lie outside the planning horizon and so are not factored into decision making. This promotes policies that favour spraying and inaction (Fig.3a,b) unless a longer planning horizon is used (Fig.3c).

Effect of spray persistence upon the integration of spraying and sterile releases

Varying the persistence of sprays, as proxied by α the instantaneous rate at which adult medfly mortality from spray exposure declines, influences the ease with which lookahead policies are able to effectively integrate spraying with sterile releases. Figures 1-3 utilised a decay rate of $\alpha = 0.35$ such that spray toxicity decays to trivial values within the space of one simulated week, minimising the residual killing of sterile insects at the point of release. Figure 4 illustrates the cumulative costs and implemented action frequencies for the management of a medfly population with either a myopic policy or one of two lookahead policies when the insecticide decay rate is doubled ($\alpha = 0.7$, parameters are otherwise identical to those of Figure 1).

Reducing spray persistence substantially increases costs for all policies (Fig4.a) but especially for a myopic policy where costs accumulate rapidly. All policies adopt spraying as their most frequent action (Fig.4 b-d) but only the lookahead policies make multiple uses of sterile releases, which contribute towards reducing management costs relative to the myopic policy which utilises only a single sterile release. The increase in spraying frequency with reduced spray

persistence appears contradictory but emphasises the role of an initial knockdown in rendering self-limiting releases cost-effective over short time scales. Fewer adults are killed by each spraying event and so more will survive to enter the mated stage and produce offspring which themselves now have a greater likelihood of surviving to reproductive maturity. Thus, greater effort is required to reduce the adult population to a level for which, relative to the current planning horizon, releases appear cost-effective. The preference for larger release ratios may facilitate the persistence of the sterile population at a non-trivial level in the face of recurrent episodes of insecticide application.

Figure 5 illustrates cost trajectories and action frequencies when $\alpha = 0.175$ and all other parameters are identical to those in Figure 4. Reducing the insecticide decay rate reduces overall management costs for all policies, while still permitting effective integration of sprays with sterile releases. The reduction in costs is intuitive, if each insecticide spray persists at non-trivial concentrations for longer, fewer sprays are required for effective knockdown (Fig.5,b-c). All three policies, including the myopic policy now make more frequent use of sterile releases than foliar sprays but lookahead policies achieve significant cost reductions by favouring larger release ratios (Fig.5, e-g). Given that the persistence of the spray necessarily entails that some sterile insects will be lost to residue exposure subsequent to their release, it is striking that sterile releases are still identified as effective actions irrespective of the potential for overlap between releases and insecticide residues. The utility of releases in conjunction with a persistent spray is likely due to the improved knockdown potential of the spray rendering releases cost-effective even to a myopic policy. The preference of the lookahead policies for larger release ratios is likely promoted by the increased affordability of such ratios when knockdown from spraying is effective, and by the need to offset the loss of sterile insects to spray residues.

The contrast between persistent ($\alpha = 0.175$) and transient ($\alpha = 0.7$) insecticides is further explored in Figure 6, which illustrates the action sequences for each policy shown in Figures 4 and 5. For a persistent spray, the sequential application of sprays does not occur: spraying events are spread across the season and may be interspersed with sequential sterile releases (Fig.6a-c). The inverse is true of a transient spray: releases are periodic and, when using a lookahead policy, spread across the season, while sprays are applied sequentially (Fig.6d-f).

Lookahead policies and sub-optimality

For the pest management problem considered here a lookahead policy always outperforms a myopic policy but actions suggested by a lookahead policy are optimal only with respect to the lookahead model which derived them and are not guaranteed to constitute optimal actions with respect to the base model. Figure 7 shows the accumulation of costs over time for three lookahead policies ($H = \{1,2,3\}$) and three levels of insecticide persistence ($\alpha = \{0.175,0.35,0.7\}$) applied to medfly populations with instantaneous damage rates of $\gamma = \{0.001,0.01\}$. For any given set of conditions, the most basic lookahead policy, $H = 1$, is always outperformed by an alternative lookahead policy with a longer planning horizon. However, extending the planning horizon is not a consistent guarantee of improved performance: a lookahead policy with a planning horizon of $H = 3$ underperforms relative to a policy with a shorter horizon of $H = 2$ in Fig.7 a,d and e while it provides additional cost reductions in Fig. 7b,c and f. This discrepancy can be understood in terms of the distribution of costs over time for each scenario and the implementation of the lookahead policy itself. When the costs of an error are concentrated in the near-term and insecticides offer effective knockdown (Fig.7d,e) such that sterile releases can be effectively evaluated over short time windows (Fig.7a) then extending the planning horizon to $H = 3$ produces action sequences less effective than those for $H = 2$.

Actions taken in solving a lookahead model are intended to improve the decision made in the current time period. Extending the time horizon too far when costs/benefits are concentrated in the near-term future may, therefore, lead to decisions being based upon information that is less relevant to the current period and actions that are optimal with respect to a future that never comes to pass. However, for instances where the costs of a poor decision are concentrated in the future (Fig.7b) or sprays provide ineffective knockdown then extending the planning horizon is beneficial.

Discussion

We approached the management of a hypothetical continuously reproducing, stage-structured pest population inspired by the Mediterranean fruit fly using a lookahead policy that identifies appropriate actions by solving a temporally truncated approximation of the full problem. A lookahead policy consistently returned action sequences with superior integration of both self-limiting insect releases and foliar sprays than those of a myopic policy (*e.g.* Fig.1), even in instances where effective knockdown was difficult to achieve (Fig.4) or insecticide residues were persistent and potentially threatened self-limiting insects (Fig.5). This is congruent with the findings of similar models (Nicol et al. 2010, Nicol and Chadès 2011) and a promising indicator that there is scope for lookahead policies to provide a valuable tool for the management of populations with state spaces that are difficult to capture within a traditional dynamic programming framework.

While lookahead policies always outperformed a myopic policy, the cost-efficiency of a given lookahead policy was determined by the distribution of costs over time relative to the extent of the planning horizon H (Fig.1, Fig.2, Fig.4, Fig.5, & Fig.7). Recall that costs arriving outside of the planning horizon cannot be accounted for by an unmodified lookahead policy (Powell 2011,

2014). When the costs resulting from a poor decision are mostly deferred into the future (γ is low) extending the planning horizon was consistently beneficial. In contrast, extending the planning horizon too far when costs were concentrated in the near-term future (γ is high) can be counter-productive (Fig.7d, e). That is, incorporating the values of more distant states into the decision process when costs are imminent can produce lower-quality decisions. This could likely be remediated by the introduction of a discount rate. Discounting of future costs would devalue the terminal states of a given lookahead model, placing greater emphasis upon on the short-term behaviour of the system while still potentially enabling any benefits to a longer horizon to persist (Nicol et al. 2010). The effects of insecticide persistence can be understood in similar terms. Inefficient knockdown effectively emphasises the costs of control in the short-term relative to the long-term costs accruing from larvae which is equivalent to displacing the costs of errors into the future. Thus, transient sprays require longer planning horizons to effectively evaluate the benefits of an action, especially when larvae increase costs rapidly (Fig.7f). The opposite reasoning applies when spray persistence is increased; the costs of control are deemphasised and costs from the pest dominate decision-making.

The sensitivity of lookahead policy performance to the degree of overlap between the chosen planning horizon and the distribution of costs over time has implications for the further application of lookahead policies to the management of insect pest populations. The extent of the planning horizon will need to be chosen with reference to the rate at which costs from the damaging life stage(s) accrue and the life history of the pest. The former determines how many insects will be deemed too many, and the latter determines both the duration of the damaging stage(s) and the rapidity with which additional individuals are recruited into the damaging stage. Furthermore, consideration will need to be given to how the negative effects of the damaging life

stage accumulate. For example, the timescale for the negative effects of an agricultural pest will be intimately connected to the duration of the crop-damaging life stage but, if the pest also vectors pathogens then negative effects may persist beyond the lifetime of any one pest. Similar issues apply to the management of insect vectors of human pathogens, such as mosquitoes, where the benefits of management are frequently expressed in terms of the number of disease cases averted (Alphey et al. 2011). For such problems, a longer planning horizon may be required to ensure that decisions at least partially account for the benefits of averting additional disease cases or, where the costs of the disease and the rate of spread are well-understood and predictable, a hybrid lookahead policy (Powell 2014, Powell and Meisel 2016) could be employed using a cost function approximation to summarize the costs of cases arising beyond the planning horizon.

For clarity, we assumed that sterile releases comprised self-limiting insects with an early-acting bi-sex lethal effect: the simplest form of sterile release to model. However, self-limiting methods are diverse with respect to both their mode of action and the timing of any lethal effects (Alphey et al. 2010, Bourtzis et al. 2016, Leftwich et al. 2016, Harvey-Samuel et al. 2017). For example, male-selecting self-limiting strains for the suppression of medfly have been developed and are more likely to be available for medfly management in the near future than a bi-sex lethal strain (Gong et al. 2005, Morrison et al. 2009, Leftwich et al. 2014). For such male-selecting strains, given that the male offspring from self-limiting matings survive, the suppression potential of a male-selecting strain is reduced relative to that of a bi-sex self-limiting strain (Alphey et al. 2009). Furthermore, females inheriting the lethal allele are not killed until the pupal stage (Gong et al. 2005, Leftwich et al. 2014). This may increase the length of the planning horizon required for a lookahead policy to usefully evaluate such a release. However, while we did not here

consider the potential for the evolution of resistance to insecticides (Vontas et al. 2011), male-selecting insect releases may offer additional utility in instances where the evolution of resistance to other pest management tools is considered a threat to future management success (Alphey et al. 2007, 2009, Harvey-Samuel et al. 2015).

In addition to their ability to at least partially capture the downstream effects of contemporary decisions, there are four additional reasons to recommend lookahead policies for ecological management problems of high dimensionality. First, they do not depend upon functional approximations and so readily accommodate time-dependent processes and non-stationary parameters or functions (Powell 2014). For example, the current model could be extended to model changes in the vulnerability of fruits to attack over the course of a season (Zhang and Swinton 2012). Second, by explicitly modelling future state transitions lookahead policies naturally account for non-separable interactions between state variables which would otherwise be difficult to capture with an approximation based upon a parametric function (Powell 2014). Third, they readily incorporate forecasts of future events, even where the forecasts are updated or revised over time (Powell 2014). Taken together with the first property, an indifference to non-stationarity, this renders the combined online application of lookahead policies and adaptive management to information-dense problems a possibility. Lookahead policies are agnostic to the particular set of equations and parameters used to model the focal population and so can accommodate the updates required to account for *e.g.* structural uncertainty (Boettiger et al. 2015, Chadès et al. 2017). Finally, the time required to solve a problem using a lookahead policy is independent of the magnitude or complexity of the state space: state spaces of arbitrary complexity can theoretically be accommodated (Nicol et al. 2010, Powell et al. 2012). The runtime is instead determined by the extent of the planning horizon, the number of alternative

actions and the range of random outcomes. While we considered a deterministic problem, lookahead policies can account for stochasticity (Nicol et al. 2010, Nicol and Chadès 2011, Powell 2011, Powell and Meisel 2016) and, utilising methods from stochastic programming, can accommodate continuously valued random variables (Powell 2014). The structural agnosticism of lookahead policies entails that it would also be possible to extend this framework to include delays between the selection and implementation of an action.

With respect to our chosen example of a medfly suppression problem, we have presented a relatively basic representation for ease of understanding. However, the flexibility of lookahead policies offers the potential for substantive refinements which would augment the utility of the model to pest management professionals. Given a specific crop and locale, reliable estimates for both the costs of actions and potential costs of medfly induced losses could be utilised.

Furthermore, restrictions could be placed upon the timing and frequency of insecticide applications, particularly in relation to the timing of the harvest. The susceptibility of the crop to medfly could be allowed to vary over the course of the season, potentially resulting in periods in which medfly are present but cannot yet attack the crop. Additionally, forecasts of pertinent environmental variables such as temperature or precipitation and how they interact with the ecology of the pest, the growth of the crop and the utility of available actions could be incorporated. Alternatively, where such detailed models are unwarranted or the data for their formulation is lacking then the basic framework provided here could instead be built upon to provide intuition regarding how, for example, different forms of transgenic release (see above) could be utilised in conjunction with other management tools.

However, lookahead policies also carry several key disadvantages. In its basic form, a lookahead policy is a brute-force method, proceeding by rote application of the same procedure to each

decision period, which is indifferent to the structure of the problem and so cannot utilise advantageous structural features to accelerate computation or improve decision quality as, for example, a value function approximation could (Powell et al. 2012, Powell 2014). The requirement that a new lookahead model be solved for each time period may also slow computation relative to alternative methods. Lookahead policies will only explore states that are directly connected to the input state and will not generalise beyond these states without augmentation. Thus, for infinite horizon problems where the global value of a given state is pertinent it may be necessary to augment a lookahead policy with a means of more fully enumerating the state space (for example, by incorporating machine learning methods (Nicol and Chadès 2011)). Finally, we reiterate that a decision which is optimal with respect to a lookahead model is unlikely to be optimal with respect to the base model (Powell 2014).

Lookahead policies are imperfect solutions to imperfect problems. They are not a replacement for exact methods or myopic heuristics but an extension: they represent an additional tool to be applied when we are required to address management questions beyond the scope of exact methods. They are also not an invitation to complicate a problem for its own sake: a given problems state should remain a “minimally dimensioned function of history” (Powell 2014) and ADP methods should be invoked where this minimally dimensioned function exceeds the bounds of exact methods. If we remain aware of their shortcomings, ADP methods such as lookahead policies broaden the range of questions which can be asked and could represent a valuable tool for ecologists seeking to work with managers to produce good, if sub-optimal, solutions to detail-rich problems.

Acknowledgements

Sean Hackett is supported by BBSRC industrial CASE PhD studentship BB/K011936/1. The work was also supported, in part, by BBSRC grant BB/L00948X/1. We thank N.I.Morrison and M.Koukidou for their valuable input in developing the medfly component of the model. The authors declare no competing interests.

Literature Cited

- Alphey, L., M. Benedict, R. Bellini, G. G. Clark, D. a Dame, M. W. Service, and S. L. Dobson. 2010. Sterile-Insect Methods for Control of Mosquito-Borne Diseases: An Analysis. *Vector-Borne and Zoonotic Diseases* 10:295–311.
- Alphey, N., L. Alphey, and M. B. Bonsall. 2011. A Model Framework to Estimate Impact and Cost of Genetics-Based Sterile Insect Methods for Dengue Vector Control. *PLoS ONE* 6:e25384.
- Alphey, N., M. B. Bonsall, and L. Alphey. 2009. Combining pest control and resistance management: synergy of engineered insects with Bt crops. *Journal of Economic Entomology* 102:717–32.
- Alphey, N., P. G. Coleman, C. a Donnelly, and L. Alphey. 2007. Managing insecticide resistance by mass release of engineered insects. *Journal of Economic Entomology* 100:1642–1649.
- Barclay, H. 2005. Mathematical models for the use of sterile insects. Pages 147–174 *in* V. Dyck, J. Hendrichs, and A. S. Robinson, editors. *Sterile Insect Technique*. 1st edition. Springer Netherlands.
- Barclay, H. J. 2016. Determination of the sterile release rate for stopping growing age-structured populations. *International Journal of Pest Management* 62:40–54.

- Boettiger, C., M. Mangel, and S. Munch. 2015. Avoiding tipping points in fisheries management through Gaussian process dynamic programming. *Proceedings of the Royal Society B* 282:8–11.
- Bourtzis, K., R. S. Lees, J. Hendrichs, and M. J. B. Vreysen. 2016. More than one rabbit out of the hat: Radiation, transgenic and symbiont-based approaches for sustainable management of mosquito and tsetse fly populations. *Acta Tropica* 157:115–130.
- Carey, J. R. 2011. Biodemography of the Mediterranean fruit fly: Aging, longevity and adaptation in the wild. *Experimental Gerontology* 46:404–411.
- Chadès, I., S. Nicol, T. M. Rout, M. Péron, Y. Dujardin, J. B. Pichancourt, A. Hastings, and C. E. Hauser. 2017. Optimization methods to solve adaptive management problems. *Theoretical Ecology* 10:1–20.
- Clark, C. W., and M. Mangel. 2000. *Dynamic State Variable Models in Ecology: Methods and Applications*. Oxford University Press.
- Diamantidis, A. D., J. R. Carey, C. T. Nakas, and N. T. Papadopoulos. 2011. Population-specific demography and invasion potential in medfly. *Ecol Evol* 1:479–488.
- Diamantidis, A. D., N. T. Papadopoulos, C. T. Nakas, S. Wu, H. G. Müller, and J. R. Carey. 2009. Life history evolution in a globally invading tephritid: Patterns of survival and reproduction in medflies from six world regions. *Biological Journal of the Linnean Society* 97:106–117.
- Enkerlin, W. R. 2005. Impact of fruit fly control programmes using the sterile insect technique. Pages 651–676 *in* V. Dyck, J. Hendrichs, and A. S. Robinson, editors. *Principles and Practice*

in Area- Wide Integrated Pest Management. Springer, Dordrecht.

Gong, P., M. J. Epton, G. Fu, S. Scaife, A. Hiscox, K. C. Condon, G. C. Condon, N. I. Morrison, D. W. Kelly, T. Dafa'alla, P. G. Coleman, and L. Alphey. 2005. A dominant lethal genetic system for autocidal control of the Mediterranean fruitfly. *Nature Biotechnology* 23:453–456.

Harvey-Samuel, T., T. Ant, and L. Alphey. 2017. Towards the genetic control of invasive species. *Biological Invasions* 19:1683–1703.

Harvey-Samuel, T., N. I. Morrison, A. S. Walker, T. Marubbi, J. Yao, H. L. Collins, K. Gorman, T. G. E. Davies, N. Alphey, S. Warner, A. M. Shelton, and L. Alphey. 2015. Pest control and resistance management through release of insects carrying a male-selecting transgene. *BMC Biology* 13:49.

Kraaijeveld, K., and T. Chapman. 2004. Effects of male sterility on female remating in the Mediterranean fruitfly, *Ceratitis capitata*. *Proceedings of the Royal Society B-Biological Sciences* 271:S209–S211.

Leftwich, P. T., M. Bolton, and T. Chapman. 2016. Evolutionary biology and genetic techniques for insect control. *Evolutionary Applications* 9:212–230.

Leftwich, P. T., M. Koukidou, P. Rempoulakis, H.-F. Gong, A. Zacharopoulou, G. Fu, T. Chapman, A. Economopoulos, J. Vontas, and L. Alphey. 2014. Genetic elimination of field-cage populations of Mediterranean fruit flies. *Proceedings of the Royal Society B* 281:20141372-.

Marescot, L., G. Chapron, I. Chades, P. L. Fackler, C. Duchamp, E. Marboutin, and O. Gimenez.

2013. Complex decisions made simple: A primer on stochastic dynamic programming. *Methods in Ecology and Evolution* 4:872–884.
- Morrison, N. I., D. F. Segura, K. C. Stainton, G. Fu, C. A. Donnelly, and L. S. Alphey. 2009. Sexual competitiveness of a transgenic sexing strain of the Mediterranean fruit fly, *Ceratitidis capitata*. *Entomologia Experimentalis et Applicata* 133:146–153.
- Nicol, S. C., I. Chadès, S. Linke, and H. P. Possingham. 2010. Conservation decision-making in large state spaces. *Ecological Modelling* 221:2531–2536.
- Nicol, S., and I. Chadès. 2011. Beyond stochastic dynamic programming: A heuristic sampling method for optimizing conservation decisions in very large state spaces. *Methods in Ecology and Evolution* 2:221–228.
- Powell, W. B. 2011. *Approximate Dynamic Programming*. 2nd ed. John Wiley & Sons, Inc., Hoboken.
- Powell, W. B. 2014. Clearing the Jungle of Stochastic Optimization. *Bridging Data and Decisions*:109–137.
- Powell, W. B. 2016. A Unified Framework for Optimization under Uncertainty. *Page Informs Tutorials in Operations Research* 2016.
- Powell, W. B., and S. Meisel. 2016. Tutorial on Stochastic Optimization in Energy—Part II: An Energy Storage Illustration. *IEEE Transactions on Power Systems* 31:1468–1475.
- Powell, W. B., H. P. Simao, and B. Bouzaiene-Ayari. 2012. Approximate dynamic programming in transportation and logistics: a unified framework. *EURO Journal on Transportation and Logistics* 1:237–284.

- Shoemaker, C. 1982. Optimal integrated control of univoltine pest populations with age structure. *Operations Research* 30:40–61.
- Szyniszewska, A. M., and A. J. Tatem. 2014. Global assessment of seasonal potential distribution of mediterranean fruit fly, *Ceratitis capitata* (Diptera: Tephritidae). *PLoS ONE* 9:1–13.
- Vontas, J., P. Hernández-Crespo, J. T. Margaritopoulos, F. Ortego, H.-T. Feng, K. D. Mathiopoulos, and J.-C. Hsu. 2011. Insecticide resistance in Tephritid flies. *Pesticide Biochemistry and Physiology* 100:199–205.
- Wilson, K. A., M. F. McBride, M. Bode, and H. P. Possingham. 2006. Prioritizing global conservation efforts. *Nature* 440:337–340.
- Zhang, W., and S. M. Swinton. 2009. Incorporating natural enemies in an economic threshold for dynamically optimal pest management. *Ecological Modelling* 220:1315–1324.
- Zhang, W., and S. M. Swinton. 2012. Optimal control of soybean aphid in the presence of natural enemies and the implied value of their ecosystem services. *Journal of Environmental Management* 96:7–16.

Data Availability

The R code generated by this project will be uploaded to the Dryad Digital Repository.

Tables

Table 1: Summary of model parameters, their meanings and, where applicable, their origin.

Instantaneous mortality rates for juvenile life stages may be derived using equation [18]. The mortality rate for adult stages is found as the inverse of adult female longevity.

Parameter	Meaning	Point	Citation
		estimate/mean	
λ	Female eggs laid per		
	female per day	5	Diamantidis et al 2009
τ_e	Egg stage duration		
	(days)	2	Diamantidis et al 2011
σ_e	Egg stage		
	survivorship	0.89	Diamantidis et al 2011
τ_l	Larval stage duration		
	(days)	5.92	Diamantidis et al 2011
σ_l	Larval stage		
	survivorship	0.865	Diamantidis et al 2011
τ_p	Pupal stage duration		
	(days)	10.4	Diamantidis et al 2011
σ_p	Pupal stage		
	survivorship	99.3	Diamantidis et al 2011
τ_{fpre}	Pre-ovipositional		
	stage duration (days)	10	Diamantidis et al 2009
Adult female longevity			
	(days)	50	Diamantidis et al 2009

	Point cost of insecticide application	4000	N/A
c^{spray}			
	Point cost per self-limiting male released	0.003	N/A
c^{ste}			
	Instantaneous insecticide decay rate	0.175, 0.35, 0.7	N/A
α			
	Maximum Instantaneous insecticide mortality rate	0.9	N/A
φ_{max}			
	Instantaneous rate at which costs/losses from larval medfly accrue	0.001, 0.01	N/A
γ			
W	Season length (weeks)	16	N/A
	Number of time increments comprising a week	700	N/A
t_w			

Figure Legends

Figure 1: Comparisons of (a) the cumulative cost, (b-d) frequency of implemented actions and (e-g) sterile release ratios for three alternate policies attempting to manage a hypothetical medfly population) feeding upon a susceptible crop, $\gamma = 0.01$, over a 16-week season using insecticidal foliar sprays and sterile insect releases. Only a single action may be implemented in any given week. Black depicts a myopic policy, $H = 0$, while red and blue depict lookahead policies with planning horizons of $H = 1$ and $H = 2$, respectively where H is measured in weeks. Costs have no units and adult medfly mortality from insecticide exposure decays at instantaneous rate $\alpha = 0.35$. The population was initiated with 10,000 mated adult female medfly and a 1:1 sex ratio was assumed.

Figure 2: Comparisons of (a) the cumulative cost, (b-d) frequency of implemented actions and (e-g) sterile release ratios for three alternate policies attempting to manage a hypothetical medfly population) feeding upon a resilient crop, $\gamma = 0.001$, over a 16-week season using insecticidal foliar sprays and sterile insect releases. Only a single action may be implemented in any given week. Black depicts a myopic policy, $H = 0$, while red and blue depict lookahead policies with planning horizons of $H = 1$ and $H = 2$, respectively where H is measured in weeks. Costs have no units and adult medfly mortality from insecticide exposure decays at instantaneous rate $\alpha = 0.35$. The population was initiated with 10,000 mated adult female medfly and a 1:1 sex ratio was assumed.

Figure 3: Sequences of implemented actions for three alternative policies for the management of a simulated medfly population over 16 weeks. For all panels the full range of alternative actions available in each decision period are listed on the y-axis and adult medfly mortality from insecticide exposure decays at instantaneous rate $\alpha = 0.35$. Costs from larval feeding accrue at

instantaneous rate $\gamma = 0.001$ (a resilient crop) in the left column (a-c) and $\gamma = 0.01$ (a susceptible crop) in the right column (d-f). Black points indicate actions implemented by a myopic policy ($H = 0$) while red and blue points illustrate actions implemented by lookahead policies with planning horizons of $H = 1$ and $H = 2$, respectively where H is measured in weeks. The population was initiated with 10,000 mated adult female medfly and a 1:1 sex ratio was assumed.

Figure 4: Comparisons of (a) the cumulative cost, (b-d) frequency of implemented actions and (e-g) sterile release ratios for three alternate policies attempting to manage a hypothetical medfly population) feeding upon a susceptible crop, $\gamma = 0.01$, over a 16-week season using transient insecticidal foliar sprays and sterile insect releases. Only a single action may be implemented in any given week. Black depicts a myopic policy, $H = 0$, while red and blue depict lookahead policies with planning horizons of $H = 1$ and $H = 2$, respectively where H is measured in weeks. Costs have no units and adult medfly mortality from insecticide exposure decays at instantaneous rate $\alpha = 0.7$. The population was initiated with 10,000 mated adult female medfly and a 1:1 sex ratio was assumed.

Figure 5: Comparisons of (a) the cumulative cost, (b-d) frequency of implemented actions and (e-g) sterile release ratios for three alternate policies attempting to manage a hypothetical medfly population) feeding upon a susceptible crop, $\gamma = 0.01$, over a 16-week season using persistent insecticidal foliar sprays and sterile insect releases. Only a single action may be implemented in any given week. Black depicts a myopic policy, $H = 0$, while red and blue depict lookahead policies with planning horizons of $H = 1$ and $H = 2$, respectively where H is measured in weeks. Costs have no units and adult medfly mortality from insecticide exposure decays at

instantaneous rate $\alpha = 0.175$. The population was initiated with 10,000 mated adult female medfly and a 1:1 sex ratio was assumed.

Figure 6: Sequences of implemented actions for three alternative policies for the management of a simulated medfly population over 16 weeks. For all panels the full range of alternative actions available in each decision period are listed on the y-axis and costs from larval feeding accrue at instantaneous rate $\gamma = 0.01$. Adult medfly mortality from insecticide exposure decays at instantaneous rate $\alpha = 0.175$ (a persistent insecticide) in the left column (a-c) and $\alpha = 0.7$ (a transient insecticide) in the right column (d-f). Black points indicate actions implemented by a myopic policy ($H = 0$) while red and blue points illustrate actions implemented by lookahead policies with planning horizons of $H = 1$ and $H = 2$, respectively where H is measured in weeks. The population was initiated with 10,000 mated adult female medfly and a 1:1 sex ratio was assumed.

Figure 7: Accumulation of costs over time for three alternate lookahead policies attempting to manage a hypothetical medfly population (initiated with 10,000 mated adult females) over a 16-week season using insecticidal foliar sprays and sterile insect releases. Line colour indicates the extent of the planning horizon in weeks utilised by a given lookahead policy: red for $H = 1$, blue for $H = 2$ and orange for $H = 3$. Costs from larval feeding accrue at instantaneous rate $\gamma = 0.001$ (a resilient crop) in the left column (a-c) and $\gamma = 0.01$ (a susceptible crop) in the right column (d-f). Adult medfly mortality from insecticide exposure decays at instantaneous rate $\alpha = 0.175$ (a persistent insecticide) in the top row (a,d), $\alpha = 0.35$ in the middle row (b,e) and $\alpha = 0.7$ (a transient insecticide) in the bottom row (c,f).