Optimisation of multiple-phase systems

Graeme J. Doole

School of Agricultural and Resource Economics,

Faculty of Natural and Agricultural Sciences,

University of Western Australia, 35 Stirling Highway,

Crawley, Western Australia 6009.

October 12, 2005

gdoole@cyllene.uwa.edu.au

Abstract:

Many important economic problems concern a dynamic choice between alternate phases. Examples are determining the optimal time to switch between alternative energy sources or multiple crops. This significance has motivated a substantial theoretical literature generalising the necessary conditions of Optimal Control Theory to multiple-phase problems. However, gaining detailed insight into the practical management of these systems is difficult because suitable numerical methods are not available. In particular, traditional gradient techniques are ineffective because of the piecewise definition of the performance index. This paper resolves this deficiency through the presentation of a flexible and efficient optimisation algorithm based on a set of necessary conditions derived for finite-time multiple-phase problems. Its effectiveness is demonstrated in an application to a multiple crop system in Western Australia.

I would like to thank the Cooperative Research Centres for Australian Weed Management and Plant-Based Management of Dryland Salinity for funding this research.

I. Introduction

The Maximum Principle of Optimal Control Theory (Pontryagin et al. 1962) has been utilised extensively in economics (Arrow and Kurz 1970; Seierstad and Sydsaeter 1987; Kamien and Schwartz 1991) because of its intuitive economic interpretation (Dorfman 1969) and the significant methodological extensions to this theory developed in other fields of study, such as engineering. However, despite this broad application, there has been limited treatment of multiple-phase systems. These consist of multiple alternate regimes, each characterised by its own dynamical system, of which only one may be active at each point in time. Selecting between individual crops to plant on a given area of land is one example (Mueller, Schilizzi, and Tran 1999). Other examples are determining the optimal time to switch between alternative energy sources (Tomiyama 1985; Tomiyama and Rossana 1989) and identifying the optimal time for a government to abolish a policy, such as a capital control (Makris 2001). In actual fact, many economic decisions may be studied more precisely if cast as multiple-phase problems. For example, in production theory, these are a natural means of representing choices between the alternative technologies available to a firm, such as natural and artificial recovery of petroleum (Amit 1986).

Piecewise-constant control variables can be used to switch between multiple stages if the ordering of phases is freely determined. However, this approach is only suitable for small problems given its inherent combinatorial complexity (Papadimitriou and Steiglitz 1982). In addition, the computation of switching times can be difficult if transition costs exist. These limitations have motivated the analysis of multiple-phase systems in which the sequence of stages is pre-assigned. This approach is relevant to many important economic problems, such as the alternative technology or government policy examples discussed above. This form of multiple-phase problem requires optimisation over two levels. The first level concerns the determination of the optimal duration of each regime through calculation of the optimal transition times and also, perhaps, the terminal time. The second level entails optimisation of a phase between its endpoints. If there is no control exercised during the duration of a stage, the second level is not required and the multiple-phase system may, instead, be studied in a financial options framework (Dixit and Pindyck 1994).

In contrast, generalisation of the necessary conditions of standard optimal control (Pontryagin et al. 1962; Kamien and Schwartz 1991) is required if control variables are specified within independent phases. Such conditions have been derived for two-stage systems with costless transition (Tomiyama 1985; Tomiyama and Rossana 1989) and switching costs (Amit 1986), even though the former may be treated utilising standard control, as indicated above. The framework of Amit (1986) has also been extended to include three stages (Mueller, Schilizzi, and Tran 1999) and an infinite horizon (Makris 2001). Makris (2001) also discussed the solution of a system with *n* regimes and an infinite horizon. This included switching conditions but did not consider the finite-time case.

Although this significant body of theory exists, it is difficult to study the practical management of multiple-phase problems because there appears to be no suitable optimisation algorithms available to economists. Gradient-based methods have been widely employed to solve control problems in discrete-time. However, these are difficult to implement within a multiple-phase system because the objective functional has, by definition, discontinuous derivative(s) with respect to the control variable(s) (see Section II). Simultaneous optimisation of phase duration is also problematic because the corresponding gradients are not readily available. Dynamic programming, at least conceptually, is suited to solving the latter problem. However, it will be prohibitively large in most circumstances because of the need to represent uncertain phase transitions.

This paper contributes to theory through the derivation of a set of necessary conditions for a finite-time multiple-phase system with *n* regimes, positive switching costs, and different endpoint constraints. These conditions are used to construct a flexible optimisation algorithm for the solution of multiple-stage problems. Its effectiveness is demonstrated in a multiple crop problem of significant complexity. This algorithm appears to be the first in Economics to solve general multiple-phase problems[1] and provides practitioners with the opportunity to study these systems in considerable detail, a luxury not afforded in the analytical constructs to which they have previously been restricted.

The model and necessary conditions are presented in Section 2. Section 3 describes the numerical algorithm and discusses strategies to help implementation. An application of this algorithm to a multiple crop problem is presented in Section 4. Section 5 presents conclusions

[1] Mueller, Schilizzi, and Tran (1999) solved a three-stage problem utilising numerical methods. However, their approach is limited to control systems where the constituent differential equations can be solved explicitly. The algorithm presented here does not face such restrictions (see Section III).

and recommendations for further research. The derivation of the necessary conditions is described in an Appendix.

II. Model and Necesssary Conditions

This section formally defines a model for a multiple-phase system and presents a set of necessary conditions required for its solution.

DEFINITION 2.1. *A general multiple-phase system can be assumed to incorporate an m-dimensional state vector* $x(t)=\{x^1(t),x^2(t),...,x^m(t)\}$ *of continuous functions, piecewise continuous differentiable over the time interval* $t=[t_0,...,t_n]$ *and belonging to* $X \in R^m$*, and a v-dimensional vector of control functions* $u(t)=\{u^1(t),u^2(t),...,u^v(t)\}$*, piecewise continuous in* $t=[t_0,...,t_n]$ *and belonging to* $U \in R^v$*. The state variables fixed at the initial time are* x_0^i *for i=[1,2,…,c]. In addition, the state variables fixed at the terminal time are* x_n^i *for i=[1,2,…,d].*

□

This model concerns multiple-phase systems with a given switching sequence, as discussed in the introduction. In addition, the total number of stages is fixed. Relaxing this assumption adds significant complexity but would be a valuable extension of this work. The following definition is loosely based on the hybrid system defined in Branicky, Bortar, and Mitter (1998).

DEFINITION 2.2. *A* multiple-phase switching system *is defined as* $\Xi=\{T,K,\vartheta\}$ *where,*

1. *T is a set of discrete controls known as* switching times *that dictate the termination of one phase and the start of the next,*
2. $K=\{k_1,k_2,...,k_n\}$ *is a finite, fixed, and exogenously determined sequence of discrete (integer) states that index individual continuous dynamical systems,* $\vartheta=\{\vartheta_k\}_{k\in K}$*, where* $\vartheta_k=[X,f_k,U]$*. The ordinal ranking of sequences is defined over the closed interval* $j=[1,2,...,n]$*,*
3. X *is a continuous state space where* $X \in R^m$*,*
4. f_k *is a vector field defining the dynamical law for each stage k, and*
5. U *is an open set of admissible controls lying in* R^v*.* □

It is assumed that the number of state and control variables and their relevant spaces are identical between stages for ease of exposition. However, these assumptions are not critical to the following derivation and consequently may be relaxed, if required.

The terminal time is defined as a control variable in the model for generality. However, the necessary conditions required for fixed endpoint problems are discussed specifically given their importance.

DEFINITION 2.3. *A* control input *for an embedded switching system* Ξ *consists of a set of vectors* $\chi_\Xi = \{t, u\}$ *where,*

1. $t = \{t_1, t_2, ..., t_{n-1}\}$ *is a sequence of real numbers denoting switching times, the moment* t_j *at which stage* k_j *is terminated and the stage* k_{j+1} *becomes active. It follows that regime* k_j *is active over the semi-open interval* $[t_{j-1}, t_j)$,
2. $t = t_n$ *is a freely determined terminal time, and*
3. $u = \{u_1, u_2, ..., u_n\}$ *is a collection of control functions defined for each stage in sequence K.*

□

It is possible for switching times to accummulate in this model. Consequently, not all regimes in the predefined sequence must be activated. For example, it may be optimal for two consecutive switching times, such as t_j and t_{j+1}, to coalesce (that is, $t_j = t_{j+1}$), in which case, movement from k_j to k_{j+2} will occur without the activation of k_{j+1}. This allows for the case where the operation of a stage or number of stages in sequence K is not contained within the optimal solution.

The state variable is continuous at the switching times in this model. However, jumps within the state variable (Vind 1967) are easily accommodated (see Bryson and Ho 1975, p. 106-108).

DEFINITION 2.4. *A* trajectory *(*Γ*) for a multiple-phase switching system* Ξ *and control sequence* χ_Ξ *is* admissable *over the interval* $t = [t_0, t_1, ..., t_{n-1}, t_n]$ *if it satisfies Definition 2.1 and the continuous dynamics* $\dot{x} = f_j(x(t), u_j(t))$, *for* $[t_{j-1}, t_j)$ *and* $j \in J$, *for a predefined switching sequence* $K = \{k_1, k_2, ..., k_n\}$. □

These definitions permit the classification of a general multiple-phase optimal control problem.

PROBLEM 2.1. *For a multiple-phase system* Ξ *identify an admissible trajectory that maximises the objective functional,*

$$J = e^{-rt_n} G(x(t_n), t_n) + \sum_{j=1}^{n-1} e^{-rt_j} C_j(x(t_j)) + \sum_{j=1}^{n} \left[\int_{t_{j-1}}^{t_j} \left[e^{-rt} F_j(x(t), u_j(t)) \right] dt \right], \tag{1}$$

subject to:

$$\dot{x} = f_j(x(t), u_j(t)), \textit{ for } [t_{j-1}, t_j) \textit{ and } j = [1,2,...,n] \textit{ given } K = \{k_1, k_2, ..., k_n\}, \tag{2}$$

$$\Lambda = \{t_0, x_0^i\} \textit{ fixed for } i = [1,2,...,c], x_n^i \textit{ fixed for } i = [1,2,...,d], \textit{ and} \tag{3}$$

$$\Theta_j = \{t_j, x(t_j)\} \textit{ free for } j = [1,2,...,n]. \tag{4}$$

where r is a discount rate, $G(x(t_n), t_n)$ *is a terminal reward function,* $C_j(x(t_j))$ *is a switching cost function for the jth phase, and* $F_j(x(t), u_j(t))$ *is a reward function on* $X^m \times U^v$ *for the jth phase. Functions G, C and F are all real-valued functions that are once continuously differentiable. The terminal value function is defined for* $x_n^i(t_n)$ *where i=[1,2,…,q]. Terminal state variables* $x_n^i(t_n)$ *are free in (4) for* $i = [d+1,...,m]$. □

The terminal reward function $G(x(t_n), t_n)$ is often defined as a salvage value function in economic applications of optimal control (Seierstad and Sydsaeter 1987). The switching cost function defines a sunken cost accruing to the termination of one stage and the initialisation of another. (These can be understood as terminal value functions for individual regimes.) They are a pertinent feature of many multiple-phase systems. For example, it can be costly to remove one crop and establish another (Mueller, Schilizzi, and Tran 1999) or invest in the productive capacity required for the artificial recovery of petroleum (Amit 1986). Both the terminal value function $G(\cdot)$ and the switching cost function $C(\cdot)$ are dependent on the state variable ($x(t_j)$). The former is a standard assumption in optimal control. The latter is included because such a relationship is likely to exist in a number of important multiple-phase problems. For example, the herbicide dose required for the establishment or removal of a crop may be dependent on weed density. Or, investing in a new production technology may require an initial outlay that is dependent on the current size of the existing firm.

THEOREM 2.1. *Consider a multiple-phase system* Ξ *described by Definitions 2.1-2.4. For* $j=[1,2,...,n]$ *and switching sequence* $K=\{k_1,k_2,...,k_n\}$*, let* $(x^*(t),u_j^*(t),t_j^*)$ *denote the admissible trajectory that maximises the value of J in Problem 2.1. This is the optimal trajectory* Γ^*.

Define a Hamiltonian function for each regime k_j *as,*

$$H_j(x(t),u_j(t),\lambda_j(t),t)=e^{-rt}F_j(x(t),u_j(t))+\lambda_j(t)f_j(x(t),u_j(t),t), \tag{5}$$

across the interval $[t_{j-1},t_j)$.

An optimal trajectory Γ^* *requires,*

i) *initial condition* $x_0^i=x(t_0)$ *for fixed initial state variable(s)* x_0^i *for i=[1,2…c],* (6a)

ii) *initial condition* $\lambda_1^T=0$ *for free initial state variable(s)* x_0^i *for i=[c+1,…,m],* (6b)

iii) *n m-dimensional vectors of real-valued, piecewise continuous adjoint functions* $\lambda_j(t)=\{\lambda_j^1(t),\lambda_j^2(t),...,\lambda_j^m(t)\}$*, defined across* $j=[1,2,...,n]$ *and piecewise continuously differentiable over the interval* $[t_{j-1},t_j)$*, that satisfy,*

$$\dot{\lambda}_j^T(t)=-\frac{\partial H_j(x(t),u_j(t),\lambda_j(t),t)}{\partial x(t)}, \tag{7}$$

where $\lambda_j^T(t)$ *denotes the transpose of the n adjoint vectors,*

iv) *control function(s) for each regime* k_j *that satisfy, for* $j=[1,2,...,n]$,

$$\frac{\partial H_j(x(t),u_j(t),\lambda_j(t),t)}{\partial u_j(t)}=0, \tag{8}$$

v) *for state variables free at the terminal time, an optimal trajectory requires an adjoint vector* $\lambda_n(t_n)$ *that satisfies, for* $x_n^i(t_n)$ *where i=[1,2,…,q],*

$$\lambda_n^T(t_n) = \frac{\partial e^{-rt_n} G(x(t_n), t_n)}{\partial x(t_n)}, \text{ and} \tag{9a}$$

a terminal adjoint vector $\lambda_n(t_n)$ *that satisfies, for* $x_n^i(t_n)$ *where i=[q+1,…,m],*

$$\lambda_n^T(t_n) = 0, \tag{9b}$$

NOTE: $x_n^i = x(t_n)$ *replaces (9a) or (9b) for* x^i *where i=[1,2,…d],*

vi) $$H_n(x(t), u_n(t), \lambda_n(t), t)\Big|_{t_n} + \frac{\partial e^{-rt_n} G(x(t_n), t_n)}{\partial t_n} = 0, \tag{10a}$$

NOTE: if no terminal value function is defined, then the equivalent of (10a) is,

$$H_n(x(t), u_n(t), \lambda_n(t), t)\Big|_{t_n} = 0, \tag{10b}$$

if, instead, the terminal time is fixed, then no additional necessary condition is required as $t = t_f$,

vii) *adjoint vectors that satisfy the boundary conditions,*

$$\lambda_j^T(t_j) + \frac{\partial e^{-rt_j} C_j(x(t_j))}{\partial x(t_j)} = \lambda_{j+1}^T(t_j), \tag{11}$$

at switching times $t = \{t_1, t_2, ..., t_{n-1}\}$ *and* $j = [1,2,..., n-1]$,

viii) $$H_j(x(t), u_j(t), \lambda_j(t), t)\Big|_{t_j} - \frac{\partial e^{-rt_j} C_j(x(t_j))}{\partial t_j} = H_{j+1}(x(t), u_{j+1}(t), \lambda_{j+1}(t), t)\Big|_{t_j}, \tag{12}$$

for those switching times in $t = \{t_1, t_2, ..., t_{n-1}\}$ *for which* $t_{j-1} < t_j < t_{j+1}$ *holds,*

ix) $$H_j(x(t), u_j(t), \lambda_j(t), t)\Big|_{t_j} - \frac{\partial e^{-rt_j} C_j(x(t_j))}{\partial t_j} \le H_{j+1}(x(t), u_{j+1}(t), \lambda_{j+1}(t), t)\Big|_{t_j}, \tag{13}$$

for those switching times in $t = \{t_1, t_2, ..., t_{n-1}\}$ *for which* $t_{j-1} = t_j < t_{j+1}$ *holds , and*

x) $$H_j(x(t),u_j(t),\lambda_j(t),t)\Big|_{t_j} - \frac{\partial e^{-rt_j} C_j(x(t_j))}{\partial t_j} \geq H_{j+1}(x(t),u_{j+1}(t),\lambda_{j+1}(t),t)\Big|_{t_j}, \quad (14)$$

for those switching times in $t = \{t_1, t_2, ..., t_{n-1}\}$ *for which* $t_{j-1} < t_j = t_{j+1}$ *holds.* □

PROOF. *See Appendix.*

Necessary conditions (6)-(10) are identical to those utilised in single phase problems incorporating a salvage value when the Hamiltonian function is differentiable with respect to (w.r.t) the control variable, $u_j^*(t)$, and this variable does not take a boundary value. This follows the definition of a multiple-phase problem as a set of *n* dynamical systems.

In contrast, switching conditions (11)-(14) are not found in standard optimal control problems. These describe how individual systems are linked over time under optimal management. These conditions appear in similar form in the models of Amit (1986), Mueller, Schilizzi, and Tran (1999), and Makris (2001). It is therefore established here that they generalise to a finite-time multiple-phase model with *n* regimes, positive switching costs, and different endpoint constraints. Equations (11) and (12) are also equivalent to the smooth pasting and value-matching conditions found in applications of stochastic control in finance (Brekke and Oksendal 1994; Dixit and Pindyck 1994).

Equation (11) determines the optimal level of the state variable(s) at each switching time. This condition states that it is optimal to switch when the marginal value of a change in the state variable is equivalent between stages. The shadow price variables, $\lambda_j^T(t_j)$ and $\lambda_{j+1}^T(t_j)$, represent the marginal adjustment in optimal value accruing to a change in the state variable, within the corresponding stage, at switching time t_j. The definition of state-dependent transition costs introduces an additional marginal effect for the active regime *j* in (11), a marginal cost term $[e^{-rt_j} C_j(x(t_j))]_{x(t_j)}$ (where $[\cdot]_x$ denotes the derivative of the term enclosed in square brackets with respect to x).

Switching conditions (12)-(14) describe the management of optimal switching times given the relative values of alternate stages. The value of the Hamiltonian function for a given regime at t_j represents the shadow price of altering the length of this phase by one instant. The term

$[e^{-rt_j} C_{s(j)}(x(t_j))]_{t_j}$ in (12) is the rate at which transition costs within regime *j* change over time. Equation (12) therefore states that it is optimal to switch to the subsequent regime at time t_j if the rate at which the capital value of each stage changes over time is equal at that point. Regime *j* should not be activated if its total value, reflected through its Hamiltonian and switching cost functions, is dominated at each potential switching time by that of the successive regime. This is described in (13). Moreover, the successive regime should not be adopted if there is no time t_j at which its capital value matches that earned within the active phase. This is stated in equation (14). Only one of conditions (12)-(14) will hold for a given stage *j*.

Necessary conditions (11)-(14) are not needed if T is empty. In this instance, Theorem 2.1 collapses to represent the necessary conditions required for the optimisation of a standard single-regime optimal control problem. The state variable(s) could be fixed for a given switching time t_j. In this instance, equation (11) is no longer required for determination of $x(t_j)$. Alternatively, the control input may consist of fixed switching times. Necessary conditions (12)-(14) are not required in this case.

The boundary conditions are obviously affected if switching cost functions $e^{-rt_j} C_j(x(t_j))$ and their relevant derivatives are not defined. If switching costs do not exist or are independent of the state vector, condition (11) requires equality between the adjoint variables of stages *j* and *j+1*. That is, $\lambda_j^T(t_j) = \lambda_{j+1}^T(t_j)$. Likewise, equation (12) simplifies to a requirement of equality between the total capital value of each regime at the switching time; that is, $H_j(\cdot)\big|_{t_j} = H_{j+1}(\cdot)\big|_{t_j}$; if switching costs are not defined or are independent of time. (Switching costs will not be independent of time in many economic problems because of discounting.) These conditions state rather unequivocally that it is optimal to switch when there is no benefit to remaining in the current phase. These results are analogous to the Weierstrass-Erdmann corner conditions (Seierstad and Sydsaeter 1987) from variational calculus, which are also required when state and/or control variables are subject to inequality constraints (Pontryagin et al. 1962). This equivalency highlights the close symmetry between multiple-phase problems with fixed versus free mode sequencing, if the latter is incorporated utilising piecewise constant controls. This symmetry outlines, rather intuitively, that it will also be optimal to switch modes in a free sequencing framework when there is no benefit to remaining in the current stage.

III. Algorithm

Theorem 2.1 may be used to study pedagogical multiple-phase systems using analytical methods. However, closed form solutions are notoriously difficult to obtain in all but the simplest control problems. The study of meaningful real-world systems consequently requires the construction of suitable numerical algorithms.

Shooting methods are iterative solution techniques for two-point boundary value problems (Keller 1968), which commonly arise in the context of solving the necessary conditions of an optimal control model. Shooting methods involve a series of steps; (a) guessing an unknown boundary value(s) of the state or adjoint variable(s), (b) integrating the state and adjoint equations either forward or backward (depending on whether the boundary value is an initial or terminal point) using an appropriate Initial Value Problem (IVP) method, (c) simultaneously calculating the optimal control utilising the optimality condition (see equation (8)), (d) updating the initial guess utilising a nonlinear equation method and (e) repeating the process until the boundary conditions are satisfied to sufficient accuracy (Ascher, Mattheij, and Russell 1995). This approach and Theorem 2.1 suggest an intuitive algorithm for the solution of multiple-phase problems.

ALGORITHM 3.1 REGIME-SHOOTING ALGORITHM

PURPOSE: Identify an optimal control sequence χ_{Ξ} for the multiple-phase system Ξ.

INITIALISATION: Determine a fixed stage sequence *S*. Define the maximum number of permissible iterations (*maxit*). Define the stopping tolerance ε. (Alternatively, different stopping tolerances may be defined for each component in Step 4.) Set the number of iterations *iter* to 1. Define a set of initial conditions $\Lambda = \{t_0, x_0^i\}$ for $i = [1,2,...,c]$. Provide initial guesses for the optimal switching times (t_j for $j = [1,2,...,n-1]$) and the state vector at these times ($x(t_j)$ for $j = [1,2,...,n-1]$).

STEP 1. Increase iteration count by 1. Optimise the first *n-1* stages independently of each other. Each phase is a fixed endpoint control problem, with the boundary conditions for these phases entirely determined by Λ and estimates of t_j and $x(t_j)$ from the previous iteration. (Note here that a switching time is a fixed terminal time for one phase but the initial time for another.)

Identify the value of the adjoint variable(s) and Hamiltonian at each boundary following the solution of each stage.

STEP 2. Optimise the terminal stage as a free-time control problem with t_{n-1} and $x(t_{n-1})$ as the initial time and state respectively. Following solution, identify the value of the adjoint variable(s) and Hamiltonian at t_{n-1}.

STEP 3. STOP and print output if $t_j^{iter} - t_j^{iter-1} < \varepsilon$ and $x(t_j^{iter}) - x(t_j^{iter-1}) < \varepsilon$ or $\lambda_j^T(t_j) + \dfrac{\partial e^{-rt_j} C_j(x(t_j))}{\partial x(t_j)} - \lambda_{j+1}^T(t_j) < \varepsilon$ and $H_j(\cdot)\big|_{t_j} - \dfrac{\partial e^{-rt_j} C_j(x(t_j))}{\partial t_j} - H_{j+1}(\cdot)\big|_{t_j} < \varepsilon$ for $j = [1,2,...,n-1]$. Otherwise, go to Step 4.

STEP 4. For $j = [1,2,...,n-1]$, use a non-linear equation method to identify improved estimates for $x(t_j)$ and t_j through the solution of $\lambda_j^T(t_j) + \dfrac{\partial e^{-rt_j} C_j(x(t_j))}{\partial x(t_j)} - \lambda_{j+1}^T(t_j) = 0$ and $H_j(\cdot)\big|_{t_j} - \dfrac{\partial e^{-rt_j} C_j(x(t_j))}{\partial t_j} - H_{j+1}(\cdot)\big|_{t_j} = 0$ respectively. Adjoint and Hamiltonian values are those calculated in Steps 1 and 2.

STEP 5. If *iter=maxit* then STOP and report progress; else go to Step 1.

The boundary conditions for each individual control problem in Step 1 are well-defined following the prior definition of the switching times and the state variables at these points. It is natural to question whether the designation of these fixed points will affect the optimality condition (8), as the variation δu in equation (A.11) (see Appendix) is no longer entirely arbitrary but must now satisfy these endpoint constraints. However, it may be shown that (8) holds despite this induced restriction (see Kamien and Schwartz 1991, Section II.6).

Solution of independent phases in the first and second steps may be achieved utilising Theorem 2.1 and gradient or shooting methods. Although more robust than shooting methods, the convergence of gradient algorithms is heavily dependent on the quality of the initial control history (i.e., the initial guess) (Bryson 1999). Standard shooting methods may be used but simultaneous integration of the state and adjoint equations may lead to significant instability in the optimal trajectories given sensitivity to initial conditions (Stoer and Bulirsch 1980).

Application of both methods is also complicated through the need to derive necessary conditions, which limits flexibility and may be problematic in models of even moderate complexity. (Although this last limitation is difficult to circumvent, a hybrid algorithm incorporating both methods may be the most efficient for the solution of simple problems.)

Multiple shooting aims to increase the robustness of standard shooting methods through division of the problem into multiple intervals that reduce the length of each integration (Keller 1968; Osborne 1969; Lipton et al. 1982). This method may be adapted to analyse multiple-phase problems (see Bulirsch and Chudej 1995 for a two-phase example). However, together with the need to correctly derive and code necessary conditions for each application, the computational burden is increased above that required for Algorithm 3.1, *ceteris paribus*, because the non-linear equation solver must also enforce the continuity of each state variable at the switching time (Pesch 1994).

An alternative method for Step 1 and 2 optimisation involves the discretisation of a control problem and solution through non-linear programming (NLP) (Goh and Teo 1988; Teo, Goh, and Wong 1991; Hull 1997). This direct transcription exploits the efficiency of modern NLP codes and has better covergence characteristics than other standard approaches (Betts 1999). It is also very flexible and suited to complex applications because it does not require the analytical derivation of necessary conditions. (However, their coding is recommended to increase the speed of many such algorithms, as this will avoid the computational expense associated with the evaluation of derivatives through the calculation of finite differences.) The efficiency of this method motivates its use for Step 1 and 2 optimisation in the following application. It is implemented using a variant of the MISER3.2 optimal control software (Teo, Goh, and Wong 1991), which is engineered to operate more efficiently in an iterative scheme.

A number of root-finding methods may be used in Step 4 to update the estimates of $x(t_j)$ and t_j (Stoer and Bulirsch 1980). Newton methods are difficult to implement in this instance because the first derivatives of the switching conditions are not readily available. Important alternatives are the secant and bisection methods that do not require these evaluations. The bisection method will converge at a linear rate, which is slower than both of the others. However, it is adopted in the following application because, unlike the secant or Newton methods, the search for a root is limited to an initial bracket. This will help to limit the search among plausible alternatives, provided that the interval is thoughtfully constructed and is wide enough for a meaningful search to be implemented.

Phases may be bypassed in this algorithm if equation (12) is satisfied for consecutive switching times at a single moment. (This corresponds to either equation (13) or (14) holding with equality.) However, this algorithm does not cater for the situation where (13) and (14) hold as inequalities because of the complexity that this introduces. This simplification is unlikely to introduce significant bias in meaningful multiple-phase systems. Nonetheless, these constraints may be incorporated utilising mathematical programming if the differential equations within each phase are explicitly solvable.

There are a number of ways to improve the efficiency of Algorithm 3.1. Firstly, solution time may be decreased through using optimal trajectories from the previous iteration as initial guesses for the next. For example, this method decreases solution time by over half in the first scenario of the application presented in Section IV. However, this process must be carefully implemented to ensure that a poor result from one iteration does not detrimentally affect the convergence of further runs. Secondly, parallel processing may be used to solve the independent subproblems in Steps 1 and 2. Optimal trajectories may be stored for each iteration if convergence or the sensitivity of these time paths to different endpoints is of interest. However, lastly, the implementation of the algorithm may be improved if this matrix is, instead, updated asynchronously. This is of particular relevance to large problems; for example those incorporating many control variables and a fine level of control discretisation, if direct transcription is used.

The following application is programmed in MATLAB version 7.0.4 (Miranda and Fackler 2002). The MISER3.2 software is implemented using the FMINCON function in the MATLAB Optimisation Toolbox version 3.0.2. A loose stopping criterion ($\varepsilon = .0001$) is utilised for the bisection method in the outer iteration so that numerical errors generated in the optimisation phase do not detrimentally affect convergence (Judd 1998).

IV. Application

Herbicides have been used extensively in Australian dryland agriculture since the advent of reduced tillage systems in the late 1970s (Pratley 2000). However, major crop weeds, such as annual ryegrass (*Lolium rigidum*) (Llewelyn and Powles 2001) and wild radish (*Raphanus raphanistrum*) (Walsh et al. 2001), have developed resistance to a number of herbicides since this time. Annual ryegrass, in fact, is now established as the world's most herbicide resistant cropping weed (Pannell et al. 2004) given resistance to multiple modes of herbicide action (Hall et al. 1994). For example, Llewellyn and Powles (2001) identified that nearly half of the ryegrass populations that they sampled in the Western Australian wheat belt exhibited resistance to diclofop-methyl herbicides, 64 percent were resistant to chlorsulfuron, and 28 percent were susceptible to both. Such levels of resistance incur significant costs on producers, both through the reduction of crop yield and higher control costs associated with a need to utilise alternative treatments (Powles and Bowran 2000).

Continuous cropping is commonplace in the Western Australian wheat belt given the higher profitability of cereals relative to livestock activities, the continued enhancement of reduced tillage technology, and the widespread adoption of crop legumes that permit crop phases of significant length (Pannell 1995; Poole et al. 2002). However, the inclusion of regular pasture phases has the potential to delay or help to minimise the effects of herbicide resistance through permitting the use of a wide range of weed control strategies (Powles et al. 1997). Examples of such strategies are grazing, competition from pasture plants, non-selective herbicides, winter-cleaning, spray-grazing, green and brown manuring, hay or silage making, and pasture topping through mechanical means. The economics of herbicide resistance and the utilisation of non-chemical treatments has been investigated previously (Gorddard et al. 1995, 1996; Pannell et al. 2004). Yet, the optimal management of multiple stages, phase length, and pasture treatments have not been studied given that significant methodological difficulties are predicted (see, for example, Gorddard et al. 1995, p. 73). These may be overcome, however, through the adoption of the framework presented in this paper.

It is assumed that a producer is interested in determining the optimal management of a single field in the eastern wheat belt of Western Australia. The aim of the producer is to maximise the value of the field between t_0 and the variable terminal time t_n. The initial seed population is fixed at $x_0 = \{x_0^s, x_0^h\}$. It is assumed that crop yield is detrimentally affected by the population of a single weed, annual ryegrass. Multiple weeds may be incorporated but only one weed is studied here because this permits the identification of important components while maintaining tractability. The densities of the crop and weed populations, together with the intensity of weed treatments, are assumed to be distributed uniformly. We may therefore focus on one hectare of the field in order to simplify the calculation of model quantities. Spatial variability is not treated, however, this issue is examined in Dorr and Pannell (1992).

The farmer is assumed to manage the field in a wheat crop - lucerne pasture rotation for five phases ($j = \{1,2,3,4,5\}$). Both the initial and terminal stages are cropping phases. The stage sequence is therefore crop – pasture – crop – pasture – crop. There are four switching times ($t = \{t_1, t_2, t_3, t_4\}$) and the planning horizon ends at t_5. The field is to be sold at the end of the last phase, with the value of the land asset dependent on the weed burden.

Establishment costs for wheat crops are incurred annually *within* a phase and are therefore not characterised here as switching costs. In contrast, lucerne requires establishment and removal at the beginning and end of each pasture phase respectively. Lucerne is expensive to establish because of high seeding and weed control costs. Its effective removal also requires careful grazing management and non-selective herbicides (Bee and Laslett 2002). Significant switching costs for lucerne establishment and renewal are therefore incorporated in the model. The switching cost representing lucerne establishment is $e^{-rt_j} C_{le}(x(t_j))$ for $j = \{1,3\}$ and that for lucerne removal is $e^{-rt_j} C_{lr}(x(t_j))$ for $j = \{2,4\}$. The switching cost may be a function of the state variable(s). For example, in some systems, the herbicide dose required for weed control at pasture establishment may be positively related to weed density (i.e., more herbicide is required when a high weed population is present). However, the switching cost is independent of weed density in this specification given that a high rate of non-selective herbicide is used, in accordance with standard practice. This independence of the state variable will remove the switching cost component from the equivalent of (11) for each switching time (i.e., the switching condition (11) becomes $\lambda_j^T(t_j) = \lambda_{j+1}^T(t_j)$ for $j = \{1,2,3,4\}$).

Two state variables are required to represent the weed seed population because of herbicide resistance (Gorddard et al. 1995, 1996). First, x^s is the population of annual ryegrass seeds that following germination are susceptible to the selective Group A (diclofop-methyl) herbicide (Preston 2000). Second, x^h is the population of seeds that following germination are resistant to this herbicide. In the following, these two state variables are referred to as *susceptible seeds* and *resistant seeds* respectively for ease of exposition. Resistance to only one herbicide is studied to focus attention on the intertemporal management of herbicide resistance.

Cereal phase dynamics

The objective functional for the cereal phases $j = \{1,3,5\}$ are,

$$J = \int_{t_{j-1}}^{t_j} e^{-rt}\left[py_0(1-\eta u^h(t))\left[(1-z)+z\left(\frac{b}{b+kW(t)}\right)\right] - c^h u^h(t) - \frac{\gamma u^n}{(1-u^n)}\right]dt\,, \qquad (15)$$

where r is a discount rate, p is a constant price, y_0 is weed-free yield, η is the proportion of yield lost to phytotoxic damage for a given dosage (kg/ha) of diclofop-methyl herbicide u^h, z is the maximum proportion of grain yield lost at high weed density, b is a crop-dependent density parameter, k is a constant representing the competitiveness between the weed population and the wheat crop, $W(t)$ represents the total weed population, c^h is the cost of herbicide, and γ is the cost of achieving 50 percent weed kill utilising alternative treatments u^n (Gorddard et al. 1995). The latter is a composite control variable representing a number of diverse treatments, such as cultivation, high seeding rates, burning, green manuring, and haymaking. The cost function for alternative treatments reflects an increasing marginal cost associated with their use. High levels of weed kill may only be achieved with cultural methods through green-manuring and the making of hay or silage, all of which require a significant cost as all crop revenues are sacrificed. In contrast, moderate mortality of seeds or weeds may be achieved through utilisation of comparatively inexpensive treatments. Examples are the burning of crop stubble and the use of seed catchers at harvesting time to decrease the return of weed seed to the field. The yield function in (15) is a constant crop density (100 plants m^{-2}) analogue of that found in Pannell et al. (2004).

The weed population is defined $W(t) = W^s(t) + W^h(t)$, where W^s is the susceptible weed population and W^h is the herbicide resistant weed population. These weed populations are related to the susceptible and resistant seed populations through $W^s = x^s g(1-M_w)e^{-yu^h}(1-u^n)$ and $W^h = x^h g(1-M_w)(1-u^n)$ respectively, where g represents the rate of germination, M_w represents the natural mortality of germinated weed seeds, and y is a parameter designating the strength of the relationship between ryegrass mortality and herbicide dose. Note that the alternative treatments kill both susceptible and resistant weeds, while herbicide obviously affects only the susceptible population. The last cereal phase ($j = 5$) also has a terminal value of $e^{-rt_5}e^{\alpha}(x_5^s)^{\beta}(x_5^h)^{\chi}$, where $\{\alpha, \beta, \chi\}$ are parameters.

The seed population declines with germination (g) and the natural mortality of ungerminated seeds (M_S). It increases with total seed production, which is simply the number of weeds setting seed multiplied by the mean seed production (R) of each individual plant.

The motion equation for the susceptible seed population is consequently,

$$\dot{x}^s = x^s\left(-v_1 + v_2 e^{-yU^h}(1-u^n)R\right), \qquad (16)$$

where $v_1 = g + (1-g)M_S$ and $v_2 = g(1-M_W)$.

The motion equation for the herbicide resistant seed population is,

$$\dot{x}^h = x^h\left(v_1 + v_2(1-u^n)R\right). \qquad (17)$$

Pasture phase dynamics

Revenue from the pasture enterprise will consist solely of animal production, given the poor economics of hay and silage making in this dryland environment. Even though the stocking density has important implications for ryegrass control, it is likely to be determined by farm-level factors that are not incorporated here. The pasture enterprise is therefore defined as a cost minimisation problem with exogenous definition of the interaction between grazing and weed density.

The selective herbicide may also be used during the pasture phase. Annual ryegrass may also be controlled through control of seed-set through spray-topping with a non-selective herbicide (Jones et al. 1984). This treatment is included here and is illustrative of the effective and affordable weed control treatments available during a pasture phase.

The producer's problem in the lucerne phases $j = \{2,4\}$ is,

$$\min J = e^{-rt}(c^h u^h + c^m u^m), \qquad (18)$$

subject to,

$$\dot{x}^s = x^s\left(-v_1 + v_2 e^{-yu^h} e^{-qu^m}(1-gra)R\right), \qquad (19)$$

$$\dot{x}^h = x^h\left(-v_1 + v_2 e^{-qu^m}(1-gra)R\right), \tag{20}$$

where c^m is the cost of a non-selective herbicide, u^m is the dosage (kg/ha) of this non-selective herbicide, q is a parameter designating the strength of the relationship between ryegrass mortality and this dose rate, and gra is the level of ryegrass control achieved through grazing.

Parameter values

This section describes the parameter values used for the numerical solution of the optimal control model. All monetary values are expressed in 2004 Australian dollars.

Weed-free yield (y_0) is 1.3 tonnes (Pannell et al. 2004). The relationship between herbicide dose and the proportion of yield lost to phytotoxic damage is taken from Gorddard et al. (1995, 1996). This relationship is $D = \eta u^h$ where $\eta = 0.1448$ (*s.e.*=0.026, $n = 96$, and R^2=0.36). Herbicide efficacy is described through $y = 7.451$ for the selective herbicide (Gorddard et al. 1995) and $q = 0.025$ for glyphosate (estimated from data in Wakelin, Lorraine-Colwill, and Preston 2004). The values of the other parameters within the yield function are $z = 0.6$, k=0.33, and b=105 (Pannell et al. 2004). The annual germination of each seed population (g) is 0.8 (or 80 percent) (Gill 1996a). Natural seed mortality is significant, particularly over summer and early autumn when seeds may perish from disease or die from dehydration if autumn rains are not consistent. The parameter representing total seed mortality (M_s) is 0.55. In contrast, the natural mortality of plants is low. The estimated rate (M_w) is 0.05. The average rate of seed production by an individual plant (R) is assumed to be 20,000 seeds (Gramshaw 1972). Parameters describing the population dynamics for each seed population are the same. This introduces little bias given that there is little evidence of differences in relative fitness between resistant and susceptible populations. Grazing effectiveness (gra) is set at 0.5, a conservative estimate that considers spray-topping effects and the rotational grazing of lucerne (C. Revell, personal communication).

The price received for a tonne of wheat is \$165 (DAWA 2004). The real discount rate is 5 percent (Pannell et al. 2004). The cost per kilogram of active ingredient (diclofop methyl) of herbicide is \$40. This follows from the cost of the Hoegrass® 375 herbicide, which is \$15 per litre, and its load of .375 kg of active ingredient per litre (DAWA 2004). Formulating an estimate of the cost of non-chemical control is problematic because this represents a number of weed control methods. Realistic estimates of the relationship between weed control and control costs are achieved when $\gamma = 10$. The cost of the glyphosate herbicide used for spray-topping is \$12.50 for a kilogram of active ingredient (DAWA 2004). Switching cost estimates are taken from DAWA (2004). That for lucerne establishment is $e^{-rt_j}C_{le}$, where $j = \{1,3\}$ and $C_{le} = \$58.50$. That for lucerne removal is $e^{-rt_j}C_{lr}$, where $j = \{2,4\}$ and $C_{lr} = \$21.25$.

No information is available that would permit the accurate estimation of a suitable terminal value function, although survey evidence confirms a relationship beteen land value and herbicide resistance (R. Llewellyn, personal communication). Inclusion of a terminal value function is equivalent to optimisation across an infinite time horizon because this function represents the expected long-term value of agricultural land under perfect information (McConnell 1983). It is consequently estimated through multiple runs of the cereal phase over a 100 year period[2] utilising different combinations of initial seed densities and a terminal point of 500 seeds per square metre to permit comparison. Regression of the resultant data using a Cochrane-Orcutt technique (Cochrane and Orcutt 1949) with AR(1) errors identified that $\alpha = 9.4295$ (s.e.=0.159), $\beta = 0.0514$ (s.e.=0.019), and $\chi = 0.232$ (s.e.=0.011) ($n = 50$, $R^2 = 0.958$).

Model output

The first scenario involves an initial susceptible seed population of 400 seeds m^{-2} and an initial herbicide resistant seed population of 100 seeds m^{-2}. The model solves after ten iterations and 32 minutes of solution time on a desktop computer incorporating a Pentium 4 2GHz processor and 1GB RAM. The optimal trajectories for both seed populations are shown in Figure 1. Here, the switching times are denoted with arrows and labels *tj*. For example, the first switching time is denoted *t1*.

[2] This is considered a sufficient proxy for an infinite horizon because technical progress over such a length of time seems likely to produce changes in farming systems that will overcome profit losses associated with resistance to Group A selective herbicides.

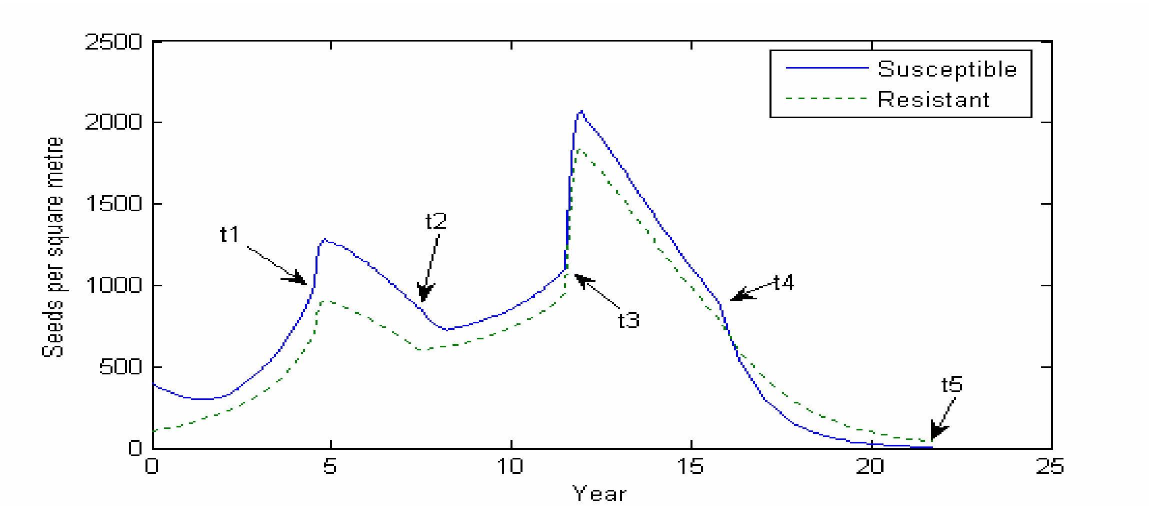

Figure 1. Optimal trajectories of susceptible and resistant annual ryegrass seeds over five phases of a wheat crop - lucerne rotation.

The susceptible seed population declines at the beginning of the first and third cropping phases (i.e., following *Year 0* and *t2* in Figure 1). The selective herbicide is applied at light rates at this time to decrease the susceptible population. However, most in-crop control arises from alternative treatments, which are used at 90 percent intensity over these phases. This heavy use is necessary because these treatments are effective against both populations, while the selective herbicide only controls the susceptible weeds. Heavy reliance on alternative treatments and their effectiveness against either population explains the similarity between the two trajectories in Figure 1.

The value of a pasture phase in the optimal rotation is clear because the seed population is lower after their use. For example, the seed populations are both lower at *t2* than at *t1*. The weed populations increase sharply at the beginning of each pasture phase. (This occurs at the moment immediately following *t1* and *t3* in Figure 2.) This leads to the increase in seed burdens observed at the corresponding times in Figure 1. The high effectiveness of spray-topping permits these increases to occur early in the pasture phase but not continue into the subsequent cereal crop. These high weed burdens do not detrimentally affect pasture profitability in this model. In reality, they may decrease animal production through promoting disease (e.g., ryegrass toxicity) (Pearce and Holmes 1976) or suppressing legume growth. However, annual ryegrass is also an important component of grazed pastures in many farming systems in Western Australia, particularly in early winter when the growth of legumes is reduced (Gill 1996b). Relating ryegrass density with animal production would therefore be an interesting extension of this work if information of sufficient quality were available.

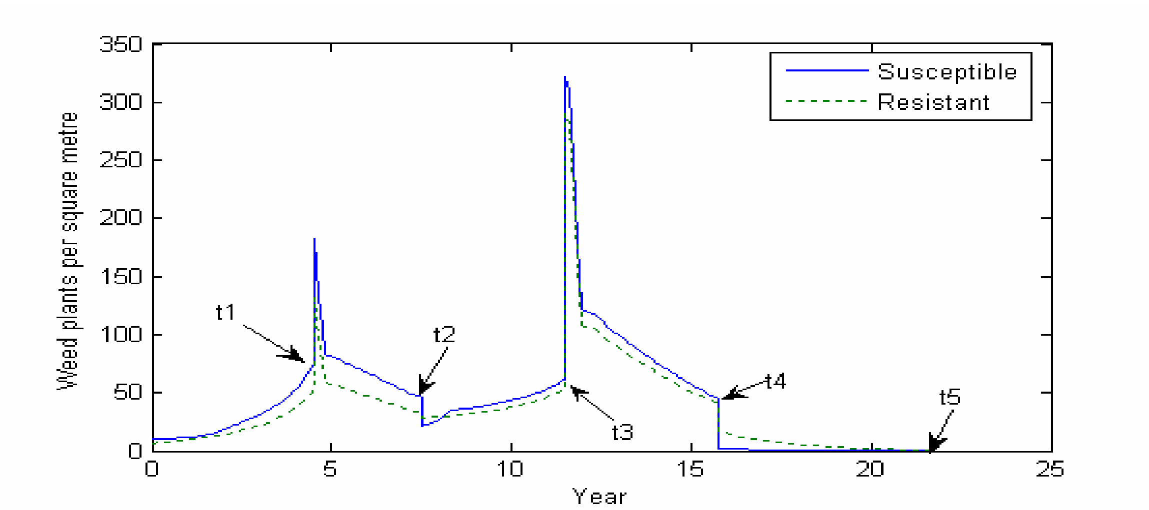

Figure 2. Optimal trajectories of susceptible and resistant annual ryegrass plants over five phases of a wheat crop - lucerne rotation.

The terminal cropping phase involves the use of alternative treatments at 97 percent intensity and continuous applications of selective herbicide at between 0.2 and 0.3 kilograms of active ingredient per hectare. This is the only time that the intensity of the selective herbicide is sustained across an entire phase in the first scenario. These practices result in a sharp decline in both the susceptible and resistant populations (see Figures 1 and 2). In addition, the heavy level of selective herbicide use causes the susceptible population to fall below that of the resistant population for the first time across the planning horizon. These changes in management reflect the impending sale of the farm and an associated additional benefit to weed control, a significant inverse relationship between land value and the terminal weed population.

The second scenario involves an initial susceptible seed population of 400 seeds m^{-2} and no herbicide resistance. This causes cropping income to increase by 67.5 percent over the planning horizon, reflecting the significant cost of resistance. This is higher than an estimate of 35 percent identified by Pannell and Zilberman (2001); however, their estimate includes pasture income, a planning horizon that is half of the length of that used here, and no terminal land value.

The optimal trajectories for the susceptible seed population for the “with resistance” and “without resistance” scenarios are shown in Figure 3. The switching times and terminal point for these scenarios are denoted *tj* and *hj* (for $j = \{1,2,3,4,5\}$) respectively. The optimal susceptible seed population is lower across most of the time horizon when no herbicide resistance is present (Figure 3). This is intuitive because the marginal value of selective herbicide application is higher, as resistant weeds do not need to be controlled simultaneously.

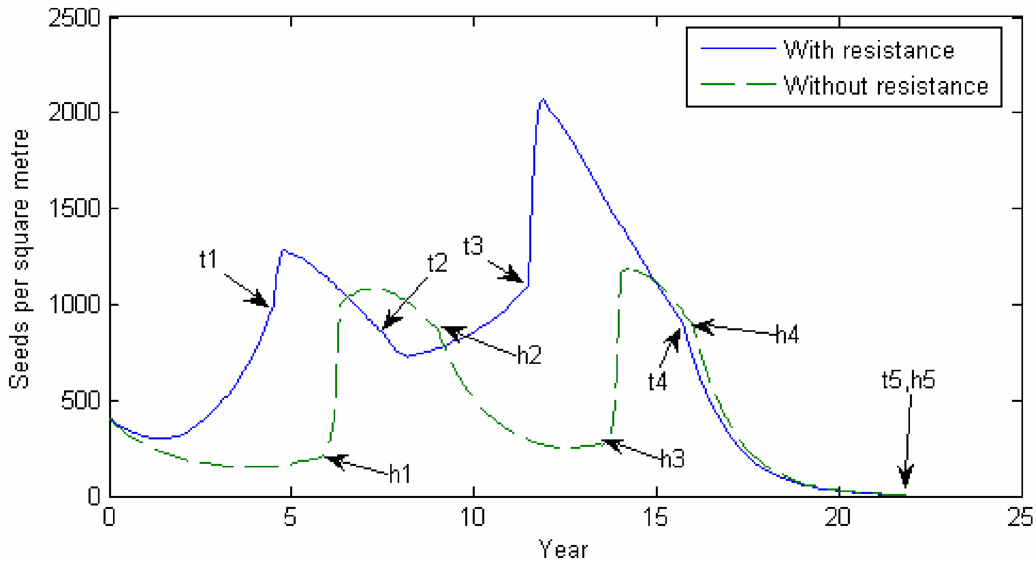

Figure 3. Optimal trajectories for the susceptible seed population with and without herbicide resistance over five phases of a wheat crop - lucerne rotation.

The utilisation of the selective herbicide permits the susceptible weed population to be effectively controlled during crop phases when there is no herbicide resistance. This is demonstrated in the decline of the seed population in Figure 3 between the initial time and *h1* and between *h2* and *h3*. This reflects the increased value of the selective herbicide, which is used intensively across each crop (see Figure 4(a)). (Figure 4 also clearly demonstrates the piecewise definition of control variables given the definition of multiple phases.) In comparison, the value of alternative treatments declines significantly with no herbicide resistance. This is reflected in their optimal intensity in the second scenario (see Figure 4(b)), relative to the 90 percent intensity utilised in the first. The value of pasture phases for weed control also declines significantly when there is no herbicide resistance. Both pasture phases begin later (*h1*>*t1* and *h3*>*t3*) and are also shorter given no herbicide resistance. This reflects that prolonged cropping phases are more profitable when selective herbicides are available to decrease in-crop competition.

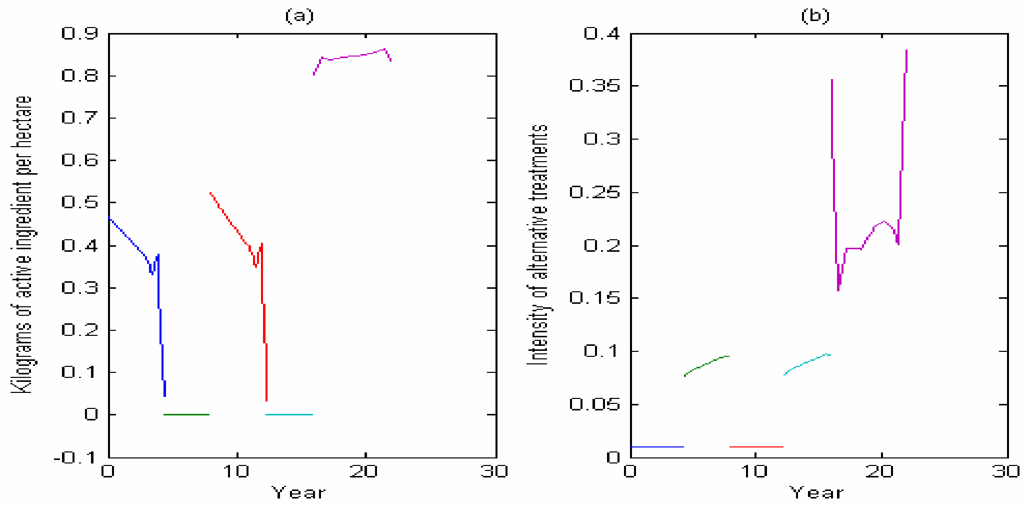

Figure 4. (a) Kilograms of active ingredient applied per hectare and the (b) intensity of alternative treatments under optimal management without herbicide resistance. (The alternative treatment in the second and fourth stages in (b) is the glyphosate spray-topping treatment, with its intensity measured in kilograms of active ingredient per hectare.)

The susceptible seed population follows a similar trajectory for both scenarios in the terminal phase (Figure 3), reinforcing the significant relationship between the future productivity of land and the terminal weed burden. However, the terminal phase is more profitable without herbicide resistance, as all weeds are susceptible to the efficient selective herbicide, which is applied heavily across the terminal phase (Figure 4(a)).

V. Conclusions

There appears to be no general framework for the numerical optimisation of multiple-phase systems in Economics. This is a significant limitation because these arise in many important situations, such as determining the optimal time to switch between production technologies, energy sources, or land uses. The optimisation algorithm presented in this paper offers a flexible and efficient platform for the solution of multiple-phase problems in which the number and sequence of phases is pre-assigned. Removing these latter restrictions is a key area for further work. To this end, we believe that direct transcription and the theory of piecewise constant controls has much to offer.

Appendix

This Appendix describes the derivation of Theorem 1. These conditions are derived utilising the approach used by Bryson et al. (1963) for the optimisation of a system incorporating inequality constraints and its extension by Bryson and Ho (1975) to deal with state variable discontinuities (pp. 106-108). The derivation of switching conditions for the Hamiltonian functions follows Amit (1986).

The problem is to identify those vectors that solve the following problem,

$$\max_{u_j, t_j} J = e^{-rt_n} G(x(t_n), t_n) + \sum_{j=1}^{n-1} e^{-rt_j} C_j(x(t_j)) + \sum_{j=1}^{n} \left[\int_{t_{j-1}}^{t_j} \left[e^{-rt} F_j(x(t), u_j(t)) \right] dt \right], \qquad \text{(A.1)}$$

subject to:

$$\dot{x} = f_j(x(t), u_j(t)), \; for \; [t_{j-1}, t_j) \; and \; j = [1,2,...,n] \; given \; K = \{k_1, k_2, ..., k_n\}, \qquad \text{(A.2)}$$

$$\Lambda = \{t_0, x_0^i\} \; fixed \, for \; i = [1,2,...,c], \; x^i \; fixed \, for \; i = [1,2,...,d], \; and \qquad \text{(A.3)}$$

$$\Theta_j = \{t_j, x^i(t_j)\} \; free \, for \; i = [d+1,...,m] \; and \; j = [1,2,...,n]. \qquad \text{(A.4)}$$

First, adjoin the n constraint(s) (A.2) to the objective functional (A.1) using n m-dimensional vectors of adjoint multipliers, $\lambda_j(t)$, to form the augmented functional, ϕ. This yields (A.5),

$$\phi = e^{-rt_n} G(x(t_n), t_n) - \sum_{j=1}^{n-1} e^{-rt_j} C_j(x(t_j)) + \sum_{j=1}^{n} \int_{t_{j-1}}^{t_j} [e^{-rt} F_j(x(t), u_j(t)) + \lambda_j^T(t)(f_j(x(t), u_j(t), t) - \dot{x}(t))] dt.$$

where $\lambda_j^T(t)$ denotes the transpose of the adjoint vector. Define a Hamiltonian function for each stage j,

$$H_j(x(t), u_j(t), \lambda_j(t), t) = e^{-rt} F_j(x(t), u_j(t)) + \lambda_j^T(t) f_j(x(t), u_j(t), t) \, . \qquad \text{(A.6)}$$

Substitute equation (A.6) into equation (A.5),

$$\phi = e^{-rt_n}G(x(t_n),t_n) - \sum_{j=1}^{n-1} e^{-rt_j}C_j(x(t_j)) + \sum_{j=1}^{n}\int_{t_{j-1}}^{t_j}[H_j(x(t),u_j(t),\lambda_j(t),t) - \lambda_j^T(t)\dot{x}(t)]dt. \quad \text{(A.7)}$$

Integrate the final term in the square brackets in equation (A.7) by parts,

$$\sum_{j=1}^{n} -\int_{t_{j-1}}^{t_j} \lambda_j^T(t)\dot{x}(t)dt = \sum_{j=1}^{n}\left(-\lambda_j^T(t_j)x(t_j) + \lambda_j^T(t_{j-1})x(t_{j-1}) + \int_{t_{j-1}}^{t_j} \dot{\lambda}_j^T(t)x(t)dt \right). \quad \text{(A.8)}$$

Substitute equation (A.8) into equation (A.7) to obtain,

$$\begin{aligned}\phi &= e^{-rt_n}G(x(t_n),t_n) - \sum_{j=1}^{n-1} e^{-rt_j}C_j(x(t_j)) + \sum_{j=1}^{n}\left(-\lambda_j^T(t_j)x(t_j) + \lambda_j^T(t_{j-1})x(t_{j-1})\right) + \\ &\sum_{j=1}^{n}\int_{t_{j-1}}^{t_j}[H_j(x(t),u_j(t),\lambda_j(t),t) + \dot{\lambda}_j^T(t)\dot{x}(t)]dt\end{aligned} \quad \text{(A.9)}$$

Consider *differential* changes in the discrete controls (the switching times and the endpoint of the terminal stage) and *infinitesimal variations* in the continuous controls ($u_j(t)$). These will produce equation (A.10), the first variation of equation (A.9),

$$\begin{aligned}d\phi &= \frac{\partial e^{-rt_n}G(x(t_n),t_n)}{\partial x(t_n)}dx(t_n) + \frac{\partial e^{-rt_n}G(x(t_n),t_n)}{\partial t_n}dt_n - \sum_{j=1}^{n-1}\left(\frac{\partial e^{-rt_j}C_j(x(t_j))}{\partial x(t_j)}dx(t_j) + \frac{\partial e^{-rt_j}C_j(x(t_j))}{\partial t_j}dt_j\right) \\ &+ \sum_{j=1}^{n}\left(H_j(\cdot)\Big|_{t_j}dt_j - H_j(\cdot)\Big|_{t=t_{j-1}}dt_{j-1}\right) + \sum_{j=1}^{n}\int_{t_{j-1}}^{t_j}\left\{\left(\frac{\partial H_j(\cdot)}{\partial x(t)} + \frac{\partial \dot{\lambda}_j^T(t)x(t)}{\partial x(t)}\right)\delta x + \frac{\partial H_j(\cdot)}{\partial u_j(t)}\delta u_j\right\}dt \\ &+ \sum_{j=1}^{n}\left(-\lambda_j^T(t_j)\delta x(t_j) + \lambda_j^T(t_{j-1})\delta x(t_{j-1}) - \lambda_j^T(t_j)\dot{x}(t_j)dt_j + \lambda_j^T(t_{j-1})\dot{x}(t_{j-1})dt_{j-1}\right)\end{aligned} ,$$

where $d\phi$, dt, and $dx(t_n)$ are differential changes in the performance index, time, and the state variable at the final moment, respectively, and δx and δu represent variations in the state and control trajectories respectively.

The expression $\sum_{j=1}^{n}\left(-\lambda_j^T(t_j)\delta x(t_j) + \lambda_j^T(t_{j-1})\delta x(t_{j-1}) - \lambda_j^T(t_j)\dot{x}(t_j)dt_j + \lambda_j^T(t_{j-1})\dot{x}(t_{j-1})dt_{j-1}\right)$ in (A.10) can be simplified to $\sum_{j=1}^{n}[-\lambda_j^T(t_j)(\delta x(t_j) + \dot{x}(t_j)dt_j) + \lambda_j^T(t_{j-1})(\delta x(t_{j-1}) + \dot{x}(t_{j-1})dt_{j-1})]$. This can be simplified further to $\sum_{j=1}^{n}[-\lambda_j^T(t_j)dx(t_j)+\lambda_j^T(t_{j-1})dx(t_{j-1})]$ using the identity $dx(t_j)\equiv\delta x(t_j)+\dot{x}(t_j)dt_j$ and its analogue for *j-1*.

$\sum_{j=1}^{n}[-\lambda_j^T(t_j)dx(t_j)+\lambda_j^T(t_{j-1})dx(t_{j-1})]$ is equivalent to the expression $\left[\lambda_1^T(t_0)dx(t_0)-\lambda_n^T(t_n)dx(t_n)+\sum_{j=1}^{n-1}\left(-\lambda_j^T(t_j)dx(t_j)+\lambda_{j+1}^T(t_j)dx(t_j)\right)\right]$. Similarly, the analogue of $\sum_{j=1}^{n}\left(H_j(\cdot)\big|_{t_j}dt_j-H_j(\cdot)\big|_{t=t_{j-1}}dt_{j-1}\right)$ is $H_n(\cdot)\big|_{t_n}dt_n+\sum_{j=0}^{n-1}\left(H_j(\cdot)\big|_{t_j}dt_j-H_{j+1}(\cdot)\big|_{t_j}dt_j\right)$. (Note the different ranges of the summation signs for the shadow price variables and the Hamiltonian functions.)

Substitution of these relationships into (A.10) and the collection of terms yields (A.11),

$$
\begin{aligned}
d\phi=\lambda_1^T(t_0)dx(t_0)&+\left(H_n(\cdot)\big|_{t_n}+\frac{\partial e^{-rt_n}G(x(t_n),t_n)}{\partial t_n}\right)dt_n+\sum_{j=0}^{n-1}\left(H_j(\cdot)\big|_{t_j}-H_{j+1}(\cdot)\big|_{t_j}-\frac{\partial e^{-rt_j}C_j(x(t_j))}{\partial t_j}\right)dt_j\\
&+\sum_{j=1}^{n-1}\left(-\lambda_j(t_j)+\lambda_{j+1}(t_j)-\frac{\partial e^{-rt_j}C_j(x(t_j))}{\partial x(t_j)}\right)dx(t_j)+\left(-\lambda_n(t_n)+\frac{\partial e^{-rt_n}G(x(t_n),t_n)}{\partial x(t_n)}\right)dx(t_n)\\
&+\sum_{j=1}^{n}\int_{t_{j-1}}^{t_j}\left\{\left(\frac{\partial H_j(\cdot)}{\partial x(t)}+\frac{\partial\dot{\lambda}_j(t)x(t)}{\partial x(t)}\right)\delta x+\frac{\partial H_j(\cdot)}{\partial u_j(t)}\delta u\right\}dt
\end{aligned}
$$

where $H_j(\cdot)=H_j(x(t),u_j(t),\lambda_j(t),t)$ for $j=\{1,2,...,n\}$.

The standard approach of deriving necessary conditions for a given optimal control problem using the variational approach requires the determination of variations in the state variable $\delta x(t)$ arising from perturbations in the control variable denoted by $\delta u(t)$. A more concise and intuitive method is utilised here. It is standard knowledge that the functional ϕ is extremal if it is stationary with respect to arbitrary perturbations. Therefore,

i) State variables x^i for $i=[1,2,...,c]$ are fixed. Admissible variations must satisfy $dx^i(t_0)=0$ in order for stationarity of ϕ to be guaranteed, so no conditions are required to identify the optimum initial state(s). State variables x^i for $i=[c+1,...,m]$ are free. Here, $dx^i(t_0)\neq 0$ so $\lambda_1^T(t_0)=0$ is needed for ϕ to remain stationary following perturbations in the initial state variable(s).

ii) If the terminal time is fixed, then $dt_n = 0$ so no condition is required to identify the optimum terminal time. If, instead, the terminal time is free, then stationarity of ϕ is only guaranteed if $H_n(x(t), u_n(t), \lambda_n(t), t)\big|_{t_n} + \frac{\partial e^{-rt_n} G(x(t_n), t_n)}{\partial t_n} = 0$ given $dt_n \neq 0$. This is modified accordingly for those state variables x^i for *i=[1,2,…,q]* for which a terminal value function is not defined.

iii) If the level of the state variable(s) are fixed at the switching times $t = [t_1, t_2, ..., t_{n-1}]$ then $dx(t_j) = 0$ for $j = [1,2,...,n-1]$ and no condition is required to identify them. Otherwise, stationarity of ϕ is only guaranteed if $\lambda_j^T(t_j) + \frac{\partial e^{-rt_j} C_j(x(t_j))}{\partial x(t_j)} - \lambda_{j+1}^T(t_j) = 0$ given $dx(t_j) \neq 0$.

iv) If the level of the state variable(s) are fixed at the terminal time then $dx(t_n) = 0$ and no condition is required to identify them. Otherwise, stationarity of ϕ is only guaranteed if $\lambda_n^T(t_n) - \frac{\partial e^{-rt_n} G(x(t_n), t_n)}{\partial x(t_n)} = 0$ given $dx(t_n) \neq 0$.

v) The adjoint functions are selected so $\frac{\partial H_{s(j)}(x(t), u_{s(j)}(t), \lambda_{s(j)}(t), t)}{\partial x(t)} + \dot{\lambda}_{s(j)}^T(t) = 0$ for $j = [1,2,...,n]$. This guarantees stationarity of ϕ in relation to arbitrary variations δx.

vi) Similarly, the control functions are selected so $\frac{\partial H_j(x(t), u_j(t), \lambda_j(t), t)}{\partial u_j(t)} = 0$ for $j = [1,2,...,n]$. This guarantees stationarity of ϕ in relation to arbitrary variations δu.

vii) If all of these necessary conditions are satisfied, then (A.11) becomes,

$$d\phi = \sum_{j=0}^{n-1} \left(H_j(\cdot)\big|_{t_j} - H_{j+1}(\cdot)\big|_{t_j} - \frac{\partial e^{-rt_j} C_j(x(t_j))}{\partial t_j} \right) dt_j . \qquad \text{(A.12)}$$

If the optimal solution includes $t_{j-1} < t_j < t_{j+1}$, then feasible modification dt_j is freely variable. Stationarity of (A.18) is only then guaranteed if,

$$H_j(\cdot)\big|_{t_j} - H_{j+1}(\cdot)\big|_{t_j} - \frac{\partial e^{-rt_j} C_j(x(t_j))}{\partial t_j} = 0 \text{, or, alternatively,} \qquad \text{(A.13)}$$

$$H_j(\cdot)\Big|_{t_j} - \frac{\partial e^{-rt_j} C_j(x(t_j))}{\partial t_j} = H_{j+1}(\cdot)\Big|_{t_j} \; . \tag{A.14}$$

If the optimal solution includes $t_{j-1} = t_j < t_{j+1}$, then feasible modification dt_j is instead non-negative. Stationarity of (A.18) is only then guaranteed if,

$$H_j(\cdot)\Big|_{t_j} - \frac{\partial e^{-rt_j} C_j(x(t_j))}{\partial t_j} \le H_{j+1}(\cdot)\Big|_{t_j} \; . \tag{A.15}$$

If the optimal solution includes $t_{j-1} < t_j = t_{j+1}$, then feasible modification dt_j is instead non-positive. Stationarity of (A.18) is only then guaranteed if,

$$H_j(\cdot)\Big|_{t_j} - \frac{\partial e^{-rt_j} C_j(x(t_j))}{\partial t_j} \ge H_{j+1}(\cdot)\Big|_{t_j} \; . \tag{A.16}$$

If, instead, the switching times are fixed, it follows that $dt_j = 0$ for $t = [t_1, t_2, ..., t_{n-1}]$ and no condition is required to identify them. □

References

Amit, R. "Petroleum Reservoir Exploitation: Switching from Primary to Secondary Recovery." *Operations Research* 34, July-August (1986): 534-49.

Arrow, K. J., and M. Kurz. *Public Investment, the Rate of Return, and Optimal Fiscal Policy*. Baltimore: John Hopkins University Press, 1970.

Ascher, U. M., R. M. M. Mattheij, and R. D. Russell. *Numerical Solution of Boundary Value Problems for Ordinary Differential Equations*. Philadelphia: SIAM, 1995.

Bee, G. A., and G. Laslett. "Development of a Rainfed Lucerne-Based Farming System in the Mediterranean Climatic Region of Southwestern Australia." *Agricultural Water Management* 53 (2002): 111-16.

Betts, J. T. "A Direct Approach to Solving Optimal Control Problems." *Computing in Science and Engineering* 1, no. 3 (1999): 73-75.

Branicky, M. , V. Borkar, and S. Mitter. "A Unified Framework for Hybrid Control: Model and Optimal Control Theory." *IEEE Transactions on Automatic Control* 43, no. 1 (1998): 31-45.

Brekke, K. A., and B. Oksendal. "Optimal Switching in an Economic Activity under Uncertainty." *SIAM Journal on Control and Optimisation* 32 (1994): 1021-36.

Bryson, A. H., W. Denham, and S. Dreyfus. "Optimal Programming Problems with Inequality Constraints I: Necessary Conditions for Extremal Solutions." *American Institute of Aeronautics and Astronautics Journal* 1 (1963): 2544-50.

Bryson, A. H., and Y. Ho. *Applied Optimal Control: Optimisation, Estimation, and Control*. New York: Hemisphere Publishing, 1975.

Bryson, A. H. *Dynamic Optimization*. New York: Addison-Wesley, 1999.

Bulirsch, R., and K. Chudej. "Combined Optimisation of Trajectory and Stage Separation of a Hypersonic Two-Stage Space Vehicle." *Zeitschrift fur Flugwissenschaften und Weltraumforschung* 19 (1995): 55-60.

Cochrane, D., and G. H. Orcutt. "Application of Least Squares Regressions to Relationships Containing Autocorrelated Error Terms." *Journal of American Statistical Association* 44 (1949): 32-61.

DAWA. *2004 Farm Budget Guide*. Perth: Western Australian Primary Industry Press, 2004.

Dixit, A. K., and R. S. Pindyck. *Investment under Uncertainty*. Princeton: Princeton University Press, 1994.

Dorfman, R. "An Economic Interpretation of Optimal Control Theory." *American Economic Review* 59 (1969): 817-31.

Dorr, G. J., and D. J. Pannell. "Economics of Improved Spatial Distribution of Herbicide for Weed Control and Crops." *Crop Protection* 11, no. 4 (1992): 385-91.

Gill, G. S. "Ecology of Annual Ryegrass." *Plant Protection Quarterly* 11, S(1) (1996a): 195-98.

———. "Why Annual Ryegrass Is a Problem in Australian Agriculture." *Plant Protection Quarterly* 11, S(1) (1996b): 193-94.

Goh, C. J., and K. L. Teo. "Control Parameterisation: A Unified Approach to Optimal Control Problems with General Constraints." *Automatica* 24 (1988): 3-18.

Gorddard, R. J., D. J. Pannell, and G. Hertzler. "An Optimal Control Model for Integrated Weed Management under Herbicide Resistance." *Australian Journal of Agricultural Economics* 39 (1995): 71-87.

———. "Economic Evaluation of Strategies for Management of Herbicide Resistance." *Agricultural Systems* 51 (1996): 281-89.

Gramshaw, D. "Germination of Annual Ryegrass Seeds (*Lolium rigidum*) as Influenced by Temperature, Light, Storage, Environment, and Age." *Australian Journal of Agricultural Research* 23 (1972): 779-87.

Hall, L. M., J. A. M. Holtum, and S. B. Powles. "Mechanisms Responsible for Cross Resistance and Multiple Resistance." In *Herbicide Resistance in Plants: Biology and Biochemistry*, edited by S. B. Powles and J. A. M. Holtum, 243-62. Boca Raton: Lewis Publishers, 1994.

Hull, D. G. "Conversion of Optimal Control Problems into Parameter Optimisation Problems." *Journal of Guidance, Control and Dynamics* 20, no. 1 (1997): 57-60.

Jones, S. M., W. M. Blowes, P. England, and P. K. Fraser. "Pasture Topping Using Roundup Herbicide." *Australian Weeds* 3, no. 4 (1984): 150-51.

Judd, K. L. *Numerical Methods in Economics*. Cambridge: MIT, 1998.

Kamien, M. I., and N. L. Schwartz. *Dynamic Optimisation: The Calculus of Variations and Optimal Control in Economics and Management*. Edited by C. J. Bliss and M. D. Intriligator, *Advanced Textbooks in Economics*. New York: Elsevier, 1991.

Keller, H. B. *Numerical Methods for Two-Point Boundary Value Problems*. London: Blaisdell, 1968.

Lipton, D., J. Poterba, J. Sachs, and L. Summers. "Multiple Shooting in Rational Expectations Models." *Econometrica* 50 (1982): 1329-34.

Llewellyn, R. S., and S. B. Powles. "High Levels of Herbicide Resistance in Rigid Ryegrass (*Lolium rigidum*) in the Wheatbelt of Western Australia." *Weed Technology* 15 (2001): 242-48.

Makris, M. "Necessary Conditions for Infinite-Horizon Discounted Two-Stage Optimal Control Problems." *Journal of Economic Dynamics and Control* 25 (2001): 1935-50.

McConnell, K. E. "An Economic Model of Soil Conservation." *American Journal of Agricultural Economics* 65 (1983): 83-89.

Miranda, M. J., and P. L. Fackler. *Applied Computational Economics and Finance*. Cambridge: MIT, 2002.

Mueller, U., S. Schilizzi, and T. Tran. "The Dynamics of Phase Farming in Dryland Salinity Abatement." *Australian Journal of Agricultural and Resource Economics* 43, no. 1 (1999): 51-67.

Osborne, M. R. "On Shooting Methods for Boundary Value Problems." *Journal of Mathematical Analysis and Applications* 27 (1969): 417-33.

Pannell, D. J. "Economic Aspects of Legume Management and Legume Research in Dryland Farming Systems of Southern Australia." *Agricultural Systems* 49 (1995): 217-36.

Pannell, D. J., and D. Zilberman. "Economic and Sociological Factors Affecting Growers' Decision-Making on Herbicide Resistance." In *Herbicide Resistance and World Grains*, edited by D. L. Shaner and S. B. Powles, 251-77. Boca Raton: CRC Press, 2001.

Pannell, D. J., V. Stewart, A. Bennett, M. Monjardino, C. Schmidt, and S. B. Powles. "RIM: A Bioeconomic Model for Integrated Weed Management of *Lolium rigidum* in Western Australia." *Agricultural Systems* 79 (2004): 305-25.

Papadimitriou, C., and K. Steiglitz. *Combinatorial Optimisation: Algorithms and Complexity*. New Jersey: Prentice-Hall, 1982.

Pearce, G. A., and J. E. Holmes. "The Control of Annual Ryegrass." *Journal of Agriculture Western Australia* 17 (1976): 77-81.

Pesch, H. J. "A Practical Guide to the Solution of Real-Life Optimal Control Problems." *Control and Cybernetics* 23 (1994): 7-60.

Pontryagin, L. , V. Boltyanskii, R. Gamkrelidze, and E. Mishchenko. *The Mathematical Theory of Mathematical Processes*. New York: Wiley, 1962.

Poole, M. L., N. C. Turner, and J. M. Young. "Sustainable Cropping Systems for High Rainfall Areas of South-Western Australia." *Agricultural Water Management* 53 (2002): 201-12.

Powles, S. B., C. Preston, I. B. Bryan, and A. R. Jutsum. "Herbicide Resistance: Impact and Management." *Advances in Agronomy* 58 (1997): 57-93.

Powles, S. B., and D. G. Bowran. "Crop Weed Management Systems." In *Australian Weed Management Systems*, edited by B. M. Sindel, 287-306. Melbourne: R. G. and F. J. Richardson, 2000.

Pratley, J. E. "Tillage and Other Physical Methods." In *Australian Weed Management Systems*, edited by B. M. Sindel, 105-23. Melbourne: R. G. and F. J. Richardson, 2000.

Preston, C. "Herbicide Mode of Action and Herbicide Resistance." In *Australian Weed Management Systems*, edited by B. M. Sindel, 209-25. Melbourne: R. G. and F. J. Richardson, 2000.

Seierstad, A., and K. Sydsaeter. *Optimal Control Theory with Economic Applications*. Amsterdam: North Holland, 1987.

Stoer, J., and R. Bulirsch. *Introduction to Numerical Analysis*. New York: Springer, 1980.

Teo, K. L., C. J. Goh, and K. H. Wong. *A Unified Computational Approach for Optimal Control Problems*. New York: Longman Scientific and Technical, 1991.

Tomiyama, K. "Two-Stage Optimal Control Problems and Optimality Conditions." *Journal of Economic Dynamics and Control* 9, no. 3 (1985): 317-87.

Tomiyama, K., and R. Rossana. "Two-Stage Optimal Control Problems with an Explicit Switch Point Dependence." *Journal of Economic Dynamics and Control* 13, no. 3 (1989): 319-37.

Vind, K. "Control Systems with Jumps in the State Variables." *Econometrica* 35 (1967): 273-77.

Wakelin, A. M., D. F. Lorraine-Colwill, and C. Preston. "Glyphosate Resistance in Four Different Populations of *Lolium rigidum* is Associated with Reduced Translocation of Glyphosate to Meristematic Zones." *Weed Research* 44 (2004): 453-59.

Walsh, M. J., R. D. Duane, and S. B. Powles. "High Frequency of Chlorsulfuron Resistant Wild Radish (*Raphanus Raphanistrum* L.) Populations across the Western Australian Wheatbelt." *Weed Technology* 15 (2001): 199-203.