

The spread of a biological invasion in space and time: Modelling active and passive surveillance

Hester, S.M.¹ and O.J. Cacho¹

¹ *School of Business, Economics and Public Policy, University of New England, Armidale, New South Wales*
Email: shester@une.edu.au

Abstract: Invasive species are an important threat to global biodiversity and cause considerable economic losses. Modelling the spread of invaders can assist in mitigating the impacts of biological invasions by allowing us to identify strategies that are most likely to be effective in slowing or reversing their spread. In many situations, the main constraint to controlling or eradicating invaders is finding them rather than eliminating them after they are located. Once an invasion is found it can be treated and killed with a high probability of success. Searching large areas actively is expensive and therefore enlisting the help of the public through ‘passive surveillance’ is increasingly being used by pest-management agencies.

The roles of active and passive surveillance and their interaction are investigated here using a spatially-explicit simulation model of the spread of an invasive species. The landscape is represented as a raster map consisting of square cells. Each cell in the landscape is characterised by various attributes, including habitat suitability and ownership type (private or public). The probability that a given site will be invaded depends on both habitat suitability and the number of propagules landing on it. Dispersal of propagules across the landscape is assumed to follow a Cauchy kernel. Long-distance dispersal may also occur independently, such as when propagules are transported by road or water. An invasion may be detected as a result of a report from the public or through active searching by a pest-control agency. Over time, the pest control agency uses passive detections, repeat searches and information about cell attributes to undertake additional searches in an attempt to eradicate the invader.

The model is applied to a hypothetical invasion. Measures of success, such as cost and probability of eradication, are incorporated as fitness measures within an evolutionary algorithm that identifies optimal search and control strategies. Strategies are defined in terms of search effort applied per cell, the size of the neighbouring radius that is searched when an infestation is discovered, and the number of repeat visits to previously treated sites. Results demonstrate that increases in passive detection can reduce eradication costs and increased the probability of eradication.

Although it is impossible to ensure that the global optimum is identified for a given scenario, the evolutionary algorithm helps identify quasi-optimal solutions that may be difficult to find through trial and error.

Keywords: *Passive detection, Pest eradication, Spatially-explicit simulation model, Dispersal, Evolutionary algorithm*

1. INTRODUCTION

Invasive species are recognised as an important threat to global biodiversity (Vitousek et al., 1996) and are responsible for large economic losses (Liebman et al., 2001, Liebhold et al., 1995). Modelling the spread of invaders to assist in mitigating the ecological and commercial impacts of biological invasions allows us to make best estimates of which strategies are likely to be most effective in slowing or reversing the spread of invaders. In most situations, outside of agriculture, the main constraint to eliminating invaders is not killing them but finding them. Once an invasion is found it can be treated and killed with often high probability.

The probability of finding an organism through active search by a pest-control agency in a target area is affected by the detectability of the organism and the amount of search effort applied per unit area. With a limited budget, application of more search effort per site results in less sites able to be searched, and hence a higher risk that the invasion will spread outside the containment area. There is a clear tradeoff between intensity of search and the extent of area searched. Understanding this tradeoff can lead to improved chances of controlling or eradicating invasions.

As an option to avoid spending large amounts of money searching for invaded sites over a large area, pest management agencies are now using ‘passive surveillance’ - reports from members of the public of encounters with pests - to assist in surveillance and control. The Ministry of Agriculture and Forestry Biosecurity New Zealand (MAFBNZ) describes passive surveillance as ‘[s]urveillance that relies on members of the public, industry groups, plant or animal health professionals and/or laboratories reporting suspected cases of plant or animal disease or the presence of a pest at their discretion.’ (MAFBNZ 2008, p. 29). Passive detections are often the method by which an invader is first recognised in a country or region. For example, the initial detection of the European wasp (*Vespula germanica*) in Western Australia occurred following a private submission of a wasp for identification (Davis and Wilson, 1991). This initial report subsequently led to the discovery of five nests and an eradication campaign. Another example is the initial discovery of the red imported fire ant (*Solenopsis invicta*) in Australia, following submission of two separate samples of the ant by members of the public. The initial reports led to the discovery of two epicentres of infestation and the establishment of the Red Imported Fire Ant Eradication Program (Jennings, 2004). The recent passive detection of Khapra beetle in Perth and mango leaf gall midge on Horn Island in Queensland are two more examples of the important role the community can play in the early detection of invasive species (Beale et al., 2008).

In addition to its value in the initial detection of a pest, passive surveillance during eradication programs has also proven to be very important: records from the campaign to eradicate the European wasp from Western Australia show that the public are responsible for finding 90% of the infestations in new areas (Davis and Wilson, 1991); reports of fire ants by the public have resulted in detections of half the outlying populations of the ant (Jennings, 2004); and the Four Tropical Weeds Eradication Program (4TWP) in Queensland obtained information for more than a quarter of the locations of weeds in the program due to detection by members of the public (Brooks and Galway, 2008).

Given the difficulty and expense of finding organisms through active search, and the potential importance of passive detections, an understanding of both types of surveillance can help develop useful decision tools. Cacho et al. (2006, 2007) adapted search theory (a technique used in search and rescue and military operations) to the control of biological invasions. Their model assumed a homogeneous search area where population density could increase, but did not consider spatial spread. In this paper we extend their model by applying the search algorithm to multiple sites within a landscape and representing spread through space as well as time. In addition, we incorporate passive surveillance. We apply the model to a hypothetical invasion and determine the probability and potential costs of eradication, for various active and passive surveillance strategies. A genetic algorithm (GA) is then applied to find better search strategies in terms of search effort, search radius and repeat visits for given rates of passive detection. Results provide useful insights into the design of efficient pest-control programs.

2. THE MODEL

The model applied in this paper is implemented in Matlab (The Mathworks, 2002). The model is described in detail by Cacho et al. (in review). Only a brief overview of the model is presented here. The landscape is represented as a matrix of dimensions $n_r \times n_c$ whose elements are square cells of 1 ha each. Cells have four attributes, each contained in a different map layer:

- Habitat suitability (α_i): the probability that a propagule landing on cell i will become established, taking a value of 0 (will not become established) or 1 (will become established).

- Detectability (λ_i), the effective sweep width of cell i , measured in distance (m) from the search path (explained below).
- Search speed (s_i), the speed (m/h) at which cell i can be traversed following standard search procedures.
- Ownership type (o_i), a binary variable indicating whether cell i is privately owned ($o_i = 1$) or publicly owned ($o_i = 0$).

For convenience all maps are represented as column vectors of dimension $n \times 1$, (where $n = n_r \times n_c$). This speeds up execution and simplifies coding of numerical operations. These vectors can be easily mapped back to the original matrix (their elements are arranged vertically down the rows and then across the columns of the matrix).

The state of a cell i is given by its invasion status, represented by binary variable x_i (1=presence, 0=absence) and contained in vector \mathbf{x} (the state vector). An invaded cell produces w propagules per time period, and these propagules spread to neighbouring cells. The distance between cells (d_{ij}) determines the proportion of propagules from cell i that reach cell j according to a dispersal kernel. The dispersal kernel is used to create an adjacency matrix (\mathbf{A}) of dimensions $n \times n$, whose element A_{ij} represents the expected proportion of propagules dispersing from cell i to cell j . Element A_{ij} is calculated based on the distance between cell i and cell j , applying a Cauchy kernel. Long distance dispersal can occur with probability p_L independently of the dispersal kernel, as may occur when propagules are transported by road, water, or other means.

Stochastic simulation of dispersal through time is executed by the matrix multiplication

$$\mathbf{y} = (\mathbf{x}' w) \mathbf{A} \tag{1}$$

where \mathbf{y} represents the number of propagules landing in each cell. The habitat suitability (α_i) of each cell is then used to determine the probability that the site will become infested and new infestations are selected through random sampling.

An invaded cell can be detected through passive surveillance with probability (p_p) or through active search with probability (p_a). In passive surveillance the public detects an invader and reports it to the relevant agency. The probability of passive detection depends on the ownership attributes (o_i) of a cell.

The agency in charge of managing the invasion invests search effort (M) in the following order of priority: (i) by searching sites where treatment has occurred in the recent past (repeat search); (ii) by searching sites in response to reports from the public (follow-up search); and (iii) through independent searching in public land not previously searched during (i) or (ii) (active search). The probability that an invasion will be detected (p_{ai}) depends on the search effort applied:

$$p_{ai} = 1 - \exp\left(-\left(\frac{s_i \lambda_i m_i}{a}\right)\right) \tag{2}$$

The expression within the inner brackets represents coverage, defined as the ratio of the area actually searched over the total area of the cell; a is the cell area (10 000 m²), s_i and λ_i are the map attributes defined earlier, and m_i is the search effort applied (hr) in cell i . The effective sweep (λ_i) is a measure of the detectability of the target and is affected by target characteristics, environmental conditions and the capability of the searcher (see Cacho et al., 2006; 2007 for details). Equation (2) represents the probability of detection if the invader is present, as a function of search effort (Figure 1). Any invasions detected, whether by active or passive surveillance, are treated and killed with probability p_k .

We assume that, to encourage passive detections, the public is offered a bounty payment (C_B) for each detection reported to the relevant agency. The total cost of any given scenario therefore depends on the number of passive reports, the amount of search undertaken by the agency and the cost of treatment. Total cost (C) is calculated in terms of

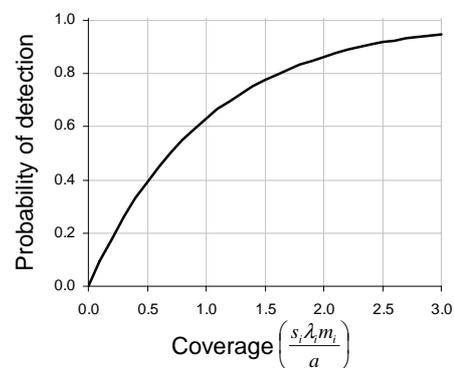


Figure 1. Probability of detection as a function of coverage, given in equation (2).

present value as:

$$C = \sum_t \left[(N_{pt} C_B + N_{at} m C_m + N_{Tt} C_T) + \sum_{\tau=t-S_R}^{t-1} N_{a\tau} C_m \right] (1 + \beta)^{-t} \tag{3}$$

where N_{pt} is the simulated number of cells where a passive find is reported in year t , N_{at} is the simulated number of cells where search takes place, N_{Tt} is the number of cells treated as a result of the three types of search, and C_B , C_m and C_T are the bounty payment (\$/report), the cost of searching (\$/ha) and the cost of treatment (\$/ha), respectively. S_R is the number of years repeat searching should occur and β is the discount rate. The second summation term in (3) represents the cost of repeat searches to ensure that cells treated in previous years have not been reinvaded.

If a passive detection occurs in a private parcel the invasion is destroyed, but not necessarily reported. If the passive detection occurs in a public parcel it is destroyed only when reported. This covers the situation where a person may eliminate a pest from their backyard but they may not inform the authorities and therefore no follow-up searches would occur in adjacent areas. The probability that an invasion will be reported is given by p_B .

We assumed that active search by the pest-control agency occurs only in public land unless a detection is made. In this case, an intensive search is conducted in all parcels within a specific radius (r_m) of the detection site, regardless of ownership type.

3. THE SIMULATION

An arbitrary world of dimension $n=16,641$ (hectares) ($n_r=n_c=129$) was created using a mid-point displacement fractal algorithm (Saupe, 1988); in this world 0.7 of the area was under private ownership and the remaining 0.3 was under public ownership. A base model run (*Base*) using parameter values that reflect the spread of a hypothetical but plausible insect pest (Table 1) was compared with an optimal solution generated using a genetic algorithm (*Base GA*). In both simulations the probability of passive detection on private land (p_{p1}) was set at 0.3. In a subsequent run, (*Passive GA*), p_{p1} was increased to assess the effects on cost and time to eradication of an increase in passive surveillance. All simulations consisted of 100 Monte Carlo iterations with a planning horizon (T) of 15 years.

Eradication was defined as absence of invaded sites by the end of the simulation.

Let

$$X_t = \sum_i x_{it} \tag{4}$$

represent the total area invaded at time t . X_t is used as a measure of performance and to calculate eradication probabilities. In the set of simulations reported below an initial invasion was generated randomly and used as the initial state for all simulations. In each simulation the total effort available for active search was set at 6 554 hours, enough to apply about 1.3 hours of search per public cell. This is in addition to the time required for repeat searching and following up on passive detections. Additional parameter values used in the simulations are presented in Table 1.

3.1. The Genetic Algorithm

A binary-string genetic algorithm was applied to identify optimal control strategies.

Table 1. Parameter values used in the base case simulation

Parameter	Value	Description
w	100	Propagules produced by invaded cells (no.)
p_{p1}	0.3	Probability of passive detection, (private)
p_{p0}	0.1	Probability of passive detection, (public)
p_L	0.02	Probability of long-distance jump
M	6554	Total effort available (hr)
α_i	0.02	Habitat suitability
λ_i	5	Effective sweep width (m)
s_i	1,000	Search speed (m/hr)
m_i	2	Minimum search effort per cell (hr)
t_D	5	Time period when invasion is discovered
γ	3.95	Dispersal kernel parameter
p_k	1	Probability of killing treated invasions
p_B	1	Proportion of passive detections reported
r_m	5	Search radius for reported sites (no. cells)
S_R	3	Number of repeat searches
C_B	500	Cost of bounty (\$ per find)
C_m	30	Cost of search (\$/hr)
C_T	100	Cost of treatment (\$/ha)
β	0.06	Discount rate
a	10,000	Cell area (m ²)
T	15	Planning horizon (y)

This type of evolutionary algorithm has been applied to a wide range of problems in the area of agriculture and natural resource management (Mayer, 2002) including to the management of invasive species (Taylor and Hastings, 2004). The aim of our GA is to identify optimal search and control strategies that minimise the cost of management the pest while maximising the likelihood of eradication. The chromosomes are contained in vectors consisting of three parameters: search effort applied per cell (m), the size of the neighbouring radius that is searched when an infestation is discovered (r_m) and the number of repeat visits to previously treated sites (S_R). The last two variables are integers and therefore the binary-string GA is well suited to the problem.

Table 2. Values of GA parameters

Description	Value
Population size	20
Number of generations	50
Lower, upper bound of m	0, 10
Lower, upper bound of r_m	0, 7
Lower, upper bound of S_R	0, 3
Mutation	0.6
Crossover	0.6
Number of new individuals	20

Each individual in the GA population has a measure of fitness that incorporates both the total cost (C from equation 3) of a particular strategy and the invasion size measured as the area invaded at the end of the time period (X_{15}), measured in years. The fitness function (F) to be maximised is:

$$F = \left(\frac{\bar{X}_1 - \bar{X}_{15}}{C} \right) \tag{5}$$

Pairs of individuals were selected for reproduction using roulette-wheel sampling (Goldberg, 1989). Additional information about the GA is given in Table 2. The base-case values reported in Table 1 were used to start the optimisation run. Although it is impossible to ensure that the global optimum is identified for a given scenario, the evolutionary algorithm helps identify quasi-optimal solutions that may be difficult to find through trial and error.

4. RESULTS AND DISCUSSION

Comparing the cumulative distribution functions (CDF) of final area invaded for the *Base* and *Base GA* simulations provides useful insights for planning purposes, as well as evidence that the GA produces a considerable improvement over the base simulation (Figure 2A). The initial invasion size in the simulation was 206 ha. The base case resulted in partial control of the invasion, as the probability of final area invaded being less than 206 ha was virtually 1.0. However this strategy did not result in eradication, as the *Base* curve intersects the horizontal axis at a positive value (19 ha) implying that the probability that 0 ha will be invaded by year 15 is zero (Figure 2A). The optimal solution for *Base GA* resulted in much better control, and a 0.7 probability that the invasion would be eradicated by year 15, and this was achieved at a lower cost than in the base case (Figure 2B). The entire CDF of cost for the GA solution is to the left of the CDF of cost for the base case (Figure 2B). These results indicate that the GA strategy dominates the base strategy by a considerable margin, in terms of both costs and probability of eradication.

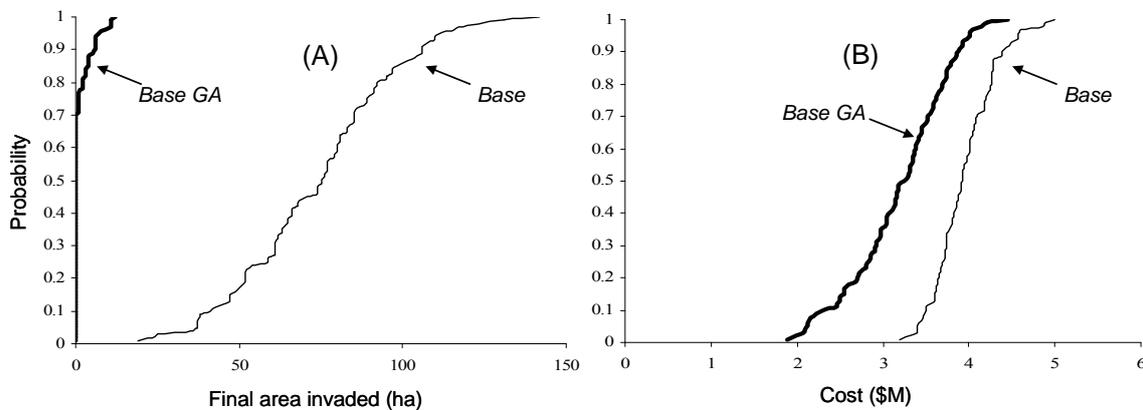


Figure 2. Cumulative distribution functions of final area invaded (A) and Cost (B) showing the base case solution (*Base*) and the optimal solution for the base genetic algorithm (*Base GA*).

Results in Table 3 indicate that the main difference between the base simulation and the GA solution is in the intensity of search per cell (m). In the base case m is 2 hours per cell, whereas in the base GA, m is 8.06 hours

per cell. Given the size of the public area to be searched and the total amount of active search effort available (M), the base strategy would allow about 66% of public land to be searched, whereas the GA strategy would allow only 16% of public land to be searched. Applying the parameter values from Table 1 to equation (2) we find that the base strategy results in only 0.63 probability of detection, whereas the GA strategy results in a 0.98 probability of detection. Active search occurs in public lands, where the probability of passive detection is low (0.1) and the results indicate that, if a site is to be searched, it is important to ensure a high probability of detection.

Other results in Table 3 indicate that the increased search effort per cell should be accompanied by a reduction in the search radius around detections (r_m) from 5 to 4 cells, and repeat visits to previously treated cells (S_R) should occur for only one year rather than three. Using the *GA Base* strategy the cost of eradication is \$3.19 million, compared to \$3.95 million, nearly a 20% reduction. In both cases most of the cost is attributed to searching for the pest. The median time to eradication is 13 years under *Base GA* while it never occurs within the 15-year time frame under *Base*.

When p_{p1} is increased to 0.8 (*GA Passive*) optimal values of m , r_m and S_R fall compared to those under *Base GA*. The optimal amount of time spent searching (5.48 hr/cell) is still higher than under the base case, and would allow 24% of public land to be searched with a probability of detection of 0.94. An interesting effect of increasing the probability of passive detection in private land to 0.8 in *GA Passive*, is that the search radius decreases to 2 cells and repeat searches are eliminated. The total cost decreases substantially, to \$1.90 million, and the median time to eradication is 8 years.

Our results suggest that increased use of public surveillance makes the invasion easier and cheaper to eradicate. The cost of bounties paid to the public are small compared to search costs, ranging between \$80,000 and \$210,000 in our example, and representing less than 10% of total costs.

In summary it appears that it is cost effective to engage the public in the search for pests, even if monetary rewards are involved. However, we increased passive detection probability exogenously (from 0.3 to 0.8) and did not capture all the costs of this shift. We included bounty costs, but not other costs, such as investment in public awareness campaigns as well as the time required to respond to false-positive reports. These costs are unknown and their estimation may require a behavioural model, involving the response of the public to information campaigns and monetary incentives. This is an interesting topic for future research. In the meantime, we can calculate the maximum amount we should be willing to pay for these information campaigns. By comparing *Base GA* and *GA Passive* in Table 3, we see that increasing the passive detection probability in private land from 0.3 to 0.8 results in \$1.29 million savings. This is the minimum amount we should pay to achieve such an increase in passive detection. The actual amount may be higher than this if the value of earlier eradication is factored in (e.g. as a reduction in the risk that the invasion will escape out of the surveillance area).

5. CONCLUSION

We developed a spatially-explicit model that represents the spread of a biological invasion and the process of detection by pest-control agencies and the public. The model was used to simulate the spread of a pest through a landscape. A GA was used to identify quasi-optimal strategies for controlling the invasion. The optimal parameter values for search effort, search radius and repeat visits selected by the GA algorithm led to improvements in the feasibility of eradication and the total cost of the control program.

Table 3. Results of base case and the two GA strategies.

Scenario:	<i>Base</i>	<i>Base GA</i>	<i>GA Passive</i>
Assumptions:			
Passive detection probability private land, p_{p1}	0.3	0.3	0.8
Results from GA: (^a indicates base values).			
m	2.00 ^a	8.06	5.48
r_m	5 ^a	4	2
S_R	3 ^a	1	0
Results from all simulations (means of 100 iterations):			
Total cost (\$M)	3.95	3.19	1.90
Bounty cost (\$M)	0.21	0.08	0.18
Search cost (\$M)	3.73	3.10	1.70
Treatment cost (\$M)	0.01	0.01	0.01
Final area invaded, $E(X_t)$	74	1	5
Median year eradicated	--	13	8
Probability of eradication	0.00	0.61	0.62
Mean passive detections (no.)	608	1889	405
Mean passive effort per year (h)	3,293	3,464	1,574
Mean active effort per year (h)	6,427	6,552	6,535
Mean repeat effort per year (h)	3,381	1,492	0

It was shown that passive surveillance may be a critically important source of information for public pest management agencies. Increased probability of passive detection resulted in considerable cost savings and higher probability of eradication. Active surveillance was also shown to be important, particularly when the probability of passive detection in public lands is low. The GA results indicated that, under the assumptions of our model, it is important to apply relatively high search effort per unit area, to achieve probabilities of detection > 0.9 , even at the expense of a smaller total area being searched.

This study was based on a number of arbitrary (although plausible) assumptions regarding the characteristics of the landscape and the invasive organism. The next step is to apply this model to a real pest/landscape scenario. Future research on whether changes in pest characteristics and landscape features may affect optimal strategies will be useful and contribute further to designing more efficient eradication programs.

ACKNOWLEDGMENTS

This research was funded by the Australian Centre for Excellence in Risk Analysis (ACERA). This work has benefited from discussions with Daniel Spring, Paul Pheloung, Dane Panetta and participants in a workshop for ACERA project 0806: *Application of search theory to invasive-species control programs*.

REFERENCES

- Beale, R., Fairbrother, J., Inglis, A. and Trebeck, D. (2008). *One Biosecurity: A working Partnership*. The independent review of Australia's quarantine and biosecurity arrangements report to the Australian Government. Commonwealth of Australia.
- Brooks, S.J., Galway, K.E. (2008). Processes leading to the detection of tropical weed infestations during an eradication program. In: Van Klinken, R.D., Osten, V.A., Panetta, F.D., Scanlan, J.C. (Eds.), 16th Australian Weeds Conference, ISBN 9780646488196, pp. 424-426.
- Cacho, O.J., Spring, D. and Hester, S. (in review). *Environmental Modelling and Software*
- Cacho, O.J., Hester, S.M., Spring, D. (2007). Applying search theory to determine the feasibility of eradicating an invasive population in natural environments. *Australian Journal of Agricultural and Resource Economics* 51, 425-433.
- Cacho, O.J., Spring, D., Pheloung, P., Hester, S. (2006). Evaluating the feasibility of eradicating an invasion. *Biological Invasions* 8, 903-917.
- Davis, P.R. and Wilson, P.L. (1991). Report on European wasps in Western Australia with special reference to the 1990-91 season. Department of Agriculture, Western Australia.
- Goldberg, D.E. (1989). *Genetic algorithms in search, optimization and machine learning*. Addison-Wesley, Reading.
- Jennings, C. (2004). A brief history of the red imported fire ant eradication program. *The Australian Journal of Emergency Management* 19 (3), 97-100.
- Liebholt AM, Macdonald WL, Bergdahl D and Mastro VC (1995). Invasion by exotic forest pests: a threat to forest ecosystems. *Forest Science Monographs* 30, 49 pp.
- Liebman M, Mohler CL and Staver CP (2001). *Ecological management of agricultural weeds*. Cambridge University Press, Cambridge, UK, 532 pp.
- MAFBNZ (Ministry of Agriculture and Forestry Biosecurity New Zealand) (2008). Review of the current state of the biosecurity surveillance system. <http://www.biosecurity.govt.nz/files/pests/surv-mgmt/surv/mafbnz-surv-strategy-current-state.pdf>.
- Mayer, D. (2002). *Evolutionary algorithms and agricultural systems*. Kluwer Academic Publishers, Boston.
- Saupe, D. (1988). Algorithms for random fractals. In Peitgen, H.-O., Saupe, D., (Eds.), *The science of fractal images*. Springer-Verlag New York Inc: New York.
- Taylor and Hastings, (2004). Finding optimal control strategies for invasive species: a density-structured model for *Spartina alterniflora*. *Journal of Applied Ecology* 41, 1049-1057.
- The Mathworks (2002). *Using Matlab*, The Mathworks, Inc, Matick, MA.
- Vitousek PM, D'Antonio CM, Loope LL and Westbrooks R (1996). Biological invasions as global environmental change. *American Scientist* 84: 468-478.