SoftwareX 21 (2023) 101321

Contents lists available at ScienceDirect

SoftwareX

journal homepage: www.elsevier.com/locate/softx

Original software publication

Modelling biocontrol of invasive insects using WaspSim: A MATLAB simulation model

Oscar J. Cacho [a], Susan M. Hester [a,b,*]

[a] UNE Business School, University of New England, Armidale, New South Wales, Australia
[b] Centre of Excellence for Biosecurity Risk Analysis, The University of Melbourne, Parkville, Victoria, Australia

ARTICLE INFO

Article history:
Received 26 October 2022
Received in revised form 17 January 2023
Accepted 19 January 2023

Dataset link: https://github.com/ocacho/WaspSim

Keywords:
Invasive species
Classical biological control
Vespula germanica
Control feasibility
Simulation model

ABSTRACT

Globally, governments spend significant amounts of money attempting to prevent or mitigate the negative impacts of invasive species using different strategies, including classical biological control for species that are widespread in the landscape. Arriving at the point of biocontrol release requires significant investment of time and funds for research and testing, and success of the biocontrol agent is uncertain. We present a numerical model designed to assist in this process. The model is based on well-established equations of population growth and spread. Technical feasibility of a biocontrol programme is assessed based on the population parameters for the pest and biocontrol agent.

Code metadata

Current code version	1.0
Permanent link to code/repository used for this code version	https://github.com/ElsevierSoftwareX/SOFTX-D-22-00344
Permanent link to reproducible capsule	–
Legal code license	MIT License.
Code versioning system used	none
Software code languages, tools and services used	Matlab
Compilation requirements, operating environments and dependencies	MATLAB codes to run in MATLAB 2019a or later
If available, link to developer documentation/manual	https://github.com/ocacho/WaspSim
Support email for questions	ocacho@une.edu.au

1. Motivation and significance

While insects are an integral part of ecosystem processes, they also have the capacity to become invasive, impacting human health, infrastructure, agriculture, and the environment. The negative impacts of invasive insects are enormous — Bradshaw et al. [1] estimate they cost at least US$70.0 billion per year globally, while associated health costs exceed US$6.9 billion per year. Decision analysis using simulation models has made significant contributions to the management of invasive species in recent years [2–6]. Models based on ecological principles allow biosecurity agencies to determine the conditions for success of management programmes prior to potentially significant investment occurring.

The spread of an invasion is a complex spatio-temporal and stochastic process which is not easy to model in an operational setting. The software described in this paper, WaspSim, provides a tool for modelling the spread of two species and their interaction as pest and biocontrol agent. The WaspSim model was developed as part of a project aimed at evaluating whether an invasive insect would be a good candidate for classical biological control (CBC) [7], where the aim is permanent establishment of an agent to control a pest. CBC can be an economically feasible option to control pests that are widespread in the landscape, but long time frames and large budgets are typically required for screening,

* Corresponding author.
E-mail addresses: ocacho@une.edu.au (Oscar J. Cacho), shester@une.edu.au (Susan M. Hester).

https://doi.org/10.1016/j.softx.2023.101321

Table 1
Model equations.

N	Equation	Description
(1)	$W_{a,t} = \alpha_W \left(1 - E_{\alpha,t}\right) W_{s,t} e^{-\beta_W W_{s,t}}$	Wasp growth, autumn
(2)	$W_{s,t+1} = W_{a,t} \left(1 - E_{\mu,t}\right)$	Wasp growth, spring
(3)	$B_{a,t} = \alpha_B B_{s,t} e^{-\beta_{B,t} B_{s,t}}$	Biocontrol growth, autumn
(4)	$B_{s,t+1} = B_{a,t} \left(1 - \mu_B\right)$	Biocontrol growth, spring
(5)	$E_{\alpha,t} = min\left(\rho_B B_{s,t}, 1\right)$	Biocontrol effect on wasp growth
(6)	$E_{\mu t} = min\left(\varphi_B B_{s,t}, 1\right)$	Biocontrol effect on wasp mortality
(7)	$\beta_W = \frac{\ln \alpha_W}{\kappa_W \theta_W}$	Growth exponent, wasp
(8)	$\beta_{B,t} = \frac{\ln \alpha_B}{\kappa_B W_{s,t}}$	Growth exponent, biocontrol
(9)	$p_{i,j} = \frac{1}{\pi\gamma\left(1+\left(\frac{d_{i,j}}{\gamma}\right)^2\right)}$	Dispersal kernel
(10)	$W_d = P_W W_t$	Distribution after dispersal, wasp
(11)	$B_d = P_B B_t$	Distribution after dispersal, bioc.
(12)	$p_d = 1 - e^{-\delta_W W_{a,t} H}$	Probability of detection
(13)	$K_t = p_d W_{a,t}$	Wasp nests destroyed
(14)	$\mathrm{W_{a,t}} = \mathrm{W_{a,t}} \left(1 - \mathrm{p_d}\right)$	Wasp nests remaining
(15)	$f\left(u\right) = \omega_N \left(N_{obs} - \overline{N}_{pred}\left(u_i\right)\right)^2 + \omega_A \left(A_{obs} - \overline{A}_{pred}\left(u_i\right)\right)^2$	Genetic Algorithm likelihood function

testing, breeding and releasing biocontrol agents. Technical feasibility is a pre-requisite for economic feasibility and if a biocontrol agent is not technically feasible, then an economic evaluation is not needed. Technical feasibility in the CBC sense means that the pest and the biocontrol populations reach an equilibrium at low density, where the damage caused by the pest is kept low at no additional expense. WaspSim allows for assessment of technical feasibility and will therefore be most beneficially used in the context of early assessments of biocontrol agent performance when screening potential candidate species.

The model of Barlow et al. [8] is modified in WaspSim to represent the growth of the wasp and biocontrol agent populations and their interaction, and to introduce dispersal equations for spatial spread of both species. WaspSim was recently applied by Cacho and Hester [9] who focused on the invasion of European wasp (*Vespula germanica*), in south-eastern Australia. The European wasp is a social insect, reliant on one or more queens for reproduction, with its population represented in the model as nests, or colonies, per unit area. The input data consists of maps of the area of interest, representing the presence of the pest over time, human population density, and habitat suitability for the pest.

The model is designed for social insects in general, where a colony (nest) has a fixed location, and spread over the landscape through time occurs via a queen. This is similar to the case of an invasive plant, where the adult has a fixed location and spread occurs through seed dispersal. In both cases, the core population (nests or plants) is fixed in space and dispersal occurs through propagules (queens or seeds). The model could be applied to non-social invasive insects, or other animals, as long as the spatial range of the core population (the variable that measures population density) is within a cell on the map.

2. Software description

WaspSim is a free MATLAB package for use in evaluating whether an invasive insect would be a good candidate for a classical biological control programme. It can be used to estimate uncertain population dynamics parameters based on public data, and can be extended to evaluate impacts based on the location of affected industries, human populations, and environmental assets. Examples are provided as stand-alone scripts. The mathematical model is explained in detail in Cacho and Hester [9] and details are not repeated here. The focus of this paper is on the application of the numerical model.

Matlab was chosen because it provides an integrated modelling environment based on matrices, it has several useful toolboxes with common syntax, and it is fast to solve large problems. The disadvantage is that accessing the Matlab environment has a cost, as opposed to other languages such as Python.

2.1. Software architecture

The model equations are listed in Table 1. These are conventional equations of population dynamics and spread. The effects of the biocontrol on the pest are assumed to be linear (see equations 5 and 6), but they can be easily adjusted if there is evidence that other functional forms are more appropriate.

The biocontrol agent in the model affects the pest population in two ways, by reducing the growth rate of new queens, and by increasing the mortality of nests. Model parameters can be adjusted to represent any biocontrol agent (or combination of agents) that act in this way. Other modes of action for the biocontrol could be represented by targeting other growth and dispersal parameters for the pest. The full code for the model is publicly available, so these changes could be made by users if needed.

The equations from Table 1 are packaged within a series of MATLAB functions to conduct stochastic spatiotemporal simulations of the pest and biocontrol populations. Running the simulation model requires access to the MATLAB Base software. To speed up execution the model will run in parallel if the Parallel Computing Toolbox is available. In addition, the Statistics and Machine Learning Toolbox is required to run some of the functions that process model outputs. Fig. 1 shows the steps of a simulation run as a flow diagram. This diagram represents the strategy used to run the model for the analysis conducted by Cacho and Hester [9]. Other approaches to conducting simulations can be followed by manipulating different parameters in the model and creating different experimental designs.

The first step (Fig. 1) is to select the biocontrol release strategy based on two decision variables: the spatial coverage of the release (x_c), and the intensity of release per site (x_p). Both x_c and x_p are expressed as percentages, x_c is the percentage of sites on the map that will be inoculated with the biocontrol based on density of the pest per site, and x_p is the proportion of nests that will be inoculated in each site. The second step is to

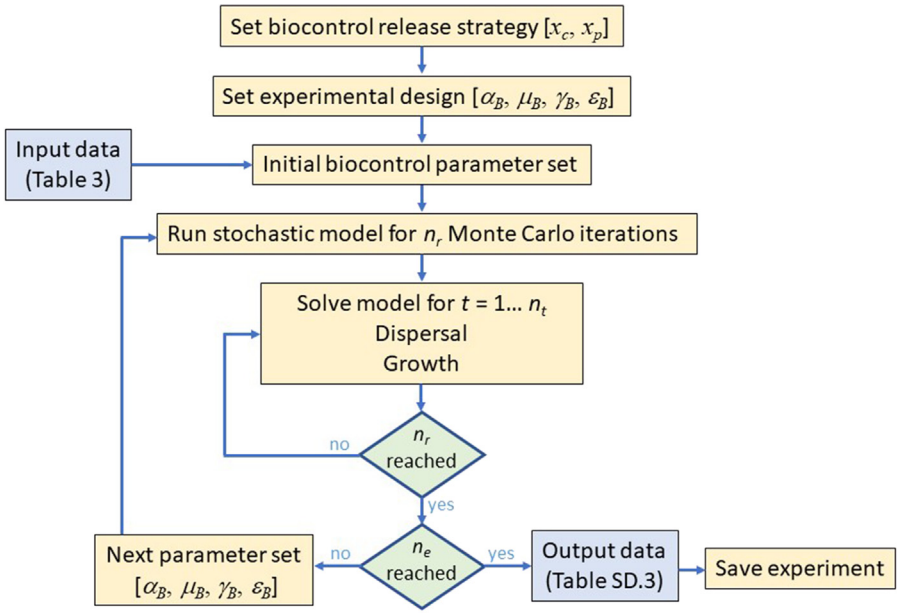

Fig. 1. Flow diagram of the wasp biocontrol simulation model.

setup the experimental design, which consists of selecting model parameters and combinations of values to be tested. In Cacho and Hester [9] we used a full factorial design with three levels for each factor, resulting in $n_e = 3^4 = 81$ experimental treatments. Each treatment is a row in the experimental matrix ($\mathbf{b}_{parm}$), which has dimensions 81×4.

The next series of steps consists of two loops, setting the biocontrol parameters based on each row of ($\mathbf{b}_{parm}$), and running a Monte Carlo simulation for the desired number of stochastic iterations (n_r). Within each stochastic iteration, a simulation is run for the desired number of time periods (n_t). Each simulation produces output matrices for the pest and the biocontrol populations ($\mathbf{w}_t$ and $\mathbf{b}_t$ respectively). These matrices have dimensions $n_c \times n_t$, where n_c is the number of grid cells in the map for the area of interest. Essentially, each column of $\mathbf{w}_t$ and $\mathbf{b}_t$ is a map of the state of the invasion at time t for the corresponding stochastic run, expressed as density per unit area for each cell. The results of stochastic simulations are saved in tri-dimensional matrices $\mathbf{W}_t$ and $\mathbf{B}_t$, of dimensions $n_c \times n_t \times n_r$, where each matrix along the third dimension represents a realisation of $\mathbf{w}_t$ and $\mathbf{b}_t$.

Using this process, we produce 81 sets of $\mathbf{W}_t$ and $\mathbf{B}_t$ matrices (see Example 1 below), one for each row of the experimental design matrix $\mathbf{b}_{parm}$. At this stage the analyst needs to decide how to package these results. Options include:

(1) Introducing a fourth dimension to contain the results for each parameter set.
(2) Using an array of n_e *Struct* or *Cell* variables within MATLAB to contain the three-dimensional matrices for the whole experiment;
(3) Producing a two-dimensional dataset containing individual observations of w_t and b_t per map cell, time period, stochastic run, and experimental treatment; or
(4) Calculate the summary statistics across stochastic runs and package the summary results within new *Struct* variables.

The best option depends partly on personal preferences and skills of the analyst, partly on further analysis needs, and partly on memory and storage space available. It would also be influenced by the intended audience. Many terabytes of data can be produced quickly depending on the size and resolution of the map (n_c), the desired time horizon (n_t), the number of rows of the experimental design matrix (n_e), the number of stochastic iterations (n_r), and the number of biocontrol release strategies to be simulated.

In the example below, we use option (4) to save storage space and gain faster access to the data saved on disk for further analysis. Strategy 3 is well suited to produce output files in csv format that can be read into other software (such as R or Python) for further analysis of results. Options 1 and 2 allow the full set of results to be saved but will result in very large binary (.mat) files that can be slow to access later on, particularly when stored in remote servers.

We save the summary results from stochastic runs in one output file for each experiment, so each biocontrol-release strategy has its own file. The outputs of the model are explained in more detail in the Supplementary Data (SD) (see Table SD.3). The results are saved as data structures with the expected value (*ev*), the standard deviation (*sd*) and the probability of wasp and biocontrol presence in each cell on the map (*prob*). These three variables have the same structure (Figure SD.4), containing tri-dimensional matrices for wasps ($\mathbf{W}_t$) and biocontrol ($\mathbf{B}_t$), with dimensions $n_c \times n_t \times n_e$.

2.2. Software functionalities

The components of the MATLAB model are listed in Table 2. There are 4 scripts and 12 functions. Most of the work happens within the functions, but the average user only needs to become familiar with the scripts to run the model effectively. Users who wish to introduce more sophisticated changes, such as modifying the nature of the biocontrol effects on the pest (equations 5 and 6) will need to modify code within functions. To simplify this process, the functions in Table 2 are cross-referenced to the equations in Table 1, making it easier to locate the function that needs to be modified for a particular purpose.

Running the model

To get started, open *Sim_cluster_script.m* in the MATLAB environment. This script allows you to generate the results reported

Table 2
MATLAB model components.

Filename	Description
Scripts	
Sim_cluster_script.m	Main script to run the model
Create_W0.m	Creates initial random invasion matrix
Wasp_GA_script.m	Main script to run the GA parameter estimation process
Plot_figs.m	Script to read result files and plot figures presented in this paper.
Functions	
Disp_Kernel.m	Dispersal kernel *(Eq. 9)*
Extract_Results.m	Packages summary results from stochastic wasp simulation to save memory
Grow_Ricker.m	Ricker growth function *(Eqs. 1, 2, 3, 4)*
Lik_fn.m	Likelihood function to be minimised by GA to estimate wasp parameters *(Eq. 15)*
Rand_Spread.m	Random spread function using a loop for invaded sites *(Eqs. 10, 11)*
Rand_Spread_02.m	Random spread function using a full matrix (slow) *(Eqs. 10, 11)*
Run_experiment.m	Runs an experiment of simulation model for a set of treatments with varying biocontrol parameters
Run_Wasp_GA.m	Runs one iteration the GA for a given random seed
Sim_Biocontrol.m	Runs stochastic simulation of wasp-biocontrol for given parameter set
Solve_W0.m	Estimates initial number of wasp infestations based on presence reports and probability of detection *(Eq. 12)*
Spread_Stoch_Biocontrol.m	Runs one step of wasp-biocontrol simulation for n_t years *(Eqs. 5, 6, 7, 8, 12, 13, 14)*
Spread_Stoch_control.m	Runs one step of wasp simulation for n_t years with no biocontrol but with destruction of nests that are detected *(Eqs. 7, 12, 13, 14)*
Data	
base_data_nsw.mat base_data_vic.mat	base input data sets used to initialise the wasp invasion
ga_data_nsw.mat ga_data_vic.mat	input datas used in the GA estimation
nsw_0.mat vic_0.mat	counterfactual (no biocontrol) simulation results
nsw_20_10_1.mat vic_20_10_1.mat nsw_50_01_1.mat vic_50_01_1.mat	Simulation results for biocontrol release strategies xp = proportion release per site (20% or 50%) xc = spatial coverage (10% or 1%)
W0_prm_nsw.mat W0_prm_vic.mat	Input data for main simulation, based on initial random invasion matrix W0 previously generated with Create_W0.m

in Cacho and Hester [9]. Before running the code, consider that it may take several days to complete the simulations, depending on the computer used. You may wish to change the experimental design and simulate only a few cases for a single spatial cluster to get a feel for the model.

There are two input data files available (*W0_prm_nsw.mat* and *W0_prm_vic.mat*) that represent wasp invasions in the Australian states of Victoria (VIC) and New South Wales (NSW). Loading either of these files in the MATLAB environment allows users to inspect the variables required to run the model. Becoming familiar with these variables is a required step for those who wish to apply the model to other regions or to a different species of pest.

The workspace for these two input files is shown in Fig. 2. Notice the symmetry between the variables for the two clusters, with the only difference being in the dimensions of matrices that represent maps. There are 389 grid cells (n_c) in NSW (Fig. 2 A) and 658 in VIC (Fig. 2B). Simulations will take longer to solve for VIC compared to NSW due to the larger dimensions of the matrices.

Going back to the *Sim_cluster_script.m* file, note the correspondence of the code to the steps illustrated in Fig. 1. In step (1) the biocontrol release strategy is set using a factorial design for 25 combinations of x_c and x_p. In step (2) the experimental design is set using a factorial experiment for 81 combinations of biocontrol parameter values. In step (3) the experiment is run, first for VIC and then for NSW. This last step calls *Run_experiment.m* within a loop and saves a file for each biocontrol release strategy in the *Results* folder.

The number of stochastic runs (Monte Carlo iterations) is predetermined by the number of columns in the initial state matrix *W0*, which was created earlier using the script *Create_W0.m*. Fewer iterations can be run by truncating columns of *W0* if desired, or running *Create_W0.m* again for the desired number of iterations. In the latter case, it is a good idea to save the file with a different name to avoid confusion. For example, if a new version of W0 is created for 1200 stochastic runs the files could be named *W0_prm_nsw_1200.mat* and *W0_prm_vic_1200.mat*.

3. Illustrative examples

Example 1. Biocontrol feasibility

In this example we show how to replicate some of the results reported in Cacho and Hester [9]. The main interest in that study was to find sets of biocontrol parameter values that would result in a feasible programme. Feasibility, in the classical biological control sense, means that the pest and biocontrol populations reach some form of equilibrium where the pest is kept at low density, and for negligible cost into the future after the initial investment.

In this example we test four parameters for the biocontrol agent (step 2 in Fig. 1): the growth rate (α_B) the carrying capacity per wasp nest (μ_B), the median dispersal distance (γ_B), and the effectiveness of the biocontrol in suppressing wasp nests (ε_B). We used a full factorial design with three levels for each factor, resulting in $n_e = 3^4 = 81$ experimental treatments. Each

A

Workspace

Name	Value
D	389x389 double
garea	389x1 double
gmap	389x1 struct
gpoly	389x1 polyshape
hh_dens	389x1 double
prm	1x1 struct
W0	389x1000 double
wp_mat	20x4 double
wp_samp	1000x4 double

B

Workspace

Name	Value
D	658x658 double
garea	658x1 double
gmap	658x1 struct
gpoly	658x1 polyshape
hh_dens	658x1 double
prm	1x1 struct
W0	658x1000 double
wp_mat	20x4 double
wp_samp	1000x4 double

Fig. 2. The workspace after loading the input files *W0_prm_nsw.mat* (A) and *W0_prm_vic.mat* (B). The meanings of variables in the workspace are given in Table SD.2.

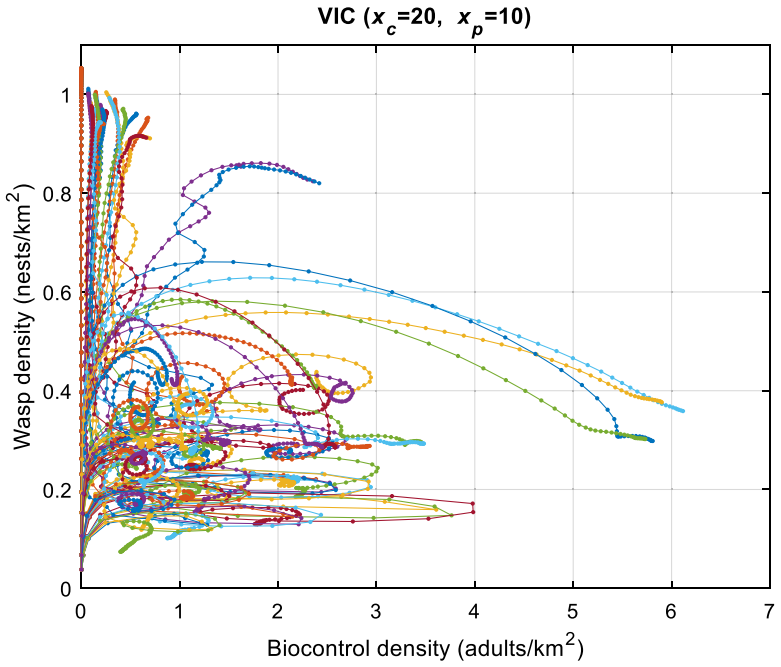

Fig. 3. Simulation results presented as a phase diagram for the VIC cluster, with biocontrol release coverage (x_c) of 20% and release intensity (x_p) of 10%. Each simulation begins at the top left of the diagram, corresponding to high wasp density and low parasitoid density, and general tracks down and to the right towards a steady state. Colours represent the (81) different time trajectories.

treatment is a row in the experimental matrix ($\mathbf{b}_{parm}$), which has dimensions 81 × 4.

A useful way of finding feasible solutions is to plot phase diagrams showing the trajectories through time of the pest and the biocontrol against each other (Fig. 3). The trajectories that end in equilibrium towards the bottom of the diagram (at low wasp density) are feasible in a control sense. Recall that each output file produced by *Sim_cluster_script.m* represents an experiment with a different biocontrol release strategy. Each experiment would be associated with a phase diagram, so in this case we would have 81 diagrams for each of the two clusters (VIC and NSW). This link between phase diagram and experiment becomes obvious by inspecting the MATLAB code in Listing 1, used to construct Fig. 3. Note that the file *vic_20_10.mat* is loaded to create the plot. This instruction can be replaced with a different file name (and the title of the plot adjusted) to represent any experiment.

Listing 1: Code to plot Figure 3

```
clear;
load('vic_20_10.mat');
Wmean = permute(sum(ev.W .* garea)./sum(garea),[3,2,1]);
Bmean = permute(sum(ev.B .* garea)./sum(garea),[3,2,1]);
plot(Bmean', Wmean','.-');
axis([0,7,0,1.1]);
xlabel('Biocontrol density (adults/km^2)');
ylabel('Wasp density (nests/km^2)');
title('VIC ({\itx_c}=20, {\it x_p}=10)')
grid
clear;
load('vic_20_10.mat');
```

The MATLAB code in Listing 1 shows one way of accessing the simulation results using the *ev* (expected value) data structure explained before. The phase diagram is generated with the command plot(*Bmean*', *4*', '.-'); the variables *Bmean* and *Wmean* are

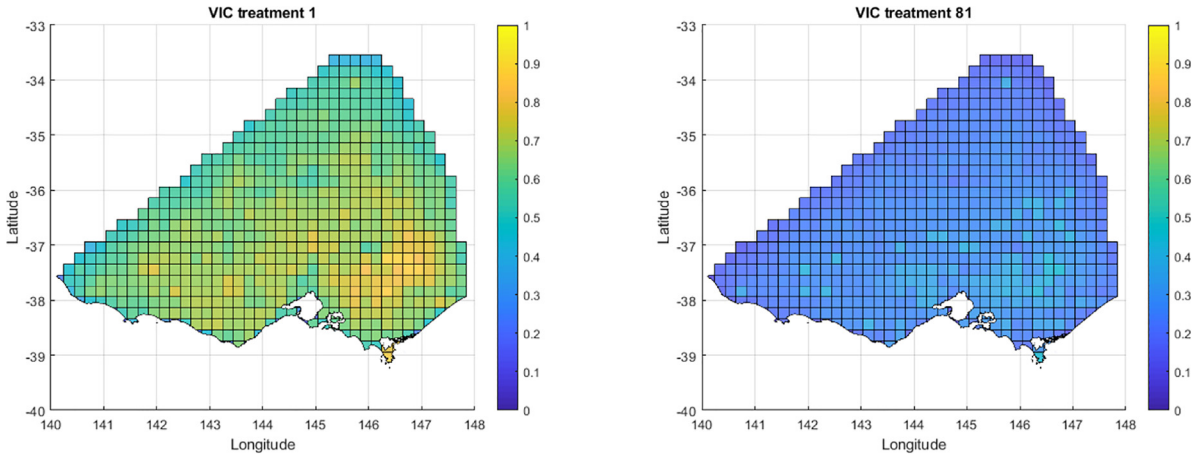

Fig. 4. Wasp invasion maps for two different experimental treatments in the VIC cluster.

the mean density of biocontrol and wasp populations over time for the whole map. The command that creates the latter variable is:

Wmean=permute(sum(ev.W. ∗ garea)./sum(garea), [3, 2, 1]);

The key variables here are *ev.W* and *ev.B*, which contain the means of 1000 stochastic simulations packaged as a tri-dimensional matrices with dimensions (658 × 60 × 81), equivalent to $(n_c \times n_t \times n_e)$ for the VIC cluster. To calculate the mean wasp density across the 658 cells on the map, we first convert to total nests per cell, sum across all cells and divide by total area of the map. We need to do this because not all cells are the same area (due to truncation along the coastline). The command sum(ev.W .* garea)./sum(garea) calculates a weighted average of wasp density based on the areas of the cells in the map. This command results in a matrix of dimensions (1 × 60 × 81), which is converted to two dimensions (81 × 60) by swapping the first and third dimensions using permute. The same operation is applied to *ev.B* to obtain *Bmean*. The rows of *Wmean* and *Bmean* represent experimental treatments and the columns are years. Plotting the transpose of these matrices results in the 81 time trajectories in Fig. 3.

Above we aggregated over the whole map to work with mean population densities and did not capture spatial variation in the data. We can also use *ev.W* to look at spatial patterns in wasp density using the command Wmap = permute(mean(ev.W,2), [1–3]). In this case we estimate the mean across the second dimension (time) to calculate the average wasp density per cell on the map over the whole time horizon, and the permute command swaps the second and third dimensions. This results in a matrix *Wmap* of dimensions (658 × 81), where each column is the map of average wasp density for a different experimental treatment. Fig. 4 displays the maps for the first and last treatments in the experiment, showing the difference between failure and success based on average wasp density over time. The code to plot these maps is in the *Plot_figs.m* function, starting on line 17.

Example 2. Estimating wasp parameter values using a genetic algorithm

Parameter values for any species could be estimated using different estimation techniques depending on data available and the spatial scale of the analysis. This could be done by embedding the simulation model within other types of algorithms to fit the model to data.

In the case study of Cacho and Hester [9], parameter values for the population dynamics of European wasp in Australia were not available, but realistic ranges were known for the four key parameters: growth rate (α_W), carrying capacity (κ_W), median dispersal distance (γ_W), and probability of detection (δ_W). In this example we use the model to estimate values for these parameters using a genetic algorithm (GA) combined with occurrence data from Atlas of Living Australia (ALA) (see Table SD.1). Details of the GA estimation are provided in the SD.

The parameter distributions for the two clusters in our case study are compared against each other in Fig. 5. See the *Plot_figs.m* script for the code to generate this figure.

4. Impact

WaspSim was designed as a tool to evaluate the technical feasibility of a biological control programme. It will be a useful tool for biosecurity practitioners prior to investing significant time and funds into CBC programmes that could be technically unfeasible, and which often fail [10]. The original research was commissioned by a national biosecurity agency for this purpose. The first application of the model was published recently [9], but we expect other applications to follow as practitioners become aware of the tool and learn to use it. The model can be run easily by anyone with access to MATLAB. The data used to calibrate and populate the model are publicly available for many species (see Table SD.1) and other datasets would be available to biosecurity agencies with interest in applying the model. The innovation in WaspSim is in its integration of data and code to provide a useful tool for a complex spatiotemporal problem.

Example 1 involved assessment of technical feasibility for a biocontrol programme, based on biological parameters for pest and biocontrol species. This in itself can lead to considerable savings by focusing only on candidate species most likely to succeed. The assessment can be extended to evaluate economic feasibility. By overlaying maps of industries and environments affected by the pest, we can estimate damages avoided over time for any pest control scenario.

The model can be applied to other social insect species via modification of population parameter values. It can also be used for spread modelling of a single species, independently of the biocontrol component. This is needed for delimiting the likely area of an incursion, designing containment strategies and conducting time to eradication modelling. All these are actions that biosecurity agencies need to undertake on a regular basis with less than ideal analytical tools. WaspSim fills an important need and we expect it to have impact within the biosecurity and invasive species communities.

Application of the model requires information on likely values of parameters for the population dynamics of the biocontrol agent and the pest, which could come from the literature or from fitting

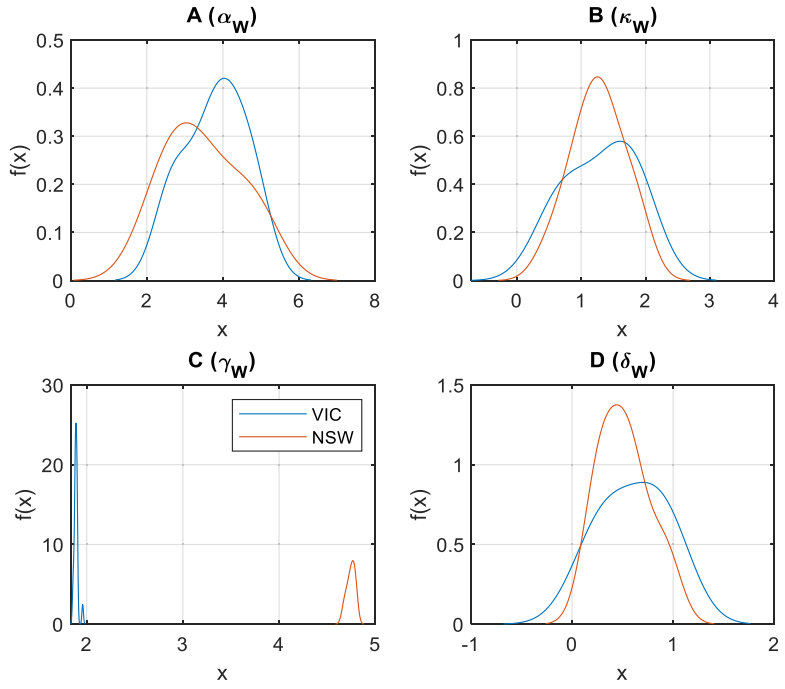

Fig. 5. Comparison of wasp population parameter results for two geographical clusters (VIC and NSW) expressed as density distributions for growth rate (A), carrying capacity (B), median dispersal distance (C) and detection probability (D). Values estimated using a GA with 20 different random seeds.

equations to data. This means that the model is applicable once one or more potential biocontrol agents have been identified, and some knowledge of their population dynamics exists. The model could be used to determine which species should be selected for further investment.

5. Conclusions

Implementing classical biological control programmes can take significant amounts of time and money, and success is not guaranteed. Mathematical models of biological control can give early indications about the technical feasibility of the agent, using available information on pest spread and likely agent performance in the environment into which it will be introduced. WaspSim has been developed for this purpose. We have shown its application to control of European wasp given available data on pest spread and assumptions about biocontrol agent performance. The feasibility that the biocontrol agent will establish and spread; and the effectiveness of the agent in suppressing growth and spread of the pest, are key to the success of the biological control programme. Understanding the conditions that would make a biocontrol agent effective can help screen potential candidates and focus efforts only on those that are likely to succeed.

CRediT authorship contribution statement

Oscar J. Cacho: Conceptualization, Methodology, Software, Formal analysis, Writing – original draft, Writing – review & editing. **Susan M. Hester:** Data curation, Writing – original draft, Writing – review & editing, Project administration.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The links to the research data/code are in https://github.com/ocacho/WaspSim.

Acknowledgements

This research was funded by the Centre of Excellence for Biosecurity Risk Analysis, Australia at the University of Melbourne via CEBRA Project 190804 '*Re-evaluating management of established pests including the European wasp,* Vespula germanica, *using biocontrol agents*'.

Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.softx.2023.101321.

References

[1] Bradshaw CJA, Leroy B, Bellard C, et al. Massive yet grossly underestimated global costs of invasive insects. Nature Commun 2016;7:12986. http://dx.doi.org/10.1038/ncomms12986.

[2] Cacho OJ, Spring D, Hester SM, Mac Nally R. Allocating surveillance effort in the management of invasive species: a spatially-explicit model. Environ Model Softw 2010;25(4):444–54. http://dx.doi.org/10.1016/j.envsoft.2009.10.014.

[3] Carrasco LR, Baker R, MacLeod A, Knight JD, Mumford JD. Optimal and robust control of invasive alien species spreading in homogeneous landscapes. J R Soc Interface. 2010;7:529–40. http://dx.doi.org/10.1098/rsif.2009.0266.

[4] Epanchin-Niell RS, Hastings A. Controlling established invaders: integrating economics and spread dynamics to determine optimal management. Ecol Lett. 2010;13(4):528–41. http://dx.doi.org/10.1111/j.1461-0248.2010.01440.x.

[5] Portela R, Vicente JR, Roiloa SR, Cabral JA. A dynamic model-based framework to test the effectiveness of biocontrol targeting a new plant

invader– the case of Alternanthera philoxeroides in the Iberian Peninsula. J Environ Manag 2020;264:110349. http://dx.doi.org/10.1016/j.jenvman.2020.110349.

[6] Yemshanov D, Haight RG, Koch FH, Lu B, Venette R, Fournier RE, et al. Robust surveillance and control of invasive species using a scenario optimization approach. Ecol Econ. 2017;133:86–98. http://dx.doi.org/10.1016/j.ecolecon.2016.11.018.

[7] Cacho O, Hester S, Tait P. Re-evaluating management of established pests including the European wasp. In: Centre of Excellence for Biosecurity Risk Analysis; 2021 Nov. Report for Project 190804. Melbourne (VIC: Vespula germanica using biocontrol agents, https://cebra.unimelb.edu.au/__data/assets/pdf_file/0007/3949891/190804-Final-Report-updated-November-2021.pdf.

[8] Barlow ND, Moller H, Beggs JR. A model for the effect of Sphecophaga vesparum vesparum as a biological control agent of the common Wasp in New Zealand. J Appl Ecol 1996;33(1):31–44. http://dx.doi.org/10.2307/2405013.

[9] Cacho OJ, Hester SM. Modelling biocontrol of invasive insects: An application to European Wasp (*Vespula germanica*) in Australia. Ecol Model 2022;467:109930. http://dx.doi.org/10.1016/j.ecolmodel.2022.109939, Retrieved from https://www.sciencedirect.com/science/article/pii/S0304380022000618.

[10] Cock MJW, Murphy ST, Kairo MTK, Thompson E, Murphy RJ, Francis AW. Trends in the classical biological control of insect pests by insects: an update of the BIOCAT database. BioControl. 2016;61(4):349–63. http://dx.doi.org/10.1007/s10526-016-9726-3.